JOURNAL OF
ENVIRONMENTAL HYDROLOGY

The Electronic Journal of the International Association for Environmental Hydrology
On the World Wide Web at http://www.hydroweb.com

VOLUME 16 2008

A COMBINED NEURAL-WAVELET MODEL FOR PREDICTION OF WATERSHED PRECIPITATION, LIGVANCHAI, IRAN

Vahid Nourani
Mohammad Taghi Alami
Mohammad Hossein Aminfar

Faculty of Civil Engineering
University of Tabriz
Tabriz, Iran

The first step in a river management program is precipitation modeling over the watershed. Considering the high stochastic property of the process, many models are still being developed to define this complex phenomenon. The Artificial Neural Network (ANN), a non-linear inter-extrapolator, is extensively used by hydrologists for rainfall modeling as well as in other fields of hydrology. In this research, wavelet analysis was linked to the ANN concept for prediction of Ligvanchai watershed precipitation at Tabriz, Iran. The main time series was decomposed to some multi-frequency time series by wavelet theory, then these time series were imposed as input data to the ANN to predict precipitation one month ahead. The results show the proposed model can predict both short and long term precipitation events by using multi-scale time series as the ANN input layer.

INTRODUCTION

The true prediction of hydrological signals such as precipitation can give effective information for city planning, land use, flood and water resource management for a watershed. It also plays an important role in the mitigation of impacts of drought on water resources systems.

Classic time series models such as Auto Regressive Moving Average (ARMA) are widely used for hydrological time series forecasting (Salas et al., 1980). However, they are basically linear models assuming that data are stationary, and have a limited ability to capture non–stationarities and non–linearities in hydrologic data.

Nowadays, the Artificial Neural Network (ANN), as a self–learning and self–adaptive function approximator, has shown great ability in modeling and forecasting nonlinear hydrologic time series. ANNs offer an effective approach for handling large amounts of dynamic, non–linear and noisy data, especially when the underlying physical relationships are not fully understood. This makes them well suited to time series modeling problems of a data–driven nature.

In spite of suitable flexibility of ANN in modeling hydrologic time series, sometimes it is not adequate when signal fluctuations are highly non-stationary and physical hydrologic processes operate under a large range of scales varying from one day to several decades. In such a situation, ANNs may not be able to cope with non–stationary data if pre–processing of the input and/or output data is not performed (Cannas et al., 1980).

Recently, wavelet transform analysis has become a popular analysis tool due to its ability to elucidate simultaneously both spectral and temporal information within the signal. This overcomes the basic shortcoming of Fourier analysis, which is that the Fourier spectrum contains only globally averaged information. Therefore, a data pre–processing can be conducted by time series decomposition into its subcomponents using wavelet transform analysis. Wavelet transforms provide useful decompositions of main time series, so that wavelet–transformed data improve the ability of a forecasting model by capturing useful information on various resolution levels. Hence a hybrid ANN–wavelet model which uses multi–scale signals as input data may present more probable forecasting rather than a single pattern input.

The ANN–wavelet conjunction model was first presented by Aussem et al. (1998) for financial time series forecasting. Zhang and Dong (2001) proposed a short–term load forecast model based on ANN and the multi–resolution wavelet decomposed. In hydrology, Wang and Ding (2003) applied a wavelet–network model to forecast shallow groundwater level and daily discharge. Kim and Valdes (2003) proposed a conjunction model based on dyadic wavelet transforms and ANNs to forecast droughts for the Conches river basin in Mexico; they used ANN to forecast sub-signals from wavelet decomposition and also to reconstruct the main signal from the forecast sub–signals. In both cited researches the “a trous” algorithm for the discrete dyadic wavelet transform accompanied by three–layered feed forward neural networks was used in order to predict hydrological time series.

Cannas et al. (2006) investigated the effects of data pre–processing on the ANN model performance using continuous and discrete wavelet transforms. The results showed that networks trained with pre-processed data, performed better than networks trained on undecomposed, noisy raw signals.

In this paper the sensitivity of the pre-processing to the wavelet type and decomposition level is examined. In order to accomplish this objective, the monthly precipitation time series of the Ligvanchai river basin was decomposed into sub–signals at various resolution levels; then these sub-signals were entered into the ANN model to reconstruct the original forecast time series. Finally to evaluate the model ability, the proposed model was compared with the individual ANN and a classic time series model.

WAVELET TRANSFORM

The wavelet transform has increased in usage and popularity in recent years since its inception in the early 1980s, yet still does not enjoy the widespread usage of the Fourier transform.

In the field of earth sciences, Grossmann and Morlet (1984), who worked especially on geophysical seismic signals, introduced the wavelet transform application. A comprehensive literature survey of wavelets in geosciences can be found in Foufoula–Georgiou and Kumar (1995) and the most recent contributions are cited by Labat (2005). As there are many good books and articles introducing the wavelet transform, this paper will not delve into the theory behind wavelets and only the main concepts of the transform are briefly presented. Recommended literature for the wavelet novice includes Mallat (1998) or Labat et al. (2000).

