

## Forecasting of meteorological drought using Wavelet-ANFIS hybrid model for different time steps (case study: southeastern part of east Azerbaijan province, Iran)

Bagher Shirmohammadi · Hamidreza Moradi · Vahid Moosavi ·  
Majid Taie Semiromi · Ali Zeinali

Received: 5 September 2012 / Accepted: 7 May 2013 / Published online: 22 May 2013  
© Springer Science+Business Media Dordrecht 2013

**Abstract** Drought is accounted as one of the most natural hazards. Studying on drought is important for designing and managing of water resources systems. This research is carried out to evaluate the ability of Wavelet-ANN and adaptive neuro-fuzzy inference system (ANFIS) techniques for meteorological drought forecasting in southeastern part of East Azerbaijan province, Iran. The Wavelet-ANN and ANFIS models were first trained using the observed data recorded from 1952 to 1992 and then used to predict meteorological drought over the test period extending from 1992 to 2011. The performances of the different models were evaluated by comparing the corresponding values of root mean squared error coefficient of determination ( $R^2$ ) and Nash–Sutcliffe model efficiency coefficient. In this study, more than 1,000 model structures including artificial neural network (ANN), adaptive neural-fuzzy inference system (ANFIS) and Wavelet-ANN models were tested in order to assess their ability to forecast the meteorological drought for one, two, and three time steps (6 months) ahead. It was demonstrated that wavelet transform can improve meteorological drought modeling. It was also shown that ANFIS models provided more accurate predictions than ANN models. This study confirmed that the optimum number of neurons in the hidden layer could not be always determined using specific formulas; hence, it should be determined using a trial-and-error method. Also, decomposition level in wavelet transform should be delineated according to the periodicity

---

B. Shirmohammadi · H. Moradi (✉) · V. Moosavi · M. T. Semiromi · A. Zeinali  
Department of Watershed Management Engineering, Faculty of Natural Resources,  
Tarbiat Modares University, Noor 46417-76489, Iran  
e-mail: moosavi\_v66@yahoo.com; hrmoradi@modares.ac.ir

B. Shirmohammadi  
e-mail: b.shirmohammadi@gmail.com

V. Moosavi  
e-mail: Vahid\_mou2006@yahoo.com

M. T. Semiromi  
e-mail: majid.taie@gmail.com

A. Zeinali  
e-mail: ali\_zeinali631@yahoo.com

and seasonality of data series. The order of models with regard to their accuracy is as following: Wavelet-ANFIS, Wavelet-ANN, ANFIS, and ANN, respectively. To the best of our knowledge, no research has been published that explores coupling wavelet analysis with ANFIS for meteorological drought and no research has tested the efficiency of these models to forecast the meteorological drought in different time scales as of yet.

**Keywords** Natural hazards · Drought · Forecasting · Wavelet-ANN model · East Azerbaijan province

## 1 Introduction

Drought is a major natural hazard having severe consequences in regions all over the world. The range of drought impacts is related to drought occurring in different stages of the hydrologic cycle, and usually, different types of droughts are distinguished. The origin is a meteorological drought, which is defined as a deficit in precipitation. A meteorological drought can develop into a soil moisture drought, which may reduce agricultural production and increase the probability of forest fires. It can further develop into a hydrologic drought defined as a deficit in surface water and groundwater, for example, reducing water supply for drinking water, irrigation, industrial needs, and hydropower production, causing death of fish and hampering navigation in some countries (Yahiaoui et al. 2009). Temperatures, high winds, low relative humidity, timing and characteristics of rains, including distribution of rainy days during crop growing seasons, intensity and duration of rain, and onset and termination play a significant role in the occurrence of droughts (Mishra and Singh 2010). Natural hazards, such as drought, tsunami, hurricane, flood, wildfire, and earthquake, are likely to become ever more costly in human lives and economic development (World Bank Independent Evaluation Group 2006). The lessons learned from the 2004 Tsunami in the Indian Ocean and 2005 Hurricane Katrina in the southern United States indicate an urgent need to develop new means for early warning and response systems to enhance human collaborative capabilities in coping with large-scale natural hazards (UNDP 2005; The White House 2006). While modern remote sensing, spatial modeling, and geographic information technologies help users to detect, simulate, and forecast environmental changes, such technology is not yet well integrated with multilevel social cooperative responses. The new techniques such as artificial neural networks (ANN), fuzzy logic (FL), and ANFIS have been recently accepted as an efficient alternative tool for modeling of complex hydrologic systems and widely used for forecasting. For instance, some special applications of ANN in hydrology including modeling rainfall-runoff process (Jeong and Kim 2005; Kumar et al. 2005), hydrologic time series modeling (Jain and Kumar 2007), sediment concentration estimation (Nagy et al. 2002), estimation of heterogeneous aquifer parameters (Mantoglou 2003), runoff, and sediment yield modeling (Agarwal et al. 2006). Morid et al. (2007) examined the effectiveness of ANN approach for medium and long-term forecasting of both the probability of drought events and their severity.

