



# Application of a novel technique of the multi-discrete wavelet transforms in hybrid with artificial neural network to forecast the daily and monthly streamflow

Sadegh Momeneh<sup>1</sup> · Vahid Nourani<sup>2</sup>

Received: 8 February 2022 / Accepted: 28 March 2022 / Published online: 16 April 2022  
© The Author(s), under exclusive licence to Springer Nature Switzerland AG 2022

## Abstract

In this study, a technique for accurate forecasting of streamflow, which is of great importance to reduce the risks of floods and water resources management, in the catchment area of Gamasiab River, located in western Iran, is presented. A hybrid of the artificial intelligence (AI) model called artificial neural network (ANN) models, with data preprocessing, includes discrete wavelet transform (DWT) and multi-discrete wavelet transform (M-DWT), were used. These hybrid models were used to forecast the flow of the study basin on a daily and monthly time scale with 1, 2, 3, and 7 steps ahead. The data used in this study include daily and monthly streamflow, precipitation, and temperature data for 31 years (23 September 1986–22 September 2017), which is a time series of delayed data used as an input signal to the models. The criteria of Nash–Sutcliffe efficiency (NSE), root mean square error (RMSE), and correlation coefficient ( $R$ ), were used to evaluate the performance of the models. The results indicated that the hybrid M-DWT-ANN models significantly reduced the forecasting error compared to the unit models DWT-ANN and usual models ANN for the time steps ahead. M-DWT-ANN model performed better than other models in forecasting the current for the 1, 2, 3, and 7 days and months ahead. For example, the RMSE ( $m^3/s$ ) values of the ANN, DWT-ANN, and M-DWT-ANN models for streamflow forecasting the 7-daily-ahead in the verification period are 14.10, 9.96, and 0.26, respectively. In general, the results showed that using the M-DWT method as preprocessing of input data is a valuable tool to increase the accuracy and performance improvements of the predictive model. The findings of this study showed the potential of M-DWT-AI hybrid models to improve streamflow forecasting.

**Keywords** Artificial intelligence · Artificial neural network · Gamasiab River basin · Streamflow forecasting · Wavelet transform

## Introduction

The management and forecasting of hydrological processes and their complexities are fundamental issues in different geographical areas. Streamflow forecasting plays an essential role in formulation strategies and sustainable management of water resources. It can also be considered, for various

reasons, including helping with planning and optimizing the water resources system and reducing flood risks. Forecasting very accurate and reliable flow fluctuations due to the environmental structure and the intermittent nature of the streamflow in semi-arid watersheds are of particular significance. Most rivers flowing in semi-arid and arid catchments due to climatic conditions such as seasonal precipitation and high evaporation rate have rebellious behavior and seasonal and non-permanent flow and play a vital role in meeting the water demand of these areas. These happenings occur in most parts of Iran. Since in catchments, it is not possible to measure all the observational quantities required to model and analyze flow fluctuations. Therefore, essential to choose a model that can, despite the simple structure, requires minimal parameters for forecast runoff and flow caused by rainfall in the basin with high accuracy.

✉ Sadegh Momeneh  
sadeghmomeneh@gmail.com

Vahid Nourani  
vnourani@yahoo.com

<sup>1</sup> Department of Civil Engineering, Razi University, Kermanshah, Iran

<sup>2</sup> Department of Water Resources Engineering, University of Tabriz, Tabriz, Iran

There are three approaches to modeling streamflow, including physical (process-oriented), empirical/metric (data-based and statistics), and conceptual (based on hypotheses) approaches (Beck 1991; Ahooghalandari et al. 2016). The physical and conceptual models often require significant inputs, including various hydrological, geometric and structural, and climatic data from the basin. In addition, calibration of these models is complicated and time-consuming (Tokar and Markus 2000; Panda et al. 2010; Shi et al. 2011; Arnold et al. 2012). If sufficient input data is not available, metric (data-based) models perform better and more reliably than conceptual models (Carcano et al. 2008). Several studies have compared traditional physical and conceptual models, including SWAT, IHACRES, MIKE 11, and HEC-HMS, with data-based models such as ANN for flow forecasting and often found that ANN models provide forecasting better (e.g., Carcano et al. 2008; Panda et al. 2010; Kim and Pachepsky 2010; Rezaeianzadeh et al. 2013; Ahooghalandari et al. 2016; Young et al. 2017; Jimeno-Sáez et al. 2018; Ahmadi et al. 2019; Wagena et al. 2020).

At the same time, obtaining accurate forecasting is often more important than understanding the process and recognizing the mechanisms that create it; thus, simple data-based models can be a suitable alternative (Ebrahimi and Rajaei 2017; Jha and Sahoo 2015). In streamflow forecasting programs, data-based or data-driven hydrological methods have become increasingly popular and used due to their rapid development time and minimal information requirements. The data-based approach involves mathematical equations derived not from the physical process of the basin but from time series analysis (Solomatine and Ostfeld 2008). In other words, data-based models can learn and generalize trends based on functional relationships in data by developed algorithms (Coulibaly et al. 2000; Zhang et al. 2001). When the observations and field data are not enough, and accurate estimation is more critical than understanding physics, that's the time a black-box or data-driven model can perform well. Although they may not provide a physical interpretation and performance of the catchment processes, they nevertheless forecast the relatively accurate streamflow. In data-based flow forecasting, linear models such as multiple linear regression (MLR) and autoregressive integrated moving average (ARIMA) are traditionally used to forecast streamflow. Linear models have limitations and do not perform well enough when modeling hydrological time series. Because these processes are often non-linear, the dynamic behavior of the hydrological system changes over time (Bierkens 1998; Tokar and Johnson 1999; Nourani et al. 2014a). In later years, the non-linear models and artificial intelligence (AI) data-based models, including the artificial neural network (ANN) and support vector regression (SVR), were introduced for flow forecasting applications. Kang et al. (1993) and Hsu et al. (1995) were among the first

to use ANN programs in river flow and rainfall-runoff forecasting and compared the ANN model and linear models. They found that ANN is practical and suitable for forecasting river flow. Gradually AI models, including ANN, SVR, radial function base network (RBF), adaptive neural-fuzzy inference system (ANFIS) models, and other AI models, to find essence relationships and identify patterns in a complex system in between streamflow and various hydrological variables, were used without having to build a conceptual model and understanding of the mechanism physics in complex systems and relationships. Thus, in recent years, AI models such as ANN, RBF, SVR, and ANFIS as alternative and efficient tools are accepted for modeling complex hydrological systems and used effectively and extensively for streamflow forecasting (e.g., Zealand et al. 1999; Campolo et al. 1999; Kumar et al. 2004; El-Shafie et al. 2007; Pramanik and Panda 2009; Kagoda et al. 2010; Meng et al. 2015; Kasiviswanathan et al. 2016; Modaresi et al. 2018; Ateeq-ur-Rauf et al. 2018; Ali and Shahbaz 2020).

Although essential features of AI methods are their ability to identify patterns in a complex system, if the inputs are very non-stationary often do not cope with such data if the inputs are not preprocessed (Cannas et al. 2006). Here the combination of preprocessing with AI models can play an auxiliary role. In the last decade, wavelet analysis has been used as a data preprocessor in water resources engineering and hydrology in various issues, and non-stationary data management has been very effective. Wavelet transforms (WT) provide functional decompositions of original time series (input data). The data decomposed by the WT addresses the potential shortcomings of the model forecasting by capturing and placing valuable information at different resolution levels. The use of WT as a mining tool can reveal and extract various concealed features in the physical structure of the data. The WT is an effective and well-known tool in non-stationary and noisy data analysis that for modeling and forecasting has been used in a wide range of water resources management issues (Nourani et al. 2014a). Wavelet has proven to be an efficient mathematical tool (Adamowski 2008; Partal 2009).

Have been accepted Wavelet-AI hybrid models in recent years as a potentially helpful method for modeling hydrological processes. These hybrid models in various applications have been used, including forecasting' streamflow, rainfall-runoff, precipitation, water quality, groundwater level, water temperature, evaporation, and sedimentation (Nourani et al. 2014b; Raghavendra and Deka 2014). Preprocessing data before using it as input for ANN, ANFIS, RBF, and SVR networks (or other data-driven models) can significantly improve the performance of these models (Moosavi et al. 2014; Liu et al. 2014). The discrete wavelet transform (DWT)/wavelet analysis (WA) method can with ANN, SVR, RBF, ANFIS, and other models be combined

to create a hybrid model entitled DWT/WA-ANN, DWT/WA-SVR, DWT/WA-RBF, and DWT/WA-ANFIS. Many researchers have shown in their studies that wavelet-based coupling models, especially ANN hybrid models, perform better than conventional models such as ANN, ANFIS, and RBF. For example, Cannas et al. (2006) used data processing by WT as input to the ANN model to forecast river flow. They found which trained networks with preprocessed data, better performance in forecast compared to trained networks with raw data and without preprocessing. Kisi (2009) used the wavelet and ANN hybrid technique to study daily river flow forecasting and compared it with the results of individual ANN models. His results showed that wavelet preprocessing could significantly increase the accuracy of ANN forecasting in daily streamflow forecasting. Adamowski and Sun (2010) developed the WA-ANN model to improve the accuracy of streamflow forecasting at lead times of 1 and 3 days for two different non-perennial rivers for semi-arid catchments in Cyprus. Liu et al. (2014) proposed to improve the accuracy of the DWT-SVR hybrid model and used the daily and monthly streamflow data at two stations in Indiana, USA, to evaluate the model's forecasting skills. The results showed that the DWT-SVR hybrid models performed better than the SVR model for daily and monthly streamflow forecasting. Zhu et al. (2016) modeled the streamflow of watersheds have in the upper reaches of the Yangtze River, China. After analyzing the data was used by wavelet transform and applying these time series as monthly input data to SVR to forecast streamflow. The results showed that the forecasting of DWT-SVR models improved compared to SVR models. Hadi and Tombul (2018) used WT as a preprocessor of input data to AI models to predict the streamflow for seven days ahead in a basin in southwestern Turkey. The results indicate an increase in the accuracy of WT-coupled AI models in streamflow prediction. Tayyab et al. (2019) data-based ANN and RBF models hybrid with DWT (i.e., DWT-ANN and DWT-RBF) were used to rainfall-runoff in the river basin in China. The results showed that hybrid models provided predictions better. Freire et al. (2019) used a synthetic WA-ANN composition to forecast daily flows to the Sobradinho Reservoir in Brazil for seven days ahead. Their studies results showed the superiority of the WA-ANN model over the single ANN model. Dalkılıç and Hashimi (2020) evaluated the ANN, ANFIS, and DWT-ANN models in the Büyük Menderes River in Western Anatolia and found that the DWT-ANN model performed best.

In the study, the novel technique called the multi-discrete wavelet transform (M-DWT) for preprocessing the inputs data was used and achieving high-precision streamflow forecasting. Combined models, including M-DWT-ANN, have been developed for this purpose. To prove the M-DWT preprocessing efficiency, were comparisons between the performance of the models developed by this method and the

performance of the DWT-ANN unit models and separate ANN models for forecasting daily and monthly flows in the Gamasiab River basin. The study area is in parts of the Kermanshah and Hamedan provinces, located in a catchment area with a semi-arid climate in western Iran.

## Materials and methods

### Study area and data

In the present study, the flow in the catchment area of the Gamasiab River in western Iran evaluates (Fig. 1). This basin has 10,935 square kilometers, with a semi-arid climate and a semi-humid climate in the highlands. This basin in the geographical area with coordinates of latitude  $33^{\circ}49' N$  to  $34^{\circ}57' N$  and longitude  $47^{\circ}06' E$  to  $49^{\circ}10' E$ , is located. The maximum and minimum altitude of the region is 3450 m and 1272 m, respectively. It is from the above mean sea level (AMSL). The average height of the basin is 1873 m, and the perimeter is 636 km. The compactness coefficient (Gravelius method) of the catchment is 1.7. the shape is almost elongated. The length of the longest main waterway is 221 km. The slope of the basin varies from 0.1 to 53.1%. The slope of the canal is 2.91%, and the average slope of the plain is 7.96%. This basin has vegetation and land use in the middle areas, and lowlands are mainly horticultural and agricultural (irrigated, rainfed, and rainfed). In the highlands, the vegetation of the rangelands is semi-dense, dense, and poor density, respectively. A small part includes forest cover, barren lands, water, watercourse, mountainous (rocky), urban, and residential. Due to the density of wells in different parts of the area because of insufficient surface irrigation networks, groundwater is used to meet the water needs of agricultural products. The alluvial (porous) aquifer in this region, which consists of a complex distribution of gravel, sand, grit, silt, and clay, is an example of many sedimentary systems of aquifers in Iran. The study basin is surrounded by heights of the Zagros Mountains and in the middle of a rugged hilly area and plain. Where feeding is the primary source of groundwater, and inclusive rainfall, infiltration from rivers in the region, return flow from irrigation. Also, the main discharge parameters in this catchment are unauthorized exploitation and evaporation. Eventually, is discharged the water flows in the catchment by the outlet at the southwestern part of the catchment. The catchments of arid and semi-arid regions, including Iran, are flooded rivers, of which the Gamasiab River is no exception.

