Using hydrologic model ensembles to better understand the impact of climate change on the hydrology of large river basins

Oriana Shackell Chegwidden

A dissertation

submitted in partial fulfillment of the

requirements for the degree of

Doctor of Philosophy

University of Washington

2020

Reading Committee:

Bart Nijssen, Chair

David Rupp

Faisal Hossain

Program Authorized to Offer Degree:

Civil and Environmental Engineering

University of Washington

Using hydrologic model ensembles to better understand the impact of climate change on the hydrology of large river basins

Oriana Shackell Chegwidden

Chair of the Supervisory Committee:
Professor WOT Bart Nijssen
Civil and Environmental Engineering

Abstract

Whether at the scale of a small watershed or a large multinational basin, it has become common practice for water managers to use ensembles of projections to plan for hydrologic change. Better understanding these ensembles can help improve the design of future hydrologic modeling studies. In this dissertation I will describe three uses of hydroclimate ensembles to support water resource planning efforts. In Chapter 2 I present a large ensemble of hydrologic climate change projections for the Columbia River basin within the hydroclimatically diverse Pacific Northwestern United States and Canada (PNW). I show how methodological decisions in the modeling process variously affect the projections of change depending on hydroclimatic regime and metric of interest. In Chapter 3 I delve deeper into the PNW to examine the impactful metric of changes in floods, determining how dominant flood generating processes will evolve under climate change. I also calculate first-order sensitivities of high flows to changes in climate. In Chapter 4, I apply the lessons learned from the first two studies, conducted within the transboundary Columbia River basin, to transboundary rivers around the world. I present a study identifying hot spots of changes in water availability and hydropolitical risk for over 80 rivers (esp. transboundary rivers) around the world as projected by results from the Coupled Model Intercomparison Project Phase 6. Finally, I present how the findings from this dissertation can contribute to improved hydroclimate impacts assessments.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

I first want to acknowledge that I wrote this dissertation while on the occupied land of the coast Salish people.

I would like to thank my advisor Bart Nijssen for teaching me two key lessons. First, when you're building something it's helpful to first create a simple prototype. Once it works, you can always add bells and whistles. Second, once you've finished whatever you wanted to make, don't just chuck it at people; tell them a story about why it matters.

I thank the members of my committee who offered sage academic and career advice: David Rupp, Jeff Arnold, Ann Bostrom, and Faisal Hossain. I am deeply indebted to the current and former members of the Computational Hydrology Group at the University of Washington. In particular I thank Joe Hamman who helped me become a Pythonista. My many professors at the University of Washington, particularly those in the Program on Climate Change, provided me an excellent interdisciplinary base.

I am exceedingly grateful for the opportunity to meet and learn from people living and working with water and climate throughout the Pacific Northwest. The creative and effective members of the Climate Impacts Research Consortium taught me how to make sure my science useful and usable. I am particularly appreciative of Phil Mote, Katherine Hegewisch, and John Abatzoglou. I thank the many members of the Columbia River Intertribal Fish Commission, the River Management Joint Operating Committee and resource agencies throughout the PNW. I also thank the many, many people I chatted with (or will chat with) while on my perpetual search for what to be when I grow up.

I thank the National Science Foundation and the Institute for Pure and Applied Mathematics for the opportunities to attend, respectively, the American Meteorological Society Summer Policy Colloquium and the Women in Public Policy and Mathematics conference. Both of these programs afforded me a window into science for policy and policy for science.

Angie Malorni provided mac and cheese and radical analysis whenever called for, Rosa Gimson supplied celebrity gossip and a moral compass, and Rhiannon Bronstein taught me how to tackle fear when roped or unroped. I thank Andrew Bennett for suggesting that I acquire a watercolor set. I thank the following people, in reverse alphabetical order, for their senses of

humor: Diana Windemuth, Hazel Scott, Yixin Mao, Harper Hubbeling, Audrey Hendricks, Jiawei He, Yifan Cheng, Liz Clark, and Katie Brennan. The Seattle Solidarity Network taught me how to facilitate a meeting and the power of collective action. I thank anyone who has ever attended an event hosted by Dr. Party Planner. Andy Glaser taught me the freedom of going on a walk with nothing in my pockets.

I thank my family for making 90% of me. My mom taught me to love books. My dad carried me on his shoulders at demonstrations when I was young. The rest of my family raised me to appreciate a commotion.

This work was funded in part by the Bonneville Power Administration's Technology and Innovation Program under grant TIP 304 to the University of Washington and Oregon State University. Additional support was provided by the United States Army Corps of Engineers Climate Preparedness and Resilience Programs and the Bureau of Reclamation under Cooperative Agreement R17AC00024 to the University of Washington. Computing resources were provided by Pangeo and the Hyak Supercomputer at the University of Washington. This dissertation contains modified Copernicus Climate Change Service Information (2020). Neither the European Commission nor the European Centre for Medium-range Weather Forecasts is responsible for any use that may be made of the Copernicus Information or Data it contains. I acknowledge the World Climate Research Programme's Working Group on Coupled Modelling, which is responsible for CMIP, and I thank the climate modeling groups (listed in the Appendices) for producing and making available their model output. For CMIP the U.S. Department of Energy's Program for Climate Model Diagnosis and Intercomparison provides coordinating support and led development of software infrastructure in partnership with the Global Organization for Earth System Science Portals.

DEDICATION

For my family and friends

Chapter 1 INTRODUCTION

1.1. BACKGROUND

Climate change is extremely likely to alter critical freshwater resources (Asadieh & Krakauer, 2017; Gleick, 1989; Hoegh-Guldberg et al., 2018). Accordingly, the climate science community has devoted a great amount of effort toward investigating the projected impacts of climate change on hydrology. Countless studies agree that climate change is very likely to intensify the water cycle, creating a "the rich get richer" cycle in which wet areas become wetter and dry areas become drier (Stocker et al., 2013). These changes, combined with increased temperatures, are very likely to alter hydrology, including a loss of snowpack (Arnell, 1999), changes in the timing of snowmelt periods (Musselman et al., 2017), and an intensification of hydrologic extreme events (Kundzewicz et al., 2014). Climate change will have a significant, consequential impact on water resources.

The science driving the hydrologic climate change projections has grown sufficiently robust that water resources managers are increasingly willing, and in many cases eager, to incorporate climate change information into their planning efforts (Megdal et al., 2017). Indeed, it has become widespread for governments, ranging in size and complexity from small municipal water districts to multinational coalitions, to incorporate hydrologic climate impacts assessments into water resources planning (Bernardini et al., 2012; Kirchhoff et al., 2015). However, merely reporting projections of hydrologic change is insufficient to bridge the gap between hydroclimate impacts science and practice (Turnhout et al., 2013).

When converting the abstract science of hydroclimatology into the application of water resources engineering and planning, we apply real world constraints to physical findings. This

combination means that for scientific findings to be actionable, they must satisfy a number of criteria from the practitioner community (Vano et al., 2018). For example, the hydroclimate community has established that numerical land surface models can produce a wide spread of projections of hydrologic impacts. As a result, a water manager with limited resources would value a better understanding of how to tailor their hydrologic impact study to their specific questions of interest to reveal the uncertainty within the projections. A deeper understanding of the science and uncertainty within hydrologic projections supports the development of improved future impacts assessments.

This dissertation aims to strengthen the understanding of hydroclimate impacts studies, through robust analyses of scientific methods and large hydrologic climate change ensembles. These findings will be of use to both scientists and practitioners interested in engaging in hydrologic climate change research.

1.2. Research Objectives And Overview

This dissertation improves our understanding of hydrologic climate change impacts to support water resources planning. Specifically, the dissertation answers the following questions:

1) How do model decisions affect projections of hydrologic climate change impacts?
2) How will climate change impact the magnitudes and causes of high streamflow events in headwaters basins?
3) Can we identify hot spots of changes in water availability in international river basins at risk of hydropolitical tension?

The first two studies described in this dissertation are based upon an existing dataset of ensemble hydrologic climate change projections (Chegwidden et al., 2017). This dataset, driven by results from the Coupled Model Intercomparison Project Phase 5 (CMIP5; Taylor et al., 2012), provides streamflow projections for 400 locations across the hydroclimatically diverse Columbia River Basin in the Pacific Northwestern United States and Canada. Chapter 2 evaluates key metrics of changes in streamflow at these locations. Using the analysis of variance (ANOVA) statistical technique I disentangle the ensemble to determine which methodological choices have the strongest influence on changes in the projections. In Chapter 3 I delve deeper into the particularly consequential topic of floods and determine how flood generating processes are likely to be altered by climate change. I then apply methods similar to those of Vano et al. (2015) to derive first-order sensitivities of high flows to changes in annual climate. Lastly, inspired by the multinational status of the Columbia River, I take the lessons from Chapters 2 and 3 and apply them to multinational rivers across the globe. Using an ensemble of results from the CMIP6 experiment (Eyring et al., 2016) I assess projected changes in water availability within multinational basins. I focus on comparing changes between upstream and downstream countries sharing a basin. I summarize the results by combining the projections with documented assessments of hydropolitical risks (De Stefano et al., 2010).

Throughout the dissertation I work with ensembles of hydrologic projections. These ensembles are constructed through permutations of different elements of the impact modeling chain (Bosshard et al., 2013) or output directly from multiple GCMs from the CMIP project. I emphasize the models' and findings' strengths and limitations to promote trust and uptake within the planning community. In summary, this dissertation supports the development of rigorous hydroclimate impact science to support resilient water resource planning.

Chapter 2 HOW DO MODELING DECISIONS AFFECT THE SPREAD AMONG HYDROLOGIC CLIMATE CHANGE PROJECTIONS?

This chapter is published in the journal *Earth's Future*. © American Geophysical Union. Used with permission. The supplemental material for this chapter is provided in Appendix A.

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

2.1. INTRODUCTION

Robust analysis of hydrologic climate change requires an understanding of the spread among the projections of the estimated impacts. To quantify and assess this spread, the field of hydrologic projection has moved away from single simulation or small ensemble studies of future traces. Instead, it is increasingly popular to create large ensembles representing an array of possible hydrologic futures (Addor et al., 2014; Bosshard et al., 2013; Dobler et al., 2012; Finger et al., 2012; Harding et al., 2012; Hattermann et al., 2018; Lafaysse et al., 2014; Mendoza et al., 2015; Mizukami et al., 2016; Prudhomme & Davies, 2009; Vidal et al., 2016; Wilby & Harris, 2006). The string of models used to create these ensembles has been termed the "impact modeling chain" (Bosshard et al., 2013). These chains consist of a series of sequential modeling steps which ingest climate information and produce estimates of hydrologic impacts. An example of one such chain (the one implemented for this paper) is shown in Figure 2.1. Each step in the chain requires the selection of a model from a variety of modeling options. By implementing multiple options for

every modeling decision, we can investigate the ways in which modeling decisions affect our projections.

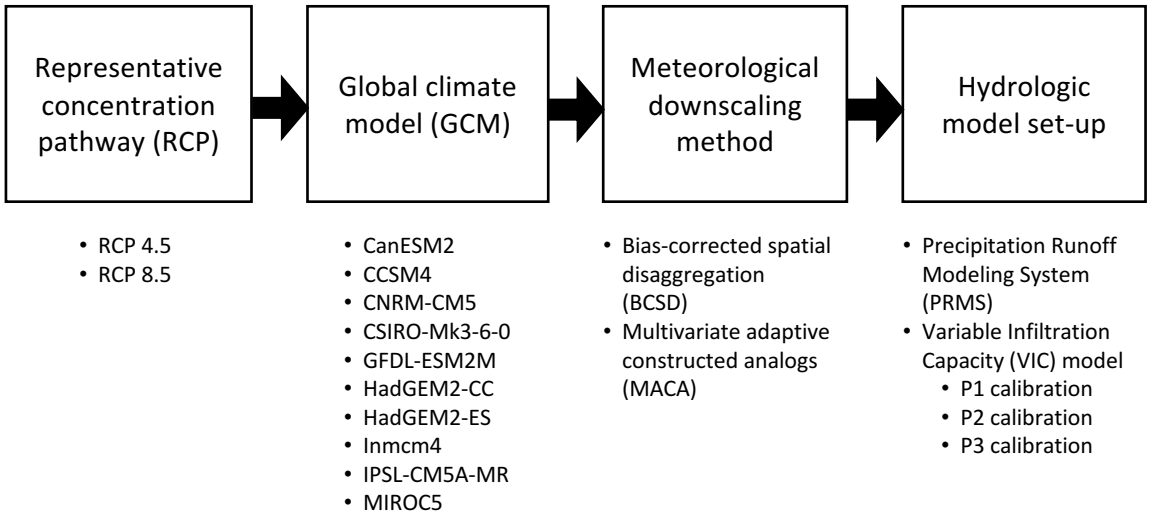


Figure 2.1 The study described in this paper consisted of four model decision points represented by the boxes above. Every permutation of the bulleted list of modeling options was evaluated to create the study's ensemble.

Using ensemble approaches to assess the impact of model choice is well established. Hawkins and Sutton (2009, 2011) attributed uncertainty in regional temperature and precipitation projections to choices of global climate model (GCM), representative concentration pathway (RCP), and internal variability (IV). The choice of hydrologic model implementation has been shown to have a strong influence within the hydrologic modeling impact chain in a variety of smaller basin studies (Addor et al., 2014; Jung et al., 2012; Najafi et al., 2011; Wilby & Harris, 2006), though Chen et al. (2011) found relatively smaller influences of hydrologic model choice compared to the choice of GCM. Hatterman et al. (2018) investigated the relative contributions of steps of the impact modeling chain on changes in streamflow across nine river basins across the globe. They found that the largest determinants of spread in changes in mean and high streamflow

volumes are GCM and RCP, but that changes in low flows were more likely driven by hydrologic model choice. However, their analyses did not extend to the effect of hydrologic model parameterization and did not account for IV. Addor et al. (2014) noted the potential benefits of conducting an uncertainty analysis across many diverse basins or a larger domain but acknowledge the computational constraints of an analysis of that scale. In summary, while comprehensive studies are common within smaller basins, it is difficult to obtain large-scale experiments that allow for consistent comparisons across modeling decisions.

In this paper, we conduct an experiment across a large and hydroclimatically diverse region: the Pacific Northwest of North America (PNW, Figure 2.2). We analyze an ensemble of future hydrologic scenarios using multiple options for four modeling decisions (see Figure 2.1). By combining every option from each decision in a factorial experiment we created 160 different possible hydrologic futures for the region (Chegwidden et al., 2017). We also explicitly account for the role of internal variability, or the natural fluctuations of the earth system which deviate from an externally-forced long-term trend. Lafaysse et al. (2014) and Alder and Hostetler (2018) emphasize the importance of this line of analysis, though it has otherwise often been neglected in hydrologic impacts studies.

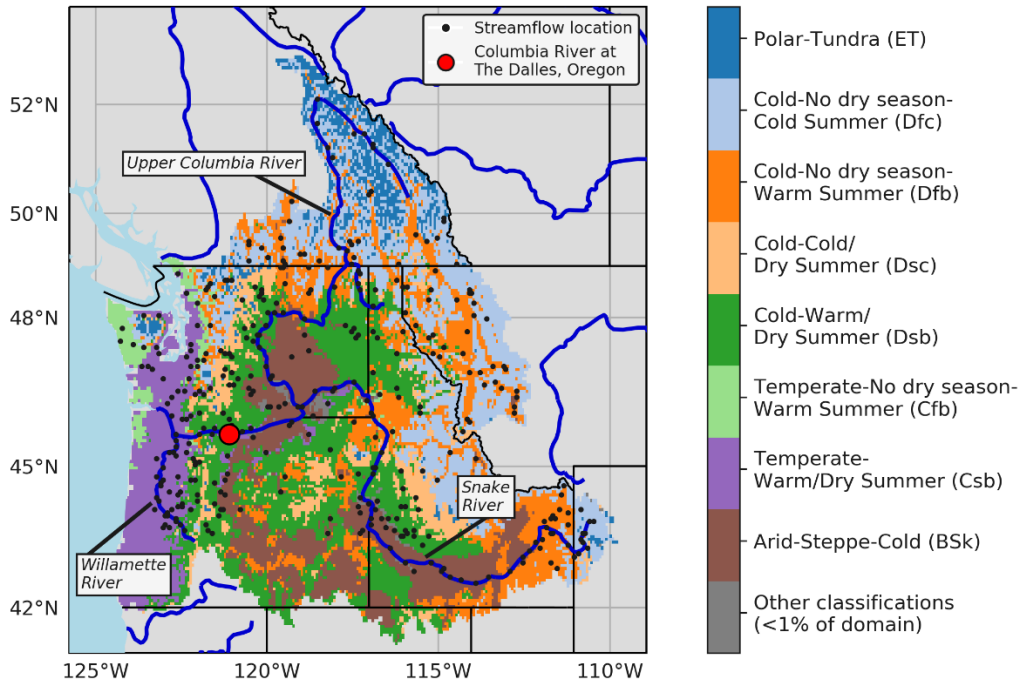


Figure 2.2 The study domain is shown above classified by Köppen-Geiger climatic region. Streamflow locations are shown as black dots.

Anthropogenic climate change and its hydrological impacts in the PNW have been extensively studied (Elsner et al., 2010; Schnorbus et al., 2014; Werner et al., 2013). Abatzoglou et al. (2014) have noted that anthropogenic climate change has already caused statistically-significant increases in temperature during the second part of the 20th century across the PNW. The changes, and their associated observed impacts on hydrology (Casola et al., 2009; Hidalgo et al., 2009; Mote et al., 2005, 2018; Stoelinga et al., 2010) are projected to continue during the 21st century (Rupp et al., 2017).

In this paper, we specifically use the term "spread" as opposed to "uncertainty" to emphasize that our ensemble is a subsample of the true distribution of all possible future scenarios. As a result, our numerical results are specific to our ensemble. Nevertheless, by incorporating multiple modeling options and conducting our analyses across a hydroclimatically diverse domain, we aim for our spread to capture much of that uncertainty space and thus make our findings more generally applicable.

Understanding the spread among projected changes is of critical interest given the extensive management of the Columbia River. These infrastructure projects are used for flood risk management, hydropower generation, irrigation water needs, recreation, and navigation. The dams also impact fish and ecosystem services, with notable relevance to First Nations. Further, because the basin straddles the border between the United States and Canada, the projected streamflow impacts have particular political significance. The two nations signed the Columbia River Treaty in 1964 outlining how the river would be managed and operated economically (U.S Army Corps of Engineers, 2012). In 2013 the Columbia River Management Joint Operating Committee (RMJOC) called for research to investigate climate change impacts for the Columbia River through the end of the 21st century, in part to inform the ongoing regional and international management of the river for the 21st century. Our study responds to these interests, with particular emphasis on the spread among the projected changes.

This paper answers two main questions:

(1) How does model spread depend on choice of evaluation metric and location? The diversity of climatic regimes and 150-year temporal extent of our study aid in answering this question.

(2) What effect do our modeling choices have on projections of hydrologic change? The factorial construction of this ensemble allows us to draw robust conclusions about the contribution of methodological choices to the spread among projections.

2.2. METHODS

2.2.1 Study domain

We conducted our study over the Columbia River Basin (drainage area 668,000 km^2) in the PNW as well as the United States (US) coastal Pacific Ocean drainages (Figure 2.2). The Pacific Northwest is characterized by contrasting climatic regimes. The western slopes of the Cascade and Olympics Mountains in the Pacific coastal drainages experience moderate temperatures and receive some of the highest precipitation in North America, more than 6000 mm annually (see the left panels of Figures A1.1 and 2.3). Conversely, some locations in the interior of the Columbia Basin receive an average of less than 100 mm annually. The Columbia Basin is bounded to the east and north by the cold, moist Rocky Mountains. Historically, precipitation within the mountainous regions has been winter-dominant, causing the region's hydrology to be snowmelt-driven.

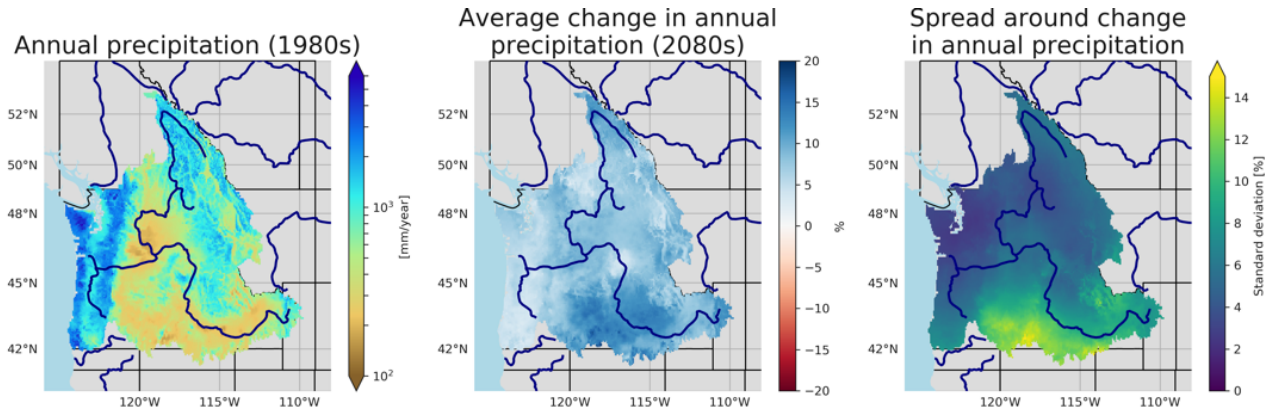


Figure 2.3 Historical mean annual precipitation (left), changes by the 2080s (center) and the spread around those changes (right) according to all 10 GCMs for RCP 8.5.

2.2.2 Historical datasets

A gridded meteorological forcing dataset (1/16th degree, ~6 km; Livneh et al., 2013) encompassing the PNW domain was used as a basis for meteorological bias-correction and two of the four hydrologic model calibrations. Reference streamflow data was taken from the "no regulation, no irrigation" (NRNI) streamflow dataset (RMJOC, 2017) which provides time series at 190 sites throughout the Pacific Northwest for the water years (WY) 1928-2008 (a water year in the western US is the period from October 1 in the previous year through September 30 of the current year). The NRNI dataset removes the effects of regulation, irrigation, and reservoir evaporation from measured streamflow, and can be used as a proxy for naturalized streamflow to support hydrologic model calibration.

2.2.3 Methodological choices

2.2.3.1 Representative concentration pathways (RCPs)

We selected climate simulations from the repository compiled by the Coupled Model Intercomparison Project Phase 5 (CMIP5; Taylor et al., 2012). For this study, we used the RCPs

corresponding to a high emissions scenario (RCP 8.5) and a midrange mitigation emissions scenario (RCP 4.5) as described in Taylor et al. (2012). As in Addor et al. (2014) we chose to include the RCP as another modeling choice to investigate how global changes affect local hydrology.

2.2.3.2 Global climate models (GCMs)

We selected ten GCMs (Figure 2.1) for this study by ranking each model's ability to best reproduce a variety of observed metrics over the historical period within the Pacific Northwest (Rupp et al., 2013). We further based our GCM selection on their ability to represent mesoscale atmospheric dynamics, which we assessed by evaluating 500 hPa heights in the control simulations.

2.2.3.3 Meteorological forcing downscaling methods (DSMs)

Two statistical DSM techniques were used to downscale each of the twenty RCP/GCM combinations to the $1/16^{th}$ degree (~6 km) grid resolution. The bias correction, spatial disaggregation (BCSD, Wood et al., 2004) method was implemented at the monthly time step. The multivariate adaptive constructed analogs (MACA) approach was implemented at the daily time step as described in Abatzoglou & Brown (2012). The Livneh et al. (2013) dataset from 1950-2005 was used as the reference for the bias-correction step of the DSMs. The resulting statistically-downscaled GCM data include control simulations (1950-2005) and climate change projections (2006-2099).

Table 2.1 Descriptions of the four hydrologic models (columns) referenced in the rightmost column of Figure 2.1.

	VIC-P1	PRMS-P1	VIC-P2	VIC-P3
Calibrated parameters	b_i, depth of soil layers 2 and 3, Ksat, Nijssen et al. (2001) parameters D1, D2, D3	slowcoef_sq, sat_threshold, pref_flow_den, gwflow_coef, ssr2gw_rate, snowinfil_max, soil2gw_max, soil_moist_max	See Oubeidillah et al. (2014)	See Mizukami et al. (2017)
Calibration methodology	Inverse calibration	Inverse calibration	Lumped basin calibration	Calibrated parameter transfer functions
Reference meteorological dataset	Livneh et al. (2013)	Livneh et al. (2013)	Daymet (Thornton et al., 1997)	Livneh et al. (2013)
Reference streamflow dataset	No-regulation, no-irrigation flows (RMJOC, 2013)	No-regulation, no-irrigation flows (RMJOC, 2013)	USGS WaterWatch gauges at monthly timestep (Brakebill et al., 2011)	Hydro-Climate Data Network basins (Newman et al., 2014)

2.2.3.4 Hydrologic model implementations

The downscaled CMIP5 meteorology was used to force four distinct hydrologic model implementations (Table 1). All hydrologic modeling was conducted on the same 1/16th degree grid with results available at a daily time step from 1950-2099. Two different hydrologic model codes were used in this project: the Variable Infiltration Capacity (VIC; Liang et al., 1994) model and the Precipitation Runoff Modeling System (PRMS; Leavesley et al., 1983). Both are process-based, energy balance models. The VIC model code included a newly-developed glacier model (Hamman & Nijssen, 2015) which uses a volume-area scaling relationship (Bahr et al., 1997) to estimate the areal coverage of a glacier within a grid cell. When the areal coverage of the glacier exceeds the area allocated to a given elevation zone, the model moves ice to lower, warmer elevation zones enabling glacier melt. We used three unique VIC implementations, each with independently derived parameter sets, denoted as P1, P2, and P3. Each VIC simulation was initialized with a 255-year spin-up run to allow for sufficient initialization of the glacier model.

The PRMS simulations involved a 20-year spin-up. Within this paper, "hydrologic model" (HM) will refer to the combination of model code and parameter set. Historical performance of the four HMs is displayed in Figure A1.2.

The VIC-P1 and PRMS-P1 implementations (Table 1) were developed specifically for this study. For both implementations, we calibrated model parameters in the Columbia Basin upstream of Bonneville Dam and the Willamette River Basin. We selected the NRNI streamflow timeseries at 60 sites to create a spatially-distributed set of runoff timeseries using the approach detailed in Pan et al. (2013). Each grid cell in the VIC-P1 and PRMS-P1 HMs was then calibrated independently. We implemented the shuffled complex evolution calibration method (Duan et al., 1993) maximizing the weekly Kling-Gupta Efficiency (KGE) statistic for WY 1992-2001 (Gupta et al., 2009). For coastal drainages we used parameters from a previous Columbia River climate change study (Hamlet et al., 2013) for VIC-P1. We developed the non-calibrated soil parameters for the PRMS-P1 implementation by merging the State Soil Geographic Database (STATSGO) for the United States portion of the domain and the Canadian Soil Survey (Soil Classification Working Group, 1998) and followed methods recommended in the PRMS documentation to develop PRMS parameters (USGS, 2009).

The VIC-P2 implementation used the same vegetation parameter sets as those used in VIC-P1, but with soil and snow band parameters as described by Oubeidillah et al., (2014). The simulated runoff at each HUC8 subbasin was calibrated to USGS WaterWatch runoff at the monthly time step using the Daymet meteorological forcing dataset (Brakebill et al., 2011; Thornton et al., 1997), in contrast to the other three set-ups which used Livneh et al. (2013). Parameters were used as-is except for the area upstream of Hungry Horse Dam which was calibrated to NRNI for that gage to improve model performance for that subbasin.

