Final Report to "Expansion of the PCIC Climate Explorer - Online Design Value Projection Tool and Future Initiatives Planning"

Agreement Number # 820LA0009

March 31, 2024

Markus A. Schnorbus, Mohamed Ali Ben Alaya, and Narges Dehkordi

Pacific Climate Impacts Consortium

Prepared for: Engineering Services Branch, Integrated Transportation and Infrastructure Services Division, Ministry of Transportation and Infrastructure, Government of British Columbia.

University of Victoria

Table of Contents

1 Introduction

Based on the BC Ministry of Transportation and Infrastructure (MoTI) led synthesis of Vulnerability Assessments, completed using the Public Infrastructure Engineering Vulnerability Committee (PIEVC) Protocol, climate change has increased risk to transportation infrastructure in BC (BC MoTI et al., 2014). To address this increased risk, the BC MoTI released directives and guidance for incorporating climate adaptation into engineering designs in its T-04/19 Technical Circular (BC MoTI, 2019). This guidance document stipulates transportation engineering design projects should "incorporate information, analyses and projections of the impact of future climate change and weather extremes". It also lists a few sources of climate change information such as the Pacific Climate Impacts Consortium's (PCIC's) analysis tools, including the Climate Explorer and Plan2Adapt.

To support the requirement for guidance on climate change impacts related to peak flow, MoTI supported the Pacific Climate Impacts Consortium (PCIC) in a pilot project to estimate historical and future design flow values for the upper Fraser watershed, a 34200 km^2 section of the Fraser River basin upstream of Prince George, BC (Schoeneberg et al., 2021). Historic and future streamflow used results derived from VIC-GL hydrologic projections forced with the CanESM2 50-member initial-conditions ensemble (CanESM2-LE). These design flood values were estimated on the 0.0625° x 0.0625° VIC-GL computational grid and were provided as a prototype online design flow tool that has been integrated into the PCIC Climate Explorer (PCEX). Following a review by MoTI staff, PCIC was asked to revise and upgrade the tool, including expanding the study domain to include the entire Fraser and Peace basins, provide design flow estimates as a ratio of future value to historic value, and provide changes factors at 10-year intervals using a 30-year sliding window, centered on 2015, 2025, 2035, 2045, 2055, 2065, 2075, and 2085 based on CanESM2 Large Ensemble with RCP8.5 (Schnorbus and Sun, 2022; Schnorbus and Ben Alaya, 2023).

Following from this work, MoTI has as expressed that attention be given to a range of issues. Firstly, MoTI has expressed a need to continue to expand the spatial domain of the peak flow design values available through the PCEX. Secondly, MoTI has also identified a requirement to consider extreme low flows in the design of transportation infrastructure (e.g. with respect to ensuring safe fish passage). Thirdly, several future initiatives have been identified that include the requirement for staff training, continued exploration of low flow extremes, and exploring the opportunity to leverage off work currently underway at PCIC to provide updated extreme high and low flow design values on high-resolution routing network.

Based on the broad range of issues being addressed the project has been grouped into two main components: Tool Expansion and Future Initiatives. The 'Tool Upgrade' component of the current project has seen the spatial expansion of the peak flow design analysis to include the upper Columbia basin. As well, this project also includes a pilot study that explores and approach to develop extreme low flow metrics, using the Nicola basin (a tributary of the Thompson River) as a test case. The 'Tool Upgrade' work again takes advantage of hydrologic projections produced by PCIC using the VIC-GL hydrology model driven with the CanESM2 50-member large ensemble (CanESM2-LE) (Government of Canada, 2019; Kushner et al., 2018). The 'Future Initiatives' component of the project is intended to briefly lay out a general blueprint of options for future work based on the aforementioned emerging issues. These include several options to expand of the analysis extreme low flows, and a proposal to upgrade current and future extreme flow metrics by delivering data on an improved vector-based routing network. The

current project has also seen the development of a high-level plan for training MoTI staff in the interpretation and utilization of climate change information.

The purpose of this report is to describe the work done on Tool Upgrade and Future Initiatives components of the project. Tool Upgrade describes the spatial expansion of the domain of the peak design flow values and explores the results of our pilot study into the use of extreme low flow metrics. For Future Initiatives this report provides several options for continued work exploring and developing extreme low flow metrics, describes an option to downscale design values to a higher resolution vector-based routing network, and provide a proposal to provide MoTI staff training to promote effective use of PCIC climate and hydrologic projection knowledge and tools. In the remainder of the report, section 2 describes Tool Upgrades, which is further divided into a description the study area in section 2.1, a summary the VIC-GL simulation methodology in section 2.2, a description of the peak flow analysis in section 2.3, and a description of the low flow study in section 2.4. Results of both the peak flow analysis and low flow pilot study are provided in Section 2.5. A discussion of Future Initiatives comprises section 3. A combined model verification for the entire study domain has also been completed and is included as Appendix A.

2 Tool Upgrade

2.1 Study Area

2.1.1 Peak Flow Design Values

For this aspect of the project, the spatial domain includes the Fraser basin, the Peace River basin above Peace River, AB, and the upper Columbia River above the Priest Rapids Dam[1], for a combined drainage area of 665,000 km^2 (Figure 1). Based on PNWNAmet (Werner et al., 2019) the 1981-2000 mean annual temperature for the three basins combined is 1.0 °C. Mean annual temperature ranges from a low of -4 °C at high elevations in the Coast and Columbia Mountains (Fraser and upper Columbia) and the Cassiar Mountains (northern Peace basin), to a high of 10-11 °C in the low-lying Fraser River delta and southern Columbia (Figure 2). Due to the complex topography, precipitation in the basin exhibits high spatial variability (Figure 3). Annual precipitation is highest along the windward length of the Coast Mountains in the lower Fraser, with a secondary wet band along the Rocky and Columbia Mountain in the headwaters of the Fraser, Columbia, and Peace basins. Due to a heavy rain shadow effect, precipitation decreases dramatically in the interior of the Fraser and Columbia and east of the Rockies in the Peace. Precipitation is highest (> 4000 mm) along the windward-facing North Shore Mountains at the basin outlet, and lowest (< 300 mm) in the central interior.

2.1.2 Low Flow Pilot Study

The low flow pilot study was conducted using the Nicola basin as the test region (Figure 1). This 7100 km^2 basin is a tributary of the Thompson River and enters the Thompson 235 km upstream of its confluence with the Fraser. The Nicola's main tributaries include the Coldwater River, Guichon Creek, Moore Creek, Clapperton Creek, Stump Lake Creek, Quilchena Creek, and Spius Creek. Located leeward of the Coast and Cascade Mountains, the Nicola has a relatively arid climate with an average annual precipitation of less than 700 mm throughout most of the basin, although annual mean precipitation exceeds 1500 mm in the headwaters of the Coldwater (south-west corner of basin) (Figure 3). Annual average temperature ranges between 2 °C and 6 °C, depending upon elevation (Figure 2).

Figure 4 shows the climatological daily hydrograph at the basin outlet, which corresponds to Water Survey of Canada station 08LG006 (Nicola River near Spences Bridge). Note that streamflow is plotted on a logarithmic scale to highlight low flows. The hydrograph exhibits behavior typical of a snowmelt regime, with highest flows in the spring resulting from snowmelt runoff and two low flow periods, one in the summer (centered on day-of-year 260) and on in the winter (centered on day-of-year 365). Although mean low flow is lowest in summer, flow variability is higher in winter such that extreme low flows events are possible in either winter or summer.

[1] https://en.wikipedia.org/wiki/Priest_Rapids_Dam

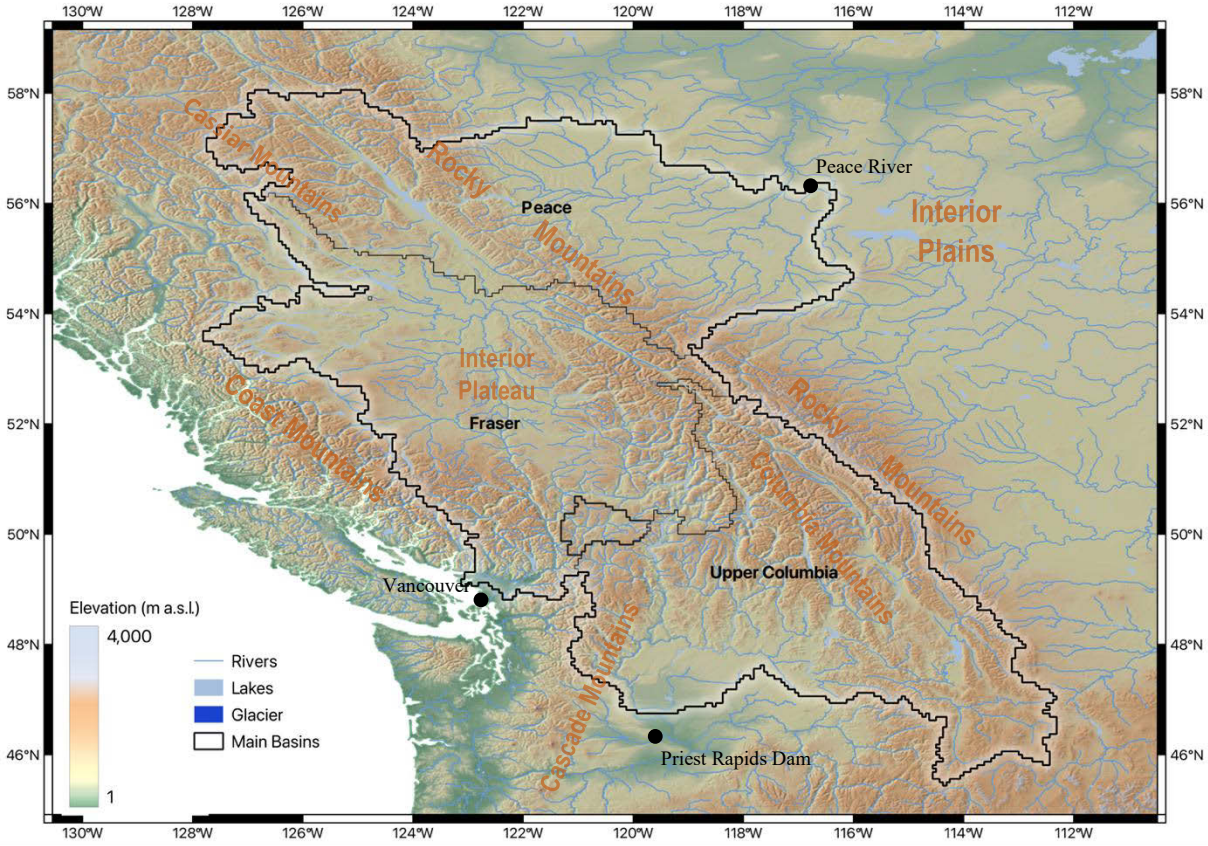

Figure 1. Project domain showing the upper Columbia basin above Priest Rapids Dam, the Fraser basin above tidewater at Vancouver, and the Peace basin above Peace River, AB. Also shown is the outline for the Nicola sub-basin.

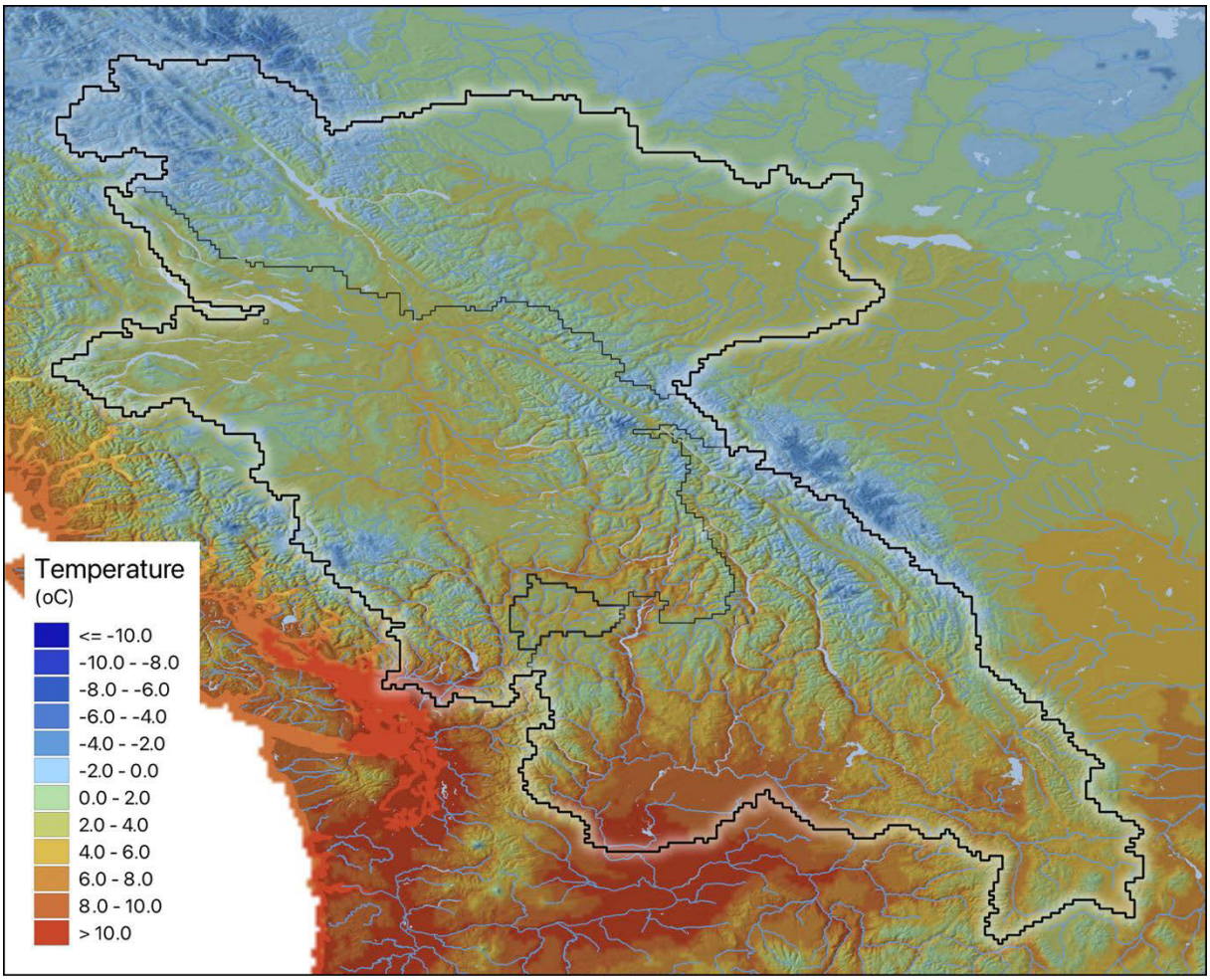

Figure 2. Annual average temperature for 1980-2010 from PNWNAmet.

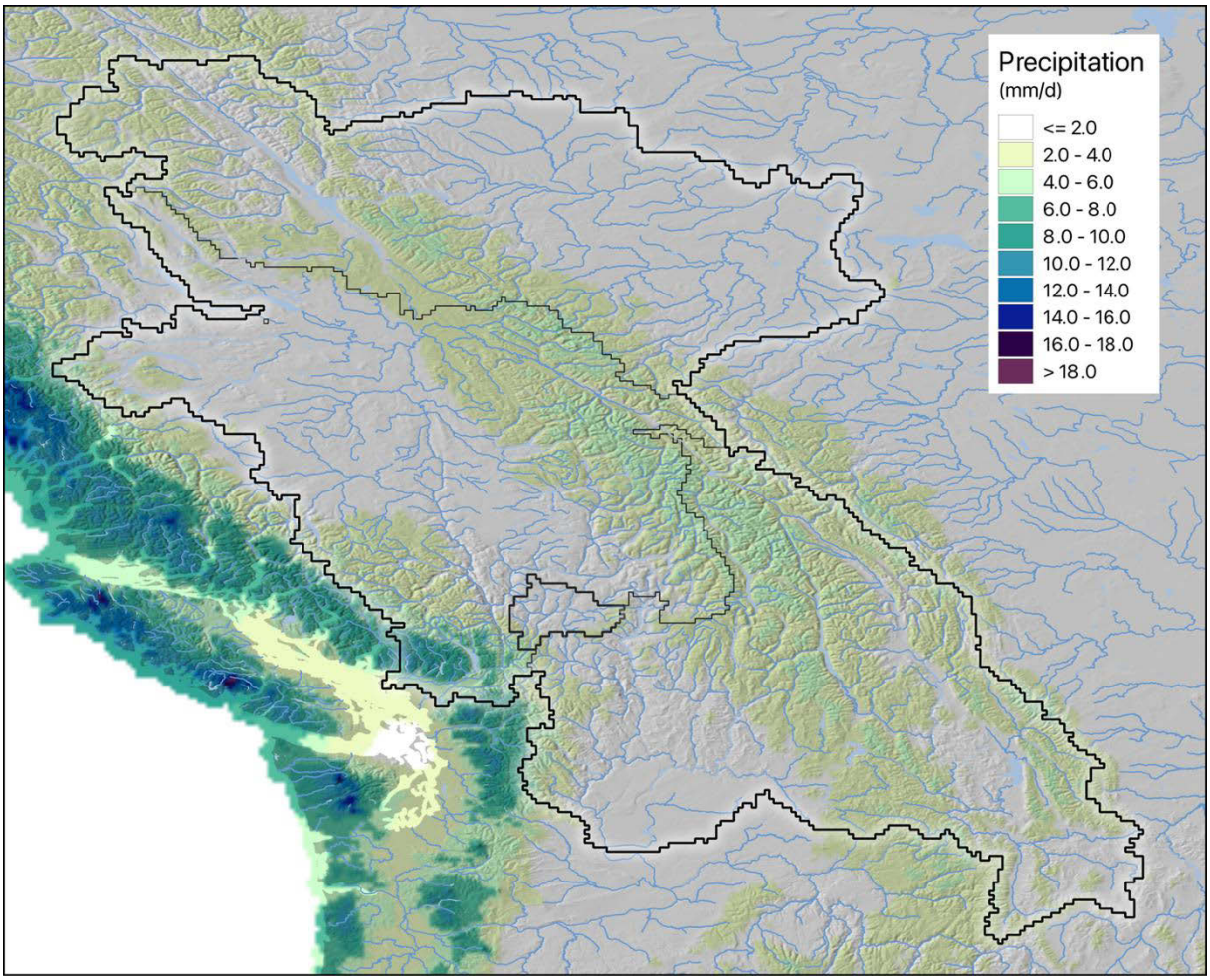

Figure 3. Annual average precipitation for 1980-2010 from PNWNAmet.

