Future Design Flood Values in the Upper Fraser River Basin Using the CanESM2-LE

Final Report to "Pilot Study for Development of Stream Flow Design Value Projections and Prototype Online Tool"

Agreement Number # 820LA006

March 29, 2021

Arelia T. Schoeneberg (M.Sc.), Dr. Qiaohong Sun (Ph.D.) and

Markus A. Schnorbus (M.ASc.)

Pacific Climate Impacts Consortium

Prepared for: Engineering Services Branch of the Engineering Systems Department of the Highway Services Department, Ministry of Transportation and Infrastructure, Government of British Columbia.

Table of Contents

1 Introduction

Based on the BC Ministry of Transportation and Infrastructure (BCMoTI) 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 (BCMoTI et al., 2014). To address this increased risk, the BCMoTI released directives and guidance for incorporating climate adaptation into engineering designs in its T-04/19 Technical Circular (BCMoTI, 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 and Western University's IDF_CC Tool.

One of the primary risks to infrastructure is the possibility of increased risk of floods, yet the aforementioned tools provide information on changes to precipitation only. While precipitation has intensified in North America (Kirchmeier-Young and Zhang, 2020; Sun et al., 2020) and will become more intense in the future (Li et al., 2020, 2019), this information is not directly translatable to changes in flood risk. BC's contemporary hydrologic regime is primarily snow-dominated and peak flow events, which are generated in spring by snow-melt (Burn and Whitfield, 2016), are not necessarily correlated with precipitation extremes. Even in a future with less snow and increased rainfall, the relationship between rainfall extremes and flooding will be scale dependent and changes in extreme rainfall will be a poor predictor of flooding in relatively large basins. In order to address this gap BCMoTI supported the Pacific Climate Impacts Consortium (PCIC) in a pilot 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. As part of this work, PCIC has also been asked to calculate and supply the Melton Ratio as a gridded product. The Melton Ratio (ratio of watershed relief to the square of the area) is used to characterize a location as being dominated by either flooding (Melton ratios <0.3) or debris flows (Melton ratios >0.6) (Wilford et al., 2004).

This study focuses on the Upper Fraser, a 34,200 km^2 region upstream of Prince George, BC, with primarily snow-dominated watersheds. Results are provided for every model grid cell in the domain, wherein 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 ~32,400 km^2. 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) (Government of Canada, 2019; Kushner et al., 2018). This large ensemble (7500 simulation years), which is based on the RCP8.5 scenario, provides sufficient peak flow samples to allow statistically robust estimation of large return-period events. The use of large-ensembles as a means to develop robust projections of changes to climate extremes and flood frequencies is well established (Alaya et al., 2020; Curry et al., 2019; Fyfe et al., 2017; Gao et al., 2020; Kirchmeier-Young et al., 2017b; Kirchmeier-Young and Zhang, 2020; Li et al., 2020, 2019).

The structure of this report is as follows: in Section 2 we describe the Study Area. Section 3 details the Methodology of the hydrologic model and its parameterization, the CanESM2 Large Ensemble (LE), the downscaling method, the flood frequency analysis and summarize the overall study design. Section 4, Results presents (4.1) projected changes in peak flow by sub-basin, (4.2) projected changes in peak flow for all grid cells, and (4.5) Melton Ratio. Uncertainties, Limitations and Conclusions are presented in Sections 5 and 6.

2 Study Area

The focus of this work is the Upper Fraser River watershed, a 34,200 km^2 region of the Fraser River upstream of Prince George, BC (Figure 1). It was chosen because it represents a large area that has no large lakes and is free of flow regulation. For modelling and presentation purposes, the study region has been divided into eight sub-basins that are delineated based on the locations of eight Water Survey of Canada gauges. The characteristics of these sub-basins are summarized in Table 1 and shown in Figure 1.

The western part of the study area (FRSRP, FRSMC, FRSHA and MCGRE) drains the Rocky and the Columbia Mountains and is bisected by the Rocky Mountain trench. This region is characterized by rugged terrain with high relief (Figure 1, Table 1) and large climatic gradients (Figure 2), where high-elevation areas are dominated by rock and ice (Figure 1, Table 1). The region encompassed by the BOWRB and WILLO sub-basins, which is composed of highlands, uplands and some isolated ridges, has more modest elevations (Figure 1 and Table 1) and precipitation amounts (Figure 2). The most eastern portion of the study area, which includes the SALMO sub-basin and that portion of the study region downstream of MCGRE, FRSHA, WILLO and SALMO, is characterized by lowland and rolling upland terrain with a more sub-continental climate with even precipitation throughout the year. This region has the lowest elevation (Figure 1, Table 1) and driest climate (Figure 2) in the study area.

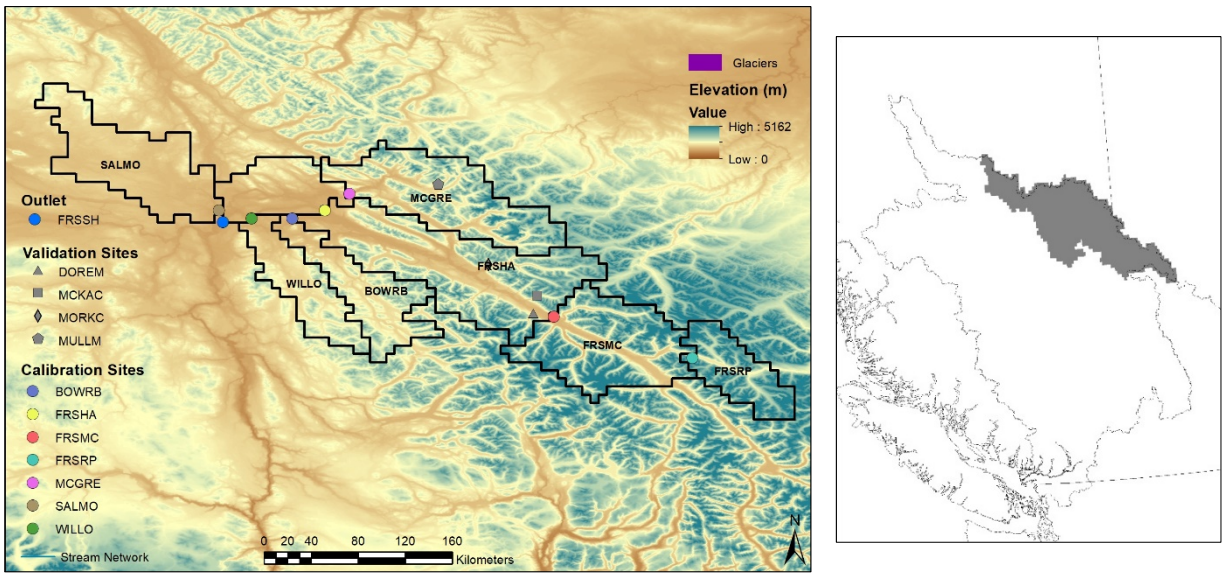

Figure 1 - Study Area for Upper Fraser

Based on PNWNAmet (Werner et al., 2019) the 1961-1990 mean annual minimum temperature for the region ranges from -0.4 to 8.3 °C, annual maximum temperature ranges from -0.8 to 12 °C and annual total precipitation varies from 430 to 2280 mm (Figure 2). Minimum temperatures are below freezing May through October and maximum temperatures are greater than 20°C on only a few days in July/August (Figure 3). Daily precipitation is greatest September through January, with a secondary peak in June/July.

Table 1 - Sub-basins of the Upper Fraser calibrated in the VIC-GL model

Basin Code	Basin Name	WSC ID	Area (km^2)	Glacier Area (km^2)* (%)	Elev. Min (m)	Elev. Max (m)	Elev. Mean (m)
FRSRP	Fraser River at Red Pass	08KA007	2,538	79 (3.1%)	1028	3307	1961
FRSMC	Fraser River at McBride	08KA005	8,025	422 (7.7%)	702	3929	1742
FRSHA	Fraser River at Hansard	08KA004	16,010	45 (0.1%)	604	2885	1318
MCGRE	McGregor River at Lower Canyon	08KB003	5,484	234 (4.3%)	627	3199	1405
WILLO	Willow River above Hay Creek	08KD006	2,844	0	589	1954	1104
BOWRB	Bowron River below Box Canyon	08KD007	3,364	0	625	2368	1190
SALMO	Salmon River near Prince George	08KC001	4,912	0	571	1558	849
FRSSH	Fraser River at Shelley	08KB001	32,400	780 (2.4%)	570	3929	1308

* Glacier area *c*. 2000

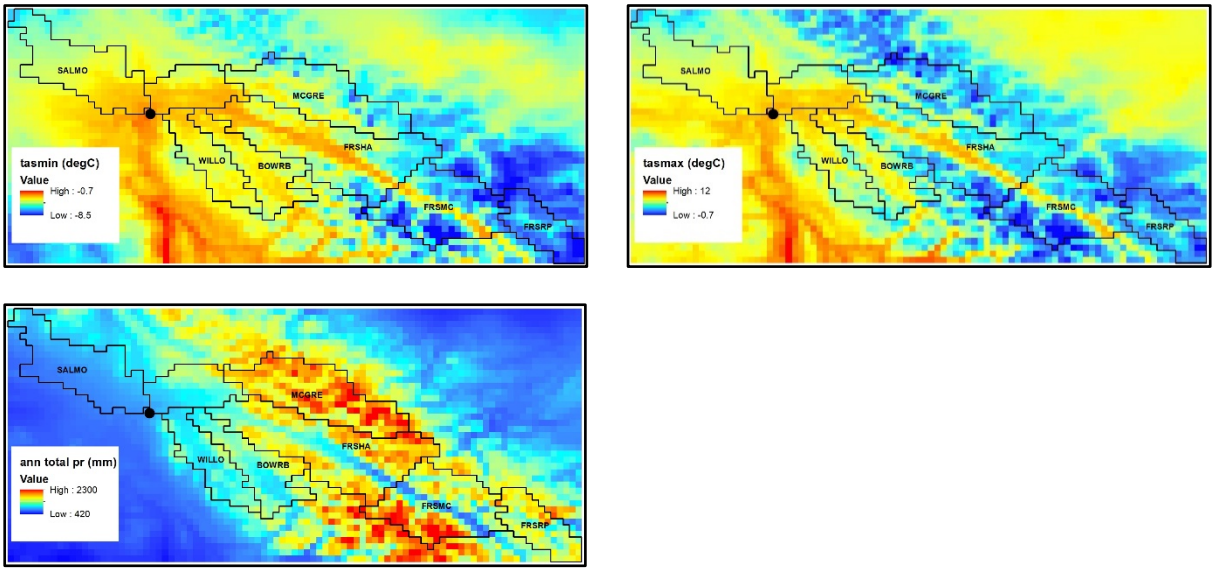

Figure 2 - Annual average minimum and maximum temperature (top left and right, respectively), and average annual total precipitation (bottom left) for 1961-1990 based on PNWNAmet.

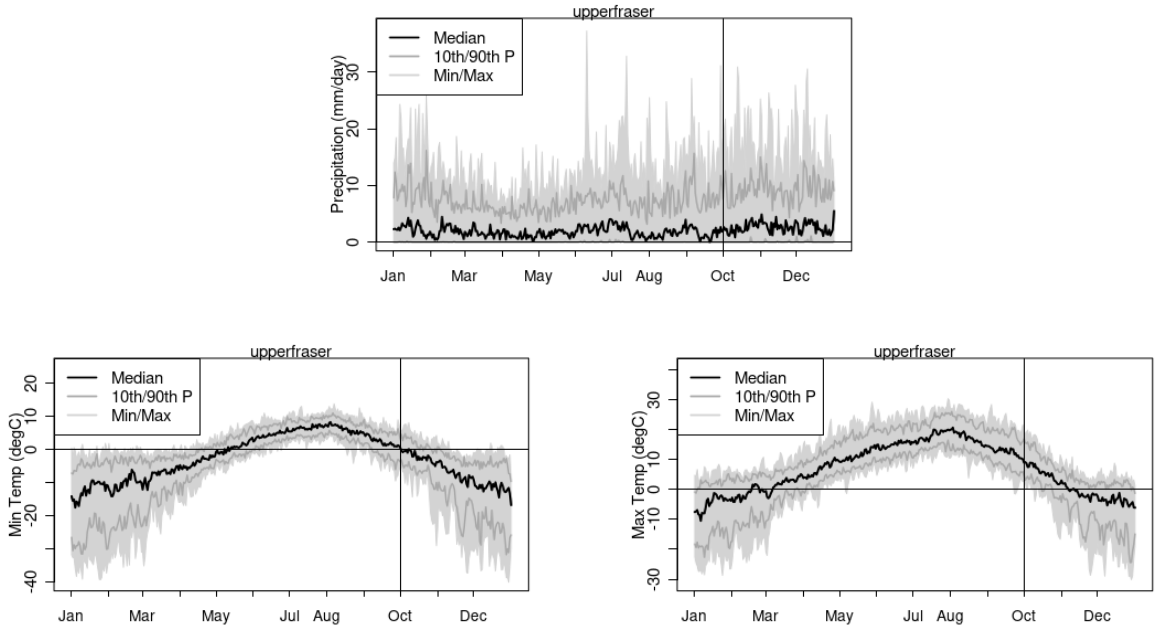

Figure 3 - Daily climatology of precipitation, minimum and maximum temperature in the Upper Fraser over 1961-1990.

Throughout the Upper Fraser the annual hydrograph is dominated by snow and glacier melt and, depending upon the presence of glaciers, can be considered nival (i.e., BOWRB, WILLO and SALMO) or glacial-nival (FRSRP, FRSMC, FRSHA). The daily annual maximum streamflow event is driven by spring snowmelt and generally occurs between the months of May and July (Figure A6). A notable exception is the region largely defined by the MCGRE sub-basin, which occasionally experiences annual maximum events in the fall driven by large rainfall events (Figure A6).

3 Models and Methodology

Projections of future annual maximum streamflow and their flood frequencies 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 4. The implications of the chosen design with respect to modelling uncertainty and limitations will be discussed in Section 5.

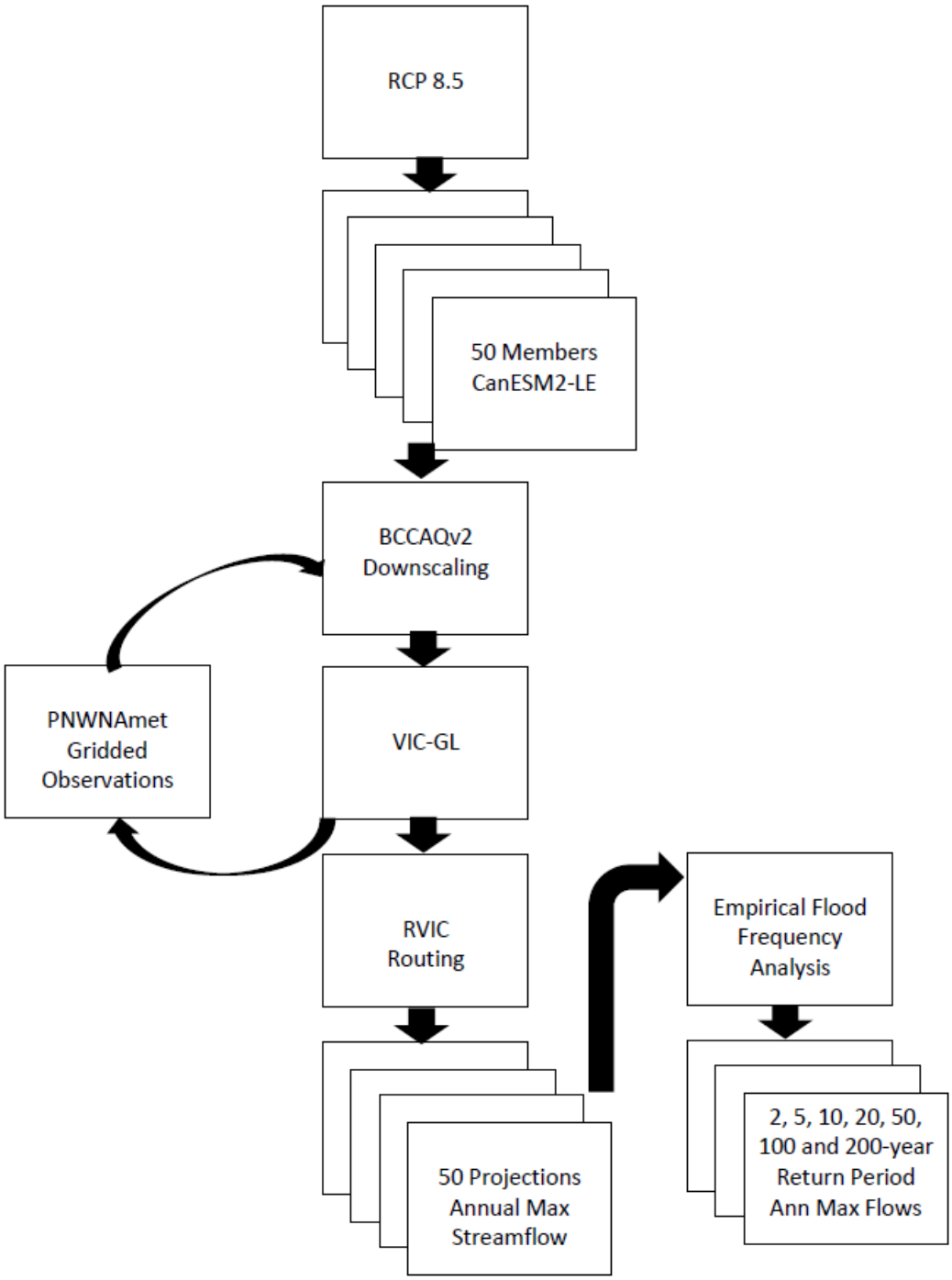

Figure 4 – Modelling workflow used to derive estimates of annual maximum streamflow quantiles from the CanESM2 large ensemble.

3.1 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., 2017a; 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.

3.2 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).

3.3 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).

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 is considered to be too homogeneous in this region and 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). For more details on parameterization of VIC-GL in these basins, please see Schnorbus (in press).

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). Verification of the simulated streamflow is provided in Appendix A, with an in-depth analysis of performance for annual maximum streamflow.

3.4 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). Runoff and baseflow generated in each VIC-GL model cell is collected and routed via a two-step process: 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 the Upper Fraser is shown in *Figure 5*.