Hydrometry and Evaporation Stations of Polechehr (or Chehr Bridge) at the outlet of the catchment area with latitude coordinates of  $34^{\circ}20' N$  latitude and  $47^{\circ}26' E$  longitudes, is located. The elevation of the Hydrometry Station is 1280 m above mean sea level, and it is on the Gamasiab

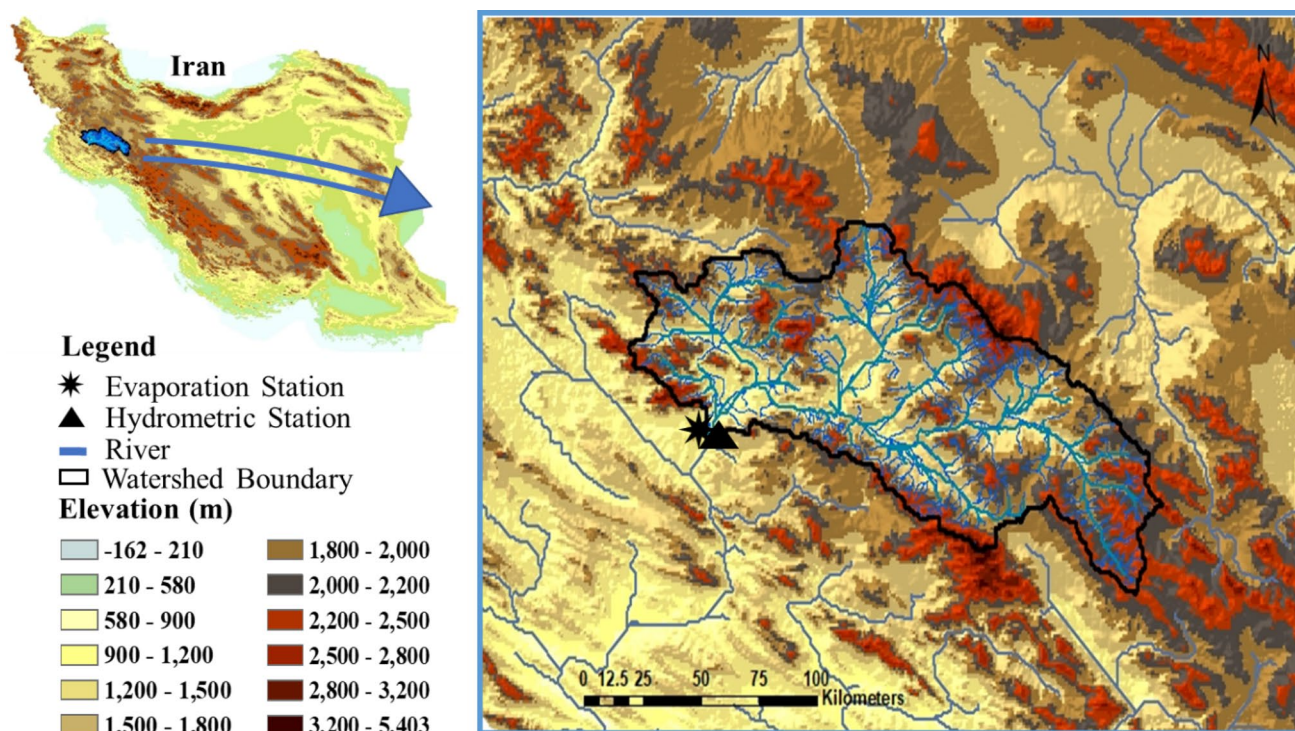


Fig. 1 Location map of the study area with national boundaries

River located in Kermanshah province in western Iran. The Gamasiab River is the tributary of the Karkheh River and part of the Persian Gulf and Oman sea catchment in terms of the catchment area. The average annual discharge of Polechehr station is  $25.82 \text{ m}^3/\text{s}$ , and the maximum recorded discharge rate for the long-term period is  $796 \text{ m}^3/\text{s}$ . The maximum discharges rate recorded is most often in November–May, when most has occurred of the rainfall. The maximum monthly average temperature and the minimum monthly temperature at Polechehr station in July and January are  $38.76 \text{ }^\circ\text{C}$  and  $-4.15 \text{ }^\circ\text{C}$ , respectively. The average annual rainfall in this period is  $384 \text{ mm}$ , and about 92% of the total annual rainfall occurs between November and May. As can be seen in Fig. 4, the river is a non-permanent river, and this region has hot and dry summers. Due to the irrigation of agricultural fields in the summer, the river becomes a seasonal and dry river. To prevent this can be used catchment management and flood control plans.

In this study, to train and verify the performance of ANN models, 31 years of daily data measured at Polechehr station, including streamflow, precipitation, and temperature data (23 September 1986–22 September 2017), were used in the catchment. Also, for monthly data was used of 31 years average monthly data (October 1986–September 2017). In most studies, it is divided into two parts and divided into two parts that can be sufficient in the modeling process (Nourani et al. 2015). The data utilized were standardized and normalized

using scaling between zero and one to confident that all variables have been paid equal attention during the training step. The first 70% of the total data set to develop the model (training), and the remaining 30% of the entire data set to evaluate (test) the developed models, were used. Meantime, is 1, 2, 3, and 7 time-steps days and months as forecasting time horizons selected. The data used in this research, the archive of the data and information of the regional water company of Kermanshah, was obtained.

### Model performance criteria

Performance evaluation of a hydrological model is performed and described, usually by comparing the error values and the differences between the observed and simulated variables. Were divided in forecasting hydrological phenomena, most data into the calibration data set (training) and verification data set (testing) to obtain correct evaluation and comparison of model performance. It is also necessary for AI models to find a suitable structure. In this study, for all models developed for streamflow forecasting, statistical criteria inclusive correlation coefficient ( $R$ ), Nash–Sutcliffe efficiency coefficient (NSE), and root mean square error (RMSE) to evaluate the statistical relationship between the forecasted value and the observed value, to assessment the forecast power of the model, and to measure the variance of the error, respectively, were used.

$$R = \frac{\sum_{i=1}^n (O_i - \bar{O}) \times (F_i - \bar{F})}{\sqrt{\sum_{i=1}^n (O_i - \bar{O})^2 \times (F_i - \bar{F})^2}}, \tag{1}$$

$$NSE = 1 - \frac{\sum_{i=1}^n (O_i - F_i)^2}{\sum_{i=1}^n (O_i - \bar{O})^2}, \tag{2}$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (O_i - F_i)^2}, \tag{3}$$

where  $n$  is the number of data set used,  $F_i$  is the forecasted values (model outputs),  $O_i$  is the observed data, and  $\bar{F}$  and  $\bar{O}$  are the average values for  $F_i$  and  $O_i$ , respectively. The best fit between the forecasted value and the observed value occurs when the values obtained from these relationships (Eqs. 1–3) reach values  $R$  and  $NSE$  to maximum one and value  $RMSE$  to minimum zero, respectively (Gong et al. 2016; Liu et al. 2014).

### Artificial neural network (ANN)

In recent decades, the ANN estimation approach as a black-box model has been a great deal of consideration from many researchers globally. It has been used widely in diverse fields such as time series forecasting, pattern and sequence

recognition, processing data, mining data, and identification and control system (Nayak et al. 2006). ANNs have performed well in the input–output function approximation such as forecasting. Hence, they have been used successfully for modeling and forecasting in the earth sciences (ASCE 2000a; b). The ANN from several artificial neural cells interconnected in several layers conforming to the specific architecture is composed. Can be using the ANNs to forecast future values of possibly noisy time series based on past histories. Were organized the neural networks for converting inputs into meaningful outputs (Adamowski and Chan 2011). In the connections between neurons are adjustable parameters located are called weight. The input signal through the network in a forward direction is transmitted. These signals are received in each neuron (node) in the input layer from external inputs and another layer from outputs from other neurons to which it is linked. Each neuron produces a result by an activation function that is a linear/non-linear static function of the weighted sum of these inputs. ANN multi-layer perceptron (MLPs), first by Rumelhart and McClelland (1986), was proposed and is one of the most widely used neural networks for hydrological modeling that can recognize latent and non-linear patterns (Nayak et al. 2006; Principe et al. 2000). Figure 2 are displayed the architecture of a typical MLP network with a hidden layer in which the logistic (sigmoid) activation function and a linear function in the output layer. The approach of feed-forward MLP to mathematical expression is as follows:

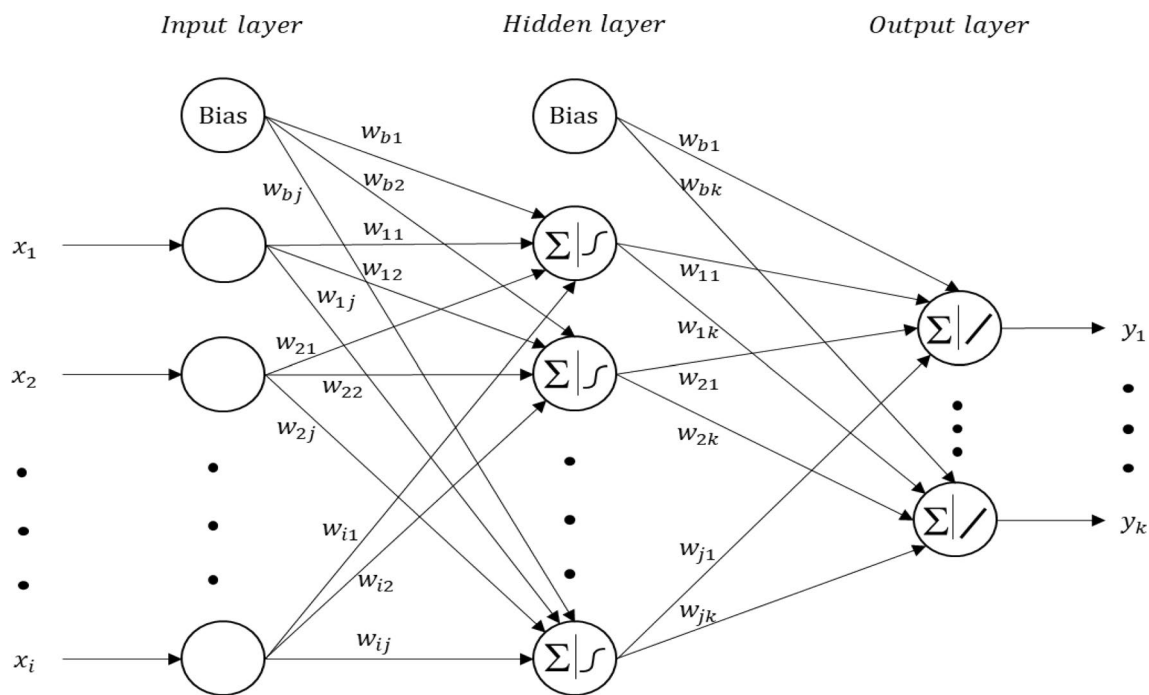


Fig. 2 ANN architecture three-layer with one hidden layer

$$y_k(t) = f_O \left[ \sum_{j=1}^m w_{jk} f_h \left( \sum_{i=1}^n w_{ij} x_i(t) + w_{bj} \right) + w_{bk} \right], \quad (4)$$

where  $n$  is the number of data set,  $m$  is the number of neurons in the hidden layer;  $w_{ij}$ ,  $w_{jk}$  = the weight that neurons have in the input and output layers, respectively;  $w_{bj}$ ,  $w_{bk}$  = bias in the hidden and output layers, respectively;  $f_h$ ,  $f_o$  = activation function of the neurons in the hidden and output layers, respectively;  $x_i(t)$ ,  $y_k(t)$  = the  $i$ -th input variables and the  $k$ -th output variables at time step  $t$ , respectively (Kim and Valdes 2003).

