METHODOLOGIES AND APPLICATION

Development and evaluation of the cascade correlation neural network and the random forest models for river stage and river flow prediction in Australia

Mohammad Ali Ghorbani[1] · Ravinesh C. Deo[2] · Sungwon Kim[3] · Mahsa Hasanpour Kashani[4] ·
Vahid Karimi[5] · Maryam Izadkhah[5]

Abstract
Accurately predicting river flows over daily timescales is considered as an important task for sustainable management of freshwater ecosystems, agricultural applications, and water resources management. In this research paper, artificial intelligence (AI) techniques, namely the cascade correlation neural networks (CCNN) and the random forest (RF) models, were employed in daily river stage and river flow prediction for two river systems (i.e., Dulhunty River and Herbert River) in Australia. To develop the CCNN and RF models, a significant 3-day antecedent river stage and river flow time series were used. 80% of the whole data were used for model training and the remaining 20% for model testing. A total of ten different model structures with different input combinations were used to evaluate the optimal model in the training phase, and the results were analyzed using statistical metrics including the root mean square error (RMSE), Nash–Sutcliffe coefficient (NS), Willmott's index of agreement (WI), and Legate and McCabe's index (E_{LM}) in the testing phase. The inter-comparison of CCNN and RF models for both river systems showed that the CCNN model was able to generate a more accurate prediction of the river stage and river flow compared to the RF model. Due to hydro-geographic differences leading to a different underlying historical data characteristics, the optimal CCNN's performance for the Dulhunty River was found to be most accurate, in terms of $E_{LM} = 0.779$, WI = 0.964, and $E_{NS} = 0.862$ versus 0.775, 0.968, and 0.885 for the Herbert River. Following the performance accuracies, the authors ascertained that the CCNN model can be taken as a preferred data intelligent tool for river stage and river flow prediction.

Keywords Australia · Cascade correlation neural networks · Prediction · Random forest · River flow

1 Introduction

River flow modeling plays a significant role in water resources management including flood control, hydropower generation, watershed management, reservoir operation,

✉ Mohammad Ali Ghorbani
mohammadalighorbani@tdtu.edu.vn

1 Sustainable Management of Natural Resources and Environment Research Group, Faculty of Environment and Labour Safety, Ton Duc Thang University, Ho Chi Minh City, Vietnam

2 School of Sciences, Centre for Applied Climate Sciences and Centre for Sustainable Agricultural Systems, University of Southern Queensland, Springfield, QLD 4300, Australia

3 Department of Railroad Construction and Safety Engineering, Dongyang University, Yeongju, Republic of Korea

4 Department of Water Engineering, Faculty of Agriculture and Natural Resources, University of Mohaghegh Ardabili, Ardabil, Iran

5 Department of Water Engineering, University of Tabriz, Tabriz, Iran

and river sediment management (Khatibi et al. 2012). The ability to generate accurate river flow rates based on antecedent (i.e., historical) time series and minimum site information requirement using an artificial intelligence (AI) technique is becoming a popular modeling choice of river system forecasters (Nayak et al. 2004). AI techniques are simple to formulate, yet they are quite robust and can handle complex and nonlinear biophysical processes embedded in hydrological data series. These techniques have a major advantage compared to the parametric modeling techniques since they do not require a priori concept on the relationships between the input variables and output data (Gocić et al. 2015), and they are mathematically less complex to design and implement.

Nowadays, a variety of AI techniques have been implemented in river flow prediction. These models include, among the others, the artificial neural networks (Tawfik et al. 1997; Jain and Chalisgaonkar 2000; Deka and Chandramouli 2003; Sudheer and Jain 2003; Wu et al. 2005; Bhattacharya and Solomatine 2005; Habib and Meselhe 2006; Alvisi et al. 2006; Chen et al. 2006; Clemmens and Wahlin 2006; Kişi 2007; Baiamonte and Ferro 2007; Firat 2008; Khatibi et al. 2011; Aggarwal et al. 2012; Ajmera and Goyal 2012; Hasanpour Kashani et al. 2015; Ghorbani et al. 2016a; Khatibi et al. 2017; Zhang et al. 2018; Kashani and Soltangheys 2018), fuzzy logic (Alvisi et al. 2006), M5 model trees (Bhattacharya and Solomatine 2005; Ajmera and Goyal 2012), adaptive neuro-fuzzy inference system (ANFIS) (Khatibi et al. 2011; Hasanpour Kashani et al. 2015; Zhang et al. 2018; Kashani and Soltangheys 2018), support vector machines (Sivapragasam and Muttil 2005; Aggarwal et al. 2012; Goel and Pal 2012; Ghorbani et al. 2016b; Zhang et al. 2018), genetic programming (Ghimire and Reddy 2010; Khatibi et al. 2011; Liu and Chung 2014; Hasanpour Kashani et al. 2015; Zhang et al. 2018), chaos theory (Khatibi et al. 2012), hybrid wavelet gene expression programming (WGEP) (Shoaib et al. 2015), wavelet artificial neural networks (WANN), hybrid ANN-PSO and ANN-GA (Zhang et al. 2018), hybrid SVM-QPSO (Ch et al. 2013), hybrid NNRF-PSO (Taormina and Chau 2014), hybrid ELM-BSO (Taormina et al. 2015), hybrid ANN-IIS-W (Prasad et al. 2017), and hybrid MLP-FFA (Khatibi et al. 2017). These research works have revealed an acceptable level of accuracy for various water management and water flow prediction tasks that were attained with many categories of AI techniques. In addition to the above-mentioned AI techniques, some new hybrid intelligent methods have been developed (Deng et al. 2017a, b; Zhao et al. 2017; Deng et al. 2018, 2019a, b) and can be applied for accurate modeling of water management issues.

In this paper, the authors utilize a suite of less-explored AI techniques, known as the cascade correlation neural network (CCNN) and random forest (RF) models. The ability of the CCNN model has led to some investigators applying the method for river flow prediction. For example, the study of Karunanithi et al. (1994) has applied the CCNN model for the prediction of the Huron River flow at the Dexter sampling station, and the study of Diamantopoulou et al. (2007) has developed the CCNN model with Kalman learning rule to forecast the 1-day ahead daily flow at Ilarionas station on the Aliakmon River, Northern Greece. In another study, Alok et al. (2013) used two neural networks models, known as Elman neural networks (ENN) and CCNN, to estimate the data of the Indian River Brahmani, revealing the good ability of the CCNN model to evaluate the river flow time series.

To benchmark the CCNN model, the random forest (RF) model, nonparametric regression and rule-based algorithm, is also used in this study. The RF model entails the predictions that are created by averaging the predicted values from multiple regression (or decision) trees, which are trained on separate bootstrapped resamples of the training dataset. Zhao et al. (2012) introduced the RF model for selecting predictor dataset from measured streamflow and a total of 74 hydro-climatic indices to predict seasonal streamflow, while the study of Nguyen et al. (2015) has investigated the application of the RF model, least absolute shrinkage and selection operator (LASSO) and support vector regression (SVR) for forecasting of daily water levels at Thakhek station on the Mekong River. Shortridge et al. (2016) used the ANN, RF, and other models to simulate monthly streamflow of five highly seasonal rivers in the highlands of Ethiopia.

