

Climate robust culvert design: Probabilistic estimates of fish passage impediments for Washington State



Prepared by the
Climate Impacts Group | University of Washington
June 2018



Guillaume Mauger, University of Washington, Climate Impacts Group
Se-Yeun Lee, University of Washington, Climate Impacts Group
Jason Won, University of Washington, Climate Impacts Group
Kyuhyun Byun, University of Notre Dame
Alan F. Hamlet, University of Notre Dame

Cover image: Sockeye salmon attempting to jump into an elevated culvert outlet
(Source: Kitsap Sun; March 29, 2013; reprinted with permission)

Acknowledgements

The authors would like to thank Anne-Marie Ou and Daniel Larsson for their invaluable help in producing the online tool. Jane Atha, George Wilhere, Kevin Lautz, Lynn Helbrecht, and Tim Quinn at the Washington Department of Fish and Wildlife (WDFW) provided helpful feedback on the project and ideas about how to coordinate WDFW and SC² efforts. Credit goes to Ingrid Tohver for supporting the initial integration of climate change data by WDFW. Larry Wasserman and Carol MacIlroy contributed important insights into how to make the tool user friendly and relevant for those considering culvert sizing in the face of climate change. Roger Fuller, Crystal Raymond, and Jon Riedel provided helpful feedback and comments on this report. This work was completed through the Skagit Climate Science Consortium (SC²) with funding from the Environmental Protection Agency (EPA).

Recommended Citation

Mauger, G.S., S.Y. Lee, J.S. Won, K. Byun, and A.F. Hamlet, 2018. Climate robust culvert design: Probabilistic estimates of fish passage impediments. Final report for the Skagit Climate Science Consortium. Climate Impacts Group, University of Washington, Seattle.

Table of Contents

EXECUTIVE SUMMARY	4
BACKGROUND	6
TASK 1: CLIMATE DATA VALIDATION	11
OBSERVATIONAL DATA	11
MODEL DATA	12
RESULTS	12
TASK 2: STREAMFLOW VALIDATION	17
OBSERVATIONAL DATA	17
MODEL DATA	18
RESULTS	19
TASK 3: ESTIMATING THE LIKELIHOOD OF CULVERT “FAILURE”	22
DATA	22
METHODS	23
Calculating Changes in Bankfull Width (BFW)	23
Estimating the Likelihood of Failure	27
RESULTS	28
CONCLUSIONS	29
REFERENCES	30

Executive Summary

This report describes a new tool that is designed to support climate-robust culvert design. In addition, the report describes an evaluation of the meteorological and streamflow datasets that are the basis for the tool, with the goal of supporting the use and interpretation of the results. Building on existing efforts, this work consisted of three tasks:

Task 1: Climate Data Validation.

We compared available historical climate datasets against independent observations for a few weather stations in or near the Skagit basin. Although limited by a lack of independent observations, the analysis suggests that monthly total and water year maximum precipitation are both well-captured by the dataset used to develop the new tool.

Task 2: Streamflow Data Validation.

We evaluated historical streamflow simulations for currently-available datasets, comparing observations with simulated flows for 11 small streams in the vicinity of the Skagit basin. The results indicate that the primary dataset used in the tool, and in previous work, performs about as well as other readily-available historical streamflow datasets. Most sites show good agreement between observed and modeled streamflow, although simulations for a few of the smallest low elevation sites differ substantially from the observations.

Task 3: Probabilistic Estimates of Culvert “Failure”.

We developed an approach to estimating the probability of culvert failure over a given design lifetime, defining “failure” to mean that there is a statistically significant increase in bankfull width (BFW) for a given year within the design lifespan, such that a culvert no longer meets the state’s established design criteria for culvert size as a function of BFW. A new online tool (Figure 1) allows users to query climate change projections and evaluate the probability of failure for different culvert sizes selected by the user. Specifically, the tool allows users to pick their location of interest, and then asks them to enter three pieces of information:

1. Current bankfull width,
2. Design lifetime,
3. Proposed culvert width

Based on these criteria, the tool calculates the probability that the proposed culvert size would fail to meet the state’s design criteria over its design lifetime.

By emphasizing relative changes in streamflow (% change) our approach minimizes the influence of model biases. The approach is also an improvement over past studies in that

it provides a specific estimate of the probability of failure during the design lifespan, integrating the range of model projections into a single number summarizing the implications for a particular culvert size and location. The tool is flexible and designed to accommodate a range of different design lifetimes and proposed culvert sizes.

The following sections describe the results of these three tasks. Additional results can be found online at the project website:

<https://cig.uw.edu/our-work/decision-support/culvert-phase-2/>

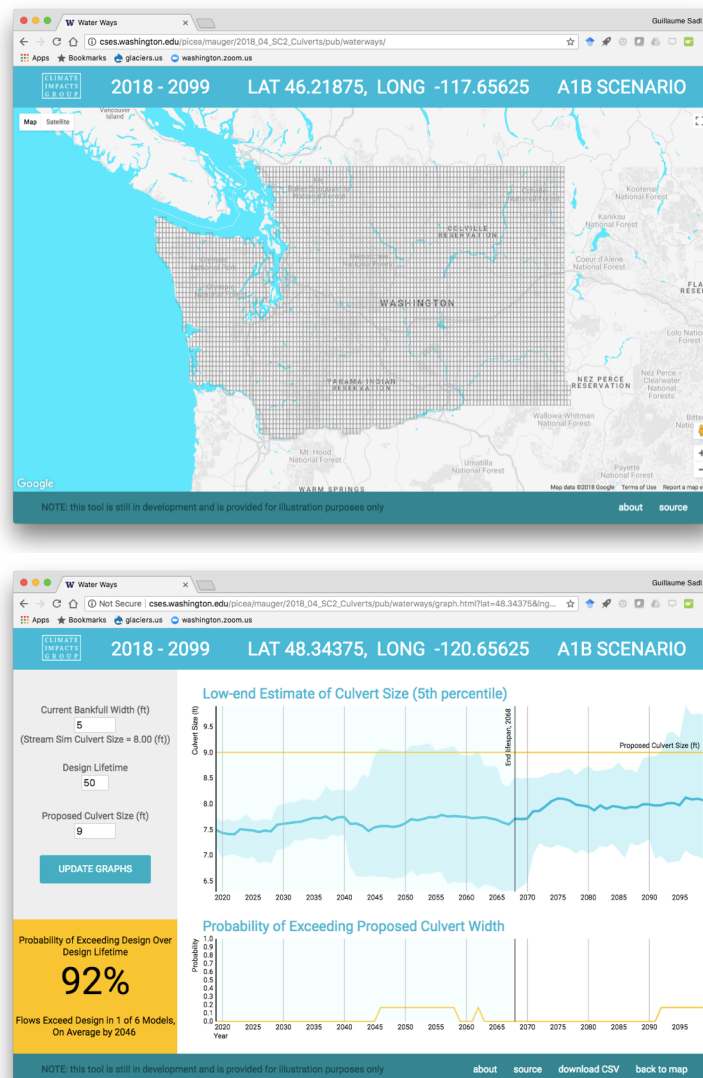


Figure 1. Screenshots of the culvert design tool. The top panel shows the first screen, where the user selects the area of interest. The bottom panel shows the subsequent screen, in which the user can enter design parameters and evaluate the likelihood that a culvert will not continue to meet the Stream Sim standard as a result of future changes in streamflow.

Background

Apart from a few special cases, Washington State culverts are required to be sized based on the Stream Simulation (or “Stream Sim”) guidance issued by the Washington Department of Fish and Wildlife (WDFW; Barnard et al. 2013, 2015). In this approach, culverts are sized based on a simple linear function of BFW:

$$\text{Culvert Width} = 1.2 \times \text{BFW} + 2 \text{ ft} \quad \text{Eq. 1}$$

This reflects a geomorphic approach to culvert design that is intended to be applicable across a large range of situations. Bankfull width is a well-established metric that is generally straightforward to measure in the field, though in some instances it can be obscured by debris flows, vegetation loss, or the presence of multiple channels.

At larger stream widths, there can be important trade-offs. For instance, larger culverts may require a wider road “prism” – the sloped dirt and rock bordering elevated roadways that provide structural support – in order to accommodate the larger height of the road relative to the bottom of the culvert. This increase in width necessitates a longer culvert, and the additional fill can further disrupt either in-stream or adjacent habitat. Similarly, WDFW recommends a bridge in lieu of a culvert for streams with BFW greater than 15 ft. This can significantly increase the cost of a stream crossing.

Future changes in BFW have previously been estimated by Wilhere et al. (2016) using hydrologic projections developed by Hamlet et al. (2013). Wilhere et al. estimated the percent change in BFW derived from projected changes in runoff. This percent change can then be applied to direct observations of channel width. This approach reduces the influence of bias in the hydrologic simulations.

There are two primary challenges to estimating future changes in BFW from streamflow simulations: (1) Current modeling cannot reliably capture the processes responsible for determining channel width (erosion, deposition, etc.), and (2) Uncertainties in climate change projections lead to a wide range among estimates of future BFW. Past work, described below, has developed technical approaches to addressing these challenges, and this report describes some new advances that facilitate use of the projections in spite of these uncertainties. In addition, users will need to decide – or obtain guidance – on what level of risk is acceptable.

Wilhere et al. (2016) address the first issue by using results from Castro and Jackson (2001), which evaluated the relationship between observed streamflow and channel geometry for a wide range of streams across the Pacific Northwest. In that study, Castro and Jackson relate BFW and peak streamflow statistics via an exponential relationship:

$$\text{BFW} = \alpha Q^\beta \quad \text{Eq. 2}$$

Where Q is the peak flow statistic for a particular return frequency (or, alternatively, the annual chance of exceedance) and α and β are fitted parameters that relate streamflow

to BFW. Different sets of parameters were estimated for three different ecoregions in Washington State: Pacific Maritime Mountains, the Western Cordillera, and the Columbia Basin (Table 1, Figure 2). In Castro and Jackson (2001), as in our analysis, streamflow statistics were estimated from the water year maxima in daily observed flows.

