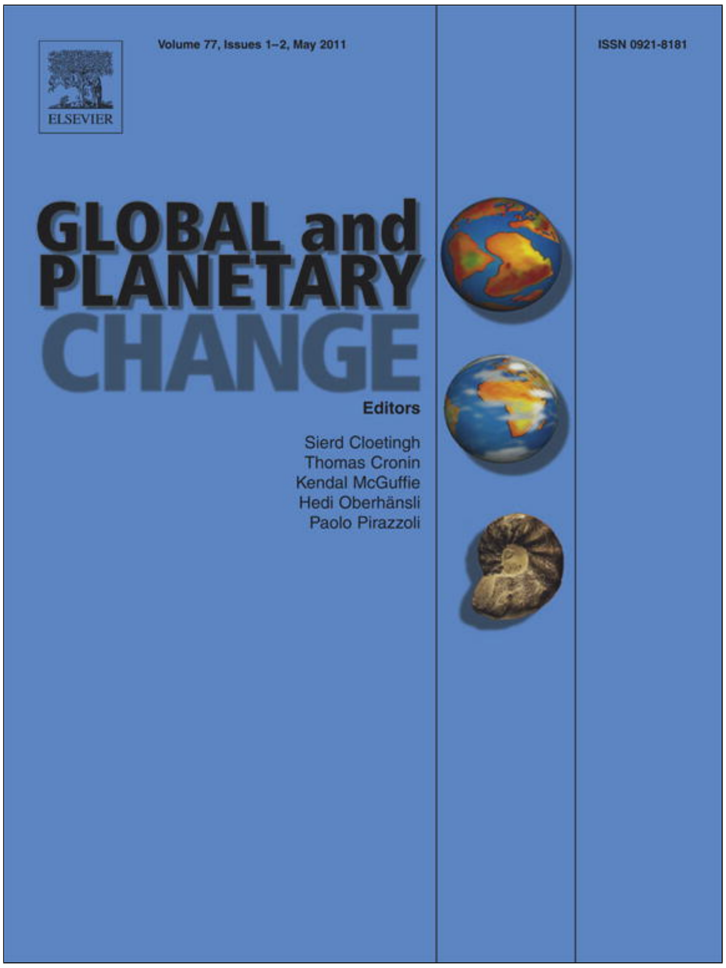
Volume 77, Issues 1–2, May 2011
ISSN 0921-8181
ELSEVIER
GLOBAL and
PLANETARY
CHANGE
Editors
Sierd Cloetingh
Thomas Cronin
Kendal McGuffie
Hedi Oberhänsli
Paolo Pirazzoli

Global and Planetary Change 77 (2011) 77–84

Contents lists available at ScienceDirect

Global and Planetary Change

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

Using MODIS snow cover and precipitation data to model water runoff for the Mokelumne River Basin in the Sierra Nevada, California (2000–2009)

Cynthia Powell *, Leonhard Blesius [1], Jerry Davis [2], Falk Schuetzenmeister [3]

Department of Geography, San Francisco State University, San Francisco, CA 94132, USA

ARTICLE INFO

Article history:
Received 23 December 2010
Accepted 31 March 2011
Available online 8 April 2011

Keywords:
MODIS
snow cover
statistical model
water supply
precipitation
Mokelumne River
climate change

ABSTRACT

Climate change will affect snowpack and water supply systems in California, and methods for predicting daily stream flow help prepare for these changes. This research provides a daily model to predict stream flow based on snow cover and precipitation in the Mokelumne River Basin in the Sierra Nevada in California. The snow cover of the Mokelumne River Basin is monitored using Moderate Resolution Imaging Spectroradiometer (MODIS) satellite images. Using data from these images as well as precipitation data from 2000 to 2009, we produced a predictive statistical model. The final results show that with an R^2 of 0.71, the true natural flow (TNF) of the Mokelumne River is based on the daily area of snow cover in each of seven equal area elevation zones according to the time lag of that zone as well as the accumulated precipitation functioning as a proxy for snow depth. The capability of this model to predict water supply suggests the potential for developing new spatial hydrologic informational products based on MODIS and the probability of improving the accuracy of the prediction of hydrologic processes for water resource managers.

1. Introduction

The Intergovernmental Panel on Climate Change (IPCC) Fourth Assessment Report (2007) outlines several expected impacts of global warming that bear consequences for US water resources. This report states that reductions in snow cover over recent decades are projected to accelerate throughout the 21st century, thus reducing water availability, hydropower potential, and changing seasonality of flows in regions supplied by mountain meltwater (IPCC, 2007).

General Circulation Models (GCMs) have been downscaled to predict more specific regional consequences for California water resources (Wilby and Wigley, 1997). One common GCM, the HadCM2 developed by the Hadley Centre in the UK, predicts increased overall regional precipitation when applied to California for all fossil fuel scenarios (Brekke et al., 2004). The PCM Parallel Climate Model produced in the US and funded by the Department of Energy, on the other hand, predicts decreased overall precipitation for all fossil fuel scenarios (Miller et al., 2003; Vanrheenen et al., 2004; Brekke et al., 2004). California's warming rates are similar in the HadCM2 and PCM models for all fossil fuel scenarios (Brekke et al., 2004). Many models, but not all, predict increased extreme precipitation events irrespective of increasing or decreasing precipitation amount (Trenberth, 1998; Kim, 2005).

Although no model consensus has been reached about changes in overall precipitation volume in California based on downscaled models (Leung and Ghan, 1999; Kim, 2001), increased temperatures will nonetheless change the current water supply systems in place due to California's geography. The Sierra Nevada stores snow until the spring or summer when snow melts into rivers which, combined with aqueducts, bring water to the Central Valley and coastal areas where the majority of the California population lives (EBMUD, 2004). Thus the mountains function as a “natural reservoir,” and most of the state's water supply derives from snowmelt runoff from these mountains. Approximately 75% of the annual discharge of most of the major streams in the Western US is from melting mountain snowpack (Palmer, 1988).

In addition to less snow overall, although not necessarily less precipitation, climate change has also already changed the timing of spring runoff; the water that was stored as snow until August may now be stored only until early summer (Mote et al., 2005; Dozier et al., 2008). This leads to water shortages later in the summer and may decrease available electricity supplied from hydropower as well (Kiparsky and Gleick, 2003; Dettinger et al., 2004; Dozier et al., 2008). Earlier California snowmelt runoff in the spring would decrease summer freshwater flow into the Sacramento–San Joaquin Delta allowing more time for summer saltwater intrusion into the Delta (Kiparsky and Gleick, 2003).

