ENHANCING STREAM FORECAST MODELING IN CALIFORNIA MOUNTAIN WATERSHEDS

John Shupe[1], Christopher Potter[2], and Steve Klooster[1]

ABSTRACT

This study describes research using the NASA Ames version of the Carnegie-Ames-Stanford Approach (CASA) ecosystem model coupled with a surface hydrologic routing scheme called the Hydrological Routing Algorithm (HYDRA) to model the system of California watersheds. To assess CASA-HYDRA's ability to estimate actual water flows in both extreme and non-extreme precipitation years, we have compared the gridded model results on a monthly time scale with gage station data throughout the state. Considering the importance of the snow pack to the state's water regime, we are particularly interested in CASA-HYDRA's performance in California's mountain watersheds. Consequently, we made additional comparisons between our model and results generated from the USDA Snowmelt Runoff Model (SRM) applied to the Merced River watershed in Yosemite, CA. The results of both comparisons highlighted the need to enhance CASA-HYDRA's snowmelt algorithm. The efficacy of the degree-day approach has been demonstrated by the consistently solid performance of the SRM, so we have integrated this algorithm into CASA-HYDRA. After making this modification and increasing the amount of winter snow accumulation that the model generates, the timing of spring runoff has been improved. This research supports the development of a near-term (4-6 weeks) forecast system for snowmelt runoff and flood potential from California mountain watersheds.

INTRODUCTION

California contains the largest diversity of climates and hydrologic conditions of any other state in the country. This diversity is a great asset to the state, but it also poses significant challenges for water resource managers (Mount, 1995). Because of global patterns of atmospheric pressure circulating over the oceans, it is typical for the west coast of continents in the middle latitudes to experience wet winters and summer droughts. California is no exception. It receives 75% of its annual precipitation between November and March (Carle, 2004). Moreover, the annual variability of its water supply is as dramatic as its seasonal variability. In recoded history, the average annual runoff in the state has been as low as 18.5 billion m^3 in 1977 and as high as 166.5 billion m^3 in 1983 (Carle, 2004). In light of California's dynamic water regime, the snow pack is instrumental as the state's largest water storage "reservoir." From the Sierra Nevada in central California to the Klamath and Trinity mountains in the northern reaches of the state, over 50% of the total runoff is derived from snow (Mount, 1995). A great amount of study is dedicated to the forecasting of the state's annual water budget, including its snow pack. Computer modeling is an important component in this effort to manage one of California's most important resources.

This study describes research using both the SRM and the NASA Ames version of the CASA ecosystem model coupled with a surface hydrologic routing scheme called HYDRA to validate and generate daily forecasts for snowmelt runoff in California mountain watersheds. CASA is a global simulation model that combines multi-year satellite, climate, and other land surface databases to estimate the biosphere-atmosphere exchange of energy, water, and trace gases from plants and soils (Potter et al., 2003). Soil water balance in NASA-CASA is controlled by vegetation cover type and soil rooting zone depth settings that are both derived from NASA satellite data sets. Soil water moves through 3-4 surface layers that freeze and thaw in cold regions according to empirical temperature algorithms. Drainage outputs below the root zone are predicted. More details are available on the NASA-CASA project page at http://geo.arc.nasa.gov/sge/casa/.

CASA's drainage and evapotranspiration outputs are used to drive the HYDRA model. HYDRA is a fully distributed computer model that simulates the regional scale flow of water (Coe, 2000). HYDRA simulates a set of physical hydrologic processes including interception, infiltration, interflow, base flow, overland routing and channel routing. More details about HYDRA are also available on the NASA-CASA project page. We have been running HYDRA over the entire state of California at 1 km spatial resolution and a monthly temporal resolution.

Paper presented Western Snow Conference 2008

[1] John Shupe, California State University at Monterey Bay, Seaside, CA, John.W.Shupe@nasa.gov ;
[2] Christopher Potter, NASA Ames Research Center, Moffett Field, CA, Chris.Potter@nasa.gov ;
[1] Steve Klooster, California State University at Monterey Bay, Seaside, CA, Steven.Klooster@nasa.gov

The time period of our study has been 1982-1990. To validate CASA-HYDRA's performance, we compare the model's results against the actual gage flow from a set of 336 U.S. Geological Survey (USGS) gage stations.

In the initial results, HYDRA suffered from diminished performance in mountainous regions where most of the precipitation falls as snow. An analysis of the drainage data from CASA revealed that not enough snow was being created in the winter and that snow was melting too quickly in the spring. With the objective of improving CASA's snowmelt algorithm, we tested the SRM's degree-day method by applying the model to the Merced River basin in Yosemite, CA. The SRM is designed to simulate and forecast daily stream flow in mountain basins where snowmelt is a major runoff factor (Martinec, 1975). The SRM requires daily input for temperature, precipitation and snow cover extent and it has eight parameters which can either be derived from measurements or estimated by hydrological judgment (Martinec et al., 1998). The SRM's convincing performance affirmed the efficacy of the degree-day approach, so we integrated this algorithm into CASA. After an additional modification was made to increase snow accumulation in CASA, the performance of CASA-HYDRA significantly improved in snow pack dominated regions, such as the Sierra Nevada.