The VIC-P3 implementation used the same vegetation and snow band parameters as VIC-P1, but with a soil parameter set developed using the multiscale parameter regionalization (MPR) method (Mizukami et al., 2017). Instead of the traditional parameter calibration, this technique calibrates the parameters of transfer functions that relate soil and landscape properties to model parameters for a subset of a domain. The transfer functions can then be used domain-wide to estimate model parameters outside of the calibration basins. Transfer function parameters were calibrated at a daily time step for unimpaired or Hydro-Climate Data Network basins and applied to the continental U.S. to produce spatially-consistent parameters for the entire domain (Newman et al., 2014).

2.2.4 Streamflow routing

Modeled runoff fields were routed to 396 sites throughout the domain (see streamflow locations in Figure 2.2; Chegwidden et al., 2017) using the RVIC routing model (Hamman et al., 2017), a source-to-sink model based on the Lohmann et al. (1996) routing model. The same routing set-up was used for all simulations. It does not account for any water management (e.g. reservoirs or withdrawals) and the resulting streamflows (and our ensuing analyses) do not address uncertainties related to water management modeling or operations.

2.2.5 Analysis of variance

We focused our analyses on spreads in projected changes in streamflow and assessed how model decisions and internal variability contribute to those spreads. We compared streamflow for a 30-year control period (the “1980s”, consisting of WY 1970-1999) with a 30-year future period (“2080s”, WY 2070-2099). For each of these periods, we calculated three streamflow metrics: 1)

centroid of timing, defined as the day of the water year when 50% of the annual flow volume has passed a streamflow location, 2) annual volume, and 3) annual minimum 7-day flow. We examined the change signals as opposed to the future values themselves to control for systematic differences in the model set-ups. We calculated the change signal for each ensemble member independently and thus created a 160-member ensemble of projected change signals from the 1980s to the 2080s.

We used this ensemble of projected changes in streamflow to evaluate the impact of methodological choices on those projected changes. We developed a methodology inspired by previous uncertainty analysis studies (e.g. Alder & Hostetler, 2018; Bosshard et al., 2013; Hawkins & Sutton, 2009) but we highlight that our methods (1) focus on changes in 30-year means and (2) explicitly account for internal variability.

To estimate internal variability, we assumed each metric responds linearly to the radiative forcing in a given RCP. Specifically, for each annual time series (1950-2099) of each metric and streamflow location, we fitted a linear model $x = af + b$, where x are the predicted annual values of the metric, f are the annual values of radiative forcing corresponding to the respective RCP, a and b are the fitted coefficients. We used the residuals ε of the fit to estimate the internal variability of a change in 30-year mean. See Section S2 in the Supporting Information for a detailed description.

We used analysis of variance (ANOVA) to assess the impacts of methodological choices on the spread of the projections of change in the 30-year means from the fitted time series $x(t)$. Our final variance analyses can be summarized as

$$\mathrm{TV} = \mathrm{IV} + \mathrm{MV} \tag{1}$$

where TV is the total variability in a change projection, IV is the internal variability of the projected change, and MV is the model variability defined as

$$\mathrm{MV} = \mathrm{RCP} + \mathrm{GCM} + \mathrm{DSM} + \mathrm{HM} + \mathrm{GCM{:}\,RCP} + \mathrm{HM{:}\,GCM} + \mathrm{Residual} \quad (2)$$

where RCP, GCM, DSM, and HM are the portions of variance explained by the corresponding methodological choices. (Note that "Residual" in Equation (2) is distinct from residual ε used to estimate IV). The interaction terms "$\mathrm{GCM{:}\,RCP}$" and "$\mathrm{HM{:}\,GCM}$" were the dominant terms in the residual error in the variance. We lumped other interaction terms into the residual term in Equation (2) to preserve degrees of freedom within the ANOVA analysis.

2.2.6 Climate classifications

We applied ANOVA at every model grid cell in the domain. The grid cells were grouped according to the updated Köppen-Geiger classification scheme (Peel et al., 2007) based on the monthly temperature and precipitation data from Livneh et al. (2013) for the historical period 1970-1999. Over 99% of grid cells were classified into eight climate regions (Figure 2.2). The remaining <1% of the grid cells were excluded from the analyses for simplicity.

2.3. HYDROLOGIC IMPACTS OF CLIMATE CHANGE AND PROJECTION SPREAD

Snow accumulation and melt are the dominant drivers of the hydrologic cycle in the Columbia River basin. As temperature increases and snow presence decreases, the role of snow storage will diminish. In this section, we briefly review the dominant climate change impacts on the hydrology of the Columbia River Basin but focus our analysis on the spread within the

ensemble of projections. The projected changes in hydrologic variables are in overall agreement with previous work in the PNW (Elsner et al., 2010; Hamlet et al., 2013).

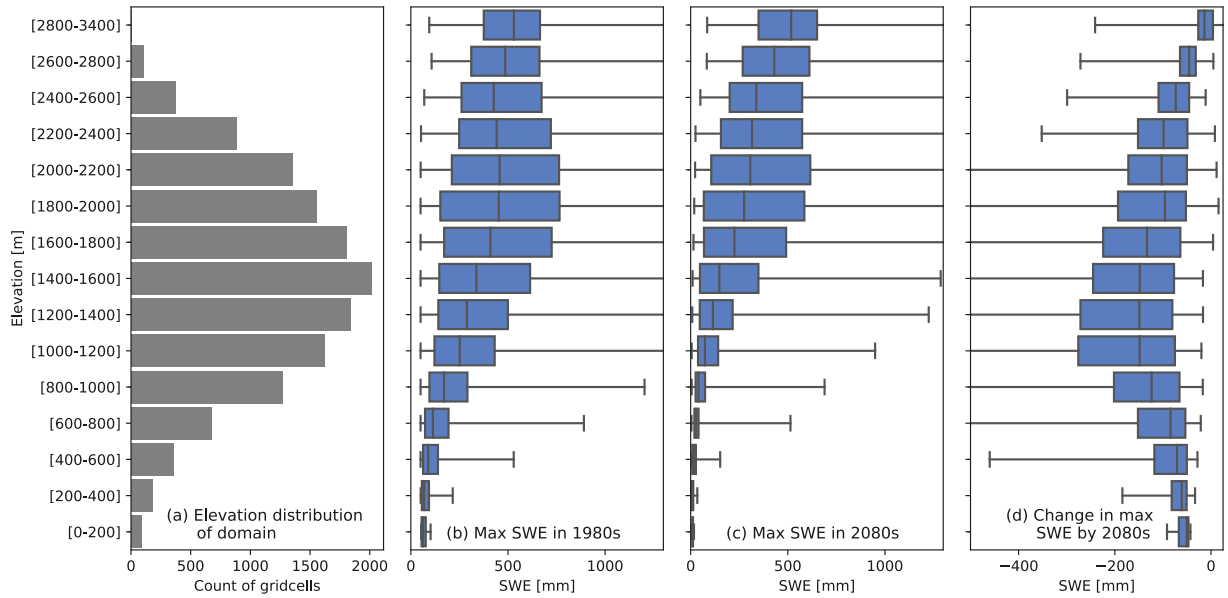


Figure 2.4 For fifteen elevation zones shown in (a), the distribution of modeled snow water equivalent (SWE) for the 1980s (b), the 2080s using RCP 8.5 (c), and the difference (d). In the last three panels, boxes and whiskers indicate minimum, median, and maximum values, and the interquartile range.

2.3.1 Changes in snowpack in the Pacific Northwest

For our snow analyses we only considered areas where snow accumulation and melt are an important hydrologic process, which we defined as grid cells with more than 50 mm of annual maximum snow water equivalent (SWE) during the historical period. The elevation profile of these areas is shown in Figure 2.4a. The areas between 1000 and 1800 m are projected to experience the highest losses of maximum SWE, with median decreases of 150 mm by the 2080s (RCP 8.5). The spread in the changes is largest for elevations between 1000-1400 m (Figure 2.4d). The spreads are narrower at the lowest and highest elevations because little snow is present at low elevations

and winter temperatures remain below freezing at high elevations despite anthropogenic warming.

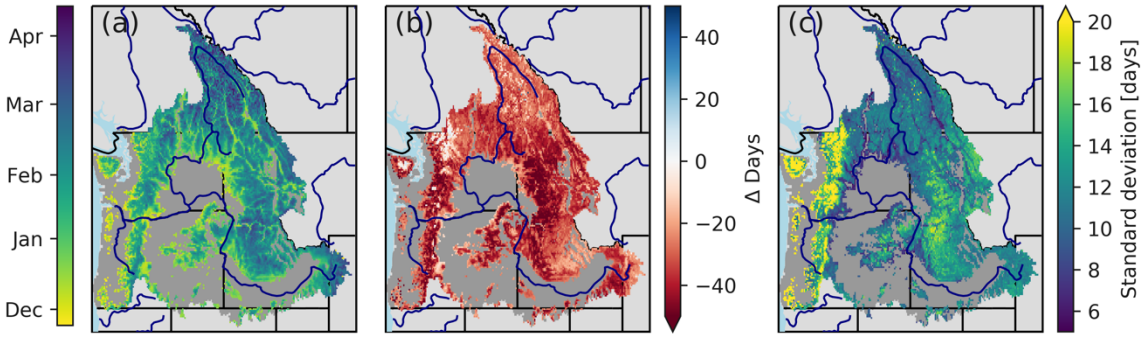


Figure 2.5 The date of median annual maximum SWE for the 1980s is shown in panel (a). The ensemble (RCP 8.5) median change in the date of peak SWE by the 2080s is shown in panel (b). The standard deviation (measure of spread) of the distribution of projected changes in date of peak SWE is shown in panel (c).

The median date of annual maximum peak SWE for the 1980s as simulated within the ensemble is shown in Figure 2.5a and the median change in the date of peak SWE is shown in Figure 2.5b (RCP 8.5). While annual maxima of peak SWE occurred historically in March, for much of the domain the date of peak SWE is projected to occur approximately 20 and 40 days earlier by the 2080s. The standard deviation of the projected changes in peak SWE date is shown in Figure 2.5c and can be interpreted as a measure of the ensemble spread. The greatest spread in change in peak SWE date occurs on the western side of the Cascades and the Olympic mountains. These areas lie in the transient snow zone and already experience significant mid-winter melt episodes during the historic period.

2.3.2 Annual streamflow changes and model agreement

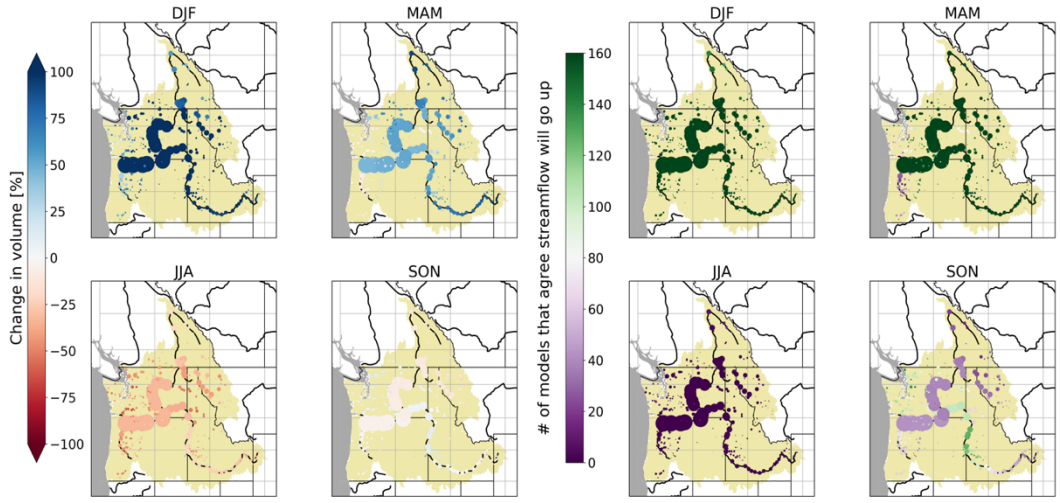


Figure 2.6 Changes (from control to future periods) in mean seasonal streamflow volumes for locations throughout the domain for winter (DJF), spring (MAM), summer (JJA), and fall (SON) are shown in the left panel. The right panel shows the model agreement on the direction of streamflow change. Streamflow location marker sizes are scaled by the mean annual historic NRNI flow.

Annual streamflow volumes are projected to increase at all locations in the domain. This is consistent with a general increase in precipitation across the PNW (Figure 2.3). Changes in seasonal streamflow between the control and future periods are shown in the left panel of Figure 2.6 for winter (December, January February), spring (March, April, May), summer (June, July, August), and autumn (September, October, November). Streamflow is projected to increase at all locations in winter and spring. Summer streamflow is, on average, projected to decrease owing predominantly to an earlier shift in snowmelt onset accompanied by a reduction in summer precipitation and increases in evaporation due to higher temperatures. Changes in streamflow across the basin are further discussed in RMJOC (2018).

We analyzed the ensemble to determine the level of agreement regarding streamflow increases or decreases between the control and future periods (four right panels in Figure 2.6).

Ensemble agreement has been shown to be a good summary indicator of overall ensemble behavior (Melsen et al., 2017). Green colors indicate that the ensemble generally agrees that streamflow will increase, while purple colors indicate a decrease. Lighter colors indicate disagreement among the ensemble members.

Models agree that streamflow will increase at almost all locations in winter and decrease at almost all locations in summer. Models agree that spring streamflow will increase for most locations as a result of the earlier shift of the snowmelt peak. Model agreement is smallest during autumn. The most notable disagreement is along the Snake River in the southeast of the study domain in the autumn where 110 ensemble members project increases and 50 project decreases. While percentage changes in streamflow are smallest in autumn, late summer/early autumn is also a critical period for many locations so small streamflow changes could still significantly impact water users.

2.4. CONTRIBUTION OF METHODOLOGICAL CHOICES TO ENSEMBLE SPREAD

In this section, we will evaluate the spread within the ensemble projections of hydrologic change in different regions and determine how methodological choices impact the spread.

2.4.1 Streamflow seasonality for the Columbia River Basin

Mean-monthly hydrographs for the 160-member ensemble for the Columbia River at The Dalles (red marker in Figure 2.2) are shown in Figure 2.7. Every line in the spaghetti plot represents a single ensemble member for a given 30-year period. The grey traces in Figure 2.5a represent the 1980s climatological period. Within the historic period the streamflows lie close to the NRNI reference flow (shown in black), indicating good skill of the modeling set-up in reproducing the

historic reference flow.

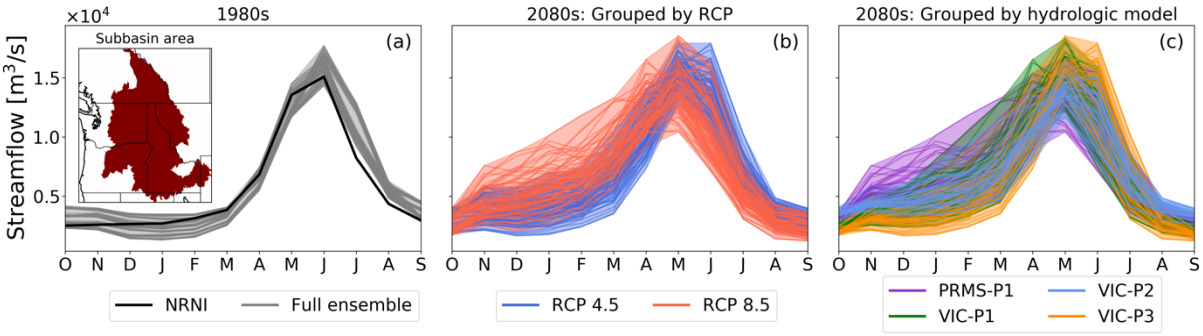


Figure 2.7 Ensemble seasonal hydrographs for the Columbia River at The Dalles. The basin area is highlighted in red in the inset map. Each trace represents a different ensemble member. Panel (a) shows the historical 1980s climatological period including the no-regulation, no-irrigation reference flow in black. Panels (b) and (c) show the projected streamflow by the 2080s, grouped respectively by RCP choice and HM choice.

The traces in Figures 2.7b and 2.7c are identical but are highlighted differently to show the effect that methodological choice has on streamflow projections. In Figure 2.7b, the ensemble is grouped by emissions scenario. In Figure 2.7c, the ensemble is grouped by HM. By the 2080s, the RCP 4.5 simulations show the snowmelt-related streamflow peak in either May or June whereas RCP 8.5 simulations peak a month earlier. Furthermore, the RCP 8.5 simulations show large increases in winter streamflow. By the 2080s, the spread in simulations widens, particularly in the winter months when shifts from snow to rain regimes differ among the models.

Comparing groupings in Figure 2.7c, PRMS-P1 simulations show the greatest sensitivity to the changed climate of the 2080s, as they exhibit the greatest increases in winter streamflow. VIC-P1 and VIC-P3 simulations exhibit shifts in the peak toward April and May but VIC-P2 shows the least sensitivity to the changing climate.

2.4.2 Impact of modeling decisions on ensemble's distribution of centroid shift

The shift in centroid of timing is an indicator of overall changes in streamflow timing. The full ensemble's distribution of changes in the centroid of timing (from control to future periods) for the Columbia River at The Dalles is shown in Figure 2.8. The Gaussian kernel density estimate (KDE) of the full ensemble is shown in the black line in every panel of Figure 2.8. The projected timing shifts range from zero to 60 days earlier. The KDE of timing shifts is decomposed according to three different methodological choices: RCP (Figure 2.8a), GCM (Figure 2.8b), and HM (Figure 2.8c). The effect of DSM choice is not shown as the distinction between BCSD and MACA had minimal impact on the timing shift for this location. Choice of RCP has a large impact on the timing shift at The Dalles, with a mean timing shift of 19-days earlier for RCP 4.5 and 35-days earlier for RCP 8.5. The timing shifts grouped by GCM are also markedly distinct, with differences between some GCMs as large as between RCPs. PRMS-P1 was associated with the largest changes in centroid of timing while simulations from VIC-P2 showed the smallest shift.

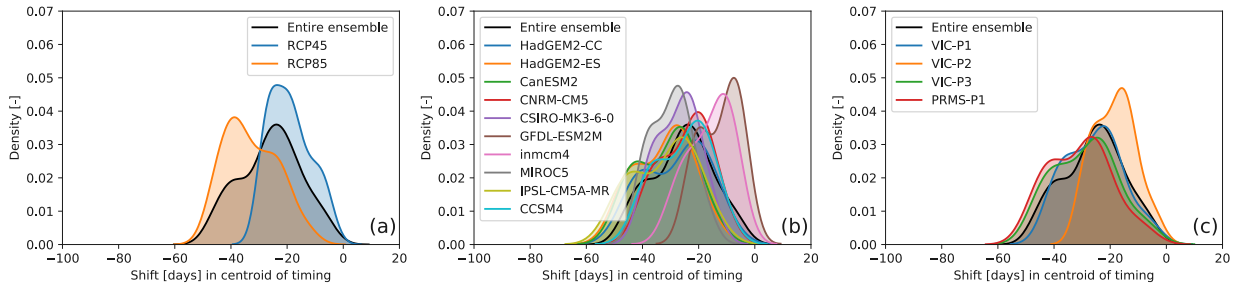


Figure 2.8 Distribution of the full ensemble's projected changes (1980s to 2080s) in centroid of timing for streamflow of the Columbia River at The Dalles, Oregon. The black line represents the full ensemble of projected changes and is identical in all three panels. In each panel the distribution is decomposed into simulations sourced from different methodological options. The distribution is separated according to (a) RCP choice, (b) GCM choice, and (c) HM choice.

2.4.3 Contribution of each modeling component to ensemble spread

The ANOVA quantifies the contribution of each modeling choice to the spread in changes. The pie charts in Figure 2.9 show, for 12 locations throughout the domain, how much each modeling choice, as well as internal variability, explains changes in different metrics: centroid of timing (Figure 2.9a), annual volume (Figure 2.9b), and low flows (Figure 2.9c). The size of each pie is proportional to the total variance in that metric at that location. Overall, choice of RCP drives the spread in the change in timing, choice of GCM contributes most toward the spread in the change in volume, and choice of HM explains most of the spread in change in the magnitude of the low flows.

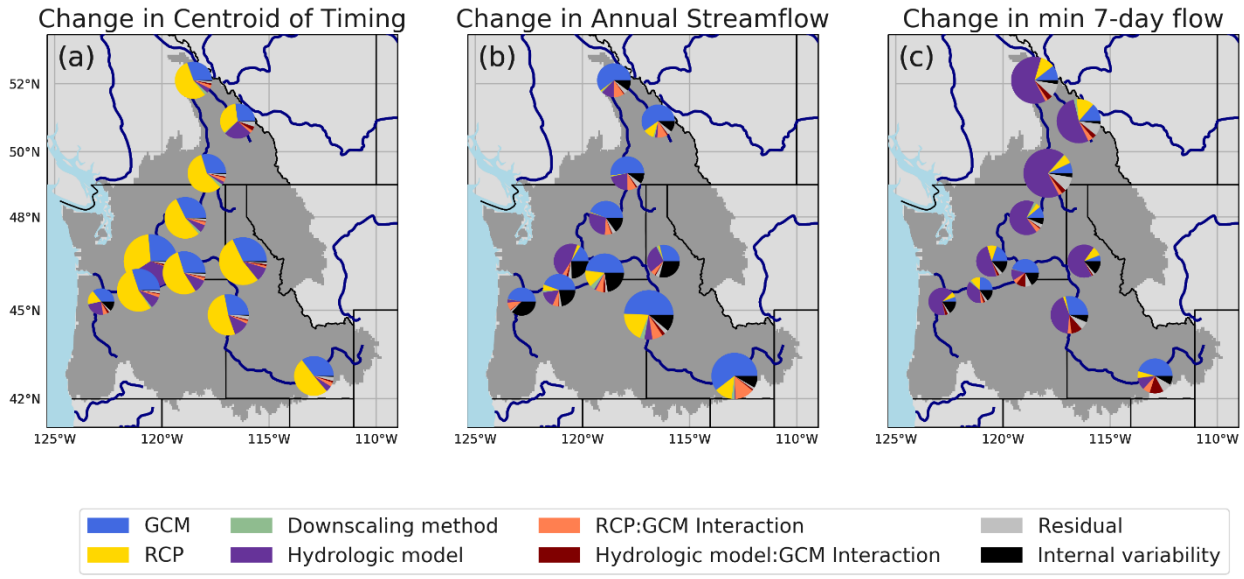


Figure 2.9 ANOVA results for the change in three metrics between the 1980s and the 2080s. The portion of variance of the change explained by each methodological choice is shown for selected sites of interest. The size of each marker is scaled by the standard deviation of the ensemble for that metric and location.

Figure 2.10 highlights the dominant driver of the spread in the change for all metrics at all locations in the domain. For downstream locations, RCP tends to be the dominant contributor to the spread in timing shift (Figure 2.10a). The choice of GCM is the dominant contributor to the spread among locations in the Upper Columbia, the Willamette, and small coastal drainages in Washington State.

Changes in annual volume (Figures 2.9b and 2.10b) are largely controlled by the choice of GCM, given annual volume's dependence on precipitation. The control of GCM on volume change is strongest along the Snake River where GCMs show a large spread in precipitation changes (Figure 2.2). In contrast, when the spread of projected precipitation change is small, as along the mainstem of the Columbia above the confluence with the Snake River, HM choice is most relevant to explaining the spread.

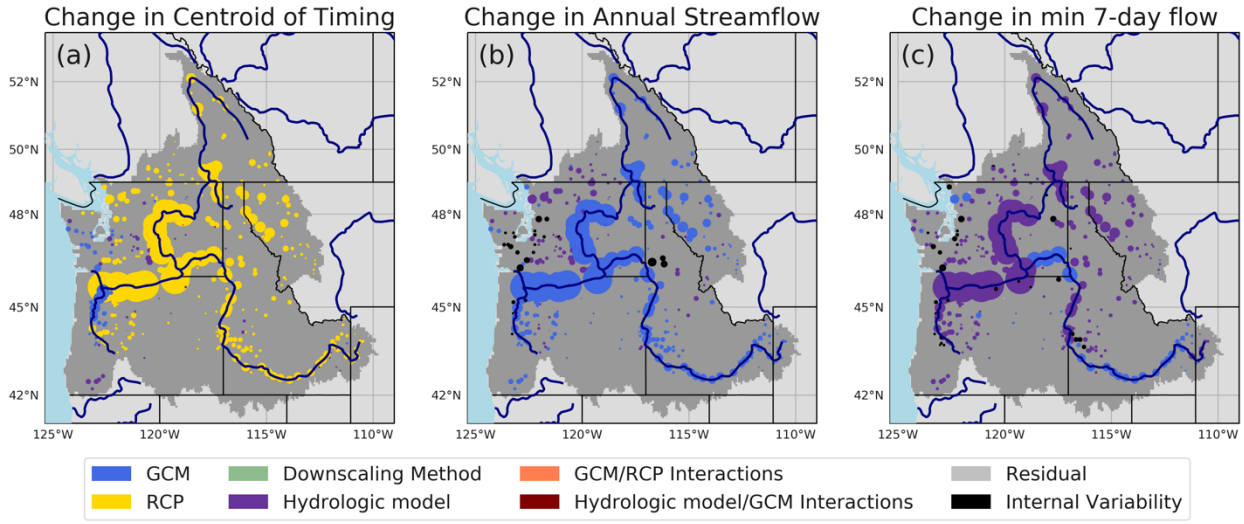


Figure 2.10 The dominant contributor to the variance of the projected changes in three metrics from the 1980s to the 2080s: (a) change in centroid of timing, (b) change in annual streamflow, and (c) change in low flows. Marker sizes are scaled by the mean annual historic NRNI flow.

The spread in changes in low flows (Figures 2.9c and 2.10c) are largely explained by choice of HM. The low flow period in the domain occurs typically in the drier early autumn when snow

and soil moisture stores are depleted. Because changes in low flows are driven by soil processes as opposed to precipitation, and each HM had distinct soil parameter datasets (VIC implementations) or different process representation (VIC versus PRMS), the baseflow from each model varies widely. The exception to this relationship appears in the Snake where the diversity among precipitation projections overrides low flow generation so that changes in low flows are controlled more by GCM than HM.

2.4.4 Generalizing results by climatic region

We conducted the ANOVA for the runoff from every grid cell in the domain and averaged the results across each of the eight climate regions (Figure 2.11). The averaging gives every grid cell equal weight, regardless of its contribution to the streamflow of the basin. In this way, we see how each region responds, but do not expect the results to agree with those conducted at streamflow locations which aggregate across upstream area, thus favoring grid cells with greater runoff contributions. We leverage the diversity of the domain to elucidate general patterns linking hydroclimatic regime, metric of interest, and model spread to inform future hydroclimate impacts studies.

