Machine learning models coupled with empirical mode decomposition for simulating monthly and yearly streamflows: a case study of three watersheds in Ontario, Canada

Peiman Parisouj, Changhyun Jun, Sayed M. Bateni, Essam Heggy & Shahab S. Band

Published online: 21 Aug 2023.

Submit your article to this journal ⬀

Article views: 254

View related articles ⬀

View Crossmark data ⬀

Taylor & Francis
Taylor & Francis Group

Machine learning models coupled with empirical mode decomposition for simulating monthly and yearly streamflows: a case study of three watersheds in Ontario, Canada

Peiman Parisouj[a,b]*, Changhyun Jun[a,c]*, Sayed M. Bateni[b]*, Essam Heggy [d,e]* and Shahab S. Band[f]*

[a]Department of Smart Cities, Chung-Ang University, Seoul, Korea; [b]Department of Civil and Environmental Engineering, and Water Resources Research Center, University of Hawaii at Manoa, Honolulu, HI, USA; [c]Department of Civil and Environmental Engineering, Chung-Ang University, Seoul, Korea; [d]Viterbi School of Engineering, University of Southern California, Los Angeles, CA, USA; [e]Jet Propulsion Laboratory, California Institute of Technology, Pasadena, CA, USA; [f]Department of Information Management, International Graduate School of Artificial Intelligence, National Yunlin University of Science and Technology, Douliou, Taiwan

ABSTRACT

This paper presents a novel approach for enhancing long-term runoff simulations through the integration of empirical mode decomposition (EMD) with four machine learning (ML) models: ensemble, support vector machine (SVM), convolutional neural networks (CNN), and artificial neural networks with backpropagation (ANN-BP). The proposed methodology uses EMD to decompose precipitation and temperature time-series into intrinsic mode functions, thereby revealing underlying data patterns. Subsequently, these components are incorporated into the ML models to simulate the runoff time-series. The effectiveness of the hybrid models is evaluated using streamflow runoff data obtained from the Grand, Winnipeg, and Moosonee Rivers in Ontario, Canada. Four widely used performance indices, namely, correlation coefficient, root mean square error (RMSE), mean absolute relative error, and Nash–Sutcliffe efficiency, are employed to assess the models' performance. The results demonstrate that the hybrid EMD-ML models exhibit significantly superior performance compared with the standalone ML methods. During the validation phase, the EMD-Ensemble, EMD-SVM, EMD-CNN, and EMD-ANN-BP models exhibit notable reductions in the RMSEs of monthly streamflow estimates for the Grand River, amounting to 11%, 22%, 8%, and 33%, respectively, compared with their non-EMD counterparts. Additionally, these hybrid models exhibit improved RMSEs for yearly simulations in the Winnipeg River, with reductions of 54%, 0.08%, 6%, and 4.5% respectively. To further enhance the accuracy of monthly and yearly streamflow estimates, an SVM-recursive feature elimination technique is employed to select a more appropriate EMD dataset in all study cases. This research underscores the potential of integrating EMD with ML models to enhance long-term runoff simulations. The outcomes highlight the superior performance of the hybrid EMD-ML models, demonstrating their ability in generating lower biases than the standalone ML methods. These findings hold significant implications for the field of computational fluid mechanics and can contribute to the understanding of hydrological processes.

Introduction

A key task in hydrological modelling studies is to accurately simulate the runoff in a basin over a long-term period, for example, monthly and yearly time-series, which affects the water supply efficiency, flood management, and water resource management (Mohammadi, 2021; Soltani et al., 2021). Several methods have been applied to simulate the runoff, including conceptual models, machine learning (ML) models, and coupled models (Kratzert et al., 2018). In particular, several researchers have developed time-series statistical models

such as moving average (MA), autoregressive (AR), and/or autoregressive moving average for runoff time-series simulation (He et al., 2019; Weeks & Boughton, 1987). Recently, ML and deep learning (DL) models have been applied for runoff modelling (Hu et al., 2018; Mallick et al., 2022; Parisouj et al., 2020, 2022).

ML is a broad field that covers artificial intelligence, probability, psychology, and statistics, among other domains. ML methods can be used to simplify and solve problems (Nasteski, 2017). Recently, DL methods, as a subset of ML methods, have received considerable

attention in hydrology and community studies (Barzegar et al., 2021; Goliatt et al., 2021; Nasteski, 2017). Artificial neural networks (ANNs) are widely used in hydrology for runoff simulation. However, their learning process is slow, and they are prone to being trapped in local minima (Parisouj et al., 2020). To address this limitation, Cortes and Vapnik (1995) proposed the support vector machine (SVM) technique to efficiently find global optimum solutions (Meng et al., 2019). Several researchers have highlighted the superiority of support vector regression (SVR) over artificial neural network with backpropagation (ANN-BP) for simulating runoff (Bafitlhile & Li, 2019; Kalteh, 2013). DL architectures, such as long short-term memory (LSTM), deep neural networks, and convolutional neural networks (CNNs) are being widely used to predict the time-series of wind, solar, and streamflow (Ghimire et al., 2021). Among these, the CNN algorithm lacks post-processing capability (Liu et al., 2022).

Other promising ML models include ensemble models such as extra tree regressor, bagging regressor, adaptive boosting regression (AdaBoost), and stack generalisation. Ensemble models combine multiple models, and the output of one model is input to another model to increase the prediction accuracy (Sagi & Rokach, 2018). Recently, ensemble models have been applied for predicting rainfall runoff (Barrera-Animas et al., 2022; Jose et al., 2022; Zhao et al., 2022). Tarfaya et al. (2022) used ensemble models for predicting the index rainfall and showed that the extra tree model yielded reasonable results. Liu et al. (2014) applied AdaBoost to enhance the accuracy of runoff prediction. Elbeltagi et al. (2022) predicted the river flow rate in the Moines watershed by applying ML models, and the bagging model was noted to exhibit acceptable performance.

The prediction accuracy of ML models can be enhanced by extracting trends and harmonics from hydrological time-series and removing noise through appropriate data preprocessing techniques, such as genetic algorithm optimisation, MA, principal component analysis, singular spectrum analysis, wavelet analysis, and gamma testing (Band et al., 2021; Bartoletti et al., 2018; Cui et al., 2021; Golshan et al., 2020).

Huang et al. (1998) proposed the empirical mode decomposition (EMD) technique for noise-assisted data analysis. Several researchers have applied EMD to extract signals from noisy nonstationary data in the analysis of several aspects, such as hydroclimatic processes, solar radiation, wind speed, speaker recognition, and ice-snow coverage (Lee & Ouarda, 2012; Metzger et al., 2020; Prasad et al., 2019; Sánchez-Martínez et al., 2022; Zhang et al., 2018). The EMD technique is self-adaptive, which makes it preferable to other traditional approaches (Tayyab et al., 2018). Temperature, precipitation, and runoff represent nonlinear and nonstationary time-series (Chen et al., 2018). Therefore, the EMD approach can be used to analyse these hydrological time-series data.

The objective of this study was to use the novel hybrid EMD-Ensemble method to estimate runoff at monthly and yearly time scales. To the best of our knowledge, none of the existing studies have used the combined EMD-Ensemble approach to estimate monthly and yearly streamflows. The hybrid EMD-Ensemble was compared with hybrid EMD-SVR, EMD-CNN, and EMD-ANN-BP. In addition, the performance of hybrid models was compared with the standalone ensemble, SVR, CNN, and ANN-BP models. Three monthly and annual runoff time-series from the Grand, Winnipeg, and Moosonee Rivers in Canada were investigated to ensure the applicability of the proposed framework. These three rivers were chosen because of differences in their drainage area (watershed), landcover, discharge, watershed topography, and climate.

Method and materials

Study area and data description

The study areas included the Grand, Winnipeg, and Moosonee River basins in Ontario, Canada, which have different sizes, geology, and hydro climatology. Figure 1 shows the locations of the three basins in Ontario.

The Grand River watershed, with a drainage area of $6965\,km^2$ and length of 280 km, is located in southern Ontario. Bahamonde et al. (2015) reported that 76% and 17% of the watershed pertain to agricultural land and forest areas, respectively. Thirty treated effluents from municipal wastewater are dumped into the Grand River basin. The elevation difference in the regions upstream and downstream of the Grand River basin is approximately 350 m. The northwest region receives more rain than the southeast, with annual precipitation levels averaging 850 mm with a peak of 1,000 mm. January and February are the driest months, and July and August are the wettest months. Additionally, the mean annual temperature varies from 5°C in the higher elevations in the north to 8°C along the lakeside (Krause et al., 2001). Great Lakes, the Arctic region, and the Gulf of Mexico affect the climate of the Grand River basin.

Most of the Winnipeg River watershed, with a drainage area of approximately $150,000\,km^2$, is located in the northwestern Ontario province, with part of it lying it the southeastern Manitoba province. The basin's main branch passes predominantly through forest areas between the northern United States and northern Canada. The Winnipeg River flows to the west, through Manitoba province, into lake Winnipeg. The elevation
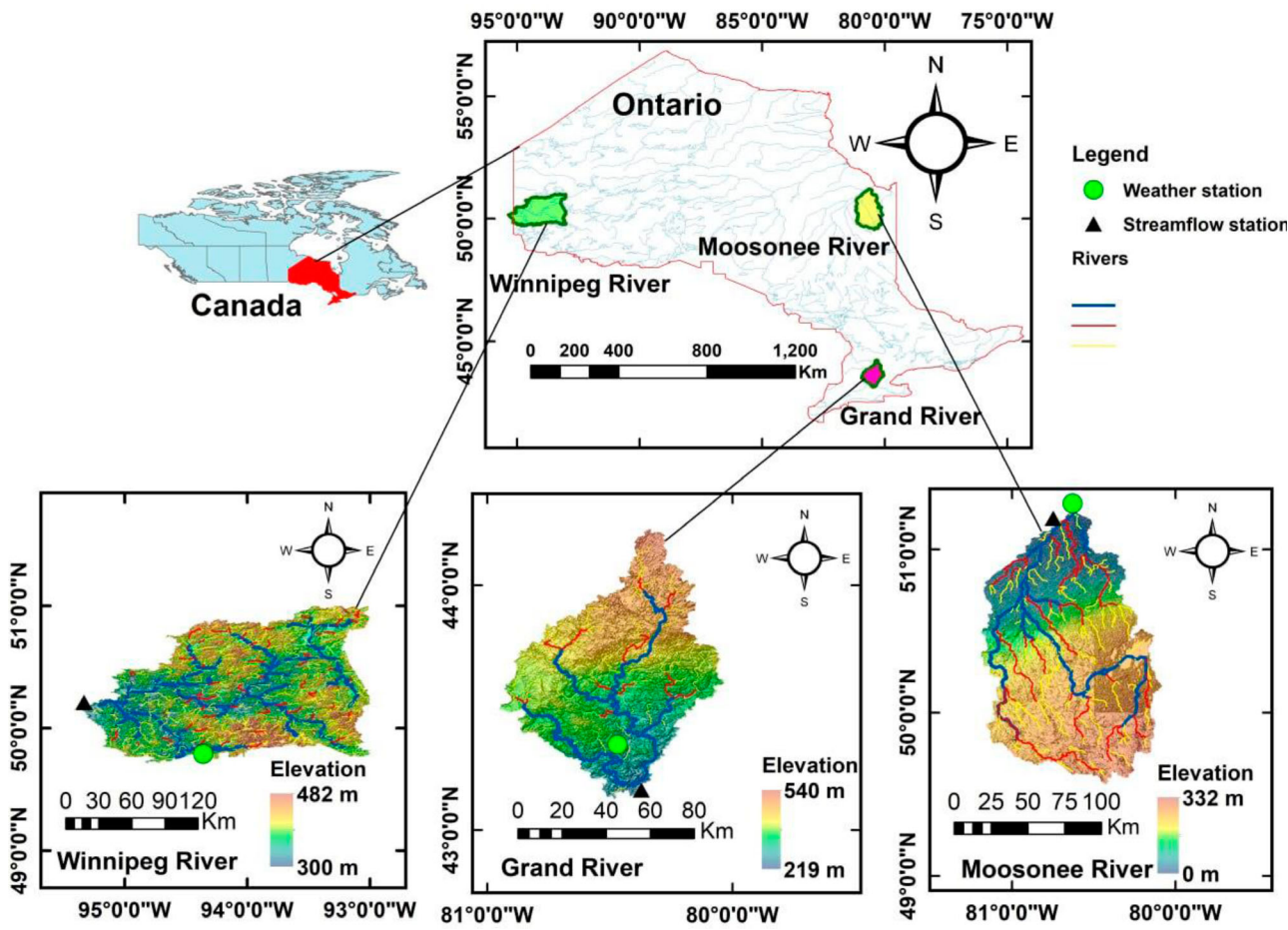
Figure 1. Study area.

difference in the regions upstream and downstream of the Winnipeg River basin is approximately 217 m. Approximately 30% of the precipitation is snowfall, with annual precipitation averaging 780 mm. Similar to the Grand River basin, January and February are the driest months, and June and July are the wettest months. The annual mean temperature varies between −18°C and 25°C at Slave Falls (St. George, 2007).

The Moosonee River watershed, with a drainage area of approximately 109,000 km², is located in northeast Ontario, southwest of the James Bay region. The watershed consists of three main tributaries, including the Missinaibi and Mattagami Rivers that constitute the Moose and Abitibi Rivers. The basin is divided into northern and southern portions. This study focused on the southern portion, which is more topographically diverse. The elevation difference in the regions upstream and downstream of the Moosonee River basin is approximately 580 m. The mean annual precipitation varies from 650 to 1,000 mm, and the mean annual temperature is between −9°C and 8°C. The correlation between precipitation and temperature is typically positive, but the wettest period is July–December. Approximately 35%

of the mean annual precipitation is snowfall, and more than half of the annual Moosonee River watershed discharge occurs in spring (Ho et al., 2005; Story & Buttle, 2001).

The daily discharge data from January 1, 1973, to December 31, 2020, was selected from the Water Survey of Canada for three measurement stations including the watersheds (https://wateroffice.ec.gc.ca/). A daily dataset including precipitation (P), maximum temperature (T_{max}), and minimum temperature (T_{min}) was selected, consistent with the discharge period from the nearest weather stations. Records of weather observations are available at https://climate.weather.gc.ca/. Figure 1 shows the locations and information of the streamflow and weather stations.

Methodology

EMD model

Huang et al. (1998) developed the EMD algorithm for nonlinear and nonstationary datasets. The core of this algorithm is that most raw time-series have multiple frequencies of different scales (Karthikeyan & Nagesh
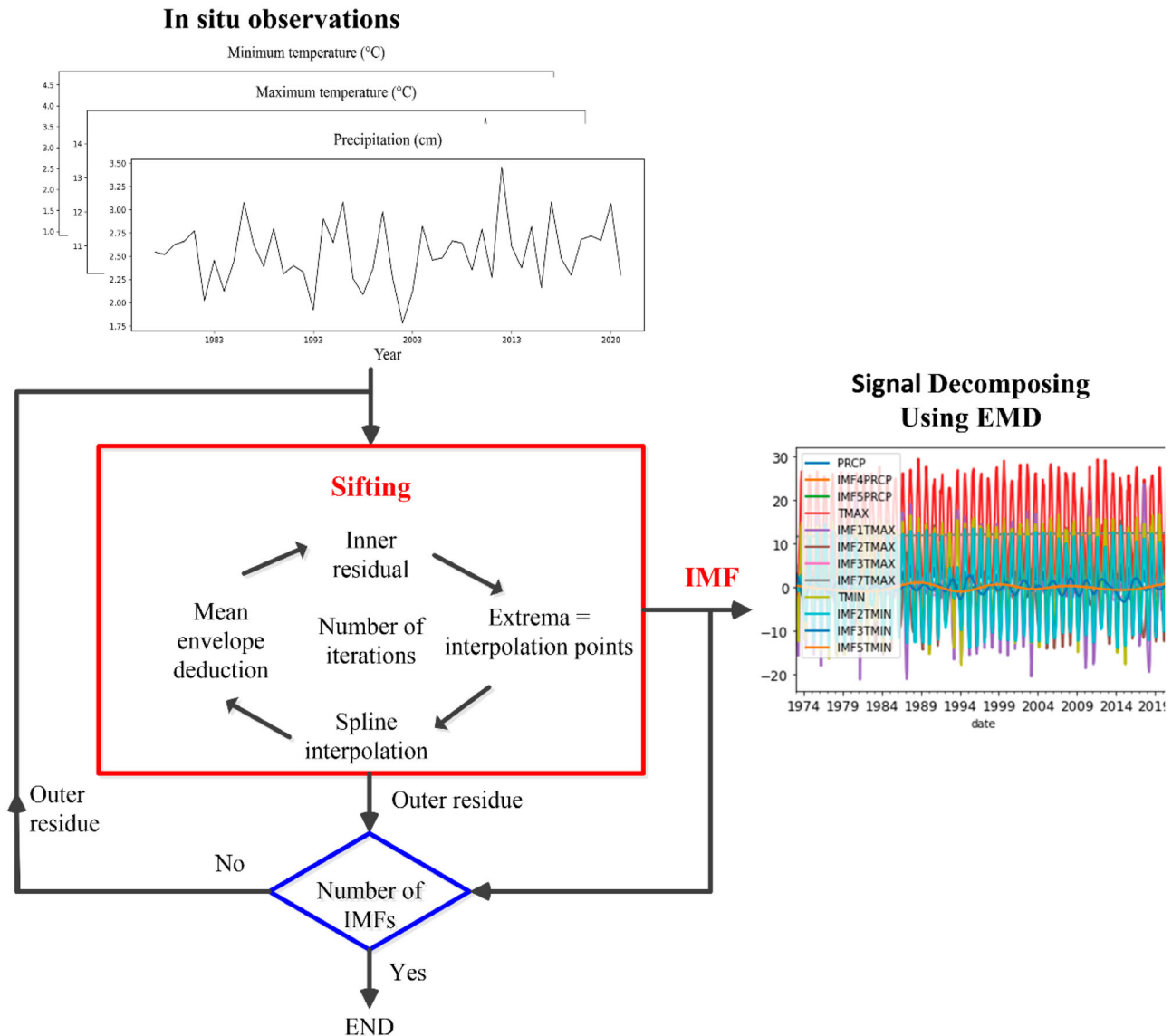
Figure 2. Process flow of empirical mode decomposition (EMD).

