A hybrid linear–nonlinear approach to predict the monthly rainfall over the Urmia Lake watershed using wavelet-SARIMAX-LSSVM conjugated model

Jamileh Farajzadeh and Farhad Alizadeh

ABSTRACT

The present study aimed to develop a hybrid model to predict the rainfall time series of Urmia Lake watershed. For this purpose, a model based on discrete wavelet transform, ARIMAX and least squares support vector machine (LSSVM) (W-S-LSSVM) was developed. The proposed model was designed to handle linear, nonlinear and seasonality of rainfall time series. In the proposed model, time series were decomposed into sub-series (approximation (a) and details (d)). Next, the sub-series were predicted separately. In the proposed model, sub-series were fed into SARIMAX to be predicted. The residual of predicted sub-series (error) of the rainfall time series was then fed into LSSVM to predict the residual components. Then, all predicted values were aggregated to rebuild the predicted time series. In order to compare results, first a classic modeling was performed by LSSVM. Later, wavelet-based LSSVM was used to capture the peak values of rainfall. Results revealed that Daubechies 4 and decomposition level 4 (db(4,4)) led to the best outcome. Due to the performance of db(4,4), it was selected to be applied in the proposed model. Based on results, it was observed that the W-S-LSSVM's performance was improved in comparison with other models.

Key words | discrete wavelet transform (DWT), LSSVM, rainfall, SARIMAX, Urmia Lake watershed

Jamileh Farajzadeh
East Azarbaijan Regional Water Authority,
Tabriz,
Iran

Farhad Alizadeh (corresponding author)
Department of Water Resources and
Environmental Engineering, Faculty of Civil
Engineering,
University of Tabriz,
29 Bahman Avenue,
Tabriz,
Iran
E-mail: *f.alizadeh.ce@gmail.com*

INTRODUCTION

In meteorological sciences, the prediction of rainfall over watersheds is an important research topic that has attracted much interest (Chang *et al.* 2016; Fadhel *et al.* 2016; Wang *et al.* 2016). Proper management of water resources involves planning, development, and distribution of water resources, which are associated with prediction of rainfall or spatial examination of the hydrologic cycle. Among hydrological cycle variables, rainfall prediction has a substantial role in runoff modeling and hence, water resource management (Valipour 2015). An accurate quantitative rainfall prediction can identify the potential for heavy rainfall and possible associated flash flooding, as well as providing information for hydrologic interests (Hall *et al.* 1998). This is particularly true in areas prone to flash flooding, where the prediction accuracy is highly dependent on the rapid availability of the rainfall distribution in advance (Ferraris *et al.* 2002).

Precipitation forecasting is a very difficult procedure. Calculating the response by applying conventional approaches in modeling rainfall time series is far from a trivial task, since the hydrologic processes are complex and involve various inherently complex predictors such as geomorphologic and climatic factors, which are as yet latent. Many researchers have carried out studies for quantitative precipitation forecasting using diverse techniques, including numerical weather prediction models and remote sensing observations (Yates *et al.* 2000; Ganguly & Bras 2003; Sheng *et al.* 2006; Davolio *et al.* 2008; Diomede *et al.* 2008), statistical models (Chu & He 1995; Chan & Shi 1999; DelSole & Shukla 2002; Munot & Kumar 2007; Li & Zeng 2008; Nayagam *et al.* 2008), non-parametric nearest neighbors method (Toth *et al.* 2000), and soft computing-based methods including artificial neural networks (ANN),

support vector regression, and fuzzy logic (Venkatesan et al. 1997; Silverman & Dracup 2000; Toth et al. 2000; Pongracz et al. 2001; Brath et al. 2002; Guhathakurta 2008; Nasseri et al. 2008; Sedki et al. 2009; Talei et al. 2010). Chang et al. (2014) developed a neuro-fuzzy networks based technique called adaptive network-based fuzzy inference system (ANFIS) in order to forecast rainfall for a watershed. Bias correction was undertaken on the data used as input into the developed technique to enhance the accuracy of the data. The results of the study demonstrated that the ANFIS technique yielded reliable and stable predictions. Valipour et al. (2013) demonstrated how the accuracy of the ANN model improved when it was combined with an autoregressive model. Accurate and reliable predictions can also be achieved by decomposing the original data into components so that the direct use of the original data, which results in less accurate predictions, can be avoided. Recently, the discrete wavelet transform (DWT) type of wavelet analysis has been used to decompose the original data into bands (sub-series) and, as a result, prediction accuracy has been improved for short lead times. Partal & Cigizoglu (2009) predicted the daily precipitation (lead time of one day) from meteorological variables including temperature, humidity, and precipitation using the wavelet–neural network method. In their study, they combined DWT, which is used to decompose the original data into bands, and ANN, which is used as a modeling tool. The study showed that the wavelet–ANN model provided a good fit with the observed data and performed significantly better than those obtained by either the conventional ANN model or a multilinear regression model. Solgi et al. (2014) combined wavelet analysis with ANN to predict daily rainfall for a one-day lead time in Verayneh, Iran, and this combined approach was compared with an adaptive neuro-fuzzy system. The results showed that the combined wavelet–neural network model has a better performance (coefficient of determination of 0.9 for one-day lead time) than the adaptive neuro-fuzzy system.

It can be inferred from the literature that the accuracy of prediction can be enhanced when wavelet transform (WT) is used as a pre-processing approach by decomposing the original data into sub-series. Despite this achievement, the predictions are with very short lead times, so models could be developed to predict rainfall accurately for longer lead times.

In this paper, a new multivariate black box model based on AI techniques and DWT is proposed for the rainfall modeling of Urmia watersheds located in West Azerbaijan, Iran, which have different climatologic characteristics. The main idea of conjugating models in predictions is to employ each model's unique features to capture different patterns in the data. Both theoretical and empirical findings suggest that combining different methods can be an efficient way to improve forecasting accuracy (Zhang 2003). One of such hybrid models is the autoregressive integrated moving average (ARIMA)-ANN model, which was developed in recent years and is still being used in engineering and financial time series forecasting. This model was first proposed and evaluated by Zhang (2003) in order to forecast a univariate time series without considering seasonal effect. This technique has enjoyed a few applications in hydrological time series modeling, which usually show highly non-stationary seasonal behavior. In the proposed model, first, monthly rainfall time series of the watershed is decomposed into sub-signals in various resolution levels and periodicity scales considering the existence of seasonality in the time series. Next, a seasonal ARIMA model with exogenous inputs (i.e., rainfall), SARIMAX, was combined with the support vector machine (SVM) approach to construct the hybrid WT-SARIMAX-LSSVM (W-S-LSSVM) model.

MATERIAL AND METHODS

Study area and used data

Urmia Lake watershed is located in the northwest of Iran between longitude 44° 07′E to 47° 53′E and latitude 35° 40′N to 38° 30′N with a total area of about 52,679 km^2 (Figure 1). Urmia Lake watershed is 51,876 km^2 and is one the six largest watersheds in Iran. It is the largest lake in Iran and one of the world's salt-saturated lakes, and has a significant role in moderating the climate of a vast area containing East and West Azerbaijan and Kurdistan provinces. According to the 39-year period of daily precipitation data (1973–2011), the annual mean precipitation over the watershed is 352 mm. The river flow is the main source to supply domestic, industrial, and agricultural consumption as well as groundwater feeding along the river, which are increasing annually. The monthly rainfall data of
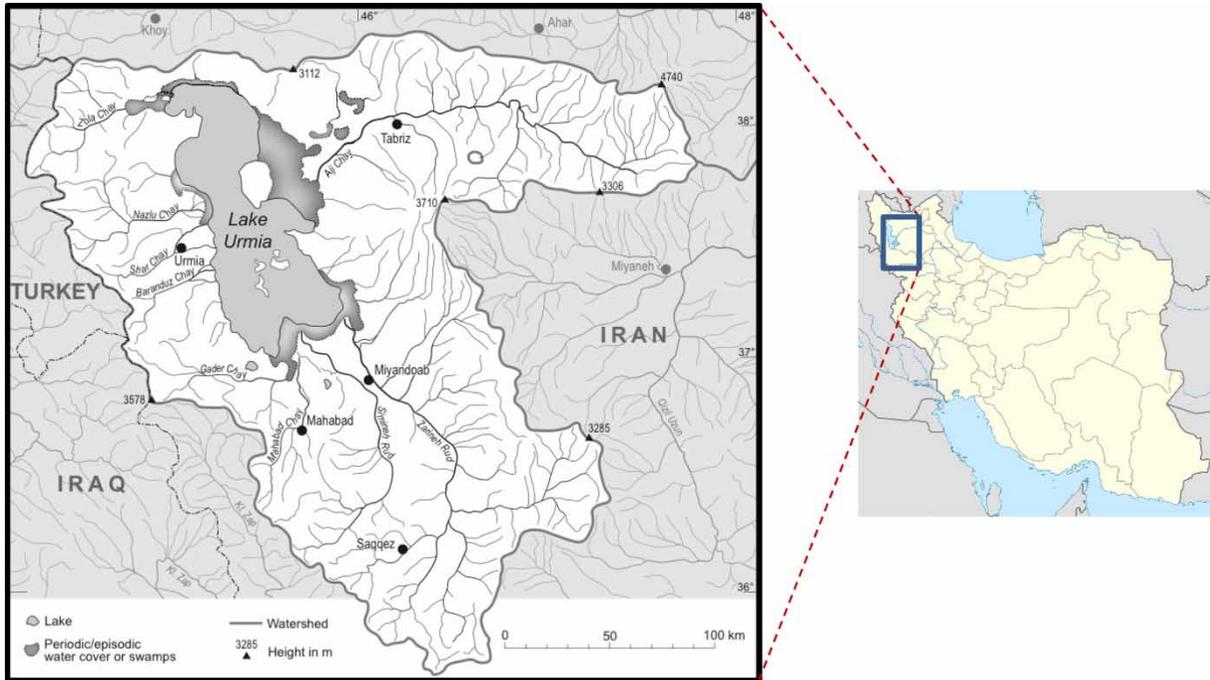
Figure 1 │ Geographic location of Urmia Lake watershed in northwest Iran with three provinces.

