Wavelet and neuro-fuzzy combination model for predicting suspended sediment transport in rivers

AMIR ALIKHANI
Department of Civil Engineering
Qom University
Old rod of Isfehan, Qom
IRAN
dramiralikhani@yahoo.com

Abstract: - In this paper prediction of the suspended sediment load (SSL) of the rivers by using the available actual data from Pecos River gauging station in the USA was done. For this purpose, result of three models were compared with the actual original data. The models were conventional sediment rating curve (SRC) model, neuro-fuzzy (NF) and conjunction of wavelet analysis and neuro-fuzzy (WNF). In the proposed WNF model, observed time series of river discharge and SSL were decomposed at different scales by wavelet analysis. Then summed effective time series of discharge and SSL were imposed as inputs to the NF model for prediction of SSL in one day ahead. Obtained results showed that the WNF model performance was better in prediction rather than the NF and SRC models. The WNF model produced reasonable predictions for the extreme values. Furthermore, the cumulative SSL estimated by this technique was closer to the actual data than the other ones. In addition to that, the model could be employed to simulate hysteresis phenomenon, while SRC method did not able to simulate this event.

Key-Words: - Time series, Wavelet decomposition, Artificial intelligence, Sediment rating curve, Hysteresis.

1 Introduction

A wavelet analysis is a set of building blocks to build or represent a signal or function. In recent years, it has increased in practice and popularity. 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. Wavelet transformed data of original time series improve the ability of a predicting model by capturing useful information on various resolution levels see 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). Recently, there has been a growing interest in combining methods. The sediment load transported in river is the most complex hydrological and environmental phenomenon. In most rivers, sediments are mainly transported as SSL. 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 is led to a good fit with the measured data. The results of their study showed that the provided model produced significantly better results than neuro-fuzzy approach.

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 using the data of Pecos gauging station in the USA. The purpose of combining the wavelet analysis with NF technique is to increase the accuracy of SSL prediction.

2 Data and area of study

The proposed NF and 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) was used for calibration and verification for all the models provided in this study.

<table>
<thead>
<tr>
<th>$C_{xx}$</th>
<th>7.458</th>
<th>5.663</th>
<th>5.312</th>
<th>6.304</th>
<th>7.075</th>
<th>6.065</th>
</tr>
</thead>
<tbody>
<tr>
<td>Min</td>
<td>0.009</td>
<td>1097.3</td>
<td>0.066</td>
<td>6359</td>
<td>0.009</td>
<td>1097.3</td>
</tr>
<tr>
<td>Max</td>
<td>79315</td>
<td>149.47×10^5</td>
<td>42768</td>
<td>137.38×10^5</td>
<td>79315</td>
<td>149.47×10^5</td>
</tr>
<tr>
<td>$R_s$</td>
<td>0.676</td>
<td>0.809</td>
<td>0.861</td>
<td>0.865</td>
<td>0.712</td>
<td>0.828</td>
</tr>
<tr>
<td>$R_2$</td>
<td>0.426</td>
<td>0.628</td>
<td>0.704</td>
<td>0.642</td>
<td>0.48</td>
<td>0.634</td>
</tr>
<tr>
<td>$R_3$</td>
<td>0.288</td>
<td>0.529</td>
<td>0.576</td>
<td>0.487</td>
<td>0.344</td>
<td>0.517</td>
</tr>
</tbody>
</table>

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

3 METHODS

3.1 Sediment Rating Curve (SRC) method

In many rivers a main part of the sediment is transported in suspension. Almost of this load contained of silt and clay, i.e. wash load. It can thus be concluded that wash load plays an important role in the sediment transport in rivers (Asselman, 2000). The primary data collected to determine the suspended sediment discharge of a river are $Q_w$ and $C$. $Q_{sw}(m^3/s)$ is the instantaneous river discharge and measured with a current meter or taken from a stage-discharge curve for the gauging station. $C$ is the instantaneous suspended sediment concentration in mg/l or ppm. Concentration is measured by analysis of water samples. Finally, the instantaneous suspended sediment discharge (ton/day) is calculated from $Q_w$ and $C$.

