Title
Flooding and atmospheric rivers in coastal Western U.S. watersheds: The role of hydrological initial conditions in a changing climate

Author
Cao, Qian

Publication Date
2020

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA

Los Angeles

Flooding and atmospheric rivers in coastal Western U.S. watersheds: The role of hydrological initial conditions in a changing climate

A dissertation submitted in partial satisfaction of the requirements for the degree Doctor of Philosophy in Geography

by

Qian Cao

2020

ABSTRACT OF THE DISSERTATION

Flooding and atmospheric rivers in coastal Western U.S. watersheds:

The role of hydrological initial conditions in a changing climate

by

Qian Cao

Doctor of Philosophy in Geography

University of California, Los Angeles, 2020

Professor Dennis P. Lettenmaier, Chair

A body of work over the last several decades has demonstrated that most major floods along the U.S. West Coast are attributable to atmospheric rivers (ARs). Recent studies suggest that observed changes in extreme precipitation associated with a general warming have not necessarily lead to corresponding changes in floods, and changes in antecedent hydrological conditions could be a primary causal mechanism. This study examines climate change impacts on AR-related floods and their modulation by antecedent soil moisture (ASM) conditions in three watersheds that form a transect along the U.S. Pacific Coast: the Chehalis River basin in Washington State, the Russian River basin in Northern California, and the Santa Margarita River basin in Southern California. All three basins are rain-dominant and frequented by ARs. I used the Distributed

Hydrology-Soil-Vegetation Model (DHSVM), a spatially distributed hydrological model, to reconstruct floods in all three basins.

For historical AR flooding, based on a combination of observed and model-simulated soil moisture in the Russian River basin, my results show that the storm runoff-precipitation ratio during extreme precipitation events is much more strongly related to ASM than to storm total precipitation. If ASM is low, extreme precipitation may not lead to extreme discharge. When I used Global Climate Model (GCM)-projected future atmospheric forcings (primarily precipitation and temperature) in the three basins, my results show that the projected fraction of AR-related extreme discharge events slightly decreases in the Chehalis basin, but increases in the Russian and Santa Margarita River basins. These changes in California are driven by increases in AR-related extreme precipitation events, as well as projected increases in year-to-year volatility of annual precipitation, which increases the likelihood of concurrent occurrence of large storms and wet ASM. I also investigated the subseasonal forecast skill of AR-related flooding using the NOAA/Climate Testbed Subseasonal Experiment (SubX) database, applied to both AR- and non-AR related floods. I found that flood forecast skill drops quickly after week 1. There is some probabilistic forecast skill in week 2, but only a hint of skill in weeks 3-4, especially for annual maximum floods, notwithstanding some probabilistic skill for smaller floods in weeks 3-4.

The dissertation of Qian Cao is approved.

Yongkang Xue

Thomas Welch Gillespie

Steven Adam Margulis

Fred Martin Ralph

Dennis P. Lettenmaier, Committee Chair

University of California, Los Angeles

2020

This dissertation is dedicated to my parents and my aunt.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

I would like to thank my advisor Prof. Dennis Lettenmaier for his guidance, support, and encouragement throughout my studies at the University of Washington and UCLA. I have learned much from his insightful and invaluable instructions. More so, his actions speak louder than his words. He arrives on campus before 8 am every day over all the years. He shows me that to pursue an academic career, one needs not only curiosity and enthusiasm but also self-discipline and dedication.

I would like to acknowledge my thesis committee members Yongkang Xue, Tom Gillespie, Steve Margulis, and Marty Ralph for their support. I have also learned from other faculty members in our department. I would like to thank especially Daniela Cusack, Eric Sheppard, Glen MacDonald, and Greg Okin for their instruction. For my dissertation work, I received help as well from Sasha Gershunov, Tamara Shulgina, and Mike DeFlorio at the Scripps Institution of Oceanography, and from Shrad Shukla at UC Santa Barbara as well.

I would like to thank each and every one of my group members. I am grateful that I was able to start the first year of my Ph.D. program at the University of Washington with instruction from Ning Sun (now at Pacific Northwest National Laboratory), John Yearsley, and Bart Nijssen. I also want to thank Yixin Mao (now at Salesforce) and Xiaodong Chen (now at PNNL) for their help during the first year of my Ph.D. studies at UW. I also thank Liz Clark (now at King County Surface Water Management) and Ali Mehran (now at the University of North Georgia) for their help during my first several years at UCLA. I would like to thank Mu Xiao especially from whom I have learnt so much. I would also like to thank other current UCLA Land Surface Hydrology Research

Group members Dongyue Li, Kim Wang, Ruth Engel, Solomon Vimal, Zhaoxin Ban, Lu Su, and Emilie Tarouilly, who make our lab a very supportive and comfortable place. Each of them is so unique. Outside of our group, I thank my friends Yuxiao Niu, Huilin Huang, Yufei Liu, and especially Ting Yuan for their support. I also thank our department staff for their help, especially Kasi McMurray and Brian Won.

Finally, I especially owe great thanks to my father Hua Cao, my mother Shiju Yu, and my aunt Shirong Yu who brought me up, as well as all the rest family members. My father teaches me the value of hard work, and my mother shows me the strength of optimism. They are the most amazing and kindhearted people in my world. Without their continuous encouragement and unconditional support, I would not be who I am today.

VITA

EDUCATION

C.Phil.	Geography, University of California, Los Angeles, 2016
M.S.	Water Resources and Hydropower Engineering, Wuhan University, 2013
B.E.	Agricultural Water Resources Engineering, Wuhan University, 2011

PROFESSIONAL EXPERIENCE

Graduate Research Assistant: University of California, Los Angeles	2014-current
Graduate Research Assistant: University of Washington	2013-2014

REFEREED PUBLICATIONS

Published:

Cao, Q., A. Mehran, F.M. Ralph, and D.P. Lettenmaier, 2019: The role of hydrological initial conditions on Atmospheric River floods in the Russian River basin, *Journal of Hydrometeorology*, 20, doi: 10.1175/JHM-D-19-0030.1

Cao, Q., E.A. Clark, Y. Mao, and D.P. Lettenmaier, 2019: Trends and interannual variability in terrestrial water storage over the eastern United States, 2003-2016. *Water Resources Research*, 55, doi: 10.1029/2018WR023278

Cao, Q., T.H. Painter, W.R. Currier, J.D. Lundquist, and D.P. Lettenmaier, 2018: Estimation of Precipitation over the OLYMPEX Domain during Winter 2015/16. *Journal of Hydrometeorology*, 19, 143-160, doi: 10.1175/JHM-D-17-0076.1

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's Central Valley lose during the 2012–2016 drought? *Geophysical Research Letters*, 44, 4872-4879, doi: 10.1002/2017GL073333.

Sun N., J. Yearsley, M. Baptiste, **Q. Cao**, D.P. Lettenmaier, and B. Nijssen, 2016: A spatially distributed model for assessment of the effects of changing land use and climate on urban stream quality, *Hydrological. Processes*, 30, 4779-4798, doi: 10.1002/hyp.10964.

Cusack, D.F., J. Karpman, D. Ashdown, **Q. Cao**, M. Ciochina, S. Halterman, S. Lydon, and A. Neupane (2016), Global change effects on humid tropical forests: Evidence for biogeochemical and biodiversity shifts at an ecosystem scale, *Review of Geophysics*, 54, 523-610, doi: 10.1002/2015RG000510.

Cao, Q., N. Sun, J. Yearsley, B. Nijssen, and D.P. Lettenmaier, 2016: Climate and land cover effects on the temperature of Puget Sound streams. *Hydrological. Processes*, 30, 2286–2304, doi: 10.1002/hyp.10784.

Henn, B., **Q. Cao**, D.P. Lettenmaier, C.S. Magirl, C. Mass, J. Brent Bower, M.S. Laurent, Y. Mao, and S. Perica, 2015: Hydroclimatic Conditions Preceding the March 2014 Oso Landslide. *Journal of Hydrometeorology*, 16, 1243–1249, doi: 10.1175/JHM-D-15-0008.1.

Submitted or in preparation:

Cao, Q., Shraddhanand Shukla, Michael J. DeFlorio, F. Martin Ralph, and Dennis P. Lettenmaier, 2020: Evaluation of the subseasonal forecast skill of atmospheric river floods in coastal Western U.S. watersheds. *Journal of Hydrometeorology*, (*in prep*).

Hatchett, B.J., **Q. Cao**, P.B. Dawson, C.J. Ellis, C.W. Hecht, and Coauthors, 2020: Observations of an extreme atmospheric river storm with a diverse sensor network. *Earth and Space Science*, https://doi.org/10.1002/essoar.10502163.1, (in review).

Cao, Q., A. Gershunov, T. Shulgina, F.M. Ralph, N. Sun, and D.P. Lettenmaier, 2020: Floods due to atmospheric rivers along the U.S. West Coast: The role of hydrological initial conditions in a warming climate. *Journal of Hydrometeorology*, (in revision).

SELECTED PRESENTATIONS

Cao, Q., Shraddhanand Shukla, Michael J. DeFlorio, F. Martin Ralph, and Dennis P. Lettenmaier: Evaluation of the subseasonal forecast skill of atmospheric river floods along the coastal Western U.S. Oral presentation, AGU Fall Meeting, San Francisco, CA, Dec 2019

Su, L., **Q. Cao**, M. Xiao, D.P. Lettenmaier, D. Li, M. Barlage, D.M. Mocko: Drought Variability and Trends over the Conterminous United States over the Past Century. Oral presentation, AGU Fall Meeting, San Francisco, CA, Dec 2019

Cao, Q., E.A. Clark, Y. Mao, and D.P. Lettenmaier: Trends and Interannual Variability in Terrestrial Water Storage Over the Eastern United States, 2003–2016. Oral presentation, NLDAS teleconference (Invited), March 2019

Cao, Q., A. Gershunov, T. Shulgina, A. Mehran, F.M. Ralph, and D.P. Lettenmaier: The role of hydrological initial conditions on Atmospheric River floods along the U.S. West Coast in a warming climate. Oral presentation, AGU Fall Meeting, Washington, D.C., Dec 2018

Cao, Q., A. Mehran, F.M. Ralph, F. Cannon, and D.P. Lettenmaier: The role of hydrological initial conditions on Atmospheric River floods in the Russian River basin. Poster session, AGU Fall Meeting, New Orleans, LA, Dec 2017

Cao, Q., D.P. Lettenmaier, T.H. Painter, J.D. Lundquist, and W.A. Peterson: Estimation of precipitation over the OLYMPEX domain during winter 2015-2016 using radar, gauge precipitation and ASO snow estimates. Oral presentation, AGU Fall Meeting, San Francisco, CA, Dec 2016

AWARDS & HONORS

2019 UCLA Department of Geography Graduate Student Publication Award
2018 UCLA Department of Geography Conference Travel Stipend
2016 UCLA Department of Geography Graduate Student Publication Award
2013 Outstanding Master's Student in Wuhan University
2011 Second Prize of China National Post-Graduate Mathematic Contest in Modeling
2011 Outstanding Graduates in Wuhan University
2010 China National Undergraduate Scholarship

Chapter 1. Introduction

1.1 Background

Floods are a pervasive natural hazard. Over the last 40 years, floods lead to an average annual property damage of $1.1 billion in the western U.S., and atmospheric rivers (ARs) account for 84% of these damages (Corringham et al., 2019). It is now well understood that most of the extreme precipitation events that lead to major floods along the U.S. Pacific Coast are associated with ARs (e.g. Ralph et al., 2006; Dettinger et al., 2011; Neiman et al., 2011; Barth et al, 2017). ARs are long, narrow, and transient corridors of anomalously strong horizontal water vapor transport (Zhu and Newell, 1998; Ralph et al., 2018). Landfalling ARs greatly impact the hydrology of the U.S. West Coast (USWC) particularly where they interact with complex topography and lead to copious amounts of orographic precipitation (Gershunov et al., 2017). In recent years, an increasing number of studies have quantified how ARs move water vapor with the help of ground or satellite-based observations. At the same time, a number of AR detection methods have been developed and improved (e.g. Ralph et al., 2004, 2005, 2006; Neiman et al., 2009; Dettinger, 2011; Ralph and Dettinger, 2011; Lavers et al., 2012; Rutz et al. 2014; Guan and Waliser, 2015; Gershunov et al., 2017; Guan et al. 2018). They are now primarily based on integrated water vapor (IWV) and IWV transport (IVT) (Shields et al., 2018), quantities with more potential forecast improvement than for prediction of precipitation directly (e.g. Lavers et al., 2016).

With the advent of AR detection tools, it is now possible to identify decades-long historical sequences of AR events and their landfall locations, at regional and global scales, through application of AR detection algorithms to reanalysis datasets, and in turn

to examine relationships with flood events. For example, Lamjiri et al. (2017) found that landfalling ARs account for 60-100% of extreme storms (with precipitation-total return intervals greater than 2 years) along the USWC. Similarly, by applying AR detection methods to Global Climate Model (GCM) output, the climate change impact on AR land-falling activities can be examined, and changes in extreme precipitation associated with ARs can be projected. For example, the number of landfalling AR days is projected to substantially increase by the end of 21st century along the USWC (e.g. Espinoza et al., 2018).

Previous studies suggest that AR landfalls, AR-related storms, and runoff response to ARs along the USWC show strong seasonal and geographic signatures. Gershunov et al. (2017) examined the climate-scale variability of AR landfalls during the period 1948-2017. They showed that AR landfalls have a marked seasonal progression from the Pacific Northwest in the late fall to northern California in early winter. However, climatological AR intensities do not entirely follow AR frequency, with peak intensity occurring in December in far Northern California (Gershunov et al., 2017). Lamjiri et al. (2017) examined AR-related storm characteristics over the conterminous U.S. (CONUS) using gridded hourly precipitation observations from 1948 to 2002. They found that storm precipitation totals along the USWC, in contrast to the southeastern U.S., were modulated more by storm durations than hourly intensities, indicating the importance of AR persistence to extreme storms. They also found that average storm duration generally decreases from the north to the south along the USWC. Konrad and Dettinger (2017) examined the runoff response to ARs over the Western U.S. during 1949-2015 using U.S. Geological Survey (USGS) daily mean stream gauge records. They found that high

runoff occurred mostly during ARs and was restricted to the coastal ranges and the western Sierra Nevada and Cascade Ranges, reflecting the spatial distribution of ARs and their increasing high-elevation intensity terrain (Konrad and Dettinger, 2017).

One primary goal of AR studies is to better predict extreme precipitation events associated with ARs and thus to better predict flooding. However, it is well known that antecedent soil moisture (ASM) conditions are a key factor linking precipitation and runoff. Previous studies used in situ observations to examine the critical threshold of ASM to differentiate high or low runoff ratios (the amount of storm runoff divided by the amount of storm precipitation), but limited to very small catchments (e.g. Penna et al., 2011; Radatz et al., 2013). Using satellite-based soil moisture, Crow et al. (2017) examined the relationship between ASM and observed storm runoff ratios in the South Central U.S. They found that the runoff ratio was more strongly correlated to ASM than storm total precipitation.

Some studies have examined the impact of ASM on AR-related flooding, but mostly for a single event (e.g. Leung and Qian, 2009; Neiman et al., 2014). Ralph et al. (2013b) examined the impact of ASM, from in situ observations, on AR-related runoff events in a sub-basin (drainage area 163 km^2) of California's Russian River Basin from 2004 to 2010 when there were soil moisture observations. They found that AR-induced heavy precipitation did not lead to significant streamflow when ASM was dry and that extreme floods usually occurred when ASM was quite wet.

However, observation-based analysis alone is limited by its spatial and/or temporal coverage. Simulations from hydrologic models are also essential to better understand the role of ASM in (long-term) historic and especially future flood events. Recent studies

suggest that observed changes in extreme precipitation associated with a general warming have not necessarily lead to commensurate changes in floods, and changes in antecedent hydrological conditions could be a primary missing link (Sharma et al., 2018). For example, Wasko and Sharma (2017) used station observations to examine the sensitivity of extreme daily precipitation and discharge to changes in daily temperature. They found that changes in heavy rainfall events linked with observed warming did not lead to similar changes in streamflow in most regions globally possibly due to initial moisture conditions. Although the number of AR-related extreme precipitation events is projected to increase, ASM can have a large impact on AR-related floods especially in a warming climate.

Subseasonal weather forecasts (1-4 week lead times) have not received much attention until recently. The past several years have witnessed a joint effort from the weather and climate communities to bridge the weather-climate prediction gap at subseasonal to seasonal lead times (Mariotti et al., 2018). A few subseasonal weather forecast databases have been developed, such as the NOAA/Climate Testbed Subseasonal Experiment (SubX) project (Pegion et al., 2019). The subseasonal prediction skill of ARs has been evaluated in both forecasts (DeFlorio et al., 2019) and reforecasts (e.g. DeFlorio et al., 2018; Mundhenk et al., 2018; Nardi et al., 2018). Yet, the usefulness of these databases on AR-related flooding has not been examined despite its considerable socioeconomic value.

1.2 Science questions

Here, I seek to examine climate change impacts on AR-related floods and their modulation by hydrological initial conditions in coastal Western U.S. watersheds. The science questions I intend to address are:

1) What is the role of hydrologic initial conditions on the interaction between surface climate forcings associated with ARs and the flood response at river basin scale in the current (historical) climate?

2) How will climate change impact the role of hydrologic initial conditions on AR-related floods along the U.S. West Coast?

3) What is the subseasonal forecast skill (at 1-4 week lead times) of AR-related floods in coastal Western U.S. watersheds?

I address these questions are addressed in the following three chapters. I select three coastal watersheds frequented by ARs as my study domain: the Chehalis River basin in Washington State, the Russian River basin in Northern California, and the Santa Margarita River basin in Southern California. They form a transect along the U.S. Pacific Coast and their locations reflect seasonal and geographic signatures of AR landfalls. Besides, all three basins are rain-dominant. I implement the Distributed Hydrology-Soil-Vegetation Model (DHSVM) (Wigmosta et al., 1994) in each of the three basins. Below, I summarize the three core chapters of this dissertation (Chapters 2-4), which finishes with overall conclusions in Chapter 5.

Chapter 2 (published as Cao et al., 2019) focuses on the Russian River basin, which is the site of NOAA's Hydrometeorology Testbed (HMT) soil moisture observation network (Ralph et al., 2013a) and has long-term records. This chapter

examines the interaction of ASM with AR storm characteristics in the current (historical) climate, addressing Question 1 using both observation- and simulation-based analyses. **Chapter 3** (in revision as Cao et al., 2020a) examines how those relationships will change in the future across three basins based on model runs driven by downscaled Global Climate Model (GCM) forcings, which addresses Question 2. Question 3 is addressed in **Chapter 4** (to be submitted to *Journal of Hydrometeorology* as Cao et al., 2020b) via a series of modeling experiments using downscaled NOAA's SubX forcings.

References

Barth, N. A., G. Villarini, M. A. Nayak, and K. White, 2017: Mixed populations and annual flood frequency estimates in the western United States: The role of atmospheric rivers. *Water Resour. Res.*, **53**, 257-269. https://doi.org/10.1002/2016wr019064.

Crow, W. T., F. Chen, R. H. Reichle, and Q. Liu, 2017: L band microwave remote sensing and land data assimilation improve the representation of prestorm soil moisture conditions for hydrologic forecasting. *Geophys. Res. Lett.*, **44**, 5495-5503, https://doi.org/10.1002/2017gl073642.

Cao, Q., A. Mehran, F.M. Ralph, and D.P. Lettenmaier, 2019: The Role of Hydrological Initial Conditions on Atmospheric River Floods in the Russian River Basin. *J. Hydrometeor.*, **20**, 1667–1686, https://doi.org/10.1175/JHM-D-19-0030.1.

——, A. Gershunov, T. Shulgina, F.M. Ralph, N. Sun, and D.P. Lettenmaier, 2020a: Floods due to atmospheric rivers along the U.S. West Coast: The role of hydrological initial conditions in a warming climate. *Journal of Hydrometeorology*, (in revision).

——, Shraddhanand Shukla, Michael J. DeFlorio, F. Martin Ralph, and Dennis P. Lettenmaier, 2020b: Evaluation of the subseasonal forecast skill of atmospheric river floods in coastal Western U.S. watersheds. *Journal of Hydrometeorology*, (*in prep*).

Corringham, T.W., F.M. Ralph, A. Gershunov, D.R. Cayan, and C.A. Talbot, 2019: Atmospheric Rivers Drive Flood Damages in the Western United States. *Sci. Adv.*, **5**, 2375-2548, https://doi.org/10.1126/sciadv.aax4631

DeFlorio, M.J., D.E. Waliser, B. Guan, F.M. Ralph, and F. Vitart, 2018: Global evaluation of atmospheric river subseasonal prediction skill. *Climate Dyn.*, **52**, 3039–3060, https://doi.org/10.1007/s00382-018-4309-x.

——, and coauthors, 2019: Experimental Subseasonal-to-Seasonal (S2S) Forecasting of Atmospheric Rivers Over the Western United States. *J. Geophys. Res. Atmos*., 124. https://doi.org/10.1029/2019JD031200

Dettinger, M., 2011: Climate Change, Atmospheric Rivers, and Floods in California - A Multimodel Analysis of Storm Frequency and Magnitude Changes. *J. Am. Water Resour. Assoc.*, **47**, 514-523, https://doi.org/10.1111/j.1752-1688.2011.00546.x.

——, F. M. Ralph, T. Das, P. Neiman, and D. Cayan, 2011: Atmospheric rivers, floods and the water resources of California. *Water*, **3**, 445–478, https://doi.org/10.3390/w3020445.

Espinoza, V., D. E. Waliser, B. Guan, D. A. Lavers, and F. M. Ralph, 2018: Global analysis of climate change projection effects on atmospheric rivers. *Geophys. Res. Lett*., **45**, 4299–4308. https://doi.org/10.1029/ 2017GL076968.

Gershunov, A., T. Shulgina, F. M. Ralph, D. A. Lavers, and J. J. Rutz, 2017: Assessing the climate-scale variability of atmospheric rivers affecting western North America. *Geophys. Res. Lett*., **44**, 7900-7908, https://doi.org/10.1002/2017gl074175.

Guan, B., and D. E. Waliser, 2015: Detection of atmospheric rivers: Evaluation and application of an algorithm for global studies. *J. Geophys. Res.: Atmos*., **120**, 12514-12535, https://doi.org/10.1002/2015jd024257.

——, D.E. Waliser, and F.M. Ralph, 2018: An Intercomparison between Reanalysis and Dropsonde Observations of the Total Water Vapor Transport in Individual

Atmospheric Rivers. *J. Hydrometeor*., **19**, 321–337, https://doi.org/10.1175/JHM-D-17-0114.1

Konrad, C. P., and M.D. Dettinger, 2017: Flood runoff in relation to water vapor transport by atmospheric rivers over the western United States, 1949–2015. *Geophys. Res. Lett*., **44**, 11,456–11,462, https://doi.org/10.1002/2017GL075399.

Lamjiri, M. A., M. D. Dettinger, F. M. Ralph, and B. Guan, 2017: Hourly storm characteristics along the US West Coast: Role of atmospheric rivers in extreme precipitation. *Geophys. Res. Lett*., **44**, 7020-7028, https://doi.org/ 10.1002/2017gl074193.

Lavers, D. A., G. Villarini, R. P. Allan, E. F. Wood, and A. J. Wade, 2012: The detection of atmospheric rivers in atmospheric reanalyses and their links to British winter floods and the large-scale climatic circulation. *J. Geophys. Res.: Atmos*., **117**, 13, https://doi.org/10.1029/2012jd018027.

——, D. E. Waliser, F. M. Ralph, and M. D. Dettinger, 2016: Predictability of horizontal water vapor transport relative to precipitation: Enhancing situational awareness for forecasting western US extreme precipitation and flooding. *Geophys. Res. Lett*., **43**, 2275-2282, https://doi.org/10.1002/2016gl067765.

Leung, L. R., and Y. Qian, 2009: Atmospheric rivers induced heavy precipitation and flooding in the western US simulated by the WRF regional climate model. *Geophys. Res. Lett*., **36**, 6, https://doi.org/10.1029/2008gl036445.

Mariotti, A., P. M. Ruti, and M. Rixen, 2018: Progress in subseasonal to seasonal prediction through a joint weather and climate community effort. *npj Climate Atmos. Sci*., **1**, 4, https://doi.org/10.1038/s41612-018-0014-z.

Mundhenk, B.D., E.A. Barnes, E.D. Maloney, and C.F. Baggett, 2018: Skillful empirical subseasonal prediction of landfalling atmospheric river activity using the Madden-Julian oscillation and quasi-biennial oscillation. *npj Climate Atmos. Sci*., **1**, 20177. https://doi.org/10.1038/s41612-017-0008-2

Nardi, K.M., E.A. Barnes, and F.M. Ralph, 2018: Assessment of Numerical Weather Prediction Model Reforecasts of the Occurrence, Intensity, and Location of Atmospheric Rivers along the West Coast of North America. *Monthly Weather Review*, **146**, 3343–3362, https://doi.org/10.1175/MWR-D-18-0060.1

Neiman, P. J., A. B. White, F. M. Ralph, D. J. Gottas, and S. I. Gutman, 2009: A water vapour flux tool for precipitation forecasting. *P. I. Civil Eng.-Wat. M*., **162**, 83-94, https://doi.org/ 10.1680/wama.2009.162.2.83.

——, L. J. Schick, F. M. Ralph, M. Hughes, and G. A. Wick, 2011: Flooding in Western Washington: The Connection to Atmospheric Rivers. *J. Hydrometeor.*, **12**, 1337-1358, https://doi.org/10.1175/2011jhm1358.1.

——, F. M. Ralph, B. J. Moore, and R. J. Zamora, 2014: The Regional Influence of an Intense Sierra Barrier Jet and Landfalling Atmospheric River on Orographic Precipitation in Northern California: A Case Study. *J. Hydrometeor.*, **15**, 1419-1439, https://doi.org/10.1175/jhm-d-13-0183.1.

Pegion, K., and coauthors, 2019: The Subseasonal Experiment (SubX): A multi-model subseasonal prediction experiment. *Bull. Amer. Meteor. Soc*., **100**, 2043–2060, https://doi.org/10.1175/BAMS-D-18-0270.1

Penna, D., H. J. Tromp-van Meerveld, A. Gobbi, M. Borga, and G. Dalla Fontana, 2011: The influence of soil moisture on threshold runoff generation processes in an alpine

headwater catchment. *Hydrol. Earth Syst. Sci.*, **15**, 689-702, https://doi.org/10.5194/hess-15-689-2011.

Radatz, T. F., A. M. Thompson, and F. W. Madison, 2013: Soil moisture and rainfall intensity thresholds for runoff generation in southwestern Wisconsin agricultural watersheds. *Hydrol. Processes*, **27**, 3521-3534, https://doi.org/10.1002/hyp.9460.

Ralph, F.M., P.J. Neiman, and G. Wick, 2004: Satellite and CALJET aircraft observations of atmospheric rivers over the eastern north pacific ocean during the winter of 1997/98. *Mon. Wea. Rev.*, **132**, 1721-1745, https://doi.org/10.1175/1520-0493(2004)132<1721:SACAOO>2.0.CO;2.

——, ——, and R. Rotunno, 2005: Dropsonde observations in low-level jets over the northeastern Pacific Ocean from CALJET-1998 and PACJET-2001: Mean vertical-profile and atmospheric-river characteristics. *Mon. Wea. Rev.*, **133**, 889-910, https://doi.org/10.1175/MWR2896.1.

——, ——, G. Wick, S. Gutman, M. Dettinger, D. Cayan, and A. White, 2006: Flooding on California's Russian River: Role of atmospheric rivers. *Geophys. Res. Lett*., **33**, L13801, https://doi.org/10.1029/2006GL026689.

——, and M. Dettinger, 2011: Storms, floods, and the science of atmospheric rivers. *Eos Trans. AGU*, **92**, 265, https://doi.org/10.1029/2011EO320001.

——, and Coauthors, 2013a: The Emergence of Weather-Related Test Beds Linking Research and Forecasting Operations. *Bull. Amer. Meteor. Soc*., 94, 1187–1211, https://doi.org/10.1175/BAMS-D-12-00080.1

——, T. Coleman, P. Neiman, R. Zamora, and M. Dettinger, 2013b: Observed Impacts of Duration and Seasonality of Atmospheric-River Landfalls on Soil Moisture and

Runoff in Coastal Northern California. *J. Hydrometeor.*, **14**, 443-459, https://doi.org/10.1175/JHM-D-12-076.1.

——, M. D. Dettinger, M. M. Cairns, T. J. Galarneau, and J. Eylander, 2018: DEFINING "ATMOSPHERIC RIVER" How the Glossary of Meteorology Helped Resolve a Debate. *Bull. Amer. Meteor. Soc.*, **99**, 837-839, https://doi.org/10.1175/bams-d-17-0157.1.

Rutz, J. J., W. J. Steenburgh, and F. M. Ralph, 2014: Climatological Characteristics of Atmospheric Rivers and Their Inland Penetration over the Western United States. *Mon. Wea. Rev.*, **142**, 905-921, https://doi.org/ 10.1175/mwr-d-13-00168.1.

Sharma, A., C. Wasko, and D. P. Lettenmaier, 2018: If precipitation extremes are increasing, why aren't floods? *Water Resour. Res.*, **54**. https://doi.org/10.1029/2018WR023749.

Shields, C.A., and coauthors, 2018: Atmospheric River Tracking Method Intercomparison Project (ARTMIP): Project goals and experimental design. *Geosci. Model Dev.*, **11**, 2455-2474, https://doi.org/10.5194/gmd-11-2455-2018.

Wasko, C., and A. Sharma, 2017: Global assessment of flood and storm extremes with increased temperatures. *Sci. Rep.*, **7**, 8, https://doi.org/10.1038/s41598-017-08481-1

Wigmosta, M. S., L. W. Vail, and D. P. Lettenmaier, 1994: A distributed hydrology-vegetation model for complex terrain. *Water Resour. Res.*, **30**, 1665-1679, https://doi.org/10.1029/94WR00436.

Zhu, Y., and R. Newell, 1998: A proposed algorithm for moisture fluxes from atmospheric rivers. *Mon. Wea. Rev.*, **126**, 725-735, https://doi.org/10.1175/1520-0493(1998)126<0725:APAFMF>2.0.CO;2.

Chapter 2. The Role of Hydrological Initial Conditions on Atmospheric River Floods in the Russian River Basin

This chapter has been published in its current form in the *Journal of Hydrometeorology*.
 The supplemental material for this chapter is provided in Appendix A.

Cao, Q., A. Mehran, F.M. Ralph, and D.P. Lettenmaier, 2019: The Role of Hydrological Initial Conditions on Atmospheric River Floods in the Russian River Basin. *J. Hydrometeor.*, 20, 1667–1686, https://doi.org/10.1175/JHM-D-19-0030.1

Abstract

A body of work over the last decade or so has demonstrated that most major floods along the U.S. West Coast are attributable to Atmospheric Rivers (ARs). Recent studies suggest that observed changes in extreme precipitation associated with a general warming of the Western U.S. have not necessarily led to corresponding changes in floods, and changes in antecedent hydrological conditions could be a primary missing link. Here we examine the role of antecedent soil moisture (ASM) conditions on historical AR flooding on California's Russian River Basin, a coastal watershed whose winter precipitation extremes are dominated by ARs. We examined the effect of observed warming on ASM for the period 1950-2017. We first constructed an hourly precipitation product at $1/32^{o}$ spatial resolution. We used the Distributed Hydrology-Soil-Vegetation Model (DHSVM) to estimate storm total runoff volumes and soil moisture. We found that up to 95% of Peaks Over Threshold (POT) extreme discharge events were associated with ARs. The storm runoff-precipitation ratio generally increased with wetter pre-storm conditions, and

the relationship was stronger as drainage area increased. We found no trends in extreme precipitation but weak downward trends in extreme discharge. The latter were mostly consistent with weak downward trends in the first 2-day storm precipitation. We found no trends in ASM, however, ASM was significantly correlated with peak flow. The ASM was affected more by antecedent precipitation than evapotranspiration and hence temperature increases had weak effects on ASM.

2.1 Introduction

A body of work over the last decade or so has demonstrated that most major floods along the U.S. West Coast are attributable to Atmospheric Rivers (ARs) (e.g. Ralph et al., 2006; Dettinger et al., 2011; Neiman et al., 2011; Barth et al, 2017), which are long, narrow, and transient corridors of anomalously strong horizontal water vapor transport (Zhu and Newell, 1998; Ralph et al., 2018b). Accompanied by warm air temperatures and strong low-level winds, AR landfalls may lead to enhanced precipitation when interacting with the complex topography (Neiman et al., 2002). ARs make up 30-50% of annual precipitation on the U.S. Pacific Coast and thus contribute to the region's water supply and water resources (Guan et al., 2010; Dettinger et al., 2011; Lavers and Villarini, 2015; Lamjiri et al., 2017). However, strong AR events can result in heavy precipitation and lead to disastrous flooding such as the coastal flooding in Southern California during late March 2018. Landfalling ARs account for 60-100% of extreme storms (with precipitation-total return intervals greater than 2 years) along the U.S. West Coast (Lamjiri et al., 2017) and 40–75% of extreme wind and precipitation events (exceeding the 98th percentile) over 40% of the coastlines worldwide particularly in the mid-latitudes

(Waliser and Guan, 2017). Moreover, ARs contribute more than half of the mean annual runoff and more than 80% of annual peak flows along the Pacific Northwest and Northern California coast (Paltan et al., 2017; Barth et al., 2017).

Recent years have witnessed the rapid development of AR detection methods based on integrated water vapor (IWV) and IWV transport (IVT) (Shields et al., 2018) as well as increased ability to forecast ARs (e.g. Nayak et al., 2014; Lavers et al., 2016; Cordeira et al., 2017; Deflorio et al., 2018; Martin et al., 2018). A number of studies have been carried out to examine the key factors linking ARs and precipitation, such as the AR duration (Ralph et al., 2013b; Lamjiri et al., 2017; Nayak and Villarini, 2018), (IVT) intensity, and direction with respect to regional terrain orientation (Ralph et al., 2003; Neiman et al., 2011, 2013; Hughes et al., 2014; Hecht and Cordeira, 2017). The first two factors can be used as a scale to categorize AR events and their potential hydrologic impacts at a given location (Ralph et al., 2019).

Hydrologic initial conditions play an important role in the natural links between precipitation and floods. Previous studies examined the critical threshold of antecedent soil moisture (ASM) to differentiate high or low runoff ratios (the amount of storm runoff divided by the amount of storm precipitation) based on in situ observations. For example, Penna et al. (2011) assessed the effect of ASM in the upper 30 cm layer (of clay loam/silty clay loam) on 40 runoff events with precipitation greater than 6 mm during June 2005-October 2006 in an alpine headwater catchment with a drainage area of 1.9 km^2. They found that the runoff ratio was mostly below 0.09 when soil saturation (percentage of porosity) was below 70%. Similarly, Radatz et al. (2013) found that runoff ratios were near zero when soil saturation was below 80% in the upper 30 cm layer of silt

loam during March 2004-September 2007 across 6 small agricultural watersheds in southwestern Wisconsin with areas ranging within 0.06-0.17 km^2. Due to limitations of in situ observations, these studies focused on very small catchments.

The development of satellite-based large-scale soil moisture monitoring products using L band microwave radiometry in recent years, such as the European Space Agency's Soil Moisture and Ocean Salinity (SMOS) mission (Kerr et al., 2010) and the NASA's Soil Moisture Active Passive (SMAP) mission (Entekhabi et al., 2010), have enabled studies at larger scales. Crow et al. (2017) examined the relationship between ASM from the Level-4 SMAP (SMAP_L4) product, which is based on the assimilation of SMAP brightness temperature observations into the Catchment Land Surface Model (LSM), and the storm runoff ratio from precipitation and streamflow observations in the South Central U.S. They found that the runoff ratio showed a much stronger (rank) correlation with pre-storm surface soil moisture than storm total precipitation.

Sensitivity analyses using model simulations can help to better understand the role of ASM on floods. Castillo et al. (2003) conducted a stochastic sensitivity analysis in three small catchments (0.06-0.24 km^2) in semiarid southeast Spain. By examining the simulation results from a combination of soil moisture and storm precipitation scenarios, they showed that the ASM was an important controlling factor of runoff during low and medium intensity storms but not for high intensity storms, during which floods were dominated by the infiltration-excess mechanism. A sensitivity analysis in the Fella Basin (623 km^2) in Italy showed similar results (Nikolopoulos et al., 2011). Their study used a spatially distributed hydrologic model to examine the sensitivity of flash flood response to initial wetness conditions by adjusting water table depth and thus the initial soil

moisture profile. In addition, Nikolopoulos et al. (2011) found that the sensitivity of flood response to initial wetness increased for increasing basin scale by examining the sensitivity across catchments with areas of 24, 165 and 329 km^2. Thomas et al. (2016) examined the sensitivity of simulated peak flows to antecedent soil saturation in a 45 km^2 catchment in Iowa. They found that ASM became less important to peak flows as rainfall depth increased.

Given the awareness of the importance of prestorm wetness conditions on runoff, some studies have examined the impact of ASM on AR flooding. Leung and Qian (2009) ran a 20-year simulation of regional climate with the Weather Research and Forecasting (WRF) model to examine AR-induced heavy precipitation and flooding events over the western U.S. from 1980 to 2000. They found that for two selected events with similar amounts of total precipitation, different ASM conditions could lead to a difference of more than 0.3 in the runoff ratio. Neiman et al. (2014) closely examined a single AR flood event in Northern California using in situ observations and found that flooding occurred shortly after the soil water content exceeded its field capacity during the heaviest rains at one study site. Ralph et al. (2013b) examined the impact of precursor soil moisture conditions on the streamflow in a sub-basin (drainage area 163 km^2) of the Russian River Basin in Northern California during 91 AR events from 2004 to 2010. They used hourly observations of upslope IWV flux from the local AR observatory, precipitation and streamflow from gauges, as well as soil moisture from NOAA's Hydrometeorology Testbed (HMT) (Ralph et al., 2013a). They found that AR-induced heavy precipitation did not lead to significant streamflow when precursor soil moisture

was below 20% (volumetric water content) and that extreme floods might occur when ASM was greater than about 35%.

Recent studies have examined the role of changes in soil moisture on floods in a warming world. For example, Woldemeskel and Sharma (2016) examined trends in annual maximum precipitation and accompanying antecedent soil moisture (with antecedent precipitation index (API) as a surrogate) over the past century globally. In North America, they found positive trends in annual maximum rainfall but no trend in API, which they argued partially explained the lack of positive trends in annual maximum flows in this region. Berghuijs et al. (2016) argued that soil moisture changes are particularly important in areas, such as California, where flooding is caused by a large single precipitation excess event. By analyzing historical station records of precipitation and streamflow, Wasko and Sharma (2017) found that increased heavy rainfall events caused by warming did not lead to similar increases in the streamflow in most regions globally, suggesting the importance of initial moisture conditions in these areas. They found that the intensity of extreme precipitation (exceeding the 99^{th} percentile) has decreased in the U.S. Northwest (including northern California) as temperatures have increased; this is in contrast to the general increasing pattern in the subtropics and temperate regions. Similar decreasing trends in both the annual maximum precipitation and the annual frequency of heavy precipitation in these regions were found by Mallakpour and Villarini (2017), based on the Climate Prediction Center (CPC) gridded daily precipitation product during the period 1948-2012. The extreme streamflow in these areas showed even greater negative response than extreme precipitation did to increased temperatures (Wasko and Sharma, 2017), indicating the influence of ASM conditions.

Yet, it is unclear whether the greater negative response of extreme discharge was caused by increased temperature or antecedent precipitation conditions (Sharma et al., 2018).

Given this background, we address here the following motivating questions:

1) What is the role of ASM on historical AR flooding in California's Russian River Basin?

2) How has observed warming during the period 1950-2017 affected ASM and thus extreme discharge events?

We selected the Russian River Basin as our study domain because it is a coastal watershed where AR events frequently occur. Snow is rare, so the basin's response to extreme precipitation generally is not conflated with a rain-on-snow contribution. Based on observations, Ralph et al. (2006) found that all of the seven largest floods during the period 1997-2006 in this basin were due to heavy orographic precipitation caused by ARs. Topographic variations have little effect on the spatial variability of precipitation in this basin, thus avoiding the added complexity of the influence of snowmelt on stream flow that more commonly occurs in mountainous basins. Most importantly, the Russian River basin is the site of NOAA's HMT soil moisture observation network which facilitates the evaluation of model simulations.

2.2 Study region

The Russian River Basin is located in Northern California. It is bounded by the Coast Range both to the east and west (see Figure 2.1a). It has a drainage area of about 3850 km^2 at its mouth. Elevation ranges from sea level to 1324 m at the top of Mount Saint Helena. It is a rain-dominant basin with basin-average midwinter (December,

January and February) temperatures above 7.5 ^{o}C. It has dry summers and wet winters, with over 96% of the annual total precipitation (on average over our domain) falling between October and April. Annual precipitation ranged from 423 mm in water year (WY) 1977 to 2052 mm in WY 1983 with a mean of 1061 mm during the period 1950-2017.

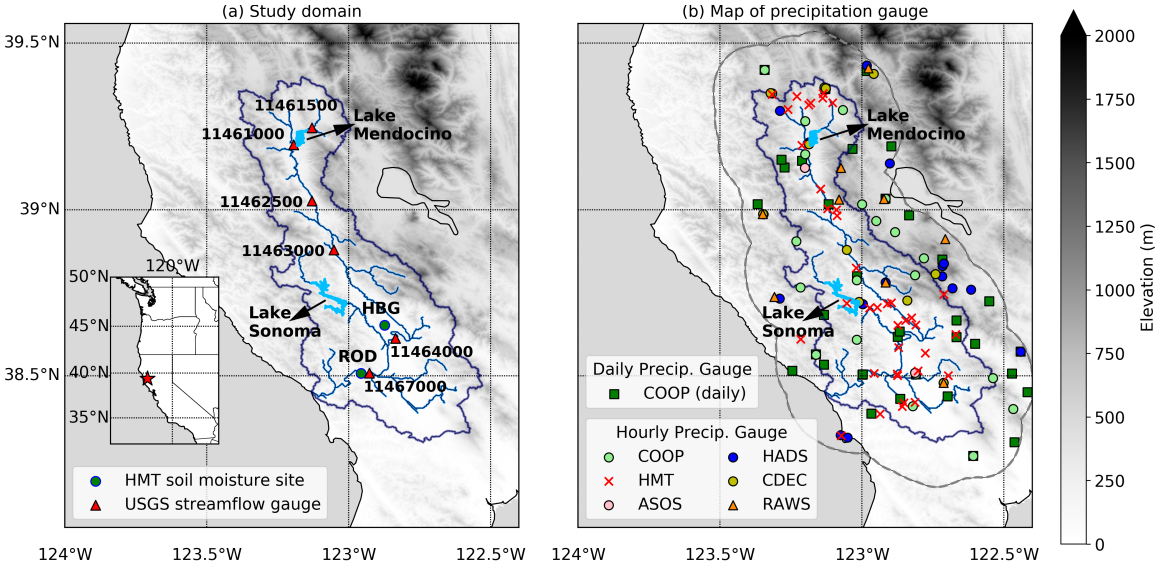

Figure 2.1 (a) Study domain, with locations of USGS stream gauges and HMT soil moisture observation sites shown, and (b) precipitation gauge locations. The 15-km buffer for the precipitation gauges selection is shown in gray dashed line in (b).

The main stem of the Russian River flows southward to its confluence with Mark West Creek north of Forestville, where it turns westward and flows into the Pacific Ocean. There are two reservoirs within the basin that are operated primarily for flood control: Coyote Dam (Lake Mendocino) on the East Fork Russian River near Ukiah, and Warm Springs Dam (Lake Sonoma) on Dry Creek west of Healdsburg (see Figure 2.1a).

2.3 Data and methods

We first constructed an hourly gridded precipitation product for the study period. We implemented the Distributed Hydrology-Soil-Vegetation Model (DHSVM)

(Wigmosta et al., 1994) over the entire basin and ran it at an hourly time step to obtain hydrographs and ASM for historical storm events during the period WY 1950-2017. We identified Peaks Over Threshold (POT) extreme precipitation events and extreme discharge events that were coincident with AR events. Using the hourly gridded precipitation and hourly simulated runoff and soil moisture, we investigated the role of ASM on historical AR flooding by examining the relationship between precursor soil moisture and storm runoff-precipitation ratios. Finally, we assessed the effect of warming using the model simulations where we detrended the meteorological forcings for long-term temperature changes.

a. Gridded hourly precipitation

We derived hourly gridded precipitation data at a spatial resolution of $1/32^{o}$ for the period WY 1950-2017. We first quality controlled the gauge data, then gridded the gauge daily precipitation, including the gauge hourly precipitation aggregated to daily, using the Mountain Mapper (MM) method (Schaake et al., 2004), and finally interpolated the gridded daily precipitation to hourly using the nearest hourly gauge.

We took daily and hourly precipitation data from NOAA's Cooperative Observer Program (COOP) network; Remote Automatic Weather Stations (RAWS); the Automated Surface Observing System (ASOS); the NOAA Hydrometeorological Automated Data System (HADS); the California Data Exchange Center (CDEC); and NOAA's Hydrometeorolgy Testbed (HMT). COOP stations have both hourly and daily precipitation records, some of which go back at least to WY 1950. We excluded COOP daily gauges without time of observation records. Gauges from the other networks all have hourly data with the earliest record going back to 1985. We selected gauges within

a buffer of 15 km from the basin and excluded gauges with overlap among different networks. We used a total of 133 gauges, including 35 COOP daily gauges, 20 COOP hourly gauges, 38 HMT gauges, 2 ASOS gauges, 18 HADS gauges, 11 CDEC gauges, and 9 RAWS gauges (see Figure 2.1b).

We performed a quality control (QC) for all hourly gauges following Cao et al. (2018). After the QC, we aggregated gauge hourly precipitation to daily. We considered daily precipitation as occurring between 0000 and 2400 Pacific Standard Time (PST; UTC-0800). Similarly, we used the aggregated daily data to QC the COOP daily gauges after we proportioned them to PST 24 h according to their observation time using the nearest hourly gauges.

We used monthly precipitation from the Parameter-Elevation Regressions on Independent Slopes Model (PRISM; Daly et al. 1994, 2008) as a background precipitation distribution map. Following the MM method, we calculated ratios between daily gauge precipitation and the PRISM monthly climatologies at station grid nodes. We interpolated the ratios onto the $1/32^{o}$ grids using the synergraphic mapping system (SYMAP) algorithm (Shepard 1984). Gridded daily precipitation were obtained by multiplication of the gridded ratios with the PRISM monthly climatology. Finally, we interpolated the gridded daily precipitation to hourly. For a given day in a grid cell, we first searched for the nearest precipitation gauge with complete and valid hourly data on that day, and then multiplied the hourly ratios (hourly precipitation divided by daily sum) by the daily precipitation in that grid cell.

b. Model implementation

In order to obtain full hydrographs of historical storm events from sub-daily data and historical soil moisture conditions, we implemented the DHSVM model (Wigmosta et al., 1994) at a spatial resolution of 150 m over the entire Russian River Basin and ran it at an hourly time step for the (WY) 1950 to 2017 period. Given that there were two regulated reservoirs within the basin and naturalized streamflow was needed for our analysis, we used a version of DHSVM that includes a reservoir module (DHSVM-res) (Zhao et al., 2016).

1) METEOROLOGICAL DRIVING DATA

In addition to precipitation data, DHSVM also requires meteorological inputs including air temperature, wind speed, relative humidity, downward solar and longwave radiation at the model's hourly time step. Similar to Cao et al. (2016), we calculated the last three using the Mountain Microclimate Simulation Model (MTCLIM) algorithms as described and implemented by Bohn et al. (2013). We took wind speed data from the lowest atmospheric level in the NCEP-NCAR reanalysis output (Kalnay et al., 1996). We obtained daily maximum and minimum temperature data from 11 COOP stations with long-term temperature records. We interpolated the station temperature anomalies to a spatial resolution of $1/32^{o}$ based on the PRISM 1981-2010 climatology. Details of the gridding approach can be found in Maurer et al. (2002) and Wood and Lettenmaier (2006).

2) MODEL EVALUATION DATA

There are 29 USGS gauges within the basin at which instantaneous (15-min) streamflow records are available, with the longest record dating from October 1987. We

selected six stream gauges with records longer than ~30 years for model calibration, including the upstream- and downstream-most ones (see Figure 2.1a). Two upstream gauges are free of reservoir effects, while the four downstream gauges are affected by reservoir regulation. We obtained hourly reservoir storage and elevation data from CDEC, with records starting from December 1988. There are 12 HMT soil moisture sites within the basin. We selected two sites, HBG and ROD, which had the longest records (starting from December 2006). The surface layer observation was at 10 cm depth.

