Title

Climatic Controls on Streamflow and Snowpack over the Colorado River Basin

Permalink

https://escholarship.org/uc/item/6tc180xp

Author

Xiao, Mu

Publication Date

2020

Supplemental Material

https://escholarship.org/uc/item/6tc180xp#supplemental

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA

Los Angeles

Climatic Controls on Streamflow and Snowpack over the Colorado River Basin

A dissertation submitted in partial satisfaction of the

requirements for the degree Doctor of Philosophy

in Geography

by

Mu Xiao

2020

ABSTRACT OF THE DISSERTATION

Climatic Controls on Streamflow and Snowpack over the Colorado River Basin

by

Mu Xiao

Doctor of Philosophy in Geography

University of California, Los Angeles, 2020

Professor Dennis P. Lettenmaier, Chair

The Colorado River is the main source of surface water for the Southwestern U.S. and Mexico. It is heavily regulated by two large reservoirs, and many smaller ones. Although summer precipitation is about the same amount as winter precipitation averaged over the basin, runoff is mainly generated from the melting of snowpack accumulated during the cold season. Therefore, fresh water availability is challenged by warming temperatures that have occurred over the last few decades. A noticeable downward trend in the basin's naturalized streamflow appears to be related to the increasing temperatures, as precipitation changes have been small. In this dissertation, I evaluate the effects of climatic controlling factors on two major hydrologic components (runoff and snow water equivalent) of the region's water cycle. The primary tool I use is offline land surface simulations from macro-scale hydrological models. In particular, this dissertation comprises three studies that have or will be published as journal articles. First, I employed the

Variable Infiltration Capacity (VIC) model to investigate the causes of the century-long decreasing trend in Colorado River streamflow. I separated the influences of warming temperatures and unevenly (spatially) distributed (and mostly small) precipitation changes by conducting a parallel set of experiments that isolated the various effects. My experiments suggest that more than half of the downward trend in streamflow is attributable to warming temperatures. Compared with an earlier drought period (1953-1968) caused mainly by insufficient precipitation over the UCRB, about half of runoff losses during the Millennium Drought (2000-2014) is attributed to warm temperature. Second, I explored the factors that control snow ablation processes across the West and differences in their representation in a multi-model suite. I selected ten USDA Snow Telemetry (SNOTEL) stations distributed across the mountainous ranges of the Western U.S. Consistent with earlier studies, I find that during the ablation period net radiation generally has stronger effects on melt rates than does air temperature. However, estimates of melt rates vary greatly across the models, in part because of differences in the way they represent the effects of vegetation on the surface energy balance. The canopy effect of each model on snow melting is also evaluated with parallel experiments. Finally, I reconstructed snowpack in the Upper Colorado River Basin (UCRB) over the last 67 years (1949-2015) using the macroscale VIC model implemented at 1/16º latitude-longitude spatial resolution. I then investigated the storms that were associated with accumulation of snow water equivalent (SWE) using 86 SNOTEL stations distributed across the UCRB. In particular, I classified storms associated with SWE accumulation into Atmospheric River (AR) related and non-AR events. The storms I identified (both AR and non-AR) during the study period account for an average of 78.2% of annual peak SWE. On average, 69% of the storms are AR-related; they contribute 56.3% of the annual snowpack maxima. I find no statistically significant basin-wide trends in the number of storms (of either type) or their contributions to SWE. However,

in the middle of the basin, there are a number of grid cells with significant upward trends in the storm contributions to snow, which suggests some movement of snow accumulation towards the UCRB mid-zone over the last few decades.

The dissertation of Mu Xiao is approved.

Glen Michael MacDonald

Yongkang Xue

Rong Fu

Kingtse Mo

Dennis P. Lettenmaier, Committee Chair

University of California, Los Angeles

2020

Table of Contents

List of Figures

List of Tables

Acknowledgement

I would like to thank my family and my friends, who have kindly supported me throughout all these years. I am very grateful for all of the advice and help I've received from faculty members and colleagues within the Department of Geography at UCLA. I also greatly appreciate assistance from Kasi McMurray, Brian Won, Matt Zebrowski and other staff members of the department, who dedicate themselves to helping us students achieve our goals. I thank Professor Qingyun Duan of Hohai University and Professor Bart Nijssen of University of Washington for their indispensable guidance and advice prior to my Ph.D. program. I would further like to acknowledge my Doctoral Committee's contribution to the completion of this dissertation. Finally, I want to express my gratitude to my advisor, Professor Dennis P. Lettenmaier, for guiding me through this challenging but rewarding process.

Curriculum Vitae

Educations

M.S.E., Civil and Environmental Engineering, University of Washington, Seattle   2014

B.S., Mathematics, Beijing Normal University, Beijing   2012

Selected Publications

Chegwidden O., B. Nijssen, D. Rupp, J. Arnold, M. Clark, J. Hamman, S. Kao, Y. Mao, P. Mote, M. Pan, E. Pytlak, **M. Xiao**, 2019: How do modeling decisions affect the spread among hydrologic climate change projections?, *Earth's Future*. doi:10.1029/2018EF001047

Mishra V., A.D. Tiwari, R. Shah, **M. Xiao**, D.S. Pai, and D.P. Lettenmaier, 2019: Drought and Famine in India, 1870-2016. *Geophysical Research Letters*. doi:10.1029/2018GL081477

Xiao M., B. Udall, and D.P. Lettenmaier, 2018: On the causes of declining Colorado River streamflow, *Water Resources Research*. doi:10.1029/2018WR023153

Mote, P.W., S. Li, D.P. Lettenmaier, **M. Xiao**, and R. Engel, 2018: Dramatic declines in snowpack in the western U.S., *NPJ Climate and Atmospheric Science*. doi:10.1038/s41612-018-0012-1

Marlier M.E., **M. Xiao**, R. Engel, B. Livneh, J. Abatzoglou and D.P. Lettenmaier, 2017: The 2015 drought in Washington State: a harbinger of things to come?, *Environmental Research Letters*. doi:10.1088/1748-9326/aa8fde

Xiao M., A. Koppa, Z. Mekonnen, B.R. Pagán, S. Zhan, Q. Cao, A. Aierken, H. Lee and D.P. Lettenmaier, 2017: How much groundwater did California lose during the 2012-2016 drought?, *Geophysical Research Letters*. doi:10.1002/2017GL073333

Wanders N., A. Bachas, X.G. He, H. Huang, A. Koppa, Z.T. Mekonnen, B.R. Pagán, L.Q. Peng, N. Vergopolan, K.J. Wang, **M. Xiao**, S. Zhan, D.P Lettenmaier and E.F. Wood, 2016: Forecasting the hydroclimatic signature of the 2015-16 El Niño event on the

western U.S., *Journal of Hydrometeorology*. doi:10.1175/JHM-D-16-0230.1

Mote, P.W., D.E. Rupp, S. Li, D.J. Sharp, F. Otto, P.F. Uhe, **M. Xiao**, D.P. Lettenmaier, H. Cullen, M.R. Allen, 2016: Perspectives on the causes of exceptionally low 2015 snowpack in the western United States, *Geophysical Research Letters*. doi:10.1002/2016GL069965

Xiao, M., B. Nijssen, and D. P. Lettenmaier, 2016: Drought in the Pacific Northwest, 1920-2013, *Journal of Hydrometeorology*. doi:10.1175/JHM-D-15-0142.1

Tao, Y., Q. Duan, A. Ye, W. Gong, Z. Di, **M. Xiao**, and K. Hsu, 2014: An evaluation of post-processed TIGGE multimodel ensemble precipitation forecast in the Huai river basin, *Journal of Hydrology*. doi:10.1016/j.jhydrol.2014.04.040

Chapter 1 Introduction

The Colorado River Basin, which drains an area of approximately 650,000 km^2 of seven states (WY, UT, CO, NV, NM, AZ and CA) and Mexico, is the primary water source for the U.S. Southwest. The Colorado River is heavily regulated by Hoover and Glen Canyon Dams, which provide storage capacity of over four times the mean annual flow of the river. More than 90% of the natural (unregulated) streamflow in the basin is generated in the Upper Colorado River Basin (UCRB), defined as the area upstream of the U.S. Geological Survey stream gauge at Lees Ferry, AZ (USGS 09380000). Although the climate in much of the Colorado River Basin is warm, ~70% of the streamflow originates from snowpack accumulated during the winter in its mountainous headwaters (Li et al., 2017), which makes the entire system vulnerable to the warming temperature. The basin's natural flow has a decreasing trend over the last century and warming temperatures have been shown to play a role in this decline (Hoerling et al., 2019; Milly & Dunne, 2020; Udall & Overpeck, 2017). Vano et al. (2014) report that substantial additional reduction in Colorado River streamflow could occur by the end of the 21st century due to the combined effects of increasing temperature and projected declines in precipitation. Given the significance of the Colorado River to the Southwest and the challenges of ongoing climate change, understanding the hydroclimatic characteristics and drivers within the basin is of great interest both scientifically, and to stakeholders within the basin.

Among the components of the land surface hydrological cycle, runoff and snowpack are the main focus of this dissertation. Runoff (locally generated as the excess of supply to the soil column over infiltration) contributes directly to streamflow via aggregation in the channel network (Lohmann et al., 1996). Snowpack, as noted above, supplies much of the basin's natural flow. During the period of rapidly warming temperatures that have been prevalent over the last half-

century across the Western U.S. (Hansen et al., 2001; G J McCabe & Wolock, 2002), mountain snowpack has substantially decreased (Mote et al. 2005; 2016; 2017), which has led to a shift in the timing of runoff earlier in the year (Stewart et al., 2005) and reduction in UCRB runoff (Woodhouse et al., 2016).

My dissertation seeks to provide a better understanding of the effects of hydroclimatic drivers with primary focus on the UCRB's hydrological behavior. I address in particular the following three questions:

1) What are the causes of Colorado River streamflow decline over the last century?

2) What are the factors that control snow ablation processes, and how well or poorly are they represented in models?

3) What are the characteristics of winter storms that contribute to UCRB snowpack, and how spatially coherent are these storms? What role do Atmospheric Rivers play in SWE accumulation in the UCRB?

I address these questions in three core chapters of this dissertation. The primary tool I employ is macroscale land surface hydrologic models, including the Variable Infiltration Capacith model VIC (Liang & Lettenmaier, 1994), Simplified SiB (Y. Xue et al., 1991), Catchment (Koster et al., 2000) and Noah-MP (Niu et al., 2011). In **Chapter 2** I employ the VIC model to separate the effects of warming temperatures on long-term reductions in UCRB streamflow from other factors. I isolated attribution of the decreasing streamflow from a hydrological perspective. This chapter was published in *Water Resources Research* (Xiao et al., 2018). **Chapter 3** utilizes a multi-model approach to explore snow ablation behavior over the western U.S. and assesses the effects of variables (temperature, radiation and canopy cover) that control snow ablation processes. This chapter has been submitted to the *Journal of Hydrometeorology* and currently is in revision.

Chapter 4 reconstructs the snowpack in the Upper Colorado River Basin over the last seven decades and evaluates storm contributions (AR- and non-AR) to SWE. This chapter will be submitted to *Water Resources Research*.

Chapter 2 On the Causes of declining Colorado River streamflows

Abstract: The Colorado River is the primary surface water resource in the rapidly growing U.S. Southwest. Over the period 1916-2014, the Upper Colorado River Basin naturalized streamflow declined by 16.5%, despite the fact that annual precipitation in the UCRB over that period increased slightly (+1.4%). In order to examine the causes of the runoff declines, we performed a set of experiments with the Variable Infiltration Capacity (VIC) hydrology model. Our results show that the pervasive warming has reduced snowpacks and enhanced evapotranspiration (ET) over the last 100 years; over half (53%) of the long-term decreasing runoff trend is associated with the general warming. Negative winter precipitation trends have occurred in the handful of highly productive sub-basins that account for over half of the streamflow at Lee's Ferry. We also compared a mid-century drought with the (ongoing) post-Millennium Drought, and find that whereas the earlier drought was caused primarily by pervasive low precipitation anomalies across UCRB, higher temperatures have played a large role in the post-Millennium Drought. The post-Millennium Drought has also been exacerbated by negative precipitation anomalies in several of the most productive headwater basins. Finally, we evaluate the UCRB April-July runoff forecast for 2017, which decreased dramatically as the runoff season progressed. We find that while late winter and spring 2017 was anomalously warm, the proximate cause of most of the forecast reduction was anomalous late winter and early spring dryness in UCRB, which followed exceptionally large (positive) early winter precipitation anomalies.

1.0 Introduction

The Colorado River is the largest river in the Southwestern U.S. It is the source of drinking water for many of the Colorado River Basin's 40 million people, and provides irrigation water to ~13,000 km2 of crops in the U.S. and Mexico (Cohen et al., 2013). It is a lifeline for the population and agricultural economy of parts of seven U.S. states (WY, UT, CO, NV, NM, AZ and CA) and the Mexican states of Sonora and Baja California. The river's naturalized streamflow (see Section 2.2 for discussion of naturalized streamflows) at Imperial Dam (the downstream-most long-term gauging station) has averaged about 20.7 km3/yr (16.8 MAF/yr) over the last century, approximately 90% of which is generated in the Upper Colorado River Basin (Gregory J. McCabe & Wolock, 2007), defined as the ~289,000 km2 of drainage area upstream of the U.S. Geological Survey stream gauge at Lees Ferry, AZ (USGS 09380000). Snowpack stored in the high elevation Rocky Mountain headwater basins contributes about 70% of the annual streamflow (Niklas S. Christensen et al., 2004).

The Colorado River is heavily regulated, mostly by Glen Canyon Dam (Lake Powell) and Hoover Dam (Lake Mead), with combined reservoir storage capacity of 67.5 km2 (54.7 maf). The importance of these reservoirs, which can store close to four times the natural annual flow at Lees Ferry, AZ, has become especially evident during the so-called Millennium Drought, which began about 2000. This drought has coincided with increases in water demand (Rajagopalan et al., 2009), which resulted in Lake Mead reaching its lowest level on record in October 2016. Lakes Mead and Powell dropped precipitously from 2000-2004 due to very low flows (71%, 74%, 41%, 71% and 64% of average, respectively), and have not recovered due to continued high demands equal to inflows and a lack of high flow years. Indeed, only four of the last 18 years have had above average river discharge, limiting reservoir refill opportunities.

A pronounced warming trend across the Colorado River Basin (CRB) since the 1970s (Dawadi & Ahmad, 2012) has further contributed to the post-2000 imbalance between CRB runoff and water demand. Vano et al. (2012) evaluated the temperature sensitivity (annual average streamflow change per 1°C temperature change) and found that the average sensitivity of annual runoff at Lees Ferry was around -5% per °C), suggesting that warming over the last ~50 years may account for a 5-10% reduction in annual streamflow over that period.

Several studies have investigated the effects of ongoing warming on the flow of the Colorado River. Barnett and Pierce (2009) concluded that anthropogenic climate change would reduce CRB runoff by 10%-30% by 2050. Reynolds et al. (2015) predicted that minimum streamflows will decline as warming of the basin continues. Woodhouse et al. (2016) report an increase in the frequency of warm years with low streamflow since 1988. McCabe et al. (2017) found that increases in temperature since the late 1980s have decreased runoff generation efficiency, reducing streamflows by 7%. Udall and Overpeck (2017) similarly found temperature-induced streamflow decreases of approximately 6% during 2000-2014 and projected large mid-century temperature-induced declines of 20% or more should precipitation not change.

Here, we utilize a hydrological model applied for the period 1916-2014 (all data are for water years if not specified otherwise) to evaluate the spatial and temporal signature of the Millennium Drought in the CRB. Along with a baseline simulation forced by gridded observations, we perform a T-detrend experiment, in which we remove the long-term temperature trend from the model forcings, to investigate the role of the warming on streamflow declines both over the long term and during the recent drought. We analyze runoff in each of 20 sub-basins of the CRB, which allows us to study spatial variations in runoff generation and anomalies. We also analyze the historical 1953-68 drought in an attempt to shed light on how the hydrologic response to climate

variations has changed in recent decades, and during the Millennium Drought in particular. Finally, we dissect the 2017 April-July streamflow forecast to understand the role of late winter and early spring precipitation and temperature in the substantial seasonal forecast reductions that occurred as water year 2017 progressed.

2.0 Data and approach

2.1 VIC model and forcings

The Variable Infiltration Capacity (VIC) model is a physically-based, semi-distributed hydrological model which represents the land surface water and energy budgets over a grid mesh (here 1/16th degree spatial resolution) and routes runoff through a prescribed river network to produce streamflow estimates at specified river nodes (Liang & Lettenmaier, 1994). We applied the model at a daily time step, using what is termed full-energy balance mode, meaning that the model iteratively solves the surface energy budget by estimating the effective surface temperature at each time step. Therefore, the daily average surface temperature produced by VIC is not the average of the forcing temperatures, i.e. 0.5*(daily maximum+daily minimum). Unless stated otherwise, the temperatures we report here are outputs from the VIC simulations.

Similar to other land surface models, the fundamental water balance equation in VIC can be summarized as Runoff (RO) = Precipitation (P) - Evapotranspiration (ET) – changes in Soil Moisture (ΔSM) - changes in Snow Water Equivalent (ΔSWE). Groundwater is not represented in the version of VIC we used; Rosenberg et al. (2013) found that inclusion of a parameterzation of groundwater had little effect on the model's streamflow simulations in the CRB. It is important to note that VIC represents snowpack sublimation within its winter ET. Sublimation is sparsely measured but nonetheless is important to some aspects of our study (Andreadis et al., 2009); we describe the model's performance with respect to sublimation in section 4.2. The VIC model has

7

been successfully applied previously in a number of hydrological studies over the CRB and the U.S. Southwest (Nadia S Christensen & Lettenmaier, 2007; Niklas S. Christensen et al., 2004; Mote et al., 2005, 2018; Vano et al., 2012, 2014)

The VIC model simulates surface hydrological processes with parameterizations of sub-grid vegetation, soil variability and topography, and has provided plausible representations of CRB surface water conditions in the above-referenced studies. We forced the model with an updated version of the Hamlet and Lettenmaier (hereafter H&L) dataset (Hamlet & Lettenmaier, 2005) at 1/16-degree resolution for the period water years 1916-2014. We chose the H&L data set because its long-term variability is indexed to the U.S. Historical Climatology Network (HCN; Easterling et al. 1996) stations in the region, which have been carefully quality controlled for effects that could otherwise result in spurious trends, such as station moves and instrument changes (e.g., the shift to MMTS temperature sensors in the 1980s). As described in Hamlet and Lettenmaier (2005) the H&L dataset uses HCN station data to constrain decadal variability (and hence long-term trends), hence is in our view most appropriate for exploration of the causes of century-scale streamflow declines over our study period 1916-2014.

2.2 Naturalized streamflows

To evaluate our model simulation results, we used naturalized streamflow data for the Colorado River produced by the U.S. Bureau of Reclamation (USBR); see https://www.usbr.gov/lc/region/g4000/NaturalFlow/current.html for details. The naturalized streamflows are derived from USGS historical streamflow observations by a process of adjustments that compensate for anthropogenic effects including consumptive uses of water, reservoir storage, trans-basin diversions, and other effects (see USBR, 1983). The naturalized streamflow data sets are produced for 29 well-distributed tributary stations across the CRB (as

well as the main stem) for the period 1906 through 2015. Others (Prairie & Callejo, 2005) have noted that USBR has improved the quality of the naturalized flow data set after 1971 and the estimates may be somewhat better after that time.

2.3 Sub-basin analysis

We performed our analyses for the Colorado River above Imperial Dam, as well as for the 20 sub-basins delimited by USGS WaterWatch gauges (see Figure 1) which are a subset of the 29 naturalized streamflow points noted above. The river channel network dataset we used is from Wu et al. (2012), based on which we determined the masks for each of the 20 sub-basins. The Wu et al. sub-basins are similar to, but slightly different from the more familiar 6-digit Hydrologic Unit Codes normally used in the Basin. Detailed information about each sub-basin is reported in the SI. It is important to note that our analysis excludes the Gila River given its distinct hydrological and legal characteristics. The Gila River joins the Colorado River below Imperial Dam just upstream of the U.S. border with Mexico, and in recent years has been mostly dry at its mouth due to upstream uses by Arizona. Since 1964, the U.S. Supreme Court has excluded it from administration under the Colorado River Compact. Although the Gila is an important basin, its absence from this study is logical given its unique status.

2.4 Model testing and evaluation

Table 1 summarizes the long-term runoff contribution percentages from nine major sub-basins at which naturalized streamflows are available, and for which we also produced VIC simulations. The runoff contribution percentages from the model and naturalized flows generally are in good agreement. The Upper Basin (UCRB; defined as the drainage area above Lees Ferry, AZ) produces more than 90% of the flow at Imperial Dam. Therefore, we mainly focus on the UCRB here, acknowledging unusual Lower Basin (LCRB) conditions when noteworthy.

Figure 2a shows the annual time series of naturalized streamflow (NFL) and VIC simulations at Lees Ferry, AZ. Both the annual naturalized streamflows and VIC simulations (r2=0.75), and their trends over the period of record (NFL: -3.3 km3/yr, VIC: -3.4 km3/yr) are similar, suggesting that the VIC model provides a plausible representation of natural conditions (i.e., those responding primarily to climate forcings) and long-term hydrologic change in the basin. Hereafter, we mainly focus on VIC results in our analysis of UCRB sub-basin long-term (1916-2014) trends and comparison between the 1953-1968 and the Millennium drought. The annual precipitation and average temperature (calculated by VIC as noted in section 2.1) time series plots are also presented in Figure 2.

3.0 Results

3.1 Basin-wide trend analysis

Table 2 summarizes long-term linear (regression) trends for the UCRB for four hydrological variables (precipitation, evapotranspiration, runoff and April 1 snow water equivalent) from the baseline VIC simulation, and the temperature detrended (T-detrend) simulation. We also computed trends using the non-parametric Theil-Sen slope estimator (Sen, 1968; Theil, 1950) and found that they generally are in close agreement (Table S1). Therefore, we refer to the linear trends hereafter for convenience. The T-detrend simulation uses the same forcings as the baseline, except that annual linear trends in the daily temperature maxima and minima are removed. We also disaggregated summer season (Apr-Sep) and winter season (Oct-Mar) for each variable (all summers and winters mentioned hereafter are so defined).

Over the simulation period 1916-2014 the UCRB annual precipitation increased by +1.5 km3 (1.4%) whereas winter precipitation, which is the main source for Apr 1st snow water equivalent and streamflow in the spring and summer, had only a very small (not statistically significant)

negative trend (long-term ΔP is -0.1 km3, -0.2%). In our baseline simulation, the long-term linear change of annual runoff (ΔRO) in the UCRB is -3.4 km3 (-16.5%) and long-term change in annual evapotranspiration (ΔET) is +4.2 km3 (+4.7%). Apr 1st SWE decreased significantly (ΔSWE -9.1 km3, -39.0%), which reduces warm season streamflow from the Upper Basin, as evidenced by summer RO decreases (-3.8 km3, 23.3%) even given a positive trend in summer precipitation (ΔPsummer is +1.6 km3). As summer RO makes up more than 3/4 of the annual RO in the UCRB, the long-term annual ΔRO is negative as noted above, although summer RO decreases are slightly compensated by increasing winter RO (ΔROwinter +0.4 km3, 10.4%).

We performed the T-detrend simulation using the same precipitation as the baseline simulation but with the temperature trend removed from the forcing dataset on a grid cell by grid cell basis. In this no-warming-trend scenario, the long-term decreasing trend in annual runoff is reduced to -1.6 km3 (-7.7%), from -3.4 km3 but not eliminated. It suggests that 53% (-1.8 / -3.4) of the annual runoff trend is attributable to the annual warming temperature. The increase in ET in the T-detrend simulation is smaller by 1.9 km3 (baseline: +4.2 km3, T-detrend: +2.3 km3), which explains the increase in runoff (1.8 km3) to within 0.1 km3.

The numbers in Table 2 also show that the effects of the temperature trend on winter RO (baseline: +0.4 km3, T-detrend: +0.4 km3) and summer ET (baseline: -0.8 km3, T-detrend: -0.6 km3) are small. Increasing temperatures cause a decrease in summer RO (baseline: -3.8 km3, T-detrend: -1.9 km3), and an increase in annual ET (baseline: +4.2 km3, T-detrend: +2.3 km3) that comes mostly in the winter (baseline: +4.9 km3, T-detrend: +2.9 km3). On a percentage basis, both of these increasing winter trends in ET are substantial over the 1906-2014 period: a 30% increase in the baseline ET and an 18% increase in the T-detrend simulation ET. The summer ET changes of -1.1% and -0.8% are comparatively small. It is worth noting that the long-term trend in

UCRB winter ET is positive in the T-detrend simulation even given no significant trend in winter precipitation. The positive trend in winter ET is mainly caused by increased snow sublimation. Although sublimation is strongly controlled by surface temperature, other factors also contribute as well (see Section 4).

The remaining -1.6 km3 (-7.7%) decrease in RO in the T-detrend simulation is curious given the increasing summer precipitation (ΔPsummer +1.6 km3, 3.0%) and negligible winter precipitation change (ΔPwinter -0.1 km3, -0.2%). In addition, although the SWE anomaly in the T-detrend simulation is less compared with that of the baseline simulation (baseline: -9.1 km3, T-detrend: -5.6 km3), the long-term 1906-2014 SWE trend is still negative in the T-detrend simulation (-23.9%). Winter ΔET in the T-detrend simulation is only +2.9 km3 as reported in Table 2, which cannot explain all of the SWE anomaly. One possible answer is that while the overall basin-wide precipitation changes over time are small, precipitation declines in the most productive basins while increasing in the less productive basins. We explore the effects of such spatial variations below.

3.2 Sub-basin conditions

Figure 1 shows that there are four sub-basins in the upper CRB (denoted by red numbers) that produce most of the UCRB runoff: the Yampa River, Colorado River near Cameo, Gunnison River and San Juan River (from north to south, respectively). The most productive sub-basin is the Colorado River near Cameo (USGS 09095500) in the northeastern part of the UCRB. This sub-basin produces almost one quarter of the total naturalized runoff of the UCRB. It contains not only the mainstem, but also several large tributaries, including the Eagle, the Roaring Fork, and the Blue. A little more than 30% of the UCRB flow is produced by the other three subbasins, and

in total about 55.5% of the total discharge of the UCRB is attributable to these four tributaries. Below, we discuss the nature of the long-term changes in these critical sub-basins.

Figure 4 shows annual precipitation, ET and runoff changes for all sub-basins over the 1916-2014 study period. The top row is extracted from our baseline simulation and the bottom row is from the T-detrend simulation. We note that although some sub-basins appear similar between baseline and original maps, the numbers are more different than they might appear by visual inspection of the maps (Tables S3 and S4). We calculated the changes relative to the initial value of each linear fit, shown in Table 2. Figure 4a shows a noteworthy east-west dipole in the precipitation changes over time in the UCRB. In the UCRB, precipitation decreases have occurred mainly in the high runoff generating northeastern part of the basin while several sub-basins in the northwestern part of UCRB show long-term annual precipitation increases.

Precipitation declines have also occurred in the LCRB where little runoff occurs. These decreases in precipitation led to declines in ET and little change in sub-basin runoff (Figures 4c and 4f), with negligible impact on total basin runoff (e.g., at Imperial Dam).

There are two sub-basins in the northeastern part of the UCRB, that have relatively large annual precipitation decreases of -2.3 km3 (Colorado River above Cameo) and -0.7 km3 (Gunnison River) with a combined runoff decrease of -2.9 km3 (SI). These are the same highly productive sub-basins shown in Figure 1 and are a major driver of the overall annual runoff decline. Four basins in the northwestern part of UCRB with increasing precipitation (the Green River downstream portion along with its San Rafael River and Duchesne River tributaries; colored in deeper blues in Figure 4) have partially offset these long-term runoff declines by about 1.0 km3. Figures 5 and 6 are similar to Figure 4 but for winter (Oct-Mar) and summer (Apr-Sep), respectively. Winter runoff changes are small for both the baseline and T-detrend simulations, as

most runoff occurs during the summer season. Although the total precipitation amounts are similar during warm and cold seasons, winter precipitation is much more important to the UCRB's runoff. Summer precipitation mainly contributes to ET rather than runoff, as high summer temperatures lead to large ET, especially at lower elevations. Winter precipitation in mountain headwater regions accumulates as snowpack and contributes mostly to RO rather than ET, when it melts.

April 1 SWE trend plots for all the sub-basins (Figures 3a/b) show that the four highly productive sub-basins (Yampa River, Colorado River near Cameo, Gunnison River, and San Juan River) in the northeastern part of the basin that contribute much of the runoff losses in the UCRB have all experienced substantial SWE decreases. Those sub-basins are also snow-dominant regions as indicated by Figure 3c. Figure 5a shows that winter precipitation has declined in all of the northeast UCRB sub-basins except for the San Juan River, which shows a positive winter precipitation trend. Nonetheless, both SWE (Figure 3a) and annual RO (Figure 4c) in the San Juan Basin are decreasing. The reason is that winter ET has increased substantially: ΔP_{winter} is +0.4 km3 while long-term ΔET_{winter} is +1.1 km3, with SWE decreasing by -0.7 km3, or 30.1%. Declines in SWE in the other three basins, all of which experience declines in precipitation, are more severe and range from -46% to -49%. The increased winter ET, along with reductions in precipitation in these basins explain the strongly decreasing SWE and substantially explain the declines in sub-basin runoff.

As noted above, 53% (1.8 of 3.4) of the long-term runoff trend in the UCRB is related to warming temperatures. To dissect the remaining -1.6 km3 (47%) in the T-detrend simulation, we performed a P- and T-detrend experiment, in which we removed both the temperature and winter precipitation trend from the original input dataset. Importantly, under this experiment the northeast UCRB basins see increased winter precipitation while the northwest basins see decreased winter

precipitation relative to the baseline and T-detrend simulations. Note, also, that we do not modify the summer precipitation, which increased over the study period. Under the P&T-detrend simulation, the UCRB's long-term runoff losses become -0.6 km3 (1.0 km3 less than the pure T-detrend and 2.8 km3 less than the baseline). The residual -0.6 km3 loss over the 1916-2014 period is attributable to increased winter ET. Section 4.2 below evaluates why ETwinter shows a positive trend given no-P-trend and no-T-trend. The total runoff decline of -3.4 km3 can thus be attributed to warming (-1.8 km3), insufficient P in the northeast part of CRB (-1.0 km3) and increased winter ET (-0.6 km3).

Summer precipitation and summer ET trend spatial plots (Figure 6a-6d vs 6b-6e) show similar patterns for both the baseline and T-detrend simulations: negative trends have occurred over the LCRB and the eastern UCRB while some increases have occurred in the northwestern headwaters. The spatial patterns confirm that in the summer increases in precipitation drive increases in ET while decreases in precipitation drive decreases in ET over both the LCRB and UCRB when surface air temperatures are relatively high.

In the UCRB the baseline simulation Apr-Sep runoff (Figure 6c), which constitutes almost three quarters of the CRB annual total, shows spatial patterns similar to the SWE spatial plots in Figure 3. Taken together, the figures show where water is stored as snow in the UCRB during winter in the cold, high-elevation headwaters regions and how SWE then contributes to runoff in the following spring and summer. Over the last century, warming temperatures, reduced winter precipitation in the most productive mountain sub-basins in the UCRB and slight increases in winter ET (Figure 5b) lead to reduced SWE and consequently reduced runoff.

In the LCRB, the annual precipitation, ET and runoff plots show mostly P decreases, ET increases and small RO changes (Figure 4). In winter, some P increases occur in the NW portion,

ET increases everywhere except in the south, and RO has little change (Figure 5). Summer shows decreasing P, increasing ET and little RO change (Figure 6).

3.3 Drought comparisons

In order to examine the causes of the Millennium Drought, we compare the recent dry period from 2000-14 (D2) with the 1953-68 drought (D1). Figure 7 shows the time series of UCRB annual streamflow volume. Long-term averages are marked as the black horizontal baseline and the LOWESS-smoothed VIC streamflows are plotted in red. We report basin-wide (CRB, UCRB and LCRB) annual average anomalies for four selected variables (P, SWE, ET and RO) over the two drought periods in Table 3. Spatial anomaly plots by sub-basin of P, SWE, ET and RO for the 1953-1968 and 2000-2014 periods are shown in Figures 8 and 9.

