

# **Exhibit B**

# Water Resources Research



## RESEARCH ARTICLE

10.1029/2018WR024209

### Key Points:

- Multiple sensitivity analyses varying initial parameters provide a frugal method to account for nonlinearity of integrated hydrologic models
- Calibration results from integrated hydrologic models of irrigated landscapes may be sensitive to parameters of the crop water demand system
- Weakly coupled integrated models are a viable and efficient approach for reproducing observed groundwater-surface water interactions

### Supporting Information:

- Figure S1

### Correspondence to:

T. Harter,  
thharter@ucdavis.edu


### Citation:

Tolley D., Foglia, L., & Harter, T. (2019). Sensitivity analysis and calibration of an integrated hydrologic model in an irrigated agricultural basin with a groundwater-dependent ecosystem. *Water Resources Research*, 55(10), 10.1029/2018WR024209. <https://doi.org/10.1029/2018WR024209>

Received 2 OCT 2018

Accepted 19 JUN 2019

Accepted article online 27 JUN 2019

© 2019 The Authors. 

©2019. The Authors.

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

## Sensitivity Analysis and Calibration of an Integrated Hydrologic Model in an Irrigated Agricultural Basin With a Groundwater-Dependent Ecosystem

D. Tolley<sup>1</sup> , L. Foglia<sup>1</sup>, and T. Harter<sup>1</sup> 

<sup>1</sup>Department of Land, Air, and Water Resources, University of California, Davis, CA, USA

**Abstract** While sensitivity analysis and calibration are common practice in integrated hydrologic modeling, little work has been done to understand how the design of the sensitivity analysis and calibration affects the simulation outcome in these often highly nonlinear models. This is especially true for irrigated agricultural basins with a strong connection between land use, groundwater, and surface water. Using a range rather than a single set of initial parameter values, multiple sensitivity analyses, calibrations, and linearity tests were performed using UCODE\_2014 on the Scott Valley Integrated Hydrologic Model. Calibration results show that parameters related to crop demand and applied irrigation water are most sensitive. Influence statistics show that low streamflow observations provide the most information during model calibration, indicating preference should be given to these observations during model development, sensitivity analysis, and calibration. Importantly, due to the nonlinearity of the integrated model, significant differences are found in results when initial parameter values are sampled from within their respective expected ranges. Estimates for some parameters varied up to an order of magnitude between calibrations, while all produced similar final objective function values, groundwater elevations, and stream flow. Confidence intervals for individual sensitivity analyses and calibration runs only spanned a fraction of the ensemble estimated parameter range across multiple runs. Our work suggests that a calibration design with multiple sensitivity analyses and calibrations of integrated hydrologic models, each using one of several widely varying sets of initial values, provides a frugal approach to identify parameters across the global parameter space.

### 1. Introduction

Groundwater and surface water resources are increasingly being stressed due to changes in population, land use, management practices, and climate (Van Roosmalen et al., 2009; Hanson et al., 2012; Taylor et al., 2012; Kløve et al., 2014; Dettinger et al., 2015). In order to gain insight and understanding of system behavior and complex feedbacks inherent between groundwater, the landscape, and surface water, numerical models that approximate physical flow processes are typically used. Although interactions between groundwater and surface water and between groundwater and the irrigated landscape have been known since the inception of numerical modeling, these systems have traditionally been handled separately with little to no connection between them. This was primarily due to computational limitations and different response times and spatial scales between surface and subsurface routed water (Prudic, 1989; Brunner et al., 2010; Unland et al., 2013; Singh, 2014).

A variety of models have been developed in the last two decades that simulate the flux of water between the surface and subsurface to varying degrees. These include (1) analytical or spreadsheet models (S. S. Papadopoulos and Associates, Inc, 2000; Manghi et al., 2012; Foglia, McNally, & Harter, 2013; Lane et al., 2015), (2) iteratively coupled models like the Integrated Water Flow Model (IWF; California Department of Water Resources, 2016a, 2016b), MODFLOW (Harbaugh, 2005) with the stream package (STR; Prudic, 1989), streamflow routing (SFR) package (Prudic et al., 2004), and/or farm process package (FMP2; Schmid et al., 2000; Schmid & Hanson, 2005), MODFLOW One-Water Hydrologic Model (MF-OWHM; Hanson, Boyce, et al., 2014), and GSFLOW (Markstrom et al., 2008), and (3) fully coupled models such as ParFlow (Ashby & Falgout, 1996; Kollet & Maxwell, 2006) and Hydrogeosphere (Therrien & Sudicky, 1996; Brunner & Simmons, 2012). Iteratively coupled models solve multiple systems of equations (e.g., saturated flow, unsaturated flow, and surface flow) at each time step and iteratively pass fluxes between these systems until

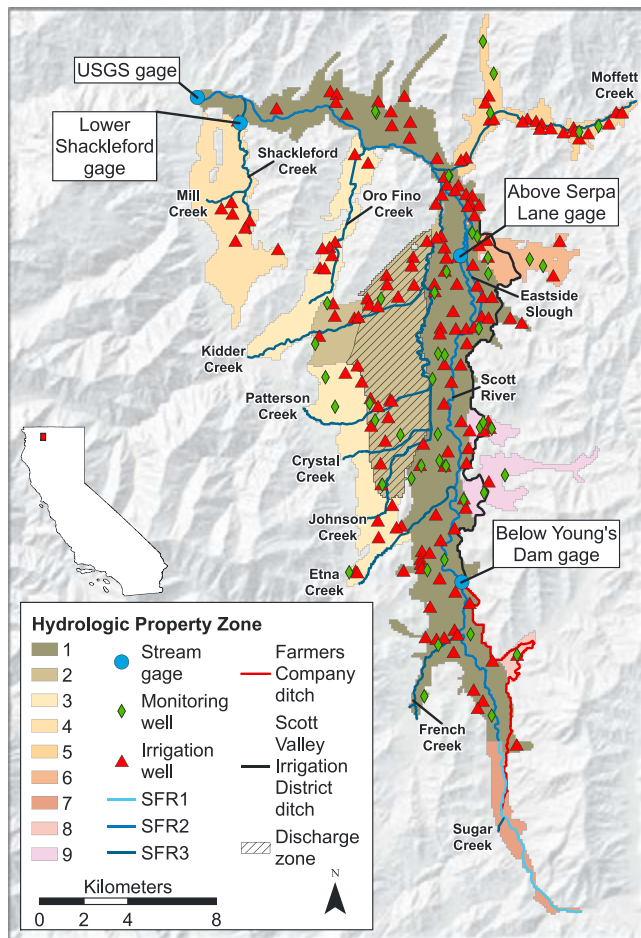
convergence is achieved. Fully coupled models assemble the entire system into a single system of equations. This is computationally much more expensive and generally requires a greater degree of parameterization than simpler iteratively coupled models. Superiority of full coupling over iterative coupling is largely application dependent, since full coupling may lead to numerical difficulties resulting from the different nature of the equations used to describe flow (Furman, 2008). Effective water management requires an accurate but also efficient method to represent the hydrologic system with sufficient detail, including appropriate coupling of subsystems, to answer specific questions, yet simple enough to reflect data availability, economic cost, and serve stakeholder needs (La Vigna et al., 2016).

In areas with Mediterranean climate (wet winters and dry summers) and dominated by irrigated agriculture, the connection between groundwater and surface water is highly pronounced due to significant alteration and seasonality of the water resources availability and use in the landscape. Application of surface water for irrigation generally increases groundwater recharge in the spring and early summer (Roark & Healy, 1998; McMahon et al., 2003). Groundwater pumping for irrigation and urban water use can result in streamflow depletion (Chen & Yin, 2001; Barlow & Leake, 2012) and adversely impact groundwater-dependent ecosystems by decreasing flows and increasing temperatures in critical fish habitat (Stark et al., 1994).

Rates of groundwater extraction and recharge in irrigated agricultural areas often lack historic data and are not commonly measured even though they are significant portions of a basin's water budget (Ruud et al., 2004; Yin et al., 2011). Where metering data are not available, groundwater pumping is usually estimated as the difference between crop water demands and applied surface water and precipitation (Ramireddygarri et al., 2000; Pokhrel et al., 2015; California Department of Water Resources, 2016a). Groundwater recharge is estimated by closure of the water balance and may include consideration of vadose zone flux constraints (Ruud et al., 2004; De Silva & Rushton, 2007). These estimates are sometimes made a priori and used as boundary conditions for numerical groundwater-surface-water models. Some integrated hydrologic models allow for a dynamic calculation of groundwater pumping for irrigation and recharge, as they include an iterative coupling between a crop water demand model and a hydrologic model within a time step. This is available in software such as MODFLOW with FMP2, MF-OWHM, and IWFm. Coupling the crop water, vadose zone, groundwater, and streamflow models allows shallow groundwater to influence crop water demands, which can ultimately affect groundwater pumping, recharge, and evapotranspiration (ET) rates in the model. Irrigated agriculture accounts for about 45% of consumptive water use in the United States and about 78% in western states like California (Dieter et al., 2018). Despite this large proportion of consumptive water use and due to lack of regulatory pressures, there are few studies focusing on sensitivity analysis and calibration of integrated groundwater-surface-water models in agricultural areas when compared, for example, to the large body of literature on groundwater contaminant models (Miller & Pinder, 2004; Jousma et al., 2012; Singh, 2014) developed primarily for regularly compliance. Regional integrated models of California's Central Valley have been developed by James M. Montgomery Consulting Engineers (1990), Faunt et al. (2009), and Brush and Dogrul (2013). Basin-scale models are available for the Pajaro Valley near Santa Cruz, CA (Hanson, Schmid, et al., 2014), and Butte Basin near Chico, CA (CDM, 2008), among others.

As for groundwater models developed at contaminant sites, sensitivity analysis and calibration are an essential part of integrated hydrologic model development (Hill & Tiedeman, 2007). However, unlike many contaminant site groundwater flow (and transport) models, integrated hydrologic models include numerous nonlinear cross dependencies between model subsystems (e.g., stream-aquifer flux is a function of the head difference, but heads in the stream and aquifer are also a function of stream-aquifer flux). When multiple sources of nonlinearity are present in a model, their effects are compounded (Cooley, 2004). Also, in significant departure from the site modeling practice, basin models typically use both, larger horizontal cell dimensions (on the order of  $10^3$  m) and coarser time steps (days and months). Perhaps because of the longer-standing, dominant groundwater modeling practice for contaminant sites, integrated hydrologic model development has mostly utilized the same calibration tools (Huntington & Niswonger, 2012; Hanson, Schmid, et al., 2014).

When applied to nonlinear models, however, these tools may be problematic, leading to nonglobal minima in parameter estimation and to ill-defined uncertainty predictions (Hill & Tiedeman, 2007). Clark and Kavetski (2010) and Kavetski and Clark (2010) have shown that hydrologic model nonlinearities make sensitivity analyses and calibration more difficult. New statistical tools in sensitivity analysis and calibration to address some of these difficulties have been developed and demonstrated (e.g., Foglia et al., 2009;



**Figure 1.** Map of Scott Valley showing model domain, major hydrologic features, irrigation wells, and monitoring wells used in the model. Aquifer properties were distributed among nine zones and streambed properties were distributed to tailings reaches (SFR1), nontailings Scott River (SFR2), and tributaries (SFR3). USGS = U.S. Geological Survey; SFR = streamflow routing.

Rakovec et al., 2014). The work shows the need to not only select from alternative parameters but also evaluate alternative model structures, including alternative structures for designing sensitivity analysis and calibration (Clark et al., 2008; Mendoza et al., 2015; Borgonovo et al., 2017). Using frugal methods to overcome potential computational inefficiencies in such a complex model development process has been shown to be a necessary and potentially successful alternative for integrated hydrologic models (Hill et al., 2016).

This paper contributes to the small but increasing collection of integrated hydrologic models developed to support management of irrigated agricultural groundwater basins with groundwater-dependent ecosystems in interconnected streams (e.g., Faunt et al., 2009; Brush & Dogrul, 2013; Hanson, Schmid, et al., 2014; Phillips et al., 2015; Hanson et al., 2018). Specifically, this paper investigates two related questions: What is the relative importance of and uncertainty about a diversity of parameters across the physical submodels in these systems? And how does the design of the sensitivity analysis and calibration affect the assessment of parameter importance and uncertainty?

The Scott Valley Integrated Hydrologic Model (SVIHM) consists of a soil water budget model (SWBM) weakly coupled to a groundwater-surface-water model. UCODE\_2014, a universal inverse modeling software suite, was used to perform sensitivity analyses and calibrate the model, yielding information about parameter importance and uncertainty. In a frugal approach to test the sensitivity analysis and calibration design, we focus on examining the influence of the initial parameter values. Varying initial parameter values across a physically realistic range would potentially provide a conceptually simple and computationally efficient alternative for evaluating highly nonlinear integrated hydrologic models as a best practice approach when using such models to inform water management and policy decisions.

## 2. Study Area

The Scott Valley (Figure 1) in Northern California was chosen as a type case for an agricultural groundwater basin with Mediterranean climate and a groundwater-dependent ecosystem. Part of the larger Klamath Basin watershed that straddles the California-Oregon border, the Scott River watershed drains 2,100 km<sup>2</sup> and provides key spawning habitat for native anadromous fish species, including *Oncorhynchus tshawytscha* (Chinook salmon) and the threatened *Oncorhynchus kisutch* (coho salmon) (NCRWQCB, 2005). A large portion of the Scott River has been mapped as medium to high ranking on a groundwater-dependent ecosystem index by Howard and Merrifield (2010), indicating a high degree of connectivity between groundwater and surface water.

The montane valley, at over 800 m above mean sea level, is approximately 200 km<sup>2</sup> and formed during the Pleistocene by extension along a steep normal fault that dips to the east and strikes to the northwest (Mack, 1958). The surrounding uplands are part of the Klamath Mountains province, a sequence of accreted terranes and granitic intrusions associated with subduction of the Farallon plate beneath the North American plate (Irwin, 1990). Valley floor aquifer sediments are highly heterogeneous fluvial and alluvial deposits composed of gravels, sands, silts, and clays. Thickness reaches a maximum of more than 120 m (390 ft) in the central portion of the valley between Etna and Greenview and thins toward the valley margins. However, below 76 m (250 ft) the aquifer is not productive (Foglia, McNally, Hall, et al., 2013). Groundwater levels show interannual variation depending on the water year type but do not indicate long-term overdraft in the basin (University of California at Davis, 2016).

Climate in the area is Mediterranean with cold, wet winters and warm, dry summers. Mean annual precipitation on the valley floor is about 500 mm and accumulates predominantly during the winter and early

spring months (October–May). The mountains surrounding the valley receive higher precipitation rates due to their elevation. Mean temperature for January and July are 0 °C (32 °F) and 21 °C (70 °F), respectively. Mean annual runoff from Scott Valley, measured at the U.S. Geological Survey (USGS) stream gage (11519500) located in the Scott River Canyon just below the valley, is  $543 \times 10^6 \text{ m}^3$  (U.S. Geological Survey, 2015). Winter and spring flows (December–May) average about  $28.3 \text{ m}^3/\text{s}$  ( $1,000 \text{ ft}^3/\text{s}$ ) but have peaked at  $1,120 \text{ m}^3/\text{s}$  ( $39,500 \text{ ft}^3/\text{s}$ ). Mean summer streamflow is about  $0.85 \text{ m}^3/\text{s}$  ( $30 \text{ ft}^3/\text{s}$ ) but commonly drops below  $0.57 \text{ m}^3/\text{s}$  ( $20 \text{ ft}^3/\text{s}$ ) in the late summer and early fall. Maintaining sufficient instream flow during this critical low flow period is a key policy driver for water management in Scott Valley.

Land use in the valley is predominantly agricultural. There is a nearly even split between alfalfa hay/grain and pasture, which together account for  $136 \text{ km}^2$  (68%) of land use (California Department of Water Resources, 2000). Both surface water and groundwater are used as sources of irrigation water. Reliance on groundwater increased following the 1976–1977 drought with the widespread introduction of pressurized wheel line sprinkler systems and eventually center pivots to replace flood irrigation, although all three irrigation methods are still used in the valley (Van Kirk & Naman, 2008). The growing season typically lasts from mid-April to mid-September but varies depending on the year. The southern, narrow upstream portion of the valley has been heavily modified by dredging operations that left behind tailings 6 m (20 ft) high, forming a zone of highly permeable open-framework gravels. The main stream channel was straightened in many parts of the valley by the Army Corps of Engineers in the early 1930s for flood control purposes (U.S. Department of War, 1938), which has resulted in channel incision at some locations.

The Scott River flows from south to north through the valley, fed by 10 major tributaries (Figure 1). Two large irrigation diversions, Farmer's Ditch and the Scott Valley Irrigation District (SVID) Ditch, have their point of diversion located near river km 87 and 74, respectively. These two ditches are primarily active from April to July and run along the east side of the valley. In the north central area of the valley is a drainage slough ("eastside slough") that collects agricultural tailwater for discharge back into the Scott River. Several other minor irrigation ditches exist on most tributaries in the valley.

### 3. Methods

#### 3.1. SVIHM Overview

The SVIHM simulates hydrologic conditions in the Scott Valley from 1 October 1990 to 30 September 2011 by integrating three different models representing four subsystems: the upper watershed, and the alluvial basin landscape water, groundwater, and surface water. A statistical model is used to estimate tributary inflows at the valley margins when upper watershed flow data are unavailable ("streamflow regression model"; Foglia, McNally, Hall, et al., 2013). A land use/crop-soil water budget model ("soil water budget model") simulates agricultural practices in the valley to estimate hydrologic fluxes at the individual field scale using a tipping bucket approach (Foglia, McNally, Hall, et al., 2013), including determination of recharge and agricultural pumping rates. A finite difference groundwater-surface-water model simulates spatial and temporal groundwater and surface water conditions in the valley overlying the alluvial basin ("MODFLOW model"). The SVIHM is weakly coupled in that fluxes are passed from the SWBM to the MODFLOW model, but there are no direct feedbacks from the MODFLOW model to the SWBM.

Workflow for SVIHM involves running the streamflow regression model to generate average monthly streamflows at the valley margins. This is a preprocessing step as we assume no feedback from the valley to the upper watershed, and therefore, estimated values do not change unless the regression is modified. In a typical SVIHM run the SWBM is called which writes the necessary MODFLOW input files for streamflow, recharge, pumping, and ET. The MODFLOW model is then called, which calculates groundwater heads and streamflow in the valley. Flow into tile drains (see section 3.4) is extracted from the MODFLOW output, and the SWBM and MODFLOW model are run once more to route flows from tile drains to the surface water system. A summary of input data, key parameters, and model outputs is available in the supporting information. The SWBM takes about 10 min to run, and the MODFLOW model takes less than 2 hr using a desktop computer with Intel® Core™ i7-4770 @3.4-GHz processors with 16 GB of RAM for a total run time of typically less than 4 hr. Tighter integration of the landscape, groundwater, and stream subsystems is available in some software (e.g., IWFM or OWHM) but was not considered necessary for this study and avoided due to run time concerns.

### 3.2. Streamflow Regression Model

Upper watershed stream inflow data are available for limited and differing time periods for most tributaries as either daily mean flow values or monthly volumes (see the supporting information for date ranges and data available for each tributary). The streamflow regression model is used to fill in data gaps for tributary inflows to the valley, which are used as model boundary conditions. Data from all tributary gages are regressed against the USGS gaging station, where continuous daily streamflow data are available. When tributary flow data are available, measured values are used as model inputs. When tributary inflow data are unavailable, monthly streamflow is estimated using the regression model. Two of the tributaries, Johnson Creek and Crystal Creek, do not have any streamflow observations and therefore could not be included in the regression. Inflows for these two tributaries are calculated by scaling the estimated values for nearby Patterson Creek using the ratio of the subwatershed areas (Foglia, McNally, Hall, et al., 2013).

### 3.3. SWBM

The purpose of the SWBM is to estimate the unknown rates of groundwater pumping and recharge using a mass balance approach that incorporates local agricultural management practices (Foglia, McNally, Hall, et al., 2013). Fluxes of water in the shallow vadose zone are simulated on a daily basis using a tipping bucket style approach for 2041 fields identified from the California Department of Water Resources land use survey (California Department of Water Resources, 2000). Field areas vary from 5.2E3 to 6.6E6 m<sup>2</sup> (1.3 to 1,600 acres), with a median area of 6.9E4 m<sup>2</sup> (17 acres). The daily water budget for each field is calculated according to

$$\theta_k = \max(0, \theta_{k-1} + P_k + AW_k - ET_k - R_k) \quad (1)$$

$$AW_k = \begin{cases} \max\left(0, \frac{ET_k - P_k}{AE}\right) & \text{when } AE < 100\% \\ \max\left(0, \frac{ET_k - P_k}{1 + \text{SMDF}}\right) & \text{when } AE = 100\% \end{cases} \quad (2)$$

$$\text{SMDF} = \frac{\sum \text{soil moisture depletion}}{\sum AW} \quad (\text{during irrigation season}) \quad (3)$$

$$R_k = \max(0, \theta_{k-1} + P_k + AW_k - ET_k - \theta_{\max}) \quad (4)$$

$$ET_k = ET_{0,k} * K_c \quad (5)$$

$$P_k = \begin{cases} 0 & \text{when } P_k \leq 0.2 * ET_0 \\ P_k & \text{when } P_k > 0.2 * ET_0 \end{cases} \quad (6)$$

where  $\theta$  is available soil moisture,  $\theta_{\max}$  is the soil field capacity,  $P$  is effective precipitation,  $AW$  is applied water,  $ET$  is actual evapotranspiration,  $ET_0$  is the reference evapotranspiration,  $K_c$  is the crop coefficient,  $R$  is recharge,  $AE$  is application (irrigation) efficiency,  $\text{SMDF}$  is the soil moisture depletion factor, and the subscript  $k$  denotes the day. The SWBM takes into account land use, irrigation method, water source, and soil storage properties for each field. Irrigation demand (equation (2)) for each field is driven by daily reference  $ET$  ( $ET_0$ ) in excess of daily effective precipitation ( $P_k$ ), land use crop coefficient ( $K_c$ ), and soil moisture depletion factor ( $\text{SMDF}$ ) specific to the land use and irrigation type. The  $\text{SMDF}$  allows for actual  $ET$  to exceed applied water to account for deficit irrigation, for contributions from deep soil moisture below the depth of the simulated soil zone (Foglia, McNally, Hall, et al., 2013), or for other generic sources of water (Dogrul et al., 2018). For fields with access to groundwater, pumping is assigned to the nearest well. Groundwater is assumed to be available at all times as there have been no reports of wells going dry in the valley, even during the 2012–2015 drought. Runoff from fields is considered to be negligible. Soil moisture in excess of field capacity at the end of each day is assumed to recharge groundwater. Daily values of pumping and recharge are converted to monthly average rates to match the stress period of the MODFLOW model.

### 3.4. Groundwater-Surface-Water Model

Groundwater fluxes, heads, groundwater-surface-water exchange, and streamflow are simulated with MODFLOW and the SFR package using monthly stress periods and daily time steps. The domain consists of 440

rows, 210 columns, and two layers with 100-m (328 ft) lateral resolution ranging from 0 to 61 m (0–200 ft) thick. A combination of remotely sensed elevation data using a digital elevation model and light detection and ranging (LiDAR) with horizontal resolutions of 10 and 1 m, respectively, is averaged within each model cell to determine ground surface elevation. LiDAR data from the valley has a  $2\sigma$  relative accuracy of 4 cm (Watershed Sciences, 2010) and covers more than 90% of the model domain. Bedrock surrounding and outcropping within the valley is assumed to be impermeable relative to the valley sediments. Hydraulic conductivity and storage properties for the valley aquifer are spatially distributed between nine hydrogeologic zones (Figure 1), similar to those proposed by Mack (1958). The model simulates unconfined flow with variable storage and transmissivity using the Newton formulation of MODFLOW (MODFLOW-NWT; Niswonger et al., 2011) to allow for rewetting of cells that go dry during the simulation, especially along the valley margin. Newton solver variables are set to default values corresponding to the “COMPLEX” option defined in the user manual, as “SIMPLE” and “MODERATE” result in shorter run times but unsatisfactory numerical errors.

Groundwater pumping is simulated at 164 agricultural wells (Figure 1) located in the second layer using the well (WEL) package (Harbaugh, 2005). The wells and their locations were identified through well logs, stakeholder feedback, aerial photography, and field surveys when possible. Due to the low population density of the valley, domestic pumping is not included in the model as it is a small portion of total groundwater extractions (Mack, 1958). The default value of 0.05 for PHIRAMP in the WEL package is used for reducing pumping rates in cells when there was not enough water to satisfy the applied pumping rate. The weak coupling between the SWBM and MODFLOW can result in a mass balance error between the two submodels, as the SWBM is not aware of pumping reductions that happen within MODFLOW. Pumping reductions range from 0% to 7.6% on a monthly basis, with a mean of 3.7%, so it is not considered to be a significant limitation of SVIHM at this time.