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

$$T(a,b)=\frac{1}{\sqrt{a}}\int_{-\infty}^{+\infty} g^{*}\left(\frac{t-b}{a}\right)x(t).dt \tag{1}$$

Where * corresponds to the complex conjugate and *g(t)* is called the wavelet function or mother wavelet. The parameter *a* acts as a dilation factor, while *b* corresponds to a temporal translation of the function *g(t)*, which allows the study of the signal around *b*. The main property of the wavelet transform is to provide a time-scale localization of processes, which derives from the compact support of its basic function. This is opposed to the classical trigonometric function of the Fourier analysis. The wavelet transform searches for correlations between the signal and wavelet function. This calculation is done at different scales of *a* and locally around the time of *b*. The result is a wavelet coefficient (*T(a,b)*) contour map known as a scalogram. In order to be classified as a wavelet, a function must have finite energy, and it must satisfy the following "admissibility conditions":

$$\int_{-\infty}^{+\infty} g(t)dt=0\,,\;\; C_g=\int_{-\infty}^{+\infty}\frac{\left|\hat{g}(w)\right|^2}{|w|}dw<\infty \tag{2}$$

where $\hat{g}(w)$ is Fourier transform of g(t); i.e. the wavelet must have no zero frequency component.

In order to obtain a reconstruction formula for the studied signal, it is necessary to add "regularity conditions" to the previous ones:

$$\int_{-\infty}^{+\infty} t^k g(t)dt=0 \quad \text{where } k= 1,2,\ldots, n\text{-}1 \tag{3}$$

So the original signal may be reconstructed using the inverse wavelet transform as:

$$x(t)=\frac{1}{c_g}\int_{-\infty}^{+\infty}\int_{0}^{\infty}\frac{1}{\sqrt{a}}a\left(\frac{t-b}{a}\right)T(a,b)\frac{da.db}{a^2} \tag{4}$$

For practical applications, the hydrologist does not have at his or her disposal a continuous–time signal process but rather a discrete–time signal. A discretization of Equation (1) based on the trapezoidal rule may be the simplest discretization of the continuous wavelet transform. This transform produces N^2 coefficients from a data set of length N; hence redundant information is locked up within the coefficients, which may or may not be a desirable property (Addison et al., 2001).

To overcome this redundancy, logarithmic uniform spacing can be used for the a scale discretization with correspondingly coarser resolution of the b locations, which allows for N transform coefficients to completely describe a signal of length N. Such a discrete wavelet has the form:

$$g_{m,n}(t) = \frac{1}{\sqrt{a_0^m}} g\left(\frac{t - nb_0a_0{}^m}{a_0^m}\right) \tag{5}$$

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 and must be greater then zero. The most common and simplest choice for parameters are a_0= 2 and b_0=1.

This power–of–two logarithmic scaling of the translation and dilation is known as the dyadic grid arrangement. The dyadic wavelet can be written in more compact notation as:

$$g_{m,n}(t) = 2^{-m/2} g(2^{-m}t - n) \tag{6}$$

Discrete dyadic wavelets of this form are commonly chosen to be orthonormal; i.e.:

$$\int_{-\infty}^{+\infty} g_{m,n}(t) g_{m',n'}(t) dt = \delta_{m,m'} \delta_{n,n'} \tag{7}$$

where δ is the Kronecker delta.

This allows for the complete regeneration of the original signal as an expansion of a linear combination of translates and dilates orthonormal wavelets.

For a discrete time series, x_i , the dyadic wavelet transform becomes:

$$T_{m,n} = 2^{-m/2} \sum_{i=0}^{N-1} g(2^{-m} i - n) x_i \tag{8}$$

Where $T_{m,n}$ is wavelet coefficient for the discrete wavelet of scale a=2^m and location b=$2^m n$. Equation (8) considers a finite time series, x_i , i = 0,1,2, … , N-1; and N is an integer power of 2 : N= 2^M. This gives the ranges of m and n as, respectively, 0<n<2^{M-m}-1 and 1 < m < M. At the largest wavelet scale (i.e., 2^m where m=M) only one wavelet is required to cover the time interval, and only one coefficient is produced. At the next scale (2^{m-1}), two wavelets cover the time interval, hence two coefficients are produced, and so on down to m=1. At m=1, the a scale is 2^1 , i.e. 2^{M-1} or N/2 coefficients are required to describe the signal at this scale. The total number of wavelet coefficients for a discrete time series of length N=2^M is then 1+2+4+8+ …+2^{M-1} = N-1.

In addition to this, a signal smoothed component, $\overline{T}$, is left, which is the signal mean. Thus, a time series of length N is broken into N components, i.e., with zero redundancy. The inverse discrete transform is given by:

$$x_i = \overline{T} + \sum_{m=1}^{M} \sum_{n=0}^{2^{M-m}-1} T_{m,n} 2^{-m/2} g(2^{-m} i - n) \tag{9}$$

or in a simple format as:

$$x_i = \overline{T}(t) + \sum_{m=1}^{M} W_m(t) \tag{10}$$

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

The wavelet coefficients, $W_m(t)$ (m=1,2,...,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.

Because of simplicity of $W_1(t)$, $W_2(t)$, ..., $W_M(t)$, $\overline{T}(t)$, some interesting characteristics, such as period, hidden period, dependence and jump can be diagnosed easily through wavelet components.