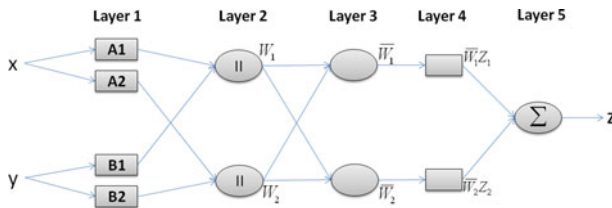
The standardized precipitation index (SPI) is a tool which was developed principally for defining and monitoring drought. It allows a researcher to determine a drought at a given time scale (temporal resolution) of interest for any rainfall station with historic data. The SPI is not a drought prediction tool. McKee et al. (1993) developed the SPI to calculate the

precipitation shortage for multiple time scales, reflecting the impact of precipitation deficiency on the accessibility of different water supplies. The SPI provides a quick and useful approach to drought analysis. Other advantages of this approach are its relative ease and minimal data requirements. Mishra and Desai (2006) applied the feed-forward recursive neural network and ARIMA models for drought forecasting using SPI series as a drought index. The results have demonstrated that neural network method can be successfully applied for drought forecasting. Wu et al. (2008) applied the neural network method to establish a risk evaluation model of heavy snow disaster using back-propagation artificial neural network (BP-ANN). On the other hand, several studies have also been carried out using FL in hydrology and water resources planning (Nayak et al. 2005; Altunkaynak et al. 2005). In recent years, adaptive neuro-fuzzy inference system (ANFIS), which is integration of ANN and FL methods, has been used in the modeling of nonlinear engineering and water resources problems (Sen and Altunkaynak 2006; Firat and Gungor 2007, 2008). Moreover, Chou and Chen (2007) have used the neuro-fuzzy computing technique for the development of drought early warning index. For this aim, an approach has been proposed to develop drought early warning index (DEWI) for Southern Taiwan to detect the drought in advance for setting up proper plans to mitigate the water shortage impacts. Hana et al. (2010) developed the AR (1) model for VTCI series in order to drought forecasting in the Guanzhong Plain. Several other studies were performed on drought indices, especially SPI such as Moreira et al. (2006, 2008), Canon et al. (2007), Mishra et al. (2007), Duggins et al. (2010), Kao and Govindaraju (2010), and Dupuis (2010). The main objective of this research is to develop Wavelet-ANFIS and Wavelet-ANN models to forecast 6-month SPI for different prediction periods. In this research, Wavelet-ANFIS and Wavelet-ANN models were developed and compared with ANFIS and ANN models in the Ajabshir Plain, Iran.

## 2 The adaptive neuro-fuzzy inference system (ANFIS)

The ANFIS is a universal estimator and is able to approximate any real continuous function on a compact set to any degree of accuracy (Jang et al. 1997). The basic structure of the type of fuzzy inference system could be seen as a model that maps input characteristics to input membership functions. Then, it maps input membership function to rules and rules to a set of output characteristics. Finally, it maps output characteristics to output membership functions and the output membership function to a single-valued output or a decision associated with the output (Jang et al. 1997). Each fuzzy system contains three main parts including fuzzifier, fuzzy database, and defuzzifier. Also, Fuzzy database consists of two main parts containing fuzzy rule base, and inference engine.

Figure 1 represents a typical ANFIS architecture. In layer one, every node is an adaptive node with a node function such as a generalized bell membership function or a Gaussian membership function. In layer two, every node is a fixed node representing the firing strength of each rule and is calculated by the fuzzy and connective of the ‘product’ of the incoming signals. In layer three, every node is a fixed node showing the normalized firing strength of each rule. The *i*th node calculates the ratio of the *i*th rule’s firing strength to the summation of two rules firing strengths. In layer four, every node is an adaptive node with a node function indicating the contribution of *i*th rule toward the overall output. In layer five, the single node is a fixed node indicating the overall output as the summation of all incoming signals (Jang and Sun 1995).



**Fig. 1** A typical ANFIS architecture (Jang 1993). Here,  $x$  and  $y$  are the inputs and  $z$  is the final output; A1, A2, B1, and B2 are the linguistic label (small, large, etc.) associated with this node function, and  $w_i$  is the normalized firing strength that is the ratio of the  $i$ th rule's firing strength ( $W_i$ ) to the summation of the first and second rules' firing strengths ( $W1$  and  $W2$ ) and  $\Pi$  is the node label

### 3 Artificial neural network (ANN)

ANNs are powerful nonlinear modeling approaches based on the function of the human brain. They can identify and learn correlated patterns between input data sets and target values. Neural networks can be described as a network of simple processing nodes or neurons, interconnected to each other in a specific order, performing simple numerical manipulations (See and Openshaw 1999). A common three-layered neural network is composed of several elements namely nodes. These networks are made up of an input layer consisting of nodes representing different input variables, the hidden layer consisting of many hidden nodes, and an output layer consisting of output variables (Haykin 1999).

Feed-forward neural networks are successfully applied in many different problems. This network architecture and the corresponding learning algorithm can be viewed as a generalization of the popular least-mean-square (LMS) algorithm (Haykin 1999). Fully recurrent networks, introduced by Elman (1988), feed the outputs of the hidden layer back to itself. Partially recurrent networks start with a fully recurrent net and add a feed-forward connection that bypasses the recurrence, effectively treating the recurrent part as a state memory. Recurrent networks are the state of the art in nonlinear time series prediction, system identification, and temporal pattern classification.