River flow forecasts are reasonably accurate for any water year (October 1–September 30) that is close to the mean of the historical record (Davis et al., 1999). With climate change, fewer water years

* Corresponding author. Tel.: +1 415 699 6773.
E-mail addresses: cynthiap@sfsu.edu (C. Powell), lblesius@sfsu.edu (L. Blesius), jerry@sfsu.edu (J. Davis), schuetzenmeister@berkeley.edu (F. Schuetzenmeister).
[1] Tel.: +1 415 405 2144; fax: +1 415 338 6243.
[2] Tel.: +1 415 338 2983.
[3] Tel.: +1 510 604 8010.

0921-8181/$ – see front matter
doi:10.1016/j.gloplacha.2011.03.005

will be close to the historical record and new methods for assessing water supply will need to be developed (Dozier et al., 2008). While monthly forecasting models may be more useful than an annual model, a daily model would allow water supply managers to improve decisions regarding when and how much water to release from the reservoir each day. The purpose of this research was to develop a daily runoff forecasting tool for East Bay Municipal Utility District (EBMUD), the agency responsible for supplying much of the Eastern San Francisco Bay population with Mokelumne River water (EBMUD, 2004).

EBMUD currently maintains tools for forecasting annual and monthly runoff into the Pardee Reservoir based on historical records of precipitation and surveyed snow water equivalent; however, these tools have limited use at sub-monthly time scales. The objective of this research was to develop an improved water supply model for water resource agencies such as EBMUD. The improved model described in this paper integrates precipitation data with MODIS (Moderate Resolution Imaging Spectroradiometer) satellite data showing the extent of snow covered area (SCA). Model parameters such as lag times were set for conditions in the Mokelumne River watershed, but can be adjusted for other watersheds. Thus, it is anticipated that the model can be used for other watersheds, taking local conditions into account.

2. Study area

The Mokelumne River is located on the west side of the central Sierra Nevada and flows into the San Joaquin River. The river drains a total area of approximately 1700 km^2. The upper basin to the east contains forested mountains and the western part of the basin contains grassy foothills and agricultural floodplains. The Mokelumne's channel is an average of 30 m wide (Pasternack et al., 2004).

The Upper Mokelumne River Basin located above the Pardee Reservoir (Fig. 1) is the focus of this study since it is the only part of the Mokelumne River Basin that is used for water supply (EBMUD, 2008). This upper basin supplies 1.3 million people with water in Alameda and Contra Costa counties, including the cities of Oakland and Berkeley (EBMUD, 2007). This area consists of approximately 1500 km^2 and ranges from 150 to 3,150 m in elevation (National Elevation Dataset, 2009).

Heavily influenced by a Mediterranean climate, the basin receives most of its precipitation between October and April. Over 50% of this precipitation is in the form of snow (Jeffres et al., 2006). Average precipitation from 2000 to 2009 was approximately 125 cm (50 in.) per yr. From October to April, precipitation ranges from approximately 120 cm in the headwaters to 51 cm in the lowlands (Pasternack et al., 2004).

Spring snowpack in the Sierra Nevada declined in the past half century, especially at lower elevations (Dozier et al., 2008). Overall precipitation may have increased, but some of the precipitation that used to fall as snow is now falling as rain (EBMUD, 2008). Because mountain snowpack serves as a natural water reservoir, the amount of water supplied by snowpack to this river will likely change with climate change. Thus, it is helpful to water supply engineers to be able to predict how much water the river can provide on a daily basis throughout the year. This information is valuable because climate change affects the amount, as well as the duration, of water available from this basin.

The upper watershed is contained within national forest or other undeveloped lands. Mokelumne River water from the upper part of the basin is stored at the Pardee Reservoir (drinking water) and the Camanche Reservoir (irrigation water). Both of these reservoirs are located 140 km (90 miles) east of Oakland, California. Water from the Pardee Reservoir is conveyed via three aqueducts to EBMUD's service area.

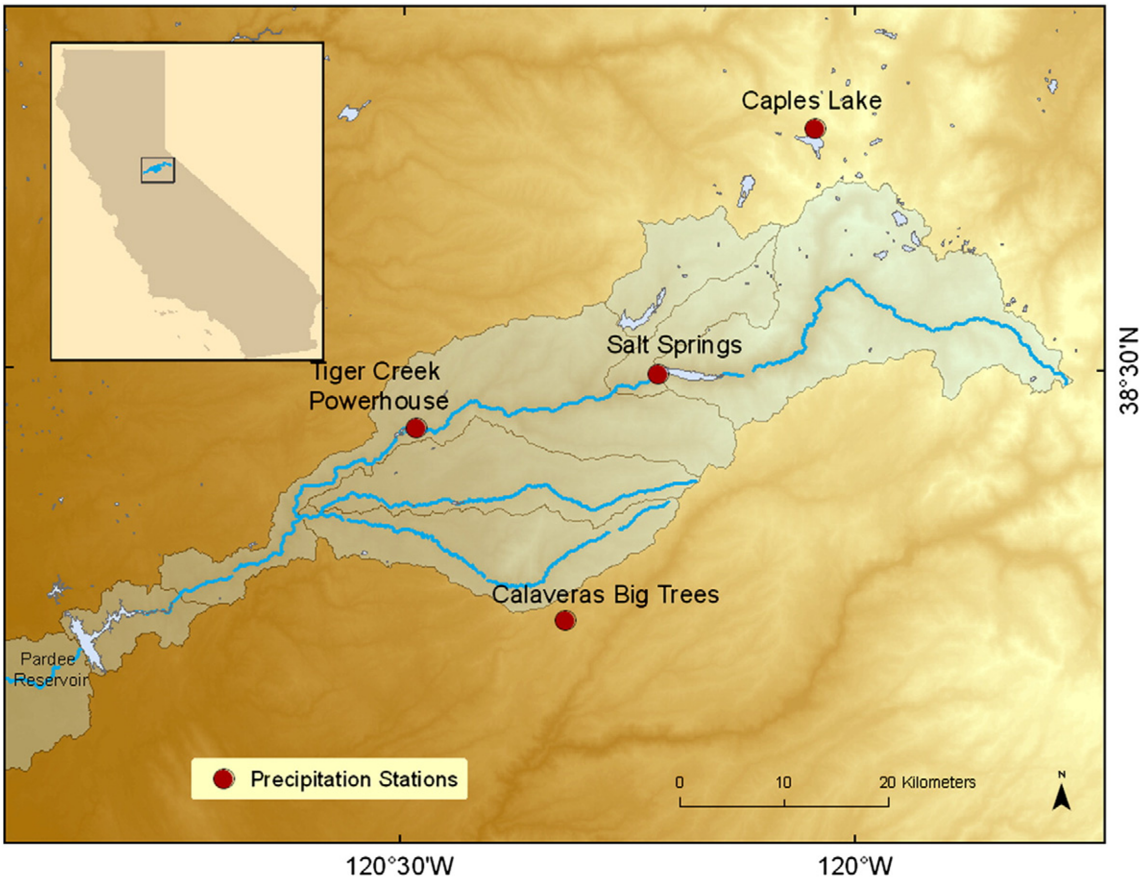

Fig. 1. Upper Mokelumne River Basin and precipitation stations.

