Evaluating the ability of a hydrologic model to replicate hydro-ecologically relevant indicators

Rajesh R. Shrestha, Daniel L. Peters, & Markus A. Schnorbus

2014

Original citation:

Shrestha, R. R., Peters, D. L., & Schnorbus, M. A. (2014). Evaluating the ability of a hydrologic model to replicate hydro-ecologically relevant indicators. *Hydrological Processes*, *28*(14), 4294–4310. https://doi.org/10.1002/hyp.9997

HYDROLOGICAL PROCESSES
Hydrol. Process. **28**, 4294–4310 (2014)
Published online 2 September 2013 in Wiley Online Library
(wileyonlinelibrary.com) DOI: 10.1002/hyp.9997

Evaluating the ability of a hydrologic model to replicate hydro-ecologically relevant indicators

Rajesh R. Shrestha,[1]* Daniel L. Peters[2] and Markus A. Schnorbus[1]

[1] *Pacific Climate Impacts Consortium, University of Victoria, Victoria, BC, Canada*
[2] *Environment Canada at Water and Climate Impacts Research Centre, University of Victoria, Victoria, BC, Canada*

Abstract:

It is a common practice to employ hydrologic models for assessing alterations to streamflow as a result of anthropogenically driven changes, such as riverine, land use, and climate change. However, the ability of the models to replicate different components of the hydrograph simultaneously is not clear. Hence, this study evaluates the ability of a standard hydrologic model set-up: Variable Infiltration Capacity (VIC) hydrologic model for two headwater sub-basins in the Fraser River (Salmon and Willow), British Columbia, Canada, with climate inputs derived from observations and statistically downscaled global climate models (GCMs); to simulate six general water resource indicators (WRIs) and 32 ecologically relevant indicators of hydrologic alterations (IHA). The results show a generally good skill of the observation-driven VIC model in replicating most of the WRIs and IHAs. Although the WRIs, including annual volume, centre of timing, and seasonal flows, and the IHAs, including maximum and minimum flows, were reasonably well replicated, statistically significant differences in some of the monthly flows, number and duration of flow pulses, rise and fall rates, and reversals were noted. In the case of GCM-driven results, additional monthly, maximum, and minimum flow indicators produced statistically significant differences. A number of issues with the model input/output data, hydrologic model parametrization and structure as well as downscaling methods were identified, which lead to such discrepancies. Therefore, there is a need to exercise caution in the use of model-simulated indicators. Overall, the WRIs and IHAs can be useful tools for evaluating changes in an altered hydrologic system, provided the skill and limitations of the model in replicating these indicators are understood. © 2013 Her Majesty the Queen in Right of Canada. Hydrological Processes © 2013 John Wiley & Sons Ltd.

KEY WORDS indicators of hydrologic alteration; water resource indicators; model evaluation; modelling uncertainties

Received 8 January 2013; Accepted 1 August 2013

INTRODUCTION

Hydrologic models are powerful tools for assessing alterations to streamflow resulting from natural and anthropogenically driven changes such as land use, riverine, and climate change. Models are generally employed to simulate baseline/reference conditions and altered states, and the degree of changes to the system are assessed using a range of hydrologic indicators. For instance, a survey of the literature on hydrologic impacts of climate change (e.g. Dibike and Coulibaly, 2005; Merritt *et al*., 2006; Toth *et al*., 2006; Chang and Jung, 2010; Shrestha *et al*., 2012a,2012b) revealed a focus on the analysis of water resource indicators (WRIs), such as monthly, seasonal, and annual flows; magnitude and timing of peak and low flows; and centre of timing of annual flow.

Hydrologic indicators have also been a focus for ecological flow needs (EFN) research and recommendations. With the advent of the natural flow paradigm for river conservation and restoration in the 1990s (Poff *et al*., 1997), it has become clear that protection of the river ecosystem requires examination of temporal variation in key streamflow hydrograph components (Richter *et al*., 1996, 1997; Lytle and Poff, 2004; Sanford *et al*., 2007) beyond the typical WRIs. This is especially true for cold-climate countries, such as Canada, where the bulk of flow regimes is characterized by winter low flows and high spring flows, with potentially highly variable summer and fall flow periods (Monk *et al*., 2011). Richter *et al*. (1996) proposed a suite of 32 ecologically relevant indicators of hydrologic alteration (IHA), which expands on the typical WRIs to include a range of statistical magnitude and timing of events (e.g. annual maximum/minimum 3-, 7-, 30-, and 90-day mean flows) and critical fluvial components outlined in the natural flow paradigm that characterizes intra-annual and inter-annual variability in flow conditions (e.g. rise/fall rate and

*Correspondence to: Rajesh R. Shrestha, Pacific Climate Impacts Consortium, University House 1, PO Box 3060 STN CSC, University of Victoria, Victoria, BC, Canada V8W 3R4.
E-mail: rshresth@uvic.ca
Reproduced with the permission of the Minister of Environment Canada

number of reversals in the hydrograph). A detailed discussion on the IHAs and their ecological significance is beyond the scope of this paper; excellent descriptions are readily found in original IHA (Richter *et al*., 1996, 1997; Poff *et al*., 1997) and subsequent user studies (e.g. Monk *et al*., 2011).

An international focused group (Poff *et al*., 2010) and a Canadian focused group (Peters *et al*., 2012) recently proposed frameworks for developing EFN guidelines to protect river systems from instream (e.g. damming for hydroelectric generation) and out-of-stream development (e.g. water abstraction). The application of the EFN framework (e.g. to assess the degree of hydrological alteration on a flow regime and propose sustainable ecological flow recommendations) necessitates the use of hydrologic models to simulate anticipated impacts of interest, whereas both natural and impacted streamflow simulations are required in the case of ungauged basins. Assessment of projected climate change impacts on hydro-ecology is another area where the hydrologic-model-simulated IHAs have been used (e.g. Gibson *et al*., 2005; Kim *et al*., 2011). With such recognition of the importance of IHAs for EFN and climate change impact assessments, there will be a growing dependence on IHAs extracted from hydrological model simulation.

However, the ability of hydrological models to replicate the complete suite of ecologically relevant IHAs in a single simulation effort is not entirely clear. This lack of clarity is highlighted by recent studies that employed simple rainfall–runoff and statistical modelling approaches (Carlisle *et al*., 2010; Knight *et al*., 2012; Murphy *et al*., 2013), which showed positive or negative biases in the modelled indicators. A key factor affecting a hydrologic model's ability to reproduce observed variations in the hydrologic indicators is the inherent uncertainties, which arise from the following (Beven, 2006; Renard *et al*., 2010): (i) input uncertainty, e.g. errors in precipitation measurement, and interpolation uncertainties due to elevation and orographic effects; (ii) output uncertainty, e.g. errors in observed discharge data due to hydrometric data uncertainties (Pelletier, 1990; Hamilton, 2008; Hamilton and Moore, 2012) and stage–discharge relationship uncertainties (Shrestha *et al*., 2007); (iii) structural uncertainty, arising from simplified and/or incomplete representation of hydrologic processes in the model; and (iv) parameter uncertainty, arising from the need to calibrate parameter values based on observed streamflow. Difficulties in reproducing some of the hydrologic indicators using the goodness-of-fit (GOF) measures [coefficient of determination, R^2, and Nash–Sutcliffe coefficient of efficiency (NSE)] (Nash and Sutcliffe, 1970) have been documented by previous studies, such as simultaneous representation of low flow and high flow with a single set of parameter values (Wagener, 2003; Fenicia *et al*., 2007). Furthermore, because the GOF measures such as NSE and R^2 are mainly sensitive to peak flows (Legates and McCabe, 1999; Krause *et al*., 2005; Pushpalatha *et al*., 2012), models calibrated with such measures may not be able to accurately represent low-flow conditions. Estimation of the model parameters may be further complicated by the use of only an instream hydrometric station for model calibration, which leads to an equifinality problem (Beven, 2006), as several different parameter combinations within a chosen model structure can yield similar model performance. Low flows are poorly reproduced by several hydrological models (Staudinger *et al*., 2011) because model structural uncertainty plays a major role in low-flow conditions (Najafi *et al*., 2011), which in many cases may not be fully represented by the model.

Climate scenarios used to project future climate impacts on streamflow generation contribute to further to uncertainties. Such uncertainties arise from future greenhouse gas emissions, global climate model (GCM)/regional climate model structure, downscaling method, and natural variability of the climate system (Kay *et al*., 2008). In particular, previous studies (Kay *et al*., 2008; Prudhomme and Davies, 2008a,2008b; Bennett *et al*., 2012) indicated that a GCM structure is the largest source of uncertainty in projected hydrologic impacts, whereas other studies (Maurer *et al*., 2010; Quintana Seguí *et al*., 2010; Chen *et al*., 2011) suggested that the downscaling methods can considerably add to the uncertainty.

Given the aforementioned uncertainties and limitations, an important question arises concerning the confidence we can have in the model-simulated elements of the hydrologic cycle (Blöschl and Montanari, 2010). Therefore, there is an obvious need for a rigorous evaluation of the replicability of the WRIs and IHAs, before using the model-simulated indicators for impact assessments. Although it could be argued that the model agreement with observations (using such rigorous evaluation) does not guarantee reliable predictions of altered states, such agreement with the baseline condition is currently the best way to assign model confidence, with the underlying assumption that a model that accurately describes baseline condition will make a better projection of the altered condition (Reichler and Kim, 2008).

The primary objective of this study is to provide insights on the suitability of a hydrologic model simulating WRIs and IHAs to conduct the EFN and climate change assessments. The suitability was assessed via comparison of indicators extracted from the observations to the simulated flows from a commonly used hydrologic model. A secondary objective is to assess the replicability from a typical climate change studies perspective, i.e. by utilizing a hydrologic model driven by statistically downscaled GCMs.

METHODS

Hydrologic model set-up

This study employed the macro-scale VIC hydrologic model version 4.0.7 (Liang *et al*., 1994, 1996, 2003), which was previously applied to the Fraser River Basin in British Columbia, Canada, to investigate the impacts of mountain pine beetle infestation and climate change on streamflow generation (Schnorbus *et al*., 2010; Shrestha *et al*., 2012b). VIC represents soil moisture processes in three soil layers and explicitly accounts for snow accumulation and ablation by using a two-layer snowpack medium individually for each land cover/elevation tile. Simulated energy and mass fluxes, as well as the state variables for each grid cell, are calculated as the area averages of the tiles (Liang *et al*., 2003). VIC is distinguished from other macro-scale models by explicit consideration of the sub-grid variability of land surface vegetation classes and topography. The spatial variability of infiltration and runoff generation is simulated using the variable infiltration curve (Liang *et al*., 1994), and baseflow is represented using the empirically based Arno conceptual curve (Todini, 1996). Surface runoff from the upper two soil layers is generated when the moisture exceeds the storage capacity of the soil. A detailed description of the VIC model is available in Liang *et al*. (1994, 1996, 2003).