Kumar, 2013). The algorithm defines the decomposition dataset to a group of bands by several intrinsic mode functions (IMFs) and a residual function that is calculated for all IMFs and remains constant. Owing to its simplicity, EMD has been widely applied in hydrology (Karthikeyan & Nagesh Kumar, 2013). Kamath and Senapati (2021) coupled the EMD model with ANN to predict 24-h wind speed. The results revealed that EMD-ANN outperformed ANN. Sibtain et al. (2021) predicted the runoff and explored the effect of EMD on the ANN-BP model. The EMD-ANN-BP was noted to outperform the standalone ANN-BP. Yuan et al. (2021) incorporated ensemble EMD (EEMD) in an LSTM to predict daily runoff and proved that the use of the EEMD model output as input data to the LSTM model enhanced the model performance. In general, EMD intermittently extracts the scales of the IFM time-series. The IMFs must satisfy two requirements:

(1) The absolute value of the total local minima and local maxima minus zero-crossing equals zero or one.
(2) The mean between the upper and lower envelopes in each time-series point equals zero.

To satisfy these requirements, the IMF must be generated by making the function smooth with respect to zero. The EMD breaks down time-series into IMFs through 'sifting'. Chu and Huang (2020)_ENREF_11 demonstrated that EMD can satisfactorily treat nonstationary time-series in hydrological analyses. The process flow of the EMD algorithm is illustrated in Figure 2.

SVR model

SVR, a ML model introduced by Cortes and Vapnik (1995), has been widely applied in hydrology (Achite et al., 2022; Kolachian & Saghafian, 2021; Mozaffari et al., 2022; Sun et al., 2021). Unlike methods such as ANN, which implement empirical risk minimisation, SVR implements the concept of structural risk minimisation. The SVR algorithm for regression computation can be expressed as follows:

$$f(x) = a. k(x) + b \tag{1}$$

where a represents the weight vectors or coefficients, $k(x)$ is the kernel function, and b is a bias term. In this study, the radial basis function (RBF) kernel is used to solve Equation (1).

$$\text{maximize} \begin{cases} -\dfrac{1}{2} \displaystyle\sum_{i,j=1}^{l} (\alpha_i - \alpha_i^*)(\alpha_j - \alpha_j^*) k(x_i - x_j) \\[2em] +\displaystyle\sum_{i,j=1}^{l} y_i(\alpha_i - \alpha_i^*) \end{cases} \tag{2}$$

$$\text{subject to} \begin{cases} \displaystyle\sum_{i=1}^{l} (\alpha_i - \alpha_i^*) = 0 \\[1.5em] \displaystyle\sum_{i=1}^{l} (\alpha_i - \alpha_i^*) \leq Cvl \\[1.5em] \alpha_i, \alpha_j \quad \in [0, C] \end{cases} \tag{3}$$

where l is the sample size, α and α^* are Lagrange multipliers, C is the cost of the kernel function, y_i is the output, and $k(x_i - x_j)$ is the kernel function. Equation (2) must satisfy the Karush–Kuhn–Tucker requirements to map the dataset, which can be defined as follows:

$$\begin{cases} \alpha_i^*(f(x_i) - y_i - \eta - \omega_i^*) = 0 \\ \alpha_i(y_i - f(x_i) - \eta - \omega_i) = 0 \\ \alpha_i^*\alpha_i = 0; \omega_i^*\omega_i = 0 \\ (C - \alpha_i^*)\omega_i^* = 0; (C - \alpha_i)\omega_i = 0 \end{cases} \tag{4}$$

where η, ω_i^*, and ω_i are slack variables. Finally, the SVR can be solved using the following set of equations:

$$\sum_{i,j=1}^{l} (\alpha_i - \alpha_i^*) \cdot k(x_i - x_j) + b, \tag{5}$$

$$b = y_i + \eta - \sum_{i,j=1}^{l} (\alpha_i - \alpha_i^*) \cdot k(x_i - x_j) + b \tag{6}$$

CNN model

ML research focuses on DL modelled on the human brain. ANNs are representative computational systems in this context. Neural networks in DL must train computers to have the same functionality as that of the neural system in the human brain. LeCun et al. (1998) designed CNNs, which are DL models that have been widely used in classification and regression tasks in several fields, especially hydrology (Hussain et al., 2020; Sadeghi et al., 2019; Tu et al., 2021). Unlike traditional neural networks, CNNs incorporate multiple architectures such as pooling, local connections, and shared weights. CNNs work on the principle that the input dataset consists of images or data that can be represented as images. Consequently, the processing time and number of parameters are reduced. A CNN typically includes convolutional layers, pooling layers, and fully connected layers. Convolutional layers, as key components, include filters known as kernels that apply convolutional functions to the input dataset and prepare pixels for the next process. Pooling layers help the CNN model control overfitting, thereby limiting the required computation and parameters by reducing the representation size in convolutions (Tu et al., 2021). The LeakyRelu function can accelerate the convergence of the CNN model and facilitate the learning of the neuron weights, even if the input includes zero values. Several structures have been developed based on the type of input data and research objectives, such as InceptionV3 (Szegedy et al., 2016), VGG16 (Simonyan & Zisserman, 2014), ResNet50 (He et al., 2016), Xception (Chollet, 2017), and InceptionResNetV2 (Szegedy et al., 2017), and the corresponding layers, learning parameters, and training process have been elucidated.

ANN-BP model

ANN-BP is a three-layer traditional ANN with feedforward that uses BP in the training dataset (Parisouj et al., 2020, 2022). ML algorithms can be divided into supervised and unsupervised methods. The ANN-BP is a supervised learning method in which data passes through the input layers. The weights minimise the error through hidden layers and add biases to the calculation. The output is generated by the output layer (Sudheer et al., 2002). BP is implemented by assigning a gradient of the loss function to each weight in the chain rule. The gradient is computed one layer at a time, iterating backward from the last layer to avoid superfluous intermediate term calculations. Stochastic gradient descent is a representative learning algorithm that uses BP to compute a gradient.

Ensemble model

Ensemble learning is an ML algorithm that uses a group of base learners to assess and solve real-world issues. Meta-learning pertains to learning from base learners, and ensemble ML (EML) techniques are meta-learning methods (Sagi & Rokach, 2018; Tyralis et al., 2021; Zhang & Ma, 2012) that merge two or more models to enhance the generalisation and performance.

AdaBoost: AdaBoost is an effective ensemble technique that adapts a sequence of base learners to a larger dataset using a more recent dataset. To reproduce the final prediction, base learner predictions are merged using a weighted summation (Idris et al., 2012). The training set is enhanced by the addition of weights ω_1, $\omega_2, \ldots, \omega_N$ in each boosting iteration. The initial boosting iteration uses the same weights and data. Subsequently, the learner algorithm is applied to the newly weighted data. In subsequent rounds, the weights of the incorrectly (correctly) predicted training data are increased (decreased). Eventually, each poor learner is forced to focus on the samples missed by the previous learners (Liu et al., 2014).

Bagging regressor: The bagging regressor is a bootstrap aggregation-based ensemble meta-estimator. There exist m bootstrap copies of a sample data point drawn with replacement, and the base learner is used for each bootstrap sample. Finally, the results of each base learner are averaged or voted on. In certain cases, the base learner is a regression or a classification algorithm. Aggregation helps reduce the variance of an individual base learner (Breiman, 1996; Meddage et al., 2021; Singh et al., 2022). Notably, research on runoff prediction using the bagging regressor is limited at present.

Extra tree regressor: Geurts et al. (2006) presented the extra tree model as a decision tree based on an ML algorithm. The model builds a group of decision trees that do not have to be pruned (the trees grow in a top-to-down configuration). The advantages of the extra tree model can be summarised as follows: (1) can easily avoid overfitting, (2) is robust to noise, and (3) can efficiently handle high-dimensional data without feature selection. Similar to other tree-based ensemble approaches, the extra tree regressor produces a collection of decision trees but emphasises randomisation to reduce variance without increasing bias (Eslami et al., 2020; Geurts et al., 2006). The extra tree regressor technique generates random split nodes, allowing it to be implemented faster than other decision-tree-based approaches. To prevent any subsequent increase in bias, three parameters are imported: (1) number of randomly selected attributes at each node (random state), (2) minimum number of samples required to split an internal node (min samples split), and (3) number of trees in the forest (n estimators).

Stack generalisation: Wolpert (1992) introduced the stacking generalisation method, in which several models are assembled to develop an efficient meta-learner. This model takes advantage of singular models to enhance the generalisation. Several models are used as estimators, and one model is used as the final estimator. The results of estimators are used as the input of the final estimator. In this manner, the stacking generalisation method can enhance the performance of the final estimator model. To avoid overfitting, the meta-model does not directly learn the outputs of the base models. The model can be mathematically expressed as follows:

$$\hat{y}(x) = \sum_{i=1}^{m} \omega_i h_i(x) \tag{7}$$

where ω_i is the weight determined for each base learner, and $h_i(x)$ is the model prediction.

The optimal final prediction is defined by minimising the set of stacking weights using mean square linear regression. Equation (8) mathematically represents the least-squares regression:

$$\begin{cases} \omega^* = argmin_\omega \sum_{j=1}^{n} (y(x_j) - \sum \omega_i h_i^{(-i)}(x_j))^2 \\ \omega_i \geq 0 \\ \sum_{i=1}^{M} \omega_i = 1 \end{cases} \tag{8}$$

where n is number of samples, $y(x_j)$ denotes the observed values, $h_i^{(-i)}(x_j)$ represents the output or prediction of the ith base learner for the jth data point, and $\omega^* = (\omega_1, \omega_2, \ldots, \omega_m)$ is the set of weights assigned to the base learners.

Model development and input

The ability of four methods (ANN-BP, SVR, CNN, and ensemble stacking model) in simulating the monthly and yearly runoff in the three selected basins was evaluated. To enhance the performance, the IMFs of P, $T_{\min}$, and $T_{\max}$ were applied as input variables. Different numbers of IMFs were set for the monthly and yearly simulations. The optimal group of variables of the input dataset was selected by applying two feature algorithms: SVM-recursive feature elimination (SVM-RFE) and random forest-Boruta (RF-Boruta) feature selection algorithm for all three basins and models. For a more comprehensive

Table 1. (a) Variables selected for each method for the Grand River basin. (b) Variables selected for each method for the Winnipeg River basin. (c) Variables selected for each method for the Moosonee River basin.

Time scale	Method	Feature selection	Variables
(a)			
Monthly	SVR-RFE	Main variables and IMFs	$P, P'4, P'5, T_{max}, T_{max}'1, T_{max}'2, T_{max}'3, T_{max}'7, T_{min}, T_{min}'2, T_{min}'3, T_{min}'5$
		IMFs	$P'1, P'2, P'3, P'4, P'5, T_{max}'1, T_{max}'2, T_{max}'3, T_{min}'2, T_{min}'3$
	RF-Boruta	Main variables and IMFs	$P, T_{max}, T_{max}'5, T_{min}$
		IMFs	$P'1, P'2, P'4, P'5, T_{min}'2$
Yearly	SVR-RFE	Main variables and IMFs	$P, P'1, T_{max}, T_{max}'1, T_{max}'2, T_{min}'1, T_{min}'4$
		IMFs	$P'1, P'2, P'4, T_{max}'1, T_{min}'3, T_{min}'4$
	RF-Boruta	Main variables and IMFs	$P, P'1, T_{max}, T_{max}'1$
		IMFs	$P'1, T_{max}'1$
(b)			
Monthly	SVR-RFE	Main variables and IMFs	$P'2, P'4, P'7, P'8, T_{max}, T_{max}'2, T_{max}'3, T_{max}'4, T_{min}, T_{min}'1, T_{min}'6, T_{min}'7$
		IMFs	$P'3, P'4, P'7, P'8, T_{max}'2, T_{max}'3, T_{max}'4, T_{min}'3, T_{min}'6, T_{min}'7$
	RF-Boruta	Main variables and IMFs	$P'3, P'4, P'5, P'6, P'7, P'8, T_{max}'1, T_{max}'2, T_{max}'4, T_{max}'5, T_{max}'6, T_{max}'7, T_{min}, T_{min}'1, T_{min}'2, T_{min}'3, T_{min}'4, T_{min}'5, T_{min}'7$
		IMFs	$P'3, P'4, P'5, P'6, P'7, P'8, T_{max}'1, T_{max}'2, T_{max}'4, T_{max}'5, T_{max}'6, T_{max}'7, T_{min}'1, T_{min}'2, T_{min}'3, T_{min}'4, T_{min}'5, T_{min}'7$
Yearly	SVR-RFE	Main variables and IMFs	$P, P'2, P'3, T_{max}, T_{max}'1, T_{min}, T_{min}'1$
		IMFs	$P'2, P'3, T_{max}'1, T_{max}'2, T_{max}'3, T_{min}'1$
	RF-Boruta	Main variables and IMFs	$P, P'2, T_{max}, T_{min}'1$
		IMFs	$P'3, T_{max}'1$
(c)			
Monthly	SVR-RFE	Main variables and IMFs	$P, P'2, P'3, P'4, P'6, T_{max}'2, T_{max}'5, T_{min}, T_{min}'1, T_{min}'2, T_{min}'5, T_{min}'6$
		IMFs	$P'1, P'2, P'3, P'4, P'6, T_{max}'1, T_{max}'5, T_{max}'6, T_{min}'1, T_{min}'2, T_{min}'5$
	RF-Boruta	Main variables and IMFs	$P, P'1, T_{max}, T_{max}'2, T_{min}$
		IMFs	$T_{max}'1, T_{max}'2, T_{min}'1, T_{min}'2$
Yearly	SVR-RFE	Main variables and IMFs	$P, P'1, P'2, T_{max}, T_{max}'2, T_{max}'4, T_{min}$
		IMFs	$P'1, P'2, P'3, P'4, T_{max}'2, T_{max}'4$
	RF-Boruta	Main variables and IMFs	$P, P'1$
		IMFs	$P'1, P'2$

Note: P.: precipitation; T.: temperature; min.: minimum; max.: maximum; the symbol ′ represents the IMF.

understanding of the SVM-RFE and RF-Boruta feature selection algorithms, please refer to the works of Ahmadpour et al. (2021), Parisouj et al. (2022), Jamei et al. (2023), Farhana et al. (2023), Kursa and Rudnicki (2010), and Maguire et al. (2022). Five groups of input variables were considered for each model: (1) P, T_{min}, and T_{max}, and their IMFs selected by SVM-RFE; (2) P, T_{min}, and T_{max}, and their IMFs selected by RF-Boruta feature selection; (3) IMFs of the main values obtained using SVM-RFE; (4) IMFs of the main values obtained using RF-Boruta feature selection; and (5) only the three main variables (P, T_{min}, and T_{max}). Table 1(a–c) presents the input variables of each method for the Grand, Winnipeg, and Moosonee River basins at monthly and yearly time scales, respectively.

The average monthly and yearly P, T_{min}, and T_{max} and their IMFs were applied as the model input, and the observed streamflow variable was used to evaluate the prediction accuracy of each model. Data from January 1973 to August 2006 were used for model training and those from September 2006 to December 2020 were used for model testing. The normalisation method was used to normalise input variables to improve the model's learning ability. The mean and standard deviation (μ and σ, respectively) were used in the training phase, defined as in Equation (9).

$$\text{normalized } x = \frac{x - \mu}{\sigma}, \qquad (9)$$

where x represent the original value.

Model parameterisation

To avoid overfitting, the 10-fold cross-validation method was applied to minimise the root mean square error (RMSE) to determine the optimal parameters of the ANN-BP, SVR, CNN, and ensemble models for training each model. The optimal hyperparameters were applied to build the optimal model for simulating the runoff for the training and testing phases. Python 3.8 was used for preparing scripts.

The ANN-BP model was built based on a three feedforward layer configuration with the limited-memory Broyden–Fletcher–Goldfarb–Shanno solver. The hidden layer consisted of the logistic sigmoid function that delineates the input weights that are transformed into an output from the nodes. The learning rate and max-iteration of the ANN-BP were 0.07 and 10,000,

respectively. The optimal hyperparameters of ANN-BP were defined by random generalisation and 10-fold cross-validation for the training model. Three hidden layers were applied to generate the highest runoff accuracy by minimising the RMSE function. Therefore, the network structure was (number of inputs for each state):5:1 for the input, hidden, and output layers.