228 stations located inside and outside of Urmia Lake watershed were used to analyze and model the rainfall time series. The rainfall data used herein are on a monthly scale and cover the period from 1972 to 2014. For the goals of calibration and verification, time series were separated into calibration (70%) and verification (30%) purposes.

Least squares support vector machine (LSSVM)

LSSVM is the least squares formulation of a standard SVM. It was first proposed by Suykens & Vandewalle as a classifier in 1999. Unlike the inequality constraints introduced in the standard SVM, LSSVM proposed equality constraints in the formulation (Suykens & Vandewalle 1999; Noori *et al.* 2011). This results in the solution being transformed from one of solving a quadratic program to a set of linear equations known as the linear Karush–Kuhn–Tucker systems. The regression computation using LSSVM was later proposed in 2002. Multi-class classification can be considered as regression computation with multiple predefined threshold values. The LSSVM toolbox is also available at the web site <http://www.esat.kuleuven.be/sista/lssvmlab>. Suppose $\{(X_t, y_t)\}$ for $t = 1$ to N is a given set of data where $X_t = (x_{t1}, x_{t2}, \dots, x_{tk})$ is

the input vector with k multiple variables and y_t is the corresponding price data at time t which could be defined as:

$$y_t = f(x_t) = \langle W, \varphi(x_t) \rangle + b \qquad (1)$$

where $\langle , \rangle$ denotes the dot product, W is the weight vector, b is the bias, and $\varphi()$ is the mapping function that transfers input vector X_t into a much higher dimensional feature space (could be infinite). The corresponding optimization problem for LSSVM is formulated as:

$$\text{minimize} \quad \frac{1}{2}\|W\|^2 + \frac{1}{2}\gamma \sum_{t=1}^{N} e_t^2 \qquad (2)$$

$$\text{subjected to } y_t = \langle W, \varphi(x_t) \rangle + b + e_t$$

where e_t is the error variable at time t and γ is a regulation constant. The Lagrange function can be obtained as:

$$L(W, e, \alpha) = \frac{1}{2}\|W\|^2 + \frac{1}{2}\gamma \sum_{t=1}^{N} e_t^2$$
$$- \sum_{t=1}^{N} \alpha t(\langle W, \varphi(x_t) \rangle + b + e_t + y_t) \qquad (3)$$

where α_t is the Lagrange multipliers. The dual problem is then:

$$
\begin{cases}
y_t = \sum_{t,l=1}^{N} \alpha_t \varphi(x_t)\varphi(x_l) + b + \dfrac{\alpha_t}{\gamma} \\
\sum_{t=1}^{N} \alpha_t = 0
\end{cases}
\tag{4}
$$

An alternative formulation of (4) can be expressed as:

$$
\begin{bmatrix} 0 & 1_N^T \\ 1_N & \Omega + I_{N\gamma-1} \end{bmatrix} \begin{bmatrix} b \\ \alpha \end{bmatrix} = \begin{bmatrix} 0 \\ Y \end{bmatrix}
\tag{5}
$$

where $\Omega_{tl} = \langle \varphi(x_t), \varphi(x_l) \rangle$ is called the kernel function and $\alpha = \sum_{T=1}^{N} \alpha_t$. The final LSSVM for nonlinear functions can be written as:

$$
y_t = f(x_t) = \sum_{t=1}^{N} \alpha_t K(x, x_t) + b
\tag{6}
$$

where $K(x, x_t) = \langle \varphi(x_t), \varphi(x_l) \rangle$ is called the kernel function. Gaussian radial basis kernel is the most effective one in nonlinear function estimation and expressed as:

$$
K(x, x_t) = \exp\left(-\frac{\|x - x_t\|^2}{2\sigma^2} \right)
\tag{7}
$$

Basic concept of WT

WT has enjoyed widespread usage in the recent period since its inception in the early 1980s, and has some advantages in comparison to Fourier transform (Grossmann & Morlet 1984). Fourier analysis has a major deficiency; losing time information in transformation to the frequency domain. Wavelet analysis authorizes the use of long time spaces where more precise low-frequency information and shorter regions are necessary when high-frequency information is required. The main characteristic of WT is to process a time-scale localization of process, deduced from the dense support of its basic function. The WT searches for correlations between the signal and wavelet function. In real hydrological problems, the time series are usually in the discrete format rather than continuous and, therefore, the discrete WT in the following form is usually used (Mallat 1998):

$$
g_{m,n}(t) = \frac{1}{\sqrt{a_0^m}} g \times \left(\frac{t - nb_0 a_0^m}{a_0^m} \right)
\tag{8}
$$

where m and n are integers that control the wavelet dilation and translation, respectively; a_0 is a specified fined dilation, step greater than 1; and b_0 is the location parameter which must be greater than zero. The most common and simplest choice for parameters is $a_0 = 2$ and $b_0 = 1$. This power-of-two logarithmic scaling of the dilation and translation is known as the dyadic grid arrangement. The dyadic wavelet can be written in more compact notation as (Mallat 1998):

$$
g_{m,n}(t) = 2^{-m/2} g(2^{-m}t - n)
\tag{9}
$$

For a discrete time series, x_i, the dyadic WT becomes (Mallat 1998):

$$
T_{m,n}(t) = 2^{-m/2} \times \sum_{i=0}^{N=1} g(2^{-m}i - n)x_i
\tag{10}
$$

where $T_{m,n}$ is wavelet coefficient for the discrete wavelet of scale $a = 2^m$ and location $b = 2^{mn}$. Equation (10) considers a finite time series, x_i, $i = 0,1,2, \ldots, N_1$; and N is an integer power of 2: $N = 2M$. This gives the ranges of m and n as, $0 < n < 2^{M-m-1}$ and $1 < m < M$, respectively. The inverse discrete transform is given by Mallat (1998):

$$
x_i = \bar{T} + \sum_{m=1}^{M} \sum_{n=0}^{M-m-1} T_{m,n} 2^{-(m/2)} g(2^{-m}i - n)
\tag{11}
$$

or in a simple format as (Mallat 1998):

$$
x_i = \bar{T}(t) + \sum_{m=1}^{M} w_m(t)
\tag{12}
$$

where $\overline{T}(t)$ is called approximation sub-signal at level M and $W_m(t)$ are details of sub-signals at levels $m = 1, 2, \ldots, M$.

The wavelet coefficients, $W_m(t)$ $(m = 1, 2, \ldots, M)$, provide the detail signals, which can capture small features of

interpretational value in the data; the residual term, $\overline{T}(t)$, represents the background information of data.

Selection of mother wavelets

One of the advantages of the AI-based wavelet conjunction model compared with the AI method is its ability to identify data components in a time series such as irregular components with multi-level wavelet decomposition. For this purpose, the rainfall signal was separated into different temporal scales (levels) by DWT. Selection of an appropriate wavelet function poses significant challenges and is governed largely by the problem at hand and some of the distinctive properties of the wavelet function, such as i) its region of support and ii) the number of vanishing moments (Maheswaran & Khosa 2012a).

Haar wavelet

These are symmetric and non-continuous wavelets, and the number of vanishing moments in this case is equal to one (1). Haar wavelets have been found appropriate for applications in signals that have sharp changes because of the relatively narrow span over which their energy is distributed.

Daubechies wavelet

These are a family of wavelets and include the Haar wavelet, written as db1, as a special case. These are compactly supported wavelets with extreme phase and highest number of vanishing moments for a given support width. The Daubechies wavelets have associated minimum-phase scaling filters, are both orthogonal and biorthogonal, and do not have an explicit analytic expression except for the db1 (or Haar) form.

Sym wavelet

Sym wavelet (or, in a compact form, symlet) is only near symmetric, and its other properties are similar to those of the general Daubechies wavelets, dbNs. Symlet has the highest number of vanishing moments for a given support width, and its associated scaling filters are near linear-phase filters.

Order N can be 2, 3, y. These wavelets can be both orthogonal and biorthogonal and provide compact support.

Coif wavelet

Coiflets are discrete wavelets that have scaling functions with vanishing moments. The wavelets are near symmetric, their wavelet functions have $N/3$ vanishing moments and scaling functions $N/3-1$, and they have been used in many applications.

Meyer wavelet

The Meyer wavelet is an orthogonal wavelet that is indefinitely differentiable with infinite support.