3. Data acquisition

Three types of data were used to develop this model. (1) Mokelumne River true natural flow (TNF) data, or the amount of water that would flow into the Pardee Reservoir without human interference, were used to calibrate the model. (2) Precipitation data collected from four locations in and near the watershed were used to create a proxy for snow depth. (3) MODIS satellite snow cover data were used directly as inputs in this model to see how the TNF changed based on area of snow cover. Details of these datasets are as follows.

3.1. True Natural Flow (TNF)

In order to calibrate or "train" the model, the calculated TNF of the Mokelumne River was used. The flow of the Mokelumne River is altered before it reaches the Pardee Reservoir; Pacific Gas and Electric Company (PG&E) stores and releases water to generate electricity, and Amador and Calaveras County water agencies divert water before it reaches the Pardee Reservoir. EBMUD water supply engineers measure the actual amount of water that flows into the Pardee Reservoir, but they do not know exactly what this amount would have been if not for the diversions and store-and-release practices that occur while the water is en route. Because this research is based on a natural system (snowpack and melt), EBMUD-calculated TNF values were used to estimate what the flow would have been if not for the diversions and store-and-release practices.

EBMUD calculates the daily TNF just upstream of the Pardee Reservoir at a float gauge station. There are negative numbers in the TNF as well as extremely high outliers. To address the negative flow numbers, the average of the day before and the day after was used when a negative number appeared. If one of those values was negative, then the following day and the previous day was used. To address the extremely high outliers, a 3-day moving average was used after the negative values were omitted.

3.2. Precipitation

Daily precipitation data were calculated using the mean of four locations (Fig. 1): Caples Lake (Department of Water Resources), Salt Springs Powerhouse (Pacific Gas & Electric), Tiger Creek Power House (Pacific Gas & Electric), and Calaveras Big Trees (EBMUD). These four stations were chosen because they are well-spaced in and near the basin, represent a span of elevations (respectively 2430 m; 1100 m; 910 m; 1400 m) (National Elevation Dataset, 2009), and have a history of accurate records dating to 1930 (EBMUD, 2008). They are also the four stations used by EBMUD.

3.3. MODIS images

MODIS images are distributed through the National Snow and Ice Data Center in Boulder, Colorado (http://www.nsidc.org/) as a gridded (sinusoidal map projection) HDF file (hierarchical data format). The MODIS/Terra Snow Cover Daily L3 Global 500 m Grid was used for this research; The Mokelumne River Basin is located entirely within the H08V05 tile of the MOD10A1 version 5, the most current daily 500 m resolution product available in August of 2009. Each tile is approximately 1200 km × 1200 km (Riggs et al., 2006). MODIS Data Products from both Terra and Aqua include swaths of 500 m or 1 km resolution, snow cover daily, snow cover 8-day, and sea ice extent (http://www.nsidc.org 2010).

A MODIS instrument is aboard NASA's Terra and Aqua satellites. Actual data from Terra MODIS are available starting in February of 2000 and in June of 2002 for Aqua. Terra's orbit passes from north to south across the equator in the morning, while Aqua passes from south to north over the equator in the afternoon. Terra MODIS and Aqua MODIS view the entire Earth's surface every 1 to 2 days and acquire data in 36 discrete and narrow spectral bands (NASA, 2009). These spectral bands range in wavelength from 0.4 μm to 14.4 μm. The instrument's field of view is ±55° from nadir (Lopez et al., 2007).

MODIS Data Products from both Terra and Aqua include swaths of 500 m or 1 km resolution, snow cover daily, snow cover 8-day, and sea ice extent (http://www.nsidc.org 2010). Only Terra products were used for this research because approximately 70% of the band-6 detectors failed shortly after launch of the Aqua MODIS. Although there are alternative algorithms that can be used on other Aqua bands to extract snow cover, this may affect the accuracy of the Aqua snow cover products (Hall and Riggs, 2007).

3.3.1. Normalized Difference Snow Index (NDSI) to create MODIS daily snow map

There are four data fields of snow data within this product derived using NDSI: snow cover, fractional snow cover, snow albedo, and quality assessment for the data product. For this research, the snow cover field was used, which is essentially a daily snow cover map for the region.

The daily snow cover map is constructed by examining the many observations acquired from 1 day and mapped to each cell of the grid by a scoring algorithm (Riggs et al., 2006; Liang et al., 2008). The scoring algorithm is used to select an observation for 1 day (Tiat et al., 2001). The object of the scoring is to select the observation closest to local noon time (highest solar elevation angle), nearest to nadir, and with greatest coverage that was mapped into the grid cell in order to prevent viewing angle problems (Riggs et al., 2006). No information on the snow cover is obtained when it is dark because the visible spectrum is required (Tiat et al., 2001).

The scoring algorithm is:

$$0.5^{*}(\text{solar elevation}) + 0.3^{*}(\text{distance from nadir}) + 0.2^{*}(\text{observation coverage}) \quad (1)$$

(Riggs et al., 2006).

After this algorithm has been applied and the final daily image selected, the various types of land cover are coded.

The MODIS snow mapping algorithm applies the Normalized Difference Snow Index (NDSI), similar to the commonly used Normalized Difference Vegetation Index (NDVI), to map snow cover:

$$\text{NDSI} = (\text{Band4} - \text{Band6}) / (\text{Band4} + \text{Band6}) \quad (2)$$

Band 4: from 0.545 to 0.565 μm (visible)
Band 6: from 1.628 to 1.652 μm (short-wave infrared)

Pixels with an NDSI of equal to or greater than 0.4 are initially considered snow-covered. After this initial assignment, dense coniferous forests and water bodies are also corrected because they may fall into this NDSI range as well. Some snow may actually have an NDSI of less than 0.4 and these errors are corrected using a method combining NDSI and NDVI to distinguish between snow-covered and snow-free forests (Tiat et al., 2001; Klein and Barnett, 2003).

The MODIS Users Guide (Riggs et al., 2006) includes the numerical codes for daily snow cover data. The following codes were used in the data extraction process for this study: snow ("200"), no snow ("25"), cloud ("50"), and no data ("250") (Riggs et al., 2006). Even after this process, some misclassifications are not discovered and the final MODIS SCA output may contain errors. To address these errors, the accuracy of the product in the Mokelumne River Basin needed to be evaluated first.

