

The combined use of wavelet transform and black box models in reservoir inflow modeling

Umut Okkan^{1*}, Zafer Ali Serbes²

¹ Department of Civil Engineering, Balıkesir University, Balıkesir, Turkey.

² Department of Farm Structures and Irrigation, Ege University, Izmir, Turkey; E-mail: zafer.ali.serbes@ege.edu.tr

* Corresponding Author. Phone: 0090 0266 612 11 94-95. E-mail: umutokkan@balikesir.edu.tr

Abstract: In the study presented, different hybrid model approaches are proposed for reservoir inflow modeling from the meteorological data (monthly precipitation, one-month-ahead precipitation and monthly mean temperature data) by the combined use of discrete wavelet transform (DWT) and different black box techniques. Multiple linear regression (MLR), feed forward neural networks (FFNN) and least square support vector machines (LSSVM) were considered as the black box methods. In the modeling strategy, meteorological input data were decomposed into wavelet sub-time series at three resolution levels and ineffective sub-time series were eliminated by Mallows' C_p based all possible regression method. As a result of all possible regression analyses, 2-months mode of time series of monthly temperature (D1_ T_t), 8-months mode of time series (D3_ T_t) of monthly temperature and approximation mode of time series (A3_ T_t) of monthly temperature were eliminated. Remained effective sub-time series were used as the inputs of MLR, FFNN and LSSVM. When the performances of the training and testing periods were compared, it was observed that the DWT-FFNN conjunction model has better results in terms of mean square errors (MSE) and determination coefficients (R^2) statistics. The discrete wavelet transform approach also increased the accuracy of multiple linear regression and least squares support vector machines.

Keywords: Hybrid model; Discrete wavelet transform; Mallows' C_p method; Multiple linear regression; Feed forward neural networks; Least squares support vector machines; Reservoir inflow modeling.

INTRODUCTION

Analyzing inflow data of reservoir can give significant statistical information for both past and future characteristics of a dam. Therefore, recording and modeling of inflow data have highly considerable roles in planning, reservoir operation studies and water resources management. There are three main approaches that have been used flow modeling which include white-box model (physically based distributed models), gray-box model (conceptual models) and the black-box model. The white and gray-box models aim to simulate the physical mechanisms underlying each component in the transformation of rainfall into flow, such as surface, subsurface and groundwater flow. Yet, due to their data requirements, uncertainties and complexities, these models may not be readily used in some applications (Abbott and Refsgaard, 1996).

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 types and stochastic models were commonly used. There are some important references which provide detailed descriptions of these conventional statistical models for hydrology context (Salas et al., 1980).

Over the past decade, artificial intelligence methods have been widely used as black-box technique in the modeling studies. Especially, artificial neural networks (ANN), which is a nonlinear computing artificial intelligence technique inspired by the learning function of a human brain, have been used in many applications. There are numerous ANN applications with different algorithms in hydrology context (Campolo et al., 1999; Cigizoglu, 2005; Coulibaly et al., 2000; Razavi and Araghinejad, 2009). In addition to ANN, support vector machines (SVM) which is one of the artificial intelligence meth-

ods has been also successfully applied (Bray and Han, 2004; Cimen, 2008; Lin et al., 2006; Liong and Sivapragasam, 2002; Okkan, 2012).

Although some artificial intelligence methods had been used extensively as useful tools for modeling, it has also some problems to deal with non-stationary data. As the hydrological time series includes several frequency components and have nonlinear relationships, some hybrid models which include different data-preprocessing and statistical techniques (set pair analysis, principle component analysis, clustering methods) have been used to raise the modeling performance (Wang et al., 2006; Wu et al., 2008).

Recently, an alternative data-preprocessing technique called wavelet transform has been found to be popular in modeling studies due to its advantages over conventional methods. There are several applications of the combined use of wavelet transform-neural networks (Anctil and Tape, 2004; Wang and Ding, 2003; Wang et al., 2009), wavelet transform-multiple regression (Kucuk and Agiralioglu, 2006), wavelet transform - support vector machine (Kisi and Cimen, 2011). In addition to flow modeling, wavelet transform was also applied to the analysis of lake levels. Cengiz (2011) investigated the variability of lake levels in four lakes in the Great Lakes region by using wavelet transform and global spectra method. It was showed that lake levels periodicities were generally the annual cycle.

In this study, three hybrid models are proposed for reservoir inflow modeling from the monthly meteorological data by using wavelet transform and some black box approaches (multiple linear regression, feed forward neural networks and least square support vector machines). Firstly, meteorological data were decomposed into wavelet sub-time series. Later, explanatory sub-time series were determined by using all possible regression method. These predictors constituted the inputs of these black box models.