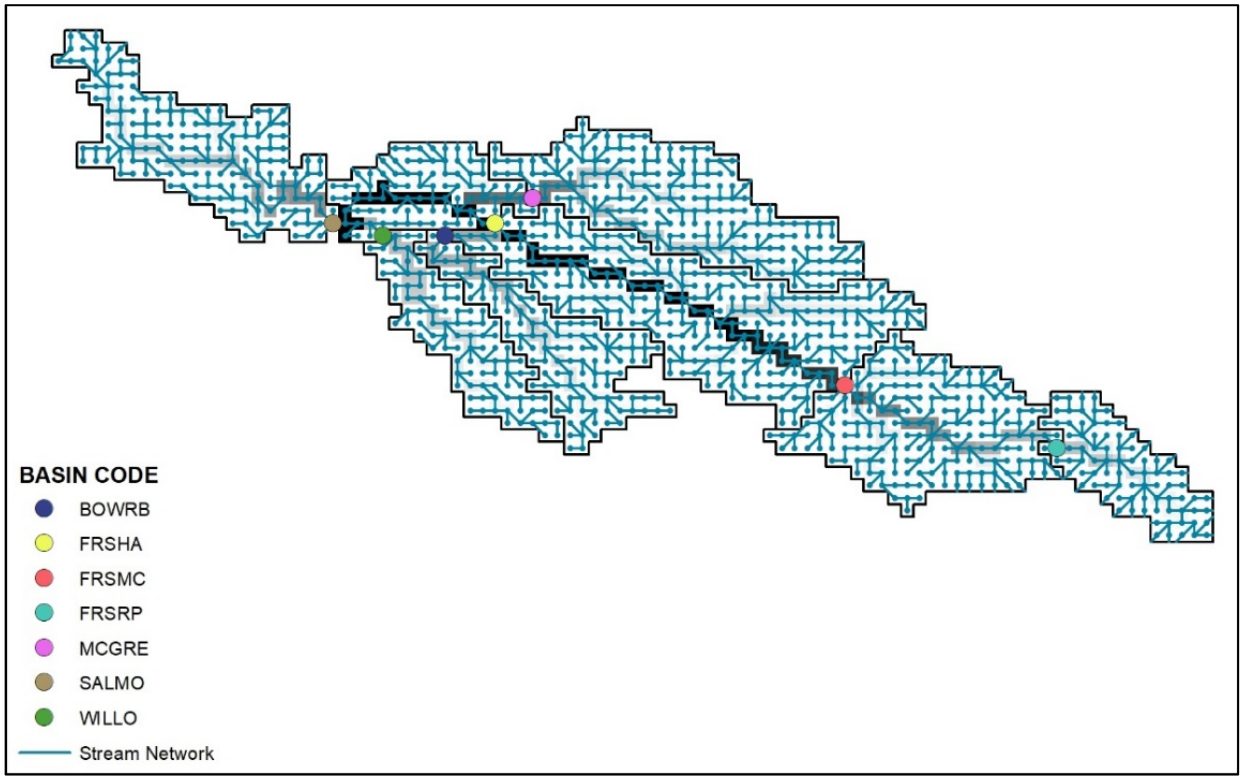

Figure 5 –Upper Fraser flow network with flow accumulation (as number of upstream cells) and sub-basin pour points.

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.

3.5 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. For example, for a given 30-year period, the large ensemble simulation of the Fraser provides 1500 years of streamflow data, and thus 1500 realizations of annual maximum streamflow, that is representative of that period.

Streamflow extremes were analyzed in non-overlapping thirty-years blocks 1961-1990, 2010-2039, 2040-2069, and 2070-2099. Annual maxima were extracted for each year and each model run of the large ensemble, resulting in 1500 values for each thirty-year period (30 years by 50 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 1500 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 1500 samples, which corresponds to the 2-, 5-, 10-, 20-, 50-, 100- and 200-year return levels, respectively. We define the best quantile estimates as those derived from the full sample. The confidence intervals for each quantile are quantified using the following bootstrapping approach: (1) randomly sample with replacement 30 years from each thirty-year block to generate a new sample of size n=1500; (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.

4 Results

4.1 Projected Changes in Peak Flow by Sub-basin

Here we present the flood frequency results for seven locations corresponding to the outlets of the FRSRP, FRSMC, FRSHA, MCGRE, BOWRB, WILLO and SALMO sub-basins as well as the outlet of the entire study area (Upper Fraser). Results are summarized graphically as flood frequency curves (Figure 6) and relative changes for select return periods (Figure 7). Results indicate that the projected flood frequency response to climate change varies by sub-basin. As the projected flood frequency response is largely a function of changing snow dynamics, much of the spatial variation of basin response appears related to elevation. In the higher elevation regions of the FRSRP, FRSMC and MCGRE the flood frequency shows a largely unambiguous response to climate change, wherein flood magnitudes are expected to increase for all return periods for all future periods (Tables B1, B2 and B4). In contrast, the lower elevation basins WILLO, BOWRB and SALMO display a different response to climate change. For these three sites, changes in flood magnitude are projected to be relatively small early in the century. By mid-and end-century large events (100- to 200-year return levels) are projected to increase in magnitude whereas smaller events are projected to decrease in magnitude relative to base period values (Figure 6 and Figure 7; Tables B5, B6 and B7). For the higher-order basins Upper Fraser and FRSHA that contain a combination of high- and low-elevation regions, results are more mixed, with increases in flood magnitude projected for early- and mid-century for most return periods (Figure 6 and Figure 7). By the end of the century however, flood magnitudes are projected to decrease for frequent events ($T \leq$20 years)

below base period values in these basins, whereas large events (100- and 200-year return levels) are still projected to be larger than baseline (Figure 7; Tables B3 and B8).

Examining the timing of annual maximum streamflow gives some indication of its physical driver, such as snowmelt freshet or fall rains. This is explored in Figure 8 and Figure 9, which show for each sub basin a polar plot of annual maximum peak flow magnitude versus occurrence (as the day of year). For clarity, results are plotted for single CanESM2 run only (CanESM2_run1_r1i1p1). In the FRSRP and FRSMC basins, annual maximums shift to earlier in the year and decrease in magnitude with each consecutive future 30-year period (Figure 8). Historically, maximum flows occur between July and August. In 2010-2039 (2020s) and 2040-2069 (2050s) annual maximums are projected to occur between June and July and are higher in magnitude than the historical period. By 2070-2999 (2080s) annual maximum events are projected to occur as early as May 1st, and although higher than historical, not higher than the previous two future periods (2020s and 2050s). Results are similar in FRSHA, except discharge magnitudes are not projected to increase as much as in FRSRP and FRSMC (Figure 8). Large events can occur in October in the projections for MCGRE in the 2080s. The occurrence of peak streamflow events shifts primarily to fall in WILLO and BOWRB by the 2080s (Figure 9). The projected magnitude of annual maximum events progressively decreases and occurrence shifts to earlier in the year in SALMO, tending to occur primarily between March 1st and May 1st, although events still occur in November/December.

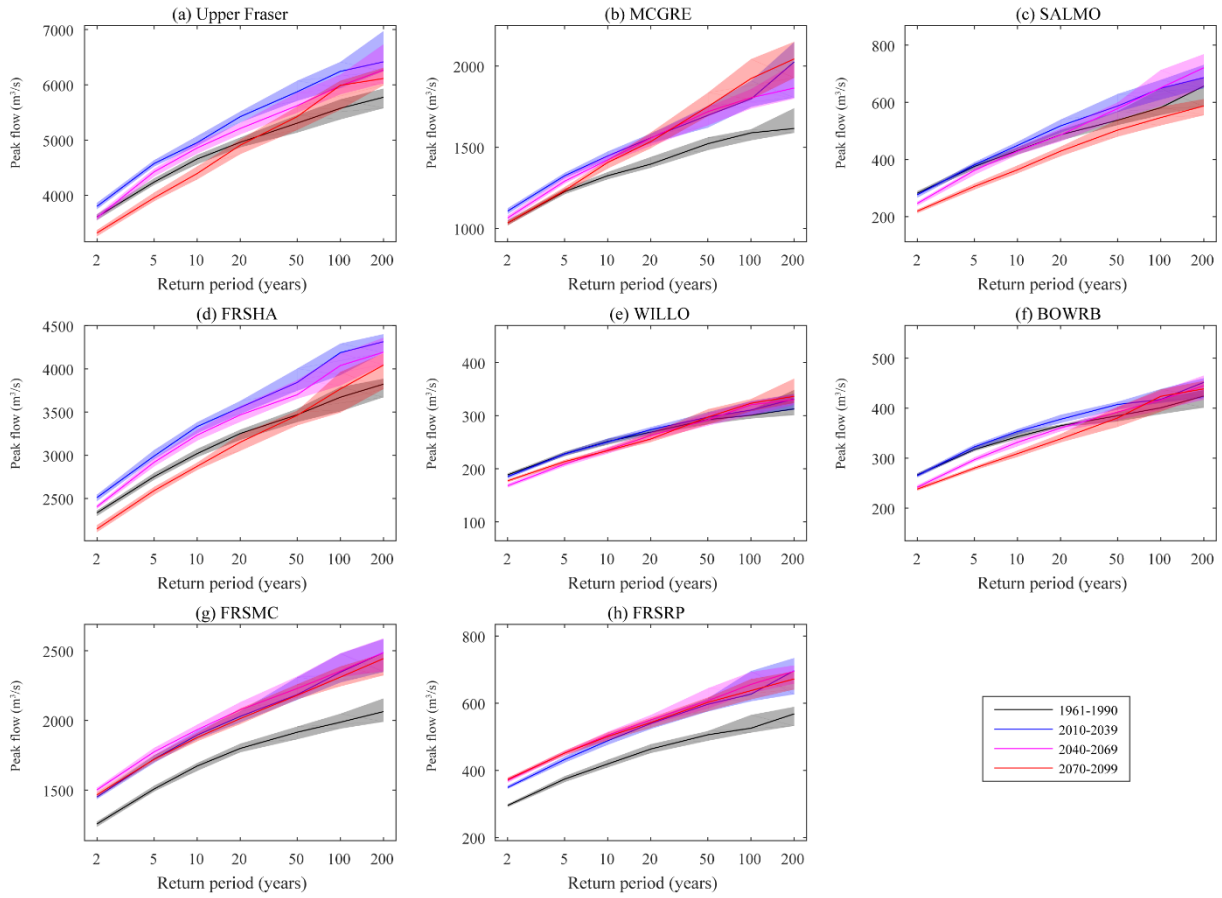

Figure 6 - Flood frequency plot of annual maximum daily streamflow for seven sub-basins and four decades. Best estimates are given by the solid lines and the ribbons show the 95% confidence intervals.

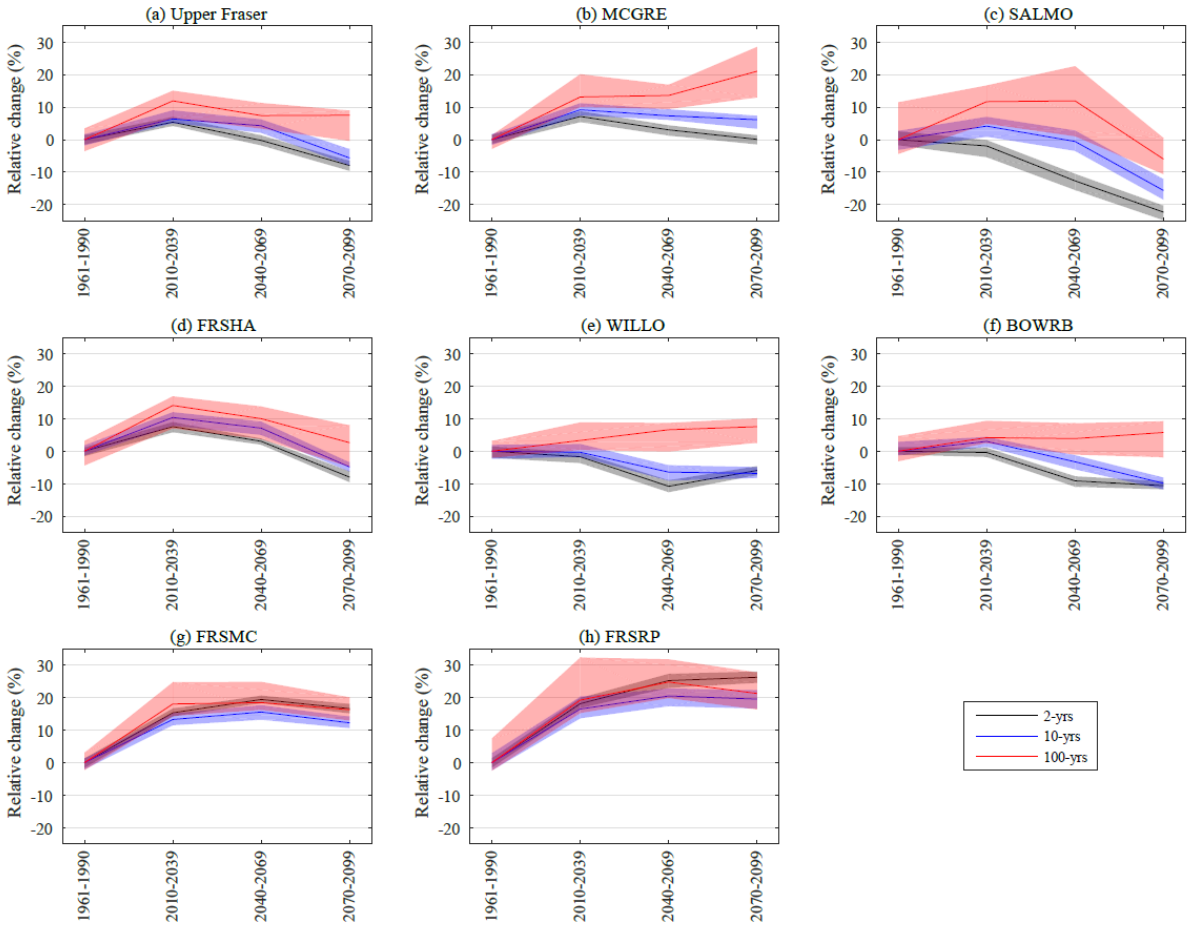

Figure 7- Projected relative change in peak flow magnitude for seven sub-basins for three return periods and three future periods. All changes are relative to the 1961-1990 period.

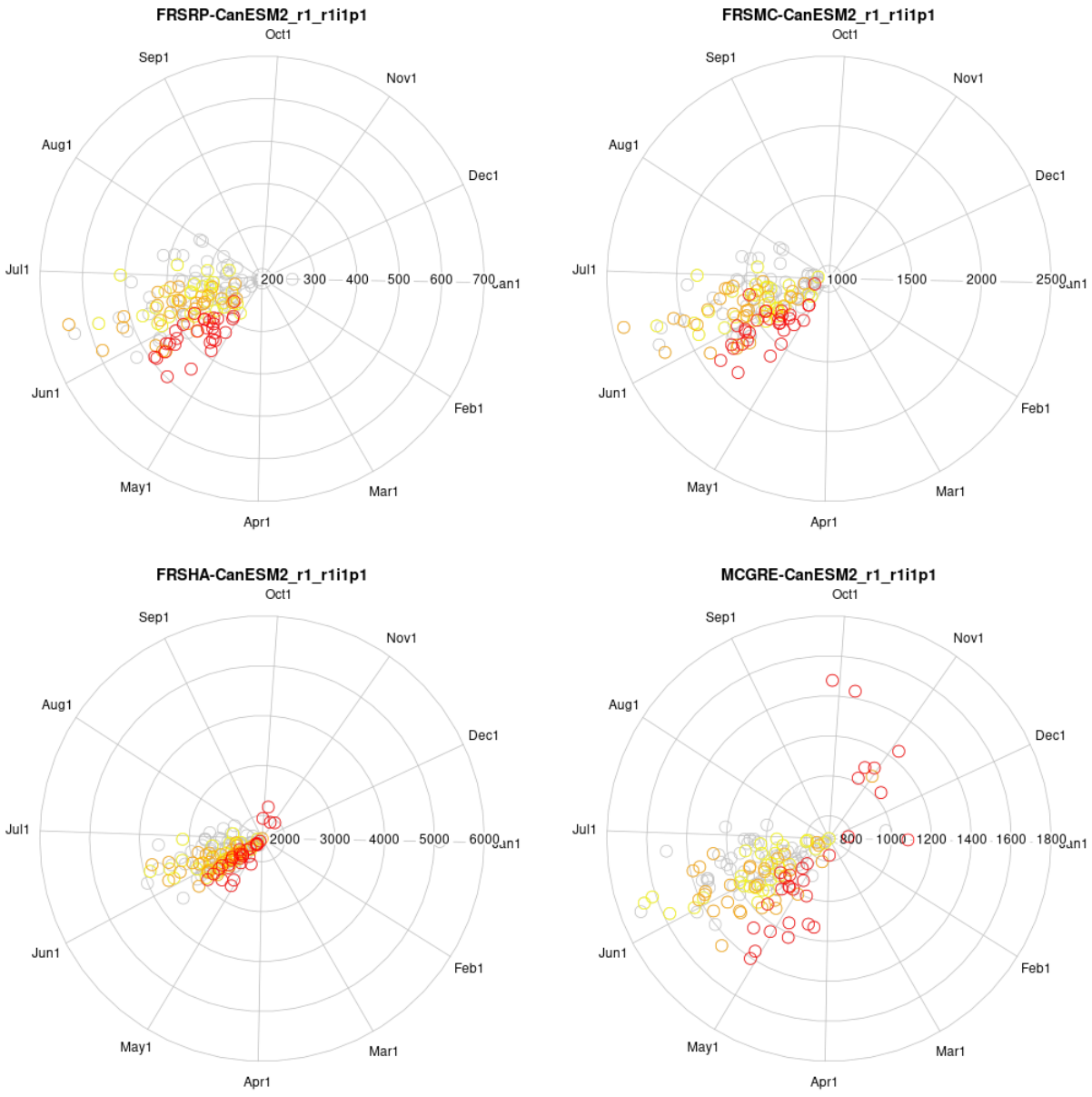

Figure 8 - polar plots of annual maximum daily peak flow occurrence and magnitude in the FRSRP, FRSMC, FRSHA and MCGRE for the single ensemble member CanESM2-run1-r1ir1p1 for the historical (1971-1990 – grey) and three future periods 2010-2039 (yellow), 2040-2069 (orange) and 2070-2099 (red). The behavior seen in run CanESM2-run1-r1ir1p1 is representative of that seen in all other runs.

