Streamflow Projections to Inform Climate-Adapted Culvert Designs in Washington State

Prepared by the Climate Impacts Group | University of Washington

July 2023

Matthew Rogers | UW Climate Impacts Group

Anne Thebo | UW Climate Impacts Group

Guillaume Mauger | UWClimate Impacts Group

Cover Image Source:

Washington Department of Natural Resources (link to photo)

Acknowledgements

Phase 4 was funded by the Washington Department of Fish and Wildlife. We thank Oriana Chegwidden and Bart Nijssen for their help obtaining the dataset and taking the time to answer our questions about it. We also thank George Wilhere and Timothy Quinn for the considerable time and effort they took to provide feedback on this report.

Citation

Rogers, M., Thebo, A., Mauger, G. (2023). Streamflow Projections to Inform Climate-Adapted Culvert Designs in Washington State. Report prepared forWashington Department of Fish and Wildlife. Climate Impacts Group, University of Washington. https://doi.org/10.6069/YKZE-PG92.

Data Availability

All results from this work can be found at the following link: https://data.cig.uw.edu/picea/culverts/Phase_4/

Table of Contents

INTRODUCTION..... 4
Background/Previous Work..... 4
PROJECT APPROACH..... 5
Overview..... 5
Background on RMJOC-II data..... 5
Description of PRMS and VIC calibration..... 6
Project Tasks..... 7
Task 1: Calculate peak flows, percent change, for all projections..... 7
Task 2: Investigate the statistical properties of the peak flow results..... 9
Task 3: Recommend an approach for synthesizing projections for each grid cell..... 9
Task 4: Test approaches for regionalizing the individual grid cell projections..... 10
RESULTS AND DISCUSSION..... 10
Changes in Peak Flows and the Range Among Projections..... 10
Exploring Approaches for Synthesizing BFQ Projections..... 16
Evaluating the Central Tendency..... 16
Ensemble Spread Evaluation..... 20
Recommendations..... 22
RCP 4.5 vs RCP 8.5..... 22
GCMs..... 22
Downscaling Method - BCSD vs. MACA..... 23
Hydrologic Model - PRMS vs. VIC..... 23
Regionalizing the Projections..... 24
CONCLUSIONS AND NEXT STEPS..... 30
Central Tendency and Range..... 30
Regionalization..... 31
Future Work..... 31
References..... 33
Analysis of RCP 4.5 Projections..... 34
Additional RCP 8.5 Analyses..... 40

INTRODUCTION

Background/Previous Work

Restoring or maintaining habitat connectivity is essential to the viability of most fish and wildlife species. Habitat connectivity within lotic ecosystems of Washington State has been severely degraded by poorly designed or improperly installed culverts and other artificial hydraulic structures: at present, over 10,000 culverts block or impede the movement of fish. The loss of habitat connectivity due to culverts is one form of habitat degradation that has led to the listing of salmon ESUs as threatened or endangered.

The impacts of future climate change on habitat connectivity are a critical issue in contemporary fish and wildlife management. In lotic ecosystems, climate change is projected to increase annual peak discharge and alter channel morphology across much of Washington State (Wilhere et al. 2017). The design and construction of climate-adapted culverts, which accommodate projected future changes in annual peak discharge and channel morphology, will help to maintain habitat connectivity in lotic ecosystems in the face of climate change.

Fish and wildlife managers attempting to anticipate and address the impacts of future climate change often rely on quantitative models that project future climate-related impacts to habitats. However, the best way to develop, evaluate, and apply such models remains an open question. This project addresses one of the common challenges encountered when using projections of future climate-related impacts for the management of fish or wildlife habitats: how to grapple with many competing but plausible models and the concomitant uncertainty. More specifically, we address this challenge in the context of climate-adapted culvert design.

A dataset is available to assess changes in streamflow across the Pacific Northwest: The River Management Joint Operating Committee's version 2 hydrologic projections (RMJOC-II, https://www.hydro.washington.edu/CRCC, Chegwidden et al. 2019). This is a comprehensive dataset that provides estimates of future 'naturalized' streamflow (i.e., not accounting for dams or withdrawals) for the entire Columbia River basin and coastal drainage basins in Oregon and Washington. Culverts, along with stormwater infrastructure, are frequently designed based on specific peak flows (e.g., the 100-year event). The purpose of this study is to conduct supplementary analysis of the RMJOC-II data to estimate future changes in peak flows for use in these contexts, and to provide recommendations for how to interpret such projections.

The primary goals of this project include:

1. Update existing projections of future changes in peak flows and bankfull width for Washington State;
2. Explore and select statistics for expressing central tendency and uncertainty when faced with an extreme multiplicity of models; and
3. Recommend a process for applying this information in climate-adapted culvert design.

PROJECT APPROACH

Overview

Culverts and other instream structures (e.g., bridges, stream-bank armoring) are frequently designed based on specific peak flows (e.g., the 100-year event). The purpose of this study is to estimate future changes in peak flows for use in these contexts, and to provide recommendations for how to interpret such projections. It is not always practical or even warranted – given model uncertainties – to report projections of the effects of future climate with high spatial resolution. As such, this project also assesses the potential for regionalizing the results to more accurately convey the spatial precision of the projections, thereby simplifying the climate change information used in culvert design. The general workflow and tasks for this project are outlined in Figure 1 and detailed below.

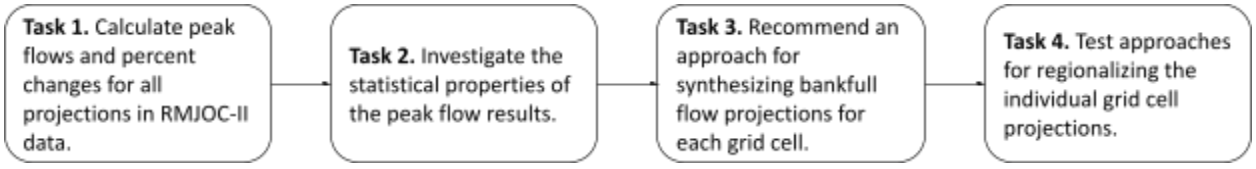

Figure 1. Project tasks and workflow.

Background on RMJOC-II data

The RMJOC-II dataset (sometimes alternatively referred to as the "Columbia River Climate Change", or CRCC, dataset) builds on previous hydrologic projections by updating to newer models and providing a more comprehensive representation of the uncertainty space (Figure 2). Specifically, global climate model (GCM) projections were obtained from the newer Climate Model Intercomparison Project, phase 5 (CMIP5; Taylor et al., 2012). Ten GCMs were selected based on Rupp et al. (2013), who evaluated and ranked global climate models based on their ability to reproduce the climate of the Pacific Northwest. For each GCM, two greenhouse gas scenarios were evaluated: RCP 4.5, a low-end scenario that has emissions peaking in mid-century and declining thereafter, and RCP 8.5, a high-end

scenario that has emissions increasing through the end of the 21st century (Van Vuuren et al. 2011).

The GCM projections were statistically downscaled using two approaches: (1) the Multivariate Adaptive Constructed Analog technique (MACA, Abatzoglou and Brown 2012), and (2) the Bias-Correction, Spatial-Disaggregation technique (BCSD, Wood et al. 2004). All projections provide daily maximum temperature, minimum temperature, precipitation, and wind speed, for the years 1950 through 2099.

The hydrologic modeling is further delineated by using two hydrologic models and three approaches to model calibration. The three calibration approaches were used with the Variable Infiltration Capacity (VIC) hydrologic model, one of which was also applied to the Precipitation Runoff Modeling System (PRMS) hydrologic model, making for a total of four hydrologic model-calibration pairs. The VIC model includes a simple glacier model.

Combining the two greenhouse gas scenarios, 10 GCM projections, two downscaling approaches, and four hydrologic model-calibration pairs, there are a total of 160 climate change projections (80 per greenhouse gas scenario) for each 1/16-degree grid cell across the entire Pacific Northwest model domain (Figure 2). The model domain includes 5168 grid cells within Washington State. Each projection includes daily estimates of surface and subsurface runoff for the years 1950-2099.

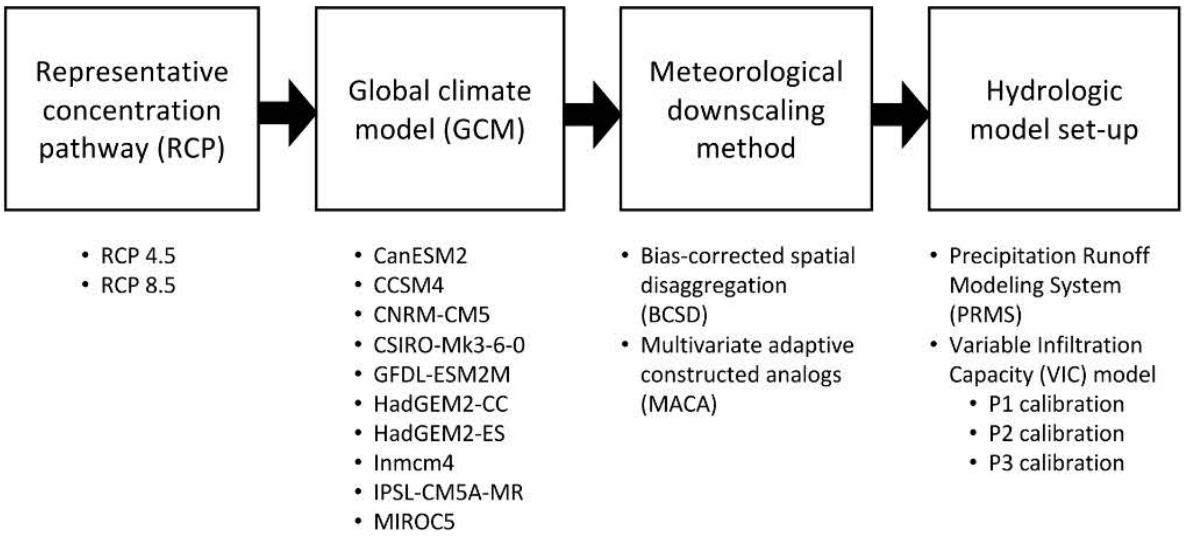

Figure 2. Components of RMJOC-II Model Ensemble. (Figure source: Chegwidden et al. 2019)

Description of PRMS and VIC calibration

The RMJOC-II data analyzed in this project combines calibration approaches from UW (PRMS-P1 and VIC-P1), NCAR (VIC-P2), and ORNL (VIC-P3). Each of these differed with regard

to the parameters calibrated, calibration methodology, and reference meteorological and streamflow datasets used (Table 1). Additional details are available in Chegwidden et al. (2019).

Table 1. Description of the four hydrologic models included in the RMJOC-II data.
Table source: Chegwidden et al. (2019).

	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)

Project Tasks

Task 1: Calculate peak flows, percent change, for all projections

Calculation of Peak Flows

VIC and PRMS produce gridded fields of surface and subsurface runoff, in this case with a spatial resolution of 1/16-degree. Flow extremes were estimated following the approach described in Tohver and Hamlet (2014): using the sum of surface and subsurface runoff, then taking the maximum flow in each water year and using L-moments to fit a generalized extreme value distribution to the maximum flows. Extreme statistics were calculated for bankfull flows (BFQ) based return intervals in Castro and Jackson (2001, Table 2) and also for the 100-year event. Peak flows were evaluated for all 160 projections, for the following two time periods: 1990-2019 ("2000s") and 2070-2099 ("2080s"). The current analysis focuses on the percent changes for the 2080s relative to historical. All results are available online via the CIG website (project page link). The analysis included both the RCP 4.5 and RCP 8.5 scenarios, though only the results from the RCP 8.5 scenario are included in this report. We chose to focus on RCP 8.5 because the results are more demonstrative of the differences among hydrologic models-calibration pairs and downscaling methods due to the larger increase in future temperatures. The results for the RCP 4.5 scenario simulations can be found in the appendices.

Estimating Bankfull Width

Bankfull widths (BFW) were calculated using the approach of Wilhere et al. (2016) including the empirical relationships developed by Castro and Jackson (2001) (Equation 1).

$$BFW = \alpha Q^{\beta} \qquad \text{Eq. 1}$$

Where Q was estimated as bankfull flow (BFQ) and α and β are fitted parameters relating streamflow to bankfull width (BFW) (Castro and Jackson (2001). $\boldsymbol{\tau}$ indicates the bankfull discharge recurrence interval from Castro and Jackson (2001). Fit parameters were determined for three ecoregions in Washington State: Pacific Maritime Mountains, Western Cordillera, and Columbia Plateau (Figure 3, Table 2).

Table 2. Parameters used to relate bankfull width to peak streamflow statistics. Source: Castro and Jackson (2001).

Ecoregion	**α**	**β**	**τ**
Pacific Maritime Mountains	2.37	0.50	1.2-yr
Western Cordillera	3.50	0.44	1.5-yr
Columbia Plateau	0.96	0.60	1.4-yr

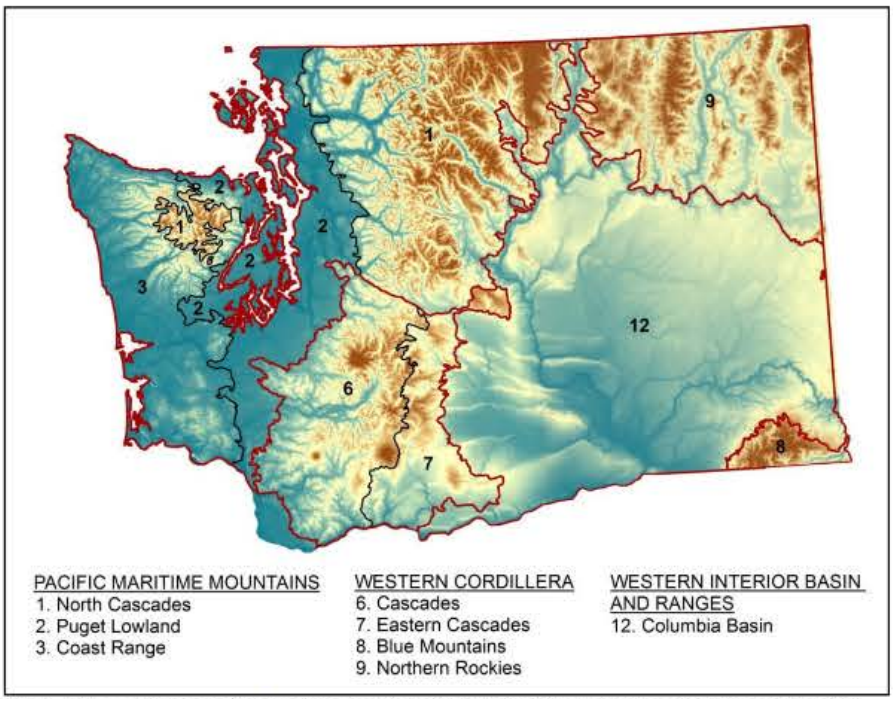

Figure 3. Ecoregions used to define the parameters relating bankfull width (BFW) to peak streamflow statistics across Washington State. Source: Wilhere et al. 2016.

