Int. J. Environ. Sci. Tech., 7 (1), 93-110, Winter 2010
ISSN: 1735-1472

Prediction of daily suspended sediment load using wavelet and neuro-fuzzy combined model

[1]T. Rajaee; [1]*S. A. Mirbagheri; [2]V. Nourani; [3]A. Alikhani

[1]Department of Civil Engineering, K.N. Toosi University of Technology, Tehran, Iran

[2]Faculty of Civil Engineering, University of Tabriz, Tabriz, Iran

[3]Deptartment of Civil Engineering, University of Qom, Qom, Iran

Received 27 December 2008; revised 18 April 2009; accepted 1 November 2009; available online 1 December 2009

ABSTRACT: This study investigated the prediction of suspended sediment load in a gauging station in the USA by neuro-fuzzy, conjunction of wavelet analysis and neuro-fuzzy as well as conventional sediment rating curve models. In the proposed wavelet analysis and neuro-fuzzy model, observed time series of river discharge and suspended sediment load were decomposed at different scales by wavelet analysis. Then, total effective time series of discharge and suspended sediment load were imposed as inputs to the neuro-fuzzy model for prediction of suspended sediment load in one day ahead. Results showed that the wavelet analysis and neuro-fuzzy model performance was better in prediction rather than the neuro-fuzzy and sediment rating curve models. The wavelet analysis and neuro-fuzzy model produced reasonable predictions for the extreme values. Furthermore, the cumulative suspended sediment load estimated by this technique was closer to the actual data than the others one. Also, the model could be employed to simulate hysteresis phenomenon, while sediment rating curve method is incapable in this event.

Keywords: *Artificial intelligence; Hysteresis; Modeling; Sediment rating curve; Wavelet decomposition*

INTRODUCTION

The sediment load transported in river is the most complex hydrological and environmental phenomenon due to the large number of obscure parameters such as spatial variability of basin characteristics and river discharge patterns. In most rivers, sediments are mainly transported as suspended sediment load (SSL) (Morris and Fan, 1997). Many models have been provided to simulate this phenomenon (Mirbagheri *et al.*, 1988a, b; Verstraeten and Poesen, 2001). Due to large number of obscure parameters involved in this phenomenon, the theoretical governing equations may not be of much advantage in gaining knowledge of the overall process. Studies have been made to develop artificial intelligence techniques for simulation processes with limited adequate knowledge of the physics (Rajaee and Mirbagheri, 2009; Nourani *et al.*, 2008a). In the previous years, the fuzzy logic has been used in the simulation of uncertainties in the water resources and environmental engineering such as river pollution management (Nasiri *et al.*, 2007), centralized return centers location evaluation in a reverse logistics network (Tuzkaya and Gülsün, 2008) environmental performance evaluation of suppliers (Tuzkaya *et al.*, 2009) and integrated water systems modeling (Nguyen *et al.*, 2007). Artificial neural networks have been successfully applied to many tasks in environmental engineering (Bandyopadhyay and Chattopadhyay, 2007; Rene *et al.*, 2008). Neuro-fuzzy modeling is another method that refers to the approach of applying deferent learning algorithms developed in the neural network literature to fuzzy modeling or a fuzzy inference system (FIS) (Brown and Harris, 1994). Neuro-fuzzy model (NF) has been applied to a number of problems in water resources and environmental engineering, including river flow modeling (Zounemat-Kermani and Teshnehlab, 2008), predicting and identifying traffic hot spots (Hadji Hosseinlou and Sohrabi, 2009), hydrological time series modeling (Firat and Gungor, 2008) and ecological status modeling in surface waters (Ocampo *et al.*, 2007). There are a few researches in employment of fuzzy inference system and neuro-fuzzy approaches in suspended sediment modeling. Tayfur *et al.* (2003) provided a fuzzy logic method using the rainfall intensity and slope data to predict sediment loads from bare soil surfaces. The research showed that the fuzzy approach performed better under very high rainfall intensities over different

