Scientia Iranica A (2012) 19 (6), 1445–1455

Sharif University of Technology

Scientia Iranica

Transactions A: Civil Engineering

www.sciencedirect.com

Wavelet neural network model for reservoir inflow prediction

U. Okkan *

Department of Civil Engineering, Balikesir University, Balikesir, Turkey

Received 10 February 2012; revised 4 July 2012; accepted 2 September 2012

KEYWORDS

Wavelet neural network model;
Discrete wavelet transform;
Levenberg-Marquardt algorithm;
Reservoir inflow prediction.

Abstract In this study, a Wavelet Neural Network (WNN) model is proposed for monthly reservoir inflow prediction by combining the Discrete Wavelet Transform (DWT) and Levenberg-Marquardt optimization algorithm-based Feed Forward Neural Networks (FFNN). The study area covers the basin of Kemer Dam which is located in the Aegean region of Turkey. Monthly meteorological data were decomposed into wavelet sub-time series by DWT. Ineffective sub-time series have been eliminated by using all possible regression method and evaluating the Mallows' Cp coefficients to prevent collinearity. Then, effective sub-time series components have been used as the new inputs of neural networks. DWT has been also integrated with multiple linear regressions (WREG) within the study. The results of Wavelet Neural Network (WNN) model and WREG have been compared with conventional Feed Forward Neural Networks (FFNN) and multiple linear regression (REG) models. When the statistical-based criteria are examined, it has been observed that the DWT method has increased the performances of feed forward neural networks and regression methods. The results determined in the study indicate that the WNN is a successful tool to model the monthly inflow series of dam and can give good prediction performances than other methods.

1. Introduction

Analyzing hydrological data can give significant statistical information for both past and future characteristics of a basin. Especially, recording and modeling of reservoir inflow data have highly considerable roles in reservoir operation studies and water resources planning. In addition to the physical-based models and conceptual models, a basin can also be represented by black-box models which associate basin inputs and desired outputs without detailed considerations on the physical processes. In this context, conventional statistical models such as linear-nonlinear regression and stochastic models were commonly used.

Over the past years, artificial intelligence methods have been widely used as black-box models in predicting of hydrological variables. Especially, Artificial Neural Network (ANN), which is a nonlinear computing approach inspired by the learning process of the brain, has been accepted as one of the effective tools for modelling a complex hydrologic system [1–6].

Although ANN methods had been used extensively as useful tools for prediction of hydrological variables, it has also some problems to deal with non-stationary data [7,8]. Since the hydrological time series includes several frequency components and have nonlinear relationships, hybrid model approaches which include different data-preprocessing and combine techniques have been used to raise the prediction performance of neural networks.

Chaotic neural networks [9], Set Pair Analysis (SPA) and principle component analysis-based (PCA) neural networks [10–12], threshold neural networks [10], and cluster-based hybrid neural networks [13] were successfully applied to hydrological variables.

In the last years, wavelet transform, which is an alternative data-preprocessing technique, showed excellent performance in hydrological modeling due to its ability to analyze a signal in both time and frequency. This approach overcomes the basic drawbacks of conventional Fourier transform. Nourani et al. (2009, 2011) showed that the wavelet transform provided effective decompositions of time series so that decomposed data increased the performance of hydrological prediction model by capturing useful information on different resolution levels. Hence a wavelet neural network model which uses

* Tel.: +90 0266 612 11 94/95.
E-mail address: umutokkan@balikesir.edu.tr.

Peer review under responsibility of Sharif University of Technology.

Production and hosting by Elsevier

1026-3098 © 2012 Sharif University of Technology. Production and hosting by Elsevier B.V. Open access under CC BY-NC-ND license.
doi:10.1016/j.scient.2012.10.009

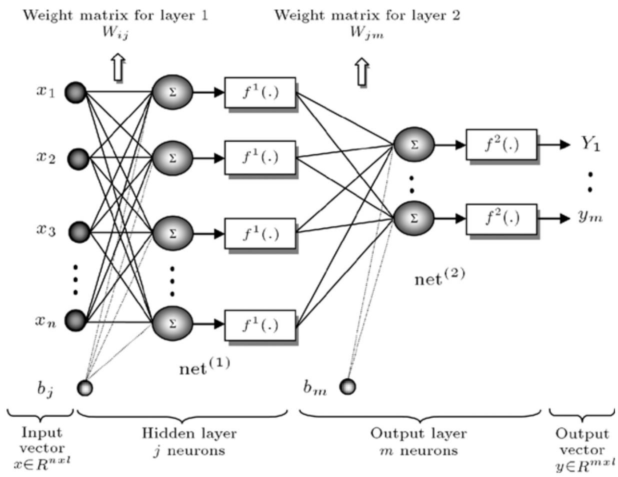

Figure 1: FFNN structure.

multi-scale signals as input data can present more suitable prediction performance rather than a single pattern input [14,15]. There are some successful applications of wavelet neural network models which are prepared by the combined use of wavelet transform and neural networks [7,8,14,15]. The wavelet transform is also integrated with multiple linear regression [16–18] and support vector machine approach [19].

The main purpose of the study presented is to examine the applicability of a wavelet neural network model for the prediction of the monthly reservoir inflow values of a study region which is an important water resource for the Menderes Basin/Turkey, and to compare its performance with single feed forward neural networks, multiple linear regression (REG) and the wavelet-MLR combined model (WREG)

Firstly, meteorological data (monthly areal precipitation and monthly areal temperature data) have been decomposed into wavelet sub-time series. Later, effective sub-time series have been determined by using all possible regression method [20] and the evaluation of Mallows' C_p coefficients to prevent collinearity. These new time series have constituted the inputs of neural network and multiple linear regression models. Some favorite performance evaluation measures are employed to assess developed models.

2. Materials and methods

2.1. Feed forward neural networks

Depending on the techniques to train a Feed Forward Neural Network (FFNN) model (Figure 1), different back propagation algorithms were developed. In this study, the Levenberg–Marquardt back propagation algorithm, which is a simplified version of Newton method, has been used in training of the FFNN. This algorithm is a second-order nonlinear optimization technique that is usually faster and more reliable than any other back propagation techniques [21,22].

The training process can be viewed as finding a set of weights that minimize the error (e_p) for all samples in the training set (T). The performance function is a sum of squares of the errors as follows [22]:

$$E(W) = \frac{1}{2}\sum_{p=1}^{P}(d_p - y_p)^2 = \frac{1}{2}\sum_{p=1}^{P}(e_p)^2, \quad P = mT, \tag{1}$$

where T is the total number of training samples, m is the number of output layer neurons, W represents the vector containing all the weights in the network, y_p is the actual network output, and d_p is the desired output.

When training with the Levenberg–Marquardt algorithm, the changing of weights ΔW can be computed as follows [22]:

$$\Delta W_k = -[J_k^T J_k + \mu_k I]^{-1} J_k^T e_k. \tag{2}$$

Then, the update of the weights can be adjusted as follows:

$$W_{k+1} = W_k + \Delta W_k, \tag{3}$$

where J is the Jacobian matrix, I is the identify matrix, e is the network error, μ is the Marquardt parameter which is to be updated using the decay rate β depending on the outcome. In particular, μ is multiplied by the decay rate $\beta (0 < \beta < 1)$ whenever $E(W)$ decreases, while μ is divided by β whenever $E(W)$ increases in a new step [22].