### 4 Wavelet transform

Wavelet transform is a time-dependent spectral analysis that decomposes time series in the time–frequency space and provides a time-scale illustration of processes and their relationships (Daubechies 1990). In this method, the data series are broken down by the transformation into its ‘wavelets’, a scaled and shifted version of the mother wavelet (Grossman and Morlet 1984). Wavelet analysis allows the use of long-time intervals for low-frequency information and shorter intervals for high-frequency information and is capable of revealing aspects of data like trends, breakdown points, and discontinuities that other signal analysis techniques might miss.

The discrete wavelet transform (DWT) is used to decompose the time series. The classical continuous wavelet transform (CWT) requires a significant amount of computation time and data (Partal 2009; Adamowski 2007). So, in this study, the discrete wavelet transform is preferred because it requires less computation time and data (Christopoulou et al. 2002). The DWT operates two sets of functions: high-pass and low-pass filters. The original time series is passed through high-pass and low-pass filters, and detailed

coefficients and approximation series are obtained. In fact, it decomposes the signal into a group of functions (Cohen and Kovacevic 1996):

$$\psi_{j,k}(x) = 2^{j/2} \psi_{j,k}(2^j x - k) \tag{1}$$

where  $\psi_{j,k}(x)$  is produced from a mother wavelet  $\psi(x)$  which is expanded by  $j$  and translated by  $k$ . The discrete wavelet function of a signal  $f(x)$  can be calculated as follows:

$$c_{j,k} = \int_{-\infty}^{\infty} f(x) \psi_{j,k}^*(x) dx$$

$$f(x) = \sum_{j,k} c_{j,k} \psi_{j,k}(x) \tag{2}$$

where  $c_{j,k}$  is the approximate coefficient of a signal. The mother wavelet is formulated from the scaling function  $\varphi(x)$  as:

$$\varphi(x) = \sqrt{2} \sum h_0(n) \varphi(2x - n)$$

$$\psi(x) = \sqrt{2} \sum h_1(n) \varphi(2x - n) \tag{3}$$

$$h_1(n) = (-1)^n h_0(1 - n)$$

Different sets of coefficients  $h_0(n)$  can be found corresponding to wavelet bases with various characteristics. In the DWT, coefficients  $h_0(n)$  play a critical role (Rajaei 2011). In summary, although, several studies have been conducted on application of ANN and FL techniques to find out how accurately they can predict weather and climate processes, few researches have integrated ANN and FL as a new method so-called ANFIS so that previous studies have shown that ANFIS model has more advantages and less weaknesses than other computational intelligence techniques. So, the present study is programmed to evaluate the capability of this method for meteorological drought forecasting in the part of East Azerbaijan Province, Iran.

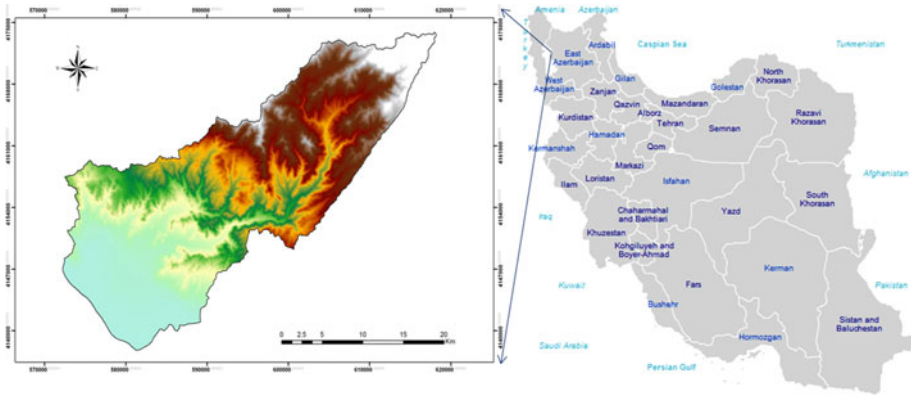
## 5 Materials and methods

### 5.1 Study area

The study site located in Tabriz, East Azerbaijan province, Iran, lies from 45 30' to 46 30' E longitude and from 37 to 38 N latitude (Fig. 2). The average of yearly precipitation and temperature of Ajabshir plain is 272.3 mm and 13.3 °C, respectively. In recent years, several drought events have caused many agricultural and social loses on account of most of people here are strongly dependent on agricultural productions. The climate of the study area is Mediterranean and most of rainfall occurs in half of year, and there are two wet and dry periods in each year. This is the reason that the 6-month SPI is used in this study.

### 5.2 SPI

The time series data from 1952 to 2011 were used in this study (the first 40 years for training and the last 20 years for verification). The SPI index was calculated using the



**Fig. 2** Location of the study area in East Azerbaijan province, Iran

precipitation data with the DIP (Drought Indices Package) software based on the following equation:

In most cases, the distribution that best models observational precipitation data is the Gamma distribution. The density probability function for the gamma distribution is given by the expression:

$$g(x) = \frac{1}{\beta^\alpha \Gamma(\alpha)} x^{\alpha-1} e^{-\frac{x}{\beta}} \quad \text{for } x > 0 \tag{4}$$

where  $\alpha > 0$  is the shape parameter,  $\beta > 0$  is the scale parameter, and  $x > 0$  is the amount of precipitation.  $\Gamma(\alpha)$  is the value taken by the standard mathematical function known as the Gamma function defined by the integral:

$$\Gamma(\alpha) = \lim_{n \rightarrow \infty} \prod_{v=0}^{n-1} \frac{n! n^{y-1}}{y+v} = \int_0^\infty y^{\alpha-1} e^{-y} dy \tag{5}$$