3.3.2. Establishing MODIS accuracy

MOD10A1 has a reported accuracy of 93%, but this accuracy rate varies by land type, season, and snow conditions (Hall and Riggs, 2007). Cloud cover discrimination problems play a large role in MODIS errors due to spectral similarities (Hall and Riggs, 2007). For example, thin cirrus clouds that contain ice make snow particularly difficult to discriminate accurately (Hall et al., 2002; Klein and Barnett, 2003; Dozier et al., 2008).

There is no single way of assessing MODIS accuracy, but one common way is to compare in-situ measurements, such as snow pillows or snow courses, to MODIS snow cover data (Klein and Barnett, 2003). The stations used for the course and pillow data are ground measurements collected by the California Department of Water Resources (DWR) and located either within the Mokelumne Basin or not more than 7000 m (~3 miles) away from it. Some stations just outside the actual basin were used in order to increase the number of data points available for establishing MODIS accuracy. Comparisons between ground stations and MODIS data must be made with caution, however, due to the scale mismatch between point samples and MODIS grids (Garen and Marks, 2005).

Snow pillows are large (approximately 3 m^2) automated devices for calculating snow depth and snow water equivalent (SWE). The depth and SWE calculations are based on hydrostatic pressure created by overlying snow. They may be less accurate than the manually-determined snow courses but they do provide a continuous record unlike snow courses. A snow course is a series of approximately 10 snow sampled points spaced at approximately 30-m intervals. The snow courses in the Mokelumne River basin are measured irregularly throughout the snow season, approximately once per two or three months. Snow depth and/or SWE at these points are determined manually on-site and then simply averaged to arrive at one value for the whole snow course point. Thus, snow pillow snow depth and SWE values may or may not match the snow course values at a specific location (Government of Alberta, Canada, 2009).

Out of 129 points with snow course data from the years 2000 to 2009 January through April, 119 had snow and 10 did not. MODIS images identified 109 accurately as snow, 10 accurately as no snow, 10 as no snow when there was snow, and none as snow when there was no snow for an overall accuracy in this case of 92% (see Table 1).

For the second part of the MODIS accuracy testing, snow pillow data was also used in order to increase the number of data points tested. Adding an additional 117 snow pillow points to the snow course points for a total of 246 data points, MODIS images identified 222 accurately as snow, 11 accurately as no snow, 13 as no snow when there was snow, and none as snow when there was no snow for an overall accuracy in this case of 95% (see Table 1).

Thus, overall MODIS accuracy in the Mokelumne River Basin for the years 2000–2009 was 92–95% with MODIS unlikely to over represent snow. The mismatched points tend to be places that have little snow cover and/or late season snow cover (late March through May). Other researchers have also found this to be the case, particularly with the Terra satellite (Riggs et al., 2006; Liang et al., 2008; Parajka and Blöschl, 2008). In one study, MODIS accuracy was 96% when no snow was present, but only 41% with trace (<10 mm depth) amounts of snow (Ault et al., 2006). Four of the ten inaccurate MODIS readings when compared to snow course data occurred when there was <6.4 cm of snow (0.5 cm, 2.5 cm, 3.8 cm, 6.1 cm); this likely means that much of the MODIS pixel was indeed *not* covered in snow.

Table 1
DWR snow course and snow pillow data compared to MODIS snow cover data.

		MODIS		
		Snow	No snow	Total
DWR snow courses	Snow	109	10	119 (92%)
	No snow	0	10	10 (100%)
	Total	109 (100%)	20(50%)	129 (92%)
DWR snow courses & pillows	Snow	222	13	235 (94%)
	No snow	0	11	11 (100%)
	Total	222 (100%)	24 (46%)	246 (95%)

Some authors have found that MODIS accuracy depends on elevation, perhaps due to forest canopy (Hall et al., 2002; Jain et al., 2008). Canopies may cause inaccuracies in the data because some snow may land below the canopy and be reported as “no snow cover” when the canopy itself does not hold snow (Andreadis et al., 2009). The Mokelumne River is below the tree line; however, there was no evidence that the ten discrepancies were affected by station location or elevation.

4. Model development

The first step in the process of developing a linear regression model of TNF was to divide the river basin into seven equal area elevation zones. Then, two scripts were developed to extract the SCA (snow covered area) data from the daily MODIS images. Next, the missing daily snow-covered pixel values in an elevation zone due to clouds or noise were interpolated. Finally, a linear regression model was developed to determine which factors influence the TNF of the Mokelumne River and to predict short-term inflow (see Fig. 2).

4.1. Elevation zones

Dividing the river basin into elevation zones allowed separate analysis of the cloud cover within each zone on each day. Thus, for 1 day it is still possible to extract measured data for some low-cloud zones while not extracting data for zones covered by clouds. The reduction of the amount of interpolated data resulted in a more accurate model (cf. Parajka and Blöschl, 2008).

Based on distance and friction, water managers assumed it would take 7 days for melted snow in the upper zone to reach the reservoir, 6 days from the next zone, etc. For this reason the area was divided into seven zones (see Fig. 3). Zones of equal area were chosen to create equal spatial likelihood of snowmelt or accumulation. Conveniently, the equal area approach generated similar elevation differences in zones two through six. The 7-day linear estimation did not prove to be accurate; by experimenting and maximizing R^2, snow cover in the highest elevation zone has the maximum influence on TNF within an average of 11 days before reaching the reservoir.

4.2. Data extraction

Two Python ArcGIS (Esri, 1999–2010) geoprocessing scripts were used to extract the daily snow cover data. The first script extracted the Upper Mokelumne River basin from each daily MODIS swath; the boundaries of the basin extent were provided by EBMUD.

The second script used the output from the initial script as input to extract daily snow or cloud cover data within each of the seven elevation zones. This second script also assigned Julian dates to the input and provided daily snow and cloud cover data for each zone in a CSV (comma-separated values) format.