Although there have been different research works to predict water stage and river flow using ANN-based models, the applications of the CCNN and RF models, as attempted in this paper, have been relatively limited compared to the conventional ANN-based models. In view of this deficit from current knowledge, the novelty of present study is to design for the first time and to investigate the capability of CCNN and RF models for river flow prediction and also to broaden their usage in hydrologic time series predictions where modeling data for two major river systems in Australia (i.e., Dulhunty and Herbert Rivers) are employed.

2 Methodology

2.1 Cascade correlation neural networks (CCNN)

The CCNN model, developed by Fahlman and Lebiere (1990), is a special case of the artificial neural networks (ANN) model where a parallel information processing system is used, consisting of a set of neurons arranged in

input, hidden, and output layers. These neurons provide suitable conversion functions for weighted inputs. The CCNN model, as opposed to classical ANN models, can automatically synthesize the best neural networks architecture as part of the model's training process. The CCNN model is an efficient constructive neural networks model combining the idea of incremental structure and learning algorithm during its training. Training starts with minimal networks consisting of an input and an output layer without a hidden layer. If the training can no longer reduce the residual error, then this phase of training is stopped and enters the next phase for the training of the potential hidden node. The potential hidden node has associated connection weights from the input layer and all preexisting hidden nodes but not toward the output layer (Karunanithi et al. 1994; Diamantopoulou et al. 2007; Kim et al. 2014; Zounemat-Kermani et al. 2019).

In a CCNN model, the connection weights associated with the potential hidden nodes are optimized by the gradient ascent method to maximize the correlation between its output and the residual error of the CCNN model. When a potential hidden node is trained, connection weights associated with the output layer are kept unchanged. When a potential hidden node is added to the CCNN's structure, it becomes a new hidden node, and its incoming connection weights are fixed for the remainder of training. After installing a hidden node successfully, the training updates all of the connection weights, which directly feed the output layer. The CCNN model automatically constructs a suitable structure for a given problem (Karunanithi et al. 1994; Thirumalaiah and Deo 1998). Figure 1a shows the structure of the CCNN model.

2.2 Random forest (RF)

The RF model consists of an ensemble of randomized regression trees that operate by constructing a multitude of regression trees and then aggregating them to yield a single prediction (Fig. 1b). The RF model provides reliable error estimated by using the out-of-bag (OOB) data and estimates covariate importance by changing the order of arrangement/arranging the values of each covariate in the OOB sample and predicting OOB samples using the permuted variable (Zhao et al. 2012). The change in OOB error is then an indication of the importance of that covariate in the dataset. The RF model depends only on two user-defined parameters, the mtry (the number of ancillary data in each random subset) and ntree (the number of trees in the forest), which are optimized by iterating mtry values from 1 to 10 (the total number of covariates) and ntree values from 100 to 10,000 by increments of 100 (Hengl et al. 2015; Were et al. 2015).

2.3 Study area, data available, and performance criteria

The daily mean river stage and river flow time series data of Dulhunty and Herbert Rivers in Australia were used in this study (Table 1). The data are arbitrarily divided into two parts for training and testing. The training datasets were chosen at 80% of the length of the time series (from 2014/05/22 up to 2016/10/14), and the testing datasets covered their remaining 20% (from 2016/10/15 up to 2017/05/21). Figure 2a, b shows a plot of the observed daily river stage and discharge, respectively, with training and testing period. The statistical parameters of river stage and discharge data are given in Table 2. In the table, the X_{mean}, S_x, C_v, C_{sx}, X_{max}, and X_{min} denote the mean, standard deviation, variation coefficient, skewness, maximum, and minimum, respectively.

To validate the performance of the models, diagnostic plots and statistical score metrics were employed in the testing phase. These metrics are described as:

I. Root mean square error (RMSE) (Willmott and Matsuura 2005) is expressed as:

Fig. 1 Structure of the models: a CCNN model (Kim et al. 2014) and b RF model (Rodriguez-Galiano and Atkinson 2016)

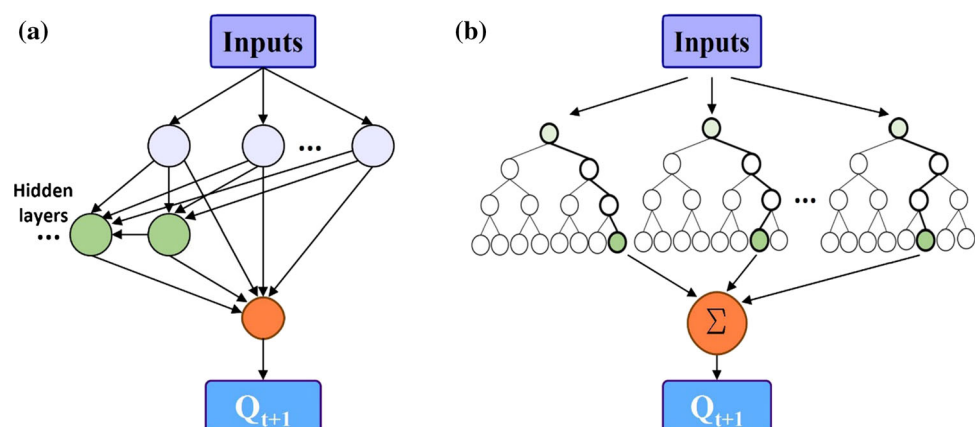**Table 1** Physiographical characteristics of river stage and discharge measurement stations

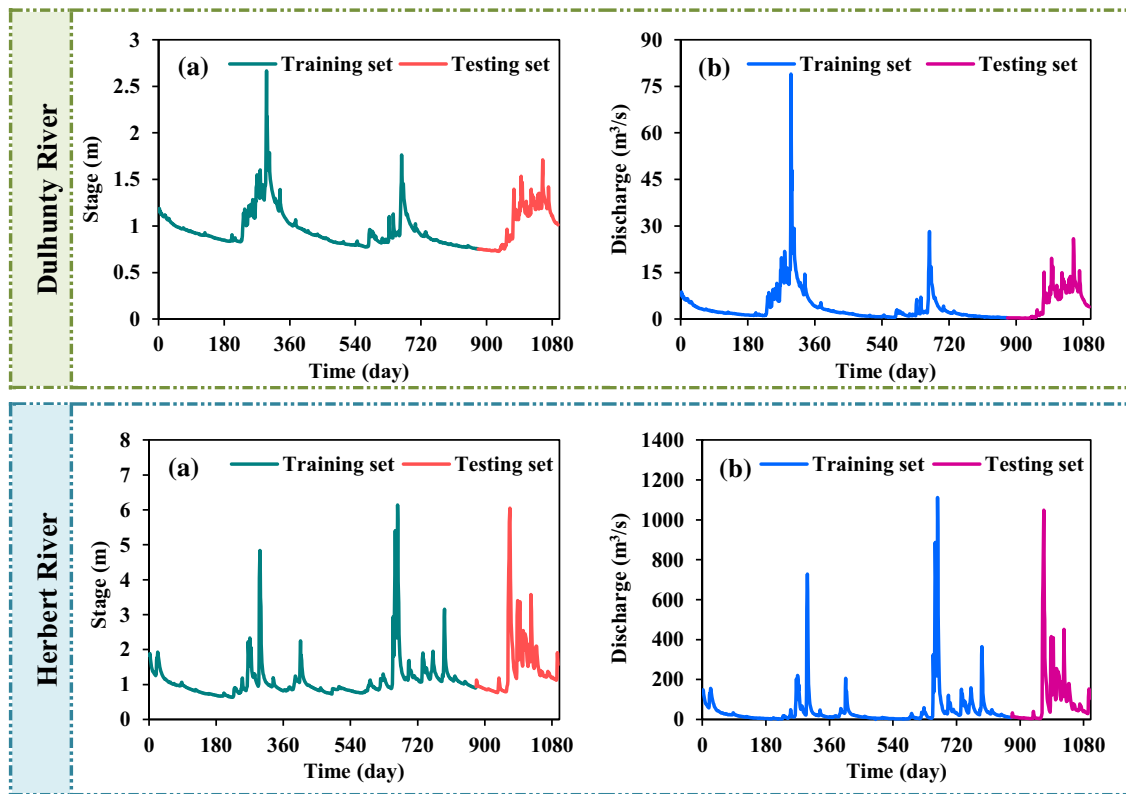| River	Latitude (decimal degree)	Longitude (decimal degree)	Elevation (m)	Record period	Number of data
Dulhunty	11.832684	142.421975	51	2014/05/22–2017/05/21	1096
Herbert	18.63275	146.14267	16.5	2014/05/22–2017/05/21	1096

