Long-Term Vulnerability Assessment and Adaptation Planning for the San Francisco Public Utilities Commission Water Enterprise

Technical Report 2: Hydrologic Modeling Module

Prepared for:

Water Research Foundation and San Francisco Public Utilities Commission

Prepared by:

University of Massachusetts Amherst, Hydrosystems Research Group

August 2021

Table of Contents

LIST OF TABLES

LIST OF FIGURES

Purpose

This report describes the methods and calibration of the hydrologic module component of the Long-Term Vulnerability Assessment and Adaptation Planning for the San Francisco Public Utilities Commission Water Enterprise. The hydrologic module of the Vulnerability Assessment is comprised of hydrologic models spanning three hydrologic regions (i.e., Upcountry, Alameda (or Eastbay), Peninsula systems) and using two hydrologic modeling methods: the Sacramento Soil Moisture Accounting (SAC-SMA) model and the Precipitation Runoff Modular System (PRMS).

All of the modeling components for the three regions have been completed. The input data and the computational tools (all relevant scripts) needed to run the model have been transferred to SFPUC.

Integration of the hydrologic models into the Vulnerability Assessment will be straightforward, with scripts written in R or Python as needed to facilitate translating climate assumptions to inflows into reservoirs for use in the dependent water system model.

1. Introduction

One of the dominant operational drivers of a water system is water availability, which varies in both time and space, and at multiple temporal scales. Water availability in the San Francisco and Sierra Nevada region in particular varies significantly within the year, from year to year, and is increasingly expected to change over the long term as climate changes, yet in uncertain ways. Understanding and representing ranges of potential future hydrologic conditions is thus a core component to the Long-Term Vulnerability Assessment and Adaptation Planning for the San Francisco Public Utilities Commission (SFPUC) Water Enterprise project ("Vulnerability Assessment"). Representations of hydrologic processes is needed specifically to map uncertain climatic conditions into uncertain reservoir inflow conditions and inflow estimation capabilities.

The SFPUC Regional Water System (RWS) operations are directly impacted by hydrologic processes in three regions: the Upper Tuolumne River watershed in the Central Sierra Nevada ("Upcountry"); Alameda Creek watershed in the East Bay ("Alameda"); and the Peninsula watersheds ("Peninsula"). The principal subwatersheds comprising each of these three regions are shown in Figure 1-1. To represent hydrologic processes in each of these contributing subwatersheds for the Vulnerability Assessment, a suite of hydrologic models was developed for each region. Together, these models comprise the Hydrology Module of the Vulnerability Assessment, as described in the Detailed Analytical Design and Work Plan.

For each of these regions, this report summarizes the importance of the region, hydrologic modeling methods, calibration results, and hydrologic drought analysis. In the Vulnerability Assessment, the hydrologic models will be used to evaluate the hydrologic response of the regions − streamflow in particular − to various climate assumptions. The integration of hydrologic models into the Vulnerability Assessment are discussed last.

The organization of the hydrology module into three different regions with different hydrologic models is a matter of convenience and legacy efforts to re-configure the previous Upcountry hydrologic model. Organizing around three different regions is convenient in that the model development process can be divided amongst hydrologists. It also recognizes the unique hydro-geographic characteristics of each region. The use of different modeling methods (described below) resulted from earlier preliminary investigations of modeling options. It was decided to retain and extend the existing model of the Upcountry region (i.e., PRMS) and develop new models for the Alameda and Peninsula regions using SAC-SMA. Though this regionalization is worth noting, from a modeling perspective it is of little practical importance, as all models will be integrated into the hydrologic module when used with the water system model. The climate inputs (precipitation and temperature) and model outputs (inflows to reservoirs) will be uniform and consistent.
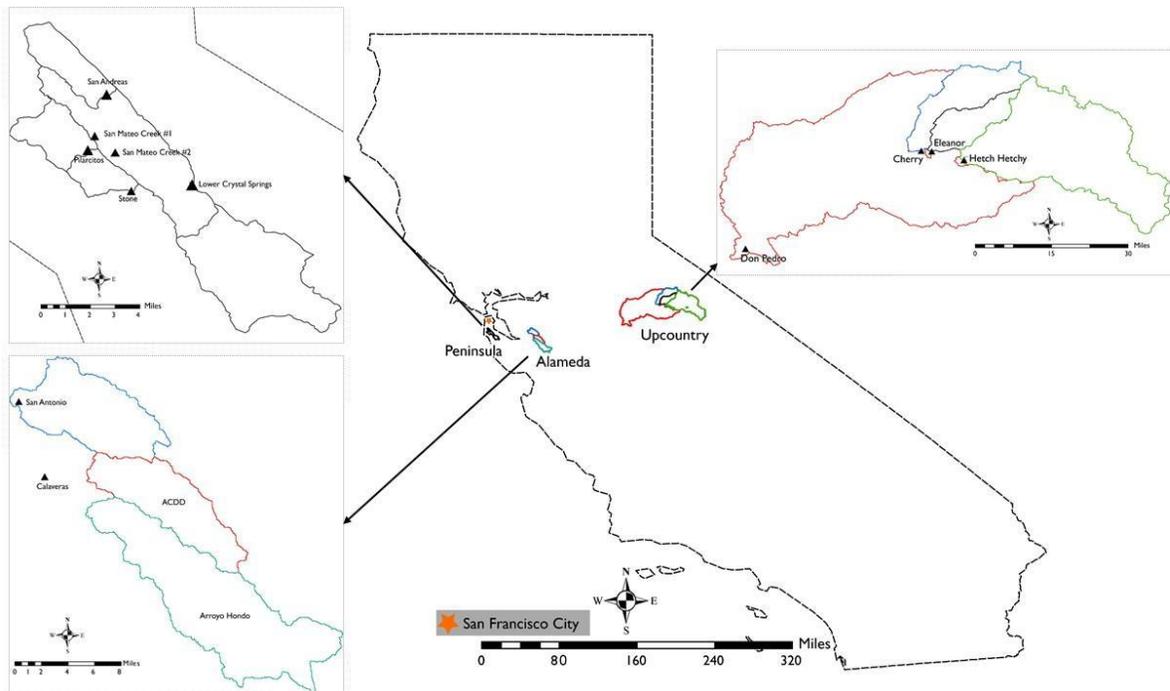
Figure 1-1 Map showing watersheds for the three regions that SFPUC receives waters from.

2. Overview of methods

Precipitation-Runoff Modeling System (PRMS) (Markstrom, et al., 2015) models were developed for the Upcountry region, while Sacramento Soil Moisture Accounting (SAC-SMA) (NOAA, 2002) models were developed for the Alameda and Peninsula regions, as described below. Though these modeling tools differ in their methodological approaches, their overall structure and purpose are similar. Hydrologic processes are calculated for discrete areas, called Hydrologic Response Units (HRUs), within which physical characteristics are assumed homogeneous. Inputs to the hydrologic models (i.e., to each HRU) include meteorological conditions (daily precipitation and temperature, and other conditions depending on

the specific model), land use / land cover information (which may be implicit in a model parameter rather than explicitly accounted for).

For each model, necessary inputs were obtained from historical records as available, with model calibration used to estimate unknown parameters based on observed runoff. Outputs include several hydrologic response variables, including surface/subsurface flow, base flow, actual evapotranspiration, and snow water equivalent. Simulated streamflow can then be used directly or aggregated in a range of systems related processes and decisions (e.g., inflow to reservoirs, input to determining water year type, etc.).

Calibration of each model is a significant and important part of the hydrologic model development process. There are often many unknown parameters in a hydrologic model, which must be estimated through some systematic approach. Typically, optimization is employed for this model parameter estimation process, as optimization methods, by design, can help identify the best combinations of values to achieve some performance objective. In each of the models developed, evolutionary algorithms were used for the calibration process. With the evolutionary algorithm approach, initial random sets of hydrologic parameters are successively improved by identifying which combinations of hydrologic parameters result in a better performing hydrologic model. Several evolutionary algorithms exist, which differ primarily by their strategy for improving parameter values through successive iterations.

In each calibration process used, the goal is to estimate the parameter values that result in the best match between observed and simulated runoff at locations with historical observations. Hydrologic calibration often uses some combination of various measures of model performance. Other metrics were also used in the calibration process, but metrics differ between model types.

Though metrics used for calibration targets differed between models, calibration results are described consistently using Nash-Sutcliffe Efficiency (NSE) and Kling-Gupta Efficiency (KGE), which indicates overall fit, and percent bias, which provides some measure of whether the model produces more (positive bias) or less (negative bias) compared to observations. Further explanation of calibration metrics is available in appendix F.

3. Upcountry Hydrologic Model
3.1. Introduction
The Tuolumne River and its tributaries provide the water in the Upcountry watersheds in the Sierra Nevada. Three major reservoirs located in the Upcountry watersheds − Cherry, Eleanor and Hetch Hetchy − account for most of the storage in the SFPUC system. The SFPUC is the junior water right holder on the Tuolumne River at Don Pedro Reservoir, while the senior water rights holders at Don Pedro are the Modesto and Turlock Irrigation Districts (MID and TID). An agreement with MID and TID allowed for the creation of the Water Bank, a virtual reservoir in Don Pedro Reservoir wherein the SFPUC can meet its water rights obligations in advance of inflows. A hydrologic model of the Upcountry region, partially built by SFPUC and extended by UMass, will help identify vulnerabilities not only to immediate water supplies but also to Water Bank operations.

3.2. Methodology

3.2.1. PRMS

The Precipitation-Runoff Modeling System (PRMS) is a deterministic, distributed-parameter, physical-process-based hydrologic model (Figure 3-1) developed and maintained by the U.S. Geological Survey (USGS). PRMS is developed with modules that are used to simulate various hydrologic processes. A given process can be represented with several modules, where each module represents an alternative conceptualization or approach to simulation of that specific process.

PRMS-IV, the version adopted in this study and abbreviated as PRMS here, accepts three files as input: a control file, a parameter file, and a data file. The control file initializes the PRMS model by activating necessary modules and contains the absolute paths for the parameter file and the data file. The parameter file specifies dimensions and parameters required for a PRMS simulation (i.e., the parameter values for each HRU). The data file contains the time-series data used as the input for the model; i.e., precipitation and temperature time series. Figure 3-2 shows the computational sequence within PRMS.

Input to the PRMS model includes fixed parameters and time series input data for each HRU. There are options for more than one hundred parameters in PRMS for each HRU. For the Upcountry PRMS models, 43 parameters were selected as important for calibration. Input time series data include daily precipitation (P), minimum temperature (T_{min}), and maximum temperature (T_{max}). Table 3-1 shows area and number of HRUs associated with each Upcountry subwatershed.

To facilitate running the PRMS models for both model calibration and the many climate scenarios anticipated for the Vulnerability Assessment, a wrapper library for PRMS called *prmsR* was developed using the R scripting language. *prmsR* generates the necessary input text files (control file and parameter file), runs PRMS, and parses the generated output files. In addition, *prmsR* can parse control files and parameter files from previously-run scenarios, such as those generated previously by SFPUC for the Hetch Hetchy/Cherry-Eleanor regions. This tool will be critical for both the model calibration and a large number of PRMS runs with climate change scenarios in the Vulnerability Assessment.
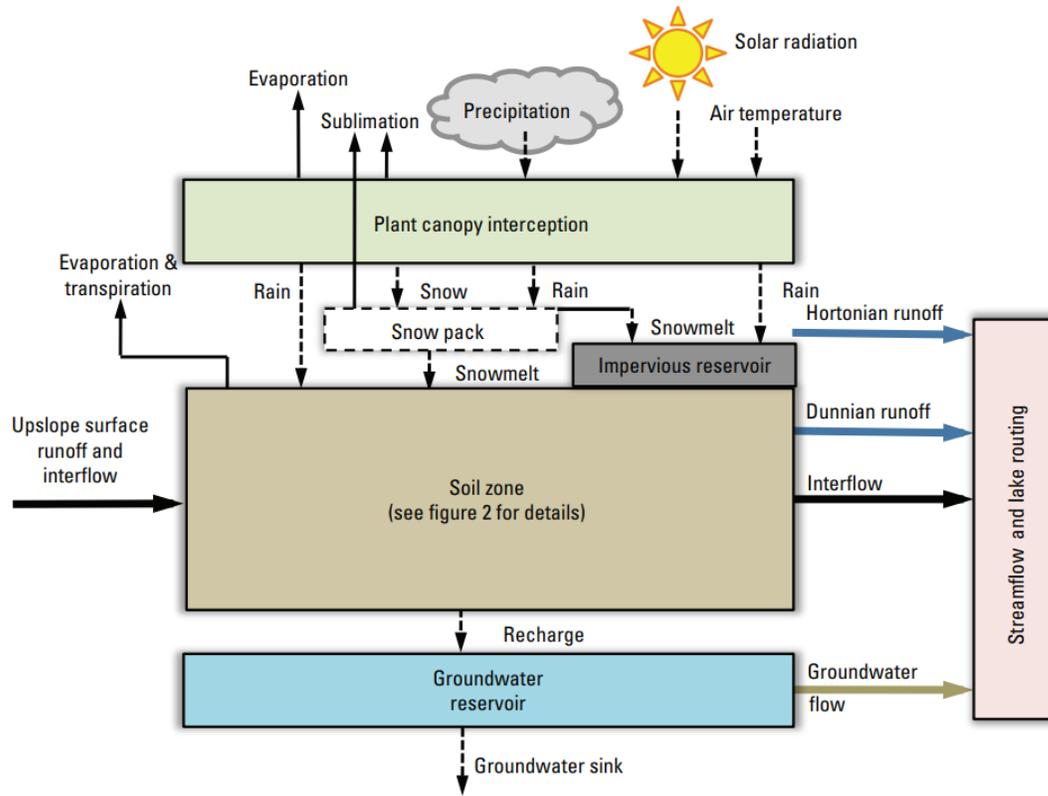
Figure 3-1 Hydrologic Processes simulated within PRMS (Markstrom, et al., 2015).
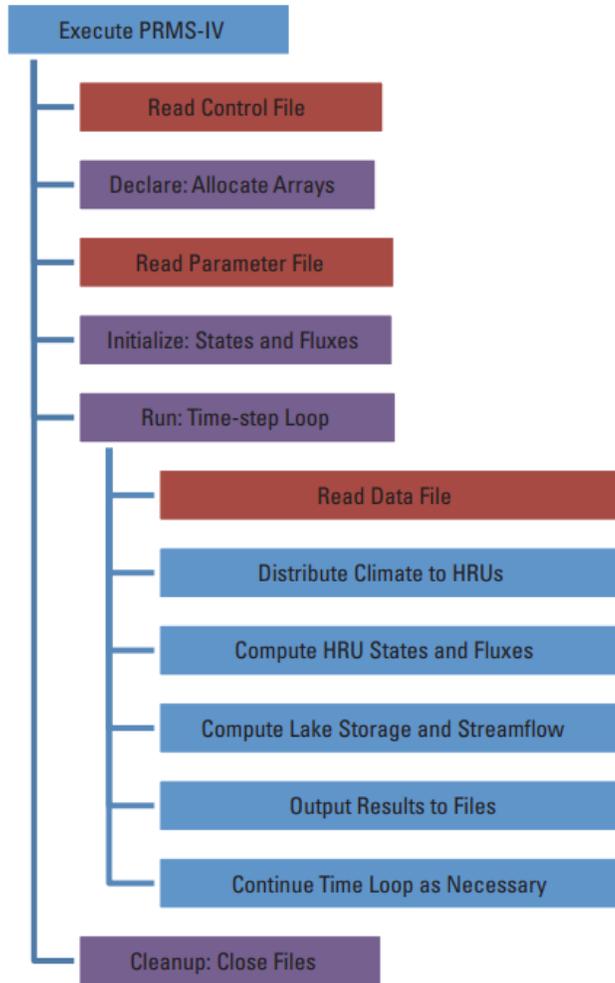
Figure 3-2. Computational sequence of PRMS (Markstrom, et al., 2015)

Table 3-1 Upcountry HRUs

Subwatershed	Area (thousand acres)	# of HRUs
Hetch Hetchy	290	280
Cherry-Eleanor	126	211
Don Pedro	564	574

3.2.2. Calibration

The PRMS was used to develop three models for the Upcountry region: one for inflows to Hetch Hetchy, one for combined inflows to the Cherry and Eleanor reservoirs, and a third for the region between Don Pedro reservoir and those three upstream reservoirs. The first two of these models (Hetch Hetchy and Cherry-Eleanor) were developed by SFPUC, while the UMass team developed a new PRMS model for the Don Pedro sub-region. To have a unified and consistent calibration within the Upcountry region, UMass referred to the Hetch Hetchy/Cherry-Eleanor models to calibrate the new Don Pedro model; the model parameters being calibrated were kept same and some of parameters of the Hetch Hetchy/Cherry-Eleanor PRMS models were adopted by the Don Pedro PRMS (e.g., snow and PET module parameters). The Upcountry subwatersheds are shown in Figure 3-3, as are the HRUs for the Don Pedro subwatershed.

As the Hetch Hetchy and Cherry-Eleanor PRMS models were developed by SFPUC,

The calibration process entails estimating 43 hydrologic parameters for each HRU in each region. Particle Swarm Optimization (PSO), a kind of evolutionary algorithm (Marini & Walczak, 2015), was selected to perform the calibration for the Don Pedro PRMS, while a Monte Carlo experiments were performed to find the best performing parameter sets of the Hetch Hetchy/Cherry-Eleanor PRMS models.

PSO is inspired by the collective behavior of social animals such as birds and fish. Particles—in this case, a possible set of hydrologic parameter values—are randomly placed in the search space of the problem and each evaluates the objective function at its current location. Each particle then determines its movement through the search space through the combination of the history of its own current and best-fitness locations with those from the members of the swarm with some randomness. Then the next iteration begins after all particles (possible parameter sets) have been moved. Eventually, all the particles are likely to move close to an optimum of the fitness function (Poli, et al., 2007). The PSO algorithm is depicted generally in Figure 3-4.

In the Upcountry region calibration, the calibration objective (fitness function) for each region is to minimize Root Mean Squared Error (RMSE). The calibration process of the Don Pedro model is to maximize Kling-Gupta Efficiency (KGE; Gupta et al., 2009), and has been finalized within an R routine developed to facilitate the PRMS calibration task. KGE can range between $-\infty$ and 1; $-\infty$ means no fitness between discharge simulation and observation, while 1 means perfect fitness between modeled and observed stream flows.
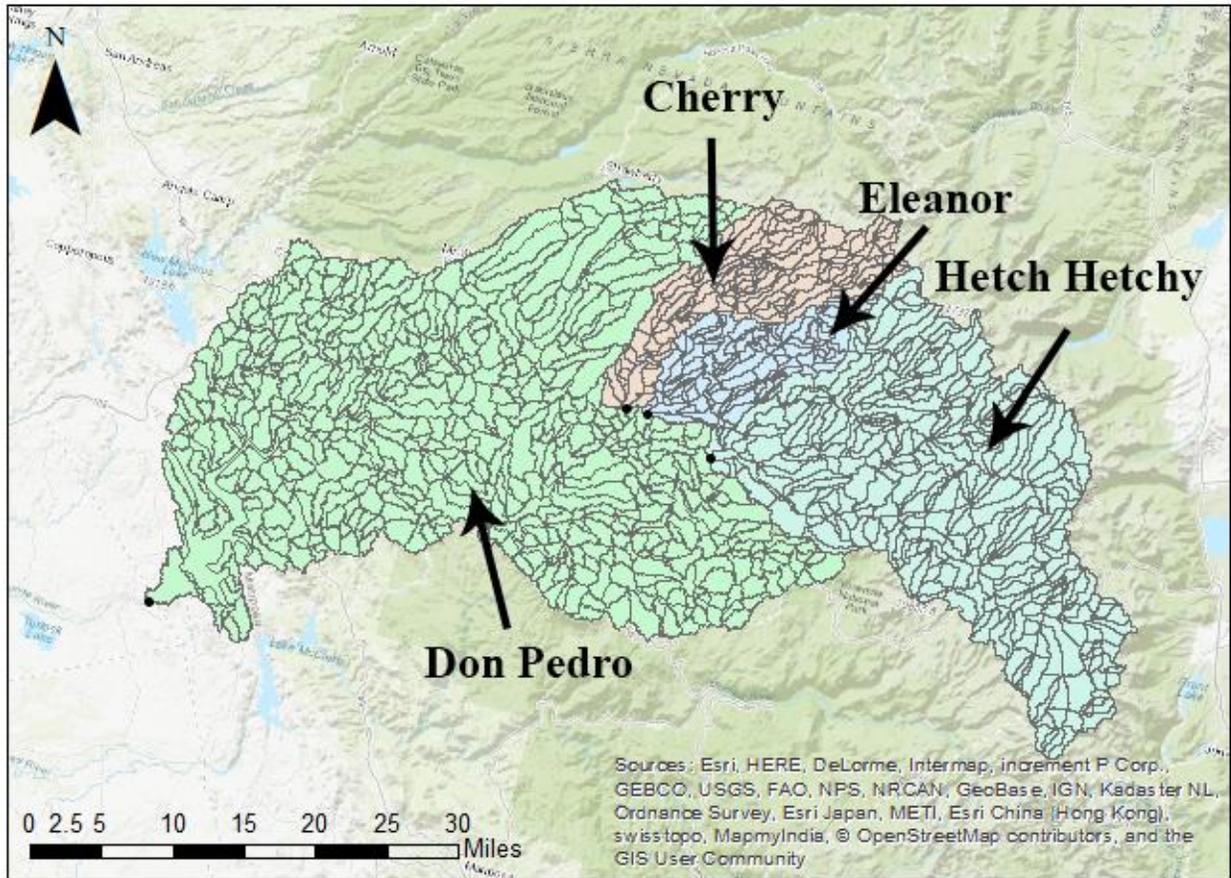
Figure 3-3. Upcountry hydrologic region subwatersheds and HRUs (Don Pedro subwatershed only). Subwatershed outlets are shown as black dots.
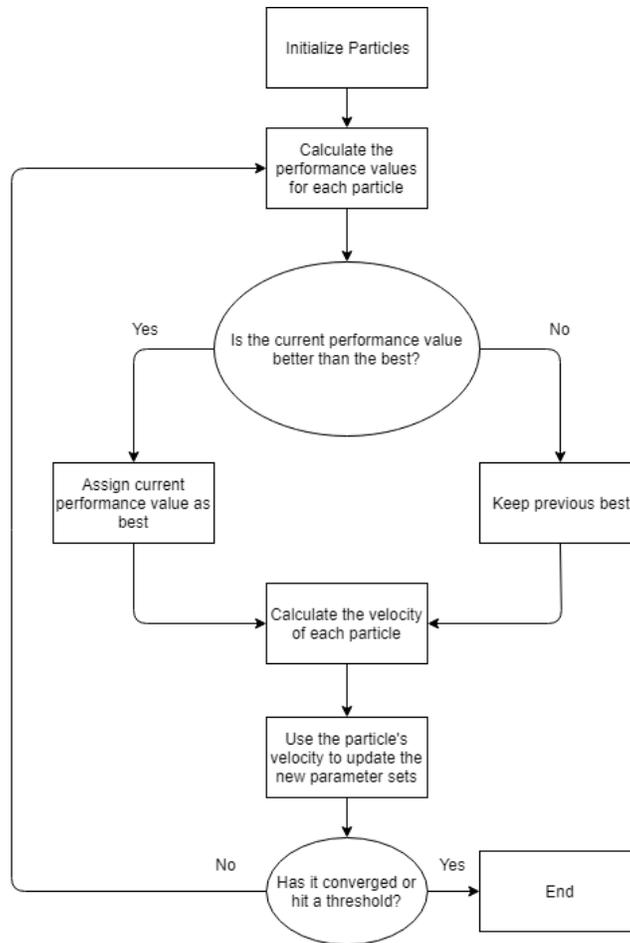
Figure 3-4. Particle Swarm Optimization Schematic Flowchart

3.2.3. Climate Data

The PRMS models are driven by daily maximum and minimum temperature and precipitation. The weather stations collected for the up-country PRMS study are shown in Figure 3-5, with the blue square representing the 9 temperature stations and the green triangle representing the 9 precipitation stations. As shown in the figure, six stations provide both precipitation and temperature.
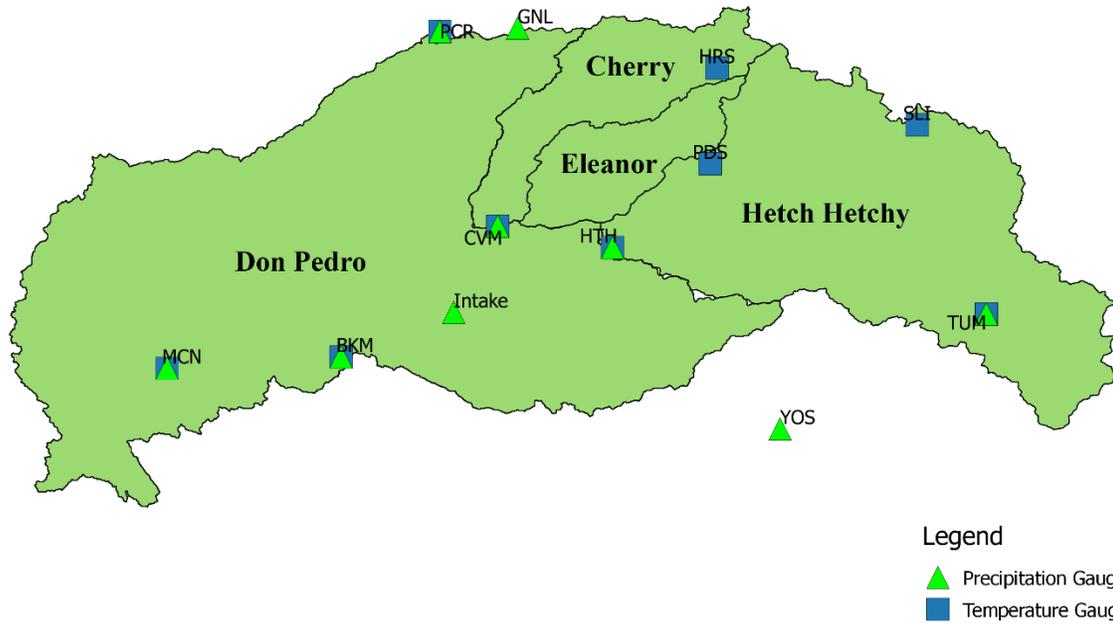
Figure 3-5 Up Country weather stations

The Upcountry PRMS models use a subset of the weather stations selected based on proximity, data quality, and length of data record. Table 3-2 provides a description of the nine precipitation stations and nine temperature stations used to drive the Up-country PRMS models.

Table 3-2 Stations used in PRMS

Station	CDEC ID	Precipitation	Temperature	Duration of Data Record
Hetch Hetchy	HEM	Yes	Yes	1/1/1930 to present
Buck Meadows	BKM	Yes	Yes	7/19/1999 to present
Tuolumne Meadows	TUM	Yes	Yes	10/15/1985 to present
Cherry Valley Dam	CVM	Yes	Yes	12/25/1952 to present
Moccasin	MCN	Yes	Yes	1/1/1930 to present
Paradise Meadow	PDS	No	Yes	10/1/1985 to present

24

Horse Meadow	HRS	No	Yes	10/1/1985 to present
Slide Canyon	SLI	No	Yes	10/1/1985 to present
Pinecrest	PCR	Yes	Yes	7/10/1996 to present
Gianelli Meadow	GNL	Yes	No	5/24/1988 to present
Yosemite	YYV	Yes	No	12/10/1998 to present
Early Intake	EIN	Yes	No	1/1/1930 to present

While not direct inputs to the PRMS models, a total of 22 snow pillow and snow survey locations were used to calibrate the PRMS snow outputs. Figure 3-6 shows the geographic locations of the snow observation points. In addition to the snow pillow and survey data, the PRMS snow module was also calibrated to the Airborne Snow Observatory's basin-wide average snow measurement.

In the following sub-sections, the use of precipitation and temperature data in the PRMS is described in more detail.
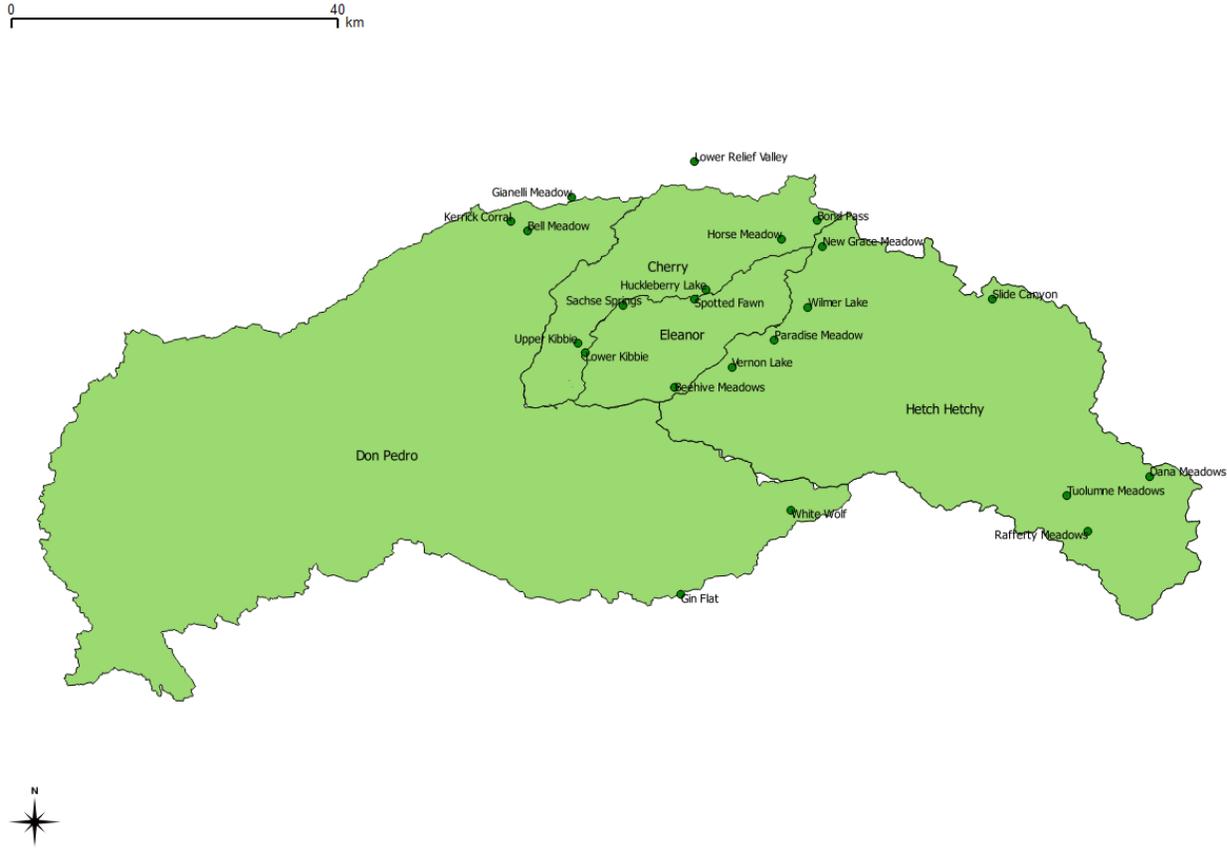
Figure 3-6 Snow pillow and survey sites for the Up Country region

3.2.3.1. Interpolation of Precipitation Station

As the climate station data is often missing from one period to another, the precipitation stations Buck Meadows (BKM), Tuolumne Meadows (TUM), Pinecrest (PCR), and Gianelli Meadow (GNL) were created by monthly (12) multivariate regressions using the precipitation stations Hetch Hetchy (HTH), Cherry Valley (CVM), Moccasin (MCN), Yosemite (YOS), and Early Intake (INTAKE). The multivariate regression used to estimate daily precipitation at the gage i from available gages k during the month m is given by Equation 3-1.

$$P_i(t) = \sum_k a_{k,i,m} P_k(t) + b_{k,i,m}, \qquad \text{3-1}$$

with $P_i(t)$ the estimated precipitation at the gage i for the time step t during the month m. $P_k(t)$ is the daily precipitation at the gage $k \in$ {HTH, CVM, MCN, YOS, INTAKE}. $a_{k,i,m}$ and $b_{k,i,m}$ are the slopes and intercepts of the multivariate regression. For the sake of model parsimony, the use of a single model that could predict precipitation for all months was investigated. However, such a model leads to a significantly bias in the seasonal pattern of precipitation with substantially lower precipitation during winter months and larger estimated precipitation during summer months.

Note that application of Equation 3-1 leads to two issues that need to be address prior using the interpolated precipitation time series as input to PRMS model. First, since the slope parameters $a_{k,i,m}$ can take negative values, the predicted daily precipitation can take negative values too. For these rare instances, the negative values are replaced by 0. The second issue is when all predictor values $P_k(t)$ are 0 for the time step t, which lead the predicted precipitation $P_i(t)$ to be equal to the intercepts $b_{k,i,m}$. To prevent creating artificially precipitation events while all predictors were dry days, the monthly volume of precipitation falling during days where predictors are all null is redistributed over rainy days within the same month. The distribution is not even as days with the large rainfall gets more extra rainfall than days with only few rainfall.

The regression parameters were calibrated over the period of 1999-10-01 through 2016-10-01 and validated over the period of 1969-10-01 through 1999-10-01. The calibrated parameters are shown in appendix A. These parameters were calibrated for the period of 1999-10-01 through 2016-10-01 and while validated for the period of 1969-10-01 through 1999-10-01. The estimated precipitations at BKM, TUM, PCR, and GNL were then checked based on all rain gauge predictors, where no precipitation is measured the rainfall assigned are set to 0. These precipitation volumes were then re-distributed to other precipitation events to preserve mass-balance.

Figure 3-7 and Figure 3-8 shows the calibration and validation scatterplot for the monthly multivariate regression. This methodology was accepted to extend out the stochastic weather generator (CliWxGen) stations for the Up Country region.
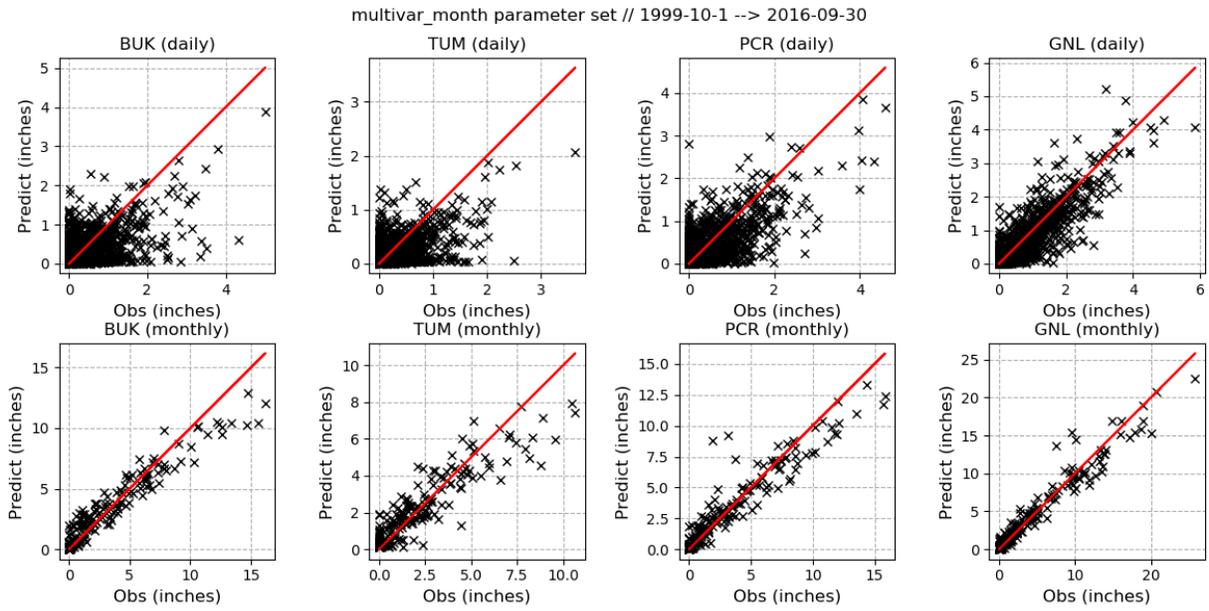
Figure 3-7. Calibration scatterplot for the monthly multivariate regression
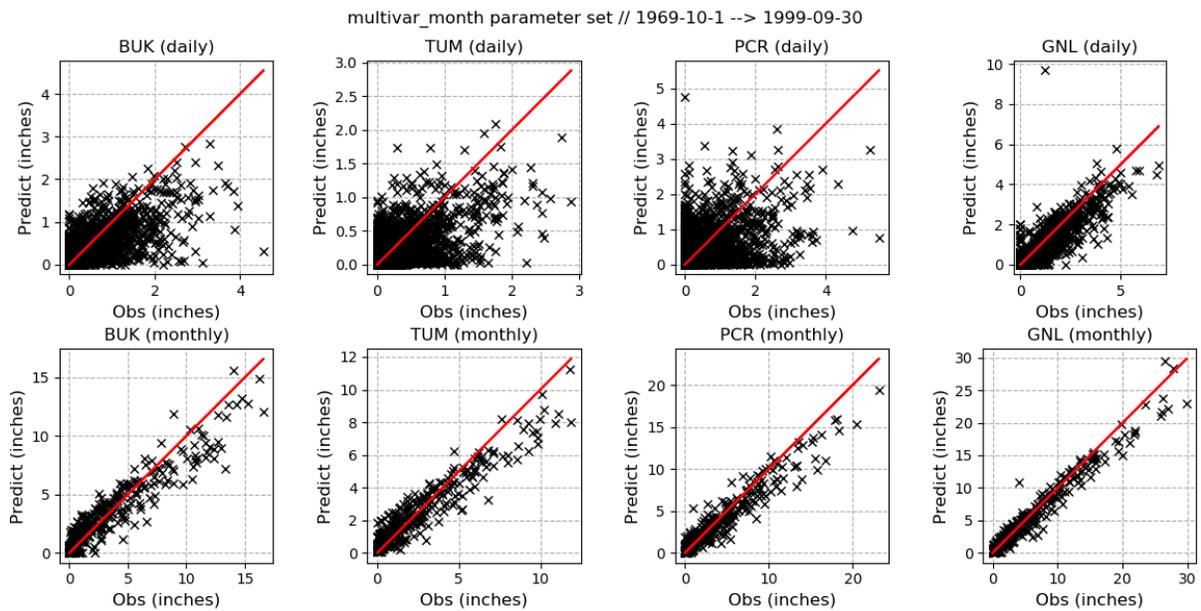
Figure 3-8. Validation scatterplot for the monthly multivariate regression

3.2.3.2. Interpolation of Temperature Station

Temperature interpolation was done by SFPUC for Hetch Hetchy and Cherry-Eleanor PRMS Hydrology models. The same input maximum and minimum temperatures were used, but the specific steps done by SFPUC are available in Appendix A.

3.2.3.3. Bias Correction of Temperature Station

The temperature dataset was bias corrected using quantile mapping with normal distributions fit for each month on a daily time-step. The bias correction was done using maximum and minimum observed temperature from 1969-10-01 through 2016-09-30 and the input climate time-series from the baseline maximum and minimum temperature from the climate stress test.

3.3. Don Pedro Hydrology Calibration Results

The Don Pedro PRMS model contains 30 parameters that no data was available, therefore they were calibrated using particle swarm optimization. The particle swarm optimization function was applied to the Don Pedro watershed with 150 particles (i.e., 150 parameter sets) improved over 50 iterations. The Don Pedro model contains 76 universal parameters and 53 distributed parameters. Out of the 129 parameters, a total of 30 parameters were calibrated through the particle swarm optimization scheme (Figure 3-4). Below, Table 3-3 lists the parameters being calibrated.

Table 3-3 PRMS parameters calibrated with particle swarm optimization

Parameter	Distributed	Description
Rad_trncf	No	Transmission coefficient for short-wave radiation through the winter vegetation canopy
Cecn_coef	No	Monthly (January to December) convection condensation energy coefficient
Imperv_stor_max	No	Maximum impervious area retention storage for each HRU
Pref_flow_den	No	Fraction of the soil zone in which preferential flow occurs for each HRU
Smidx_coef	No	Coefficient in non-linear contributing area algorithm for each HRU
Smidx_exp	No	Exponent in non-linear contributing area algorithm for each HRU
Snowinfil_max	No	Maximum snow infiltration per day for each HRU
Soil_moist_init	No	Initial value of available water in capillary reservoir for each HRU

Soil_moist_max	No	Maximum available water holding capacity of capillary reservoir from land surface to rooting depth of the major vegetation type of each HRU
Soil_rechr_init	No	Initial storage for soil recharge zone for each HRU
Soil_rechr_max	No	Maximum storage for soil recharge zone for each HRU
Transp_tmax	No	Temperature index to determine the specific date of the start of the transpiration period
Radadj_intcp	No	Intercepts in air temperature range adjustment to solar radiation equation
Radadj_slope	No	Slope in air temperature range adjustment to degree-day equation
Jh_coef	No	Monthly (January to December) air temperature coefficient
K_coef	Yes (segments)	Travel time of flood wave from one segment to the next downstream segment
X_coef	No	The amount of attenuation of the flow wave
Carea_max	No	Maximum possible area contributing to surface runoff expressed as a portion of the HRU area
Fastcoef_lin	No	Linear coefficient in equation to route preferential-flow storage down slope for each HRU
Fastcoef_sq	No	Non-linear coefficient in equation to route preferential flow storage down slope for each HRU
Sat_threshold	No	Water holding capacity of the gravity and preferential flow reservoirs; difference between field capacity and total soil saturation for each HRU
Slowcoef_lin	No	Linear coefficient in equation to route gravity reservoir storage down slope for each HRU
Slowcoef_sq	No	Non-linear coefficient in equation to route gravity-reservoir storage down slope for each HRU
Soil2gw_max	No	Maximum amount of the capillary reservoir excess that is routed directly to the GWR for each HRU
Ssr2gw_exp	No	Non-linear coefficient in equation used to route water from the gravity reservoirs to the GWR for each HRU

Ssr2gw_rate	No	Linear coefficient in equation used to route water from the gravity reservoir to the GWR for each HRU
Sstor_init	No	Initial storage of the gravity and preferential-flow reservoirs for each HRU
Gwflow_coef	No	Linear coefficient in the equation to compute groundwater discharge for each GWR
Gwsink_coef	No	Linear coefficient in the equation to compute outflow to the groundwater sink for each GWR
Gwstor_init	No	Storage in each GWR at the beginning of a simulation

Because the Don Pedro PRMS model domain is affected by upstream dam operations, reconstructed naturalized flow have been derived as calibration target using data from the Hetch Hetchy streamflow gauge, Cherry/Eleanor gauge, and the La Grange streamflow gauge:

$$Target\ Streamflow = La\ Grange - (Hetch\ Hetchy + "\text{Cherry/Eleanor}"\ Gauge)$$

The above natural target streamflow calculation results in some negative values. The negative streamflow values were not calculated when factoring in for the performance metric of the Kling-Gupta Efficiency.

The calibrated parameter sets returned a KGE of 0.70 on a monthly scale. The KGEs for other temporal scales along with other performance metrics (NSE and RMSE) are also provided in Table 3-4. **Error! Reference source not found.** shows the streamflow hydrograph for various temporal scales as well as monthly scatterplot and flow duration curve for La Grange gauge, which is a summation of Hetch Hetchy, Cherry-Eleanor, and Don Pedro hydrology models. When reviewing the daily streamflow plot, the modelled Don Pedro model is observed to underestimate the peaks of the streamflow. However, on the monthly aggregate time step, the modelled peaks line up much better with the natural target streamflow. Based on the annual hydrograph comparison, the overall calibrated model fits are fair for monthly streamflow as the model underestimating for the wet years and overestimates during the dry years. However, the annual fit is not very good with underestimation of streamflow over wetter periods and the overestimation for the drier periods. The scatterplot in **Error! Reference source not found.** has a fitted line which shows the hydrology models underestimates observed monthly streamflow. The flow duration curve shows the models underestimates during low flow periods. This has resulted in requiring an additional correction model, which was developed and described in detail on section 3.5.

Table 3-4 Performance indices of Don Pedro accretions calibration

	KGE (Target)	NSE	Percent Bias (%)
Daily (Target)	0.65	0.56	
			2.2
Monthly	0.70	0.70	

Yearly	0.56	0.78
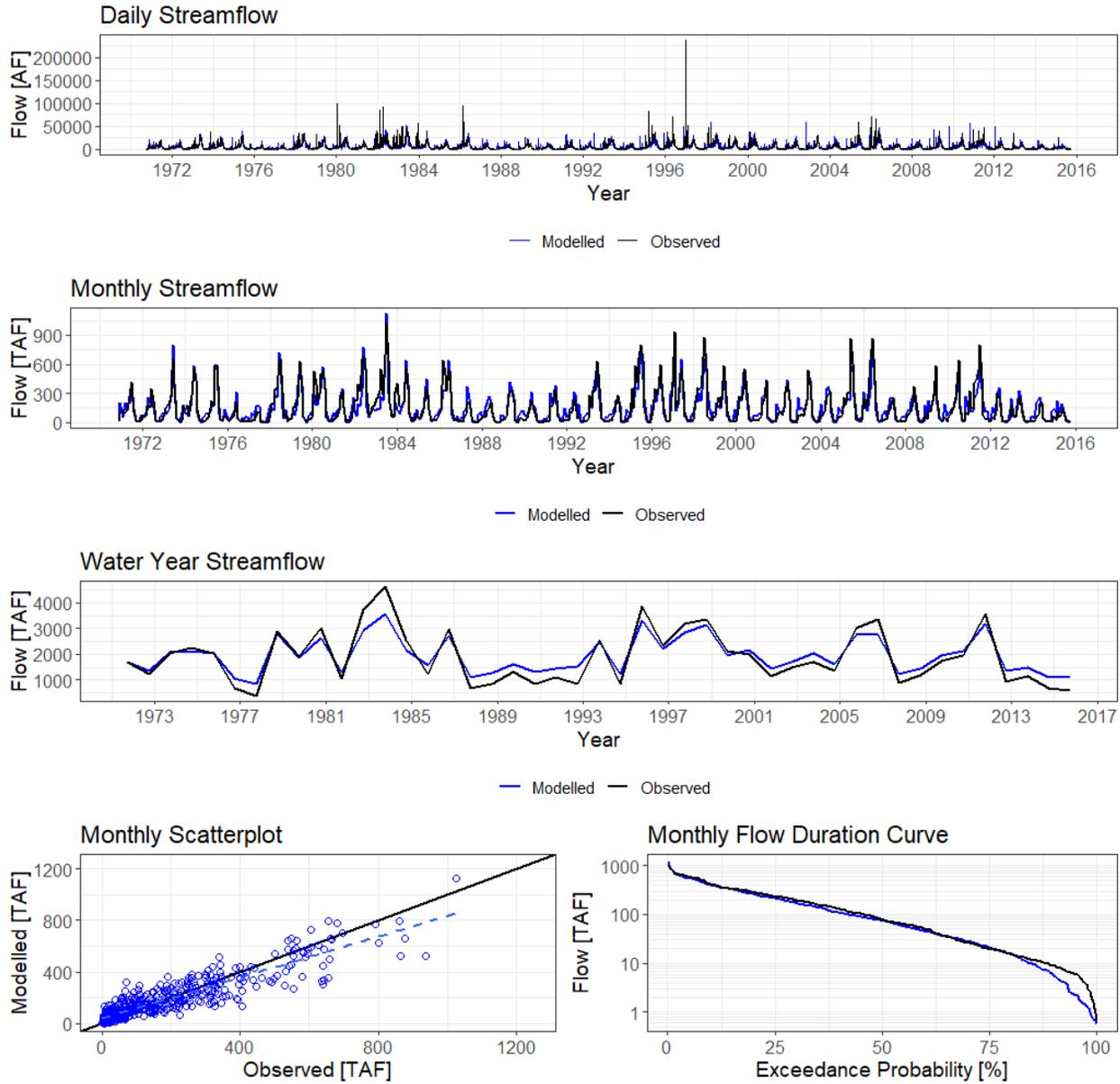
Figure 3-9. Daily, Monthly, and Water Year hydrograph, monthly scatterplot, and monthly flow duration curve for La Grange Streamflow at New Don Pedro Reservoir. Flow Duration Curve has the Y-axis in log-based scale.

3.4. Definition of Water Available to the City (WAC)

A good reproduction of the unimpaired Tuolumne flow at La Grange is key for the RWS as the Modesto and Turlock Irrigation Districts (MID and TID) entitlements and Water Available to the City (WAC) values are defined. The performance of the combined PRMS model simulations through the mass balance

equation somewhat averages out the performance obtained for the three calibrated PRMS models. Similar to its upstream components, the simulated flow at La Grange presents a bias in variance at annual time scale, leading to overestimation of dry years and underestimation of wet years.

TID and MID have senior water rights on the water of the Tuolumne River. WAC represents the water from the unimpaired flow in the Tuolumne River at La Grange that belongs to SFPUC. The Raker Act specifies how much and when Tuolumne river flows must be available to MID and TID (district entitlements). On any given day, the actual district entitlements consist of the lesser of i) 4,792 AF/day (2,416 cfs) from June 14th through April 14th or 8,064 AF/day (4,066 cfs) from April 15th through June 13th (aka: maximum district entitlement), and ii) unimpaired flow in the Tuolumne River at La Grange. The implication is that water available to SFPUC from the Tuolumne River (i.e., WAC) is any unimpaired flow at La Grange (below Don Pedro) that is greater than actual district entitlements. Figure 3-10 illustrates the WAC for an extremely wet water year (1983) and an extremely dry water year (1987), which highlights the large variability of both the unimpaired flow of the Tuolumne River at La Grange and the WAC. Regarding the calculated WAC from the PRMS simulations, it is noted that the underestimation of the inter-annual variability of the unimpaired flow at La Grange leads to similar bias for WAC. An overestimation of the WAC during dry years is a concern as it prevents proper representation of droughts.
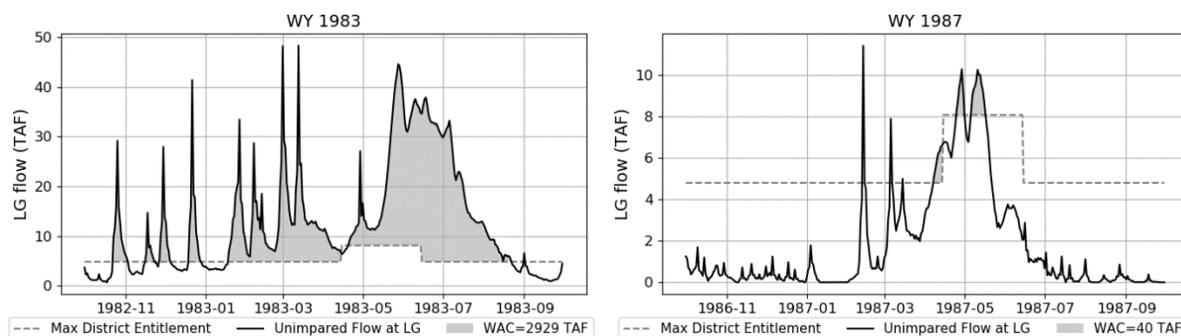
Figure 3-10. Water Available to the City (WAC) for an extreme wet year (left, 1983) and a dry year (right, 1987). The year 1987 is the first year of the historical drought 1987-1992. The shaded area shows the WAC and the dash-grey line shows the maximum district entitlement. Note that the y-axis differs for the two subplots.

3.5. Current Precipitation Index (CPI)

Given the importance of a good reproduction of the WAC, UMASS collaborated with SFPUC to correct the PRMS simulations across the Upcountry region. A post-processing model was developed by SFPUC to correct the PRMS daily streamflow. The considered method attends to correct the residual model errors using meteorological indices. For precipitation variable, an index called "Current Precipitation Index" is used to account for the basin wetness. For temperature variable, a heat index similar to a degree-day is used to account from the basin warmness. Note that the considered post-processed correction is a parametric method (i.e., it requires some parameters to be calibrated). The calibration of these parameters was done with the main objective to better reproduce the dry years. The complete writeup of the post-processing model, including calibration and validation, are provided in Appendix J along with the calibrated parameter sets.

The comparison of the raw PRMS simulations (i.e., prior correction) with the post-processed simulations (i.e., after correction) is shown in Table 3-5 and in Figure 3-11 and Figure 3-12. The results clearly indicate a significant improvement of the inter-annual variability of the simulated streamflow, with significant improvement of the KGE and NSE scores at annual time scale, especially for Cherry Reservoir/Lake Eleanor watershed and Don Pedro accretion watershed. It also results in a much better reproduction of the inter-annual variability of the annual Tuolumne flow at La Grange and of the WAC. The flow duration curve (Figure 3-11and the scatterplots (Figure 3-12) figures illustrate the significant improvement of the simulations during the low flow years/months. The downside of the post-processed model is that it worsens the bias across the region. At La Grange, the overall bias shifted from +3.2% prior correction to -6.8% after correction. Regarding WAC, the reduction of the bias in inter-annual variability (and a significant improvement of the low WAC years) comes with a significant improvement of the high WAC years (Figure 3-11 and Figure 3-12). Table 3-5 also shows the improvement of the KGE and NSE scores at all temporal scales, with a major improvement at annual scale, which, however, leads to a slight increase in the negative bias (from -8.2% to -8.8%).

Table 3-5. PRMS streamflow calibration metrics for the Upcountry region before and after correction. Performance criteria are estimated over the period 1972-2015.

	Raw PRMS simulations (prior correction)			Post-processed PRMS simulations (after correction)		
	KGE	NSE	Percent Bias (%)	KGE	NSE	Percent Bias (%)
Hetch Hetchy Reservoir inflow						
Daily	0.85	0.70		0.87	0.74	
Monthly	0.92	0.86	3.9	0.92	0.89	-3.8
Water Year	0.76	0.91		0.83	0.91	
Cherry Reservoir/Lake Eleanor inflow						
Daily	0.78	0.56		0.85	0.73	
Monthly	0.91	0.83	3.0	0.90	0.88	7.3
Water Year	0.67	0.86		0.91	0.92	
Accretion flow at Don Pedro Reservoir						
Daily	0.65	0.56		0.67	0.63	
Monthly	0.70	0.70	2.2	0.76	0.77	-20.8
Water Year	0.56	0.78		0.77	0.86	
Naturalized Tuolumne flow at La Grange						
Daily	0.79	0.68		0.87	0.79	
Monthly	0.84	0.83	3.2	0.91	0.90	-6.8
Water Year	0.66	0.86		0.87	0.94	
Water Available to the City (WAC)						
Daily	0.71	0.56		0.81	0.70	
Monthly	0.78	0.75	-8.2	0.88	0.85	-8.8
Water Year	0.61	0.82		0.84	0.93	
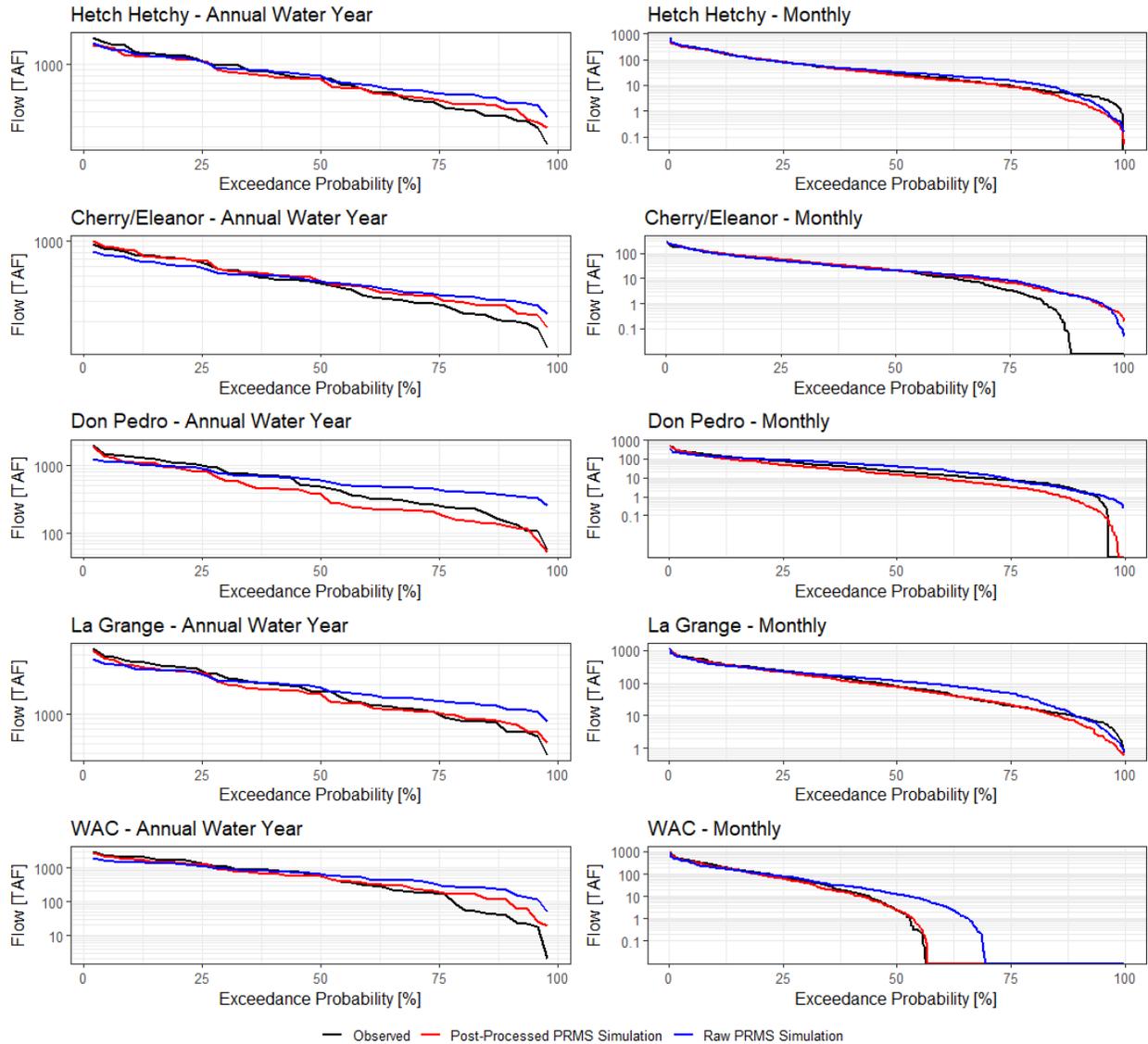
Figure 3-11. Flow duration curve for Tuolumne Flow at La Grange, Hetch Hetchy Reservoir inflow, Cherry Reservoir and Lake Eleanor inflows, accretion flow to Don Pedro Reservoir, and WAC for annual water year and monthly temporal scale. The black, blue and red lines show the flow duration curve for the historic observed flow, the raw PRMS simulation (i.e., prior correction), and the post-processed PRMS simulation (i.e., after correction). A logarithm scale is used for the streamflow to ease visualization of the low flow years/months.
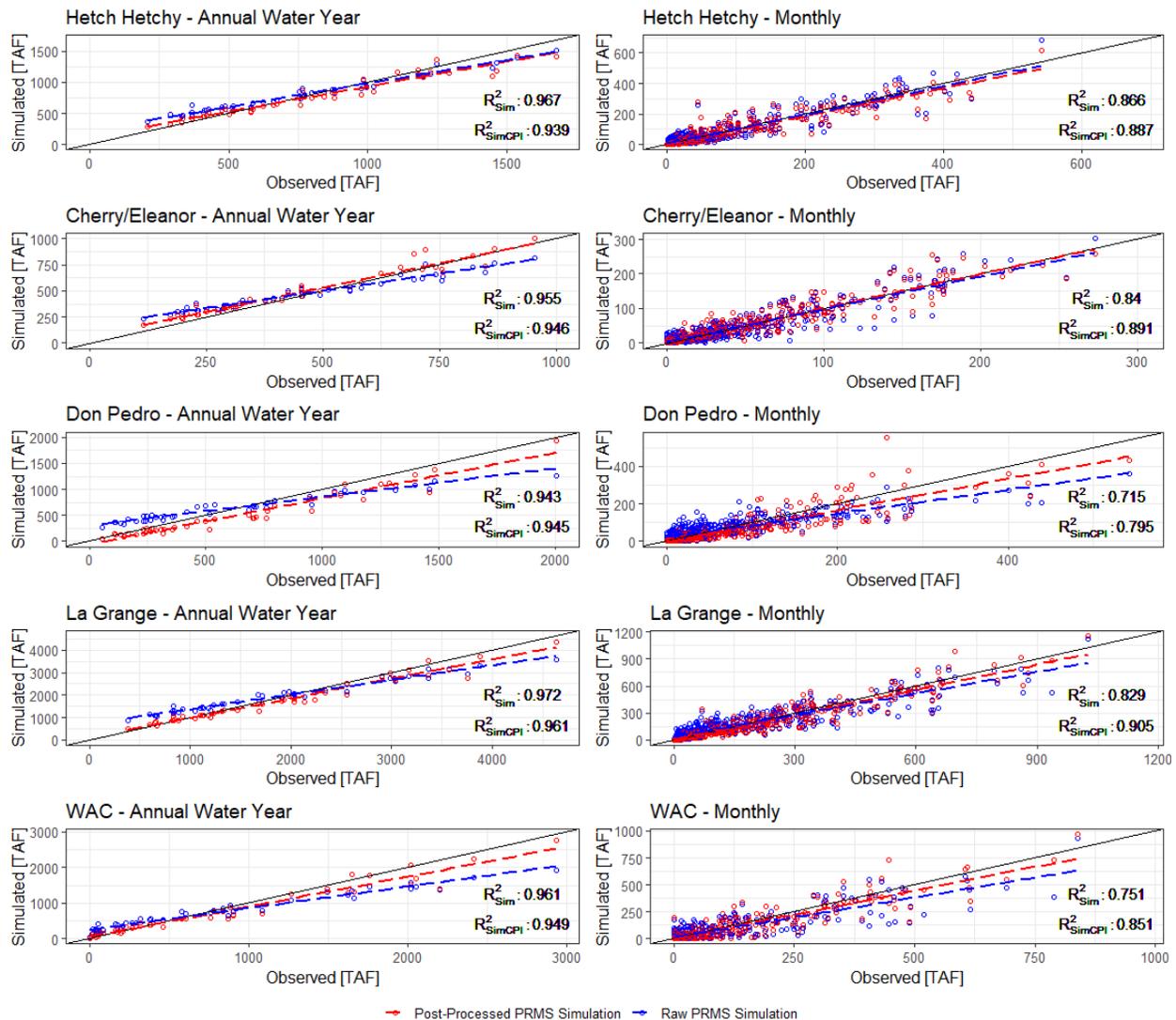
Figure 3-12. Scatterplot for Tuolumne Flow at La Grange, Hetch Hetchy Reservoir inflow, Cherry Reservoir and Lake Eleanor inflows, and accretion flow to Don Pedro Reservoir for annual water year and monthly temporal scales. The black line represents the identity line whereas blue and red lines represent the linear regression between the observed flow and the raw and post-processed PRMS simulations, respectively.