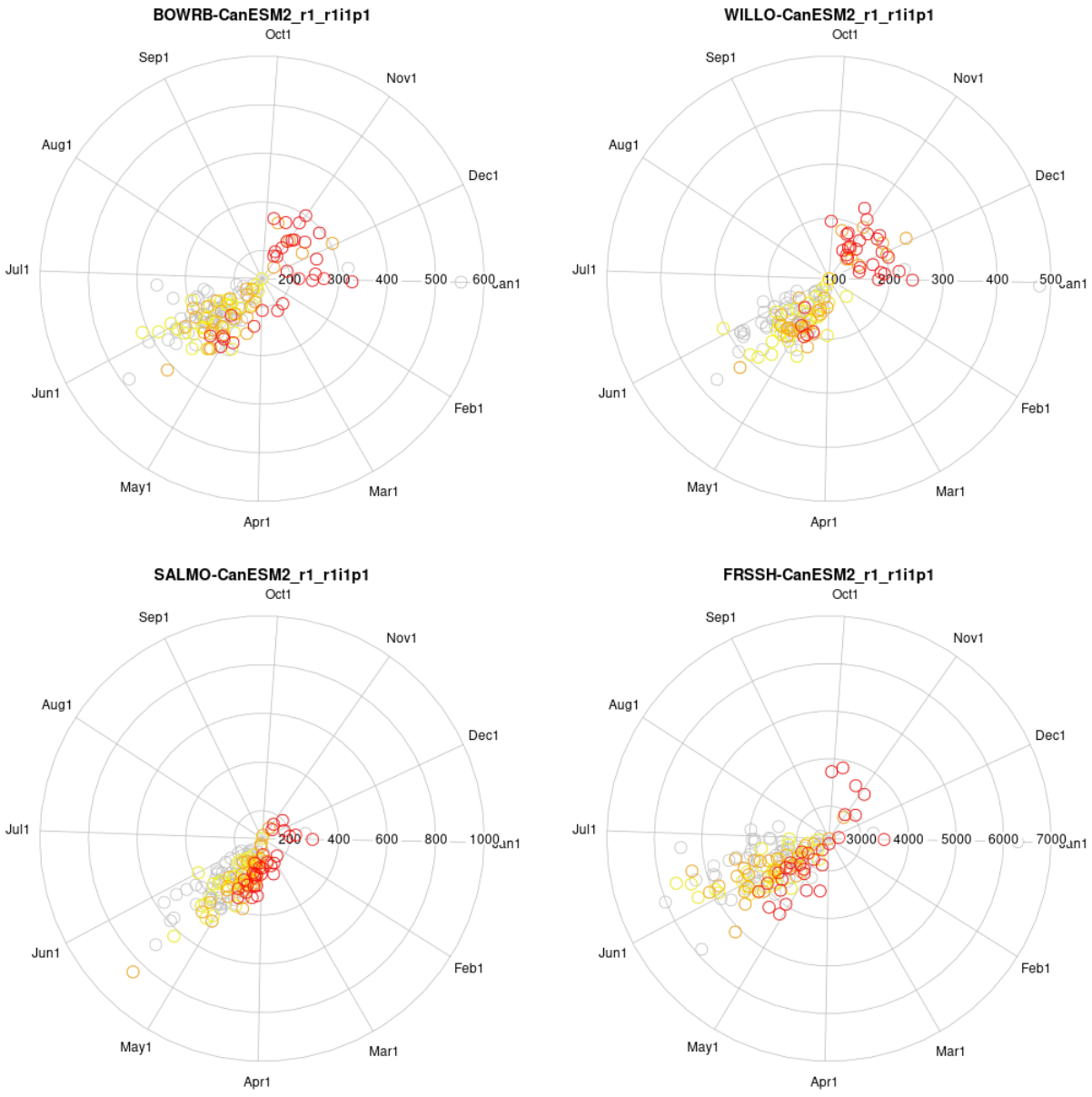

Figure 9 - polar plots of annual maximum daily peak flow occurrence and magnitude in the BOWRB, WILLO, SALMO and FRSSH for the single ensemble member CanESM2-run1-r1ir1p1 for the historical (1971-1990 – grey) and three future periods 2010-2039 (yellow), 2040-2069 (orange) and 2070-2099 (red). The behavior seen in run CanESM2-run1-r1ir1p1 is representative of that seen in all other runs.

4.2 Projected Changes in Peak Flow by Grid Cell

In this section we describe the projected changes in design flood values over the 1229 basins defined at each individual grid cell. We display results for the 2-year and 100-year events in *Figure* 10 and Figure 11, respectively. Results for the 5-, 20-, 50- and 200-year design flood values are presented graphically in Appendix C.

Consistent with the topology depicted in Figure 4, the magnitude of peak flow increases with increasing drainage area such that larger values occur where flow concentrates in the main valley-bottom channels and the largest values occur along the main stem of the Fraser River downstream of the confluence with the McGregor River (*Figure* 10 and Figure 11). 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. Consequently, a slight climatological gradient is also apparent in the gridded peak flow values, where for a given size of drainage area, peak flow magnitude is smaller in the lower-elevation and drier SALMO, WILLO and BOWRB sub-basins than in the wetter FRSRP, FRSHA and MCGRE regions.

Relative 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 any widespread clear relationship to either topography, climate, or network topology (*Figure* 10 and Figure 11, respectively). Nevertheless, for both the 2- and 100-year events, the influence of climate and elevation can be seen at a more regional scale. For the 2-yr return period event the response at each grid cell within the low-elevation WILLO, SALMO and BOWRB sub-basins tends to be a decrease in event magnitude (*Figure* 10; Table 1). Conversely, in the higher elevations of the FRSRP and FRSMC sub-basins, the response tends towards increasing event magnitude. Results are more mixed for FRSHA and MCGRE. For the 100-year event (Figure 11), projections for the region encompassed by the FRSRP, FRSMC, and MCGRE sub-basins suggest increased flood magnitude at both mid- and end-century. For this more extreme event, the area within the BOWRB and WILLO is not projected to experience very large changes, whereas results for the SALMO are more mixed. Results for the 5-, 10-, 20- and 50-year events are intermediate to the 2- and 100-year response (Appendix C; Figures C1 through C4). The projected response of the 200-year event has a similar spatial structure to that of the 100-year event, although relative changes (whether positive or negative) have slightly larger magnitude (Figure C5).

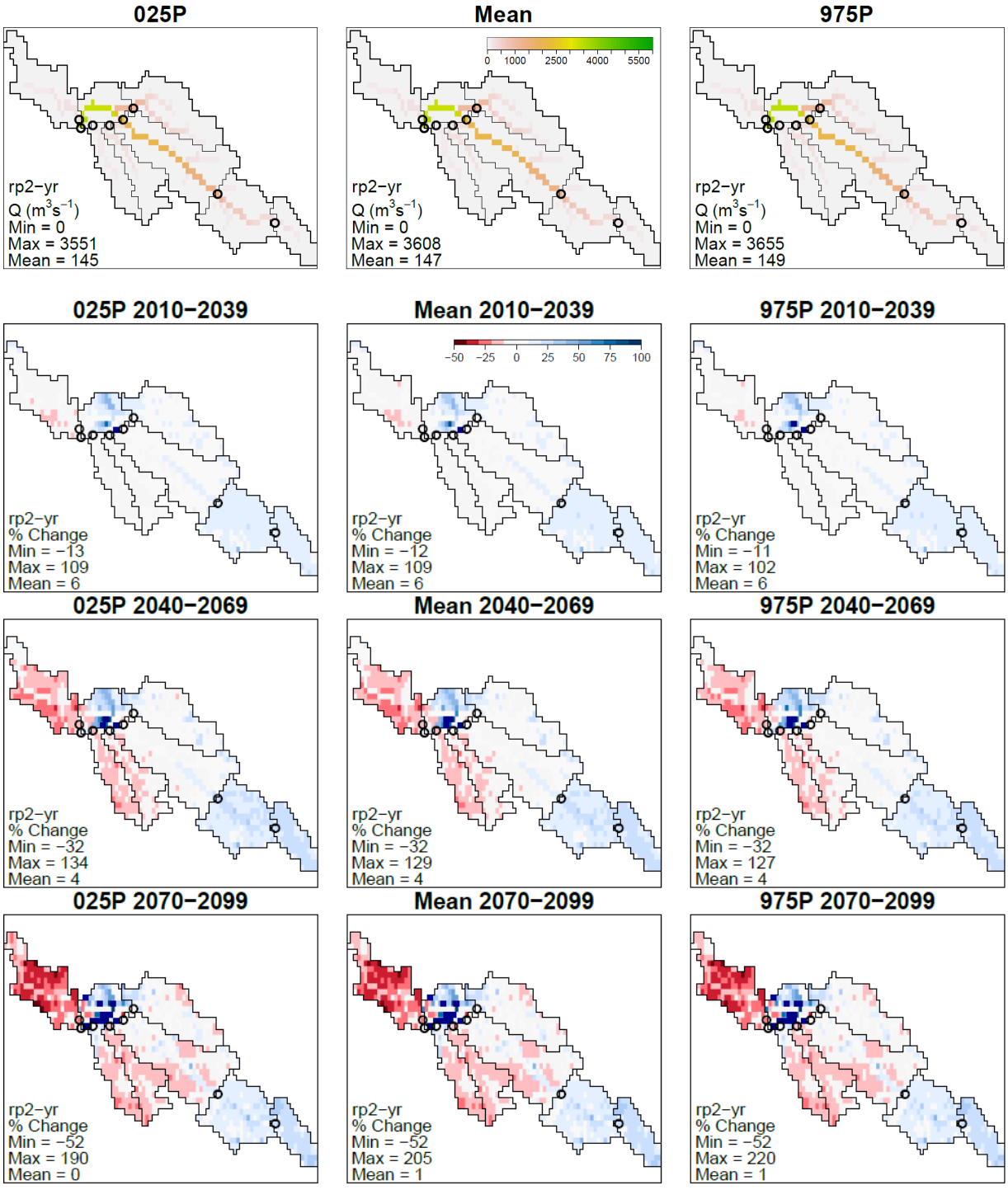

Figure *10* - Absolute (m^3/sec) 1961-1990 (top) and projected relative change (%) in annual peak flow magnitude for 2-yr return period events for 2010-2039, 2040-2069 and 2070-2099 versus the baseline period (1961-1990). Results are shown for the 2.5th percentile (left), best estimate (middle) and 97.5th percentile (right). The black circles correspond to outlets of the FRSRP, FRSMC, FRSHA, MCGRE, BOWRB, WILLO and SALMO sub basins.

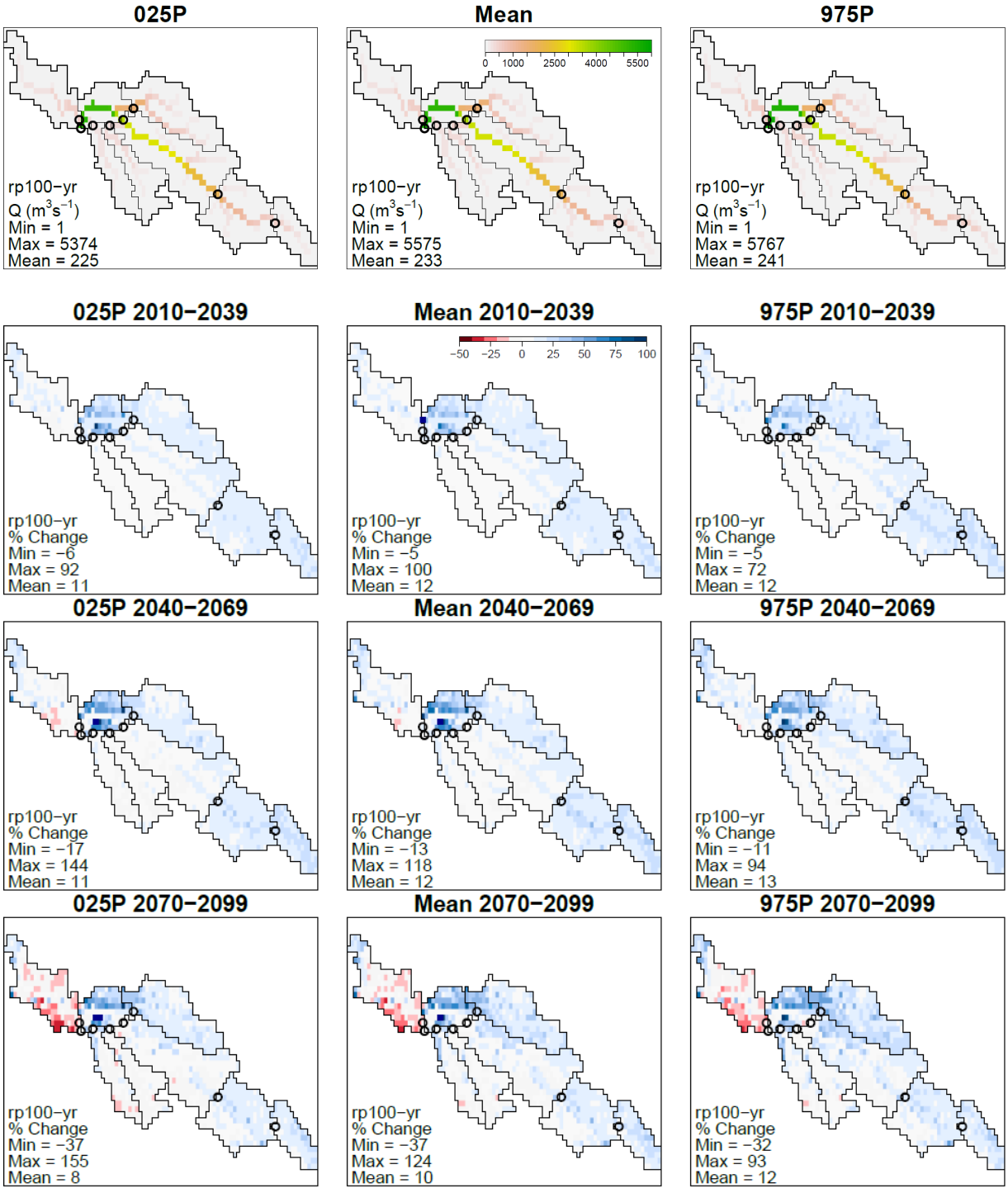

Figure 11 - Absolute (m^3/sec) 1961-1990 (top) and projected relative change (%) in annual peak flow magnitude for 100-yr return period events for 2010-2039, 2040-2069 and 2070-2099 versus the baseline period (1961-1990). Results are shown for the 2.5th percentile (left), best estimate (middle) and 97.5th percentile (right). The black circles correspond to outlets of the FRSRP, FRSMC, FRSHA, MCGRE, BOWRB, WILLO and SALMO sub basins.

4.3 Melton Ratio

The Melton Ration, which is defined as the watershed relief (km) divided by the square root of watershed area (km), is used "to differentiate watersheds prone to flooding from those subject to debris flows and debris floods" (Melton, 1957; Wilford et al., 2004). The Melton Ratio has been incorporated into PCIC's watershed API. Derived Melton Ratios of each sub-basin in the upper Fraser are summarized in Table 2. Flood watersheds have Melton ratios <0.3 and debris flows watersheds have Melton ratios >0.6 (Wilford et al., 2004). Melton ratios are all substantially less than 0.3 at the sub-basin scale (>2,500 km^2) in the upper Fraser (Table 2).

Table 2 - Melton ratio (km/km), minimum and maximum elevation (masl) and area (km^2) summarized by sub-basin.

Basin ID	Melton ratio (km/km)	Min elev. (masl)	Max elev. (masl)	Area (km^2)
FRSRP	0.045	1028	3307	2538
FRSMC	0.036	702	3929	8023
FRSHA	0.023	604	3807	19370
MCGRE	0.035	627	3199	5483
WILLO	0.026	589	1954	2843
BOWRB	0.030	627	2368	3363
SALMO	0.014	571	1558	4912
Outlet	0.018	570	3929	35119

5 Discussion, Uncertainties and Limitations

Hydrologic projections are subject to uncertainties arising from the need to make choices regarding the various components of the modelling chain (Figure 4). 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). Like all statistical approaches, BCCAQv2 also contains assumptions of stationarity, specifically in the quantile-mapping relationship and in the fine-scale spatial patterns of the temperature and precipitation fields, that may not be valid in a climate change context. In addition, errors in the chosen target data set may introduce artefacts into the downscaling process.

VIC-GL's overall performance (streamflow, evapotranspiration and snow cover) is quite strong, and the accurate representation of the seasonality of high flows in general (Figure A2), and timing of peak flows in particular (Figure A6) lends confidence that the physical mechanisms that generate peak flow events in the Upper Fraser are simulated with reasonable accuracy. Further, despite the biases in simulated annual maximum flows (Appendix A), the simulated peak flow changes reflect a physically realistic representation of the hydrologic response to climate change. Nevertheless, implementation of a hydrology and routing model requires a range of 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). Different choices and combinations can lead to a wide range in future projections (Chegwidden et al., 2019).

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 well the range in CanESM2-based hydrologic projections compares to the range produced if a larger set of GCMs were used. We explored this issue by comparing annual maximum peak flows generated from the CanESM2 large ensemble (50 realizations) with those simulated using the PCIC6 ensemble, which is composed of six GCMs, some with multiple runs, for total of 15 ensemble members. The PCIC6 ensemble explores both GCM structural variability (different GCMs) and, to some degree, internal variability (multiple runs per GCM). The six GCMs that comprise the PCIC6 were selected specifically so as to best capture the full spread in climate change projections available from the full set of CMIP5 GCMs using the method of Cannon (2015). A comparison of simulated annual maximum streamflow produced by both ensembles is plotted by sub basin in Appendix D . Despite the fact that the CanESM2 large ensemble is designed only to address internal climate variability, results show that the spread of the CanESM2 large ensemble is comparable to (and even larger than) the spread captured by the PCIC6 ensemble. One can also note from the figures in Appendix E that magnitude of peak flow (as indicated by the respective ensemble medians) is higher in CanESM2 than PCIC6. Also, the direction of trend of the ensemble medians of annual maximum flow over the coming century is identical between ensembles in all sub basins, although and CanESM2 has larger positive and smaller (less negative) increasing and decreasing trends, respectively, than PCIC6. Therefore, we suggest that range of results provided by the large ensembles from CanESM2 coupled with the high-end RCP8.5 scenario can be considered to provide an upper-limit or worst-case estimate of the range of future design flow changes in the Upper Fraser.

The resolution of the VIC-GL model also offers challenges with regard to the interpretation of streamflow values and peak flow changes. Each grid cell can only have a single flow direction and a single channel, therefore substantial simplification of the drainage system is imposed. 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.

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 1500 identically distributed peak flow events. A comparison of quantile estimation using both the GEV approach, and the direct empirical approach indeed confirms that the parametric GEV approach produces biased estimates, particularly for large events (see Appendix E, Figure E1). 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 estimates (Figure E2 and Figure E3).

6 Conclusions

PCIC has completed a pilot 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. As part of this work, PCIC has also provided the Melton Ratio as a gridded product. This pilot study uses on the Upper Fraser, a 34,200 km^2 region upstream of Prince George, BC, as the study region. 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 ~32,400 km^2. Design flood values are provided for the thirty-year blocks 1961-1990, 2010-2039, 2040-2069, and 2070-2099, respectively representing the historical, future early-, future mid- and future end-century periods. 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.

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. Nevertheless, despite being limited to only a single hydrology mode, the performance of VIC-GL is such that the simulated peak flow changes reflect a physically realistic representation of the hydrologic response to climate change. However, based on the sensitivity of the peak flow response from CanESM2 and the large spread in the CanESM2 large ensemble as compared to other GCMs, coupled with the RCM8.5 scenario, the range of results provided by the large ensemble can be considered an upper-limit, or worst-case estimate, of the range of future design flow changes in the Upper Fraser.

7 References

Alaya, M.A.B., Zwiers, F., Zhang, X., 2020. An Evaluation of Block-Maximum-Based Estimation of Very Long Return Period Precipitation Extremes with a Large Ensemble Climate Simulation. J. Clim. 33, 6957–6970. https://doi.org/10.1175/JCLI-D-19-0011.1

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.

Burn, D.H., Whitfield, P.H., 2016. Changes in floods and flood regimes in Canada. Can. Water Resour. J. Rev. Can. Ressour. Hydr. 41, 139–150. https://doi.org/10.1080/07011784.2015.1026844

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., 2015. Selecting GCM Scenarios that Span the Range of Changes in a Multimodel Ensemble: Application to CMIP5 Climate Extremes Indices. J. Clim. 28, 1260–1267. https://doi.org/10.1175/JCLI-D-14-00636.1

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.

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

Gao, C., Booij, M.J., Xu, Y.-P., 2020. Assessment of extreme flows and uncertainty under climate change: disentangling the contribution of RCPs, GCMs and internal climate variability. Hydrol. Earth Syst. Sci. Discuss. 1–28. https://doi.org/10.5194/hess-2020-25

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

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., Zhang, X., 2020. Human influence has intensified extreme precipitation in North America. Proc. Natl. Acad. Sci. https://doi.org/10.1073/pnas.1921628117

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

Kirchmeier-Young, M.C., Zwiers, F.W., Gillett, N.P., Cannon, A.J., 2017b. Attributing extreme fire risk in Western Canada to human emissions. Clim. Change 1–15. https://doi.org/10.1007/s10584-017-2030-0

Klemes, V., 1986. Operational testing of hydrological simulation models. Hydrol. Sci. 31, 13–24.