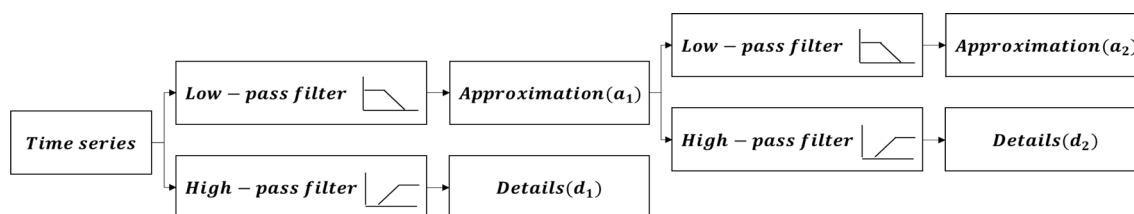
For ANN model development, the designation of the structure and the model's training algorithm is critical. An algorithm is needed that provides proper performance in the forecast. Commonly used training algorithms include Levenberg–Marquardt (LM), scaled conjugate gradient (SCG), gradient descent with momentum, adaptive learning rate (GDX), and Bayesian regularization (BR) (Mohanty et al. 2010; Gong et al. 2016). Among these, the best BR back-propagation algorithm was selected, according to performance criteria (i.e., lowest RMSE and highest NSE and  $R$ ), to train and develop the ANN models. In these algorithms, the error between the intended and forecasted output is back-propagation through the network, and weights linking the neurons in the learning phase through a training algorithm are updated. MLPs can perform well in function approximation, provided that there are sufficient neurons in the hidden layer of the network enough amount of data is existed (Cybenko 1989; Principe et al. 2000). In this study, preliminary results showed that a hidden layer is sufficient to approximate the relationship between observed and forecasted streamflow. By the trial-and-error method was determined the optimum number of neurons hidden layers. The number of layers and neurons has been selected, with the lowest RMSE values as the appropriate number. Meantime, to ensure the optimal performance of the network, a cross-verification method, be used to choose the best network architecture (Principe et al. 2000).

In the present study, used data sets of the precipitation, temperature, and streamflow into two subsets are divided. The first subset is the training data set used to calculate the error gradient and update the weight and bias of neurons in different

layers of the network. The second subset is the test suite that an independent data set employed to verify the efficiency and performance of the model. Primary stopping criteria were used based on cross-validation are applied during the training of the neural networks, including the Mu (equal  $1.00e^{+10}$ ), Gradient (equal  $1.00e^{-7}$ ), and Maximum Iteration (equal 1000 epoch) criteria. The optimal structure of the ANN network and parameter tuning were designated using a trial-and-error method. In such a way that the optimal value of the parameters and variables based on performance criteria, including the RMSE,  $R$ , and NSE by sensitivity analysis, was determined.

## Discrete wavelet transforms (DWT)

The wavelet transform (WT) is a mathematical tool that is a time-dependent spectral analysis that analyzes signals in a time–frequency space and provides a time-scale illustration of processes and their relationships (Daubechies 1990). The WT is a valuable and essential derivative of the Fourier transform (FT). Fourier analysis has a primary disadvantage and the loss of time information in transforming into a frequency domain. At the same time, the WT includes a poly-resolution decomposition in the time and frequency domains (Tiwari and Chatterjee 2011). One type of WT is the discrete wavelet transform (DWT), which is used widely due to its simplicity and low data generation, and the need for short computational time. At the same time, with its concise and valuable analysis, it still produces a very efficient and precise analysis (Partal and Kucuk 2006). The DWT using the different filters and various mother wavelets possesses long-time distances for low-frequency data and short-time distances for high-frequency data and can reveal some properties and hidden aspects of the time series. The DWT is particularly beneficial when the signal contains various embedded information, jumps, or shifts (Nalley et al. 2012). DWT often is used for time series analysis in natural hydrological problems (Nourani et al. 2014b). As shown in Fig. 3, the DWT has two sets of functions, the original time series passing through high-pass (detail) and low-pass (approximate) filters, and decomposes at different scales. Eventually, are shown fast events and trends in Fig. 8. Wavelets retain the characteristics of the frequency domain and time



**Fig. 3** The process of decomposition of a time series by the DWT

domain described by the wavelet function (called the mother wavelet) and the scaling function (called the father wavelet). The mother wavelet mathematically is expressed as follows (Percival and Walden 2006):

$$\psi_{a,b}(t) = \frac{1}{\sqrt{a}} \psi\left(\frac{t-b}{a}\right), \quad (5)$$

where the coefficient “ $a$ ” is a positive number and the parameter “ $b$ ” is any real number.  $\psi_{a,b}(t)$  = wavelet function;  $a$  = frequency or scale (or dilated) parameter;  $b$  = translation or shifted parameter.

In the DWT, the scales “ $a$ ” and shift times “ $b$ ” in the mother wavelet is considered power-of-two, i.e., scale  $a = 2^m$  and location  $b = n2^m$ . For a discrete time-series  $x(t)$  decomposed into several finite subsets, which happens at a discrete-time  $t$ , DWT can have been calculated as follows (Mallat 1989):

$$\text{DWT}_{m,n}(t) = 2^{-m/2} \int_{-\infty}^{\infty} x(t) \psi^*\left(\frac{t-n2^m}{2^m}\right) dt, \quad (6)$$

where wavelet with dilation is by “ $m$ ,” and it is shifted by “ $n$ ,” and this way wavelet tune and control.

### Multi-discrete wavelet transforms (M-DWT), and its hybrid with ANN models

The principal purpose of using DWT as preprocessing is to provide more information to increase the understanding and accuracy of model forecasting (Maheswaran and Khosa 2012). In the last decade, hybrid modeling by wavelets-AI techniques has expanded significantly. Studies' results wide range of researchers represented the superior performance of coupling models compared with single models in accurately forecasting streamflow (Nourani et al. 2014b). The hybrid wavelet-AI model to achieve the ability has been designed to model non-linear. Choosing a suitable mother wavelet has an essential and significant role in wavelet-AI modeling. Time series of the hydrological phenomena have different characteristics due to the complexity and being affected by many parameters and can have long-term, short-term features, or various combinations. Hence, an appropriate mother wavelet can cover those processes compact or broad so that models provide better forecasts (Maheswaran and Khosa 2012). Thus, it seems that a combination of different mother wavelets for time series decomposition is more proper and can have better covering and more compatibility with other time series shapes.

On the other hand, hydrological phenomena have inherently intricate processes, and the sampled and observational data of these processes often contain noise and redundant information. The input data of a network must so was organized and prepared to obtain trusted multi-time-step ahead streamflow

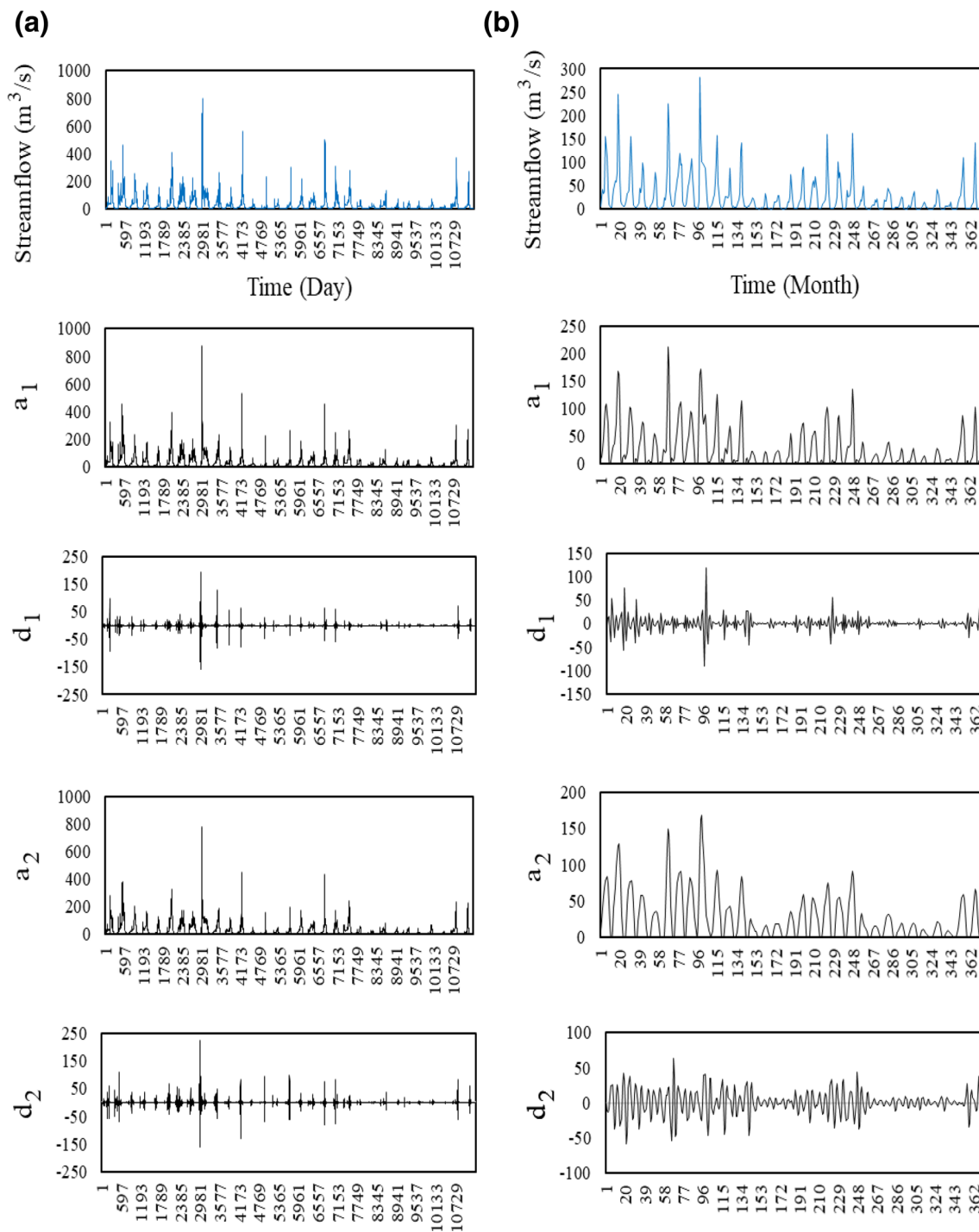
forecasting and such a way that it can adequately encompass wholly the information related to the desired data. Thus, decomposing or eliminating data noise is another fundamental step in modeling hydrological processes. The DWT-based method with decomposition and noise decrease can improve the performance of models if it has a suitable mother wavelet adequate decomposition level.

Thus, the importance of selecting and combining the appropriate mother wavelet increases in the utilization of multiple wavelets simultaneously. Accordingly, in this study, to obtain high-accuracy results, applied the DWT as a preprocessor was. Their effect on the performance of the models was evaluated and compared. DWT processes the original signal through high-pass and low-pass filters and decomposes them into subsets. Finally, have been used the sub-signals as input to ANN models. For example, Fig. 4 shows the approximations and details originating from the streamflow time series decomposed by the db7 mother wavelet at level 2. For each input data delay to the model was formed a time series. Then, each time series in each wavelet transform (WT) at decomposition level two was decomposing into four subsets including,  $a_1$ ,  $a_2$ ,  $d_1$ , and  $d_2$ . As represented in Fig. 5, the use of several wavelets simultaneously with different scaling and filters lengths can contain various parts of the signal (streamflow time series data). A suitable combination of them can lead to increased understanding and accuracy of data-driven models such as ANN.

In the present study, data preprocessing has been used to construct hybrid streamflow models both as single wavelets and simultaneously with several mother wavelets (Daubechies; db, Symlet; sym, Coiflet; coif, Biorthogonal; bior, and Fejer-Korovkin; fk). Then the decomposed data by DWT in several combinations were imported to the ANN models and as the DWT-AI and M-DWT-AI. The matter is means were used several different wavelets for data decomposition, and then all the decomposed data by these wavelets were fed together as input to the ANN model. In other words, this technique can consider as a manner of mining and fusion data with different features. The decomposition level selection and the type and number of wavelets in DWTs were designated using a trial-and-error method. In such a way that the optimal value of the parameters and variables based on performance criteria, including the RMSE,  $R$ , and NSE by sensitivity analysis, was determined. Meanwhile, all the ANN models, combinations of its, and DWTs, have been coded in MATLAB R2018 software.