METHODS

SRM Study Area

The basin of the Merced River spans 832 km^2 and is situated on the western slope of the Sierra Nevada (Figure 1). The elevation range of the Merced basin is from 1178 m at the USGS gage station at Pohono Bridge to 3997 m at Mount Lyell. After dropping 871 m in 31.9 km with an average stream gradient of 27.3 m/km (Mast and Clow, 2000), the Merced River flows into the eastern end of Yosemite Valley at the Happy Isles USGS Hydrologic Benchmark Network gage station. In 1967 the HBM Network was created to provide long-term data on water quality and discharge for small undeveloped watersheds (Clow et al., 1996). The Merced River merely drops an additional 46 m as it meanders 11 km (4.2 m/km average stream gradient) though Yosemite Valley to the Pohono Bridge gage station at the western end of the valley (USGS Water, 2008).

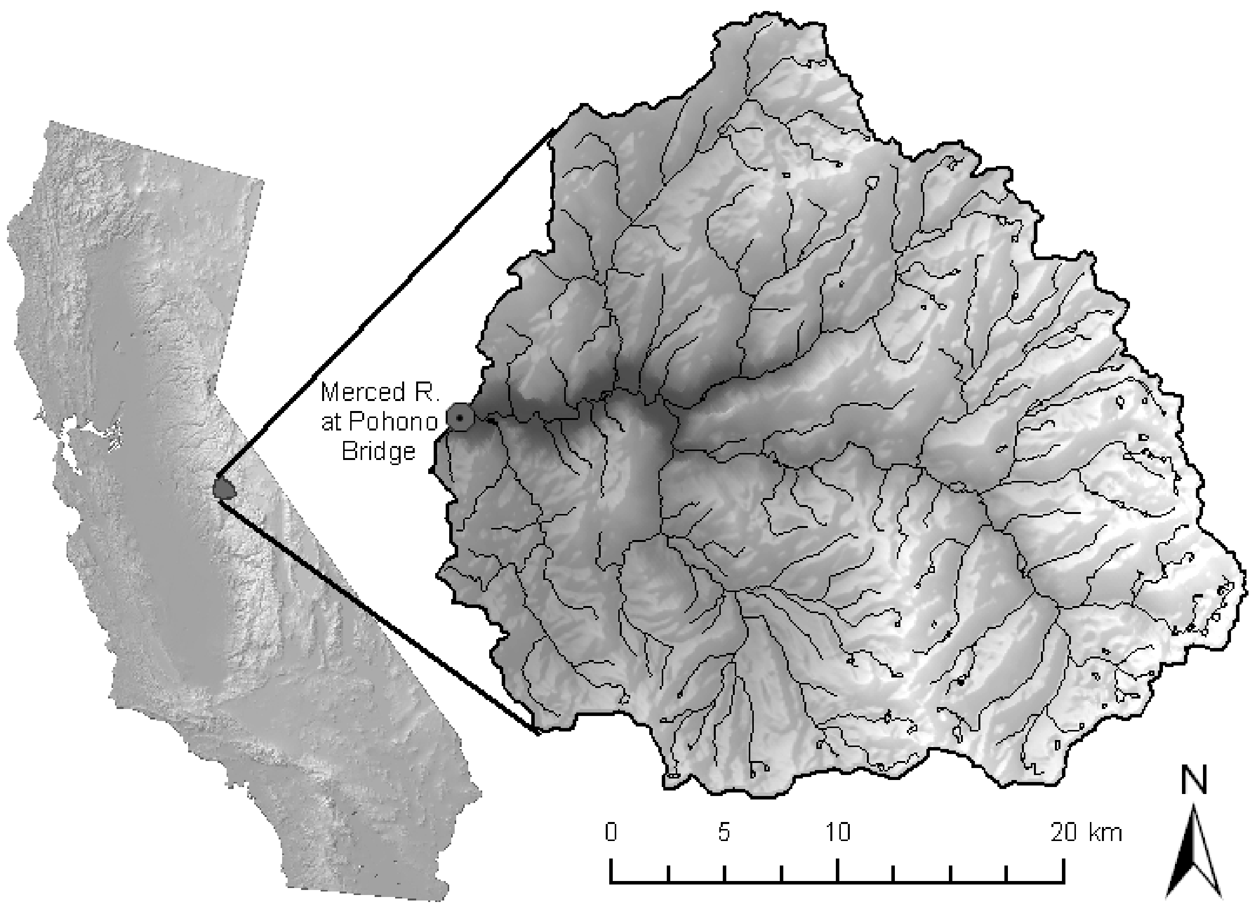

Figure 1. Merced River study area for the Snowmelt Runoff Model

The terrain of the Merced River basin is characterized by broad glaciated valleys and steep faces of exposed granite, while alpine cirques are prevalent in the headwaters. Less than one half of the basin area is forested and the remainder is covered by lakes, wetlands, meadows, rock outcrops, and talus slopes (Mast and Clow, 2000). Conifers are the most prevalent vegetation type, although there is a small percentage of deciduous trees, low-lying tundra plants and alpine meadow vegetation (Christensen et al., 2008). The basin's usually shallow and highly permeable soils are typically Inceptisols, which mostly developed on alluvial deposits and glacial tills derived from the local bedrock (Mast and Clow, 2000).