Fig. 2 Time series plot for the observed stage and discharge data with a period of 2014/05/22 to 2017/05/21): **a** Dulhunty River and **b** Herbert River

Table 2 Statistical characteristics of the river daily stage and discharge data sets

River	Dataset	Data type	Number of data	X_{mean}	S_x	C_v	C_{sx}	X_{max}	X_{min}
Dulhunty	Training	Stage (m)	877	0.95	0.19	0.20	2.62	2.67	0.75
		Discharge (m³/s)	877	3.67	5.34	1.45	5.63	78.95	0.42
	Testing	Stage (m)	219	1.01	0.23	0.23	0.17	1.71	0.73
		Discharge (m³/s)	219	5.40	4.95	0.91	0.75	25.91	0.29
Herbert	Training	Stage (m)	877	1.10	0.54	0.49	4.69	6.14	0.63
		Discharge (m³/s)	877	42.84	90.76	2.12	6.57	1111.78	0.00
	Testing	Stage (m)	219	1.47	0.84	0.57	2.54	6.05	0.77
		Discharge (m³/s)	219	96.94	147.53	1.52	3.44	1048.79	3.64

$$\text{RMSE} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(P_i - O_i)^2} \qquad (1)$$

$$E_{NS} = 1 - \left[\frac{\sum_{i=1}^{N}(O_i - P_i)^2}{\sum_{i=1}^{N}\left(O_i - \overline{O_i}\right)^2}\right], \quad -\infty < E_{NS} \le 1$$

II. Nash–Sutcliffe coefficient (E_{NS}) (Nash and Sutcliffe 1970) is expressed as:

$$(2)$$

III. Willmott's index of agreement (WI) (Willmott et al. 2012) is expressed as:

$$WI = 1 - \left[\frac{\sum_{i=1}^{N} (O_i - P_i)^2}{\sum_{i=1}^{N} \left(|P_i - \overline{O_i}| + |O_i - \overline{O_i}| \right)^2} \right],$$
$$0 < WI \leq 1 \tag{3}$$

IV. Legate and McCabe's index (E_{LM}) (Legates and Davis 1997; Legates and McCabe 1999, 2013) is expressed as

$$E_{LM} = 1 - \left[\frac{\sum_{i=1}^{N} |O_i - P_i|}{\sum_{i=1}^{N} |O_i - \overline{O_i}|} \right], \quad 0 < E_{LM} \leq 1 \tag{4}$$

where O_i and P_i are the observed and predicted ith value of the Q, and $\bar{O}$ is the average of observed Q value.

3 Results and discussion

This study has used daily river stage and river flow data from the Dulhunty and Herbert Rivers in Australia. As stated previously, the evaluation and comparison for the performance of CCNN and RF models are the core point of the paper based on the prediction of daily river flow.

3.1 Development of CCNN model

Various combinations of river stage and river flow variables for the Dulhunty and Herbert Rivers were applied as the input variables for CCNN and RF models to determine the best input variables. Therefore, the CCNN and RF models are a priori fed with water stage (H). It was adopted as the minimum number of input combinations represented by the CCNN 1 and RF 1 for choosing the optimal input combinations. Various input combinations of river stage and river flow variables to estimate the river flow are shown in Table 3.

A difficult task with ANN modeling is to choose its optimal architecture and determine the number of hidden layers and nodes using a trial and error method. The network geometry and architecture rely on the addressed problem (Kişi 2007; Kim et al. 2012; Seo et al. 2015). This study started using one hidden layer for the construction of the CCNN model since one hidden layer can be enough to represent the specific nonlinear relationships (Kumar et al. 2002). The number of hidden nodes was determined using a trial and error method for the CCNN model with the different input combinations based on the statistical criteria.

Table 4 shows a summary of the statistical indices of each CCNN model during the training and test phases. It can be given from Table 4 that the CCNN 6 model, whose input variables are flow discharges at times t and $t - 1$ (Q_t, Q_{t-1}), produced the most accurate results among the other input combinations for the Dulhunty River. Moreover, the CCNN 10 model, whose input variables are river stage and river flow at times t until $t - 3$ (H_t, H_{t-1}, H_{t-2}, H_{t-3}, Q_t, Q_{t-1}, Q_{t-2}, Q_{t-3}), provided the most accurate results among the other input combinations for the Herbert River. Here, the optimum structure of the CCNN 6 (2, 3, 1) denotes a CCNN model comprising two input, three hidden, and one output nodes, respectively. Also, the optimum structure of the CCNN 10 (8, 1, 1) denotes a CCNN model comprising eight input, one hidden, and one output nodes, respectively. Figure 3a, b shows observed and predicted river flow values and their corresponding scatter plots during the test phase for the CCNN 6 and CCNN 10 models for the Dulhunty and Herbert Rivers, respectively. Figure 3 shows that although there can be found little errors in peak flows prediction, the CCNN 6 and CCNN 10 models can estimate nonlinear river stage and river flow values efficiently. This is in agreement with the previous reports provided by Alok et al. (2013).

Table 3 Different input combinations of the models

No.	Input combination	Output	RF	CCNN
1	H_t	Q_{t+1}	RF1	CCNN1
2	H_t, H_{t-1}	Q_{t+1}	RF2	CCNN2
3	H_t, H_{t-1}, H_{t-2}	Q_{t+1}	RF3	CCNN3
4	$H_t, H_{t-1}, H_{t-2}, H_{t-3}$	Q_{t+1}	RF4	CCNN4
5	Q_t	Q_{t+1}	RF5	CCNN5
6	Q_t, Q_{t-1}	Q_{t+1}	RF6	CCNN6
7	Q_t, Q_{t-1}, Q_{t-2}	Q_{t+1}	RF7	CCNN7
8	$Q_t, Q_{t-1}, Q_{t-2}, Q_{t-3}$	Q_{t+1}	RF8	CCNN8
9	$H_t, H_{t-1}, Q_t, Q_{t-1}, Q_{t-2}, Q_{t-3}$	Q_{t+1}	RF9	CCNN9
10	$H_t, H_{t-1}, H_{t-2}, H_{t-3}, Q_t, Q_{t-1}, Q_{t-2}, Q_{t-3}$	Q_{t+1}	RF10	CCNN10

Table 4 Result of the CCNN model for different input combinations of the Dulhunty and Herbert Rivers