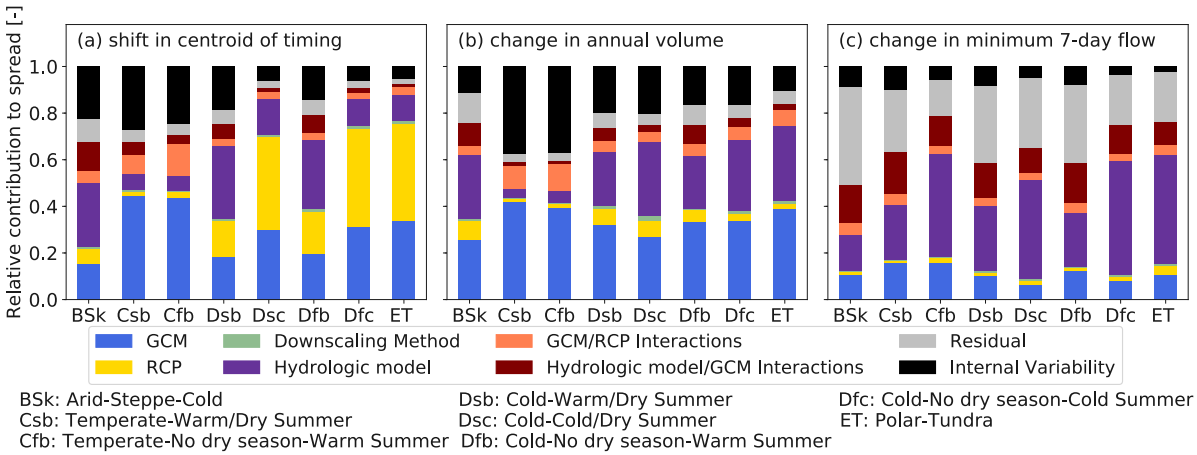


Figure 2.11 The ANOVA results grouped by climatic region are shown above. The relative size of each color band denotes how much each methodological choice explains the spread in the change in each region.

2.4.4.1 RCP choice impactful for timing in areas susceptible to large changes in snowpack

As seen in Figure 2.11a, mid-elevation cold areas (Dfc) comprise the region affected most by RCP choice for the change in timing. The Dfc region is also projected to experience significantly larger losses in snowpack according to RCP 8.5 as compared to RCP 4.5 (Figure 2.3), meaning that the choice of RCP would have a pronounced impact on the resulting changes for these regions. The same relationship does not apply to the lower elevation Dfb regions which are expected to lose nearly all of their snowpack by the 2080s making the choice of RCP less important for those regions (Figure 2.11a).

2.4.4.2 Choice of HM more important in water-limited regimes

When evaporative and baseflow processes account for a larger portion of the water balance (whether in low flow periods or in arid locations), the choice of HM becomes a more impactful decision, in agreement with Vidal et al. (2016). The dominance of HM in the change in low flows seen in Figures 2.9c and 2.10c is again evident across climatic regime (Figure 2.11c). Choice of

HM was the dominant driver of the change in centroid of timing for arid sections of the domain (Figure 2.11a). Across all climatic regimes, GCM-HM interactions contributed more to the spread in the low flow metric than to the other two metrics.

2.4.4.3 GCM choice and IV more important for energy-limited regimes

The majority of the streamflow in the Pacific Northwest originates in energy-limited environments, as seen by the overlap of the high-precipitation areas in Figure 2.3 and the cold, wet regions (Dfc, Dfb) in Figure 2.2. In these regions, changes in precipitation are strongly correlated with changes in annual streamflow volumes. Choice of GCM is a much larger contributor to change in precipitation than is RCP. Thus, the choice of GCM is the dominant contributor to spread in annual volume, as confirmed by Figure 2.10b.

The impact of GCM choice is enhanced when the GCMs lack consensus about the projected changes in precipitation, which aligns with findings from Melsen et al. (2017). For example, in the Willamette River Basin, precipitation is projected to increase between 2 and 5%, but with a standard deviation of about 5% (Figure 2.2). Thus, some GCMs project increases in precipitation and others decreases. This same region is the basin with the most disagreement on the direction of change in annual streamflow (Figure A1.3). Accordingly, the choice of GCM accounted for about 50% of the total spread in projections for the Willamette River Basin, with IV and GCM-RCP interactions accounting for nearly the entire remainder.

IV is more important (and sometimes becomes the dominant contributor to the spread) when precipitation changes are the most significant driver. For example, across the three metrics, IV is most important for changes in annual streamflow. Further, IV is more important for the temperate coastal regions in the domain (Csb and Cfb in Figure 2.11), even for changes in centroid

of timing, where rain- (as opposed to snow-) dominance results in a strong dependence of centroid of timing on changes in precipitation rather than temperature. The IV contribution is smaller for our results than those reported by Alder and Hostetler (2018), who used a shorter 11-year averaging window instead of our 30-year averaging window.

2.4.4.4 Changes in annual streamflow explained most by GCM selection

Figures 2.9b and 2.10b show that the choice of GCM along with its interaction with RCP account for the largest portion of the spread in projections of change in annual volume for this ensemble. This is particularly true in the temperate regions of the domain (Csb and Cfb) where choice of HM has negligible impact. In these regions, the runoff ratio is so high that, with precipitation as the main driver, GCM and its interaction with RCP account for nearly the entire spread in the projected changes.

2.4.4.5 Choice of DSM did not explain spread for the regions and metrics studied

The difference between the two statistical DSMs (BCSD and MACA) was not an important contributor to the spread in these metrics or domains. Choice of DSM contributed most to the spread in projections of changes in annual volume in the arid, smaller basins of the domain, in agreement with Jiang et al. (2018). However, this paper does not address extreme high flows where the two statistical DSMs are expected to differ more substantially in their projections of precipitation changes given that the two techniques were implemented at different time steps: BCSD at monthly and MACA at daily. Nevertheless, our ensemble did use the same meteorological dataset for training both DSMs and, as Alder and Hostetler (2018) note, the training dataset can have a strong influence on projected changes in snowpack. We acknowledge that DSM

choice might have contributed more to ensemble spread had our modeling chain included additional DSMs (e.g. additional statistical techniques or dynamical downscaling). This is supported by the work of Gutmann et al. (2014) who evaluated differences among a large number of statistical downscaling methods and established particularly strong impacts on precipitation estimates.

2.4.4.6 Residual most relevant in water-limited situations

The spread is most explained by the residual when water is limited, either by climate (i.e. arid regions) or by our analysis focusing on low flows. For example, in Figure 2.11c we see that the residual is one of the strongest contributors to low flows at the gridcell level. In some of the smaller arid basins of the southeastern domain the residuals even become the dominant contributor to the spread (Figure 2.10c).

2.5. CONCLUSIONS

Our findings about hydrologic changes in the PNW in response to climate change are generally consistent with other studies (Elsner et al., 2010; Hamlet et al., 2013). However, the spread among the projections is considerably wider as a result of expanding the number of modeling options used to create the ensemble. While the ensemble projects decreasing peak SWE throughout the domain, the spread among those projections is widest in mid-elevation areas where changes are projected to be greatest. Further, the ensemble is in good agreement that Canadian portions of the basin are likely to retain a large part of its snowpack through the end of the 21st century. This shift in snow retention has significant implications for international agreements regarding water resources planning. The ensemble agrees overwhelmingly that winter streamflows

will increase across the domain. There is also overwhelming agreement that summertime streamflows will decrease. Most disagreement exists in projections of autumn streamflow, especially in smaller tributaries. The disagreement in these locations highlights that there is generally less consistency among projections of climate change impacts on smaller upstream areas.

Our analyses offer guidance for future hydrologic climate impacts studies. The extent to which each modeling decision contributes to the spread in projections of change depends on both location and metric of interest. Choice of RCP is most impactful when enhanced snowmelt will affect the region or metric. For situations in which the GCMs disagree about changes in precipitation, and in particular in energy-limited environments, GCM selection is most important. Choice of HM is a larger contributor for situations in which soil processes are more active (e.g. low-flows, arid environments). The dominance of the residual in water-limited situations necessitates further work involving the ways that modeling components interact. This is of particular interest given the acute stresses that these low-flow periods have upon water users and ecosystems.

Our findings are largely robust to the internal variability in the system. Overall, the role of IV is strongest when precipitation is a driving factor, moderate when temperature is the driving factor, and smallest when the HM is the driving factor. Nevertheless, for the 30-year period over which we calculated our metrics, model variance exceeds that of IV. Thus, modeling choices have a greater impact on our results than the natural variability in the system. As Hawkins and Sutton (2009) note, the portion of variance contributed by IV is irreducible since it comes from the natural climate system. On the other hand, the model variance highlights areas where model refinement could improve hydrologic projections of climate change impacts.

These findings can inform future modeling efforts by leading to a select set of recommendations. While increasing the number of ensemble end members improves the understanding of the spread in projections, based on our results it may be possible to tailor a future study to the questions of interest. One could emphasize analysis and resources within the model decision step which accounts for the most spread in the ensemble for any metric of interest. For example, for adequately understanding climate change impacts on low flows, it may be an efficient use of resources to evaluate how and why hydrologic models differ in their sensitivities of changes in low flows. Within this domain and this ensemble, the influence of DSM choice was negligible compared to that of GCM choice, but we only investigated two statistical DSMs and thus likely underestimate DSM contribution to model spread. Therefore, it may be useful to shift efforts toward incorporating more GCMs into an ensemble, particularly in areas with high runoff ratios where changes in precipitation will be the driving factor determining metrics like changes in annual volume.

To best aid water resources preparedness, it is critical that modeling efforts capture the most representative range of possible outcomes. Given perennial computational restraints, we might derive more confidence in future projections by emphasizing model diversity at steps in the chain most relevant for our research questions.

Chapter 3 CLIMATE CHANGE ALTERS FLOOD MAGNITUDES AND MECHANISMS IN CLIMATICALLY-DIVERSE HEADWATERS ACROSS THE NORTHWESTERN UNITED STATES

This chapter is in review for publication in the *Environmental Research Letters*. The supplemental material for this chapter is provided in Appendix B.

Chegwidden, Oriana S., Rupp, David E., and B. Nijssen, 2020: Climate change alters flood magnitudes and mechanisms in climatically-diverse headwaters across the northwestern United States. Environmental Research Letters, in review.

3.1. INTRODUCTION

The hydrology of river systems is expected to change under projected human-caused changes in air temperature and precipitation. Changes to high flows are of particular interest since they can lead to damaging flooding events (Smith & Katz, 2013). A large body of literature assesses how high flows have changed in magnitude in observational records (Mallakpour & Villarini, 2015; Matti et al., 2017; Yuan et al., 2018) as well as their response to projected climate change (Bürger et al., 2011; Knighton et al., 2017; Köplin et al., 2014; Mandal & Simonovic, 2017; Mantua et al., 2010; Taye et al., 2011; Vormoor et al., 2015). A growing number of studies recognize the importance of incorporating upstream metrics and processes (e.g. temperature, snow-rain partitioning, soil moisture) into their assessments of changes in high flow magnitudes (Davenport et al., 2020; Wasko & Nathan, 2019; Wasko & Sharma, 2017; Yan et al., 2019) and the relative prevalence of different flood generating processes (FGPs; e.g. Berghuijs *et al* 2016,

2019). However, these studies were all based on the observational record which has limitations for extrapolating under a changing climate. First, applying historic relationships may not hold under a future climate that differs markedly from the observed climate. For example, Diffenbaugh (2020) showed that extrapolation based upon even the most recent historical records can be tenuous. Second, the dominant FGPs may change in an altered climate, particularly when warming causes a substantial change in precipitation phase. No study to date has explicitly calculated sensitivities of high flows, classified according to their dominant FGP, to changes in climate. Our study responds to the grand challenge posed by Sharma et al. (2018) calling for a better understanding of the impacts of precipitation changes on floods.

Modeling studies can address changes in the FGP, allowing for a process-based analysis of projected changes in high flows. For example, Musselman *et al* (2018) examined changes in the rain-on-snow (ROS) FGP across the western U.S. due to projected climate change and calculated the change in ROS-generated rain and snowmelt water available for runoff. However, they highlighted as potential future work the use of a hydrologic model to account for the role that soil plays in changing high flows. Sharma *et al* (2018) also encouraged further evaluation of the role of the soil column in mediating extreme runoff, particularly under climate change.

This paper aims to understand how climate change will impact both the prevalence of different FGPs and the magnitude of high flows generated by these FGPs. Furthermore, this paper seeks to quantify the "sensitivity" of extreme streamflow generated by distinct FGPs to changes in climate. The streamflow response to changes in temperature or precipitation has been called "sensitivity" (Milly et al., 2018; Nijssen et al., 2001; Vano et al., 2012) or "scaling" (Wasko & Sharma, 2017). In this paper we will use the term "sensitivity." Sensitivities are easy-to-interpret estimates of the response of a chosen variable (in our case streamflow) to climate change. Vano et

al. (2015) applied the sensitivity approach to annual streamflow across the Pacific Northwest (PNW). Their study provided first-order approximations of hydrologic impacts to aid in scenario selection for decision making. These model experiments did not evaluate sensitivities of extreme high flows. Expanding the sensitivity approach to encompass extreme events can aid in climate change planning for water resources management.

All analyses are based on a large ensemble of climate change projections to provide a more robust analysis of changes in extreme events (van der Wiel et al., 2019). Previous studies detail how modeling choices can strongly affect climate change projections of mean streamflow values (Addor et al., 2014; Alder & Hostetler, 2019; Chegwidden et al., 2019; Velázquez et al., 2013). Here we use multiple modeling set-ups to better understand the effect of modeling decisions on high flows.

Our study covers a variety of hydroclimatic regions in the Columbia River Basin (Figure 3.1), which allows for the assessment of the effect of projected changes on different dominant FGPs. There is also a strong socioeconomic motivation for studying the Columbia River Basin, which is intensely managed by a variety of stakeholders with often competing interests. Operationally, the system of dams in the region prioritizes flood risk management under international treaty obligations that are under renegotiation at the time of this writing (Bankes, 2012).

3.2. METHODS

The methods here are built on Berghuijs *et al* (2016), Vano *et al* (2012, 2015), Musselman *et al* (2018), and Curry and Zwiers (2018). Berghuijs *et al* (2016) classified extreme flow events as originating from snowmelt, extreme precipitation, or excess soil moisture. We used similar flood classifications as Berghuijs *et al* (2016). Unlike their methods that related mean timing of

floods to mean timing of upstream processes, we linked individual high flow events to upstream hydrologic conditions at or before each specific event. We based our classifications on a set of hydrologic simulations distributed across headwater basins and evaluated how these classifications change under climate change. We classified high flows according to their FGP, anticipating that different FGPs might respond differently to changes in climate.

3.2.1. Base dataset

We used a large ensemble of hydrologic and climate change simulations for the Pacific Northwestern United States spanning the years 1950 to 2099 (Chegwidden et al., 2017) as our base dataset. As described in Chegwidden *et al* (2019) and River Management Joint Operating Committee (RMJOC; 2018), the simulations represent a multi-model ensemble of hydrologic projections, which is useful for representing uncertainty in projections of the hydrologic response to climate change (Bosshard et al., 2013; Vano et al., 2018; Velázquez et al., 2013; Vormoor et al., 2015). The climate projections from two representative concentration pathways (RCPs; mid-range greenhouse gas concentrations, RCP 4.5; high concentrations, RCP 8.5) and 10 global climate models (GCMs) were selected from the Coupled Model Intercomparison Project Phase 5 (Taylor et al., 2012). Climate data were downscaled to the spatial resolution of the hydrologic models (~6 km) using the Multivariate Adaptive Constructed Analogs (MACA) downscaling method (Abatzoglou & Brown, 2012). The four hydrologic models include three implementations of the Variable Infiltration Capacity model (i.e. VIC-P1, VIC-P2, VIC-P3) and one implementation of the Precipitation Runoff Modeling System (PRMS-P1). The resulting hydrologic simulations do not account for changes in land use or landcover, which can influence hydrologic extremes (Blöschl et al., 2007; Bronstert et al., 2002; Hurkmans et al., 2009). Further

details on the construction of the dataset can be found in the Supplementary Materials Section S2.1.

Most analyses to date have used gridded runoff fields taken directly from climate or hydrologic model simulations without routing this runoff through a stream network. We used routed streamflow instead to assess the aggregate impact at locations in the stream network. We analyze headwaters basins (most smaller than 4500 km^2 and all smaller than 10,000 km^2) because floods in small basins likely result from a single dominant FGP. Although, our results are based on unregulated flows and do not account for local channel and floodplain hydraulics, the findings can still aid in informing future management.

3.2.2. Study domain

We conducted our study across 21 headwater basins located in a variety of hydroclimates within the Pacific Northwestern United States (PNW, see Figure 3.1). Broadly, this region is characterized by winter-dominant precipitation (Hamlet et al., 2013). The precipitation falls on regions of complex topography and temperature regimes, creating a spectrum of snow-rain partitioning across the region. The range of hydroclimates creates a variety of ways in which snow mediates the timing of runoff in the region.

As a representative sample, we distributed the 21 headwaters basins among three different hydroclimatic regions: (1) temperate: seven basins in lowland, historically rain-dominant regions, (2) transition: seven basins on the leeward side of the Cascade Range, characterized as basins where precipitation is projected to transition from snow to rain under climate change, and (3) montane: seven high-elevation basins in the Rocky Mountains where snow is generally less at risk under a changing climate (Chegwidden et al., 2019; Hamlet et al., 2013).

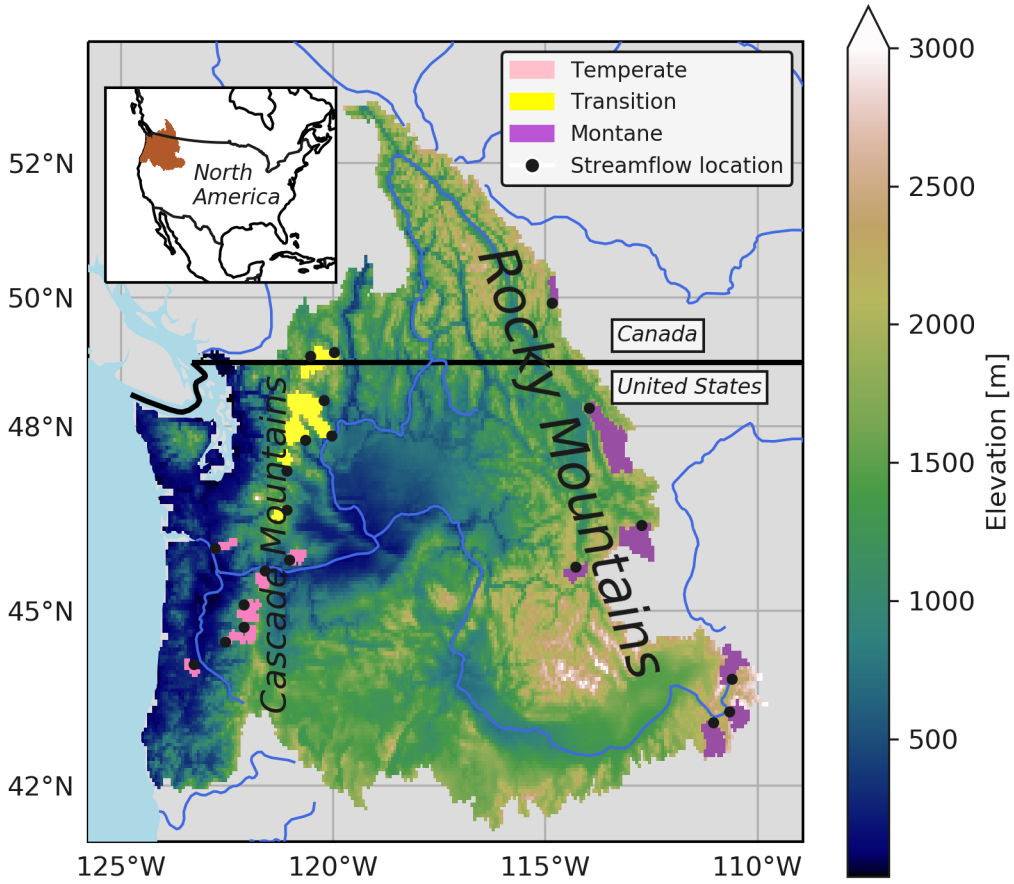


Figure 3.1. The 21 headwater study basins.

3.2.3. Classifying annual maxima by flood generating process (FGPs)

Our analyses focus on annual maximum flow (AMF) from 1951-2099. We calculated AMF as the highest daily flow to occur in any given water year (WY, October 1 – September 30). We specifically selected AMF as opposed to other high flow metrics to create a single value for each year in the study period. As if often done in the literature, we will use the term "flood" when discussing high flow, although we recognize that a river's AMF will not necessarily exceed its banks. Bankfull flow can be approximated by the one- to two-year return flow, and so the AMF is reasonably representative of flood stage.

We based our FGP classification generally on Berghuijs *et al* (2016) but modified the classification scheme to better suit our objectives. We determined the FGP directly for each AMF event, which is in contrast to Berghuijs *et al* (2016) who inferred FGP by aligning the mean timing of streamflow events with the mean timing of precipitation and snowmelt events. Determining the FGP accounts for potential changes in upstream climate and hydrologic states, which is important in a climate change study.

We first collated the AMFs with their corresponding basin-average states and processes (i.e. precipitation and snowmelt) for each model simulation during the period (1950 –2099). We then classified each AMF according to its dominant FGP via a decision tree. The definitions of each FGP and the classification tree are outlined fully in Section S2.3. The FGP classes were: (1) Precipitation driven, (2) snowmelt driven, (3) Rain-on-snow (ROS) driven, (4) Other. We limited our analyses to a small number of classes to create a sufficiently large sample size for each classification in our sensitivity analysis (Section 2.4). While many hydrologic events result from a diversity of upstream processes (e.g. rain-on-snow at high elevations combined with rain at low elevations), we identified a single, dominant FGP for each high flow event. Accordingly, we selected small headwater basins so that a single upstream FGP was likely to be dominant. We then identified how these FGPs changed under climate change.

For regional analyses, we calculated the portion of events for all simulations at all sites in each region caused by each FGP for each year. This paper does not focus on the "Other" class as it played a minor role overall in our study basins.

3.2.4. Calculating sensitivities

Vano *et al* (2012) estimated sensitivities of annual streamflow to changes in annual temperature and precipitation. We instead estimated sensitivities of high flows to changes in annual temperature and precipitation and took the additional step of calculating sensitivities of high flow conditional upon their FGPs.

For each location, we calculated the ensemble's projected sensitivities to changes in annual temperature and precipitation according to the methods described in Section S2.4 of the Supplementary Materials. Changes in climate were calculated as changes in mean annual precipitation and temperature between the control (WY 1970-1999; "1980s") and future (WY 2070-2099; "2080s") periods for the area upstream of each streamflow location for the years classified by each FGP.

We purposefully base our analyses on annual changes in temperature and precipitation for three reasons. First, we want to provide a first-order understanding of the sensitivity of high streamflow to climate change to support stakeholders with limited access to computing resources or costly modeling studies. These first-order understandings also align well with the existing sensitivity literature which is often with respect to mean annual changes. A similar approach was used by Curry and Zwiers (2018) to investigate linear relationships between high flows and hydroclimate predictors based on the observational record. Second, we reiterate the reasoning of Curry and Zwiers (2018) that the annual precipitation signal is less variable than the daily extremes and thus the sensitivity approach is more robust when calculated upon the annual values.

3.3. Results and Discussion

3.3.1. High flows will increase in magnitude with smaller subbasins experiencing larger increases

As detailed in Queen *et al* (in review), AMFs are projected to increase across the PNW at almost all locations under any scenario, in general agreement with (Tohver et al., 2014). Projected changes in AMFs from all 396 streamflow locations in the Chegwidden *et al* (2019) ensemble are shown in Figure A2.1 in the Supplementary Materials. As AMF magnitude (and, by proxy, basin size) decreases, the increases in AMF grow larger. The contrast between small and large basins aligns with the suggestion from Sharma *et al* (2018) that extremes will intensify more strongly in headwater basins. The significance of headwaters basins as being most susceptible to changes in flooding further motivates our focus on smaller basins in this study.

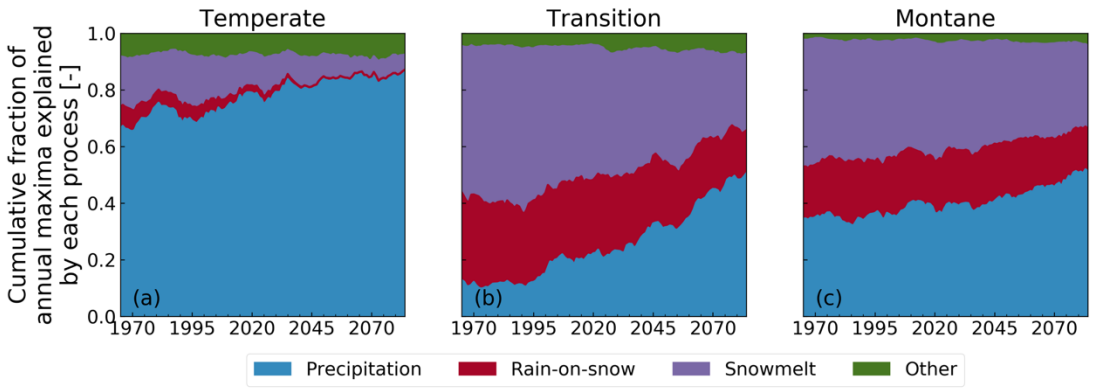


Figure 3.2 Changing proportion of annual maxima daily flow (AMF) attributable to each flood generating process under RCP 8.5. Proportions shown represent fractions of all AMFs from 30-year windows (centered on the year shown in the plot) across simulations from 10 GCMs and 4 hydrologic models from seven headwaters basins in each study region.

3.3.2. High flows will shift from snowmelt-driven to precipitation-driven

Snow accumulations are projected to decrease throughout the 21st century with projections of increased temperatures (Chegwidden et al., 2019; Elsner et al., 2010; Gergel et al., 2017; Hamlet et al., 2013). Accordingly, the overall role of snow in the hydrology of extreme events is projected to diminish, thus affecting the distribution of FGPs.

Changes in dominant FGP over time demonstrate the changing role that different processes play in the generation of AMF (Figure 3.2). Figure 3.2a shows that in the 20th century AMFs in temperate basins were mostly precipitation driven, with about 25% caused either by snowmelt or rain-on-snow (though mostly snowmelt). By the 2080s under RCP 8.5, ROS-driven AMFs nearly disappear and snowmelt-driven AMFs account for only about 5% of the total.

The transition basins will experience a large shift from snowmelt to precipitation-driven AMFs (Figure 3.2b). In the 20th century, precipitation is responsible for only 13% of AMFs but by the end of the 21st century it is responsible for 50% of AMFs. Transition basins also experience a small reduction in the frequency of AMFs caused by ROS (23% to 18%). When compared to the results for RCP 4.5 shown in Figure A2.2, we see that emission mitigation has its greatest impact in transition basins, where mitigation could reduce the share of precipitation-driven AMFs from 50% to 35% by the 2080s.

Our findings contrast with those of Musselman *et al* (2018) who found that, based on one set of simulations from one high-resolution climate model, ROS incidence would decrease at low elevation basins in the PNW but increase in higher elevation basins. In our study, we see a near disappearance of ROS AMFs in the temperate (lower elevation) region, and a small decrease in higher elevation regions (Figure 3.3c), likely owing to the overall reduction in snow presence. Besides climate model differences, a key distinction between our studies is that the analysis here addressed all FGPs but only for AMF, while Musselman *et al* (2018) investigated ROS events only but for all event magnitudes. Changes in smaller magnitude ROS events are not captured by our analyses, which focus on the mechanisms behind the largest events.

As noted in the Introduction, the choice of hydrologic model can substantially impact model results. Figure 3.3 shows the proportion of AMFs triggered by each of the three major FGPs

for each of the three hydro-climatic regions and separated by hydrological models. The control and future (RCP 8.5) periods are shown with open circles and filled triangles, respectively. Simulations from every GCM-hydrologic model-headwater basin permutation are then plotted on the ternary graphs. The results from the four different models are shown in different colors, with the marginal distributions of control (dashed lines) and future (solid lines) displayed along the axes of the ternary diagrams. The mean behavior of the four different hydrologic model results is shown in the differently colored bold circles and triangles with connecting lines.