METHODS

Feed forward neural networks (FFNN)

There are some books which provide a detailed description of the FFNN (Ham and Kostanic, 2001), and hence only a brief description of FFNN is given here. The running procedure of FFNN involves typically two phases; forward computing and backward computing.

In forward computing, each layer uses a weight matrix associated with all the connections made from the previous layer to the next layer.

The hidden layer has the weight matrix W_{ij} and activation function $f^{(1)}$; the output layer has the weight matrix W_{jm} and activation function $f^{(2)}$. Given the network input vector $x \in R^{nx1}$, the output of the output layer, which is the response (output) of the network $y \in R^{mx1}$, can be written as:

$$y_m = f^{(2)} \left\{ \sum_{j=1}^m \left[f^{(1)} \left(\sum_{i=1}^n x_i W_{ij} + b_j \right) \right] W_{jm} + b_m \right\}. \quad (1)$$

After the phase of forward computing, backward computing, which depends on the algorithms to adjust weights, is carried out. The process of adjusting these weights to minimize the differences between the actual and desired output values is called training or learning of network. If these differences (errors) are higher than the desired values, the errors are passed backwards through the weights of the network. In ANN terminology, this phase is also called the back propagation.

Depending on the techniques to train FFNN models, different back propagation algorithms have been developed. In this study, the Levenberg-Marquardt back propagation algorithm was used in training of the FFNN. The Levenberg-Marquardt back propagation algorithm is a second-order nonlinear optimization technique that is usually faster and more reliable than any other back propagation techniques (Coulibaly et al., 2000). The Levenberg-Marquardt optimization algorithm represents a simplified version of Newton method applied to the training of FFNN (Hagan and Menhaj, 1994). The training process can be viewed as finding a set of weights that minimize the error (e_p) for all samples in training set (T). The performance function is a sum of squares of the errors as follows:

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

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:

$$\Delta W_k = - [J_k^T J_k + \mu_k I]^{-1} J_k^T e_k. \quad (3)$$

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

$$W_{k+1} = W_k + \Delta W_k, \quad (4)$$

where J is the Jacobian matrix, I is the identify matrix, 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 β ($0 < \beta < 1$) whenever $E(W)$ decreases, while μ is divided by β whenever $E(W)$ increases in a new step (Coulibaly et al., 2000; Ham and Kostanic, 2001).

Least squares support vector machines (LSSVM)

The details of standard SVM method were given by Vapnik (1998). In this study, least squares support vector machine (LSSVM), which method is a simplified version of standard SVM, is used in modeling. LSSVM method can be described as follow according to the studies proposed by Suykens et al. (2002).

Like ANN, a LSSVM model is made up neurons which are organized in two layers. However, unlike ANN, it is based on theory in statistical learning and different error optimization approach (Suykens et al., 2002).

Given n -dimensional input vector $x_k \in R^n$, the output of the LSSVM model $y_k \in R$, can be written as:

$$y = W^T \phi(x) + b, \quad (5)$$

where W is weight vector, $\phi(x)$ nonlinear mapping function ($\phi: R^n \rightarrow R^{nh}$ mapping to high dimensional feature space) and b is the bias term.

Given training set $\{x_k, y_k\}_{k=1}^N$, the error optimization problem is defined as:

$$\min J(W, e) = \frac{1}{2} W^T W + \gamma \frac{1}{2} \sum_{k=1}^N e_k^2, \quad (6)$$

subject to equality constraints

$$y_k = W^T \phi(x_k) + b + e_k, \quad k = 1, \dots, N, \quad (7)$$

where e_k is error term and γ regularization parameter.

Because of high dimensional feature space, solution of this optimization problem cannot be obtained by using classical numerical methods. The solution of the optimization problem is obtained by considering the Lagrangian as:

$$L(W, b, e, \alpha) = J(W, e) - \sum_{k=1}^N \alpha_k \left\{ W^T \phi(x_k) + b + e_k - y_k \right\}, \quad (8)$$

where α_k are Lagrange multipliers. Conditions for optimality can be obtained by differentiating with respect to W , b , e_k and α_k , i.e.

$$\frac{\partial L}{\partial W} = 0 \rightarrow W = \sum_{k=1}^N \alpha_k \phi(x_k), \quad (9a)$$

$$\frac{\partial L}{\partial b} = 0 \rightarrow \sum_{k=1}^N \alpha_k = 0, \quad (9b)$$

$$\frac{\partial L}{\partial e_k} = 0 \rightarrow \alpha_k = \gamma e_k, \quad k = 1, \dots, N, \quad (9c)$$

$$\frac{\partial L}{\partial \alpha_k} = 0 \rightarrow W^T \phi(x_k) + b + e_k - y_k, \quad k = 1, \dots, N. \quad (9d)$$