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

Li, C., Zwiers, F., Zhang, X., Chen, G., Lu, J., Li, G., Norris, J., Tan, Y., Sun, Y., Liu, M., 2019. Larger Increases in More Extreme Local Precipitation Events as Climate Warms. Geophys. Res. Lett. 46, 6885–6891. https://doi.org/10.1029/2019GL082908

Li, C., Zwiers, F., Zhang, X., Li, G., Sun, Y., Wehner, M., 2020. Changes in annual extremes of daily temperature and precipitation in CMIP6 models. J. Clim. 1, 1–61. https://doi.org/10.1175/JCLI-D-19-1013.1

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

Melsen, L.A., Addor, N., Mizukami, N., Newman, A.J., Torfs, P.J.J.F., Clark, M.P., Uijlenhoet, R., Teuling, A.J., 2018. Mapping (dis)agreement in hydrologic projections. Hydrol. Earth Syst. Sci. 22, 1775–1791. https://doi.org/10.5194/hess-22-1775-2018

Melton, M.A., 1957. An analysis of the relations among elements of climate, surface properties, and geomorphology. Office of Nav Res Dept Geol Columbia Univ, NY. Tech Rep 11. https://doi.org/10.7916/d8-0rmg-j112

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.

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

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., in press. VICGL Model Deployment Report (PCIC Internal Report). Pacific Climate Impacts Consortium, Victoria, BC.

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.

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.

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

Shrestha, R.R., Schnorbus, M.A., Peters, D.L., 2016. Assessment of a hydrologic model's reliability in simulating flow regime alterations in a changing climate. Hydrol. Process. 30, 2628–2643. https://doi.org/10.1002/hyp.10812

Shrestha, R.R., Schnorbus, M.A., Werner, A.T., Zwiers, F.W., 2014. Evaluating Hydroclimatic Change Signals from Statistically and Dynamically Downscaled GCMs and Hydrologic Models. J. Hydrometeorol. 15, 844–860. https://doi.org/10.1175/JHM-D-13-030.1

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

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

Sun, Q., Zhang, X., Zwiers, F., Westra, S., Alexander, L.V., 2020. A Global, Continental, and Regional Analysis of Changes in Extreme Precipitation. J. Clim. 34, 243–258. https://doi.org/10.1175/JCLI-D-19-0892.1

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

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

Wilford, D.J., Sakals, M.E., Innes, J.L., Sidle, R.C., Bergerud, W.A., 2004. Recognition of debris flow, debris flood and flood hazard through watershed morphometrics. Landslides 1, 61–66. https://doi.org/10.1007/s10346-003-0002-0

Appendix A - Verification of VIC-GL in the Upper Fraser

We summarize the performance of VIC-GL in simulating daily and annual maximum streamflow, snow covered area, evapotranspiration and basin-average glacier surface mass balance in seven sub-basins. This evaluation helps to understand the uncertainties inherent in projecting future annual max streamflow. Both the model calibration and validation process are conducted by forcing the VIC-GL model with gridded observations of daily minimum and maximum temperature and precipitation from PNWNAmet (Werner et al., 2019).

A.1 VIC-GL Calibration

Model calibration is governed by the desire to exploit the spatially distributed nature of the VIC-GL model. However, streamflow is the primary variable for water resources planning and management, thus calibration works around the availability of discharge data. In addition to discharge (*Q*), selected parameter sets for sub-basin were constrained using snow covered area (*SCA)*, evapotranspiration (*ET*) and basin-average glacier surface mass balance (*B*). Calibration and validation performance is evaluated using the metrics and evaluation periods described in Table A1. Performance metrics include the Kling-Gupta efficiency (KGE; Gupta et al. 2009), the Nash-Sutcliffe efficiency for log-transformed discharge (LNSE; Nash and Sutcliffe 1970) and the bell membership function (*BMF*; Zhao and Bose 2002). The possible value ranges for the various metrics are -∞ to 1 for *KGE* (1 is best), -∞ to 1 for *LNSE* (1 is best) and 0 to 1 for *BMF* (1 is best). Seven sub-basins had sufficient data over the collective calibration period (1991-2000). See Table 1 for a description of these stations, Figure 1 and *Figure 5* for their locations.

Table A1 - Calibration and validation metrics, and evaluation periods

	Discharge (*Q*)	Evapotranspiration (*ET*)	Snow Cover (*SCA*)	Glacier Mass Balance (*B*)
Statistics	*KGE, LNSE*	*BMF*	*KGE*	*BMF*
Calibration Period	1991-2000	1991-2000	2000-2005	1985-1999
Validation Period	2001-2007	2001-2005	2006-2010	1985-1999[$]

Model performance is generally high over the calibration period (Table A2, Figure A1). KGE_Q values are all greater than zero, with a median value of 0.86. This indicates that the model shows high correlation, low bias and variability that is well-matched to the observations. $LNSE_Q$ values are greater than zero for the majority of sub-basin. However, two sub-basins have $LNSE_Q$ scores less than zero, which is worse performance than the mean. Hydrologic models are known to struggle in simulating low-flows (Her et al., 2019; Melsen et al., 2018; Shrestha et al., 2014, 2016; Werner and Cannon, 2016), so this result is not altogether surprising.

VIC-GL simulates evaporation well in these basins, based on the BMF_{ET} scores, with values are 0.43 or greater in the calibration period. Snow covered area is modelled extremely well with KGE_{SCA} scores of 0.74 or greater. For the sub-basins that contain glaciers, FRSRP, FRSMC, FRSHA and MCGRE, all except FRSRP simulate glacier mass balance well based on BMF_B with values that exceed 0.8 and therefore simulated glacier mass balance is within the range of the upper and lower uncertainty bounds of the geodetic mass balance observations.

Table A2 - Calibration results for streamflow (Q), evaporation (ET), snow covered area (SCA) and glacier mass balance (B)

Basin Code	Calibration				
	KGE_Q	$LNSE_Q$	BMF_{ET}	KGE_{SCA}	BMF_B
FRSRP	0.67	-0.15	0.43	0.93	0.48
FRSMC	0.89	0.86	0.43	0.83	1.00
FRSHA	0.92	0.88	0.49	0.85	1.00
MCGRE	0.83	0.78	0.46	0.83	0.88
WILLO	0.88	0.66	0.55	0.74	
BOWRB	0.86	0.32	0.59	0.76	
SALMO	0.80	-0.08	0.52	0.82	

A.2 VIC-GL Validation

We used a split-sample approach to model validation (Klemes, 1986), where the available observed data is split into two periods, one of which is used for calibration and the other for validation. Validation of model performance with independent data is necessary to ensure the model parameterization is robust, which is important to establish confidence in the model's ability to extrapolate to climatic conditions unobserved during the historical period. During model calibration, simulated streamflow at the calibration site is composed simulated local runoff and observed upstream inflow, all routed to the outlet site (i.e., any errors from upstream basins are removed by using observed data). However, simulated streamflow for the validation runs is based on simulated runoff data from the entire upstream domain. Glacier area is held static during model calibration but allowed to evolve dynamically during the validation runs. As such, the validation diagnostics represent a very stringent test of VIC-GL performance.

Table A3 - Validation results for streamflow (Q), evaporation (ET), snow covered area (SCA) and glacier mass balance (B)

Basin	Validation				
	KGE_Q	$LNSE_Q$	BMF_{ET}	KGE_{SCA}	BMF_B
FRSRP	0.50	-0.09	0.37	0.95	0.98
FRSMC	0.52	0.85	0.40	0.87	0.60
FRSHA	0.76	0.63	0.41	0.88	0.95
MCGRE	0.83	0.74	0.46	0.86	0.90
WILLO	0.85	0.57	0.50	0.82	
BOWRB	0.85	0.41	0.56	0.81	
SALMO	0.74	-0.03	0.55	0.83	

Performance metrics for streamflow decrease moderately in the validation period (Table A4). This is expected due to the more stringent nature of the validation test (compared to calibration) and the fact that some over-fitting to the calibration period is unavoidable. KGE_Q values are 0.74 or greater for all basins except FRSRP (0.50) and FRSMC (0.52). FRSRP was already weakly calibrated in comparison to the other basins, while FRSMC decreased from 0.89 to 0.52. LNSE is 0.63 or greater for four of seven basins, 0.41 in BOWRB and less than zero in FRSRP and SALMO. FRSRP and SALMO already had poor performance in the calibration period for this metric (Table A2). Simulated and observed daily streamflow during the calibration and validation periods are shown in Figure 5. Although the Fraser River at Shelley (FRSSH; WSC gauge 08KB001) data was not used for model calibration, it is included for

streamflow validation purposes and presented in Figures A1 through A7. Model streamflow validates well at FRSSH with a KGE_Q of 0.79 and $LNSE_Q$ of 0.61. The simulation of glaciers improves for FRSRP, but worsens for FRSMC, dropping to 0.60 from 1 for BMF_B. Evaporation performance is fairly stable, decreasing only minimally, and continues to be high in the validation period, based on BMF_{ET} (Table A4). Snow covered area is simulated as well or better than it was in the calibration period based on KGE_{SCA}.

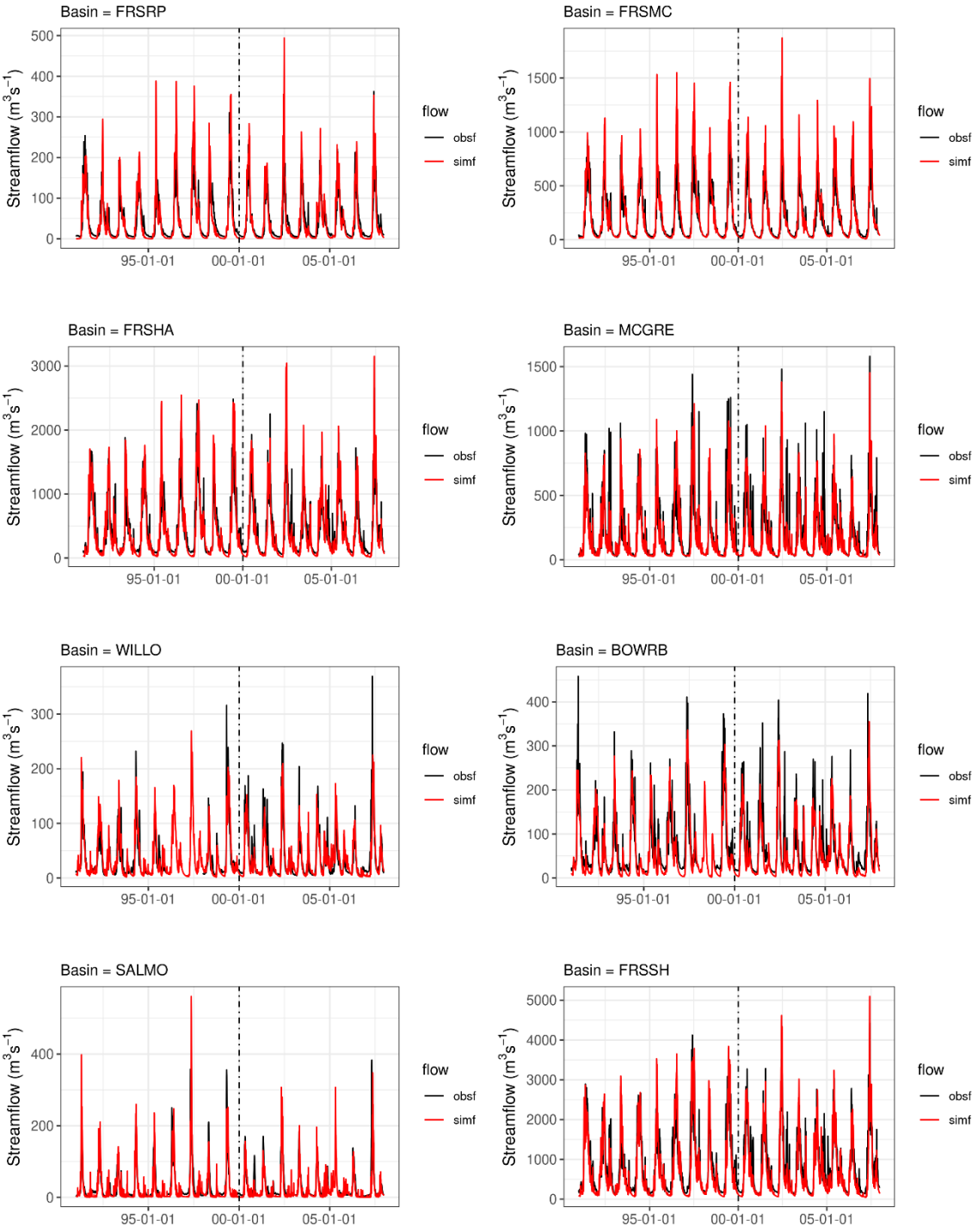

Figure A1 - Daily streamflow for calibration (1991-2000) and validation (2001-2007) periods..

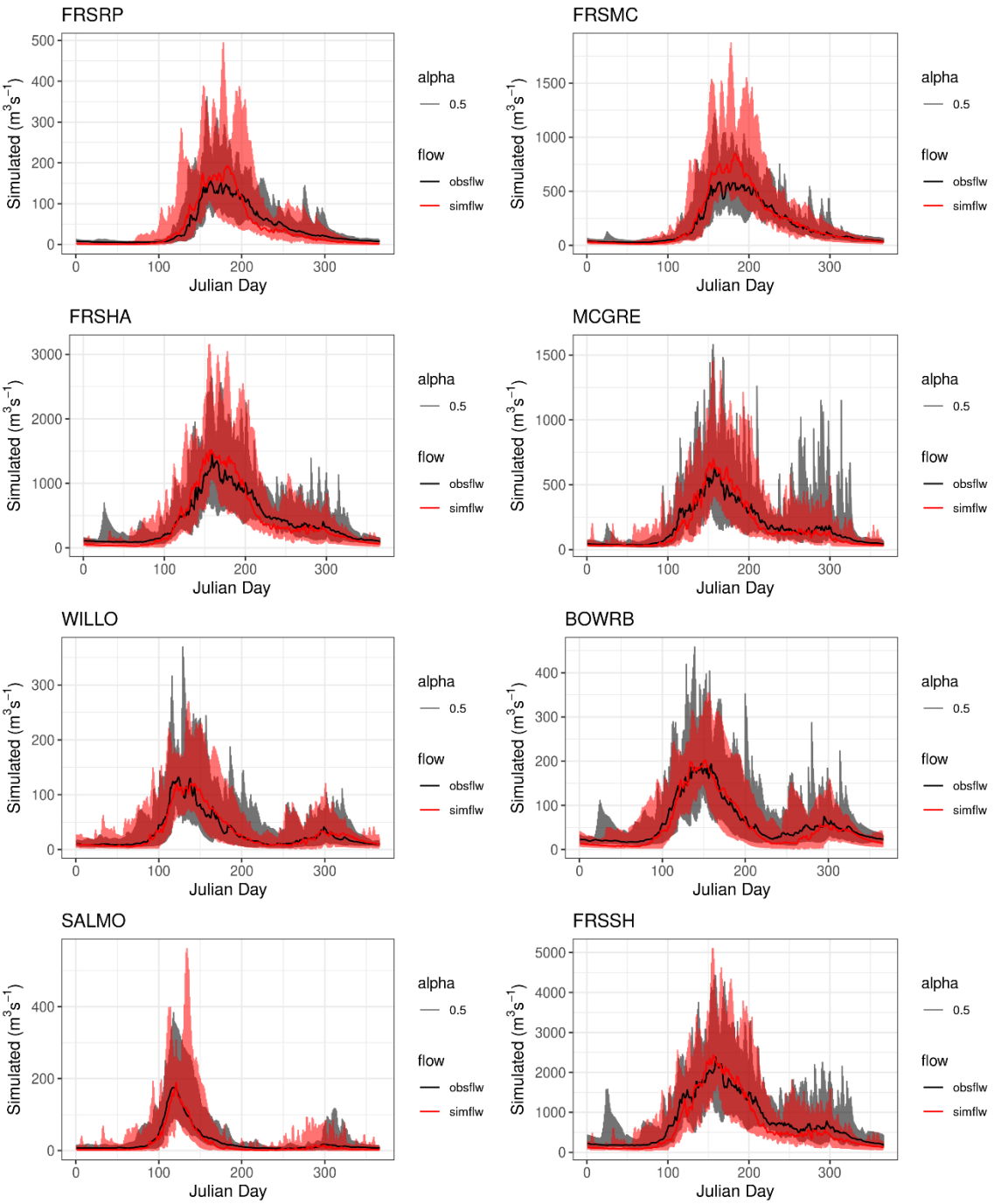

Figure A2 - Minimum and maximum (polygon) and median (solid line) of simulated versus observed daily streamflow by Julian day over 1990-2007.

A.3 VIC-GL Annual Maximum Flow

As peak flows are the emphasis of this study, we evaluate the performance of VIC-GL specifically for this metric. Simulated annual daily maximum peak flows are higher than observed in FRSRP, FRSMC, FRSHA, SALMO and FRSSH, and lower in magnitude in MCGRE, WILLO and BOWRB (Table A4, Figure A3). Relative bias, RBIAS, ranges from 24% to -26%, RMSE from 48% to 24% and MAE from 40% to 18% of Mn.O. Although simulated and observed peaks flows show significant positive correlation at the 5% significance level (Table A4, Figure A3 and A4) in all basins, the distribution of the observed and simulated peak flow events can only be considered statistically identical at the 5% significance level in the FRSHA, MCGRE, WILLO and FRSSH basins (Table A4, Figure A5).

The timing of peak flows in the simulated and observed are compared with polar plots (Figure A6 and Figure A7). Annual maximum peak flows are generated almost exclusively during the snowmelt freshet and occur predominately between May 1st and July 1st in all basins, except in lower-elevation SALMO, BOWRB and WILLO, where events can occur as early as mid-April. The MCGRE is the only site that exhibits a bi-modal distribution of annual maximum peak flows, with a minority of events occurring after the freshet period during the fall, likely as a result of extreme rainfall events. The seasonality, or timing, of annual maximum flow events is similar between the observed and simulated at all validation sites, including the MCGRE. This gives us confidence that the dominant physical mechanisms generating freshet peak flows are accurately represented by the VIC-GL model.

Table A4 – Performance summary for simulated versus observed annual daily maximum peak flow events (1950-2010). Statistics include observed (Mn.O) and simulated (Mn.S) mean, observed and simulated standard deviation (SD.O and SD.S, respectively), observed and simulated coefficient of variation (Cv.O and Cv.S, respectively), bias (RBIAS), root mean square error (RMSE), mean absolute error (MAE), Spearman rank correlation (RS) and Kolmogorov-Smirnov test statistic (KS). RBIAS, RMSE and MAE are presented as a percentage relative to Mn.O. Values of RS and KS that are significant at the 5% level (two-sided test) are shown in bold. The number of observations available over the 1950-2010 period are given as percent records (% Recs).