The basin's Mediterranean climate has cool, wet winters and warm, dry summers. Annual precipitation ranges from about 94 cm on the floor of Yosemite Valley to about 140 cm at treeline (~3,200 m), where the majority of the precipitation falls as snow (Mast and Clow, 2000). The snow pack dominates the water regime of the Merced basin. Accumulation typically occurs between October and April and the snow melts from April to June (Christensen et al., 2008). The mean annual discharge for the water years 1917-2007 on the Merced River at Pohono Bridge is 17.8 m^3/s. The watershed experiences a pronounced seasonal variation with a mean monthly discharge over the same period of time ranging from a low of 1.7 m^3/s in October to a high of 67.1 m^3/s in May (USGS Water, 2008).

California rivers are notorious for their tremendous range in discharge (Carle, 2004) and the Merced River is no exception. The lowest mean annual discharge at the Pohono gage station, for example, was only 3.6 m^3/s in the 1977 water year, which was one of the worst drought years in California. In contrast, the 1983 water year witnessed the largest mean annual discharge of 41.5 m^3/s. As further evidence of discharge variability, the Pohono station's minimum discharge of 0.15 m^3/s was measured on October 26, 1977, while the river's maximum discharge of 696.6 m^3/s was recorded on January 3, 1997 (USGS Water, 2008).

SRM Data

We ran the model with data from 1990, which enabled a comparison with the last year of results from our CASA-HYDRA study period (1982-1990). The SRM can be run for an entire year or for just a snowmelt season. We choose the later option, considering that our primary interest was how snowmelt is handled. 1990 was a dry year in California and the snowmelt season for the Merced basin occurred from mid-March through June.

The SRM documentation advises that watersheds should be split up into zones that encompass between 500-750 m of elevation (Martinec et al., 1998). Consequently, the roughly 2800 m elevation range of the Merced basin was divided into 4 zones. The extent of zone 1 (1178-1880 m) is limited to Yosemite Valley, which is 5% of the watershed. Zones 2 (1880-2580 m) and 3 (2580-3280 m) encompass 46% and 43%, respectively. The basin's headwaters are contained in zone 4 (3280-3997 m), which comprises 6% of the study area.

The SRM requires three input variables, which are temperature, precipitation, and snow cover extent. The daily snow cover maps that we used in our study were derived from AVHRR and GOES data. The processing of the satellite data was performed by a team at the National Operational Hydrologic Remote Sensing Center (Szeliga et al., 1995). These maps delineate the binary presence or absence of snow at 1 km resolution. We selected ten cloud free images that were evenly distributed throughout the snowmelt season and used them to generate snow depletion curves for each SRM elevation zone. Percentages of daily snow cover calculated from these depletion curves were then used as input into the model.

Ordinarily, the SRM is run with temperature and precipitation data from the nearest weather station. SRM interpolates these values to the hypsometric mean elevation of each zone by simply using lapse rates that the user provides as model input (0.65 ºC/100m is the default value). A different approach was used in this study. Preferring the interpolation algorithm used by DAYMET (Thornton et al., 1997), we used daily, 1 km resolution DAYMET temperature and precipitation data sampled along the hypsometric mean of each elevation zone to drive the model. For instance, using a NED 30 m digital elevation model (USGS DEM, 2008), we identified six points in zone 1 that corresponded to the hypsometric mean elevation of 1510 m (+/-5 m). We then collected the DAYMET data for each of these six points and used the average temperature and precipitation values as the model input for that zone.

SRM Results

The SRM overestimates the total water flow by less than 5% and the Nash-Sutcliffe efficiency is 0.859. Moreover, the hydrograph in Figure 2 also reflects a good correspondence of both flow timing and magnitude. The model's performance on the Merced watershed is consistent with most other published results, including a study done in the relatively near by Kings River watershed (Rango and Martinec, 1995). In light of the SRM's consistently good results, we felt comfortable using its degree-day algorithm to improve CASA's snowmelt performance.

