

Forecast of streamflows to the Arctic Ocean by a Bayesian neural network model with snowcover and climate inputs

Kabir Rasouli, Bouchra R. Nasri, Armina Soleymani,
Taufique H. Mahmood, Masahiro Hori and Ali Torabi Haghighi

ABSTRACT

Increasing water flowing into the Arctic Ocean affects oceanic freshwater balance, which may lead to the thermohaline circulation collapse and unpredictable climatic conditions if freshwater inputs continue to increase. Despite the crucial role of ocean inflow in the climate system, less is known about its predictability, variability, and connectivity to cryospheric and climatic patterns on different time scales. In this study, multi-scale variation modes were decomposed from observed daily and monthly snowcover and river flows to improve the predictability of Arctic Ocean inflows from the Mackenzie River Basin in Canada. Two multi-linear regression and Bayesian neural network models were used with different combinations of remotely sensed snowcover, *in-situ* inflow observations, and climatic teleconnection patterns as predictors. The results showed that daily and monthly ocean inflows are associated positively with decadal snowcover fluctuations and negatively with interannual snowcover fluctuations. Interannual snowcover and antecedent flow oscillations have a more important role in describing the variability of ocean inflows than seasonal snowmelt and large-scale climatic teleconnection. Both models forecasted inflows seven months in advance with a Nash–Sutcliffe efficiency score of ≈ 0.8 . The proposed methodology can be used to assess the variability of the freshwater input to northern oceans, affecting thermohaline and atmospheric circulations.

Key words | Arctic ocean, Bayesian neural network, climate teleconnections, Mackenzie River Basin, snowcover extent, streamflow forecast

Kabir Rasouli (corresponding author)
Meteorological Service of Canada, Environment
and Climate Change Canada,
2121 Trans Canada Route, Dorval, QC, H9P 1J3,
Canada
and
Department of Geoscience,
University of Calgary,
Calgary, AB,
Canada
E-mail: kabir.rasouli@usask.ca

Bouchra R. Nasri
GERAD and Department of Mathematics and
Statistics,
McGill University,
Montreal, QC,
Canada

Armina Soleymani
Department of Systems Design Engineering,
University of Waterloo,
ON,
Canada

Taufique H. Mahmood
Harold Hamm, School of Geology and Geological
Engineering,
University of North Dakota,
Grand Forks, ND,
USA

Masahiro Hori
Earth Observation Research Center,
Japan Aerospace Exploration Agency,
Japan

Ali Torabi Haghighi
Water, Energy and Environmental Engineering
Research Unit,
University of Oulu,
Finland

INTRODUCTION

The thermohaline circulation is a key element of the global climate system. The freshwater released from ice and snowmelt to oceans in the high latitudes sustains the global

thermohaline circulation. Runoff from northern rivers (38%), inflow through Bering Strait from the Pacific Ocean (30%), net precipitation on the ocean (24%), and other sources (8%) contribute to the Arctic Ocean freshwater system (Serreze *et al.* 2006). Circumpolar river basins discharge about $3,300 \text{ km}^3$ per year of freshwater to the Arctic Ocean (Rachold *et al.* 2004) that sustains the ocean

This is an Open Access article distributed under the terms of the Creative Commons Attribution Licence (CC BY 4.0), which permits copying, adaptation and redistribution, provided the original work is properly cited (<http://creativecommons.org/licenses/by/4.0/>).

doi: 10.2166/nh.2020.164

stratification, with the upper 200-m layer being cold and fresh and the lower layer being warm and salty, similar to the Atlantic Ocean (Aagaard & Carmack 1989). Streamflows in the northern rivers have increased significantly (Syed *et al.* 2007; Rood *et al.* 2017; Durocher *et al.* 2019) in response to the Arctic amplification of temperature change in the last decades (Johannessen *et al.* 2016). The increased water flowing to the Arctic Ocean has caused a general freshening tendency over the Siberian shelf as a result of increasing discharge from Eurasian rivers (Polyakov *et al.* 2008) and the Canadian Arctic as a result of increasing discharge from North American rivers (Fichot *et al.* 2013). Peterson (2006) found that the increasing river discharge and excess net precipitation on the ocean contributed $\approx 20,000 \text{ km}^3$ of freshwater to the high-latitude oceans from the 1960s to the 1990s. The increases in Arctic freshwater flux have affected the Atlantic Meridional Overturning Circulation, an important component of ocean thermohaline circulation (Yang *et al.* 2016). Macdonald *et al.* (1999) found that the high freshening at the Surface Heat Budget of the Arctic Ocean (SHEBA) site in 1997 was associated with an exceptionally high inflow from the Mackenzie River Basin (MRB). Increasing river discharges to the Arctic Ocean with increased water temperatures (Padilla *et al.* 2015) transfer a large heat flux from the land water to the northern oceans. Yang *et al.* (2014) reported an estimated heat flux of 790×10^9 megajoules to the Arctic Ocean only from the MRB. The heat flux in the MRB does not necessarily follow the flow regime and the largest heat flux ($3,100 \times 10^9$ megajoules) occurs in July with warm water temperatures, while the highest flow occurs one month earlier (Yang *et al.* 2014).

The freshwater input and its transport from the Arctic Ocean to the North Atlantic have high interannual variability (Serreze *et al.* 2006), which has a large effect on regional and global climate variability (Proshutinsky *et al.* 2002). Freshwater discharge from large snow-dominated rivers varies with weather patterns, seasonality of snowcover extent (SCE), and large-scale climatic teleconnections such as the Arctic Oscillation (AO) and Antarctic Oscillation (AAO) (Wang *et al.* 2006; Rasouli 2010). Large-scale atmospheric circulations like El-Niño and La-Niña (Cohen & Entekhabi 1999) affect SCE. Studies show that there is a strong correlation between SCE and winter Pacific–North

American (PNA) and North Atlantic Oscillation (NAO) teleconnections in North America and Canada (Gutzler & Rosen 1992; Brown & Goodison 1996). A strong link between NAO and SCE was also found in Europe and central Asia (Bojariu & Gimeno 2003). For instance, the unusually cold conditions over the last decade of the 20th century near Greenland and the eastern Mediterranean region and the anomalous warmth over northern Europe, Russia, and much of Central Asia are related to NAO (Hurrell 1995). The link between SCE and large-scale atmospheric circulations has been intensified in the last decades as winter and early spring temperatures have increased in western North America (Cohen & Entekhabi 1999; Folland *et al.* 2001) and the Arctic (Eythorsson *et al.* 2019).

Changes in SCE affect local-scale ground heat fluxes, snow redistribution by blowing wind (Pomeroy & Gray 1995; Liston 1999), soil moisture, and soil thermal regime (Park *et al.* 2015). Changes in the soil thermal regime alter thawing and freezing mechanisms, which affect water storage in soils and runoff patterns in cold regions and, hence, the rate of discharge to the Arctic Ocean (Mackay & MacKay 1974; Bao *et al.* 2011). An increase in ground temperatures, especially in northern regions, has increased the global temperature in response to anomalies in the atmospheric circulations (Coulibaly & Burn 2004). Snow recedes quickly, and snowmelt is intensified under increased ground temperatures, which can affect infiltration and streamflow, especially when rain-on-snow events occur. Burn *et al.* (2004) showed an increasing trend in the annual minimum flow and a decreasing trend in the date of occurrence of the spring freshet in two main subbasins of MRB attributed to changes in snow processes under climate changes. In a relatively mild climate in the southern parts of MRB, the annual streamflow of the rivers feeding Lake Athabasca showed decreasing trends over the last half-century (Rasouli *et al.* 2013, 2015a). As circumpolar drainage rivers are the main contribution to the freshwater in the northern latitudes, it is crucial to study the predictability and variability of discharges from these rivers under spatial and temporal variations of climate and snowcover extent. Variability of snow and streamflow regimes in small river basins have been linked to climatic teleconnections in the literature (e.g., Bojariu & Gimeno 2003; Rasouli *et al.* 2012). Ek & Ritchie (1996) forecasted surface water and

energy fluxes over MRB using the Canadian global spectral operational forecast model and found that forecasts are strongly sensitive to a net accumulation of precipitation minus evaporation. The model, however, was not skillful in long-range forecasts when extended to one month. Temimi *et al.* (2004) studied the near real-time floods over MRB using passive microwave data. Spence *et al.* (2007) applied a combination of canonical correlation and multiple regression analyses to forecast streamflows in MRB and investigate the effects of discontinuing monitoring gauges on forecasting the ungauged streamflows.

To our knowledge, there is no previous research on forecasting the Arctic Ocean inflows from the land-rivers incorporating snowcover and climate variations. The contribution of this study is to fill this knowledge gap by providing short-term and long-term forecasts and identifying the role of snow, climate, and streamflow memory of the northern rivers on decadal to sub-seasonal time scale in the variability of the freshwater input to the Arctic Ocean. The objectives of this research are to (1) identify the linkages of Arctic Ocean inflow with SCE, antecedent inflows and climate teleconnection patterns, and (2) improve the predictability of ocean inflows from the Mackenzie River at multiple temporal scales. More specifically, we differentiate the roles of antecedent river flow conditions, climatic teleconnection patterns, and multi-decadal to monthly variations of snowcover in forecasting inflows of the Arctic Ocean one day to seven months in advance. Linking regional-scale SCE and large-scale climate variations to ocean inflows from the large river basins has important implications in predicting the freshwater input to the Arctic Ocean, estimating rates of freshwater transport to the North Atlantic, and better understanding climate variability and predictability in the near future in regional and global scales.

STUDY AREA AND DATA SOURCES

The case study in this paper is MRB in western Canada, which is fed by Great Slave Lake, Great Bear Lake, and Liard and Peace rivers, and it releases water at a rate of $9,910 \text{ m}^3 \text{ s}^{-1}$ into the Beaufort Sea in the Arctic Ocean. The basin drainage area is $1,679,100 \text{ km}^2$. The Mackenzie River yields the largest North American source of freshwater discharge into the

Arctic Ocean. Figure 1 shows MRB, its subbasins, and drainage systems. There is a delta near the Beaufort Sea at the mouth of the basin. In this paper, combinations of data for antecedent flow in MRB, remotely sensed snowcover extent, and climatic indices such as NAO, AO, AAO, and PNA are used to forecast daily and monthly inflows of the Arctic Ocean from the MRB. Data cover 38 years, from January 1979 to December 2016 with daily time steps. The remote sensing data acquired by well-calibrated instruments for snow cover are not available before 1979. Therefore, the dataset used in this study is limited to post-1979 data. Data for the Arctic Ocean inflows were used from the gauge on MRB at Arctic Red River (# 10LC014), recording streamflows at the mouth of the Mackenzie River, and were obtained from the website of the Water Survey of Canada (<https://wateroffice.ec.gc.ca>). The Peel River is one of the MRB subbasins, but it joins the MRB after the gauge. For consistency between the streamflow and snowcover, the domain for the remotely sensed snowcover products (Figure 1) was matched with the domain of MRB, excluding the drainage area of the Peel River. The climatic indices data were downloaded from the National Oceanic and Atmospheric Administration Earth System Research Laboratory (NOAA) website (<http://www.esrl.noaa.gov>). The remotely sensed snowcover extent data were obtained from the Japanese Satellite Monitoring for Environmental Studies (JASMES) website (<https://kuroshio.eorc.jaxa.jp/JASMES/index.html>). The JASMES SCE products were used in this study because: (1) these products have 5 km spatial resolution with daily time steps; (2) they were obtained from historical optical sensors on polar-orbiting satellites; (3) the *in-situ* measured snow data were used to evaluate the accuracy of snowcover and to correct SCEs; and (4) they used radiance products from both the Advanced Very High-Resolution Radiometer (AVHRR) and the Moderate Resolution Imaging Spectroradiometer (MODIS) products to accurately obtain snowcover extents (Hori *et al.* 2017). The accuracy of the JASMES products was compared against *in-situ* measurements of snow depth and the SCE maps from the National Oceanic and Atmospheric Administration Climate Data Record (NOAA-CDR). The evaluation showed that the JASMES product is more accurate than NOAA products (Hori *et al.* 2017). The JASMES snowcover image is a coded raster. All types of snow, including dry and wet snow and an indication of snow accumulation on ice or