In general, the gamma function is evaluated either numerically or using the values tabulated depending on the value taken by parameter  $\alpha$ . In order to model the data observed with a gamma distributed density function, it is necessary to estimate appropriately parameters  $\alpha$  and  $\beta$ . Different methods have been suggested in the literature for the estimate of these parameters, for example, in Edwards and McKee (1997), the Thom (1958) approximation is used for the maximum probability.

$$\hat{\alpha} = \frac{1}{4A} \left( 1 + \sqrt{1 + \frac{4A}{3}} \right) \tag{6}$$

$$\hat{\beta} = \frac{\bar{x}}{\hat{\alpha}} \tag{7}$$

where for  $n$  observations

$$A = \ln(\bar{x}) - \frac{\sum \ln(x)}{n} \tag{8}$$

The estimate of the parameters can be further improved by using the interactive approach suggested in Wilks (1995). After estimating coefficients  $\alpha$  and  $\beta$ , the density of

probability function  $g(x)$  is integrated with respect to  $x$  and we obtain an expression for cumulative probability  $G(x)$  that a certain amount of rain has been observed for a given month and for a specific time scale.

$$G(x) = \int_0^x g(x)dx = \frac{1}{\hat{\beta}^{\hat{\alpha}} \Gamma(\hat{\alpha})} \int_0^x x^{\hat{\alpha}-1} e^{-x/\hat{\beta}} dx \tag{9}$$

Taking  $t = x/\hat{\beta}$ , this equation becomes the incomplete gamma function

$$G(x) = \frac{1}{\Gamma(\hat{\alpha})} \int_0^x t^{\hat{\alpha}-1} e^{-t} dt \tag{10}$$

The gamma function is not defined by  $x = 0$ , and since there may be no precipitation, the cumulative probability becomes:

$$H(x) = q + (1 - q)G(x) \tag{11}$$

where  $q$  is the probability of no precipitation. The cumulative probability is then transformed into a normal standardized distribution with null average and unit variance from which we obtain the SPI index. For the details, see Edwards and McKee (1997) or Lloyd-Hughes and Saunders (2002). The above approach, however, is neither practical nor numerically simple to use if there are many grid points or many stations on which to calculate the SPI index. In this case, an alternative method was described in Edwards and McKee (1997) using the technique of approximate conversion developed in Abramowitz and Stegun (1965) that converts the cumulative probability into a standard variable  $Z$ .

The SPI index is then defined as:

$$Z = \text{SPI} = - \left( t - \frac{c_0 + c_1 t + c_2 t^2}{1 + d_1 t + d_2 t^2 + d_3 t^3} \right) \quad \text{for } 0 < H(x) < 0.5 \tag{12}$$

$$Z = \text{SPI} = + \left( t - \frac{c_0 + c_1 t + c_2 t^2}{1 + d_1 t + d_2 t^2 + d_3 t^3} \right) \quad \text{for } 0.5 < H(x) < 1 \tag{13}$$

$$t = \sqrt{\ln \left[ \frac{1}{(H(x))^2} \right]} \quad \text{for } 0 < H(x) < 0.5 \tag{14}$$

$$t = \sqrt{\ln \left[ \frac{1}{(1 - H(x))^2} \right]} \quad \text{for } 0.5 < H(x) < 1 \tag{15}$$

where  $x$  is precipitation,  $H(x)$  is the cumulative probability of precipitation observed, and  $c^0, c^1, c^2, d^0, d^1, d^2$  are constants with the following values:

$$c^0 = 2.515517, c^1 = 0.802853, c^3 = 0.010328$$

$$d^0 = 1.432788, d^1 = 0.189269, d^3 = 0.001308$$

### 5.3 ANN models

The ANN models for meteorological drought forecasting were developed using MATLAB R2010 software. In this paper, two networks were constructed including feed-forward

neural network (FNN) and Elman or recurrent neural network (RNN). Also, 2, 3, 4, 5, 6, 7, 8, 9, and 10 nodes were tested as the number of nodes in the hidden layer. The numbers of nodes were selected using a common trial-and-error approach. There are several training algorithms for ANNs such as gradient descent with momentum and adaptive learning rate back-propagation (GDX), Levenberg–Marquardt (LM), and Bayesian regularization (BR) (as were used in this research). GDX uses back-propagation to calculate derivatives of performance cost function with respect to the weight and bias variables of the network. Each variable is adjusted according to the gradient descent with momentum. For each step of the optimization, if performance decreases the learning rate is increased. This is probably the simplest and most common way to train a network (Haykin 1999). Similarly, the LM method is a modification of the classic Newton algorithm for finding an optimum solution to a minimization problem. The BR is an algorithm that automatically sets optimum values for the parameters of the objective function. In the approach used, the weights and biases of the network are assumed to be random variables with specified distributions. In order to estimate regularization parameters, which are related to the unknown variances, statistical techniques are being used. The advantage of this algorithm is that whatever the size of the network, the function will not be over-fitted (Maier and Dandy 1998). The SPI data for the current time were imported as inputs, and the SPI for one, two, and three time steps (6 months) ahead were considered as target.