While the post-processing model improved the results, additional work needs to be performed to reduce the bias. The hydrologic model used to simulate the streamflow on the Tuolumne watershed in response to precipitation and temperature overestimates streamflow during dry years. For example, Table 3-6 shows the flow computed at the Tuolumne River at La Grange is overestimated, and therefore WAC is also overestimated by about 482,000 acre-feet during the drought sequence 1987-1992 (observed is 813,000 acre-feet versus simulated is 1,295,000 acre-feet).

Table 3-6. Comparison of observed and simulated WAC over 1987-1992 drought using PRMS hydrology models

	Water available to the City (acre-feet)		
Water Year	Observed	Simulated	Error

35

1987	40,439	174,868	134,428
1988	23,521	112,425	88,905
1989	379,305	383,159	3,854
1990	46,632	114,957	68,325
1991	269,733	302,719	32,986
1992	53,525	207,221	153,697
1987-92 totals	813,155	1,295,349	482,194

3.6. Climate Stress Test

3.6.1. Hydrology

This section presents the results of climate stress test for all Upcountry subwatershed systems using the weather generator output that reflect a number of scenarios for changes in precipitation and temperature.

As the PRMS models require precipitation and temperature datasets as inputs, the weather generator for the Up Country region generated 1,360 climate scenarios using statistical modeling with 8 different temperature changes (i.e., temperature increases relative to the baseline ranging from 0°C to 7°C, with an increment level of 1°C), 17 different precipitation changes (from -40% to +40%, with 5% increments), and 10 realizations. Therefore, single set of climate stress test reflects 136 different climate change scenarios.

The climate stress test were applied to all three watersheds of the Up Country region (Hetch Hetchy, Cherry & Eleanor, and Don Pedro) over the period of 2021 through 2070 and the response surface was generated based on annual mean streamflow. Figure 3-13 shows the response surface for the Hetch Hetchy, Cherry & Eleanor, and Don Pedro watersheds with the dotted line representing the historical baseline of annual flows. In addition, the dots represent different GCM projections of precipitation and temperature changes under RCP 8.5. The Upcountry annual flow response surfaces show that annual flows for Upcountry regions are not significantly sensitive to changes in temperature as demonstrated by the vertical contour lines in the response surfaces. On the other hand, precipitation changes greatly influence the changes in annual flows of the Upcountry watershed systems.
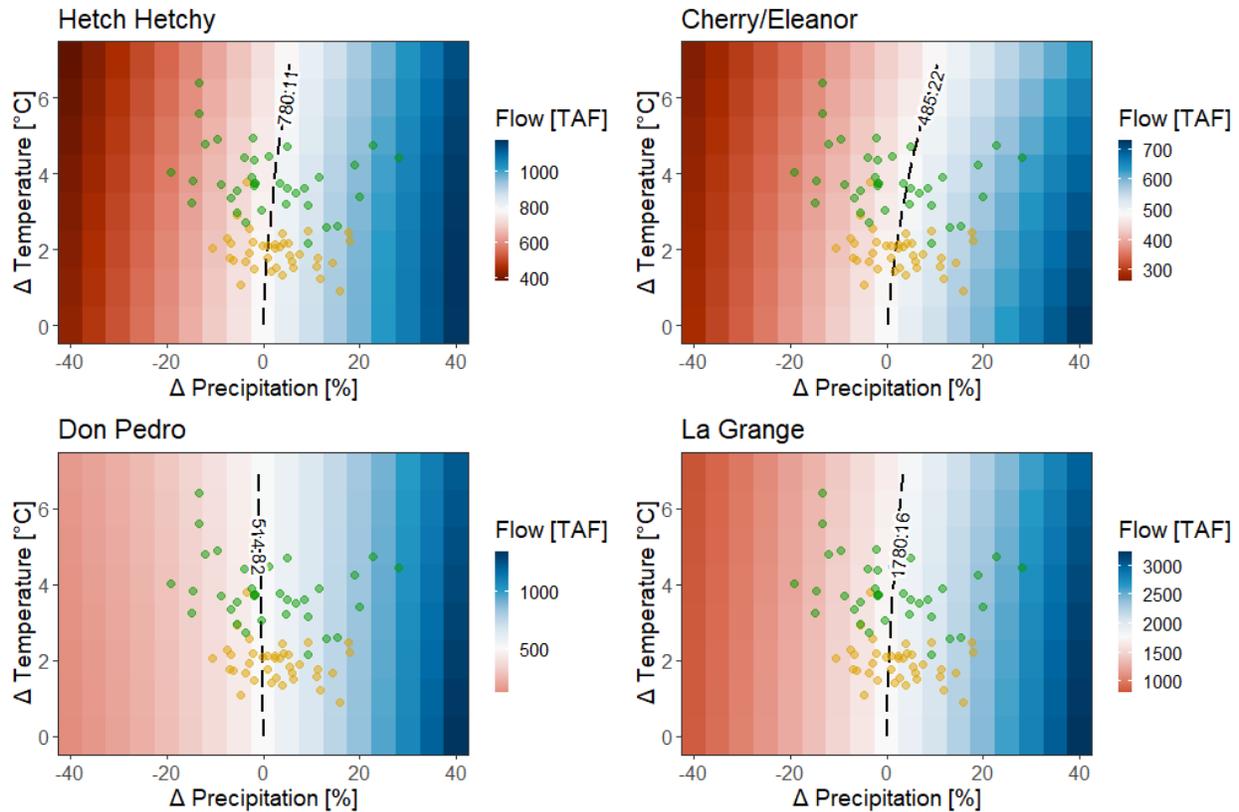
Figure 3-13. Climate response surface of annual flow at Hetch Hetchy reservoir, Cherry Reservoir/Lake Eleanor, Don Pedro accretion, and Tuolumne River at La Grange. The white color shows no change in streamflow while blue and red shows increase or decrease in annual flow, respectively. The yellow and green dots over the response surface shows CMIP5 projections under RCP 8.5 for two 30-year long periods: 2040 (2026-2055) and 2070 (2056-2095). Baseline is 1986-2005.

3.6.2. Water Available to the City (WAC)

The modifications of the Tuolumne River flow at La Grange under changing precipitation and warming temperature affects the annual values of WAC and its distribution through the years. Figure 3-14 illustrates the effects of warming temperature (left) and changing precipitation (right) on the Tuolumne River flow at La Grange and WAC. Following the response of the Tuolumne River flow at La Grange, precipitation change is the main driver of change of annual WAC. For example, a decrease in precipitation by 20% leads to a reduction of the annual average WAC from roughly 750 TAF to 400 TAF (i.e., a decrease by roughly 45%), while an increase by 20% precipitation leads to an increase of WAC from 750 TAF to 1260 TAF (i.e., +68%). At the annual scale, warming temperature slightly reduces the WAC. For instance, a +5°C warming decreases the annual average WAC from roughly 750 TAF to 729 TAF (i.e., a decrease by -2.8%). Figure 3-15 shows the modification of the annual average WAC to change in both temperature and precipitation. It highlights no significant compound effect on the annual WAC.

a) Effect of temperature change   b) Effect of precipitation change
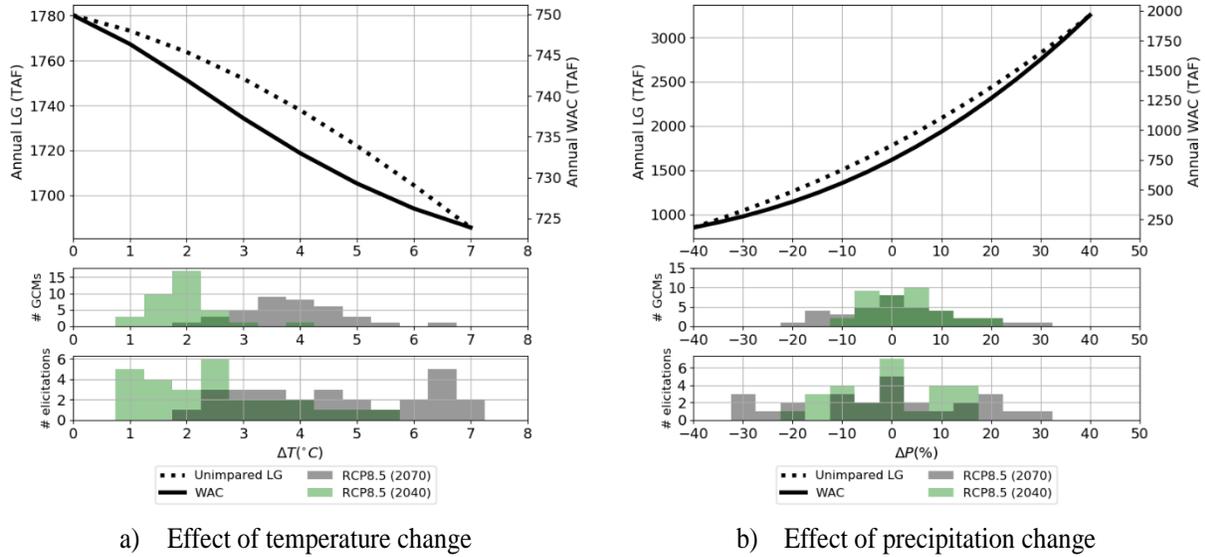
Figure 3-14 Effect of temperature and precipitation change on the Unimpaired flow at La Grange and Water Available for the City (WAC). Top panel shows unimpaired flow at La Grange (solid line) and WAC (dash line) with changes in temperature (ΔT) (a) and precipitation (ΔP) (b). Middle and bottom panels show the distribution of changes in temperature from CMIP5 projections (RCP8.5) and expert elicitations and for two 30-yr long periods centered in 2040 and 2070.
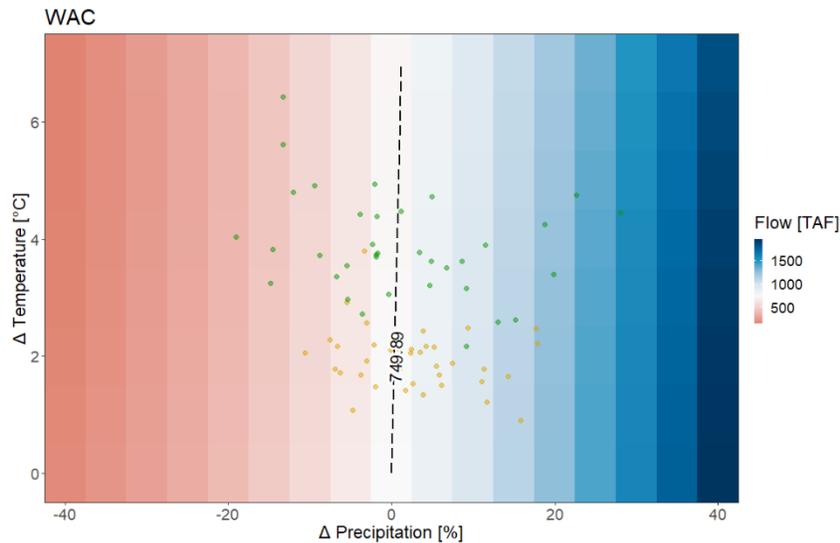
Figure 3-15. Climate response surface of the annual WAC in respect with change in precipitation (x-axis) and temperature (y-axis). See Error! Reference source not found. caption for details.

It is important to note that behind the non-significant change in annual WAC that would follow from warming temperature (Figure 3-14, left), the WAC is actually expected to decrease for roughly 60% of the years, and increase for 40% of the years (Figure 3-16). As illustrated Figure 3-16, the distribution of change in annual WAC is left skewed, which means that absolute values of reduction in WAC can get much larger values than for the increases in WAC.

38
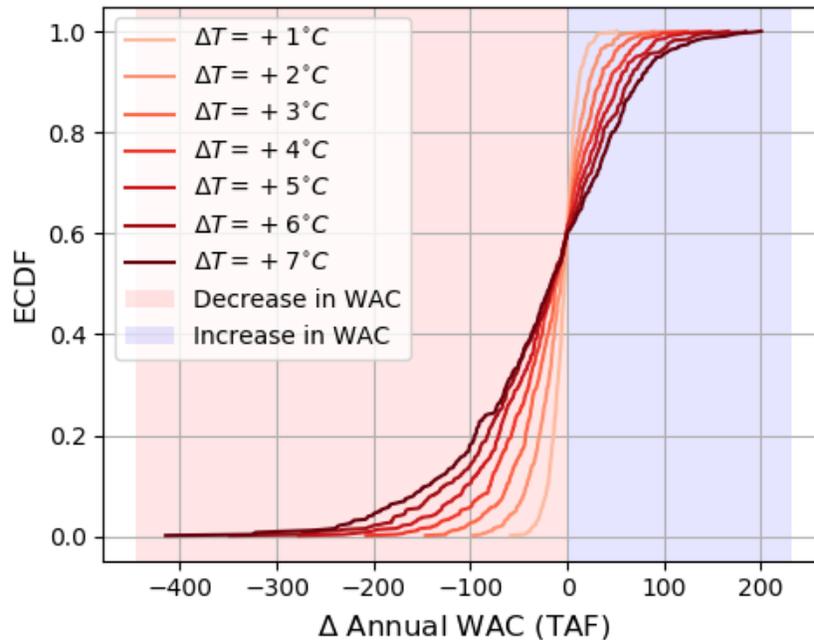
Figure 3-16. Empirical Cumulative Distribution Function (ECDF) of change in annual WAC resulting from warming temperature. Positive values (blue panel) shows years for which WAC increases and negative values (red panel) shows years for which WAC decreases. Results are obtained across the 9 stochastic realizations plus the historical realization.

Figure 3-17 illustrates the modification of the distribution of annual WAC under changing precipitation and temperature. Note that changes in all quantiles of the WAC distribution are not same. For instance, when compared the 20% precipitation reduction scenario (light red, Figure 3-17 left) with the baseline precipitation scenario (black, Figure 3-17 left), it is noted that the median value (i.e., 50% exceedance probability) decreases from 600 TAF to 300 TAF, which corresponds to a decrease by 50%. However, the quantile 10 of the distribution (i.e., 10% exceedance probability) decreases from 100 TAF to 40 TAF, which corresponds to a decrease by 60%. Figure 3-17 (left) also shows that under baseline climate conditions, roughly 25% of the years have a WAC values lower than 269 TAF while a reduction in precipitation by 20% increases this number to almost 45%, which also explains the significant increase in frequency of droughts.

The distribution of annual WAC obtained from the 510 realizations under warming conditions almost lines up with the one obtained under baseline conditions (Figure 3-17, right). Despite the lack of significant change in annual WAC, the lower percentiles of the distribution (i.e., lower than 15th) slightly increases (Figure 3-17, right), which likely explains the reduction is drought frequency (i.e., increase in return period).
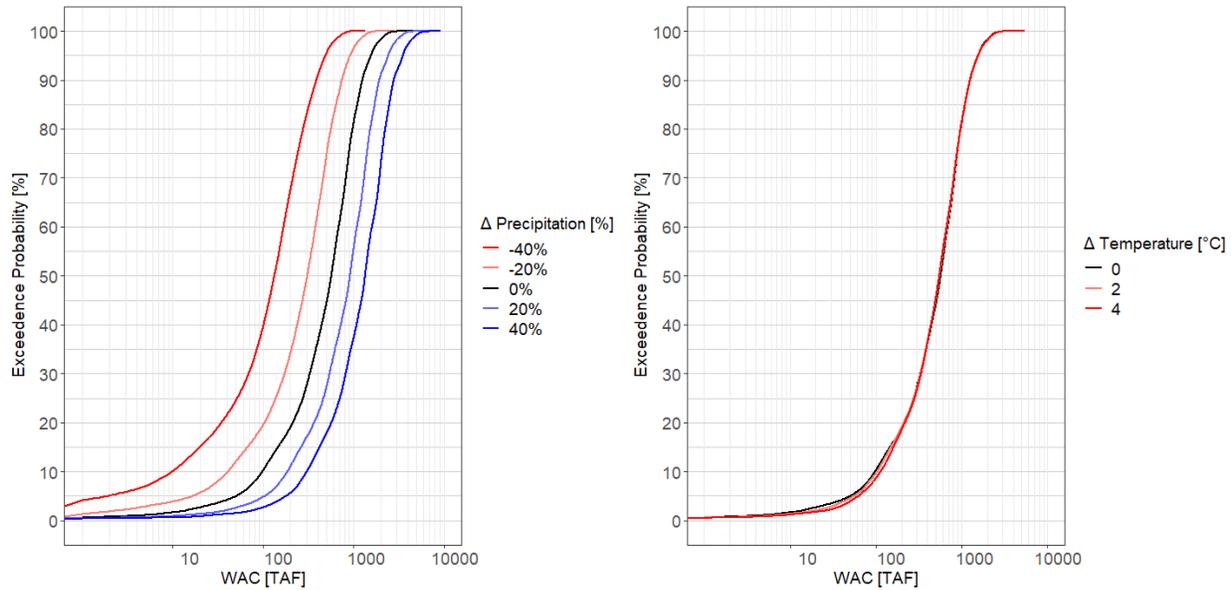
Figure 3-17. Effect of precipitation (left) and temperature (right) changes on the distribution of annual WAC. Annual WAC values are based on water year. Values are shown across the 'stochastic realizations' dataset composed by 509 stochastic realization plus the historical realization.

Figure 3-18 shows the simulated hydrograph for six selected water years obtained from the realization 10 under baseline climate and +7°C warming. An extreme warming was considered in this figure for illustration only. It highlights that both increase and decrease in annual WAC can happened during normal, dry and wet WAC years. It also shows that change in WAC roughly follows from three mechanisms:

- The spring runoff volume tends to be smaller with warmer temperatures due to a reduced snowpack. The spring runoff volume tends also to spread over longer period. Both effect tends to reduce WAC during spring season.
- Spring runoff arrive earlier in the year because above freezing temperatures are seen earlier in the year. Given the higher maximum district entitlement from April 15th through June 13th, either a reduction or an increase in WAC is possible during spring / summer seasons.
- The winter floods tend to be more frequent and with larger magnitude because more precipitation fall as rain in winter, leading to an increase in WAC during this period.

All these mechanisms combine in a non-linear fashion and lead to either an increase or decrease in annual WAC. The change in WAC distribution within years illustrated with by example years in Figure 3-18 is summarized in Figure 3-19 and Figure 3-20. Both figures show scatterplots for each calendar month between monthly WAC values obtained for the current climate (i.e., no change in precipitation and temperature) and +3°C or +7°C warming, respectively. These figures show that winter months will likely get larger WAC values as temperature increase while summer months will see a significant reduction in WAC. As an example, under the extreme +7°C warming, August months are likely to have null WAC, while WAC during July months will be close to null value too.

40

"Normal" WAC Years
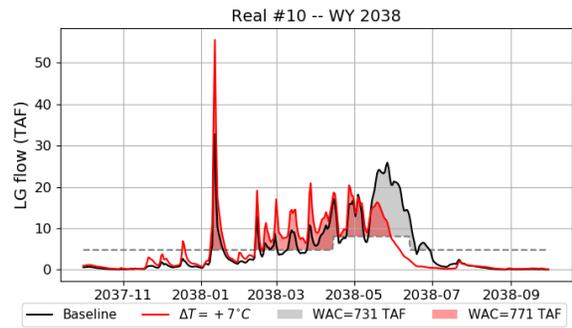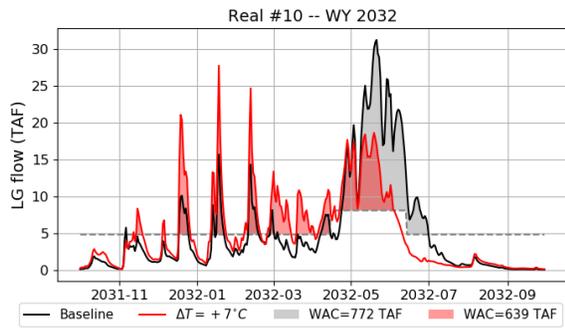
"Dry" WAC Years
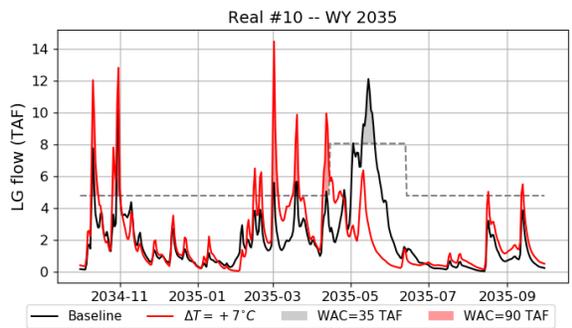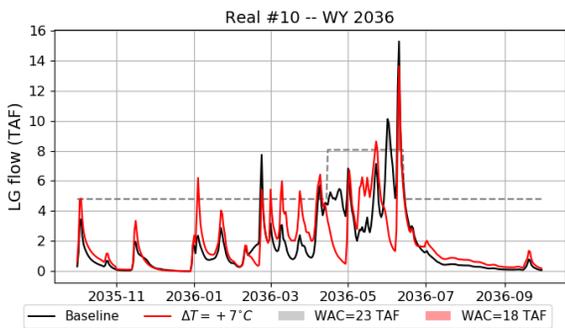
"Wet" WAC Years
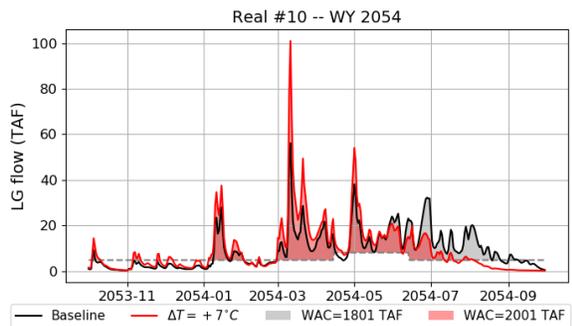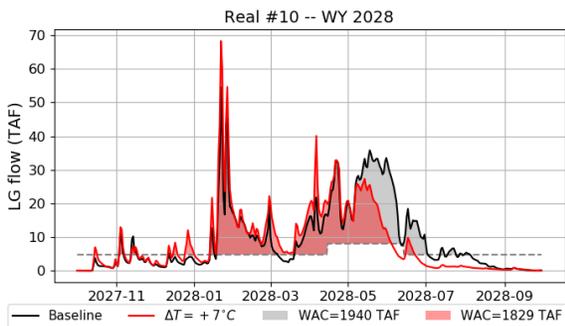
Figure 3-18. Illustration of the effect of temperature change in the temporality and cumuli of WAC. The example presented in this figure are all from the realization R10. Black color show the Tuolumne River flow at La Grange under historical climate while the red color shows the results obtained for an extreme warming of +7°C. Black and red shaded areas shows the WAC. Left and right columns show years for which WAC either decreases or decreases with a +7°C warming, respectively. The first, second and third row illustrate years for which the annual WAC is close to the average across the 10 realization (≈750 TAF), low and large, respectively.
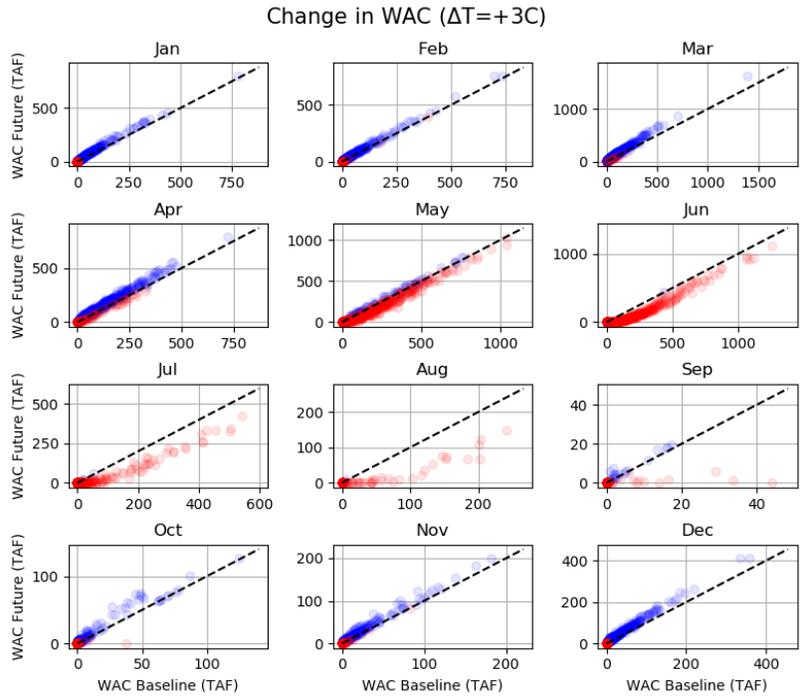
Figure 3-19. Change in monthly distribution of WAC resulting from a warming by +3°C. The panels show for each calendar month the scatter plot between monthly WAC obtained under baseline climate (x-axis) and +3°C warming (y-axis). The black dash-line is the 'no change' line. Blue and red colors are used to highlight months for which the WAC either increases or decreases, respectively. Results shown are obtained across the 10 realizations, meaning that each scatterplot shows 500 data points (50 years x 10 realizations)
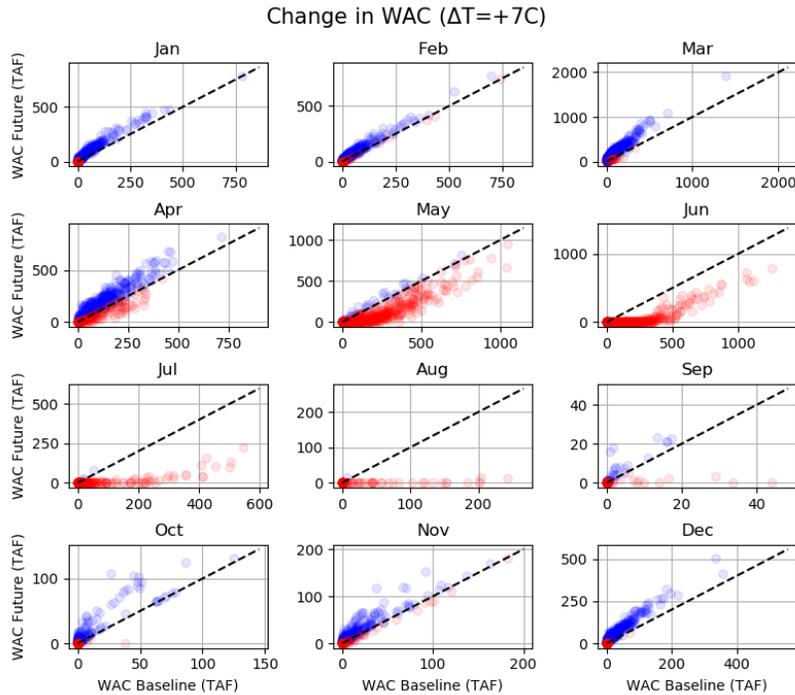
Figure 3-20. Same as Figure 3-19 but for a temperature scenario of +7°C

In summary, precipitation change has an effect on the mean annual WAC volume.

- By 2040, the median projections of +2°C warming combined with 0% change in mean annual precipitation results in no significant change in mean annual WAC volume but WAC would decrease in 60% of the years and in 20% of them, the decrease would be more than 30 TAF.
- By 2040, most projections and elicitations of precipitation change fall between -5% and +5% which would correspond to a change in mean annual WAC between a decrease of 110 TAF and an increase of 120 TAF. Also, most projections and elicitations of warming are between +1°C and +4°C. At +4°C, WAC would decrease by more than 50 TAF in 20% of the years.
- WAC would decrease from May through August and would increase in the other months.
- By 2070 RCP8.5, the median projections of about +4°C combined with 0% change in mean annual precipitation results in a 2.5% decrease in mean annual WAC volume. Most projections and elicitations of precipitation change are between -15% and +15% resulting in change in mean annual WAC volume between a decrease of 40% and an increase of 45%. . Most projections and elicitations of warming range between +3°C and +6°C. At +6°C, WAC would decrease by more than 80 TAF in 20% of the years.

Note that the PRMS hydrologic model used to simulate the streamflow on the Tuolumne watershed in response to precipitation and temperature overestimates streamflow during dry years. For example, the flow computed at the Tuolumne River at La Grange (foot note: the location where flow is allocated between the Districts and San Francisco) is overestimated, and therefore San Francisco's allocation is also overestimated by about 482,000 acre-feet during the drought sequence 1987-1992 (observed is 813,000 acre-feet versus simulated is 1,295,000 acre-feet). This volume of water is significant and is about equal to the volume of rationing that was required by customers during the six-year period, and so overstating the RWS water supply reliability.

3.7. NCAR vs CliWxGen

The section provides a summary of the task of comparing synthetic sequences of daily climate generated by NCAR-WG and CliWxGen. Although both NCAR and CliWxGen generators produced new sequences of daily precipitation and daily maximum and minimum temperature across the Upcountry region, the comparison described in this section focuses on the precipitation variable only. The comparison is carried out using simulated precipitation at nine stations spread over the Upcountry region. Precipitation time series at these nine locations are required inputs to PRMS, the hydrology model used for the LTVA to simulate streamflow across the region. The nine considered precipitation stations are Hetch Hetchy Reservoir, Buck Meadows, Tuolumn Meadows, Cherry Valley Dam, Moccasin, Pinecrest, Yosemite, Gianlelli and Early Intake). Prior to comparing the NCAR and CliWxGen weather generator outputs, it is worth highlighting few differences between these two tools.

As described in more details in the previous sections, CliWxGen is a stochastic weather generator that combines a wavelet autoregressive model with the method of fragment and the k-nearest neighbors (Knn) approach. As such, CliWxGen is only driven by the low frequency components that are identified from the historical observed rainfall. In this study, only one significant low frequency component was identified (≈15 years frequency). CliWxGen simulated daily precipitation time series for only five out of the nine stations that are used in PRMS model. Precipitation time series at the four missing locations are obtained via interpolation from the five available stations; see section 3.2.3.1 for more details regarding the interpolation). Nine 50-year long stochastic realizations of daily precipitation are used for the comparison. These nine stochastic realizations built by CliWxGen are deemed consistent with the historical period used, which is the period 1956-2011.

The NCAR weather generator is a non-parametric stochastic weather generator that combines a Markov Chain Model (MCM) with the Knn approach. MCM is used to create daily stochastic sequences of dry, wet and very wet days over arbitrary long sequences (in this case, 30-year long periods). Knn is used to randomly select a date from the historic record that satisfies the sequence state (dry-to-dry, dry-to-wet, dry-to-very wet, etc) that has been simulated with MCM. To account for seasonal effects, candidate days are sampled within a moving window that is centered on the current Julian day. For each selected dates, precipitation, maximum and minimum temperature are sampled at once to keep consistency among the weather variables. For future periods, i) a similar temperature trend as observed in the GCM projections is added to the synthetically generated temperature time series; ii) the magnitude of extremes is corrected to account for the change in extreme precipitation between historic and future periods, as seen in the GCM projections. Four different GCMs are used to condition the NCAR weather generator (CCSM, CESM, GFDL, MPI). Up to 30 stochastic realizations were generated for each GCM-forcing. The dataset used to build NCAR weather generator is the dataset that SPUC uses as input data to their PRMS models for Upcountry. This dataset spans from 1969 to 2015. More details regarding the NCAR weather generator are given in the Technical Report: Climate Change Storylines (NCAR, 2018).

Figure 3-21 shows the results of the comparison of the simulated annual precipitation obtained from the NCAR weather generator (boxplot) and CliWxGen (colored symbols). Although the comparison is meant to be carried out under current climate conditions, simulated future periods are shown for NCAR-WG. As compared to the historical mean of the Upcountry annual precipitation (i.e., roughly 950 mm; 37.4 in), all nine CliWxGen realizations well reproduced the historical average. The historic realization (brown square symbol) is from five available gauges for which available rainfall data covers the period used by

CliWxGen (i.e., 1956-2011) plus precipitation at the missing four gauges were obtained via interpolation (see Section 3.2.3.1). The black square shows the annual average from historic observed ground stations that are used in PRMS. Each NCAR-WG produces a range of mean statistics that varies from 1981-2010 through 2041-2070. It is observed CCSM, CESM, and GFDL-informed realizations' mean decreases from 1981-2010 to 2011-2040 before increasing for 2041-2070. Historical (1981-2010) CCSM, CESM, GFDL, and MPI-informed realizations centered closely on the historical mean (based on the median) and captures the range of the nine stochastic climate realizations. However, CESM tends to produce lower annual total precipitation than the historical (one realization of CESM goes as low as about 800 mm).
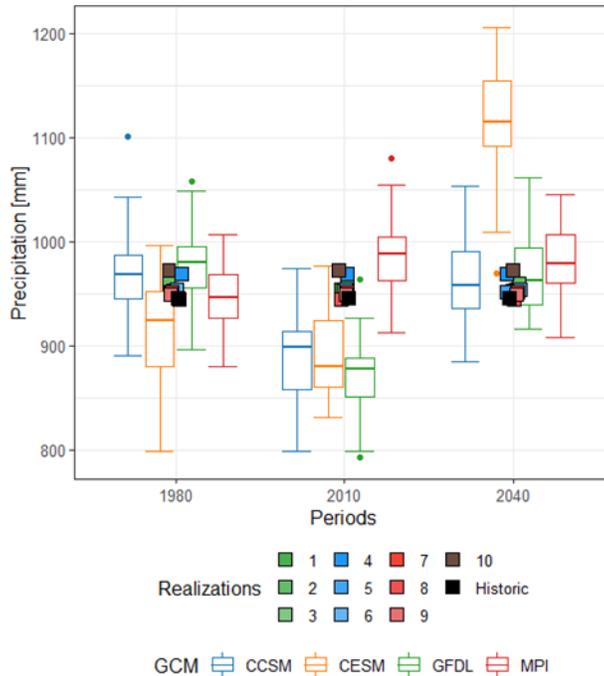
Figure 3-21. Comparison of the simulated annual average precipitation across the Upcountry region (average of 9 stations). Results from the NCAR-WG are shown using boxplots. Each color shows a different GCM used to conditions the NCAR-WG. Each boxplot summarizes the distribution across the 30 stochastic realizations. Results are shown for the baseline period (1981-2010, labeled '1980') and two futures periods (2011-2040 and 2041-2070, respectively labeled '2010' and '2040'). Results obtained from each stochastic realization (9 total) simulated via CliWxGen are shown with colored square symbols. CliWxGen realizations are meant to represent the 1956-2011 period (to ease the reading, results are repeated for each period). Note that a random noise was added to the x-axis of the square to ease reading of the figure. The historical realization is shown in brown color and the black squares show the annual average calculated using PRMS inputs (1970-2016). In the context of this comparison, the latter is considered being the 'truth', although the time periods used by each model vary.

Figure 3-22 illustrates the comparison results between NCAR-WG and CliWxGen in regards with variability of the precipitation variable. Variability at both annual and monthly scales is analyzed. Results show that simulations from NCAR-WG tend to overestimate the variability of the annual precipitation. It is noted for instance that the variability of the historical annual precipitation (black squares in Figure 3-22, left) is below the lower inter-quartile of the annual precipitation distribution across the ensemble of 30 realizations for each forcing GCM. The nine stochastic realizations from CliWxGen are closer to the historical value. Some slightly overestimate while some slightly underestimate the inter-annual variability

45

of the historical precipitation. As an ensemble, the nine stochastic realizations represent correctly the inter-annual variability of the annual historical precipitation across the Upcountry region.

The results of the comparison for the monthly temporal scale is significantly different than the one discussed for the annual scale. It is noted that simulations from NCAR-WG significantly overestimate the variability of the monthly precipitation across the region. On the other hand, the stochastic realizations from CliWxGen underestimate the precipitation variability at monthly scale. Note that the historical realization (brown) is close to the historic precipitation. A similar results is obtained a daily temporal scale, although not shown in this report.

Figure 3-22. Comparison of the standard deviation of the simulated precipitation across the Upcountry region (average of 9 stations). Left: annual temporal scale. Right: monthly temporal scale. See Figure 3-21 for more caption details.

4. East Bay region

4.1. Hydro-meteorological context

Three sub-watersheds of the Alameda Creek watershed (Figure 4-1) have been modeled for the need of the long term vulnerability assessment of the Hetch Hetchy Regional Water System (HHRWS); the Arroyo Hondo watershed (77.1 mi^2), the Upper Alameda Creek watershed (aka Alameda Creek Diversion Dam watershed or ACDD; 33.3 mi^2), and San Antonio watershed (37 mi^2). Flows from each of these catchments are collected, at least partially for ACDD, by SFPUC owned facilities; the Calaveras reservoir for Arroyo Hondo, the Alameda Creek diversion dam for ACDD watershed (which diverts water to the Calaveras Reservoir), and the San Antonio reservoir for the San Antonio watershed. Note that the Arroyo Hondo watershed outlet, which is the USGS gage 11173200, is located upstream the Calaveras reservoir.

Figure 4-1. Alameda hydrologic region subwatersheds. Subwatershed outlets are shown as black dots.

Despite the relatively small size of the East Bay region, precipitation vary considerably within the region due to a sharp gradient in elevation (ranging from few dozen of feet above sea level to nearly 4,400 feet at the top of the Mont Hamilton). Rainfall regime is strongly seasonal with most of the rainfall events occurring during the Fall and Winter seasons, followed by a pretty consistent meteorological drought extending from late April through September (Figure 4-2). In addition, precipitation during fall and winter months highlight strong inter-annual variability. For instance, Figure 4-2 illustrates that observed rainfall during winter months (i.e., December, January, February) at Calaveras and Sunol stations range from null values to more than 250 mm (i.e., 9.8 inches), while the interquartile spans from +/- 50m from the monthly mean.

Figure 4-2. Rainfall climatology (1956-10-01→2011-09-30) at Calaveras (left) and Sunol (right) rain gages. Bold black curves show the average monthly rainfall (mm). The gray areas show the deviation between the quantiles 75 and 25 of the monthly rainfall. Dash black curves show the minimum and maximum monthly rainfall for the period.

Given the significant rainfall variability above described, streams are dry during the summer drought period while flash floods occur frequently during winter months due to convective storms passing through the region and hitting the elevation ranges of the basin.

Figure 4-3 illustrates the hydrological regimes observed across the three watersheds and the significant inter-annual variability (cf. the difference between a wet year (i.e., water year 1998) and a dry year (i.e., water year 2007)). Note that snow pack dynamic has no influence in the region.
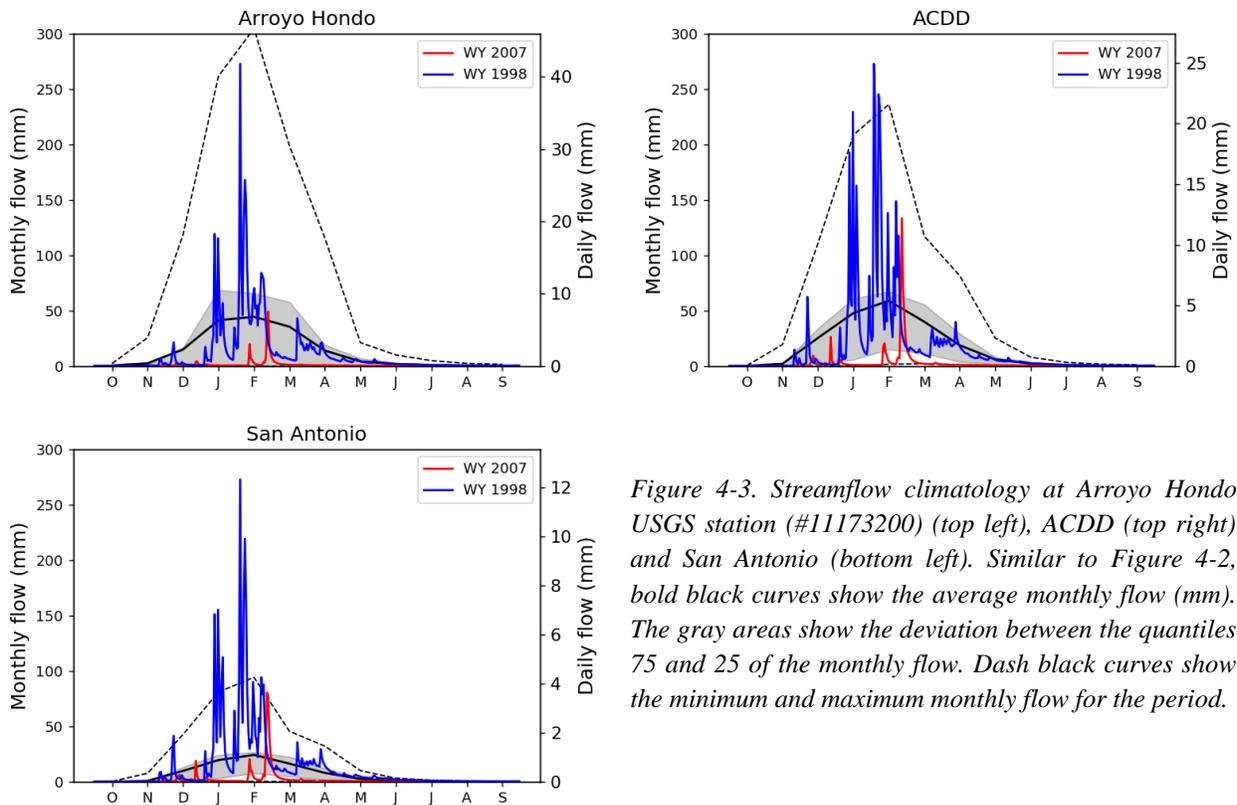
Figure 4-3. Streamflow climatology at Arroyo Hondo USGS station (#11173200) (top left), ACDD (top right) and San Antonio (bottom left). Similar to Figure 4-2, bold black curves show the average monthly flow (mm). The gray areas show the deviation between the quantiles 75 and 25 of the monthly flow. Dash black curves show the minimum and maximum monthly flow for the period.

48

4.2. SACramento Soil Moisture Accounting model (SAC-SMA)

For each of the three sub-watersheds in the East Bay region, a distributed version of the Sacramento Soil Moisture Accounting (SAC-SMA model; Anderson and McDonnell, 2005) model was built to generate inflows at their outlet. The SAC-SMA model is a well-known conceptual hydrology model that has been used for flood forecasting by the National Weather Service. It is often considered to benchmark model performance, such as for the CAMELS project in United-States (e.g., Newman et al. 2015). While a distributed version of SAC-SMA is used to model the considered East Bay catchments, the performance of the lumped SAC-SMA when compared to other hydrologic models across the United States is deemed average (Brunner et al. 2020; Kratzert et al. 2020). However, Brunner et al. (2020) show that SAC-SMA simulated peak flows are strongly related to precipitation, which, given the flashy feature of the considered catchments, makes SAC-SMA suitable for modeling the hydrology in East Bay region. In addition, the coupling between SAC-SMA and the Lohmann et al. (1978) river routing model makes SAC-SMA fully distributed and suitable for the topographic and climatic heterogeneity of the region. We use the acronym SAC-SMA-DS to refer to the distributed version of this coupled model. In addition to the Lohmann river routing model, SAC-SMA-DS consists of modules representing soil moisture accounting, evapotranspiration and runoff routing through a Nash-cascade unit hydrograph, depicted in Figure 4-4. Potential evapotranspiration (PET) in the SAC-SMA-DS method is calculated using the Hamon method, which uses daily mean temperature and daylight hours to determine daily PET. For river channel routing, the linearized Saint-Venant Equation is used. Appendix A includes further details, including equations and references.
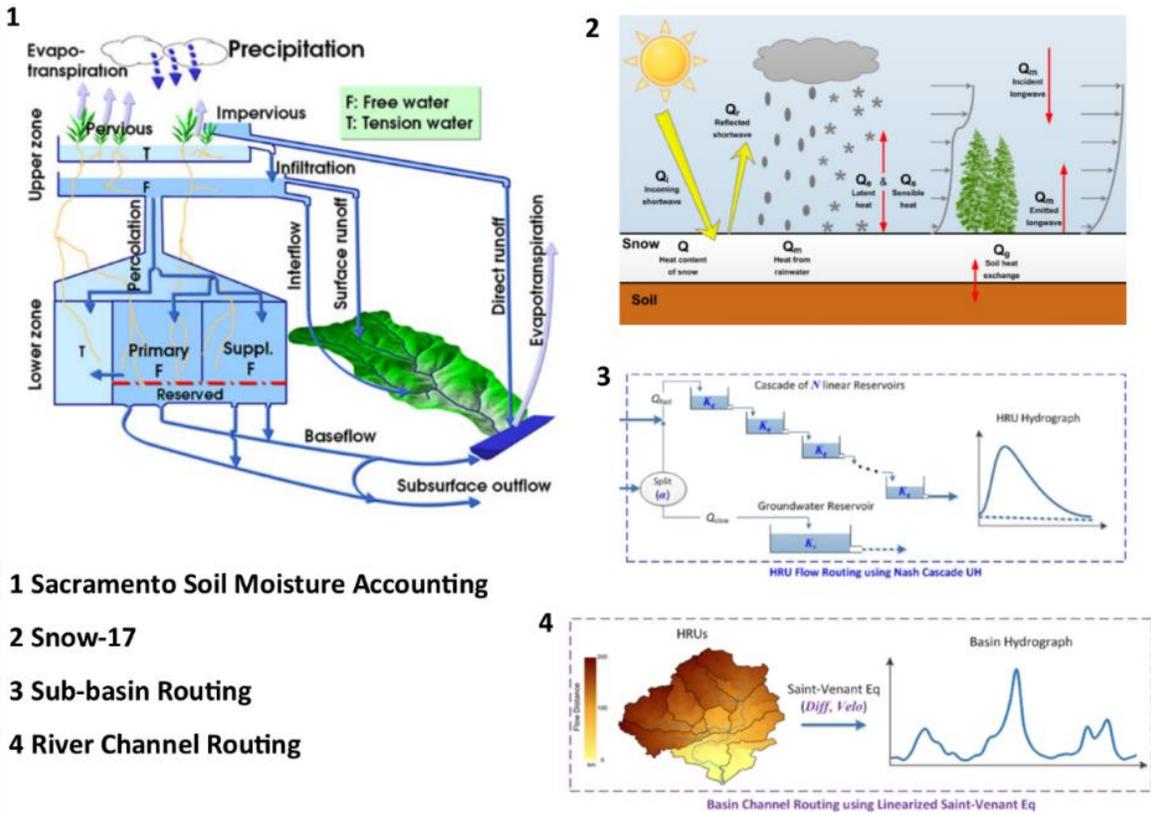
Figure 4-4. Schematic of the distributed Sacramento Soil Moisture Accounting model (SAC-SMA-DS). Note that the Snow-17 module (Anderson, 2006) is turned-off as snowpack dynamic has no influence in East Bay region.

49

4.3. Meteorological forcing

4.3.1. Precipitation

Given the significant gradient in elevation within the East Bay region, the rainfall is expected to vary significantly through space, although the rain gages with long records are sparse through the catchment, and thus, are not likely catching this variability. Three rain gages with long temporal records have been identified through the construction of the stochastic weather generator CliWxGen (HRG TR1, 2018); they are Calaveras, Sunol and Mont Hamilton rain gages. To assess whether the spatial coverage from the three rain gages is enough to represent the rainfall variability across the considered catchments, the high-resolution (~4 km^2) PRISM gridded dataset was used (http://prism.oregonstate.edu/). The result of this analysis is illustrated in Figure 4-5.
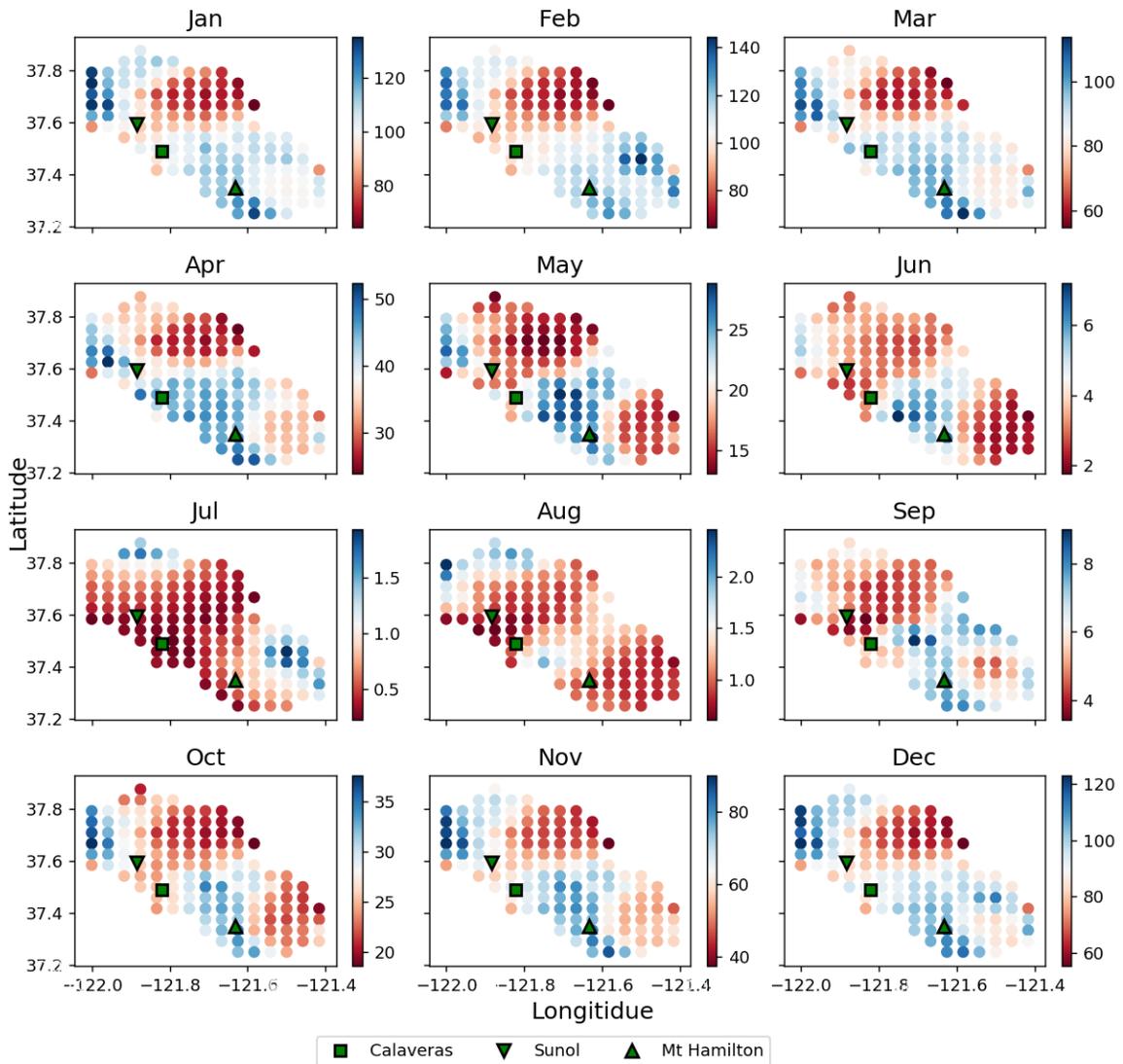
Figure 4-5. Monthly long term average of the PRISM precipitation (1981-2015). Monthly precipitation are indicated in mm (color bar). Each dot represents the monthly precipitation obtained for a PRISM grid cell within the simulated watersheds. PRISM spatial resolution is nearly 4km^2. The location of the three rain gages with long term records that are used by the weather generator CliWxGen are indicated with green symbols.

50

Figure 4-5 illustrates the long-term monthly precipitation cumuli at 4km^2 resolution from PRISM dataset. Only the PRISM grid cells within the considered catchments are illustrated. Two main results can be drawn out of this figure. First, the variability of the precipitation is significant across the region, as expected given the gradient in elevation. The precipitation variability through space is especially important during the raining season (i.e, during fall and winter months). For instance, precipitation across the San Antonio reservoir appear to be significantly lower than the one across the Arroyo Hondo and ACDD catchments. The variability within the Arroyo Hondo and ACDD catchments is also important, with high elevation (i.e., where the Mont Hamilton rain gage is) receiving more water than the lower part of the Arroyo Hondo catchment or even the upper region of the ACDD catchment. The second result one can draw out of the figure is that the three available rain gages, illustrated with green symbols in Figure 4-5, do not capture well the observed variability within PRISM dataset.

When reliable gage observations are available, they are commonly deemed more trustful than any interpolation-based gridded dataset such as PRISM. However, it seems important to account for the spatial distribution of the rainfall across the region, which, given the available stations, is not possible. For doing so, the observation at the gages were used to estimate the precipitation at the PRISM grid cell level by mean of a scaling factor that accounts for the difference in long-term averages between the PRISM grid cell where the rain gage is, and the PRISM grid cell where one want to estimate the rainfall:

$$P_k(t) = P_{i_k}(t) \frac{\overline{PRISM_{j_i}}(m(t))}{\overline{PRISM_k}(m(t))}, \qquad \textit{4-1}$$

with $P_k(t)$ the estimated precipitation for the day t at the PRISM grid cell k, P_{i_k} the precipitation measured at the gage i being the closest to the PRISM grid cell k (i could be any of the available rain gage). The Euclidian distance using the latitude and longitude coordinate is used to find the closest station and/or grid cell[1]. $\overline{PRISM_{j_i}}(m(t))$ and $\overline{PRISM_k}(m(t))$ are the long-term averages (i.e., 1981-2015) precipitation obtained from PRISM for the grid cells j_i and k and for the current month m. Note that the grid cell j_i is the PRISM grid cell that in which the gage i_k is located. The use of the ratio between the PRISM precipitation at the gage location and at the grid cell where the precipitation is estimated is justified by the pretty good representation of the long-term precipitation at the gage by the PRISM dataset, as illustrated in Figure 4-6. Only the precipitation during the wetter months at Mont Hamilton seems slightly overestimated by PRISM. For the Calaveras and Sunol gages, the match is very good for all months.

---

[1] As discussed further down in the section, the rain gage at Mont Hamilton was disregarded. As such, SAC-SMA-DS models for Arroyo Hondo and ACDD sub-watersheds only used Calaveras rain gage. For San Antonio sub-watershed, 4 out of 12 PRISM grid cells were assigned with Sunol rain gage, while the remaining 8 grid cells were assigned with Calaveras rain gage.
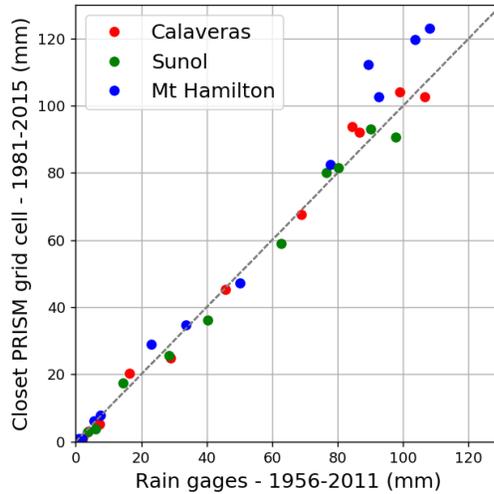
Figure 4-6. Comparisons between long term average monthly precipitation at each rain gage and PRISM precipitation. Each dots represent one month.

Note that during the revision of the LTVA and hydrology reports, discussion with SFPUC personnel leads to disregard the Mont Hamilton station for two reasons. First, as illustrated in Figure 4-7, a significant trend in annual precipitation is observed at Mont Hamilton station. This trend is questionable because no trend is observed at Calaveras and Sunol gages. Second, the presence of such trend is not compatible with the CliWxGen stochastic weather generator that, by construction, requires stationary rainfall time series.
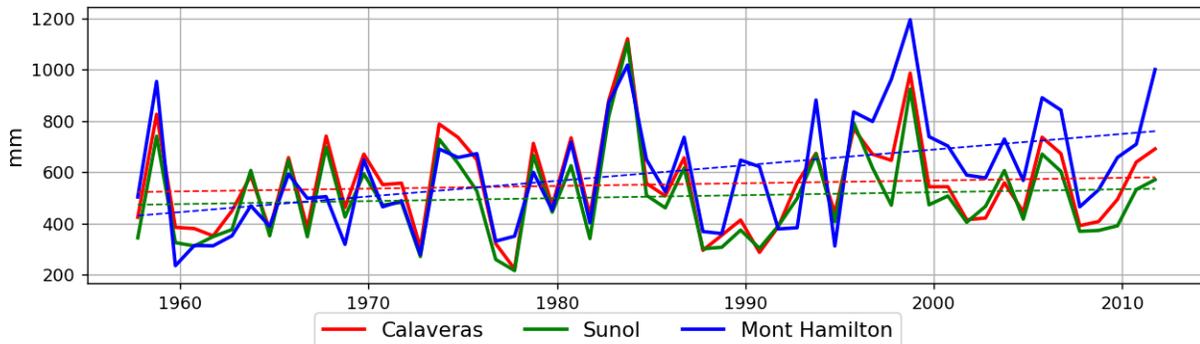
Figure 4-7. Annual precipitation observed at Calaveras (red), Sunol (green) and Month Hamilton (blue) gages. Annual values are given in mm. Colored dash curves show linear trend obtained for each station.

4.3.2. Temperature

A comparison between the available temperature stations across the region and the Livneh et al. (2013) gridded dataset (aka CONUS) has been carried out during the development of the CliWxGen stochastic generator. It has been found that CONUS dataset represents well enough the temperature across the region. Following this development and findings, it has been decided in agreement with SFPUC personnel to use CONUS temperature gridded dataset as input to SAC-SMA-DS model. More details regarding this comparison is available in Section 2.1 of the Weather Generator report (HRG TR1, 2018). The temperature from the CONUS grid elevation is adjusted to the PRISM grid elevation using a temperature lapse equal to 6.2ºC /km.

4.4. Model calibration across the three considered East Bay catchments

This section presents the results of the calibration/validation of SAC-SMA-DS model at the outlet of Arroyo Hondo, ACDD and San Antonio sub-watersheds. The calibration of SAC-SMA-DS model parameters was carried out using the Genetic Algorithm described in Appendix E. The objective function for the calibration is the Kling-Gupta Efficiency (KGE; Gupta et al. 2009) obtained from the daily simulated and observed runoff time series. The calibration and validation periods shown in Table 4-1 have been chosen given i) the availably of the observed/reconstructed discharges and ii) the goal of keeping dry conditions during the validation period to check whether the calibrated model parameters are able to correctly represent dry years. Note that the most recent drought between fall 2011 and fall 2015 could not be included in the calibration/validation of any model because the temperature forcing dataset ends in 2011.

Arroyo Hondo model was calibrated over the most recent period (i.e., from 1995/10/01 to 2011/09/30) and validated across an older period (i.e., from 1969/10/01 to 1982/09/30). The validation period includes the 1976-1977 drought, which provides a good test-bed to validate the Arroyo Hondo model during dry conditions. Note that because no observed discharge is available in-between 1982 and 1995, this period could not be used for calibration/validation. The time period for which observed/reconstructed discharge time series at the outlet of the ACDD and San Antonio sub-watersheds is significantly shorter and only includes the recent period spanning from 1996/10/01 to 2011/09/30. For both catchments, the calibration was carried out through the period 1996/10/01 to 2006/09/30, leaving the period 2006/10/01 to 2011/09/30 for validation. This split has the advantage of having the driest year of the available record (i.e., water year 2007) for validation, which helps ensuring that calibrated models are able to reproduce dry conditions.

Table 4-1. Calibration and validation period used for each catchment in East Bay region

Basin name	Area (square mile)	Calibration period	Validation period	Objective function	Nature of the target
Arroyo Hondo	77.1	1995/10/01 → 2011/09/30	1969/10/01 → 1982/09/30	KGE	USGS gage number 11173200
ACDD	33.3	1996/10/01 → 2006/09/30	2006/10/01 → 2011/09/30	KGE	Reconstructed by SFPUC
San Antonio	37	1996/10/01 → 2006/09/30	2006/10/01 → 2011/09/30	KGE	Reconstructed by SFPUC

Calibration and validation results are illustrated below using two set of figures and one set of tables. Figure 4-8, Figure 4-10 and Figure 4-12 summarize the performance for Arroyo Hondo, ACDD and San Antonio watersheds, respectively. These figures are collection of four plots; a) the daily observed and simulated runoff time series obtained during calibration and validation periods; b) the observed and simulated annual time series during calibration and validation periods; c) a scatter plot of the monthly simulated versus observed runoff; and d) the flow duration curves of the observed and simulated runoff. These plots are further used below to illustrate the performance of the SAC-SMA-DS models at the daily, monthly and annual time scales. Figure 4-9, Figure 4-11 and Figure 4-13 are used to illustrate for each water year the match between the observed and simulated time series for each catchment. Performance metrics, including the KGE and NSE estimated for both daily and monthly temporal scales, together with

the percent bias, are given in Table 4-2, Table 4-3 and Table 4-4 for Arroyo Hondo, ACDD and San Antonio sub-watersheds, respectively.