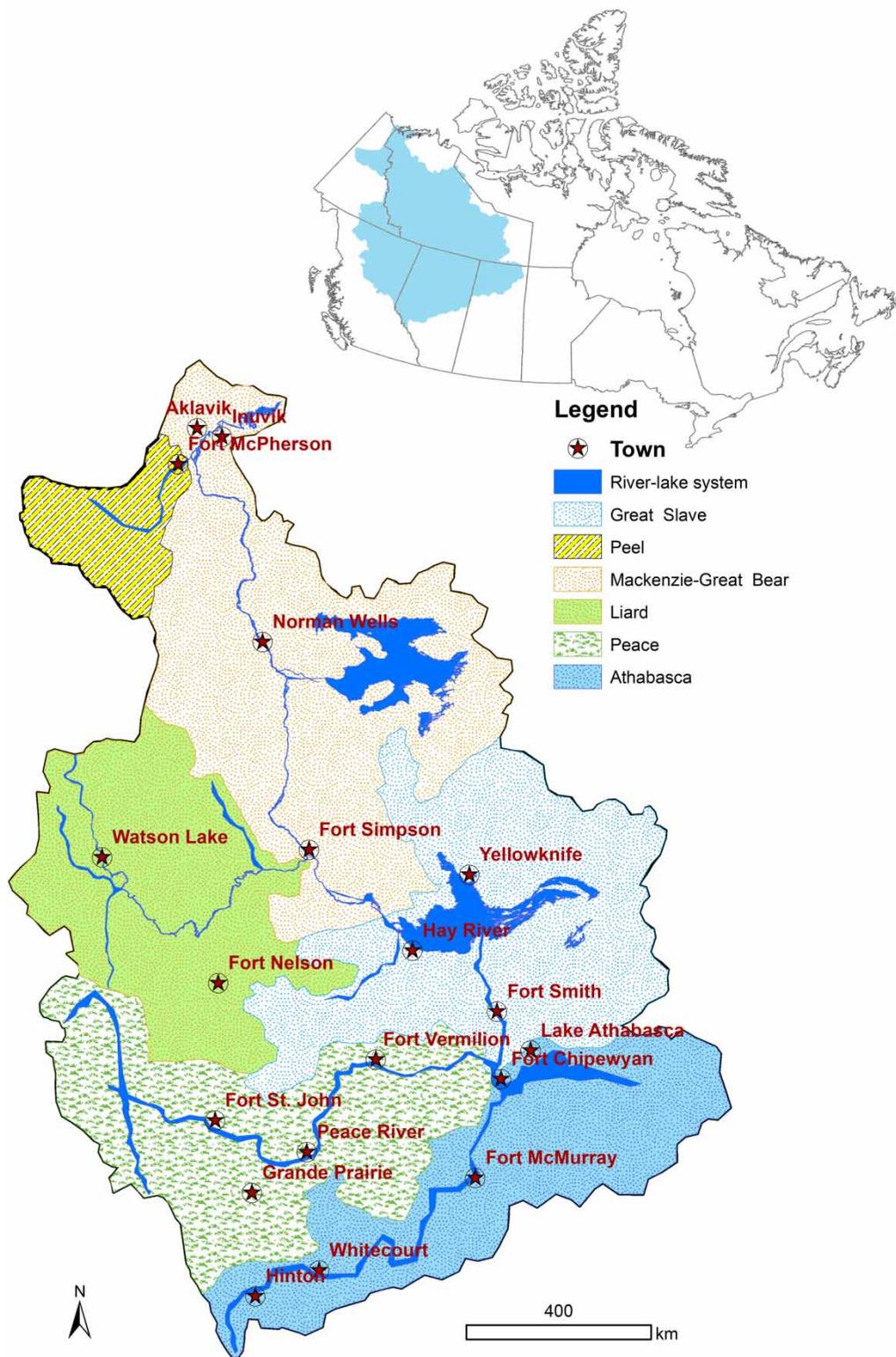


Figure 1 | Mackenzie River Basin and its subbasins located in northern Canada. The hydrometric station that records the Arctic Ocean inflows from this basin is located at the Arctic Red River station (# 10LC014) close to the Beaufort Sea delta.

land, are used to estimate the fraction of SCE for the entire basin. The SCE fraction can be greater than 90% in October and less than 10% in July in the study area. A threshold of 50% SCE fraction is used to separate snow-covered and non-snow-covered grids. Those grids that have SCE fractions greater than 50% are used to estimate the total fraction of the basin covered by snow.

METHODOLOGY

The role of climatic teleconnection patterns and fluctuations of snowcover and streamflows on decadal to sub-seasonal time scale in the predictability of ocean inflows was assessed. The performance of two linear and nonlinear statistical models in representing the interrelationship between snowcover variation modes and inflows from a large-scale watershed was evaluated. For this purpose, daily SCE products with 5 km spatial resolutions obtained from optical remote sensing satellites by Japan Aerospace Exploration Agency were processed and applied to predict Arctic Ocean inflows from the MRB in Canada. To extract space-time patterns of SCE and inflow, singular spectrum analysis (SSA) was conducted. Multi-linear regression and nonlinear Bayesian neural networks were applied to forecast the Arctic Ocean inflows. Daily and monthly snowcover and inflow observations, obtained respectively from remotely sensing products and a gauge at the outlet of the MRB, were used to study spectral information and extract low frequency (decadal) and high frequency (monthly) modes of oscillation in snowcover and the ocean inflow. Four temporal variation modes of SCE and four temporal variation modes of inflows, covering a wide range of variability from multi-decades to months, were obtained and regressed against inflows from the outlet of the basin to forecast inflows one day to seven months in advance.

Linear and nonlinear forecasting methods

To identify the relevant hydro-climatic predictors of the Arctic Ocean inflows from the MRB, two forecasting methods were used at lead times of one to seven days and one to seven months. These models are the multi-linear regression (MLR) and the Bayesian neural network

((BNN) MacKay 1992). The reasons for choosing these two linear and nonlinear models in this study are that we can compare the performance of both linear and nonlinear methods in large-scale streamflow forecasts, and BNN has shown a better performance than other machine learning methods in streamflow forecasting (e.g., Rasouli et al. 2012). Other nonlinear regression models can also be applied, including generalized additive models with smoothing link functions. A multiple model intercomparison, however, was beyond the scope of this study. The information extracted from SCE, climatic teleconnection patterns, and antecedent flows were used to forecast inflows of the Arctic Ocean from MRB. Let $\mathbf{x} = (x_1, \dots, x_m)$ be the vector of m predictor variables in time t and y is the forecast variable of interest in time $t + p$, with $p \in \{1, \dots, 7\}$. Given a dataset of $(y_{t+p}, x_{t1}, \dots, x_{tm})_{t=1}^n$ with n sample size and m predictors, the relationship between y_{t+p} and \mathbf{x}_t is established by the MLR model assuming a linear relationship and through an error variable, ε_t . The error variable follows a normal distribution with mean 0 and variance σ^2 . Hence, the MLR model can be written as:

$$y_{t+p} = \beta_0 + \beta_1 x_{t1} + \dots + \beta_m x_{tm} + \varepsilon_t, \quad \varepsilon_t \sim N(0, \sigma^2) \quad (1)$$

$$t = 1, \dots, n \quad j = 0, \dots, m$$

where β_j and σ^2 are the parameters that are estimated. After estimating these parameters, a stepwise method was used to select the best model with the highest parsimony. For the nonlinear model, we used the Bayesian neural network which is described as a nonlinear parameterized mapping from an input, \mathbf{x} to an output, y through a nonlinear function $g(\mathbf{x}_t; \mathbf{w}, \mathbf{b}, \beta)$. In fact, the relationship between the variable of interest and the predictors are given by:

$$y_{t+p} = g(\mathbf{x}_t; \mathbf{w}, \mathbf{b}, \beta) + \varepsilon_t$$

$$= \sum_{k=1}^s w_k g_k \left(b_k + \sum_{j=1}^m x_{tj} \beta_j^k \right) + \varepsilon_t, \quad t = 1, \dots, n \quad (2)$$

where $\varepsilon_t \sim N(0, \sigma_\varepsilon^2)$ and σ_ε^2 is the variance of the error terms. s is the number of neurons, w_k is the weight of the k -th neuron, b_k is a bias for the k -th neuron, β_j^k is the weight of the j -th input, and g_k is the activation function. In this study, a singular spectrum analysis on SCE and ocean

inflows was conducted to remove or minimize the noises and extract short-term and long-term variability modes of SCE and ocean inflows from MRB. We then used multilinear regression and Bayesian neural network models to determine the major contributors to the variability of the Arctic Ocean freshwater using combined remotely sensed and *in-situ* observations.

The predictors of Arctic ocean inflow

In this study, the forecasted variable of interest is ocean inflow with lead times of one day to seven months, and the predictors' vector is composed of the climatic indices (AO, NAO, AAO, and PNA), and the antecedent inflow and snowcover extent variation modes (Q_t and SCE_t). The antecedent inflow and SCE conditions are characterized by decadal and interannual variation modes, representing the basin memory. The reasons for using antecedent flow and snowcover conditions in this study are that the hydrological cycle and water budget may not be closed at an annual scale, and drivers of water storage variability are required to be studied on decadal to sub-seasonal time scale to better understand variations of the large-scale river flows linked with regional climatic patterns. Since both SCE_t and the Q_t show a seasonality and strong serial dependency (Figure 2), a singular spectrum analysis is used to decompose these variation modes. There are other methods to break down time series into meaningful components in the literature (e.g., wavelet transform (Huang *et al.* 1998) and empirical mode decomposition (Daubechies 1992)). According to Alexandrov *et al.* (2012), however, the SSA method showed the highest performance. Therefore, the SSA method was used in this study. The basic mechanism of SSA is to decompose time series into a sum of interpretable components such as trend, periodic components, and noise. It is based on the principal component analysis of the auto-covariance matrix for lags $\in \{1, \dots, L\}$. According to seasonality shown in Figure 2, for the analysis of daily data, we chose $L = 365$ and for monthly data we chose $L = 12$. The SSA algorithm can be divided into two steps: decomposition and reconstruction. In the first step, the one-dimensional time series were decomposed into multi-dimensional series by adding lagged variables and then creating a so-called trajectory matrix, which is used to calculate the principal components. In the second step, similar

components were regrouped, and disjoint subsets were created to form the final signal decomposition. For more information about SSA, refer to Hsieh (2009). Main SSA components of SCE_t and Q_t instead of the original series were used as inputs to the MLR model as in Equation (1) and the BNN model as in Equation (2).