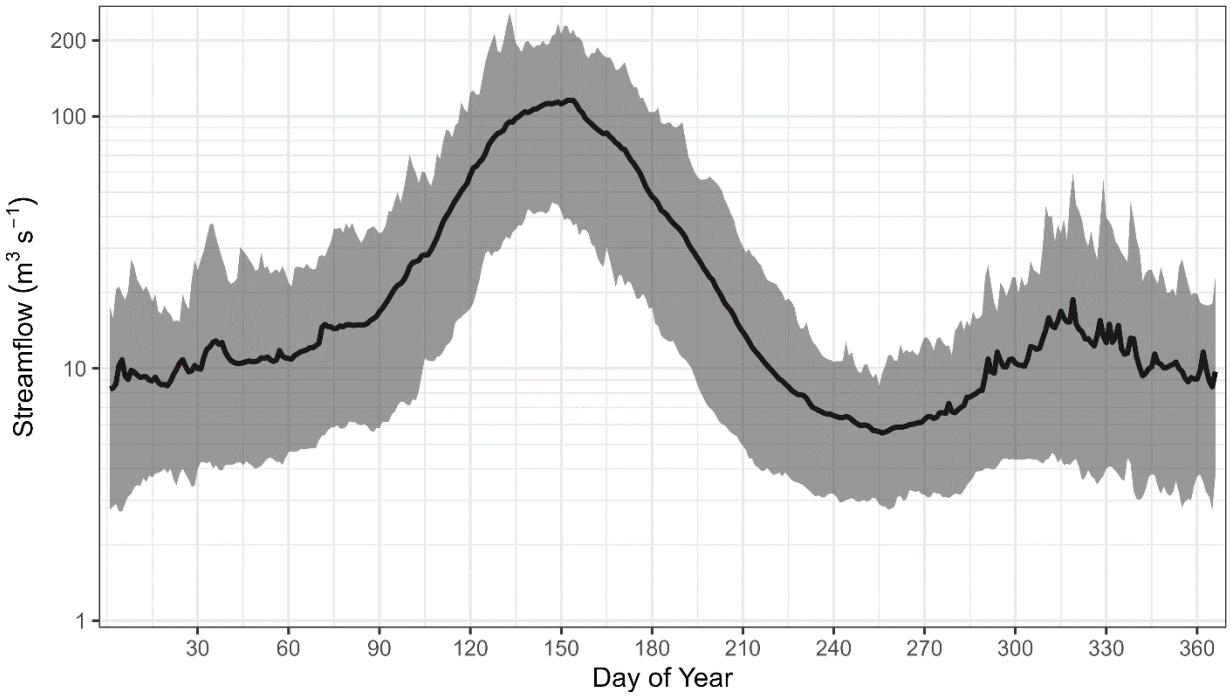

Figure 4: Hydrograph of climatological daily discharge for the Nicola River near Spence's Bridge (Water Survey of Canada station 08LG006). Thick black line is the climatological mean for each day of the year and the shading shows the range between the 5th- and 95th-percentile. The climatological period is 1957 to 2021.

2.2 Hydrologic Projection Methodology

Both the peak flow frequency analysis (Section 2.3) and the extreme low flow study (Section 2.4) were performed on daily streamflow simulated using the VIC-GL model driven by the CanESM2 Large Ensemble, which was forced with the RCP8.5 scenario. Both analyses were performed on a cell-by-cell basis on the VIC-GL computational grid. The projections of future streamflow and subsequent analysis of extremes requires a chain of models. The following provides details on each modelling component and summarizes the overall study design. A graphical summary of the study design is given in Figure 5. The implications of the chosen design with respect to modelling uncertainty and limitations will be discussed in Section 2.6.

2.2.1 CanESM2 – Large Ensemble

Future climate projections at the regional scale start with model simulations of the global climate system. Because these are models, there is inherent uncertainty related to (1) the assumptions about how the greenhouse gases (GHG) will evolve, (2) the climate model and how it represents the physical processes and (3) the internal variability, e.g. the natural variability that we experience as weather or El Nino events, which is irreducible (Arora and Cannon, 2018; Cannon et al., 2020). The hydrologic projections were produced using climate simulations from CanESM2, which is a coupled Earth system model developed and run by the Canadian Centre for Climate Modelling and Analysis (Arora et al., 2011). CanESM2 is part of the World Climate Research Programs (WCRP) fifth Coupled Model Inter-comparison Project (CMIP5) (Taylor et al., 2011). The CMIP5 submission of CanESM2 included five ensemble members run with historical forcings from 1850 to 2005. A much larger ensemble of climate projections was produced by expanding the ensemble to 50 members, each spanning the period 1950 to

2100 (Government of Canada, 2019; Kirchmeier-Young et al., 2017; Kushner et al., 2018). A random number generator with a pre-set seed was used to perturb slightly the initial state of each of the 50 ensemble members. Thereby, quasi-independent climate change realizations were generated without any change to the model dynamics, physics or structure (Fyfe et al., 2017). The resulting ensemble represents 50 equally plausible realizations of the evolution of the global weather and climate that are consistent with the observed emissions over the period 1950 to 2005 and the RCP 8.5 emissions scenario from 2006 to 2100. In the RCP8.5 scenario, emissions continue to rise throughout the 21st century and this scenario is often used as the basis for worst case climate change.

2.2.2 Downscaling

The climate response to a prescribed RCP scenario that is obtained from a climate model is of too coarse a spatial resolution, with individual grid cells typically encompassing 10,000 km^2, to be used directly in driving a hydrology model. For example, GCM output at this resolution does not reflect the detailed spatial variation in climate due to local orography and variations in land surface properties that are necessary for simulating surface hydrology well. Therefore, to model changing hydrologic conditions at local and regional scales, daily values of minimum temperature, maximum temperature and precipitation have been statistically downscaled to the resolution of VIC-GL. This downscaling used the Bias Correction/Constructed Analogues with de-trended Quantile mapping reordering downscaling technique (BCCAQv2) (Hiebert et al., 2018) with PNWNAmet (Werner et al., 2019) as the reference meteorology. BCCAQv2 is a hybrid method that combines results from bias-corrected constructed analogs (BCCA) (Maurer et al., 2010) and de-trended quantile mapping (QMAP) (Gudmundsson et al., 2012). BCCA obtains spatial information from a linear combination of historical analogues for daily large-scale fields. QMAP applies quantile mapping to daily climate model outputs interpolated to the high-resolution grid using the climate imprint method of Hunter and Meentemeyer (2005). The BCCAQv2 method includes a revision to the quantile mapping procedure that better preserve changes in quantiles and extremes (Cannon et al., 2015) as compared to its original implementation. BCCAQv2 works well for hydrologic extremes because of its ability to resolve event-scale spatial gradients (Werner and Cannon, 2016). For more information on BCCAQv2 see (Cannon et al., 2015; Hiebert et al., 2018; Sobie and Murdock, 2017; Werner and Cannon, 2016).

2.2.3 VIC-GL Model Summary

Streamflow was simulated with VIC-GL, an upgraded version of the Variable Infiltration Capacity (VIC) model that explicitly models glacier mass balance (accumulation, melt and runoff) and glacier dynamics (change in glacier area) (see Schnorbus 2018 for details). VIC is a spatially distributed macro-scale hydrologic model that calculates water and energy balances in each grid cell. Spatial variability in soil properties within a drainage basin is modelled by sub-dividing the model domain into a computational grid with a spatial resolution of 0.0625° latitude by 0.0625° longitude (approximately 6 km x 5 km within the study region). The variability of land cover and topography within individual grid cells is further described using hydrologic response units (HRUs) which characterize land surface properties as a function of elevation. VIC runs at a 3-hour temporal resolution and output is aggregated to daily values. Soil moisture processes are represented by three-soil layers, spatial heterogeneity of runoff generation with variable infiltration curves, and subsurface flow generation using the Arno conceptual model (Todini, 1996). Surface runoff is generated when the moisture exceeds the storage capacity of the soil. Water fluxes are computed for a range of hydrologic processes such as evapotranspiration, snow accumulation, snowmelt, infiltration, soil moisture and surface and subsurface runoff. A detailed description of the baseline VIC model is available in Liang et al. (1996, 1994) and Cherkauer et al. (2003).

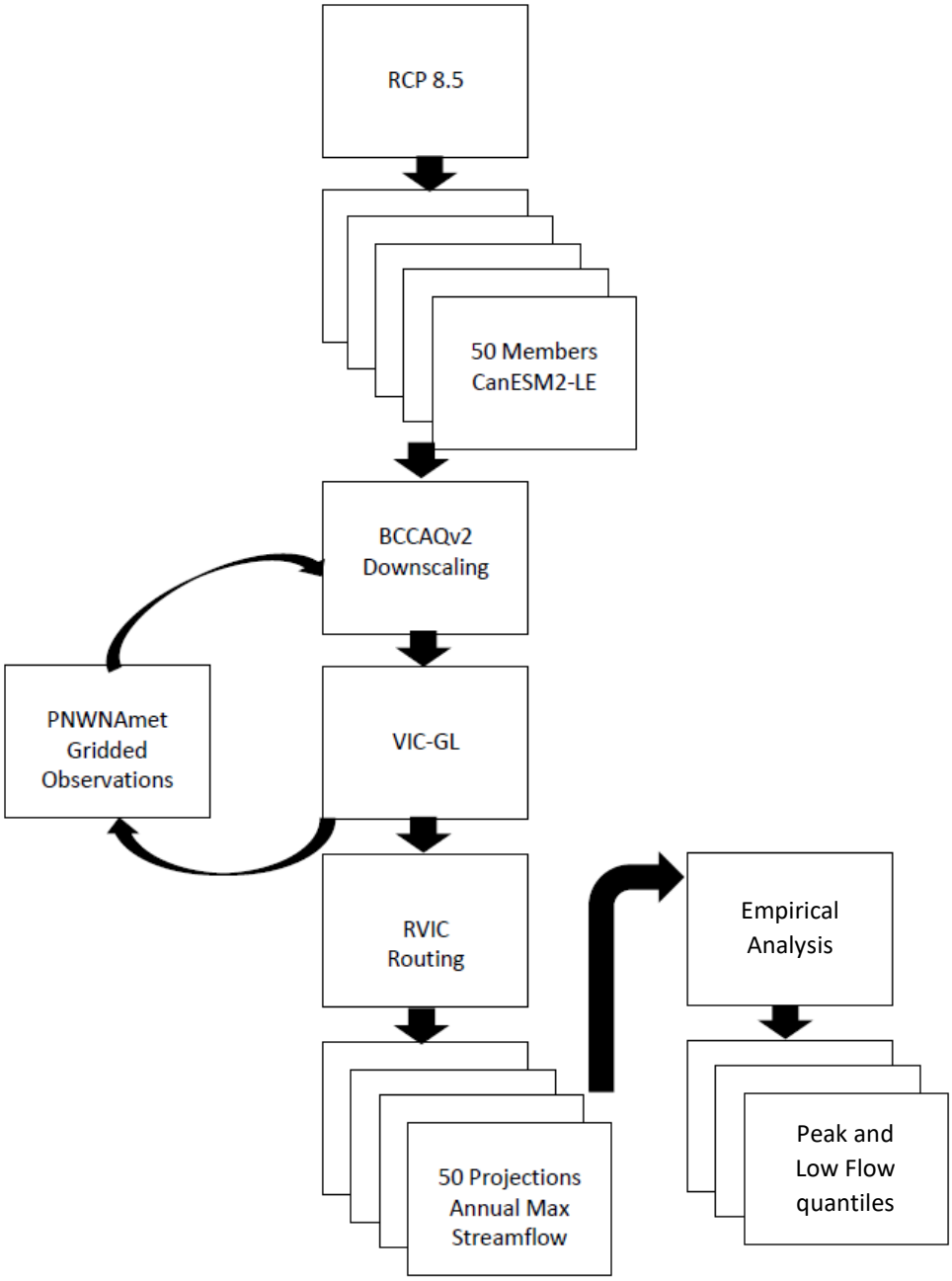

Figure 5. Modelling workflow used to derive estimates of annual maximum and 7-day minimum streamflow quantiles from the CanESM2 large ensemble.

VIC-GL uses several parametrization strategies to describe the influence of topography and vegetation cover. Sub-grid elevation is described using 200-m elevation bands derived from the GMTED2010 digital elevation model (Danielson and Gesch, 2011). Vegetation classification utilizes the North America Land Cover dataset, edition 2 (Natural Resources Canada / The Canada Centre for Mapping and Earth

Observation 2013) produced as part of the North America Land Change Monitoring System (NALCMS). The NALCMS land cover data set divides North America into 19 classes representing circa 2005 conditions, with most forest areas in the region for which VIC-GL has been parameterized being included in a single class, the temperate or sub-polar needle-leaf forest class. This homogeneous region has therefore been further subdivided based on vegetation height and leaf area index. Leaf are index data is from the GEOV1 global time series dataset (Baret et al., 2013; Camacho et al., 2013). Vegetation height is based on global mapping using space borne light detection and ranging (LIDAR) (Simard et al., 2011). The final land cover classification, with needle-leaf forest further sub-divided, contains 22 land cover classes. Although an Ice class exists in the NALCMS-based land cover inventory, the extent and location of glaciers and ice fields was updated using the Randolph Glacier Inventory (RGI) version 3.2 (Pfeffer et al., 2014). Soil classification and parameterization relies on physical soil data from the Soils Program in the Global Soil Data Products CD-ROM (Global Soil Data Task, 2014).

Calibration is the process whereby certain model parameters are adjusted such that simulated output is in close agreement with observations. During the calibration process, VIC-GL was forced with the PNWNAmet gridded meteorological data set (Werner et al., 2019). Model calibration used a multi-objective approach that constrained the model using observations of streamflow, evaporation, snow cover and glacier mass balance (estimated from thinning rates). For more details on calibration process, see Schnorbus (2017). An evaluation of VICGL overall performance for the Peace and Fraser is reported in Schnorbus (2021) and direct verification of simulated annual maximum streamflow for the Peace basin is provided as Appendix A.

2.2.4 Surface Routing

Surface water routing is applied as a post-processing step in the modelling chain using the RVIC model (https://rvic.readthedocs.io/en/latest/), which is based on the numerical schemes described in Lohmann et al. (1998). Note that streamflow simulated by the VICGL-RVIC models only represents natural flows. Hence the results will not reflect the effects of flow regulation in those portions of the Peace, Fraser and upper Columbia that are affected by reservoir storage and releases. A two-step process is used to rout runoff and baseflow generated in each VIC-GL model cell: in-grid routing and channel routing. In-grid routing conceptually moves surface runoff through the sub-grid drainage network to the main channel using a transfer function that essentially describes the time distribution for runoff reaching the outlet of a grid box. The transport of water within the channel is modelled using a one-dimensional diffusive wave approximation to the full Saint Venant equations. The channel system is defined by assigning one of eight flow directions (N, NE, E, SE, S, SW, W or NW) to each cell. The resultant channel network and flow accumulation (the number of upstream cells) for all three main basins in the study area is shown in Figure 6.

In-grid routing is parametrized by specifying the unit hydrograph for each grid cell. Channel routing requires specification of the channel length, wave celerity and diffusivity for each grid cell. Based on manual calibration, wave celerity is assumed constant at 2.0 m/s and channel diffusivity is set to 1300 m. Channel length is estimated as the cell height and cell width for north-south, east-west flow directions, respectively, and as the cell diagonal for all remaining flow direction.

For this project it was determined that hydrologic projections for one ensemble member could not be used due to incomplete spatial coverage. Hence, only 49 ensemble members were used in subsequent analysis.