4.4.1. Arroyo Hondo

The performance of the SAC-SMA-DS model at the Arroyo Hondo gage is very good. Daily NSE and KGE criteria are high (i.e., larger than 0.69 and larger than 0.84 respectively). The model is not biased with even less than 5% bias obtained for the validation period. We also note a good performance at reproducing the dry years during the validation period (i.e., 1972, 1976 and 1977). However, the model tends to underestimate the peak discharge, which is common for small and flashy basins. Higher temporal resolution (i.e., hourly) could be require to better reproduce peak flows. Overall, based on the various performance criteria, the model performance is satisficing for the need of the LTVA.
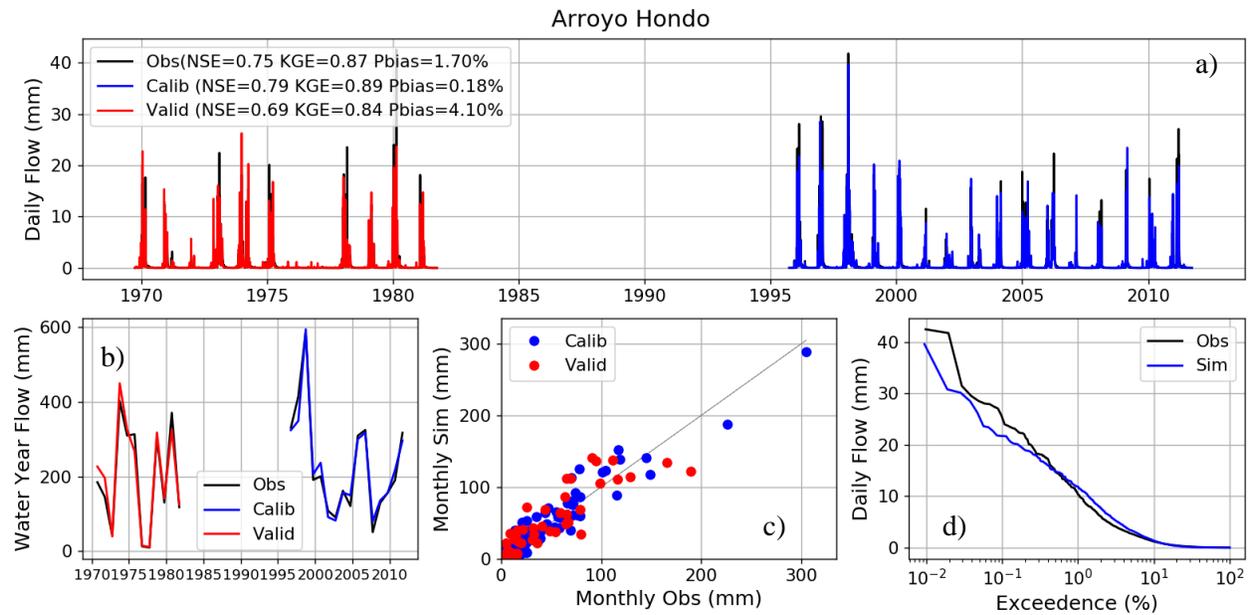
Figure 4-8. Calibration results for the Arroyo Hondo watershed. a) Comparison of the daily observed (black) and simulated time series during calibration (blue) and validation (red) period. Nash-Sutcliffe (NSE) and Kling-Gupta (KGE) efficiency and the percent bias (Pbias, %) are given for the calibration, validation and the entire periods; b) Same as a) but for the annual scale (i.e., water year from 10/01 to 09-30). The data point for the WY XXXX is located at the date 09/30/XXXX; c) Scatter plot showing the monthly simulated vs. observed discharges. Calibration and validation periods are illustrated in blue and red color respectively. d) Simulated (blue) and observed (black) flow duration curve. Here, the simulated flow duration curve include both calibration and validation periods.

Table 4-2. Summary of the SAC-SMA-DS performance for the Arroyo Hondo sub-watershed. Nash-Sutcliffe and Kling-Gupta efficiencies are given for both daily and monthly temporal scales.

	NSE_daily	KGE_daily	Pbias(%)	NSE_monthly	KGE_monthly
Full period	0.75	0.87	1.70	0.83	0.87
Calib (1996/10/01→2011/09/30)	0.79	0.89	0.18	0.93	0.96
Valid (1969/10/01→1981/09/30)	0.69	0.84	4.10	0.84	0.91
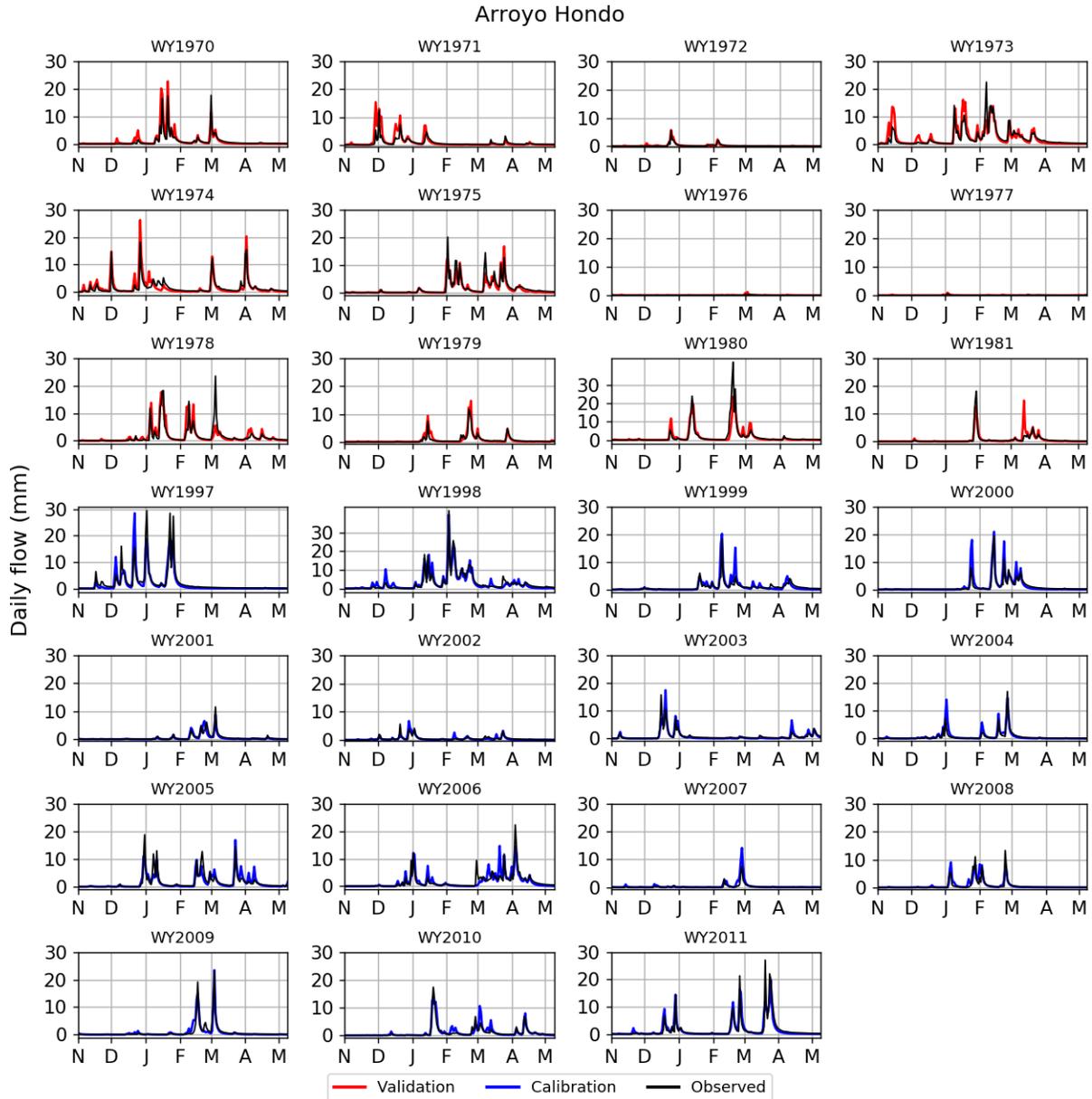
Figure 4-9. Comparison of the observed (black) and simulated flow at the outlet of Arroyo Hondo catchment. Blue and red curves are used to highlighted calibration and validation periods. Only November through early May periods are shown for each water year to increase readability of the figure.

4.4.2. ACDD

Similar to Arroyo Hondo sub-watershed, SAC-SMA-DS performance is high for the ACDD sub-watershed. Daily NSE and KGE are respectively larger than 0.78 and 0.82 for both calibration and validation periods. The percent bias is low for calibration and the whole periods (0.86% and 3.85% respectively) but it is larger for the validation period (i.e., 13.4%). This rather large bias during the validation period seems to result from the water year 2010 (Figure 4-11), which could follow from an overestimation of the precipitation during this year. However, since the driest year (i.e., WY2007) is well reproduced, the model is expected to represent sufficiently well the streamflow conditions during dry

years. Given the calibration and validation results, the model at ACDD is deemed satisfying for the need of the LTVA.
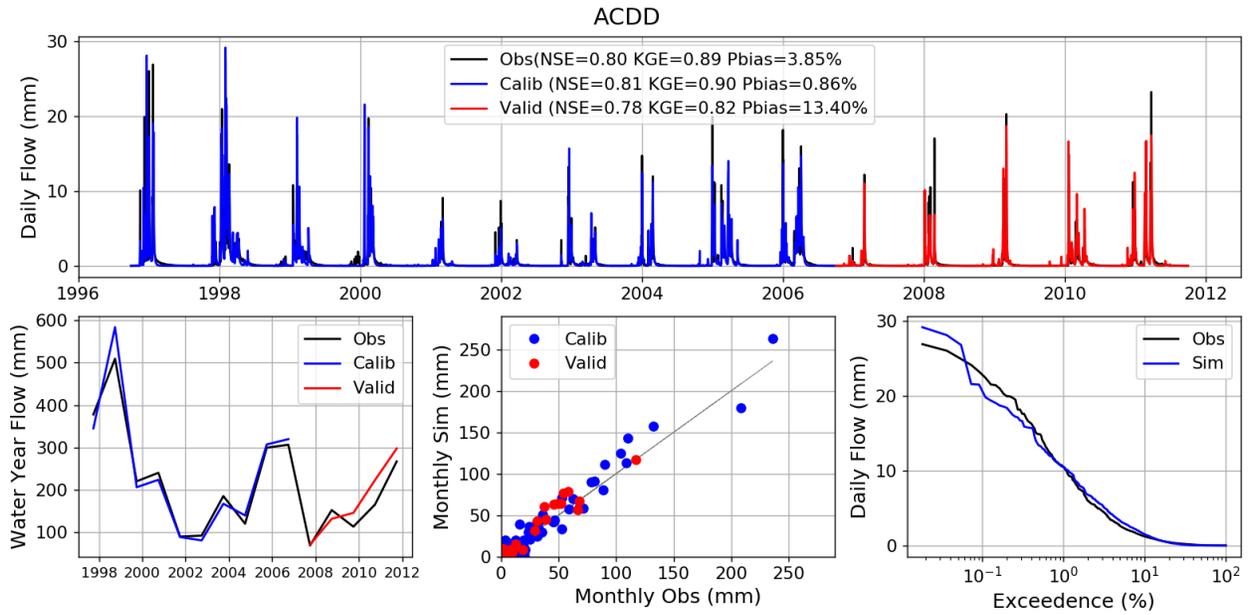
Figure 4-10. Same as Figure 4-8 but for ACDD sub-watershed.

Table 4-3. Summary of the SAC-SMA-DS performance for the ACDD sub-watershed. Nash-Sutcliffe and Kling-Gupta efficiencies are given for both daily and monthly temporal scales.

	NSE_daily	KGE_daily	Pbias(%)	NSE_monthly	KGE_monthly
Full period	0.80	0.89	3.85	0.95	0.90
Calib (1996/10/01➔2006/09/30)	0.81	0.90	0.86	0.95	0.92
Valid (2006/10/01--➔2011/09/30)	0.78	0.82	13.40	0.91	0.80
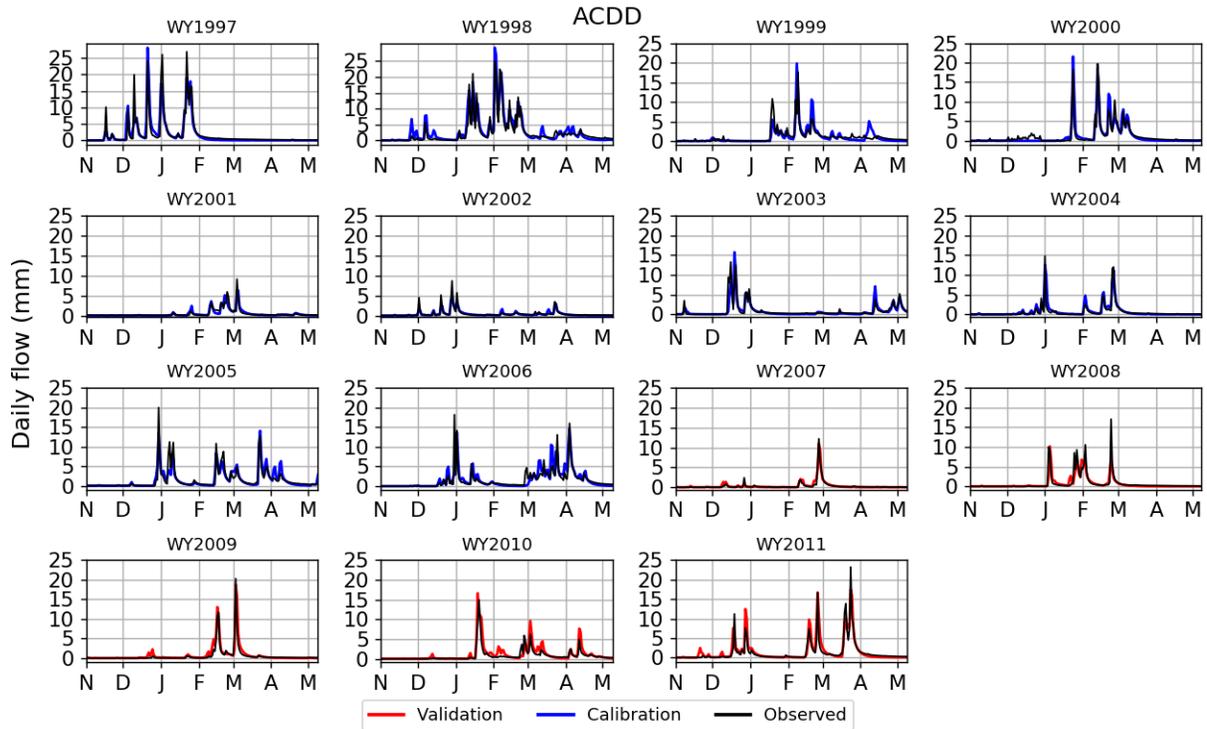
Figure 4-11. Comparison of the observed (black) and simulated flow at the outlet of ACDD catchment. Blue and red curves are used to highlighted calibration and validation periods. Only November through early May periods are shown for each water year to increase readability of the figure.

4.4.3. San Antonio

Similar to the Arroyo Hondo and ACDD sub-watersheds, SAC-SMA-DS model performance is very good with daily NSE and KGE larger than 0.8 and 0.84 for both calibration and validation periods. Overall bias is low with, however, a slightly larger value obtained for the validation period (i.e., 12.4%), which, similarly to ACDD, seems to mostly result from the water year 2010. The performance during the dry water year 2007 is also very good, giving confidence that the model represents correctly the dry years.
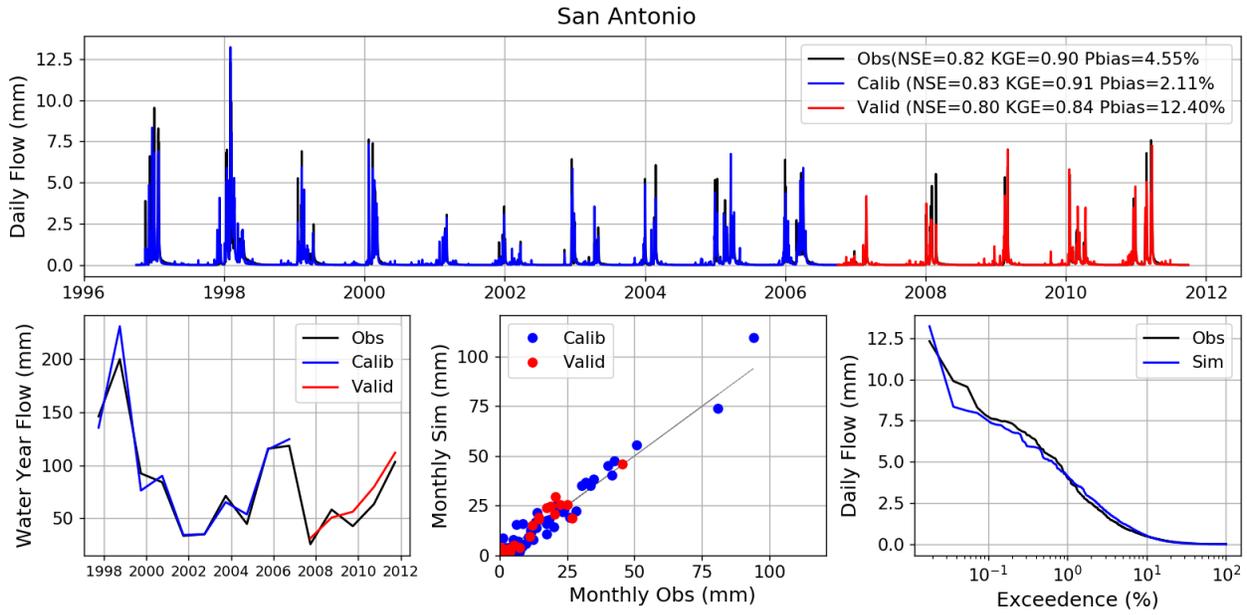
Figure 4-12. Same as Figure 4-8 but for San Antonio sub-watershed.

Table 4-4. Summary of the SAC-SMA-DS performance for the San Antonio sub-watershed. Nash-Sutcliffe and Kling-Gupta efficiencies are given for both daily and monthly temporal scales.

	NSE_daily	KGE_daily	Pbias (%)	NSE_monthly	KGE_monthly
Full period	0.82	0.90	4.55	0.96	0.92
Calib (1996/10/01→2006/09/30)	0.83	0.91	2.11	0.96	0.93
Valid (2006/10/01→2011/09/30)	0.80	0.84	12.40	0.93	0.86
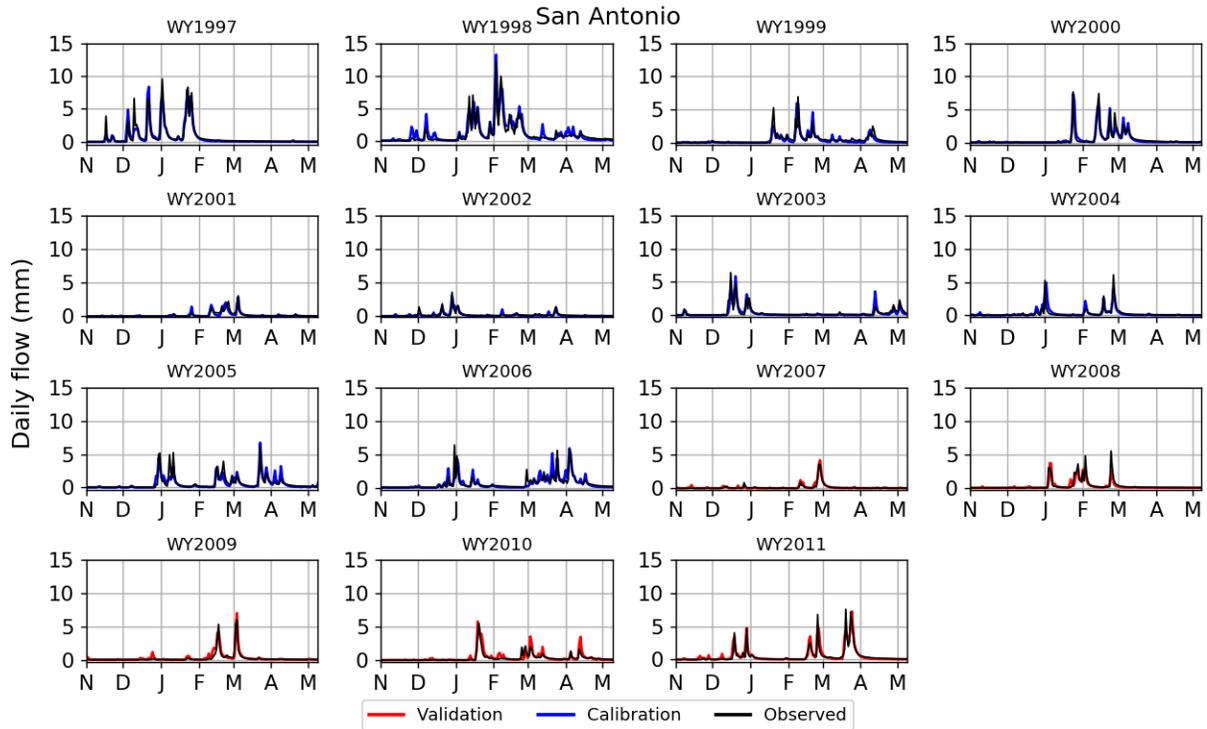
Figure 4-13. Comparison of the observed (black) and simulated flow at the outlet of San Antonio catchment. Blue and red curves are used to highlighted calibration and validation periods. Only November through early May periods are shown for each water year to increase readability of the figure.

4.5. SAC-SMA-DS simulations forced by the stochastic weather realizations

This section presents the results of the SAC-SMA-DS simulations when forced by the output of the CliWxGen weather generator. First, the forcing dataset (i.e., simulated temperature and precipitation time series) obtained through the weather generator are compared with the observed weather variables to detect any potential bias. Results show a minor bias in precipitation, which has been corrected prior forcing the SAC-SMA-DS model (Section 4.5.1). Second, the bias in runoff variables are discussed for the three sub-watersheds (Section 0).

4.5.1. Bias correction of the weather generator outputs

Figure 4-14 illustrates for Arroyo Hondo (top), ACDD (middle) and San Antonio (bottom) the relative deviation in average between the stochastic realizations obtained from the weather generator CliWxGen and the historical weather. The description of the nine realizations can be found in the technical report describing the weather generator CliWxGen (HRG TR1, 2018). Note that in Figure 4-14, precipitation and temperature time series used to estimate the deviation between stochastic and historical time series are average across each basin after the pre-processing detailed in Section 4.3.

59
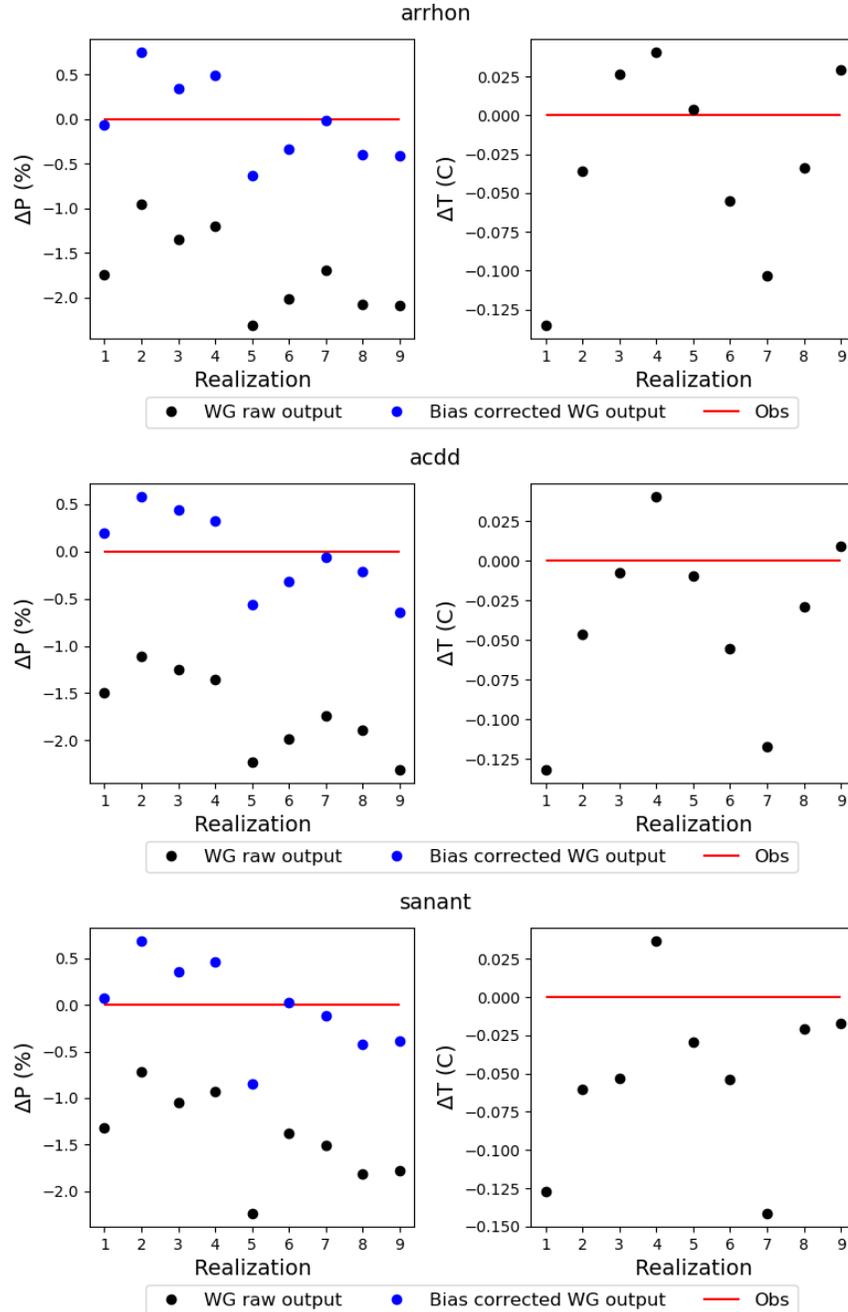
Figure 4-14. Comparison between the SAC-SMA-DS forcing dataset for the LTVA with the historical forcing dataset. Black dots show for each realization (x-axis) the average deviation between the historical record for the precipitation (left column) and temperature (right column) variables and the raw output of the weather generator CliWxGen mapped to the PRISM grid cells (see Equation 4-2). Blue dots show the deviations from the historical obtained with the bias corrected realizations. The y-axis show the deviation from the historical values, relative for precipitation and absolute (in ^{o}C) for temperature.

Two main results are drawn out from Figure 4-14. First, none of the nine stochastic realization is biased regarding the temperature variable (right column). The largest deviation from the historical average is lower than 0.15ºC, which is negligible. Second, all stochastic realizations show a dry bias ranging from 1

to 2.5% reduction when compared to the historical average. The bias is rather consistent across all three sub-watersheds. To ensure that overall, the baseline used for the climate stress test have a similar long-term precipitation average than the historical record, the weather generator precipitation output must be corrected prior altering their average throughout the climate stress test.

Let us recall that the weather generator CliWxGen outputs precipitation time series at the Calaveras and Sunol gage locations. Let us refer to these time series as $P_{i,CliWxGen}$ (i being an index to either refer to Calaveras or Sunol rain gage). The precipitation time series from the weather generator are first mapped to the each PRISM grid k using Equation 4-1, which becomes:

$$P_{k,CliWxGen}(t) = P_{i_k,CliWxGen}(t)\frac{\overline{PRISM_{J_i}}(m(t))}{\overline{PRISM_k}(m(t))}, \qquad 4\text{-}2$$

with similar notation as the ones used for Equation 4-1. The precipitation time series from the weather generator mapped to the each PRISM grid cell k are then corrected by multiplying their daily values $P_{k,CliWxGen}(t)$ by a correction factor b_w that is specific to each sub-watershed w. Note that for a given sub-watershed w, the correction factor b_w is common across all realizations:

$$P_{k,CliWxGen,corr}(t) = b_w P_{k,CliWxGen}(t), \qquad 4\text{-}3$$

with $P_{k,CliWxGen,corr}$ the bias corrected precipitation time series obtained from the weather generator and mapped to the PRISM grid cell k. Bias correction factors b_w for the sub-watershed w is estimated from the deviations illustrated in Figure Figure 4-14 (left column) betweeh the historical precipitation (red line) and the raw weather generator output mapped to the PRISM gridd cells (black dots):

$$b_w = \frac{\overline{P_k}(t)}{\overline{P_{k,CliWxGen}}}, for\ k\ \epsilon\ w \qquad 4\text{-}4$$

with $P_k(t)$ the estimated precipitation for the day t at the PRISM grid cell k (obtained from Equation 4-1). In Equation 4-4, the average of the precipitation from the weather generator mapped to the PRISM grid cells $\overline{P_{k,CliWxGen}}$ is calculated across all nine stochastic realizations. The bias correction factors are given in Table 4-5 for each sub-watershed.

Table 4-5. Bias correction factors to be multiplied with the precipitation time series obtained from Equation 4-1.

Arroyo Hondo	San Antonio	ACDD
1.017	1.014	1.017

4.5.2. SAC-SMA-DS simulations for the LTVA

The long-term averages of observed and simulated runoff are compared in Table 4-6. The average monthly cycles are illustrated in Figure 4-15. It is important to recall that long-term averages and average cycles presented below are obtained for different time periods, which might explain some differences. The averages for the observed historical records vary from one catchment to another (cf. periods in Table 4-6 caption). For all three sub-watersheds, the length is significantly shorter than the period for which runoff are simulated with SAC-SMA-DS forced by the stochastic realization. The long-term averages for the simulated historical is obtained for the period spanning from 1956/10/01 to 2011/09/30. As described in the weather generator technical report (HRG TR1, 2018), each of the nine stochastic realization is 50-

year long and has been obtained from simulation of the CliWxGen model that has been built using the observed weather from 1956/10/01 to 2011/09/30. Although each realization has been obtained through use of resampling approaches, which introduce somewhat randomness, they should be representative of this period; the deviation from one realization to another is deemed noise stemming from climate variability and sampling uncertainty.

For Arroyo Hondo, both observed and simulated historical long-term averages fall within the range of long-term averages obtained from the ensemble of 9 realizations. When compared to the observed historical, one may note a slight wet bias that averaged out to 1.2%. On can also note that the average simulated runoff peak during the rainy season is larger than the one from the historical record. This difference could be explained by the fact that the wetter year (i.e., 1984), together with some other wet years (i.e., 1959 and 1983) are not present in the historical runoff record, although they are used by the weather generator, and are thus somewhat included within the 9 stochastic realizations.

For San Antonio sub-watershed, one can note that the long-term averages obtained with the nine stochastic realizations are slightly larger than the observed historical long-term average. The positive bias ranges from 1.2% (realization #3) to 9.2% (realization 6). The average bias across realization equals 4.8%. This positive bias could be explained by the shorter observed historical period. As the historical record is short (i.e., 19 years) its long-term average may be more sensitive to low frequency climate variability. As such, the number of dry years within the observed historical time series, including the start of the most recent drought (i.e., 2011 through 2014), might introduce a dry bias into the long-term average historical estimate when compared to its theoretical true (but unknown) value that could be better estimated using a longer record. Also, three out of the four wettest years on record are prior 1996 (i.e., water years 1959, 1982, 1983 as seen in Figure 4-7). As such, it is expected that some years from the ensemble of 9 realizations to have larger streamflow values than the ones reconstructed at the San Antonio catchment outlet, which could also contribute to explain the apparent positive bias when compared to the 1996-2014 historical record.

For ACDD, a similar result than for San Antonio sub-watershed is observed. In this case, the positive bias obtained from the 9 realization ranges from 3.2% to 9.7% (with an average equal to 5.8%).

The above assumption that the long-term averages obtained for San Antonio and ACDD sub-watersheds are potentially underestimating the true value is supported by the fact that the model does not overestimate the runoff at Arroyo Hondo sub-watershed, for which the historical long-term average estimate was obtained using a longer period (i.e., 35 years).

Table 4-6. Long-term averages for observed runoff ('Observed historical'), simulated runoff forced with historical weather ('Simulated historical') and simulated runoff forced with the stochastic weather realizations (indexed from 1 through 9). Runoff values are given in mm. Note that 'Historical' averages are given for the longest available period (i.e., 1969/10/01→ 2018/09/30 for Arroyo Hondo (with missing values from 1981/10/01 to 1995/09/30); 1995/10/01→ 2013/09/30 for San Antonio; 1995/10/01→2014/09/30 for ACDD). The period used for simulated historical is the one used to generate the stochastic realization (i.e., 1956/10/01 → 2011/09/30).

Realization id	Arroyo Hondo	San Antonio	ACDD
Observed historical	0.576	0.223	0.552
Simulated historical	0.574	0.227	0.573

1	0.576	0.229	0.581
2	0.609	0.244	0.606
3	0.567	0.226	0.570
4	0.586	0.234	0.588
5	0.575	0.230	0.578
6	0.571	0.231	0.573
7	0.598	0.241	0.597
8	0.576	0.227	0.579
9	0.589	0.234	0.590

Regarding the seasonal cycles illustrated in Figure 4-15, their representation is fairly good for all three sub-watersheds. One may notice, however, that simulated dry period starts and ends earlier than in the historical records. This difference could be explained by several factors that have not been thoughtfully investigated because deemed not critical for the performance of the LTVA. Note that difference is only about few mm which represents, given the size of the sub-watersheds, a tiny fraction of the annual runoff.
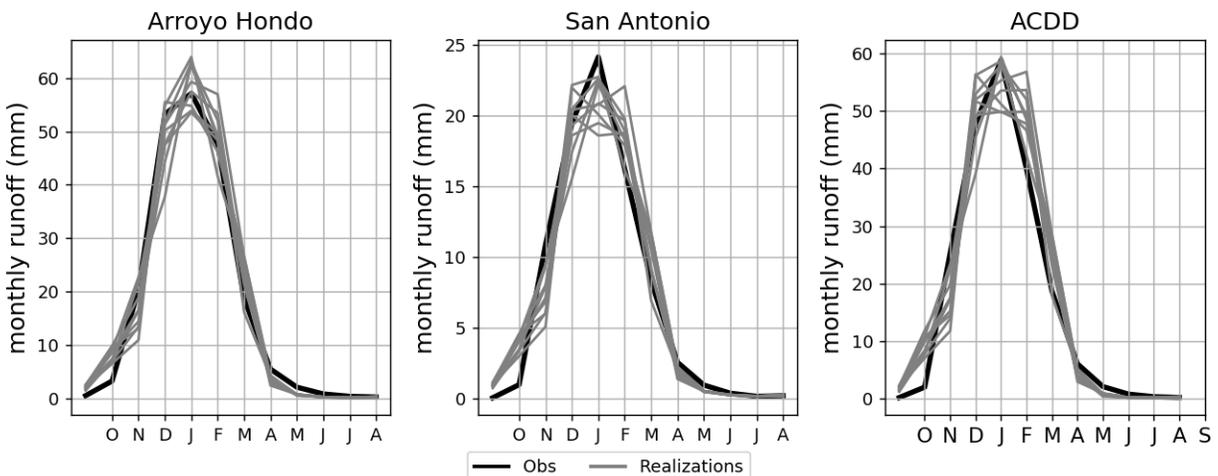
Figure 4-15. Comparison of the observed historical (black) and LTVA simulated (grey) runoff for each sub-watershed. All nine SAC-SMA-DS simulations driven by the weather generator outputs are represented in this figure (grey curves). Runoff is given in mm.

4.6. Generation of the input for the San Francisco Water System Model

Note first that the San Francisco Water System Model (SFWSM) uses acre feet by day (AF/day) as unit. As such, everything single SAC-SMA-DS outputs is converted from mm to AF/day.

4.6.1. Runoff for the ungagged sub-watersheds

In addition to the streamflow time series simulated at the outlet of the Arroyo Hondo, ACDD and San Antonio sub-watersheds, SFWSM may require up to two additional runoff time series.

One required time series is for the intermediate watershed in-between the Arroyo Hondo gage, where the SAC-SMA-DS model is calibrated, and the Calaveras reservoir. Following discussion with SFPUC personnel, runoff from this intermediate sub-watershed is obtained by scaling the simulated time series at Arroyo Hondo by an adjustment factor equal to 0.21. In other word, the actual runoff entering Calaveras

reservoir (not accounting from the diversion from ACDD tunnel) equals 1.21 times the simulated runoff at the Arroyo Hondo gage.

The second time series is optional and is only required for the SFWSM version that accounts for the Recapture Alameda Creek project. In this case, a time series of runoff is required to represent the contribution from the catchment in-between the Calaveras reservoir and the pit F2 entrance. Following discussion with SFPUC personnel, it has been decided to scale the simulated time series at the San Antonio reservoir by the ratio of the catchment areas. This ratio was found to be equal to 0.975.

4.6.2. Evaporation and precipitation over Calaveras and San Antonio reservoirs
SFWSM requires precipitation and evaporation rates over the reservoirs. The simulated potential evapotranspiration by SAC-SMA-DS and the precipitation at the grid cell that includes the outlet of the Arroyo Hondo and San Antonio sub-watersheds were outputted for that purposed.

4.7. SAC-SMA-DS simulations forced by the stochastic weather generator climate stress test
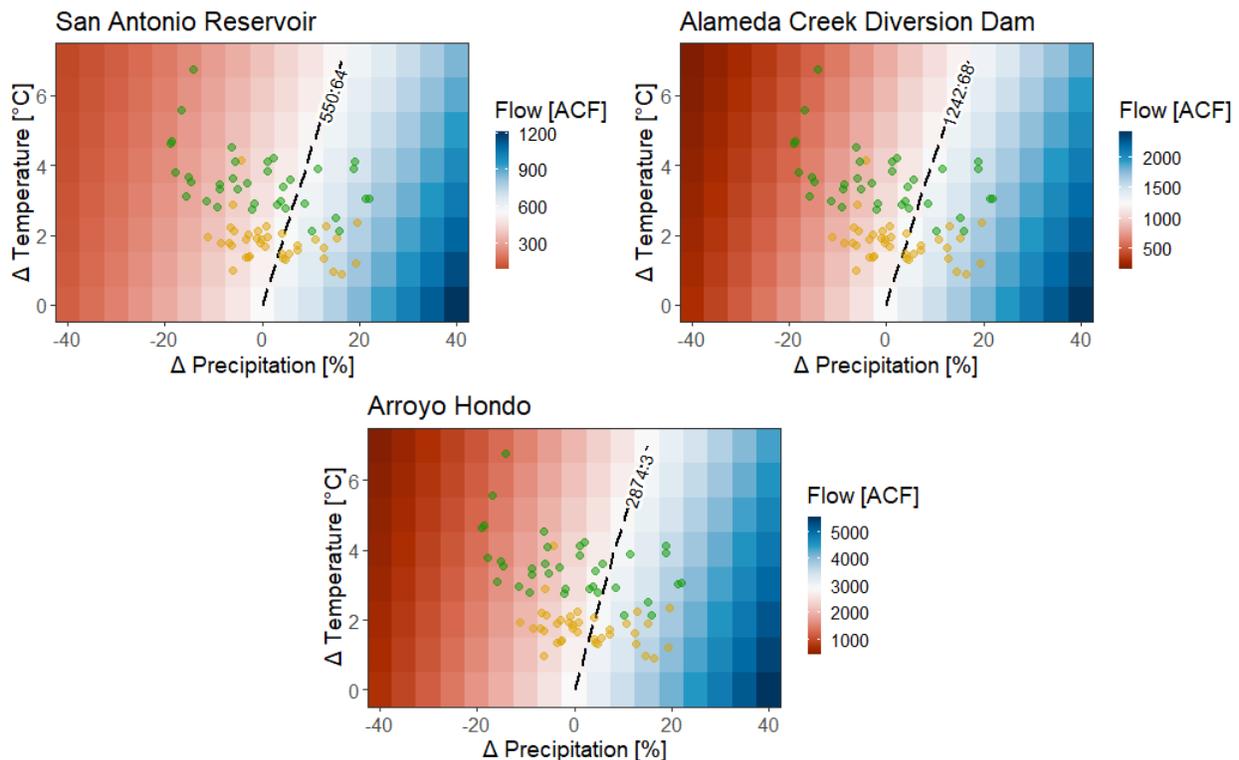
Figure 4-16 Total annual stream flow (acre-feet) under temperature and precipitation changes respectively for San Antonio Reservoir, Alameda Creek Diversion Dam, and Arroyo Hondo sub-watersheds. The yellow and green dots over the response surface shows CMIP5 projections under RCP 8.5 for two 30-year long periods: 2040 (2026-2055) and 2070 (2056-2095). Baseline is 1986-2005. Historic mean observed is shown in white and follows the dashed line.

64

Figure 4-16 shows the total annual stream flow under temperature and precipitation changes. The higher precipitation generates more flows while higher temperature produces less flows. This change is due to additional water collected in the watershed from precipitation events and more evapotranspiration due to increase in temperature. Being a smaller watershed than the Upcountry region, changes in evapotranspiration has a larger impact on the annual streamflow. Results below are presented for inflow at Calaveras reservoir (Arroyo Hondo) the largest watershed in the East Bay.

- By 2040, the median projections of +2°C warming combined with 0% change in mean annual precipitation results in 9% reduction in mean annual inflow.
- By 2040, most projections and elicitations of warming are between +1°C and +4°C and for precipitation change between -5% and +5% which would correspond to a maximum change in mean annual inflow between a decrease of 27% and an increase of 7%.
- By 2070 RCP 8.5, the median projections of about +4°C combined with 0% change in mean annual precipitation results 17% decrease in mean annual inflow. Most projections and elicitations of warming range between +3°C and +6°C and of precipitation change between -15% and +15% resulting in change in mean annual inflow between a decrease of 50% and an increase of 33%.

5. Peninsula Hydrologic Model

5.1. Introduction

The Peninsula Watershed, almost completely owned by the SFPUC, is located south of the City of San Francisco in central San Mateo County. The Watershed contains three reservoirs, San Andreas, Crystal Springs and Pilarcitos. In addition to serving as storage facilities for the water draining the Peninsula Watersheds, these reservoirs also store water from the Upcountry system.

The Peninsula watershed has the area of 23,000 acres, located in the central San Mateo County and at the South of San Francisco City and County (shown in Figure 5-1) (San Francisco Planning Department, 2001).
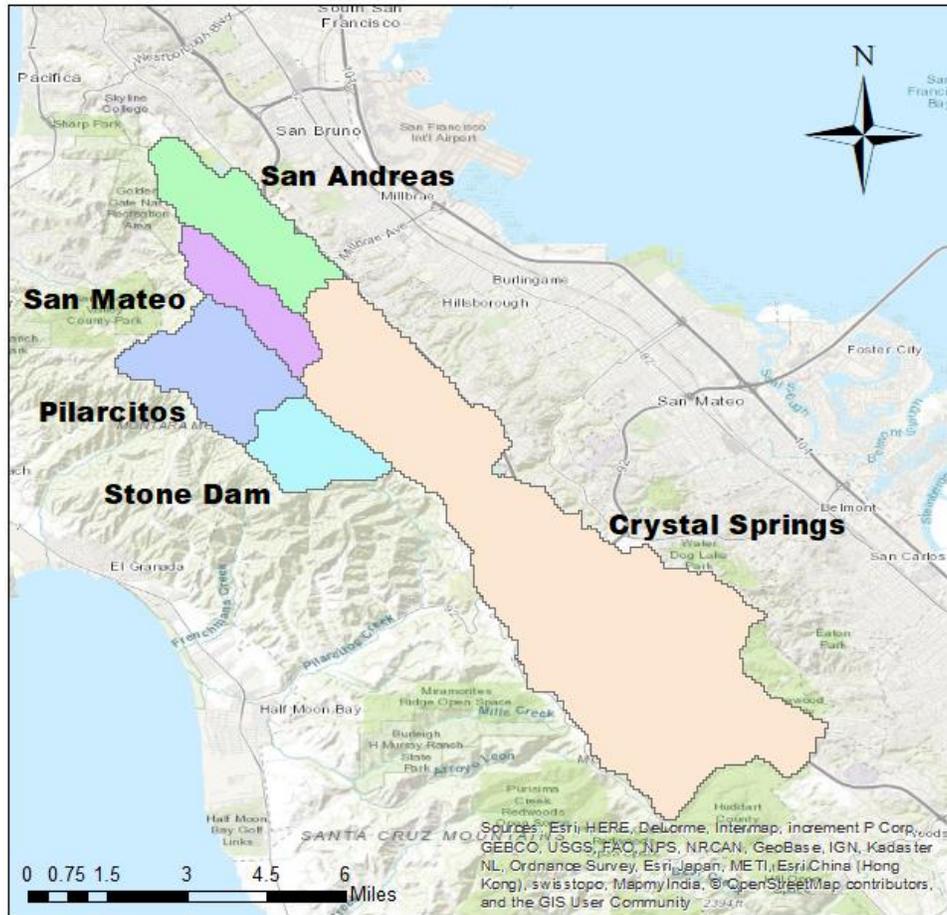
Figure 5-1 Peninsula watersheds

5.1.1. Geology and soils

The Peninsula watershed belongs to the natural region of California where there are northwest-trending faults, mountain ranges, and valleys as consequences from geologically complex and seismically active zone. (San Francisco Planning Department, 2001)

The lowest elevation in the Peninsula watershed is about 300 feet above mean sea level (msl) along the shoreline of Crystal Springs reservoir. The highest elevation is around 1,900 feet above msl at the crest of Montara Mountain (at North Peak and Scarper Peak) and over 2,000 feet above msl near Kings Mountain. Northern slopes are approximately steeper than southern slopes (San Francisco Planning Department, 2001).

Pilarcitos Fault shown in the Figure 5-2 divides the Peninsula watershed into two parts (west and east). In the west, soils are underlain by granitic rocks and sedimentary rocks. These soils drain well, and their depths vary from shallow to deep. They can be erosive dramatically. In the east, soils are underlain by the varied rocks of the Franciscan complex and the Butano Sandstone. Their characteristics are the same as soils in the west (San Francisco Planning Department, 2001).
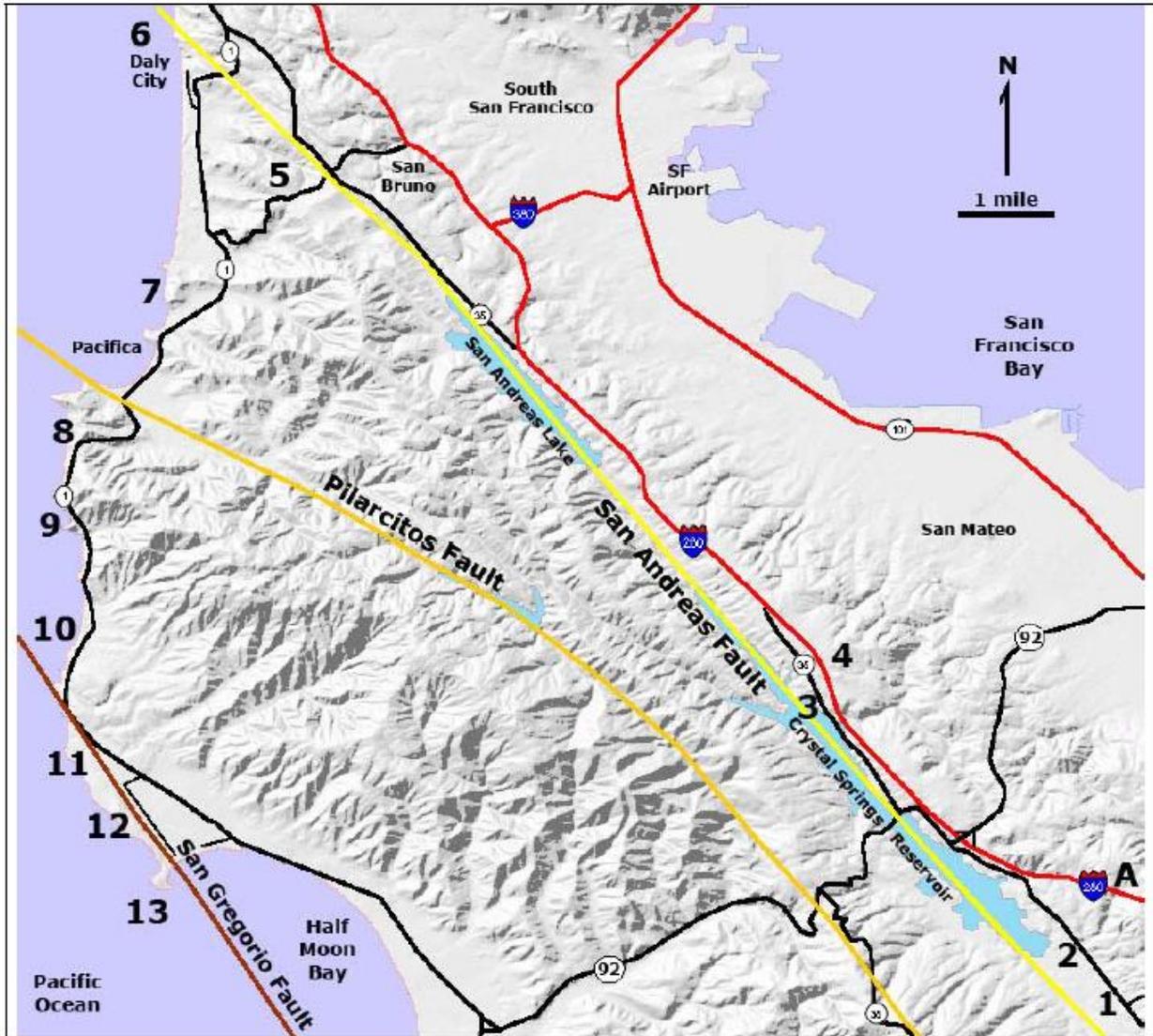
Figure 5-2 Map of the Peninsula showing major faults in the northern Santa Cruz Mountains in San Mateo County. Stops include: A) I-280 Vista Point, 1) Filoli Center, 2) Pulgas Water Temple, 3) Crystal Springs Dam, 4) I-280 Rest Area, 5) Milagro Ridge, 6) Mussel Rock Park, 7) Pacifica Quarry, 8) San Pedro Mountain and Devil's Slide, 9) Montara Mountain, 10) Montara Beach, 11) James V. Fitzgerald Marine Preserve, 12) Half Moon Bay Airport, and 13) Pillar Point and Mavericks (The United States Geological Survey, 2005).

5.1.2. Hydrology

There are three main creeks in the Peninsula watershed. They are San Mateo, Pilarcitos, and San Andreas creeks, which are natural drainages for this watershed. Nevertheless, the watershed also has artificial drainage systems from San Francisco Public Utilities Commission (SFPUC). Three major drainage areas include (1) Upper and Lower Crystal Springs reservoirs, (2) Pilarcitos reservoir and Creek, and (3) San Andreas reservoir (San Francisco Planning Department, 2001). However, in order to provide system model inflows at important discharge gauges, this study decides to delineate the Peninsula watershed into five sub-watersheds which are San Andreas reservoir watershed, San Mateo Creek watershed, Pilarcitos reservoir watershed, Stone Dam watershed, Crystal Springs reservoir watershed relating to five discharge

67

gauges with the same names (shown in Figure 5-3). Table 5-1 gives the summary of these discharge gauges.

Table 5-1 Discharge gauges in the Peninsula watershed

Full name	Short name	Used to delineate	Outlet to sub-watershed	Latitude	Longtitude
San Mateo Creek at Mud Dam 1	San Mateo Creek	YES	San Mateo Creek	37.55668778	-122.4180597
San Mateo Creek below the junction box	San Mateo Creek below the junction box	NO		37.556247	-122.418027
Filoli Main Bridge above Upper Crystal Springs	Filoli Main Bridge	NO		37.476412	-122.312184
Crystal Springs reservoir Inflow by Mass Balance	Crystal Springs reservoir	YES	Crystal Springs reservoir	37.528917	-122.362348
San Andreas reservoir Inflow by Mass Balance	San Andreas reservoir	YES	San Andreas reservoir	37.580239	-122.411845
Pilarcitos Creek above Stone Dam	Stone Dam	YES	Stone Dam	37.527173	-122.398355
Pilarcitos reservoir	Pilarcitos reservoir	YES	Pilarcitos reservoir	37.54805556	-122.4286111
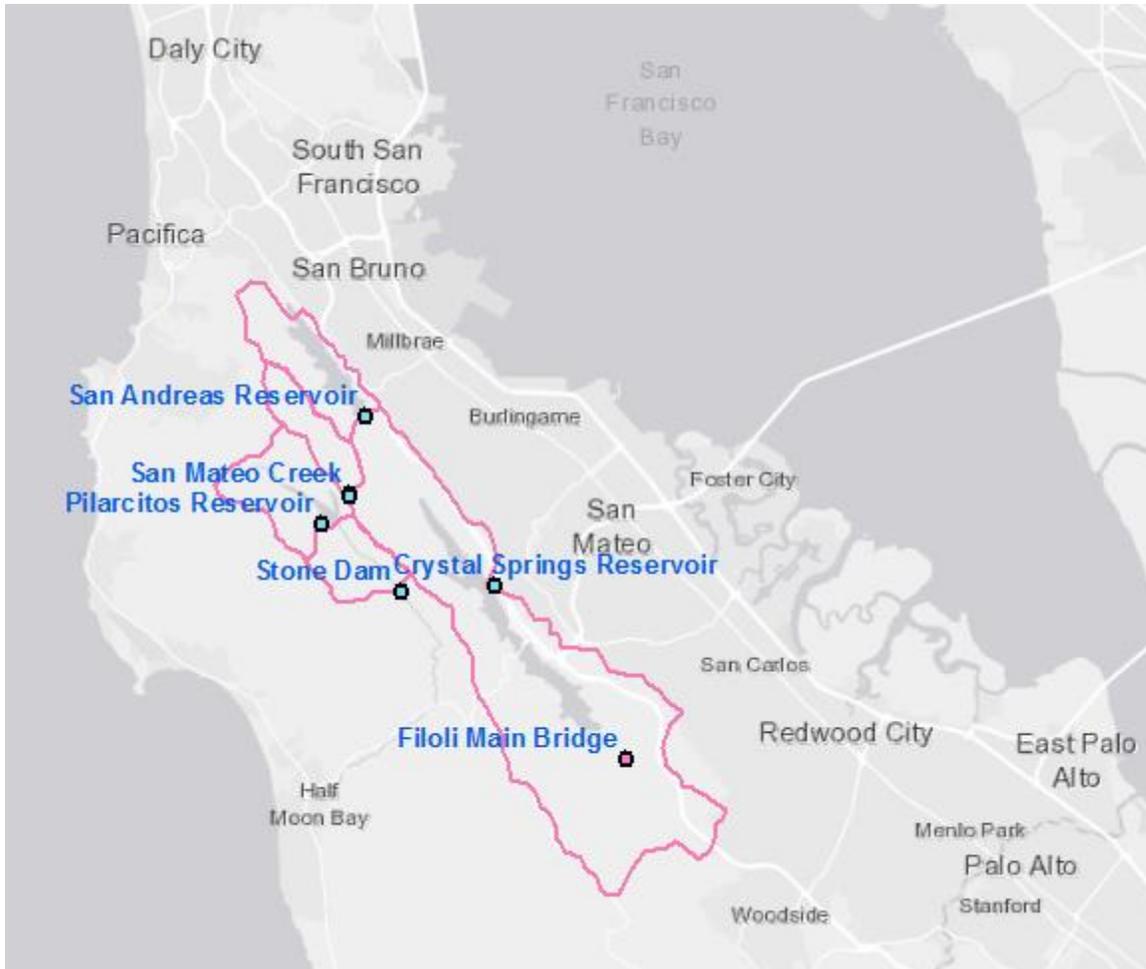
Figure 5-3 Peninsula watershed with five sub-watersheds including San Andreas reservoir watershed, San Mateo Creek watershed, Pilarcitos reservoir watershed, Stone Dam watershed, Crystal Springs reservoir watershed having the outlets at five discharge gauges with the same name representing in blue points on the map. Two discharge gauges Filoli Main Bridge and San Mateo Creek below the junction box are not used in the watershed delineation, representing in pink points on the map. The discharge gauge "San Mateo Creek below the junction box" is adjacent to the discharge gauge "San Mateo Creek" so that it does not appear on the map.

The Peninsula watershed was protected from the urbanization because of its collection and storage function. Four reservoirs (Pilarcitos reservoir, San Andreas Dam reservoir, Upper Crystal Springs reservoir, and Lower Crystal Springs reservoir) were constructed respectively in 1864, 1870, 1877, and 1890 in order to store water in the Peninsula watershed (San Francisco Planning Department, 2001). Five sub-watersheds contain reservoirs as their names; for instance, San Andreas Reservoir watershed has San Andreas Reservoir. Only San Mateo Creek watershed contains flows from San Mateo Creek, not a reservoir.

Aside from local runoff originated within the Peninsula watershed, water from Hetch Hetchy system blended with treated water from the Alameda Watershed is conveyed through Bay Division Pipelines to Crystal Springs reservoir and then flows to San Andreas reservoir. Some of the water in Pilarcitos reservoir is conveyed to the Coastside County Water District (CCWD). Surplus water from Pilarcitos

reservoir is transfered to San Andreas or Crystal Springs reservoirs. Water in San Andreas and Crystal Springs reservoirs is treated at the Harry W. Tracy Water Treatment Plant (Tracy WTP) before distributing to water users (San Francisco Planning Department, 2001). Figure 5-4 represents the water system in the Peninsula watershed.

Stream flows of the Peninsula watershed divert to the oceans by two directions. First, from the upstream of San Andreas reservoir and the upstream of San Mateo Creek, streams flow to the South and direct into San Mateo Creek in the East towards San Francisco Bay. In addition, streams flow from the South of Crystal Springs reservoir towards the North of Crystal Springs reservoir and divert into San Mateo Creek at the same location as stream flows from the upstream of San Andreas reservoir and the upstream of San Mateo Creek. Second, stream flows from the upstream of Pilarcitos reservoir direct into Pilarcitos Creek towards Pacific Ocean (shown in Figure 5-6 and Figure 5-6).
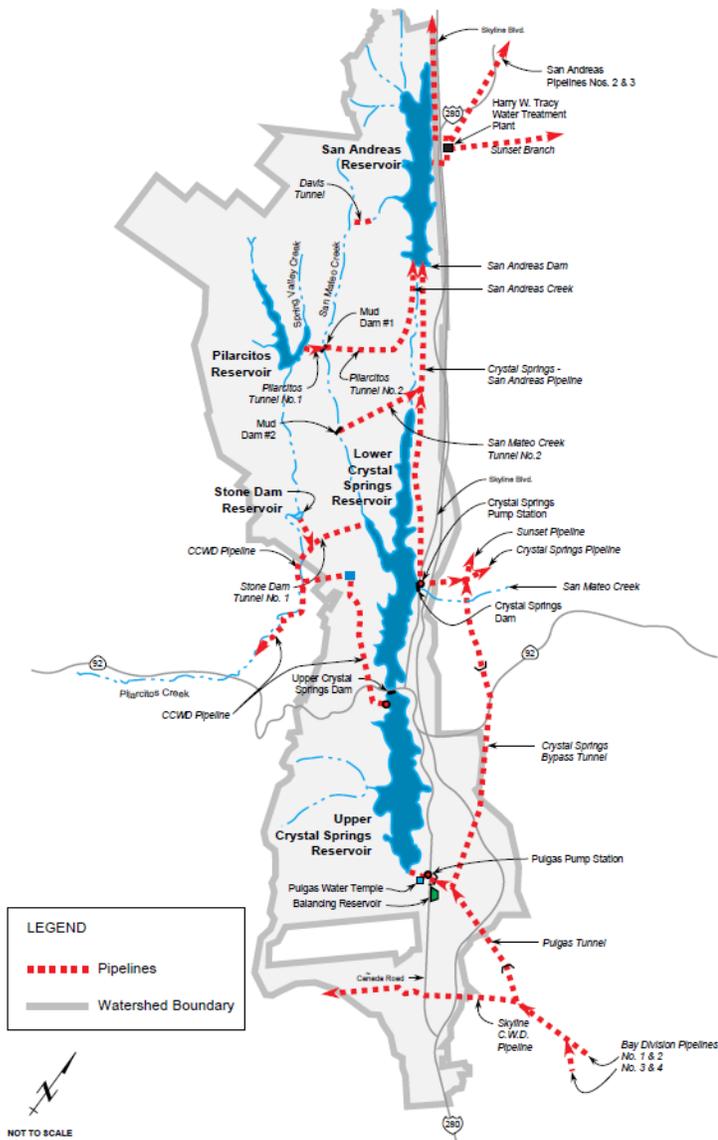
Figure 5-4 Flows in the Peninsula watershed (San Francisco Planning Department, 2001).

70
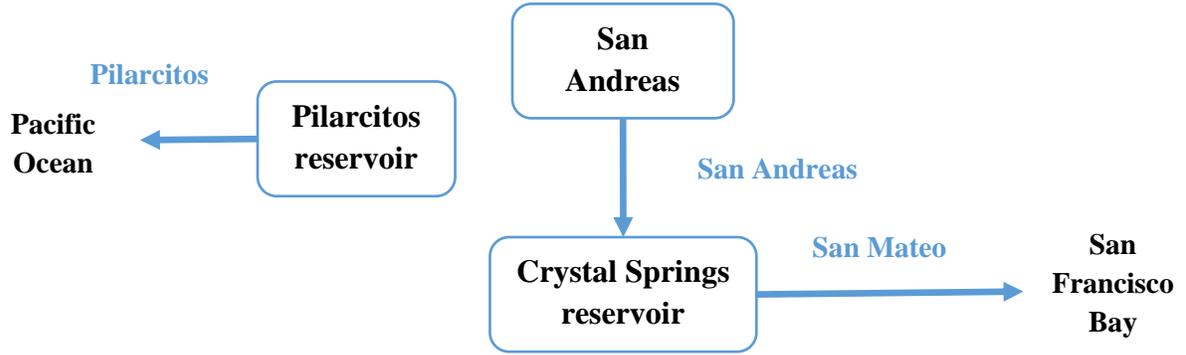
Figure 5-5 The scheme of main unregulated flows in the Peninsula watershed
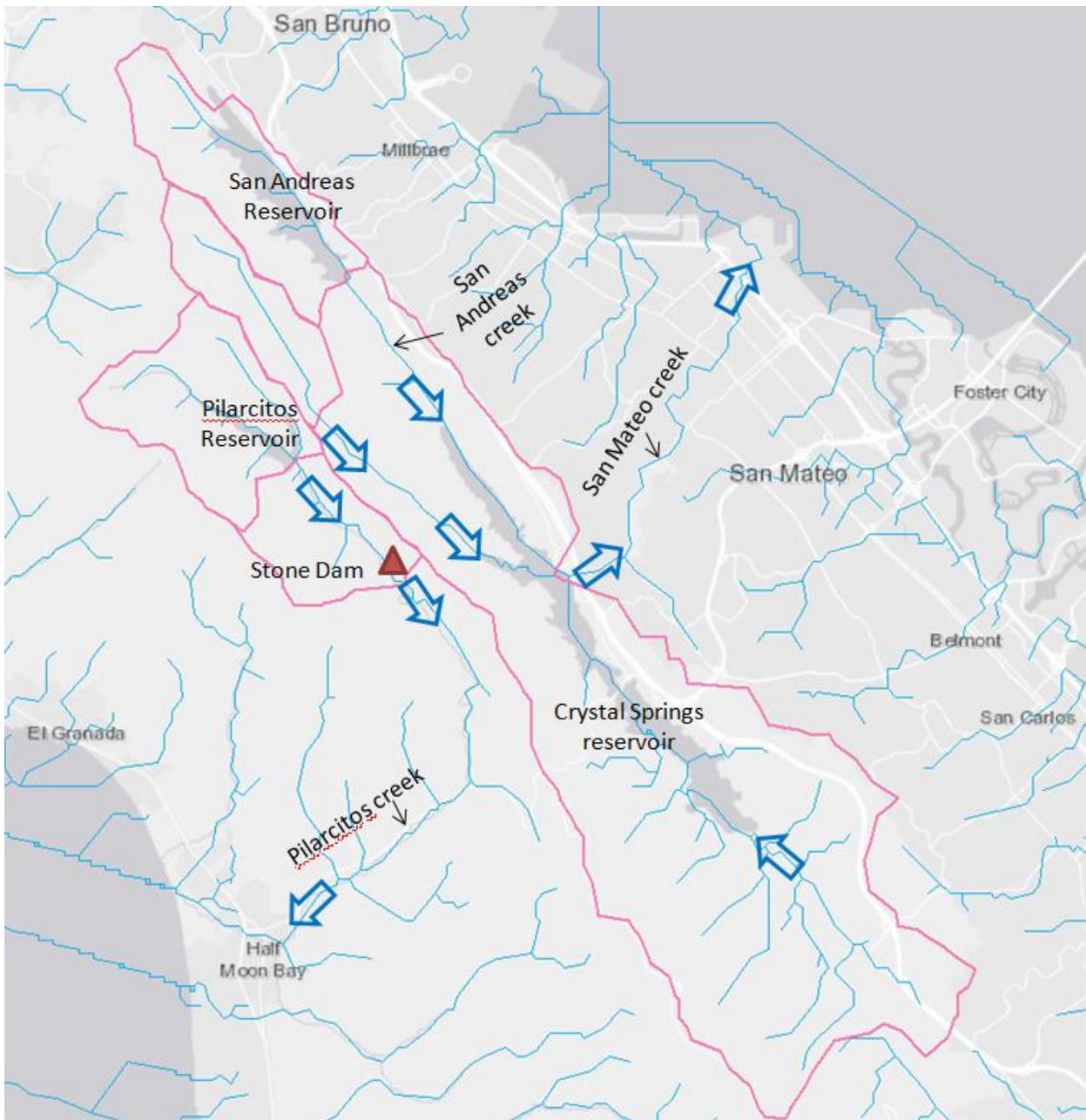
Figure 5-6 Unregulated flows in the Peninsula watershed

Figure 5-7 Regulated and unregulated flows in the Peninsula watershed (San Francisco Planning Department, 2005)

San Andreas and Crystal Springs reservoirs are final reservoirs in SFPUC water system so that they do not only store local runoff from the Peninsula watershed, but they also store water from the Tuolumne River, the Alameda creek watershed, and Pilarcitos creek. Thus, their storage capacity is bigger than the capacity relating to the water storage purpose for local runoff in the Peninsula watershed. San Andreas and Crystal Springs reservoir can reduce flows from San Andreas and San Mateo creeks immediately below the dams which can prevent flooding, except for occasional spills or releases. Despite flood control is not the original purpose for constructing Crystal Springs reservoir, it can diminish peak flows most of time. When flood storage capacity is filled, it can release uncontrolled flows over the spillway at Lower

73

Crystal Springs Dam or controlled releases at the outlets. Before releasing water from this reservoir, SFPUC always regards downstream effects. (San Francisco Planning Department, 2005)

In term of drought control, there are no required releases from Crystal Springs, San Andreas, and Pilarcitos reservoirs to maintain minimum stream flows in San Mateo, San Andreas, and Pilarcitos Creeks. (San Francisco Planning Department, 2005)

5.2. Methodology

The reservoir design is based on the natural stream flows in the past 50 years. Any amount of future discharges outside of the stream flows variations in the past 50 years can cause adverse hydrologic effects on SFPUC's water system such as flooding, dewatering, erosion, or drainage alternation. (San Francisco Planning Department, 2005). Thus, this study aims to create a hydrologic model which can represent naturalized stream flows in the Peninsula watershed. Then, it will use different weather scenarios from the weather generators to create a range of future naturalized stream flows at the outlets of five sub-watersheds. Likewise, a climate stress test will determine a range of stream flow alteration under precipitation and temperature changes.