There are limitations to this approximation. First, the Castro and Jackson analysis is focused on the relationship between flow and channel width in large rivers. This approximation is unlikely to be as pertinent for small culvert streams, particularly given the relatively larger influence of vegetation on smaller channels. Second, channel evolution

Region	Q	α	β
Pacific Maritime Mountains	1.2-yr	2.37	0.50
Western Cordillera	1.5-yr	3.50	0.44
Columbia Basin	1.4-yr	0.96	0.60

Table 1 Parameters used to relate bankfull width (BFW) to peak streamflow statistics. The parameter α is included even though it is not used in our analysis. *Source:* Castro and Jackson (2001).

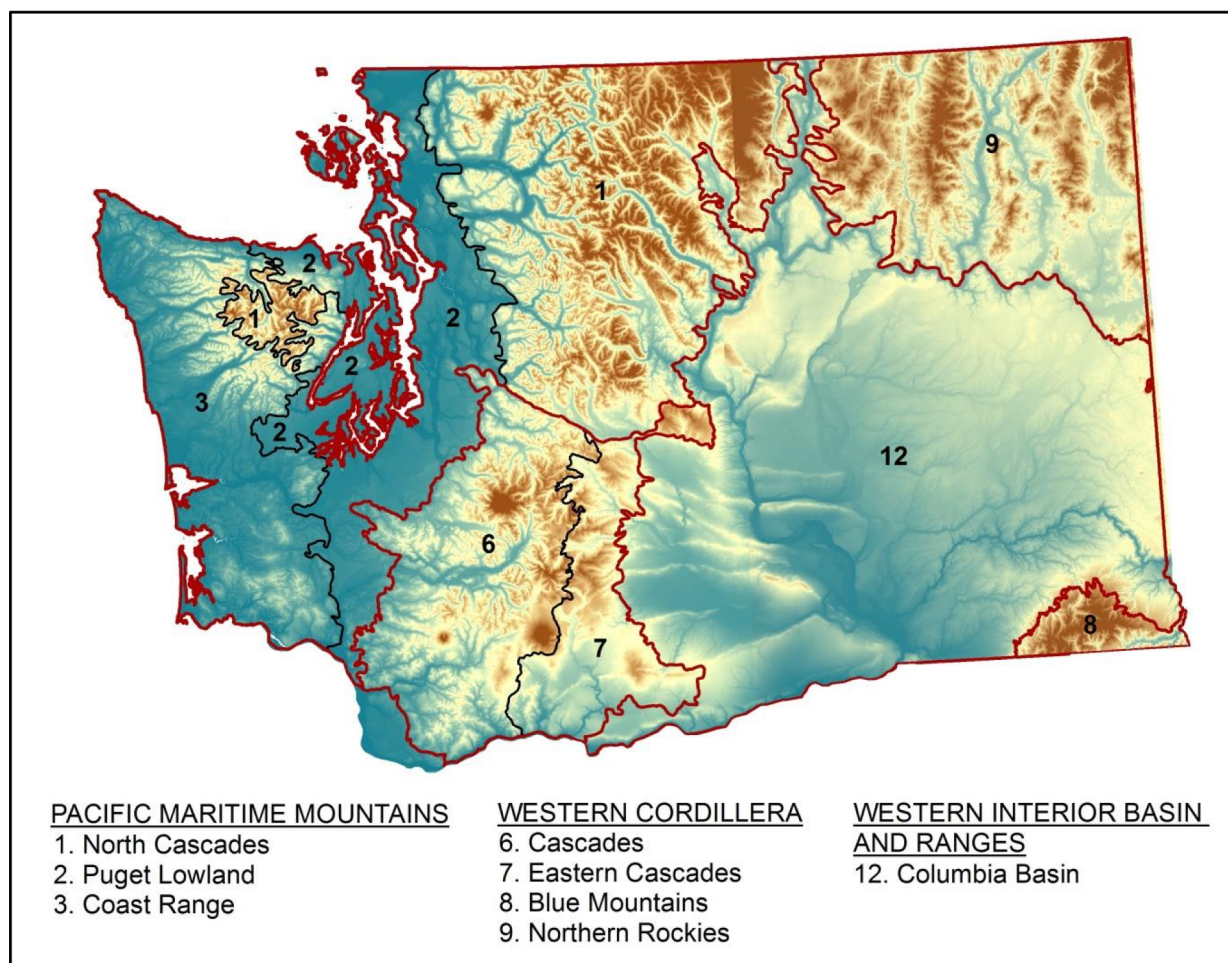


Figure 2. Ecoregions used to define the parameters relating bankfull width (BFW) to peak streamflow statistics across Washington State (Castro and Jackson 2001). *Source:* Copied, with permission, from Wilhere et al. 2016.

is complex and depends on many factors, including the channel slope, substrate, land cover, the influence of groundwater, and the history of human modification (e.g., Buffington 2012, Hession 2003, Booth 1990, Booth 2015). These issues highlight the need for future work to evaluate the Castro and Jackson (2001) approximation and determine if a better approach is needed.

The second challenge – dealing with the uncertainty in climate change projections – relates to three primary sources of uncertainty:

1. *Natural variability in climate.* Fluctuations in climate can systematically skew the estimates of BFW, particularly given its dependence on rare events, on time scales ranging from years to a few decades.
2. *Uncertainty in greenhouse gas scenarios.* Future emissions of greenhouse gases depend on a range of possible changes in human behavior, population growth, and technological innovations (see Section 1 of Mauger et al. 2015).
3. *Differences among model projections.* Global climate models, “downscaling” techniques, and hydrologic models can all differ in terms of the way they represent climate, weather, and hydrology. As a result there is a range among projections of future BFW.

Natural variability is typically addressed by using 30-year periods to calculate statistics, and this time frame is considered adequate for capturing extreme statistics near the mean annual flood, such as the 1.5-year flood.

Since likelihoods are not assigned to greenhouse gas scenarios and climate models project different 21st century changes, many studies have attempted to assess the range of outcomes from a large group of GCM projections and a smaller number of greenhouse gas scenarios (Hamlet et al. 2013). Users of the tool can then evaluate the range among projections to make a decision about how they want to manage risks. For example, more risk averse users may choose to more heavily weight high-impact scenarios, whereas more risk tolerant users may more strongly value consensus in the models, and/or make use of conservative estimates of impacts (e.g. use estimates of change which are exceeded in 90% of the scenarios).

Model uncertainty is generally assessed by considering projections from a range of models. It is important to note, however, that this is an approximation of the uncertainty, since GCMs are rarely independent and hydrologic datasets only rarely include multiple “downscaling” approaches and hydrologic models. In addition, GCMs are not the only component of model uncertainty. GCMs represent the climate at coarse spatial scales, and therefore need to be “downscaled” to estimate the corresponding changes at local scales. There are two general approaches to downscaling: (1) Statistical, in which empirical relationships between historical GCM simulations and surface weather observations are extrapolated forward in time, and (2) Dynamical, in which a GCM is used to drive a finer-scale regional climate model. Statistical approaches are inexpensive to implement and readily available, but are less reliable in topographically complex areas,

especially where few observations are available. Dynamical approaches capture important physical processes that are absent from statistical downscaling (Salathé et al. 2010), but are restricted to only a few GCM projections, and have not been as thoroughly tested. For example, although there are a few hydrologic datasets based on dynamically downscaled projections (e.g., Salathé et al. 2010), there are currently no ensemble hydrologic datasets derived from dynamical downscaling that encompass the full range of GCM projections.

The statistical downscaling approach that Wilhere et al. (2016) used results in a 92-year record that replicates the statistics of each time period (Hybrid Delta, from Hamlet et al. 2013; hereafter referred to as “HB2860”).¹ Projections are based on a single greenhouse gas scenario that represents a middle estimate of future emissions (A1B, Nakicenovic et al. 2000). Wilhere et al. (2016) evaluated changes in BFW using two 30-year periods centered on the 2040s and 2080s. They suggest using two metrics for evaluating changes in BFW: (1) percent change in BFW, and (2) number of models projecting an increase in BFW. As an example, they define “actionable” locations as those for which the mean increase in BFW is >5% and for which at least 6 of the 10 models show an increase.

As described in Section 3, we took a similar approach to Wilhere et al. (2016). Specifically, we used the same definition that culvert “failure” occurs when the Stream Simulation design standard is exceeded by a statistically significant amount. In practice, culverts can fail due to a wide variety of issues related to debris, scour, etc. Instead of attempting to characterize each failure mechanism, we take the same approach as Wilhere et al. (2016) and define failure based on the state’s regulatory standard for culvert width.

We also used the same dataset of streamflow projections as Wilhere et al. (2016): HB2860. There are three reasons we chose to use the HB2860 projections:

1. There are currently no publically available dynamically downscaled streamflow projections for Washington State that provide a sufficiently large ensemble of projections (e.g., comparable to HB2860),
2. Recent work has shown that newer statistically downscaled datasets (e.g., Mote et al. 2017) have significant biases (Mauger et al. 2016), and
3. Wilhere et al. (2016) used HB2860 and are familiar with this product. (i.e. by using the same dataset, our results are more consistent with their analysis.)