STUDY AREA

The Ligvanchai watershed is located in northwest Iran in Ajarbaijan province and its main channel is a sub–branch of the Ajichai river which discharges to Urmieh lake (Figure 1). The watershed area is 75 km^2 and its elevation varies between 2140 m to 3620 m above sea level. The monthly precipitation time series for 29 years (1973-1999, 348 months), which was used in this research, is presented in Figure 2.

The monthly mean and maximum precipitations are 27.1 mm and 158.5 mm respectively in the study duration and a strong seasonality can be obviously seen in the time series. Since the normalized data is usually entered to the ANN model, the data is firstly normalized between 0 to 1.

RESULTS AND DISCUSSION

At the first, the Multi Layer Perceptron (MLP) feed forward ANN model without any data pre–processing was used to model the watershed monthly precipitation. This kind of ANN model accompanied by back propagation training algorithm is extensively used in hydrologic modeling (ASCE, 2000). At this step the model efficiency criterion (determination coefficient, R^2) showed the low performance of the model (R^2=0.31) even when a chain format of the previous months data were used as input neurons. This is probably because of significant fluctuations of the data around the mean value, so that the short term regression between data is reduced.

In the next step the pre–processed data were entered in the ANN model in order to improve the model accuracy. For this purpose the dyadic discrete wavelet transforms were used. Wavelet algorithms process data at different temporal scales (levels), thereby permitting gross and small features of a signal to be separated.

In this study the effects of the wavelet type used as well as the decomposition level on the model

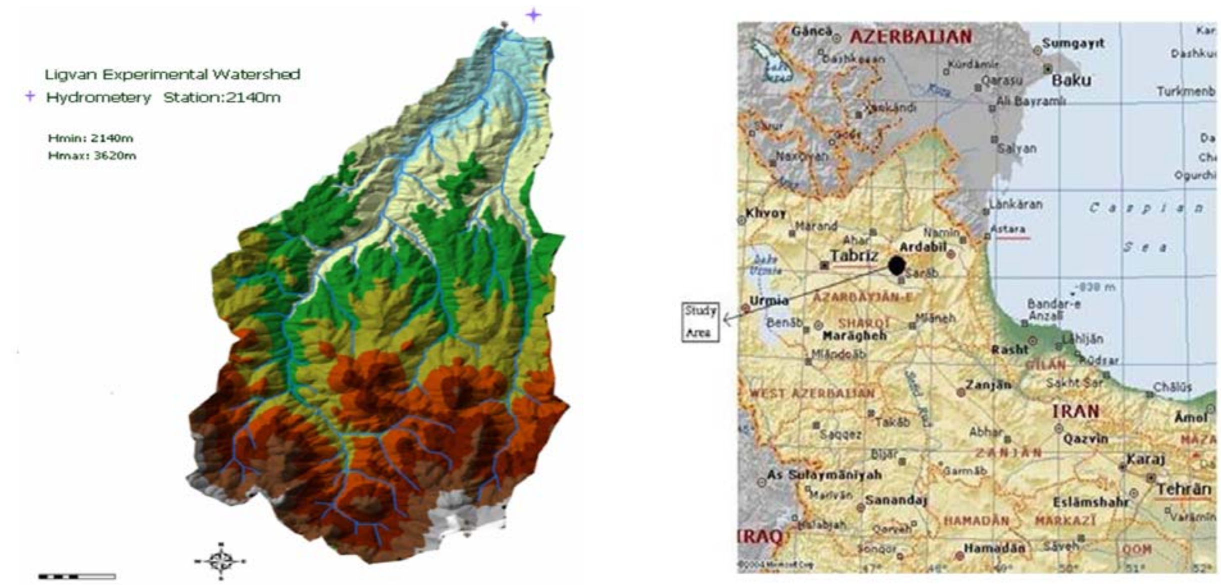

Figure 1. Study area.