Task 2: Investigate the statistical properties of the peak flow results

In order to know how to best summarize the projections of peak-flow statistics, we need to know how the projections are distributed statistically. For example, results from one hydrologic model–calibration pair may not overlap with those from another, meaning that averaging over the two may give a misleading result. To address this need, UW evaluated the 160 projections per grid obtained in Task 1 to better understand how 100-yr flows, BFQ, and BFW vary within grid cells across Washington State. This included evaluating the extreme-flow statistics for each time period and the projected percent changes in those statistics. Statistics were calculated separately for each hydrologic model-calibration pair (PRMS-P1, VIC-P1, VIC-P2, VIC-P3) and downscaling method (MACA, BCSD). Statistics computed for peak flow results include:

- 10-model ensemble median percent change;
- Difference between ensemble maximum and minimum percent change (range);
- Difference between ensemble 90th percentile and 10th percentile percent change; and
- Count of models with a positive sign of change.

We selected the difference between the 90th percentile and 10th percentile percent change to determine the sensitivity of the range results by excluding the most extreme model results without significantly reducing the sample size in our measure of uncertainty. Initial analyses were performed on a few select grid cells, then computed across the entire state with results summarized as a map series showing variations across the state.

Task 3: Recommend an approach for synthesizing projections for each grid cell

In Task 2, we computed four statistics evaluating the range and central tendency across the subsets of the model ensemble, as described above. Task 3 interprets the findings from Task 2 as they relate to four key decision points in the models and implications of those decisions in assessing future hydrologic conditions:

- Emissions scenario (RCP 4.5, RCP 8.5);
- Global climate model;
- Meteorological downscaling method; and
- Hydrologic model (PRMS-P1, VIC-P1, VIC-P2, VIC-P3).

The four statistics we computed for each grid cell include:

- Ensemble average;
- Ensemble median;
- Ensemble maximum - ensemble minimum (range); and
- Ensemble 90th percentile - ensemble 10th percentile.

The first two of these statistics evaluate the central tendency of the ensemble and the final two statistics evaluate the model range at each grid cell.

Task 4: Test approaches for regionalizing the individual grid cell projections

Regionalization is a geostatistical approach used to aid interpretation of trends in data across space, lessen the impacts of outliers, and simplify noisy or complex data into operational units such as watersheds. Results of adjacent grid cells are combined where similar changes are expected. Any regionalization approach will by definition involve reducing the information content of the projections. The extent of smoothing or clustering depends on both the range among the projections and the level of precision that is needed for culvert design. Professional judgment is needed to inform both -- model uncertainties are not quantifiable on a grid cell basis because this implies a false level of precision, and culvert design could require more precision than regional uncertainty under some circumstances.

In Task 4, four classes of regionalization approaches were evaluated:

- Ecoregions (Omernik level III and IV)
- Watersheds (HUC 8 and HUC 12)

These regionalization approaches were evaluated for BFQ and BFW for each of the four hydrologic model-calibration pairs (PRMS-P1, VIC-P1, VIC-P2, VIC-P3) and both downscaling methods (BCSD, MACA). These four approaches were selected to include a range of geostatistical and geographically based approaches. Other methods were considered but ultimately not pursued due to their complexity and disconnect with geographical and planning-based needs (e.g., statistical smoothing, clustering) when applied over the entire domain. Benefits and tradeoffs of these approaches are discussed in subsequent sections of this report.

RESULTS AND DISCUSSION

Changes in Peak Flows and the Range Among Projections

We first select a series of points across Washington state at which to analyze the statistics at the grid cell level. To capture a variety of different ecological regions across the state, we selected a point in each of the 8 Omernik level 3 ecological regions that comprise Washington state (Figure 4). The specific latitude and longitudes of these points can be found in Table 3.

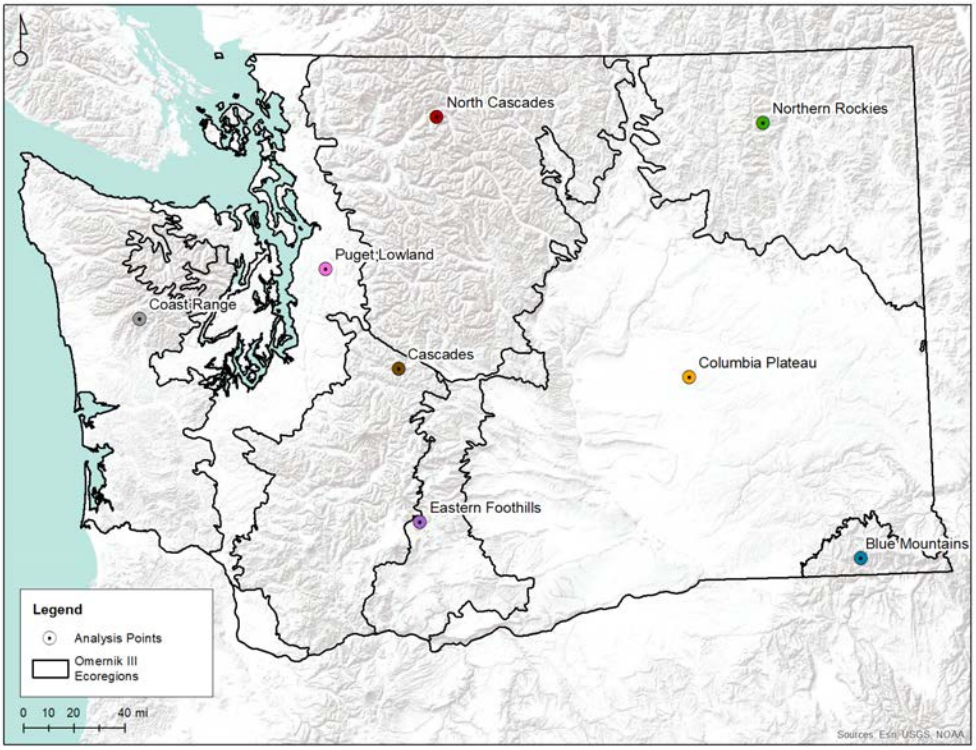

Figure 4. Grid cells used in initial extreme flow analysis. Points are color-coded by ecoregion.

Table 3. Latitude and longitude of selected grid cells across the 8 Omernik level 3 ecoregions used in the point analysis.

Region	Latitude	Longitude
Blue Mountains	46.09375	-117.65625
Columbia Plateau	47.15625	-119.03125
Northern Rockies	48.59375	-118.34375
North Cascades	48.65625	-121.15625
Eastern Foothills	46.34375	-121.28125
Cascades	47.21875	-121.46875
Puget Lowland	47.78125	-122.09375
Coast Range	47.46875	-123.65625

For these eight points, we first demonstrate the performance of the GEV fit when applying the L-moments method as described in Tohver and Hamlet (2013). Figure 5 compares the cumulative distribution function (CDF) from the model results and the cumulative distribution function using the generalized extreme value parameters at five of these points for the high greenhouse gas scenario in the 2080s. The model results are derived from water year maxima of the sum of daily surface and subsurface runoff. In general, the CDFs based on the GEV fit capture the shape of the empirical CDFs while lending confidence to the use of the L-moments method and supporting the results discussed hereafter. However, the magnitudes of runoff at these points vary considerably between the PRMS and VIC hydrologic model simulations, especially for the grid cells east of the Cascades such as the Eastern Foothills, Blue Mountains, and Columbia Plateau.

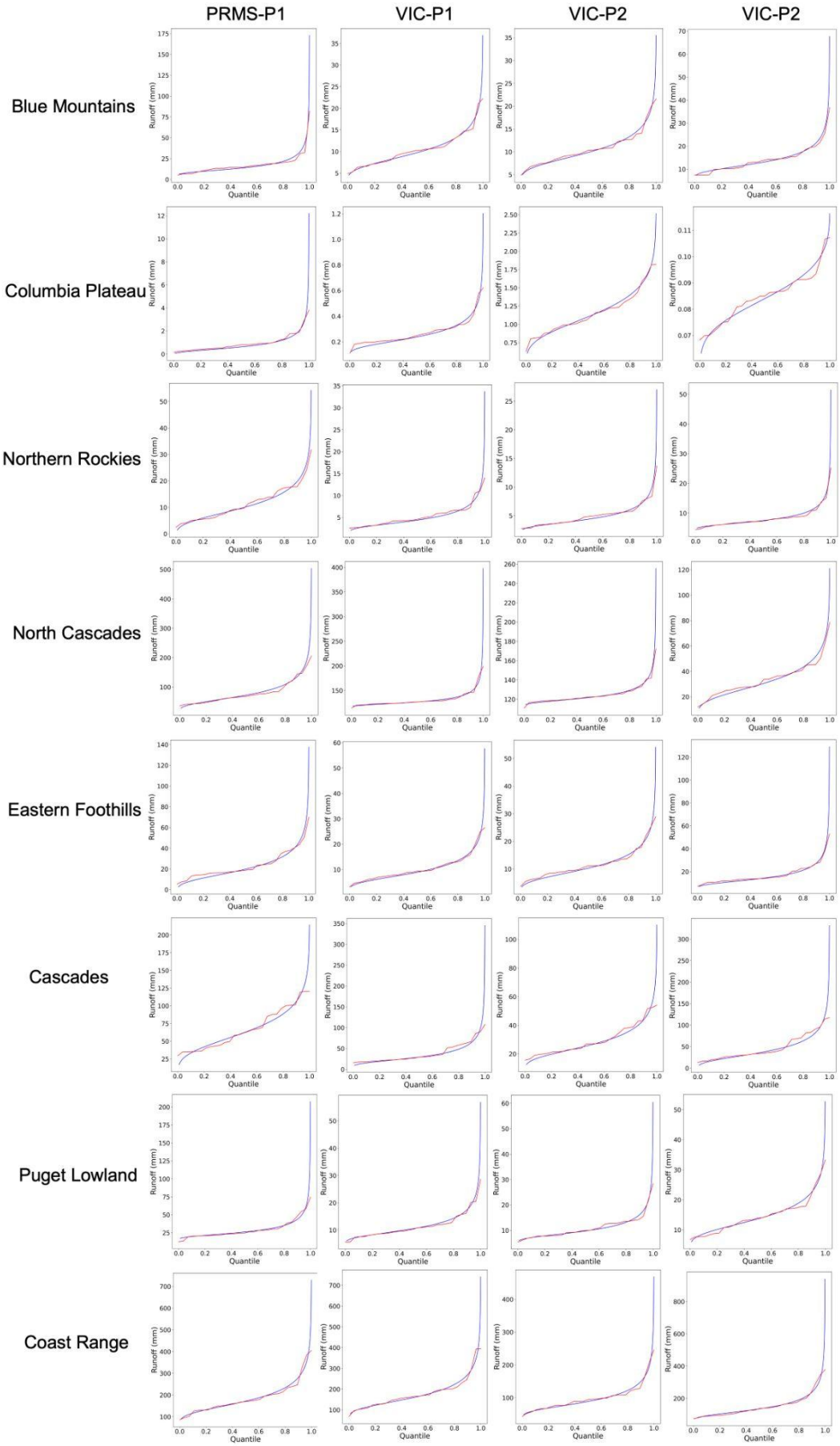
PRMS-P1
VIC-P1
VIC-P2
VIC-P2
Blue Mountains
Columbia Plateau
Northern Rockies
North Cascades
Eastern Foothills
Cascades
Puget Lowland
Coast Range
Runoff (mm)
Quantile

Figure 5. Comparison of the cumulative distribution functions (CDFs) of the grid cells listed in Table 3 in the 2080s (2070-2099) for the MACA downscaling scheme in the 2080s for the CCSM4 model. Blue lines represent the modeled CDFs using the L-moments method, and red lines represent the CDFs derived directly from the modeled data. Columns are the various hydrologic model - calibration pairs.

Next, we compare the model spread in percent change in BFQ for the 2080s for each of the hydrologic model-calibration pairs in Figures 6 (MACA simulations) and 7 (BCSD simulations) at each point in Table 3. The high degree of similarity between the results for each hydrologic model-calibration pair suggests that the selected downscaling method does not have a large influence on the results at the selected points. Rather, the largest differences in results for each point appears to be caused by the selection of the hydrologic model, and in some cases, the spread among GCMs. Chegwidden et al. (2020) did not investigate changes in peak flows, but this finding is similar to what they found for low flow projections. Among the hydrologic models-calibration pairs, the three VIC calibrations show relatively minor differences with each other (e.g., VIC-P3 Cascades point has a larger spread than the others), whereas the PRMS-P1 results for both the MACA and BCSD simulations diverge from the VIC results, particularly for points in mountainous locations (e.g., North Cascades, Eastern Foothills, and Northern Rockies). Further spatial analysis, below, shows that these differences between the PRMS-P1and VIC results are representative of what we find across the domain.

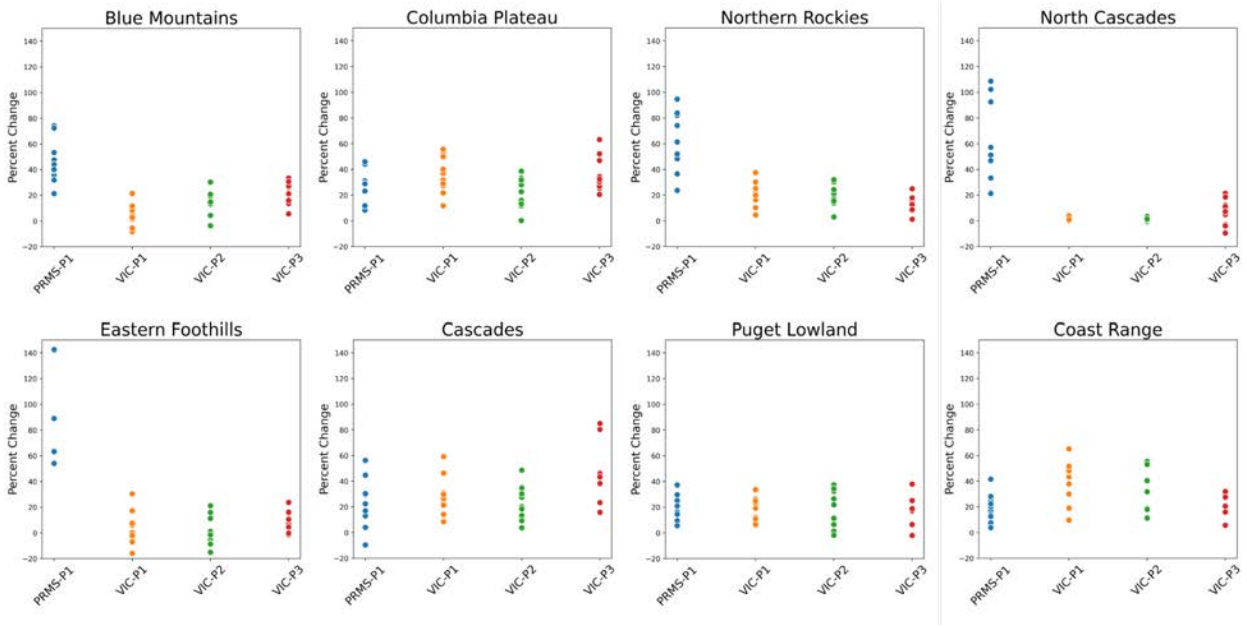

Figure 6. Percent change in BFQ for the MACA downscaled simulations under the high-end greenhouse gas scenario (RCP 8.5) for the 2080s (2070-2099) relative to the 2000s (1990-2019) at the 8 grid cells shown in Fig 4 and listed in Table 3. There are 10 points (i.e., models) per grid cell but some points are obscured by overlap.