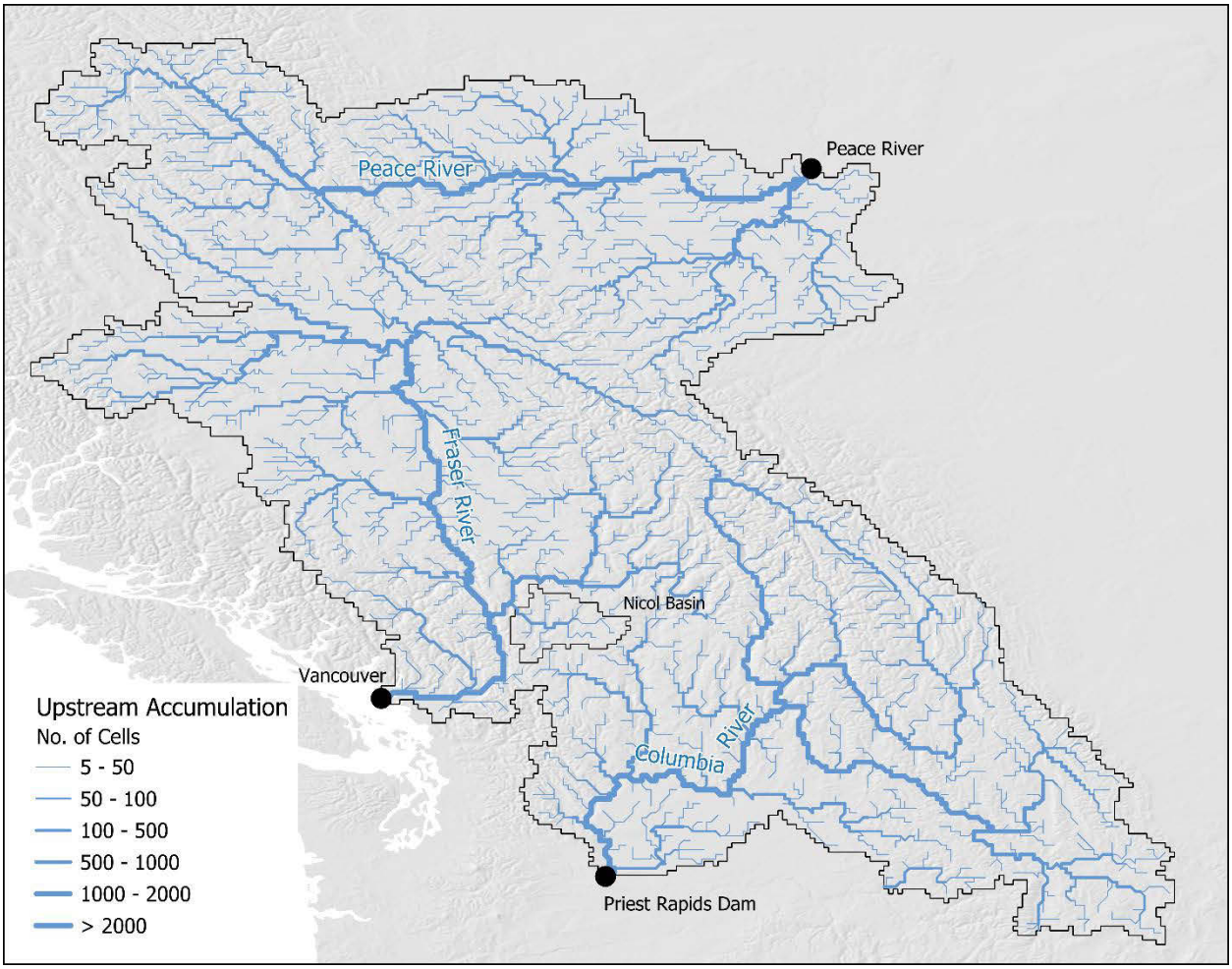

Figure 6. VICGL flow network for the upper Columbia, Fraser, and Peace with flow accumulation (as the number of upstream cells). Also shown is the Nicola sub-basin. The flow threshold has been set to five grid cells (i.e. channels with upstream flow accumulation less than five cells are excluded from the image).

2.3 Flood Frequency Analysis

The design of roads, bridges, culverts, and other structures often requires estimates of peak flow quantiles that correspond to return periods varying from 2-years up to perhaps 200-years, depending on the application. When reliant on small sample sizes to estimate peak flow quantiles, as is typically the case when using observed data, these quantiles often correspond to return periods that are substantially longer than the sample. In these cases, one must generally resort to parametric flood frequency approaches so that one can extrapolate beyond the data to produce the necessary quantile estimates, such as using the Generalized Extreme Value, or GEV, distribution to describe the available peak flow data. Unfortunately, parametric approaches may suffer from some lack-of-fit that can result in biased quantile estimates, particularly for infrequent events. The benefit of using the CanESM2 large ensemble is that it provides enough samples of identically distributed annual maximum events that the quantiles, even for very large design values such as the 200-year event, can be estimated directly from the empirical cumulative distribution.

Streamflow extremes were analyzed at ten-year intervals using a 30-year sliding window centered on 2015 (2001- 2030), 2025 (2011-2040), 2035 (2021-2050), 2045 (2031-2060), 2055 (2041- 2070), 2065 (2051-2080), 2075 (2061-2090), and 2085 (2071-2100). Annual maxima were extracted for each year and pooled across model runs of the large ensemble, resulting in 1350 values for each thirty-year period (30 years by 45 ensemble members) from streamflow simulated at each grid point. Quantiles were estimated empirically using the function "quantile" from the R stats package (R Core Team, 2019). We used the default quantile algorithm where the p^{th} percentile from the sample is estimated as:

$$x_{\lfloor h \rfloor} + (h - \lfloor h \rfloor) \times (x_{\lceil h \rceil} - x_{\lfloor h \rfloor}) \quad (1)$$

where $h = (N-1)p + 1$ and N is the number of samples. Note that for very high quantiles, empirical estimates can be somewhat biased depending on the choice of plotting position formula. However, because 200-year return levels (99.5 percentiles) are estimated from 1350 years of data we don't consider this as a point of concern. We estimate the 50th, 80th, 90th, 95th, 98th, 99th, and 99.5th percentile of the 1350 samples, which corresponds to the 2-, 5-, 10-, 20-, 50-, 100- and 200-year return levels, respectively. We define the median quantile estimates as those derived from the full sample. The confidence intervals for each quantile are quantified using the following bootstrapping approach: (1) for each 30-year block randomly sample with replacement a sample of size *n*=1350 from the pooled results across all 45 ensemble members; (2) estimate the quantiles from the new sample using equation 1; (3) repeat steps (1) and (2) 1000 times, and (4) estimate the 2.5% and 97.5% percentiles from the 1000 quantile estimates at each return level to obtain the 95% confidence intervals.

Projections of change were assessed by converting the absolute streamflow design values into change factors. Change factors are the ratio of future value to baseline value. Change factors utilize a baseline period of 1951-2000. This means that the quantiles for the baseline period are estimated from samples of size 50x45 = 2250 years. The confidence intervals for each quantile are estimated in a similar bootstrapping fashion as the underlying quantiles using the following approach: (1) for each 30-year block randomly sample with replacement a sample of size *n*=1350 from the values pooled across all ensemble members; (2) randomly sample with replacement 45 years across all ensemble members from the baseline period for sample of size *n*=2250; (3) estimate quantiles from each thirty-year block and the baseline block and calculate change factor for each quantile; (4) repeat steps (1), (2), and (3) 1000 times; and (5) estimate the 2.5% and 97.5% change factor percentiles from the 1000 change factor estimates at each return level to obtain the 95% confidence intervals. As the change factor is a ratio of two random variables, this approach explicitly considers the underlying sampling variability of both quantiles (future and baseline) when estimating the uncertainly of the change factors.

2.4 Extreme Low Flow Metrics

Several low-flow statistics can be used to characterize the low-flow regime at a given location, but frequency analysis of extreme low flows is perhaps the most common approach. This approach uses the quantiles nQ_y of the lowest mean discharge over a consecutive *n*-day period corresponding to a recurrence interval (return period) of *y*-years (Ouarda et al., 2008; Smakhtin, 2001; Tallaksen et al., 2004). As opposed to peak flow design values, extreme low flow quantiles express the flow magnitude that has a given probability of not being exceeded. In other words, the hazard to be mitigated in this case is the occurrence of annual minimum flows *below* a given magnitude. Averaging intervals of 7, 10 or 30 days are common (Tallaksen et al., 2004), however intervals as high as 284 days have been reported (Smakhtin, 2001). However, longer intervals have the potential to generate dependency (serial correlation) in the time series of annual minima. The $7Q_{10}$ index is perhaps the most widely used index in the United States and Canada, although other indices such as $7Q_2$, $7Q_5$, $7Q_{20}$, $30Q_5$ and $30Q_{10}$ have also

been adopted (Ouarda et al., 2008; Pyrce, 2004). Therefore, our initial exploration of extreme low flow indices focusses on the development of a range of $7Q_y$ metrics using the Nicola basin as a test case (see Figure 1).

In British Columbia many streams exhibit two low flow periods. The storage of precipitation as snow in the winter, possibly accompanied by frozen soil, results in a winter low-flow season. Summer also corresponds to a low-flow season when snowmelt is exhausted and rainfall is relatively scarce. Therefore, as summer and winter low flow are driven by different processes, it is often appropriate to consider separate indices for each season. However, for the sake of simplicity we only consider an annual $7Q_y$ index in the current pilot study.

As is the case with flood frequency analysis, use of observed flow typically results in small sample sizes and a reliance on parametric flood frequency approaches that one can use to extrapolate beyond the data to produce the necessary quantile estimates. For low flow analysis, the two-parameter Weibull distribution (the EVIII form of the GEV distribution for minima, which is bounded below by zero) is a common choice. Again, the benefit of using the CanESM2 large ensemble is that it provides enough samples that quantiles can be estimated directly from the empirical cumulative distribution. Generation of the large ensemble streamflow projections is described in Section 2.2.

For frequency analysis to be valid the sample data must be independent, i.e. the data cannot exhibit serial correlation. Such persistence could occur as a result of memory in the hydrologic system resulting from large storages in groundwater, lakes, reservoirs and glaciers. Due to the dependence of low flow on storage depletion (i.e. low flows tend to occur during periods that are absent rainfall or snowmelt input), extreme low flow may be particularly susceptible to long-term persistence. Although lakes and reservoirs are not present in VIC-GL, the effect of groundwater and glacier storage may still impart serial correlation in the simulated low flows.

Also unique to low flow analysis is the occurrence of years with zero values, which is particularly common in arid regions, but may also occur during the dry season in temperate regions. Samples that contain zero values are generally treated as censored data (i.e. the river may actually be dry, however it is more often the case that discharge often falls below a recording limit). In these situation it is recommended to use a conditional probability model (Stedinger et al., 1993). However, this situation is largely avoided in hydrology models which do not have a recording threshold.

As per the peak flow analysis, the change in low flow extremes due to climate change will be addressed by conducting LFFA on a 1951-2000 baseline and for eight future periods at ten-year intervals using a 30-year sliding window centered on 2015 (2001- 2030), 2025 (2011-2040), 2035 (2021-2050), 2045 (2031-2060), 2055 (2041- 2070), 2065 (2051-2080), 2075 (2061-2090), and 2085 (2071-2100). Daily streamflow data were subjected to a 7-day moving average calculation. Annual 7-day average minima were then extracted for each year and pooled across model runs of the large ensemble, resulting in 1470 values for each thirty-year period (30 years by 49 ensemble members) from streamflow simulated at each grid point.

Quantiles were estimated empirically for the 50^{th}, 20^{th}, 10^{th}, 5^{th}, 2^{nd}, and 1^{st} percentiles of the 1470 annual minima samples, corresponding to below-threshold return levels of 2-, 5-, 10-, 20-, 50- and 100-years. Quantiles were estimated in Python using the quantile function in the numpy package using the default ‘linear’ method (Equation 1). We define the median quantile estimates as those derived from the full sample. The confidence intervals for each quantile were estimated using the bootstrapping approach in a manner identical to that employed for the peak flow design values (Section 2.3).

2.5 Results

2.5.1 Spatial Pattern of Projected Changes in Peak Flow

We describe the projected changes in design flood values at each individual grid cell in the Fraser, Peace, and upper Columbia domains. Change factors for the 2-year and 100-year events are displayed in Figure 7 and Figure 8, respectively. Specific runoff (runoff/unit area) is higher in wetter climates than it is in drier climates, so for an equal size drainage area, wetter climates produce larger magnitude flood events. Therefore, at the local scale (about 100 km^2), the spatial distribution of peak flow magnitude is broadly consistent with precipitation climatology (Figure 3). However, as runoff is integrated over increasingly larger scales, peak flow magnitude in larger basins tends also to increase with increasing drainage area. Consistent with the topology depicted in Figure 6, the largest peak flow values occur where flow concentrates along major tributaries and the main stem of the Fraser, Peace and Columbia Rivers. The result is that flow design values span several orders of magnitude across the basin. For example, the 2-year design event ranges from <1.0 m^3/s for headwater cells in the drier interior regions to over 9000 m^3/s, at the outlet of the Peace and over 12000 m^3/s at the outlets of the Fraser and upper Columbia, respectively (Figure 7). For the 100-year event the range is <1.0 m^3/s to over 14000 m^3/s, 17000 m^3/s and 19000 m^3/s at the outlets of the Peace, Fraser, and upper Columbia, respectively.

Changes in peak flow magnitude at each grid cell will be influenced by numerous factors, the relative importance of which vary as a function of spatial scale. For cells draining very small drainage areas, individual grid cell changes will be affected more by local elevation, relief, and changes to the local climate, whereas peak flow changes in cells draining larger areas are likely influenced by changes occurring in distant (potentially wetter) upstream locations. Consequently, at the individual cell scale the spatial patterns of peak flow change for the 2- and 100-year events displays a rather heterogeneous pattern, without a clear widespread relationship to either topography, climate, or network topology (Figure 7 and Figure 8, respectively). Spatial heterogeneity is also affected by noise resulting from model error and calibration artifacts. Model uncertainty is discussed in detail in Section 2.6.

For the 2-year design event, the projected median change factors over the basin show a mixed response at mid-century (2041-2070), where both increases and decreases in peak flow are projected in the Fraser and upper Columbia, but predominantly increasing flows are projected in the Peace (Figure 7). The Fraser and upper Columba are projected to experience increased design flows along the Coast and Rocky Mountains, but predominantly decreased flows in the interior of each basin. The same general spatial pattern repeats in the Fraser and Upper Columbia at end-century (2071-2100), although a more mixed response emerges in the Peace. For all three basins the magnitude of the change, whether increase or decrease, tends to be larger at end-century than at mid-century. The consensus on the direction of change is shows little spatial variability as judges by comparison of the 2.5% and 97.5% change factor percentiles for each period.

For the 100-year event the median ensemble change factors show that most of the domain is expected to experience an increase in design flow at both mid- and end-century (Figure 8). Nevertheless, several regions, notably the interior upper Columbia and portions of the interior Fraser, are expected to experience declining design flow values in mid- and end-century (Figure 8). However, unlike the 2-year design flow event, the level of consensus between ensemble members regarding the direction of change is not as great i.e. there are large regions where the direction of change is opposite between the 2.5% and 97.5% percentile change factor. Also notable is that the magnitude of change factors, especially where design flow is projected to increase, are larger for the 100-year event than for the 2-year event.

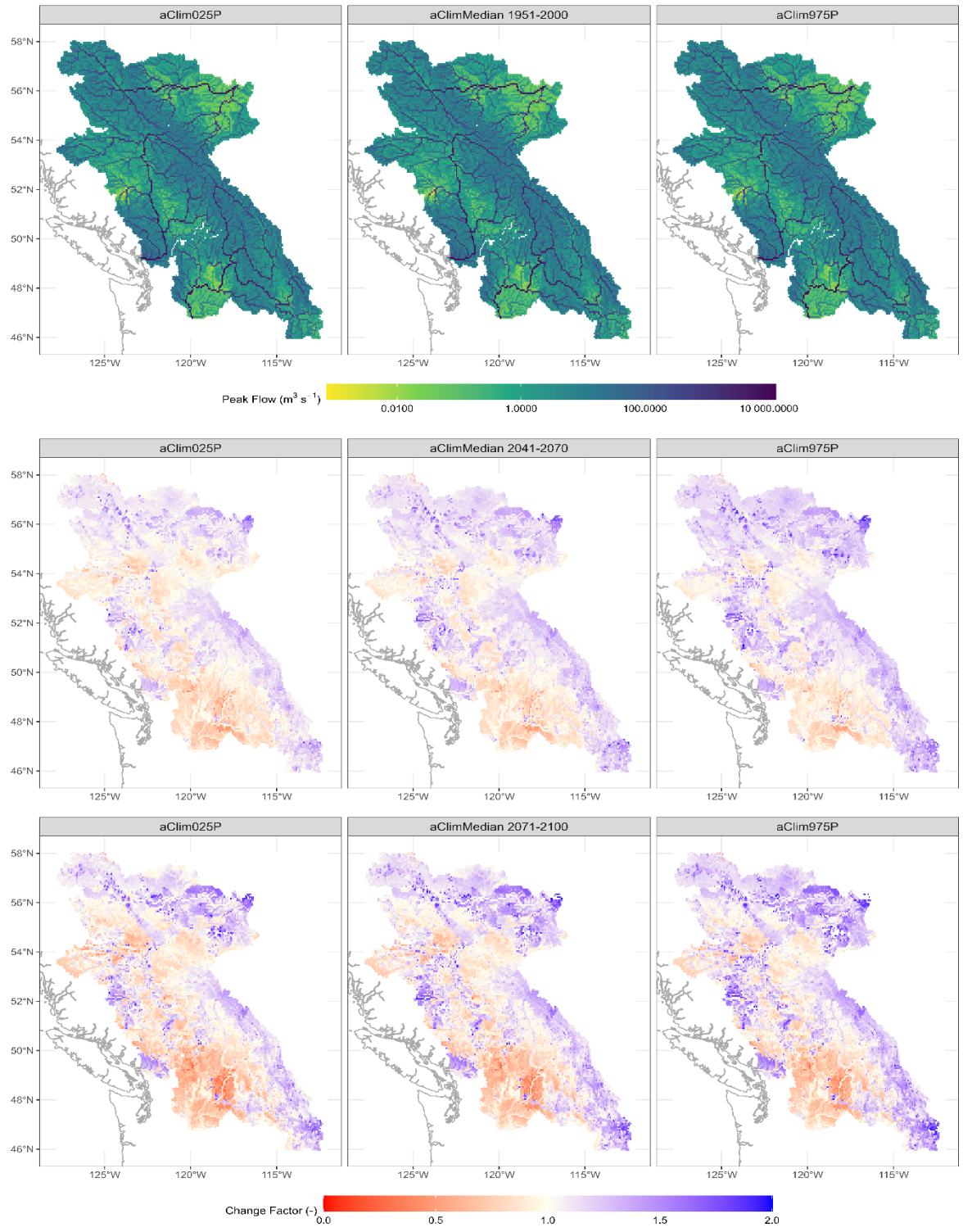

Figure 7. The absolute value (m^3/s) of annual peak flow during the 1951-2000 period (plotted on logarithmic scale) (top) and change factors in annual peak flow magnitude for 2-yr return period events for 2041-2070 and 2071-2100 versus the baseline period (1951-2000). Results are shown for 2.5th percentile (left), median (middle) and 97.5th percentile (right).