Seepage from the SVID and Farmers Company ditches (Figure 1) have been determined using field seepage experiments. They are represented using injection wells with rates of  $1.8 \times 10^{-2} \text{ m}^3/\text{s}$  per km ( $1.0 \text{ ft}^3/\text{s}$  per mile) and  $8.8 \times 10^{-3} \text{ m}^3/\text{s}$  per km ( $0.5 \text{ ft}^3/\text{s}$  per mile), respectively (Echols, 1991; S. S. Papadopoulos and Associates, Inc, 2012). Water is diverted from the stream using “ghost” SFR segments at the respective points of diversion at the same rate it is injected via the WEL package in order to conserve mass. Ditches in the valley are generally active from April to July and primarily used for stock watering and limited irrigation (S. Sommarstrom, personal communication, June 21st, 2017). As most of the flow is needed to generate sufficient head for water to reach the end of the ditches (P. Harris, personal communication, June 21st, 2017), diversions from the Scott River into ditches are greater than the ditch seepage rate supplied to the WEL package to account for evaporation losses and stockwater use. These consumptive losses are assumed to be  $5.7\text{E}-2 \text{ m}^3/\text{s}$  for each ditch, or 12.5% and 25% of the Farmers Company ditch and SVID ditch diversion rates, respectively. Mountain front recharge (MFR), used here to describe the diffuse portion of recharge to adjacent basins from surrounding mountains (Wilson & Guan, 2013), was simulated only along the western model boundary using injection wells placed in the first layer. Rates and spatial distribution of MFR segments were estimated by S. S. Papadopoulos and Associates, Inc. (2012) using a water balance approach. No MFR occurs along the eastern valley boundary as these mountain ranges are lower and in the rain shadow of the western ranges of the watershed.

The SFR package was used to simulate the surface water system (Figure 1) and its interactions with groundwater. Bed elevations for each SFR node were extracted from the elevation data sets using the nearest thalweg data point to the stream node. Inflows for each tributary and the main stem of the Scott River are specified at the model boundaries for each stress period using the streamflow regression model described above. Surface water used by SWBM is subtracted from the inflow estimated by the streamflow regression model at the model boundary, as most diversions from tributaries occur near or upstream of the model boundary. The stream channel is assumed to be rectangular with flow calculated using Mannings equation. Stream properties (e.g., bed conductivity and roughness) were assigned using one of three groups: (1) tailings reaches, (2) nontailings Scott River, and (3) tributaries (Figure 1).

Drains were placed at the land surface within the Discharge Zone shown in Figure 1. This area is known to have a water table very near or at the land surface (Mack, 1958), resulting in lateral fluxes of water between fields that SWBM does not account for. Water intercepted by these drains is routed into a nearby stream segment, approximating overland flow. In the Discharge Zone, crop water demands are met primarily through

subsurface irrigation (direct uptake from groundwater). A tight integration of landscape/crop-soil water and groundwater subsystems would be most suitable in this zone. Here, a work-around was applied: Fields with subirrigation were not irrigated in SWBM, resulting in “dry” soils and underestimation of crop ET. To compensate, the MODFLOW ET package was employed instead to simulate ET from these fields, using an extinction depth of 0.5 m. The applied ET rate for these fields is equal to the average potential ET rate for the specific stress period and crop, scaled by the fraction of time per month when the field was “dry” (and therefore ET was not occurring) in SWBM. For example, a field within the Discharge Zone that had 0 m/day of ET in SWBM simulation for 5 days in July because available soil moisture dropped to 0 m<sup>3</sup> would be simulated in MODFLOW with an applied ET rate equal to the average potential ET rate for that month multiplied by 0.16 (5 of 31 days). This ensures that ET is neither double counted between SWBM and MODFLOW, nor significantly underestimated, given the water table depth.

### 3.5. Sensitivity Analysis and Calibration

Sensitivity analysis and calibration of SVIHM was performed using the universal inverse modeling software suite UCODE\_2014 (Poeter & Hill, 1998; Poeter et al., 2014), which compares observed and simulated values to create an objective function given by

$$\phi = \sum_{i=1}^{ND} (y_i - y'_i) w_i^{1/2} \quad (7)$$

where  $\phi$  is the objective function value,  $y_i$  is the observed value,  $y'_i$  is the simulated value,  $w_i$  is the observation weight, and ND is the number of observations. Given the nonlinear structure of SVIHM, global methods for sensitivity analysis and calibration would provide the most rigorous approach. However, this is impractical due to the large number of model runs required and the long ( $\gg 1$  hr) single-run CPU time for SVIHM. Gradient-based perturbation methods are more efficient, but computed parameter sensitivities can vary as a function of the starting values (Hill & Tiedeman, 2007). As an alternative design for the sensitivity analysis and to account for model nonlinearity, multiple forward difference sensitivity analyses using different starting parameter combinations (referred to as parameter sets 1–5) were performed. This is similar to the distributed evaluation of local sensitivity analysis methodology proposed by Rakovec et al. (2014) and the use of multiple sensitivity indices (e.g., variance based and local sensitivity) by Borgonovo et al. (2017), although less rigorous to preserve computational efficiency. Starting values for the different sensitivity runs (Table 1) were selected either from a previous model calibration using the river package (RIV; Harbaugh, 2005; Foglia et al., 2018) or from within an expected range based on professional judgment. Parameters were log-transformed and increased by 1% from their starting values for all sensitivity analyses.

Measured groundwater elevations at 55 wells (Figure 1) accounted for 2,197 observations (47% of total). The majority of these were collected monthly beginning in 2006 as part of the voluntary Scott Valley Community Groundwater Monitoring Program. Weights for groundwater head observations were set to the inverse of the measurement error variance (Hill & Tiedeman, 2007) given by

$$w_i = \frac{1}{\sigma_i^2} \quad (8)$$

where  $\sigma_i^2$  is the measurement error variance and assumed to be 1.0 m<sup>2</sup> for all head observations. This accounts for measurement errors in both well reference point elevation and depth to water.

Streamflow data at four gage locations (Figure 1) were also included as calibration targets in the objective function and separated into three categories: (1) below  $2.44 \times 10^5$  m<sup>3</sup>/s (100 ft<sup>3</sup>/s; 68% flow exceedance probability), (2) between  $2.44 \times 10^5$  and  $2.44 \times 10^6$  m<sup>3</sup>/s (100–1,000 cfs), and (3) greater than  $2.44 \times 10^6$  m<sup>3</sup>/s (1,000 cfs; 22% flow exceedance probability). These reflect low, medium, and high streamflow rates for the Scott River, respectively. A total of 2485 streamflow observations, consisting of 1,385 at the USGS Fort Jones gage, 500 at the Lower Shackelford Creek gage, 300 at the Above Serpa Lane (AS) gage, and 300 at the Below Young's Dam (BY) gage, was randomly selected from data available during the model simulation period so the total number of streamflow observations was similar to the number of groundwater head observations. Streamflow observation weights were determined using the equation

$$w_i = \frac{1}{(y_i * CV_i)^2}, \quad (9)$$



**Table 1**  
*Parameters Adjusted During Sensitivity Analysis With Initial Starting Values*

Parameter	Description	Initial value				
		Parameter Set 1	Parameter Set 2	Parameter Set 3	Parameter Set 4	Parameter Set 5
Kx1		100	20	60	200	250
Kx2		11	60	2	100	100
Kx3		100	80	100	50	100
Kx4	Hydraulic	20	2	50	10	30
Kx5	conductivity	10	1	80	10	25
Kx6	(m/day)	30	200	70	10	50
Kx7		1,000	500	50	1,000	500
Kx8		30	90	5	10	10
Kx9		60	2	100	10	20
Kvar1		100	10	50	20	71
Kvar2		100	10	50	50	73
Kvar3		100	10	50	100	92
Kvar4	Vertical	100	10	50	80	95
Kvar5	anisotropy	100	10	50	40	10
Kvar6	(—)	100	10	50	30	77
Kvar7		1	1	1	10	55
Kvar8		100	10	50	60	94
Kvar9		100	10	50	50	46
Sy1		0.1	0.15	0.05	0.1	0.12
Sy2		0.15	0.05	0.12	0.07	0.08
Sy3		0.15	0.2	0.1	0.1	0.11
Sy4	Specific	0.1	0.12	0.11	0.1	0.1
Sy5	yield	0.15	0.15	0.1	0.1	0.13
Sy6	(—)	0.15	0.08	0.1	0.1	0.09
Sy7		0.3	0.25	0.15	0.3	0.25
Sy8		0.15	0.11	0.06	0.1	0.05
Sy9		0.15	0.12	0.15	0.1	0.07
Ss1		1.00E−05	5.00E−04	2.00E−04	6.40E−05	1.50E−05
Ss2		1.00E−05	5.00E−04	2.00E−04	5.90E−05	9.00E−05
Ss3		1.00E−05	5.00E−04	2.00E−04	2.70E−05	2.60E−04
Ss4	Specific	1.00E−05	5.00E−04	2.00E−04	8.30E−05	3.10E−04
Ss5	Storage	1.00E−05	5.00E−04	2.00E−04	1.30E−05	8.00E−05
Ss6	(1/m)	1.00E−05	5.00E−04	2.00E−04	2.00E−05	2.00E−05
Ss7		1.00E−05	5.00E−04	2.00E−04	6.50E−05	7.00E−05
Ss8		1.00E−05	5.00E−04	2.00E−04	2.90E−04	3.00E−04
Ss9		1.00E−05	5.00E−04	2.00E−04	8.00E−04	1.00E−05
Wel5		2.34E+02	2.34E+03	1.17E+04	2.34E+03	8.19E+02
Wel6		6.30E+03	1.05E+04	5.25E+02	1.05E+04	2.10E+03
Wel7	Mountain front	1.20E+04	8.60E+02	2.58E+04	1.72E+03	5.16E+03
Wel8	recharge	2.24E+04	8.94E+03	2.98E+03	1.12E+04	2.98E+04
Wel9	(m <sup>3</sup> /day)	2.00E+03	2.00E+04	4.00E+04	6.00E+03	4.00E+03
Wel10		3.86E+03	1.93E+04	3.86E+02	9.65E+02	2.90E+03
Wel11		2.43E+03	2.70E+02	2.70E+01	1.89E+03	3.24E+03

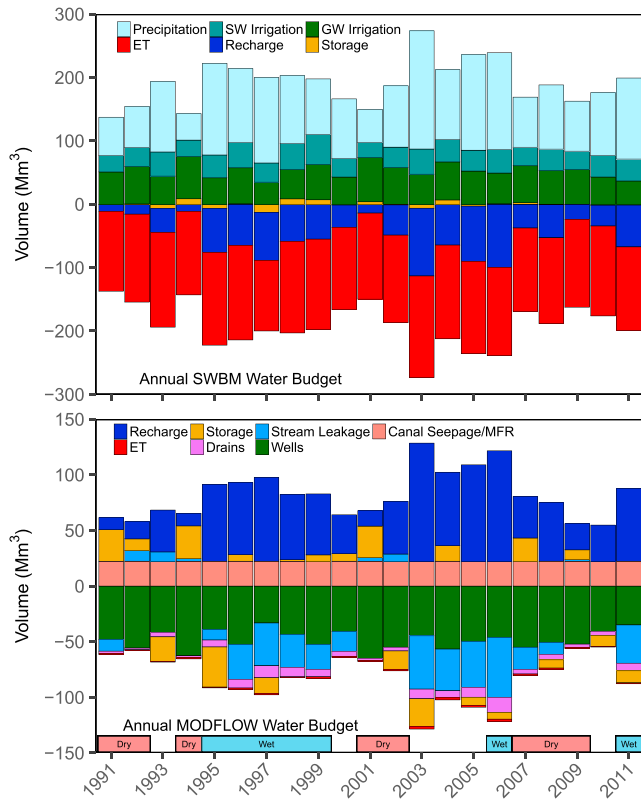
Table 1 (continued)

Parameter	Description	Initial value				
		Parameter Set 1	Parameter Set 2	Parameter Set 3	Parameter Set 4	Parameter Set 5
Wel20	Ditch seepage	1.45E+04	2.42E+04	6.44E+03	1.61E+04	2.42E+04
Wel21	(m <sup>3</sup> /day)	7.02E+04	4.68E+04	1.17E+04	2.34E+04	7.02E+04
BedK1	Streambed	10	0.1	10	10	10
BedK2	conductivity	10	0.1	10	5	10
BedK3	(m/day)	10	0.1	10	15	10
Rough1	Streambed	0.035	0.03	0.04	0.035	0.035
Rough2	Roughness	0.035	0.03	0.04	0.035	0.035
Rough3	(d/m <sup>1/3</sup> )	0.035	0.03	0.04	0.035	0.035
RD_Mult		1	0.9	1.1	1	1.2
S MDF_Flood	SWBM	0.7	0.75	0.55	0.7	0.55
S MDF_WL_LU25	parameters	1.05	1.09	1.01	1.09	1.1
S MDF_CP_LU25	(—)	1.1	1.15	1.01	1.15	1.1
S MDF_WL_LU2		0.85	0.81	0.99	0.81	0.95
S MDF_CP_LU2		0.95	0.91	0.91	0.95	0.99
Kc_alfalfa_mult		1	1.04	0.96	1	1.05
Kc_grain_mult		1	0.96	1	1.05	0.98
Kc_pasture_mult		1	1	0.96	1.04	0.95
Kc_noirr		0.6	0.7	0.79	0.65	0.65

Note. Sensitivity analysis was performed using a 1% forward difference perturbation. Note that the reported soil moisture depletion factor (S MDF) is the sum of the application efficiency (AE) and S MDF but is reported as a single value for convenience.

where  $CV_i$  is the coefficient of variation. The low, medium, and high streamflow categories were assigned coefficients of variation equal to 10%, 20%, and 40%, respectively. Low flows at the non-USGS gages were the only exception as they included observations at or very near 0, and weights approach infinity when observation values approach 0 using the coefficient of variation weighting method. The median weight of the USGS low flow observations, equal to  $1 \times 10^{-8}$  d<sup>2</sup>/m<sup>2</sup>, was assigned to the low flow category for the non-USGS gages to prevent weights from becoming too large and dominating the objective function.

Physical hydrologic properties or fluxes were represented by 61 parameters (Table 1) contained within SWBM and MODFLOW portions of SVIHM. These include seasonal crop coefficient multipliers (e.g., Kc\_ Alfalfa\_Mult) and aquifer parameters for each zone such as hydraulic conductivity (e.g., Kx1) and storage (e.g., Sy1 and Ss1). Specific cuttings of alfalfa were not represented using a variable crop coefficient as cutting times vary across the valley due to distributed ownership, management practices, and climate conditions. Instead, crop coefficients for alfalfa and pasture were set to seasonal averages of 0.9 during the growing season and zero otherwise. A variable  $K_c$  was used for grain due to growers having similar management practices and ranged from 0 to 1.15 with an average of 0.62 over the 4-month growing period. Effective root zone depth was assumed to be 2.44 m (8 ft) for alfalfa and 1.22 m (4 ft) for grain and pasture (Weaver, 1926), adjusted by the use of a single root zone depth multiplier (RD\_Mult). Each of the five simulated combinations of land use (alfalfa/grain, A/G; pasture, P) and irrigation (flood; wheel line sprinkler, WL; center pivot sprinkler, CP) was assigned specific AE and S MDF values (Table 1). Channel roughness and bed conductivity (Figure 1) parameters were distributed among the three stream segment classifications (see section 3.4). Seepage fluxes from MFR and ditches were also included as parameters. Recharge and groundwater pumping were not explicitly included as calibration parameters because they are accounted for with SWBM parameters. SWBM parameterization also affects surface water inflows to the simulated stream network via simulated surface water diversions from streams at the model boundary. In contrast, daily total precipitation across the valley was considered to have relatively small measurement error and was not considered in the sensitivity analysis.



**Figure 2.** Annual water budgets for SWBM (top) and MODFLOW (bottom) portions of Scott Valley Integrated Hydrologic Model. Values are water year totals, with colors along the bottom of the lower plot indicating dry (red) and wet (blue) years. Positive and negative storage values correspond with decreases and increases in storage, respectively. ET = evapotranspiration; SW = surface water; GW = groundwater; SWBM = soil water budget model; MFR = mountain front recharge.

Composite scaled sensitivity (CSS), a measure of the importance of observations as a whole to a single parameter, is calculated for each parameter from the sensitivity analyses according to

$$CSS_j = \left[ \frac{\sum_{i=1}^{ND} \left[ \left( \frac{\partial y'_i}{\partial b_j} \right) b_j w_i^{1/2} \right]^2}{ND} \right]^{1/2} \quad (10)$$

where  $CSS_j$  is the composite scaled sensitivity of the  $j$ th parameter and  $b$  is the vector of parameters. These values were normalized in order to compare between sensitivity analyses using

$$CSS_{j,s, norm} = \frac{CSS_{j,s}}{\max [CSS_s]} \quad (11)$$

where  $CSS_{j,s}$  is the composite scaled sensitivity of the  $j$ th parameter for set  $s$  (1–5 in the study) and  $CSS_s$  is the array of CSS values for parameter  $j$ . Calibration parameters were selected by ranking the CSS values for each parameter set and identifying parameters that consistently showed the greatest sensitivity across all five sensitivity analyses. Multiple calibrations (Runs 1–5) were then performed, again with different starting values for adjustable parameters to test for model uniqueness. Values of fixed parameters (i.e., those excluded from the calibration process) were selected from the first sensitivity analysis parameter set (Table 1). Adjustable parameters were modified by UCODE\_2014 in an attempt to minimize the objective function and therefore provide the best match between observed and simulated values. Convergence was met when either parameter values did not vary by more than 1% (TolPar = 0.01), or the objective function did not change by more than 1% for three consecutive iterations (TolSOSC = 0.01). Nonlinearity of SVIHM was evaluated using the modified Beale's measure (Cooley & Naff, 1990) calculated by the program MODEL\_LINEARITY, available in the UCODE\_2014 distribution (<https://igwmc.mines.edu/ucode/>).

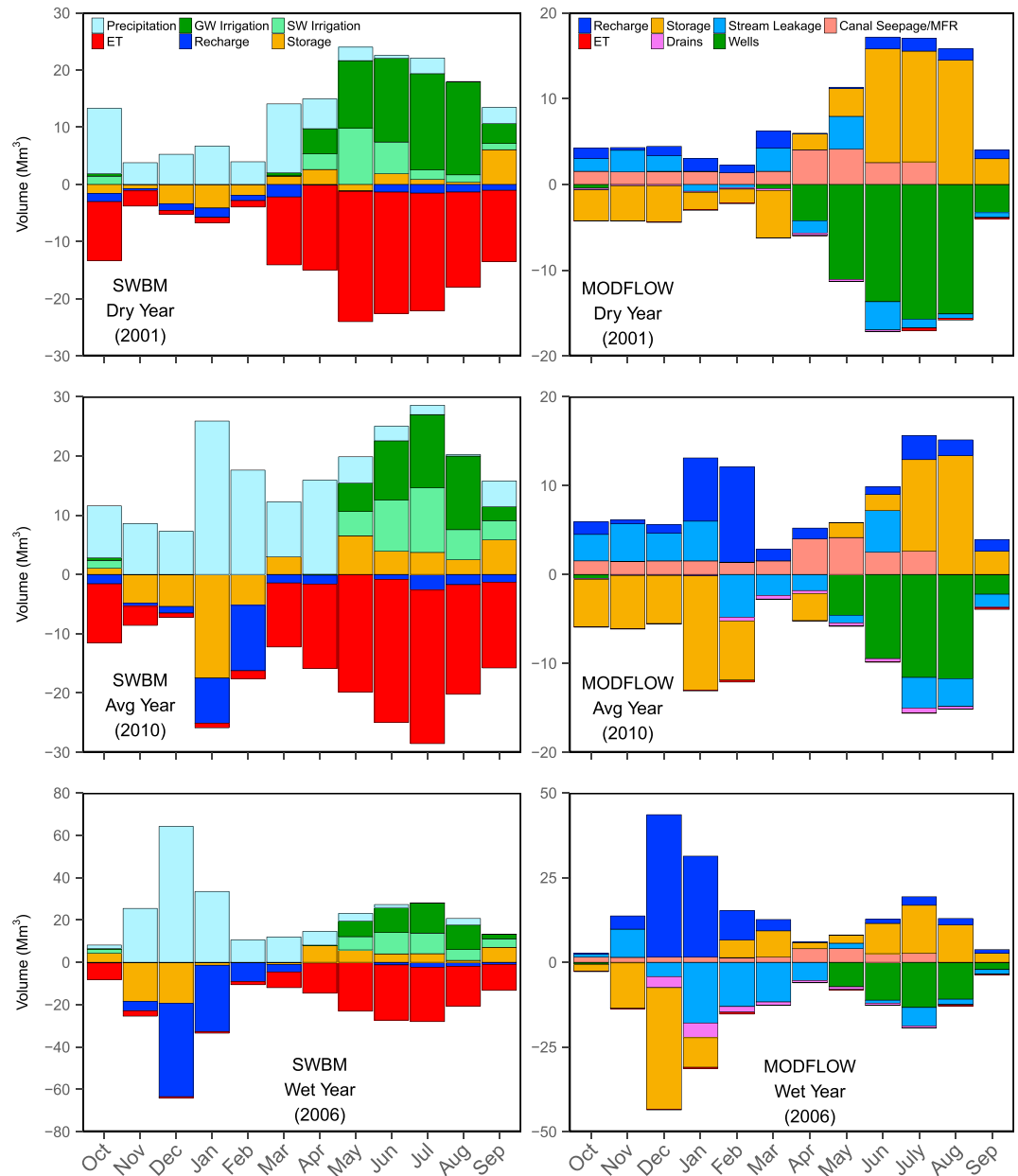
The influence that observations exerted during the calibration process were evaluated using the DFBETAS and Cook's D statistics. The DFBETAS statistic provides information about the influence of an observation on each calibration parameter, with more influential observations having greater absolute values. Cook's D measures how influential an observation is on the entire parameter set by calculating how much regression estimates would change if the observation was omitted. Like DFBETAS, greater values indicate greater influence. Both statistics have critical values for defining when observations are considered influential. For DFBETAS and Cook's D this is equal to  $4/(ND + NPR)$  and  $2/(ND + NPR)^{1/2}$ , respectively, where ND is the number of observations and NPR is the number of prior information equations (Yager, 1998; Hill & Tiedeman, 2007). Prior information equations were not used in this version of SVIHM.

### 3.6. Streamflow Matching

Comparison of simulated streamflow to observed values at gages was done both graphically and using a modified version of the Nash-Sutcliffe model efficiency coefficient (NSE; Nash & Sutcliffe, 1970)

$$NSE = 1 - \frac{\sum_{i=1}^n (\log [y_i - \log y'_i])^2}{\sum_{i=1}^n (\log [y_i - \log \bar{y}_i])^2}, \quad (12)$$

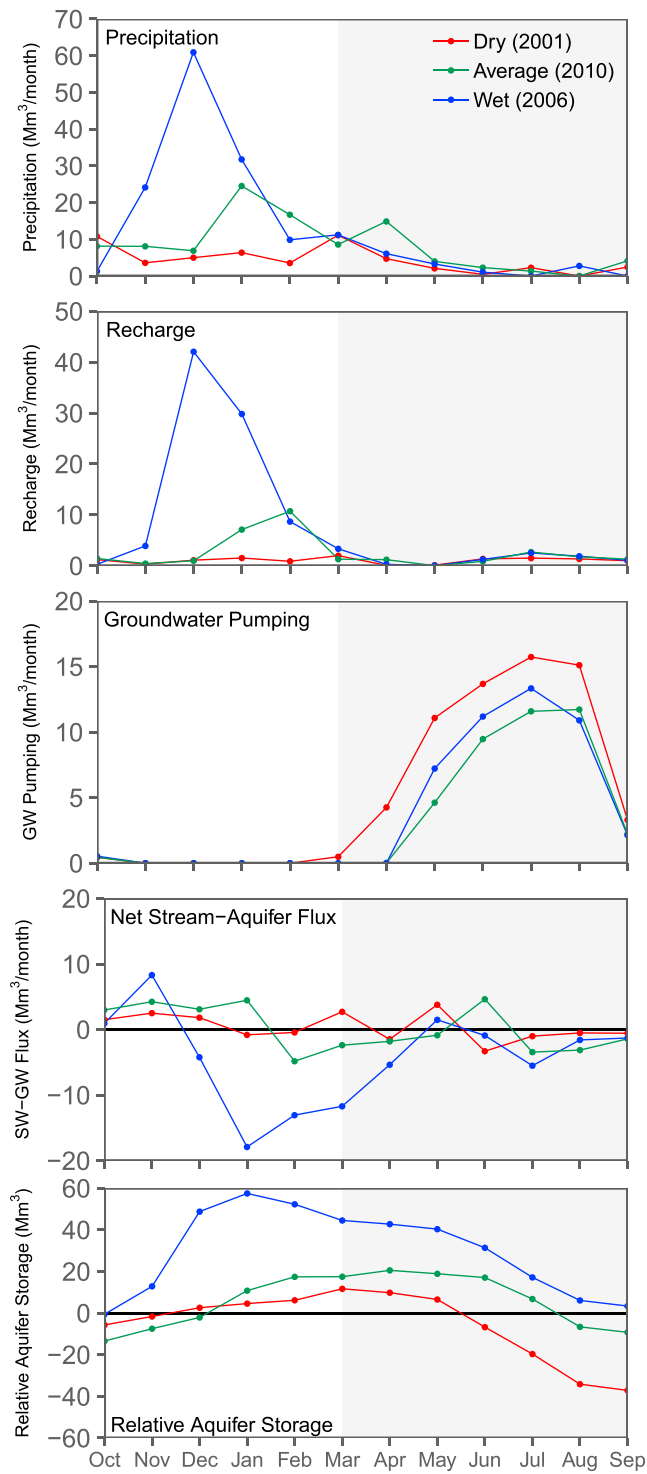
where  $n$  is the number of streamflow observations for the gage. A NSE of 1.0 indicates the model perfectly matches observations, while a value of 0.0 means the model is no more accurate than predicting the mean value. Streamflow data were log-transformed because they span nearly 4 orders of magnitude in the Scott Valley and large variance can produce high NSE values even if model fit is relatively poor (Jain & Sudheer, 2008). Therefore, NSE values presented in this paper are conservative.