River	Model	Model structure	Training set				Testing set			
			RMSE (m^3/s)	E_{NS}	WI	E_{LM}	RMSE (m^3/s)	E_{NS}	WI	E_{LM}
Dulhunty	CCNN1	1-1-1	2.274	0.819	0.946	0.754	1.966	0.841	0.962	0.685
	CCNN2	2-3-1	2.190	0.832	0.952	0.794	1.943	0.845	0.958	0.760
	CCNN3	3-2-1	2.209	0.829	0.950	0.825	1.933	0.847	0.961	0.750
	CCNN4	4-1-1	2.235	0.825	0.948	0.784	1.959	0.842	0.961	0.719
	CCNN5	1-5-1	2.170	0.835	0.952	0.802	1.955	0.843	0.959	0.771
	CCNN6	2-3-1	2.239	0.824	0.949	0.839	1.836	0.862	0.964	0.778
	CCNN7	3-9-1	1.934	0.869	0.963	0.822	1.895	0.853	0.962	0.768
	CCNN8	4-1-1	2.268	0.820	0.947	0.830	1.833	0.862	0.963	0.765
	CCNN9	6-0-1	2.299	0.815	0.946	0.802	1.892	0.853	0.961	0.755
	CCNN10	8-0-1	2.236	0.825	0.949	0.805	1.858	0.858	0.963	0.753
Herbert	CCNN1	1-4-1	42.107	0.785	0.937	0.697	59.441	0.837	0.953	0.719
	CCNN2	2-0-1	37.689	0.827	0.952	0.634	54.551	0.863	0.960	0.744
	CCNN3	3-1-1	35.194	0.849	0.958	0.710	53.533	0.868	0.963	0.749
	CCNN4	4-1-1	35.741	0.845	0.956	0.716	53.158	0.870	0.963	0.758
	CCNN5	1-3-1	42.163	0.784	0.937	0.726	59.068	0.839	0.953	0.732
	CCNN6	2-2-1	38.221	0.822	0.950	0.736	52.223	0.874	0.964	0.769
	CCNN7	3-0-1	37.932	0.825	0.951	0.751	52.644	0.872	0.964	0.775
	CCNN8	4-1-1	36.975	0.834	0.953	0.780	51.612	0.877	0.965	0.780
	CCNN9	6-0-1	36.702	0.836	0.954	0.736	51.710	0.877	0.965	0.774
	CCNN10	8-1-1	33.389	0.865	0.963	0.761	49.925	0.885	0.968	0.775

Fig. 3 Comparative plots of the observed and predicted flow of the best CCNN models and their corresponding scatter plots during the testing phase: **a** Dulhunty River and **b** Herbert River
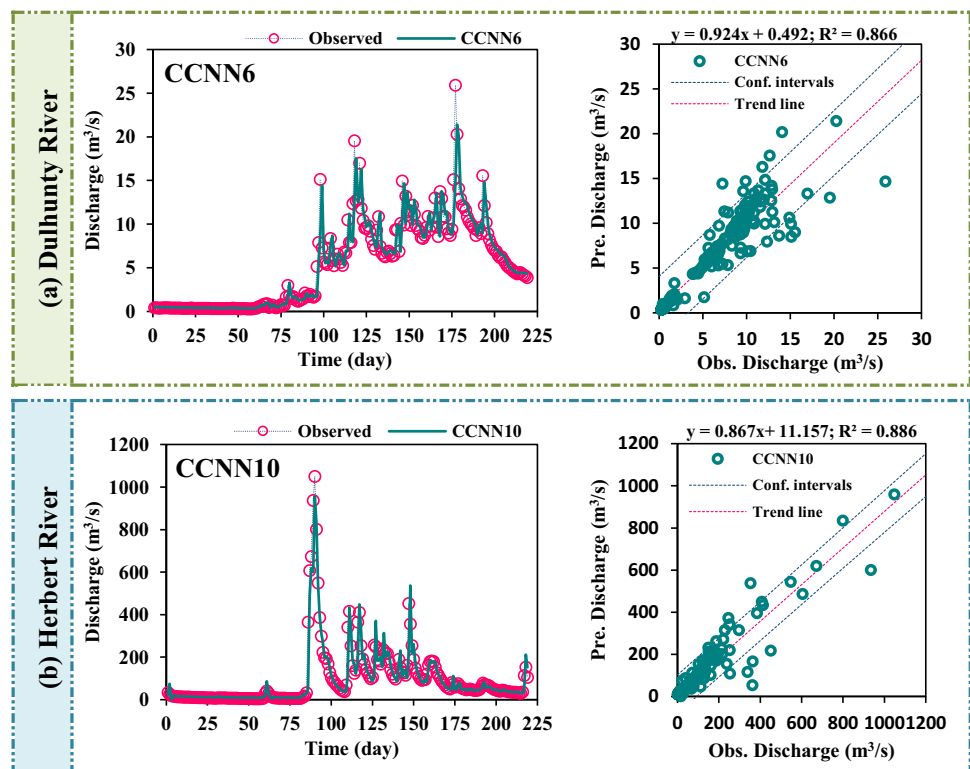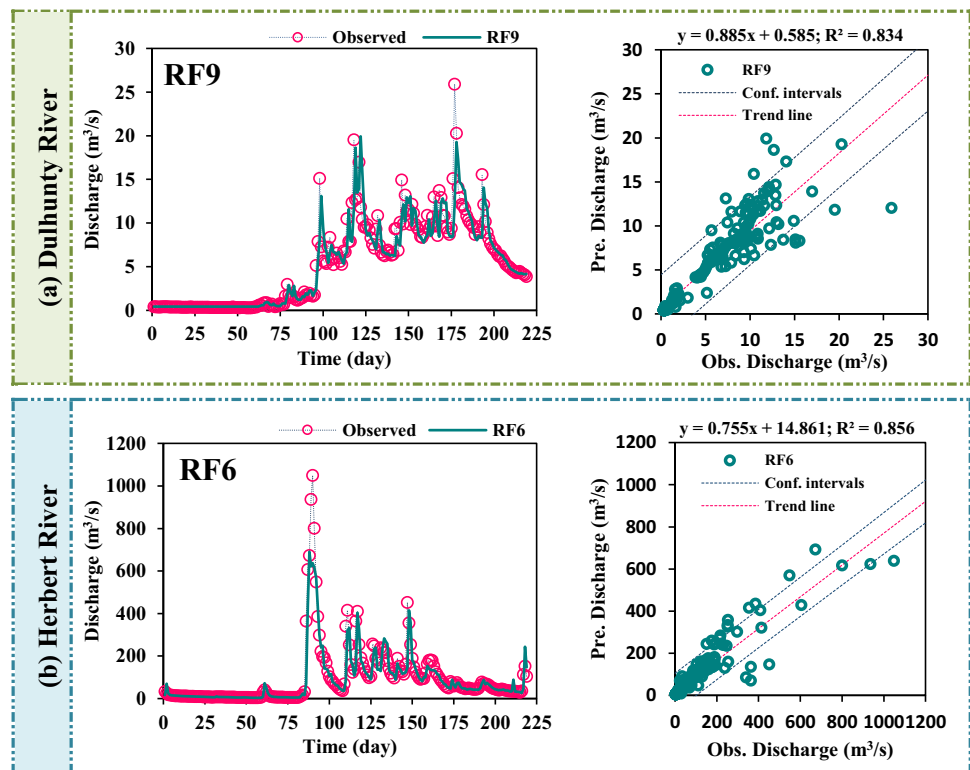
Table 5 Result of the RF model for different input combinations of the Dulhunty and Herbert Rivers