Generally, the hydrologic models agree on their patterns of shifting from snow-driven to precipitation-driven AMF, as seen by the shift of every circle (control) toward the lower right part of the plot for its corresponding triangle (future) in the mean. Nevertheless, we note several nuances. First, there can be a large spread in the FGP partitioning depending on the region. For example, in the montane region while the means of the different hydrologic models are very tightly clustered, the spreads of the simulations are wide. The transition region shows the greatest hydrologic model diversity, with PRMS behaving markedly differently from the three VIC implementations with more precipitation- and ROS-driven AMFs at the expense of snowmelt events.

We also note that hydrologic model can have a strong impact on the attribution of the process causing the AMF. For example, in every region the PRMS implementation consistently exhibits fewer snowmelt and ROS AMFs in favor of precipitation-driven events. This is a result of a simpler elevation representation in PRMS compared to VIC which allows for high elevation snowfall in regions of complex topography. As another example, the VIC-P1 simulations in the temperate region show more snowmelt-driven AMFs, due to a less flashy hydrograph in those simulations, decreasing the peak of the precipitation events causing the snowmelt peak to be the

year's highest event. Broadly, however, the relative behaviors of the models are preserved under climate change: the relative positions of each marker on the ternary plot is consistent in control and future periods. For example, the PRMS simulations show less snow in all regions in the control period and continue to in the future.

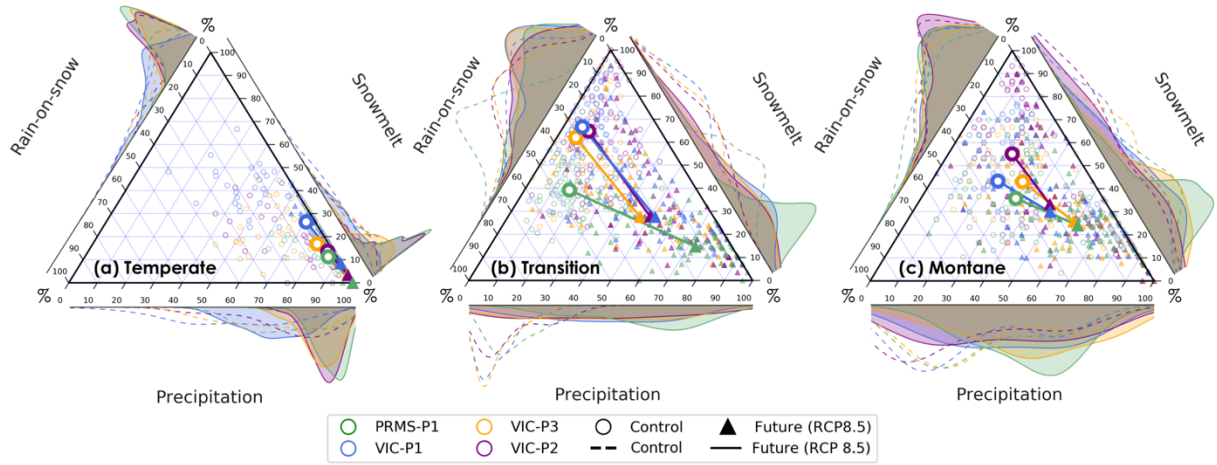


Figure 3.3. Partitioning of annual maximum daily flows for the control (1980s) and future (2080s; RCP 8.5) by their respective flood generation processes. Every marker represents a different (GCM-hydrologic model-headwater basin) permutation, with the larger markers indicating the median of each hydrologic model's simulations. The marginal distributions of those markers are shown along the axes. The marginal distributions were smoothed using a Gaussian kernel density estimator.

3.3. Projected changes in high flows: Precipitation-driven high flows will increase in magnitude more than snowmelt-driven AMFs

The changes in AMF for each FGP and their relationship to changes in climate are plotted in the scatter plots of Figure 3.4. Each row corresponds to a different hydro-climatic region and each column corresponds to a different FGP. Every dot represents, for each (ensemble member-hydrologic model-basin) combination, the mean change in temperature and precipitation along on the x-and y-axes, respectively, between control (1980s) and future (2080s) periods, with dots colored by the corresponding change in AMF magnitude. The mean changes in AMF for each FGP/region combination are shown in the upper left bar chart.

Grouping all hydro-climatic regions, precipitation-driven AMFs show the largest increases in flow magnitude, ranging between 29% and 36%. Increased AMF magnitude and increases in annual precipitation are connected because precipitation-driven AMFs are the result of fast-response runoff. With increased precipitation, soils become increasingly saturated, promoting fast-response flow so an individual precipitation event will be converted preferentially into runoff. ROS-driven AMFs show the smallest changes in magnitude (between -9 and +10%), likely because smaller snowpacks counteract the influence of elevated soil moisture that increases the magnitude of precipitation-driven AMFs.

Changes in snowmelt-driven AMFs depend highly on the hydroclimatic regime: while in the transition and montane regions we see average increases of 12% in magnitude, in temperate regions, AMFs increase by 33%. We can explain this difference using the same logic of the interpretation of the precipitation-driven AMFs above. The general pattern for a snowmelt-driven event includes a period of elevated soil moisture due to the melting of an upstream snowpack. Then, for any year's AMF to satisfy the classification of snowmelt-driven, elevated streamflow due to snowmelt must follow a (series of) precipitation event(s) or warm period triggering intense melt. Because snow plays a smaller role in the temperate region, the AMFs are more likely caused by precipitation events, and thus the sensitivities of the temperate snowmelt-driven AMFs behave more similarly to the precipitation-driven floods. The snowmelt-driven floods in the transition and montane regions experience less of the influence of elevated precipitation and show smaller average increases in AMF (~12%).

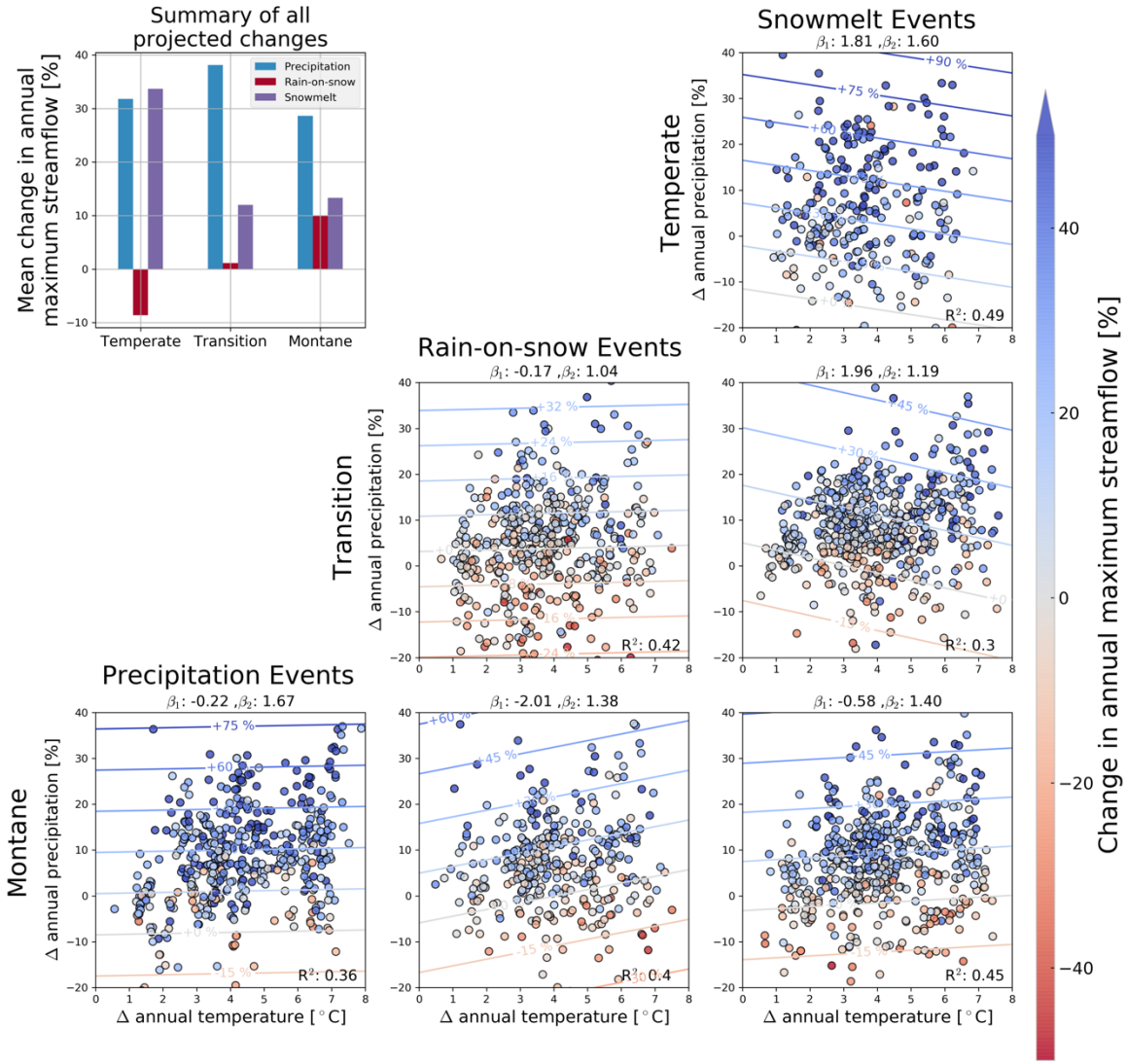


Figure 3.4. The dependence of changes in annual maximum daily flow on changes in mean annual temperature and precipitation between the control and future period. Columns show floods from different FGPs. Rows indicate different study regions. Every dot corresponds to a different climate forcing (RCP-GCM) permutation for a different streamflow basin. Isolines indicate a linear regression from Equation (3) with coefficients for temperature (β_1) and precipitation (β_2) shown in plot titles. Plots only shown in which regressions had an R^2 of greater than 0.3. The mean change of all region/FGP combinations is shown in the upper left bar chart.

3.4. Sensitivities of high flows to changes in climate: Up to 2% increase in flood magnitudes in response to 1% increase in annual precipitation

The contours on the panels of Figure 3.4 reflect the multilinear regression of changes in AMF (ΔAMF) on changes in mean annual temperature (ΔT) and annual precipitation (ΔP)

$$\Delta AMF = \beta_0 + \beta_1 \Delta T + \beta_2 \Delta P. \quad (3)$$

The coefficients β_1 and β_2 for each FGP/region combination (see Figure 3.4 panel titles) represent, respectively, the relative increase in AMF magnitude given a 1 °C increase in annual temperature and a 1% increase in annual precipitation. The contours in each panel represent the linear combination of the response in AMF magnitude over a range of changes in temperature and precipitation from Equation 3. As a simple linear regression, Equation 3 ignores interactions among predictor variables and non-linearities in the response of AMF to changes in climate, but aids in interpretability of the coefficients β_1 and β_2.

Sensitivities are positive for annual precipitation for all FGPs and regions. The region/FGP combinations with higher changes in AMF magnitude also have higher sensitivities. For example, precipitation-driven events in montane regions (average increase 28%; upper left panel) show a high flow sensitivity (β_2) of 1.67 %/% (bottom left panel) compared to the 1.40 %/% sensitivity of ROS-driven AMF in montane regions (which only increase on average by 9%). We note that the temperature sensitivity is more variable across FGPs and regions, with the strongest coefficient being -2.0 % / °C. Because the units of changes in temperature [°C] and precipitation [%] are not directly comparable, we cannot say that one scenario is "twice as sensitive to changes in precipitation as changes in temperature." Nevertheless, we can conclude that increases in annual precipitation show a much stronger relationship with increases in AMF than changes in annual temperature. Comparing the slopes of the contours between panels in Figure 3.4 allows us to see whether changes in AMF are more strongly related to temperature for one FGP compared to another, or in one hydroclimate compared to another.

Changes in mean climate explain, at most, 49% of the variance in changes in AMF (as seen by the R^2 values in Figure 3.4), leaving the majority of the variance unexplained. Still, we are able to derive useful first-order relationships between changes in high flows and changes in mean climate. For example, the sensitivity approach explains ~40% of the variance in montane regions and changes in snowmelt events. We also find that annual precipitation change is a stronger driver of changes in AMF than changes in annual temperature. For example, the montane region shows almost no temperature sensitivity, potentially because AMFs occur during wet, cold times of the year when ET is low and the changes in temperature are not sufficient to cause a large change in precipitation phase. The two sensitivity regressions with the strongest temperature components (see precipitation-driven events in temperate and transition basins) have among the lowest R^2 of those studied. The one example of a relationship with a stronger temperature sensitivity and an R^2 above 0.3 is snowmelt events in temperate basins.

3.4. CONCLUSIONS

While AMFs are projected to increase almost universally across the Pacific Northwest, there are distinct differences among flood generating processes (FGP). With reductions in snow, temperate and transition zones will experience AMFs which are precipitation-driven instead of snowmelt-driven. These precipitation-driven events show the greatest sensitivity in their magnitudes with changes in climate and will on average be 35% higher in magnitude. The reasoning behind this higher sensitivity is intuitive: precipitation-driven floods originate from fast-response runoff which has less opportunity to be delayed by other mediating processes like infiltration or evapotranspiration and sublimation. This supports Wasko and Nathan (2019) who found that while soil moisture trends can affect less extreme high flow events, higher flow events (including AMF) are mainly controlled by precipitation changes. By the end of the 21st century,

ROS events will likely have greater magnitudes but will have a smaller role in controlling AMFs because of a reduced prevalence of ROS. The patterns of shifts from snow to precipitation-driven dominance are largely robust across hydrologic models, though the ability of the models to represent sub-grid scale snow processes, such as multiple elevation zones within each model grid cell, influences the rate of this shift. We also show that a simple linear regression of changes in high flows against changes in mean annual temperature and precipitation can explain about 40% of the variance in the high flow changes. This suggests that a simple linear relationship can provide a first-order analysis of changes in high flows. Further, because of the interpretability of the linear regression framework, we can deduce sensitivities of flood magnitudes ranging between 1 and 2% increases for every 1% increase in annual precipitation. We believe that these are the first published numerical estimates of flood magnitude sensitivity to changes in annual climate and contribute toward the grand challenge posed by Sharma *et al* (2018).

Finally, a comparison of the differences between the changes in prevalence of each FGP of RCP 4.5 and RCP 8.5 climate scenarios shows the potential for climate change mitigation to limit the risk of increased high flows. Based upon these analyses, a warmer world will experience increased precipitation-induced flooding, which also shows a greater sensitivity to changes in climate. This finding presents two interrelated benefits of emissions mitigation: (1) we mitigate flood risk by preserving snowmelt-triggered floods which (2) have a lesser sensitivity to the warming that does occur.

Chapter 4 CLIMATE CHANGE HOT SPOTS OF WATER STRESS AND HYDROPOLITICAL CONFLICT

This chapter is in preparation to be submitted to Nature Climate Change. The supplemental material for this chapter is provided in Appendix C.

Chegwidden, O. S., N. Cristea, and B. Nijssen 2020: **Climate change hot spots of water stress and hydropolitical conflict.** *Nature Climate Change*, Manuscript in Preparation.

4.1. INTRODUCTION

Anthropogenic climate change is very likely to alter timing and volumes of water availability in basins around the world and to impact water resources (Arnell, 1999; Schewe et al., 2014; Vörösmarty et al., 2000). Anticipating these changes will help society sustainably adapt by implementing conservation measures, building water reuse infrastructure, considering additional storage capacity, and building alliances among users of the same water source (Vano et al., 2018).

Three hundred and ten large basin in the world straddle at least one international border (McCracken & Wolf, 2019). Many of these international basins have experienced hydropolitical discord ranging from treaty disagreements to outright military conflict (De Stefano et al., 2010; TFDD, 2016). Future changes in available freshwater from these basins have the potential to strain international relations (De Stefano et al., 2017). To mitigate risks from this threat, the United Nations recommended collaborative water management in their sixth Sustainable Development Goal (United Nations, 2016).

The earth science modeling community has focused extensive research efforts on projecting water availability under climate change. Many studies have assessed changes in precipitation as a proxy for changes in water availability (Trenberth et al., 2003, 2014). While this kind of analysis may suffice at locations far upstream in a basin, water resources are not always local, with rivers potentially impacting downstream communities thousands of kilometers from the source precipitation. Further, while precipitation delivers water to the land surface, water availability is ultimately determined by an interplay of land surface processes including evapotranspiration, snow processes, and runoff. Accordingly, it is necessary to analyze not just changes in precipitation, but changes in runoff, as some studies have begun to do (Cook et al., 2020; Ukkola et al., 2020).

Spatially-distributed runoff from the land surface is subject to two key overlapping geographic boundaries: (1) basin boundaries that determine via which river runoff is conveyed from land to oceans and (2) political boundaries that divide the river according to national borders. Both geographic boundaries are relevant to determining how changes in runoff might impact future hydropolitical tension. A handful of hydropolitical studies have assessed the susceptibility of river basins to climate change (Best, 2019; Farinosi et al., 2018; De Stefano et al., 2017). However, their analyses did not achieve the spatial specificity necessary to account for international, subbasin differences in hydrologic responses and upstream/downstream connections. Further, no hydropolitical analysis has been performed on an ensemble of climate change projections, which has been shown to be critical for a robust assessment of climate change impacts (Vano et al., 2018).

With the release of results from the Coupled Model Intercomparison Project Phase 6 (CMIP6; Eyring et al., 2016) the global climate modeling (GCM) community has provided a rich source of hydrologic climate impacts information. The number and sophistication of GCMs has grown, providing a diversity of ways of representing the land surface and its hydrologic sensitivity to

changes in climate. Model choices can have a large influence on the projected changes in hydrology (Addor et al., 2014; Chegwidden et al., 2019; Clark et al., 2016; Prudhomme & Davies, 2009; Vano et al., 2018). Analyzing the large ensemble of projections from CMIP6 supports robust hydrologic impact assessments. This is of critical importance, particularly since early results indicate that CMIP6 has a higher climate sensitivity indicating a wider range of possible climate impacts in the 21st century (Zelinka et al., 2020).

Accordingly, this paper will address two key questions:

(1) According to runoff estimates from CMIP6, how will climate change impact water availability through the 21st century?
(2) How are these projected changes overlaid on existing hydropolitical tensions?

With this study we aim to highlight regions meriting more localized study using finer resolution climate projections and well-calibrated hydrologic modeling. We do not intend to provide projections of changes in hydrology for planning purposes. We restrict our analyses to surface water availability and do not address water quality, water demand, or water management, all of which greatly influence practical water availability. However, this analysis is intended to represent a broad study of the water quantity as expressed through changes in water availability. Given that deteriorations of water quality (Whitehead et al., 2009) and increasing demand (Kummu et al., 2016) are likely to exacerbate access to water, we see our results as a lower bound of potential risks. Water management may alleviate some challenges of water availability for some countries, but can also elevate hydropolitical tensions, as discussed in this study. We also showcase the use

of the Pangeo data sharing and analysis architecture which was used for all analysis in this study (Eynard-Bontemps et al., 2019).

4.2. METHODS

We assessed runoff from 22 different GCMs as part of the Coupled Model Intercomparison Project Phase 6 (CMIP6; Eyring et al., 2016). Given the more limited availability of daily runoff data, we decided to use monthly-averaged runoff (i.e. 12 values per year) as the variable of interest. To test the robustness of the methods using monthly runoff, we evaluated the metrics on a subset of the GCM runoff data which were available at both monthly and daily timesteps. We found that the differences between monthly-averaged and 30-day running means were negligible (not shown).

4.2.1 Runoff metrics of interest

In this study we investigated the change in a variety of metrics which characterize the mean state and variability of the hydrology of every pixel in the global domain. We used the following metrics as measures of hydrology, all based upon the monthly averaged runoff.

- Mean annual runoff [m/year] – This metric was used to measure the overall water availability in a basin. From a hydrologic perspective, ignoring water reuse or groundwater extraction, changes in this quantity are a hard cap on the total amount of water available to basin water users annually.
- Hydrologic range (R_h) of monthly runoff [-] – This metric was used to represent how dynamic a river system is and is defined as follows:

$$R_h = \frac{1}{N}\sum_y^n \frac{\max_m(RO_{y,m}) - \min_m(RO_{y,m})}{\sum_{m=1}^{12} RO_{y,m}} \quad (1)$$

Runoff (*RO*) was analyzed for every year *y* at the monthly timestep (m=1,2,..12) resulting in 12 values for each year. The maximum and minimum monthly runoff were selected for each year *y* (i.e. evaluated across the *m* dimension) and then scaled by the annual sum of runoff for that year. Those values were then averaged across each period of length *N*=30.

Each metric was calculated as an average of each of the historical (1986-2015) and future (2070-2099) climatological periods. For the analyses shown here, all changes in metrics were calculated as differences between the future and historical periods. Changes between historical and future periods were reported as percentage differences (%) for annual mean and absolute differences [-] for hydrologic range. Thresholds for "no change" for annual mean were +/- 10% and for R_h were +/- 0.05. Otherwise, the changes were labeled as "increase" or "decrease".

4.2.2 Basin spatial averaging analyses

We conducted analyses for two different kind of polygon areas: river basin polygons and basin-country-unit (BCU) polygons. We used basin polygons from The World Bank (The World Bank, 2017). BCUs were constructed as follows: each river basin area was overlaid onto a map of national boundaries ("Countries WGS84," 2015). Each basin was then divided according to the boundaries of the member nations within that basin into its different national components called basin-country units (BCUs; De Stefano et al., 2017).

Basin-mean or BCU-mean hydrologic indices were calculated by spatially-averaging the gridded monthly runoff to a single timeseries for each polygon. We highlight that we use averaged runoff and do not route the streamflow. This is defensible since the travel time for runoff across even the largest basin is shorter than one month. GCMs were not regridded to a common grid before conducting the spatial averaging. All spatial mapping from grid to basin or BCU polygons

was conducted using the native GCM grid to preserve each individual GCM's response within each BCU polygon. Some polygons for certain BCUs were too small to cover an entire gridcell. For these cases, the single GCM gridcell closest to the centroid of the BCU was assigned to that BCU.

4.2.3 Use ERA5-Land Reanalysis as reference dataset

The ERA5-Land reanalysis product (Copernicus Cilmate Change Service (C3S), 2020) was used to determine the basins' runoff ratios in Figure 4.2. For this, the gridded monthly runoff and precipitation from 1986-2015 was spatially averaged to produce a single value of runoff and precipitation for each basin. The ERA5-Land product was also used as a reference dataset for ground-truthing of the annual runoff (Figure S1) and hydrologic range (Figure S2) metrics as calculated across CMIP6.

4.2.4 Upstream/downstream analyses for international basins

Of the 86 basins analyzed in Figures 4.1 and 4.2, 30 were selected for analysis within an international basin context. We selected the basins by first including those that were marked as "high risk" according to De Stefano et al. (2017) and then adding other international basins to provide at least three basins on every continent other than Australia.

By qualitative visual analysis, we assigned every BCU to a node in a tree network for each of the 30 international basins. Each BCU node is connected to a single downstream BCU. The most downstream BCU is considered the outlet BCU and has no downstream countries. Each BCU can have multiple immediate upstream countries and upstream BCUs are considered recursively (e.g. every upstream BCU is considered upstream of the outlet BCU).

The hydrologic analyses were then conducted across each BCU separately. For example, the hydrologic response of the Brazilian portion of the Amazon was calculated by conducting a spatial average of all gridcells from each GCM falling within both Brazil and the Amazon polygons.

For each of the 22 GCM results, the change in each BCU was calculated and then classified as follows. The change in mean annual runoff was classified as an "increase" or "decrease" if it changed by +/- 10%. A change in hydrologic range was classified as an "increase" or "decrease" if it changed by +/- 0.05. Otherwise, the BCU was determined to receive "No change". Then, the change in each BCU was compared with the changes in upstream or downstream BCUs. Each pairwise connection was then assigned to one of the boxes in the matrix detailed in the upper left of Figure 4.3. For example, within a given basin as determined by a given GCM, an increase in an upstream BCU with a decrease in a downstream GCM would be added to the lower left of the matrix. This analysis was conducted on every upstream/downstream connection for each GCM in a given basin.

Then, the counts in each cell of each matrix were normalized such that the total of the matrix cells sums to 1. Thus, we present a single matrix for each metric for each transboundary basin, describing the portion of upstream/downstream connections in each basin with different orientations of projected hydrologic changes. The matrices describing four basins for each metric are shown in Figure 4.3, with the full set of 30 transboundary basins shown in Figures S1 and S2.

4.2.5 Integration with hydropolitical risk dataset

Hydropolitical risk information was acquired from the Transboundary Water Assesment Programme's (TWAP) River Basin online data portal on April 16, 2020 (UNEP-DHI, 2016). The

risk indicator "Hydropolitical tension" (TWAP Risk Indicator 11 in the Summary Indicator spreadsheet) was used for the analyses of existing hydropolitical tension in Figure 4.4. Basin populations for Figure 4.4 were also derived from the TWAP basin database.

4.3. Results and Discussion

4.3.1 Most basins will be wetter with more variable water

We analyzed changes in area-averaged surface water ("runoff") for international basins around the world based on results from the shared socioeconomic pathways (SSP) 5-85 from CMIP6, which offers continuity with the RCP 8.5 scenario from CMIP5(O'Neill et al., 2016; Taylor et al., 2012). The SSP 5-85 scenario represents unmitigated emissions for the 21st century with a high level of competition among nation states. O'Neill et al. (2016) recommend prioritizing this scenario for impacts assessments.

We conducted our analyses of water availability in basins because rivers aggregate upstream responses. From an engineering perspective, runoff generated just outside of a basin boundary is irrelevant for inhabitants inside that basin. Simultaneously, as an example, while the headwaters of the Nile are thousands of miles from downstream Egypt, what occurs in these headwaters is relevant for downstream users. Thus, it is critical to assess hydrologic changes within the context of basin areas. Using monthly runoff estimates from the outputs of 22 global climate models (those for which monthly runoff was available at time of analysis) forced by the SSP 5-85 scenario, we assessed how water availability will change in a selection of 86 large basins around the globe (as defined by having a basin area covering at least 10 GCM gridcells according to the 1°x1.25° resolution CESM2 GCM grid).

We assessed changes in spatially-averaged mean annual runoff and hydrologic range. In this paper we define the metric of "hydrologic range", R_h, to represent the overall dynamism of the seasonal runoff cycle in a basin. As a unitless quantity, it is the difference between the maximum and minimum basin-averaged runoff as a fraction of the total annual runoff. For example, an R_h-value of 1 would mean that all runoff occurs in a single month, while an R_h-value of 0 indicates that runoff is constant throughout the year. Together, these two metrics provide a general understanding of the seasonal cycle of a basin's water resources. Changes in these two metrics might require a re-evaluation of current water management practices, as decreases in annual mean runoff suggest potential water shortages and increases in hydrologic range indicate increased floods and droughts and/or may require additional storage capacity. For both annual runoff and hydrologic range, the results of changes within a basin are spatially-averaged so as to weight changes in wetter portions of the basin more strongly than drier portions of the basin. See the Methods section for complete details on how to replicate this study.

4.3.2 Europe, the Americas, and southern Africa are hot spots

Median changes in annual runoff and hydrologic range are shown in Figure 4.1. The median historical (average of 1985-2015) runoff estimates are shown in the left panel of each figure. Runoff is higher in the tropics, aligning with the intertropical convergence zone, where runoff can exceed 2 m per year.