The model training and validating strategy

The parameters of the MLR and weights and biases of the Bayesian network were adjusted so as to minimize the cost/loss/error function. The method of least squares was used to train the coefficients of the MLR model (Equation (1)). Different activation functions have been suggested in the literature for neural networks. The most widely used functions are linear, sigmoid, and rectified linear unit (ReLU), among others. Yonaba *et al.* (2010) compared a wide range of activation functions for neural networks with Bayesian regularization and found that the hyperbolic (sigmoid) tangent is a more pertinent activation function for daily streamflow forecasting than linear and logistic functions. Parviz & Rasouli (2019) also studied monthly precipitation forecast sensitivity to the choice of activation function and found that the hyperbolic tangent function improved the precipitation forecast skills. The choice of activation function has less important consequences in the forecast accuracy. In this paper, we applied the hyperbolic tangent function that is used in the *brnn* R package, developed by Rodriguez & Gianola (2018), and it is more suitable for streamflow forecasts. This function is given by $g_k(z) = \tanh(z)$. The Bayesian framework (MacKay 1992) was used to estimate the unknown parameters of the BNN model (Equation (2)) using a CRAN package developed by Rodriguez & Gianola (2018). The leave-one-out cross-validation method (Vehtari *et al.* 2017) was applied to the entire study period to train and evaluate the MLR and BNN models to take advantage of all the data and increase the model evaluation period, especially for forecasts with monthly time steps. The leave-one-out cross-validation method includes these steps: first, the entire data period was divided into multiple chunks. Then, one chunk of the entire data was left out at a time and used to validate the model, while the rest was used for training. The validation chunk was selected each time so that all of the divided

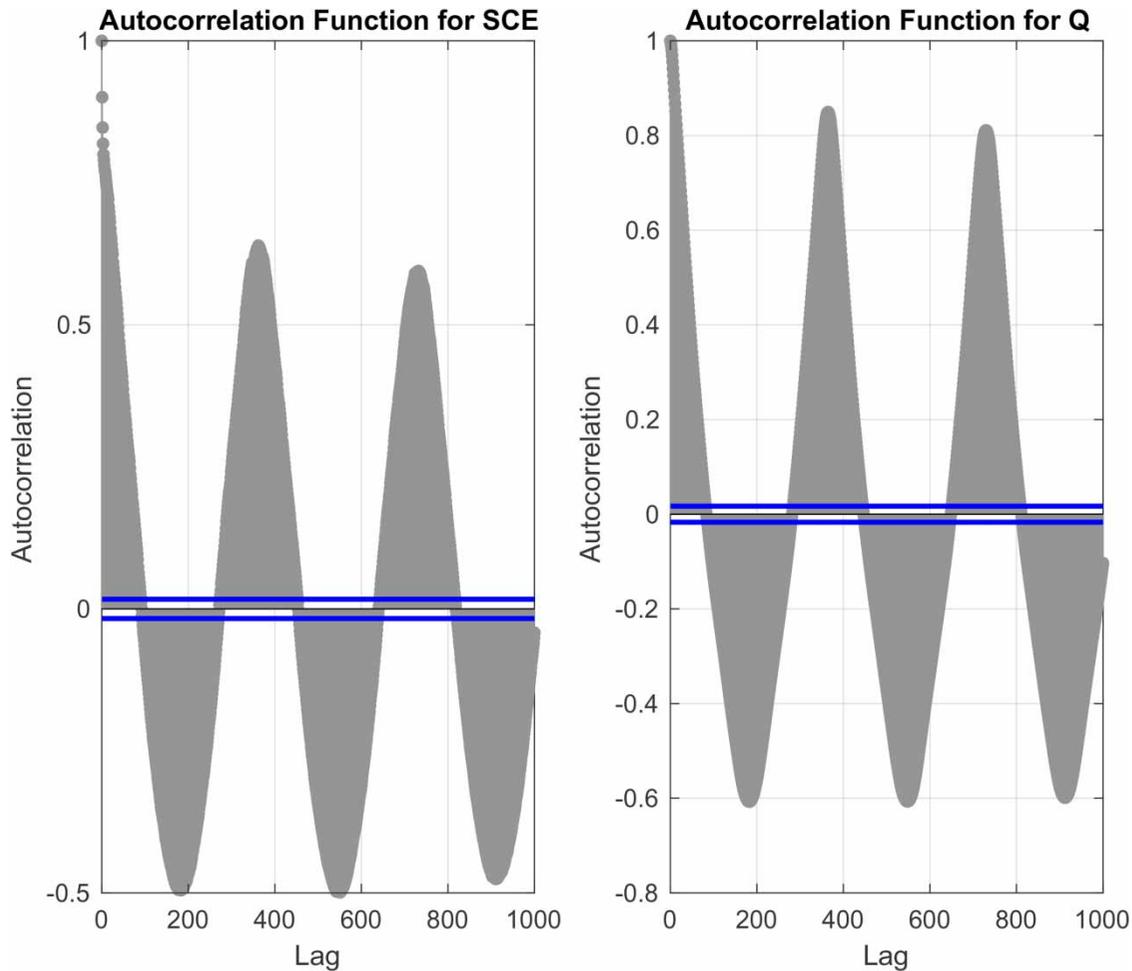


Figure 2 | The autocorrelation functions for snowcover extent and the Arctic Ocean inflows from the Mackenzie River Basin calculated for 1,000-day lag time, illustrating seasonality.

chunks were once used for validating the model. Four different metrics were used to assess the forecasting performance of the daily and the monthly inflows using the linear and nonlinear models: the bias (*BIAS*), the root mean squared error (*RMSE*), the normalized root mean squared error (*NRMSE*), and the Nash–Sutcliffe efficiency (*NSE*).

RESULTS

The linkage between snowcover extent and the Arctic Ocean inflows

The remotely sensed daily SCE products were processed and long-term time series of SCE fractions were obtained for

MRB with a large drainage area. Hydrometric data were used from the records at the Arctic Red River station, the outlet of MRB, to estimate the basin discharge. The SCE fractions range between 0 and 100, where 0 denotes no snow, and 100 shows complete snow coverage of the basin. Figure 3(a) and 3(b) demonstrate the monthly SCE fractions in MRB and the Arctic Ocean inflows from the outlet of this basin. The fraction of the MRB drainage area that is covered by snow in summer (mid-May through July) is very small. SCE gradually increases starting in August and reaches its annual peak in late October. It gradually decreases after February and reaches its minimum in July (Figure 3(a)). The Arctic Ocean inflow from the basin is very low and almost constant in winter (October through January). It increases gradually after February and reaches its peak in July (Figure 3(b)). The interannual variations of

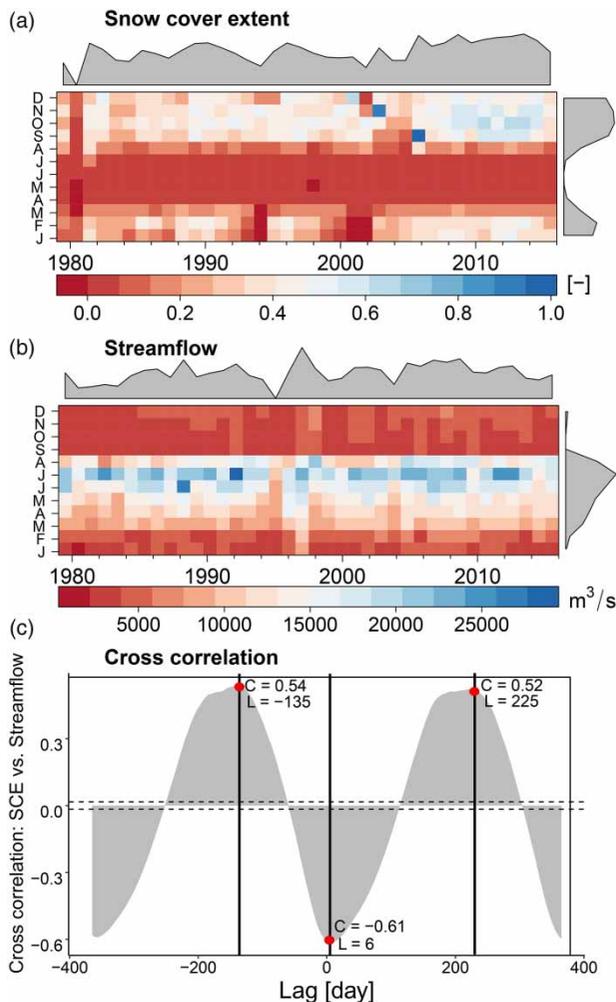


Figure 3 | Multi-decadal and seasonal cycles of (a) remotely sensed snowcover extent and (b) Arctic Ocean inflow from the outlet of the Mackenzie River Basin for the period of 1979–2016, and (c) their correlations. Shaded areas in (a) and (b) show the interannual and monthly variation of snowcover extent and discharge between 1979 and 2016.

both SCE and inflows are very high. Unlike low interannual variability of inflows in winter, flows are highly variable in spring and summer (Figure 3(b)).

The high variability of SCE in spring is partly due to rapid changes in weather systems and rain-on-snow events. Inflows also respond to convective storms in summer months, which can cause a rise in summer flows. A correlation analysis between discharge and SCE (Figure 3(c)) indicates that the decrease of inflows during the snow accumulation period (fall to early spring) and the increase of inflows during snowmelt period (spring and summer)

are closely related to the remotely sensed SCE variations in MRB for the period of 1979–2016. A lag time between inflow and SCE is evidenced almost every year of the study period. Figure 3(c) shows cross-correlation values between inflow and SCE with 365 days lag time ahead of and before January 1. As expected, there is a lag between the time of the maximum SCE and time of the peak inflow associated with the large size of the basin and the high variability of the snowcover in the MRB. The lag time (L) of zero-day shows the correlation coefficient (C) between SCE fractions and inflows on January 1. There is a strong correlation between SCE and inflow. For example, inflow on January 6 ($L = 6$ days) is negatively correlated to SCE on January 1 ($C = -0.61$), while inflow on August 13 ($L = 225$ days) is positively correlated to SCE on January 1 ($C = 0.52$, Figure 3(c)).

Time series of snowcover extent fractions were decomposed into main temporal variation modes, which explain long-term (decadal), medium-range (interannual and seasonal), and short-term (monthly) variations for the period of 1979–2016 (Figure 4(a)). Four modes obtained from SSA were chosen as predictors for the ocean inflow from the basin. Figure 4(b)–4(e) illustrate the modes of the SCE temporal variation, ranging from decades to months. Low-frequency modes (Figure 4(b) and 4(c)), which represent decadal and interannual cycles of the SCE in MRB, show an important fraction of the variance and are found to be significantly different from the noise. This is an important aspect of the snow regime that is neglected if only seasonal snow in an annual hydrological cycle is considered in streamflow forecasts. Figure 4(b) shows the decadal variation of SCE fluctuations that is linked to the climate teleconnection patterns. Figure 4(d) and 4(e) shows seasonal and monthly variations of SCE, which are linked to regional climatic conditions. Four temporal modes of SCE variations explain 98% of the variance of the SCE, and there is a good agreement between original SCE time series and modeled snowcover values reconstructed based on the main decadal, interannual, seasonal, and monthly components. This clearly indicates that SCE has distinct cycles and variation modes, which can be decomposed from the original time series. Extracting the main variability modes from the original snowcover time series helps to understand their linkage to inflow variations. Inflows were also decomposed

