Statistical emulation of streamflow projections from a distributed hydrological model: Application to CMIP3 and CMIP5 climate projections for British Columbia, Canada

Markus A. Schnorbus & Alex J. Cannon

2014

Pacific Climate Impacts Consortium (PCIC)

PCIC Publications

Original citation:

Schnorbus, M. A., & Cannon, A. J. (2014). Statistical emulation of streamflow projections from a distributed hydrological model: Application to CMIP3 and CMIP5 climate projections for British Columbia, Canada. *Water Resources Research*, *50*(11), 8907–8926. https://doi.org/10.1002/2014WR015279

AGU PUBLICATIONS

Water Resources Research

RESEARCH ARTICLE
10.1002/2014WR015279

Statistical emulation of streamflow projections from a distributed hydrological model: Application to CMIP3 and CMIP5 climate projections for British Columbia, Canada

Markus A. Schnorbus[1] and Alex J. Cannon[1]

[1]Pacific Climate Impacts Consortium, University of Victoria, Victoria, British Columbia, Canada

Key Points:
- Statistical emulation is an effective means of producing streamflow projections
- Regularized multiple linear regression provides a robust emulation model
- CMIP5-RCP and CMIP3-SRES streamflow projections are qualitatively similar

Supporting Information:
- Readme
- Figure fs01
- Table ts01
- Table ts02

Correspondence to:
M. A. Schnorbus,
mschnorb@uvic.ca

Citation:
Schnorbus, M. A., and A. J. Cannon (2014), Statistical emulation of streamflow projections from a distributed hydrological model: Application to CMIP3 and CMIP5 climate projections for British Columbia, Canada, *Water Resour. Res.*, *50*, 8907–8926, doi:10.1002/2014WR015279.

Received 9 JAN 2014
Accepted 20 OCT 2014
Accepted article online 27 OCT 2014
Published online 19 NOV 2014

Abstract A recent hydrological impacts study in British Columbia, Canada, used an ensemble of 23 climate change simulations to assess potential future changes in streamflow. These Coupled Model Intercomparison Project Phase 3 (CMIP3) simulations were statistically downscaled and used to drive the Variable Infiltration Capacity (VIC) hydrology model over several watersheds. Due to computational restrictions, the 23 member VIC ensemble is a subset of the full 136 member CMIP3 archive. Extending the VIC ensemble to cover the full range of uncertainty represented by CMIP3, and incorporating the latest generation CMIP5 ensembles, poses a considerable computing challenge. Thus, we extend the VIC ensemble using a computationally efficient statistical emulation model, which approximates the combined output of the two-step process of statistical downscaling and hydrologic modeling, trained with the 23 member VIC ensemble. Regularized multiple linear regression links projected changes in monthly temperature and precipitation with projected changes in monthly streamflow over the Fraser and Peace River watersheds. Following validation, the statistical emulator is forced with the full suite of CMIP3 and CMIP5 climate change projections. The 23 member VIC ensemble has a smaller spread than the full ensemble; however, both ensembles provide the same consensus estimate of monthly streamflow change. Qualitatively, CMIP5 shows a similar streamflow response as CMIP3 for snow-dominated hydrologic regimes. However, by end-century, the CMIP5 worst-case RCP8.5 has a larger impact than CMIP3 A2. This work also underscores the advantage of using emulation to rapidly identify those future extreme projections that may merit further study using more computationally demanding process-based methods.

1. Introduction

According to the Intergovernmental Panel on Climate Change (IPCC) Fifth Assessment Report, it is now "extremely likely" that observed widespread warming of the atmosphere and oceans since the mid-twentieth century is due to anthropogenic net increases in atmospheric greenhouse gas concentrations [*IPCC*, 2013]. Recent hydro-climatic trends in western North America have been attributed to climate change, mainly in the form of increased regional warming [*Barnett et al*., 2008; *Bonfils et al*., 2008; *Pierce et al*., 2008]. Within the snow-dominated regimes of western North America, documented hydrologic trends over recent decades generally include decreasing snowpack, earlier onset of spring melt and decreasing summer flow [*Whitfield and Cannon*, 2000; *Foreman et al*., 2001; *Barnett et al*., 2005; *Regonda et al*., 2005; *Hamlet et al*., 2007; *Déry et al*., 2009; *Luce and Holden*, 2009; *Stewart*, 2009; *Fleming and Weber*, 2012; *Hatcher and Jones*, 2013; *Bawden et al*., 2014]. These trends will persist with continued emissions of greenhouse gases, and further changes in regional temperature and precipitation patterns are expected to affect the regional hydrologic cycle [*Morrison et al*., 2002; *Barnett et al*., 2005; *Toth et al*., 2006; *Stahl et al*., 2008; *Kerkhoven and Gan*, 2011; *Shrestha et al*., 2012; *Schnorbus et al*., 2014], with possible impacts to various water-related resources and water-dependent activities [*Cohen et al*., 2000; *Payne et al*., 2004; *Hamlet et al*., 2010; *Mantua et al*., 2010; *Vano et al*., 2010a, 2010b; *Jung et al*., 2011]. Effective water resource management and planning requires an understanding of the impacts of climate change in order to accurately assess the volume and timing of future water supplies.

Possible future changes in hydrology are typically assessed using projections based on climate change scenarios. The common approach is to choose a climate change projection, which is a combination of a global climate model (GCM) driven by an emission scenario specifying trajectories of greenhouse gas emissions or

radiative forcings, downscale the climate change projection to the regional or catchment scale, force a hydrologic model with the downscaled climate projection, and compare model simulations from both current and future climates [e.g., *Elsner et al*., 2010; *Schnorbus et al*., 2014].

Each step introduces a component of uncertainty, the contributions of which are typically assessed in an ensemble framework using some combination of multiple GCMs, emission scenarios, downscaling techniques, and hydrology models. Many such studies indicate that GCM structure is the primary source of uncertainty for the evaluation of hydrologic impacts [*Kay et al*., 2009; *Prudhomme and Davies*, 2009; *Najafi et al*., 2011; *Bennett et al*., 2012; *Surfleet and Tullos*, 2012; *Vano et al*., 2014]. Therefore, eight GCMs that contributed to the World Climate Research Programme's Coupled Model Intercomparison Project Phase 3 (CMIP3) multimodel data set [*Meehl et al*., 2007] were selected, based on their historical performance over the globe and more locally over western North America [*Werner*, 2011] (based on the work of *Gleckler et al*. [2008] and *Radić and Clarke* [2011]), as the basis for hydrological impacts studies recently conducted at the Pacific Climate Impacts Consortium (PCIC) for several watersheds in British Columbia (BC), Canada [*Shrestha et al*., 2012; *Schnorbus et al*., 2014]. Based on this work, an ensemble of 23 climate change projections, consisting of multiple GCMs and the A2, A1B, and B1 emissions scenarios (Special Report on Emissions Scenarios (SRES)) [*Nakićenović and Swart*, 2000], were statistically downscaled and used to drive the Variable Infiltration Capacity (VIC) hydrologic model.

The 23 member ensemble used by PCIC is a subset of the full CMIP3 archive, from which PCIC has obtained 136 projections (22 GCMs, 3 SRES scenarios, and up to 7 ensemble members per GCM/forcing scenario combination). In addition to CMIP3 outputs, which contributed to IPCC Fourth Assessment Report, the latest generation of climate projections from the Coupled Model Intercomparison Project Phase 5 (CMIP5) multimodel data set [*Taylor et al*., 2012], which are based on the new Representative Concentration Pathways (RCP) scenarios [*van Vuuren et al*., 2011a], are now also available and have contributed to the IPCC AR5.

Given the large role of GCM uncertainty, are streamflow projections based on the 23 member VIC-simulated ensemble representative of the full range of uncertainty from the 136 member CMIP3 ensemble? Further, do the more recent CMIP5 projections, which are based on different emission scenarios, and, in most cases, new models, present a different picture of future streamflow changes? The most direct means of answering these questions would be to force VIC with the additional climate change projections and then evaluate the mean and spread of the resulting streamflow projections ensemble. However, the computational burden and storage demands to run more than just a handful of hydrologic projections using an analytical model such as VIC can be impractical and expensive for many organizations, particularly if such models are run at high resolution over large study domains.

As an alternative, a computationally efficient statistical model could be used to emulate the streamflow response simulated by the VIC model for the 23 CMIP3 climate projections. One could then force the statistical emulator with additional climate change projections from both CMIP3 and CMIP5. The use of statistical emulation (also called surrogate modeling or metamodeling) is common in many areas of science, where large complex computer simulation models are replaced with computationally efficient models that emulate well the behavior of the simulation model [see *Barton*, 1998; *O'Hagan*, 2006; *Kleijnen*, 2008; *Carnevale et al*., 2012; *Villa-Vialaneix et al*., 2012, for a general discussion]. Although one could fit a statistical model directly to hydro-climatic variation observed over the last decades [e.g., *Cannon and Whitfield*, 2002], such a model would have to be extrapolated unreasonably far beyond the bounds of the calibration data to investigate potential future hydrologic change under radiative forcing as high as $\sim$8 W/m^2 above preindustrial levels (e.g., emissions scenarios SRES A2 and RCP 8.5). Instead, future hydrologic change is extrapolated using a more robust process-based hydrologic simulation approach, exploring a wide range of emissions futures. By then fitting the statistical model to the simulated output, exploring the hydrologic response to the full range of available climate projections is more an exercise of interpolation, and when extrapolation is required, it falls closer to the range of the calibration data. For example, *Holden and Edwards* [2010], *Li et al*. [2012], and *Castruccio et al*. [2013] employed emulation of global and regional climate models to expand projections of temperature and precipitation change. Except for *Vano and Lettenmaier* [2013], who used a sensitivity-based approach to emulate the cumulative distribution functions of simulated annual streamflow change based on CMIP3 projections, there are no examples of similar work in the context of hydrologic projections. In the water resources field, model emulation is most commonly applied to problems of optimization and design [e.g., *Aly and Peralta*, 1999; *Broad et al*., 2005; *Yan and Minsker*, 2006],

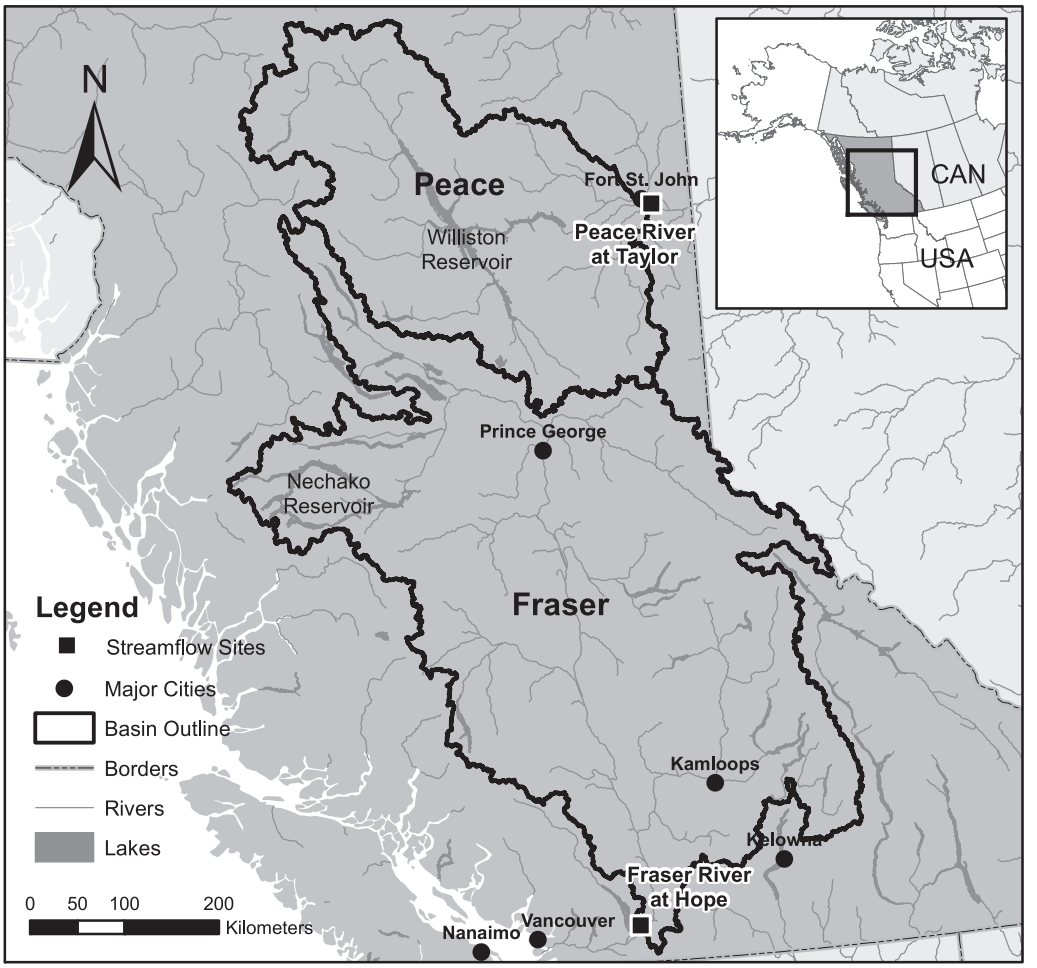

Figure 1. Study areas showing the streamflow sites and upstream drainage areas for the Fraser River at Hope and Peace River at Taylor in British Columbia, Canada. The inset map shows the study area in relation to western North America.

automatic model calibration [e.g., *Zou et al*., 2009; *Song et al*., 2012], and uncertainty and sensitivity analysis [e.g., *Schultz et al*., 2004, 2006; see also *Razavi et al*., 2012a, 2012b, for recent reviews]. To the best of our knowledge, an emulation approach has not been explored to date in the context of analyzing seasonal and monthly hydrologic impacts due to climate change for both CMIP3 and CMIP5 projections.