We note, however, that our approach differs somewhat from Wilhere et al. in that we used a different set of projections from the Hamlet et al. dataset. Instead of Hybrid Delta, we used the monthly Bias Correction and Spatial Downscaling (BCSD) projections from HB2860. This was done to allow a unique probability distribution to be constructed for each design year in the future. In addition, we developed an approach to estimating the probability that a culvert’s designed size (either proposed or “as built”) will be significantly

¹ “HB2860” refers to the Washington State legislative bill that funded the project. Results are available online at: <http://warm.atmos.washington.edu/2860/>

exceeded as a result of climate change over a given design lifetime. These new results are incorporated in a user-friendly tool that is designed to be easily integrated into the process of selecting an appropriate culvert design. Our projections should generally show comparable changes to those of Wilhere et al. (2016), but with a new perspective afforded by the probabilistic approach.

Task 1: Climate Data Validation

Surface weather observations – in particular long-term high-quality records – are typically sparse in spatial coverage, and tend to be biased towards low elevation areas near population centers. This means that the majority of point observations are generally sub-optimal for use in hydrologic modeling studies, in which the manifestation of large-scale weather patterns may be very different from one part of the watershed to another – in particular in areas with complex terrain. In addition, hydrologic models often employ a gridded spatial structure, meaning that daily weather conditions must be interpolated from each observation point onto a grid in order to run the model.

As with statistically- and dynamically-downscaled climate projections, gridded historical datasets can also be generated using these two distinct approaches. In the statistical approach, station observations are interpolated onto a grid and then adjusted to account for proximity and topography (e.g., Livneh et al. 2013, 2015). The dynamical approach, in contrast, uses large-scale observed weather conditions as boundary conditions for a finer-scale regional climate model simulation (e.g., Salathé et al. 2010).

The purpose of the current evaluation is to support the interpretation of our probabilistic projections (Section 3) in the Skagit river basin. Although the primary focus is on the HB2860 dataset, other historical datasets are included in order to better understand the relative performance of each.

Observational Data

In this task we provide a few spot comparisons between observed weather and a selection of currently-available gridded historical datasets. A key detail here is that the validation data must be independent of the data used to create the gridded datasets. This limits the analysis to only a few stations. In this study, we focus on three observational networks: (1) the NOAA Climate Reference Network (CRN, <https://www.ncdc.noaa.gov/crn/>), (2) the

Table 2. Weather stations sites used in the analysis. Sites are listed from lowest to highest elevation.

Name	ID	Lat.	Long.	Elev (m)	Source	Years
Fir Island	FirIsland	48.360	-122.420	0	AgNet	2008-2017
Mt. Vernon	MtVernon	48.440	-122.390	7	AgNet	1993-2017
Sakuma	Sakuma	48.497	-122.378	9	AgNet	2006-2017
Sedro Wooley	SWYW1	48.522	-122.224	52	RAWS	2013-2017
Darrington 221 NNE*	451998	48.540	-121.446	124	CRN	2004-2017
Finney Creek	FIFW1	48.392	-121.818	658	RAWS	2000-2017
Gold Hill	GHFW1	48.243	-121.546	1021	RAWS	2000-2017

*Although labeled “Darrington”, this station is actually located about 20 miles north of Darrington in Marblemount.

WSU Ag Weather Net, a network of weather stations aimed at monitoring weather in agricultural areas (<https://weather.wsu.edu/>), and (3) the National Interagency Fire Center's (NIFC) Remote Automated Weather Stations (RAWS, <https://raws.nifc.gov/>).

Weather stations were selected to cover a range of elevations, all in the vicinity of the Skagit watershed (Table 2). To ensure robust statistics, we required that at least 25 days of valid daily data be present for each month for the monthly statistics, and that all 12 months meet this criteria in order to record the water year statistics. This is similar to the approach taken by NOAA in processing the Cooperative Observer (COOP) station records. For some stations, these criteria significantly limited the number of valid months.

Model Data

Daily historical weather data were compared for five different datasets (Table 3), representing both the statistical and dynamical approaches described above. Additional information about each of these datasets can be found in the references provided in the table. All results were aggregated to daily data for the comparisons. Data were extracted from each dataset for the grid cell closest to each weather station.

Table 3. Historical meteorological datasets that were evaluated in the current study.

Short Name	Resolution	Time Step	Years	Type	Citation
HB2860	1/16-deg.*	1 day	1915-2006	Statistical	Hamlet et al. 2013
bcLivneh	1/16-deg.*	1 day	1950-2013	Statistical	Livneh et al. 2015, Mauger et al. 2016
Livneh	1/16-deg.*	1 day	1950-2013	Statistical	Livneh et al. 2015
WRF-NNRP	12 km	6 hours	1950-2010	Dynamical	Dulière et al. 2011, Salathé et al. 2010
WRF-PNNL	6 km	1 hour	1979-2015	Dynamical	Ruby Leung, personal communication.

* about 5 x 7 km.

Results

Overall, these data proved very limited for evaluating the gridded historical datasets, primarily due to a significant amount of missing or invalid data in the observations. After filtering for months and years that had a sufficient number of valid observations (i.e., a minimum of 25 days of valid data per month, with all 12 months meeting this criteria for water year statistics), very little data remained with which to compare against the gridded datasets. In part this is due to the length of the records: the longest observational record among all 7 stations is 25 years, and the shortest is just 5 years long. In addition, the gridded datasets end as early as 2006, which limits the number of overlapping years available for comparison. In part this limitation is by construction: in order to maximize the information available to the interpolation, the gridded datasets strive to incorporate as

much available data as can be obtained. This makes it challenging to identify independent observations that can be used to validate the gridded datasets.

In spite of these limitations, we present two comparisons showing the performance of the gridded datasets for the seven sites listed in Table 2. Given the large amount of missing data, the water year statistics are particularly sparse. As a result, Figure 3 shows a

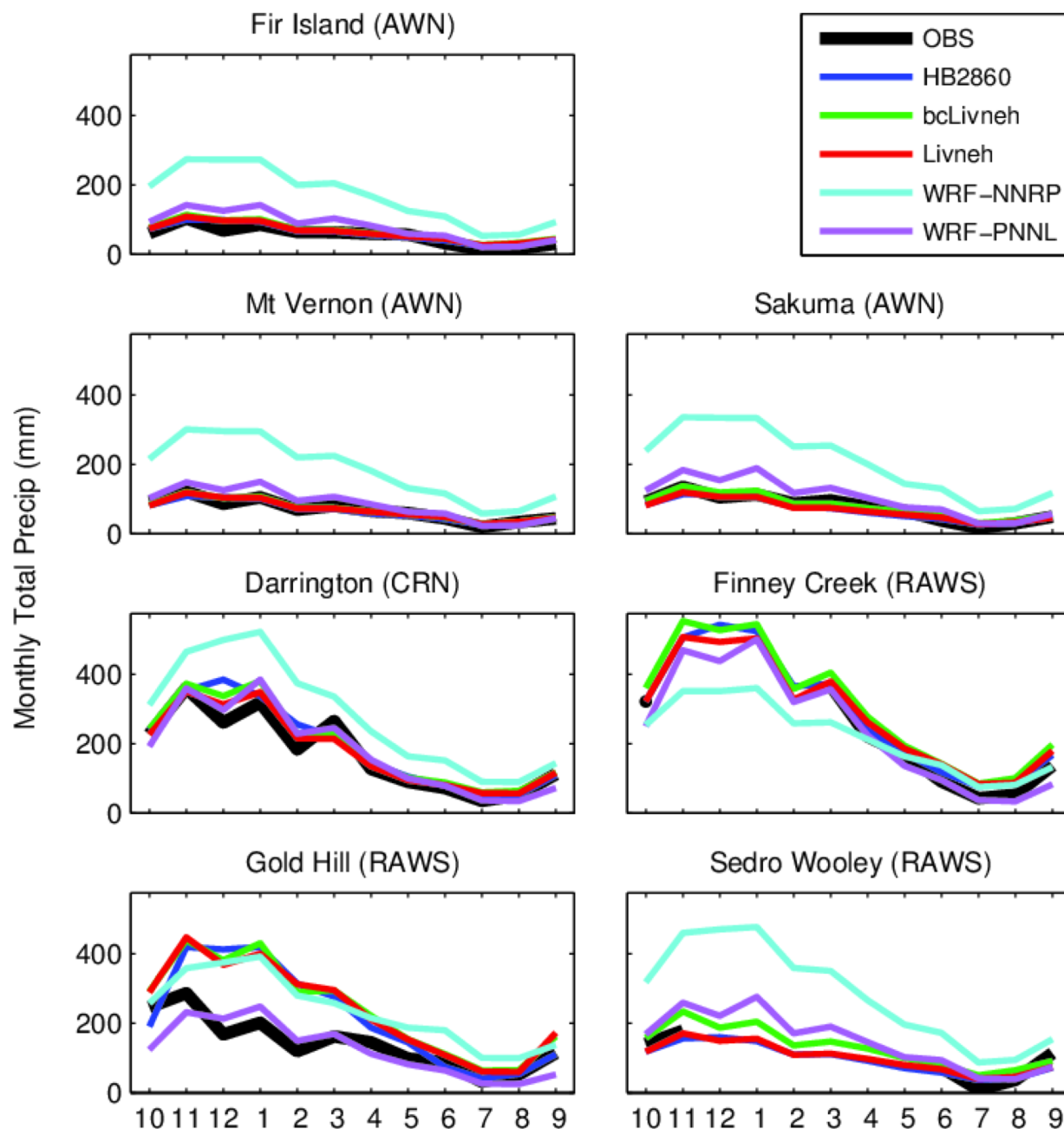


Figure 3. Comparing average monthly precipitation for the observations (thick black lines) and gridded historical datasets (colored lines). Due to missing data, some of the months are absent from the observations.

comparison of average monthly precipitation for each site. In this case, the averages include any years and months for which there is valid data, meaning that the time periods may not always overlap. In addition to showing a wide range of precipitation among the seven sites, the results show a wide range among gridded datasets. Most of the datasets agree well with the observations, with two exceptions: (1) the WRF-NNRP simulation generally over-predicts precipitation, and (2) only the WRF-PNNL appears to capture precipitation at the Gold Hill location. This latter finding is interesting, given the high elevation location of Gold Hill and the known challenges for the interpolated datasets in mountainous areas.

