A dynamical climate model–driven hydrologic prediction system for the Fraser River, Canada

Rajesh R. Shrestha, Markus A. Schnorbus, & Alex J. Cannon

2015

Pacific Climate Impacts Consortium (PCIC)

PCIC Publications

Original citation:

Shrestha, R. R., Schnorbus, M. A., & Cannon, A. J. (2015). A dynamical climate model–driven hydrologic prediction system for the Fraser River, Canada. *Journal of Hydrometeorology*, *16*(3), 1273–1292. https://doi.org/10.1175/JHM-D-14-0167.1

A Dynamical Climate Model–Driven Hydrologic Prediction System for the Fraser River, Canada

RAJESH R. SHRESTHA, MARKUS A. SCHNORBUS, AND ALEX J. CANNON

Pacific Climate Impacts Consortium, University of Victoria, Victoria, British Columbia, Canada

(Manuscript received 9 September 2014, in final form 11 January 2015)

ABSTRACT

Recent improvements in forecast skill of the climate system by dynamical climate models could lead to improvements in seasonal streamflow predictions. This study evaluates the hydrologic prediction skill of a dynamical climate model–driven hydrologic prediction system (CM-HPS), based on an ensemble of statistically downscaled outputs from the Canadian Seasonal to Interannual Prediction System (CanSIPS). For comparison, historical and future climate traces–driven ensemble streamflow prediction (ESP) was employed. The Variable Infiltration Capacity model (VIC) hydrologic model setup for the Fraser River basin, British Columbia, Canada, was used as a test bed for the two systems. In both cases, results revealed limited precipitation prediction skill. For streamflow prediction, the ESP approach has very limited or no correlation skill beyond the months influenced by initial hydrologic conditions, while the CM-HPS has moderately better correlation skill, attributable to the enhanced temperature prediction skill that results from CanSIPS's ability to predict El Niño–Southern Oscillation (ENSO) and its teleconnections. The root-mean-square error, bias, and categorical skills for the two methods are mostly similar. Hydrologic modeling uncertainty also affects the prediction skill, and in some cases prediction skill is constrained by hydrologic model skill. Overall, the CM-HPS shows potential for seasonal streamflow prediction, and further enhancements in climate models could potentially to lead to more skillful hydrologic predictions.

1. Introduction

Predictability of seasonal streamflow response is crucial for assessing water availability in river basins and for managing extremes such as floods and drought. A popular method for seasonal streamflow prediction is to use historical climate traces with an appropriate model initialization at the time of forecast, usually referred to as the ensemble streamflow prediction (ESP; Franz et al. 2003; Shi et al. 2008; Shukla and Lettenmaier 2011). An improvement on the ESP approach could be to condition the historical traces based on climate states at the time of forecast, for example, the phase of El Niño–Southern Oscillation (ENSO; Wang et al. 2011; Yuan et al. 2013). Other possibilities of improving the forecast could be to assimilate data for initial hydrologic conditions, for example, soil moisture and snow water equivalent (SWE; e.g., Mahanama et al. 2008; Li et al. 2009; Shukla and Lettenmaier 2011), and postprocessing of the ESP outputs (e.g., Shi et al. 2008; Bohn et al. 2010). However, predictive skill improvement using such data assimilation and postprocessing methods will always be limited by the variability of historical climate data and its suitability to represent the climate state at the time of forecast.

The dynamical climate prediction system based on coupled global climate models (GCMs) could potentially provide an alternative for seasonal streamflow prediction, especially given recent advancements in the representation of large-scale climate variability in these models. For instance, Saha et al. (2006) evaluated the skill of the National Centers for Environmental Prediction (NCEP) forecast system and found useful skill of the model in predicting ENSO-related winter precipitation. Younas and Tang (2013) evaluated the predictability of the Pacific–North American (PNA) pattern using three climate prediction systems and found that the PNA pattern, which originates mainly from the ENSO forcing, is one of the most predictable patterns in climate variability over the Northern

Corresponding author address: Dr. Rajesh R. Shrestha, Pacific Climate Impacts Consortium, University House 1, P.O. Box 1700 STN CSC, University of Victoria, Victoria BC V8W 2Y2, Canada. E-mail: rshresth@uvic.ca

DOI: 10.1175/JHM-D-14-0167.1

Hemisphere. Merryfield et al. (2011, 2013a) evaluated the skill of the Canadian Seasonal to Interannual Prediction System (CanSIPS) and found consistently high anomaly correlations for predictions of the mean Niño-3.4 index (i.e., ENSO prediction skill) for certain initialization months. An intercomparison of the predictive skills of 12 dynamical (GCM based) and 8 statistical models by Barnston et al. (2012) further illustrates higher ENSO prediction skills of the dynamical models.

Such advances in the representation of large-scale climate states could be especially useful for seasonal hydrologic prediction in the western North America region that is known to be influenced by global atmospheric teleconnection patterns. Particularly, previous studies in the region indicate that ENSO teleconnections are a key driver of seasonal and annual hydroclimatic variability, such as snowpack evolution (Clark et al. 2001; McCabe and Dettinger 2002; Seager et al. 2010) and streamflow response (Cayan et al. 1999; Hidalgo and Dracup 2003; Fleming et al. 2007; Whitfield et al. 2010). However, predictive skill of the climate model–driven hydrologic prediction system (CM-HPS)—which typically consists of a dynamical climate model, a statistical and/or dynamical downscaling scheme, and a hydrologic model—remains an open question. For instance, Wood et al. (2002, 2005) employed the NCEP Global Spectral Model (GSM), a bias-corrected spatial disaggregation (BCSD) downscaling scheme, and the Variable Infiltration Capacity model (VIC) hydrologic model for seasonal streamflow prediction over the Pacific Northwest basins. They found negligible skill improvement for the regionally averaged hydroclimatic hindcasts compared to ESP. Yuan et al. (2013) used the NCEP Climate Forecast System, version 2 (CFSv2), a Bayesian downscaling scheme, and VIC over the conterminous United States and found limited added skill of the CM-HPS compared to climatology beyond the first month. Another study (Bastola et al. 2013) also found limited skill of the CM-HPS when they used Florida State University seasonal hindcasts, downscaled with a quantile-based approach, driving multiple hydrologic models over the southeastern United States. Nevertheless, recent studies have found improvements in the skill of the GCMs, such as the simulation of Pacific climate modes and their teleconnections for phase 5 of the Coupled Model Intercomparison Project (CMIP5) generation of GCMs compared to the previous phase (CMIP3; Polade et al. 2013; Weare 2013). Such improvement should also improve the predictive skill of the CM-HPS.

The results obtained from a CM-HPS also need to be considered in the context of associated uncertainties in the modeling chain. For instance, the impact of downscaling can be considerable, especially given that archived hindcasts from the climate models are usually available for a limited number of years (~20 years), and the choice of the calibration period and its length affects the statistical characteristics of the downscaled scenarios (Winkler et al. 1997; Wilby et al. 2002). Additionally, hydrologic models are affected by different sources of uncertainties, such as input and output data, model structure, and parameter uncertainties (Beven 2006). These uncertainties affect the ability of a hydrologic model in replicating the observed variability of streamflow (Shrestha et al. 2014), which could also affect streamflow prediction skill.

This study explores the possibility of using climate outputs from the Canadian Climate Centre for Modelling and Analysis (CCCma) Third and Fourth Generation Canadian Coupled Global Climate Models (CanCM3 and CanCM4; Merryfield et al. 2011, 2013a) as inputs to a seasonal hydrologic prediction system for lead times up to 11 months. The primary objective of this paper is to provide insights into the hydrologic prediction skills of a CM-HPS in comparison to a traditional ESP in western Canada. The secondary objective is to assess the predictive skills in the context of ENSO teleconnections. A test bed CM-HPS was set up by using the CanCM3- and CanCM4-driven VIC hydrologic model (Liang et al. 1994, 1996, 2003) for the Fraser River basin (FRB), Canada. A statistical downscaling method, the BCSD (Wood et al. 2002, 2004), was employed to downscale the precipitation, wind speed, and air temperature data from CanCM3 and CanCM4 (10-member ensembles for each). For comparison, an ESP setup with inputs derived from the historical and future traces of daily climate observations was employed.

2. Study basin

The FRB study area covers a drainage area of approximately 230 000 km^2, with relief varying from sea level to about 4000 m (Fig. 1). The river has its headwaters in the Rocky Mountains near Jasper, Alberta, and its major tributaries include the Stuart River, Nechako River, Quesnel River, Chilcotin River, Thompson River, and Harrison River. The Fraser River basin is geographically diverse, and its main stem and tributaries flow through 12 ecoregions and 9 biogeoclimatic zones before discharging into the Strait of Georgia (Schnorbus et al. 2010).

The hydroclimatology of the FRB reflects the interaction of the dominant westerly atmospheric circulation with mountain ranges (Moore 1991). As summarized in Table 1, the subbasins of the FRB exhibit highly variable precipitation and runoff responses, with generally higher precipitation at higher elevation ranges. Mean annual temperature and precipitation vary between −5° and 10°C and 200 and 5000 mm, respectively. Although the hydrologic response of the entire FRB is predominantly nival (snow dominated), responses of the areas within the FRB

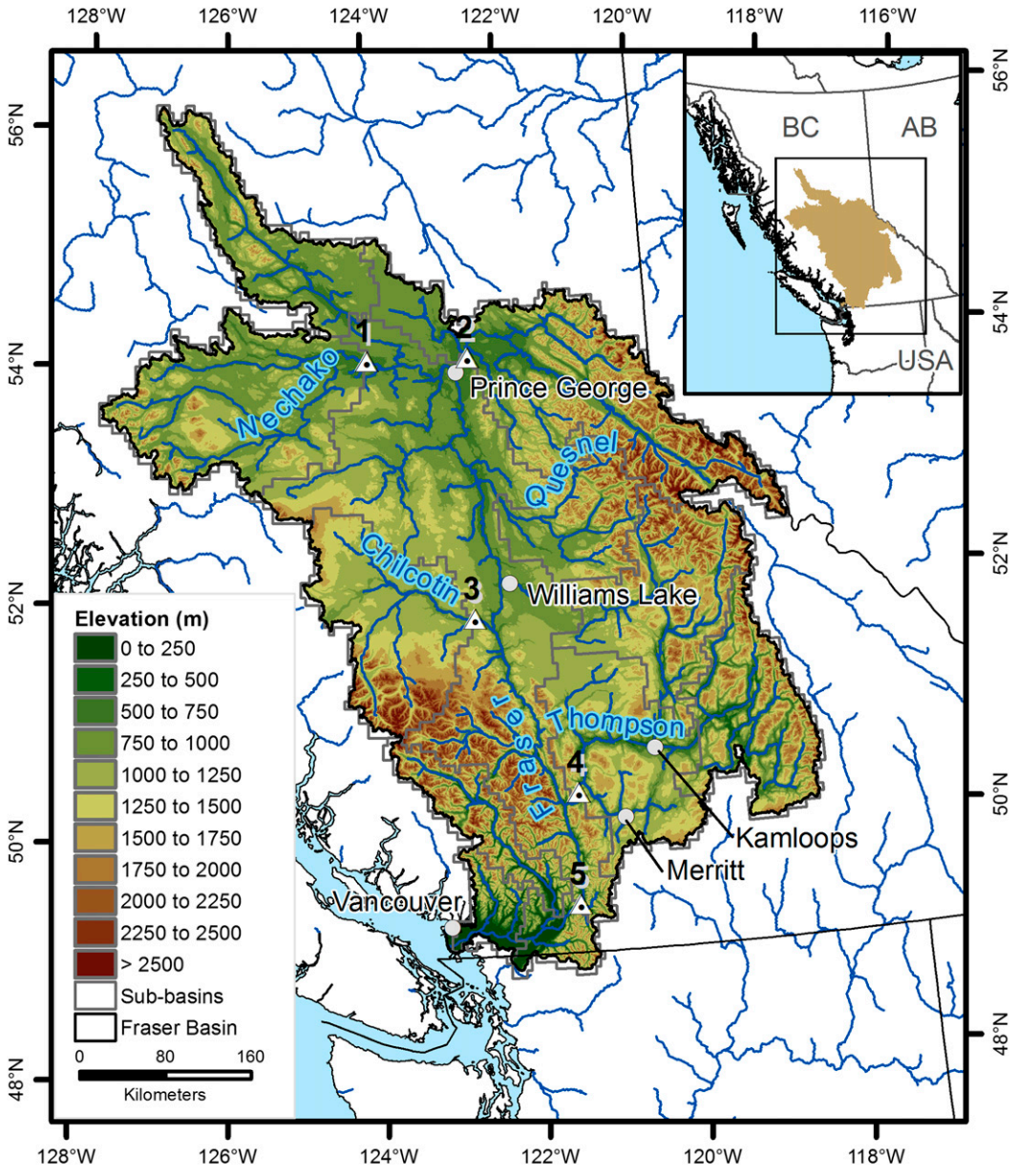

FIG. 1. Location map and elevation range of the FRB. Numbers 1–5 are the outlets of the subbasins (Table 1) considered in this study.

vary from nival to hybrid (rain and snow) to pluvial (rain dominated; Shrestha et al. 2012). The mean annual discharge at the Fraser–Hope hydrometric station is about 2850 $m^3 s^{-1}$ (1961–90), with the maximum monthly mean discharge in June (7600 $m^3 s^{-1}$) and minimum monthly mean discharge in February (1000 $m^3 s^{-1}$). About 65% of the total flow occurs during the snowmelt season from May to August.