This paper presents a study of climate-driven changes in streamflow for two watersheds in BC. The goal is to demonstrate, using a relatively simple proof-of-concept, the general efficacy of using an emulation approach to efficiently generate large ensembles of hydrologic projections. The main strategy is to construct a statistical emulator that accurately captures the monthly streamflow response as simulated by the original analytical VIC model. We then propose using the emulator to facilitate the exploration of the streamflow response to the full range of available CMIP3 and CMIP5 climate projections. The remainder of the paper is organized as follows: section 2 describes the study areas; section 3 describes the methodology, including data processing and the emulation model; results and discussion are contained in section 4; and concluding remarks are in section 5.

2. Study Areas

Streamflow was modeled for two locations in BC, the Peace River near Taylor and the Fraser River at Hope (Figure 1). The Peace River site drains 101,000 km^2 of north-eastern BC and corresponds to the Water Survey of Canada (WSC) gauge (07FD002) located near the town of Taylor (just upstream of the provincial border between BC and Alberta). The drainage area has an elevation range of 400–2800 m. The Peace River basin

has a continental climate [*Demarchi*, 1996], with basin-wide average temperature ranging from −12.0°C in January to 12°C in July, and precipitation follows a seasonal pattern of summer maximum and spring minimum. The Peace River has a nival hydrologic regime, with average monthly discharge that peaks in June at 4750 m^3/s and reaches a minimum of 300 m^3/s in February and March. The Fraser River site drains a 217,000 km^2 area of the Fraser River basin in south-central BC and the site corresponds to the WSC gauge (08MF005) located at the town of Hope (approximately 120 km upstream from the city of Vancouver, where the Fraser discharges into the Pacific Ocean). Elevation in the basin varies from near sea level to about 4000 m. The Fraser River basin has a mixed coastal-continental climate [*Moore*, 1991]; basin-wide monthly average temperature varies from −10°C in January to 15°C in July and precipitation exhibits some seasonal variation, with maximum precipitation occurring during the winter, and minimum precipitation during the summer in the western coastal mountains, and during spring in the central and eastern portions of the basin [*Shrestha et al*., 2012]. Streamflow at the Fraser River site also has a nival regime, where average monthly discharge ranges from 7000 m^3/s in June to only 900 m^3/s throughout January, February, and March. Both the Peace and Fraser basins are regulated for hydro-power generation. The Peace River is impounded by the W.A.C. Bennett Dam to form the Williston Reservoir. While the main stem of the Fraser remains undammed, the Nechako River, a tributary of the Fraser, is impounded by the Kenney Dam to form the Nechako Reservoir, from which water is diverted westward directly to the BC coast. More details on the study areas can be found in *Shrestha et al*. [2012] and *Schnorbus et al*. [2014]. These two watersheds were deliberately chosen for their tractable nature, i.e., snow-dominated runoff, minimal glacier influence at the basin outlets (see section 3.2), large enough basin areas to robustly capture the regional GCM climate changes (i.e., integrated over multiple GCM grid cells), and, when using naturalized discharge, not affected by flow regulation.

3. Methods

3.1. Data and Processing

GCM output of historical and future projected monthly mean temperature and precipitation was obtained from the CMIP3 and CMIP5 data archives at the Program for Climate Model Diagnosis and Intercomparison (http://www-pcmdi.llnl.gov/). CMIP3 outputs include data from 22 GCMs, which are summarized in supporting information Table S1 (full set). Three future emissions scenarios are considered for the CMIP3 models, namely SRES A1B, A2, and B1, of which 51, 38, and 44 projections are available, respectively. At the time of writing, output for CMIP5 included results from 27 models, which are summarized in supporting information Table S2. Future emissions for the CMIP5 model output are based on 38, 48, and 48 projections from RCP scenarios 2.6, 4.5, and 8.5, respectively [*van Vuuren et al*., 2011a]. Baseline climatologies for both CMIP3 and CMIP5 model runs uses data temporally averaged over the 1961–1990 period (referred to as the 1970s) based on historical twentieth century simulations. The historical simulations represent natural and anthropogenic forcings over the periods 1850–2000 and 1850–2005 for CMIP3 and CMIP5, respectively. The CMIP3 and CMIP5 historical runs differ in that the CMIP3 runs contain a less complete and complex description of historical radiative forcings. Less than half of CMIP3 models include time-varying natural radiative forcings (e.g., solar, volcanic, and land-use change) and most CMIP3 models only incorporate direct aerosol effects [*Rind et al*., 2009; *Knutson et al*., 2013]. GCM projections for 2001–2100 for the respective emissions scenarios (SRES and RCP for CMIP3 and CMIP5, respectively) are used to derive 30 year average monthly temperature and precipitation changes (with respect to the 1970s baseline) for three future periods centered on the 2020s (2011–2040), 2050s (2041–2070), and 2080s (2071–2100). Basin-average values of monthly temperature and precipitation change for each future period were extracted by a weighted spatial averaging of GCM grid cells that overlap the drainage areas upstream of each respective site.

Projected daily streamflow data for both the Fraser and Peace come from the work of *Shrestha et al*. [2012] and *Schnorbus et al*. [2014], and are based on simulations of the Variable Infiltration Capacity (VIC) model, a spatially distributed process-based and macroscale hydrology model [*Liang et al*., 1994, 1996]. A total of 23 simulations is available, in which the VIC model was forced with downscaled climate projections from eight CMIP3 GCMs and the SRES A1B, A2, and B1 scenarios representing future emissions (see PCIC Sub-Set in supporting information Table S1). Using transient output of monthly average temperature and total precipitation from each GCM, statistical downscaling with the Bias-Correction Spatial Disaggregation (BCSD) approach [*Salathé*, 2005] was used to generate daily values of maximum and minimum temperature and

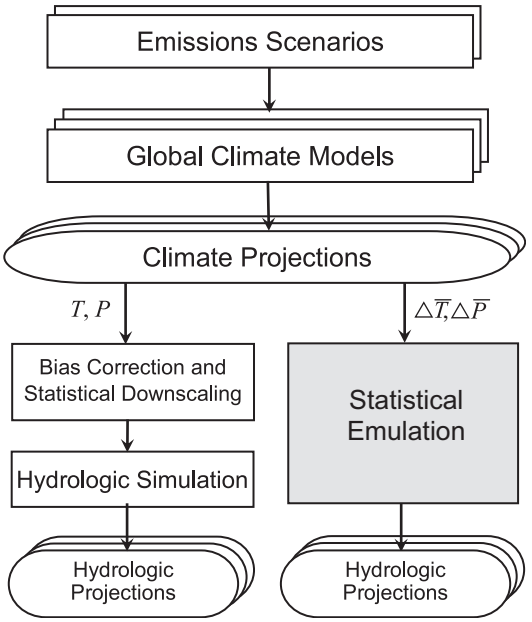

Figure 2. Flowchart of hydrologic simulation and statistical emulation.

precipitation at 1/16° spatial resolution for the period 1950–2100, as described in *Schnorbus et al*. [2014]. Forced with these boundary conditions, VIC calculates the one-dimensional (i.e., vertical) water and energy balance in each grid cell, where excess moisture is routed through a surface drainage network using the routing scheme of *Lohmann et al*. [1996]. These simulations were performed at a daily time step. Additional meteorological drivers, such as daily solar radiation (direct and diffuse), longwave radiation, and dewpoint temperature, are not supplied directly as input and were calculated by the VIC model at runtime from the downscaled time series of daily maximum and minimum temperature and precipitation using techniques described by *Maurer et al*. [2002]. Therefore, the fundamental climate forcing for the twentieth century hydrologic projections is GCM monthly average temperature and total precipitation.

Daily streamflow data were simulated from 1950 to 2100 for all climate projections. Note that discharge simulated for both the Peace and Fraser Rivers represents naturalized discharge (i.e., absent the effects of regulation and diversion) [see *Shrestha et al*., 2012; *Schnorbus et al*., 2014, for details] based on unpublished data from BC Hydro (Peace) and the BC Ministry of Environment (Fraser). We then calculated 30 year averages of monthly streamflow change between the 1961 and 1990 baseline period and the three future periods centered on the 2020s, 2050s, and 2080s for each site.

3.2. Emulator Description

An emulator is a statistical approximation of a simulator, where a simulator is a deterministic input-output computer model (or chain of models) [*O'Hagan*, 2006]. In this study, an emulator has been developed to approximate the combined output of the two-step process of statistical downscaling and hydrologic modeling (the simulator) (Figure 2). In this way, we aim to emulate streamflow changes, as would be simulated by the VIC model in response to downscaled projected climate forcings, but with a much more computationally efficient surrogate model. The emulator is based directly on the statistical relationship between basin-wide projected changes in monthly temperature and precipitation at the GCM resolution and VIC-simulated changes in monthly streamflow at the Peace and Fraser River sites.

We approach the design of the emulator from the perspective of the monthly water balance, which is given by the following relationship:

$$R_m = P_m - E_m - \Delta S_m, \tag{1}$$

where R_m is runoff, P_m is precipitation, E_m is evaporation, and ΔS_m is the change in storage over month m [where Δ is the difference operator, such that $\Delta f(x) = f(x+1) - f(x)$]. We assume a closed basin such that all runoff exits the single basin outlet as streamflow. There is a modest routing effect in both the Fraser and Peace basins, such that approximately 75% of runoff generated within a month reaches the basin outlet as streamflow (estimated assuming uniform daily runoff and convolution of the derived unit hydrograph for each basin). However, the structure of the statistical emulator model will accommodate any lag between runoff and streamflow, therefore, for simplicity we take streamflow as equivalent to runoff in any given month m. On an annual basis, ΔS is typically taken as zero [e.g., *Oki et al*., 1995], thus greatly simplifying the relationship in equation (1). For instance, using Gravity Recovery and Climate Experiment (GRACE) data in the form of monthly mass grids for land (obtained from the Jet Propulsion Laboratory website at http://grace.jpl.nasa.gov/data/gracemonthlymassgridsland/) shows that the long-term (October 2002 to September 2010) change in total water storage over the Fraser-Peace (318,000 km^2) is negligible (0.3 ± 2 cm water

equivalent) [*Swenson and Wahr*, 2006; *Landerer and Swenson*, 2012]. On a monthly basis, however, the change in storage is nonnegligible and must be retained due to the seasonal nature of snow and soil moisture. With respect to some baseline period, the change in runoff in a future period is given as

$$\Delta R_m = \Delta P_m - \Delta E_m - \Delta^2 S_m. \tag{2}$$

Instead of applying equation (2) directly to estimate streamflow changes, we use a simple statistical model which directly uses the monthly temperature (T) and precipitation outputs from the GCMs. In this context, ΔT is a proxy for ΔE and ΔT and ΔP are proxies for $\Delta^2 S$, particularly as temperature and precipitation relate to snow accumulation and melt. The presence of the storage term in equation (2) indicates that the runoff change in any given month must not only consider P and T changes in the same month, but also P and T changes in prior months. For example, the volume of snow which accumulates over the winter strongly affects the volume of snowmelt runoff that occurs over the succeeding spring, a fact that is often readily exploited in water supply forecasting [e.g., *Wood and Lettenmaier*, 2006; *Mahanama et al*., 2012]. Thus we replace equation (2) by taking streamflow change as linear function of temperature and precipitation changes in the current and preceding 11 months, which is formally described using the multiple linear regression model

$$\Delta R_m = a_m + \sum_{i=1}^{12} b_{mi} \Delta T_i + \sum_{i=1}^{12} c_{mi} \Delta P_i + \varepsilon_m, \tag{3}$$

where ΔR_m (unitless) is the relative change in average streamflow for target month m, ΔT_i and ΔP_i are the changes in GCM monthly temperature (absolute; °C) and precipitation (relative; %), respectively, for input month i, ε_m is a random error of mean 0 and variance $\sigma^2_{\varepsilon_m}$, and a_m, b_{mi}, and c_{mi} are the regression model parameters. ΔT_i and ΔP_i are spatially averaged over the area upstream of the respective stream gauge location. Using all 12 months in the regression, we explicitly consider the seasonal effect of any potential changes in water storage. Model inputs are GCM 30 year average monthly temperature and precipitation changes for the 2020s, 2050s, and 2080s and model output is the corresponding 30 year average monthly streamflow change for the 2020s, 2050s, and 2080s at each gauge location. One model is fit for each month for each streamflow site, for a total of 24 models. Projected future streamflow is estimated as:

$$Q_m^{(f)} = Q_m^{(ref)} (1 + \Delta R_m), \tag{4}$$

where $Q_m^{(ref)}$ is the simulated reference streamflow (1961–1990). When using equation (4) to make hydrologic projections based on climate projections that have not been downscaled and processed through the VIC model, no baseline (i.e., 1961–1990) streamflow exists for that specific projection. Hence, a common reference streamflow simulation is used for all projections, namely streamflow simulated by VIC forced with observed historical meteorological data (which is available for the period 1950–2006).