Similar to the long-term trends discussed in section 3.3, comparison of the annual anomalies of precipitation, ET and runoff during both droughts in Table 3 confirms that the UCRB dominates total basin-wide runoff production during drought periods as in the long term. In the Millennium Drought annual precipitation decreased more in the LCRB, which substantially reduced ET, but not runoff. This is a very large part of the overall basin-wide ET loss (-7.9/8.7 km3), but the LCRB ET does not make much difference to streamflow because most Lower Basin precipitation is converted to ET, drought or no drought. Since our primary interest is on the causes of declining runoff, we again focus on the UCRB.

Table 3 summarizes climate and hydrological differences and similarities between the two drought periods. In particular, UCRB RO anomalies for the two drought periods are quite similar (-2.4 vs -2.6 km3; all the numbers are D1 vs D2 in this paragraph) whereas the SWE decrease is much greater in the Millennium Drought (-2.7 vs -4.4 km3). Although the basin-wide annual (negative) precipitation anomaly in 1953-1968 is much less than the Millennium Drought (-8.8 vs

16

-11.4 km3), this order is reversed in the UCRB (-6.1 vs -3.2 km3), where most runoff is generated. In the UCRB, the earlier 1953-68 drought has less average annual precipitation than the Millennium Drought (104.6 vs 107.5 km3), especially in winter when precipitation in the UCRB differentially contributes to runoff production, as discussed in section 3.2. Winter precipitation in the UCRB is 51.5 km3 and 54.5 km3 for 1953-1968 and 2000-2014 respectively; whereas summer precipitation is nearly identical (53.1 km3 and 53.0 km3; Table S6). Much higher temperatures (+0.1 vs +1.0°C), less SWE (-2.7 vs -4.4 km3) and more winter ET (+0.4 vs +1.8 km3) are indicative of additional key differences between the two droughts.

Table 3 combined with Figures 8a-e shows that the 1953-1968 drought mainly resulted from a spatially widespread and consistent negative precipitation anomaly across most of the UCRB. Temperatures were within 0.1 °C of the climatological mean. The corresponding ET and runoff anomalies therefore mostly reflect the precipitation reductions in each sub-basin. SWE clearly decreases uniformly in almost all parts of the UCRB, as does runoff. Note that Figures 8a and 8c-e all have similar patterns.

Interpretation of anomalies during the Millennium Drought is more complicated due to spatially heterogeneous conditions. Pervasive anomalously high temperatures, resulting in part from the long-term warming-trend, which emerged around the 1970s and exacerbated by drought-specific warming, play a substantial role (Figure 9b). In addition, D2 average ET in the UCRB (Table 3, Figure 9d), only 0.8 km3 less than the climatological mean (despite drier conditions), combined with precipitation reductions in the most highly productive sub-basins (Figure 9a) caused large runoff reductions in those key basins. In the UCRB, the western sub-basins experienced positive precipitation anomalies with commensurate increases in ET. The northeastern sub-basins where snow dominates and most of the UCRB runoff originates (Figure

3c), experienced negative precipitation anomalies but without commensurate decreases in ET, which acted to amplify the SWE reductions. Thus, Figures 9c and 9e show substantial declines in SWE and RO from these northeastern basins along with smaller declines and even some increases in SWE and RO from the northwestern basins. Unlike the 1950s drought, the spatial patterns in Figures 9a and 9c-e are highly complex.

Eight basins – four from the highly productive northeast, and four from the less productive northwest -- provide additional insights into how spatially heterogenous precipitation, ET, and SWE combined to produce spatially variable runoff in the Millennium drought. The four most highly productive sub-basins (marked with red numbers in Figure 1) contributed more than 83% of the total -2.6 km3/yr RO anomaly in the Millennium Drought; their contribution was only 34% of the -2.4 km3/yr RO anomaly during 1953-1968 (numbers of each sub-basin are provided in the Supporting Information). Four sub-basins on the western side of the UCRB (draining the Uinta and Central Utah Mountains) had positive annual precipitation anomalies during 2000-2014 (leading to 0.5 km3/yr positive RO anomaly) but that positive anomaly was more than canceled by other runoff-losing sub-basins (-2.4 km3 for the 4 highly productive northeastern basins). Compared with 1953-68, precipitation anomalies were much more uneven in the Millennium Drought. The relatively evenly distributed positive +1°C temperature anomalies lead to more winter ET (+0. 8 km3, 3.7% of the annual streamflow) and reduced SWE (-4.8 km3, 23.0% of the annual streamflow), exacerbating the precipitation reductions over the UCRB.

By combining our T-detrend and P&T-detrend simulations we can gain additional insights into the Millennium Drought when used in a similar fashion to our long-term trend analysis. Comparing the two simulations suggests that the temperature anomaly was responsible for -1.4 km3 of the Millenium Drought runoff loss (total is -2.6 km3), while the precipitation deficit caused

-1.0 km3 of the remaining -1.2 km3 runoff loss. The average runoff in the P&T-detrend results is quite close to the long-term climatology (P&T-detrend: 18.7 km3, climatology: 18.9 km3, less than 1% difference), suggesting that the model precipitation and temperature changes are faithfully capturing the drought causes.

4.0 Interpretation and discussion

4.1 Long-term Trends

The Colorado River is snow-dominated, although only about 18% of the entire basin area accumulates enough SWE to produce substantial spring and summer RO (see Apr 1st SWE climatology >50mm as shown in Figure 3c). Basin-wide Apr 1 SWE is approximately 20 km3, which is close to the annual runoff at Lees Ferry. Li et al. (2017) show that for the UCRB, SWE accounts for 71% of annual runoff on average. Summer (Apr-Sep) RO constitutes almost ¾ of the total annual RO in both the UCRB and the entire basin. Clearly then, winter precipitation (and hence spring SWE) are closely linked to annual runoff changes. Although the overall winter precipitation trend from 1916-2014 is not significant over the entire UCRB (-0.2%, Table 2), uneven spatial distribution causes important winter precipitation decreases in several of the snow-dominant most runoff-productive headwater sub-basins. Warming temperatures over our nearly hundred-year period of record in the UCRB (annual long-term ΔT is 1.8°C as in Figure 2) induce -1.8 km3 (53%) of the annual runoff losses totaling -3.4 km3. The remaining -1.6 km3 results from negative winter precipitation anomalies, mostly in the northeastern sub-basins of UCRB (-1.0 km3) and increasing winter ET (-0.6 km3).

4.2 Winter ET and Sublimation

We found that increasing winter ET in both the baseline (4.9 km3) and the T-detrend (2.9 km3, Table 2) comes mainly from snow sublimation. In the T-detrend simulation, the November

to February long-term change of UCRB sublimation is 2.2 km3 (75.9% of the 2.9 km3 ΔET winter increase) with the remaining 0.7 km3 from increased evaporation in March. A possible cause of these trends in individual months was our approach using annual rather than seasonal (e.g., monthly) trend removal. Therefore, we performed another simulation with temperature detrended on a grid cell by grid cell basis for each month, instead of annually. This resulted in a considerable decrease in the March ET trend, which apparently was caused primarily by the increasing annual temperature trend. However, snow sublimation from October to February still showed increasing trends in this monthly T-detrend simulation. We were therefore left to explain the positive trends in snow sublimation over Oct-Feb given neither temperature nor precipitation trends.

We considered other factors that can influence the sublimation process in VIC. We found that the winter months had positive trends in surface aerodynamic resistance (AR), which leads to positive trend in surface snow sublimation. The AR trend was traced to the wind forcings in our VIC input dataset, which are based on NCEP/NCAR reanalysis, the record for which starts in 1949. Following Livneh et al. (2013), absent wind data prior to 1948, the earlier values were set to their monthly climatological averages. Although this approach didn't result in a trend in wind over the 1916-2014 period, the non-linear relationship between AR and wind speed results in larger AR values occurring after 1948 and thus results in the long-term increasing sublimation trend. While the resulting overall RO negative trend associated with this effect was modest (-0.6 km3), we changed our pre-1949 wind values by randomly sampling from the later (post-1948) record. This resulted in the long-term UCRB annual RO trend becoming essentially zero in a new P&T detrend simulation. Livneh et al., [2013] reported that using wind climatology had only small impacts on their long-term mean RO, but in the case of the relatively dry CRB, the abrupt change in wind variability created artificial sublimation that was not negligible.

4.3 Drought Comparisons

Compared to the 1953-68 drought, the causes of the Millennium Drought are more complicated. During the 1953-1968 drought, annual precipitation anomalies were negative across the entire CRB (Figure 8a) and temperature was close to its long-term mean (Figure 8b). Sub-basin runoff anomalies, as well as SWE and ET anomalies, all responded primarily to the precipitation deficits. In contrast, the upper and lower part of CRB behaved much differently during the Millennium Drought. In the UCRB, both winter and summer precipitation during 2000-2014 are just slightly below their climatologies (54.4 km3/winter compared to 55.8 km3/winter long-term mean) and 53.0 km3/summer (compared to 55.0 km3/summer long-term mean). The UCRB received approximately normal (slightly negative anomalies) winter precipitation, which was clearly higher than Pwinter during 1953-1968 as noted in Section 3.3, but produced less annual runoff (16.3 km3/yr vs 16.5 km3/yr).

The situation is reversed, however, if the temperature trend is removed. In this case the 1953-1968 drought becomes worse than the Millennium Drought. In the T-detrend simulation, the average annual runoff for the UCRB during 1953-1968 and 2000-2014 was 17.2 km3/yr and 17.7 km3/yr, respectively (baseline annual runoff climatology is 18.9 km3). Therefore, the warming temperature accounts for 54% of the annual runoff anomaly during the Millennium Drought (-1.4 km3/yr of -2.6 km3/yr), which is very close to its 53% contribution to the long-term decreasing runoff trend. The other half of the runoff deficit was caused by UCRB's negative winter precipitation anomalies in the northeastern part of the basin where the highest runoff-generating sub-basins are. The winter ΔP over 2000-2014 in those four highly productive sub-basins was -2.4 km3/yr, much larger (in absolute value) than ΔPwinter over 1953-1968, -0.9 km3/yr. Exacerbated by above normal winter temperature in the baseline simulation, the UCRB winter ET anomaly

over 2000-2014 was 1.8 km3/yr and ΔSWE is -4.4 km3/yr (23.7% less compared to the climatology).

These results demonstrate that warming temperature was a major driver for the UCRB's runoff shortage over the Millennium drought, in agreement with Udall and Overpeck (2017). In the Lower Basin, annual precipitation had very serious negative anomalies across the entire LCRB as shown in Figure 9a: all sub-basins exhibited pronounced negative anomalies. While temperatures were also higher across the LCRB, there is no need to invoke a temperature forcing to explain the drought. As noted above, though, these LCRB precipitation anomalies have little effect on RO.

Using the Millennium Drought anomalies, we can estimate the runoff-precipitation-elasticity relationships as follows: the baseline average annual runoff for the UCRB is 18.9 km3 and the T-detrend runoff is 17.7 km3, therefore, the 1.2 km3 runoff decrease apparently is attributable to precipitation. Over 2000-2014 annual precipitation in the UCRB was 107.5 km3 and the climatology was 110.8 km3/yr, so ΔP/P is -0.029. The implied elasticity is 2.12 (ΔRO • RO-1 • ΔP-1 • P = -0.0616/-0.0291), which is in good agreement with Vano et al. (2012).

4.4 Uncertainties

The results and analysis we've presented to this point are based on VIC simulations forced by the extended H&L dataset. The robustness of the conclusions is potentially dependent on both the forcings and model performance. In order to examine the robustness of our results, we performed an exploratory uncertainty analysis of both the model forcings and hydrological model. First, we compared the H&L forcings to two other widely-used gridded climate datasets: Precipitation Regressions on Independent Slope Method (PRISM; Di Luzio et al. 2008) and Livneh (Livneh et al., 2013). Over the UCRB, trends in annual precipitation of these three datasets (H&L,

PRISM and Livneh) have long-term annual trends ranging from -6% to +2%, and for winter precipitation from -10% to +6%. As for the temperature, on an annual basis the positive trend over UCRB range from 1.0˚C to 1.4˚C, and for winter temperature from 1.0˚C to 1.6˚C. As noted in the Supplement (Table S7), the H&L temperature trends generally are larger than for the other two datasets (also see section 2.1; the VIC temperature trend is not the same as the H&L trend but rather is somewhat larger, approximately 0.4˚C, as it results from energy budget closure in the model). The relatively large negative precipitation trend in Livneh is mostly attributable to large annual precipitation early in the record, and in likelihood is traceable to the relatively liberal criterion that data set uses to allow entry of stations with relatively short record lengths.

Our choice of the H&L data set is based on its relationship with the HCN station data (to which its decadal variability is controlled; see Hamlet and Lettenmaier, 2005). The HCN data have been carefully quality controlled, and in this sense arguably are more appropriate for trend-related studies than are the other two data sets (or for that matter, other data sets we might have chosen). We evaluated the H&L long-term temperature trend over UCRB (1.4˚C increase) in comparison with the simple average over all HCN stations in the UCRB (also 1.4˚C increase; identical to two significant figures). On this basis, and given the criteria used in construction of the H&L data set, we believe it is most appropriate for our purposes. We do not believe that other methods that, for instance, might use multiple ensembles and effectively average either inputs to our outputs from our hydrological model would be appropriate given the objectives of our analysis.

As for hydrological models, we extracted the Noah-MP and VIC results from the UCLA Drought Monitor (Xiao et al., 2016) for model comparison (note that the forcings for the UCLA Drought Monitor are different than H&L, but are common to the two models). Over the entire Upper Basin and the four most productive sub-basins we identified, the long-term trends in Noah-

MP and VIC runoff are generally consistent. For instance, for the entire UCRB, (VIC: -3.5 km3/yr; Noah-MP: -4.3 km3/yr); see also sub-basin trends shown in Figure S2. Although different models would no doubt produce somewhat different results, the fact that VIC and Noah-MP, which have essentially no common heritage, produce similar trends gives us some confidence that our results are reasonable model-independent.

This uncertainty analysis improves the confidence in our conclusions. Nonetheless, more work could be done along these lines. For example, there is substantial uncertainty in the gridded forcing data sets we used, which are sparse and especially rare at high-elevations. More sophisticated methods could be used to represent the uncertainty in the gridded data sets (aside from testing sensitivity to different data sets, as we have done). Furthermore, land surface models, which simulate complex systems, contain approximations and uncertainties that produce errors that are difficult to represent in analyses such as ours. Thus, given computational constraints, less than complete understanding of physical processes and limited observation resolutions, state-of-the-art land surface models will inevitably produce somewhat uncertain results. We acknowledge these uncertainties, which no doubt will motivate future work. We nonetheless argue that our results in the larger sense transcend the effects of these uncertainties, in particular given their robustness with respect to models and model forcing data sets.

5.0 2017 Streamflow forecast

The Colorado Basin River Forecast Center (CBRFC) produces seasonal (Apr-Jul) streamflow forecasts starting about January 1 with monthly updates for the CRB using its Ensemble Streamflow Prediction (ESP) approach (Werner & Yeager, 2013) based on the Sacramento Soil Moisture Accounting model (Burnash et al., 1973). General characteristics of Sacramento and VIC simulations, and hence ESP forecasts, are roughly similar (Vano et al., 2012).

The CBRFC forecast utilizes historical meteorological forcings for 1981-2010 to generate an ensemble of future streamflow series given hydrological conditions (soil moisture and SWE) on the forecast initiation date (e.g., April 1), which are taken from an historical model simulation. We analyzed the forecasts issued on the first day of each month in 2017 from January through June. The official CRBFC forecast for the UCRB 2017 Apr-Jul streamflow (natural flow at Lees Ferry) decreased dramatically from much above normal on January 1 as the runoff season progressed. Some media reports attributed these decreases to anomalously warm late winter and spring conditions, and drew parallels between water year 2017 conditions, and the long-term trends analyzed above, especially in temperature.

We evaluated the causes of the changes in the 2017 forecasts using the same ESP approach as used by CRBFC, but using the VIC rather than the Sacramento model. Because the ESP method requires near real-time records and meteorological forcings, we used the UCLA/UW Drought Monitor data set (see Xiao et al. 2016) to perform the retrospective ensemble forecasts. Figure 10 shows the predicted naturalized streamflow at Lees Ferry for each forecast initialized on the first day of each month. The red line shows the official forecasts produced by CRBFC and the blue line is the average of the ensemble predictions generated using the UCLA/UW drought monitor dataset. The green line shows the streamflow predictions that would have been made with a perfect precipitation forecast (they come from a VIC simulation with observed 2017 precipitation) but with temperature ensembles taken from observations for 1981-2010.

We performed this experiment to separate the effect of precipitation and temperature on the ESP results. In interpreting the forecasts, it is important to note that the forecast period is the same (April-July) for all forecasts, even though for post-April 1 forecasts, part of the forecast period has already occurred, and some of the water literally has already gone "under the bridge".

25

It is clear that both the red and blue curves exhibit peaks around February-March with forecasts declining later. The CBRFC forecasts are higher than those made with VIC, which most likely is attributable to a different hydrologic model and different model forcing data sets; however, both sets of forecasts have the same general patterns. Also, both sets of forecasts are still above climatology for the last forecast (June 1), due to anomalously high SWE early in the forecast period. From the green line we can infer that the differences between the perfect precipitation forecasts initialized at each time and climatology are considerably smaller than the differences between either of the ESP forecast sets and climatology.

Given perfect precipitation forecasts, the forecasts vary from 100% to 110% of the mean, which are close to the true value (observed flow relative to climatology) of 105.3%. Anomalously warm temperatures in February and March 2017 (plots are not shown here) caused some error in the forecasts: the streamflow forecasts initialized on February 1st and March 1st are both higher than observed because the climatology is cooler, but the differences are modest. In general, warm temperatures lead to less runoff and vice versa but this appears not to be the primary explanation for the rapid decrease in the two ESP ensemble means through the winter and early spring.

Figure 11 shows the monthly time series plots of precipitation, SWE, runoff and soil moisture change (P-ΔSWE-ET-RO) for the UCRB for both 2017 and climatology from 1981-2010. The precipitation plot (Figure 11a) shows that the UCRB received anomalously high precipitation in January and February (with the highest anomaly in January), but the precipitation later in the forecast period was less than climatology. The direct effect is that in February 2017 there was a large positive SWE anomaly (Fig. 11(b)), but the anomaly decreased thereafter. This explains why the ESP forecast peak was in February.

The RO time series plot in Figure 11(c) is more complicated: RO production was anomalously high in March, April and June, but lower than climatology in May. The question of interest is, where did the snowpack that accumulated in Jan and Feb go? From Figure 11(b), about 5 km3 of SWE melted in February and March. However, precipitation anomalies were in the range (negative) 2-4 km3 for each month from Mar 2017 on. The fact that 2017 ET during the forecast period was close to climatology (plot not shown) suggests that enhanced early season snowmelt supplied water to the soil column, but reduced subsequent precipitation hindered runoff production. Figure 11(d) shows the modeled water balance for the soil column (P-ET-RO-ΔSWE). Figure 11(d) shows that ΔSMs in Mar and Apr are larger than climatology, but not by much. Furthermore, runoff generation (Figure 11c) is above climatology during that period. However, as the precipitation deficit persisted into late spring and summer, SM began to decrease substantially. The RO actually produced was less than the early forecasts (initialized in February and March) because the ESP ensemble mean effectively corresponds to normal precipitation, which is higher than actually occurred from late winter on in 2017. In summary, the sharp reduction in forecasts through late winter and spring appears to be primarily related to negative anomalies in late winter precipitation, with anomalously warm late winter temperatures having a secondary effect.

6.0 Summary and Conclusions

Both long-term (~100 years) trends in streamflow and comparisons of two major drought periods (1953-68 and Millennium) point to ongoing changes in the relative control of precipitation and temperature on the river's runoff. Udall and Overpeck (2017) have argued that a transition is occurring, which is especially evidenced by the different responses of the 1953-68 and ongoing Millennium drought to precipitation and temperature anomalies. We find that, while there is strong evidence for such a transition, the situation is complicated by spatial variations across the sub-

basins that contribute most to both long-term trends and drought variations in the basin, as well as to seasonal differences in temperature and precipitation trends and anomalies. Specifically, we conclude that:

Over the UCRB (which produces about 90% of the entire basin's runoff), the long-term 1916-2014 decreasing trend of annual runoff is -3.4 km3 (or 16.5% over the entire record). The increasing trend in annual temperature averaged over the basin over the same period has been 1.8 °C. When the annual temperature trend is removed, the negative trend in annual runoff becomes -1.6 km3, which suggests that warming caused a little over half (1.8 km3 or 53%) of the annual runoff trend. Four snow-dominated, sub-basins in the northeast part of the basin that in combination account for over half of the UCRB runoff have experienced modest declines in winter precipitation, which account for a substantial part of the UCRB runoff trend (-1.0 km3) that is not attributable to warming. The remainder of the runoff loss (-0.6 km3) is mostly associated with increased winter ET (mainly snow sublimation).

Compared to the 1953-1968 drought, which was caused by a basin-wide precipitation deficit, the Millennium Drought reflects a strong influence of warmer temperatures. The UCRB experienced low streamflow (2.6 km3/yr below average, slightly more severe than the 2.4 km3/yr negative anomaly for 1953-1968) during the Millennium Drought years (2000-2014 in our analysis). The four sub-basins in the northeastern part of the UCRB with the largest negative long-term trends are also the major contributors to Millennium Drought runoff anomalies. The decrease of runoff for the Colorado River near Cameo was especially prominent - it alone accounts for over half of the 2000-2014 runoff anomalies. Although sub-basins with positive runoff anomalies on the south side of Uinta Mountains such as the Duchesne and San Rafael Rivers counteract some

of the deficit, UCRB Millennium Drought runoff was well below normal due primarily to deficits in the northeastern sub-basins.

During the Millennium Drought years, the UCRB's precipitation was close but slightly below the long-term climatology (annual: 107.5 vs 110.8 km3/yr; winter: 54.5 vs 55.8 km3/yr). However, Millennium Drought annual precipitation was higher than the average for 1953-1968 (104.6 km3/yr). Winter precipitation during the Millennium Drought was also higher than in the 1953-68 drought; only summer precipitation was slightly lower. However, the highly productive sub-basins in the northeastern portion of the UCRB had comparatively large winter precipitation deficits during 2000-2014, which resulted in 1.0 km3/yr of the UCRB streamflow total reductions (2.6 km3/yr) that were not attributable to warming. Warming temperatures caused 1.4 km3/yr runoff losses.

By reforecasting the 2017 Apr-Jul natural streamflow at Lees Ferry using the same ESP approach used by CBRFC, we reproduce similar reductions in forecasted runoff to the CBREFC forecasts through the forecast season in what started as a large positive forecast anomaly in Apr-Jul runoff forecast on Jan 1. The Apr-Jul forecast peaked around March 2017 due to abundant SWE in the UCRB induced by high early winter precipitation. Anomalously high snowmelt increased runoff in March and April. However, precipitation from March on continued below normal, and the forecast trended downward in the later months, eventually ending with only modestly above normal Apr-Jul runoffs. Anomalously warm temperatures from late winter on in 2017 aggravated the situation but appear not to be the major cause of the forecast declines, which rather was relatively dry conditions from mid-winter on.

Given the importance of the Colorado River Basin to the rapidly growing U.S. Southwest, others likely will address the causes of the both the long-term and recent changes in CRB runoff,

and the future implications of these findings as the 21st century continues to warm . As we note in section 4.4, our results and conclusions are tightly linked with the forcing dataset, and the model(s) we used. The gridded forcings (for precipitation and temperature, as well as other variables derived from them) propagate through the hydrologic modeling and in turn our diagnosis of runoff changes. We opted to use the Hamlet and Lettenmaier (2005) model forcing data set because it is closely linked to the U.S. Hydroclimatic Network (HCN; Easterling et al., 1996), which is based on a set of stations with relatively complete long-term records that have been corrected for station moves and instrument changes. Nonetheless, the stations included in HCN are predominantly at low elevations, and various avenues (e.g., assimilation of available surface and/or satellite observations into a coupled land/atmosphere model) could be pursued to better represent the role of high elevation climatic changes, which may well not have occurred in concert with changes at lower elevations. We also note in section 4.4 (and explore, via limited experiments with a second model, Noah-MP) the possible sensitivity of our results to the form of the LSM, but much more could be done in this respect. Finally, we note that all of our experiments are off-line, hence we partition CRB runoff changes into those associated with warming temperatures and other factors (mostly precipitation changes); however, these multivariate changes may well be linked in ways that we have not explored. For instance, the modest changes in precipitation that we examined may be coupled with temperature changes, and/or changes in the atmospheric radiative balance, and such linkages certainly are worth exploring.

Chapter 2 Table 1: Naturalized (NFL) and VIC runoff contribution percentages for selected USGS gauges. Values are computed relative to the annual streamflow climatology at the Imperial Dam, AZ-CA. The percentages are relative to long-term averages for water year 1971-2014.

Station name	NFL	VIC
COLORADO RIVER NEAR CAMEO (09095500)	22.8%	21.8%
GUNNISON RIVER NEAR GRAND JUNCTION (09152500)	14.6%	10.3%
SAN JUAN RIVER NEAR BLUFF (09379500)	12.4%	10.9%
GREEN RIVER NEAR GREENDALE (09234500)	12.2%	10.7%
WHITE RIVER NEAR WATSON (09306500)	3.5%	3.5%
DUCHESNE RIVER NEAR RANDLETT (09302000)	4.8%	6.5%
YAMPA RIVER AT DEERLODGE PARK (09260050)	8.0%	9.2%
COLORADO RIVER AT LEES FERRY LEEFY (09380000)	91.8%	91.0%
COLORADO RIVER ABOVE IMPERIAL DAM (09429490)	100%	100%

Chapter 2 Table 2: UCRB annual and seasonal changes in water balance variables over water-years 1916-2014 in km3/yr (km3 for SWE) and percentages relative to the starting value of the fit. P is precipitation, T is temperature in Celsius, ET is evapotranspiration, RO is total runoff and SWE is Apr 1 snow water equivalent. Dashed "D" denotes results from T-detrend simulation. Winter period is Oct-Mar and summer period is Apr-Sep.

	P	T	ET	ET-D	RO	RO-D	SWE	SWE-D
Annual	1.5(1.4%)	1.8	4.2(4.7%)	2.3(2.6%)	-3.4(-16.5%)	-1.6(-7.7%)	-9.1 (-39.0%)	-5.6 (-23.9%)
Winter	-0.1(-0.2%)	1.9	4.9(30.5%)	2.9(18.0%)	0.4(10.4%)	0.4(9.0%)	Na	Na
Summer	1.6(3.0%)	1.7	-0.8(-1.1%)	-0.6(-0.8%)	-3.8(-23.3%)	-1.9 (-11.9%)	Na	Na

Chapter 2 Table 3: Annual average anomalies during the mid-century drought D1 (1953-1968) and Millennium Drought D2 (2000-2014) for CRB, UCRB and LCRB. Long-term climatologies are also provided. Results are relative to the 1916-2014 baseline simulation (Table 2); units are km^3 (except temperature is Celsius). The climatologies are extracted from the baseline simulation. (Table S6 includes the summer and winter anomalies for UCRB).