Lakes and tributaries within the FRB also provide spawning habitat for all five species of eastern Pacific salmon (Schnorbus et al. 2010), and the basin is a particularly important spawning ground for sockeye and chinook salmon, accounting for majority of the Canadian stock (Morrison et al. 2002).The FRB is also home to 63% of British Columbia's population, including Vancouver, Canada's third-largest city. Approximately 170 km from the mouth, the river emerges onto the lower Fraser valley region, where it flows through braided channels now partially constrained by flood defenses (Rice et al. 2009). The risk of flooding is greatest in the lower Fraser because of a large population and significant residential, commercial, industrial, utilities, and transportation infrastructure

TABLE 1. Water survey of Canada (WSC) hydrometric stations and corresponding study subbasin characteristics. The subbasin mean annual precipitation and runoff are for 1961–90, with runoff obtained by normalizing the hydrometric station discharges by subbasin areas. Runoff for Nechako and Fraser–Hope subbasins are based on estimated naturalized discharges.

Subbasin No.	Station name	Subbasin name	WSC ID	Subbasin area (km^2)	Elev range (m)			Precipitation ($mm\ yr^{-1}$)	Runoff ($mm\ yr^{-1}$)
					Min	Mean	Max		
1	Nechako River at Vanderhoof	Nechako	08JC001	25 100	383	1060	2740	716	300
2	Fraser River at Shelley	Fraser–Shelly	08KB001	32 400	569	1308	3928	1163	804
3	Chilcotin River below Big Creek	Chilcotin	08MB005	19 300	531	1268	2891	490	158
4	Thompson River near Spences Bridge	Thompson–Spences	08LF051	54 900	196	1747	3193	844	441
5	Fraser River at Hope	Fraser–Hope	08MF005	217 000	29	1330	3928	805	422

in the floodplain (Fraser Basin Council 2010). Two reservoir systems, Nechako (890 MW) and Bridge–Seton (492 MW), regulate the flows of the tributaries of the Fraser River for hydroelectric power generation. Additionally, the Nechako reservoir diverts about 40% of total naturalized discharge out of the Nechako River system (based on 1961–90 naturalized discharge at Nechako–Vanderhoof); however, the effect of flow diversion gradually diminishes downstream and constitutes about 3.5% of the total Fraser River discharge at the Fraser–Hope station. This study does not consider regulated flow from the reservoir systems, but instead uses naturalized flow (British Columbia Ministry of Environment 2008, unpublished data) for hydrologic model calibration and evaluation of the streamflow prediction skill. The main stem of the Fraser River is not regulated, and the capacity to store flood water in the system is limited. Hence, the ability to predict streamflow is crucial for water resources management in the region.

3. Data and methods

a. Dynamical climate models

This study used climate hindcasts from CanSIPS (Merryfield et al. 2013a), which is Environment Canada's operational two-model seasonal forecasting system, combining ensemble predictions from CanCM3 and CanCM4. The two models are part of the North American Multi-Model Ensemble (NMME) prediction experiment focused on seasonal-to-interannual time scales (Kirtman et al. 2014). The models were employed by previous studies for forecasting Arctic sea ice area (Sigmond et al. 2013) and global drought onset (Yuan and Wood 2013). These two models share common ocean, land, and sea ice components and run at T63 horizontal resolution (1.875° at the equator) for the atmosphere. The atmospheric components of the two models differ. CanCM3 has 31 vertical levels (Scinocca et al. 2008), while CanCM4 has 35 vertical levels (Arora et al. 2011). The CanSIPS seasonal hindcasts consist of 20 ensemble members, 10 each for CanCM3 and CanCM4 that are initialized at the beginning of each month from January 1979 to December 2009 and extend 1 year into the future. Forecast initial conditions were provided by assimilation runs, with atmospheric variables, sea surface temperatures, and sea ice concentrations constrained near observational values and sea ice thickness initialized by relaxing to the model climatology (Merryfield et al. 2013a). A detailed description of CanSIPS is available in Merryfield et al. (2013a,b) and Sigmond et al. (2013).

b. Downscaling of dynamical climate models

In this CM-HPS approach, the BCSD statistical downscaling method (Wood et al. 2002, 2004) was used to downscale the coarse-resolution CanCM3 and CanCM4 outputs to the resolution of the hydrologic model (1/16°) (Fig. 2). Previous studies (e.g., Maurer and Hidalgo 2008; Bürger et al. 2012) have demonstrated the effectiveness of BCSD in capturing many aspects of historical daily variability and found the method to be competitive with many other statistical downscaling methods.

BCSD performs downscaling by developing relationships between large-scale monthly average climate variables—in this case, precipitation, wind speed, minimum temperature, and maximum temperature—and the corresponding local-scale daily climate variables. In this study, BCSD was implemented, in a slightly modified form, following Maurer and Hidalgo (2008). BCSD consists of the following steps: (i) bias correction of monthly climate model outputs using quantile mapping, (ii) spatial disaggregation of the climate model outputs to the finescale observational grid, and (iii) temporal disaggregation to a daily time step. In the first step, bias correction using quantile mapping adjusts the monthly climate model values at each grid point so that they match the distribution of spatially aggregated gridded observations. Each mapping is defined over the ranges of the historical dataset. If a GCM forecast value falls

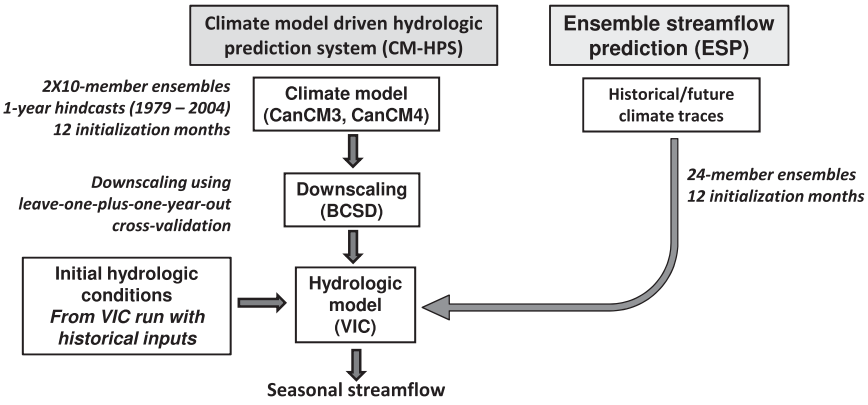

FIG. 2. Illustration of the methods used for CM-HPS and ESP.

outside the range of the historical simulated values, then extrapolation by the "delta method" of Hamlet et al. (2010) was used. For precipitation, this involves applying multiplicative changes relative to maximum values of the historical GCM-simulated values directly to the maximum observed values. For example, if a GCM forecast value is 10% higher than the maximum simulated value over the historical period, then the bias-corrected forecast value is set to be 10% higher than the observed historical maximum value. For temperature, changes are instead applied additively with respect to the observed extremes. In the second step, the bias-corrected monthly values of precipitation, wind speed, and temperature are disaggregated spatially using high-resolution scaling adjustments for precipitation and wind speed and offset adjustments for temperature. As suggested by Bürger et al. (2012), the algorithm used in this study incorporates monthly minimum and maximum temperature directly as predictors to improve the resolution of the diurnal temperature range. Finally, in the third step, the spatially disaggregated monthly values are temporally disaggregated by selecting a random month from the observed historical record and then adjusting its daily values, multiplicatively for precipitation and additively for temperature, to reproduce the monthly value. Following Maurer and Hidalgo (2008), if a multiplicative scaling factor for precipitation is greater than three, then the selected month must have more than two wet days to avoid simulating unrealistically large daily precipitation amounts. A detailed description and evaluation of BCSD over British Columbia is given by Werner (2011).

To apply BCSD, predictand variables were taken from the gridded observation dataset of daily maximum and minimum air temperature, daily total precipitation, and daily wind speed used to calibrate the VIC hydrologic model. The wind datasets based on interpolated estimates of 10-m wind speed from the National Centers for Environmental Prediction–National Center for Atmospheric Research (NCEP–NCAR) reanalysis (Kalnay et al. 1996). For temperature and precipitation, observation data were generated on a 1/16° grid for the entire area of British Columbia, primarily from the Environment Canada climate station observation network (Schnorbus et al. 2010). These gridded observation data are available from 1950 to 2004, while CanCM3 and CanCM4 data are available from 1979 to 2010. This leaves 26 years (1979–2004) of overlapping data for applying BCSD and evaluating the performance of the hydrologic prediction systems. A leave-one-plus-one-year-out cross-validation strategy was employed in this study. The strategy uses the entire 26-yr dataset, except for the hindcast year, and the following year after model initialization as the reference BCSD period. For instance, for the prediction year from May 1985 to April 1986, a 24-yr of dataset consisting of historical years from May 1979 to April 1985 and "future" years from May 1987 to April 2004 was used in BCSD to develop the quantile mapping relationships and for sampling of months in the temporal disaggregation step. The additional year following the hindcast year was excluded to prevent interannual serial correlation of climate variables from affecting validation results. Although the use of future years is not strictly hindcasting, all these years of data will be available if the model is to be run in an operational forecast mode for the present time; therefore, it was considered appropriate to use all available data. Leave-one-out cross-validation has also been used to evaluate seasonal streamflow forecasts in the neighboring Columbia River basin (Gobena et al. 2013).

BCSD was run separately for each combination of initialization month (January–December), climate model (CanCM3 and CanCM4), ensemble member (1–10), and variable (precipitation, maximum and minimum temperature, and wind speed). The resulting downscaled outputs were used to drive the VIC hydrologic model for each year and initialization month (Fig. 2).

c. Hydrologic model and initialization

The spatially distributed macroscale VIC hydrologic model, version 4.0.7 (Liang et al. 1994, 1996, 2003), was employed in a water balance mode for the FRB. The model was previously applied to the FRB to investigate the impacts of mountain pine beetle infestation and climate change on streamflow generation (Schnorbus et al. 2010; Shrestha et al. 2012). VIC computes water fluxes for a range of hydrologic processes such as evapotranspiration, snow accumulation, snowmelt, infiltration, soil moisture, and surface and subsurface runoff. The model represents soil moisture processes in three soil layers and represents subgrid variability of land surface vegetation classes and topography by partitioning each grid cell into a number of mosaic-type land cover tiles and elevation (snow) bands, respectively. VIC explicitly accounts for snow accumulation and ablation by using a two-layer snowpack individually for each land cover and elevation tile. The model uses variable infiltration curves for runoff generation and the Arno conceptual model (Todini 1996) for subsurface flow generation. Surface runoff from the upper two soil layers is generated when the soil moisture content exceeds the storage capacity of the soil. Simulated fluxes as well as state variables for each grid cell are calculated as area averages of subgrid values. An offline routing model (Lohmann et al. 1998) is used for generating the streamflow. The routing model collects surface runoff and base flow from the VIC and routes within grid cells using the unit hydrograph approach and to the outlet using the linearized Saint Venant equation in a flow network.