## Results and discussion

This study aims to apply artificial neural network (ANN) models to provide more accurate streamflow forecasting by introducing a new technique (i.e., M-DWT) in the short-term up to 7 days and long-term up to 7 months beyond data records. For that purpose, after selecting the best network

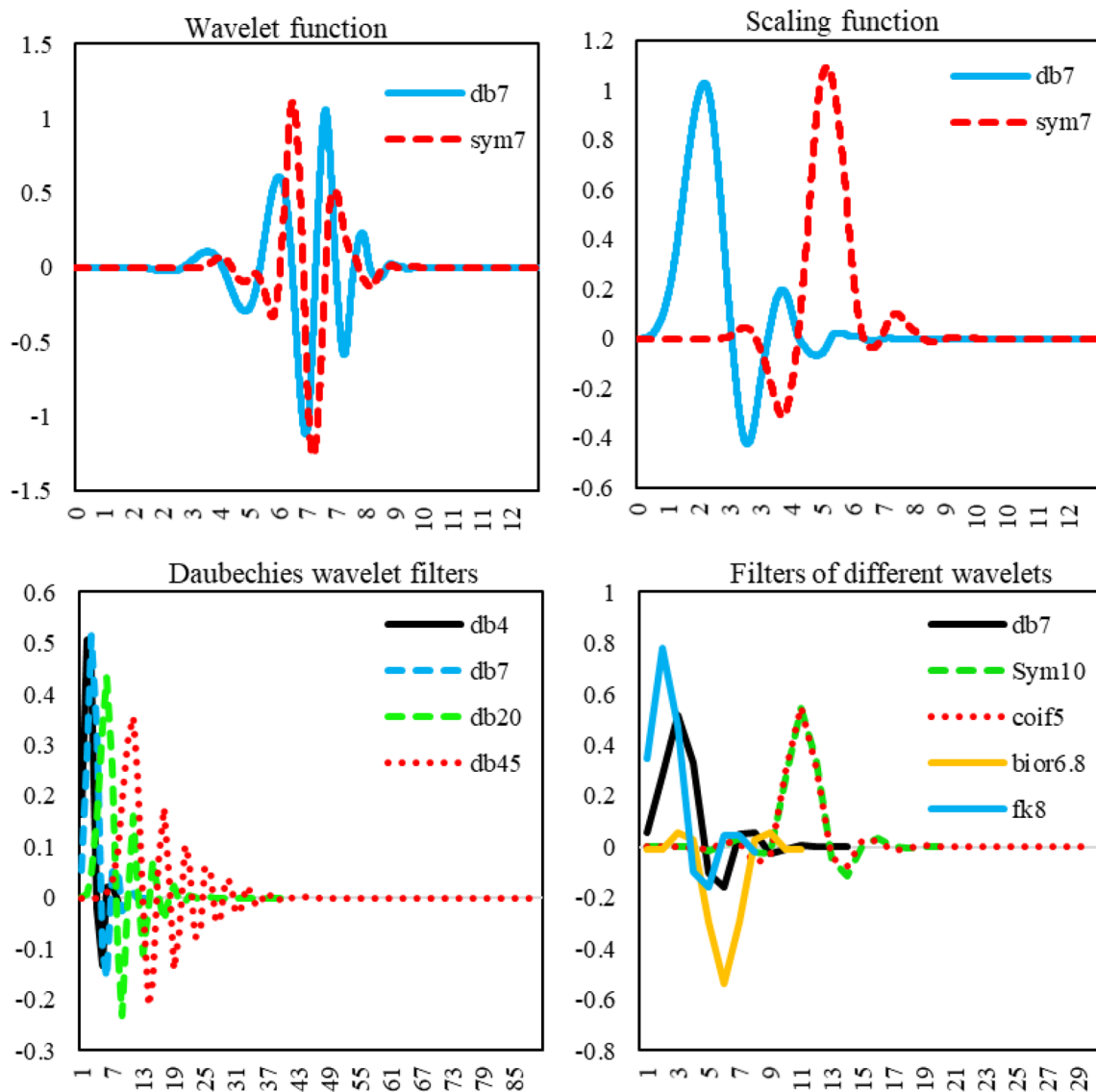


**Fig. 4** Approximation ( $a_1$ ,  $a_2$ ) and detail ( $d_1$ ,  $d_2$ ) sub-signals of streamflow time series with the time scale **a** daily, **b** monthly, decomposed by db7 wavelet at level two

architecture of the model and the respective input composition, it has been used to forecast the streamflow fluctuations of 1-, 2-, 3-, and 7-time-step ahead for daily and monthly time scales. Monthly flow rate forecasts are more used to manage water resources, and the daily flow rate forecasts are more applied to reduce flood risks.

Initially, the ANN model was used to model and forecast streamflow of catchment with daily and monthly time scales without any preprocessing of input data. Table 1 represents the RMSE (m<sup>3</sup>/s) for several different ANN architectures as an instance. Each time in the model has been placed five to 20 neurons in a single hidden layer





**Fig. 5** Daubechies; db, Symlet; sym, Coiflet; coif, Biorthogonal; bior, and Fejer-Korovkin; fk wavelets with different filter lengths and scaling functions

for network training and was selected the optimal architecture according to the better performance of the model based on RMSE, *R*, and NSE performance criteria in both calibration and verification stages. The optimal number of neurons in the hidden layer, i.e., five neurons, was used to predict at different time stages. As shown in Table 1, was presenting several combinations of input data and various architectures for the ANN model as instances also optimal values of the artificial network structure and input data are highlighted. On a daily time scale, the best network structure, i.e., 22-5-1, and the best input combination, i.e., a synthesis with a delay of 1 day for precipitation, temperature, and 1–20 days for streamflow, is presented. The model was selected with the most negligible value RMSE

and the most value NSE and *R* in the verification stage as the optimal network (i.e., RMSE = 4.79, NSE = 0.96, and *R* = 0.98). Also, on a monthly time scale, given the slightest error in the verification stage, the best configuration (i.e., 20-5-1) for a combination set of inputs and networks is to forecast the streamflow one month ahead. This structure, the best input combination, i.e., the combination with delays of 1–20 months, and the most appropriate number of neurons in the hidden layer, i.e., five neurons, is for forecasts at different time steps (see Table 1). After determining the most appropriate ANN configuration based on the performance criteria, was completed the network calibration. The weights obtained from the network neurons in the calibration stage generated in the network are

**Table 1** RMSE (m<sup>3</sup>/s) results of the ANN model and a hybrid with different mother wavelets for one-time-step-ahead streamflow forecasted in calibration and verification periods

Model inputs with delay (Wavelet <sub>type</sub> -Inputs)	Structure network	Daily time scale		Monthly time scale	
		RMSE (m <sup>3</sup> /s) cali- bration	RMSE(m <sup>3</sup> /s) verification	RMSE(m <sup>3</sup> /s) calibration	RMSE(m <sup>3</sup> /s) verification
$Q(t-1)$	1-20-1	14.80	6.08	31.62	14.84
$Q(t-1)$	1-10-1	15.21	6.13	32.47	16.81
$Q(t-1)$	1-5-1	15.50	6.01	32.84	16.15
$Q(t-1, -2, -3, -4, -5)$	5-5-1	14.04	6.40	29.18	17.12
<b><math>Q(t-1, -2, \dots, -20)</math></b>	<b>20-5-1</b>	10.67	9.52	<b>24.81</b>	<b>14.54</b>
$Q(t-1, -2, \dots, -20)$	20-10-1	10.32	7.39	23.60	14.68
$Q(t-1, -2, \dots, -20)$	20-20-1	7.79	20.11	26.11	14.88
$Q(t-1), T(t-1), P(t-1)$	3-5-1	13.33	4.94	29.16	18.12
$Q(t-1), T(t-1), P(t-1)$	3-10-1	11.28	8.49	29.17	18.09
$Q(t-1, -2, -3), T(t-1), P(t-1)$	5-5-1	13.22	7.31	28.94	17.73
$Q(t-1, -2, \dots, -10), T(t-1), P(t-1)$	12-5-1	12.86	5.02	18.47	15.29
<b><math>Q(t-1, -2, \dots, -20), T(t-1), P(t-1)</math></b>	<b>22-5-1</b>	<b>8.52</b>	<b>4.79</b>	25.28	15.78
$Q(t-1, -2, -3), T(t-1, -2, -3), P(t-1, -2, -3)$	9-5-1	10.75	5.37	20.88	15.75
$Q(t-1, -2, \dots, -7), T(t-1, -2, \dots, -7), P(t-1, -2, \dots, -7)$	21-5-1	9.06	27.80	20.02	15.66
$Q(t-1, -2, \dots, -7), T(t-1, -2, \dots, -7), P(t-1, -2, \dots, -7)$	21-10-1	7.49	18.81	1.83	44.97
$Q(t-1, -2, \dots, -20), T(t-1, -2, \dots, -20), P(t-1, -2, \dots, -20)$	60-5-1	7.25	14.62	0.001	27.85
$W_{db45}-Q(t-1), W_{db45}-T(t-1), W_{db45}-P(t-1)$	12-5-1	6.65	3.88	10.56	11.91
$W_{db45}-Q(t-1), W_{db45}-T(t-1), W_{db45}-P(t-1)$	12-10-1	5.22	4.23	5.49	14.96
$W_{sym10}-Q(t-1), W_{db7}-T(t-1), W_{db7}-P(t-1)$	12-5-1	6.66	5.94	10.39	13.09
$W_{db45}-Q(t-1), W_{sym4}-T(t-1), W_{sym4}-P(t-1)$	12-5-1	6.79	4.97	7.78	11.49
$W_{db45}-Q(t-1, -2, -3), W_{db45}-T(t-1, -2, -3), W_{db45}-P(t-1, -2, -3)$	36-5-1	2.68	2.85	2.23	6.02
$W_{db8}-Q(t-1, -2, \dots, -7), W_{sym8}-T(t-1, -2, \dots, -7), W_{sym8}-P(t-1, -2, \dots, -7)$	84-5-1	1.52	1.93	1.47	7.30
<b><math>W_{db45}-Q(t-1, -2, \dots, -7), W_{db45}-T(t-1, -2, \dots, -7), W_{db45}-P(t-1, -2, \dots, -7)</math></b>	<b>84-5-1</b>	<b>0.55</b>	<b>0.59</b>	<b>0.27</b>	<b>4.86</b>
$W_{db45}W_{sym10}-Q(t-1, -2, \dots, -20), W_{db7}-T(t-1), W_{db7}-P(t-1)$	168-5-1	0.057	0.041	0.057	3.08
$W_{db45}W_{sym10}-Q(t-1, -2, \dots, -20), W_{db45}-T(t-1), W_{db45}-P(t-1)$	168-5-1	0.059	0.037	0.005	2.88
$W_{db45}W_{sym10}-Q(t-1, -2, \dots, -20), W_{db4}-T(t-1), W_{db4}-P(t-1)$	168-5-1	0.054	0.035	0.17	3.31
$W_{db45}W_{sym4}W_{coif5}W_{bior5.5}W_{fk6}-Q(t-1), W_{db45}-T(t-1), W_{db45}-P(t-1)$	28-5-1	5.04	4.22	5.75	6.73
$W_{db8}W_{sym4}W_{coif5}W_{bior5.5}W_{fk6}-Q(t-1), W_{db4}-T(t-1), W_{db4}-P(t-1)$	28-5-1	4.92	3.89	4.65	6.75
$W_{db8}W_{sym4}W_{coif5}W_{bior5.5}W_{fk6}-Q(t-1), W_{sym8}-T(t-1), W_{sym8}-P(t-1)$	28-5-1	4.90	3.68	4.84	6.69
$W_{db8}W_{sym4}W_{coif5}W_{bior5.5}W_{fk6}-Q(t-1, -2, -3), W_{sym8}-T(t-1, -2, -3), W_{sym8}-P(t-1, -2, -3)$	84-5-1	0.54	0.54	0.40	1.18
$W_{db8}W_{sym4}W_{coif5}-Q(t-1, -2, -3), W_{db8}W_{sym4}W_{coif5}-T(t-1, -2, -3), W_{db8}W_{sym4}W_{coif5}-P(t-1, -2, -3)$	108-5-1	0.956	1.30	0.52	4.37
$W_{db8}W_{sym4}W_{coif5}W_{bior5.5}W_{fk6}-Q(t-1, -2, \dots, -7), W_{sym8}-T(t-1, -2, \dots, -7), W_{sym8}-P(t-1, -2, \dots, -7)$	196-5-1	0.031	0.030	0.004	1.75
$W_{db8}W_{sym20}W_{coif5}W_{bior6.8}W_{fk8}-Q(t-1, -2, \dots, -8), W_{db45}-P(t-1)$	164-5-1	0.018	0.013	0.022	0.52
$W_{db8}W_{sym20}W_{coif5}W_{bior6.8}W_{fk8}-Q(t-1, -2, \dots, -6), W_{db45}-P(t-1)$	124-5-1	0.050	0.037	0.016	0.27

**Table 1** (continued)

Model inputs with delay (Wavelet <sub>type</sub> -Inputs)	Structure network	Daily time scale		Monthly time scale	
		RMSE (m <sup>3</sup> /s) cali-bration	RMSE(m <sup>3</sup> /s) verification	RMSE(m <sup>3</sup> /s) calibration	RMSE(m <sup>3</sup> /s) verification
$W_{db7}W_{sym10}W_{coif5}W_{bior6.8}W_{fk8}-Q(t-1, -2, \dots, -6), W_{db45}-P(t-1)$	<b>124-5-1</b>	0.037	0.021	<b>0.0061</b>	<b>0.15</b>
$W_{db8}W_{sym4}W_{coif5}W_{bior5.5}W_{fk6}-Q(t-1, -2, \dots, -10), W_{db4}-T(t-1), W_{db4}-P(t-1)$	208-5-1	0.00198	0.00125	0.015	0.86
$W_{db8}W_{sym4}W_{coif5}W_{bior5.5}W_{fk6}-Q(t-1, -2, \dots, -20), W_{db4}-T(t-1), W_{db4}-P(t-1)$	408-5-1	0.00113	0.00087	0.0001	1.87
$W_{db45}W_{sym4}W_{coif5}W_{bior5.5}W_{fk6}-Q(t-1, -2, \dots, -20), W_{sym8}-T(t-1), W_{sym8}-P(t-1)$	<b>408-5-1</b>	<b>0.00042</b>	<b>0.00038</b>	0.0008	1.98
$W_{db20}W_{sym8}W_{coif2}W_{bior5.5}W_{fk14}-Q(t-1, -2, \dots, -20), W_{db4}-T(t-1), W_{sym8}-P(t-1)$	408-5-1	0.00049	0.00042	0.0004	2.43

Input variables are *P* precipitation, *T* temperature, *Q* streamflow data sets

stored, and these weights in the verification (testing) stage be applied.