3) REMOVING RESERVOIR EFFECTS FROM STREAMFLOW OBSERVATIONS

The reservoir module divides each reservoir into an inactive pool, a conservation pool (for water supply), a flood control pool and a surcharge pool, the elevations of which were obtained from the Sonoma County Water Agency (SCWA) for Lakes Mendocino and Sonoma. The release scheme used by DHSVM-res is based primarily on real-time water levels and predefined water demands. We estimated real-time water levels from storage through an empirical relationship based on historical CDEC records. Similarly, we estimated the real-time surface area (for the calculation of reservoir evaporation) from storage as based on CDEC storage observations and historical surface area time series from LandSat data.

We evaluated the model simulations of streamflow at gauges throughout the basin and simulations of reservoir storage. We then calculated the difference of simulated streamflow without and with the reservoir module at each stream gauge affected by the reservoir regulations. We then added the difference back to the observed streamflow in order to create an estimate of naturalized flows.

c. AR-related POT extreme events

We used the POT method to select extreme events, which samples observations above a given threshold value and considers a wider range of events than the block maxima approach (e.g. Lang et al., 1999; Begueria et al., 2011; Mallakpour and Villarini, 2017). We first selected POT extreme precipitation events and POT extreme discharge events separately based on daily observations. We then identified the extreme events that were coincident with AR events.

1) EXTREME PRECIPITATION EVENTS

For the six long-term stream gauges on which we focused, we calculated the daily precipitation at each averaged over its upstream drainage area. We selected POT extreme precipitation events based on daily precipitation data. We defined events based on consecutive days with peak daily precipitation exceeding a given threshold. We selected thresholds to result in 1, 2, 3, 5 and 7 events per year on average, which we denote as POT_{N1}, ... POT_{N7}. Events were separated from each other by at least one day with daily precipitation below the threshold value, as in Mondal and Mujumdar (2015).

We used the 6-hourly AR catalog of Gershunov et al. (2017) who applied an AR detection method based on both IWV and IVT to the National Centers for Environmental Prediction-National Center for Atmospheric Research (NCEP-NCAR) reanalysis data, starting with 1948. Ralph et al. (2018a) evaluated the performance of a diverse set of AR detection tools (ARDTs) in the Russian River Basin in comparison with the local AR observatory, including the three applied to the NCEP-NCAR data set: Rutz et al. (2014), Guan and Waliser (2015) and Gershunov et al. (2017). The three tools had similar parameters and geometric characteristics and hence showed similar performance in

capturing AR frequency, duration and intensity. All three tools inferred slightly higher AR contribution to precipitation compared with other ARDTs (Ralph et al., 2018a) due to their less stringent geometric criteria. We used the AR catalog of Gershunov et al. (2017) primarily because it spans the entire period of the NCEP-NCAR reanalysis up to near real-time. We extracted the NCEP-NCAR grid cells (at a relatively coarse 2.5^{o} spatial resolution) that intersected the Russian River Basin and identified landfalling ARs. We examined the intersection of AR events from this catalog and POT extreme precipitation events identified as above.

2) EXTREME DISCHARGE EVENTS

For the selection of POT extreme discharge events, we first applied the independence criteria from the U.S. Water Resources Council (USWRC, 1982) to daily streamflow. According to these criteria the second flood peak of two consecutive events must be rejected if:

$$N_{\text{interval}} < 5 \text{ days} + \log(A) \text{ or } X_{\min} < (3/4) \cdot \min[Q_1, Q_2] \qquad (2.1)$$

where N_{interval} is the number of interval days between two peaks, A is the basin area in square miles, X_{min} is the minimum intermediate flow between two peaks, and Q_1 and Q_2 are two consecutive peak values. Same as extreme precipitation events, we set thresholds for the observed daily streamflow at each gauge after removing the reservoir effect to result in 1, 2, 3, 5 and 7 extreme discharge events per year on average.

3) ESTIMATION OF STORM RUNOFF RATIOS

The selections of POT extreme precipitation events and extreme discharge events were both based on daily observations. However, we needed the sub-daily data for the full hydrographs in order to estimate storm runoff ratios. Hence we calculated the runoff

ratios using hourly gridded precipitation from observations and hourly runoff from model simulations (readily available hourly discharge observations are too short for a period to be useful).

For each POT extreme precipitation event, we calculated its runoff ratio, which is equal to the storm total runoff volume divided by storm total precipitation. We took the beginning of a precipitation event as the first hour with precipitation greater than 0.3 mm (to avoid small spikes in COOP hourly data). The event ends once the hourly precipitation becomes lower than 0.3 mm or when the AR event ends, if it is one, and exceeds the first criterion. The summation of precipitation over the event hours is the storm total precipitation. For a runoff event, we determined the start and end of events in the hourly simulated streamflow as follows: 1) An event starts with the rise of the hydrograph; 2) If there is no following event, the end of the event is determined using the constant-*k* method of Blume et al. (2007), in which *k* is the recession coefficient, but no longer than three days after the peak hour; 3) If there is an immediate following event before the end of this event, the event ends with the start of the following one. The summation of streamflow over the event hours is the storm total runoff. The estimated runoff ratio of a storm event may be larger than 1.0 since we did not separate the baseflow from the streamflow time series, and the criteria we used to determine the start and end of individual events might not be sophisticated enough for time series at such a high (i.e. hourly) temporal resolution especially for small storm events. To take POT_{N3} events as an example, runoff ratios larger than 1 occurred less than 10% and 15 % of the time at downstream and upstream gauges, respectively. The ASM is defined as the

minimum value of the hourly surface-layer soil moisture within the 24 hours prior to the start of a precipitation event.

d. Examination of warming effect

The domain-average annual daily maximum temperature (T_{max}) and daily minimum temperature (T_{min}) increased by 0.8 ^{o}C and 1.3 ^{o}C respectively from 1950 to 2017, determined by linear regression. T_{min} generally increased over the entire domain in all seasons, with slight decreases during fall and winter months in the southwestern part, while T_{max} generally increased over the southern half of the domain and slightly decreased in the northern part especially during summer and fall months (see Figure A1).

To evaluate the effect of temperature change on ASM and floods since 1950, we constructed scenarios representing the temperature conditions in 1950 and 2017. Following Hamlet and Lettenmaier (2007), we detrended the temperature time series for each 1/32^{o} grid cell and each calendar month by removing the linear trends in the monthly average T_{max} and T_{min} over the period of WY 1950-2017 relative to the pivot year 1950 (denoted as "T1950") and 2017 (denoted as "T2017") based on the following equation.

$$T_{adj}[month][year] = T_{orig}[month][year] + Trend[month] \cdot (pivot\_year - year) \quad (2.2)$$

Under both scenarios, we assume that precipitation is not affected by temperature and stays unchanged. By comparing the ASM and flood response between the scenarios T1950 and T2017, we evaluate the impact of temperature increases without the influence of precipitation change over the past 68 years.

2.4 Results

a. Model evaluation

We first assessed the precipitation gridding method. We then evaluated the model performances for reservoir storage, streamflow, storm total runoff volume and soil moisture. We also compared the composite time series of hourly streamflow during extreme precipitation events. We summarize each of these evaluations below.

1) GRIDDED PRECIPITATION

We assessed the accuracy of the precipitation gridding method by systematically removing individual stations within the basin one at a time and evaluating the gridded product at the station grid in comparison with the removed station following Cao et al. (2018). The predicted precipitation at station locations showed reasonable matches with the available observations during the period WY 1950-2017, with the coefficient of determination (R^2) ranging from 0.83 to 1.0 and root-mean-square error (RMSE) ranging from 0.3 mm/day to 6.5 mm/day (mostly smaller than 3 mm/day (see Figure A2)).

2) STREAMFLOW AND STORM TOTAL RUNOFF VOLUME

We evaluated model performance for reservoir storage using R^2, RMSE and bias. For the hourly storage calibration, the R^2 values were 0.72 and 0.77, the RMSEs were 15.0 and 16.0 million m^3, and the biases were 11.3 and 11.8 million m^3 for Lake Mendocino and Lake Sonoma respectively (see Table 2.1 and Figure A3). In terms of elevation, the RMSEs were 2.2 m and 1.7 m, and the biases were 1.6 m and 1.2 m respectively. Model performance in the validation periods was similar to the calibration periods at Lake Mendocino but slightly degraded at Lake Sonoma.

Table 2.1 Streamflow [cubic meters per second (cms)] and reservoir storage calibration statistics. DA is drainage area and KGE is the Kling-Gupta efficiency. All statistics are for hourly data. The unit of RMSE and bias for reservoir storage is in million cubic meters and the unit for reservoir elevation is in meters.

USGS Gauge	Location	DA [mi^2]	Reservoir impact	Calibration periods (2005-2014)			Validation periods (1991-2000)		
				KGE	RMSE [cms]	Bias [cms]	KGE	RMSE [cms]	Bias [cms]
11461500	Calpella	92	N	0.67	7.5	3.7	0.60	8.2	4.5
11461000	Ukiah	100	N	0.69	8.2	2.7	0.72	8.0	2.7
11462500	Hopland	362	Y	0.84	19.9	7.6	0.86	20.3	8.2
11463000	Cloverdale	503	Y	0.90	22.0	8.6	0.91	23.0	9.3
11464000	Healdsburg	793	Y	0.93	31.3	11.9	0.92	32.5	12.5
11467000	Guerneville	1338	Y	0.87	48.2	19.3	0.90	49.2	19.6
Reservoir	Location	DA [mi^2]	Variable	Calibration periods (2005-2014)			Validation periods (1991-2000)		
				R^2	RMSE	Bias	R^2	RMSE	Bias
Coyote	Lake Mendocino	105	Storage	0.72	15	11.3	0.73	13	9.2
			Elevation	0.72	2.2	1.6	0.73	1.9	1.4
Warm Springs	Lake Sonoma	130	Storage	0.77	16	11.8	0.73	17	11.7
			Elevation	0.77	1.7	1.2	0.73	1.8	1.2

We used the Kling-Gupta efficiency (KGE) (Gupta et al., 2009), RMSE and bias to evaluate the goodness-of-fit between hourly streamflow observations and hourly simulations at the six selected stream gauges throughout the basin (see Figure 2.1a for gauge location). The simulations at gauges affected by reservoir regulations were from the model runs with the reservoir module activated. For the calibration period, the KGE ranged from 0.67 to 0.93; the RMSE ranged from 7.5 to 48.2 cms, and the bias ranged from 2.7 to 19.3 cms (see Table 2.1 and Figure A4). Model performance in the validation periods was similar to the calibration periods. The KGE was the highest at the downstream gauge 11464000.

In addition, we compared the observed (with reservoir effects removed) and simulated storm total runoff volumes calculated from the hourly data at six stream gauges for POT_{N3} (three events per year on average) extreme precipitation events during WY

1988-2017, which is the period when there were available observations for hourly streamflow (see Figure A5). The R^2 ranged from 0.68 to 0.95, with better match downstream.

3) SOIL MOISTURE

We evaluated the model performance of surface-layer soil moisture simulations at the HBG and ROD sites using R^2, RMSE and bias (see Figure 2.2). For hourly soil moisture, the R^2 values were 0.72 and 0.86, the RMSEs were 7.1% (volumetric water content) and 4.2%, and the biases were 5.2% and 3.2% for HBG and ROD respectively during late fall and winter months (October-March) from WY 2007 to WY 2017, the period when there were available observations.

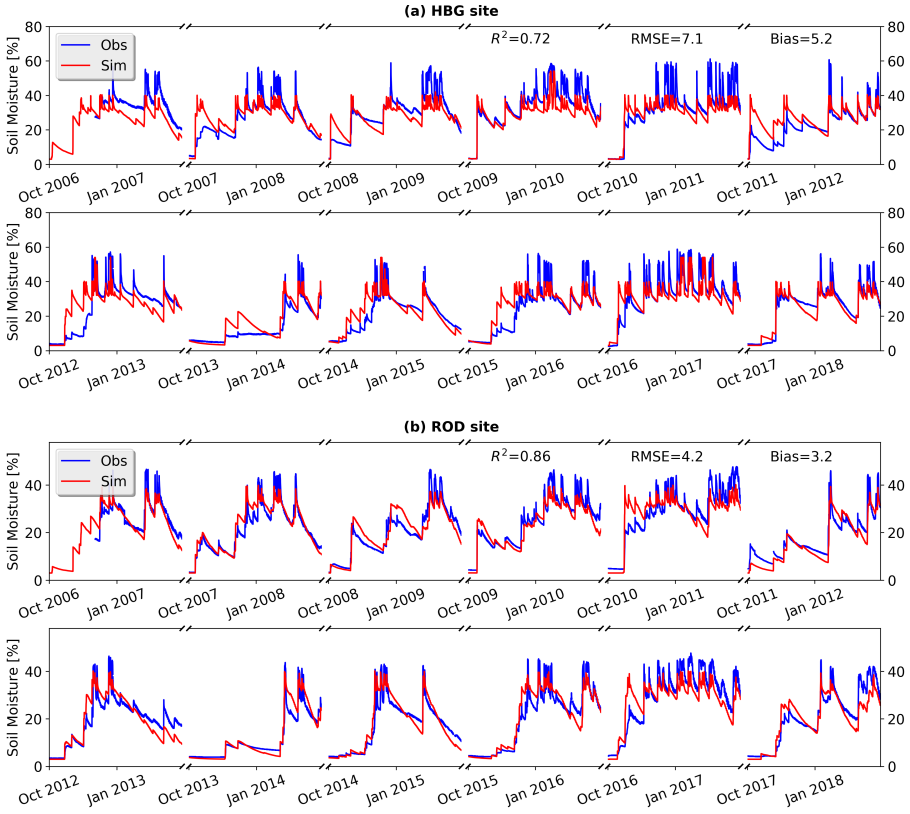

Figure 2.2 Comparison of simulated and observed surface-layer soil moisture at (a) HBG and (b) ROD sites during late fall and winter months (October–March) of WY 2007–17. The simulation is from the 150-m (model spatial resolution) pixel nearest to the observation site.

4) PEAK HOUR OF STREAMFLOW

Figure 2.3 shows the composite time series of POT_{N3} extreme precipitation events during WY 2007-2017. The simulated hourly streamflow showed reasonable matches with the observed hourly streamflow at downstream gauges. However, its performance was not as good at upstream gauges, which have relatively small drainage areas and where the simulated streamflow arguably is more sensitive to the quality of the precipitation forcings than at downstream gauges. Furthermore, there are very few long-term precipitation gauges in the headwater areas. The peak hour of observed streamflow gradually increased from 14 hours at the upstream-most gauge (11461500) to 40 hours at the downstream-most gauge (11467000). The simulated streamflow showed the best match for peak hour at gauge (11463000) in the middle of the basin.

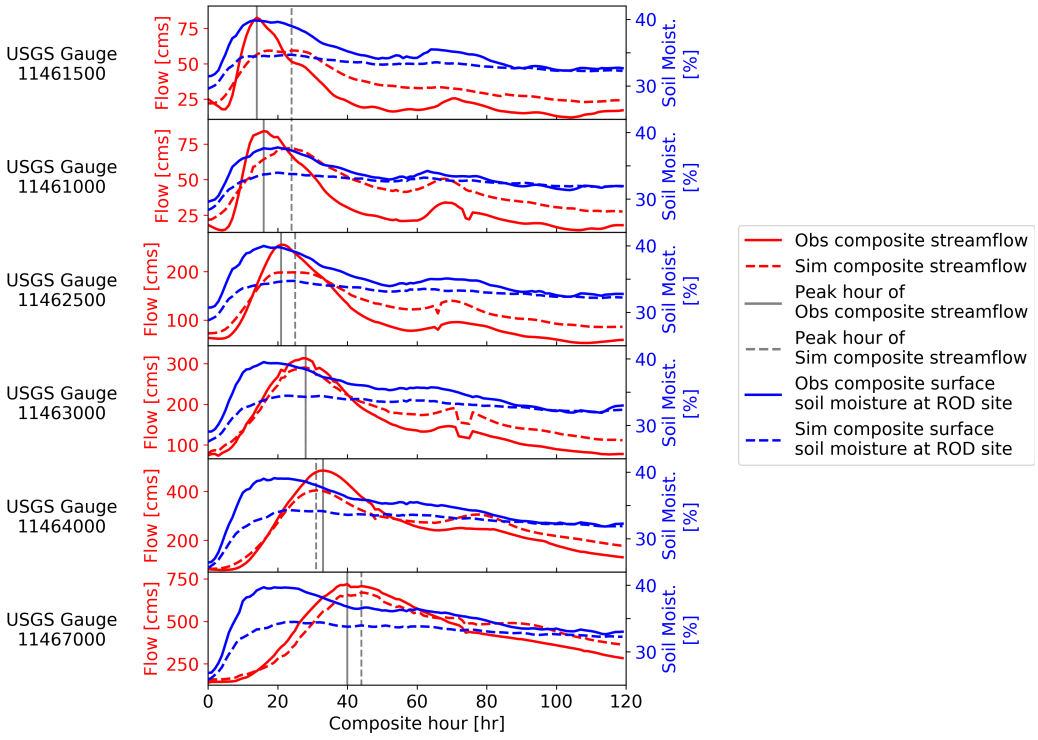

Figure 2.3 Composite time series of observed and simulated hourly streamflow and surface-layer soil moisture at ROD site of POT_{N3} extreme precipitation events during WY 2007–17 at six USGS stream gauges ranked by drainage area. Peak hours of the composite observed and simulated streamflow are shown in gray solid line and gray dashed line, respectively.

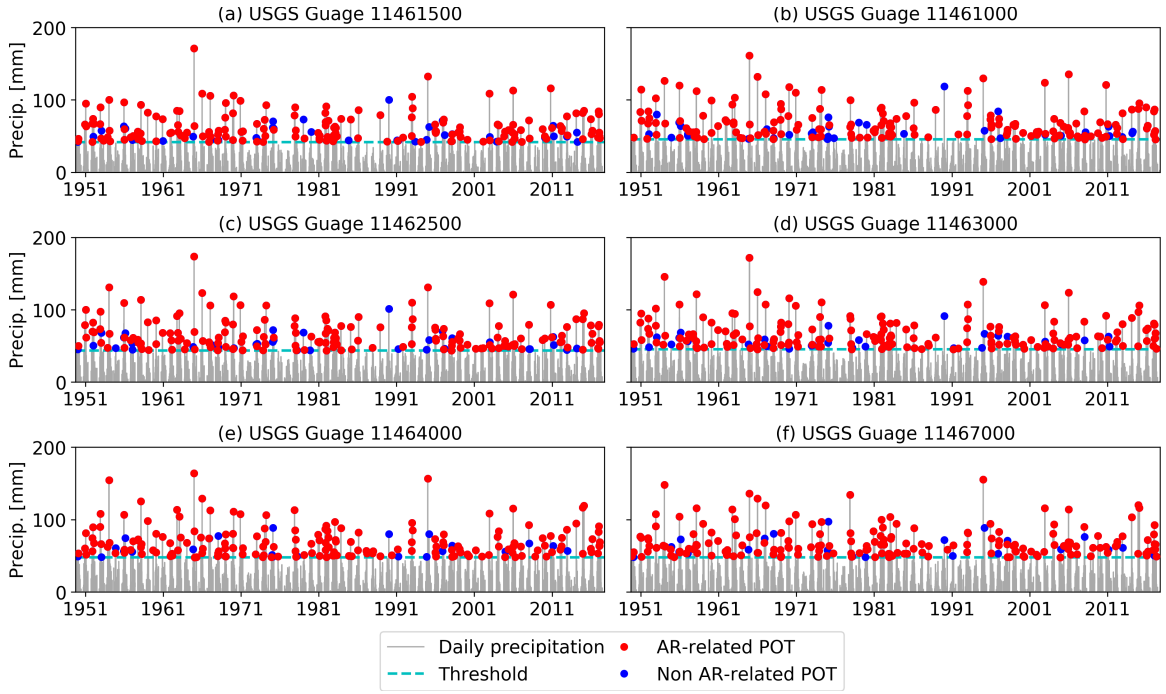

Figure 2.4 POT_{N3} extreme precipitation events (with threshold set to three events per year on average) based on gridded daily precipitation averaged over the upstream drainage area at six USGS stream gauges during WY 1950–2017.

b. The role of AR in extreme events

We set thresholds of POT extreme precipitation events to make sure there were 1, 2, 3, 5 and 7 events per year on average. Figure 2.4 shows the POT extreme precipitation events when the threshold was set to three events per year at six stream gauges from upstream most to downstream most throughout the basin during WY1950-2017, in which most of the events were AR-related. Table 2.2 summarizes the percentage of events that are related to ARs with different threshold values. We can see that the percentage of AR-related precipitation events increased as the threshold increased. The AR contribution increased (from 81.9%-83.0% with a mean of 82.5% to 95.6%-100.0% with a mean of 98.8%) across the gauges when the threshold increased from POT_{N7} to POT_{N1}. Also, no more than one non AR-related POT event across six gauges was in the upper 25th

percentile of all extreme precipitation events when the threshold was higher than three events per year. Even when the threshold was set to five or seven events per year, no more than four of them fell into the upper 25th percentile (see Table 2.2). The median (peak) precipitation of AR-related POT extreme precipitation events was greater than non AR-related events by 7.8%-23.2% across six gauges with a threshold of POT_{N3} and by 23.1%-34.9% with a threshold of POT_{N7} (see Table A1).

Table 2.2 Statistics of POT extreme precipitation and discharge events at six USGS stream gauges. Variable *N* is the number of events and $N_{upper25th}$ is the number of events in the upper 25th percentile.

POT extreme precipitation events	Threshold [N/yr.]	USGS stream gauge					
		11461500	11461000	11462500	11463000	11464000	11467000
AR-related percentage [%]	1	100.0	100.0	100.0	100.0	97.1	95.6
	2	96.3	95.6	97.1	96.3	96.3	94.9
	3	91.7	93.1	92.6	93.6	94.6	95.6
	5	86.5	88.5	87.1	87.6	86.8	87.4
	7	81.9	83.0	82.6	83.0	82.1	82.6
$N_{upper\ 25th}$ (AR/Non AR)	1	17/0	17/0	17/0	17/0	17/0	17/0
	2	34/0	34/0	34/0	34/0	34/0	34/0
	3	51/0	51/0	51/0	51/0	50/1	50/1
	5	84/1	85/0	85/0	84/1	83/2	82/3
	7	115/4	116/3	116/3	116/3	115/4	115/4
POT extreme discharge events	Threshold [N/yr.]	USGS stream gauge					
		11461500	11461000	11462500	11463000	11464000	11467000
AR-related percentage [%]	1	95.6	92.6	95.6	97.1	97.1	91.2
	2	90.4	94.9	92.6	93.4	93.4	85.3
	3	83.9	88.8	88.2	88.7	88.7	80.4
	5	75.3	78.8	73.2	74.4	73.5	68.5
	7	69.0	69.7	65.5	65.1	64.3	59.9
$N_{upper\ 25th}$ (AR/Non AR)	1	17/0	17/0	17/0	17/0	17/0	17/0
	2	33/1	33/1	34/0	34/0	34/0	33/1
	3	48/3	49/3	51/0	51/0	51/0	45/6
	5	79/6	80/5	78/7	81/4	81/4	75/10
	7	110/9	113/6	110/9	112/7	112/7	104/15

We set the same thresholds for POT extreme discharge events. The AR contribution increased from 59.9%-69.7% with a mean of 65.6% to 91.2%-97.1% with a mean of 94.9% across the gauges when the threshold increased from POT_{N7} to POT_{N1}. No more than one non AR-related POT event across six gauges was in the upper 25th percentile of all extreme discharge events when the threshold was higher than two events per year (see Table 2.2). The median (peak) discharge of AR-related POT extreme discharge events was greater than non AR-related events by 34.6%-44.8% across six gauges with a threshold of POT_{N3} and by 80.0%-213.2% with a threshold of POT_{N7} (see Table A1).

c. The role of ASM in historical AR flooding

1) EXTREME PRECIPITATION EVENTS

(i) Observation-based examination

We examined the relationship between the ASM and runoff ratio in the POT_{N3} extreme precipitation events based on observations at six stream gauges throughout the basin during WY 2007-2017 (the period when hourly soil moisture observations were available at two of the HMT sites and hourly streamflow observations were available at USGS gauges). Following Crow et al. (2017), we used the Spearman rank correlation coefficient R_s to evaluate the strength of the potentially nonlinear relationship between ASM and the runoff ratio. Figure 2.5a shows the observed runoff ratio versus the observed ASM at the ROD site. The R_s values generally were higher at downstream gauges than upstream gauges, ranging from 0.71 to 0.82, with the highest value at the downstream-most gauge 11467000. A similar pattern was found for the relationships between the simulated runoff ratio and simulated ASM at the ROD site, with R_s ranging

from 0.76 to 0.88 (see Figure 2.5b). Similarly, the ranges of R_s for the observation versus observation and model versus model comparisons at the HBG site are 0.63-0.75 and 0.55-0.73, respectively (see Figure 2.5c-d).

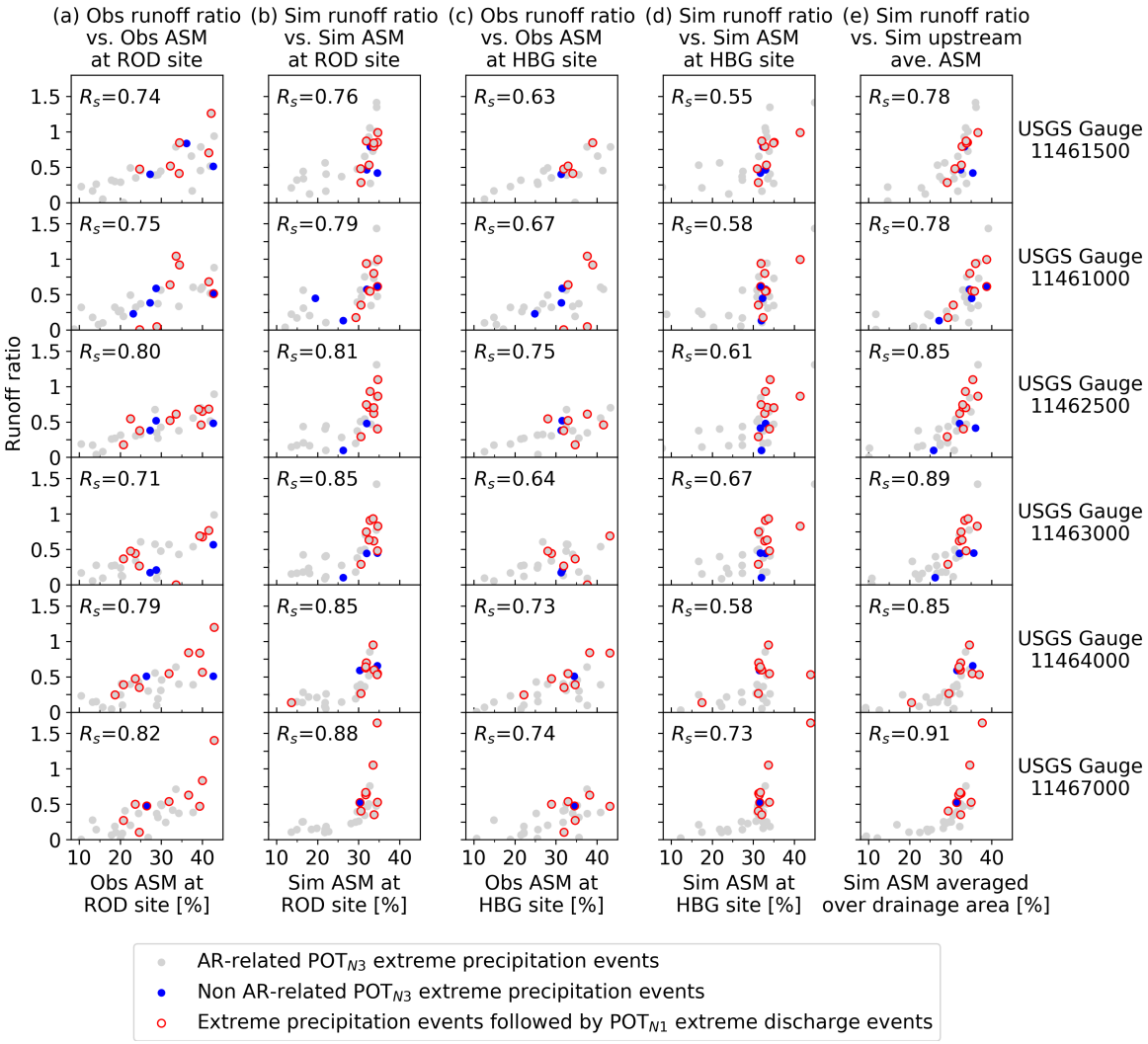

Figure 2.5 Comparison of the observation-based and simulation-based relationships between runoff ratio and ASM of POT_{N3} extreme precipitation events (with threshold set to three events per year on average) during WY 2007–17 (period of available soil moisture observations) at six USGS stream gauges ranked by drainage area, including (a) observed runoff ratio vs. observed ASM at ROD site; (b) simulated runoff ratio vs. simulated ASM at ROD site; (c), (d) as in (a), (b), but at HBG site; and (e) simulated runoff ratio vs. simulated ASM averaged over upstream drainage area of the stream gauge. The observed runoff ratio is calculated based on gridded hourly precipitation and observed hourly streamflow with reservoir effects removed. The simulated runoff ratio is calculated based on gridded hourly precipitation and simulated hourly streamflow.

Relatively long soil moisture observation records were available only at the HBG and ROD sites; however, we simulated soil moisture for the entire basin. For each stream gauge, we calculated the average ASM over its upstream drainage area. The R_s between simulated runoff ratios and simulated upstream average ASM ranged from 0.78 to 0.91, higher than the above R_s values calculated based on ASM at one single site (see Figure 2.5e). The R_s across gauges generally increased as drainage area increased. Besides, the POT_{N3} extreme precipitation events followed by POT_{N1} extreme discharge events generally had a wet prestorm condition (see Figure 2.5e). Also, the top POT_{N1} extreme discharge events were all related to ARs during WY 2007-2017.

(ii) Simulation-based examination

Given that the model produced plausible reproductions of observed streamflow and soil moisture, we examined the relationship between the ASM and runoff ratio in the POT_{N3} extreme precipitation events related to ARs based on simulated hourly streamflow and simulated hourly soil moisture at six stream gauges throughout the basin during WY 1950-2017. Figure 2.6 shows boxplots of the simulation-based runoff ratios versus possible influencing factors including maximum precipitation intensity, average precipitation intensity, storm total precipitation and ASM. The plots show that runoff ratio was barely affected by either the maximum or average precipitation intensity (see Figure 2.6a-b). It generally increased, in terms of the median value, as storm total precipitation increased, but with a wide range regardless of the total precipitation (Figure 2.6c).

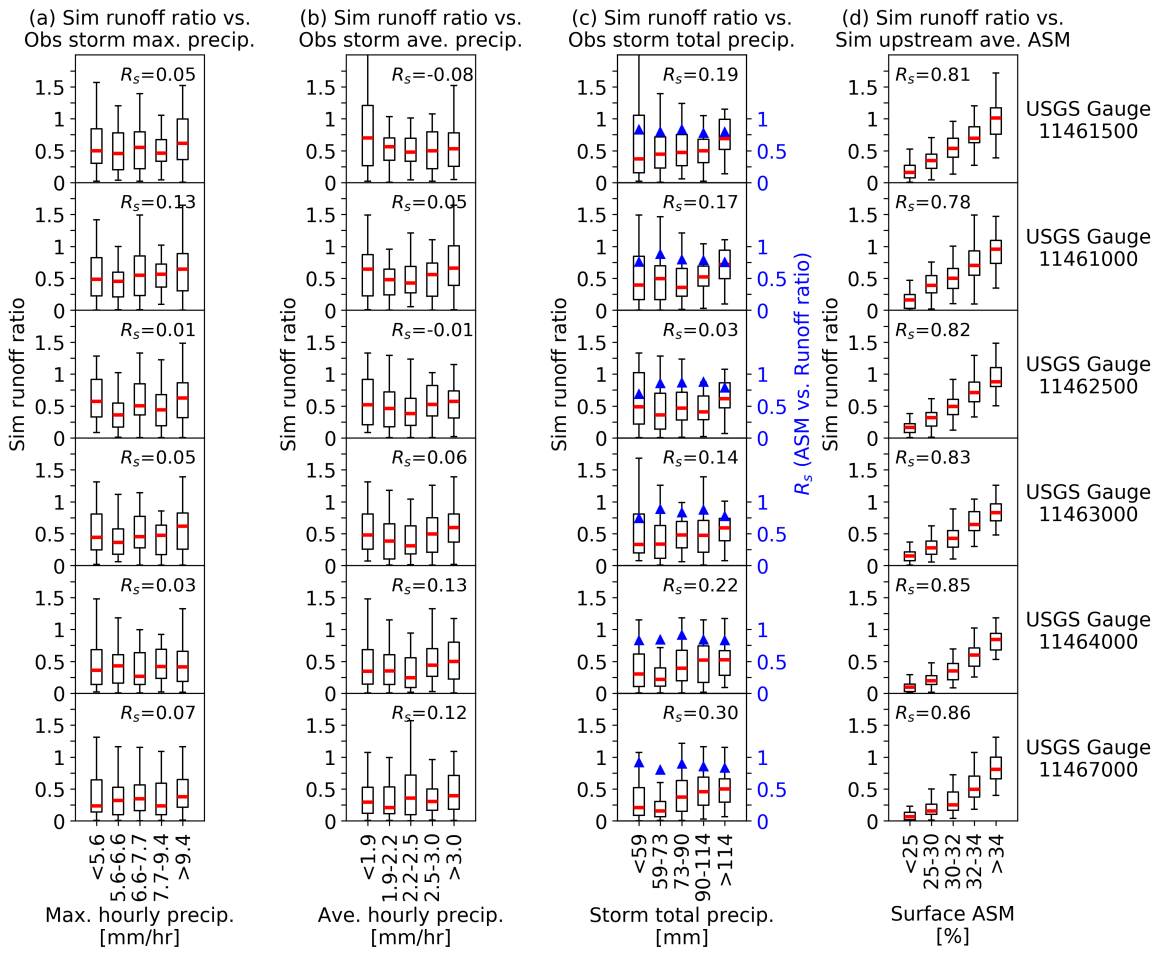

Figure 2.6 Boxplots with an interval of 20th percentile of the simulation-based runoff ratio vs. the (a) maximum precipitation intensity, (b) average precipitation intensity, (c) storm total precipitation, and (d) ASM of POT_{N3} extreme precipitation events (with threshold set to three events per year on average) during WY 1950–2017 at six USGS stream gauges ranked by drainage area. The Spearman's rank coefficient (R_s) between the runoff ratio and ASM vs. the storm total precipitation is shown as blue triangles in (c).

In contrast, the runoff ratio is much more strongly related to ASM, with R_s increasing slightly from 0.78 at upstream gauge 11461000 to 0.86 at downstream gauge 11467000 for all POT_{N3} extreme precipitation events (Figure 2.6d). When ASM was in the lower 20th percentile, the runoff ratios were low regardless of precipitation intensity. ASM, however, had a smaller effect on the runoff ratios as storm total precipitation increased, especially at downstream gauges (Figure 2.6c).

In order to compare the relationship between the ASM and runoff ratio for AR- and non-AR events, we set the threshold lower to include more events, so that we have enough non-AR events to get reasonable statistics. We set the threshold of POT extreme precipitation events to make sure there were 10 events per year on average. For 680 POT extreme precipitation events during the period WY 1950-2017, there were 506-515 AR events and 165-174 non-AR ones across six gauges. The runoff ratio was generally higher for AR events than non-AR events given the same threshold of ASM since storm precipitation was generally larger for AR events (see Figure 2.7).

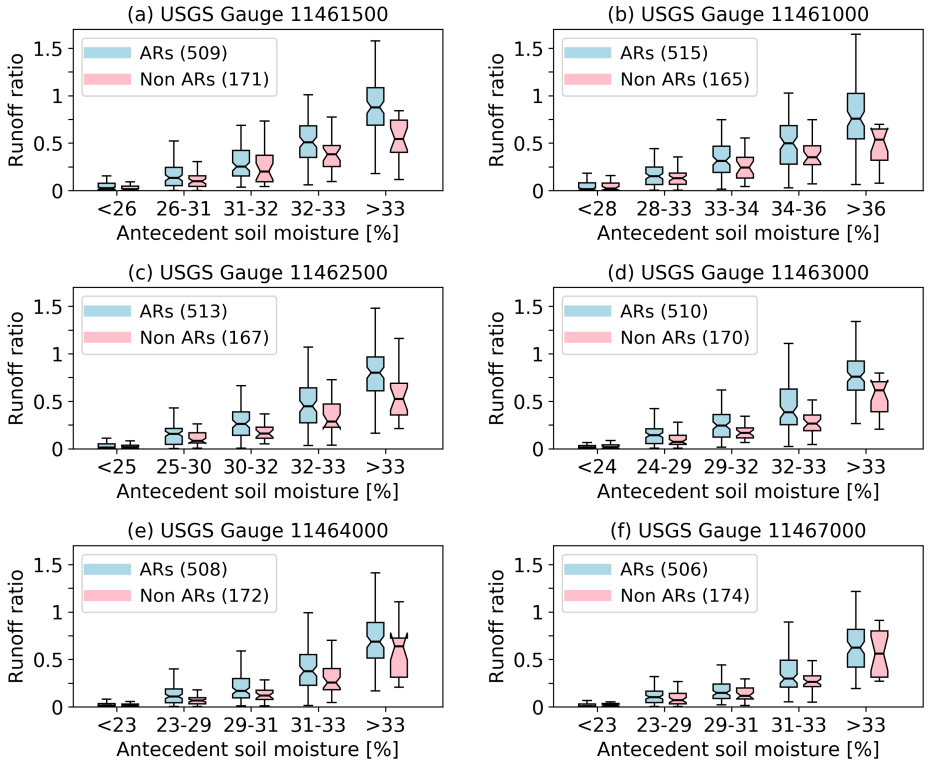

Figure 2.7 Boxplots with an interval of 20th percentile of the simulation-based runoff ratio vs. ASM of POT_{N10} extreme precipitation events (with threshold set to 10 events per year on average) during WY 1950–2017 at six USGS stream gauges. The POT events are categorized by ARs and non ARs, with number of events shown in legends.

2) EXTREME DISCHARGE EVENTS

Our analysis above shows that extreme precipitation events lead to extreme discharge events when ASM is wet or when storm total precipitation is large enough (see

section 5 for more discussion). In this section, we examine the relationship between the ASM, storm precipitation and peak daily flow in the POT_{N1}, POT_{N2}, and POT_{N3} extreme discharge events based on observed daily streamflow and simulated hourly soil moisture at six stream gauges throughout the basin during WY 1950-2017.

First, we examined the relationship between accumulated storm precipitation and observed peak daily flow given antecedent soil moisture. We used durations ranging from 6 hours to 72 hours since the start of storm precipitation during each extreme discharge event so that we could determine which duration of precipitation most strongly affects the peak flow (see Figure 2.8 for results). The correlation generally increased as duration increased and peaked around 48 hours for most gauges in POT_{N2} and POT_{N3} events. The correlation peaked at a longer duration for some gauges in POT_{N1} events. For consistency, we used the first 2-day (48-hour) accumulated storm precipitation (denoted as SP_{2d}) in the analysis described below.

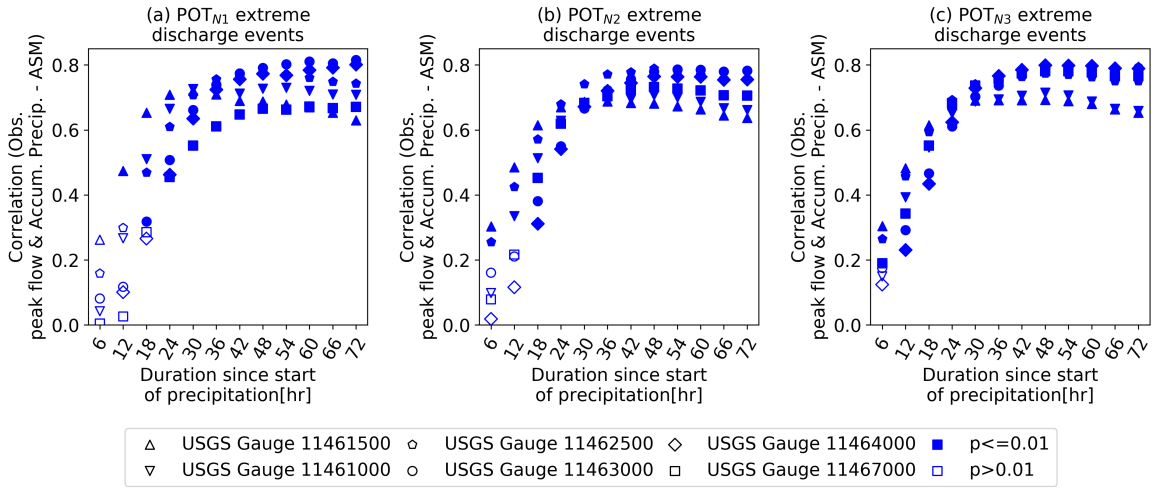

Figure 2.8 Relationship between accumulated precipitation and observed peak daily flow given ASM in (a) POT_{N1} extreme discharge events (with threshold set to one event per year on average), (b) POT_{N2} extreme discharge events (two events per year on average), and (c) POT_{N3} extreme discharge events (three events per year on average) during WY 1950–2017 at six USGS stream gauges.

We then examined the partial correlation between the SP_{2d} and observed peak daily flow (denoted as PF) given ASM (denoted as $r_{SP2d\&PF\text{-}ASM}$), and the partial correlation between ASM and observed PF given SP_{2d} (denoted as $r_{ASM\&PF\text{-}SP2d}$) (see Figure 2.9a). SP_{2d} had higher correlations with observed PF, ranging from 0.69 to 0.80 across six gauges in POT_{N3} events, than ASM with observed PF, ranging from 0.40 to 0.67 across six gauges in POT_{N3} events, but all correlations are statistically significant with p<=0.01. $r_{ASM\&PF\text{-}SP2d}$ generally increased with drainage areas. Both correlations were higher when using the simulated PF since the ASM was from the model (see Figure 2.9b). The effects of POT threshold on $r_{SP2d\&PF\text{-}ASM}$ and $r_{ASM\&PF\text{-}SP2d}$ varied by gauge location.

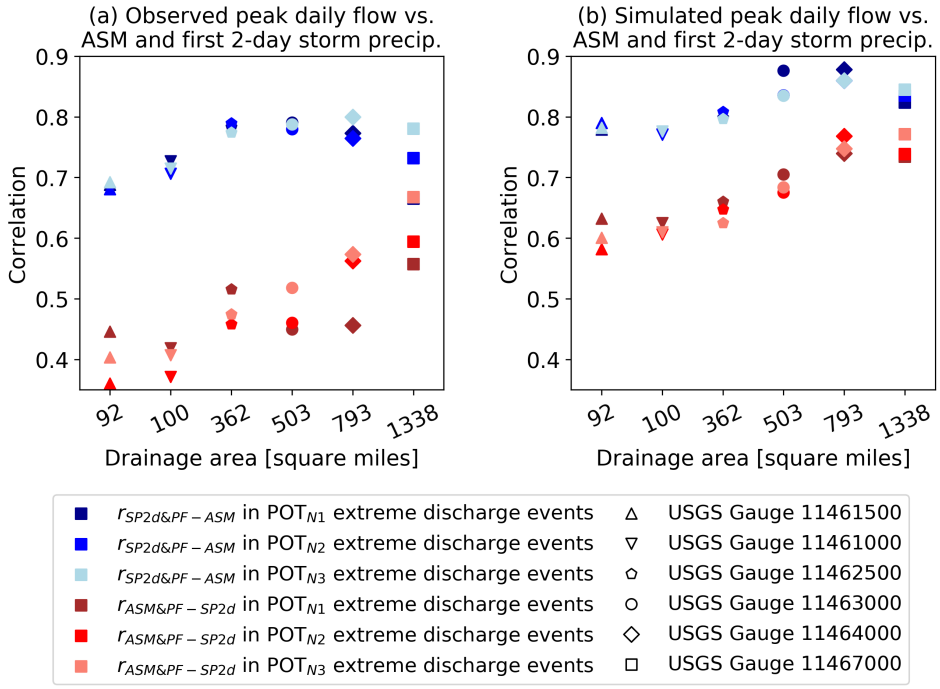

Figure 2.9 (a) Correlation between first 2-day accumulated storm precipitation and observed peak daily flow given ASM (denoted as $r_{SP2d\&PF\text{-}ASM}$), and correlation between ASM and observed peak daily flow given first 2-day accumulated storm precipitation (denoted as $r_{ASM\&PF\text{-}SP2d}$) in POT_{N1}, POT_{N2}, and POT_{N3} extreme discharge events (i.e., one, two, and three events per year on average) during WY 1950–2017 at six USGS stream gauges. (b) As in (a), but with simulated peak daily flow. All correlations are statistically significant with $p \leq 0.01$.

d. Warming effect

1) TRENDS IN HISTORICAL POT EXTREME EVENTS

We examined the trend in POT extreme precipitation events (based on gridded daily precipitation averaged over drainage area) and POT extreme discharge events (based on observed daily streamflow with reservoir effects removed) at six stream gauges throughout the basin with thresholds POT_{N1}, POT_{N2} and POT_{N3} during WY 1950-2017 (see Table A2). Trends in POT extreme discharge events are not necessarily in agreement with trends in POT extreme precipitation events since whether the extreme precipitation events lead to extreme discharge events depends strongly on ASM (and more weakly on storm total precipitation).

No significant trends were found in extreme precipitation. However, weak (but statistically significant) downward trends were found at some gauges for extreme discharge, particularly in POT_{N2} extreme discharge events (see Figure 2.10a), which are consistent with weak (but statistically significant) downward trends in the SP_{2d} of extreme discharge events at most gauges (see Figure 2.10c). However, statistically significant downward trends in the SP_{2d} do not necessarily lead to significant decreases in PF despite their high correlation due to the influence of ASM, which is most obvious at downstream gauges. One hypothesis for the stronger trends in SP_{2d} as contrasted with PF shown in Figure 2.10 is that the absence of trends in ASM (see Figure 2.10b) is (somewhat) modulating the trends in precipitation. The absence of trends in ASM could be related to a) cancellation of increased evaporative demand due to warmer temperature by slight increases in small precipitation amounts, which would need additional analysis,

b) timing of the events, which is pretty critical for fall vs. winter. This is a topic that could be investigated in detail in future work.