VIC for the FRB was set up at a $1/16^\circ$ spatial resolution and daily temporal resolution. Geospatial data for the VIC were derived from (i) a 90-m resolution digital elevation model from the NASA Shuttle Radar Topography Mission (Jarvis et al. 2008), (ii) a 25-m resolution land cover dataset from the Earth Observation for Sustainable Development of Forests (Wulder et al. 2003), and (iii) a $1/12^\circ$ resolution soil classification and parameterization based on the Global Soil Data Products (Global Soil Data Task 2000). The geospatial data were processed and gridded to match the resolution of VIC ($1/16^\circ$). The model was driven by gridded observation data consisting of daily precipitation, maximum and minimum air temperature, and wind speed for calibration–validation. The FRB was subdivided into 66 subbasins based on Water Survey of Canada hydrometric station locations. Daily subbasin streamflow was calibrated by using the multiobjective complex evolution (MOCOM) technique (Yapo et al. 1998) with the goodness-of-fit measures: Nash–Sutcliffe coefficient of efficiency (NSE), NSE of log-transformed discharge (LNSE), and water balance error (WBE) as objective functions. For model calibration, 6 years (1985–90) of data were used. For model validation, 5 years (1991–95) were used. Five runoff generation parameters and an adjustment factor for precipitation were used for calibration. Additionally, fresh snowpack albedo and its decay rate during snowmelt were adjusted to replicate the snowmelt response in the British Columbia river basins (Schnorbus et al. 2010, 2011). A detailed description of VIC set up, calibration, and evaluation of model performance for the FRB is available in Schnorbus et al. (2010) and Shrestha et al. (2012).

Initial hydrologic conditions (i.e., soil moisture and snowpack storage and its thermal condition) required for running the VIC in hindcast (1979–2004) mode were captured for the first day of each month from an observation-based VIC run for the basin (Shrestha et al. 2012). The captured initial conditions were used to "warm start" simulations in the hindcast mode with the same initial state applied to each ensemble member. Before running the routing model, the hindcast runs were concatenated with the observation-driven VIC results from previous years. This prevents "dry" grid cells at the initial hindcast time step of the routing model run.

d. Ensemble streamflow prediction

To compare CM-HPS with a traditional seasonal streamflow prediction system, ESP with meteorological inputs derived from the historical and future climate traces of precipitation, maximum and minimum temperature, and wind speed was used. In the ESP approach, VIC for each hindcast year and initialization month was driven by an ensemble of historical and future climate traces, excluding the hindcast year and the following year. For instance, for the hindcast year from May 1985 to April 1986, 24 1-yr traces including 6 historical years (from May 1979 to April 1985) and 18 future years (from May 1987 to April 2004) were used as inputs. From all ESP runs, only the runs corresponding to the ENSO state at the time of model initializations were selected for a separate analysis and hereafter referred to as ESP-ENSO. In this case, each initialization month was assigned an ENSO state according to the Oceanic Niño Index data from the National Weather Service's Climate Prediction Center (see www.cpc.ncep.noaa.gov/products/analysis_monitoring/ensostuff/ensoyears.shtml for data), and the historical and future ENSO runs matching the ENSO state at the initialization month were selected. However, the ENSO state at the time of model initialization may not persist through the entire year for which the ESP-ENSO results are considered. Additionally, given that the hydrologic response of the FRB is predominantly snowmelt driven, and runoff response to winter precipitation only occurs in spring–summer, teleconnections to the ENSO states may not be always be reflected in monthly streamflows. Given these

issues, the ESP-ENSO results may not always be representative of the actual evolution of ENSO states for the entire evaluation year. In this approach too, the ESP runs for the hindcast year and the following year (regardless of the ENSO state) were excluded.

e. Methods of evaluation

The results obtained from the two approaches—CM-HPS and ESP based (ESP and ESP-ENSO)—were compared with observations (observation-driven VIC results in the case of SWE) and their predictive skills were analyzed using four deterministic statistical metrics: (i) Pearson correlation coefficient, (ii) root-mean-square error (RMSE) normalized by mean observed flow (RMSE/observed flow), (iii) volume bias, and (iv) percent correct categorical forecast skill (Table 2). The first three metrics use the ensemble mean result while the fourth metric uses the entire ensemble for the evaluation of skills.

The correlation coefficient measures the strength of linear association between observations and simulated results. It is insensitive to the model's mean bias and is scale independent, and it gives the proportion of the total variability of observations that can be explained, in a linear sense, by the forecasts (Kharin et al. 2001). The higher the correlation coefficient is, the better the model performance skill is, and for the 26-yr data used in this study, a correlation value larger than 0.387 would be statistically significant at a 5% significance level (using Fisher's z transformation; see Wilks 2006). The RMSE skill provides a measure of the average error of the prediction model, while the volume bias provides a measure of the model's ability to simulate the water balance. The lower the RMSE skill or volume bias is, the better the model performance is. The percent correct categorical forecast skill evaluates the skill of the ensembles relative to the tercile boundaries of the observations, that is, values below the 33rd percentile are below normal, values between the 33rd and 67th percentiles are near normal, and values above the 67th percentile are above normal. The percentage correct score, which gives a combined proportion of ensemble members falling in the three observed categories (Table 2), varies between 0% and 100%, and the higher the percent correct value, the better the model performance. Since it is based on the discrete tercile boundaries, a 33% skill could be achieved by chance alone and a percent correct skill >33% implies better model performance than chance.

TABLE 2. Skill metrics formulae, where n is the number of years; X_i and Y_i are the observed and simulated values; $\overline{X}$ and $\overline{Y}$ are the mean of the observed and simulated values; and X_v and Y_v are the observed and simulated volumes (seasonal or annual), respectively. The variable p_k is the number of correct forecasts at each of the kth category (m = three categories in this case: <33rd percentile = below normal; >33rd and < 67th percentiles = near normal; and >67th percentile = above normal), and p_t is the total number of ensemble members.

Skill metric	Formula
Pearson correlation coefficient	$\dfrac{\sum_{i=1}^{n}(X_i-\overline{X})(Y_i-\overline{Y})}{\sqrt{\sum_{i=1}^{n}(X_i-\overline{X})^2}\sqrt{\sum_{i=1}^{n}(Y_i-\overline{Y})^2}}$
RMSE/observed flow	$\dfrac{\sqrt{\dfrac{\sum_{i=1}^{n}(X_i-Y_i)^2}{n}}}{\overline{X}}$
Volume bias	$\sum_{i=1}^{n}\dfrac{\lvert X_{v,i}-Y_{v,i}\rvert}{X_{v,i}}$
Percent correct categorical forecast skill	$\dfrac{1}{n}\sum_{i=1}^{n}\dfrac{\sum_{k=1}^{m}p_k}{p_t}$

4. Results and discussion

a. VIC calibration and validation results

Figure 3 depicts the VIC validation results obtained for five subbasins of the FRB. The Fraser–Hope subbasin covers about 94% of the total drainage area, while Nechako, Chilcotin, and Thompson–Spences are located at three major tributaries, and Fraser–Shelly is located on the main stem covering the headwater drainage area. Hence, these subbasins could be considered as representative for the entire basin (Fig. 1). The results provide a general indication of the hydrologic model's ability to represent the dynamics of the observed 1991–95 hydrograph. The results reveal generally good agreement between the observed and simulated discharges, particularly for the annual major hydrologic event driven by the spring snowmelt. The goodness-of-fit measures obtained from the calibration and validation results (Table 3) 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 modeled results can be seen from visual inspection of the plotted results (Fig. 3). In particular, there are considerable discrepancies in the magnitude of the peaks for some of the subbasins. Uncertainties in the input data, model structure, parameters, and observed streamflow data affect the ability of a hydrologic model to replicate different elements of the hydrograph (Shrestha et al. 2014). In this particular case, the use of naturalized flow for the Nechako subbasin (and to a small extent for Fraser–Hope) affects the calibration and validation results. Given that the hydrologic modeling uncertainty can be a limiting factor in the skill of the prediction models, the gridded observation-driven VIC results are

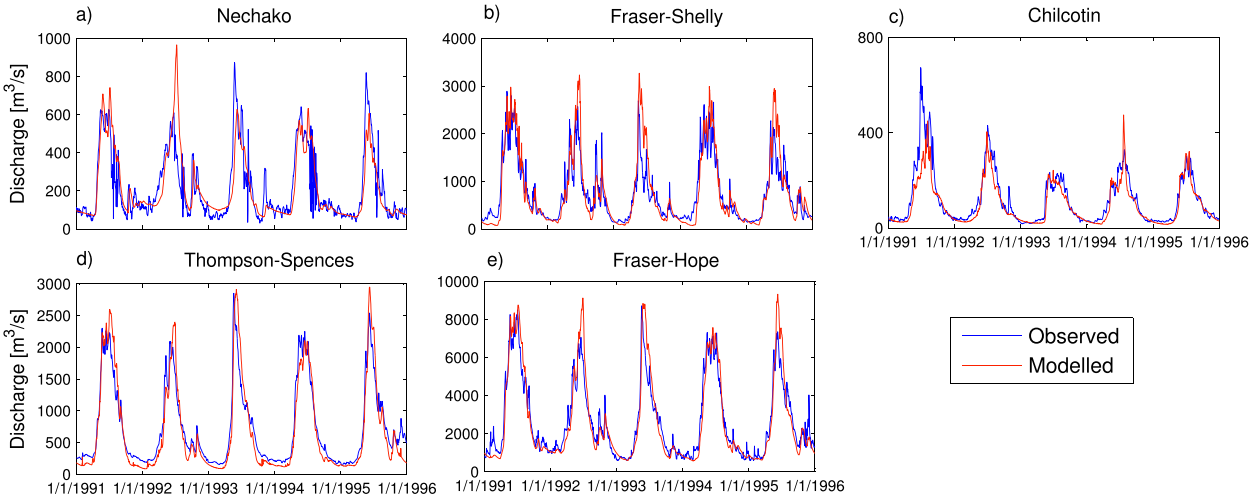

FIG. 3. Observed and VIC-simulated streamflow ($m^3 s^{-1}$) for validation period.

presented together with the prediction models in the subsequent sections. With regard to the skill metrics calculated by comparing with observations, the skill of the observation-driven VIC can generally be considered as the upper limit for correlation and the lower limit for RMSE and volume bias. While it is technically possible for the prediction models to exceed these limits in skill, occurrences are unlikely to be attributable to anything but chance. In the case of percent correct categorical skill, since only a single member is available for the observation-driven VIC simulation, its categorical skill was not evaluated and was not compared with the prediction results.

b. *Precipitation and temperature prediction skill*

Figures 4a and 4b depict correlations of monthly precipitation and temperature, respectively, obtained from the BCSD downscaled combined CanCM3 and CanCM4 (referred to as CanCM3/CanCM4) 20-member ensemble means for the entire FRB (1979–2004) compared to the gridded observation data. It is to be noted that downscaling does not improve the correlation skill, and the plots of the raw CanCM3/CanCM4 ensemble means (without downscaling) are very similar to the downscaled results (not shown). The ENSO-conditioned mean precipitation and temperature derived from historical and future climate traces are not expected to have correlation skills and are presented as the reference forecasts (Figs. 4c,d) for comparing with the CanCM3/CanCM4 results. The ENSO-conditioned output includes the ensemble means of the historical and future climate data corresponding to the ENSO state at given months/years, excluding the hindcast year and the following year (regardless of the ENSO state). Given that the prediction for the first month is derived from day 1 to end of month prediction, it is considered as a zero lead.

The results revealed negative or close to zero correlations for most monthly CanCM3/CanCM4 precipitation predictions and for all initialization months beyond the zero lead month. Such difficulty in replicating monthly precipitation fields is not unique to the FRB, as Merryfield et al. (2011) noted limited precipitation prediction skill of CanCM3/CanCM4 for all of Canada. Nevertheless, compared to the reference ENSO-conditioned mean precipitation, correlations of the CanCM3/CanCM4 are slightly higher for most months.

Similarly, compared to the reference ENSO-conditioned mean temperature, correlations of the BCSD-downscaled CanCM3/CanCM4 temperatures are higher and positive (Figs. 4b,d). An important factor for the predictability of temperature in the region is its teleconnection with ENSO, with the ENSO influence on western Canadian temperatures tending to be strongest in winter and persisting into early spring (Shabbar 2006). Evidence of this is seen in the banded structure in Fig. 4b. Irrespective of the initialization month, correlations for CanCM3/CanCM4 predictions in the late winter and spring are mostly positive, whereas correlations for the fall and early winter are mostly near zero or negative. Similar correlation patterns are also present in the ENSO-conditioned results, but the correlation values are lower than CanCM3/CanCM4. Additionally, correlation values for some of the CanCM3/CanCM4 late winter and spring predictions are higher than the statistical significance level (0.387, at a 5% significance level). This is consistent with the findings of previous studies using the same climate models by Merryfield et al. (2011, 2013a). They found higher temperature prediction skill for most of North America, with much of the seasonal predictability, particularly in winter and early spring, attributable to the teleconnected influence of ENSO. Such improved

TABLE 3. NSE, LNSE, and WBE for the VIC calibration (1985–90) and validation (1991–95) results. Values in the parentheses are validation results.