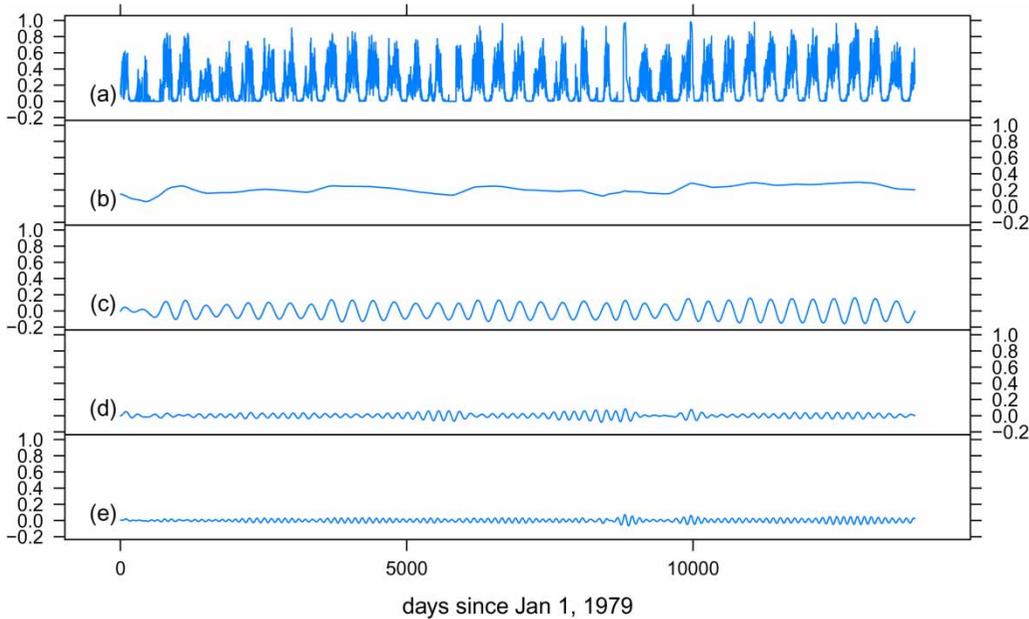


Figure 4 | Original (a), and decadal, (b), interannual (c), seasonal (d), and monthly (e) variation modes of the snowcover extent fractions (*unitless*) in the Mackenzie River Basin obtained from a singular spectrum analysis over the period of 1979–2016.

into main temporal variation modes, similar to snowcover extent for the period of 1979–2016 in MRB (Figure 5(a)). Four modes of inflow obtained from SSA were also chosen as predictors for inflow predictions at short-term (one- to seven-days) and long-term (one- to seven-months) lead times. Figure 5(b)–5(e) illustrate the modes of the inflow variation, ranging from decadal to monthly. Figure 5(b) shows the decadal variation of inflow fluctuations, which has a lower frequency relative to the decadal variation mode of SCE (Figure 5(b)). Interannual (Figure 5(c)) and seasonal (Figure 5(d)) cycles of ocean inflow are similar to those of SCE, except that there is a lag in the timing of the peak snowcover in the basin and maximum inflow.

The results show that the interannual variations of SCE and inflow have a primary role in the fluctuation of daily (Figure 6(a)) and monthly (Figure 6(b)) inflows from MRB to the Arctic Ocean. Daily and monthly inflows are correlated negatively to interannual variation of SCE and positively to interannual variation of inflow. The decadal variation mode of inflow is positively correlated to the decadal variation mode of SCE, while interannual and seasonal variation modes are negatively correlated to interannual and seasonal variation modes of SCE (Figure 6). There is also a

strong and negative correlation between inflow and climate teleconnection patterns such as NAO and PNA (Figure 6), suggesting that large-scale climatic patterns play a secondary role in inflow fluctuations. This confirms that not only the seasonal cycle of snow accumulation and snowmelt affects short-range inflows, but also decadal and interannual fluctuations of SCE play an important role in inflow variations. Therefore, extracting the main temporal modes of variations in SCE and inflow is essential for better characterizing the interrelationship between snow and the Arctic Ocean inflow from large-scale watersheds such as MRB.

Arctic Ocean inflow forecast by linear and nonlinear models

A pool of climatic teleconnection patterns and four principal components of SSA on both SCE fractions and inflows were used to forecast the Arctic Ocean inflow from MRB at one-day to seven-month lead times. To differentiate the role of SCE, climatic teleconnections, and antecedent inflow variations, four sets of predictors were designed: (1) only SCE variation modes, (2) both SCE variation modes and climate teleconnection patterns, (3) both SCE and

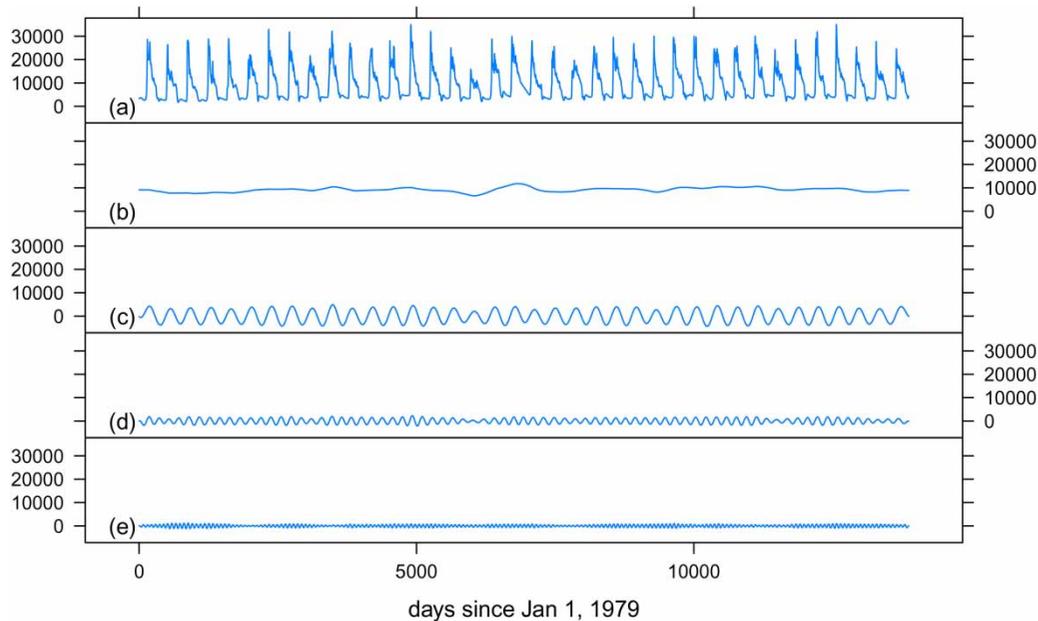


Figure 5 | Original (a), and decadal, (b), interannual (c), seasonal (d), and monthly (e) variation modes of the Arctic Ocean inflows from the Mackenzie River Basin (m^3s^{-1}) obtained from a singular spectrum analysis over the period of 1979–2016.

inflow variation modes, and (4) all SCE and inflow variation modes and climate teleconnection patterns. The leave-one-out cross-validation method was applied to the entire period of 1979–2016 with 500-day subsets for daily forecasts and 18-month subsets for monthly forecasts to train and evaluate the MLR and BNN models. The leave-one-out cross-validation method is especially useful for short monthly data records with a sample size of $n = 456$ in this study. Since the four principal components of SCE and inflow explained over 98% of the total variance, they were selected to compare the role of SCE, antecedent flow conditions, and climate patterns in forecasting ocean inflows from MRB. Reconstructing snowcover from its main variation modes had a minimal modeling error and showed the strength of the SSA method in extracting spectral information from the snowcover data. The stepwise linear regression method (Derksen & Keselman 1992) was used to select the appropriate number of predictors for all four input sets of both daily and monthly forecasting models (Table 1). Despite a low cross-correlation between climate teleconnection patterns (e.g., AO and NAO) as predictands (Figure 6), all of the teleconnections were selected as predictors of inflows

at lead times of one to seven days and one month when they were used in the pool of linear and nonlinear models inputs (Table 1). The climatic teleconnection patterns, however, showed less linkage to monthly inflows at forecasting lead times of two to seven months. All of the decadal, interannual, seasonal, and monthly variation modes of SCE and inflow (Table 1) were selected as predictors of daily inflows, while interannual, seasonal, and monthly (and not decadal) variation modes of SCE were selected as predictors of monthly inflows (Table 1). Strong correlations between the predictors (e.g., decadal variation modes of SCE and inflow in Figure 6) can cause multicollinearity in regression analysis (Farrar & Glauber 1967). A forward and backward stepwise regression method was used to select the relevant predictors for both MLR and BNN models based on the F -test. For daily forecasts of the Arctic Ocean inflow, the correlation coefficients between the predictors (Figure 6(a)) were not statistically significant, and all of the SCE and streamflow variation modes were selected based on the stepwise regression method and the F -test (Table 1). For monthly forecasts, however, the correlation coefficients were significant for some variables, and therefore, the correlated predictors were discarded (Table 1).

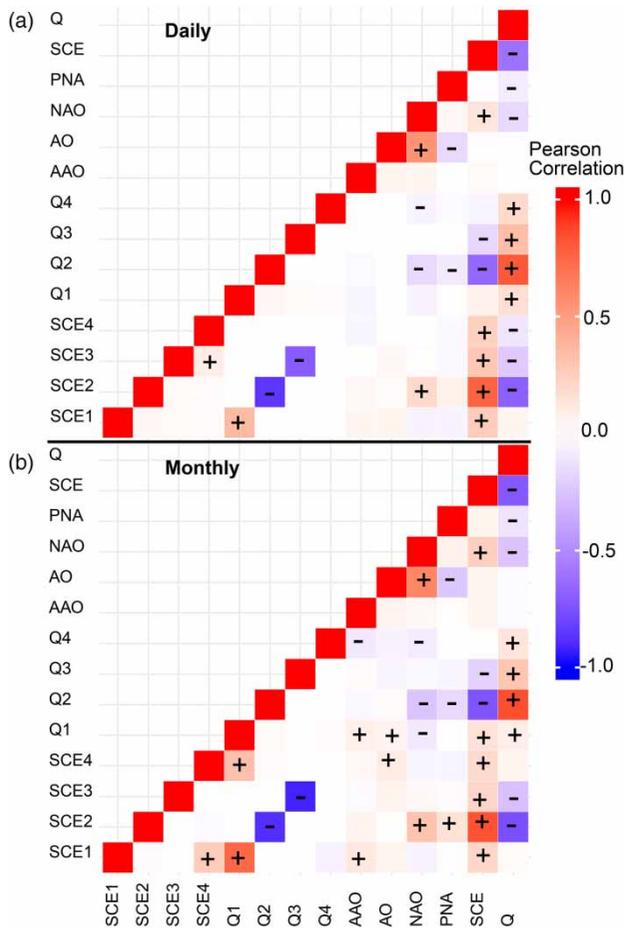


Figure 6 | The Pearson correlation coefficient among (a) daily and (b) monthly large-scale climate teleconnection patterns and four decadal, interannual, seasonal, and monthly variation modes of snowcover extent (SCE1–SCE4), and Arctic Ocean inflows from Mackenzie River Basin (Q1–Q4).

The performance of both linear and nonlinear models was found promising in forecasting daily inflows with a different combination of SCE and inflow variation modes and climatic teleconnections as predictors. Evaluation of the MLR model showed that there is a small bias between predicted and observed inflows (Figure 7), for both linear and non-linear models ($BIAS < 3\%$) at lead times of one to seven days over 1979–2016. The Nash–Sutcliffe score is above 0.8 when all snow, inflow, and climate predictors were used, which showed a relatively good agreement between model outputs and observations. Even though there is an overestimation in high flows, the overall performance of the MLR model is reasonably good, and inflow simulation error is within an acceptable range. The performance of the BNN model for the same period is also shown

in Figure 7. Overall, the performance of the nonlinear model of BNN is similar to the performance of the linear model in both high and low flow periods. Both models are powerful tools to capture not only medium-range inflows but also high flow events in the MRB. Figure 8 summarizes $BIAS$, $RMSE$, $NRMSE$, and prediction skill score of Nash–Sutcliffe for linear and nonlinear models. The $RMSE$ scores were reported on million cubic meters (MCM) of water released to the Beaufort Sea at the mouth of the MRB. Except for a few cases, all of the model performance criteria for BNN are similar to those for the MLR model (Figure 8), which implies that both models performed well in relating the main temporal variation modes of the snowcover extent to streamflow magnitudes in a large-scale watershed.