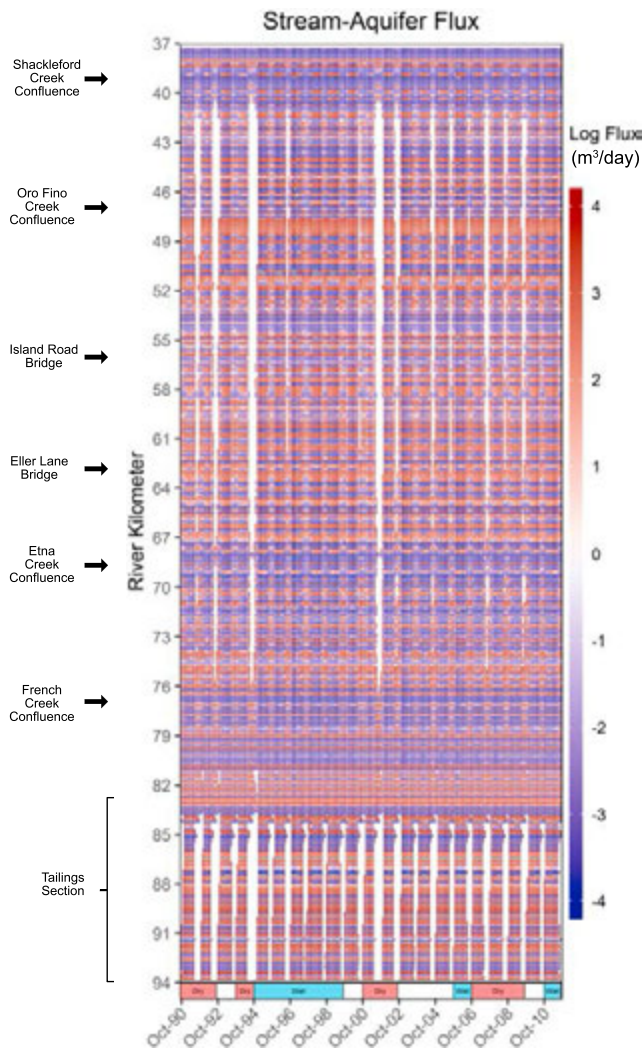
**Figure 3.** Monthly water budgets for SWBM (left column) and MODFLOW (right column) portions of SVIHM for dry (top row), average (middle row), and wet (bottom row) years. Positive and negative storage values correspond with decreases and increases in storage, respectively. Colors along the bottom of the plot indicate dry/critical (red) and wet (blue) water year types according to the Sacramento Valley water year hydrologic classification. ET = evapotranspiration; GW = groundwater; SW = surface water; SWBM = soil water budget model; MFR = mountain front recharge.

### 3.7. Qualitative Model Validation

To validate the usefulness of SVIHM for informing water management decisions related to the critical late summer period when stream connectivity is most compromised, a model validation was performed to test SVIHM's capability to predict late summer dry reaches. Data obtained from the U.S. Department of Agriculture's National Agriculture Imagery Program (NAIP) in combination with direct observations by landowners and resource professionals were used to map dry/disconnected stream reaches in the Scott Valley. Imagery and direct observations from August 2005 and 2014, representing average and dry water year conditions in the Scott Valley, respectively, were digitized using ArcGIS. Water year 2014 was used as a proxy for water year 2001 in the model since NAIP data only extend back to 2003 and SVIHM terminates in 2011. Both



**Figure 4.** Monthly fluxes of selected basin water budget components for dry (2001), average (2010), and wet (2006) years. Negative values for net stream-aquifer flux correspond with groundwater discharge to surface water. Relative aquifer storage is the cumulative change in groundwater storage from initial conditions. Gray shaded area indicates growing season. GW = groundwater; SW = surface water.



**Figure 5.** Spatiotemporal heat map of fluxes between groundwater and surface water for the Scott River with geographic locations noted. Fluxes are highly spatially variable, despite relatively homogeneous parameterization of the stream. White areas indicate dry reaches. Colors along the bottom of the plot indicate dry/critical (red) and wet (blue) water year types according to the Sacramento Valley water year hydrologic classification. The absolute value is the magnitude of the flux, while the sign indicates flux direction: Red and blue indicate losing and gaining reaches, respectively.

years show similarities in timing and magnitudes of streamflow. Reaches were assigned one of four categories: dry/disconnected, flowing, questionable, and no data. These mapped sections were compared with streamflow values produced by SVIHM at the end of the corresponding month. Modeled stream reaches with flows less than  $2.8 \times 10^{-2} \text{ m}^3/\text{s}$  ( $1 \text{ ft}^3/\text{s}$ ) were considered dry/disconnected, whereas all other reaches were considered flowing.

NAIP imagery is taken during the growing season but generally before field observations were collected, typically at the seasonally lowest streamflow in late August or September. Where only NAIP imagery is available, the data may therefore be biased toward wetter conditions than would have been observed later in the year. However, it was considered to be a useful data set for qualitative model validation.

## 4. Results

### 4.1. Water Budget and Groundwater-Surface-Water Interactions

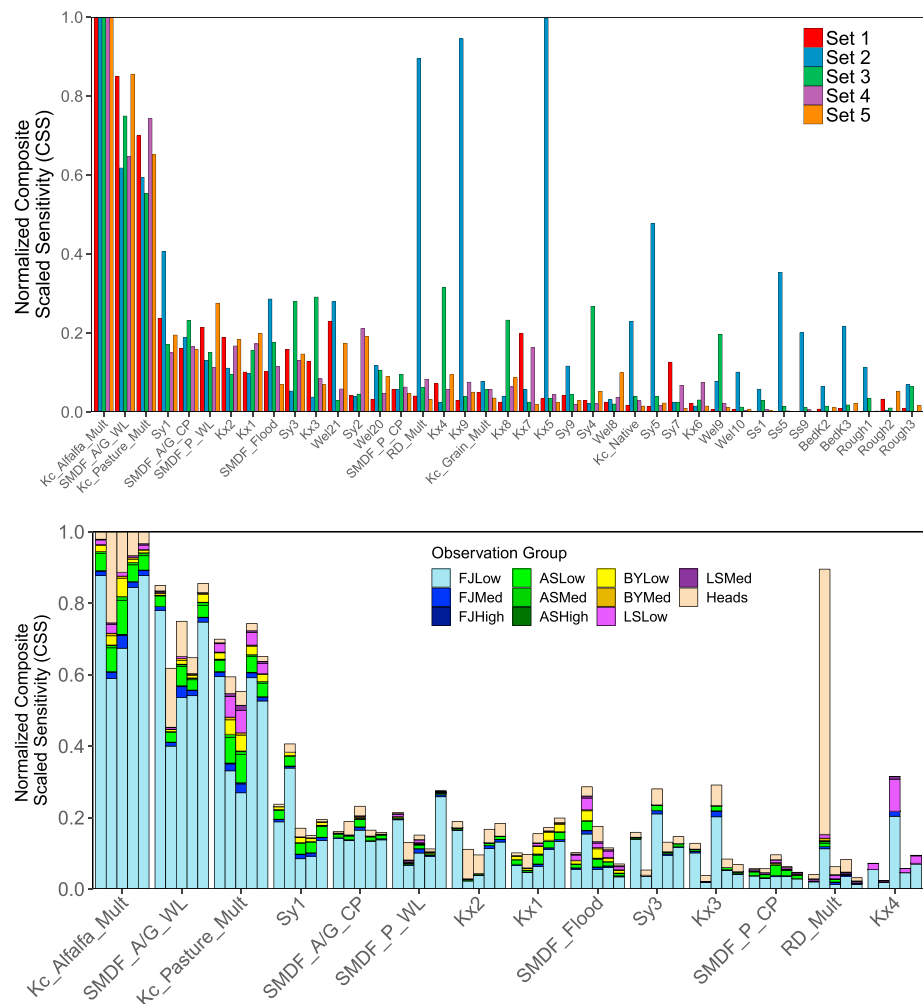
In order to provide context for the sensitivity analysis and calibration results, we first present model results from one of the calibrated parameter sets. Run 4 was chosen because it had the lowest objective function value. Annual water budgets were computed separately for the soil landscape (SWBM) and groundwater (MODFLOW) systems (Figure 2). Landscape inflows consist of precipitation and irrigation, while outflows are to ET and groundwater recharge. Groundwater inflows include landscape recharge and stream recharge. Groundwater outflows include pumping, drains, discharge to streams, and ET within the discharge zone and along the riparian corridor. Precipitation varies considerably from year to year, ranging from 22.5 cm (8.9 inches) to 98.1 cm (38.6 inches). Applied irrigation from groundwater and surface water in the valley also varies interannually from 31.7 cm (12.5 inches) to 50.4 cm (19.8 inches). Groundwater pumping accounts for about 50% of irrigation water in wet years but nearly 75% in dry years. Evapotranspiration has small interannual variability and is the largest flux out of the soil zone due to the dominant presence of agriculture in the valley. Given that interannual soil moisture storage changes are small (0.2% to 15% relative to field capacity), interannual fluctuations in recharge therefore follow those of precipitation, with which it is highly correlated ( $R^2 = 0.89$ ). Monthly and annual water budgets are available in the supporting information.

At the basin scale, landscape recharge is the largest inflow to the aquifer on an annual basis. In the current version of the model, canal seepage and MFR are constant and constitute the second largest inflow to the aquifer.

Dry years show reductions in groundwater storage. At the valley scale,

groundwater is a net contributor to streamflow in all but the driest years. In those years, streams are net contributors to the groundwater budget but only after significant depletion of groundwater storage in the preceding year (Figure 2). Groundwater pumping and discharge to the stream are the two largest annual outflows from the aquifer and show large variations, with 50% more groundwater pumping in dry years than wet years. Drains in the Discharge Zone account for approximately 4% of the annual aquifer outflow during dry years and nearly 15% during wet years. A small volume of water is removed from the aquifer directly via ET due to shallow water table conditions in the Discharge Zone (Figure 1, see section 3.4). Annual change in aquifer storage is highly variable and largely dependent on water year type.

Monthly water budgets for dry (2001), average (2010), and wet (2006) water type years (State Water Resources Control Board, 1999; Deas, 2006) demonstrate significant intra-annual, seasonal variations that drive the system (Figure 3). Generally, recharge is high and both, soil and groundwater storage increase during the



**Figure 6.** Normalized composite scaled sensitivities (CSS) for each parameter set (top) and the 14 most sensitive parameters in Scott Valley Integrated Hydrologic Model (bottom). Parameters with a value less than 5% for all sensitivity runs have been excluded for plotting purposes in the top graph. Colors represent either the parameter set (top) or the contribution of different observation groups to parameter sensitivity (bottom). Low streamflow observations, particularly those at the U.S. Geological Survey gage, are the most sensitive to changes in model parameter values.

winter months. Storage decrease, groundwater pumping, surface water irrigation, and high ET dominate from spring to early fall, leading to the lowest aquifer storage each year by late summer, although some recharge may again occur during the irrigation season (Figure 4).

While the general patterns persist, timing and magnitude of some fluxes contribute to significant differences in seasonal patterns between water year types. In a dry year, winter precipitation is lacking, soil moisture is not refilled, and pumping starts at the beginning of the growing season in March, whereas it is delayed until May during an average and wet year. Groundwater pumping is similar between most year types during the summer months, but dry years cause significant spring abstractions of groundwater, while summer recharge from irrigation return water is lower (Figures 3 and 4). In a wet year, cloud cover and precipitation days during the growing season can significantly reduce ET (Figure 3). Seasonal aquifer storage trends are very similar between a wet and a dry water year types, with similar reduction in storage between March and September. But with a dry winter, aquifer storage is significantly more depleted in March when compared to a wet year. In an average year, winter storage increases are not as large as in a wet year, while summer depletion is not as significant as in a dry year. Hence, an average year type tends to see the least groundwater storage change between March and October (Figure 4).

**Table 2**  
*Initial and Final Values for Five Calibration Runs*

Parameter	Adjustable parameter values										Calibration Bounds
	Calibration Run 1		Calibration Run 2		Calibration Run 3		Calibration Run 4		Calibration Run 5		
	Initial	Final	Initial	Final	Initial	Final	Initial	Final	Initial	Final	
Kx1	100	40	20	26	60	74	200	194	250	46	$10^{-7}$ – $10^3$
Kx2	11	4	60	4	2	6	100	7	100	9	$10^{-7}$ – $10^3$
Kx3	100	56	80	116	100	9	50	10	100	7	$10^{-7}$ – $10^3$
Kx4	20	24	2	25	50	24	10	16	30	23	$10^{-7}$ – $10^3$
Sy1	0.1	0.09	0.15	0.08	0.05	0.13	0.1	0.13	0.12	0.09	0.01–0.35
Sy3	0.15	0.03	0.2	0.06	0.1	0.04	0.1	0.05	0.11	0.04	0.01–0.35
RD_Mult	1	2	0.9	1.7	1.1	1.9	1	1.4	1.2	1.2	0.5–2
SMDF_Flood	0.7	0.67	0.75	0.74	0.55	0.8	0.7	0.72	0.55	0.68	0.5–0.8
SMDF_A/G_WL	1.05	1.02	1.09	1.04	1.01	1.05	1.14	1	1.03	1.17	1–1.2
SMDF_A/G_CP	1.1	1.11	1.15	1.01	1.01	1	1.15	1.16	1.1	1.13	1–1.2
SMDF_P_WL	0.85	1	0.81	0.8	0.99	0.99	0.9	0.85	0.95	1	0.8–1
SMDF_P_CP	0.95	0.9	0.91	0.85	0.91	0.91	0.95	0.87	0.99	1	0.8–1
Kc_Alalfa_Mult	1	0.97	1.04	1.05	0.96	1.05	1	1.05	1.05	1.01	0.95–1.05
Kc_Pasture_Mult	1	1.05	1	1.01	0.96	0.98	1.04	1.05	0.95	1.02	0.95–1.05
Objective function value	9.33E+04		9.57E+04		9.08E+04		8.90E+04		9.66E+04		

*Note.* The only differences between the five runs was the starting value of the adjustable parameters. Note that the reported soil moisture depletion factor (SMDF) is the sum of the application efficiency and SMDF but is reported as a single value for convenience.

Simulated stream-aquifer fluxes are highly variable along the stream profile (Figure 5). Exchange fluxes between groundwater and the stream vary from tens to thousands of cubic meters per day. Gaining stream reaches alter with losing stream reaches at a rate of typically 200 to about 1,000 m.

The lower 10 km of the Scott River (river km 37–47) is mostly gaining reaches interspersed with small segments of losing reaches. The same general pattern is observed for the 17 km (river km 67–84) of the Scott River below the Tailings section. In contrast, the 20-km-long midsection of the river, from just downstream of the confluence with Etna Creek to the confluence with Oro Fino Creek (river km 47–67), and much the Tailings section in the uppermost 10 km of the Scott River (river km 84–94) are dominated by losing reaches.

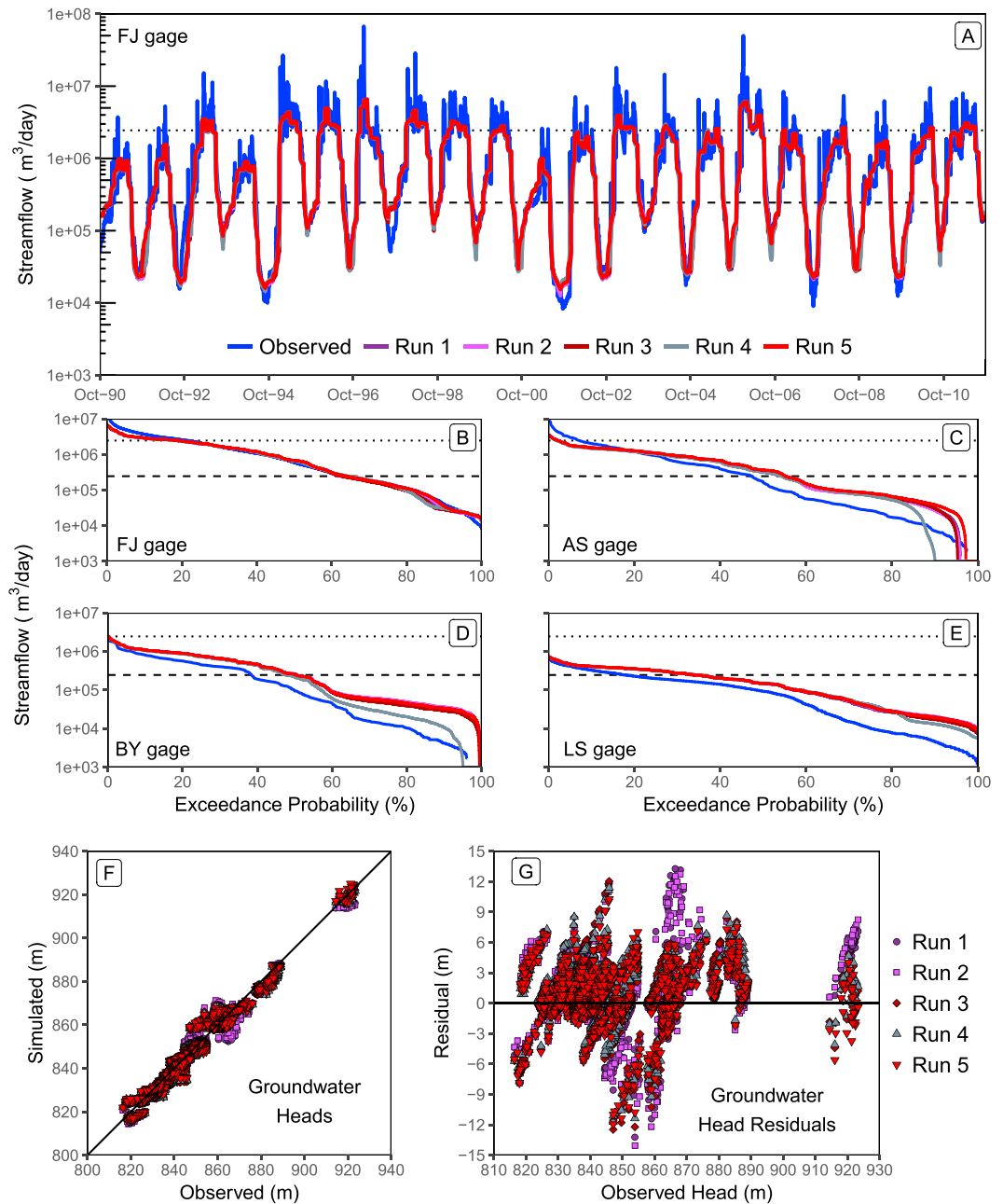
In contrast to the high variability of groundwater-surface water fluxes along the stream profile, local fluxes remain relatively constant over the 21-year simulation period. Some seasonal variations are observed in the simulation, involving either a slight upstream or downstream translocation of a gaining/losing reach transition (Figure 5). Transitions are often consistent between seasons, year after year, regardless of water year type. Some reaches show seasonal expansion/contraction patterns or reaches, with a longer area recharging the aquifer during the winter and spring months and longer sections of groundwater discharging to the stream in the summer and fall. Reaches consistently showing either gains or losses during the entire simulation period each accounted for about 25% of the length of the Scott River (Figures 5 and 10).

#### 4.2. Sensitivity Analysis

CSS showed considerable variation across the parameter sets (Figure 6). The alfalfa crop coefficient multiplier (Kc\_Alalfa\_Mult) was the most sensitive, having the highest value for each parameter set tested. The soil moisture depletion factor for wheel line irrigation of alfalfa and grain (SMDF\_A/G\_WL) and the pasture crop coefficient (Kc\_Pasture\_Mult) also showed a large degree of sensitivity for all runs. Aside from these three parameters, ordering of CSS values between sets was highly variable. Twenty-two parameters did not show sensitivity within 5% of Kc\_Alalfa\_Mult (normalized CSS < 0.05) for any of the runs. A few parameters (e.g., Kx5, Kx9, and Kc\_Native) showed relatively high sensitivity for a single run but were insensitive for other combinations of parameters.

Low streamflow at the USGS gage was the most sensitive observation group to changes in parameter values (Figure 6), followed by low flow observations at the other gages and then by groundwater heads. Medium and high streamflow observation groups did not show large sensitivities to parameter perturbations. Using

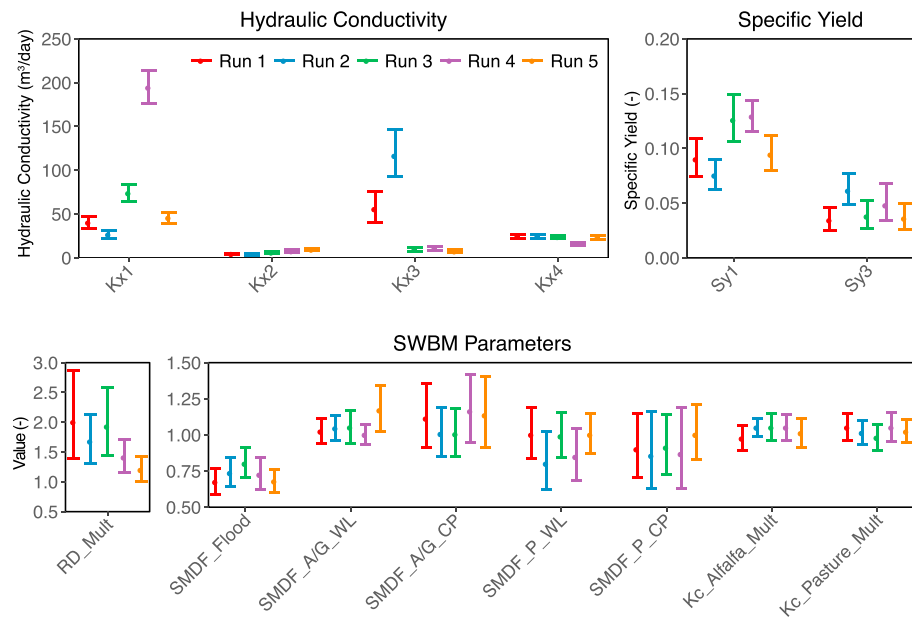




**Figure 7.** (a) Observed and simulated values of streamflow at the U.S. Geological Survey FJ gage using optimized values from all five calibration runs. (b–e) Observed and simulated streamflow exceedance probabilities at the four gaging stations in the valley. (f and g) Observed and simulated groundwater heads and head residuals for the five calibrations. Low, medium, and high flow categories are below the dashed line, between the dashed and dotted line, and above the dotted line, respectively. FJ = Fort Jones; AS = Above Serpa Lane; BY = Below Young’s Dam; LS = Lower Shackleford Creek.

the same weight for all three stream flow categories (CV = 10%) does not affect the much larger sensitivity of the low streamflow observation group when compared to the medium or high streamflow observation groups (not shown).

Parameter correlation coefficients exceeding 0.95, the general threshold at which parameters become nonuniquely estimable (Hill & Tiedeman, 2007), were only observed in two of the parameter sets. A negative correlation was found between specific storage and specific yield for hydraulic property Zones 6 and 8 in Parameter Set 2, which is expected given their relation to storativity. In Parameter Set 4 there was a strong



**Figure 8.** Optimized parameter values for the five calibration runs. Error bars show 95% linear confidence intervals calculated by UCODE\_2014. SWBM = soil water budget model.