Subbasin No.	Subbasin name	Statistical performance calibration (validation)		
		NSE	LNSE	WBE
1	Nechako	0.72 (0.66)	0.71 (0.73)	−0.08 (0.00)
2	Fraser–Shelly	0.85 (0.75)	0.73 (0.70)	−0.05 (0.05)
3	Chilcotin	0.86 (0.78)	0.89 (0.85)	−0.06 (−0.17)
4	Thompson–Spences	0.92 (0.89)	0.84 (0.76)	−0.03 (−0.07)
5	Fraser–Hope	0.91 (0.88)	0.87 (0.89)	−0.02 (−0.01)

representation of the winter–spring temperature due to ENSO teleconnections could potentially provide improved representation of the snow and streamflow responses for the FRB.

c. *SWE prediction skill*

Given that the hydrologic regime of the FRB is predominantly nival, SWE prediction skill provides a useful indicator of hydrologic prediction skill. Figure 5 depicts the correlation and volume bias of the CM-HPS and ESP-based (ESP and ESP-ENSO) mean April SWE with respect to the observation-driven VIC results. Since all SWE values are VIC-simulated results, the effect of hydrologic modeling uncertainties are not reflected in the correlation and bias calculations, and as a result, the skill values are likely inflated. Additionally, part of the SWE prediction skill for December–April initializations arises from initial hydrologic conditions, mainly in the form of snow storage. Nevertheless, slightly higher correlations of the CM-HPS compared to the ESP-based results for the December and January initializations, which have limited snow storage memory, indicate the effect of better temperature prediction skill (Fig. 4b). Furthermore, given the positive correlations of precipitation for the zero lead months (Fig. 4a), additional experimental runs using a combination of the ESP precipitation and CM-HPS temperature was carried out for the January initialization (not shown). The results revealed a slight deterioration of the SWE correlation skill for the combined inputs (0.55) compared to the CM-HPS (0.62). This suggested that the precipitation prediction skill also contributed to the SWE prediction skill. Additionally, for the May–November initializations, where lead times decrease from 11 to 5 months and the effect of model initial conditions is small, CM-HPS exhibits positive correlations, some of which are statistically significant at a 5% significance level; a correlation value of 0.4 means that the predictions can explain 16% of variance in the observation-driven VIC results. On the other hand, correlations of most ESP-based results are close to zero (no skill). Such improved SWE predictability by the CM-HPS can be attributed to the better winter–spring temperature predictability by CanCM3/CanCM4. A comparison of the ESP and ESP-ENSO SWE correlations revealed no improvement in the prediction skills when the ESPs are conditioned with the ENSO states. As stated earlier, the historical and future ENSO evolution corresponding to an initialization month may not always match the actual evolution of the ENSO state and hence may not be reflected in the April snow storage results. This highlights the difficulty in conditioning inputs according to the ENSO states at the time of model initialization.

A comparison of the volume biases from the two approaches mostly revealed similar errors, with low biases for the January–April initializations when compared to the rest of the year, again illustrating the effect of the hydrologic model's initial conditions. In this case also, the bias values are small, likely due to the fact that the prediction results are compared with the observation-driven VIC-simulated results. Nevertheless, compared to the ESP and ESP-ENSO results, the CM-HPS results depict slightly lower errors for most months (Fig. 5b), possibly because of the enhanced temperature prediction skill.

d. *Streamflow prediction skill*

Figure 6 compares correlations of the CM-HPS and ESP-based (ESP and ESP-ENSO) streamflow predictions with observed streamflow at the Fraser–Hope station. Additionally, correlations of the observation-driven VIC results are shown. In this case, prediction skills for the initial months mostly derive from the hydrologic model's initial conditions, such as soil moisture and snow storage. In particular, for the January–June initializations, the memory of snow storage contributes to the streamflow prediction skill, leading to statistically significant (at a 5% significance level) correlations for the first few months. Overall, while the correlations of the observation-driven VIC are high for all months (0.8–0.9), the correlations of the ESP-based and CM-HPS predictions tend to drop quickly after the initial months. This implies that the hydrologic model skill is not a limiting factor for the correlation skill of prediction models. Additionally, as described earlier, while the precipitation skill for the zero lead month appear to have contributed

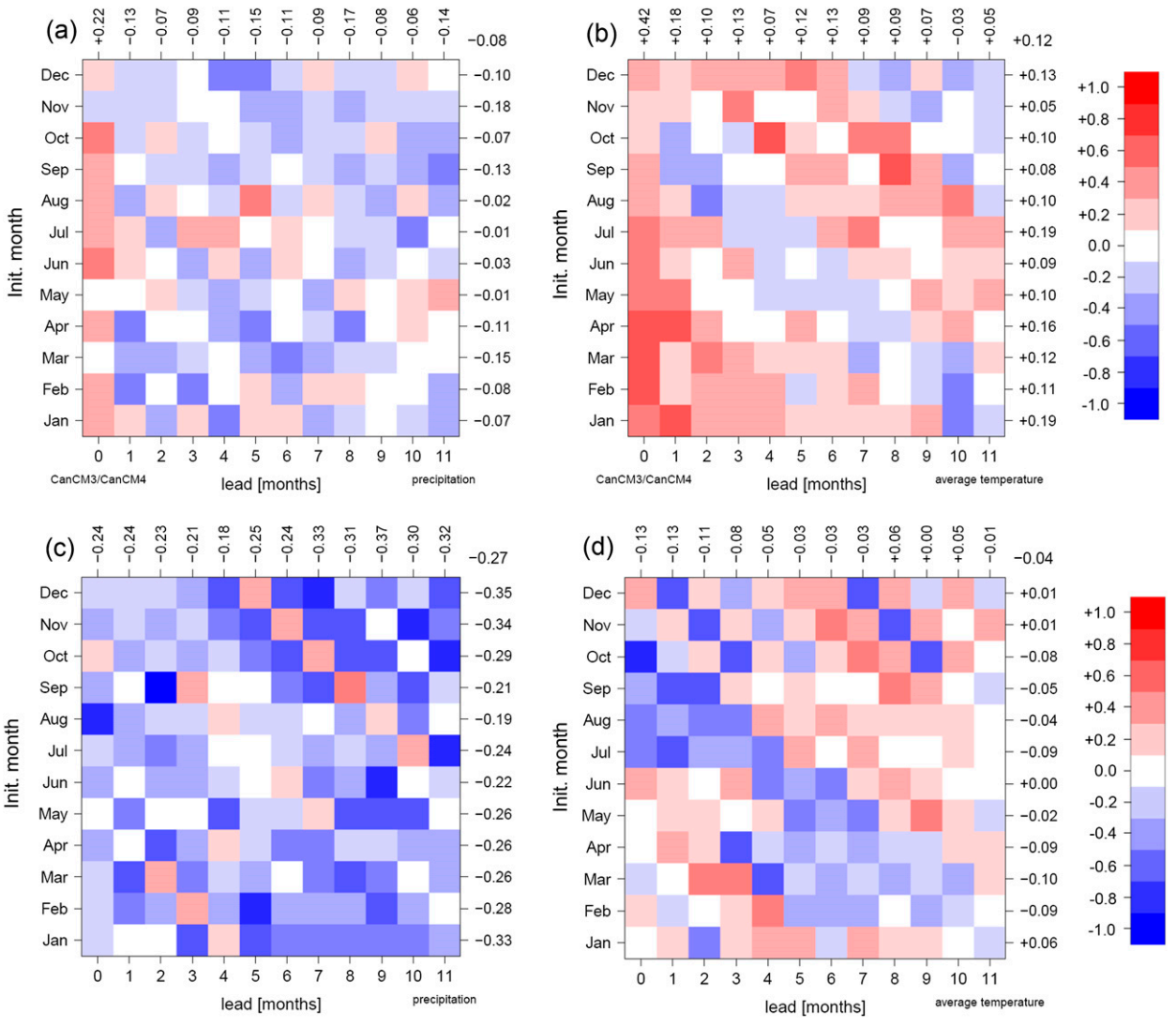

FIG. 4. Correlation skills of CanCM3/CanCM4 ensemble mean (BCSD downscaled) for (a) precipitation and (b) temperature and of ENSO-conditioned mean for (c) precipitation and (d) temperature with gridded observed data. Each row shows the correlation values for 0–11 lead months for a given initialization month, the values at right (above) are column (row) means, and upper-right corner is the mean of all columns and rows.

to the higher SWE prediction skill for January, such skill improvement was not found for the streamflow prediction (not shown), probably because precipitation from only a part of the basin causes streamflow in January. Comparing the correlations beyond the months affected by the initial conditions, the CM-HPS results exhibit positive correlations for some of the months, and small improvements with respect to the ESP-based results, particularly for the June–September initializations. Specifically, CM-HPS winter–spring correlations are noticeably superior, which are related to the better temperature (Fig. 4) and SWE (Fig. 5) predictive skills. Given that the winter–spring temperature prediction skill arises from the ENSO prediction skill of CanCM3/CanCM4, better streamflow prediction skill (albeit weak) can be linked to the predictability of ENSO and ENSO teleconnections. Comparing the ESP and ESP-ENSO results revealed some skill improvement for ESP-ENSO over ESP. Specifically, while the correlations of ESP tend to drop below zero (e.g., for June–October initializations) after few initial months, the ESP-ENSO skill for some of the months showed positive skills (e.g., September–November initializations). As described earlier, the difficulty in conditioning the ENSO states for the entire prediction period likely restricted a substantial improvement in the skill. It is also interesting to note that ESP results showed strong negative skill (<-0.387, significance level; Fig. 6). This is due to the strong negative correlations of the ensemble

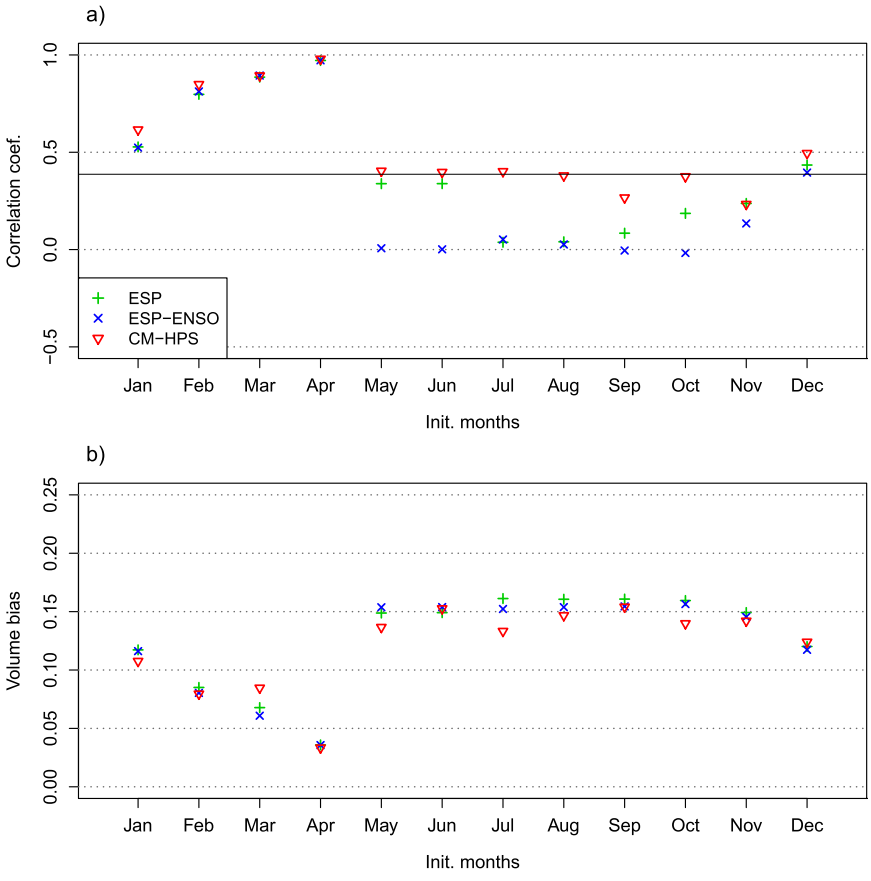

FIG. 5. Correlation coefficient and volume bias of mean April SWE for given initialization months. Ensemble means of the VIC-simulated SWE values from the CM-HPS, ESP, and ESP-ENSO are compared with the gridded observation-driven VIC results. The horizontal line at 0.387 correlation is the level of statistical significance (at a 5% significance level).

mean of the ESP input data (precipitation or temperature) with the observed data. This issue becomes apparent in the streamflow correlation as the effect of the initial conditions diminishes; hence, the strong negative correlation does not represent predictive skill of the model.