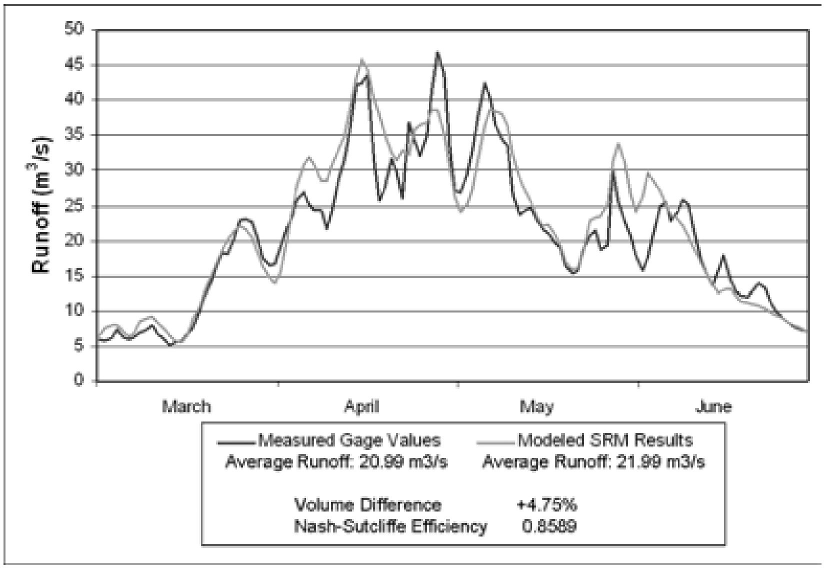

Figure 2. Measured vs. computed runoff for the Merced River

CASA Modifications

1. Degree-day Method. To improve CASA's snowmelt algorithm, we integrated the following degree-day calculation into the model;

$$S_m = a*T$$

Where S_m is potential snowmelt (cm), *a* is the degree-day factor (cm/ºC·day) and *T* is the degree-days (ºC·day). Essentially, the degree-day factor converts the number of degree-days into a potential snowmelt depth. For the degree-day factor, we are using values calculated by Dr. Bob Rice (2007), based on his studies in the Tuolumne and Merced watersheds.

CASA runs with monthly temperature data, which makes it problematic to calculate degree-days because using average monthly temperatures has the propensity to underestimate the true degree-day total. This is due to the fact that a temperature below the melt threshold will not subtract from the cumulative total of degree-days but it can lower the average temperature. In this study, we refer to this potential underestimation as the degree-day difference.

To develop a method for estimating the degree-day difference, we analyzed daily DAYMET temperature data for the months of November through February during the years 1980-2003. We collected data at four points from the Merced River watershed and five points from the American River watershed, which is a Sacramento River tributary that is west of Lake Tahoe. The points in each watershed were chosen so that they represented the elevation range of each basin, which is about 1,200 m to 4,000 m for the Merced and about 800 m to 3,000 m for the American. Additionally, the basins were chosen so that they captured the variability of the Sierra Nevada. The American River is in the northern Sierra where elevations are lower but precipitation is more plentiful. The Merced River watershed is in the southern Sierra, which

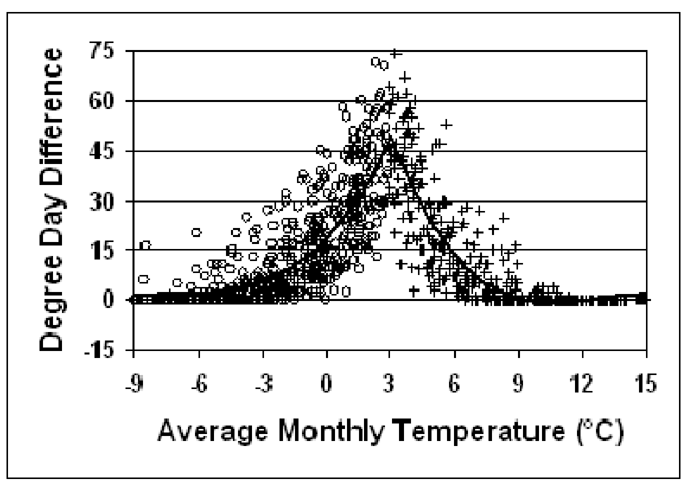

Figure 3. Estimating degree-day difference with monthly average temperature

experiences less precipitation but it contains the highest elevations seen in the continental United States.

Our analysis revealed that two independent variables explained most of the variance in the degree-day difference. These covariates are the average monthly temperature and the monthly temperature range. Instead of a multivariate analysis, we did a straightforward approach of looking at these independent variables individually. The scatter plot in Figure 3 shows the correlation between average monthly temperature and the degree-day difference. There is an obvious point of inflection at 3 ºC, which is the threshold that we used to calculate degree-days since it is the typical temperature at which snow melts early in the season (Martinec et al., 1998). Average monthly temperature by itself explains approximately 0.62 of the variance to the left of the inflection point and 0.74 to the right of it.

It would have been natural to use standard deviation as the second independent variable, but since CASA already has minimum and maximum monthly temperatures as model inputs, the monthly temperature range is used as a proxy. Figure 4 shows that, by breaking up the scatter plot in Figure 3 by monthly temperature range, we are able to estimate the degree-day difference with greater accuracy. Figure 4a contains data for the points that have a monthly temperature range between 0-10 °C. Similarly, the monthly temperature ranges for Figures 4b and 4c are 11-15 °C and 16-20 °C, respectively. The pattern that these graphs reveal is that the degree-day difference is more pronounced as the average monthly temperature range increases.