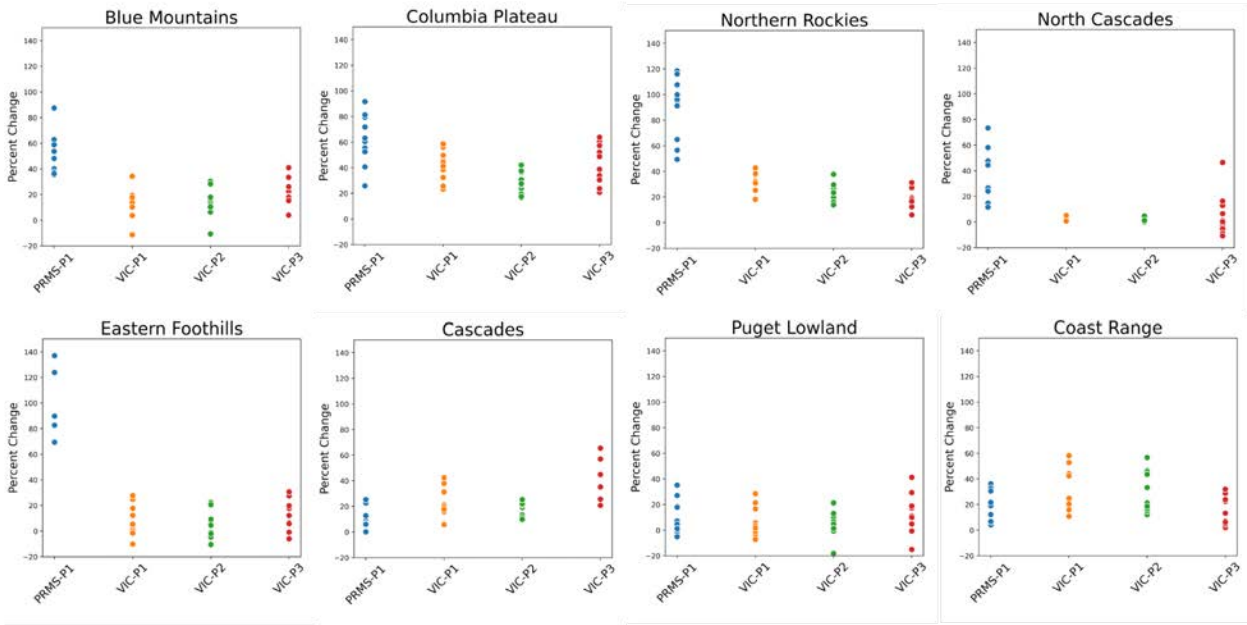

Figure 7. As in Figure 6 except showing results for the BCSD downscaling.

Exploring Approaches for Synthesizing BFQ Projections

Evaluating the Central Tendency

In order to evaluate potential approaches to synthesizing the 80 projections, we compared several measures of the range and central tendency in the hydrologic models-calibration pairs and two different downscaling methods that comprise this dataset (totaling 10 models per ensemble, or sub-ensembles hereafter). We did this to identify whether a specific hydrologic model-calibration pair, downscaling method, or global climate model introduces the most to the ensemble spread, and whether we can explore excluding any outliers from the larger ensemble and the estimates of the range, our proxy for uncertainty in the projections. Additionally, assessing the sub-ensemble spread for multiple measures of central tendency and range gives insight into which measure will be more practical and accurate for applications to culvert design. Figures 8 and 9 show the first measure of central tendency among the sub-ensembles, the average percent change, for the MACA and BCSD downscaling methods, respectively.

For the 100-year event, the differences in the ensemble average are stark across the hydrologic model-calibration pairs. We expect greater differences in the 100-year change estimates, because a 30-year sample requires extrapolation to estimate the magnitude of this extreme event. There are also important differences between the two downscaling methods, with generally larger changes projected with the MACA dataset. The PRMS-P1 results show larger differences, with much of western Washington showing opposite signs of change for the MACA and BCSD downscaled projections. This contrasts somewhat from the point analysis, where the differences among downscaling approaches were not as apparent.

The subplots for BFQ and BFW show much more consistency across the hydrologic model sub-ensembles than the 100-year event. As noted above, this is expected given that the flow extreme corresponding to the BFQ (which ranges from a 1.2 to 1.5-year event; see Table 2) is adequately sampled with a 30-year record. In this case, the different VIC calibrations are quite similar in their patterns, both across calibration approaches and downscaling methods. The PRMS-P1 results, compared to the three VIC calibrations, consistently project larger increases for both BFQ and BFW. This is consistent with the point analysis; the maps show that the differences between the PRMS-P1and VIC simulations are not limited to the points assessed in Task 2.

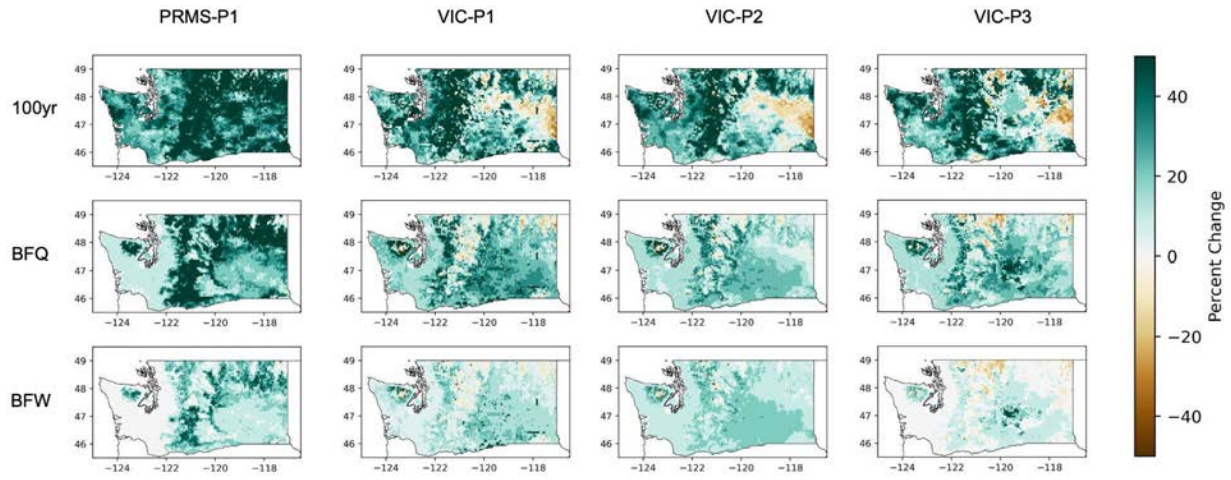

Figure 8. 10-model ensemble average percent change for the MACA downscaled simulations under the high-end greenhouse gas scenario (RCP 8.5) for the 2080s (2070-2099) relative to the 2000s (1990-2019). The columns are the different hydrologic model-calibration pairs. The rows are the results for the 100-year event (top), bankfull flow (BFQ, middle), and bankfull width (BFW, bottom).

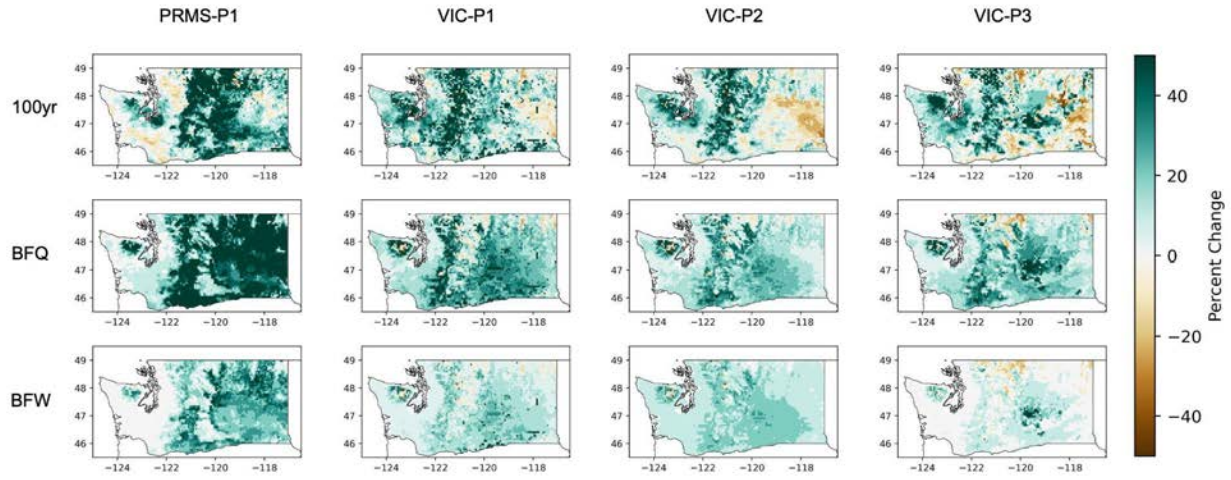

Figure 9. As in Figure 8 except showing results for the BCSD downscaled projections.

Figure 10 shows the ensemble median percent change for the MACA downscaled simulations and Figure 11 shows the ensemble median percent change for the BCSD downscaled simulations. In general, the results for the three VIC calibrations are similar to those for the ensemble mean, with projected increases in the 100-year runoff event magnitude for the Cascade mountain range and much of western Washington and sporadic decreases throughout eastern Washington. The spatial patterns for BFQ and BFW are similar in the different VIC calibrations as well, with all three showing broad increases through Washington state, excluding some areas in the eastern foothills of the Cascades

and northeast Washington. As with the ensemble average, the ensemble median percent changes for PRMS-P1are quite different from the three VIC calibrations, although the differences are not quite as pronounced. The results for PRMS-P1 show much larger increases in the Cascades, eastern foothills of the Cascades and Northern Rockies. These results are consistent with the results for the ensemble average (Fig 8) the point analyses (Fig. 7), which all show that the PRMS-P1 results are quite different from the three VIC calibrations.

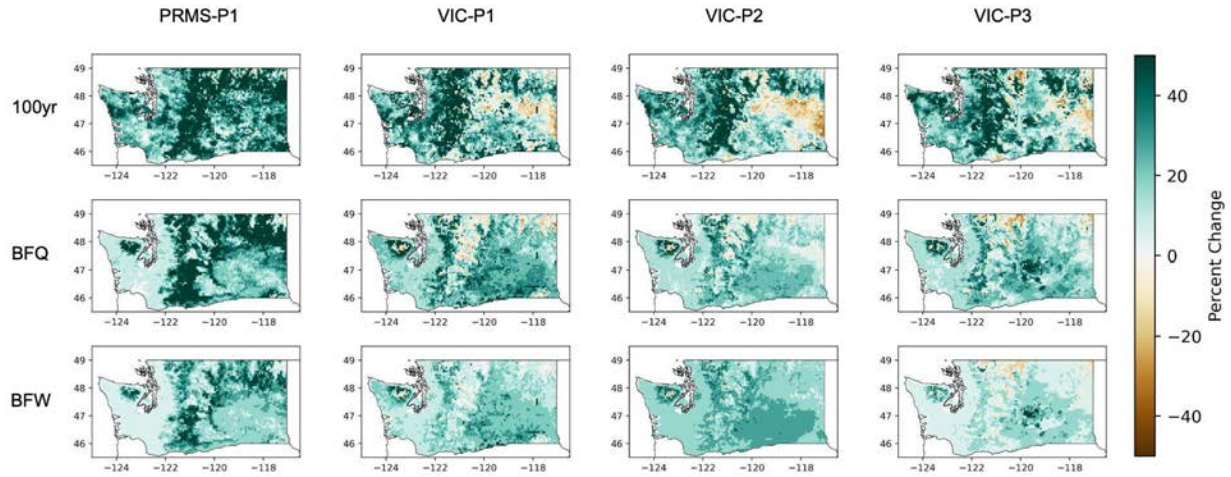

Figure 10. 10-model ensemble median percent change for the MACA downscaled simulations under the high-end greenhouse gas scenario (RCP 8.5) for the 2080s (2070-2099) relative to the 2000s (1990-2019). The columns are the different hydrologic model-calibration pairs. The rows are the results for the 100-year event, bankfull flow (BFQ), and bankfull width (BFW).

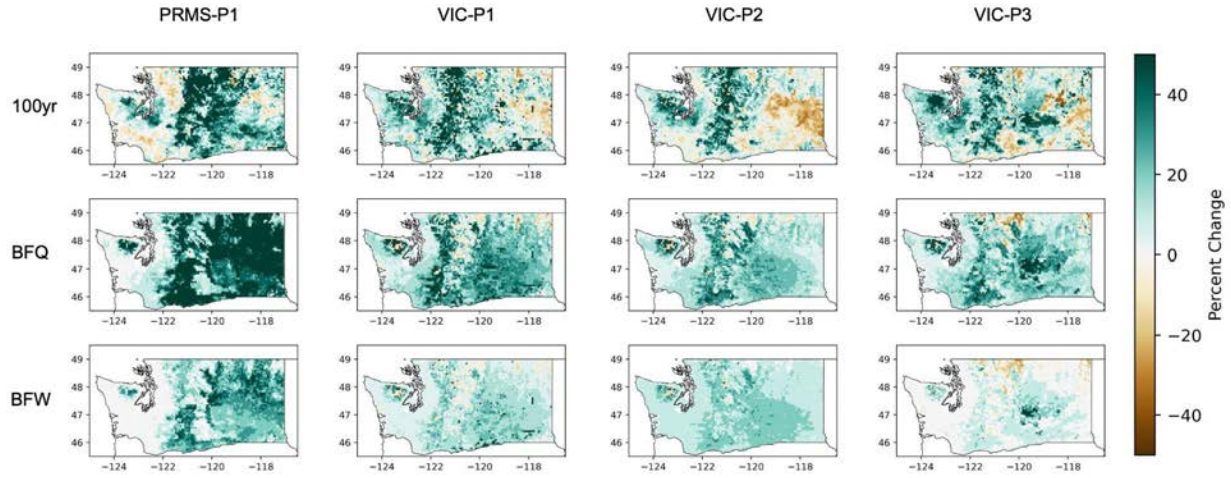

Figure 11. As in Figure 10 except showing results for the BCSD downscaled projections.