River	Model	Model structure	Training set				Testing set			
			RMSE (m³/s)	E_{NS}	WI	E_{LM}	RMSE (m³/s)	E_{NS}	WI	E_{LM}
Dulhunty	RF1	No. of trees: 52; leaf size: 5	1.364	0.935	0.982	0.893	2.286	0.785	0.942	0.736
	RF2	No. of trees: 51; leaf size: 5	1.283	0.942	0.985	0.908	2.270	0.788	0.945	0.741
	RF3	No. of trees: 50; leaf size: 5	1.325	0.938	0.983	0.909	2.205	0.800	0.947	0.739
	RF4	No. of trees: 50; leaf size: 5	1.184	0.951	0.987	0.911	2.034	0.830	0.954	0.756
	RF5	No. of trees: 52; leaf size: 5	1.214	0.948	0.986	0.914	2.105	0.818	0.952	0.747
	RF6	No. of trees: 51; leaf size: 5	1.299	0.941	0.984	0.918	2.107	0.818	0.951	0.756
	RF7	No. of trees: 50; leaf size: 5	1.371	0.934	0.982	0.917	2.079	0.822	0.953	0.750
	RF8	No. of trees: 50; leaf size: 5	1.069	0.960	0.989	0.921	2.199	0.801	0.948	0.742
	RF9	No. of trees: 49; leaf size: 5	1.238	0.946	0.985	0.914	2.030	0.831	0.954	0.758
	RF10	No. of trees: 48; leaf size: 5	1.152	0.953	0.988	0.918	2.086	0.821	0.952	0.753
Herbert	RF1	No. of trees: 52; leaf size: 5	25.952	0.918	0.977	0.817	66.273	0.797	0.940	0.697
	RF2	No. of trees: 51; leaf size: 5	31.145	0.882	0.965	0.844	62.638	0.819	0.942	0.748
	RF3	No. of trees: 50; leaf size: 5	32.533	0.871	0.962	0.836	67.436	0.790	0.934	0.732
	RF4	No. of trees: 50; leaf size: 5	33.403	0.864	0.959	0.827	67.461	0.790	0.930	0.720
	RF5	No. of trees: 52; leaf size: 5	34.903	0.852	0.956	0.838	64.366	0.809	0.941	0.734
	RF6	No. of trees: 51; leaf size: 5	31.221	0.882	0.965	0.870	58.869	0.840	0.950	0.766
	RF7	No. of trees: 50; leaf size: 5	31.489	0.879	0.964	0.863	63.236	0.815	0.941	0.747
	RF8	No. of trees: 50; leaf size: 5	32.676	0.870	0.961	0.853	72.061	0.760	0.921	0.713
	RF9	No. of trees: 49; leaf size: 5	33.082	0.867	0.960	0.861	64.121	0.810	0.937	0.765
	RF10	No. of trees: 48; leaf size: 5	18.046	0.960	0.989	0.894	65.926	0.799	0.937	0.747

Fig. 4 Comparative plots of the observed and predicted flow of the best RF models and their corresponding scatter plots during the testing phase: **a** Dulhunty River and **b** Herbert River

3.2 Development of RF model

The performance of the RF model for different input combinations is presented in Table 5 based on the statistical measures. In the RF technique, two parameters (i.e., no. of trees and leaf size) need optimizing at first using a trial and error method. Table 5 presents the optimal parameters for the RF model. The results indicated that the best performance of the RF model could be achieved with the low no. of trees (49) and leaf size of 5 (RF 9, whose input variables are H_t, H_{t-1}, Q_t, Q_{t-1}, Q_{t-2}, Q_{t-3}) and higher no. of trees (51) and leaf size of 5 (RF 6, whose input variables are Q_t, Q_{t-1}) for the Dulhunty and Herbert Rivers, based on the statistical criteria.