The spatially distributed VIC model was considered appropriate for the Fraser basin because of the basin's large area (230 000 km^2) and physiographic and hydro-climatic heterogeneity (Shrestha *et al*., 2012b). The suitability of the VIC model for hydrologic and climate change assessment of similar large basins has been successfully demonstrated by a number of previous studies in the Pacific Northwest region (e.g. Hidalgo *et al*., 2009; Elsner *et al*., 2010; Wenger *et al*., 2010; Schnorbus *et al*.,). The 1/16° model resolution used in this study is consistent with similar basin-scale applications (e.g. Elsner *et al*., 2010; Wenger *et al*., 2010).

The VIC model for the Fraser River basin can be described as a 'standard' application for a macro-scale basin. The Fraser river basin was sub-divided into 66 sub-basins (based on the Water Survey of Canada hydrometric station network) for the representation of spatially heterogeneous hydrologic responses (Schnorbus *et al*., 2010; Shrestha *et al*., 2012b). The VIC model was driven by gridded observation data of daily maximum and minimum air temperature, daily total precipitation, and daily average wind speed, derived primarily from the Environment Canada climate station observation network. Model calibration also followed a standard approach: A multi-objective complex evolution (MOCOM) (Yapo *et al*., 1998) method was used with three commonly used GOF measures as objective functions: (i) NSE; (ii) NSE of log-transformed discharge (LNSE); and (iii) water balance error (WBE). A set of five runoff generation parameters were used for calibration: variable infiltration curve parameter (*Bi*), fraction of maximum soil moisture where nonlinear baseflow occurs (*Ws*), maximum velocity of baseflow (mm/day) (*Dsmax*), fraction of *Dsmax* where nonlinear baseflow begins (*Ds*), and variation of saturated hydraulic conductivity with soil moisture (*Expn*). The *Bi* parameter controls the partitioning of net precipitation into surface and infiltration (and ultimately baseflow). The *Ds*, *Ws*, and *Dsmax* parameters control the rate of baseflow discharge as a function of soil moisture in the lowest (third) soil layer and influence the overall magnitude and timing of the hydrograph. The baseflow curve also affects the rate of soil moisture storage change over time, which indirectly affects the volume of moisture available for transpiration. The *Expn* parameter controls the variation of hydraulic conductivity as a function of soil moisture, which governs the rate of vertical percolation between the three soil layers (Schnorbus *et al*., 2011). Demaria *et al*. (2007) found discharge simulation to be most sensitive to *Bi* and *Expn*. Additionally, in view of the uncertainties in the precipitation data (such as low station density at high elevation and orographic effects) (Stahl *et al*., 2006; Neilsen *et al*., 2010), an adjustment factor for precipitation was used for the VIC model calibration. Five independent MOCOM runs were performed for the sub-basins. From the model performance with respect to the three objective functions, the model parameters with the best overall performance were selected by using the fuzzy preference selection methodology (Shrestha and Rode, 2008), which is based on the composite degree of fulfilment of the multiple-objective functions. As in the case of most hydrologic modelling studies (which use 5- to 10-year calibration periods), 6 years (1985–1990) of observed discharge was used for model calibration and an additional 5 years (1991–1995) was used for model validation. Calibration was conducted after a 3-year spin-up period to exclude the effects of initial conditions on hydrologic simulations.

Additionally, given the uncertainties in driving GCMs and downscaling, this study analysed the VIC model results driven by eight GCMs participating in phase 3 of the Coupled Model Intercomparison Project (Meehl *et al*., 2007a) (Table I), each for a Climate of the 20th Century run. The GCM data, consisting of monthly precipitation, and minimum and maximum air temperature were bias corrected, spatially disaggregated (to a 1/16° grid) and temporally disaggregated (to a daily time step) using a statistical method called bias correction spatial disaggregation (BCSD) (Wood *et al*., 2004). The BCSD downscaling was calibrated to reproduce the distribution of observations over 1950–1990 and validated over

Table I. Eight GCMs participating in phase 3 of the Coupled Model Intercomparison Project (Meehl *et al*., 2007a) used in this study

Model ID	Modelling centre	Atmospheric resolution	Primary reference
CCSM3	National Center for Atmospheric Research (USA)	T85 L26	Collins *et al*. (2006)
CGCM3 (T47)	Canadian Centre for Climate Modelling and Analysis (Canada)	T47 L31	Scinocca *et al*. (2008)
CSIRO-Mk3.0	Commonwealth Scientific and Industrial Research Organisation (Australia)	T63 L18	Rotstayn *et al*. (2010)
ECHAM5	Max Planck Institute for Meteorology (Germany)	T63 L32	Roeckner *et al*. (2006)
GFDL-CM2.1	NOAA Geophysical Fluid Dynamics Laboratory (USA)	N45 L24	Delworth *et al*. (2006)
HADCM3	Hadley Centre for Climate Prediction and Research (UK)	T42 L19	Collins *et al*. (2001)
HADGEM1	Hadley Centre for Climate Prediction and Research (UK)	N96 L38	Martin *et al*. (2006)
MIROC3.2	Center for Climate Systems Research (Japan)	T42 L20	K-1 model developers (2004)

1991–2000. The GCMs provide the long-term variability of climate; and daily/monthly or annual variability of the observed precipitation/temperature is not replicated. Thus, the downscaled-GCM-driven hydrologic model is not anticipated to replicate the daily/monthly or annual dynamics of the observed streamflow. A detailed description of the BCSD is available in Wood *et al*. (2004), and its application for climate projections over BC is available in Werner (2011).

From the 66 sub-basins individually calibrated in the Fraser basin, the Salmon and Willow sub-basins (Figure 1) were selected for this study because of key characteristics of the sub-basins that minimize uncertainties (i.e. no significant instream flow alteration, abstraction or regulation, and/or glaciers or large lakes that are not explicitly accounted for in the VIC model) and facilitate a rigorous evaluation of the hydrologic model performance. However, as in the case of most other sub-basins in the Fraser basin, both Salmon and Willow sub-basins have experienced forest disturbances from insect infestations (mountain pine beetle) and harvesting (Lin and Wei, 2008; Schnorbus *et al*., 2010). The study by Lin and Wei (2008) indicated increases in mean and peak flows over annual and spring periods and no change in the low flow

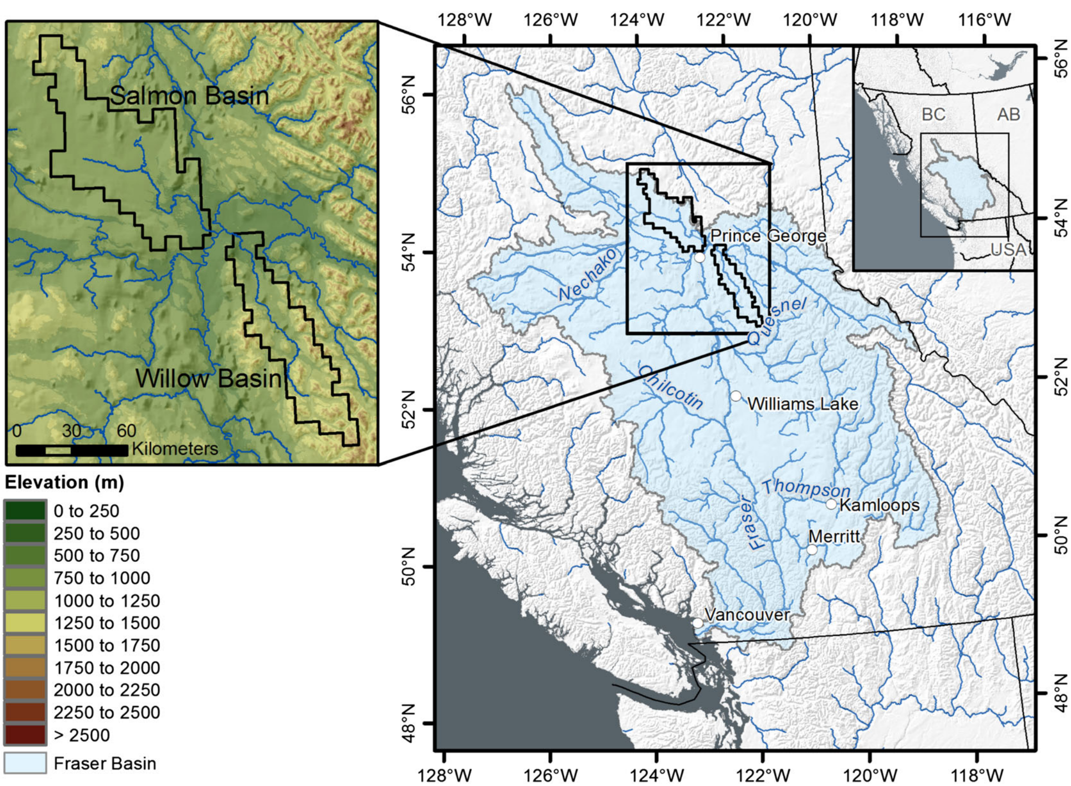

Figure 1. Fraser basin and study sub-basins. Elevation shown is from Shuttle Radar Topography Mission digital elevation model

due to forest harvesting (annual forest harvesting ranges from 0.6% to 1.2% of total sub-basin area) in the Willow sub-basin, whereas Schnorbus *et al.* (2010) found no change in the 20-year peak flow due to insect infestations in both sub-basins. The impact of forest cover change on the hydrologic model performance is not considered in this study.

The two study sub-basins are located in the central plateau region of the Fraser River basin, contributing flow to the river main stem near the Shelly hydrometric station. The Salmon River originates in the vicinity of Tahaetkun and Bouleau Mountains, whereas the Willow River originates in the Cariboo Mountains. Forests are the dominant land cover in both sub-basins. Although located in close proximity of one another, the physiographic and hydro-climatic characteristics of the two sub-basins differ considerably. Specifically, the Salmon sub-basin, which is located in the lower-elevation ranges, receives lower precipitation and generates less runoff compared with the Willow sub-basin (Table II). Mean monthly temperature varies between −10 °C (January) and 14 °C (July) in the Salmon sub-basin and −8.5 °C (January) and 13 °C (July) in the Willow sub-basin. Both sub-basins are dominated by snowmelt-driven runoff (nival regime), with peak flow generally occurring in the spring.

It should be noted that this study is not intended as an evaluation of the two sub-basin characteristics or the VIC hydrologic model used in this study. By using a 'standard' model set-up, which has been found suitable for this particular region, this study is intended to provide general insights on the suitability of hydrologic-model-simulated indicators for the EFN and climate change assessment.

Methods of evaluation

The ability of the VIC hydrologic model to replicate the characteristics of the observed streamflow was analysed using the six WRIs and 32 IHAs that are listed in Table III. It is to be noted that some of the IHAs (i.e. monthly median flows and magnitude and timing of maximum and minimum daily flows) are also commonly considered in water resource applications. The WRIs and IHAs were extracted from the observed daily streamflow dataset, as well as from the VIC model results driven by observed historical climatology and a suite of eight downscaled-GCM datasets. The IHA software version 7.1 (The Nature Conservancy, 2009) was used for the extraction of IHAs. At least 15 years of discharge records is recommended for hydrologic index analyses (Kennard *et al.*, 2010) and ~30 years when considering long-term climate change impacts (Carter *et al.*, 2007). In this study, the 30-year (1971–2000) and the available 25-year (1976-2000) periods for the Salmon River and Willow River, respectively, were used for the analyses of WRIs and IHAs. For the comparison of the simulated results with observations, three statistical tests were employed: (i) Pearson correlation test; (ii) hydrologic alteration factor (HAF); and (iii) Kolmogorov–Smirnov (KS) test.