The SVR model was built based on the RBF kernel, and random generalisation was applied for each parameter to determine the optimal hyperparameters. The three hyperparameters of the RBF kernel were set to have the following ranges to tune the model: γ (30 values from 0.0001 to 1) C (100 values from 2000 to 5000), and ε (21 values from 0.5 to 3).

The CNN architecture included two one-dimensional CNN layers, a dropout layer for regulation, and a pooling layer. CNN layers are frequently created in pairs to aid the model in learning properties from the input data. The dropout layer aims to lower the learning rate of the CNNs to produce a more accurate final model. The pooling layer decreases the size of the learned features to a quarter of their original size, allowing them to be concentrated on the most relevant aspects. Next, the flatten

function is used to flatten learned features as a vector and pass them into a fully connected layer. The fully connected layer prevents the learned features and target value to define the learned features before making a prediction. In this study, 512 and 256 filter features with kernel sizes 5 and 2 were used for the first and second CNN layers, respectively. RMSprop was used as the network optimiser. LeakyRelu was used as an activation function. A random generalisation process and 10-fold cross-validation were applied to determine the parameters.

The AdaBoost model involves three key factors: estimators, learning rate, and loss. The estimator and learning rate were set to range as follows: 30 values from 1 to 200 and from 0.0001 to 1, respectively. In the training phase, linear, square, and exponential loss functions were evaluated. In the bagging regressor and extra tree regressor models, the optimal estimator values were identified from a candidate set of thirty variables ranging from 1 to 200 and from 300 to 500, respectively, during training through 10-fold cross-validation. Stacking generation was incorporated into the models to enhance the accuracy of target data. The bagging regressor and extra tree regressor models were selected as estimators,

Table 2. (a) Monthly model performance indicators during the calibration and validation phases for the Grand River basin (combination of *P*, *T*, and EMD dataset). (b) Monthly model performance indicators during the calibration and validation phases for the Winnipeg River basin (combination of *P*, *T*, and EMD dataset). (c) Monthly model performance indicators during the calibration and validation phases for the Moosonee River basin (combination of *P*, *T*, and EMD dataset).

		Calibration				Validation			
Model	Input Selection	RMSE (m^3s^{-1})	R	NSE	MARE	RMSE (m^3s^{-1})	R	NSE	MARE
					(a)				
SVR-EMD	SVM-RFE	30.95	0.79	0.57	25.96	33.57	0.71	0.49	37.87
	RF-Boruta	34.36	0.7	0.47	38.9	40.09	0.57	0.27	50.14
ANN-BP-EMD	SVM-RFE	11.55	0.97	0.94	12.96	40.22	0.63	0.27	50.87
	RF-Boruta	22.52	0.88	0.77	25.99	44.86	0.52	0.09	56.75
CNN-EMD	SVM-RFE	37.16	0.67	0.38	53.36	39.4	0.67	0.3	39.64
	RF-Boruta	36.2	0.64	0.41	54.44	38.86	0.58	0.31	52.65
Ensemble-EMD	SVM-RFE	21.47	0.9	0.79	26.09	34.62	0.75	0.46	47.22
	RF-Boruta	15.19	0.97	0.9	21.42	37.28	0.62	0.37	50.1
					(b)				
SVR-EMD	SVM-RFE	285.17	0.77	0.58	29.14	329.62	0.67	0.44	30.53
	RF-Boruta	286.5	0.77	0.57	28.34	348.17	0.62	0.37	37.36
ANN-BP-EMD	SVM-RFE	205.01	0.88	0.78	18.4	356.43	0.63	0.34	35.92
	RF-Boruta	101.18	0.97	0.95	5.41	374.72	0.55	0.27	40.50
CNN-EMD	SVM-RFE	316.15	0.71	0.48	29.78	493.46	0.64	−0.26	63.27
	RF-Boruta	333.89	0.68	0.42	32	431.72	0.6	0.03	52.38
Ensemble-EMD	SVM-RFE	188.36	0.92	0.82	22.8	345.85	0.7	0.38	39.81
	RF-Boruta	258.93	0.81	0.65	25.67	340.3	0.66	0.40	39.84
					(c)				
SVR-EMD	SVM-RFE	62.78	0.79	0.60	64.39	71.48	0.74	0.51	68.87
	RF-Boruta	50.5	0.87	0.74	47.04	75.41	0.68	0.46	82.43
ANN-BP-EMD	SVM-RFE	60.56	0.79	0.63	78.62	70.56	0.73	0.52	79.04
	RF-Boruta	59.39	0.80	0.64	83.48	67.43	0.76	0.57	88.47
CNN-EMD	SVM-RFE	70.88	0.71	0.49	86.15	79.06	0.65	0.40	91.24
	RF-Boruta	71.88	0.70	0.48	106.92	81.30	0.61	0.37	116.04
Ensemble-EMD	SVM-RFE	49.86	0.87	0.75	47.43	66.35	0.77	0.58	56.83
	RF-Boruta	31.03	0.95	0.90	35.76	72.26	0.71	0.50	88.26

and the AdaBoost model was chosen as the final estimator to simulate the runoff values. To enhance the model performance, 10,000 runs of each model were performed, and the RMSE was used to evaluate the highest accuracy among training periods.

Model evaluation

The coefficient of correlation (R), RMSE, mean absolute relative error (MARE), and Nash–Sutcliffe efficiency (NSE) were applied to assess the runoff accuracy for ANN-BP, SVR, CNN, and ensemble model for the training and testing periods. The definitions of these statistical

indices are presented in the following equations:

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (s_i - o_i)^2} \tag{10}$$

$$R = \frac{\left(\frac{1}{n}\right) \sum_{i=1}^{n} (o_i - \bar{o})^2 (s_i - \bar{s})}{\sqrt{\left(\frac{1}{n}\right) \sum_{i=1}^{n} (o_i - \bar{o})^2} \times \sqrt{\left(\frac{1}{n}\right) \sum_{i=1}^{n} (s_i - \bar{s})^2}} \tag{11}$$

$$NSE = 1 - \frac{\sum_{i=1}^{n} (s_i - o_i)^2}{\sum_{i=1}^{n} (s_i - \bar{o})^2} \tag{12}$$
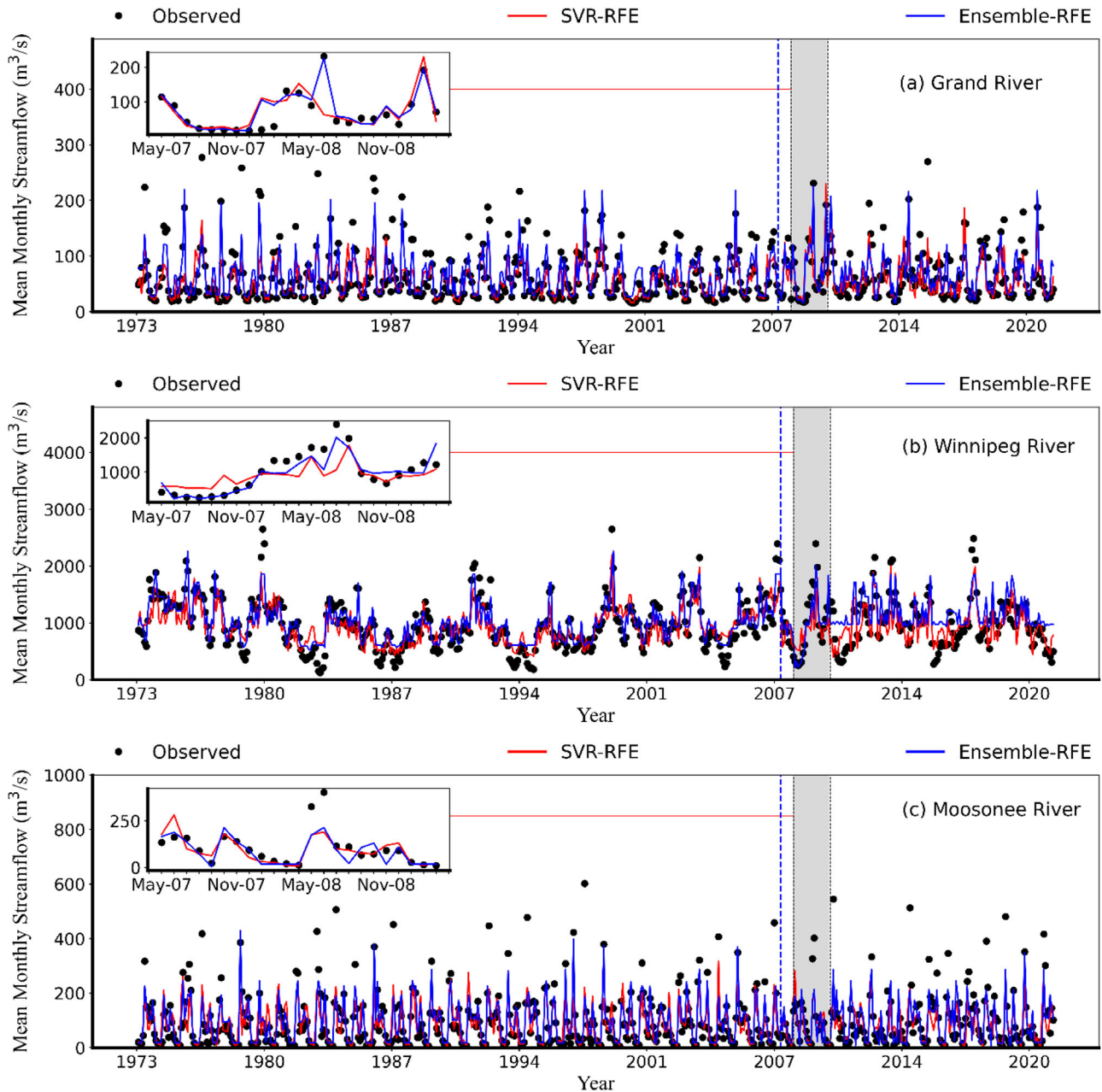
Figure 3. Best monthly line-graph using a combination of *P*, *T*, and EMD dataset.

$$\text{MARE} = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{s_i - o_i}{o_i}\right| \times 100 \qquad (13)$$

where o_i and s_i refer to the observed and estimated values, respectively; and $\bar{o}$ and $\bar{s}$ are the average observed and estimated values, respectively.
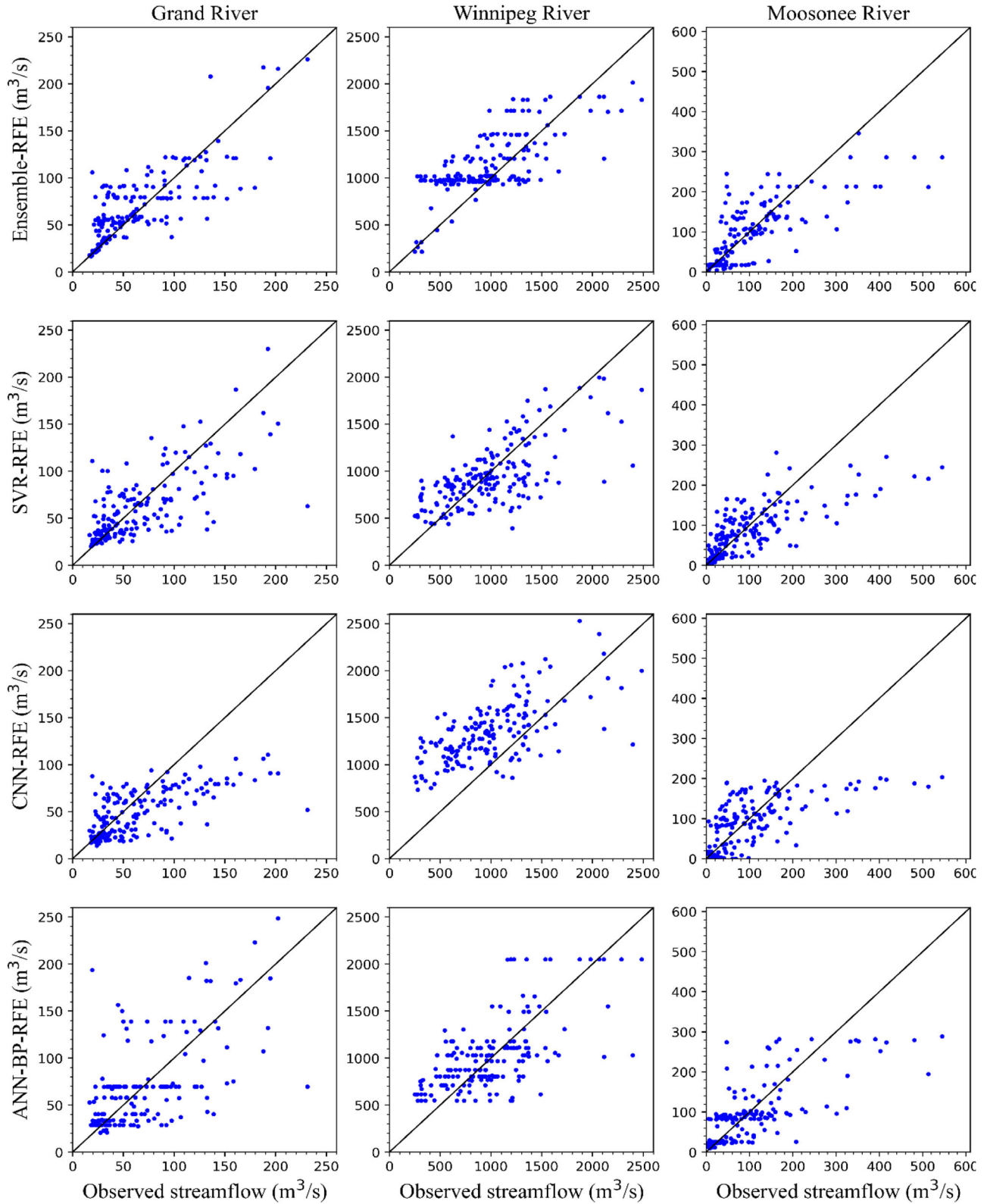
Figure 4. Scatter plots of monthly streamflow simulation using a combination of *P, T,* and EMD dataset in the validation period. The different rows present the results of different models.

Results and discussion

EMD was performed to simulate the monthly and yearly streamflow based on five scenarios: (1) using precipitation (P), maximum temperature (T_{max}), and minimum temperature (T_{min}); (2) using the RF-Boruta algorithm to select appropriate variables among decomposed variables; (3) using the RF-Boruta algorithm to select appropriate variables among decomposed variables including P, T_{max}, and T_{min}; (4) using the SVM-RFE algorithm to select appropriate variables among decomposed variables; (5) using the SVM-RFE algorithm to select appropriate variables among decomposed variables including P, T_{max}, and T_{min}. All the models were applied for the Grand, Winnipeg, and Moosonee River basins. To clarify the influence of EMD on the simulation performance, simulation models both with and without the EMD dataset were prepared. Additionally, the results for the different feature selection methods were compared to identify the most accurate model for simulating the runoff in Canadian basins.

Decomposing monthly and yearly runoff time-series using EMD

The EMD approach was used to decompose the runoff time-series for the three basins into IMFs at the monthly and yearly scales for the precipitation and maximum and minimum temperatures. Notably, several researchers have applied decomposition methods for analysing the runoff fluctuations of rivers and investigated the causes of cyclical changes in hydrological data and corresponding occurrence mechanisms (Pekárová et al., 2003; Wang et al., 2015; Williams, 1961). Decomposition can help enhance the prediction ability by transforming nonlinear and nonstationary time-series into stationary time-series.

Simulation results using a combination of P, T, and EMD dataset

Monthly time-series

Table 2(a–c) summarises the calibration and validation phase results for Grand, Winnipeg, and Moosonee

Table 3. (a) Yearly model performance indicators during the calibration and validation phases for the Grand River basin (combination of P, T, and EMD dataset). (b) Yearly model performance indicators during the calibration and validation phases for the Winnipeg River basin (combination of P, T, and EMD dataset). (c) Yearly model performance indicators during the calibration and validation phases for the Moosonee River basin (combination of P, T, and EMD dataset).