Although the monthly averages provide a valuable first look at the results, heavy precipitation events are of primary interest in the current analysis since changes in bankfull width are thought to be driven by high flow events. In addition, given our emphasis on the relative changes in streamflow, it makes more sense to consider the correlation of each dataset with the observations. We computed correlations for all time series that overlapped by at least five years. Although five years is not sufficient to capture the effects of long-term variability, this was chosen as a compromise given the short records and lack of overlap among the available data.

Table 4. Correlations between the observed maximum in daily precipitation for each water year and the corresponding values in each gridded dataset. Blank squares indicated that there were less than 5 years of overlapping data from which to compute a correlation.

site name	<i>Obs Start Year</i>	HB2860	bcLivneh	Livneh	WRF-NNRP	WRF-PNNL
<i>Model End Year</i>		2006	2013	2013	2015	2010
Fir Island (AWN)	2009	–	0.59	0.21	0.12	–
Mt Vernon (AWN)	1999	0.93	0.78	0.71	-0.09	0.02
Sakuma (AWN)	2008	–	0.80	0.60	0.04	–
Sedro Wooley (RAWS)	2015	–	–	–	–	–
Darrington (CRN)	2005	–	0.72	0.75	0.82	0.51
Finney Creek (RAWS)	2014	–	–	–	–	–
Gold Hill (RAWS)	2009	–	–	–	–	–

Table 5. As in Table 4 except showing correlations for the water year total in precipitation.

site name	Obs Start Year	HB2860	bcLivneh	Livneh	WRF-NNRP	WRF-PNNL
Model End Year		2006	2013	2013	2015	2010
Fir Island (AWN)	2009	–	0.92	0.92	0.61	–
Mt Vernon (AWN)	1999	0.98	0.78	0.45	0.49	0.69
Sakuma (AWN)	2008	–	0.11	0.80	0.52	–
Sedro Wooley (RAWS)	2015	–	–	–	–	–
Darrington (CRN)	2005	–	0.94	0.97	0.65	0.84
Finney Creek (RAWS)	2014	–	–	–	–	–
Gold Hill (RAWS)	2009	–	–	–	–	–

Table 4 shows the correlations for the water year maximum in daily precipitation. For comparison, Table 5 shows the correlations for total water year precipitation. As can be seen from the tables, there are many comparisons for which there were less than the minimum five years needed to estimate a correlation. Unfortunately, this is particularly limiting for HB2860: the primary dataset used in the present study. Although limited, the results do provide two potential insights. First, the correlations for total annual precipitation are much better than those for extreme precipitation. This is not surprising, given that the entire region tends to vary in concert at seasonal and longer time steps. However it is worth noting since the differences are so dramatic – in one case even going from a negative to a positive correlation. Second, the dynamical datasets appear to underperform at capturing the extremes in precipitation. This is surprising, given that the anticipated benefit of the dynamical approach is that it better captures the conditions away from existing weather stations, particularly for the extremes. One explanation for this could be that the only stations that we have been able to evaluate are either at relatively low elevations or near existing observations that are already ingested in the interpolated datasets (e.g., the Darrington CRN station is likely co-located with the existing COOP station that HB2860 and Livneh use in their interpolations).

Overall there is insufficient data from which to draw a conclusion about the performance of these datasets. The sites included here are too few and cover too small an area to reliably draw conclusions about each of the datasets. Nonetheless, the comparisons do indicate general agreement with the observations while also showing notable differences in performance. Specifically, this analysis suggests that monthly total and water year maximum precipitation are both well-captured by the statistically-generated gridded

datasets, including HB2860. Additional work would be needed to confirm that this pattern is present elsewhere in the state and at higher elevations.

Task 2: Streamflow Validation

Few validations have evaluated hydrologic simulations over a number of small catchments as opposed to large river sites. In this section we evaluate the results of the hydrologic simulations for a number of small streams with long-term gauges. As with the climate data validation, this evaluation is simply intended to support the interpretation of our probabilistic projections (Section 3) in the Skagit river basin. The primary focus is on the HB2860 dataset, which is the basis for our projections. Other historical datasets are included in order to compare the performance of each approach.

Observational Data

There are no streamflow estimates that are specific to any particular culvert in the Skagit basin. However, several streamflow gauges exist for small creeks in the area (Table 6). Sites were chosen to encompass a range of elevations and emphasize smaller catchments. Although most are larger than a typical culvert catchment, these were the smallest streams with available observations in the vicinity of the Skagit basin. For example, the catchment area for Thunder Creek is 105 square miles; much larger than a typical culvert stream, which tend to range from about 0.5 to 2 square miles. We nonetheless evaluate results for this site because it is small compared to mainstem Skagit sites, and there are no other long-running gauges at similar elevations in the area.

Table 6. Streamflow sites used in the analysis. As in Table 2, sites are listed from lowest to highest elevation.

Name	ID	Lat.	Long.	Elv. (ft)	Area (sq. mi)	Source	Years
Fishtrap Cr at Front St at Lynden	12212050	48.939	-122.478	54	37.8	USGS	1998-2018
Hansen Cr. Near Sedro Wooley	03J100	48.531	-122.201	89	7.02	ECY	2005-2018
Anderson Cr at Smith Rd Nr Goshen	12210900	48.833	-122.338	200	8.96	USGS	1998-2018
Anderson Cr Nr. Bellingham	12201950	48.674	-122.266	315	4.13	USGS	1967-2018
Olsen Creek Nr. Bellingham	12202300	48.751	-122.352	315	3.78	USGS	1967-2018
Euclid Cr at Euclid Ave at Bellingham	12202400	48.749	-122.408	320	0.54	USGS	2001-2018
Carpenter Cr Nr Bellingham	12202310	48.754	-122.353	320	1.17	USGS	2002-2018
Silver Beach Cr at Maynard Pl at Bellingham	12202450	48.769	-122.405	320	1.20	USGS	2001-2018
Brannian Cr at S. Bay Dr Nr Wickersham	12201960	48.669	-122.279	330	3.36	USGS	2001-2018
Skookum Cr Abv Diversion Nr Wickersham	12209490	48.672	-122.138	410	23.0	USGS	2008-2018
Thunder Cr Nr Newhalem	12175500	48.673	-121.072	1,220	105	USGS	1930-2018

Table 7. Historical streamflow datasets that are evaluated in the current study.

Short Name	Years	Type	Citation
HB2860	1915-2006	Statistical	Hamlet et al. 2013
bcLivneh	1950-2013	Statistical	Livneh et al. 2015, Mauger et al. 2016
WRF-NNRP	1950-2010	Dynamical	Dulière et al. 2011, Salathé et al. 2010
Livneh	1950-2013	Statistical	Livneh et al. 2015
CRCC-PRMS-1	1950-2011	Statistical	Chegwidden et al. 2018, Livneh et al. 2013
CRCC-VIC-1	1950-2011	Statistical	Chegwidden et al. 2018, Livneh et al. 2013
CRCC-VIC-2	1950-2011	Statistical	Chegwidden et al. 2018, Livneh et al. 2013
CRCC-VIC-3	1950-2011	Statistical	Chegwidden et al. 2018, Livneh et al. 2013

Model Data

Although not all of the climate datasets evaluated in Section 1 have been used to produce hydrologic simulations, several do include streamflow estimates that can be used in the current validation. In several instances the same meteorological dataset has been applied to different combinations of hydrologic models and calibration approaches, resulting in different streamflow estimates for the same dataset.

Table 7 lists the datasets used in our comparisons. All are produced at a resolution of 1/16-degree (about 5x7 km). As in Table 3, the “Type” column refers to the approach: “statistical” datasets are created by interpolating from surface weather station observations, while “dynamical” datasets require a regional climate model simulation, driven by observations, to develop historical meteorological estimates. In all cases, the meteorological estimates are then run through a hydrologic model to obtain streamflow estimates. Streamflow was estimated by computing an area-weighted sum over all grid cells that overlap with the contributing basin for each streamflow gauge. Larger rivers often require an extra post-processing step in which runoff estimates are “routed” through the stream network based on an assumed time distribution for flow (e.g. Hamlet et al. 2013). In this study, routing is not necessary since the basins are small enough that the time required for water to travel to the outlet of each catchment is less than the model’s time step of one day.

Although additional streamflow estimates could presumably be obtained from recent fine-scale hydrologic modeling using the Distributed Hydrology Soil Vegetation Model (DHSVM; Bandaragoda et al. 2015), doing so would require modifying the DHSVM model to include additional streamflow locations at all culvert sites of interest, then re-running the simulations. This was not deemed feasible for the current study.