Figures 10 and 11 show the ensemble median percent change for the MACA and BCSD sub-ensembles, respectively. Comparing the two, we see that the downscaling method does not change the general pattern of ensemble median percent change in the 100-year event, BFQ, and BFW for the three different VIC calibrations much (Figs. 10,11), similar to the ensemble average percent change figures (Figs. 8,9). Much like the ensemble average, the differences between PRMS-P1 for the MACA and BCSD downscaling methods are much larger. For example, much of western Washington shows an increase in magnitude of the 100-year flood event in the simulations, while the BCSD simulations show a decrease.

The ensemble average (Figs. 8, 9) and the ensemble median (Figs. 10, 11) projections show very similar patterns for each hydrologic model-calibration pair, downscaling method, and metric (100-yr, BFQ, BFW). However, although the patterns are similar the projected changes for the ensemble median are generally slightly smaller than those for the ensemble average, owing to the median placing less weight on the tails of the ensemble distributions. Since both tend to capture the pattern of projected changes in the 100-year runoff event, BFQ, and BFW, the choice to use either the ensemble average or ensemble median is most influenced by the desire to weight the tails of the ensemble distribution in the measure of central tendency. Current practice in climate science favors using the median to measure the central tendency since it is more resistant to outliers in the dataset.

The results from the point analysis (Fig. 7), the ensemble average change (Figs. 8, 9), and ensemble median change (Figs. 10, 11), suggest that the spread within the larger 80-simulation ensemble (i.e., all hydrologic model-calibration pairs, GCMs, and statistical downscaling methods of the high-end greenhouse gas scenario) are introduced through the different hydrologic model-calibration pairs and not the GCMs or choice of statistical downscaling method. PRMS-P1 is the exception to this, with a much larger GCM spread at several points (Fig. 7) and differences in results between downscaling methods (Fig. 8, 9, 10, 11).

Comparing these results with Wilhere et al. (2016), we note that the ensemble average percent change in the 100-year flood event, BFQ, and BFW patterns differ from previous projections. Namely, the decrease in BFQ and BFW around the boundaries of the Columbia Plateau in Wilhere et al. (2016) are not apparent in these results, especially for the PRMS-P1 simulation. The reason for this is unknown and could be investigated in future work

Ensemble Spread Evaluation

Figures 12 and 13 show the model agreement on the sign of change at each grid cell for the MACA and BCSD downscaling methods, respectively. Consistent with previous results, larger areas of white shading (maximum ensemble disagreement on the sign of change) in the 100 year events subplots suggests more spread in the ensemble for this metric. For BFQ and BFW, there is a high degree of agreement in the sign of change for most grid cells, regardless of downscaling method or hydrologic model-calibration pair. PRMS-P1 includes high levels of agreement in positive change in regions where all three VIC simulations project high certainty in negative change. This means that the central tendency results for the individual simulations of the PRMS-P1sub-ensemble are consistent with each other, and differences between PRMS-P1and VIC are not due to a larger model spread or bias in the PRMS-P1 results. The choice of statistical downscaling method does not change the results in most locations, except for a few locations in western Washington.

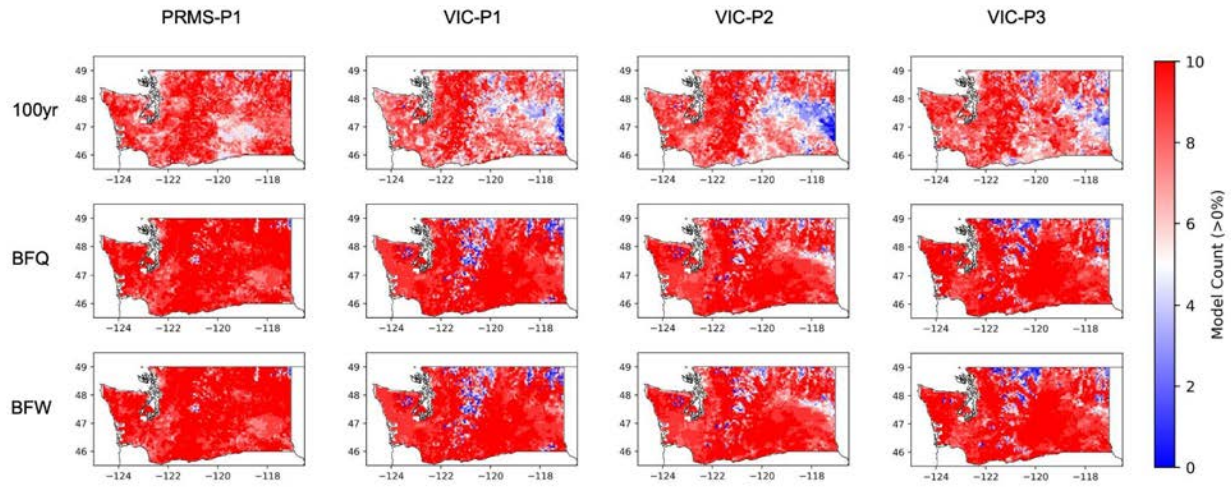

Figure 12. 10-model ensemble agreement in sign of change for the MACA downscaled simulations under the high-end greenhouse gas scenario (RCP 8.5) for the 2080s relative to the 2000s (1990-2019). Red shading denotes most models indicating a positive change, while blue shading denotes negative change. The columns are the different hydrologic model-calibration pairs. The rows are the results for the 100-year event, bankfull flow (BFQ), and bankfull width (BFW).

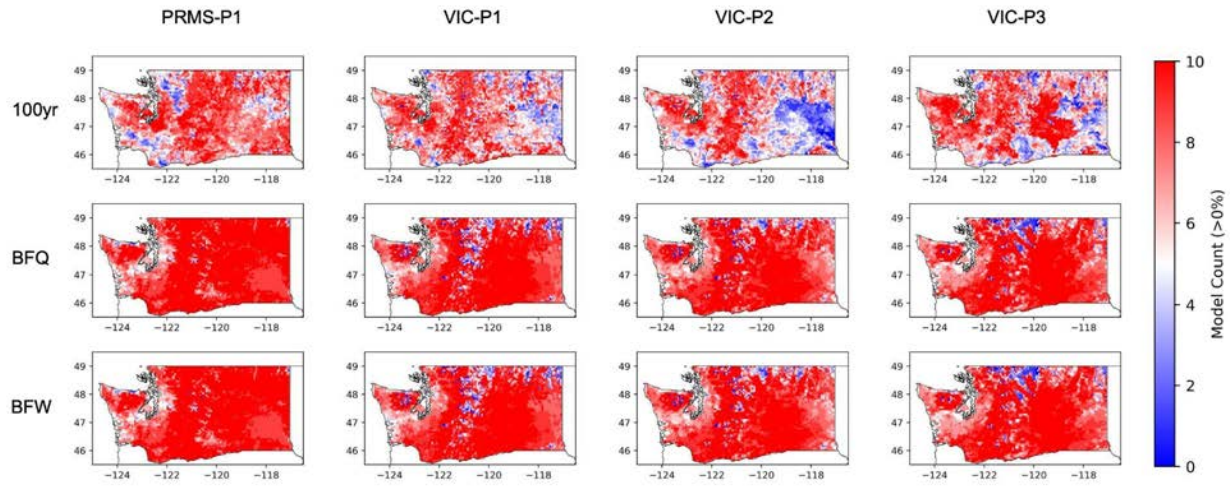

Figure 13. As in Figure 12 except showing results for the BCSD projections.

Figure 14 shows the difference between the ensemble 90th percentile and 10th percentile in percent change for the 100-year peak flow, BFQ, and BFW for the 2080s, based on the MACA downscaling. Figure 15 shows the same but for BCSD downscaling. The range for the 100-year event is particularly large due to the sampling considerations outlined above. The range for BFQ and BFW are much smaller, with most regions in Washington showing differences from 0% to 50%.

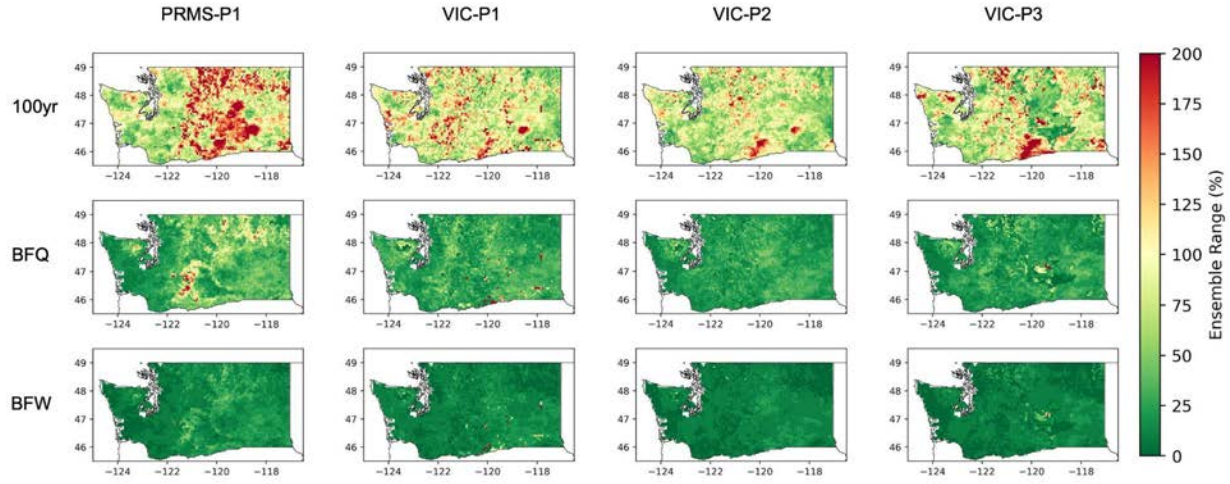

Figure 14. 10-model ensemble 90th percentile - 10th percentile range, in percent change, for the MACA downscaled simulations under the high-end greenhouse gas scenario (RCP 8.5) for the 2080s (2070-2099) relative to the 2000s (1990-2019). The columns are the different hydrologic model-calibration pairs. The rows are the results for the 100-year event, bankfull flow (BFQ), and bankfull width (BFW).

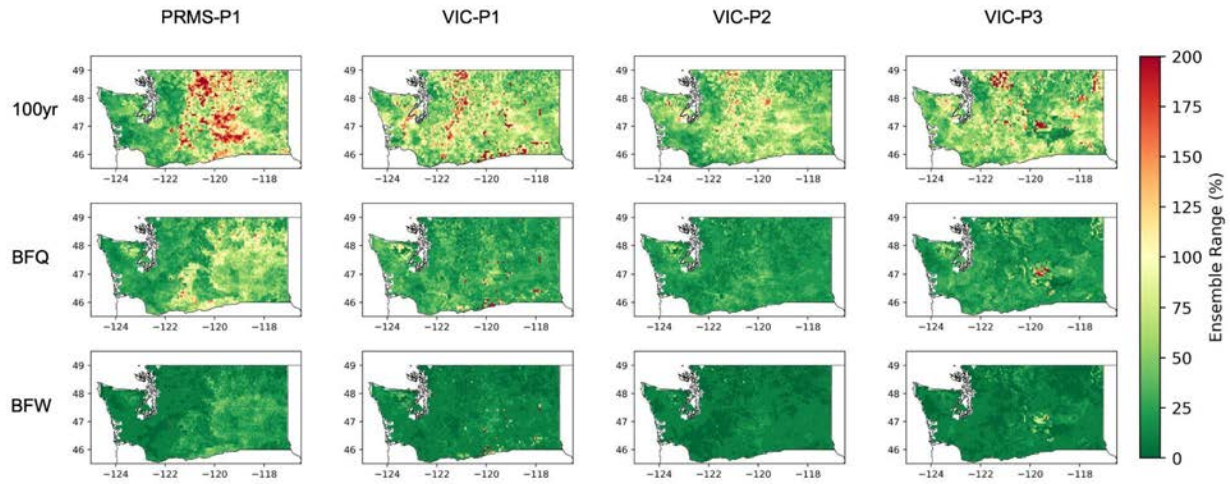

Figure 15. As in figure 14. Except showing results for the BCSD downscaled projections.

Recommendations

RCP 4.5 vs RCP 8.5

The decision of which future emission scenario to use is dependent on the time frame of the decision and the level of acceptable risk. Before 2050, the choice of scenario does not have a major effect on the projections. Later in the 21st century, the RCP 8.5 generally shows larger increases in BFQ and BFW across Washington state. Using results from the RCP 4.5 scenario to inform decisions is therefore a less conservative choice (i.e., less likely to overestimate BFW in the future), whereas RCP 8.5 is more conservative.

Recent research indicates that the RCP 8.5 scenario may be high, that is, future emissions are unlikely to be high enough to match this scenario (e.g. Hausfather and Peters 2020). RCP 4.5, in contrast, is considered to be a more plausible low-end scenario. Nonetheless, a lower-than-expected scenario is more likely to underestimate actual changes that may be experienced in the future. Since there are consequences to both under-design and over-design, the choice of scenario is ultimately a policy decision.

GCMs

We recommended including the entire ensemble of GCMs when using this data. Research consistently shows that using an ensemble of GCMs is the best way to accurately estimate the central tendency in the projections (Kharin and Zwiers 2002). In addition, individual GCMs do not consistently inhabit the low-end, middle, or high-end of the range among the ensemble. For example, in assessing the distribution of projections for the points selected

in Task 1, no single GCM was a consistent outlier in the ensemble (Figs. 6, 7). Finally, the range among GCMs is our best estimate of the uncertainty in climate change projections.

Downscaling Method - BCSD vs. MACA

Compared to the differences in spread and central tendency between the VIC and PRMS-P1 simulations, the two different downscaling methods do not significantly alter the results. There may be justification to use MACA data over BCSD data given evidence suggesting it better captures extreme precipitation changes (Abatzoglou and Brown 2012), but further validation of these results is required before we can definitively say one downscaling method provides better results than another. In addition, the larger uncertainty in downscaling likely comes from the fact that both methods employ a statistical approach, whereas research indicates that dynamical downscaling is needed to accurately estimate changes in precipitation extremes (Salathé et al. 2014). As such, we recommend including both approaches in a larger ensemble of runoff projections.

Hydrologic Model - PRMS vs. VIC

The PRMS-P1 simulation showed results for the 100-year event, BFQ, and BFW that are quite different from the three VIC calibrations. Given the regions with the greatest differences, we hypothesized this could be a result of how the two hydrologic models simulate snowpack. PRMS simulates snowpack for a single elevation – the grid cell average, which can lead to biases in grid cells that have large gradients in elevation. Conversely, the VIC model divides each grid cell into four elevation bands, simulating snowpack independently for each elevation. Comparing the annual maximum snow water equivalent (SWE) for the historical period (1990-2019) in the PRMS-P1 and VIC-P1 simulations, we find PRMS-P1 is systematically larger than VIC-P1 (Fig. S15). This may contribute to the larger increases in the 100-year event, BFQ, and BFW for the 2080s seen in the PRMS-P1 results (Figs. 8, 9, 10,11). This may not be the only distinction between the simulations, and diagnosing the cause of the differences between PRMS-P1 and VIC-P1 is outside the scope of this project. Given the increased weighting to the VIC hydrologic model since the ensemble includes three separate calibrations and only one from PRMS, and that previous model validation does not indicate large biases in the PRMS-P1 streamflow results (Chegwidden et al., 2019), we recommend including the PRMS-P1 results until further evaluation indicates otherwise.