2.2. Wavelet transform

2.2.1. Continuous wavelet transform

The wavelet transform, developed during the last years, is an effective decomposition method. This method provides an analyzing way of a signal in both time and frequency and appears to be more successful than the conventional Fourier transforms that do not provide time-frequency analysis for the variables involving non-stationary signals [23].

The time-scale wavelet transform of a continuous time signal, $x(t)$, is defined as [8,24,25]:

$$\mathrm{W}(s,\ \tau) = |s|^{-1/2}\int_{-\infty}^{\infty} \psi^*\left(\frac{t-\tau}{s}\right) x(t)\, dt$$
$$\tau \in R,\ s \in R,\ s \neq 0, \tag{4}$$

where $*$ corresponds to the complex conjugate and $\psi\ (t)$ is wavelet function or mother wavelet; s is the scale or frequency factor, τ is the time factor, R is the domain of real number.

Eq. (4) describes that wavelet transform is the decomposition of $x(t)$ under different resolution scale [8,24,25]. The original series can be reconstructed using the inversion of transform.

In this study, the Haar mother wavelet (simple wavelet) has been used because it is conceptually simple, fast and memory efficient [25].

The Haar mother wavelet is defined as [25]:

$$\psi(t) = \begin{cases} 1 & 0 \le t \le 1/2 \\ -1 & 1/2 \le t \le 1 \\ 0 & \text{otherwise.} \end{cases} \tag{5}$$

2.2.2. Discrete wavelet transform

For practical applications in hydrology, researchers have access to a discrete time signal rather than to a continuous time signal [26]. A discretization of Eq. (4) based on trapezoidal rule is perhaps the simplest discretization of the Continuous Wavelet Transform (CWT). This transform produces N^2 coefficients from a data set of length N; hence unnecessary information is locked up within the coefficients, which may or may not be a desirable property [24,26,27]. To overcome this difficulty, Discrete Wavelet Transforms (DWT) which present power of two logarithmic scaling of the translations can be used in practical applications. Because it reduces the computational complexity of the CWT and the redundancy of the CWT, there is an advantage for preferring the DWT over the CWT [27].

In this study, the Mallat algorithm was used for the DWT of monthly time-series. According to the Mallat algorithm, the discrete wavelet transform of discrete time series x_t is defined as [27]:

$$W_{j,k} = 2^{-j/2} \sum_{t=0}^{N-1} \psi(2^{-j}t - k)x_t, \tag{6}$$

where t is integer time steps, j and k are integers that control, respectively, the scale and time; $W_{j,k}$ is the wavelet coefficient for the scale factor, $s = 2^j$, and the time factor, $\tau = 2^j k$.

In DWT method, the time series (x_i) passes through two filters and are decomposed into wavelet sub-time series components, which can be computed by using Eq. (6), without losing the information about the instant of the element occurrence. The DWT converts a signal into father and mother wavelets. Father wavelets represent the high-scale, low frequency components (approximation (A) components). Mother wavelets are representations of the low-scale and high frequency components (detail (D) components). Thus, DWT allows one to study different investing behaviors in different time scales independently [17,26].

2.3. Study area and data

To assess the usefulness of wavelet neural network model, a similar study area and data which were considered by Okkan and Dalkilic [28] have been used in this study. This application area covers the drainage basin of Kemer Dam, which is located in the Aegean region of Turkey (Figure 2). The basin is fed by four rivers and the streamflow values are observed by four streamflow gauging stations (Calikoy/EIE-730, Yemisendere/EIE-731, Degirmenalani/ EIE-732, and Goktepe/EIE-733) located at the upstream of the dam (Figure 3). These data were collected from the records of two institutes of Turkey: XXI. Regional Directorate of State Hydraulic Works, and Operational Directorate of Kemer Dam Power Plant which is a part of the Electrical Works Authority. Thus, the collected reservoir inflow data were prepared for the period between January 1980 and December 2005. In addition to inflow data, the monthly data of precipitation and temperature at Denizli and Mugla meteorological stations were obtained from the State Meteorological Organization of Turkey. Next, Thiessen weighted areal precipitation values and arithmetical mean temperature (areal temperature) values were prepared for monthly time scale, using records available at both stations.

In modeling studies, the selection of appropriate input variables plays an important role. In this study, modeling strategy that predicts inflow data from inputs based on monthly areal precipitation and areal temperature data. In addition to the concurrent values of monthly areal precipitation data, areal precipitation values at various lags have been also considered. A statistical approach has been employed by Sudheer et al. (2002) to determine the appropriate order of precipitation lag. The approach is based on the heuristic that the potential influencing variables corresponding to different time lags can be identified through statistical analysis of the data series that uses cross correlation between the variables [29]. According to the cross correlation results, the Cross Correlation Function (CCF) between the reservoir inflow and areal precipitation values at various lags have showed significant correlation at concurrent areal precipitation (P_t) and one month of areal precipitation lag (P_{t-1}) on the inflow. Thus, three input data

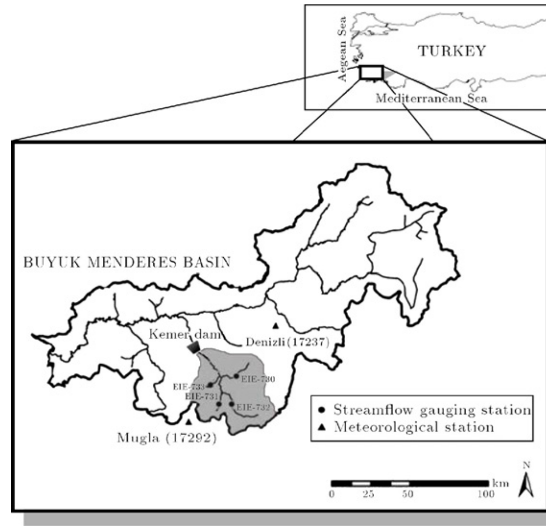

Figure 2: Kemer Dam, the streamflow gauging stations and meteorological stations within the study area.

(P_t, P_{t-1} and T_t), have been prepared for the same periods of the monthly inflow records (P_t: monthly areal precipitation; P_{t-1}: one-month-ahead areal precipitation; T_t: monthly areal temperature).

2.4. Modeling strategy developed in the study

In the study, Discrete Wavelet Transform (DWT) has been linked to feed forward neural network and multiple linear regression models for monthly reservoir inflow prediction. First, the input data (P_t, P_{t-1}, T_t) of training and testing periods have been decomposed into a certain number of sub-time series components by DWT. The selection of the optimal decomposition level is one of the keys to determine the performance of the model in wavelet domain. Decomposition level is generally based on signal characteristics and experiences to selection. For example, Chou and Wang (2004) showed that using only one decomposition level to model the streamflow time series does not easily represent the process [30]. Kisi and Cimen (2011) used three decomposition levels in their monthly streamflow forecasting study [19].

After various trials for this study, the best results have been obtained by three decomposition levels. Thus, meteorological input data have been decomposed using the Haar wavelet function and twelve sub-time series components (time series of 2-month mode (D1), 4-month mode (D2), 8-month mode (D3) and approximation mode (A3)) have been obtained for the training and testing period. The three levels decomposition of the P_t, P_{t-1} and T_t signals that yield four sub-signals by the Haar wavelet are shown in Figure 3, Figure 4 and Figure 5.