#### 5.4 ANFIS models

This model was trained and tested using the data set same as those were used in ANN models. Four different membership functions (MFs) were tested for ANFIS models in this work, that is, Gaussian (MFgauss), bell-shaped (MFgbell), triangular (MFtri), and spline-based (MFspi), or Piduetoits shape. ANFIS models with different types of MF were run with 2, 3, 4, and 5 MFs and with 50, 100, 150, 200, 250, 300, and 400 iterations for each node of input data (Rajae 2011). Training characteristics of the ANFIS model are presented in Table 1. The SPI data for the current time were imported as inputs, and the SPI for on, tow, and three time steps (12 months) ahead were considered as target.

#### 5.5 Wavelet-ANN and Wavelet-ANFIS

In order to build hybrid Wavelet-ANN and Wavelet-ANFIS model, sub-series elements which are derived from the use of the discrete wavelet transform on the original time series data were used as inputs for neural network models. Each sub-series element plays a unique role in the original time series, and the performance of each sub-series is distinct. In the first step, the original SPI data were decomposed into a series of details using discrete wavelet transform. The decomposition process was iterated with successive approximation signals being decomposed in turn, so that the original time series was broken down into

**Table 1** Training parameters of the ANFIS models

Parameters	Type
AND method	Prod
Or method	Maximum
Imp. method	Prod
Aggr. method	Maximum
Defuzzification method	Wtaver



many lower resolution components (Adamowski and Fung Chan 2011). All of the mentioned variables were decomposed to 1, 2, 3, and 4 levels by seven different kinds of wavelets, that is, db4, bior1.1, bior1.5, rboi1.1, rboi1.5, coif2, and coif4 wavelets.

### 5.6 Performance comparison of models

Coefficient of determination ( $R^2$ ) and root mean squared error (RMSE) were used to compare the performance of models and select the best one.

$$R^2 = 1 - \frac{\sum_{i=1}^N (y_o - y_e)^2}{\sum_{i=1}^N (y_o - \bar{y}_o)^2} \tag{16}$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^N (y_o - y_e)^2}{N}} \tag{17}$$

$$E = 1 - \frac{\sum_{i=1}^n (y_o - y_e)^2}{\sum_{i=1}^n (y_o - \bar{y}_o)^2} \tag{18}$$

where  $y_o$ ,  $y_e$ , and  $N$  are observed and estimated SPI, and number of data, respectively. In the Nash–Sutcliffe model efficiency coefficient, an efficiency of one corresponds to a perfect match of forecasted data to the observed data. An efficiency of zero indicates that the model predictions are as accurate as the mean of the observed data (Pulido-Calvo and Gutierrez-Estrada 2009).

## 6 Results and discussion

### 6.1 ANN models

In this study, several different networks were tested. Table 2 shows the  $R^2$  and RMSE for different ANN models. The best number of neurons in the hidden layer, that is, was determined 4 neurons and for the different time step predictions. As shown in this table, FNN-BR (1 4 1) was the best network. This result is in contrast to the suggestion that was expressed by Kavzoglu and Mather (2003). They suggest that the number of neurons in hidden layer should be set as twice as the input nodes plus one. It confirms that the number of neurons in hidden layer should be determined using a trial-and-error method. All models forecasted the meteorological drought for one and two time steps ahead appropriately, but for 3 and 4 months ahead, the  $R^2$  and Nash–Sutcliffe model efficiency coefficient decreased and RMSE increased meaningfully. In other words, more time step duration less accuracy is expected (Table 2).

### 6.2 ANFIS models

There is not any basic rule to delineate the number of membership functions (MFs) of ANFIS models, and they usually determined by trial-and-error approach. To select the number of MFs, a modeler should avoid using a large number of membership functions or parameters to save time and calculation effort (Keskin et al. 2004). Four different types of membership function (MF) were tested in this study including Gaussian (MFgauss), Bell-shaped (MFg-bell), Triangular (MFtri), and Spline-based (MFspi), or Piduetoits shape (Jang 1993). ANFIS

**Table 2**  $R^2$  and RMSE for different ANN models with the best number of neurons in the hidden layer (4 neurons) and for the different predicted time steps (verification data set)

Training algorithm	Transfer function in the hidden layer	Predicted time steps								
		$SPI_{t+1}$			$SPI_{t+2}$			$SPI_{t+3}$		
		RMSE	$R^2$	$E$	RMSE	$R^2$	$E$	RMSE	$R^2$	$E$
TRAINLM	Tansig	1.81	0.89	0.80	1.91	0.85	0.80	3.02	0.78	0.71
TRAINSCG	Logsig	2.2	0.83	0.79	2.41	0.81	0.77	3.59	0.76	0.65
TRAINRP	Tansig	2.1	0.81	0.76	2.2	0.80	0.75	3.34	0.76	0.67
TRAINGDx	Logsig	2.17	0.84	0.78	2.29	0.83	0.76	3.46	0.77	0.66

**Table 3** Results of ANFIS models with the best number of membership functions (2 MF), the best iteration (150), and different predicted time steps (verification data set)