Using the output from these two scripts, a dataset was created that contained snow cover in each elevation zone for each day when cloud cover was less than 18% within at least one zone. For days when the cloud cover was more than or equal to 18%, the snow cover values were linearly interpolated for each zone. This interpolation was necessary because, from November to March, a specific elevation zone might have more than 18% cloud cover for a week or longer. Limiting cloud cover to 15% left stretches of more than 10 days with no data

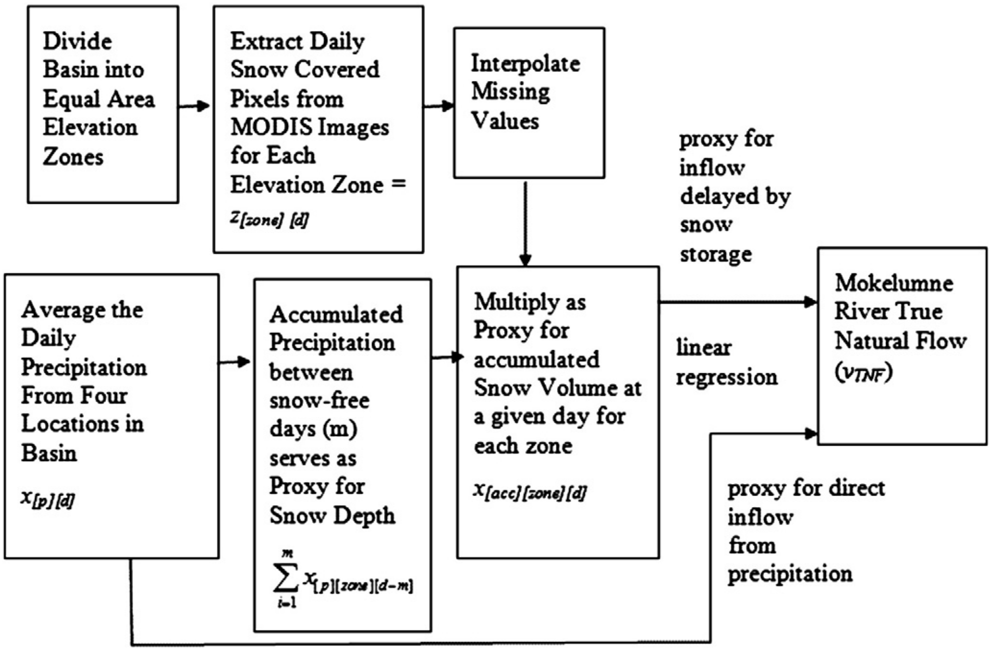

Fig. 2. Flowchart of Model.

due to cloud cover. 18% was the first cut-off where the "50" or cloudy stretches were less than or equal to 10 days long from 2000 to 2009.

4.3. Interpolation

Interpolation may be used to estimate missing data points, e.g. snow depth, snow water content, precipitation, or air temperature may be interpolated based on data from ground stations (Parajka and Blöschl, 2008). When less than or equal to 10 days in an elevation zone were cloudy, interpolation was used to replace those values based on data from clear days (cf. Dozier et al., 2008).

Linear interpolation was used for this model; for a series of cloudy days or days for which MODIS had no data, the snow cover value before the beginning of the cloud cover and after the end of the cloud cover were considered 'truths' and the values on cloudy days distributed evenly and equally between those two truths.

Two temperature values were used to validate the interpolation, the maximum temperatures from Salt Springs Powerhouse (1100 m) and Caples Lake (2430 m) (Fig. 2). When the interpolation created a series of days of melt or accumulation, temperature values were checked to ensure that either melt or accumulation was possible (cf. Lopez et al., 2007).

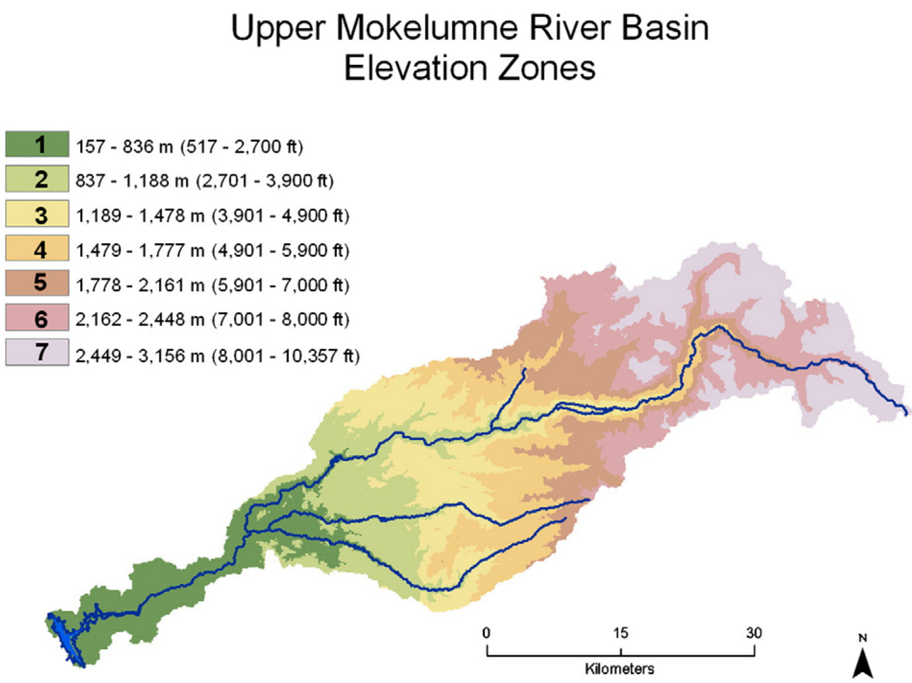

Fig. 3. Seven elevation zones in Mokelumne River Basin.

4.4. Regression Model

After the interpolated data for the years 2000–2009 were compiled, a statistical script using R software (R Development Team 2009) was developed to determine which factors contributed to TNF. The linear regression model formula is as follows:

$$y = \sum_{i=1}^{n} b_i x_i + c \tag{3}$$

where

- y is the value to predict—in our case TNF (y_{TNF}),
- b_i refers to the model factors to be fit
- x_i stands for the different independent variables that determine y_{TNF}, and
- c is an intercept representing Mokelumne flow not related to snow cover (summer flow).

The parameters precipitation (x_{p}), snowmelt or buildup ($x_{\text{melt/buildup}}$), snow covered area represented by pixels (x_{SCA}), and accumulated precipitation (as a proxy for snow depth, x_{acc}) were all tested individually in the model, as well as other meaningful combinations.

This partial regression helped to determine which factors were most important for determining TNF and fitting our model to the data by maximizing R^2 as a measure of explanatory power.

After developing this basic regression, the precipitation that was not stored as snow cover but that immediately ran off was considered. Since a lag time of 11 days had been determined (see Eq. (4)), all precipitation within this time adds to the TNF. However, only the current day precipitation ($x_{\text{[p][d]}}$) and the previous day precipitation ($x_{\text{[p][d-1]}}$) had a significant influence on TNF. Precipitation from 2 days ago ($x_{\text{[p][d-2]}}$) to precipitation from 10 days ago ($x_{\text{[p][d-10]}}$)were also examined.

There is rarely summer precipitation in this part of the Sierra Nevada. Therefore, a baseline precipitation model separately from the complex dynamics of the natural reservoir of snow cover could not be constructed. For this reason, precipitation alone had very little explanatory power ($R^2 = 0.05$).