Regionalizing the Projections

There are two reasons to consider regionalizing spatial data. First, this approach increases sample sizes (where n = number of grid cells) for each designated region, which improves the precision of the estimates by averaging out randomly distributed sources of error. Second, regionalization simplifies decision-making. In the case of culverts, this would translate to a simpler design process for planners and engineers. A potential pitfall of regionalization is that it may lump together areas that are meaningfully different. As a result, it's important to consider whether the areas that are grouped together are similar enough to treat as equal, and how much the spatial pattern of the original results is altered in any particular regionalization scheme. Finally, we only consider regionalization approaches that are based on spatially cohesive units. Although other approaches, like clustering, may improve accuracy by averaging across similar conditions, they will rarely be encompassed by simple boundaries and therefore would not simplify the design process. Future work could explore the potential gains in accuracy by nonetheless averaging across grid cells with common response characteristics

In this section we assess how different regionalization schemes impact the central tendency and spread of results within each sub-ensemble, focusing on the practicality and accuracy of each option. Figure 17 shows the BFW regionalization results for geographically-based planning areas using the spatial median of the ensemble median as the measure of central tendency. We tested four schemes: Omernik III and Omernik IV ecological regions, and the HUC8 and HUC12 watershed boundaries. Naturally, we find a trade-off between the similarities between gridded and regionalized data based on the size of each region. Both the Omernik IV results and HUC12 results are more similar to the original gridded results, owing to their smaller regions. The Omernik III and HUC8 regions, while simpler since they include fewer distinct regions, smooth out many distinct features found within the original gridded data. For example, the large gradient in percent change in BFW near the Olympic Mountains is all but missing within both – and within the Omernik IV results as well. This means it is important to consider these results through the lens of culvert replacements: where culverts tend to be located, and how much variation in projections is within the design tolerance. This is likely most important in areas where the risk of underdesign is greatest, or near the BFW threshold that might warrant a bridge over a culvert.

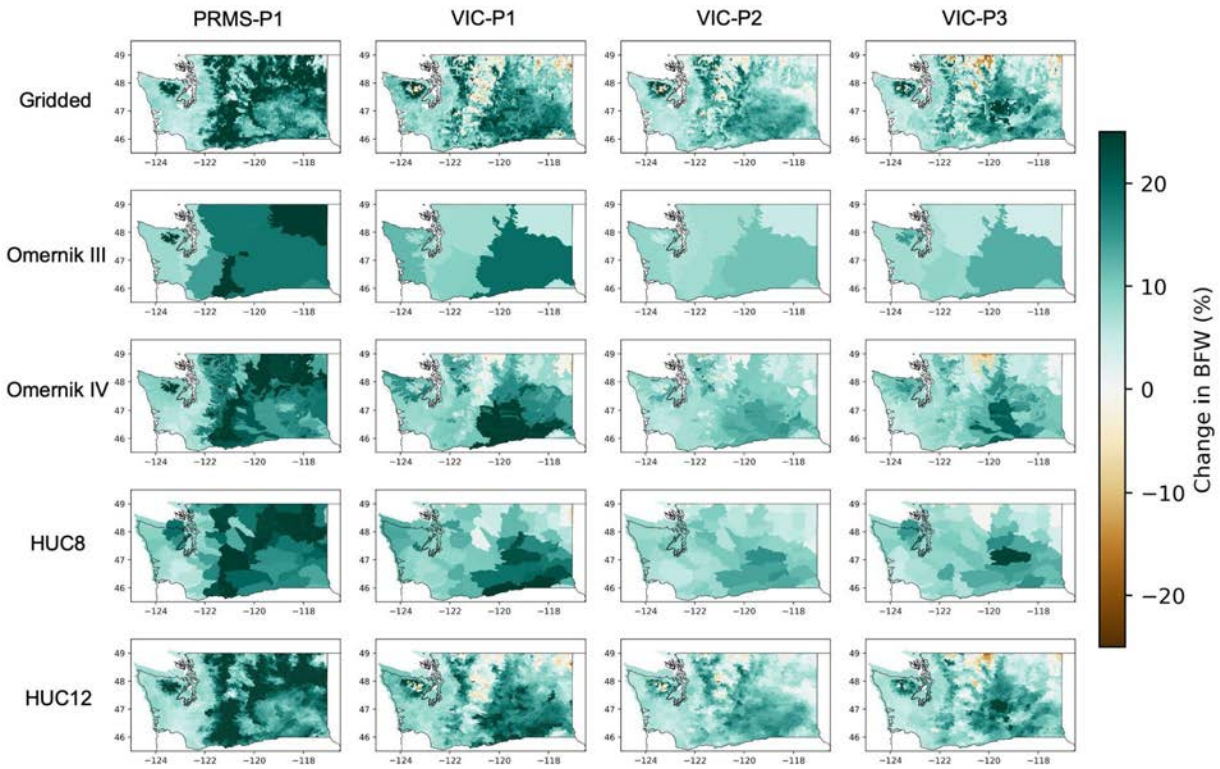

Figure 16. Geographically based ensemble median of the area average regionalization schemes for percent change in BFW for the MACA downscaled simulations under the high-end greenhouse gas scenario (RCP 8.5) for the 2080s (2070-2099) relative to the 2000s (1990-2019). The columns are the different hydrologic model-calibration pairs. The rows are the results for no regionalization (Gridded) and four different schemes.

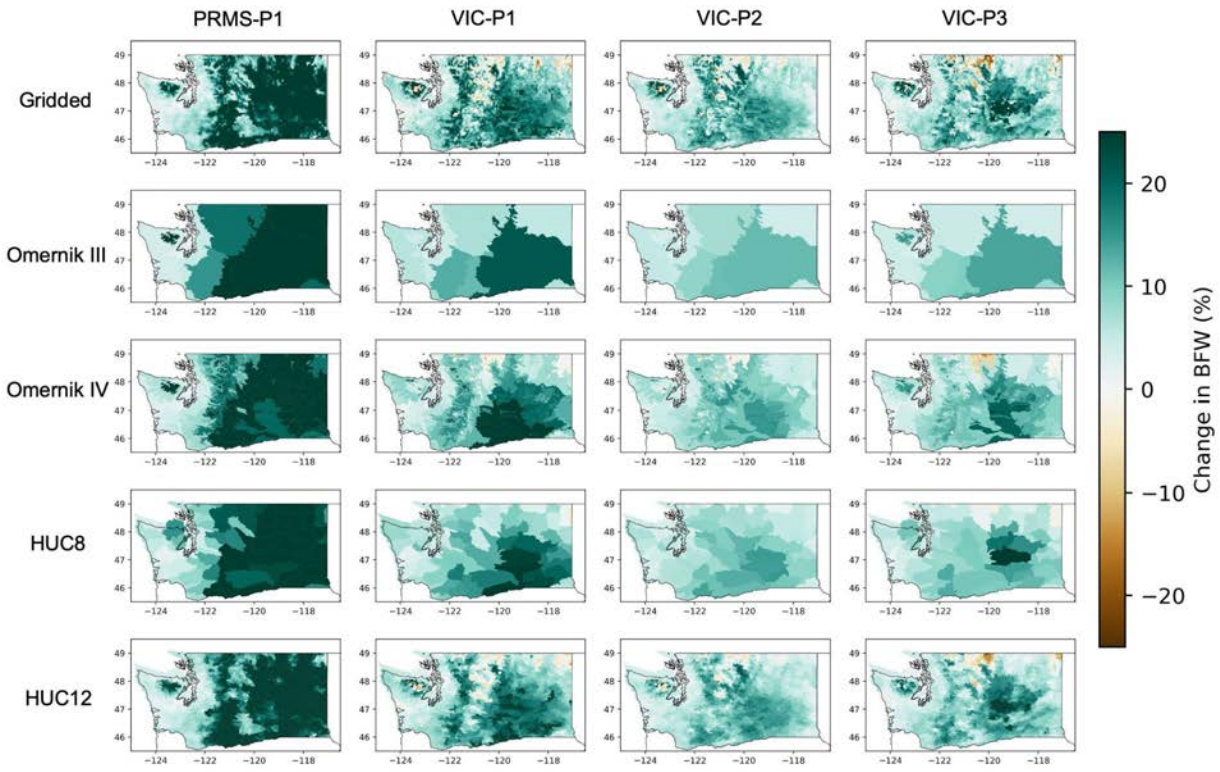

Figure 17. As in figure 16 except showing results for the BCSD downscaled projections.

Another measure of central tendency that is more resistant to intra-regional outliers than the area average is the area median (Figures 18, 19). In general, this method reduces the spatial variability in the results, since the median is unaffected by grid cells that are at the tails of the intra-region distribution. Some regions in the areal average approach (Fig. 16,17) show larger increases in BFW than the areal-median (Fig.18, 19); e.g., south-central Washington).

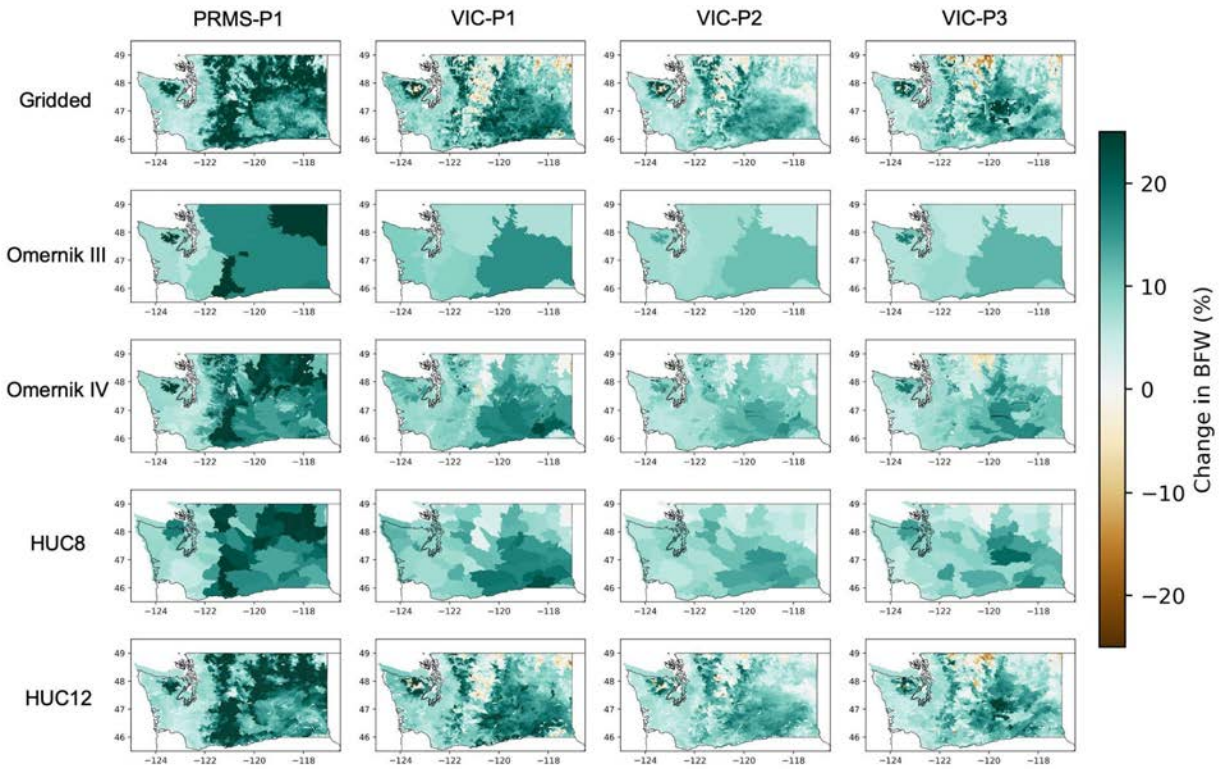

Figure 18. Geographically based ensemble median of the area median regionalization schemes for percent change in BFW for the MACA downscaled simulations under the high-end greenhouse gas scenario (RCP 8.5) for the 2080s (2070-2099) relative to the 2000s (1990-2019). The columns are the different hydrologic model-calibration pairs. The rows are the results for no regionalization (Gridded) and four different schemes.

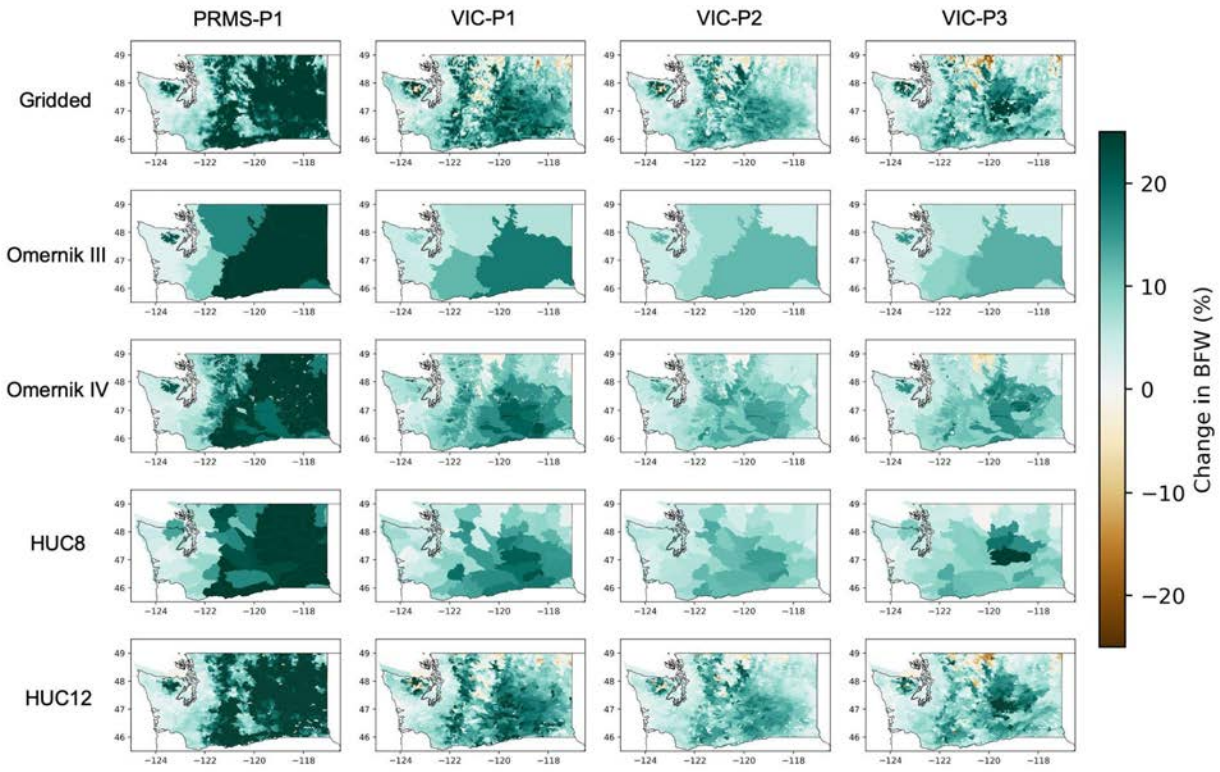

Figure 19. As in figure 18 except showing results for the BCSD downscaled projections.