Membership function	Time step prediction								
	$SPI_{t+1}$			$SPI_{t+2}$			$SPI_{t+3}$		
	RMSE	$R^2$	$E$	RMSE	$R^2$	$E$	RMSE	$R^2$	$E$
Gauss	0.089	0.90	0.79	0.092	0.89	0.77	0.094	0.89	0.69
Gbell	0.090	0.89	0.77	0.091	0.88	0.76	0.093	0.88	0.67
<b>pi</b>	<b>0.089</b>	<b>0.91</b>	<b>0.83</b>	<b>0.091</b>	<b>0.89</b>	<b>0.82</b>	<b>0.092</b>	<b>0.90</b>	<b>0.73</b>
Triangular	0.091	0.89	0.75	0.092	0.88	0.75	0.092	0.88	0.68

**Table 4** Results of Wavelet-ANN models for the different predicted time steps (verification data set)

Network	Time step prediction								
	$L_{t+1}$			$L_{t+2}$			$L_{t+3}$		
	$R^2$	RMSE	$E$	$R^2$	RMSE	$E$	$R^2$	RMSE	$E$
db 4	0.94	0.303	0.86	0.95	0.373	0.84	0.94	1.027	0.74
Bior 1.1	0.92	0.32	0.83	0.91	0.36	0.83	0.9	0.923	0.75
Bior 1.5	0.93	0.303	0.84	0.89	0.343	0.86	0.87	1.087	0.75
Rboi 1.1	0.94	0.347	0.84	0.91	0.363	0.83	0.90	0.953	0.74
<b>Rboi 1.5</b>	<b>0.95</b>	<b>0.227</b>	<b>0.89</b>	<b>0.93</b>	<b>0.233</b>	<b>0.88</b>	<b>0.92</b>	<b>0.657</b>	<b>0.76</b>
Coif 2	0.93	0.363	0.86	0.93	0.393	0.80	0.91	0.837	0.74
Coif 4	0.92	0.383	0.84	0.94	0.383	0.83	0.92	0.880	0.75

models with different types of MF were run with 2, 3, 4, and 5 MFs and with 50, 100, 150, 200, 250, 300, and 400 iterations for each node of input data (Table 3).

In general, ANFIS performs more efficiently than ANN models on account of the effect of fuzzification of the input through membership functions.

### 6.3 Wavelet-ANN and Wavelet-ANFIS models

In this part, Wavelet-ANN and Wavelet-ANFIS models were tested. The best ANN model was selected to make hybrid Wavelet-ANN models. The data were grouped into 1, 2, 3,

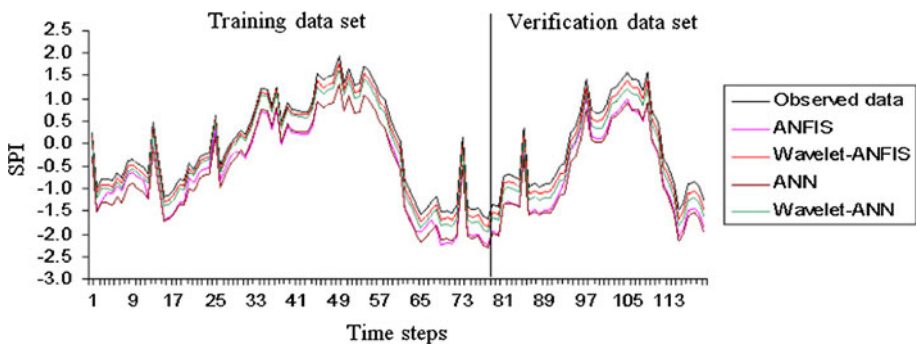
**Table 5** Results of Wavelet-ANFIS models for the different predicted time steps (verification data set)

Network	Predicted time steps								
	$L_t + 1$			$L_t + 2$			$L_t + 3$		
	$R^2$	RMSE	$E$	$R^2$	RMSE	$E$	$R^2$	RMSE	$E$
db 4	0.96	0.10	0.88	0.94	0.173	0.87	0.92	0.23	0.78
Bior 1.1	0.94	0.12	0.87	0.92	0.16	0.85	0.90	0.22	0.79
Bior 1.5	0.95	0.10	0.87	0.93	0.143	0.87	0.91	1.24	0.79
Rboi 1.1	0.97	0.14	0.90	0.95	0.163	0.88	0.92	0.25	0.80
<b>Rboi 1.5</b>	<b>0.98</b>	<b>0.097</b>	<b>0.93</b>	<b>0.96</b>	<b>0.10</b>	<b>0.91</b>	<b>0.95</b>	<b>0.19</b>	<b>0.86</b>
Coif 2	0.95	0.16	0.89	0.94	0.19	0.88	0.92	0.26	0.78
Coif 4	0.94	0.13	0.90	0.93	0.18	0.89	0.92	0.27	0.79

and 4 levels. The best Wavelet-ANN models were found to forecast SPI more accurately than both chosen ANN and ANFIS models.

Table 4 shows the results of Wavelet-ANN models with the best network (i.e. FNN-BR) and for different predicted time steps. Moreover, Table 5 shows the results of Wavelet-ANFIS models with the best number and type of MF (i.e. 4 membership function and bell-shape membership function) and for different predicted time steps.