	P anomaly	P-Climatology	SWE anomaly	SWE-Climatology	ET anomaly	ET-Climatology	RO anomaly	RO-Climatology	T anomaly	T-Climatology
CRB-D1	-8.8	163.8	-2.9	19.3	-6.0	143.0	-2.7	20.7	0.0	8.5
CRB-D2	-11.4		-4.8		-8.7		-2.8		1.0	
UCRB-D1	-6.1	110.8	-2.7	18.7	-3.7	91.9	-2.4	18.9	0.1	5.6
UCRB-D2	-3.2		-4.4		-0.8		-2.6		1.0	
LCRB-D1	-2.7	53.0	-0.2	0.6	-2.3	51.1	-0.3	1.8	-0.2	13.0
LCRB-D2	-8.2		-0.4		-7.9		-0.2		1.0	
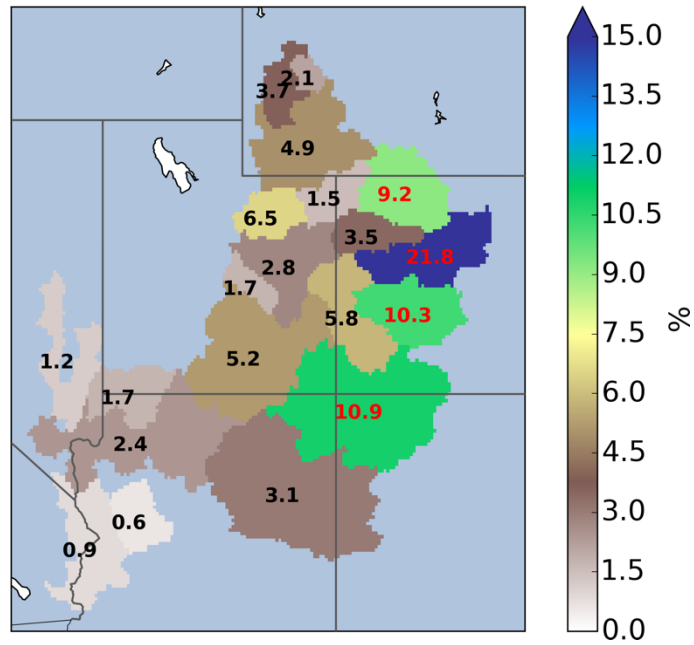
Chapter 2 Figure 1: Percent of total CRB runoff (at Imperial Dam) originating from 20 sub-basins, calculated based on long-term average from VIC simulation for water years 1971-2014. The sub-basins shown in Figure 1 were extracted from a published dataset by Wu et al. (2012).
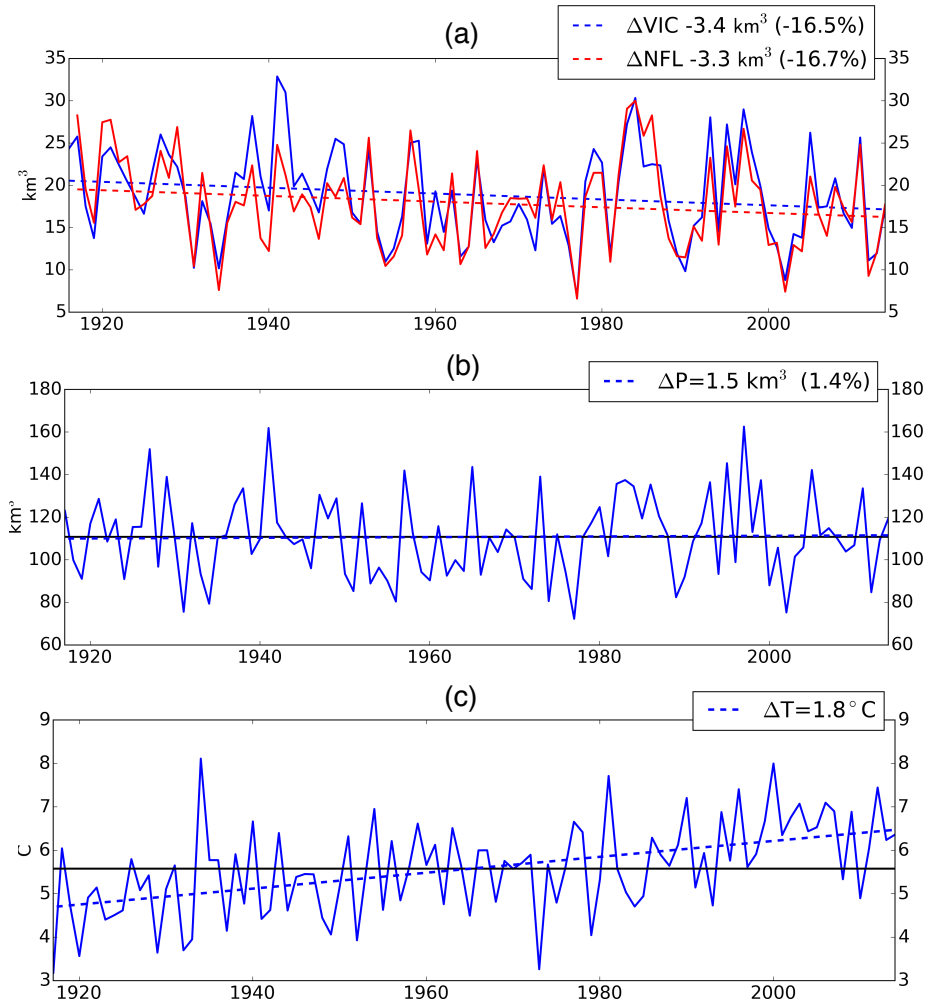
Chapter 2 Figure 2: Annual time series and linear regression trend plots for Colorado River Basin above Lees Ferry; (a) annual (naturalized) runoff, (b) annual precipitation and (c) annual average surface temperature calculated by VIC. Changes are calculated relative to the starting value of the fit. Note that precipitation (b) is from an extended version of the Hamlet and Lettenmaier (2005) data set at $1/16^{th}$ degree spatial resolution while temperature (c) is calculated from VIC and is approximately $0.4°C$ warmer than the Hamlet and Lettenmaier input temperature.
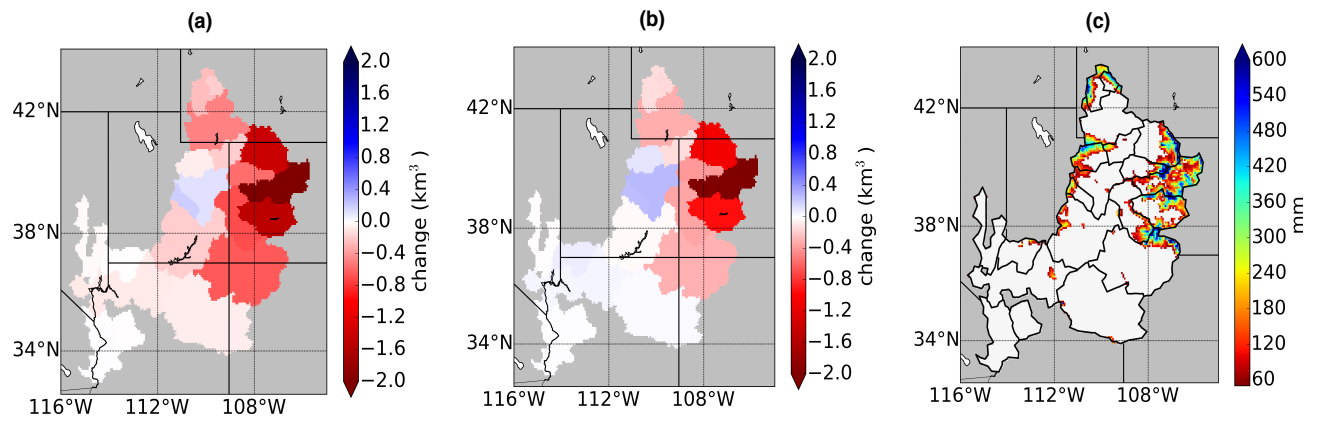
Chapter 2 Figure 3: Spatial plots of Apr 1 SWE trends for (a) baseline simulation; (b) T-detrend simulation over each sub-basin. The changes over 1916-2014 are calculated relative to starting value of the linear regressions; (c) Long-term average Apr 1st SWE
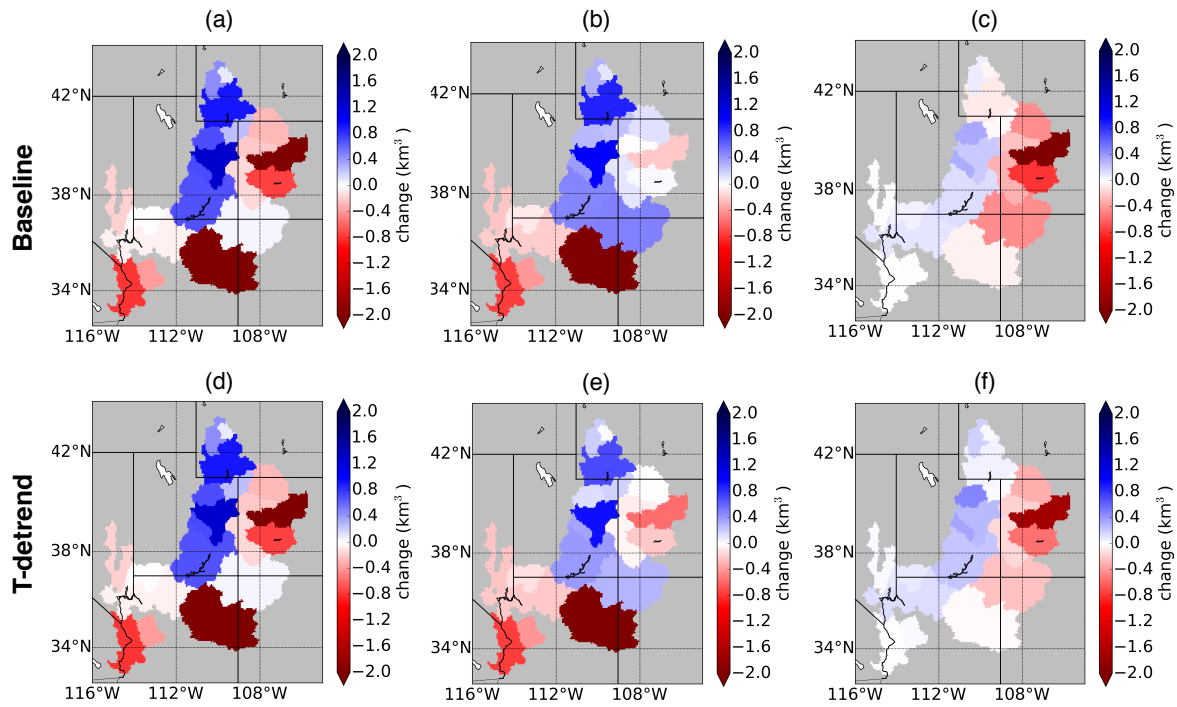
Chapter 2 Figure 4: Spatial changes of (a) annual precipitation from gridded observations; (b) ET; and (c) runoff from baseline VIC simulation over 1916-2014 for CRB above Imperial Dam. Changes are calculated relative to the starting value of linear fits. Panels (d)- (f) are the same as (a)-(c) but variables are extracted from the T-detrend simulation. Panels (a) and (d) are identical.
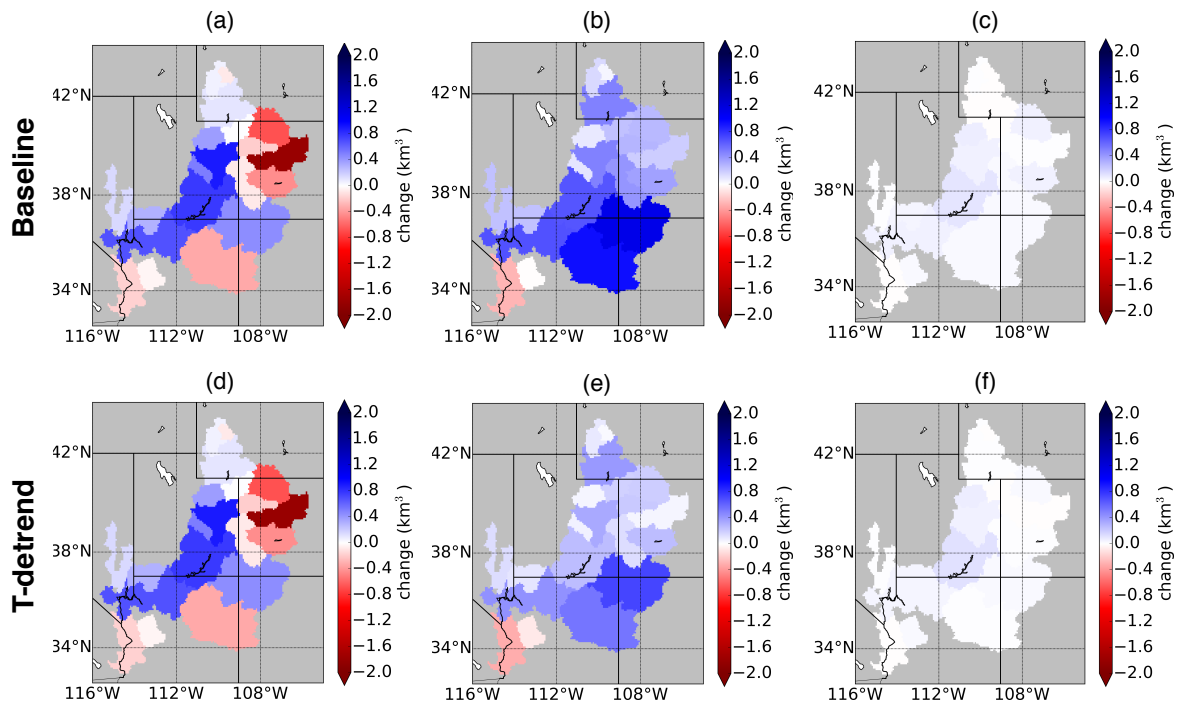
Chapter 2 Figure 5: Same as Figure 4 but for winter (Oct-Mar).
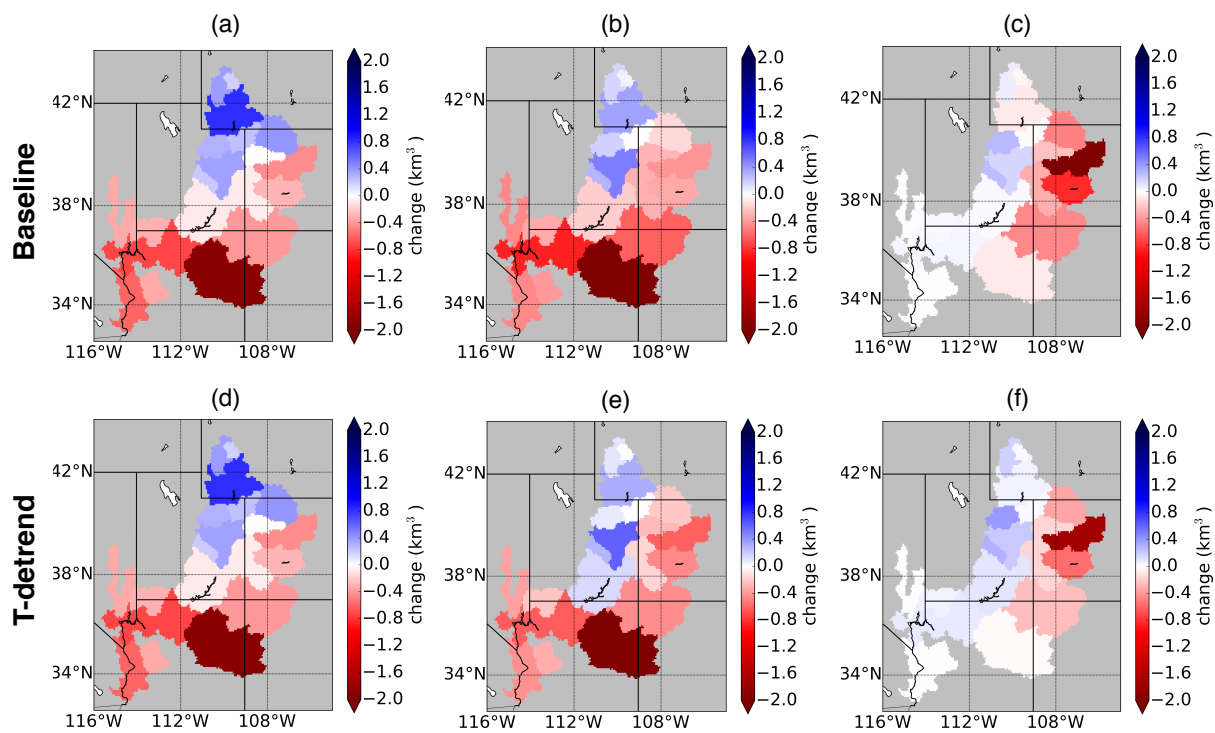
Chapter 2 Figure 6: Same as Figure 4 but for summer (Apr-Sep).
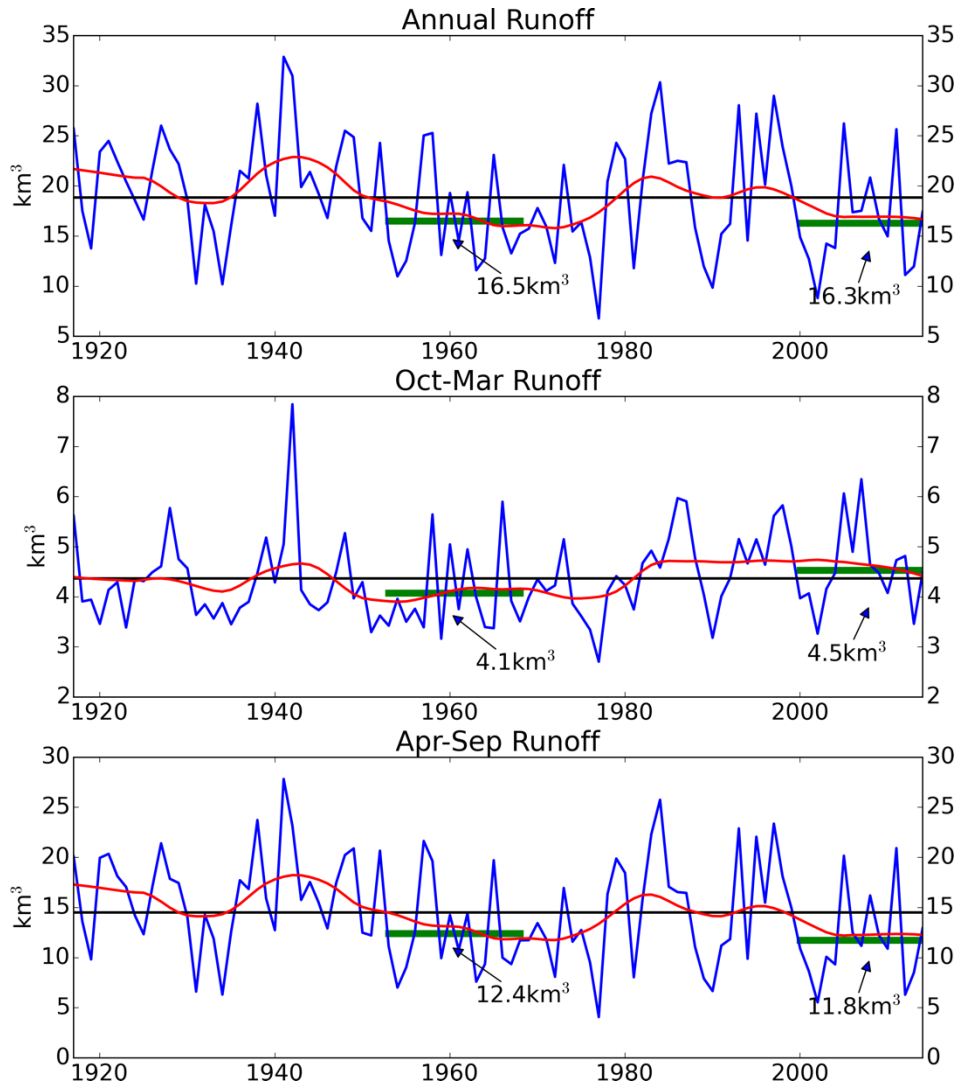
Chapter 2 Figure 7: Time series of VIC simulations of annual runoff (top), winter runoff (middle) and summer runoff (bottom) at Lees Ferry (UCRB). The black horizontal lines are the long-term means, and red lines result from LOWESS filtering of VIC results.
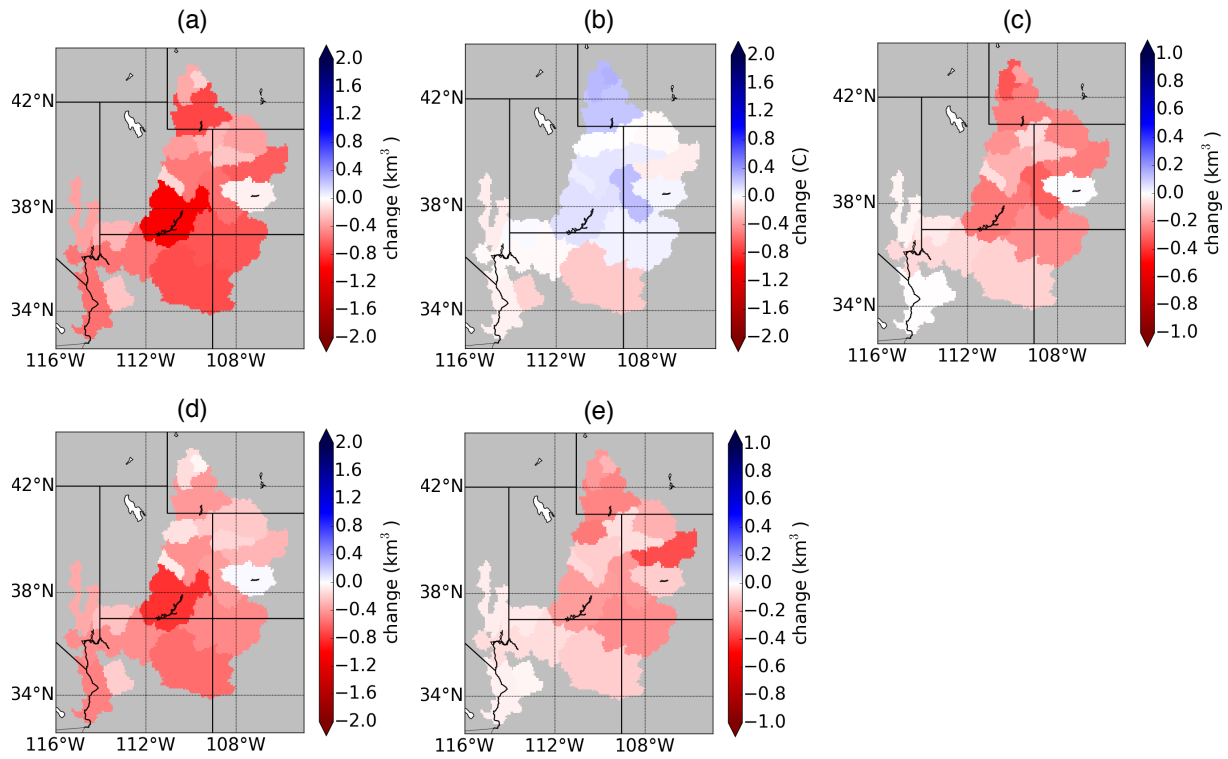
Chapter 2 Figure 8: Average annual anomaly plots for each sub-basin during the drought period 1953-1968. The variables in each panel are: (a) precipitation; (b) temperature; (c) SWE; (d) ET; and (e) runoff (panels (c)-(e) are from VIC simulations).
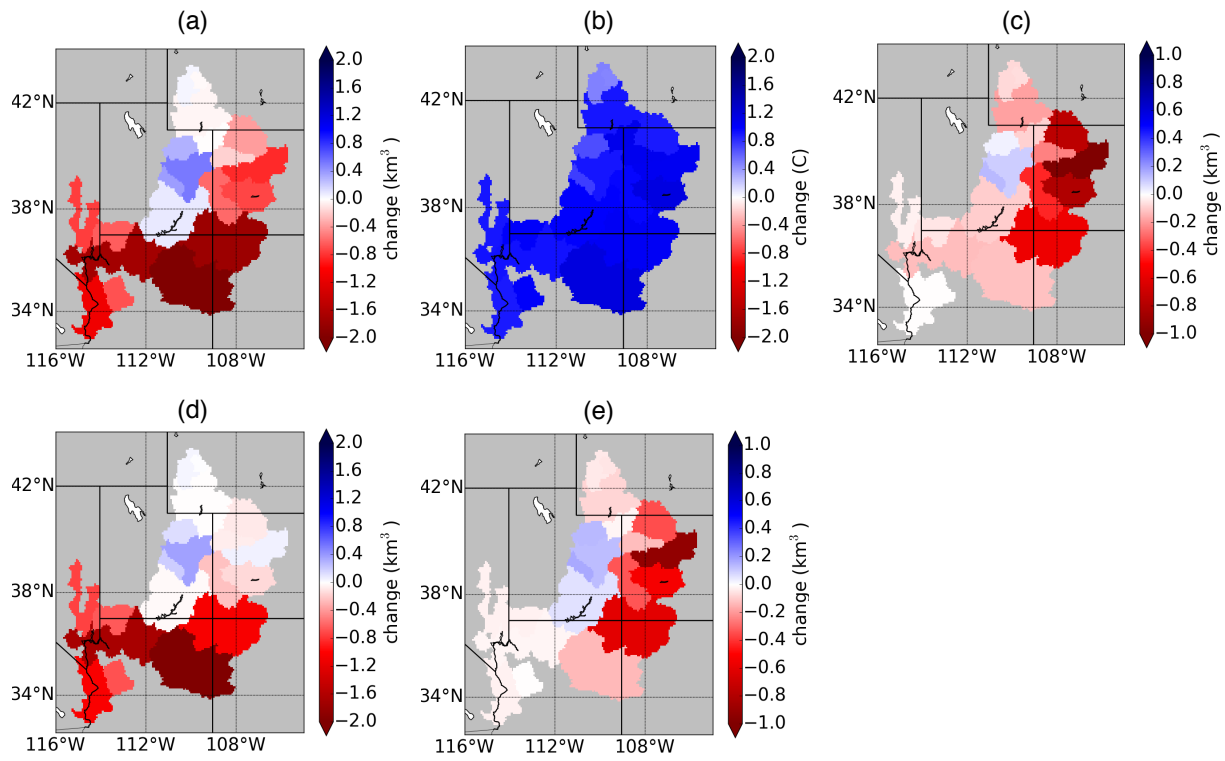
Chapter 2 Figure 9: Same as Figure 8 but for 2000-2014 Millennium Drought.
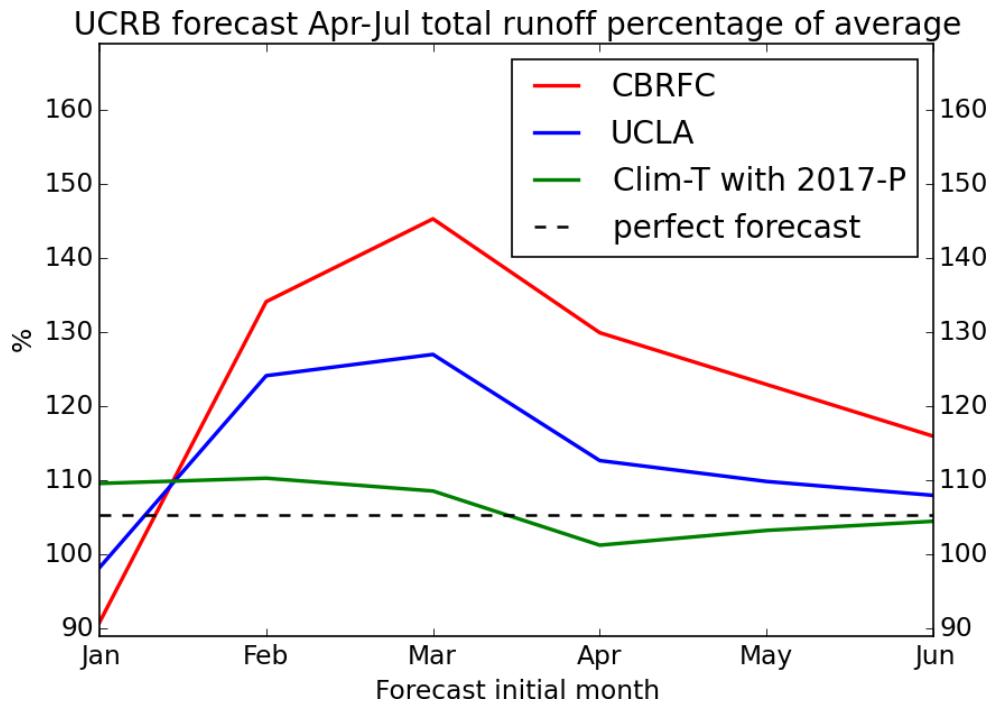
Chapter 2 Figure 10: Apr-Jul 2017 streamflow forecasts at Lees Ferry initialized on the first day of each month expressed as percentages relative to 1981-2010 climatology. Red line represents the official forecasts published by CBRFC; blue line represents equivalent VIC reforecasts; green line is forecast with perfect precipitation forecast and temperature climatology. The horizontal dashed line is from a forecast with perfect precipitation and temperature.
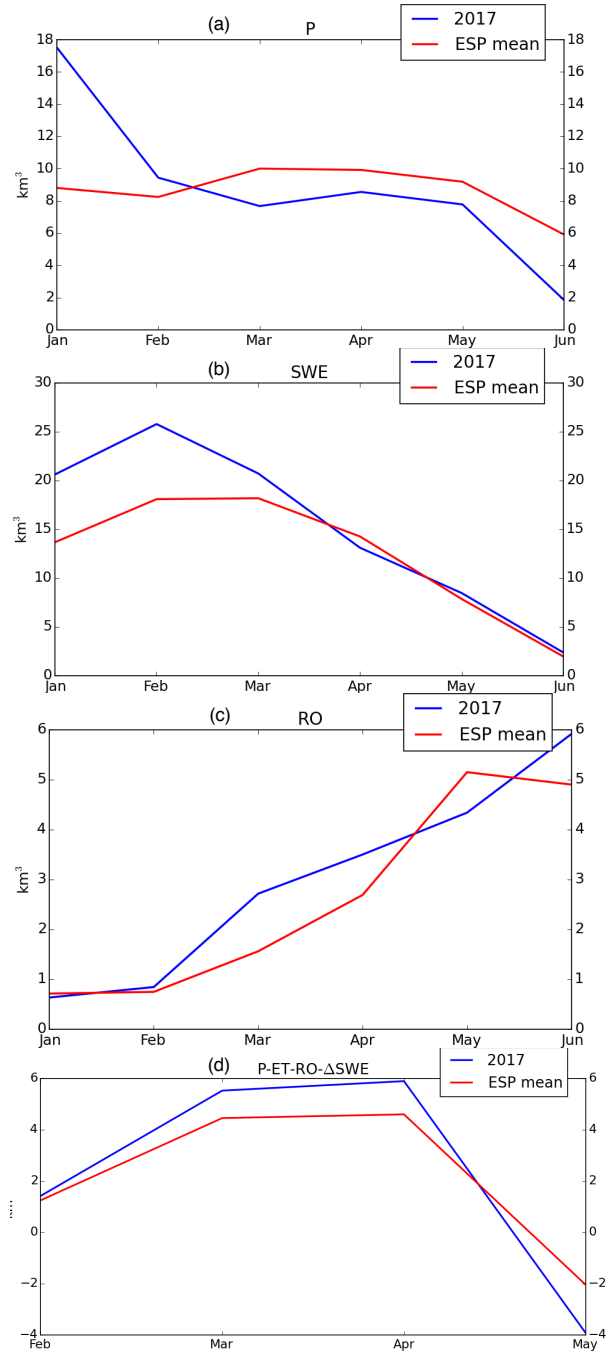
Chapter 2 Figure 11: Monthly time series plots of (a) precipitation, (b) SWE, (c) runoff, and (d) soil moisture change. The blue line is the 2017 forecast; red line is historical climatology.

Chapter 3 Snow Ablation over the Western United States Mountains: Patterns and Controlling Factors

Abstract: Snow accumulation over the mountainous Western U.S. is usually reasonably well predicted by widely-used hydrological models. However, there is a much greater divergence among otherwise "good" models in their simulation of snow ablation process. Here, we explore differences in the performance of VIC, Noah-MP, Catchment and SSiB3 in their ability to reproduce observed snow water equivalent (SWE) during the ablation season at ten SNOTEL stations with over 20 years of record. During the ablation period net radiation generally has stronger correlations with observed melt rates than does air temperature. Average ablation rates tend to be higher (in both model predictions and observations) at stations with large accumulated SWE, where the snowpack remains on average as the downward solar radiation approaches its seasonal peak. Of the four models, VIC and Noah-MP simulate higher net radiation with a larger portion allocated to canopy upward sensible heat (i.e., heat flux from the surface to the overlying air). In Catchment and SSiB3, the sensible heat tends to be downward during the ablation period which enhances the melt energy. If we manually change the surface cover to bare soil in all the models, the magnitude of sensible heat in VIC and Noah-MP decreases dramatically, as in these two models a large portion of the sensible heat flux during the snow season comes from canopy. Catchment predicts decreased sensible heat under the bare soil situation, as it does not include attenuation of wind speed in its calculation and removing the canopy only reduces the surface roughness.

1.0 Introduction

Snow is a dominant aspect of the land surface hydrological cycle of the Western U.S., especially in the headwaters of the major river basins. Snowpacks store precipitation during the cold season and release water via melt during the following warm season, effectively providing a natural reservoir that shifts the timing of peak runoff relative to precipitation by several months. In most Western U.S. river basins, snow is the largest (seasonally varying) water storage component (Mote et al., 2005). Li et al. (2017) found that 53% of the runoff over the Western U.S. originates from melting snowpacks, a number that increases to 70% in the mountainous parts of the region. In relatively dry and heavily-populated Southern California, more than half the water supply is derived from snowmelt from remote mountainous sources (Waliser et al., 2011). As temperatures have warmed in recent decades, snowpack behavior and corresponding hydrological processes have been severely affected. For instance, Mote et al. (2018) report that over 90% of the snow monitoring stations across the Western U.S. with long-term records have shown declines over 1955-2014. As temperatures continue to warm, Rauscher et al. (2008) estimate that snowmelt-driven runoff over the West could occur as much as two months earlier than it has historically.

Despite its importance to surface water hydrology, determining representations of the complicated mechanisms that govern snowpack accumulation and ablation in hydrologic models remain challenging. Given both the scientific challenges and practical implications, Dozier et al. (2016) have argued that estimation of the spatial distribution of SWE over mountainous areas is the most important unsolved issue in snow hydrology. The problem is complicated by the fact that snow depth variability can be caused by a mix of multiple process at various spatial scales (Clark et al., 2011). On the other hand, snow accumulation over the Western U.S. is usually well predicted by the accumulated precipitation occurring during the winter at temperatures below a threshold

(typically slightly greater than 0°C on daily average). For instance, Figure 1(a) shows that SWE estimated using a very simple rule that approximates the seasonal maximum SWE as the accumulation of all precipitation that occurs during the winter season below a fixed (daily average) temperature predicts maximum winter snow accumulations reasonably accurately. Figure 1 also shows that different land surface models reproduce observed SWE maxima that are reasonably close to the observations over the Western U.S. Where there is a much greater divergence among otherwise "good" models is in their predictions of snow ablation. Figure 1(c) shows, when the models are initialized with the observed seasonal SWE maxima, the variations in ablation rates are substantial, and can lead to variations in the predicted date of last SWE that exceed one month.

Here, we explore, in off-line simulations, the ablation season performance of four energy-based snow models that are widely used in macroscale hydrologic models and coupled land-atmosphere models. In particular, we examine their ability to reproduce observed snow ablation rates at selected Snow Telemetry (SNOTEL) sites (snow pillows operated by Natural Resources Conservation Service (NRCS)) across the Western U.S. We examine differences among the snow models (and between models and observations) during the ablation period by analyzing a range of factors that control snow ablation. The remainder of the paper is organized as the follows: section 2 describes the data and models used in the comparisons. We report results in section 3, with interpretation and discussion in section 4. Finally, our conclusions are presented in section 5.