After the decomposition process, effective sub-time series components should be determined since some correlated sub-time series components may reduce the generalization capabilities of the models. In this study, this was carried out using Mallows' C_p based on all possible regression method as this is an effective way to determine the subset of variables in cases where there are a large number of potential predictor variables [20].

Mallows' C_p is a measure of the error in the best subset model, relative to the error incorporating all variables. Adequate models are those for which C_p is roughly equal to the

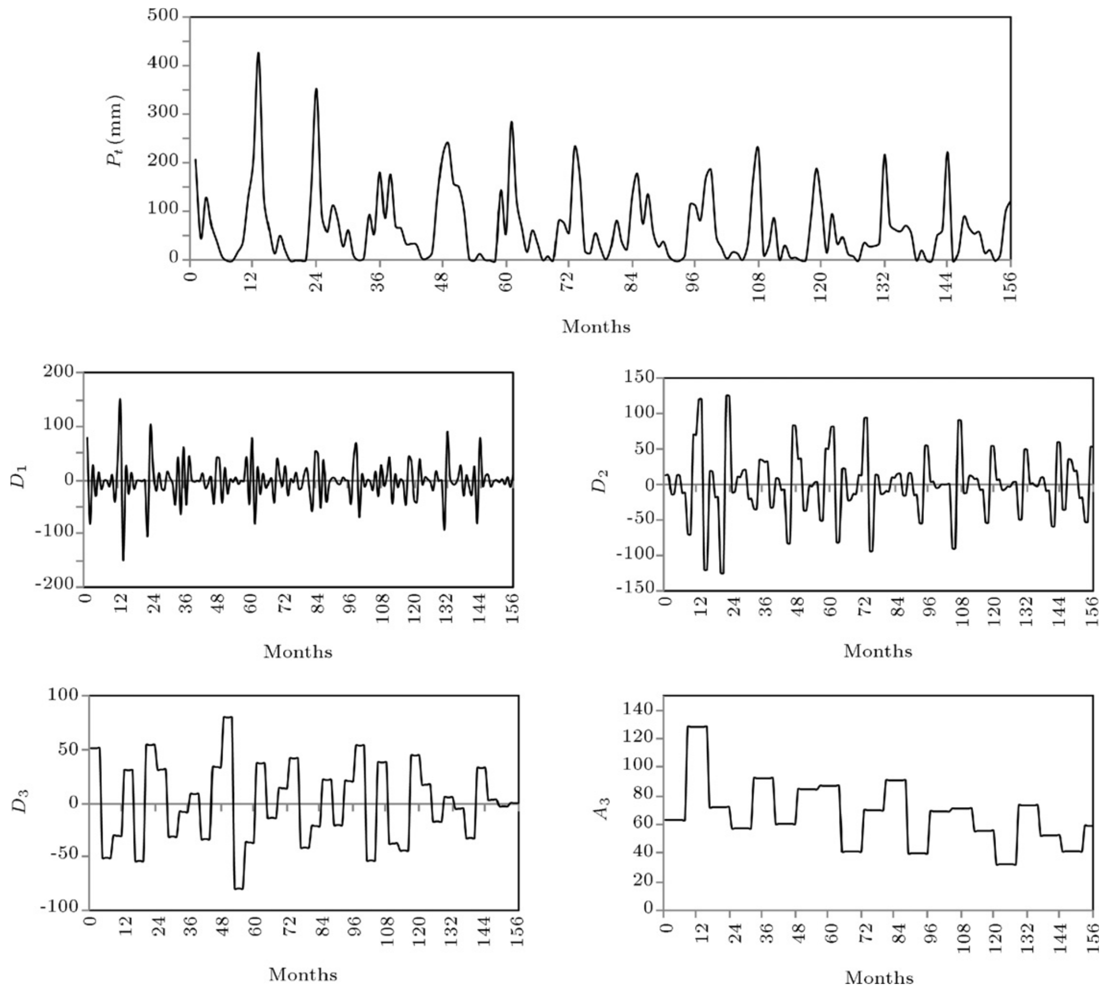

Figure 3: Original time series, 2-month mode of time series (D1), 4-month mode of time series (D2), 8-month mode of time series (D3) and approximation mode of time series (A3) of monthly areal precipitation for training period.

number of variables in the model. The C_p values can be computed as [20]:

$$C_p = (N - k)\frac{MSE_i}{MSE_F} - (N - 2i - 1), \tag{7}$$

where N is the number of data, MSE_i is the mean of residual squares in the model with $i$variable, MSE_F is the mean of residual squares in the full model with k variable.

These new time series have constituted the inputs of models. In some studies, new series which are obtained by summing the effective sub-time series were used as input to the models [18,19]. Unlike these studies, components of sub-time series determined with all the possible regression method have been used as individual separate model inputs in this study. The modeling strategy developed in this study is summarized in Figure 6.

2.5. Data normalization and evaluation of model performances

The input and output data are normalized to prevent the model from being dominated by the variables with large values, as is commonly used in artificial intelligence models. In this study, the normalization processes of all data have been carried out using Eq. (8).

$$z_i = \frac{x_i - x_{\min}}{x_{\max} - x_{\min}}. \tag{8}$$

Some favorite measures are suggested for hydrological time series prediction accuracy evaluation according to literature related to the training and testing of models. In this study, three popular performance measures have been considered.

All prepared models with optimum structures of them have provided the best training result in terms of the minimum Mean Squared Error (MSE) or the minimum Root Mean Squared Error (RMSE), and the maximum determination coefficient (R^2) have been also employed for the testing period (Eqs. (9) and (10)).

$$\mathrm{MSE} = \frac{1}{N}\sum_{t=1}^{N}(Q_{g,t} - Q_{m,t})^2, \qquad \mathrm{RMSE} = \sqrt{\mathrm{MSE}}, \tag{9}$$

$$R^2 = \left[\frac{\sum_{t=1}^{N}(Q_{g,t} - \overline{Q_g})(Q_{m,t} - \overline{Q_m})}{\sqrt{\sum_{t=1}^{N}(Q_{g,t} - \overline{Q_g})^2}\sqrt{\sum_{t=1}^{N}(Q_{m,t} - \overline{Q_m})^2}}\right]^2, \tag{10}$$

where N is the number of training or testing samples, $Q_{m,t}$ is the model output, $Q_{g,t}$ is the observed runoff data in the tth time