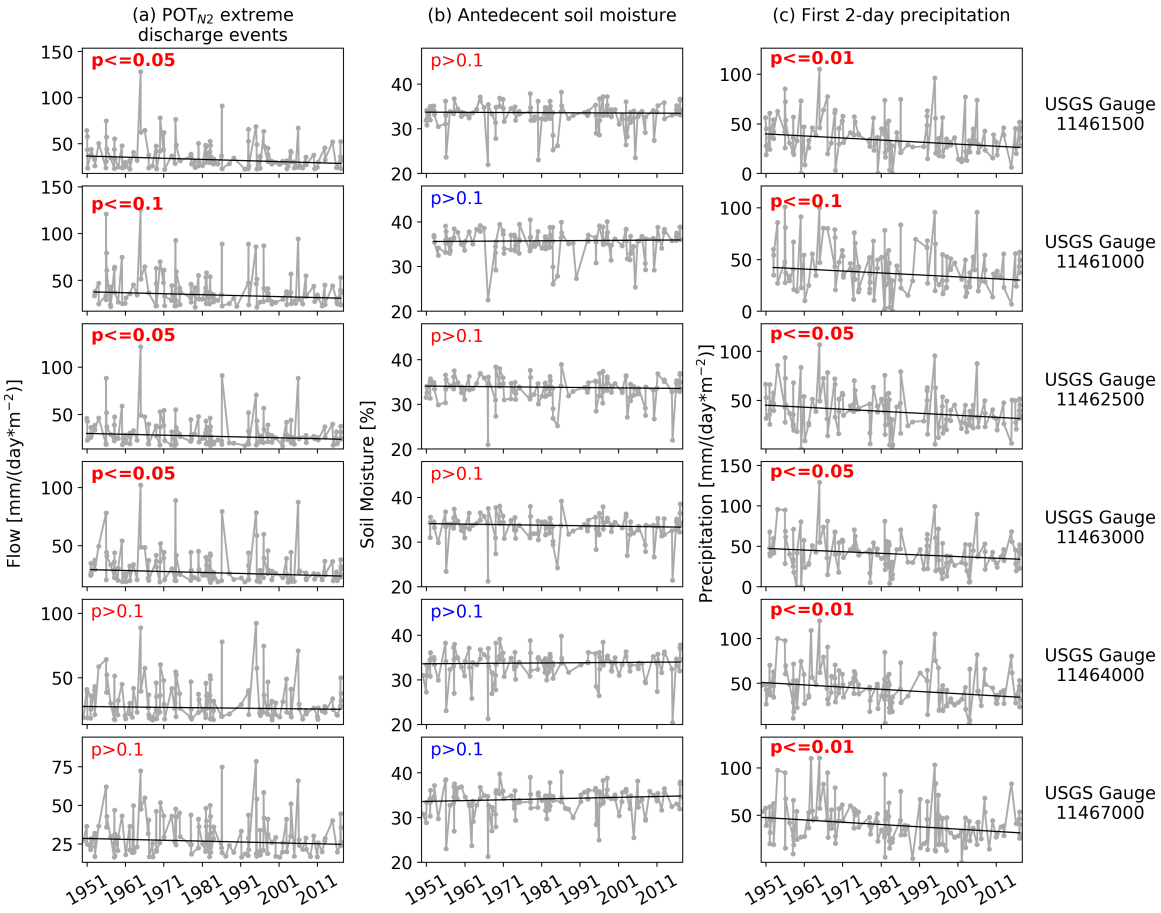

Figure 2.10 Trends in (a) observed peak daily streamflow and (b) ASM from model hourly simulation, and (c) first 2-day accumulated storm precipitation of POT_{N2} extreme discharge events (threshold set to two events per year on average). The p values are shown in plots, categorized by $p \leq 0.01$, $0.01 < p \leq 0.05$, $0.05 < p \leq 0.1$, $p \leq 0.1$, and $p > 0.1$, with positive trends are marked in blue and negative in red. The p values not greater than 0.1 are marked in bold font.

Both extreme discharge and its corresponding SP_{2d} generally decreased at all gauges in POT_{N1}, POT_{N2} and POT_{N3}. In contrast, ASM only decreased at 2, 3 and 4 gauges in POT_{N1}, POT_{N2} and POT_{N3}, but none of the trends in ASM were significant except the one ($p<0.1$) at Gauge 11462500 in POT_{N3} events (see Table A2).

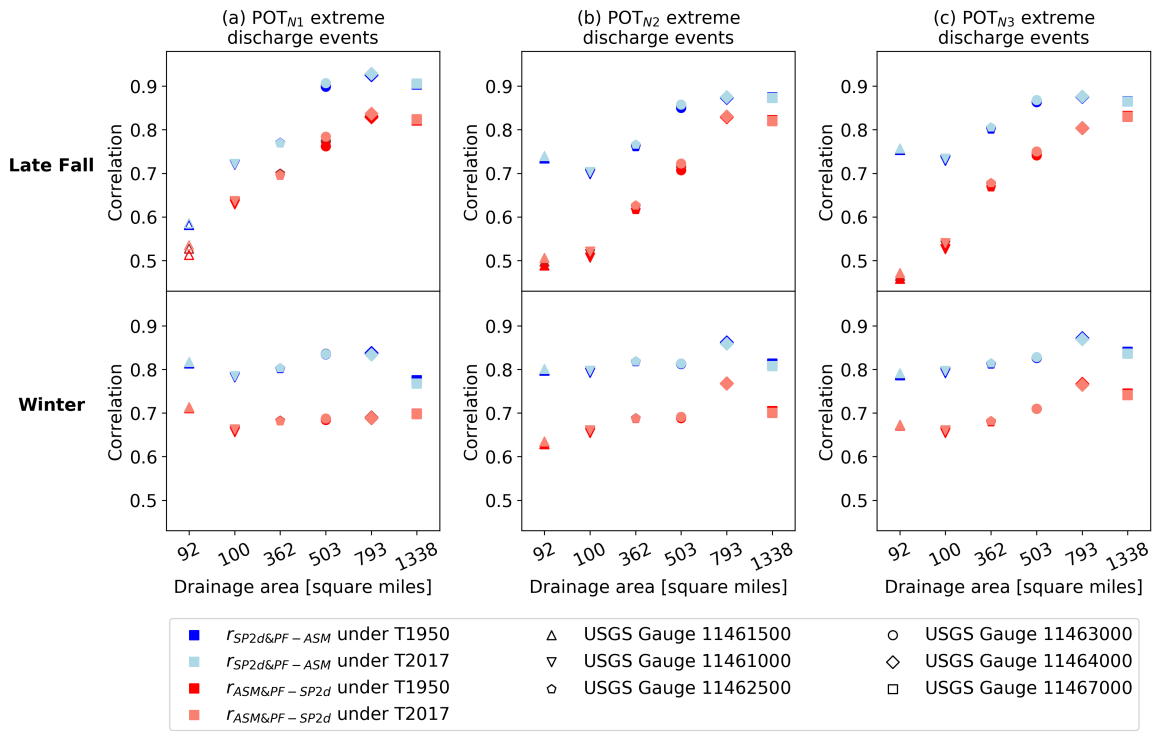

Figure 2.11 Correlation between first 2-day accumulated storm precipitation and observed peak daily flow given ASM ($r_{SP2d\&PF-ASM}$), and correlation between ASM and observed peak daily flow given first 2-day accumulated storm precipitation ($r_{ASM\&PF-SP2d}$) in (a) POT_{N1}, (b) POT_{N2}, and (c) POT_{N3} extreme discharge events (i.e., one, two, and three events per year on average) during WY 1950–2017 at six USGS stream gauges. (top panel) Events in late fall (October–December) and (bottom panel) events in winter (January–March). All correlations are statistically significant with $p \leq 0.05$. The ones with $p \leq 0.01$ are shown by solid symbols.

2) TEMPERATURE EFFECT

In order to assess the effect of increased temperature alone, we compared the trends in extreme discharge under temperature scenarios T1950 and T2017 (see Table A2). The relative change in trends of extreme discharge was very small: from -2.4%~5.7% across six gauges in POT_{N2} extreme discharge events (in all cases, extreme discharge was on average smaller for T2017 than T1950 especially at downstream gauges, presumably because of increased evapotranspiration and hence decreased ASM). We also examined the correlation between ASM and observed PF in extreme discharge events given SP_{2d} under T1950 and T2017 in late fall (Oct-Dec) and winter (Jan-Mar) months (see Figure

2.11). ASM had a larger impact on extreme discharge in winter than in late fall months at the upstream gauges, and the reverse was true at downstream gauges. Warming slightly increased the effect of ASM on extreme discharge in the late fall months but had little effect in the winter months, possibly due to that there usually are enough minor storms in winter and ASM is fairly wet.

2.5 Discussion

We note that there are some limitations of our study based on the available data. One is the low spatial resolution of the NCEP-NCAR reanalysis dataset that the AR date catalog is based on. Other limitations are the representaton of reservoir impacts on streamflow, and systematic modeling error (especially at upstream gauges). We used the AR catalog based on the NCEP-NCAR dataset because of its long record. Despite of its coarse spatial resolution, others have found that AR detection algorithms based on the NCEP-NCAR reanalysis identify (mostly) the same storms as other reanalysis datasets with finer spatial resolution (but shorter records) -- e.g. ERA-Interim and MERRA2 -- particularly the stronger and longer ones (Ralph et al., 2018a). We used a reservoir module, DHSVM-res, to remove the reservoir impacts on streamflow at downstream gauges to estimate naturalized flows (see Table 2.1 and Figures A3-A5 for model performance). Finally, to mitigate the potential impacts of the systematic modeling error, we used observations in the analysis where they were available as much as possible such as the long-term trend analysis of the peak flow in section 4.4.

As mentioned above, we selected both POT extreme precipitation events and extreme discharge events based on daily observations, as in many previous studies (e.g.

Lang et al., 1999; Begueria et al., 2011; Mallakpour and Villarini, 2017). However, runoff ratios as well as peak daily discharges are more strongly related to storm total precipitation than peak daily precipitation. In other words, POT extreme precipitation events may not necessarily lead to POT extreme discharge events of the same category. Figure 2.12 shows corresponding categories of peak observed daily flow led by POT_{N1}, POT_{N2}, and POT_{N3} extreme precipitation events during WY 1950-2017 at downstream-most USGS Gauge 11467000 (Russian River near Guerneville). In general, almost all extreme precipitation events led to extreme discharge events of the same category when upstream average ASM was wetter than 32%, corresponding to the 60th percentile of ASM conditions in all POT_{N3} extreme precipitation events during WY1950-2017. When ASM was very dry (lower than 25%, corresponding to the 20th percentile), POT_{N1}, POT_{N2}, and POT_{N3} extreme precipitation events with storm total precipitation exceeding 209 mm, 204 mm and 70 mm, respectively, might still be able to lead to extreme discharge events of the same category.

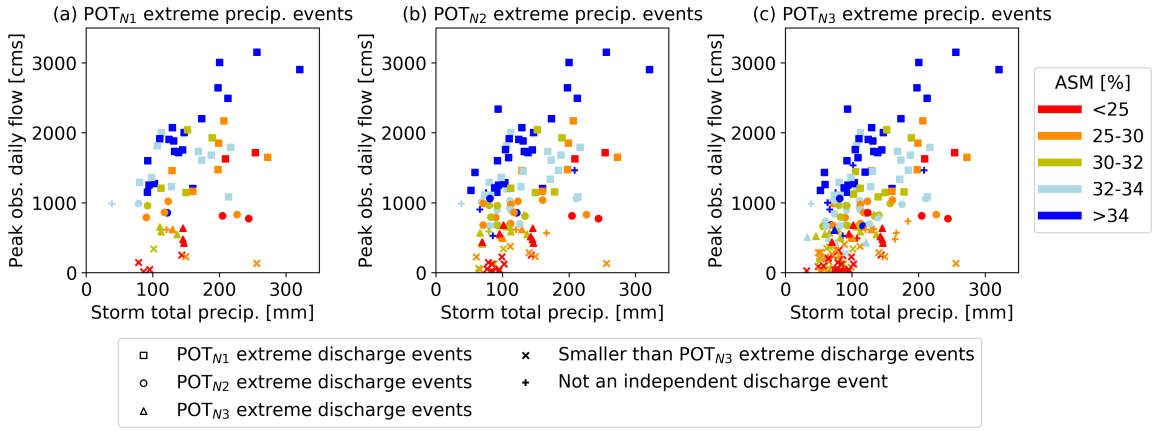

Figure 2.12 Corresponding categories of peak observed daily flow led by (a) POT_{N1}, (b) POT_{N2}, and (c) POT_{N3} extreme precipitation events (i.e., one, two, and three events per year on average) during WY 1950–2017 at downstream-most USGS gauge 11467000 (Russian River near Guerneville).

Table 2.3 Simulation-based runoff ratio and ASM of discharge events with crest exceeding major flood stage (40 ft) at the downstream-most USGS gauge 11467000 (Russian River near Guerneville) during WY 1950–2017. The NWS flood stage observations are measured at the USGS gauge 11467002 (Russian River at Guerneville), a few kilometers away from the USGS gauge 11467000.

NWS event rank	Date	Observed crest of flood stage [ft]	Simulated hourly peak [cms]	Simulation-based runoff ratio	Simulated ASM [volumetric water content expressed as %]	Simulated ASM percentile in historical POT_{N3} precipitation events [%]
1	2/18/1986	49.5	3545	1.15	40.2	100.0
2	1/10/1995	48.0	4046	0.88	36.0	94.8
3	12/23/1955	47.6	4004	1.02	38.4	98.4
4	12/23/1964	47.4	4491	0.76	34.3	85.3
5	1/1/1997	45.0	2326	1.16	38.3	97.9
6	1/5/1966	42.5	3134	0.88	35.0	90.6
7	1/1/2006	41.8	3169	0.67	34.0	80.6
8	3/10/1995	41.5	1780	0.25	28.5	31.9
9	1/24/1970	41.3	3067	1.09	39.7	99.5
10	2/1/1963	41.1	1963	0.89	33.6	77.0
11	1/17/1974	40.7	2555	0.83	36.0	95.3
12	1/27/1983	40.4	2699	0.76	34.6	86.4
13	2/25/1958	40.2	2201	0.86	34.6	86.9
	Median		3067	0.88	35.0	90.6
	Max		4491	1.16	40.2	100.0
	Min		1780	0.25	28.5	31.9

For the same gauge, we further examined its runoff ratios and ASM for the largest historical extreme discharge events with crests exceeding flood stage (see Table 2.3). For these flood events, the runoff ratio ranged from 0.25 to 1.16 with a median value of 0.88. The lowest value of 0.25 for the event on 3/10/1995 was due to a consecutive dependent small peak in the hourly streamflow data, leading to an early end of the streamflow event according to the predefined event separation criteria. Aside from this event (classification of which arguably is an artifact of the storm precipitation identification procedure), all of the historical major flood events had simulation-based runoff ratios greater than 0.67 and ASM greater than 33.6%, the latter corresponding to the 77^{th} percentile of ASM

conditions in all POT_{N3} extreme precipitation events during WY1950-2017. Ralph et al. (2019) examined the relationship between the AR scale and major flooding at this gauge (see Table 5 in their paper). They found that 6 out of 10 AR events in the highest category (in terms of IVT intensity and AR duration) were associated with major flooding during 1980-2017. They hypothesized that the three that did not lead to major flooding (one occurred when streamflow data were not available) were related to dry soils. We showed here (see Table A3) that the ASM values were less than 20% (volumetric water content expressed as %) for these three events, which corroborates their hypothesis. Furthermore, 5 out of the 6 AR events that led to flooding had ASM values greater than 32%. The remaining one was the smallest flood event among the six, with an ASM of 22%.

Long-term trends in temperature alone had little effect on the ASM in POT extreme discharge events, indicating that the changes in ASM were largely caused by antecedent precipitation. We examined the relative effects of temperature and antecedent precipitation on ΔASM by calculating the partial correlation between ΔASM and accumulated precipitation given accumulated evapotranspiration (ET) (denoted as $r_{\Delta ASM\ \&P\text{-}ET}$), and partial correlation between ΔASM and ET given P (denoted as $r_{\Delta ASM\ \&ET\text{-}P}$) under different pre-event durations (from 2 days to 12 weeks) in late fall (Oct-Dec) and winter (Jan-Mar) (see Figure A6a-b). The accumulated P had a larger effect on ΔASM than ET, and its relative effect decreased as pre-event duration lengthened. The influence of ET on ΔASM was larger in late fall than winter. We further examined the impact of warming on ΔASM. Warming slightly increased the effect of ET on ΔASM in late fall

especially when the pre-event duration exceeded one month. Warming barely affected the influence of ET on ΔASM in winter.

2.6 Conclusions

We first assessed the contribution of ARs to the extreme precipitation and extreme discharge events based on daily observations in California's Russian River Basin using the POT method with varying thresholds (hence, number of events) during WY 1950-2017. We then examined the role of ASM on historical AR flooding to answer two questions: 1) Under what ASM conditions would extreme precipitation events lead to extreme discharge events of the same POT category? 2) For extreme discharge events, what is the role of ASM on peak flow? Recent studies suggest that observed changes in extreme precipitation associated with a general warming have not necessarily lead to corresponding changes in floods, and changes in ASM could be a primary missing link. We finally examined how the observed warming during the past 68 years had affected ASM and thus extreme discharge events. Based on our analysis, we find that:

1) Most extreme precipitation and flood events in the Russian River Basin are associated with ARs – up to 99% of extreme precipitation events and up to 95% of flood events, depending on the threshold used. The non-AR events were mostly of minor severity, with very few falling in the upper 25^{th} percentile of all extreme precipitation or discharge events.

2) The runoff ratio during extreme precipitation events is much more strongly related to ASM than to storm total precipitation. When ASM is wet (e.g. upstream average ASM greater than 32% at the downstream-most stream gauge 11467000, Russian

River near Guerneville) or storm precipitation is sufficiently large, almost all extreme precipitation events lead to extreme discharge events of the same POT category. Among the extreme discharge events, however, the first 2-day storm precipitation had a greater effect on the peak flow than did ASM, but the effects of ASM on peak flow increases as drainage area increases.

3) We found no trends in the magnitude of extreme precipitation but weak downward trends in the magnitude of extreme discharge at some stream gauges in the Russian River Basin during the period WY 1950-2017. We found no trends in ASM for extreme discharge events despite the fact that ASM was significantly correlated with peak flow; rather downward trends in extreme discharge are caused mostly by changes in the first 2-day storm precipitation. The ASM for extreme discharge events (which occur mostly in late fall and winter) were affected more by antecedent precipitation than evapotranspiration and hence temperature increases had weak effects on ASM.

Acknowledgement

We thank Huilin Gao and Gang Zhao from Texas A&M University for providing the DHSVM-res codes and time series of reservoir surface areas extracted from Landsat data. The research sup- ported herein was funded by the Center for Western Weather and Water Extremes (CW3E) at the Scripps Institution of Oceanography via AR Program Phase I, 4600010378 TO#15 Am 22, sponsored by the California Department of Water Resources.

References

Barth, N. A., G. Villarini, M. A. Nayak, and K. White, 2017: Mixed populations and annual flood frequency estimates in the western United States: The role of atmospheric rivers. *Water Resour. Res.*, **53**, 257-269. https://doi.org/10.1002/2016wr019064.

Begueria, S., M. Angulo-Martinez, S. M. Vicente-Serrano, J. I. Lopez-Moreno, and A. El-Kenawy, 2011: Assessing trends in extreme precipitation events intensity and magnitude using non-stationary peaks-over-threshold analysis: a case study in northeast Spain from 1930 to 2006. *Int. J. Climatol.*, **31**, 2102-2114, https://doi.org/10.1002/joc.2218.

Berghuijs, W. R., R. A. Woods, C. J. Hutton, and M. Sivapalan, 2016: Dominant flood generating mechanisms across the United States. *Geophys. Res. Lett.*, **43**, 4382-4390. https://doi.org/10.1002/2016gl068070.

Blume, T., E. Zehe, and A. Bronstert, 2007: Rainfall-runoff response, event-based runoff coefficients and hydrograph separation. *Hydrol. Sci. J.*, **52**, 843-862, https://doi.org/10.1623/hysj.52.5.843.

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

Cao, Q., N. Sun, J. Yearsley, B. Nijssen, and D. P. Lettenmaier, 2016: Climate and land cover effects on the temperature of Puget Sound streams. *Hydrol. Processes*, **30**, 2286-2304, https://doi.org/10.1002/hyp.10784.

——, T. H. Painter, W. R. Currier, J. D. Lundquist, and D. P. Lettenmaier, 2018: Estimation of Precipitation over the OLYMPEX Domain during Winter 2015/16. *J. Hydrometeor*., **19**, 143-160, https://doi.org/10.1175/jhm-d-17-0076.1.

Castillo, V. M., A. Gomez-Plaza, and M. Martinez-Mena, 2003: The role of antecedent soil water content in the runoff response of semiarid catchments: a simulation approach. *J. Hydrol.*, **284**, 114-130, https://doi.org/10.1016/s0022-1694(03)00264-6.

Cordeira, J.M., F.M. Ralph, A. Martin, N. Gaggini, J.R. Spackman, P.J. Neiman, J.J. Rutz, and R. Pierce, 2017: Forecasting Atmospheric Rivers during CalWater 2015. *Bull. Amer. Meteor. Soc.*, **98**, 449–459, https://doi.org/10.1175/BAMS-D-15-00245.1.

Crow, W. T., F. Chen, R. H. Reichle, and Q. Liu, 2017: L band microwave remote sensing and land data assimilation improve the representation of prestorm soil moisture conditions for hydrologic forecasting. *Geophys. Res. Lett*., **44**, 5495-5503, https://doi.org/10.1002/2017gl073642.

Daly, C., R. Neilson, and D. Phillips, 1994: A statistical topographic model for mapping climatological precipitation over mountainous terrain. *J. Appl. Meteor*., **33**, 140–158, https://doi.org/10.1175/1520-0450(1994)033,0140:ASTMFM.2.0.CO;2.

——, M. Halbleib, J. I. Smith, W. P. Gibson, M. K. Doggett, G. H. Taylor, J. Curtis, and P. P. Pasteris, 2008: Physiographically sensitive mapping of climatological temperature and precipitation across the conterminous United States. *Int. J. Climatol*., **28**, 2031–2064, https://doi.org/10.1002/joc.1688.

DeFlorio, M. J., D. E. Waliser, B. Guan, D. A. Lavers, F. M. Ralph, and F. Vitart, 2018: Global Assessment of Atmospheric River Prediction Skill. *J. Hydrometeor*., **19**, 409-426, https://doi.org/10.1175/jhm-d-17-0135.1.

Dettinger, M., F. Ralph, T. Das, P. Neiman, and D. Cayan, 2011: Atmospheric rivers, floods and the water resources of California. *Water*, **3**, 445–478, https://doi.org/10.3390/w3020445.

Entekhabi, D., and Coauthors, 2010: The Soil Moisture Active Passive (SMAP) Mission. *Proc. IEEE*, **98**, 704-716, https://doi.org/10.1109/jproc.2010.2043918.

Gershunov, A., T. Shulgina, F. M. Ralph, D. A. Lavers, and J. J. Rutz, 2017: Assessing the climate-scale variability of atmospheric rivers affecting western North America. *Geophys. Res. Lett*., **44**, 7900-7908, https://doi.org/10.1002/2017gl074175.

Guan, B., N. P. Molotch, D. E. Waliser, E. J. Fetzer, and P. J. Neiman, 2010: Extreme snowfall events linked to atmospheric rivers and surface air temperature via satellite measurements. *Geophys. Res. Lett*., **37**, 6, https://doi.org/10.1029/2010gl044696.

——, and D. E. Waliser, 2015: Detection of atmospheric rivers: Evaluation and application of an algorithm for global studies. *J. Geophys. Res.: Atmos*., **120**, 12514-12535, https://doi.org/10.1002/2015jd024257.

Gupta, H. V., H. Kling, K. K. Yilmaz, and G. F. Martinez, 2009: Decomposition of the mean squared error and NSE performance criteria: Implications for improving hydrological modeling. *J. Hydrol.*, **377**, 80–91, https://doi.org/10.1016/j.jhydrol.2009.08.003.

Hamlet, A. F., and D. P. Lettenmaier, 2007: Effects of 20th century warming and climate variability on flood risk in the western U.S. *Water Resour. Res*., **43**, W06427, https://doi.org/10.1029/2006WR005099.

Hecht, C. W., and Cordeira, J. M., 2017: Characterizing the influence of atmospheric river orientation and intensity on precipitation distributions over North Coastal California. *Geophys. Res. Lett*., **44**, 9048-9058, https://doi.org/10.1002/2017GL074179.

Hughes, M., K. M. Mahoney, P. J. Neiman, B. J. Moore, M. Alexander, and F. M. Ralph, 2014: The Landfall and Inland Penetration of a Flood-Producing Atmospheric River in Arizona. Part II: Sensitivity of Modeled Precipitation to Terrain Height and Atmospheric River Orientation. *J. Hydrometeor*., **15**, 1954-1974, https://doi.org/10.1175/jhm-d-13-0176.1.

Kalnay, E., and Coauthors, 1996: The NCEP/NCAR 40-year reanalysis project. *Bull. Amer. Meteor. Soc.*, **77**, 437-471, https://doi.org/10.1175/1520-0477(1996)077<0437:tnyrp>2.0.co;2.

Kerr, Y. H., and Coauthors, 2010: The SMOS Mission: New Tool for Monitoring Key Elements of the Global Water Cycle. *Proc. IEEE*, **98**, 666-687, https://doi.org/10.1109/jproc.2010.2043032.

Lamjiri, M. A., M. D. Dettinger, F. M. Ralph, and B. Guan, 2017: Hourly storm characteristics along the US West Coast: Role of atmospheric rivers in extreme precipitation. *Geophys. Res. Lett*., **44**, 7020-7028, https://doi.org/ 10.1002/2017gl074193.

Lang, M., T. Ouarda, and B. Bobee, 1999: Towards operational guidelines for over-threshold modeling. *J. Hydrol.*, **225**, 103-117, https://doi.org/10.1016/s0022-1694(99)00167-5.

Lavers, D. A., and G. Villarini, 2015: The contribution of atmospheric rivers to precipitation in Europe and the United States. *J. Hydrol.*, **522**, 382-390, https://doi.org/10.1016/j.jhydrol.2014.12.010.

——, D. E. Waliser, F. M. Ralph, and M. D. Dettinger, 2016: Predictability of horizontal water vapor transport relative to precipitation: Enhancing situational awareness for forecasting western US extreme precipitation and flooding. *Geophys. Res. Lett.*, **43**, 2275-2282, https://doi.org/10.1002/2016gl067765.

Leung, L. R., and Y. Qian, 2009: Atmospheric rivers induced heavy precipitation and flooding in the western US simulated by the WRF regional climate model. *Geophys. Res. Lett.*, **36**, 6, https://doi.org/10.1029/2008gl036445.

Mallakpour, I., and G. Villarini, 2017: Analysis of changes in the magnitude, frequency, and seasonality of heavy precipitation over the contiguous USA. *Theor. Appl. Climatol.*, **130**, 345-363, https://doi.org/10.1007/s00704-016-1881-z.

Martin, A., and Coauthors, 2018: Evaluation of Atmospheric River Predictions by the WRF Model Using Aircraft and Regional Mesonet Observations of Orographic Precipitation and Its Forcing. *J. Hydrometeor.*, **19**, 1097-1113, https://doi.org/10.1175/jhm-d-17-0098.1.

Maurer, E. P., A. W. Wood, J. C. Adam, D. P. Lettenmaier, and B. Nijssen, 2002: A long-term hydrologically based dataset of land surface fluxes and states for the

conterminous United States. *J. Clim*., **15**, 3237-3251, https://doi.org/ 10.1175/1520-0442(2002)015<3237:althbd>2.0.co;2.

Mondal, A., and P. P. Mujumdar, 2015: Modeling non-stationarity in intensity, duration and frequency of extreme rainfall over India. *J. Hydrol.*, **521**, 217-231, https://doi.org/10.1016/j.jhydrol.2014.11.071.

Nayak, M.A., G. Villarini, and D.A. Lavers, 2014: On the skill of numerical weather prediction models to forecast atmospheric rivers over the central United States, *Geophys. Res. Lett*., **41**, 4354-4362, https://doi.org/10.1002/2014GL060299.

——, and ——, 2018: Remote sensing-based characterization of rainfall during atmospheric rivers over the central United States. *J. Hydrol.*, **556**, 1038-1049, https://doi.org/10.1016/j.jhydrol.2016.09.039.

Neiman, P. J., F. M. Ralph, A. B. White, D. E. Kingsmill, and P. O. G. Persson, 2002: The statistical relationship between upslope flow and rainfall in California's coastal mountains: Observations during CALJET. *Mon. Wea. Rev.*, **130**, 1468-1492, https://doi.org/10.1175/1520-0493(2002)130<1468:tsrbuf>2.0.co;2.

——, L. J. Schick, F. M. Ralph, M. Hughes, and G. A. Wick, 2011: Flooding in Western Washington: The Connection to Atmospheric Rivers. *J. Hydrometeor.*, **12**, 1337-1358, https://doi.org/10.1175/2011jhm1358.1.

——, F. M. Ralph, B. J. Moore, M. Hughes, K. M. Mahoney, J. M. Cordeira, and M. D. Dettinger, 2013: The Landfall and Inland Penetration of a Flood-Producing Atmospheric River in Arizona. Part I: Observed Synoptic-Scale, Orographic, and Hydrometeorological Characteristics. *J. Hydrometeor.*, **14**, 460-484, https://doi.org/ 10.1175/jhm-d-12-0101.1.

——, ——, ——, and R. J. Zamora, 2014: The Regional Influence of an Intense Sierra Barrier Jet and Landfalling Atmospheric River on Orographic Precipitation in Northern California: A Case Study. *J. Hydrometeor.*, **15**, 1419-1439, https://doi.org/10.1175/jhm-d-13-0183.1.

Nikolopoulos, E. I., E. N. Anagnostou, M. Borga, E. R. Vivoni, and A. Papadopoulos, 2011: Sensitivity of a mountain basin flash flood to initial wetness condition and rainfall variability. *J. Hydrol.*, **402**, 165-178, https://doi.org/10.1016/j.jhydrol.2010.12.020.

Paltan, H., D. Waliser, W. H. Lim, B. Guan, D. Yamazaki, R. Pant, and S. Dadson, 2017: Global Floods and Water Availability Driven by Atmospheric Rivers. *Geophys. Res. Lett.*, **44**, 10387-10395, https://doi.org/10.1002/2017gl074882.

Penna, D., H. J. Tromp-van Meerveld, A. Gobbi, M. Borga, and G. Dalla Fontana, 2011: The influence of soil moisture on threshold runoff generation processes in an alpine headwater catchment. *Hydrol. Earth Syst. Sci.*, **15**, 689-702, https://doi.org/10.5194/hess-15-689-2011.

Radatz, T. F., A. M. Thompson, and F. W. Madison, 2013: Soil moisture and rainfall intensity thresholds for runoff generation in southwestern Wisconsin agricultural watersheds. *Hydrol. Processes*, **27**, 3521-3534, https://doi.org/10.1002/hyp.9460.

Ralph, F.M., P.J. Neiman, D.E. Kingsmill, P.O. Persson, A.B. White, E.T. Strem, E.D. Andrews, and R.C. Antweiler, 2003: The Impact of a Prominent Rain Shadow on Flooding in California's Santa Cruz Mountains: A CALJET Case Study and Sensitivity to the ENSO Cycle. *J. Hydrometeor.*, **4**, 1243–1264, https://doi.org/10.1175/1525-7541(2003)004<1243:TIOAPR>2.0.CO;2.

——, ——, G. Wick, S. Gutman, M. Dettinger, D. Cayan, and A. White, 2006: Flooding on California's Russian River: Role of atmospheric rivers. *Geophys. Res. Lett*., **33**, L13801, https://doi.org/10.1029/2006GL026689.

——, and Coauthors, 2013a: The Emergence of Weather-Related Test Beds Linking Research and Forecasting Operations. *Bull. Amer. Meteor. Soc*., 94, 1187–1211, https://doi.org/10.1175/BAMS-D-12-00080.1

——, T. Coleman, P. Neiman, R. Zamora, and M. Dettinger, 2013b: Observed Impacts of Duration and Seasonality of Atmospheric-River Landfalls on Soil Moisture and Runoff in Coastal Northern California. *J. Hydrometeor.*, **14**, 443-459, https://doi.org/10.1175/JHM-D-12-076.1.

——, and Coauthors, 2018a: ARTMIP-early start comparison of atmospheric river detection tools: how many atmospheric rivers hit northern California's Russian River watershed? *Clim. Dyn*., 1-22, https://doi.org/10.1007/s00382-018-4427-5.

——, M. D. Dettinger, M. M. Cairns, T. J. Galarneau, and J. Eylander, 2018b: DEFINING "ATMOSPHERIC RIVER" How the Glossary of Meteorology Helped Resolve a Debate. *Bull. Amer. Meteor. Soc.*, **99**, 837-839, https://doi.org/10.1175/bams-d-17-0157.1.

——, J.J. Rutz, J.M. Cordeira, M. Dettinger, M. Anderson, D. Reynolds, L.J. Schick, and C. Smallcomb, 2019: A Scale to Characterize the Strength and Impacts of Atmospheric Rivers. *Bull. Amer. Meteor. Soc.*, **100**, 269–289, https://doi.org/10.1175/BAMS-D-18-0023.1.

Rutz, J. J., W. J. Steenburgh, and F. M. Ralph, 2014: Climatological Characteristics of Atmospheric Rivers and Their Inland Penetration over the Western United States. *Mon. Wea. Rev.*, **142**, 905-921, https://doi.org/ 10.1175/mwr-d-13-00168.1.

Schaake, J., A. Henkel, and S. Cong, 2004: Application of PRISM climatologies for hydrologic modeling and forecasting in the western U.S. *18th Conf. on Hydrology*, Seattle, WA, Amer. Meteor. Soc., 5.3. [Available online at https://ams.confex.com/ams/84Annual/techprogram/paper_72159.htm.]

Sharma, A., C. Wasko, and D. P. Lettenmaier, 2018: If precipitation extremes are increasing, why aren't floods? *Water Resour. Res.*, **54**. https://doi.org/10.1029/2018WR023749.

Shepard, D. S., 1984: Spatial statistics and models. *Computer Mapping: The SYMAP Interpolation Algorithm*, G. L. Gaile and C. J. Willmott, Eds., D. Reidel, 133–145.

Shields, C.A., and coauthors, 2018: Atmospheric River Tracking Method Intercomparison Project (ARTMIP): Project goals and experimental design. *Geosci. Model Dev.*, **11**, 2455-2474, https://doi.org/10.5194/gmd-11-2455-2018.

Thomas, N. W., A. A. Amado, K. E. Schilling, and L. J. Weber, 2016: Evaluating the efficacy of distributed detention structures to reduce downstream flooding under variable rainfall, antecedent soil, and structural storage conditions. *Adv. Water Resour.*, **96**, 74-87, https://doi.org/10.1016/j.advwatres.2016.07.002.

USWRC, 1982. Guidelines for Determining Flood Flow Frequency. Bulletin 17B, United States Water Resources Committee, Washington, DC, USA.

Waliser, D., and B. Guan, 2017: Extreme winds and precipitation during landfall of atmospheric rivers. *Nat. Geosci*., **10**, 179-U183, https://doi.org/10.1038/ngeo2894.

Wasko, C., and A. Sharma, 2017: Global assessment of flood and storm extremes with increased temperatures. *Sci. Rep*., **7**, 8, https://doi.org/10.1038/s41598-017-08481-1.

Wigmosta, M. S., L. W. Vail, and D. P. Lettenmaier, 1994: A distributed hydrology-vegetation model for complex terrain. *Water Resour. Res.*, **30**, 1665-1679, https://doi.org/10.1029/94WR00436.

Woldemeskel, F., and A. Sharma, 2016: Should flood regimes change in a warming climate? The role of antecedent moisture conditions. *Geophys. Res. Lett*., **43**, 7556-7563. https://doi.org/10.1002/2016gl069448.

Wood, A. W., and D. P. Lettenmaier, 2006: A test bed for new seasonal hydrologic forecasting approaches in the western United States. *Bull. Amer. Meteor. Soc.*, **87**, 1699-+, https://doi.org/10.1175/bams-87-12-1699.

Zhao, G., H. L. Gao, B. S. Naz, S. C. Kao, and N. Voisin, 2016: Integrating a reservoir regulation scheme into a spatially distributed hydrological model. *Adv. Water Resour.*, **98**, 16-31, https://doi.org/10.1016/j.advwatres.2016.10.014.

Zhu, Y., and R. Newell, 1998: A proposed algorithm for moisture fluxes from atmospheric rivers. *Mon. Wea. Rev.*, **126**, 725-735, https://doi.org/10.1175/1520-0493(1998)126<0725:APAFMF>2.0.CO;2.

Chapter 3. Floods due to atmospheric rivers along the U.S. West Coast: The role of hydrological initial conditions in a warming climate

This chapter has been submitted in its current form to the *Journal of Hydrometeorology* and now in revision. The supplementary materials for this chapter are provided in Appendix B.

Cao, Q., A. Gershunov, T. Shulgina, F.M. Ralph, N. Sun, and D.P. Lettenmaier, 2020: Floods due to atmospheric rivers along the U.S. West Coast: The role of hydrological initial conditions in a warming climate, *Journal of Hydrometeorology*, (in revision).

Abstract

Precipitation extremes are projected to become more frequent along the U.S. West Coast due to increased atmospheric river (AR) activity, but the frequency of less intense precipitation events may decrease. Antecedent soil moisture (ASM) conditions can have a large impact on flood responses especially if decreased pre-storm precipitation and increased antecedent evaporative demand in a warming climate result in reduced soil moisture at the onset of extreme precipitation events. We examine the impact of ASM on AR-related floods in a warming climate in three basins that form a transect along the U.S. Pacific Coast: the Chehalis River basin in Washington, the Russian River basin in Northern California, and the Santa Margarita River basin in Southern California. We ran the Distributed Hydrology-Soil-Vegetation Model (DHSVM) over each of the three river basins using forcings downscaled from 10 Global Climate Models (GCMs). We examined the dynamic role of ASM by comparing the changes of the largest 50, 100 and 150 extreme events in two periods, 1951-2000 and 2050-2099. In the Chehalis basin, the

projected fraction of AR-related extreme discharge events slightly decreases. In the Russian basin, this fraction increases, however, and more substantially so in the Santa Margarita basin. This is due to increases in AR-related extreme precipitation events, as well as the fact that the relationship of AR-related extreme precipitation to extreme discharge is strengthened by projected increases in year-to-year volatility of annual precipitation in California, which increases the likelihood of concurrent occurrence of large storms and wet ASM conditions.

3.1 Introduction

Atmospheric rivers (ARs) are responsible for most of the storm events leading to extreme precipitation and runoff along the U.S. West Coast (e.g. Ralph et al., 2006; Dettinger et al., 2011; Neiman et al., 2011). Most floods and flood damages in the West have been explicitly linked to ARs (Barth et al, 2017; Konrad and Dettinger, 2017, Corringham et al., 2019). Over the past decade, several studies have examined the potential impact of climate change on AR land-falling activity in this region, in order to better project the changes in extreme precipitation associated with ARs. Using seven Global Climate Models (GCMs), Dettinger (2011) projected that the number of landfalling ARs in California would increase by ~30% in the 21st century. The peak AR intensity, storm temperature and the length of AR season would increase as well. Espinoza et al. (2018) cross-compared their results with previous studies focused on the U.S. West Coast (Hagos et al., 2015; Gao et al., 2015; Payne and Magnusdottir, 2015; Warner et al., 2015; Shields and Kiehl, 2016). Despite great discrepancies in the projected landfalling AR frequency changes, all of these studies showed that the number

of landfalling AR days would substantially increase by the end of 21st century along the U.S. West Coast under the Representative Concentration Pathway (RCP) 8.5 emission scenario. So would the AR intensity, in terms of the integrated water vapor transport (IVT) (Warner et al., 2015; Espinoza et al., 2018). These changes in ARs will increase both the number of extreme precipitation days and the peak precipitation intensity along the U.S. West Coast where they interact with complex topography and lead to orographic precipitation (Hagos et al., 2015; Warner et al., 2015).

Although extreme precipitation associated with ARs will increase in a warming climate, a commensurate increase in high flows may not be expected on the basis of theory and observational evidence (Sharma et al., 2018). Antecedent hydrological conditions play an important role in the linkage between extreme precipitation and flooding. By analyzing historical station records of precipitation and streamflow over the contiguous U.S. during 1950-2000, Ivancic and Shaw (2015) found that extreme (99^{th} percentile) precipitation only led to extreme (99^{th} percentile) discharge 36% of the time. When the antecedent soil moisture (ASM) conditions were wet, the percentage increased to 62%. Furthermore, Wasko and Sharma (2017) examined the sensitivity of extreme daily precipitation and discharge to changes in daily temperature based on station records. They found that changes in heavy rainfall events linked with observed warming did not lead to similar changes in streamflow in most regions globally possibly due to initial moisture conditions. Wasko and Nathan (2019) assessed the influence of changes in rainfall and ASM on trends in flooding at stream gauges in Australia. They used a quantile regression method in order to assign a recurrence interval to trends beyond the mean sampling frequency. They found that only for the most extreme flood flows with

return period larger than 40 years, the trend in peak flows was similar in magnitude to the trend in extreme precipitation, indicating that ASM is especially important in modulating the response of floods that are not very rare. Bennett et al. (2018) examined the effect of ASM (using antecedent precipitation as a proxy) on flood volume in 100 Australian catchments with sub-daily streamflow and precipitation observations. They found that although flood-producing precipitation was the dominant driver of flood magnitudes, the elasticity of flow to antecedent precipitation was about one third of the elasticity to flood-producing precipitation, however the influence of antecedent precipitation weakened as event magnitudes increased.

Changes in ASM and their impact on flood response are especially important in areas like the U.S. West Coast where flooding is typically caused by large single precipitation excess events (Berghuijs et al., 2016) and where precipitation is strongly winter-dominant, with many potentially flood-inducing events occurring relatively early in the wet season when soils tend to be dry. Cao et al. (2019) examined the role of ASM in historical AR flooding on California's Russian River Basin, a coastal watershed whose winter precipitation extremes are dominated by ARs. They showed that low ASM was an offsetting factor for the three AR Category 5 (the highest category of ARs on the scale defined by Ralph et al. (2019)) events that did not lead to major flooding during 1980-2017. They also found that the ASM was affected more by antecedent precipitation than evapotranspiration (ET) and hence temperature increases had relatively modest effects on ASM.

Warming is expected to become increasingly pronounced by the end of 21st century especially under the RCP8.5 global emissions scenario (Rogelj et al., 2012) as

will its likely impact on ASM. Meanwhile, studies suggest that although the frequency of extreme precipitation will increase, the frequency of low to medium intensity precipitation will decrease and offset the change in annual precipitation along the U.S. West Coast (e.g. Pierce et al. 2013b; Polade et al., 2017; Gershunov et al., 2019). By analyzing the ensemble average of daily precipitation of 28 GCMs from the Phase 5 of Coupled Model Intercomparison Project (CMIP5), Polade et al. (2014) found that there would be slightly more dry days (with precipitation < 1 mm) in this region by the end of the 21st century under the RCP8.5 emissions scenario.

In California, the climate projections indicated robust increases in the frequency of heavy and extreme precipitation, but decreases in the frequency of low-medium intensity precipitation, with the latter mainly occurring in the fall and spring months (Pierce et al. 2013b; Polade et al., 2017). Gershunov et al. (2019) examined the role of ARs in future precipitation regime change over the western U.S. using projections from 16 CMIP5 GCMs. They identified five GCMs that they deemed most realistic in terms of their performance in capturing the statistics of historical AR events and AR contribution to total annual precipitation (they denoted these GCMs as "Real-5"). They found that the Real-5 GCMs projected, with a significant reduction of projection uncertainty in comparison with the full ensemble, less frequent but more intense precipitation from ARs. This in turn translated into more volatility of year-to-year total annual precipitation, especially in California. This suggests a more flood- and drought-prone future precipitation regime and may affect the importance of ASM considerations.

Along the U.S. West Coast, Gershunov et al. (2019) found (based on the Real-5 GCMs) the frequency of the heavy (90^{th}-99^{th} percentile) and extreme (>99^{th} percentile)

precipitation generally increased due to increased AR activity, with the greatest increase projected in northern California, while the frequency of medium-intensity (30^{th}-90^{th} percentile) precipitation decreased due to non-AR events. These changes in low-medium intensity precipitation would lead to lengthening dry periods (punctuated by stronger precipitation extremes); taken together with increased temperature, they may exert a synergic effect on ASM thus mitigating flood response to increased extreme precipitation.

A proliferation of previous studies has assessed projected future changes in floods over the western U.S. Much of the work has focused on mountainous regions with snow cover since snow-dominant basins are more sensitive than rain-dominant basins to increased temperatures where warming tends to mitigate the effects of increased precipitation on flooding by causing reductions in spring snowpack (Hamlet and Lettenmaier, 2007). Some previous work evaluated coastal river basins, most of which have little to modest snow effect. Such basins will be the focus of our study. Most of the earlier studies used the Variable Infiltration Capacity (VIC) land surface model (Liang et al., 1994) driven by downscaled GCM forcings to examine the projected changes in extreme flows. They projected that flood risks, in terms of both magnitude and frequency of floods, will increase along the U.S. West Coast over the remainder of the 21^{st} century due to projected increases in winter precipitation (e.g. Salathé et al., 2014; Tohver et al., 2014; Pagán et al., 2016; Mallakpour et al., 2018; Maurer et al., 2018; Naz et al., 2016, 2018). All of these studies focused on assessing the projected changes in floods, yet none examined the changes in the context of the AR contribution or the role of ASM.

Although the number of AR-related extreme precipitation events is projected to increase in a warming climate in many GCMs (and notably those that best reproduce historic storm statistics, e.g., Real-5), ASM can have a large impact on AR-related floods especially if decreased pre-storm low-medium intensity precipitation frequency and increased antecedent evaporative demand result in reduced soil moisture at the onset of extreme precipitation events. Given this background, we address here the following motivating questions:

1) How will ARs' contribution to floods change along the U.S. West Coast by the end of 21st century?

2) How will climate change affect ASM and what is its role in modulating flood response in a warming climate?

3.2 Study region

We selected three watersheds that form a transect along the U.S. Pacific Coast: the Chehalis River basin in Washington State, the Russian River basin in Northern California, and the Santa Margarita River basin in Southern California. The three river basins have drainage areas of 5400 km^2, 3850 km^2 and 1870 km^2, respectively (see Figure 3.1). Gershunov et al. (2019) used the first two and another one next to the Santa Margarita River basin to represent projected precipitation regimes in the western U.S. coastal domain. We selected these three basins as our study domain because 1) they are coastal watersheds frequented by ARs and their geographical locations reflect different AR landfalling signatures; 2) all of them are rain-dominant basins with relatively modest topographic variations (elevation ranges of 0-1429 m, 0-1324 m and 143-1736 m,

respectively), which avoids the added complexity of the influence of snowmelt on stream flow that more commonly occurs in mountainous basins; 3) they are somewhat less developed (although there are two dams in the Russian River basin) in comparison with surrounding heavily developed and urbanized basins.

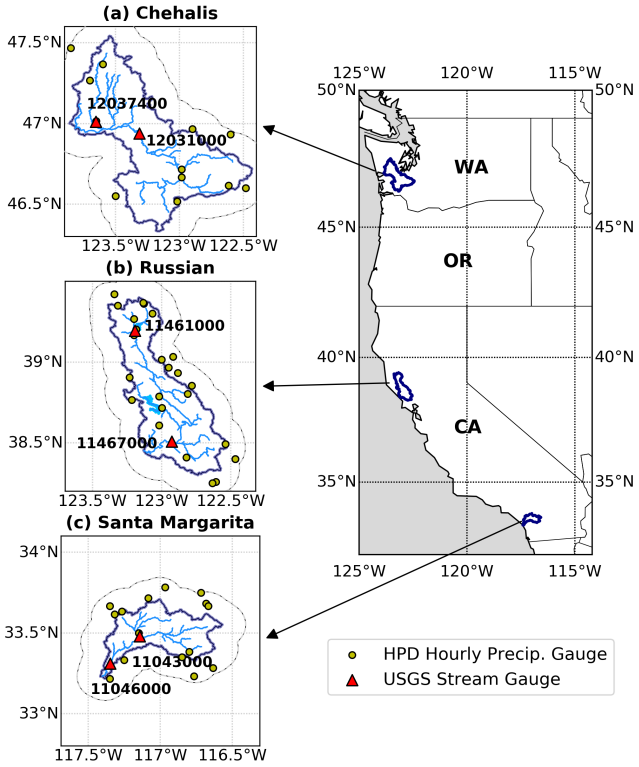

Figure 3.1 Map of study region including a) the Chehalis River basin in Washington State, b) the Russian River basin in Northern California, and c) the Santa Margarita River basin in Southern California.

The Pacific Coastal region has strongly winter-dominant precipitation and mostly dry summers. The annual precipitation of the Chehalis, Russian and Santa Margarita basins ranged between 1500-3100 mm, 500-2200 m and 145-905 mm, respectively, during water years (WY) 1951-2000, with 85%, 95% and 91% of precipitation falling between October and April.

3.3 Data and methods

We implemented the Distributed Hydrology-Soil-Vegetation Model (DHSVM) (Wigmosta et al., 1994), the parallel version (William et al. 2019), in the three basins with essentially the same model setup as in Cao et al. (2019). We ran it at an hourly time step in order to obtain hydrographs and ASM. Our analyses include two parts: 1) we examine the AR contribution to historical floods and the role of ASM on historical AR flooding in each of the basins during WY 1951-2000 using meteorological forcings from gridded observations; 2) we assess the changes in AR contribution to future flood events and the changes in ASM using downscaled GCM forcings (under the RCP8.5 scenario) by comparing simulated flood statistics for the periods WY 1951-2000 with WY 2050-2099.