		Calibration				Validation			
Model	Input Selection	RMSE (m³s⁻¹)	R	NSE	MARE	RMSE (m³s⁻¹)	R	NSE	MARE
					(a)				
SVR-EMD	SVM-RFE	3.4	0.96	0.92	2	10.81	0.91	0.59	12.19
	RF-Boruta	6.61	0.84	0.71	7.93	11.88	0.76	0.5	15.91
ANN-BP-EMD	SVM-RFE	0	1	1	0	12.44	0.76	0.45	15.68
	RF-Boruta	0	1	1	0	24.16	0.26	−1.07	28.3
CNN-EMD	SVM-RFE	10.01	0.81	0.33	12.52	11.29	0.78	0.55	15.96
	RF-Boruta	8.4	0.8	0.53	10.51	18.26	0.47	−0.18	21.4
Ensemble-EMD	SVM-RFE	5.17	0.91	0.82	6.04	12.35	0.88	0.46	13.73
	RF-Boruta	8.64	0.73	0.5	11.4	12.88	0.79	0.41	16.51
					(b)				
SVR-EMD	SVM-RFE	148.00	0.93	0.75	13.67	191.23	0.87	0.52	16.49
	RF-Boruta	156.90	0.85	0.72	12.29	244.92	0.55	0.21	21.62
ANN-BP-EMD	SVM-RFE	27.14	1.00	0.99	1.26	261.63	0.71	0.10	20.38
	RF-Boruta	92.07	0.95	0.90	6.04	432.03	0.19	−1.45	37.14
CNN-EMD	SVM-RFE	183.33	0.84	0.61	18.76	221.49	0.80	0.36	17.90
	RF-Boruta	221.51	0.75	0.43	21.66	515.22	0.42	−2.49	49.29
Ensemble-EMD	SVM-RFE	36.56	0.99	0.98	2.70	96.22	0.94	0.88	7.73
	RF-Boruta	117.19	0.93	0.84	11.97	230.10	0.62	0.30	18.55
					(c)				
SVR-EMD	SVM-RFE	6.38	0.95	0.88	6.06	16.71	0.82	0.36	14.54
	RF-Boruta	11.08	0.81	0.65	9.95	14.33	0.73	0.53	11.41
ANN-BP-EMD	SVM-RFE	0.00	1.00	1.00	0.00	19.15	0.53	0.16	15.80
	RF-Boruta	0.00	1.00	1.00	0.00	23.10	0.29	−0.23	18.60
CNN-EMD	SVM-RFE	9.80	0.86	0.72	8.36	13.37	0.79	0.59	11.83
	RF-Boruta	11.82	0.78	0.60	9.06	15.06	0.70	0.48	12.67
Ensemble-EMD	SVM-RFE	0.00	1.00	1.00	0.00	12.40	0.83	0.65	9.96
	RF-Boruta	1.21	1.00	1.00	0.63	13.14	0.79	0.60	11.35

Rivers. When combined with EMD and SVM-RFE or RF-Boruta for input selection, the ANN-BP and ensemble models are more efficient during calibration. For example, the ANN-BP model with SVM-RFE outperforms the RF-Boruta in the Grand River calibration phase by 48% and 22% in terms of the RMSE and NSE, respectively. In the calibration phase for Winnipeg River, the ensemble model using SVM-RFE outperforms RF-Boruta with a 27% lower RMSE and 26% higher NSE. In most cases, SVM-RFE performs better than RF-Boruta during calibration. However, in the Moosonee River watershed, the situation is reversed. In this case, the SVR-EMD model

optimised with RF-Boruta exhibits a 24% higher RMSE and 19% lower NSE compared with SVM-RFE.

During the validation phase, the SVR model outperforms the other models in terms of the RMSE, NSE, and MARE for the Grand River, although its R value is inferior. The SVR model using SVM-RFE has an 81% higher NSE and a 16% lower RMSE, which confirms (Luo et al., 2022) that SVM-RFE typically selects better variables than RF-Boruta. The ensemble model outperforms other models for Winnipeg and Moosonee Rivers, regardless of its integration with SVM-RFE or RF-Boruta. Compared with RF-Boruta, SVM-RFE and EMD exhibit 1.3% and
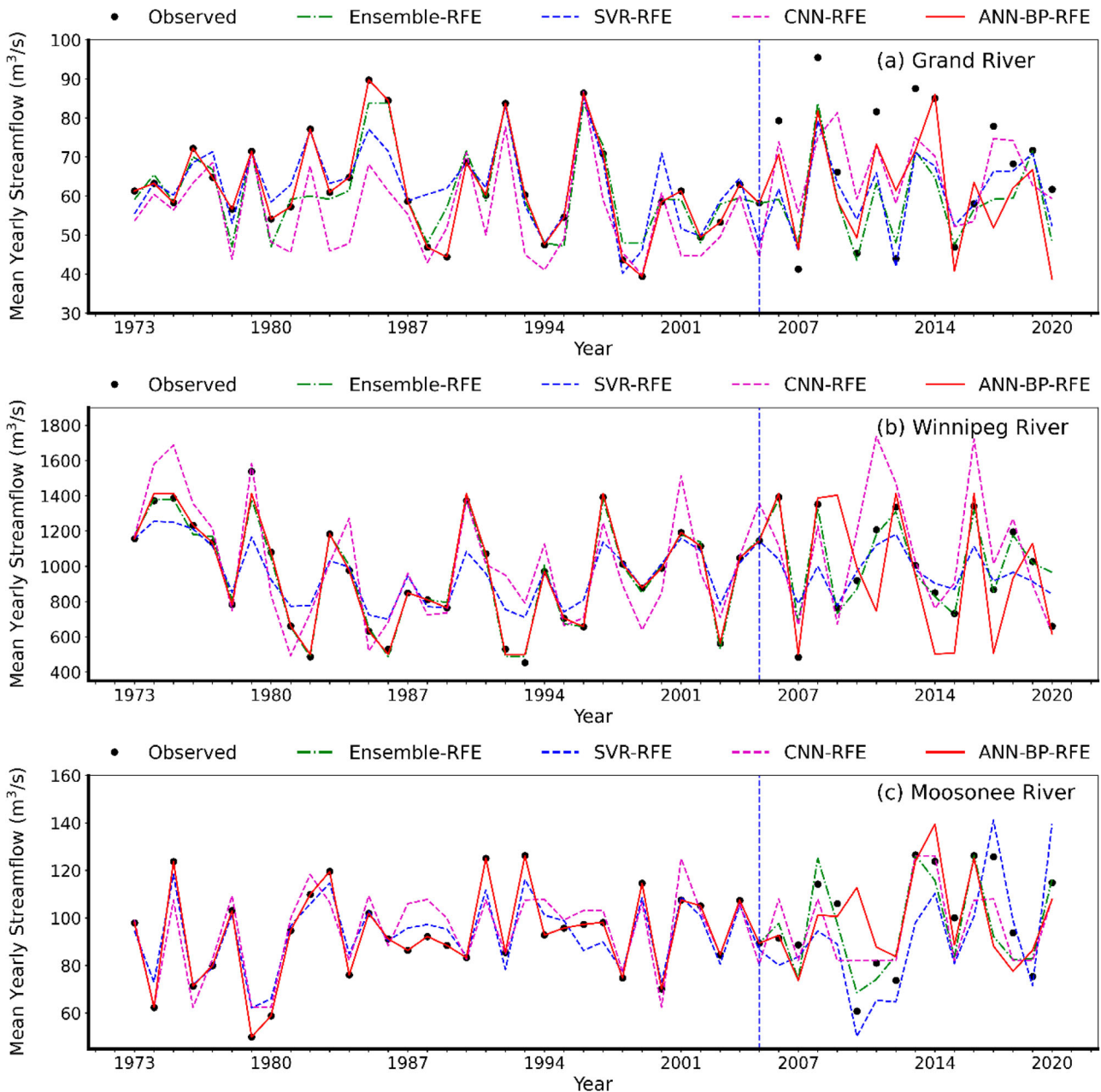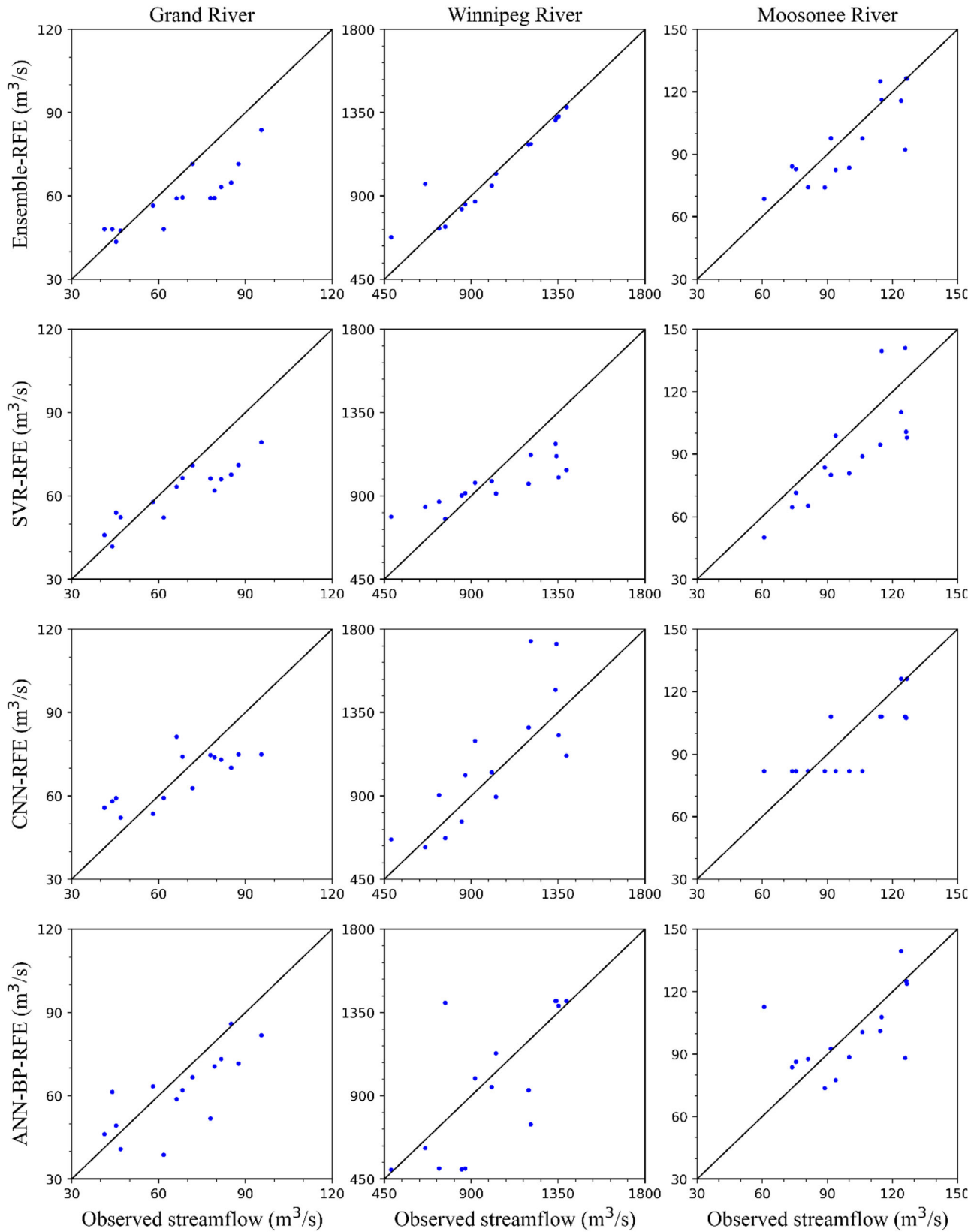

Figure 5. Yearly line-graph using a combination of *P*, *T*, and EMD dataset.

Figure 6. Scatter plots of yearly streamflow simulation using a combination of *P*, *T*, and EMD dataset in the validation period.

8.1% lower RMSEs and 5% and 16% higher NSEs, respectively, for the Winnipeg and Moosonee Rivers. The ANN-BP model outperforms the CNN model in the Moosonee River, with a 14% lower RMSE and 20% higher NSE.

The R values of the ensemble model match those of the SVR model, suggesting a similar correlation between the observed and modelled values. SVR and ensemble models exhibit reduced biases, indicating improved generalizability. The SVR model utilises a structural risk minimisation approach, resulting in a superior solution, whereas the ensemble model adopts multiple weak learners, yielding accurate predictions (Gizaw & Gan, 2016; Htike, 2017; Kumar et al., 2019; Shrestha & Shukla, 2015). Despite the overall superior performance of the ensemble model against SVR, ANN-BP, and CNN models, the SVR model is significantly superior in the Grand River basin.

Figures 3 and 4 show that the simulated and observed streamflows are consistent for both the calibration and validation phases, especially for the Moosonee River. Figure 4 demonstrates that the values predicted by the CNN and ANN-BP models lie in ranges of 50–250 m³s⁻¹ for the Grand River, 550–2,500 m³s⁻¹ for the Winnipeg River, and 50–300 m³s⁻¹ for the Moosonee River. These variances indicate that the ensemble and SVR models are more flexible and accurate in their predictions. In summary, the combination of ensemble and SVR models with SVM-RFE yields superior results, rendering them promising alternatives for streamflow prediction and similar tasks. This observation is consistent with previously reported findings and highlights exciting possibilities for future work.

Yearly time-series

Table 3(a–c) summarises the performance metrics of various models for the Grand, Winnipeg, and Moosonee River basins at the yearly scale. The findings, especially those of the NSE and MARE, emphasise the advantage of SVM-RFE over RF-Boruta in feature selection across these basins. For the Grand River basin, the SVR-EMD model using SVM-RFE exhibits a 30% higher NSE and lower MARE during the calibration phase compared with RF-Boruta. This enhancement is also observed in the validation phase, with an 18% higher NSE. Similar trends can be observed in the cases of the Winnipeg and Moosonee River basins.

Figures 5 and 6 present line graphs for the calibration and testing periods, and scatter plots for the

Table 4. (a) Monthly model performance indicators during the calibration and validation phases for the Grand River basin (using only the EMD dataset). (b) Monthly model performance indicators during the calibration and validation phases for the Winnipeg River basin (using only the EMD dataset). (c) Monthly model performance indicators during the calibration and validation phases for the Moosonee River basin (using only the EMD dataset).

		Calibration				Validation			
Model	Input Selection	RMSE (m³s⁻¹)	R	NSE	MARE	RMSE (m³s⁻¹)	R	NSE	MARE
					(a)				
SVR-EMD	SVM-RFE	16.13	0.95	0.88	3.82	42.97	0.43	0.16	56.08
	RF-Boruta	38.96	0.59	0.32	41.56	37.76	0.62	0.35	43.80
ANN-BP-EMD	SVM-RFE	12.05	0.97	0.93	13.59	60.02	0.22	−0.63	80.10
	RF-Boruta	22.46	0.88	0.77	31.74	52.57	0.38	−0.25	73.35
CNN-EMD	SVM-RFE	39.34	0.56	0.3	61.72	42.94	0.42	0.16	61.89
	RF-Boruta	42.43	0.45	0.19	64.98	42.34	0.45	0.19	60.49
Ensemble-EMD	SVM-RFE	23.5	0.92	0.75	18.04	39.07	0.56	0.31	54.81
	RF-Boruta	33.61	0.73	0.49	40.81	37.72	0.60	0.35	56.76
					(b)				
SVR-EMD	SVM-RFE	284.63	0.76	0.58	32.20	389.54	0.56	0.21	45.50
	RF-Boruta	290.59	0.77	0.56	28.15	377.18	0.56	0.26	42.59
ANN-BP-EMD	SVM-RFE	238.26	0.84	0.71	26.12	400.18	0.51	0.17	39.47
	RF-Boruta	66.97	0.99	0.98	4.49	635.00	0.40	−1.09	58.28
CNN-EMD	SVM-RFE	324.91	0.69	0.45	32.00	381.12	0.55	0.25	42.65
	RF-Boruta	325.32	0.69	0.45	30.95	479.92	0.47	−0.20	42.69
Ensemble-EMD	SVM-RFE	215.90	0.88	0.76	23.18	390.41	0.59	0.21	48.69
	RF-Boruta	230.15	0.86	0.73	25.09	347.61	0.65	0.37	40.82
					(C)				
SVR-EMD	SVM-RFE	75.58	0.71	0.42	52.35	92.05	0.51	0.19	94.85
	RF-Boruta	68.25	0.75	0.53	53.95	88.16	0.53	0.26	101.44
ANN-BP-EMD	SVM-RFE	89.53	0.46	0.19	92.30	100.95	0.46	0.03	97.01
	RF-Boruta	52.14	0.85	0.72	69.92	100.95	0.46	0.03	97.01
CNN-EMD	SVM-RFE	87.58	0.61	0.22	78.45	97.43	0.49	0.09	86.46
	RF-Boruta	78.72	0.62	0.37	121.53	87.54	0.53	0.27	122.22
Ensemble-EMD	SVM-RFE	57.50	0.86	0.66	41.53	89.01	0.51	0.24	120.60
	RF-Boruta	57.50	0.86	0.66	41.53	89.01	0.51	0.24	120.60

testing period in the three Canadian basins. The ANN-BP-EMD model paired with SVM-RFE achieves perfect calibration across all studied basins. However, its inconsistent outcome in the validation phase may suggest potential overfitting. In contrast, the SVR-EMD and ensemble-EMD models combined with SVM-RFE yield more consistent results in both calibration and validation phases. Specifically, the ensemble-EMD model paired with SVM-RFE exhibits robust performance across all basins. For example, in the case of the Winnipeg River basin, it achieves an NSE of 0.98 and a MARE of 2.70 during calibration, significantly surpassing both CNN-EMD and SVR-EMD. This high performance is retained in the validation phase, with an NSE of 0.88. Similar trends are observed for the Moosonee River basin. Overall, in the annual simulations involving scenario 5, SVM-RFE consistently outperforms RF-Boruta in feature selection across all basins, and the ensemble-EMD model outperforms the other models.