The Sacramento soil moisture accounting model (SAC-SMA) is used to build Peninsula hydrologic model. SAC-SMA is described in depth on section 4.2. For Peninsula Hydrology, SAC-SMA was built with a lumped configuration, which removes the Lohmann routing module. In addition, Genetic Algorithm (GA) is used to calibrate the hydrologic model, which is described in depth in Appendix E.

5.3. Input data

5.3.1. Digital elevation model (DEM)

SRTM 90m digital elevation data produced by The National Aeronautics and Space Administration (NASA) are used in this project. SRTM is the short name of The Shuttle Radar Topography Mission, which was flown aboard the space shuttle *Endeavour* February 11-22, 2000. During Endeavour's 11-day mission, SRTM collected radar data over 80% of the Earth's land surface between 60° north and 56° south latitude. There are two spatial resolution: 1 arc-second for global coverage (~30 meters) and 3 arc-seconds for global coverage (~90 meters). In this study, 90-meter resolution is used for delineating the watersheds and calculating average ground elevation to msl.

5.3.2. Climate data

Temperature and precipitation datasets are used as input climatology to the SAC-SMA hydrology models. As the SAC-SMA hydrology models for the Peninsula regions are lumped models, each watershed uses a single input time-series for each climate dataset.

5.3.2.1. Temperature

Temperature datasets are gathered from gridded dataset from (Livneh et al., 2015) over a period of 1950 to 2011. Figure 5-8 shows the Livneh temperature grid cells overlaid on top of the watersheds with various ground temperature stations used to validate the dataset. A lapse rate is introduced to adjust the gridded dataset elevation temperature to ground elevation using the rate of 6.2℃/km. Figure 5-9 shows the elevation-adjusted Livneh gridded data time-series with Spring Valley and Pulgas temperature stations.
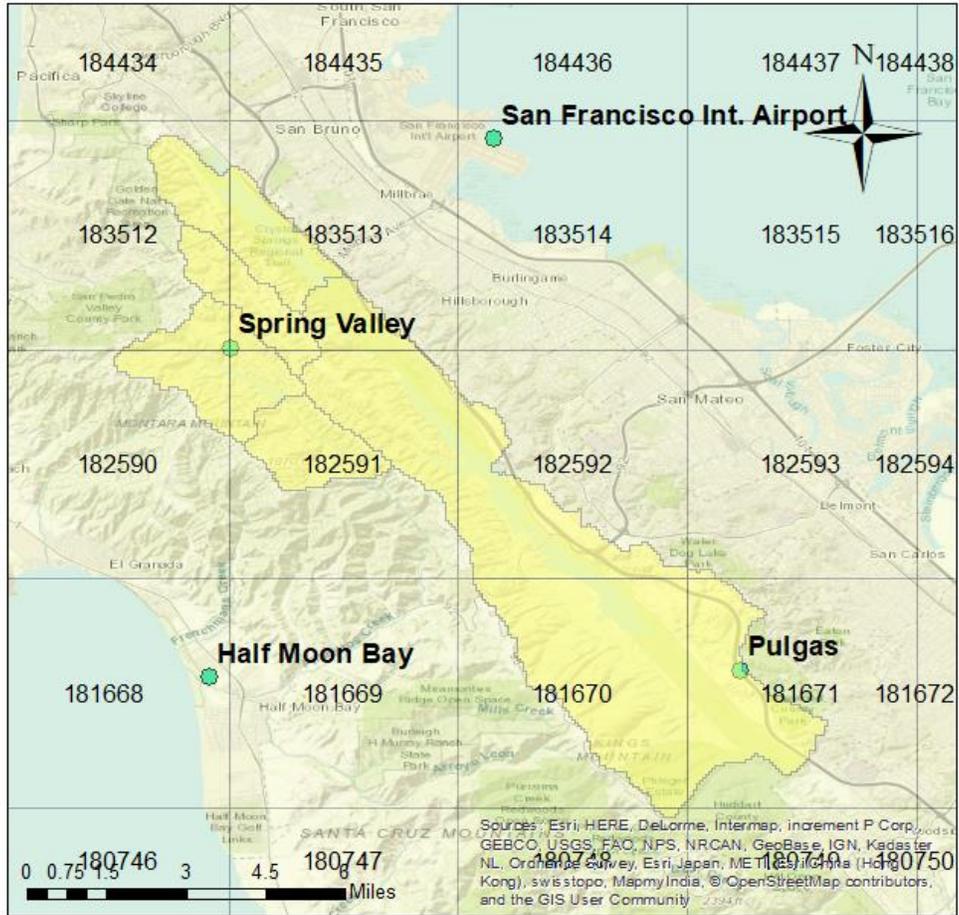
Figure 5-8. Temperature ground station with Livneh, et. al 2015 grid cells with LOCA ID numbers overlaid on top of the Peninsula watersheds.

(a)

(b)
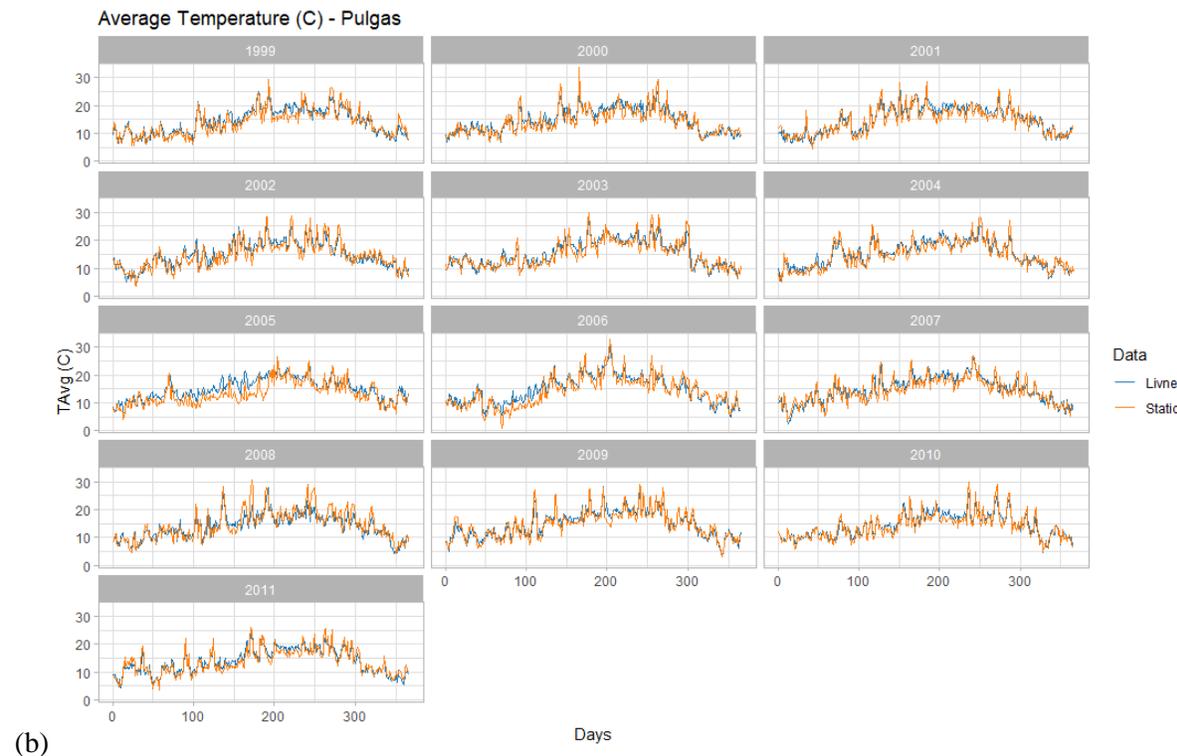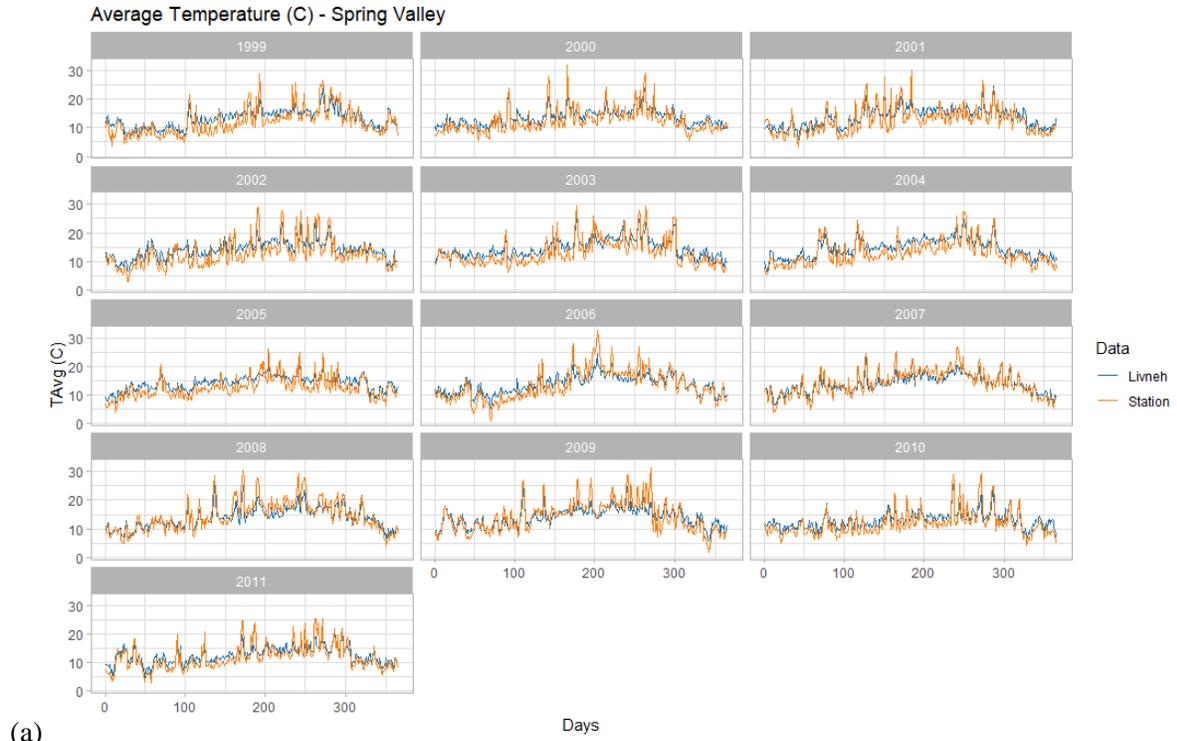
Figure 5-9. Temperature data comparison of elevation adjusted Livneh, et. al. 2015 respective ground temperature stations (a) Spring Valley and (b) Pulgas.

5.3.2.2.   Precipitation

The precipitation dataset is gathered from three ground stations inside the study area. A Thiessen polygon is generated for the watersheds to create an area-weighted average precipitation time-series to drive the hydrology models. Table 5-2 outlines the duration and locations of the precipitation stations used with Figure 5-10 visually shows the coverage areas for each watershed.

Table 5-2. Precipitation station latitude and longitude

Station Name	Latitude	Longitude	Start Date	End Date
Pilarcitos	37.55305556	-122.4286111	1956-10-01	2011-09-30
San Andreas Reservoir	37.57916667	-122.4088889	1956-10-01	2011-09-30
Upper Crystal Springs	37.51222222	-122.3541667	1956-10-01	2011-09-30
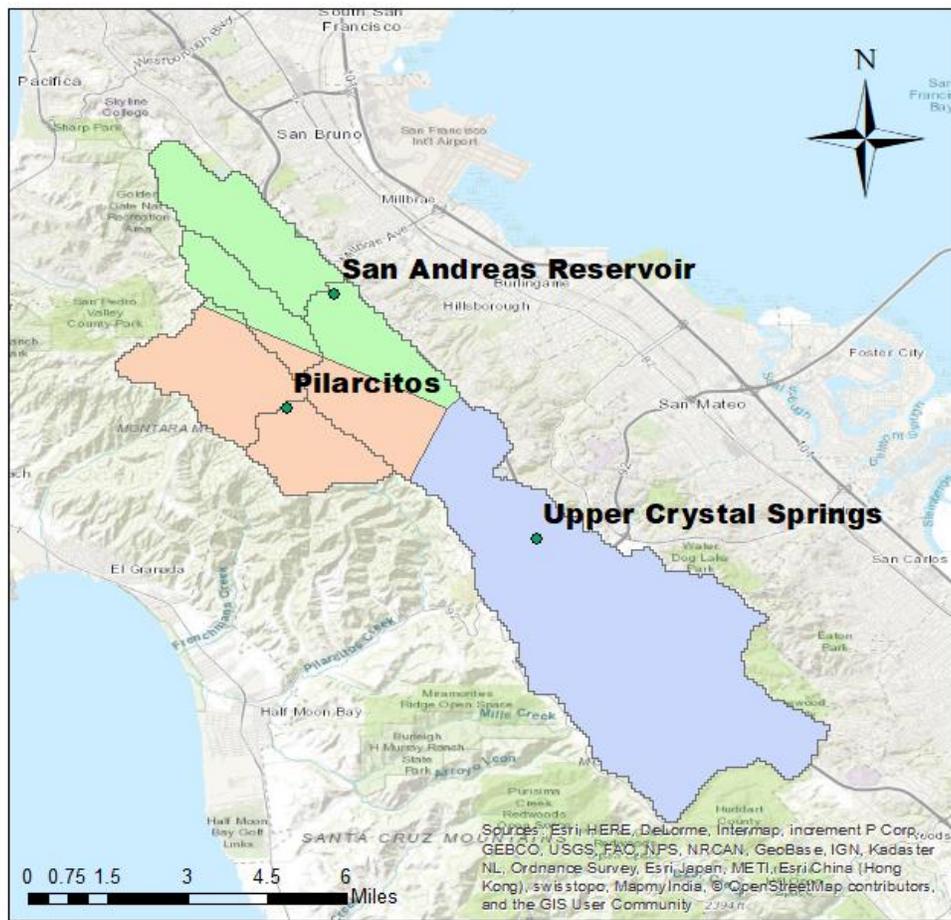
Figure 5-10. Precipitation stations with Theissen polygon over Peninsula watersheds

5.3.3. Stream gauges

Locations of stream gages are illustrated in Figure 5-3 and are from gauges or back-calculated naturalized flows. The streamflow datasets were received from SFPUC and the locations are listed in Table 5-3. The watershed areas with streamflow data periods are listed in Table 5-4.

Table 5-3. Stream gages

Sub-watershed	Stream gage	Latitude	Longtitude
Pilarcitos Reservoir	Pilarcitos Creek at Pilarcitos Dam	37.54805556	122.4286111
Stone Dam	Pilarcitos Creek above Stone Dam Accretion	37.527173	-122.398355
San Andreas Reservoir	San Andreas Reservoir Inflow by Mass Balance	37.580239	-122.411845
Crystal Springs Reservoir	Crystal Springs Reservoir Inflow by Mass Balance	37.528917	-122.362348
San Mateo Creek	San Mateo Creek at Mud Dam 1	37.55668778	-122.4180597

Table 5-4. Watershed area and streamflow data range

Watershed	Start Date	End Date	Area (km^2)
Crystal Springs	2007-10-01	2017-09-22	62.119596
Stone Dam	2011-11-22	2017-07-31	6.102815
Pilarcitos	1999-10-01	2018-06-19	10.064031
San Mateo	2011-02-28	2017-06-23	5.514076
San Andreas	1979-01-22	2017-09-22	9.636758

Limitations in gridded climate time-series limited the period of streamflow available for calibration. Specifically, these were concerns for San Mateo and Stone Dam watersheds, as the streamflow time-series did not overlap enough for calibration purposes. Therefore, the correlation was calculated for the San Mateo watershed streamflow over 2011-10-01 through 2016-09-30 and Stone Dam watershed over 2012-10-01 through 2016-09-30. Table 5-5 shows the calculated correlation of the various watersheds and shows the highest correlation streamflows were Pilarcitos. Using the Pilarcitos streamflow, a basin-area ratio calculation was applied generate streamflow from 1999-10-01 through 2011-09-30 for each respective region.

Table 5-5. Correlation calculated for (a)San Mateo watershed over 2011-10-01 to 2016-09-30 and (b)Stone Dam watershed over 2012-10-01 through 2016-09-30.

(a) San Mateo Watershed	San Mateo	Pilarcitos	San Andreas	Crystal Springs
San Mateo	1.000	**0.927**	0.698	0.757
Pilarcitos	**0.927**	1.000	0.692	0.779
San Andreas	0.698	0.692	1.000	0.665
Crystal Springs	0.757	0.779	0.665	1.000
(b) Stone Dam Watershed	**Stone Dam**	**Pilarcitos**	**San Andreas**	**Crystal Springs**
Stone Dam	1.000	**0.841**	0.716	0.714
Pilarcitos	**0.841**	1.000	0.783	0.785
San Andreas	0.716	0.783	1.000	0.688
Crystal Springs	0.714	0.785	0.688	1.000

5.4. Model Calibration

Figure 5-11 shows the process of hydrologic modelling. There are four stages. The first stage is the sub-watershed delineation from DEM and locations of stream gauges. The second stage is the climate data preparation which combines gridded data and gauged data to create datasets of precipitation and temperature. These stages are conducted simultaneously. Then, precipitation and temperature are input data for hydrologic model in the simulation. Finally, the hydrologic model is calibrated for daily KGE using genetic algorithm.
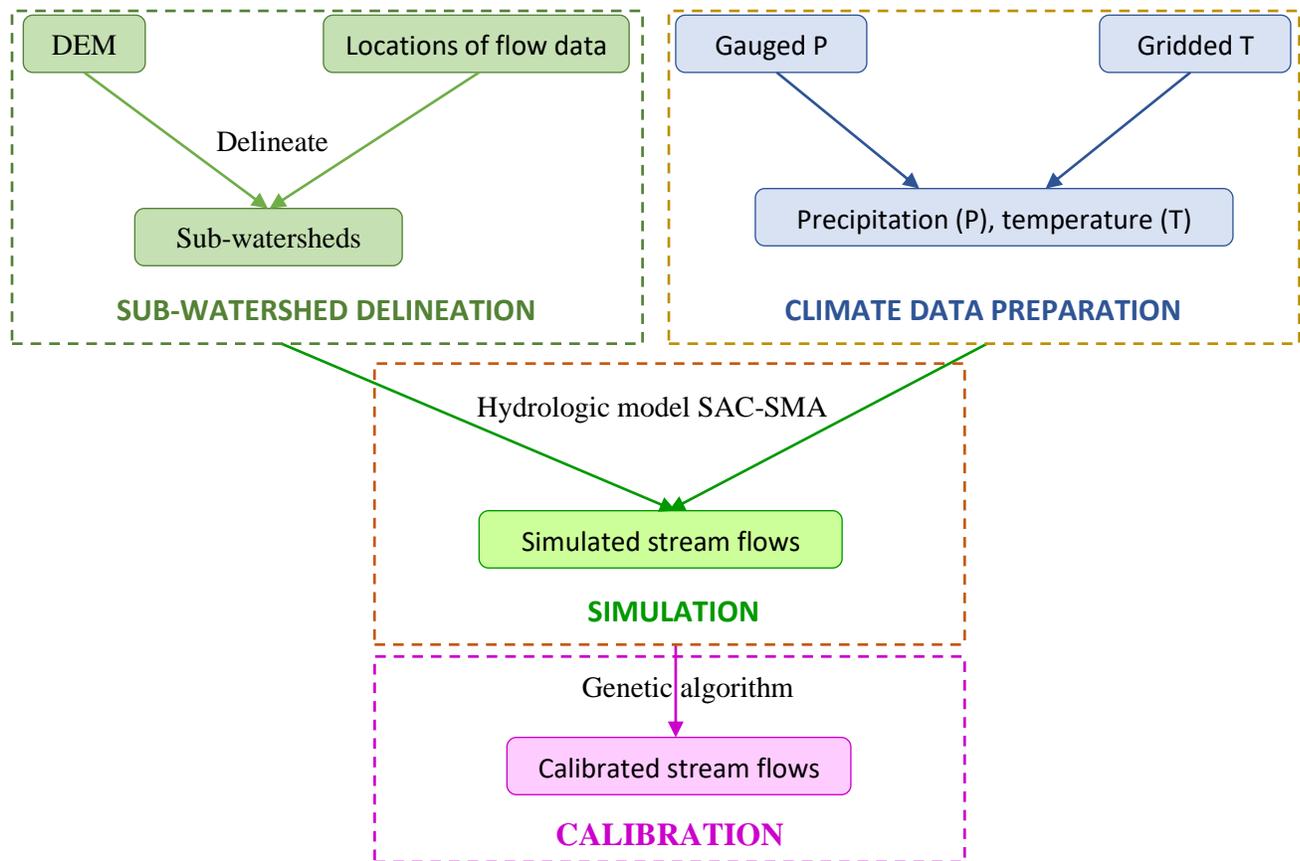
Figure 5-11 Work flow of hydrologic modelling

Each watershed is calibrated based on the configurations defined in Table 5-6 and validated over 2010-10-01 through 2011-09-30.

Table 5-6. Watershed calibration and validation periods with additional comments

Watershed	Start Date	End Date	Validation Period	Note
San Mateo	1999-10-01	2010-09-30		Basin-area ratio adjusted Pilarcitos streamflow as calibration target
San Andreas	1999-10-01	2010-09-30		
Crystal Spring	2007-10-01	2010-09-30	2010-10-01 – 2011-09-30	2006-10-01 to 2007-09-30 used as warmup period
Stone Dam	1999-10-01	2010-09-30		Basin-area ratio adjusted Pilarcitos streamflow as calibration target
Pilarcitos	1999-10-01	2010-09-30		

5.5. Calibration Results

Table 5-7 shows performance metrics for calibration and validation of all five sub-watersheds in Peninsula watershed. All but San Andreas sub-watersheds have good daily KGE calibration values (above 0.8), but San Andreas KGE is very close to 0.8. While maximizing daily KGE is the target for calibration, other performance metrics were also compared. Daily NSE is seen to be within acceptable levels (above 0.7) for all sub-watersheds except for San Andreas, but with significantly higher monthly KGE and monthly NSE scores and fairly small percent bias values during the calibration period. Validation metrics shows similar results to calibration metrics, except for the increase in percent bias for Crystal Spring and decrease for San Mateo. These will be discussed further below.

Table 5-7. Calibration and validation metrics for Peninsula watersheds

Calibration					
Watershed	**Daily KGE**	**Daily NSE**	**PBias (%)**	**Monthly KGE**	**Monthly NSE**
San Mateo	0.873	0.775	-3.0	0.912	0.870
San Andreas	0.794	0.594	1.0	0.630	0.760
Crystal Spring	0.872	0.743	0.4	0.906	0.912
Stone Dam	0.860	0.722	0.1	0.895	0.832
Pilarcitos	0.871	0.770	-2.9	0.925	0.877
Validation					
Watershed	**Daily KGE**	**Daily NSE**	**PBias (%)**	**Monthly KGE**	**Monthly NSE**
San Mateo	0.810	0.750	-14.7	0.808	0.871
San Andreas	0.675	0.328	-4.9	0.565	0.665
Crystal Spring	0.740	0.737	22.5	0.774	0.937
Stone Dam	0.873	0.766	0.2	0.798	0.864
Pilarcitos	0.838	0.766	1.1	0.748	0.825
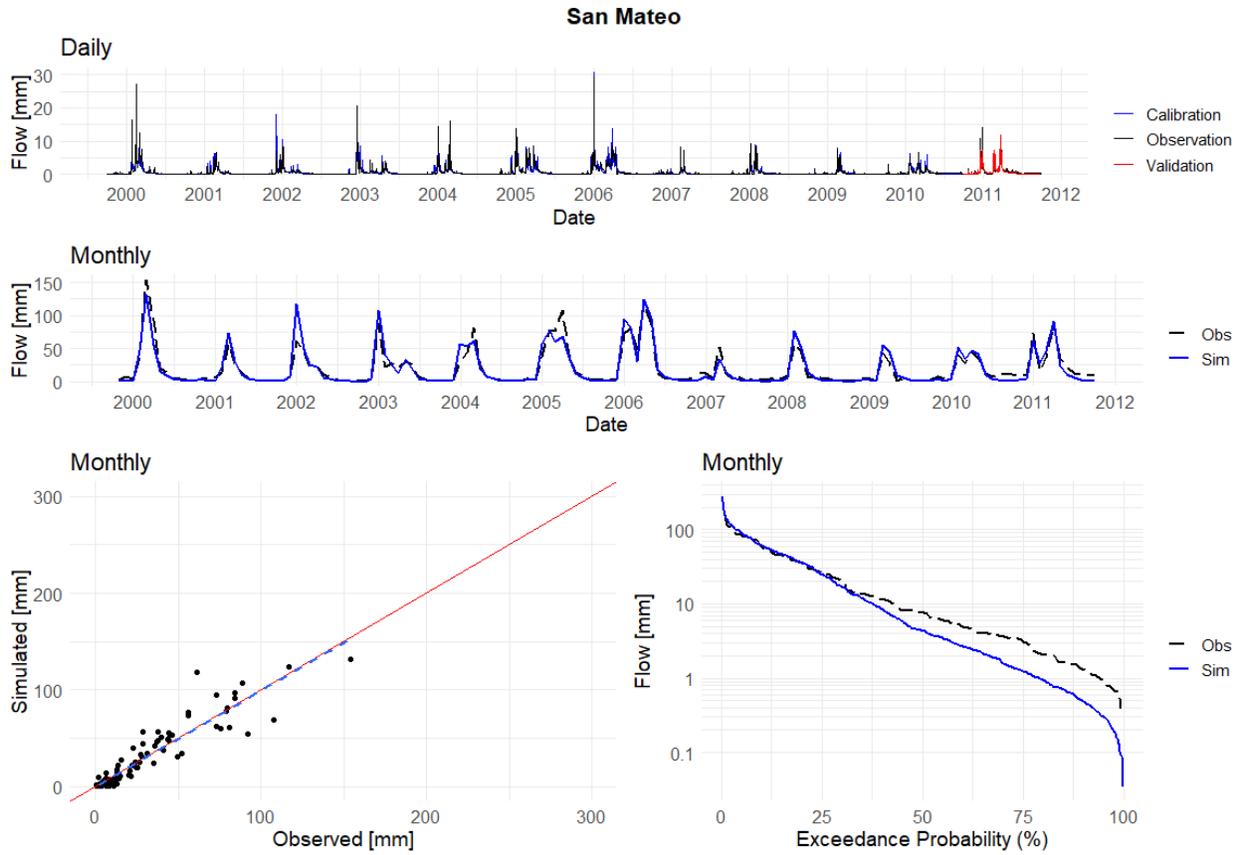
Figure 5-12. San Mateo sub-watershed daily and monthly streamflow hydrographs, monthly scatterplot, and monthly flow duration curve.

San Mateo calibration results are shown in Figure 5-12 with daily and monthly streamflow hydrographs, monthly scatterplot, and monthly flow duration curves. The streamflow is measured in the plot as depth in millimeters. A conversion can be made from millimeters to cubic feet per second by multiplying the watershed area. The plots all show a good model fit with the flow duration curve showing an underestimation of low flows compared to historic observed values. However, this estimate is assumed to be acceptable as the historic observed value is the basin-area ratio of Pilarcitos streamflow. The validation period from 2010-10-01 through 2011-09-30 shows a percent bias of -14.7%, which is due to the underestimation of the flows over those periods, however the monthly hydrograph and scatterplot shows minimal issues during the validation periods.
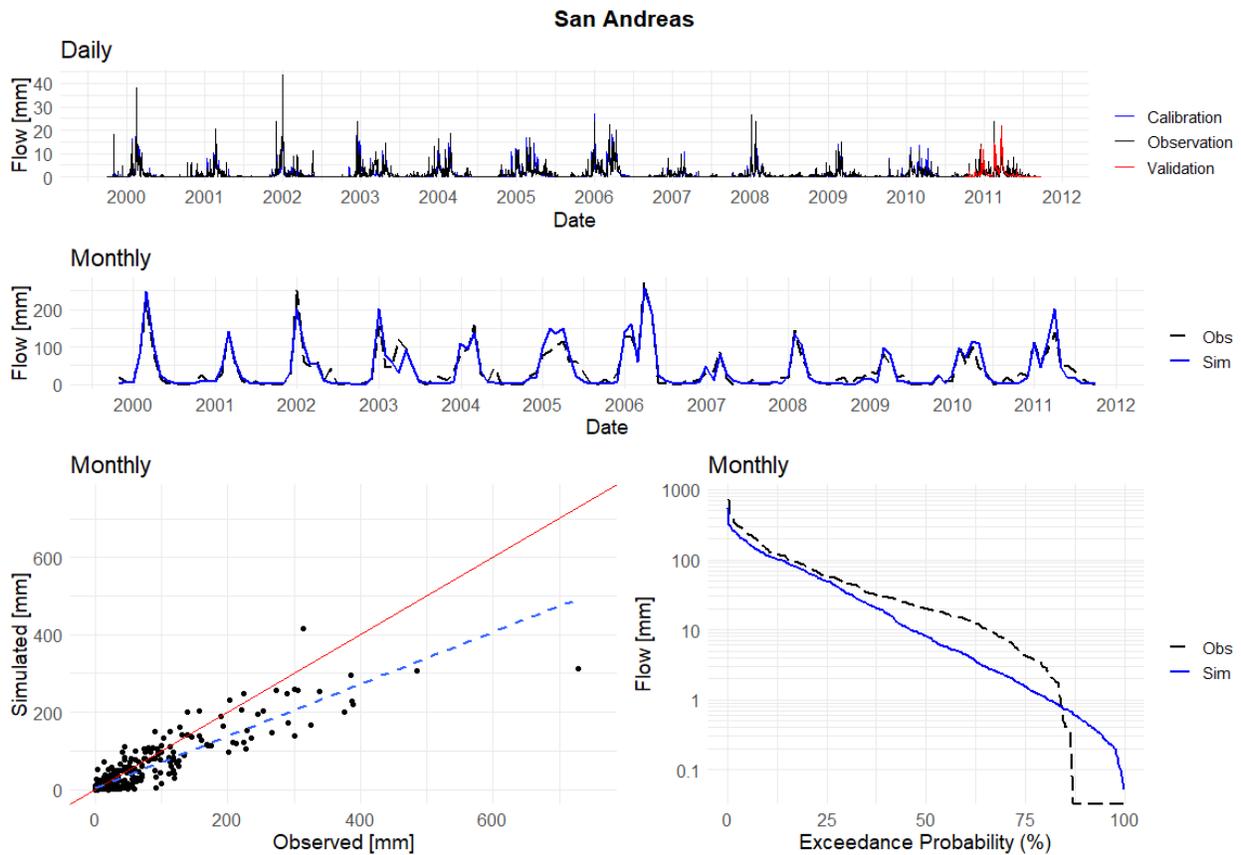
Figure 5-13. San Andreas sub-watershed daily and monthly streamflow hydrographs, monthly streamflow scatterplot, and flow duration curve.

San Andreas calibrations, and all the other hydrologic model calibrations, were performed similarly to San Mateo model calibrations. San Andreas is shown to have a more difficult model fit due to observed zero flow events. This is shown in Figure 5-13 where the monthly streamflow scatterplot shows an underestimation of flows. In addition, the monthly flow duration curve shows the creek as dry for a few months while the hydrology model has no dry flow months.
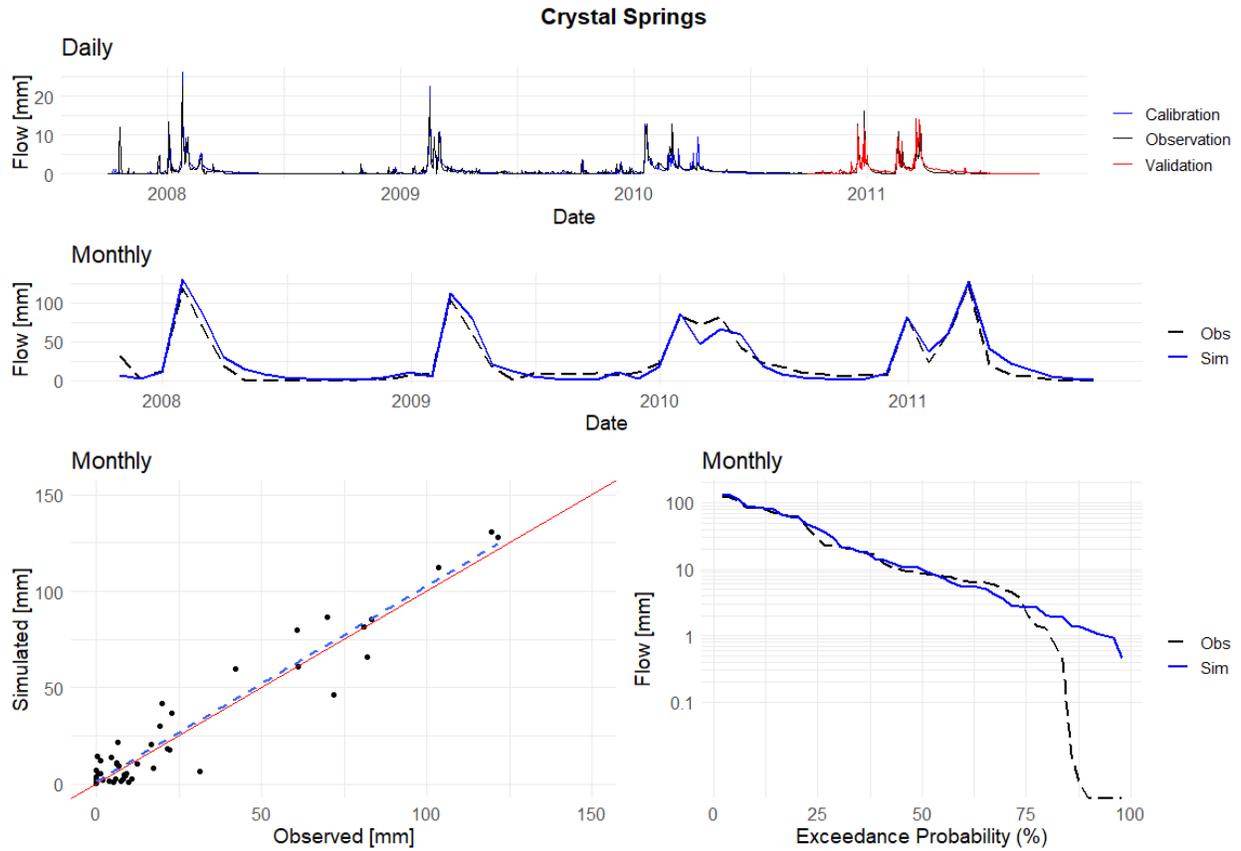
Figure 5-14. Crystal Springs sub-watershed daily and monthly streamflow hydrographs, monthly streamflow scatterplot, and flow duration curve.

Crystal Springs calibration was severely limited due to the limited streamflow dataset available. To work with the limited data available, the hydrology model warmup was performed repeating the first year's worth of data. However, even with limited data availability, the hydrology model is shown to have a great model fit based on Figure 5-14. The daily and monthly streamflow hydrographs are shown to have a good model fit as well as the monthly streamflow hydrograph. The monthly flow duration curve plot shows a good fit for high to low flows but continues to have minimal flow during no-flow periods. Another point to address is the 22.5% bias increase during the validation period of 2010-10-01 through 2011-09-30. Figure 5-15 shows a daily hydrograph over the validation period, so the increase in percent bias can be explained in the large amount of overestimation during low flow event.
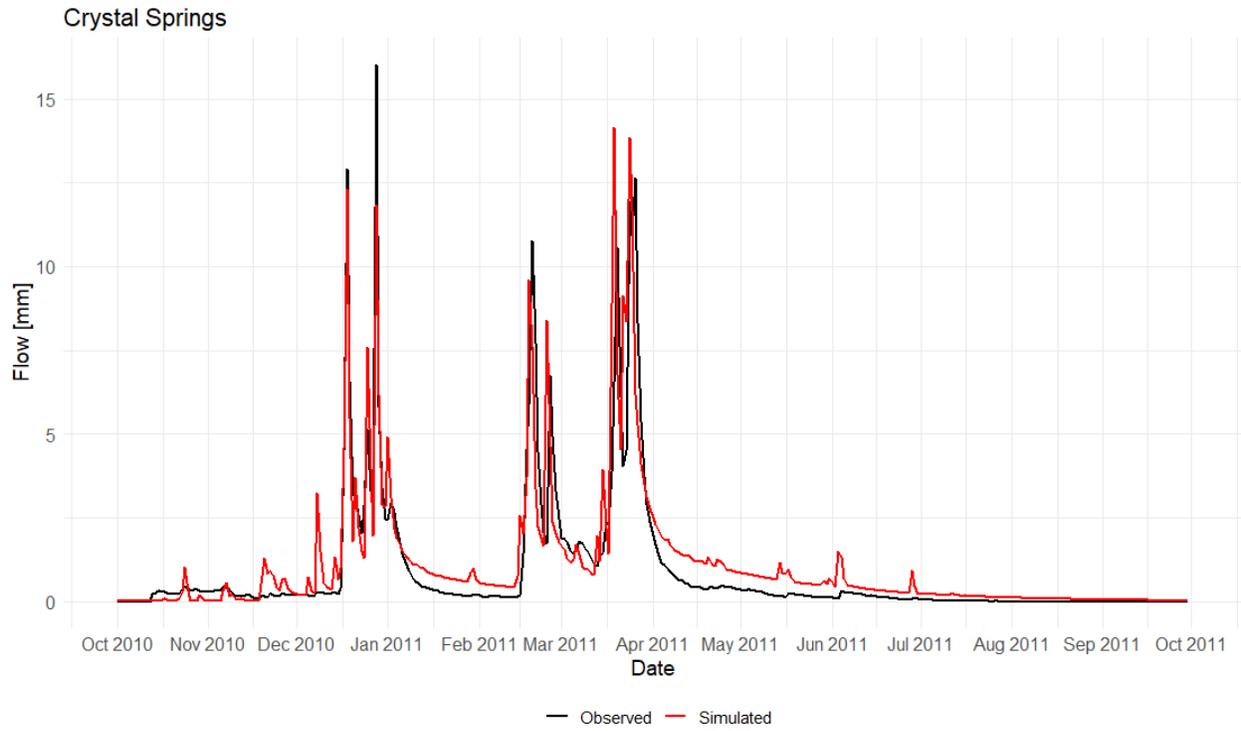
Figure 5-15. Daily streamflow hydrograph over a period of 2010-10-01 through 2011-09-30 for Crystal Springs
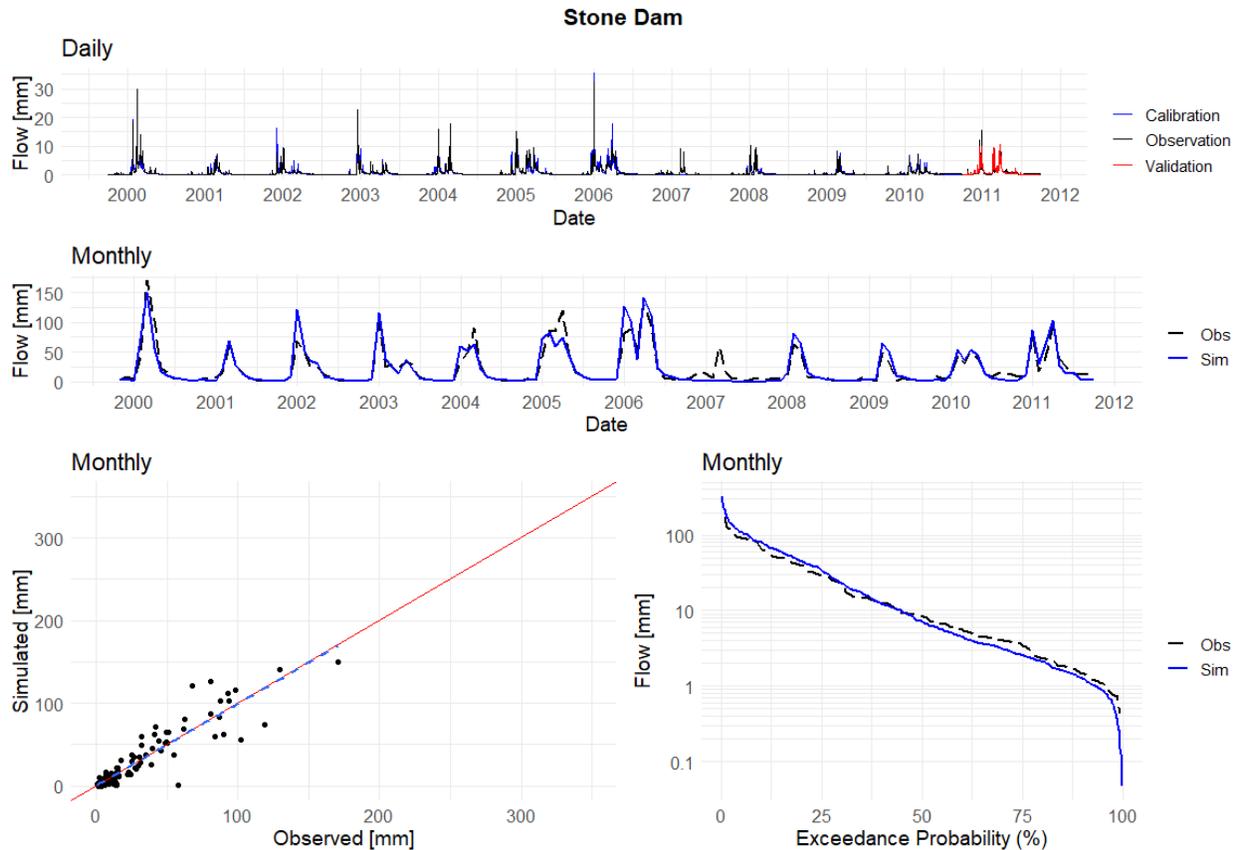
Figure 5-16. Stone Dam sub-watershed daily and monthly streamflow hydrographs, monthly streamflow scatterplot, and flow duration curve.

Stone Dam sub-watershed calibration is observed to be a good fit for the hydrology data. Figure 5-16 shows acceptable daily and monthly streamflow hydrographs, monthly scatterplot, and flow duration curve. The Stone Dam model has high daily and monthly KGE values with 0.860 and 0.895 respectively. The 2007 streamflow period is from the basin area-ratio streamflow from Pilarcitos watershed where the Pilarcitos streamflow shows flow during 2007 but there is little precipitation as input to the hydrology model as shown in Figure 5-17.
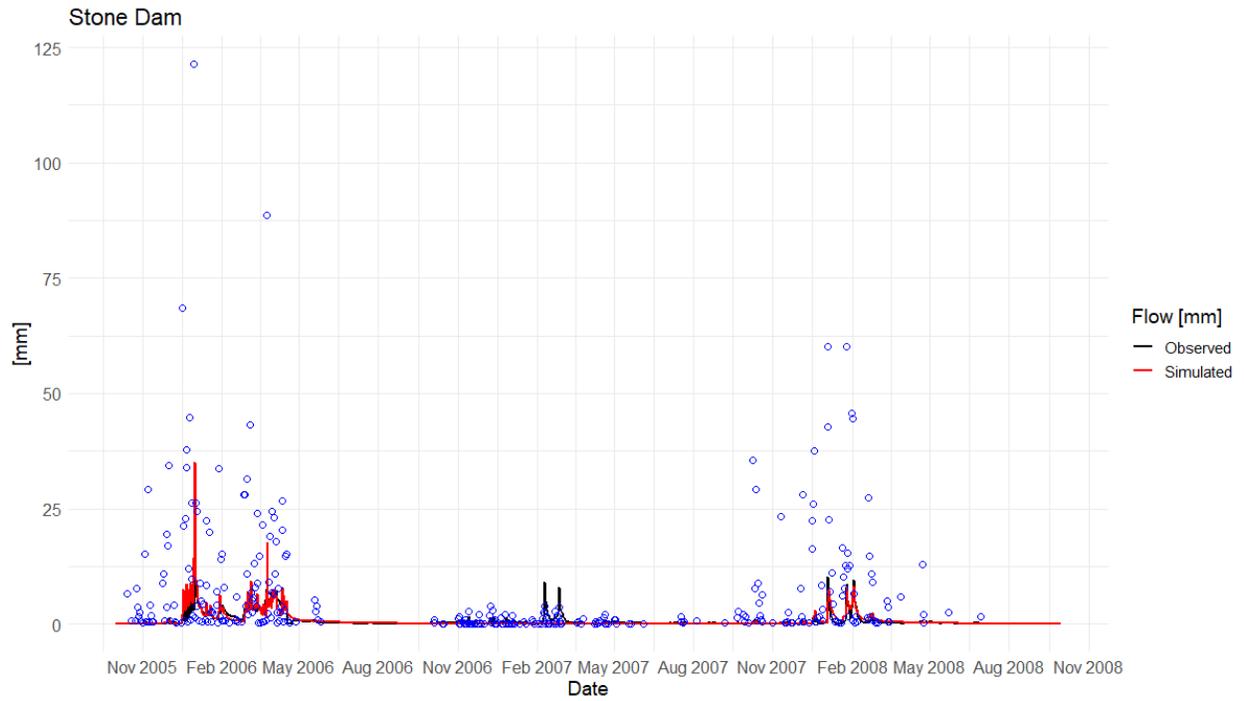
Figure 5-17. Observed and simulated streamflow (black and red respectively) with precipitation (blue) from 2005-10-01 through 2008-09-30.
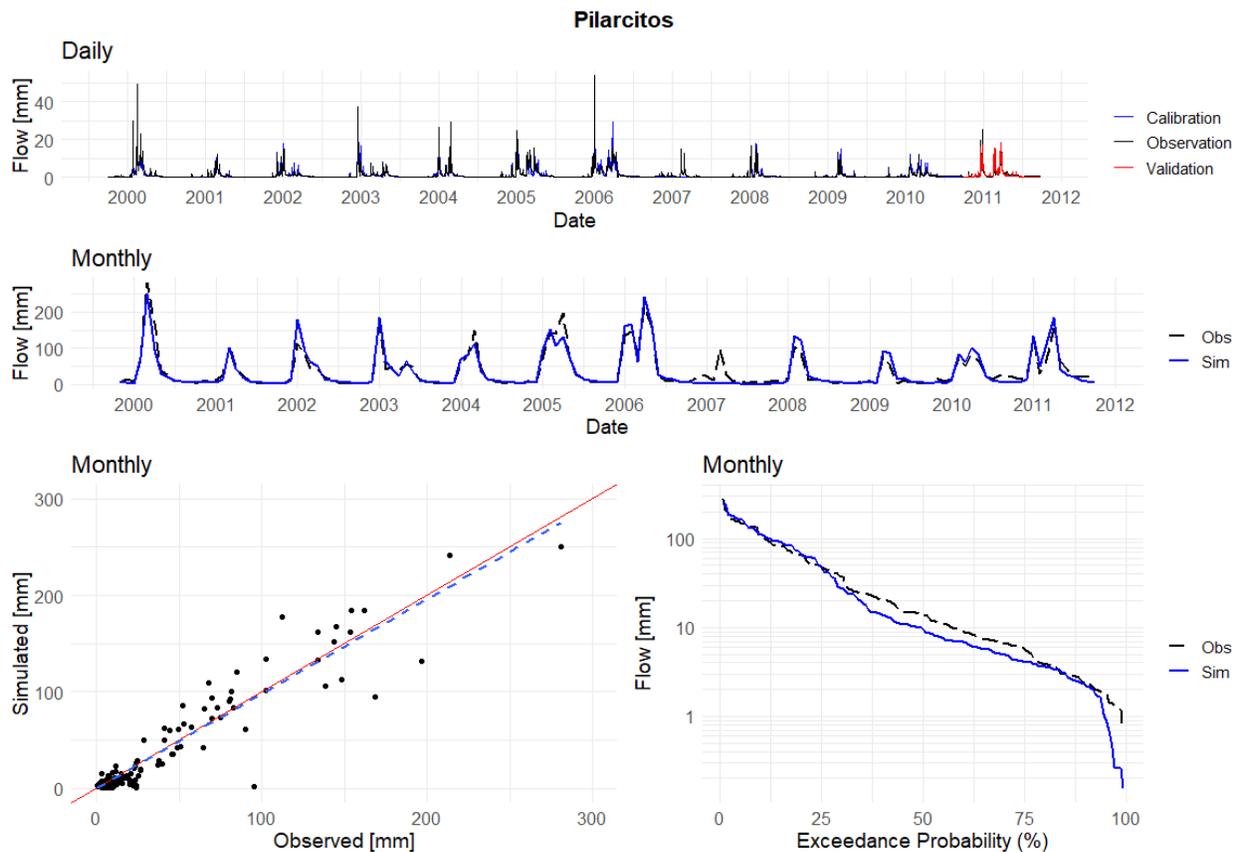
Figure 5-18. Pilarcitos sub-watershed daily and monthly streamflow hydrographs, monthly streamflow scatterplot, and flow duration curve.

Pilarcitos calibrations were performed with daily observed streamflow and precipitation driven from the Pilarcitos climate station. The calibration resulted in a daily KGE and NSE values of 0.871 and 0.770 respectively, with monthly KGE and NSE values of 0.925 and 0.877 respectively. Calibration of Pilarcitos is shown to have an underestimation of 2.9% as also shown by the fitted line in the monthly scatter plot. The flow duration curve also shows an underestimation of streamflow. The 2007 year event is not captured by the hydrology model due to limited precipitation data during that period, similar to Figure 5-17. However, the model still performed well during the validation period with daily KGE and NSE of 0.838 and 0.766 respectively and monthly KGE and NSE of 0.748 and 0.825.

5.6. SAC-SMA-L simulations forced by the stochastic weather generator climate stress test
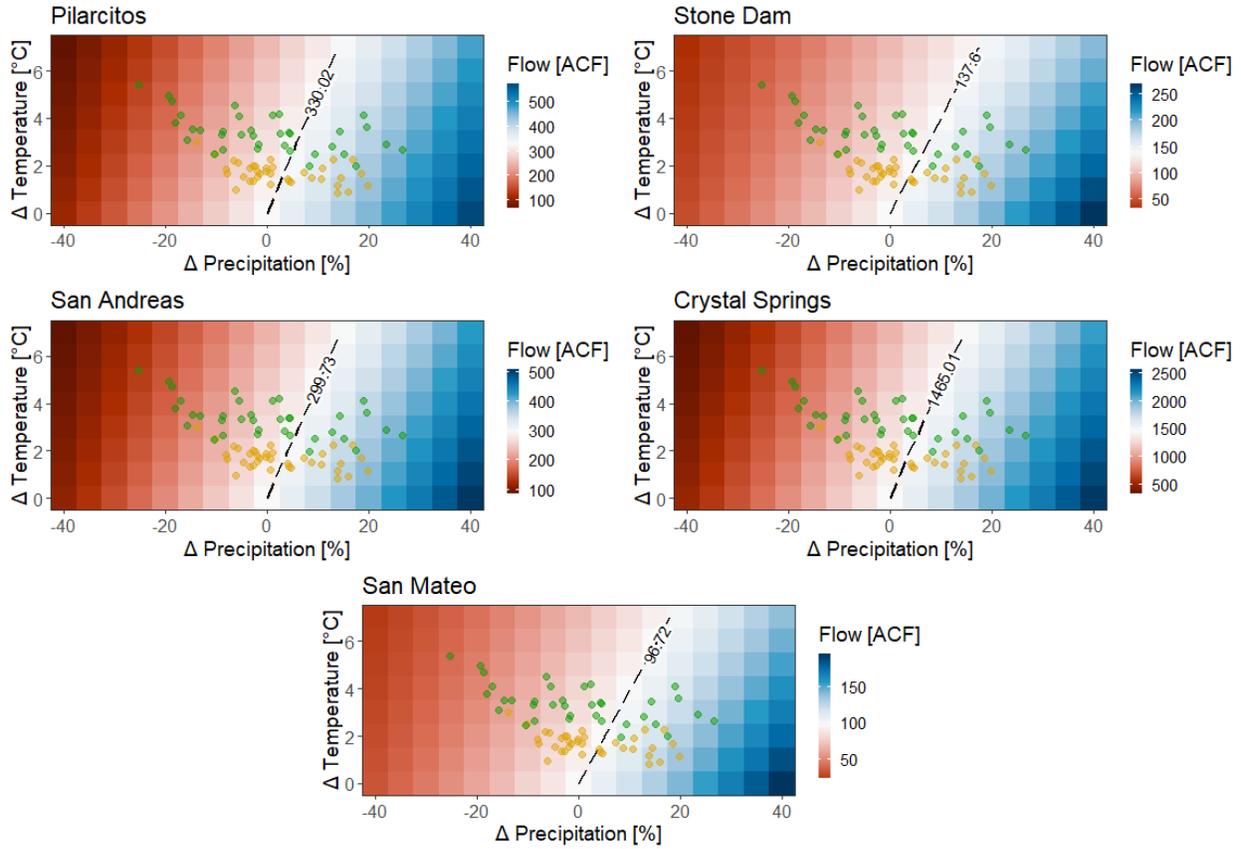
Figure 5-19 Total annual stream flow (acre-feet) under temperature and precipitation changes respectively for Crystal Springs Reservoir, San Andreas Reservoir, San Mateo Creek, Pilarcitos Reservoir, and Stone Dam sub-watersheds. The yellow and green dots over the response surface shows CMIP5 projections under RCP 8.5 for two 30-year long periods: 2040 (2026-2055) and 2070 (2056-2095). Baseline is 1986-2005. Historic mean observed is shown in white and follows the dashed line.

Figure 5-19 shows the total annual stream flow under temperature and precipitation changes. The pattern of total stream flow is similar to the East Bay climate stress test. The higher precipitation generates more flows while higher temperature produces less flows. This change is due to additional water collected in the watershed from precipitation events and more evapotranspiration due to increase in temperature. Results below are presented for inflow at Crystal Springs reservoir the largest watershed in the Peninsula.

- By 2040, the median projections of +2°C warming combined with 0% change in mean annual precipitation results in 7% reduction in mean annual inflow.
- By 2040, most projections and elicitations of warming are between +1°C and +4°C and for precipitation change between -5% and +5% which would correspond to a maximum change in mean annual inflow between a decrease of 23% and an increase of 10%.
- By 2070 RCP8.5, the median projections of about +4°C combined with 0% change in mean annual precipitation results 14% decrease in mean annual inflow. Most projections

and elicitations of warming range between +3°C and +6°C and of precipitation change between -15% and +15% resulting in change in mean annual inflow between a decrease of 46% and an increase of 29%.

6. Upcountry Hydrologic Drought Analysis

The aim of this drought analysis is to estimate how sensitive the frequency of droughts in the Tuolumne River is to changes in climate, and so it will focus on hydrologic drought which is the deficit in streamflow below a specified threshold. The analysis is limited in scope to the Upcountry region as it represents roughly 85% of yield from the RWS. This analysis is strictly looking at the availability of water for SFPUC on the Tuolumne River in sequences of dry years using multiple datasets of streamflow: historical, paleo-reconstruction and simulated with PRMS hydrologic model using stochastic weather generated inputs. Figure 6-1 presents historical deviations of annual flow from the mean on the Tuolumne River at La Grange. Consecutive years with negative deviation can be considered droughts. In recent history, three droughts are especially remarkable for SFPUC: 1976-1977, 1987-1992 and the most recent 2012-2016.

This analysis focuses on the frequency of two characteristics of droughts: the severity – cumulative flow deficit – and the duration. The frequency of occurrence of events such droughts or floods is often expressed as a return period, which gives the estimated time interval between events of similar characteristics (same severity, same duration or same severity and duration). The return period between occurrences is the inverse of the average frequency of occurrence. For example, a 100-year flood has a $1/100 = 0.01$ or 1% chance of being exceeded in any one year. This does not mean that if a 100-year flood occurs, then the next will occur in about one hundred years' time - instead, it means that, in any given year, there is a 1% chance that it will happen, regardless of when the last similar event was. The same reasoning is applicable to drought severity and duration.
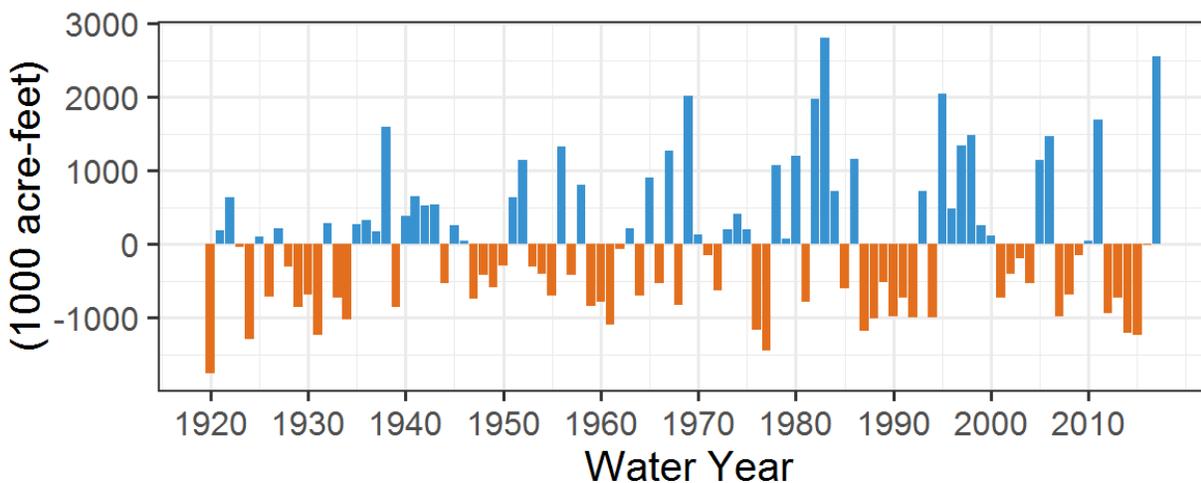
Figure 6-1. Unimpaired annual flow of the Tuolumne River at La Grange. Deviations from historical mean (1924-2017).

First, the analysis estimates the distribution of severity and duration, separately. However, there is a strong dependence between severity and duration (really severe droughts tend to be longer). The separate analysis of severity and duration does not reveal the significant correlation relationship between them. Therefore, a better approach for describing drought characteristics of severity and duration is to derive the

joint distribution of severity and duration. This analysis will allow to evaluate a change in frequency of drought severity and duration with changes in mean annual temperature and precipitation.

6.1. Definition of hydrologic drought

The scientific literature has proposed several indices to characterize drought episodes (Stahl et al. 2020). In this study, drought characteristics are derived using the theory of run (Yevjevich, 1967) by calculating the cumulative deficit and duration of shortfall of flow below a certain threshold. The considered 'flow' variable to be compared with the drought threshold is the Water Available to the City, or WAC. WAC is defined as the difference between the unimpaired Tuolumne River flow at La Grange Q_{LG}, and the actual irrigation district entitlement Q_{DE}:

$$WAC(t) = Q_{LG}(t) - Q_{DE}(t). \tag{6-1}$$

Q_{DE} is the minimum between Q_{LG} and the maximum irrigation district entitlements Q_{DE}^{max}:

$$Q_{DE}(t) = \min(Q_{LG}(t), Q_{DE}^{max}(t)), \tag{6-2}$$

where the maximum district entitlement is 4,800 AF/day (2,416 cfs) from June 13th to April 15th and 8,065 AF/day (4,066 cfs) from April 15th through June 13th. The cumulative deficit D during the water year Y is defined as:

$$D(Y) = \min(D(Y-1) + \textbf{Threshold} - WAC(Y), 0), \text{ where } D(Y = Y_i) = 0, \tag{6-3}$$

where Y_i is the first year of the time series. It is noted that equations (6-1) and (6-2) are solved at daily time step so that the calculation of WAC can account for the variations in the maximum district entitlement Q_{DE}^{max}. An illustration of the calculation of WAC is given in Figure 3-18 for the water year 1983 (wet) and 1987 (dry). When streamflow data are not available at daily resolution, a regression is used to estimate the annual actual district entitlement Q_{DE} (Section 6.2) and Equation (6-3) is solved at annual time scale. Two drought thresholds are considered below. The first threshold is 269 TAF, which is the largest annual transfer in the past decade from the Upcountry region to the East Bay and Peninsula regions. It occurred during the fiscal year 2012-2013. The second threshold is 365 TAF, which is the maximum annual transfer capacity from the Upcountry region based on the capacity of the San Joaquin Pipelines. This second threshold is used to assess the upper bound, given the current transfer capacity from the Upcountry region, of what the drought distribution may look like in a hypothetical reality in where the San Joaquin Pipelines were used at full capacity.