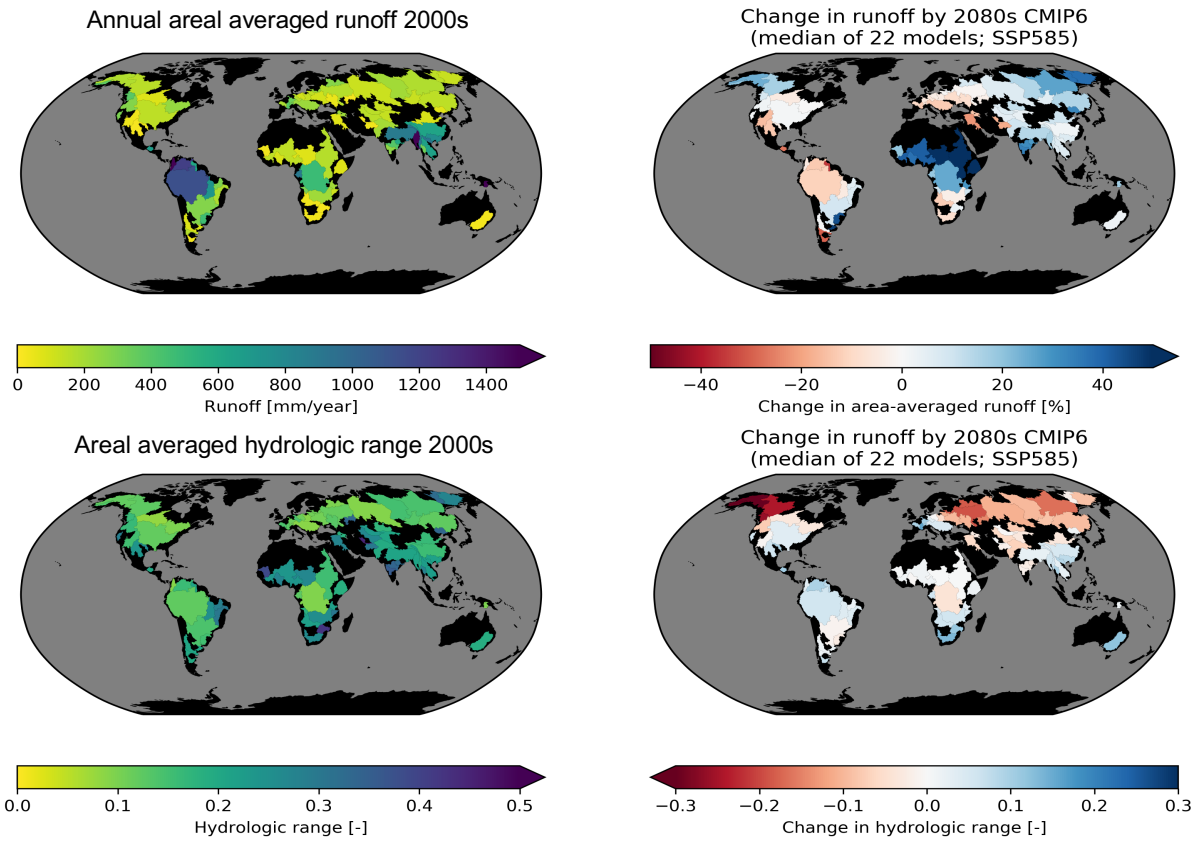


Figure 4.1 Mean annual runoff ([m/year]; top row) and hydrologic range ([-], bottom row). Baseline (mean of 1986-2015) shown on left with median projected change from 22 GCM models in the CMIP6 project (SSP 5-8.5) by the end of the 21st century (on right mean 2070-2099).

The response of basins by the end of the century varies throughout the world but certain spatial patterns arise. Runoff is projected to increase in Central Africa, the high latitudes and southeast Asia and projected to decrease in South America, North America and Europe. These changes in runoff fall largely within the +/-25% range, excepting the endorheic Lake Chad basin where the ensemble projects a doubling of runoff. This basin is very arid and thus a large percentage increase in runoff does not necessarily indicate a great increase in amount of water.

4.3.3 Hydrologic range decreases in the Arctic

The historical representation of the hydrologic range within the 22 GCMs is shown in the bottom left of Figure 4.1, with the projected changes by the end of the 21st century in the bottom right. High latitudes and cold mid-latitude basins see a distinct decrease in their hydrologic range by the end of the 21st century. This is a direct result of a lengthening of the Arctic summer resulting in a longer hydrologically active season with diminished snow cover (Dankers & Christensen, 2005). With a smaller snow influence under a warmer climate we see distinctly less dynamic river systems in cold basins, aligning with other studies of Arctic rivers (Dankers & Christensen, 2005). Elsewhere, particularly in tropical basins, we see increases in the hydrologic range. This is likely a result of the intensification of the hydrologic cycle resulting in elevated precipitation intensity and an increasingly intense monsoon (Ha et al., 2020; Kutzbach, 1981; Xin et al., 2020).

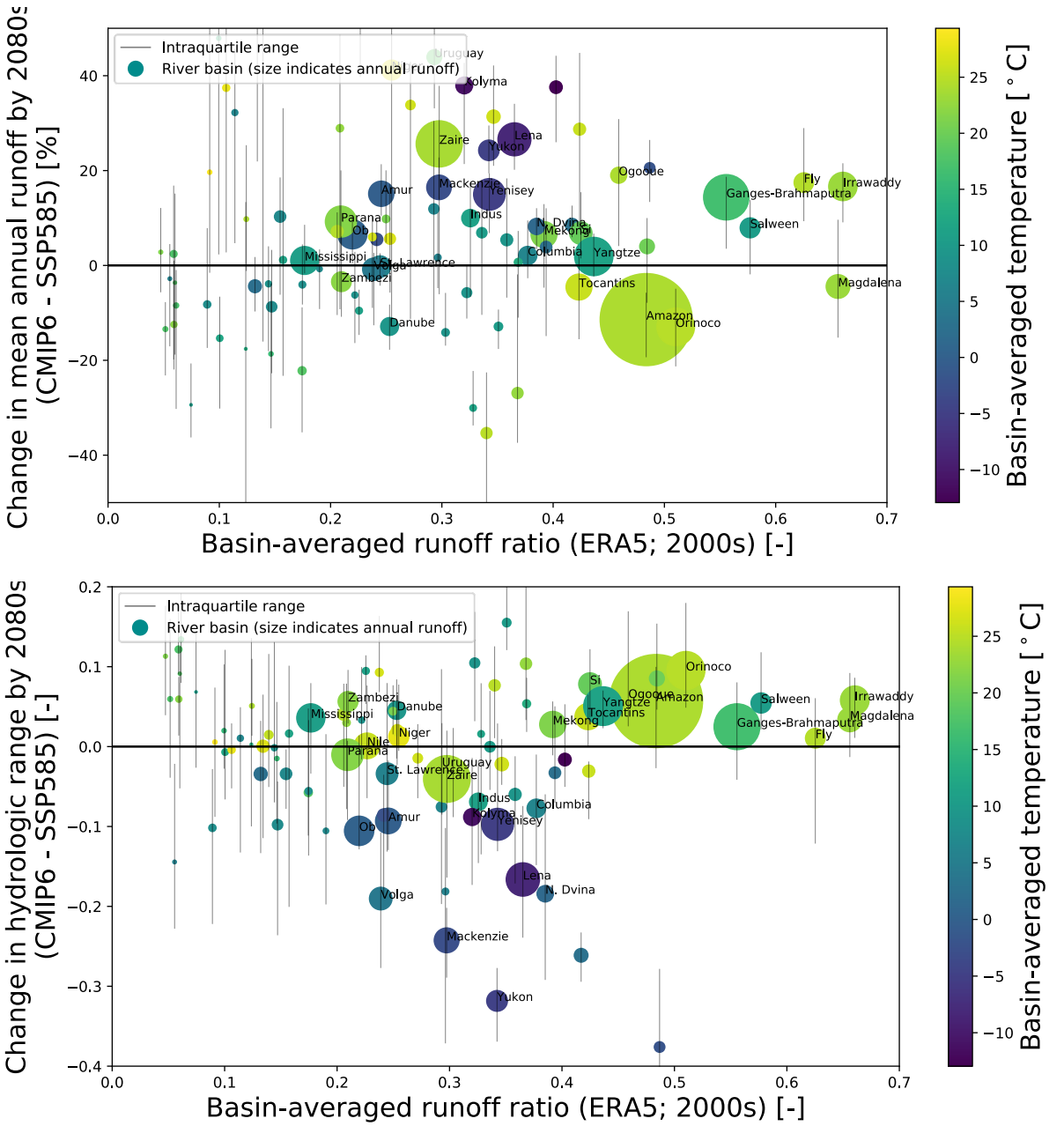


Figure 4.2 Full ensemble of changes in basin mean annual runoff (top, [%]) and hydrologic range (bottom, [-]). Dots are located vertically at the median of the ensemble, with the ensemble interquartile range represented by a grey vertical line. Dots are located horizontally according to their basin-averaged runoff ratio according to the ERA5-Land reanalysis product, with moisture-limited basins toward the left of the plots. The size of the dot corresponds to the total annual volumetric runoff from the basin and the color indicates the basin areally-averaged runoff from ERA5. The 30 basins with the highest annual runoff are labeled.

4.3.4 Changes in runoff are less certain in drier climates

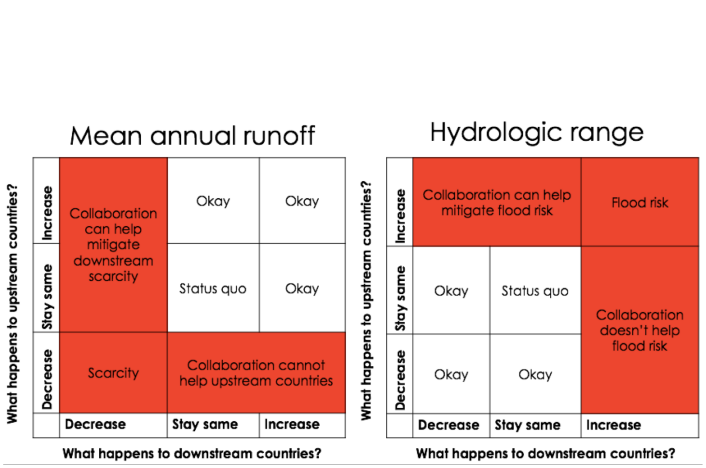
All of the analyses above were of median changes in runoff from an ensemble of 22 GCMs. While there are patterns in the ensemble's median response, the interquartile range can be quite large. In Figure 4.2, we see the interquartile range of percentage changes in annual runoff (top) and hydrologic range (bottom) for each of the river basins in Figure 4.1. Each basin is represented by a dot, its color indicating the mean basin-averaged temperature (as determined by the ERA5-Land (Copernicus Cilmate Change Service (C3S), 2020)) and its size representing the total basin runoff (basin area × annual area-averaged runoff). The horizontal location of each basin represents its runoff ratio (the fraction of annual precipitation which results in runoff), which indicates how efficiently the basin produces runoff from its input precipitation. We see that basins to the left in the plot show much larger error bars, indicating a high sensitivity of hydrology in dry, moisture-limited basins.

4.3.5 Most basins respond uniformly across national boundaries

Rivers in international basins flow across international borders. A variety of regional changes can affect hydropolitical tension such as changes in water demand or tensions arising from non water-related factors. In this analysis we assess changes in water resources within each country subregion within a transboundary basin (basin-country unit; BCU). The upstream/downstream orientation of the hydrologic response of different BCUs can be consequential. For example, tension-inducing changes (in this paper defined as a decrease in annual mean runoff or an increase in hydrologic range) in an upstream country will affect downstream countries, though not the reverse. Further, increased flood risk (represented here by an increased hydrologic range) in an upstream country might be eased by passing high water to a downstream country. On the contrary, increased water

scarcity (as represented by decreased annual mean) in an upstream country cannot be eased by a downstream country.

We calculated median hydrologic changes of the GCM ensemble in each BCU, assign those changes to either an "increase", "decrease", or "no change" category, and then compare changes between upstream/downstream countries (described in greater detail in the Methods section). We then assigned the changes from each upstream/downstream pair of BCUs to one of the nine boxes in the change matrices shown in the upper left of Figure 4.3. The position within the box corresponds to the change within each of the upstream/downstream pairs. The resulting matrices explain the overall orientation of projected changes in hydrology in the study basins. The key for the matrices in Figure 4.3 provides a qualitative explanation of what the overall changes could cause (e.g. if both upstream and downstream countries experience increases in hydrologic range, management of high flows could become more difficult leading to flooding). We expect that responses in the comparison matrices would cluster along the diagonal from lower left to upper right with the upstream and downstream BCUs responding similar. Deviations from the diagonal cells suggest potential challenges of international negotiation or conflict.



	Annual mean	Hydrologic range
Basin-wide increase	**"More water"** Mekong, Nile, Irrawaddy, Zaire, Jubba-Shebelle, Ganges-Brahmaputra, Niger, Indus, Uruguay	**"More variable hydrograph"** Amazon, Rhine, Essequibo, Irrawaddy, Salween, San Juan, Senegal, Parana, Danube
Basin-wide decrease	**"Less water"** Amazon, Colorado, Orinoco, Maritsa, San Juan, Vardar, Tigris-Euphrates, Danube, Essequibo,	**"Flatter hydrograph"** Ob

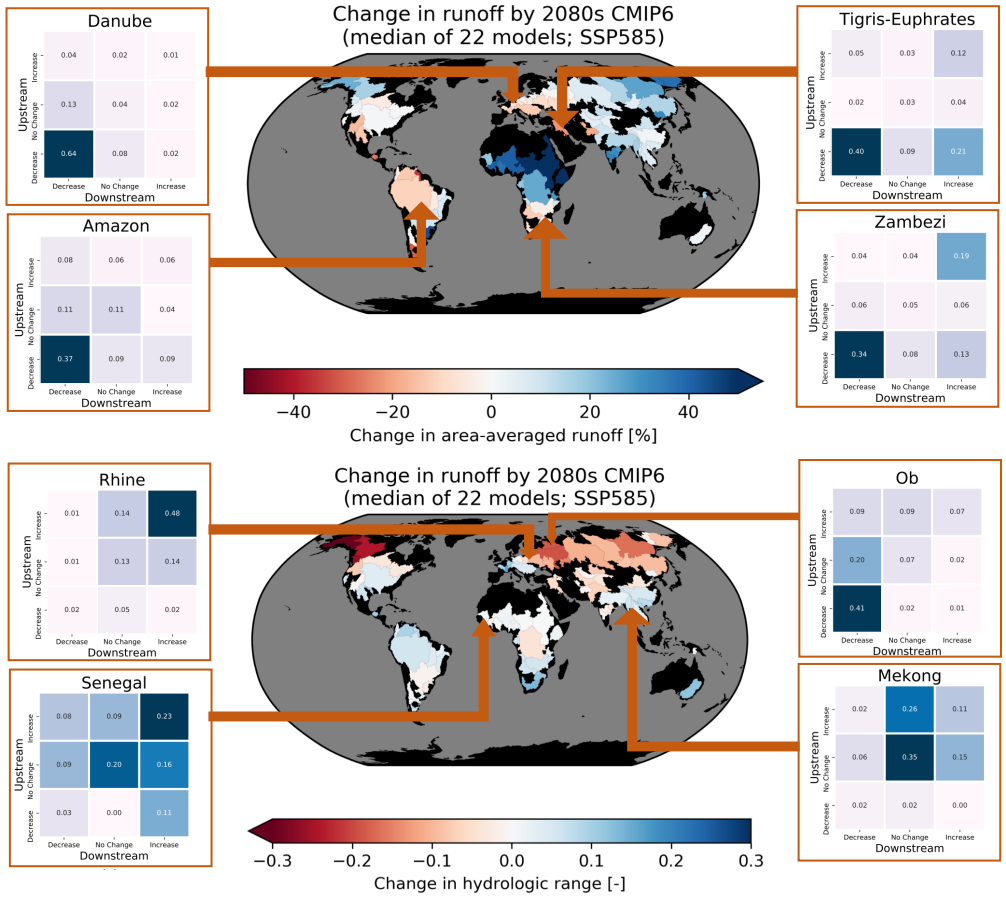


Figure 4.3 Cartoon explaining upstream/downstream connection matrices in upper left. Locations in the matrix corresponding to adverse changes are highlighted in red. Matrices from four sample basins for each hydrologic change metric: annual runoff (top two rows) and hydrologic runoff (bottom two rows). Cells are colored darker blue to indicate a higher proportion of that basin's ensemble categorized in that cell. Maps of changes in basins are same as those in Figure 2. Summary table at lower left categorizes basins with strong agreement of changes in upstream/downstream connections.

Four highlighted change matrices for each of the metrics are highlighted on the right side of Figure 4.3, annotating the maps in the same figure (see Figures S1 and S2 for the change matrices for all basins). The darkest blue cells correspond to the most common combinations among all BCUs in a basin and among all 22 GCMs. The CMIP6 ensemble projects that decreases will occur in both upstream and downstream countries within the Amazon and Danube basins. This homogeneity is in contrast to the Tigris-Euphrates basin and the Zambezi, where some downstream BCUs show projected increases in annual mean. Within the Zambezi those increases coincide with increases in upstream basins. However, within the Tigris-Euphrates, some downstream countries show increases while the upstream decreases.

Figure 4.3 bottom right shows the same analysis for hydrologic range. Some basins exhibit increases both upstream and downstream (the Rhine) or decreases both upstream and downstream (the Ob). The former would qualitatively suggest increased flood risk, while the latter indicates a flatter hydrograph with reduced chances of flood. The Mekong shows upstream nations exhibiting higher increases in the hydrologic range than downstream BCUs, which show “no change”. This pattern would suggest that downstream countries could perhaps shoulder some of the increased hydrologic range upstream. The Senegal shows a more varied response, generally filling the “increase” squares of the comparison matrix.

Any basin with a large portion of their upstream/downstream connections in the red boxes are worthy of further study. While this coarse resolution of study doesn’t provide actionable estimates of change, it can be useful for highlighting regions worthy of additional study.

4.3.6 Identifying hot spots of hydropolitical tension risk

The Transboundary Water Assessment Programme (TWAP) conducted a comprehensive analysis of the risks of hydropolitical tension in transboundary basins globally (UNEP-DHI, 2016). Their analysis combined information about (1) current vulnerability (as measured via lack of treaties or collaborative multinational relationships) with (2) documented future plans of large scale dam and water diversion projects (TFDD, 2016). The resulting numerical scale indicates the level of potential hydropolitical tension risk.

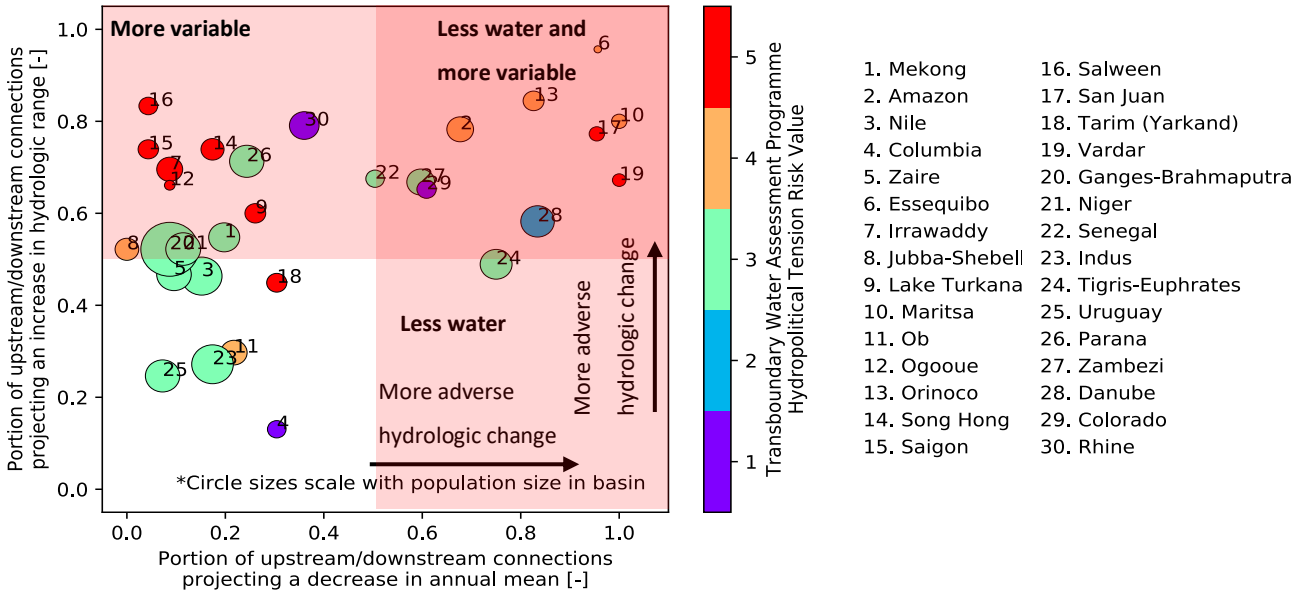


Figure 4.4 Portion of upstream/downstream connections from the full GCM ensemble which show adverse changes in hydrology. Location horizontally indicates the prevalence of projected decreases in annual runoff (prompting water scarcity) and location vertically indicates prevalence of increases in hydrologic range (prompting flood risk). Circle sizes scale with basin populations and colors indicate risk of hydropolitical tension as determined by the Transboundary Water Assessment Programme. A risk of 1 (purple) indicates a largely collaborative transboundary basin and 5 (red) indicates a basin with a high risk of internal conflict. If a basin falls within a red shaded portion of the plot, at least half of its ensemble of upstream/downstream connections are projected to experience an adverse hydrologic change (e.g. either a decrease in mean runoff or an increase in hydrologic range). If a basin falls within the upper right red portion, it is expected to experience less annual runoff with a higher hydrologic range.

In Figure 4.4 we combine the hydropolitical tension risk metric from the Transboundary Water Assessment Programme's (TWAP) with our analyses of projected changes in hydrology to identify transboundary hot spots. The location of each circle represents the portion of each basin's BCU pairs that is projected to experience adverse changes in hydrology (change in hydrologic range versus change in annual mean). The color of each basin's circle represents the TWAP hydropolitical risk score. For example, the Essequibo, Maritsa, Orinoco, San Juan, and Vardar basins all have high risks of hydropolitical tension. They also all show projections of decreased annual runoff and a more variable seasonal hydrograph. The increases in hydrologic range suggest a greater need for water management to regulate streamflow. Decreases in annual volume would impose an additional strain on the system, threatening water scarcity.

Meanwhile, the Song Hong, Saigon, Salween, and Irrawaddy Rivers also have high risks of hydropolitical tension but results from CMIP6 suggest that these basins will experience increases in hydrologic range and not decreases in annual runoff. Relatively safe spots also emerge from Figure 4.4: some medium-risk basins (e.g. the Indus and the Uruguay) are not at risk of adverse hydrologic change. These basins may, as a result, warrant less focused assessment from the standpoint of hydropolitical tension.

For all but seven of the basins studied, the hydrologic range is projected to increase in a majority of the ensemble of BCU pairs. In other words, it is very likely that most of the basins studied will experience a greater need for water management in the future due to an increased hydrologic range. Moreover, these changes are likely to occur in regions with high risks of hydropolitical tension, potentially exacerbating that tension in the future. As an additional aggravating factor, we note that some of these basins, notably the Vardar and San Juan, are highlighted by the TWAP program as being highly dependent economically on water (see Figure

S5 in supporting materials). The complicating factor of economic dependence could further elevate risk of tension.

Finally, we highlight that some basins (notably in southeast Asia) are expected to experience increased hydrologic range with increased annual runoff. In contrast to the adverse hydrologic changes we discussed above which promoted water scarcity, this combination of projected responses could exacerbate flood risk. These hot spots of flooding are worthy of finer scale hydrologic and reservoir modeling which better capture the spatiotemporal scales applicable for flood risk.

4.4. CONCLUSIONS

This study established hot spots of hydrologic change, and then contextualized those changes within existing risks hydropolitical tension. Using CMIP6 we revealed robust signals of adverse changes in hydrology. Most basins are projected to experience elevated total surface water accompanied by an increase in hydrologic range. Arctic river basins are a notable exception where the seasonal cycle of runoff is likely to dampen due a longer Arctic summer and reduced role of snow.

Combining potential changes in hydrology with existing potential risk reveals areas that are likely to experience exacerbated hydropolitical tension. Of note are European as well as Central and South American basins where increases in hydrologic range and decreases in total runoff are overlaid on existing high risks of tension. Several high-risk basins in Southeast Asia are expected to experience increases in hydrologic range.

We stress that our findings reveal patterns of changes, without providing actionable estimates for any one basin. Rather, we recommend that our results provide guidance of regions

worthy of more rigorous hydrologic climate impacts analyses. With enough lead time, actionable estimates of hydrologic change could support collaborative planning efforts, thereby reducing the risk of hydropolitical conflict.

Chapter 5 CONCLUSIONS AND RECOMMENDATIONS

5.1. CONCLUSIONS

This dissertation used ensembles of hydrologic climate change projections to evaluate changes in hydrology in multinational river basins around the world. Beginning in the Pacific Northwestern Columbia River, I conducted two rigorous hydrologic climate change assessments to better understand the nuances of a state-of-the-art hydrologic climate change ensemble. I then leveraged the most recent iteration of global climate model results to conduct a global analysis of changing water availability in transboundary basins. It is my intention that the findings from this last study (Chapter 4) will cycle back to inspire additional studies like those in Chapters 2 and 3 within other multinational river basins.

The findings of the dissertation, which answer the questions outlined in the introduction, are summarized in brief below. In Chapter 2 I evaluated the influence of modeling decisions on results from a large ensemble of hydrologic climate change projections. I found that in matters of timing and volume, climate projections (i.e. GCM and emissions pathways) were the most influential choices. In contrast, when assessing changes in low flows, the choice of hydrologic model can greatly influence projections. In Chapter 3, I dove deeper into the hydrologic climate change impacts within headwaters basins of the Pacific Northwest. I classified mechanisms causing flood events in the future and confirmed the diminishing role of snow as a flood generating process. I also calculated a first order sensitivity of high flows in headwaters streams to changes in climate, determining that increases of each 1% in annual precipitation can elicit up to ~1.4% increase in annual flow magnitudes. In Chapter 4 I expanded the hydrologic climate change analyses to the global scale, assessing impacts to runoff from multinational river basins. We identified hot spots of elevated hydropolitical risk worthy of more rigorous climate change study

as was done for the Columbia River Basin (Chegwidden et al., 2017; Hamlet et al., 2013; River Management Joint Operating Committee (RMJOC), 2018).

Throughout this dissertation I have showed spreads among climate projections of water resource availability and change. I note consistently that those spreads are wider in more arid regions. The consistency of this finding supports the claim that arid regions pose a more challenging environment for water resources planning. I stress that there are still examples in the literature of climate studies using only one model, but the material in this dissertation provides a strong warning against using single model studies. At the same time, I acknowledge that water resource planning agencies are often understaffed and under resourced. The findings in this dissertation address that challenge. Through the dissertation I have shown that it is possible to distill large modeling studies into useful, useable scientific findings and smaller, more translatable pieces of information, whether through supporting scenario selection (Chapter 2), developing cheaper, simpler flood relationships (Chapter 3), or by prioritizing future studies (Chapter 4).