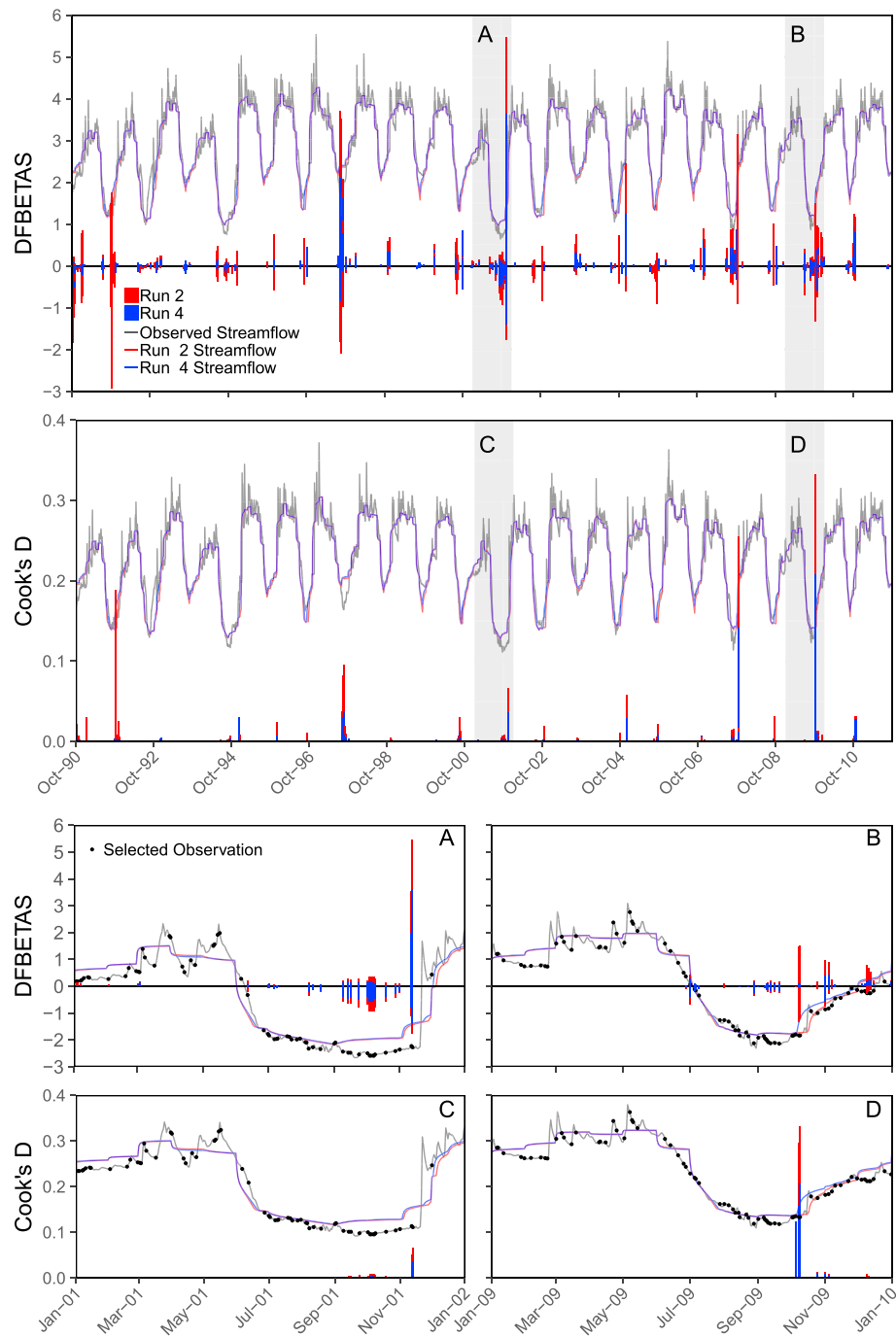
negative correlation between Wel5 and Wel7, two parameters that control MFR along the eastern boundary of the southern portion of the valley. While correlated, the parameters themselves were found to have low CSS values indicating that they were not suitable for calibration. None of the parameters selected for the calibration were correlated, suggesting that unique calibration estimates may be obtained.

### 4.3. Calibration

Based on information obtained from the sensitivity analyses, 14 parameters (Table 2) were selected for calibration. Streamflow and groundwater heads simulated by all five calibrated models show very good agreement with observed values (Figure 7). Both interannual and intra-annual streamflow variations are captured at all gage locations, with log-transformed NSEs ranging from 0.61 to 0.91 at all gages for the five calibration runs. Streamflow values are consistently underpredicted at the USGS gage during the winter and spring months when high flow events typically have a time scale much shorter than the monthly stress period. Low flows simulated at gages other than the USGS gage tend to be overestimated during the summer and fall (Figure 7). Simulated groundwater heads show a very strong correlation ( $R^2 \geq 0.98$ ) with observed values, having root-mean-square errors between 2.28 and 2.78 m. Residuals less than or equal to 1, 2, and 3 m accounted for approximately 50%, 70%, and 80% of head observations, respectively.

All five of the calibration runs converged because the objective function did not change more than 1% for three consecutive iterations (TolSOSC convergence). Final objective function values varied between about 1% and 8% difference of each other (Table 2). Although the objective function reached a similar value for all runs, estimates of several parameters varied significantly between calibrations (Figure 8). The largest variations were observed in Kx1, Kx3, and Sy1, which ranged over an order of magnitude for hydraulic conductivity and varied up to 50% for specific yield. Parameters contained within SWBM showed similar variations across runs but with much less variability due to tighter imposed constraints. None of the parameters were calibrated to unreasonable values, with only a few limited by upper or lower calibration bounds.

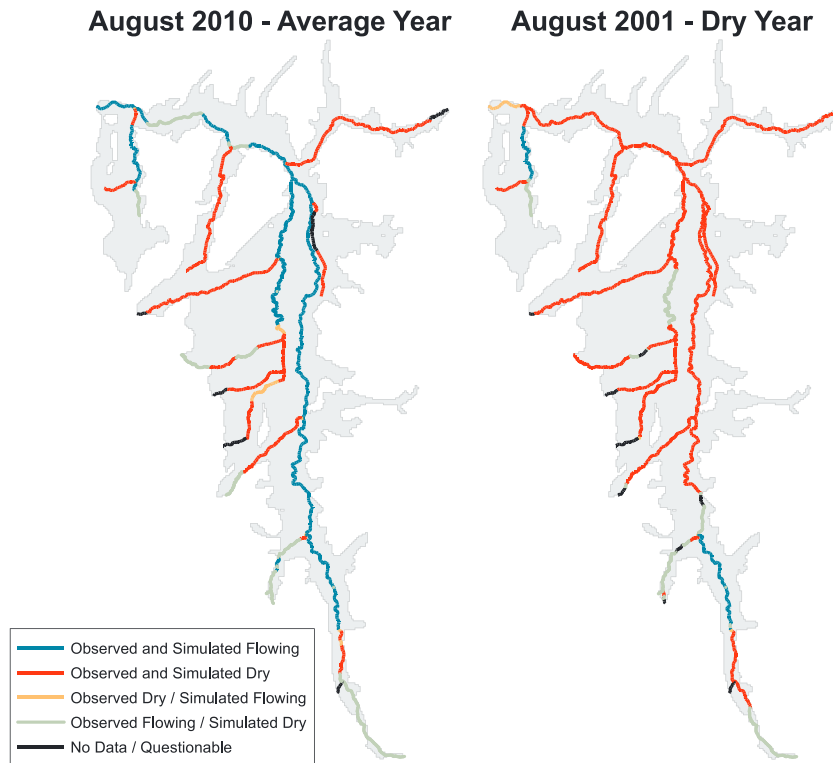
Linear 95% confidence intervals estimated by UCODE\_2014 are relatively narrow for the MODFLOW parameters when compared to the range of estimated values across all five calibration runs. For example, the largest confidence interval for Kx1 was observed in Calibration Run 4 and spanned 38 m/day. This is only 22% of the range in estimated values across all five calibration runs. Conversely, individual linear confidence intervals for SWBM parameters compared to the range in ensemble estimated parameter values are much greater, ranging from 52% to 386%. These linear confidence intervals may not reflect the true parameter



**Figure 9.** Influence statistics DFBETAS (top row) and Cook's D (bottom row) for streamflow observations at the U.S. Geological Survey gage. Lines show observed (gray) and simulated (red and blue) streamflow values that have been log-transformed and scaled to fit the DFBETAS and Cook's D axes to provide timing context. DFBETAS values are only differentiated by calibration run and not by parameter for plotting purposes. The most influential observations generally occur during or immediately following the lowest streamflow period of the year.

confidence intervals, as the modified Beale's measure ranged from 1.6 to 2.9 times the upper critical value and indicates a highly nonlinear model.

Values of DFBETAS and Cook's D (Figure 9) show that timing of the most influential observations occurs during or immediately following the lowest period of streamflow during the year. Although the seasonal timing of influential observations between the two statistics is similar, the most influential observations



**Figure 10.** Observed and simulated stream status from Calibration Run 4 at all locations for the month of August in an average and dry water year. Simulated streamflow status shows strong agreement with observed values, indicating the model provides a good representation of streamflow spatially in addition to temporally at the streamflow calibration points (Figure 8).

for the two are located in different years highlighted by the shaded regions in Figure 9. The percentage of observations exceeding critical values ranged from about 15% to 18% for DFBETAS and 6% to 7% for Cook's D. Streamflows accounted for about 64–79% of the observations exceeding critical values, which were predominantly (>90%) low streamflow observations.

#### 4.4. Model Validation

Qualitative streamflow observations of dry and flowing reaches for August in an average (2010) and dry (2001) year were matched by simulated conditions along 81% and 87% of the total stream length, respectively (Figure 10). This binary, spatially continuous information was not included during the calibration process and provides an independent check of model performance. Nearly the entire length of the main stem Scott River and most of the tributaries are accurately predicted for flowing and dry conditions. Some reaches, such as French Creek and portions of the tailings section, Patterson Creek, Big Slough, and Shackelford Creek, appear to have flows that are underpredicted by the model for both water year types (Figure 10). The qualitative, binary validation information suggests SVIHM not only is an excellent prediction tool for flow hydrographs at few existing stream gages but also performs well in predicting the spatial distribution of dry stream sections, the most critical low streamflow predictor for water managers.

### 5. Discussion

The water budget produced by the model shows the highly dynamic nature of the Scott Valley both seasonally and interannually. This is especially true for groundwater pumping and net groundwater-surface-water fluxes. While we typically associate wet years with less groundwater pumping in agricultural areas (Faunt et al., 2009), this is not always the case. Annual volume of precipitation is important from an overall water budget perspective, but precipitation amounts during the early spring, when the rainy season may overlap with the growing season, influence groundwater pumping rates. For example, 2006 had approximately 50% more total precipitation but also 12% more pumping compared with 2010 (Figure 4). This is because more

precipitation in 2010 fell during the growing season and helped satisfy crop ET demand, thereby reducing the amount of irrigation. Similar instances of this occurrence have been shown in Ruud et al. (2004) and Faunt et al. (2009).

Net fluxes between groundwater and surface water in the first half of the growing season appear to be largely controlled by precipitation and timing of runoff entering the valley from the upper watershed, while the second half is controlled by aquifer storage and groundwater pumping. Seasonal decrease in relative aquifer storage (i.e., cumulative change of groundwater storage relative to the first stress period) over the growing season (March–October) are similar in wet and dry years (Figure 4) due to the Mediterranean climate. In wet years, large gains in storage during the winter contribute to groundwater pumping and increased baseflow during the late summer. In dry years, initial groundwater storage is less and then reduced further by groundwater pumping. In the example average water year type shown, summer pumping affects groundwater storage but recovery begins already in late summer due to early arrival of storms in October (Figure 4). Similar patterns of storage change are observed in other Mediterranean, irrigated groundwater basins (Fleckenstein et al., 2004; CDM, 2008; Brush & Dogrul, 2013).

For highly nonlinear models like SVIHM, sensitivity analysis results show that CSS values and therefore the ranking of parameters by CSS vary as a function of the parameter starting value (Figure 6, top). Since selection of calibration parameters is done by CSS ranking, choice of calibration parameters varies, sometimes drastically, between sensitivity analyses. Simultaneously considering the CSS rankings from multiple sensitivity analyses provides additional information that overcomes the dependency of the (local) sensitivity analysis on the starting values. The very different parameter ranking that appeared in only one of the five sensitivity analyses (Parameter Set 2, Figure 6, top) was considered an outlier. In their more rigorous sensitivity analysis using distributed evaluation of local sensitivity analysis, Rakovec et al. (2014) found that model parameters (including those identified to be most sensitive generally) were insensitive for about 20% of the tested combinations. Here we see the reverse, with parameters that are insensitive in other runs showing greater sensitivity.

The seemingly anomalous sensitivity results for Parameter Set 2 are likely a consequence of overestimation of groundwater heads for one well in Zone 9 and two wells in Zone 5. Although the initial hydraulic conductivity for these zones in Parameter Set 2 is still relatively high at 1–2 m/day (Table 1), they are the lowest of the ensemble. The final calibration values indicate highly conductive aquifer sediments in these zones. The reason for increased sensitivity of some parameters in set three (e.g., Kx3, Kx4, Sy3, and Sy4) is not clear, as there does not appear to be a strong relationship between observations in zones three and four and their initial parameter values.

The most sensitive parameters in SVIHM are crop coefficients for alfalfa and pasture, which control water demand (ET), and the SMDF for alfalfa/grain fields, which affects how much irrigation water is applied and therefore recharge rates for that land use type. At first this appears to be counterintuitive since these parameters are contained in SWBM and subsequently filtered through the MODFLOW model. However, local differences between recharge and pumping (net extraction values) in SVIHM are highly dependent on these parameters. Head distributions resulting from groundwater pumping cannot fully explain this, as head observations generally contribute less than 20% to the CSS values. Instead, a likely driver for this sensitivity to net extraction is streamflow depletion due to groundwater pumping as low flows are proportionally most affected and those observations show the greatest contribution to CSS values. This suggests that groundwater-surface-water models in agricultural areas that fix typically unmeasured pumping rates or adjust them uniformly instead of estimating them based on spatially distributed crop demands and irrigation efficiencies (or SMDF if deficit irrigating) are missing important parameters that significantly influence calibration results. Uncertainty estimates of predicted outcomes obtained from models that do not include the most sensitive model parameters are not likely to span the true range of possible outcomes.

Seasonal average crop coefficient values for alfalfa estimated during model calibration range from 0.88 to 0.95, similar to previously published values of 0.94 for the Scott Valley (Hanson et al., 2011). The SMDF factors for alfalfa also agree with preliminary results from recent research in the Scott Valley that show alfalfa is deficit irrigated in large part due to cutting schedules (S. Orloff, personal communication, July 14th 2017) and irrigation events do not provide enough water to fully satisfy demand. Instead, storage within the effective root zone is continuously depleted as the growing season progresses. This shuts off groundwater recharge under those fields during the growing season. It is not clear whether such deficit irrigation in alfalfa

is common elsewhere. When present, not accounting for deep soil moisture depletion during the growing season would overestimate recharge from alfalfa fields (Luthin & Bianchi, 1954; Orloff & Hanson, 2000; Sanden et al., 2003).

Results from the various calibration runs indicate the model is generally converging to two locations in parameter space. Runs 1, 2, and 5 show lower values for  $K_{x1}$  and  $S_{y1}$  compared to those of Runs 3 and 4 (Figure 8). A heuristic explanation for this may be found by considering that the calibration is trying to match low streamflow values during the late summer: If storage near the Scott River (Zone 1) is low, then the hydraulic conductivity must be low as well to keep heads in the aquifer high enough in the late summer to provide a positive gradient to the stream. If heads in the aquifer drop too low streamflow is underestimated. Likewise, if storage near the river increases then the discharge per unit gradient between the aquifer and the stream increases. To compensate, near-stream hydraulic conductivity must increase to reduce the gradient between the aquifer and the stream; otherwise, low streamflow will be overestimated. This correlation between storage and hydraulic conductivity was not observed during the sensitivity analyses, possibly due to the lack of higher-order observations used to define the objective function. This may provide some insight for integrated groundwater-surface-water models in areas where knowledge of aquifer properties is limited and/or higher-order observations are not readily available.

Seasonality of fluxes between groundwater and surface water is driven by high winter and spring flows that recharge the aquifer, which in turn sustain baseflow during the summer and fall. The model produces a large amount of spatial heterogeneity in fluxes between groundwater and surface water for the Scott River, despite reaches having nearly homogenous hydraulic and cross-sectional (conductance) parameterization. Instead, the large spatial heterogeneity is due to undulations in the longitudinal streambed profile: Streambed elevations in the model were assigned using observed thalweg elevations obtained from high-resolution LiDAR and digital elevation model data and are the only stream parameters that have significant spatial variation. Even in a predominantly flat agricultural area like the Scott Valley, the profile of streambed elevations is not uniformly sloping. At the 100-m resolution of the model grid, significant nonuniformity in streambed elevations reflects typical variations in stream geomorphology. This leads to highly variable groundwater-surface water interactions, such as those seen in models that take into account geologic heterogeneity (Fleckenstein et al., 2006) or streambed geomorphic factors at multiple scales (Cardenas, 2009; Stonedahl et al., 2012). Assigned streambed elevations therefore exert a large control on interactions between groundwater and surface water in numerical models.

Influential observations were consistently representing periods of low streamflow. The low sensitivity of the model to medium and high streamflow observations stems from their origin: Medium and high streamflows are largely controlled by specified inflow to the SFR package and not by aquifer properties. During high streamflow, upper watershed streamflow input determines streamflow out of the Scott Valley at the downgradient boundary. It is orders of magnitude larger than baseflow and therefore provides little to no information about aquifer properties. During periods when baseflow dominates (summer and fall), streamflow provides significant information about aquifer flow and storage properties since flow in the stream predominantly reflects groundwater contributions. By extension, low flows also provide information about SWBM parameters that affect groundwater contributions during this time (e.g., crop coefficient and SMDF).

The ensemble calibration results show nonunique parameter estimations when calibration runs start at different locations in the parameter space. This occurred even though the inverse problem is well posed for SVIHM as the number of observations far exceed the number of parameters and no significant correlations were found between calibration parameters. Simulated streamflow and groundwater heads are similar across calibration runs despite some parameter values varying up to an order of magnitude.

Uncertainty estimates from an individual calibration run are lacking as they do not span the entire parameter space over which similar heads and streamflow are produced. Linear confidence intervals are often the only feasible measure of parameter uncertainty for models with significantly long run times as nonlinear analysis methods such as Markov chain Monte Carlo require far too many runs to be of practical use. The minimum number of model runs for a combined sensitivity analysis and calibration in this study was 202, which took a total of 5.8 days to complete when parallelized using five processors. A Markov chain Monte Carlo analysis, assuming (conservatively) a minimum of  $1 \times 10^4$  nonparallelizable model runs that take 2 hr each, would require 2.3 years. Our results suggest that integrated hydrologic models need to be rigorously evaluated as much as feasibly possible for nonlinear behavior and sensitivity relationships to more

accurately capture sensitivities and resulting uncertainties in parameter estimation and predicted values. This is particularly important in basin-scale simulation models that will ultimately be used to inform decision makers and establish uncertainty ranges as part of project risk analyses and sustainability compliance. The results presented here do not provide a conclusive, rigorous evaluation of sensitivity, such as those provided in other methods (Rakovec et al., 2014; Borgonovo et al., 2017), due to long model runs times. Instead, running a limited number of sensitivity analyses and calibrations using a wide range of starting values provides a frugal heuristic method to capture some key uncertainties associated with nonlinearity inherent in basin-scale integrated groundwater-surface-water models.

## 6. Conclusions

Agricultural demands and groundwater-surface water conditions for the Scott Valley were simulated over 21 years by weakly coupling a streamflow regression model, a soil-water budget model, and a groundwater-surface-water model. The soil-water budget model operates at the field scale, providing high spatial resolution of groundwater recharge and pumping across the model domain. Coupling the groundwater-surface-water model with the soil-water budget model provided a simple, efficient, and transparent method of estimating groundwater pumping and recharge, two of the biggest forcings for agricultural groundwater basins. Multiple sensitivity analyses show that the most sensitive model parameters are those that control pumping and recharge, which are typically unmeasured and therefore often only available by employing estimation methods. Our work suggests it is important to embed the pumping and recharge estimation model into the sensitivity and calibration process for a *a posteriori* estimation rather than using values estimated *a priori*. This provides a quantitative measure of parameter importance across submodels and a more realistic representation of spatially distributed groundwater pumping and recharge across the agricultural landscape. Streamflow observations, particularly during the driest times of the year, provide the most information about model parameters. We did not include higher-order observations (e.g., drawdowns and streamflow differences) in this analysis, but their contributions and if/how they alter the parameter rankings should be explored in future work.

Performing multiple (local) sensitivity analyses and calibrations of this nonlinear integrated model was essential to develop a more comprehensive, quantitative understanding of model parameter importance and uncertainty. As complexity, and likely nonlinearity, of hydrologic models increases, sensitivity analyses and model calibrations must explore more of the parameter space by using multiple sets of initial parameters to gain a better understanding of which model parameters are sensitive overall as opposed to sensitive within a particular area of the parameter space. Calibration of the Scott Valley Integrated Hydrologic Model generally converges to two different areas in parameter space, with both producing results that show good agreement with observations. The use of multiple parallel calibrations revealed greater uncertainty in some model parameters that a single calibration could not detect. Optimized parameter values indicate that groundwater recharge from alfalfa fields is negligible during the growing season due to deficit irrigation. Linear 95% confidence intervals of parameter values calculated within each calibration run are generally very different than the range in estimated parameter values obtained across the five calibrations. Further research into parameter and prediction uncertainty of these weakly coupled, highly nonlinear models may utilize more rigorous nonlinear methods to more precisely define parameter uncertainty.

The Scott Valley Integrated Hydrologic Model shows that a weakly coupled, computationally efficient model can be successfully employed in lieu of an iteratively or fully coupled integrated model to simulate highly dynamic groundwater-surface-water interactions in an agricultural watershed. Computational efficiency and the ability to adjust model structure is an important consideration when developing models to meet the needs of various water managers and stakeholders. The model needs to be complex enough to capture salient hydrologic processes yet usable, modifiable, and capable of having its results communicated to and understood by a broad audience. California is an example of where models like this are likely to see expanded use, as integrated groundwater basin modeling and stakeholder outreach is an important component of the Sustainable Groundwater Management Act regulation that is currently being implemented.

## Notation

$\theta$	Available soil water
$\theta_{\max}$	Field capacity

$P$	Precipitation
ET	Evapotranspiration
$ET_0$	Reference evapotranspiration
AW	Applied water
$R$	Recharge
AE	Application efficiency (irrigation efficiency)
SMDF	Soil moisture depletion factor
$K_c$	Crop coefficient
$k$	Subscript denoting the day
$\phi$	Objective function value
$y_i$	Observed value
$y_i'$	Simulated value
$w_i$	Observation weight
$i$	Observation number
ND	Total number of observations
$\sigma_i^2$	Observation error variance
$CV_i$	Coefficient of variation for streamflow observation
$\bar{y}_i$	Mean of observations
$n$	Number of observed-simulated data pairs
$b_j$	$j$ th parameter
$b$	vector containing initial parameter values
$CSS_{j,s}$	Composite scaled sensitivity of the $j$ th parameter for parameter set $s$
$CSS_s$	Array of CSS values for parameter $j$

#### Acknowledgments

This research was funded through the California North Coast Regional Water Quality Control Board (SWRCB Agreement 14-020-110). We would like to thank Jim Morris, Sari Sommarstrom, Preston Harris, Erich Yokel, and the Scott Valley groundwater advisory committee. We would also like to thank the reviewers who provided very helpful comments and greatly improved the quality of the manuscript. A link to the SVIHM GitHub repository, which contains all of the model files and postprocessing scripts, can be found in the supporting information. This paper is dedicated to the memory of our wonderful friend, colleague, and researcher, Steve Orloff.