2.0 Data and Methods

2.1 Snow observations and ablation estimate

The USDA Natural Resources Conservation Service (NRCS) Snow Survey and Water Supply Forecasting (SSWSF) Program (https://www.wcc.nrcs.usda.gov/) has a network of more than 750 automated SNOTEL stations in the Western States. Starting in the early 1980s, the

SNOTEL stations began to report daily snow water equivalent (SWE) using snow pillows (which essentially weigh the accumulated snowpack continuously in time), as well as (most sites) daily precipitation, and daily maximum and minimum temperature. We selected 10 SNOTEL stations distributed over the Western United States (Figure 2) whose data are of high quality (missing values less than 5%). These stations form the basis for our analyses. The station names and elevations are given in Table 1.

In order to evaluate snow ablation characteristics, we first need to define the ablation process and melt rates. Dyer and Mote (2007) defined a snow ablation event as a period with a decrease in snow depth between two successive days. They assessed trends in ablation events over North America accordingly. However, our focus does not require such a short temporal scale, as our main objectives are to explore the behavior and the controlling factors during the (entire) snow melt season and to determine the bias and uncertainty among the models in estimating SWE during this period. Therefore, we use a broader definition of the ablation period, which is: for each water year (Oct-Sep), the ablation period is the time from the date of maximum SWE to the last day of snow existence (SWE>0). Further, we extract the 20th-80th-quantile of the ablation period, which we define as the period from the date when 80% of the maximum accumulated SWE remains to the date when 20% of SWE remains. Based on our exploratory analysis, focusing on this central portion of the melt period seems to provide a representation of the ablation process that is free from unusual conditions near the beginning and end of the melt period (e.g., occasional accumulation events early in the melt period, and very warm conditions with partial snow cover late in the melt period). We performed some comparisons (not reported) that showed that our results were not very sensitive to modest changes in our definition of the ablation period. Therefore, In the analyses we report below, our results are based on the 20th-80th-quantile definition unless

stated otherwise. Accordingly, we calculate snow ablation rates for each year the 80th-quantile of SWE minus the 20th-quantile of SWE divided by the number of days between the corresponding dates.

2.2 Land surface models

We examined simulations of SWE using four Land Surface Models (LSMs) : Variable Infiltration Capacity (VIC), Noah Multi-Parameterization (Noah-MP), Catchment, and Simplified SiB version 3 (SSiB3), all of which have been applied previously in numerous snow-related studies (e.g. Tan et al. 2011; Shi et al. 2013; Chen et al. 2014; Newman et al. 2014; Xia et al. 2016; Magand et al. 2013; Xue et al. 2018; Oaida et al. 2015; and Cortés et al. 2016 among many others). The relevant archival references for the snow algorithms in the four models are: VIC (Andreadis et al., 2009); Noah-MP (Niu et al., 2011); Catchment (Stieglitz et al., 2001); and SSiB3 (Sun et al., 1999; Yongkang Xue et al., 2003). The key features of the snow algorithms in each of the model are summarized in Table 2. We also provide brief descriptions of each model below.

VIC is a physical-based, macroscale hydrologic model with an energy-based snow module that explicitly accounts for snow accumulation and ablation in the vegetation canopy (Andreadis et al., 2009; Liang & Lettenmaier, 1994). It represents two layers in the vertical (one for thin snowpacks) – a relatively thin surface layer, and a deeper pack layer. The VIC snow model is capable of simulating sub-grid variability in vegetation canopies and the effects of topography on snow accumulation and ablation via "tile" and elevation band representations, respectively. It also has a parameterization for subgrid redistribution of SWE (e.g., via wind).

Noah-MP has much different physics than the original Noah LSM (Chen & Dudhia, 2001; Ek et al., 2003) to the extent that it essentially is a different model. Regarding the snowpack modeling, the Noah-MP snow model partitions the snowpack into up to three layers according to

the snow depth and snow cover fraction as determined by snow density, snow depth and ground roughness length. Within each grid cell, Noah-MP utilizes a "semi-tile" scheme to calculate the energy balance and solves for the snow temperature over the vegetated and bare fractions separately.

Catchment incorporates a three-layer snow module to account for snowpack growth and ablation (Stieglitz et al., 2001). Catchment determines the net solar radiation flux using estimates of surface albedo; this albedo is calculated separately for the snow-covered and snow-free fractions of the land element, and vegetation "sticking out" of the snowpack modifies the albedo in the snow-covered fraction. The model calculates the heat flow within the snowpack via linear diffusion, with thermal conductivity a function of snow density. Snow can melt in the upper snow layers and, following percolation, can refreeze in lower layers. Snowmelt water that leaves the snowpack either infiltrates the soil or is removed from the system as runoff. Turbulent fluxes into the air (including sublimation) are determined as part of the energy balance calculations performed for the top (~8 cm) snow layer. Catchment redistributes the heat contents and mass of snow into the three layers at every time step. Catchment does not separate downward solar radiation according to vegetated and bare-soil surfaces, i.e. it does not use a two-stream scheme as do other three models. Instead, it would first calculate the tile-average surface albedo (with and without snow) and compute the net solar radiation for the entire tile.

SSiB3 uses the snow-atmosphere-soil transfer (SAST) model of Sun et al. (1999). SAST uses up-to three layers to represent snow in vegetation-free areas and under forest canopies. Each tile is divided into canopy and bare soil partitions according to the vegetation fraction in the same way that SSiB does for snow-free tiles. The snow energy fluxes and surface soil temperature are solved simultaneously to guarantee energy conservation at each time step.

2.3 Forcings and experimental set-up

We extracted daily meteorological observations (daily precipitation and temperature maxima and minima) at the selected SNOTEL sites. We used wind speed from the Livneh data set (Livneh et al., 2013) which is interpolated from the lowest layer of the NCEP/NCAR reanalysis (Kalnay et al., 1996). We applied the Mountain Climate (MTCLIM) algorithms (Hungerford et al., 1989) as incorporated in the VIC model (Bohn et al., 2013) at each station to produce hourly downward solar and longwave radiation, pressure and humidity forcings. Our study period is from 1991 to 2012, which was determined by the availability of the SNOTEL meteorological observations and the temporal coverage of the Livneh dataset.

To evaluate the magnitude and nature of differences in ablation rates among the models, we manually adjusted the SWE predictions for all models to match the SNOTEL annual maxima for each water year (i.e. within every year, when the SNOTEL observation reached its annual maximum, we replaced the simulated SWE on that day with the observed value). We then continued the model simulations through the date of last snow, and repeated the process for the next water year. This procedure allowed us to focus entirely on the models' predictions of snow ablation, without confounding them with differences in snow accumulation.

3.0 Results

3.1 Ablation rates

Figure 3 shows the average ablation rates calculated as described in section 2.1 at each of the 10 SNOTEL sites for the 21-year study period 1992-2012. Overall, the Catchment model produced the best estimates as compared with observations in terms of Mean Absolute Error (MAE). SSiB had slightly higher MAEs than Catchment. VIC and Noah-MP both generally had melt rates that were biased low with one or two exceptions (e.g. site 10 for VIC, sites 7 and 9 for

Noah-MP), where the estimated ablation rates exceeded those from observations by up to 105%. The overall bias across all models is slightly negative (the observations have higher ablation rates than the simulations) although SSiB has generally positive biases. The multi-model ensemble-average yielded melt rates with MAEs that were slightly higher than the best model. The station-averaged errors (model minus observed) in the estimated last day of the ablation period are 9.3 (VIC), 3.6 (Noah-MP), -1.6 (SSiB), -0.1 (Catchment) and 2.8 (model-average) days, respectively.

Table 3 summarizes the climatologies of the 10 SNOTEL sites in terms of average temperature and maximum annual SWE. Considering the ablation rates in Figure 3 and the maximum SWE values in the table, the stations that have the highest SWE accumulations also tend to experience faster melt rates. Figure 4 reports the correlation coefficients between average annual maximum SWE and average ablation rates for the observations and modeled results across all 10 stations. Linear regression relationships are also plotted in the figure. The results from observations are highly correlated (r=0.97) as are the Catchment results. Only VIC is an outlier with a (relatively) small r-value of 0.78. One possible reason to explain the correlations is that the low SWE stations melt their snow before the period of highest available energy (late spring and early summer). As the downward solar radiation increases seasonally, only those stations with higher SWE remain snow covered. The snowpack at these high SWE stations receives more downward shortwave radiation later in the year, and thus tends to have higher ablation rates. Musselman et al. (2017) argue that in a warmer climate, snow ablation rates in the western U.S. will decrease for this reason, which is generally consistent with our results.

3.2 Dependence on temperature and net-radiation

Figure 5 shows the results of linear regressions of the computed ablation rates on the average temperature during the melt season along with the correlation coefficients for observed

and simulated results. Overall, the correlations between ablation and temperature are high, with values from observations ranging from 0.51 to 0.92 with an average of 0.73. The model results also show more or less linear dependences, with only 10% of the r-values across all stations and models less than 0.6). Although there are some deviations for individual models, the model-averaged results in general capture the observed relationships between temperature and ablation rates at each of the SNOTEL sites.

Figure 6 is similar to Figure 5, except with temperature replaced with net radiation. There is no observation-based net-radiation, instead we used the average net radiation from the four LSMs as a surrogate for observations. The correlation coefficients in Figure 6 generally are higher than in Figure 5. In particular, the station average for both observation-based (0.92 in the last subplot of Fig 6) and model-averaged (0.94 in the last subplot of Fig 6) both are substantially higher than those in Figure 5 (0.73 for observed analysis and 0.81 for model average). Statistically, 63.6% of the r-values in Figure 5 are greater than 0.8, and this percentage increases to 78.8% in the Figure 6 net-radiation correlation results. This result should not be surprising as net radiation is the dominant source of melt energy, and temperature appears only in the net longwave radiation component of net radiation (which generally is much smaller than net shortwave during the melt season). This result is consistent with Painter et al. (2018) who show, in the different context of the role of dust on snowmelt rates that radiative forcings are a much more important determinant of snowmelt rates that cause the rising limb of the hydrograph in Upper Colorado's spring runoff than is temperature. One could in fact argue that the only reason that the temperature correlations in Figure 5 are as high as they are is that high temperatures tend to be correlated with clear sky conditions during the melt period, which in turn are associated with high downward solar radiation.

We also performed a similar test of the relationship between wind speed and ablation rate. We found that correlations were weak in most cases. Only three SNOTEL sites have statistically significant ($p<0.05$) correlations between wind speed and ablation rate (Figure S0). At those three sites, there is a (weak) inverse relationship between net-radiation and wind speed, which likely leads to the apparent relationship with wind speed. We do note that the source of our wind speed data is the surface level wind in the NCEP-NCAR reanalysis (Kalnay et al., 1996) which is a coarse scale product (2.5 degrees latitude by longitude) which is unable to capture local scale variations in wind speed. However, a larger factor likely is that wind speed is a determinant of turbulent fluxes (latent and sensible heat) which generally are of opposite sign during the ablation period, and therefore tend to be small in magnitude relative to net radiation. During rain-on-snow events (which do occur occasionally during the ablation period) latent heat flux can be an important contributor to melt (Guan et al., 2016; Moore & Owens, 1984). However, such events occur infrequently enough, and are of small enough magnitude during the melt period, that they appear not to have a major effect on ablation.

4.0 Discussion

4.1 Energy components

To better understand the factors that control snowmelt, we need to identify the sources of melt energy. The surface energy budget equation (which is represented directly in all four of the LSMs), can be expressed as:

$$Q_M = Rn + SH + LH + GH + Q_A,$$

where Q_M is the energy absorbed by the snowpack (melt energy), Rn is the net radiation, SH is the sensible heat flux, LH is the latent heat flux, GH is the ground heat flux, and Q_A is the energy advected to the snowpack by precipitation (the directions of these energy terms in the equation are

all downward). GH and Q_A are usually small during the melt season and we neglect them. We focus here on Rn, SH, LH, and their residual Qr (Rn + SH + LH) which accounts for most of the melt energy.

We show simulated net radiation, sensible heat and latent heat fluxes for each model and station in Figure 7. Net shortwave, net longwave and net downward radiation are shown in Figure 8. In Figure 7, the white circles indicate Qr, the melt energy. VIC and Noah-MP exhibit similar behaviors, with large negative sensible heat during the ablation period (i.e. the surface warms the air) except for VIC at Schofield Pass (site 10). SSiB and Catchment generally have positive sensible heat fluxes, which means that energy is transferred from the air to the surface. Having upward sensible heat flux over snow-covered site in the forest is not unrealistic, as shown by ground observations reported a previous study (Fig.9 in Chen et al. 2014). Of the four models, Noah-MP produces the most net radiation. However, its ablation rate is not the highest, as it also has large negative sensible heat flux. Generally, SSiB has the largest melt energy Qr, and hence it produces the highest ablation rates among the models. However, we also note that there are few exceptions where these relationships among models are reversed, e.g. Noah-MP vs SSiB at site 7 and site 9, which implies that SSiB may be allocating more energy to ground heat flux there.

4.2 Vegetation effects

During the ablation process, if present, the vegetation canopy, can play an important role in energy transfer to the snowpack. Furthermore, each model determines the vegetation cover types on computational surface tiles on their own way using various global data sets etc. Usually (although not always, see below) SNOTEL sites are located in clearings surrounded with at most short vegetation that is covered by snow for most of the ablation season. Each model's vegetation cover mechanism is distinct as is its representation of the interaction between canopy and land

surface and snow on and under vegetation. For example, VIC uses pre-defined sub-tiles to represent different types of canopy cover and the final result is an area-weighted average (Liang et al., 1994). Noah-MP utilizes a dynamic vegetation cover fraction, which is related to the LAI value (Niu et al., 2011). In Catchment, each computational tile is assigned a single vegetation type and the overall surface albedo of the tile is then determined as the weighted average of snow free and snow-covered fractions. Catchment's snow free parameterization is designed to match MODIS climatological mean albedo at the location at any given time. The snow parameterization in Catchment (Stieglitz et al. 2001) uses a 13 mm threshold of SWE to compute the snow-covered fraction within the tile, i.e. if SWE is greater than or equal to 13 mm, the entire tile is assumed to be snow covered. SSiB employs a monthly-varying parameters for vegetation cover fraction, leaf area index, and other vegetation properties dependent on vegetation type (Sellers et al., 1996).

Furthermore, the models have different representations of how much snow can be intercepted by the vegetation canopy and the energetics of snow on and below the canopy. Their representations of the effects of the canopy on absorption and re-radiation of solar radiation, as well as the effects of the canopy on wind, and hence under-canopy turbulent fluxes also vary. Arguably the first consideration (snow interception) is less important during the ablation season than is the second (vegetation effects on under-canopy net radiation and turbulent fluxes.

In order to evaluate the canopy effects and corresponding model behaviors, we performed a parallel set of simulations, whose setup was the same as the baseline described above but with the vegetation cover removed. For comparison purposes, we give the vegetation type of each model and some key vegetation parameters for the baseline simulation in Table 4.

Figure 9 shows the ablation rates that resulted from the no vegetation experiment (note that the melt rates calculated from the observations are identical to the results shown in Figure 3 as

they require no assumptions about vegetation). From Figure 9, we see that without the canopy cover, the ablation rate in Noah-MP increases substantially. VIC's response is similar in direction but the magnitude of the changes is much smaller. Melt rates for both Catchment and SSiB are reduced relative to their baseline runs when the vegetation is removed. Overall, removal of vegetation results in large degradation of Noah-MP's performance relative to observations (MAE increases to 17.0 mm/day from 6.7 mm/day in the baseline experiment). VIC and SSiB have smaller MAEs in the no canopy condition relative to the baseline. The MAE of Catchment increases very slightly in the no vegetation simulation (likely because the baseline simulation assumes only short vegetation; see Table 4). We do note that at some of the sites (Olallie Meadows, Banner Summer, Blue Mountain Spring, and Silver Creek in particular) a review of photos of the SNOTEL sites shows the presence of some vegetation in the vicinity of the snow pillow, i.e., the no vegetation assumption may not be entirely appropriate. In those cases, the no vegetation assumption is best interpreted as an end point for comparison with the vegetated base runs.

To explain the cause and effect of different model behaviors, we need to analyze the energy components in the no vegetation simulations and relate them to the models' own algorithms. Figure 10 shows the energy terms and Figure 11 presents the breakdown of tile-wide net radiation (net shortwave and net longwave) for all models from the no vegetation simulations. The downward net longwave radiation decreases in the no vegetation scenario for all models (24.6% averaged over the four models), as removing the canopy eliminates the contribution of longwave re-radiated from the canopy (which originates as solar radiation absorption). The net shortwave radiation in VIC, Noah-MP and SSiB all decrease in the no vegetation experiment while Catchment shows a slight increase. We explore the causes of Catchment's behavior below.

In Catchment, the overall net shortwave and net longwave radiation consists of two parts: energy from snow-covered and non-snow parts of each tile. When the SWE in Catchment is greater than 13 mm (which is almost always true during our 20-80-quantile ablation period), the model considers the tile to be fully snow covered and applies the snow surface albedo to the entire tile. The simulated net shortwave and net longwave are almost identical when it is fully covered by snow (as Figure 12 shows). However, the vegetation substantially affects Catchment's calculation of sensible heat. Vegetation cover increases the surface roughness and thus decreases the aerodynamic resistance. With no wind attenuation effect, removing the canopy decreases the surface roughness and thus reduces the sensible heat. Therefore, the snowpack receives more melt energy and the ablation period becomes shorter in the no vegetation experiment. The reduced snow season leads to less net radiation during ablation, because the snowpack is gone before most of the seasonal increase in downward solar radiation increases (as Figure 11 and Figure 12 shows).

The Rn, SH and LH terms show different responses to canopy removal as Figure 10 indicates. Compared to the baseline experiment, the overall behaviors of Qr (defined in section 4.1; as indicated by the white marker on the bars) are that VIC and Noah-MP experience increases, whereas SSiB and Catchment decrease, consistent with the ablation rate responses of each model. It is worth noting that all the sensible heat flux terms are positive for each model in the no vegetation experiment, i.e. heat is being transferred from the air to the snowpack in this situation. In the presences of canopy cover, VIC and Noah-MP have negative (upward) sensible heat fluxes (which means the tile is warming the air) as shown in Figure 7. Although VIC and Noah-MP can absorb more energy in the baseline experiments (with vegetation), a large portion of that absorbed energy (41.7% in VIC and 42.5% in Noah-MP) is distributed to heat the air and hence not to melt snow (the behaviors of sensible heat flux in VIC and Noah-MP is further explored below). In the

no vegetation experiment, although the net radiation in VIC and Noah-MP is reduced, the sensible heat switches from negative to positive and the overall effect leads to more melt energy.

VIC and Noah-MP represent wind attenuation effects on under-canopy turbulent heat fluxes. Therefore, it is somewhat counter-intuitive that the sensible heat fluxes in both models decrease substantially in the no vegetation experiment. The reason of this behavior is that the vegetation component of sensible heat dominates total sensible heat flux in both models when vegetation is present (total sensible = sensible from canopy + sensible from ground). Figures S1 and S2 show time-series of the sensible heat components (canopy and ground) during the ablation season for both models. In the no vegetation experiment, the sensible heat from the snow surface does in fact increase in both models relative to under-canopy sensible heat when vegetation is present (the increase is larger in VIC than Noah-MP). But the increase is much too small to cancel the loss of sensible heat from the canopy, which of course isn't present in the no vegetation experiments. When vegetation is present, the trees would absorb energy and transfer much more heat to the air, which is the main contributor of sensible heat. Therefore, the overall magnitude would decrease if we remove canopy cover in VIC and Noah-MP.

All the models follow the general rule that the snow albedo is greater than that of bare soil and vegetation cover. Therefore, if we remove the canopy in the simulation, the net downward shortwave radiation decreases in VIC, Noah-MP and SSiB as Figures 8 and 11 show. The only exception is the Catchment model, which treats the entire tile as fully covered by snow when SWE is greater than 13 mm and employs snow albedo to calculate absorbed solar energy (as we have discussed above).

5 Summary and Conclusions

We employed four widely used energy-based LSM snow models in offline simulations to explore differences in melt season ablation rates at 10 SNOTEL stations across the Western United States. We extracted precipitation and temperature data from in-situ observations at each of the SNOTEL sites. We manually adjusted the maximum annual SWE value each year to match the in-situ observations for the purpose of focusing on differences in model performance during the ablation periods. We assessed the linear dependence of the ablation rate on two major atmospheric factors: temperature and radiation. We also performed a no vegetation scenario to study the effects of vegetation on ablation rates at each of the SNOTEL sites. Based on these experiments, we conclude that:

1) On average, the four LSMs produce ablation rates that match observations at the SNOTEL sites in the baseline experiments plausibly well. The average MAE for all models is 4.3 mm/day (22% of the observed average ablation rate across the 10 stations), ranging from 3.6 mm/day (Catchment) to 6.7 mm/day (Noah-MP). SSiB is the only model that has positive bias (higher ablation rate than observations) in the baseline experiments. The multi-model average of the estimated last day of the ablation period has a bias of about half a week (last day of snow on average 2.8 days later than in observations). In experiments where we removed the canopy cover, the MAE values averaged over models becomes 34% of the observed station-average ablation rate. The MAE of each individual model in the no-vegetation simulations is close to the baseline results: SSiB has a tiny improvement while VIC and Catchment produce slightly higher values. One model (Noah-MP) is an exception; it has a large increase in MAE in the no-vegetation scenario.

2) The modeled ablation rates are highly correlated with accumulated maximum SWE in part because high SWE stations have their ablation periods at a time of year (generally later in spring than low SWE sites) when downward solar radiation, and hence net radiation, is higher. Net radiation is highly correlated with ablation rates (more so than is temperature), which is consistent with other published studies. Wind speed is not a strong predictor of ablation rates during the melting process.

3) The effects of vegetation canopy cover vary substantially across the models. The presence of a vegetation canopy increases the average ablation rates in two models (VIC and Noah-MP), but decreases ablation in SSiB and Catchment. When the canopy is removed, the simulated sensible heat reverses direction and its magnitude decreases substantially (in absolute value) in VIC and Noah-MP. The direction of sensible heat is unchanged in Catchment and SSiB, but the magnitude of the former decreases and of the the latter increases. The difference among models is attributable to changes with removal of vegetation in the fate of absorbed solar energy (due to lower albedo relative to the snowpack) of vegetation, surface albedo representations, parameterizations of the attenuation of wind speed by canopies, and how much of the absorbed radiation is transformed to sensible heat (which warms the air) as contrasted with re-radiated longwave (much of which becomes melt energy). The differences in model parameterizations that lead to these inter-model differences in vegetation effects should be a topic for further development in the modeling community.

Chapter 3 Table 1 Site locations and attributes for the selected SNOTEL sites.

Site #	Station name	Lon	Lat	State	Elevation (m)
1	Olallie Meadows	-121.44	47.37	WA	1228
2	Hand Creek	114.84	48.31	MT	1535
3	Pike Creek	-113.33	48.30	MT	1808
4	Hemlock Butte	-115.63	46.48	ID	1771
5	Banner Summit	-115.23	44.30	ID	2146
6	Blue Mountain Spring	-118.52	44.25	OR	1789
7	Silver Creek	-121.18	42.96	OR	1750
8	Central Sierra Snow Laboratory	-120.37	39.33	CA	2101
9	Leavitt Lake	-119.61	38.28	CA	2194
10	Schofield Pass	-107.05	39.02	CO	3261

Chapter 3 Table 2: Key features of the snow-related physics in the four Land Surface Models.

	VIC	Noah-MP	SSiB	Catchment
Snow albedo decay	Yes	Yes	Yes	Yes
Canopy interception	Liquid and snow	Liquid and snow	Liquid and snow	Liquid and snow
Canopy radiation transfer	Two streams	Two streams	Two streams	Tile average
Max snow layers	2-layer	3-layer	3-layer	3-layer
Canopy attenuation of solar radiation	Yes	No	Yes	Yes
Canopy attenuation of wind	Yes	Yes	Yes	No

Chapter 3 Table 3: Climatology of average April-July daily temperature (T), annual maximum SWE and average temperature during ablation as defined in section 2.1 at selected stations over 1992-2012.

#	Station name	Avg Apr-Jul T (˚C)	Avg SWE (mm)	Avg T during melt period (˚C)
1	Olallie Meadows	7.3	1362.4	8.1
2	Hand Creek	8.7	272.8	4.6
3	Pike Creek	8.1	572.9	7.0
4	Hemlock Butte	8.8	1117.8	9.0
5	Banner Summit	7.4	614.6	6.5
6	Blue Mountain Spring	8.9	396.7	5.0
7	Silver Creek	9.7	293.9	3.8
8	Central Sierra Snow Laboratory	9.0	973.0	5.1
9	Leavitt Lake	6.1	314.6	8.2
10	Schofield Pass	5.4	910.9	6.5

Chapter 3 Table 4: Vegetation cover type for the LSMs at the selected SNOTEL sites. The corresponding LAI are reported in Table S1.

#	Station name	VIC	Noah-MP	SSiB	Catchment
1	Olallie Meadows	Evergreen needleleaf and mixed cover	Evergreen Needleleaf	Needleleaf with ground cover	Ground cover
2	Hand Creek	Evergreen needleleaf and woodland	Evergreen Needleleaf	Evergreen needleleaf	Ground cover
3	Pike Creek	Evergreen needleleaf	Evergreen Needleleaf	Evergreen needleleaf	Ground cover
4	Hemlock Butte	Evergreen needleleaf and mixed cover	Evergreen Needleleaf	Needleleaf with ground cover	Ground cover
5	Banner Summit	Evergreen needleleaf, woodland and grasslands	Evergreen Needleleaf	Broadleaf shrubs with ground cover	Ground cover
6	Blue Mountain Spring	Woodland	Evergreen Needleleaf	Broadleaf shrubs with ground cover	Ground cover
7	Silver Creek	Woodland	Evergreen Needleleaf	Broadleaf shrubs with ground cover	Ground cover
8	Css Lab	Evergreen needleleaf, woodland and grasslands	Evergreen Needleleaf	Broadleaf shrubs with ground cover	Ground cover
9	Leavitt Lake	Evergreen needleleaf, woodland and grasslands	Evergreen Needleleaf	Dwarf trees with ground cover	Ground cover
10	Schofield Pass	Woodland and Grasslands	Mixed forest	Grassland	Ground cover
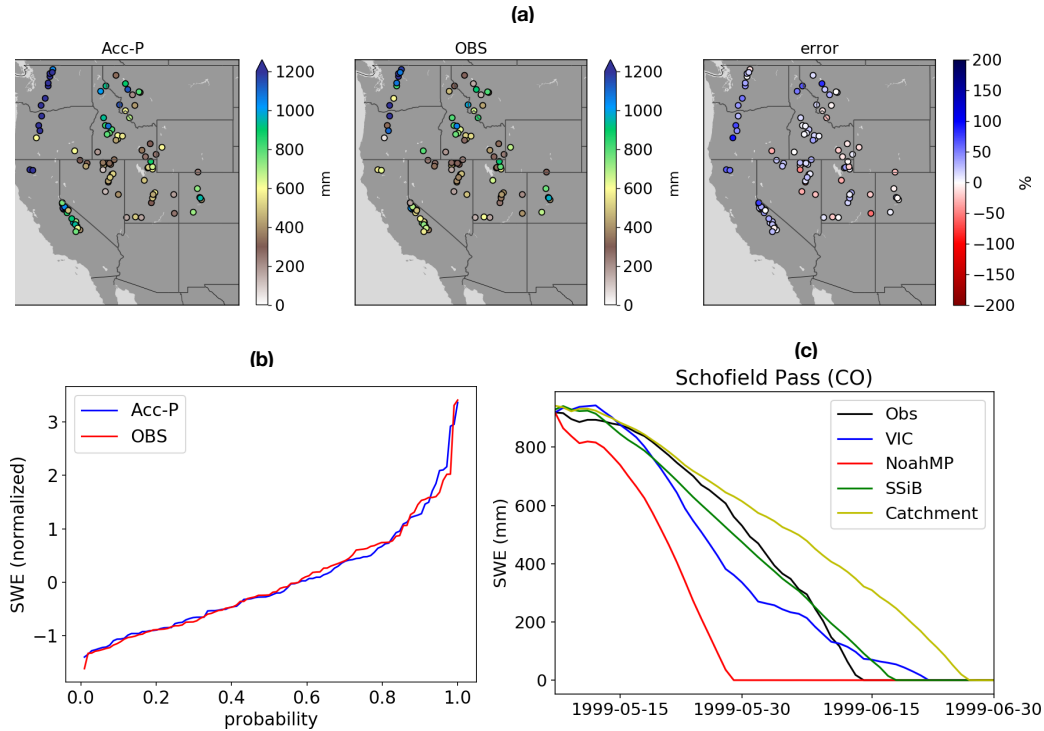
Chapter 3 Figure 1: (a) Climatology of annual maximum SWE estimated by accumulated precipitation (Acc-P), observations (OBS) and error percentage over 1986-2005 averaged over ~100 SNOTEL stations. (b) Empirical cumulative probability curves for annual maximum SWE from observations (OBS) and accumulated precipitation (ACC-P) over all the stations in (a). Both the red and blue lines are normalized. (c) Observed and simulated SWE time-series plot for Schofield Pass, CO. The models are initialized with the observed annual maximum SWE.

Chapter 3 Figure 2: Selected NRCS SNOTEL stations over the Western U.S. The names and index numbers correspond to the information given in Table 1.
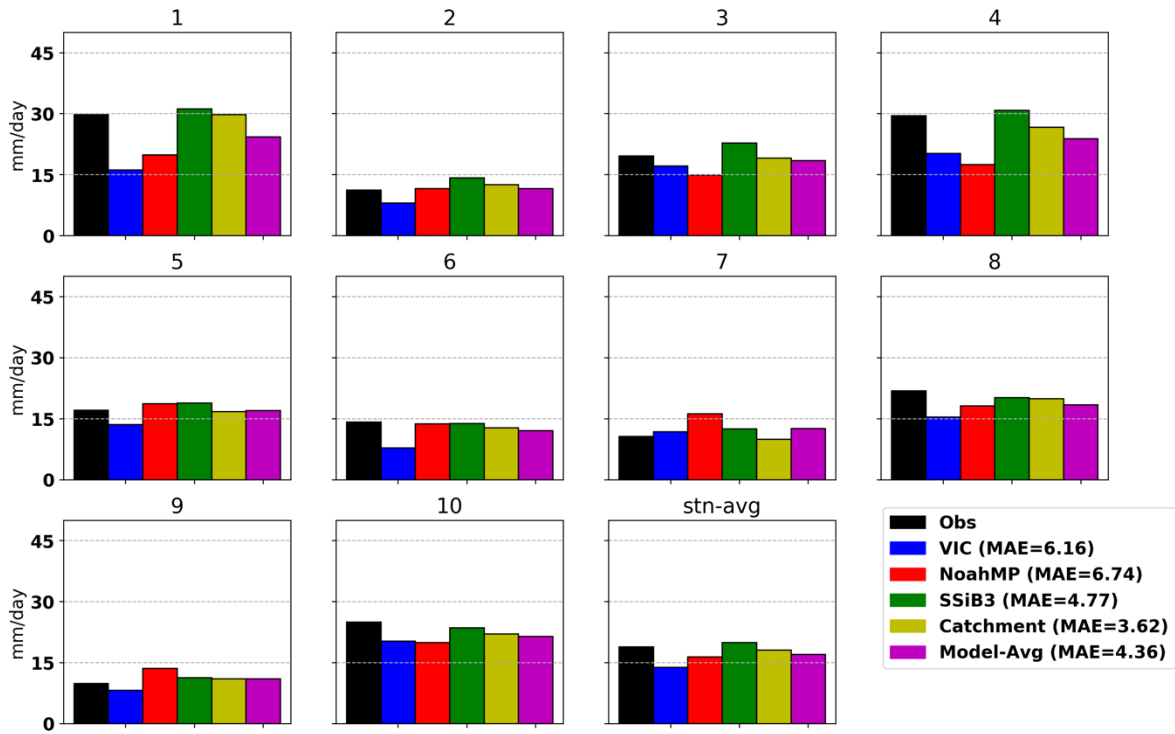
Chapter 3 Figure 3: Snow ablation rates at the 10 SNOTEL sites averaged over study period.