3.3.1 Model implementation

We implemented the DHSVM at a spatial resolution of 150 m as in Cao et al. (2019) over the Chehalis, the Russian and the Santa Margarita river basins. We ran the model at an hourly time step in order to obtain full hydrographs of storm events from sub-daily data and soil moisture conditions. The model requires meteorological driving data at its hourly time step, the preparation of which is described in section 3.2. The model inputs include DEM, soil class, land cover type, soil depth and flow direction. The first three were taken from the NASA Shuttle Radar Topography Mission (SRTM) 90-m product, the USDA STATSGO2 data, and the US Geological Survey (USGS) 30-m GAP/LANDFIRE land cover map based on the 2001 imagery, respectively, which were resampled to model's 150 m resolution. Soil depths and flow directions were determined

using scripts included in DHSVM as described in Cao et al. (2016). In this study, we did not consider future land cover change.

We first calibrated the model in each basin using meteorological forcings from gridded observations, for which we selected USGS stream gauges with relatively complete records that go back to WY 1951 in each basin. We then ran the model and evaluated simulations of historical flood events and GCM-projected changes.

3.3.2 Meteorological forcing data

3.3.2.1 Historical events

For historical events, we used the meteorological forcings of Livneh et al. (2015) for the period of WY 1951-2000. It is a daily dataset of gridded observations at a spatial resolution of 1/16^{o}, with variables including daily precipitation, daily maximum and minimum temperature. Its wind data is from the lowest level of the National Centers for Environmental Prediction and National Centers for Atmospheric Research (NCEP/NCAR) reanalysis (Kalnay et al., 1996). DHSVM requires meteorological inputs including precipitation, wind speed, air temperature, relative humidity, downward solar and longwave radiation at the model's hourly time step. The hourly calculation of the last four variables was performed using the Mountain Microclimate Simulation Model (MTCLIM) algorithms as described and implemented by Bohn et al. (2013). Wind speed was taken to be constant throughout a day. Given that fine-timescale precipitation data is important to hydrologic predictions in small watersheds, we separately describe the hourly disaggregation of precipitation in section 3.2.3.

3.3.2.2 GCM projected changes

We used downscaled forcings from 10 GCMs (see Table B1) under the RCP8.5 emission scenario for the control period of WY 1951-2000 and the future period of WY 2050-2099. Pierce et al. (2014) statistically downscaled the CMIP5 GCM daily minimum and maximum temperature, and daily precipitation to 1/16^{o} using Localized Constructed Analogs (LOCA), with Livneh et al. (2015) dataset as the observed training dataset. As mentioned in the Introduction, Gershunov et al. (2019) evaluated the performance of 16 GCMs that archived variables sufficient to identify ARs in reproducing the key statistical features of historical landfalling AR activity and its contribution to total precipitation in comparison with an AR catalog (Gershunov et al., 2017) derived from the NCEP/NCAR reanalysis dataset. They identified five GCMs, the Real-5, that were most realistic in capturing the historical AR events. We selected 10 GCMs for this study based on their evaluation results, including the Real-5 (see Table B1). These 10 GCMs, among the 16 GCMs used in Gershunov et al. (2019), also showed relatively good credibility in reproducing observed metrics of precipitation at the seasonal, annual and decadal scales in the Southwest (Rupp et al., 2013). The disaggregation of daily to hourly meteorological inputs was the same as for historical events.

3.3.2.3 Hourly disaggregation of precipitation

Following Westra et al. (2012), we disaggregated the gridded daily precipitation in each basin to hourly using a regionalized method of fragments (denoted as "MoF") algorithm. We first collected hourly precipitation data from NOAA's Hourly Precipitation Data (HPD) database. We selected stations within a buffer of 15 km from each basin boundary and with records longer than 5 years. There are 12, 25 and 15 stations that met our criteria in the Chehalis, Russian and Santa Margarita basins,

respectively. Some stations have records that go back at least to WY 1951. We performed quality control (QC) for the hourly stations following Cao et al. (2018).

Based on the MoF method, for a given 1/16^{o} grid cell and for a given wet day, we searched for days within a moving window of ±15 days centered on that day across all years of record and across 4 nearby HPD stations, from which we selected the wet days from station data with the same previous- and next-day wetness state (i.e. precipitating or not) as the grid data to account for continuity. The selected daily station precipitation was then ranked by its absolute deviation from the gridded daily precipitation. We found up to 10 nearest neighbors with absolute deviations less than 10% of the gridded precipitation, and randomly drew one of them with probabilities determined by the previous ranking. The selected fragment (hourly ratio of station data to its daily precipitation) was then multiplied to the grid daily precipitation.

3.3.3 AR-related extreme events

We first used the Peaks Over Threshold (POT) method to identify extreme precipitation and extreme discharge events. For the latter, we also examined the annual maximum flow (AMF) events as this is a common characterization of flood flows. We then examined ARs' contribution to extreme events by identifying the ones that were coincident with AR events.

We selected POT extreme precipitation and extreme discharge events based on daily data. The POT method samples observations above a given threshold value and considers a wider range of events than the block maxima approach (e.g. Lang et al., 1999; Begueria et al., 2011; Mallakpour and Villarini, 2017). We first applied independence criteria to the data as in Cao et al. (2019). For precipitation events, they were separated

from each other by at least one day with daily precipitation below the threshold value. We selected thresholds to result in 1, 2 and 3 events per year on average, which we denote as POT_{N1P}, POT_{N2P}, and POT_{N3P}. Because we are interested in the dynamic role of ASM on floods in a changing climate, we used separate thresholds for the historical climate and future climate periods, which results in same numbers of events for the two periods and keeps a focus on the largest ones. For extreme discharge, we used criteria on the interval between two peaks and a relative threshold on the intermediate flow from the U.S. Water Resources Council (USWRC, 1982). Similar to precipitation events, we set thresholds (separately for the two climate periods) for daily streamflow at each stream gauge to result 1, 2 and 3 extreme discharge events per year on average, which we denote as POT_{N1D}, POT_{N2D} and POT_{N3D}.

To examine the AR contribution to extreme events, we identified the POT events as well as AMF that were coincident with AR events. For historical events, we used the AR catalog of Gershunov et al. (2017), which is based on the NCEP/NCAR reanalysis (denoted as "SIO-R1"). For the GCM projected changes, we used the AR catalogs of Gershunov et al. (2019) which they developed through application of the same automated AR detection scheme that they applied to historical observations to daily GCM output. For each catalog, we extracted the grid cells that intersected each basin and identified landfalling ARs.

3.3.4 The role of ASM

We examined the role of ASM on historical floods by evaluating the relationship between precursor soil moisture and storm runoff-precipitation ratios, which we took as storm total runoff volume divided by storm total precipitation. We used the definition of

precipitation and runoff events following Cao et al. (2019). For a POT extreme precipitation event selected based on daily data, we calculated the storm total precipitation based on hourly forcings. We took the beginning of a precipitation event as the first hour with precipitation exceeding a certain threshold and the end of event as the hour with precipitation dropping below that threshold. The summation of precipitation over the event hours is the storm total precipitation. For a POT extreme runoff event selected based on daily data, we calculated the storm total runoff using hourly model simulations. The start of a runoff event is the hour of the rise of the hydrograph and the end of event is determined by the constant-*k* method of Blume et al. (2007) but no longer than three days after the peak hour or the start of the following event. We defined ASM as the minimum value of the hourly surface-layer soil moisture from model simulations within the 24 hours prior to the start of a precipitation event.

We examined the role of ASM in modulating flood response in a warming climate in two ways. We first examined the changes in the connection between extreme precipitation and extreme discharge events during two periods, WY 1951-2000 and WY 2050-2099. Specifically, we examined changes in the probability of extreme precipitation events leading to extreme discharge events of the same POT threshold, which we denote as $Pr(POT_{N1D}|POT_{N1P})$, $Pr(POT_{N2D}|POT_{N2P})$ and $Pr(POT_{N3D}|POT_{N3P})$ for events with threshold set to 1, 2 and 3 events per year on average. We then examined the role of ASM in modulating flood response to extreme precipitation by examining changes in the magnitude of peak flows. We focused on the AMF and its related storm precipitation and ASM. Since the patterns in the changes are similar during the POT extreme discharge events, the results of the POT event analysis were shown in the supplement.

3.4 Results

3.4.1 Model evaluation

For the streamflow evaluation, we selected stream gauges with relatively complete daily records going back to the 1950s throughout each basin, including 6, 6 and 3 gauges in the Chehalis, Russian and Santa Margarita basins respectively. We used the Kling-Gupta efficiency (KGE) (Gupta et al., 2009), normalized root-mean-square error (NRMSE) and relative bias to evaluate the goodness-of-fit between daily streamflow observations and aggregated daily simulations at each gauge (see Table B2).

In the Chehalis basin, KGE ranged from 0.69 to 0.89; NRMSE ranged from 0.36 to 0.49, and the relative bias ranged from -10% to 23% across the six gauges during the calibration period (1986-2000). Model performance in the verification period (1971-1985) was similar to that during the calibration period. In the Russian basin, the downstream gauges are influenced by two reservoirs. We obtained the naturalized flows at these gauges by calculating the difference between simulated streamflow without and with implementation of the DHSVM reservoir module (Zhao et al., 2016) at each gauge and then adding the difference back to the observations, following Cao et al. (2019). After doing so, KGE ranged from 0.68 to 0.93; and NRMSE ranged from 0.24 to 0.46, and the relative bias ranged from -6% to 28% across the six gauges during the calibration period. KGE was highest at the downstream-most gauge. In the Santa Margarita basin, KGE ranged from 0.53 to 0.74; NRMSE ranged from 0.66 to 0.68, and the relative bias ranged from -32% to 3% across three gauges during the calibration period. The NRMSE was larger in this basin compared with the other two partly due to its smaller magnitude

of streamflow. We also evaluated model performance for peak flows (see Figure 3.1 for stream gauge locations). Figure 3.2 shows the distribution of simulated peak (daily) discharge in POT_{N3D} in comparison with observations during the period of WY 1951-2000 at selected upstream and downstream USGS gauges. The simulations matched the observations reasonably well in each of the basins, except for some general underestimation at the upstream gauge in the Russian and some overestimation for the upper tail at the downstream gauge in the Santa Margarita.

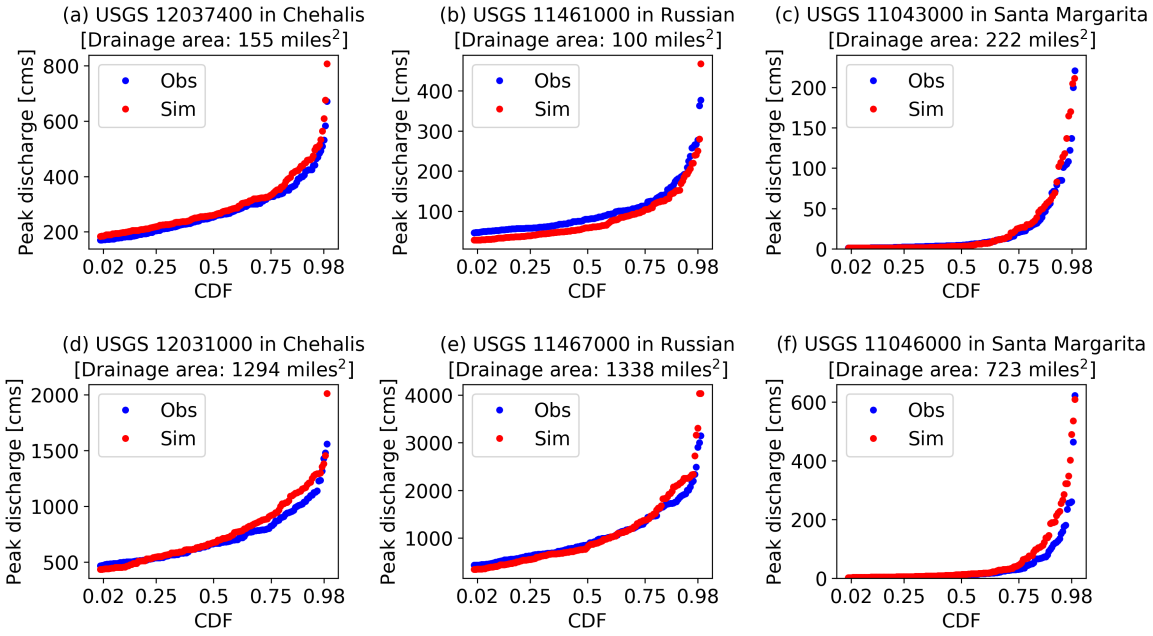

Figure 3.2 Comparison of simulated and observed POT_{N3D} (extreme discharge events with threshold set to three events per year) at selected upstream (upper row) and downstream (bottom row) USGS gauges in three basins during the period WY 1951-2000.

3.4.2 Historical events

3.4.2.1 Role of ARs

We set thresholds of POT extreme events during WY 1951-2000 to make sure there were 1, 2 and 3 events per year on average. For extreme precipitation, 60%-74%

(depending on the threshold used) of the POT events were coincident with ARs in the Chehalis River basin, based on the SIO-R1 AR date catalog (see Table B3). In the Russian River basin, the AR-related percentages were 95%-98%, while in the Santa Margarita River basin, the AR-related percentages were 60%-78%. The AR-related percentages increased as the POT threshold increased in all three basins, and the percentages were highest in the Russian River basin, which is located in the vicinity of the most intense AR activity along the U.S. west coast. In particular, ARs have the greatest contribution to total annual precipitation along the U.S. west coast in this area (Gershunov et al., 2017). The AR-related percentages were lower for the future climate projections based on the ensemble mean of all 10 GCMs, 51%-60%, 65%-75% and 31%-41% respectively in three basins (see Table 3.1). The AR-related percentages based on the mean of Real-5 GCMs matched with the observations slightly better than the full ensemble except in the Santa Margarita basin (see Table B3).

Table 3.1 Fraction [%] of AR-related POT extreme precipitation events, POT extreme discharge events, and annual maximum flow (AMF) events in three river basins based on the ensemble average of 10 GCMs.

River Basin	Period	Extreme precipitation events			Extreme discharge events			
		POT_{N1P}	POT_{N2P}	POT_{N3P}	POT_{N1D}	POT_{N2D}	POT_{N3D}	AMF
Chehalis	WY 1951-2000	60	55	51	58	52	47	51
	WY 2050-2099	58	54	50	58	50	45	51
	Change	-2	-1	-1	0	-2	-2	0
Russian	WY 1951-2000	75	70	65	67	56	48	59
	WY 2050-2099	77	72	65	71	62	54	63
	Change	2	2	0	4	6	6	4
Santa Margarita	WY 1951-2000	41	36	31	34	29	25	30
	WY 2050-2099	46	41	38	43	37	33	38
	Change	5	5	7	9	8	8	8

The AR contribution to extreme discharge also increased as the POT threshold increased. 66%-80% of the POT events were coincident with ARs in the Chehalis basin, based on the SIO-R1 AR date catalog (see Table B4). In the Russian basin, the AR-related percentages were 85%-98%, while in the Santa Margarita basin, the AR-related percentages were 51%-72%. Figure 3.3 shows the POT_{N3D} events that were coincident with ARs in three basins based on the SIO-R1 catalog. In terms of AMF events, they were 70%, 86% and 60% coincident with landfalling ARs in the three basins respectively. The AR-related percentages of POT events were lower based on the ensemble mean of all 10 GCMs, which were 47%-58%, 48%-67% and 25%-34% respectively in three basins (see Table 3.1). The AR-related percentages of AMF events were 51%, 59% and 30% respectively in the three basins, based on the ensemble mean of all 10 GCMs. Similar to extreme precipitation events, the AR-related percentages based on the mean of the Real-5 GCMs matched the observations slightly better than the full ensemble except in the Santa Margarita basin (see Table B4). In summary, the GCM-based results capture the geographic pattern of historical relative contributions of ARs to extreme precipitation and discharge in the three basins, although the models tend to underestimate the magnitudes of these contributions.

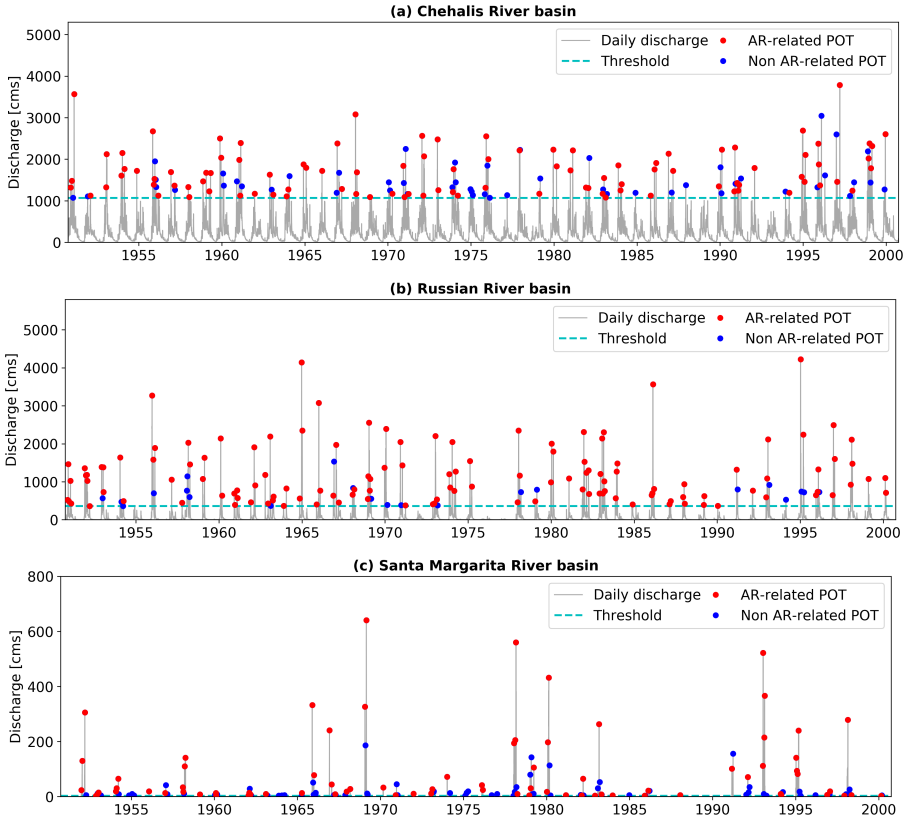

Figure 3.3 AR-related POT_{N3D} (extreme discharge events with threshold set to 3 events per year on average) based on simulated daily streamflow at basin outlets of a) Chehalis River basin, b) Russian River basin, and c) Santa Margarita River basin during WY 1951-2000.

3.4.2.2 Role of ASM

We examined the relationship between the ASM and runoff ratio in the POT_{N3P} events. Our previous analysis in the Russian basin (Cao et al., 2019) showed that the runoff ratio was much more strongly related to ASM than to storm precipitation. The same was found in the other two basins (see Figure 3.4). Following Crow et al. (2017), we used the Spearman rank correlation coefficient *Rs* to evaluate the strength of the potentially nonlinear relationship between ASM and the runoff ratio, which were 0.76, 0.81 and 0.74 in the three basins respectively.

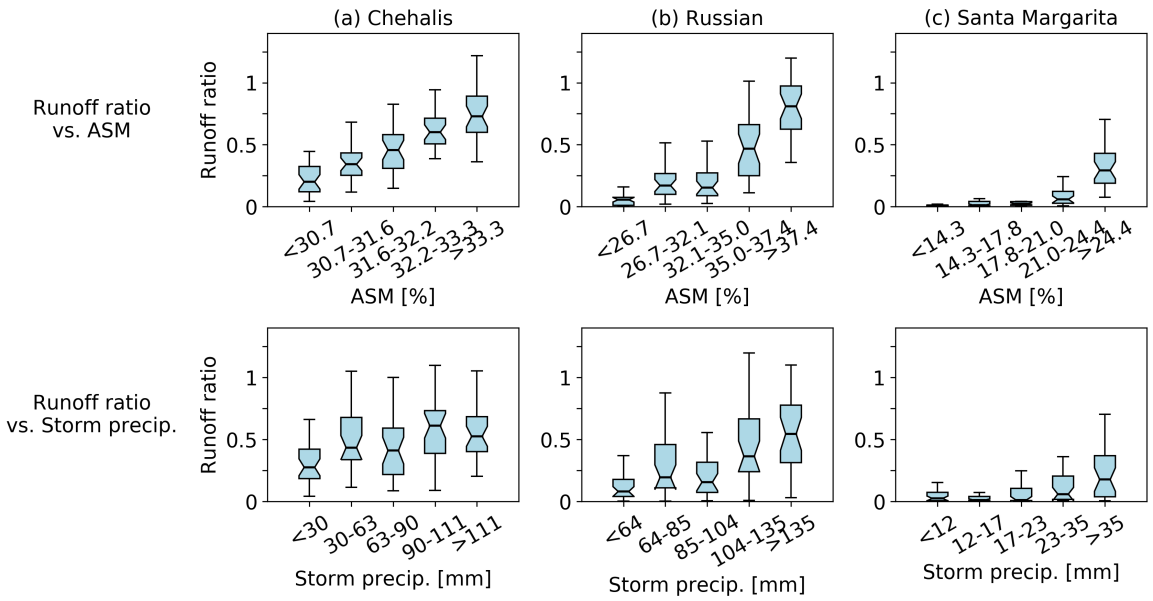

Figure 3.4 Box plots with an interval of 20th percentile of the simulation-based runoff ratio versus the antecedent soil moisture (ASM) and storm total precipitation of POT_{N3P} (extreme precipitation events with threshold set to 3 events per year on average) in three basins during WY 1951-2000.

In our previous analysis, we also showed that when ASM is high or storm precipitation is sufficiently large, extreme precipitation events can lead to extreme discharge events of the same POT threshold (see Figure 12 in Cao et al. (2019)). On the other hand, if the ASM is low, extreme precipitation may not lead to extreme discharge of the same POT threshold. In the following section, we examine whether the future increases in storm precipitation could outweigh the effects of future changes in ASM and thus enhance the connection of extreme precipitation and extreme discharge.

3.4.3 GCM projected changes

3.4.3.1 Changes in AR contribution to extreme precipitation and discharge events

We examined changes in AR contributions to extreme precipitation and extreme discharge events by comparing the periods WY 1951-2000 with WY 2050-2099 based on the ensemble mean of all 10 GCMs (see Table 3.1; the results based on the Real-5 GCMs

are given in Table B5). The percentage of AR-related POT_{N1P}, POT_{N2P} and POT_{N3P} precipitation events changed by -2%~1% in the Chehalis basin, 0%~2% in the Russian basin, and 5%~7% in the Santa Margarita basin. However, the changes in percentage of AR-related discharge events were different from that of extreme precipitation events. The percentage of AR-related POT_{N1D}, POT_{N2D} and POT_{N3D} events changed by -2%~0% in the Chehalis basin, 4%~6% in the Russian basin, and 8%~9% in the Santa Margarita basin. Similar to POT extreme discharge events, the percentage of AR-related AMF events stayed unchanged in the Chehalis basin based on the mean of the full ensemble, but increased by 4% and 8% in the Russian and Santa Margarita basins, respectively. Nonetheless, the POT thresholds generally increased for both extreme precipitation and extreme discharge events in all basins (see Tables B3-B4).

3.4.3.2 Changes in the relationship of extreme precipitation to extreme discharge

Given that there were mismatches in the changes in AR contributions to extreme precipitation and extreme discharge events in the three basins, the relationship between extreme precipitation and extreme discharge might have changed. We examined future changes in the probability of extreme precipitation events leading to extreme discharge events of the same POT threshold by comparing the full GCM ensemble during the periods WY 1951-2000 and WY 2050-2099 (see Figure 3.5). Following Maurer et al. (2018), we used the Wilcoxon signed-rank (non-parametric) test to determine whether there is any significant change in the relevant probabilities among various GCM projections. In the Santa Margarita basin, $Pr(POT_{N1D}|POT_{N1P})$, $Pr(POT_{N2D}|POT_{N2P})$ and $Pr(POT_{N3D}|POT_{N3P})$ are all projected to significantly increase ($p<0.1$) at the basin outlet. In the Russian basin, the probabilities of all three thresholds are projected to increase at

the basin outlet but only the increase of $Pr(POT_{N1D}|POT_{N1P})$ is statistically significant. However, there is no statistically significant change at the Chehalis basin outlet for any threshold, and in fact, the $Pr(POT_{N1D}|POT_{N1P})$ and $Pr(POT_{N2D}|POT_{N2P})$ are projected to slightly decrease.

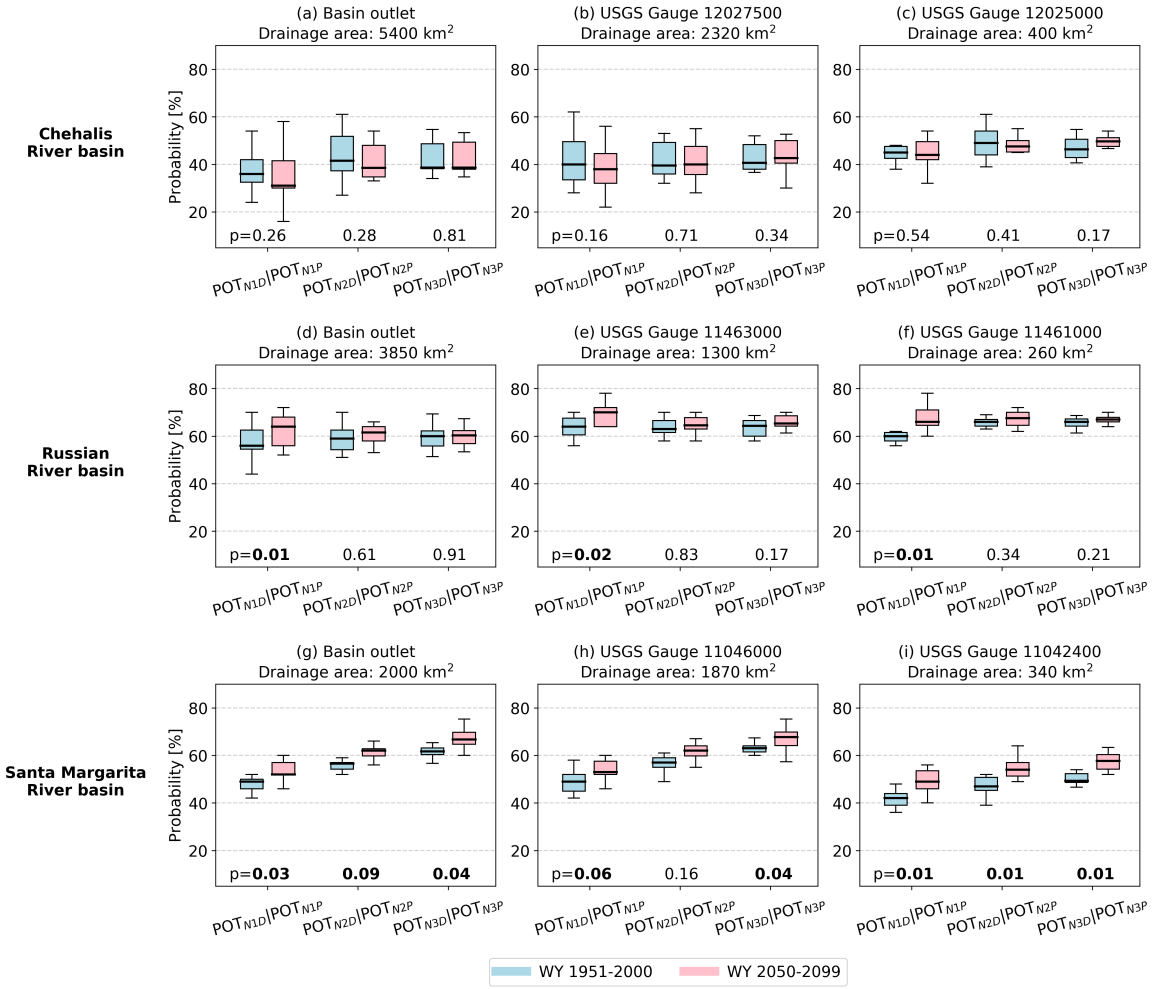

Figure 3.5 Box plots of the probability of extreme precipitation events leading to extreme discharge events of the same POT thresholds based on GCM ensembles during the periods of WY 1951-2000 and WY 2050-2099. The *p* values of the Wilcoxon signed-rank test are shown in plots, with values not greater than 0.1 marked in bold font.

To eliminate the impact of drainage area (denoted as DA) differences in the three basins (as well as the potential impact of the relationship between peak daily precipitation and storm total precipitation), we further examined the probability of changes at the sub-basin level. Specifically, we grouped the sub-basins into small and large categories by

DAs below or above 1000 km^2, a threshold used in previous studies (e.g. Ivancic and Shaw, 2015; Wasko and Sharma, 2017; Wasko and Nathan, 2019). Figures 3.5b, e, and h show results for sub-basins with drainage areas of greater than 1000 km^2, and Figures 3.5c, f, and i for sub-basins with DA less than 1000 km^2). The pattern of changes across the three basins was generally similar to those at the basin outlets, but changes tended to be more significant for the smaller drainage areas.

Based on the analysis of historical events in the Russian River basin (Cao et al., 2019), we hypothesize that there are two ways that the future relationship between extreme precipitation and extreme discharge may be strengthened. One is through increases in storm precipitation; hence we first examined the role of ARs since they are usually associated with the most extreme precipitation events. The other is that large storms and wet ASM conditions are more likely to be concurrent; hence we examined the ASM and timing of extreme precipitation events.

a. Role of ARs

Figure 3.6 shows the probabilities conditioned on ARs (i.e. when extreme precipitation events were associated with ARs). The probabilities were higher in all three basins during each of the two periods (WY 1951-2000 and WY 2050-2099) in comparison with no conditioning. For the projected changes in $Pr(POT_D|POT_P)$, there was still no statistically significant change in the Chehalis basin, but the $Pr(POT_{N1D}|POT_{N1P})$ slightly increased in terms of the median value. In the Russian basin, the increase of $Pr(POT_{N1D}|POT_{N1P})$ was still statistically significant at the basin outlet, and there were more statistically significant changes in sub-basins for POT events with lower thresholds, indicating that the changes in AR-related storm precipitation were

sufficiently large to lead to more extreme discharge of the same POT threshold. In the Santa Margarita basin, the probabilities still increased both at the basin outlet and within sub-basins, but less significant increases were possibly due to the relatively small number of AR-related extreme precipitation events in this basin.

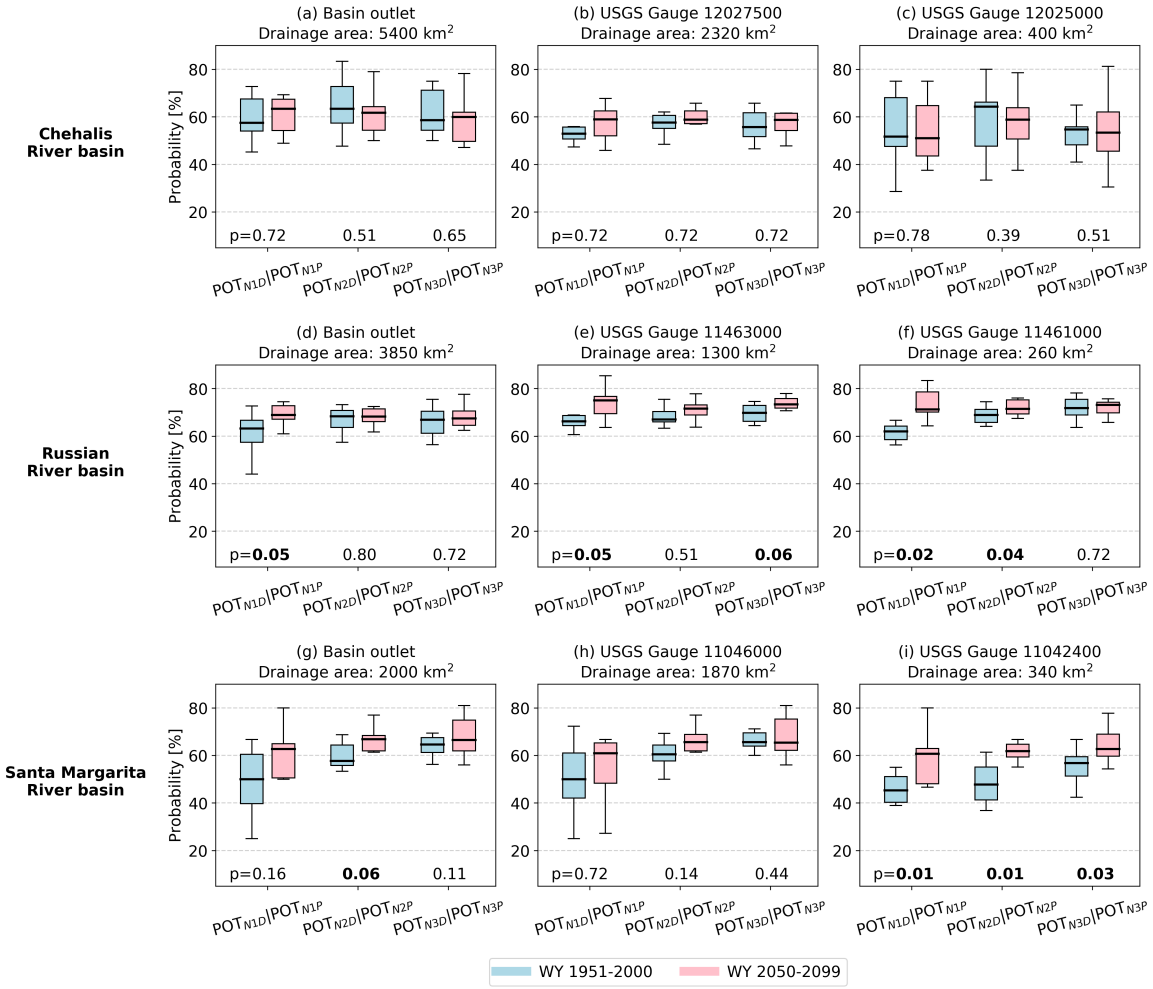

Figure 3.6 Same as Figure 3.5 but for extreme precipitation events conditioned on ARs.

b. Role of ASM

Whether an extreme precipitation event could lead to an extreme discharge event depends not only on storm total precipitation but also on ASM. We first examined the average changes in storm precipitation and ASM (see Figure 3.7). Storm precipitation

generally increased in all three basins across all three POT thresholds, but the average changes in ASM varied among basins. When we sorted the POT_P events by the POT_D threshold of their peak flows, both the storm precipitation and ASM are higher on average (see Figure B1) for the POT_P events that lead to the POT_D of the same threshold.

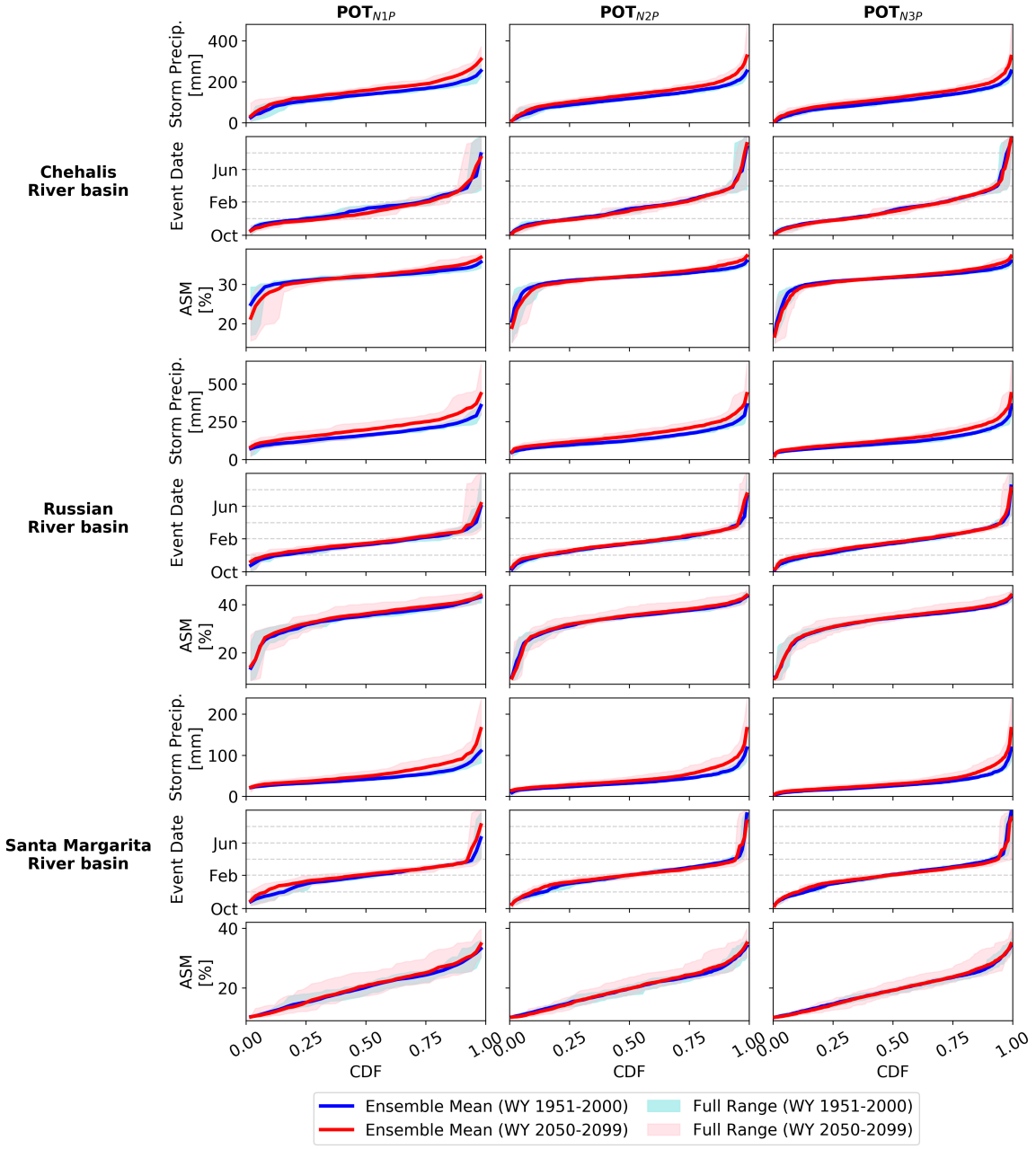

Figure 3.7 CDF of storm total precipitation, storm occurrence dates and antecedent soil moisture (ASM) conditions in three basins during POT_{N1P}, POT_{N2P} and POT_{N3P} events (i.e. extreme precipitation events with thresholds set to 1, 2 and 3 events per year on average).

We further examined the seasonal timing of the POT events (see Figure 3.7 and B2). The mean occurrence date of historical POT_{N3D} events became later progressing from north to south along the coast: around mid January in the Chehalis basin, late January in the Russian basin, and early February in the Santa Margarita basin. Extreme discharge events historically occurred from mid-November to mid-March in the Chehalis basin when the soil was relatively wet and when it was close to the peak of the AR season in the Pacific Northwest region (Gershunov et al. 2017). Sorting the POT_P events by their occurrence dates shows that low ASM occurs more often during the shoulder seasons (see Figure B2), which is projected to become even lower, possibly due to precipitation frequency loss (see Figure B3 for the seasonal cycle of ASM). Large storms are projected to occur more often in the late fall in the Chehalis River basin (see Figure B2), which will be increasingly strongly affected by carry-over (dry) ASM from the previous summer, and are increasingly unlikely to lead to extreme runoff events. In contrast, large storms are projected to occur more often in the winter in the Russian and Santa Margarita River basins due to the delayed onset of winter precipitation in California (Pierce et al. 2013b).

We examined the *Rs* between storm precipitation and ASM during POT extreme precipitation events (see Figure 3.8). There is no clear pattern of changes in *Rs* in the Chehalis River basin. The *Rs* is generally projected to increase, however, in the Russian and Santa Margarita River basins for POT_{N2P} and POT_{N3P} events. Nearly half of the *Rs* values are projected to be significant ($p<0.1$) among 10 GCMs, indicating that a large storm is projected to be more likely to follow a wet ASM condition as the year-to-year volatility of annual precipitation is projected to increase in California.

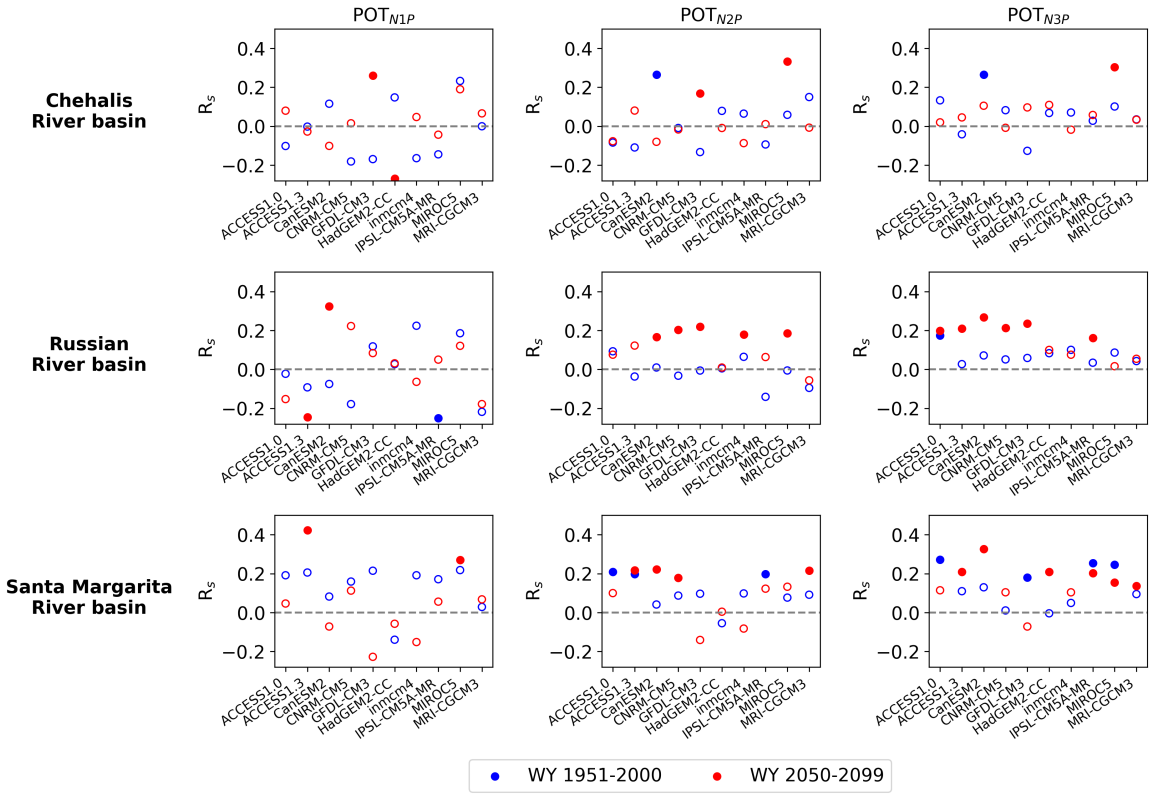

Figure 3.8 The Spearman's rank correlation (*Rs*) between storm precipitation and antecedent soil moisture (ASM) during POT_{N1P}, POT_{N2P} and POT_{N3P} events (i.e. extreme precipitation events with thresholds set to 1, 2 and 3 events per year on average). The correlations with p values not greater than 0.1 are shown as solid symbols.

3.4.3.3 Changes in AMF flows

We next examine changes in the magnitude of AMF events, as well as storm precipitation and related ASM based on the ensemble mean of all 10 GCMs (see Figure 3.9). The pattern of changes in these variables was similar to POT_{N1D} events (see Figure B4), which hence are not shown here. The magnitude of AMF and its storm precipitation showed an overall increase in all three basins. The AMF (averaged over 50 years) increased by 21%, 29% and 48% respectively in the Chehalis, Russian and Santa Margarita basins. Storm precipitation increased by 17%, 26% and 38%. However, the ASM preceding future AMF events is not projected to change much -- 2%, 1% and -2% on average in the three basins. Small increase in ASM occurred over all quartiles in the

Chehalis basin, for the upper three quartiles in the Russian basin, and only over the upper quartile in the Santa Margarita basin. When we sort the variables by the CDF of AMF events, we can see that the influence of changes in ASM on AMF varies among individual events (see Figure B5).

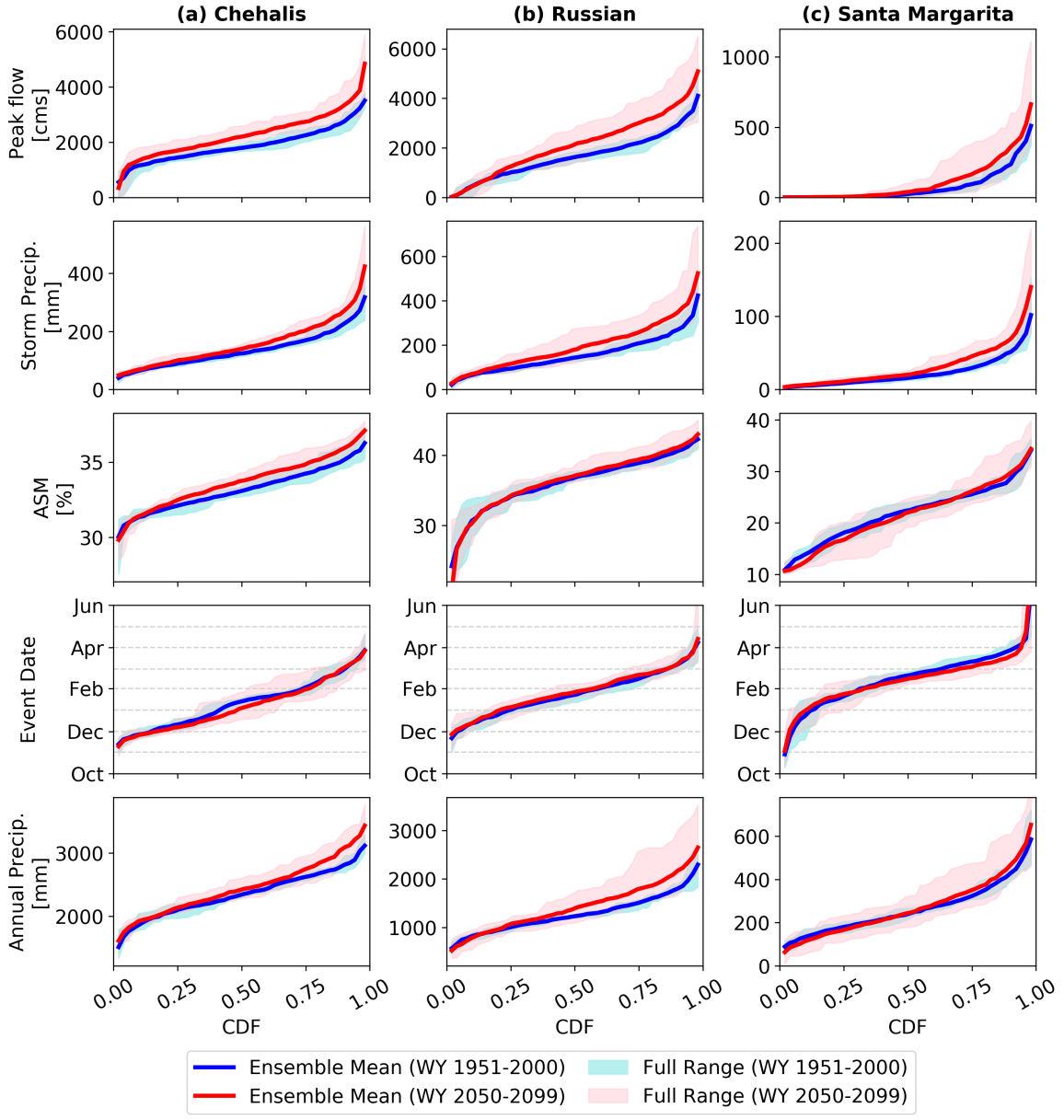

Figure 3.9 CDF of peak daily flow, storm total precipitation, antecedent soil moisture (ASM) conditions, and event occurrence dates during annual maximum flow (AMF) events, as well as annual precipitation in a) Chehalis River basin, b) Russian River basin, and c) Santa Margarita River basin based on the ensemble of 10 GCMs.

We further examined the correlation between storm precipitation and AMF given ASM (denoted as "$r_{SP\&AMF\text{-}ASM}$"), and correlation between ASM and AMF given storm precipitation (denoted as "$r_{ASM\&AMF\text{-}SP}$") during WY 1951-2000 and WY 2050-2099 (see Figure B6). The $r_{ASM\&AMF\text{-}SP}$ is projected to slightly decrease ($p<0.1$ in the Wilcoxon test) among GCMs, yet still significant, in the Russian River basin where storm precipitation is projected to increase most among three basins caused by ARs. No statistically significant changes were found in the other two basins.