#### References

- Ashby, S. F., & Falgout, R. D. (1996). A Parallel Multigrid Preconditioned Conjugate Gradient Algorithm for Groundwater Flow Simulations. *Nuclear Science and Engineering*, *124*(1), 145–159.
- Barlow, P. M., & Leake, S. A. (2012). Streamflow depletion by wells—Understanding and managing the effects of groundwater pumping on streamflow. U.S. Geological Survey Circular 1376.
- Borgonovo, E., Lu, X., Plischke, E., Rakovec, O., & Hill, M. C. (2017). Making the most out of a hydrological model data set: Sensitivity analyses to open the model black-box. *Water Resources Research*, *53*, 7933–7950. <https://doi.org/10.1002/2017WR020767>
- Brunner, P., & Simmons, C. T. (2012). HydroGeoSphere: A fully integrated, physically based hydrological model. *Ground Water*, *50*(2), 170–176. <https://doi.org/10.1111/j.1745-6584.2011.00882.x>
- Brunner, P., Simmons, C. T., Cook, P. G., & Therrien, R. (2010). Modeling surface water-groundwater interaction with MODFLOW: Some considerations. *Ground Water*, *48*(2), 174–180. <https://doi.org/10.1111/j.1745-6584.2009.00644.x>
- Brush, C. F., Dogrul, E. C., & Kadir, T. N. (2013). *Development and calibration of the California Central Valley groundwater-surface water simulation model (C2VSim), version 3.02-CG*. California Department of Water Resources: Bay-Delta Office.
- California Department of Water Resources (2000). 2000 Siskiyou County land use survey data.
- California Department of Water Resources (2016a). *Integrated Water Flow Model: IWFm-2015, Theoretical documentation*. California Department of Water Resources: Bay-Delta Office.
- California Department of Water Resources (2016b). *IWFm demand calculator: IDC-2015, Theoretical documentation and user's manual*. California Department of Water Resources: Bay-Delta Office.
- Cardenas, M. B. (2009). Stream-aquifer interactions and hyporheic exchange in gaining and losing sinuous streams. *Water Resources Research*, *45*, W06429. <https://doi.org/10.1029/2008WR007651>
- CDM (2008). Butte Basin Groundwater Model Update (Phase II Report, Technical Report).
- Chen, X., & Yin, Y. (2001). Streamflow depletion: Modeling of reduced baseflow and induced stream infiltration from seasonally pumped wells. *Journal of the American Water Resources Association*, *37*(1), 185–195. <https://doi.org/10.1111/j.1752-1688.2001.tb05485.x>
- Clark, M. P., & Kavetski, D. (2010). Ancient numerical daemons of conceptual hydrological modeling: 1. Fidelity and efficiency of time stepping schemes. *Water Resources Research*, *46*, W10510. <https://doi.org/10.1029/2009WR008894>
- Clark, M. P., Slater, A. G., Rupp, D. E., Woods, R. A., Vrugt, J. A., Gupta, H. V., et al. (2008). Framework for Understanding Structural Errors (FUSE): A modular framework to diagnose differences between hydrological models. *Water Resources Research*, *44*, W00B02. <https://doi.org/10.1029/2007WR006735>
- Cooley, R. L. (2004). *A theory for modeling ground-water flow in heterogeneous media*, U.S. Geological Survey Professional Paper 1679 (pp. 56). <https://doi.org/10.3133/pp1679>
- Cooley, R. L., & Naff, R. L. (1990). Regression modeling of ground-water flow, *Techniques of water-resources investigations of the United States Geological Survey, Book 3, Chapter B4* (pp. 241).
- De Silva, C. S., & Rushton, K. R. (2007). Groundwater recharge estimation using improved soil moisture balance methodology for a tropical climate with distinct dry seasons. *Hydrological Sciences Journal*, *52*(5), 1051–1067. <https://doi.org/10.1623/hysj.52.5.1051>
- Deas, M. L. (2006). Water supply indices: Year types for the Scott River Basin (Technical Report), 29 pages, Watercourse Engineering: Davis, CA.
- Dettinger, M. D., Udall, B., & Georgakakos, A. (2015). Western water and climate change. *Ecological Applications*, *25*(1), 2069–2093. <https://doi.org/10.1890/15-0938.1>



- Dieter, C., Maupin, M., Caldwell, R., Harris, M., Ivahnenko, T., Lovelace, J., et al. (2018). *Estimated use of water in the United States in 2015* (pp. 1441). U.S. Geological Survey Circular. <https://doi.org/10.3133/cir1441>
- Dogruel, E. C., Kadir, T. N., & Brush, C. H. (2018). *IWFM demand calculator, IDC-2015, Revision 68, theoretical documentation and user's manual* (pp. 325). California Department of Water Resources: Bay-Delta Office. <https://doi.org/10.3133/cir1441>
- Echols, K. G. (1991). Scott River flow augmentation study (Technical report): California Department of Water Resources.
- Faunt, C. C., Hanson, R. T., Belitz, K., Schmid, W., Predmore, S. P., Rewis, D. L., & McPherson, K. (2009). *Groundwater availability of the Central Valley Aquifer California* (pp. 246): U.S. Geological Survey Professional Paper 1766.
- Fleckenstein, J., Anderson, M., Fogg, G., & Mount, J. (2004). Managing surface water-groundwater to restore fall flows in the Cosumnes River. *Journal of Water Resources Planning and Management*, 130(4), 301–310. [https://doi.org/10.1061/\(ASCE\)0733-9496\(2004\)130:4\(301\)](https://doi.org/10.1061/(ASCE)0733-9496(2004)130:4(301))
- Fleckenstein, J., Niswonger, R. G., & Fogg, G. E. (2006). River-aquifer interactions, geologic heterogeneity, and low-flow management. *Ground Water*, 44(6), 837–852. <https://doi.org/10.1111/j.1745-6584.2006.00190.x>
- Foglia, L., Hill, M. C., Mehl, S. W., & Burlando, P. (2009). Sensitivity analysis, calibration, and testing of a distributed hydrological model using error-based weighting and one objective function. *Water Resources Research*, 45, W06427. <https://doi.org/10.1029/2008WR007255>
- Foglia, L., McNally, A., Hall, C., Ledesma, L., & Hines, R. (2013). Scott Valley Integrated Hydrologic Model: Data collection, analysis, and water budget (Technical report). Davis: University of California.
- Foglia, L., McNally, A., & Harter, T. (2013). Coupling a spatiotemporally distributed soil water budget with stream-depletion functions to inform stakeholder-driven management of groundwater-dependent ecosystems. *Water Resources Research*, 49, 7292–7310. <https://doi.org/10.1002/wrcr.20555>
- Foglia, L., Neumann, J., Tolley, D. G., Orloff, S. B., Snyder, R. L., & Harter, T. (2018). Modeling guides groundwater management in a basin with river-aquifer interactions. *California Agriculture*, 72(1), 84–95. <https://doi.org/10.3733/ca.2018a0011>
- Furman, A. (2008). Modeling coupled surface-subsurface flow processes: A review. *Vadose Zone Journal*, 7(2), 741–756. <https://doi.org/10.2136/vzj2007.0065>
- Hanson, R. T., Boyce, S. E., Schmid, W., Hughes, J. D., Mehl, S. M., Leake, S. A., et al. (2014). One-Water Hydrologic Flow Model (MODFLOW-OWHM), U.S. Geological Survey Techniques and Methods 6-A51.
- Hanson, R. T., Flint, L. E., Flint, A. L., Dettinger, M. D., Faunt, C. C., Cayan, D., & Schmid, W. (2012). A method for physically based model analysis of conjunctive use in response to potential climate changes. *Water Resources Research*, 48, W00L08. <https://doi.org/10.3133/sir20145111>
- Hanson, B., Orloff, S., & Putnam, D. (2011). Drought irrigation strategies for Alfalfa. *Agricultural and Natural Resources (University of California)*, 8448, 1–10.
- Hanson, R. T., Ritchie, A. B., Boyce, S. E., Galanter, A. E., Ferguson, I. A., Flint, L. E., & Henson, W. R. (2018). Rio Grande transboundary integrated hydrologic model and water-availability analysis. New Mexico and Texas, United States, and Northern Chihuahua, Mexico: U.S. Geological Survey Open-File Report 2018-1091.
- Hanson, R., Schmid, W., Faunt, C. C., Lear, J., & Lockwood, B. (2014). Integrated hydrologic model of Pajaro Valley, Santa Cruz and Monterey Counties, California. *U.S. Geological Survey Scientific Investigations Report*, 2014–5111. <https://doi.org/10.3133/sir20145111>
- Harbaugh, A. W. (2005). MODFLOW-2005, The U.S. Geological Survey modular ground-water model—The ground-water flow process: U.S. Geological Survey Techniques and Methods 6-A51.
- Hill, M. C., & Tiedeman, C. R. (2007). *Effective groundwater model calibration: With analysis of data, sensitivities, predictions, and uncertainty*. Hoboken, NJ: John Wiley & Sons, Inc.
- Hill, M. C., Kavetski, D., Clark, M., Ye, M., Arabi, M., Lu, D., et al. (2016). Practical use of computationally frugal model analysis methods. *Groundwater*, 54(2), 159–170. <https://doi.org/10.1111/gwat.12330>
- Howard, J., & Merrifield, M. (2010). Mapping groundwater dependent ecosystems in California. *PLoS ONE*, 5(6), e11249. <https://doi.org/10.1371/journal.pone.0011249>
- Huntington, J. L., & Niswonger, R. (2012). Role of surface-water and groundwater interactions on projected summertime stream-flow in snow dominated regions: An integrated modeling approach. *Water Resources Research*, 48, W11524. <https://doi.org/10.1029/2012WR012319>
- Irwin, W. P. (1990). The San Andreas Fault System, California: U.S. Geological Survey Professional Paper 1515.
- Jain, S., & Sudheer, K. (2008). Fitting of hydrologic models: A close look at the Nash-Sutcliffe index. *Journal of Hydrologic Engineering*, 10(13), 981–986. [https://doi.org/10.1061/\(ASCE\)1084-0699\(2008\)13:10\(981\)](https://doi.org/10.1061/(ASCE)1084-0699(2008)13:10(981))
- James M. Montgomery Consulting Engineers (1990). Central Valley ground-surface water model, Central Valley, California (Technical Report).
- Jousma, G., Bear, J., & Haines, Y. Y. (2012). *Groundwater contamination: Use of models in decision-making, Proceedings of the International Conference on Groundwater Contamination, Amsterdam, The Netherlands, 26–29 October 1987 Organized by the International ground water modeling center (IGWMC)*. Indianapolis - Delft: Springer Science and Business Media.
- Kavetski, D., & Clark, M. P. (2010). Ancient numerical daemons of conceptual hydrological modeling: 2. Impact of time stepping schemes on model analysis and prediction. *Water Resources Research*, 46, W10511. <https://doi.org/10.1029/2009WR008896>
- Klove, B., Ala-Aho, P., Bertrand, G., Gurdak, J. J., Kupfersberger, H., Kværner, J., et al. (2014). Climate change impacts on groundwater and dependent ecosystems. *Journal of Hydrology*, 518, 250–266. <https://doi.org/10.1016/j.jhydrol.2013.06.037>
- Kollet, S. J., & Maxwell, R. M. (2006). Integrated surface-groundwater flow modeling: A free-surface overland flow boundary condition in a parallel groundwater flow model. *Advances in Water Resources*, 29(7), 945–958. <https://doi.org/10.1016/j.advwatres.2005.08.006>
- La Vigna, F., Hill, M. C., Rossetto, R., & Mazza, R. (2016). Parameterization, sensitivity analysis, and inversion: An investigation using groundwater modeling of the surface-mined Tivoli-Guidonia basin (Metropolitan City of Rome, Italy). *Hydrogeology Journal*, 24(6), 1423–1441. <https://doi.org/10.1007/s10040-016-1393-z>
- Lane, B. A., Sandoval-Solis, S., & Porse, E. C. (2015). Environmental flows in a human-dominated system: Integrated water management strategies for the Rio Grande/Bravo Basin. *River Research and Applications*, 31. <https://doi.org/10.1002/rra.2804>
- Luthin, J. N., & Bianchi, W. (1954). Alfalfa and water table levels. *California Agriculture*, 8(5), 4–5.
- Mack, S. (1958). Geology and ground-water features of Scott Valley, Siskiyou County, California, U.S.: Geological Survey water-supply paper 1462.
- Manghi, F., Mortazavi, B., Crother, C., & Hamdi, M. R. (2012). Estimating regional groundwater recharge using a hydrological budget method. *Water Resources Management*, 23(12), 2475–2489. <https://doi.org/10.1007/s11269-008-9391-0>
- Markstrom, S. L., Niswonger, R. G., Regan, R. S., Prudic, D. E., & Barlow, P. M. (2008). SFLOW-coupled ground-water and surface water flow model based on the integration of the Precipitation-Runoff Modeling System (PRMS) and the Modular Ground-Water Flow Model (MODFLOW-2005): U.S. Geological Survey, Techniques and Methods 6-D1.

- McMahon, P., Dennehy, K., Michel, R., Sophocleous, M., Ellett, K., & Hurlbut, D. B. (2003). Water movement through thick unsaturated zones overlying the Central High Plains Aquifer, Southwestern Kansas, 2000-2001. U.S. Geological Survey Water-Resources Investigations Report 03-4171.
- Mendoza, P. A., Clark, M. P., Barlage, M., Rajagopalan, B., Samaniego, L., Abramowitz, G., & Gupta, H. (2015). Are we unnecessarily constraining the agility of complex process-based models? (1). <https://doi.org/10.1002/2014WR015820>
- Miller, C. T., & Pinder, G. F. (2004). *Computational methods in water resources: Volume 2: Proceedings of the XVth International Conference on Computational Methods in Water Resources (CMWR XV)*. Chapel Hill, USA: NC.
- NCRWQCB (2005). *Staff report for the action plan for the Scott River watershed sediment and temperature total maximum daily loads*. Santa Rosa, CA: North Coast Regional Water Quality Control Board.
- Nash, J. E., & Sutcliffe, J. V. (1970). River flow forecasting through conceptual models Part 1—A discussion of principles. *Journal of Hydrology*, 10, 282–290. [https://doi.org/10.1016/0022-1694\(70\)90255-6](https://doi.org/10.1016/0022-1694(70)90255-6)
- Niswonger, R. G., Panday, S., & Ibaraki, M. (2011). MODFLOW-NWT, A Newton formulation for MODFLOW-2005, U.S. Geological Survey Techniques and Methods 6–A37.
- Orloff, S. B., & Hanson, B. (2000). Monitoring alfalfa water use with soil moisture sensors. In *proceedings of the 30th California Alfalfa Symposium*.
- Phillips, S. P., Lewis, D. L., & Traum, J. A. (2015). Hydrologic model of the Modesto Region, California, 1960–2004, U.S. Geological Survey Scientific Investigations Report.
- Poeter, E. P., & Hill, M. C. (1998). Documentation of UCODE, A computer code for universal inverse modeling, U.S. Geological Survey Water-Resources Investigations Report 98-4080, 457–462.
- Poeter, E. P., Hill, M. C., Lu, D., & Tiedeman, C. R. (2014). *UCODE\_2014, with new capabilities to define parameters unique to predictions, calculate weights using simulated values, estimate parameters with SVD, evaluate uncertainty with MCMC, and more, Integrated Groundwater Modeling Center*. IGWMC: Colorado School of Mines.
- Pokhrel, Y. N., Koirala S., Yeh, P. J. F., Hanasaki, N., Longuevergne, L., Kanae, S., & Oki, T. (2015). Incorporation of groundwater pumping in a global Land Surface Model with the representation of human impacts. *Water Resources Research*, 51, 78–96. <https://doi.org/10.1002/2014WR015602>
- Prudic, D. (1989). Documentation of a computer program to simulate stream-aquifer relations using a modular, finite-difference ground-water flow model, U.S. Geological Survey Open-File Report 88-729.
- Prudic, D. E., Konikow, L. F., & Banta, E. R. (2004). A new streamflow-routing (SFR1) package to simulate stream-aquifer interaction with MODFLOW-2000: U.S. Geological Survey Open-File Report 2004-1042, 104 pages.
- Rakovec, O., Hill, M. C., Clark, M. P., Weerts, A. H., Teuling, A. J., & Uijlenhoet, R. (2014). Distributed Evaluation of Local Sensitivity Analysis (DELSA), with application to hydrologic models. *Water Resources Research*, 50, 409–426. <https://doi.org/10.1002/2013WR014063>
- Ramireddygar, S. R., Sophocleous, M. A., Koelliker, J. K., Perkins, S. P., & Govindaraju, R. S. (2000). Development and application of a comprehensive simulation model to evaluate impacts of watershed structures and irrigation water use on streamflow and groundwater: The case of Wet Walnut Creek Watershed, Kansas, USA. *Journal of Hydrology*, 236, 223–246. [https://doi.org/10.1016/S0022-1694\(00\)00295-X](https://doi.org/10.1016/S0022-1694(00)00295-X)
- Roark, M., & Healy, D. F. (1998). Quantification of deep percolation from two flood irrigated alfalfa fields, Roswell basin, New Mexico: U.S. Geological Survey Water-Resources Investigations Report 98-4096, 39 pages.
- Ruud, N., Harter, T., & Naugle, A. (2004). Estimation of groundwater pumping as closure to the water balance of a semi-arid, irrigated agricultural basin. *Journal of Hydrology*, 297, 51–73. <https://doi.org/10.1016/j.jhydrol.2004.04.014>
- S. S. Papadopoulos and Associates, Inc. (2000). Middle Rio Grande Water Supply Study (Technical Report), p. 338.
- S. S. Papadopoulos and Associates, Inc. (2012). Groundwater conditions in Scott Valley, California (Technical Report), p. 130.
- SWRCB (1999). Water Right Decision 1641, California State Water Resources Control Board, California Environmental Protection Agency, 207 pages.
- Sanden, B., Poole, G., & Hanson, B. (2003). Soil moisture monitoring in alfalfa: Does it pay?. In *Proceedings of the 33rd California Alfalfa Symposium*.
- Schmid, W., & Hanson, R. (2005). The Farm Process Version 2 (FMP2) for MODFLOW-2005: Modifications and upgrades to FMP1, U.S. Geological Survey Techniques and Methods 6-A-32. p. 116.
- Schmid, W., Hanson, R., Maddock, T. I., & Leake, S. A. (2000). User Guide for the Farm Process (FMP1) for the U.S. Geological Survey's modular three-dimensional finite-difference ground-water flow model, MODFLOW-2000 U.S. Geological Survey Techniques and Methods 6-A17. p. 140.
- Singh, A. (2014). Groundwater resources management through the applications of simulation modeling: A review. *Science of the Total Environment*, 414–423. <https://doi.org/10.1016/j.scitotenv.2014.05.048>
- Stark, J. R., Armstrong, D. S., & Zwilling, D. R. (1994). Stream-aquifer interactions in the Straight River area, Becker and Hubbard Counties, Minnesota (U.S. Geological Survey Water-Resources Investigation Report 94-4009). p. 92.
- Stonedahl, S. H., Harvey, J. W., Detty, J., Aubeneau, A., & Packman, A. I. (2012). Physical controls and predictability of stream hyporheic flow evaluated with a multiscale model. *Water Resources Research*, 48, W10513. <https://doi.org/10.1029/2011WR011582>
- Taylor, R. G., Scanlon, B., Döll, P., Rodell, M., van Beek, R., Wada, Y., et al. (2012). Ground water and climate change. *Nature Climate Change*, 3, 1–9. <https://doi.org/10.1038/NCLIMATE1744>
- Therrien, R., & Sudicky, E. A. (1996). Three-dimensional analysis of variably-saturated flow and solute transport in discretely-fractured porous media. *Journal of Contaminant Hydrology*, 23(95), 1–44. [https://doi.org/10.1016/0169-7722\(95\)00088-7](https://doi.org/10.1016/0169-7722(95)00088-7)
- U.S. Department of War (1938). Report of the Chief of Engineers, U.S. Army (Technical Report): U.S. Army Corps of Engineers.
- University of California at Davis (2016). Scott Valley, Siskiyou County, California: Voluntary Private Well Water Monitoring Program, Spring 2006 - January 2016 (PowerPoint Slides). retrieved from <http://groundwater.ucdavis.edu/Research/ScottValley/>
- Unland, N. P., Cartwright, I., Andersen, M. S., Rau, G. C., Reed, J., Gilfedder, B. S., et al. (2013). Investigating the spatio-temporal variability in groundwater and surface water interactions: A multi-technique approach. *Hydrology and Earth System Sciences*, 17(9), 3437–3453. <https://doi.org/10.5194/hess-17-3437-2013>
- U.S. Geological Survey (2015). National Water Information System data available on the World Wide Web (USGS Water Data for the Nation). accessed [October 11, 2015], at URL [<http://waterdata.usgs.gov/nwis/>], U.S. Geological Survey.
- Van Kirk, R. W., & Naman, S. W. (2008). Erratum to Van Kirk and Naman (2008). Relative effects of climate and water use on base-flow trends in the Lower Klamath Basin. *Journal of the American Water Resources Association*, 44(4), 1053–1054. <https://doi.org/10.1111/j.1752-1688.2008.00235.x>

- Van Roosmalen, L., Sonnenborg, T. O., & Jensen, K. H. (2009). Impact of climate and land use change on the hydrology of a large-scale agricultural catchment. *Water Resources Research*, 45, W00A15. <https://doi.org/10.1029/2007WR006760>
- Watershed Sciences (2010). LiDAR remote sensing data collection, Scott Valley, California (*Technical Report*). Submitted to Tetra Tech on November 11th, 2010, 28 pages.
- Weaver, J. E. (1926). *Root development of field crops* (pp. 167). New York, NY: McGraw-Hill.
- Wilson, J. L., & Guan, H. (2013). Mountain-block hydrology and mountain-front recharge. In J. Hogan, F. Phillips, & B. Scanlon (Eds.), *Groundwater recharge in a desert environment: The southwestern United States* (pp. 113–137). Washington, DC: American Geophysical Union.
- Yager, R. M. (1998). Detecting influential observations in nonlinear regression modeling of groundwater flow. *Water Resources Research*, 34(7), 1623–1633.
- Yin, L., Hu, G., Huang, J., Wen, D., Dong, J., Wang, X., & Li, H. (2011). Groundwater-recharge estimation in the Ordos Plateau, China: Comparison of methods. *Hydrogeology Journal*, 19(8), 563–157. <https://doi.org/10.1007/s10040-011-0777-3>

# **Exhibit C**

**RELATIVE EFFECTS OF CLIMATE AND WATER USE ON  
BASE-FLOW TRENDS IN THE LOWER KLAMATH BASIN<sup>1</sup>***Robert W. Van Kirk and Seth W. Naman<sup>2</sup>*

**ABSTRACT:** Since the 1940s, snow water equivalent (SWE) has decreased throughout the Pacific Northwest, while water use has increased. Climate has been proposed as the primary cause of base-flow decline in the Scott River, an important coho salmon rearing tributary in the Klamath Basin. We took a comparative-basin approach to estimating the relative contributions of climatic and non-climatic factors to this decline. We used permutation tests to compare discharge in 5 streams and 16 snow courses between “historic” (1942-1976) and “modern” (1977-2005) time periods, defined by cool and warm phases, respectively, of the Pacific Decadal Oscillation. April 1 SWE decreased significantly at most snow courses lower than 1,800 m in elevation and increased slightly at higher elevations. Correspondingly, base flow decreased significantly in the two streams with the lowest latitude-adjusted elevation and increased slightly in two higher-elevation streams. Base-flow decline in the Scott River, the only study stream heavily utilized for irrigation, was larger than that in all other streams and larger than predicted by elevation. Based on comparison with a neighboring stream draining wilderness, we estimate that 39% of the observed 10 Mm<sup>3</sup> decline in July 1-October 22 discharge in the Scott River is explained by regional-scale climatic factors. The remainder of the decline is attributable to local factors, which include an increase in irrigation withdrawal from 48 to 103 Mm<sup>3</sup>/year since the 1950s.

(KEY TERMS: surface water hydrology; climate variability/change; rivers/streams; Klamath River; salmon; permutation tests.)

Van Kirk, Robert W. and Seth W. Naman, 2008. Relative Effects of Climate and Water Use on Base-Flow Trends in the Lower Klamath Basin. *Journal of the American Water Resources Association (JAWRA)* 44(4):1035-1052. DOI: 10.1111/j.1752-1688.2008.00212.x

**INTRODUCTION**

Snowmelt is an important contributor to discharge in nearly all major rivers of the western United States (U.S.). Analyses of hydrometeorological data from this region show that climate warming has decreased the percentage of precipitation falling as snow and accelerated snowpack melt, resulting in

earlier peak runoff and lower base flows (Hamlet *et al.*, 2005; Mote *et al.*, 2005; Regonda *et al.*, 2005; Stewart *et al.*, 2005; Mote, 2006). These trends may have begun nearly a century ago but are well documented to have occurred over the past 60 years (Hamlet *et al.*, 2005; Mote, 2006). Climate patterns in the Pacific Northwest over this time period have been affected both by long-term, systematic warming and by decadal-scale oscillations (Hamlet *et al.*, 2005;

<sup>1</sup>Paper No. JAWRA-07-0074-P of the *Journal of the American Water Resources Association (JAWRA)*. Received June 12, 2007; accepted December 12, 2007. © 2008 American Water Resources Association. **Discussions are open until February 1, 2009.**

<sup>2</sup>Respectively, Associate Professor, Department of Mathematics, Idaho State University, 921 S. 8th Ave., Stop 8085, Pocatello, Idaho 83209; and Research Assistant, Department of Fisheries Biology, Humboldt State University, Arcata, California 95521 (E-Mail/Van Kirk: rob.vankirk@gmail.com).

Regonda *et al.*, 2005; Stewart *et al.*, 2005). In particular, the Pacific Decadal Oscillation (PDO) cycled through a cool phase (increased snowpack and streamflow) from the mid-1940s to 1976 and through a warm phase (decreased snowpack and streamflow) from 1977 through at least the late 1990s (Minobe, 1997; Mote, 2006). Regardless of the degree to which climatic trends since the 1940s reflect short-term *vs.* long-term processes, base flow in Pacific Northwest rain-snow systems is strongly dependent on timing and amount of snowmelt, which is reflected by April 1 snow water equivalent (SWE) (Gleick and Chalecki, 1999; Leung and Wigmosta, 1999; McCabe and Wolock, 1999). Trends in April 1 SWE appear to be driven primarily by temperature, which, along the Pacific Coast, is a function of elevation and latitude (Knowles and Cayan, 2004; Mote, 2006), and secondarily by precipitation (Hamlet *et al.*, 2005; Mote *et al.*, 2005; Stewart *et al.*, 2005).