5.2. RECOMMENDATIONS

Findings from this dissertation suggest that there is more work to be done in understanding model ensembles. All three chapters in this dissertation used model ensembles equally weighting each ensemble member. Other studies have incorporated clever schemes to weight each ensemble member according to its ability to replicate historical observations (Hawkins & Sutton, 2011). Such a weighting scheme could be useful in further hydrologic impacts work. Nevertheless, there persists one assumption at the root of the climate change analyses I present here. The hydrology models used in Chapters 2 and 3 were calibrated on the historical record and are assumed to

provide meaningful sensitivities for changes in climate. A fruitful area of research would be in probing the question of whether this assumption holds: are land surface model calibrations based upon the historical record appropriate for use under a different climate?

Chapter 4 presents an example of a push to begin using runoff fields directly from GCMs. Along with a handful of others, this approach is still in its nascent stages. Lehner *et al* (2019) propose one method of constraining runoff sensitivities using regional streamflow observations. Alternatively, expanding the results from Chapter 4 to include multiple ensemble members from different initial conditions from GCMs could provide a method of understanding sensitivities to different climates. Regardless, a rigorous assessment of hydrologic land surface representations within GCMs would support the use of hydrologic outputs moving forward. Lastly, Chapter 4 suggests a number of high-risk basins and regions each worthy of its own hydrologic climate change impacts assessment.

This dissertation brings to light a contrast between the wealth of resources available to study basins in richer parts of the world (e.g. the Columbia River Basin) compared to basins in regions with fewer resources (e.g. the Zambezi River Basin). This dissertation aims to resolve some of those challenges in resource availability and access. Chapter 2 provides guidance to stakeholders with fewer resources who might not be able to afford to evaluate a large ensemble of climate change projections. Chapter 3 provides a back-of-the-envelope assessment of changes in high flows based only on changes in climate, a shortcut without running a computationally expensive hydrologic model. And lastly, Chapter 4 highlights basins around the world where projected adverse changes in hydrology coincide with hydropolitical risk. In a world of limited computational resources, these findings might never come to light. I hope that these findings will

support, in particular, planning by communities which may not otherwise have the resources to do so.

BIBLIOGRAPHY

Abatzoglou, J. T., & Brown, T. J. (2012). A comparison of statistical downscaling methods suited for wildfire applications. *International Journal of Climatology*, *32*(5), 772–780. https://doi.org/10.1002/joc.2312

Abatzoglou, J. T., Rupp, D. E., & Mote, P. W. (2014). Seasonal climate variability and change in the pacific northwest of the united states. *Journal of Climate*, *27*(5), 2125–2142. https://doi.org/10.1175/JCLI-D-13-00218.1

Addor, N., Rössler, O., Köplin, N., Huss, M., Weingartner, R., & Seibert, J. (2014). Robust changes and sources of uncertainty in the projected hydrological regimes of Swiss catchments. *Water Resources Research*, *50*(10), 7541–7562. https://doi.org/10.1002/2014WR015549

Alder, J. R., & Hostetler, S. W. (2019). The Dependence of Hydroclimate Projections in Snow-Dominated Regions of the Western United States on the Choice of Statistically Downscaled Climate Data. *Water Resources Research*, *55*(3), 2279–2300. https://doi.org/10.1029/2018WR023458

Arnell, N. W. (1999). Climate change and global water resources. *Global Environmental Change*, *9*(1), S31–S49. https://doi.org/10.1016/S0959-3780(99)00017-5

Asadieh, B., & Krakauer, N. Y. (2017). Global change in flood and drought intensities under climate change in the 21st century. *Hydrology and Earth System Sciences Discussions*, 1–22. https://doi.org/10.5194/hess-2017-253

Bahr, D. B., Meier, M. F., & Peckham, S. D. (1997). The physical basis of glacier volume-area scaling. *Journal of Geophysical Research*, *102*(B9), 20355–20362. https://doi.org/doi:10.1029/97JB01696

Bankes, N. (2012). The Flood Control Regime of the Columbia River Treaty: Before and After 2024. *Washington Journal of Environmental Law & Policy*, *1*, 1–72.

Berghuijs, W. R., Harrigan, S., Molnar, P., Slater, L. J., & Kirchner, J. W. (2019). The relative importance of different flood-generating mechanisms across Europe. *Water Resources Research*, 2019WR024841. https://doi.org/10.1029/2019WR024841

Berghuijs, W. R., Woods, R. A., Hutton, C. J., & Sivapalan, M. (2016). Dominant flood generating mechanisms across the United States. *Geophysical Research Letters*, *43*(9), 4382–4390. https://doi.org/10.1002/2016GL068070

Bernardini, F., Enderlein, R., Koeppel, S., & Lipponen, A. (2012). Europe and North America. In *The United Nations World Water Development Report* (pp. 655–670).

Best, J. (2019). Anthropogenic stresses on the world's big rivers. *Nature Geoscience*, *12*(1), 7–21. https://doi.org/10.1038/s41561-018-0262-x

Blöschl, G., Ardoin-Bardin, S., Bonell, M., Dorninger, M., Goodrich, D., Gutknecht, D., … Szolgay, J. (2007). At what scales do climate variability and land cover change impact on flooding and low flows? *Hydrological Processes*, *21*(9), 1241–1247. https://doi.org/10.1002/hyp.6669

Bosshard, T., Carambia, M., Goergen, K., Kotlarski, S., Krahe, P., Zappa, M., & Schär, C. (2013). Quantifying uncertainty sources in an ensemble of hydrological climate-impact projections. *Water Resources Research*, *49*(3), 1523–1536. https://doi.org/10.1029/2011WR011533

Brakebill, J. W., Wolock, D. M., & Terziotti, S. E. (2011). Digital Hydrologic Networks Supporting Applications Related to Spatially Referenced Regression Modeling. *Journal of the American Water Resources Association*, *47*(5), 916–932. https://doi.org/10.1111/j.1752-1688.2011.00578.x

Bronstert, A., Niehoff, D., & Brger, G. (2002). Effects of climate and land-use change on storm runoff generation: Present knowledge and modelling capabilities. *Hydrological Processes*, *16*(2), 509–529. https://doi.org/10.1002/hyp.326

Bürger, G., Schulla, J., & Werner, A. T. (2011). Estimates of future flow, including extremes, of the Columbia River headwaters. *Water Resources Research*, *47*(10), 1–18. https://doi.org/10.1029/2010WR009716

Casola, J. H., Cuo, L., Livneh, B., Lettenmaier, D. P., Stoelinga, M. T., Mote, P. W., & Wallace, J. M. (2009). Assessing the impacts of global warming on snowpack in the Washington cascades. *Journal of Climate*, *22*(10), 2758–2772. https://doi.org/10.1175/2008JCLI2612.1

Chegwidden, O. S., Nijssen, B., Rupp, D. E., & Mote, P. W. (2017). Hydrologic Response of the Columbia River System to Climate Change [Data set]. Zenodo. Retrieved from doi:10.5281/zenodo.854763.

Chegwidden, Oriana S., Nijssen, B., Rupp, D. E., Arnold, J. R., Clark, M. P., Hamman, J. J., … Xiao, M. (2019). How Do Modeling Decisions Affect the Spread Among Hydrologic Climate Change Projections? Exploring a Large Ensemble of Simulations Across a Diversity of Hydroclimates. *Earth's Future*, *7*(6), 623–637. https://doi.org/10.1029/2018EF001047

Chen, J., Brissette, F. P., Poulin, A., & Leconte, R. (2011). Overall uncertainty study of the hydrological impacts of climate change for a Canadian watershed. *Water Resources Research*, *47*(12), 1–16. https://doi.org/10.1029/2011WR010602

Clark, M. P., Wilby, R. L., Gutmann, E. D., Vano, J. A., Gangopadhyay, S., Wood, A. W., … Brekke, L. D. (2016). Characterizing Uncertainty of the Hydrologic Impacts of Climate Change. *Current Climate Change Reports*, *2*(2), 55–64. https://doi.org/10.1007/s40641-016-0034-x

Cook, B. I., Mankin, J. S., Marvel, K., Williams, A. P., Smerdon, J. E., & Anchukaitis, K. J. (2020). Twenty-First Century Drought Projections in the CMIP6 Forcing Scenarios. *Earth's Future*, *8*(6), e2019EF001461. https://doi.org/10.1029/2019EF001461

Copernicus Cilmate Change Service (C3S). (2020). C3S ERA5-Land reanalysis. *Copernicus Climate Change Service*.

Countries WGS84. (2015). UNIGIS Geospatial Education Resources.

Curry, C. L., & Zwiers, F. W. (2018). Examining controls on peak annual streamflow and floods in the Fraser River Basin of British Columbia. *Hydrology and Earth System Sciences*, *22*(4), 2285–2309. https://doi.org/10.5194/hess-22-2285-2018

Dankers, R., & Christensen, O. B. (2005). Climate Change Impact on Snow Coverage, Evaporation and River Discharge in the Sub-Arctic Tana Basin, Northern Fennoscandia. *Climatic Change*, *69*(2), 367–392. https://doi.org/10.1007/s10584-005-2533-y

Davenport, F. V, Herrera-Estrada, J. E., Burke, M., & Diffenbaugh, N. S. (2020). Flood Size Increases Nonlinearly Across the Western United States in Response to Lower Snow-Precipitation Ratios. *Water Resources Research*, *56*(1). https://doi.org/10.1029/2019WR025571

Diffenbaugh, N. S. (2020). Verification of extreme event attribution: Using out-of-sample observations to assess changes in probabilities of unprecedented events. *Science Advances*, *6*(12), eaay2368. https://doi.org/10.1126/sciadv.aay2368

Dobler, C., Hagemann, S., Wilby, R. L., & StÃtter, J. (2012). Quantifying different sources of uncertainty in hydrological projections in an Alpine watershed. *Hydrology and Earth System Sciences*, *16*(11), 4343–4360. https://doi.org/10.5194/hess-16-4343-2012

Duan, Q. Y., Gupta, V. K., & Sorooshian, S. (1993). Shuffled complex evolution approach for effective and efficient global minimization. *Journal of Optimization Theory and Applications*, *76*(3), 501–521. https://doi.org/10.1007/BF00939380

Elsner, M. M., Cuo, L., Voisin, N., Deems, J. S., Hamlet, A. F., Vano, J. A., … Lettenmaier, D. P. (2010). Implications of 21st century climate change for the hydrology of Washington State. *Climatic Change*, *102*(1–2), 225–260. https://doi.org/10.1007/s10584-010-9855-0

Eynard-Bontemps, G., Abernathey, R., Hamman, J., Ponte, A., & Rath, W. (2019). The Pangeo Big Data Ecosystem and its use at CNES. In *Proc. of the 2019 conference on Big Data from Space (BiDS'2019)* (pp. 49-52 EUR 29660 EN). Luxembourg: Publications Office of the European Union. https://doi.org/10.2760/848593

Eyring, V., Bony, S., Meehl, G. A., Senior, C. A., Stevens, B., Stouffer, R. J., & Taylor, K. E. (2016). Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organization. *Geoscientific Model Development*, *9*(5), 1937–1958. https://doi.org/10.5194/gmd-9-1937-2016

Farinosi, F., Giupponi, C., Reynaud, A., Ceccherini, G., Carmona-Moreno, C., De Roo, A., … Bidoglio, G. (2018). An innovative approach to the assessment of hydro-political risk: A spatially explicit, data driven indicator of hydro-political issues. *Global Environmental Change*, *52*(March), 286–313. https://doi.org/10.1016/j.gloenvcha.2018.07.001

Finger, D., Heinrich, G., Gobiet, A., & Bauder, A. (2012). Projections of future water resources and their uncertainty in a glacierized catchment in the Swiss Alps and the subsequent effects on hydropower production during the 21st century. *Water Resources Research*, *48*(2), 1–20. https://doi.org/10.1029/2011WR010733

Gergel, D. R., Nijssen, B., Abatzoglou, J. T., Lettenmaier, D. P., & Stumbaugh, M. R. (2017). Effects of climate change on snowpack and fire potential in the western USA. *Climatic Change*, *141*(2), 287–299. https://doi.org/10.1007/s10584-017-1899-y

Gleick, P. H. (1989). Climate change, hydrology, and water resources. *Reviews of Geophysics*, *27*(3), 329–344. https://doi.org/10.1029/RG027i003p00329

Gupta, H. V., Kling, H., Yilmaz, K. K., & Martinez, G. F. (2009). Decomposition of the mean

squared error and NSE performance criteria: Implications for improving hydrological modelling. *Journal of Hydrology*, *377*(1–2), 80–91. https://doi.org/10.1016/j.jhydrol.2009.08.003

Gutmann, E., Pruitt, T., Clark, M. P., Brekke, L., Arnold, J. R., Raff, D. A., & Rasmussen, R. M. (2014). An intercomparison of statistical downscaling methods used for water resource assessments in the United States. *Water Resources Research*, *50*(9), 7167–7186. https://doi.org/10.1002/2014WR015559

Ha, K.-J., Moon, S., Timmermann, A., & Kim, D. (2020). Future Changes of Summer Monsoon Characteristics and Evaporative Demand Over Asia in CMIP6 Simulations. *Geophysical Research Letters*, *47*(8), e2020GL087492. https://doi.org/10.1029/2020GL087492

Hamlet, A. F., Elsner, M. M., Mauger, G. S., Lee, S.-Y., Tohver, I., & Norheim, R. A. (2013). An Overview of the Columbia Basin Climate Change Scenarios Project: Approach, Methods, and Summary of Key Results. *Atmosphere-Ocean*, *51*(4), 392–415. https://doi.org/10.1080/07055900.2013.819555

Hamman, J., & Nijssen, B. (2015). VIC 4.2.glacier. Retrieved from https://github.com/UW-Hydro/VIC/tree/support/VIC.4.2.glacier

Hamman, J., Nijssen, B., Roberts, A., Craig, A., Maslowski, W., & Osinski, R. (2017). The coastal streamflow flux in the Regional Arctic System Model. *Journal of Geophysical Research: Oceans*, *122*(3), 1683–1701. https://doi.org/10.1002/2016JC012323

Harding, B. L., Wood, A. W., Prairie, J. R., & Service, N. W. (2012). The implications of climate change scenario selection for future streamflow projection in the Upper Colorado River Basin, 3989–4007. https://doi.org/10.5194/hess-16-3989-2012

Hattermann, F. F., Vetter, T., Breuer, L., Su, B., Daggupati, P., Donnelly, C., … Krysnaova, V. (2018). Sources of uncertainty in hydrological climate impact assessment: A cross-scale study. *Environmental Research Letters*, *13*(1). https://doi.org/10.1088/1748-9326/aa9938

Hawkins, E., & Sutton, R. (2009). The potential to narrow uncertainty in regional climate predictions. *Bulletin of the American Meteorological Society*, *90*(8), 1095–1107. https://doi.org/10.1175/2009BAMS2607.1

Hawkins, E., & Sutton, R. (2011). The potential to narrow uncertainty in projections of regional precipitation change. *Climate Dynamics*, *37*(1), 407–418. https://doi.org/10.1007/s00382-010-0810-6

Hidalgo, H. G., Das, T., Dettinger, M. D., Cayan, D. R., Pierce, D. W., Barnett, T. P., … Nozawa, T. (2009). Detection and Attribution of Streamflow Timing Changes to Climate Change in the Western United States. *Journal of Climate*, *22*(13), 3838–3855. https://doi.org/10.1175/2009JCLI2470.1

Hoegh-Guldberg, O., Jacob, D., Taylor, M., Bindi, M., Brown, S., Camilloni, I., … Zhou, G. (2018). Impacts of 1.5ºC Global Warming on Natural and Human Systems. In *Global Warming of 1.5°C. An IPCC Special Report on the impacts of global warming of 1.5°C above pre-industrial levels and related global greenhouse gas emission pathways, in the context of strengthening the global response to the threat of climate change,*.

Hurkmans, R. T. W. L., Terink, W., Uijlenhoet, R., Moors, E. J., Troch, P. A., & Verburg, P. H.

(2009). Effects of land use changes on streamflow generation in the Rhine basin. *Water Resources Research*, *45*(6), 1–15. https://doi.org/10.1029/2008WR007574

Jiang, Y., Kim, J. B., Still, C. J., Kerns, B. K., Kline, J. D., & Cunningham, P. G. (2018). Inter-comparison of multiple statistically downscaled climate datasets for the Pacific Northwest, USA. *Scientific Data*, *5*(February 2018). https://doi.org/10.1038/sdata.2018.16

Jung, I. W., Moradkhani, H., & Chang, H. (2012). Uncertainty assessment of climate change impacts for hydrologically distinct river basins. *Journal of Hydrology*, *466–467*, 73–87. https://doi.org/10.1016/j.jhydrol.2012.08.002

Kirchhoff, C. J., Esselman, R., & Brown, D. (2015). Boundary organizations to boundary chains: Prospects for advancing climate science application. *Climate Risk Management*, *9*, 20–29. https://doi.org/10.1016/j.crm.2015.04.001

Knighton, J., Steinschneider, S., & Walter, M. T. (2017). A Vulnerability-Based, Bottom-up Assessment of Future Riverine Flood Risk Using a Modified Peaks-Over-Threshold Approach and a Physically Based Hydrologic Model. *Water Resources Research*, *53*(12), 10043–10064. https://doi.org/10.1002/2017WR021036

Köplin, N., Schädler, B., Viviroli, D., & Weingartner, R. (2014). Seasonality and magnitude of floods in Switzerland under future climate change. *Hydrological Processes*, *28*(4), 2567–2578. https://doi.org/10.1002/hyp.9757

Kummu, M., Guillaume, J. H. A., de Moel, H., Eisner, S., Flörke, M., Porkka, M., … Ward, P. J. (2016). The world's road to water scarcity: shortage and stress in the 20th century and pathways towards sustainability. *Scientific Reports*, *6*(1), 38495. https://doi.org/10.1038/srep38495

Kundzewicz, Z. W., Kanae, S., Seneviratne, S. I., Handmer, J., Nicholls, N., Peduzzi, P., … Sherstyukov, B. (2014). Flood risk and climate change: global and regional perspectives. *Hydrological Sciences Journal*, *59*(1), 1–28. https://doi.org/10.1080/02626667.2013.857411

Kutzbach, J. E. (1981). Monsoon Climate of the Early Holocene: Climate Experiment with the Earth's Orbital Parameters for 9000 Years Ago. *Science*, *214*(4516), 59–61. https://doi.org/10.1126/science.214.4516.59

Lafaysse, M., Hingray, B., Mezghani, A., Gailhard, J., & Terray, L. (2014). Internal variability and model uncertainty components. *Water Resources Research*, 3317–3341. https://doi.org/10.1002/2013WR014897.Received

Leavesley, G. H., Lichty, R. W., Troutman, B. M., & Saindon, L. G. (1983). Precipitation-runoff modeling system- User's manual: U.S. Geological Survey Water-Resources Investigations Report, *83*(4238), 207.

Lehner, F., Wood, A. W., Vano, J. A., Lawrence, D. M., Clark, M. P., & Mankin, J. S. (2019). The potential to reduce uncertainty in regional runoff projections from climate models. *Nature Climate Change*, *9*(12), 926–933. https://doi.org/10.1038/s41558-019-0639-x

Liang, X., Lettenmaier, D. P., Wood, E. F., & Burges, S. J. (1994). A simple hydrologically based model of land surface water and energy fluxes for general circulation models. *Journal of Geophysical Research: Atmospheres*, *99*(D7), 14415–14428. https://doi.org/10.1029/94JD00483

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

Lohmann, D., Nolte-Holube, R., & Raschke, E. (1996). A large-scale horizontal routing model to be coupled to land surface parametrization schemes. *Tellus, Series A: Dynamic Meteorology and Oceanography*, *48*(5), 708–721. https://doi.org/10.3402/tellusa.v48i5.12200

Mallakpour, I., & Villarini, G. (2015). The changing nature of flooding across the central United States. *Nature Climate Change*, *5*(3), 250–254. https://doi.org/10.1038/nclimate2516

Mandal, S., & Simonovic, S. P. (2017). Quantification of uncertainty in the assessment of future streamflow under changing climate conditions. *Hydrological Processes*, *31*(11), 2076–2094. https://doi.org/10.1002/hyp.11174

Mantua, N., Tohver, I., & Hamlet, A. (2010). Climate change impacts on streamflow extremes and summertime stream temperature and their possible consequences for freshwater salmon habitat in Washington State. *Climatic Change*, *102*(1–2), 187–223. https://doi.org/10.1007/s10584-010-9845-2

Matti, B., Dahlke, H. E., Dieppois, B., Lawler, D. M., & Lyon, S. W. (2017). Flood seasonality across Scandinavia—Evidence of a shifting hydrograph? *Hydrological Processes*, *31*(24), 4354–4370. https://doi.org/10.1002/hyp.11365

McCracken, M., & Wolf, A. T. (2019). Updating the Register of International River Basins of the world. *International Journal of Water Resources Development*, *35*(5), 732–782. https://doi.org/10.1080/07900627.2019.1572497

Megdal, S. B., Eden, S., & Shamir, E. (2017). Water governance, stakeholder engagement, and sustainable water resources management. *Water (Switzerland)*, *9*(3), 1–7. https://doi.org/10.3390/w9030190

Melsen, L., Addor, N., Mizukami, N., Newman, A., Torfs, P., Clark, M., … Teuling, R. (2017). Mapping (dis)agreement in hydrologic projections. *Hydrol. Earth Syst. Sci Disc*, (September). https://doi.org/10.5194/hess-2017-564

Mendoza, P. A., Clark, M. P., Mizukami, N., Newman, A. J., Barlage, M., Gutmann, E. D., … Arnold, J. R. (2015). Effects of Hydrologic Model Choice and Calibration on the Portrayal of Climate Change Impacts. *Journal of Hydrometeorology*, *16*(2), 762–780. https://doi.org/10.1175/JHM-D-14-0104.1

Milly, P. C. D., Kam, J., & Dunne, K. A. (2018). On the Sensitivity of Annual Streamflow to Air Temperature. *Water Resources Research*, *54*(4), 2624–2641. https://doi.org/10.1002/2017WR021970

Mizukami, N., Clark, M. P., Gutmann, E. D., Mendoza, P. A., Newman, A. J., Nijssen, B., … Brekke, L. D. (2016). Implications of the Methodological Choices for Hydrologic Portrayals of Climate Change over the Contiguous United States: Statistically Downscaled Forcing Data and Hydrologic Models. *Journal of Hydrometeorology*, *17*(1), 73–98. https://doi.org/10.1175/JHM-D-14-0187.1

Mizukami, N., Clark, M. P., Newman, A. J., Wood, A. W., Gutmann, E. D., Nijssen, B., …

Samaniego, L. (2017). Towards seamless large-domain parameter estimation for hydrologic models. *Water Resources Research*, *53*(9), 8020–8040. https://doi.org/10.1002/2017WR020401

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

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

Musselman, K. N., Clark, M. P., Liu, C., Ikeda, K., & Rasmussen, R. (2017). Slower snowmelt in a warmer world. *Nature Climate Change*, *7*(3), 214–219. https://doi.org/10.1038/nclimate3225

Musselman, K. N., Lehner, F., Ikeda, K., Clark, M. P., Prein, A. F., Liu, C., … Rasmussen, R. (2018). Projected increases and shifts in rain-on-snow flood risk over western North America. *Nature Climate Change*, *8*(9), 808–812. https://doi.org/10.1038/s41558-018-0236-4

Najafi, M. R., Moradkhani, H., & Jung, I. W. (2011). Assessing the uncertainties of hydrologic model selection in climate change impact studies. *Hydrological Processes*, *25*(18), 2814–2826. https://doi.org/10.1002/hyp.8043

Newman, A., Sampson, K., Clark, M. P., Bock, A. R., Viger, R. J., & Blodgett, D. L. (2014). A large-sample watershed-scale hydrometeorological dataset for the contiguous USA. Boulder, CO: UCAR/NCAR. https://doi.org/doi:10.5065/D6MW2F4D

Nijssen, B., O'Donnell, G. M., Hamlet, A. F., & Lettenmaier, D. P. (2001). Hydrologic Sensitivity of Global Rivers to Climate Change. *Climatic Change*, *50*(1), 143–175. https://doi.org/10.1023/A:1010616428763

O'Neill, B. C., Tebaldi, C., Van Vuuren, D. P., Eyring, V., Friedlingstein, P., Hurtt, G., … Sanderson, B. M. (2016). The Scenario Model Intercomparison Project (ScenarioMIP) for CMIP6. *Geoscientific Model Development*, *9*(9), 3461–3482. https://doi.org/10.5194/gmd-9-3461-2016

Oubeidillah, A. A., Kao, S. C., Ashfaq, M., Naz, B. S., & Tootle, G. (2014). A large-scale, high-resolution hydrological model parameter data set for climate change impact assessment for the conterminous US. *Hydrology and Earth System Sciences*, *18*(1), 67–84. https://doi.org/10.5194/hess-18-67-2014

Peel, M. C., Finlayson, B. L., & McMahon, T. A. (2007). Updated world map of the Köppen-Geiger climate classification. *Hydrology and Earth System Sciences*, *11*(5), 1633–1644. https://doi.org/10.5194/hess-11-1633-2007

Prudhomme, C., & Davies, H. (2009). Assessing uncertainties in climate change impact analyses on the river flow regimes in the UK. Part 2: Future climate. *Climatic Change*, *93*(1–2), 197–222. https://doi.org/10.1007/s10584-008-9461-6

River Management Joint Operating Committee (RMJOC). (2017). NRNI Flows 1929-2008 Corrected 04-2017. Bonneville Power Administration. Retrieved from

https://www.bpa.gov/p/Power-Products/Historical-Streamflow-Data/Pages/No-Regulation-No-Irrigation-Data.aspx

River Management Joint Operating Committee (RMJOC). (2018). *Climate and Hydrology Datasets for RMJOC Long-Term Planning Studies, Second Edition: Part 1 - Hydroclimate Projections and Analyses.* Retrieved from https://www.bpa.gov/p/Generation/Hydro/Pages/Climate-Change-FCRPS-Hydro.aspx

Rupp, D. E., Abatzoglou, J. T., & Mote, P. W. (2017). Projections of 21st century climate of the Columbia River Basin. *Climate Dynamics*, *49*(5–6), 1783–1799. https://doi.org/10.1007/s00382-016-3418-7

Rupp, D. E., Abatzoglou, J. T., Hegewisch, K. C., & Mote, P. W. (2013). Evaluation of CMIP5 20 th century climate simulations for the Pacific Northwest USA. *Journal of Geophysical Research: Atmospheres*, *118*(19), 10,884-10,906. https://doi.org/10.1002/jgrd.50843

Schewe, J., Heinke, J., Gerten, D., Haddeland, I., Arnell, N. W., Clark, D. B., … Kabat, P. (2014). Multimodel assessment of water scarcity under climate change. *Proceedings of the National Academy of Sciences of the United States of America*, *111*(9), 3245–3250. https://doi.org/10.1073/pnas.1222460110

Schnorbus, M., Werner, A., & Bennett, K. (2014). Impacts of climate change in three hydrologic regimes in British Columbia, Canada. *Hydrological Processes*, *28*(3), 1170–1189. https://doi.org/10.1002/hyp.9661

Smith, A. B., & Katz, R. W. (2013). US billion-dollar weather and climate disasters: Data sources, trends, accuracy and biases. *Natural Hazards*, *67*(2), 387–410. https://doi.org/10.1007/s11069-013-0566-5

"Soil Classification Working Group." (1998). *The Canadian System of Soil Classification*.