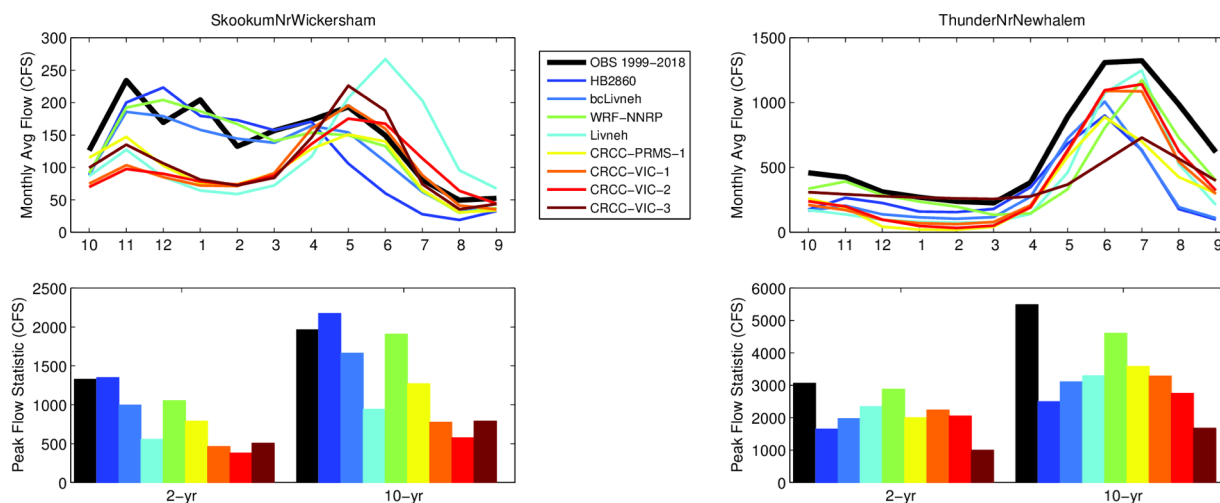


Figure 4. Example results for two of the 11 streamflow sites included in the analysis. In each panel, the top plot shows a comparison of monthly average flows for each dataset, while the bottom panel shows results for the 2- and 10-year flood.

Results

In order to evaluate the implications for culverts, the analysis is focused on peak flows. Peak flow statistics are computed as in previous studies (e.g., Tohver et al. 2014). Specifically, a “block maximum” approach is taken, in which the maximum daily flow value is extracted for each water year (Oct-Sep). These peak flow values are used to evaluate the correlations with each observationally-based dataset (Table 7). Flow estimates for specific recurrence intervals (e.g., 2-year flood) are calculated by fitting the peak daily flows to a Generalized Extreme Value (GEV) distribution, then extracting the appropriate quantile from the fitted distribution (e.g., the 2-year flood corresponds to the 50% chance of exceedance while the 10-year corresponds to the 10% chance event).

All comparisons are made available on the project website; here we provide a few examples. Figure 4 shows the results for the Skookum and Thunder Creek sites. As is expected, the datasets that include the Livneh cold bias (discussed below) tend to overestimate the importance of snow, resulting in a corresponding underestimate of winter runoff. This distinction is not detectable for the Thunder Creek site, reflecting the greater sensitivity of the Skookum Creek simulations given its closer proximity to the snowline.

Correlations, taken from the overlapping period between the observations and each model dataset, reflect a similar heterogeneity in model skill (Table 8). The table shows the correlation, for each streamflow site, with each of the model datasets. Correlations are taken for the overlapping period between the observations at each gauge site and each of the modeled streamflow datasets. Although for Thunder Creek the correlations are based on more than 60 years of data, the sample size is generally confined to 10 years or less for the other sites.

There is substantial variability in the correlations for different sites, due to both the shorter records and the fact that model performance may differ from one site to the next (e.g., Figure 4). However, correlations for the bcLivneh and WRF-NNRP datasets are generally among the highest for the eight datasets evaluated.

Correlations emphasize the degree to which the models track the variability in the observations. This is the primary measure of interest in the current study, since the emphasis is on percent changes in bankfull flows. Nonetheless, it is important to also evaluate the absolute biases that are not captured in a correlation. Figure 5 shows the mean percent bias for the 2-year flood, for all sites and all datasets. As with the

Table 8. Correlations between the observed peak in daily flows for each water year and the corresponding estimates from each gridded dataset.

site name	Obs Start Year	HB2860	bcLivneh	WRF-NNRP	Livneh	CRCC-PRMS-1	CRCC-VIC-1	CRCC-VIC-2	CRCC-VIC-3
Model End Year		2006	2013	2013	2010	2011	2011	2011	2011
Fishtrap Cr at Front St at Lynden	1999	0.60	0.64	0.54	0.34	0.65	0.50	0.63	0.49
Hansen Cr. Near Sedro Wooley	2005	–	0.08	-0.06	-0.18	-0.13	0.01	-0.02	-0.17
Anderson Cr at Smith Rd Nr Goshen	1999	0.87	0.84	0.65	0.11	0.69	0.44	0.68	0.72
Anderson Cr Nr. Bellingham	2008	–	0.84	0.73	0.96	0.58	0.69	0.65	0.46
Olsen Creek Nr. Bellingham	2002	0.84	0.53	0.33	0.16	0.27	0.05	0.23	0.34
Euclid Cr at Euclid Ave at Bellingham	2002	0.19	0.65	0.48	0.66	0.47	0.39	0.53	0.32
Carpenter Cr Nr Bellingham	2002	0.92	0.51	0.37	0.14	0.32	0.30	0.29	0.18
Silver Beach Cr at Maynard Pl at Bellingham	2002	0.22	0.54	0.52	0.35	0.58	0.42	0.60	0.39
Brannian Cr at S. Bay Dr Nr Wickersham	2002	0.88	0.91	0.85	0.12	0.81	0.83	0.82	0.76
Skookum Cr Abv Diversion Nr Wickersham	1999	0.41	0.63	0.27	0.01	0.41	0.42	0.46	0.49
Thunder Cr Nr Newhalem	1951*	0.62	0.59	0.26	0.46	0.58	0.34	0.36	0.29

* Although the observational record for Thunder Creek starts in 1931, the correlation was assessed starting in 1951 to allow for greater consistency across gridded datasets.

correlations, there are considerable differences among datasets. In general, the HB2860 dataset performs as well, or better, than the other datasets evaluated.

It is worth noting the particularly poor results for Hanson Creek. One potential issue is the particularly short record, given that observations did not start until 2005. Although it should not affect the results, this is the only gauge site retrieved for the WA Department of Ecology network. Additional investigation would be needed to better diagnose the reason behind the biases for this site.

Overall, this analysis suggests that the HB2860 dataset performance is comparable to that of other readily-available historical streamflow datasets. Although some sites do show large biases, most show general agreement in both the magnitude and variation in monthly average and extreme flows. As in the previous section, it is worth noting that the scope of this analysis is relatively limited, and a more thorough assessment would be needed to determine if the same results hold for the region as a whole.

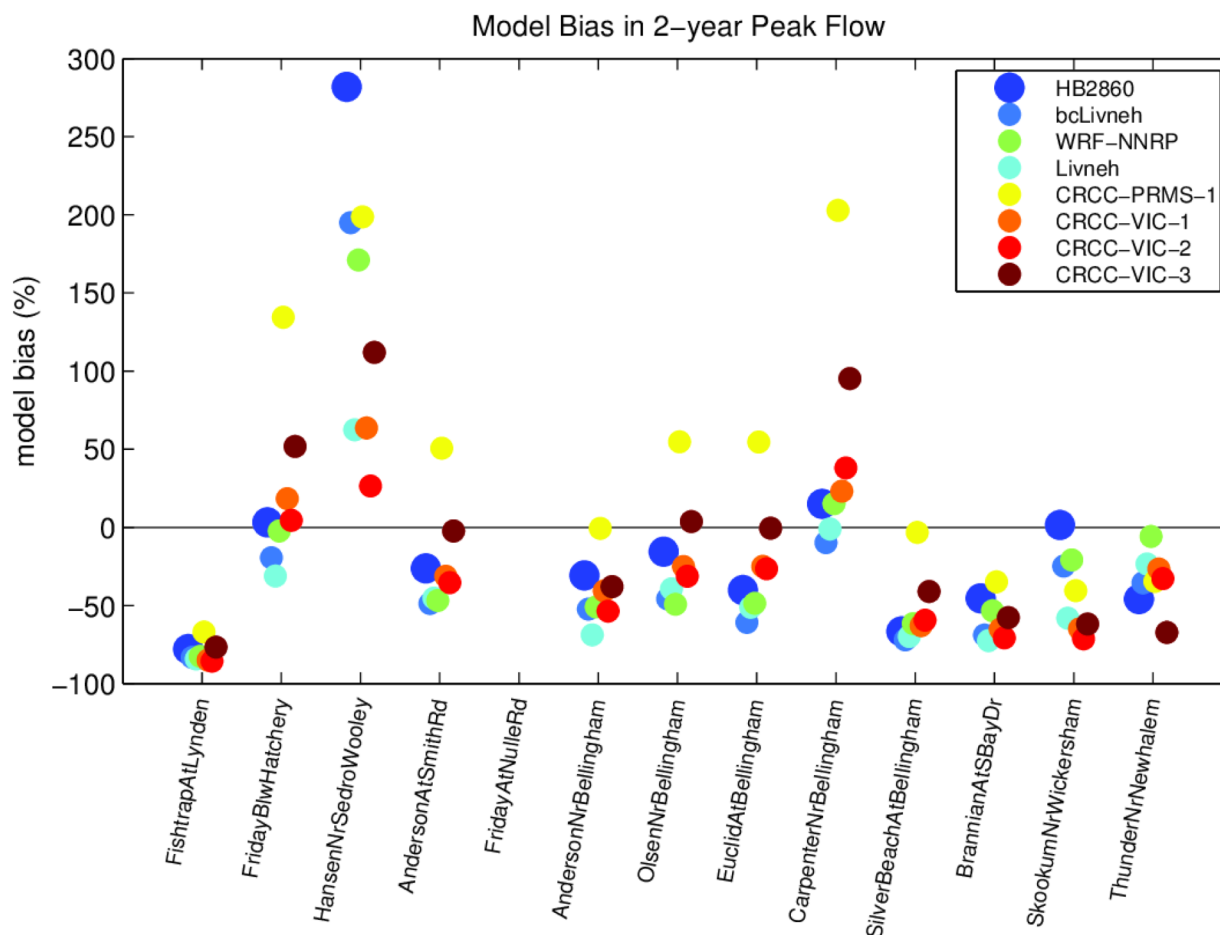


Figure 5. Percent bias, relative to observations, in the historical 2-year flood. Results for the HB2860 dataset, used in Section 3 of this report, are highlighted with the large blue circle. Sites are ordered from lowest elevation on the left to highest elevation on the right.