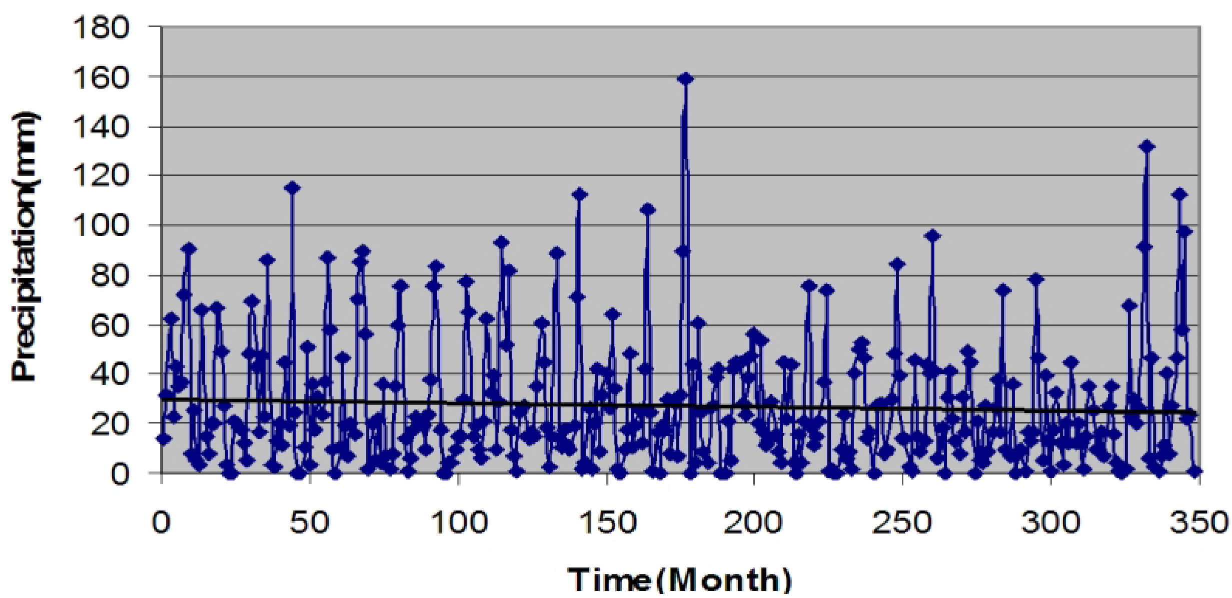

Figure 2. Precipitation time series.

efficiency was investigated. To achieve this purpose the time series was normalized and divided into two parts, calibration (24 years) and verification (5 years) data sets. Then they were decomposed to one, two and three levels by three different kinds of wavelet transforms, i.e. the 1-Haar wavelet, a simple wavelet, the 2-Daubechies-4(db4) wavelet, and a most popular wavelet, the 3-Meyer wavelet, a complex wavelet (Mallat ,1998). These wavelets are shown in Figure 3.

As examples, the level 2 decomposition of the main signal which yields 3 sub-signals (approximation at level 2 and details at levels 1,2) by the db4 wavelet and the level 3 decomposition of the signal by the Meyer wavelet are presented in Figures 4 and 5.

In continue, for each of the 9 cases mentioned, the precipitation values of a distinct month forming each sub-signal of the calibration data set were considered as input layer neurons to

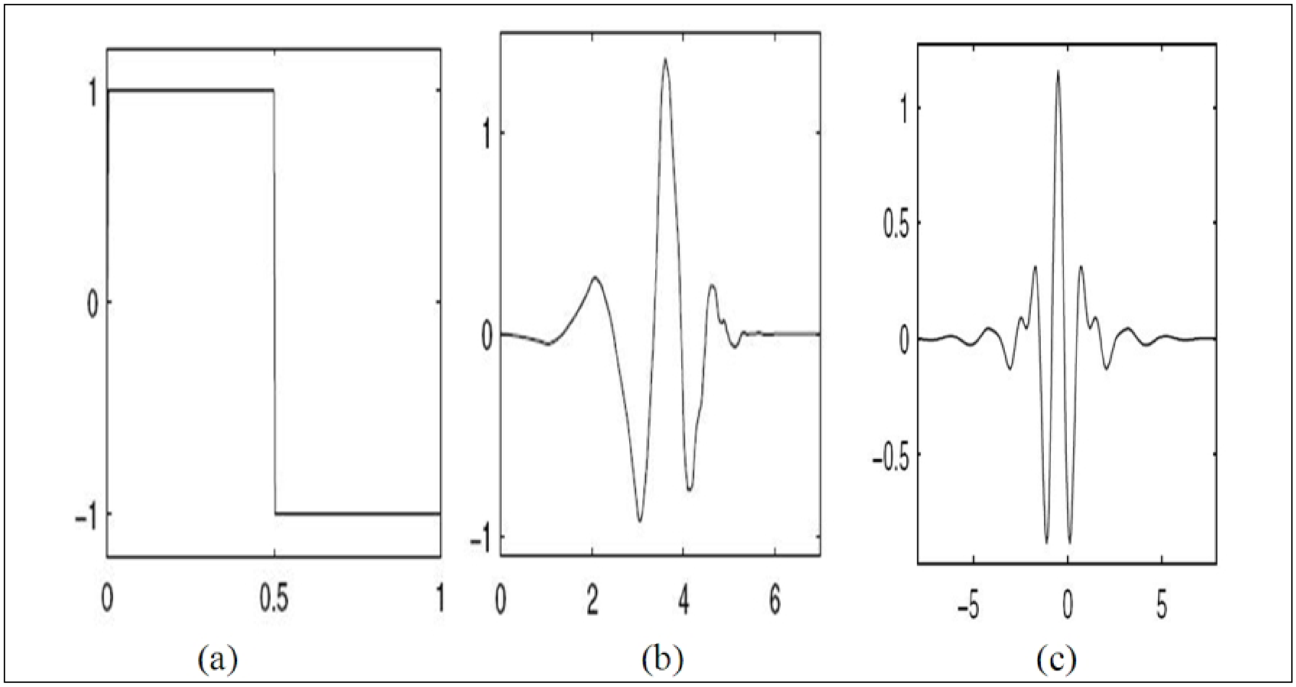

Figure 3. a) Haar wavelet b) db4 wavelet c) Meyer wavelet.

predict the precipitation one month ahead (as output layer neuron) via a 3 layers feed forward ANN model which was trained by a back propagation algorithm. Then the trained modal was validated by the verification data set. In ANN modeling two points are important and more attention must be paid to them. Firstly appropriate selection of the ANN architecture and secondly training iteration number (epoch) prevents the ANN model to be over trained. The obtained results of the study are added up in Table 1 for all cases.

When multilevel sub-signals are entered in the model as input neurons, the applied weights by ANN will be different at different levels, so that high weights will be applied to the valid level of the signal.