By calibrating to the VIC-simulated streamflow changes (i.e., the simulator output), equation (3) emulates the behavior of the VIC model in translating downscaled GCM outputs into monthly streamflow changes. We note that equation (3) is not a function of the VIC model parameters and strictly reflects the behavior of the VIC model as calibrated to a given set of parameter values. Further, model equation (3) does not emulate the dynamic state-dependent behavior of the VIC model (i.e., wherein the current value of the underlying process is modeled as dependent upon the state in the previous time step(s)) [see, for instance, *Reichert et al*., 2011; *Young and Ratto*, 2011; *Castelletti et al*., 2012; *Fraser et al*., 2013]. Hence, we treat each climatologically summarized VIC output as deriving from a static and independent response to climatologically summarized inputs. As this application assumes no state-dependence, streamflow changes in a given period are considered independent of those in prior periods and interannual dependence in basin water storage is considered negligible. In BC, this assumption may be invalidated for watersheds that generate substantial glacier runoff, which could potentially be subject to long-term (i.e., multidecadal) storage and runoff trends under a warming climate [*Stahl et al*., 2008]. By combining observed glacier thinning rates in BC from 1985 to 1999 [*Schiefer et al*., 2007] with glacier area from the Randolph Glacier Inventory [*Pfeffer et al*., 2014], contemporary annual glacier runoff volume is estimated as 17.5×10^8 and 1.35×10^8 m^3/yr for the Fraser at Hope and the Peace near Taylor, which results in an annual runoff contribution of 2% and $<$1%, respectively. If glacier runoff is assumed to be concentrated in the months of July, August, and September, this ratio changes to 6% and 1%, respectively. For convenience, we therefore assume that the

glacier runoff contribution is small and that potential future changes in glacier storage can be neglected at the far downstream locations considered in this study. However, credible projections at upstream or headwater locations would require careful modeling of glacier melt and change processes [e.g., *Stahl et al*., 2008; *Nolin et al*., 2010; *Jost et al*., 2011].

Future climate change may also cause long-term groundwater trends [e.g., *Green et al*., 2011; *Taylor et al*., 2013; *Allen et al*., 2014]. Hydrogeological characteristics can also in turn affect seasonal streamflow trends [*Tague et al*., 2008; *Tague and Grant*, 2009; *Safeeq et al*., 2013]. The VIC model represents the shallow ($<$3 m) local groundwater system (i.e., soil water that is recharged, stored, and discharged within the area represented by an individual computational grid ($\sim$30 km^2)); however, VIC does not represent deeper regional groundwater or water stored in bedrock. The extent and volume of groundwater resources in BC and the degree of interaction of slow-feedback regional and bedrock aquifers in the annual hydrologic cycle are poorly quantified [*BC Environment*, 1994]. Hence, any potential long-term streamflow changes due to groundwater trends are neglected in this study. Following calibration and verification, the final emulator model is fit to the complete data set (i.e., 2020s, 2050s, and 2080s under A1B, A2, and B1 forcing).

3.3. Emulator Calibration

Calibration of the emulator requires estimation of the 25 regression coefficients in equation (3) based on information provided by the 23 member VIC simulation ensemble. This leads to a "small n/large p" estimation problem, i.e., where the number of predictors approaches or exceeds the number of samples, and hence measures to avoid over fitting and generate a robust linear model are required. In this context, some form of regularization or predictor selection will be needed to provide adequate generalization performance, especially given that the emulator may be asked to extrapolate based on CMIP5 predictors that fall outside the ranges found in the SRES scenarios. The exact form that this takes, whether ridge regression, stepwise predictor selection, etc., is an open problem, and a wide variety of solutions to the "small n/large p" problem have been explored in the literature [*Johnstone and Titterington*, 2009]. Here we adopt a combination of regularized linear regression, in this case based on a modified form of gradient descent with early stopping [*Skouras et al*., 1994], and bootstrap aggregation [*Breiman*, 1996]. Note that this is a linear version of the ensemble neural network training procedure used by *Cannon and Whitfield* [2001] to model transient water quality events, and by *Cannon and Whitfield* [2002] to downscale daily streamflow based on climate variables. Adopting this calibration algorithm allowed for direct comparison between the proposed linear emulator and a neural network model [*Maier and Dandy*, 2000; *Hsieh*, 2009; *Maier et al*., 2010], which is an example of a flexible nonlinear modeling approach used successfully in the hydrologic modeling studies cited above.

Parameter estimation in a bootstrap aggregated ensemble model, whether for linear regression or a neural network, follows the same basic steps. In bootstrap aggregation, an ensemble of training data sets is generated by sampling with replacement from the available pool of calibration data. Regression models are then calibrated on the resampled data sets using an iterative nonlinear optimization algorithm, in this case a Newton-type method [*Schnabel et al*., 1985], with results from individual ensemble members averaged to yield final regression predictions. As bootstrap resampling is done with replacement, approximately 37% of training cases are not included in each of the bootstrapped sets. These "out-of-bootstrap" cases are used to estimate the generalization error of each regression model during each iteration of the nonlinear optimization algorithm, thereby allowing the calibration to be stopped at a point prior to overfitting. Rather than choosing the regression coefficients that maximize performance on the bootstrapped training set, final coefficients are instead chosen to maximize performance on the out-of-boostrap cases. In this study, the ensemble size was set to 200. Thus, even though each individual model only sees a portion of the data available for calibration, the ensemble as a whole will see the calibration data set in its entirety. Regression coefficients and confidence intervals are estimated from the ensemble following *Heskes* [1997] and *Carney et al*. [1999].

3.4. Emulator Verification

In order to verify that the linear regression model provides a robust means of estimating future streamflow changes, the emulation model was initially trained on the 2020s A1B, A2, and B1 data and then verified using the out-of-sample 2080s A2 projections. Figure 3, using the Fraser study area as an example, clearly indicates that the 2080s A2 input data (ΔP_i and ΔT_i) used for model verification are well outside the range

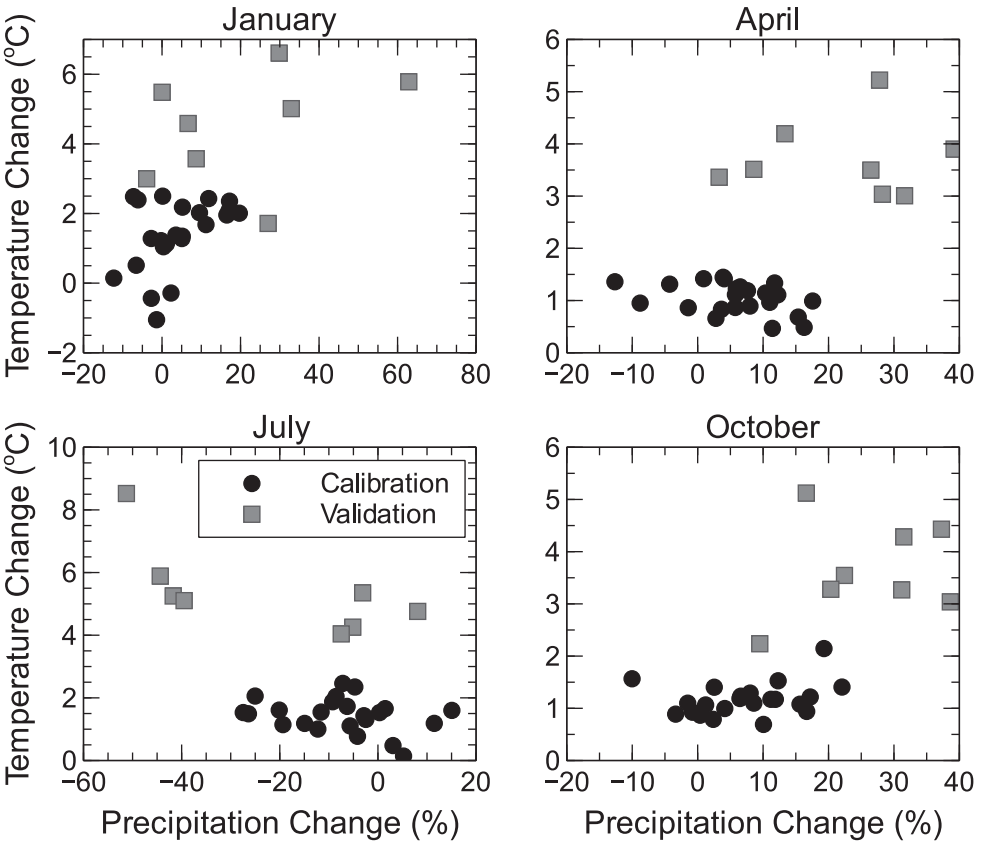

Figure 3. Comparison of range of calibration and validation inputs for the Fraser River study area.

of the 2020s data used for model calibration and, hence, should provide a robust and strict test of the model's effectiveness in emulating future streamflow changes. The degree of extrapolation is similar for the Peace River watershed (not shown). Because extrapolation behavior is sensitive to the underlying functional form of the regression model [e.g., see *Hsieh*, 2009, Figure 12.5], results from the regularized linear regression model are compared to results from three other models: ridge regression [*Hoerl and Kennard*, 1970; *Skouras et al*., 1994], which is an alternative means of regularizing ill posed linear regression problems; a nonlinear ensemble neural network model following *Cannon and Whitfield* [2001, 2002]; and linear trend. In addition, variants of regularized linear regression and ridge regression models were tested to see if nonlinear terms might improve emulation and extrapolation performance; in both cases, squares (ΔT_i^2 and ΔP_i^2) and cross-products ($\Delta T_i \Delta P_i$) terms were added to equation (3). For ridge regression, optimum values of the ridge penalty parameter were selected over the range 2^{-8} to 2^{16} via split-sample validation on bootstrap/out-of-bootstrap samples from the calibration data. Neural networks with one and two hyperbolic tangent hidden nodes were trained and verified using the same procedure as regularized linear regression. In the linear trend model, ΔR_m for the 2080s A2 period was estimated by extrapolating the linear trend in monthly streamflow between the 1961–1990 baseline and the 2011–2040 A2 scenario to the 2080s.

Finally, a second verification test was performed, this time with the emulation models trained on both the 2020s and 2050s A1B, A2, and B1 data and then verified, again, on the out-of-sample 2080s A2 projections. Model performance on the second verification test is expected to be better than on the first, as the training data set includes twice as many samples and less extrapolation is required when projecting to the 2080s based on data from the 2050s. Model performance for both verification tests is summarized in Table 1. Root mean squared errors (RMSE) for out-of-sample predictions of ΔR_m from the regularized linear regression model are encouraging and are universally superior to those for ridge regression, neural network, and linear trend, as well as those for the regularized linear regression model with squares and cross-products terms. The linear form of equation (3) provides robust extrapolation outside the range of the calibration data relative to the nonlinear models, and the gradient descent with early stopping approach to regularization

Table 1. Root Mean Square Error (RMSE) Values for Predictions of ΔR_m (All Months Combined) by the Regularized Linear Regression Model for the Fraser (FR) and Peace (PR) Projected Streamflow Data Compared to Alternative Models[a]

				RMSE (ΔR_m, Relative Change)	
	Watershed:	Fraser	Peace	Fraser	Peace
	Calibration:	2020s	2020s	2020s, 2050s	2020s, 2050s
Model (2080s A2 Verification)					
Reg. linear regression		**0.258**	**0.386**	**0.179**	**0.268**
Linear trend		0.583	0.554	0.365	0.362
Ridge regression		0.394	0.537	0.237	0.309
Neural network, 1 node		0.620	0.597	0.375	0.397
Neural network, 2 nodes		0.622	0.625	0.377	0.380
Reg. linear regression; $\Delta T_i^2, \Delta P_i^2, \Delta T_i \Delta P_i$		0.467	0.539	0.196	0.367
Ridge regression; $\Delta T_i^2, \Delta P_i^2, \Delta T_i \Delta P_i$		0.563	0.623	0.279	0.388

[a]Values for the best performing model in each watershed are shown in bold; values for models that perform worse than linear trend are underlined.

outperforms ridge regression. Within-sample RMSE values for the full regularized linear regression model (i.e., trained using all VIC-simulated data) are 0.076 and 0.123 for Fraser and Peace, respectively.

For context, model performance for the regularized linear model trained on data from the 2020s is verified graphically in Figure 4, where VIC-based and emulated projected monthly streamflow changes for the 2080s A2 are compared for both the Fraser and Peace. For the most part, emulated ΔR_m is in general agreement with simulated results. The seasonal variation in emulated ΔR_m is consistent with the simulated values in both basins, showing a pattern typical for nival regimes throughout western North America [*Morrison et al.*, 2002; *Toth et al.*, 2006; *Elsner et al.*, 2010; *Bürger et al.*, 2011; *Kerkhoven and Gan*, 2011; *Schnorbus et al.*, 2014; *Shrestha et al.*, 2012; *Hamlet et al.*, 2013], with increased streamflow in November through May and decreased streamflow in July through September. At both locations, the simulated and emulated results show that the largest streamflow increase and decrease is projected to occur in March and August, respectively. In the majority of months, the emulated results tend, however, to have dampened variability between GCMs, which is typical of linear regression [*Bürger et al.*, 2012, Appendix A]. Exceptions are the low flow months of August and September (Fraser) and July and August (Peace), where the spread between GCMs in the emulated flow is larger than that in the VIC-based projections. The emulated results also tend to underestimate streamflow increases, particularly for Peace, and slightly overestimate streamflow decreases, particular for Fraser. Nevertheless, results are encouraging and indicate that regularized linear regression is a robust approach to emulating the VIC-based projected streamflow changes for both study sites.