According to the monthly time-scale resolution, in the subgroup level 2, there are 2 subsets ( $2^1$  month mode which is nearly yearly mode and  $2^2$  month). In fact, the duration and time interval of data series should be considered in order to determine the appropriate number of decomposition levels. This result is in accordance with Nourani et al. (2011) who evaluated some artificial intelligence methods to model rainfall-runoff process. Generally, Wavelet-ANFIS and Wavelet-ANN models were found to provide more accurately predicted meteorological drought than ANN and ANFIS models. Also, it was reached that Wavelet-ANFIS hybrid model is able to forecast SPI more accurately than Wavelet-ANN. Since Wavelet-ANFIS models consisted of ANN, FL, and wavelet transform, they could model SPI more accurately. Figure 3 shows the best ANN, ANFIS, Wavelet-ANN, and Wavelet-ANFIS models that could predict SPI in a good agreement with observed ones for both subbasins. As graphically illustrated by this figure, all models often underestimated SPI. The order of models according to their accuracy is as following: Wavelet-ANFIS, Wavelet-ANN, ANFIS, and ANN, respectively.



**Fig. 3** Observed and simulated SPI using selected models

## 7 Conclusions

The capability of coupled Wavelet-ANFIS models in comparison with ANN, ANFIS, and Wavelet-ANN models for one, two and three months ahead of forecasted SPI was assessed in this study. Wholly, mentioned time series are characterized by high nonlinearity, non-stationary, and seasonality behavior. In this study, the wavelet transform, ANN, and ANFIS approaches were combined in order to develop two hybrid models to forecast SPI for different time steps. At first, ANN and ANFIS models were used without any pre-processing. Results showed that these models may be unable to cope with the nonlinearity and seasonality behavior of data. In the second step, wavelet transform was performed on the data and then, the pre-processed data were used as input for ANN and ANFIS models. This research demonstrated that pre-processed data can improve SPI forecasting. Also, results showed that Wavelet-ANFIS hybrid model had the best performance. Additionally, it was concluded that all modeling approaches are capable nearly of forecasting SPI accurately.

**Acknowledgments** The authors are grateful to the East Azerbaijan regional water authority for their constructive suggestions and data that have improved the scientific quality of this paper.

## References

- Abramowitz M, Stegun A (1965) Handbook of mathematical formulas, graphs, and mathematical tables. Dover Publications, Inc, New York
- Adamowski J (2007) Development of a short-term river flood forecasting method based on wavelet analysis. Polish Academy of Sciences Publication, Warsaw, p 172
- Adamowski J, Fung Chan H (2011) A wavelet neural network conjunction model for SPI index forecasting. *J Hydrol* 407:28–40
- Agarwal A, Mishra SK, Ram S, Singh JK (2006) Simulation of runoff And sediment yield using artificial neural networks. *Biosyst Eng* 94(4):597–613
- Altunkaynak A, Ozger M, Cakmakci M (2005) Water consumption prediction of Istanbul City by using fuzzy logic approach. *Water Resour Manag* 19:641–654
- Canon J, Gonzalez J, Valdes J (2007) Precipitation in the Colorado River Basin and its low frequency associations with PDO and ENSO signals. *J Hydrol* 333:252–264
- Chou FNF, Chen BPT (2007) Development of drought early warning-index: using neuro-fuzzy computing technique. In: 8th international symposium on advanced intelligence systems 2007 Korea. Paper No: A 1469
- Christopoulou EB, Skodras AN, Georgakilas AA (2002) The “Trous” wavelet transform versus classical methods for the improvement of solar images. In: Proceedings of 14th international conference on digital signal processing 2:885–888
- Cohen A, Kovacevic J (1996) Wavelets: the mathematical background. *Proc IEEE* 84:514–522
- Daubechies I (1990) The wavelet transform, time-frequency localization and signal analysis. *IEEE Trans Inf Theory* 36:961–1005
- Duggins J, Williams M, Kim DY, Smith E (2010) Change-point detection in SPI transition probabilities. *J Hydrol* 388:456–463
- Dupuis D (2010) Statistical modeling of the monthly palmer drought severity index. *J Hydrol Eng* 15(10):796–807
- Edwards DC, McKee TB (1997) Characteristic of 20th century drought in the United States at multiple timescales. Colorado State University: Fort Collins, Climatology Report
- Elman JL (1988) Finding structure in time. CRL Technical Report 8801. Centre for Research in Language, University of California at San Diego
- Firat M, Gungor M (2007) River flow estimation using adaptive neuro-fuzzy inference system. *Math Comput Simul* 75(3–4):87–96
- Firat M, Gungor M (2008) Hydrological time-series modeling using adaptive neuro-fuzzy inference system. *Hydrol Process* 22(13):2122–2132