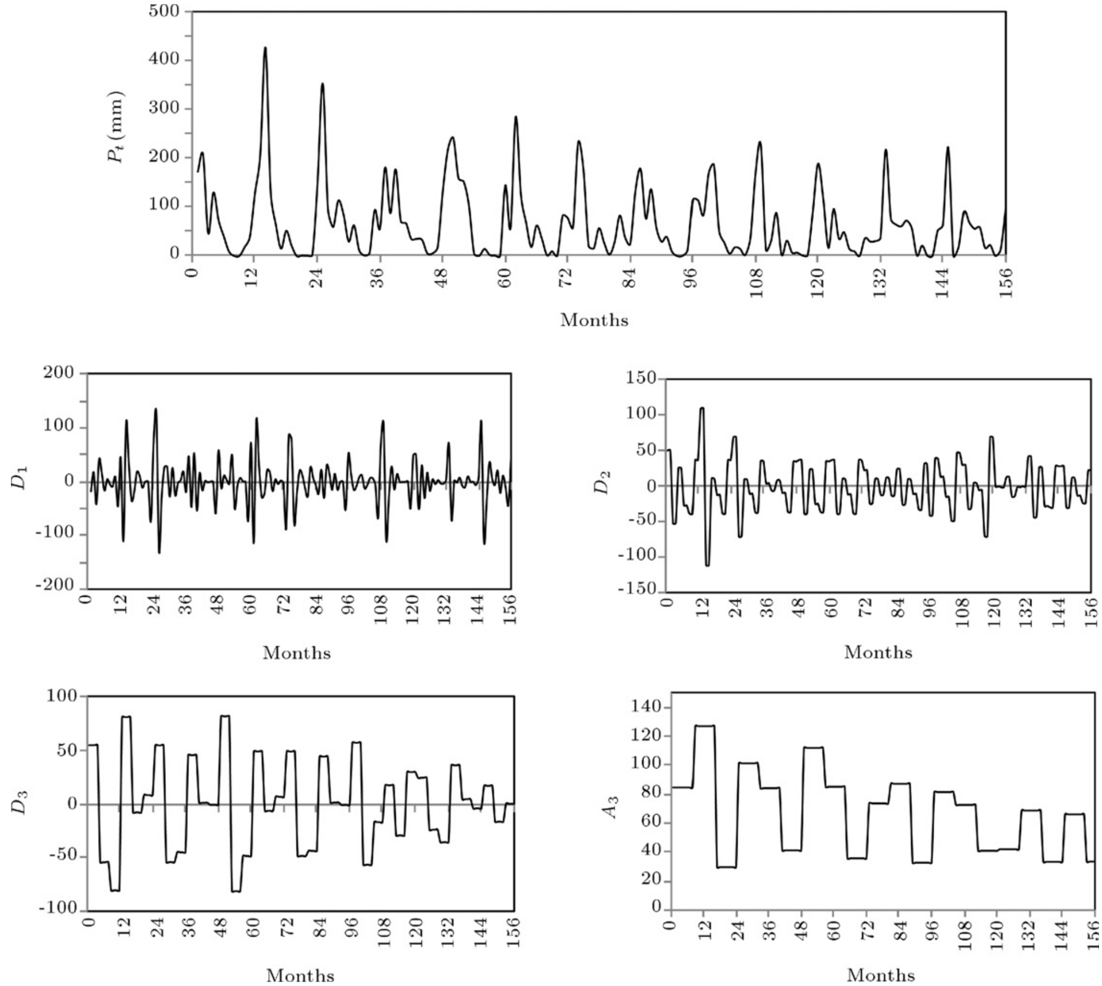

Figure 4: Original time series, 2-month mode of time series (D1), 4-month mode of time series (D2), 8-month mode of time series (D3) and approximation mode of time series (A3) of one-month-ahead areal precipitation for training period.

period, $\overline{Q_g}$ and $\overline{Q_m}$ are the mean over the observed and modeled series, respectively.

In addition to these favorite statistical measures, homogeneities of the model predictions have been also examined with Mann–Whitney U (M–W) test to present more evidence on application and success of the models. This non-parametric statistical test is used to analyze two comparing groups to identify whether they have the same distribution or not [31]. M–W is based on bringing together and arranging of two groups (e.g., predicted and observed values). When the lining up of these group members is done, for each member, a line number is assigned. The membership status of these members (to which group they belong) is ignored. Then, all these line numbers are summed up. The sum of the members of the first group is R_1 and sum of the members of the second group is R_2. Then, the U values can be calculated as:

$$U_i = N_1 N_2 + \frac{N_i(N_i+1)}{2} - R_i, \quad (i = 1, 2). \tag{11}$$

After the calculation for $i = 1$ and $i = 2$, U_1 and U_2 are obtained, and the bigger is chosen (U^*).

$$z = \frac{U^* - \frac{N_1 N_2}{2}}{\sqrt{\frac{N_1 N_2 (N_1+N_2+1)}{12}}}, \tag{12}$$

where N_1 and N_2 are the numbers of data for the groups compared.

The z value is compared with 0.05 significance level ($z_{cr} = 1.96$). For these values $z < 1.96$, it means that there is no significant difference between the measured data and model predictions. The asymptotic significance of z test statistics has been also used in this study.

3. Results

In the study, a MATLAB code which involves Mallat's DWT algorithm has been used. Levenberg–Marquardt algorithm-based FFNN model has been also prepared by a MATLAB code. To evaluate the generalization capability of all models, the data has been divided into two subsets; a training dataset to construct the models, and a testing dataset to estimate the performances of models.

Researchers have used different data division between testing and training datasets and it generally varies with problem types. There is no certain rule for data division between training and testing. In this study, adopted data division between training and testing is determined as 50% (i.e. 50% of the available data has been used for training). This data division has given the best results for the different performance measures (e.g., MSE, RMSE, R^2, the z values of M–W test).