Index numbers correspond to Table 1; "stn-avg" is the mean over all stations.
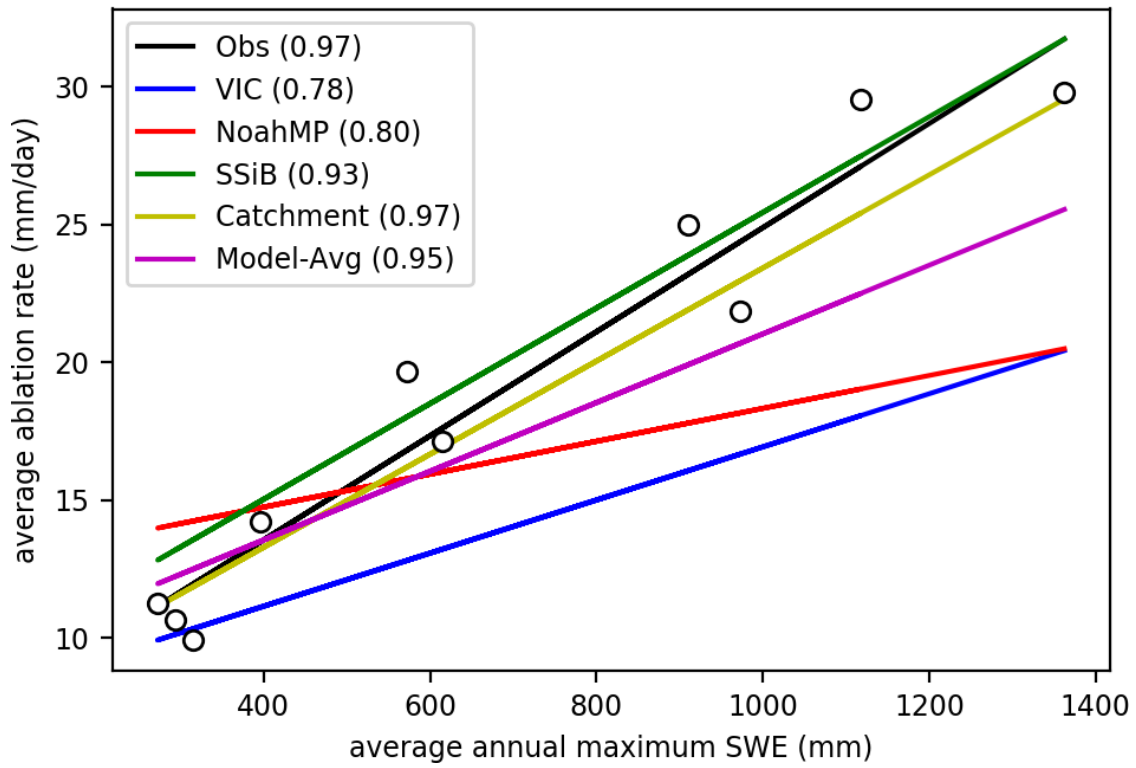
Chapter 3 Figure 4: Linear regressions between annual maximum SWE climatology and average melt rates over the 10 sites. The legend provides the correlation coefficients. The circles are the mean observed melt rate vs mean observed SWE.
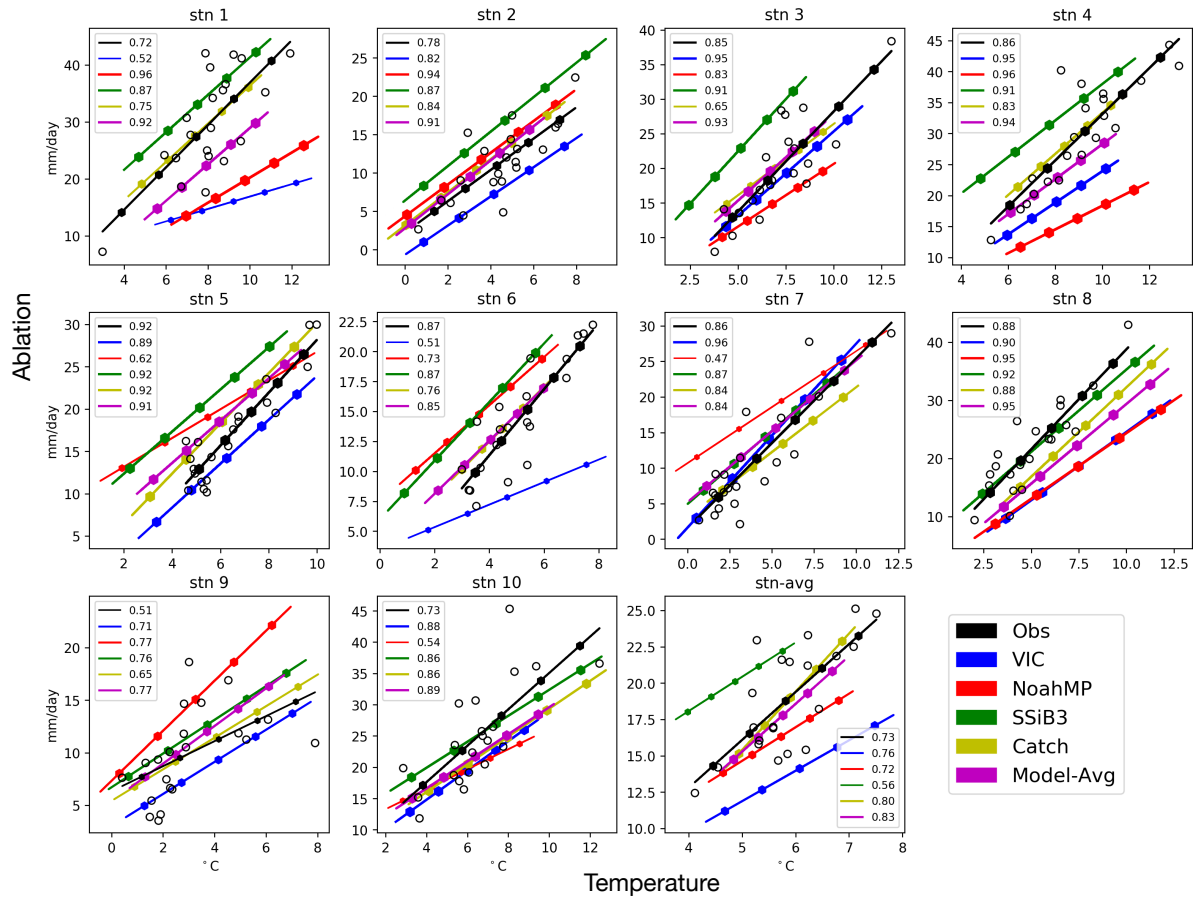
Chapter 3 Figure 5: Linear regressions of melt rate against average temperature during the melt period across all stations for both observed and simulated data (correlation coefficients are given in the legend). The black circles are the observed ablation rates. The ablation units are mm/day (temperature in °C). Larger plot symbols indicate higher r-values.
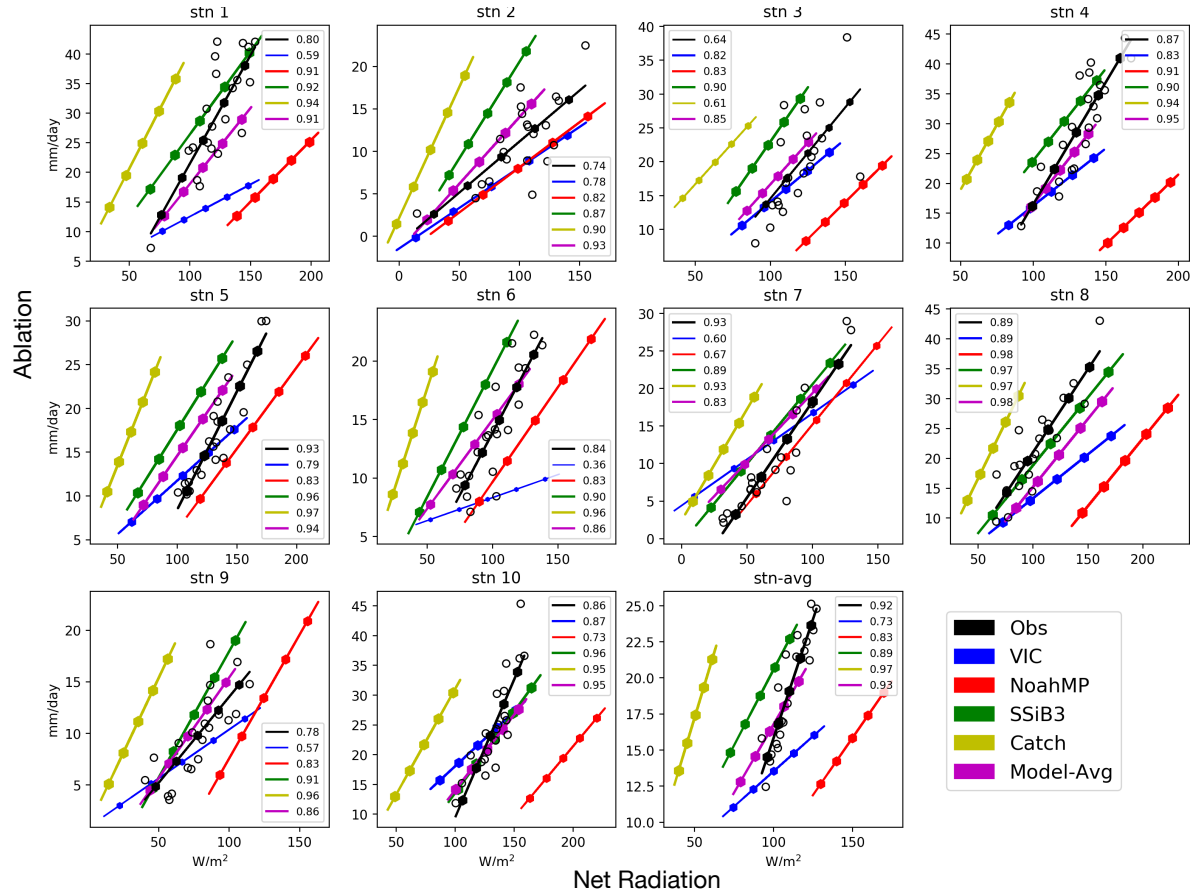
Chapter 3 Figure 6: Same as Figure 5 but the temperature is replaced by net radiation. For the 'Obs' curves we use model-averaged net radiation as a surrogate for observations. Net radiation units are W/m².
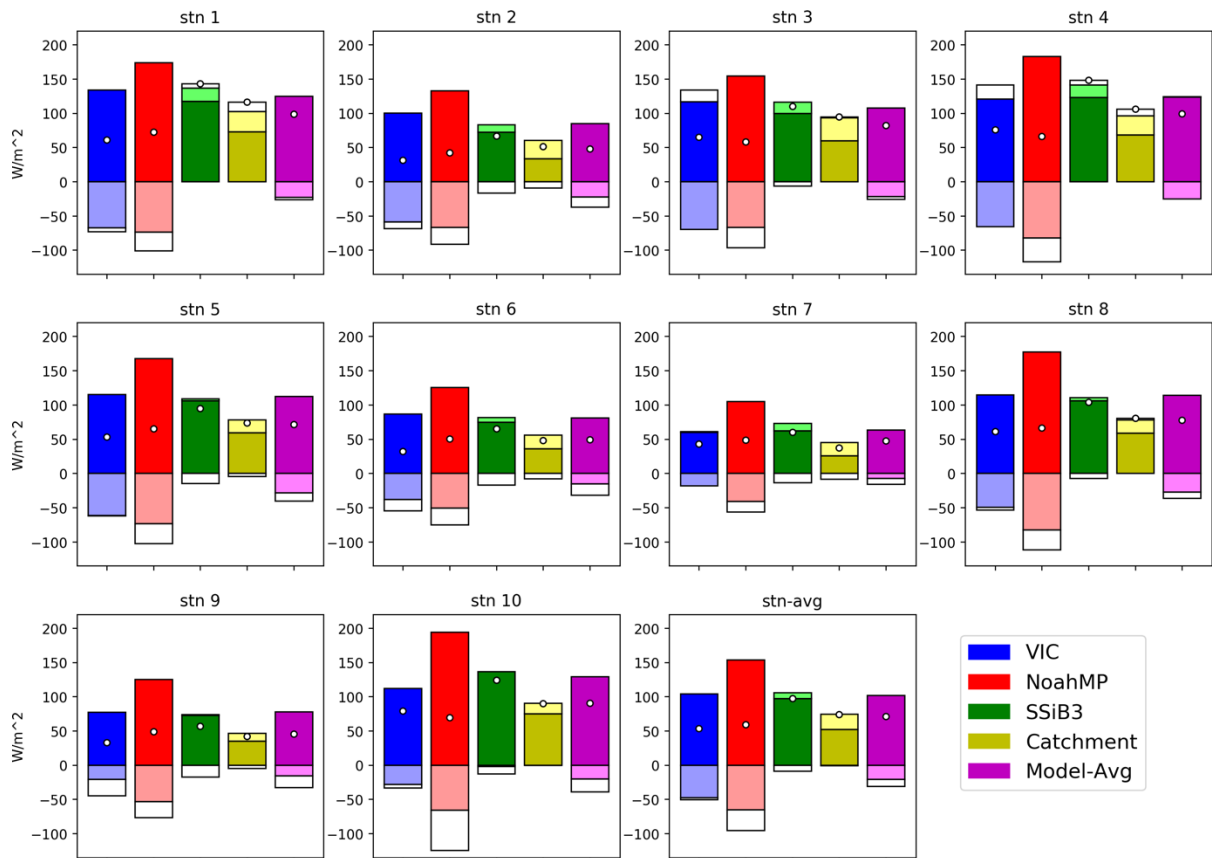
Chapter 3 Figure 7: Energy components for each of the 10 SNOTEL stations. The deep colored bars indicate net radiation (Rn), the white bars are the latent heat (LH), and the shaded bars are the sensible heat (SH). The white dots indicate the energy difference term, Qr (Rn-LH-SH). All nits are W/m^2.
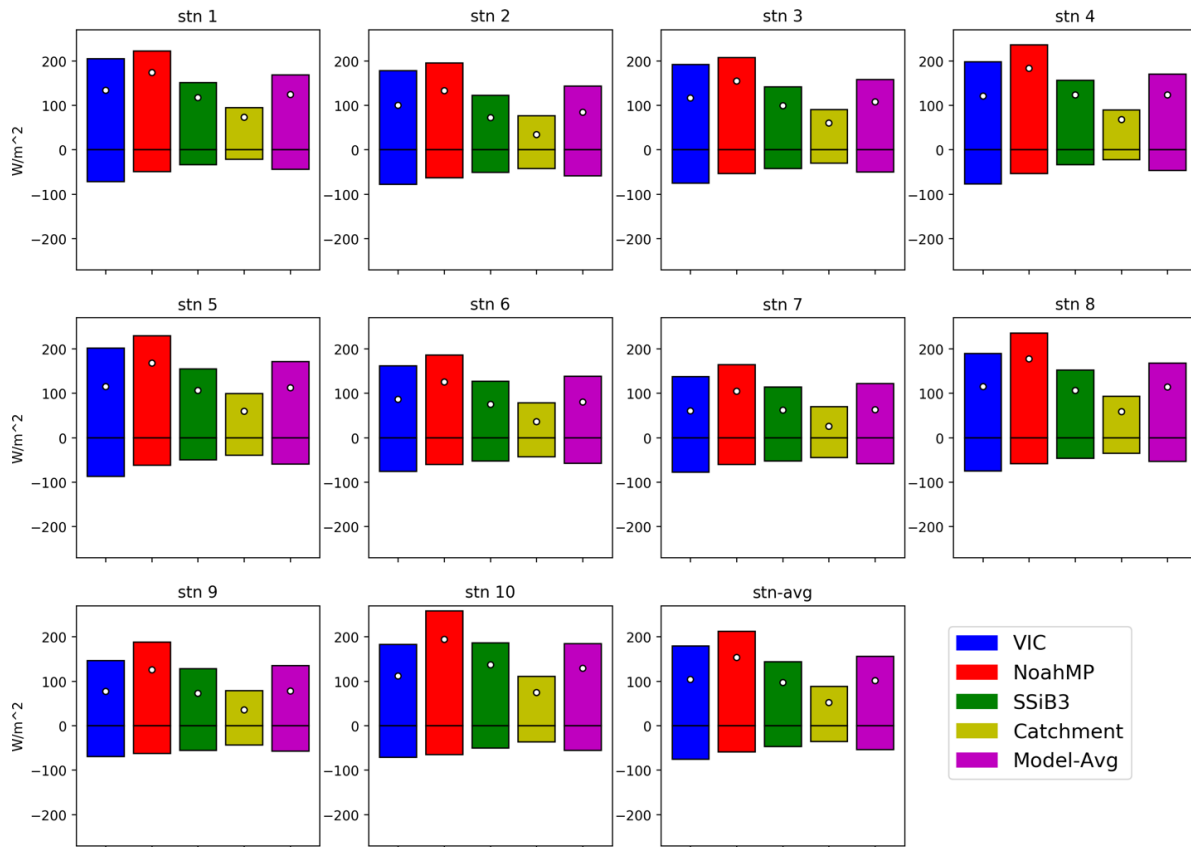
Chapter 3 Figure 8: Tile-wide downward net shortwave (positive) and net longwave (negative) radiation in W/m^2 over all the SNTOEL sites. White circles indicate the net radiation (i.e. net shortwave minus net longwave) term.
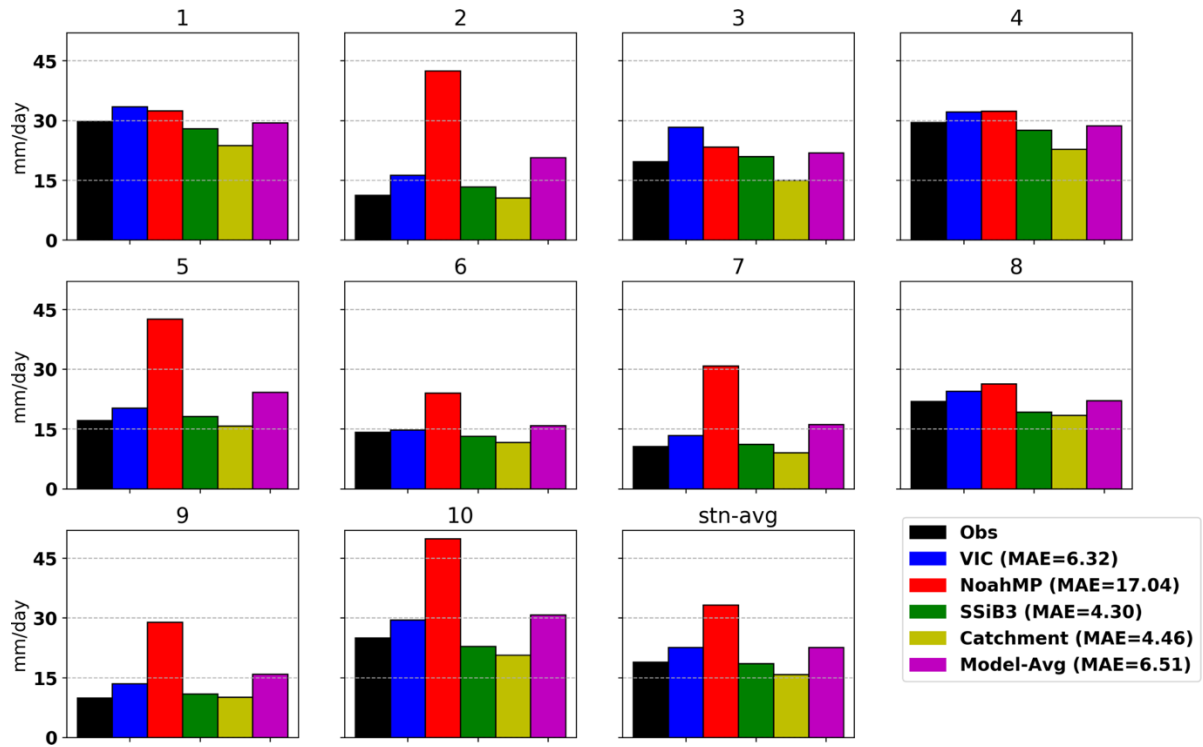
Chapter 3 Figure 9: same as Figure 3 but for the no vegetation simulation.
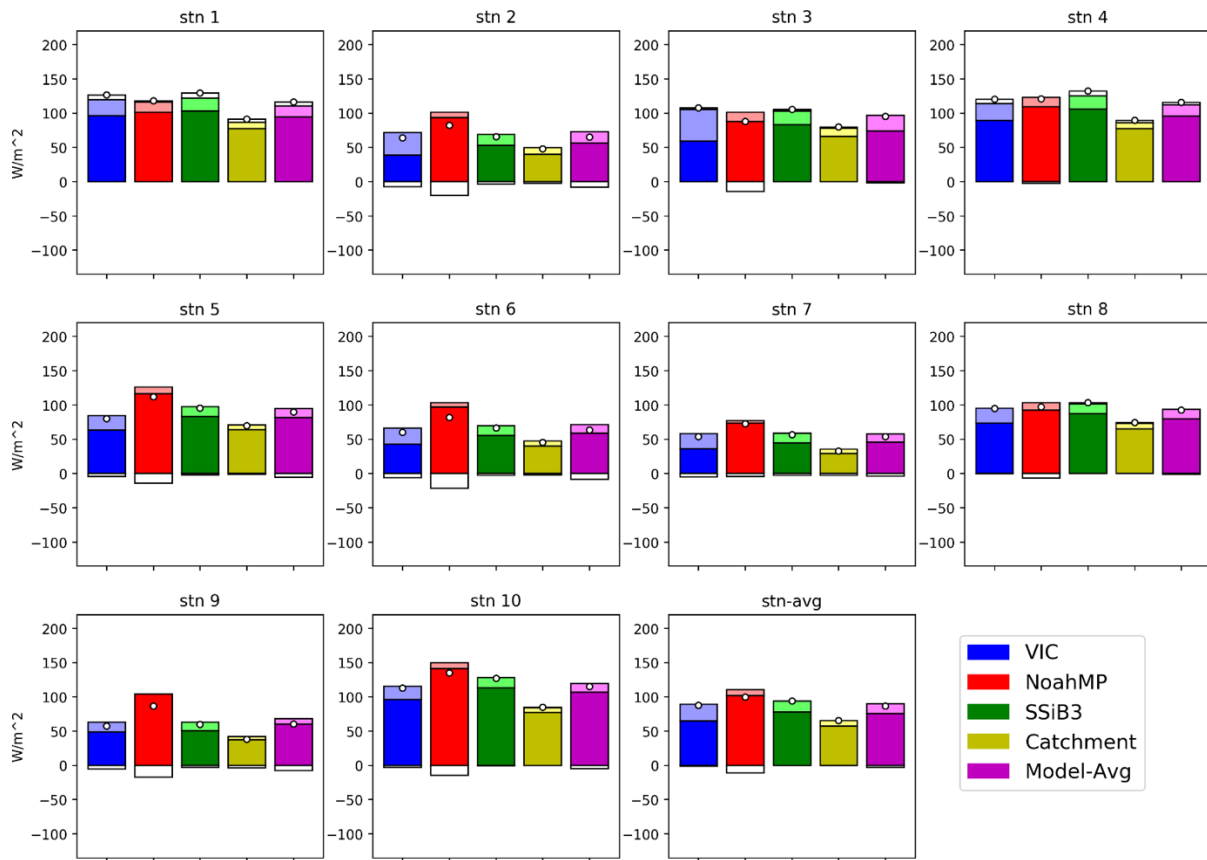
Chapter 3 Figure 10: same as Figure 7 but for the no vegetation simulation.
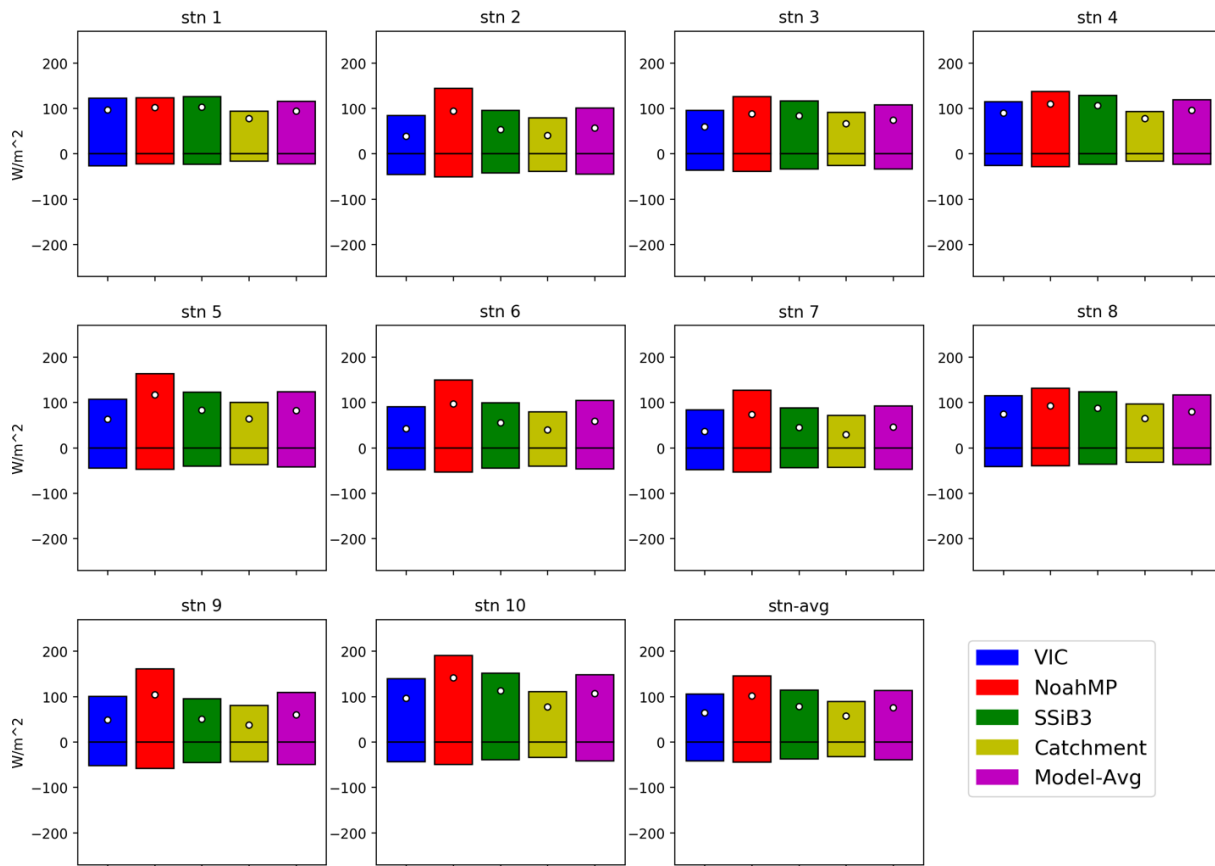
Chapter 3 Figure 11: same as Figure 8 but for no vegetation simulation.
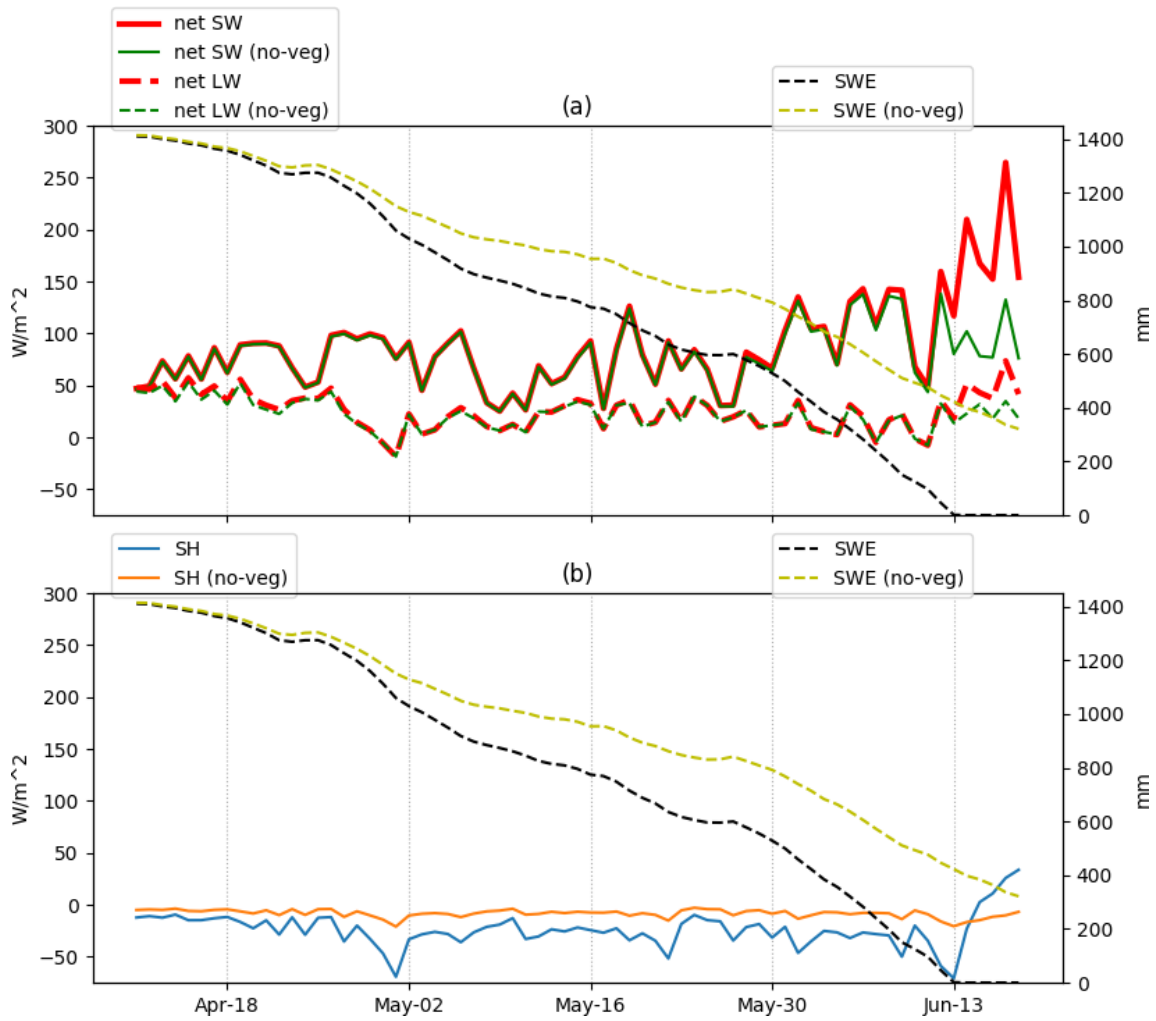
Chapter 3 Figure 12:Time-series plot of (a) net shortwave (net SW), net longwave (net LW) and

(b) sensible heat (SH) at Olallie Meadows station in 1998 for both baseline and no-veg simulations.

The Snow Water Equivalent (SWE) are plotted on a secondary scale in both panels to indicate the

ablation season.

Chapter 4 Identification of winter storm contributions to snowpack in the Upper Colorado Basin

(This Chapter is going to be submitted to *Water Resources Research*)

Abstract: We used the Variable Infiltration Capacity (VIC) macroscale hydrology model to reconstruct daily snowpack records in the Upper Colorado River Basin headwaters for the 67-water-year period 1949-2015 with focus on the accumulation season. We applied a snowfall-based storm identification method to the reconstructed data to attribute the sources of the accumulated snow as either Atmospheric River (AR) (based on an AR catalog) and non-AR. Over our study period, we find that there are on average 37.4 days during which snow accumulates each year, consisting of an average of 16.2 storms per water year. These storms account for an average of 78.2% of annual peak SWE. This number is higher (86.1%) in wet years and lower (70.3%) during dry years. 69% of all storms on average are AR-related they contribute 56.3% of the annual snowpack peak. Although there are no significant basin-wide trends in AR-storm days or storm days per year over our study period, we found that there were parts of the basin (mostly in the middle latitudes) with significant upward trends in the contributions of AR-days and storms to accumulated Snow Water Equivalent (SWE) in the mid-latitudes of the basin.