Task 3: Estimating the likelihood of Culvert “Failure”

In this study we adopt the same Castro and Jackson (2001) approach to estimating changes in BFW, and follow the Wilhere et al. (2016) definition that culvert “failure” occurs when the Stream Simulation design standard is exceeded. We also use the same dataset of streamflow projections.

As in previous studies (e.g., Tohver et al. 2014, Wilhere et al. 2016), our analysis is focused on the projected change in BFW as opposed to its absolute magnitude, or scale. By only considering the relative change – the ratio of future BFW to historical BFW – our approach is not sensitive to absolute biases in the models.

Our approach differs from Wilhere et al. (2016) in the treatment of uncertainty. Building on the work of Byun and Hamlet (in prep), we take a new approach that we believe is more robust to uncertainty, by estimating the uncertainty in extreme statistics and evaluating changes over time instead of for a few discrete decades. In addition, we estimate the probability that a culvert’s designed size (either proposed or as built) will be exceeded as a result of climate change over a given design lifetime. This “probability of under design” represents the likelihood that a culvert’s design parameters will be significantly exceeded when calculated using projected conditions in the future. All of the results are integrated into a user-friendly tool that is designed to be easily integrated in culvert design.

Data

For consistency with the Wilhere et al. (2016) study, we use results from the same dataset (HB2860, Hamlet et al. 2013). However, since our analysis is predicated on evaluating the time evolution of changes, we use the transient projections from that dataset, produced using the monthly Bias Correction and Spatial Downscaling (BCSD) approach. This approach results in a continuous time series of daily temperature and precipitation spanning from 1950-2099 for each global model and greenhouse gas scenario. As with all projections in this dataset, the BCSD are based on global climate model simulations archived as part of the Coupled Model Intercomparison Project, Phase 3 (CMIP3, Meehl et al. 2007) multi-model database. In the current project we evaluated projections for six global models (CCSM3, CGCM3.1-t47, CNRM-CM3, ECHAM5, ECHO-G, and PCM1) and one greenhouse gas scenario (the moderate “A1B” scenario, Nakicenovic et al. 2000; for more on climate scenarios see section 1 of Mauger et al. 2015). Hydrologic projections were produced using the Variable Infiltration Capacity (VIC) macroscale model (Liang et al. 1996, 1998; specifics on the model configuration and performance can be found in Hamlet et al. 2013). The result is a set of gridded estimates of daily future hydrology, including both surface and subsurface runoff, at a spatial resolution of 1/16-degree (about 5x7 km).

It is important to note that newer datasets are available. For example, the recent Integrated Scenarios (Mote et al. 2016) and Columbia River Climate Change

(Chegwidden et al. 2018) datasets include a number of advances in statistical downscaling and hydrologic modeling approaches, and are also based on the new CMIP5 (Taylor et al. 2012) global models and greenhouse gas scenarios. However, recent evaluations of these datasets (e.g., Mauger et al. 2016) have shown substantial temperature biases in the Livneh et al. (2013) dataset that these use as a basis. Similarly, recent research has shown that dynamical downscaling is needed to accurately capture changes in streamflow extremes (Salathé et al. 2014). Although such projections may prove to be superior in the long run, at present these have not been produced for an adequate number of large-scale forcing scenarios to adequately characterize the uncertainty in projections, nor have they been used to produce new calibrated hydrologic model projections for the region. Based on these considerations and the desire to maintain consistency with Wilhere et al. (2016), we used the HB2860 dataset in the current study.

Methods

Calculating Percent Changes in Future Bankfull Width (BFW) and Associated Culvert Design Widths

The methods used in this section are an extension of Monte Carlo flood analysis approach developed by Byun and Hamlet (in prep.). We use the daily streamflow estimates from the HB2860 dataset to estimate annual variations in bankfull width for each 1/16-degree grid cell in Washington State. We take a number of steps to ensure that our estimates are robust to uncertainties in extreme statistics and the confounding effect of natural variability on estimates of long-term trends. Specifically, we processed the data as follows:

1. *Calculate BFW statistics using a moving 30-year window.*

The window is left-adjusted, so that the statistics for any given year are calculated from the model results for the previous 30 years (e.g., for 2027, statistics were calculated based on the 30 years from 1998-2027). We chose to consider only prior years given the anticipated adjustment time for BFW.

2. *Bootstrap the probability distribution BFW.*

As described above, BFW is approximated as an exponential function of the 1.2, 1.4, or 1.5-year peak flow (Table 3.1). In order to ensure a robust estimate, we take a Monte Carlo approach, repeating the peak flow estimate by randomly selecting 20-year subsamples, with replacement, for each 30-year sample. For each sub-sample we recalculate the corresponding extreme statistic, repeating this process 1,000 times to develop a probability distribution for the BFW flow statistic.

3. *Calculate the change relative to a common baseline.*

In order to minimize the effect of model biases, we consider only the relative change in BFW. This also allows us to scale the results to different catchment sizes of interest, since the basin area for each culvert will be different and will rarely correspond to the area of a 1/16-degree grid cell. In addition, this ensures that our approach is not sensitive to absolute model biases.

To do this, we divide all estimates of BFW by the estimated BFW in the first year in the analysis. In order to provide a contemporary baseline, we evaluate BFW for 2015-2099, dividing all future values by the BFW estimate for 2015 (estimated from the years 1986-2015). Given the bootstrap approach in Step #2, the result is a set of 1,000 BFW ratios for each year.

4. *Extract the 5th percentile estimate.*

All subsequent analyses are based on the 5th percentile of the 1,000 BFW ratios obtained in Step #3. We use the 5th percentile to correspond to a 95% confidence interval. That is, we can say with 95% confidence that the “true” ratio is at or above that 5th percentile value, based on the uncertainty in the sample statistics established in the Monte Carlo simulations. Note that in this case we are assuming that a one-sided significance test is most appropriate, based on the reasoning that decreases in bankfull flows are not of interest.

(The choice of a 95% confidence level was deemed appropriate for diagnosing a statistically significant change in BFW and culvert design size. This choice is arbitrary and could be changed based on a policy decision about the level of change in BFW that is meaningful.)

5. *Identify statistically significant changes in BFW and calculate future culvert widths*

We compared the 5th percentile value of the BFW ratio derived above (expressed as a % change in BFW) to a threshold of 0 (no change in BFW). If at any time during the design lifespan the 5th percentile values for a given year exceeds a percent change of 0, this is classified as a statistically significant change in BFW for that year and GCM projection (see comparison in Figure 6).

In addition, based on the 5th percentile percent change in BFW, the current BFW (provided by the user), and the WDFW culvert design standard (eq 1), the projections can be translated from percent changes in BFW (i.e., the 5th percentile ratios from the Monte Carlo analysis) to culvert design width.

6. *Specify design value and identify failure/no failure for each simulation year*

We begin by specifying a new threshold for the culvert width associated with a proposed culvert design (provided by the user), and then check to see if the 5th percentile future values exceed this value. If the 5th percentile value of simulated future culvert width (as outlined in the last part of 5 above) exceeds the specified culvert design threshold, then the proposed culvert design results in a “failure” for that year, GCM projection, and proposed culvert design width.

7. Repeat steps 1-6 for all six global climate models.

The first six steps were applied to the six BCSD projections that are currently available (CCSM3, CGCM3.1-t47, CNRM-CM3, ECHAM5, ECHO-G, and PCM1), all based on the moderate A1B greenhouse gas scenario. The result is six separate time series of bankfull width for each grid cell, with each time series starting in 2015 and ending in 2099. These can be used to assess the likelihood of failure for a particular change in culvert size, as described below.

We tested a number of alternatives to this methodology, and found the approach above to be most robust. Specifically, we found that a bootstrap approach based on 1,000 sub-samples was sufficient to characterize the probability distribution in bankfull flows. Similarly, we tested using smaller sub-samples of 10 and 15 years instead of the 20 year subsamples chosen. Although both the 15-year and 20-year subsamples gave similar results, we found a marked increase in the range of estimates when only 10-years were used to calculate the extreme statistics. We interpret this to be a result of under-sampling and chose to use the 20-year sample size in the current analysis.

We also evaluated the approach to estimating the 5th percentile in the ratio of historical to future bankfull flows. We first tested if the bootstrap approach could be applied to just the first 30-year sample (1986-2015), while only using a single 30-year sample to estimate bankfull flows for every subsequent year. This approach results in a much narrower probability distribution of bankfull flows, and a resulting overestimate of the change in the 5th percentile estimate (i.e., because the distribution is narrower, the 5th percentile value is higher). Since this approach neglects the uncertainty in future bankfull flow estimates, we interpreted this to mean that the latter approach under-samples the probability distribution for changes in bankfull flows. Similarly, we tested whether or not the 5th percentile values could be extracted from each distribution before calculating the ratio relative to the first year in the record. In this case we also found that the ratio of the 5th percentiles resulted in an overestimate relative to first calculating the ratios before extracting the 5th percentile in the ratios of bankfull flow. This is again a consequence of narrowing the distribution, which results in a higher 5th percentile estimate. As above, this is likely a consequence of under-sampling and so was deemed less accurate than our current approach of taking ratios for each of the 1,000 random sub-samples for each future year, relative to the same for historical.