Thus, it was assumed that the daily differences in snow cover would be a good predictor for TNF; melt would increase flow and the extension of SCA would slow it down:

$$x_{\text{[melt/buildup][zone]}} = z_{\text{[zone]}[d]} - z_{\text{[zone]}[d-1]} \quad \text{(for each elevation zone)} \tag{4}$$

where z is the number of snow-covered pixels in an elevation zone extracted from MODIS images, d is a given day, and $d-1$ refers to the previous day.

These values were calculated for each elevation zone. In the model for a given day, values from previous days for the different elevation zones according to the lag times were determined by maximizing R^2. This meant using pixels that had melt or buildup in elevation zone seven ($x_{\text{[melt/buildup][7]}}$) 11 days prior to the prediction date, pixels that had melt or buildup in elevation zone six ($x_{\text{[melt/buildup][6]}}$) 8 days prior the prediction date, etc. However, a model built upon these assumptions had an R^2 of only 0.08 even when precipitation that flowed directly into the Pardee Reservoir was taken into account.

The results so far could be interpreted to mean that both precipitation and snow melt typically occur during warmer, cloudy periods of time (Beniston et al., 2003; Räisänen, 2008). Because of this, differences in SCA can increase and decrease within a few hours.

Also, weather often changes significantly during the 11 day period when water is traveling down from the highest elevation zone; melt water could refreeze on a hillslope on its way down (Shirazi et al., 2009).

Learning from these model runs and hesitant to overfit the model with a large number of factors, a new approach was needed. Assuming most precipitation is not absorbed and flows through the watershed, reaching the Pardee Reservoir eventually, then snow simply slows down the water flow. Some water evaporates or drains into the ground, establishing the residual c.

In order to account for the accumulation of water in the area, a proxy for snow depth was used. To do this, the daily precipitation was summed (the average for the whole basin calculated from the four rain gauges, see Fig. 1) for each elevation zone. In order to estimate the stored water volume this proxy was multiplied with the number of snow-covered pixels determined from MODIS:

$$x_{[acc][zone][d]} = x_{[SCA][zone][d]} \sum_{i=1}^{m} x_{[p][zone][d-m]} \tag{5}$$

where m is the number of days on which snow was present without interruption within an elevation zone and m resets to zero when there is no snow cover in that elevation zone, $x_{[acc][zone][d]}$ is accumulated precipitation (proxy for depth) in an elevation zone on a given day, $x_{[SCA][zone][d]}$ is the snow covered area in number of pixels in an elevation zone on a given day, and $x_{[p][zone][d-m]}$ is the precipitation in an elevation zone for days on which snow was present.

Since the average was used as a first approximation, the amount of precipitation added for every elevation zone was exactly the same. This solution became necessary since there was not precipitation data available for each elevation zone (see Fig. 1). When MODIS satellite images determined that this particular elevation zone was snow-free as defined by zero snow covered pixels, the accumulated water was gone. Therefore, value for this zone was set to zero while others still accumulated snow.

Next, to improve the model further, increasing temperature and reduced the accumulated precipitation was taken into account. However, this strategy did not improve the model but instead overfit it because of the high intercorrelation among the variables. In addition, the reduction in snow volume was represented by the shrinking SCA. Snow melt is a highly complex process depending on solar radiation, air temperature, wind, aspect, albedo, etc., which require very thorough and locally specific studies (Marks and Dozier, 1992). In contrast, this research addressed whether it is possible to make reasonable predictions about water flow from snowmelt on the basis of general and easily available data.

The overall model presented by the formula:

$$y_{\text{TNF}} = \sum_{i=0}^{1} b_{i+1} x_{[p][d-i]} + \sum_{j=1}^{7} b_{j+2} x_{[acc][j]} + c \tag{6}$$

where

- y is the value to predict—in our case TNF (y_{TNF}),
- b_i refers to the model factors to be fit
- x_i stands for the different independent variables that determine y_{TNF}, and
- c is an intercept representing Mokelumne flow not related to snow cover (summer flow).

yields a reasonable R^2 of 0.71 (the snow depth proxy alone generates an R^2 of 0.67) if trained over the entire period where data are available.

5. Results and discussion

Table 2 contains the determined lag time as well as the b coefficients for the overall model predicting TNF.

These results show that the change in SCA has a somewhat different meaning in different elevation zones. In zone 1, snow cover

Table 2
Coefficients and time lag.

b_1 (precipitation on the day of predicted TNF)	4.097×10^{-1}	$p<0.001$
b_2 (precipitation 1 day before predicted TNF)	6.505×10^{-1}	$p<0.001$
b_3 (accumulated snow in zone 1, 1 day prior predicted TNF)	-2.091×1^{-4}	$p<0.01$
b_4 (accumulated snow in zone 2. 1 day prior predicted TNT)	1.267×10^{-5}	not significant
b_5 (accumulated snow in zone 3, 1 day prior predicted TNF)	1.749	$p<0.01$
b_6 (accumulated snow in zone 4,3 days prior predicted TNF)	-3.704×10^{-5}	$p<0.001$
b_7 (accumulated snow in zone 5, 5 days prior predicted TNT)	-2.035×10^{-5}	not significant
b_8 (accumulated snow in zone 6, 8 days prior predicted TNT)	-6.456×10^{-6}	not significant
b_9 (accumulated snow in zone 7, 11 days prior predicted TNT)	9.929×10^{-5}	$p<0.001$
c (intercept in m^3/second)	4.933×10^{1}	$p<0.001$

Table 3
Partial model results.

Precipitation only ($p<0.001$)	$R^2=0.05$
Precipitation and change in snow cover (melt and build-up)	$R^2=0.08$
Precipitation and change in snow cover and snow covered area (includes lag time)	$R^2=0.41$
Accumulated snow (proxy for depth) which includes SCA from MODIS	$R^2=0.67$
Precipitation and accumulated snow (proxy for depth) which includes SCA from MODIS	$R^2=0.71$

indicates very cold weather. For this reason, as assumed above, it signifies a reduction in TNF which is true for most elevation zones. Not all elevation zones have a significant influence on TNF. This is an effect of the high intercorrelation of the snow cover values in the elevation bins. Interestingly, zones 3 and 7 have a positive effect on TNF. They represent the stored water masses that change TNF in the spring time significantly.

This final model result in Fig. 4, uses the established elevation zones and time lags to establish on what day to use the snow cover from each elevation zone, e.g. the TNF for today will use the snow cover in $zone_7$ from 11 days ago in addition to the snow cover in elevation $zone_6$ from 8 days ago, etc. Based on these lag times and the fact that each elevation zone consists of an equal area within the river basin, the TNF is not a linear function. This is expected since TNF is based ultimately on snow depth, which is a volume and thus this function could be cubic. The large Pardee inflow surges (peaks in red) can be explained by large precipitation events (in blue) later in the season when precipitation is likely to be in the form of rain.