The HAF was designed to evaluate human perturbations to the hydrological regime by considering a target range of variability for each hydrologic indicator (Richter *et al.*, 1997). The HAF has been extensively employed for the IHA analysis (Mathews and Richter, 2007; Yang *et al.*, 2008; Zolezzi *et al.*, 2009; Suen, 2010) and provides a useful test of replicability of the observed distribution by the simulated distribution, which is expressed as

$$\mathrm{HAF} = (\text{Simulated frequency} - \text{Observed frequency}) / \text{Observed frequency} \quad (1)$$

where the observed and simulated frequencies are defined as the number of samples between a target range of the 25th and 75th percentiles (inter-quartile range) of the observed dataset following the previous applications of the HAF by Mathews and Richter (2007) and Suen (2010). Although it is customary to also consider the target ranges between the 0th and 25th percentiles and between the 75th and 100th percentiles, such ranges are not considered in this application because of a limited number of sample points. The HAF values range between -1 and 1, with $\mathrm{HAF} = 0$ implying perfect agreement between observed and simulated frequencies (within the observed target range) and $\mathrm{HAF} > 0$ ($\mathrm{HAF} < 0$) implying higher (lower) frequency of simulated values than of observed values. Consistent with previous work by Suen (2010), HAF values in the range of ±0.33 (i.e. simulated frequencies are within ±33% of the observed frequencies) are considered indicative of acceptable model accuracy.

Table II. Water survey of Canada (WSC) hydrometric stations and corresponding study sub-basin characteristics. The sub-basin mean annual temperature precipitation and runoff are for 1971-2000 (Salmon), and 1976-2000 (Willow). Runoff is obtained by normalizing the hydrometric station discharges by sub-basin area

Station name	Sub-basin name	WSC ID	Sub-basin area (km^2)	Elevation range (m)			Prec. (mm/year)	Runoff (mm/year)
				Min.	Mean	Max.		
Salmon River near Price George	Salmon	08KC001	4379	571	847	1572	667	200
Willow River above Hay creek	Willow	08KD006	2855	586	1102	1981	793	363

Table III. Water resources indicators (WRIs) and indicators of hydrologic alterations (IHAs) analysed in this study (adapted from Richter *et al.*, 1996; Sanford *et al.*, 2007; Monk *et al.*, 2011)

Hydrologic Indicators	No. of indicators	Examples of hydrologic influence	Examples of ecological influence
Water resources indicators (WRIs)			
Annual volume (km^3), centre of timing of annual flow (day of occurrence of 50% annual flow between October and September) (day), median seasonal flow: Winter (December–January–February); Spring (March–April–May); Summer (June–July–August) and Autumn (September–October–November) (m^3/s)	6	Annual water balance, magnitude and timing of seasonal conditions	Availability and suitability of habitat for aquatic organisms
Indicators of hydrologic alterations (IHAs)			
Monthly flow (m^3/s)	12	Magnitude of monthly water availability	Suitable habitat availability; influence on secondary variables e.g. water temperature, oxygen
Annual mean 1-day, 3-day, 7-day, 30-day, 90-day minimum and maximum (m^3/s), and baseflow (7-day minimum/mean annual flow)	11	Magnitude of annual flood and drought conditions	Duration of stressful conditions (high and low flows)
Day of each annual 1-day minimum and maximum (water year) (day)	2	Timing of annual flood and drought conditions	Spawning cues for fish; compatibility with life cycles of organisms
No of low (annual median – 25th percentile) and high (annual median + 25th percentile) pulses in a year, median duration of low and high pulses within each year (day)	4	Frequency and duration of high and low-flow conditions	Availability of floodplain habitat; influences channel morphology e.g. bed load transport
Rise/flow rate (median of all positive/negative changes in flow between consecutive days), number of reversals (no. of switches between rising and falling period)	3	Rate and frequency of hydrograph changes	Drought (falling levels), flooding (rising levels) or desiccation stress for low mobility organisms

Additionally, a two-sample nonparametric KS test was employed to compare the distributions of observation-driven and GCM-driven WRIs and IHAs with observations. Because pooling together the ensemble of eight GCM members into one sample for statistical significance test was considered problematic (Von Storch and Zwiers, 2012), each member of the GCM-driven ensemble was compared separately with observations. The test was formulated with the null hypothesis that the observed and simulated samples come from the same population and an alternate hypothesis that the samples come from different populations. Statistically significant differences were identified when the null hypothesis is rejected at a 5% significance level. It is worth noting that the observed precipitation and temperature data serve as the downscaling target (for BCSD calibration period: 1951–1990), and observed streamflow data (1991–1995) serve as target for the VIC model calibration; therefore, the assumption of the independent samples is not completely fulfilled for the KS test. The KS test is also sensitive to the serial correlation of the datasets. However, because lag 1 correlation coefficient values for 35 (Willow) and 34 (Salmon) (out of 38) WRIs and IHAs lie within the critical values at a 5% significance level (Table of Critical Values for Pearson *r*; Sheskin, 2004), the effect of serial correlation on the KS test is considered noncritical. It should be noted too that for each test with a significance level of 5%, there is a 5% chance of rejecting the null hypothesis by accident. Therefore, there is also a chance of rejecting the null hypothesis by accident when eight GCM-driven WRIs and IHAs are considered together, which is 0.279 and 0.057 for one and more than one rejections, respectively (also at a 5% significance level) by using the binomial cumulative distribution function (Von Storch and Zwiers, 1999). Thus, in other words, there is a one in three chance of falsely rejecting the null hypothesis, and a reasonable global test of significance is more than one rejection of the null hypothesis at the 5% significance level.

RESULTS AND DISCUSSION

VIC model calibration and validation results

The calibration and validation results for the Salmon and Willow sub-basins (Figure 2) provide a general indication of

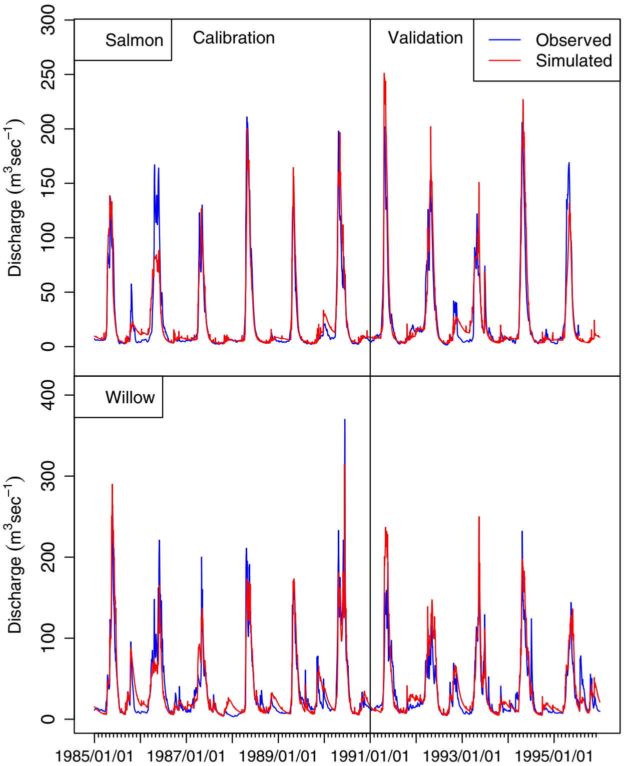

Figure 2. Observed and VIC simulated discharge (m^3/s) for calibration and validation period

the hydrologic model performance. The results show generally good agreement between the observed and simulated discharges, particularly for the annual major hydrological event driven by the spring snowmelt. The three GOF measures presented in Table IV [NSE (>0.7), LNSE (>0.7), and WBE (<0.15)] are mostly similar to those obtained in other VIC model calibration/validation studies of the Fraser River sub-basin (Schnorbus *et al.*, 2010; Shrestha *et al.*, 2012b) and are generally considered good for the hydrologic model performance rating (Moriasi *et al.*, 2007). However, the ability of the model to replicate specific elements of the hydrograph is not clear from these measures. For instance, discrepancies between the observed and modelled results can be seen from visual inspection of the plotted results (Figure 2). In particular, the magnitude of peaks shows considerable differences for both sub-basins, and the low flows look poorly replicated, especially for the Willow sub-basin. Such discrepancies between the observed and modelled results emphasize the need for a more rigorous evaluation of the model performance and consider the inherent sources of uncertainties, especially when a particular element of the hydrograph is of interest.

Performance of WRIs and IHAs

Given that the hydrologic regime of the Willow and Salmon sub-basins is nival, snowmelt has a major

Table IV. Goodness of fit measures of the VIC calibration and validation results

Sub-basin name	Statistical performance calibration (validation)		
	NSE	LNSE	WBE
Salmon	0.87 (0.83)	0.86 (0.81)	0.00 (-0.04)
Willow	0.85 (0.71)	0.86 (0.76)	-0.02 (-0.14)

influence on the dominant hydrograph event of the water year (i.e. spring rising limb, summer falling limbs of the hydrograph, and peak flow) and the WRIs and IHAs associated with it (i.e. centre of volume, spring and summer seasonal and monthly flows, magnitude and timing of maximum flows, number and duration of high pulses, rise and fall rate, and reversals). Annual low flows occur primarily in winter (especially in the Willow sub-basin), when most precipitation falls as snow (Moore and Wondzell, 2005) and the streamflow consists mainly of the baseflow component, which affects the WRIs and IHAs associated with the low flow (i.e. seasonal and monthly winter flows, magnitude and timing of minimum flows, number and duration of flow pulses, and baseflow). Hence, in this specific case, the WRIs and IHAs are largely affected by the two key processes (spring freshet-driven high flows and winter low flows) and their representation in the VIC model. Besides replication of IHAs is affected by model input/output uncertainties and model calibration, which are discussed in the following paragraphs.

The Pearson correlation coefficient between the observed and simulated results (Figure 3) shows mixed model performance. Most WRIs and monthly, maximum, and minimum flow IHAs show relatively high correlation coefficient values (>0.5), whereas most IHAs related to the number and duration of flow pulses, rise and fall rates, and reversals show low correlation values (<0.5). Such result is common to both sub-basins, with most correlation values close to each other.