Solution

$$\begin{bmatrix} 0 & \bar{1}^T \\ \bar{1} & \Omega + \gamma^{-1}I \end{bmatrix} \begin{bmatrix} b \\ \alpha \end{bmatrix} = \begin{bmatrix} 0 \\ y \end{bmatrix} \quad (10)$$

with $y = [y_1; \dots; y_N]$, $\bar{1} = [1; \dots; 1]$, $\alpha = [\alpha_1; \dots; \alpha_N]$ and by applying Mercer's theorem (Mercer, 1909).

$$\Omega_{km} = \phi(x_k)^T \phi(x_m) = K(x_k, x_m), \quad k, m = 1, \dots, N. \quad (11)$$

Resulting LSSVM model for function estimation can be expressed as:

$$y = \sum_{k=1}^N \alpha_k K(x_k, x) + b, \quad (12)$$

where $K(x_k, x)$ is the Kernel function.

The Kernel functions treated by standard SVM and LSSVM are the functions with linear, polynomial, Gaussian radial basis, exponential radial basis, splines etc. (Okkan, 2012). The Gaussian radial basis function used in this study can be defined as

$$K(x_k, x) = \exp\left(-\frac{\|x - x_k\|^2}{2\sigma^2}\right), \quad (13)$$

where σ is the Gaussian radial basis Kernel function width.

Wavelet transform

The wavelet transform, developed during the last decades, is an effective decomposition method. This method provides an analyzing way of a signal in both time and frequency and appears to be a more successful than the conventional Fourier transforms that do not provide time-frequency analysis for the variables involve non-stationary signals (Daubechies, 1990; Kucuk and Agiralioglu, 2006).

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

$$W(s, \tau) = |s|^{-1/2} \int_{-\infty}^{\infty} \psi^* \left(\frac{t - \tau}{s} \right) x(t) dt \quad \tau \in R, s \in R, s \neq 0, \quad (14)$$

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

Eq. (14) describes that wavelet transform is the decomposition of $x(t)$ under different resolution scale. The original series can be reconstructed using the inversion of transform (Rajae et al., 2011).

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

The Haar mother wavelet is defined as:

$$\psi(t) = \begin{cases} 1 & 0 \leq t \leq 1/2 \\ -1 & 1/2 \leq t \leq 1 \\ 0 & \text{otherwise} \end{cases} \quad (15)$$

For practical applications in hydrology, researchers have access to a discrete time signal, rather than to a continuous time signal. A discretization of Eq. (14) based on trapezoidal rule is perhaps the simplest discretization of the continuous wavelet transform (CWT). It 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. 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 (Rajae et al., 2011).

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 (Mallat, 1989):

$$W_{j,k} = 2^{-j/2} \sum_{t=0}^{N-1} \psi(2^{-j}t - k)x_t, \quad (16)$$

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_t) passes through two filters and are decomposed into wavelet sub-time series components, which can be computed by using Eq. (16), 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, high frequency components (detail (D) components). Thus, DWT allows one to study different investing behaviors in different time scales independently (Rajae et al., 2011). These computations can be carried out by wavelet toolbox for use with MATLAB.

STUDY AREA AND DATA

The study area covers the drainage basin of Demirkopru Dam (total drainage area of 6590 km²), which is located in the Gediz basin and Aegean region in Turkey. The reservoir of the dam is fed by four streams. Streamflow is observed at four gauging stations (Demirci/EIE-522, Deliinis/EIE-515, Selendi/EIE-514, and Gediz/EIE-523) located upstream of the dam (Fig. 1).

For the study area, the data sets of streamflow (10⁶ m³) were acquired from the records of the II. Regional Directorate of State Hydraulic Works of Turkey, and were arranged for the period between January 1977 and December 2006. In addition to inflow data, the data sets of precipitation at Demirci, Icikler, Kiransih, Fakili and Gediz meteorological stations were obtained from the State Meteorological Organization of Turkey and the General Directorate of State Hydraulic Works of Turkey. Next, the monthly mean areal precipitation values were calculated from these meteorological stations by Thiessen polygons. The monthly data sets of temperature at Demirci and Gediz meteorological stations were also obtained from the State Meteorological Organization of Turkey and the monthly mean

areal temperature values were computed by using arithmetical mean values from these meteorological stations for the period from January 1977 to December 2006.