The regression coefficients are negative for every elevation zone except for the highest zone seven and mid-range zone three. For zone seven, this could be because snow actually slows down the water flowing into the Mokelumne River except at the highest elevation where the snowmelt flowing into the Mokelumne tends to be later in the season and sudden. Zone three is more difficult to explain and could be due to the intercorrelation among variables.

The y-intercept describes the Mokelumne flow not explained by snow, or the summer stream flow. This flow likely originates from groundwater. This factor was highly significant.

The partial results of the model (see Table 3) show which parameters determined TNF as well as which ones did not determine TNF. All of the partial models have significant factors; each factor alone is significant, but when combined some are not significant due to covariance (or intercorrelation).

Table 3 was created by trying different combinations of input to establish which ones influence TNF and which ones do not. As shown in Table 3, TNF is better explained by snow cover than by snowmelt (melt is defined by change in snow-covered pixels); the spatial extent of snow cover and lag time is more important in TNF than melt. This could be because melt and precipitation often occur simultaneously; cloudy days are not as cold as clear days in winter. It could also be because a MODIS pixel (500 m × 500 m) is too large to use to establish melt for specific pixels.

6. Conclusion

Climate change will affect snowpack and water supply systems in California, and this method for predicting daily snowmelt runoff is intended to help prepare for these changes for one river basin in the western Sierra Nevada. Using MODIS satellite, precipitation, and TNF data, we created a predictive model to forecast the flow into the Pardee Reservoir along the Mokelumne River.

The objective of this study was to see if MODIS snow cover images and precipitation data could be used to create a useful and repeatable

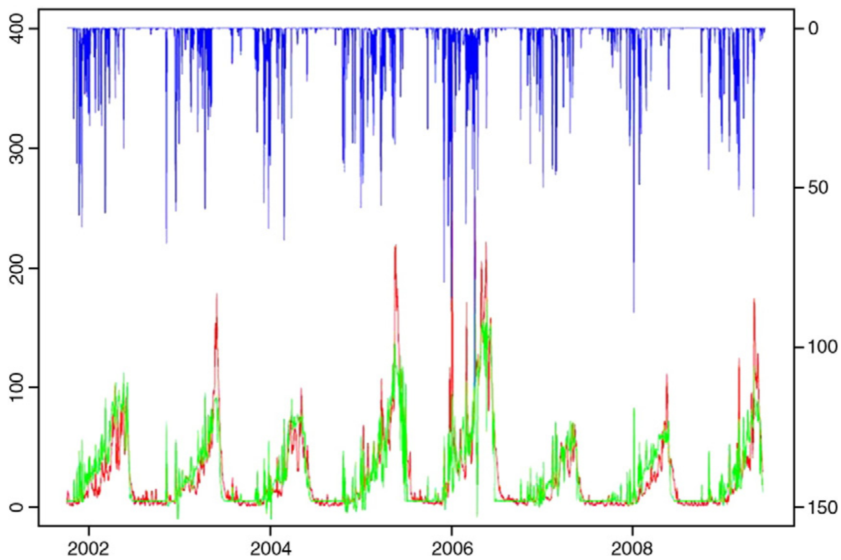

Fig. 4. TNF (red) and model predicted flow (green). Precipitation is in blue.

statistical hydrologic model predicting TNF in the Mokelumne River Basin. The investigation was based on comparing true natural flow (TNF) to the model output based on MODIS daily snow cover in seven different elevation zones and overall average precipitation for the entire river basin. The main implication of this research is that it is possible for water resource managers in the Mokelumne River Basin to use a daily model to help plan and support water operations; this is a particularly important goal with climate change altering our snow-based water supply systems, making historical records unreliable as a guide for future conditions.

MODIS accuracy was also calculated. Comparing the MODIS output at the grid level to California Department of Water Resources point location snow-depth measurements in this river basin for the years 2000 to 2009, MODIS was 92–95% accurate.

Unexpectedly, the original hypothesis that MODIS change in snow cover (melt and accumulation) would predict TNF did not hold true. Instead, the specific spatial extent of snow in each elevation zone combined with lag time and using precipitation accumulation as a proxy for snow depth predicts TNF. We hypothesize that this is because melt and accumulation occur at the same time—on cloudy days; thus, the fluctuation of the snow line was not meaningful in terms of TNF.

Next steps would include applying this model to a non-modified river such as the Cosumnes River just north of the Mokelumne River where the TNF would be actual and not calculated. Also, continuing to assess the model's ability to accurately predict the TNF by comparing model predictions with actual flow data would help determine the model's accuracy. Applying this model to other modified rivers to establish repeatability would be another worthwhile step, as would changing the threshold for snow in a zone to something larger than one snow covered pixel in the entire zone.

Acknowledgements

We would like to thank Seth Hiatt for MODIS and mapping support and Kevin Richards and Steve Wollmer at EBMUD for data and insights.

References

Andreadis, K.M., Storck, P., Lettenmaier, D.P., 2009. Modeling snow accumulation and ablation processes in forested environments. Water Resources Research 45.

Ault, T., Czajkowski, K.P., Benko, T., Coss, J., Struble, J., Spongberg, A., Templin, M., Gross, C., 2006. Validation of the MODIS snow product and cloud mask using student and NWS cooperative station observations in the Lower Great Lakes Region. Remote Sensing of Environment 105, 341–353.

Beniston, M., Keller, F., Goyette, S., 2003. Snow pack in the Swiss Alps under changing climatic conditions: an empirical approach for climate impacts studies. Theoretical and Applied Climatology 74 (1), 19–31.

Brekke, L.D., Miller, N.L., Bashford, K.E., Quinn, N., Dracup, J.A., 2004. Climate change impacts uncertainty for water resources in the San Joaquin River Basin, California. Journal of the American Water Resources Association 40 (1), 149–164.

Davis, R.E., Pangburn, T., Daly, S., Ochs, E., Hardy, J.P., Bryant, E., Pugner, P., 1999. Can satellite snow maps, ground measurements, and modeling improve water management and control in the Kings River Basin, California?—Efforts toward finding the answer. Western Snow Conference Proceedings, April 1999, Lake Tahoe, California, U.S. 1999, pp. 54–61.

Dettinger, M.D., Cayan, D.R., Meyer, M.K., Jeton, A.E., 2004. Simulated HydrologicResponses to Climate Variations and Change in the Merced, Carson, and American River Basins, Sierra Nevada, California, 1900–2099. Climatic Change 62, 283–317.

Dozier, J., Painter, T.H., Rittger, K., Frew, J.E., 2008. Time–space continuity of daily maps of fractional snow cover and albedo from MODIS. Advances in Water Resources 31, 1515–1526.