3.5 Discussion

We examined changes in the magnitude of extreme floods with return periods of 4, 10, 20, 50 and 100 years by fitting the Generalized Extreme Value (GEV) distribution to the AMF from each GCM (see Figure B7). Mallakpour et al. (2019) evaluated different distributions for the AMF in multiple basins in California and found the GEV performed best in most cases. We used the Wilcoxon test to determine whether there is any significant change in the percent change of flood magnitude between WY 2050-2099 and WY 1951-2000 (see Figure 3.10). The full GCM ensemble projects significant increases ($p<0.05$) in the magnitude of floods with all return periods in all three basins. The Real-5 GCMs (arguably the most realistic in terms of capturing historical AR event statistics; see SM1 in the supplemental material) project significant increases ($p<0.05$) in the 4, 10 and 20-year floods in the Chehalis River basin, the 4-year floods in the Russian River basin (despite of much higher median values for all extreme floods in comparison with the full GCM ensemble), and the 100-year floods in the Santa Margarita River basin, where the fraction of AR-related AMF events increases most.

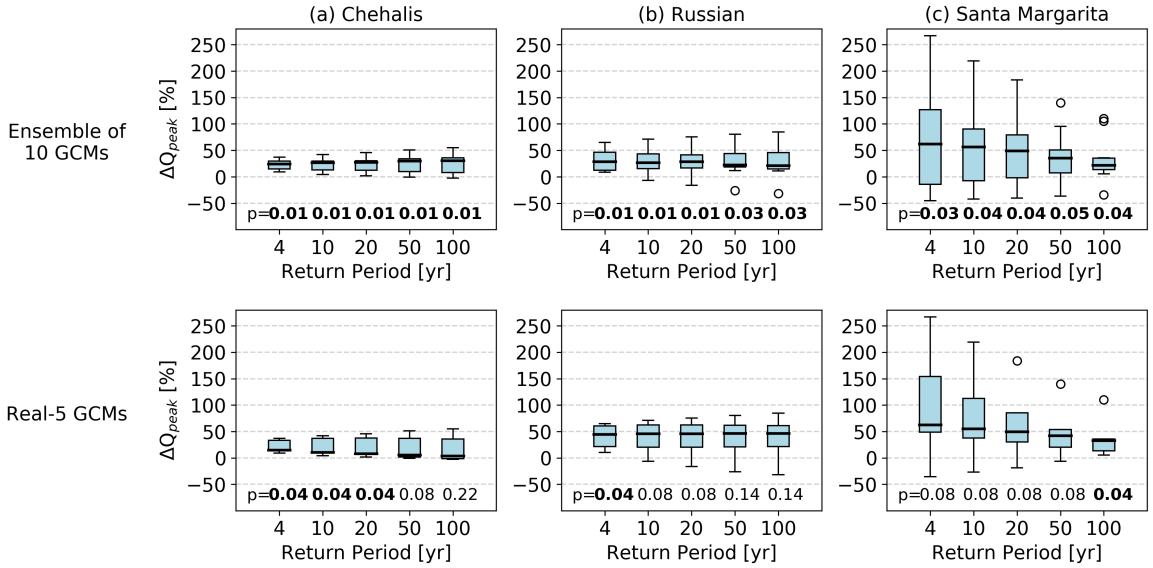

Figure 3.10 Percent change in 4, 10, 20, 50 and 100-year recurrence interval flow between WY 2050-2099 and WY 1951-2000. The upper row is based on the ensemble of 10 GCMs and the lower row is based on the Real-5 GCMs. The p values not greater than 0.05 are marked in bold font.

As mentioned above, the ASM associated with AMF events generally (slightly) increased in the Chehalis and Russian basins, while they slightly decreased in the Santa Margarita basin. We examined what caused the changes in ASM in the three basins, specifically changes in temperature vs. changes in antecedent precipitation. Figure 3.11 shows the correlation between the changes in soil moisture (ΔSM) and antecedent precipitation (P) given evapotranspiration (ET) (denoted as "$r_{\Delta SM\&P\text{-}ET}$"), and the correlation between ΔSM and ET given P (denoted as "$r_{\Delta SM\&ET\text{-}P}$") for different pre-event durations (from 2 days to 12 weeks). The figure shows that the ASM are affected more by P than ET in each of the basin during both periods, except for the Chehalis basin during WY 1951-2000 due to the influence of seasonality of events. When they are examined in separate seasons (i.e. late fall and winter), the relative influence of P is larger than ET. For the Santa Margarita basin in Southern California where the projected loss of (non-AR) heavy precipitation is most pronounced (Polade et al. 2014, 2017; Gershunov

et al. 2019), the correlation between ΔSM and P given ET is much higher than in the other two basins, indicating that precipitation regime change is a relatively more important driver of projected ASM changes in this basin. In the meantime, the relative influence of ET on ΔSM is smallest in the Chehalis basin, intermediate in the Russian basin, and largest in the Santa Margarita basin. Warming shows a greater impact on ASM of AMF events in the Santa Margarita basin, where the change in precipitation regime may exert a synergetic effect that makes the ASM more vulnerable to warming. This may be the case for other Mediterranean climate regimes as well, where precipitation frequency loss is even more pronounced (Polade et al. 2017).

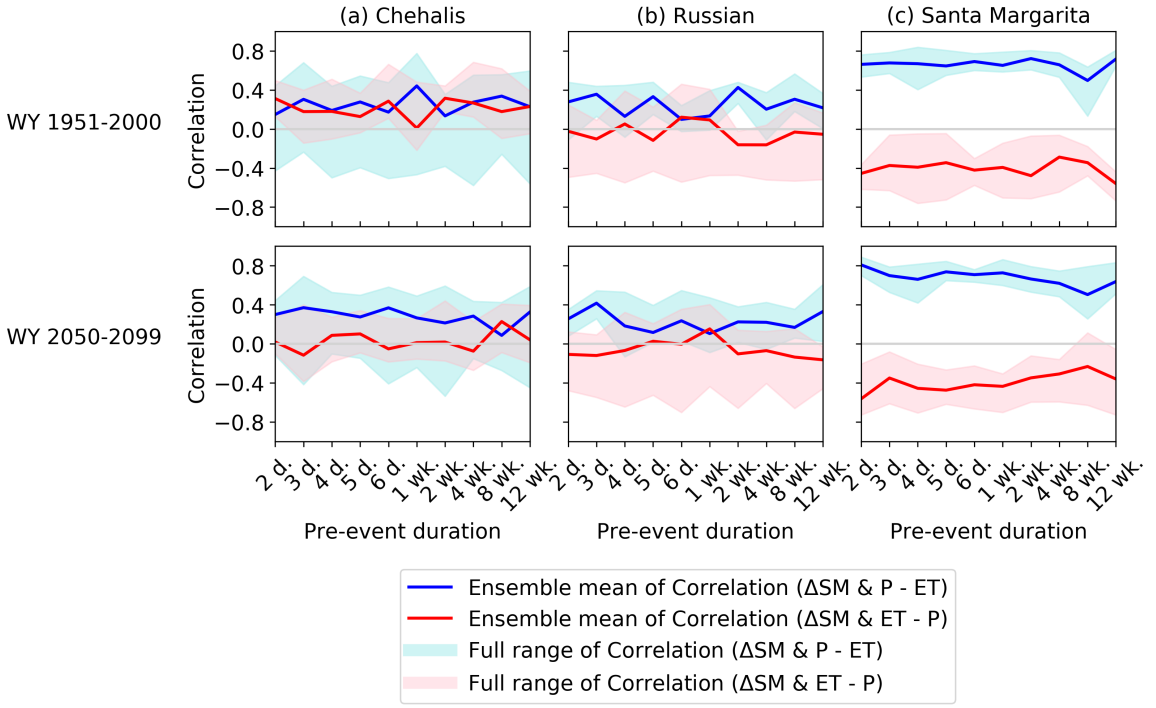

Figure 3.11 Effects of temperature and antecedent precipitation conditions on changes of antecedent soil moisture (ΔSM) of annual maximum flow events in a) the Chehalis River basin, b) the Russian River basin, and c) the Santa Margarita River basin, which are examined as correlation between ΔSM and accumulated precipitation (P) given accumulated evapotranspiration (ET), and correlation between ΔSM and ET given P under different pre-event duration. The upper row is for the period of WY 1951-2000 and the bottom row is for the period of WY 2050-2099.

We note here some limitations of our study. One is that there are uncertainties in the AR-related percentages of extreme events we reported here, partly related to the low spatial resolution of both the NCEP–NCAR reanalysis dataset and the GCMs that the AR date catalogs are based on. However, Ralph et al (2018) showed that the derived landfalling AR catalogs were not sensitive to reanalysis spatial resolution (this was specifically for the Russian River basin) but were much more sensitive to the detection methodology used. Also, the SIO-R1 catalog we used here was highlighted by Ralph et al (2018) as relevant for precipitation-related studies as it was specifically developed with precipitation applications in mind and evaluated with independent precipitation data (Gershunov et al. 2017). In terms of the uncertainties in percentages based on GCMs, rather than their spatial resolution, a greater uncertainty appears to be GCM biased in IVT — mostly wet, but some unrealistically dry, as is discussed in Gershunov et al. (2019). Hence we focused more on changes in the percentages than their absolute values. Other limitations are related to meteorological forcings. For example, Pierce et al. (2013a) showed that the algorithm we used here to estimate humidity might not preserve the original global model's humidity trends, but they also pointed out that coastal areas were less biased than the interior areas. Additionally, some studies have shown the impact of warming on the intensification of precipitation at sub-daily time scales (e.g. Westra et al., 2013, 2014; Meredith et al., 2019), indicating that atmospheric states may need to be considered in the precipitation disaggregation for the future period, which will be included in our future work.

3.6 Conclusions

Recent studies have projected large increases in AR activity along the U.S. West Coast, which are projected to be associated with increases in extreme precipitation events during the 21st century (e.g. Hagos et al., 2015; Warner et al., 2015; Gershunov et al., 2019). However, ASM is a primary link between precipitation and flooding, and might decrease with projected decreases in pre-storm low-medium intensity precipitation frequency, and increased antecedent evaporative demand associated with warming. This would mitigate the flood response to increases in extreme precipitation caused by ARs. Here, we first examined changes in AR contributions to the largest 50, 100 and 150 extreme discharge events in three river basins along the U.S. West Coast in two periods, WY 1951-2000 and WY 2050-2099 for three rainfall-dominated watersheds (only the Chehalis has modest contributions of snowmelt to flood runoff). We then examined how ASM is likely to change in the future and assessed its effect on changes in flood response, particularly on the relationship of extreme precipitation to extreme discharge, and on the magnitude of AMF. Based on our analysis, we find:

1) Historically most extreme discharge events in all three river basins have been AR-related (specific fractions depend on the POT threshold used and the specific river basin). In a warmer climate the projected fraction of AR-related extreme discharge events will decrease slightly in the northernmost Chehalis River basin, primarily due to slight decreases in the fraction of AR-related extreme precipitation events. However, in the Russian River basin and even more in the southernmost Santa Margarita River basin the fraction of AR-related extreme discharge events will increases, partly due to increases in the fraction of AR-related extreme precipitation events, and partly due to changes in the

relationship between extreme precipitation and extreme discharge (discussed below). The changes in fraction slightly increase as the threshold decreases.

2) Changes in the fraction of AR-related extreme discharge events differ from changes in the (fraction of) AR-related extreme precipitation events, due to changes in the relationship of extreme precipitation to extreme discharge. The relationship is projected to become slightly weaker in the Chehalis River basin, possibly due to a) modest increases in storm precipitation, b) increases in the number of extreme precipitation events that occur in the shoulder seasons when ASM is relatively dry (and is projected to become drier as a result of precipitation frequency loss), and c) little change in precipitation volatility relative to the other two basins. The relationship is projected to become significantly stronger in the Russian River basin for POT events of the highest threshold, and likewise in the Santa Margarita River basin for POT events of all thresholds, partly due to large storms being more likely to follow wet ASM conditions as the year-to-year volatility of annual precipitation is projected to increase in California. When extreme precipitation events are conditioned on ARs, the relationship is projected to become stronger in all three basins, especially in the Russian River basin.

3) The ASM associated with AMF events is projected to slightly increase (on average) in the Chehalis and Russian River basins, while the ASM is projected to slightly decrease (on average) in the Santa Margarita River basin, where the loss of non-AR precipitation frequency is most pronounced and hence it may exert a synergetic effect that makes the ASM more sensitive to warming. The influence of changes in ASM on AMF varies among individual events in three basins, but the relative effect of ASM on AMF given storm precipitation is projected to become slightly weaker (yet still

significant) in the Russian River basin where storm precipitation is projected to increase most among three basins caused by ARs.

4) In terms of the general changes in extreme floods, the full GCM ensemble projects significant increases in the magnitude of 100-year floods in all three basins. The "Real-5" (five most realistic GCMs in terms of their reproduction of historical AR event statistics) project relatively small increases in the magnitude of the 100-year flood in the Chehalis River basin but greater increases in the Russian and Santa Margarita River basins than the full GCM ensemble. Their projections show that, however, the increase in the magnitude of 100-year flood is only significant in the Santa Margarita River basin, where the number of AR-related AMF events increases most and the year-to-year volatility of annual precipitation is projected to increase most. Differences in the results indicate that caution is needed in the selection of GCMs for future flood analyses in these basins.

Acknowledgement

The research supported herein was funded by the Center for Western Weather and Water Extremes (CW3E) at the Scripps Institution of Oceanography via AR Program Phase II, sponsored by the California Department of Water Resources.

References

Barth, N. A., G. Villarini, M. A. Nayak, and K. White, 2017: Mixed populations and annual flood frequency estimates in the western United States: The role of atmospheric rivers. *Water Resour. Res.*, **53**, 257-269. https://doi.org/10.1002/2016wr019064.

Begueria, S., M. Angulo-Martinez, S. M. Vicente-Serrano, J. I. Lopez-Moreno, and A. El-Kenawy, 2011: Assessing trends in extreme precipitation events intensity and magnitude using non-stationary peaks-over-threshold analysis: a case study in northeast Spain from 1930 to 2006. *Int. J. Climatol.*, **31**, 2102-2114, https://doi.org/10.1002/joc.2218.

Bennett, B., M. Leonard, Y. Deng, and S. Westra, 2018: An empirical investigation into the effect of antecedent precipitation on flood volume. *J. Hydrol.*, **567**, 435–445. https:// doi.org/10.1016/j.jhydrol.2018.10.025.

Berghuijs, W. R., R. A. Woods, C. J. Hutton, and M. Sivapalan, 2016: Dominant flood generating mechanisms across the United States. *Geophys. Res. Lett.*, **43**, 4382-4390, https://doi.org/10.1002/2016gl068070.

Blume, T., E. Zehe, and A. Bronstert, 2007: Rainfall-runoff response, event-based runoff coefficients and hydrograph separation. *Hydrol. Sci. J.*, **52**, 843-862, https://doi.org/10.1623/hysj.52.5.843.

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

Cao, Q., N. Sun, J. Yearsley, B. Nijssen, and D. P. Lettenmaier, 2016: Climate and land cover effects on the temperature of Puget Sound streams. *Hydrol. Processes*, **30**, 2286-2304, https://doi.org/10.1002/hyp.10784.

——, T. H. Painter, W. R. Currier, J. D. Lundquist, and D. P. Lettenmaier, 2018: Estimation of Precipitation over the OLYMPEX Domain during Winter 2015/16. *J. Hydrometeor*., **19**, 143-160, https://doi.org/10.1175/jhm-d-17-0076.1.

——, A. Mehran, F.M. Ralph, and D.P. Lettenmaier, 2019: The Role of Hydrological Initial Conditions on Atmospheric River Floods in the Russian River Basin. *J. Hydrometeor.*, **20**, 1667–1686, https://doi.org/10.1175/JHM-D-19-0030.1.

Corringham, T.W., F.M. Ralph, A. Gershunov, D.R. Cayan, and C.A. Talbot, 2019: Atmospheric Rivers Drive Flood Damages in the Western United States. *Sci. Adv*., **5**, 2375-2548, https://doi.org/10.1126/sciadv.aax4631

Crow, W. T., F. Chen, R. H. Reichle, and Q. Liu, 2017: L band microwave remote sensing and land data assimilation improve the representation of prestorm soil moisture conditions for hydrologic forecasting. *Geophys. Res. Lett*., **44**, 5495-5503, https://doi.org/10.1002/2017gl073642.

Dettinger, M., 2011: Climate Change, Atmospheric Rivers, and Floods in California – A Multimodel Analysis of Storm Frequency and Magnitude Changes. *J. Am. Water Resour. Assoc.*, **47**, 514-523, https://doi.org/10.1111/j.1752-1688.2011.00546.x.

——, F. M. Ralph, T. Das, P. Neiman, and D. Cayan, 2011: Atmospheric rivers, floods and the water resources of California. *Water*, **3**, 445–478, https://doi.org/10.3390/w3020445.

Espinoza, V., D. E. Waliser, B. Guan, D. A. Lavers, and F. M. Ralph, 2018: Global analysis of climate change projection effects on atmospheric rivers. *Geophys. Res. Lett*., **45**, 4299–4308. https://doi.org/10.1029/ 2017GL076968.

Gao, Y., J. Lu, L. R. Leung, Q. Yang, S. Hagos, and Y. Qian, 2015: Dynamical and thermodynamical modulations on future changes of landfalling atmospheric rivers over western North America, *Geophys. Res. Lett*., **42**, 7179-7186, https://doi.org/10.1002/2015GL065435.

Gershunov, A., T. Shulgina, F. M. Ralph, D. A. Lavers, and J. J. Rutz, 2017: Assessing the climate-scale variability of atmospheric rivers affecting western North America. *Geophys. Res. Lett*., **44**, 7900-7908, https://doi.org/10.1002/2017gl074175.

——, and Coauthors, 2019: Precipitation regime change in Western North America: The role of Atmospheric Rivers. *Sci. Rep*., **9**, 9944, https://doi.org/10.1038/s41598-019-46169-w.

Gupta, H. V., H. Kling, K. K. Yilmaz, and G. F. Martinez, 2009: Decomposition of the mean squared error and NSE performance criteria: Implications for improving hydrological modeling. *J. Hydrol.*, **377**, 80–91, https://doi.org/10.1016/j.jhydrol.2009.08.003.

Hagos, S. M., L. R. Leung, J.-H. Yoon, J. Lu, and Y. Gao, 2016: A projection of changes in landfalling atmospheric river frequency and extreme precipitation over western North America from the Large Ensemble CESM simulations, *Geophys. Res. Lett*., **43**, 1357– 1363, https://doi.org/10.1002/2015GL067392.

Hamlet, A. F., and D. P. Lettenmaier, 2007: Effects of 20th century warming and climate variability on flood risk in the western U.S. *Water Resour. Res.*, **43**, W06427, https://doi.org/10.1029/ 2006WR005099.

Ivancic, T.J., and S.B. Shaw, 2015: Examining why trends in very heavy precipitation should not be mistaken for trends in very high river discharge. *Clim. Change*, **133**. 681-693, https://doi.org/10.1007/s10584-015-1476-1.

Kalnay, E., and Coauthors, 1996: The NCEP/NCAR 40-year reanalysis project. *Bull. Amer. Meteor. Soc.*, **77**, 437-471, https://doi.org/10.1175/1520-0477(1996)077<0437:tnyrp>2.0.co;2.

Konrad, C. P., and M.D. Dettinger, 2017: Flood runoff in relation to water vapor transport by atmospheric rivers over the western United States, 1949–2015. *Geophys. Res. Lett.*, **44**, 11,456–11,462, https://doi.org/10.1002/2017GL075399.

Lang, M., T. Ouarda, and B. Bobee, 1999: Towards operational guidelines for over-threshold modeling. *J. Hydrol.*, **225**, 103-117, https://doi.org/10.1016/s0022-1694(99)00167-5.

Liang, X., D. P. Lettenmaier, E. Wood, and S. Burges, 1994: A simple hydrologically based model of land-surface water and energy fluxes for general-circulation models. *J. Geophys. Res.*, **99**, 14 415–14 428, https://doi.org/10.1029/94JD00483.

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

Mallakpour, I., and G. Villarini, 2017: Analysis of changes in the magnitude, frequency, and seasonality of heavy precipitation over the contiguous USA. *Theor. Appl. Climatol.*, **130**, 345-363, https://doi.org/10.1007/s00704-016-1881-z.

——, A. AghaKouchak, and M. Sadegh, 2019: Climate-induced changes in the risk of hydrological failure of major dams in California. *Geophys. Res. Lett.*, **46**, 2130–2139. https://doi.org/10.1029/2018GL081888.

——, M. Sadegh, and A. AghaKouchak, 2018: A new normal for streamflow in California in a warming climate: Wetter wet seasons and drier dry seasons. *J. Hydrol.*, **567**, 203– 211, https://doi.org/10.1016/j.jhydrol.2018.10.023.

Maurer, E.P., G. Kayser, L. Doyle, and A.W. Wood, 2018: Adjusting flood peak frequency changes to account for climate change impacts in the Western United States. *J. Water Resour. Plan. Manag.* **144**, 05017025. https://doi.org/10.1061/(ASCE)WR.1943- 5452.0000903.

Meredith, E. P., U. Ulbrich, and H. W. Rust, 2019: The diurnal nature of future extreme precipitation intensification. *Geophys. Res. Lett.*, **46**, 7680-7689, https://doi.org/10.1029/2019GL082385.

Naz, B. S., S.-C. Kao, M. Ashfaq, D. Rastogi, R. Mei, and L. C. Bowling, 2016: Regional hydrologic response to climate change in the conterminous United States using high-resolution hydroclimate simulations. *Global Planet. Change*, **143**, 100– 117, https://doi.org/10.1016/j.gloplacha.2016.06.003.

——, ——, ——, H. Gao, D. Rastogi, and S. Gangrade, 2018: Effects of climate change on streamflow extremes and implications for reservoir inflow in the United States. *J. Hydrol.*, **556**, 359-370, https://doi.org/10.1016/j.jhydrol.2017.11.027.

Neiman, P. J., L. J. Schick, F. M. Ralph, M. Hughes, and G. A. Wick, 2011: Flooding in Western Washington: The Connection to Atmospheric Rivers. *J. Hydrometeor*., **12**, 1337-1358, https://doi.org/10.1175/2011jhm1358.1.

Pagán, B. R., M. Ashfaq, D. Rastogi, D. R. Kendall, S. C. Kao, B. S. Naz, R. Mei, and J. S. Pal, 2016: Extreme hydrological changes in thesouthwestern US drive reductions in water supply to Southern California by mid century, *Environ. Res. Lett*., **11**, 094026, https://doi.org/10.1088/1748-9326/11/9/094026.

Payne, A. E., and G. Magnusdottir, 2015: An evaluation of atmospheric rivers over the North Pacific in CMIP5 and their response to warming under RCP 8.5. *J. Geophys. Res. Atmos.*, **120**, 11,173– 11,190, https://doi.org/10.1002/2015JD023586.

Pierce, D. W., A. L. Westerling, and J. Oyler, 2013a: Future humidity trends over the western United States in the CMIP5 global climate models and variable infiltration capacity hydrological modeling system. *Hydrol. Earth Syst. Sci*., **17**, 1833–1850, https://doi.org/10.5194/hess-17-1833-2013.

——, and Coauthors, 2013b: Probabilistic estimates of future changes in California temperature and precipitation using statistical and dynamical downscaling. *Climate Dyn.*, **40**, 839–856, doi:10.1007/s00382-012-1337-9.

——, D.R. Cayan, and B.L. Thrasher, 2014: Statistical Downscaling Using Localized Constructed Analogs (LOCA). *J. Hydrometeor*., **15**, 2558–2585, https://doi.org/10.1175/JHM-D-14-0082.1.

Polade, S., D. W. Pierce, D. R. Cayan, A. Gershunov, M. D. Dettinger, 2014: The key role of dry days in changing regional climate and precipitation regimes. *Sci. Rep*., **4**, 4364, https://doi.org/10.1038/srep04364.

——, A. Gershunov, D. R. Cayan, M. D. Dettinger, and D. W. Pierce, 2017: Precipitation in a warming world: Assessing projected hydro-climate changes in California and other Mediterranean climate regions. *Sci. Rep*., **7**, 10 783, https://doi.org/10.1038/s41598-017-11285-y.

Ralph, F.M., P.J. Neiman, G. Wick, S. Gutman, M. Dettinger, D. Cayan, and A. White, 2006: Flooding on California's Russian River: Role of atmospheric rivers. *Geophys. Res. Lett*., **33**, L13801, https://doi.org/10.1029/2006GL026689.

——, and Coauthors, 2018: ARTMIP-early start comparison of atmospheric river detection tools: how many atmospheric rivers hit northern California's Russian River watershed? *Clim. Dyn*., 1-22, https://doi.org/10.1007/s00382-018-4427-5.

——, J.J. Rutz, J.M. Cordeira, M. Dettinger, M. Anderson, D. Reynolds, L.J. Schick, and C. Smallcomb, 2019: A Scale to Characterize the Strength and Impacts of Atmospheric Rivers. *Bull. Amer. Meteor. Soc.*, **100**, 269–289, https://doi.org/10.1175/BAMS-D-18-0023.1.

Rogelj J., M. Meinshausen, and R. Knutti, 2012: Global warming under old and new scenarios using IPCC climate sensitivity range estimates. *Nat Clim Chang*, 2, 248-253, https://doi.org/10.1038/nclimate1385.

Rupp, D. E., J. T. Abatzoglou, K. C. Hegewisch, and P. W. Mote, 2013: Evaluation of CMIP5 20th century climate simulations for the Pacific Northwest USA. *J. Geophys. Res. Atmos*., **118**, 10 884–10 906, https://doi.org/10.1002/jgrd.50843.

Salathé, E.P., A.F. Hamlet, C.F. Mass, S. Lee, M. Stumbaugh, and R. Steed, 2014: Estimates of Twenty-First-Century Flood Risk in the Pacific Northwest Based on

Regional Climate Model Simulations. *J. Hydrometeor.*, **15**, 1881–1899, https://doi.org/10.1175/JHM-D-13-0137.1.

Sharma, A., C. Wasko, and D. P. Lettenmaier, 2018: If precipitation extremes are increasing, why aren't floods? *Water Resour. Res.*, **54**. https://doi.org/10.1029/2018WR023749.

Shields, C. A., and J. T. Kiehl, 2016: Atmospheric river landfall-latitude changes in future climate simulations. *Geophys. Res. Lett.*, **43**, 8775–8782, https://doi.org/10.1002/2016GL070470.

Tohver, I. M., A. F. Hamlet, and S.-Y. Lee, 2014: Impacts of 21st-Century Climate Change on Hydrologic Extremes in the Pacific Northwest Region of North America. *J. Amer. Water Resour. Assoc.*, **50**, 1461-1476, https://doi.org/10.1111/jawr.12199.

USWRC, 1982. Guidelines for Determining Flood Flow Frequency. Bulletin 17B, United States Water Resources Committee, Washington, DC, USA.

Warner, M.D., C.F. Mass, and E.P. Salathé, 2015: Changes in Winter Atmospheric Rivers along the North American West Coast in CMIP5 Climate Models. *J. Hydrometeor.*, **16**, 118–128, https://doi.org/10.1175/JHM-D-14-0080.1.

Wasko, C., and R. Nathan, 2019: Influence of changes in rainfall and soil moisture on trends in flooding. *J. Hydrol.*, **575**, 432-441, https://doi.org/10.1016/j.jhydrol.2019.05.054.

——, and A. Sharma, 2017: Global assessment of flood and storm extremes with increased temperatures. *Sci. Rep.*, **7**, 8, https://doi.org/10.1038/s41598-017-08481-1.

Wigmosta, M. S., L. W. Vail, and D. P. Lettenmaier, 1994: A distributed hydrology-vegetation model for complex terrain. *Water Resour. Res.*, **30**, 1665-1679, https://doi.org/10.1029/94WR00436.

William, A.P., Z. Duan, N. Sun, M.S. Wigmosta, M.C. Richmond, X. Chen, L.R. Leung, 2019: Parallel Distributed Hydrology Soil Vegetation Model (DHSVM) using global arrays. *Environmental Modelling & Software*, **122**, 104533, https://doi.org/10.1016/j.envsoft.2019.104533.

Westra, S., R. Mehrotra, A. Sharma, and R. Srikanthan, 2012: Continuous Rainfall Simulation: 1—A regionalised sub-daily disaggregation approach, *Water Resour. Res.*, **48**, W01535, https://doi.org/10.1029/2011WR010489.

——, J. P. Evans, R. Mehrotra, and A. Sharma, 2013: A conditional disaggregation algorithm for generating fine time-scale rainfall data in a warmer climate, *J. Hydrol.*, **479**, 86–99, https://doi.org/10.1016/j.jhydrol.2012.11.033.

——, and Coauthors, 2014: Future changes to the intensity and frequency of short-duration extreme rainfall. *Rev. Geophys.*, **52**, 522–555, https://doi.org/10.1002/2014RG000464.

Zhao, G., H. L. Gao, B. S. Naz, S. C. Kao, and N. Voisin, 2016: Integrating a reservoir regulation scheme into a spatially distributed hydrological model. *Adv. Water Resour.*, **98**, 16-31, https://doi.org/10.1016/j.advwatres.2016.10.014.

Chapter 4. Evaluation of the subseasonal forecast skill of atmospheric river floods in coastal Western U.S. watersheds

This chapter will be submitted to the *Journal of Hydrometeorology* as

Cao, Q., Shraddhanand Shukla, Michael J. DeFlorio, F. Martin Ralph, and Dennis P. Lettenmaier, 2020: Evaluation of the subseasonal forecast skill of atmospheric river floods in coastal Western U.S. watersheds. *Journal of Hydrometeorology,* (*in prep*).

The supplementary materials for this chapter are provided in Appendix C.

Abstract

Atmospheric rivers (ARs) are responsible for most of the storm events leading to extreme precipitation and runoff along the U.S. West Coast. The time scale of subseasonal forecasting (days to about a month) is critical to proactive disaster mitigation efforts, including responses to AR-related flooding. The NOAA/Climate Testbed Subseasonal Experiment (SubX) project has produced a reforecast data base that facilities evaluation of the potential to forecast AR-related floods. Here, we examine the SubX precipitation and temperature forecast skill, as well as resulting flood forecast skill, with particular attention to the role of hydrologic factors, particularly antecedent soil moisture (ASM) that modulate the relationship between meteorological and hydrological forecast skill. We focus in particular on SubX forecasts of AR-related flooding along the coastal Western U.S. with lead times of 1-4 weeks using the Distributed Hydrology-Soil-Vegetation Model (DHSVM). We study three watersheds that form a transect along the U.S. Pacific Coast: the Chehalis River Basin in Washington, the Russian River Basin in Northern California, and the Santa Margarita River Basin in Southern California. We find

that the SubX forecast skill of extreme discharge events drops quickly after week 1, during which there is relatively strong deterministic forecast skill. We find some probabilitistic forecast skill in week 2, but negligible skill in weeks 3-4, especially for annual maximum floods (one event per year on average), notwithstanding some probabilistic skill for smaller floods in weeks 3-4. We also find that forecast skill is strongly influenced by ASM, with higher forecast skill when ASM is wet especially at shorter lead times and for lower flood thresholds. Using ensemble streamflow prediction (ESP) and reverse-ESP experiments, we find that at short forecast lead times, ASM dominates streamflow forecast skill, while SubX forecast skill dominates at longer lead times. The length of lead times for which ASM has a stronger influence on flood forecast skill is longer in the northern-most (Chehalis) basin, and generally decreases north to south.

4.1 Introduction

Atmospheric rivers (ARs) are responsible for most of the storm events leading to extreme precipitation and runoff along the coastal Western U.S. (e.g. Ralph et al., 2006; Dettinger et al., 2011; Neiman et al., 2011; Barth et al, 2017; Konrad and Dettinger, 2017). Recently, the subseasonal prediction skill of ARs has been evaluated in both forecasts (DeFlorio et al., 2019) and reforecasts (e.g. DeFlorio et al., 2018; Mundhenk et al., 2018; Nardi et al., 2018) using newly emerged subseasonal weather forecast data sets. To date, however, the usefulness of these forecasts for AR-related flooding has not been examined.

Although the time scale of subseasonal weather forecasts is critical to proactive disaster mitigation efforts, such as reservoir operations for flood control, it has not received much attention until recently (Vitart et al., 2017). The past several years have witnessed a joint effort from the weather and climate communities to bridge the weather-climate prediction gap at the subseasonal to seasonal (S2S) range (Mariotti et al., 2018), which typically is defined by lead times ranging from about one to four weeks. A few subseasonal forecast data bases have been developed, such as the World Weather Research Programme (WWRP)/World Climate Research Program (WCRP) Subseasonal to Seasonal (S2S) Prediction Project (Vitart et al., 2017) and the NOAA/Climate Testbed Subseasonal Experiment (SubX) project (Pegion et al., 2019).

Several recent studies have evaluated hindcasts of precipitation from the WWRP/WCRP S2S database. For example, Lin et al. (2018) examined the forecast skill of 11 S2S models for extreme precipitation at a lead time of about 2 weeks for a 2017 flood event in eastern Canada. They found that most of the models predicted above-normal precipitation during the flood event but most underestimated precipitation amounts in comparison with observations, possibly due to the underestimation of the amplitudes of the Madden-Julian Oscillation (MJO) teleconnections in boreal winter. Pan et al. (2019) evaluated S2S precipitation prediction skill over the western U.S. WWRP/WCRP S2S data base. They found that the best-performing models had useful deterministic skill at week 2 but beyond that only their probabilistic skill was useful. They also found that the model performances varied by the phases of El Niño–Southern Oscillation (ENSO) and MJO, which are two dominant tropical variations affecting precipitation, especially for the West Coast of North America.

In addition, a few studies have examined the potential application of the WWRP/WCRP S2S precipitation forecasts for streamflow forecast. For example, Schick et al. (2019) examined the forecast skill of the European Centre for Medium-Range Weather Forecasts (ECMWF) model, one of the eleven S2S models, in predicting monthly average streamflow at lead times of zero and 20 days in 16 European catchments. They used the model output statistics (MOS) method to regress observed streamflow on ECMWF hindcast data. They found that prediction skill of monthly streamflow was frequently absent in comparison with streamflow climatology, especially at the lead time of 20 days. Furthermore, forecast skill varied greatly among the predictor combinations, catchments and dates of prediction. Li et al. (2019) evaluated the precipitation forecasts of eight S2S models for streamflow simulation, particularly for extreme events, using a hydrologic model applied to four basins in China, with drainage areas of 3,312-52,150 km^2. They found that the S2S models had some prediction skill for daily precipitation up to a lead time of 11 days. They also found that bias-corrected streamflow simulations outperformed bias-corrected precipitation in simulating the amount of maximum continuous streamflow and peak flows.

All the above studies used the WWRP/WCRP S2S database. The NOAA's SubX differs from it by having a research-to-operations focus, and hence includes operational as well as research models and produces forecasts in near real-time (Pegion et al., 2019). Here, we instead use NOAA's SubX reforecast precipitation and temperature subseasonal reforecasts given its almost immediate availability of forecasts, with specific attention to AR-related storms and flooding along the coastal Western U.S. Pegion et al. (2019) evaluated the skill of the week 3 averages (average of days 15-21 of the forecast period)

of the seven SubX models globally. They found greater skill in temperature as contrasted with precipitation forecasts for lead three weeks. Baker et al. (2019) found similar results in an evaluation of the skill of the Climate Forecast System version 2 (CFSv2), one of the SubX models, over the conterminous United States (CONUS) domain. Despite the fact that precipitation skill dropped quickly by weeks 2-3, the West Coast showed the highest skill over the CONUS during the winter months (Baker et al., 2019).

The forecast skill of meteorological forcings (particularly precipitation) is an important determinant of flood prediction skill, however antecedent hydrological conditions play an important role as well (e.g. Mahanama et al., 2008). For instance, low antecedent soil moisture (ASM) (as is often the case along the Pacific Coast early in the winter season) has been shown to be an offsetting factor for several extreme historical AR events in California's Russian River basin that otherwise would have lead to major flooding (Cao et al., 2019).

Ensemble streamflow prediction (ESP) and reverse-ESP (revESP) experiments have been used in previous studies to examine the relative importance of initial hydrological conditions (denoted as "IHCs") and climate forecast error as sources of streamflow forecast uncertainty at seasonal time scales (e.g. Wood and Lettenmaier, 2008; Li et al., 2009; Shukla and Lettenmaier, 2011). In ESP, a hydrologic model with assumed perfect IHCs is forced by an ensemble of meteorological forcings resampled from past observations. In contrast in revESP, the model is forced with assumed perfect meteorological forecasts with an ensemble of resampled IHCs. Here we used the ESP/revESP method(s) to partition the relative contributions of ASM and meteorological forecast skill to errors in flood forecasts at subseasonal time scales.

In addition to the ESP/revESP construct for partitioning the role of IHCs and meteorological forecast skill in hydrologic forecasts, ESP can be used as a baseline for determining the contribution of meteorological forecast skill to hydrological (e.g. streamflow) forecast skill (Li et al., 2009). For example, Monhart et al. (2019) compared a traditional ESP approach with subseasonal forecasts from ECMWF (after statistical downscaling) in three alpine catchments with areas of 80-1,700 km^2. They found the ECMWF forecasts could provide added value relative to ESP especially at shorter lead times.

Given this background, our motivating questions are:

1) What is the subseasonal forecast skill (at 1-4 week lead times) of AR-related flooding in coastal Western U.S. watersheds?

2) What are the relative contributions of ASM and subseasonal weather forecast uncertainties to errors in flood forecasts? Are SubX-based streamflow forecasts more skillful than traditional ESP?

4.2 Study region

We focused on three watersheds that form a transect along the U.S. Pacific Coast: the Chehalis River basin in Washington State, the Russian River basin in Northern California, and the Santa Margarita River basin in Southern California. These are the same watersheds used in our previous study, Cao et al. (2020), where we examined future climate impacts on the role of ASM in AR-related floods. AR landfalls show a marked seasonal progression from the Pacific Northwest in the late fall when they are most frequent to northern California in early winter (Gershunov et al., 2017). The geographical

locations of these three watersheds reflect different AR landfalling signatures. Their drainage areas are roughly similar (5400 km^2, 3850 km^2 and 1870 km^2, respectively for the Chehalis, Russian, and Santa Margarita) (see Figure 4.1). The precipitation in all three basins is strongly winter-dominant, varying, on an annual basis, from 1560-2700 mm, 320-1580 m and 160-750 mm in the Chehalis, Russian and Santa Margarita River basins, respectively, during 1999-2016. During these years, 79%, 87% and 83% of precipitation fell between October and March.

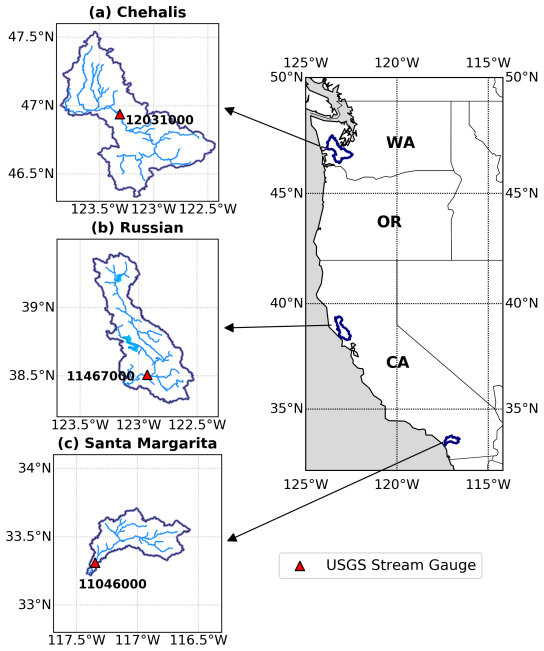

Figure 4.1 Map of study region including a) the Chehalis River basin in Washington State, b) the Russian River basin in Northern California, and c) the Santa Margarita River basin in Southern California.

4.3 Data and methods

We first downscaled the SubX output to a finer spatial resolution, given its coarse native resolution of $1^o \times 1^o$ with respect to our study domain, and the high spatial resolution of our hydrological model. We implemented the Distributed Hydrology-Soil-Vegetation Model (DHSVM) (Wigmosta et al., 1994; see Chapter 3 for details of the

model implementation) in the three basins and ran the model with the downscaled SubX forcings. We then evaluated the skill of streamflow forecasts relative to observations. We also examined how the relative contribution of meteorological forcings and ASM to errors in flood forecasts evolved with lead time.

4.3.1 Downscaling of meteorological forcings

We used 5 out of 7 models in the SubX database (available at http://iridl.ldeo.columbia.edu/SOURCES/.Models/.SubX/), excluding one model with only one ensemble member. There are 26 ensemble members in total for the five models we used (see Table 4.1). The reforecasts (retrospective forecasts) cover the period 1999-2016. We only used data during winter months (Oct-Mar) when most precipitation events occur. The initialization interval of the models is at least once a week and the lead time is at least 32 days. The spatial resolution of the SubX output is $1^{o}\times1^{o}$ and its temporal resolution is daily. We downscaled the SubX output to $1/16^{o}\times1/16^{o}$ using the gridded observation dataset of Livneh et al. (2013) (extended to 2018 as described in Su et al., 2020) as the training dataset.

Table 4.1 List of SubX models used in this study

Model	Ens Members	Init Interval [days]	Forecast period [days]	Reference(s)
ECCC-GEPS5	4	7	32	Lin et al. (2016)
EMC-GEFS	11	7	35	Zhou et al. (2016, 2017); Zhu et al. (2018)
ESRL-FIMr1p1	4	7	32	Sun et al. (2018a,b)
GMAO-GEOS_V2p1	4	5	45	Koster et al. (2000); Molod et al. (2012); Reichle and Liu (2014); Rienecker et al. (2008)
RSMAS-CCSM4	3	7	45	Infanti and Kirtman (2016)

We used two statistical downscaling methods: a) Bias Correction and Spatial Downscaling (BCSD) (Wood et al., 2004) and b) Localized Constructed Analogs (LOCA) (Pierce et al., 2014). We implemented BCSD at a daily time scales. Daily BCSD has been shown to be an effective approach for removing bias (e.g. Monhart et al., 2018; Baker et al., 2019) in atmospheric model output. We applied daily BCSD to precipitation, maximum daily temperature (Tmax), minimum daily temperature (Tmin) and wind speed.

Some analyses have indicated that constructed analog-based techniques may outperform daily BCSD in capturing the magnitude of extremes (e.g. Abatzoglou and Brown, 2012). Therefore we also implemented the constructed analog-based method, LOCA, for the downscaling of precipitation, arguably the most important hydrologic forcing. We summarize these two approaches briefly below.

4.3.1.1 Daily BCSD

There are two steps in daily BCSD: 1) spatial (bilinear) interpolatation of the $1^{o}\times1^{o}$ daily SubX model output to $1/16^{o}\times1/16^{o}$; 2) for each $1/16^{o}$ grid cell, application of the quantile mapping (QM) method (Wood et al., 2002) for bias correction. The training period we used was 1999-2016. We applied both steps to each of the five SubX models.

When applying the QM, we pooled the reforecast days and observation days with similar climatology (see Figure C1 for illustration). In so doing, we wanted to preserve the forecast skill with respect to lead times. Hence we performed the QM in a lead-time-dependent manner similar to Monhart et al. (2018) in their bias correction of ECMWF model output. The general steps in our processing included: a) for a given forecast initialization date, we first found the initialization falling within a 15-day window

centered on the given forecast initialization date over the 18 year reforecast period; b) for each lead on the given initialization date, we selected the reforecast days with the same lead time from all ensemble members of the same model; c) we pooled the climatology (observations) based on a 15-day window centered on reforecast calendar days over 18 years; d) we replaced the reforecast value with a value from the observation climatology with the same quantile based on their empirical distributions.

When the percentile of a reforecast value was outside the range of the empirical percentile of observations (based on Weibull plotting positions), we fit theoretical probability distributions to the data. Following Wood et al. (2002), we used the Gumbel distribution for upper tails and the Weibull distribution for lower tails for precipitation. Additionally, we examined the effect of populating the sample distribution in step b) by pooling the following seven days (denoted as "BCSD_7d") versus using one day only (denoted as "BCSD_1d").

Similar to precipitation, we applied daily BCSD to Tmax, Tmin and wind speed. For temperature, we used a normal distribution for the tails. For wind speed, we used only the empirical distribution.

4.3.1.2 LOCA

We also used used LOCA (Pierce et al., 2014) for the downscaling of precipitation. We used the same 1999-2016 as for daily BCSD. Before applying LOCA, we applied bias correction of the coarse SubX grids, using the same QM steps as bor BCSD (section 3.1.1).

The general procedures of LOCA include: a) after daily observations are coarsened to $1^{o}\times1^{o}$, a seasonal spatial mask is generated for each coarse grid cell by excluding those

with no spatial correlation. This step aims to limit the analog pool domain at the regional scale; b) select 30 analog days for each coarse grid cell at the regional scale, by minimizing the root mean squared error (RMSE) between a target day and calendar days within a 45-day window over the training period excluding the target year; c) after the SubX data and the coarsened observation are interpolated to $1/16^{o}\times1/16^{o}$, the single best matching day is determined at the local scale by minimizing the RMSE over a region of 21×21 $1/16^{o}$ grid cells between a target day and the 30 analog days. The first three steps only area based on coarsened observations; d) calculate a scaling factor from c) and scale the fine scale observation data.

After these steps, we added an additional bias correction step following Abatzoglou and Brown (2012). We applied a final QM procedure to the downscaled output to ensure statistical moments of the downscaled data conform to observations. This step is basically the same as the QM in Section 3.1.1. When applying LOCA, we also compared the pooling of the following seven days (denoted as "LOCA_7d") versus using one day only (denoted as "LOCA_1d") in the QM for both coarse and fine grid cells.

4.3.2 Model implementation

We implemented DHSVM in the three basins with essentially the same model setup as in Cao et al. (2020). DHSVM requires meteorological inputs including precipitation, wind speed, air temperature, relative humidity, downward solar and longwave radiation. In order to run the model at an hourly time step, we disaggregated the daily data to hourly using the Mountain Microclimate Simulation Model (MTCLIM) algorithms as described and implemented by Bohn et al. (2013). We did this for both Livneh et al. (2013) data and downscaled SubX data.

After model calibration, we ran the model using the Livneh et al. (2013) forcings for the period 1999-2016 as a control run. The initialization interval for most SubX models is seven days, but different models are initialized on different days. One primary purpose of the control run was to provide IHCs for each model run using SubX forcings. A multi-model ensemble is usually generated by averaging all forecasts from the same start date, has been termed a lagged average ensemble (Vitart et al., 2017; Pegion et al., 2019). Following this method, we output model states for the 7^{th}, 14^{th}, 21^{st} and 28^{th} of each month. This results in 432 IHC dates in total over 18 years. For each IHC date, we identified the latest SubX model initialization date within the previous week. For each SubX ensemble member and each initialization date, we ran DHSVM for 28 days (four-week forecast).

4.3.3 Assessment of streamflow forecast skill

4.3.3.1 Identification of AR-related extreme events

We used the Peaks Over Threshold (POT) method to identify extreme discharge events as in Cao et al. (2020). We first applied the event independence criteria from USWRC (1982) to daily streamflow data in order to identify independent discharge events. We set thresholds at each stream gauge that resulted in 1, 2 and 3 extreme events per year on average, which we denote as POT_{N1}, POT_{N2} and POT_{N3}.

We examined AR contributions to extreme events by identifying the flood events that were coincident with AR events. The AR date catalog we used is based on ECMWF Reanalysis-Interim (ERA-Interim) data set, from Guan and Waliser (2015). We extracted the grid cells from the catalog that intersected each basin and identified them as potential AR-related floods.