Figure 4 depicts the spread of observation-driven and GCM-driven HAF values. For the observation-driven VIC model, the HAF values for WRIs and IHAs predominately (28 for Salmon and 26 for Willow, out of 38) lie between −0.33 and +0.33, indicating good model replication of the frequencies (number of samples between the 25th and 75th percentiles). Larger deviations in frequencies for the observation-driven results can be seen in some of the monthly IHAs, such as November and December (HAF < −0.67) for Willow, implying that the simulated range lies predominantly outside of the observed range, and August (HAF > 0.67) for Salmon, implying that the simulated range is narrower and lies

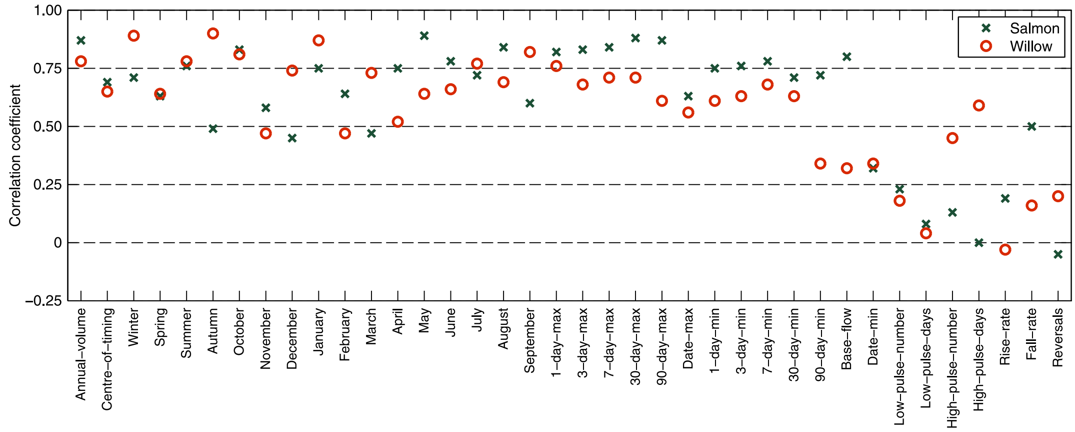

Figure 3. Correlation coefficients between observed and VIC simulated WRIs and IHAs for the Salmon (1971–2000) and Willow (1976–2000) sub-basins

mostly within the observed range. In the case of the 1- to 90-day maximum flows, the frequencies are mostly well represented, whereas the 1- to 90-day minimum flows show larger divergence from the observed frequencies for the Salmon sub-basin. Such HAF values (mostly >0.5) imply smaller spreads of the simulated minimum flows. For most observation-driven flow pulses, rise and fall rate, and number of reversals, the HAF values are mostly negative and smaller than -0.33, indicating that the simulated ranges lie predominantly outside of observed range. For most GCM-driven results, the divergence of the HAF values from zero is greater than that of observation-driven results, with some of the seasonal, monthly maximum, and monthly minimum flows showing larger deviations, especially for the Salmon sub-basin. The HAF values for some of the WRI and IHA values show large spreads; therefore, considerable differences occur when different downscaled-GCM data are used. Specifically, larger deviations of some of the monthly and flow pulse-related IHAs illustrate the difficulties in replicating the intra-annual variability. Such discrepancies cannot be discounted, especially if the modelled HAF values are used in the EFN and climate change studies. It is also important to understand the sources of uncertainty for such discrepancies, which are discussed in the following paragraphs.

Ranges of the WRIs obtained from the observed streamflows and observation-driven and GCM-driven VIC streamflow simulations are shown in Figure 5, with numbers at the bottom specifying the number of results with statistically significant differences using the KS test ($p < 0.05$). The results generally show good skill of the observation-driven VIC model in replicating annual flow volumes and centre of timing, with the median values and distributions represented reasonably well. In the case of seasonal flows, whereas the spring and autumn flows are well simulated, winter flow is over-predicted, and summer flow is under-predicted. Additionally, despite the fact that part of the observation-driven VIC streamflow is used for VIC model calibration and therefore is not independent of observed streamflow, statistically significant differences for winter (Salmon and Willow) and summer (Salmon) are obtained. Such discrepancies illustrate the difficulties in correctly simulating the seasonal streamflow values, even when annual cycles are well replicated. In the case of GCM-driven results, additional uncertainties induced by the downscaled-GCM data contribute to further discrepancies. Specifically, in addition to the statistically significant differences obtained from the observation-driven results, the centre of timing and seasonal flows obtained from some of the GCM-driven results also exhibit statistically significant differences.

The comparison of the observation-driven median monthly flows with observations also shows mixed results (Figure 6). Although the distributions of high-flow months are mostly well reproduced, the flows during low-flow months show statistically significant differences. Such discrepancies in replicating monthly flow characteristics can be attributed to inherent uncertainties in the data and the model. Firstly, errors in the observed discharge data (i.e. streamflow gauging and stage–discharge relationship uncertainties) will lead to a certain degree of deviations between the observed and simulated results. In particular, the ability to obtain reliable low-flow measurements is complicated by factors such as hyporheic flow exchange and the presence of river ice and aquatic vegetation (Hamilton, 2008), leading to considerable uncertainties in the measured discharge. The fact that over 35% of the observed data used are obtained under ice conditions (Water Survey of Canada, hydrometric data: http://www.wsc.ec.gc.ca/applications/H2O/index-eng.cfm) implies considerable uncertainty in the measured low-flow values. In turn, the lack of consideration of hydraulic effects of river ice on streamflow in the

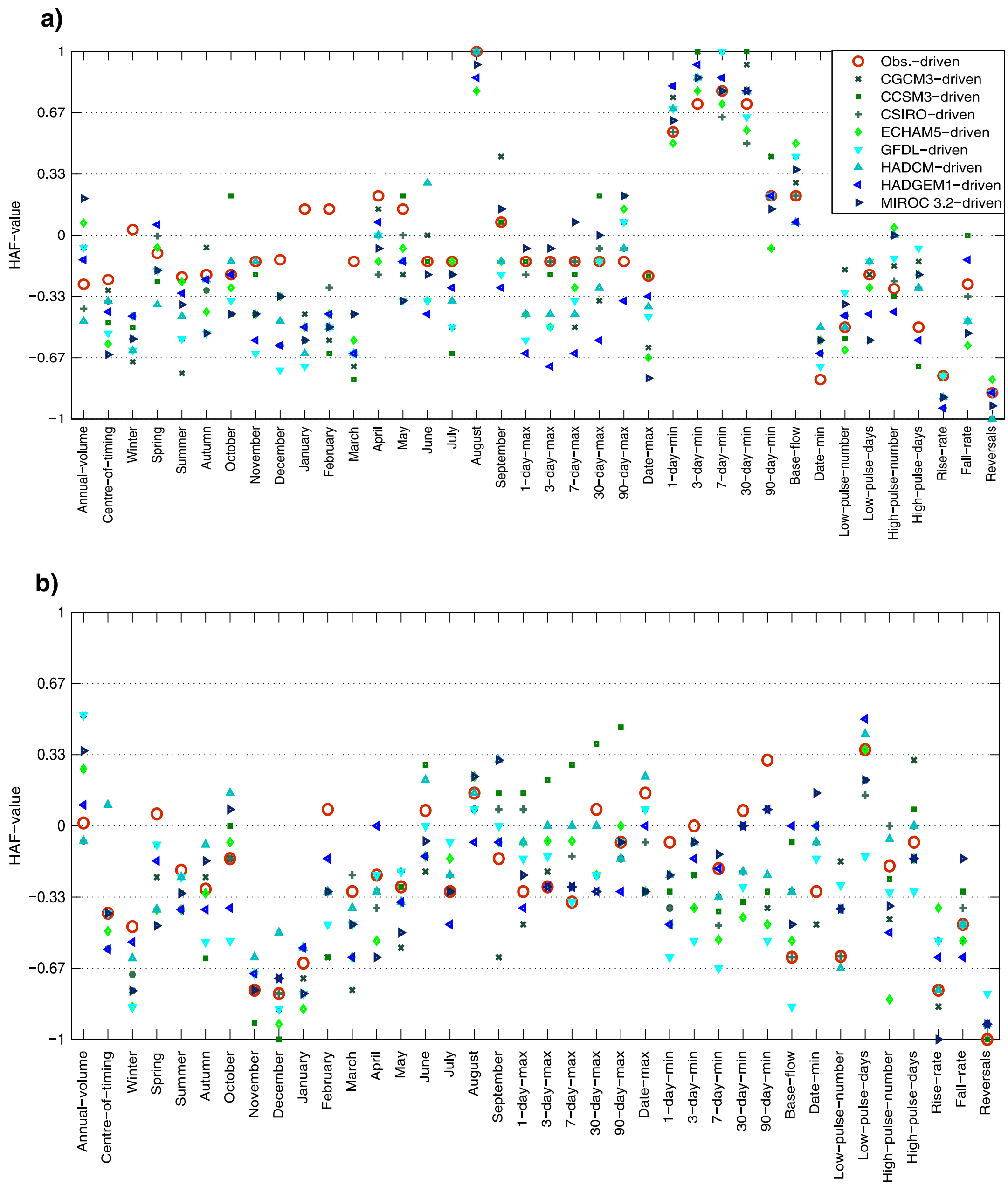

Figure 4. Hydrologic alteration factor (HAF) values for observation-driven VIC and GCM-driven VIC results for (a) Salmon (1971–2000) and (b) Willow (1976–2000) sub-basins. The HAF values range between −1 and 1, with HAF = 0 implying perfect agreement, HAF > 0 implying that the simulated values are more frequent (within the target range) and have narrower spread than the observed, and HAF < 0 implying that simulated values are less frequent and lie outside the observed target range

VIC model (this is also the case for other hydrologic models) also leads to uncertainty. Secondly, errors in precipitation and temperature data (i.e. low station density at high-elevation areas) affect the runoff generation processes, such as snow storage, melt and soil moisture, and resulting monthly flow response. Thirdly, the GOF measures, NSE, WBE, and to a lesser extent LNSE, are predominantly tuned to the high-flow period and do not directly consider the intra-annual variability such as monthly distribution of flow. Fourthly, structural uncertainty of the hydrologic model, no doubt, also contributes to the low-flow discrepancies. Although such problems in replicating the low flows are not unique to the VIC model, previous studies (e.g. Warrach *et al.*, 2002; Demaria *et al.*, 2007; Surfleet *et al.*, 2012) also demonstrated difficulties in reproducing low flows by the VIC model, which may be partly related to the conceptual representation of soil moisture processes. Furthermore, fractions of the two sub-basins are covered by wetlands (4.5% and 1.8% of the Salmon and Willow sub-basins, respectively,