Finally, we visualize the tradeoffs of the different regionalization schemes in figures 20 and 21, which show the intra-region ranges for each region in each scheme. The larger regions, such as those in the Omernik III and HUC8 schemes, show much larger percent change range within their regions regardless of the downscaling method or hydrologic model - calibration pair. Conversely, the Omernik IV and HUC12 schemes show more cohesion in each of their regions because they are smaller in area and do not encompass as much varied terrain.

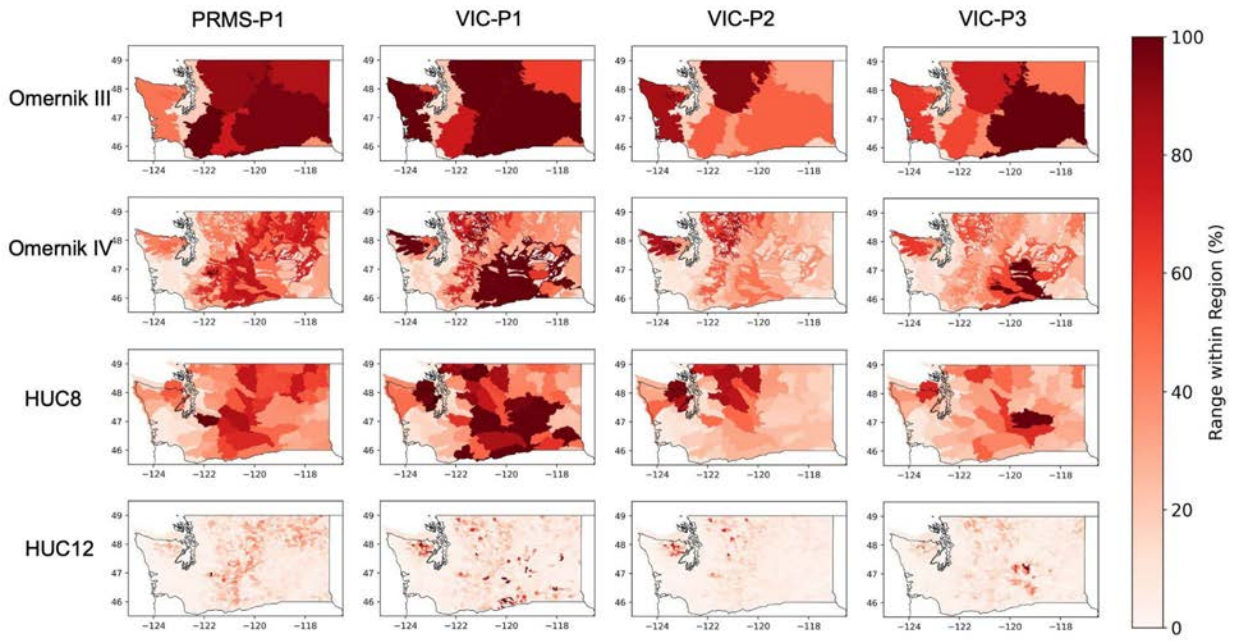

Figure 20. Geographically based range within each region (max - min) of the regionalization schemes for percent change in BFW for the MACA downscaled simulations under the high-end greenhouse gas scenario (RCP 8.5) for the 2080s (2070-2099) relative to the 2000s (1990-2019). The columns are the different hydrologic model-calibration pairs. The rows are the results for the four different schemes.

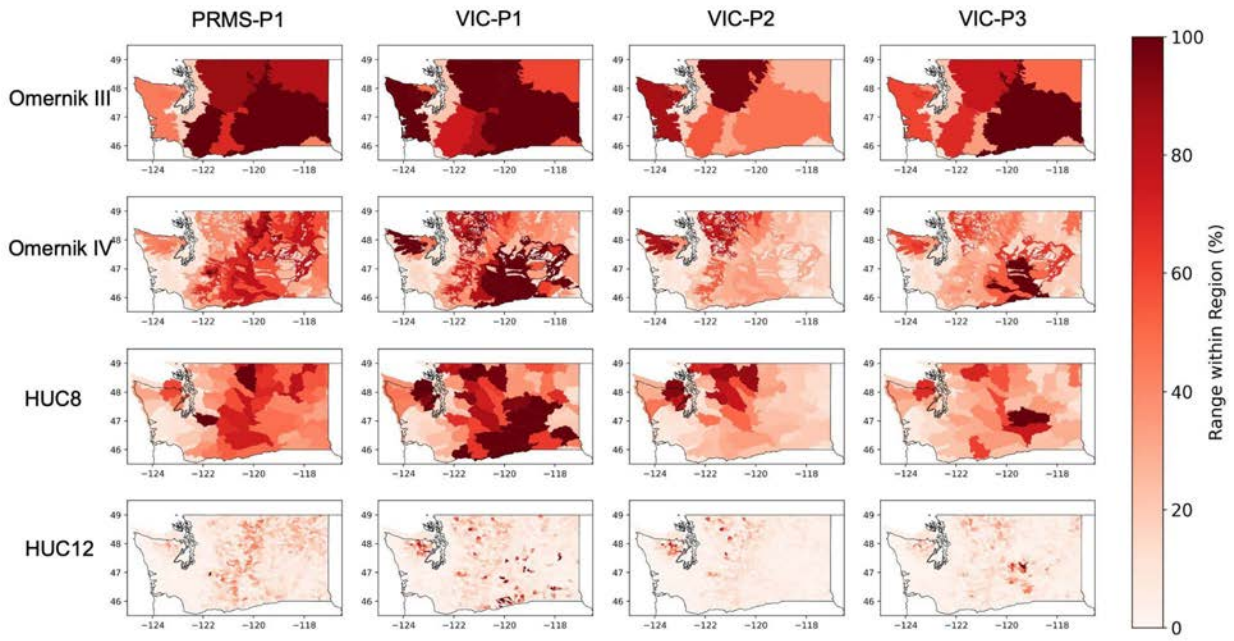

Figure 21. As in figure 22 except showing results for the BCSD downscaled projections.

CONCLUSIONS AND NEXT STEPS

Central Tendency and Range

The spatial patterns of both the ensemble average and ensemble median percent change are similar, and yet the range among hydrologic model-calibration pairs can be quite large. This suggests that the differences between the PRMS-P1and VIC results are not sufficient to skew the results. Nonetheless, given the potential for outliers, we recommend using the ensemble median to evaluate the central tendency in the projections.

We also considered the ensemble spread as an estimate of the uncertainty in the projections. This included the difference between the ensemble 90th percentile and 10th percentile, and the difference between the ensemble maximum and minimum. In situations where a culvert is being scoped or updated with no development upstream, if the design is likely to perform adequately under a wide range of future flow conditions, and in a location where there are minimal risks to the ecosystem, then planning to a lower peak flow may suffice. Conversely, other projects may warrant planning to a higher peak flow percentile or even the ensemble maximum.

In analyzing the different components of the ensemble, we find that the hydrologic model contributes the most to the range among models. PRMS-P1generally projects larger increases in the 100-year peak flow, BFQ, and BFW, compared to VIC. We hypothesize that this may be due in part to how PRMS-P1 simulates snowpack for a single grid cell average

elevation, whereas VIC considers up to four elevation bands within each grid cell. However, determining the exact cause is outside the scope of this project. Despite the difference in snow simulations, we do not currently have evidence to indicate the PRMS-P1 simulations should be excluded from the ensemble.

Regionalization

We explored several different potential regionalization schemes, including Omernik level IV ecoregions, and 8- and 12-digit Hydrologic Unit Code (HUC8, HUC12) basins. Not surprisingly, the coarser schemes (e.g., Omernik level III and HUC8) provide a simpler picture of the changes across Washington State. However, this reduction in spatial resolution could be at the expense of potentially useful regional detail (e.g., projected decreases in the North Cascades). The HUC12 and Omernik IV regionalization schemes show the smaller features with large gradients of change in BFW at the expense of more complexity and potentially a false level of precision.

Future Work

There are a number of additional analyses that could build on the current work. These are enumerated in the list below:

- Some model projections may be outside of the range of what is plausible, either based on observational constraints or professional judgment. Comparing the BFW results here with precipitation projections in the same regions, can help estimate a cap on rain-dominated grid cells that have unrealistic increases (e.g., identify grid cells that have a larger BFW increase than makes sense with the precipitation projections). This may be a helpful criteria for diagnosing model inaccuracies and excluding specific projections.
- These models have not been validated at the scale needed for culvert design. For instance, Chegwidden et al. (2019) focused on big river locations such as the Columbia, Snake, and Willamette. Additional validation is needed to better understand model performance, and how that varies spatially, across Washington State.
- Additional model diagnosis and sensitivity testing is needed to understand the mechanisms governing some of the unexpected aspects of these and previous projections. Specifically: (a) why the PRMS-P1 results differ considerably from the three VIC calibrations, and (b) why the Wilhere et al. (2016) projections show decreases along the margins of the Columbia Plateau, whereas the Chegwidden et al. (2019) results do not. This work could be further expanded to evaluate the

sensitivities of the projections more broadly, thereby identifying areas where uncertainties in the meteorology or calibrations may be most important.

- The current projections are all based on statistical downscaling, which previous research has shown does not provide accurate estimates of changes in precipitation extremes (Salathé et al. 2014). Additional analysis is needed to understand the differences between existing statistical and dynamically downscaled projections, and how those may affect the results. Further work could develop a new set of hydrologic model projections based on dynamically downscaled data.
- The current projections are also based on the previous generation of climate model projections: CMIP5. A new generation is now available: CMIP6. Additional analysis could be performed to understand how the CMIP6 projections differ from those of CMIP5, and if those differences are relevant to climate-resilient culvert design.
- This study, as with all related studies since Wilhere et al. (2016), is based on the empirical relationships between bankfull flow and bankfull width derived by Castro and Jackson (2001). That analysis was generally focused on larger rivers, which are less relevant to culvert design. In addition, more than 20 years of new observations are now available from which to evaluate new locations and obtain better statistics from places with a continuous record. Initial analyses should first evaluate the sensitivity of the projections to uncertainties in these empirical relationships. If these uncertainties prove important, then it may be beneficial to initiate a new study, aimed at developing empirical relationships that are more tailored to culvert design.
- Additional analyses could be performed using statistical smoothing or clustering within each region of a geographically-based regionalization scheme. This could prevent unintentional smoothing over important regional details in BFW changes, while also avoiding false levels of precision.

Although likely not comprehensive, these additional analyses illustrate the breadth of studies that could further inform climate-adapted culvert design in Washington State.

REFERENCES

Abatzoglou J.T. and Brown T.J., 2012: A comparison of statistical downscaling methods suited for wildfire applications. International Journal of Climatology, doi:10.1002/joc.2312.

Chegwidden, O. 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.

Hausfather, Zeke & Peters, Glen. (2020). Emissions – the 'business as usual' story is misleading. Nature. 577. 618-620. 10.1038/d41586-020-00177-3.

Kharin, V. V., & Zwiers, F. W. (2002). Climate Predictions with Multimodel Ensembles, *Journal of Climate*, *15*(7), 793-799. doi: https://doi.org/10.1175/1520-0442(2002)015<0793:CPWME>2.0.CO;2

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

Salathé, E. P., Hamlet, A. F., Mass, C. F., Lee, S. Y., Stumbaugh, M., & Steed, R. (2014). Estimates of twenty-first-century flood risk in the Pacific Northwest based on regional climate model simulations. Journal of Hydrometeorology, 15(5), 1881-1899.

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.

Van Vuuren, D. P., Edmonds, J., Kainuma, M., Riahi, K., Thomson, A., Hibbard, K., ... & Rose, S. K. (2011). The representative concentration pathways: an overview. Climatic change, 109(1), 5-31.

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, doi:10.1023/B:CLIM.0000013685.99609.9e.

Wilhere, G.F., J.B. Atha, T. Quinn, I. Tohver, and L. Helbrecht. (2017). Incorporating climate change into culvert design in Washington State, USA. Ecological Engineering 104:67-79.

SUPPLEMENTAL INFORMATION

Analysis of RCP 4.5 Projections

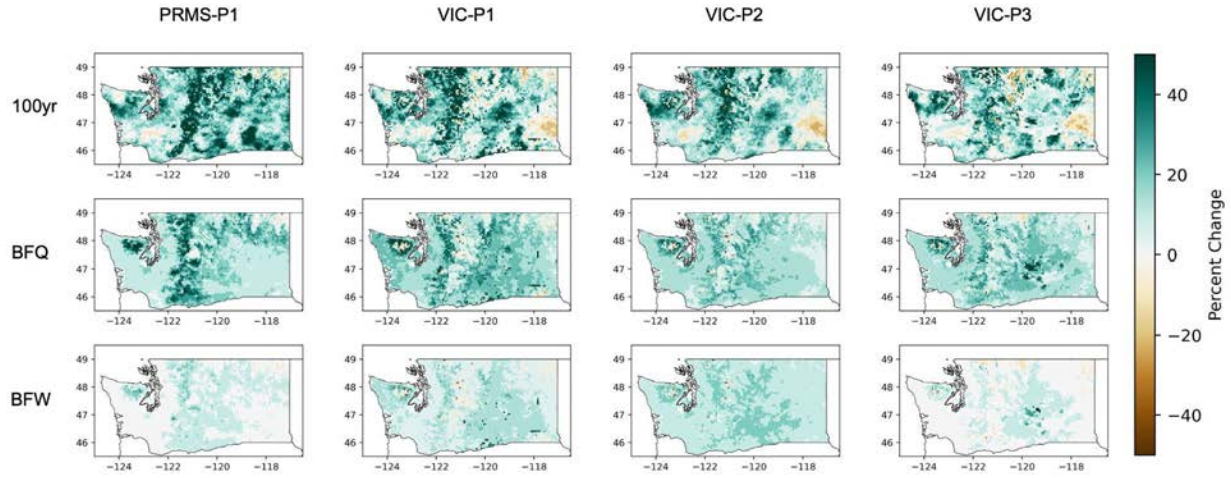

Figure S1. 10-model ensemble average percent change for the MACA downscaled simulations under the low-end greenhouse gas scenario (RCP 4.5) for the 2080s relative to the 2000s (1990-2019). The columns are the different hydrologic model-calibration pairs. The rows are the results for the 100-year event (top), bankfull flow (BFQ, middle), and bankfull width (BFW, bottom).