Figure 4a, b shows the observed and predicted river flow values and their corresponding scatter plots during the test
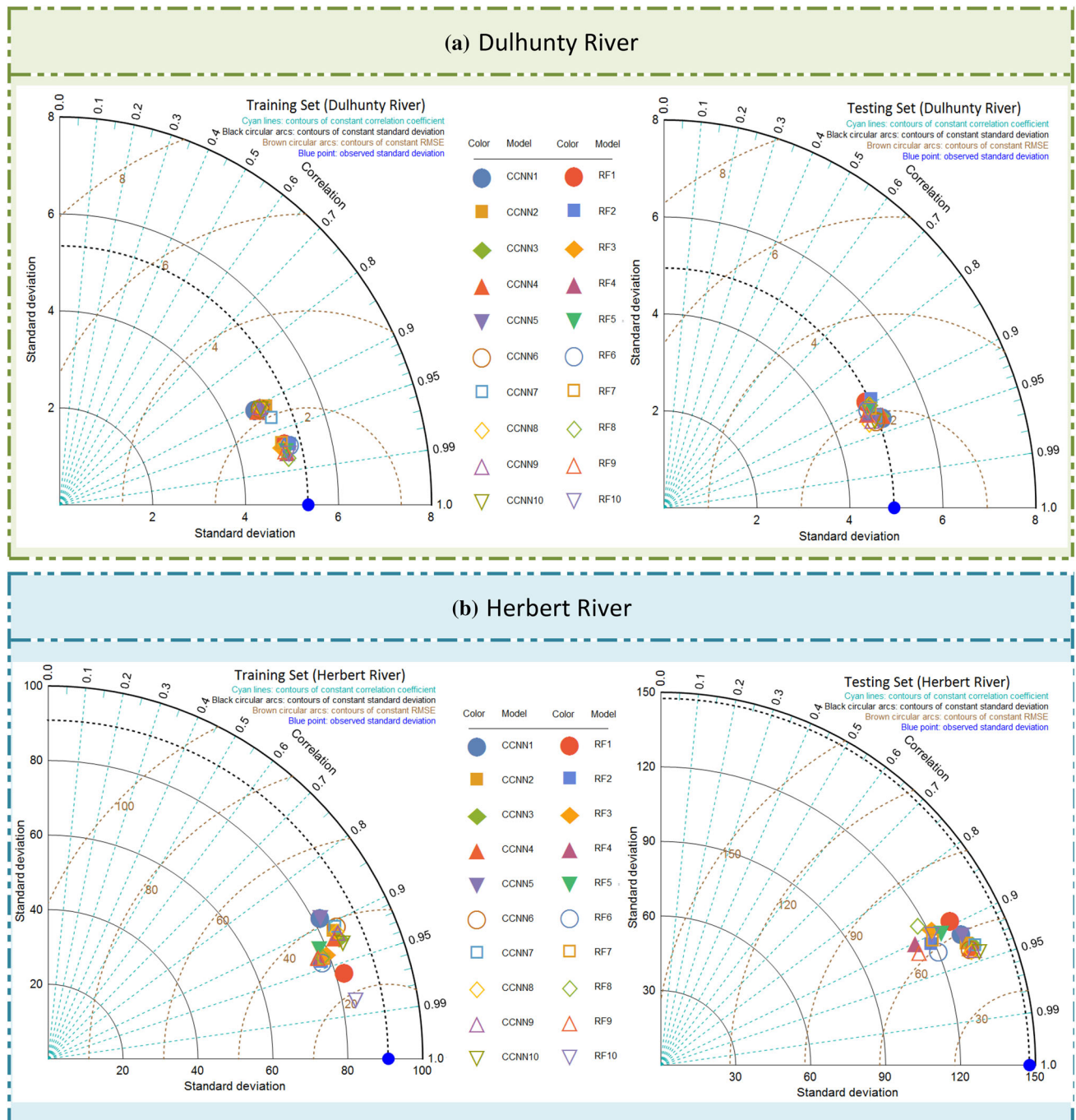
Fig. 5 Taylor diagrams of RF and CCNN models for predicting the daily river flow over the training and testing phases: **a** Dulhunty River and **b** Herbert River
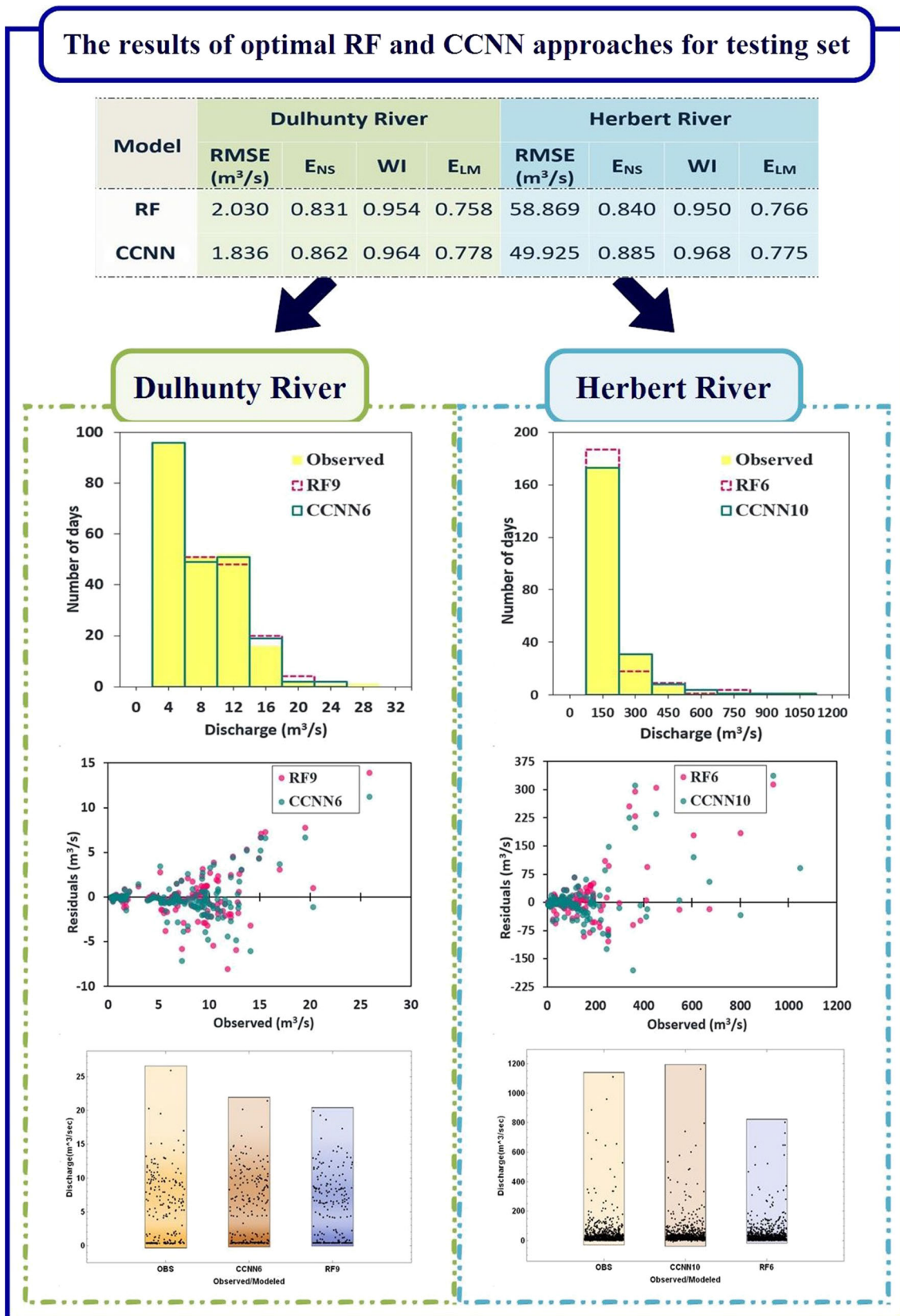
Fig. 6 Result of optimal RF and CCNN models for predicting the daily river flow in Dulhunty and Herbert Rivers over the testing phase

phase for the RF 9 and RF 6 models in the Dulhunty and Herbert Rivers, respectively. It can be found in Fig. 4 that the RF 9 and RF 6 models can provide the nonlinear river stage and river flow values successfully. This result is in agreement with the former paper obtained by Zhao et al. (2012). Like the CCNN model, it provided some errors for predicting the peak flows. This is also following the outputs obtained by Shortridge et al. (2016).

3.3 Comparisons and discussions of results

In this chapter, the performances of two applied models (i.e., CCNN and RF) were compared based on different figures. Figure 5a, b provides Taylor diagrams for the results generated by the CCNN and RF models, respectively. Here, the Taylor diagrams represent a simple graphical comparison that the similarity between the observed and predicted river flow values in terms of the correlation coefficient and standard deviations (Taylor 2001) has been used to investigate the efficiency of applied models visualized using the points as a polar plot. The ratio of variance can be calculated to produce the relative depths of predicted and observed variations in the testing phase (Taylor 2001; Gleckler et al. 2008). In this study, the Taylor plot has been used to outline the proposed models (i.e., CCNN and RF) to represent the degree prediction accuracy where the distance from the observed point can be measured from the centered RMSE (Taylor 2001). It can be seen from these visualizations that the model denoted as CCNN 10 and RF 6 generated the results that were the closest to the observed point compared to other models for the Herbert River. However, for the Dulhunty River, it is difficult to judge which one of these models is superior in their performance.

Figure 6 shows the performance evaluation criteria for the two models. Given the obtained results in Fig. 6, it indicated that all models have efficient performances in flow river prediction. As a comparison result of RMSE, N_{SE}, WI, and E_{LM} values for predicting daily river flow, it can be concluded that the CCNN model provided an accurate performance compared with the RF model for both the Dulhunty and Herbert Rivers. A direct comparison of CCNN and RF models can be also illustrated in Fig. 6 in terms of residual plots and histograms for the two rivers during the testing phase. The residual plots showed that the residuals (error) of the CCNN model were less than those of the RF model. In determining the best input combination, all models produced almost the different results for both the Dulhunty and Herbert Rivers. This indicates that the model architecture can be considered as an important factor to determine the most effective input variables

It can be seen from the results that the CCNN and RF models could estimate the river flow within

acceptable ranges. Since the application of CCNN and RF models could not be found in hydrologic time series modeling fields (e.g., streamflow, sediment, rainfall, evaporation, and groundwater), they can provide the diverse development and application processes based on the different data groups. Besides, two models (i.e., CCNN and RF) among the different ANNs-based models are a minimum level to select the optimal input combination in general. The clear selection for optimal input combinations based on the specific environments (i.e., river, watershed, lake, and reservoir) can depend on the number of developed and applied models. Therefore, the continuous researches are required to select and specify the optimal input combination using the different ANNs-based models.