Fig. 1. Gediz Basin and Demirköprü Dam (● flow gauging stations; ▲ meteorological stations).

In hydrological modeling, the determination of appropriate input variables would play an important role. In this study, modeling strategy that predicts inflow data from inputs based on monthly rainfall and temperature data. In addition to concurrent values of input data, precipitation values at various lags were also considered. A statistical approach proposed by Sudheer et al. (2002) was employed 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. The cross correlation function (CCF) between the reservoir inflow and precipitation values at various lags showed significant correlation at concurrent precipitation and one month of precipitation lag on the inflow. Thus, three input data (P_t, T_t, P_{t-1}) were prepared for the same periods of the monthly inflow records (P_t : monthly precipitation; P_{t-1} : one-month-ahead precipitation; T_t : monthly temperature).

APPLICATION AND RESULTS

In the study presented, DWT was linked to different black box models (MLR, FFNN and LSSVM) for the reservoir inflow modeling. In the modeling strategy, the input data (P_t, P_{t-1}, T_t) of training and testing periods were decomposed into a certain number of sub-time series components by DWT. In order to decompose the input data, a decomposition level must be selected. For the monthly scaled hydrological modeling studies, three decomposition levels are able to represent the related time series (Kisi and Cimen, 2011). Similarly, three decomposition levels were used in this study. Thus, meteorological data were decomposed and twelve sub-time series components (time series of 2-months mode (D1), 4-months mode (D2), 8-months mode (D3) and approximation mode (A3)) were obtained for the training and testing periods by proportions of 2/3 (January 1977–December 1996) and 1/3 (January 1997–December

2006), respectively. For example, the three levels decomposition of the P_t that yield four sub-signals by the Haar wavelet are presented in Fig. 2.

After the decomposition processes, the effective sub-time series components should be determined as some correlated sub-time series components may reduce the performances of the black box models. In this study, the effective sub-time series components were determined by Mallows’ C_p based all possible regression method. This statistical method is an effective way to determine the subset of variables in cases where there are a large number of potential predictor variables. In Mallows’ C_p approach, adequate models are those for which C_p is roughly equal to the number of variables in the model (Mallows, 1973). The C_p values can be computed as:

$$C_p = (N - k) \frac{MSE_i^2}{MSE_F^2} - (N - 2i - 1), \tag{17}$$

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.

Once the effective variables were selected, based on the C_p values for training period were calculated as shown in Tab. 1. Performances of the models with the nine variables (seen as bold and underlined characters in Tab. 1) are nearly the same as those of the model with twelve variables; that is, the explained variance of the monthly reservoir inflow data by the model with nine variables, which has the minimum C_p value, is nearly equal to that explained the full linear model with twelve variables. According to the analyses, 2-months mode of time series of monthly temperature (D1_ T_t), 8-months mode of time series of monthly temperature (D3_ T_t), and approximation mode of time series of monthly temperature (A3_ T_t) were eliminated.

Before presenting the input data to FFNN and LSSVM, the all data were normalized in the range [0–1] to prevent the models from being dominated by the variables with the extreme values.

Models with optimum structures provided the best training result in terms of the minimum mean square errors (MSE), and the maximum determination coefficients (R^2) were also employed for the testing period. In the application, some codes which involve discrete wavelet transform and multi-linear regression (DWT-MLR), Levenberg-Marquardt algorithm based feed forward neural networks (DWT-FFNN), least squares support vector machines (DWT-LSSVM) were written in MATLAB. The results of DWT-MLR, DWT-FFNN and DWT-LSSVM models were also compared with conventional MLR, FFNN and LSSVM models. MLR and DWT-MLR coefficients were computed by using the variables of training period (January 1977–December 1996). Thus, the testing of models was carried out using determined regression equations.

In FFNN applications, three widely used transfer functions, namely tangent sigmoid, linear, and log-sigmoid are evaluated in FFNN structure trials. The best results were achieved by using the log-sigmoid function. The optimum values of LSSVM and DWT-LSSVM parameters were determined by using the trial and error approach. The best configurations of all models are presented in Table 2.