Figure 6-2 illustrates the simulated cumulative deficit for the two considered thresholds (middle and bottom panels). The unimpaired annual streamflow time series at La Grange used in Figure 6-2 (top panel) is a reconstructed time series that was provided by SFPUC. Drought events start when the cumulative deficit value gets negative and end when they reach their local maximum (hereafter denoted as severity). The effect of setting a larger threshold is clear on this figure: larger drought threshold leads to more frequent, longer and more severe droughts, despite some nonlinearities (cf. the one year long drought the happens right after the 1987-92 drought for 269 TAF threshold but does not exist with 365 TAF threshold). Table 6-1 lists the drought events recorded using both thresholds together with their

severity and duration. Note that all historical droughts highlighted in Figure 6-2 are correctly identified; and their duration and severity depend on the considered threshold.
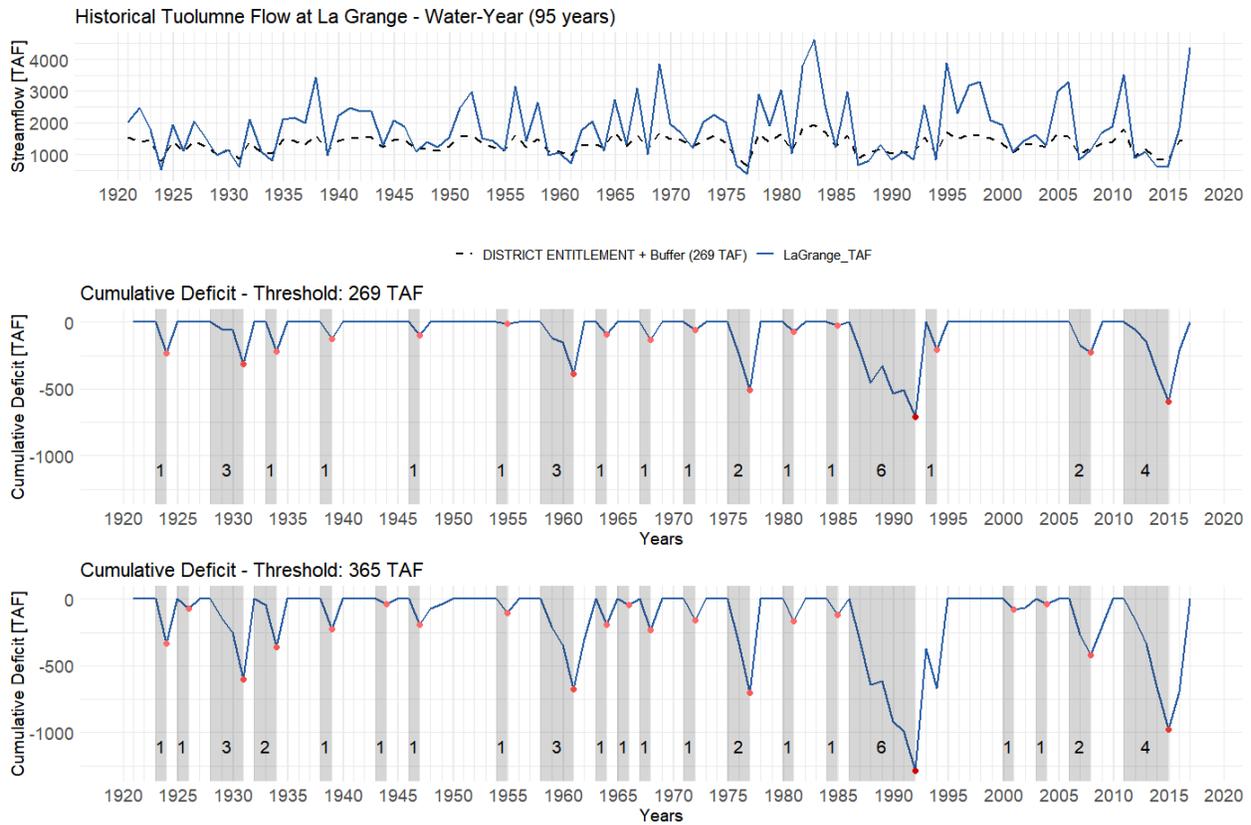
Figure 6-2. Historical Tuolumne Flow at La Grange with District Entitlement and threshold at 269 TAF (top) with the corresponding cumulative deficit plot (middle) and cumulative deficit plot for threshold at 365 TAF (bottom). The grey shaded areas show the identified droughts and the numbers tell their duration in years. The red dots show the associated severity to each drought event.

Table 6-1. Extracted drought events from historical Tuolumne flow at La Grange for two different thresholds. For each threshold, the drought events are sorted by decreasing severity.

	Threshold: 269 TAF			Threshold: 365 TAF	
Year Drought ends	**Severity [TAF]**	**Duration of Deficit [Years]**	**Year Drought Ends**	**Severity [TAF]**	**Duration of Deficit [Years]**
1992	707.39	6	1992	1283.39	6
2015	594.35	4	2015	978.35	4
1977	510.18	2	1977	702.18	2
1961	389.44	3	1961	677.44	3
1931	312.14	3	1931	600.14	3
1924	233.66	1	2008	418.98	2
2008	226.98	2	1934	357.10	2
1934	218.34	1	1924	329.66	1
1994	204.77	1	1968	229.06	1
1968	133.06	1	1939	223.20	1
1939	127.20	1	1947	190.42	1

1947	94.42	1	1964	189.19	1
1964	93.19	1	1981	165.90	1
1981	69.90	1	1972	154.99	1
1972	58.99	1	1985	118.42	1
1985	22.42	1	1955	104.96	1
1955	8.96	1	2001	75.15	1
			1926	72.70	1
			1966	45.69	1
			1944	37.45	1
			2004	37.09	1

6.2. Streamflow Datasets

Following the methodology described in the previous section, the detection of drought events and estimation of their severity and duration only requires a time series of unimpaired flow at La Grange. Three different data sources are used to provide such time series.

The first dataset is an annual time series of reconstructed unimpaired streamflow at La Grange from 1921 to 2011. This time series was provided by SFPUC and illustrated in Figure 6-3. The temporal resolution of this time series is annual, which prevents from calculating the actual district entitlements using the above equations. However, an annual time series of actual district entitlements was provided by SFPUC. This dataset will be used as reference for the severity and duration of the historical droughts.

The second dataset considered for the drought analysis is a collection of simulated streamflow time series at La Grange obtained from PRMS model when forced by climate realizations generated with CliWxGen. In addition to the 10 realizations, 500 other realizations are considered. More details about the input climatology is available in the Technical Report 1 (HRG TR1, 2021). The objective of using such a large number of stochastic realizations is to create a large collection of droughts to increase the robustness of the statistical inference of the distribution parameters for the severity and duration models.

The third dataset used for the drought analysis is a paleo record for the Tuolumne River at La Grange. This record is from the Northern California Tree Ring Study by California Department of Water Resources. This record spans from the year 900 to 2012 by creating a correlation based on tree-ring width to annual observed streamflow or precipitation data. The drought events extracted from the paleo dataset is of interest because it provides insights on the effect of the natural climate variability on the frequency, severity, and duration of droughts. Both paleo and reconstructed records are compared over the period 1921-2012 in Figure 6-3, allowing for an evaluation of the paleo streamflow. Over the period 1921-2012, the annual average of the paleo streamflow is 1,918.5 TAF, while the observed is 1,858.5 TAF. The result suggests a slight overestimation of paleo streamflow (+3.3%) with lower variability as outlined in Table 6-2. Therefore, quantile mapping bias correction was applied to the full paleo dataset using two fitted gamma distributions, which significantly reduced the bias in average and variability over the period 1921-2012 as indicated in Table 6-2.
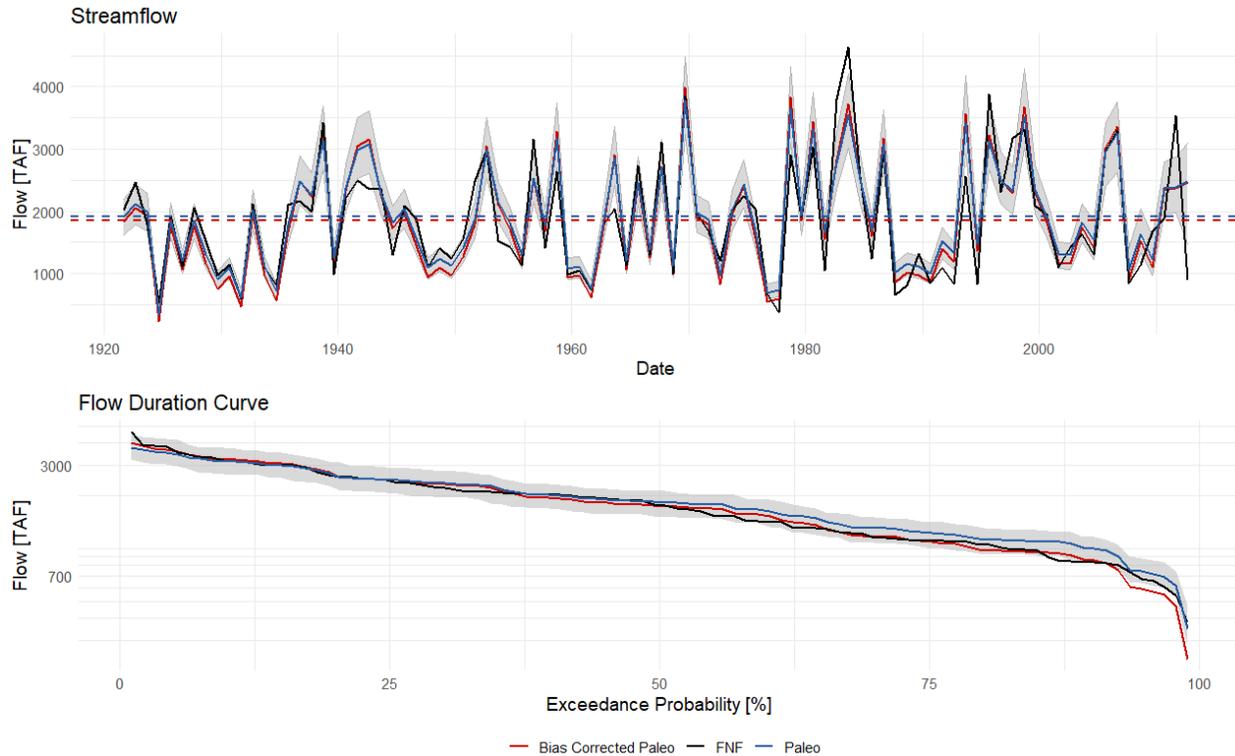
Figure 6-3. Comparison of available Paleo record data with observed historical full-naturalized flow of Tuolumne River at La Grange. These annual water-year time-series are shown from 1921 through 2012. The grey shaded region is the 50% confidence interval for the Paleo record. The dashed lines represent the long-term average for full naturalized flow at La Grange (black), Paleo dataset (blue) and bias-corrected Paleo (red).

Table 6-2. Summary statistics for Historic, Paleo, and Bias-corrected Paleo time-series

Dataset (1921-2012)	Mean (TAF)	Standard Deviation (TAF)
Historic	1858.5	907.9
Paleo	1918.5	810.8
Bias Corrected Paleo	1859.4	901.9

The paleo record at La Grange is available at annual temporal scale and, contrary to the historical record reconstructed by SFPUC, it does not come with a reconstruction of the actual district entitlements. As such, the actual district entitlements for the paleo records were estimated using the regression presented in Figure 6-4. This regression has been estimated using the historical reconstruction of the flow at La Grange and actual district entitlements provided by SFPUC. It uses annual unimpaired flow at La Grange as predictor to predict the annual actual district entitlement. The fit of the regression model was deemed satisfying for the drought analysis.

95
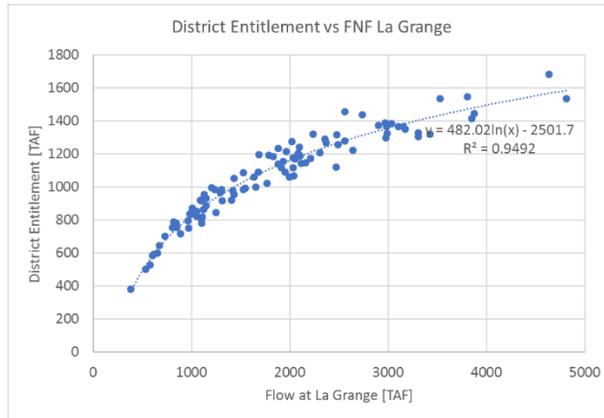
Figure 6-4. Regression between actual district entitlements and unimpaired flow at La Grange.

6.3. Fitting of distributions for severity and duration

This section presents first the identified drought events using the methodology described in Section 6.16.2 and the datasets presented in Section 6.2. Then, statistical distributions are selected, and their parameters calibrated for modeling the frequency of drought severity and duration. Return periods for historical droughts are provided for both severity and duration. Last, copula functions are investigated to model the relationship between severity and duration and provide return periods accounting for the joint distribution of severity and duration of droughts.

Drought events were extracted from 1,113 years of paleo streamflow records (900 – 2012), 100 years of historic Tuolumne Flow at La Grange (1921 – 2020), 49 years from each of the 10 climate realizations from the CliWxGen-WG, and 49 years from the 500 drought realizations selected from the remaining realizations in CliWxGen-WG (only 49 years from the 50-year long simulations were used because the first year serves as warm-up period for PRMS). When put together, the different datasets accumulate to a total of 26,110 years. This latter dataset is further denoted as 'combined' dataset. Note that when combined, the years in the paleo records after 1920 were disregarded because this period is available from the reconstructed historical streamflow provided by SFPUC.

Figure 6-5 shows the distribution of drought severity and duration across the identified events using the combined dataset (numbers of identified events for each dataset are given in Table 6-3). Using the combined dataset and a drought threshold set to 269 TAF, a total 4,318 drought events have been extracted (average duration is 1.61 years and average severity is 217 TAF). When using a threshold equal to 365 TAF, this number grows to 4,351 (average duration is 2.04 years and average severity is 346 TAF). It is interesting to note that the number of events does not increase significantly with a larger threshold (i.e., +33 events). However, the average severity increases significantly by +129 TAF, which corresponds to an increase by almost 60%. The severity of the identified drought events ranges from roughly 0 to 1,218 TAF or 1,985 TAF, whether the 269 or 365 TAF is used. Similar to severity, the duration of the simulated drought events is significantly influenced by the chosen thresholds. Using a 269 TAF threshold, the duration ranges from 1 to 9 years. Only 4 droughts have a duration of 8 years or more (difficult to see on the figure due to scale and resolution). Using the 365 TAF threshold, the duration ranges from 1 to 14 years.
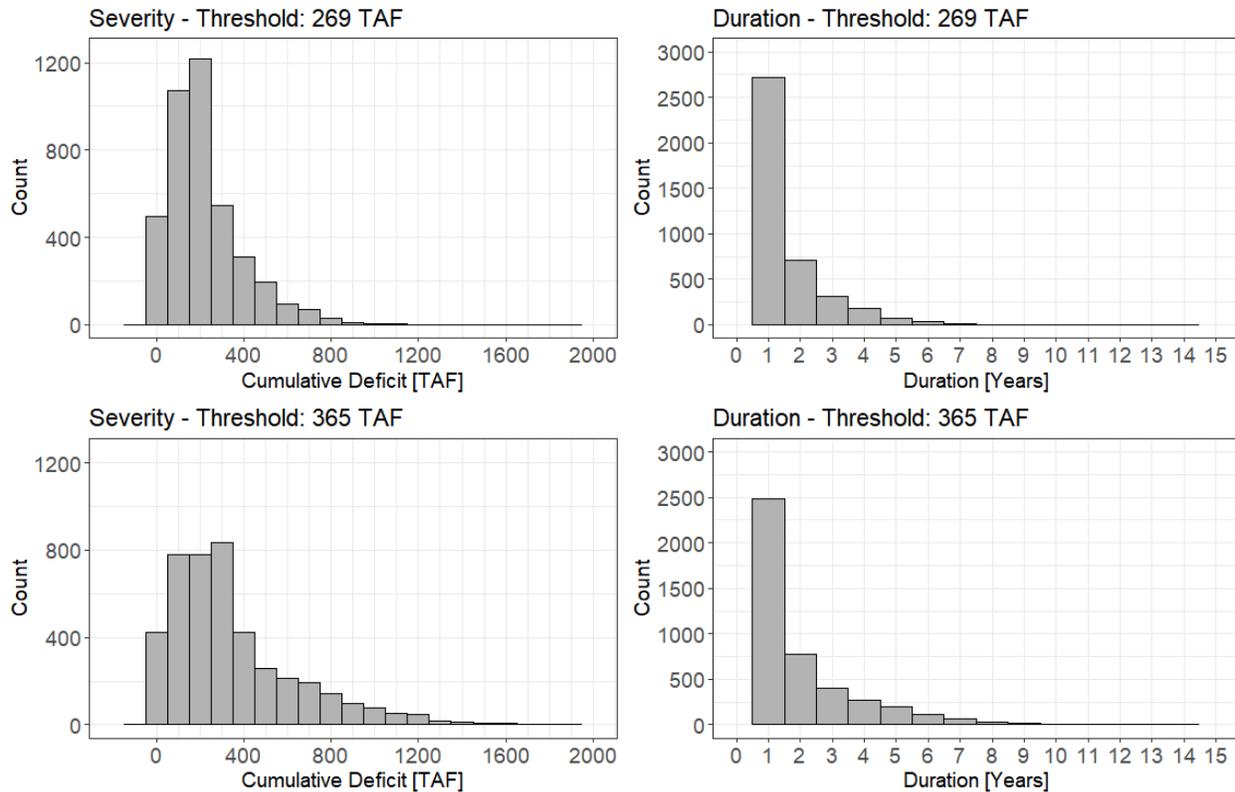
Figure 6-5. Distribution of the severity (left) and duration (right) of the identified drought events using the 269 TAF (top) and 365 TAF (right) thresholds with the combined dataset.

Table 6-3. Number of identified drought events using theory of run for each dataset. The combined dataset is a combined dataset with drought events from historical (1921-2020), Paleo (900-1920), and 510 stochastic realizations.

Dataset	Number of years	Average Severity (TAF)		Number of Drought Events	
		Threshold: 269 TAF	Threshold: 365 TAF	Threshold: 269 TAF	Threshold: 365 TAF
Historical (1921 – 2020)	100	236	333	17	21
Paleo (900 – 2012)	1,113	147	222	135	173
Bias Corrected Paleo (900 – 2012)	1,113	188	294	171	190
Stochastic Realizations	24,990	220	351	3,902	4,172
Combined	26,110	217	346	4,318	4,351

Several statistical distributions have been investigated to model the frequency of the severity and duration of the identified droughts. The considered distributions to model severity and duration are given in Table 6-4 and Table 6-5, respectively. Continuous distributions were tried for severity and both continuous and discrete distributions were used for duration. The method to calibrate the model parameters is the Maximum Likelihood Estimation (MLE). The fitted models were evaluated based on their visual fit to

97

density plots, quantile (Q-Q) plots, cumulative distribution function (CDF) plots, and probability (P-P) plots. Figure 6-6 shows these plots for the Weibull distribution fit, the selected distribution, to cumulative drought deficit for a threshold of 269 TAF with the data fitting the empirical fitted distribution as closely as possible.
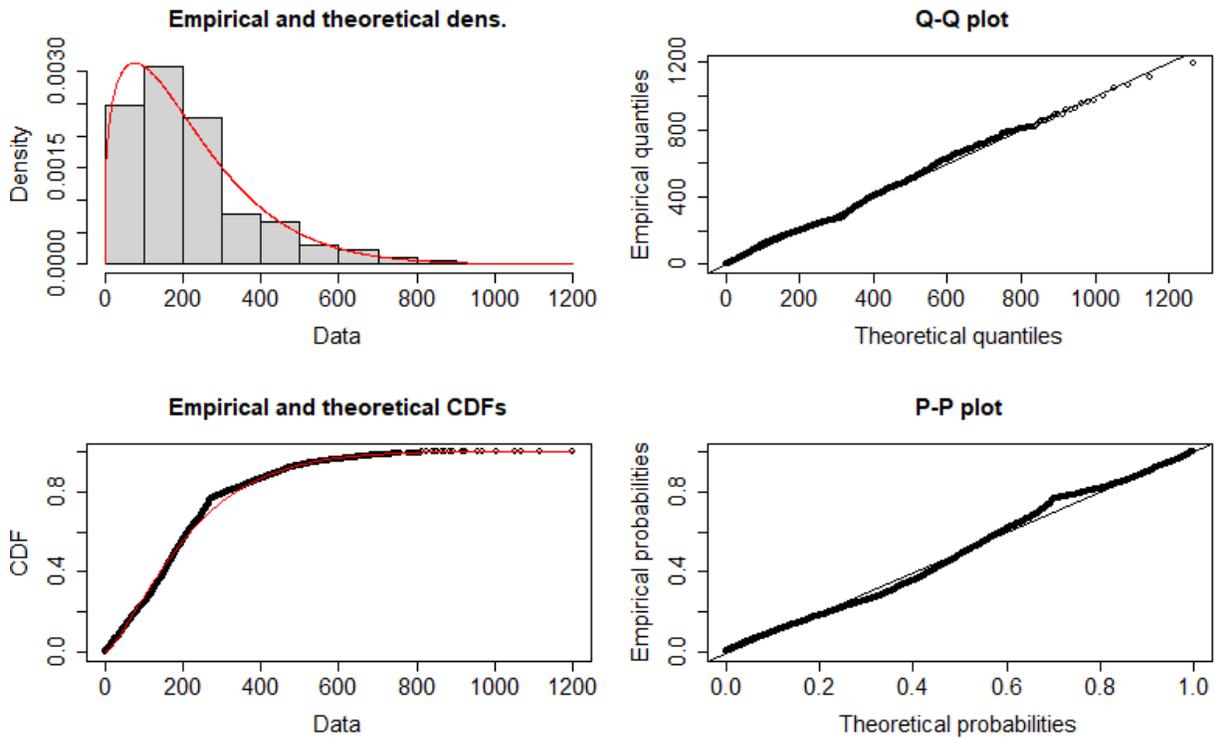
Figure 6-6. Weibull distribution fit on cumulative deficit for a threshold of 269 TAF. The figures are distribution fit to the density plot (top-left), Q-Q plot (top-right), cumulative distribution function plot compared to the fitted distribution (bottom left), and the P-P plot (bottom-right).
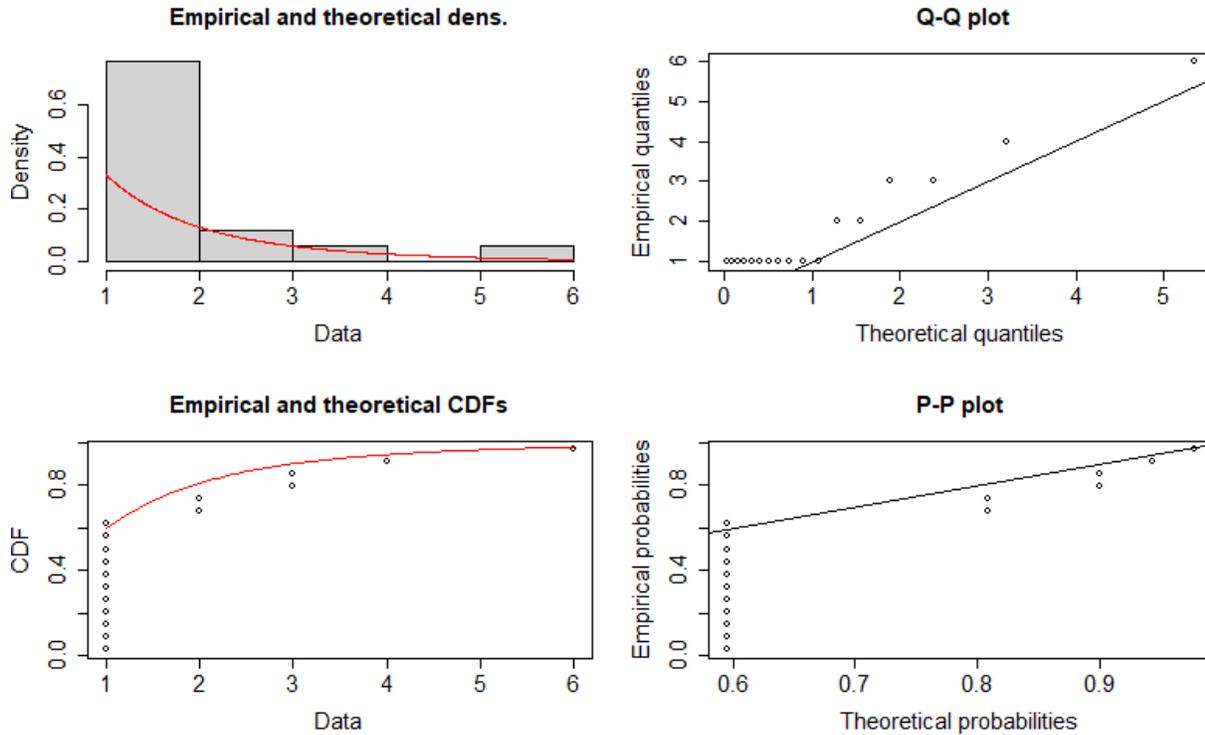
Figure 6-7. Generalized Pareto distribution (GPD) fit on duration for a threshold of 269 TAF. The figures are distribution fit to the density plot (top-left), Q-Q plot (top-right), cumulative distribution function plot compared to the fitted distribution (bottom left), and the P-P plot (bottom-right).

The fitted distributions were evaluated considering both the Akaike and Bayesian information criteria (AIC and BIC, respectively), in addition to the visual inspection of the goodness of fit of the empiric and simulated cumulative distributions for severity and duration. AIC and BIC criteria are commonly used to compare the goodness-of-fit across multiple distributions. They are both negatively oriented (low values are better than high values). For the combined dataset, Table 6-4 shows that a Weibull distribution provides the best fit to model the frequency of the severity of the identified drought events. Still for the combined dataset, Table 6-5 shows that the Log-Pearson Type-III distribution provides the minimum AIC score, although the visual fit for the drought duration cumulative distribution is bad (Figure 6-8). For this reason, the selection of the distribution to model the frequency of the drought duration was based on the visual fit of the data. As such, the Generalized Pareto Distribution (GPD) was selected to model the frequency of the drought duration (Figure 6-8). Weibull (severity) and GPD (duration) were also found to be suitable distributions for the other datasets. Note that the choice of either 269 or 365 TAF for threshold does not influence the model selection. The calibrated parameters of the fitted distribution are given in Table 6-6.

Table 6-4. Cumulative drought deficit distribution fit criterions for thresholds 259 TAF and 365 TAF. AIC and BIC are obtained for the combined dataset. Distributions are sorted by increasing AIC. The * symbol indicates the selected model for the drought severity.

Distribution	Threshold: 269 TAF		Threshold: 365 TAF	
	AIC	BIC	AIC	BIC
Weibull*	49689	49702	59611	59624
Gamma	49792	49804	59665	59662
Log-Pearson Type III	49868	49887	59759	59778
Generalized Pareto (GPD)	49872	49884	59650	59662
Generalized Extreme Value (GEV)	49898	49917	60061	60080
Exponential	50147	50153	59789	59795

Table 6-5. Drought duration distribution fit criterions for thresholds 259 TAF and 365 TAF. AIC and BIC are obtained for the combined dataset. Distribution are sorted by increasing AIC. The * symbol indicates the selected model for the drought duration. As indicted in the text, the selection of the model for drought duration was based on the visual fit of the empirical distribution rather than the AIC score. 'Inf' stands for Infinity.

Distribution	Threshold: 269 TAF		Threshold: 365 TAF	
	AIC	BIC	AIC	BIC
Log-Pearson Type III	5959	5978	10139	10158
Gamma	9320	9333	13521	13534
Weibull	10099	10112	14044	14057
Poisson	11290	11297	15103	15109
Negative Binomial	11292	11305	15102	15109
Exponential	11635	11641	14964	14970
Generalized Pareto (GPD)*	12618	12624	16716	16722
Geometric	13666	13673	16820	16826
Generalized Extreme Value (GEV)	*Inf*	*Inf*	12463	12482
Binomial	*Inf*	*Inf*	*Inf*	*Inf*
Hypergeometric	*Inf*	*Inf*	*Inf*	*Inf*
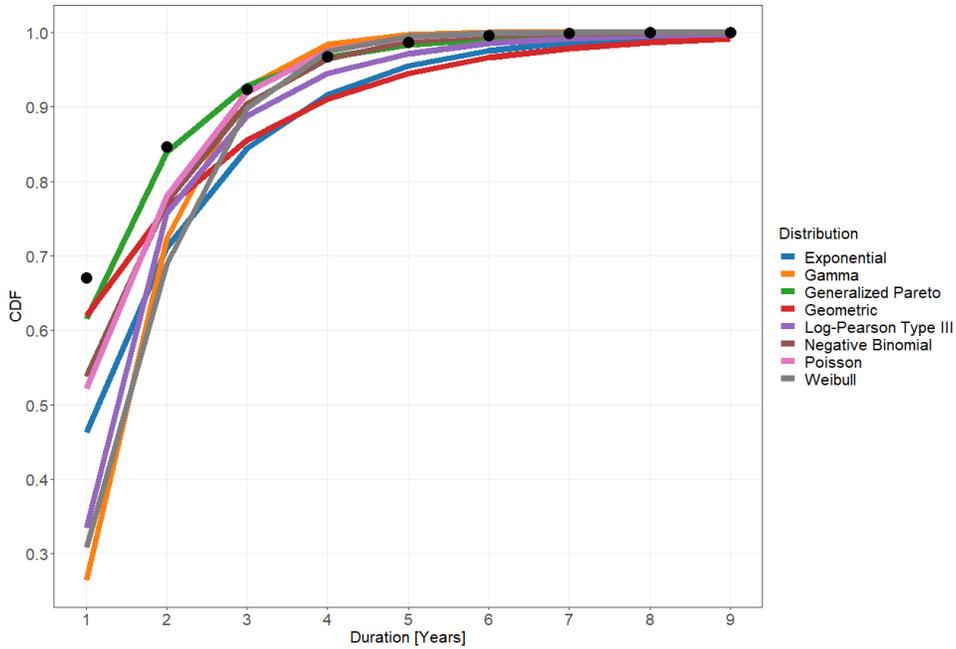
Figure 6-8. Empirical (black dots) and theoretical (color lines) cumulative distribution functions for drought duration.

Table 6-6. Distribution parameters for fitted baseline scenario distributions for severity and duration.

Variable	Distribution	Dataset	Threshold: 269 TAF		Threshold: 365 TAF	
			Shape	**Scale**	**Shape**	**Scale**
Severity	Weibull	Historical	1.14	246.94	1.06	341.59
		Paleo	1.23	156.90	1.31	240.39
		Bias Corrected Paleo	1.39	205.59	1.29	316.43
		Stochastic Realizations	1.31	237.76	1.18	370.83
		Combined	1.30	234.85	1.18	365.58
Duration	GPD	Historical	0.22	0	0.17	0
		Paleo	-0.14	0	0.023	0
		Bias Corrected Paleo	-0.07	0	0.17	0
		Stochastic Realizations	0.10	0	0.33	0
		Combined	0.063	0	0.32	0

The return period T for a specific severity or duration value x is calculated based on:

$$T = \frac{\mu}{1 - F_x(x)},$$

(6-4)

where $F_x(x)$ is the cumulative distribution function value for the variable x, and μ is the mean inter-arrival time (Gräler, 2013). The mean inter-arrival time is estimated as the ratio between the total numbers of years and the number of drought events. As such, it varies depending on the considered threshold and the considered dataset. The mean inter-arrival time values for each dataset are given for each dataset and threshold in Table 6-7.

Table 6-7. Mean inter-arrival time obtained for the considered datasets and the two drought thresholds

Dataset	Mean inter-arrival time μ (years)	
	Threshold: 269 TAF	Threshold: 365 TAF
Historic	4.76	5.88
Paleo	8.24	6.43
Bias Corrected Paleo	6.51	5.86
Stochastic Realizations	6.40	5.99
Combined	6.05	6.00

Figure 6-9 illustrates the drought severity and duration frequency curves obtained with the considered datasets and the 269 TAF threshold. Figure 6-10 shows similar frequency curves but obtained using 365 TAF threshold. The sampling uncertainty is shown using 95% confidence intervals obtained with bootstrap method (5,000 random samples were used). The three most recent major historical droughts are indicated on these figures. The estimated return periods for the historical droughts estimated from each dataset and using the two considered thresholds are given in Table 6-8.
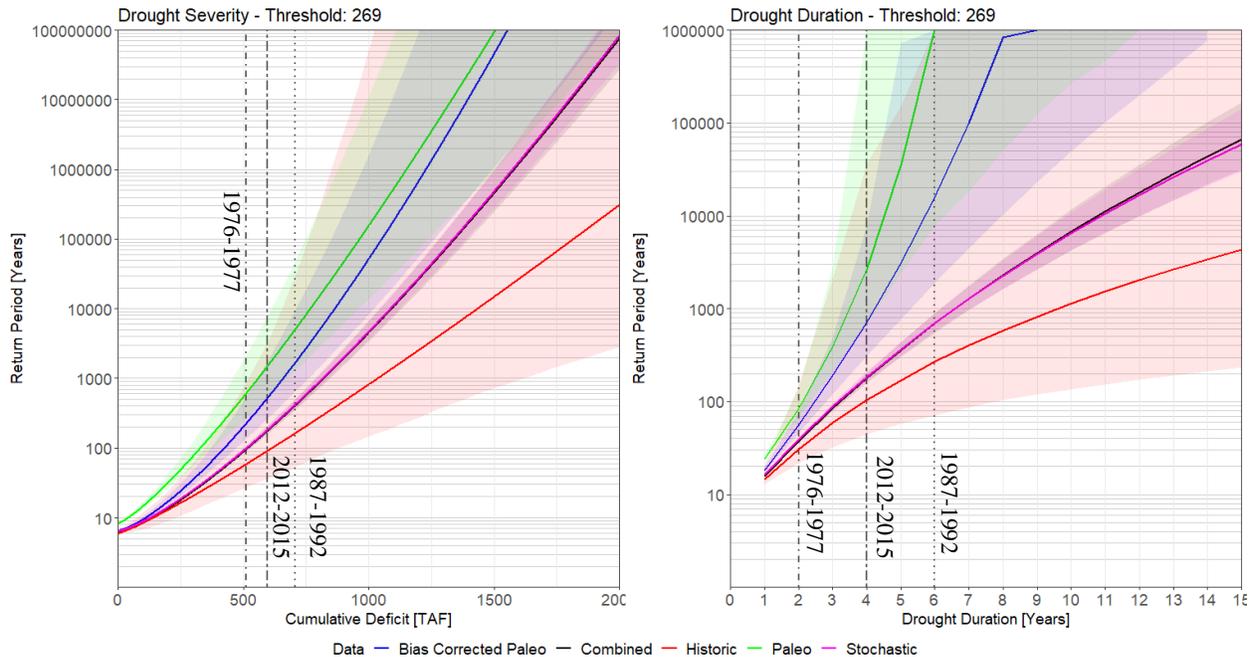
Figure 6-9. Drought frequency curves for severity (left) and duration (right). Results are obtained for the 269 TAF threshold. The color lines show the frequency curves obtained for the considered datasets: bias corrected paleo

(blue), paleo (green), historic (red), 510 stochastic realizations (purple) and the combined dataset (black). The shaded areas show the 95% confidence intervals obtained using the bootstrap method for each model. Severity and duration for the historic drought events are shown and labeled as vertical lines.
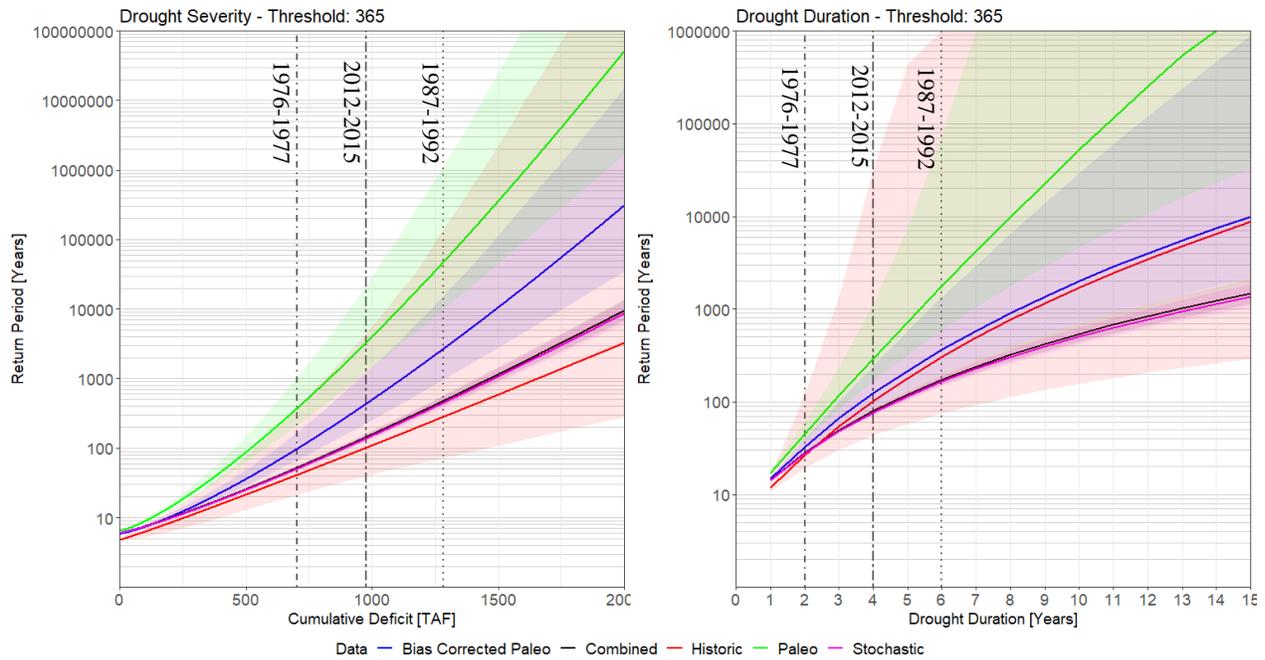
Figure 6-10. Drought frequency curves for severity (left) and duration (right). Results are obtained for the 269 TAF threshold. Caption details are similar than Error! Reference source not found..

Table 6-8. Estimated return periods of drought severity and duration for the historic drought events.

Drought Event	Dataset	Threshold: 269 TAF Return Period [Years]		Threshold: 365 TAF Return Period [Years]	
		Severity	Duration	Severity	Duration
1976-1977	Historic	58	31	41	27
	Paleo	599	85	370	45
	Bias Corrected Paleo	221	56	97	33
	Stochastic Realizations	98	40	50	28
	Combined	95	38	52	28
1987-1992	Historic	163	267	280	305
	Paleo	5,033	*Inf*	47,385	1,769
	Bias Corrected Paleo	1,673	15,838	2,672	364

	Stochastic Realizations	420	698	447	165
	Combined	407	694	475	172
	Historic	89	103	101	102
	Paleo	1,455	2,644	3,324	295
2012-2015	Bias Corrected Paleo	508	714	436	124
	Stochastic Realizations	178	184	137	77
	Combined	173	179	144	79

Estimated return periods for drought severity and duration for historic droughts are shown to be highly dependent on the underlying datasets used to fit the statistical models. For instance, the return period estimates for the severity of the 1976-77 and 1987-92 droughts range from 58 to 599 years, and from 163 to 5,033 years, respectively (Table 6-8). Similar comments can be made for the drought duration (Table 6-8). One important explanation for the large estimate range across datasets is the difference in number of years and identified drought events within each datasets (Table 6-3). Return periods calculated with a limited number of identified drought events corresponding are expected to be unreliable, especially for severity or duration significantly beyond the range observed in the dataset (Bonaccorso, et. al. 2003). The sampling uncertainty can be visualized in Figure 6-10 with the 95% confidence intervals surrounding the frequency curve obtained from each dataset (i.e., shaded areas in color). Confidence intervals were obtained via the bootstrap method using 5,000 random samplings. Confidence intervals are large for datasets with a limited number of events available to fit the distribution (e.g., historic and paleo), while they are small for datasets with a large number of droughts.

6.4. Joint distribution of severity and duration

As aforementioned, the separate analysis of severity and duration does not reveal the significant correlation relationship between them. Therefore, a better approach for describing drought characteristics of severity and duration is to derive the joint distribution of severity and duration. This analysis will allow to evaluate a change in frequency of drought severity and duration with changes in mean annual temperature and precipitation.

Figure 6-11 illustrates the relationship between drought severity and duration as seen across the identified events and for the two considered thresholds. The 1976-77 and 1987-92 droughts are labeled on the plot. Several observations can be made from this figure.

First, it is noted that the severity tends to increase with the duration of the events, although the severity varies significantly for each duration. For example, the severity of 2-year long drought events ranges from about 0 to about 550 TAF or 750 TAF for drought thresholds set to 269 or 365 TAF, respectively. Note that the historical 1976-77 drought is on the very high-end of this range.

Another highlight shown in Figure 6-11 is that droughts with rather short duration (e.g., less than 3 years) can be as severe, if not more severe, than droughts with significantly longer duration. For example,

considering a 269 TAF threshold, the severity of the most severe 2-year long droughts is similar to the median severity of the 4-year long droughts, which is also twice as much as the severity of the less severe 6-year long drought.
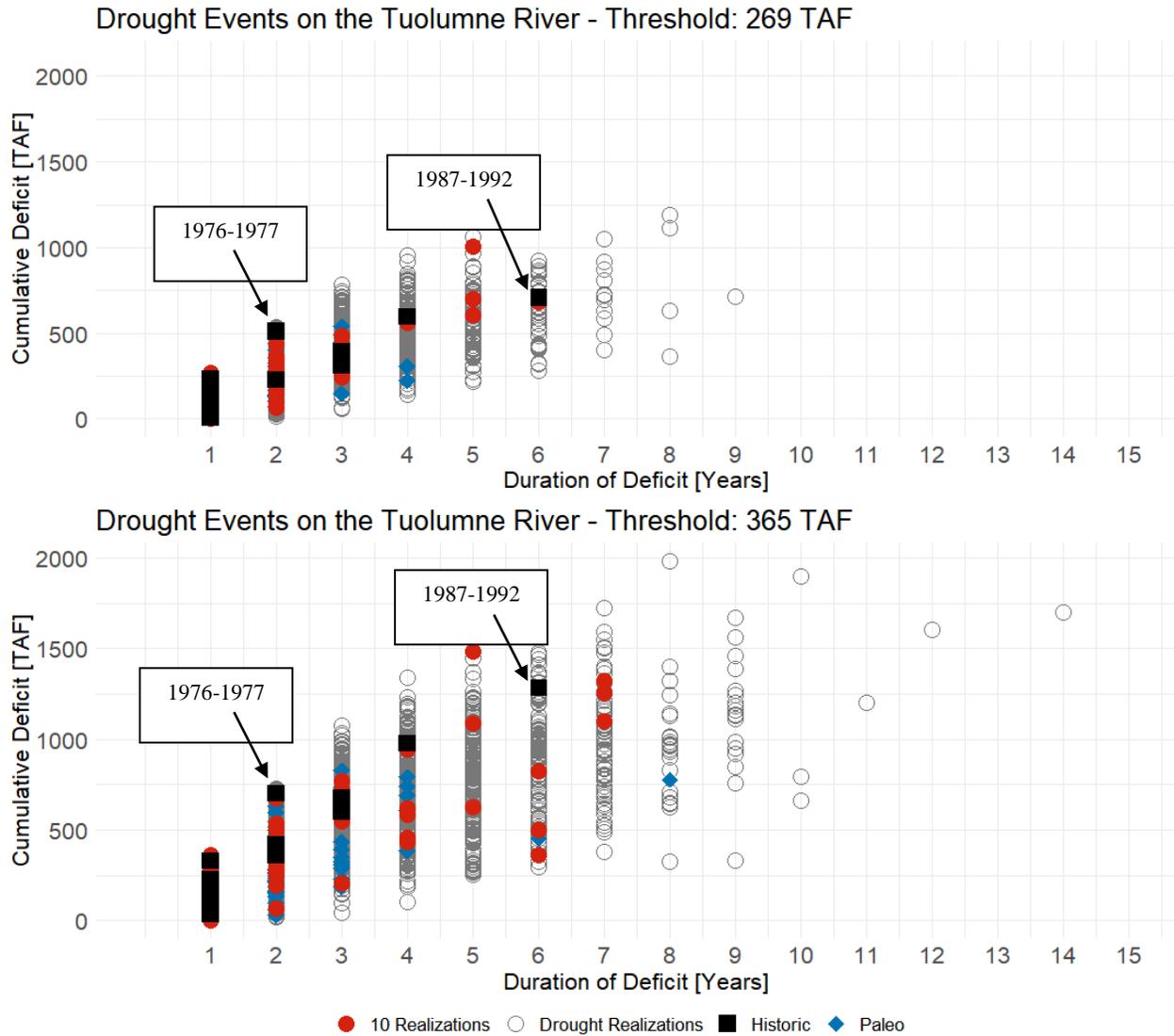
Figure 6-11. Extracted drought events across all 10 climate realizations (red circle), 500 drought realizations (grey circle), historic data (black square), and paleo (blue diamond) records for the two thresholds; 269 TAF (top) and 365 TAF (bottom).

The complex relationship between drought duration and severity illustrated above highlights the importance of accounting for the joint probability of the severity and duration for assessing drought risk. To provide insight into the relationship between drought severity and duration, Kendall's rank correlation coefficient, Spearman's rank correlation coefficient, and Pearson's linear correlation coefficient are shown in Table 6-9. All correlation coefficients are positive implying that longer droughts tend to be more severe. Pearson's correlation indicates an explained variance slightly above either 61% or 66% when using either 269 TAF or 365 TAF threshold, indicating a significant relationship between severity

and duration (i.e., the explained variance is obtained by taking the squared value of the Pearson' correlation coefficient).

Table 6-9. Correlation coefficients between droughts severity and duration

Correlation	Threshold: 269 TAF	Threshold: 365 TAF
Kendall's Rank Correlation Coefficient	0.564	0.628
Spearman's Rank Correlation Coefficient	0.694	0.773
Pearson's Linear Correlation Coefficient	0.782	0.817

In statistics, the dependence between two random variables, say Y and Z, can be modeled via their joint distribution using multivariate distributions, such as the bivariate normal distribution for instance. However, one major limitation of the multivariate distributions is that the marginal distributions must follow the same theoretical distribution. For instance, in the case of the bivariate normal distribution, the two marginal distributions are normal. This is a significant limitation for the drought analysis conducted for the LTVA because the drought severity and duration of the identified drought events are distributed following two different distributions; a Weibull distribution for the severity and the GPD for the duration (Section 6.3).

Following Sklar's theorem (Sklar, 1959), however, any multivariate joint distribution $F(y, z)$ can be specified from the univariate marginal distributions for Y and Z and a copula function C that describes the dependence structure between the random variables Y and Z (e.g., Genest and Favre, 2007):

$$F(y, z) = \Pr(Y < y, Z < z) = C(F_Y(y), F_Z(z), \vartheta) \tag{6-5}$$

where ϑ is a parameter that controls the degree of dependence between Y and Z. In the case developed here, Y and Z are the duration and severity of the identified drought events. Various copula functions were fit using maximum likelihood estimation, and the selection was made based on the Akaike Information Criteria (AIC). The survival Clayton (Clayton, 1978) and Gumbel (Gumbel, 1960) copula functions were found to fit best the severity and duration joint distribution, respectively when using 269 TAF or 365 TAF thresholds (Table 6-10).

The structure of the Clayton copula is defined as:

$$C_\vartheta^{Cl}(u_1, \dots, u_n) = \left(\sum_i \left(u_i^{-\vartheta} - 1 \right) + 1 \right)^{1/\vartheta}, \text{where } \vartheta \geq 0 \text{ and } 0 \leq u_i \leq 1, i = 1, \dots, n \tag{6-6}$$

The survival Copula is a 180° rotation of the input pseudo-observation of the covariate u_n, which is shown in equation (6-7 with survival copula $\check{C}$ and the copula $\bar{C}$.

$$\check{C}(u_1, \ldots, u_n, \ldots, u_N) = \bar{C}(1 - u_1, \ldots, 1 - u_n, \ldots, 1 - u_N) \tag{6-7}$$

The structure of the Gumbel copula is defined as:

$$C_\vartheta^{Gu}(u_1, \ldots, u_n) = \exp\left(\left(\sum_i (-\log u_i)^\vartheta\right)^{1/\vartheta}\right), \text{where } \vartheta \geq 1 \text{ and } 0 \leq u_i \leq 1, i \tag{6-8}$$
$$= 1, \ldots, n$$

For the application of the above copula functions, $n=2$ as only 2 co-variates are considered (i.e., severity and duration).

Table 6-10. Copula model fit performance for various copula families. Lower the value, the better the fit. Bold values show the minimum AIC values that indicate the copula function to use for each threshold.

Copula	Threshold: 269 TAF		Threshold: 365 TAF	
	AIC	BIC	AIC	BIC
Survival Clayton	**-2909.63**	-2903.33	-3622.42	-3616.04
Clayton	-2888.01	-2881.70	-3668.83	-3662.46
Joe	-2852.59	-2846.28	-3622.42	-3616.04
Gumbel	-2828.46	-2822.15	**-3746.15**	-3739.77
Gaussian	-2541.45	-2535.15	-3552.63	-3546.25
Student t-Copula	-2539.78	-2527.17	-3544.26	-3531.51
Frank	-2485.88	-2479.58	-3593.14	-3586.77

The fitted copula parameters are respectively $\vartheta = 2.54$ and $\vartheta = 2.54$ for the survival Clayton (269 TAF) and Gumbel (365 TAF) copula functions. The estimates of the Kendall's rank correlation coefficients from the fitted copulas respectively equal 0.537 and 0.606, which indicates that the fitted copula functions fairly well represent the observed rank correlation (0.564 and 0.628, respectively for 269 and 365 TAF threshold; Table 6-9).

The joint return period of an event with severity x and duration y is calculated as:

$$T_{joint} = \frac{\mu}{1 - F(x) - F(y) + C(x,y)}, \tag{6-9}$$

with $F(.)$ the cumulative distribution function for each margin and μ the average inter-arrival time.
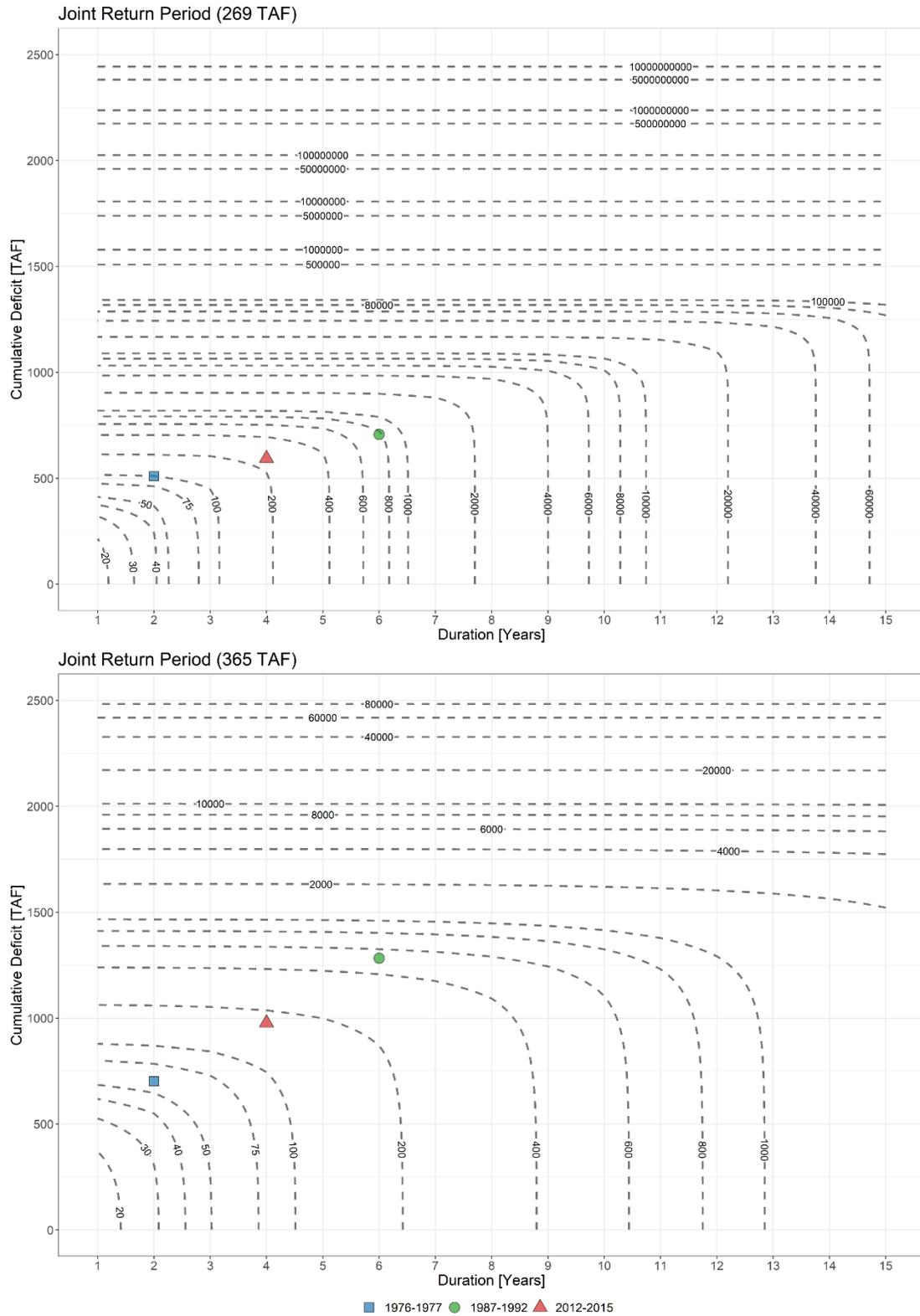
Figure 6-12. Joint return period for thresholds 269 TAF (top) and 365 TAF (bottom) on the Tuolumne River. Selected historical droughts are highlighted in the figure. Contour lines are shown as dashed black lines with labels.

108

Table 6-11. Calculated joint return period for thresholds 269 and 365 TAF.

Drought Event	Threshold: 269 TAF Joint Return Period [Year]	Threshold: 365 TAF Joint Return Period [Year]
1976-1977	98	61
1987-1992	772	537
2012-2015	236	176

The estimates of the return period for the most important historical droughts are shown in Table 6-11. Figure 6-12 shows the contour return periods based on severity and duration. The 1987-92 drought is shown to be the historical drought with the largest return period (772 and 537 years for 269 TAF and 365 TAF thresholds). The return period of the 1976-77 drought was estimated to 98 years with 269 TAF threshold, and to 61 years for 365 TAF threshold. Finally, the return period of the most recent 2012-2015 drought was estimated to 236 and 176 years, respectively for 269 and 365 TAF threshold.

6.5. Drought return period under climate change

Impact of changes in precipitation and temperature on drought severity and duration are presented in this section.

6.5.1. Changes in Severity

Figure 6-13 illustrates the change in drought severity for a sample of scenarios of precipitation changes (i.e., -10% and -20%) and temperature changes (i.e., +2°C and +4°C), and for the two considered drought thresholds. Table 6-12 gives the estimated return periods for the historical droughts for the precipitation and temperature scenarios shown in.
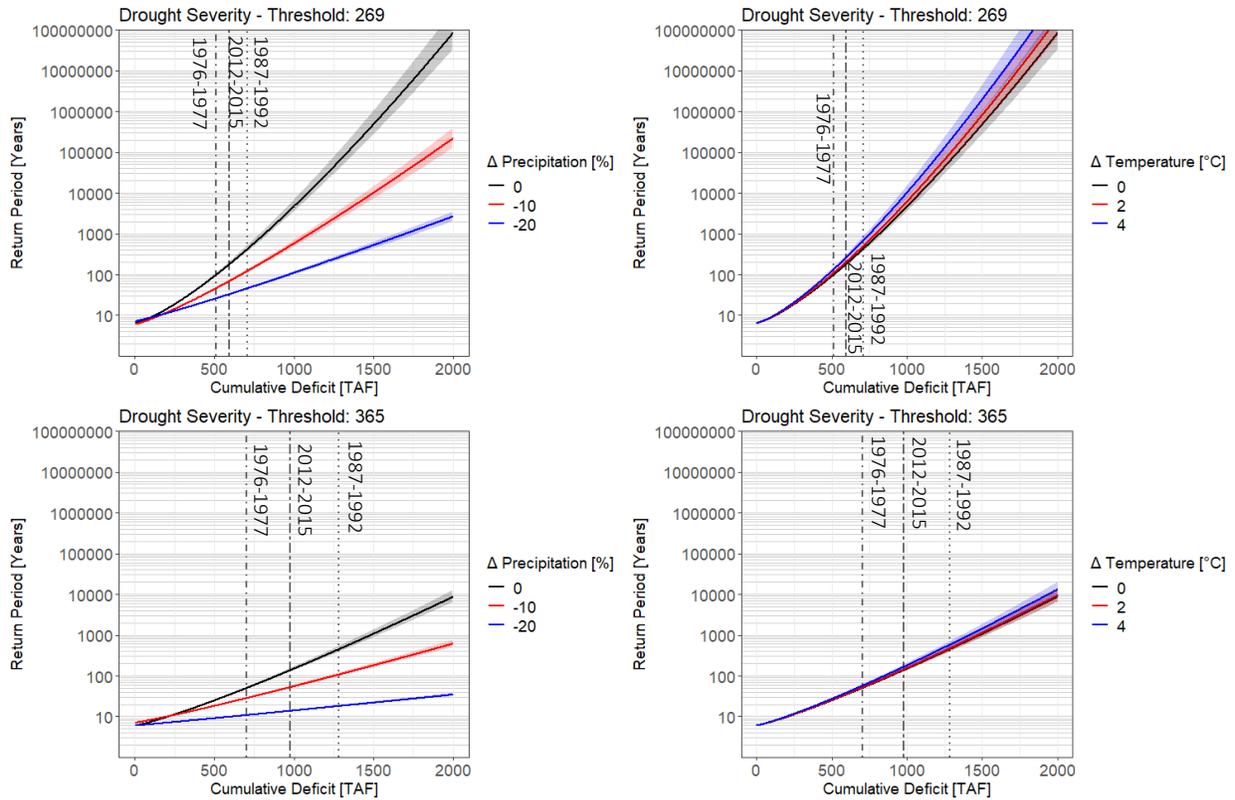
Figure 6-13. Effect of precipitation (left) and temperature (right) change on the drought severity frequency. Results are shown with 269 TAF (top) and 365 TAF (bottom) drought threshold. The x-axis shows that the absolute value of the cumulative deficit and the y-axis shows the estimated return period (years). The 95% confidence intervals are shown in shaded areas and are obtained using the bootstrap method. Vertical lines shows the severity of the historical droughts obtained using the historical dataset.

Table 6-12. Effect of precipitation and temperature change on the return periods associated with the severity of the historic droughts. Return periods are round off to the nearest 5 years.

Threshold [TAF]	Drought Event	Changes in Precipitation			Changes in Temperature [°C]		
		0%	-10%	-20%	0	+2	+4
	1976-1977	100	45	25	100	105	130
269	1987-1992	420	120	45	420	495	675
	2012-2015	180	70	35	180	200	260
	1976-1977	50	30	10	50	50	60
365	1987-1992	445	105	20	445	470	575
	2012-2015	135	55	15	140	145	165

Precipitation change is a significant driver of change for the severity of drought. Any reduction in precipitation leads to an increase in frequency of drought severity (i.e., a reduction of the return period of

110

any severity level). For instance, considering the threshold of 269 TAF and a 10% precipitation reduction, the severity associated with the 1976-77 becomes roughly twice as more frequent (from 98 to 46 years return period). Similarly, the return period associated with the 2012-2015 drought drops from 178 to 69 years. The changes in return periods for the severities associated with the historical events under a 20% precipitation reduction scenario are more drastic. All three highlighted historical droughts see their return period fall below 50 years, including the severity associated with the 1987-1992 droughts whose estimated return period under baseline climate scenario is larger than 400 years.

The effect of warming temperature appears to be opposite with an increase in return period for all historical droughts. For instance, the return period of the severity associated with the 1987-92 drought increase from 420 to 494 years under 2°C warming.

Similar results are obtained for the 365 TAF threshold, although the change following from a reduction in precipitation are more important, and in the case of temperature warming, less important.