1. Introduction

The Colorado River is the largest river in the U.S. Southwest, and the region's most important surface water source. Although the area of the entire Colorado River Basin (CRB) is approximately 637,000 km^2, more than 90% of the natural streamflow is generated in the Upper Colorado Basin (UCRB) above Lees Ferry, AZ. The river is highly influenced by snowpack in the Rocky Mountain headwaters sub-basins, which accounts for over 70% of the river's annual flow (Li et al., 2017). The Colorado River is one of the most heavily regulated rivers in the world, owing to municipal and agricultural water demands in the Lower Basin (below Lees Ferry) where some 13,000 km^2 of agricultural lands are irrigated with river water (Cohen et al., 2013), and to the transfer of an additional ~20% of the river's main flow to California for agricultural and urban water supply. The ability of the river to meet these water demands is aided by two large reservoirs, Lakes Powell and Mead, which have a combined storage capacity in excess of four times the mean flow at Lees Ferry. Given the exceptionally high use of the river's water and the need to efficiently manage it in the face of a warming climate, better understanding of the hydrological behavior and patterns within the basin are of great interest both to the scientific and water management communities.

Despite the significance of the snowpack in the UCRB headwaters, the long-term climatology of winter storm contributions to the snowpack have not been carefully explored. It is known that differences in climatic conditions strongly affect the snowpack variability over the mountainous parts of the UCRB (Trujillo & Molotch, 2014). Snow observations come mostly from the NRCS SNOTEL Snow Water Equivalent (SWE) network with about 80 stations across the UCRB mostly in operation since the 1980s and 1990s) and predecessor manual snow course observations. Some previous studies have attempted to reconstruct the snowpack in the basin with

a variety of data sources and tools. Schneider & Molotch (2016) used SNOTEL SWE data combined with Moderate Resolution Imaging Spectroradiometer (MODIS) satellite snow areal extent imagery to improve the real-time snowpack estimate in the Colorado River Basin. Timilsena & Piechota (2008) analyzed tree-ring chronologies for the period 1500-1980, and reconstructed SWE at a set of snow course sites in the UCRB. Several model-based experiments have also reconstructed snowpack over the UCRB. Barlage et al. (2010) improved the snow simulation in the Noah land surface model (Ek et al., 2003) and reported improved performance of the updated model's ability to simulate the magnitude and timing of seasonal maximum SWE over the UCRB headwaters. Ikeda et al. (2010) evaluated seasonal variations in UCRB snowpack using the Weather Research and Forecasting (WRF) regional climate model. The implications of future warming over the UCRB, including snowpack changes, were studied using WRF by (Rasmussen et al., 2011). Chen et al. (2014) employed several well-known hydrological models to simulate SWE over the UCRB. However, none of the previous published work, however, has evaluated the contribution of winter storms (and in particular, Atmospheric Rivers) to SWE in the UCRB.

In contrast, several recent studies have evaluated the characteristics of storms that contribute to snowpacks in the Sierra-Nevada (Dettinger, 2016; Eldardiry et al., 2019; Huning & Margulis, 2017; Margulis et al., 2016). These studies are potentially relevant to the UCRB as well, notwithstanding that there are important differences in winter storm patterns in the two regions. California winter precipitation is highly dependent on large storms, as the wettest 5% of precipitating days contribute around 1/3 of the total precipitation (Dettinger, 2016). Huning & Margulis (2017a) analyzed a high-resolution reanalysis SWE dataset (Margulis et al., 2015, 2016) for the Sierra Nevada and found that more than half of the snowpack in the region is from less than 3 large storms. They defined snowstorms as periods during which basin-wide SWE accumulates

(grid cells at 90-m resolution higher than 75th percentile of the elevation distribution show positive SWE changes) with increases greater than 1% of the total annual maximum SWE (ΔSWE>1%). They found that at least 50% of the accumulated snow (averaged over the Sierra Nevada) comes from no more than three large storms. Eldardiry et al., (2019) used WRF reconstructions of hydroclimatic variables along the Pacific Coast of the western U.S. and found that high positive net snow accumulation during winter accompany AR events at the same time.

Here, we utilize the physically-based, semi-distributed Variable Infiltration Capacity (VIC) hydrological model forced with the Livneh dataset to reconstruct SWE over the headwaters of the UCRB (Fig.1) for water years 1949-2014 (hereafter any reference to years implies water years unless stated otherwise). We then use the simulated SWE data to identify storms and assess their spatial patterns and origins, including storms (and storm days) that are associated with ARs. The remainder of this paper is organized as follows: Section 2 describes the methods and dataset, section 3 presents the major results, more details and discussion are provided in section 4, and summary and conclusions are in section 5.

2. Dataset and Methods

2.1 Hydrologic model and meteorological forcings

We used the Variable Infiltration Capacity (VIC) model (Liang & Lettenmaier, 1994) version 4.2.d as our primary tool to reconstruct snowpack over the UCRB during water years 1949-2014. We focused on the accumulation season, which we define as the period from Oct-1st to the domain-average peak SWE date in each water year, where we defined our domain as all 1/16th degree grid cells in the Colorado River Basin where long-term average Apr-1st SWE exceeded 50 mm (see Figure 1). The VIC model requires gridded meteorological variables as forcings. We used daily gridded records (at 1/16th degree spatial resolution) of precipitation, temperature maximum,

temperature minimum, wind speed from (Livneh et al., 2013). We applied the Mountain Climate (MTCLIM) algorithm (Bohn et al., 2013) to produce downward longwave and shortwave radiation, surface air pressure and humidity.

2.2 Snow observation and AR catalogs

SNOTEL stations (of which there are 86 within our domain) collect daily SWE, air temperature, and precipitation observations dating back to the 1980s (and in some cases 1990s) over the Western U.S. The SWE observations reported by SNOTEL are measured by automated snow pillows which essentially weigh the overlying snow mass. The 86 SNOTEL sites we used all have data availability from 1991 or earlier. We downloaded all the available records for each of the 86 sites for further analysis.

We used the AR catalog of (Guan & Waliser, 2015) which is based on the NCEP-NCAR reanalysis. This global AR catalog covers the 1948-2015 (calendar year) period on a 6-hourly basis. Details of the detection algorithm and the AR catalog can be found in (Guan & Waliser, 2015) and therefore are not discussed here.

2.3 Storm identification

We followed a similar approach of (Huning & Margulis, 2017) which defines storms based on SWE volume to capture the relationship between snowfall and snowpack accumulation across our study domain. In this study, a day would be identified as a storm day if that day experiences a basin-wide SWE increase equal to or greater than 1% of the long-term average of the domain's annual SWE maximum (in our case here, 1% of 270 mm is 2.70 mm). We aggregated consecutive storm days into storms, which accounts for the possibility that snowfall events can be longer than one day (Serreze et al., 2001). One common concern about this identification approach is that it may miss storms that partially cover the domain. In order to address this issue, we tested a cell-

based storm identification criterion, and then defined basin-scale storm days as occurring when more than 30% of the grid cells in our domain had SWE increases larger than the 2.7 mm threshold on the given day. The number of storms and AR-storms identified by the two methods are quite similar (less than 10% difference) as shown in Supplement Figure S1. The consistency of the two methods indicates that we are not missing major storms that cover only part of the domain. Therefore, we use the basin-average threshold in our subsequent analysis. As we apply our identification algorithm, the storm identification threshold is a fixed value taken as the average over the entire domain (2.7 mm/d). We use this criterion to analyze spatial diversity of storm contributions to SWE across the domain, as well as the contributions in drought and wet years.

Apart from identifying major snowfall events, we further classify storms into AR and non-AR categories using the (Guan & Waliser, 2015) catalog. The AR catalog is derived from 6-hourly global atmospheric products from the NCEP/NCAR reanalysis (Kalnay et al., 1996) for calendar years 1948-2015. For each storm identified as described above, we then check whether an AR event occurred in the domain on the same date (as well as one day before and one day after). Following this approach, we classified all storms into AR-related and non-AR types for further evaluation.

3. Results and Discussion

3.1 Snowpack reconstruction verification

We used the VIC model to reconstruct the snowpack over our 67-year study period. The VIC model has been successfully applied in numerous previous studies of hydrological conditions and associated water resources of the Colorado River basin (e.g. T. P. Barnett et al., 2005; Tim P Barnett et al., 2008; Christensen et al., 2004; Christensen & Lettenmaier, 2007; Koster et al., 2010; Vano et al., 2012, 2014; Xiao et al., 2018; and others). More specifically, several previous studies

have used the VIC model to address snow-related issues in the CRB. For instance, Mote et al. (2005, 2018) employed both in-situ measurements and VIC simulations to assess long-term snow declines in the mountainous Western U.S., and they found that the trends estimated by the two approaches were in good agreement in the UCRB. Painter et al., (2010) examined the effects of dust radiative forcing on runoff responses in the UCRB using VIC model simulations. Deems et al., (2013), in a follow-up study, utilized the VIC model to estimate the combined influences of dust and regional warming on snowmelt and streamflow timing in the CRB. Li et al., (2017) performed VIC model simulations over the mountainous Western U.S. and used the results to evaluate the contribution of snowpack to annual streamflow across. In summary, the VIC model has been widely applied in the UCRB and elsewhere in the Western U.S. to reconstruct long-term variations in snowpack, in a manner similar to our application here.

The Livneh et al. (2013) data set we used likewise has been successfully used in a number of previous studies of the UCRB, including those mentioned above as well as (Corringham & Cayan, 2019; Dierauer et al., 2018; Gautam & Mascaro, 2018; Hoerling et al., 2019; McAfee et al., 2019; and Yan et al., 2019). The Livneh data set is observation (and model) based, and hydrologically-consistent. It was derived from precipitation and temperature records from approximately 20,000 NOAA Cooperative Observer (COOP) stations. The Livneh dataset is an update an extension of the Maurer data set (Maurer et al., 2002). The methods used in the Livneh data set are based on (Maurer et al., 2002) but with higher spatial resolution and longer temporal coverage.

Notwithstanding the widespread use of the VIC model and the Livneh et al. data set, we nonetheless evaluated the performance of both the model and data set. We extracted the daily precipitation record during the accumulation season in each water year at all of the SNOTEL sites

(see Figure 1 for locations) as well as the Livneh temperature and wind speed data to run the VIC model. The purpose of utilizing SNOTEL observed precipitation is to reduce the inconsistency between SNOTEL (point) and gridded values, as well as the effects of topographic differences between point observations and interpolated gridded data. Figure 2 shows the cumulative distribution functions (CDFs) of VIC-simulated and observed annual SWE maximum values for all years. The SWE values at each site are normalized (by subtracting the mean and dividing by the standard deviation) before they are incorporated in the "All sites" plot. CDFs of actual SWE peak values at five individual stations, which are geographically distributed in the domain, are also included in Figure 2. The locations of these 5 selected stations are shown in Figure S2 and detailed information is reported in Table S1. The simulated and observed CDFs of normalized SWE annual peak values for all sites are in good agreement. By comparing the CDFs extracted at individual sites, we find that the CDFs of VIC and observations have some inconsistencies, which might be caused by the fact that VIC simulations are for the entire 1/16˚ grid cell and SNOTEL observations are effectively at points. We also compared the time-series of the mean observed and simulated SWE values across the 86 SNOTEL during the accumulation seasons (Figure S3). In general, the simulated results and the observed snow records agree quite well.

3.2 Basin-wide storm contribution

We applied the methods described in Section 2.3 to produce VIC-simulated SWE records using the Livneh forcings to identify storms responsible for SWE increases and the subset of AR-related storms. We show time series plots of individual storm days, number of storms and number of AR-related storms each year in Figure 3. Over the entire study domain, there are on average 37.4 storm days per year. The mean number of storms is 16.2 per year of which 69% (11.2) are AR-related. The average duration of the storms we identified is 2.3 days long.

After identifying the storms in each accumulation season over the entire record, we calculated each storm's contribution to basin peak SWE for that water year. Figure 4 shows the contribution of storms to annual maximum SWE for each year. We also show the contributions from each AR storm in the same figure. We note that we only include storm days within the AR-window (as described in section 2.3) in our calculation of AR-storm contributions. We use the same term "AR-storm" hereafter to denote the storm days within the AR-window. This definition is different from "AR-related" storms, although the difference is rather limited (only 10% of the storm days belonging to an AR-related storm are outside the AR-window). Finally, we calculated the contributions from all days when precipitation yields snow increases (denoted as "all precipitation" hereafter) in the accumulation season (Figure 4). In some cases, the estimates can exceed 100%. This can occur because we compared the accumulated precipitation to annual peak SWE over the entire domain, and some low-elevation areas can experience mid-season melt and sublimation. We find that the average contribution of AR-storms to annual peak snowpack is 63.3% over the entire record, and the average contribution from all storm days is 78.2%, which indicates that a large portion (~80%) of the snowpack in the Upper Colorado mountains originates from moderate to heavy snow storms. Huning & Margulis (2017)used a similar approach to estimate the range of snowstorm contribution in the Sierra-Nevada and found a range of 83%-93%. Compared to the Sierra Nevada, the values are smaller in the UCRB (primarily because the distance from the coast is greater) but nonetheless is still quite high. We also find that about 75% of all individual storm days are AR-related, and they produce 63% of the total maximum snow accumulation. The average contribution of all precipitation days to the grid cell maximum accumulation averaged over all grid cells in our domain is 116.8%, which implies that the excess (16.8% of the SWE maxima) melts (or is sublimated) before the domain's peak SWE occurs.

3.3 Wet, dry, warm and cold years

We selected the 10 most extreme years in each category (wet, dry, warm and cold) and investigated the contributions of storm days in each of these categories to SWE. We defined wet, dry warm, and cold based on the total precipitation amount or average temperature during the accumulation season (beginning from Oct 1st to peak SWE date) averaged over our domain. Table 1 reports the annual average number of storm days, storms and AR-storms in each category. The number of storms and storm days (both AR and non-AR) is higher during wet and cold years compared with dry and warm years. The differences between the number of AR-storms in each of the extreme climatic categories are relatively small, given the fact that only about 10 AR-storms occur per year on climatological average. However, the differences in terms of storm days are larger – 54.9 vs. 23.0 days per year for wet vs dry, and 41.5 vs 36.2 cold vs warm, respectively.

Table 2 gives the percent contributions from storms and precipitation days to the maximum annual SWE for the four climatic categories, as well as the climatology (all years). Comparison of the statistics in wet and dry years suggests that while storms play a more important role in snow accumulation during the wet years, the contribution percentages from all precipitation in wet years are nonetheless lower than for dry years. The reason for this is that the actual amount of accumulated SWE is much smaller during dry years, which makes contribution percentages rise. The SWE losses (difference between total accumulated SWE and the annual maxima) are similar for wet (2.4 km^3/yr) and dry (2.6 km^3/yr) years (but as a percentage of peak SWE, much larger in dry years). These results suggest that during dry years, relatively small snowfall events become more important to the accumulated snowpack in UCRB. Nonetheless, the dominant contribution to SWE is from storms in both wet and dry cases. During dry years, not only are there are fewer storms, but the precipitation amount per storm also is less. The average SWE increase is 0.76 km^3

per storm for dry years and 1.57 km^3 per storm for wet years (reported in Table 2). On the other hand, the contribution percentages of AR-storms, all storms and all precipitation days are all higher in warm years than cold ones, and the accumulated maximum SWE decreases in warm years. The contribution from all precipitation in cold years are lowest as expected, arguably the result of less mid-season SWE loss by melting or sublimation (only 1.8 km^3/yr, 4% of the climatology in cold years, is eliminated during the mid-season). In cold and wet years, snowfall contributes more efficiently to maximum SWE (less midwinter loss) and the contributions from storms are higher (including AR storms). The flip side of that is that in warm and dry years, more of the total snowfall comes from minor events. Overall, 72.7% of all storms' contribution to annual peak SWE is attributed to AR-storms in all years, as high as 76.5% for wet years but still 70.7% in dry years (5th row in Table 2).

Figure 5 shows the same bar plots as Figure 4 with wet and dry years highlighted. We estimated the distribution of the contributions to peak SWE for all the 67-year-long records using Weibull plotting positions (see Figure 5). Based on the plots of the contributions, we notice that generally both AR and non-AR storm contributions tend to be higher in wet years and lower in dry years. For the contribution of all precipitation, the results are somewhat different: the contributions (of storms to peak SWE) tend to be higher in dry years and lower in wet years. The reason for dry years having a higher contribution percentage is that max SWE in those years is small. More mid-season SWE loss in dry years also has some effects, but the main difference between dry and wet years (with respect to mid-season snowpack loss) is not large enough (2.7 vs 2.2 km^3/yr) to be the dominant cause.

Similar to Figure 5, Figure 6 shows the same bars and distributions with warm and cold years highlighted. During warm years, because mid-season SWE loss effect is the largest amongst

the four climatic situations, the average percentages are high for AR-storm, all storm, and all precipitation. On the other hand, both the numbers in Table 2 and the distribution plots in Figure 6 show that storm contributions during cold years are not much smaller than for all years. This suggests that although lower winter temperatures can provide greater snow accumulation (as the last row of Table 2 reports), the percentage contribution from storms is not substantially affected. The major sources of snowpack are still snowfall during storms and thereby determined by precipitation availability.

3.4 Spatial analysis

Although we defined storms as basin-wide events, most storms do not cover the entire domain. Therefore, for all event measures (storms, AR-storms and storm days) we performed an analysis of SWE changes at each grid cell to determine whether that specific grid cell was influenced by the particular event. If the cell's SWE increased by over 0.5 mm after the event (ΔSWE>0.5 mm), we considered that grid cell to have been influenced by the event. Applying this 0.5 mm threshold for all the events, we determined each storm's coverage and number of events that each grid cell experienced. On average, each storm affects 84.9% of the entire domain and each AR-storm covers 85.7% of the domain, which indicates that the AR-storm's scale is very slightly larger. Figure 7 shows the cumulative contribution of (AR-) storms (y-axis) as a function of storm cover fraction (x-axis), i.e. given a certain value, μ, the y-axis reports how much SWE is provided by (AR-) storms that cover less than μ of the domain. Of all the contribution from AR-storms to annual peak SWE, 71.7% is attributable to AR-storms that affect more than 90% of the entire region. The contribution from AR-storms that cover less than 70% of the domain is only 6.2%. The remaining 22.1% (100%-71.7%-6.2%) is contributed by AR-storms that cover between 70-90% of the domain. If we do the same calculation for all storms, we find that storms that cover

at least 90% of the entire basin provide 67.9% of all storms' contribution to the SWE annual maxima. Storms that cover less than 70% of the basin would yield 6.1% of the total contribution, which means the remaining 26.0% is attributed from storms that cover 70-90% of the domain. Overall, the contributions of both AR- and all storms are mainly attributable to events that cover much of the domain.

Figure 8 shows the multi-year average number of AR-storms, storms and storm days on a grid cell basis averaged over the entire record (note that, as in Figure 1, we only consider grid cells with > 50 mm average Apr-1st SWE). We also show sub-basin boundaries for reference (more detailed information about the sub-basin analysis is included in the Supplement). In general, Figure 1 shows that the eastern part of the basin has more storms and storm days on average than the western part of the basin. Furthermore, grid cells with more storm days also have higher snow accumulation (see Figure 1). Notwithstanding the west to east trend, spatial variations in storm statistics across the domain generally are modest.

If we compute ΔSWE for each event divided by the basin average peak SWE for each year, we can form a time series of the contribution of that grid cell to the basin's total snowpack. We do so in Figure 9, which shows the average contributions (over the entire study period) from AR-storms, all storms and all precipitation to basin total snowpack in each grid. The AR-storm (Fig.9 left panel) and storm (Fig.9 middle panel) maps generally show consistent spatial patterns: the highest numbers are in the east and the smallest contributions are the grid cells with lowest SWE climatology (see Figure1). Nonetheless, if we take all precipitation into consideration, the northwestern part of the domain (around 42.5˚N) also makes large contributions to the basin snowpack (Figure 9 right panel). Because the number of storms days in the northwestern part of the basin are smaller than the eastern (Figure 8 right panel), the high contribution in the plot

illustrate that small-scale snowfall events play an important role in that (northwestern part of UCRB) region.

We also extracted average AR-storm, storm and all precipitation contributions for warm and cold years (defined as described in section 3.3), results of which are shown as spatial maps in Figure 10. Figure 11 shows similar information but for wet and dry years. The spatial patterns of AR-storm and all-storm contributions during wet and cold years are mostly similar to the long-term climatology (Figure 8), where larger contributions tend to occur at those cells with more events. The northwest part (~42.5˚N) of the basin shows showing uncommonly high snowfall contributions (as do cool or wet years), which indicates that for warm and dry conditions minor snowfall events still are especially important as compared with the rest of UCRB.

3.5 Domain-wide trends

We performed the non-parametric Mann-Kendall (MK) test (Mann, 1945; Kendall, 1957) on the time series of basin-wide number of AR-related storms, all storms and number of storm days per year and found no trends at the 0.05 significance level. Further, the contributions of AR-storm, all storms and all precipitation reported in Figure 4 also failed to pass the MK-test at the 0.05 significance level. As for the basin-wide analysis, we found no trends in either the number of storms or SWE per storm, and for either AR- or all storms. Generally, no significant trends were detected by the MK-test (5% significance level) over the 1949-2015 study period. However, there are statistically significant trends detected by MK if we include earlier snowpack records. We tested the annual trends in number of storm days and storms over a longer period, 1916-2015, with VIC snow data generated using the same methods (AR-related trends cannot be extended because the AR catalog is not available before 1948). The annual time series of both are show downward trends, which reveals that there are fewer storms in recent decades compared to the early 1900s.

Nonetheless, the storms' annual contribution percentage to peak SWE does not showing any (decreasing) trend over 1916-2015 as does the storm number, which suggests that the average contribution percentage per storm might be increasing. However, we checked the trend in ΔSWE per storm and found that there is no significant trend over the same period as above. Therefore, the increasing contribution percentage per storm must be the result of decreasing annual peak SWE.

We then applied the MK-test at each grid cell in the domain to evaluate the spatial pattern of trends. Figure 12 shows the grid cells with statistically significant trends in the number of AR-storms, storms and storm days. There are only 4.6% and 8.4% of total valid cells (long-term Apr 1st SWE > 50 mm) that have downward trends in the number of AR-storm and all-storms trends. The numbers of cells diagnosed as showing upwards in the domain are negligible: no annual upward trend detected in AR-storms, one cell for number of all storms and four cells for number of individual storm days. Overall, the number of events in the basin does not show obvious trends over the study period as Figure 12 presents.

Although there are no statistically significant trends in any basin-average storms contributions (AR-storm, all-storm, and all precipitation), a number of individual grid cells in the domain have statistically significant trends as shown in Figure 13. The percentage of each type are summarized in Table 3. The grid cells with increasing contributions are primarily in the middle-latitude zone of the domain. In the northwest and southeast part of the domain, ~10% of the total cells have a significant downward trend in contributions of AR-storm, all-storm and all precipitation to the snowpack. Figure 14 shows the trend detected by the MK-test in temperature and precipitation during the accumulation season over all years. The spatial patterns in Figure 13 panel (b) and (c) match well with the pattern of trends in precipitation (Figure 14 right panel), which suggests that trends in precipitation likely are the primary factor. These maps suggest that

over the entire study period, the snowpack source is moving towards the mid-zone of the domain from the northern and southern extremes.

Finally, we applied field significance tests to investigate whether the trends detected at each individual cell would be statistically significant at the domain level. We followed the approach proposed by (Livezey & Chen, 1983) in conducting field significance tests. We determined the degree of freedom (number of independent sites) following the Chi-square-distribution method proposed by (Wang & Shen, 1999). The results show that there are too few cells with trends in number for all three types of events (see Figure 12 presents) to pass the field test, i.e. the trends in number of events are not field significant. However, the percentage of cells with trends in contribution are high enough (Figure 13) to be statistically significant as determined by the field significant test. Therefore, we reject the null hypothesis that the trends at those cells can be attributed to chance and declare the trends in contributions at those cells are statistically significant. The fact that the basin-average results do not show statistically significant trends (discussed above) may be the result of upward and downward cells cancelling over the domain.

4. Summary and Conclusion

We applied the VIC model forced with the Livneh dataset to reconstruct snowpack in the Upper Colorado mountainous region for the last six decades. On average, the simulated daily SWE time series successfully capture the major characteristics of surface observations during the accumulation season. Using the reconstructed SWE and meteorological data, we employed a snowfall-oriented definition to identify storms and further investigate the storms variations and contributions over the domain. Specifically, we conclude that:

1. The average number of days identified as being associated with snowfall storms is 37.4 per year, consisting of an average of 16.2 storms that contribute to the majority (78.2%) of the annual peak SWE. Atmospheric Rivers in the UCRB affect ~70% of these storms and supply 56.9% of the accumulated snowpack's peak value. Compared to the Sierra Nevada region (Huning & Margulis, 2017), the values are smaller in the UCRB

(primarily because the distance from the coast is greater) but nonetheless is still quite high.

2. In the mountainous parts of the UCRB, moderate and heavy storms are still the prominent source of SWE for all four climatic conditions we studied. In wet and cold years, snowfall contributes more efficiently to annual peak SWE because of less mid-season melt or sublimation. More minor snowfall events occur under dry and warm scenarios, and they contributed to 48.4% and 35.8% peak SWE value during the accumulation season (compared with 21.0% and 27.6% during wet and cold years).

3. The eastern part of the basin tends to have more storms (and AR-storms) and higher storm contributions to snow accumulation than the western part. Small-scale snowfall events have the greatest effect on snow accumulation in the northwestern part of the basin. By investigating the coverage and contribution of each AR- and non-AR storm, we found that ~70% for the snowpack is attributable to events that cover at least 90% of the domain. In other words, of all the (AR-) storms, domain-wide events make the main contribution to SWE.

4. On a basin-wide basis, there are no statistically significant trends in the total number of storms, number of AR-storms, or in total storm days over 1949-2015. However, the number of storms does show a statistically significant downward trend over a longer period (1916-2015). On the other hand, there are statistically significant trends for some (less than 1/3 of total number) individual grid cells. Upward trends mainly are in the mid-latitude mountainous portion of the basin and grid cells with downward trends are mostly in the northwestern and southeastern portions of the basin.

Chapter 4 Table 1: Long-term mean number of storm days, storms and AR-storms in one year as described in section 3.3. All-water year climatology is also provided for reference.

	Wet years	Dry years	Warm years	Cold years	All years
Storm days	54.9	23.0	36.2	41.5	37.4
Storms	18.7	12.8	14.8	17.6	16.2
AR-storms	13.2	8.0	10.9	11.9	11.2

Chapter 4 Table 2: Average contributions of AR-storm, all storms and all precipitation to annual peak SWE for wet, dry, warm and cold years. The last column presents the climatology of the basin annual SWE maximum under each category.

	Wet years	Dry years	Warm years	Cold years	All years
(a) AR-storm	65.9%	49.7%	61.7%	56.0%	56.9%
(b) All storms	86.1%	70.3%	84.0%	76.9%	78.2%
Total	107.1%	118.7%	119.8%	104.5%	110.6%
(a)/(b)	76.5%	70.7%	73.5%	72.8%	72.8%
SWE (km^3)	34.0	13.8	20.3	27.1	23.2
ΔSWE per AR-storm (km^3)	1.97	1.22	1.37	1.66	1.51
ΔSWE per storm (km^3)	1.57	0.76	1.15	1.18	1.12

Chapter 4 Table 3: Percentage of grid cells that have trends in annual contribution of AR-storm, all storm and all precipitation (Total) at 0.05 significant level over the domain.