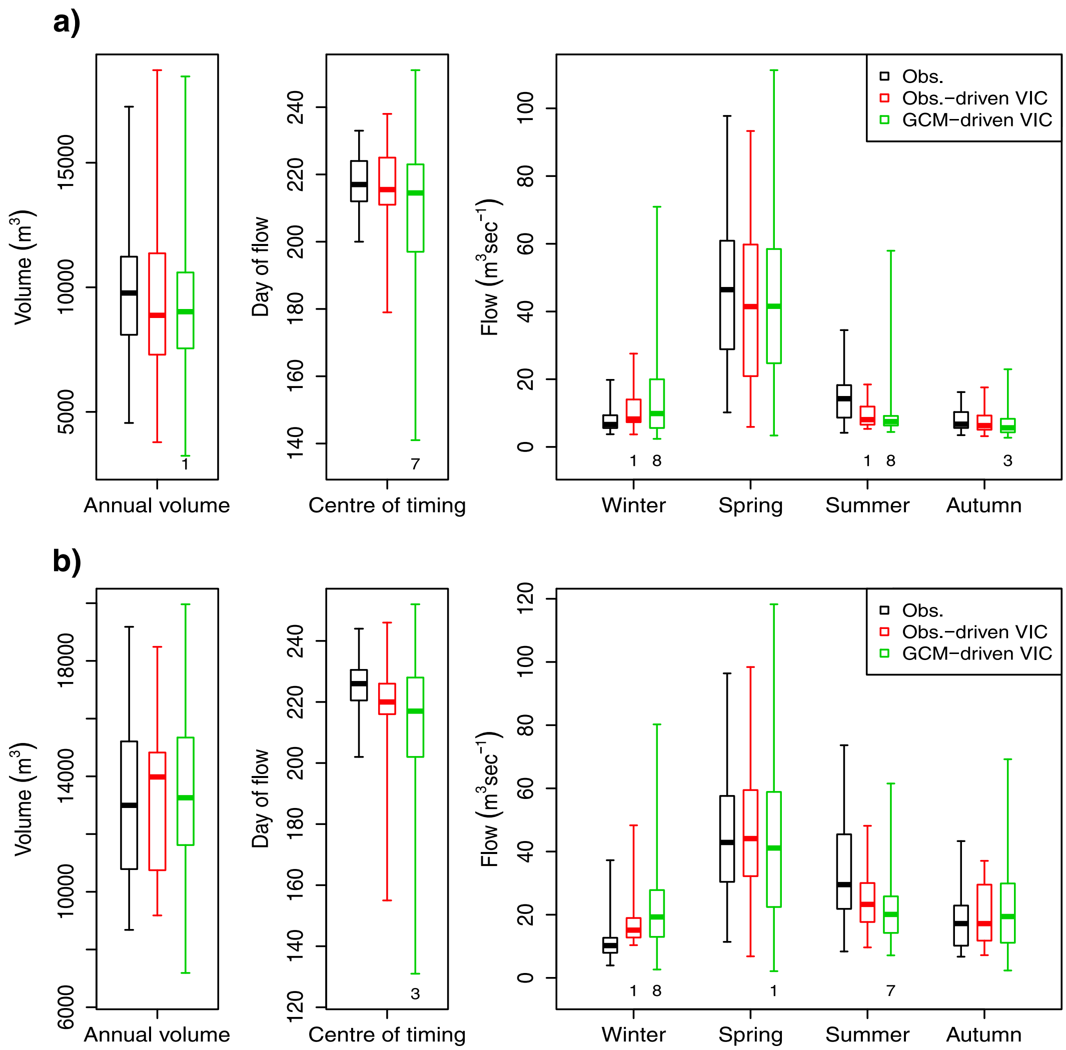

Figure 5. Observations, observation-driven VIC and GCM-driven VIC water resource indicators for (a) Salmon (1971–2000) and (b) Willow (1976–2000) sub-basins. Each box plot illustrates the median and inter-quartile range, and the whiskers represent the upper and lower limits. The numbers at the bottom refer to the number of observation-driven (out of one) and GCM-driven (out of eight) results with statistically significant differences, and empty fields refer to those with no statistically significant differences based on the KS test

calculated from BC Land and Resource Data, http://archive.ilmb.gov.bc.ca/lrdw/), and the lack of explicitly considering wetlands in the version of VIC model used could have some influence on the simulated flows. Bowling and Lettenmaier (2010) implemented a dynamic lake/wetland model in the VIC model version 4.1.1 and found improvement in the timing and shape of the snowmelt-induced hydrograph (in watersheds with 50% and 70% seasonal flooding). The replication of the timing of snowmelt-induced hydrograph is not an issue for the two sub-basins considered, but wetlands could affect annual water balance, and the effect of wetlands on hydrologic response should be considered in future studies. Another factor that may cause such divergence of modelled results is the impact of the short calibration period. Six years (1985–1990) of data covering 64% and 55% of the range (difference between maximum and minimum) of 25 years (Willow) and 30 years (Salmon) was used; therefore, the variability of the analysis period (25–30 years) was not fully represented.

Furthermore, as in the case of seasonal distribution, the results also show discrepancies in the monthly distribution of the flow, with overestimation during winter months and underestimation during summer months (especially for the Willow sub-basin). Such discrepancies indicate the presence of systematic bias, which is also supported by high correlation coefficients between the observed and observation-driven VIC results where statistically significant differences in the monthly IHAs occur (i.e. Salmon: February 0.64, July 0.72, August 0.84, September 0.60; Willow: November 0.47, December 0.74, January 0.87, September 0.82; Figure 3). This suggests a possibility of

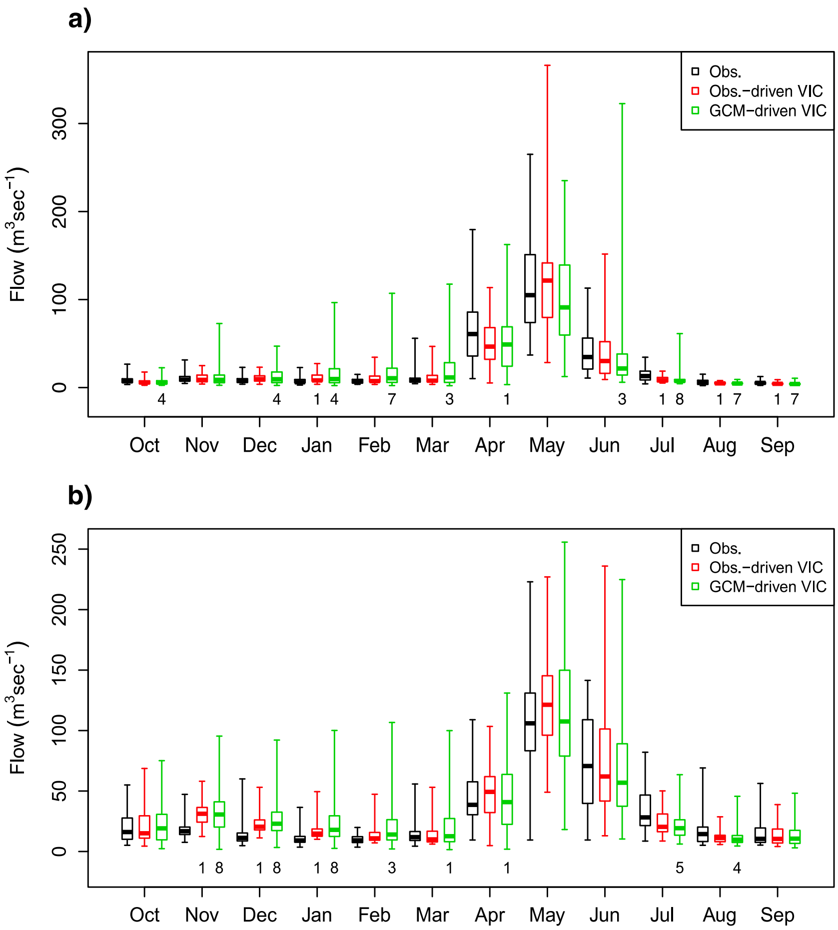

Figure 6. Observations, observation-driven VIC and GCM-driven median monthly flows for (a) salmon (1971–2000) and (b) willow (1976–2000) sub-basins. Each box plot illustrates the median and inter-quartile range, and the whiskers upper and lower limits. The numbers at the bottom refer to the number of observation-driven (out of one) and GCM-driven (out of eight) results with statistically significant differences, and empty fields with no statistically significant differences based on the KS test

improving the model performance, e.g. by using the monthly flows as objective functions for hydrologic model calibration. In the case of GCM-driven results, more than two ensemble members (which is considered a threshold for the global null hypothesis test) exhibit statistically significant differences for additional winter, spring, and summer months, further illustrating the role of downscaled-GCM uncertainties. Specifically, although the BCSD downscaling method is effective at capturing many aspects of historical daily temperature and precipitation variability and is competitive in this respect with many other well-known statistical downscaling methods (Bürger *et al*., 2012), some uncertainties in extremes will no doubt remain, especially given that the BCSD performs downscaling at a monthly time step and disaggregates into daily time step.

Figure 7 shows the comparison of the IHAs related to the annual maximum and minimum flows, their timings, and baseflow ratio (7-day minimum flow/mean annual flow). As previously stated, the observed maximum and minimum flow data (especially for smaller duration events) can be subject to considerable uncertainties, which can affect the ability of the model to replicate these events. Nevertheless, the range and median values of the annual maximum and minimum flows are reproduced reasonably well by the observation-driven VIC model for both sub-basins. The 1-day maximum flow date is also replicated reasonably well. Such plausible reproduction of magnitude and timing of the peak flow implies the capability of the VIC model in representing the snowmelt-driven hydrologic response. The agreement of the observation-driven minimum flow is achieved despite problems in replicating IHAs for low-flow months (Figures 6 and 7). A comparison of the 1-day minimum flow date (Figure 7) provides a possible explanation; the median simulated 1-day minimum flow occurs later than the observed flows (especially for the Willow sub-basin), causing discrepancies in the low-flow months. This apparent mismatch in the magnitude and timing of the minimum flows also provides an explanation for the systematic model bias in the simulated monthly flows. This also points to a limitation of the model