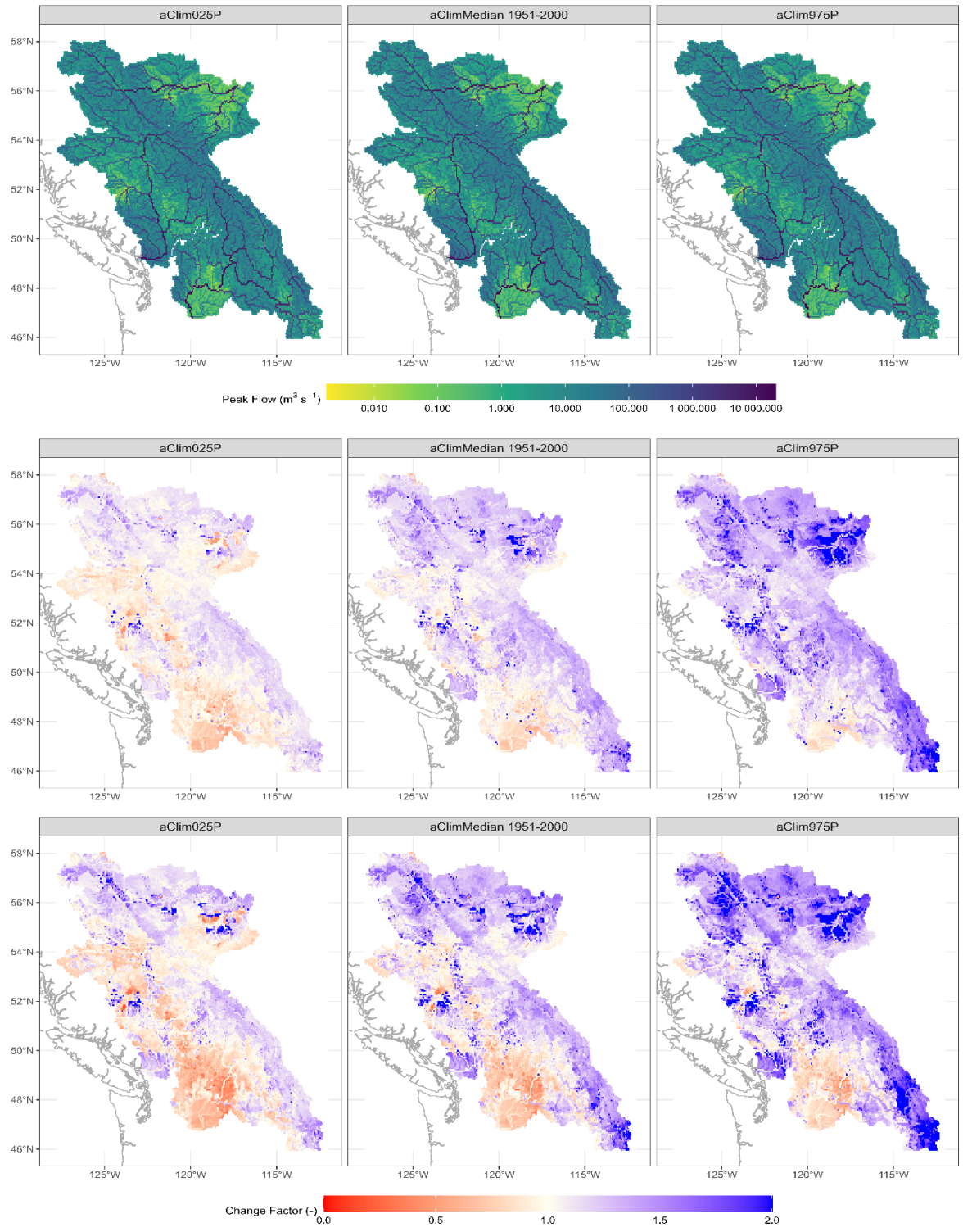

Figure 8. The absolute value (m^3/s) of annual peak flow during the 1951-2000 period (plotted on logarithmic scale) (top) and change factors in annual peak flow magnitude for 100-yr return period events for 2041-2070 and 2071-2100 versus the baseline period (1951-2000). Results are shown for 2.5th percentile (left), median (middle) and 97.5th percentile (right).

2.5.2 Spatial Pattern of Projected Changes in Extreme Low Flow

The gridded results for $7Q_{10}$ for the Nicola watershed are shown in Figure 9. As per peak flows, the magnitude of extreme low flows events is a function of both climate and drainage area. At local scale the spatial distribution of $7Q_{10}$ is broadly dictated by climate, where the wetter climate in the south of the watershed produces larger magnitude events. However, over increasingly larger scales, low flow magnitude tends to increase with increasing drainage area, and the largest events occur along the Nicola main stem and its various main tributaries (see top row of Figure 9).

Gridded projected changes in $7Q_{10}$ for mid- and end-century are shown in the bottom two rows on Figure 9. In the wettest portions of the Nicola basin (upper reaches of the Nicola, Coldwater and Spius) the $7Q_{10}$ is projected to become more extreme (i.e. drier; negative change factor) or change is projected to be negligible, whereas in the remainder of the basin the $7Q_{10}$ is projected to become to less extreme (i.e. wetter, positive change factor). In general, the trend, whether it is positive or negative change, increases in magnitude from mid- to end-century. Figure 9 reflects changes in the $7Q_{10}$ based on the annual minima. However, the season in which the streamflow minima occur (and the mechanisms that drive extreme low flow) may vary spatially within the domain such that some spatial variability in the climate response is likely attributable to the variation in low flow regime. This is discussed in greater detail in the following section.

2.5.3 Projected Change in Extreme Low Flow by Sub-basin

Results for four streamflow locations, the Nicola River near Spences Bridge (NICOS), the Nicola River at the outlet of Nicola Lake (NICOO), the Coldwater River at Merritt (COLDW), and Spius Creek near Canford (SPIUS) are shown in Figure 10. The extreme low flow varies between sites and between$7Q$ quantiles. Much of the variability, particularly that between quantiles, is possibly die to the hybrid nature of the low flow regime in snow-dominated basins (see Section 2.1.2 and Figure *4*). Specifically, the frequency curves in Figure *10* are a mix of summer and winter low flow events. The most straightforward response occurs at station NICOO, which drains the upper Nicola basin. At this location extreme low flow events for all return periods are projected to become more extreme (i.e. drier), with the magnitude of change increasing further into the future. This suggests that extreme low flows at this location occur, and will continue to occur, in summer and are expected to respond to projected decrease in summer precipitation. The response for SPIUS is almost the opposite of that for NICOO. At this location, extreme low flows for are projected to become less extreme (i.e. wetter) in the future, which is counterintuitive. This would suggest that extreme low flows in this region of the Nicola basin occur during winter, and will continue to occur in winter, and will respond to projected increases rainfall (due to both increasing precipitation and increasing temperature). However, beyond mid-century the less extreme events ($T < 5$ years) will become drier than the historical period, indicating the development of a mixed low flow regime. The repose for the NICOS is very similar to SPIUS. The response of the COLDW suggest a mixed low flow regime that will evolve to a predominantly summer low flow regime further into the future. Early in the coming century low flows for $T > 5$ years are projected to become less extreme, however, as time progresses, the more frequent low flow events ($T < 10$ for mid-century and $T < 50$ for end-century) are expected to become drier.

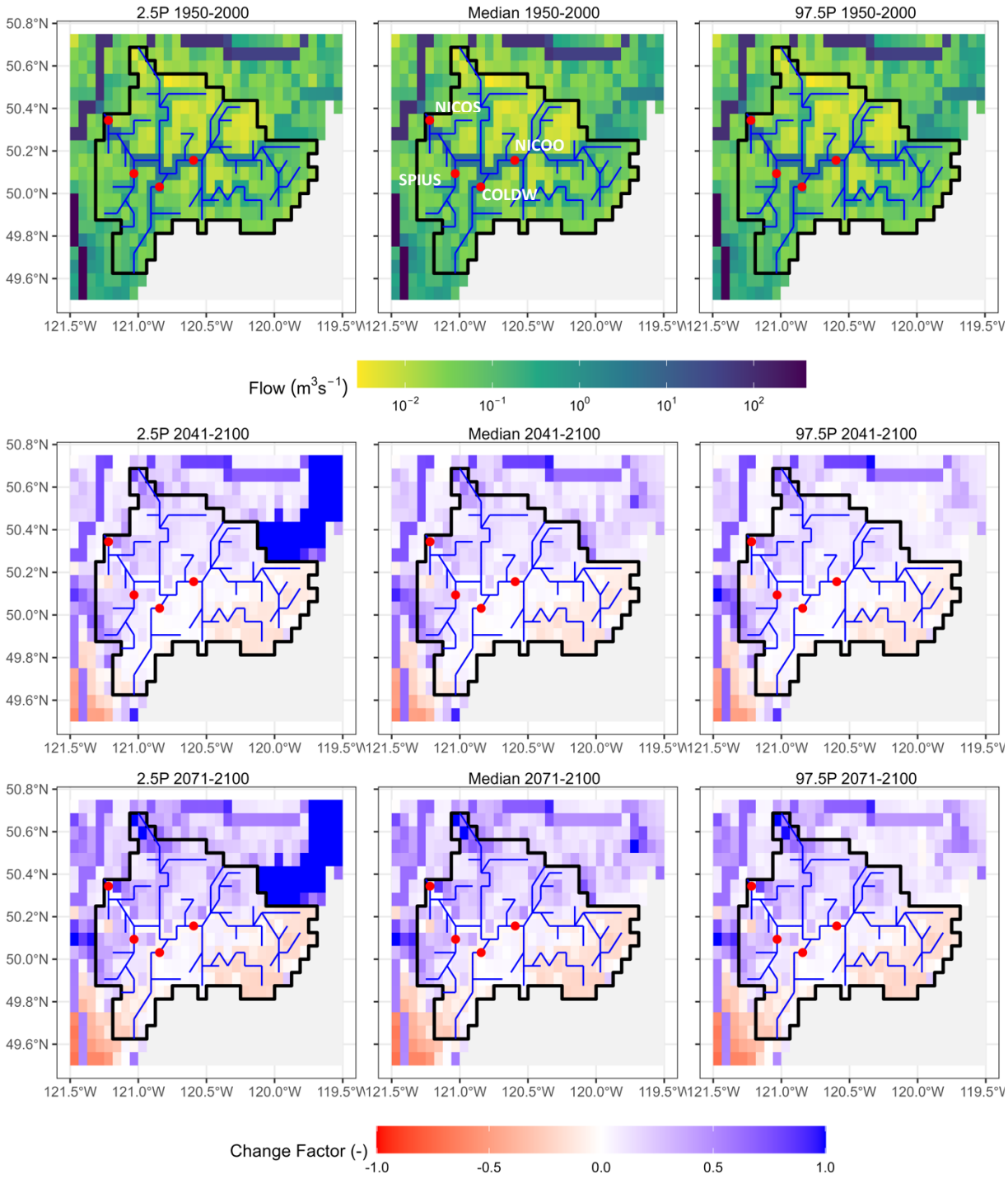

Figure 9. The absolute value (m^3/s) of annual minimum 7-day mean flow during the 1951-2000 period (plotted on logarithmic scale) (top) and change factors in annual minimum 7-day mean flow magnitude for 10-yr return period events for 2041-2070 and 2071-2100 versus the baseline period (1951-2000). Results are shown for 2.5th percentile (left), median (middle) and 97.5th percentile (right). Results are shown for the Nicola basin.

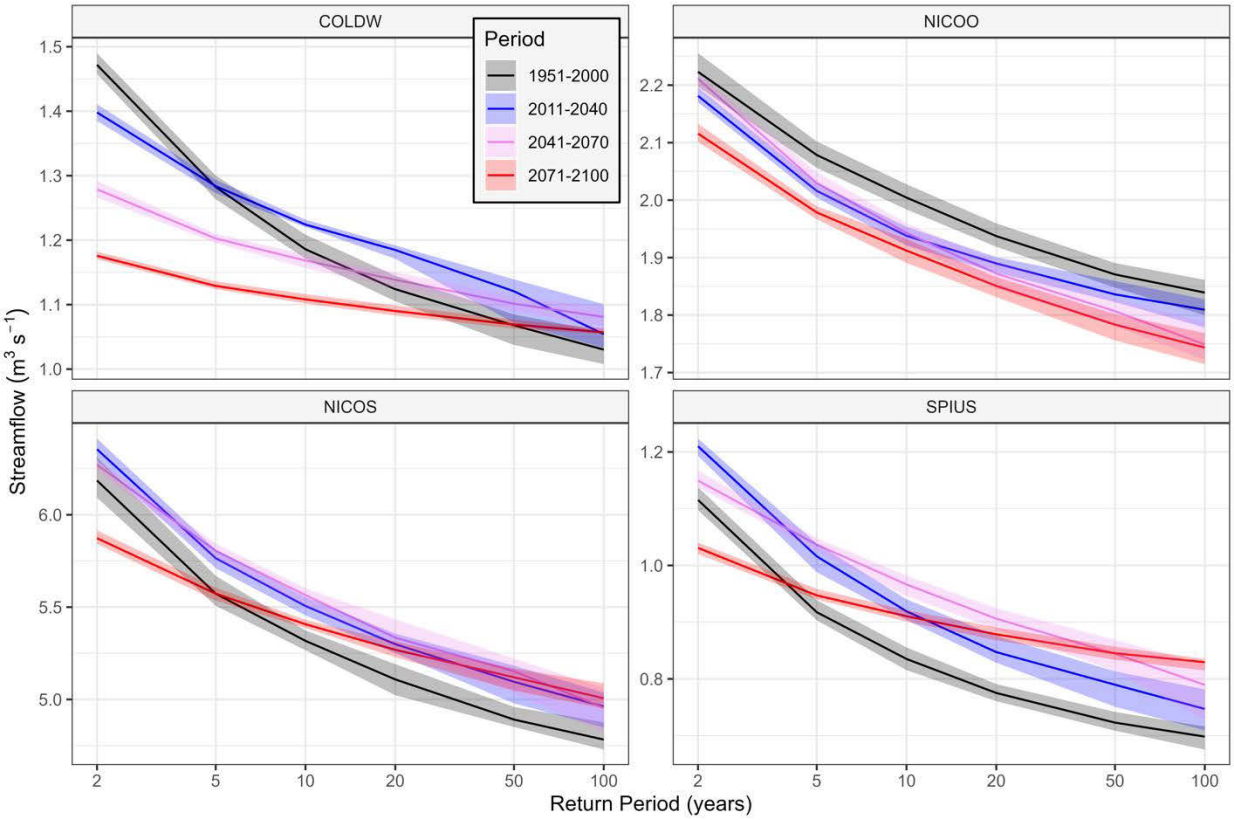

Figure 10. Low flow frequency curves for four periods at four locations in the Nicola basin: the Coldwater River at Merritt (COLDW), the Nicola River at the outlet of Nicol Lake (NICOO), the Nicola River near Spences Bridge (NICOS), and Spius Creek near Canford (SPIUS).

2.6 Discussion, Uncertainties and Limitations

2.6.1 Uncertainty in the Modelling Chain

Hydrologic projections are subject to uncertainties arising from the need to make choices regarding the various components of the modelling chain (Figure 5). And results are affected by the choice of emissions trajectory, GCM, downscaling approach, hydrology model structure, model calibration approach, methodology for estimating flood quantiles, and the various data sets used for model parametrization and calibration (Chegwidden et al., 2019; Curry et al., 2019; Queen et al., 2021; Schnorbus and Curry, 2019; Schoeneberg and Schnorbus, 2020). When using results from this study one should take careful consideration of the uncertainties and limitations inherent in the various modelling choices, which we discuss in the remainder of this section.

The actual trajectory of greenhouse gas emission that will occur during the 21st century is unknown. This uncertainty is addressed by using scenarios to capture a range of plausible emissions trajectories, which have been represented by the Representative Concentration Pathways (RCPs), of which there are four, RCP 2.6, 4.5, 6.0 and 8.5 (where the numbers refer to their peak radiative forcing at the end of the 21st century in W/m^2). The CanESM2 Large Ensemble is based solely on RCP 8.5, which has the highest radiative forcing of the four RCPs available.

Modelling the global climate response to radiative forcing includes two sources of uncertainty which can affect the possible range in future extremes. Differences between GCM structure (model numerics, resolution, process representation, parametrization, etc.) results in a range of climate responses to a given radiative forcing, resulting in so-called structural uncertainty. A second source of uncertainty is internal or natural climate variability, which refers to climate variations over time resulting from natural causes. We generally concern ourselves with unforced variations, which are internally generated redistributions of energy within the system that occur without changes in external factors, such as manifested by the El Niño/ Southern Oscillation. By only using ensemble results for the CanESM2 model, the spread of different runs can be attributed to internal variability only (Dai and Bloecker, 2019; Mahmoudi et al., 2021) and thus the generated hydrologic ensemble does not address GCM structural uncertainty.

The native-scale outputs of climate models and climate model output is of too coarse a resolution for most hydrologic applications and must be downscaled, often using statistical methods. Users must choose from a wide number of algorithms and target data sets, where the representation of precipitation and hydrologic extremes can be sensitive to the choice of downscaling method (Gutmann et al., 2014; Werner and Cannon, 2016). Users should also be aware of the limitations that result from assumptions that need to be made in the statistical downscaling process. Like all statistical downscaling approaches, BCCAQv2 assumes that the quantile-mapping relationships that are established based on an observed target dataset and climate model simulated data for the historical training period continue to be valid in a climate change context. In addition, errors in the chosen target data set may introduce artefacts into the downscaling process.