To clarify the influence of the combined P, T, and EMD on the model efficiency, simulation models using only the EMD technique were evaluated, as described in the following section.
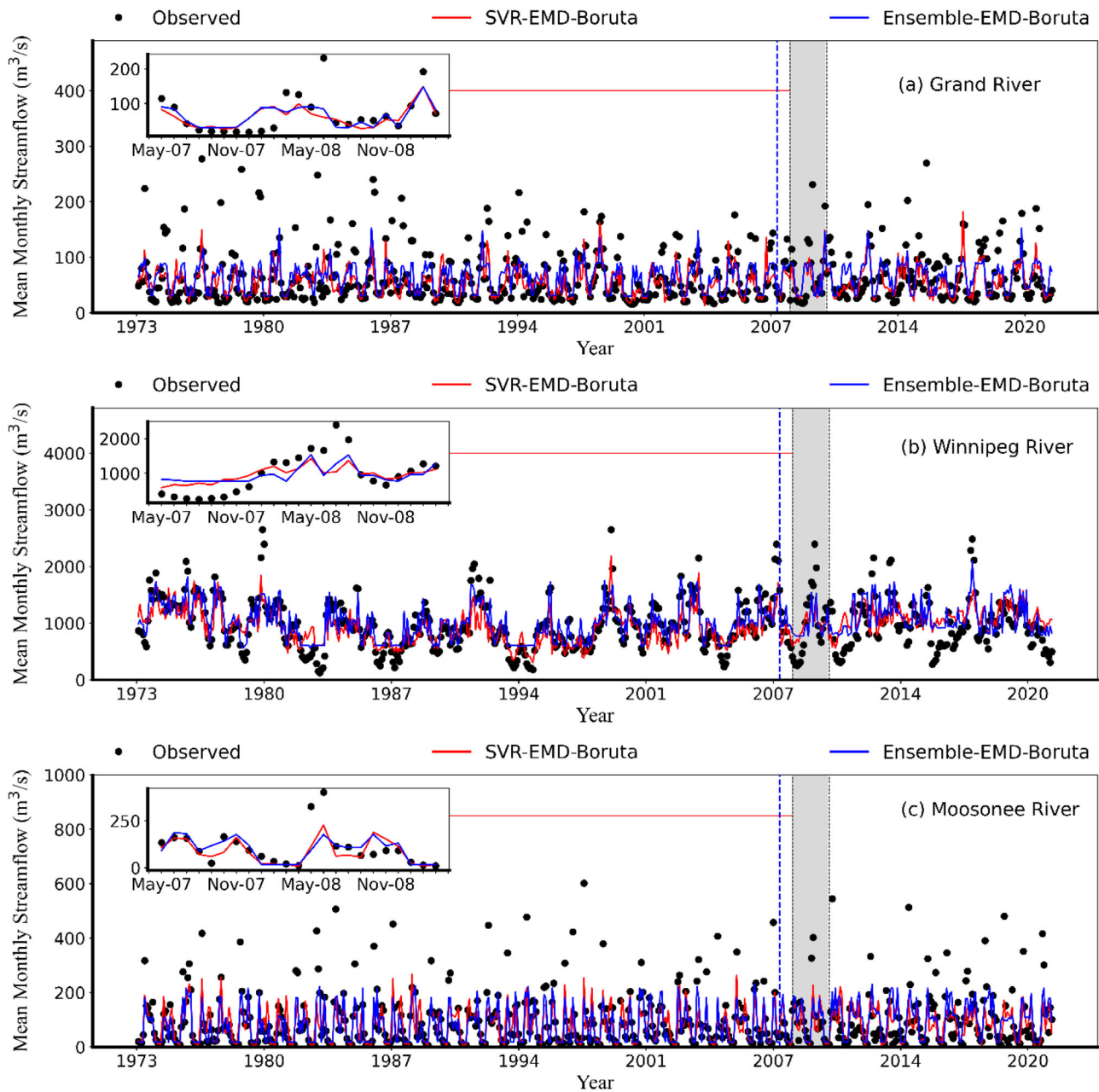
Figure 7. Best monthly line-graph using the EMD dataset.
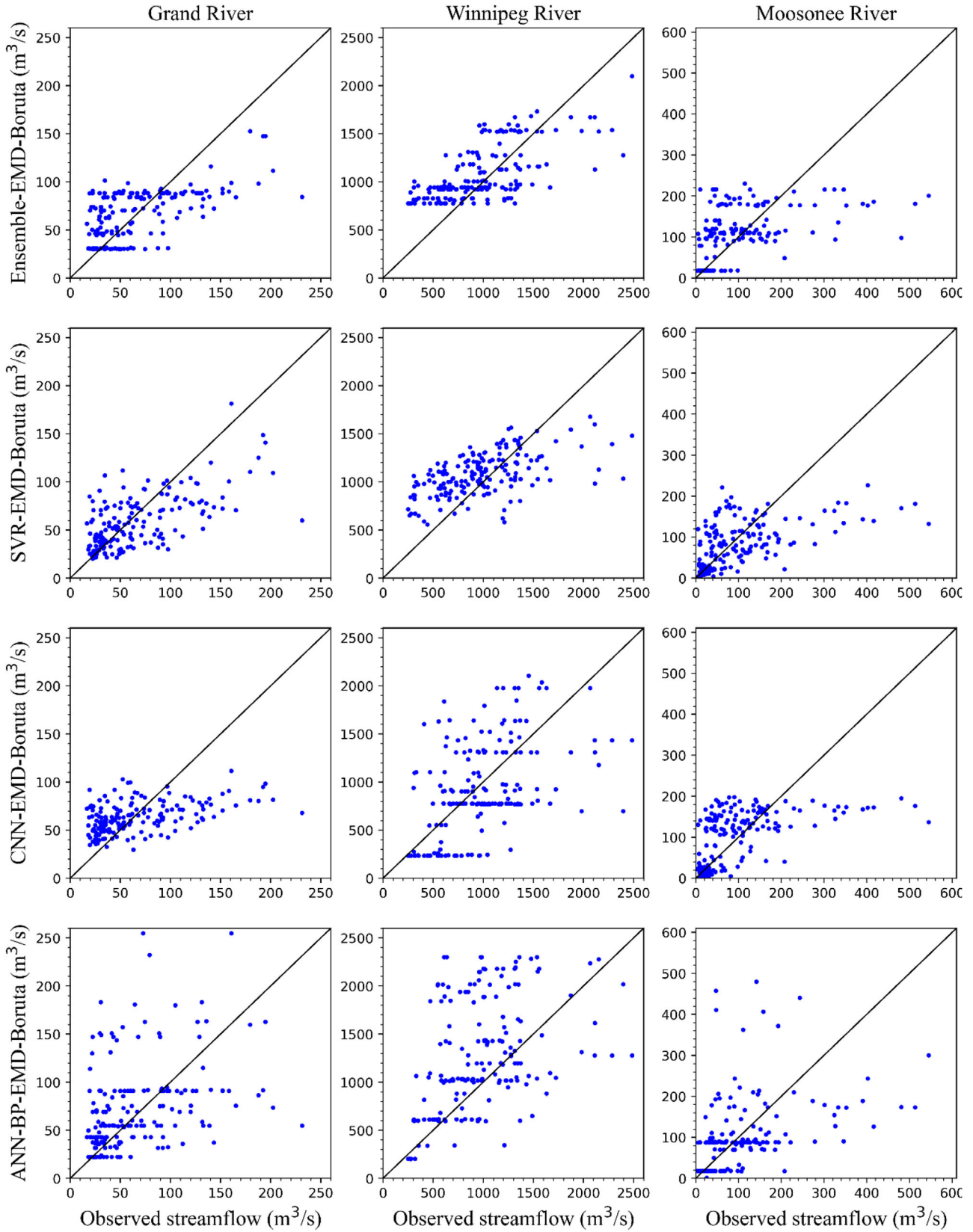
Figure 8. Scatter plots of monthly streamflow simulation using the EMD dataset in the validation period.

Simulation results using the EMD dataset

Monthly time-series

Valuable insights could be derived from the analysis of models across the three river basins. In the case of the Grand River basin (Table 4(a)), the ANN-BP-EMD model with SVM-RFE selection achieves the highest NSE of 0.93 during calibration, but this value sharply deteriorates to −0.63 in the validation phase, indicating potential overfitting. In contrast, the ensemble-EMD model with RF-Boruta consistently performs well in both stages with an NSE of 0.49 and 0.35 in the calibration and validation phases, respectively.

In the case of the Winnipeg River basin (Table 4(b)), the ANN-BP-EMD model with RF-Boruta demonstrates satisfactory calibration performance with an NSE of 0.98. However, its performance deteriorates during the validation phase. In comparison, the ensemble-EMD model with RF-Boruta consistently exhibits a strong performance, with an NSE of 0.73 and 0.37 during calibration and validation, respectively. Notably, the ensemble model outperforms all other models, and the worst performance corresponds to the ANN-BP model using RF-Boruta.

For the Moosonee River basin (Table 4(c)), the ANN-EMD model with RF-Boruta for feature selection demonstrates strong performance during calibration (NSE = 0.72), but its performance deteriorates in the validation phase (NSE = 0.03). In contrast, the ensemble-EMD model exhibits a high performance in both phases (NSE = 0.66 and 0.24 in calibration and validation, respectively), regardless of the feature selection method used (SVM-RFE or RF-Boruta).

In general, the inclusion of main variables such as precipitation and temperature in the EMD dataset improves model performance, except for the SVR model with RF-Boruta in the Grand River basin, where the improvement is minimal. This observation is consistent with that reported by Zhu and Pierskalla (2016), which emphasises the potential bias introduced by a large number of input features. Figures 7 and 8 present the performance metrics of models using the EMD dataset with RF-Boruta across the three Canadian basins. The model ranking in calibration and validation is ANN-BP > ensemble > SVR > CNN and ensemble > SVR > CNN > ANN-BP, respectively.

Yearly time-series

Table 5(a–c) presents the RMSE, R, NSE, and MARE metrics for the model during the training and validation

Table 5. (a) Yearly model performance indicators during the calibration and validation phases for the Grand River basin (using only the EMD dataset). (b) Yearly model performance indicators during the calibration and validation phases for the Winnipeg River basin (using only the EMD dataset). (c) Yearly model performance indicators during the calibration and validation phases in the Moosonee River basin (using only the EMD dataset).

		Calibration				Validation			
Model	Input Selection	RMSE ($m^3 s^{-1}$)	R	NSE	MARE	RMSE ($m^3 s^{-1}$)	R	NSE	MARE
					(a)				
SVR-EMD	SVM-RFE	7.43	0.8	0.63	9.01	15.2	0.88	0.18	16.95
	RF-Boruta	8.8	0.71	0.49	10.95	10.18	0.85	0.63	11.66
ANN-BP-EMD	SVM-RFE	0	1	1	0	18.41	0.66	−0.2	18.57
	RF-Boruta	0	1	1	0	18.03	0.53	−0.15	20.51
CNN-EMD	SVM-RFE	19.69	0.8	−1.58	28.97	21.74	0.7	−0.67	29.51
	RF-Boruta	10.82	0.57	0.22	14.98	13	0.65	0.4	18.63
Ensemble-EMD	SVM-RFE	7.49	0.8	0.63	9.7	10.06	0.82	0.64	13.91
	RF-Boruta	8.66	0.78	0.50	11.78	12.81	0.66	0.42	18.12
					(b)				
SVR-EMD	SVM-RFE	139.57	0.93	0.78	12.23	191.39	0.85	0.52	16.90
	RF-Boruta	155.03	0.86	0.72	9.87	211.75	0.73	0.41	17.83
ANN-BP-EMD	SVM-RFE	0.01	1.00	1.00	0.00	274.01	0.68	0.01	22.05
	RF-Boruta	45.85	0.99	0.98	2.48	271.54	0.45	0.03	22.69
CNN-EMD	SVM-RFE	130.80	0.90	0.80	12.25	235.64	0.72	0.27	18.63
	RF-Boruta	170.93	0.82	0.66	17.33	231.12	0.60	0.30	22.68
Ensemble-EMD	SVM-RFE	42.08	0.99	0.98	3.48	213.66	0.76	0.40	15.47
	RF-Boruta	187.61	0.79	0.59	14.95	185.94	0.77	0.55	15.37
					(c)				
SVR-EMD	SVM-RFE	10.21	0.85	0.70	7.34	16.67	0.71	0.36	13.97
	RF-Boruta	13.76	0.71	0.45	11.65	16.71	0.67	0.36	14.85
ANN-BP-EMD	SVM-RFE	9.21	0.87	0.75	7.46	18.21	0.53	0.24	15.35
	RF-Boruta	9.49	0.86	0.74	8.37	22.52	0.47	−0.17	18.85
CNN-EMD	SVM-RFE	16.10	0.73	0.25	11.62	34.71	0.68	−1.77	31.36
	RF-Boruta	15.68	0.63	0.29	11.15	25.28	0.55	−0.47	21.03
Ensemble-EMD	SVM-RFE	12.45	0.76	0.55	11.83	14.71	0.72	0.50	12.50
	RF-Boruta	8.18	0.94	0.81	4.72	16.11	0.71	0.40	11.74

periods. Because of the inherent structure of neural network models, their performance during the validation phase is inferior to that during calibration. In the case of the Grand River basin, the SVR model using RF-Boruta and the ensemble model using SVM-RFE exhibit excellent performances, with the highest R values of 0.85 and 0.82, respectively, during validation. The corresponding RMSE values are 10.18 and 10.06 m^3s^{-1}, and the NSE values are 0.63 and 0.64. Table 5(a) reveals that although the models combined with RF-Boruta generally outperform those combined with SVM-RFE, the ensemble model that uses SVM-RFE for feature selection demonstrates the best performance among all models.