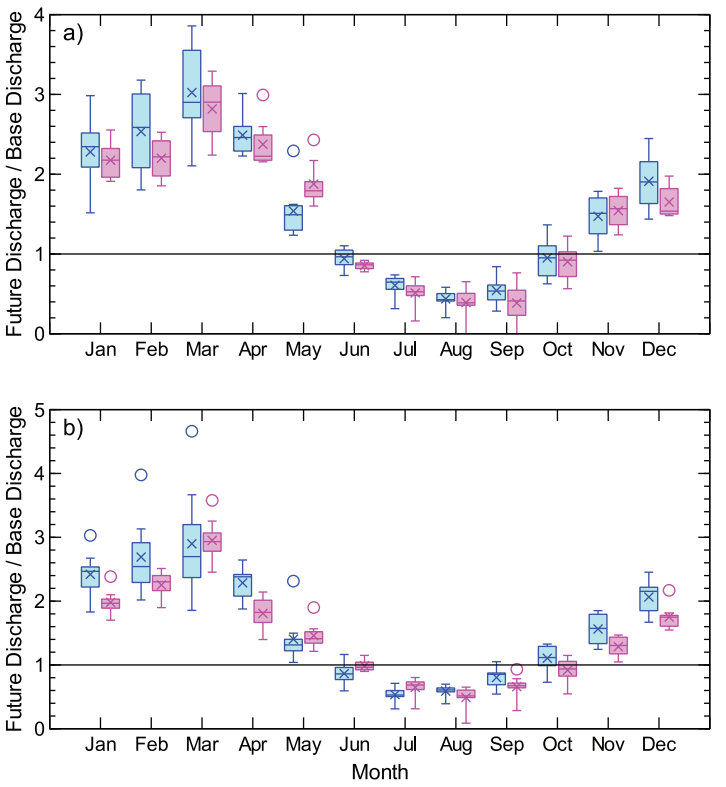

Figure 4. Out-of-sample verification comparing simulated (blue) to emulated (pink) box plots of projected monthly discharge changes for the 2080s A2 ensemble (8 × GCMs) for (a) the Fraser River at Hope and (b) the Peace River at Taylor.

Coefficient values (excluding intercepts) for the final model

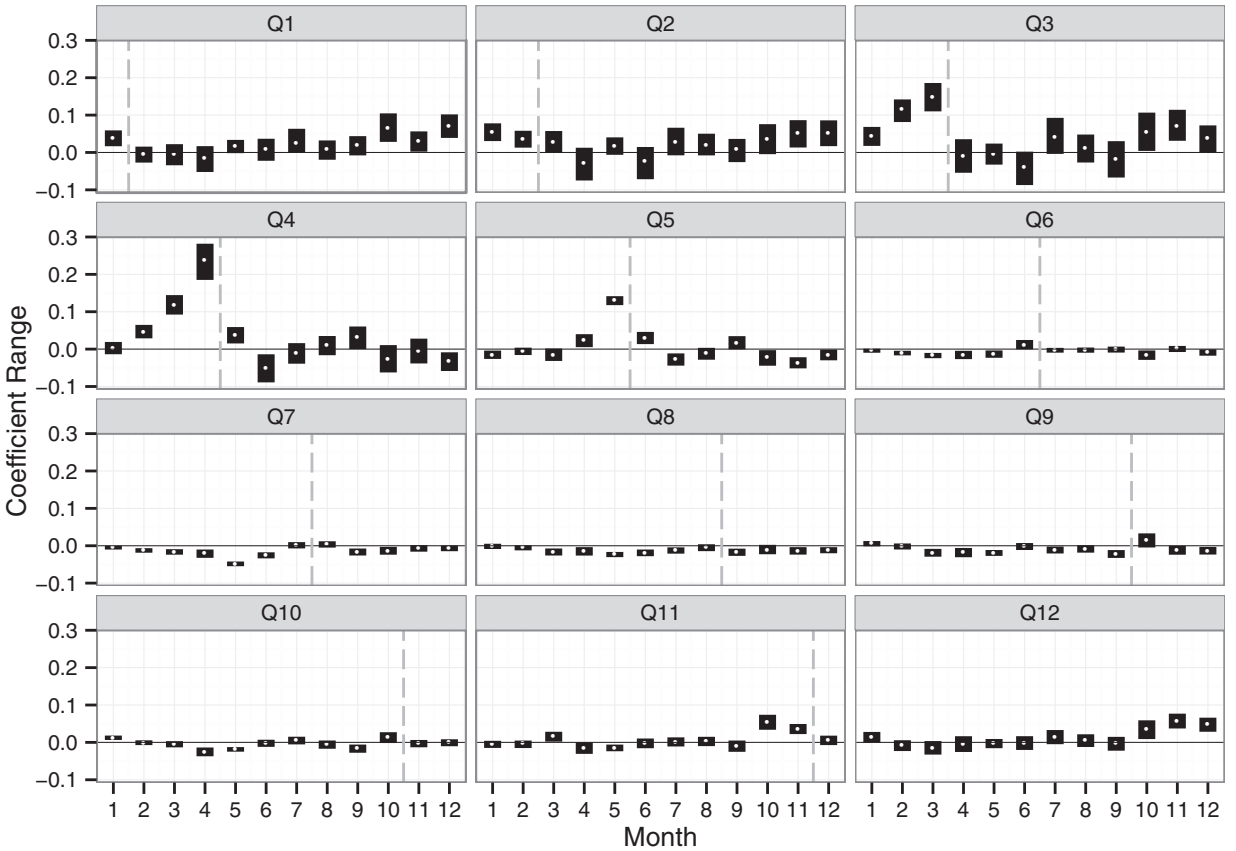

Figure 5. Full regression model monthly coefficient values for temperature change for given target month (Q1 = January, Q2 = February, etc.) for the Fraser River at Hope. Symbols show parameter values (white dot) with 2.5–97.5% uncertainty range (vertical bars).

are presented in Figures 5 and 6 for temperature and precipitation changes, respectively. For brevity, values are only shown for the Fraser. Although a physical interpretation of the individual model coefficients is not necessarily straightforward, we can make some general interpretations of how temperature and precipitation changes affect streamflow in the emulator. For temperature (Figure 5), the coefficient values can be interpreted as the relative change in monthly average streamflow for the target month resulting from a 1°C increase in monthly average temperature in any given month. Consistent with the Peace and the Fraser being nival regimes, temperature appears to have the strongest association with streamflow during the winter and spring months, when increasing temperature in the current and preceding months predicts increased streamflow during the target month. For the most part, streamflow only shows a significant response to temperature changes (interpreted based on the 95% confidence interval on the individual parameters) in the 2–4 months prior to the target month. Streamflow changes in the summer and fall months, although significant, do not exhibit as large a sensitivity to temperature changes. For precipitation effects (Figure 6), the coefficient values can be interpreted as the relative change in monthly average streamflow for the target month resulting from a 1% increase in monthly average precipitation in any given month. For these nival watersheds emulated streamflow changes are dominated by changes in temperature (supporting information Figure S1), such that the resultant coefficient values for precipitation are noisy and have physically unreasonable signs in certain months (e.g., occurrence of negative slope values in April, May, and November).

The emulator was constructed based on an a priori assumption that runoff change in a given month would be sensitive to temperature and precipitation changes in the current and all (i.e., 11) antecedent months. It is clear from the results that this is not necessarily the case, particularly for precipitation, indicating that the base model specified by equation (3) may be overparameterized for the two example basins. Nevertheless, as our primary goal is to emulate streamflow changes, and the emulator model is demonstrated to be a robust predictor of monthly streamflow change, all 12 months for both ΔT and ΔP are retained as predictors

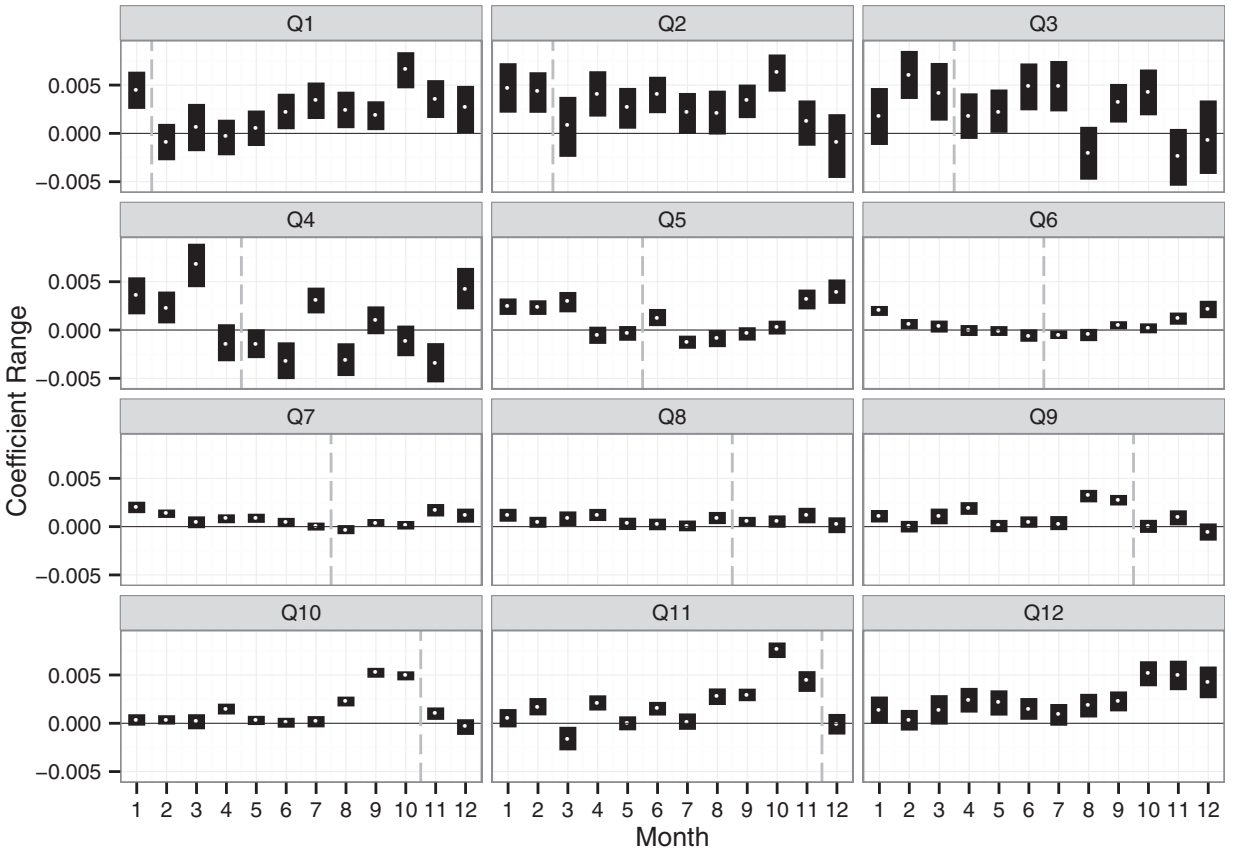

Figure 6. Same as Figure 5, but for precipitation change.

for sake of generality. Despite the large number of predictors in the base model, the final emulator does not suffer from overfitting because the effective number of parameters is reduced due to regularization via early stopping and ensemble averaging. As pointed out by *Marconato et al.* [2013], regularization techniques like early stopping can reduce overall model complexity without necessarily setting any of the original parameters exactly to zero. Given the dominant influence of ΔT relative to ΔP on streamflow changes in the emulator (supporting information Figure S1), the noisy but nonzero coefficient values associated with precipitation changes do not lead to reductions in model performance. Nevertheless, a more parsimonious baseline model was also specified and fit to determine if a reduction in the number predictors, in particular for precipitation, might lead to improvements in emulator performance. In this alternate model, streamflow changes are a function of ΔT and ΔP in months m and $m-1$, indicative of coincident and fast reaction to antecedent changes, as well as cumulative temperature and precipitation changes in months $m-6$ to $m-2$, indicative of lagged reactions to cumulative antecedent conditions:

$$\Delta R_m = a_m + b_m \Delta T_m + c_m \Delta P_m + b_{m-1} \Delta T_{m-1} + c_{m-1} \Delta P_{m-1} + b_s \sum_{i=2}^{6} \Delta T_{m-i} + c_s \sum_{i=2}^{6} \Delta P_{m-i} + \varepsilon_m \quad (5)$$

The total number of parameters is thus reduced from 25 to 7. As above, parameters in this reduced model were fit based on 2020s A1B, A2, and B1 data and performance was then verified using the out-of-sample 2080s A2 projections. Resulting RMSE values for the reduced model are found to be 5.0% worse than the full model (cf. Table 1) for the Fraser and 5.2% worse for the Peace, which lends further support to the general validity of an emulator based on equation (3).