The modelling chain used in this study is only designed to specifically address uncertainty due to climate variability, as it utilizes only a single emissions trajectory, GCM, downscaling scheme (trained to a single target data set), hydrology model (with one attempt at calibration) and routing model. However, recent research concludes that the choice of emissions scenario and GCM structural uncertainty tends to provide the largest source of uncertainty in hydrologic projections, and that the remaining sources of uncertainty are relatively small in comparison (Chegwidden et al., 2019; Hattermann et al., 2018; Her et al., 2019; Queen et al., 2021; Schoeneberg and Schnorbus, 2020; Sharma et al., 2018). In this context, it is instructive to understand how results from the CanESM2-based hydrologic projections compare to those produced by a larger set of GCMs. In a pilot study Schoeneberg et al. (2021) explored this issue by comparing annual maximum peak flows generated from the CanESM2 large ensemble (45 realizations) with those simulated using the PCIC6 ensemble, which is composed of six GCMs, some with multiple runs, for total of 15 ensemble members. Comparison were conducted using annual maximum streamflow at the outlets of a number of different sub-basins. It was determined that the two ensembles strongly agree on the overall direction of change over the coming century. However, the single-model CanESM2 ensemble produces larger increasing trends and smaller (less negative) decreasing trends than the full PCIC6 ensemble. In other words, change factors for the single-model CanEMS2 ensemble tend to be higher than the multi-model PCIC6 ensemble.

With access to sufficiently large sample sizes, one can estimate quantile values directly from the empirical density function that are unbiased. This is the distinct advantage of using the CanESM2 large ensemble as the basis for this work, wherein quantiles for each analysis period are estimated using a sample of 1350 identically distributed peak flow events. A comparison of quantile estimation using both the GEV approach and the direct empirical approach was conducted during the pilot study (Schoeneberg et al., 2021) and confirmed that the parametric GEV approach produces biased estimates, particularly for large events. Nevertheless, the tradeoff for unbiased estimates is that the empirical quantiles have more variance and require a larger confidence interval than the corresponding GEV-based.

2.6.2 Interpreting Results and VIC-GL Limitations

Simulation of annual maximum peak flow events represents a considerably harder challenge as accurate simulation over short times scales is very sensitive to model error and uncertainty in streamflow routing, forcing data (rainfall intensity or melt energy), snow dynamics, and storage dynamics (snow and soil). This is apparent when comparing model performance for specific peak flow metrics (see Appendix A). Differences between observed and simulated peak flow magnitude can be very large, and VIC-GL is only able to match observed peak flow distributions at roughly 50% of the evaluated sites. However, the large uncertainties associated with observations of peak flow events, for which errors can exceed ±30% (Di Baldassarre and Montanari, 2009; Horner et al., 2018), and the fact that many sites are affected by flow regulation, particularly in the Columbia, also affects model calibration and makes the interpretation of evaluation metrics and the objective assessment of model performance very difficult.

The physically based nature of the VIC-GL model lends some confidence that, over most of the combined Columbia, Fraser and Peace study area the simulated peak and low flow changes reflect a physically realistic representation of the hydrologic response to climate change. However, due to the possible lack of representation of important hydrologic processes uniquely affecting runoff on the Interior Plains (see Schnorbus and Ben Alaya, 2023 for discussion), confidence is lower regarding the ability of VIC-GL to correctly represent the impact of climate change on peak flows in this region. One should also be aware that different choices regarding the model structure (level of abstraction, grid resolution, model physics, etc.), model parametrization, and calibration method (including data used to both force and constrain the model) can lead to a wide range in future projections (Chegwidden et al., 2019).

Streamflow simulated by the VICGL-RVIC models only represents natural flows and peak design flow results will not reflect the effects of flow regulation. Consequently, in those portions of the Peace, Fraser and upper Columbia that are affected by reservoir storage and releases, the simulated design flow values and change factors may grossly misrepresent actual conditions. This is particularly the case for the Nechako River below the Skins Lake spillway, the Peace River main stem below Williston Lake, the Columbia River main stem below Kinbasket Lake, and the Kootenay River below Lake Koocanusa. In addition, there are many additional smaller tributaries and basins throughout the domain affected by flow regulation that potential users should be aware of. The Nicola basin and many of its tributaries are also regulated such that the results of the low flow analysis are also subject to unknown error.

For calibration purposes the Columbia, Peace and Fraser basins were divided into sub-basins based on the location of hydrometric sites. This sub-division represents a trade-off between number of calibration sites and available record lengths; longer record lengths (but with fewer sites) include more hydro-climate variability to train the model robustly whereas more sites (with shorter records) allow for a more realistic spatial variation in the model parameters. Calibration and parameter selection was conducted independently on each individual sub-basin. This approach, although computationally efficient, introduces artifacts into the spatial pattern in peak flow change. These calibration artifacts appear as abrupt changes in the sign of the response between grid cells that align with sub-basin boundaries used for model calibration (e.g., Figure 8). These artifacts tend to be more pronounced in the drier parts of the Columbia, Peace and Fraser basins (where absolute changes may be small, but relative changes are large). This reinforces that relative peak flow changes are sensitive to model parameter choices and that parameter selection and calibration set-up introduce noise, notably at the model grid scale.

The resolution of the VIC-GL model also offers challenges in the interpretation of streamflow values, peak flow changes, and low flow changes. Each grid cell can only have a single flow direction and a single channel, which imposes substantial simplification of the drainage system. This means that as

drainage area decreases the modelled channel network (and resultant streamflow) becomes increasingly more abstract in terms of representing the detailed spatial structure of the drainage network. Also, with increasingly smaller drainage areas, the relative coarseness of the model resolution increases and the ability to accurately represent basin morphology and area (and, hence, runoff volume) degrades.

As mentioned in the previous section, a further limitation is that while the results are based on a large ensemble of climate change simulations, it is an ensemble from a single climate model using a single future forcing prescription. In addition, the climate model simulations have been downscaled using a single downscaling method trained against a particular gridded training dataset. Consequently, the full range of uncertainty in the derived design flows and change factors is certainly wider than that considered in this work. Nevertheless, despite the limitations, the information on projected design flow changes that is obtained should help to illuminate the potential impacts of climate change on design flow changes We reiterate that the physically based nature of the model does provide some level of confidence that it is responding in a plausible way to changes in the water balance and accompanying hydrological regime to projected climate changes. It would not perhaps be prudent for an engineer to alter design values from current values based directly on these results, but due diligence presumably imposes the requirement to think seriously about the implications of projected design value changes, even if uncertain.

3 Future Initiatives

3.1 Expansion of Extreme Flow Analysis

3.1.1 Expansion of Low Flow Study

Considering the promising results from the pilot study, we offer some options for continuation of the low flow study for future consideration. Future work could see continued development of the demonstrated low flow frequency analysis methodology. Other options to consider are to explore additional metrics to quantify extreme low flow, which generally fall into two categories: a) those based on the flow duration curve (FDC), and b) threshold level methods. Another broad consideration is the need to adopt a seasonal approach to the analysis of extreme low flows.

One option for expanding the low flow study is to continue exploring and developing the low flow frequency analysis (LFFA) presented in Section 2.4. This could involve a) exploring seasonal nQ_y metrics (as discussed below), b) considering alternative *n*-day averaging periods, and c) spatially expanding the nQ_Y analysis to include the Peace, Fraser and upper Columbia.

Expansion of the low flow analysis can also adopt a flow duration curve approach. +The flow duration curve (FDC), which is the complement of the cumulative distribution function, illustrates the percentage of time a given streamflow was equaled or exceeded during a specified period of time. The FDC has seen widespread use in water management (Vogel and Fenessey, 1995), and is also well suited for describing low flow characteristics (Smakhtin, 2001). Various low flow indices can be estimated from 'low-flow' portion of the FDC. Widely used indices include *n*-day discharges exceeded 70-99% of the time. Indices of extreme low flow would, of course, be defined using the higher percentile range such as 90-99% time exceeded.

A drawback of both LFFA and FDC analysis is that they have no regard to the sequence of occurrence of low flow episodes nor do they provide information about the length of continuous periods below a particular flow value of interest. They also give no indication of a possible deficit of flow, which is accumulated during a continuous low-flow event. If metrics or information on the continuous availability of prescribed flow is required (e.g. for fish migration), then metrics based on the "threshold level" method may prove informative. In this approach streamflow drought duration (continuous time below the flow threshold) and deficit volume (total volume below threshold) are derived from streamflow time series using a pre-determined flow threshold (Hisdal et al., 2004; Tallaksen et al., 1997; Van Loon, 2015). One can then conduct an analysis of extreme values of drought duration and deficit, either treating each variable as independent or treating them as jointly occurring.

Abnormally low streamflow can result from several processes, which can vary seasonally in watersheds throughout British Columbia. Historically, many locations can have multiple low flow seasons. A winter low flow season can occur when runoff is delayed because of precipitation being stored as snow. A summer low flow season occurs due to low precipitation, possibly following the cessation of snow melt. There are also many pathways for extreme low flows to change due climate change. In summer low flows may become more extreme due to increasing rainfall deficits, earlier depletion of the snowpack, and loss of glacier melt. In winter, low flows may experience an increase from both increased precipitation and a greater proportion of precipitation occurring as rain. For future consideration it would be valuable to adopt a seasonal approach to low flow, regardless of the type of metric considered.

3.1.2 Vector-based Routing

The purpose of this initiative is to overcome the constraints imposed by the rather coarse spatial resolution of the VIC-GL streamflow projections. As discussed in Section 2.6.2, the output from VIC-GL

occurs on a simplified drainage network that, at small scale, does not accurately represent the fine spatial detail of the actual drainage network. Under this initiative PCIC would downscale runoff from the previously produced VIC-GL large hydrologic ensembles onto a highly detailed vector-based routing network. The resultant streamflow could then be used to support site-specific design flow analysis at the scale of individual stream reaches. Specifically, the broad-scale VIC-GL hydrologic projections would be used to force a routing model that will be deployed at very fine spatial resolution. The streamflow routing would utilize Raven (http://raven.uwaterloo.ca), which is a flexible hydrologic modelling framework that sits atop a robust and extendible software architecture. This initiative would leverage off work currently underway at PCIC that will result in the deployment of a high-resolution vector-based routing network in the Fraser and BC Coast domains. An example deployment is given in Figure 11, which compares the VIC-GL-based routing network to the proposed high-resolution vector-based network in and around Quesnel Lake. The proposed vector-based routing will provide a much more accurate representation of the river and lake network, leading to more faithful representation of streamflow at the scale of individual channels reaches and lakes.

An associated activity would be the building and deployment of a web-based tool to deliver reach-scale design flows. The vector-based nature of the results will require a non-standard data format, such as non-binary tree or a directed acyclic graph (DAG) (stored as a list of nodes and edges), which will require custom implementation.

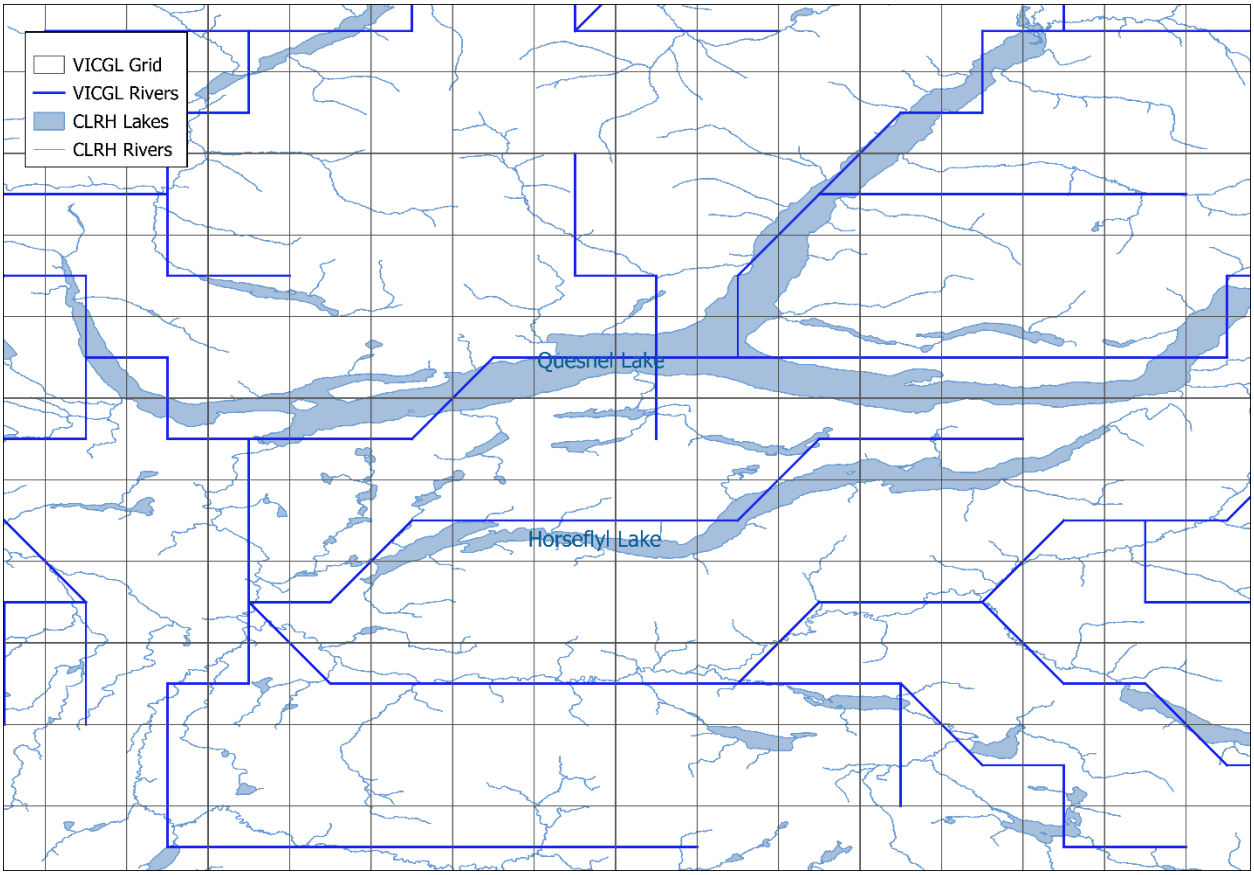

Figure 11. Comparison of the coarse-resolution VIC-GL grid-based routing network versus the proposed (CLRH) high-resolution vector-based routing network. This example depiction is based on a region centered on Quesnel Lake.

3.2 Ministry Staff Training

A wide range of data and tools are available through PCIC, and a continuing objective is to facilitate more effective use of resources among users, including MoTI staff. Thus, there is an opportunity to support the Ministry to achieve more consistent and effective integration of PCIC climate change information into engineering design processes. Key challenges appear to include limited awareness of the full suite of available resources, as well as some uncertainty around how to apply them in practice. This initiative aims to address these challenges by providing training that increases awareness of available tools and builds confidence in their effective use, ultimately supporting broader and more consistent utilization.

We propose introductory training plan that would deliver on the following broad objectives: a) showcase the resources that are available through PCIC for supporting design flow analysis, and b) provide guidance on interpretation and utilization of projected climate and hydrologic changes available from PCIC. The training would take the form of a one-off online workshop targeted at ministry engineering and technical staff. We foresee that, as a secondary goal, such a workshop would also identify potential requirements or opportunities for more tailored training, and foster discussion on additional data or information requirements in support of transportation infrastructure design and planning.

The proposed training assumes that the audience has a basic level of climate literacy (i.e. understand anthropogenic influence on climate and how climate influences hydrology) and would, therefore, focus on more applied concerns regarding utilization of existing climate change information and resources. Note that if the intended audience does not have the prerequisite climate literacy, PCIC does have other training that can address this issue.

The proposed workshop would require no more than a single day and would include the following topics:

- Review of the hydrologic modelling workflow (i.e. data provenance), including:
    - GCM and scenario selection/availability
    - Statistical downscaling
    - VIC-GL model
- Demonstration of climate and hydrology data available from PCIC data portals (https://pacificclimate.org/data), specifically
    - Statistically downscaled climate scenarios
    - Gridded hydrologic model output
    - Station hydrologic model output
- Demonstration of the PCIC Climate Explorer, specifically the ‘Extreme Streamflow’ tool (https://services.pacificclimate.org/pcex/app/#/data/flood/frapce).
- Presentation and discussion on interpretation and utilization of climate change information, with particular attention to:
    - Sources of uncertainty
    - Ensemble selection
    - Scenario selection.

4 Conclusions

PCIC has completed a project to quantify design flood values (2-, 20-, 50-, 100- and 200-year events) for historical and future periods and make them accessible as a gridded product via PCIC's Climate Explorer tool. This work represents a continuation of the work reported by Schoeneberg et al. (2021), Schnorbus and Sun (2022), and Schnorbus and Ben Alaya (2023), and uses a spatial domain that includes the Peace, Fraser and upper Columbia basins (665,000 km^2). Results are provided for every model grid cell in the domain at a spatial resolution of 0.0625°, and design flood values for each grid are based on streamflow routed from the area upstream of the selected cell. Hence, watershed areas range from ~30 km^2 (i.e., the area of a single headwater grid cell) to 665,000 km^2. Flood design values have been provided at a temporal discretization based on 30-year sliding windows centered every 10-years on 2015 (2001-2030), 2025 (2011-2040), 2035 (2021-2050), 2045 (2031-2060), 2055 (2041-2070), 2065 (2051-2080), 2075 (2061-2090), 2085 (2071-2100). This work takes advantage of hydrologic projections produced by PCIC using the VIC-GL hydrology model driven with the CanESM2 50-member large ensemble (CanESM2-LE), which allows for statistically robust estimation of large return-period events. Change factors, ratios of future value to historic value, to represent projected changes in peak flow design values. Change factors can then be applied to scale design values derived directly from streamflow observations.