Proposed model

In the proposed model, the rainfall signal was first decomposed into sub-signals with different scales, i.e., a large scale sub-signal and several small scale sub-signals in order to obtain temporal characteristics of the input time series. For a given time series, the time series corresponding to $a(t)$ (i.e., $Ia(t)$) is approximation sub-signal (large scale) of the original signal, and ith detailed sub-signal (small scale) is identified by I (i.e., $Idi(t)$), where i is the decomposition levels of the rainfall ($I(t)$) time series. In this study, some irregular mother wavelets such as Haar, db2 and db4 (Daubechies wavelet of order 2 and 4), sym2 and coif2 and Meyer were used. Mallat (1998) can be referred to for more information about these wavelets. Since rainfall time series are measured in point (such as rain gauges) and in a discrete form, in this study dyadic DWT was used. Also, 'a trous' is another alternative decomposition and reconstruction algorithm and can be considered for the modeling. It is a simple and fast algorithm, but it is a redundant algorithm, and its key is to determine an appropriate low-pass filter (usually a spline is selected); its determination is not usually easy for the non-expert.

The seasonal ARIMA with exogenous repressor (SARIMAX or structural SARIMA) process differs from the SARIMA process, ostensibly because it takes cognizance of an exogenous input, which consists of additional exogenous variables that could explain the behavior of

the dependent variable. For construction of the wavelet-SARIMAX-LSSVM (W-S-LSSVM) model, the proposed methodology by Zhang (2003) in developing the ARIMA-ANN model is followed in this paper, but also considering two other components. First, because of the seasonality characteristics of the hydrological time series, and second, in order to employ the relevant variables (i.e., approximation and detail sub-series of precipitation) in the modeling, wavelet-SARIMAX-LSSVM is proposed to take advantage of SARIMAX for precipitation modeling instead of the ARIMA model, which considers just autoregressive property of the time series in the modeling, and also its formulation contains only one variable. According to the complex behavior of the hydrological time series (e.g., rainfall time series), they might include both linear and nonlinear components. Therefore, application of an individual linear model (such as ARIMA or SARIMA) for such signals may not be adequate. On the other hand, with all the advantages of the LSSVM, these models are not universal models suitable for all circumstances and may yield mixed results in modeling linear problems. Hence, it is not wise to apply LSSVM blindly to any type of data without any data pre/post-processing. For example, in a process which includes seasonal and non-stationary characteristics, temporal data pre-processing may improve the efficiency of the modeling (Zhang 2003; Cannas *et al.* 2006); but if the process shows trend in space, spatial data pre-processing such as spatial clustering should be employed prior to the main modeling (Nourani & Kalantari 2010). Therefore, since it is difficult to understand the characteristics of the data in a real problem completely, hybrid modeling can be a useful methodology for practical use, and by combining different models, different aspects of the underlying patterns may be captured.

The proposed hybrid W-S-LSSVM model consists of three steps. First, the rainfall time series are decomposed by DWT. In the next step, a SARIMAX model is initially fitted to model the linear components of DWT (approximations and details) of the rainfall time series so that the other input data are also used as external (regressor) variable in this exogenous modeling. The result of this stage is the estimation of actual value of rainfall time series for one-time ahead (i.e., $\widehat{I}(t+1)$) via the used SARIMAX.

Now, considering the actual time series to be composed of a linear structure and a nonlinear component:

$$I(t+1) = \widehat{I}(t+1) + e(t+1)$$
(i.e. for decomposed time series:
$$Ia(t+1) = \widehat{Ia}(t+1) + eIa(t+1))$$
(13)

The SARIMAX model just yields the linear estimation for rainfall $(\widehat{I}(t+1))$ and cannot handle the nonlinear patterns; hence, the residual time series of the SARIMAX (i.e., $e(t+1)$) may contain the nonlinear relationship, which relation can be detected by an LSSVM via the second step of the W-S-LSSVM modeling. (In the proposed model, the decomposed sub-series are predicted; as an instance, $Ia(t+1)$ is the approximation signal of rainfall as target data.) With N_N input vectors in the LSSVM which include residual values of N_N previous months and according to Equation (13), the LSSVM model for the residuals will be as follows:

$$e(t+1) = f(e(t), \ldots, (t-N_N)) + \varepsilon_{t+1}$$
(14)

where $f(t)$ is a nonlinear function determined by the LSSVM and $e(t+1)$ is the random error of sub-series. Consequently, considering Equations (13) and (14), the forecasted value for the rainfall at time $t+1$ will be computed by Equation (15). This procedure is illustrated in Figure 2.

$$I(t+1) = \widehat{I}(t+1) + f(t) + \varepsilon_{t+1}$$
(i.e. for decomposed time series:
$$Ia(t+1) = \widehat{Ia}(t+1) + f(t) + \varepsilon Ia(t+1))$$
(15)

Efficiency criteria

In this study, the total data were separated into calibration and verification sets. Two different criteria were selected to evaluate the efficiency of the proposed forecasting methods; the root mean square error (*RMSE*) and the determination coefficient (*DC*). The *RMSE* and *DC* were applied to exhibit discrepancies between predicted and observed values (Adamowski & Chan 2011). *RMSE* is the standard deviation of the residuals (prediction errors). Residuals are a measure of how far from the regression line data points are. In other words, it describes how concentrated the
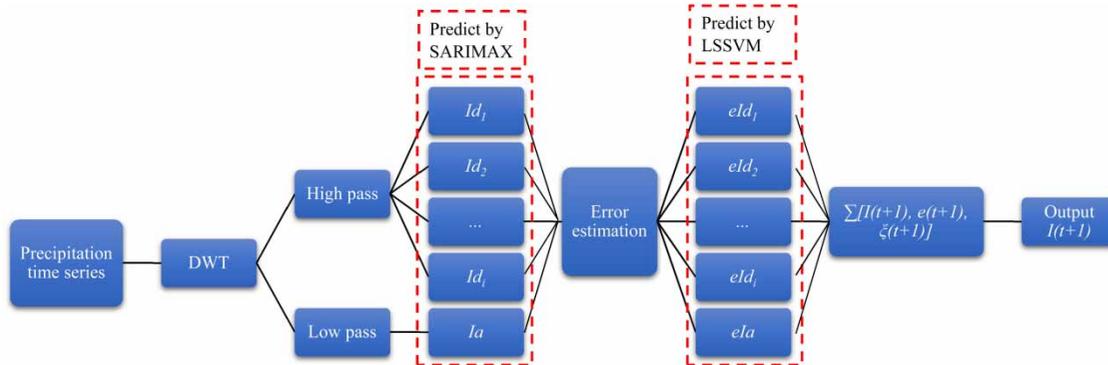
Figure 2 | Schematic of modeling via W-S-LSSVM to predict rainfall time series.

data are around the line of best fit. The *RMSE* is utilized to measure forecast exactness, which gives a positive value by squaring the errors. The *RMSE* decreases from large positive values for poor performance to zero for perfect forecasts.

On the other hand, the *DC* was used to quantify the modeling accuracy by generating a positive value. The *DC* grows from zero for perfect forecasts through large positive values as the differences between forecasts and observations become increasingly large. It describes how many data points fall within the results of the line formed by the regression. The higher the coefficient, the higher the percentage of points the line passes through when the data points and line are plotted. Values of 1 or 0 would indicate that the regression line represents all or none of the data, respectively. A higher coefficient is an indicator of a better goodness of fit for the observations. The usefulness of *DC* is its ability to find the likelihood of future events falling within the predicted outcomes. The idea is that if more samples were added, the coefficient would show the probability of a new point falling on the line.

Obviously, a high value for *DC* (up to 1) and a small value for *RMSE* indicate high efficiency of the model. Legates & McCabe (1999) indicated that a hydrological model can be sufficiently evaluated by these statistics. These measures are not oversensitive to extreme values (outliers) and are sensitive to additive and proportional differences between model predictions and observations. Therefore, correlation-based measures (e.g., *DC* statistic) can indicate that a model is a good predictor (Legates & McCabe 1999).

Input and output variables are normalized by scaling between 0 and 1 to eliminate their dimensions and to ensure that all variables receive equal attention during calibration of the model. Therefore, *RMSE* values are transformed between 0 and 1 (Nourani *et al.* 2016). Descriptions of *DC*, *RMSE*, and *Normalization* are presented as follows:

$$DC = 1 - \frac{\sum_{i=1}^{N} \left(I_{obs_i} - I_{com_i}\right)^2}{\sum_{i=1}^{N} \left(I_{obs_i} - \bar{I}_{obs}\right)^2} \qquad (16)$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^{N} \left(I_{obs_i} - I_{com_i}\right)^2}{N}} \qquad (17)$$

$$l_i = \frac{L_i - L_{min}}{L_{max} - L_{min}} \qquad (18)$$

where N, I_{obs_i}, I_{com_i} and $\bar{I}_{obs}$ are the number of observations, observed data, computed values, and mean of observed data, respectively. L_i is the actual value and l_i is the respective normalized value. L_{min} and L_{max} are the minimum and maximum of all used values, respectively.