- Grossman A, Morlet J (1984) Decompositions of hardy functions into square integrable wavelets of constant shape. *SIAM J Math Anal* 15:723–736
- Hana P, Wang PX, Zhang ShY, Zhu DH (2010) Drought forecasting based on the remote sensing data using ARIMA models. *Math Comput Model* 51:1398–1403
- Haykin S (1999) *Neural networks, a comprehensive foundation*, 2nd edn. Prentice-Hall, Englewood Cliffs, NJ, pp 135–155
- Jain A, Kumar AM (2007) Hybrid neural network models for hydrologic time series forecasting. *Appl Soft Comput* 7:585–592
- Jang JSR (1993) ANFIS: adaptive network based fuzzy interface system. *Proc IEEE Trans Syst Man Cybern* 23:665–685
- Jang JSR, Sun CT (1995) Neuro-fuzzy modeling and control. *Proc IEEE* 83:378–406
- Jang JSR, Sun CT, Mizutani E (1997) *Neuro-fuzzy and soft computing: a computational approach to learning and machine intelligence*. Prentice-Hall, Eaglewood cliffs, NJ, pp 665–685
- Jeong D, Kim YO (2005) Rainfall-run off models using artificial neural networks for ensemble stream flow prediction. *Hydrol Process* 19:3819–3835
- Kao SC, Govindaraju RS (2010) A copula-based joint deficit index for droughts. *J Hydrol* 380:121–134
- Kavzoglu T, Mather PM (2003) The use of back-propagating artificial neural networks in land cover classification. *Int J Remote Sens* 24:4907–4938
- Keskin ME, Terzi O, Taylan D (2004) Fuzzy logic model approaches to daily pan evaporation estimation: in western Turkey. *Hydrol Sci* 49:1001–1010
- Kumar ARS, Sudheer KP, Jain SK, Agarwal PK (2005) Rainfall-runoff modeling using artificial neural networks: comparison of network types. *Hydrol Process* 19:1277–1291
- Lloyd-Hughes B, Saunders MA (2002) A drought climatology for Europe. *Int J Clim* 22:1571–1592
- Maier HR, Dandy GC (1998) The effect of internal parameters and geometry on the performance of back-propagation neural networks: an empirical study. *Environ Model Softw* 13:193–209
- Mantoglou A (2003) Estimation of heterogeneous aquifer parameters from piezometric data using ridge functions and neural networks. *Stoch Environ Res Risk Assess* 17:339–352
- McKee TB, Doesken NJ, Kleist J (1993) The relationship of drought frequency and duration to time scales. In: *Preprints, 8th conference on applied climatology*. 17–22 Jan, Anaheim, California, pp 179–184
- Mishra AK, Desai VR (2006) Drought forecasting using feed-forward recursive neural network. *Ecol Model* 198:127–138
- Mishra AK, Singh VP (2010) A review of drought concepts. *J Hydrol* 391:202–216
- Mishra A, Desai V, Singh V (2007) Drought forecasting using a hybrid stochastic and neural network model. *J Hydrol Eng* 12(6):626–638
- Moreira EE, Paulo AA, Pereira LS, Mexia JT (2006) Analysis of SPI drought class transitions using loglinear models. *J Hydrol* 331:349–359
- Moreira EE, Coelho CA, Paulo AA, Pereira LS, Mexia JT (2008) SPI-based drought category prediction using loglinear models. *J Hydrol* 354:116–130
- Morid S, Smakhtin V, Bagherzadeh K (2007) Drought forecasting using artificial neural networks and time series of drought indices. *Int J Climatol* 27:2103–2111
- Nagy HM, Watanabe K, Hirano M (2002) Prediction of sediment load Concentration in rivers using artificial neural network model. *J Hydraul Eng* 128:588–595
- Nayak PC, Sudheer KP, Ramasastri KS (2005) Fuzzy computing Based rainfall-runoff model for real time flood forecasting. *Hydrol Process* 19:955–968
- Nourani V, Kisi Z, Mehdi K (2011) Two hybrid artificial intelligence approaches for modeling rainfall-runoff process. *J Hydrol* 402:41–59
- Partal T (2009) River flow forecasting using different artificial neural network algorithms and wavelet transform. *Can J Civil Eng* 36:26–38
- Pulido-Calvo I, Gutierrez-Estrada JC (2009) Improved irrigation water demand forecasting using a soft-computing hybrid model. *Biosyst Eng* 102(2):202–218
- Rajae T (2011) Wavelet and ANN combination model for prediction of daily suspended sediment load in rivers. *Sci Total Environ* 409:2917–2928
- See L, Openshaw S (1999) Applying soft computing approaches to river level forecasting. *Hydrol Sci J des Sci Hydrol* 44:763–777
- Sen Z, Altunkaynak A (2006) Prediction of Istanbul City by using fuzzy logic approach. *Water Runoff coefficient and runoff estimation*. *Hydrol Process* 20:1993–2009
- The White House (2006) The federal response to Hurricane Katrina: lessons learned. <http://www.whitehouse.gov/reports/katrina-lessons-learned/>
- Thom HCS (1958) A note on the gamma distribution. *Month Wat Rev* 86(4):117–122

- UNDP (2005) Reducing risks from tsunamis: disaster and development, bureau for crisis prevention and recovery, United Nations Development Programme. <http://www.undp.org/cpr/disred/documents/tsunami/undp/rdrtsunamis.pdf>
- Wilks DS (1995) Statistical methods in the atmospheric sciences. An introduction. Academic Press, London
- World Bank Independent Evaluation Group (2006) Hazards of nature, risks to development. World Bank Publications, Washington, DC
- Wu JD, Li N, Yang HJ, Li CH (2008) Risk evaluation of heavy snow disasters using BP artificial neural network: the case of Xilingol in Inner Mongolia. *Stoch Environ Res Risk Assess* 22:719–725. doi: [10.1007/s00477-007-0181-7](https://doi.org/10.1007/s00477-007-0181-7)
- Yahiaoui A, Touaïbia B, Bouvier C (2009) Frequency analysis of the hydrological drought regime. Case of oued Mina catchment in western of Algeria, *Revue Nature et Technologie*, pp 3–15