Table 5(b) shows that for the Winnipeg River basin, the SVR model based on SVM-RFE outperforms the other models, with the highest R values. Moreover, the NSE is 0.52 and RMSE is 191.39 m^3s^{-1}. For the Moosonee River basin (Table 5(c)), the ensemble model exhibits excellent performance during the validation period, achieving the highest R values of 0.72 and 0.71 when coupled with the SVM-RFE and RF-Boruta selection methods, respectively. The corresponding NSE values are 0.50 and 0.40, and the RMSE values are 14.71 and 16.11 m^3s^{-1}. These metrics are highly improved compared with the other models, and in particular, the RMSE is 5% lower.
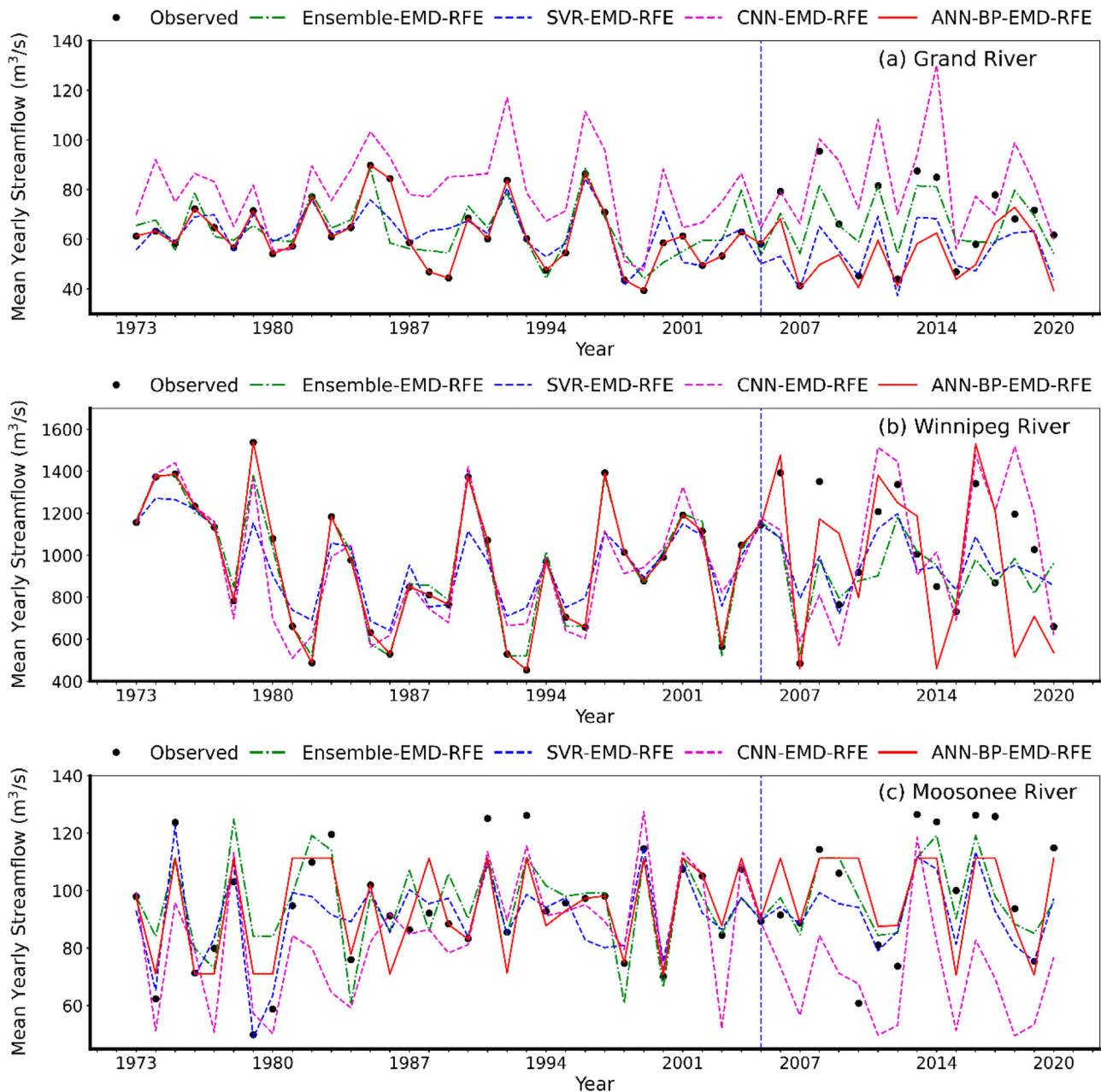
Figure 9. Best yearly line-graph using the EMD dataset.
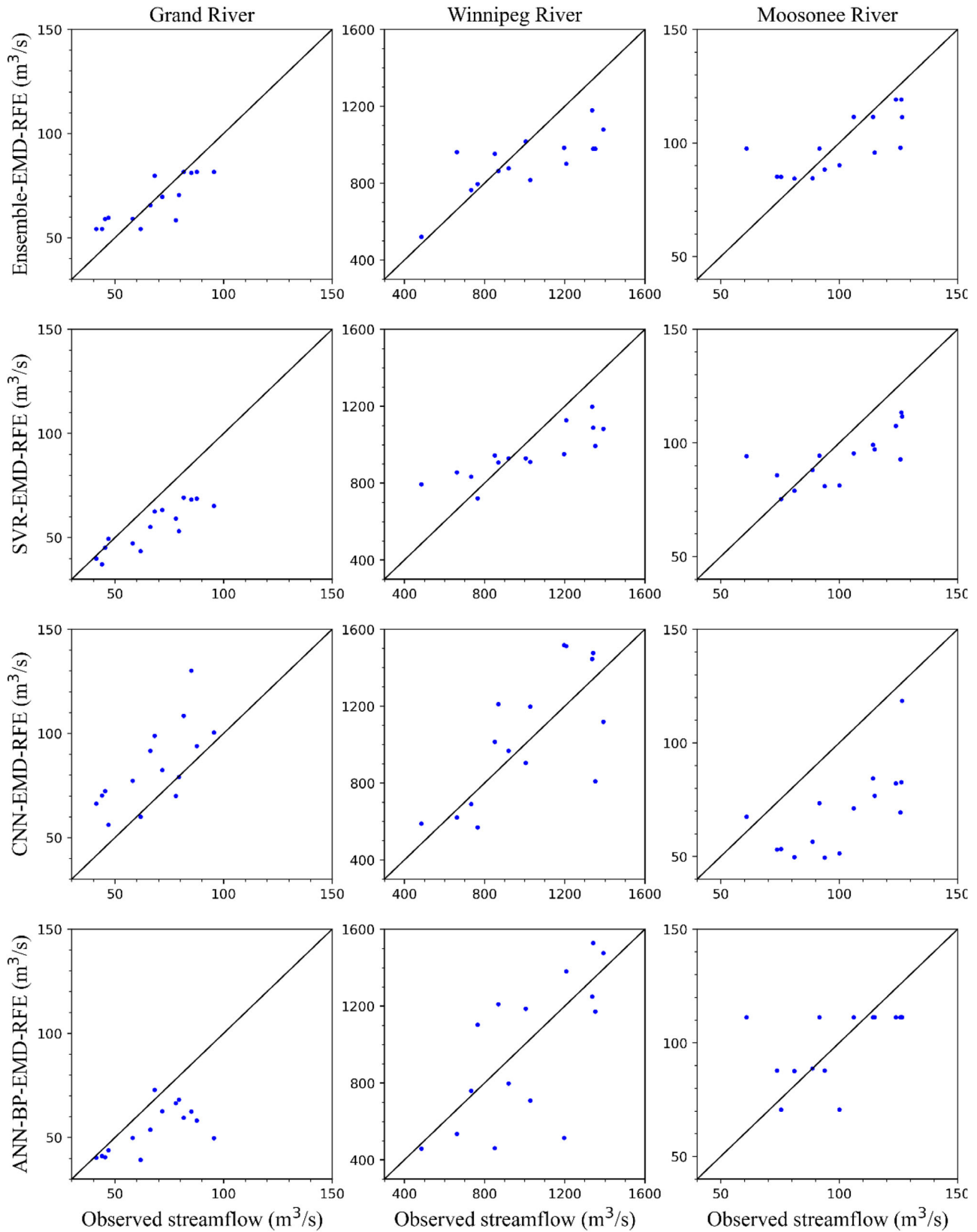
Figure 10. Scatter plots of yearly streamflow simulation using the EMD dataset in the validation period.

Figures 9 and 10 show the plots of the observed and simulated yearly streamflows. The simulation results obtained using the SVR and ensemble models are closer to the observed streamflows than those of the CNN and ANN-BP, which tend to overestimate the yearly streamflow values. The overall performance ranking of the models can be summarised as follows: SVR is comparable or superior to the ensemble model, which outperforms CNN, which in turn is superior to ANN-BP.

Simulation results using P and T

Monthly time-series

Table 6(a–c) summarises the performance metrics of the models in the training and testing phases. In both phases, the ensemble model consistently outperforms the other models, with the highest R, RMSE, NSE, and MARE values across the Grand, Winnipeg, and Moosonee Rivers. For the Grand River basin, the ensemble model shows an RMSE of 36.27 m^3s^{-1} in the validation phase, approximately 15% lower than that of the SVR (42.44), indicating reduced prediction errors. Moreover, its NSE is nearly two times higher, indicating an increase of almost 100%, and its MARE is approximately 8% lower than that of the SVR model. Table 6(b) reveals that the trends for the Winnipeg River are similar: The RMSE is 329.46 m^3s^{-1}, which is approximately 40% lower than that of the SVR (543.86 m^3s^{-1}). The NSE of the ensemble model (0.44) is significantly higher than that of the SVR (−0.53), and its MARE is nearly 53% lower (72.29). In Table 6(c), similar

Table 6. (a) Monthly model performance indicators during the calibration and validation phases for the Grand River basin (using P and T). (b) Monthly model performance indicators during the calibration and validation phases for the Winnipeg River basin (using P and T). (c) Monthly model performance indicators during the calibration and validation phases for the Moosonee River basin (using P and T).

	Calibration				Validation			
	(a)							
Model	RMSE (m^3s^{-1})	R	NSE	MARE	RMSE (m^3s^{-1})	R	NSE	MARE
SVR	32.23	0.74	0.53	33.72	42.44	0.51	0.18	46.73
ANN-BP	22.02	0.88	0.78	29.56	54.76	0.31	−0.36	61.01
CNN	39.03	0.65	0.32	40.42	43.23	0.62	0.15	51.4
Ensemble	20.17	0.91	0.82	19.6	36.27	0.65	0.4	42.94
	(b)							
SVR	323.72	0.69	0.46	31.34	543.86	0.60	−0.53	72.29
ANN-BP	238.26	0.84	0.71	26.12	485.98	0.22	−0.23	48.71
CNN	349.49	0.61	0.37	40.46	490.70	0.62	−0.25	64.34
Ensemble	286.20	0.77	0.58	34.19	329.46	0.66	0.44	34.03
	(c)							
SVR	68.08	0.76	0.53	53.42	84.31	0.59	0.32	122.38
ANN-BP	46.81	0.90	0.78	109.95	111.97	0.50	−0.20	183.62
CNN	78.98	0.67	0.37	84.24	79.03	0.68	0.40	80.03
Ensemble	42.53	0.91	0.82	34.29	72.42	0.71	0.50	85.59

Table 7. (a) Yearly model performance indicators during the calibration and validation phases for the Grand River basin (using P and T). (b) Yearly model performance indicators during the calibration and validation phases for the Winnipeg River basin (using P and T). (c). Yearly model performance indicators during the calibration and validation phases for the Moosonee River basin (using P and T).

	Calibration				Validation			
	(a)							
Model	RMSE (m^3s^{-1})	R	NSE	MARE	RMSE (m^3s^{-1})	R	NSE	MARE
SVR	7.37	0.81	0.64	8.62	13.62	0.88	0.34	14.52
ANN-BP	0	1	1	0	17.3	0.55	−0.06	20.2
CNN	8.8	0.73	0.48	11.2	12.33	0.71	0.46	16.82
Ensemble	7.76	0.79	0.6	9.03	11.43	0.77	0.54	14.03
	(b)							
SVR	155.75	0.85	0.72	11.65	196.04	0.72	0.50	19.60
ANN-BP	101.17	0.94	0.88	4.67	298.07	0.53	−0.17	20.35
CNN	212.12	0.70	0.48	23.02	232.46	0.58	0.29	23.06
Ensemble	142.28	0.91	0.77	9.78	187.45	0.79	0.54	17.48
	(c)							
SVR	10.82	0.82	0.66	8.55	15.19	0.72	0.47	13.09
ANN-BP	0.00	1.00	1.00	0.00	24.18	0.55	−0.35	23.30
CNN	14.33	0.73	0.41	11.70	21.99	0.67	−0.11	17.91
Ensemble	11.95	0.77	0.59	11.94	15.79	0.68	0.43	13.21
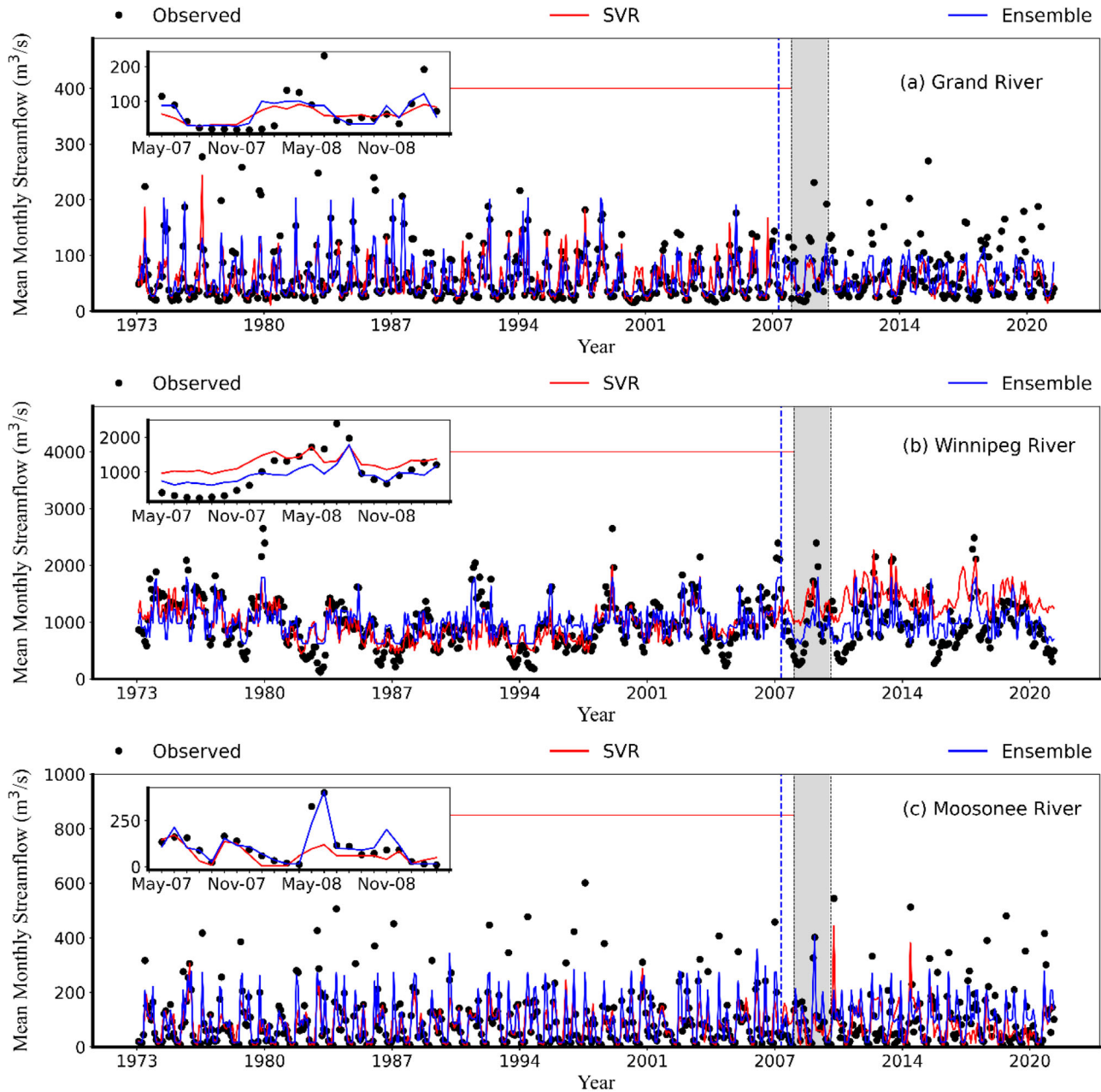
Figure 11. Best monthly line-graph using *P* and *T*.

observations can be made for the Moosonee River. The ensemble model exhibits a 14% lower RMSE and 56% higher NSE compared with the SVR model.

Figure 11 compares the observed and predicted streamflows by the SVR and ensemble models, highlighting the superior performance of these two models. As shown in Figure 12, during the validation phase, the ensemble model outperforms the CNN, SVR, and ANN-BP across all river basins. In the validation phase (Figure 12), the ensemble model outperforms the CNN, SVR, and ANN-BP across all river basins. Therefore,

the ensemble model is noted to be the most reliable for predicting streamflow.