6.5.2. Changes in Duration

Figure 6-14 illustrates the change in drought duration for a sample of scenarios of precipitation changes (i.e., -10% and -20%) and temperature changes (i.e., +2°C and +4°C), and for the two considered drought thresholds. Table 6-13 gives the estimated return periods for the historical droughts for the scenarios shown in Figure 6-14.
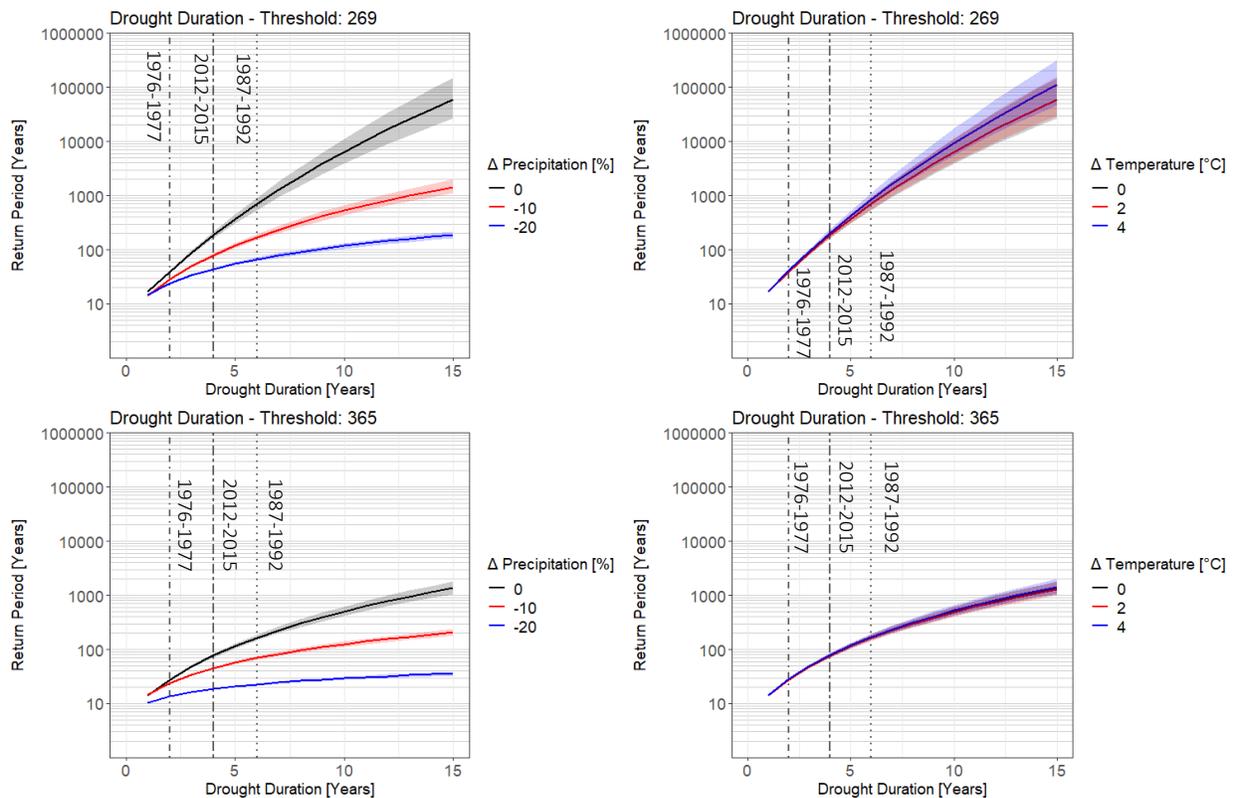
Figure 6-14. Effect of precipitation (left) and temperature (right) change on the drought duration frequency. Results are shown with 269 TAF (top) and 365 TAF (bottom) drought threshold. The x-axis shows that the drought duration

(years) and the y-axis shows the estimated return period (years). The 95% confidence intervals are shown in shaded areas and are obtained using the bootstrap method. Vertical lines shows the duration of the historical droughts obtained using the historical dataset.

Table 6-13. Effect of precipitation and temperature change on the return periods associated with the duration of the historic droughts.

Threshold [TAF]	Drought Event	Changes in Precipitation			Changes in Temperature [°C]		
		0%	-10%	-20%	0	+2	+4
	1976-1977	40	30	25	40	40	40
269	1987-1992	700	170	65	700	700	840
	2012-2015	185	80	45	185	185	205
	1976-1977	30	25	15	30	30	30
365	1987-1992	165	70	25	165	160	170
	2012-2015	75	45	20	75	75	80

Like drought severity, precipitation is the main driver of change in drought duration. Any decrease in precipitation makes droughts of any duration more frequent, especially the long ones. For instance, with 269 TAF threshold, the return period of 2-yr long drought (i.e., 1976-1977) drops from 40 years to 28 years or 24 years if precipitation decreases by either 10 or 20%, respectively. For long droughts, for instance of the 1987-1992 (i.e., 6 years), the return period drops from 700 years to 170 or 70 years with a reduction in precipitation of either 10 or 20%, respectively.

Warming temperature tends to slightly reduce the frequency of droughts with specific duration, although the signal is hardly significant. For instance, using 269 TAF threshold and a warming scenario of +4°C, the return period of the 1987-92 drought duration (i.e., 6 years) increases from 700 years to 850 years.

Similar results are obtained for the 365 TAF threshold: a major increase in frequency when precipitation decreases, and an insignificant signal under warming conditions.

6.5.3. Changes in Joint Return Period

The modification of the frequency of drought when accounting for the dependence between severity and duration was conducted using copula functions. The survival Clayton copula was used with the 269 TAF threshold while the Gumbel copula was used for the 365 TAF threshold. Note that the return periods obtained under baseline climate (i.e., no change in precipitation and temperature) are slightly different from the ones presented in Table 6-11 as the 'combined' dataset was used in section 6.4 rather than the 'stochastic realizations' dataset'.

Table 6-14 details for the considered climate scenarios and drought threshold the modification of the dependence (i.e., correlation) between drought severity and duration by mean of the Kendall's rank correlation coefficient (τ^{data}). It is noted that decreasing precipitation increases the dependence between

drought severity and duration (i.e. τ^{data} increases) while increasing precipitation tends to decrease the dependence between drought severity and duration (i.e., τ^{data} decreases). Warming temperature tends to slightly decrease the correlation, although the changes are not as significant as for precipitation changes. Although the fitted copula functions for each climate scenarios and thresholds reproduce the Kendall's rank correlation well, they consistently underestimate the correlation (τ^{cop} compared to τ^{data}).

Table 6-14. Effect of precipitation and temperature change on the dependence between drought severity and duration. The Kendall's rank correlation coefficients obtained from the fitted copula functions ($\boldsymbol{\tau^{cop}}$) and estimated from the identified drought events ($\boldsymbol{\tau^{data}}$) are used to assess the dependence between drought severity and duration. The fitted copula parameter is given for information only.

Threshold	Precipitation Scenario	Temperature Scenario	Copula	τ^{cop}	τ^{data}	Copula Parameter
269	ΔP=-20%	ΔT=0°C	Survival Clayton	0.674	0.719	2.272
	ΔP=-20%	ΔT=4°C		0.685	0.726	2.387
	ΔP=0%	ΔT=0°C		0.464	0.567	1.732
	ΔP=0%	ΔT=4°C		0.452	0.557	1.651
	ΔP=+20%	ΔT=0°C		0.366	0.451	1.738
	ΔP=+20%	ΔT=4°C		0.334	0.420	1.005
365	ΔP=-20%	ΔT=0°C	Gumbel	0.758	0.782	2.93
	ΔP=-20%	ΔT=4°C		0.754	0.772	2.64
	ΔP=0%	ΔT=0°C		0.558	0.632	2.26
	ΔP=0%	ΔT=4°C		0.552	0.625	2.23
	ΔP=+20%	ΔT=0°C		0.421	0.522	2.10
	ΔP=+20%	ΔT=4°C		0.399	0.498	1.81

The effect of precipitation and temperature change on the joint return period of drought severity and duration is presented in Table 6-15 and shown as a plot in Figure 6-15. The results are consistent with the results discussed for drought severity and duration separately. As was expected, increasing precipitation decreases the frequency of drought significantly (i.e., increases the return period). With a 269 TAF threshold and under 20% increase precipitation, the reduction in frequency is such that the return periods associated with the considered historical droughts become all roughly larger than 20,000 years. However, decreasing precipitation leads to significant increase in drought frequency. Under 20% precipitation reduction scenario, the return period of all historical droughts gets below 100 years.

The effect on temperature change on the drought frequency appears to not be linear regardless of it's associated to either a decrease or an increase in precipitation. When combined with a decrease in precipitation, the effect of temperature is not significant.

The main results remain broadly consistent when using 365 TAF threshold. Note that the droughts are more frequent using this threshold and the effect of temperature change remain low when compared with the increase in frequency following from a reduction in precipitation.
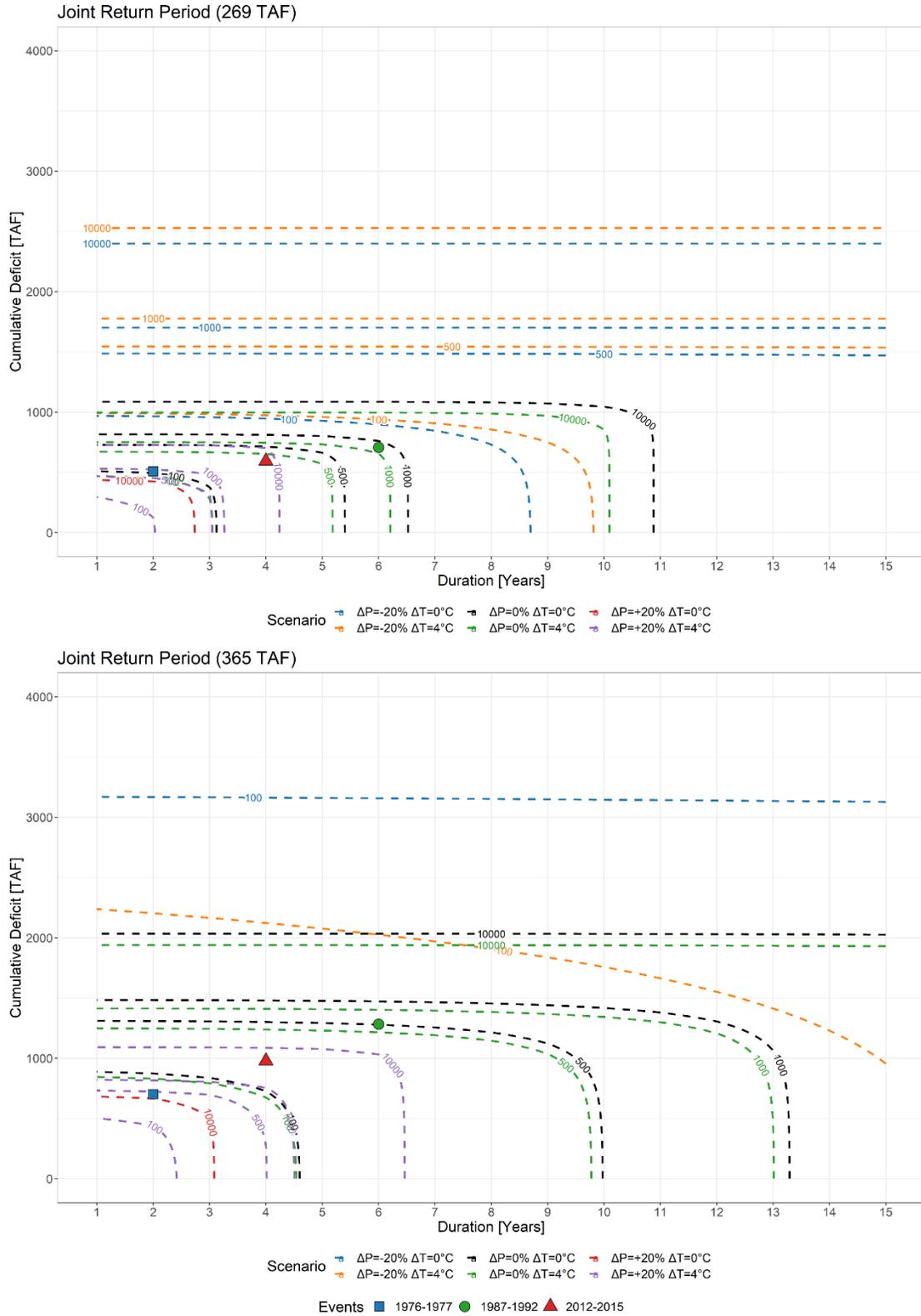
Figure 6-15. Joint return period contours for changes in precipitation and temperature for thresholds 269 TAF and 365 TAF. Contour colors are different climate scenarios with the historic drought events shown as points.

Table 6-15. Effect of precipitation change (ΔP, %) and temperature change (ΔT, °C) on the joint return period (years) associated with the severity and duration of the historic droughts for drought thresholds of 269 TAF and 365 TAF.

| Threshold | Event | Joint Return Period (years) | | | | | |
| | | ΔP=-20% | | ΔP=0% | | ΔP=+20% | |
		ΔT=0°C	ΔT=4°C	ΔT=0°C	ΔT=4°C	ΔT=0°C	ΔT=4°C
	1976-1977	35	35	110	140	19,300	855
269	1987-1992	75	70	855	1,160	728,510	1,532,685
	2012-2015	50	50	270	355	61,685	4,655
	1976-1977	15	45	60	65	12,075	425
365	1987-1992	25	75	505	635	702,410	62,225
	2012-2015	20	60	170	195	78,515	4,085

In summary, models indicate that hydrologic drought severity on the Tuolumne River for SFPUC will significantly increase with reduction in precipitation but not necessarily with temperature increase. Note that the models used (combination of weather generator and hydrologic model) tend to underestimate both the frequency of occurrence and severity of droughts, as compared to the observed record (not enough droughts occur in the simulated record compared to the historical record). Therefore, it is assumed that the frequency of occurrence and severity of hydrologic droughts in the simulated record, with climate change, is also underestimated, overstating the RWS water supply reliability.

References

Anderson, M. G., & McDonnell, J. J. (Eds.). (2005). Sacramento Soil Moisture Accounting Model (SAC-SMA). In Encyclopedia of Hydrological Sciences. Chichester, UK: John Wiley & Sons, Ltd. Retrieved from http://doi.wiley.com/10.1002/0470848944.hsa279

Anderson, E. (2006). Snow Accumulation and Ablation Model - SNOW-17 (p. 61). NOAA. Retrieved from http://www.nws.noaa.gov/oh/hrl/nwsrfs/users_manual/part2/_pdf/22snow17.pdf

Gupta, H. V., Kling, H., Yilmaz, K. K. & Martinez, G. F., 2009. Decomposition of the mean squared error and NSE performance criteria:Implications for improving hydrological modelling. Journal of Hydrology, 377(1-2), pp. 80-91.

HRG TR1 (2018). Technical Report 1: Weather Generator Module. Hydrosystems Research Group, University of Massachusetts, Amherst, Amherst, Massachusetts.

Jain, A. & Srinivasulu, S., 2009. Hydrologic Model Calibration Using Evolutionary Optimisation. In: R. J. Abrahart, L. M. Lee & D. P. Solomatine, eds. *Practical Hydroinformatics.* Berlin: Springer, Berlin, Heidelberg.

Marini, F. & Walczak, B., 2015. Particle swarm optimization (PSO). A tutorial. *Chemometrics and Intelligent Laboratory Systems,* pp. 153-165.

Markstrom, S. L. et al., 2015. PRMS-IV, the Precipitation-Runoff Modeling System, Version 4: U.S. Geological Survey Techniques and Methods, book 6, chap. B7. In: Reston: U.S. Geological Survey.

Moriasi, D. N. et al., 2006. Model evaluation guidelines for systematic quantification of accuracy in watershed simulations. Transactions of the ASABE, 50(3), pp. 885-900.

Newman, A., Clark, M. P., Sampson, K., Wood, A., Hay, L. E., Bock, A., Viger, R. J., Blodgett, D., Brekke, L., Arnold, J. R., Hopson, T., & Duan, Q. (2015). *Development of a large-sample watershed-scale hydrometeorological data set for the contiguous USA : data set characteristics and assessment of regional variability in hydrologic model performance*. 209–223. https://doi.org/10.5194/hess-19-209-2015

NOAA, 2002. *Conceptualization of the Sacramento Soil Moisture Accounting Model,* Silver Sping: National Oceanic and Atmospheric Administration.

Poli, R., Kennedy, J. & Blackwell, T., 2007. Particle swarm optimization: An overview. *Swarm Intelligence,* pp. 35-57.

Livneh, B., Bohn, T.J., Pierce, D.W., Munoz-Arriola, F., Nijssen, B., Vose, R., Cayan, D.R., Brekke, L., 2015. A spatially comprehensive, hydrometeorological data set for Mexico, the U.S., and Southern Canada 1950–2013. Sci. Data 2, 150042. https://doi.org/10.1038/sdata.2015.42

Lohmann, D., R. Raschke, B. Nijssen, and D. P. Lettenmaier (1998), Regional scale hydrology: I. Formulation of the VIC-2L model coupled to a routing model, Hydrolog. Sci. J., 43, 131-141.

Lu, J., Sun, G., McNulty, S.G., Amatya, D.M., 2005. A Comparison of Six Potential Evapotranspiration Methods for Regional Use in the Southeastern United States1. JAWRA J. Am. Water Resour. Assoc. 41, 621–633. https://doi.org/10.1111/j.1752-1688.2005.tb03759.x

McCall, J., 2005. Genetic algorithms for modelling and optimisation. J. Comput. Appl. Math. 184, 205–222. https://doi.org/10.1016/j.cam.2004.07.034

San Francisco Planning Department, 2005. San Francisco Peninsula Streams and Reservoirs [WWW Document]. URL http://sf-planning.org/sites/default/files/FileCenter/Documents/7999-2005.0159E_vol3_sec5-5_wsip_finalpeir.pdf (accessed 8.21.18).

San Francisco Planning Department, 2001. Peninsula Watershed Management Plan [WWW Document]. URL https://www.sfwater.org/Modules/ShowDocument.aspx?documentID=4343 (accessed 8.21.18).

The United States Geological Survey, 2005. The San Andreas and San Gregorio Fault Systems in San Mateo County [WWW Document]. URL https://pubs.usgs.gov/of/2005/1127/chapter8.pdf (accessed 8.21.18).

World Bank, 2002. Conceptualization of the Sacramento soil moisture accounting model [WWW Document]. URL http://www.nws.noaa.gov/oh/hrl/nwsrfs/users_manual/part2/_pdf/23sacsma.pdf?referrer=Baker (accessed 8.23.18).

A. Appendix – Parameters for Climate Data Interpolation

Each temperature station was analyzed to develop regressions between stations using the data from 1999-10-01 through 2016-10-01. The relationships were used to fill the missing data in the period between 1969 and the installation of the station. The relationships were established by first analyzing the correlation between the stations. The station with the highest correlation value is then selected and then a basic regression is used to estimate maximum and minimum temperature at the gage. i is the selected station with target station s and is calculated based on equation A-1. This task was completed by HHWP team.

$$T_s(t) = a_{i,s}T_i(t) + b_{i,s},$$
A-1

A.1. Maximum and Minimum Temperature Stations Parameters

Table A-1 Maximum temperature coefficients between stations

	Hetchy	Buck	Tuolumne	Cherry	Moccasin	Paradise	Horse	Slide	Pinecrest
Hetchy	1.000	0.894	0.853	0.959	0.923	0.813	0.850	0.846	0.924
Buck	0.894	1.000	0.827	0.922	0.884	0.838	0.849	0.826	0.887
Tuolumne	0.853	0.827	1.000	0.894	0.818	0.896	0.931	0.934	0.891
Cherry	0.959	0.922	0.894	1.000	0.932	0.868	0.898	0.892	0.957
Moccasin	0.923	0.884	0.818	0.932	1.000	0.790	0.824	0.810	0.891
Paradise	0.813	0.838	0.896	0.868	0.790	1.000	0.933	0.972	0.872
Horse	0.850	0.849	0.931	0.898	0.824	0.933	1.000	0.965	0.902
Slide	0.846	0.826	0.934	0.892	0.810	0.972	0.965	1.000	0.905
Pinecrest	0.924	0.887	0.891	0.957	0.891	0.872	0.902	0.905	1.000

Table A-2 Relationship between stations (slope and intercept). Dependent variable is in column heading. For example: $Hetchy = 0.924 \times Buck + 1.045$

(a) Slope

	Hetchy	Buck	Tuolumne	Cherry	Moccasin	Paradise	Horse	Slide	Pinecrest
Hetchy	1.000	0.894	0.853	0.959	0.923	0.813	0.850	0.846	0.924
Buck	0.894	1.000	0.827	0.922	0.884	0.838	0.849	0.826	0.887
Tuolumne	0.853	0.827	1.000	0.894	0.818	0.896	0.931	0.934	0.891
Cherry	0.959	0.922	0.894	1.000	0.932	0.868	0.898	0.892	0.957
Moccasin	0.923	0.884	0.818	0.932	1.000	0.790	0.824	0.810	0.891

Paradise	0.813	0.838	0.896	0.868	0.790	1.000	0.933	0.972	0.872
Horse	0.850	0.849	0.931	0.898	0.824	0.933	1.000	0.965	0.902
Slide	0.846	0.826	0.934	0.892	0.810	0.972	0.965	1.000	0.905
Pinecrest	0.924	0.887	0.891	0.957	0.891	0.872	0.902	0.905	1.000

(b) Intercept

	Hetchy	Buck	Tuolumne	Cherry	Moccasin	Paradise	Horse	Slide	Pinecrest
Hetchy	0.000	6.603	0.703	-3.044	10.342	-2.031	-3.542	-4.781	0.297
Buck	1.045	0.000	-0.722	-4.560	9.036	-5.079	-5.741	-6.752	-1.081
Tuolumne	9.125	13.119	0.000	3.734	18.540	-0.999	-2.400	-3.703	5.558
Cherry	5.558	9.788	2.913	0.000	14.409	-0.237	-1.535	-3.254	3.316
Moccasin	-4.213	0.063	-3.859	-9.439	0.000	-6.877	-8.299	-10.367	-5.363
Paradise	14.582	16.923	6.588	9.069	23.798	0.000	1.256	-4.963	10.397
Horse	13.865	17.059	6.123	8.639	23.160	2.238	0.000	-0.720	9.953
Slide	15.337	19.947	7.130	10.953	26.106	6.277	2.458	0.000	11.874
Pinecrest	4.819	9.080	1.391	-0.574	13.888	-1.954	-3.259	-5.265	0.000

Table A-3 Minimum temperature coefficients between stations

	Hetchy	Buck	Tuolumne	Cherry	Moccasin	Paradise	Horse	Slide	Pinecrest
Hetchy	1.000	0.860	0.786	0.910	0.870	0.804	0.795	0.810	0.886
Buck	0.860	1.000	0.778	0.853	0.828	0.789	0.785	0.784	0.832
Tuolumne	0.786	0.778	1.000	0.807	0.720	0.884	0.905	0.927	0.792
Cherry	0.910	0.853	0.807	1.000	0.809	0.848	0.822	0.854	0.859
Moccasin	0.870	0.828	0.720	0.809	1.000	0.723	0.740	0.717	0.851
Paradise	0.804	0.789	0.884	0.848	0.723	1.000	0.916	0.966	0.807
Horse	0.795	0.785	0.905	0.822	0.740	0.916	1.000	0.954	0.800
Slide	0.810	0.784	0.927	0.854	0.717	0.966	0.954	1.000	0.799
Pinecrest	0.886	0.832	0.792	0.859	0.851	0.807	0.800	0.799	1.000

Table A-4 Relationship between stations (slope and intercept). Dependent variable is in column heading. For example: $Hetchy = 0.957 \times Buck + 2.675$

(a) Slope

	Hetchy	Buck	Tuolumne	Cherry	Moccasin	Paradise	Horse	Slide	Pinecrest
Hetchy	1.000	0.898	0.956	0.970	0.925	1.010	1.009	0.976	0.834
Buck	0.957	1.000	0.983	0.969	0.932	1.033	1.036	0.981	0.834
Tuolumne	0.822	0.792	1.000	0.847	0.780	0.982	0.998	0.990	0.731
Cherry	0.938	0.880	0.953	1.000	0.877	1.020	1.010	0.993	0.807
Moccasin	0.941	0.889	0.923	0.922	1.000	0.966	0.982	0.934	0.824
Paradise	0.796	0.764	0.900	0.831	0.749	1.000	0.962	0.981	0.707
Horse	0.787	0.758	0.907	0.815	0.754	0.953	1.000	0.961	0.700
Slide	0.830	0.799	0.936	0.860	0.768	0.985	0.992	1.000	0.734
Pinecrest	1.062	0.997	1.084	1.064	1.033	1.143	1.143	1.089	1.000

(b) Intercept

	Hetchy	Buck	Tuolumne	Cherry	Moccasin	Paradise	Horse	Slide	Pinecrest
Hetchy	0.000	3.610	-20.074	0.295	5.438	-17.531	-19.808	-18.838	-0.494
Buck	2.675	0.000	-20.387	1.131	5.934	-17.657	-20.096	-18.381	0.188
Tuolumne	25.843	25.646	0.000	24.281	28.923	5.294	2.598	2.301	20.067
Cherry	3.671	5.318	-18.945	0.000	8.429	-16.919	-18.798	-18.879	1.486
Moccasin	0.555	2.088	-20.605	0.375	0.000	-17.690	-20.747	-18.837	-1.857
Paradise	22.504	22.521	-2.245	20.546	25.934	0.000	-1.323	-4.257	17.131
Horse	24.574	24.440	-0.297	22.891	27.564	3.484	0.000	0.388	18.944
Slide	24.000	24.061	-0.536	22.562	27.474	5.176	0.756	0.000	18.866
Pinecrest	5.514	7.010	-17.224	4.443	8.747	-14.447	-16.788	-15.680	0.000

B. PRMS Modules for Upcountry watershed

The summary of PRMS modules selected for representing various hydrologic processes of the Upcountry watershed systems is provided in Table B-1. The modules highlighted in bold are ones selected over other available modules, details of which are further discussed in the following sections.

Table B-1 Summary of hydrologic processes and modules for the Upcountry hydrologic models (modules in bold are the selected modules over other available modules)

Hydrologic Process	Used Module	Description
Basin Definition Process	basin	Defines watershed-wide and HRU-based physical parameters and variables
Cascading Flow Process	cascade	Determines computational order of HRUs for routing flow.
Solar Table Process	soltab	Computes potential solar radiation and sunlight hours for each HRUs for each day of the year
Time Series Data Process	obs	Reads and stores observed data from all specified measurement stations
Combined Climate Distribution Process	**xyz_dist**	Distributes precipitation and temperature to each HRU by using a multiple linear regression of measured data
Solar Radiation Distribution Process	**ddsolrad**	Distributed solar radiation to each HRU and estimates missing solar radiation data using maximum temperature per degree-day relations
Potential Evapotranspiration Process	**potet_jh**	Computes the potential evapotranspiration by using the Jensen-Haise formulation
Canopy Interception Process	intcp	Computes volume of intercepted precipitation, evaporation from intercepted precipitation, and throughfall that reaches the soil or snowpack
Snow Process	snowcomp	Simulates snowpack and snow accumulation and depletion process by using an energy-budget approach
Surface Runoff Process	**srunoff_smidx**	Computes surface runoff and infiltration for each HRU by using a nonlinear variable-source-area method allowing for cascading

		flow
Soil-zone Process	soilzone	Computes inflow and outflows from soil zone of each HRU including infiltration, groundwater, and upslope HRUs, outflows to gravity drainage, interflow, and surface runoff to down-slope HRUs
Groundwater Process	gwflow	Sums inflow and outflow from PRMS groundwater reservoirs
Streamflow Process	**muskingum**	Routes water between segments in the system using Muskingum routing

B.1. Combined Climate Distribution Module (xyz_dist)

The xyz_dist module uses a three-dimensional, multiple-linear regression based on longitude, latitude, and elevation to distribute temperature and precipitation data from two or more stations. It was initially developed as a method to statistically downscale precipitation and temperature data from atmospheric models for each HRU.

The module develops multiple linear-regression relationships for each climate variable based on stations, grid cell elevation, longitude, and latitude, to calculate precipitation, maximum air temperature, and minimum air temperature. A total of three parameters (ppt_lapse, max_lapse, and min_lapse, for precipitation, maximum, and minimum air temperature respectively) that are distributed by three dimensions (latitude, longitude, and elevation) and monthly mean values from the climate stations located near the watershed (Jan through Dec) are used in the climate distribution module to distribute the climate variables across all HRUs. The general equation below describes a plane in three dimensional space with multiple linear regression parameters that intersects the climate variable (CV) axis at b_0.

$$CV = (lapse_x \times x_{sta}) + (lapse_y \times y_{sta}) + (lapse_z \times z_{sta}) + b_0$$

Where:

$lapse_x$ = Value of ppt_lapse$_{x_{month}}$/max_lapse$_{x_{month}}$/min_lapse$_{x_{month}}$

$lapse_y$ = Value of ppt_lapse$_{y_{month}}$/max_lapse$_{y_{month}}$/min_lapse$_{y_{month}}$

$lapse_z$ = Value of ppt_lapse$_{z_{month}}$/max_lapse$_{z_{month}}$/min_lapse$_{z_{month}}$

The climate variable equation is calculated from a set of stations specified by the psta_nuse or tsta_nuse and a different set of stations can be used for each day. If no stations have valid values on a particular day, then the mean monthly values are used. b_0 is calculated according to the following equation and climate variable (for equation example, precipitation is shown but this is the same for temperature).

$$b_0 = ppt\_mean - \left(ppt\_lapse_{x,month} \times x_{mean}\right) - \left(ppt\_lapse_{y,month} \times y_{mean}\right)$$
$$- \left(ppt\_lapse_{z,month} \times z_{mean}\right)$$

Then based on the climate variable, the values can be distributed for each HRU by:

$$hru_{ppt_{HRU}} = \left(ppt\_lapse_{x,month} \times hru\_x_{HRU}\right) + \left(ppt\_lapse_{y,month} \times hru\_y_{HRU}\right)$$
$$+ \left(ppt\_lapse_{z,month} \times hru\_z_{HRU}\right) + b_0$$

The dependent and independent variables in the regression are then transformed by subtracting the mean (ppt_add, x_add, y_add, and z_add) and dividing by the standard deviation (ppt_div, x_div, y_div, z_div) to reduce the effects of units, magnitude, and inconsistency of the study area.

B.2. Solar Radiation Distribution Module (ddsolrad)

The ddsolrad module computes shortwave solar radiation with a modified degree-day method. The ratio of the actual-to-potential radiation for horizontal surface is obtained based on the relationship between the monthly maximum air temperature and the degree-day coefficient. The short wave radiation for each HRU is calculated by:

$$swrad_{HRU} = \frac{solf_{HRU} \times soltab\_potsw_{HRU}}{\cos(\arctan(hru\_slope_{HRU}))}$$

For days with precipitation greater than the monthly parameter value ppt_rad_adj, sward is adjusted according to the following equation:

$$swrad_{HRU} = swrad_{HRU} \times rad\_adj$$

Where:

rad_adj is $radj\_sppt$ for summer days and $radj\_wppt$ for winter.

B.3. Potential Evapotranspiration Module (potet_jh)

The potet_jh module uses the modified Jensen-Haise formulation to compute the potential evapotranspiration for each HRU. This is computed as a function of air temperature, solar radiation, and the coefficients jh_coef and jh_coef_hru, which can be estimated by using regional air temperature, elevation, and saturation vapor pressure. This is calculated per HRU by:

$$potet_{HRU} = jh\_coef_{month} \times (tavgf_{HRU} - jh\_coef\_hru_{HRU}) \times \frac{swrad_{HRU}}{2.54 \times \lambda_{HRU}}$$

$$\lambda_{HRU} = 597.3 - (0.5653 \times tavgf_{HRU})$$

Where:

λ_{HRU} is the latent heat of vaporization on the HRU.

The air temperature parameter (jh_coef_hru) used in Jensen-Haise can be estimated for each HRU:

$$jh\_coef\_hru_{HRU} = 27.5 - \left[0.25 \times \left(\rho_{high\;temp} - \rho_{low\;temp}\right)\right] - \frac{hru\_elev_{HRU}}{1000}$$

Where:

$\rho_{high\;temp}$ is the saturated vapor pressure for the mean maximum air temperature for the warmest month of the year in millibars.

$\rho_{low\;temp}$ is the saturated vapor pressure for the mean minimum air temperature for the warmest month of the year in millibars.

B.4. Surface Runoff Module (srunoff_smidx)

The module srunoff_smidx computes the surface runoff from infiltration excess and soil saturation by using a non-linear, variable-source-area method, where the runoff generating area of the watershed surface varies in location and size over time. The module configured for the Up Country region has two computational sections, which is broken down into impervious storage and evaporation section and Hortonian surface runoff and infiltration section.

- *Impervious Storage and Evapotranspiration*

When the sum of the rain throughfall, snowmelt, and the antecedent impervious storage (avail_water) exceeds the retention storage capacity on the impervious portion of an HRU for a time step, the impervious Hortonian surface runoff is calculated. Water beyond the impervious storage capacity (imperv_stor_max) is retained until it is evaporated and the Hortonian surface runoff from the impervious portion of an HRU (hru_sroffi) for each time step is calculated according to:

$$avail\_water = imperv\_stor_{HRU}^{t-1} + net\_rain_{HRU} + snowmelt_{HRU}$$

If $avail\_water > imperv\_stor\_max_{HRU}$ is true, t hen the surface runoff for a single HRU is calculated.

$$hru\_sroffi_{HRU} = (avail\_water - imperv\_stor\_max_{HRU}) \times hru\_percent\_imperv_{HRU}$$

However, if the condition above is not true,

$$hru\_sroffi_{HRU} = 0$$

Evaporation for the impervious section of the HRU is based on the available water and unsatisfied potential evapotranspiration volumes. They're calculated by:

$$avail\_water = imperv\_stor_{HRU}^{t-1} + net\_rain_{HRU} + snowmelt_{HRU} - \frac{hru\_sroffi_{HRU}}{hru\_percent\_imperv_{HRU}}$$

$$avail\_et = potet_{HRU} - snow\_evap_{HRU} - hru\_intcpevap_{HRU} - dprst\_evap\_hru_{HRU}$$

If $avail\_et \geq avail\_water$, the evaporation from the impervious portion of the HRU is computed by:

$$hru\_impervevap_{HRU} = avail\_water \times (1 - snowcov\_area_{HRU}) \times hru\_percent\_imperv_{HRU}$$

If $avail\_et < avail\_water$, then the evaporation from the impervious portion for an HRU for the time step is calculated by:

$$hru\_impervevap_{HRU} = avail\_et \times (1 - snowcov\_area_{HRU}) \times hru\_percent\_imperv_{HRU}$$

Storage on the impervious portion of the HRU is calculated by:

$$hru\_impervstor_{HRU} = hru\_impervstor_{HRU}^{t-1} - hru\_sroffi_{HRU} - hru\_impervevap_{HRU} +$$
$$(net\_rain_{HRU} + snowmelt_{HRU}) \times hru\_percent\_imperv_{HRU}$$

- *Pervious Hortonian Surface Runoff and Infiltration*

The infiltration excess of each HRU happens when the throughfall and snowmelt for the infiltration exceeds the capacity of the soil. Therefore, the Hortonian surface runoff of an HRU is calculated by:

$$hru\_sroffp_{HRU} = ca\_fraction \times (upslope\_hortonian_{HRU} + net\_rain_{HRU} + snowmelt_{HRU})$$

where:

$cafraction$ is the fractional variable-source area for the previous portion of an HRU

For the non-linear surface runoff model, the antecedent soil-moisture content is computed of the capillary reservoir by:

$$avail\_et = potet_{HRU} - snow\_evap_{HRU} - hru\_intcpevap_{HRU} - dprst\_evap\_hru_{HRU}$$

$$ca\_fraction = smidx\_coef_{HRU} \times (10)^{smidx\_exp_{HRU} \times smidx}$$

If $ca\_fraction > carea\_max_{HRU}$, then $ca\_fraction = carea\_max_{HRU}$.

In situations where no snowpack exists, infiltration to the area is calculated as:

$$infil_{HRU} = (upslope\_hortonian_{HRU} + net\_rain_{HRU} + snowmelt_{HRU} - hru\_sroffi_{HRU}$$
$$- hru\_sroffp_{HRU}) \times (1 - hru\_percent\_imperv_{HRU})$$

In situations where there is a snowpack, surface runoff and infiltration of the previous portion of the HRU are adjusted on the basis of the parameter snowinfil_max and the capillary reservoir is represented as:

$$capacity = soil\_moist\_max_{HRU} - soil\_moist_{HRU}^{t-1}$$

$$avail\_water = upslope\_hortonian_{HRU} + snowmelt_{HRU}$$

$$excess = avail\_water - capacity$$

Any additional surface runoff is computed as:

$$excess\_infil = \max(0, avail\_water - capacity - snowinfil\_max_{HRU})$$

If the *excess_infil* value is greater than 0, the value is also added to hru_sroffp and the amount of water that infiltrated into the capillary reservoir for the HRU is defined by:

$$infil_{HRU} = snowinfil\_max_{HRU} + capacity_{HRU}$$

The total Hortonian surface runoff for each HRU is then calculated as:

$$hortonian\_flow_{HRU} = hru\_sroffp_{HRU} + hru\_sroffi_{HRU}$$

B.5. Streamflow Module (Muskingum)

The Muskingum routing module uses stream networks as a conceptualized single-direction sequence of connected stream segments as specified by the parameter tosegment. A stream segment is associated with each one-plane HRU or the left and right bank HRUs as specified by hru_segment parameter. The Muskingum routing equation assumes a linear relationship between storage and the segment inflow characteristics (seg_inflow) and segment outflow (seg_outflow). Storage in a stream segment for the internal time step is calculated by:

$$storage_{segment}^{t} = k\_coef_{segment}$$
$$\times \left((x\_coef_{segment} \times seg\_inflow_{segment}^{t}) + (1 - x\_coef_{segment}) \times seg\_outflow_{segment}^{t} \right)$$

The routing function assumes the average flow during the internal time step is equal to the average flow at the start and the end times. Therefore, the continuity equation is then expressed as:

$$\Delta storage_{segment}^{t} = storage_{segment}^{t} - storage_{segment}^{t-1}$$

$$\Delta storage_{segment}^{t}$$
$$= \left(\frac{seg\_inflow_{segment}^{t} + seg\_inflow_{segment}^{t-1}}{2} \right) \times \Delta t_{segment}$$
$$- \left(\frac{seg\_outflow_{segment}^{t} + seg\_outflow_{segment}^{t-1}}{2} \right) \times \Delta t_{segment}$$

The segment outflow is then solved for the internal time step by:

$$segoutflow_{segment}^{t}$$
$$= \left(c0_{segment} \times seg\_inflow_{segment}^{t} \right) + \left(c1_{segment} \times seg\_inflow_{segment}^{t-1} \right)$$
$$+ \left(c2_{segment} \times seg\_outflow_{segment}^{t-1} \right)$$

Where:

$$c0_{segment} = \frac{-(k\_coef_{segment} \times x\_coef_{segment}) + \frac{\Delta t_{segment}}{2}}{(k\_coef_{segment}) - (k\_coef_{segment} \times x\_coef_{segment}) + \frac{\Delta t_{segment}}{2}}$$

$$c1_{segment} = \frac{\left(k\_coef_{segment} \times x\_coef_{segment}\right) + \frac{\Delta t_{segment}}{2}}{\left(k\_coef_{segment}\right) - \left(k\_coef_{segment} \times x\_coef_{segment}\right) + \frac{\Delta t_{segment}}{2}}$$

$$c2_{segment} = \frac{k\_coef_{segment} - \left(k\_coef_{segment} \times x\_coef_{segment}\right) - \frac{\Delta t_{segment}}{2}}{k\_coef_{segment} - \left(k\_coef_{segment} \times x\_coef_{segment}\right) + \frac{\Delta t_{segment}}{2}}$$

The internal time step (Δt) is calculated for each stream segment according to:

$$\Delta t_{segment} = \left\lceil \frac{24}{\left\lceil \frac{24}{k\_coef_{segment}} \right\rceil} \right\rceil$$

The travel time (in hours) is rounded down to an even divisor of 24 hours due to the PRMS system being restricted to daily time steps. This also means the travel time can never be greater than 24 hours.

C. Appendix – Results of Hetch Hetchy and Cherry-Eleanor PRMS Hydrology Model Calibrations

The implementation of PRMS used by HHWP has 129 parameters, controlling snow accumulation and melt, evapotranspiration, water storage and flux, and streamflow routing. Of these parameters, 56 are distributed amongst the HRUs, 6 characterize the temperature sensors, 6 characterize the precipitation gauges, 36 are universal across the model domain, 17 vary monthly, 7 are distributed by subbasin or river segment, and 1 characterizes the snow depletion curves. A number of data sources and methods were used to determine parameter values in PRMS.

1) A **Digital Elevation Model (DEM)** of the basin draining into the Hetch Hetchy Reservoir was used to delineate the HRUs based on surface topography. Once the HRUs were delineated, HRU area, and HRU average slope, aspect, latitude and longitude were determined from analysis of the DEM.
2) **Real time measurements** were used to determine some parameters, including earliest and latest dates of snow melt initiation (melt_look and melt_force) and temperature and precipitation lapse rates.
3) **ASO snow on and snow off acquisitions** were used to determine a set of parameters, including the vegetation cover types, summer and winter cover densities, and the percent impervious of each HRU.
4) The **PRMS calibration of the Merced Basin**, the closest basin to the south of the Tuolumne was used to determine some parameters. These parameters were assumed to be the same in the Tuolumne as in the Merced.

5) Some parameters were assumed to match the **default values**. These were parameters that either did not greatly affect flows or snow pack (as determined in the sensitivity analysis above), or were difficult to determine / distribute / calibrate.

6) **Calibration to snow pillow and snow survey data** was used to determine parameters that were not distributed, and controlled snow accumulation and melt.

7) **Calibration to ASO** snow data from 20113-2015 was used to determine distributed parameters controlling snow accumulation and melt.

8) **Calibration to annual inflows into Hetchy Hetchy Reservoir** was used to determine parameters that affect seasonal and annual water balance.

9) **Calibration to storm to seasonal streamflow patterns** was used to determine the final set of parameters.

C.4. Parameters Derived from DEMs

18 parameters were directly derived from Lidar based DEMs of the Tuolumne Basin (**Error! Reference source not found.**). All parameters denoted hru_XXX were derived using ArcGIS and the GISWeasel by Bruce McGurk. The parameter jh_coef_hru was calculated as a function of HRU elevation, using the formula presented in the PRMS manual (Equation 1-52). The snarea_thresh parameter was derived as a function of elevation, using the relationship seen at the Merced watershed. The XXX_div and XXX_add parameters were determined using the hru_x, hru_y and hru_z parameters.

Table C-1: Parameters directly derived from DEMs

Parameter	Dimension	Units	Value Range
x_add	one	meters	-5,801.1
x_div	one	meters	2,847.5
y_add	one	meters	-2E+06
y_div	one	meters	14,608.6
z_add	one	meters	2,048,673
z_div	one	meters	27,604.6
tosegment	nsegment	none	0 to 16
subbasin_down	nsub	none	0 to 16
hru_area	nhru	acres	276 to 3,446
hru_aspect	nhru	Cardinal direction	21 to 299
hru_elev	nhru	feet	3,778 to 11,458
hru_lat	nhru	Degrees	37.748 to 38.165
hru_segment	nhru	none	1 to 16
hru_slope	nhru	feet/feet	0.02 to 1.01
hru_subbasin	nhru	none	1 to 16
hru_type	nhru	none	0 to 1
hru_x	nhru	meters	-2,048,719 to-2,002,321
hru_y	nhru	meters	1,881,827 to 1,934,207
jh_coef_hru	nhru	1/F	9.2 to 16.88

C.5. Parameters Derived from Real Time Measurements

20 parameters were derived from real time measurement stations (Table C-2), including temperature probes, precipitation gauges and snow pillows. These include the locations of the temperature and precipitation sensors, which were taken from the cdec.gov website.

Table C-2: Parameters directly derived from Real Time Measurements

Parameter	Dimension	Units	Value Range
den_init	one	grams/cm^3	0.20
den_max	one	grams/cm^3	0.58
ppt_div	one	inches	0.35316
ppt_add	one	inches	-0.1013
tmax_div	one	F	10.7361
tmax_add	one	F	-60.639
tmin_div	one	F	9.99494
tmin_add	one	F	-33.315
max_lapse	nmonths, nlapse	none	-1.4589 to -0.9620
min_lapse	nmonths, nlapse	none	-1.4364 to -0.8836
ppt_lapse	nmonths, nlapse	none	0.232 to 0.451
tsta_elev	ntemp	feet	938 to 9,200
tsta_month_max	ntemp	F	36.16 to 100.53
tsta_month_min	ntemp	F	9.26 to 62.23
tsta_x	ntemp	meters	-2,095,987 to -2,013,855
tsta_y	ntemp	meters	1,897,611 to 1,946,363
psta_elev	nrain	feet	938 to 8,600
psta_month_ppt	nrain	inches	0
psta_x	nrain	meters	-2,095,987 to -2,013,855
psta_y	nrain	meters	1,884,654 to 1,946,363
melt_force	nhru	day	55 to 130
melt_look	nhru	day	31 to 73

C.6. Parameters Derived from ASO data

Four parameters were derived from ASO data (Table C-3). The ASO lidar flights provided percent of each HRU with vegetation greater than 0, 1, 3 and 5 meters. This data was used to approximate the vegetation type (cov_type) and summer cover density (covden_sum). Winter cover density (covden_win) was assumed to be a function of summer cover density. The percent of each HRU with impervious bedrock (hru_percent_imperv) was assumed to be equal to the fraction of the HRU with vegetation less than 1 m tall.

Table C-3: Parameters directly derived from ASO Data

Parameter	Dimension	Units	Value Range

cov_type	nhru	none	0 to 3
covden_win	nhru	none	0 to 0.6
covden_sum	nhru	none	0 to 0.9939
hru_percent_imperv	nhru	none	0 to 0.3914

C.7. Parameters Taken from Merced Calibration

Seven Parameters were taken directly from the Merced Calibration (USGS – detailed in PRMS Manual) (Table C-4). These parameters control solar radiation and evaporation characteristics and are assumed to be consistent between the Merced and Tuolumne basins.

Table C-4: Parameters Taken from Merced Calibration

Parameter	Dimension	Units	Value Range
albset_rna	one	fraction	0.8
albset_rnm	one	fraction	0.6
albset_sna	one	fraction	0.05
albset_snm	one	fraction	0.2
radmax	one	fraction	0.8
settle_constant	one	fraction	0.1
solrad_elev	one	elev_units	1208
adjmix_rain	nmonths	fraction	1
dday_intcp	nmonths	dday	-50.095 to 5.877
dday_slope	nmonths	dday/F	0.305 to 0.790
tmax_index	nmonths	F	60.800 to 82.583
snarea_thresh	nhru	inches	5 to 37.2385

C.8. Parameters Determined Using Professional Expertise

31 Parameters were set using the professional expertise of the developers of the model. Some are straightforward, such as the measurement units (XXX_units). Others were set to the default value, due to the difficulties measuring and / or calibrating. Other parameters were set at default because model outputs were not sensitive to the parameter value.

Table C-5: Parameters Determined using Professional Expertise

Parameter	Dimension	Units	Value Range
conv_flag	one	none	0
elev_units	one	none	0
outlet_sta	one	none	1
parameter_check_flag	one	none	0
precip_units	one	none	0
print_freq	one	none	5
print_type	one	none	2
radj_sppt	one	decimal fraction	0.44
radj_wppt	one	decimal fraction	0.50
runoff_units	one	none	0
temp_units	one	none	0
psta_freq_nuse	nrain	none	0 or 1
psta_nuse	nrain	none	0 or 1
tsta_nuse	ntemp	none	0 or 1
adjust_rain	nmonths	inches	0
adjust_snow	nmonths	inches	0
ppt_rad_adj	nmonths	inches	0
rain_code	nmonths	none	5
tstorm_mo	nmonths	month	0 or 1
obsin_segment	nsegment	none	0
segment_flow_init	nsegment	cfs	0
segment_type	nsegment	none	0
snarea_curve	ndeplval	decimal fraction	0.00 to 0.99
gwstor_min	ngw	inches	0.0
hru_deplcrv	nhru	none	1
snowpack_init	nhru	inches	0
soil_type	nhru	none	1
tmax_adj	nhru	F	0
tmin_adj	nhru	F	0
transp_beg	nhru	month	4
transp_end	nhru	month	10

C.9. Parameters Derived from Calibration to Snow Pillows and Surveys

Five parameters were calibrated using Snow Pillow and Survey data. 300,000 Monte Carlo simulations were run, using parameter values as determined in sections 7.1-7.5 above, and varying the parameters listed in Table C-6. At 300,000 runs, and 5 parameters, each parameter space was dissected ~12.5 times. The parameters tmax_allsnow_dist and tmax_allrain_dist are relics of earlier versions of PRMS – it is unclear which are actually used in the current version, so both were included and set to identical values as tmax_allsnow and tmax_allrain, respectively. After the Monte Carlo runs, the top 100 parameter sets were extracted for each Snow Survey and Snow Pillow. The parameter set was in the top 100 for the most sites was chosen (parameter set 65,877).

Closer inspection showed poor fits to the Tuolumne and Dana Meadows snow pillows and Survey sites. It was determined that this was due to the well-established rain shadow in the southeast corner of the Tuolumne watershed. To increase the fits at these sites, and incorporate the institutional knowledge about the rain shadow, the parameter ppt_lapse was manually adjusted until Tuolumne and Dana Meadow pillow and survey fits improved (i.e. Figure C-1). The north-south lapse was increased, while the east-west lapse was decreased. The vertical lapse, as determined using the observed vertical lapse rate in measured precipitation, was not adjusted. Total precipitation volume in the watershed was kept constant. These edits to the lateral lapse rates were incorporated into the Hetch Hetchy PRMS model but were not incorporated into the Cherry / Eleanor model, as the Cherry / Eleanor model domain does not include areas influenced by the Tuolumne rain shadow.

Table C-6: Parameters derived from calibration to snow pillows and surveys

Parameter	Dimension	Units	Value Range
emis_noppt	one	fraction	0.757798
freeh2o_cap	one	fraction	0.198345
potet_sublim	one	fraction	0.22336
tmax_allrain	one	F	39.69274
tmax_allsnow	nmonths	F	33.67106
ppt_lapse (east/west and north/south)	nlapse X nmonths	none	-0.3 & 0.3

Table C-7: Model fits to Snow Survey and Snow Pillow sites

Site	Lowest RMSE, All Simulations	RMSE, Best Fit Scenario	RMSE, Best Fit with ppt_lapse Edit
Beehive Meadows	17.3	17.4	17.7
Bell Meadow	10.7	10.9	9.4
Bond Pass	6.0	6.0	6.1
Dana Meadows	7.4	13.9	8.9
Gianelli Meadow	18.9	19.1	13.1
Gin Flat	19.9	19.9	24.8
Horse Meadow	16.3	16.5	15.3
Huckleberry Lake	18.7	18.7	17.7
Kerrick Corral	13.5	13.5	11.7
Lower Kibbie	17.1	17.2	16.6
Lower Relief Valley	13.6	13.7	10.4
New Grace Meadow	12.0	12.2	12.1
Paradise Meadow	20.0	20.0	21.3
Rafferty Meadows	6.0	8.7	12.6
Sachse Springs	13.2	13.3	11.2
Spotted Fawn	21.8	21.9	20.7
Tuolumne Meadows	7.1	13.3	8.2
Upper Kibbie	12.3	12.5	11.8
Vernon Lake	14.3	14.5	15.1
Wilmer Lake	18.6	18.6	20.0
Dana Meadows	5.4	11.5	5.2
Gianelli Meadow	10.3	10.3	7.7
Gin Flat	8.9	8.9	11.3
Horse Meadow	10.0	10.1	9.5
Lower Relief Valley	8.8	8.9	6.8
Paradise Meadow	10.3	10.4	10.9
Slide Canyon	6.7	6.8	9.0
Tuolumne Meadows	3.9	9.5	3.5
White Wolf	4.6	4.6	7.6
Average	12.2	13.2	12.3

Figure C-1: Model fit to Horse Meadow snow survey and pillow data. Subplots show same day measurement and model values for pillow and survey.

Figure C-2: Zoom in on model fit to Horse Meadow survey and pillow data. Note occasional discrepancy between pillow and survey data – this is expected due to differences in measurement volume and measurement techniques.
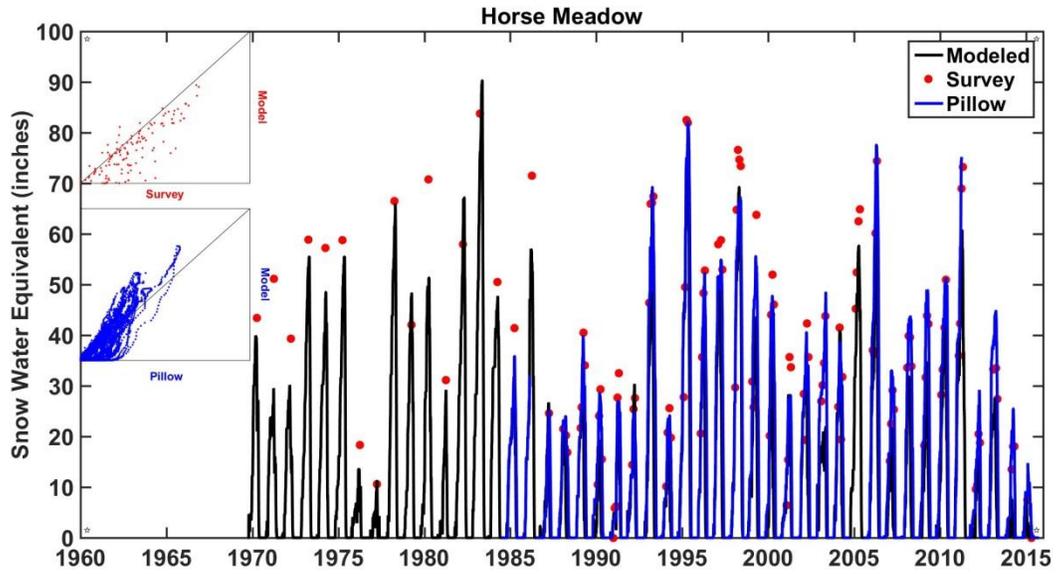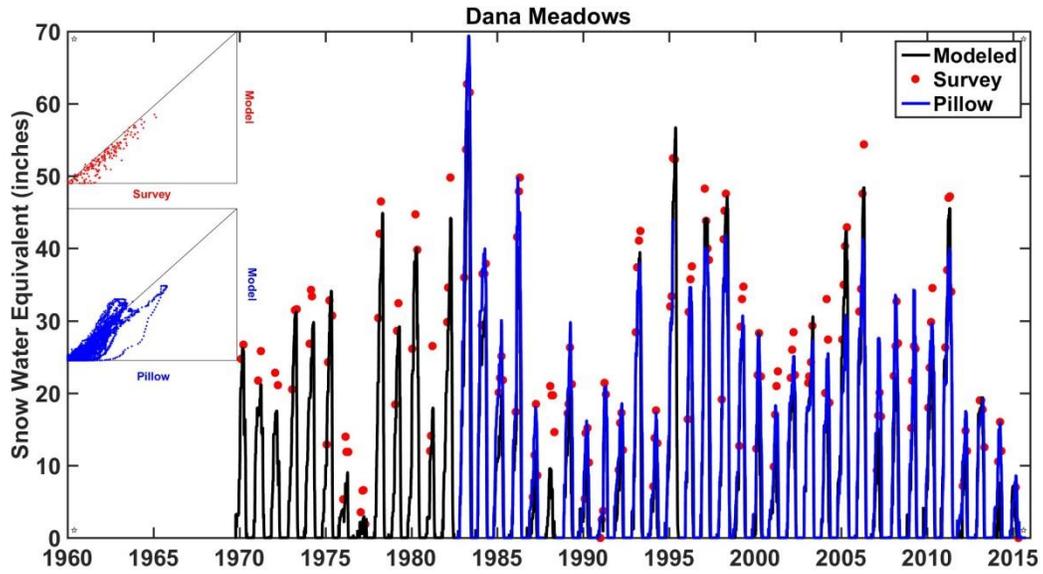
Figure C-3: Model fit to Dana Meadows snow survey and pillow data. Subplots show same day measurement and model values for pillow and survey.
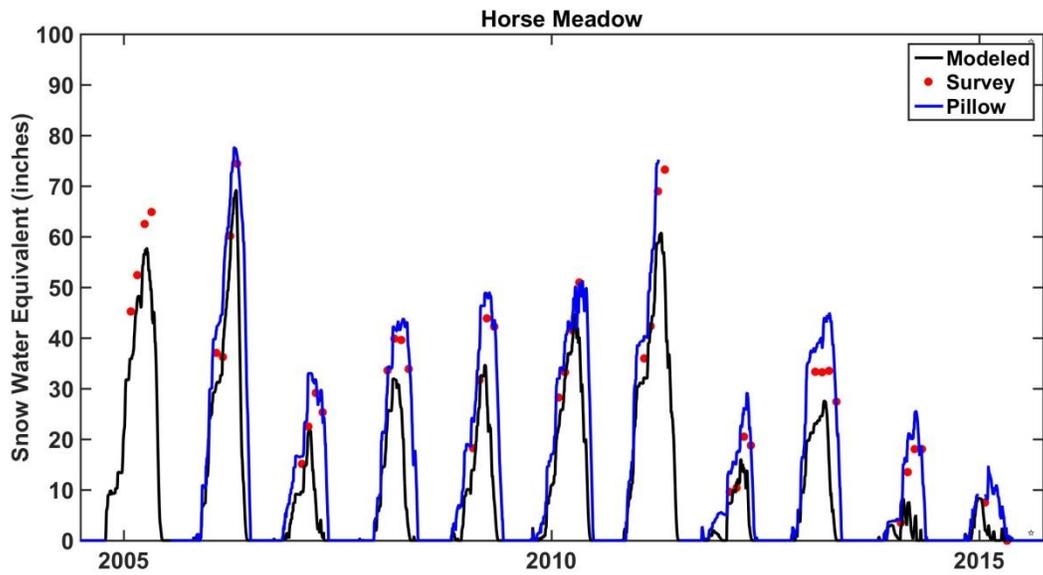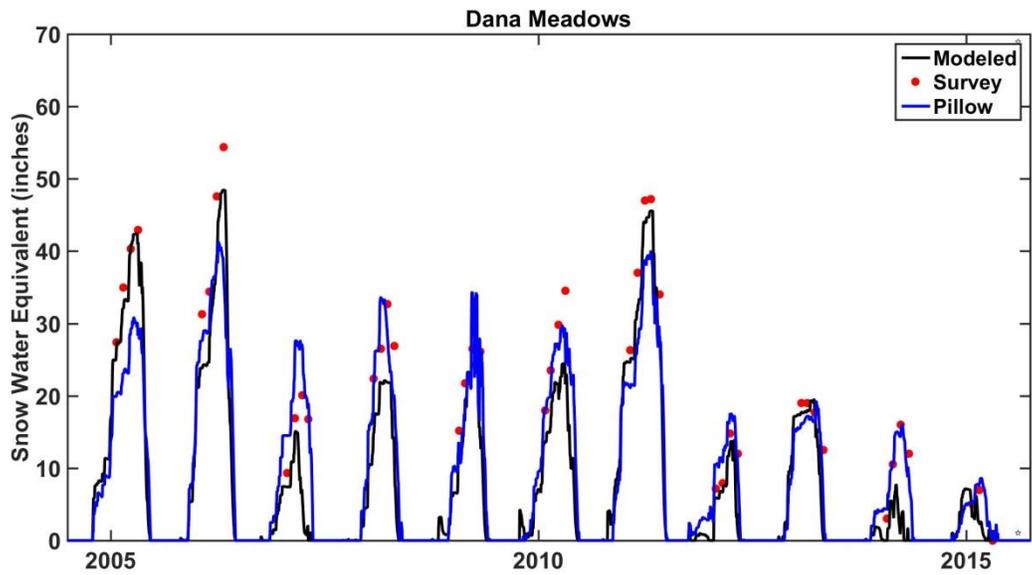

Figure C-4: Zoom in on model fit to Dana Meadows survey and pillow data. Note good fit to snowmelt timing.

C.10.   Parameters Derived from Calibration to ASO Data

Three parameters were calibrated using Airborne Snow Observatory (ASO) data. 100,000 Monte Carlo simulations were run, using parameter values as determined in sections 7.1-7.6 above, and varying the parameters listed in Table C-6. At 100,000 runs, and 3 parameters, each parameter space was dissected ~46.4 times. The parameters were all determined to be functions of the winter cover density, as determined by analysis of ASO snow off flights. The interception factors were both directly correlated with winter cover density, while the radiation transmission coefficient was set to one value throughout the domain.

Table C-8: Parameters derived from calibration to ASO Data

Parameter	Dimension	Units	Value Range
rad_trncf	nhru	decimal fraction	0.6
snow_intcp	nhru	inches	0.10933 to 0.28993
wrain_intcp	nhru	inches	0.10757 to 0.28527

C.11. Parameters Derived from Calibration to Seasonal Streamflow

12 parameters were calibrated using Seasonal (annual and April through July) Hetch Hetchy Inflows. 386,000 Monte Carlo simulations were run, using parameter values as determined in sections **Error! R eference source not found.**-C.10 above, and varying the parameters listed in Table C-6. At 386,000 runs, and 12 parameters, each parameter space was dissected ~2.9 times. These parameters were calibrated to the Hetch Hetchy seasonal streamflow, and then carried over to the Cherry / Eleanor model. Model fits to Cherry / Eleanor were deemed sufficiently good to not require additional calibration.