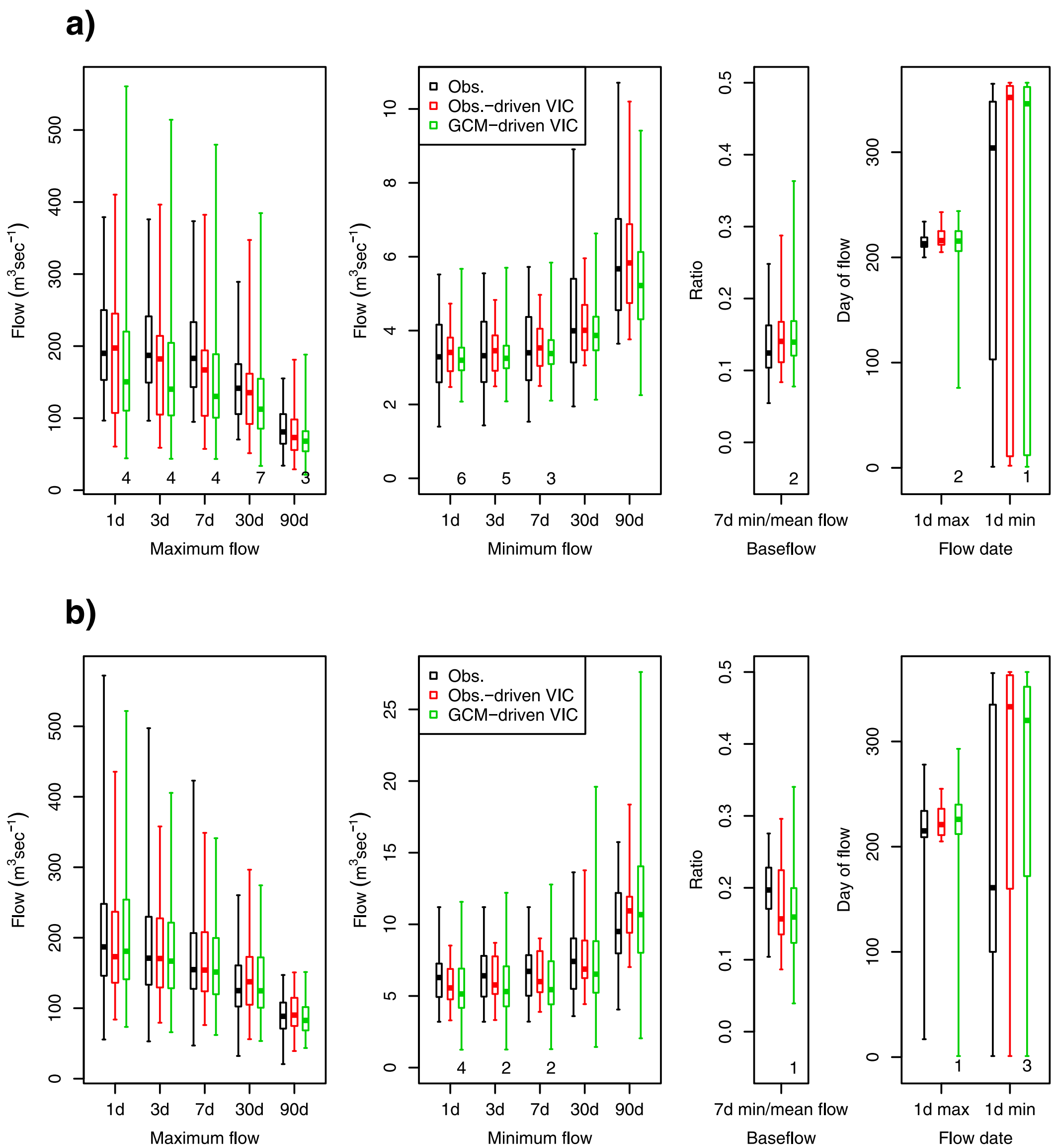

Figure 7. Observations, observation-driven VIC, and GCM-driven VIC maximum and minimum flows for (a) Salmon (1971–2000) and (b) Willow (1976–2000) sub-basins. Each box plot illustrates the median and inter-quartile range, and the whiskers represent the upper and lower limits. The numbers at the bottom refer to the number of observation-driven (out of one) and GCM-driven (out of eight) results with statistically significant differences, and empty fields refer to those with no statistically significant differences based on the KS test

calibration, which although considers the magnitude of low flow (i.e. by using the LNSE) does not consider the associated timing. A possible method for improving the model performance for a particular WRI or IHA (such as monthly flow) might be to use an objective function best suited for that WRI or IHA, but that would potentially lead to multiple calibrations for a suite of WRIs and IHAs. In the case of baseflow ratio, the distribution is mostly well reproduced for the Salmon sub-basin, whereas for the Willow sub-basin, most values show a negative bias. The baseflow ratio correlation for the Willow sub-basin is also low (Figure 3), indicating poor replication of interannual variability. A closer examination of the results reveals that biases in the annual mean flow (mostly positive bias) (Figure 5) and 7-day low flow (mostly negative bias) (Figure 7) affect the annual values of baseflow ratios and lead to low correlation with observed baseflow ratio. A lower correlation was also obtained for the 90-day minimum flow (which occurred in winter months) for the Willow sub-basin (Figure 3), which arises from the larger error in winter flows (Figure 6). For the GCM-driven maximum and minimum flows, their timing, and their baseflow ratios, the results depict statistically significant differences ($p < 0.05$) for more than two members of the ensemble (especially for the Salmon sub-basin), again illustrating the role of downscaled GCMs on the simulated outputs.

In the case of flow pulses, rise and fall rates, and reversals, both observation-driven and GCM-driven VIC model results differ considerably from observations (Figure 8). Specifically, a low correlation (Figure 3) and statistically significant ($p < 0.05$) differences in reproducing the number of low-pulse (defined as annual median − 25th percentile) events and their durations further emphasize the problem in modelling low flows. The model exhibits better skill in reproducing high-flow-pulse (defined as annual median + 25th percentile) events and their durations, which is related to the good model skill in replicating spring snowmelt-generated streamflow period (Figure 4). But the problems in reproducing hydrograph rise (median of all positive changes) and fall rate (median of all negative changes) and the number of hydrograph reversals (number of switches between rising and falling period) illustrate another limitation. Specifically, the simulated rise and fall rates and number of reversals (at a daily time step) are higher than those seen in the observed streamflows, and Pearson correlation coefficients between observed and observation-driven VIC results for these seven IHAs (i.e. number and duration of low-flow and high-flow pulses, rise and fall rate, and reversals) are mostly low (Figure 3). Although the issues related to hydrologic model structure and parameter uncertainties play a part in such mismatches, the uncertainty in hydrometric data can also be an important factor for such deviations. Specifically, given that over 35% of the discharge data used are obtained under ice

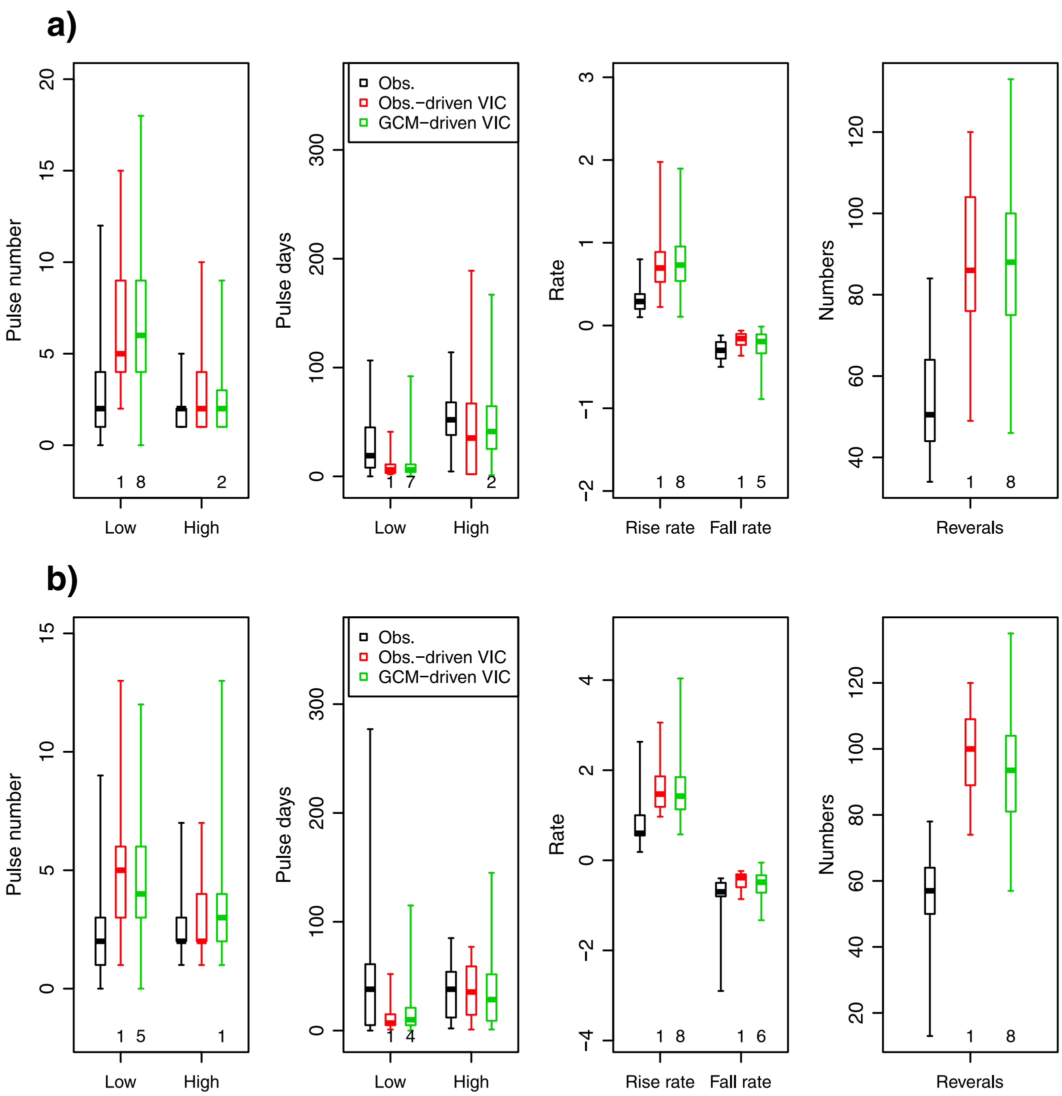

Figure 8. Observations, observation-driven VIC, and GCM-driven VIC low and high pulse numbers, low and high pulse days, rise and fall rates, and number of reversals for (a) Salmon (1971–2000) and (b) Willow (1976–2000) sub-basins. Each box plot illustrates the median and inter-quartile range, and the whiskers represent the upper and lower limits. The numbers at the bottom refer to the number of observation-driven (out of one) and GCM-driven (out of eight) results with statistically significant differences, and empty fields refer to those with no statistically significant differences based on the KS test

conditions, recorded daily fluctuations (and therefore the number and duration of low-flow pulses, rise and fall rates, and reversals) may not be representative of true daily fluctuations, thereby causing divergence of the modelled results.

Conclusions and future work

This study evaluated the general replicability of six WRIs and 32 IHAs that are proposed for ecological flow needs and climate change assessments. A typical hydrologic model set-up (VIC) for two headwater sub-basins (Salmon and Willow) of the Fraser river basin was used for the evaluation. This study is not intended to be an evaluation of the capabilities of the VIC model (and the processes it represents) or the two sub-basin characteristics. Going beyond a typical calibration/validation framework, we focused on providing insights on the general capabilities and uncertainties surrounding the replications of WRIs and IHAs. Overall, the replicability of the WRIs and IHAs ranges from good to limited. The results showed good performance for the annual flow and snowmelt-driven monthly and peak flows and limited performance for low flow, number and duration of flow pulses, rise and fall rates, and reversals. A number of factors that cause such discrepancies were identified, such as the lack of explicit consideration of these indicators in model calibration, errors in meteorological inputs, and errors in observed and/or simulated discharge data. For example, in this particular case, sub-basin-specific factors such as uncertainties in the discharge data during river ice conditions likely contribute to the discrepancies for some of the IHAs such as minimum flows, number and duration of low-flow pulses, rise and fall rates, and reversals. Additionally, model structural uncertainties affect the model's ability to replicate observed flows (especially low flow) and contribute to the discrepancies. The observation-driven VIC model generally performed better compared with the GCM-driven VIC model (especially for the maximum and minimum flows), indicating the role of additional errors induced by using downscaled-GCM data. In the comparison of the model performance for two sub-basins, the results are mostly similar. Similar problems in replicating some of the indicators (i.e. monthly flows, number and duration of flow pulses, rise and fall rates, and reversals) imply that the errors are independent of sub-basin characteristics and are related to model input and output, parameter, and/or structural uncertainties.