Yearly time-series

Table 7(a–c) indicate that during the calibration phase, the ANN-BP model exhibits excellent performance across all basins. However, its performance deteriorates in the validation phase, particularly for the Grand River basin (Table 7(a)). This deterioration is evidenced by the *R* value, which decreases from 1 to 0.55, corresponding to a 45% reduction, and its RMSE, which increases from
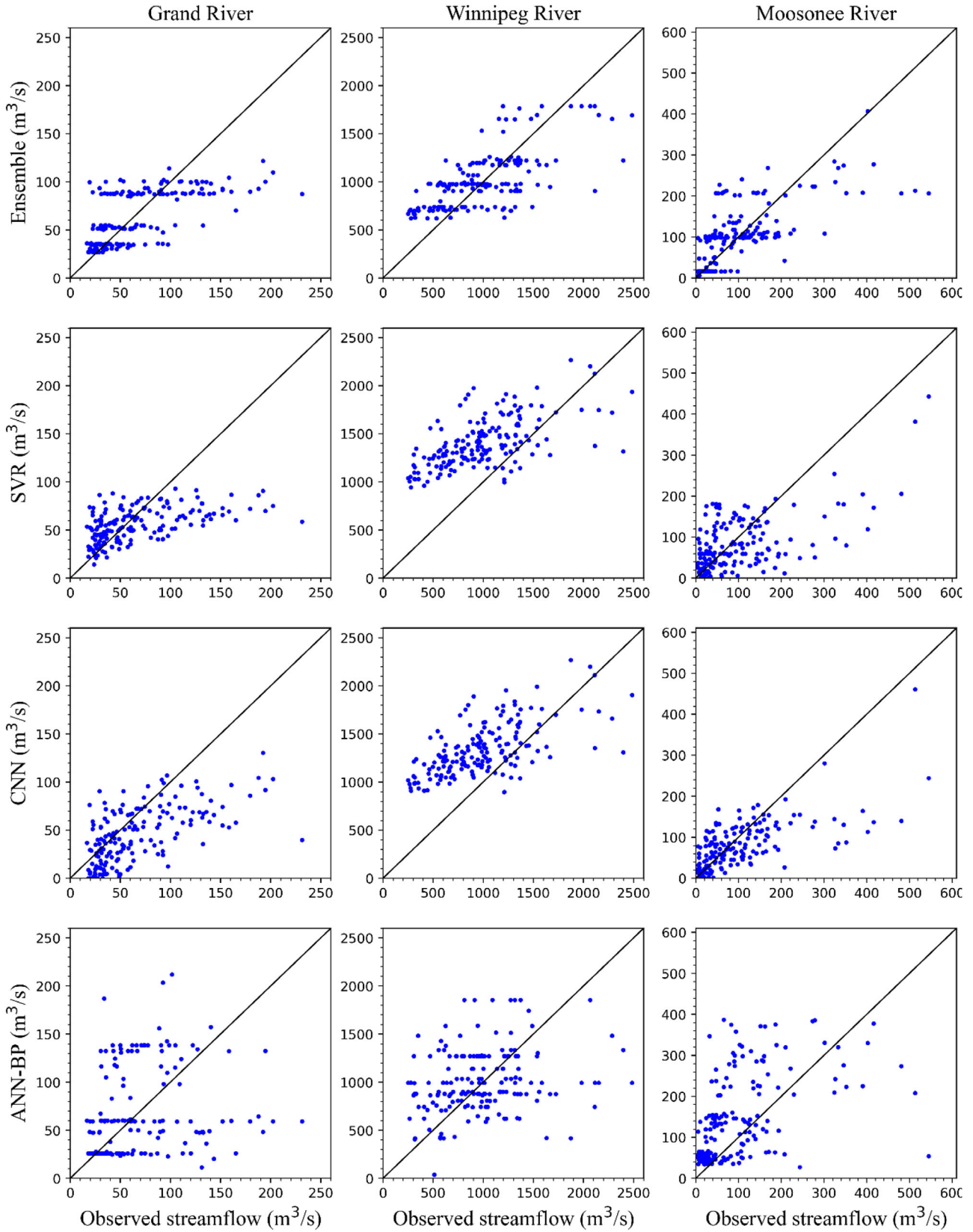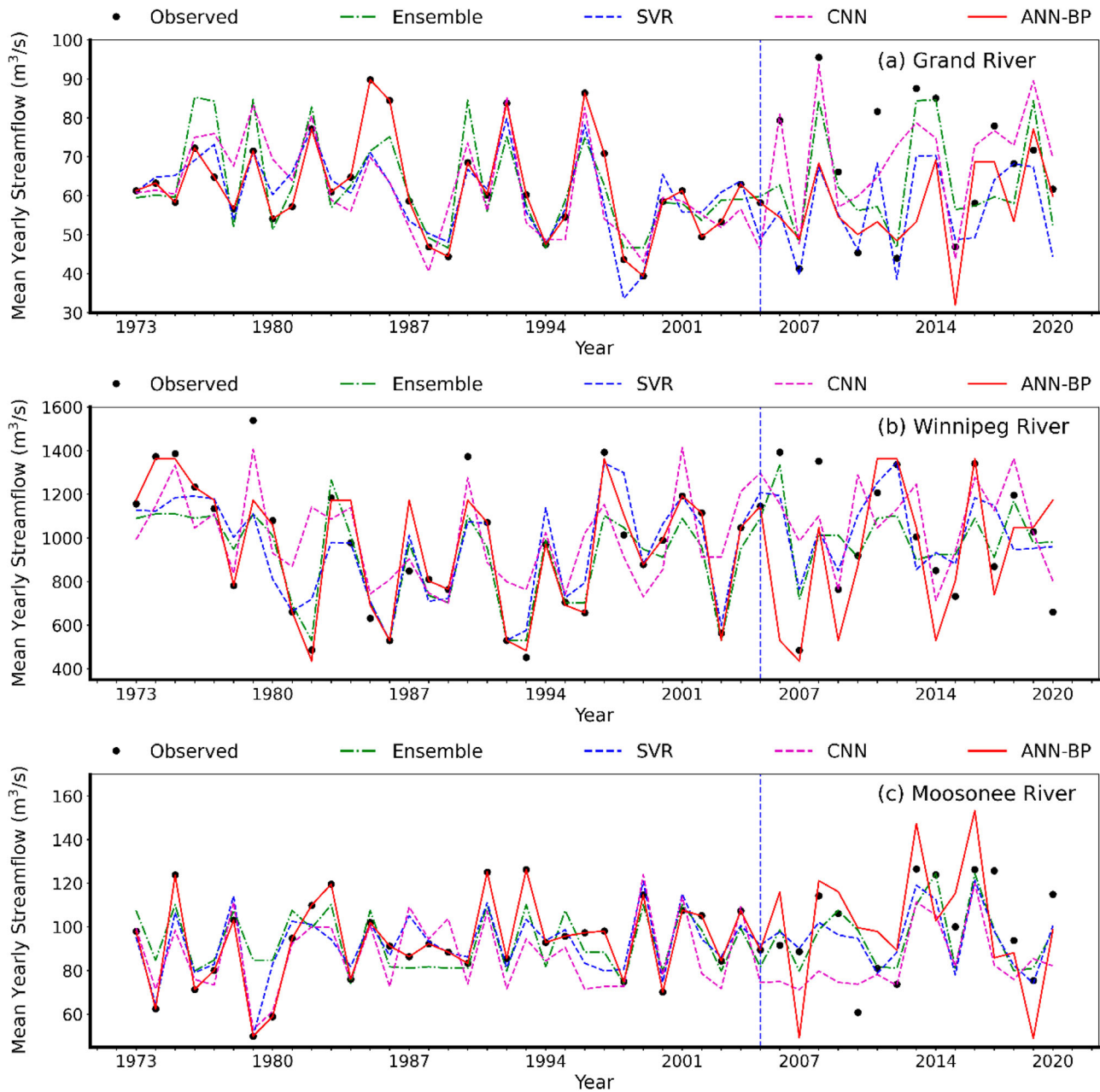
Figure 12. Scatter plots of monthly streamflow simulation using *P* and *T* in the validation period.

Figure 13. Yearly line-graph using *P* and *T*.

0 to 17.3 m^3s^{-1}. Similar trends are observed for the Winnipeg River basin (Table 7(b)): The RMSE of ANN-BP increases by nearly 195%, from 101.17 to 298.07 m^3s^{-1}. Similarly, in the case of the Moosonee River basin (Table 7(c)), the model performance deteriorates, with the RMSE increasing from 0 to 24.18 m^3s^{-1} and the R value decreasing from 1 to 0.55, corresponding to a 45% reduction. These findings highlight the inconsistency of the model performance in the calibration and validation phases.

In the validation phase, the SVR and ensemble models outperform the ANN-BP across all basins. In the case of the Grand River basin, the ensemble records a 30% lower

RMSE (11.43 m^3s^{-1}) and a 40% higher *R* value (0.77). Similarly, in the case of the Winnipeg River basin, the validation RMSE of the ensemble model is 37% lower (187.45 m^3s^{-1}) and its *R* value is 49% higher (0.79). Finally, in the case of the Moosonee River basin, the SVR model exhibits superior performance with a 37% lower RMSE (15.19 m^3s^{-1}) and 31% higher *R* value (0.72). The SVR model, which has an RMSE only 4% lower than that of the ensemble model, can be considered slightly inferior to the ensemble model.

Figures 13 and 14 show the results of simulations using the three main variables as the predictor variables. Figure 13 shows that the performance of most
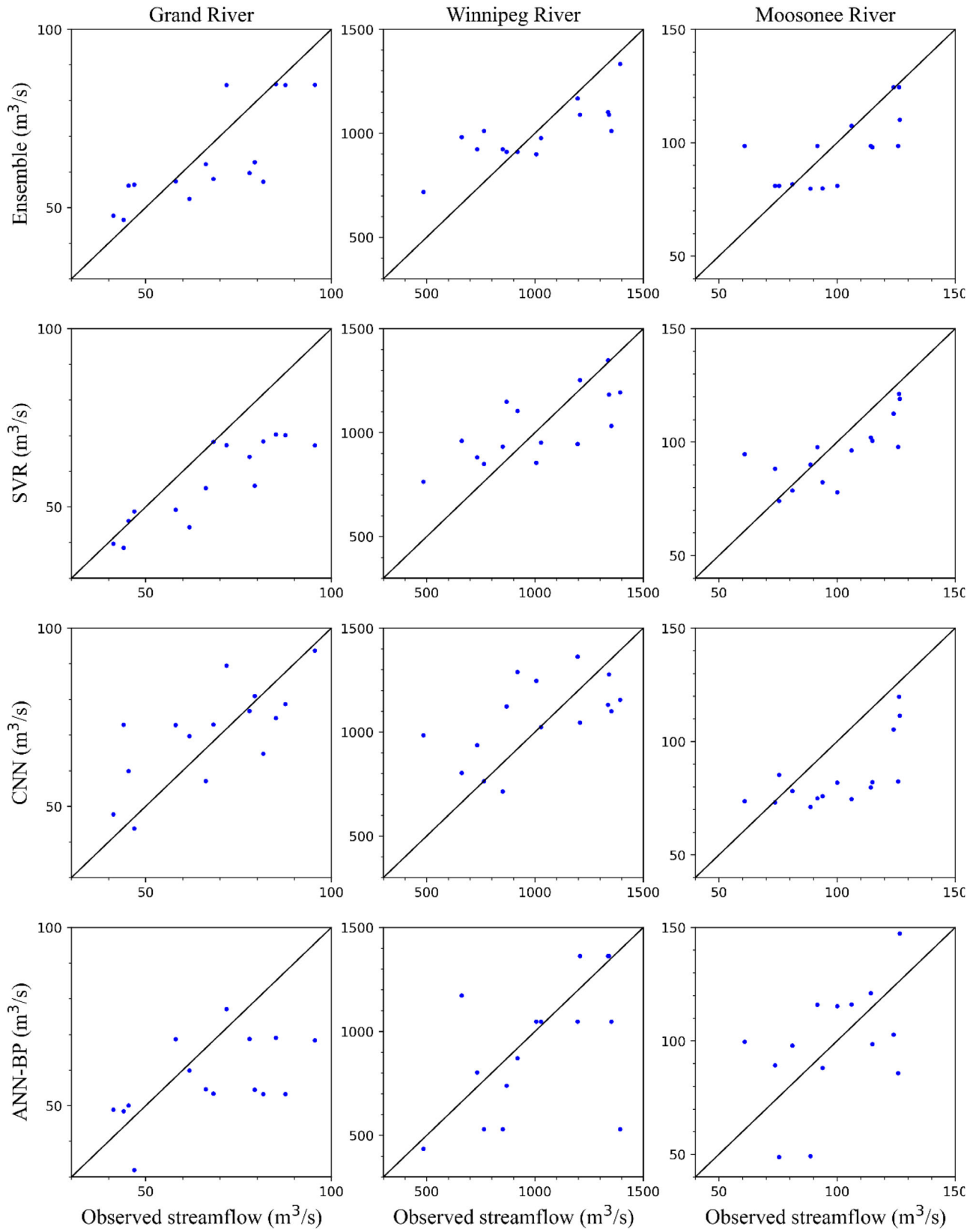
Figure 14. Scatter plots of yearly streamflow simulation using *P* and *T* in the validation period.

models in the case of the Grand River basin is higher than those for the Winnipeg and Moosonee River basins. Figure 14 shows the scatter plots for each model for the three Canadian basins in the validation period. For all three basins, all four models, especially ANN-BP, overestimate or underestimate the streamflow variable.

Conclusions

This paper proposes a hybrid model based on EMD and four ML models: ensemble, SVR, CNN, and ANN-BP models, to simulate the monthly and yearly runoff time-series and increase the simulation accuracy of long-term runoff. ML models based on the original variables (monthly and yearly time-series of the precipitation and maximum and minimum temperatures) were developed as a comparative standard. Monthly and yearly runoff data from the Grand, Winnipeg, and Moosonee Rivers in Canada were used, and four statistical metrics (RMSE, MARE, R, and NSE) were adopted to assess the model performances. The results demonstrated that the EMD increases the simulation precision, and the proposed EMD-ML models are superior to the standalone ML models in monthly and annual runoff time-series modelling. The proposed hybrid approach can be applied in future research for simulating monthly and annual runoff. The advantages of the proposed technique can be summarised as follows: First, despite the simplicity of the EMD, it can offer valuable insights into the characteristics of the monthly and yearly runoff time-series. Second, the monthly scale is associated with a lower accuracy than the yearly scale. Third, the ensemble model outperforms the SVR, CNN, and ANN-BP for the Grand, Winnipeg, and Moosonee River basins. Finally, the proposed models do not necessitate intricate decision-making regarding the explicit form for each instance. Overall, a hybrid simulation model with EMD can yield precise and consistent simulation results, and it is thus a valuable tool for studies that focus on hydrological time-series simulations to address various problems associated with reservoir management.

Disclosure statement

Funding

ORCID

Essam Heggy 🔟 http://orcid.org/0000-0001-7476-2735

References

Achite, M., Jehanzaib, M., Elshaboury, N., & Kim, T. W. (2022). Evaluation of machine learning techniques for hydrological drought modeling: A case study of the Wadi Ouahrane Basin in Algeria. *Water (Switzerland)*, *14*(3), 431. https://doi.org/10.3390/w14030431

Ahmadpour, H., Bazrafshan, O., Rafiei-Sardooi, E., Zamani, H., & Panagopoulos, T. (2021). Gully erosion susceptibility assessment in the Kondoran watershed using machine learning algorithms and the Boruta feature selection. *Sustainability*, *13*(18), 10110. https://doi.org/10.3390/su131810110

Bafitlhile, T. M., & Li, Z. (2019). Applicability of ϵ-support vector machine and artificial neural network for flood forecasting in humid, semi-humid and semi-arid basins in China. *Water*, *11*(1), 85. https://doi.org/10.3390/w11010085

Bahamonde, P. A., Fuzzen, M. L., Bennett, C. J., Tetreault, G. R., McMaster, M. E., Servos, M. R., Martyniuk, C. J., & Munkittrick, K. R. (2015). Whole organism responses and intersex severity in rainbow darter (Etheostoma caeruleum) following exposures to municipal wastewater in the Grand River basin, ON, Canada. Part A. *Aquatic Toxicology*, *159*, 290–301. https://doi.org/10.1016/j.aquatox.2014.11.023

Band, S. S., Heggy, E., Bateni, S. M., Karami, H., Rabiee, M., Samadianfard, S., Chau, K. W., & Mosavi, A. (2021). Groundwater level prediction in arid areas using wavelet analysis and Gaussian process regression. *Engineering Applications of Computational Fluid Mechanics*, *15*(1), 1147–1158. https://doi.org/10.1080/19942060.2021.1944913

Barrera-Animas, A. Y., Oyedele, L. O., Bilal, M., Akinosho, T. D., Delgado, J. M. D., & Akanbi, L. A. (2022). Rainfall prediction: A comparative analysis of modern machine learning algorithms for time-series forecasting. *Machine Learning with Applications*, *7*, 100204. https://doi.org/10.1016/j.mlwa.2021.100204

Bartoletti, N., Casagli, F., Marsili-Libelli, S., Nardi, A., & Palandri, L. (2018). Data-driven rainfall/runoff modelling based on a neuro-fuzzy inference system. *Environmental Modelling & Software*, *106*, 35–47. https://doi.org/10.1016/j.envsoft.2017.11.026

Barzegar, R., Aalami, M. T., & Adamowski, J. (2021). Coupling a hybrid CNN-LSTM deep learning model with a boundary corrected maximal overlap discrete wavelet transform for multiscale lake water level forecasting. *Journal of Hydrology*, *598*, 126196. https://doi.org/10.1016/j.jhydrol.2021.126196

Breiman, L. (1996). Bagging predictors. *Machine Learning*, *24*(2), 123–140. https://doi.org/10.1007/bf00058655

Chen, X., Li, F. W., & Feng, P. (2018). A new hybrid model for nonlinear and non-stationary runoff prediction at annual and monthly time scales. *Journal of Hydro-Environment Research*, *20*, 77–92. https://doi.org/10.1016/j.jher.2018.05.004

Chollet, F. (2017, January). Xception: Deep learning with depthwise separable convolutions. In *Proceedings of the 30th*

IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2017 (pp. 1800–1807). https://doi.org/10.1109/CVPR.2017.195

Chu, T. Y., & Huang, W. C. (2020). Application of empirical mode decomposition method to synthesize flow data: A case study of Hushan Reservoir in Taiwan. *Water*, *12*(4), 927. https://doi.org/10.3390/W12040927

Cortes, C., & Vapnik, V. (1995). Support-vector networks. *Machine Learning*, *20*(3), 273–297. https://doi.org/10.1007/BF00994018

Cui, Z., Qing, X., Chai, H., Yang, S., Zhu, Y., & Wang, F. (2021). Real-time rainfall-runoff prediction using light gradient boosting machine coupled with singular spectrum analysis. *Journal of Hydrology*, *603*, 127124. https://doi.org/10.1016/j.jhydrol.2021.127124

Elbeltagi, A., Di Nunno, F., Kushwaha, N. L., de Marinis, G., & Granata, F. (2022). River flow rate prediction in the Des Moines watershed (Iowa, USA): a machine learning approach (Iowa, USA): A machine learning approach. *Stochastic Environmental Research and Risk Assessment*, *36*(11), 3835–3855. https://doi.org/10.1007/s00477-022-02228-9

Eslami, E., Salman, A. K., Choi, Y., Sayeed, A., & Lops, Y. (2020). A data ensemble approach for real-time air quality forecasting using extremely randomized trees and deep neural networks. *Neural Computing and Applications*, *32*(11), 7563–7579. https://doi.org/10.1007/s00521-019-04287-6

Farhana, N., Firdaus, A., Darmawan, M. F., & Ab Razak, M. F. (2023). Evaluation of Boruta algorithm in DDoS detection. *Egyptian Informatics Journal*, *24*, 27–42. https://doi.org/10.1016/j.eij.2022.10.005

Geurts, P., Ernst, D., & Wehenkel, L. (2006). Extremely randomized trees. *Machine Learning*, *63*(1), 3–42. https://doi.org/10.1007/s10994-006-6226-1

Ghimire, S., Yaseen, Z. M., Farooque, A. A., Deo, R. C., Zhang, J., & Tao, X. (2021). Streamflow prediction using an integrated methodology based on convolutional neural network and long short-term memory networks. *Scientific Reports*, *11*(1), 17497. https://doi.org/10.1038/s41598-021-96751-4

Gizaw, M. S., & Gan, T. Y. (2016). Regional flood frequency analysis using support vector regression under historical and future climate. *Journal of Hydrology*, *538*, 387–398. https://doi.org/10.1016/j.jhydrol.2016.04.041

Goliatt, L., Sulaiman, S. O., Khedher, K. M., Farooque, A. A., & Yaseen, Z. M. (2021). Estimation of natural streams longitudinal dispersion coefficient using hybrid evolutionary machine learning model. *Engineering Applications of Computational Fluid Mechanics*, *15*, 1298–1320. https://doi.org/10.1080/19942060.2021.1972043

Golshan, M., Kavian, A., Esmali, A., & Ziegler, A. D. (2020). Runoff and sediment yield modeling in data-sparse catchments in the Garehsoo River basin, northern Iran. *Environmental Earth Sciences*, *79*(14), 351. https://doi.org/10.1007/s12665-020-09084-2

He, K., Zhang, X., Ren, S., & Sun, J. (2016, December). Deep residual learning for image recognition. In *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition* (pp. 770–778). https://doi.org/10.1109/CVPR.2016.90.