Finally, as noted above we chose to use the HB2860 dataset (Hamlet et al. 2013) in the present analysis, since the newer statistically downscaled datasets have substantial temperature biases and there are currently no dynamically downscaled hydrologic projections that can be used for these purposes. Nonetheless, here we include a comparison between an adjusted version of the Integrated Scenarios dataset (Mote et al. 2016), in which a simple correction has been applied to minimize the effect of the temperature biases (hereafter referred to as “bcMACA”; for additional detail see Mauger et al. 2016). Results are shown for two model grid cells: one near Mt Vernon (48.46875N, 122.40625W) and another near Marblemount (48.53125N, 121.46875W). The comparison (Figure 6) shows that the results for both datasets are in general agreement. Although the scope of the current project does not allow for a more in-depth comparison, this suggests that the results of the current study are likely representative of what might be found from other statistically downscaled projections.

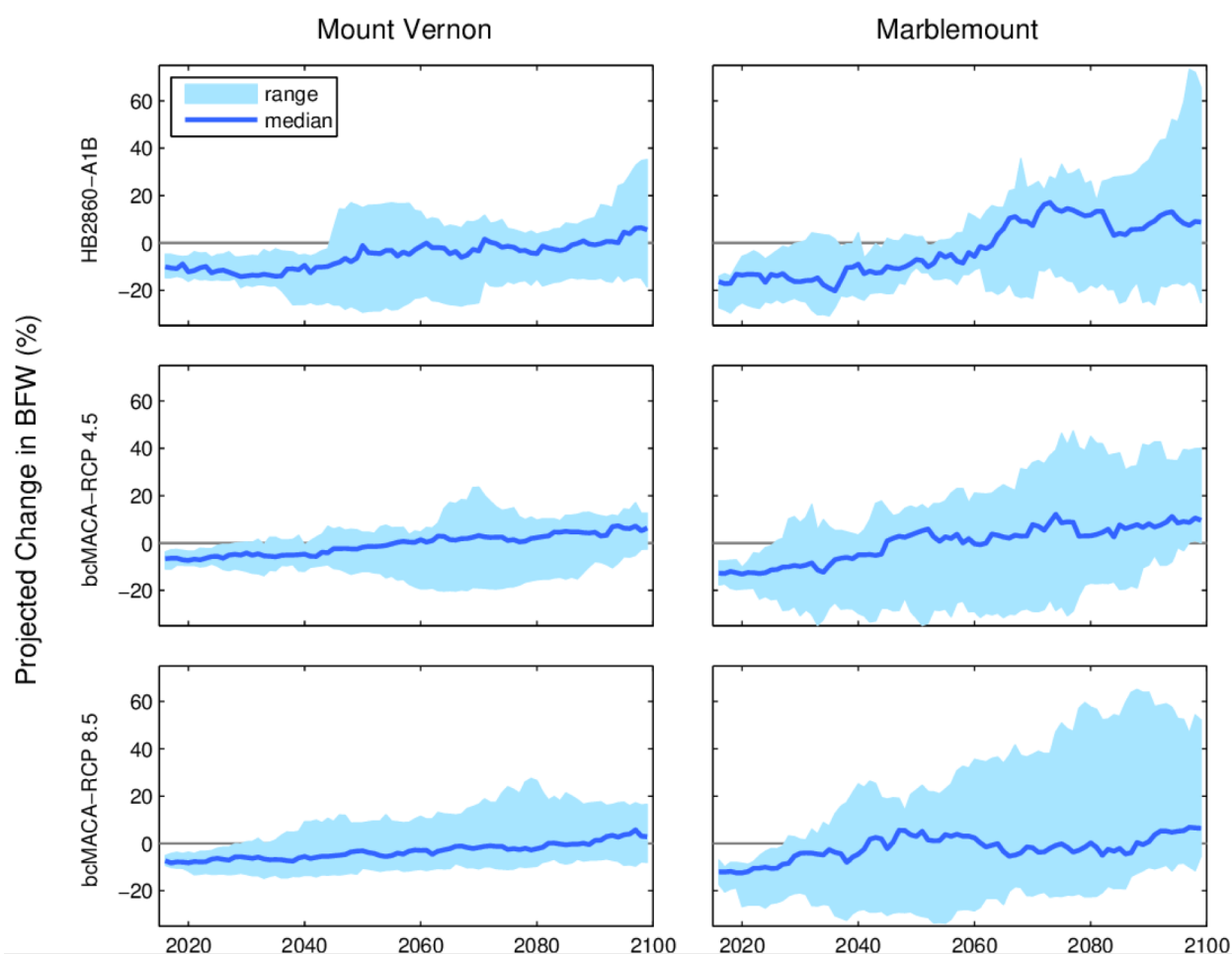


Figure 6. Comparing bankfull width projections, following the methodology described in this report, for the HB2860 dataset (A1B scenario; Hamlet et al. 2013) with those from the bcMACA dataset (RCP 4.5 and 8.5 scenarios; Mote et al. 2016, Mauger et al. 2016). The current study uses results from HB2860; as shown in the figure, these are in qualitative agreement with those from the bcMACA dataset.

Estimating the Likelihood of Failure

The above approach results in a time series of the projected change in the 5th percentile estimate of BFW (expressed as the corresponding culvert width) for each year and each of the six projections. We can then ask if, and when, the estimated culvert width exceeds some particular threshold (as outlined in step 6 above). For example, an engineer may propose to build a culvert 10% wider than is currently required under the Stream Sim guidelines. If the 5th percentile values of BFW ratio established in the Monte simulations (again expressed as culvert width, as outlined in step 5 above) exceed the proposed culvert design width, that year is identified as a failure for the proposed design. The result is a simple binary measure of failure or non-failure for each year and for each climate change projection.

We then combine the results for all projections by taking the fraction of projections that showed a culvert failure in a particular year, using this to define the probability that it has failed. For example, if two out of six models show a failure in 2037, then the probability of failure for that year is 2/6 or 33.3%.

Once we have an estimate of the probability of failure for each year, we can combine it over the relevant years to estimate the total probability of one or more failures over the design lifetime of the culvert. This is done by estimating the likelihood of non-failure, since doing so greatly simplifies the calculation. The lifetime probability is obtained by taking the product of the probability of non-failure for each year in the design lifetime. The probability of failure can then be estimated by simply subtracting the probability of non-failure from one:

$$P_{lifetime} = 1 - \prod_{yr=1}^N (1 - p_{yr}) \quad \text{Eq. 3}$$

where $P_{lifetime}$ is the probability of failure over the design lifetime, and p_{yr} is the probability of failure for one particular year (see Byun and Hamlet in prep. for additional details.).

It is important to note that this approach to estimating likelihoods is limited in two ways. First, different likelihood estimates must be estimated separately for each greenhouse gas scenario. This is because likelihoods cannot be assigned to future emissions of greenhouse gases. In the current study, we focus on a moderate scenario (A1B). Different results would be obtained for a higher or lower scenario of future emissions. Second, the six global climate model projections only provide an approximation of the probability distribution of future climatic conditions. In particular, the minimum and maximum changes projected by the six models are unlikely to represent the full range of possible future conditions. In the current project we make the simplifying assumption that zero out of six models means there is no probability of failure, while six out of six corresponds to a 100% probability. In reality, each probably corresponds to a more moderate definition of the extremes of the distribution (e.g., in IPCC 2013, the spread among projections is

often assumed to represent the 5th and 95th percentiles). Future work could add results for additional greenhouse gas scenarios and develop a more nuanced approach to estimating the probability distribution among different climate model projections.

Results

The primary purpose of this work was to produce a tool that could be used by engineers and planners to evaluate the implications of climate change for culvert design (Figure 1). To do this we calculated changes in bankfull width (BFW) for each 1/16-degree grid cell in Washington State, following the methodology described above. Our approach is intentionally designed to estimate the ratio of future to historical BFW, since catchment size, and the corresponding BFW, is different for each culvert. Since the WDFW “Stream Sim” standard requires that culverts be sized relative to the observed BFW, we ask the user to enter this value and use it to scale the ratios obtained from the projections. Specifically, the 5th percentile of the BFW ratio projections is multiplied by the observed BFW, and this value is then converted to a culvert dimension for each year using the WDFW regulatory standard (Eq. 1).

The likelihood calculation depends on both the design lifetime of the culvert and its proposed size. As a result, users are also prompted to enter the design lifetime for their project and a proposed size for the new culvert. The likelihood of “failure” is then calculated by evaluating the probability that changes in BFW (after scaling, as described above) exceed the proposed culvert size over its design lifetime.

The tool (Figure 1) and can be accessed via the CIG website (<https://cig.uw.edu/our-work/decision-support/culvert-phase-2/>). In the first screen, users select the grid cell that encompasses the area of interest. This then takes the user to the second screen, where she can enter the observed bankfull width and design lifetime, then try out different proposed culvert sizes to evaluate the likelihood of failure. On this screen, the total probability of failure, for a particular size and design lifetime, is shown in the orange box in the lower-left corner of the screen. The upper plot shows the time evolution in the mean and range of culvert size projections, scaled to match the current bankfull width measurement that was entered by the user. The bottom plot shows the probability of failure for each year, given a particular design lifetime and proposed culvert size. Figure 1 shows screen shots of both the first and second screen, using results for the grid cell closest to Stehekin, WA as an example.