Concurrent with the observed declines in April 1 SWE over the past 60 years, water use in the Pacific Northwest has increased substantially. Total water withdrawal in California, Idaho, Oregon, and Washington increased 82% between 1950 and 2000, with irrigation accounting for nearly half of this increase (MacKichan, 1951; Hutson *et al.*, 2004). Accordingly, declines in streamflow over the past half century could be caused by a combination of continental-scale climatic factors and watershed-scale increases in water use rather than by climatic factors alone. Although climate models diverge with respect

to future trends in precipitation over this region, there is widespread agreement that the trend toward lower SWE and earlier snowmelt will continue (Leung and Wigmosta, 1999; McCabe and Wolock, 1999; Miller *et al.*, 2003a; Snyder *et al.*, 2004; Barnett *et al.*, 2005; Zhu *et al.*, 2005; Vicuna *et al.*, 2007). Thus, availability of water resources under future climate scenarios is expected to be most limited during the late summer (Gleick and Chalecki, 1999; Miles *et al.*, 2000). Development and implementation of appropriate water management strategies to deal with these shortages will require distinction between the component of late-summer flow decrease attributable to large-scale climatic factors and that attributable to local-scale changes in water use. Management actions implemented at the watershed or basin scale have the potential to reverse declines in streamflow that have been caused by increased water use but will not reverse those caused by continental-scale climatic factors.

The lower Klamath Basin in northern California (Figure 1) provides an important example of the need to distinguish the effects of climate on observed declines in base flow from those of water use. The Klamath River and its tributaries support populations of anadromous fish species with economic, ecological, and cultural importance. Of these, coho salmon (*Oncorhynchus kisutch*, Southern Oregon/Northern California Coasts Evolutionarily Significant Unit) are listed as threatened under the U.S. Endangered Species Act (Good *et al.*, 2005). In addition,

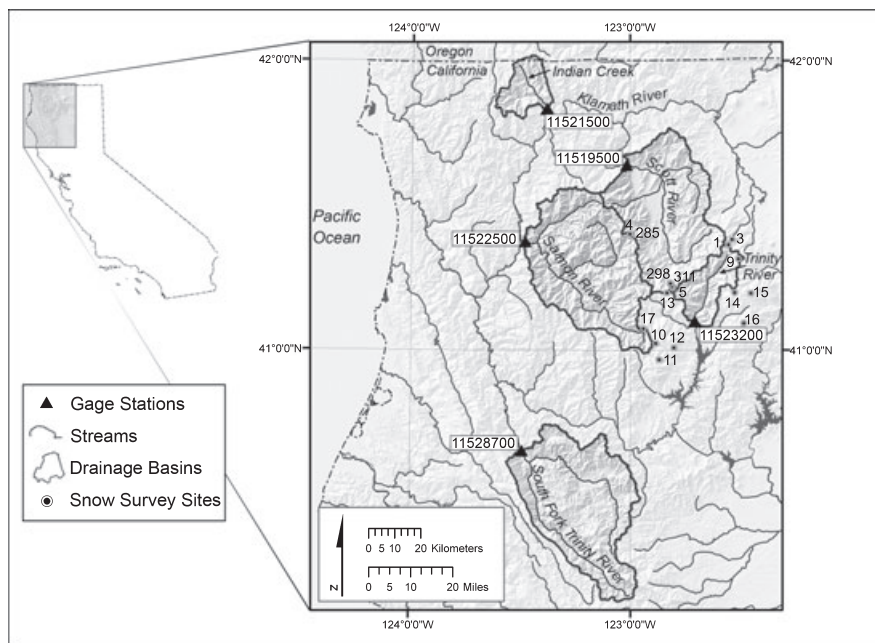


FIGURE 1. Map of Lower Klamath Basin, California, Showing Study Watersheds, Stream Gages, and Snow Courses Used in This Study. Snow course and stream gage numbers correspond to those listed in Tables 1 and 2.

steelhead trout (*Oncorhynchus mykiss*) and Chinook salmon (*Oncorhynchus tshawytscha*) in the lower Klamath Basin are of special concern or are at risk of extinction (Nehlsen *et al.*, 1991). Habitat degradation, over-exploitation, and reductions in water quality and quantity have been implicated in declines of these species (Nehlsen *et al.*, 1991; Brown *et al.*, 1994; Good *et al.*, 2005). In particular, low late-summer and early fall streamflow in several tributaries is a major factor limiting survival of juvenile coho salmon (NRC, 2003; CDFG, 2004). Increasing late-summer tributary flow is a major objective of coho salmon recovery efforts, particularly in the Scott River (Figure 1), the most important coho salmon spawning and rearing stream in the basin (Brown *et al.*, 1994; NRC, 2003; CDFG, 2004). If reduction in Scott River base-flow has been caused primarily by climatic factors, as has been proposed by Drake *et al.* (2000), then flow objectives for coho salmon recovery may not be attainable through local management, and the success of other recovery objectives (e.g., habitat restoration) may be limited by continued low base flows. On the other hand, if reduction in base flow is due in substantial part to changes in amount, timing and source of water withdrawal, then at least that particular component of flow reduction caused by water-use factors could be mitigated through local management actions.

### *Research Approach and Objectives*

The goal of this study is to distinguish the relative effects of regional-scale climatic factors from those of local-scale factors on trends in late-summer and early fall flows in lower Klamath tributaries, with particular emphasis on the Scott River. We aim to provide water and fisheries managers with information they need to develop realistic and attainable base-flow objectives for fisheries recovery. Ideally, such a study would analyze water-use data, including location and timing of withdrawals, source of water withdrawn (ground *vs.* surface), and rate of consumptive use. Furthermore, in agricultural settings, it is desirable to analyze the type of crops irrigated, method of irrigation application, amount of return flow, and pathways (ground *vs.* surface) by which return flow enters stream channels. Unfortunately, almost no data of these types are available for the watersheds of the lower Klamath Basin, including that of the Scott River, where a large amount of irrigated agriculture occurs. Thus, as an expeditious, first-order attempt to distinguish between effects of climate *vs.* water use on base flow declines, we use statistical analysis of existing SWE and streamflow data from across the basin. Results of this study can then be used to

prioritize future data collection and modeling efforts focused more specifically on mechanisms that could explain the observed statistical trends and on the predicted effects of possible management strategies.

We begin with the operating hypothesis that declines in base flow that have been observed in the Scott River are caused primarily by climate trends, as expected based on the large body of climate literature cited above and on the results of Drake *et al.* (2000), the only published study we could find that has addressed this problem. According to this hypothesis, trends in base flow observed in the Scott River should be consistent with those observed in other streams in the lower Klamath Basin, across which climate is relatively uniform. Further, we expect to observe differences in base-flow trends among these streams because of variation in elevation and latitude, which directly influence SWE. Secondary differences in streamflow trends among streams in the basin can then be attributed to local, watershed-scale factors such as land and water use. Although applied here to a specific basin, our methodology has applicability to any river system in which there are at least a few gaged streams unregulated by storage reservoirs. We use permutation tests for our statistical hypothesis tests, but this is not a methodological study intended to compare the results and applicability of these types of tests to those of other types of statistical tests. However, because permutation tests are not widely applied in water resources research, we provide sufficient detail in statistical methods so that they can be adopted by researchers in other basins.

The objectives of this paper are to (1) quantify basin-scale trends in streamflow and SWE in the lower Klamath Basin, (2) analyze the dependence of base flow and SWE trends on elevation and latitude, (3) compare relative change in base flow among different streams in the basin using a paired-basin approach, and (4) use paired-basin correlation analysis to estimate the component of decline in Scott River base-flow that is attributable to regional-scale climatic factors. The difference between this component and the total decline in base flow is attributable to local-scale factors, which we discuss. We also compare our results with those of Drake *et al.* (2000) and discuss implications for fisheries management.

### STUDY AREA

We define the lower Klamath Basin as the drainage of Klamath River downstream of the

Oregon-California state line (Figure 1). This coincides approximately with the location of Irongate Dam, which blocks upstream migration of anadromous fish, as well as the point at which the river exits the low-relief, volcanic geology of the Cascade Mountains and enters the high-relief, geologically complex Klamath Mountain and Coast Range provinces. This point is also roughly at the transition between the ocean-influenced climate to the west and the arid, intermountain climate to the east.

Elevations in the study area range from sea level to 2,500 m. Annual precipitation ranges from 50 cm in the eastern valleys to over 200 cm at higher elevations. Nearly all precipitation falls from October through April. Precipitation occurs almost exclusively as rain at elevations below 500 m and almost exclusively as snow above 2,000 m. Snowpack generally accumulates throughout the mid-winter to late-winter at elevations exceeding 1,500 m. High relief and impermeable bedrock geology contribute to rapid runoff of both rainfall and snowmelt from upland areas, and ground-water storage is generally limited to relatively small alluvial aquifers immediately adjacent to major streams. Correspondingly, stream hydrographs in the study area are of the rain/snow type (Poff, 1996), characterized by rapidly increasing discharge at the onset of the rainy season, a broad peak lasting most of the winter and spring, and recession beginning in June, once maximum snowmelt has occurred

(Figure 2). Base flow, which is generally 1.5 orders of magnitude lower than peak flow, occurs during late summer and early fall. Variability in this pattern across catchments is driven by the relative contribution of rain and snowmelt to runoff, which, in turn, is determined primarily by elevation and latitude, and to a lesser degree by distance from the coast and local topographic features.

To focus on changes in streamflow related to climate change, we limited our analysis to streams that have a continuous record of discharge dating back at least 40 years from the present and are unaffected by storage reservoirs. Only five streams in the lower Klamath Basin met these criteria: the Scott, Salmon, Trinity (upstream of reservoirs), and South Fork Trinity rivers and Indian Creek (Figure 1, Table 1).

All five of the study watersheds are sparsely populated, although population is increasing in some locales, particularly in the South Fork Trinity watershed. Uplands are mountainous areas managed by the U.S. Forest Service. Substantial timber harvest has occurred in all five watersheds, although it has been more limited in the Salmon and Trinity watersheds because of large amounts of federally designated wilderness. Rugged terrain and a preponderance of federal land limit most human activities to narrow river corridors in the Indian, Salmon, and Trinity watersheds. Additionally,

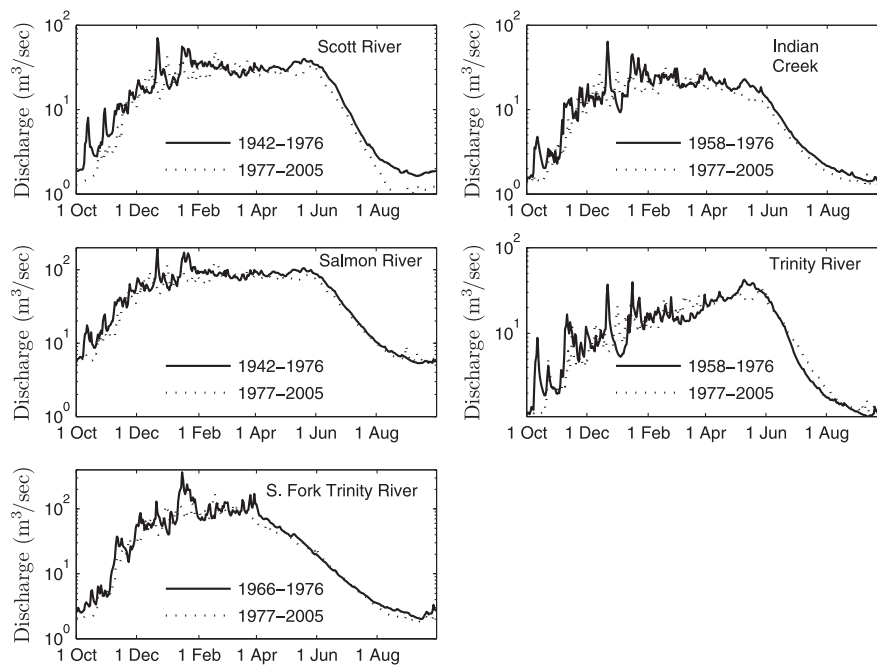


FIGURE 2. Mean Historic-Period and Modern-Period Hydrographs for the Five Study Streams. All streams display a rain-snow hydrologic regime with base flow period during late summer. Discharge is shown on a logarithmic scale to facilitate visual comparison of modern and historic periods at low discharge values. However, statistical comparison of annual and late-summer discharge between periods was performed on untransformed data.



TABLE 1. Study Basin Descriptions and Flow Statistics.

	Scott River	Indian Creek	Salmon River	South Fork Trinity River	Trinity River
USGS stream gage	11519500	11521500	1522500	11528700	11523200
Drainage area (km <sup>2</sup> )	1,691	311	1,945	1,979	386
Mean basin elevation (m)	1,688	1,220	1,386	1,378	1,734
Latitude of basin centroid (°N)	41.479	41.904	41.293	40.468	41.228
Earliest year analyzed	1942	1958	1942	1966	1958
Mean annual historic-period discharge (Mm <sup>3</sup> )	605.7	403.1	1,744	1,420	385
Mean annual modern-period discharge (Mm <sup>3</sup> )	514	345.3	1,517	1,175	361.1
<i>p</i> -value: historic and modern annual discharges equal	0.127	0.116	0.113	0.163	0.294
Mean late summer historic-period discharge (Mm <sup>3</sup> )	10.96	9.193	37.04	14.77	7.273
Mean late summer modern-period discharge (Mm <sup>3</sup> )	6.541	8.274	37.47	12.08	8.024
<i>p</i> -value: historic and modern late summer discharges equal	<0.001	0.055	0.629	0.049	0.799

Notes: Flow data are from the USGS National Water Information System, <http://www.waterdata.usgs.gov/nwis>, accessed December 2006. Historic period ends in 1976; modern period is 1977 through 2005; *p*-values are reported for the one-sided alternative hypothesis that modern-period discharge is less than historic-period discharge.

topography prevents substantial agricultural development. The South Fork Trinity watershed supports some agriculture, primarily fruit and vegetable farms, vineyards, and cattle grazing operations. Because agricultural development in the South Fork Trinity watershed is relatively small in scale, few if any data on irrigation withdrawals are available.

Only the Scott watershed contains large areas of private, non-mountainous land that support large-scale agriculture; about 120 km<sup>2</sup> of pasture, grain, and alfalfa are irrigated in the Scott watershed. A typical western-U.S. system of water rights based on the doctrine of prior appropriation governs withdrawal and delivery of surface water for irrigation in the Scott Valley (California Superior Court, 1950, 1958, 1980). Under this type of water rights system, surface water diverted from streams is delivered to water users in order of decreed water right priority date (date on which the claim to put the water to beneficial use was first made; these are typically dates in the mid to late 19th Century in California). Early in the irrigation season, when streamflows are high, all users receive their full allocation of water. As streamflow declines throughout the irrigation season, those users with junior (i.e., more recent) priority dates must cease diversion to leave the available water to users with more senior rights. By the end of a typical irrigation season, only users with the most senior rights continue to divert surface water. The California Department of Water Resources (CDWR) collects some data on irrigation use in the Scott Valley. However, CDWR does not provide watermaster service to account for distribution of decreed surface rights in all areas of the Scott watershed, and withdrawal and distribution of ground water is unregulated.

## METHODS

Streamflow and SWE data were available in our study area from the mid-1940s to the present. Given our working hypothesis regarding climate effects and the natural division of this time period into two distinct phases of the PDO (cool from mid-1940s to 1976, warm from 1977 on), we used a two-step comparison approach to analysis of temporal trends (Helsel and Hirsch, 1992). Because streamflow data for the Scott River were first collected in water year 1942, we defined the “historic” period as 1942-1976 and the “modern” period as 1977-2005. We then analyzed differences in SWE and streamflow between these two time periods. We used permutation tests (Ramsey and Schafer, 2002; Good, 2005; see Appendix A) to perform all statistical hypothesis tests. We performed these tests at the  $\alpha = 0.05$  significance level.

All of the hypothesis tests involved comparing values of a particular SWE or discharge variable between the historic and modern periods. Although use of permutation tests does not require the data to meet any distributional assumptions, it does require independence of observations (Good, 2005). Thus, we first corrected the data for dependence caused by first-order serial autocorrelation using the correction as

$$x_t = u_t - ru_{t-1}, \quad (1)$$

where  $x_t$  is the corrected value of the variable for year  $t$ ,  $u_t$  is the uncorrected value for year  $t$ , and  $r$  is first-order serial autocorrelation coefficient (i.e., the Pearson correlation coefficient between  $u_t$  and  $u_{t-1}$  (Neter *et al.*, 1989; Ramsey and Schafer, 2002). We then calculated the test statistic as

$$T = \frac{\bar{x}_1 - \bar{x}_2}{SE}, \quad (2)$$

where  $\bar{x}_1$  is the mean of the corrected daily discharge values over Group 1,  $\bar{x}_2$  is the mean over Group 2, and

$$SE = s\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}, \quad (3)$$

where  $s$  is the pooled standard deviation,  $n_1$  is the number of years in Group 1 and  $n_2$  is the number of years in Group 2. Groups 1 and 2 refer to the complementary subsets into which the data are divided according to a given permutation (see Appendix A). To calculate the value of the test statistic obtained from the data as they occurred in the observed permutation, Group 1 is taken to be the collection of data observed over the modern period of years, and Group 2 is that observed over the historic period, that is,

$$T_{\text{observed}} = \frac{\bar{x}_{\text{modern}} - \bar{x}_{\text{historic}}}{SE} \quad (4)$$

Although (Equation 4) is the test statistic of the standard two-sample  $t$ -test, we use it instead in permutation tests and as the response variable in regressions. Thus, we refer to it as a generic “ $T$ ”-statistic.

As we wanted to focus our analysis on the period of days during the base flow period over which declines in discharge in the Scott River have been most apparent, we defined the “late summer” period of base flow based on analysis of the Scott River data at the daily scale instead of defining this period based on visual examination of hydrographs or on a convenient calendar designation (e.g., August and September). We first  $\log_{10}$ -transformed daily discharge for each individual day between June 1 and November 30. The transformation was performed not to meet the assumptions of the hypothesis test but rather to prevent rare but extreme daily flow events from exerting excessive influence over group mean. We then compared the mean of the transformed discharge between historic and modern periods of years with a permutation test on the  $T$ -statistic (see Appendix A). We performed these tests with a two-sided alternative. This analysis showed that the mean of  $\log_{10}$ -transformed daily discharge (equivalently, the geometric mean) differed significantly between the historic and modern periods on every day of the period August 2 through October 5. We thus defined “late summer” to be this period of consecutive days.

## Streamflow and SWE Trends

We tested for differences in total late-summer discharge between historic and modern periods at all five stream gages. For streams on which gaging began after 1942, we defined the historic period to begin with the first year in the period of record (Table 1). Because of the smoothing inherent in averaging daily discharge over the 65-day late-summer period, we did not transform the raw discharge data. These tests were performed with the one-sided alternative that late-summer discharge during the modern period was less than that during the historic period, in accordance with what would be expected based on climate change. We also performed this analysis on annual water-year discharge at each stream gage and on April 1 SWE at all 16 snow courses in the study area for which at least 40 years of data were available (Figure 1, Table 2). For these tests, we also used a one-sided alternative, for consistency with the late-summer for analysis.

## Dependence of Base Flow and SWE Trends on Elevation and Latitude

To quantify dependence of change in SWE and streamflow on elevation, we performed permutation regression analysis (see Appendix A) of the observed  $T$ -statistic (Equation 4) as a function of elevation. In this case,  $T_{\text{observed}}$  serves as a dimensionless measure of change in SWE or streamflow between historic and modern periods and thus allows direct comparison of the regression line for streamflow to that for SWE. To incorporate the effect of latitude, we used Mote’s (2006) estimate that winter isotherms along the Pacific Coast of North America increase southward at a rate of 137 m in elevation per degree of latitude. We referenced latitude to that of Indian Creek, the furthest north of the study watersheds, and defined latitude-adjusted elevation of a given snow course or study watershed to be

$$E_{\text{adjusted}} = E - 137(L_{\text{Indian}} - L), \quad (5)$$

where  $E$  is the actual elevation of the snow course or watershed (mean over the watershed),  $E_{\text{adjusted}}$  is the adjusted elevation,  $L_{\text{Indian}}$  is the watershed-centroid latitude of the Indian Creek watershed, and  $L$  is the latitude of the snow course or watershed centroid. Centroids and mean elevations of the drainage basins were computed in a Geographic Information System from Digital Elevation Models. For the SWE analysis, we regressed dimensionless change in April 1 SWE

TABLE 2. Snow Course Descriptions and April 1 Snow Water Equivalent (SWE) Statistics.

Course Number	Elevation (m)	Latitude (°N)	Earliest Year of Record	Mean Historic-Period April 1 SWE (cm)	Mean Modern-Period April 1 SWE (cm)	<i>p</i> -Value: Historic and Modern April 1 SWE Equal
17	1,554	41.077	1946	40.3	30.2	0.021
14	1,646	41.150	1947	84.7	90.2	0.666
285	1,676	41.397	1951	104.2	68.2	0.001
15	1,722	41.197	1947	66.2	52.0	0.022
298	1,737	41.233	1956	49.4	44.5	0.224
3	1,783	41.382	1942	37.0	30.0	0.059
4	1,798	41.400	1951	95.0	52.1	<0.001
16	1,838	41.093	1942	55.5	51.5	0.261
13	1,875	41.200	1949	91.1	91.1	0.482
311	1,890	41.225	1949	71.1	72.5	0.568
12	1,951	41.008	1947	127.8	126.7	0.434
11	1,981	40.967	1947	95.2	101.2	0.704
5	2,012	41.217	1946	80.8	81.4	0.524
1	2,042	41.367	1942	95.3	88.4	0.218
10	2,042	41.023	1946	111.7	112.7	0.542
9	2,195	41.318	1946	84.2	86.9	0.634

Notes: Table is sorted by elevation for ease of interpretation. Data are from the California Department of Water Resources snow course database, <http://www.cdwr.water.ca.gov/misc/SnowCourses.html>, accessed May 2007. Historic period is earliest year of record through 1976; modern period is 1977 through 2005; *p*-values are reported for the one-sided alternative hypothesis that modern-period SWE is less than historic-period SWE.

against latitude-adjusted snow course elevation. We performed an analogous regression for change in late-summer discharge against latitude-adjusted mean watershed elevation for the five study streams.

#### Comparison of Relative Base-Flow Decline Among Study Streams

To compare base-flow trends among the five study streams, we used a before after control-impact-pairs analysis (Stewart-Oaten *et al.*, 1986). For each of the 10 ( ${}^5C_2 = \frac{5!}{2!3!} = 10$ ) unique pairwise combinations (*a,b*) of the five study streams and for each year in the intersection of the periods-of-record of the two streams, we computed the ratio  $\frac{Q_a}{Q_b}$ , where  $Q_a$  is the total late-summer discharge in stream *a* for the given year and  $Q_b$  is the total late-summer discharge in stream *b*. To prevent small values in the denominator from producing extremely large values of the ratio, we chose stream *b* to be the stream in each pair with the larger mean late-summer discharge during the modern period. We then compared the mean of these annual ratios  $\frac{Q_a}{Q_b}$  between modern and historic periods using the permutation method. We used two-sided alternatives because the purpose of the paired-basin tests was to assess differences in streamflow response among the study streams, and if factors other than climate change affected this response, we would not know *a priori* which stream

in a given pair should have the lower relative streamflow during the modern period.

#### Component of Scott River Base-Flow Decline Attributable to Climate

We estimated the component of base-flow decrease in the Scott River due to climate by comparing daily flow in the Scott River with that of a reference stream. Based on geographic proximity and lack of substantial changes in anthropogenic effects on water resources over the past half-century, either the Salmon or Trinity could serve as the reference stream for this estimate. Although the Trinity watershed is closer in elevation to that of the Scott, we chose the Salmon as the reference watershed because it is much closer in size to that of the Scott (Table 1) and because the hydrograph of the Salmon River is more similar to that of the Scott than to any of the other study streams (Figures 2 and 3). Furthermore, because the latitude-adjusted elevation of the Salmon River watershed is lower than that of the Scott River, comparison with the Salmon River provides an overestimate of the effect of climate and hence an underestimate of the effect of local-scale factors on Scott River base-flow. We used the line of organic correlation (Helsel and Hirsch, 1992) to determine the linear relationship between daily Scott River discharge and daily Salmon River discharge. Because the relationship was used for prediction and not for hypothesis

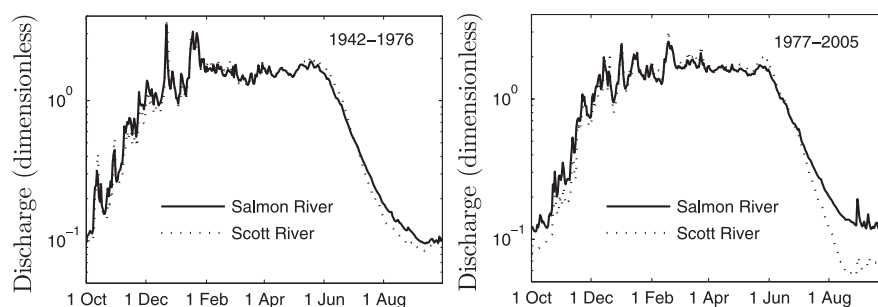


FIGURE 3. Period-of-Record Mean Dimensionless Hydrographs for the Salmon and Scott Rivers, Historic and Modern Periods. Dimensionless discharge is daily discharge divided by mean period-of-record discharge. Note that the hydrographs were nearly identical during the historic period but that during the modern period, Scott River discharge was much lower than Salmon River discharge from early July through late October.

testing, we did not correct daily values for serial autocorrelation. In this analysis we used all daily flow values from July 1 through October 22 during each of the calendar years in the historic period. This period of days was chosen because it was the time period over which the relationship between Scott and Salmon river hydrographs differed most between the historic and modern periods (Figure 3). We applied the organic linear relationship to modern-period Salmon River daily discharge values to estimate what discharge would have been in the Scott River during the modern period if response of flows in the Scott River to regional climate change had been the same as that of flows in the Salmon River. Because the line had a negative intercept, predicted discharge on a small percentage of days was slightly negative, and discharge on these days was set to zero. The difference between this estimated modern-period discharge and the observed modern-period discharge was our estimate of the component of Scott River summer discharge decrease due factors other than climate. For comparison, we also determined the line of organic correlation relating Scott and Salmon river discharge over the modern period.