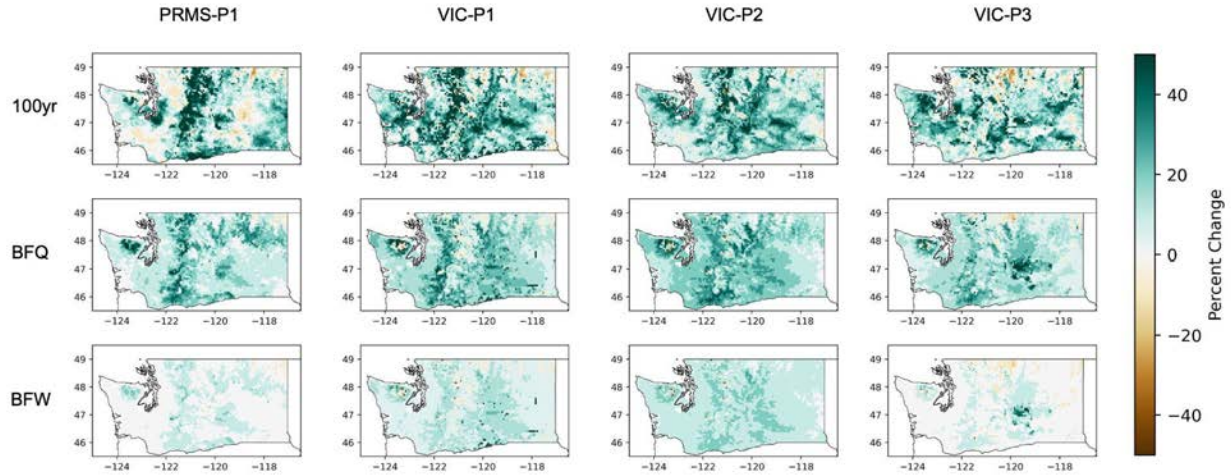

Figure S2. As in figure S1 except showing results for the BCSD downscaled projections.

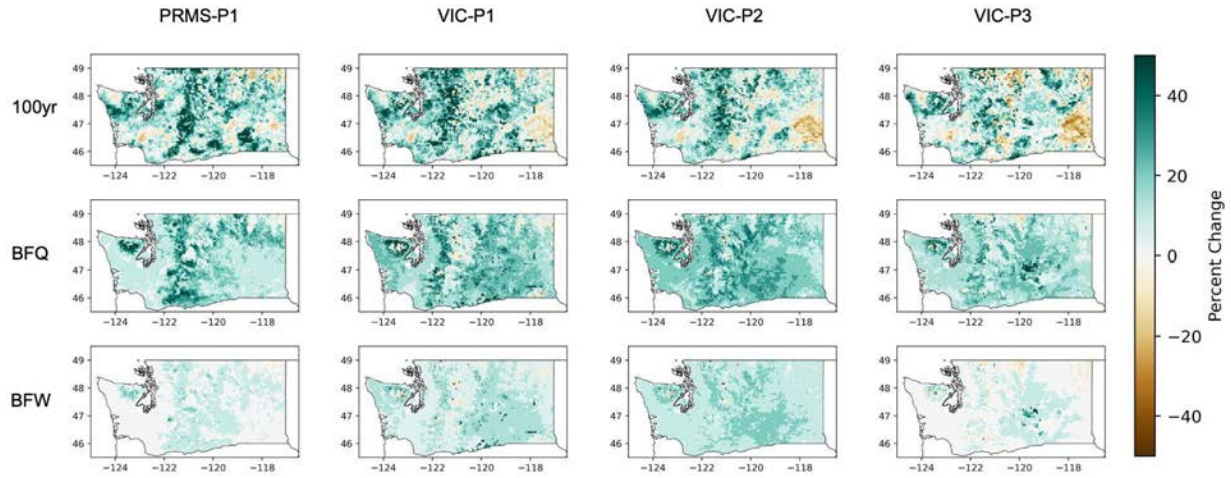

Figure S3. 10-model ensemble median percent change for the MACA downscaled simulations under the low-end greenhouse gas scenario (RCP 4.5) for the 2080s relative to the 2000s (1990-2019). The columns are the different hydrologic model-calibration pairs. The rows are the results for the 100-year event (top), bankfull flow (BFQ, middle), and bankfull width (BFW, bottom).

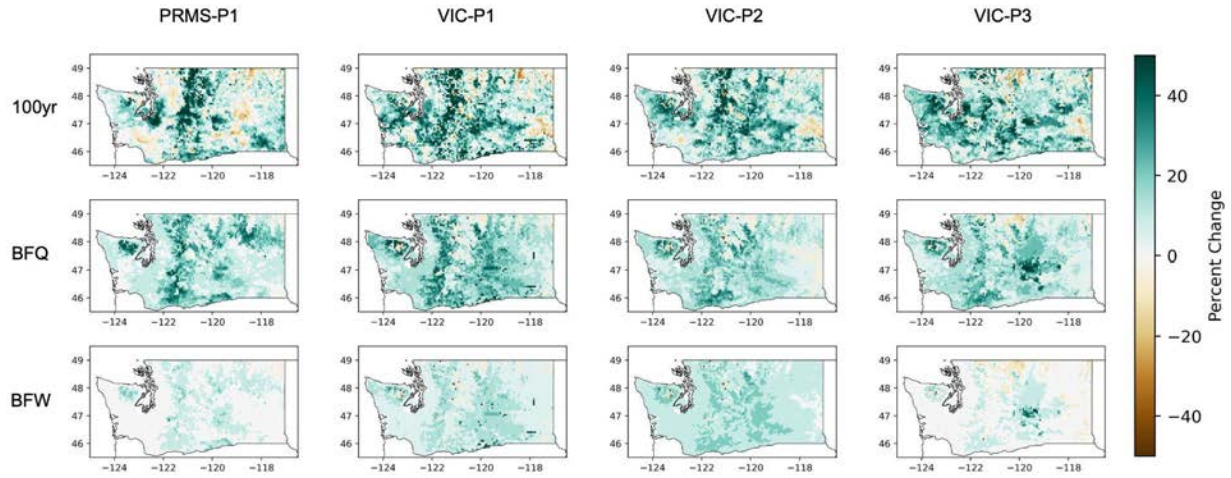

Figure S4. As in figure S3 except showing results for the BCSD downscaled projections.

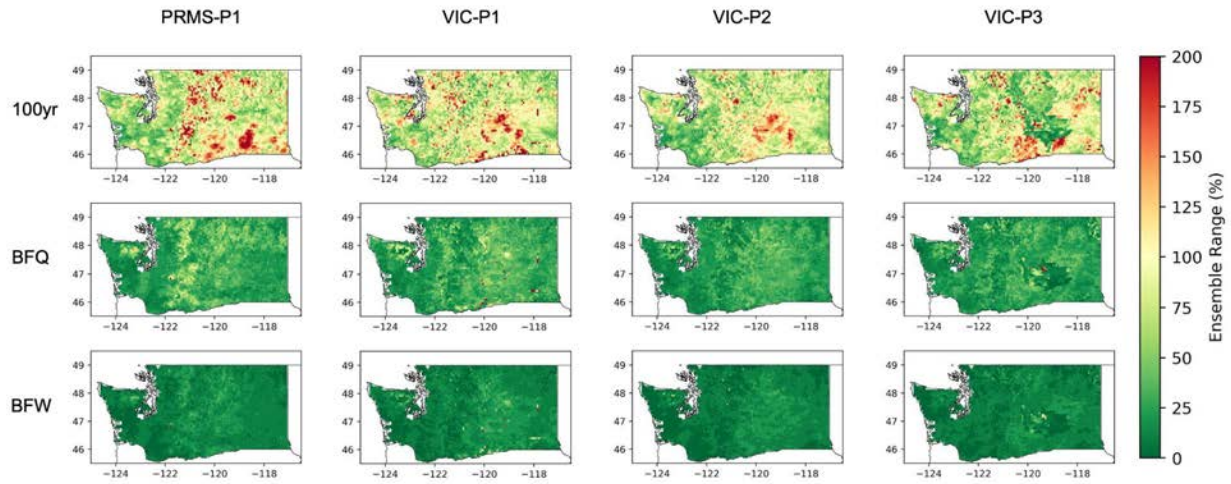

Figure S5. 10-model ensemble 90th percentile - 10th percentile range, in percent change, for the MACA downscaled simulations under the low-end greenhouse gas scenario (RCP 4.5) for the 2080s (2070-2099) relative to the 2000s (1990-2019). The columns are the different hydrologic model-calibration pairs. The rows are the results for the 100-year event, bankfull flow (BFQ), and bankfull width (BFW).

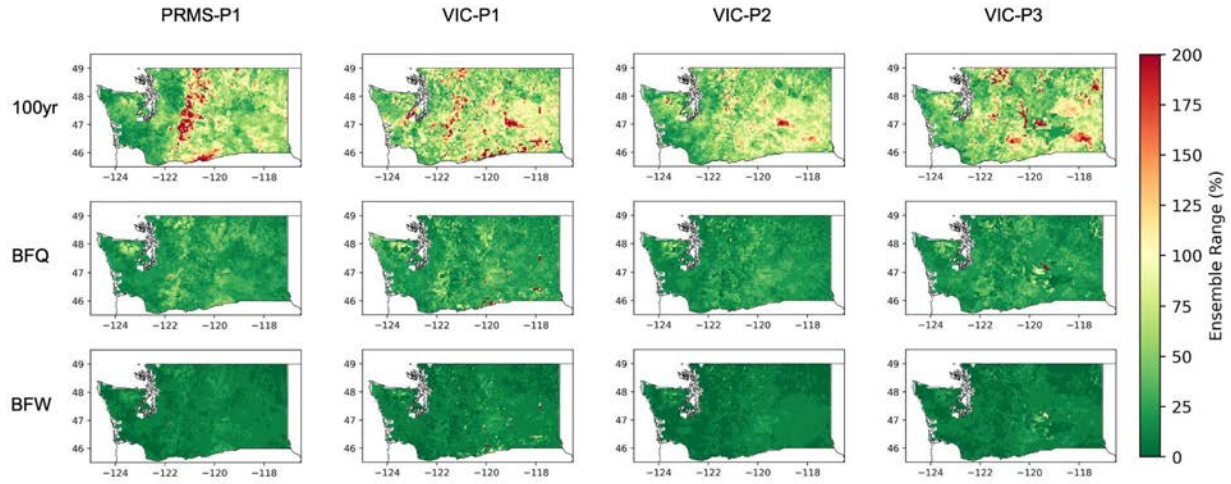

Figure S6. As in figure S5 except showing results for the BCSD downscaled projections.

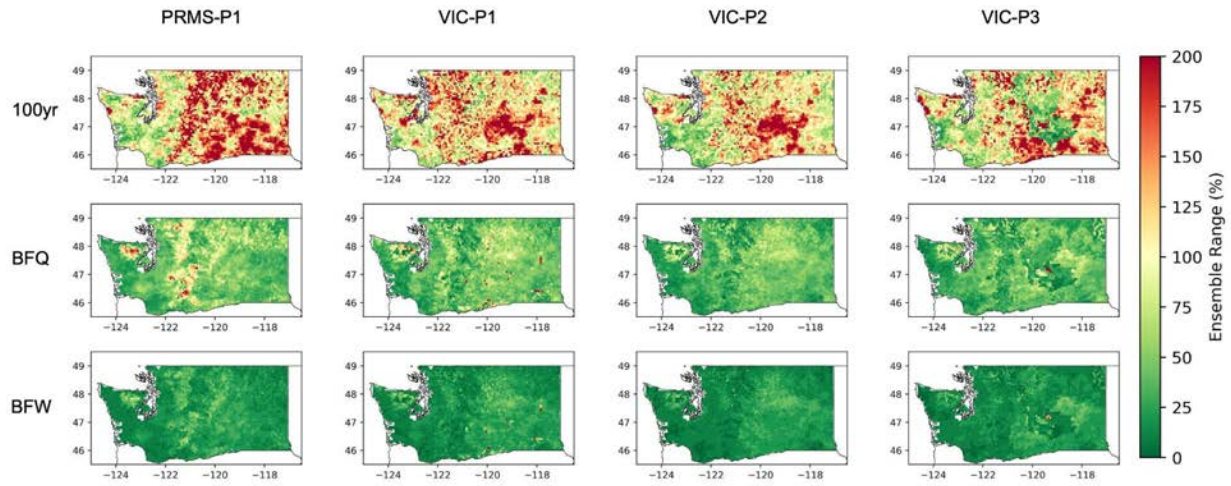

Figure S9. 10-model ensemble range in percent change for the MACA downscaled simulations under the low-end greenhouse gas scenario (RCP 4.5) for the 2080s relative to the 2000s (1990-2019). The columns are the different hydrologic model-calibration pairs. The rows are the results for the 100-year event (top), bankfull flow (BFQ, middle), and bankfull width (BFW, bottom).

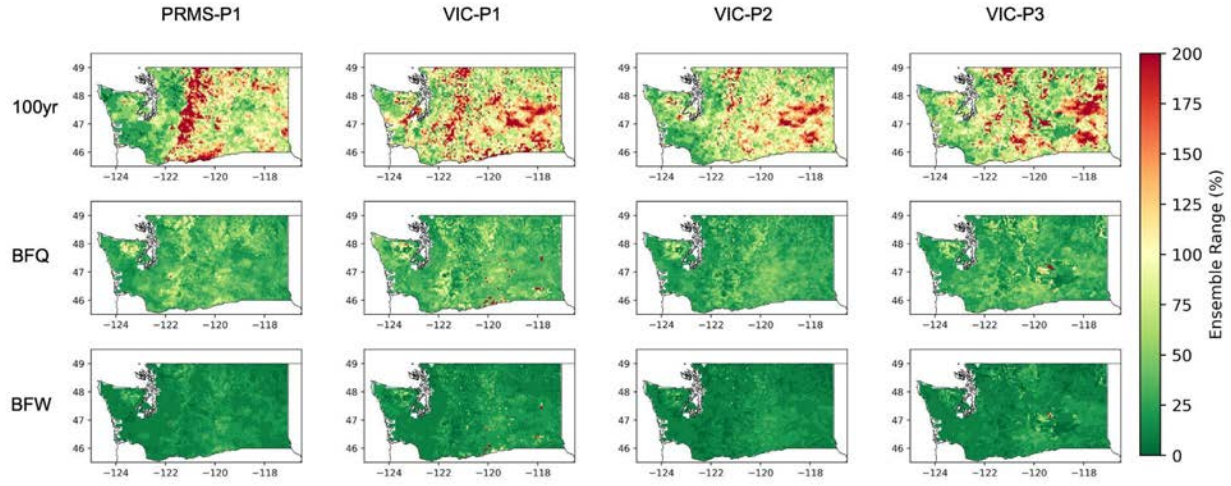

Figure S10. As in figure S7 except showing results for the BCSD downscaled projections.