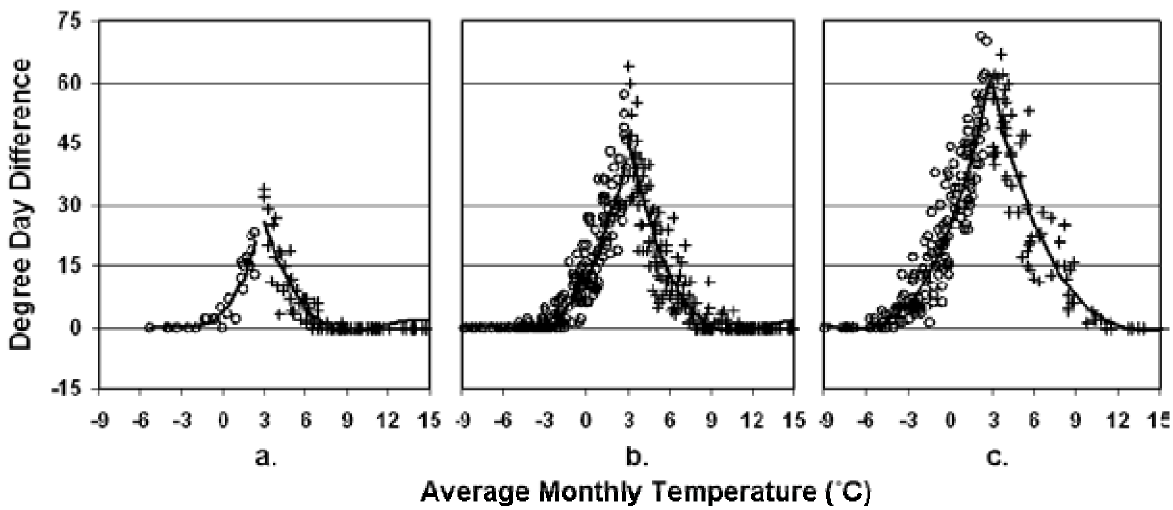

Figure 4. Estimating degree-day difference with two covariates

Between 0.7518 and 0.8702 of the variance in the degree-day difference is explained by taking monthly temperature range into account in conjunction with average monthly temperature. These results and the root mean squared error values are summarized in Table 1. Table 1 also includes the exponential functions that we use to calculate the degree-day difference in our model. Although it is not represented in Figure 4, there is an additional category for a monthly temperature range over 21 °C. Such a wide range in temperatures only occurred in months that averaged less that 3 °C, so there is no data for the scenario where the average monthly temperature is greater than 3 °C.

Table 1. Estimating the degree-day difference

		Less than 3°C			Greater than 3°C		
		Degree-day Difference =	DF Adj. R^2	RMSE	Degree-day Difference =	DF Adj. R^2	RMSE
Monthly temp. range (°C)	0-10	$-1.09 + 5.3e^{(x/1.68)}$	0.847	2.94	$-0.52 + 201.48e^{(x/-1.59)}$	0.8481	2.92
	11-15	$-3.02 + 15.18e^{(x/2.65)}$	0.8275	5.62	$-0.8 + 222.38e^{(x/-2.09)}$	0.8702	5.12
	16-20	$-5.65 + 29.53e^{(x/3.4)}$	0.8694	6.59	$-4.83 + 142.21e^{(x/-3.89)}$	0.8351	8.92
	21+	$-7.5 + 41.33e^{(x/5.39)}$	0.7518	7.33	N/A		

In summary, the average monthly temperature is used to make an initial degree-day calculation. Then a set of regression formulas are used to compute what degree-day difference needs to be added to the initial degree-day calculation so that it better approximates the true degree-day amount (*T*). This degree-day total can then be multiplied by the degree-day factor (*a*) in order to calculate the potential melt depth (S_m). CASA keeps track of the snow pack depth (as snow water equivalence) for each pixel, so the potential melt depth is then subtracted from this amount and the result is added to the runoff total for that pixel.

2. Snow-Rain Threshold. After making the modification to the snowmelt algorithm, a further analysis of CASA-HYDRA's results indicated that the model was not creating a large enough snow pack during the winter months. We first addressed this problem by increasing CASA's snow-rain threshold temperature from 0 ˚C to 1 ˚C. Research by Auer (1974) suggests that we could have gone as high as 2.2 ˚C. However, we are in the preliminary stages of this work so we decided upon a more conservative increase.