## RESULTS

### *Streamflow and SWE Trends*

Mean daily hydrographs showed relatively small differences between historic and modern periods, with the exception of substantially lower modern-period discharge during late summer and early fall in the Scott River (Figure 2). Mean annual discharge in all five study streams was lower during the modern period, but none of the differences were significant (Table 1). The Scott River showed by far the greatest

decrease in late summer discharge between the two time periods (40.3% decrease,  $p < 0.001$ ), followed by the South Fork Trinity (18.2% decrease,  $p = 0.049$ ) and Indian Creek (10.0% decrease,  $p = 0.055$ ). Late-summer discharge increased slightly in the Salmon (1.2% increase,  $p = 0.629$ ) and Trinity (10.3% increase,  $p = 0.799$ ) rivers between historic and modern periods.

Mean April 1 SWE was lower in the modern period at all seven snow courses below 1,800 m, and these differences were significant at four of these courses and marginally significant at a fifth (Table 2). Mean April 1 SWE was higher in the modern period at five of the nine courses with elevations above 1,800 m, but none of these differences were significant.

### *Dependence of Base Flow and SWE Trends on Elevation and Latitude*

Change in April 1 SWE between historic and modern periods showed a significant, positive dependence on latitude-adjusted snow-course elevation (Figure 4). There was no significant dependence of change in late summer streamflow on latitude-adjusted drainage-basin elevation among the five study watersheds, but this dependence was significant when the Scott River was removed from the analysis (Figure 4). The slopes of the SWE and the significant (i.e., Scott River not included) flow regression lines were similar (0.00427/m for change in SWE, and 0.00539/m for change in late summer flow). Under the null hypothesis that the SWE and significant flow regressions are independent of each other, permutation analysis showed that the probability of obtaining a linear relationship between change in SWE and elevation as significant as that observed and a relative difference between the slope of the two lines this small is  $p = 0.00203$  (see Appendix A). This provides strong evidence that the similarity in slopes of these two

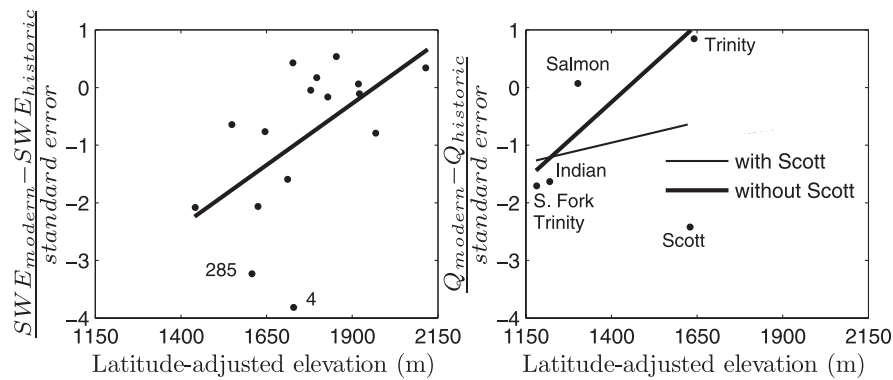


FIGURE 4. Change in April 1 Snow Water Equivalent (SWE; left) and Late-Summer Flow (right) Between the Historic and Modern Periods as a Function of Latitude-Adjusted Elevation. Decrease in both parameters is measured by the dimensionless *T*-statistic (Equation 4). Snow course Numbers 4 and 285 are identified in the left panel. Change in April 1 SWE showed a significant dependence on elevation ( $y = 0.00427x - 8.39, p = 0.028$ ). Change in late-summer flow showed no significant dependence on elevation with all data included ( $y = 0.00141x - 2.39, p = 0.700$ ) but showed significant dependence on elevation when the Scott River was removed from the analysis ( $y = 0.00539x - 7.80, p = 0.042$ ).

regression lines cannot be caused by chance alone, that is, that the dependence of change in streamflow on elevation is linked with that of change in SWE, as expected based on the underlying hydrologic processes.

*Comparison of Relative Base-Flow Decline Among Study Streams*

Late-summer flow in the Scott River declined between historic and modern periods relative to all four of the other study streams, and all of the differences in discharge ratio involving the Scott River were significant (Table 3). Decline in base flow in the Scott River was greatest relative to the Trinity River, followed by that relative to the Salmon River, Indian Creek, and the South Fork Trinity River, respectively. Late-summer flow in the South Fork Trinity declined

relative to all study streams except the Scott, and these differences were all significant. Late-summer flow in Indian Creek declined relative to the Trinity and Salmon rivers, but only the decline relative to the Trinity was significant. As mentioned above, late-summer discharge in the Salmon and Trinity rivers increased slightly between the historic and modern periods, and the paired-basin test showed that the increase observed in the Trinity River was significantly greater relative to that in the Salmon River.

*Component of Scott River Base-Flow Decline Attributable to Climate*

Scott River daily discharge from July 1 to October 22 was much lower relative to Salmon River discharge during the modern period than during the

TABLE 3. Paired-Basin Tests of the Null Hypothesis That the Ratio of Late Summer (August 2 through October 5) Discharge Is Equal Between Modern and Historic Periods.

Pair	Mean Ratio of Late-Summer Discharge (historic)	Mean Ratio of Late-Summer Discharge (modern)	Stream With Lower Relative Late Summer Discharge in Modern Period	<i>p</i> -Value: Historic and Modern Ratios Equal
Scott/Trinity	1.65	0.602	Scott	<0.001
Scott/Salmon	0.136	0.063	Scott	0.001
Scott/Indian	0.961	0.589	Scott	0.003
Scott/South Fork Trinity	0.599	0.400	Scott	0.010
Trinity/South Fork Trinity	0.397	0.590	South Fork Trinity	0.007
South Fork Trinity/Salmon	0.334	0.272	South Fork Trinity	0.035
Indian/ South Fork Trinity	0.590	0.747	South Fork Trinity	0.018
Trinity/Indian	0.803	0.973	Indian	0.001
Indian/Salmon	0.237	0.223	Indian	0.174
Trinity/Salmon	0.172	0.193	Salmon	0.045

Notes: Mean ratios of late-summer discharge are shown here; means of late summer discharge for each basin are given in Table 1; *p*-values are reported for the two-sided alternative hypothesis.

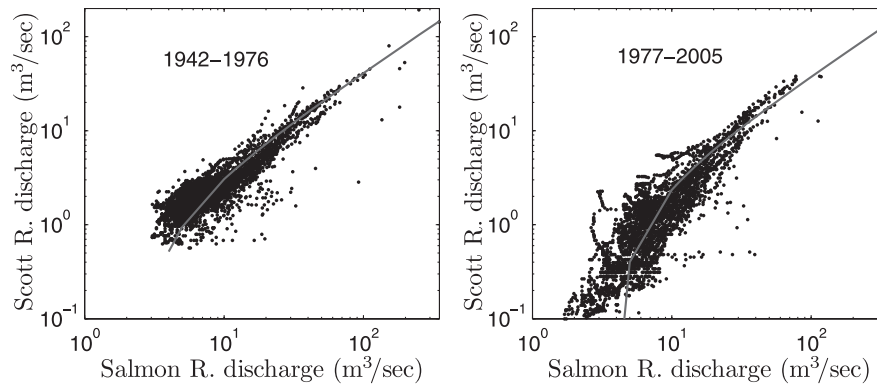


Figure 5. Scatterplots and Lines of Organic Correlation Relating Scott River Daily Discharge ( $y$ ) and Salmon River Daily Discharge ( $x$ ) for July 1 Through October 22, Historic and Modern Periods. Lines of organic correlation are  $y = 0.422x - 1.17$  for the historic period and  $y = 0.398x - 1.62$  for the modern period. Discharge is plotted on logarithmic scales to show detail at low discharge values; however, the lines of organic correlation and all analyses were performed on the untransformed data. Note that daily discharge in the Scott River never fell below  $0.566 \text{ m}^3/\text{s}$  during the historic period but fell below this value on 28.6% of all days between July 1 and October 22 during the modern period.

historic period (Figures 3 and 5). Furthermore, whereas the magnitudes of daily discharge in the Salmon River showed little difference between the historic and modern periods, daily discharge in the Scott River showed a large decrease in mean (from  $3.23$  to  $2.15 \text{ m}^3/\text{s}$ ). During the historic period, discharge in the Scott River was less than  $1 \text{ m}^3/\text{s}$  on 4.3% of all days from July 1 through October 22, whereas during the modern period, flows were less than  $1 \text{ m}^3/\text{s}$  on 46.2% of these days. Applying the historic-period organic linear relationship to modern-period Salmon River daily discharge produced an estimate of Scott River daily flow under the influence of regional-scale climate trends alone (Figure 6). The estimated mean hydrograph differed very little from the observed his-

toric-period hydrograph from July 1 through early August, but estimated modern-period discharge was lower over most of August, September, and October. Observed July 1 through October 22 discharge in the Scott River averaged  $31.8 \text{ Mm}^3/\text{year}$  over the historic period and  $21.3 \text{ Mm}^3/\text{year}$  over the modern period. Our estimate of July 1 through October 22 discharge under the influence of regional-scale climate trends alone averaged  $27.8 \text{ Mm}^3/\text{year}$  over the modern period. Thus, the component of decrease in Scott River discharge caused by factors other than regional-scale climate is estimated at  $6.5 \text{ Mm}^3/\text{year}$ , 61% of the observed decrease.

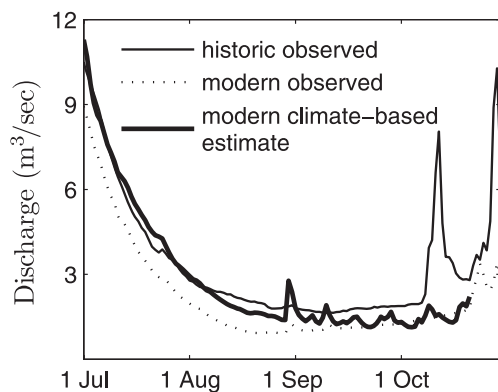


FIGURE 6. Mean July Through October Hydrographs for the Scott River, Showing Observed Historic-Period and Modern-Period Discharge and Estimated Modern-Period Discharge Based on Correlation With the Salmon River (climate-based estimate). Estimated modern-period flows show little deviation from historic-period flows during July and early August but are lower than historic-period flows from mid-August through late October.

## DISCUSSION

### *Streamflow and SWE Trends and Dependence on Elevation and Latitude*

Base flow and April 1 SWE in the lower Klamath Basin follow general trends toward lower April 1 SWE and lower base flows observed throughout the Pacific Northwest over the past 60 years (Hamlet *et al.*, 2005; Mote *et al.*, 2005; Regonda *et al.*, 2005; Stewart *et al.*, 2005; Mote, 2006). Models indicate that global warming may increase precipitation over the Pacific Northwest (Leung and Wigmosta, 1999; McCabe and Wolock, 1999; Salathé, 2006) so that at the highest elevations, April 1 SWE may actually increase because of increased winter-time precipitation, despite the trend toward higher temperatures. In the lower Klamath Basin, SWE has decreased significantly at lower-elevation snow courses but has

increased slightly at several higher-elevation courses (Table 2). Thus, our results are consistent with regional-scale analyses and reflect trends in both temperature and precipitation. The patterns of base-flow change between the historic and modern periods in the South Fork Trinity, Indian, Salmon and Trinity watersheds are exactly as predicted by SWE-elevation-latitude relationships. The within-basin analysis (Table 1), the paired-basin analysis (Table 3), and the regression analysis (Figure 4) all showed that when compared with that of the historic period, late-summer discharge in the modern period in each stream, both independently and relative to the other streams, followed the order predicted by latitude-corrected elevation and by the SWE patterns. Base flow decreased in the two watersheds with the lowest latitude-adjusted elevation (South Fork Trinity River and Indian Creek), and the decrease was greatest in the South Fork Trinity, which has the lowest latitude-adjusted elevation of any of the study streams. Base flow increased in the Trinity and Salmon rivers, and the increase was greatest in the Trinity River, which has the highest latitude-adjusted elevation of any of the study streams. The increases in late-summer flow observed in the Salmon and Trinity watersheds have occurred despite moderate decreases in total annual flow in these streams, suggesting effects from finer-scale patterns in temperature and precipitation that we did not analyze.

#### *Base-Flow Decline in the Scott River Relative to the Other Streams*

Base-flow trends in the Scott River clearly do not follow those of the other four streams. The latitude-corrected elevation of the Scott River watershed is only 31.5 m less than that of the Trinity River watershed (Figure 4), but base flows in the Scott River showed by far a greater decrease between historic and modern periods than those in any of the other four watersheds. The paired-basin analyses (Table 3), regression relationships (Figure 4), and Salmon River comparison (Figures 3 and 5) provide strong evidence that base flow in the Scott River has responded to regional-scale climate in a much different way than the other four streams and/or that factors other than climate have contributed to changes observed in Scott River base-flow since the late 1970s.

Certainly, some of the trends in Scott River base-flow are caused by the same climatic factors that have affected the other study streams. Decreases in mean annual discharge between historic and modern periods were 6.2% in the Trinity River, 13.0% in the Salmon River, 14.3% in Indian Creek, 15.1% in the

Scott River, and 17.0% in the South Fork Trinity River (Table 1). The *p*-values for the significance of these declines were remarkably similar for all but the Trinity River (Table 1). Furthermore, the paired-basin analysis showed no significant trends in total annual discharge among the study streams. Differences in response of the Scott River relative to the other streams appear to be limited only to base flow trends because at the annual scale, response of the Scott River to climatic differences between the two time periods was indistinguishable from those of the other study streams.

#### *Factors Affecting Scott River Base-Flow*

Geographic factors may be partially responsible for the large apparent difference in base-flow response between the Scott River and the other study streams. Although not the furthest east of the study basins, the Scott watershed does lie partially within a precipitation shadow formed by the large region of high-elevation terrain to the west of the watershed, contributing to a drier, more continental climate than that of the other four study watersheds. The Scott watershed has by far the smallest basin yield (discharge per unit watershed area, Table 1), an indication of both lower precipitation and higher evapotranspiration, the latter of which includes a large amount of irrigation not present in the other watersheds. The elevation dependence exhibited by base-flow change in the other streams predicts an increase in base flow in the Scott River between historic and modern periods (Figure 4). However, the comparison with the Salmon River predicts a decrease, albeit one only about 40% as large as that observed. The two snow courses with the largest decreases in April 1 SWE were Courses 4 and 285, located on the western side of the Scott watershed (Table 2, Figures 1 and 4). Although these are two of the lower-elevation snow courses in the study area, their decline is disproportionate with their elevation (Figure 4). The large decreases in April 1 SWE at these courses could be caused by local geography (e.g., the precipitation shadow), but a snow survey technician who has conducted measurements at these courses noted that forest vegetation has encroached on the courses, reducing accumulation of snowpack on the courses themselves (Power, 2001; J. Power, personal communication). Furthermore, none of the other courses in the Scott basin (Numbers 5, 298, and 311) show patterns inconsistent with the rest of the courses, and SWE has increased slightly at Courses 5 and 311 (Table 2).

Additional data provide evidence that part of the observed decrease in Scott River base-flow since the

1970s is likely caused by an increase in withdrawal of water for irrigation in the Scott Valley. Although data on water use in the Scott Valley are sparse and difficult to obtain, those that we were able to acquire show that irrigation withdrawals in the Scott Valley increased by 115% between 1953 and the period over which modern data are available (1988-2001; Figure 7). We were unable to locate data from the 1960s and 1970s to determine when the majority of the increase occurred, but across the western U.S. as a whole, the largest increase in irrigation withdrawal between 1950 and 2000 occurred in the 1970s (Hutson *et al.*, 2004). This increase in irrigation withdrawal accompanied an 89% increase in irrigated land area (Figure 7). In 1953, 77 cm of irrigation was applied over the growing season, and Mack (1958) reported that application rates in the 1940s averaged about 76 cm per year. Average application rate over the period 1988-2001 was 88 cm per year, a 15% increase over historic values. The limited data available show no change in crop types since the 1950s; irrigation has been applied primarily to alfalfa, grain, and pasture through both the historic and modern periods. Climatic factors could have influenced the increase in irrigation application rate; a warmer climate could result in a longer growing season and in higher evapotranspiration rates. However, the 15% increase in application rate is small compared the observed increases of 89% in irrigated land area and

115% in irrigation withdrawal between the historic and modern periods.

A second important trend in irrigation practices in the Scott Valley is that most irrigation in the Scott Valley is currently applied with sprinklers, and conveyance occurs in a pipe network. Recharge of ground water resulting from former flood irrigation practices has been largely eliminated, as has been observed in other locations around the western U.S. (Johnson *et al.*, 1999; Venn *et al.*, 2004). Mack (1958) estimated that during water year 1953, recharge to the alluvial aquifers in the Scott Valley was provided by precipitation (about 25 Mm<sup>3</sup>), tributary inflow (unspecified amount), and irrigation seepage (about 21 Mm<sup>3</sup>). Thus, in 1953, of the 48 Mm<sup>3</sup> withdrawn for irrigation, only about 27 Mm<sup>3</sup> (56%) was used consumptively. This efficiency is typical of flood irrigation systems with ditch conveyance (Battikhi and Abu-Hammad, 1994; Venn *et al.*, 2004). Conversion from flood to sprinkler irrigation has been reported to increase efficiencies to about 70% (Venn *et al.*, 2004), implying that while withdrawal of irrigation water in the Scott Valley has increased 115% since the 1950s, consumptive use may have increased by as much as 167%. Venn *et al.* (2004) reported that after conversion from flood to sprinkler irrigation in an alluvial valley in Wyoming, streamflow decreased significantly in the late summer and early fall because of decreased recharge of ground water, and this same mechanism could be acting in the Scott Valley as well.

A third important change is that ground water replaced surface water as the dominant source of irrigation water between 1990 and 2000 (Figure 7), reflecting trends observed across the western U.S. (Hutson *et al.*, 2004). Even if recharge from precipitation and tributary inflow have remained unchanged since the 1950s, change in irrigation conveyance and application methods and increased pumping of ground water in the Scott Valley could have resulted in decline of aquifer water levels. These alluvial aquifers discharge to the Scott River and its tributaries (Mack, 1958), and thus decline in aquifer levels could result in lowered base flows in the Scott River. In the upper Snake River basin of Idaho, where ground water-surface water interactions in an irrigation system have been extensively studied, conversion from flood to sprinkler irrigation and increase in pumping of ground water have resulted in significant declines discharge from the aquifer into the Snake River (Johnson *et al.*, 1999; Miller *et al.*, 2003b). Because of lag times inherent in ground water responses, withdrawal of ground water in the middle of the irrigation season can affect stream base-flow into the late summer and early fall. Furthermore, ground water provides a source of irrigation water late in the season

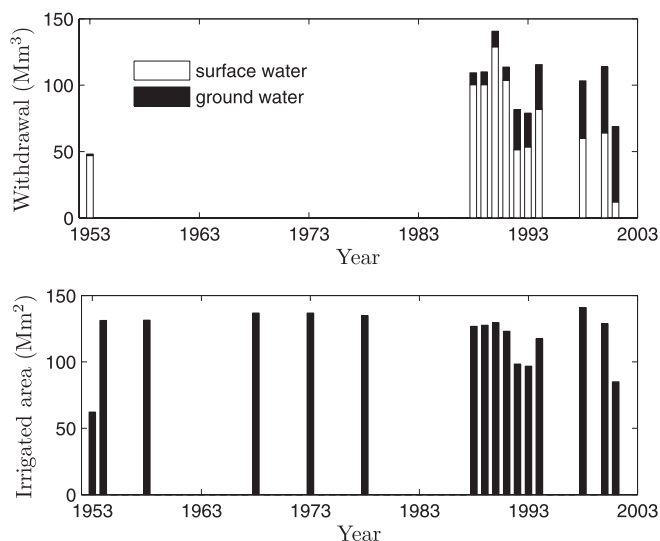


FIGURE 7. Annual Irrigation Withdrawal (top) and Irrigated Land Area (bottom) in the Scott River Basin From 1953 to 2001. Note that ground water made up less than 3% of total withdrawals in 1953 and more than 80% in 2001. Total annual withdrawal increased from 48 Mm<sup>3</sup> in 1953 to an average of 103 Mm<sup>3</sup> over the period 1988-2001, in close proportion to increase in irrigated area (62 Mm<sup>2</sup> in 1953, average of 117 Mm<sup>2</sup> over 1988-2001). Data for 1953 are from Mack (1958). All other data were provided by the California Department of Water Resources upon request.



when streamflow is low and availability of surface water is limited. Thus, transition from an irrigation system based primarily on diversion of surface water from streams to one with a large capacity to pump ground water allows more water to be used late in the irrigation season. Finally, because ground-water pumping in the Scott Valley is unregulated, actual withdrawal amounts could differ from those reported on an annual basis by CDWR, and there is a general lack of data that is sufficient in spatial and temporal extent to perform the mechanistic modeling of interactions between ground and surface water that would be necessary to quantify the effect that changes in irrigation practices have had on streamflow in the Scott River.

#### *Comparison With Drake et al. (2000)*

Our estimate that 39% of the decrease in Scott River base-flow is due to climatic factors is contrary to that of Drake *et al.* (2000), who concluded that 78% of the decrease is due to decline in April 1 SWE. The disparity in these conclusions is easily explained by analysis methods. First, Drake *et al.* (2000) analyzed hydrologic data from the Scott River watershed alone, whereas our study employed a comparative approach using other watersheds in the basin. Secondly, they did not use any variables related to water use, which clearly show substantial changes over the same time period during which base flows have decreased (Figure 7). Finally, Drake *et al.* (2000) based their conclusion on decrease in April 1 SWE at Snow Courses 4 and 285 and a single term representing this SWE decrease in a multiple regression equation explaining September discharge in the Scott River. Their regression equation was

$$\begin{aligned}
 Q = & (2.5 + 1.18 \times \text{annualprecip.} + 8.6 \\
 & \times \text{Augustprecip.} - 6.7 \times \text{Julyprecip.} + 0.48 \\
 & \times \text{Course 285 SWE} + 0.25 \\
 & \times \text{Course 5 SWE})^2, \quad (5)
 \end{aligned}$$

where  $Q$  is September discharge, annual and monthly precipitation are as recorded on the Scott Valley floor, and April 1 values were used for the SWE terms. Because SWE at Snow Courses 4 and 285 were highly correlated, Snow Course 285 was chosen to represent these courses in the regression equation. Snow Course 5 was used to represent SWE at Courses 5 and 298, two highly correlated courses at which April SWE exhibited little temporal trend. The regression analysis did not include SWE at the other

snow course in the Scott River watershed (Course 311) nor at courses near the Scott River drainage basin divide in adjacent watersheds (Courses 1 and 13; Figure 1). April 1 SWE at these courses showed no significant decrease between historic and modern periods (Table 2). The estimate that 78% of the decline in Scott River base-flow is due to climate was based on the  $r^2$ -value of 0.78 for the regression Equation (5).

Based on mean values for the explanatory variables in the regression equation, the annual precipitation term is six times greater in magnitude than the August precipitation term and over 10 times greater in magnitude than the July precipitation term. Thus, July and August precipitation contribute relatively little to September discharge. The annual precipitation term is about 1.5 times greater than the Snow Course 285 term and about three times greater than the Snow Course 5 term. Mean annual precipitation at the Ft. Jones weather station, located near the Scott River gage, was 55.9 cm during the historic period and 54.8 cm during the modern period. April 1 SWE at Course 5 averaged 80.8 cm during the historic period and 81.4 cm during the modern period. These two variables show almost no change between historic and modern periods, and the sum of their respective terms in the regression equation is over twice as large as the Snow Course 285 term. Therefore, the conclusion of Drake *et al.* (2000) is based on a single term that accounts for less than one-third of the total magnitude of the variable terms in the regression equation.