In the next step, has been used the single wavelet (i.e., DWT) to preprocess the input data to the ANN model. The results showed that applying DWT data can significantly improve the performance of the models. The results of the present study are consistent and the findings of other researchers, including Adamowski and Sun (2010) and Tayyab et al. (2019). Therefore, due to the usefulness of DWT hybrids with AI models such as ANN, in this study, the multiple wavelet simultaneous technique (i.e., M-DWT) to better decompose the data was used and increased the streamflow forecasting accuracy.

The M-DWT was used to preprocess the input time series to the ANN model. As shown in Table 1, was presenting several combinations of input data and different architectures for the DWT-ANN and M-DWT-ANN models as examples and are highlighted optimal values of the artificial network structure and input data. Indeed, the bolded part of this table shows the most accurate models for forecasting streamflow. For the daily time scale, the most optimal network structure, i.e., is 408-5-1 for the M-DWT-ANN model is from was obtained 1–20 delays for streamflow data and one delay for temperature and precipitation data. The best multi-mother wavelets simultaneously for this model include db45, sym4, coif5, bior5.5, and fk6 wavelets, which provided more accurate forecasting of the streamflow of one and several steps ahead compared to other combinations based on RMSE performance criteria (see Table 1). Also, for the monthly time scale, the most optimal network structure, i.e., 124-5-1 for the M-DWT-ANN model, was presented with 1–6 delays for streamflow data and one delay for precipitation data. The best results for preprocessing by multi-mother wavelets for the M-DWT-ANN model are the db7, sym10, coif5, bior6.8, and fk8 wavelets for the streamflow

time series and the db45 wavelet for the precipitation time series. After selecting the optimal developed models were utilized for different time steps.

It is worth noting that selecting an optimal architecture for ANN models is an essential step in modeling because improper architecture can lead to under/over-fitting and under/over-computing more problems. Furthermore, in modeling data-driven models such as ANN, specific attention be paid to the appropriate selection of inputs, which can upgrade the model's efficiency in both calibrations (training) and verification (testing) stages. The effects of the cases mentioned in the model performance can have been seeing in Table 1. It is worth noting that some network structures had better performance in the calibration stage but were weaker performance in the verification stage according to RMSE values (see Table 1). The results obtained can be related to the model performance is that has been over-fitting with the target data in the training stage. The error of the testing period is usually more significant than the error of the training period because unknown values for the model for evaluation in the test period are used (Tapoglou et al. 2014). As shown in Table 1, this principle is reversed in this study due to the difference in amplitude of streamflow fluctuations in the training and test periods in some results according to the RMSE criterion (i.e., the amount of test error is less than the training error). This performance is due to the dependence of the RMSE criterion on the scale of variables. But this has been correctly assessed according to NSE and *R* criteria.

Finally, the performance of the ANN model and its hybrids with the wavelet analysis were compared and evaluated in streamflow forecasting. As shown in the figures and tables in this section, models calibrated with simultaneous multi-wavelet preprocessing perform much better than single-wavelet models and models without preprocessing.

According to the obtained results, the models developed by the M-DWT technique have high efficiency. Although this technique increases the amount and time of calculations, it dramatically improves the performance of the models. As Fig. 6 clearly shows, this betters the performance of the ANN model in the training period, and the hybrid models, in particular, coupled with M-DWT, had lower RMSE error rates in the number of iterations (epochs) higher. The use of wavelet analysis, especially multi-mother wavelets, increases the model input information and also increases the model's understanding of the data behavior patterns. In the following, are presented and analyzed the obtained results.

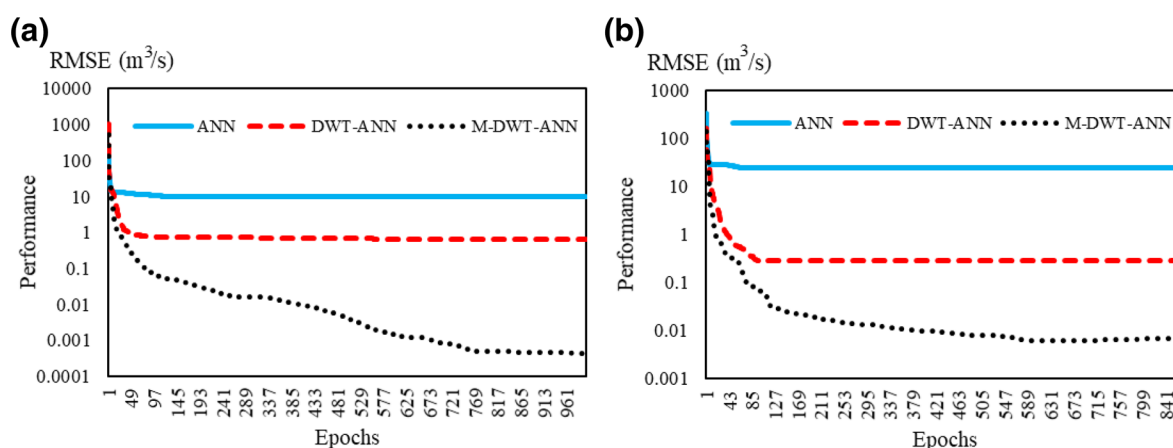
### Results of ANN models and their hybrid with DWT and M-DWT, with daily time scale

In this study, various input combinations, and different network structures for each of these inputs, were used. ANN models use past data, including streamflow, precipitation, and temperature and as input to forecast streamflow. The hybrid ANN model with one or more wavelets decomposition (i.e., the DWT-ANN and M-DWT-ANN hybrid models) is composed to reduce the forecasting error. The noise and various information contained in the data were separated and preprocessed by the DWT were utilized as input to forecast the streamflow. Figure 7 shows the results for the best ANN structure and the best DWT-ANN and M-DWT-ANN structure to forecast daily streamflow for one-time ahead. As can be seen in this figure, multi-wavelet hybrid ANN models can model and forecast the flow discharge peaks well more accurately than the single model. In Fig. 8 are shown the scatter and time series for comparing the observed and calculated streamflow. As shown in this figure, the output of the M-DWT-ANN model at different time steps has less scatter, and its values are more compact in the proximity of the direct-line than the DWT-ANN and ANN models, which

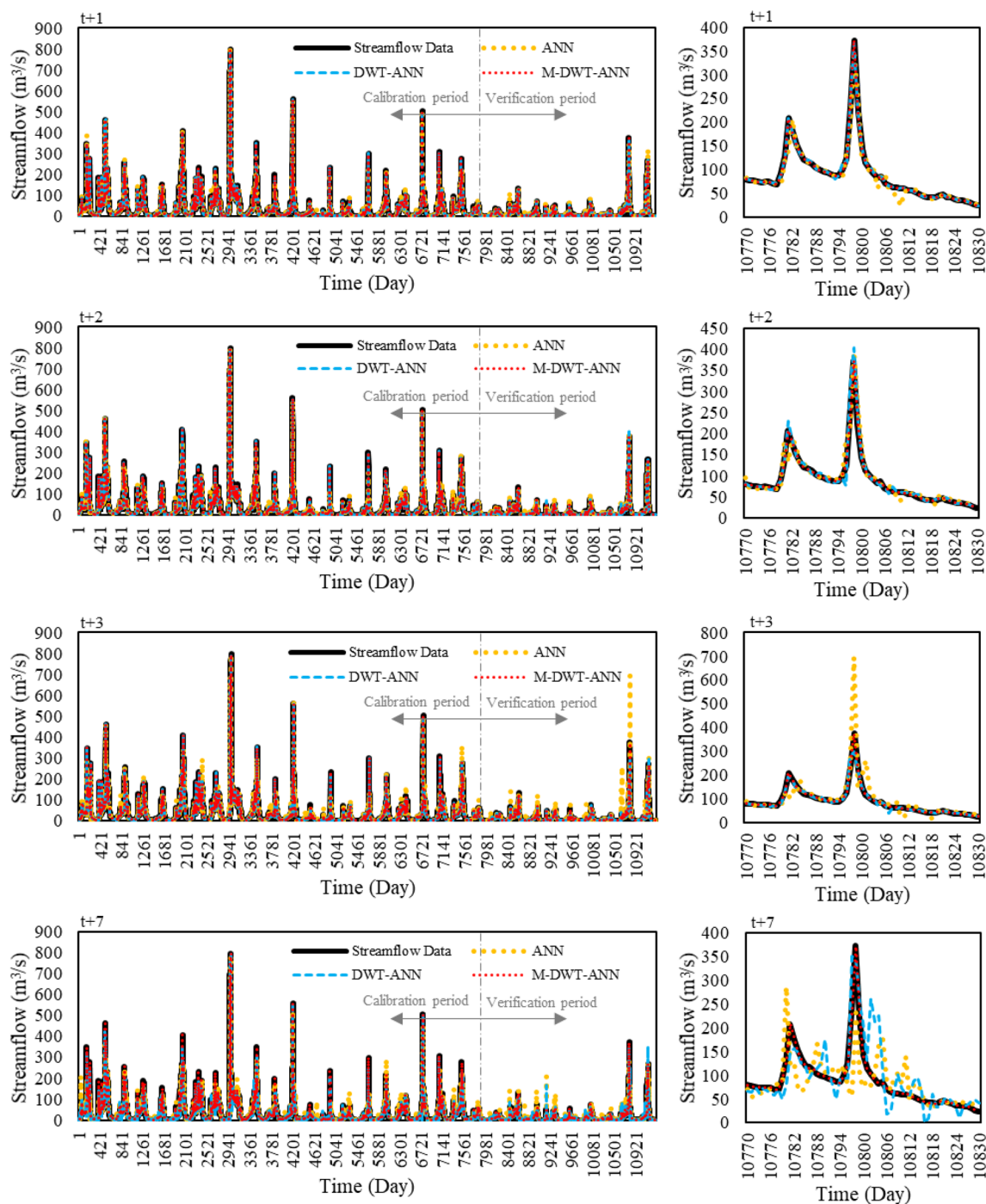
indicate the better performance of this model is. The RMSE values of the ANN, DWT-ANN, and M-DWT-ANN models are represented in Table 1 for the forecast one step ahead of time. Table 2 presents the R, RMSE, and NSE values of the ANN, DWT-ANN, and M-DWT-ANN models for forecasting different time steps. The best models with the slightest error and the most efficiency and correlation were used to forecast streamflow. Figure 6 clearly shows the proficiency of the wavelet-neural-network model to learn the non-linear relationship between input and target data. Results in the verification stage, for the 1-, 2- and 3-day ahead, the NSE and R values are entire close to 1, and the RMSE values are less than 3 for the DWT-ANN model and less than 0.01 for the M-DWT-ANN model. Increasing the forecast intervals affected adversely on forecasts, the R and NSE decreased, and RMSE increased. Yet, the forecasting results at 1-, 2-, 3-, and 7-day ahead time scales for used ANN and DWT-ANN models and especially M-DWT-ANN models are acceptable. For example, the NSE criterion values for the ANN model in the 1, 2, 3, and 7 steps ahead are 0.96, 0.90, 0.75, and 0.66, respectively. Comparing the results of the models, the M-DWT-ANN model performs better than the ANN and DWT-ANN models based on the R, RMSE, and NSE (see Table 2). For instance, the RMSE value in the verification stage for the best ANN, DWT-ANN, and M-DWT-ANN models, to forecast streamflow two time-steps 7.83 ( $\text{m}^3/\text{s}$ ), 2.32 ( $\text{m}^3/\text{s}$ ), and 0.0056 ( $\text{m}^3/\text{s}$ ), respectively, were obtained. Overall, M-DWT-ANN model forecasts in different time steps are superior to ANN and DWT-ANN models.