A pilot study to explore and develop extreme low flow metrics has also been completed. Using low flow frequency analysis (LFFA), a process that is analogous to flood frequency analysis, we have created a prototype set of gridded 7Q_y extreme flow metrics for the Nicola Basin. This work also takes advantage of the CanESM2 50-member large ensemble to produce robust empirical estimates of rare extreme low flow events. To capture potential non-stationarity in the extreme low flow metrics the approach uses the same temporal discretization as used for the peak flow frequency analysis (i.e. 30-year windows centered every 10 years). There are several advantages to the tested approach. It produces metrics that have the same familiar return-period interpretation as is currently employed with extreme high flows, the 7Q_y metrics (particularly the 7Q_{10}) are already widely adopted in many jurisdictions throughout North America, and calculation is straightforward and easily implemented within the workflow designed for the peak flow analysis. Nevertheless, other approaches exist for developing extreme low flow metrics, and these have also been identified for future consideration.

By using the CanESM2 large ensemble global climate simulations, but with only a single emissions trajectory, GCM, downscaling scheme (trained to a single target data set), hydrology model (with one attempt at calibration), and routing model, the modelling chain used in this study is only designed to specifically address uncertainty due to climate variability. Despite being limited to only a single hydrology model, the performance of VIC-GL is such that the simulated peak flow changes generally reflect a physically realistic representation of the hydrologic response to climate change over much of the combined study domain. However, the noted poor model performance in the Interior Plains physiographic region (i.e., Peace River east of the Rocky Mountains) requires caution when interpreting results from this portion of the domain; Appendix A). Results in general can be subject to high uncertainty, and results at the individual grid scale can be subject to considerable noise.

For this project three future initiative have also been identified that would provide improved guidance on the climate change impacts related to extreme flow events. Firstly, the report provides several options for expanding the exploration and development of low flow extreme metrics beyond the scope of the completed pilot study. Secondly, the report describes an opportunity to leverage work currently underway at PCIC that would upgrade the resolution of current and future design flow metrics onto a high-resolution vector network. This higher level of granularity would allow PCIC to deliver design flow results at reach-level detail. Thirdly, the report presents a plan for how PCIC might be able to help train

Ministry staff in the interpretation and utilization of information about projected climate and hydrologic change. We foresee that, as a secondary goal, such training would also identify potential requirements or opportunities for more tailored training, and foster discussion on additional data or information requirements in support of transportation infrastructure design and planning.

5 References

Arora, V.K., Cannon, A.J., 2018. A brief background on climate models: the source of future climate information. In: P. Mukhopadhyaya (ed.), 1st International Conference on New Horizons in Green Civil Engineering (NHICE-01), Victoria, BC, Canada, April 25-27, 2018. ISBN: 978-1-55058-620-6. p. 348-356.

Arora, V.K., Scinocca, J.F., Boer, G.J., Christian, J.R., Denman, K.L., Flato, G.M., Kharin, V.V., Lee, W.G., Merryfield, W.J., 2011. Carbon emission limits required to satisfy future representative concentration pathways of greenhouse gases. Geophys. Res. Lett. 38. https://doi.org/10.1029/2010GL046270

Baret, F., Weiss, M., Lacaze, R., Camacho, F., Makhmara, H., Pacholcyzk, P., Smets, B., 2013. GEOV1: LAI and FAPAR essential climate variables and FCOVER global time series capitalizing over existing products. Part1: Principles of development and production. Remote Sens. Environ. 137, 299–309. https://doi.org/10.1016/j.rse.2012.12.027

BCMoTI, 2019. Technical Circular T-04/19. Resilient Infrastructure Engineering Design - Adaptation to the Impacts of Climate Change and Weather Extremes.

BCMoTI, Nodelcorp Consulting Inc., Pacific Climate Impacts Consortium, 2014. Review and Analysis of Climate Change Vulnerability Assessments of Canadian Water Management and Drainage Infrastructure. Revision 2.

Camacho, F., Cernicharo, J., Lacaze, R., Baret, F., Weiss, M., 2013. GEOV1: LAI, FAPAR essential climate variables and FCOVER global time series capitalizing over existing products. Part 2: Validation and intercomparison with reference products. Remote Sens. Environ. 137, 310–329. https://doi.org/10.1016/j.rse.2013.02.030

Cannon, A.J., Jeong, D.I., Zhang, X., Zwiers, F.W., 2020. , in: Chapter 2 in CLIMATE-RESILIENT BUILDINGS & CORE PUBLIC INFRASTRUCTURE: An Assessment of the Impact of Climate Change on Climatic Design Data in Canada. Environment and Climate Change Canada. DRAFT. Victoria, BC, Canada, p. 15 pp.

Cannon, A.J., Sobie, S.R., Murdock, T.Q., 2015. Bias Correction of GCM Precipitation by Quantile Mapping: How Well Do Methods Preserve Changes in Quantiles and Extremes? J. Clim. 28, 6938–6959. https://doi.org/10.1175/JCLI-D-14-00754.1

Chegwidden, O.S., Nijssen, B., Rupp, D.E., Arnold, J.R., Clark, M.P., Hamman, J.J., Kao, S.-C., Mao, Y., Mizukami, N., Mote, P.W., Pan, M., Pytlak, E., 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. Earths Future 7, 623–637. https://doi.org/10.1029/2018EF001047

Cherkauer, K.A., Bowling, L.C., Lettenmaier, D.P., 2003. Variable infiltration capacity cold land process model updates. Glob. Planet. Change 38, 151–159.

Curry, C.L., Islam, S.U., Zwiers, F.W., Déry, S.J., 2019. Atmospheric Rivers Increase Future Flood Risk in Western Canada's Largest Pacific River. Geophys. Res. Lett. 46, 1651–1661. https://doi.org/10.1029/2018GL080720

Dai, A., Bloecker, C.E., 2019. Impacts of internal variability on temperature and precipitation trends in large ensemble simulations by two climate models. Clim. Dyn. 52, 289–306. https://doi.org/10.1007/s00382-018-4132-4

Danielson, J.J., Gesch, D.B., 2011. Global Multi-resolution Terrain Elevation Data 2010 (GMTED2010). U.S. Geological Survey Open-File Report 2011–1073. U.S. Department of the Interior, U.S. Geological Survey, National Geospatial-Intelligence Agency.

Di Baldassarre, G., Montanari, A., 2009. Uncertainty in river discharge observations: a quantitative analysis. Hydrol. Earth Syst. Sci. 13, 913–921. https://doi.org/10.5194/hess-13-913-2009

Fyfe, J.C., Derksen, C., Mudryk, L., Flato, G.M., Santer, B.D., Swart, N.C., Molotch, N.P., Zhang, X., Wan, H., Arora, V.K., Scinocca, J., Jiao, Y., 2017. Large near-term projected snowpack loss over the western United States. Nat. Commun. 8, 14996. https://doi.org/10.1038/ncomms14996

Global Soil Data Task, 2014. Global Soil Data Products CD-ROM Contents (IGBP-DIS). ORNL DAAC, doi:https://doi.org/10.3334/ORNLDAAC/565.

Government of Canada, 2019. The Canadian Earth System Model Large Ensembles [WWW Document]. Opendat Can. URL https://open.canada.ca/data/en/dataset/aa7b6823-fd1e-49ff-a6fb-68076a4a477c (accessed 7.19.19).

Gudmundsson, G.H., Krug, J., Durand, G., L., Favier., Gagliardini, O., 2012. The stability of grounding lines on retrograde slopes. Cryosphere Discuss. 6, 2597–2619.

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

Hattermann, F.F., Vetter, T., Breuer, L., Su, B., Daggupati, P., Donnelly, C., Fekete, B., Flörke, F., Gosling, S.N., Hoffmann, P., Liersch, S., Masaki, Y., Motovilov, Y., Müller, C., Samaniego, L., Stacke, T., Wada, Y., Yang, T., Krysnaova, V., 2018. Sources of uncertainty in hydrological climate impact assessment: a cross-scale study. Environ. Res. Lett. 13, 015006. https://doi.org/10.1088/1748-9326/aa9938

Her, Y., Yoo, S.-H., Cho, J., Hwang, S., Jeong, J., Seong, C., 2019. Uncertainty in hydrological analysis of climate change: multi-parameter vs. multi-GCM ensemble predictions. Sci. Rep. 9, 4974. https://doi.org/10.1038/s41598-019-41334-7

Hiebert, J., Cannon, A., Murdock, T., Sobie, S., Werner, A., 2018. ClimDown: Climate Downscaling in R, Journal of Open Source Software. https://doi.org/10.21105/joss.00360

Hisdal, H., Tallaksen, L.M., Clausen, B., Peters, E., Gustard, A., 2004. Chapter 5 - Hydrological Drought Characteristics, in: Tallaksen, L.M., Lanen, H.A.J. van (Eds.), Hydrologic Drought, Developments in Water Science. Elsevier, Boston, pp. 139–197.

Horner, I., Renard, B., Le Coz, J., Branger, F., McMillan, H.K., Pierrefeu, G., 2018. Impact of Stage Measurement Errors on Streamflow Uncertainty. Water Resour. Res. 54, 1952–1976. https://doi.org/10.1002/2017WR022039

Hunter, R.D., Meentemeyer, R.K., 2005. Climatologically Aided Mapping of Daily Precipitation and Temperature. J. Appl. Meteorol. 44, 1501–1510. https://doi.org/10.1175/JAM2295.1

Kirchmeier-Young, M.C., Zwiers, F.W., Gillett, N.P., 2017. Attribution of Extreme Events in Arctic Sea Ice Extent. J. Clim. 30, 553–571. https://doi.org/10.1175/JCLI-D-16-0412.1

Kushner, P.J., Mudryk, L.R., Merryfield, W., Ambadan, J.T., Berg, A., Bichet, A., Brown, R., Derksen, C., Déry, S.J., Dirkson, A., Flato, G., Fletcher, C.G., Fyfe, J.C., Gillett, N., Haas, C., Howell, S., Laliberté, F., McCusker, K., Sigmond, M., Sospedra-Alfonso, R., Tandon, N.F., Thackeray, C., Tremblay, B., Zwiers, F.W., 2018. Canadian snow and sea ice: assessment of snow, sea ice, and related climate processes in Canada's Earth system model and climate-prediction system. The Cryosphere 12, 1137–1156. https://doi.org/10.5194/tc-12-1137-2018

Liang, X., Lettenmaier, D.P., Wood, E.F., Burges, S.J., 1994. A simple hydrologically based model of land-surface water and energy fluxes for general-circulation models. J. Geophys. Res.-Atmospheres 99, 14415–14428. https://doi.org/10.1029/94JD00483

Liang, X., Wood, E.F., Lettenmaier, D.P., 1996. Surface soil moisture parameterization of the VIC-2L model: Evaluation and modification. Glob. Planet. Change, Soil Moisture Simulation 13, 195–206. https://doi.org/10.1016/0921-8181(95)00046-1

Lohmann, D., Raschke, E., Nijssen, B., Lettenmaier, D.P., 1998. Regional scale hydrology: I. Formulation of the VIC-2L model coupled to a routing model. Hydrol. Sci. J. 43, 131–141. https://doi.org/10.1080/02626669809492107

Mahmoudi, M.H., Najafi, M.R., Singh, H., Schnorbus, M., 2021. Spatial and temporal changes in climate extremes over northwestern North America: the influence of internal climate variability and external forcing. Clim. Change 165, 14. https://doi.org/10.1007/s10584-021-03037-9

Maurer, E.P., Hidalgo, H.G., Das, T., Dettinger, M.D., Cayan, D.R., 2010. The utility of daily large-scale climate data in the assessment of climate change impacts on daily streamflow in California. Hydrol Earth Syst Sci 14, 1125–1138. https://doi.org/10.5194/hess-14-1125-2010

Natural Resources Canada/ The Canada Centre for Mapping and Earth Observation (NRCan/CCMEO), United States Geological Survey (USGS), Insituto Nacional de Estadística y Geografía (INEGI), Comisión Nacional para el Conocimiento y Uso de la Biodiversidad (CONABIO), and Comisión Nacional Forestal (CONAFOR), 2013. 2013: 2010 North American Land Cover at 250 m spatial resolution.

Ouarda, T.B.M.J., Charron, C., St-Hilaire, A., 2008. Statistical Models and the Estimation of Low Flows. Can. Water Resour. J. Rev. Can. Ressour. Hydr. 33, 195–206. https://doi.org/10.4296/cwrj3302195

Pfeffer, W.T., Arendt, A.A., Bliss, A., Bolch, T., Cogley, J.G., Gardner, A.S., Hagen, J.-O., Hock, R., Kaser, G., Kienholz, C., Miles, E.S., Moholdt, G., Mölg, N., Paul, F., Radić, V., Rastner, P., Raup, B.H., Rich, J., Sharp, M.J., Consortium, T.R., 2014. The Randolph Glacier Inventory: a globally complete inventory of glaciers. J. Glaciol. 60, 537–552. https://doi.org/10.3189/2014JoG13J176

Pyrce, R., 2004. Hydrological Low Flow Indices and their Uses (WSC Report No. No. 04-2004). Watershed Science Centre, Peterborough, Ontario.

Queen, L.E., Mote, P.W., Rupp, D.E., Chegwidden, O., Nijssen, B., 2021. Ubiquitous increases in flood magnitude in the Columbia River basin under climate change. Hydrol. Earth Syst. Sci. 25, 257–272. https://doi.org/10.5194/hess-25-257-2021

R Core Team, 2019. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.

Schnorbus, M., 2018. VIC Glacier: Description of VIC model changes and updates (PCIC Internal Report). Pacific Climate Impacts Consortium, Victoria, BC.

Schnorbus, M., 2017. VICGL Model Calibration (PCIC Internal Report). Pacific Climate Impacts Consortium, Victoria, BC.

Schnorbus, M.A., 2021. Model Validation Report for the Fraser River Watershed. Pacific Climate Impacts Consortium, University of Victoria, Victoria, BC.

Schnorbus, M.A., Ben Alaya, M.A., 2023. Future Design Flood Values in the Fraser and Peace River Basins Using the CanESM2-LE: Final Report to “Revision and Expansion of Extreme Streamflow Design Value Projection Online Tool". Pacific Climate Impacts Consortium, Victoria, B.C.

Schnorbus, M.A., Curry, C.L., 2019. Climate Change Scenario Modelling for the Fraser River Watershed Phase 2 - Final Report. Pacific Climate Impacts Consortium, University of Victoria.

Schnorbus, M.A., Sun, Q., 2022. Future Design Flood Values in the Fraser River Basin Using the CanESM2-LE: Final Report to "Revision and Expansion of Extreme Streamflow Design Value Projection Online Tool ". Pacific Climate Impacts Consortium, University of Victoria, Victoria, B.C.

Schoeneberg, A.T., Schnorbus, M.A., 2020. Exploring the Strength and Limitations of PCIC’s CMIP5 Hydrologic Scenarios. Pacific Climate Impacts Consortium, University of Victoria, Victoria, BC, Canada.

Schoeneberg, A.T., Sun, Q., Schnorbus, M.A., 2021. Future Design Flood Values in the Upper Fraser River Basin using the CanESM2-LE: Final Report to “Pilot Study for Development of Streamflow Design Value Projections and Prototype Online Tool.” Pacific Climate Impacts Consortium, University of Victoria, Victoria, BC.

Sharma, T., Vittal, H., Chhabra, S., Salvi, K., Ghosh, S., Karmakar, S., 2018. Understanding the cascade of GCM and downscaling uncertainties in hydro-climatic projections over India. Int. J. Climatol. 38, e178–e190. https://doi.org/10.1002/joc.5361

Simard, M., Pinto, N., Fisher, J.B., Baccini, A., 2011. Mapping forest canopy height globally with spaceborne lidar. J. Geophys. Res. Biogeosciences 116. https://doi.org/10.1029/2011JG001708

Smakhtin, V.U., 2001. Low flow hydrology: a review. J. Hydrol. 240, 147–186. https://doi.org/10.1016/S0022-1694(00)00340-1

Sobie, S.R., Murdock, T.Q., 2017. High-Resolution Statistical Downscaling in Southwestern British Columbia. J. Appl. Meteorol. Climatol. 56, 1625–1641. https://doi.org/10.1175/JAMC-D-16-0287.1

Stedinger, J.R., Vogel, R.M., Foufoula-Georgiou, E., 1993. Frequency analysis of extreme events, in: Maidment, D.R. (Ed.), Handbook of Hydrology. McGraw-Hill, United States, p. 1424.

Tallaksen, L.M., Madsen, H., Clausen, B., 1997. On the definition and modelling of streamflow drought duration and deficit volume. Hydrol. Sci. J. 42, 15–33. https://doi.org/10.1080/02626669709492003

Tallaksen, L.M., Madsen, H., Hisdal, H., 2004. Chapter 6 - Frequency Analysis, in: Tallaksen, L.M., Lanen, H.A.J. van (Eds.), Hydrologic Drought, Developments in Water Science. Elsevier, pp. 199–271.