For this study, the emulator was calibrated to an ensemble of opportunity which, in this case, provided good fitting results and a robust model. However, legitimate questions could be raised regarding the

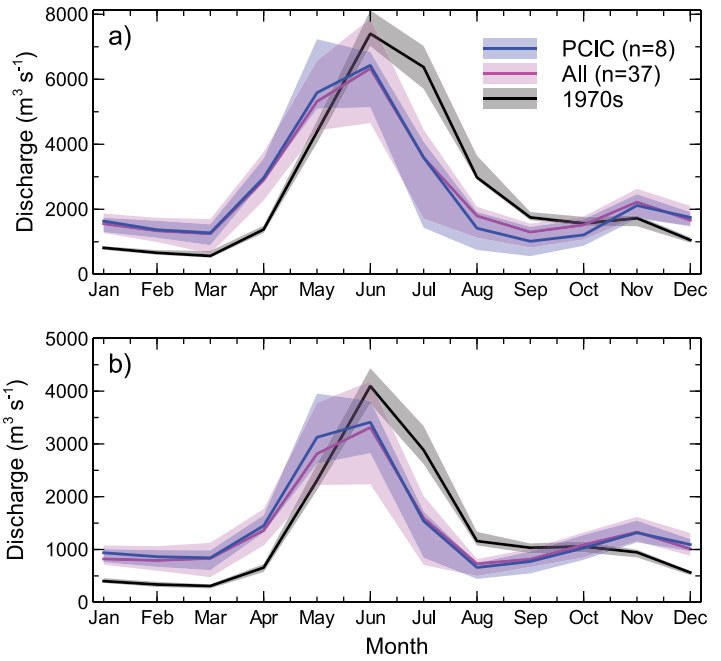

Figure 7. Monthly discharge for CMIP3 2080s A2 PCIC comparing the PCIC subset (blue) versus all CMIP3 (pink) for (a) the Fraser River at Hope and (b) the Peace River at Taylor. Dark lines show ensemble median and bands show the 5–95% percentile range (1970s range based on bootstrap estimate).

number of process-based simulations required to achieve a robust emulator calibration. Although it is beyond the scope of this study to determine an objective way of establishing an optimal sample size and type of process-based simulations, one can speculate that the choice will depend on the degree of spread among the available climate projections for the study area and the hydro-climatic regime being investigated (rainfall versus snowfall or precipitation versus temperature sensitivity).

4. Results and Discussion

4.1. CMIP3

In this section, we compare streamflow projections for the original PCIC subset of CMIP3 to the full suite of available CMIP3 climate projections (including multiple runs from the same GCM). Results are discussed for the 2080s A2 projections, which experiences the largest changes of the three emissions scenarios and periods under investigation. Projections for the PCIC subset are based on eight climate projections while those using all CMIP3 results are based on 37 climate projections. In this and all subsequent sections, future streamflow projections are plotted showing the ensemble median and the 5–95% range of the ensemble spread. For the 1970s period, ensemble spread cannot be represented for anything but the original 23 CMIP3-based hydrologic projections (8, 8, and 7 for the SRES A1B, A2, and B1 emissions scenarios, respectively). Hence, we plot the 1970s historical period using the 30 year median of the reference streamflow with 5–95% percentile range estimated using bootstrap resampling (sample size of 10,000) of the 30 median values for each month. For both the Fraser and the Peace, using a subset of available climate projections has little visible effect on the ensemble median monthly streamflow projections (Figure 7). However, the spread of the PCIC subset projections is smaller than the spread of projections for the full ensemble for the winter and spring periods.

The importance of this result depends upon the context in which GCM projections are chosen for impacts and adaptation studies. The use of multi-GCM ensembles generated from coordinated modeling experiments has become the standard to produce climate projections [*Meehl et al*., 2007; *Taylor et al*., 2012]. These ensembles sample at least some of the uncertainties due to emissions scenario, model structure, and indirect forcing effects and provide a framework for understanding and estimating projection uncertainties [*Tebaldi and Knutti*, 2007; *Knutti et al*., 2010a]. In the context of "model democracy" (i.e., where all GCMs are assumed to give plausible climate responses to greenhouse gas forcings) [*Knutti*, 2010], this result suggests that using a subset of GCMs can underestimate projection uncertainty as represented by the full ensemble. However, if models are instead chosen based on skill in simulating historical climate [e.g., *Gleckler et al*., 2008; *Rupp et al*., 2013], as was the case with the PCIC subset [*Werner*, 2011], then one may view the additional model spread as simply arising from less plausible GCMs. However, there is little agreement on metrics to separate good and bad models and there is not, as of yet, a great deal of evidence to indicate that historical simulation skill relates to the magnitude of projected changes [*Knutti et al*., 2010b; *Flato et al*., 2013]. Regardless, streamflow projections for the 2080s based on the consensus estimate (taken as the median of each respective A2 ensemble) give consistent trends with respect to projected streamflow changes, namely increased discharge in the winter and spring and reduced discharge during the summer.

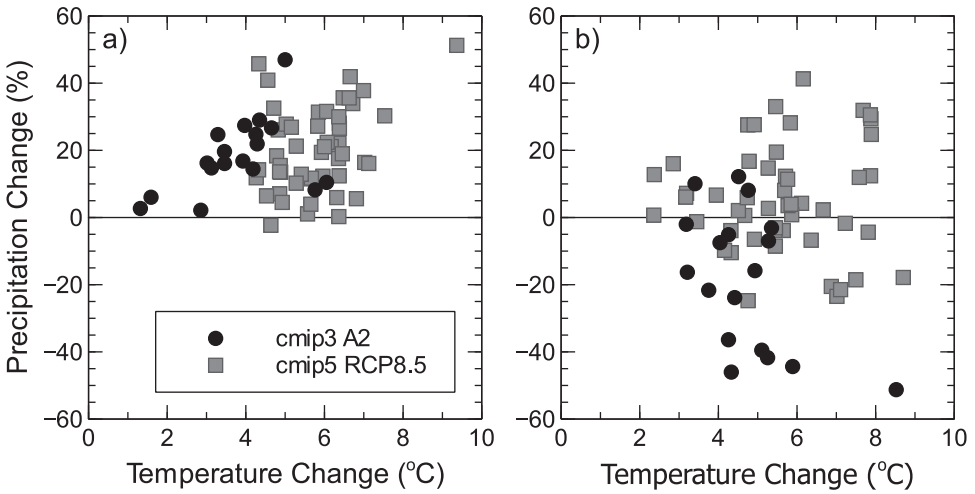

Figure 8. Projected 2080s temperature and precipitation changes for the Fraser River for the months (a) June and (b) July.

Nevertheless, agreement between the ensemble medians does not imply confidence in the overall accuracy of the projected climate and subsequent streamflow changes [*Collins et al.*, 2013].

4.2. CMIP3 Versus CMIP5

Results for this section compare streamflow changes using temperature and precipitation changes derived from both CMIP3 and CMIP5. This comparison is based on projections using the SRES A2 and the RCP8.5 emissions scenarios for the CMIP3 and CMIP5 models, respectively. The SRES A2 scenario represents a high emissions scenario, with diagnosed radiative forcing of 8–9.5 W m^{-2} over preindustrial levels by the end of the 21st century (based on the mean plus-or-minus one standard deviation from a simple climate model tuned to 19 CMIP3 GCMs) [*IPCC*, 2007]. The RCP8.5 scenario is also representative of high emissions scenarios in which no climate policies have been implemented and which represents the worst-case of the four RCP scenarios. RCP8.5 specifies a global radiative forcing over preindustrial levels of 8.5 W m^{-2} by the end of the 21st century [*van Vuuren et al.*, 2011a]. Despite similar radiative forcing by 2100, the emissions trajectories and composition of greenhouse gasses and pollutants prescribed by the two scenarios are not identical and are, therefore, not expected to generate an identical climate response [e.g., *Knutti and Sedlácek*, 2013].

Projected changes in temperature and precipitation for the Fraser are given in Figure 8 for the months of January (winter) and July (summer). The CMIP5 RCP8.5 climate projections produce greater warming than CMIP3 A2 in both January and July. Both the CMIP3 A2 and CMIP5 RCP8.5 ensembles show increased precipitation for January, but for July CMIP5 RCP8.5 is wetter (less dry) than the CMIP3 A2 ensemble. CMIP5 RCP8.5 precipitation changes for July are ambiguous (split between increased and decreased precipitation), whereas the CMIP3 A2 ensemble predominantly shows decreased precipitation. Results are generally similar for the Peace River basin for temperature in both months and January precipitation (not shown). In July, the CMIP5 RCP8.5 ensemble is again wetter than the CMIP3 A2 ensemble, however, in the case of Peace this translates to increased precipitation for CMIP5 RCP8.5 and predominantly decreased precipitation for CMIP3 A2 (not shown). In general, CMIP5 RCP8.5 generally projects warmer and slightly wetter (or less dry) conditions than CMIP3 A2.

Figure 9a compares 2080s streamflow projections for the Fraser. Although the CMIP5 RCP8.5 results are qualitatively similar to the CMIP3 A2 results, CMIP5 RCP8.5 clearly generates larger streamflow changes, namely larger increases in the winter and spring and larger decreases in the summer. In both cases, however, despite the seasonal shift in streamflow, the Fraser still essentially remains a nival regime as the annual hydrograph remains dominated by the spring freshet. Differences in projected streamflow changes represent differences in both greenhouse gas and pollutant emissions between the SRES A2 and RCP8.5 scenarios as well as differences in GCM model structure and process representation between CMIP3 and the

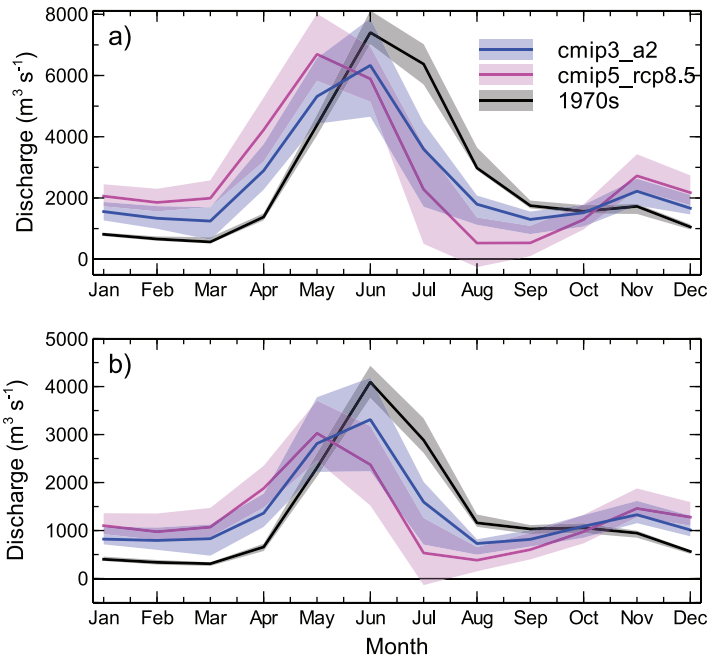

Figure 9. Monthly discharge for CMIP3 2080s A2 versus CMIP5 RCP8.5 for (a) the Fraser River at Hope and (b) the Peace River at Taylor. Dark lines show ensemble median and bands show the 5–95% percentile range (1970s range based on bootstrap estimate).

newer generation CMIP5 models (inter alia higher resolution and the inclusion of carbon cycling). The effect of these two differences cannot be diagnosed individually.

The results for the Fraser demonstrate some limitations of the statistical emulation approach, namely the projection of negative streamflow for some ensemble members during the month of August (for 12 of 50 experiments). Despite the verification of the model's robustness against CMIP3 data, the statistical model lacks physical constraints and can project physically implausible results in some extreme cases that are outside the range of the original calibration data (Figure 8). In the month of August, streamflow derives from precipitation and depletion of storage, mainly from the soil as seasonal snow is generally depleted by late summer [*Shrestha et al.*, 2012]. Projection of negative streamflow in August suggests that evaporation consumes in excess of all precipitation and available soil moisture storage, which is physically impossible. Our model, given by equation (2), implies that evaporation is linearly associated with changes in temperature, and assumes that, over a monthly integration period, temperature is an indicator of the evaporative potential of the atmosphere [e.g., *McKenney and Rosenberg*, 1993; *Shuttleworth*, 1993]. Model verification confirms that such a relationship is physically plausible in the range of the verification data, and we suspect that these conditions were indicative of energy limited conditions (i.e., evaporation restricted by available energy) or only moderate summer water stress (i.e., evaporation restricted by available water) [*Bonan*, 2008; *Roderick and Farquhar*, 2011]. The CMIP5 results suggest the emergence of increasingly moisture-limited conditions during summer in the 2080s for some scenarios, which imposes an increasingly stringent physical constraint on actual evaporation which is not explicitly incorporated in our

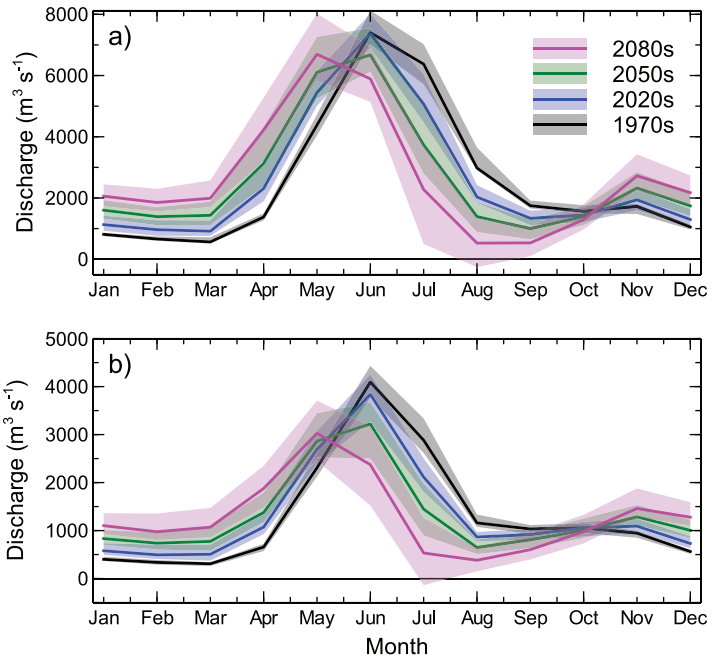

Figure 10. Monthly streamflow projections for CMIP5 RCP8.5 2020s, 2050s, and 2080s for (a) the Fraser River at Hope and (b) the Peace River at Taylor. Dark lines show ensemble median and bands show the 5–95% percentile range (1970s range based on bootstrap estimate).