The calibration and verification time series of level 3 decomposition by Meyer wavelet, which is reconstructed via trained ANN, are shown in Figure 6.

By comparing the results, it can be clearly seen in the calibration phase, level 1, 2 and 3 decompositions all give relatively equal performance, but in the verification phase, by increasing the decomposition level, the model efficiency is increased. However this increase is not so perceptible from level 2 to 3, therefore level 2 can be considered as a suitable decomposition level for the data. This is in agreement with other research which offers the following formula to determine the decomposition level (Aussem et al., 1998; Wang and Ding, 2003):

$$L=\text{int}\,[\log (N)] \tag{11}$$

in which L and N are decomposition level and time series length respectively. For the study at hand N = 348, so L=2.

In addition, according to the simple structure of the Haar wavelet, Figure 3, the signal features could not be completely captured, especially at peak values, and it yielded comparatively low efficiency.

In order to evaluate the proposed model ability in comparison with other classic models, the

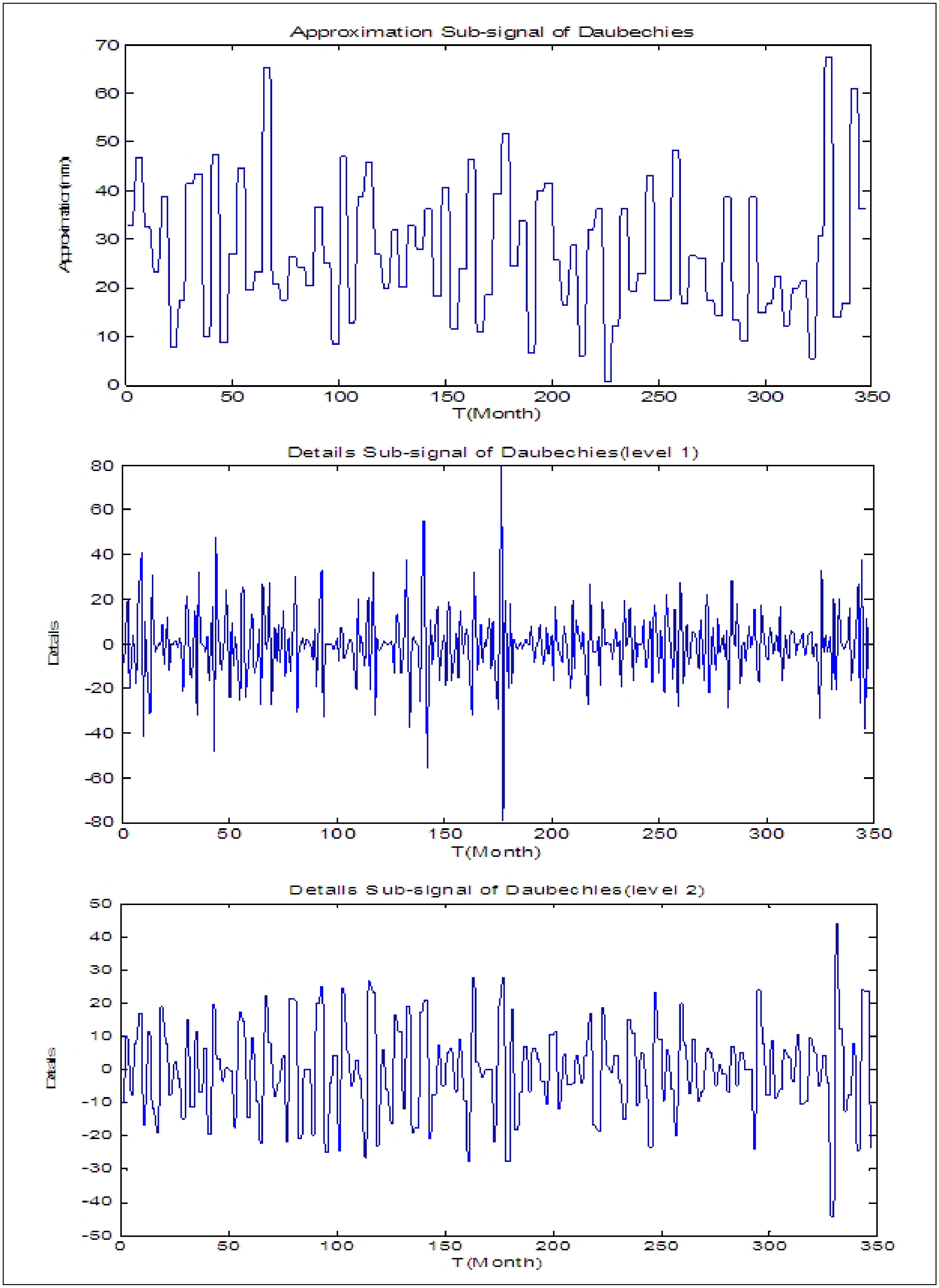

Figure 4. Approximation and details sub-signals of db4 (level 2).