Conclusions

This report describes a new tool that is designed to support climate-robust culvert design. The tool incorporates a new approach that estimates the probability of culvert failure over a given design lifetime, defining “failure” to mean that a culvert no longer meets the state’s design criteria. Covering all of Washington State, users are prompted to select a location of interest, enter their measurement of BFW and desired design lifetime, then test out different culvert sizes and obtain the likelihood of failure for each.

In addition to the tool, the report describes an evaluation of the meteorological and streamflow datasets that are the basis for the tool, with the goal of supporting the use and interpretation of the results. We compared available historical climate and streamflow datasets against independent observations in the vicinity of the Skagit basin. Although we only evaluated a few sites, the comparisons show that the primary dataset used in this study performs as well as or better than other currently-available datasets.

Our approach minimizes the effect of model biases by focusing on the relative changes in projected streamflow, evaluated in terms of the ratio of future to historical BFW. The approach is also an improvement over past studies in that it provides a specific estimate of the probability of failure over the design lifespan, integrating the range of model projections into a single number summarizing the implications for a particular culvert size and location. The tool is flexible and designed to facilitate quick evaluations in support of climate-resilient culvert design.

REFERENCES

- 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. <http://onlinelibrary.wiley.com/doi/10.1002/joc.2312/full>
- Booth, D. B. (1990). Stream - channel Incision Following drainage - basin Urbanization. *JAWRA Journal of the American Water Resources Association*, 26(3), 407-417.
- Booth, D. B., & Fischenich, C. J. (2015). A channel evolution model to guide sustainable urban stream restoration. *Area*, 47(4), 408-421.
- Buffington, J. M. (2012). Changes in channel morphology over human time scales. *Gravel-Bed Rivers: Processes, Tools, Environments*, 433-463.
- (Bulletin 17-B) U.S. Interagency Advisory Committee on Water Data, 1982, Guidelines for determining flood flow frequency, Bulletin 17-B of the Hydrology Subcommittee: Reston, Virginia, U.S. Geological Survey, Office of Water Data Coordination, [183 p.]. [Available from National Technical Information Service, Springfield VA 22161 as report no. PB 86 157 278 or from FEMA at http://www.fema.gov/mit/tsd/dl_flow.htm
- Byun, K., A.F. Hamlet, Monte Carlo techniques for quantifying flood risks and establishing infrastructure design standards in a non-stationary environment, (*in preparation*).
- Castro, J. M., & Jackson, P. L. (2001). *Bankfull discharge recurrence intervals and regional hydraulic geometry relationships: Patterns in the Pacific Northwest, USA*. *JAWRA Journal of the American Water Resources Association*, 37(5), 1249-1262. <https://doi.org/10.1111/j.1752-1688.2001.tb03636.x>
- Hamlet, A. F. and D. P. Lettenmaier (2005) *Production of Temporally Consistent Gridded Precipitation and Temperature Fields for the Continental United States* *Journal of Hydrometeorology*, 6, 330-336.
- Hamlet, A.F., M.M. Elsner, G.S. Mauger, S-Y. Lee, I. Tohver, and R.A. Norheim. 2013. An overview of the Columbia Basin Climate Change Scenarios Project: Approach, methods, and summary of key results. *Atmosphere-Ocean* 51(4):392-415, doi: 10.1080/07055900.2013.819555.
- Hession, W. C., Pizzuto, J. E., Johnson, T. E., & Horwitz, R. J. (2003). Influence of bank vegetation on channel morphology in rural and urban watersheds. *Geology*, 31(2), 147-150.
- Liang X, Wood EF, Lettenmaier DP (1996) Surface soil moisture parameterization of the VIC-2L model: Evaluation and modifications. *Glob Planet Chang* 13: 195-206

- Liang X, Wood EF, Lohmann D, Lettenmaier DP, and others (1998) The project for intercomparison of land-surface parameterization schemes (PILPS) phase-2c Red-Arkansas River basin experiment: 2. Spatial and temporal analysis of energy fluxes, *J Glob Planet Chang* 19: 137-159
- Livneh B., T.J. Bohn, D.S. Pierce, F. Munoz-Ariola, B. Nijssen, R. Vose, D. Cayan, and L.D. Brekke, 2015: A spatially comprehensive, hydrometeorological data set for Mexico, the U.S., and southern Canada 1950-2013, *Nature Scientific Data*, 5:150042, doi:10.1038/sdata.2015.42.
- Livneh B., E.A. Rosenberg, C. Lin, B. Nijssen, V. Mishra, K.M. Andreadis, E.P. Maurer, and D.P. Lettenmaier, 2013: A Long-Term Hydrologically Based Dataset of Land Surface Fluxes and States for the Conterminous United States: Update and Extensions, *Journal of Climate*, 26, 9384–9392.
- Mauger, G.S., J.S. Won, K. Hegewisch, C. Lynch, R. Lorente Plazas, E. P. Salathé Jr. (2018). New Projections of Changing Heavy Precipitation in King County. Report prepared for the King County Department of Natural Resources. Climate Impacts Group, University of Washington, Seattle.
- Mauger, G.S., S.-Y. Lee, C. Bandaragoda, Y. Serra, J.S. Won, (2016). Effect of Climate Change on the Hydrology of the Chehalis Basin. Report prepared for Anchor QEA, LLC. Climate Impacts Group, University of Washington, Seattle. doi:10.7915/CIG53F4MH
- Mauger, G.S., J.H. Casola, H.A. Morgan, R.L. Strauch, B. Jones, B. Curry, T.M. Busch Isaksen, L. Whitely Binder, M.B. Krosby, and A.K. Snover (2015). State of Knowledge: Climate Change in Puget Sound. Report prepared for the Puget Sound Partnership and the National Oceanic and Atmospheric Administration. Climate Impacts Group, University of Washington, Seattle. <https://doi.org/10.7915/CIG93777D>
- Meehl, G. A., Covey, C., Delworth, T., Latif, M., McAvaney, B., Mitchell, J. F., ... & Taylor, K. E. (2007). The WCRP CMIP3 multimodel dataset: A new era in climate change research. *Bulletin of the American meteorological society*, 88(9), 1383-1394.
- Mote, P., J. Abatzoglou, D. Lettenmaier, D. Turner, D. Rupp, D. Bachelet, D. Conklin (2016). *Final Report for Integrated Scenarios of climate, hydrology, and vegetation for the Northwest*. <https://climate.northwestknowledge.net/IntegratedScenarios/>
- Nakicenovic, N. et al., 2000. Special Report on Emissions Scenarios: A Special Report of Working Group III of the Intergovernmental Panel on Climate Change, Cambridge University Press, Cambridge, U.K., 599 pp. Available online at: <http://www.grida.no/climate/ipcc/emission/index.htm>

- Nick, M., Das, S. and Simonovic, S.P. 2011. The Comparison of GEV, Log-Pearson Type 3 and Gumbel Distributions in the Upper Thames River Watershed under Global Climate Models, the University of Western Ontario Department of Civil and Environmental Engineering, Report No:077. <https://ir.lib.uwo.ca/wrrr/40/>
- Pytlak, E., C. Frans, K. Duffy, J. Johnson, B. Nijssen, O. Chegwiddden, D. Rupp (2018). Second Edition: Climate and Hydrology Datasets for RMJOC Long-Term Planning Studies (RMJOC-II). Part I: Hydroclimate Projections and Analyses. <http://hydro.washington.edu/CRCC/>
- Rahman, A., Weinmann, P.E. and Mein, R.G. (1999). At-site flood frequency analysis: LP3-product moment, GEV-L moment and GEV-LH moment procedures compared. In: Proceeding Hydrology and Water Resource Symposium, Brisbane, 6–8 July, 2, 715–720.
- Rahman, A., Karin, F, and Rahman, A. 2015. Sampling Variability in Flood Frequency Analysis: How Important is it? 21st International Congress on Modelling and Simulation, Gold Coast, Australia, Nov 29-Dec 4, 2015, 2200-2206.
- Salathé, E.P., Leung, L.R., Qian, Y. et al. 2010. Regional climate model projections for the State of Washington, Climatic Change,102:51. <https://doi.org/10.1007/s10584-010-9849-y>
- Salathé Jr, E. P., Hamlet, A. F., Mass, C. F., Lee, S. Y., Stumbaugh, M., & Steed, R. (2014). Estimates of 21st century flood risk in the Pacific Northwest based on regional climate model simulations. *Journal of Hydrometeorology*, (2014).
- Taylor, K. E. et al., 2012. An overview of CMIP5 and the experiment design. *Bulletin of the American Meteorological Society*, 93(4), 485-498, doi:10.1175/BAMS-D-11-00094.1
- 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. *JAWRA Journal of the American Water Resources Association*, 50(6), 1461-1476. <https://doi.org/10.1111/jawr.12199>
- Vogel, R.M., McMahon, T.A. and Chiew, F.H.S. (1993). Flood flow frequency model selection in Australia, *Journal Hydrology*, 146, 421-449. [https://doi.org/10.1016/0022-1694\(93\)90288-K](https://doi.org/10.1016/0022-1694(93)90288-K)
- Warner, M. D., Mass, C. F., & Salathé Jr, E. P. (2015). Changes in winter atmospheric rivers along the North American west coast in CMIP5 climate models. *Journal of Hydrometeorology*, 16(1), 118-128. <https://doi.org/10.1175/JHM-D-14-0080.1>
- George Wilhere (WDFW), Jane Atha (WDFW), Timothy Quinn (WDFW), Lynn Helbrecht (WDFW) and Ingrid Tohver (Climate Impacts Group) (2016). *Incorporating Climate*

Change into the Design of Water Crossing Structures: Final Project Report.
Washington Department of Fish and Wildlife. Olympia, Washington.

Wilks, D. (1995). Statistical Methods in Atmospheric Sciences: An Introduction. Academic Press.