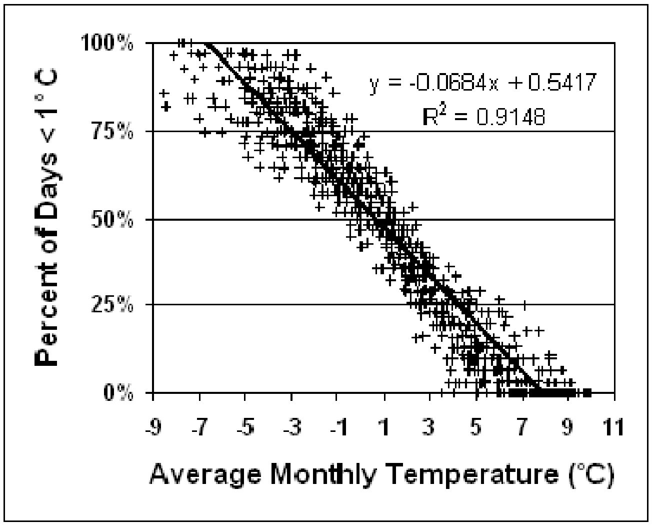

Figure 5. Calculating a percentage of snow and rain

We added a further modification to CASA's snow-rain threshold algorithm based on a regression analysis of average monthly temperatures and the percent of days within each month that are below the snow-rain threshold. The temperature data that we used is what was described in the previous section of this paper. Previously, the model would designate all the precipitation in a month as either all snow or all rain, depending on the average monthly temperature. Using a linear regression formula derived from the scatter plot in Figure 5, the model now divides the monthly precipitation into a percentage of snow and rain. Precipitation that occurs near or slightly above freezing naturally falls as a combination of snow and rain (Auer, 1974) and this new approach loosely mimics that dynamic. We are, however, also exploring ways to incorporate the Auer curve into CASA's snow-rain threshold algorithm.

An assumption that is inherent in using the percent of days below and above the snow-rain threshold to partition precipitation into rain and snow is that precipitation has an equal chance of falling on any day of the month. However, there can be patterns that violate this assumption. For instance, a cursory analysis of DAYMET precipitation data showed that only 10% of precipitation events in November occur within the first week of the month. California's rainy season ramps up during November, so it is not surprising that precipitation events would be more frequent later in the month. Another pattern we encountered is that daily average temperatures are roughly 0.5 ˚C (in December) to 2.5 ˚C (in November) cooler on days that have precipitation events.

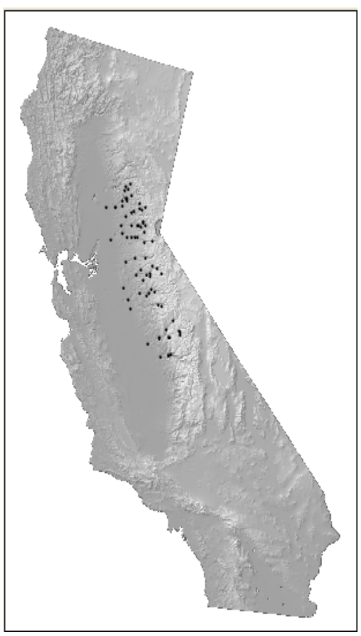

Figure 6. Sierra Nevada

RESULTS

Although CASA's snow-rain threshold algorithm could benefit from further refinements, our initial modifications have significantly improved the model's performance. With new drainage data from CASA, HYDRA performs much better in the snow dominated Sierra Nevada mountain range. The Sierra Nevada occupies one fifth of the state and it accounts for the majority of the snow pack (Carle, 2004). We used 73 USGS gage stations to validate CASA-HYDRA's results in this region (Figure 6).

The hydrograph in Figure 7 shows CASA-HYDRA's original results for the Sierra Nevada study area. The value on the y-axis is the average flow for the 73 gages stations. A statistical analysis yielded a coefficient of determination (R^2) of 0.504 and a Nash-Sutcliffe efficiency (E) of 0.426. Additionally, the relative root mean squared error (RMSE) is 84%. CASA-HYDRA underestimates the combined flow of all nine study years (1982-1990) by 7%.

An inspection of the hydrograph reveals that the model's discharge peaks are occurring too early in the year. This issue is the result of CASA producing too small of a snow pack in the winter and melting it away too quickly in the spring. The premature peaks are particularly obvious in 1987-1990, however this pattern is also evident in other years. In 1984, for instance, CASA-HYDRA overestimated the peak that occurred very early in the winter and, as a result, there was not enough runoff available to match the secondary peak that occurred later in the spring.