The RMSE/observed flow skill of the three prediction models are, in most cases, inferior to the observation-driven VIC (Fig. 7). For the February–July initializations, RMSE skills for up to two lead months are close to the observation-driven VIC results, showing the influence of the initial conditions. Since it is a relative measure, the error values are amplified for low-flow months (e.g., March and April). The CM-HPS RMSE skills are mostly similar to the ESP and ESP-ENSO skills, while the ESP and ESP-ENSO results mostly overlap for all initialization months and lead times, again illustrating no particular advantage in conditioning inputs according to the ENSO state during initialization months.

Figure 8 compares the seasonal [December–February (DJF), March–May (MAM), June–August (JJA), and September–November (SON)] and annual volume biases for the three prediction systems and five subbasins of the FRB. The results revealed considerable biases from the observation-driven VIC results, particularly for the seasonal volumes. As outlined previously, uncertainties in the input, parameter, and model structure affect the modeled streamflow and the model's ability to replicate observed streamflow. Furthermore, these findings emphasize results of the previous study (Shrestha et al. 2014) that good representation of goodness-of-fit metrics does not necessarily mean a good representation of the monthly and seasonal streamflows. Such hydrologic model biases also affect the seasonal and annual flow volumes obtained from the CM-HPS and ESP-based results. Specifically, for some of the subbasins (i.e., Chilcotin and Thompson–Spences), the biases closely follow the observation-driven VIC biases, indicating limitations imposed by the hydrologic modeling uncertainties. The winter streamflow volumes for most subbasins (especially for CM-HPS Nechako and Fraser–Shelly subbasins) exhibit larger bias, which is partly due to the fact that it is a relative measure, and the error values are amplified for

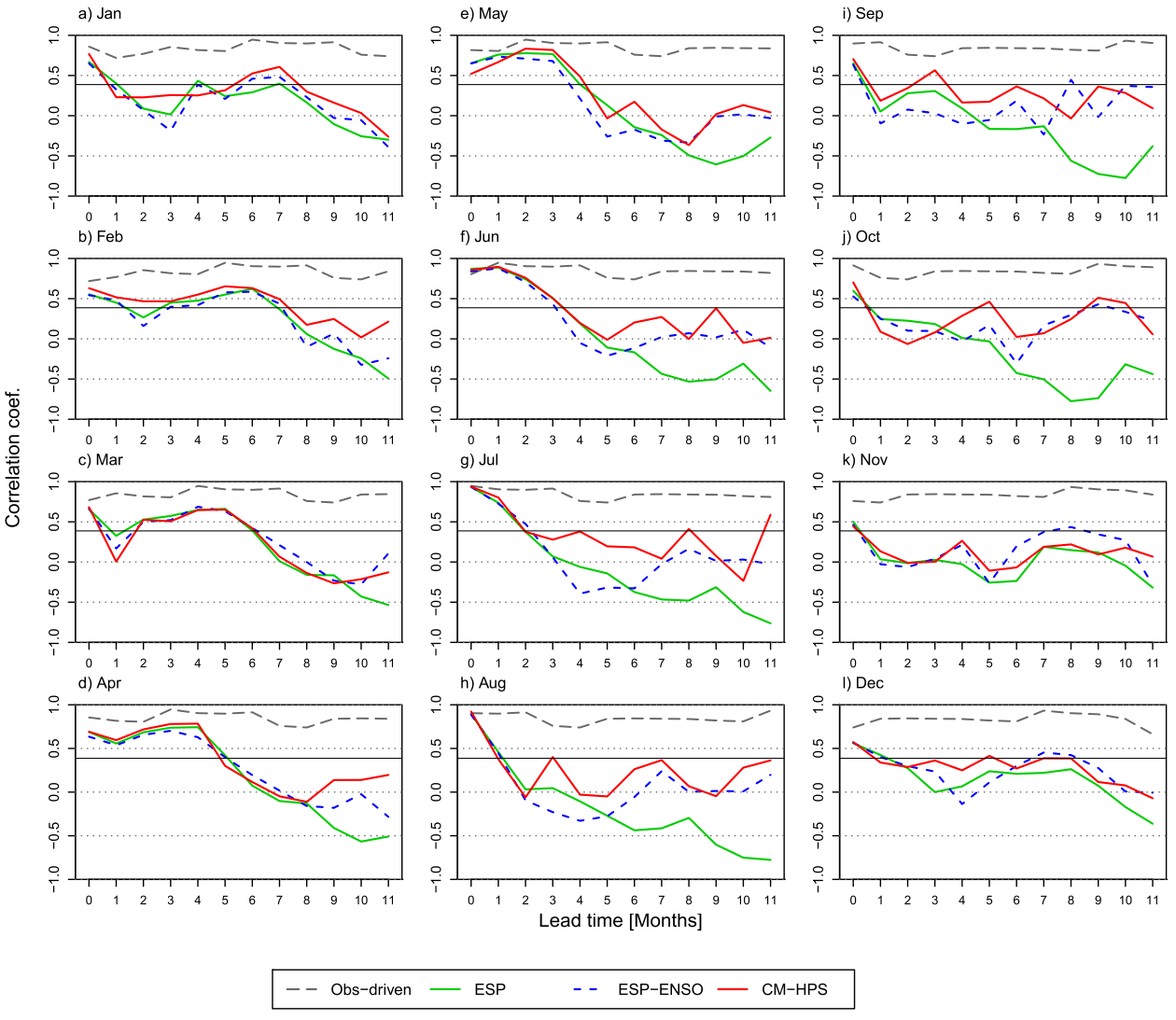

FIG. 6. Correlations of the observation-driven and ensemble means of ESP, ESP-ENSO, and CM-HPS monthly streamflows with observations. Results of the runs initialized at the beginning of indicated months are shown. The horizontal line at 0.387 correlation is the level of statistical significance (at a 5% significance level). All results are for the Fraser–Hope station.

low-flow months. Comparing the results of the three prediction models, the results differ between different subbasins, with slightly higher biases for the CM-HPS for most subbasins. In this case also, the ESP and ESP-ENSO results are mostly indistinguishable.

The combined percent correct score relative to the three observed categories—above normal (>67th percentile), near normal (between 33rd and 67th percentiles), and below normal (<33rd percentile)—are shown in Fig. 9. The skill for most months ranges between 30% and 60%, with generally greater agreement during the first few months that are affected by initial conditions. As the lead time increases, the skill declines for most models. Nevertheless, the percent correct skill for some of the long-lead months are above the 33% level, particularly for the September and December initializations, indicating skill beyond chance. Overall, there is negligible or no percent correct skill improvement by using the CM-HPS, compared to the ESP and ESP-ENSO.

The differences in prediction skill among the different skill metrics illustrate the need to employ multiple metrics for the evaluation of prediction models. While it may be possible to improve the RMSE and bias skill by postprocessing (e.g., Shi et al. 2008), provided that the biases are systematic, it is generally not possible to improve the correlation skill by postprocessing. Therefore, the improved correlation skill (albeit small) of the CM-HPS over the ESP-based methods can be considered as an asset for the seasonal hydrologic prediction.

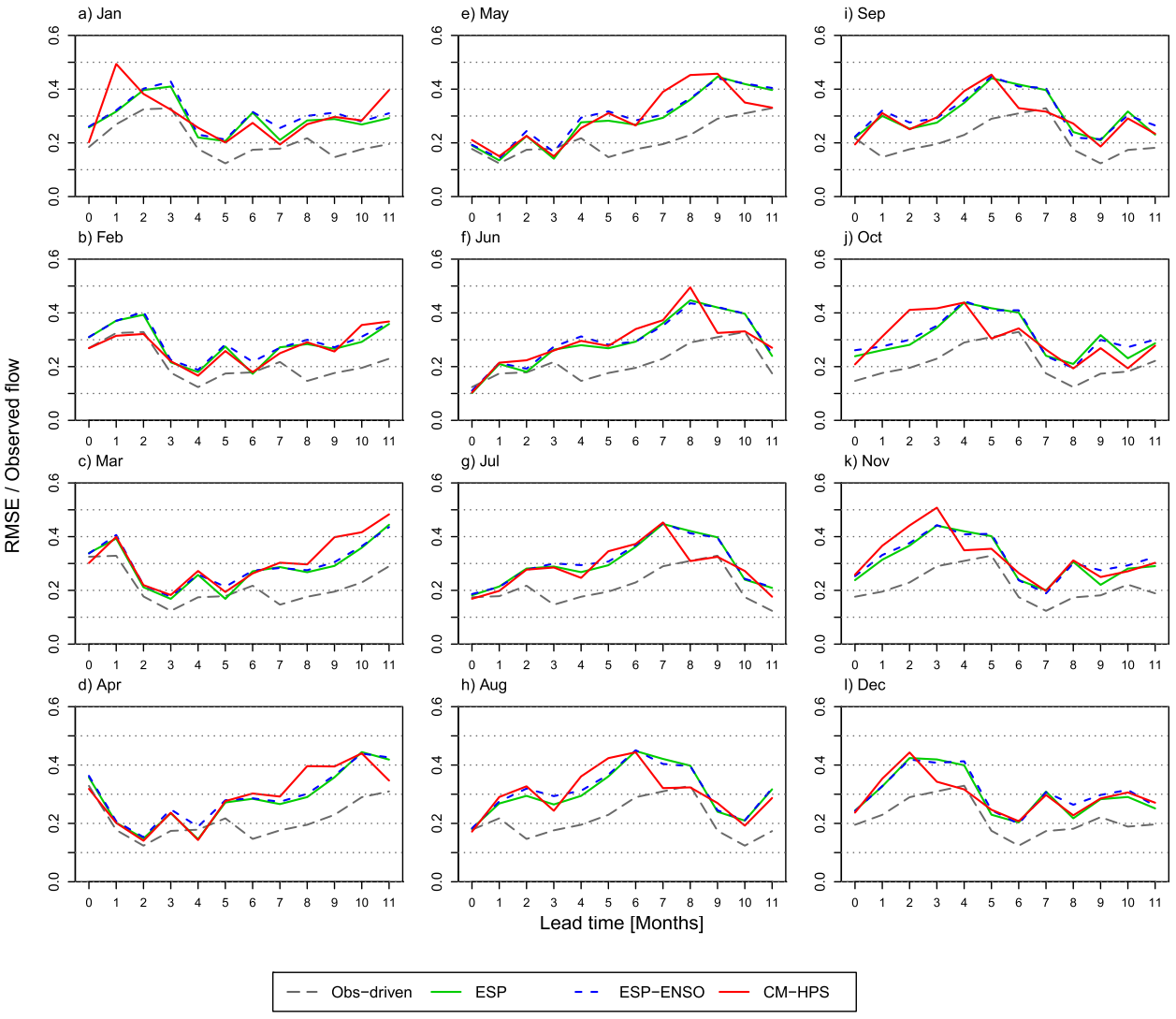

FIG. 7. As in Fig. 6, but for RMSE/observed flow.

e. *Maximum and minimum streamflow prediction skills*

The ability of a model system to predict extreme events, such as peak flow and low flow, is another important criterion for assessing seasonal streamflow prediction. However, given the limited skill of the prediction models, only the maximum and minimum monthly flows are considered (Figs. 10, 11). Given that the maximum (or minimum) flows in the observed and simulated results could occur at different months, the simulated flows were considered for the same months when the observed maximum (or minimum) flows occur.