4.3.3.2 Evaluation metrics

We used the evaluation metrics from DeFlorio et al. (2019) where they assessed the subseasonal forecast skill of ARs, including 1) debiased Brier skill score (BSS) (Weigel et al., 2007), which is used to evaluate the skill of probabilistic forecasts; 2) Relative Operating Characteristic (ROC; Hanley and McNeil, 1982)-like diagrams, which account for both hit rate and false alarm rate in a lead-dependent manner. We discuss these two skill measures and our application of them briefly below.

a. BSS

The BSS is a relative measure of probabilistic skill that is sensitive to small ensemble sizes. Following DeFlorio et al. (2019), we used the debiased BSS, which adds a correction term in the denominator of BSS to overcome the small ensemble size issue. The debiased BSS can be calculated as follows:

$$\mathrm{BSS} = 1 - \frac{\mathrm{BS}}{\mathrm{BS}_{\mathrm{ref}} + \mathrm{D}} \tag{4.1}$$

$$\mathrm{BS} = \frac{1}{N}\sum_{i=1}^{N}\left(P_i - O_i\right)^2 \tag{4.2}$$

$$\mathrm{BS}_{\mathrm{ref}} = \frac{1}{N}\sum_{i=1}^{N}\left(P_{\mathrm{clim}} - O_i\right)^2 \tag{4.3}$$

$$D = \frac{1}{M}P_{\mathrm{clim}}\left(1 - P_{\mathrm{clim}}\right) \tag{4.4}$$

where N is the number of reforecast samples during Oct-Mar; Pi represents the forecast ability of a particular level of discharge events, which is the fraction of ensemble members that predict maximum discharge falling into a particular category during a week-long lead period for a single reforecast event; Oi is the binary representation of whether the observed discharge fell into that category (1 if yes, 0 if no); M is the ensemble size; and $Pclim$ is the probability of the reference climatology.

This metric can be sensitive to the choice of the reference climatology (Bartholmes et al., 2009). Here, we focused on POT extreme discharge events. For POT_{N1} events (with threshold set to one event per year on average), the *Pclim* is set to 1/24, corresponding to one extreme event occurring over 6 months (i.e. 24 weeks). For POT_{N2} events (with threshold set to two events per year on average), *Pclim* is set to 2/24, and so forth. BSS ranges from $-\infty$ to 1. Values above 0 indicate that the reforecast skill is higher than skill from a forecast using reference climatology.

b. ROC-Like Diagrams

A ROC-Like diagram shows the ensemble mean, hit rates, and false alarm rates. The hit rate is calculated as the number of hits divided by the total number of hits + misses. The false alarm rate is calculated as the number of false alarms divided by the total number of false alarms + correct rejections. For each POT category and each week of lead time, the terms comprising these rates are defined as below:

Hit = an event is observed and it is forecasted.

Miss = an event is observed but it is not forecasted.

False alarm = an event is forecasted but it is not observed.

Correct rejection = an event is not forecasted and it is not observed.

4.3.4 ESP and revESP implementation

In our implementation of ESP, we used IHCs from Section 3.2. ESP considers the IHCs to be "true" and the model is forced with resampled gridded observations. For each IHC date, we extracted the following 28-days of observed forcings for the same calendar day period over the 18 years of the reforecast period. We then forced the model with 18

ensemble members starting from the IHC date for a period of 28 days. We compared the SubX-based forecasts with ESP to see which was most skillful.

Similarly, revESP samples IHCs from climatology to initialize the model, which is forced with "true" observations. For a given forecast date, the revESP experiments sampled 18 IHCs for the same calendar day over 18 years to initialize the model. The model was then forced with the true (observed) forcings for a period of 28 days.

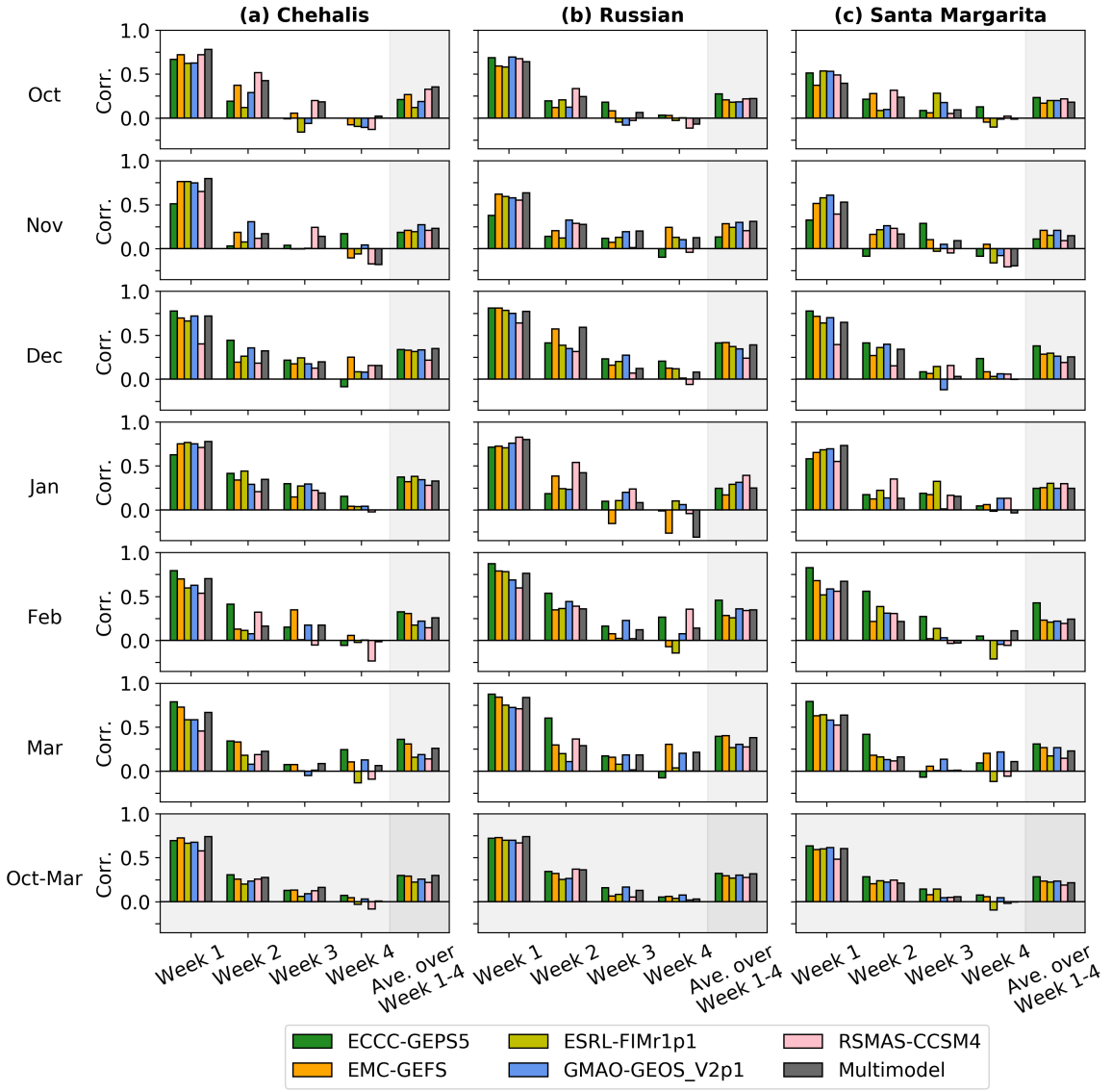

Figure 4.2 Precipitation skill (in correlation) of SubX models averaged over each basin and each month before bias correction.

4.4 Results

4.4.1 Evaluation of SubX reforecast focings

4.4.1.1 Precipitation and temperature skill

We examined the precipitation skill of the individual SubX models, as well as the multimodel ensemble mean (denoted as "MME"), at lead times of 1-4 weeks in each basin (see Figure 4.2). We examined each month during the Oct-Mar period separately as well. Figure 4.2 shows that precipitation skill (as measured by correlation with observations) drops quickly after week 1. In week 2, almost all models have positive correlations in all months, but by week 3, some models show no skill in certain months. Over all months (Oct-Mar; see bottom panel in Figure 4.2), the precipitation skill of individual models in the Chehalis River basin are 0.58~0.73, 0.20~0.31, 0.06~0.13, and -0.08~0.07 at lead times 1-4 weeks, respectively, while those of MME are 0.74, 0.28, 0.16, and 0.01. The precipitation skill of individual models in the Russian River basin are 0.67~0.73, 0.25~0.37, 0.05~0.17, and 0.02~0.08, while those of MME are 0.74, 0.36, 0.13, and 0.03. In the Santa Margarita River basin, the skills of individual models are 0.48~0.63, 0.21~0.28, 0.05~0.14, and -0.10~0.08, while those of MME are 0.60, 0.21, 0.06, and 0.00. Overall, model skills in the Chehalis and Russian River basins are similar, but are somewhat lower in the Santa Margarita basin.

In terms of skills averaged over weeks 1-4, the skills of individual models are 0.22~0.30, 0.27~0.32, and 0.19~0.28 in the three basins respectively, while those of MME are 0.30, 0.32 and 0.22. Among individual models, ECCC-GEPS5 performed the best in the three basins. It had similar skill to MME in the Chehalis and Russian River basins, and even higher skill than MME in the Santa Margarita River basin. However,

model performances varied by lead time, month, and basin. No individual model had consistently better performance than others in weeks 3-4 across the three basins.

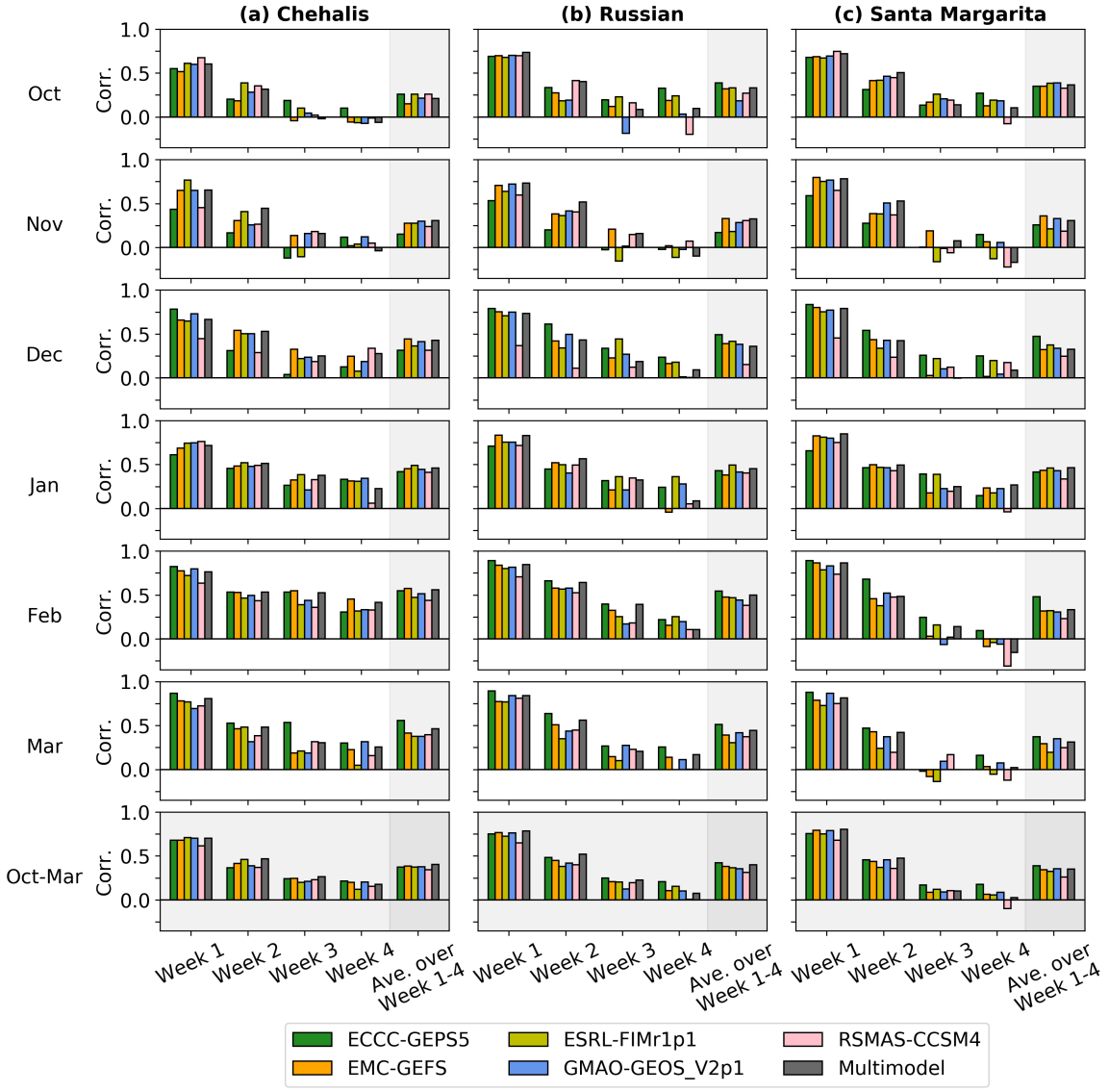

Figure 4.3 Maximum daily temperature (Tmax) skill (in correlation) of SubX models averaged over each basin and each month before bias correction.

We also examined the temperature forecast skill. Figure 4.3 shows the results for Tmax. (Since the patter for Tmin is similar, it is shown in supplement Figure C2). Similar to precipitation, Tmax skill drops quickly after week 1. However, there are fewer

negative correlations for Tmax in weeks 3-4 in comparison with precipitation. In general, Tmax generally has higher skill than precipitation in week 2-4.

Over all months (Oct-Mar; see bottom panel in Figure 4.3), the Tmax skills of individual models were 0.61~0.71, 0.37~0.46, 0.20~0.25, and 0.12~0.21 at lead times of week 1-4 respectively for the Chehalis River basin, while those of MME were 0.70, 0.47, 0.27 and 0.18. The Tmax skills for individual models in the Russian River basin were in the range 0.65~0.77, 0.38~0.48, 0.13~0.25, and 0.01~0.21, while those of MME were 0.79, 0.52, 0.22 and 0.08. In the Santa Margarita River basin, the skills of individual models were in the range 0.68~0.79, 0.36~0.46, 0.09~0.17, and -0.10~0.18, while those of MME are 0.80, 0.48, 0.10 and 0.03. Averaged over weeks 1-4, the skills of individual models were 0.34~0.38, 0.31~0.42, and 0.26~0.39 in three basins respectively, while those of MME are 0.42, 0.40 and 0.35. Overall, temperature skills were roughly similar in the Russian and Santa Margarita River basins, and slightly lower in the Chehalis.

4.4.1.2 Performance of downscaling methods

We examined the performance of daily BCSD and LOCA. Figure 4.4 shows the precipitation skill of one SubX model, EMC-GEFS, before and after applying these two methods. The difference in precipitation skills before and after applying either daily BCSD or LOCA is very small, which meets our expectation since the QM is done in a lead-time-dependent manner. For the averaged skill over week 1-4 and over all months, the difference is within 0.00, 0.01 and 0.03 in the Chehalis, Russian, and Santa Margarita River basins, respectively, with the largest difference associated with LOCA_1d (LOCA with the pooling of one day only in each init).

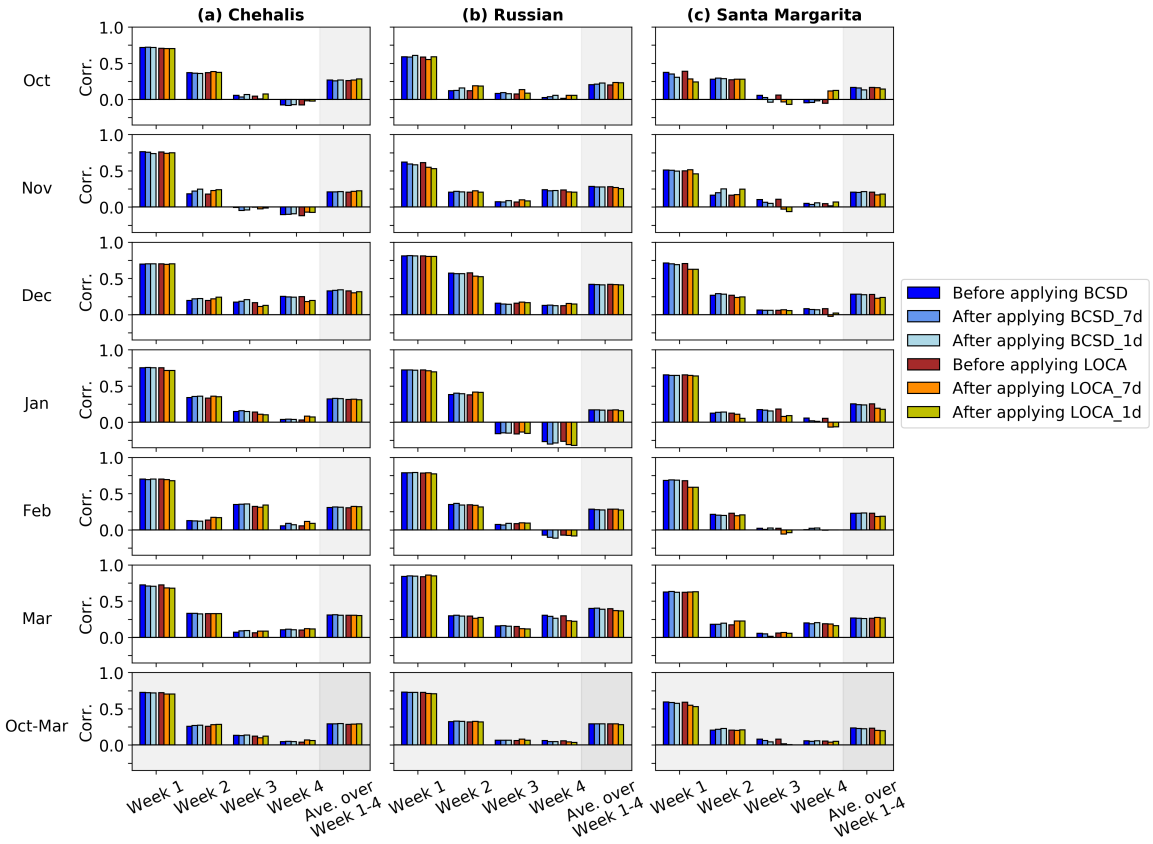

Figure 4.4 Basin-average precipitation skill (correlation) in one of SubX models, EMC-GEFS, before and after applying daily BCSD and LOCA in three basins.

Figure 4.5 shows the basin-average relative bias of precipitation forecasts in EMC-GEFS, before and after applying daily BCSD and LOCA in three basins. Before applying daily BCSD, the basin-average relative biases over all months were -21%, -3% and -32% in the three basins respectively. They were reduced to 1%, -2% and 1% after applying BCSD_7d, and 1%, 1% and 8% after applying BCSD_1d. The change in relative bias is small in the Russian River basin because the values reported here are averaged over the basin and over all months. The change in each month as well as in spatial patterns is much larger. It is more spatially coherent after bias correction (see Figure C3-C5). Similarly, before applying LOCA, the basin-average relative biases over all months were -25%, -1% and -26% in the three basins respectively. They changed to -1%, -2% and 1% after applying BCSD_7d, and 0%, 2% and 5% after applying BCSD_1d.

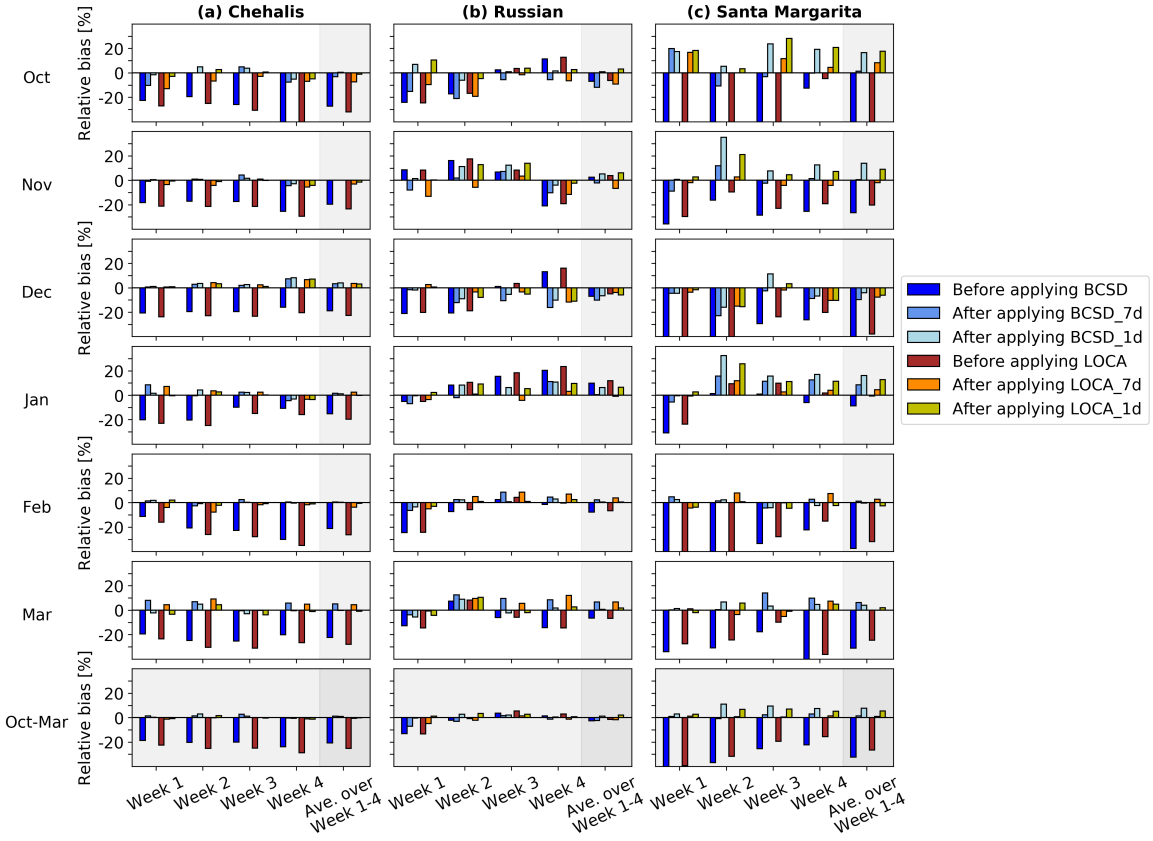

Figure 4.5 Basin-average relative bias of precipitation in one of SubX models, EMC-GEFS, before and after applying daily BCSD and LOCA in three basins.

For each month (from Oct to Mar) and each lead week, the relative biases were reduced to less than ±10% after bias correction in most cases in the Chehalis and Russian River basins, while in the Santa Margarita River basin, the relative biases after bias correction were larger, especially in dry months. In most cases, they are within ±20%. The difference between BCSD_7d and LOCA_7d, or BCSD_1d and LOCA_1d, is generally smaller than the difference between BCSD_7d and BCSD_1d, or LOCA_7d and LOCA_1d. Across basins, LOCA performed slightly better in terms of bias in comparison with BCSD in the Santa Margarita River basin, but the performance of the two methods was similar performance in the other two basins. Given that the

computational time and effort of LOCA are several times higher than BCSD, we report the streamflow forecast skill only for BCSD below.

4.4.2 Hydrologic model evaluation

We evaluated simulated streamflow from the control run, forced by the Livneh et al. (2013) data. We used the Kling-Gupta efficiency (KGE) (Gupta et al., 2009), normalized root-mean-square error (NRMSE) and relative bias to evaluate the goodness-of-fit between daily streamflow observations and aggregated daily simulations at the downstream-most USGS stream gauge in each basin (for other gauges, see Table C1). The KGE, NRMSE and relative bias were 0.94, 0.26 and 3.5% respectively in the Chehalis River basin during the calibration period 1999-2007. They were 0.89, 0.27 and 9.2% in the Russian River basin, and 0.56, 0.57 and 27.4% in the Santa Margarita River basin. Reduced performance of the model in the Santa Margarita basin likely is related to the more variable precipitation regime there. Model performance deteriorated slightly during the validation period (2008-2016) in the Russian River basin, but was similar in calibration and validation periods in the other two basins.

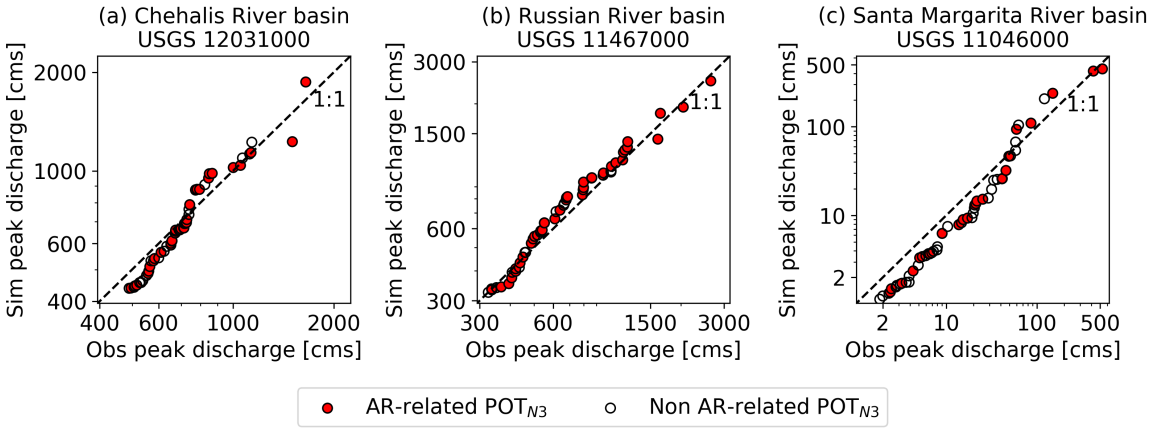

Figure 4.6 Evaluation of hydrologic model performance: simulated versus observed daily peak flow of POT_{N3} extreme discharge events (with threshold set to 3 events per year on average) during the period of 1999-2016. The events associated with ARs are marked by circles filled with red color.

We examined model performance during extreme discharge events given our interest in flood forecasting. Figure 4.6 compares simulated and observed peak flows for POT_{N3} extreme discharge events during the SubX period. The simulated peak flows generally align with observations in three basins, with some underestimation for smaller events in the Chehalis and Santa Margarita River basins. In Figure 4.6, we also show model performance for AR events. The percentages of POT_{N3} extreme discharge events that were coincident with ARs during 1999-2016 are 52%, 74% and 41% respectively in the Chehalis, Russian and Santa Margarita River basins. We can see that most of the largest extreme discharge events were AR-related, especially in the Russian River basin.

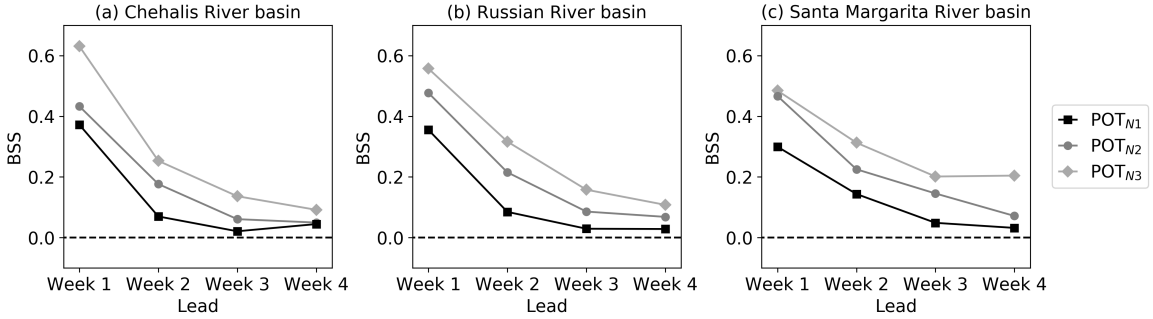

Figure 4.7 Brier skill score (BSS) over weeks 1-4 lead time for POT_{N1}, POT_{N2} and POT_{N3} extreme discharge events (with threshold set to 1, 2 and 3 events per year on average).

4.4.3 Assessment of streamflow forecast skill

Figure 4.7 shows the BSS values for week 1-4 lead time for POT_{N1}, POT_{N2} and POT_{N3} extreme discharge events (with thresholds set to 1, 2 and 3 events per year on average) in the three basins. For the POT_{N1} events, the BSS drops quickly after week-1 and is close to 0 after week-2 in the Chehalis and Russian River basins. BSS is a little higher in week-3 in the Santa Margarita River basin in comparison with the other two basins, but this could be due to large uncertainty in this basin. When we lower the

threshold to POT_{N2} and POT_{N3}, BSS increases in all weeks, which may be partly due to the influence of ASM.

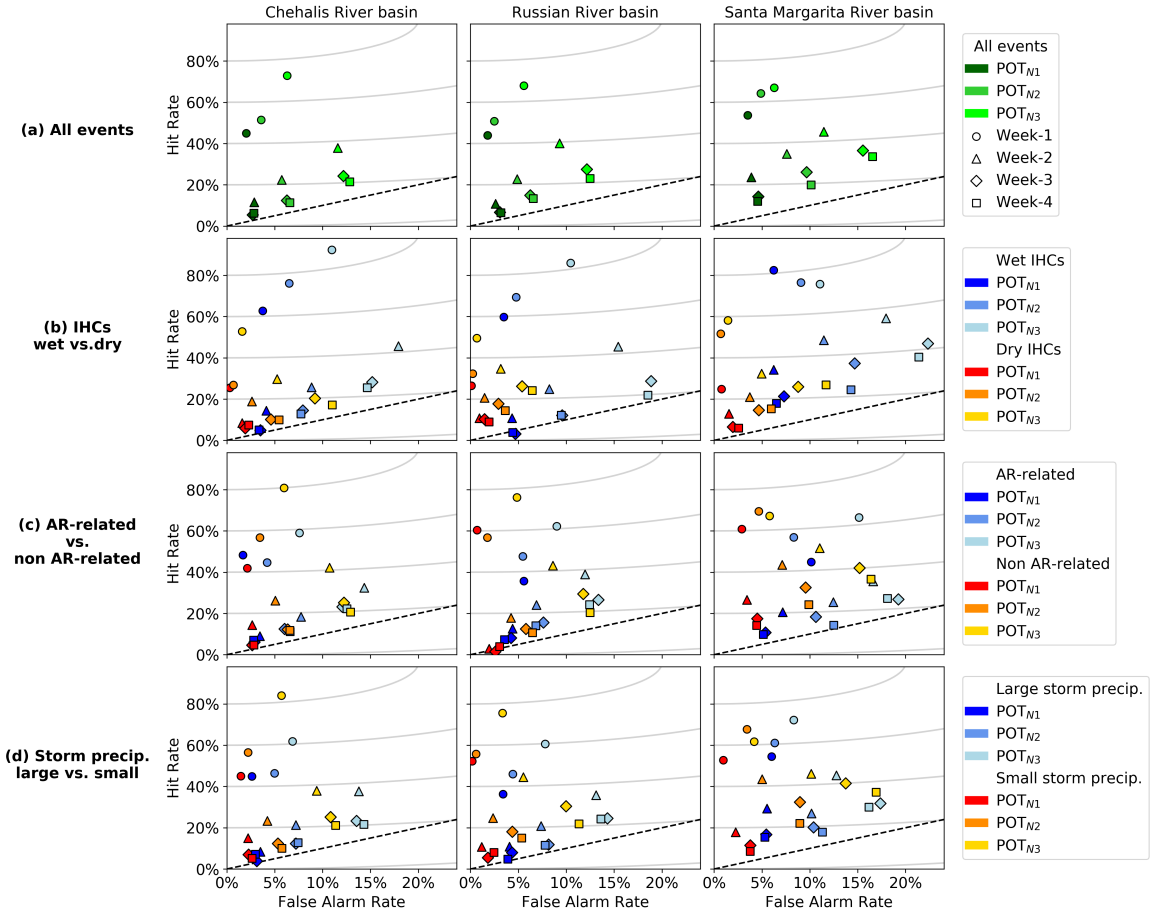

Figure 4.8 ROC-like diagrams (ensemble mean hit rate vs. false alarm rate) of POT_{N1}, POT_{N2} and POT_{N3} extreme discharge events in three basins, for a) all events; b) events with wet initial hydrological conditions (IHCs) vs. events with dry IHCs, separated by median values; c) AR-related events vs. non AR-related ones; and d) events with large storm precipitation vs. events with small storm precipitation, separated by median values.

We examined the hit rate and false alarm rate in ROC-Like diagrams (see Figure 4.8a). For POT_{N1} events, the hit rate drops quickly from week-1 to week-2 in all three basins and the false alarm rate slightly increases. After week-2, the positive skill is very low. As the threshold is lowered to POT_{N2} and to POT_{N3}, the hit rate increases but so does the false alarm rate. Across basins, the hit rate in the Santa Margarita River basin is higher after week-1 in comparison with the other two basins, but its false alarm rate is

also higher, indicating the large natural variability (and hence forecast uncertainty) in this basin.

4.4.4 The role of ASM in streamflow forecast

Streamflow forecast errors and lead times for ESP and revESP experiments are shown in Figure 4.9. The RMSE for ESP generally increased with lead time in all three basins but with large variations in the Santa Margarita River basin. The RMSE for revESP generally decreased with lead times in all three basins but also with large variations in the Santa Margarita River basin. The RMSE for revESP was larger than that for ESP in the first few days, indicating that ASM dominates the streamflow forecast skill at shorter lead times. We also compared ESP with SubX-based forecasts. The RMSE for SubX eas lower than ESP over all lead times in all three basins, with a few exceptions in the Santa Margarita River basin. This indicates that SubX forecasts are more skillful than ESP. In addition, we compared the effect of populating the sample distribution for QM in daily BCSD by pooling the following seven days of a model day versus using one day only. We found that the RMSE of BCSD_7d was slightly lower than the latter for BCSD_1d, but the difference increased with lead times.

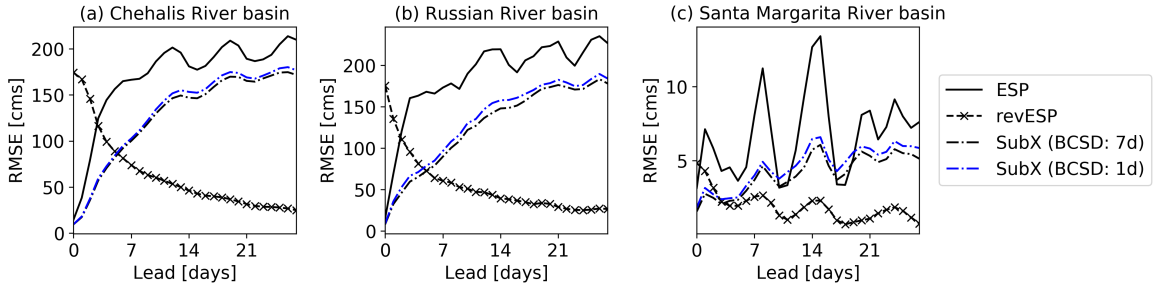

Figure 4.9 RMSE of streamflow forecasts for ESP, revESP, and SubX-based forecasts.

We further examined the RMSE ratios to evaluate the relative contributions of ASM and subseasonal weather forecast skill to streamflow forecast skill. Figure 4.10

shows the ratios by month. The ESP-based ratio is generally higher than the SubX-based ratio, but the difference becomes smaller with lead time. If the RMSE ratio is less than one, we infer that ASM dominates the streamflow forecast skill and vice versa.

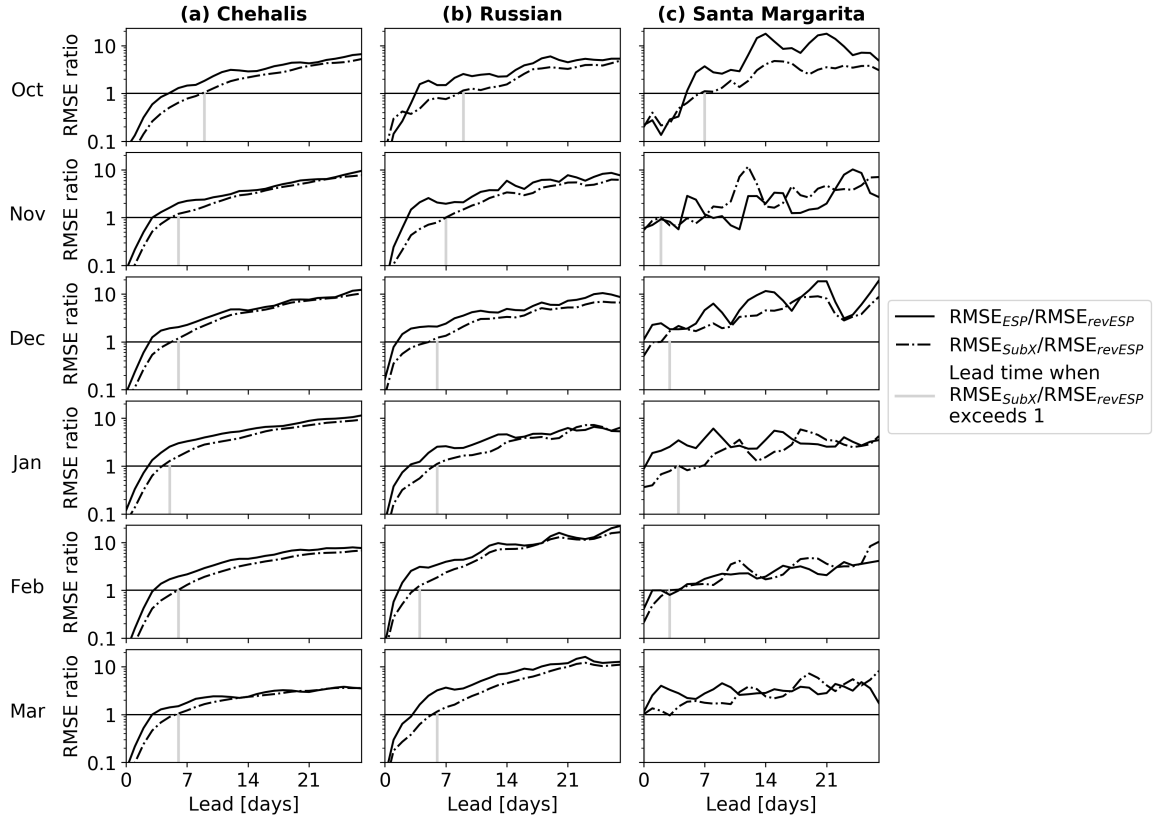

Figure 4.10 Variation of RMSE ratios ($RMSE_{ESP}/RMSE_{revESP}$ and $RMSE_{SubX}/RMSE_{revESP}$) with lead time in three river basins in each month from October to March. The lead time when $RMSE_{SubX}/RMSE_{revESP}$ exceeds one is marked by a vertical gray line.

Figure 4.10 shows that the lead time when the SubX-based ratio exceeds one generally decreases from October to March in all three basins, suggesting that ASM dominates the streamflow forecast skill in fall months, but that the effect is reduced through the winter. This makes sense, as early fall soil moisture is dominated by the prolonged preceding summer dry period. Across the three basins, the lead time when SubX forecast skill starts to dominate streamflow forecast skill generally decreases from

north to south. Overall, the SubX forecast skill starts to dominate streamflow forecast skill after week-1 except in October months.

4.5 Discussion

In Section 4.3, we showed that the forecast skill of extreme discharge events generally increased at all leads when the POT threshold was reduced (see Figure 4.7 and Figure 4.8a) – meaning that our event population becomes less dominated by the largest floods. Here we further examine the potential influencing factors, such as IHCs, ARs, and storm precipitation. We separated all POT events shown in Figure 4.8a into two groups for each factor. The three pairs for comparison are: 1) events with wet IHCs vs. events with dry IHCs, separated by median values; 2) AR-related events vs. non AR-related events; 3) events with large storm precipitation vs. events with small storm precipitation, separated by median values (see Figure 4.8b-d).

Figure 4.8b shows that forecast skill is generally higher when IHCs are wet than dry for all weeks and across all basins, except for weeks 3-4 in the Russian River basin. However, the difference between the two groups decreases as the POT threshold increases and lead time increases. Also, forecast skill is generally higher for the non AR-related events than for AR-related ones (see Figure 4.8c). If we consider the POT_{N3} events, which have a larger sample size (and hence smaller magnitudes) than the other two POT categories, this tendency (higher skill for non-AR events) is stronger in weeks 1-2 in the Chehalis and Russian River basins, and for all lead times in the Santa Margarita River basin. Events associated with ARs generally have higher storm precipitation than non-AR events, so this suggests that forecast skill is higher for less

extreme events, which is not particularly encouraging for hazard mitigation aspects of extended weather prediction. The pattern for the skill of events separated by storm precipitation indicates similar message to that by ARs (see Figure 4.8d). The forecast skill of extreme discharge events is higher when storm precipitation is small. Despite the fact that AR vs non-AR, and groups separated storm precipitation amounts show the above differences (higher forecast skill for smaller events), they are not as distinct as the effects of IHCs -- especially at shorter lead times and for lower POT thresholds.

4.6 Conclusions

We examined the performance of SubX reforecasts for subseasonal forecasts of AR-related flooding in three watersheds along the coastal Western U.S. We first evaluated SubX forecast skill for precipitation and temperature. After the statistical downscaling and bias correction of the forcings, we ran the DHSVM model in each of the three basins, with focus on peak-over-threshold events in the period 1999-2016. We further evaluated the relative contributions of ASM and subseasonal weather forecast skill to streamflow forecast skill with leads from one to four weeks using ESP and revESP experiments. Based on our analysis, we find

1) Over all months (Oct-Mar), SubX precipitation skill drops quickly after week 1 lead, but still has useful skill at weeks 2-3, while at week 4, most models show negligible skill. ECCC-GEPS5 overall performed best among individual models for all basins, with performance that is comparable with MME. Precipitation forecast skill is generally higher in the Chehalis and Russian River basins than in the Santa Margarita River basin (towards the southern end of our domain). Generally, there is higher skill in temperature

than precipitation forecasts, with all models showing usable skill through lead 3 weeks, but all showing little or no skill at week 4. For precipitation downscaling and bias correction, the difference between the two methods we examined, daily BCSD and LOCA, was smaller than the difference resulted from choices in the pooling of days (7 days vs. 1 day) with similar climatology.

2) SubX-based forecast skill for extreme discharge events drops quickly after week 1, with minimal forecast skill by week 3 for the largest (POT_{N1}) events. Forecast skill was slightly higher for smaller events (lower POT thresholds). Furthermore, forecast skill was generally lower for AR storms than for non-AR storms. However, forecast skill is influenced more strongly by IHCs than by storm magnitude with lower skill when IHCs are dry, especially at shorter lead times and with lower POT thresholds (forecast skill is diminished less by IHCs for large as contrasted with small storms).

3) SubX-based streamflow forecast skill is higher than that of ESP-based forecasts in all cases with usable precipitation forecast skill (most meaning leads of three weeks or less). Across months, the length of lead times for which ASM dominates streamflow forecast skill generally decreases through the wet season months. Across basins, the lead time when the SubX forecast skill starts to dominate streamflow forecast skill generally decreases from the north to south, but generally occurs at leads no longer than one week except in October months.

Acknowledgement

We acknowledge NOAA, NASA, and the U.S. Navy, the agencies that supported the SubX experiment and archive, and we thank the climate modeling groups

(Environment Canada, NASA, NOAA/NCEP, NRL and University of Miami) for producing and making available their model output. NOAA's Modeling, Analysis, Predictions, and Projections (MAPP) Program, the Office of Naval Research, NASA, and the NOAA National Weather Service jointly provided coordinating support and led development of the SubX system. Access to the AR catalog was provided by Dr. Bin Guan of NASA' Jet Propulsion Laboratory via https://ucla.box.com/ARcatalog. Development of the AR detection algorithm used to produce the AR catalog (by Dr. Guan and colleagues) was supported by NASA. The research supported herein was funded by the Center for Western Weather and Water Extremes (CW3E) at the Scripps Institution of Oceanography via AR Program Phase II, sponsored by the California Department of Water Resources.

References

Abatzoglou, J. T., and T. J. Brown, 2012: A comparison of statistical downscaling methods suited for wildfire applications. *Int. J. Climatol.*, **32**, 772–780, doi:https://doi.org/10.1002/joc.2312.

Baker, S.A., A.W. Wood, and B. Rajagopalan, 2019: Developing Subseasonal to Seasonal Climate Forecast Products for Hydrology and Water Management. *J. Am. Water Resour. Assoc.*, **55**, 1024-1037, https://doi.org/10.1111/1752-1688.12746.

Barth, N. A., G. Villarini, M. A. Nayak, and K. White, 2017: Mixed populations and annual flood frequency estimates in the western United States: The role of atmospheric rivers. *Water Resour. Res.*, **53**, 257-269. https://doi.org/10.1002/2016wr019064.

Bartholmes, J. C., J. Thielen, M.H. Ramos, and S. Gentilini, 2009: The european flood alert system EFAS-Part 2: Statistical skill assessment of probabilistic and deterministic operational forecasts. *Hydrol. Earth Syst. Sci.*, **13**, 141–153, https://doi.org/10.5194/hess-13-141-2009.

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

Cao, Q., A. Mehran, F.M. Ralph, and D.P. Lettenmaier, 2019: The Role of Hydrological Initial Conditions on Atmospheric River Floods in the Russian River Basin. *J. Hydrometeor.*, **20**, 1667–1686, https://doi.org/10.1175/JHM-D-19-0030.1.

——, A. Gershunov, T. Shulgina, F.M. Ralph, N. Sun, and D.P. Lettenmaier, 2020: Floods due to atmospheric rivers along the U.S. West Coast: The role of hydrological initial conditions in a warming climate. *J. Hydrometeor.*, in revision.

DeFlorio, M.J., D.E. Waliser, B. Guan, F.M. Ralph, and F. Vitart, 2018: Global evaluation of atmospheric river subseasonal prediction skill. *Climate Dyn.*, **52**, 3039–3060, https://doi.org/10.1007/s00382-018-4309-x.

——, and coauthors, 2019: Experimental Subseasonal-to-Seasonal (S2S) Forecasting of Atmospheric Rivers Over the Western United States. *J. Geophys. Res. Atmos*., 124. https://doi.org/10.1029/2019JD031200

Dettinger, M., F. Ralph, T. Das, P. Neiman, and D. Cayan, 2011: Atmospheric rivers, floods and the water resources of California. *Water*, **3**, 445–478, https://doi.org/10.3390/w3020445.

Gershunov, A., T. Shulgina, F. M. Ralph, D. A. Lavers, and J. J. Rutz, 2017: Assessing the climate-scale variability of atmospheric rivers affecting western North America. *Geophys. Res. Lett*., **44**, 7900-7908, https://doi.org/10.1002/2017gl074175.

Guan, B., and D. E. Waliser, 2015: Detection of atmospheric rivers: Evaluation and application of an algorithm for global studies, *J. Geophys. Res. Atmos*., **120**, 12514-12535, https://doi.org/10.1002/2015JD024257

Gupta, H. V., H. Kling, K. K. Yilmaz, and G. F. Martinez, 2009: Decomposition of the mean squared error and NSE performance criteria: Implications for improving hydrological modeling. *J. Hydrol.*, **377**, 80–91, https://doi.org/10.1016/j.jhydrol.2009.08.003.

Hanley, J. A., and B.J. McNeil, 1982: The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiol.*, **143**, 29-36, https://doi.org/10.1148/radiology.143.1.7063747

Infanti, J. M., and B. P. Kirtman, 2016: Prediction and predictability of land and atmosphere initialized CCSM4 climate forecasts over North America. *J. Geophys. Res. Atmos.*, **121**, 12 690–12 701, https://doi.org/10.1002/2016JD024932.

Konrad, C. P., and M.D. Dettinger, 2017: Flood runoff in relation to water vapor transport by atmospheric rivers over the western United States, 1949–2015. *Geophys. Res. Lett.*, **44**, 11,456–11,462, https://doi.org/10.1002/2017GL075399.

Koster, R. D., M. J. Suarez, A. Ducharne, M. Stieglitz, and P. Kumar, 2000: A catchment-based approach to modeling land surface processes in a general circulation model: 1. Model structure. *J. Geophys. Res.*, **105**, 24 809–24 822, https://doi.org/10.1029/2000JD900327.

Li, H., L. Luo, E.F. Wood, and J. Schaake, 2009: The role of initial conditions and forcing uncertainties in seasonal hydrologic forecasting. *J. Geophys. Res.*, **114**, D04114, https://doi.org/10.1029/2008JD010969.

Li, W., J. Chen, L. Li, H. Chen, B. Liu, C. Xu, and X. Li, 2019: Evaluation and bias correction of S2S precipitation for hydrological extremes. *J. Hydrometeor.*, **20**, 1887–1906, https://doi.org/10.1175/JHM-D-19-0042.1

Lin, H., N. Gagnon, S. Beauregard, R. Muncaster, M. Markovic, B. Denis, and M. Charron, 2016: GEPS-based monthly prediction at the Canadian Meteorological Centre. *Mon. Wea. Rev.*, **144**, 4867–4883, https://doi.org/10.1175/MWR-D-16-0138.1.