	AR-storm	All storm	Total
Upward trend	8.9%	16.1%	20.1%
Downward trend	8.0%	9.7%	11.2%
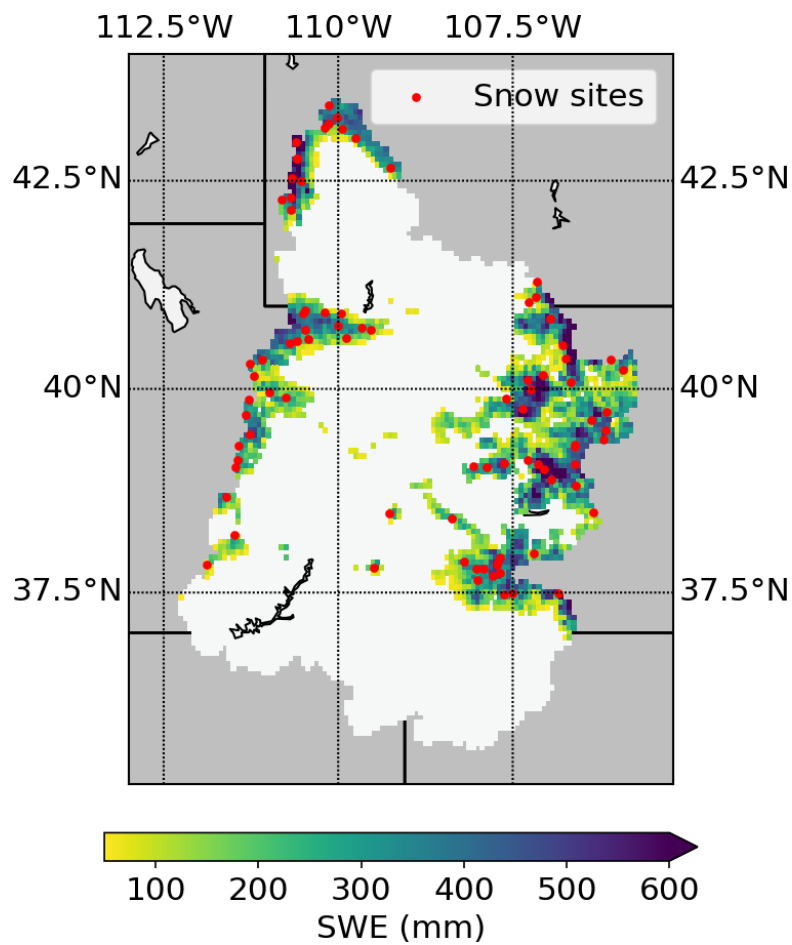
Chapter 4 Figure 1: Headwater regions in the Upper Colorado River Basin. Only those grid cells with long-term average Apr 1st SWE>50mm are shown. Red dots mark the 86 SNOTEL stations location inside the domain.
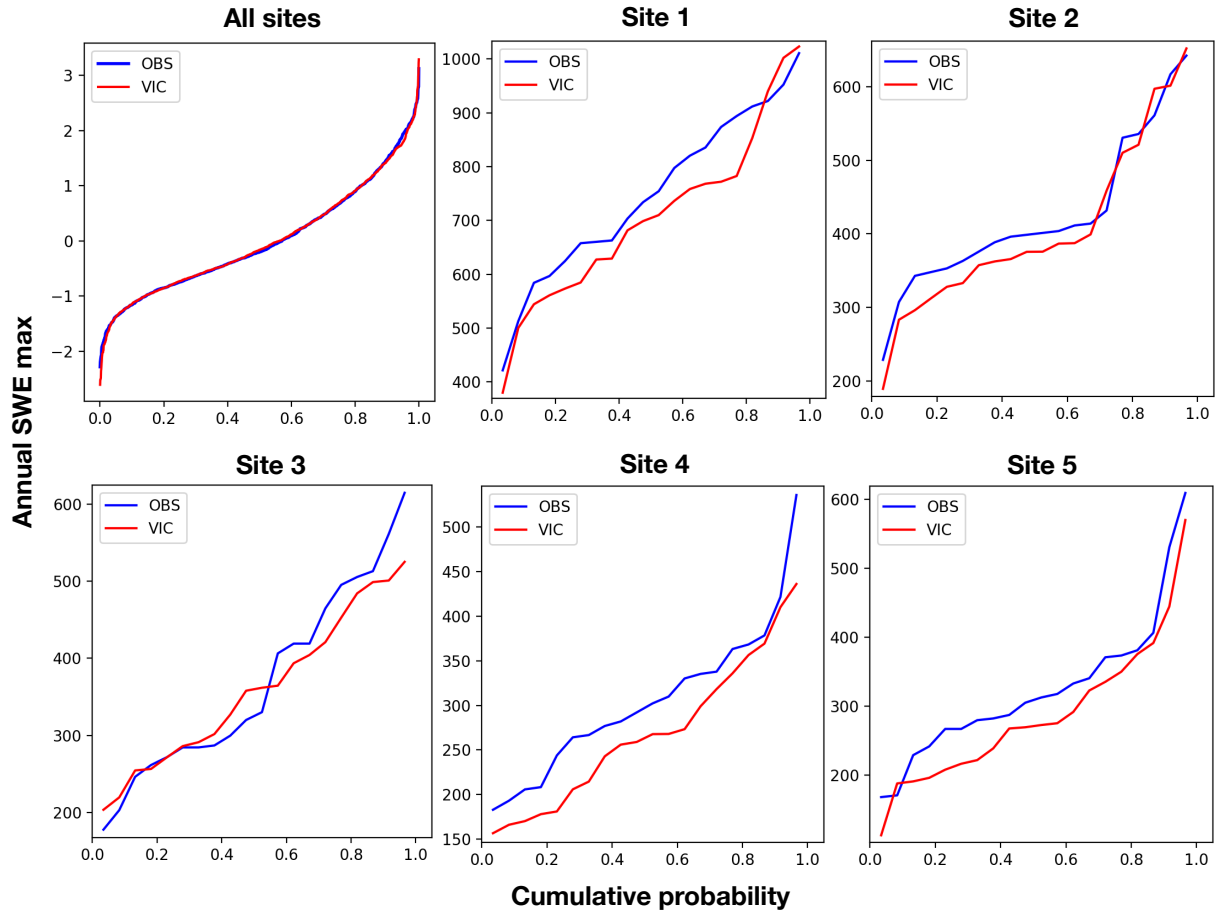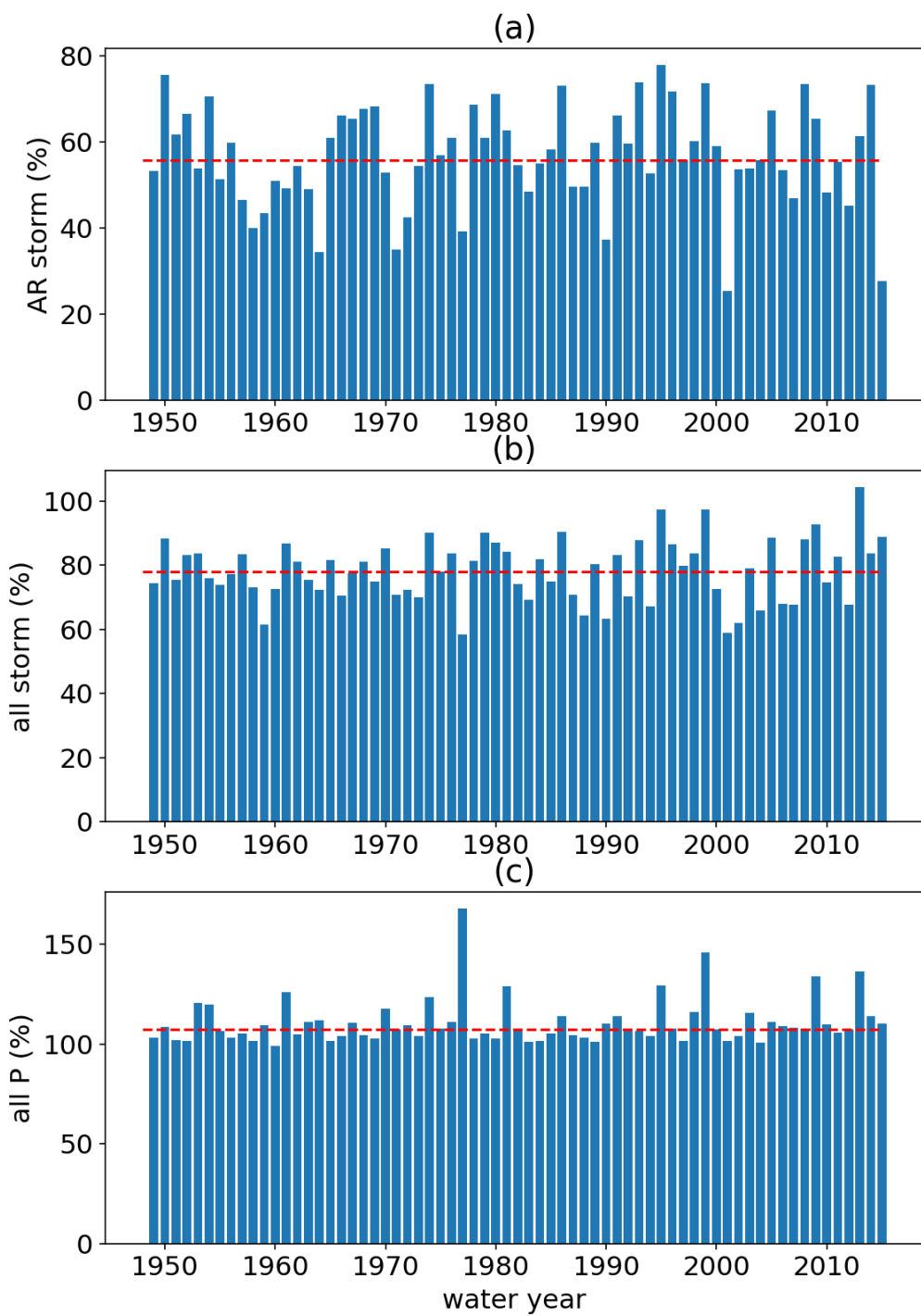
Chapter 4 Figure 2: CDFs of simulated (red) and observed (blue) annual SWE max values in all years. The first panel incorporates normalized SWE values from all 86 SNOTEL sites. The other panels reported data at 5 selected stations (detailed information of these 5 sites is provided in the Supplement). The units for SWE is mm except the first panel.
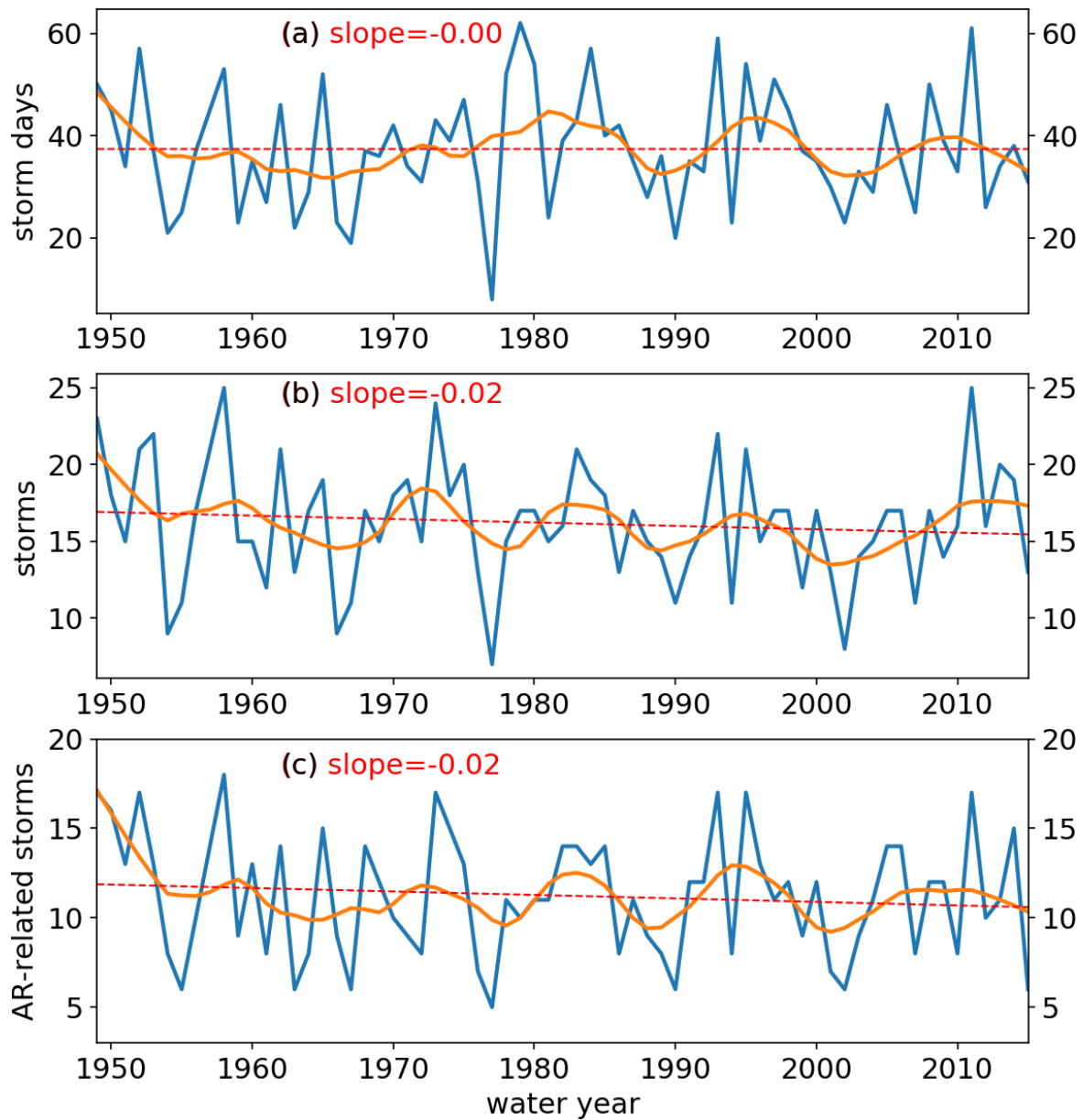
Chapter 4 Figure 3: Time series plots of number of storm days (top), number of storms (middle) and number of AR-related storms (bottom) for 1949-2015. The red dashed line is the linear regression against time (although none is statistically significant). The slope is reported in red. The orange line is smoothed using a Lowess fitter (fraction = 0.17).

Chapter 4 Figure 4: The contribution of (a) AR-storms, (b) all storms and (c) all precipitation to basin-wide SWE in each year. The red dashed line indicates the long-term mean.
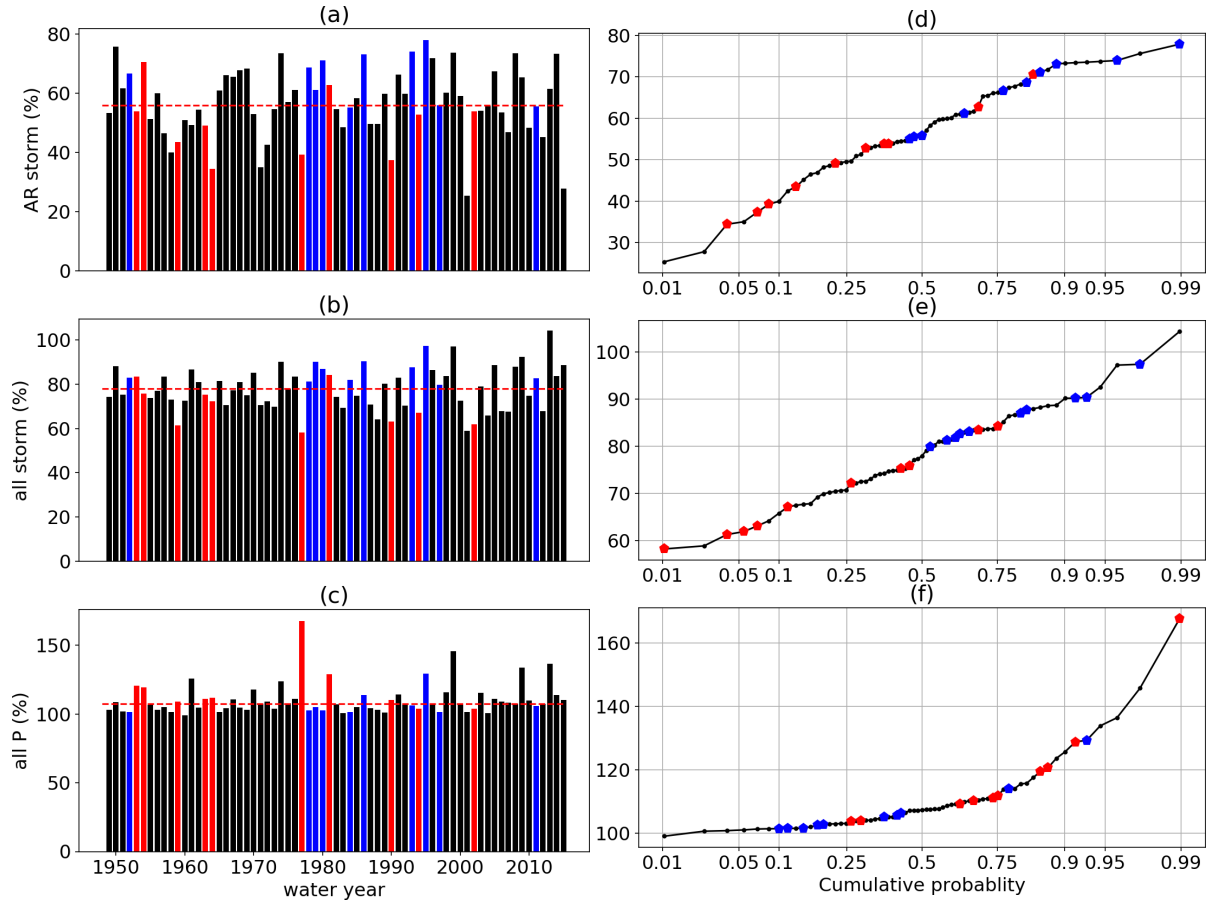
Chapter 4 Figure 5: Bar plots (left column) and empirical distributions (right column) of the contribution to peak SWE of AR storms, all storms, and all precipitation over the study period. Wet years are highlighted with blue and dry years are with red. The left column bars are the same as in Figure 4.

100
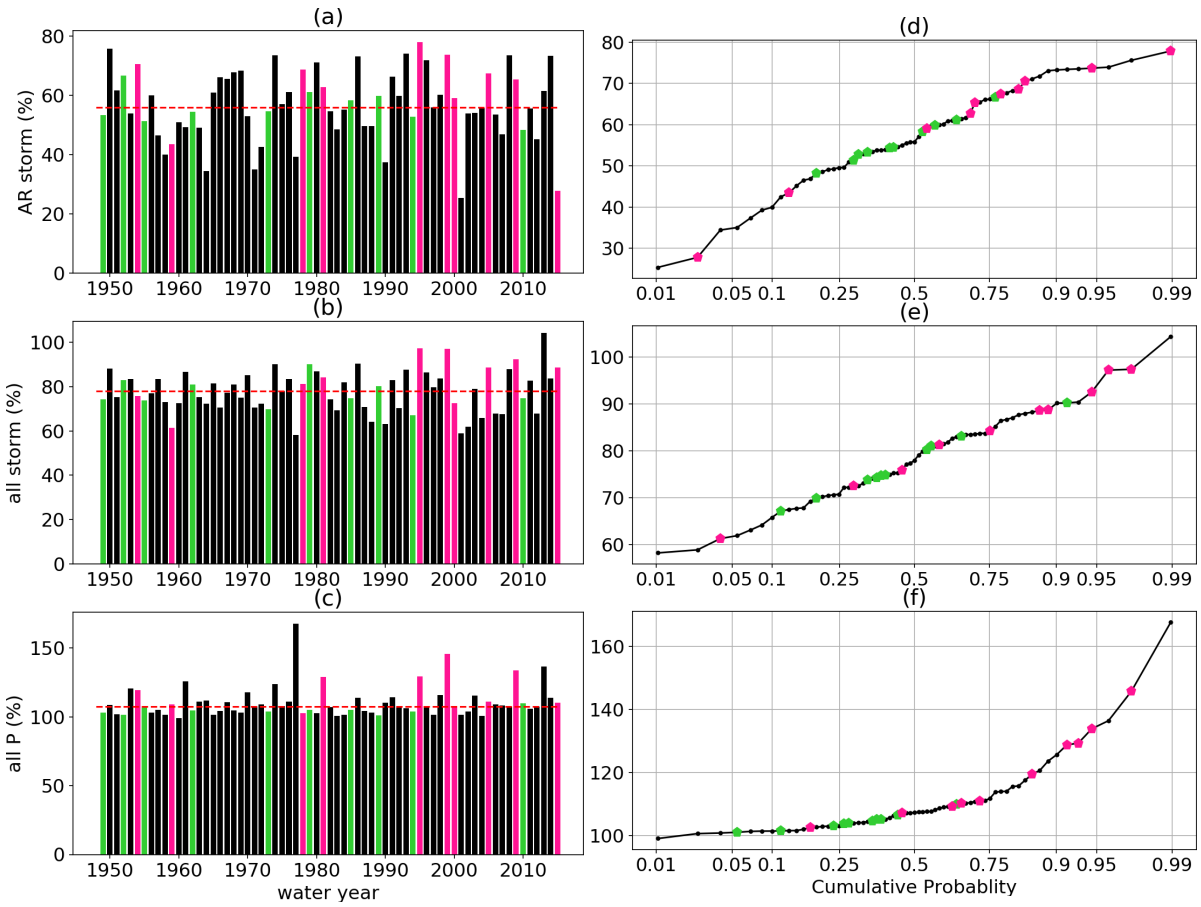
Chapter 4 Figure 6: Same as Figure 5 but for warm years (pink) and cold years (green).
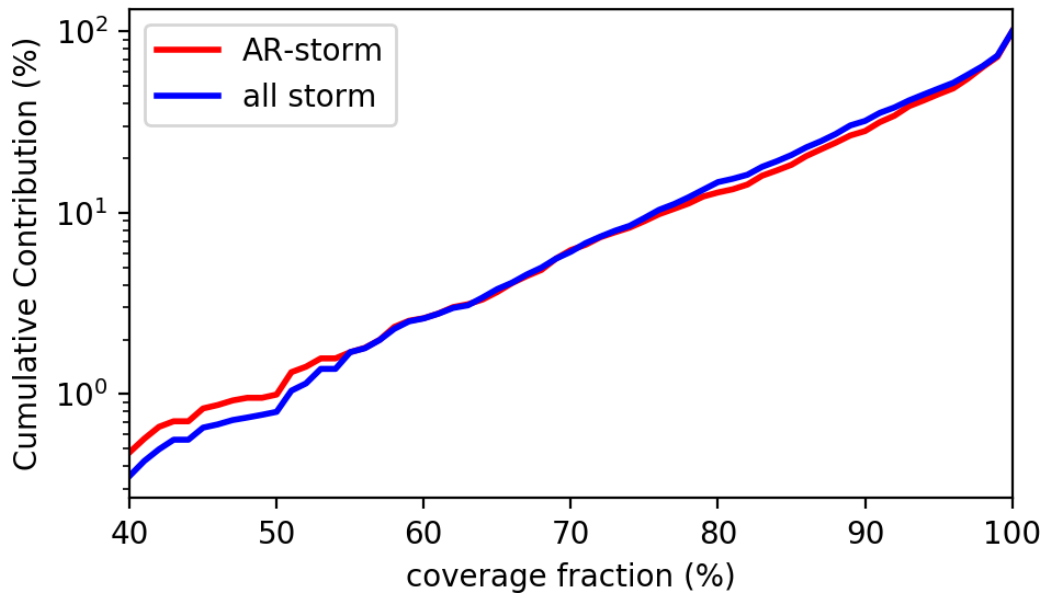
Chapter 4 Figure 7: Coverage area fraction vs cumulative contribution to snowpack of AR-storms (red) and all storms (blue). The y-axis is in log scale.
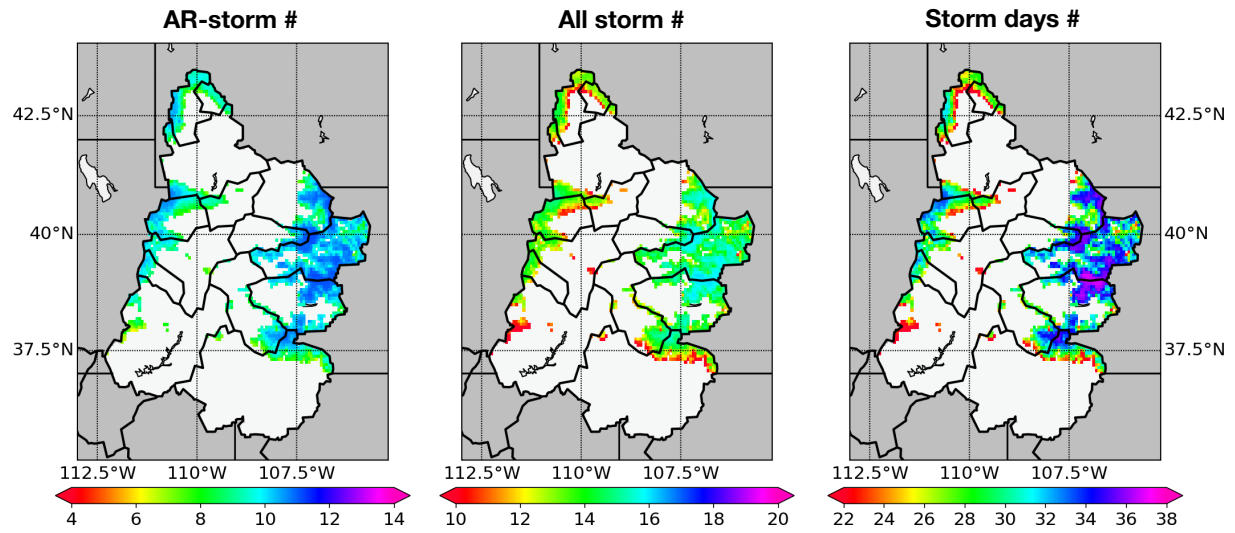
Chapter 4 Figure 8: Multiyear average number of AR-storms (left), all storms (middle) and storm days (right) for all grid cells. Note that the color scales are different in each panel.
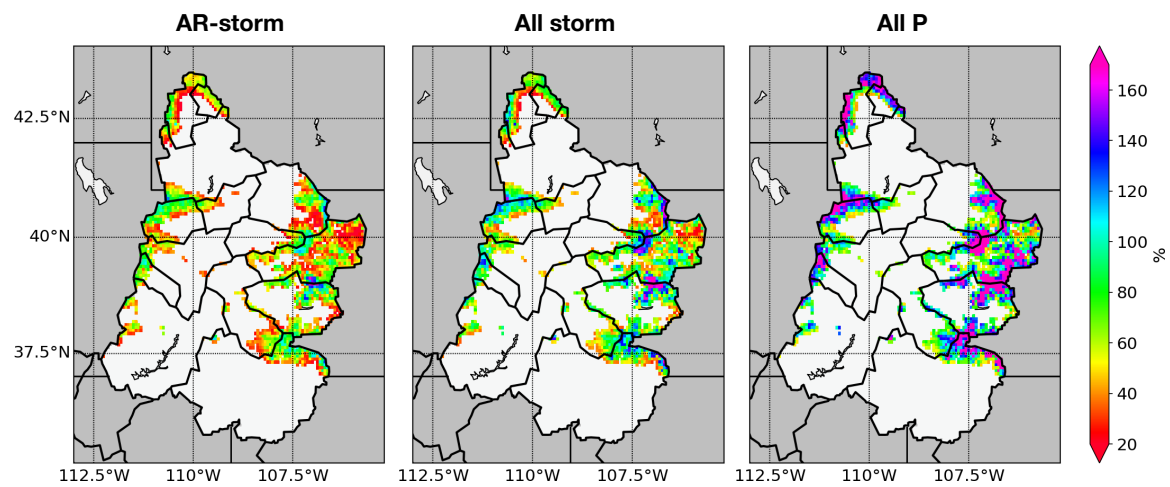
Chapter 4 Figure 9: AR-storm (left), all storms (middle) and all precipitation (right) average contribution to annual snowpack maximum over the study period.
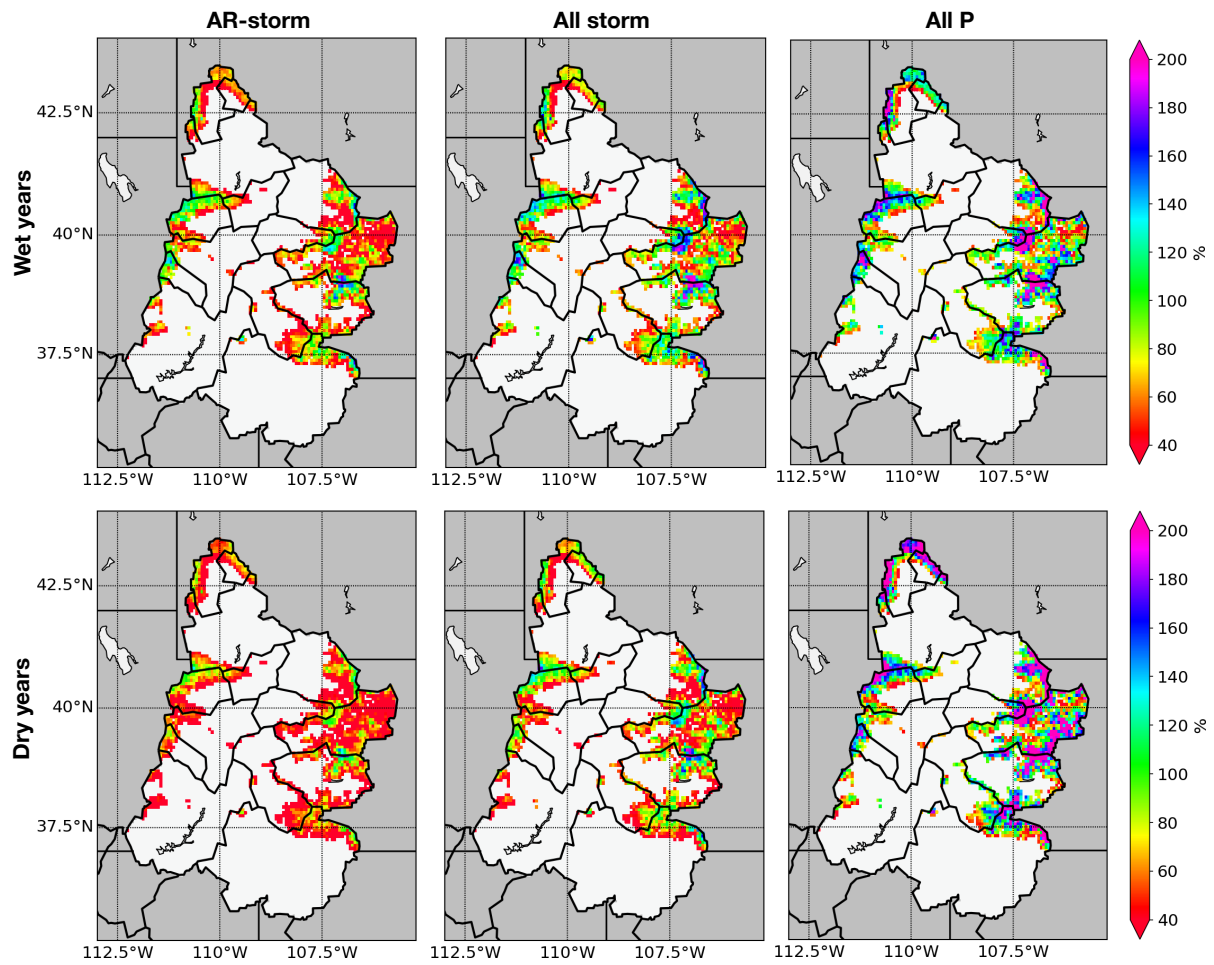
Chapter 4 Figure 10: Average contribution of AR-storm, all storms and all precipitation to annual SWE maximum over the selected wet (top row) and dry (bottom row) years for each individual grid cell.
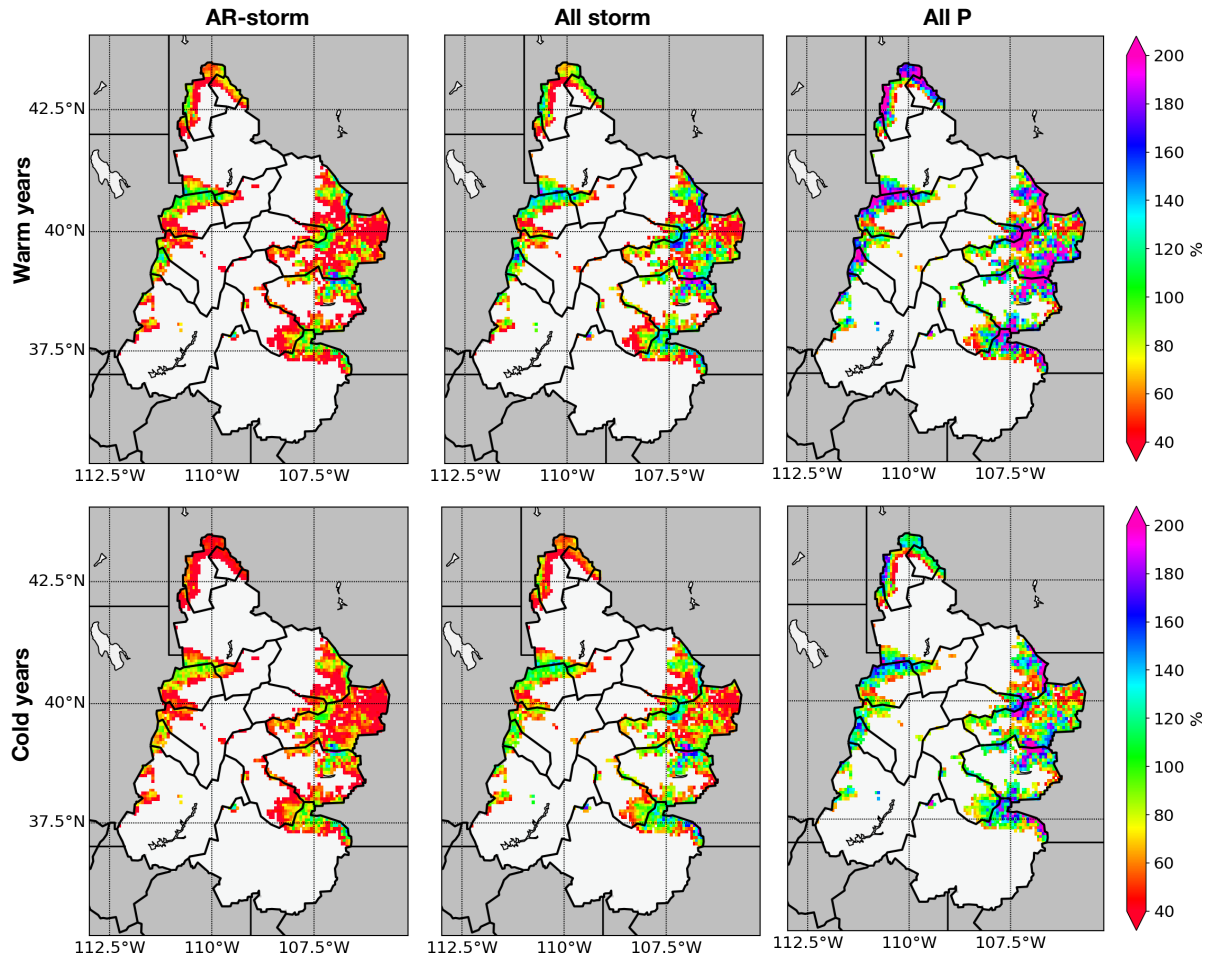
Chapter 4 Figure 11: Same as Figure 9 but for warm (top row) and cold (bottom row) years.
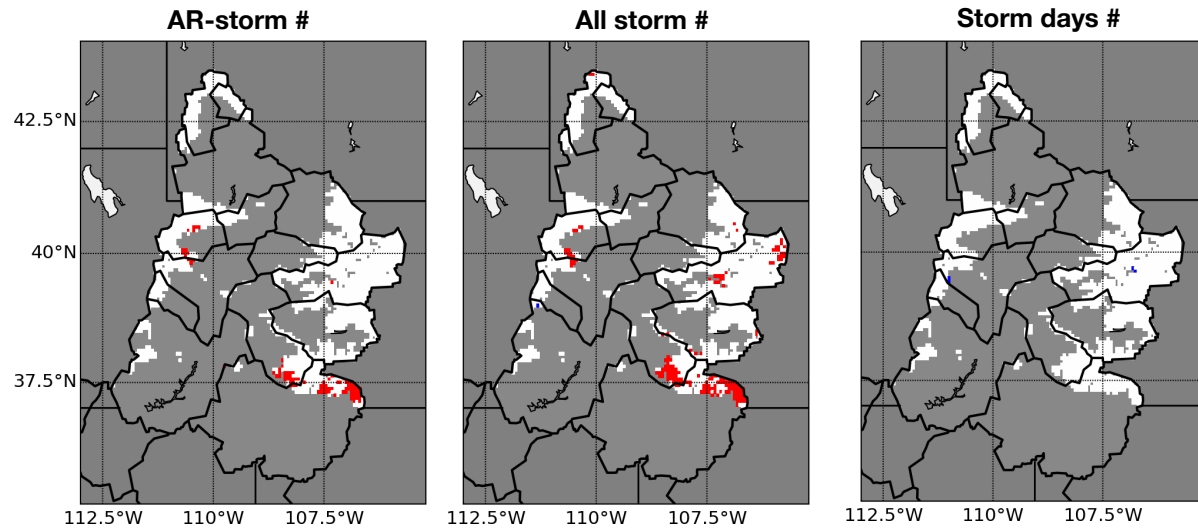
Chapter 4 Figure 12: Annual trends in number of AR-storms (left), all storms (middle) and individual storm days (right) at all grid cells. Blue indicates upward trend, red is downward trend and white represents no significant trend. Only the cells with long-term Apr-1st SWE > 50 mm are shown. The trend is determined by MK-test at 0.05 significant level.
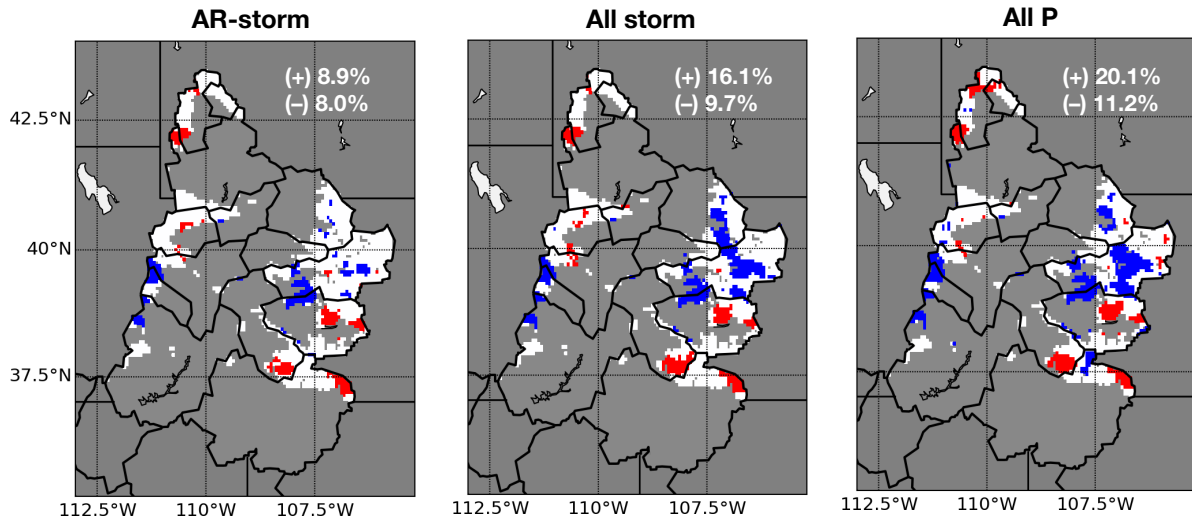
Chapter 4 Figure 13: Annual trends in contributions of AR-storms (left), all storms (middle) and all precipitation (right) to annual maximum SWE. Blue indicates upward trend and red is downward trend. The trend is determined by MK-test at 0.05 significant level.
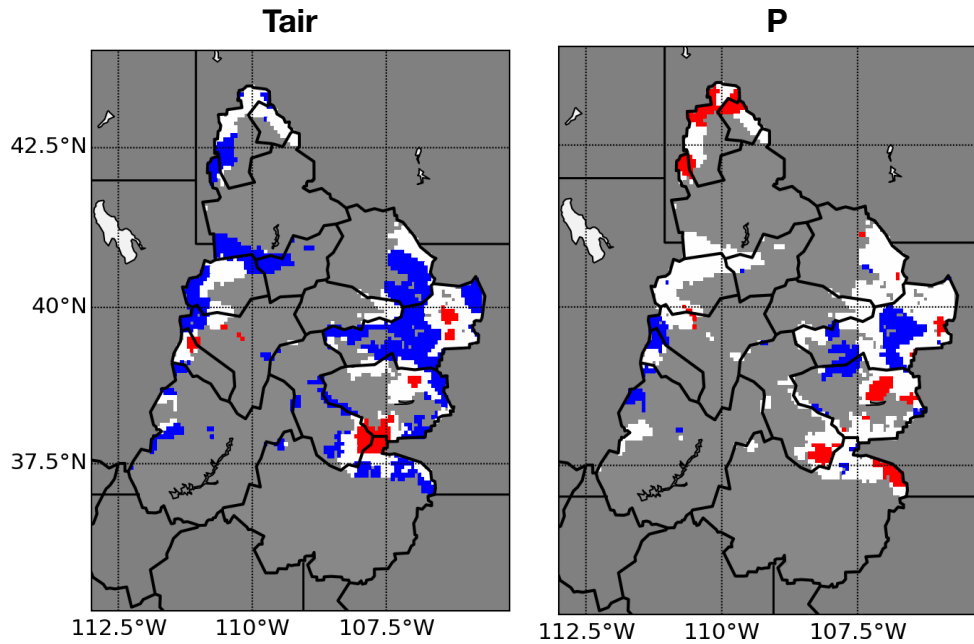
Chapter 4 Figure 14: MK-trend test results for temperature (left) and total precipitation during the accumulation season at each single grid over the 1949-2015. Blue indicates statistically significant upward trend and red indicates statistically significant downward trend at 0.05 significance level.