DISCUSSIONS AND CONCLUSIONS

Many of the activities associated with planning and management of water resources require appropriate hydrological

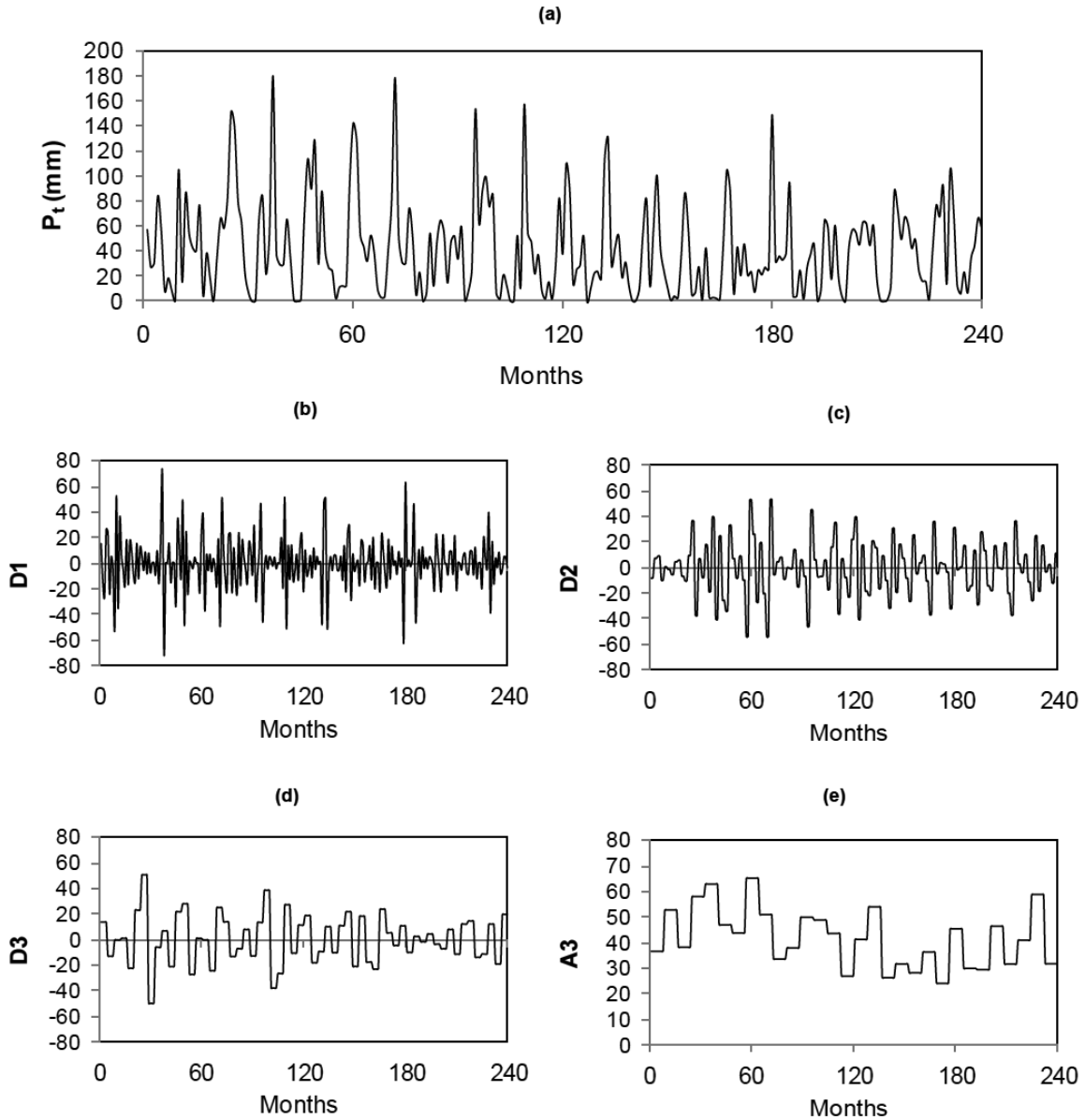


Fig. 2. Original time series (a), 2-months mode of time series (b), 4-months mode of time series (c), 8-months mode of time series (d) and approximation mode of time series (e) of monthly precipitation for training period.

models in practice. In particular, there is a necessity for modeling of flow values to optimize the system or to plan for future variations. The current study presented the applications of DWT based black box models (DWT-MLR, DWT-FFNN and DWT-LSSVM) compared with MLR, FFNN and LSSVM models used undecomposed data, for modeling of monthly inflows of Demirkopru Dam, based on the meteorological data.

In the study presented, monthly meteorological data (P_t, P_{t-1}, T_t) were decomposed into wavelet sub-time series by using the discrete wavelet transform (DWT). Three resolution levels were considered. Thus, twelve sub-time series components (time series of 2-months mode (D1), 4-months mode (D2), 8-months mode (D3) and approximation mode (A3) for each input data) were determined for the training period and testing period.