Taylor, K.E., Stouffer, R.J., Meehl, G.A., 2011. An Overview of CMIP5 and the Experiment Design. Bull. Am. Meteorol. Soc. 93, 485–498. https://doi.org/10.1175/BAMS-D-11-00094.1

Todini, E., 1996. The ARNO rainfall—runoff model. J. Hydrol. 175, 339–382. https://doi.org/10.1016/S0022-1694(96)80016-3

Van Loon, A.F., 2015. Hydrological drought explained. Wiley Interdiscip. Rev. Water 2, 359–392. https://doi.org/10.1002/wat2.1085

Vogel, R.M., Fenessey, N.M., 1995. Flow duration curves ii: a review of applications in water resources planning. J. Am. Water Resour. Assoc. 31, 1029–1039. https://doi.org/10.1111/j.1752-1688.1995.tb03419.x

Werner, A.T., Cannon, A.J., 2016. Hydrologic extremes – an intercomparison of multiple gridded statistical downscaling methods. Hydrol. Earth Syst. Sci. 20, 1483–1508. https://doi.org/10.5194/hess-20-1483-2016

Werner, A.T., Schnorbus, M.A., Shrestha, R.R., Cannon, A.J., Zwiers, F.W., Dayon, G., Anslow, F., 2019. A long-term, temporally consistent, gridded daily meteorological dataset for northwestern North America. Sci. Data 6, 180299. https://doi.org/10.1038/sdata.2018.299

Appendix A: Model Verification

Method

Given that the purpose of the simulated streamflow is to estimate design flow values verification focuses on comparison based on the cumulative distribution function of annual maximum streamflow. Specifically, using annual maximum peak flow data we construct empirical distribution functions from both observed and simulated data. The performance of the VIC-GL – RVIC model chain in accurately reproduced the distribution of extreme flow events is quantified using metrics that measure the distance between the observed and simulated empirical distribution functions.

Probability Error

If $F_x(\cdot)$ and $F_y(\cdot)$ are the empirical distribution function for the observed and simulated annual maximum flow samples, respectively, then the mean probability error is estimated as

$$MPE = \frac{1}{99}\sum_{k=1}^{99} F_x(q_k) - F_y(q_k)$$

and the mean absolute probability error is

$$MAPE = \frac{1}{99}\sum_{k=1}^{99} \left|F_x(q_k) - F_y(q_k)\right|$$

for the set of quantiles q_k given by

$$q_k = F_x^{-1}(p_k) \;\; for \; (p_k)_1^{99} \; and \; p_k = 0.01k.$$

Essentially q_k are the set of 99 quantiles estimated from the observed empirical distribution function sampled at regular probability intervals of 0.01. *MPE* is analogous to the bias in the simulated distribution function, such that if MPE is less than zero then the simulated annual maximum values will tend to be smaller than the observed annual maximum values (i.e. for a given quantile the simulated probabilities will tend to be higher than the observed probabilities). Values for MPE are in the range (-∞, ∞), with a value of zero indicating perfect performance. *MAPE* is a measure of the total deviation between the simulated and observed distribution functions with values in the range [0,∞), where 0 indicates perfect model performance. The empirical distribution functions and quantiles were estimated in R using the *stats*::*ecdf* and the *stats*::*quantile* functions, respectively.

Kolmogorov-Smirnov Test

The Kolmogorov–Smirnov (KS) statistic also quantifies the distance between the observed and simulated empirical distribution functions. In this case the KS test statistics, *D*, measures the maximum distance between the empirical distribution functions of the two samples as

$$D_{x,y} = \sup_u \left|F_x(u) - F_y(u)\right|.$$

The KS test statistic (*KS.D*) and associated *p*-value (*KS.P*) were estimated in R using the *stats*::*ks.test* function. The null hypothesis that both the observed and simulated data derive from the same population is rejected at a given probability level α if

$$D_{x,y} > \sqrt{-\ln\left(\frac{\alpha}{2}\right) \cdot \frac{1+\frac{m}{n}}{2m}},$$

where *m* and *n* are the respective sizes the two samples. Values for *D* have a range of [0,1] and the smaller the value of *D*, the better the model performance.

Data

The accuracy of the VIC-GL – RVIC model chain is verified at 157 sites located throughout the Peace, Fraser, and upper Columbia basins. These sites correspond to streamflow gauges or water management project sites and derives from several sources:

- Water Survey of Canada (WSC) for streamflow gauging sites in Canada
- US Geological Survey (USGS) for streamflow gauging sites in the United States
- BC Hydro (BCH) project sites in British Columbia (naturalized flow data)
- Bonneville Power Authority (BPA) for project sites in the US portion of the upper Columbia (naturalized flow data)

As indicated above, for sites where streamflow is affected by water management the observations represent naturalized flows. These are observed flows that have been adjusted via modelling to remove the effects of irrigation withdrawals, flow diversions and flow regulation. Fifty such sites are present in the validation data (see **Error! Reference source not found.**). Simulated streamflow data derives from a c alibrated version of VICGL-RVIC (Schnorbus 2017, 2020) driven by the PNWNAmet observation-based meteorological data set (Werner et al. 2019). Hence, these flows represent our 'best estimate' of actual streamflow that occurred at the verification locations during the period 1945 to 2012.

Results

A boxplot summary of the *KS.D*, *MPE* and *MAPE* statistics for all 156 sites is provided in Figure A1 Values of *MPE*, *MAPE*, *KS.D* and *KS.P* for all stations is provided as Table A1. Values for *MPE* range from a low of -0.5 (station MUC) to a high of 0.5 (station BCHWL) (Table A1). The median MPE across all sites is -0.06, suggesting an overall tendency for VICGL to underestimate the magnitude of peak flow events. The best *MPE* performance (i.e. values closest to zero) occur at stations BARRM (-0.002) and STHOM (0.003). Values for *MAPE* range from a low of 0.02 (station CLEAS) to a high of 0.50 (station MUC), with a median of 0.14 (Table A1 and Figure A2). Values for the *KS.D* statistic ranges from a low of 0.07 (BruneauR) to a high of 1.00 (BCHWL), with a median value of 0.29 (Table A1). At a significance level of 10% the null hypothesis that the observed and simulated distribution functions are the same is rejected at eighty-nine stations (i.e. simulated results fail to reproduce the observed distribution of annual maximum peak flows). At a more stringent significance level of 5% this number drops to 80 sites (51% of sites), and at a significance level of 1% the number of stations is 63 (40% of sites) (Table A1).

The values for *MPE* and KS.D are mapped in Figure A2 and Figure A3, respectively. *MAPE* and *KSD* are highly correlated, so mapped results for *MAPE* are omitted. Although values of *MPE* appear to show no clear spatial pattern, the Peace and upper Columbia (excepting the headwaters of the Kootenay and Columbia Rivers) experience predominantly negative values, whereas as the sign of *MPE* is mixed in the Fraser (Figure A2). The spatial pattern in KS.D values is also somewhat mixed, although that portion of the Peace located on the Interior Plains (i.e. leeward of the Rocky Mountains), excepting the Peace River main stem, demonstrates consistently poor performance.

The observed and simulated empirical distribution functions for several sample sites is provided as Figure A4. These sites were chosen to reflect the range in model performance across all sites, from very good

(station CLEAS) to very poor (station MUC) and are intended as an aid in interpreting the values of the various metrics. Clearly, values of *MPE*, *MAPE* and *KS.D* close to zero indicate very good model performance (e.g. CELAS, STHOM, and FLAWE). The magnitude of these values increases as the degree of separation between the observed and simulated empirical distribution curves increases. Negative values of *MPE* indicate that the simulated quantiles are smaller than the corresponding observed quantiles (the simulated distribution curve is shifted 'left' of the observed distribution curve in Figure A4.; e.g. BOWRB, RRNRG, and MUC). Conversely, positive values of *MPE* indicate that the simulate quantiles are generally larger than the corresponding observed quantiles (the simulated distribution curve is shifted 'right of the observed distribution curve in Figure A4.; e.g. KICHO and ARNT7).

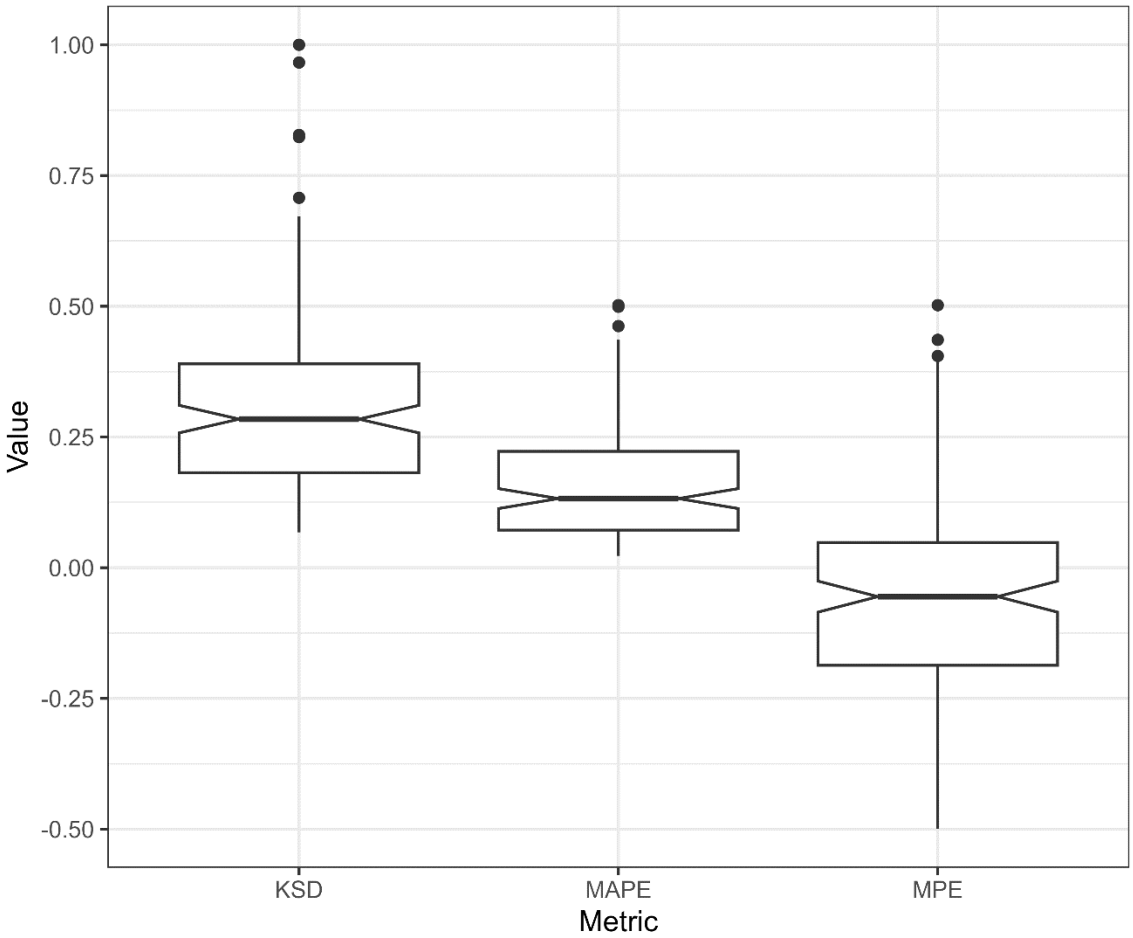

Figure A1. Boxplots of the KS test statistic (KSD), mean absolute probability error (MAPE) and mean probability error (MPE) for all validation sites in the Peace, Fraser, and upper Columbia.

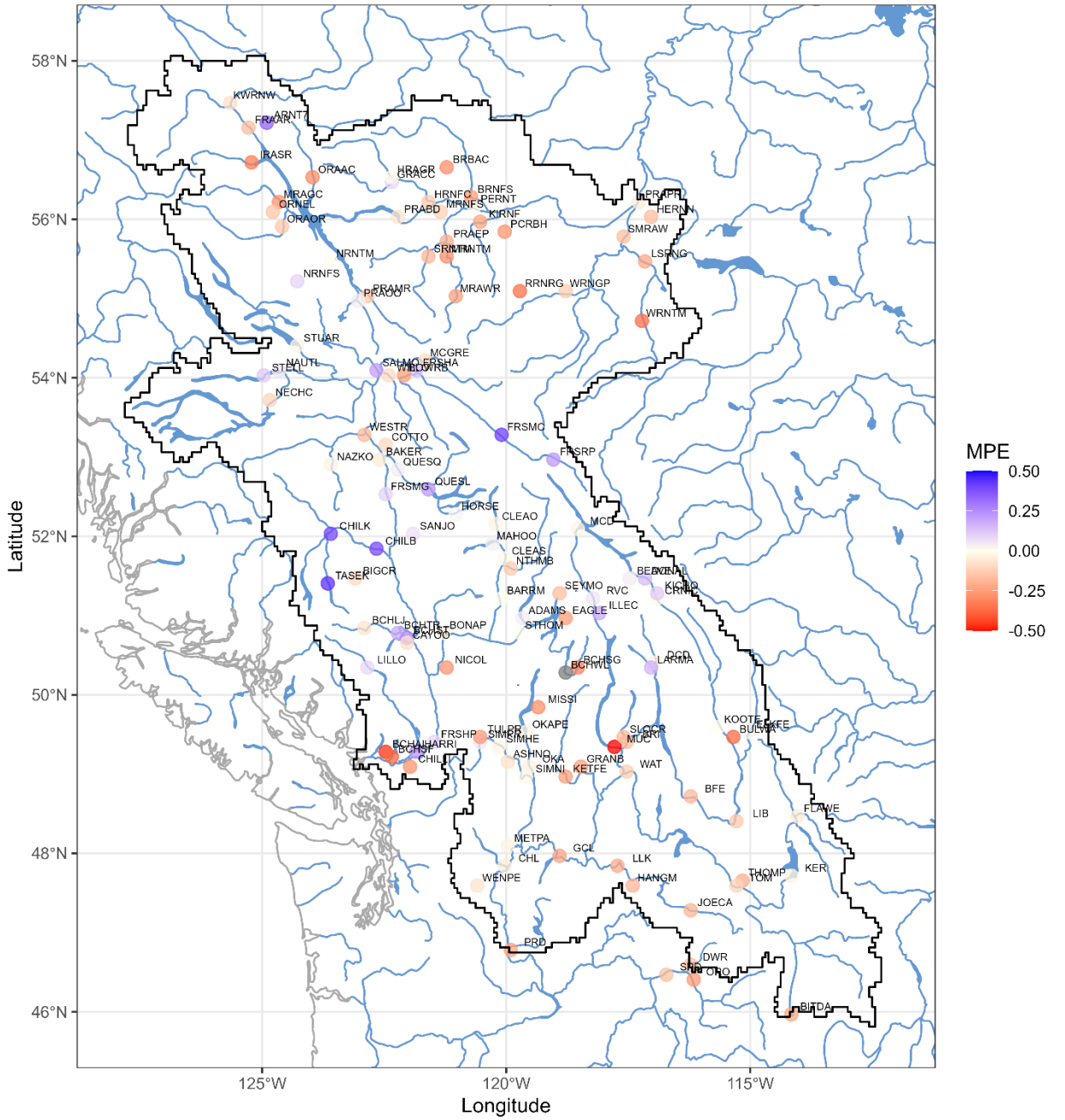

Figure A2. Map of mean probability error (MAPE) for annual maximum peak flow for the 1950 to 2012 validation period.

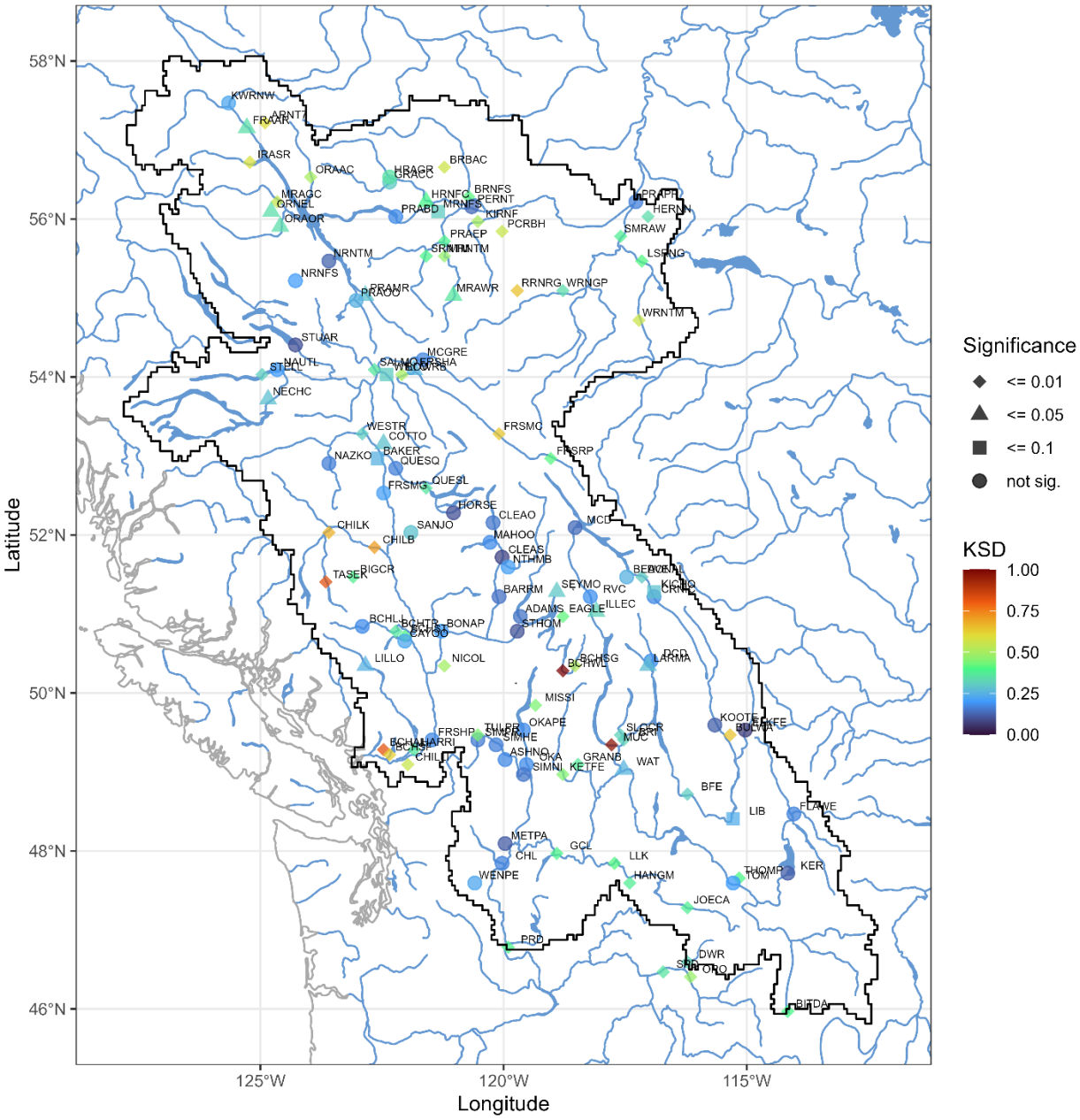

Figure A3. Map of Kolmogorov-Smirnov test statistics and associated probability for annual maximum peak flow for the 1950 to 2012 validation period. The color scale ranges from good performance (blue) to poor performance (red). Values have been organized by *p*-value into four categories: $p \leq 0.01$, $p \leq 0.05$, $p \leq 0.1$, and $p > 0.1$ (not significant).