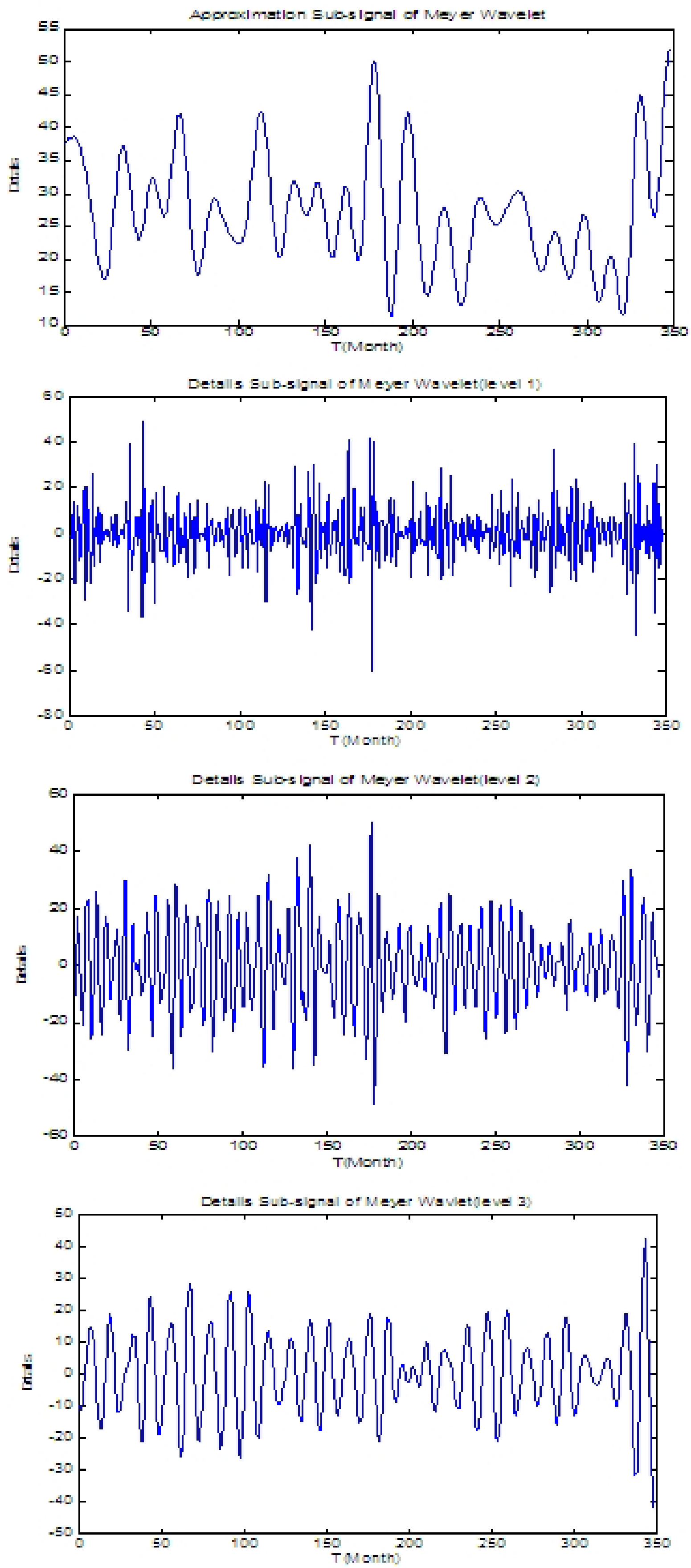

Figure 5. Approximation and details sub-signals of Meyer (level 3).

Table 1. Model characteristics and efficiency criteria.

Wavelet Type	Decomposition Level	ANN Architecture	Training Epoch	Calibration R^2	Verification R^2
Haar	1	2-11-1	350	0.76	0.61
Haar	2	3-16-1	400	0.8	0.7
Haar	3	4-20-1	450	0.78	0.72
db1	1	2-9-1	350	0.9	0.78
db1	2	3-16-1	400	0.93	0.85
db1	3	4-18-1	450	0.93	0.87
Meyer	1	2-9-1	300	0.91	0.8
Meyer	2	3-16-1	400	0.93	0.84
Meyer	3	4-24-1	450	0.935	0.89

SARIMA (Seasonal Auto Regressive Integrated Moving Average) time series model (Salas et al., 1980) was also used to model the Ligvanchai watershed precipitation. The PACF (Partial Auto Correlation Function) and final result of the model as a scatter plot are presented in Figures 7 and 8. Although the SARIMA model efficiency (R^2=0.64) shows the model acceptability because of using seasonal decomposition, it still has low performance in comparison with the proposed model. This may be due to the linear nature of the SARIMA model, but when ANN is used to reconstruct sub-signals, its non-linear property can help the model to detect and catch the non-linear features of the modeled process.

CONCLUSIONS

In this study the wavelet transform, which can capture the multi scale features of signals, was used to decompose the Ligvanchai precipitation time series. Then the sub-signals were used as input to the ANN model to predict the precipitation one month ahead. Because of using the ANN model for reconstruction of the signal, the developed model has a non-liner kernel so that it can simulate the non-linear behavior of the phenomenon more accurately than other linear models such as the SARIMA.