Biases between forecasted and observed daily inflows are below 9%, and NSE is above 0.45 for both linear and non-linear models used in this study (Figure 8). With only SCE variation modes as daily inflow predictors, all ranges of ocean inflow were predicted reasonably well, except for high flows (Figure 7). Over-estimation of high inflows forecasted by the BNN model degraded the performance scores (Figure 8) when the SCE variation modes were used as predictors for lead times of 2–5 days (Figure 7). The prediction skill scores, however, improved when inflow variation modes were also used as predictors. For instance, NSE for the one-day lead time forecasts increased from 0.45, when only SCE variation modes were used as predictor, to 0.95, when both SCE and inflow variation modes were selected as predictors (Figure 8). The $NRMSE$ score also improved from 75% to 65%. The prediction skill scores substantially improved when climate teleconnections were used as predictors. For instance, NSE for the one-day lead time forecasts increased from 0.45, when only SCE variation modes were used as predictors, to 0.61 when both SCE variation modes and climate teleconnections were used as predictors (Figure 8). The $NRMSE$ score also improved from 75% to 23%. Results showed that the linear model was not capable of improving forecast skill scores when climatic teleconnections were selected as predictors. This is shown in Figure 8 with similar values of the model performance scores for the cases with only SCE variation modes as predictors and both SCE variation modes and climatic teleconnections as predictors.

Table 1 | List of predictor variables, selected by the stepwise regression in each of the four predictor cases: (1) only snowcover extent (SCE), (2) snowcover extent and climate teleconnection patterns, (3) snowcover extent and antecedent inflows – Q, and (4) snowcover extent, climate teleconnection patterns, and antecedent inflows, to force the forecasting models with short-term (one- to seven-day) lead times, denoted by ×, and long-term (one- to seven-month) lead times, denoted by √

Predictor	Time	SCE variability modes				Inflow variability modes				Climate variability indices			
		Decade	Year	Season	Month	Decade	Year	Season	Month	AAO	AO	NAO	PNA
SCE	1	×√	×√	×√	×√								
	2	×	×√	×√	×√								
	3	×	×√	×√	×√								
	4	×	×√	×√	×√								
	5	×	×√	×√	×√								
	6	×√	×√	×√	×√								
	7	×	×√	×√	×√								
SCE + Climate	1	×√	×√	×√	×√					×	×	×	×
	2	×	×	×√	×√					×	×√	×√	×
	3	×	×√	×√	×√					×	×	×	×
	4	×	×√	×√	×√					×	×	×	×
	5	×√	×√	×√	×√					×	×	×	×√
	6	×√	×√	×√	×√					×√	×	×	×√
	7	×	×√	×√	×√					×	×	×√	×
SCE + Q	1	×	×√	×	×√	×√	×√	×√	×√				
	2	×	×√	×	×√	×√	×√	×√	×√				
	3	×√	×√	×√	×√	×√	×√	×√	×√				
	4	×	×√	×√	×√	×√	×√	×√	×√				
	5	×	×√	×	×√	×√	×√	×√	×√				
	6	×	×	×√	×√	×√	×√	×√	×√				
	7	×√	×√	×	×√	×√	×√	×√	×				
SCE + Q + Climate	1	×	×√	×	×√	×√	×√	×√	×√	×√	√	×	√
	2	×	×√	×	×√	×√	×√	×√	×√	×		×	
	3	×√	×√	×√	×√	×√	×√	×√	×√	×		×	
	4	×	×√	×	×√	×√	×√	×√	×√		×√	×√	×
	5	×	×√	×	×√	×√	×√	×√	×√		×	×	×√
	6	×	×	×√	×√	×√	×√	×√	×√	√	×	×√	×
	7	×	×√	×	×√	×√	×√	×√	×		×	×√	×√

The linear and nonlinear models performed well in forecasting monthly inflows with a different combination of SCE and inflow variation modes, and climatic teleconnections as predictors. In monthly forecasting, the BNN model outperformed the MLR model for all of the four combinations of predictors (Figure 9). Similar to daily forecasts of inflow, monthly forecasts improved when variation modes of inflow were used as predictors. The forecast skill score of *NSE* increased gradually from the lead time of one month to five months and degraded after a five-month lead time. It is encouraging that monthly inflows from MRB can be forecasted with high confidence ($NSE > 0.80$, $NRMSE < 30\%$) seven months in advance (Figure 10). This is partly because of a seven-month lag time between peak SCE and peak

inflow (Figure 3(c)), which was included in the predictor set using monthly variation modes of SCE and inflow. The BNN model performed well when the climatic teleconnections were used as predictors in addition to the SCE variation modes to forecast one month in advance (Figure 10). Including the climatic teleconnections in the predictor pool did not improve the forecasts for lead times beyond one month. Differentiating the role of SCE and inflow variation modes, and climatic teleconnections in forecasting both daily and monthly inflows revealed that the multi-decadal and interannual variations in SCE and inflow had a primary role and the climatic teleconnection patterns had a secondary role in the predictability of Arctic Ocean inflows from MRB.

○ Bayesian neural network △ Multi-linear regression

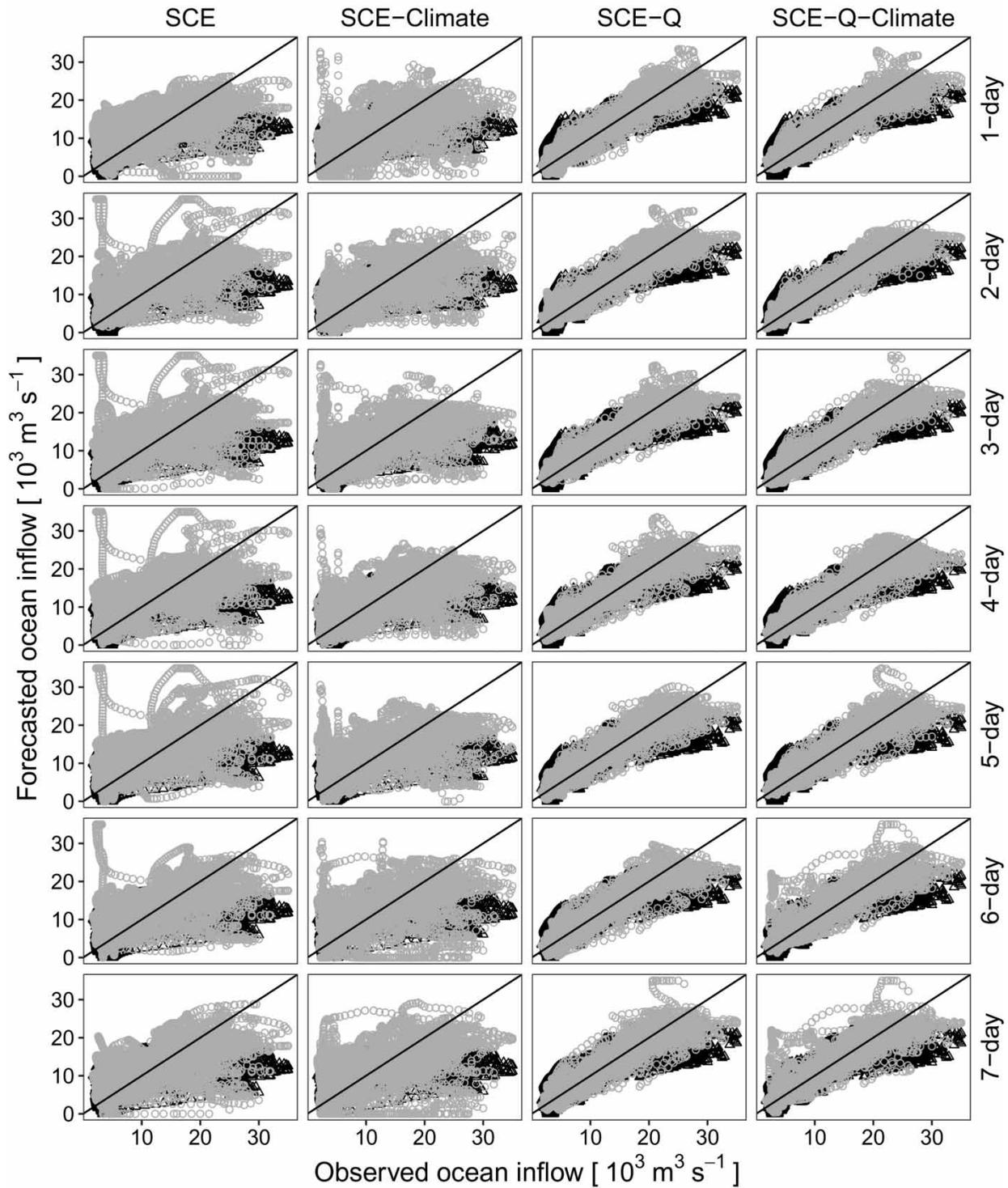


Figure 7 | Performance of the multi-linear regression (MLR, black) and Bayesian neural network (BNN, gray) models in forecasting daily inflows from the outlet of the Mackenzie River Basin with one- to seven-day lead time and under four predictor cases: (1) only snowcover extent, (2) snowcover extent and climate teleconnection patterns, (3) snowcover extent and antecedent inflows, and (4) snowcover extent, climate teleconnection patterns, and antecedent inflows. The black lines show the 1:1 slopes.

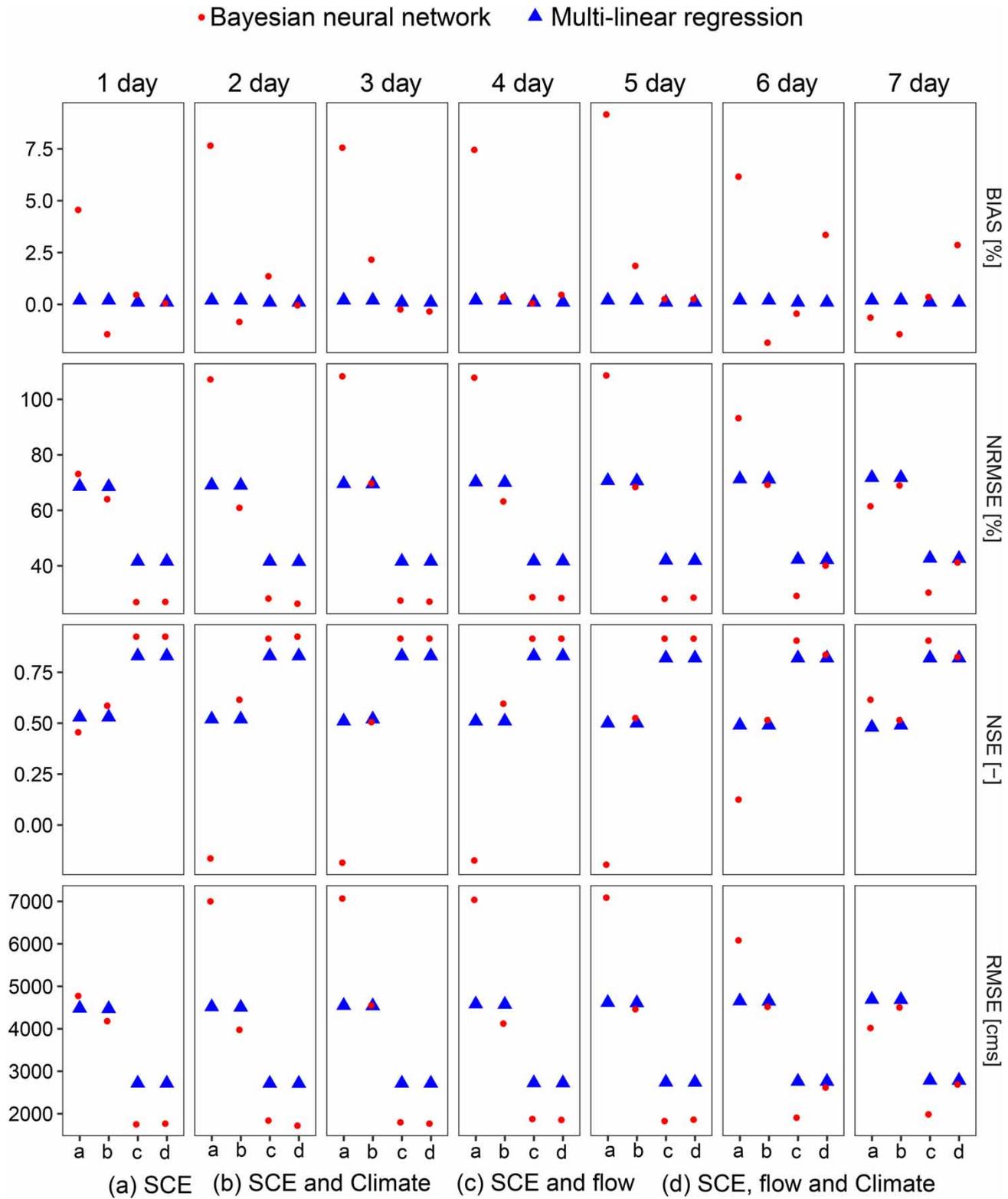


Figure 8 | Metrics for the assessment of the two multi-linear regression and Bayesian neural networks in forecasting the Arctic Ocean inflows from the Mackenzie River Basin, with lead times of one to seven days.