Given that the FRB is predominantly snowmelt driven, the maximum flow for any given year is the basin response to the accumulated snow storage. Prediction results for the maximum monthly flow, which usually occurs in June, are also influenced by the initial hydrologic conditions. Specifically, the correlation and percent correct categorical skill increase progressively and RMSE/observed flow decrease progressively from January to June initializations (Fig. 10), as the effect of initial conditions (i.e., soil moisture and snow storage) becomes more pronounced. As lead times increase after the July initialization, the correlation and percent correct categorical skill decreases and the RMSE/observed flow skill increases for most months. Nevertheless, statistically significant correlations for some of the subbasins for July–December initializations suggest some ability of the models to represent observed variability beyond model initialization. Percent correct categorical skill beyond July initialization ranges between 30% and 50%, also indicating moderate skill of the model, beyond what could be achieved by chance. Comparing the results of the three methods, none stands

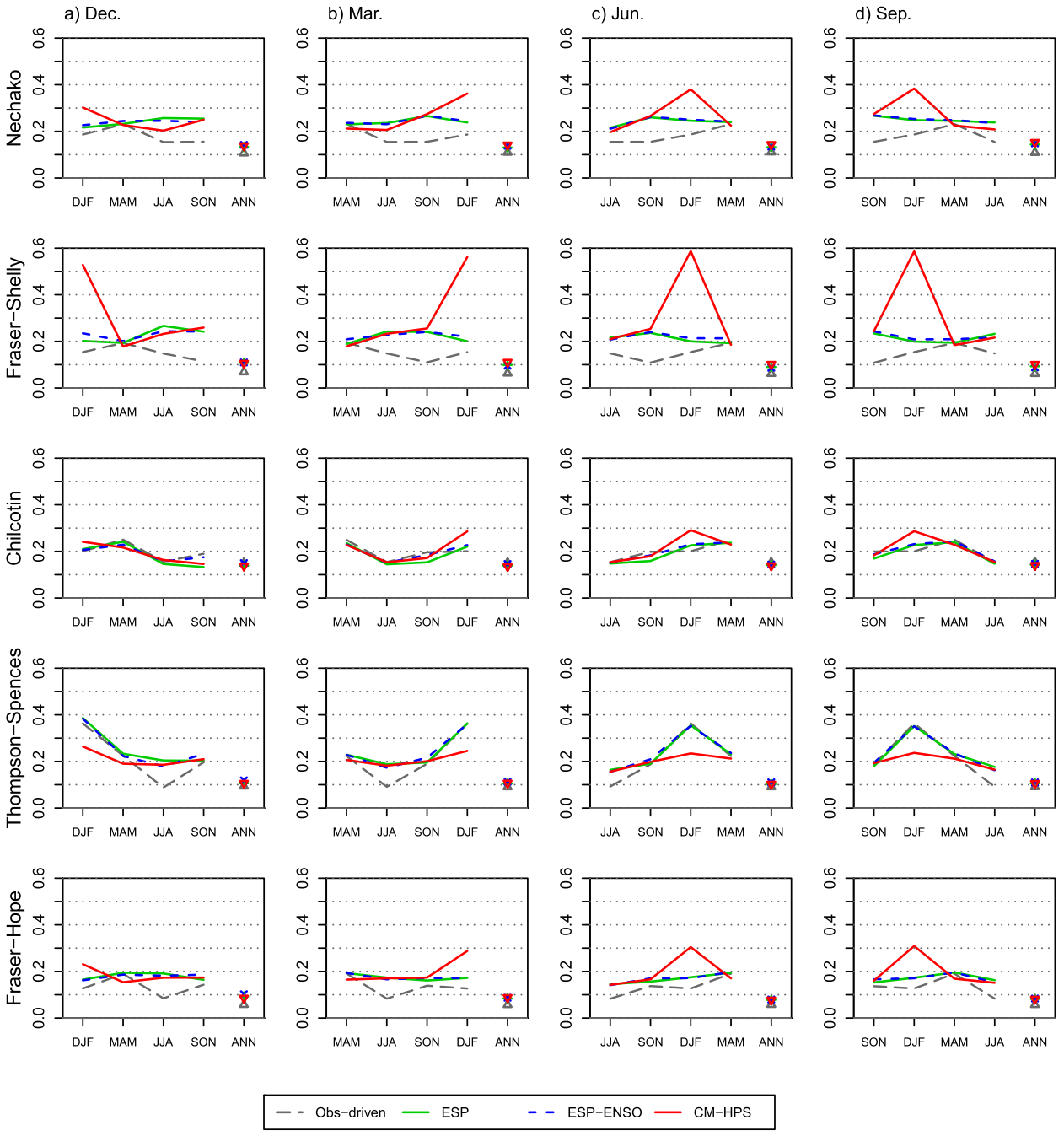

FIG. 8. Mean seasonal and annual streamflow volume biases for five subbasins of the FRB for (a) December, (b) March, (c) June, and (d) September initializations.

out with consistently better skill, indicating no particular advantage in using any of the three approaches for maximum monthly flow prediction. The RMSE/observed flow skills for the Fraser–Shelly and Chilcotin subbasins are close to the observation-driven values, indicating limitations imposed by the hydrologic modeling uncertainties.

In the case of monthly minimum flows, which usually occur in February–March, the difficulties in adequately replicating the observed values by the hydrologic model is apparent, especially for the Nechako and Chilcotin subbasins, where some of the observation-driven values fall below the threshold for statistical significance (Fig. 11). Monthly minimum flows occur because of rainfall and base flow from groundwater and, as a result, are sensitive to the representation of these variables in the prediction systems and hydrologic model (model structure). All three

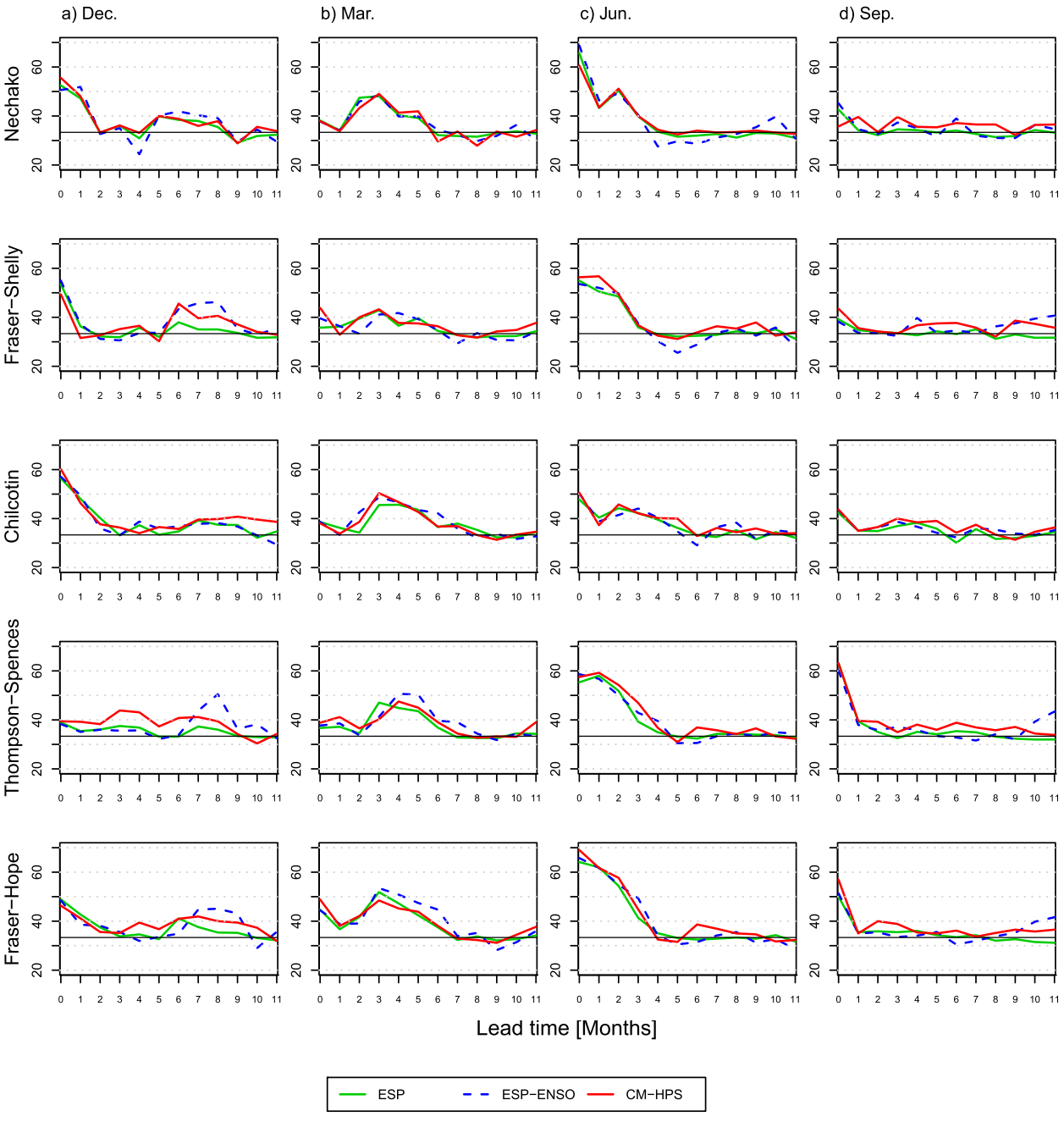

FIG. 9. Percent correct categorical skill score for five subbasins of the FRB for (a) December, (b) March, (c) June, and (d) September initializations. The skill summarizes the proportion of ensemble members falling in the category corresponding to observations. The horizontal line at 33% is the level that could be achieved by chance alone.

prediction models exhibit similarly low correlation values, which are mostly below the statistically significant level, illustrating difficulties in prediction. The normalized RMSE values for the CM-HPS results are larger than the ESP-based methods, except for the Thompson–Spences subbasin. In order to isolate the source of this error from the hydrologic model uncertainty, the RMSE skill was calculated by comparing prediction model results with the observation-driven VIC results (not shown). The patterns of the normalized RMSE remained the same even when compared to the observation-driven VIC results, that is, a larger minimum flow RMSE for the CM-HPS results compared to the ESP-based methods. Larger RMSE for the CM-HPS were also obtained in the Thompson–Spences subbasin. These results imply that the higher RMSE in the CM-HPS originates from the driving climate models,

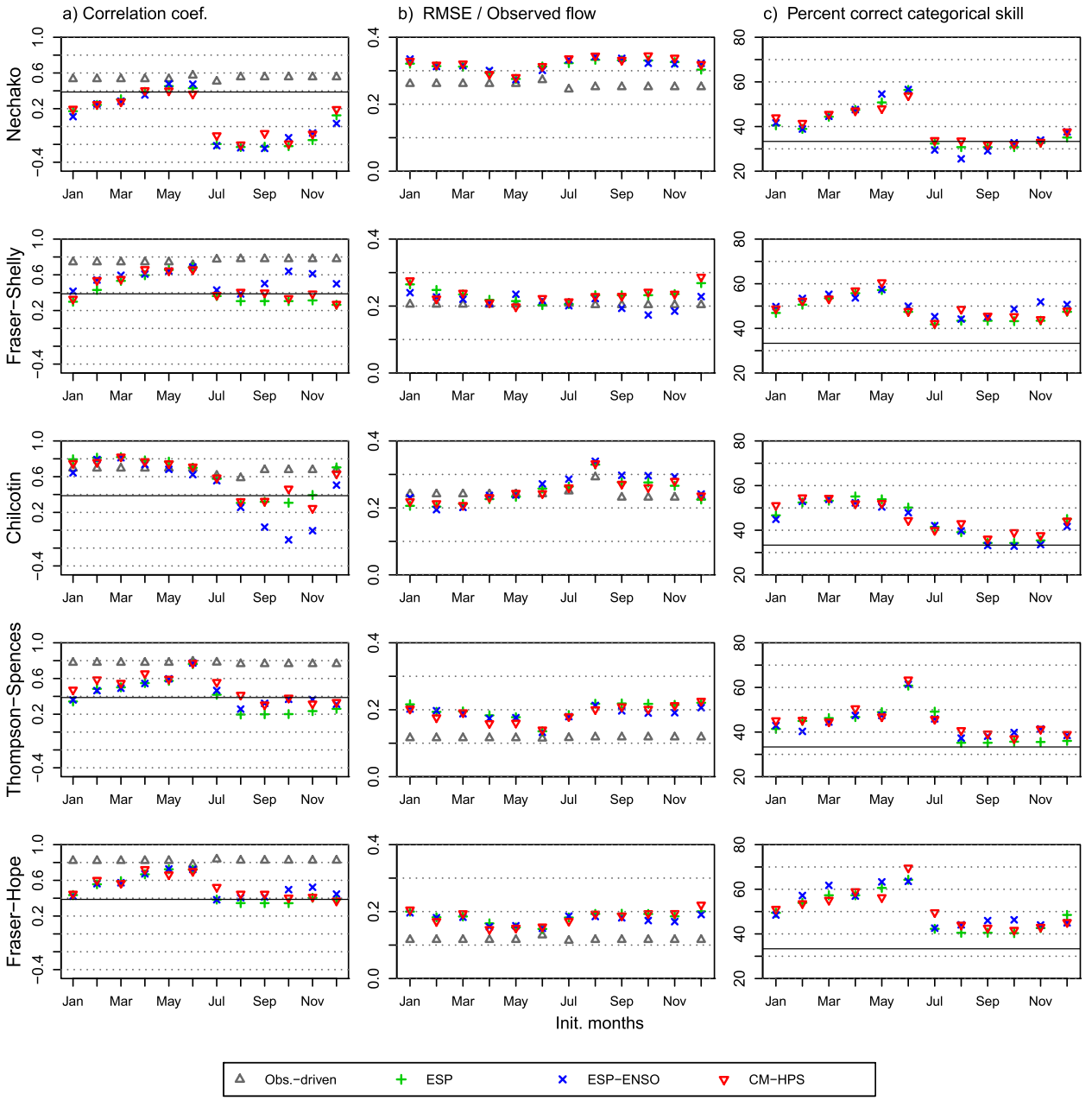

FIG. 10. (a) Correlation coefficient, (b) RMSE/observed flow, and (c) percent correct categorical skills of max monthly flows for five subbasins of the FRB. The horizontal line at 0.387 correlation is the level of statistical significance (at a 5% significance level), and 33% percent correct categorical skill is the level that could be achieved by chance alone. Modeled flows are considered for the months when the observed max flow occur.

including the Thompson–Spences subbasin, where it appears that the hydrologic modeling uncertainty compensates the RMSE when compared to the observed minimum flows. In the case of percentage correct categorical skill, the values from the three prediction models for the five subbasins mostly lie around the 33% level, with none of the prediction models exhibiting consistently superior skill. Overall, the low correlation values and the high RMSE values for all subbasins and all prediction models, part of which arise from the hydrologic modeling uncertainties, reinforce the general lack of skill in predicting low flows.