Because correlated sub-time series components reduced the generalization capabilities of the models, effective sub-time series components were determined by using Mallows C_p based all possible regression method. As a result of all possible regression analyses, effective sub-time series components were

selected. Thus, these variables that represented the new inputs of the MLR, ANN and LSSVM were treated so as to reduce collinearity.

Table 2 presents the statistical measures performed by the different kind of black box models. Because the phenomenon of the hydrological time series have the inherent complexities and nonlinearities, the performance of single MLR model was not suitable. On the contrary, DWT approach increased the performances of the MLR, FFNN and LSSVM in terms of MSE and R^2 values for training and testing periods (Okkan, 2011, 2012). The model structures with the best accuracy were also provided in Tab. 2. Furthermore, when $y = ax + b$ fitted lines are examined, one can see that for DWT-MLR, DWT-FFNN, and DWT-LSSVM models, the slope “ a ” gets closer to one, and “ b ” gets closer to zero, compared to the conventional MLR, FFNN and LSSVM models.

In terms of the best accuracy, DWT-FFNN model with Haar mother wavelet resulted in MSE reductions and, R^2 increases relative to that of the other models. This approach has generali-

Table 1. Summary of all possible regression analyses.

Number of variables	R^2 [%]	Adj. R^2 [%]	C_p	D1_	D2_	D3_	A3_	D1_	D2_	D3_	A3_	D1_	D2_	D3_	A3_
				P_t	P_t	P_t	P_t	P_{t-1}	P_{t-1}	P_{t-1}	P_{t-1}	T_t	T_t	T_t	T_t
1	25.0	24.7	327.8												
1	19.0	18.7	373.0			•									
1	17.8	17.5	382.1												•
2	42.8	42.3	195.9												•
2	39.4	38.9	221.5												•
2	37.5	37.0	236.0												•
3	55.3	54.7	104.2			•									•
3	52.2	51.6	127.6												•
3	51.9	51.3	129.7												•
4	63.9	63.3	41.1	•	•										•
4	60.8	60.2	64.5	•											•
4	60.5	59.9	66.7	•	•										•
5	66.1	65.4	27.0	•	•										•
5	64.9	64.2	35.6	•	•	•									•
5	64.8	64.0	37.0	•											•
6	67.1	66.2	21.5	•	•	•									•
6	66.9	66.0	22.9	•	•										•
6	66.9	66.0	23.0	•	•										•
7	67.9	66.9	17.4	•	•	•									•
7	67.9	66.9	17.5	•	•	•									•
7	67.7	66.7	18.9	•	•										•
8	68.7	67.6	13.4	•	•	•									•
8	68.6	67.5	14.0	•	•	•									•
8	68.4	67.3	15.4	•	•										•
9	69.4	68.2	9.9	•	•	•									•
9	68.9	67.6	14.1	•	•	•									•
9	68.8	67.6	14.3	•	•	•									•
10	69.6	68.3	10.6	•	•	•									•
10	69.6	68.3	10.7	•	•	•									•
10	69.5	68.1	11.5	•	•	•									•
11	69.8	68.3	11.4	•	•	•									•
11	69.7	68.2	12.2	•	•	•									•
11	69.6	68.2	12.3	•	•	•									•
12	69.8	68.2	13.0	•	•	•									•

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

Models	Model structures	R^2 (%)		MSE ($10^6 m^6$)	
		Training	Testing	Training	Testing
MLR	–	61.4	59.1	1890.7	2054.2
DWT-MLR	–	69.4	66.5	1498.5	2047.7
FFNN	$j = 3; \mu_0 = 0.001; \beta = 0.1; k = 10$	79.2	75.3	1020.9	1181.6
DWT-FFNN	$j = 9; \mu_0 = 0.001; \beta = 0.01; k = 7$	96.7	84.6	161.9	886.7
LSSVM	$\gamma = 16; \sigma^2 = 0.97$	81.1	73.6	932.2	1228.2
DWT-LSSVM	$\gamma = 45.5; \sigma^2 = 6.51$	96.5	81.3	180.9	894.4

j is number of neurons in hidden layer, k is iteration number, μ_0 is initial Marquardt parameter, β is decay rate of Levenberg-Marquardt algorithm, γ is regularization parameter, σ is Gaussian radial basis kernel function width.