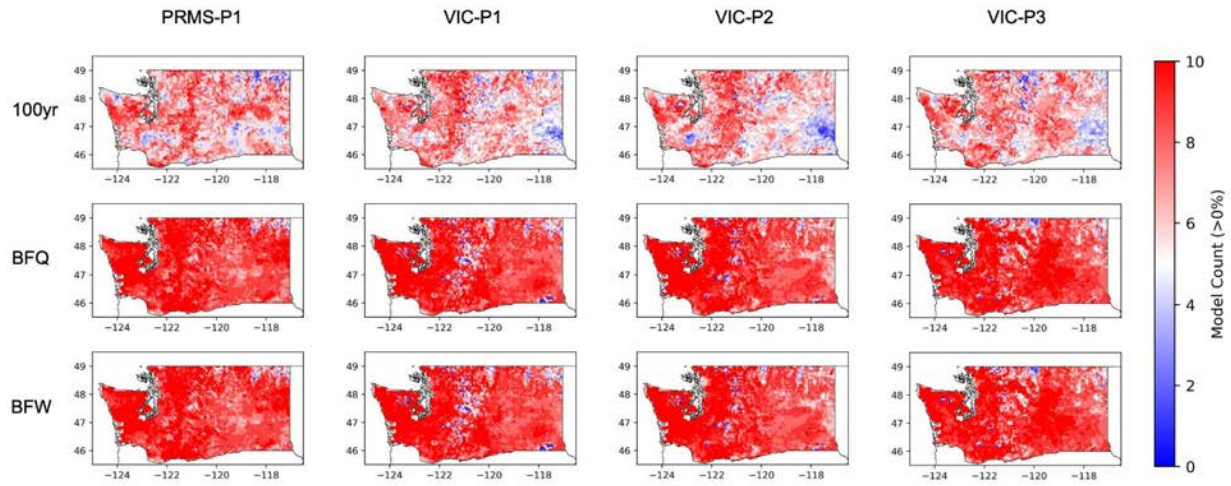

Figure S11. 10-model ensemble agreement in sign of change for the MACA downscaled simulations under the low-end greenhouse gas scenario (RCP 4.5) for the 2080s relative to the 2000s (1990-2019). Red shading denotes most models showing a positive change, while blue shading denotes negative change. The columns are the different hydrologic model-calibration pairs. The rows are the results for the 100-year event (top) (top), bankfull flow (BFQ, middle), and bankfull width (BFW, bottom).

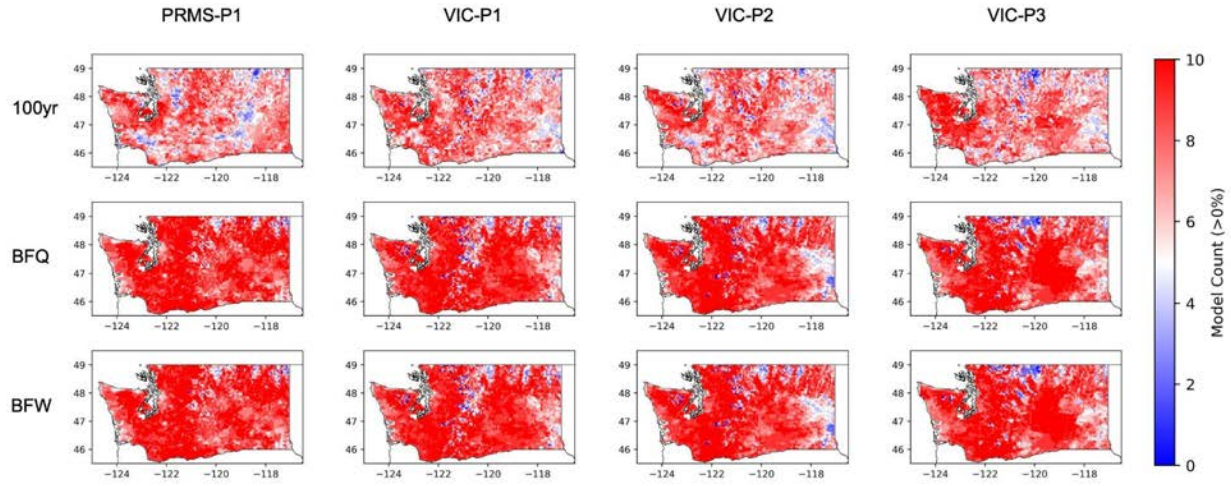

Figure S12. As in figure S9 except showing results for the BCSD downscaled projections.

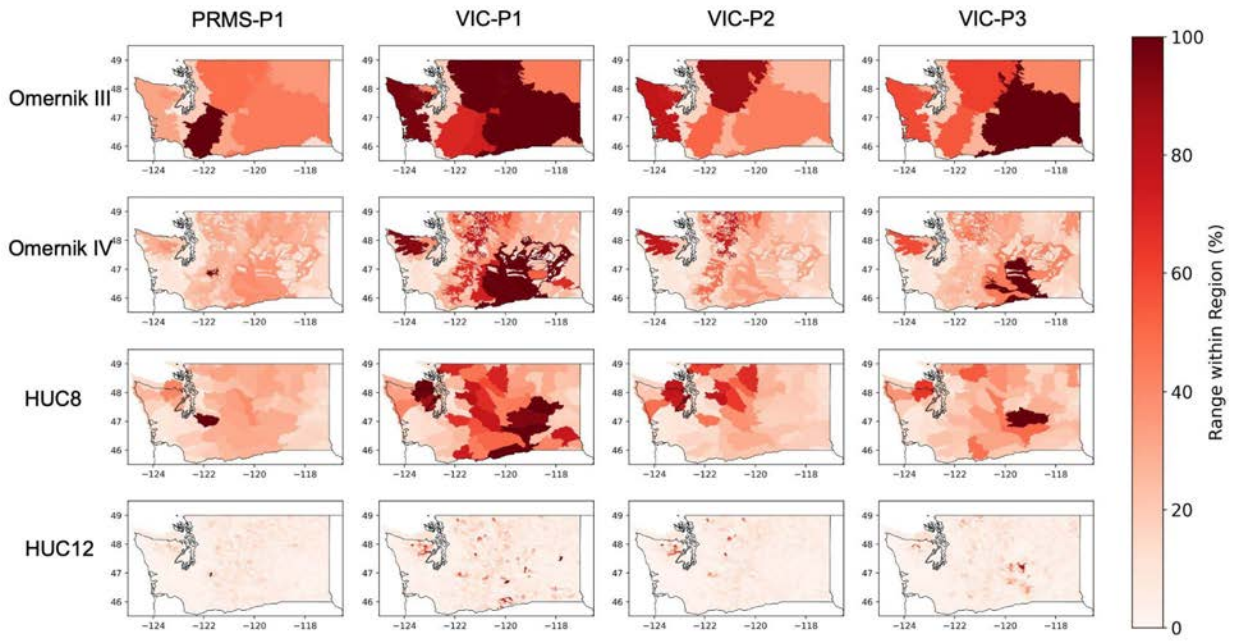

Figure S13. Geographically based range within each region (max - min) of the regionalization schemes for percent change in BFW for the MACA downscaled simulations under the high-end greenhouse gas scenario (RCP 4.5) for the 2080s (2070-2099) relative to the 2000s (1990-2019). The columns are the different hydrologic model-calibration pairs. The rows are the results for the four different schemes.

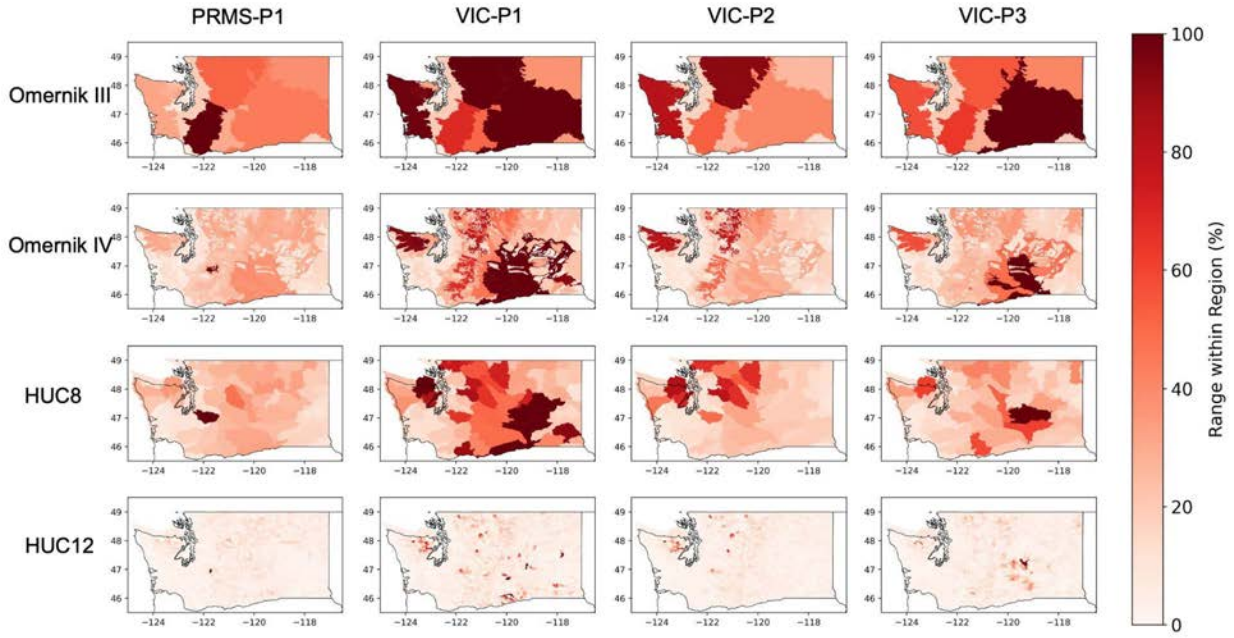

Figure S14. As in figure S13 except showing results for the BCSD downscaled projections.

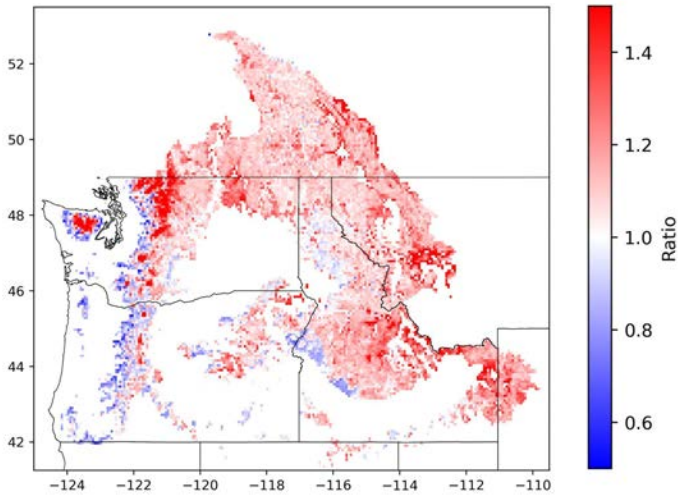

Figure S15. Ratio of the 10-model ensemble median annual maximum snow-water equivalent (SWE) for the PRMS-P1 results to the VIC-P1 model results for the 2000s (1990-2019). Red shading denotes positive bias in annual maximum SWE and blue denotes negative bias.

Additional RCP 8.5 Analyses

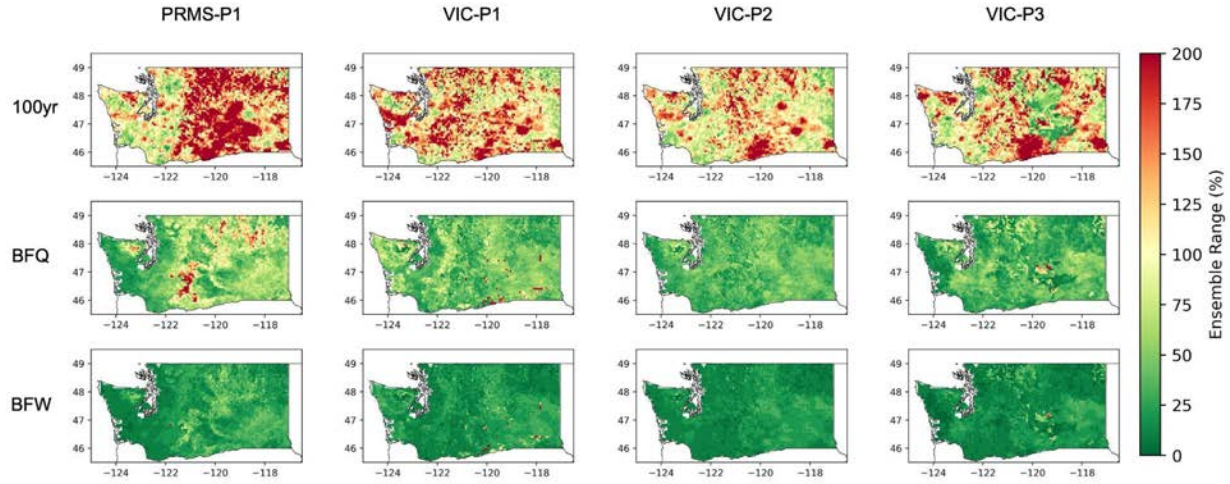

Figure S16. Model range in percent change for the MACA downscaled simulations under the high-end greenhouse gas scenario (RCP 8.5) for the 2080s relative to the 2000s (1990-2019). The columns are the different hydrologic model-calibration pairs. The rows are the results for the 100-year event (top), bankfull flow (BFQ, middle), and bankfull width (BFW, bottom).

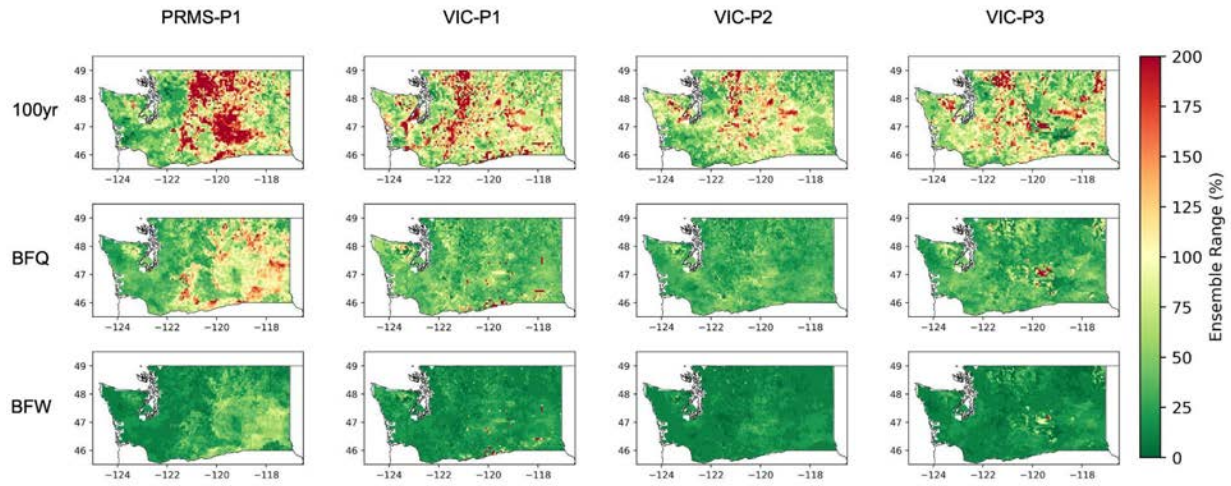

Figure S17. As in figure S11 except showing results for the BCSD downscaled projections.