East Bay Municipal Utility District (EBMUD), 2004. Pretreatment and pollution prevention report. http://www.ebmud.com/sites/default/files/images/Wastewater/ww_pollution_prevention_annualreport_2004%5B1%5D.pdf (last accessed 4/15/2011).

EBMUD, 2007. All About EBMUD. http://ebmud.com/sites/default/files/pdfs/allaboutebmud.pdf (last accessed 4/15/2011).

EBMUD, 2008. Mokelumne Watershed Master Plan. http://ebmud.com/sites/default/files/pdfs/MWMPP1_0.pdf (last accessed 4/15/2011).

Esri, 1999–2010. ArcGIS 9.3. Esri, Inc., Redlands, California. Copyright (c) 1999–2010.

Garen, D.C., Marks, D., 2005. Spatially distributed energy balance snowmelt modeling in a mountainous river basin: Estimation of meteorological inputs and verification of model results. Journal of Hydrology 315, 126–153.

Government of Alberta, Canada, 2009. Environment Division, Edmonton, Alberta, Canada T5K 2G8. http://environment.alberta.ca/forecasting/FAQ/definitions.html (last accessed 10/18/09).

Hall, D.K., Riggs, G.A., 2007. Accuracy assessment of the MODIS snow products. Hydrological Processes 21, 1534–1547.

Hall, D.K., Riggs, G.A., Salomonson, V.V., DiGirolamo, N.E., Bayr, K.J., 2002. MODIS snow-cover products. Remote Sensing of Environment 83, 181–194.

Intergovernmental Panel on Climate Change (IPCC), 2007. Report. http://www.ipcc.ch/ (last accessed 12/12/2010).

Jain, S.K., Goswami, A., Saraf, A.K., 2008. Accuracy Assessment of MODIS, NOAA, and IRS data in snow cover mapping under Himalayan conditions. International Journal of Remote Sensing 29 (20), 5863–5878.

Jeffres, C.A., Klimley, A.P., Merz, J.E., Cech Jr., J.J., 2006. Movement of Sacramento sucker, Catostomus occidentalis, and hitch, Lavinia exilicauda, during a spring release of water from Camanche Dam in the Mokelumne River, California. Environmental Biology of Fishes 75, 365–373.

Kim, J., 2001. A nested modeling study of elevation-dependent climate change signals in California induced by increased atmospheric CO2. Geophysical Research Letters 28, 2951–2954.

Kim, J., 2005. A projection of the effects of the climate change induced by increased CO2 on extreme hydrologic events in the Western U.S. Climatic Change 68, 153–168.

Kiparsky, M., Gleick, P., 2003. Climate Change and California Water Resources: A Survey and Summary of the Literature. Pacific Institute for Studies in Development. http://www.pacinst.org/reports/climate_change_and_california_water_resources.pdf (last accessed 10/18/09).

Klein, A.G., Barnett, A.C., 2003. Validation of daily MODIS snow cover maps of the Upper Rio Grande River Basin for the 2000–2001 year. Remote Sensing of Environment 86, 162–176.

Leung, L.R., Ghan, S.J., 1999. Pacific northwest climate sensitivity simulated by a regional climate model driven by a GCM. Part II: 2×CO2 Simulations. Journal of Climate 12, 2031–2053.

Liang, T., Zhang, X., Xie, H., Wu, C., Feng, Q., Huang, X., Chen, Q., 2008. Toward improved daily snow cover mapping with advanced combination of MODIS and AMSR-E measurements. Remote Sensing of Environment 112, 3750–3761.

Lopez, P., Sirguey, P., Arnaud, Y., Pouyaud, B., Chevallier, P., 2007. Snow cover monitoring in the Northern Patagonia Icefield using MODIS satellite images (2000–2006). Global and Planetary Change 61, 103–116.

Marks, D., Dozier, J., 1992. Climate and energy exchange at the snow surface in the alpine region of the Sierra Nevada, 2, Snow cover energy balance. Water Resources Research 28, 3043-3054.

Miller, N.L., Bashford, K.E., Strem, E., 2003. Potential impacts of climate change on California hydrology. Journal of the American Water Resources Association 39 (4), 771–784, doi:10.1111/j.1752-1688.2003.tb04404.x.

Mote, P.W., Hamlet, A.F., Clark, M.P., Lettenmaier, D.P., 2005. Declining mountain snowpack in western North America. Bulletin of the American Meteorological Society 86 (1), 39–49, doi:10.1175/BAMS-86-1-39.

NASA (National Aeronautics and Space Administration), 2009. http://modis.gsfc.nasa.gov/(last accessed 10/18/09).

National Elevation Dataset, 2009. http://ned.usgs.gov/ (last accessed 3/02/10).

Palmer, P.L., 1988. The SCS snow survey water supply forecasting program: Current operations and future directions. *Proceedings Western Snow Conference*, pp. 43–51. Kalispell, MT, Colorado State University. Fort Colins, Colorado.

Parajka, J., Blöschl, G., 2008. The value of MODIS snow cover data in validating and calibrating conceptual hydrologic models. Journal of Hydrology 358, 240–258.

Pasternack, G.B., Wang, C.L., Merz, J.E., 2004. Application of a 2D hydrodynamic model to design of reach-scale spawning gravel replenishment on the Mokelumne River, California. River Research and Applications 20 (2), 205–225, doi:10.1002/rra.748.

Räisänen, J., 2008. Warmer climate: less or more snow? Climate Dynamics 30 (2), 307–319.

Riggs, G.A., Hall, D.K., Salomonson, V.V., 2006. MODIS Snow Products User Guide Collection 5. http://modis-snow-ice.gsfc.nasa.gov/sugkc2.html.

Shirazi, T., Allen, D.M., Quinton, W.L., Pomeroy, J.W., 2009. Estimating soil thaw energy in sub-Alpine tundra at the hillslope scale, Wolf Creek, Yukon Territory, Canada. Hydrology Research 40 (1), 1–18.

Tiat, A.B., Barton, J.S., Hall, D.K., 2001. A prototype MODIS-SSM/I snow-mapping algorithm. International Journal of Remote Sensing 22 (17), 3275–3284.

Trenberth, K.E., 1998. Atmospheric moisture residence times and cycling: implications for rainfall fates and climate change. Climatic Change 39, 667–694.

Vanrheenen, N.T., Wood, A.W., Palmer, R.N., Lettenmaier, D.P., 2004. Potential Implications of PCM Climate Change Scenarios for Sacramento-San Joaquin River Basin Hydrology and Water Resources. Climatic Change 62, 257–281.

Wilby, R., Wigley, T., 1997. Downscaling general circulation model output: a review of methods and limitations. Progress in Physical Geography 21 (4), 530–548, doi:10.1177/030913339702100403.