5. Summary and conclusions

This study evaluated the possibility of using climate outputs from the CCCma's coupled dynamical climate

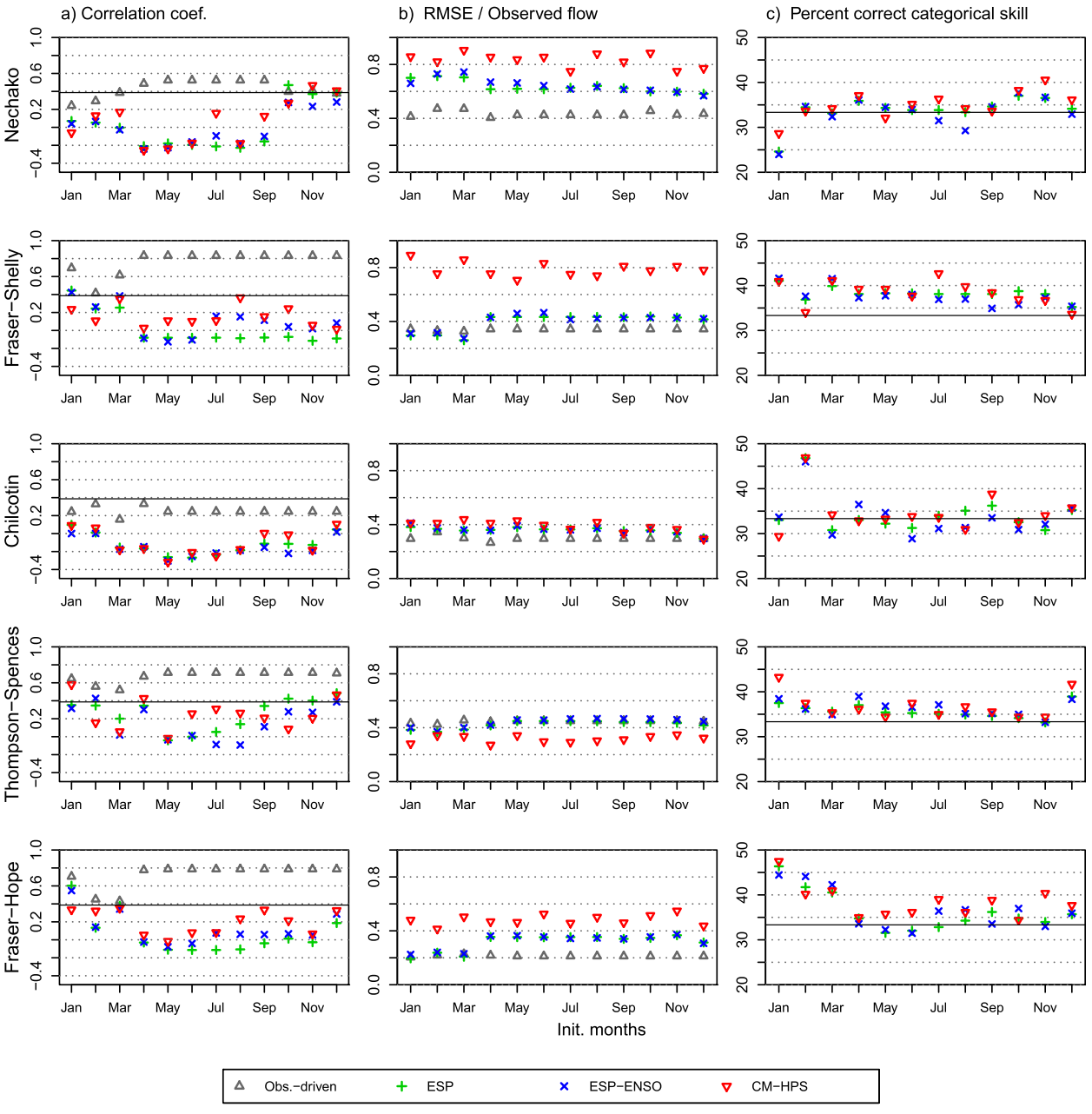

FIG. 11. As in Fig. 10, but for min flows.

models (CanCM3 and CanCM4) as the basis for seasonal streamflow prediction. A test bed climate model–driven hydrologic prediction system (CM-HPS) for a 26-yr (1979–2004) hindcast period was set up using the VIC hydrologic model for the Fraser River basin (FRB), Canada. The BCSD downscaling algorithm was employed to statistically downscale the daily precipitation and air temperature from CanCM3 and CanCM4. For comparison, two ESP-based setups with daily temperature and precipitation derived from (i) 26-yr (1979–2004) historical and future traces (ESP) and (ii) the 1979–2004 historical and future traces conditioned with the ENSO state at the time of hindcast (ESP-ENSO) were employed.

The comparison of precipitation correlations from the two methods revealed limited skill for both CanCM3/CanCM4 and ENSO-conditioned mean hindcast beyond the zero lead month. Compared to precipitation, temperature correlations were found to be relatively superior, especially considering the results from CanCM3/CanCM4 for the months of January–April. As in the case of previous studies (Merryfield et al. 2011, 2013a),

the modest winter–spring temperature prediction skill can be linked to the teleconnected influence of ENSO.

Evaluating the effect of the ENSO predictability on hydrologic prediction skills is complicated by initial hydrologic conditions, especially for the spring months when winter snow storage from model initializations influences the results. While the hydrologic predictability for these months is mainly an initial state problem, the role of the initial state becomes less important after the snowmelt (i.e., July), and the hydrologic predictability mainly derives from the precipitation and temperature prediction skills. Considering the SWE correlation for months not influenced by the initial hydrologic conditions (May–November) revealed no skill for the ESP-based results and moderate skill for the CM-HPS. In the case of monthly streamflow, while the correlation skills from the three models remain low beyond the months affected by the initial conditions, modest improvements in the skill for the CM-HPS over ESP-based methods were noted, especially for the June–September initializations, attributable to the improved temperature and SWE prediction skills. The RMSE, bias, and percent correct skills from the three models are mostly similar. All three prediction models have similar maximum flow prediction skills, implying no particular advantage in using any of the three methods. In the case of minimum flows, the skill values are generally low, illustrating an overall lack of skill in predicting low flows. In some cases, hydrologic prediction skill was found to be affected by the skill of the hydrologic model itself. Specifically, for some subbasins, the RMSE and model bias reach the limits imposed by the VIC. Comparing the two ESP-based (ESP and ESP-ENSO) methods, there was little skill improvement for ESP-ENSO. This highlights the difficulty in conditioning inputs according to the ENSO states at the time of model initializations. Specifically, the ENSO state at the time of model initialization may not persist for the entire year used for prediction, thereby causing the model performance skill to remain low.

Given these findings, an important question arises on the usefulness and reliability of the CM-HPS. Overall, the skill in using the current generation of CCCma models (CanCM3/CanCM4) for streamflow prediction remains low. Although the climate model's ability to simulate the ENSO states provides additional skill, the improvement over the ESP-based methods is modest. Hence, it may be worthwhile to combine the CM-HPS with the ESP-based methods, for example, for conditioning each target month of the ESP–ENSO approach according to the ENSO state prediction of the CM-HPS approach. Given that some of the skill metrics are constrained by the hydrologic model skill, improvement in the hydrologic model skill (e.g., by improving the hydrologic model structure) could potentially improve some of the prediction skills. Additionally, as in the case of the previous study by Yuan et al. (2013), postprocessing of the results could improve some of the skill metrics (e.g., RMSE and model bias). Furthermore, given that CanCM3 and CanCM4 were part of the nine-model North American Multi-Model Ensemble (NMME) (Kirtman et al. 2014), it will be of interest to the scientific community to evaluate the hydrologic prediction skill of the multimodel ensembles. These points could be foci for future research toward improved seasonal hydrologic prediction. Looking ahead, further enhancement in the representation of climate state, variability, and teleconnections can be expected in future versions of dynamical climate models, which could potentially lead to enhanced performance of hydrologic prediction systems.

Acknowledgments. Funding for this project was provided by a Grants and Contribution Agreement with the Canadian Centre for Climate Modelling and Analysis (CCCma) of Environment Canada. F.W. Zwiers (PCIC) gave valuable feedback on the manuscript. We would also like to thank three anonymous reviewers for their comments, which led to an improved version of the manuscript.

REFERENCES

Arora, V. K., and Coauthors, 2011: Carbon emission limits required to satisfy future representative concentration pathways of greenhouse gases. *Geophys. Res. Lett.,* **38,** L05805, doi:10.1029/2010GL046270.

Barnston, A. G., M. K. Tippett, M. L. L'Heureux, S. Li, and D. G. DeWitt, 2012: Skill of real-time seasonal ENSO model predictions during 2002–11: Is our capability increasing? *Bull. Amer. Meteor. Soc.,* **93** (Suppl.), ES48–ES50, doi:10.1175/BAMS-D-11-00111.2.

Bastola, S., V. Misra, and H. Li, 2013: Seasonal hydrological forecasts for watersheds over the southeastern United States for the boreal summer and fall seasons. *Earth Interact.,* **17,** doi:10.1175/2013EI000519.1.

Beven, K., 2006: A manifesto for the equifinality thesis. *J. Hydrol.,* **320,** 18–36, doi:10.1016/j.jhydrol.2005.07.007.

Bohn, T. J., M. Y. Sonessa, and D. P. Lettenmaier, 2010: Seasonal hydrologic forecasting: do multimodel ensemble averages always yield improvements in forecast skill? *J. Hydrometeor.,* **11,** 1358–1372, doi:10.1175/2010JHM1267.1.

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. Climate,* **25,** 4366–4388, doi:10.1175/JCLI-D-11-00408.1.

Cayan, D. R., K. T. Redmond, and L. G. Riddle, 1999: ENSO and hydrologic extremes in the western United States. *J. Climate,* **12,** 2881–2893, doi:10.1175/1520-0442(1999)012<2881:EAHEIT>2.0.CO;2.

Clark, M. P., M. C. Serreze, and G. J. McCabe, 2001: Historical effects of El Niño and La Niña events on the seasonal evolution of the montane snowpack in the Columbia and Colorado

River basins. *Water Resour. Res.,* **37,** 741–757, doi:10.1029/2000WR900305.

Fleming, S. W., P. H. Whitfield, R. D. Moore, and E. J. Quilty, 2007: Regime-dependent streamflow sensitivities to Pacific climate modes cross the Georgia–Puget transboundary ecoregion. *Hydrol. Processes,* **21,** 3264–3287, doi:10.1002/hyp.6544.

Franz, K., H. C. Hartmann, S. Sorooshian, and R. Bales, 2003: Verification of National Weather Service ensemble streamflow predictions for water supply forecasting in the Colorado River basin. *J. Hydrometeor.,* **4,** 1105–1118, doi:10.1175/1525-7541(2003)004<1105:VONWSE>2.0.CO;2.