Basin	Mn.O (m^3s^{-1})	Mn.S (m^3s^{-1})	SD.O (m^3s^{-1})	SD.S (m^3s^{-1})	Cv.O (none)	Cv.S (none)	RBIAS (%)	RMSE (%)	MAE (%)	RS (none)	KS (none)	%Recs (%)
FRSRP	244	303	52	88	0.21	0.29	24	38	30	**0.48**	0.40	90
FRSMC	899	1261	155	269	0.17	0.21	40	48	40	**0.46**	0.64	87
FRSHA	2006	2232	386	537	0.19	0.24	11	27	20	**0.43**	**0.25**	93
MCGRE	1100	1007	265	227	0.24	0.23	-8	22	18	**0.45**	**0.16**	82
WILLO	235	174	118	38	0.50	0.22	-26	46	31	**0.81**	**0.33**	49
BOWRB	326	251	93	55	0.28	0.22	-23	32	24	**0.75**	0.52	54
SALMO	213	264	74	105	0.35	0.40	24	41	32	**0.73**	0.30	75
FRSSH	3209	3475	658	800	0.21	0.23	8	24	18	**0.51**	**0.25**	100

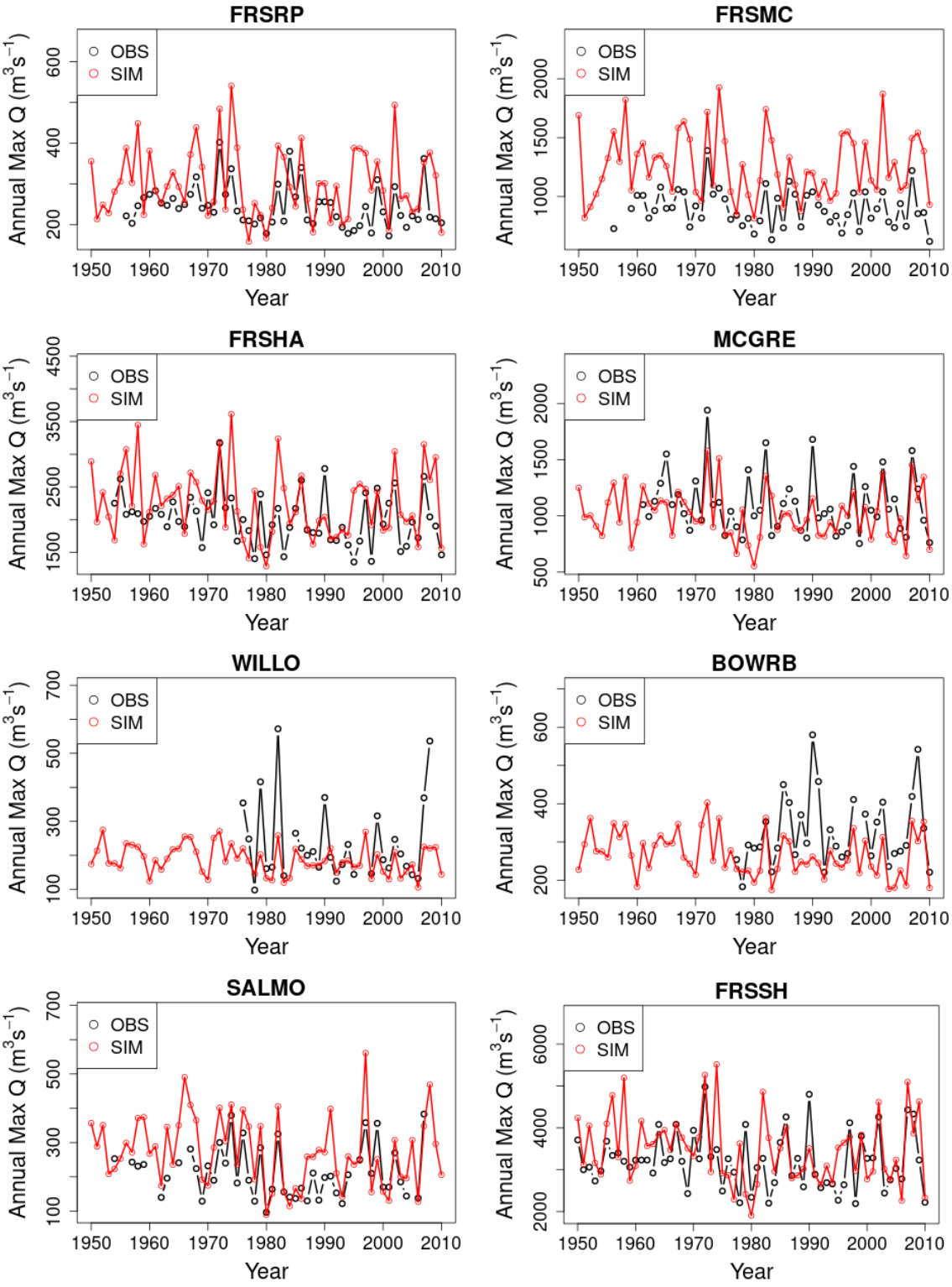

Figure A3 – time series of annual daily max streamflow for simulated and observed 1950-2010.

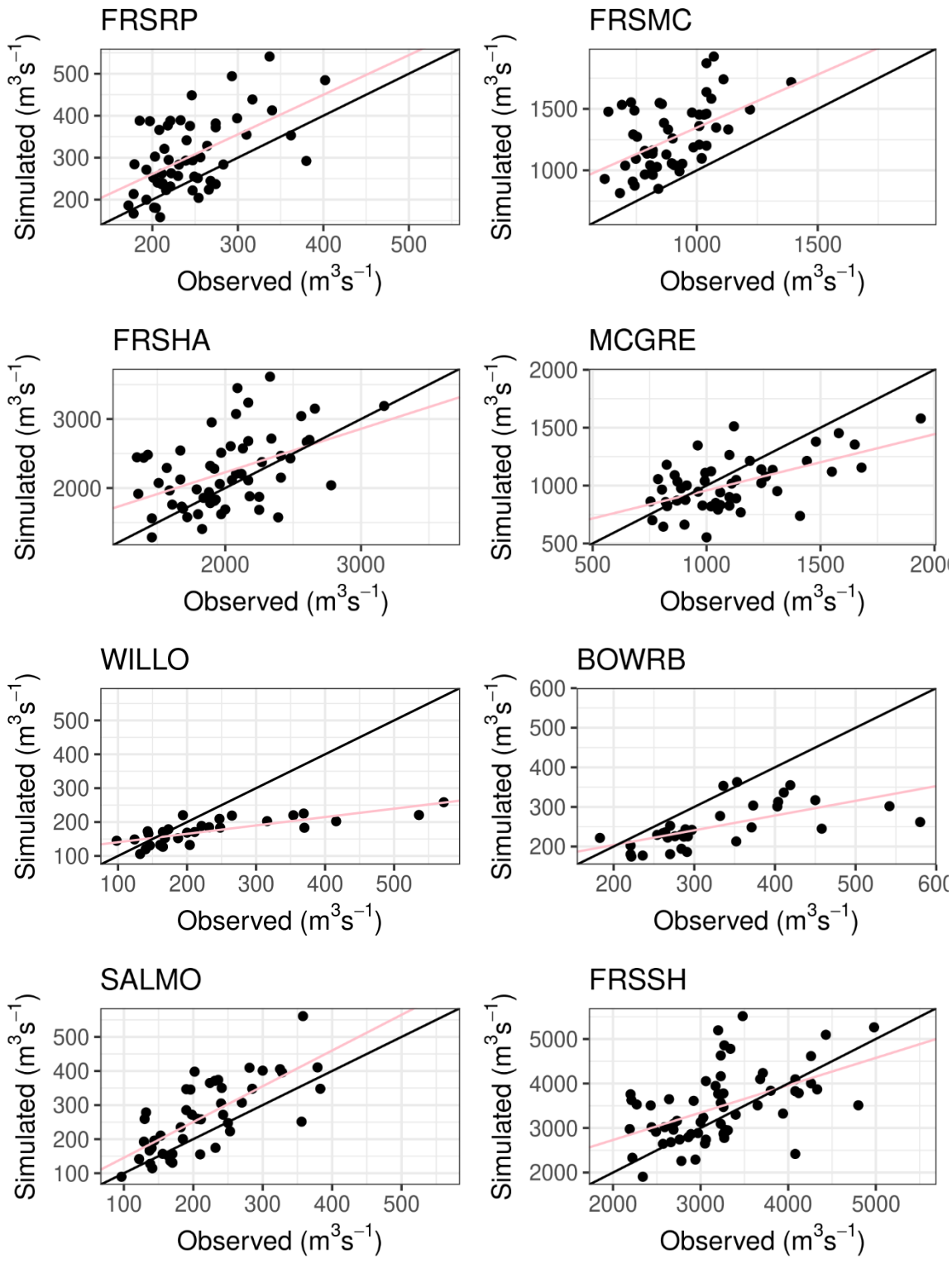

Figure A4 - Simulated annual max flow versus observed by sub-basin for 1950-2010. One-to-one line (black) and simple linear regression line for best-fit (pink)..

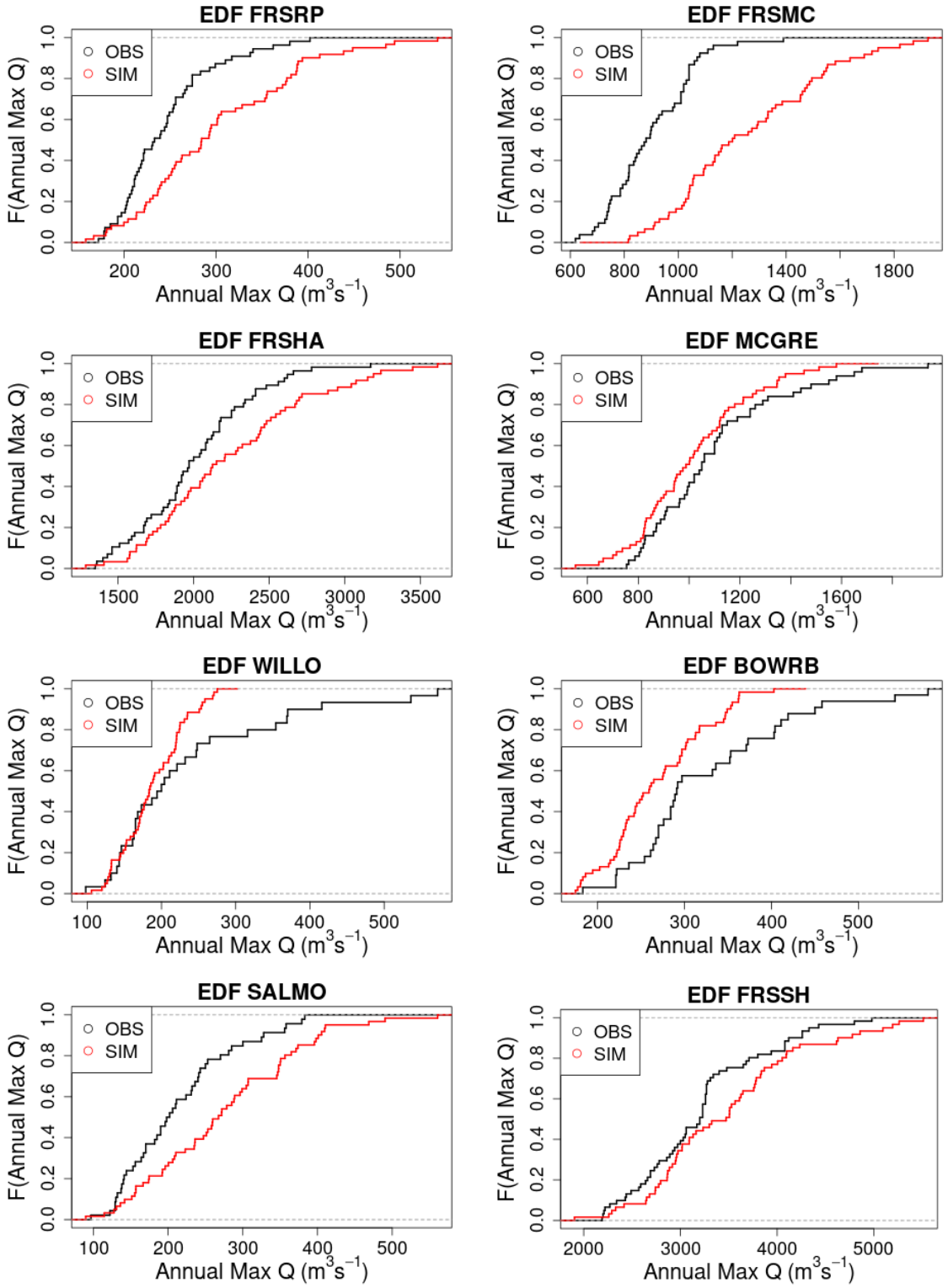

Figure A5 - empirical distribution function for simulated and observed annual max streamflow by basin for 1950-2010.

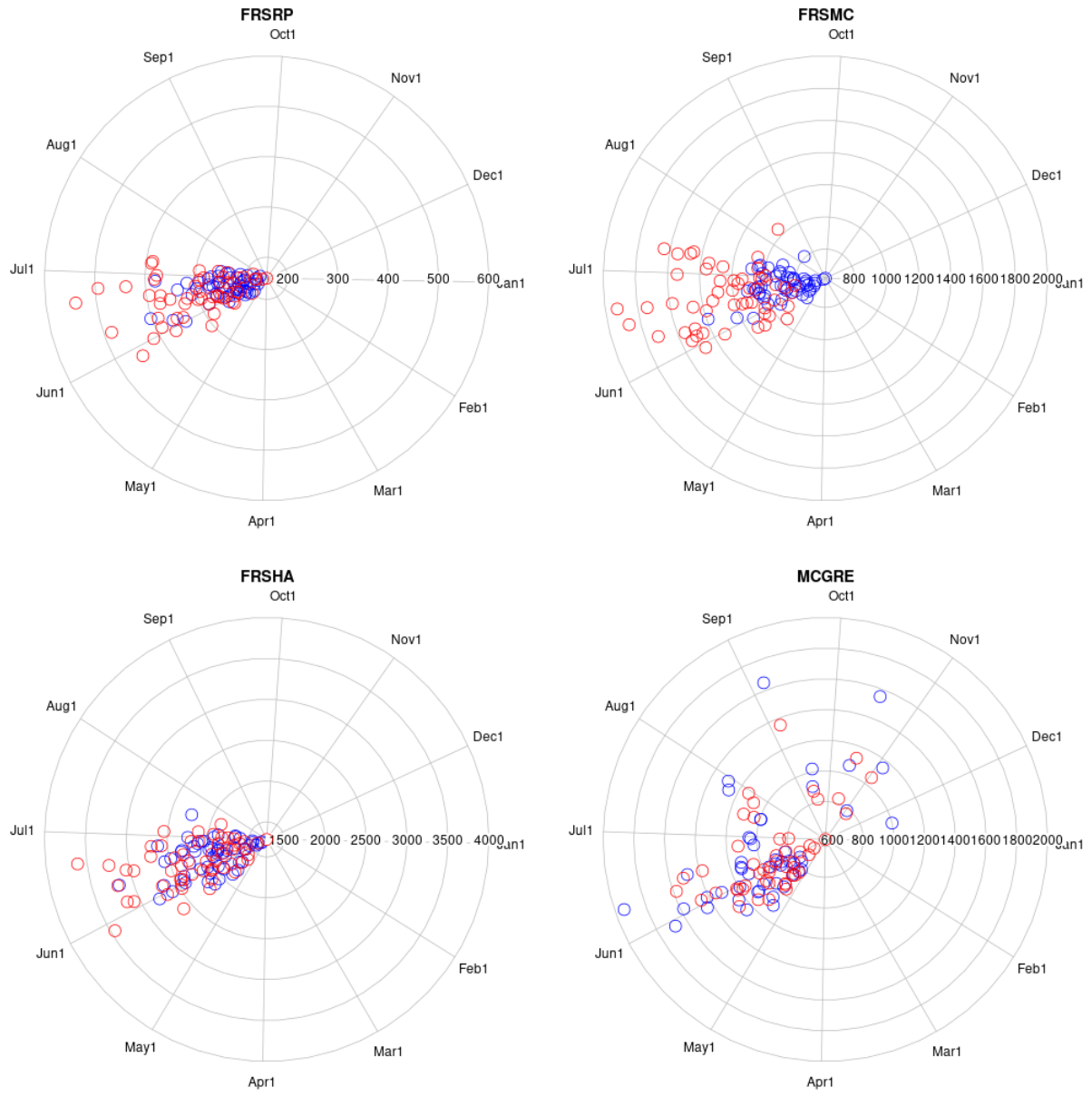

Figure A6 - polar plots comparing simulated (red) and observed (blue) annual daily max peak flow timing and magnitude in the FRSRP, FRSMC, FRSHA and MCGRE for 1950-2010.

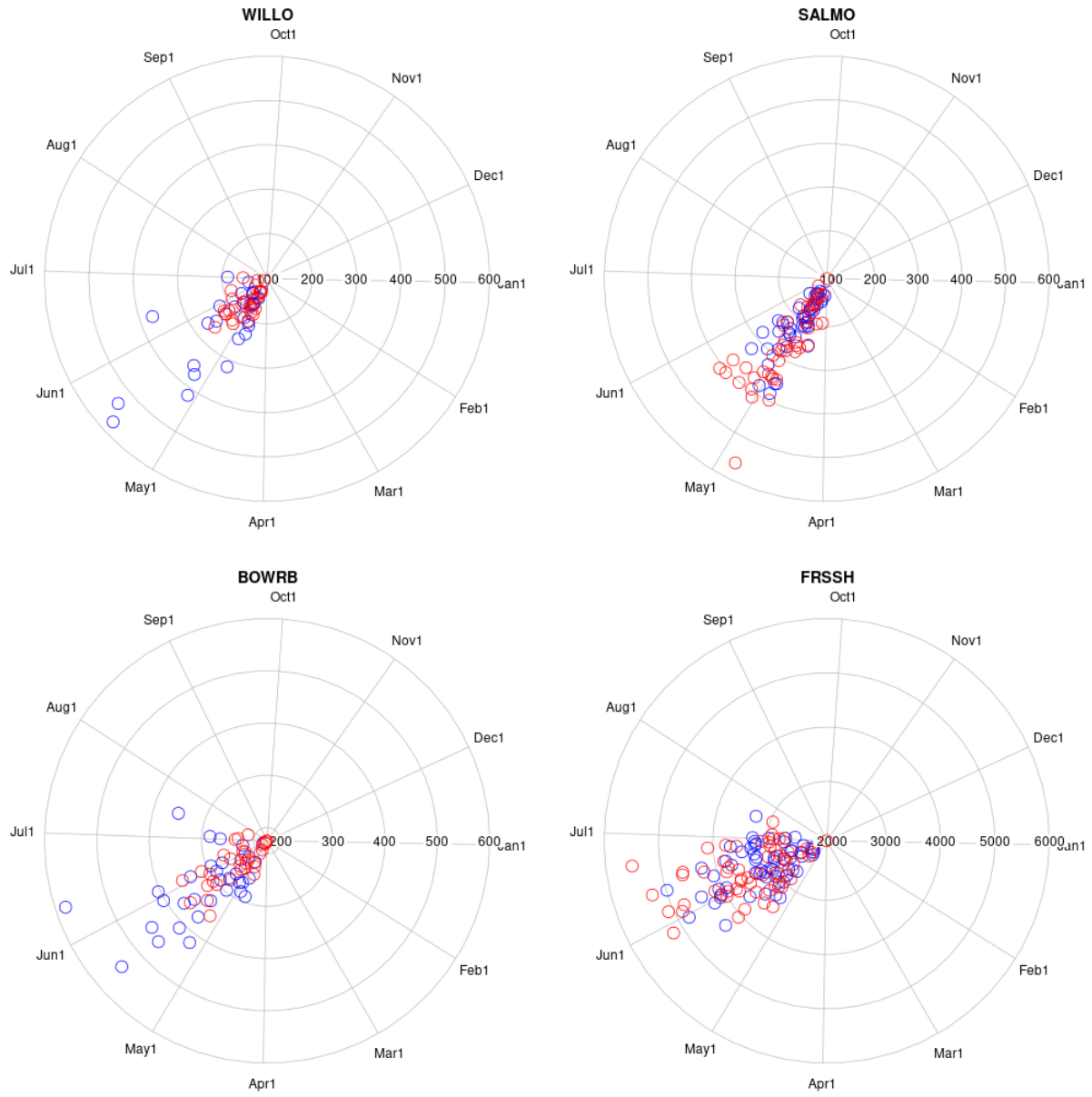

Figure A7 - same as Figure 18 but for WILLO, SALMO, BOWRB and FRSSH.