Chapter 5 Conclusions

My dissertation has sought to provide a better understanding of the influences of climatic controls on runoff and snowpack in the Colorado River Basin, using land surface modeling as the primary tool. In particular, I posed three questions in **Chapter 1** that my dissertation addresses:

1) What are the causes of Colorado River streamflow decline over the last century?

2) What are the factors that control snow ablation processes, and how well or poorly are they represented in models?

3) What are the characteristics of winter storms that contribute to UCRB snowpack, and how spatially coherent are these storms? What role do Atmospheric Rivers play in SWE accumulation in the UCRB?

In order to address these questions, I applied macroscale land surface models (LSM) in the Colorado River Basin and used gridded climate data, as well as observations from selected USDA SNOTEL sites over the UCRB headwaters, as well as other mountainous parts of the Western U.S.. Having verified the ability of one model (VIC) to reproduce historic streamflow and SWE variations, I analyzed the climatic drivers and effects on these key hydrological variables in the UCRB.

In **Chapter 2** I explore both century-long downtrends in streamflow in UCRB, as well as differences between recent drought years (2000-2014) and an earlier (1950s-1960s) drought. Absent strong trends in precipitation, I examine the effect of pervasive warming on the decreasing trend in streamflow. By removing the observed (annual) temperature trend and running a parallel simulation with VIC, I was able to isolate the part of the runoff trend attributable to warming temperature from other possible causes. I find that more than half of the decreasing trend is attributable to warming. This leaves the question of causes of the other (approximately) half of the

110

trend. I find that notwithstanding little or no basin-wide precipitation trend, there have been small downward trends in precipitation in the UCRB sub-basins that produce most of the runoff basin-wide, and that these local trends contributes much of the half of the long-term decreasing trend not attributable to warming. I also find that the long-term trend is mostly associated with increased winter ET (mainly snow sublimation). Through comparison of the ongoing post-Millennium drought with the drought of the 1950s and 1960s, I find that, as in the case of the long-term trend, about half of the post-Millennium drought signature is attributable to anomalous warmth over the last two decades, whereas the 1950s-60s event was associated almost entirely with anomalously low precipitation. As for the long-term trend, that part of the post-Millennium drought that is not attributable to warming is associated primarily with localized precipitation deficits, especially in high runoff-producing headwater sub-basins. Finally, I examined the unusual characteristics of the Spring 2017 UCRB runoff, where early forecasts of exceptionally high seasonal runoff were reduced through late winter and spring with the observed seasonal runoff eventually being only slightly above normal. My analysis indicates that while late winter was exceptionally warm, the main cause of the progressive reductions in the forecasts was below normal late winter and early spring precipitation.

In **Chapter 3**, I utilized a multi-model approach to study model-predicted snow ablation processes at 10 SNOTEL sites distributed across the Western U.S. (including one within the UCRB). The average Mean Absolute Error (MAE) for all models is 4.3 mm/day (22% of the observed average ablation rate across the 10 stations), ranging from 3.6 mm/day (Catchment) to 6.7 mm/day (Noah-MP). SSiB is the only model that has positive bias (higher ablation rate than observations) in the baseline experiments. I find that net radiation is highly correlated with melt rates (more so than temperature), which is consistent with other published studies. I also fid that

for the most part, wind speed is not a strong predictor of ablation rates during the snow melt period. On the other hand, the representation and effects of vegetation cover vary substantially across the models. When the canopy is removed in the models, the simulated sensible heat in VIC and Noah-MP reverses direction and its magnitude decreases substantially. The direction of sensible heat is unchanged in Catchment and SSiB, but the magnitude of the former decreases and of the latter increases. Catchment predicts decreased sensible heat under the no-vegetation situation, because it does not represent attenuation of wind speed by the canopy in its calculation and therefore the only effect of removing the canopy is to reduce the surface roughness.

Chapter 4 examines the contribution of storms (both atmospheric river (AR)-related and non-AR related) that cover substantial portions of the UCRB on basin-wide SWE. I performed examinations of winter SWE accumulation using VIC model reconstructions of UCRB snowpack for the 67-year period 1949-2015 (water years). Given the model-derived gridded snow observations, I used a snowfall-oriented definition to identify storms of both types (AR and non-AR). I find that, on average, AR-related storms in the UCRB comprise about 70% of all winter storms and supply 56.9% of the basin peak SWE value. This contribution to SWE is higher (65.9%) in wet years and lower (49.7%) during dry years. During wet and cold years, snowfall contributes more efficiently to SWE because there is a smaller proportion of mid-season melt and sublimation. More minor snowfall events play a more important role in dry and warm years. In terms of the spatial scale of storms, I find that ~70% of the snowpack is attributable to events that cover at least 90% of the domain (defined as that part of UCRB for which the mean maximum SWE exceeds 50 mm). Also, on a basin-wide basis, I find no statistically significant trends in the total number of storms over my 67-year study period. However, over the longer 1916-2015 period there is a statistically significant downward trend in the number of storms, which is related to a period of

high precipitation during the early 1920s (note that it is not possible to examine AR-related storms prior to 1949, which is the earliest year for which AR information exists). On the other hand, there are statistically significant trends (mostly upward) in the number of storms (boteh AR- and non-AR) in the mid-latitude portion of the basin, and some downward trends are mostly in the northwestern and southeastern portions of the basin, suggesting that SWE is becoming more concentrated toward the center of the basin.

This dissertation presents what I believe are important findings and analysis of climatic effects on runoff and snowpack over the CRB. Because the CRB's streamflow is heavily regulated and highly dependent on snowpack, identifying the influence of meteorological drivers of streamflow is particularly important. Notwithstanding that my work mostly addresses science questions relevant to the CRB, it has broader implications and applications to water-related issues across the Western U.S., and to snow-affected mountainous areas globally. For instance, the methods I have developed and applied here can be used to examine warming effects on streamflow in other river basins, especially in snow-dominant regions. Separating temperature effects from other factors that affect streamflow can also provide an historical reference for hydrological predictions of warming effects on streamflow over the next century. Also, despite the fact that I was able to produce plausible reconstructions of streamflow (and snowpack) in the CRB using he VIC model, my multi-model examination of factors that affect four model's predictions of snow ablation indicates that much remains to be done to improve the realism and fidelity of hydrologic model simulations. The complex physical processes in play are not always completely represented, and require a balance between a desire for modest complexity and realism. Enhancing model performance can involve decades of work which must balance model performance with theory. These undertakings will require comprehensive modeling experiments with high-quality field

observations. The spatial distribution and patterns of snow accumulation in mountainous environments link directly to streamflow and streamflow forecasts. Hence, it is necessary to improve our understanding of snowpack patterns during both accumulation and melting seasons. All of these topics are related to one or more core conclusions of my dissertation research, which should provide insights for future research on the relevant questions.

Reference

Andreadis, K. M., Storck, P., & Lettenmaier, D. P. (2009). Modeling snow accumulation and ablation processes in forested environments. *Water Resources Research*, *45*(5), 1–13. https://doi.org/10.1029/2008WR007042

Barnett, T. P., & Pierce, D. W. (2009). Sustainable water deliveries from the Colorado River in a changing climate. *Proceedings of the National Academy of Sciences of the United States of America*, *106*(18), 7334–7338. https://doi.org/10.1073/pnas.0812762106

Bohn, T. J., Livneh, B., Oyler, J. W., Running, S. W., Nijssen, B., & Lettenmaier, D. P. (2013). Global evaluation of MTCLIM and related algorithms for forcing of ecological and hydrological models. *Agricultural and Forest Meteorology*, *176*, 38–49. https://doi.org/10.1016/j.agrformet.2013.03.003

Burnash, R. J. C., Ferral, R. L., & McGuire, R. A. (1973). *A generalized streamflow simulation system – Conceptual modelling for digital computers*.

Chen, F., & Dudhia, J. (2001). Coupling an Advanced Land Surface–Hydrology Model with the Penn State–NCAR MM5 Modeling System. Part I: Model Implementation and Sensitivity. *Monthly Weather Review*, *129*(4), 569–585. https://doi.org/10.1175/1520-0493(2001)129<0569:CAALSH>2.0.CO;2

Chen, F., Barlage, M., Tewari, M., Rasmussen, R., Jin, J., Lettenmaier, D., et al. (2014). Modeling seasonal snowpack evolution in the complex terrain and forested colorado headwaters region: A model intercomparison study. *Journal of Geophysical Research*, *119*(22), 13795–13819. https://doi.org/10.1002/2014JD022167

Christensen, Nadia S, & Lettenmaier, D. P. (2007). A multimodel ensemble approach to assessment of climate change impacts on the hydrology and water resources of the Colorado

River Basin. *Hydrology and Earth System Sciences Discussions*, *11*(4), 1417–1434.

Christensen, Niklas S., Wood, A. W., Voisin, N., Lettenmaier, D. P., & Palmer, R. N. (2004). The effects of climate change on the hydrology and water resources of the Colorado River basin. *Climatic Change*, *62*(1–3), 337–363. https://doi.org/10.1023/B:CLIM.0000013684.13621.1f

Clark, M. P., Hendrikx, J., Slater, A. G., Kavetski, D., Anderson, B., Cullen, N. J., et al. (2011). Representing spatial variability of snow water equivalent in hydrologic and land-surface models: A review. *Water Resources Research*, *47*(7). https://doi.org/10.1029/2011WR010745

Cohen, M., Christian-Smith, J., & Berggren, J. (2013). *Water to Supply the Land: Irrigated Agriculture in the Colorado River Basin*. Retrieved from http://water.usgs.gov/watercensus/colorado.html

Cortés, G., Girotto, M., & Margulis, S. (2016). Snow process estimation over the extratropical Andes using a data assimilation framework integrating MERRA data and Landsat imagery. *Water Resources Research*, *52*(4), 2582–2600. https://doi.org/10.1002/2015WR018376

Dawadi, S., & Ahmad, S. (2012). Changing climatic conditions in the Colorado River Basin: Implications for water resources management. *Journal of Hydrology*, *430–431*, 127–141. https://doi.org/10.1016/j.jhydrol.2012.02.010

Dozier, J., Bair, E. H., & Davis, R. E. (2016). Estimating the spatial distribution of snow water equivalent in the world's mountains. *Wiley Interdisciplinary Reviews: Water*, *3*(3), 461–474. https://doi.org/10.1002/wat2.1140

Dyer, J. L., & Mote, T. L. (2007). Trends in snow ablation over North America. *International Journal of Climatology*, *27*(6), 739–748. https://doi.org/10.1002/joc.1426

Easterling, D. R., Karl, T. R., Mason, E. H., Hughes, P. Y., & Bowman, D. P. (1996). *United States Historical Climatology Network (US HCN), Monthly Temperature and Precipitation Data*. Retrieved from http://cdiac.esd.ornl.gov/ndps/ushcn/abstract_r3.html

Ek, M. B., Mitchell, K. E., Lin, Y., Rogers, E., Grunmann, P., Koren, V., et al. (2003). Implementation of Noah land surface model advances in the National Centers for Environmental Prediction operational mesoscale Eta model. *Journal of Geophysical Research*, *108*(D22), 8851. https://doi.org/10.1029/2002JD003296

Guan, B., Waliser, D. E., Ralph, F. M., Fetzer, E. J., & Neiman, P. J. (2016). Hydrometeorological characteristics of rain-on-snow events associated with atmospheric rivers. *Geophysical Research Letters*, *43*(6), 2964–2973. https://doi.org/10.1002/2016GL067978

Hamlet, A. F., & Lettenmaier, D. P. (2005). Production of Temporally Consistent Gridded Precipitation and Temperature Fields for the Continental United States. *Journal of Hydrometeorology*, *6*(3), 330–336. https://doi.org/10.1175/JHM420.1

Hansen, J., Ruedy, R., & Sato, M. (2001). A closer look at United States and global surface temperature change. *Journal of Geophysical Research*, *106*(23), 947–963. https://doi.org/10.1029/2001JD000354

Hoerling, M., Barsugli, J., Livneh, B., Eischeid, J., Quan, X., & Badger, A. (2019). Causes for the Century-Long Decline in Colorado River Flow. *Journal of Climate*, 8181–8203. https://doi.org/10.1175/jcli-d-19-0207.1

Hungerford, R. D., Nemani, R. R., Running, S. W., & Coughlan, J. C. (1989). MTCLIM: a mountain microclimate simulation model. https://doi.org/10.2737/INT-RP-414

Huning, L. S., & Margulis, S. A. (2017). Climatology of seasonal snowfall accumulation across

the Sierra Nevada (USA): Accumulation rates, distributions, and variability. *Water Resources Research*, *53*(7), 6033–6049. https://doi.org/10.1002/2017WR020915

Kalnay, E., Kanamitsu, M., Kistler, R., Collins, W., Deaven, D., Gandin, L., et al. (1996). The NCEP/NCAR 40-Year Reanalysis Project. *Bulletin of the American Meteorological Society*, *77*(3), 437–471. https://doi.org/10.1175/1520-0477(1996)077<0437:TNYRP>2.0.CO;2

Koster, R. D., Suarez, M. J., Ducharne, A., Stieglitz, M., & Kumar, P. (2000). A catchment-based approach to modeling land surface processes in a general circulation model 1. Model structure. *J. Geophys. Res.*, *105*(D20), 24809–24822. https://doi.org/10.1029/2000jd900327

Li, D., Wrzesien, M. L., Durand, M., Adam, J., & Lettenmaier, D. P. (2017). How much runoff originates as snow in the western United States, and how will that change in the future? *Geophysical Research Letters*, *44*(12), 6163–6172. https://doi.org/10.1002/2017GL073551

Liang, X., & Lettenmaier, D. (1994). a simple hydrologically based model of land surface water and energy fluxes for general circulation models. *Journal of Geophysical …*, *99*. Retrieved from http://onlinelibrary.wiley.com/doi/10.1029/94JD00483/full

Livneh, B., Rosenberg, E. A., Lin, C., Nijssen, B., Mishra, V., Andreadis, K. M., et al. (2013). A long-term hydrologically based dataset of land surface fluxes and states for the conterminous United States: Update and extensions. *Journal of Climate*, *26*(23), 9384–9392. https://doi.org/10.1175/JCLI-D-12-00508.1

Lohmann, D., Nolte-Holube, R., & Raschke, E. (1996). A large-scale horizontal routing model to be coupled to land surface parametrization schemes. *Tellus, Series A: Dynamic Meteorology and Oceanography*, *48*(5), 708–721. https://doi.org/10.1034/j.1600-0870.1996.t01-3-00009.x

Di Luzio, M., Johnson, G. L., Daly, C., Eischeid, J. K., & Arnold, J. G. (2008). Constructing

retrospective gridded daily precipitation and temperature datasets for the conterminous

United States. *Journal of Applied Meteorology and Climatology*, *47*(2), 475–497.

https://doi.org/10.1175/2007JAMC1356.1

Magand, C., Ducharne, A., Le Moine, N., & Gascoin, S. (2013). Introducing Hysteresis in Snow

Depletion Curves to Improve the Water Budget of a Land Surface Model in an Alpine

Catchment. *Journal of Hydrometeorology*, *15*(2), 631–649. https://doi.org/10.1175/jhm-d-

13-091.1

McCabe, G J, & Wolock, D. M. (2002). Trends and temperature sensitivity of moisture

conditions in the conterminous United States. *Climate Research*, *20*(1), 19–29. Retrieved

from http://proxy-

remote.galib.uga.edu/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=lah

&AN=20023175037&site=eds-live%5Cnemail: gmccabe@usgs.gov

McCabe, Gregory J., & Wolock, D. M. (2007). Warming may create substantial water supply

shortages in the Colorado River basin. *Geophysical Research Letters*, *34*(22), 1–5.

https://doi.org/10.1029/2007GL031764

McCabe, Gregory J., Wolock, D. M., Pederson, G. T., Woodhouse, C. A., & McAfee, S. (2017).

Evidence that recent warming is reducing upper Colorado river flows. *Earth Interactions*,

21(10). https://doi.org/10.1175/EI-D-17-0007.1

Milly, P. C. D., & Dunne, K. A. (2020). Colorado River flow dwindles as warming-driven loss

of reflective snow energizes evaporation, *9187*(February), 1–9.

Moore, R. D., & Owens, I. F. (1984). Controls on Advective Snowmelt in a Maritime Alpine

Basin. *Journal of Climate and Applied Meteorology*. https://doi.org/10.1175/1520-

0450(1984)023<0135:coasia>2.0.co;2

Mote, P. W., Hamlet, A. F., Clark, M. P., & Lettenmaier, D. P. (2005). Declining Mountain Snowpack in Western North America*. *Bulletin of the American Meteorological Society*, *86*(1), 39–49. https://doi.org/10.1175/BAMS-86-1-39

Mote, P. W., Rupp, D. E., Li, S., Sharp, D. J., Otto, F., Uhe, P. F., et al. (2016). Perspectives on the causes of exceptionally low 2015 snowpack in the western United States. *Geophysical Research Letters*, 1–9. https://doi.org/10.1002/2016GL069965

Mote, P. W., Li, S., Lettenmaier, D. P., Xiao, M., & Engel, R. (2018). Dramatic declines in snowpack in the western US. *Npj Climate and Atmospheric Science*, *1*(1), 2. https://doi.org/10.1038/s41612-018-0012-1

Musselman, K. N., Clark, M. P., Liu, C., Ikeda, K., & Rasmussen, R. (2017). Slower snowmelt in a warmer world. *Nature Climate Change*. https://doi.org/10.1038/nclimate3225

Newman, A. J., Clark, M. P., Winstral, A., Marks, D., & Seyfried, M. (2014). The Use of Similarity Concepts to Represent Subgrid Variability in Land Surface Models: Case Study in a Snowmelt-Dominated Watershed. *Journal of Hydrometeorology*, *15*(5), 1717–1738. https://doi.org/10.1175/JHM-D-13-038.1

Niu, G. Y., Yang, Z. L., Mitchell, K. E., Chen, F., Ek, M. B., Barlage, M., et al. (2011). The community Noah land surface model with multiparameterization options (Noah-MP): 1. Model description and evaluation with local-scale measurements. *Journal of Geophysical Research Atmospheres*, *116*(12), 1–19. https://doi.org/10.1029/2010JD015139

Oaida, C. M., Xue, Y., Flanner, M. G., Skiles, S. M., De Sales, F., & Painter, T. H. (2015). Improving snow albedo processes in WRF/SSiB regional climate model to assess impact of dust and black carbon in snow on surface energy balance and hydrology over western U.S. *Journal of Geophysical Research: Atmospheres*, *120*(8), 3228–3248.

https://doi.org/10.1002/2014JD022444

Painter, T. H., Skiles, S. M. K., Deems, J. S., Brandt, W. T., & Dozier, J. (2018). Variation in
Rising Limb of Colorado River Snowmelt Runoff Hydrograph Controlled by Dust
Radiative Forcing in Snow. *Geophysical Research Letters*, *45*(2), 797–808.
https://doi.org/10.1002/2017GL075826

Prairie, J., & Callejo, R. (2005). Natural Flow and Salt Computation Methods, Calendar Years
1971-1995. *Bureau of Reclamation*, (December), 1–112. Retrieved from
http://digitalcommons.usu.edu/govdocs/135/%5Cnpapers2://publication/uuid/6183E843-
86FF-4150-9535-C0407CB5477D

Rajagopalan, B., Nowak, K., Prairie, J., Hoerling, M., Harding, B., Barsugli, J., et al. (2009).
Water supply risk on the Colorado River: Can management mitigate? *Water Resources
Research*, *45*(8), 1–7. https://doi.org/10.1029/2008WR007652

Rauscher, S. A., Pal, J. S., Diffenbaugh, N. S., & Benedetti, M. M. (2008). Future changes in
snowmelt-driven runoff timing over the western US. *Geophysical Research Letters*, *35*(16),
1–5. https://doi.org/10.1029/2008GL034424

Reynolds, L. V., Shafroth, P. B., & LeRoy Poff, N. (2015). Modeled intermittency risk for small
streams in the Upper Colorado River Basin under climate change. *Journal of Hydrology*,
523, 768–780. https://doi.org/10.1016/j.jhydrol.2015.02.025

Sellers, P. J., Tucker, C. J., Collatz, G. J., Los, S. O., Justice, C. O., Dazlich, D. A., & Randall,
D. A. (1996). A Revised Land Surface Parameterization (SiB2) for Atmospheric GCMS.
Part II: The Generation of Global Fields of Terrestrial Biophysical Parameters from Satellite
Data. *Journal of Climate*, *9*(4), 706–737. https://doi.org/10.1175/1520-
0442(1996)009<0706:ARLSPF>2.0.CO;2

Sen, P. K. (1968). Estimates of the Regression Coefficient Based on Kendall ' s Tau Pranab
    Kumar Sen. *Journal of the American Statistical Association, 63*(324), 1379–1389. Retrieved
    from http://www.jstor.org/stable/10.2307/2285891

Shi, X., Déry, S. J., Groisman, P. Y., & Lettenmaier, D. P. (2013). Relationships between recent
    pan-arctic snow cover and hydroclimate trends. *Journal of Climate*, *26*(6), 2048–2064.
    https://doi.org/10.1175/JCLI-D-12-00044.1

Stewart, I. T., Cayan, D. R., & Dettinger, M. D. (2005). Changes toward Earlier Streamflow
    Timing across Western North America. *Journal of Climate*, *18*(8), 1136–1155.
    https://doi.org/10.1175/JCLI3321.1

Stieglitz, M., Ducharne, A., Koster, R., & Suarez, M. (2001). The Impact of Detailed Snow
    Physics on the Simulation of Snow Cover and Subsurface Thermodynamics at Continental
    Scales. *Journal of Hydrometeorology*, *2*(3), 228–242. https://doi.org/10.1175/1525-
    7541(2001)002<0228:TIODSP>2.0.CO;2

Sun, S., Jin, J., & Xue, Y. (1999). A simple snow-atmosphere-soil transfer model. *Journal Of
    Geophysical Research*, *104*, 587–597.

Tan, A., Adam, J. C., & Lettenmaier, D. P. (2011). Change in spring snowmelt timing in
    Eurasian Arctic rivers. *Journal of Geophysical Research*, *116*(D3), D03101.
    https://doi.org/10.1029/2010JD014337

Theil, H. (1950). A rank-invariant method of linear and polynomial regression analysis I.
    Nederlandse Akademie Wetenchappen, *53*(Series A), 386–392. https://doi.org/10.1007/978-
    94-011-2546-8

U.S. Department of Interior, Bureau of Reclamation. (1983). *DRAFT Colorado River Simulation
    System Hydrology Data Base*.

Udall, B., & Overpeck, J. (2017). The twenty-first century Colorado River hot drought and

implications for the future. *Water Resources Research*, 1–15.

https://doi.org/10.1002/2016WR019638

Vano, J. A., Das, T., & Lettenmaier, D. P. (2012). Hydrologic Sensitivities of Colorado River

Runoff to Changes in Precipitation and Temperature*. *Journal of Hydrometeorology*, *13*(3),

932–949. https://doi.org/10.1175/JHM-D-11-069.1

Vano, J. A., Udall, B., Cayan, D. R., Overpeck, J. T., Brekke, L. D., Das, T., et al. (2014).

Understanding Uncertainties in Future Colorado River Streamflow. *Bulletin of the*

American Meteorological Society, *95*(1), 59–78. https://doi.org/10.1175/BAMS-D-12-

00228.1

Waliser, D., Kim, J., Xue, Y., Chao, Y., Eldering, A., Fovell, R., et al. (2011). Simulating cold

season snowpack: Impacts of snow albedo and multi-layer snow physics. *Climatic Change*,

109(SUPPL. 1), 95–117. https://doi.org/10.1007/s10584-011-0312-5

Wang, X., & Shen, S. S. (1999). Estimation of spatial degrees of freedom of a climate field.

Journal of Climate, *12*(5 I), 1280–1291. https://doi.org/10.1175/1520-

0442(1999)012<1280:EOSDOF>2.0.CO;2

Werner, K., & Yeager, K. (2013). Challenges in Forecasting the 2011 Runoff Season in the

Colorado Basin. *Journal of Hydrometeorology*, *14*(4), 1364–1371.

https://doi.org/10.1175/JHM-D-12-055.1

Woodhouse, C. A., Pederson, G. T., Morino, K., Mcafee, S. A., & Mccabe, G. J. (2016).

Increasing influence of air temperature on upper Colorado River streamflow, 1–8.

https://doi.org/10.1002/2015GL067613.Received

Wu, H., Kimball, J. S., Li, H., Huang, M., Leung, L. R., & Adler, R. F. (2012). A new global

river network database for macroscale hydrologic modeling. *Water Resources Research*, *48*(9), n/a-n/a. https://doi.org/10.1029/2012WR012313

Xia, Y., Mocko, D., Huang, M., Cai, X., Rodell, M., Mitchell, K. E., et al. (2016). Comparison and Assessment of Three Advanced Land Surface Models in Simulating Terrestrial Water Storage Components over the United States. *Journal of Hydrometeorology*, *18*(3), 625–649. https://doi.org/10.1175/jhm-d-16-0112.1

Xiao, M., Nijssen, B., & Lettenmaier, D. P. (2016). Drought in the Pacific Northwest, 1920-2013. *Journal of Hydrometeorology*, (2014), JHM-D-15-0142.1. https://doi.org/10.1175/JHM-D-15-0142.1

Xue, Y., Sellers, P. J., Kinter, J. L., & Shukla, J. (1991). A Simplified biosphere model for global climate studies. *Journal of Climate*. https://doi.org/10.1175/1520-0442(1991)004<0345:ASBMFG>2.0.CO;2

Xue, Yongkang, Sun, S., Kahan, D. S., & Jiao, Y. (2003). Impact of parameterizations in snow physics and interface processes on the simulation of snow cover and runoff at several cold region sites. *Journal of Geophysical Research*, *108*(D22), 8859. https://doi.org/10.1029/2002JD003174

Xue, Yuan, Forman, B. A., & Reichle, R. H. (2018). Estimating Snow Mass in North America Through Assimilation of Advanced Microwave Scanning Radiometer Brightness Temperature Observations Using the Catchment Land Surface Model and Support Vector Machines. *Water Resources Research*, *54*(9), 6488–6509. https://doi.org/10.1029/2017WR022219