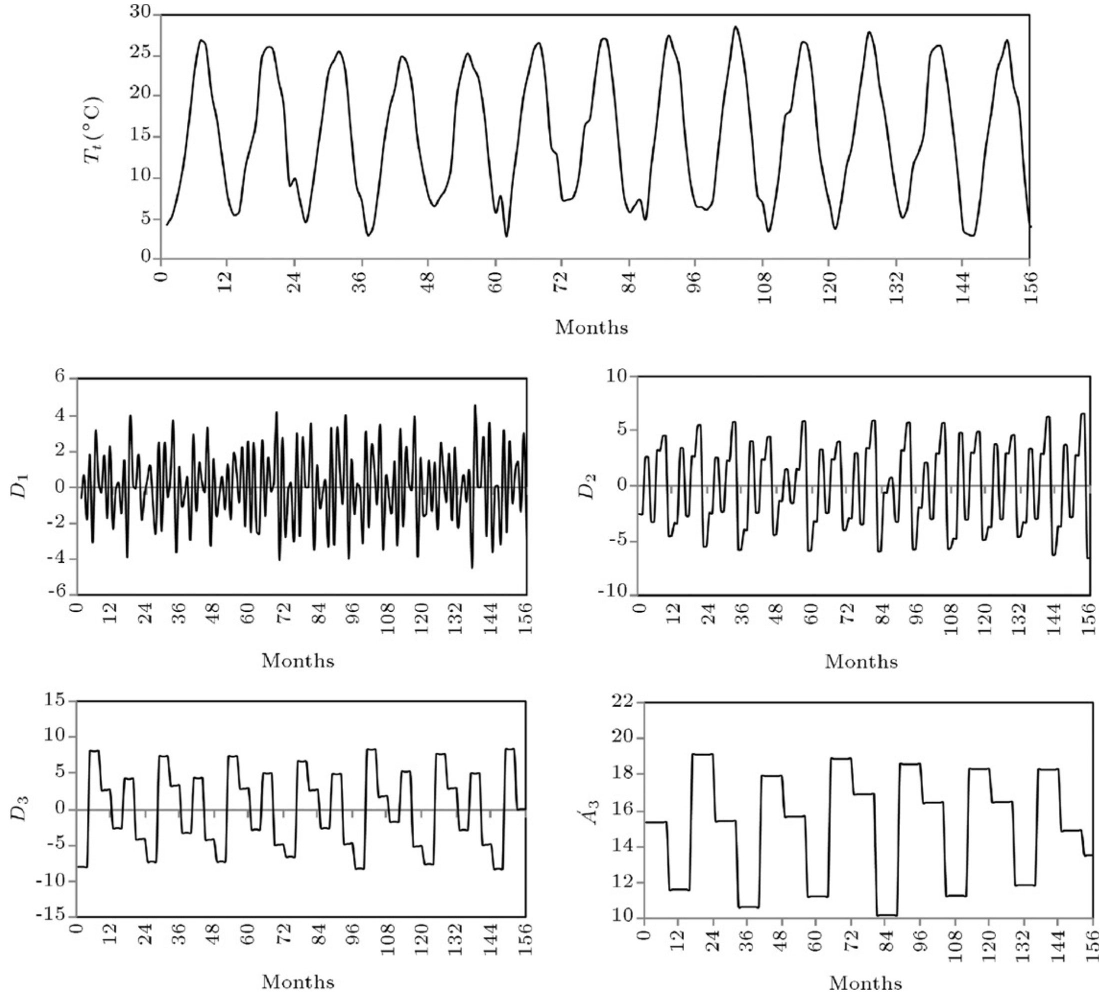

Figure 5: Original time series, 2-month mode of time series (D1), 4-month mode of time series (D2), 8-month mode of time series (D3) and approximation mode of time series (A3) of monthly areal temperature for training period.

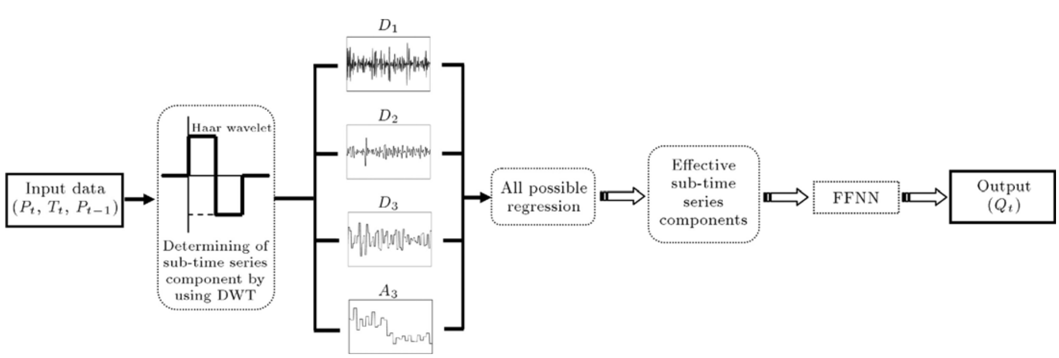

Figure 6: Structure of wavelet neural network model.

Thus, 26 years (January 1980–December 2005) input–output data have been divided into training and testing periods by proportions of 1/2 (January 1980–December 1992) and 1/2 (January 1993–December 2005), respectively. According to modeling strategy, input data of training and testing periods have been decomposed into a certain number of sub-time series components by DWT firstly. Afterward, the effective variables have been selected by the evaluation of Mallows C_p coefficients. To assess the strength and direction of the relations between twelve sub-time series components which were derived from the meteorological data, different linear regression analysis combinations in training period have been obtained using the "best subsets regression" tool in Minitab software. Although all possible regression combinations are obtained by this tool, only the best combinations are presented in Table 1.

Performances of the model with the optimum combination of the input variables (seen as bold characters in Table 1) are

Table 1: Summary of all possible regression analyses.

Input numbers	R^2 (%)	Adj. R^2 (%)	MSE (m^6/s^2)	C_p	$D1\_P_t$	$D2\_P_t$	$D3\_P_t$	$A3\_P_t$	$D1\_P_{t-1}$	$D2\_P_{t-1}$	$D3\_P_{t-1}$	$A3\_P_{t-1}$	$D1\_T_t$	$D2\_T_t$	$D3\_T_t$	$A3\_T_t$
1	28.5	28	412.3	528.1							•					
1	22.7	22.2	445.6	583.1			•									
1	17.5	17	475.7	632.8								•				
2	46	45.3	311.4	363.5							•	•				
2	44.9	44.2	318.0	374.3		•					•					
2	44.6	43.8	319.9	377.4				•			•					
3	62.4	61.7	217.1	209.6		•					•	•				
3	60.9	60.2	225.5	223.5		•		•			•					
3	59.2	58.4	235.5	240.1	•						•	•				
4	75.6	74.9	141.2	86.3	•	•					•	•				
4	74.1	73.4	149.6	100.2	•	•		•			•					
4	71.2	70.5	166.3	127.8	•					•	•	•				
5	77.5	76.7	130.2	70.1	•	•		•			•	•				
5	76.9	76.1	133.7	75.8	•	•					•	•				•
5	76.9	76.1	133.8	76	•	•	•				•	•				
6	80.5	79.8	112.7	43.1	•	•		•			•	•				•
6	78.8	77.9	122.8	59.9	•	•	•	•			•	•				
6	78.3	77.4	125.8	64.8	•	•		•		•	•	•				
7	81.8	81	105.3	32.9	•	•	•	•			•	•				•
7	81.3	80.4	108.2	37.7	•	•		•		•	•	•				•
7	81.1	80.2	109.4	39.8	•	•		•	•		•	•				•
8	82.6	81.7	100.8	27.5	•	•	•	•		•	•	•				•
8	82.5	81.6	101.3	28.3	•	•	•	•			•	•			•	•
8	82.4	81.4	102.1	29.6	•	•	•	•	•		•	•				•
9	83.6	82.6	95.3	20.4	•	•	•	•		•	•	•		•		•
9	83.3	82.3	96.8	22.9	•	•	•	•		•	•	•			•	•
9	83.2	82.1	97.6	24.2	•	•	•	•	•	•	•	•				•
10	84.3	83.2	91.2	15.7	•	•	•	•		•	•	•		•	•	•
10	84.1	83	92.0	17	•	•	•	•	•	•	•	•		•		•
10	83.9	82.7	93.6	19.6	•	•	•	•	•	•	•	•			•	•
11	**84.8**	**83.7**	**88.0**	**12.4**	•	•	•	•	•	•	•	•		•	•	•
11	84.3	83.1	91.2	17.6	•	•	•	•		•	•	•	•	•	•	•
11	84.3	83.1	91.2	17.6	•	•	•	•	•	•	•	•	•	•		•
12	85	83.7	87.2	13	•	•	•	•	•	•	•	•	•	•	•	•

(P_t: monthly areal precipitation; P_{t-1}: one-month-ahead areal precipitation; T_t: monthly areal temperature).

nearly the same as that of the full linear model with twelve variables (for $k = 12$); that is, the explained variance of the monthly reservoir inflow values, which have the minimum C_p values, are nearly equal to that explained by the full linear model. According to these results, 2-month mode of time series of monthly temperature ($D1\_T_t$) has been eliminated.

Before presenting these determined effective sub-time series as input data to FFNN, all data have been normalized using Eq. (8) to prevent the model from being dominated by the variables with the extreme values.