RESULTS AND DISCUSSION

In this study, all averages at the watershed level, for each variable, are calculated using the Thiessen (Okabe *et al.* 2000) weighted-average method, as follows:

$$I_w = \frac{1}{A_w} \sum_{i=1}^{n} A_{pi} Y_i \qquad (19)$$

where A is the area, w and p refer to the watershed and Thiessen polygon, I is the variable (i.e., precipitation), and i refers to each climatological station of the watersheds, so that I_w is the Thiessen weighted-average variable. First, monthly rainfall averages for each of the climatological stations were calculated (not shown). The watersheds' mean values per variable were then calculated using the Thiessen method (Figure 3). The monthly rainfall data of 228 stations located inside and outside of Urmia Lake watershed were used to analyze and model the rainfall time series. Due to the existence of numerous rain gauges in the study area, and stability of polygons in the period of study, the aforementioned method seems to be suitable in comparison to other interpolation methods in terms of time, cost, and accuracy. In the Thiessen method, a constant area and weight are calculated for each rain gauge during the period of study and these weights are applied to estimate the mean monthly rainfall. According to Figure 3, different shapes of polygons were calculated. The polygons inside the lake are larger, since there are fewer rain gauges covering the area inside the lake. On the other hand, polygons for the zones with more rain gauges had smaller areas, which led to a more accurate outcome. Figure 4 shows the obtained time series via Thiessen method over the Urmia watershed. Also, statistical analysis of the obtained time series is shown in Figure 4.

At first, LSSVM without any pre-processing was used to predict the monthly rainfall time series of the Urmia Lake watershed. A correct use of LSSVM is to select an appropriate kernel function and tune the related hyperparameters. A number of kernels were discussed by researchers, but studies suggested the effectiveness of radial basis function (RBF) kernel in the case of SVM in the majority of civil engineering applications (Gill *et al.*
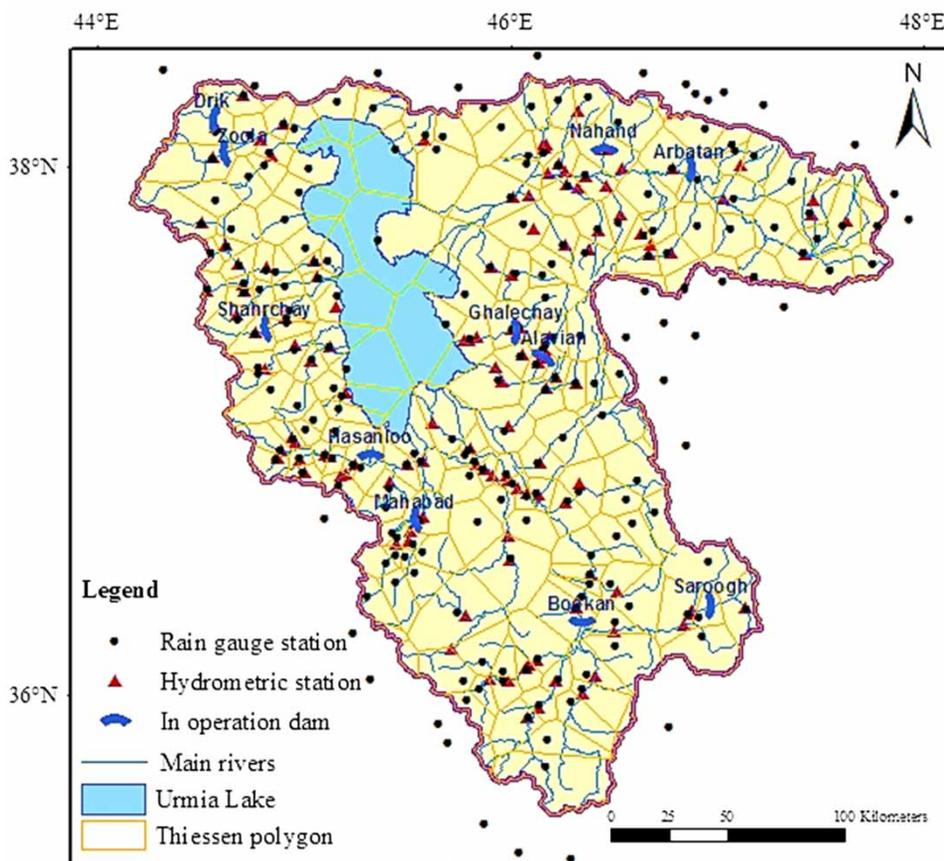

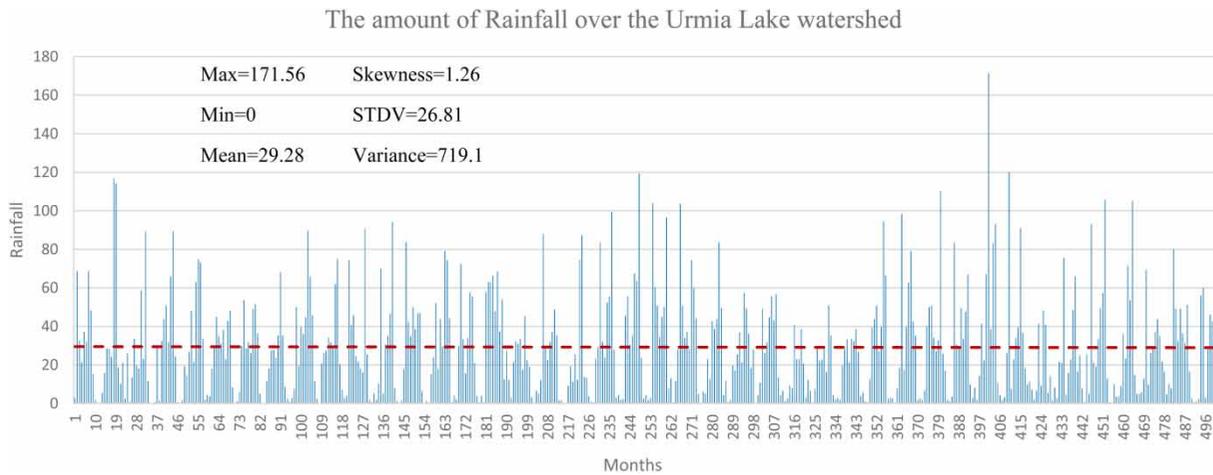**Figure 3** | Location of rain gauges of the Urmia Lake watershed with Thiessen polygons.

Figure 4 │ Calculated time series for the Urmia Lake watershed using Thiessen polygons.

2006; Goel & Pal 2009; Nourani *et al.* 2016). For a fair comparison of results, a RBF kernel where σ is a kernel specific and γ is margin parameter was used with LSSVM in this study. The tuned parameters (σ, γ) were optimized via a grid search procedure. The structure of LSSVM is demonstrated in Table 1 (for different applications). According to Table 1, minimum, maximum, and increment values of parameters were identified and then optimized through a trial-and-error process to find the best structure of LSSVM for all the models.

In order to perform classic modeling by LSSVM, some input combinations were determined according to the antecedents of rainfall dataset. Since rainfall time series might behave as a Markovian process, it is important to understand how the chronological sequence of months can influence rain prediction. In other words, the value of features at the current time step could be related to the conditions at the previous time steps. Therefore, the following shows input combinations (1–4), in which different

numbers of input variables (I) were considered as input features to predict the rainfall ($I(t + 1)$ for all stations).

Comb. (1): I(t)
Comb. (2): I(t), I(t−1)
Comb. (3): I(t), I(t−1), I(t−2)
Comb. (4): I(t), I(t−1), I(t−2), I(t−12)

For all defined combinations, t demonstrates the time step (monthly step) and I represents rainfall features. The output consists of sole variable, i.e., I at future time steps ($I(t + 1)$). Comb. 4 was designed to take advantage of seasonality of rainfall features ($t−12$).

An LSSVM through RBF kernel was trained for all input features to predict one-month-ahead rainfall values employing normalized rainfall data. In order to determine the LSSVM results, the performance evaluation results based on statistical evaluation values are demonstrated for different input combinations in Table 2. According to Table 2, Comb. 1 did not perform adequately in the calibration and verification steps. It

Table 1 │ Optimal values of user defined parameters used in this study

| Used model | RBF kernel's parameters | |
	σ^a	γ^a
LSSVM	1-50-1	0.1-10-0.5
W-LSSVM	1-100-1	0.1-20-0.5
W-S-LSSVM	1-100-0.5	0.1-20-0.25

[a]*Note:* For σ and γ, the first number is the minimum value, the second number is the maximum value, and the third number is the increment value.

Table 2 │ Results of modeling by LSSVM using antecedent-based models

| Input comb. | Calibration | | Verification | |
	RMSE[a]	DC	RMSE	DC
1.	0.05	0.7	0.054	0.65
2.	0.045	0.79	0.044	0.73
3.	0.037	0.81	0.046	0.74
4.	0.021	0.89	0.032	0.85

[a]*RMSE* values are dimensionless due to the normalization.

can be inferred that target rainfall could not be predicted with one-month lag values precisely. Of Combs. 2 and 3, Comb. 3 could predict the rainfall relatively better. Comb. 3 also employed one to three months' ago values of rainfall features. Comb. 4, which included a 12 months' lag, showed improvement in comparison to other antecedent-based combinations; nevertheless, generally, results were not assured. It was observed that Comb. 4, which was designed by considering the seasonality of the process (with a period of one year) in addition to the autoregressive characteristic of the data, could yield better results. Also, due to the poor Markovian process in monthly modeling of precipitation, it was observed that Combs. 1 to 3 led to weak outcome and proved that monthly rainfall is conditioned by seasonal components. The outcome of classic modeling by LSSVM using Comb. 4 is demonstrated in Figure 5.