Lin, H., R. Mo, F. Vitart, and C. Stan, 2018: Eastern Canada Flooding 2017 and its Subseasonal Predictions. *Atmos.-Ocean*, **57**, 195-207, https://doi.org/10.1080/07055900.2018.1547679

Livneh B., E.A. Rosenberg, C. Lin, B. Nijssen, V. Mishra, K.M. Andreadis, E.P. Maurer, and D.P. Lettenmaier, 2013: A Long-Term Hydrologically Based Dataset of Land Surface Fluxes and States for the Conterminous United States: Update and Extensions. *J. Clim*., **26**, 9384–9392, https://doi.org/10.1175/JCLI-D-12-00508.1

Mahanama, S. P. P., R. D. Koster, R. H. Reichle, and L. Zubair, 2008: The role of soil moisture initialization in subseasonal and seasonal streamflow predictability: A case study in Sri Lanka. *Adv. Water Resour*., **31**, 1333–1343, doi:10.1016/j.advwatres.2008.06.004.

Mariotti, A., P. M. Ruti, and M. Rixen, 2018: Progress in subseasonal to seasonal prediction through a joint weather and climate community effort. *npj Climate Atmos. Sci*., **1**, 4, https://doi.org/10.1038/s41612-018-0014-z.

Molod, A., L. Takacs, M. J. Suarez, J. Bacmeister, I.-S. Song, and A. Eichmann, 2012: The GEOS-5 atmospheric general circulation model: Mean climate and development from MERRA to Fortuna. *Tech. Memo. NASA/TM-2012-104606*, Vol. **28**, 115 pp., https://ntrs.nasa.gov/archive/nasa/casi.ntrs.nasa.gov/20120011790.pdf.

Monhart, S., C. Spirig, J. Bhend, K. Bogner, C. Schär, and M.A. Liniger, 2018: Skill of subseasonal forecasts in Europe: Effect of bias correction and downscaling using surface observations. *J. Geophys. Res. Atmos.*, **123**, 7999–8016, https://doi.org/10.1029/ 2017JD027923.

Monhart, S., M. Zappa, C. Spirig, C. Schär, and K. Bogner, 2019: Subseasonal hydrometeorological ensemble predictions in small- and medium-sized mountainous catchments: benefits of the NWP approach. *Hydrol. Earth Syst. Sci*., **23**, 493–513, 2019 https://doi.org/10.5194/hess-23-493-2019

Mundhenk, B.D., E.A. Barnes, E.D. Maloney, and C.F. Baggett, 2018: Skillful empirical subseasonal prediction of landfalling atmospheric river activity using the Madden-Julian oscillation and quasi-biennial oscillation. *npj Climate Atmos. Sci*., **1**, 20177. https://doi.org/10.1038/s41612-017-0008-2

Nardi, K.M., E.A. Barnes, and F.M. Ralph, 2018: Assessment of Numerical Weather Prediction Model Reforecasts of the Occurrence, Intensity, and Location of Atmospheric Rivers along the West Coast of North America. *Monthly Weather Review*, **146**, 3343–3362, https://doi.org/10.1175/MWR-D-18-0060.1

Neiman, P. J., L. J. Schick, F. M. Ralph, M. Hughes, and G. A. Wick, 2011: Flooding in Western Washington: The Connection to Atmospheric Rivers. *J. Hydrometeor.*, **12**, 1337-1358, https://doi.org/10.1175/2011jhm1358.1.

Pan, B., K. Hsu, A. AghaKouchak, S. Sorooshian, and W. Higgins, 2019: Precipitation Prediction Skill for the West Coast United States: From Short to Extended Range. *J. Climate*, **32**, 161–182, https://doi.org/10.1175/JCLI-D-18-0355.1

Pegion, K., and coauthors, 2019: The Subseasonal Experiment (SubX): A multi-model subseasonal prediction experiment. *Bull. Amer. Meteor. Soc*., **100**, 2043–2060, https://doi.org/10.1175/BAMS-D-18-0270.1

Pierce, D.W., D.R. Cayan, and B.L. Thrasher, 2014: Statistical Downscaling Using Localized Constructed Analogs (LOCA). *J. Hydrometeor.*, **15**, 2558–2585, https://doi.org/10.1175/JHM-D-14-0082.1

Reichle, R., and Q. Liu, 2014: Observation-corrected precipitation estimates in GEOS-5. *Tech. Memo. NASA/TM-2014-104606*, Vol. **35**, 18 pp., https://ntrs.nasa.gov/archive/nasa/casi.ntrs.nasa.gov/20150000725.pdf.

Ralph, F.M., P.J. Neiman, G. Wick, S. Gutman, M. Dettinger, D. Cayan, and A. White, 2006: Flooding on California's Russian River: Role of atmospheric rivers. *Geophys. Res. Lett*., **33**, L13801, https://doi.org/10.1029/2006GL026689.

Rienecker, M. M., and Coauthors, 2008: The GEOS-5 Data Assimilation System-Documentation of versions 5.0.1, 5.1.0, and 5.2.0. *Tech. Memo. NASA/TM-2008-104606*, Vol. **27**, 97 pp., https://ntrs.nasa.gov/archive/nasa/casi.ntrs.nasa.gov/20120011955.pdf.

Saha, S., and Coauthors, 2014: The NCEP Climate Forecast System version 2. *J. Climate*, **27**, 2185–2208, https://doi.org/10.1175/JCLI-D-12-00823.1.

Schick, S., O. Rössler, and R. Weingartner, 2019: An Evaluation of Model Output Statistics for Subseasonal Streamflow Forecasting in European Catchments. *J. Hydrometeor*., **20**, 1399–1416, https://doi.org/10.1175/JHM-D-18-0195.1

Shukla, S., and D.P. Lettenmaier, 2011: Seasonal hydrologic prediction in the United States: understanding the role of initial hydrologic conditions and seasonal climate forecast skill. *Hydrol. Earth Syst. Sci*., **15**, 3529–3538, https://doi.org/10.5194/hess-15-3529-2011

Su, L., Q. Cao, M. Xiao, D.P. Lettenmaier, D. Li, M. Barlage, and D.M. Mocko, 2020: Drought Variability and Trends over the Conterminous United States over the Past Century. (in prep.)

Sun, S., R. Bleck, S. G. Benjamin, B. W. Green, and G. A. Grell, 2018a: Subseasonal forecasting with an icosahedral, vertically quasi-Lagrangian coupled model. Part I: Model overview and evaluation of systematic errors. *Mon. Wea. Rev.*, **146**, 1601–1617, https://doi.org/10.1175/MWR-D-18-0006.1.

Sun, S., B. W. Green, R. Bleck, and S. G. Benjamin, 2018b: Subseasonal forecasting with an icosahedral, vertically quasi-Lagrangian coupled model. Part II: Probabilistic and deterministic forecast skill. *Mon. Wea. Rev.*, **146**, 1619–1639, https://doi.org/10.1175/MWR-D-18-0007.1.

USWRC, 1982: Guidelines for determining flood flow frequency. Bulletin 17B of the Hydrology Subcommittee, 183 pp., https://water.usgs.gov/osw/bulletin17b/dl_flow.pdf.

Vitart, F., and Coauthors, 2017: The Subseasonal to Seasonal (S2S) Prediction project database. *Bull. Amer. Meteor. Soc.*, **98**, 163–173, https://doi.org/10.1175/BAMS-D-16-0017.1.

Weigel, A.P., M.A. Liniger, and C. Appenzeller, 2007: The Discrete Brier and Ranked Probability Skill Scores. *Mon. Wea. Rev.*, **135**, 118–124, https://doi.org/10.1175/MWR3280.1

Wigmosta, M. S., L. W. Vail, and D. P. Lettenmaier, 1994: A distributed hydrology-vegetation model for complex terrain. *Water Resour. Res.*, **30**, 1665-1679, https://doi.org/10.1029/94WR00436.

Wood, A.W., and D.P. Lettenmaier, 2008: An ensemble approach for attribution of hydrologic prediction uncertainty. *Geophys. Res. Lett.*, **35**, L14401, https://doi.org/10.1029/2008GL034648.

——, E.P. Maurer, A. Kumar, and D.P. Lettenmaier, 2002: Long-range experimental hydrologic forecasting for the eastern United States. *J. Geophys. Res*., **107**, 4429, https://doi.org/10.1029/2001JD000659.

——, L.R. Leung, V. Sridhar, and D.P. Lettenmaier, 2004: Hydrologic implications of dynamical and statistical approaches to downscaling climate model outputs. *Clim. Change*, **62**, 189, https://doi.org/10.1023/B:CLIM.0000013685.99609.9e

Zhou, X., Y. Zhu, D. Hou, and D. Kleist, 2016: A comparison of perturbations from an ensemble transform and an ensemble Kalman filter for the NCEP Global Ensemble Forecast System. *Wea. Forecasting*, **31**, 2057–2074, https://doi.org/10.1175/WAF-D-16-0109.1.

Zhou, X., Y. Zhu, D. Hou, Y. Luo, J. Peng, and R. Wobus, 2017: Performance of the new NCEP Global Ensemble Forecast System in a parallel experiment. *Wea. Forecasting*, **32**, 1989–2004, https://doi.org/10.1175/WAF-D-17-0023.1.

Zhu, Y., and Coauthors, 2018: Towards the improvement of sub-seasonal prediction in the NCEP Global Ensemble Forecast System (GEFS). *J. Geophys. Res. Atmos*., **123**, 6732–6745, https://doi.org/10.1029/2018JD028506.

Chapter 5. Conclusions and recommendations for future work

5.1 Conclusions

In this dissertation, I have examined the effects of hydrologic initial conditions on atmospheric river (AR)-related floods and how they will change in a warming climate, using a transect of three rain-dominant watersheds along the U.S. Pacific Coast. The three watersheds are the Chehalis River basin in Washington State, the Russian River basin in Northern California, and the Santa Margarita River basin in Southern California. Due to their locations, floods in all three are dominated by the seasonal and geographic signatures of landfalling ARs. I applied the Distributed Hydrology-Soil-Vegetation Model (DHSVM) in each of the three basins to perform modeling experiments, which are the core of my analysis.

The first science question I posed in Chapter 1 is: "***What is the role of hydrologic initial conditions on the interaction between surface climate forcings associated with ARs and the flood response at river basin scale in the current (historical) climate?***" To address this question, Chapter 2 examines the role of antecedent soil moisture (ASM) on historical AR-related flooding in California's Russian River Basin during the period water years (WY) 1950-2017. It uses a combination of observation- and simulation-based analyses. It concludes that most extreme precipitation and flood events in this basin are associated with ARs. The runoff ratio during extreme precipitation events is much more strongly related to ASM than to storm total precipitation. Not all extreme precipitation events can lead to extreme discharge events of the same peaks over threshold (POT) category especially when ASM is low. Among the extreme discharge events, however,

the first 2-day storm precipitation has a greater effect on the peak flow than does ASM, but the effects of ASM on peak flow increase as drainage area increases.

Chapter 2 also examines trends in extreme precipitation and discharge over the past 68 years. There are no statistically significant trends in the magnitude of extreme precipitation, but there are weak downward trends in the magnitude of extreme discharge at some stream gauges. There are no statistically significant trends in ASM associated with extreme discharge events despite the fact that ASM is significantly correlated with peak flow; rather downward trends in extreme discharge are caused mostly by changes in the first two days of storm precipitation. Using a modeling experiment, I also find that the ASM for extreme discharge events in the Russian River basin is affected more by antecedent precipitation than by evapotranspiration and hence observed temperature increases have only weakly affected ASM over the past 68 years.

The second science question posed in Chapter 1 is: "***How will climate change impact the role of hydrologic initial conditions on AR-related floods along the U.S. West Coast?***" Chapter 3 addresses this question using model simulations driven by downscaled global climate model (GCM) forcings in each of the three basins. Chapter 2 shows that when ASM is wet or storm precipitation is sufficiently large, almost all extreme precipitation events lead to extreme discharge events in the same POT category. On the other hand, if the ASM is low, extreme precipitation may not lead to extreme discharge of the same POT category. However, ASM might decrease with projected decreases in pre-storm low to medium intensity precipitation, and increased evaporative demand associated with warming. Therefore, Chapter 3 examines whether future increases in storm precipitation could outweigh the effects of future changes in ASM and

thus enhance the connection of extreme precipitation and extreme discharge. It does so by examining changes in AR contributions to the largest 50, 100 and 150 extreme precipitation and discharge events in the three river basins in two periods, WY 1951-2000 and WY 2050-2099. In particular, it examines how ASM is likely to change in the future and how changes in ASM will affect the relationship of extreme precipitation and extreme discharge.

In Chapter 3, I find that the projected (consensus of 10 GCMs) fraction of AR-related extreme discharge events slightly decreases in the Chehalis basin. In the Russian River basin, this fraction increases, however, and more substantially so in the Santa Margarita basin. This is due to two effects. First, AR-related extreme precipitation events increase in frequency in the two California watersheds. Second, the relationship of AR-related extreme precipitation to extreme discharge is strengthened by projected increases in year-to-year volatility of annual precipitation in California. This increases the likelihood of concurrent occurrence of large storms and wet ASM conditions. In addition, the influence of changes in ASM on annual maximum floods (AMF) varies among individual events in three basins, but the relative effect of ASM on AMF given storm precipitation is projected to become slightly weaker (yet still significant) in the Russian River basin where storm precipitation is projected to increase most among three basins caused by ARs.

The third science question posed in Chapter 1 is: "***What is the subseasonal forecast skill (at 1-4 week lead times) of AR-related flooding in coastal Western U.S. watersheds?***" To address this question, in Chapter 4 I evaluate the performance of the NOAA/Climate Testbed Subseasonal Experiment (SubX) project reforecast forcings

(after statistical downscaling and bias correction) applied to subseasonal forecasts of AR-related flooding in the three watersheds during the period 1999-2016. I find that the SubX forecast skill for extreme discharge events drops quickly after week 1, during which there is relatively strong deterministic forecast skill. There is some probabilistic forecast skill in week 2, but negligible skill in weeks 3-4, especially for annual maximum floods (one event per year on average), notwithstanding some probabilistic skill for smaller floods in weeks 3-4.

Moreover, Chapter 4 shows that flood forecast skill is strongly influenced by ASM, with higher forecast skill when ASM is wet especially at shorter lead times and for lower flood thresholds. Via ensemble streamflow prediction (ESP) and reverse-ESP experiments, I show in Chapter 4 that that at short forecast lead times, ASM dominates streamflow forecast skill, while SubX forecast skill dominates at longer lead times. The length of lead times for which ASM has a stronger influence on flood forecast skill is longer in the northern-most (Chehalis) basin, and generally decreases north to south.

5.2 Recommendations for future work

This dissertation has demonstrated the important role of ASM in AR-related flooding in coastal Western U.S. watersheds under both current and future conditions. My focus has been on three rain-dominant watersheds that reflect AR landfall signatures. There are significant differences across basins, but they are generally systematic, and reflect the nature of storm and inter-storm characteristics. Future work could extend the analyses to other watersheds and to the regional scale, and in particular could consider the more complex set of factors that govern extreme runoff in snow-affected watersheds.

Chapter 3 examined projected changes in the magnitude of extreme (order 100-year) floods in the three basins. The full GCM ensemble (10 models) projects significant increases in extreme floods all three basins. However, the "Real-5" (five most realistic GCMs in terms of their reproduction of historical AR event statistics) project relatively small increases in the magnitude of the 100-year flood in the Chehalis River basin but greater increases in the Russian and Santa Margarita River basins than the full GCM ensemble. Understanding the nature and cause of across-GCM differences in AR projections and their implications for future flood risk along the Pacific Coast could be the subject of future studies. Across-GCM differences indicate that caution is needed in the selection of GCMs for future flood analyses. Furthermore, as the Phase 6 of Coupled Model Intercomparison Project (CMIP6) suite of models of becomes available, there will be interest in examining their implications for extreme flooding in the region.

Chapter 4 examined the subseasonal forecast skill of AR-related flooding in three watersheds using the NOAA's SubX data base. Future work of this chapter could examine the role of ASM in flood forecast for different AR categories. Ralph et al. (2019) introduced a scale based on AR duration and (integrated water vapor transport, IVT) intensity to categorize AR events and their potential hydrologic impacts at a given location. AR events can be categorized from Cat 1, which are primarily beneficial, to Cat 5, which are primarily hazardous. Corringham et al. (2019) showed that flood damages as in US dollars increased exponentially as AR category increased from Cat 2 to above, in the western U.S. during the period 1978-2017. Chapter 2 examined the role of ASM in historical AR flooding on California's Russian River Basin. I showed that low ASM was an offsetting factor for the three AR Cat 5 events that did not lead to major flooding

during 1980-2017. However, the role of ASM in subseasonal flood forecast for different AR categories needs to be further explored.

There is a great deal of future work to be explored on this topic, from both meteorological and hydrological aspects. Recent studies suggest promising windows of opportunity to improve meteorological prediction skill at subseasonal to seasonal time scales, by leveraging specific climate phenomena or conditions for a predictable signal above the weather noise, which should be examined in the context of ARs especially along the U.S. Pacific Coast (DeFlorio et al., 2019; Mariotti et al., 2020). To improve hydrologic prediction, on one hand, the ASM/runoff coupling strength in hydrologic models needs to be evaluated and improved in order to more accurately reflect observation-based relationships (Crow et al., 2018). On the other hand, the representation of hydrological initial conditions in models can be improved by integrating soil moisture measurements from in situ observations and satellite-based products. A diverse and multi-tiered observational network has been newly implemented in California (Hatchett et al., 2020). From that as well as the ongoing effort in observation and modeling, the extended-range forecast skill of AR-related flooding in coastal Western U.S. watersheds could be improved.

References

Corringham, T.W., F.M. Ralph, A. Gershunov, D.R. Cayan, and C.A. Talbot, 2019: Atmospheric Rivers Drive Flood Damages in the Western United States. *Sci. Adv*., **5**, 2375-2548, https://doi.org/10.1126/sciadv.aax4631

Crow, W. T., F. Chen, R.H. Reichle, Y. Xia, and Q. Liu, 2018: Exploiting soil moisture, precipitation, and streamflow observations to evaluate soil moisture/runoff coupling in land surface models. *Geophys. Res. Lett*., **45**, 4869-4878. https://doi.org/10.1029/2018GL077193

DeFlorio, M.J., and coauthors, 2019: Experimental Subseasonal-to-Seasonal (S2S) Forecasting of Atmospheric Rivers Over the Western United States. *J. Geophys. Res. Atmos*., **124**, https://doi.org/10.1029/2019JD031200

Hatchett, B.J., and Coauthors, 2020: Observations of an extreme atmospheric river storm with a diverse sensor network. *Earth and Space Science*, https://doi.org/10.1002/essoar.10502163.1

Mariotti, A., and Coauthors, 2020: Windows of Opportunity for Skillful Forecasts Subseasonal to Seasonal and Beyond. *Bull. Amer. Meteor. Soc*., https://doi.org/10.1175/BAMS-D-18-0326.1

Ralph, F.M., J.J. Rutz, J.M. Cordeira, M. Dettinger, M. Anderson, D. Reynolds, L.J. Schick, and C. Smallcomb, 2019: A Scale to Characterize the Strength and Impacts of Atmospheric Rivers. *Bull. Amer. Meteor. Soc.*, **100**, 269–289, https://doi.org/10.1175/BAMS-D-18-0023.1.

Appendix A

The Role of Hydrological Initial Conditions on Atmospheric River Floods in the Russian River Basin – Supplemental Material

This appendix provides supporting information for Chapter 2, which has been published in its current form in the *Journal of Hydrometeorology*. © American Meteorological Society. Used with permission.

Cao, Q., A. Mehran, F.M. Ralph, and D.P. Lettenmaier, 2019: The Role of Hydrological Initial Conditions on Atmospheric River Floods in the Russian River Basin. *J. Hydrometeor*., 20, 1667–1686, https://doi.org/10.1175/JHM-D-19-0030.1

A.1 Tables

Table A1. Statistics of AR-related and Non AR-related POT extreme precipitation and discharge events at six USGS stream gauges

POT extreme precipitation events	Threshold [N/yr.]	USGS stream gauge					
		11461500	11461000	11462500	11463000	11464000	11467000
Max. daily precip. (AR/Non AR) [mm/day]	3	171/62	161/60	174/61	172/63	164/78	156/81
Median daily precip. (AR/Non AR) [mm/day]	1	78/nan	85/nan	80/nan	82/nan	85/76	88/74
	2	64/53	69/56	68/55	67/58	71/64	71/61
	3	56/46	62/51	58/49	61/52	63/56	64/59
	5	47/39	51/42	49/39	52/41	53/43	53/41
	7	41/33	45/35	43/45	45/35	48/35	48/35
POT extreme discharge events	Threshold [N/yr.]	USGS stream gauge					
		11461500	11461000	11462500	11463000	11464000	11467000
Max. daily streamflow (AR/Non AR) [cms]	3	354/132	377/139	1318/319	1538/430	2197/640	3148/1719
Median daily streamflow (AR/Non AR) [cms]	1	121/106	138/125	410/317	564/415	919/637	1616/1354
	2	95/77	101/105	297/256	402/355	648/510	1115/935
	3	81/56	81/57	237/164	321/237	524/374	880/654
	5	57/39	61/35	182/101	259/129	402/198	621/336
	7	46/26	50/21	148/62	215/74	328/120	495/158

Note: N is the number of events.

Table A2. Trend of precipitation, discharge, and antecedent soil moisture in POT events under different temperature scenarios

POT Threshold [N/yr.]	USGS gauge	Extreme precipitation events	Extreme discharge events							
		Daily precipitation	SP_{2d}	Historical temperature			T1950		T2017	
				Obs. PF	Sim. PF	ASM	Sim. PF	ASM	Sim. PF	ASM
		[mm/yr.]	[mm/yr.]	[mm/yr.]		[$\times 10^{-2}$%/yr.]	[mm/yr.]	[$\times 10^{-2}$%/yr.]	[mm/yr.]	[$\times 10^{-2}$%/yr.]
1	11461500	0.00	**-0.24†**	-0.03	0.01	-0.08	0.00	-0.23	0.00	-0.03
	11461000	0.04	**-0.44***	**-0.12***	**-0.23†**	-0.98	**-0.24†**	-1.02	**-0.23†**	-0.88
	11462500	0.04	**-0.27***	-0.04	-0.00	0.96	-0.01	0.87	-0.01	0.93
	11463000	-0.08	**-0.22†**	-0.06	-0.08	0.12	-0.09	0.03	-0.08	0.10
	11464000	-0.02	**-0.27***	-0.10	-0.11	0.92	-0.11	1.05	-0.11	0.95
	11467000	-0.01	**-0.29***	-0.02	-0.06	2.60	-0.05	2.65	-0.05	2.69
2	11461500	-0.07	**-0.21****	**-0.12***	**-0.12†**	-0.36	**-0.13†**	-0.47	**-0.12†**	-0.38
	11461000	**-0.10†**	**-0.19†**	**-0.11†**	-0.09	0.51	-0.09	0.46	-0.09	0.49
	11462500	-0.04	**-0.21***	**-0.09***	**-0.12***	-0.78	**-0.13***	-0.90	**-0.12***	-0.83
	11463000	-0.07	**-0.20***	**-0.09***	**-0.13***	-1.20	**-0.13***	-1.15	**-0.13***	-1.24
	11464000	-0.04	**-0.25****	-0.03	**-0.10†**	0.57	-0.09	0.66	**-0.10†**	0.72
	11467000	-0.03	**-0.24****	-0.06	-0.08	1.82	-0.07	1.88	-0.08	1.94
3	11461500	-0.01	**-0.12***	0.00	-0.05	-0.06	-0.06	-0.12	-0.06	-0.08
	11461000	-0.06	-0.10	-0.02	-0.04	0.02	-0.05	-0.01	-0.05	-0.04
	11462500	-0.05	**-0.14***	-0.05	**-0.11***	**-1.47†**	**-0.11****	**-1.57†**	**-0.11****	**-1.46†**
	11463000	-0.03	**-0.13†**	-0.01	-0.05	-0.34	-0.05	-0.38	-0.06	-0.33
	11464000	-0.03	**-0.19****	-0.05	**-0.09***	-0.18	**-0.08†**	-0.12	**-0.09†**	-0.09
	11467000	-0.01	**-0.14†**	-0.02	-0.02	1.45	-0.02	1.58	-0.02	1.58

Note: N is the number of events, SP_{2d} is the first 2-day storm precipitation, PF is the peak daily flow, and ASM is antecedent soil moisture. The trend with $p<=0.01$ is marked with "**"; the trend with $0.01<p<=0.05$ is marked with "*"; the trend with $0.05<p<=0.1$ is marked with "†". Trends with $p<0.1$ are marked in bold font.

Table A3. Simulated ASM of AR Cat 5 events during Jan 1980-Apr 2017. The first 5 columns are from Table 5 in Ralph et al. (2019).

AR Cat 5 event No.	Year	Start time [UTC]	Duration [h]	Fllod stage (32 ft) at Guerneville?	Simulated ASM [volumetric water content expressed as %]	Simulated ASM percentile in historical POT_{N3} precipitation events during water years 1950-2017 [%]
1	1980	01/11-06:00	78	Yes (37 ft)	35.5	92.8
2	1983	11/09-06:00	48	-	28.1	30.3
3	1986	02/14-00:00	144	Yes (49 ft)	34.9	89.7
4	1991	03/02-12:00	60	Yes (33 ft)	34.7	87.7
5	1995	12/10-15:00	57	Yes (32 ft)	22.2	11.8
6	1996	11/16-18:00	78	No	16.7	5.1
7	1996/97	12/28-18:00	132	Yes (45 ft)	35.2	92.3
8	2010	10/23-09:00	51	No	12.2	3.6
9	2015	02/05-09:00	51	No	20.1	8.7
10	2017	01/07-06:00	48	Yes (40 ft)	32.3	62.6

A.2 Figures

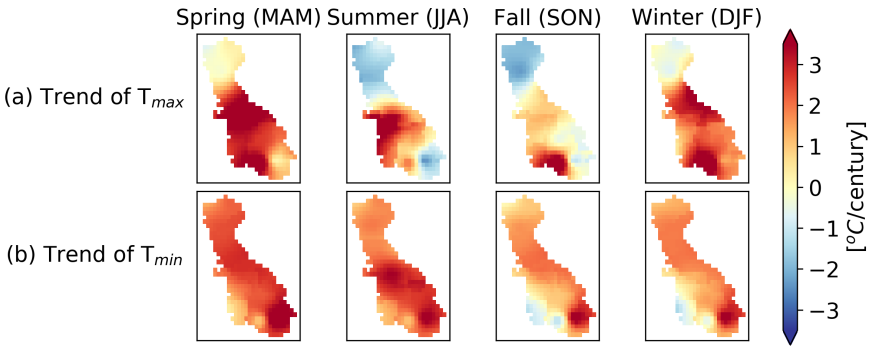

Figure A1. Seasonal trends in monthly average maximum temperature (T_{max}) and minimum temperature (T_{min}) for water years 1950-2017.

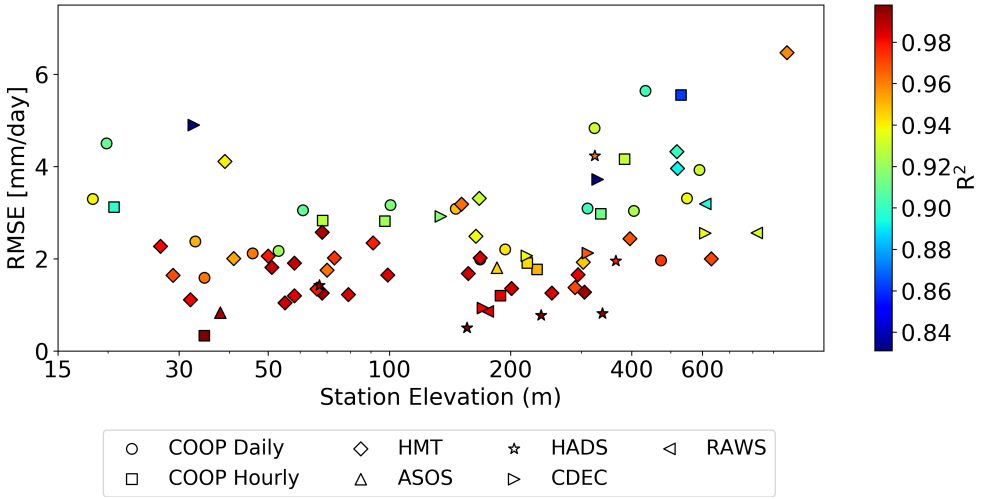

Figure A2. Assessment of the accuracy of gridding method by systematically removing individual stations within the Russian River Basin one at a time and comparing the derived grid precipitation estimate with each station removed with its available observation during the period water years 1950-2017.

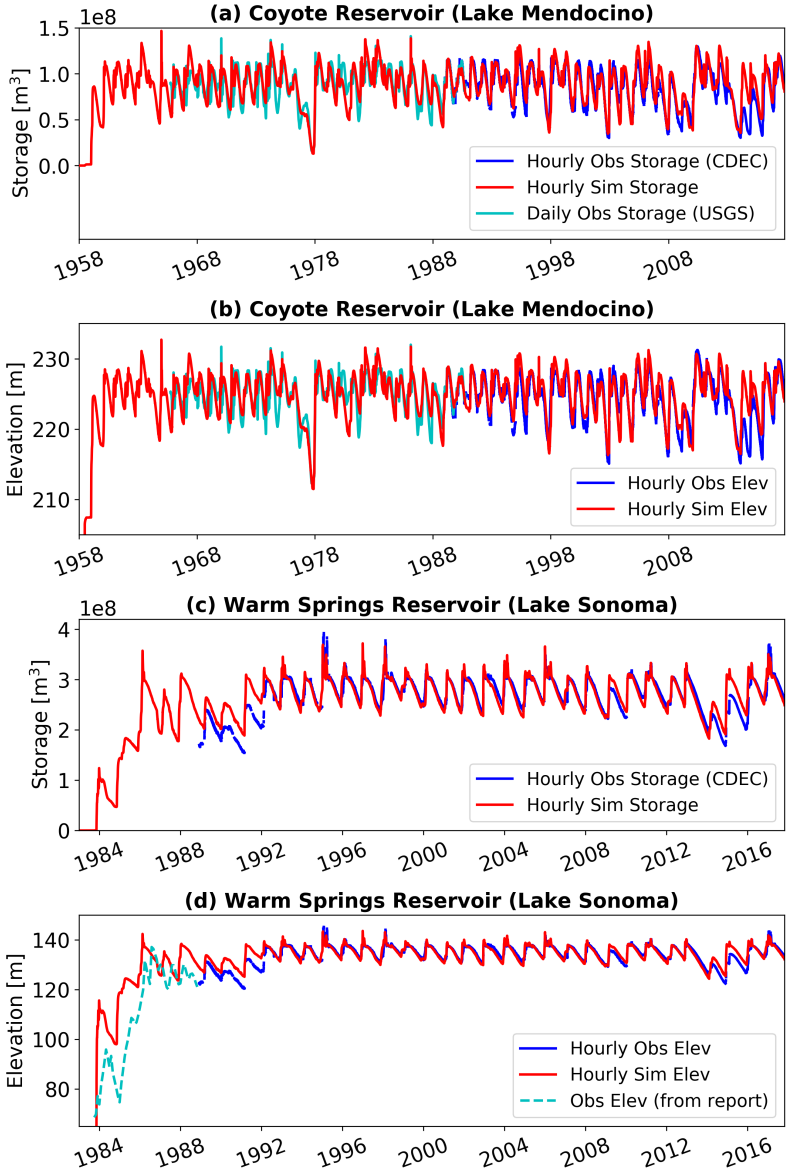

Figure A3. Comparison of observed and simulated hourly a) reservoir storage and b) lake elevation at the Coyote Reservoir (Lake Mendocino) during the period 1958 (the year when reservoir impoundment began) -2017. Same as c) and d) but for the Warm Springs Reservoir (Lake Sonoma) during the period 1983 (the year when reservoir impoundment began) -2017.

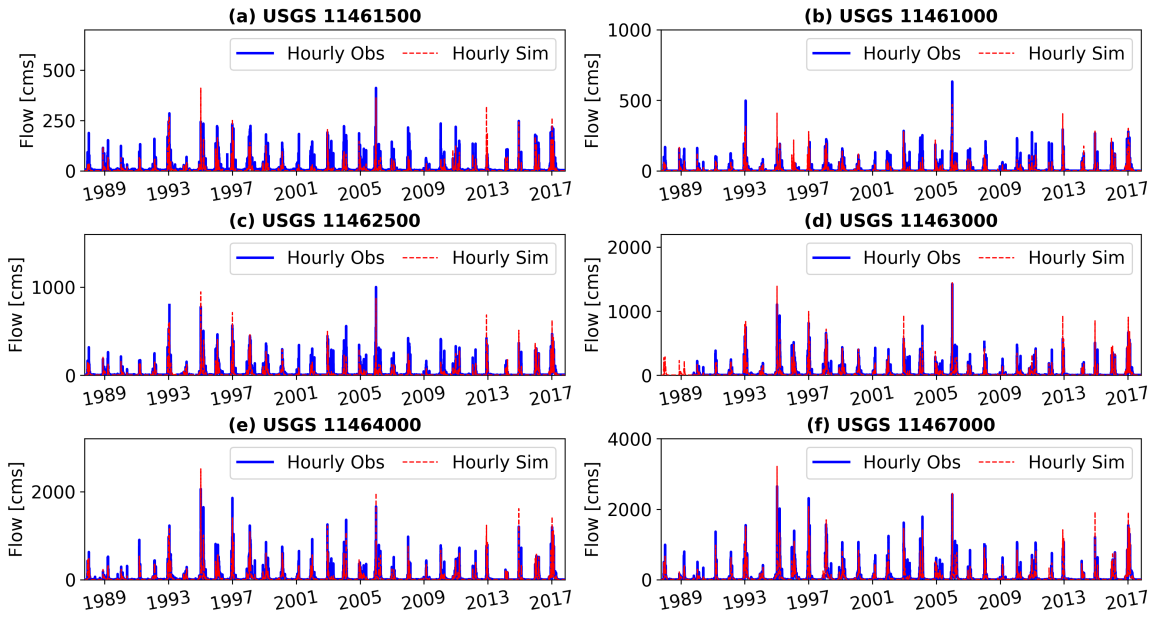

Figure A4. Comparison of observed and simulated hourly streamflow at six stream gauges during water years 1988-2017, the period when there were available observations for hourly streamflow. The simulations at gauges in c)-f) affected by reservoir regulations were from the model runs with the reservoir module activated.

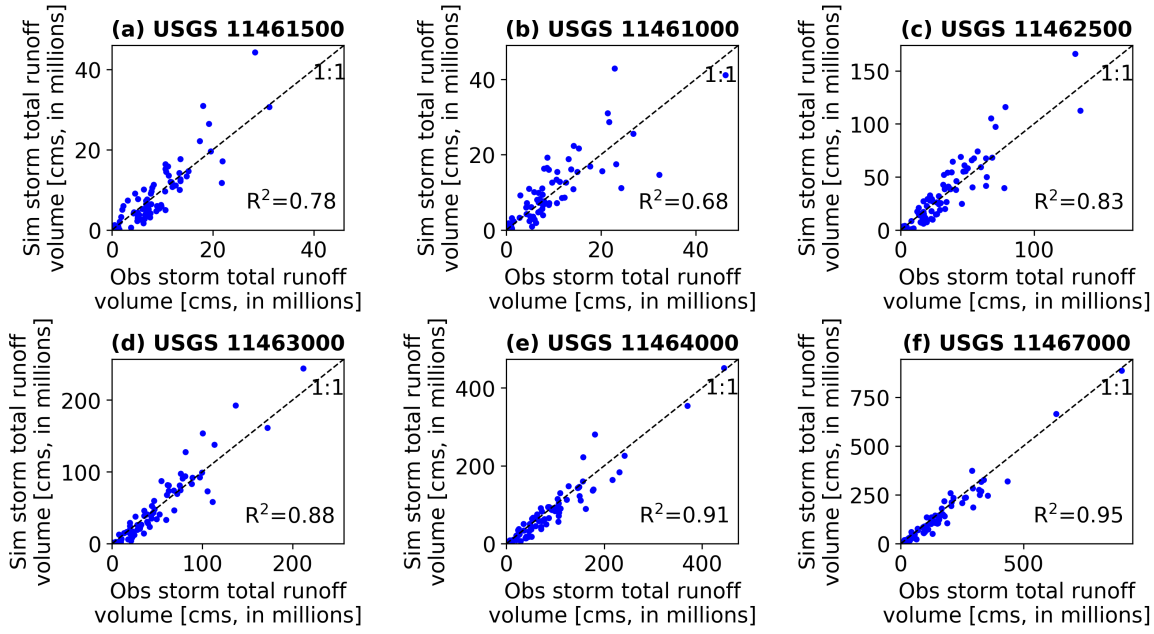

Figure A5. Comparison of observed, with reservoir effects removed in c)-f), and simulated storm total runoff volume calculated from hourly data at six stream gauges during water years 1988-2017, the period when there were available observations for hourly streamflow.

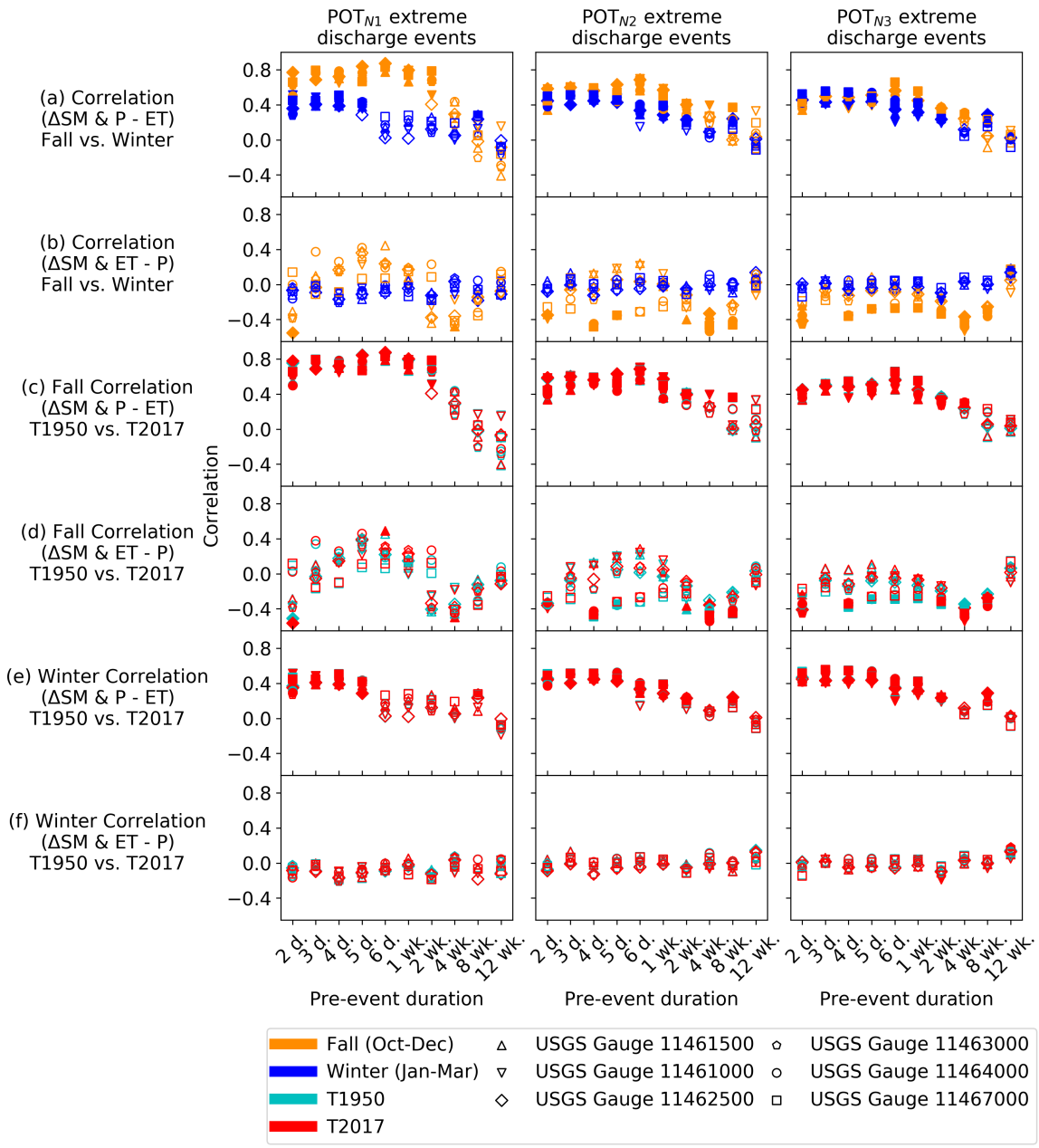

Figure A6. Effects of temperature and antecedent precipitation conditions on changes of antecedent soil moisture (ΔSM). a) Correlation between ΔSM and accumulated precipitation (P) given accumulated evapotranspiration (ET) under different pre-event duration in fall and winter months in POT_{N1}, POT_{N2} and POT_{N3} extreme discharge events. b) same as a) but for correlation between ΔSM and ET given P. c) correlation between ΔSM and P given ET in fall months under temperature scenarios T1950 and T2017. d) same as c) but for winter months. e) and f) are same as c) and d) but for correlation between ΔSM and ET given P. Correlations with p<=0.05 are marked with solid symbols.

Appendix B

Floods due to atmospheric rivers along the U.S. West Coast: The role of hydrological initial conditions in a warming climate – Supplemental Material

This appendix provides supporting information for Chapter 3, which has been submitted in its current form to the *Journal of Hydrometeorology* and now in revision.

Cao, Q., A. Gershunov, T. Shulgina, F.M. Ralph, N. Sun, and D.P. Lettenmaier, 2020: Floods due to atmospheric rivers along the U.S. West Coast: The role of hydrological initial conditions in a warming climate, *Journal of Hydrometeorology*, (in revision).

B.1 Changes in AMF based on the Real-5 GCMs (refers to main text Section 3.5)

We examined the role of ARs in projected changing magnitudes of AMF events, by comparing the ensemble mean of AMF events based on the Real-5 GCMs, with those based on the other 5 GCMs (see Figure S8). There were only small differences in the AMF projections, as well as storm precipitation of AMF events, for the two groups of GCMs in the Chehalis basin. In contrast, there was a distinct difference in the AMF series (for upper three quartiles) in the Russian basin, with increases of 39% and 19% (on average) respectively as indicated by the Real-5 versus the other 5 GCMs. The former showed increases in both storm precipitation and ASM, by 36% and 2% respectively, while the latter showed that storm precipitation would increase by 15% but there would be essentially no change in ASM. In the Santa Margarita basin, storm precipitation of AMF events increased by 51% and 26% respectively for the Real-5 and the other 5 GCMs, with most differences occurring in the upper quartile. However, changes in AMF were mostly different in the third quartile, increasing by 67% and 30% (on average)

respectively, possibly partly due to decreases in ASM, by 2% and 1% (on average) respectively for the Real-5 and the other 5 GCMs.

The difference in projected changes in the AMF series as indicated by the Real-5 and the other 5 GCMs was most distinct in the Russian basin (in terms of its overall increases despite that the percentage of changes was larger on average in the Santa Margarita basin). This is consistent with the precipitation analysis in Gershunov et al. (2019) who showed that the Real-5 GCMs had not only stronger increases in AR-related extreme precipitation but also weaker decreases in non-AR-related precipitation in comparison with other GCMs, particularly over Northern California.

B.2 Tables

Table B1. List of GCMs used in this study (refers to main text Section 3.3.2.2)

Category	CMIP5 GCM	Latitude size	Longitude size
Real-5 GCMs	ACCESS1.0 (CalWat)	1.25°	1.875°
	ACCESS1.3	1.25°	1.875°
	CanESM2 (CalWat)	2.7905°	2.8125°
	CNRM-CM5 (CalWat)	1.4007°	1.406°
	GFDL-CM3 (CalWat)	2°	2.5°
Other 5 GCMs	HadGEM2-CC (CalWat)	1.25°	1.875°
	Inmcm4	1.5°	2°
	IPSL-CM5A-MR	1.2676°	2.5°
	MIROC5 (CalWat)	1.4007°	1.40625°
	MRI-CGCM3	1.12149°	1.125°

Note: For California water resources planning, 10 out of 32 CMIP5 GCMs were selected by the Climate Action Team Research Working Group of the Fourth California's Climate Change Assessment in consultation with different scientist and organizations (e.g., Department of Water Resources, the California Energy Commission, Scripps Institution of Oceanography; Climate Change Technical Advisory Group, 2018; California Department of Water Resources, 2015), which are marked as "CalWat".

Table B2. Streamflow calibration statistics (refers to main text Section 3.4.1)

River basin	USGS Gauge	Location	DA [square miles]	Calibration period (1986-2000) Daily			Validation period (1971-1985) Daily		
				KGE	NRMSE	Relative bias [%]	KGE	NRMSE	Relative bias [%]
Chehalis	12020000	Doty	113	0.69	0.49	-10	0.88	0.43	1
	12025000	Newaukum	155	0.76	0.43	-6	0.65	0.48	-15
	12037400	Wynoocheee	155	0.74	0.49	23	0.79	0.45	19
	12035000	Satsop	299	0.89	0.38	8	0.88	0.39	8
	12027500	Grand Mound	895	0.80	0.38	-7	0.78	0.36	-9
	12031000	Porter	1294	0.88	0.36	10	0.94	0.30	3
Russian	11461500	Calpella	92	0.68	0.46	28	0.73	0.50	23
	11461000	Ukiah	100	0.87	0.40	10	0.88	0.44	-2
	11462500	Hopland	362	0.83	0.40	-6	0.81	0.40	-17
	11463000	Cloverdale	503	0.88	0.35	-3	0.85	0.39	-12
	11464000	Healdsburg	793	0.92	0.28	-2	0.89	0.33	-9
	11467000	Guerneville	1338	0.93	0.24	2	0.93	0.30	-5
Santa Margarita	11042400	Aguanga	131	0.53	0.66	-32	0.44	1.05	-23
	11043000	Temecula	222	0.74	0.68	-13	0.65	0.79	11
	11046000	Ysidora	723	0.67	0.67	3	0.43	0.82	24

Table B3. Statistics of POT extreme precipitation events in three river basins (refers to main text Section 3.4.2.1 and Section 3.4.3.1; see Section 3.3.3 in text for definitions of POT events)

River basin	Period	Model	AR-related fraction [%]			POT threshold [mm]		
			POT_{N1P}	POT_{N2P}	POT_{N3P}	POT_{N1P}	POT_{N2P}	POT_{N3P}
Chehalis	WY 1951-2000	All GCMs	60	55	51	58	48	41
		Real-5 GCMs	68	61	56	58	49	42
		SIO-R1	74	66	60	58	48	41
	WY 2050-2099	All GCMs	58	54	50	66	55	47
		Real-5 GCMs	66	60	57	67	55	47
	Change	All GCMs	-2	-1	-1	8	7	6
		Real-5 GCMs	-2	-1	1	9	6	5
Russian	WY 1951-2000	All GCMs	75	70	65	65	49	41
		Real-5 GCMs	78	71	67	65	48	40
		SIO-R1	98	98	95	65	48	40
	WY 2050-2099	All GCMs	77	72	65	77	58	46
		Real-5 GCMs	80	76	68	80	60	47
	Change	All GCMs	2	2	0	12	9	5
		Real-5 GCMs	2	5	1	15	12	7
Santa Margarita	WY 1951-2000	All GCMs	41	36	31	21	15	12
		Real-5 GCMs	38	33	27	21	14	12
		SIO-R1	78	65	60	20	15	11
	WY 2050-2099	All GCMs	46	41	38	24	16	12
		Real-5 GCMs	50	44	41	25	17	13
	Change	All GCMs	5	5	7	3	1	0
		Real-5 GCMs	12	11	14	4	3	1

Note: The AR catalog of Gershunov et al. (2017) is denoted as "SIO-R1", which is based on the NCEP/NCAR reanalysis.