Table C-9: Parameters derived from calibration to seasonal streamflow

| Parameter	Dimension	Units	Value Range
cecn_coef	nmonths	calories per degree Celsius > 0	3.652001298
epan_coef	nmonths	decimal fraction	1.0 and 1.5
imperv_stor_max	nhru	inches	0.061581224
pref_flow_den	nhru	decimal fraction	0.421961763
smidx_coef	nhru	decimal fraction	0.016358179
smidx_exp	nhru	1/inch	0.121748104
snowinfil_max	nhru	inches/day	0.738982156
soil_moist_init	nhru	inches	0.064509483
soil_moist_max	nhru	inches	0.262287965
soil_rechr_init	nhru	inches	0.029082688
soil_rechr_max	nhru	inches	0.030929524
srain_intcp	nhru	inches	0.170156311
transp_tmax	nhru	temp_units	733
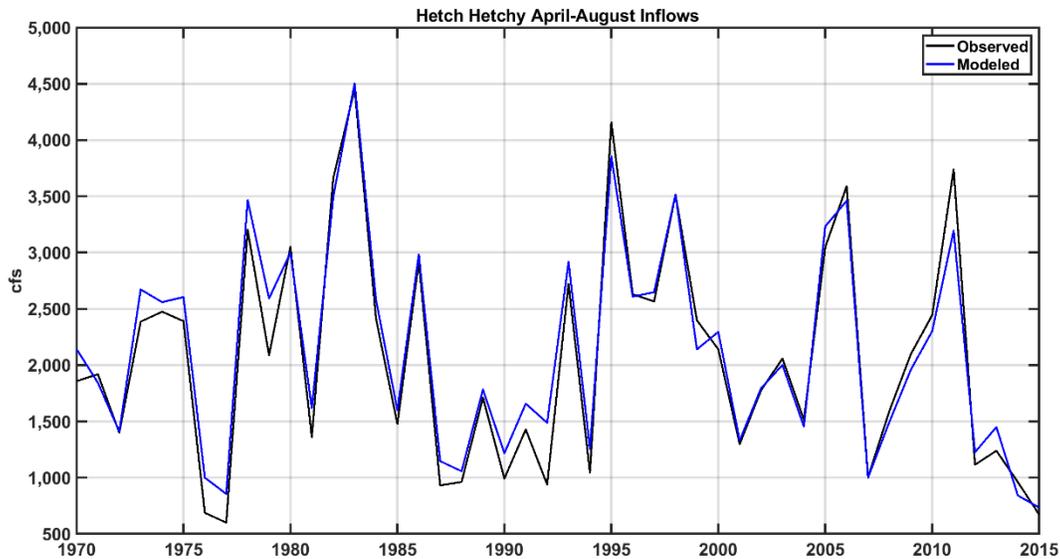
Figure C-5: Modeled and observed April to August inflows to Hetch Hetchy Reservoir.
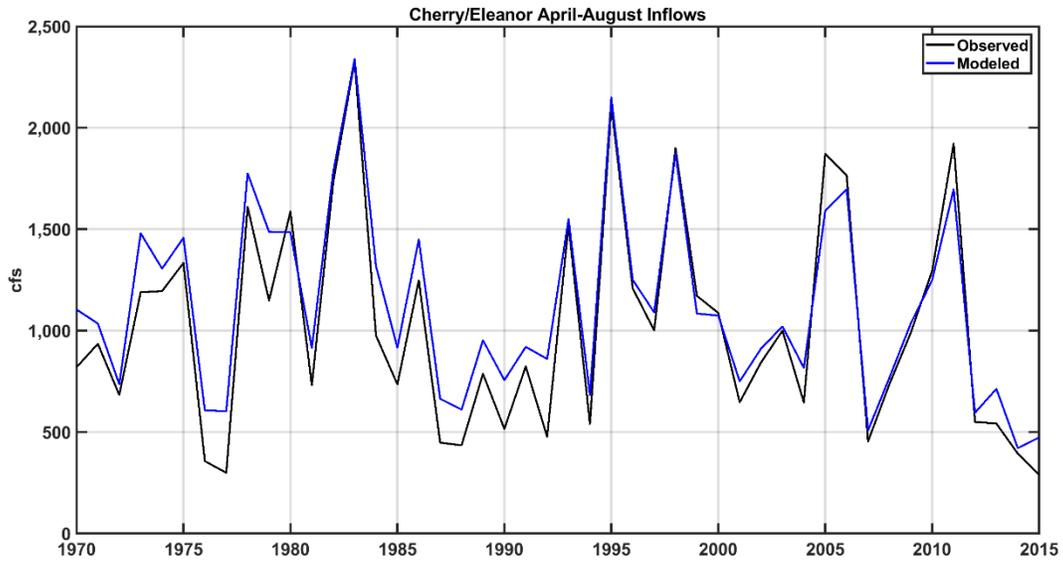
138

Figure C-6: Modeled and observed April to August inflows to Hetch Hetchy Reservoir
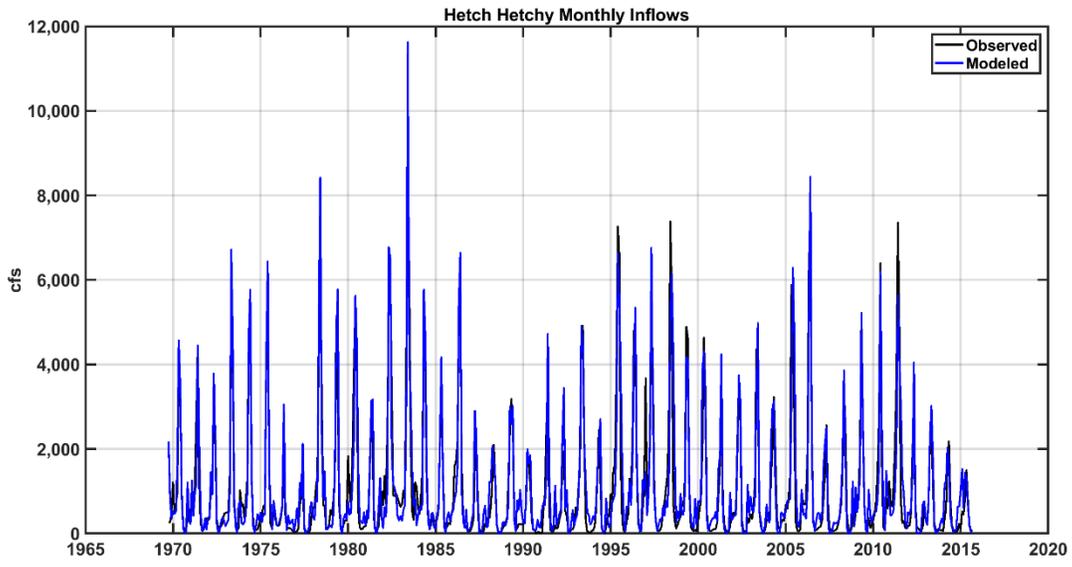

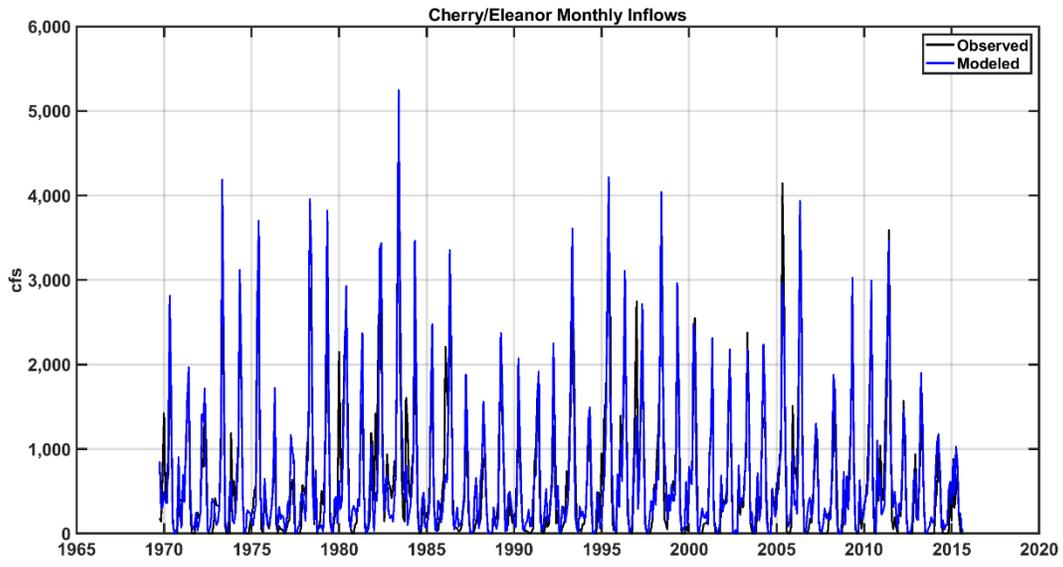*Figure C-7: Monthly modeled and observed inflows to Hetch Hetchy watershed*

Figure C-8: Monthly modeled and observed inflows to Cherry / Eleanor watershed

C.12. Parameters Derived from Calibration to Daily Streamflow

The remaining 18 parameters were calibrated using daily Hetch Hetchy Inflows. 1,000,000 Monte Carlo simulations were run, using parameter values as determined in sections above, and varying the parameters listed in Table C-6. At 1,000,000 runs, and 18 parameters, each parameter space was dissected ~2.15 times. These parameters were calibrated to the Hetch Hetchy daily streamflow, and then carried over to the Cherry / Eleanor model. Model fits to Cherry / Eleanor were deemed sufficiently good to not require additional calibration.

Table C-10: Parameters derived from calibration to daily streamflow

| Parameter	Dimension	Units	Value Range
radadj_intcp	one	dday	0.1387
radadj_slope	one	dday / temp_units	0.5948
jh_coef	nmonths	per degrees F	0.0094362
K_coef	nsegment	hours	1 to 5.5479
x_coef	nsegment	decimal fraction	0.17128
carea_max	nhru	decimal fraction	0.4787
fastcoef_lin	nhru	fraction / day	0.0439
fastcoef_sq	nhru	none	0.6073
sat_threshold	nhru	inches	40.7987
slowcoef_lin	nhru	fraction / day	0.3835
slowcoef_sq	nhru	none	0.2597
soil2gw_max	nhru	inches	2.648
ssr2gw_exp	nhru	none	2.2745
ssr2gw_rate	nhru	fraction / day	0.4241
ssstor_init	nhru	inches	10.0714
gwflow_coef	ngw	fraction / day	0.1326
gwsink_coef	ngw	fraction / day	0.0203
gwstor_init	ngw	inches	1.7452
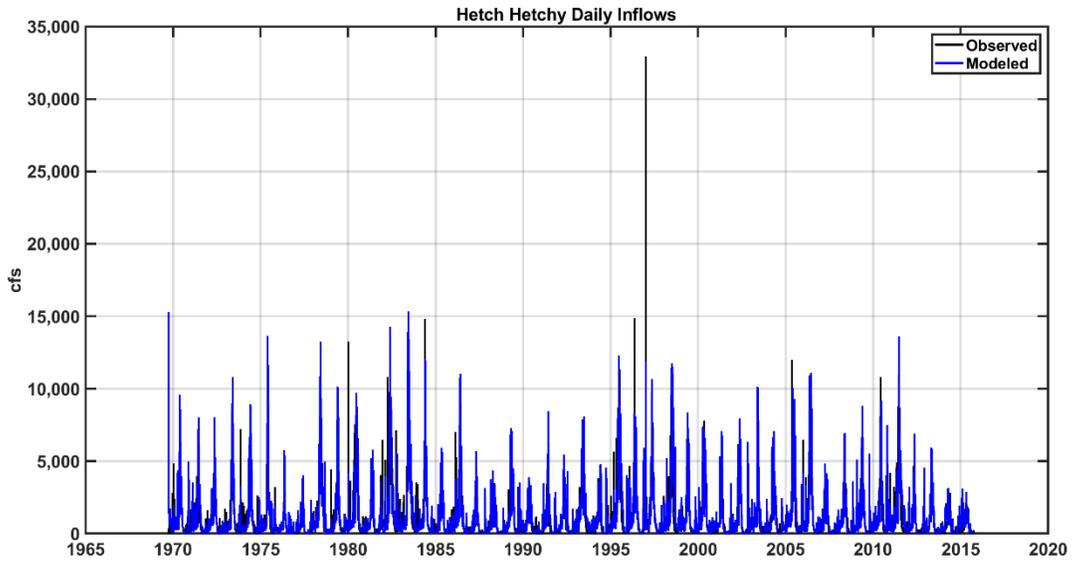
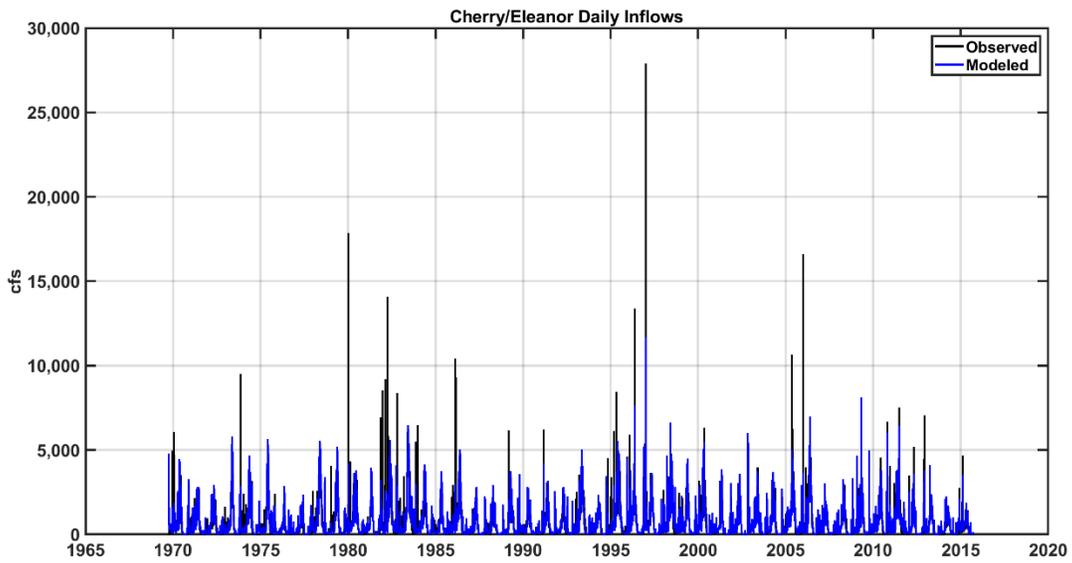*Figure C-9: Daily modeled and observed inflows to Hetch Hetchy watershed*

Figure C-10: Daily modeled and observed inflows to combined Cherry / Eleanor watershed

142
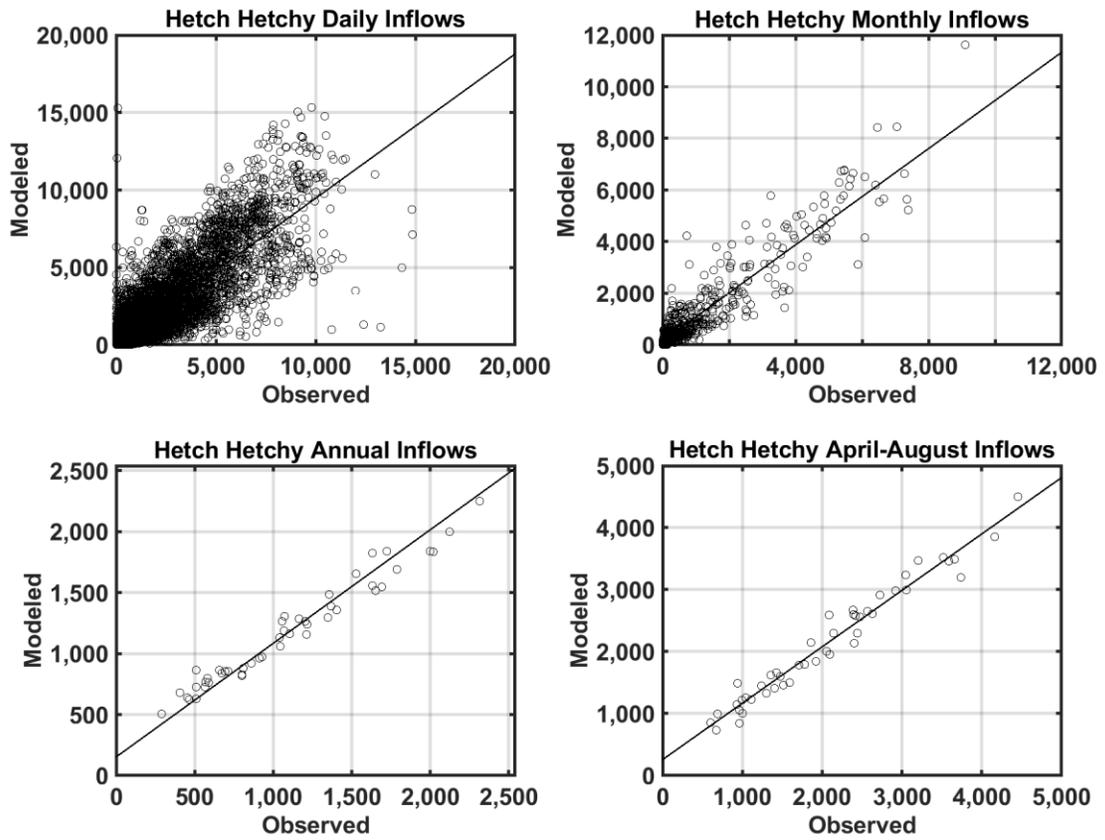
Figure C-11: Dotty Plots of Hetch Hetchy model fits at daily, monthly, annual and seasonal timescales
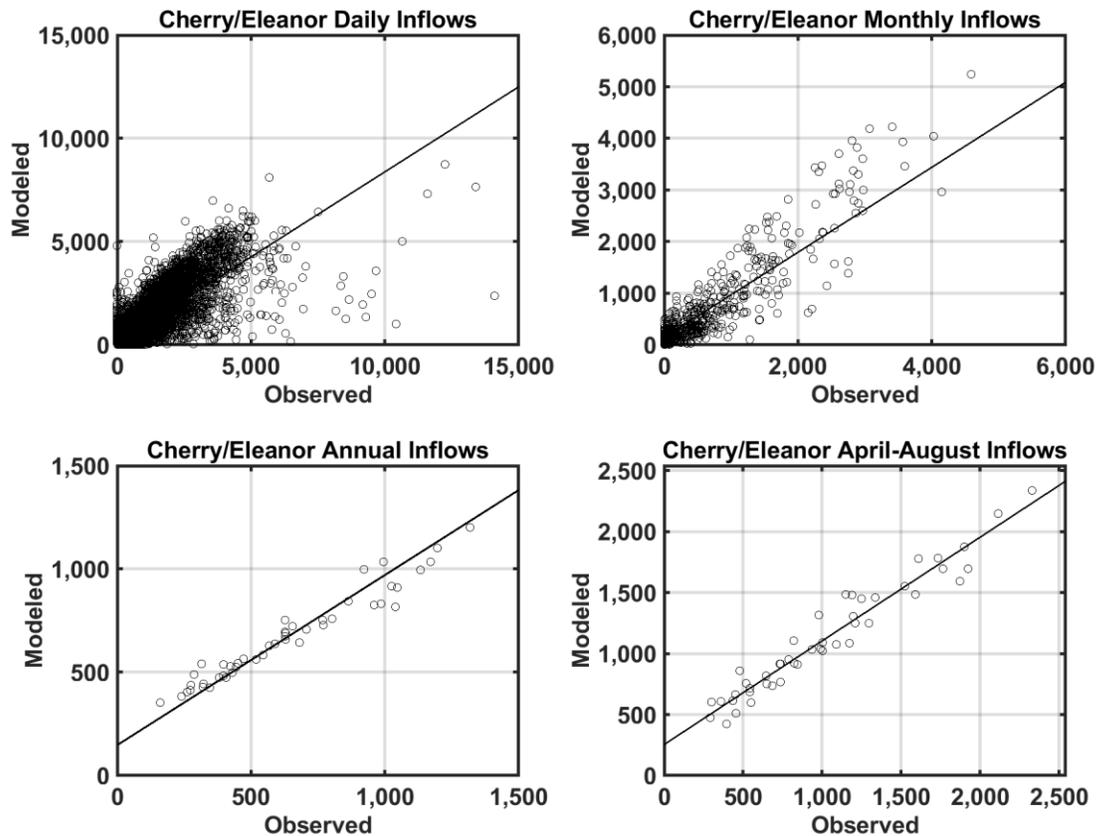
Figure C-12: Dotty Plots of Cherry/Eleanor model fits at daily, monthly, annual and seasonal timescales

D. Appendix – Description of SAC-SMA-DS Hydrologic Model

Because of its essential role in the quantification of available water on which water allocations to all water sectors is based, very high performance is required of the hydrologic model. Hydrologic model residuals propagate through the modeling chain and contribute to a cascade of uncertainty [*Wilby and Dessai*, 2010]. This section describes the development of a distributed, physically-based hydrologic model capable of supporting subsequent phases of the climate change vulnerability assessment workflow.

The amount of usable water for the CVS can be approximated as the quantity of streamflow in the twelve largest rivers flowing from the north-east into the Central Valley. These are referred to as the rim inflows. In order to estimate those twelve stream flows, the Sacramento Soil Moisture Accounting (SAC-SMA) model, a lumped conceptual hydrological model employed by the National Weather Service (NWS) of the National Oceanic and Atmospheric Administration (NOAA) to produce river and flash flood forecasts for the nation [*McEnergy et al.*, 2005], was coupled with a river routing model to be suitable for modeling a

distributed watershed system. It is here referred to as SAC-SMA_DS, denoting the distributed version of SAC-SMA. SAC-SMA_DS (Figure D-1) is composed of hydrologic process modules that represent soil moisture accounting, evapotranspiration, snow processes, and flow routing. The model operates on a daily time step and requires daily precipitation and mean temperature as input variables.

SAC-SMA_DS includes the Snow 17 module [*Anderson*, 1976] to account for snow and ice dynamics within the 12 rim sub-basins. In this study the hydrologic modeling domains for 12 rim sub-basins are spatially disaggregated using climate input grids of 1/8º resolution and 200 m interval elevation bands corresponding to the meteorological source data [*Maurer et al.*, 2002]. The runoff from each disaggregated area is weighted by its area fraction within the basin to obtain the total basin-wide runoff.

More details on the model components are provided below by focusing on the descriptions for the modules additionally introduced to develop the distributed version of SAC-SMA.
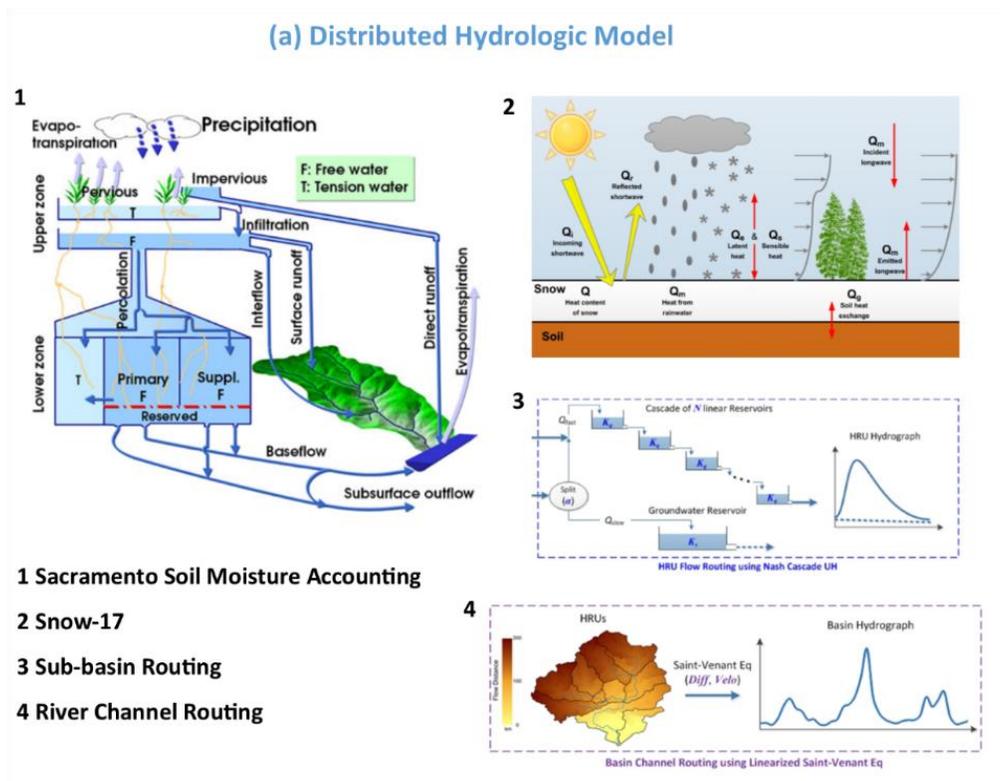
Figure D-1 Schematic of distributed hydrologic model

D.1. Hamon Evapotranspiration Calculation

The potential evapotranspiration (PET) is derived based on the Hamon method [*Hamon*, 1961], in which daily PET in millimeters (mm) is computed as a function of daily mean temperature and hours of daylight:

$$PET = \text{Coeff} \cdot 29.8 \cdot L_d \cdot \frac{0.611 \cdot \exp\left(17.27 \cdot \frac{T}{(T+273.3)}\right)}{T+273.3} \quad (2)$$

where, L_d is the daylight hours per day, T is the daily mean air temperature (ºC), and Coeff is a bias correction factor. The hours of daylight is calculated as a function of latitude and day of year based on the daylight length estimation model suggested by *Forsythe et al.* [1995].

D.2. In-grid Routing: Nash-Cascade Unit Hydrograph

The within-grid routing process for direct runoff is represented by an instantaneous unit hydrograph (IUH) [Nash, 1957], in which a catchment is depicted as a series of N reservoirs each having a linear relationship between storage and outflow with the storage coefficient of K_q. Mathematically, the IUH is expressed by a gamma probability distribution:

$$u(t) = \frac{K_q}{\Gamma(N)} (K_q t)^{N-1} \exp(-K_q t) \qquad (3)$$

where, Γ is the gamma function. The within-grid groundwater routing process is simplified as a lumped linear reservoir with the storage recession coefficient of K_s.

D.3. River Channel Routing: Linearized Saint-Venant Equation

The transport of water in the channel system is described using the diffusive wave approximation of the Saint-Venant equation [Lohmann et al., 1998]:

$$\frac{\partial Q}{\partial t} + C\frac{\partial Q}{\partial x} - D\frac{\partial^2 Q}{\partial^2 x^2} = 0 \quad (4)$$

where C and D are parameters denoting wave velocity and diffusivity, respectively.

References

Anderson, E. A. (1976), A point energy and mass balance model of a snow cover, *NOAA Tech. Rep. NSW 19, Natl. Oceanic and Atmos. Admin., Silver Spring, MD.*, 1-150.

Maurer, E. P., A. W. Wood, J. C. Adam, D. P. Lettenmaier, and B. Nijssen (2002), A long-term hydrologically based dataset of land surface fluxes and states for the conterminous United States, *J. Clim., 15*, 3237-3251.

Hamon, W. R. (1961), Estimating potential evapotranspiration, *J. Hydr. Eng. Div. -ASCE, 87*, 107-120.

Forsythe, W. C., E. J. Rykiel Jr., R. S. Stahl, H. Wu, and R. M. Schoolfield (1995), A model comparison for daylength as a function of latitude and day of year, *Ecol. Model., 80*, 87-95.

Nash, J. E. (1957), The form of the instantaneous unit hydrograph, *International Association of Science and Hydrology, 3*, 114-121.

Lohmann, D., R. Raschke, B. Nijssen, and D. P. Lettenmaier (1998), Regional scale hydrology: I. Formulation of the VIC-2L model coupled to a routing model, *Hydrolog. Sci. J., 43*, 131-141.

Wilby, R. L. and S. Dessai (2010), Robust adaptation to climate change, *Weather, 65*, 180-185.

McEnery, J., J. Ingram, Q. Duan, T. Adams, and L. Anderson (2005), NOAA'S Advanced Hydrologic Prediction Service – Building pathways for Better Science in Water Forecasting, *Bulletin of the American Meteorological Society, March*, 375-385.

E. Appendix – Description of Genetic Algorithm

John Holland proposed genetic algorithm (GA) in 1975 (McCall, 2005) in order to solve difficult optimization problems such as non-linear objective functions or mixed integer programming where some components are restricted to be integers.

This project applies GA in calibration process. GA uses the concept of biological evolution to create their algorithm. In brief, good genes in parents can inherit good genes in children; similarly in the context of optimization, the best combination of parameters towards the objective function is produced by other good combinations of parameters in previous generations. Particularly, in this study, 20 parameters in SAC-SMA can be modified to optimize the model performance or the fitness between stream flow simulation and observation. Moreover, this study uses Kling-Gupta efficiency (KGE) for assessing the performance of hydrologic model. It is between −∞ and 1; where −∞ means no fitness between discharge simulation and observation; 1 means perfect fitness between modeled stream flows and observed stream flows. KGE is computed as follow (Gupta et al. 2009).

$$KGE = 1 - ED$$

$$ED = \sqrt{(r-1)^2 + (\alpha - 1)^2 + (\beta - 1)^2}$$

ED: Euclidian distance from the ideal point

r : linear correlation between simulated flows (Qs) and observed flows (Qo)

$\alpha = \frac{\sigma_{Qs}}{\sigma_{Qo}}$: ratio between the standard deviation of simulated and observed flows

$\beta = \frac{\mu_{Qs}}{\mu_{Qo}}$: ratio between the mean simulated and mean observed flows

F. Appendix – Calibration metrics

Various calibration metrics are used to evaluate the performance of hydrologic models. This appendix will review Nash-Sutcliffe Efficiency (NSE), Kling-Gupta Efficiency (KGE), and percent bias (PBias).

Nash-Sutcliffe Efficiency (NSE) is the most widely-used metric to compare simulated and observed streamflow values in hydrologic models (Moriasi, et al., 2006). NSE is calculated as:

$$NSE = 1 - \frac{\sum_{t=1}^{T}(Q_m^t - Q_o^t)^2}{\sum_{t=1}^{T}\left(Q_o^t - \overline{Q_o}\right)^2}$$

where Q_o is observed discharge (runoff), and Q_m is modeled (simulated) discharge, t is the time step, and T is the total number of time steps. NSE measures the normalized residual variance, and can range from −∞ to 1, with NSE equal to 1 representing a perfect match between modeled and observed discharge.

Kling-Gupta Efficiency (KGE) is another widely-used metric to compare simulated and observed streamflow values in hydrologic models (Gupta, et al., 2009). KGE Is calculated as:

148

$$KGE = 1 - \sqrt{(r-1)^2 + \left(\frac{\sigma_{sim}}{\sigma_{obs}} - 1\right)^2 + \left(\frac{\mu_{sim}}{\mu_{obs}} - 1\right)^2}$$

Where r is the Pearson product-moment correlation coefficient between simulated and observed time-step, σ_{sim} and σ_{obs} are the standard deviations of simulated and observed respectively, and μ_{sim} and μ_{obs} are the mean of simulated and observed respectively.

Percent bias is calculated as:

$$pBias = \frac{\sum_{t=1}^{T}(Q_m^t - Q_o^t)}{\sum_{t=1}^{T}(Q_o^t)} \times 100$$

where terms are the same as for NSE above. Somewhat counterintuitively, a negative pBias indicates underproduction (simulated is greater than observed), whereas a positive pBias indicates overproduction (simulated is less than observed).

G. Appendix – Representative Calibration Results of SAC-SMA-DS

The figures in this Appendix represent calibration results for each of the Alameda and Peninsula hydrologic regions.

G.1. Alameda Calibration Results

G.1.1. Arroyo Hondo

Figure G-1 Monthly hydrograph of observed (dotted) and simulated flow (red) - Arroyo Hondo

Figure G-2 Comparison of total annual flow - Arroyo Hondo

Figure G-3 Comparison of maximum annual average flow (60 days) - Arroyo Hondo

G.1.2. Alameda Creek Diversion Dam

Monthly Hydrograph of Observed (dotted) and Simulated Flow (red) (ACDD)
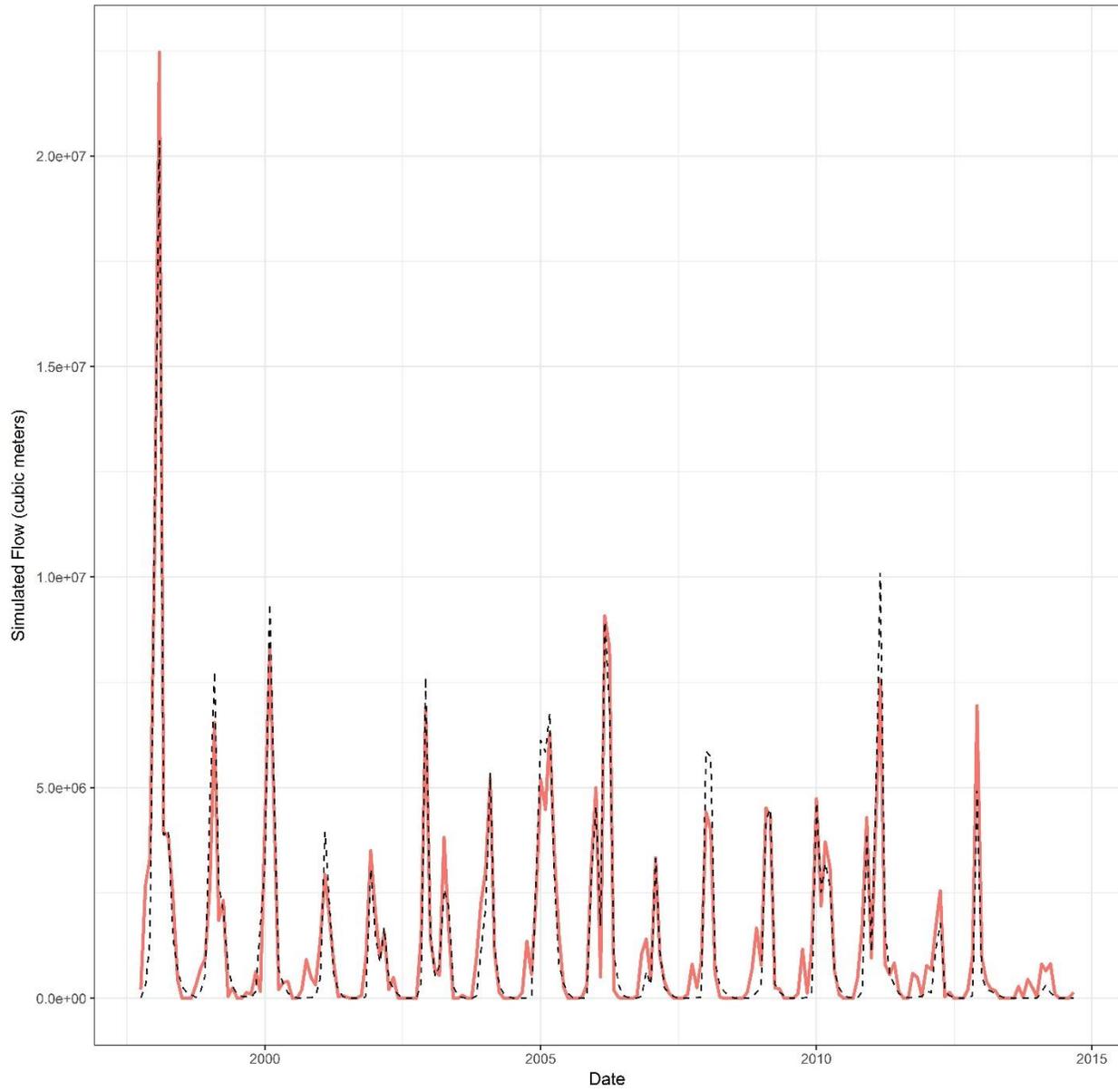
Figure G-4 Monthly hydrograph of observed (dotted) and simulated flow (red) - ACDD
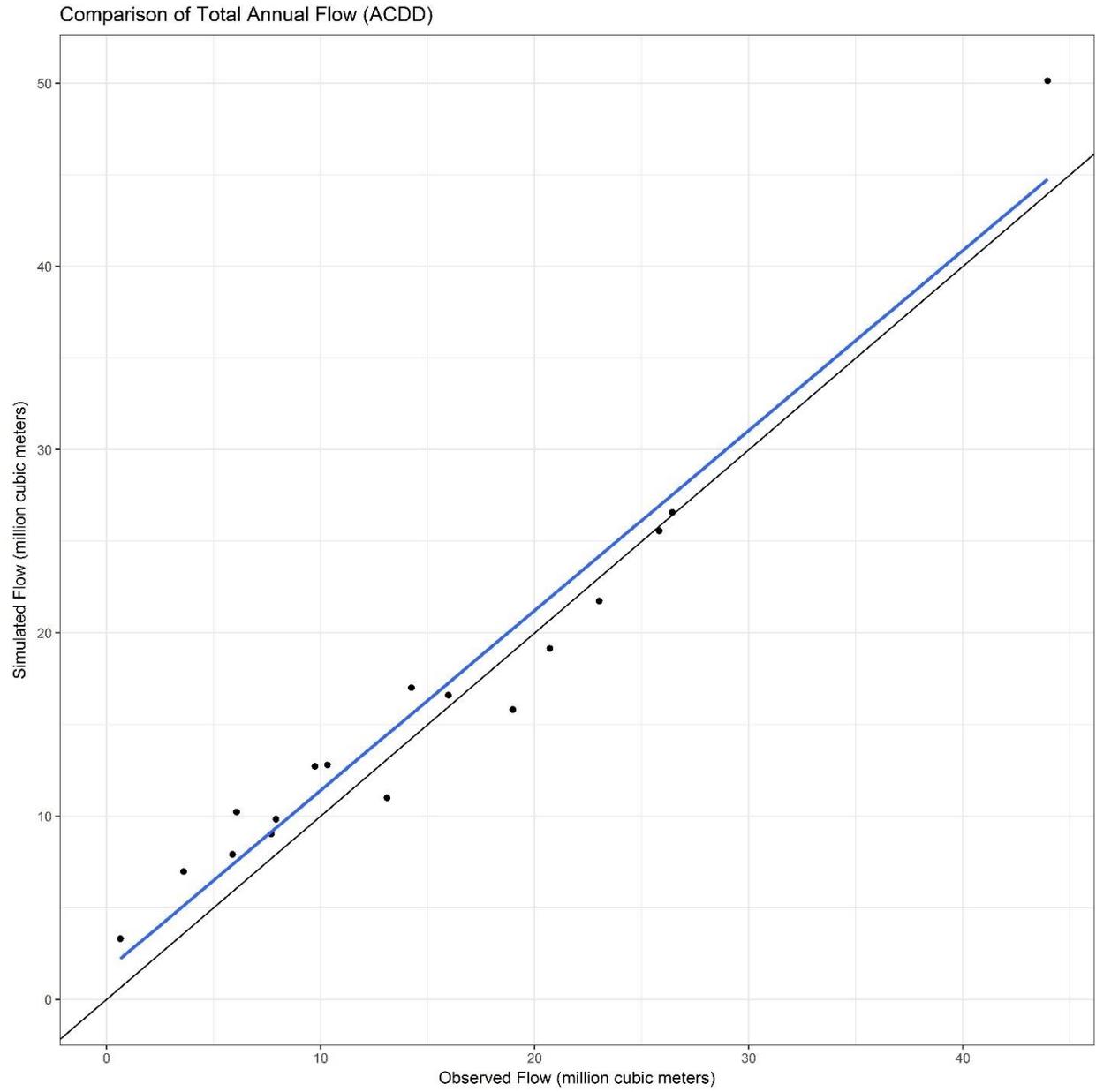
Figure G-5 Comparison of total annual flow - ACDD
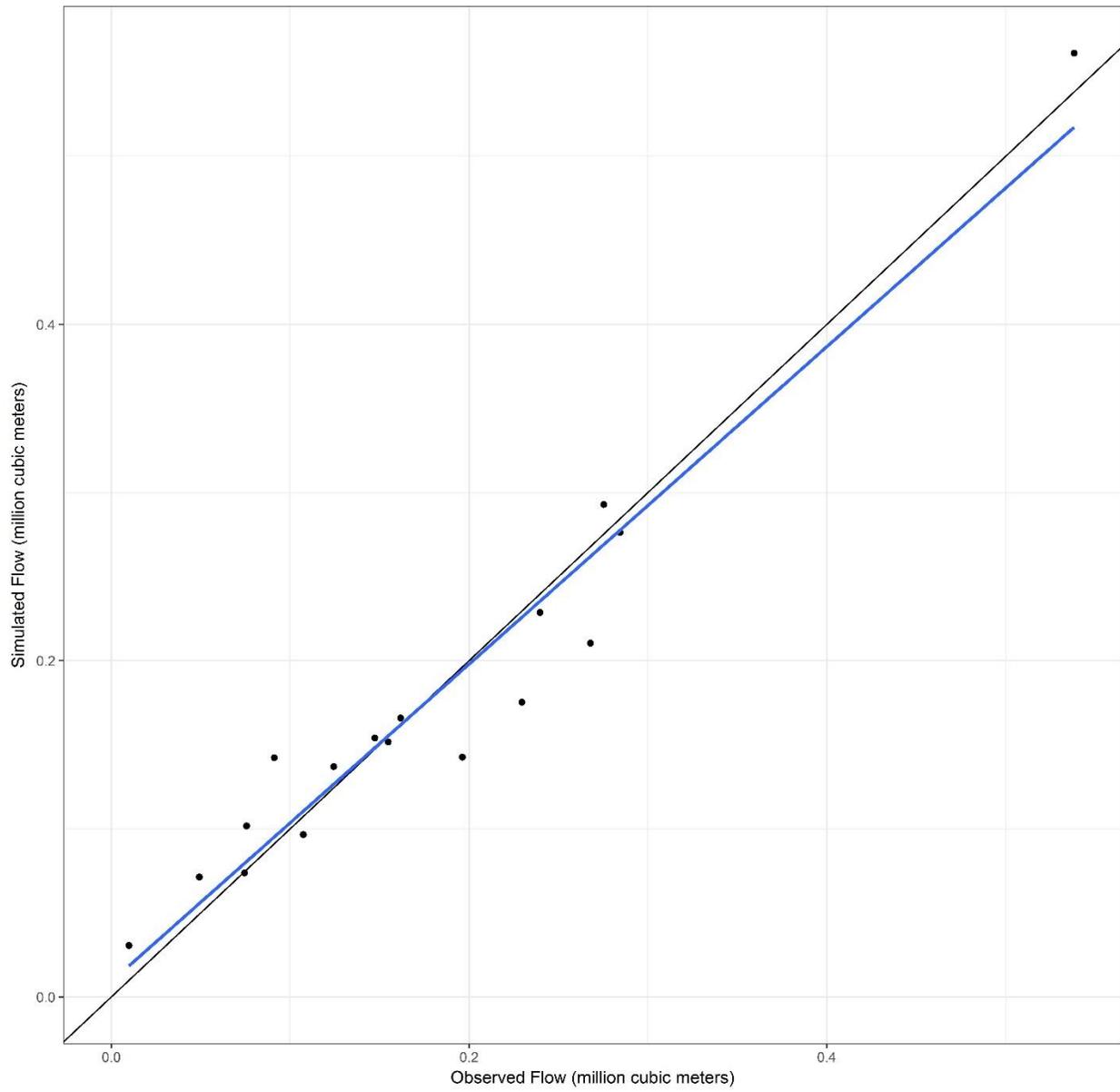
Figure G-6 Comparison of maximum annual average flow (60 days) - ACDD

G.1.3. San Antonio

Monthly Hydrograph of Observed (dotted) and Simulated Flow (red) (San Antonio)
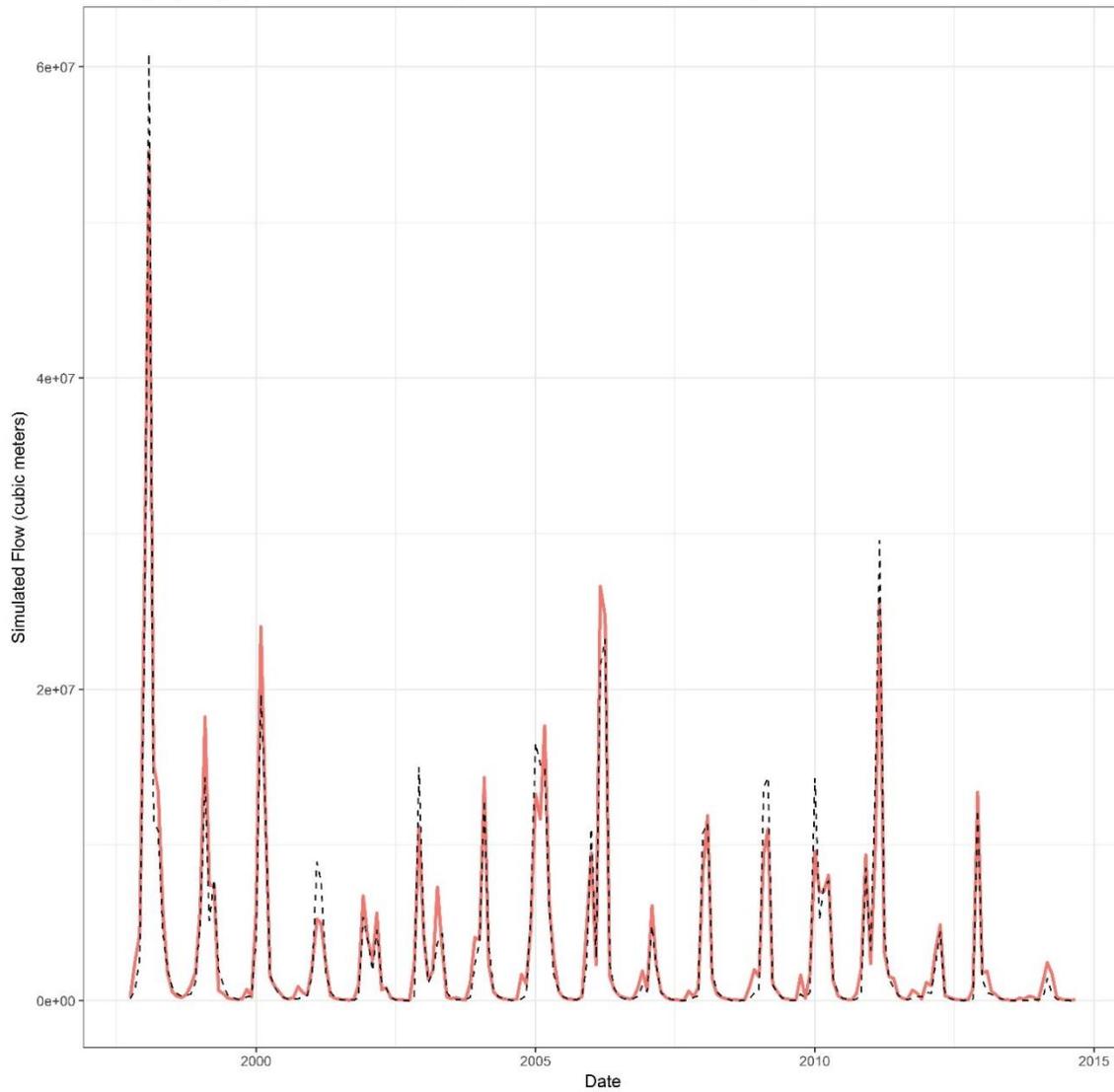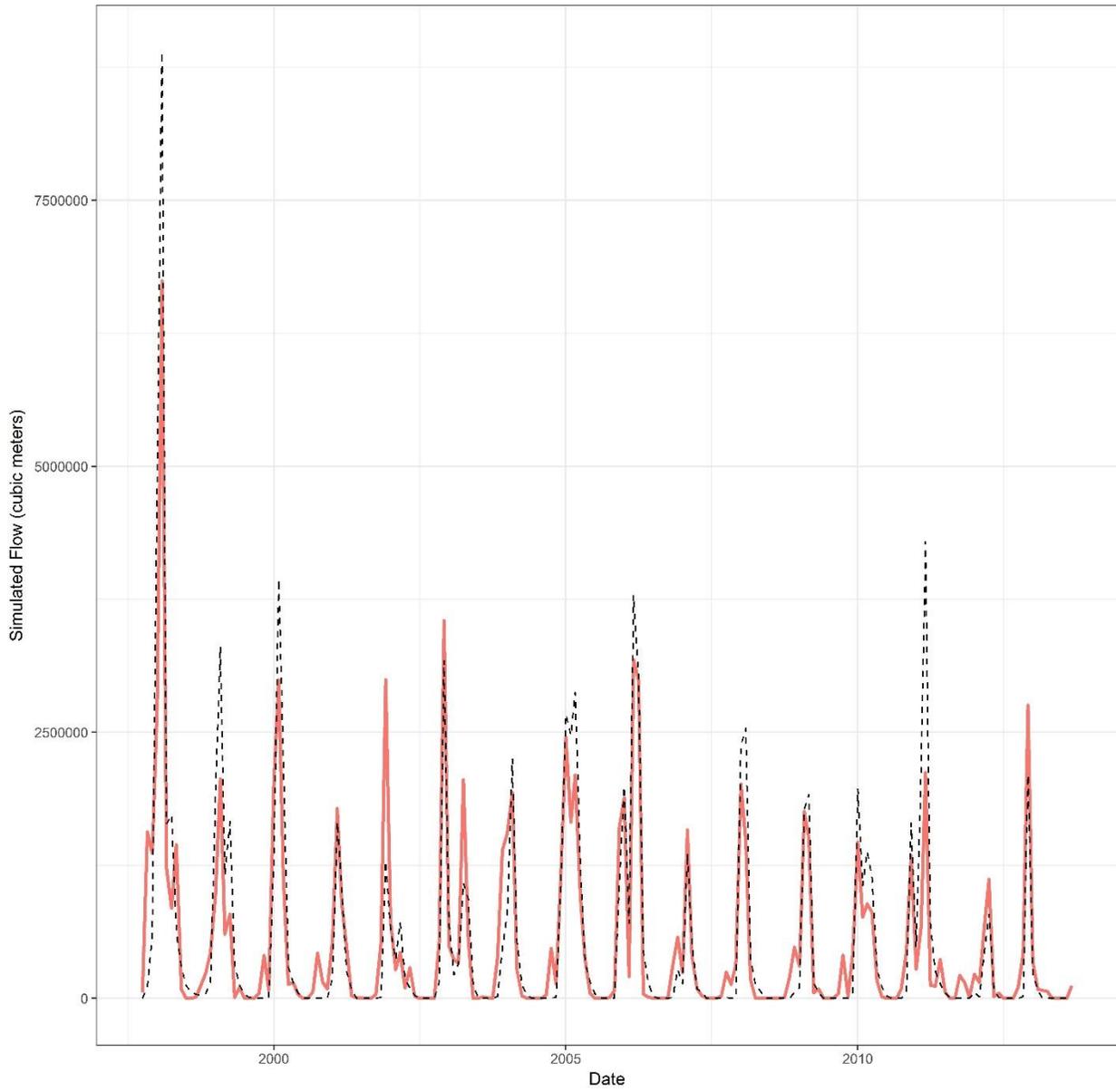

Figure G-7 Monthly hydrograph of observed (dotted) and simulated flow (red) - San Antonio
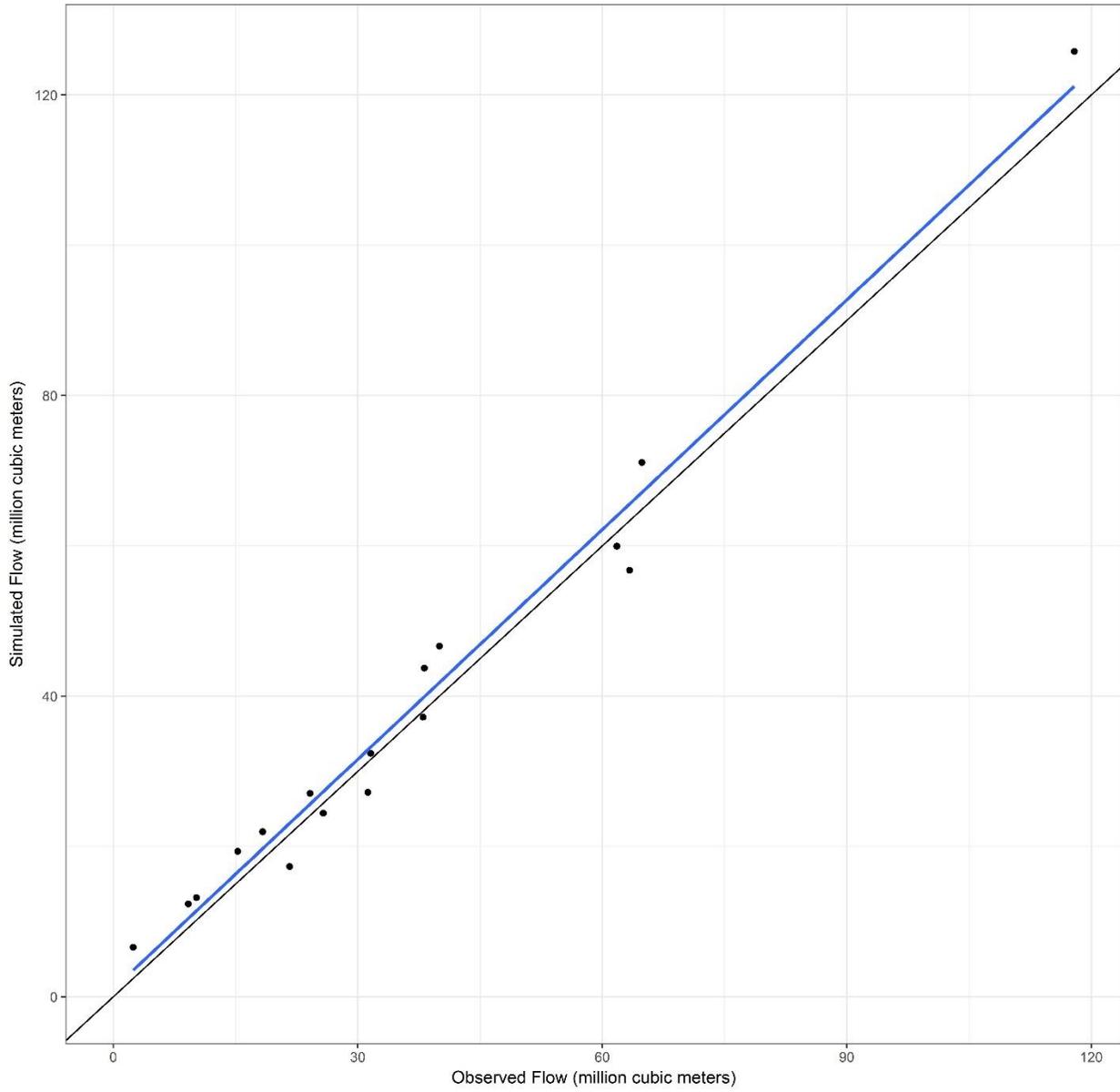
Comparison of Total Annual Flow (San Antonio)
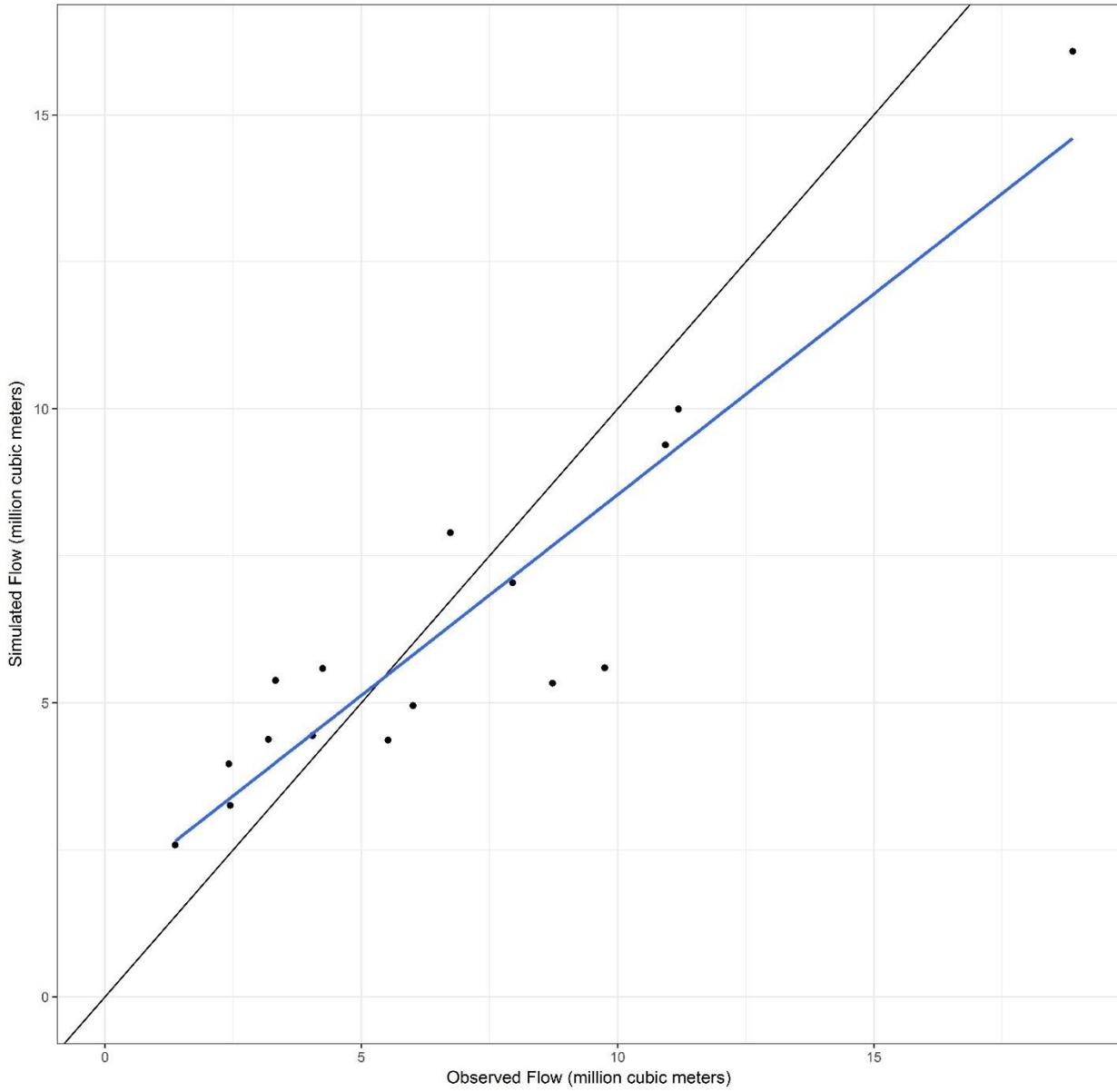
Figure G-8 Comparison to total annual flow - San Antonio
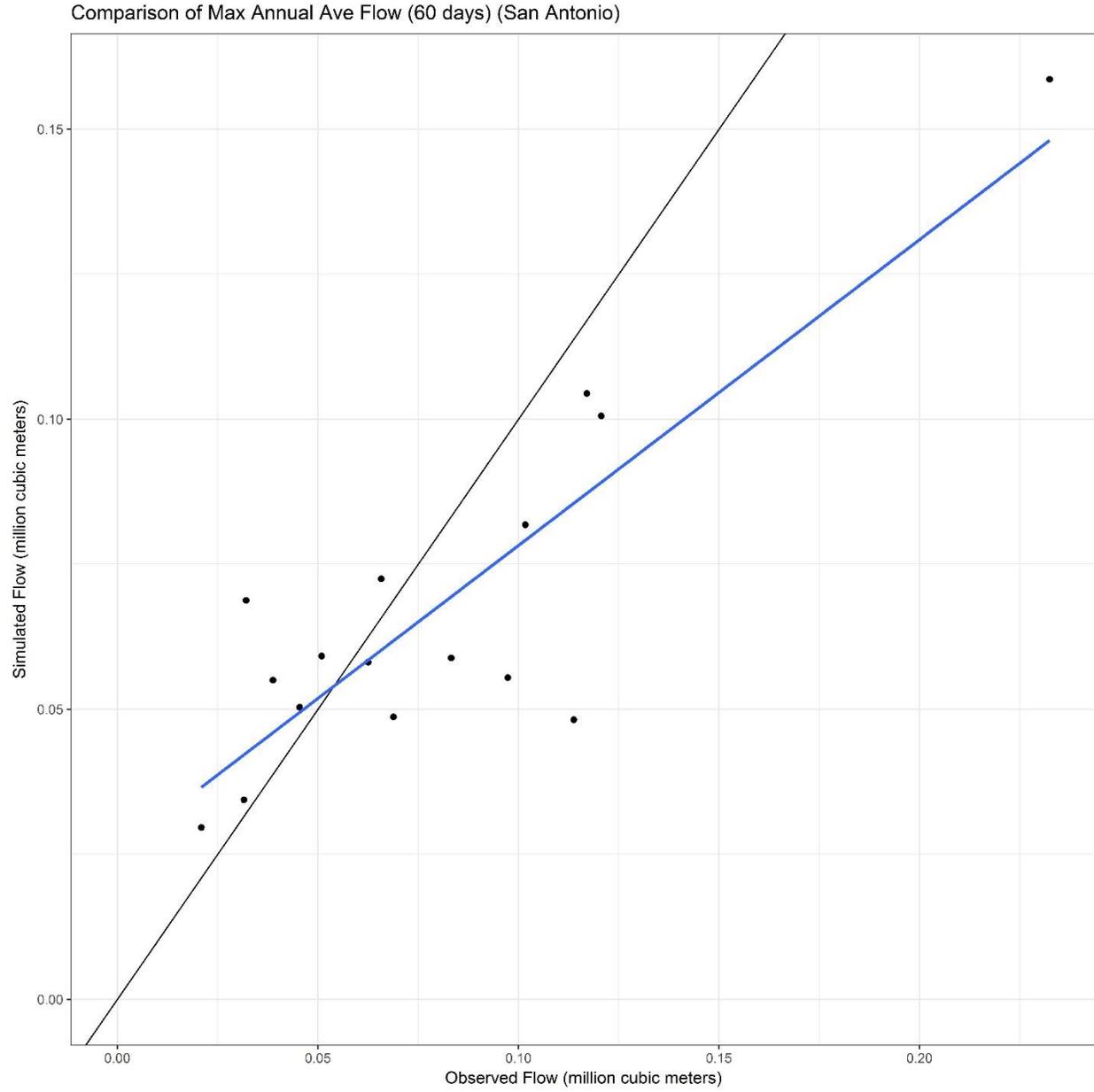
Figure G-9 Comparison of maximum annual average flow (60 days) - San Antonio

G.2. Peninsula Calibration Results
G.2.1. Basin 1 − Crystal Spring Reservoir watershed

Q1: Crystal Springs Reservoir Inflow from San Mateo Creek
Q7: San Andreas Reservoir Inflow from Watershed
Q21: Total Stream Release from Crystal Springs Reservoir
Q22: Total Stream Release from San Andreas Reservoir

Figure G-10. Basin 1 (Crystal Spring Reservoir watershed) water balance schematic.
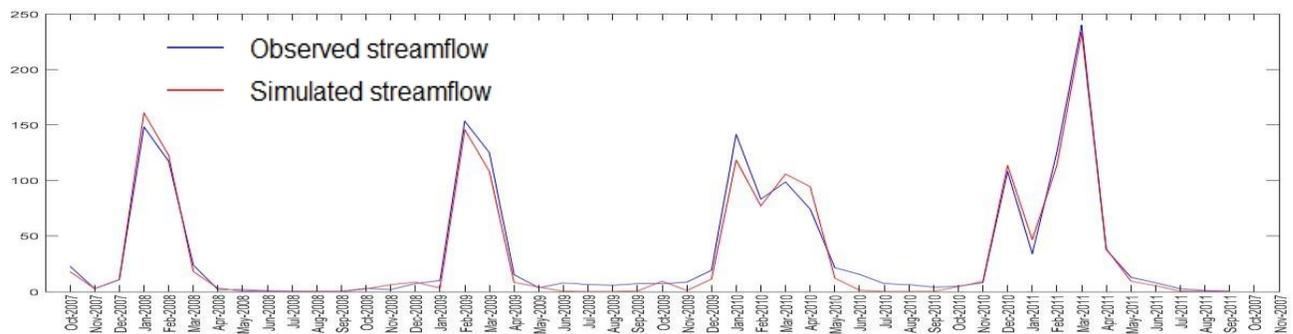
Figure G-11. Reconstructed observed and simulated streamflow for Crystal Spring Reservoir subwatershed with, NSE = 0.98.