Wavelet-based modeling results

Illustration of time series concurrent in both spectral and temporal terms could improve the performance of LSSVM to better describe the process. The WT was used to decompose rainfall time series, and the optimum decomposition level was obtained through trial-and-error procedure. In order to have a comprehensive overview of decomposition level, initially the following formula (which offers a minimum level of decomposition) was employed (Aussem *et al.* 1998; Nourani *et al.* 2009):

$$L = \text{int}[\log(N)] \tag{20}$$

where L and N are decomposition level and number of time series data, respectively. For the case study $N = 504$, thus, $L = 3$. This experimental equation was derived for fully autoregressive signals, and only considering time series length without being concerned about seasonal signatures of the hydrologic process (Nourani *et al.* 2011). Hence, by use of decomposition level equal to 3, not all seasonalities in the main time series might be taken into account. Decomposition level 3 yields three detailed sub-series (i.e., 2^1 months mode, 2^2 months mode, 2^3 months mode, which is nearly an eight-month mode), but according to the hydrological base of the time series, there might be other dominant seasonalities with longer periods. Therefore, one other decomposition level (i.e., 4) was also examined, in which decomposition level 4 led to better results via the presented modeling and was selected as the optimum decomposing level. Level 4 sub-series contains four sets of details: 2^1 months mode, 2^2 months mode, 2^3 months mode, and 2^4 months mode, which is nearly a year mode. Therefore, the seasonality of the process up to one year could be handled by the model. On the other hand, it was observed that, for example, d_4 at some time t included knowledge of past values back to $x(t-16)$; therefore, the first 15-odd data were not used in the training process (Aussem *et al.* 1998).

The treatment of boundary merits special attention when implementing wavelet-based approaches for forecasting applications. Estimation of wavelet coefficient at any time t uses observations on past states going back up to $(t-p)$ and the future, unrealized states up to $(t+p)$. In
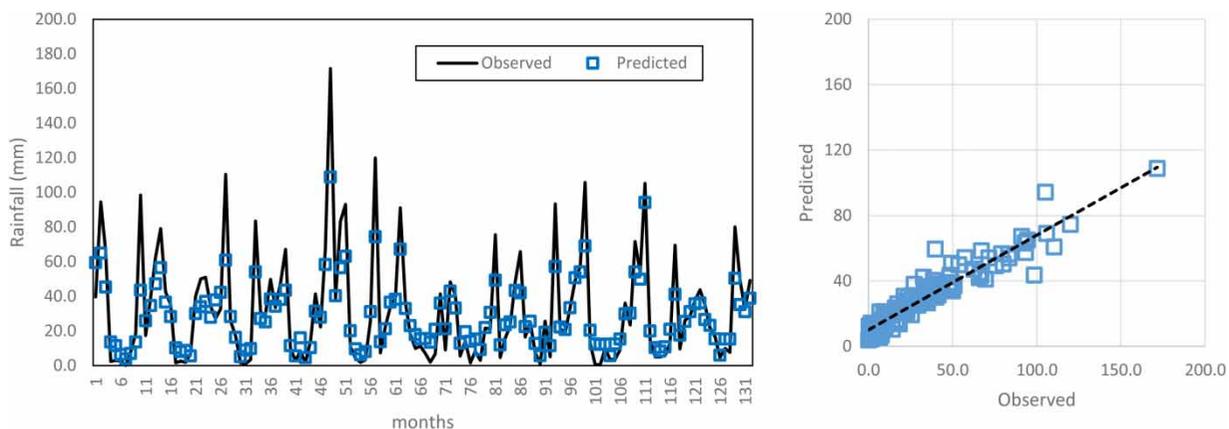

Figure 5 │ Predicted values of rainfall via LSSVM using Comb. 4.

general wavelet applications, various kinds of boundary conditions such as i) periodic boundary, ii) reflective boundary extension, and iii) constant extension are usually used for extending the series up to $(t + p)$ (Maheswaran & Khosa 2012b).

Daubechies-2 (db2), Meyer, Sym 2, and coif2 mother wavelets were applied to decompose the rainfall time series. To investigate the effect of form similarity between mother wavelet and main time series, another selection of mother wavelets was also examined. For this purpose, two other mother wavelets (Haar and Daubechies-4, db4) were chosen according to the formation of main signals. As the Haar wavelet has a pulsed shape, it could properly capture the signal features of rainfall time series and may yield comparatively high efficiency (Nourani *et al.* 2009). Also, due to the formation of a db4 wavelet that is similar to the hydrological signal, it could capture the signal features, especially peak points, efficiently and led to comparatively good results. For instance, Figure 6 shows approximation and details of sub-series of rainfall time series decomposed by the Haar and db4 mother wavelet at level 4. Boundary (edge) effect is one of the deficiencies in the application of WT, which happens due to application of the wavelet to the beginning and end of the time series (signals), at which point there are no data before and after. As a solution, the zero padding method was used (Addison *et al.* 2001). Therefore, since the time series were long enough for the case study, suitable amounts of data were neglected from the beginning and end parts of time series after wavelet application.

The time series were decomposed into three and four levels by different kinds of WTs, i.e.: 1) the Haar wavelet, a simple wavelet; 2) the Daubechies wavelet (db), the most popular wavelet; 3) the sym wavelet, with three sharp peaks; and 4) the coif1 wavelet and other mother wavelets (Mallat 1998). Results of modeling by wavelet-based LSSVM (W-LSSVM) are presented in Figure 7. Figure 7 shows that the db4 mother wavelet and decomposition level 4 had better performance, and the subsequent results of the W-LSSVM model are presented for this mother wavelet. Predicted values via db (4,4) and W-LSSVM are demonstrated in Figure 8.

It was observed that the larger errors of LSSVM were found to be in peak values, while the oscillations of the rainfall values are relatively well predicted. On the other hand, the W-LSSVM model predicted both extreme and average values more precisely in comparison to the ad hoc LSSVM model. It could be inferred that the better fitting of extreme and average values significantly reduced the *RMSE* and increased *DC* as well. Comparison of Figures 5 and 8 proved this statement; also, it was observed that the fitted trend line in Figure 8 (scatter plot) is closer to 45° (perfect prediction) in comparison to Figure 5.

Results of proposed model

Results approved the capability of db(4,4) in forecasting the rainfall time series. Therefore, db(4,4) was selected to be applied in the proposed model. Accordingly, in the first step, rainfall data were decomposed into several sub-series using DWT (db4). In the second step, various input combinations of sub-series with different time lags were tried for predicting the sub-series separately in two stages (SARIMAX and LSSVM).