In the Wavelet Neural Network model (WNN) and single FFNN applications, the optimal number of neuron in the hidden layer has been determined using a trial and error approach by varying the number of neurons from 2 to 20. Two neurons in the hidden layer of FFNN and six neurons in the hidden layer of WNN have given the best results.

In this study, three widely used transfer functions, namely tangent sigmoid, linear, and log-sigmoid are evaluated in FFNN and WNN structure trials. The best results have been achieved by log-sigmoid function for each layer. The training epochs of FFNN and WNN have been set to 25 and 15, respectively. The initial values of μ and β are selected as 10^{-3} and 0.1, respectively. The results of FFNN and WNN models have been compared with multiple linear regressions (REG) and DWT-based multiple linear regression (WREG) models.

The best fitted regression coefficients of REG and WREG models have provided the best results in terms of the minimum mean squared error for training period (January 1980–December 1992). These regression equations of REG and WREG models are presented as:

$$Q_t = -17.18 + 0.25P_t + 0.13P_{t-1} + 0.70T_t, \tag{13}$$

$$\begin{aligned} Q_t = {} & -53.18 + 0.25\ D1\_P_t + 0.19\ D2\_P_t + 0.17\ D3\_P_t \\ & + 0.32\ A3\_P_t + \cdots + 0.04\ D1\_P_{t-1} + 0.13\ D2\_P_{t-1} \\ & + 0.27\ D3\_P_{t-1} + 0.30\ A3\_P_{t-1} + \cdots \\ & + 0.96\ D2\_T_t + 0.69\ D3\_T_t + 1.96\ A3\_T_t, \end{aligned} \tag{14}$$

Table 2: The performances of model applications in the training and testing periods.

Models	R^2		MSE (m^6/s^2)		RMSE (m^3/s)	
	Training	Testing	Training	Testing	Training	Testing
REG	0.745	0.637	147.954	177.579	12.164	13.326
WREG	0.838	0.704	87.999	105.273	9.899	10.260
FFNN	0.841	0.68	92.761	94.98	9.631	9.746
WNN	**0.948**	**0.791**	**30.207**	**60.327**	**5.496**	**7.767**

Table 2 presents the results of the modeling studies, in terms of the performance measures. It can be seen from Table 2 that WNN have good performances during both training and testing, and they outperform FFNN, WREG and REG models in terms of the performance measures. In the training period, the WNN model with the Haar wavelet obtained the best MSE, RMSE and R^2 statistics of 30.207 m^6/s^2, 5.496 m^3/s, and 0.948, respectively. Investigating the results during testing period, it can be seen that the WNN model outperforms FFNN, WREG and REG models.

The scatter plot and hydrograph of all the models developed in this study during the testing period are shown in Figures 7

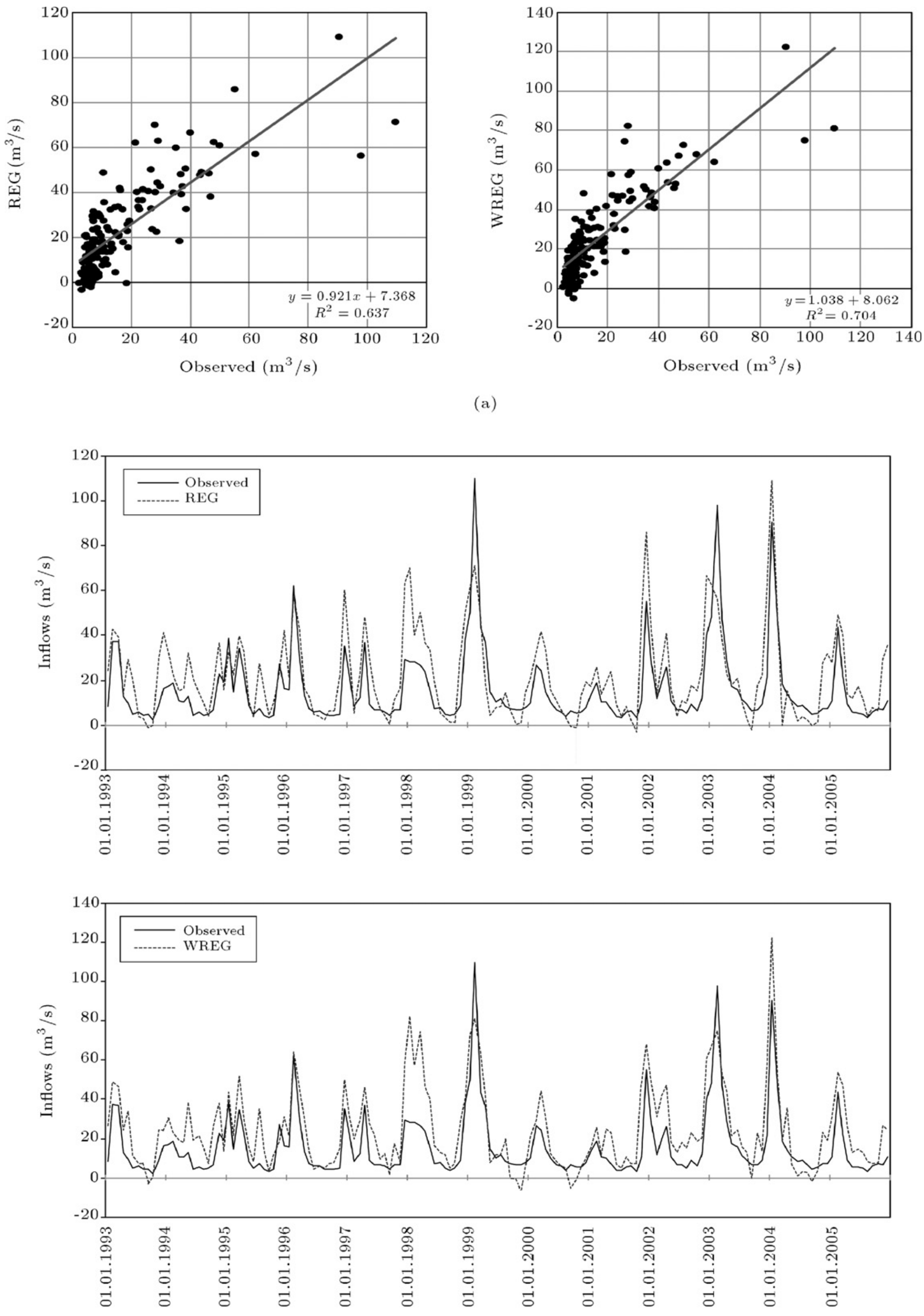

Figure 7: The scatter plots (a) and hydrographs (b) of REG and WREG models for the testing period.

and 8. When the testing period scatter graphs of the models are examined, it is observed that the standard deviations around the $y = x$ line are far less in the WNN models. In other words, when $y = ax + b$ fitted lines in graph are examined, it is observed that, in WNN model, "a" gets closer to the value 1, and "b" gets closer to the value 0, compared to the FFNN model. In addition to these, WNN model has proved itself to be precise in predicting especially the peak and low reservoir inflow values.