Appendix B – Flood Frequency Tables

Table B1- Best estimates of absolute (m^3/sec) for 1961-1990 and projected relative change (units: %) in annual peak flow magnitude by three future periods (2010-2039, 2040-2069 and 2070-2099) for different quantiles for the pour point of the FRSRP basin. The baseline period is 1961-1990.

Time Horizon	Return period events						
	2-yr	5-yr	10-yr	20-yr	50-yr	100-yr	200-yr
1961-1990	*295*	*374*	*419*	*464*	*506*	*526*	*568*
2010-2039	18	16	17	16	18	19	23
2040-2069	25	21	21	19	20	25	22
2070-2099	26	21	20	17	19	21	18

Table B2 - same as Table B1 but for FRSMC.

Time Horizon	Return period events						
	2-yr	5-yr	10-yr	20-yr	50-yr	100-yr	200-yr
1961-1990	*1258*	*1508*	*1673*	*1799*	*1915*	*1987*	*2063*
2010-2039	15	14	13	13	14	18	21
2040-2069	19	18	16	15	16	18	21
2070-2099	17	14	12	12	14	16	18

Table B3 - same as Table B1 but for FRSHA.

Time Horizon	Return period events						
	2-yr	5-yr	10-yr	20-yr	50-yr	100-yr	200-yr
1961-1990	*2334*	*2752*	*3019*	*3252*	*3469*	*3671*	*3823*
2010-2039	8	9	10	9	11	14	13
2040-2069	3	6	7	7	7	10	10
2070-2099	-8	-6	-5	-3	0	3	6

Table B4 - same as Table B1 but for MCGRE.

Time Horizon	Return period events						
	2-yr	5-yr	10-yr	20-yr	50-yr	100-yr	200-yr
1961-1990	*1033*	*1227*	*1325*	*1396*	*1521*	*1588*	*1615*
2010-2039	7	8	9	12	11	13	25
2040-2069	3	5	7	11	13	14	15
2070-2099	0	1	6	10	15	21	27

Table B5 - same as Table B1 but for WILLO.

Time Horizon	Return period events						
	2-yr	5-yr	10-yr	20-yr	50-yr	100-yr	200-yr
1961-1990	*189*	*228*	*252*	*269*	*293*	*301*	*313*
2010-2039	-2	0	0	2	2	3	6
2040-2069	-11	-9	-6	-2	0	7	6
2070-2099	-6	-6	-7	-5	2	8	8

Table B6 - same as Table B1 but for BOWRB.

Time Horizon	Return period events						
	2-yr	5-yr	10-yr	20-yr	50-yr	100-yr	200-yr
1961-1990	*266*	*317*	*343*	*365*	*385*	*401*	*424*
2010-2039	0	2	3	4	6	4	6
2040-2069	-9	-6	-3	-1	2	4	5
2070-2099	-11	-12	-10	-7	-1	6	3

Table B7 - same as Table B1 but for SALMO.

Time Horizon	Return period events						
	2-yr	5-yr	10-yr	20-yr	50-yr	100-yr	200-yr
1961-1990	*282*	*375*	*431*	*486*	*537*	*581*	*657*
2010-2039	-2	1	4	6	9	12	5
2040-2069	-13	-4	-1	0	8	12	10
2070-2099	-22	-19	-16	-12	-6	-6	-11

Table B8 - same as Table 2B1 but for Upper Fraser.

Time Horizon	Return period events						
	2-yr	5-yr	10-yr	20-yr	50-yr	100-yr	200-yr
1961-1990	*3605*	*4237*	*4656*	*4960*	*5305*	*5573*	*5776*
2010-2039	5	8	6	9	11	12	11
2040-2069	0	4	4	5	6	7	9
2070-2099	-8	-7	-6	-1	2	8	6

Appendix C - Relative Change for Different Return Period Events

Projected changes in 5-year, 20-year, 50-year and 200-year design flood values over the 1229 basins defined at each individual grid cell (Figures C1, C2, C3 and C4, respectively).

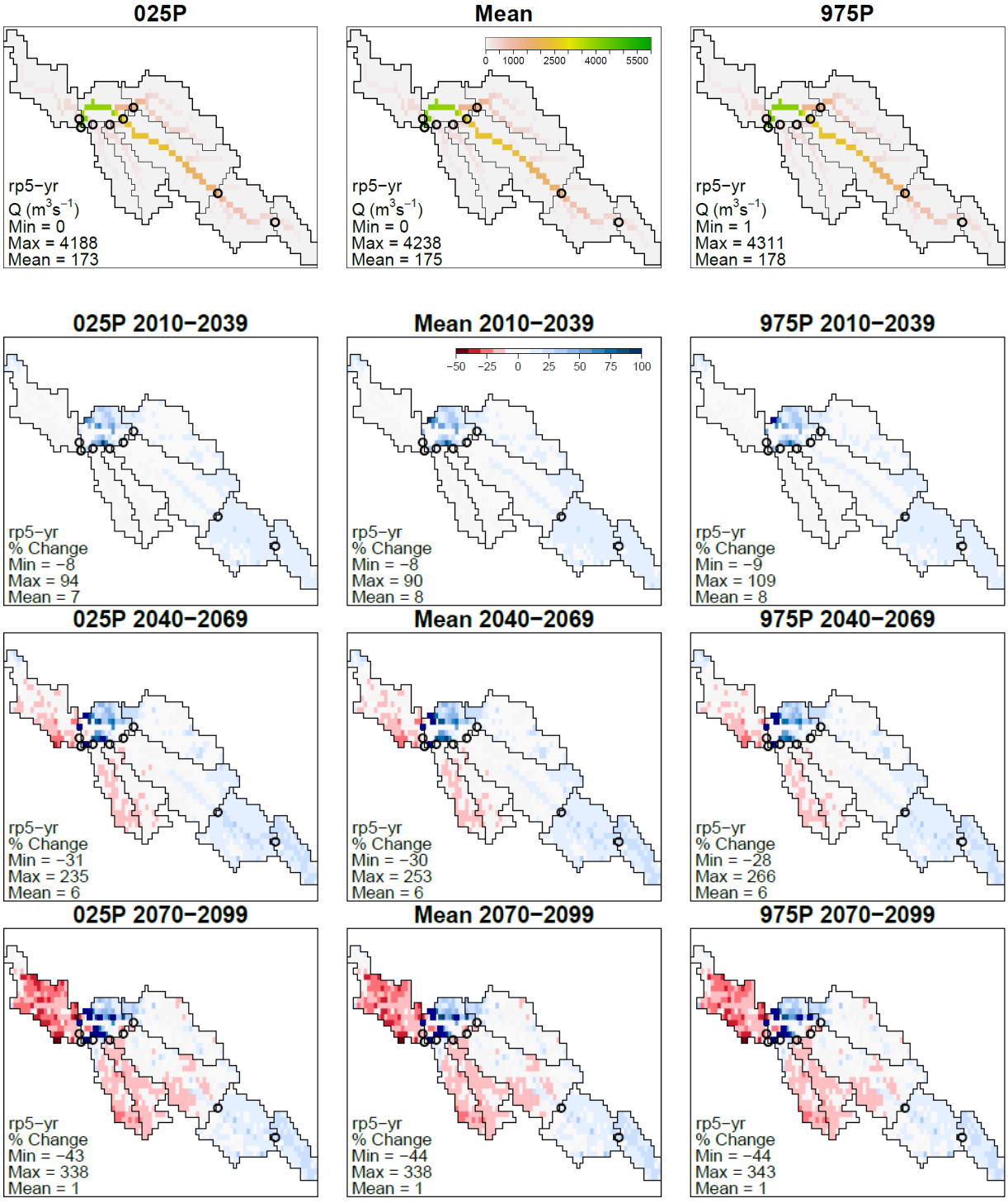

Figure C1 - Absolute (m^3/sec) and projected relative change (%) in annual peak flow magnitude for 5-yr return period events for 1961-1990 (top), 2010-2039, 2040-2069 and 2070-2099 versus the baseline period (1961-1990). Results are shown for the 2.5th percentile (left), best estimate (middle) and 97.5th percentile (right).

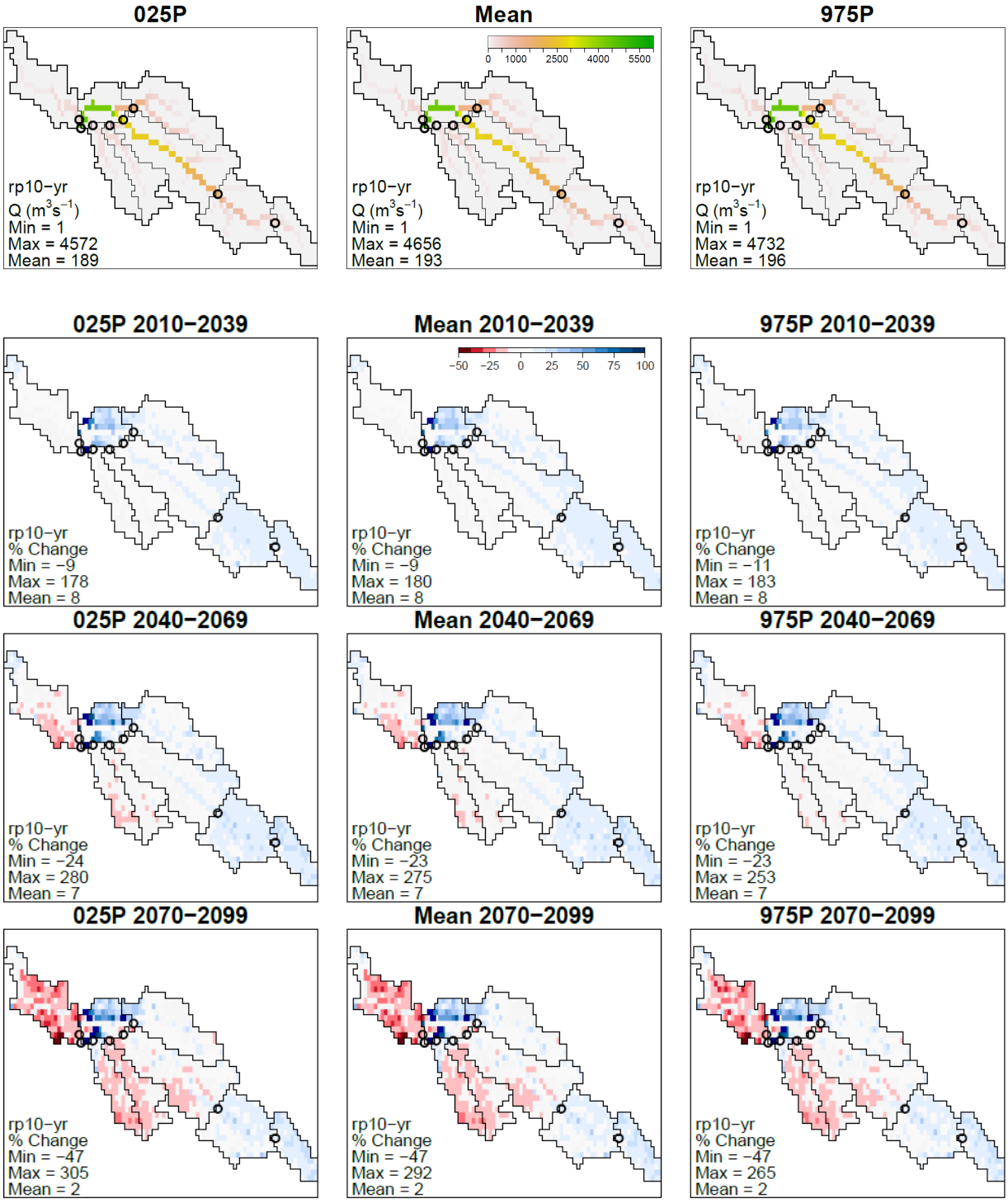

Figure C2 - Absolute (m^3/sec) and projected relative change (%) in annual peak flow magnitude for 5-yr return period events for 1961-1990 (top), 2010-2039, 2040-2069 and 2070-2099 versus the baseline period (1961-1990). Results are shown for the 2.5th percentile (left), best estimate (middle) and 97.5th percentile (right).

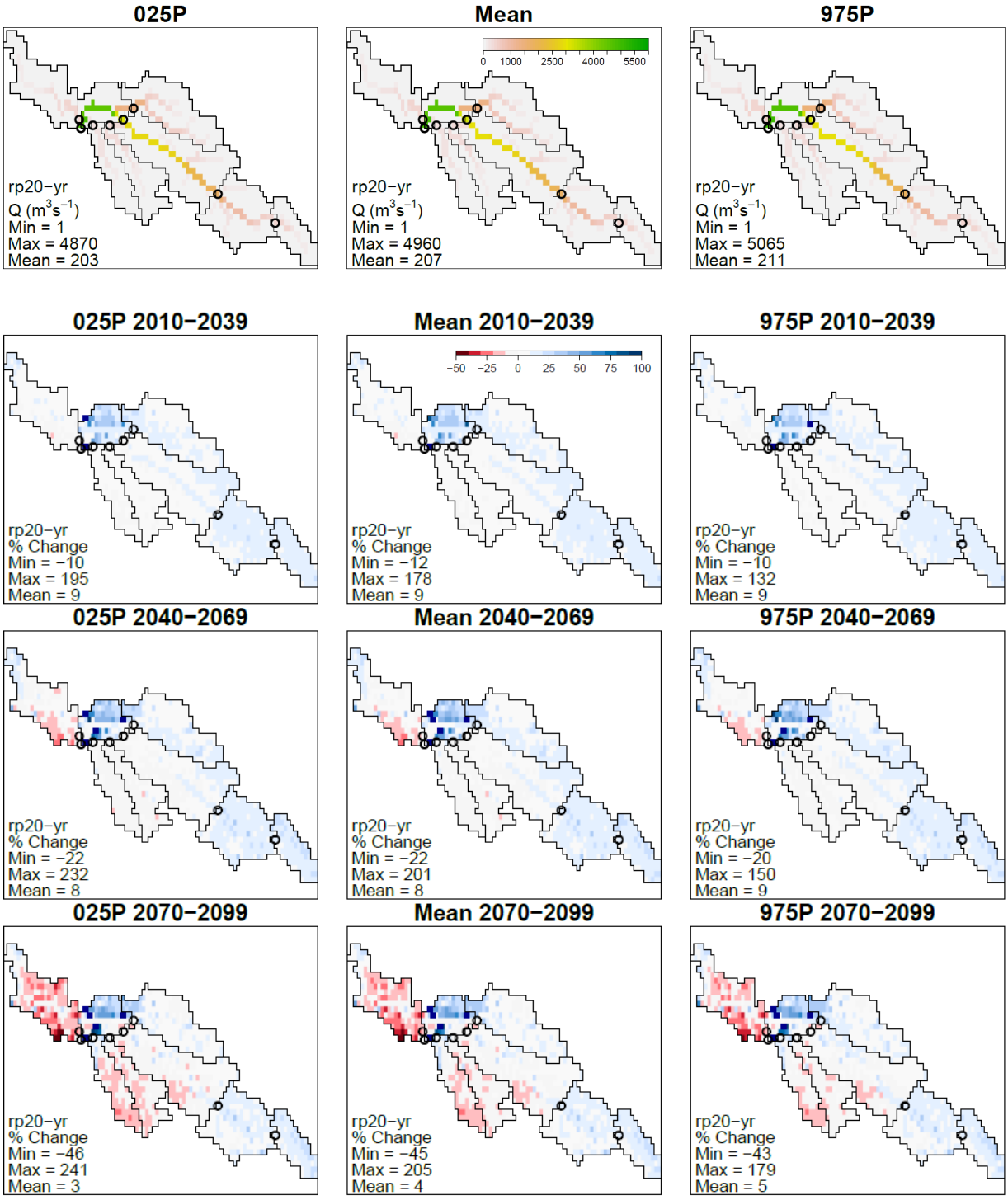

Figure C3 - Absolute (m^3/sec) and projected relative change (%) in annual peak flow magnitude for 20-yr return period events for 1961-1990 (top), 2010-2039, 2040-2069 and 2070-2099 versus the baseline period (1961-1990). Results are shown for the 2.5th percentile (left), best estimate (middle) and 97.5th percentile (right).

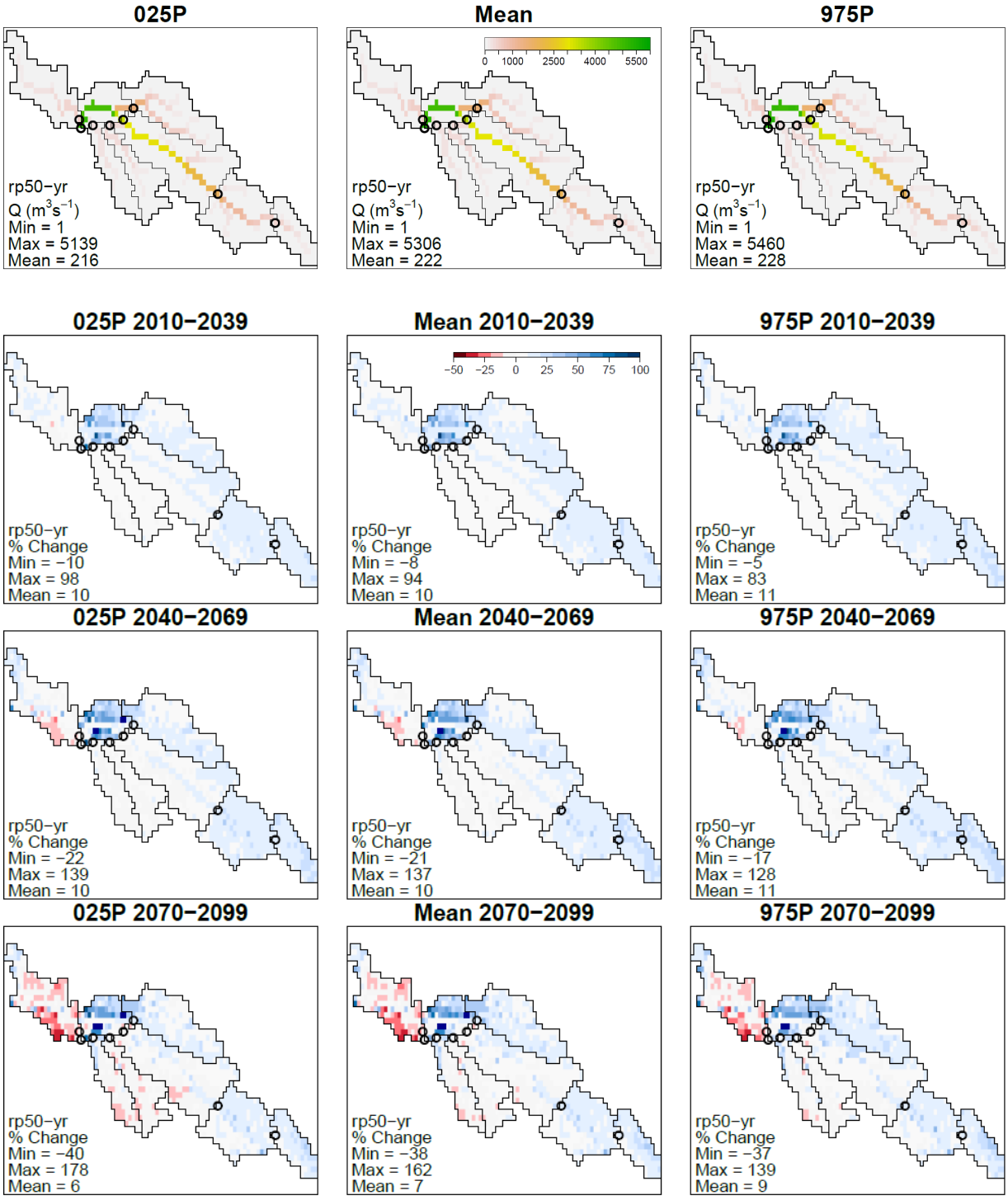

Figure C4 - Absolute (m^3/sec) and projected relative change (%) in annual peak flow magnitude for 50-yr return period events for 1961-1990 (top), 2010-2039, 2040-2069 and 2070-2099 versus the baseline period (1961-1990). Results are shown for the 2.5th percentile (left), best estimate (middle) and 97.5th percentile (right).