Table B4. Statistics of POT extreme discharge events and annual maximum flow (AMF) events in three river basins (refers to main text Section 3.4.2.1 and Section 3.4.3.1; see Section 3.3.3 in text for definitions of POT events)

River basin	Period	Model	AR-related fraction [%]				POT threshold [cms]		
			AMF	POT_{N1D}	POT_{N2D}	POT_{N3D}	POT_{N1D}	POT_{N2D}	POT_{N3D}
Chehalis	WY 1951-2000	All GCMs	51	58	52	47	1624	1254	1011
		Real-5 GCMs	60	67	59	53	1601	1248	1015
		SIO-R1	70	80	70	66	1758	1313	1070
	WY 2050-2099	All GCMs	51	58	50	45	1980	1511	1210
		Real-5 GCMs	57	66	57	52	2002	1486	1201
	Change	All GCMs	0	0	-2	-2	356	257	199
		Real-5 GCMs	-3	-1	-2	-1	401	238	186
Russian	WY 1951-2000	All GCMs	59	67	56	48	1341	775	446
		Real-5 GCMs	64	70	60	51	1314	747	420
		SIO-R1	86	98	88	85	1203	688	358
	WY 2050-2099	All GCMs	63	71	62	54	1820	1059	563
		Real-5 GCMs	67	72	63	56	1960	1115	604
	Change	All GCMs	4	4	6	6	479	284	117
		Real-5 GCMs	3	2	3	5	646	368	184
Santa Margarita	WY 1951-2000	All GCMs	30	34	29	25	38	8	3
		Real-5 GCMs	27	28	23	20	34	8	3
		SIO-R1	60	72	59	51	29	8	3
	WY 2050-2099	All GCMs	38	43	37	33	54	11	3
		Real-5 GCMs	38	47	39	34	58	11	3
	Change	All GCMs	8	9	8	8	16	3	0
		Real-5 GCMs	11	19	16	14	24	3	0

Note: The AR catalog of Gershunov et al. (2017) is denoted as "SIO-R1", which is based on the NCEP/NCAR reanalysis.

Table B5. Fraction [%] of AR-related POT extreme precipitation events, POT extreme discharge events, and annual maximum flow (AMF) events in three river basins based on the ensemble average of Real-5 GCMs (refers to main text Section 3.4.3.1; see Section 3.3.3 in text for definitions of POT events)

River Basin	Period	Extreme precipitation events			Extreme discharge events			
		POT_{N1P}	POT_{N2P}	POT_{N3P}	POT_{N1D}	POT_{N2D}	POT_{N3D}	AMF
Chehalis	WY 1951-2000	68	61	56	67	59	53	60
	WY 2050-2099	66	60	57	66	57	52	57
	Change	-2	-1	1	-1	-2	-1	-3
Russian	WY 1951-2000	78	71	67	70	60	51	64
	WY 2050-2099	80	76	68	72	63	56	67
	Change	2	5	1	2	3	5	3
Santa Margarita	WY 1951-2000	38	33	27	28	23	20	27
	WY 2050-2099	50	44	41	47	39	34	38
	Change	12	11	14	19	16	14	11

B.2 Figures

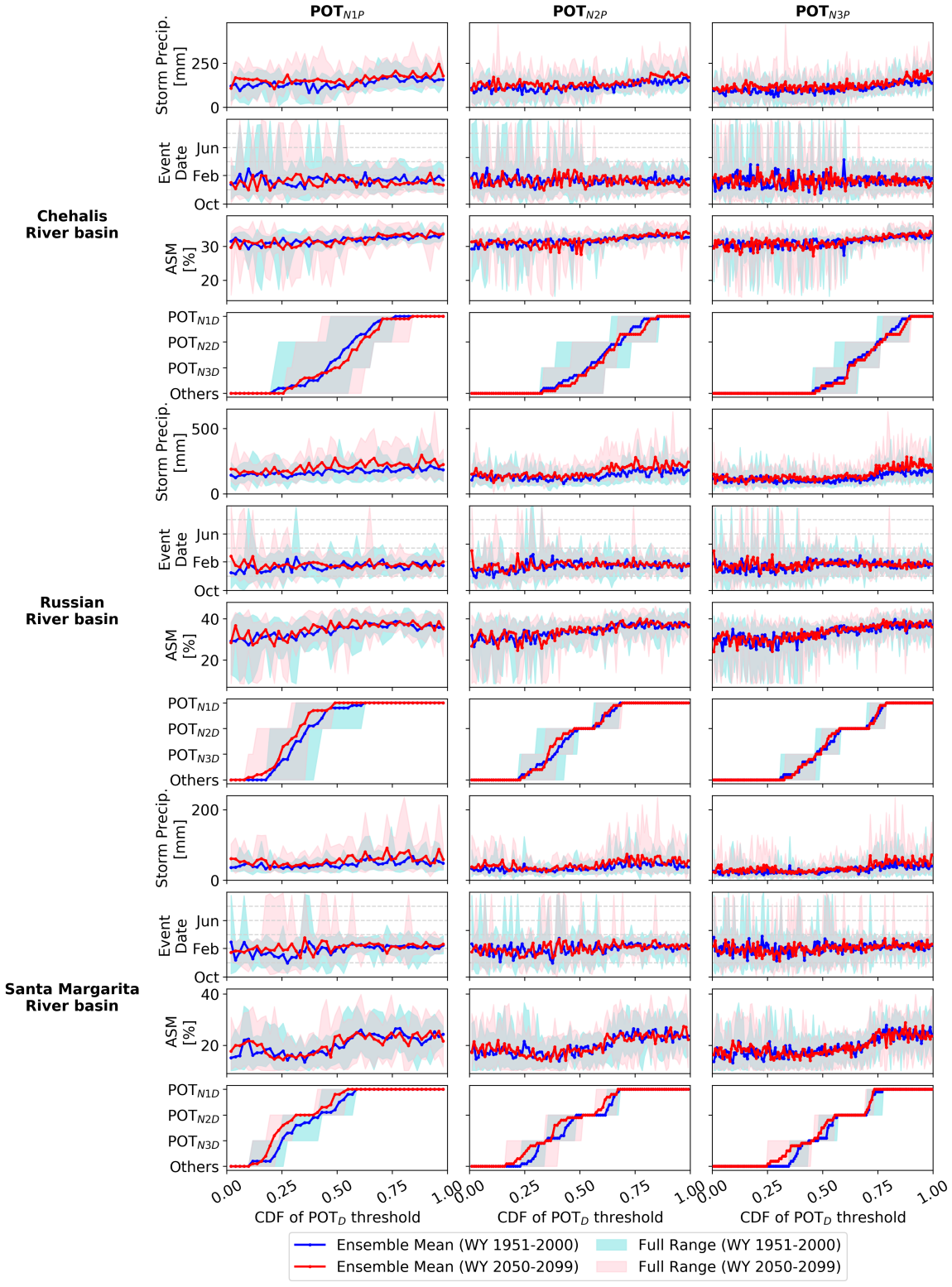

Figure B1. Same as Figure 7 but sorted by the CDF of the POT extreme discharge threshold during extreme precipitation events (refers to main text Section 3.4.3.2b).

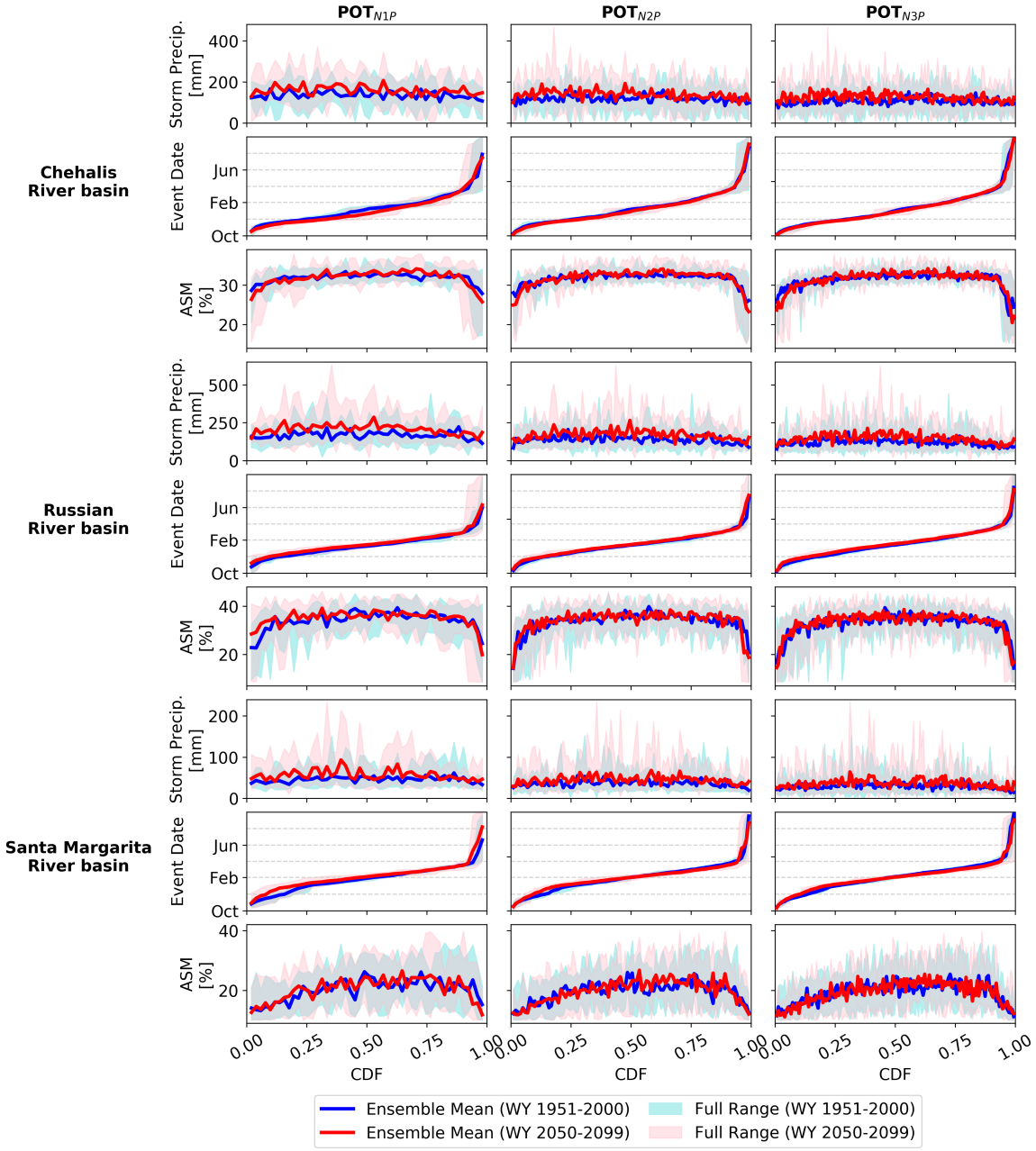

Figure B2. Same as Figure 3.7 but sorted by the CDF of the event occurrence dates (refers to main text Section 3.4.3.2b).

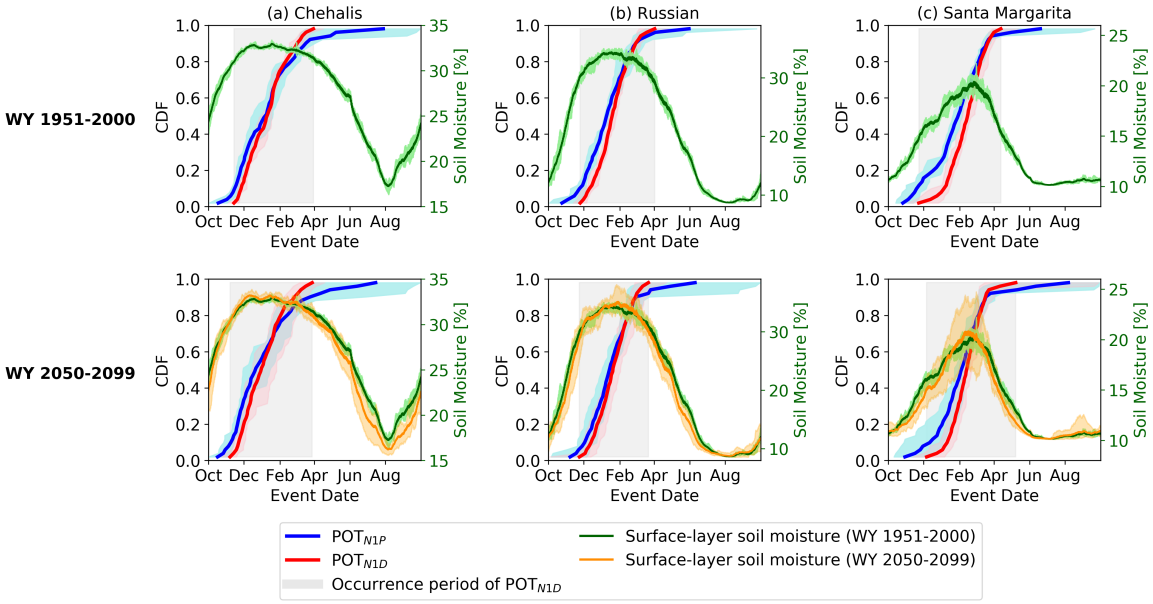

Figure B3. CDF of occurrence dates of POT_{N1P} and POT_{N1D} events (see Section 3.3.3 in text for definitions of POT_{N1P} and POT_{N1D}). The dark green line and orange line show the ensemble mean of seasonal cycle of the surface-layer soil moisture averaged over WY 1951-2000 and 2050-2099, respectively. The shadows of lines show the range of full GCM ensemble (refers to main text Section 3.4.3.2b).

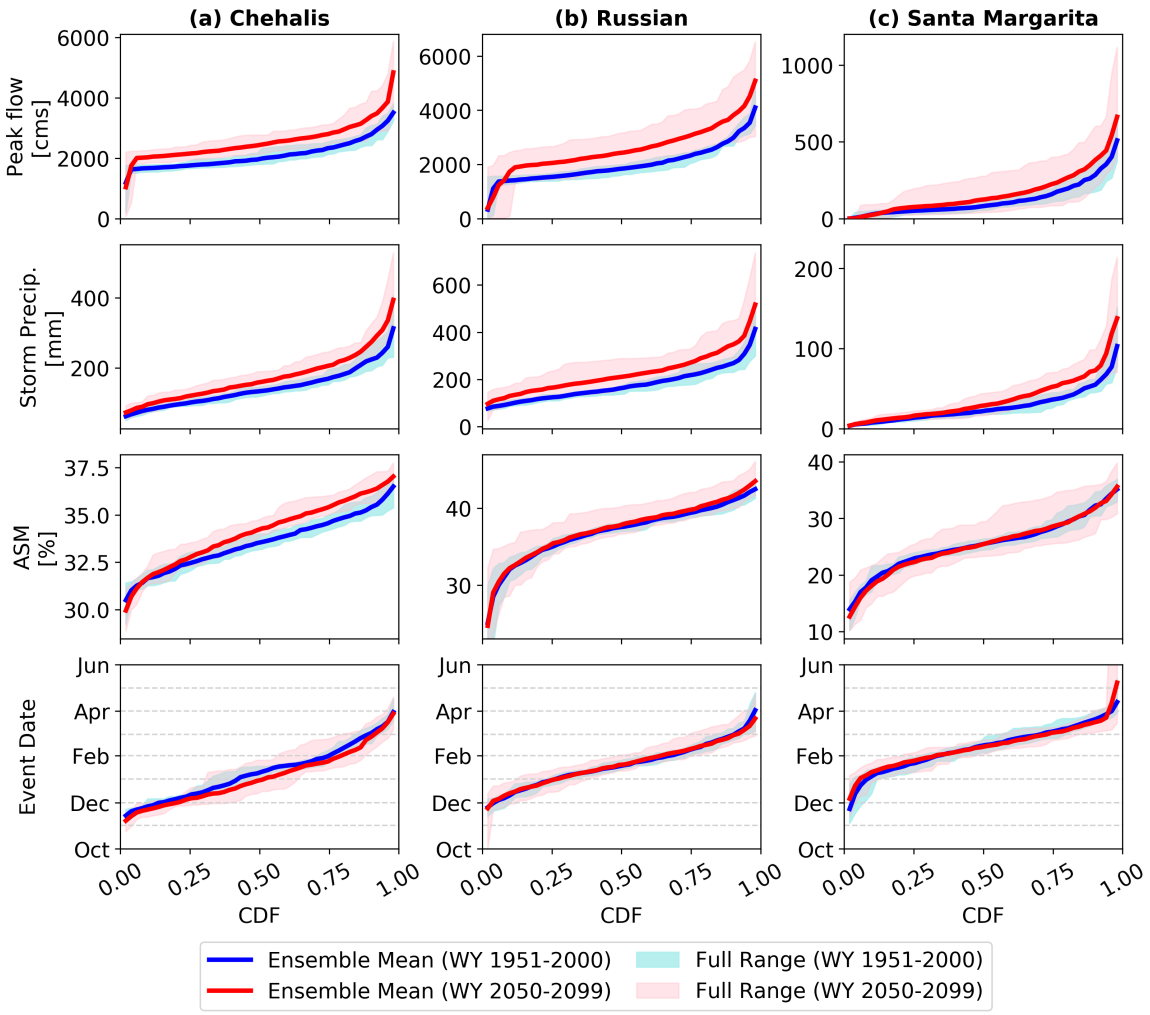

Figure B4. Same as Figure 3.8 but for POT_{N1D} events (refers to main text Section 3.4.3.3; see Section 3.3.3 in text for definitions of POT_{N1D}).

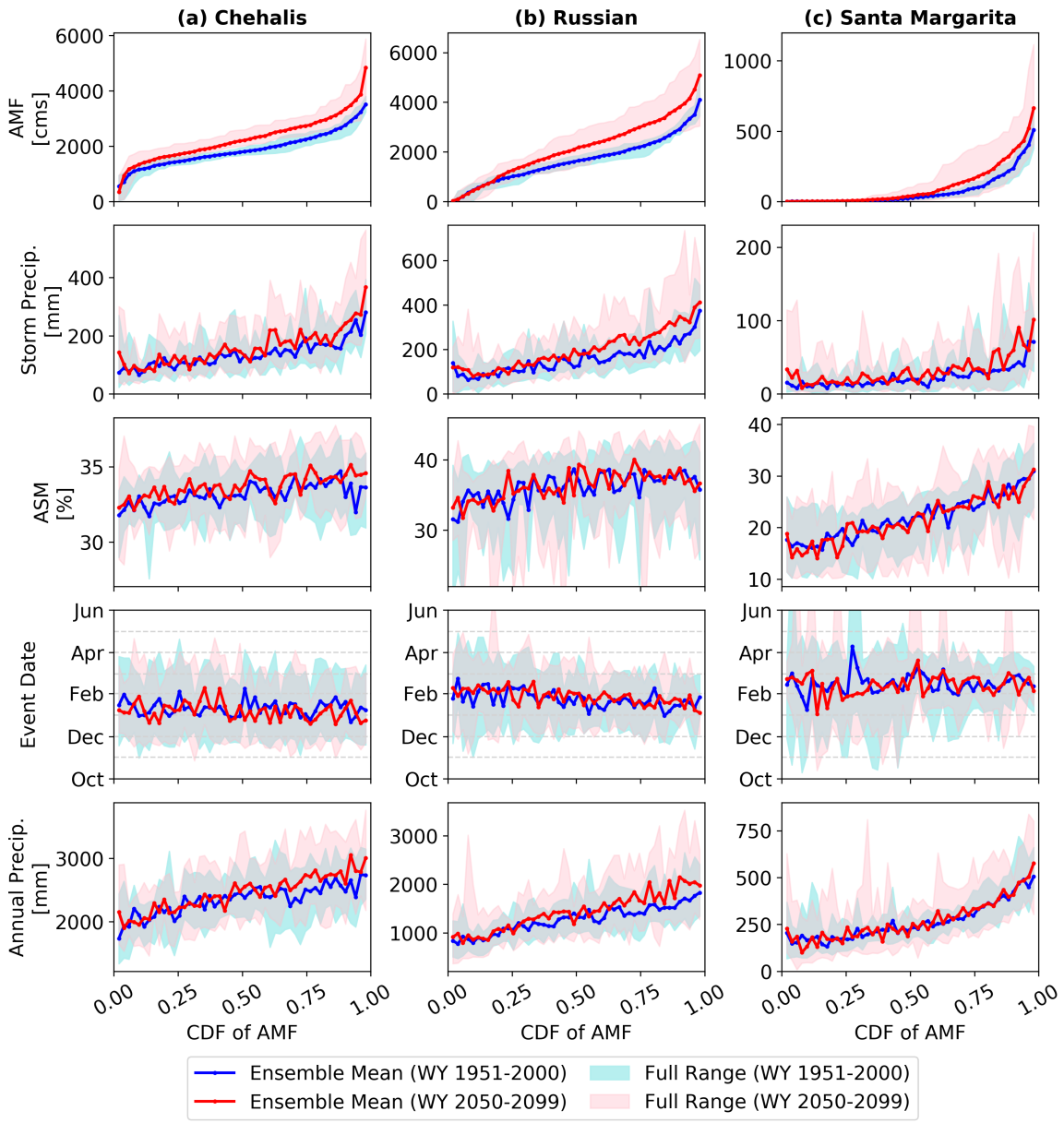

Figure B5. Same as Figure 3.8 but sorted by the CDF of AMF (refers to main text Section 3.4.3.3).

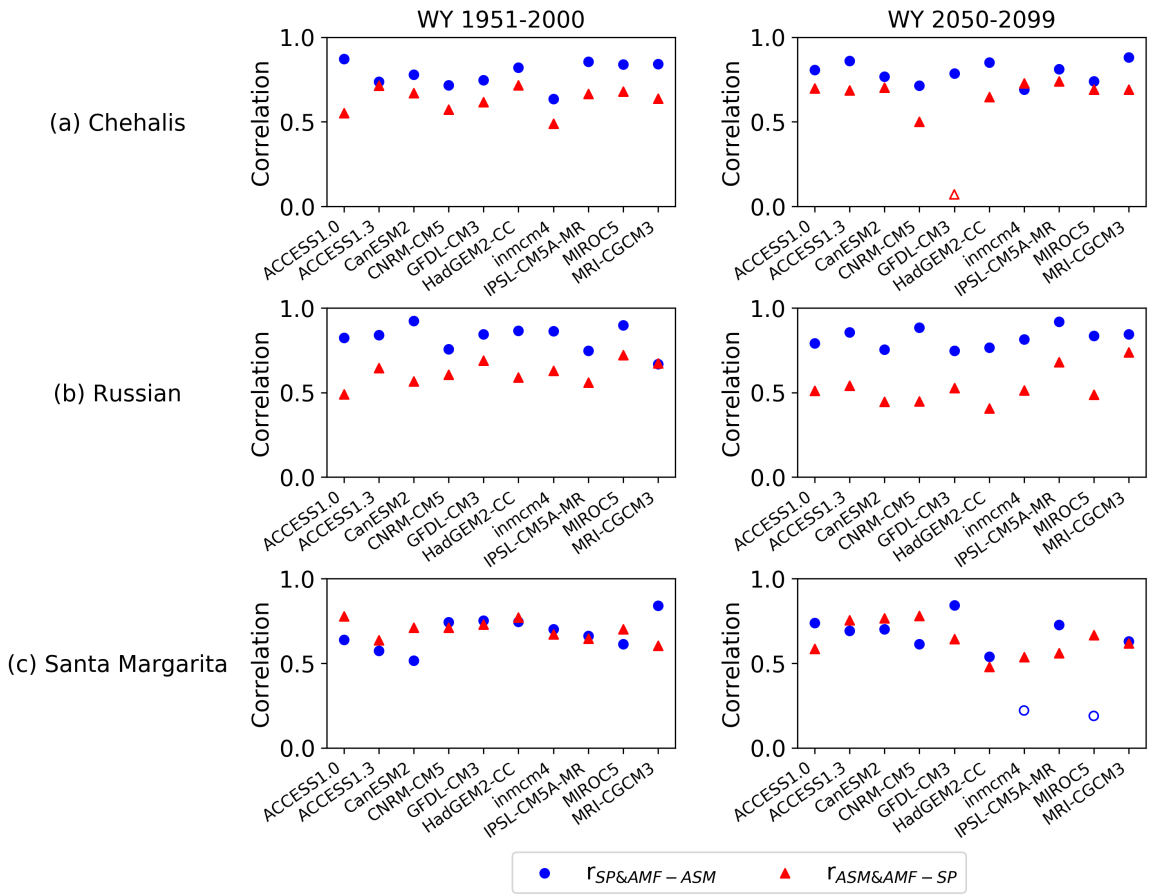

Figure B6. Correlation between accumulated storm precipitation (SP) and annual maximum flow (AMF) conditioned on antecedent soil moisture (ASM) ($r_{\text{SP\&AMF-ASM}}$), and correlation between ASM and AMF conditioned on accumulated storm precipitation ($r_{\text{ASM\&AMF-SP}}$) during WY 1951-2000 (left column) and WY 2050-2099 (right column). The correlations with p<=0.01 are shown as solid symbols (refers to main text Section 3.4.3.3).

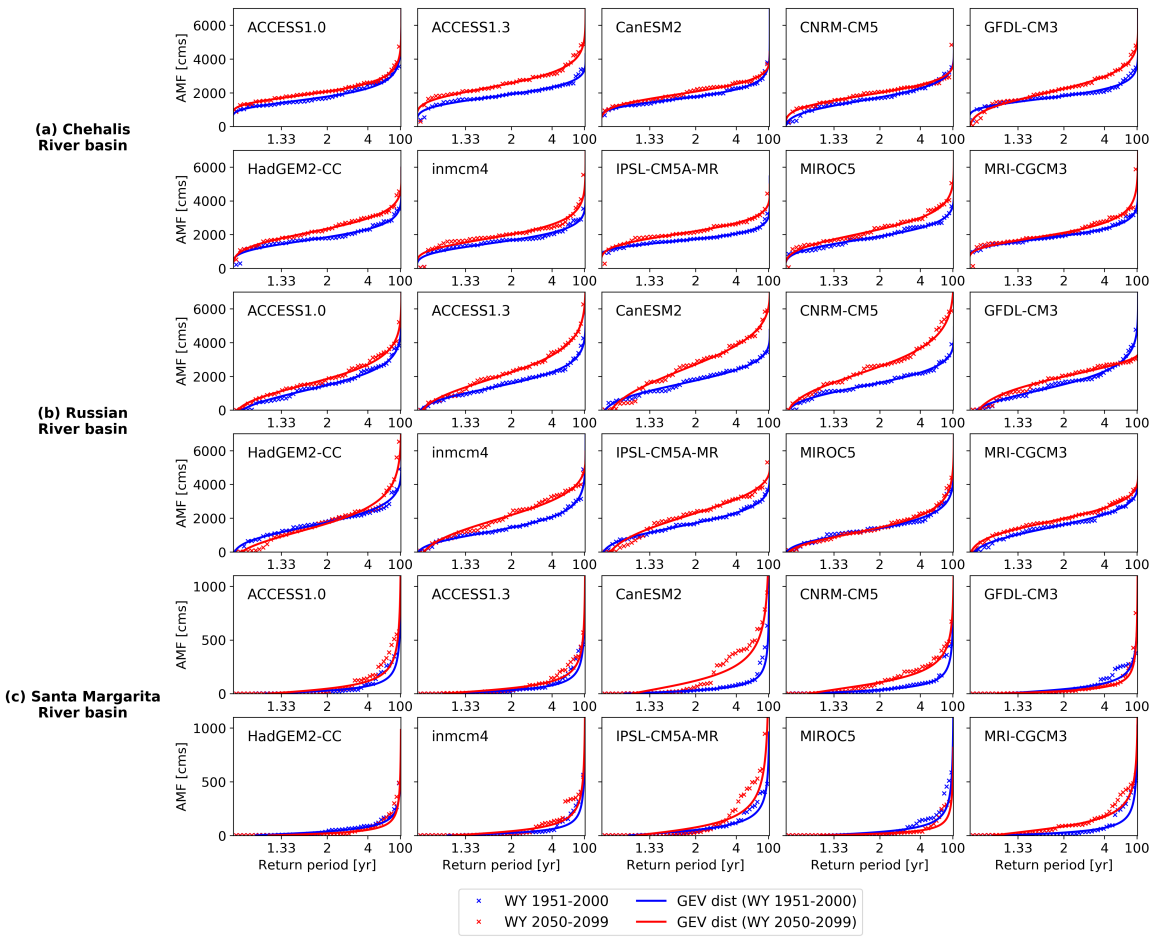

Figure B7. Fits of Generalized Extreme Value (GEV) distribution to the AMF from each GCM (refers to main text Section 3.5).

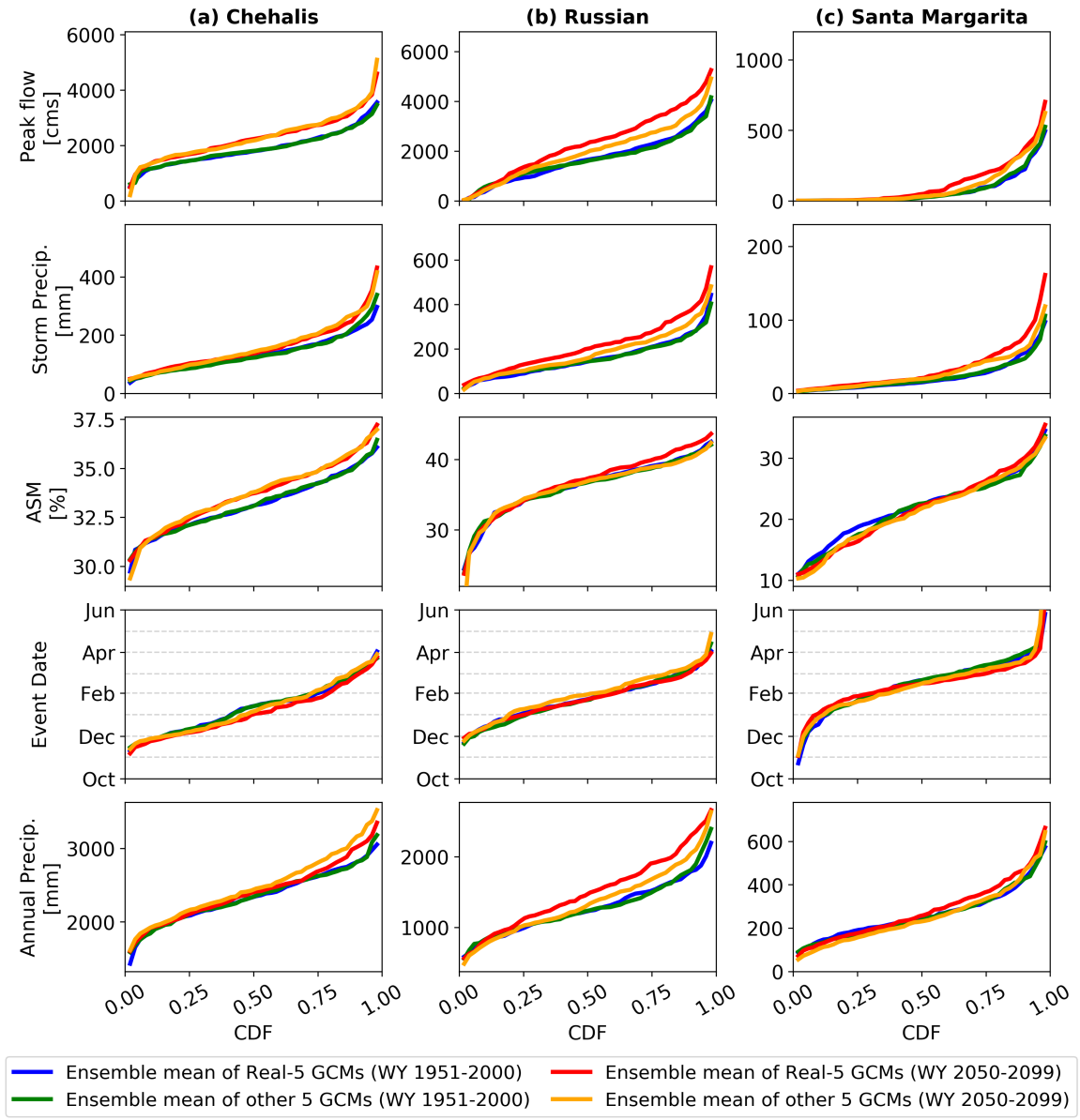

Figure B8. Same as Figure 3.8 but showing ensemble mean based for Real-5 GCMs and other 5 GCMs.

B.4 References

California Department of Water Resources, 2015: Perspectives and guidance for climate change analysis. California Department of Water Resources and Climate Change Technical Advisory Group. http://www.water.ca.gov/climatechange/docs/2015/Perspectives_Guidance_Climate_Change_Analysis.pdf

Climate Change Technical Advisory Group, 2018: Projected climate scenarios selected to represent a range of possible futures in California. A Document Developed by the Climate Action Team Research Working Group 1, California Energy Commission. http://docketpublic.energy.ca.gov/PublicDocuments/16-IEPR-04/TN215798_20170207T111409_Projected_Climate_Scenarios_Selected_to_Represent_a_Range_of_Po.pdf

Gershunov, A., and Coauthors, 2019: Precipitation regime change in Western North America: The role of Atmospheric Rivers. *Sci. Rep.*, **9**, 9944, https://doi.org/10.1038/s41598-019-46169-w.

Appendix C

Evaluation of the subseasonal forecast skill of atmospheric river floods in coastal Western U.S. watersheds – Supplemental Material

This appendix provides supporting information for Chapter 4, which will be submitted to the *Journal of Hydrometeorology* as

Cao, Q., Shraddhanand Shukla, Michael J. DeFlorio, F. Martin Ralph, and Dennis P. Lettenmaier, 2020: Evaluation of the subseasonal forecast skill of atmospheric river floods in coastal Western U.S. watersheds. *Journal of Hydrometeorology,* (*in prep*).

C.1 Tables

Table C1. Streamflow calibration statistics (refers to main text Section 4.4.2)

River basin	USGS Gauge	Location	DA [square miles]	Calibration period (1999-2007) Daily			Validation period (2008-2016) Daily		
				KGE	NRMSE	Relative bias [%]	KGE	NRMSE	Relative bias [%]
Chehalis	12020000	Doty	113	0.76	0.46	-3.2	0.79	0.48	16.6
	12025000	Newaukum	155	0.78	0.51	-7.4	0.80	0.49	-7.5
	12037400	Wynoocheee	155	0.87	0.45	7.3	0.73	0.46	-12.7
	12035000	Satsop	299	0.82	0.35	-11.9	0.74	0.42	-18.4
	12027500	Grand Mound	895	0.94	0.28	4.6	0.88	0.36	3.1
	12031000	Porter	1294	0.94	0.36	3.5	0.94	0.30	4.5
Russian	11461500	Calpella	92	0.61	0.65	-32.1	0.61	0.81	0.0
	11461000	Ukiah	100	0.75	0.47	20.2	0.65	0.62	27.7
	11462500	Hopland	362	0.79	0.46	4.6	0.45	0.72	34.9
	11463000	Cloverdale	503	0.73	0.47	8.0	0.54	0.63	30.2
	11464000	Healdsburg	793	0.90	0.36	0.4	0.68	0.48	22.0
	11467000	Guerneville	1338	0.89	0.27	9.2	0.61	0.43	31.1
Santa Margarita	11043000	Temecula	222	0.66	0.52	-30.7	0.55	0.61	-33.8
	11046000	Ysidora	723	0.56	0.57	27.4	0.79	0.46	17.9

C.2 Figures

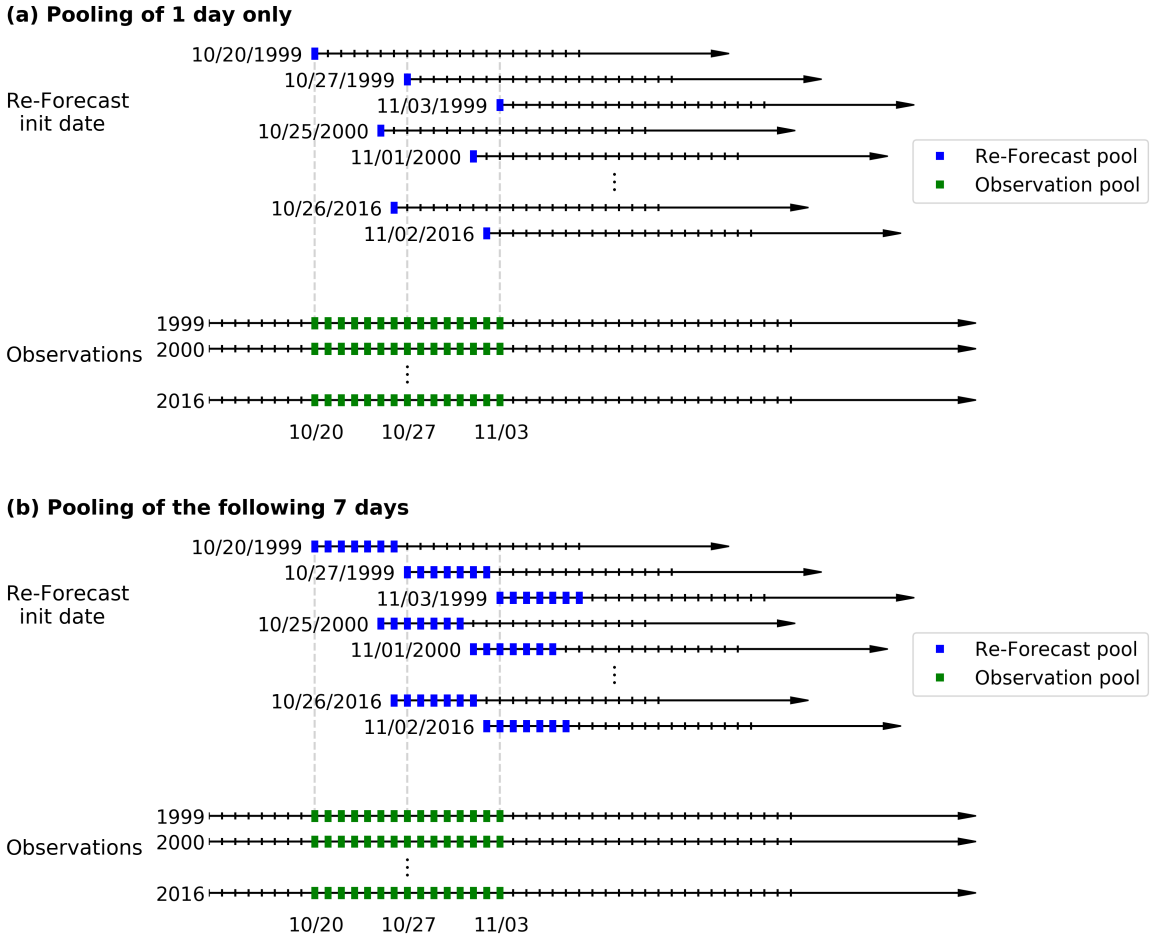

Figure C1. Schematic illustration of the pooling of days using 10/27/1999 as an example, including a) pooling of 1 day only, and b) pooling of the following 7 days.

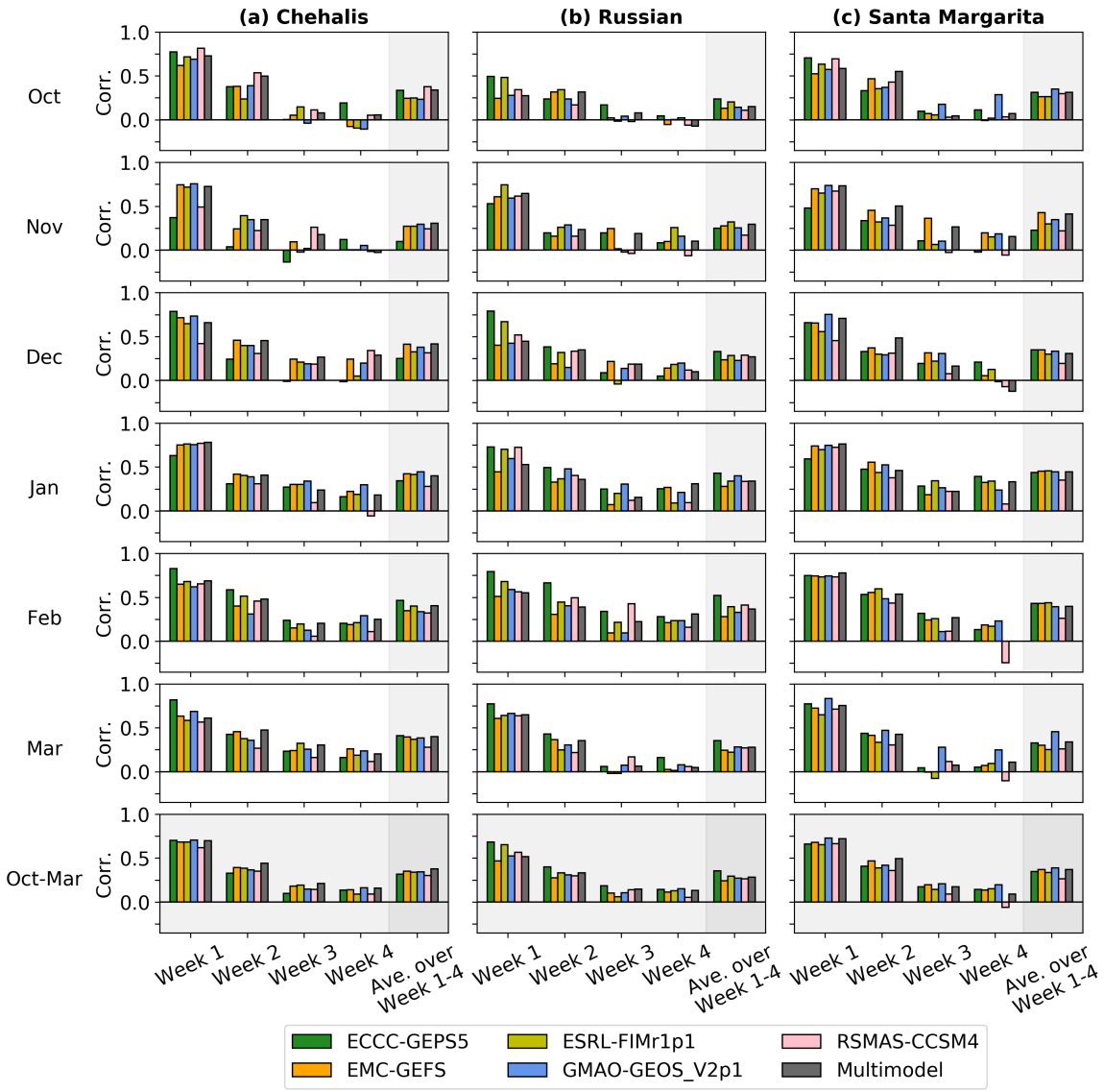

Figure C2. Maximum daily temperature (Tmin) skill (in correlation) of SubX models averaged over each basin and each month before bias correction.

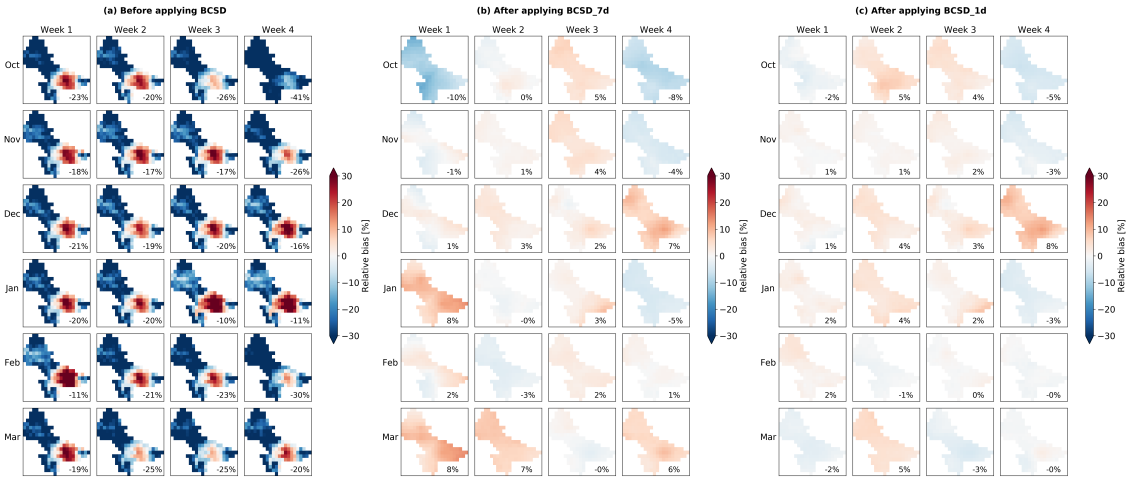

Figure C3. Spatial map of relative biases in precipitation of one SubX model, EMC-GEFS, a) before and after applying b) BCSD_7d and c) BCSD_1d in the Chehalis River basin.

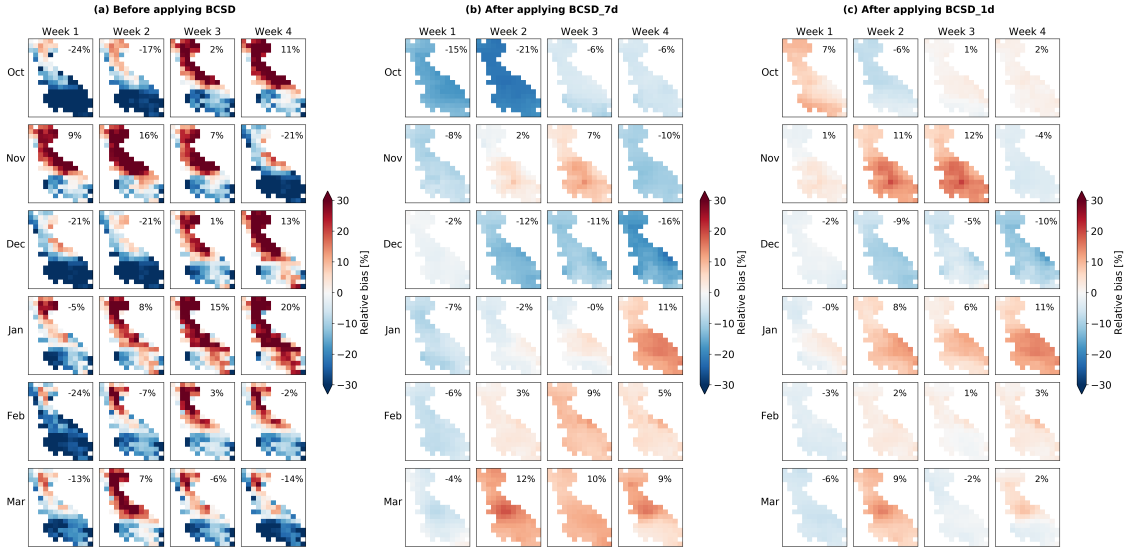

Figure C4. Spatial map of relative biases in precipitation of one SubX model, EMC-GEFS, a) before and after applying b) BCSD_7d and c) BCSD_1d in the Russian River basin.

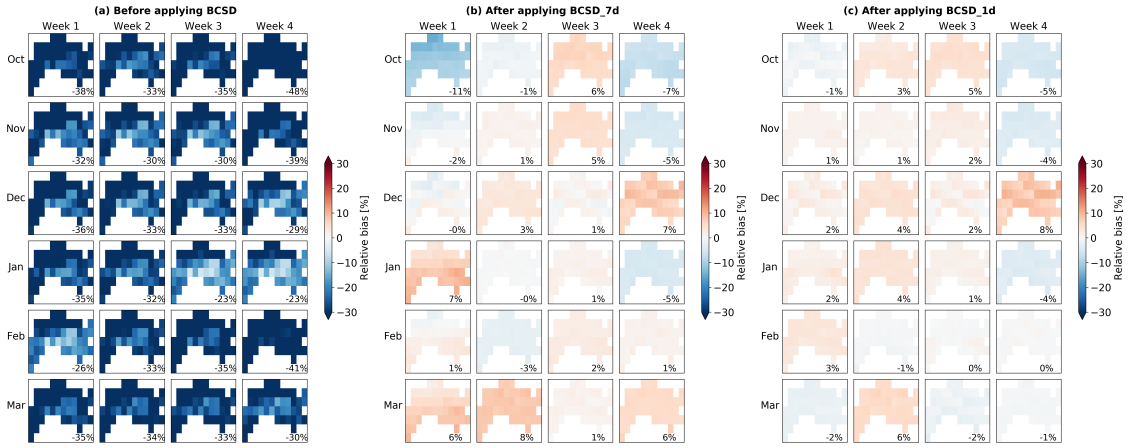

Figure C5. Spatial map of relative biases in precipitation of one SubX model, EMC-GEFS, a) before and after applying b) BCSD_7d and c) BCSD_1d in the Santa Margarita River basin.