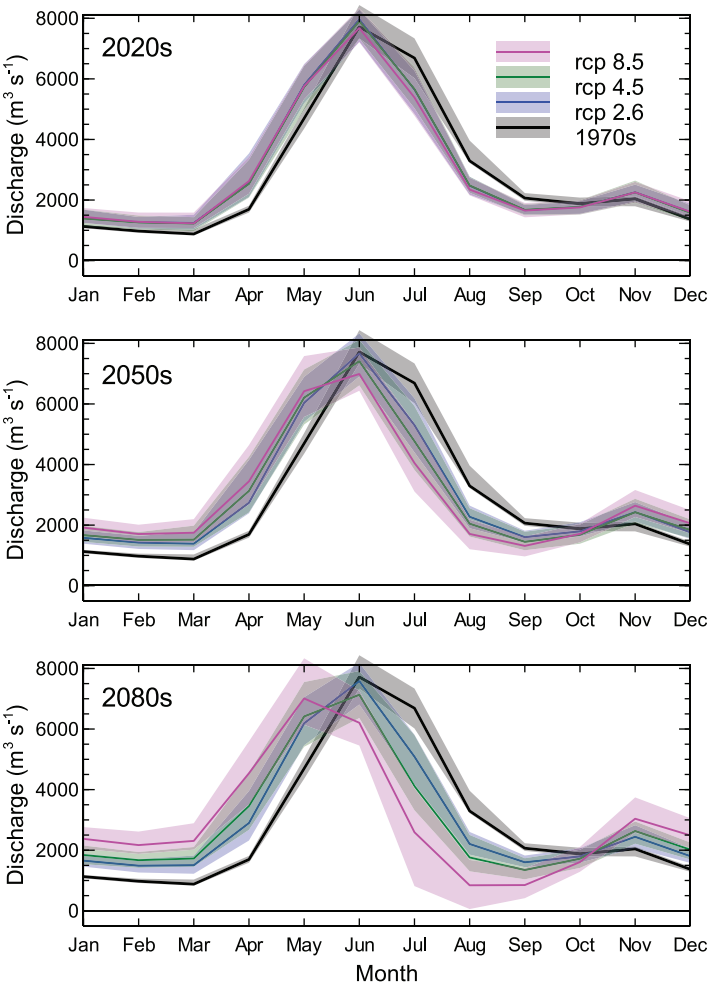

Figure 11. CMIP5 RCP2.6, RCP4.5, and RCP8.5 2020s, 2050s, and 2080s monthly streamflow projections for the Fraser River at Hope. Dark lines show ensemble median and bands show the 5–95% percentile range (1970s range based on bootstrap estimate).

statistical emulation model. Neglecting changes in glacier storage and late summer glacier melt may also be partly responsible for the error. Regardless, the results of Figure 9a are indicative of increasing aridity during the summer and the possibility for substantially reduced summer discharge. Further, we are able to demonstrate the use the emulated results to rapidly prototype the entire CMIP5 ensemble and identify those climate projections of particular interest, e.g., those indicative of high summer water stress, so they can be subject to a more thorough analysis using process-based models.

Results for the Peace are qualitatively similar (Figure 9b). Again, projected monthly streamflow changes for the 2080s under CMIP5 RCP8.5 are more severe than those projected for CMIP3 A2. As per the Fraser, 9 of 50 projections show negative streamflow for July. Curiously, the GCM projections that generate negative streamflow in the Peace and Fraser are not the same (with the exception of three). This is perhaps not surprising given the large regional variation in hydrologic response across western Canada to climate variability (e.g., El Niño-Southern Oscillation and Pacific Decadal Oscillation) [*Gobena and Gan*, 2006; *Woo and Thorne*, 2008; *Whitfield et al*., 2010; *Thorne and Woo*, 2011; *Gobena et al*., 2013], particularly the semiindependent variations in hydro-climate between northern and southern BC [*Moore and McKendry*, 1996; *Stahl et al*., 2006; *Wang et al*., 2006].

4.3. CMIP5

Monthly CMIP5 RCP8.5 streamflow projections for the 2020s, 2050s, and 2080s are given for the Fraser and Peace in Figure 10. For both locations, the streamflow response to climate change is typical of nival regimes [*Christensen et al*., 2004; *Barnett et al*., 2005; *Elsner et al*., 2010; *Schnorbus et al*., 2014; *Shrestha et al*., 2012], with increased discharge in the winter, earlier onset of the spring freshet, reduced summer discharge and lengthened summer low flow season. Changes become more severe moving from the 2020s to the 2080s and the prominence of the spring freshet progressively diminishes as more runoff is shifted to the winter. Nevertheless, even by the end of the century both study areas retain the traits of a nival regime, although snowmelt runoff is reduced in importance from that of the present climate and summer is projected to become much drier.

Streamflow projections based on three RCP emissions scenarios for the Fraser are given in Figure 11 for the periods centered on the 2020s, 2050s, and 2080s. Compared to RCP8.5, the RCP2.6 scenario represents a mitigation scenario leading to a very low radiative forcing level and RCP4.5 represents a medium

stabilization scenario resulting in 2.6 and 4.5 W m^{-2} radiative forcing, respectively, above preindustrial levels by 2100 [*van Vuuren et al*., 2011a]. During the 2020s, the projected monthly streamflow ensembles for the three scenarios are largely indistinguishable, and all three hint at increased discharge during the fall and winter, an earlier freshet onset and reduced summer discharge compared to the baseline period. These trends in monthly streamflow become larger through the 2050s and 2080s, and by the 2080s streamflow projections for the nonmitigation RCP8.5 scenario are clearly distinguishable and larger than those for either RCP2.6 or RCP4.5. However, for all three emissions scenarios there is strong consensus in all three periods of changes in monthly streamflow in every month except the seasonal transition months of May/June and October/November. Due to the long-lived nature of many greenhouse gases and the inertia of the climate system [*Collins et al*., 2013], even the high-mitigation RCP2.6 scenario (which may require a reduction in global CO_2 emissions of greater than 100% and a reduction in most non-CO_2 gases below present levels by 2100) [*van Vuuren et al*., 2011b] exhibits monthly streamflow changes throughout the century. For the 2080s, the spread in streamflow projections (representing the variation in temperature and precipitation response between the different GCMs) is largest for the RCP8.5 scenario and comparable for the RCP4.5 and 2.6 scenarios.

5. Conclusion

We have demonstrated a computationally efficient means of constructing projections of monthly streamflow change for the 2020s, 2050s, and 2080s for two streamflow sites in BC, Canada, the Peace River at Taylor and the Fraser River at Hope. The technique involves statistical emulation of streamflow projections made by the VIC hydrologic model driven with downscaled GCM data. The emulation uses regularized linear regression to model the relationship between basin-wide projected changes in monthly temperature and precipitation at the GCM resolution and VIC-simulated changes in monthly streamflow. Calibration and out-of-sample validation to a 23 member subset of the full CMIP3 multimodel ensemble (which uses the SRES A1B, A2, and B1 emissions scenarios) reveals that the approach is generally robust. Monthly projections have subsequently been made using output for larger ensembles of 136 and 134 runs of the CMIP3 and CMIP5 (which uses the RCP2.6, 4.5, and 8.5 emissions scenarios) multimodel data sets, respectively.

As GCM uncertainty is a large contributor to uncertainty in hydrological change scenarios, we considered the question of whether streamflow projections based on the 23 member VIC-simulated ensemble is representative of the full range of uncertainty from the 136 member CMIP3 ensemble. In terms of model spread, the 23 member VIC-simulated ensemble is only representative of the larger 136 member CMIP3 ensemble during the summer and fall half of the year. During the winter and spring, the smaller ensemble has a smaller spread than the larger ensemble. However, both ensembles provide essentially the same consensus estimate (taken as the ensemble median) of monthly streamflow change.

Do the more recent CMIP5 simulations, which are based on different emission scenarios, and, in most cases, new models, present a different picture of future streamflow changes? Qualitatively, CMIP5 RCP8.5 shows the same projected response as CMIP3 A2 for snow-dominated hydrologic regimes: higher winter discharge, advanced spring melt runoff, and lower summer discharge. However, the 2080s CMIP5 "worst-case" nonmitigation RCP8.5 runs have a larger impact (larger streamflow increases and decreases from the 1970s baseline) than the CMIP3 A2 runs.

Emulated results for the CMIP5 streamflow changes for the 2080s indicates that for some projections the model extrapolates to physically impossible values during the summer. In particular, emulated streamflow becomes negative for a limited number of ensemble member during July for the Peace River and during August for the Fraser River. However, this result underscores the advantage of using a simple statistical emulator to rapidly highlight and identify those future extreme projections that may merit further study using more involved, but computationally demanding, process-based methods.

As a proof-of-concept, it was beyond the scope of this project to explicitly develop a general emulation framework that would be applicable to broad range of hydro-climatic regimes. For instance, the methodology would need to be adjusted to accommodate situations requiring additional complexity where, for instance, glacier storage was a more substantial contribution to annual and monthly runoff. However, the flexibility of the model (monthly predictors for temperature and precipitation change) suggests that it is applicability to other types of hydro-climatic regimes merits investigation, so long as long-term trends in

water storage are negligible. Nevertheless, despite the relative simplicity and specificity of our application to two nival regimes in BC, this work has successfully demonstrated the efficacy of using an emulation approach to generate hydrologic projections.

Acknowledgments

We acknowledge the World Climate Research Programme's Working Group on Coupled Modeling, which is responsible for CMIP, and we thank the climate modeling groups (listed in supporting information Tables S1 and S2) for producing and making available their model output. For CMIP, the U.S. Department of Energy's Program for Climate Model Diagnosis and Intercomparison provides coordinating support and led development of software infrastructure in partnership with the Global Organization for Earth System Science Portals. GRACE land data (available at http://grace.jpl.nasa.gov) processing algorithms were provided by Sean Swenson, and supported by the NASA MEaSUREs Program. All data used for statistical emulation are available from the corresponding author by request. The authors are grateful to Francis Zwiers for his helpful comments and suggestions. We also thank two anonymous reviewers for their constructive comments.

References

Allen, D., K. Stahl, P. Whitfield, and R. Moore (2014), Trends in groundwater levels in British Columbia, *Can. Water Resour. J.*, *39*(1), 15–31, doi:10.1080/07011784.2014.885677.

Aly, A. H., and R. C. Peralta (1999), Optimal design of aquifer cleanup systems under uncertainty using a neural network and a genetic algorithm, *Water Resour. Res.*, *35*(8), 2523–2532, doi:10.1029/98WR02368.

Barnett, T. P., J. C. Adam, and D. P. Lettenmaier (2005), Potential impacts of a warming climate on water availability in snow-dominated regions, *Nature*, *438*(7066), 303–309.

Barnett, T. P., et al. (2008), Human-induced changes in the hydrology of the western United States, *Science*, *319*(5866), 1080–1083.

Barton, R. R. (1998), Simulation metamodels, in *Proceedings of the 1998 Winter Simulation Conference*, edited by D. Medeiros et al., pp. 167–174, Informs Simul. Soc., Washington, D. C.

Bawden, A. J., H. C. Linton, D. H. Burn, and T. D. Prowse (2014), A spatiotemporal analysis of hydrological trends and variability in the Athabasca River region, Canada, *J. Hydrol.*, *509*, 333–342, doi:10.1016/j.jhydrol.2013.11.051.

BC Environment (1994), *Groundwater Resources of British Columbia*, Minist. of Environ., Lands and Parks and Environment Canada, Victoria, B. C., Canada.

Bennett, K. E., A. T. Werner, and M. A. Schnorbus (2012), Uncertainties in hydrologic and climate change impact analyses in headwater basins of British Columbia, *J. Clim.*, *25*(17), 5711–5730.

Bonan, G. (2008), *Ecological Climatology: Concepts and Applications*, 2nd ed., 563 pp., Cambridge Univ. Press, N. Y.

Bonfils, C., et al. (2008), Detection and attribution of temperature changes in the mountainous western United States, *J. Clim.*, *21*(23), 6404–6424.

Breiman, L. (1996), Bagging predictors, *Mach. Learning*, *24*(2), 123–140, doi:10.1007/BF00058655.

Broad, D., G. Dandy, and H. Maier (2005), Water distribution system optimization using metamodels, *J. Water Resour. Plann. Manage.*, *131*(3), 172–180, doi:10.1061/(ASCE)0733-9496(2005)131:3(172).

Bürger, G., J. Schulla, and A. T. Werner (2011), Estimates of future flow, including extremes, of the Columbia River headwaters, *Water Resour. Res.*, *47*, W10520, doi:10.1029/2010WR009716.

Bürger, G., T. Q. Murdock, A. T. Werner, S. R. Sobie, and A. J. Cannon (2012), Downscaling extremes—An intercomparison of multiple statistical methods for present climate, *J. Clim.*, *25*(12), 4366–4388.

Cannon, A. J., and P. H. Whitfield (2001), Modeling transient pH depressions in coastal streams of British Columbia using neural networks, *J. Am. Water Resour. Assoc.*, *37*(1), 73–89, doi:10.1111/j.1752-1688.2001.tb05476.x.