*Corresponding Author Email: *mirbagheri@kntu.ac.ir*
Tel.: +9821 8877 9473; Fax: +9821 8803 5516

slopes and over very steep slopes under various rainfall intensities. Lohani *et al.* (2007) developed a fuzzy inference system to simulate the stage-discharge-sediment concentration relationship in two gauging stations in the Narmada basin in India. Results of the mentioned study showed that the fuzzy method was capable to provide much better results than rating curve method. Kisi *et al.* (2008) studied the accuracy of an adaptive neuro-fuzzy computing method in monthly suspended sediment estimation in Kuylus and Salur Koprusu stations in Turkey. The results showed that NF model produced better performance than artificial neural network (ANN) and SRC models. In Rajaee *et al.* (2009), NF, ANN, multi linear regression and SRC models were examined for daily simulation of suspended sediment concentration in two hydrometery stations. The models were trained using daily river discharge and sediment concentration data belonging to Little Black River and Salt River stations in the USA. Comparison of the models' results indicated that the NF model was more accurate in predicting sediment concentration in comparison with the other models.

A wavelet analysis is a set of building blocks to build or represent a signal or function. It has increased in practice and popularity in latest years. Wavelet analysis, which give information in both the time and frequency domains of the signal, give considerable knowledge about the physical form of the data. It supplies a time-frequency representation of a signal at many different periods in the time domain (Daubechies, 1990). Wavelet transformed data of original time series improve the ability of a predicting model by capturing useful information on various resolution levels (Kim and Valdes, 2003). An inclusive literature overview of wavelet analysis in geosciences can be found in Foufoula-Georgiou and Kumar (1995) and the most recent contributions are cited by Labat (2005). In the past few years, wavelet analysis has been employed to problems in water resources and environmental engineering, including river flow modeling (Pasquini and Depetris, 2007), meteorological pollution simulation (Osowski and Garanty, 2007), open channel wake flows analysis (Addison *et al.*, 2001) and groundwater level time series modeling (Wang and Ding, 2003).

Wavelet analysis and artificial intelligent approaches (such as FIS and NF) are indicated to be suitable when applied individually to environmental and water resources problems. Recently, there has been a growing interest in combining methods. Partal and Kisi (2007) developed a wavelet and neuro-fuzzy conjunction model for daily precipitation forecasting in Turkey. Their neuro-fuzzy model is constructed with appropriate wavelet sub-series as input and original precipitation as output. The provided wavelet-neuro-fuzzy model well fit with the measured data, particularly for zero and peaks precipitation time series. Results showed that the provided model produced significantly better results than neuro-fuzzy approach. Nourani *et al.* (2008b) proposed a combined neural-wavelet model. In their research, the wavelet analysis was linked to ANN for prediction of Ligvanchai watershed precipitation at Tabriz- Iran. For this purpose, the main time series was decomposed to some multi-frequently time series by wavelet. Then, these time series were imposed as input data to the ANN to predict the precipitation of one month ahead. The obtained results showed that the proposed model can predict both short and long term precipitation events because of using multi-scale time series as the ANN input layer. Wei *et al.* (2009) provided a wavelet network approach for modeling of a permeate flux of cross-flow membrane filtration. The aim of this research is to construct a new model based on wavelet transform and adaptive neuro-fuzzy approach for suspended sediment load prediction in Pecos gauging station in the USA. The purpose of combining the wavelet analysis with NF technique is to increase the accuracy of SSL prediction.

MATERIALS AND METHODS

Study area and statistical analysis

The proposed NF and wavelet analysis and neuro-fuzzy (WNF) models need uninterrupted time series data pertaining to river discharge (*Q*) and SSL (*S*) at a gauging station. The data obtained from the Pecos River near Artesia, NM (USGS Station No: 08396500, Basin area (sq. mi.): 15300, Latitude: 32º50'25" and Longitude: 104º 19' 23") was used for calibration and verification for all the models provided in this study. The Pecos River is situated in eastern New Mexico and western Texas. There are primarily two major water inputs, namely snowmelt from winter storms in the headwater region of the southern Rocky Mountains and runoff from warm-season monsoonal rainfall in the lower valley. This river has been the subject of investigation by Yuan *et al.*, (2007). Fig. 1 shows the state and the gauging station.