4 Conclusion

In this study, the effectiveness of CCNN and RF models is investigated for the river flow prediction. To achieve this goal, the river stage and river flow data for two gauging stations in Australia for three years (2014/05/22-2017/05/21) are used. For training and testing the models, the time series data for each station are divided into 80% and 20%, respectively. Several statistical indicators such as RMSE, N_{SE}, WI, and E_{LM} are used to compare the applied models' performances. Following the archived outcomes, the performance indicators reveal that the CCNN model is able to provide the accurate and reliable predictions in comparison with the RF model. The advantages of the RF model are: (i) there is no need for feature normalization, (ii) individual decision trees can be trained in parallel, (iii) the RF model is widely used, (iv) this method reduces overfitting, and (v) it accommodates the nonlinear relationship between input and output. Also, the RF model has some disadvantages such as (i) it is not easily interpretable and (ii) it is not a state-of-the-art algorithm.

The advantages of CCNN model are that since there is no need for a user to worry about the topology of the network, the CCNN model learns much faster than the usual learning algorithms, and training is quite robust. Also, the disadvantages of CCNN model are categorized as an extreme potential for overfitting the training data and less accurate than probabilistic neural networks on small-to-medium-size problems. To improve the current models and make them particularly useful for operational river system forecasting, further research may be warranted, using the different hydro-climatic datasets to explore the models' ability to predict river stage and river flow variables.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

Ethical approval This article does not contain any studies with human participants or animals performed by any of the authors.

References

Aggarwal SK, Goel A, Singh VP (2012) Stage and discharge forecasting by SVM and ANN techniques. Water Resour Manag 26:3705–3724. https://doi.org/10.1007/s11269-012-0098-x

Ajmera TK, Goyal MK (2012) Development of stage-discharge rating curve using model tree and neural networks: an application to Peachtree Creek in Atlanta. Expert Syst Appl 39:5702–5710. https://doi.org/10.1016/j.eswa.2011.11.101

Alok A, Patra KC, Das SK (2013) Prediction of discharge with Elman and cascade neural networks. Res J Recent Sci India 2:279–284

Alvisi S, Mascellani G, Franchini M, Bardossy A (2006) Water level forecasting through fuzzy logic and artificial neural network approaches. Hydrol Earth Syst Sci 10:1–17

Baiamonte G, Ferro V (2007) Simple flume for flow measurement in sloping open channel. J Irrig Drain Eng 133:71–78. https://doi.org/10.1061/(ASCE)0733-9437(2007)133:1(71)

Bhattacharya B, Solomatine DP (2005) Neural networks and M5 model trees in modeling water level-discharge relationship. Neurocomputing 63:381–396. https://doi.org/10.1016/j.neucom.2004.04.016

Ch S, Anand N, Panigrahi BK, Mathur S (2013) Streamflow forecasting by SVM with quantum behaved particle swarm optimization. Neurocomputing 101:18–23. https://doi.org/10.1016/j.neucom.2012.07.017

Chen SH, Lin YH, Chang LC, Chang FJ (2006) The strategy of building a flood forecast model by neuro-fuzzy network. Hydrol Process 20:1525–1540. https://doi.org/10.1002/hyp.5942

Clemmens AJ, Wahlin BT (2006) Accuracy of annual volume from current-meter-based stage discharges. J Hydrol Eng 11:489–501. https://doi.org/10.1061/(ASCE)1084-0699(2006)11:5(489)

Deka P, Chandramouli V (2003) A fuzzy neural network model for deriving the river stage—discharge relationship. Hydrol Sci J 48:197–209. https://doi.org/10.1623/hysj.48.2.197.44697

Deng W, Zhao H, Yang X, Xiong J, Sun M, Li B (2017a) Study on an improved adaptive PSO algorithm for solving multi-objective gate assignment. Appl Soft Comput 59:288–302

Deng W, Zhao H, Zou L, Li G, Yang X, Wu D (2017b) A novel collaborative optimization algorithm in solving complex optimization problems. Soft Comput 21(15):4387–4398

Deng W, Zhang S, Zhao H, Yang X (2018) A novel fault diagnosis method based on integrating empirical wavelet transform and fuzzy entropy for motor bearing. IEEE Access 6(1):35042–35056

Deng W, Xu J, Zhao H (2019a) An improved ant colony optimization algorithm based on hybrid strategies for scheduling problem. IEEE Access 7:20281–20292

Deng W, Yao R, Zhao H, Yang X, Li G (2019b) A novel intelligent diagnosis method using optimal LS-SVM with improved PSO algorithm. Soft Comput 23(7):2445–2462

Diamantopoulou MJ, Georgiou PE, Papamichail DM (2007) Performance of neural network models with Kalman learning rule for flow routing in a river system. Fresenius Environ Bull 16:1474

Fahlman SE, Lebiere C (1990) The cascade-correlation learning architecture. In: Touretzky DS (ed) Advances in neural information processing systems 2. Morgan Kaufmann Publishers Inc., San Francisco

Firat M (2008) Comparison of artificial intelligence techniques for river flow forecasting. Hydrol Earth Syst Sci 12:123–139. https://doi.org/10.5194/hess-12-123-2008

Ghimire BN, Reddy MJ (2010) Development of stage-discharge rating curve in river using genetic algorithm and model tree. In: International workshop advanced in statistical hydrology, Italy

Ghorbani MA, Khatibi R, Goel A et al (2016a) Modeling river discharge time series using support vector machine and artificial neural networks. Environ Earth Sci 75:685. https://doi.org/10.1007/s12665-016-5435-6

Ghorbani MA, Zadeh HA, Isazadeh M, Terzi O (2016b) A comparative study of artificial neural network (MLP, RBF) and support vector machine models for river flow prediction. Environ Earth Sci 75:476. https://doi.org/10.1007/s12665-015-5096-x

Gleckler PJ, Taylor KE, Doutriaux C (2008) Performance metrics for climate models. J Geophys Res Atmos. https://doi.org/10.1029/2007JD008972

Gocić M, Motamedi S, Shamshirband S et al (2015) Soft computing approaches for forecasting reference evapotranspiration. Comput Electron Agric 113:164–173. https://doi.org/10.1016/j.compag.2015.02.010

Goel A, Pal M (2012) Stage-discharge modeling using support vector machines. Int J Eng 25:1–9. https://doi.org/10.5829/idosi.ije.2012.25.01a.01

Habib EH, Meselhe EA (2006) Stage-discharge relations for low-gradient tidal streams using data-driven models. J Hydraul Eng 132:482–492. https://doi.org/10.1061/(ASCE)0733-9429(2006)132:5(482)

Hasanpour Kashani M, Daneshfaraz R, Ghorbani MA et al (2015) Comparison of different methods for developing a stage-discharge curve of the Kizilirmak River. J Flood Risk Manag 8:71–86. https://doi.org/10.1111/jfr3.12064

Hengl T, Heuvelink GBM, Kempen B et al (2015) Mapping soil properties of Africa at 250 m resolution: random forests significantly improve current predictions. PLoS ONE. https://doi.org/10.1371/journal.pone.0125814

Jain SK, Chalisgaonkar D (2000) setting up stage-discharge relations using ANN. J Hydrol Eng 5:428–433. https://doi.org/10.1061/(ASCE)1084-0699(2000)5:4(428)

Karunanithi N, Grenney WJ, Whitley D, Bovee K (1994) Neural networks for river flow prediction. J Comput Civ Eng 8:201–220. https://doi.org/10.1061/(ASCE)0887-3801(1994)8:2(201)