Cannon, A. J., and P. H. Whitfield (2002), Downscaling recent streamflow conditions in British Columbia, Canada using ensemble neural network models, *J. Hydrol.*, *259*(1–4), 136–151, doi:10.1016/S0022-1694(01)00581-9.

Carnevale, C., G. Finzi, G. Guariso, E. Pisoni, and M. Volta (2012), Surrogate models to compute optimal air quality planning policies at a regional scale, *Environ. Modell. Software*, *34*, 44–50, doi:10.1016/j.envsoft.2011.04.007.

Carney, J. G., P. Cunningham, and U. Bhagwan (1999), Confidence and prediction intervals for neural network ensembles, in *International Joint Conference on Neural Networks (IJCNN'99)*, vol. 2, pp. 1215–1218, IEEE, Piscataway Township, N. J.

Castelletti, A., S. Galelli, M. Ratto, R. Soncini-Sessa, and P. Young (2012), A general framework for dynamic emulation modelling in environmental problems, *Environ. Modell. Software*, *34*, 5–18, doi:10.1016/j.envsoft.2012.01.002.

Castruccio, S., D. J. McInerney, M. L. Stein, F. Liu Crouch, R. L. Jacob, and E. J. Moyer (2013), Statistical emulation of climate model projections based on precomputed GCM runs, *J. Clim.*, *27*(5), 1829–1844, doi:10.1175/JCLI-D-13-00099.1.

Christensen, N. S., A. W. Wood, N. Voisin, D. P. Lettenmaier, and R. N. Palmer (2004), The effects of climate change on the hydrology and water resources of the Colorado River Basin, *Clim. Change*, *62*(1–3), 337–363.

Cohen, S. J., K. A. Miller, A. F. Hamlet, and W. Avis (2000), Climate change and resource management in the Columbia River basin, *Water Int.*, *25*(2), 253–272.

Collins, M., et al. (2013), Long-term climate change: Projections, commitments and irreversibility, in *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*, edited by T. Stocker et al., pp. 1029–1136, Cambridge Univ. Press, Cambridge, U. K.

Demarchi, D. A. (1996), *An Introduction to the Ecoregions of British Columbia*., Wildlife Branch, Ministry of Environment, Lands and Parks, Victoria, B. C., Canada.

Déry, S. J., K. Stahl, R. D. Moore, P. H. Whitfield, B. Menounos, and J. E. Burford (2009), Detection of runoff timing changes in pluvial, nival, and glacial rivers of western Canada, *Water Resour. Res.*, *45*, W04426, doi:10.1029/2008WR006975.

Elsner, M. M., C. Lan, N. Voisin, J. S. Deems, A. F. Hamlet, J. A. Vano, K. E. B. Mickelson, L. Se-Yeun, and D. P. Lettenmaier (2010), Implications of 21st century climate change for the hydrology of Washington State, *Clim. Change*, *102*(1/2), 225–260.

Flato, G., et al. (2013), Evaluation of climate models, in *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*, edited by T. Stocker et al., 866 pp., Cambridge Univ. Press, Cambridge, U. K.

Fleming, S. W., and F. A. Weber (2012), Detection of long-term change in hydroelectric reservoir inflows: Bridging theory and practise, *J. Hydrol.*, *470–471*, 36–54, doi:10.1016/j.jhydrol.2012.08.008.

Foreman, M., D. Lee, J. Morrison, S. Macdonald, D. Barnes, and I. Williams (2001), Simulations and retrospective analyses of Fraser watershed flows and temperatures, *Atmos. Ocean*, *39*(2), 89–105.

Fraser, C. E., N. McIntyre, B. M. Jackson, and H. S. Wheater (2013), Upscaling hydrological processes and land management change impacts using a metamodeling procedure, *Water Resour. Res.*, *49*, 5817–5833, doi:10.1002/wrcr.20432.

Gleckler, P. J., K. E. Taylor, and C. Doutriaux (2008), Performance metrics for climate models, *J. Geophys. Res.*, *113*, D06104, doi:10.1029/2007JD008972.

Gobena, A. K., and T. Y. Gan (2006), Low-frequency variability in southwestern Canadian stream flow: Links with large-scale climate anomalies, *Int. J. Climatol.*, *26*(13), 1843–1869, doi:10.1002/joc.1336.

Gobena, A. K., F. A. Weber, and S. W. Fleming (2013), The role of large-scale climate modes in regional streamflow variability and implications for water supply forecasting: A case study of the Canadian Columbia River basin, *Atmos. Ocean*, *51*(4), 380–391, doi:10.1080/07055900.2012.759899.

Green, T. R., M. Taniguchi, H. Kooi, J. J. Gurdak, D. M. Allen, K. M. Hiscock, H. Treidel, and A. Aureli (2011), Beneath the surface of global change: Impacts of climate change on groundwater, *J. Hydrol.*, *405*(3–4), 532–560, doi:10.1016/j.jhydrol.2011.05.002.

Hamlet, A. F., P. W. Mote, M. P. Clark, and D. P. Lettenmaier (2007), Twentieth-century trends in runoff, evapotranspiration, and soil moisture in the western United States, *J. Clim.*, *20*(8), 1468–1486, doi:10.1175/JCLI4051.1.

Hamlet, A. F., L. Se-Yeun, K. E. B. Mickelson, and M. M. Elsner (2010), Effects of projected climate change on energy supply and demand in the Pacific Northwest and Washington State, *Clim. Change*, *102*(1/2), 103–128.

Hamlet, A. F., M. M. Elsner, G. S. Mauger, S.-Y. Lee, I. Tohver, and R. A. Norheim (2013), An overview of the Columbia basin climate change scenarios project: Approach, methods, and summary of key results, *Atmos. Ocean*, *51*(4), 392–415, doi:10.1080/07055900.2013.819555.

Hatcher, K. L., and J. A. Jones (2013), Climate and streamflow trends in the Columbia River basin: Evidence for ecological and engineering resilience to climate change, *Atmos. Ocean*, *51*(4), 436–455, doi:10.1080/07055900.2013.808167.

Heskes, T. (1997), Practical confidence and prediction intervals, *Adv. Neural Inf. Process. Syst.*, *9*, 176–182.

Hoerl, A. E., and R. W. Kennard (1970), Ridge regression: Biased estimation for nonorthogonal problems, *Technometrics*, *12*(1), 55–67.

Holden, P. B., and N. R. Edwards (2010), Dimensionally reduced emulation of an AOGCM for application to integrated assessment modelling, *Geophys. Res. Lett.*, *37*, L21707, doi:10.1029/2010GL045137.

Hsieh, W. W. (2009), *Machine Learning Methods in the Environmental Sciences: Neural Networks and Kernels*, 349 pp., Cambridge Univ. Press, Cambridge, U. K.

IPCC (2007), Summary for policymakers, in *Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*, edited by S. Solomon et al., pp. 1–18, Cambridge Univ. Press, Cambridge, U. K.

IPCC (2013), Summary for policymakers, in *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*, edited by T. Stocker et al., 29 pp., Cambridge Univ. Press, Cambridge, U. K.

Johnstone, I. M., and D. M. Titterington (2009), Statistical challenges of high-dimensional data, *Philos. Trans. R. Soc. A*, *367*(1906), 4237–4253.

Jost, G., R. D. Moore, B. Menounos, and R. Wheate (2011), Quantifying the contribution of glacier runoff to streamflow in the upper Columbia River basin, Canada, *Hydrol. Earth Syst. Sci. Discuss.*, *8*, 4979–5008, doi:10.5194/hessd-8-4979-2011.

Jung, I.-W., H. Chang, and H. Moradkhani (2011), Quantifying uncertainty in urban flooding analysis considering hydro-climatic projection and urban development effects, *Hydrol. Earth Syst. Sci.*, *15*(2), 617–633, doi:10.5194/hess-15-617-2011.

Kay, A. L., H. N. Davies, V. A. Bell, and R. G. Jones (2009), Comparison of uncertainty sources for climate change impacts: Flood frequency in England, *Clim. Change*, *92*(1–2), 41–63, doi:10.1007/s10584-008-9471-4.

Kerkhoven, E., and T. Y. Gan (2011), Differences and sensitivities in potential hydrologic impact of climate change to regional-scale Athabasca and Fraser River basins of the leeward and windward sides of the Canadian Rocky Mountains respectively, *Clim. Change*, *106*(4), 583–607.

Kleijnen, J. P. (2008), *Design and Analysis of Simulation Experiments, International Series in Operations Research & Management Science*, vol. 111, Springer, N. Y.

Knutson, T. R., F. Zeng, and A. T. Wittenberg (2013), Multimodel assessment of regional surface temperature trends: CMIP3 and CMIP5 twentieth-century simulations, *J. Clim.*, *26*(22), 8709–8743, doi:10.1175/JCLI-D-12-00567.1.

Knutti, R. (2010), The end of model democracy?, *Clim. Change*, *102*(3–4), 395–404.

Knutti, R., and J. Sedlácek (2013), Robustness and uncertainties in the new CMIP5 climate model projections, *Nat. Clim. Change*, *3*(4), 369–373, doi:10.1038/nclimate1716.

Knutti, R., G. Abramowitz, M. Collins, V. Eyring, P. Gleckler, B. Hewitson, and L. Mearns (2010a), Good practice guidance paper on assessing and combining multi model climate projections, in *Meeting Report of the Intergovernmental Panel on Climate Change Expert Meeting on Assessing and Combining Multi Model Climate Projections*, edited by T. Stocker et al., IPCC Working Group I Tech. Support Unit, Univ. of Bern, Bern, Switzerland.

Knutti, R., R. Furrer, C. Tebaldi, J. Cermak, and G. A. Meehl (2010b), Challenges in combining projections from multiple climate models, *J. Clim.*, *23*(10), 2739–2758, doi:10.1175/2009JCLI3361.1.

Landerer, F. W., and S. C. Swenson (2012), Accuracy of scaled GRACE terrestrial water storage estimates, *Water Resour. Res.*, *48*, W04531, doi:10.1029/2011WR011453.

Li, G., X. Zhang, F. Zwiers, and Q. H. Wen (2012), Quantification of uncertainty in high-resolution temperature scenarios for North America, *J. Clim.*, *25*(9), 3373–3389, doi:10.1175/JCLI-D-11-00217.1.

Liang, X., D. P. Lettenmaier, E. F. Wood, and S. J. Burges (1994), A simple hydrologically based model of land-surface water and energy fluxes for general-circulation models, *J. Geophys. Res.*, *99*(D7), 14,415–14,428.

Liang, X., E. F. Wood, and D. P. Lettenmaier (1996), Surface soil moisture parameterization of the VIC-2L model: Evaluation and modification, *Global Planet. Change*, *13*(1–4), 195–206.

Lohmann, D., R. NolteHolube, and E. Raschke (1996), A large-scale horizontal routing model to be coupled to land surface parametrization schemes, *Tellus, Ser. A*, *48*(5), 708–721.

Luce, C. H., and Z. A. Holden (2009), Declining annual streamflow distributions in the Pacific Northwest United States, 1948–2006, *Geophys. Res. Lett.*, *36*, L16401, doi:10.1029/2009GL039407.

Mahanama, S., B. Livneh, R. Koster, D. Lettenmaier, and R. Reichle (2012), Soil moisture, snow, and seasonal streamflow forecasts in the United States, *J. Hydrometeorol.*, *13*(1), 189–203, doi:10.1175/JHM-D-11-046.1.

Maier, H. R., and G. C. Dandy (2000), Neural networks for the prediction and forecasting of water resources variables: A review of modelling issues and applications, *Environ. Modell. Software*, *15*(1), 101–124.

Maier, H. R., A. Jain, G. C. Dandy, and K. P. Sudheer (2010), Methods used for the development of neural networks for the prediction of water resource variables in river systems: Current status and future directions, *Environ. Modell. Software*, *25*(8), 891–909.

Mantua, N., I. Tohver, and A. Hamlet (2010), Climate change impacts on streamflow extremes and summertime stream temperature and their possible consequences for freshwater salmon habitat in Washington State, *Clim. Change*, *102*(1/2), 187–223.

Marconato, A., M. Schoukens, Y. Rolain, and J. Schoukens (2013), Study of the effective number of parameters in nonlinear identification benchmarks, in *2013 IEEE 52nd Annual Conference on Decision and Control (CDC)*, pp. 4308–4313, IEEE, Piscataway Township, N. J.

Maurer, E. P., A. W. Wood, J. C. Adam, D. P. Lettenmaier, and B. Nijssen (2002), A long-term hydrologically based dataset of land surface fluxes and states for the conterminous United States, *J. Clim.*, *15*(22), 3237–3251.

McKenney, M. S., and N. J. Rosenberg (1993), Sensitivity of some potential evapotranspiration estimation methods to climate change, *Agric. For. Meteorol.*, *64*(1–2), 81–110, doi:10.1016/0168-1923(93)90095-Y.

Meehl, G. A., C. Covey, K. E. Taylor, T. Delworth, R. J. Stouffer, M. Latif, B. McAvaney, and J. F. B. Mitchell (2007), The WCRP CMIP3 multimodel dataset: A new era in climate change research, *Bull. Am. Meteorol. Soc.*, *88*(9), 1383–1394, doi:10.1175/BAMS-88-9-1383.

Moore, R. D. (1991), Hydrology and water supply in the Fraser River basin, in *Water in Sustainable Development: Exploring Our Common Future in the Fraser River Basin, Research Program on Water in Sustainable Development*, edited by A. Dorcey and J. Griggs, pp. 21–40, Wastewater Res. Cent., Univ. of B. C., Vancouver, B. C., Canada.