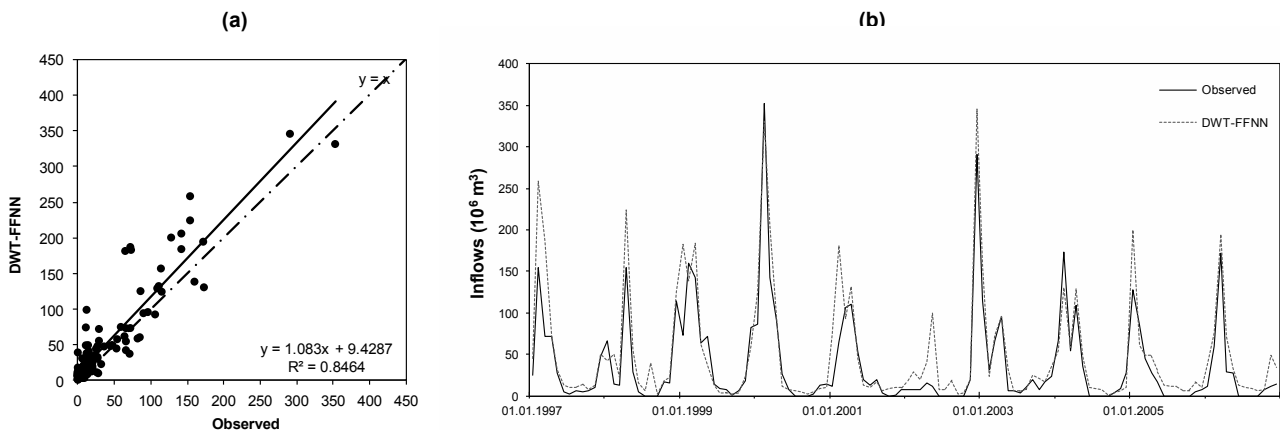


Fig. 3. The scatter plots (a) and hydrographs (b) of DWT-FFNN model for the testing period.

zation capability and thus can more easily capture inflow of Demirkopru Dam. The scatter plot and hydrograph of DWT-FFNN for the testing period are presented in Fig. 3.

Besides of the MSE and R^2 values, in terms of descriptive statistics, DWT-FFNN model proved to be successful in the training and testing periods. It was also seen that the extreme value statistics of DWT-FFNN for the testing period were quite satisfactory providing superior results compared with other conventional methods.

Although DWT-FFNN models have ability to represent complex and nonlinear relations of a basin, the structures of feed forward neural networks (FFNN) may be hard to determine and they can be determined using after many trial-and-error simulations. Therefore, DWT based MLR (DWT-MLR) and DWT based LSSVM (DWT-LSSVM) models that may be much easier for practical interpreting. Thus, these approaches can be used as alternative ways to DWT-FFNN and FFNN methods for monthly reservoir inflow modeling. For example, LSSVM and DWT-LSSVM have considerable features including the fact that requirement on Kernel function and the nature of the optimization problem results in a uniquely global optimum, generalization capability, and avoidance of converging to a local minimum solution.

The study presented uses some hybrid methodologies that are beneficial in reconstruction of reservoir inflow data. Especially, developed DWT-FFNN and DWT-LSSVM models, which are successful tools for the modeling of reservoir inflow series, can be used for future assessment of climate change impacts at the watershed scale. Thus, they can be employed to produce significant monthly scaled information for water resources management and planning problems such as irrigation planning, reservoir operation, water budget studies, watershed modeling, etc. The authors also suggest that the developed hybrid techniques can be also applied to other hydrological variables and other water resources problems to reconfirm the effectiveness of them.

In the study, the Mallat algorithm was used for the DWT of monthly meteorological time-series. For future study, the other DWT algorithms (e.g., Trous algorithm) and the other types of transform (e.g., fast wavelet transform, Fourier transform) may be used for the construction of the different neuro-computing models (e.g., other neural network algorithms, fuzzy logic). In addition to these, other data-preprocessing techniques (e.g., principle component analysis, factor analysis, set pair analysis) can be also compared with the wavelet transform for the future researches.

Acknowledgements. Authors feel responsible to thank English Instructor Cuneyt Guren and 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, which have helped to improve the paper, are also acknowledged.