Kashani MH, Soltangheys R (2018) Comparison of three intelligent techniques for runoff simulation. Civil Eng J 4(5):1095–1103

Khatibi R, Ghorbani MA, Kashani MH, Kisi O (2011) Comparison of three artificial intelligence techniques for discharge routing. J Hydrol 403(3–4):201–212

Khatibi R, Sivakumar B, Ghorbani MA et al (2012) Investigating chaos in river stage and discharge time series. J Hydrol 414–415:108–117. https://doi.org/10.1016/j.jhydrol.2011.10.026

Khatibi R, Ghorbani MA, Pourhosseini FA (2017) Streamflow predictions using nature-inspired firefly algorithms and a multiple model strategy—directions of innovation towards next-generation practices. Adv Eng Inform 34:80–89. https://doi.org/10.1016/J.AEI.2017.10.002

Kim S, Shiri J, Kisi O (2012) Pan evaporation modeling using neural computing approach for different climatic zones. Water Resour Manag 26:3231–3249. https://doi.org/10.1007/s11269-012-0069-2

Kim S, Singh VP, Seo Y (2014) Evaluation of pan evaporation modeling with two different neural networks and weather station data. Theor Appl Climatol 117:1–13. https://doi.org/10.1007/s00704-013-0985-y

Kişi Ö (2007) Streamflow forecasting using different artificial neural network algorithms. J Hydrol Eng 12:532–539. https://doi.org/10.1061/(ASCE)1084-0699(2007)12:5(532)

Kumar M, Raghuwanshi NS, Singh R et al (2002) Estimating evapotranspiration using artificial neural network. J Irrig Drain Eng 128:224–233

Legates DR, Davis RE (1997) The continuing search for an anthropogenic climate change signal: limitations of correlation-based approaches. Geophys Res Lett 24:2319–2322. https://doi.org/10.1029/97GL02207

Legates DR, McCabe GJ (1999) Evaluating the use of "goodness-of-fit" measures in hydrologic and hydroclimatic model validation. Water Resour Res 35:233–241. https://doi.org/10.1029/1998WR900018

Legates DR, McCabe GJ (2013) A refined index of model performance: a rejoinder. Int J Climatol 33:1053–1056. https://doi.org/10.1002/joc.3487

Liu WC, Chung CE (2014) Enhancing the predicting accuracy of the water stage using a physical-based model and an artificial neural network-genetic algorithm in a river system. Water (Switzerland) 6:1642–1661. https://doi.org/10.3390/w6061642

Nash JE, Sutcliffe JV (1970) river flow forecasting through conceptual models part 1—a discussion of principles. J Hydrol 10:282–290. https://doi.org/10.1016/0022-1694(70)90255-6

Nayak PC, Sudheer KP, Rangan DM, Ramasastri KS (2004) A neuro-fuzzy computing technique for modeling hydrological time series. J Hydrol 291:52–66. https://doi.org/10.1016/j.jhydrol.2003.12.010

Nguyen T-T, Huu QN, Li MJ (2015) Forecasting time series water levels on Mekong river using machine learning models. In: 2015 Seventh international conference on knowledge and systems engineering (KSE). IEEE, pp 292–297

Prasad R, Deo RC, Li Y, Maraseni T (2017) Input selection and performance optimization of ANN-based streamflow forecasts in the drought-prone Murray Darling Basin region using IIS and MODWT algorithm. Atmos Res 197:42–63. https://doi.org/10.1016/j.atmosres.2017.06.014

Rodriguez-Galiano VF, Atkinson PM (2016) Modelling interannual variation in the spring and autumn land surface phenology of the European forest. Biogeosciences 13:3305

Seo Y, Kim S, Kisi O, Singh VP (2015) Daily water level forecasting using wavelet decomposition and artificial intelligence techniques. J Hydrol 520:224–243

Shoaib M, Shamseldin AY, Melville BW, Khan MM (2015) Runoff forecasting using hybrid wavelet gene expression programming (WGEP) approach. J Hydrol 527:326–344. https://doi.org/10.1016/j.jhydrol.2015.04.072

Shortridge JE, Guikema SD, Zaitchik BF (2016) Machine learning methods for empirical streamflow simulation: a comparison of model accuracy, interpretability, and uncertainty in seasonal watersheds. Hydrol Earth Syst Sci 20:2611

Sivapragasam C, Muttil N (2005) Discharge rating curve extension—a new approach. Water Resour Manag 19:505–520. https://doi.org/10.1007/s11269-005-6811-2

Sudheer KP, Jain SK (2003) Radial basis function neural network for modeling rating curves. J Hydrol Eng 8:161–164. https://doi.org/10.1061/(ASCE)1084-0699(2003)8:3(161)

Taormina R, Chau KW (2014) Neural network river forecasting with multi-objective fully informed particle swarm optimization. J Hydroinform 17:99–113. https://doi.org/10.2166/hydro.2014.116

Taormina R, Chau KW, Sivakumar B (2015) Neural network river forecasting through baseflow separation and binary-coded swarm optimization. J Hydrol 529:1788–1797. https://doi.org/10.1016/j.jhydrol.2015.08.008

Tawfik M, Ibrahim A, Fahmy H (1997) Hysteresis sensitive neural network for modeling rating curves. J Comput Civ Eng 11:206–211. https://doi.org/10.1061/(ASCE)0887-3801(1997)11:3(206)

Taylor KE (2001) Summarizing multiple aspects of model performance in a single diagram. J Geophys Res Atmos 106:7183–7192. https://doi.org/10.1029/2000JD900719

Thirumalaiah K, Deo MC (1998) River stage forecasting using artificial neural networks. J Hydrol Eng 3:26–32. https://doi.org/10.1061/(ASCE)1084-0699(1998)3:1(26)

Were K, Bui DT, Dick OB, Singh BR (2015) A comparative assessment of support vector regression, artificial neural networks, and random forests for predicting and mapping soil organic carbon stocks across an Afromontane landscape. Ecol Indic 52:394–403. https://doi.org/10.1016/j.ecolind.2014.12.028

Willmott CJ, Matsuura K (2005) Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance. Clim Res 30:79–82. https://doi.org/10.3354/cr030079

Willmott CJ, Robeson SM, Matsuura K (2012) A refined index of model performance. Int J Climatol 32:2088–2094. https://doi.org/10.1002/joc.2419

Wu JS, Han J, Annambhotla S, Bryant S (2005) Artificial neural networks for forecasting watershed runoff and stream flows. J Hydrol Eng 10:216–222. https://doi.org/10.1061/(ASCE)1084-0699(2005)10:3(216)

Zhang Z, Zhang Q, Singh VP, Shi P (2018) River flow modeling: comparison of performance and evaluation of uncertainty using data-driven models and conceptual hydrological model. Stoch Environ Res Risk Assess 32:2667–2682. https://doi.org/10.1007/s00477-018-1536-y

Zhao TTG, Yang DW, Cai XM (2012) Predict seasonal low flows in the upper Yangtze River using random forests model. J Hydroelectr Eng 31:18–24

Zhao H, Sun M, Deng W, Yang X (2017) A new feature extraction method based on EEMD and multi-scale fuzzy entropy for motor bearing. Entropy 19(1):14

Zounemat-Kermani M, Seo Y, Kim S, Ghorbani MA, Samadianfard S, Naghshara S, Kim NW, Singh VP (2019) Can the decomposition approaches always enhance the soft computing models? Predicting the dissolved oxygen concentration in St. Johns River, Florida. Appl Sci 9(12):2534