The data from October 1, 1965 to September 30, 1972 (7 years) and the data from October 1, 1972 to September 30, 1974 (2 years) were used for calibration and verification sets, respectively. The data statistics for training and testing sets are given in Table 1, which contain the minimum, maximum, mean, standard

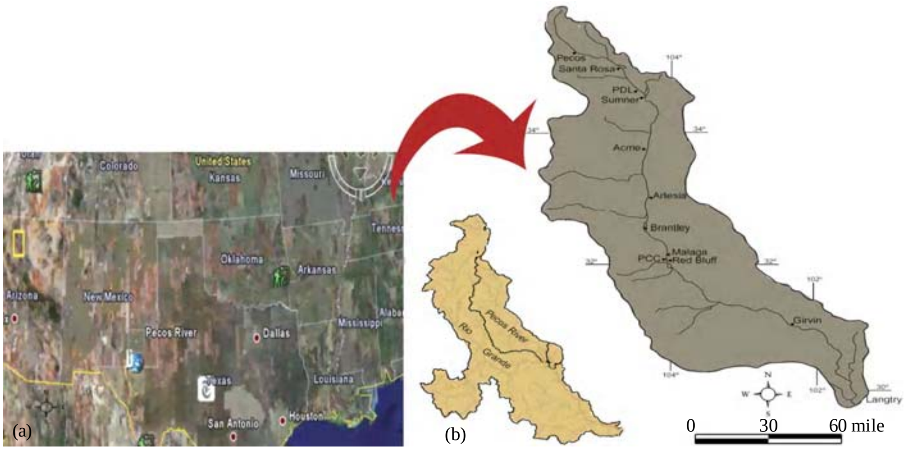

Fig. 1: (a) New Mexico State, (b) Drainage map showing the Pecos River and its adjacent areas

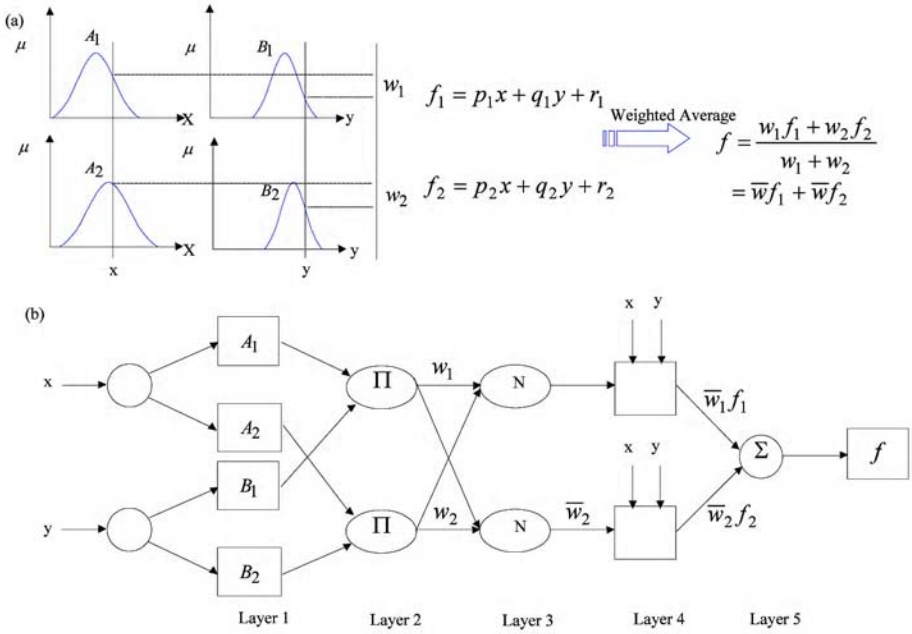

Fig. 2: Structure of ANFIS system: (a) Fuzzy inference system; (b) Equivalent ANFIS architecture

deviation (S_d), skewness coefficient (C_{sx}), lag 1 day autocorrelation coefficient (R_1), lag 2 days autocorrelation coefficient (R_2) and lag 3 days autocorrelation coefficient (R_3).