### Results of ANN models and their hybrid with DWT and M-DWT, with monthly time scale

The data set of this study, into two subsets, including calibration and validation sets, are divided. Also, different combinations of models' inputs were used similar to the daily



**Fig. 6** The performance of ANN, DWT-ANN, and M-DWT-ANN in RMSE( $\text{m}^3/\text{s}$ ) with the time scale a) Daily, b) Monthly, for training data set

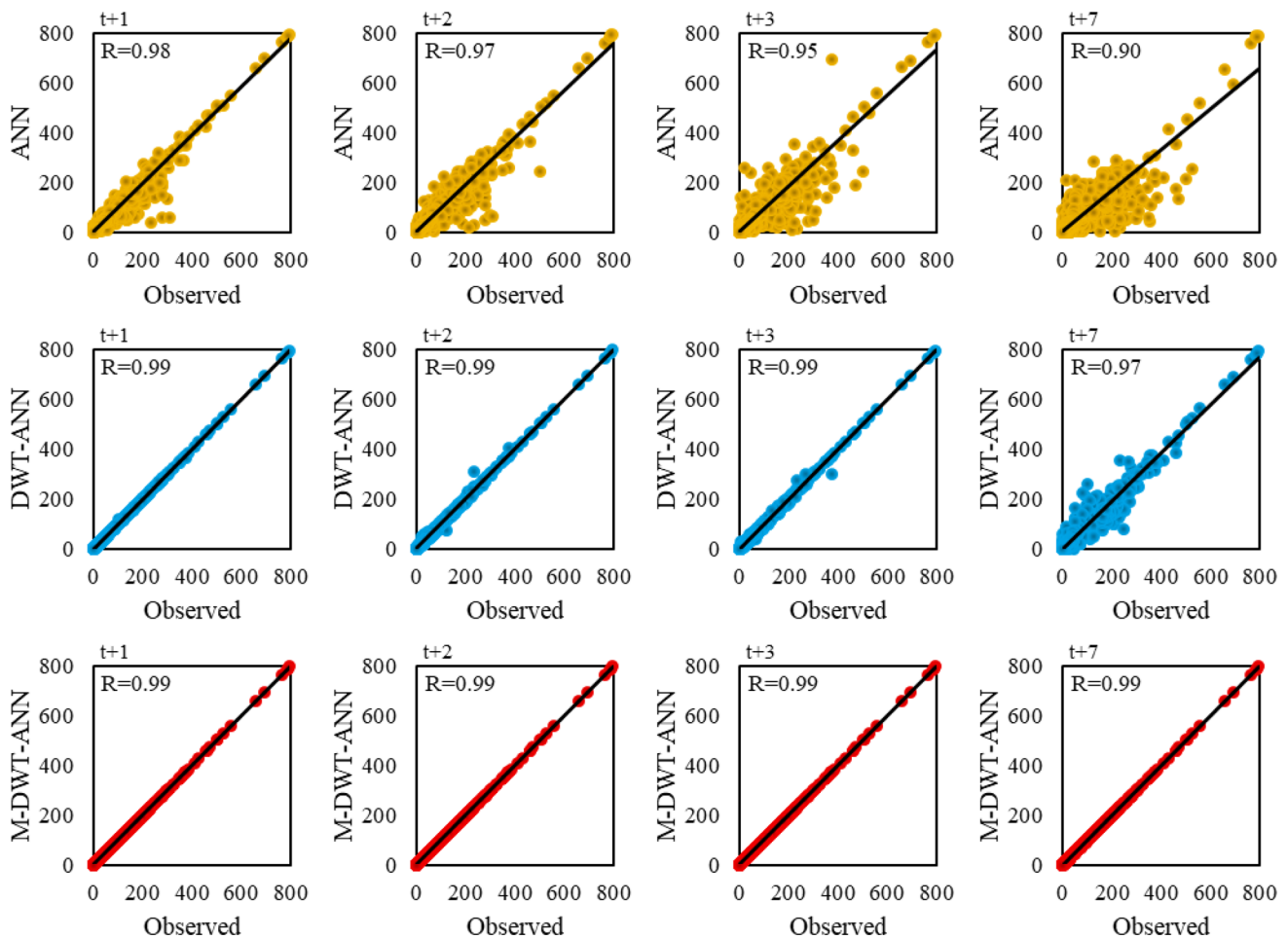


**Fig. 7** Comparing observed and forecasted streamflow using the best ANN, DWT-ANN, and M-DWT-ANN models for multi-time-step-ahead forecasted during the calibration and verification periods in the daily time scale

time scale in a monthly time scale. ANN models used the monthly average of past data, including streamflow, precipitation, and temperature, as input to forecast streamflow. Table 1 represents the RMSE ( $m^3/s$ ) for several different ANN structures as an instance. The model structure with the slightest value RMSE and the most value NSE and  $R$  was selected, In a balanced way in both the calibration and

verification stages, as the optimal structure. In Table 2, the RMSE, NSE, and  $R$  criteria values for the best ANN model were 14.54, 0.52, and 0.76, respectively, was presented.

The DWT-ANN and M-DWT-ANN hybrid models were used to minimize the forecasting error for the monthly time scale. The results indicate the usefulness of the ANN model coupled with wavelets. Figure 9 shows the superior



**Fig. 8** Scatter plot between observed and forecasted streamflow using ANN model and its hybrids with DWT and M-DWT for multi-time-step-ahead during the calibration and verification periods in the daily time scale

performance of the M-DWT-ANN hybrid model over the ANN and DWT-ANN models in modeling and forecasting streamflow. Multi-wavelet hybrid ANN models have been able to model and forecast flow discharge peaks well. The poorer performance of the ANN model is due to the inputs of noisy data. Also, Fig. 10 shows scatterplots comparing the observed and forecasted streamflow using the best M-DWT-ANN model and the best ANN and DWT-ANN models for one month ahead forecasting during the calibration (training) and verification (testing) periods. Also, estimates of the M-DWT-ANN model have fewer scattered, and its values are more compact in the proximity of the direct-line, compared to the DWT-ANN and ANN models, and means better performance of the M-DWT-ANN model. Presented in Table 1 are the RMSE values of the ANN, DWT-ANN, and M-DWT-ANN models for the forecast one-time-step-ahead.

M-DWT-ANN model is the most accurate streamflow forecast that offers one or more steps ahead of other combinations on a monthly time scale based on RMSE performance criteria (see Table 1). Table 2 presents the R, RMSE,

and NSE values of the ANN, DWT-ANN, and M-DWT-ANN models for forecasting different time steps. The best models with the slightest error and the most efficiency and correlation were used to forecast streamflow in the calibration and verification periods. Results in the verification stage for the NSE criterion values for the M-DWT-ANN model in the 1, 2, 3, and 7 steps ahead are 0.99, 0.99, 0.95, and 0.59, respectively. Increasing the forecast intervals affected unfavorably on forecasts, the R and NSE decreased, and RMSE increased. However, the forecasting results at 1-, 2-, 3-, and 7-month ahead time scales for used ANN and DWT-ANN models and especially M-DWT-ANN models are acceptable. Comparing the results of the models, the M-DWT-ANN model performs better than the ANN and DWT-ANN models based on the R, RMSE, and NSE (see Table 2). For instance, the RMSE value in the verification stage for the best ANN, DWT-ANN, and M-DWT-ANN models, to forecast streamflow two time-steps-ahead 18.18 ( $\text{m}^3/\text{s}$ ), 11.10 ( $\text{m}^3/\text{s}$ ), and 1.68 ( $\text{m}^3/\text{s}$ ), respectively, were obtained. The results indicate that ANN models coupled

**Table 2** Results of the best ANN models and their hybrids with DWT and M-DWT for one and multi-step-ahead streamflow forecasted in calibration and verification periods

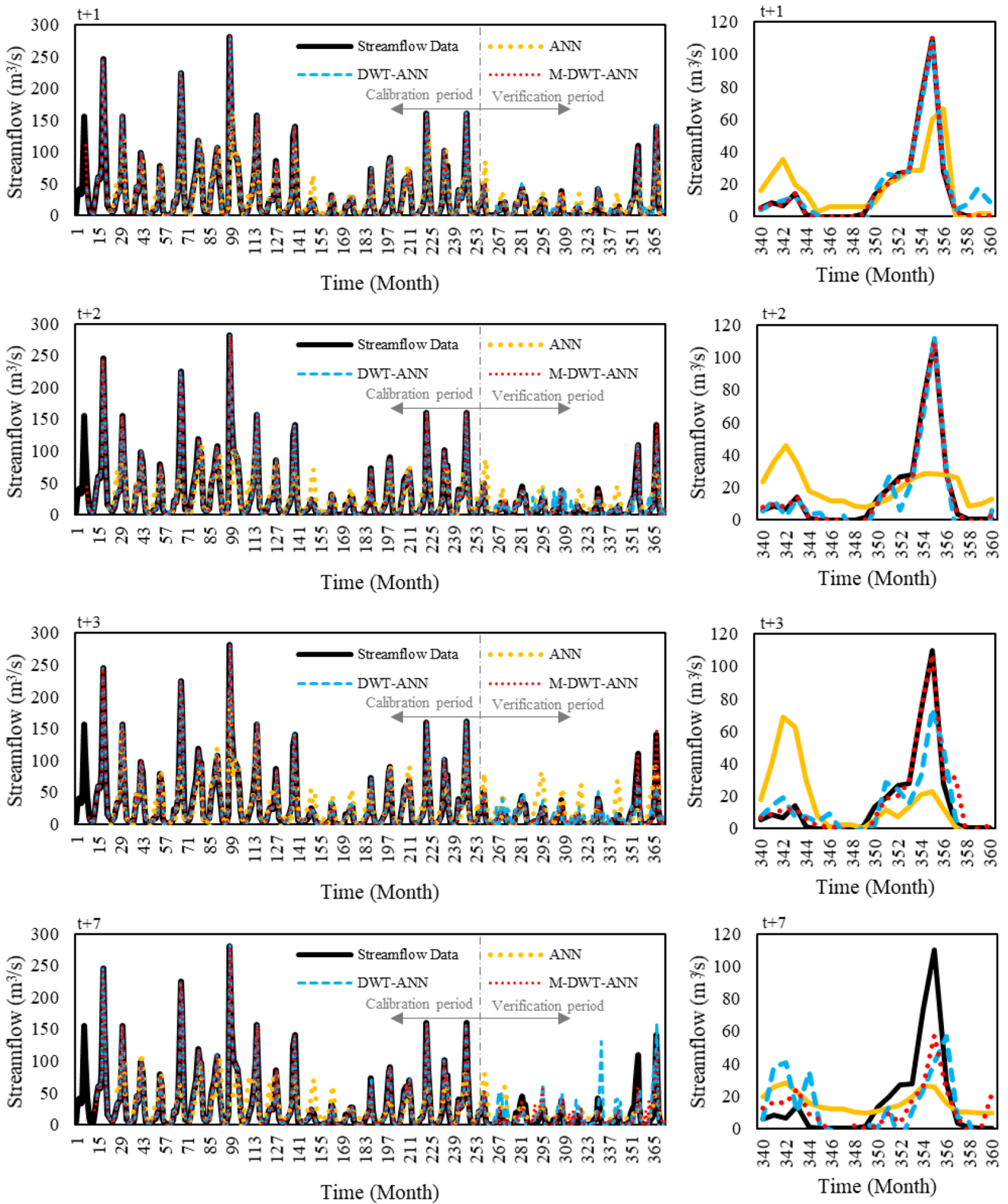
Forecasted	Model type	R		RMSE (m <sup>3</sup> /s)		NSE	
		Calibration (Train)	Verification (Test)	Calibration (Train)	Verification (Test)	Calibration (Train)	Verification (Test)
Q(t+1)	Daily time scale ANN	0.98	0.98	8.5167	4.7906	0.97	0.96
	DWT-ANN	0.99	0.99	0.5452	0.5923	0.99	0.99
	M-DWT-ANN	0.99	0.99	0.00042	0.00038	0.99	0.99
	Monthly time scale ANN	0.80	0.76	24.8136	14.5445	0.64	0.52
	DWT-ANN	0.99	0.97	0.27235	4.8576	0.99	0.95
	M-DWT-ANN	0.99	0.99	0.0061	0.15207	0.99	0.99
Q(t+2)	Daily time scale ANN	0.97	0.95	12.3416	7.8307	0.94	0.90
	DWT-ANN	0.99	0.99	1.6611	2.3172	0.99	0.99
	M-DWT-ANN	0.99	0.99	0.0055	0.0056	0.99	0.99
	Monthly time scale ANN	0.71	0.64	29.2931	18.1782	0.50	0.26
	DWT-ANN	0.99	0.86	1.4002	11.0975	0.99	0.72
	M-DWT-ANN	0.99	0.99	0.3983	1.6751	0.99	0.99
Q(t+3)	Daily time scale ANN	0.95	0.89	16.0690	12.2696	0.90	0.75
	DWT-ANN	0.99	0.99	1.8700	2.3403	0.99	0.99
	M-DWT-ANN	0.99	0.99	0.0127	0.0093	0.99	0.99
	Monthly time scale ANN	0.86	0.51	21.2007	23.2186	0.74	0.20
	DWT-ANN	0.99	0.71	1.5056	14.9969	0.99	0.48
	M-DWT-ANN	0.99	0.97	1.3977	4.6660	0.99	0.95
Q(t+7)	Daily time scale ANN	0.90	0.84	22.7003	14.1048	0.81	0.66
	DWT-ANN	0.98	0.92	10.6143	9.9572	0.96	0.83
	M-DWT-ANN	0.99	0.99	0.1716	0.2615	0.99	0.99
	Monthly time scale ANN	0.65	0.58	31.7325	18.3942	0.42	0.25
	DWT-ANN	0.99	0.58	1.7720	23.4234	0.99	-0.25
	M-DWT-ANN	0.98	0.78	2.0204	13.3593	0.99	0.59

with multi-wavelet with good accuracy simulated and forecasted flow discharge peaks. In general, M-DWT-ANN model forecasts in different time steps are better than ANN and DWT-ANN models.