De Stefano, L., Petersen-Perlman, J. D., Sproles, E. A., Eynard, J., & Wolf, A. T. (2017). Assessment of transboundary river basins for potential hydro-political tensions. *Global Environmental Change*, *45*(July 2016), 35–46. https://doi.org/10.1016/j.gloenvcha.2017.04.008

De Stefano, Lucia, Edwards, P., De Silva, L., & Wolf, A. T. (2010). Tracking cooperation and conflict in international basins: Historic and recent trends. *Water Policy*, *12*(6), 871–884. https://doi.org/10.2166/wp.2010.137

Stocker, T. F., Qin, D., Plattner, G.-K., Tignor, M., Allen, S. K., Boschung, J., … Midgley, P. M. (2013). Climate change 2013: The physical science basis.

Stoelinga, M. T., Albright, M. D., & Mass, C. F. (2010). A new look at Snowpack Trends in the Cascade Mountains. *Journal of Climate*, *23*(10), 2473–2491. https://doi.org/10.1175/2009JCLI2911.1

Taye, M. T., Ntegeka, V., Ogiramoi, N. P., & Willems, P. (2011). Assessment of climate change impact on hydrological extremes in two source regions of the Nile River Basin. *Hydrology and Earth System Sciences*, *15*(1), 209–222. https://doi.org/10.5194/hess-15-209-2011

Taylor, K. E., Stouffer, R. J., & Meehl, G. A. (2012). An overview of CMIP5 and the experiment design. *Bulletin of the American Meteorological Society*, *93*(4), 485–498.

https://doi.org/10.1175/BAMS-D-11-00094.1

TFDD. (2016). Transboundary Freshwater Dispute Database. Retrieved from http://www.transboundar-ywaters.orst.edu/index.html

The World Bank. (2017). Major River Basins of the World. Retrieved from https://datacatalog.worldbank.org/dataset/major-river-basins-world

Thornton, P. E., Running, S. W., & White, M. A. (1997). Generating surfaces of daily meteorological variables over large regions of complex terrain. *Journal of Hydrology*, *190*(3–4), 214–251. https://doi.org/10.1016/S0022-1694(96)03128-9

Tohver, I. M., Hamlet, A. F., & Lee, S. Y. (2014). Impacts of 21st-Century Climate Change on Hydrologic Extremes in the Pacific Northwest Region of North America. *Journal of the American Water Resources Association*, *50*(6), 1461–1476. https://doi.org/10.1111/jawr.12199

Trenberth, K. E., Dai, A., Rasmussen, R. M., & Parsons, D. B. (2003). The changing character of precipitation. *Bulletin of the American Meteorological Society*, *84*(9), 1205-1217+1161. https://doi.org/10.1175/BAMS-84-9-1205

Trenberth, K. E., Dai, A., van der Schrier, G., Jones, P. D., Barichivich, J., Briffa, K. R., & Sheffield, J. (2014). Global warming and changes in drought. *Nature Climate Change*, *4*(1), 17–22. https://doi.org/10.1038/nclimate2067

Turnhout, E., Stuiver, M., Judith, J., Harms, B., & Leeuwis, C. (2013). New roles of science in society: Different repertoires of knowledge brokering. *Science and Public Policy*, *40*(3), 354–365. https://doi.org/10.1093/scipol/scs114

U.S Army Corps of Engineers. (2012). Columbia River Treaty History and Treaty Review.

Ukkola, A. M., De Kauwe, M. G., Roderick, M. L., Abramowitz, G., & Pitman, A. J. (2020). Robust future changes in meteorological drought in CMIP6 projections despite uncertainty in precipitation. *Geophysical Research Letters*, *n/a*(n/a), e2020GL087820. https://doi.org/10.1029/2020GL087820

UNEP-DHI, U. (2016). Transboundary river basins: Status and trends. *United Nations Environment Programme (UNEP): Nairobi, Kenya*.

United Nations. (2016). *Transforming our world: The 2030 Agenda for Sustainable Development*. https://doi.org/10.1201/b20466-7

USGS. (2009). Modeling of Watershed Systems GSFlow Training Class Material: Instructions for GSFLOW Model Input Preparation,.

Vano, J. A., Arnold, J. R., Nijssen, B., Clark, M. P., Wood, A. W., Gutmann, E. D., … Lehner, F. (2018). DOs and DON'Ts for using climate change information for water resource planning and management: guidelines for study design. *Climate Services*, *12*(July), 1–13. https://doi.org/10.1016/j.cliser.2018.07.002

Vano, J. A., Das, T., & Lettenmaier, D. P. (2012). Hydrologic Sensitivities of Colorado River Runoff to Changes in Precipitation and Temperature*. *Journal of Hydrometeorology*, *13*(3), 932–949. https://doi.org/10.1175/JHM-D-11-069.1

Vano, J. A., Kim, J. B., Rupp, D. E., & Mote, P. W. (2015). Selecting climate change scenarios

using impact-relevant sensitivities. *Geophysical Research Letters*, *42*(13), 5516–5525. https://doi.org/10.1002/2015GL063208

Velázquez, J. A., Schmid, J., Ricard, S., Muerth, M. J., Gauvin St-Denis, B., Minville, M., … Turcotte, R. (2013). An ensemble approach to assess hydrological models' contribution to uncertainties in the analysis of climate change impact on water resources. *Hydrology and Earth System Sciences*, *17*(2), 565–578. https://doi.org/10.5194/hess-17-565-2013

Vidal, J. P., Hingray, B., Magand, C., Sauquet, E., & Ducharne, A. (2016). Hierarchy of climate and hydrological uncertainties in transient low-flow projections. *Hydrology and Earth System Sciences*, *20*(9), 3651–3672. https://doi.org/10.5194/hess-20-3651-2016

Vormoor, K., Lawrence, D., Heistermann, M., & Bronstert, A. (2015). Climate change impacts on the seasonality and generation processes of floods – Projections and uncertainties for catchments with mixed snowmelt/rainfall regimes. *Hydrology and Earth System Sciences*, *19*(2), 913–931. https://doi.org/10.5194/hess-19-913-2015

Vörösmarty, C. J., Green, P., Salisbury, J., & Lammers, R. B. (2000). Global water resources: Vulnerability from climate change and population growth. *Science*, *289*(5477), 284–288. https://doi.org/10.1126/science.289.5477.284

Wasko, C., & Nathan, R. (2019). Influence of changes in rainfall and soil moisture on trends in flooding. *Journal of Hydrology*, *575*(May), 432–441. https://doi.org/10.1016/j.jhydrol.2019.05.054

Wasko, C., & Sharma, A. (2017). Global assessment of flood and storm extremes with increased temperatures. *Scientific Reports*, *7*(1), 1–8. https://doi.org/10.1038/s41598-017-08481-1

Werner, A. T., Schnorbus, M. A., Shrestha, R. R., & Eckstrand, H. D. (2013). Spatial and temporal change in the hydro-climatology of the Canadian portion of the Columbia river basin under multiple emissions scenarios. *Atmosphere - Ocean*, *51*(4), 357–379. https://doi.org/10.1080/07055900.2013.821400

Whitehead, P. G., Wilby, R. L., Battarbee, R. W., Kernan, M., & Wade, A. J. (2009). A review of the potential impacts of climate change on surface water quality. *Hydrological Sciences Journal*, *54*(1), 101–123. https://doi.org/10.1623/hysj.54.1.101

van der Wiel, K., Wanders, N., Selten, F. M., & Bierkens, M. F. P. (2019). Added Value of Large Ensemble Simulations for Assessing Extreme River Discharge in a 2 °C Warmer World. *Geophysical Research Letters*, *46*(4), 2093–2102. https://doi.org/10.1029/2019GL081967

Wilby, R. L., & Harris, I. (2006). A framework for assessing uncertainties in climate change impacts: Low-flow scenarios for the River Thames, UK. *Water Resources Research*, *42*(2), 1–10. https://doi.org/10.1029/2005WR004065

Wood, A. W., Leung, L. R., Sridhar, V., & Lettenmaier, D. P. (2004). Hydrologic implications of dynamical and statistical approaches to downscaling climate model outputs. *Climatic Change*, *62*(1–3), 189–216. https://doi.org/10.1023/B:CLIM.0000013685.99609.9e

Xin, X., Wu, T., Zhang, J., Yao, J., & Fang, Y. (2020). Comparison of CMIP6 and CMIP5 simulations of precipitation in China and the East Asian summer monsoon. *International Journal of Climatology*, *n/a*(n/a). https://doi.org/10.1002/joc.6590

Yan, H., Sun, N., Wigmosta, M., Skaggs, R., Leung, L. R., Coleman, A., & Hou, Z. (2019). Observed Spatiotemporal Changes in the Mechanisms of Extreme Water Available for Runoff in the Western United States. *Geophysical Research Letters*, *46*(2), 767–775. https://doi.org/10.1029/2018GL080260

Yuan, X., Jiao, Y., Yang, D., & Lei, H. (2018). Reconciling the Attribution of Changes in Streamflow Extremes From a Hydroclimate Perspective. *Water Resources Research*, *54*(6), 3886–3895. https://doi.org/10.1029/2018WR022714

Zelinka, M. D., Myers, T. A., McCoy, D. T., Po-Chedley, S., Caldwell, P. M., Ceppi, P., … Taylor, K. E. (2020). Causes of Higher Climate Sensitivity in CMIP6 Models. *Geophysical Research Letters*, *47*(1). https://doi.org/10.1029/2019GL085782

Appendix A. Supplemental Materials for Chapter 2: How do modeling decisions affect the spread among hydrologic climate change projections?

S1 Figures

We include below a select number of supplementary figures which are referenced in the paper.

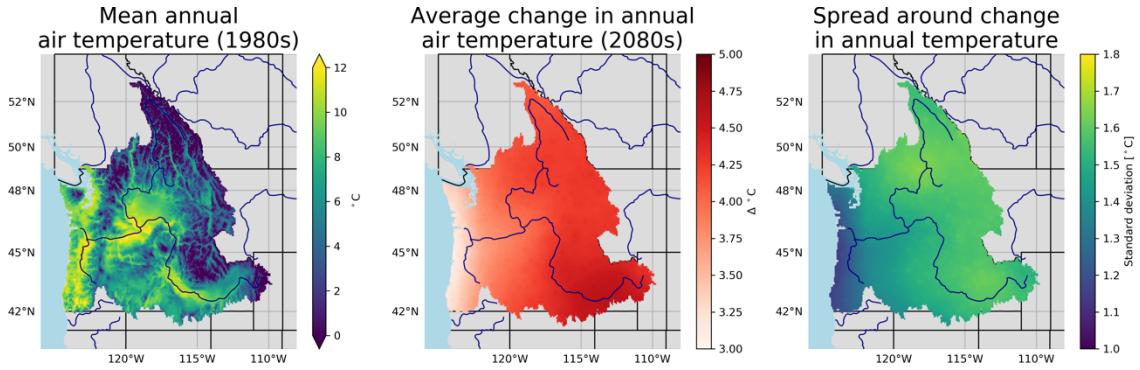


Figure S1. Mean annual air temperature for the 1980s (1970-1999) (left), changes by the 2080s (2070-2099) (center) and the standard deviation of the change (right) for the ensemble based on RCP 8.5.

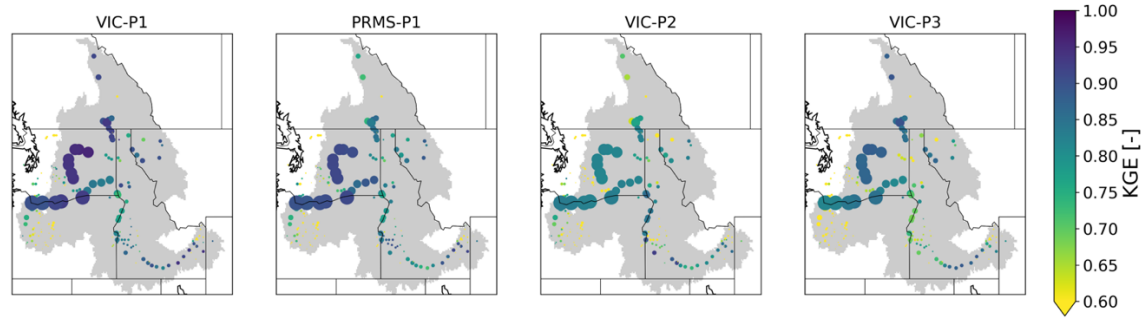


Figure S2. Historical performance (based upon weekly KGE compared to NRNI) of each of the four hydrologic models for the period WY 1960-2008, excluding the calibration period of WY 1992-2001. The grey area indicates the calibration domain. The sizes of the circles correspond to the mean annual streamflow during the 1960-2008 period. Circles only appear at locations for which NRNI exists for validation comparison.

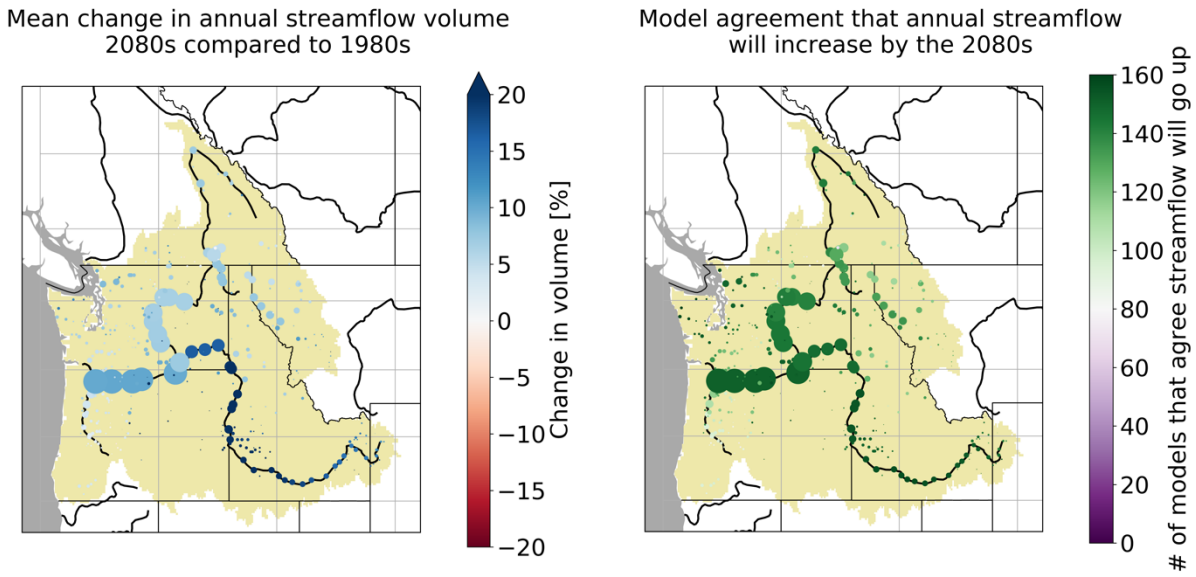


Figure S3. Projected changes in annual streamflow volumes (left) and agreement on the sign of change in streamflow at all streamflow locations in the domain. The sizes of the circles correspond to the mean annual streamflow during the 1960-2008 period.

S2 Detailed Methodology

Internal variability (IV) represents the natural fluctuations in the climate system independent of anthropogenic forces. Internal variability can cause any projected change in a variable (e.g. change in annual streamflow in the 2080s compared to the 1980s) to be either intensified or dampened. The model variability (MV), on the other hand, refers to the spread among simulations created by different modeling decisions. Together, these components comprise the total spread or TV. Thus,

$$TV = MV + IV \tag{S1}$$

Hawkins and Sutton (2009) and Alder and Hostetler (2018) achieved the separation of IV and MV by fitting a fourth order polynomial to the annual timeseries of temperature, precipitation or runoff. However, as discussed at the end of this section, our analyses reveals that the fourth order polynomial may not be appropriate for streamflow. Instead, we offer the

methodology described below, which uses a regression on the anthropogenic climate forcing itself.

Abatzoglou et al. (2014) argue that the responses of temperature and precipitation to anthropogenic forcing can be represented as a linear function of the anthropogenic forcing itself as long as the changes in the anthropogenic forcing are not very large. We propose that this assumption may be reasonable for streamflow responses as well. One variable used to quantify the anthropogenic forcing is the additional (compared to preindustrial levels) radiative forcing at the tropopause. Here, we develop a linear regression that represents the impact of the forcing (i.e. global mean radiative forcing) on a hydrometeorological variable of interest (i.e. streamflow). The predictions from this linear fit of streamflow to radiative forcing form a smoothed time series that represents changing streamflow through the 21st century without the internal variability component. We evaluate the spread among the different fitted model timeseries to calculate *MV* in equation (S1). The spread among the mean residuals from those fits represents the *IV* term in equation (S1). In doing so, we separate *IV* and *MV*. The *MV* term is then further decomposed via ANOVA into the separate components of variability owed to each model decision as follows:

$$MV = RCP + GCM + DSM + HM + GCM{:}RCP + GCM{:}HM + R \qquad \text{(S2)}$$

The *GCM:RCP* and *GCM:HM* each denote the variance explained by the choice of one of the modeling components dependent on the choice of the other. For example, the *GCM:RCP* term denotes the effect of the choice of a specific GCM when RCP is specified. *R* denotes the residual, which is the sum of the remaining interaction terms (e.g. *RCP:DSM*, *RCP:GCM:HM*, etc.). The steps are further detailed below.

We linearly regressed the metric of interest, $X_{m,t}$, against changes in radiative forcing. We performed all analyses below for every streamflow location independently. Each metric $X_{m,t}$, e.g. annual streamflow, was calculated for each water year *t* (October 1-September 30) for each individual projection *m* (n_m=160). Annual streamflow was calculated as the mean streamflow. The centroid of timing was calculated as the day of year when half of that water year's

streamflow had passed a gage. The low flow annual values for each water year *t* were calculated as the minimum 7-day flows for each year.

The metric of interest, $X_{m,t}$, was regressed upon the annual timeseries of global total anthropogenic radiative forcing from the relevant RCP (i.e. the streamflow simulations from RCP8.5 were regressed upon the radiative forcing timeseries from RCP8.5). Radiative forcing values were downloaded from http://www.pik-potsdam.de/~mmalte/rcps/index.htm (see also Meinshausen et al., 2011). An example for this step is shown in Figure S4.

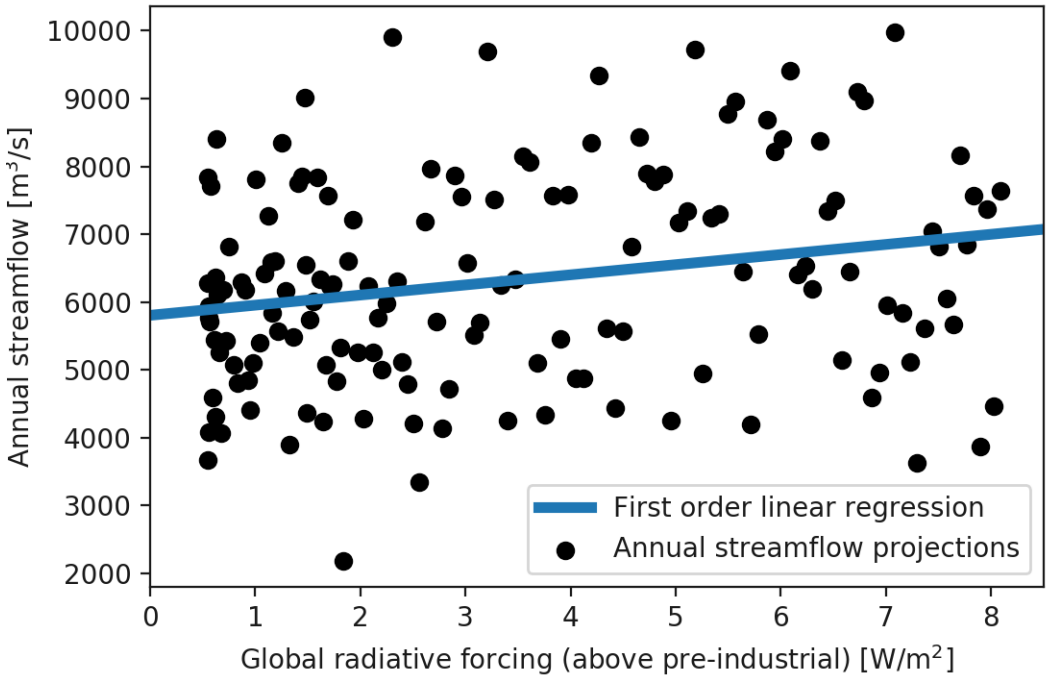


Figure S4. Example regression of annual streamflow against global radiative forcing. Every metric was regressed upon global mean additional radiative forcing at the tropopause as specified by the RCPs.

This regression was used to remove the temporal fluctuations due to natural variability. The projections thus take the form

$$X_{m,t} = x_{m,t} + \varepsilon_{m,t} \quad \text{(S3)}$$

where $x_{m,t}$ denotes the fit for each model m and timestep t, and ε denotes the residual from those fits for each model timeseries. An example of this fit is shown in Figure S5.

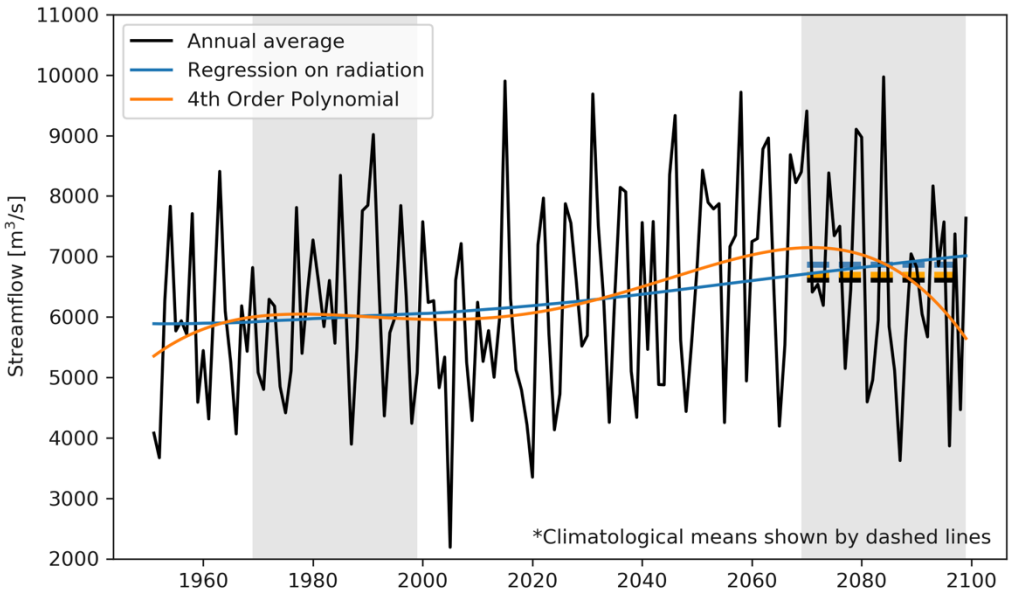


Figure S5. Example streamflow timeseries of the annual mean streamflow (black solid) with different smoothed fits. The regressions upon radiative forcing as well as (for comparison) a 4th-order polynomial are shown in blue and orange, respectively. The mean for the 2080s (2070-2099) is shown in the dashed lines for each trace. The future and reference periods are shaded in gray. The distances between the black and blue dashed lines are combined for all traces to calculate IV.

The internal variability is calculated by using the $\varepsilon_{m,t}$ residuals term from each of the 160 different model fits from Equation S3. Since our paper focuses on changes in 30-year means, we calculate the extent to which IV impacts changes in the 30-year means. We begin by calculating the mean for each period as:

$$\bar{x}_{period,m} = \frac{1}{n_t}\sum_{t_{period}} x_{t,m} \tag{S4}$$

where n_t for our case is 30, and *period* is either *reference* (1970-1999) or *future* (2070-2099). The quantity $x_{t,m}$ is expressed in units of m^3/s for annual streamflow and low flows and units of days for the centroid of timing.

The *MV* in Equation S1 is then calculated according to the variance in the change of $\bar{x}_{period,m}$ between the reference and future time periods. We want to stress that the study assesses the variances of changes in the means as opposed to the means themselves. For the relative changes (annual streamflow and low flow), we calculated *MV* as:

$$MV = var_m\left(\frac{\bar{x}_{future,m}-\bar{x}_{reference,m}}{\bar{x}_{reference,m}} * 100\%\right) \quad \text{(S5)}$$

where $var_m()$ is the variance across models m.

For absolute changes (centroid of timing), we calculated *MV* as:

$$MV = var_m\left(\bar{x}_{future,m} - \bar{x}_{reference,m}\right) \quad \text{(S6)}$$

The *IV* component of Equation S1 is calculated based upon the 30-year mean residuals from the fits denoted by:

$$\bar{\varepsilon}_{m,period} = \sum_{t_{period}} \frac{1}{n_t} \varepsilon_{m,t} \quad \text{(S7)}$$

where $\varepsilon_{t,m}$ is expressed in units of m^3/s for annual streamflow and low flows and units of days for the centroid of timing.

For relative changes (annual streamflow and low flow), we calculated IV as:

$$IV = var_m\left(\frac{\bar{\varepsilon}_{m,future}-\bar{\varepsilon}_{m,reference}}{\bar{x}_{m,reference}} * 100\%\right) \quad \text{(S8)}$$

and for absolute changes (centroid of timing), we calculated IV as

$$IV = var_m(\bar{\varepsilon}_{m,future} - \bar{\varepsilon}_{reference}) \quad \text{(S9)}$$

The *IV* estimates derived from these expressions are used to estimate the contribution of IV to the spread in projected changes in 30-year means.

Because the internal variability is calculated upon the mean residuals of the fits and the ANOVA is conducted upon the means of the fits, we are able to then sum the two sources of variance into a total spread *TV* as in equation (S1). We can then compare the relative contributions of IV and MV to the total spread.

Finally, we highlight and motivate two deviations in our analyses from the Hawkins and Sutton (2009) method. First, we chose to use the regression method shown in Figure S4 as opposed to the 4th-order polynomial suggested by Hawkins and Sutton (2009). In Figure S5 we compare the radiation regression (blue) and the fitted 4th-order polynomial (orange). The 4th-order polynomial underestimates the IV for annual streamflow changes, as seen by the smaller difference between the dashed orange and black lines as compared to the dashed blue and black lines. We decided to use the fits based upon the radiation regression as opposed to the 4th-order polynomial to best capture the potential contribution of IV to TV.

Second, we calculated IV differently from the methods described in Hawkins and Sutton (2009). They focused their IV analyses on residuals from fits of temperature to a 4th-order polynomial. They calculated the IV as the weighted multimodel mean variance across the entire annual timeseries of residuals from the fit. Thus, they assumed the variance to be constant through time. In this study, we were specifically interested in the influence of model decisions on projected changes in a future 30-year window from a reference 30-year period. Thus, our analyses of MV were specific for the 1980s to 2080s comparison. To preserve consistency between the IV and MV analyses, we decided to base the IV analyses upon the same time periods.