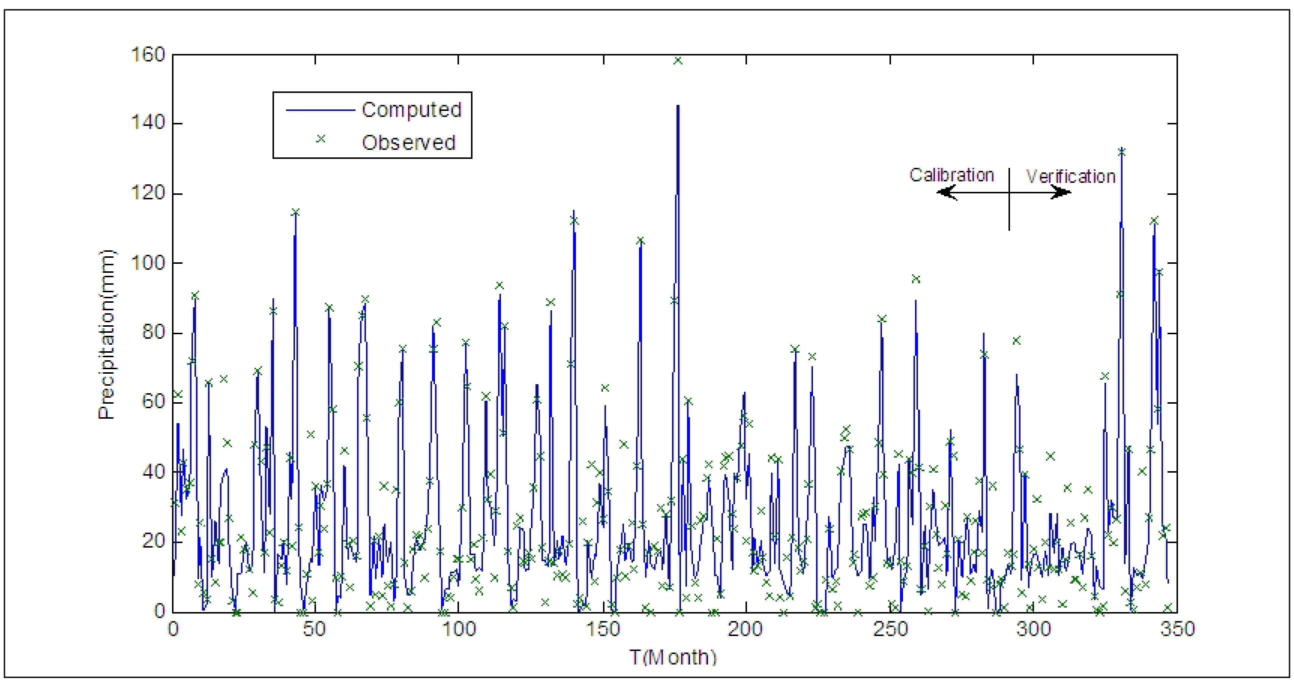

Figure 6. Computed and observed time series.

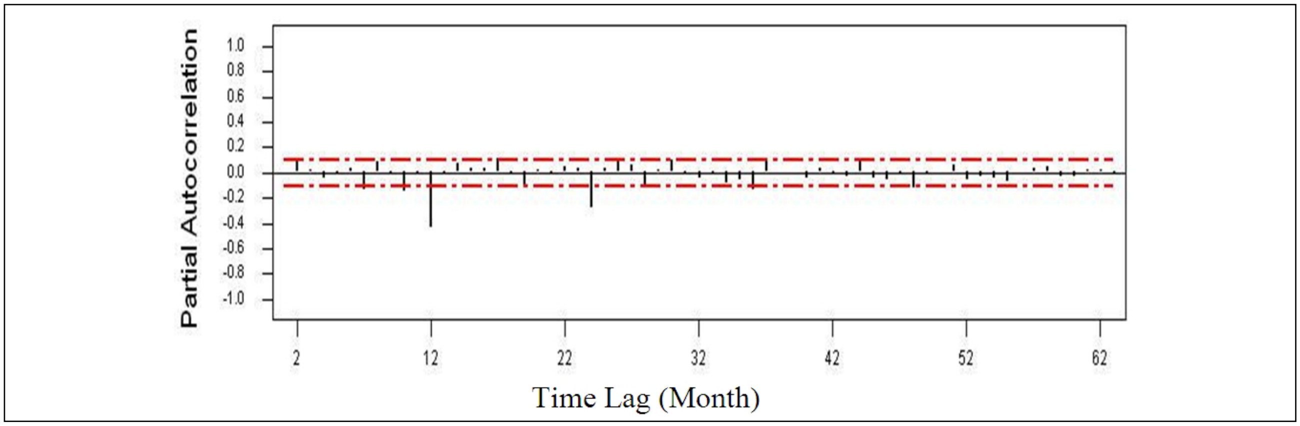

Figure 7. Partial Auto Correlation Function.