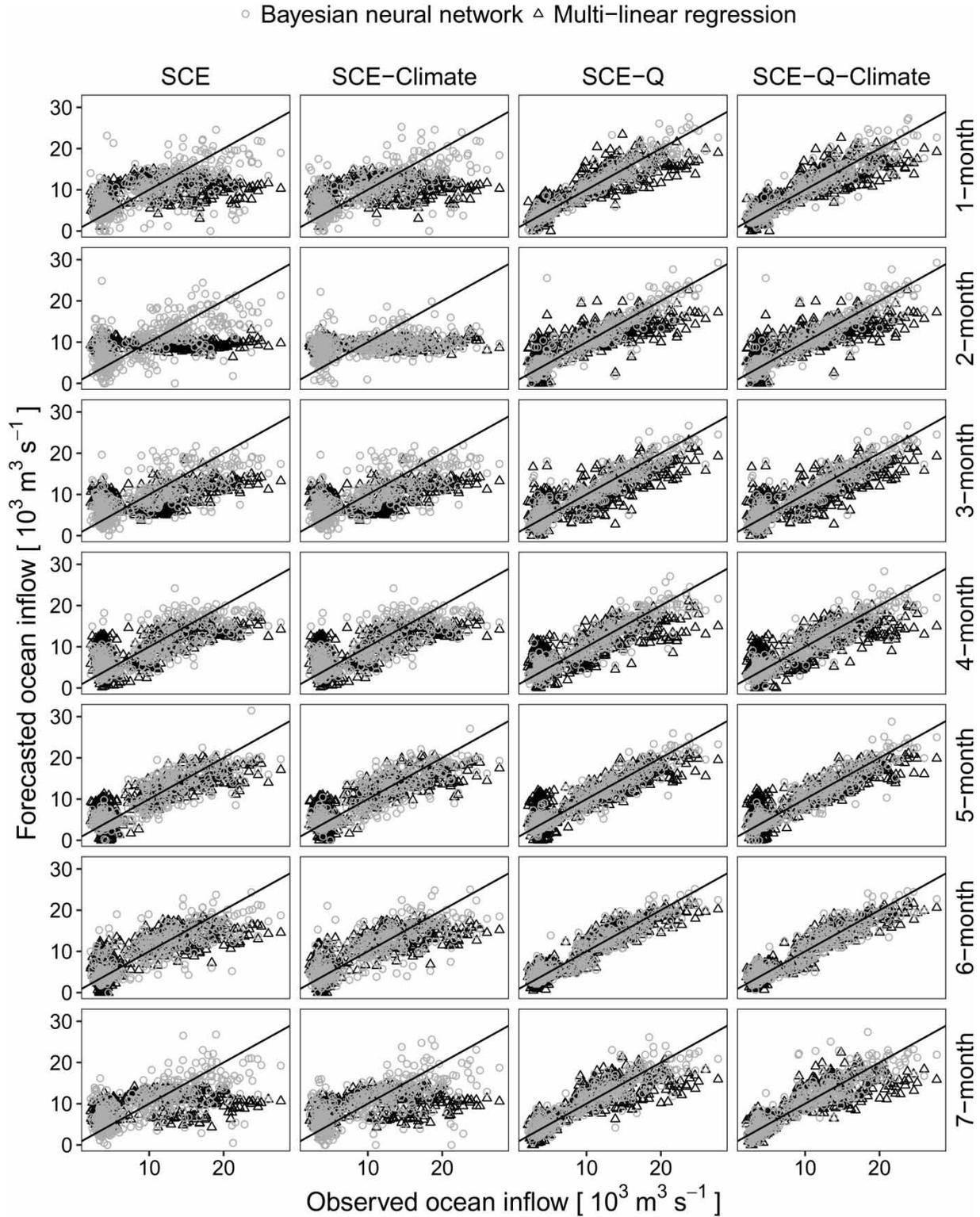


Figure 9 | Performance of the multi-linear regression (black) and Bayesian neural network (gray) models in forecasting monthly inflows from the outlet of the Mackenzie River Basin with one- to seven-month lead time and under four predictor cases: (1) only snowcover extent, (2) snowcover extent and climate teleconnection patterns, (3) snowcover extent and antecedent inflows, and (4) snowcover extent, climate teleconnection patterns, and antecedent inflows. The black lines show the 1:1 slopes.

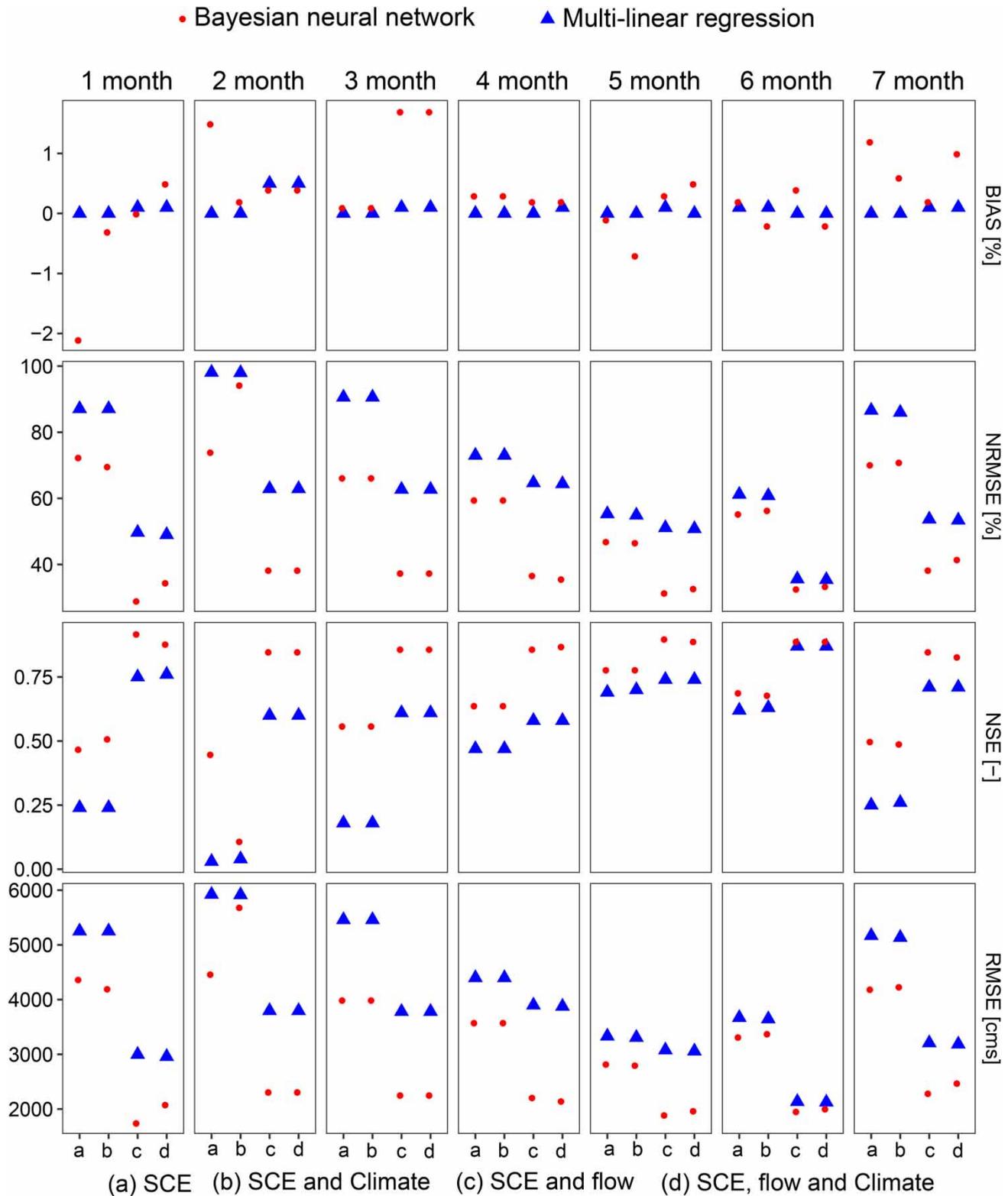


Figure 10 | Metrics for the assessment of the two multi-linear regression and Bayesian neural network in forecasting inflows of the Arctic Ocean from the Mackenzie River Basin, with lead times of one to seven months.

DISCUSSION

The Arctic Ocean inflow from MRB is highly correlated to SCE, and there is a seven-month delay in flow response to snowmelt and SCE recession (Figures 3(c) and 6), suggesting that SCE can be a good proxy to forecast seasonal inflows. The long lag time between snow depletion and inflow rise is because of the large size and complexity of the basin and also long groundwater residence time. The interrelationships between SCE and streamflow were found previously by Zhou *et al.* (2005) in the Upper Rio Grande River Basin, USA, and between snow depth and streamflow in large North American watersheds by Dyer (2008). Tong *et al.* (2010) studied snow distribution and its relationship to the hydroclimatology of MRB. Tong *et al.* (2009) also showed a similarly strong correlation between the 50% SCE and runoff ratio in the snow ablation season in the Quesnel River Basin in British Columbia, which has a small size relative to MRB. There is a negative relationship between the areal size of a snowcovered basin and its hydrograph, and streamflow is usually smoothed as the size of a watershed increases (Zhou *et al.* 2005). Due to the large drainage area of MRB, it is expected that streamflow responds very slowly to the variation of snow covered fractions, obtained from the radiance products of MODIS and AVHRR in this study.