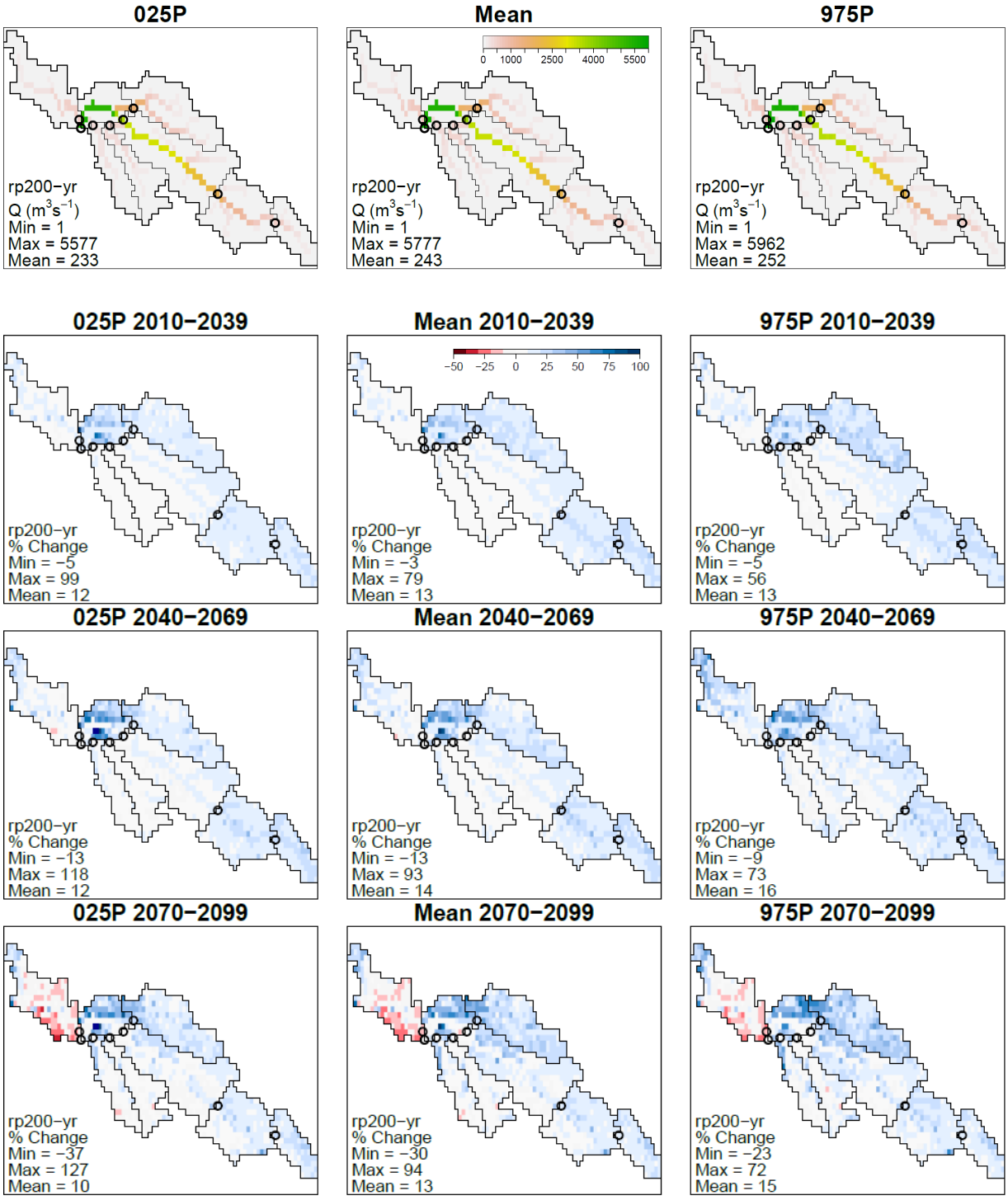

Figure C5 - Absolute (m^3/sec) and projected relative change (%) in annual peak flow magnitude for 200-yr return period events for 1961-1990 (top), 2010-2039, 2040-2069 and 2070-2099 versus the baseline period (1961-1990). Results are shown for the 2.5th percentile (left), best estimate (middle) and 97.5th percentile (right).

Appendix D – Comparison of Annual Maximum Flow between the CanESM2-LE and PCIC6 Ensembles

This section presents a comparison of simulated annual maximum peak flow by sub basin using climate projections provided by two different ensembles: the CanESM2-LE and PCIC6. The PCIC6 ensemble is composed of six GCMs (ACCESS1-0, CanESM2, CCSM4, CNRM-CM5, HadGEM2, and MPI-ESM-LR). The CanEMS2-LE is composed of 50 ensemble members based on different initializations of the CanESM2 model. The PCIC6 ensemble is composed of 15 runs made up of the six aforementioned GCMs, with multiple runs for select GMCs (5 x CanESM2, 4 x HadGEM2, 3 x MPI-ESM-LR and 2 x CCSM4). All projections are based on the RCP8.5 emissions trajectory.

Results are plotted for the outlets corresponding to the FRSRP, FRSMC, FRSHA, MCGRE, WILLO, BOWRB, SALMO and FRSHO sub basins in Figures D1 through D8, respectively. At all locations the spread of the ensemble results (indicated by the ensemble minimum and maximum) shows that the CanESM2-LE has a similar, if not slightly larger, range as the PCIC6 ensemble. Using the ensemble median, the CanESM2-LE annual maximum peak flow values tend to be larger than those obtained from the PCIC6 ensemble. Comparing trends, which are estimated by fitting a simple linear regression to the ensemble medians, we see that the direction of trend is identical between CanESM2-LE and PCIC6 at all locations. It is apparent, where the trend is positive (FRSRP, FRSMC), the CAnESM2-LE ensemble produces a stronger trend, whereas if the trend is negative (FRSHA, MCGRE, BOWRB and WILLO), CanESM2-LE produces a slightly weaker (less negative) trend.

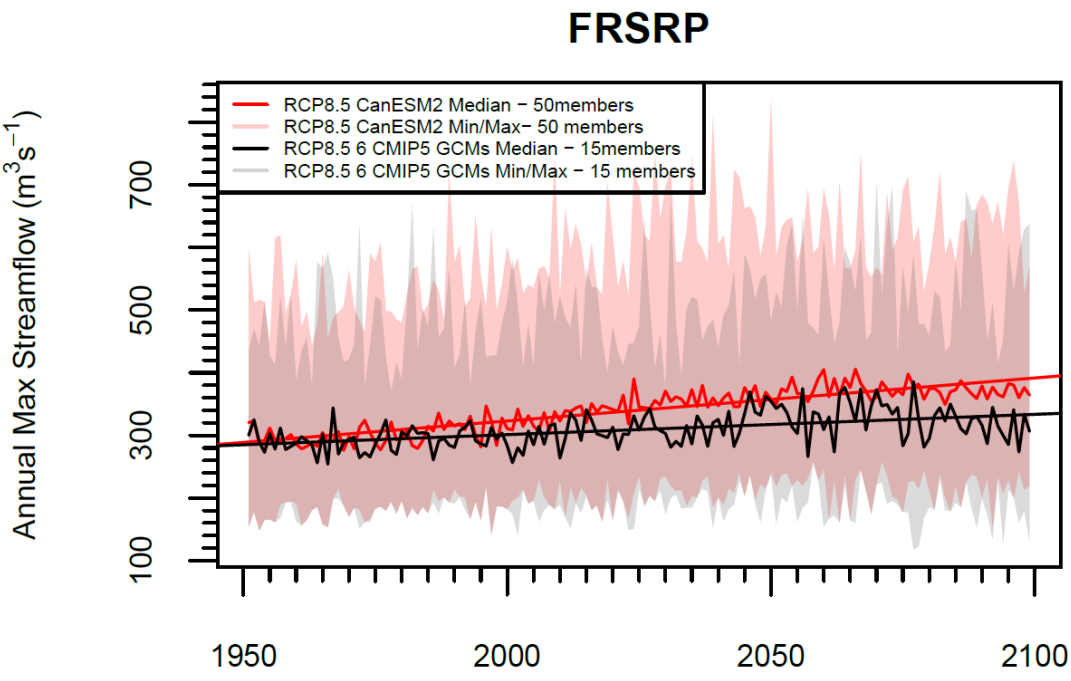

Figure D1- Annual maximum streamflow showing ensemble median (line) and minimum-maximum range (ribbon) of the CanESM2 Large Ensemble (pink/red) and the CMIP5 PCIC6 (grey/black) for RCP8.5 for FRSRP. Straight lines show fitted linear trend.

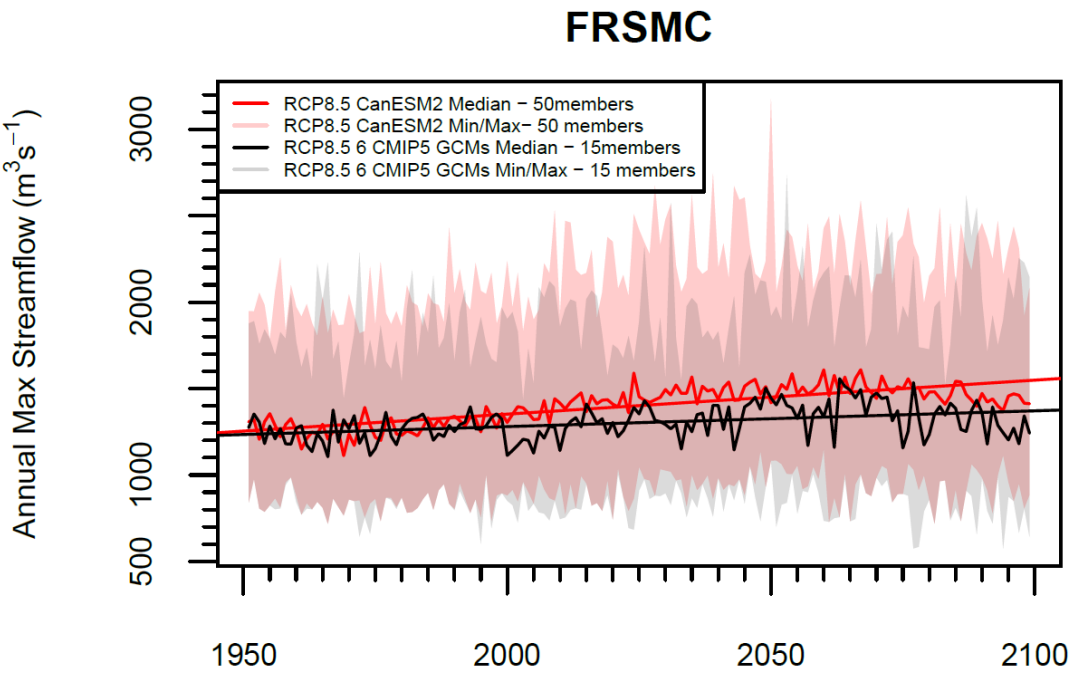

Figure D2 - Same as Figure D2, but for FRSMC.

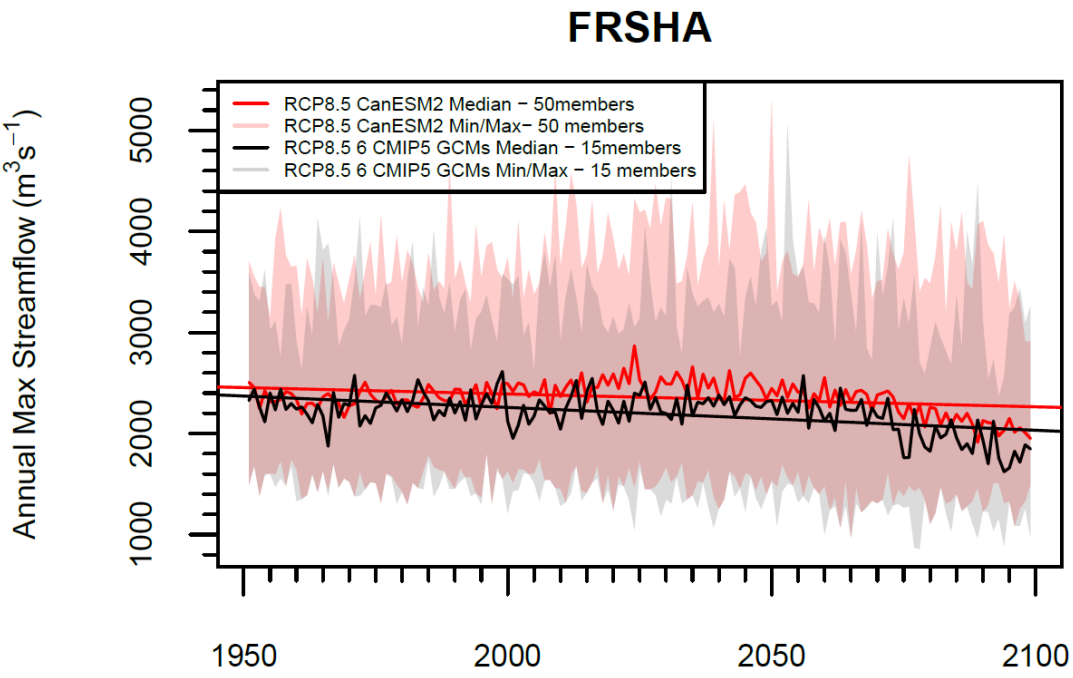

Figure D3 - Same as Figure D2, but for FRSHA.

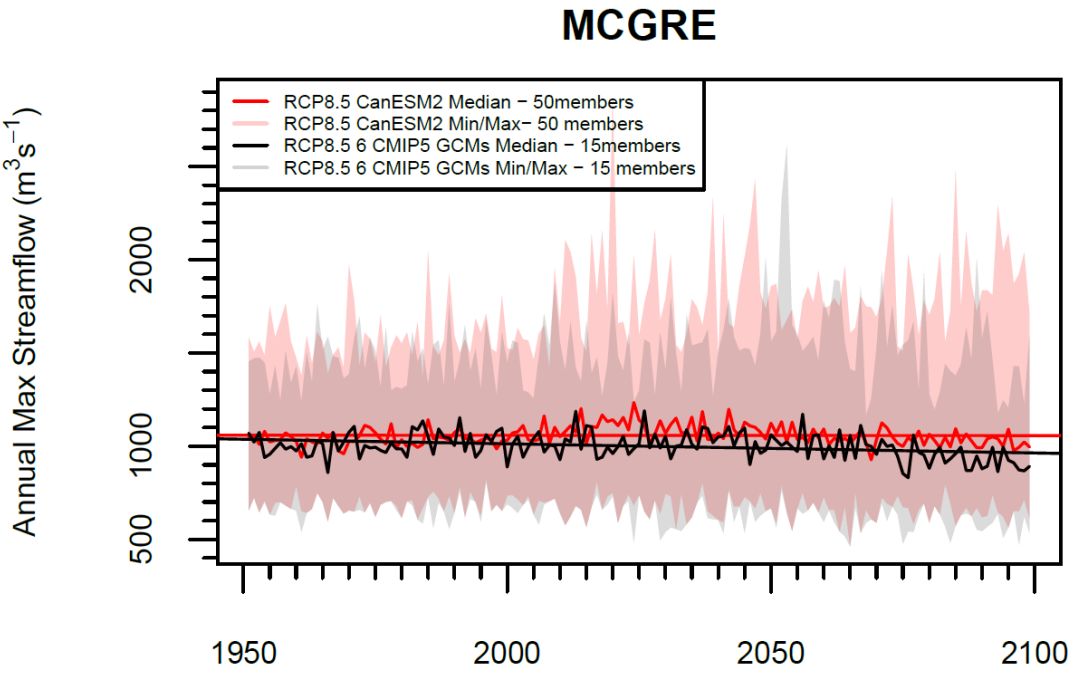

Figure D4 - Same as Figure D2, but for MCGRE.

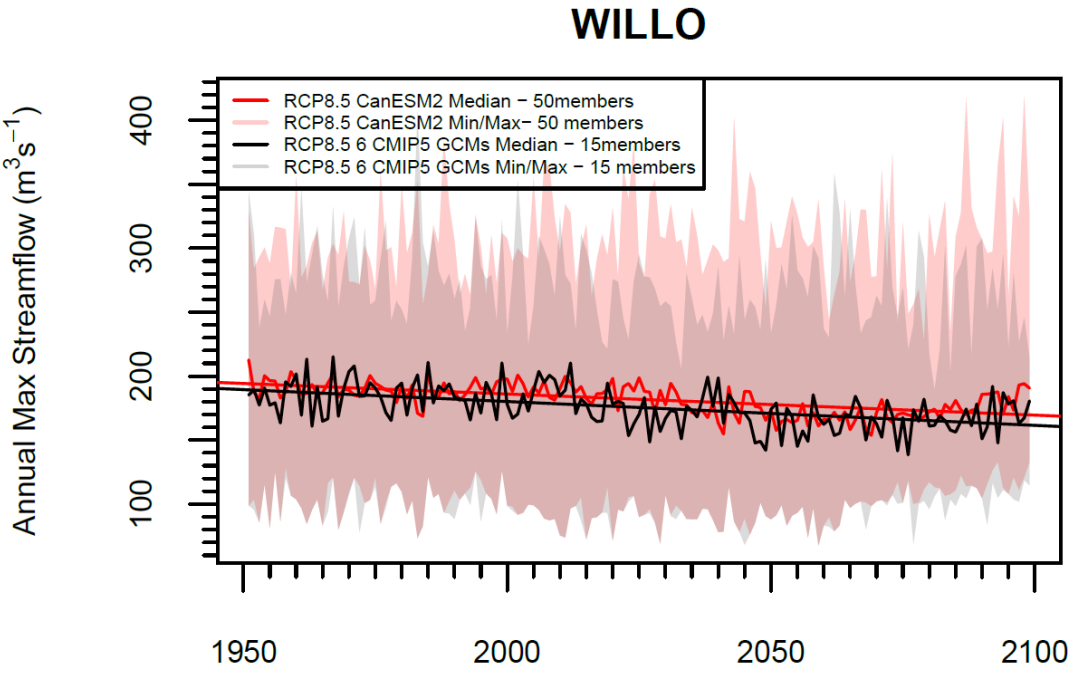

Figure D5 - Same as Figure D2, but for WILLO.