#### *Implications for Fisheries*

Based on our estimate of the component of Scott River base-flow decrease attributable to changes in water use, returning irrigation to historic-period patterns in the Scott River would, in theory, increase July 1-October 22 discharge by an average of  $0.65 \text{ m}^3/\text{s}$ . This estimate includes continued irrigation withdrawal at the pre-1970s rate of about  $50 \text{ Mm}^3$ , albeit with as much as  $21 \text{ Mm}^3$  of this returning to the aquifer and streams via canal seepage. It also accounts for decrease in streamflow caused by regional-scale climate trends. Under current conditions, streamflow in the Scott River can drop below  $0.283 \text{ m}^3/\text{s}$  in the late summer and early fall of dry years. At this discharge, some reaches of the river become a series of stagnant and disconnected pools that are inhospitable to many aquatic species. An additional  $0.65 \text{ m}^3/\text{s}$  could create a viable corridor for movement of aquatic species, decrease fluctuations in water temperature (particularly daily maxima), and maintain the functionality of cold

water seeps and tributary mouths upon which salmonids rely (Cederholm *et al.*, 1988; Sandercocock, 1991; Stanford and Ward, 1992). Bartholow (2005) observed a warming trend of 0.5°C/decade in Klamath River water temperatures over the same period of years we have analyzed, suggesting that provision of cold-water refugia for aquatic life will become even more critical as climate warming continues. Although it is not likely that irrigation sources, withdrawal amounts, and application methods in the Scott River watershed will revert back to those of the 1960s, our results at least provide evidence that observed declines in base flow have not been caused by climate trends alone and hence could be reversed to the benefit of salmon and other aquatic life through changes in water management. However, management of water resources in the Scott Valley to meet the needs of both agriculture and fish will require consistent and accurate watermaster service for the entire valley, quantification of ground-water withdrawals and their effects on surface water, and water-use data that are easily obtainable. A major research need in the Scott Valley relevant to water management and aquatic species conservation is a comprehensive study of interactions between ground water and surface water that includes mechanistic modeling of effects of ground-water withdrawal on streamflow throughout the valley.

## CONCLUSIONS

We statistically analyzed streamflow in five lower Klamath Basin streams that are unregulated by storage reservoirs as well as April 1 SWE at all 16 snow courses in the basin with long periods of record. We compared streamflow and April 1 SWE between historic (1942-1976) and modern (1977-2005) periods, which were defined based on two distinct phases of the PDO. The historic period was a cold phase, which has been associated with high snowpack and high streamflows throughout the Pacific Northwest, and the modern period was a warm phase, which has been associated with lower snowpacks and streamflows region-wide. April 1 SWE decreased significantly between historic and modern periods at low-elevation snow courses in the lower Klamath Basin. No significant trends were apparent at higher elevations. Correspondingly, base flow decreased significantly in the two study streams with the lowest latitude-adjusted elevation and increased slightly in two of the higher-elevation study streams. With the Scott River excluded from the analysis, the depen-

dence of base-flow change on adjusted elevation follows the same trend as that of SWE. Despite a latitude-adjusted elevation only 1.8% lower than the highest-elevation watershed in the study, the Scott River has experienced a much larger reduction in base flow than the other study streams. Geographic differences may account for some of the discrepancy in base flow trends between the Scott River and the other four watersheds. However, irrigation withdrawal in the Scott watershed has increased from about 48 Mm<sup>3</sup> per year to over 100 Mm<sup>3</sup> since the 1950s, and the amount of ground water withdrawn for irrigation has increased from about 1 Mm<sup>3</sup> per year to about 50 Mm<sup>3</sup>. We estimate that 39% of the observed 10 Mm<sup>3</sup> decline in July 1-October 22 discharge in the Scott River has been caused by regional-scale climatic factors and that the remaining 61% is attributable to local factors, which include increases in irrigation withdrawal and consumptive use. Even after accounting for climatic factors, returning water use to pre-1970s patterns of withdrawal sources and quantities, conveyance mechanisms, and application methods in the Scott River watershed could benefit salmon and other aquatic biota by increasing July 1-October 22 streamflow by an average of 0.65 m<sup>3</sup>/s.

If our study watersheds are representative of others in the lower Klamath Basin, climate-induced decreases in late-summer streamflow in low-elevation watersheds will, at best, complicate the recovery of anadromous salmonids and may, at worst, hinder their persistence. Sound water management and recovery efforts such as habitat and watershed restoration will be required to help offset the effects of climate warming on river ecology, particularly because both decreased base flows and increased water temperatures occur simultaneously during periods of warm climate. Because streams at lower elevations are more susceptible to decreases in base flow caused by decreases in April 1 SWE, local-scale human-induced changes associated with water and land use could have a greater affect on streamflow and water temperature in these streams than in higher-elevation streams experiencing the same continental-scale warming. The South Fork Trinity River is of particular concern. It harbors one of the few remaining stocks of wild spring Chinook salmon in the entire Klamath Basin, and the latitude and elevation of the drainage put it at particular risk of climate-induced changes that adversely affect Chinook salmon and other species. Furthermore, development and largely unquantified water use on the South Fork Trinity River and important fish bearing tributaries such as Hayfork Creek exacerbate the problem. We recommend additional gaging on streams that are susceptible to the effects of human use, such as Hayfork

Creek, and on “control” streams that drain wilderness areas, such as Wooley Creek in the Salmon River watershed and the North Fork Trinity River, to monitor future trends in water use and climate in the lower Klamath Basin.

## APPENDIX A: PERMUTATION TESTS

Standard statistical hypothesis tests are commonly used to analyze time-series data collected at precipitation and streamflow gages (e.g., Helsel and Hirsch, 1992; McCuen, 2003). Most of these tests, whether parametric or non-parametric, are based on the assumption that the data were obtained through random sampling of infinite populations. However, this assumption is generally not met by data sets collected at precipitation and stream gages. First, these types of data are not randomly selected. The locations of stream and precipitation gages are almost never randomly chosen, and the recording of data at regular intervals such as days, months, or years does not constitute random selection. Second, the data rarely constitute a sample but rather comprise the entire population. For example, if we analyze difference in annual discharge between two time periods and have discharge values for every year in both time periods, then we have the entire population at hand. There is no sampling, and hence no infinite population to which inference can be drawn. Permutation tests, often called randomization tests in experimental contexts, are appropriate statistical tests to use for analysis of these and other types of non-sampled data (Ramsey and Schafer, 2002). We refer the reader to the comprehensive texts by Edgington (1995) and Good (2005) for a full treatment of theory and methodology and here present only a brief treatment of the two permutation tests used in this paper.

The basic concept behind permutation tests is best illustrated by the example of testing for differences in mean between two groups. Consider the comparison of late-summer discharge in the Scott River between the two time periods. Once the time-series data are corrected for serial autocorrelation, the observations constitute independent, annual values for each of the 64 years between 1942 and 2005, inclusive, and hence satisfy the assumptions of permutation tests. We then measure the magnitude of difference in the mean for each of the two time periods 1942-1976 and 1997-2005, relative to variability, using the test statistic (Equation 4). This division of 64 years into the historic and modern period is only one of the  $\frac{64!}{35!29!} \approx 1.39 \times 10^{18}$  distinct ways in which this set of 64 annual values can be divided into two groups of size

35 and 29. Each of these distinct ways is called a permutation, and each has associated with it a particular value of the test statistic (Equation 2). The distribution of these test statistics is called the permutation distribution. The  $p$ -value of the permutation test is the probability that we could have selected a permutation at random for which the value of the test statistic was at least as extreme (using either one or two tails, as appropriate to the alternative hypothesis) as that of the observed grouping (i.e., division of the time period into 1942-1976 and 1977-2005 time periods).

In practice, when the number of permutations is on the order of  $10^4$  or less, one computes the test statistic for every possible permutation and obtains the exact  $p$ -value of the test. This procedure is inherently non-parametric and requires no assumptions about the distribution of the original data or the number of observations, even if one uses a test statistic such as (Equation 2) that can be used in the context of a parametric test. When the number of permutations is large, there are two choices for conducting the test. One is to randomly select a large number of permutations from among those possible and use this sample to represent the entire set of permutations (see Supplementary Material). The other is to use a standard parametric test statistic (such as the  $T$ -statistic) from an analogous sample-based hypothesis test. It has been shown that for the permutation versions of most of these basic tests, the permutation distribution approaches the sampling distribution of the test statistic asymptotically as the number of permutations becomes infinite, regardless of the distribution of the original data (Edgington, 1995; Good, 2005). In our example of  $1.39 \times 10^{18}$  permutations, the permutation distribution of (Equation 2) is in fact a  $t$ -distribution (Figure A1). Hence, we can calculate the  $p$ -value of the test by comparison of the test statistic with the standard  $t$ -distribution without having to generate any permutations. In this case, the  $p$ -value of the permutation test for difference in mean coincides with that of the two-sample  $t$ -test but the interpretation is different. In the permutation test, the  $p$ -value is the probability of having obtained a difference in *population* mean at least as extreme as that observed in a randomly selected division of the data into two *populations* of sizes 35 and 29. In the two-sample  $t$ -test, the  $p$ -value is the probability of having obtained a difference in *sample* mean at least extreme as that observed based on random selection of a *sample* of size 35 from one population and a *sample* of size 29 from a second, independent population, under the null hypothesis that the population means are the same. Thus, even though we might get the “right answer” in terms of the  $p$ -value with naïve use of a two-sample  $t$ -test, our inference would be

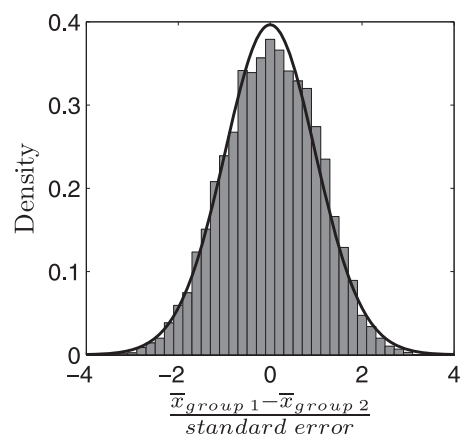


FIGURE A1. Permutation Distribution of the  $T$ -Statistic (Equation 1) for the Difference Between Historic-Period and Modern-Period Late Summer Discharge in the Scott River (Table 1). The histogram shows  $T$ -statistics from 10,000 randomly selected permutations (from among the  $1.39 \times 10^{18}$  possible), and the curve is the Student's  $t$ -distribution that would be used for the analogous  $t$ -test based on random samples from populations with unequal variances. The  $t$ -distribution has 39 degrees of freedom, as calculated using Satterthwaite's approximation (Ramsey and Schafer, 2002). In this case, the permutation and sampling distributions of the test statistic are identical.

inappropriate because our data do not constitute samples from infinite populations.

In the permutation version of linear regression, the permutations consist of all possible ways of pairing the observations of the dependent variable,  $y$ , with those of the independent variable,  $x$ . There are  $n!$  such permutations possible with a set of  $n$  ordered pairs. We perform the permutation test on the standard regression test statistic given by the ratio of regression mean square to error mean square. The observed statistic is that obtained from the data points as they were reported, and that value is compared against the values obtained from all of the other permutations. When the number of permutations is large, the permutation distribution of this test statistic is an  $F_{1,n-2}$ -distribution, identical to the sampling distribution of this test statistic. The SWE regressions used data pairs from 16 stations, so the number of permutations is  $16! \approx 2.09 \times 10^{13}$ , and use of the standard  $F$ -distribution is appropriate for computing the  $p$ -value of the permutation test. However, the number of permutations in the streamflow regressions was very small, so the standard  $F$ -distribution is not a good approximation to the permutation distribution. In the regression with the Scott River removed ( $n = 4$ ), the value of the test statistic obtained from the observed pairing of dependent and independent variables was 7.58, the largest among the 24 permutations. Thus, the  $p$ -value for this test is  $1/24 = 0.0417$  (Table A1). Regression analysis of these same four data points based on random

Table A1. Cumulative Distribution of the Test Statistic  $\frac{MSR}{MSE}$  for the Regression of Change in Streamflow *vs.* Adjusted Basin Elevation With Scott River Removed (Figure 4).

Test Statistic Value	Permutation Probability	Sampling Probability
7.5800	0.0417	0.1105
7.3102	0.0833	0.1139
2.6847	0.1250	0.2430
2.2445	0.1667	0.2728
2.1459	0.2083	0.2806
2.0749	0.2500	0.2864
1.9534	0.2917	0.2971
1.8136	0.3333	0.3104
1.2981	0.3750	0.3726
1.2497	0.4167	0.3799
1.0196	0.4583	0.4189
0.9162	0.5000	0.4395
0.9001	0.5417	0.4429
0.8407	0.5833	0.4560
0.5477	0.6250	0.5363
0.5388	0.6667	0.5393
0.4449	0.7083	0.5734
0.4289	0.7500	0.5798
0.3393	0.7917	0.6191
0.2621	0.8333	0.6596
0.2047	0.8750	0.6953
0.1894	0.9167	0.7059
0.0677	0.9583	0.8191
0.0622	1.0000	0.8263

Note: The test statistic values are those from each of the 24 possible permutations. The permutation probability is the probability of observing a test statistic at least as large from the permutation distribution, and the sampling probability is the probability of observing a test statistic at least as large from the sampling distribution, namely an  $F_{1,2}$ -distribution. The  $F$ -distribution underestimates probabilities for small values of the test statistic and overestimates them for the larger values.

sampling produces a  $p$ -value of 0.110 (Table A1). If the four study streams had been randomly selected from a large number of streams (on the order of 40 streams or more), then the probability is 0.110 of having observed a linear relationship at least this strong in a *sample* of four  $(x,y)$  pairs, under the null hypothesis that there was no linear relationship between  $x$  and  $y$  in the whole population. However, because these four streams were not selected at random (they were selected because they were streams that happened to have long periods of flow records), it is inappropriate to draw inferences to a large population from this set of four. Using permutation testing, the probability is 0.0417 of having observed a linear relationship this strong by chance assignment of the  $x$  and  $y$  values into  $(x,y)$  pairs, and we conclude that among this *population* of four study streams, there is a significant dependence of  $y$  on  $x$ .

To compare the slopes of the SWE and streamflow regressions (Figure 4), we first computed slopes  $m_1$  for each of the possible 24 permutations of the

streamflow data and slopes  $m_j$  for each permutation in a random sample of 1,000 permutations from among the 16! possible for the SWE data (see Supplementary Material). We then calculated the symmetric relative difference between the slopes given by

$$\frac{|m_i - m_j|}{0.5(|m_i| + |m_j|)} \quad (6)$$

for all possible combinations  $i, j$  as  $i$  ranged over the 24 streamflow permutations and  $j$  ranged over the 1,000 randomly selected SWE permutations. The observed relative difference was smaller than 92.61% of these differences. However, we are interested in differences in slopes not for all possible pairs of regression lines but only for those that are statistically significant to begin with. If the dependence of change in streamflow on adjusted elevation is independent of that of SWE on adjusted elevation, then the probability of randomly selecting a regression pair with a difference in slopes as small as the observed difference *and* randomly selecting a permutation of the SWE data showing as strong a linear relationship as that observed is the product of the two individual probabilities. The probability of the former event is  $1 - 0.9261 = 0.0739$ , and the probability of the latter is 0.0275. Thus, the desired probability is 0.00203. We conclude that it is extremely unlikely to have observed regression relationships this similar by chance alone if the dependence of change in streamflow on elevation is independent of that of change in SWE on elevation.

#### ACKNOWLEDGMENTS

This research was funded by the U.S. Fish and Wildlife Service. Diana Boyack of the Idaho State University Digital Mapping Laboratory created the study area map and computed drainage basin statistics. The California Cooperative Fish Research Unit, U.S. Geological Survey, provided administrative support to SWN. DeWayne Derryberry of the Idaho State University Department of Mathematics provided helpful references and comments on permutation tests. The suggestions of four anonymous reviewers greatly improved the manuscript.

#### SUPPLEMENTARY MATERIAL

Supplementary materials mentioned in the text (computer code to conduct permutation tests) are available as part of the online paper from: <http://www.blackwell-synergy.com>.

Please note: Neither AWRA nor Blackwell Publishing is responsible for the content or functionality of any supplementary materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.

#### LITERATURE CITED

- Barnett, T.P., J.C. Adam, and D.P. Lettenmaier, 2005. Potential Impacts of a Warming Climate on Water Availability in Snow-Dominated Regions. *Nature* 438:303-309, doi: 10.1038/nature04141.
- Bartholow, J.M., 2005. Recent Water Temperature Trends in the Lower Klamath River, California. *North American Journal of Fisheries Management* 25:152-162.
- Battikhi, A.M. and A.H. Abu-Hammad, 1994. Comparison Between the Efficiencies of Surface and Pressurized Irrigation Systems in Jordan. *Irrigation and Drainage Systems* 8:109-121.
- Brown, L.R., P.B. Moyle, and R.M. Yoshiyama, 1994. Historical Decline of Coho Salmon in California. *North American Journal of Fisheries Management* 14:237-261.
- California Superior Court, 1950. Shackelford Creek Decree. Siskiyou County, Yreka, California. <http://www.nd.water.ca.gov/PPAs/Watermasters/ServiceAreas/ScottRiver/index.cfm>, accessed November 2005.
- California Superior Court, 1958. French Creek Judgment. Siskiyou County, Yreka, California. <http://www.nd.water.ca.gov/PPAs/Watermasters/ServiceAreas/ScottRiver/index.cfm>, accessed November 2005.
- California Superior Court, 1980. Scott River Decree. Siskiyou County, Yreka, California. <http://www.nd.water.ca.gov/PPAs/Watermasters/ServiceAreas/ScottRiver/index.cfm>, accessed November 2005.
- CDFG (California Department of Fish and Game), 2004. Recovery Strategy for California Coho Salmon. Report to the California Fish and Game Commission, Sacramento, California.
- Cederholm, C.J., W.J. Scarlett, and N.P. Peterson, 1988. Low-Cost Enhancement Technique for Winter Habitat of Juvenile Coho Salmon. *North American Journal of Fisheries Management* 8:4338-4441.
- Drake, D.J., K.W. Tate, and H. Carlson, 2000. Analysis Shows Climate-Caused Decreases in Scott River Flows. *California Agriculture* 54(6):46-49.
- Edgington, E.S., 1995. *Randomization Tests* (Third edition). Marcel Dekker, New York, New York.
- Gleick, P.H. and E.L. Chalecki, 1999. The Impacts of Climatic Changes for Water Resources of the Colorado and Sacramento-San Joaquin River Basins. *Journal of the American Water Resources Association* 35:1429-1441.
- Good, P., 2005. *Permutation, Parametric and Bootstrap Tests of Hypotheses* (Third edition). Springer, New York, New York.
- Good T.P., R.S. Waples, and P. Adams (Editors), 2005. Updated Status of Federally Listed ESUs of West Coast Salmon and Steelhead. U.S. Department of Commerce, NOAA Technical Memorandum, NMFS-NWFSC-66, Seattle, Washington.
- Hamlet, A.F., P.W. Mote, M.P. Clark, and D.P. Lettenmaier, 2005. Effects of Temperature and Precipitation Variability on Snowpack Trends in the Western United States. *Journal of Climate* 18:4545-4561.
- Helsel, D.R. and R.M. Hirsch, 1992. *Statistical Methods in Water Resources*. Elsevier, Amsterdam.
- Hutson, S.S., N.L. Barber, J.F. Kenny, K.S. Linsey, D.S. Kumia, and M.A. Maupin, 2004. Estimated Use of Water in the United States in 2000. U.S. Geological Survey Circular 1268, Denver, Colorado. <http://pubs.usgs.gov/circ/2004/circ1268>, accessed May 2007.
- Johnson, G.S., W.H. Sullivan, D.M. Cosgrove, and R.D. Schmidt, 1999. Recharge of the Snake River Plain Aquifer: Transitioning From Incidental to Managed. *Journal of the American Water Resources Association* 35:123-131.
- Knowles, N. and D.R. Cayan, 2004. Elevational Dependence of Projected Hydrologic Changes in the San Francisco Estuary and Watershed. *Climate Change* 62:319-336.

- Leung, L.R. and M.S. Wigmosta, 1999. Potential Climate Change Impacts on Mountain Watersheds in the Pacific Northwest. *Journal of the American Water Resources Association* 35:1463-1471.
- Mack, S., 1958. *Geology and Ground-Water Features of Scott Valley Siskiyou County, California*. U.S. Geological Survey Water-Supply Paper 1462, U.S. Government Printing Office, Washington, D.C.
- MacKichan, K.A., 1951. *Estimated Use of Water in the United States – 1950*. U.S. Geological Survey Circular 115, Washington, D.C. <http://pubs.usgs.gov/circ/1951/circ115>, accessed May 2007.
- McCabe, G.J. and D.M. Wolock, 1999. General-Circulation-Model Simulations of Future Snowpack in the Western United States. *Journal of the American Water Resources Association* 35:1473-1484.
- McCuen, R.H., 2003. *Modeling Hydrologic Change*. Lewis/CRC Press, Boca Raton, Florida.
- Miles, E.L., A.K. Snover, A.F. Hamlet, B. Callahan, and D. Fluharty, 2000. Pacific Northwest Regional Assessment: The Impacts of Climate Variability and Climate Change on the Water Resources of the Columbia River Basin. *Journal of the American Water Resources Association* 36:399-420.
- Miller, N.L., K.E. Bashford, and E. Strem, 2003a. Potential Impacts of Climate Change on California Hydrology. *Journal of the American Water Resources Association* 43:771-784.
- Miller, S.A., G.S. Johnson, D.M. Cosgrove, and R. Larson, 2003b. Regional Scale Modeling of Surface and Ground Water Interaction in the Snake River Basin. *Journal of the American Water Resources Association* 39:517-528.
- Minobe, S., 1997. A 50-70 Year Climatic Oscillation Over the North Pacific and North America. *Geophysical Research Letters* 24:683-686.
- Mote, P.W., 2006. Climate-Driven Variability and Trends in Mountain Snowpack in Western North America. *Journal of Climate* 19:6209-6220.
- Mote, P.W., A.F. Hamlet, M.P. Clark, and D.P. Lettenmaier, 2005. Declining Snowpack in Western North America. *Bulletin of the American Meteorological Society* 86:39-49, doi: 10.1175/BAMS-86-1-39.
- Nehlsen, W., J.E. Williams, and J.A. Lichtowich, 1991. Pacific Salmon at the Crossroads: Stocks at Risk From California, Oregon, Idaho, and Washington. *Fisheries* 16(2):4-12.
- Neter, J., W. Wasserman, and M.H. Kutner, 1989. *Applied Linear Regression Models* (2nd edition). Irwin, Burr Ridge, Illinois.
- NRC (National Research Council), 2003. *Endangered and Threatened Fishes in the Klamath River Basin: Causes of Decline and Strategies for Recovery*. The National Academies Press, Washington, D.C.
- Poff, N.L., 1996. A Hydrogeography of Unregulated Streams in the United States and an Examination of Scale-Dependence in Some Hydrological Descriptors. *Freshwater Biology* 36:71-91.
- Power, J.H., 2001. Letters: Scott River Flows. *California Agriculture* 55(2):4.
- Ramsey, F.L. and D.W. Schafer, 2002. *The Statistical Sleuth* (2nd edition). Duxbury, Pacific Grove, California.
- Regonda, S.K., B. Rajagoplan, M. Clark, and J. Pitlick, 2005. Seasonal Shifts in Hydroclimatology Over the Western United States. *Journal of Climate* 18:372-384.
- Salathé, E.P., Jr., 2006. Influences of a Shift in North Pacific Storm Tracks on Western North American Precipitation Under Global Warming. *Geophysical Research Letters* 33: L19820, doi: 10.1029/2006GL026882.
- Sandercock, F.K., 1991. Life History of Coho Salmon (*Oncorhynchus Kisutch*). In: *Pacific Salmon Life Histories*, C. Groot and L. Margolis (Editors). University of British Columbia Press, Vancouver, British Columbia, pp. 397-445.
- Snyder, M.A., L.C. Sloan, and J.L. Bell, 2004. Modeled Regional Climate Change in the Hydrologic Regions of California: A CO<sub>2</sub> Sensitivity Study. *Journal of the American Water Resources Association* 40:591-601.
- Stanford, J.A. and J.V. Ward, 1992. Management of Aquatic Resources in Large Catchments: Recognizing Interactions Between Ecosystem Connectivity and Environmental Disturbance. In: *Watershed Management: Balancing Sustainability and Environmental Change*, R.B. Naiman (Editor). Springer-Verlag, New York, New York, pp. 91-124.
- Stewart, I.T., D.R. Cayan, and M.D. Dettinger, 2005. Changes Toward Earlier Streamflow Timing Across Western North America. *Journal of Climate* 18:1136-1155.
- Stewart-Oaten, A., W.W. Murdoch, and K.R. Parker, 1986. Environmental Impact Assessment: "Pseudoreplication" in Time? *Ecology* 67:929-940.
- Venn, B.J., D.W. Johnson, and L.O. Pochop, 2004. Hydrologic Impacts due to Changes in Conveyance and Conversion From Flood to Sprinkler Irrigation Practices. *Journal of Irrigation and Drainage Engineering ASCE* 130:192-200.
- Vicuna, S., E.P. Maurer, B. Joyce, J.A. Dracup, and D. Purkey, 2007. The Sensitivity of California Water Resources to Climate Change Scenarios. *Journal of the American Water Resources Association* 43:482-498, doi: 10.1111/j.175201688.2007.00038.x.
- Zhu, T., M.W. Jenkins, and J.R. Lund, 2005. Estimated Impacts of Climate Warming on California Water Availability Under Twelve Future Climate Scenarios. *Journal of the American Water Resources Association* 41:1027-1038.