In general, there is a negative correlation between SCE and inflow (Zhou *et al.* 2005). In this study, however, we found that the decadal variation of SCE fluctuations (Figure 4(b)), obtained from the singular spectrum analysis, surprisingly, is not negatively correlated to inflows (Figure 6). This is important as it explains a large portion of the total SCE variance. Linking decadal-monthly variations of SCE fluctuations to ocean inflows can provide a better estimation of large-scale land runoffs. Results showed that interannual variations of both SCE and inflow had the primary control on daily and monthly inflows, and climate and seasonal snow variations had a secondary role in fluctuations of the Arctic Ocean inflow from MRB (Figure 6). A strong correlation between interannual SEC and present-day (month) inflow can be explained by the large topographic gradient, large drainage area, and long residence time of the groundwater flows in MRB (McGuire *et al.* 2005). This supports our hypothesis that headwater runoffs, groundwater contributions, and snow conditions during the preceding years

or even decades can affect inflows at the current day and month. The inflow predictions in large-scale watersheds are very uncertain if only seasonal precipitation and annual hydrological cycle are considered and antecedent moisture conditions are ignored. A positive correlation between decadal variation modes of SCE and inflow and a negative correlation between shorter modes of SCE variation and inflow reveal the important role of snowmelt water storage in the basin decades and years before it releases water to the Arctic Ocean. The snowmelt water is stored in the basin in surface stream tributaries, lakes, soil moisture, and groundwater. Snowmelt is influenced by spatial variability of snowcover and heterogeneous snow redistribution and sublimation. Irradiance reaches its peak in summer, which accelerates the snow depletion and raises streamflow. Tong *et al.* (2009) found that every 1 °C increase in air temperature shifts the 50% SCE fraction ten days forward. Eythorsson *et al.* (2019) found that the number of snow-covered days has decreased by 0.91 days/year over the last two decades. The snowcover duration, which affects summer low flows, depends on precipitation phase and the ratio of snowfall to total solid and liquid precipitation amounts that vary with elevations (Meriö *et al.* 2019). With a high warming rate projected for the northern latitudes (Rasouli *et al.* 2014; Rasouli 2017) in the future, it is expected that snow will melt earlier in late spring and summer in the southern parts of the basin and contribute more to discharge in MRB. Snow, however, is expected to melt more slowly in late winter under warm conditions as the beginning of the snowmelt period moves toward the low irradiance time of year (Rasouli *et al.* 2015b). Seasonal changes in snowpack and precipitation phase change from snowfall to rainfall will result in a large change in the inflow regime under a warm climate.

An assessment of individual roles of antecedent inflow and large-scale SCE variation modes and climate teleconnection patterns showed that the antecedent inflows and SCE variation modes have a primary role and climatic teleconnection patterns have a secondary role in skillful forecasting of large-scale ocean inflows (Figures 7 and 8). In a similar study by Rasouli *et al.* (2012), climatic teleconnection patterns were found to play an important role in flow forecasting for short lead times (e.g., 5–7 days) in a small-scale watershed in British Columbia, Canada. Consistently, the results presented in this research also show that climate teleconnection patterns

affect large-scale river flows, and incorporating relevant climatic conditions in the linear and nonlinear models can improve prediction skills, especially with the BNN model. All four teleconnections of AAO, AO, NAO, and PNA were selected as predictors based on the stepwise linear regression method and the *F*-test (Table 1) and all of them improved the prediction skill scores (Figure 8). Their importance, however, degraded in the monthly forecasts as not all of the teleconnections were selected as monthly inflow predictors (Table 1), showing that daily inflow from large-scale basins such as MRB is strongly linked to climatic teleconnections. This is consistent with the previous studies showing that climatic and atmospheric teleconnection patterns (e.g., NAO) affect the interannual variability of snowcover (Derksen *et al.* 2008; Ge & Gong 2009; Bao *et al.* 2011) but do not significantly affect the snowcover trends in western Canada (Vincent *et al.* 2015; DeBeer *et al.* 2016). Durocher *et al.* (2019) found strong correlations between teleconnections (e.g., AO and NAO) and Arctic Ocean inflows from several northern rivers. Bonsal & Shabbar (2008) reported that low flow events in western Canada are associated with El Niño events and positive phases of the PNA pattern.

Increasing trends in discharge from northern rivers (Syed *et al.* 2007; Rood *et al.* 2017; Durocher *et al.* 2019) are likely linked to increasing trends in the decadal variation mode of SCE in these rivers (Figure 4(b)), and especially there is a strong correlation between decadal variation modes of SCE and inflow (Figure 6). Freshwater inflows of the northern oceans from land rivers are expected to increase by 24–31% by 2080 (Arnell 2005). These changes have the potential to alter the freshwater balance in the Arctic Ocean, which may lead to thermohaline circulation collapse (Arnell 2005) and shift of the climate system to an unbalanced condition. An accurate forecast of river water inflows to the Arctic Ocean helps to better understand the effects of released freshwater to the ocean on physical, chemical, and biological processes (Shiklomanov *et al.* 2000) and sea ice cover (Shimada *et al.* 2006).

CONCLUSION

Major drivers of variation in Arctic Ocean inflows were analyzed in this study using correlation analysis, the

information from singular spectrum analysis, and statistical models. The results show that: (i) there is a high negative correlation between SCE in January and the volume of water released to the Arctic Ocean at the outlet of the basin in August; (ii) there is a strong positive correlation between inflow and decadal cycle of the snowcover and strong negative correlations between inflow and interannual mode of snowcover variation and between inflow and seasonal mode of snowcover variation, which are decomposed by the spectrum analysis; and (iii) the antecedent discharge and SCE conditions, characterized by decadal and interannual variation modes, represent the basin memory, and show a primary role in skillful forecasting of inflows from MRB, while climatic teleconnection patterns and seasonal variation mode of SCE show a secondary role in describing the variability of the ocean inflows. This suggests that the annual hydrological cycle and water budget cannot be closed in the large river basins, and fluctuations of water storage are required to be studied on decadal to sub-seasonal time scale to better understand short-term and long-term variations of the large-scale river flows in connection with regional climatic patterns.

The results show that the performance of the Bayesian nonlinear model in inflow prediction was similar to the multi-linear regression model, with the nonlinear model performing better in forecasting monthly inflows. The simulation skill scores of Nash–Sutcliffe efficiency for both models are 0.83 for seven-day forecasts and 0.80 for seven-month forecasts. The significance of the proposed method is that it can identify major drivers of Arctic Ocean inflow variability and can link large-scale teleconnections to local-scale atmospheric and cryospheric conditions. Linking local-scale seasonal snowcover and large-scale decadal climate variations to inflows from the large river basins has important implications in predicting the freshwater input to the Arctic Ocean, estimating rates of freshwater transport from the Arctic Ocean to the North Atlantic, and better understanding of climate variability and predictability in the near future in regional and continental scales. The proposed methodology can be highly beneficial in the qualitative and quantitative assessment of low frequency flows such as groundwater and subsurface flows in large-scale watersheds and in better understanding surface and subsurface waters in the snow-dominated regions.

ACKNOWLEDGEMENTS

The authors gratefully acknowledge the financial support from the Natural Sciences and Engineering Research Council of Canada (NSERC) through Post-doctoral Fellowship to KR, NSF funded ND EPSCoR (NSF grant #IIA-135466) to THM and from the Fonds de recherche du Québec-Nature et technologies and the Canadian Statistical Sciences Institute through Post-doctoral Fellowship to BRN. The authors also thank the Earth Observation Research Center of JAXA for providing them with the JASMES snowcover extent data. The authors of this paper, hereby, confirm that they do not have any conflict of interest to declare. Datasets related to this article can be found from the website of the Water Survey of Canada (<https://wateroffice.ec.gc.ca>) for streamflow, National Oceanic and Atmospheric Administration Earth System Research Laboratory (NOAA) website (<http://www.esrl.noaa.gov>) for climate teleconnections, and Japanese Satellite Monitoring for Environmental Studies (JASMES) website (<https://kuroshio.eorc.jaxa.jp/JASMES/index.html>) for snowcover extent data.

REFERENCES

- Aagaard, K. & Carmack, E. C. 1989 The role of sea ice and other fresh water in the Arctic circulation. *Journal of Geophysical Research: Oceans* **94** (C10), 14485–14498.
- Alexandrov, T., Bianconcini, S., Dagum, E. B., Maass, P. & McElroy, T. S. 2012 A review of some modern approaches to the problem of trend extraction. *Econometric Reviews* **31** (6), 593–624.
- Arnell, N. W. 2005 Implications of climate change for freshwater inflows to the Arctic Ocean. *Journal of Geophysical Research: Atmospheres* **110** (D7), 1–9.
- Bao, Z., Kelly, R. & Wu, R. 2011 Variability of regional snow cover in spring over western Canada and its relationship to temperature and circulation anomalies. *International Journal of Climatology* **31** (9), 1280–1294.
- Bojariu, R. & Gimeno, L. 2003 The role of snow cover fluctuations in multiannual NAO persistence. *Geophysical Research Letters* **30** (4), 1–4.
- Bonsal, B. & Shabbar, A. 2008 Impacts of large-scale circulation variability on low streamflows over Canada: a review. *Canadian Water Resources Journal* **33** (2), 137–154.
- Brown, R. D. & Goodison, B. E. 1996 Interannual variability in reconstructed Canadian snow cover, 1915–1992. *Journal of Climate* **9** (6), 1299–1318.
- Burn, D. H., Abdul Aziz, O. I. & Pietroniro, A. 2004 A comparison of trends in hydrological variables for two watersheds in the Mackenzie river basin. *Canadian Water Resources Journal/Revue Canadienne des Ressources Hydriques* **29** (4), 283–298.
- Cohen, J. & Entekhabi, D. 1999 Eurasian snow cover variability and northern hemisphere climate predictability. *Geophysical Research Letters* **26** (3), 345–348.
- Coulibaly, P. & Burn, D. H. 2004 Wavelet analysis of variability in annual Canadian streamflows. *Water Resources Research* **40** (3), W03105, 1–14.
- Daubechies, I. 1992 *Ten Lectures on Wavelets (CBMS-NSF Regional Conference Series in Applied Mathematics, 61)*. SIAM, Philadelphia, PA, USA.
- DeBeer, C. M., Wheeler, H. S., Carey, S. K. & Chun, K. P. 2016 Recent climatic, cryospheric, and hydrological changes over the interior of western Canada: a review and synthesis. *Hydrology and Earth System Sciences* **20** (4), 1573.
- Derksen, S. & Keselman, H. J. 1992 Backward, forward and stepwise automated subset selection algorithms: frequency of obtaining authentic and noise variables. *British Journal of Mathematical and Statistical Psychology* **45** (2), 265–282.
- Derksen, C., Brown, R. & MacKay, M. 2008 Mackenzie basin snow cover: Variability and trends from conventional data, satellite remote sensing, and Canadian regional climate model simulations. In: *Cold Region Atmospheric and Hydrologic Studies. The Mackenzie GEWEX Experience* (M. Woo, ed.). Springer, Berlin, Heidelberg, Germany, pp. 213–239.
- Durocher, M., Requena, A. I., Burn, D. H. & Pellerin, J. 2019 Analysis of trends in annual streamflow to the Arctic ocean. *Hydrological Processes* **33** (7), 1143–1151.
- Dyer, J. 2008 Snow depth and streamflow relationships in large North American watersheds. *Journal of Geophysical Research: Atmospheres* **113** (D18), D18113, 1–12.
- Ek, N. & Ritchie, H. 1996 Forecasts of hydrological parameters over the Mackenzie River Basin: sensitivity to initial conditions, horizontal resolution and forecast range. *Atmosphere-Ocean* **34** (4), 675–710.
- Eythorsson, D., Gardarsson, S. M., Ahmad, S. K., Hossain, F. & Nijssen, B. 2019 Arctic climate and snow cover trends—Comparing global circulation models with remote sensing observations. *International Journal of Applied Earth Observation and Geoinformation* **80**, 71–81.
- Farrar, D. E. & Glauber, R. R. 1967 Multicollinearity in regression analysis: the problem revisited. *The Review of Economic and Statistics* **49** (1), 92–107.
- Fichot, C. G., Kaiser, K., Hooker, S. B., Amon, R. M. W., Babin, M., Belanger, S., Walker, S. A. & Benner, R. 2013 Pan-Arctic distributions of continental runoff in the Arctic Ocean. *Scientific Reports* **3** (1), 1053, 1–6.
- Folland, C. K., Rayner, N. A., Brown, S., Smith, T., Shen, S., Parker, D., Macadam, I., Jones, P., Jones, R. N., Nicholls, N. & Sexton, D. M. H. 2001 Global temperature change and its