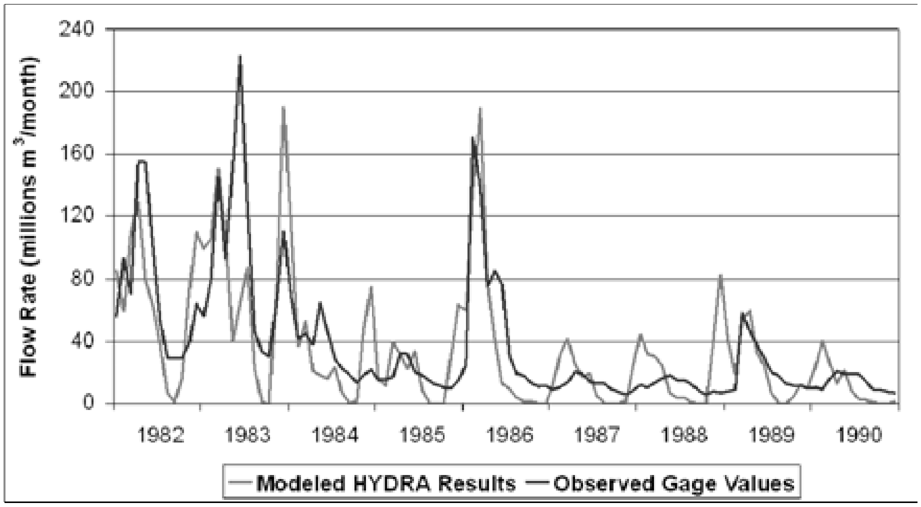

Figure 7. Hydrograph of the original CASA-HYDRA results

Figure 8 shows CASA-HYDRA's most resent results for the set of 73 gages in the Sierra Nevada. The peaks in the hydrograph are a much better fit and the statistical tests reflect that improvement. The R^2 increased to 0.719 and, similarly, E improved to 0.646. The relative RMSE dropped to 66%, which is still fairly high because of the inaccuracy in flow magnitude. Lastly, CASA-HYDRA now overestimates the combined flow of this region by 18%. This result is actually an improvement over the 7% underestimate from the original model run since the model does not currently account for human diversions, which are numerous in the Sierra Nevada.

Although the flow timing has significantly improved, the flow magnitude is still not matching up that well. One explanation is that the water resources of the Sierra Nevada are highly managed. In fact, only 5 of the 73 gages that we used in our study do not have significant upstream dams or diversions. Dams have the effect of moderating stream flow oscillations and diversions remove water from the natural channels that CASA-HYDRA simulates. It is

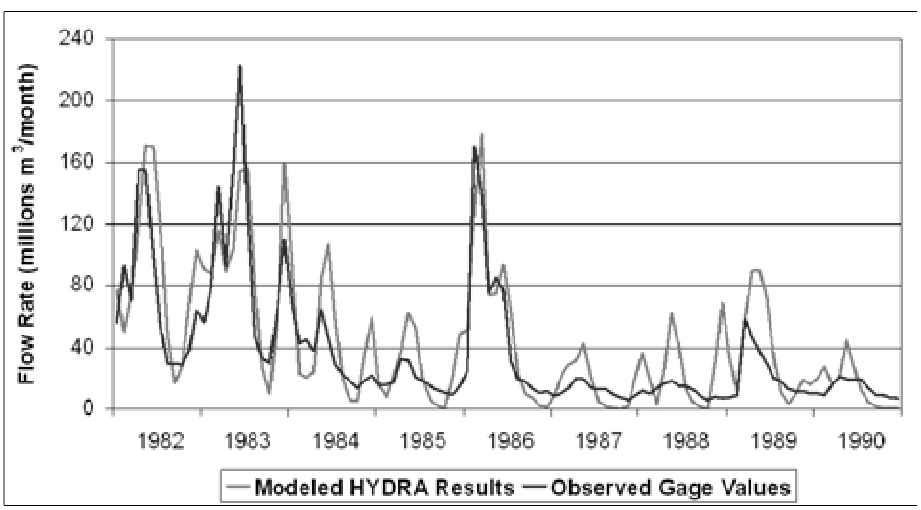

Figure 8. Hydrograph of the improved CASA-HYDRA results

particularly revealing that the magnitude issue is more pronounced during dry years, when a greater percent of water is being held back and diverted by human intervention. Because of the prevalence of dams and diversions in this region, for some select watersheds we intend to compare the model results against gage data that has been adjusted for natural discharge.

As anecdotal evidence that CASA-HYDRA performs well on natural watersheds, Figure 9 shows the hydrograph for one of the five stations that do not have significant upstream dams or diversions. This gage station on the Merced River at Pohono Bridge is the same one that we used to compare against the SRM results. The fit of the flow timing is extremely good and the flow magnitude is reasonably accurate in the majority of the study years. Correspondingly, the R^2 is 0.816 and E is 0.735. CASA-HYDRA underestimates the combined flow by 30%, which is largely due to substantial underestimates of the peak flow in just a few of the years. As a consequence, the relative RMSE is still quite high at 75%.