Appendix B. Supplemental Materials for Chapter 3: Climate change alters flood magnitudes and mechanisms in climatically-diverse headwaters across the northwestern United States

S1 Figures

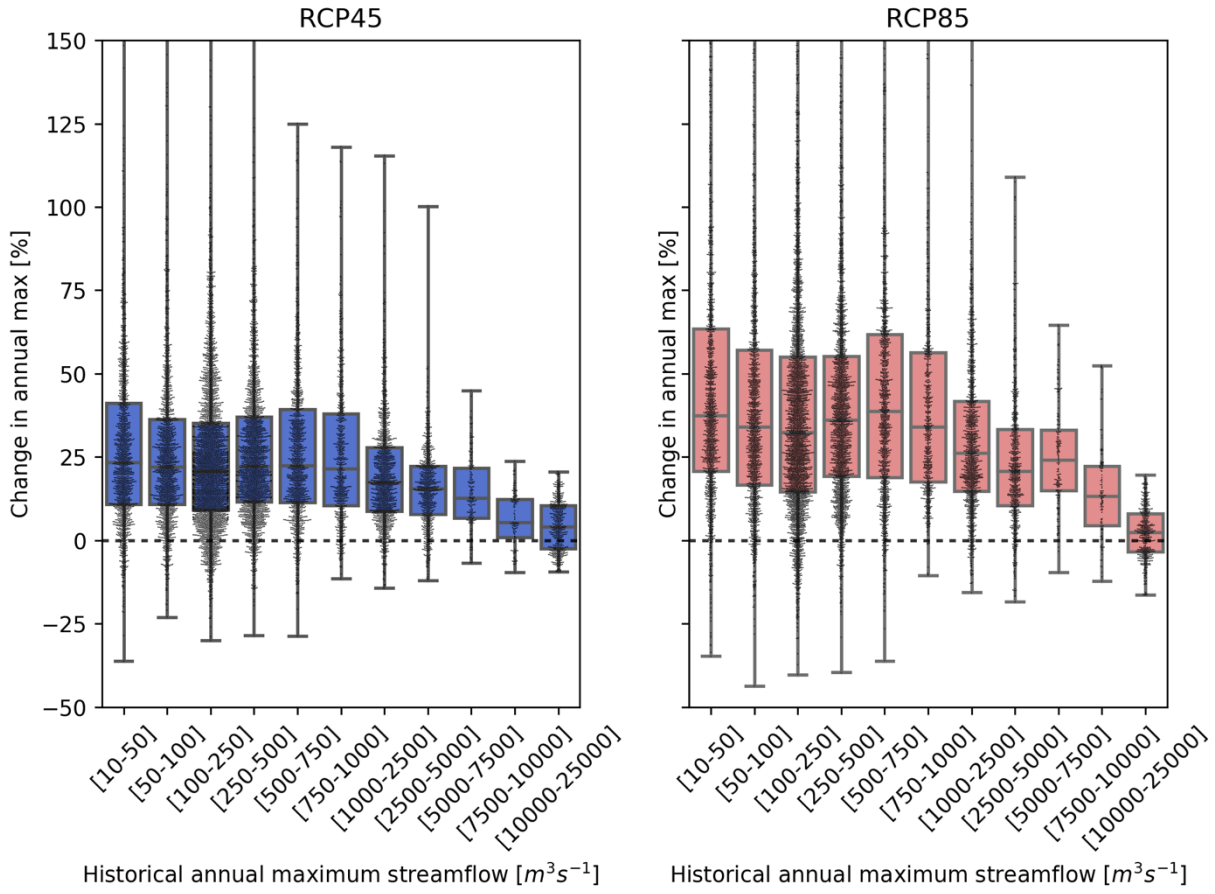


Figure S1. Changes in annual maximum streamflow (AMF) between the 1980s and the 2080s for RCP 4.5 and RCP 8.5 for every basin and simulation in the ensemble. The horizontal axis represents the magnitude of the AMF in the control simulation. Boxes represent the interquartile ranges and the whiskers represent the minima and maxima. Some extreme changes extend beyond plot boundary with the largest change at 280% for RCP 8.5, for basins with 50-100 m^3 s^{-1} historical annual maximum streamflow.

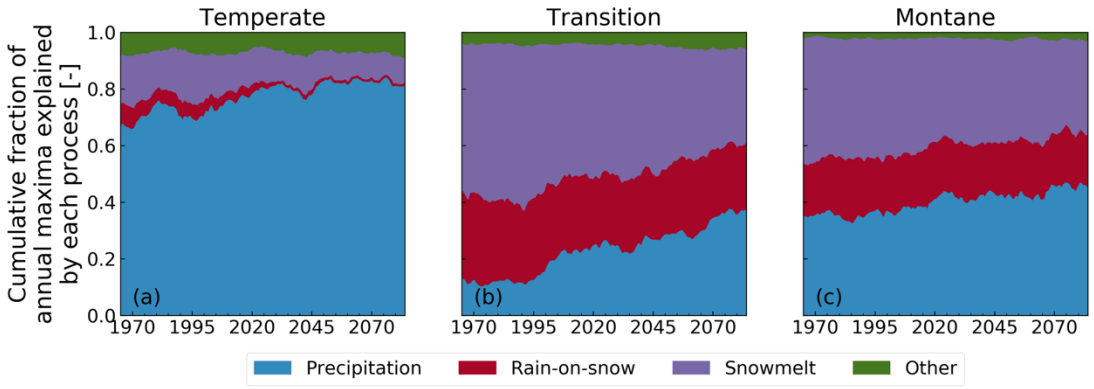


Figure S2. Changing proportion of annual maxima streamflow attributable to each flood generating process under RCP 4.5. Proportions shown represent fractions of all events from 30-year windows (centered on the year shown in the plot) across simulations of seven headwaters basins in each study region (temperate, transition, and montane) and using 10 GCMs and 4 hydrologic models per headwater basin

S2.1 Streamflow dataset

We used the Chegwidden et al (2017) dataset consisting of both routed streamflow (at ~400 point locations along the channel network) and gridded hydrologic data (e.g. spatial fields of snow water equivalent and other hydrologic variables). The methodology, model performance, and spatiotemporal domain of this dataset are described in detail in Chegwidden et al (2019) and River Management Joint Operating Committee (2018). The outlet names and locations for the basins used in this study are displayed in Table S1.

We used two representative concentration pathways (RCPs), ten global climate models (GCMs), and four hydrologic models to account the influence of modeling decisions on projected climate change impacts. As noted in Queen et al (in revision), the choice of hydrologic model can influence the projected sensitivity of extreme events, which motivated our decision to use multiple hydrologic models. The VIC simulations include sub-grid elevation information, with each 1/16 ° gridcell including 5 elevation zones to represent the distribution of temperature

with elevation within each gridcell. The PRMS simulations use a single elevation for each gridcell. VIC also distributes snow into the separate elevation zones whereas PRMS assumes snow is uniform across the gridcell.

The GCMs were selected based upon their historical performance over the Pacific Northwest as described in Rupp *et al* (2013) and River Management Joint Operating Committee (2018). While the Chegwidden et al (2017) dataset included results from two meteorological downscaling methods, we only used results based on the Multivariate Adaptive Constructed Analogs (MACA) downscaling technique. We excluded the simulations based upon the bias-corrected spatial disaggregation (BCSD) technique due to the documented superiority of MACA in simulating extreme precipitation (Abatzoglou & Brown, 2012).

S2.2 High flow selection

Streamflows were analyzed based on a water year (WY) basis (a WY is the period from October 1 in the previous calendar year through September 30 of the selected calendar year). For every year between WY 1951 and 2099, the maximum daily streamflow was identified for each individual basin and model simulation.

We considered several different streamflow metrics and selected the annual maximum daily flow (AMF) for four reasons. First, selecting the AMF is a simple, easily-reproducible way to ensure that events are independent. Second, many of the key findings in our study are based on temporal analyses in which we compare the evolution of processes from 1951 through 2099 (e.g. Figure 2). This analysis requires a continuous record of streamflow statistics, which the AMF provides. Selecting the top-n flow events for a climatological period would not provide that continuous record. Third, by selecting one event from each year, our sample sizes are consistent across climatological periods, basins, and models. Fourth, because we are comparing

changes in high flows to changes in annual temperature and precipitation, selecting the AMF from each year allows us to make a more thorough analysis of the behavior of high flows across the full climatological period.

S2.3. Classification decision tree

After calculating the AMF timeseries for each hydrologic scenario (each RCP+GCM+hydrological model combination), the "flood generating process" (FGP) was identified for each AMF. The classification tree is sequential, so each AMF is tested for each of the classifications in order (e.g. if it fulfills the conditions of the first step it will not be tested for the following steps). Each AMF receives a single classification, identifying the dominant process driving that single event. We recognize that the upstream hydrology is complex and varied and any given high flow event can result from a combination of hydrologic processes. However, the intention of this classification tree is to identify the single dominant process.

Thus, we classified each AMF using the following procedure. Spatial fields for each hydrometeorologic variable (e.g. precipitation, temperature, snow water equivalent) for the upstream areas of each of the basins were areally averaged to produce a basin-mean daily timeseries. The basin-mean quantities were paired with the streamflow at the corresponding basin outlet and fed into the following decision tree:

1) **Extreme precipitation driven:** If a single-day precipitation event within the top 99th percentile for that year occurred within four days before that WY's AMF. The four-day time window allows for a travel time from the event to the outlet. The 99th percentile criteria allowed us to include a large precipitation event that was not the absolute maximum precipitation event, but still was the dominant driver of the flow event. Note

that this differs from the criteria of Berghuijs et al. (2016) which looked for either the annual maximum single-day precipitation event or the annual maximum seven-day precipitation series.

2) **Rain on snow event:** If snow was present in the basin (>10 mm SWE averaged over the basin) at the time of a rain event (>10 mm basin-averaged in one day; classified as rain if basin-averaged temperature was > 0 °C) in the four days previous to the AMF and snowmelt accounted for greater than 20% of the sum of precipitation and snowmelt in that period. This definition is based upon Musselman *et al* (2018).
3) **Snowmelt:** If snow was present in the basin (>10 mm basin-averaged) and the snowpack reduced by at least 10% in the week previous to the AMF but the AMF did not qualify as an extreme precipitation or rain on snow event. For this classification we used a week as opposed to four days to allow more time for snowmelt to enter the soil column.
4) **Other:** All other remaining events after the previous three classifications proved false. These are events which cannot be neatly classified as driven by any of the three processes outlined above.

As a first order check of our classifications, we see that in the three domains analyzed in this study, the apportionment of the different FGPs aligns with the classifications shown in Berghuijs *et al* (2016) for the PNW for the historical period. The exception for their study domain is the presence of what Berghuijs *et al* (2016) call “soil moisture excess” as an FGP at about one quarter of the sites, particularly in the temperate region. While we do not use the “soil moisture excess” classification, its incidence in Berghuijs *et al* (2016) aligns approximately with the incidence of the “Other” category in our results.

S2.4. Sensitivity calculations

The AMF for each individual hydrologic time series was classified according to the methods detailed in Section S2.2. For each climate scenario (n=20 given 2 RCPs and 10 GCMs), the AMF were then classified according to their FGP and grouped together (e.g. all precipitation-generated floods were collected for a single timeseries). For each FGP, basin, and hydrologic model (*f*, *b,* and *h* respectively in in Equation (S1)), the group of classified AMF from control (WY 1970-1999) and future (WY 2070-2099) periods (represented as *t*) were averaged as follows:

$$AMF_{b,t,f,h} = \frac{1}{n_s * n_y} \sum_s \sum_y AMF_{b,t,f,h,s,y} \quad \text{(S1)}$$

where *b* is the number of basins in each region, *y* is the number of years, and *s* is the number of simulations (in our case 20 representing permutations of two RCPs and ten GCMs). A different $AMF_{b,t,f,h}$ is calculated for each region *r* (e.g. temperate). The AMF timeseries from every individual RCP-GCM-hydrologic model ensemble member is treated separately because the same year's simulation from a given model setup might result in the AMF being classified differently.

We then paired the change (future – control) in each $AMF_{b,t,f,h}$ to the change in annual temperature and precipitation corresponding to that basin and hydroclimate setup. Changes in climate were calculated as changes in annual average precipitation and temperature between the control and future periods for the area upstream of each of the streamflow locations for the years classified by each FGP. Changes in temperature are expressed as absolute changes (ºC) while changes in precipitation and streamflow are expressed as relative changes (%). The change in streamflow for each ensemble member from the control to the future period was then related to

that ensemble member's change in temperature and precipitation to create surfaces (i.e. contours; see Figure 4) showing the sensitivity of each classification of streamflow events to changes in temperature and precipitation.

We then calculated the change in AMF for each region *r* and FGP according to:

$$change_{r,f} = \frac{1}{n_b} \sum_b \frac{AMF_{future,r,f,b} - AMF_{control,r,f,b}}{AMF_{control,r,f,b}} * 100\% \quad \text{(S2)}$$

Equation (S2) weights all basins equally regardless of basin size. The calculation is repeated for each combination of three regions and three FGPs to create the nine mean change values in the upper left of each of the nine panels in Figure 4.

Table S1: List of streamflow locations used in this study

Location Name	**Region**	**Latitude (°N)**	**Longitude (°E)**
Green Peter	Temperate	44.46875	-122.53125
Clackamas River above Three Lynx Creek	Temperate	45.09375	-122.09375
Little Klickitat River near Wahkiacus	Temperate	45.84375	-121.03125
Fern Ridge	Temperate	44.09375	-123.28125
No. Santiam River below Boulder Creek near Detroit	Temperate	44.71875	-122.09375
Hood River	Temperate	45.65625	-121.59375
Kalama River near Kalama	Temperate	46.03125	-122.78125
Chelan	Transition	47.84375	-120.03125
Rimrock Reservoir on Tieton River	Transition	46.65625	-121.09375
Ashnola River near Keremeos	Transition	49.15625	-119.96875
Wenatchee River near Plain	Transition	47.78125	-120.65625
Pasayten River above Calcite Creek	Transition	49.09375	-120.53125
Twisp River near Twisp	Transition	48.40625	-120.21875
Cle Elum River at Cle Elum Reservoir	Transition	47.28125	-121.09375

Fording River at the mouth	Montane	49.90625	-114.84375
South Fork of Flathead River at Hungry Horse Dam	Montane	48.28125	-113.96875
West Fork Bitterroot River near Conner	Montane	45.71875	-114.28125
Clark Fork at Deer Lodge	Montane	46.40625	-112.71875
Jackson Lake Inflow	Montane	43.84375	-110.59375
Hoback River near Jackson, WY	Montane	43.28125	-110.65625
Salt River above reservoir near Etna, WY	Montane	43.09375	-111.03125

Table S2: List of Global Climate Models (GCMs) used in this study.

Global Climate Models	**Modeling Center (Country)**
CanESM2	Canadian Centre for Climate Modelling and Analysis (Canada)
CCSM4	National Center for Atmospheric Research (United States)
CNRM-CM5	Centre National de Recherches Météorologiques / Centre Européen de Recherche et Formation Avancée en Calcul Scientifique (France)
CSIRO-Mk3-6-0	Commonwealth Scientific and Industrial Research Organization in collaboration with Queensland Climate Change Centre of Excellence (Australia)
GFDL-ESM2M	NOAA Geophysical Fluid Dynamics Laboratory (United States)
HadGEM2-CC365	Met Office Hadley Centre (additional HadGEM2-ES realizations contributed by Instituto Nacional de Pesquisas Espaciais) (United Kingdom)
HadGEM2-ES365	
inmcm4	Institute for Numerical Mathematics (Russia)
IPSL-CM5A-MR	Institut Pierre-Simon Laplace (France)
MIROC5	Atmosphere and Ocean Research Institute (The University of Tokyo), National Institute for Environmental Studies, and Japan Agency for Marine-Earth Science and Technology (Japan)

Appendix C. SUPPLEMENTAL MATERIALS FOR CHAPTER 4: CLIMATE CHANGE HOT SPOTS OF WATER STRESS AND HYDROPOLITICAL CONFLICT

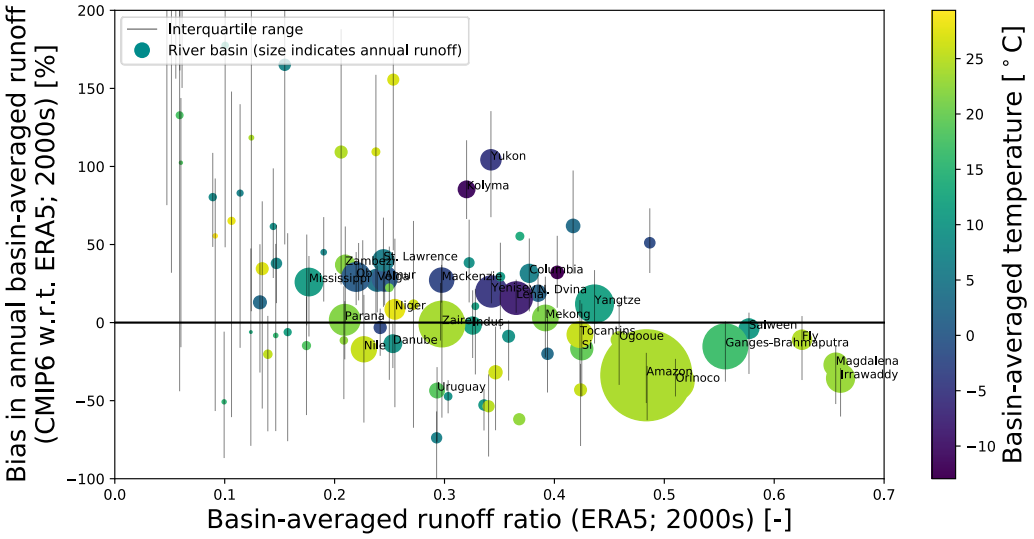


Figure S1 Biases in annual runoff from CMIP6 ensemble for global basins compared with runoff from ERA5-Land reanalysis product. Comparison for the period 1985-2015 for the 22 GCMs available.

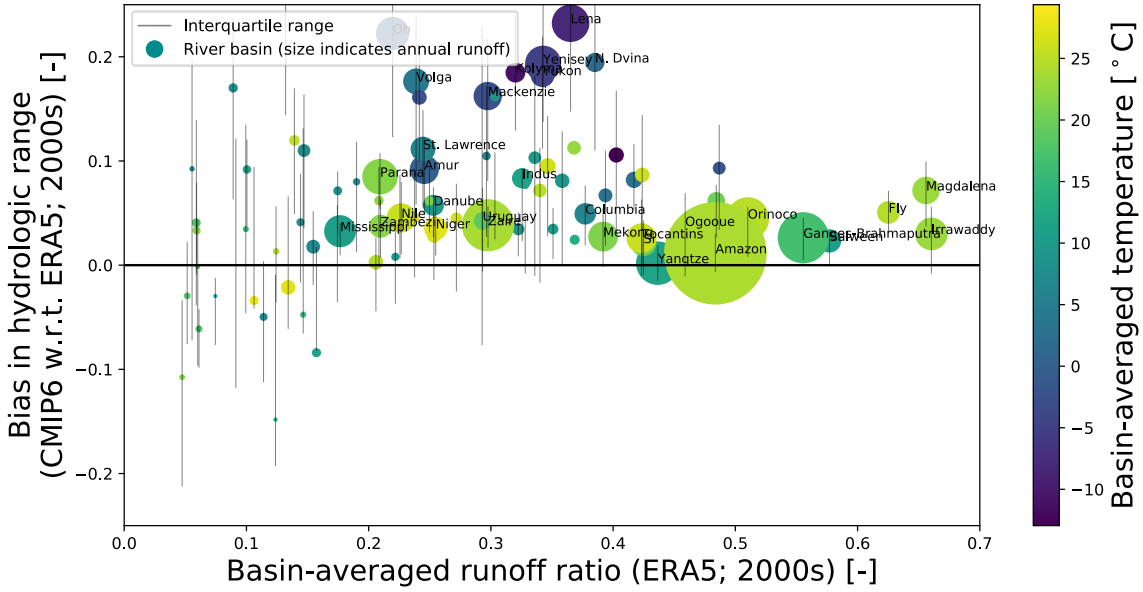


Figure S2 Biases in hydrologic range from CMIP6 ensemble for global basins compared with hydrologic range from ERA5-Land reanalysis product. Comparison for the period 1985-2015 for the 22 GCMs available.

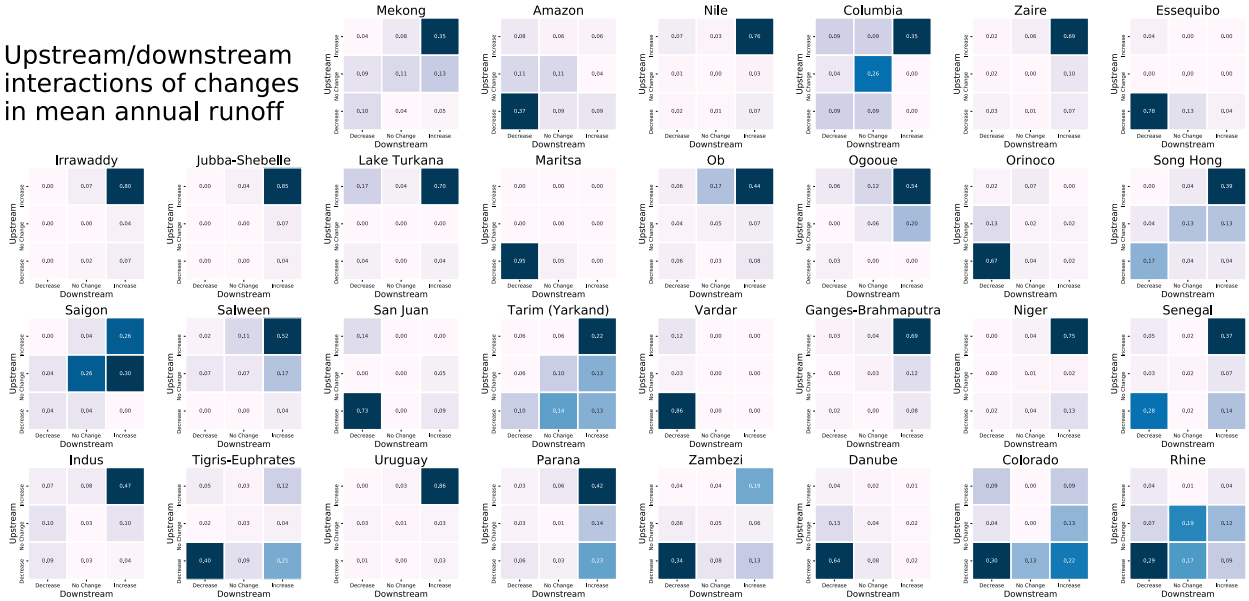


Figure S3. Matrices showing upstream/downstream pairs of changes in mean annual runoff between the control period (1986-2015) and the future period (2070-2099). Four of these matrices were shown in Figure 4 in the main manuscript.

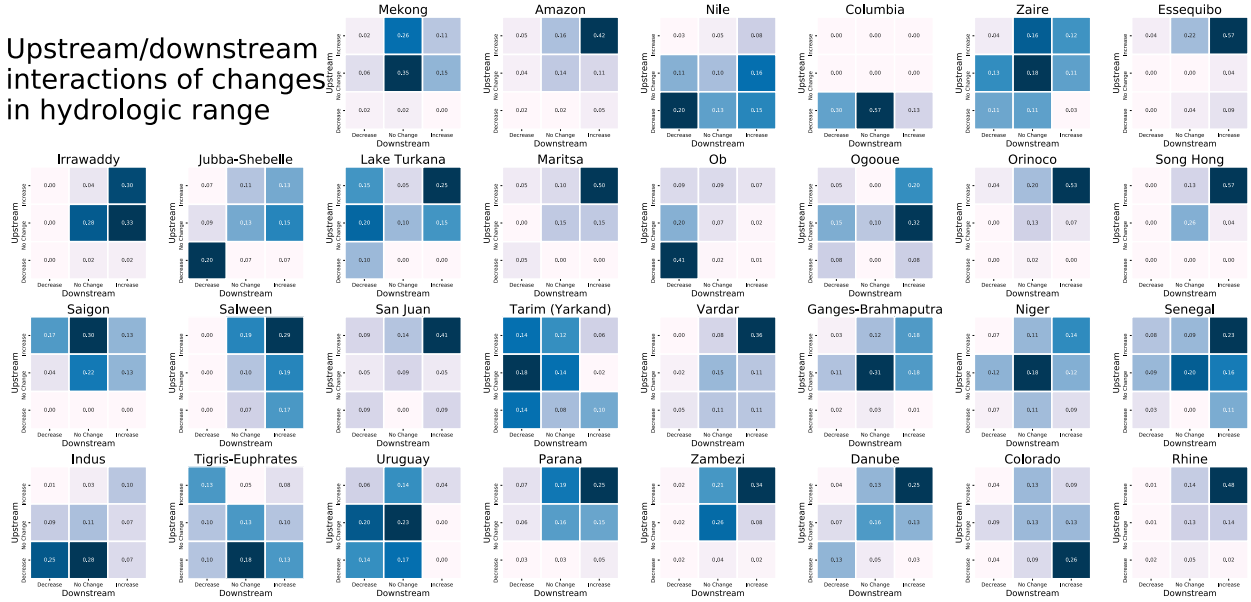


Figure S4. All matrices showing upstream/downstream pairs of changes in hydrologic range between the control period (1986-2015) and the future period (2070-2099). Four of these matrices were shown in Figure 4 in the main manuscript.

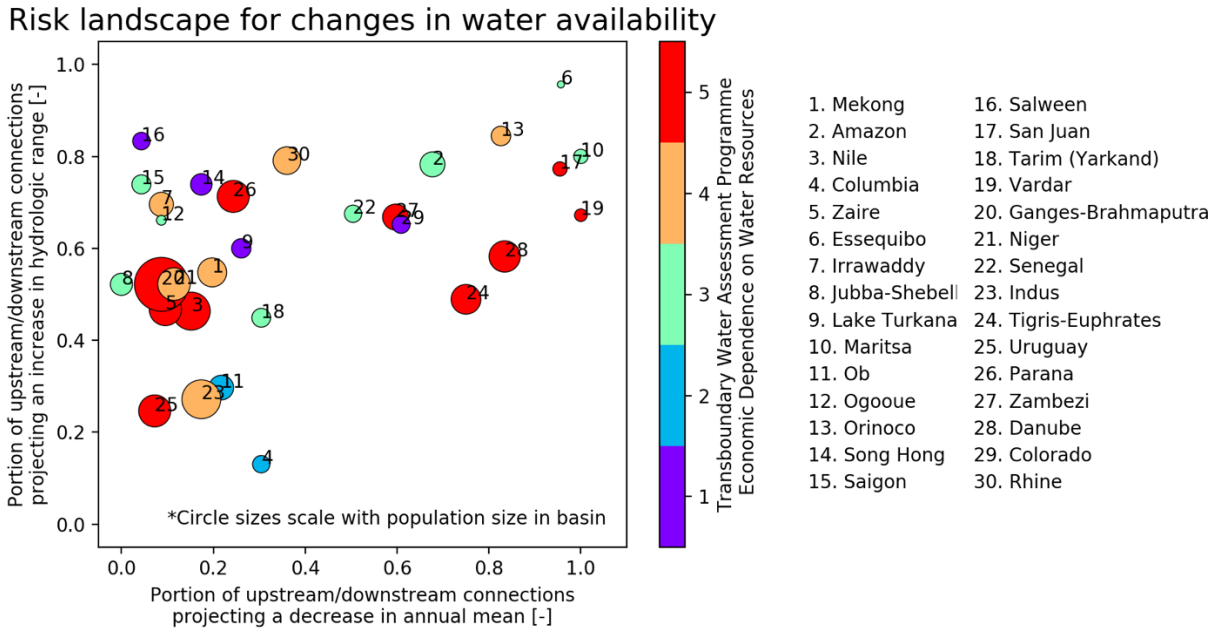


Figure S5. Same as Figure 4 in main manuscript except that color bar indicates, for each basin, the riparian nations' economic dependence on water resources. Red (dependence=5) indicates a greater dependence on water resources and purple (dependence=1) indicates a lower dependence.