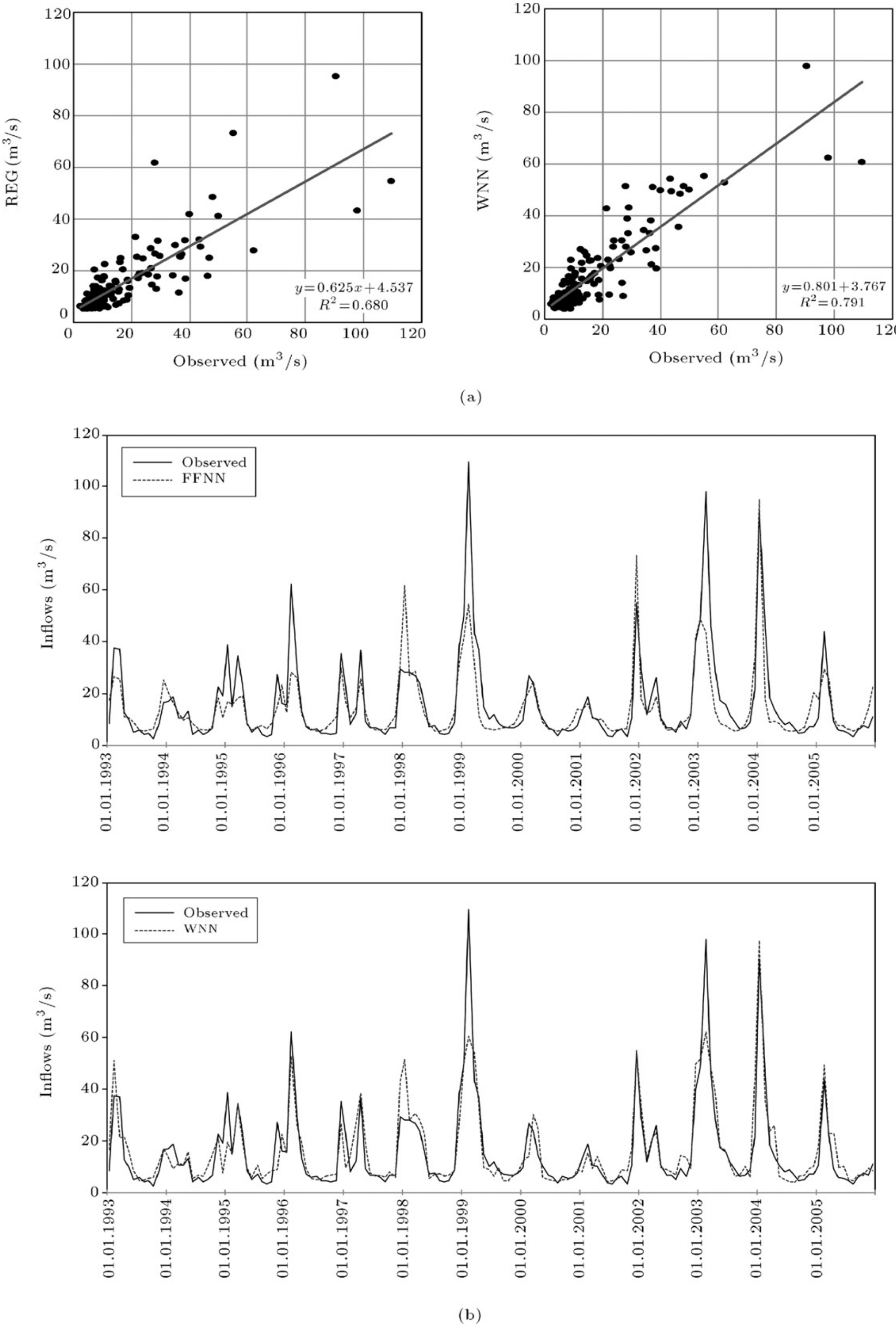

Figure 8: The scatter plots (a) and hydrographs (b) of FFNN and WNN models for the testing period.

Moreover, WREG model estimates monthly reservoir inflow values more accurate than REG model but the major weakness of these two models is related to the prediction of some small negative values for the monthly reservoir inflow which are physically meaningless. The poor operation of the least square method in calibration of the coefficients of regression models may lead to this drawback. However this drawback is resolved in the neural network approach which is artificially

Table 3: M–W statistics of all models for training and testing periods.

M–W test statistics	REG		WREG		FFNN		WNN	
	Training	Testing	Training	Testing	Training	Testing	Training	Testing
Mann–Whitney U	10906	8380	11895	9852	11857	11294	12046	11514
z	1.584	4.754	0.343	2.907	0.390	1.097	**0.153**	**0.821**
Asymptotic.		0.000		0.004		0.273		**0.412**
Sig. (two-tailed)	0.113		0.732		0.696		**0.878**	

trained by the observed values in a non-linear framework [32]. The general advantages of the neural networks over linear regression analysis were presented by Nourani and Fard [32].

When M–W test statistics were examined, it was shown that both FFNN and WNN predictions have homogeneities for training and testing set. When z statistics and asymptotic significance values are taken as a basis, it can be seen that the WNN results are better than the results of the other models (Table 3).

4. Summary and conclusion

The study presented the application of WNN and WREG models compared with FFNN and REG models used undecomposed data, for prediction of monthly reservoir inflow of Kemer Dam, and the following evaluations can be made.

- In this study, three decomposition levels have been considered for each meteorological variable (P_t, P_{t-1}, T_t), and twelve sub-time series components have been obtained for the training and testing periods. In the study presented, the Haar mother wavelet (simple wavelet) has been used with three decomposition levels. For the future studies, Daubechies's wavelets and some irregular wavelets such as Bior1.1, Rboi1.1, Coif1, Sym3, and Meyer wavelets may be used in modeling applications.
- The results of all possible regression analyses showed that highly correlated sub-time series components may reduce the generalization capability of a model. Therefore, the ineffective sub-time series components have been eliminated by all possible regression method. Thus, effective sub-time series components that represented the inputs of the neural networks and multiple regression models are treated so as to reduce collinearity as much as possible. The effective decomposed series can also be selected by other statistical approaches (e.g., stepwise regression). However, Mallows' $C_p$based on all possible regression analysis is quite an effective way to determine the subset of decomposed series in cases where there are a large number of potential predictor variables.
- The results determined in the study indicate that the DWT-based methods are successful tools to predict the monthly inflow series of Kemer Dam and can give better prediction performances than conventional models. Furthermore, modeling using wavelet transform is an effective way when it is impossible to use physical-based and conceptual models.
- When the performances of the training and testing periods are compared, it is observed that the WNN model has better results in terms of R^2, MSE, RMSE and M–W performances. Thus, it is proved with this study that wavelet neural network approach is one of the successful hybrid modeling approaches which are capable of reservoir inflow prediction.
- Although FFNN and WNN models have abilities to model complex and nonlinear relations, the structures of them may be hard to determine and they can be determined using trial-and-error approaches. Therefore, WREG models, which may be much easier to interpret, are used as an alternative way to classical neural network approaches for the monthly reservoir inflow prediction studies.
- The DWT may be used for the construction of the other machine learning models (e.g., support vector machines, least squares support vector machines, relevance vector machines and multivariate adaptive regression spline approaches) for the future studies.
- The combined use of wavelet transform and a neural network model is able to simulate nonlinear relations in this arid region which has typical Mediterranean climate characteristics. For the future studies, the proposed model structure may be used to predict reservoir inflow values of other dam basins which have different climatologic regimes.