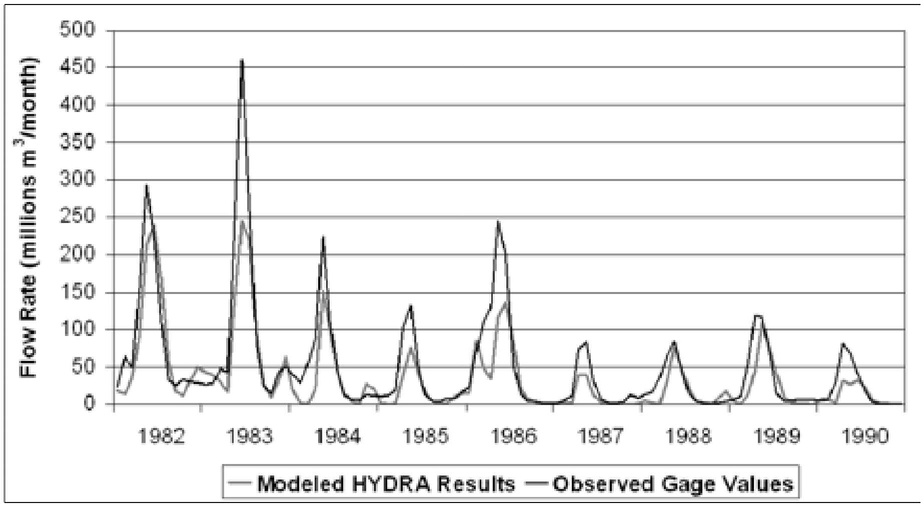

Figure 9. Hydrograph for the Merced River at Pohono Bridge

SUMMARY

To help manage California's valuable water resources, we are developing CASA-HYDRA as a near-term (4-6 weeks) forecast system for snowmelt runoff and flood potential from California mountain watersheds. California's snow pack, which is concentrated in the Sierra Nevada, is a crucial component of the annual water budget since it effectively acts as the state's largest reservoir. In this study, we have made improvements to CASA-HYDRA's snowmelt and snow-rain threshold algorithms so that it will more accurately model snow dominated regions such as the Sierra Nevada. The SRM performed well with a simple degree-day approach, which encouraged us to use the same method for CASA's snowmelt algorithm. Incorporating this method hinged upon the development of a procedure for estimating degree-days with monthly temperature data. We have also modified the snow-rain threshold algorithm so that the modeled snow pack more accurately simulates natural conditions. This is a technique that we are continuing to refine. As a consequence of these changes, CASA-HYDRA's results have significantly improved.

ACKNOWLEDGMENTS

This work is supported by NASA Gants from Applied Science, Terrestrial Ecology, and Earth Observing System Interdisciplinary Science programs.

REFERENCES

Auer, A.H. 1974. The rain versus snow threshold temperatures. Weatherwise, 27(67).

Carle, D. 2004. California Natural History Guide: Introduction to Water in California. University of California Press, Berkeley and Los Angeles, California.

Christensen, L., C.L. Tague, and J.S. Baron. 2008. Spatial pattern of simulated transpiration response to climate variability in a snow dominated mountain ecosystem. Hydrological Processes, Early View.

Clow, D.W., M.A. Mast, and D.H. Campbell. 1996. Controls on surface water chemistry in the upper Merced River basin, Yosemite National Park, California. Hydrological Processes, 10(5): 727-746.

Coe, M.T. 2000. Modeling terrestrial hydrological systems at the continental scale: Testing the accuracy of an atmospheric GCM. Journal of Climate, 13: 686-704.

Martinec, J. 1975. Snowmelt-runoff model for stream flow forecasts. Nordic Hydrology, 6(3).

Martinec, J., A. Rango, and R. Roberts. 1998. Snowmelt Runoff Model (SRM) User's Manual. Geographica Bernensia, 35.

Mast, M.A. and D.W. Clow. 2000. Environmental characteristics and water-quality of Hydrologic Benchmark Network stations in the Western United States. U.S. Geological Survey Circular 1173-D, 115 p.

Mount, J. 1995. California Rivers and Streams: The Conflict between Fluvial Process and Land Use. University of California Press, Berkeley and Los Angeles, California.

Potter, C., S. Klooster, R. Myneni, V. Genovese, P. Tan, and V. Kumar. 2003. Continental scale comparisons of terrestrial carbon sinks estimated from satellite data and ecosystem modeling 1982-1998. Global and Planetary Change, 39: 201-213.

Rango, A. and J. Martinec. 1995. Climate effects on future runoff regimes of Pacific mountain tributaries. Proceedings of the Conference on Water Resources and Environmental Hazards: Emphasis on Hydrologic and Cultural Insight in the Pacific Rim.

Rice, R., R. Bales, T. Painter, and J. Dozier. 2007. Snowcover Along Elevation Gradients in the Upper Merced and Tuolumne River Basins of the Sierra Nevada of California from MODIS and Blended Ground Data. Proceedings of the 75th Annual Western Snow Conference.

Szeliga, T., M. Allen, B. Choy, and T. Carroll. 1995. Historic Satellite Data 1990-1995. National Operational Hydrologic Remote Sensing Center.

Thornton, P.E., S.W. Running, and M.A. White. 1997. Generating surfaces of daily meteorological variables over large regions of complex terrain. Journal of Hydrology, 190: 214-251.

USGS DEM. 2008. USGS: The National Map Seamless Server. http://seamless.usgs.gov/index.php (Last Accessed: May 27, 2008).

USGS Water. 2008. USGS Real-Time Water Data for the Nation. http://waterdata.usgs.gov/nwis/uv. (Last Accessed: May 27, 2008).