### Conclusions

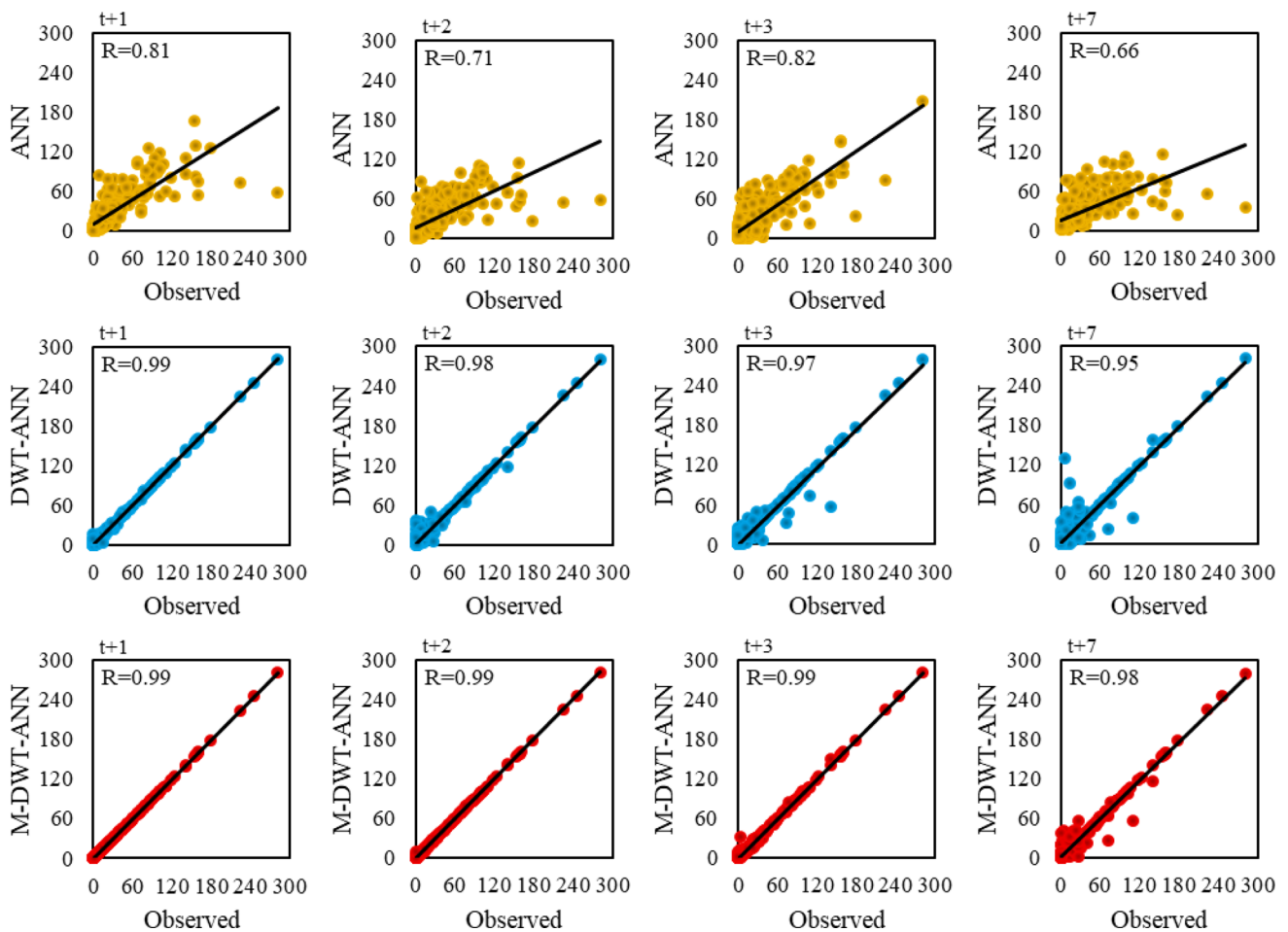
The present research suggests a technique based on a hybrid of multi-discrete wavelet transform (M-DWT) and artificial neural network (ANN) model to high accuracy streamflow forecasting. The proposed method with a better understanding of the behavioral patterns of hydrological phenomena can further help engineers and managers in floods control and sustainable management of water resources. To evaluate

and prove efficiency M-DWT-ANN model was compared to the mono-wavelet DWT-ANN model and single ANN model for streamflow forecasting. For this purpose, the data were analyzed and examined for 31 years with daily and monthly time scales for 1, 2, 3, and 7 time-steps-ahead in the catchment area of Gamasiab River located in western Iran. At first, the input data into ANN models were entered, with sans none preprocessing and in raw form. The outcomes represented that these models (i.e., ANN model) could not cope with the non-linear and complex conduct of data. In the next step, the DWT on time-series data of the streamflow, temperature, and precipitation, and then preprocessed data, as input of the ANN models, were utilized. Using DWT, each of the original time-series containing noisy data have decomposed into sub-signal sets, extracted valuable information hidden in



**Fig. 9** Comparing observed and forecasted streamflow using the best ANN, DWT-ANN, and M-DWT-ANN models for multi-time-step-ahead forecasted during the calibration and verification periods in the monthly time scale





**Fig. 10** Scatter plot between observed and forecasted streamflow using ANN model and its hybrids with DWT and M-DWT for multi-time-step-ahead during the calibration and verification periods in the monthly time scale

the data, and ultimately increased the models' understanding of the streamflow process. The present study results showed best DWT-ANN model is more accurate in forecasting future streamflow at daily and monthly time scales than the best ANN model.

Each mother wavelet has a unique feature. Therefore, using several different wavelets and passing data through various filters by separating and adding information can help understand and learn better ANN models. The reason is that each parsed time series contains hidden parts and aspects of the original time series (data). The results showed that the combination of several of them as input to ANN models led to improved training, increased accuracy, and reduced error of these models in the forecast ahead streamflow. To create and develop models from a hybrid with an appropriate selection of the multi-wavelets and the ANN models (such as db, bior, coif, sym, and fk wavelets) was used as data preprocessing under the title of M-DWT-ANN. The use of the M-DWT technique significantly improved the performance of the model. The developed models can cope well with various non-linear

characteristics of the streamflow process. This study indicated that the M-DWT-ANN model can are forecast streamflow with very high accuracy. Overall, the results showed that the best M-DWT-ANN model better performance in forecasting streamflow at daily and monthly time scales for different time steps ahead, compared to the best ANN and DWT-ANN models based on RMSE, *R*, and NSE performance criteria. It means that the M-DWT-ANN model has the lowest RMSE value and the highest *R* and NSE value compared to other models. In the flow discharge scrutiny, the flow peak discharge is of extraordinary importance, and in this regard, too M-DWT-ANN model has higher efficiency than other models. Surveys showed that forecasting streamflow with a daily time scale is more accurate than comparing a monthly time scale. This outcome may be due to the greater dependence of daily flow fluctuations on previous data and less correlation between average monthly streamflow data with earlier time delays. It is noteworthy that the number of the data (samples) has a very positive effect on the performance of data-driven models (such as ANN), and it's clear that the daily

data is more than the monthly data. Another conclusion is that with increasing the time steps intervals, less conformity was observed between the measured and forecasted data. The forecast streamflow results for the catchment of the Gamasiab River show that the M-DWT-ANN method is an effective and mighty method for streamflow forecasting by detecting hidden and important hydrological parameters. Accurate streamflow forecasting can help relevant experts and managers to sustainable exploitation and optimal management of water resources. Considering the advantages of the proposed method is recommended that in future studies, the M-DWT-AI technique in forecasting streamflow is put under consideration in other hydrological phenomena for other catchments in different geographical and climate areas.

## References

- Adamowski JF (2008) Development of a short-term river flood forecasting method for snowmelt driven floods based on wavelet and cross-wavelet analysis. *J Hydrol* 353:247–266. <https://doi.org/10.1016/j.jhydrol.2008.02.013>
- Adamowski J, Chan HF (2011) A wavelet neural network conjunction model for groundwater level forecasting. *J Hydrol* 407:28–40. <https://doi.org/10.1016/j.jhydrol.2011.06.013>
- Adamowski J, Sun K (2010) Development of a coupled wavelet transform and neural network method for flow forecasting of non-perennial rivers in semi-arid watersheds. *J Hydrol* 390:85–91. <https://doi.org/10.1016/j.jhydrol.2010.06.033>
- Ahmadi M, Moeini A, Ahmadi H, Motamedvaziri B, Zehatabian GR (2019) Comparison of the performance of SWAT, IHACRES and artificial neural networks models in rainfall-runoff simulation (case study: Kan watershed, Iran). *Phys Chem Earth* 111:65–77. <https://doi.org/10.1016/j.pce.2019.05.002>
- Ahooghalandari M, Khiadani M, Kothapalli G (2016) Assessment of Artificial Neural Networks and IHACRES models for simulating streamflow in Marillana catchment in the Pilbara, Western Australia. *Austr J Water Resour* 19:116–126. <https://doi.org/10.1080/13241583.2015.1116183>
- Ali S, Shahbaz M (2020) Streamflow forecasting by modeling the rainfall-streamflow relationship using artificial neural networks. *Model Earth Syst Env* 6:1645–1656. <https://doi.org/10.1007/s40808-020-00780-3>
- Arnold JG, Moriasi DN, Gassman PW, Abbaspour KC, White MJ, Srinivasan R, Santhi C, Harmel RD, Van Griensven A, Van Liew MW, Kannan N, Jha MK (2012) SWAT: model use, calibration, and validation. *Trans ASABE* 55:1491–1508
- ASCE Task Committee (2000a) Artificial neural networks in hydrology-I: preliminary concepts. *J Hydrol Eng* 5(2):115–123. [https://doi.org/10.1061/\(asce\)1084-0699\(2000\)5:2\(115\)](https://doi.org/10.1061/(asce)1084-0699(2000)5:2(115))
- ASCE Task Committee (2000b) Artificial neural networks in hydrology-II: hydrologic applications. *J Hydrol Eng* 5(2):124–137. [https://doi.org/10.1061/\(ASCE\)1084-0699\(2000\)5:2\(124\)](https://doi.org/10.1061/(ASCE)1084-0699(2000)5:2(124))
- Ateeq-ur-Rauf A, Ghumman AR, Ahmad S, Hashmi HN (2018) Performance assessment of artificial neural networks and support vector regression models for stream flow predictions. *Environ Monit Assess*. <https://doi.org/10.1007/s10661-018-7012-9>
- Beck MB (1991) Forecasting environmental change. *J Forecast* 10:3–19. <https://doi.org/10.1002/for.3980100103>
- Bierkens MFP (1998) Modeling water table fluctuations by means of a stochastic differential equation. *Water Resour Res* 34:2485–2499
- Campolo M, Andreussi P, Soldati A (1999) River flood forecasting with a neural network model. *Water Resour Res* 35:1191–1197. <https://doi.org/10.1029/1998WR900086>
- Cannas B, Fanni A, See L, Sias G (2006) Data preprocessing for river flow forecasting using neural networks: Wavelet transforms and data partitioning. *Phys Chem Earth* 31:1164–1171. <https://doi.org/10.1016/j.pce.2006.03.020>
- Carcano EC, Bartolini P, Muselli M, Piroddi L (2008) Jordan recurrent neural network versus IHACRES in modelling daily streamflows. *J Hydrol* 362:291–307. <https://doi.org/10.1016/j.jhydrol.2008.08.026>
- Coulibaly P, Anctil F, Bobee B (2000) Daily reservoir inflow forecasting using artificial neural networks with stopped training approach. *J Hydrol* 230:244–257. [https://doi.org/10.1016/S0022-1694\(00\)00214-6](https://doi.org/10.1016/S0022-1694(00)00214-6)
- Cybenko G (1989) Approximation by superposition of sigmoidal functions. *Math Control Signals Syst* 2:303–314. <https://doi.org/10.1007/BF02551274>
- Dalkılıç HY, Hashimi SA (2020) Prediction of daily streamflow using artificial neural networks (ANNs), wavelet neural networks (WNNs), and adaptive neuro-fuzzy inference system (ANFIS) models. *Water Supply* 20(4):1396–1408. <https://doi.org/10.2166/ws.2020.062>
- Daubechies I (1990) The wavelet transform, time-frequency localization and signal analysis. *IEEE Trans Inf Theory* 36:961–1005. <https://doi.org/10.1109/18.57199>
- Ebrahimi H, Rajaei T (2017) Simulation of groundwater level variations using wavelet combined with neural network, linear regression and support vector machine. *Glob Planet Change* 148:181–191. <https://doi.org/10.1016/j.gloplacha.2016.11.014>
- El-Shafie A, Taha MR, Noureldin A (2007) A neuro-fuzzy model for inflow forecasting of the Nile river at Aswan high dam. *Water Resour Manage* 21:533–556. <https://doi.org/10.1007/s11269-006-9027-1>
- Freire PKdMM, Santos CAG, da Silva GBL (2019) Analysis of the use of discrete wavelet transforms coupled with ANN for short-term streamflow forecasting. *Appl Soft Comput* 80:494–505. <https://doi.org/10.1016/j.asoc.2019.04.024>
- Gong Y, Zhang Y, Lan S, Wang H (2016) A Comparative study of artificial neural networks, support vector machines and adaptive neuro fuzzy inference system for forecasting groundwater levels near Lake Okeechobee, Florida. *Water Resour Manage* 30:375–391. <https://doi.org/10.1007/s11269-015-1167-8>
- Hadi SJ, Tombul M (2018) Streamflow forecasting using four wavelet transformation combinations approaches with data-driven models: a comparative study. *Water Resour Manage* 32:4661–4679. <https://doi.org/10.1007/s11269-018-2077-3>
- Hsu KL, Gupta HV, Sorooshian S (1995) Artificial neural network modeling of the rainfall-runoff process. *Water Resour Res* 31(10):2517–2530. <https://doi.org/10.1029/95WR01955>
- Jha MK, Sahoo S (2015) Efficacy of neural network and genetic algorithm techniques in simulating spatiotemporal fluctuations of groundwater. *Hydrol Process* 29:671–691. <https://doi.org/10.1002/hyp.10166>
- Jimeno-Saez P, Senent-Aparicio J, Perez-Sanchez J, Pulido-Velazquez D (2018) A comparison of SWAT and ANN models for daily runoff simulation in different climatic zones of peninsular Spain. *Water (switzerl)* 2018:10. <https://doi.org/10.3390/w10020192>
- Kagoda PA, Ndiritu J, Ntuli C, Mwaka B (2010) Application of radial basis function neural networks to short-term streamflow forecasting. *Phys Chem Earth* 35:571–581. <https://doi.org/10.1016/j.pce.2010.07.021>
- Kang KW, Kim JH, Park CY, Ham KJ (1993) Evaluation of hydrological forecasting system based on neural network model. In:

- Proceedings of the 25th Congress of the International Association for Hydraulic Research. Delft, Netherlands, pp 257–264
- Kasiviswanathan KS, He J, Sudheer KP, Tay JH (2016) Potential application of wavelet neural network ensemble to forecast streamflow for flood management. *J Hydrol* 536:161–173. <https://doi.org/10.1016/j.jhydrol.2016.02.044>
- Kim JW, Pachepsky YA (2010) Reconstructing missing daily precipitation data using regression trees and artificial neural networks for SWAT streamflow simulation. *J Hydrol* 394:305–314. <https://doi.org/10.1016/j.jhydrol.2010.09.005>
- Kim TW, Valdes JB (2003) Nonlinear model for drought forecasting based on a conjunction of wavelet transforms and neural networks. *J Hydrol Eng* 8:319–328. [https://doi.org/10.1061/\(asce\)1084-0699\(2003\)8:6\(319\)](https://doi.org/10.1061/(asce)1084-0699(2003)8:6(319))
- Kisi O (2009) Neural networks and wavelet conjunction model for intermittent streamflow forecasting. *J Hydrol Eng* 14:773–782. [https://doi.org/10.1061/\(asce\)he.1943-5584.0000053](https://doi.org/10.1061/(asce)he.1943-5584.0000053)
- Kumar DN, Raju KS, Sathish T (2004) River flow forecasting using artificial neural networks. *Water Resour Manage* 2:143–161. <https://doi.org/10.1016/j.asoc.2019.04.024>
- Liu Z, Zhou P, Chen G, Guo L (2014) Evaluating a coupled discrete wavelet transform and support vector regression for daily and monthly streamflow forecasting. *J Hydrol* 519:2822–2831. <https://doi.org/10.1016/j.jhydrol.2014.06.050>
- Maheswaran R, Khosa R (2012) Comparative study of different wavelets for hydrologic forecasting. *Comput Geosci* 46:284–295. <https://doi.org/10.1016/j.cageo.2011.12.015>
- Mallat SG (1989) A theory for multiresolution signal decomposition: the wavelet representation. *IEEE Trans Pattern Anal Mach Intell* 11(7):674–693. <https://doi.org/10.1109/34.192463>
- Meng X, Yin M, Ning L, Liu D, Xue X (2015) A threshold artificial neural network model for improving runoff prediction in a karst watershed. *Environ Earth Sci* 74:5039–5048. <https://doi.org/10.1007/s12665-015-4562-9>
- Modaresi F, Araghinejad S, Ebrahimi K (2018) A Comparative assessment of artificial neural network, generalized regression neural network, least-square support vector regression, and K-nearest neighbor regression for monthly streamflow forecasting in linear and nonlinear conditions. *Water Resour Manage* 32:243–258. <https://doi.org/10.1007/s112690171807-2>
- Mohanty S, Jha MK, Kumar A, Sudheer KP (2010) Artificial neural network modeling for groundwater level forecasting in a river island of eastern India. *Water Resour Manage* 24:1845–1865. <https://doi.org/10.1007/s11269-009-9527-x>
- Moosavi V, Vafakhah M, Shirmohammadi B, Ranjbar M (2014) Optimization of Wavelet-ANFIS and wavelet-ANN hybrid models by Taguchi method for groundwater level forecasting. *Arab J Sci Eng* 39:1785–1796. <https://doi.org/10.1007/s13369-013-0762-3>
- Nalley D, Adamowski J, Khalil B (2012) Using discrete wavelet transforms to analyze trends in streamflow and precipitation in Quebec and Ontario (1954–2008). *J Hydrol* 475:204–228. <https://doi.org/10.1016/j.jhydrol.2012.09.049>
- Nayak PC, Satyajai Rao YR, Sudheer KP (2006) Groundwater level forecasting in a shallow aquifer using artificial neural network approach. *Water Resour Manage* 20:77–90. <https://doi.org/10.1007/s11269-006-4007-z>
- Nourani V, Baghanam AH, Rahimi AY, Nejad FH (2014a) Evaluation of wavelet-based de-noising approach in hydrological models linked to artificial neural networks. In: Islam T, Srivastava PK, Gupta M, Zhu X, Mukherjee S (eds) Computational intelligence techniques in earth and environmental sciences. Springer, Netherlands, Dordrecht, vol 9789401786, pp 209–241. [https://doi.org/10.1007/978-94-017-8642-3\\_12](https://doi.org/10.1007/978-94-017-8642-3_12)
- Nourani V, Hosseini Baghanam A, Adamowski J, Kisi O (2014b) Applications of hybrid wavelet-Artificial Intelligence models in hydrology: a review. *J Hydrol* 514:358–377. <https://doi.org/10.1016/j.jhydrol.2014.03.057>
- Nourani V, Alami MT, Vousoughi FD (2015) Wavelet-entropy data pre-processing approach for ANNbased groundwater level modeling. *J Hydrol* 524:255–269. <https://doi.org/10.1016/j.jhydrol.2015.02.048>
- Panda RK, Pramanik N, Bala B (2010) Simulation of river stage using artificial neural network and MIKE 11 hydrodynamic model. *Comput Geosci* 36:735–745. <https://doi.org/10.1016/j.cageo.2009.07.012>
- Partal T (2009) Modelling evapotranspiration using discrete wavelet transform and neural networks. *Hydrol Process* 23:3545–3555. <https://doi.org/10.1002/hyp.7448>
- Partal T, Kucuk M (2006) Long-term trend analysis using discrete wavelet components of annual precipitations measurements in Marmara region (Turkey). *Phys Chem Earth* 31:1189–1200. <https://doi.org/10.1016/j.pce.2006.04.043>
- Percival DB, Walden AT (2006) Wavelet methods for time series analysis. Cambridge UP, London
- Pramanik N, Panda RK (2009) Application of neural network and adaptive neuro-fuzzy inference systems for river flow prediction. *Hydrol Sci J* 54:247–260. <https://doi.org/10.1623/hysj.54.2.247>
- Principe JC, Euliano NR, Curt Lefebvre W (2000) Neural and adaptive systems. Wiley, Hoboken
- Raghavendra S, Deka PC (2014) Support vector machine applications in the field of hydrology: a review. *Appl Soft Comput J* 19:372–386. <https://doi.org/10.1016/j.asoc.2014.02.002>
- Rezaeianzadeh M, Stein A, Tabari H, Abghari H, Jalalkamali N, Hosseinipour EZ, Singh VP (2013) Assessment of a conceptual hydrological model and artificial neural networks for daily outflows forecasting. *Int J Environ Sci Technol* 10:1181–1192. <https://doi.org/10.1007/s13762-013-0209-0>
- Shi P, Chen C, Srinivasan R, Zhang X, Cai T, Fang X, Qu S, Chen X, Li Q (2011) Evaluating the SWAT model for hydrological modeling in the Xixian Watershed and a comparison with the XAJ model. *Water Resour Manage* 25:2595–2612. <https://doi.org/10.1007/s11269-011-9828-8>
- Solomatine DP, Ostfeld A (2008) Data-driven modelling: some past experiences and new approaches. *J Hydroinf* 10:3–22. <https://doi.org/10.2166/hydro.2008.015>
- Tapoglou E, Karatzas GP, Trichakis IC, Varouchakis EA (2014) A spatio-temporal hybrid neural network-Kriging model for groundwater level simulation. *J Hydrol* 519:3193–3203. <https://doi.org/10.1016/j.jhydrol.2014.10.040>
- Tayyab M, Zhou J, Dong X, Ahmad I, Sun N (2019) Rainfall-runoff modeling at Jinsha River basin by integrated neural network with discrete wavelet transform. *Meteorol Atmos Phys* 131:115–125. <https://doi.org/10.1007/s00703-017-0546-5>
- Tiwari MK, Chatterjee C (2011) A new wavelet-bootstrap-ANN hybrid model for daily discharge forecasting. *J Hydroinf* 13:500–519. <https://doi.org/10.2166/hydro.2010.142>
- Tokar BAS, Johnson PA (1999) Rainfall-runoff modeling using artificial neural networks. *J Hydrol Eng* 4:232–239
- Tokar AS, Markus M (2000) Precipitation-runoff modeling using artificial neural networks and conceptual models. *J Hydrol Eng* 5:156–161. [https://doi.org/10.1061/\(asce\)1084-0699\(2000\)5:2\(156\)](https://doi.org/10.1061/(asce)1084-0699(2000)5:2(156))
- Wagena MB, Goering D, Collick AS, Bock E, Fuka DR, Buda A, Easton ZM (2020) Comparison of short-term streamflow forecasting using stochastic time series, neural networks, process-based, and Bayesian models. *Environ Model Softw*. <https://doi.org/10.1016/j.envsoft.2020.104669>
- Young CC, Liu WC, Wu MC (2017) A physically based and machine learning hybrid approach for accurate rainfall-runoff modeling during extreme typhoon events. *Appl Soft Comput J* 53:205–216. <https://doi.org/10.1016/j.asoc.2016.12.052>

- Zealand CM, Burn DH, Simonovic SP (1999) Short term streamflow forecasting using artificial neural networks. *J Hydrol* 214:32–48. [https://doi.org/10.1016/S0022-1694\(98\)00242-X](https://doi.org/10.1016/S0022-1694(98)00242-X)
- Zhang G, Patuwo B, Hu MY (2001) A simulation study of artificial neural networks for nonlinear time-series forecasting. *Comput Oper Res* 28:381–396. [https://doi.org/10.1016/S0305-0548\(99\)00123-9](https://doi.org/10.1016/S0305-0548(99)00123-9)
- Zhu S, Zhou J, Ye L, Meng C (2016) Streamflow estimation by support vector machine coupled with different methods of time

series decomposition in the upper reaches of Yangtze River, China. *Environ Earth Sci* 75(531):1–12. <https://doi.org/10.1007/s12665-016-5337-7>

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.