REFERENCES

- Abbott, M.B., Refsgaard, J.C., 1996. Distributed Hydrological Modeling. Kluwer Academic Publishers, Dordrecht, 17–39.
- Anctil, F., Tape, D.G., 2004. An exploration of artificial neural network rainfall-runoff forecasting combined with wavelet decomposition. *J. Environmental Engng. and Science*, 3, 21–128.
- Bray, M., Han, D., 2004. Identification of support vector machines for runoff modelling. *J. Hydroinformatics*, 6, 265–280.
- Campolo, M., Andreussi, P., Soldati, A., 1999. River flood forecasting with a neural network model. *Water Resour. Res.*, 35, 1191–1197.
- Cengiz, T.M., 2011. Periodic structures of great lakes levels using wavelet analysis. *J. Hydrol. Hydromech.*, 59, 1, 24–35.
- Cigizoglu, H.K., 2005. Generalized regression neural networks in monthly flow forecasting. *Civil Engineering and Environmental Systems*, 22, 2, 71–84.
- Cimen, M., 2008. Estimation of daily suspended sediments using support vector machines. *Hydrological Sciences Journal*, 53, 3, 656–666.
- Coulibaly, P., Anctil, F., Bobee, B., 2000. Daily reservoir inflow forecasting using artificial neural Networks with stopped training approach. *J. Hydrol.*, 230, 244–257.
- Daubechies, I., 1990. The wavelet transform, time-frequency localization and signal analysis. *IEEE Transactions on Information Theory*, 36, 5, 961–1005.
- Hagan, M.T., Menhaj, M.B., 1994. Training feed forward techniques with the Marquardt algorithm. *IEEE Transactions on Neural Networks*, 5, 6, 989–993.
- Ham, F., Kostanic, I., 2001. Principles of Neurocomputing for Science and Engineering. Macgraw-Hill. USA.
- Kisi, O., Cimen, M., 2011. A wavelet-support vector machine conjunction model for monthly streamflow forecasting. *J. Hydrol.*, 399, 132–140.
- Kucuk, M., Agiralioğlu, N., 2006. Wavelet regression technique for streamflow prediction. *J. Applied Statistics*, 33, 9, 943–960.
- Lin, J.Y., Cheng, C.T., Chau, K.W., 2006. Using support vector machines for long-term discharge prediction. *Hydrological Sciences J.*, 51, 4, 599–612.
- Liong, S.Y., Sivapragasam, C., 2002. Flood stage forecasting with support vector machines. *J. Amer. Water Resources Assoc.*, 38, 1, 173–186.
- Mallat, S.G., 1989. A theory for multi resolution signal decomposition: the wavelet representation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 11, 7, 674–693.
- Mallows, C.L., 1973. Some comments on Cp. *Technometrics*, 15, 4, 661–675.
- Mercer, J., 1909. Functions of positive and negative type and their connection with the theory of integral equations. *Philosophical Transactions of the Royal Society, London*, 209, 415–446.
- Okkan, U., 2011. Application of Levenberg-Marquardt optimization algorithm based multilayer neural networks for hydrological time series modeling. *An Int. J. Optimization and Control: Theories & Applications*, 1, 1, 53–63.
- Okkan, U., 2012. Performance of least squares support vector machine for monthly reservoir inflow prediction. *Fresenius Environmental Bull.*, 21, 3, 611–620.
- Rajae, T., Nourani, V., Mohammad, Z.K., Kisi, O., 2011. River suspended sediment load prediction: Application of ANN and wavelet conjunction model. *J. Hydrol. Engng.*, 16, 8, 613–627.
- Razavi, S., Araghinejad, S., 2009. Reservoir inflow modeling using temporal neural networks with forgetting factor approach. *Water Resour. Management*, 23, 39–55.
- Salas, J.D., Delleur, J.W., Yevjevich, V., Lane, W.L., 1980. Applied modeling of hydrologic time series. *Water Resources Pub.*, p. 484.
- Sudheer, K.P., Gosain, A.K., Ramasastri, K.S., 2002. A data-driven algorithm for constructing artificial neural network rainfall-runoff models. *Hydrol. Processes*, 16, 1325–1330.

- Suykens, J.A.K., Van Gestel, T., De Brabanter, J., De Moor, B., Vandewalle, J., 2002. Least Squares Support Vector Machines. World Science, Singapore.
- Vapnik, V., 1998. Statistical Learning Theory. John Wiley & Sons, Toronto.
- Wang, W., Ding, J., 2003. Wavelet network model and its application to the prediction of hydrology. *Nature and Science*, 1, 67–71.
- Wang, W., van Gelder, P., Vrijling, J.K., Ma, J., 2006. Forecasting daily streamflow using hybrid ANN models. *J. Hydrol.*, 324, 383–399.
- Wang, W., Jin, J., Li, Y., 2009. Prediction of inflow at Three Gorges Dam in Yangtze River with wavelet network model. *Water Resour. Management*, 23, 2791–2803.
- Wu, C.L., Chau, K.W., Li, Y.S., 2008. Predicting monthly streamflow using data-driven models coupled with data-preprocessing techniques. *Water Resour. Res.*, 45, 8, 1–23.

Received 16 January 2012

Accepted 19 June 2012