The success, or lack thereof, in capturing the statistical characteristics of WRIs and IHAs using a hydrologic model raises important questions regarding the applicability of hydrologic-model-simulated indicators in EFN and climate change assessments. The discrepancies in simulating some of the indicators emphasize a need for a detailed evaluation of modelled indicators, beyond traditional calibration/validation. Although some of the performance issues are specific to the sub-basins and/or the chosen hydrologic model (VIC), discrepancies between the simulated and observed results remain unavoidable in hydrologic modelling of river systems because of different sources of inherent uncertainties. The lack of agreement of some of the indicators (in this case, the number and duration flow pulses, rise and fall rates, and reversals) illustrates the limitation of the flow data and hydrologic model and emphasizes the need to exercise caution in the use of model-simulated indicators. Not surprisingly, the model performs better for the indicators for which it is calibrated (i.e. peak flow). Hence, if certain indicators such as low-flow months are of interest, it will be a good idea to explicitly include the indicators for calibration. In this particular case, given that the errors in some of the indicators (e.g. monthly flows and low flows) are likely due to systematic bias suggests the possibility of improving the model performance by using the monthly flows as objective functions for hydrologic model calibration. For example, formulation of an objective function that takes into account both magnitude and timing of low flow can potentially lead to improved performance for low-flow months. Model calibration with additional variables such as snow (when available) should also be considered for a better reproduction of the physical processes. Another factor not specifically considered in this study is the improvement of the hydrologic model or application of the hydrologic model best suited for a WRI or IHA of interest. For instance, the consideration of wetlands and/or use of a finer grid resolution could potentially improve the performance for some of the WRIs and IHAs.

From the climate change studies perspective, additional uncertainties induced by the downscaled-GCM data also emphasize a need for caution in the use of simulated indicators. For instance, in this study, the discrepancies in replicating the maximum and minimum flows using the GCM-driven inputs raise a question on the model's ability to simulate future changes in extremes. This is an important question given that the current generation of climate models projects a future increase in precipitation intensity (Meehl *et al.*, 2007b), which will have implications to flow extremes. An equally important issue pertains to the reliability of projected future changes that are expressed as the difference between the modelled future and baseline results, when the baseline results are significantly different from observations. These issues, such as the analysis of modelled results for two historical periods, should be a focus of future studies. In summary, hydrologic models can be powerful tools for ecological flow needs and climate change assessments, provided that the ability of the models and associated uncertainties are understood and only the reliably replicated indicators of change are included in an assessment.

ACKNOWLEDGEMENTS

The authors are thankful to Francis Zwiers (PCIC) for providing useful comments to the manuscript. The authors thank the reviewer for comments towards an improved version of the manuscript.

REFERENCES

Bennett KE, Werner AT, Schnorbus M. 2012. Uncertainties in Hydrologic and Climate Change Impact Analyses in Headwater Basins of British Columbia. *Journal of Climate* **25**(17): 5711–5730.

Beven K. 2006. A manifesto for the equifinality thesis. *Journal of Hydrology* **320**(1–2): 18–36.

Blöschl G, Montanari A. 2010. Climate change impacts – throwing the dice? *Hydrological Processes* **24**(3): 374–381.

Bowling LC, Lettenmaier DP. 2010. Modeling the effects of lakes and wetlands on the water balance of arctic environments. *Journal of Hydrometeorology* **11**(2): 276–295.

Bürger G, Murdock TQ, Werner AT, Sobie SR, Cannon AJ. 2012. Downscaling Extremes – An Intercomparison of Multiple Statistical Methods for Present Climate. *Journal of Climate* **25**(12): 4366–4388.

Carlisle DM, Falcone J, Wolock DM, Meador MR, Norris RH. 2010. Predicting the natural flow regime: models for assessing hydrological alteration in streams. *River Research and Applications* **26**(2): 118–136.

Carter TR, *et al*. 2007. General guidelines on the use of scenario data for climate impact and adaptation assessment, Task Group on Data and Scenario Support for Impact and Climate Assessment, IPCC. http://www.ipcc-data.org/guidelines/TGICA_guidance_sdciaa_v2_final.pdf (accessed May 2013).

Chang HJ, Jung IW. 2010. Spatial and temporal changes in runoff caused by climate change in a complex large river basin in Oregon. *Journal of Hydrology* **388**: 186–207.

Chen J, Brissette FP, Leconte R. 2011. Uncertainty of downscaling method in quantifying the impact of climate change on hydrology. *Journal of Hydrology* **401**: 190–202.

Collins M, Tett S, Cooper C. 2001. The internal climate variability of HadCM3, a version of the Hadley Centre coupled model without flux adjustments. *Climate Dynamics* **17**(1): 61–81.

Collins W, Bitz C, Blackmon M, Bonan G, Bretherton C, Carton J, Chang P, Doney S, Hack J, Henderson T, Kiehl J, Large W, McKenna D, Santer B, Smith R. 2006. The Community Climate System Model version 3 (CCSM3). *Journal of Climate* **19**(11): 2122–2143.

Delworth T, Broccoli A, Rosati A, Stouffer R, Balaji V, Beesley J, Cooke W, Dixon K, Dunne J, Dunne K, Durachta J, Findell K, Ginoux P, Gnanadesikan A, Gordon C, Griffies S, Gudgel R, Harrison M, Held I, Hemler R, Horowitz L, Klein S, Knutson T, Kushner P, Langenhorst A, Lee H, Lin S, Lu J, Malyshev S, Milly P, Ramaswamy V, Russell J, Schwarzkopf M, Shevliakova E, Sirutis J, Spelman M, Stern W, Winton M, Wittenberg A, Wyman B, Zeng F, Zhang R. 2006. GFDL's CM2 global coupled climate models. Part I: formulation and simulation characteristics. *Journal of Climate* **19**(5): 643–674.

Demaria EM, Nijssen B, Wagener T. 2007. Monte Carlo sensitivity analysis of land surface parameters using the Variable Infiltration Capacity model. *Journal of Geophysical Research – Atmospheres* **112**: D11113.

Dibike YB, Coulibaly P. 2005. Hydrologic impact of climate change in the Saguenay watershed: comparison of downscaling methods and hydrologic models. *Journal of Hydrology* **307**: 145–163.

Elsner MM, Cuo L, Voisin N, Deems JS, Hamlet AF, Vano JA, Mickelson KEB, Lee SY, Lettenmaier DP. 2010. Implications of 21st century climate change for the hydrology of Washington State. *Climatic Change* **102**(1–2): 225–260.

Fenicia F, Solomatine DP, Savenije HHG, Matgen P. 2007. Soft combination of local models in a multi-objective framework. *Hydrology and Earth System Sciences Systems* **11**: 1797–1809.

Gibson CA, Meyer JL, Poff NL, Hay LE, Georgakakos A. 2005. Flow regime alterations under changing climate in two river basins: implications for freshwater ecosystems. *River Research and Applications* **21**(8): 849–864.

Hamilton S. 2008. Sources of uncertainty in Canadian low flow hydrometric data. *Canadian Water Resources Journal* **33**(2): 125–136.

Hamilton AS, Moore RD. 2012. Quantifying uncertainty in streamflow records. *Canadian Water Resources Journal* **37**(1): 3–21.

Hidalgo HG, Das T, Dettinger MD, Cayan DR, Pierce DW, Barnett TP, Bala G, Mirin A, Wood AW, Bonfils C, Santer BD, Nozawa T. 2009. Detection and attribution of streamflow timing changes to climate change in the Western United States. *Journal of Climate* **22**(13): 3838–3855.

K-1 Model Developers. 2004. K-1 coupled GCM (MIROC) description. Hasumi H, Emori S (eds). K-1 Technical Report 1, Center for Climate System Research, University of Tokyo. http://www.ccsr.u-tokyo.ac.jp/kyosei/hasumi/MIROC/tech-repo.pdf (accessed January 2013).

Kay AL, Davies HN, Bell VA, Jones RG. 2008. Comparison of uncertainty sources for climate change impacts: flood frequency in England. *Climatic Change* **92**(1–2): 41–63.

Kennard MJ, Mackay SJ, Pusey BJ, Olden JD, Marsh N. 2010. Quantifying uncertainty in estimation of hydrologic metrics for ecohydrological studies. *River Research and Applications* **26**(2): 137–156.

Kim B, Kim B, Kwon H. 2011. Assessment of the impact of climate change on the flow regime of the Han River basin using indicators of hydrologic alteration. *Hydrological Processes* **25**(5): 691–704.

Knight RR, Gain WS, Wolfe WJ. 2012. Modelling ecological flow regime: an example from the Tennessee and Cumberland River basins. *Ecohydrology* **5**(5): 613–627.

Krause P, Boyle DP, Bäse F. 2005. Comparison of different efficiency criteria for hydrological model assessment. *Advances in Geosciences* **5**: 89–97.

Legates DR, McCabe GJ. 1999. Evaluating the use of "goodness-of-fit" measures in hydrologic and hydroclimatic model validation. *Water Resources Research* **35**(1): 233.

Liang X, Lettenmaier DP, Wood EF, Burges SJ. 1994. A simple hydrologically based model of land-surface water and energy fluxes for general-circulation models. *Journal of Geophysical Research – Atmospheres* **99**(D7): 14415–14428.

Liang X, Wood EF, Lettenmaier DP. 1996. Surface soil moisture parameterization of the VIC-2L model: evaluation and modification. *Global and Planetary Change* **13**(1–4): 195–206.

Liang X, Xie Z, Huang M. 2003. A new parameterization for surface and groundwater interactions and its impact on water budgets with the Variable Infiltration Capacity (VIC) land surface model. *Journal of Geophysical Research* **108**(D16): 8613.

Lin Y, Wei X. 2008. The impact of large-scale forest harvesting on hydrology in the Willow watershed of Central British Columbia. *Journal of Hydrology* **359**(1–2): 141–149.

Lytle DA, Poff NL. 2004. Adaptation to natural flow regimes. *Trends in Ecology & Evolution* **19**(2): 94–100.

Martin G, Ringer M, Pope V, Jones A, Dearden C, Hinton T. 2006. The physical properties of the atmosphere in the new Hadley Centre Global Environmental Model (HadGEM1). Part I: model description and global climatology. *Journal of Climate* **19**(7): 1274–1301.

Mathews R, Richter BD. 2007. Application of the Indicators of Hydrologic Alteration Software in Environmental Flow Setting 1. *JAWRA Journal of the American Water Resources Association* **43**(6): 1400–1413.

Maurer EP, Hidalgo HG, Das T, Dettinger MD, Cayan DR. 2010. The utility of daily large-scale climate data in the assessment of climate change impacts on daily streamflow in California. *Hydrology and Earth System Sciences Systems* **14**(6): 1125–1138.

Meehl GA, Curt Covey, Taylor KE, Delworth T, Stouffer RJ, Latif M, McAvaney B, Mitchell JFB. 2007a. THE WCRP CMIP3 multimodel dataset: a new era in climate change research. *Bulletin of the American Meteorological Society* **88**(9): 1383–1394.