Moore, R. D., and I. G. McKendry (1996), Spring snowpack anomaly patterns and winter climatic variability, British Columbia, Canada, *Water Resour. Res.*, *32*(3), 623–632.

Morrison, J., M. C. Quick, and M. G. G. Foreman (2002), Climate change in the Fraser River watershed: Flow and temperature projections, *J. Hydrol.*, *263*(1–4), 230–244, doi:10.1016/S0022-1694(02)00065-3.

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

Nakićenović, N., and R. Swart (Eds.) (2000), Emissions scenarios, Intergovernmental Panel on Climate Change (IPCC) special reports of Working Group III, 570 pp., Cambridge Univ. Press, Cambridge, U. K.

Nolin, A. W., J. Phillippe, A. Jefferson, and S. L. Lewis (2010), Present-day and future contributions of glacier runoff to summertime flows in a Pacific Northwest watershed: Implications for water resources, *Water Resour. Res.*, *46*, W12509, doi:10.1029/2009WR008968.

O'Hagan, A. (2006), Bayesian analysis of computer code outputs: A tutorial, *Reliab. Eng. Syst. Safety*, *91*(10–11), 1290–1300, doi:10.1016/j.ress.2005.11.025.

Oki, T., K. Musiake, H. Matsuyama, and K. Masuda (1995), Global atmospheric water balance and runoff from large river basins, *Hydrol. Processes*, *9*(5–6), 655–678.

Payne, J. T., A. W. Wood, A. F. Hamlet, R. N. Palmer, and D. P. Lettenmaier (2004), Mitigating the effects of climate change on the water resources of the Columbia River basin, *Clim. Change*, *62*(1–3), 233–256.

Pfeffer, W. T., et al. (2014), The Randolph Glacier Inventory: A globally complete inventory of glaciers, *J. Glaciol.*, *221*(221), 537–552, doi:10.3189/2014JoG13J176.

Pierce, D. W., et al. (2008), Attribution of declining western U.S. snowpack to human effects, *J. Clim.*, *21*(23), 6425–6444.

Prudhomme, C., and H. Davies (2009), Assessing uncertainties in climate change impact analyses on the river flow regimes in the UK. Part 2: Future climate, *Clim. Change*, *93*(1/2), 197–222.

Radić, V., and G. K. C. Clarke (2011), Evaluation of IPCC models' performance in simulating late-twentieth-century climatologies and weather patterns over North America, *J. Clim.*, *24*(20), 5257–5274, doi:10.1175/JCLI-D-11-00011.1.

Razavi, S., B. A. Tolson, and D. H. Burn (2012a), Numerical assessment of metamodelling strategies in computationally intensive optimization, *Environ. Modell. Software*, *34*, 67–86, doi:10.1016/j.envsoft.2011.09.010.

Razavi, S., B. A. Tolson, and D. H. Burn (2012b), Review of surrogate modeling in water resources, *Water Resour. Res.*, *48*, W07401, doi:10.1029/2011WR011527.

Regonda, S. K., B. Rajagopalan, M. Clark, and J. Pitlick (2005), Seasonal cycle shifts in hydroclimatology over the western United States, *J. Clim.*, *18*(2), 372–384.

Reichert, P., G. White, M. Bayarri, and E. Pitman (2011), Mechanism-based emulation of dynamic simulation models: Concept and application in hydrology, *Comput. Stat. Data Anal.*, *55*(4), 1638–1655, doi:10.1016/j.csda.2010.10.011.

Rind, D., M. Chin, G. Feingold, D. Streets, R. Kahn, S. Schwartz, and H. Yu (2009), Modeling the effects of aerosols on climate, in *Atmospheric Aerosol Properties and Climate Impacts, U.S. Climate Change Science Program and the Subcommittee on Global Change Research, Synthesis and Assessment Product*, vol. 2.3, edited by M. Chin, R. A. Kahn, and S. Schwartz, pp. 55–84, Natl. Aeronaut. and Space Admin., Washington, D. C.

Roderick, M. L., and G. D. Farquhar (2011), A simple framework for relating variations in runoff to variations in climatic conditions and catchment properties, *Water Resour. Res.*, *47*, W00G07, doi:10.1029/2010WR009826.

Rupp, D. E., J. T. Abatzoglou, K. C. Hegewisch, and P. W. Mote (2013), Evaluation of CMIP5 20th century climate simulations for the Pacific Northwest USA, *J. Geophys. Res. Atmos.*, *118*, 10,884–10,906, doi:10.1002/jgrd.50843.

Safeeq, M., G. E. Grant, S. L. Lewis, and C. L. Tague (2013), Coupling snowpack and groundwater dynamics to interpret historical streamflow trends in the western United States, *Hydrol. Processes*, *27*(5), 655–668, doi:10.1002/hyp.9628.

Salathé, E. P. (2005), Downscaling simulations of future global climate with application to hydrologic modelling, *Int. J. Climatol.*, *25*(4), 419–436.

Schiefer, E., B. Menounos, and R. Wheate (2007), Recent volume loss of British Columbian glaciers, Canada, *Geophys. Res. Lett.*, *34*, L16503, doi:10.1029/2007GL030780.

Schnabel, R. B., J. E. Koonatz, and B. E. Weiss (1985), A modular system of algorithms for unconstrained minimization, *ACM Trans. Math. Software*, *11*(4), 419–440.

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

Schultz, M., M. Small, R. Farrow, and P. Fischbeck (2004), State water pollution control policy insights from a reduced-form model, *J. Water Resour. Plann. Manage.*, *130*(2), 150–159, doi:10.1061/(ASCE)0733-9496(2004)130:2(150).

Schultz, M. T., M. J. Small, P. S. Fischbeck, and R. S. Farrow (2006), Evaluating response surface designs for uncertainty analysis and prescriptive applications of a large-scale water quality model, *Environ. Monit. Assess.*, *11*(4), 345–359, doi:10.1007/s10666-006-9043-9.

Shrestha, R. R., M. A. Schnorbus, A. T. Werner, and A. J. Berland (2012), Modelling spatial and temporal variability of hydrologic impacts of climate change in the Fraser River basin, British Columbia, Canada, *Hydrol. Processes*, *26*(12), 1840–1860, doi:10.1002/hyp.9283.

Shuttleworth, W. J. (1993), Evaporation, in *Handbook of Hydrology*, edited by D. R. Maidment, chap. 4, McGraw-Hill, N. Y.

Skouras, K., C. Goutis, and M. J. Bramson (1994), Estimation in linear models using gradient descent with early stopping, *Stat. Comput.*, *4*(4), 271–278, doi:10.1007/BF00156750.

Song, X., C. Zhan, and J. Xia (2012), Integration of a statistical emulator approach with the SCE-UA method for parameter optimization of a hydrological model, *Chin. Sci. Bull.*, *57*(26), 3397–3403, doi:10.1007/s11434-012-5305-x.

Stahl, K., R. D. Moore, and I. G. McKendry (2006), The role of synoptic-scale circulation in the linkage between large-scale ocean-atmosphere indices and winter surface climate in British Columbia, Canada, *Int. J. Climatol.*, *26*(4), 541–560.

Stahl, K., R. D. Moore, J. M. Shea, D. Hutchinson, and A. J. Cannon (2008), Coupled modelling of glacier and streamflow response to future climate scenarios, *Water Resour. Res.*, *44*, W02422, doi:10.1029/2007WR005956.
Stewart, I. T. (2009), Changes in snowpack and snowmelt runoff for key mountain regions, *Hydrol. Processes*, *23*(1), 78–94, doi:10.1002/hyp.7128.
Surfleet, C. G., and D. Tullos (2012), Uncertainty in hydrologic modelling for estimating hydrologic response due to climate change (Santiam River, Oregon), *Hydrol. Processes*, *27*(25), 3560–3576, doi:10.1002/hyp.9485.
Swenson, S., and J. Wahr (2006), Post-processing removal of correlated errors in GRACE data, *Geophys. Res. Lett.*, *33*, L08402, doi:10.1029/2005GL025285.
Tague, C., and G. E. Grant (2009), Groundwater dynamics mediate low-flow response to global warming in snow-dominated alpine regions, *Water Resour. Res.*, *45*, W07421, doi:10.1029/2008WR007179.
Tague, C., G. Grant, M. Farrell, J. Choate, and A. Jefferson (2008), Deep groundwater mediates streamflow response to climate warming in the Oregon Cascades, *Clim. Change*, *86*(1–2), 189–210, doi:10.1007/s10584-007-9294-8.
Taylor, K. E., R. J. Stouffer, and G. A. Meehl (2012), An overview of CMIP5 and the experiment design, *Bull. Am. Meteorol. Soc.*, *93*(4), 485–498, doi:10.1175/BAMS-D-11-00094.1.
Taylor, R. G., et al. (2013), Ground water and climate change, *Nat. Clim. Change*, *3*(4), 322–329, doi:10.1038/nclimate1744.
Tebaldi, C., and R. Knutti (2007), The use of the multi-model ensemble in probabilistic climate projections, *Philos. Trans. R. Soc. A*, *365*(1857), 2053–2075.
Thorne, R., and M.-k. Woo (2011), Streamflow response to climatic variability in a complex mountainous environment: Fraser River basin, British Columbia, Canada, *Hydrol. Processes*, *25*(19), 3076–3085, doi:10.1002/hyp.8225.
Toth, B., A. Pietroniro, F. M. Conly, and N. Kouwen (2006), Modelling climate change impacts in the Peace and Athabasca catchment and delta: I—Hydrological model application, *Hydrol. Processes*, *20*(19), 4197–4214.
Vano, J. A., and D. P. Lettenmaier (2013), A sensitivity-based approach to evaluating future changes in Colorado River discharge, *Clim. Change*, *122*(4), 621–634, doi:10.1007/s10584-013-1023-x.
Vano, J. A., N. Voisin, L. Cuo, A. F. Hamlet, M. M. Elsner, R. N. Palmer, A. Polebitski, and D. P. Lettenmaier (2010a), Climate change impacts on water management in the Puget Sound region, Washington State, USA, *Clim. Change*, *102*(1–2), 261–286.
Vano, J. A., M. J. Scott, N. Voisin, C. O. Stöckle, A. F. Hamlet, K. E. B. Mickelson, M. M. Elsner, and D. P. Lettenmaier (2010b), Climate change impacts on water management and irrigated agriculture in the Yakima River Basin, Washington, USA, *Clim. Change*, *102*(1/2), 287–317.
Vano, J. A., et al. (2014), Understanding uncertainties in future Colorado River streamflow, *Bull. Am. Meteorol. Soc.*, *95*(1), 59–78, doi:10.1175/BAMS-D-12-00228.1.
van Vuuren, D. P., et al. (2011a), The representative concentration pathways: An overview, *Clim. Change*, *109*(1–2), 5–31, doi:10.1007/s10584-011-0148-z.
van Vuuren, D. P., et al. (2011b), RCP2.6: Exploring the possibility to keep global mean temperature increase below 2°C, *Clim. Change*, *109*(1–2), 95–116, doi:10.1007/s10584-011-0152-3.
Villa-Vialaneix, N., M. Follador, M. Ratto, and A. Leip (2012), A comparison of eight metamodeling techniques for the simulation of N_2O fluxes and N leaching from corn crops, *Environ. Modell. Software*, *34*, 51–66, doi:10.1016/j.envsoft.2011.05.003.
Wang, J. Y., P. H. Whitfield, and A. J. Cannon (2006), Influence of Pacific climate patterns on low-flows in British Columbia and Yukon, Canada, *Can. Water Resour. J.*, *31*(1), 25–40, doi:10.4296/cwrj3101025.
Werner, A. T. (2011), *BCSD Downscaled Transient Climate Projections for Eight Select GCMs Over British Columbia, Canada*, Pac. Clim. Impacts Consortium, Univ. of Victoria, Victoria, B. C., Canada.
Whitfield, P. H., and A. J. Cannon (2000), Recent variations in climate and hydrology in Canada, *Can. Water Resour. J.*, *25*(1), 19–65.
Whitfield, P. H., R. D. Moore, S. W. Fleming, and A. Zawadzki (2010), Pacific decadal oscillation and the hydroclimatology of western Canada: Review and prospects, *Can. Water Resour. J.*, *35*(1), 1–28, doi:10.4296/cwrj3501001.
Woo, M.-K., and R. Thorne (2008), Analysis of cold season streamflow response to variability of climate in north-western North America, *Hydrol. Res.*, *39*(4), 257–265, doi:10.2166/nh.2008.
Wood, A. W., and D. P. Lettenmaier (2006), A test bed for new seasonal hydrologic forecasting approaches in the western United States, *Bull. Am. Meteorol. Soc.*, *87*(12), 1699–1712, doi:10.1175/BAMS-87-12-1699.
Yan, S., and B. Minsker (2006), Optimal groundwater remediation design using an adaptive neural network genetic algorithm, *Water Resour. Res.*, *42*, W05407, doi:10.1029/2005WR004303.
Young, P. C., and M. Ratto (2011), Statistical emulation of large linear dynamic models, *Technometrics*, *53*(1), 29–43, doi:10.1198/TECH.2010.07151.
Zou, R., W.-S. Lung, and J. Wu (2009), Multiple-pattern parameter identification and uncertainty analysis approach for water quality modeling, *Ecol. Modell.*, *220*(5), 621–629, doi:10.1016/j.ecolmodel.2008.11.021.