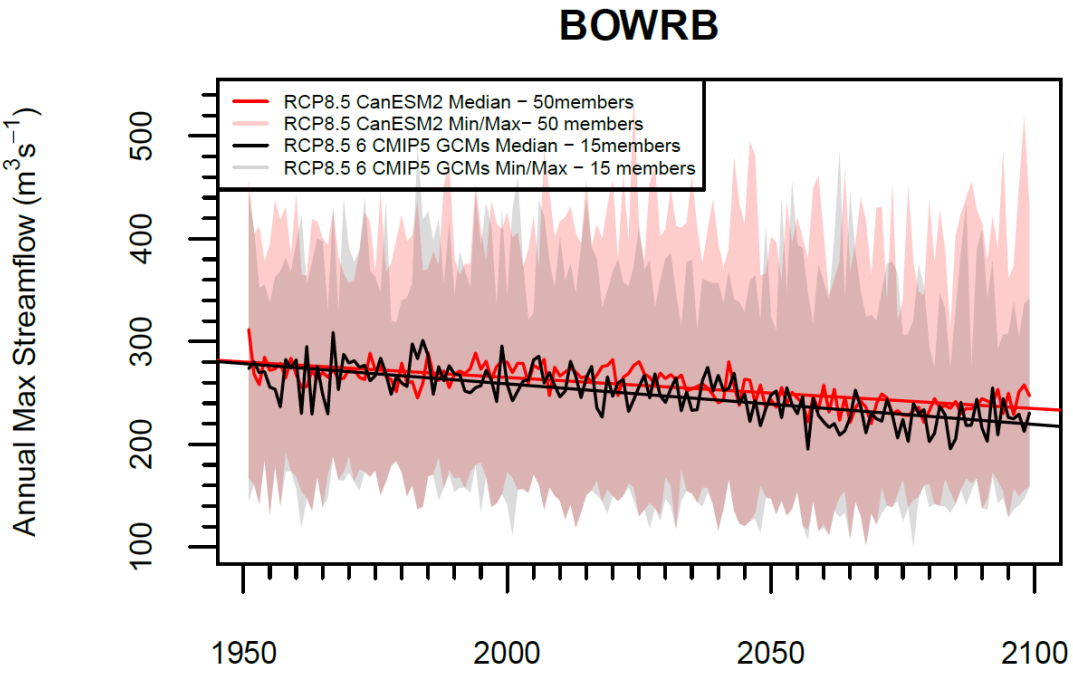

Figure D6 - Same as Figure D2, but for BOWRB.

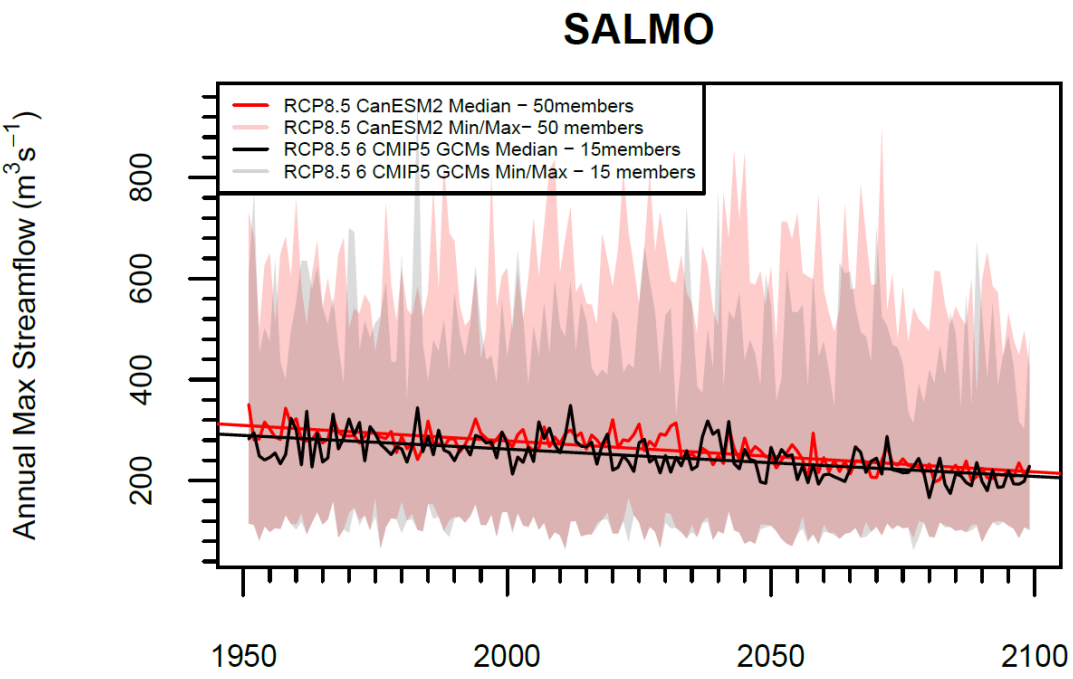

Figure D7 - Same as Figure D2, but for SALMO.

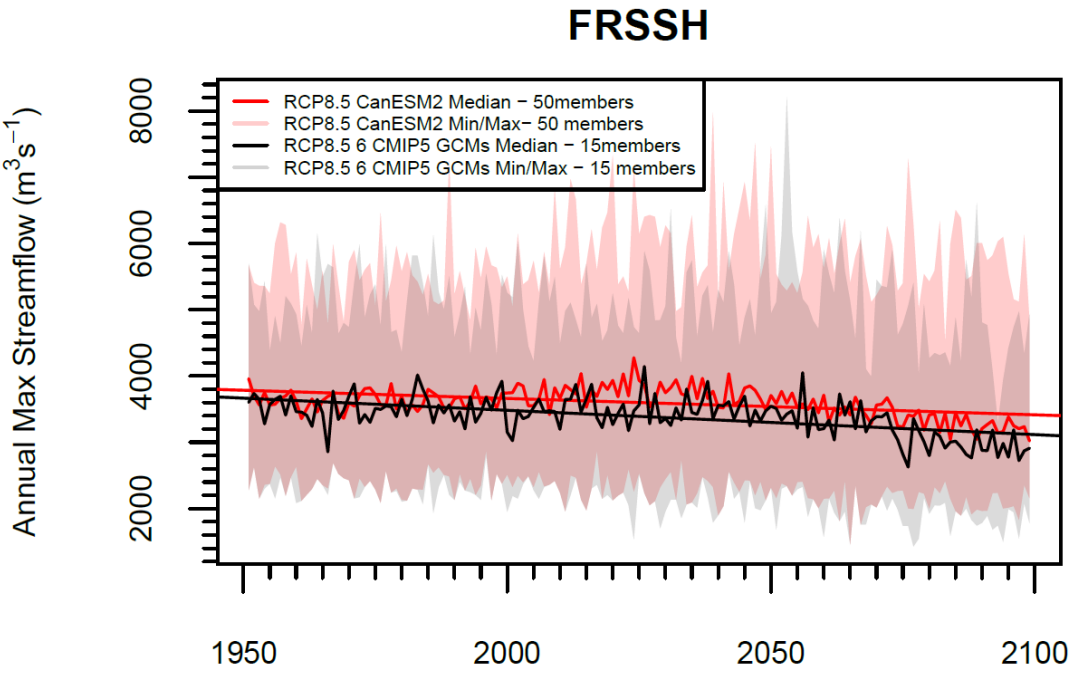

Figure D8 - Same as Figure D2, but for FRSSH.

Appendix E – Comparison of Empirical and Parametric Flood Frequency Analysis

Previous studies always fitted a parametric frequency distribution to a sample of block maximum events, such as the generalized extreme value distribution (GEV). The underlying asymptotic theory shows that the approximation should improve as the block length increases. However, there is a strong annual cycle in streamflow and daily amounts from which annual maxima are calculated are probably strongly serially correlated, resulting in that the effective block length is small. The GEV might not well approximate the upper tail of the distribution when fitted to such annual maxima.

To test this, estimated peak flow quantiles by fitting a generalized extreme value (GEV) distribution to the sample data and compared the results to those directed using the empirical estimate of the distribution function. As anticipated, the GEV-estimated quantiles show biases at high (T > 20-years) return periods compared to the empirically derived estimates, (Figure E1). Although unbiased, the tradeoff with the empirical estimates is that they have higher variance and require a wider confidence region than the GEV-based estimates (compare Figure E2 and Figure E3).

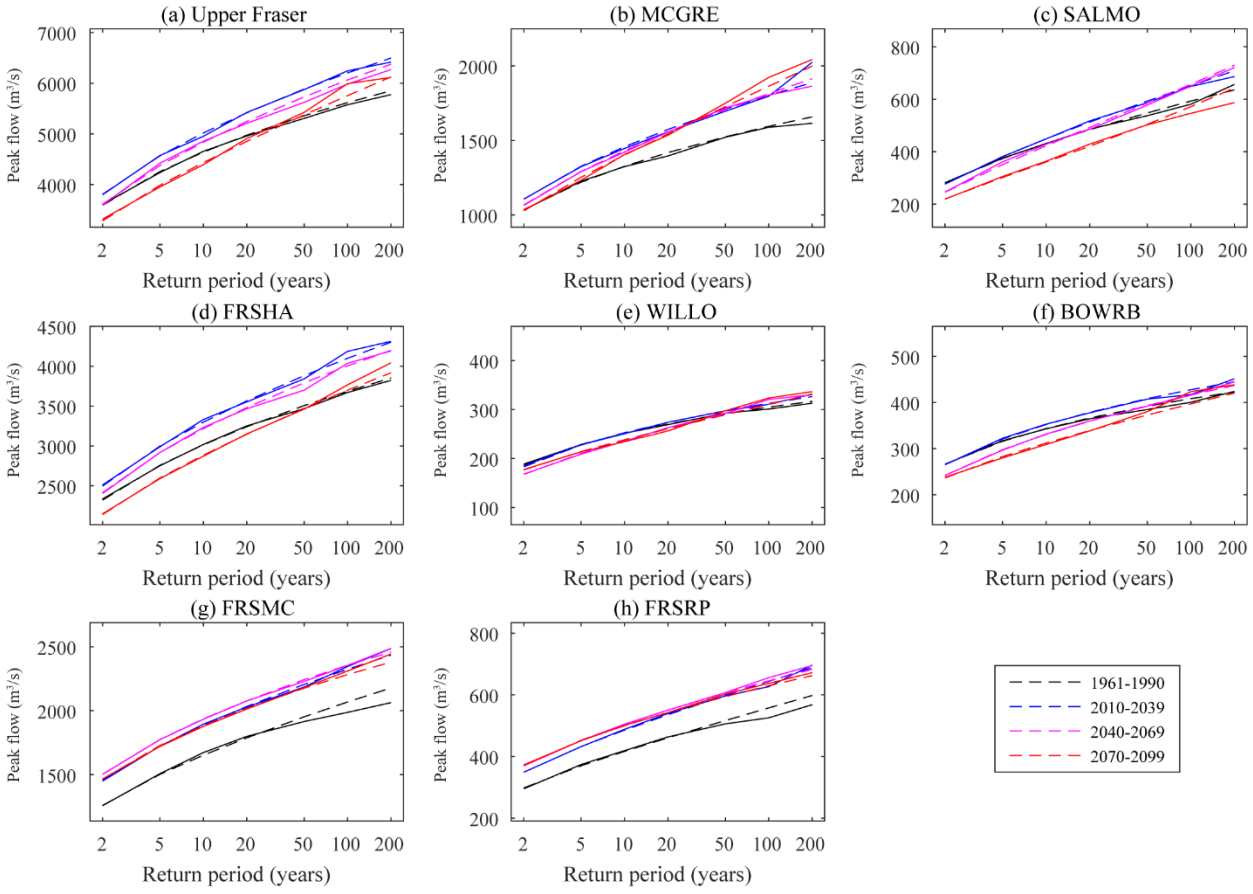

Figure E1. Flood frequency plot of results empirically (solid lines) and from GEV fitting (dotted lines) of annual peak daily flow

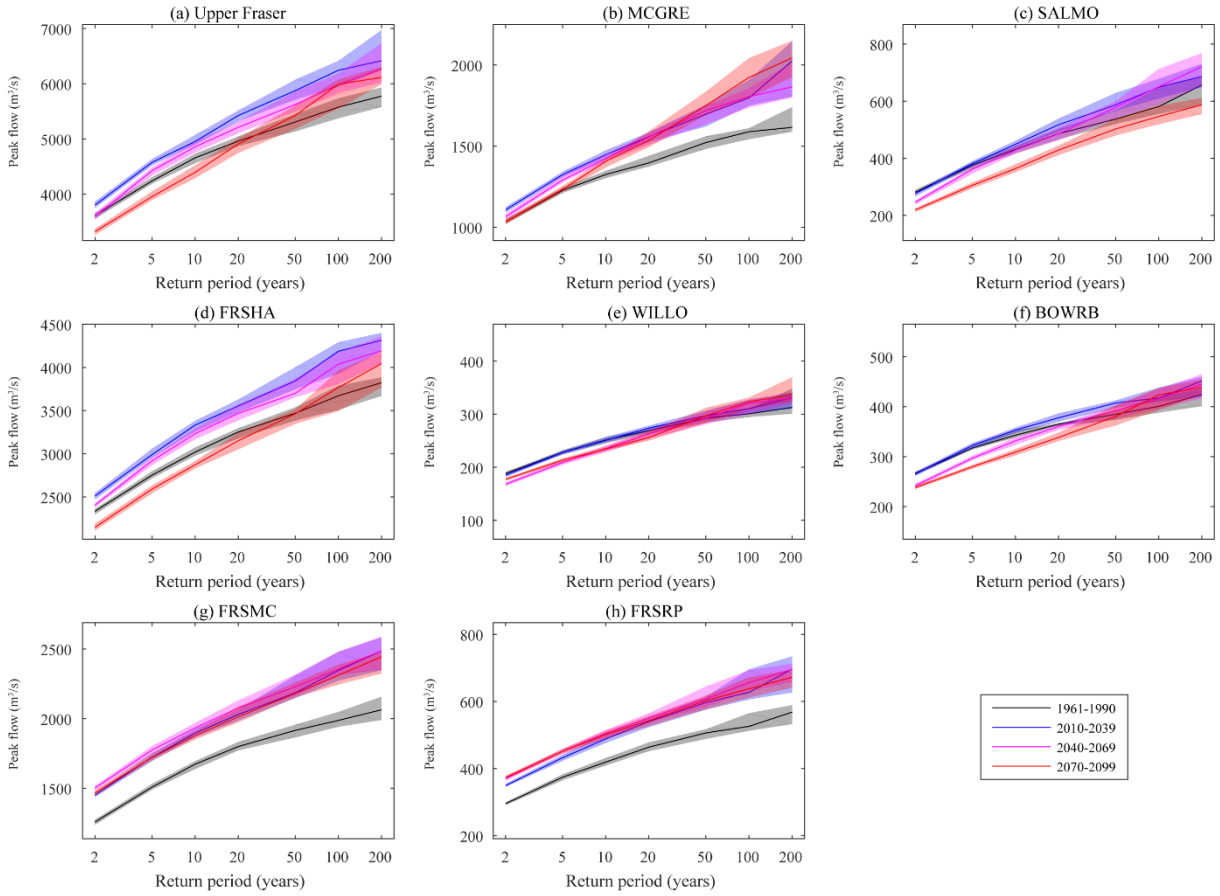

Figure E2. Flood frequency plot comparing empirical quantile estimates (solid line) with 95% confidence interval (shaded region) for four non-overlapping decades.

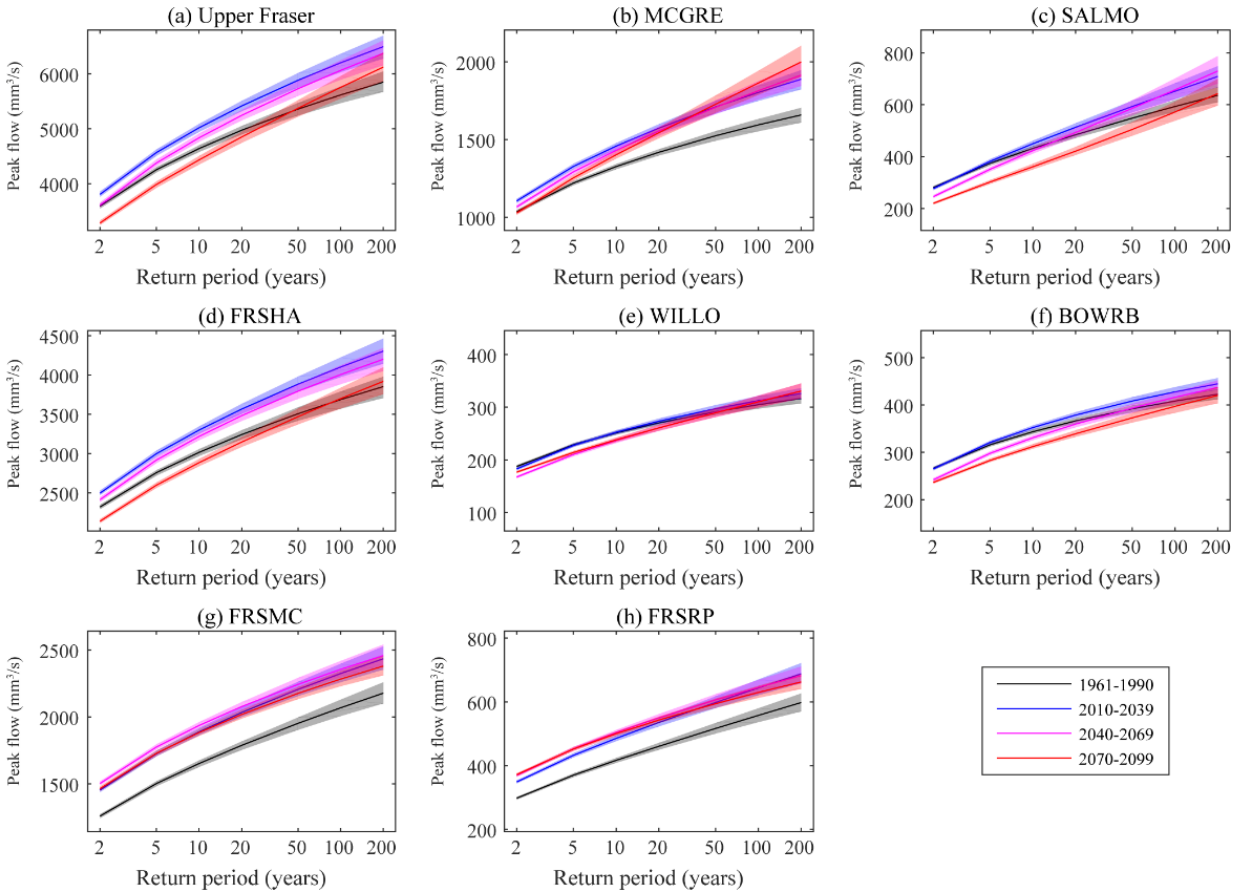

Figure E3. Flood frequency plot comparing of GEV fit quantile estimates (solid line) with 95% confidence interval (shaded region) for four non-overlapping decades.