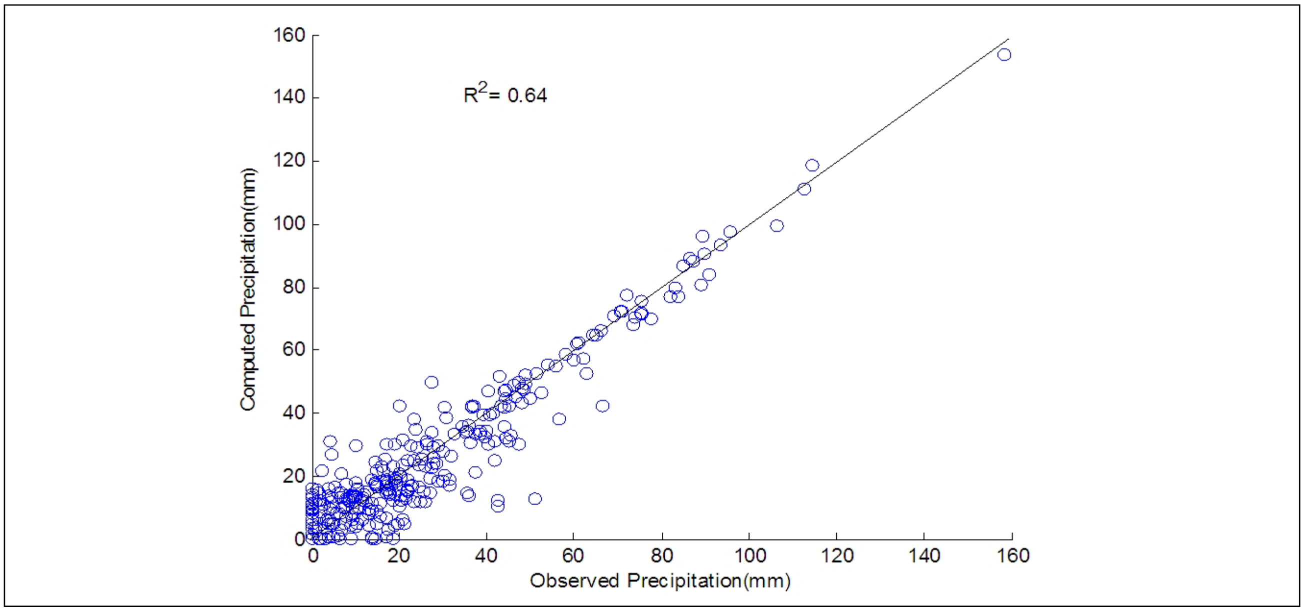

Figure 8. SARIMA output.

Furthermore, the effect of wavelet transform type on the model performance was investigated using three different kinds of wavelet transforms. The model results show the low value of the Haar wavelet in comparison with the others (i.e. db4 and Meyer) because of its simple and elementary structure.

As final result, it was deduced that although increasing decomposition level can improve the model ability, an optimum level can be chosen on the basis of the signal length.

In order to complete the current study, it is suggested to use the methodology presented here to forecast the precipitation 2,3,… months ahead and also to model the rainfall-runoff process of the watershed using rainfall and runoff time series.

Also, due to wavelet capabilities, it is proposed to use the wavelet transform for trend analysis of the watershed hydrological components. Obviously, to achieve this goal a large amount of long term data will be needed.

ACKNOWLEDGMENTS

Prof. Hassanzadeh and Dr. Farahmand Azar, academic members of Civil Eng. Faculty of Tabriz University are thanked for the paper review and useful comments.

REFERENCES

Addison, P. S., K.B. Murrary , and J.N. Watson. 2001. Wavelet transform analysis of open channel wake flows. J. of Eng. Mech., Vol. 127(1),pp. 58-70.

ASCE task Committee on application of Artificial Neural Networks in hydrology .2000. Artificial Neural Networks in hydrology 1: Hydrology application. J. of Hydrologic Eng., ASCE,Vol. 5(2),pp. 124-137.

Aussem,A., J. Campbell, and F. Murtagh. 1998. Wavelet-based feature extraction and decomposition strategies for financial forecasting. J. Compt. Intelli. Fin., Vol. 6(2),pp. 5-12.

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

Foufoula-Georgiou, E., and P. Kumar (eds). 1995. Wavelet in Geophysics. New York; Academic Press.

Grossmann,A., and J. Morlet. 1984. Decomposition of hardy function into square integrable wavelets of constant shape. SIMA Journal of Mathematical Analysis, Vol. 15, pp. 723-736.

Kim, T., and J.B. Valdes. 2003. Nonlinear model for drought forecasting based on a conjunction of wavelet transforms and neural networks. J. of Hydrologic Eng. , ASCE,Vol. 8(6) , pp. 319-328.

Labat, D . 2005. Recent advances in Wavelet analyses: part 1 –A review of concepts. J. of Hydrology, Vol. 314, pp. 275-288.

Labat, D., R. Ababou, and A. Mangin. 2000. Rainfall-runoff relation for karstic spring. Part 2 :continuous wavelet and discrete orthogonal multi resolution analyses. J. of Hydrology, Vol. 238, pp. 149-178.

Mallat, S.G. 1998. A wavelet tour of signal processing. San Diego; Academic Press.

Salas, J.D., J.W. Delleur, V. Yevjevich, and W.L. Lane. 1980. Applied Modeling of Hydrological Time Series. Colorado; Water Resources Publications.

Wang, W., and S. Ding. 2003. Wavelet network model and its application to the predication of hydrology. Nature and Science, Vol. 1(1), pp. 67-71.

Zhang, B.L., and Z.Y. Dong. 2001. An adaptive neural wavelet model for short term load forecasting. Electric Power Syst. Res., Vol. 59, pp. 121-129.

ADDRESS FOR CORRESPONDENCE
Vahid Nourani
Faculty of Civil Engineering
University of Tabriz
Tabriz , Iran

Email : nourani@tabrizu.ac.ir