Acknowledgments

Authors feel responsible to thank English Instructor R.C. Guren (from Celal Bayar University, M.Sc.), Dr. H.Y. Dalkilic (from Dokuz Eylul University, Ph.D.) and two anonymous reviewers for their valuable contribution to grammar correction of this study as well as workers of II. Regional Directorate of State Hydraulic Works and Turkish State Meteorological Service for their help with data collection. Comments from two anonymous reviewers of journal, which have helped to improve the paper, are also acknowledged.

References

[1] Hsu, K., Gupta, H.V. and Sorooshian, S. "Artificial neural network modeling of the rainfall runoff process", *Water Resources Research*, 31, pp. 2517–2530 (1995).
[2] Coulibaly, P., Anctil, F. and Bobee, B. "Daily reservoir inflow forecasting using artificial neural networks with stopped training approach", *Journal of Hydrology*, 230, pp. 244–257 (2000).
[3] Cigizoglu, H.K. "Generalized regression neural networks in monthly flow forecasting", *Civil Engineering and Environmental Systems*, 22(2), pp. 71–84 (2005).
[4] Kisi, O. "Multi-layer perceptrons with Levenberg–Marquardt training algorithm for suspended sediment concentration prediction and estimation", *Hydrological Sciences Journal*, 49(6), pp. 1025–1040 (2004).
[5] Kisi, O. "River flow forecasting and estimation using different artificial neural network techniques", *Hydrology Research*, 39(1), pp. 27–40 (2008).
[6] Derakhshan, H. and Talebbeydokhti, N. "Rainfall disaggregation in non-recording gauge stations using space–time information system", *Scientia Iranica*, 18(5), pp. 995–1001 (2011).
[7] Cannas, B., Fanni, A., See, L. and Sias, G. "Data processing for river flow forecasting using neural networks: wavelet transforms and data partitioning", *Physics and Chemistry of the Earth*, 31, pp. 1164–1171 (2006).
[8] Wang, W. and Ding, J. "Wavelet network model and its application to the prediction of hydrology", *Nature and Science*, 1, pp. 67–71 (2003).
[9] Karunasinghe, D.S.K. and Liong, S.Y. "Chaotic time series prediction with a global model: artificial neural network", *Journal of Hydrology*, 323, pp. 92–105 (2006).
[10] Wang, W., Van Gelder, P., Vrijling, J.K. and Ma, J. "Forecasting daily streamflow using hybrid ANN models", *Journal of Hydrology*, 324, pp. 383–399 (2006).

[11] Wang, H.F., Huang, W.J. and Wang, W.S. "Cuntan station of the Yangtze River annual runoff forecasting with set pair analysis method", *Journal of Heilongjiang Hydraulic Engineering College*, 33(4), pp. 3–5 (2006).
[12] Wu, C.L., Chau, K.W. and Li, Y.S. "Predicting monthly streamflow using data-driven models coupled with data-preprocessing techniques", *Water Resources Research*, 45(8), pp. 1–23 (2008).
[13] Cigizoglu, H.K. and Kisi, O. "Flow prediction by three back propagation techniques using *k*-fold partitioning of neural network training data", *Nordic Hydrology*, 36(1), pp. 49–64 (2005).
[14] Nourani, V., Kisi, O. and Komasi, M. "Two hybrid Artificial Intelligence approaches for modeling rainfall–runoff process", *Journal of Hydrology*, 402, pp. 41–59 (2011).
[15] Nourani, V., Alami, M.T. and Aminfar, M.H. "A combined neural-wavelet model for prediction of watershed precipitation, Lighvanchai, Iran", *Engineering Applications of Artificial Intelligence*, 16, pp. 1–12 (2009).
[16] Kucuk, M. and Agiralioğlu, N. "Wavelet regression technique for streamflow prediction", *Journal of Applied Statistics*, 33(9), pp. 943–960 (2006).
[17] Kisi, O. "Wavelet regression model as an alternative to neural networks for monthly streamflow forecasting", *Hydrological Processes*, 23, pp. 3583–3597 (2009).
[18] Kisi, O. "Wavelet regression model for short-term stream-flow forecasting", *Journal of Hydrology*, 389, pp. 344–353 (2010).
[19] Kisi, O. and Cimen, M. "A wavelet-support vector machine conjunction model for monthly streamflow forecasting", *Journal of Hydrology*, 399, pp. 132–140 (2011).
[20] Mallows, C.L. "Some comments on C_p", *Technometrics*, 15(4), pp. 661–675 (1973).
[21] Hagan, M.T. and Menhaj, M.B. "Training feed forward techniques with the Marquardt algorithm", *IEEE Transactions on Neural Networks*, 5(6), pp. 989–993 (1994).
[22] Ham, F. and Kostanic, I., *Principles of Neurocomputing for Science and Engineering*, McGraw-Hill, USA (2001).
[23] Daubechies, I. "The wavelet transform, time-frequency localization and signal analysis", *IEEE Transactions on Information Theory*, 36(5), pp. 961–1005 (1990).
[24] Partal, T. "Modeling evapotranspiration using discrete wavelet transform and neural networks", *Hydrological Processes*, 23, pp. 3545–3555 (2009).
[25] Bayazit, M. and Aksoy, H. "Using wavelets for data generation", *Journal of Applied Statistics*, 2(2), pp. 157–166 (2001).
[26] Rajaee, T., Nourani, V., Mohammad, Z.K. and Kisi, O. "River suspended sediment load prediction: application of ANN and wavelet conjunction model", *Journal of Hydrologic Engineering*, 16(8), pp. 613–627 (2011).
[27] Mallat, S.G. "A theory for multi resolution signal decomposition: the wavelet representation", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 11(7), pp. 674–693 (1989).
[28] Okkan, U. and Dalkilic, H.Y. "Reservoir inflows modeling with artificial neural networks: the case of Kemer Dam in Turkey", *Fresenius Environmental Bulletin*, 20(12), pp. 3110–3119 (2011).
[29] Sudheer, K.P., Gosain, A.K. and Ramasastri, K.S. "A data-driven algorithm for constructing artificial neural network rainfall–runoff models", *Hydrological Processes*, 16, pp. 1325–1330 (2002).
[30] Chou, C.M. and Wang, R.Y. "On-line estimation of unit hydrographs using the wavelet-based LMS algorithm", *Hydrological Science Journal*, 47(4), pp. 721–738 (2002).
[31] Mann, H.B. and Whitney, D.R. "On a test of whether one of two random variables is stochastically larger than the other", *Annals of Mathematical Statistics*, 18, pp. 50–60 (1947).
[32] Nourani, V. and Fard, M.S. "Sensitivity analysis of the artificial neural network outputs in simulation of the evaporation process at different climatologic regimes", *Advances in Engineering Software*, 47, pp. 127–146 (2012).

Umut Okkan is Research Assistant of Hydraulic Division in Civil Engineering Department of Balikesir University in Turkey. His main areas of interests are hydrological modeling, climatology, statistics, water resources management, G.I.S. and artificial intelligence methods. He is also reviewer of the journals "Teknik Dergi (Turkish)", "Neural Computing and Applications", "Geotechnical and Geological Engineering", "International Journal of Geotechnical Earthquake Engineering", "International Journal of Physical Sciences", "International Journal for Numerical and Analytical Methods in Geomechanics" and "Scientia Iranica".