- uncertainties since 1861. *Geophysical Research Letters* **28** (13), 2621–2624.
- Ge, Y. & Gong, G. 2009 North American snow depth and climate teleconnection patterns. *Journal of Climate* **22** (2), 217–233.
- Gutzler, D. S. & Rosen, R. D. 1992 Interannual variability of wintertime snow cover across the northern hemisphere. *Journal of Climate* **5** (12), 1441–1447.
- Hori, M., Sugiura, K., Kobayashi, K., Aoki, T., Tanikawa, T., Kuchiki, K., Niwano, M. & Enomoto, H. 2017 A 38-year (1978–2015) Northern Hemisphere daily snow cover extent product derived using consistent objective criteria from satellite-borne optical sensors. *Remote Sensing of Environment* **191**, 402–418.
- Hsieh, W. W. 2009 *Machine Learning Methods in the Environmental Sciences*. Cambridge University Press, Cambridge, UK.
- Huang, N. E., Shen, Z., Long, S. R., Wu, M. C., Shih, H. H., Zheng, Q., Yen, N.-C., Tung, C. C. & Liu, H. H. 1998 The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. *Proceedings of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences* **454** (1971), 903–995.
- Hurrell, J. W. 1995 Decadal trends in the North Atlantic Oscillation: regional temperatures and precipitation. *Science* **269** (5224), 676–679.
- Johannessen, O. M., Kuzmina, S. I., Bobylev, L. P. & Miles, M. W. 2016 Surface air temperature variability and trends in the Arctic: new amplification assessment and regionalisation. *Tellus A: Dynamic Meteorology and Oceanography* **68** (1), 28234.
- Liston, G. E. 1999 Interrelationships among snow distribution, snowmelt, and snow cover depletion: implications for atmospheric, hydrologic, and ecologic modeling. *Journal of Applied Meteorology* **38** (10), 1474–1487.
- Macdonald, R., Carmack, E., McLaughlin, F., Falkner, K. & Swift, J. 1999 Connections among ice, runoff and atmospheric forcing in the Beaufort Gyre. *Geophysical Research Letters* **26** (15), 2223–2226.
- MacKay, D. J. 1992 Bayesian interpolation. *Neural Computation* **4** (3), 415–447.
- Mackay, J. R. & MacKay, D. 1974 Snow cover and ground temperatures, Garry Island, NWT. *Arctic* **27** (4), 287–296.
- McGuire, K., McDonnell, J. J., Weiler, M., Kendall, C., McGlynn, B., Welker, J. & Seibert, J. 2005 The role of topography on catchment-scale water residence time. *Water Resources Research* **41** (5), W05002, 1–14.
- Meriö, L. J., Ala-aho, P., Linjama, J., Hjort, J., Kløve, B. & Marttila, H. 2019 Snow to precipitation ratio controls catchment storage and summer flows in boreal headwater catchments. *Water Resources Research* **55**, 4096–4109. <https://doi.org/10.1029/2018WR023031>.
- Padilla, A., Rasouli, K. & Déry, S. J. 2015 Impacts of variability and trends in runoff and water temperature on salmon migration in the Fraser river basin, Canada. *Hydrological Sciences Journal* **60** (3), 523–533.
- Park, H., Fedorov, A. N., Zheleznyak, M. N., Konstantinov, P. Y. & Walsh, J. E. 2015 Effect of snow cover on pan-Arctic permafrost thermal regimes. *Climate Dynamics* **44** (9–10), 2873–2895.
- Parviz, L. & Rasouli, K. 2019 Development of precipitation forecast model based on artificial intelligence and subseasonal clustering. *Journal of Hydrologic Engineering* **24** (12), 1–13. [https://doi.org/10.1061/\(ASCE\)HE.1943-5584.0001862](https://doi.org/10.1061/(ASCE)HE.1943-5584.0001862).
- Peterson, B. J. 2006 Trajectory shifts in the Arctic and subarctic freshwater cycle. *Science* **313** (5790), 1061–1066.
- Polyakov, I. V., Alexeev, V. A., Belchansky, G. I., Dmitrenko, I. A., Ivanov, V. V., Kirillov, S. A., Korablev, A. A., Steele, M., Timokhov, L. A. & Yashayev, I. 2008 Arctic Ocean freshwater changes over the past 100 years and their causes. *Journal of Climate* **21** (2), 364–384.
- Pomeroy, J. & Gray, D. 1995 Snowcover accumulation, relocation and management. *Bulletin of the International Society of Soil Science* **88** (2), 1–135.
- Proshutinsky, A., Bourke, R. & McLaughlin, F. 2002 The role of the Beaufort Gyre in Arctic climate variability: seasonal to decadal climate scales. *Geophysical Research Letters* **29** (23), 15–11.
- Rachold, V., Eicken, H., Gordeev, V. V., Grigoriev, M. N., Hubberten, H.-W., Lisitzin, A. P., Shevchenko, V. P. & Schirmeister, L. 2004 Modern terrigenous organic carbon input to the Arctic Ocean. In: *The Organic Carbon Cycle in the Arctic Ocean* (R. Stein & R. W. Macdonald, eds.). Springer, Berlin, Heidelberg, Germany, pp. 33–55.
- Rasouli, K. 2010 *Short Lead-Time Streamflow Forecasting by Machine Learning Methods, with Climate Variability Incorporated*. MSc thesis, University of British Columbia, Vancouver, Canada.
- Rasouli, K. 2017 *Sensitivity Analysis of Mountain Hydrology to Changing Climate*. PhD dissertation, University of Saskatchewan, Saskatoon, Canada.
- Rasouli, K., Hsieh, W. W. & Cannon, A. J. 2012 Daily streamflow forecasting by machine learning methods with weather and climate inputs. *Journal of Hydrology* **414–415**, 284–293.
- Rasouli, K., Hernández-Henrquez, M. A. & Déry, S. J. 2013 Streamflow input to Lake Athabasca, Canada. *Hydrology and Earth System Sciences* **17** (5), 1681–1691.
- Rasouli, K., Pomeroy, J. W., Janowicz, J. R., Carey, S. K. & Williams, T. J. 2014 Hydrological sensitivity of a northern mountain basin to climate change. *Hydrological Processes* **28** (14), 4191–4208.
- Rasouli, K., Hernández-Henrquez, M. A. & Déry, S. J. 2015a Reply to D. L. Peters comment on streamflow input to Lake Athabasca, Canada by Rasouli et al. (2013). *Hydrology and Earth System Sciences* **19** (3), 1287–1292.
- Rasouli, K., Pomeroy, J. W. & Marks, D. G. 2015b Snowpack sensitivity to perturbed climate in a cool mid-latitude mountain catchment. *Hydrological Processes* **29** (18), 3925–3940.
- Rodriguez, P. P. & Gianola, D. 2018 *Bayesian Regularization for Feed-Forward Neural Networks*. Available from: <https://cran.r-project.org/web/packages/brnn/index.html>

- Rood, S. B., Kaluthota, S., Philipsen, L. J., Rood, N. J. & Zanewich, K. P. 2017 [Increasing discharge from the Mackenzie River system to the Arctic Ocean](#). *Hydrological Processes* **31** (1), 150–160.
- Serreze, M. C., Barrett, A. P., Slater, A. G., Woodgate, R. A., Aagaard, K., Lammers, R. B., Steele, M., Moritz, R., Meredith, M. & Lee, C. M. 2006 [The large-scale freshwater cycle of the Arctic](#). *Journal of Geophysical Research: Oceans* **111** (C11), C11010, 1–19.
- Shiklomanov, I. A., Shiklomanov, A. I., Lammers, R. B., Peterson, B. & Vorosmarty, C. J. 2000 The dynamics of river water inflow to the Arctic Ocean. In: *The Freshwater Budget of the Arctic Ocean* (E. L. Lewis, ed.). Kluwer Academic, Dordrecht, The Netherlands, pp. 281–296.
- Shimada, K., Kamoshida, T., Itoh, M., Nishino, S., Carmack, E., McLaughlin, F., Zimmermann, S. & Proshutinsky, A. 2006 [Pacific Ocean inflow: influence on catastrophic reduction of sea ice cover in the Arctic Ocean](#). *Geophysical Research Letters* **33** (8), L08605, 1–4.
- Spence, C., Saso, P. & Rausch, J. 2007 [Quantifying the impact of hydrometric network reductions on regional streamflow prediction in northern Canada](#). *Canadian Water Resources Journal* **32** (1), 1–20.
- Syed, Z. F., Aggarwal, P., Goodall, C., Niu, X. & El-Sheimy, N. 2007 [A new multi-position calibration method for MEMS inertial navigation systems](#). *Measurement Science and Technology* **18** (7), 1897–1907.
- Temimi, M., Leconte, R., Brissette, F. & Chaouch, N. 2004 Near real time flood monitoring over the Mackenzie River basin using passive microwave data. In: *IGARSS 2004. 2004 IEEE International Geoscience and Remote Sensing Symposium*, 20–24 September 2004, Anchorage, AK, USA. IEEE, Vol. 3, pp. 1862–1865.
- Tong, J., Déry, S. J. & Jackson, P. L. 2009 [Interrelationships between MODIS/terra remotely sensed snow cover and the hydrometeorology of the Quesnel River Basin, British Columbia, Canada](#). *Hydrology and Earth System Sciences* **13** (8), 1439–1452.
- Tong, J., Déry, S. J., Jackson, P. L. & Derksen, C. 2010 [Snow distribution from SSM/i and its relationships to the hydroclimatology of the Mackenzie River Basin, Canada](#). *Advances in Water Resources* **33** (6), 667–677.
- Vehtari, A., Gelman, A. & Gabry, J. 2017 [Practical Bayesian model evaluation using leave-one-out cross-validation and WAIC](#). *Statistics and Computing* **27** (5), 1413–1432.
- Vincent, L. A., Zhang, X., Brown, R. D., Feng, Y., Mekis, E., Milewska, E. J., Wan, H. & Wang, X. L. 2015 [Observed trends in Canada's climate and influence of low-frequency variability modes](#). *Journal of Climate* **28** (11), 4545–4560.
- Wang, J. Y., Whitfield, P. H. & Cannon, A. J. 2006 [Influence of Pacific climate patterns on low-flows in British Columbia and Yukon, Canada](#). *Canadian Water Resources Journal* **31** (1), 25–40.
- Yang, D., Marsh, P. & Ge, S. 2014 [Heat flux calculations for Mackenzie and Yukon Rivers](#). *Polar Science* **8** (3), 232–241.
- Yang, Q., Dixon, T. H., Myers, P. G., Bonin, J., Chambers, D., van den Broeke, M. R., Ribergaard, M. H. & Mortensen, J. 2016 [Recent increases in Arctic freshwater flux affects Labrador Sea convection and Atlantic overturning circulation](#). *Nature Communications* **7** (1), 1–7.
- Yonaba, H., Anctil, F. & Fortin, V. 2010 [Comparing sigmoid transfer functions for neural network multistep ahead streamflow forecasting](#). *Journal of Hydrologic Engineering* **15** (4), 275–283.
- Zhou, X., Xie, H. & Hendrickx, J. M. 2005 [Statistical evaluation of remotely sensed snow-cover products with constraints from streamflow and SNOTEL measurements](#). *Remote Sensing of Environment* **94** (2), 214–231.

First received 19 October 2019; accepted in revised form 7 February 2020. Available online 27 March 2020