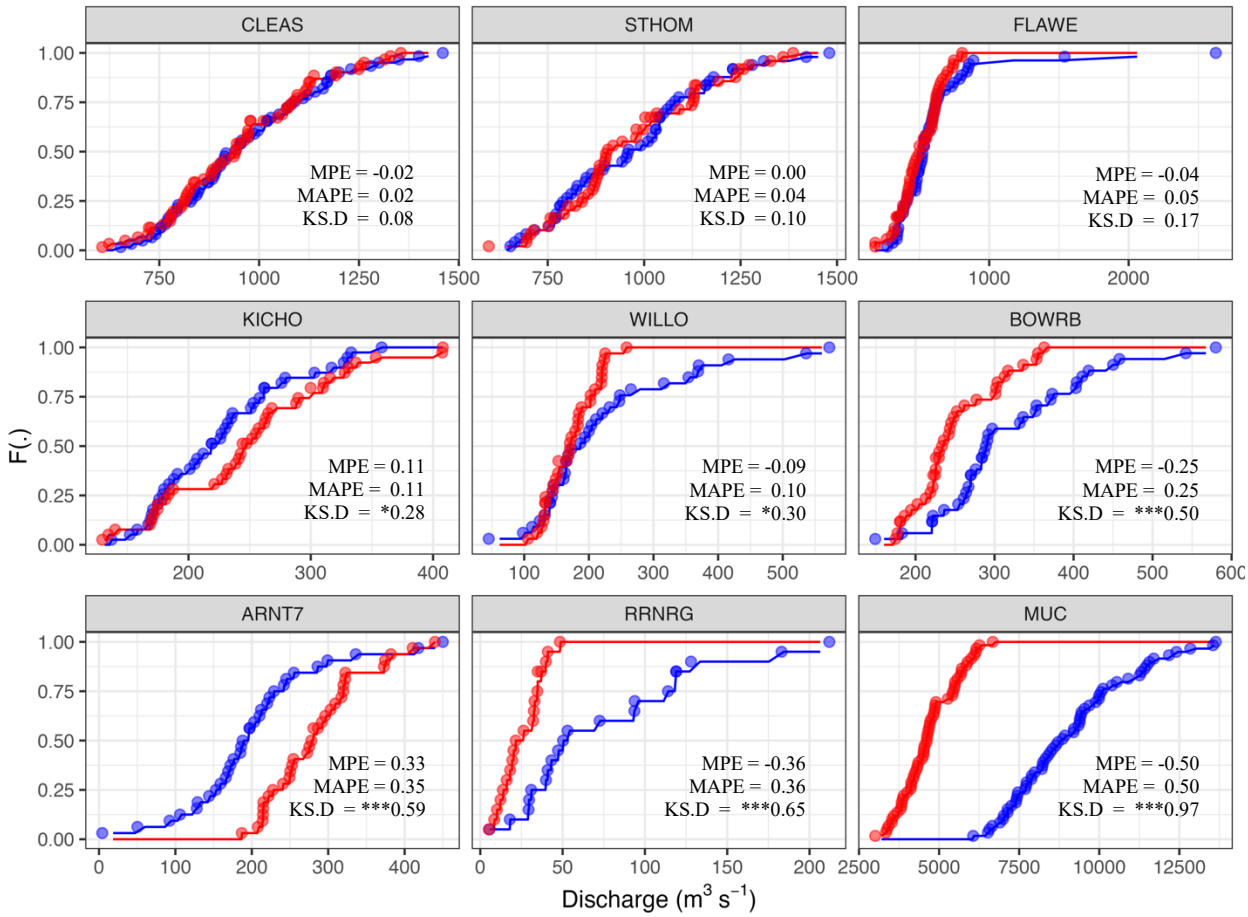

Figure A4. Comparison of observed and simulated empirical distribution functions (*edf*) for eight sample stations. The observed *edf* is in blue, and the simulated *edf* is in red. Dots show the data points from each sample and the lines show the respective edfs over the full data range (observed and simulated). Locations are ordered by increasing *KS.D* statistic (where symbols are *** for $p \leq 0.01$, ** for $p \leq 0.05$, and * for $p \leq 0.1$).

Table A1: Table of annual peak flow performance statistics for validation sites in the Peace, Fraser, and upper Columbia. Metrics are mean probability error (*MPE*), mean absolute probability error (*MAPE*), Kolmogorov-Smirnov test statistics (*KS.D*) and p-value of *KS.D* statistics (*KS.P*).

STN_ID	*MPE*	*MAPE*	*KS.D*	*KS.P*	BASIN	Observation Source
ADAMS	0.040	0.050	0.148	0.518	Fraser	WSC
ALB†	-0.041	0.049	0.153	0.502	Columbia	BPA
ANA†	-0.041	0.045	0.136	0.654	Columbia	BPA
ARNT7	0.334	0.335	0.594	***0.000	Peace	WSC
ASHNO	-0.078	0.078	0.177	0.284	Columbia	WSC
BAKER	-0.052	0.090	0.250	0.099	Fraser	WSC
BARRM	-0.002	0.037	0.145	0.607	Fraser	WSC
BCHAL†	-0.462	0.462	0.824	***0.000	Fraser	BCH
BCHLJ†	-0.071	0.071	0.180	0.396	Fraser	BCH
BCHSF†	-0.375	0.375	0.627	***0.000	Fraser	BCH
BCHSG†	-0.313	0.313	0.514	***0.000	Fraser	BCH
BCHST†	0.254	0.254	0.396	***0.001	Fraser	BCH
BCHTR†	0.183	0.183	0.360	***0.003	Fraser	BCH
BCHWL†	0.502	0.502	1.000	***0.000	Fraser	BCH
BEAVE	0.041	0.088	0.250	0.351	Columbia	WSC
BFE†	-0.186	0.186	0.305	***0.008	Columbia	BPA
BIGCR	-0.098	0.131	0.378	***0.009	Fraser	WSC
BIGI†	-0.058	0.067	0.169	0.367	Columbia	BPA
BITDA	-0.227	0.227	0.396	***0.000	Columbia	USGS
BON†	-0.221	0.221	0.356	***0.001	Columbia	BPA
BONAP	-0.040	0.077	0.205	0.389	Fraser	WSC
BOWRB	-0.247	0.248	0.500	***0.000	Fraser	WSC
BRBAC	-0.288	0.289	0.563	***0.000	Peace	WSC
BRI†	-0.170	0.170	0.322	***0.004	Columbia	BPA
BRN†	0.066	0.071	0.153	0.502	Columbia	BPA
BRNFS	-0.253	0.253	0.412	***0.000	Peace	WSC
BruneauR†	-0.014	0.023	0.068	0.999	Columbia	BPA
BULWA	-0.363	0.363	0.644	***0.000	Columbia	WSC
CAYOO	-0.081	0.099	0.229	0.161	Fraser	WSC
CHILB	0.387	0.387	0.707	***0.000	Fraser	WSC
CHILK	0.394	0.394	0.672	***0.000	Fraser	WSC
CHILLI	-0.274	0.274	0.550	***0.000	Fraser	WSC
CHL†	-0.043	0.082	0.169	0.367	Columbia	BPA
CLEAO	-0.027	0.058	0.152	0.667	Fraser	WSC
CLEAS	-0.015	0.023	0.082	0.987	Fraser	WSC
COTTO	-0.100	0.120	0.283	**0.050	Fraser	WSC
COWKO	-0.143	0.146	0.245	*0.081	Columbia	USGS
CRNIC	-0.016	0.067	0.177	0.284	Columbia	WSC

STN_ID	*MPE*	*MAPE*	*KS.D*	*KS.P*	BASIN	Observation Source
DCD†	-0.030	0.060	0.186	0.258	Columbia	BCH
DEX†	-0.144	0.144	0.254	**0.044	Columbia	BPA
DONAL	0.144	0.144	0.302	***0.006	Columbia	WSC
DWR†	-0.158	0.165	0.322	***0.004	Columbia	BPA
EAGLE	-0.247	0.248	0.417	***0.000	Fraser	WSC
ELKFE	0.020	0.024	0.070	1.000	Columbia	WSC
FLAWE	-0.044	0.051	0.170	0.430	Columbia	WSC
FRAAR	-0.167	0.179	0.343	**0.030	Peace	USGS
FRSHA	0.134	0.134	0.254	**0.041	Fraser	WSC
FRSHP	0.086	0.086	0.176	0.408	Fraser	WSC
FRSMC	0.405	0.405	0.638	***0.000	Fraser	WSC
FRSMG	0.095	0.095	0.216	0.187	Fraser	WSC
FRSRP	0.219	0.220	0.411	***0.000	Fraser	WSC
GCL†	-0.230	0.230	0.390	***0.000	Columbia	BPA
GRACC	0.084	0.108	0.300	0.129	Peace	WSC
GRANB	-0.239	0.240	0.370	***0.003	Columbia	WSC
HANGM	-0.190	0.197	0.377	***0.001	Columbia	WSC
HARRI	0.222	0.222	0.367	***0.001	Fraser	WSC
HERNN	-0.172	0.172	0.340	***0.006	Peace	WSC
HORSE	0.020	0.037	0.098	0.966	Fraser	WSC
HRAGR	-0.011	0.124	0.350	0.171	Peace	WSC
HRNFC	-0.130	0.159	0.393	**0.026	Peace	WSC
ILLEC	0.183	0.183	0.320	**0.012	Columbia	WSC
IRASR	-0.329	0.330	0.583	***0.000	Peace	WSC
JDA†	-0.221	0.221	0.356	***0.001	Columbia	BPA
JFFO	0.059	0.085	0.208	0.203	Columbia	USGS
JOECA	-0.191	0.197	0.377	***0.001	Columbia	USGS
JOHND	0.199	0.199	0.377	***0.001	Columbia	USGS
KER†	-0.028	0.038	0.119	0.805	Columbia	BPA
KETFE	-0.244	0.244	0.453	***0.000	Columbia	USGS
KICHO	0.110	0.113	0.282	*0.089	Columbia	WSC
KIMI†	0.063	0.067	0.186	0.258	Columbia	BPA
KIOW†	0.065	0.093	0.220	0.114	Columbia	BPA
KIRNF	-0.234	0.244	0.500	***0.000	Peace	WSC
KOOTE	-0.003	0.050	0.120	0.863	Columbia	WSC
KWRNW	-0.084	0.099	0.229	0.321	Peace	WSC
LARMA	0.162	0.162	0.276	**0.023	Columbia	WSC
LIB†	-0.142	0.142	0.237	*0.072	Columbia	BPA
LILLO	0.082	0.105	0.250	**0.046	Fraser	WSC
LIM†	-0.045	0.048	0.153	0.502	Columbia	BPA
LLK†	-0.215	0.220	0.407	***0.000	Columbia	BPA

STN_ID	*MPE*	*MAPE*	*KS.D*	*KS.P*	BASIN	Observation Source
LORI†	0.050	0.072	0.237	*0.072	Columbia	BPA
LSRNG	-0.228	0.228	0.407	***0.000	Peace	WSC
MAHOO	0.027	0.076	0.182	0.650	Fraser	WSC
Malheur†	0.061	0.137	0.305	***0.008	Columbia	BPA
MAY†	-0.116	0.116	0.237	*0.072	Columbia	BPA
MCD†	-0.023	0.043	0.136	0.654	Columbia	BCH
MCGRE	-0.091	0.091	0.173	0.413	Fraser	WSC
MER†	-0.099	0.124	0.237	*0.072	Columbia	BPA
METPA	-0.031	0.048	0.113	0.884	Columbia	USGS
MISSI†	-0.287	0.287	0.444	***0.000	Columbia	MISSI
MRAGC	-0.277	0.277	0.541	***0.000	Peace	WSC
MRAWR	-0.241	0.242	0.361	0.017	Peace	WSC
MRNFS	-0.119	0.125	0.303	*0.097	Peace	WSC
MRNTM	-0.278	0.278	0.472	***0.001	Peace	WSC
MUC†	-0.499	0.499	0.966	***0.000	Columbia	BPA
NAUTL	0.057	0.060	0.203	0.173	Fraser	WSC
NAZKO	-0.030	0.049	0.161	0.823	Fraser	WSC
NECHC	-0.107	0.115	0.278	**0.031	Fraser	WSC
NICOL	-0.271	0.271	0.491	***0.000	Fraser	WSC
NRNFS	0.093	0.093	0.209	0.303	Peace	WSC
NRNTM	-0.005	0.039	0.125	0.967	Peace	WSC
NTHMB	-0.133	0.133	0.216	0.186	Fraser	WSC
OKA†	-0.036	0.069	0.182	0.997	Columbia	WSC
OKAPE†	-0.026	0.060	0.182	0.997	Columbia	WSC
ORAAC	-0.292	0.292	0.500	***0.001	Peace	WSC
ORAOR	-0.143	0.143	0.342	**0.023	Peace	WSC
ORNEL	-0.152	0.152	0.344	**0.044	Peace	WSC
ORO†	-0.256	0.256	0.475	***0.000	Columbia	BPA
PCRBH	-0.270	0.283	0.550	***0.000	Peace	WSC
PERNT†	-0.017	0.040	0.133	0.958	Peace	BCH
PLEI†	-0.034	0.041	0.119	0.805	Columbia	BPA
PRABD†	-0.070	0.073	0.179	0.562	Peace	BCH
PRAEP	-0.204	0.204	0.385	***0.001	Peace	WSC
PRAMR	-0.138	0.144	0.283	**0.049	Peace	WSC
PRAOO	0.039	0.081	0.250	0.273	Peace	WSC
PRAPR	-0.047	0.073	0.167	0.808	Peace	WSC
PRD†	-0.231	0.231	0.390	***0.000	Columbia	BPA
PRVO†	-0.073	0.116	0.288	**0.015	Columbia	BPA
QUESL	0.223	0.223	0.361	***0.001	Fraser	WSC
QUESQ	0.040	0.046	0.164	0.388	Fraser	WSC
REXI†	0.096	0.097	0.203	0.175	Columbia	BPA

STN_ID	*MPE*	*MAPE*	*KS.D*	*KS.P*	BASIN	Observation Source
RML†	-0.396	0.396	0.610	***0.000	Columbia	BPA
ROMO†	-0.007	0.046	0.119	0.805	Columbia	BPA
RRNRG	-0.357	0.357	0.650	***0.000	Peace	WSC
RVC†	0.052	0.064	0.186	0.258	Columbia	BCH
SALMO	0.207	0.207	0.362	***0.001	Fraser	WSC
SANJO	0.066	0.101	0.286	0.365	Fraser	WSC
SEYMO	-0.169	0.169	0.295	**0.042	Fraser	WSC
SIMHE	-0.027	0.058	0.167	0.522	Columbia	WSC
SIMNI	-0.017	0.042	0.151	0.580	Columbia	WSC
SIMPR	0.047	0.068	0.159	0.407	Columbia	WSC
SKHI†	0.097	0.097	0.220	0.114	Columbia	BPA
SLOCR	-0.179	0.179	0.317	***0.003	Columbia	WSC
SMRAW	-0.165	0.168	0.379	***0.000	Peace	WSC
SPD†	-0.163	0.163	0.356	***0.001	Columbia	BPA
SRNTM	-0.189	0.189	0.389	***0.008	Peace	WSC
STELL	0.109	0.117	0.311	***0.005	Fraser	WSC
STHOM	0.003	0.045	0.102	0.961	Fraser	WSC
STUAR	-0.023	0.031	0.098	0.929	Fraser	WSC
SVN†	0.062	0.062	0.186	0.258	Columbia	BPA
SWAI†	0.087	0.087	0.203	0.175	Columbia	BPA
TASEK	0.436	0.436	0.828	***0.000	Fraser	WSC
THOMP	-0.207	0.207	0.396	***0.000	Columbia	USGS
TOM†	-0.106	0.106	0.203	0.175	Columbia	BPA
TULPR	-0.256	0.256	0.444	***0.000	Columbia	WSC
UMATI	0.024	0.054	0.189	0.301	Columbia	USGS
UMTW	-0.081	0.098	0.203	0.174	Columbia	USGS
WARMS	0.269	0.272	0.463	***0.000	Columbia	USGS
WAT†	-0.127	0.127	0.254	**0.044	Columbia	BPA
WAV†	-0.194	0.194	0.390	***0.000	Columbia	BPA
WEII†	0.058	0.066	0.136	0.654	Columbia	BPA
WENPE	-0.071	0.077	0.226	0.131	Columbia	USGS
WESTR	-0.177	0.178	0.305	***0.008	Fraser	WSC
WHB†	-0.180	0.180	0.271	**0.026	Columbia	BPA
WILLO	-0.091	0.103	0.303	*0.094	Fraser	WSC
WRNGP	-0.128	0.141	0.340	***0.004	Peace	WSC
WRNTM	-0.359	0.359	0.556	***0.000	Peace	WSC

† Observations based on naturalized discharge
*** Reject null hypothesis at a 1% significance level
** Reject null hypothesis at a 5% significance level
* Reject null hypothesis at a 10% significance level