The proposed hybrid model (i.e., W-S-LSSVM) was examined for the rainfall modeling of the Urmia Lake watersheds. The results of the model for db(4,4) are presented in Table 3. At first, components of db(4,4) were predicted by SARIMAX. Next, residual values (error) were fed into LSSVM to predict the error and, finally, by aggregating the sub-series that obtain from modeling, predict the rainfall values. The best W-S-LSSVM structure in monthly modeling belongs to the SARIMAX $(2, 1, 1) (1, 1, 1)_3$ $Ia(t)$ model, in which the obtained residuals were entered into the LSSVM with a window format of three successive months' residuals (i.e., $eIa(t), eIa(t-1), eIa(t-2)$) in order to find the one time step-ahead residual (i.e., $e(t+1)$). In this model, at first a SARIMAX model was fitted to the calibration rainfall dataset so that the decomposed rainfall values at previous time steps were also considered as the exogenous input of the model. The fitted model gave the estimated values for the rainfall $(\widehat{I}(t))$, and then the residual time series of the calibration data was computed by subtracting the decomposed observed data $(I(t))$ from the estimated decomposed values $\widehat{I}(t)$ (i.e., $eIa(t) = Ia(t) - \widehat{Ia}(t)$). This residual time series of the calibration dataset was then used to train an LSSVM model to detect the nonlinear relationship among the residuals. The best architecture of the LSSVM was obtained via a trial-and-error process. The trained network was then used to find the residual values
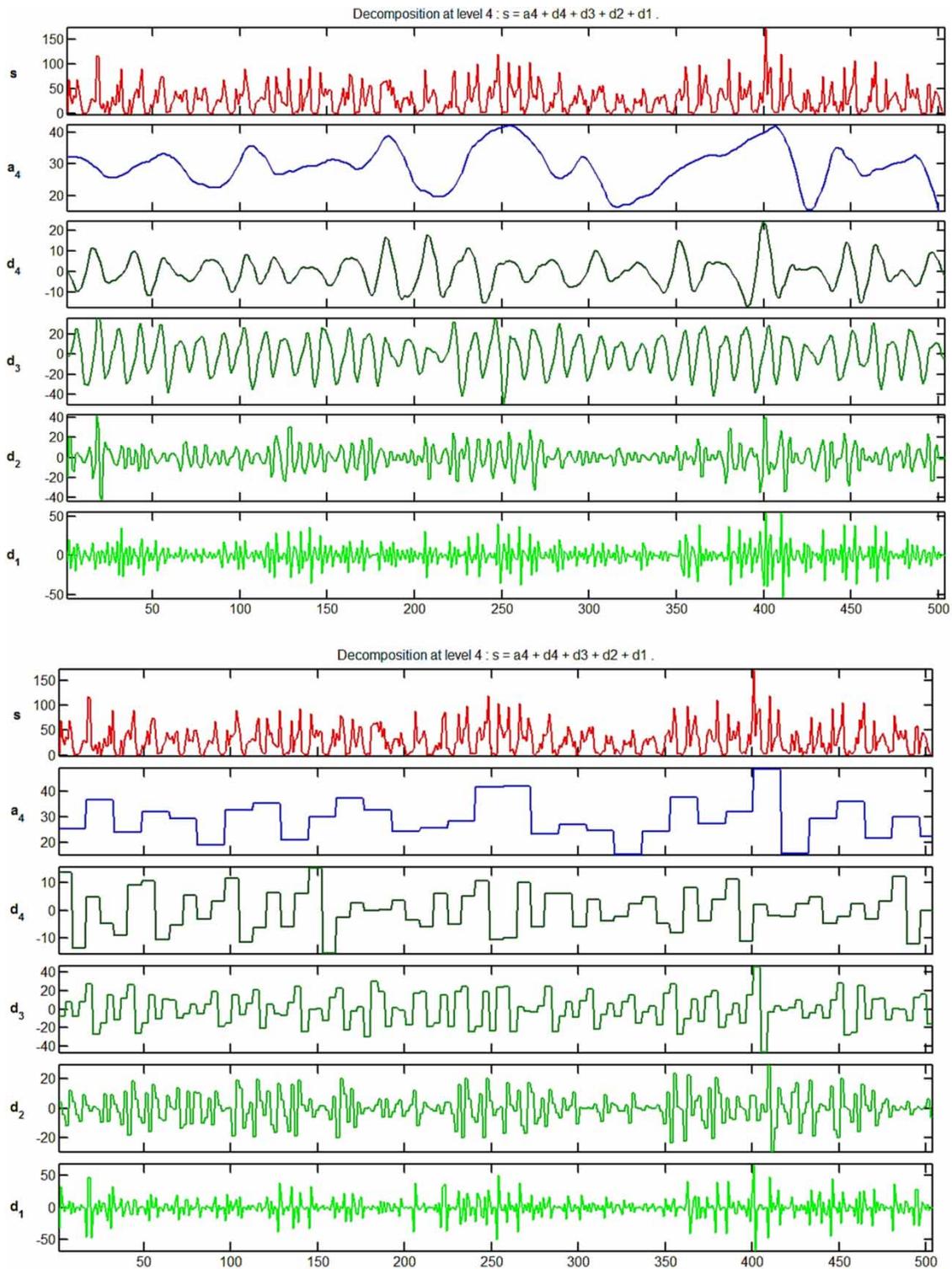
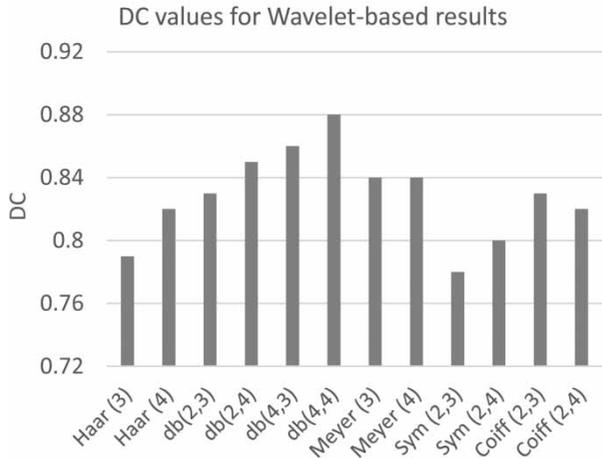**Figure 6** │ Decomposed time series of rainfall by db(4,4) and Haar (4).

Figure 7 │ Results of modeling via wavelet-based inputs. Outcomes of different mother wavelets and decomposition levels (3 and 4) are presented.

of the verification dataset month by month. On the other hand, SARIMAX $(2, 1, 1)$ $(1, 1, 1)_{12}$ $I(t)$ was separately used in order to find the linear estimation of the rainfall values of the verification dataset (Figure 9(a)). Finally, the

values of the computed rainfall for the verification dataset were obtained by Equation (15) and can be compared with the observed rainfall in terms of the *DC* and *RMSE*. In addition to the results presented in Table 3, the plot of the prediction in the verification stage is shown in Figure 9 for monthly modeling. It can be seen that the proposed model could capture the peak values, and better results were obtained in comparison to other models. Table 4 represents the overall performance of proposed models for rainfall prediction in the verification period. Based on results, it was observed that the W-S-LSSVM's performance was improved by about 7–8% in comparison to classic models and about 5–6% in comparison to the W-LSSVM model. A possible reason for this may be the fact that single models that use only rainfall data as input values for modeling perform weakly, whereas wavelet conjunction models use subseries that include various information about the studied phenomenon which improves the prediction capabilities. Using only one resolution component to model, the rainfall
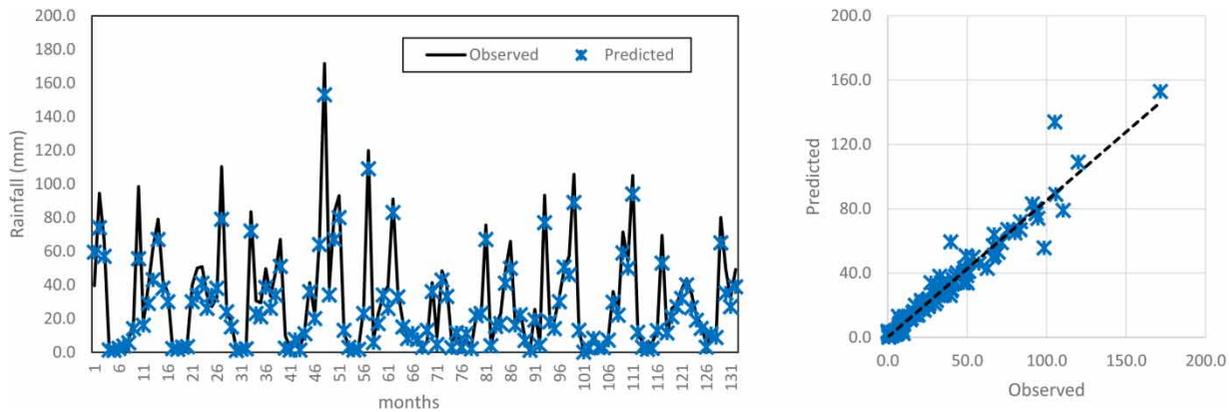

Figure 8 │ Predicted values of rainfall via LSSVM using db(4,4).

Table 3 │ Results of modeling by W-S-LSSVM to predict components of db(4,4) in the verification period

Target data	SARIMAX structure	LSSVM structure	RMSE[a]	DC
$Ia(t+1)$	$(2,1,1)(1,1,1)_3$ $Ia(t)$	$eIa(t)$, $eIa(t-1)$, $eIa(t-2)$	0.029	0.92
$Id_1(t+1)$	$(2,1,1)(1,1,1)_2$ $Id_1(t)$	$eId_1(t)$, $eId_1(t-1)$, $eId_1(t-2)$	0.028	0.9
$Id_2(t+1)$	$(2,1,1)(1,1,1)_4$ $Id_2(t)$	$eId_2(t)$, $eId_2(t-1)$, $eId_2(t-2)$	0.022	0.9
$Id_3(t+1)$	$(2,1,1)(1,1,1)_8$ $Id_3(t)$	$eId_3(t)$, $eId_3(t-1)$	0.019	0.89
$Id_4(t+1)$	$(2,1,1)(1,1,1)_{16}$ $Id_4(t)$	$eId_4(t)$, $eId_4(t-1)$	0.014	0.88