159

G.2.2. Basin 2 − San Mateo Creek watershed

Q8 − Q13 − Q4 = 0

Qoutlet1 = Qobs

No flows from system model

Q11 = 0

● **Qoutlet2**

Q10 = 0

Q4: Crystal Springs Reservoir Pumping to San Andreas Reservoir
Q8: San Andreas Reservoir Inflow from Crystal Springs, San Mateo Creek & Pilarcitos
Q10: San Andreas Reservoir Release to San Mateo Creek
Q11: San Andreas Reservoir Spill to San Mateo Creek
Q13: Pilarcitos Reservoir Release to San Andreas Reservoir

Figure G-12. Basin 2 (San Mateo Creek watershed) water balance schematic
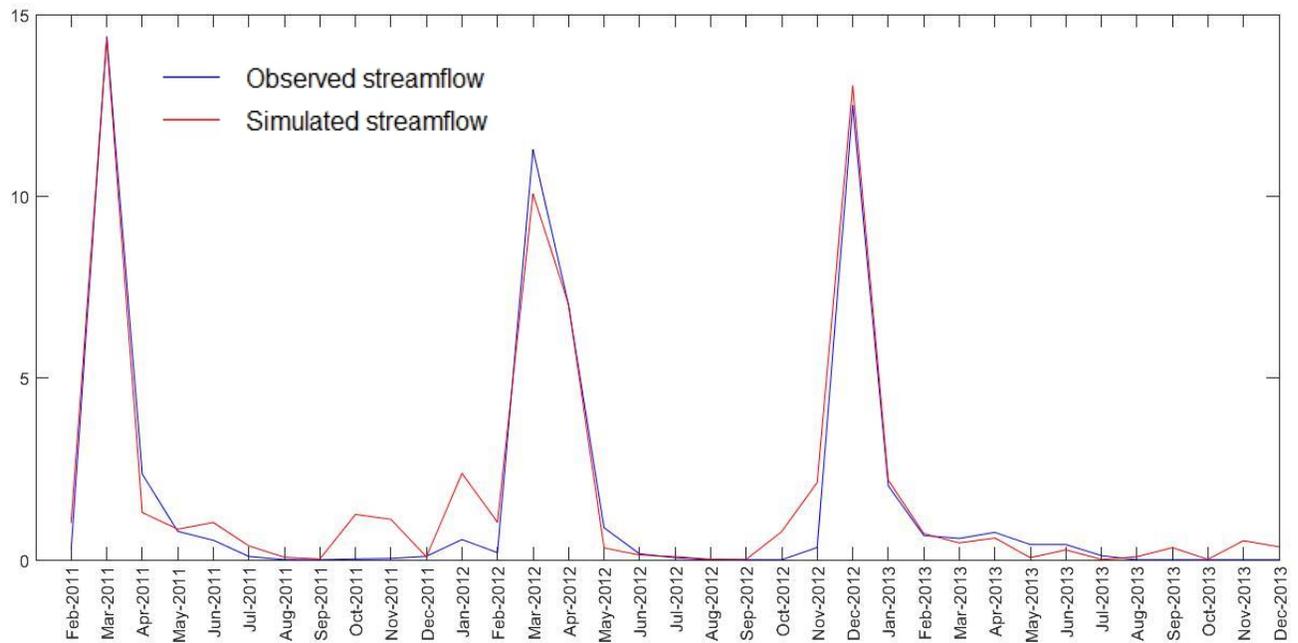
Figure G-13. Reconstructed observed and simulated streamflow for San Mateo Creek subwatershed, with NSE = 0.96.

160

G.2.3. Basin 3 − Pilarcitos subwatershed

Q12: Pilarcitos Reservoir Inflow
Q17: Stone Dam Inflow
Q23: Total Stream Release from Pilarcitos Reservoir
Q24: Total Stream Release from Stone Dam
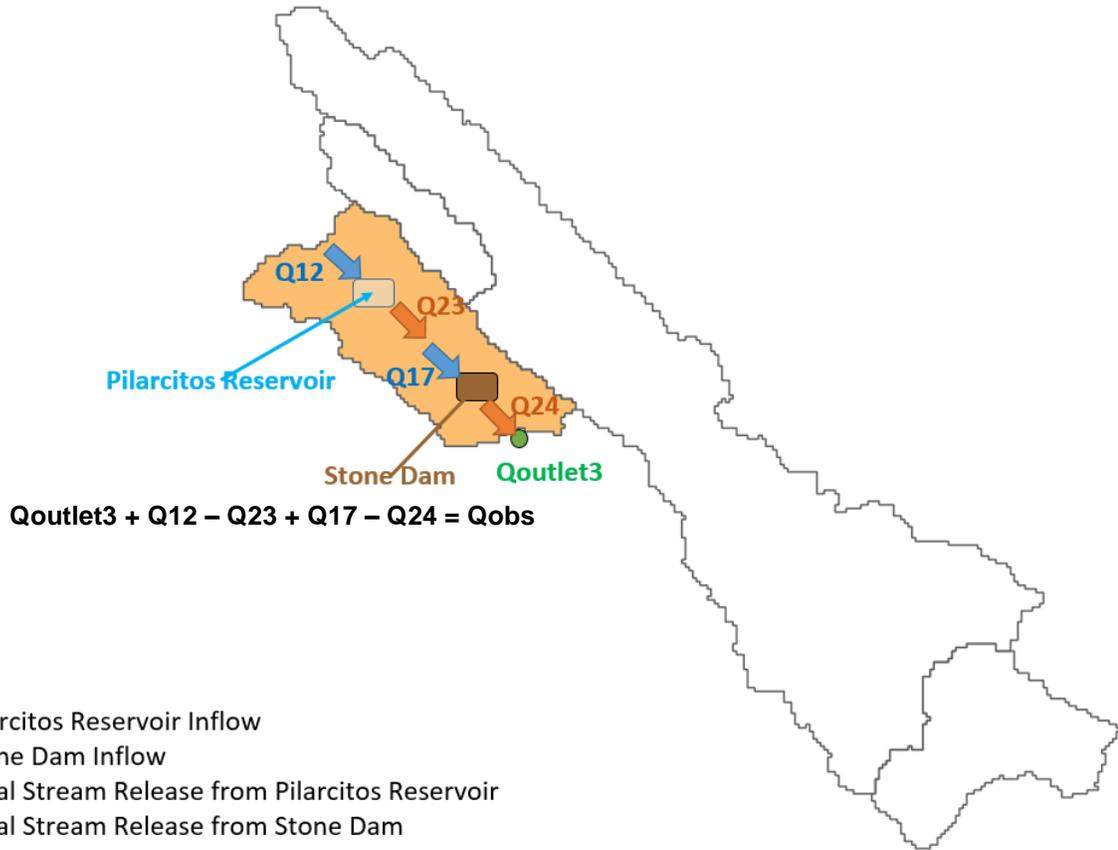
Figure G-14. Basin 3 (Pilarcitos subwatershed) water balance schematic.
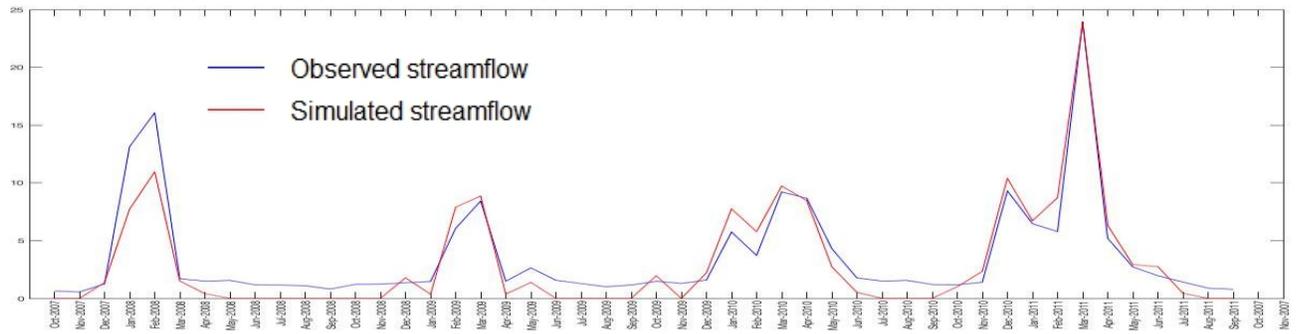
Figure G-15. Reconstructed observed and simulated streamflow for Pilarcitos subwatershed, with NSE = 0.91.

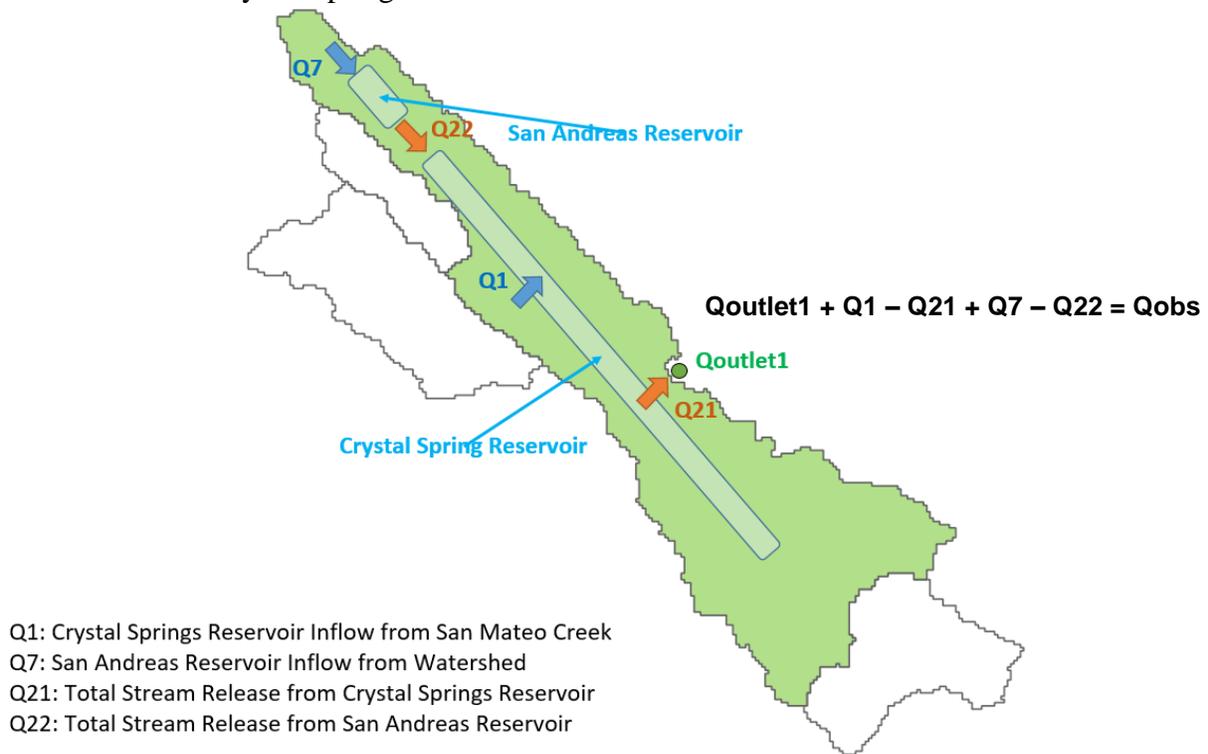## G.2.4. Basin 4 − Above Upper Crystal Spring watershed
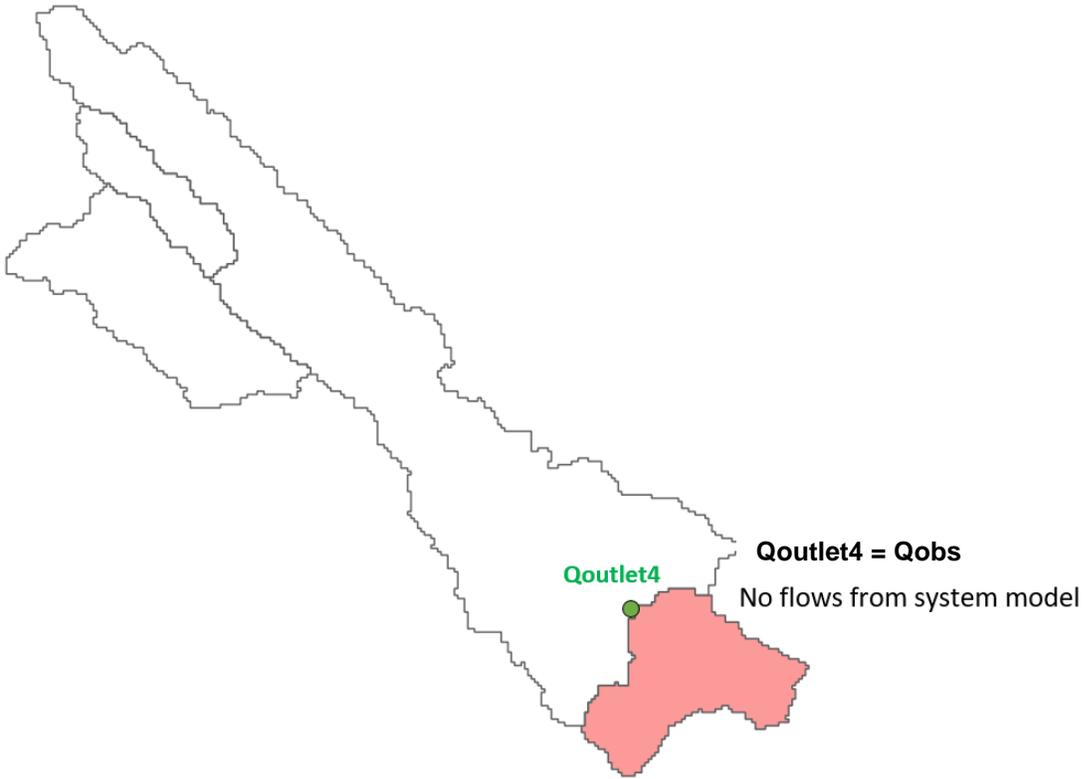
Figure G-16. Basin 4 (Above Upper Crystal Spring watershed) water balance schematic.
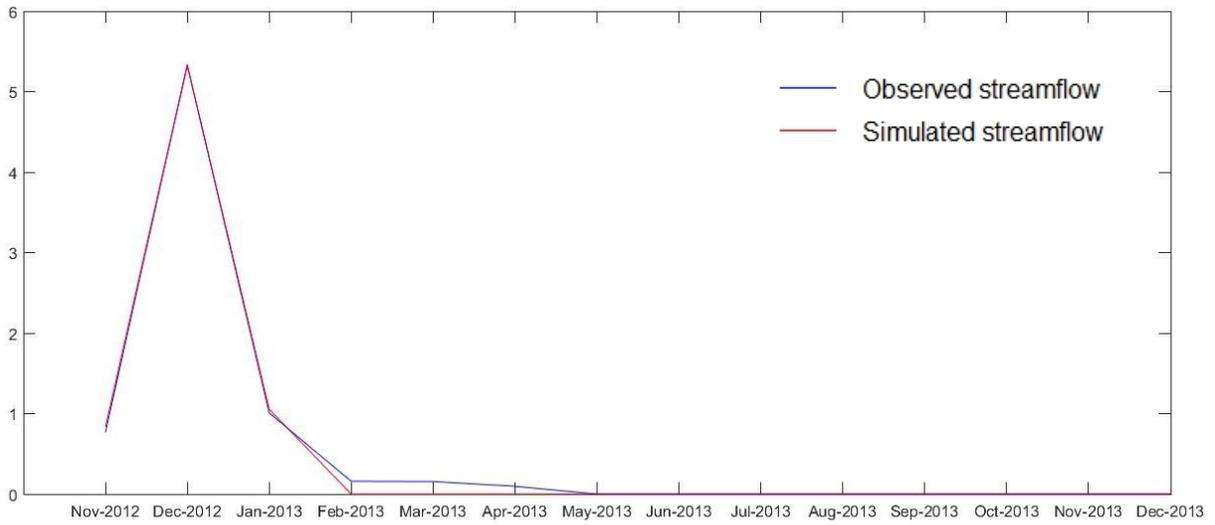
Figure G-17. Reconstructed observed and simulated streamflow for Above Upper Crystal Spring subwatershed, with NSE = 1.0.

H. Appendix – Don Pedro PRMS Parameters

H.1. Cover Type (cov_type)

The University of Maryland's Department of Geography generated a global land cover classification raster in 1998 with images from the AVHRR satellites between 1981 and 1994. The 1 kilometer pixel resolution raster was used to determine the coverage type of the entire Don Pedro watershed. With multiple land types available within each HRU, the majority type present in each HRU is considered the dominant coverage type.
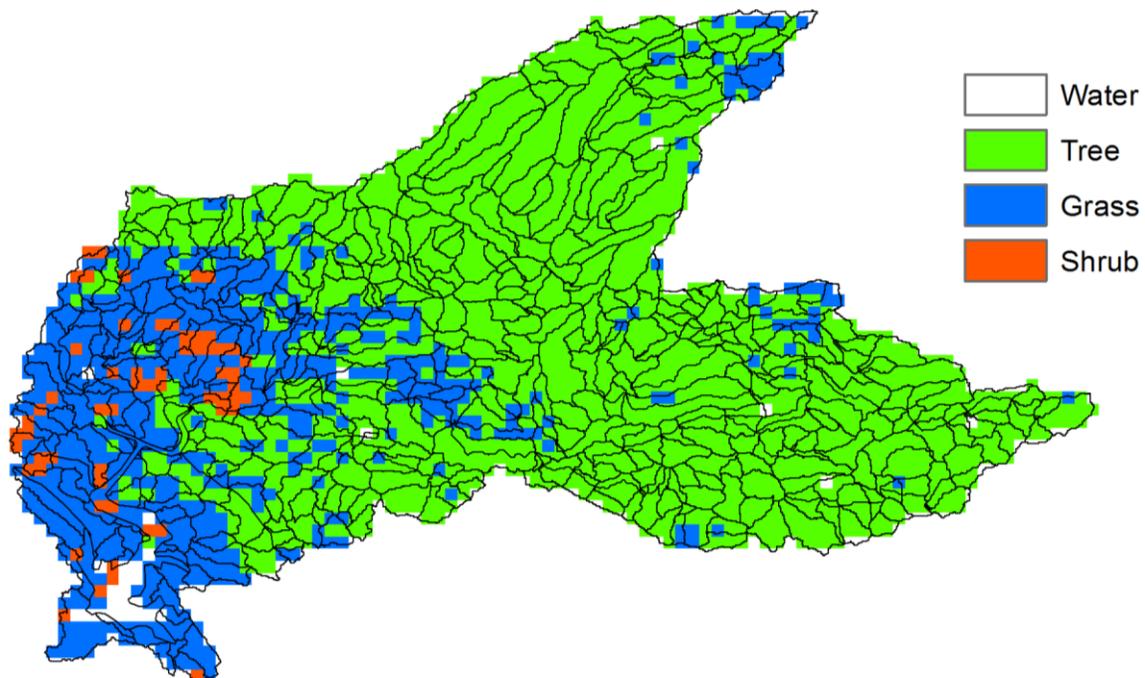
Figure H-1 Land Coverage Breakdown in Don Pedro

H.2. Cover Density Summer (covden_sum) and Winter (covden_win)

The ratio of the cover density during the summer is obtained from the ratio of the grid cells present of each type of cover within the HRU. The ratio of the area of the trees over the total area is the value of the cover density during the summer. However, as the winter cover density data is unavailable and the periods has less trees with leaves, the cover density should be less than the summer. To illustrate this relationship, a calibrated value is multiplied by the cover density of the summer, which results in the following function:

$$covden\_win = Value \times covden\_sum$$

163

H.3. Snow Intercept (snow_intcp) and Rain Intercept (wrain_intcp)

The snow_intcp parameter is the snow interception storage capacity for the major vegetation type in each HRU. The wrain_intcp is the winter rain interception storage capacity for the major vegetation type in each HRU. These parameters are difficult to collect data on, but they are related to cover density of the HRU. Therefore, a calibrated value is multiplied by the cover density for each HRU and is calculated as:

$$snow\_intcp = Value \times covden\_win$$
$$wrain\_intcp = Value \times covden\_win$$

H.4. Jensen-Haise Coefficient per HRU (jh_coef_hru)

The jh_coef_hru value is the air temperature coefficient used in Jensen-Haise potential evapotranspiration model used for each HRU. This estimation is based off of the following equation:

$$\mathbf{jh\_coef\_hru}_{HRU} = 27.5 - \left[0.25 \times \left(\rho_{high\_temp} - \rho_{low\_temp}\right)\right] - \frac{\mathbf{hru\_elev}_{HRU}}{1000}$$

Where:

$\rho_{high\_temp}, \rho_{low\_temp}$ are the saturation vapor pressure, in millibars, for the mean maximum and minimum air temperature for the warmest month of the year.

This relationship is expressed through the Tetens saturation vapor pressure equation, which is written as:

$$P = 0.61078 \exp\left(\frac{17.27T}{T + 237.3}\right)$$

Where:

T is temperature in degrees Celsius (℃)
P is in kilopascals (kPa)

ρ_{high_temp} and ρ_{low_temp} were found by using the Tetens saturation vapor pressure equation on the basin-wide average maximum and minimum temperature. This resulted in a jh_coef_hru value based on the change in HRU elevation and is compared to Hetch Hetchy values.
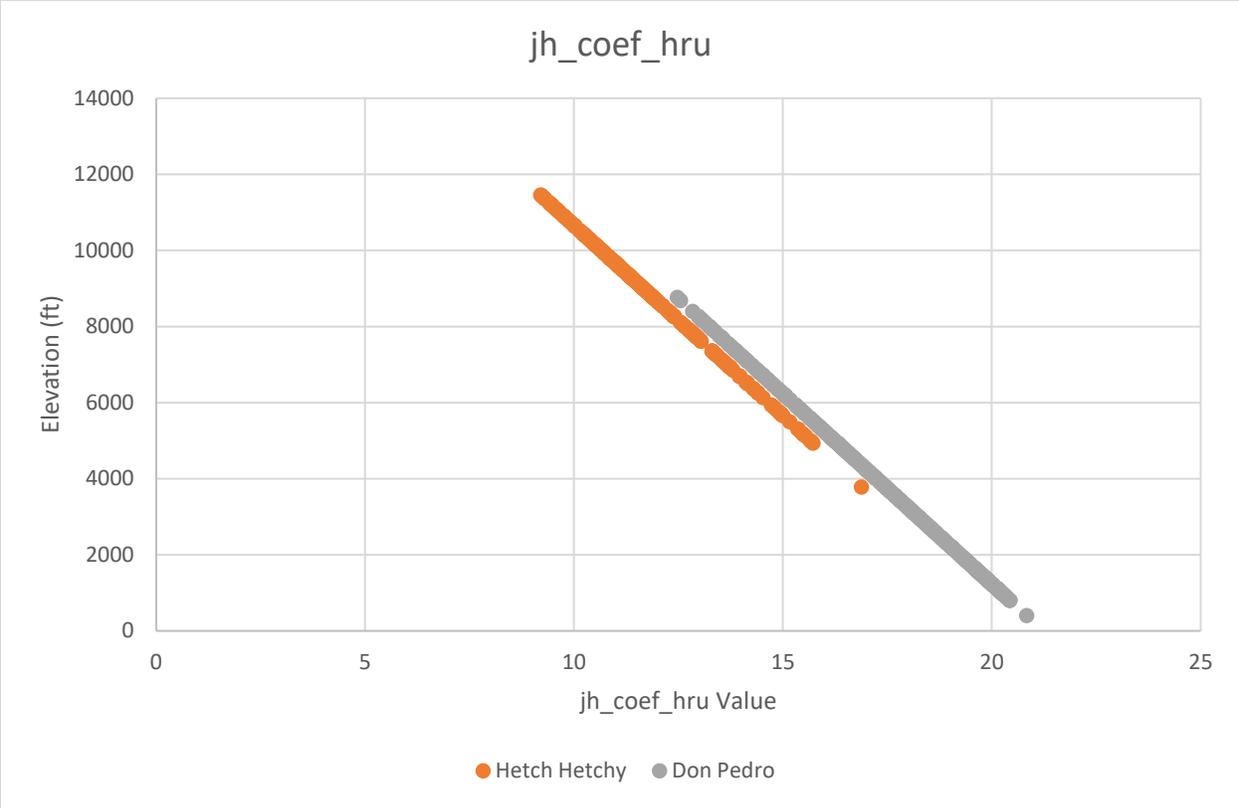
Figure H-2 Jensen-Haise coefficient for HRU based on Elevation

I. Appendix – Evapotranspiration and jh_coef_hru

The Precipitation-Runoff Modeling System (PRMS) model for the Upcountry watersheds uses the Jensen-Haise evapotranspiration formulation. This method uses two additional coefficients, **jh_coef** and **jh_coef_hru**, to calculate the potential evapotranspiration for each HRU. The **jh_coef_hru** parameters, calculated for each HRU, are based on the saturation vapor pressure. This document answers the question of what extent recalculating the **jh_coef_hru** for each different climate scenario has on the outputs of the PRMS hydrologic models.

Quick Key Findings

- Updating the $jh\_coef\_hru_{HRU}$ increases the representative scenarios to the historic average.
- However, the increased performance to representative scenarios are minor compared to other significant factors.

Jensen-Haise Formulation

The PRMS models are configured to use a modified Jensen-Haise formulation to compute potential evapotranspiration. PET is computed as a function of air temperature, solar radiation, and two parameter values labeled within the model as **jh_coef** and **jh_coef_hru**, which are estimated by using regional air temperature, elevation, and saturation vapor pressure. The calculations are defined by:

$$potet_{HRU} = jh\_coef_{month} \times (tavgf_{HRU} - jh\_coef\_hru_{HRU}) \times \frac{swrad_{HRU}}{2.54 \times \lambda_{HRU}}$$

$$\lambda_{HRU} = 597.3 - (0.5653 \times tavgf_{HRU})$$

where

λ_{HRU} is the latent heat of vaporization on the HRU, calories/gram

The air temperature parameter, **jh_coef_hru**, used in the Jensen-Haise potential evapotranspiration formula is calculated for each HRU by:

$$jh\_coef\_hru_{HRU} = 27.5 - \left[0.25 \times \left(\rho_{high\_temp} - \rho_{low\_temp}\right)\right] - \frac{hru\_elev_{HRU}}{1000}$$

where

ρ_{high_temp},   is the saturation vapor pressure, in milibars, for the mean maximum air temperature for the warmest month of the year; and

ρ_{low_temp},   is the saturation vapor pressure, in milibars, for the mean minimum air temperature for the warmest month of the year.

The saturation vapor pressure relationship is expressed through the Tetens formulation, which is written as:

$$\rho = 0.61078 \exp\left(\frac{17.27T}{T + 237.3}\right)$$

where

T is temperature in degrees Celsius (°C)
ρ is in kilopascals (kPa)

The Hetch Hetchy temperature gauge station was used to calculate the $jh\_coef\_hru_{HRU}$ coefficient for the saturation vapor pressure, where the warmest month of the year is found to be July. The mean maximum and minimum air temperature for July was calculated for the coefficient.

Methodology

A series of calibrated PRMS hydrologic model runs were completed for Hetchy Hetchy, Cherry, Eleanor, and Don Pedro watersheds with the following two scenario conditions for a single realization:

1. $jh\_coef\_hru_{HRU}$ remains consistent with historical values (is not updated as new scenarios/climates/data is entered)
2. $jh\_coef\_hru_{HRU}$ was updated with each climate scenarios and realizations (new climate inputs would update the coefficients)

The outputs were then compared to the related scenarios through three major impacted outputs from the PRMS models:

1. Potential Evapotranspiration (basin-wide and HRU-wide)

2. Streamflow (basin-wide)

Results

  To reduce the number of computational runs of the hydrologic model, the $jh\_coef\_hru_{HRU}$ were updated each PRMS run for realization 1. Climate response surfaces were generated comparing the annual potential evapotranspiration for the Hetch Hetchy model and annual streamflow for the Hetch Hetchy model and the combined Upcountry hydrologic models. Figure 1 and 2 shows the impact of the updated $jh\_coef\_hru_{HRU}$ coefficients as an increase in potential evapotranspiration. The difference between the two configurations does show a difference between the two methods, however both outputs are within reasonable observed evapotranspiration values as observed in figure 3 for the Hetch Hetchy watershed.

Figure 1. Climate response surface for annual potential evapotranspiration in Hetch Hetchy with adjustments for realization 1
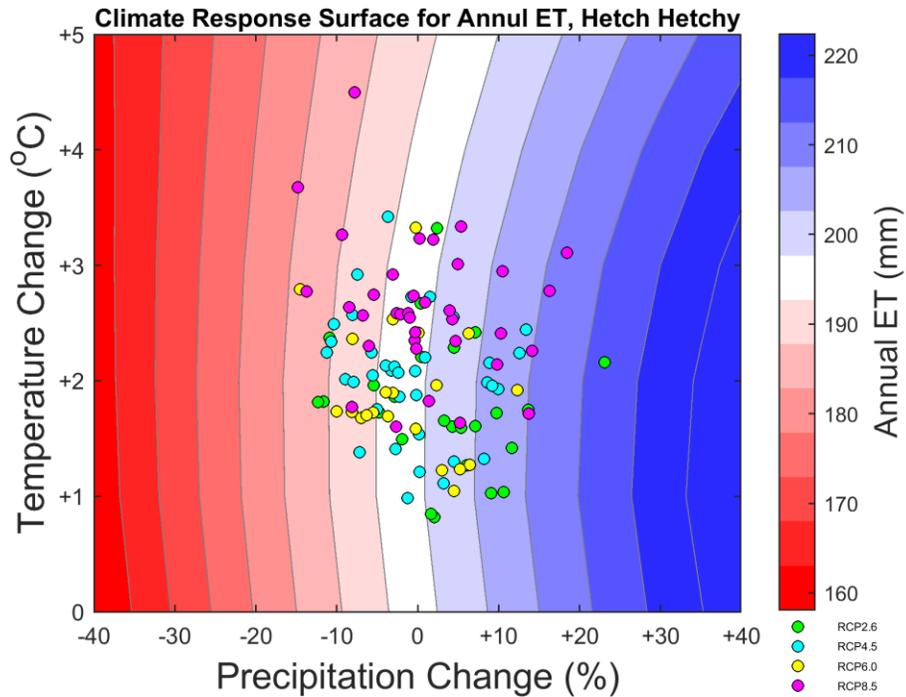
Figure 2. Climate response surface for annual potential evapotranspiration in Hetch Hetchy without adjustments for realization 1
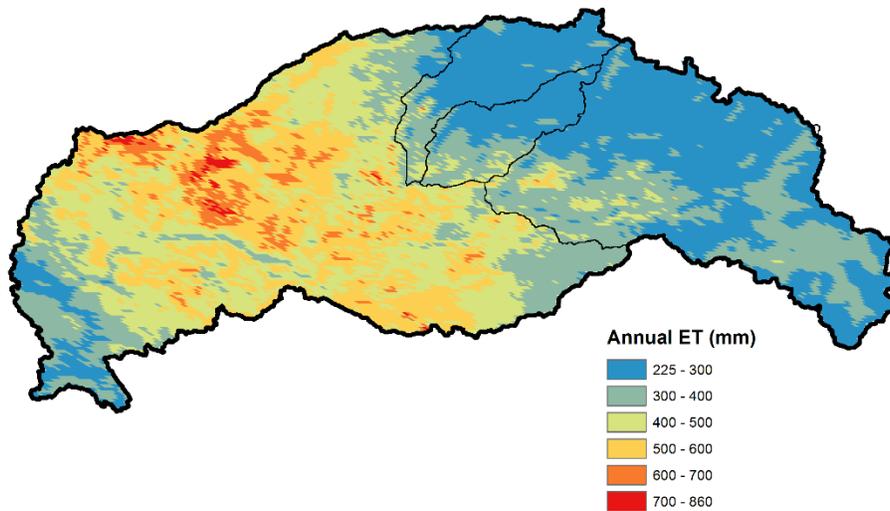
Figure 3. Historic annual potential evapotranspiration values aggregated from MODIS

When observing the impact of the updated coefficients however, the differences were observed to be much less severe. While the relative historic annual streamflow is observed to be closer to the no change scenario when the coefficient is updated, the absolute difference in streamflow is not as significant and is signified even by the buckets of each streamflow

partitions. This concludes that updates to $jh\_coef\_hru_{HRU}$ is an increase in potential evapotranspiration.

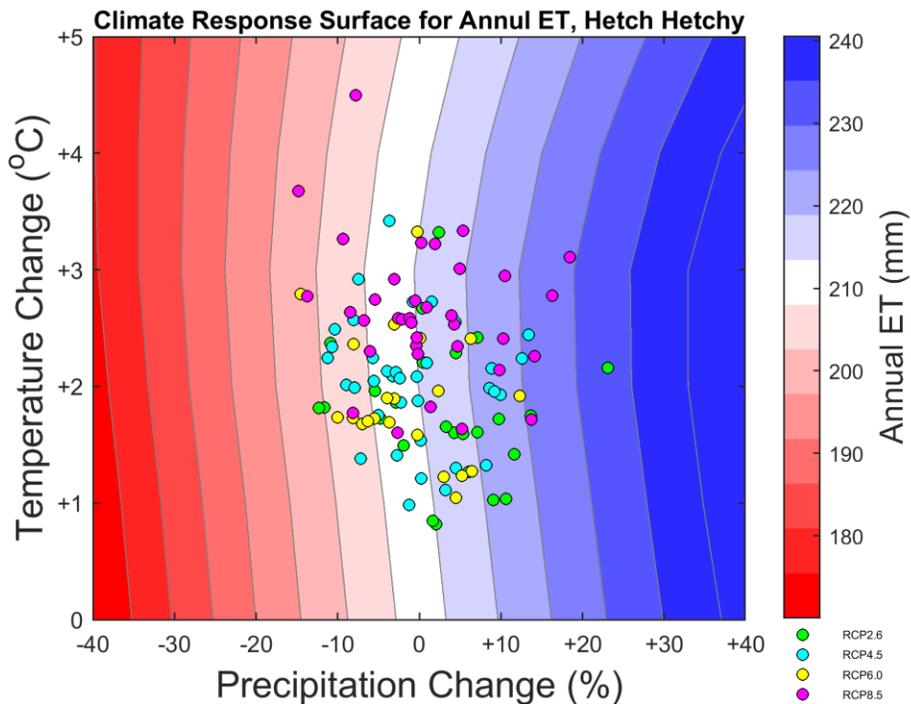**Figure 4.** Climate response surface for annual streamflow in Hetch Hetchy with adjustments for realization 1

Figure 5. Climate response surface for annual streamflow in Hetch Hetchy without adjustments for realization 1
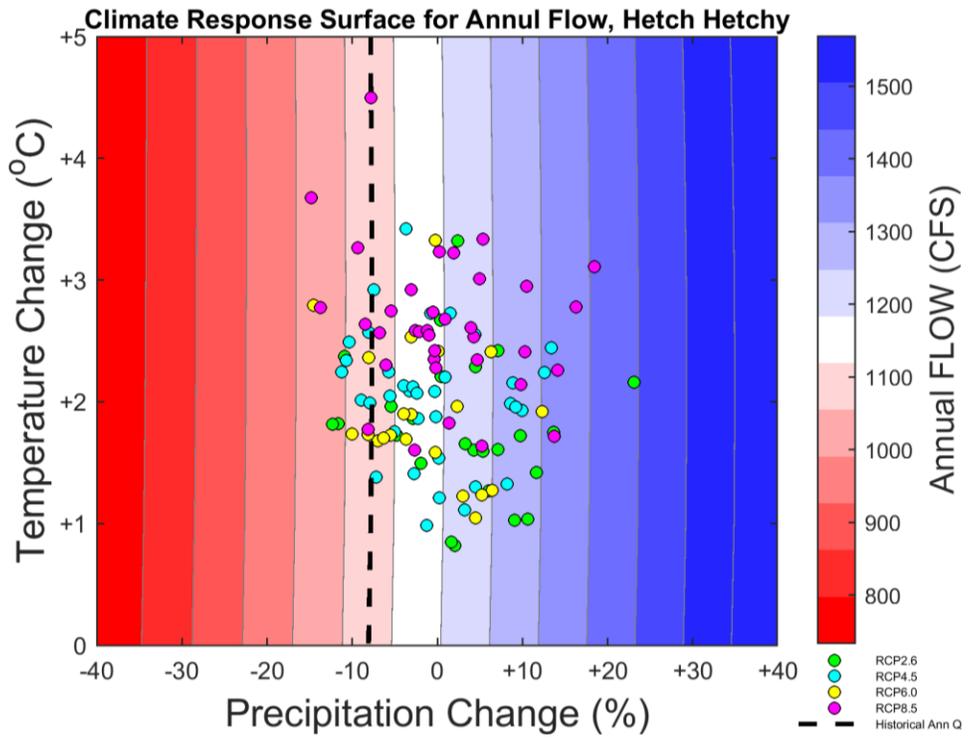
Figure 6. Climate response surface for annual streamflow in Hetch Hetchy without adjustments for all realizations

When comparing the climate response surface for annual streamflow in Hetch Hetchy in realization 1 to the climate response surface for all realizations without adjustments, the difference between the two scenarios are even more marginalized. These differences are at an even smaller scale when The Nash-Sutcliffe Efficiency (NSE) is calculated for the daily streamflow values to be 0.98. The three plots conclude that while updating the $jh\_coef\_hru_{HRU}$ may improve outputs closer to historic conditions, the improvement is marginal and is still able to reflect the natural response signal of the climate within the hydrologic models for all realizations.
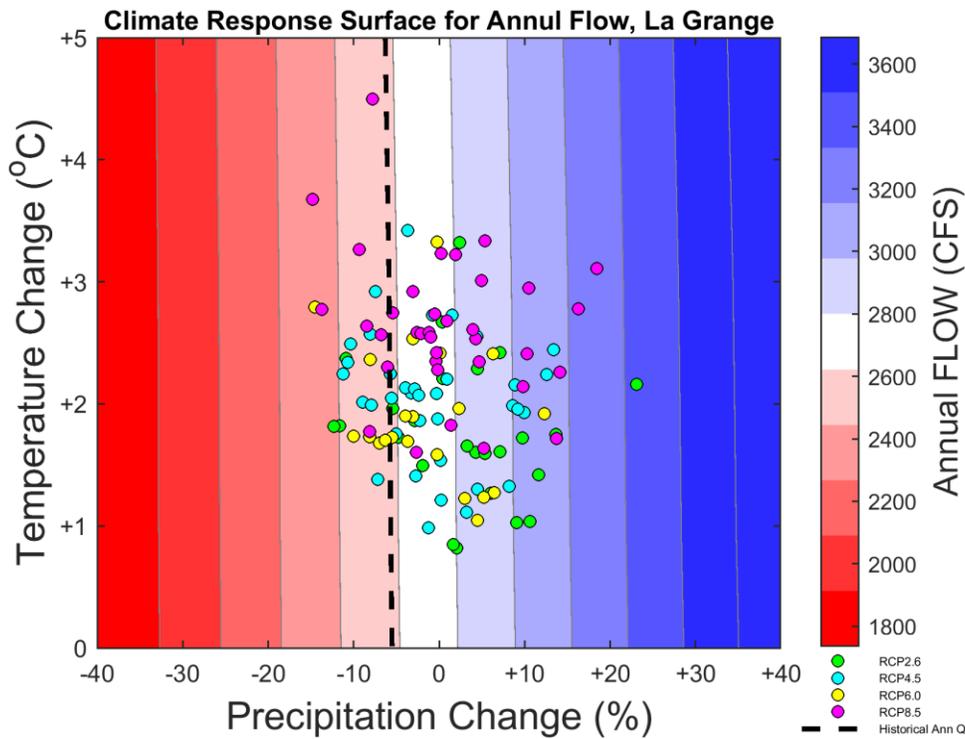
Figure 7. Climate response surface for annual streamflow in La Grange with adjustments for realization 1

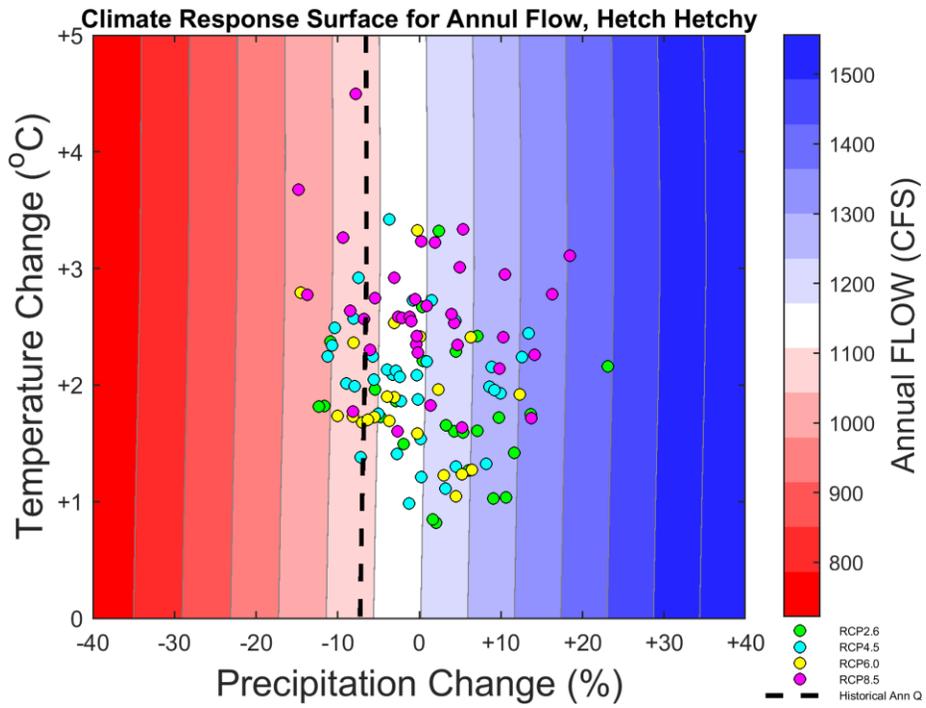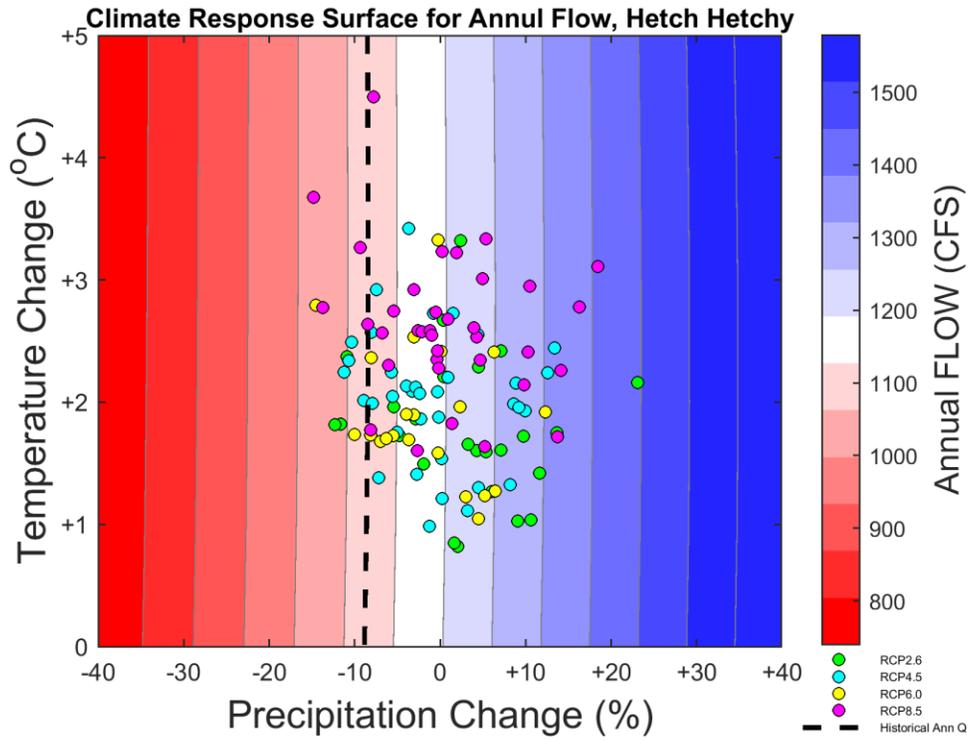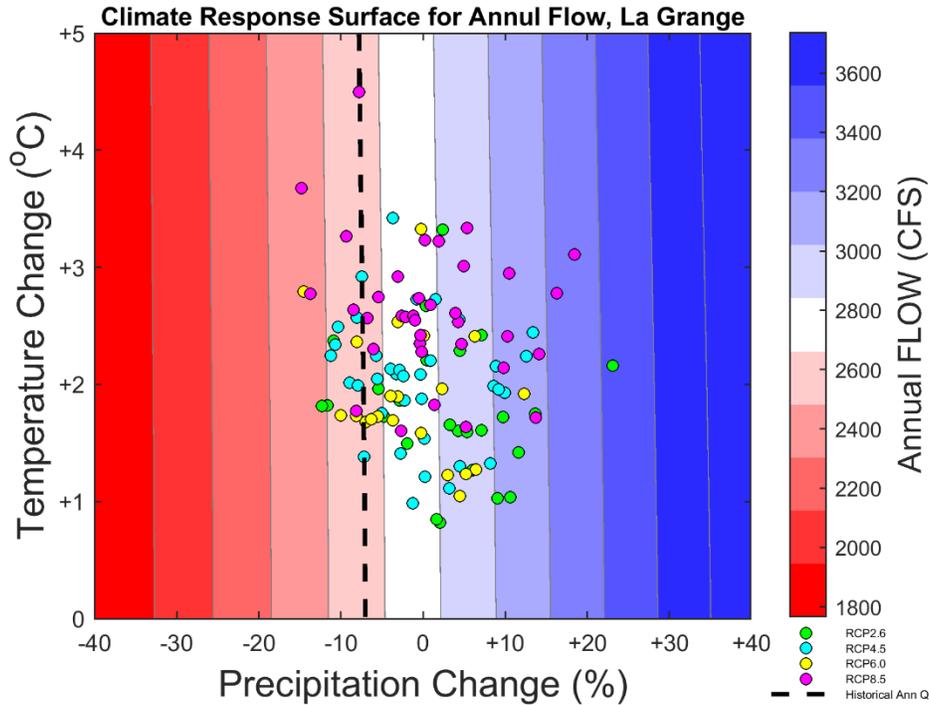**Figure 8.** Climate response surface for annual streamflow in La Grange without adjustments for realization 1

Recall the following relationship for the Upcountry hydrologic models:

$$\text{La Grange } = \text{Hetch Hetchy} + \text{Cherry and Eleanor } + \text{Between Hetchy and Don Pedro}$$

The climate response surfaces figure 7 and 8 for streamflow at La Grange continues to show similar trends in a minor increase in performance with no change in the major trends of the system. While there is an improvement to the models' performances as the $jh\_coef\_hru_{HRU}$ is updated per each climate scenario, they are negligible compared to other outside influences. However, this should be considered when interpreting the outputs of the PRMS hydrologic models.

AET sensitivity to the PRMS parameter "SM_MAX"
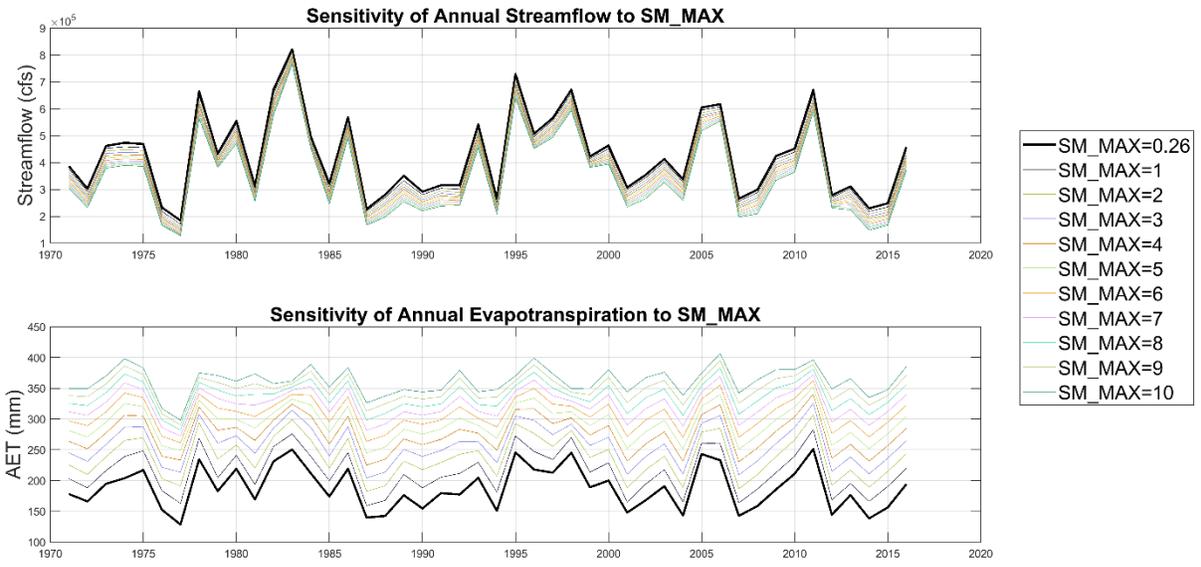
Figure 9. Sensitivity of Hetch Hetchy annual streamflow to SM_MAX.
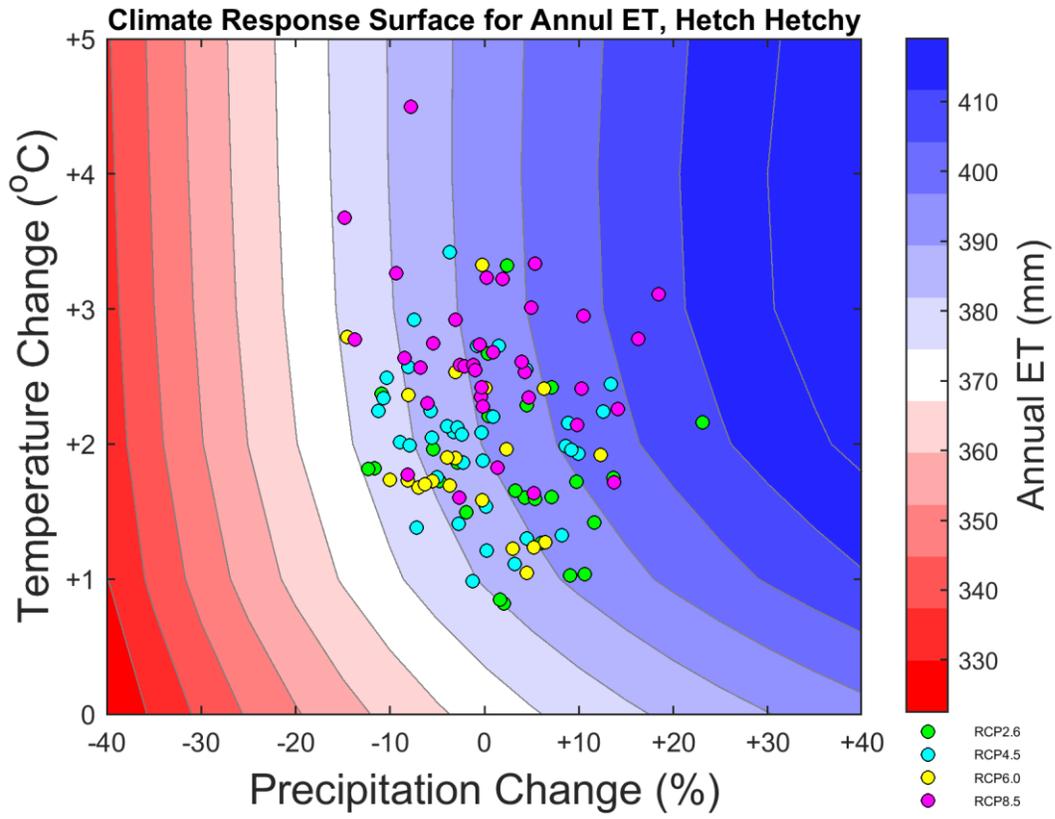
Figure 10. Climate response surface of Hetch Hetchy annual ET with the maximum value of the SM_MAX.

J. Appendix – Current Precipitation Index (CPI) Method Writeup

G.1. Introduction

San Francisco Public Utilities Commission's Hetch Hetchy Water and Power (HHWP) Division maintains and utilizes a version of the USGS Precipitation Runoff Modeling System (PRMS) for internal use in simulating natural Tuolumne River inflows to reservoirs owned and operated by the City and neighboring irrigation districts. Due to system variability, input and parameter uncertainty, long-term period of record calibration about the central tendency of inflows, and incomplete process representation in the model, biases arise when examining results against daily, monthly, and annually calculated observations of reservoir inflow. Thus, an error correction model is desired to improve model performance statistics, allowing for increased reliability for simulating system vulnerability to extreme weather variability and climate change. Here, a post-processing routine developed for such improvements is described. The post-processing model was developed by HHWP to estimate appropriate corrections to modeled streamflow and is based on average observed meteorology throughout a calibration period (Water Years 1972 to 2015). Meteorology inputs include temperature and precipitation from nine weather stations in the Sierra Nevada foothills and mountains ranging from 938 feet to 9,200 feet in elevation, within and in the periphery of the Tuolumne River Watershed from which HHWP diverts a portion of flow for municipal water supply and hydropower generation.

G.2. Tuolumne River System and Model Components

The post-processing model is derived from historical simulated residual errors in flows at four reservoirs in the Sierra Nevada: 1) Hetch Hetchy Reservoir, 2) Cherry Reservoir, 3) Lake Eleanor, and 4) Don Pedro. The first three reservoirs are considered components of the "Upcountry" system, their outlets joining at various confluences downstream before the main Tuolumne River flows into Don Pedro Reservoir. Cherry Reservoir and Lake Eleanor inflows are treated as a single node since the two bodies are connected via a tunnel through which water is frequently transferred from Eleanor to Cherry. Don Pedro is operated by the Modesto and Turlock Irrigation Districts and its outlet stream gage (La Grange) is the point at which full natural flow on the Tuolumne River is calculated and water rights partitioning is determined. The streamflow observations and simulations for Don Pedro are not direct measurements nor simulations of Don Pedro inflows explicitly, but rather they represent the intervening flow between the three upcountry reservoirs operated by San Francisco (Hetch Hetchy, Cherry, and Eleanor) and Don Pedro Reservoir downstream. Thus, the calculation for simulated unimpaired flow at La Grange is:

$$Q_{LaG} = Q_{HH} + Q_{CHEL} + Q_{DP} \qquad\qquad\qquad (G\text{-}1)$$

where Q_{LaG} represents simulated full natural flow on the Tuolumne River at La Grange, Q_{HH} is inflow to Hetch Hetchy Reservoir, Q_{CHEL} is combined inflow to Cherry Reservoir and Lake Eleanor, and Q_{DP} is the intervening flow above Don Pedro as described above. Observed natural inflows to all reservoirs, along with unimpaired flow at La Grange, are calculated daily and used for initial model calibration as well as the calibration of the post-processing routine. Observed intervening flows for Don Pedro are inferred by rearranging Eqn. 1 for the observed mass balance:

$$Q_{DP}^o = Q_{LaG}^o - Q_{HH}^o - Q_{CHEL}^o \qquad\qquad\qquad (G\text{-}2)$$

where the superscript o denotes observed flow. Residual errors are calculated as the simple difference between observed and simulated daily flows at each node and timestep:

$$\varepsilon Q_{HH} \quad = Q_{HH}^o - Q_{HH} \tag{G-3}$$

$$\varepsilon Q_{CHEL} = Q_{CHEL}^o - Q_{CHEL} \tag{G-4}$$

$$\varepsilon Q_{DP} \quad = Q_{DP}^o - Q_{DP} \tag{G-5}$$

where ε denotes the residual error between the observation and simulation estimate for each timestep. It is these errors that the post-processing model seeks to correct.

G.3.1. Post-Processing Model Input Data

The post-processing model described here attempts to correlate residual model error with observed meteorological data to nudge simulated results closer to observations, generating "corrected" reservoir inflow estimates. The model is derived from the relationship between residual error and two meteorological input indices. A key assumption is that when forced with input data from future forecasts or climate realizations, the identified error structure remains unchanged.

To begin, observed meteorological data from the PRMS input file are extracted and used to generate index temperature and precipitation time series for use in the post-processing routine. For each input station, a daily average temperature is taken as the mean of the maximum and minimum observed temperature for each timestep. An index temperature is then computed as the average across all nine stations' means:

$$T_{\mu,t} = \frac{1}{n}\sum_{j=1}^{n} \frac{T_{jmax}+T_{jmin}}{2} \tag{G-6}$$

Where T_{jmax} and T_{jmin} denote maximum and minimum daily air temperature at the jth station, respectively, n is the number of stations (9), and $T_{\mu,t}$ is the resulting index air temperature value for timestep t. To generically represent an average antecedent heat index for the basin, these temperature index values are further computed as 15-day trailing window summations:

$$T_{ss,t} = \sum_{i=t-14}^{t} T_{\mu,i} \tag{G-7}$$

where $T_{ss,t}$ denotes an index temperature summation at time t. This variable becomes the first of the two predictors that serve as inputs to the post-processing model structure.

Next, an index precipitation value is computed across all PRMS input stations at each timestep, in a similar fashion to the calculation for temperature. The simple average is computed as:

$$P_{\mu,t} = \frac{1}{n}\sum_{j=1}^{n} P_j \tag{G-8}$$

where P_j denotes the measured precipitation at station j and $P_{\mu,t}$ is the computed precipitation index value at timestep t. To generically represent basin wetness, the precipitation index time series is used to compute a "current precipitation index" (CPI) at each timestep (Smakhtin and Masse, 2000):

$$CPI_t = CPI_{t-1} \cdot \beta + P_{\mu,t} \tag{G-9}$$

where CPI_t is the current precipitation index for day t and β is a recession parameter. CPI_t becomes the second predictor for the post-processing model. The CPI is, in and of itself, a method proposed for

streamflow estimation and requires a "spin-up" period prior to calibration and simulation. CPI values were found to be insensitive to spin-up periods longer than about one year, for values of β relatively close to 1 (much shorter for smaller β). Since it is expected that model errors are often proportional to streamflow magnitude, the CPI variable is able to capture some of this heteroscedasticity, in addition to being a skillful descriptor of overall basin wetness. The CPI parameter β was fit iteratively for each model, as described later in the document.

G.3.2. Post-Processing Model Transformation and Innovation

Before quantifying the relationship between the described meteorological indices and modeled residual errors, the simulation and observation time series are first transformed to reduce heteroscedasticity and skewness in the residual error structures (Woldemskel et al., 2018), using the Box-Cox (BC) transformation (Box and Cox, 1964). For the case of observed Hetch Hetchy Reservoir inflow, the BC transformation is:

$$Z(Q_{HH}^o; \lambda, c) = \frac{(Q_{HH}^o + c)^\lambda - 1}{\lambda} \tag{G-10}$$

where c is a fixed parameter equal to one one-hundredth of the average observed Hetch Hetchy natural streamflow over the calibration period (WY 1972-2015) and λ is a power parameter that is fixed at 0.2 for all simulations, per the recommendation of McInerny et al. (2017). The transformed streamflows are computed for all observed and simulated time series as above, with values for c, variable by catchment as denoted in Table G-1.

Table G-1. c-values for three simulation catchments.

	Hetch Hetchy	Cherry-Eleanor	Don Pedro
Box-Cox c Parameter	10.322	6.718	9.863

Model residuals are then recomputed in this new transform space, with the residuals herein referred to as the innovations. The innovations for transformed Hetch Hetchy Reservoir inflow are:

$$\eta_{HH} = Z(Q_{HH}^o) - Z(Q_{HH}) \tag{G-11}$$

where η_{HH} represents the innovation.

G.3.3. Post-Processing Model Calibration

Once the observed and modeled streamflow data are transformed to reduce skewness and heteroscedasticity in the error distribution, the innovations are then fit to corresponding CPI and index temperature summation data by a two-dimensional polynomial function of the form:

$$\eta^*(x, y) = p_1 xy + p_2 y + p_3 x + p_4 \tag{G-12}$$

where η^* denotes the modeled innovation estimate, p_n denote fitting parameters for the polynomial model, and x and y are the input variables CPI and T_{ss}, respectively. Using least squares regression, parameters are estimated for three temporal intra-seasonal windows for each watershed model, allowing for dynamic parameterization for wet/dry season representativeness. These windows are defined as the monthly periods November-February, March-June, and July-October.

Once parameter sets are derived, they can be used with any combination of CPI and T_{ss} time series to generate innovation estimates for modeled streamflow. The innovation estimate is then added to a corresponding transformed streamflow simulation estimate. The post-processed streamflow is then achieved by inverting the original transformation (Eqn. G-10). Parameter estimation was repeated across ranges of candidate β values for CPI, selecting for the one producing maximum Nash-Sutcliffe Efficiency scores for each reservoir inflow model. Optimal values are shown in Table G-2.

Table G-2. Beta parameter optima by simulation catchment.

	Hetch Hetchy	Cherry-Eleanor	Don Pedro
β	0.917	0.989	0.994

Since the observation of Don Pedro intervening flow (Eqn. G-2) is estimated by differencing computed flows, its uncertainty is assumed to be higher than that for the upcountry reservoir inflow calculations, since it disregards timing uncertainties and explicit channel reach gains and losses. As such, the model parameters for Q_{HH} and Q_{CHEL} are calculated and fixed, whereas the Q_{DP} are calculated but subjected to additional computation. Since a post-processed model estimate for Q_{LaG} (Eqn. G-1) is sought, parameters for Q_{DP} producing optimized Q_{LaG} are iteratively generated via Monte Carlo simulation, sampling randomly from normal distributions about the mean parameters derived from a first pass. The new set of parameters for Q_{DP} that maximize the Nash-Sutcliffe Efficiency for Q_{LaG} are selected as the final set. Ultimately, a set of 12 parameters is established for each runoff model, four parameters for each of three temporal windows (Table G-3).

Table G-3. Post-processing model parameters by catchment and temporal window.

	Hetch Hetchy			Cherry-Eleanor			Don Pedro		
	Nov-Feb	Mar-Jun	Jul-Oct	Nov-Feb	Mar-Jun	Jul-Oct	Nov-Feb	Mar-Jun	Jul-Oct
p_1	-0.0034	0.0018	-0.0114	-0.0008	-0.0009	-0.0034	-0.0015	-0.0003	0.0000
p_2	-0.0116	-0.0012	0.0102	-0.0082	0.0105	0.0282	0.0192	0.0153	0.0164
p_3	2.5941	-0.2980	7.4146	0.7114	0.6885	3.0922	0.9963	0.4614	0.2473
p_4	2.7138	0.0350	-9.4688	1.5076	-7.8699	-25.8068	-17.7755	-17.0178	-18.7658