From Table 1, it is obvious that the extreme values of the available data are in the training set. When classifying the data into training and testing subsets, it is essential to check that the data represent the same statistical population (Masters, 1993). Discharge and SSL lag 1 day autocorrelation coefficients are very adequate and relatively same in calibration and validation data sets. Also, SSL lag 2 and 3 days autocorrelation coefficients are satisfactory. It is seen that the skewness coefficients are low in training and testing sets. It is noted that the high skewnees coefficient has a considerable negative effect on ANN

Table 1: Statistics analysis for training, testing and all data sets

	Training set		Testing set		All data set	
Statistical parameters	S (ton/day)	Q (m^3/day)	S (ton/day)	Q (m^3/day)	S (ton/day)	Q (m^3/day)
Mean	1277.3	3.89×10^5	1475.8	5.07×10^5	1321.4	4.15×10^5
S_d	5026.9	8.02×10^5	4619.2	10.5×10^5	4939.2	8.65×10^5
C_{sx}	7.458	5.663	5.312	6.304	7.075	6.065
Min	0.009	1097.3	0.066	6359	0.009	1097.3
Max	79315	149.47×10^5	42768	137.38×10^5	79315	149.47×10^5
R_1	0.676	0.809	0.861	0.865	0.712	0.828
R_2	0.426	0.628	0.704	0.642	0.48	0.634
R_3	0.288	0.529	0.576	0.487	0.344	0.517

performance (Altun *et al.*, 2007). In general, Table 1 shows satisfactory statistics characteristics between training and testing sets in terms of mean, standard deviation, skewness coefficient and correlation coefficient.

Neuro-Fuzzy approach

An special algorithm in neuro-fuzzy development is the adaptive neuro-fuzzy inference system (ANFIS). It is a network statement of Sugeno-type fuzzy models and is introduced by Jang (1993). The structure of an ANFIS is shown in Fig. 2. Fig. (2a) shows the fuzzy reasoning mechanism for the Sugeno model to derive an output function f from a given input vector $[x,y]$. The corresponding equivalent ANFIS construction is showed in Fig. (2b). As an example, a fuzzy inference system with two inputs x and y and one output f was considered. For the first order Sugeno fuzzy model, a typical rule set with two fuzzy *If-Then* rules can be expressed as:

Rule 1: If x is A_1 and y is B_1, then $f_1=p_1x+q_1y+r_1$ (1)

Rule 2: If x is A_2 and y is B_2, then $f_2=p_2x+q_2y+r_2$ (2)

Where, A_1, A_2 and B_1, B_2 are the membership functions for inputs x and y, respectively; p_1, q_1, r_1 and p_2, q_2, r_2 are the parameters of the output function. The functioning of the ANFIS is as follows:

Layer 1: The node output OP_i^l is defined by:

$$OP_i^l = \mu_{Ai}(x) \text{ for } i=1,2 \text{ or} \tag{3}$$

$$OP_i^l = \mu_{Bi-2}(y) \text{ for } i=3,4 \tag{4}$$

Where, x (or y) is the input to the node; A_i (or B_i-2) is a linguistic label (such as 'low' or 'high') associated with this node, characterized by the form of the membership functions in this node and can be any suitable functions that are continuous and piecewise differentiable such as Gaussian, trapezoidal shaped, generalized bell shaped and triangular shaped functions. The membership functions for A and B are usually described by generalized bell functions. The output OP_i^l can be computed as:

$$OP_i^l = \mu_{Ai}(x) = \frac{1}{1+\left((x-c_i)/a_i\right)^{2b_i}} \tag{5}$$

Where, $\{a_i, b_i, c_i\}$ is the parameter set.

Layer 2: Every node in this layer multiplies the incoming signals:

$$OP_i^2 = w_i = \mu_{Ai}(x)\mu_{Bi}(y), \quad , i=1,2. \tag{6}$$

Layer 3: The *ith* node of this layer calculates the normalized firing strengths as:

$$OP_i^3 = \overline{w}_i = \frac{w_i}{w_1+w_2}, \quad i=1,2. \tag{7}$$

Layer 4: Node i in this layer calculates the contribution of the *ith* rule towards the model output:

$$OP_i^4 = \overline{w}_i f_i = \overline{w}_i(p_i x + q_i y + r_i) \tag{8}$$

Where, $\overline{w}$ is the output of layer 3 and $\{p_i, q_i, r_i\}$ is the parameter set.

Layer 5: The single node in this layer calculates the overall output of the ANFIS as (Jang and Sun, 1995; Nayak *et al.*, 2004; Aqil *et al.*, 2007):

$$OP_i^5 = \sum_i \overline{w}_i f_i = \frac{\sum_i w_i f_i}{\sum_i w_i} \tag{9}$$

The optimization parameters in an ANFIS are the premise and consequent parameters. The learning algorithm is a hybrid algorithm, which is a