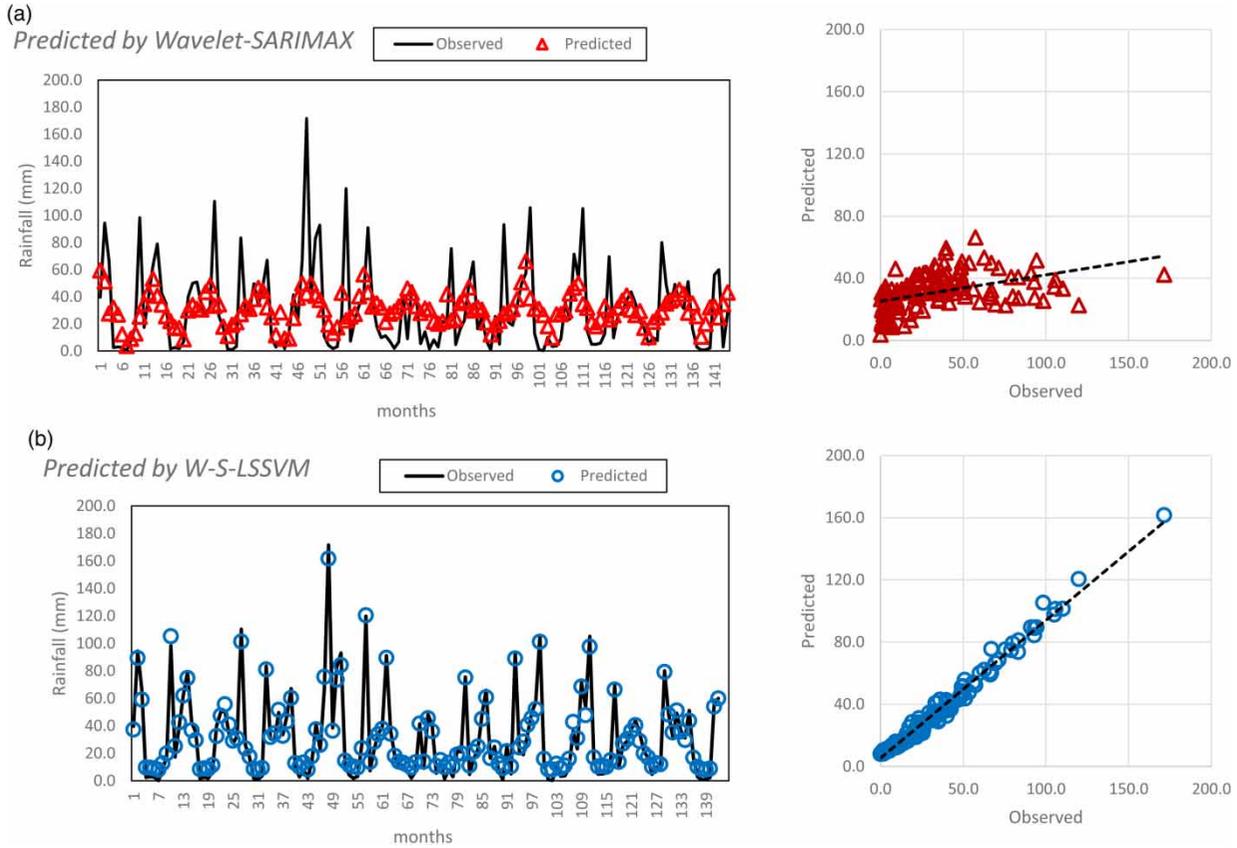[a]*RMSE* values are dimensionless due to the normalization.

Figure 9 │ Predicted values of rainfall (a) via WT-SARIMAX and (b) W-S-LSSVM using db(4,4).

Table 4 │ Comparison of results

Model	RMSE[a]	DC
LSSVM	0.032	0.85
W-LSSVM	0.028	0.87
W-S-LSSVM	0.019	0.92

[a]RMSE values are dimensionless due to the normalization.

time series does not readily explain the internal mechanism of the time series (Chou & Wang 2002).

In order to validate the stability of the proposed model, the rainfall time series were separated into two parts equally and were predicted by the proposed model. The first 50% of data (1972–1993) resulted in $DC = 0.93$ and $RMSE = 0.018$, and the other 50% (1994–2014) resulted in $DC = 0.92$ and $RMSE = 0.021$ (overall results). It was observed that even if the number of points in the times series was changed, the outcome proved to be robust.

By considering the significance of rainfall predictions in hydro-environmental systems, in this study, the ability of the proposed W-S-LSSVM model was also investigated for multi-step-ahead predictions. Two, three, six, and twelve month-ahead predictions were performed by employing W-S-LSSVM (db(4,4)) as the best model of one-step-ahead prediction. The obtained results are presented in Table 5. Results indicated that two-month-ahead prediction led to acceptable outcome

Table 5 │ Results of modeling by W-S-LSSVM to predict two, three, six, and twelve months ahead

Months ahead	Calibration		Verification	
	RMSE[a]	DC	RMSE[a]	DC
2	0.025	0.9	0.040	0.84
3	0.078	0.78	0.09	0.71
6	0.11	0.67	0.19	0.59
12	0.1	0.73	0.17	0.58

[a]RMSE values are dimensionless due to the normalization.

by the means of *RMSE* and *DC* values. Despite the results obtained for two-month-ahead prediction, the outcome of prediction for three-, six-, and twelve-month-ahead weakened relatively. It could be concluded from Table 5 that in multi-step-ahead predictions, the proposed model led to relatively poor performance in comparison to one-step-ahead modeling and lost its accuracy by increasing the prediction horizon. The reason is because in multi-step-ahead prediction, the Markovian property weakens by time step increase.

Characteristics of applied approaches in proposed models

SARIMAX is a classic time series forecasting model, and therefore was used as a hybrid model to predict the rainfall time series. Although a linear model like SARIMAX is unlikely to model a complex nonlinear hydrological process, the hybrid SARIMAX-LSSVM model could handle both linearity and nonlinearity of rainfall time series.

For the case of LSSVM, there are some main advantages. First, it has a regularization parameter to handle the over-fitting issues. Second, it uses the kernel; therefore, expert knowledge about the problem could be built via engineering the kernel. Third, an SVM is defined by a convex optimization problem (no local minima) for which there are efficient methods. Finally, it is an approximation to a bound on the test error rate, and there is a substantial body of theory behind it that suggests it to be a solid approach (Noori *et al.* 2011).

The disadvantages are that the theory only really covers the determination of the parameters for a given value of the regularization and kernel parameters and choice of kernel. In a way, the SVM moves the problem of over-fitting from optimizing the parameters to model selection. Unfortunately, kernel models can be quite sensitive to over-fitting the model selection criterion.

Results proved the capability of DWT in prediction of rainfall time series by giving incredible insight into the proposed artificial intelligence (AI)-based model. However, when applying DWT to annual precipitation data, the decimation and filtering of some of the signals at each scale resulted in signals not being time invariant. If a signal is shifted in time, the DWT does not produce a shifted version of the non-shifted transform. Redundant non-decimated WT can be used to address the issue. This algorithm essentially follows the same decomposition tree in DWT, except that instead of filtering and decimating the signals, the filters are being upsampled by inserting zeros in between prior filtering at each scale (Maheswaran & Khosa 2012b).

CONCLUSION

Data pre/post-processing are proven to be an efficient way to improve the prediction precision in most models. The current study presented an AI-based three-stage model to improve forecasting quality of rainfall time series. To validate the proposed model, the Urmia Lake watershed was selected. The monthly rainfall data of 228 stations located inside and outside of Urmia Lake watershed were used to analyze and model the rainfall time series by Thiessen polygon approach. The obtained time series, which represented the rainfall values over the watershed, was used in the modeling process. First, the time series were predicted via ad hoc LSSVM via antecedent-based inputs. Comb. 4, which included the previous months' rainfall features with lag of one, two, and twelve months, outperformed other models. Next, W-LSSVM via different mother wavelets and decomposition levels was used to amend the results. Based on outcome, db (4,4) led to the best prediction. Accordingly, db(4,4) was used in the proposed W-S-LSSVM model. In the W-S-LSSVM model, decomposed sub-series were forecast separately by ARIMAX. Next, residual values were predicted by LSSVM. Outcomes of the model were summed to build the time series of rainfall. Based on results, it was observed that the W-S-LSSVM's performance was improved by about 7–8% in comparison to classic models and about 5–6% in comparison to the W-LSSVM model. Also, due to the importance of prediction of rainfall for future time steps, two-, three-, six-, and twelve-month-ahead predictions were performed using the best model. It was concluded that the higher the prediction horizon, the lower modeling accuracy is obtained.

ACKNOWLEDGEMENTS

REFERENCES

Adamowski, J. & Chan, H. F. 2011 A wavelet neural network conjunction model for groundwater level forecasting. *Journal of Hydrology* **407**, 28–40.

Addison, P. S., Murrary, K. B. & Watson, J. N. 2001 Wavelet transform analysis of open channel wake flows. *Journal of Engineering Mechanics* **127**, 58–70.

Aussem, A., Campbell, J. & Murtagh, F. 1998 Wavelet-based feature extraction and decomposition strategies for financial forecasting. *Journal of Computational Intelligence in Finance* **6** (2), 5–12.

Brath, A., Montanari, A. & Toth, E. 2002 Neural networks and non-parametric methods for improving real time flood forecasting through conceptual hydrological models. *Hydrology Earth System Science* **6** (4), 627–640.

Cannas, B., Fanni, A., See, L. & Sias, G. 2006 Data preprocessing for river flow forecasting using neural networks: wavelet transforms and data partitioning. *Physics and Chemistry of Earth* **31** (18), 1164–1171.

Chan, J. C. L. & Shi, J. E. 1999 Prediction of the summer monsoon rainfall over South China. *International Journal of Climatology* **19** (11), 1255–1265.

Chang, F. J., Chiang, Y. M., Tsai, M. J., Shieh, M. C. & Sorooshian, S. 2014 Watershed rainfall forecasting using neuro-fuzzy networks with the assimilation of multisensor information. *Journal of Hydrology* **508**, 374–384.