Fraser Basin Council, 2010: Environmental protection in flood hazard management: A guide to practitioners. Report, 68 pp. [Available online at www.fraserbasin.bc.ca/_Library/Water/report_flood_and_environmental_protection_2010.pdf.]

Global Soil Data Task, 2000: Global Soil Data Products CD-ROM contents (IGBP-DIS). Oak Ridge National Laboratory Distributed Active Archive Center, Oak Ridge, TN, digital media, doi:10.3334/ORNLDAAC/565.

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,** 380–391, doi:10.1080/07055900.2012.759899.

Hamlet, A. F., E. P. Salathé, and P. Carrasco, 2010: Statistical downscaling techniques for global climate model simulations of temperature and precipitation with application to water resources planning studies. Final Rep. for the Columbia Basin Climate Change Scenarios Project, 27 pp. [Available online at http://warm.atmos.washington.edu/2860/r7climate/study_report/CBCCSP_chap4_gcm_final.pdf.]

Hidalgo, H. G., and J. A. Dracup, 2003: ENSO and PDO effects on hydroclimatic variations of the upper Colorado River basin. *J. Hydrometeor.,* **4,** 5–23, doi:10.1175/1525-7541(2003)004<0005:EAPEOH>2.0.CO;2.

Jarvis, A., H. I. Reuter, A. Nelson, and E. Guevara, 2008: Hole-filled SRTM for the globe version 4. SRTM 90m digital elevation database, CGIAR-CSI, Washington, DC. [Available online at http://srtm.csi.cgiar.org.]

Kalnay, E., and Coauthors, 1996: The NCEP/NCAR 40-Year Reanalysis Project. *Bull. Amer. Meteor. Soc.,* **77,** 437–471, doi:10.1175/1520-0477(1996)077<0437:TNYRP>2.0.CO;2.

Kharin, V. V., F. W. Zwiers, and N. Gagnon, 2001: Skill of seasonal hindcasts as a function of the ensemble size. *Climate Dyn.,* **17,** 835–843, doi:10.1007/s003820100149.

Kirtman, B. P., and Coauthors, 2014: The North American Multimodel Ensemble: Phase-1 seasonal-to-interannual prediction; phase-2 toward developing intraseasonal prediction. *Bull. Amer. Meteor. Soc.,* **95,** 585–601, doi:10.1175/BAMS-D-12-00050.1.

Li, H., L. Luo, E. F. Wood, and J. Schaake, 2009: The role of initial conditions and forcing uncertainties in seasonal hydrologic forecasting. *J. Geophys. Res.,* **114,** D04114, doi:10.1029/2008JD010969.

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,** 14 415–14 428, doi:10.1029/94JD00483.

——, E. F. Wood, and D. P. Lettenmaier, 1996: Surface soil moisture parameterization of the VIC-2L model: Evaluation and modification. *Global Planet. Change,* **13,** 195–206, doi:10.1016/0921-8181(95)00046-1.

——, Z. Xie, and M. Huang, 2003: A new parameterization for surface and groundwater interactions and its impact on water budgets with the variable infiltration capacity (VIC) land surface model. *J. Geophys. Res.,* **108,** 8613, doi:10.1029/2002JD003090.

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

Mahanama, S. P. P., R. D. Koster, R. H. Reichle, and L. Zubair, 2008: The role of soil moisture initialization in subseasonal and seasonal streamflow prediction—A case study in Sri Lanka. *Adv. Water Resour.,* **31,** 1333–1343, doi:10.1016/j.advwatres.2008.06.004.

Maurer, E. P., and H. G. Hidalgo, 2008: Utility of daily vs. monthly large-scale climate data: An intercomparison of two statistical downscaling methods. *Hydrol. Earth Syst. Sci.,* **12,** 551–563, doi:10.5194/hess-12-551-2008.

McCabe, G. J., and M. D. Dettinger, 2002: Primary modes and predictability of year-to-year snowpack variations in the western United States from teleconnections with Pacific Ocean climate. *J. Hydrometeor.,* **3,** 13–25, doi:10.1175/1525-7541(2002)003<0013:PMAPOY>2.0.CO;2.

Merryfield, W. J., B. Denis, J.-S. Fontecilla, W.-S. Lee, S. Kharin, J. Hodgson, and B. Archambault, 2011: The Canadian Seasonal to Interannual Prediction System (CanSIPS): An overview of its design and operational implementation. Tech. Note, Environment Canada, 51 pp. [Available online at http://collaboration.cmc.ec.gc.ca/cmc/cmoi/product_guide/docs/lib/op_systems/doc_opchanges/technote_cansips_20111124_e.pdf.]

——, and Coauthors, 2013a: The Canadian Seasonal to Interannual Prediction System. Part I: Models and initialization. *Mon. Wea. Rev.,* **141,** 2910–2945, doi:10.1175/MWR-D-12-00216.1.

——, W.-S. Lee, W. Wang, M. Chen, and A. Kumar, 2013b: Multisystem seasonal predictions of Arctic sea ice. *Geophys. Res. Lett.,* **40,** 1551–1556, doi:10.1002/grl.50317.

Moore, R. D., 1991: Hydrology and water supply in the Fraser River basin. *Water in Sustainable Development: Exploring Our Common Future in the Fraser River Basin,* A. H. J Dorcey and J. R. Griggs, Eds., Wastewater Research Centre, University of British Columbia, 21–40.

Moriasi, D. N., J. G. Arnold, M. W. Van Liew, R. L. Bingner, R. D. Harmel, and T. L. Veith, 2007: Model evaluation guidelines for systematic quantification of accuracy in watershed simulations. *Trans. ASABE,* **50,** 885–900, doi:10.13031/2013.23153.

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,** 230–244, doi:10.1016/S0022-1694(02)00065-3.

Polade, S. D., A. Gershunov, D. R. Cayan, M. D. Dettinger, and D. W. Pierce, 2013: Natural climate variability and teleconnections to precipitation over the Pacific–North American region in CMIP3 and CMIP5 models. *Geophys. Res. Lett.,* **40,** 2296–2301, doi:10.1002/grl.50491.

Rice, S. P., M. Church, C. L. Wooldridge, and E. J. Hickin, 2009: Morphology and evolution of bars in a wandering gravel-bed river; lower Fraser River, British Columbia, Canada. *Sedimentology,* **56,** 709–736, doi:10.1111/j.1365-3091.2008.00994.x.

Saha, S., and Coauthors, 2006: The NCEP Climate Forecast System. *J. Climate,* **19,** 3483–3517, doi:10.1175/JCLI3812.1.

Schnorbus, M., K. Bennett, and A. Werner, 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 Rep. BC-X-423, Canadian Forest Service, 64 pp. [Available online at http://cfs.nrcan.gc.ca/publications?id=31207.]

——, ——, ——, and A. Berland, 2011: Hydrologic impacts of climate change in the peace, Campbell and Columbia watersheds, British Columbia, Canada. Final Rep. (Part II), Pacific Climate Impacts Consortium, University of Victoria, Victoria, BC, Canada, 157 pp. [Available online at www.pacificclimate.org/sites/default/files/publications/Schnorbus.HydroModelling.FinalReport2.Apr2011.pdf.]

Scinocca, J. F., N. A. McFarlane, M. Lazare, J. Li, and D. Plummer, 2008: Technical Note: The CCCma third generation AGCM and its extension into the middle atmosphere. *Atmos. Chem. Phys.,* **8,** 7055–7074, doi:10.5194/acp-8-7055-2008.

Seager, R., Y. Kushnir, J. Nakamura, M. Ting, and N. Naik, 2010: Northern Hemisphere winter snow anomalies: ENSO, NAO and the winter of 2009/10. *Geophys. Res. Lett.,* **37,** L14703, doi:10.1029/2010GL043830.

Shabbar, A., 2006: The impact of El Niño–Southern Oscillation on the Canadian climate. *Adv. Geosci.,* **6,** 149–153, doi:10.5194/adgeo-6-149-2006.

Shi, X., A. W. Wood, and D. P. Lettenmaier, 2008: How essential is hydrologic model calibration to seasonal streamflow forecasting? *J. Hydrometeor.,* **9,** 1350–1363, doi:10.1175/2008JHM1001.1.

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,** 1840–1860, doi:10.1002/hyp.9283.

——, D. L. Peters, and M. A. Schnorbus, 2014: Evaluating the ability of a hydrologic model to replicate hydro-ecologically relevant indicators. *Hydrol. Processes,* **28,** 4294–4310, doi:10.1002/hyp.9997.

Shukla, S., and D. P. Lettenmaier, 2011: Seasonal hydrologic prediction in the United States: Understanding the role of initial hydrologic conditions and seasonal climate forecast skill. *Hydrol. Earth Syst. Sci.,* **15,** 3529–3538, doi:10.5194/hess-15-3529-2011.

Sigmond, M., J. C. Fyfe, G. M. Flato, V. V. Kharin, and W. J. Merryfield, 2013: Seasonal forecast skill of Arctic sea ice area in a dynamical forecast system. *Geophys. Res. Lett.,* **40,** 529–534, doi:10.1002/grl.50129.

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

Wang, E., Y. Zhang, J. Luo, F. H. S. Chiew, and Q. J. Wang, 2011: Monthly and seasonal streamflow forecasts using rainfall–runoff modeling and historical weather data. *Water Resour. Res.,* **47,** W05516, doi:10.1029/2010WR009922.

Weare, B. C., 2013: El Niño teleconnections in CMIP5 models. *Climate Dyn.,* **41,** 2165–2177, doi:10.1007/s00382-012-1537-3.

Werner, A. T., 2011: BCSD downscaled transient climate projections for eight select GCMs over British Columbia, Canada. Hydrologic Modelling Project Final Rep. (Part I), Pacific Climate Impacts Consortium, University of Victoria, Victoria, BC, Canada, 63 pp. [Available online at www.pacificclimate.org/sites/default/files/publications/Werner.HydroModelling.FinalReport1.Apr2011.pdf.]

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,** 1–27, doi:10.4296/cwrj3501001.

Wilby, R. L., C. W. Dawson, and E. M. Barrow, 2002: SDSM—A decision support tool for the assessment of regional climate change impacts. *Environ. Modell. Software,* **17,** 145–157, doi:10.1016/S1364-8152(01)00060-3.

Wilks, D. S., 2006: *Statistical Methods in the Atmospheric Sciences: An Introduction.* Academic Press, 650 pp.

Winkler, J. A., J. P. Palutikof, J. A. Andresen, and C. M. Goodess, 1997: The simulation of daily temperature time series from GCM output. Part II: Sensitivity analysis of an empirical transfer function methodology. *J. Climate,* **10,** 2514–2532, doi:10.1175/1520-0442(1997)010<2514:TSODTT>2.0.CO;2.

Wood, A. W., E. P. Maurer, A. Kumar, and D. P. Lettenmaier, 2002: Long-range experimental hydrologic forecasting for the eastern United States. *J. Geophys. Res.,* **107,** 4429, doi:10.1029/2001JD000659.

——, L. Leung, V. Sridhar, and D. Lettenmaier, 2004: Hydrologic implications of dynamical and statistical approaches to downscaling climate model outputs. *Climatic Change,* **62,** 189–216, doi:10.1023/B:CLIM.0000013685.99609.9e.

——, A. Kumar, and D. P. Lettenmaier, 2005: A retrospective assessment of National Centers for Environmental Prediction climate model–based ensemble hydrologic forecasting in the western United States. *J. Geophys. Res.,* **110,** D04105, doi:10.1029/2004JD004508.

Wulder, M., J. Dechka, M. Gillis, J. Luther, R. Hall, and A. Beaudoin, 2003: Operational mapping of the land cover of the forested area of Canada with Landsat data: EOSD land cover program. *For. Chron.,* **79,** 1075–1083, doi:10.5558/tfc791075-6.

Yapo, P. O., H. V. Gupta, and S. Sorooshian, 1998: Multi-objective global optimization for hydrologic models. *J. Hydrol.,* **204,** 83–97, doi:10.1016/S0022-1694(97)00107-8.

Younas, W., and Y. Tang, 2013: PNA predictability at various time scales. *J. Climate,* **26,** 9090–9114, doi:10.1175/JCLI-D-12-00609.1.

Yuan, X., and E. F. Wood, 2013: Multimodel seasonal forecasting of global drought onset. *Geophys. Res. Lett.,* **40,** 4900–4905, doi:10.1002/grl.50949.

——, ——, J. K. Roundy, and M. Pan, 2013: CFSv2-based seasonal hydroclimatic forecasts over the conterminous United States. *J. Climate,* **26,** 4828–4847, doi:10.1175/JCLI-D-12-00683.1.