Meehl GA, Stocker TF, Collins WD, Friedlingstein P, Gaye AT, Gregory JM, Kitoh A, Knutti R, Murphy JM, Noda A, Raper SCB, Watterson IG, Weaver AJ, Zhao Z-C. 2007b. *Global Climate Projections. Climate Change 2007b. Working Group I: The Physical Science Basis*. Cambridge University Press: Cambridge, United Kingdom and New York, NY, USA; 747–845.

Merritt WS, Alila Y, Barton M, Taylor B, Cohen S, Neilsen D. 2006. Hydrologic response to scenarios of climate change in sub watersheds of the Okanagan basin, British Columbia. *Journal of Hydrology* **326**: 79–108.

Monk WA, Peters DL, Allen Curry R, Baird DJ. 2011. Quantifying trends in indicator hydroecological variables for regime-based groups of Canadian rivers. *Hydrological Processes* **25**(19): 3086–3100.

Moore RD, Wondzell SM. 2005. Physical hydrology and the effects of forest harvesting in the Pacific Northwest: a review. *Journal of the American Water Resources Association* **41**(4): 763–784.

Moriasi DN, Arnold JG, Van Liew MW, Bingner RL, Harmel RD, Veith TL. 2007. Model evaluation guidelines for systematic quantification of accuracy in watershed simulations. *Transactions of the ASABE* **50**(3): 885–900.

Murphy JC, Knight RR, Wolfe WJ, S. Gain W. 2013. Predicting Ecological Flow Regime at Ungaged Sites: A Comparison of Methods. *River Research and Applications* **29**(5): 660–669.

Najafi MR, Moradkhani H, Jung IW. 2011. Assessing the uncertainties of hydrologic model selection in climate change impact studies. *Hydrological Processes* **25**(18): 2814–2826.

Nash JE, Sutcliffe JV. 1970. River flow forecasting through conceptual models part I – a discussion of principles. *Journal of Hydrology* **10**(3): 282–290.

Neilsen D, Duke G, Taylor B, Byrne J, Kienzle S, Van der Gulik T. 2010. Development and verification of daily gridded climate surfaces in the Okanagan Basin of British Columbia. *Canadian Water Resources Journal* **35**(21): 131–154.

Pelletier P. 1990. A review of techniques used by Canada and other northern countries for measurement and computation of streamflow under ice conditions. *Nordic Hydrology* **21**(4–5): 317–340.

Peters DL, Baird DJ, Monk WA, Armanini DG. 2012. Establishing standards and assessment criteria for ecological instream flow needs in agricultural regions of Canada. *Journal of Environment Quality* **41**(1): 41.

Poff NL, Allan JD, Bain MB, Karr JR, Prestegaard KL, Richter BD, Sparks RE, Stromberg JC. 1997. The natural flow regime. *BioScience* **47**(11): 769–784.

Poff NL, Richter BD, Arthington AH, Bunn SE, Naiman RJ, Kendy E, Acreman M, Apse C, Bledsoe BP, Freeman MC, Henriksen J, Jacobson RB, Kennen JG, Merritt DM, O'keeffe JH, Olden JD, Rogers K, Tharme RE, Warner A. 2010. The ecological limits of hydrologic alteration (ELOHA): a new framework for developing regional environmental flow standards. *Freshwater Biology* **55**(1): 147–170.

Prudhomme C, Davies H. 2008a. Assessing uncertainties in climate change impact analyses on the river flow regimes in the UK. Part 1: baseline climate. *Climatic Change* **93**: 177–195.

Prudhomme C, Davies H. 2008b. Assessing uncertainties in climate change impact analyses on the river flow regimes in the UK. Part 2: future climate. *Climatic Change* **93**: 197–222.

Pushpalatha R, Perrin C, Moine NL, Andréassian V. 2012. A review of efficiency criteria suitable for evaluating low-flow simulations. *Journal of Hydrology* **420–421**: 171–182.

Quintana Seguí P, Ribes A, Martin E, Habets F, Boé J. 2010. Comparison of three downscaling methods in simulating the impact of climate change on the hydrology of Mediterranean basins. *Journal of Hydrology* **383**(1–2): 111–124.

Reichler T, Kim J. 2008. How well do coupled models simulate today's climate? *Bulletin of the American Meteorological Society* **89**: 303–311.

Renard B, Kavetski D, Kuczera G, Thyer M, Franks SW. 2010. Understanding predictive uncertainty in hydrologic modeling: the challenge of identifying input and structural errors. *Water Resources Research* **46**: 22.

Richter BD, Baumgartner JV, Powell J, Braun DP. 1996. A method for assessing hydrologic alteration within ecosystems. *Conservation Biology* **10**(4): 1163–1174.

Richter B, Baumgartner J, Wigington R, Braun D. 1997. How much water does a river need? *Freshwater Biology* **37**(1): 231–249.

Roeckner E, Brokopf R, Esch M, Giorgetta M, Hagemann S, Kornblueh L, Manzini E, Schlese U, Schulzweida U. 2006. Sensitivity of simulated climate to horizontal and vertical resolution in the ECHAM5 atmosphere model. *Journal of Climate* **19**(16): 3771–3791.

Rotstayn L, Collier M, Dix M, Feng Y, Gordon H, O'Farrell S, Smith I, Syktus J. 2010. Improved simulation of Australian climate and ENSO-related rainfall variability in a global climate model with an interactive aerosol treatment. *International Journal of Climatology* **30**(7): 1067–1088.

Sanford SE, Creed IF, Tague CL, Beall FD, Buttle JM. 2007. Scale-dependence of natural variability of flow regimes in a forested landscape. *Water Resources Research* **43**: 15.

Schnorbus M, Bennett K, Werner A. 2010. *Quantifying the Water Resource Impacts of Mountain Pine Beetle and Associated Salvage Harvest Operations across a Range of Watershed Scales: Hydrologic Modeling of the Fraser River Basin*. Information Report BC-X-423, Natural Resources Canada: Canadian Forest Service, Pacific Forestry Centre, Victoria, BC.

Schnorbus MA, Bennett KE, Werner AT, Berland AJ. 2011. *Hydrologic Impacts of Climate Change in the Peace, Campbell and Columbia Watersheds, British Columbia, Canada*. Pacific Climate Impacts Consortium, University of Victoria: Victoria, BC; 157.

Schnorbus MA, Werner AT, Bennett KE. in press. Impacts of climate change in three hydrologic regimes in British Columbia, Canada. *Hydrological Processes*. doi:10.1002/hyp.9661

Scinocca JF, McFarlane NA, Lazare M, Li J, Plummer D. 2008. Technical note: the CCCma third generation AGCM and its extension into the middle atmosphere. *Atmospheric Chemistry and Physics* **8**: 7055–7074.

Sheskin D. 2004. *Handbook of Parametric and Nonparametric Statistical Procedures*. CRC Press: Boca Raton.

Shrestha RR, Rode M. 2008. Multi-objective calibration and fuzzy preference selection of a distributed hydrological model. *Environmental Modelling & Software* **23**: 1384–1395.

Shrestha RR, Bardossy A, Nestmann F. 2007. Analysis and propagation of uncertainties due to the stage-discharge relationship: a fuzzy set approach. *Hydrological Sciences Journal – Journal Des Sciences Hydrologiques* **52**: 595–610.

Shrestha RR, Dibike YB, Prowse TD. 2012a. Modelling of climate-induced hydrologic changes in the Lake Winnipeg watershed. *Journal of Great Lakes Research* **38**(3): 83–94.

Shrestha RR, Schnorbus MA, Werner AT, Berland AJ. 2012b. Modelling spatial and temporal variability of hydrologic impacts of climate change in the Fraser River basin, British Columbia, Canada. *Hydrological Processes* **26**(12): 1840–1860.

Stahl K, Moore RD, Floyer JA, Asplin MG, McKendry IG. 2006. Comparison of approaches for spatial interpolation of daily air temperature in a large region with complex topography and highly variable station density. *Agricultural and Forest Meteorology* **139**(3–4): 224–236.

Staudinger M, Stahl K, Seibert J, Clark MP, Tallaksen LM. 2011. Comparison of hydrological model structures based on recession and low flow simulations. *Hydrology and Earth System Sciences* **15**(11): 3447–3459.

Suen J-P. 2010. Potential impacts to freshwater ecosystems caused by flow regime alteration under changing climate conditions in Taiwan. *Hydrobiologia* **649**: 115–128.

Surfleet CG, Tullos D, Chang H, Jung I-W. 2012. Selection of hydrologic modeling approaches for climate change assessment: A comparison of model scale and structures. *Journal of Hydrology* **464–465**: 233–248.

The Nature Conservancy. 2009. Indicators of Hydrologic Alteration Version 7.1: User's Manual.

Todini I. 1996. The ARNO rainfall—runoff model. *Journal of Hydrology* **175**(1-4): 339–382.

Toth B, Pietroniro A, Conly FM, Kouwen N. 2006. Modelling climate change impacts in the Peace and Athabasca catchment and delta: I – hydrological model application. *Hydrological Processes* **20**(19): 4197–4214.

Von Storch H, Zwiers FW. 1999. *Statistical Analysis in Climate Research*. Cambridge University Press: Cambridge; New York.

Von Storch H, Zwiers F. 2012. Testing ensembles of climate change scenarios for "statistical significance". *Climatic Change*. doi: 10.1007/s10584-012-0551-0

Wagener T. 2003. Evaluation of catchment models. *Hydrological Processes* **17**(16): 3375–3378.

Warrach K, Stieglitz M, Mengelkamp H-T, Raschke E. 2002. Advantages of a Topographically Controlled Runoff Simulation in a Soil–Vegetation–Atmosphere Transfer Model. *Journal of Hydrometeorology* **3**(2): 131–148.

Wenger SJ, Luce CH, Hamlet AF, Isaak DJ, Neville HM. 2010. Macroscale hydrologic modeling of ecologically relevant flow metrics. *Water Resources Research* **46**: 10.

Werner AT. 2011. *BCSD Downscaled Transient Climate Projections for Eight Select GCMs over British Columbia, Canada*. Pacific Climate Impacts Consortium, University of Victoria: Victoria, BC; 63.

Wood A, Leung L, Sridhar V, Lettenmaier D. 2004. Hydrologic implications of dynamical and statistical approaches to downscaling climate model outputs. *Climatic Change* **62**(1-3): 189–216.

Yang T, Zhang Q, Chen YD, Tao X, Xu C, Chen X. 2008. A spatial assessment of hydrologic alteration caused by dam construction in the middle and lower Yellow River, China. *Hydrological Processes* **22**(18): 3829–3843.

Yapo PO, Gupta HV, Sorooshian S. 1998. Multi-objective global optimization for hydrologic models. *Journal of Hydrology* **204**(1-4): 83–97.

Zolezzi G, Bellin A, Bruno MC, Maiolini B, Siviglia A. 2009. Assessing hydrological alterations at multiple temporal scales: Adige River, Italy. *Water Resources Research* **45**: 15.