Chang, J., Wei, J., Wang, Y., Yuan, M. & Guo, J. 2016 Precipitation and runoff variations in the Yellow River Basin of China. *Journal of Hydroinformatics* **19** (1), 138–155.

Chou, C. M. & Wang, R. Y. 2002 On-line estimation of unit hydrographs using the wavelet-based LMS algorithm. *Hydrological Science Journal* **47** (5), 721–738.

Chu, P. S. & He, Y. X. 1995 Long-range prediction of Hawaiian winter rainfall using canonical correlation-analysis. *International Journal of Climatology* **14** (6), 659–669.

Davolio, S., Miglietta, M. M., Diomede, T., Marsigli, C., Morgillo, A. & Moscatello, A. 2008 A meteo-hydrological prediction system based on a multi-model approach for precipitation forecasting. *Natural Hazards and Earth System Sciences* **8** (1), 143–159.

DelSole, T. & Shukla, J. 2002 Linear prediction of Indian monsoon rainfall. *Journal of Climate* **15** (24), 3645–3658.

Diomede, T., Davolio, S., Marsigli, C., Miglietta, M. M., Moscatello, A., Papetti, P., Paccagnella, T., Buzzi, A. & Malguzzi, P. 2008 Discharge prediction based on multi-model precipitation forecasts. *Meteorology and Atmospheric Physics* **101** (3–4), 245–265.

Fadhel, S., Rico-Ramirez, A. M. & Han, D. 2016 Exploration of an adaptive merging scheme for optimal precipitation estimation over ungauged urban catchment. *Journal of Hydroinformatics* **19** (2), 225–237.

Ferraris, L., Rudari, R. & Siccardi, F. 2002 The uncertainty in the prediction of flash floods in the Northern Mediterranean environment. *Journal of Hydrometeorology* **3**, 714–727.

Ganguly, A. R. & Bras, R. L. 2003 Distributed quantitative precipitation forecasting (DQPF) using information from radar and numerical weather prediction models. *Journal of Hydrometeorology* **4** (6), 1168–1180.

Gill, M. K., Asefa, T., Kemblowski, M. W. & Makee, M. 2006 Soil moisture prediction using Support Vector Machines. *Journal of the American Water Resources Association* **42** (4), 1033–1046.

Goel, A. & Pal, M. 2009 Application of support vector machines in scour prediction on grade-control structures. *Engineering Application of Artificial Intelligence* **22** (2), 216–223.

Grossmann, A. & Morlet, J. 1984 Decomposition of Hardy function into square integrable wavelets of constant shape. *Journal of Mathematical Analysis and Applications* **5**, 723–736.

Guhathakurta, P. 2008 Long lead monsoon rainfall prediction for meteorological sub-divisions of India using deterministic artificial neural network model. *Meteorology and Atmospheric Physics* **101** (1–2), 93–108.

Hall, T., Brooks, H. E. & Doswell, C. E. 1998 Precipitation forecasting using a neural network. *Weather and Forecasting* **14**, 338–345.

Legates, D. R. & McCabe, G. J. 1999 Evaluating the use of goodness of fit measures in hydrologic and hydroclimatic model validation. *Water Resources Research* **35**, 233–241.

Li, F. & Zeng, Q. C. 2008 Statistical prediction of East Asian summer monsoon rainfall based on SST and sea ice concentration. *Journal of the Meteorological Society of Japan* **86** (1), 237–243.

Maheswaran, R. & Khosa, R. 2012a Multiscale nonlinear model for monthly streamflow forecasting: a wavelet-based approach. *Journal of Hydroinformatics* **14** (2), 424–442.

Maheswaran, R. & Khosa, R. 2012b Comparative study of different wavelets for hydrologic forecasting. *Computers and Geosciences* **46**, 284–295.

Mallat, S. G. 1998 *A Wavelet Tour of Signal Processing*, 2nd edn. Academic Press, San Diego, CA.

Munot, A. A. & Kumar, K. K. 2007 Long range prediction of Indian summer monsoon rainfall. *Journal of Earth System Science* **116** (1), 73–79.

Nasseri, M., Asghari, K. & Abedini, M. J. 2008 Optimized scenario for rainfall forecasting using genetic algorithm coupled with artificial neural network. *Expert Systems and Applications* **35** (3), 1415–1421.

Nayagam, L. R., Janardanan, R. & Mohan, H. S. R. 2008 An empirical model for the seasonal prediction of southwest monsoon rainfall over Kerala, a meteorological subdivision of India. *International Journal of Climatology* **28** (6), 823–831.

Noori, R., Karbassi, A. R., Moghaddamnia, A., Han, D., Zokaei-Ashtiani, M. H., Farokhnia, A. & GhafariGousheh, M. 2011 Assessment of input variables determination on the SVM model performance using PCA, Gamma test and forward selection techniques for monthly stream flow prediction. *Journal of Hydrology* **401**, 177–189.

Nourani, V. & Kalantari, O. 2010 Integrated artificial neural network for spatiotemporal modeling of

rainfall–runoff–sediment processes. *Environmental Engineering Science* **27** (5), 411–422.

Nourani, V., Alami, M. T. & Aminfar, M. H. 2009 A combined neural-wavelet model for prediction of Ligvanchai watershed precipitation. *Engineering Applications of Artificial Intelligence* **22**, 466–472.

Nourani, V., Kisi, O. & Komasi, M. 2011 Two hybrid artificial intelligence approaches for modeling rainfall-runoff process. *Journal of Hydrology* **402**, 41–59.

Nourani, V., Alizadeh, F. & Roushangar, K. 2016 Evaluation of a two-stage SVM and spatial statistics methods for modeling monthly river suspended sediment load. *Water Resources Management* **30**, 393–407.

Okabe, A., Boots, B., Sugihara, K. & Chiu, S-N. 2000 *Spatial tessellations: concepts and applications of Voronoi diagrams*, 2nd edn. Wiley, Chichester.

Partal, T. & Cigizoglu, H. K. 2009 Prediction of daily precipitation using wavelet-neural networks. *Hydrological Science Journal* **54** (2), 234–246.

Pongracz, R., Bartholy, J. & Bogardi, I. 2001 Fuzzy rule-based prediction of monthly precipitation. *Physics and Chemistry of the Earth, Part B: Hydrology, Oceans and Atmosphere* **26** (9), 663–667.

Sedki, A., Ouazar, D. & El Mazoudi, E. 2009 Evolving neural network using real coded genetic algorithm for daily rainfall–runoff forecasting. *Expert Systems and Applications* **36** (3), 4523–4527.

Sheng, C., Gao, S. & Xue, M. 2006 Short-range prediction of a heavy precipitation event by assimilating Chinese CINRAD-SA radar reflectivity data using complex cloud analysis. *Meteorology and Atmospheric Physics* **94** (1–4), 167–183.

Silverman, D. & Dracup, J. A. 2000 Artificial neural networks and long-range precipitation prediction in California. *Journal of Applied Meteorology and Climatology* **39** (1), 57–66.

Solgi, A., Nourani, V. & Pourhaghi, A. 2014 Forecasting daily precipitation using hybrid model of wavelet-artificial neural network and comparison with adaptive neurofuzzy inference system. Case Study: Verayneh Station, Nahavand. *Advances in Civil Engineering* **2014**, 1–12.

Suykens, J. A. K. & Vandewalle, J. 1999 Least square support vector machine classifiers. *Neural Processing Letters* **9** (3), 293–300.

Talei, A., Chua, L. H. C. & Quek, C. 2010 A novel application of a neuro-fuzzy computational technique in event-based rainfall–runoff modeling. *Expert Systems and Applications* **37** (12), 7456–7468.

Toth, E., Brath, A. & Montanari, A. 2000 Comparison of short-term rainfall prediction models for real-time flood forecasting. *Journal of Hydrology* **239** (1–4), 132–147.

Valipour, M. 2015 Long-term runoff study using SARIMA and ARIMA models in the United States. *Meteorological Applications* **22** (3), 592–598. http://dx.doi.org/10.1002/met. 1491 (June 12).

Valipour, M., Banihabib, M. E. & Behbahani, S. M. R. 2013 Comparison of the ARMA, ARIMA, and the autoregressive artificial neural network models in forecasting the monthly inflow of Dez dam reservoir. *Journal of Hydrology* **476**, 433–441.

Venkatesan, C., Raskar, S. D., Tambe, S. S., Kulkarni, B. D. & Keshavamurty, R. N. 1997 Prediction of all India summer monsoon rainfall using error-back-propagation neural networks. *Meteorology and Atmospheric Physics* **62** (3–4), 225–240.

Wang, D., Wang, X., Liu, L., Wang, D., Huang, H. & Pan, C. 2016 Evaluation of CMPA precipitation estimate in the evolution of typhoon-related storm rainfall in Guangdong, China. *Journal of Hydroinformatics* **18** (6), 1055–1068.

Yates, D. N., Warner, T. T. & Leavesley, G. H. 2000 Prediction of a flash flood in complex terrain. Part II: a comparison of flood discharge simulations using rainfall input from radar, a dynamic model, and an automated algorithmic system. *Journal of Applied Meteorology and Climatology* **39** (6), 815–825.

Zhang, G. P. 2003 Time series forecasting using hybrid ARIMA and neural network model. *Neurocomputing* **50**, 159–175.