He, X., Luo, J., Zuo, G., & Xie, J. (2019). Daily runoff forecasting using a hybrid model based on variational mode decomposition and deep neural networks. *Water Resources Management*, *33*(4), 1571–1590. https://doi.org/10.1007/s11269-019-2183-x

Ho, E., Tsuji, L. J. S., & Gough, W. A. (2005). Trends in river-ice break-up data for the western James Bay region of Canada. *Polar Geography*, *29*(4), 291–299. https://doi.org/10.1080/789610144

Htike, K. K. (2017). Efficient determination of the number of weak learners in AdaBoost. *Journal of Experimental & Theoretical Artificial Intelligence*, *29*(5), 967–982. https://doi.org/10.1080/0952813X.2016.1266038

Hu, C., Wu, Q., Li, H., Jian, S., Li, N., & Lou, Z. (2018). Deep learning with a long short-term memory networks approach for rainfall-runoff simulation. *Water (Switzerland)*, *10*(11), 1543. https://doi.org/10.3390/w10111543

Huang, N. E., Shen, Z., Long, S. R., Wu, M. C., Shih, H. H., Yen, N., 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. https://doi.org/10.1098/rspa.1998.0193

Hussain, D., Hussain, T., Khan, A. A., Naqvi, S. A. A., & Jamil, A. (2020). A deep learning approach for hydrological time-series prediction: A case study of Gilgit river basin. *Earth Science Informatics*, *13*(3), 915–927. https://doi.org/10.1007/s12145-020-00477-2

Idris, A., Khan, A., & Lee, Y. S. (2012). Genetic Programming and Adaboosting based churn prediction for Telecom. *Proceedings of the IEEE International Conference on Systems, Man and Cybernetics*, 1328–1332. https://doi.org/10.1109/ICSMC.2012.6377917

Jamei, M., Ali, M., Karimi, B., Karbasi, M., Farooque, A. A., & Yaseen, Z. M. (2023). Surface water electrical conductivity and bicarbonate ion determination using a smart hybridization of optimal Boruta package with Elman recurrent neural network. *Process Safety and Environmental Protection*, *174*, 115–134. https://doi.org/10.1016/j.psep.2023.03.062

Jose, D. M., Vincent, A. M., & Dwarakish, G. S. (2022). Management of validation of HPLC method for determination of acetylsalicylic acid impurities in a new pharmaceutical product. *Scientific Reports*, *12*(1), 1–25. https://doi.org/10.1038/s41598-021-99269-x

Kalteh, A. M. (2013). Monthly river flow forecasting using artificial neural network and support vector regression models coupled with wavelet transform. *Computers & Geosciences*, *54*, 1–8. https://doi.org/10.1016/j.cageo.2012.11.015

Kamath, P. R., & Senapati, K. (2021). Short-term wind speed forecasting using S-transform with compactly supported kernel. *Wind Energy*, *24*(3), 260–274. https://doi.org/10.1002/we.2571

Karthikeyan, L., & Nagesh Kumar, D. (2013). Predictability of nonstationary time series using wavelet and EMD based ARMA models. *Journal of Hydrology*, *502*, 103–119. https://doi.org/10.1016/j.jhydrol.2013.08.030

Kolachian, R., & Saghafian, B. (2021). Hydrological drought class early warning using support vector machines and rough sets. *Environmental Earth Sciences*, *80*(11), 390. https://doi.org/10.1007/s12665-021-09536-3

Kratzert, F., Klotz, D., Brenner, C., Schulz, K., & Herrnegger, M. (2018). Rainfall-runoff modelling using long short-term memory (LSTM) networks. *Hydrology and Earth System Sciences*, 22(11), 6005–6022. https://doi.org/10.5194/hess-22-6005-2018

Krause, P., Smith, A., Veale, B., & Murray, M. (2001). Achievements of the grand river conservation authority, Ontario, Canada. *Water Science and Technology*, 43(9), 45–55. https://doi.org/10.2166/wst.2001.0506

Kumar, P., Prasad, R., Choudhary, A., Gupta, D. K., Mishra, V. N., Vishwakarma, A. K., Singh, A. K., & Srivastava, P. K. (2019). Comprehensive evaluation of soil moisture retrieval models under different crop cover types using C-band synthetic aperture radar data. *Geocarto International*, 34(9), 1022–1041. https://doi.org/10.1080/10106049.2018.1464601

Kursa, M. B., & Rudnicki, W. R. (2010). Feature Selection with the Boruta Package. *Journal of Statistical Software*, 36), https://doi.org/10.18637/jss.v036.i11

LeCun, Y., Bottou, L., Bengio, Y., & Haffner, P. (1998). Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, 86(11), 2278–2324. https://doi.org/10.1109/5.726791

Lee, T., & Ouarda, T. B. M. J. (2012). Stochastic simulation of nonstationary oscillation hydroclimatic processes using empirical mode decomposition. *Water Resources Research*, 48(2), https://doi.org/10.1029/2011WR010660

Liu, S., Wang, J., Wang, H., & Wu, Y. (2022). Post-processing of hydrological model simulations using the convolutional neural network and support vector regression. *Hydrology Research*, 53(4), 605–621. https://doi.org/10.2166/nh.2022.004

Liu, S., Xu, J., Zhao, J., Xie, X., & Zhang, W. (2014). Efficiency enhancement of a process-based rainfall-runoff model using a new modified AdaBoost.RT technique. *Applied Soft Computing*, 23, 521–529. https://doi.org/10.1016/j.asoc.2014.05.033

Luo, C., Zhang, X., Wang, Y., Men, Z., & Liu, H. (2022). Regional soil organic matter mapping models based on the optimal time window, feature selection algorithm and Google Earth Engine. *Soil and Tillage Research*, 219, 105325. https://doi.org/10.1016/j.still.2022.105325

Maguire, T., Manuel, L., Smedinga, R. A., & Biehl, M. (2022). A review of feature selection and ranking methods. In R. Smedinga & M. Biehl (Eds.), *Proceedings of the 19th SC@RUG 2022 Proceedings 2021–2022* (pp. 15–20). Groningen: Rijksuniversiteit Groningen. https://pure.rug.nl/ws/portalfiles/portal/214074117/proceedings_2022.pdf

Mallick, J., Talukdar, S., & Ahmed, M. (2022). Combining high resolution input and stacking ensemble machine learning algorithms for developing robust groundwater potentiality models in Bisha watershed, Saudi Arabia. *Applied Water Science*, 12(4), 77. https://doi.org/10.1007/s13201-022-01599-2

Meddage, D. P. P., Ekanayake, I. U., Weerasuriya, A. U., & Lewangamage, C. S. (2021). Tree-based regression models for predicting external wind pressure of a building with an unconventional configuration. In *Proceedings of MERCon 2021-7th International Multidisciplinary Moratuwa Engineering Research Conference* (pp. 257–262). https://doi.org/10.1109/MERCon52712.2021.9525734

Meng, E., Huang, S., Huang, Q., Fang, W., Wu, L., & Wang, L. (2019). A robust method for non-stationary streamflow prediction based on improved EMD-SVM model. *Journal of Hydrology*, 568, 462–478. https://doi.org/10.1016/j.jhydrol.2018.11.015

Metzger, R. A., Doherty, J. F., Jenkins, D. M., & Hall, D. L. (2020). Approximate entropy and empirical mode decomposition for improved speaker recognition. *Advances in Data Science and Adaptive Analysis*, 12(03n04), 2050011. https://doi.org/10.1142/S2424922X20500114

Mohammadi, B. (2021). A review on the applications of machine learning for runoff modeling. *Sustainable Water Resources Management*, 7(6), 98. https://doi.org/10.1007/s40899-021-00584-y

Mozaffari, S., Javadi, S., Moghaddam, H. K., & Randhir, T. O. (2022). Forecasting groundwater levels using a hybrid of support vector regression and particle swarm optimization. *Water Resources Management*, 36(6), 1955–1972. https://doi.org/10.1007/s11269-022-03118-z

Nasteski, V. (2017). An overview of the supervised machine learning methods. *Horizons. B*, 4, 51–62. https://doi.org/10.20544/HORIZONS.B.04.1.17.P05

Parisouj, P., Mohebzadeh, H., & Lee, T. (2020). Employing machine learning algorithms for streamflow prediction: A case study of four river basins with different climatic zones in the United States. *Water Resources Management*, 34(13), 4113–4131. https://doi.org/10.1007/s11269-020-02659-5

Parisouj, P., Mokari, E., Mohebzadeh, H., Goharnejad, H., Jun, C., Oh, J., & Bateni, S. M. (2022). Physics-informed data-driven model for predicting streamflow: A case study of the Voshmgir Basin, Iran. *Applied Sciences*, 12(15), 7464. https://doi.org/10.3390/app12157464

Pekárová, P., Miklánek, P., & Pekár, J. (2003). Spatial and temporal runoff oscillation analysis of the main rivers of the world during the 19th-20th centuries. *Journal of Hydrology*, 274(1-4), 62–79. https://doi.org/10.1016/S0022-1694(02)00397-9

Prasad, R., Ali, M., Kwan, P., & Khan, H. (2019). Designing a multi-stage multivariate empirical mode decomposition coupled with ant colony optimization and random forest model to forecast monthly solar radiation. *Applied Energy*, 236, 778–792. https://doi.org/10.1016/j.apenergy.2018.12.034

Sadeghi, M., Asanjan, A. A., Faridzad, M., Nguyen, P. H. U., Hsu, K., Sorooshian, S., & Braithwaite, D. A. N. (2019). PERSIANN-CNN: Precipitation estimation from remotely sensed information using artificial neural networks–convolutional neural networks. *Journal of Hydrometeorology*, 20(12), 2273–2289. https://doi.org/10.1175/JHM-D-19-0110.1

Sagi, O., & Rokach, L. (2018). Ensemble learning: A survey. *WIRES Data Mining and Knowledge Discovery*, 8(4), e1249. https://doi.org/10.1002/widm.1249

Sánchez-Martínez, A., Ruíz-Oropeza, E. Y., Orozco-del-Castillo, M. G., Hernández-Gómez, J. J., & Yáñez-Casas, G. A. (2022). Assessment of the reduction of the icesnow coverage at the TransMexican Volcanic Belt through empirical mode decomposition on satellite imagery. In *Advances in geospatial data science: Selected papers from the international conference on geospatial information sciences 2021* (pp. 131–148). Springer International Publishing. https://doi.org/10.1007/978-3-030-98096-2_10.

Shrestha, N. K., & Shukla, S. (2015). Support vector machine based modeling of evapotranspiration using hydro-climatic variables in a sub-tropical environment. *Agricultural and*

Forest Meteorology, *200*, 172–184. https://doi.org/10.1016/j.agrformet.2014.09.025

Sibtain, M., Li, X., Bashir, H., & Azam, M. I. (2021). A hybrid model for runoff prediction using variational mode decomposition and artificial neural network. *Water Resources*, *48*(5), 701–712. https://doi.org/10.1134/S0097807821050171

Simonyan, K., & Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. *ArXiv Preprint ArXiv*, 1409–1556. https://arxiv.org/abs/1409.1556.

Singh, U. K., Jamei, M., Karbasi, M., Malik, A., & Pandey, M. (2022). Application of a modern multi-level ensemble approach for the estimation of critical shear stress in cohesive sediment mixture. *Journal of Hydrology*, *607*, 127549. https://doi.org/10.1016/j.jhydrol.2022.127549

Soltani, K., Ebtehaj, I., Amiri, A., Azari, A., Gharabaghi, B., & Bonakdari, H. (2021). Mapping the spatial and temporal variability of flood susceptibility using remotely sensed normalized difference vegetation index and the forecasted changes in the future. *Science of the Total Environment*, *770*, 145288. https://doi.org/10.1016/j.scitotenv.2021.145288

St. George, S. (2007). Streamflow in the Winnipeg River basin, Canada: Trends, extremes and climate linkages. *Journal of Hydrology*, *332*(3-4), 396–411. https://doi.org/10.1016/j.jhydrol.2006.07.014

Story, A., & Buttle, J. M. (2001). Precipitation data quality and long-term water balances within the Moose River Basin, east-central Canada. *Atmosphere-Ocean*, *39*(1), 55–69. https://doi.org/10.1080/07055900.2001.9649666

Sudheer, K. P., Gosain, A. K., & Ramasastri, K. S. (2002). A data-driven algorithm for constructing artificial neural network rainfall-runoff models. *Hydrological Processes*, *16*(6), 1325–1330. https://doi.org/10.1002/hyp.554

Sun, K., Hu, L., Guo, J., Yang, Z., Zhai, Y., & Zhang, S. (2021). Enhancing the understanding of hydrological responses induced by ecological water replenishment using improved machine learning models: A case study in Yongding River. *Science of the Total Environment*, *768*, 145489. https://doi.org/10.1016/j.scitotenv.2021.145489

Szegedy, C., Ioffe, S., Vanhoucke, V., & Alemi, A. A. (2017). Inception-v4, inception-ResNet and the impact of residual connections on learning. *Proceedings of the AAAI Conference on Artificial Intelligence*, *31*, https://doi.org/10.1609/aaai.v31i1.11231

Szegedy, C., Vanhoucke, V., Ioffe, S., Shlens, J., & Wojna, Z. (2016). Rethinking the inception architecture for computer vision. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (pp. 2818–2826). https://doi.org/10.1109/CVPR.2016.308.

Tarfaya, C., Houichi, L., & Heddam, S. (2022). Prediction of index rainfall in ungauged regions of Algeria: Survey of rule-based models using geographic predictors. *Arabian Journal of Geosciences*, *15*(7), 668. https://doi.org/10.1007/s12517-022-09944-0

Tayyab, M., Ahmad, I., Sun, N., Zhou, J., & Dong, X. (2018). Application of integrated artificial neural networks based on decomposition methods to predict streamflow at Upper Indus Basin, Pakistan. *Atmosphere*, *9*(12), 494. https://doi.org/10.3390/atmos9120494

Tu, T., Ishida, K., Ercan, A., Kiyama, M., Amagasaki, M., & Zhao, T. (2021). Hybrid precipitation downscaling over coastal watersheds in Japan using WRF and CNN. *Journal of Hydrology: Regional Studies*, *37*, 100921. https://doi.org/10.1016/j.ejrh.2021.100921

Tyralis, H., Papacharalampous, G., & Langousis, A. (2021). Super ensemble learning for daily streamflow forecasting: Large-scale demonstration and comparison with multiple machine learning algorithms. *Neural Computing and Applications*, *33*(8), 3053–3068. https://doi.org/10.1007/s00521-020-05172-3

Wang, W.-C., Chau, K.-W., Qiu, L., & Chen, Y.-B. (2015). Improving forecasting accuracy of medium and long-term runoff using artificial neural network based on EEMD decomposition. *Environmental Research*, *139*, 46–54. https://doi.org/10.1016/j.envres.2015.02.002

Weeks, W. D., & Boughton, W. C. (1987). Tests of ARMA model forms for rainfall-runoff modelling. *Journal of Hydrology*, *91*(1-2), 29–47. https://doi.org/10.1016/0022-1694(87)90126-0

Williams, G. R. (1961). Cyclical variations in world-wide hydrologic data. *Journal of the Hydraulics Division*, *87*(6), 71–88. https://doi.org/10.1061/JYCEAJ.0000668

Wolpert, D. H. (1992). Stacked generalization. *Neural Networks*, *5*(2), 241–259. https://doi.org/10.1016/S0893-6080(05)80023-1

Yuan, R., Cai, S., Liao, W., Lei, X., Zhang, Y., Yin, Z., Ding, G., Wang, J., & Xu, Y. (2021). Daily runoff forecasting using ensemble empirical mode decomposition and long short-term memory. *Frontiers in Earth Science*, *9*, 621780. https://doi.org/10.3389/feart.2021.621780

Zhang, C., & Ma, Y. (2012). *Ensemble machine learning: Methods and applications*. Springer US. https://doi.org/10.1007/978-1-4419-9326-7

Zhang, Y., Zhang, C., Sun, J., & Guo, J. (2018). Improved wind speed prediction using empirical mode decomposition. *Advances in Electrical and Computer Engineering*, *18*(2), 3–10. https://doi.org/10.4316/AECE.2018.02001

Zhao, Y., Meng, X., Qi, T., Li, Y., Chen, G., Yue, D., & Qing, F. (2022). AI-based rainfall prediction model for debris flows. *Engineering Geology*, *296*, 106456. https://doi.org/10.1016/j.enggeo.2021.106456

Zhu, J., & Pierskalla, W. P. (2016). Applying a weighted random forests method to extract karst sinkholes from LiDAR data. *Journal of Hydrology*, *533*, 343–352. https://doi.org/10.1016/j.jhydrol.2015.12.012