As the finest fraction of the SSL often is a non-capacity load it cannot be predicted using stream power related sediment transport models. Instead, empirical relations such as sediment rating curves often are applied (Asselman, 2000). The establishment of a SRC is an important object of hydrology. Since the measurement of sediment is costly and needs time, usually discharge is measured per day. The SRC is used to assess the sediment discharge corresponding to the measured flow discharge. Usually the SRC has the form $S = aQ_w^b$ where $S$ is the suspended sediment load or discharge and $a$ and $b$ are constants. Thus the SRC has important bearing on correct assessment of SSL. As a SRC can be considered a ‘black box’ type of model, the coefficients $a$ and $b$, have no physical meaning. However, some physical interpretation is often
3.2 Neuro-Fuzzy approach:
An especial 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 Figure 1. Figure (1a) shows the fuzzy reasoning mechanism for the Sugeno model to derive an output function \( f \) from a given input vector \([x, y]\).  

![Figure 1. Structure of ANFIS system: (a) Fuzzy inference system; (b) Equivalent ANFIS architecture](image)

The corresponding equivalent ANFIS construction is showed in Figure (1b).

3.3 Proposed Wavelet-Neuro-Fuzzy (WNF) model
To construct the model, firstly measured river discharge and SSL time series were decomposed to some multi-frequently time series \( Q_{d1}(t), Q_{d2}(t), ... , Q_{di}(t) \), \( Q_n(t) \) and \( S_{d1}(t), S_{d2}(t), ... , S_{di}(t) \), \( S_n(t) \) by discrete wavelet transform (DWT). Which \( Q_{d1}(t), Q_{d2}(t), ... , Q_{di}(t) \) and \( Q_n(t) \) are the details and approximation (or background) river discharge time series, respectively; \( S_{d1}(t), S_{d2}(t), ... , S_{di}(t) \) and \( S_n(t) \) are the details and approximation SSL time series, respectively; \( di \) shows the level \( i \) decomposed time series and \( a \) denotes approximation time series. These time series play various roles in the original time series and the behavior of each is distinct.

The measured river flow and SSL time series was decomposed using different scales, from 1 to 10 (i.e. into 10 wavelet decomposition levels (2-4-8-16-32-64-128-256-512-1024 days)).

The decomposed SSL time series (SDW) of the mentioned modes are illustrated in Figure 2.

The river SSL time series of 2-day mode (DDW 1), 4-day mode (DDW 2), 8-day mode (DDW 3), 16-day mode (DDW 4), 32-day mode (DDW 5), 64-day mode (DDW 6), 128-day mode (DDW 7), 256-day mode (DDW 8), 512-day mode (DDW 9), 1024-day mode (DDW 10) and approximate mode (DDW App.) are shown in this Figure. Similar figures can be shown for the river discharge time series.
The discrete wavelet transform belongs to the multi resolution analysis (Mallat, 1989; Daubechies, 1990). It decomposes the time series into a set of basis functions of various frequencies. In this study, the river discharge and SSL time series are constructed by a very irregular signal form, so an irregular wavelet, the Daubechies wavelet of order 4 (db4) (Daubechies, 1992), was selected for employing in this paper.

The aim of this research is to substitute the prediction of the observed SSL of high variability by the prediction of its summed wavelet coefficients on various levels of reduced variabilities. Remaining issues contain determining which wavelet levels are most effective in SSL prediction. In this part, in order to select the forceful wavelet components in prediction, the correlation coefficients between decomposed river discharge and SSL with original SSL time series are calculated and illustrated in Table 2.

Table 2. The correlation coefficients of the discrete wavelet components with the measured SSL ($S_t$).

<table>
<thead>
<tr>
<th>Discharge discrete wavelet components</th>
<th>Correlation between $DDW(i)_{t-1}$ and $S_t$</th>
<th>Correlation between $DDW(i)_{t-2}$ and $S_t$</th>
<th>Correlation between $SDW(i)_{t-1}$ and $S_t$</th>
<th>Correlation between $SDW(i)_{t-2}$ and $S_t$</th>
</tr>
</thead>
<tbody>
<tr>
<td>DDW 1</td>
<td>-0.087</td>
<td>-0.043</td>
<td>-0.144</td>
<td>-0.038</td>
</tr>
<tr>
<td>DDW 2</td>
<td>0.064</td>
<td>0.256</td>
<td>0.144</td>
<td>0.212</td>
</tr>
<tr>
<td>DDW 3</td>
<td>0.320</td>
<td>0.154</td>
<td>0.381</td>
<td>0.187</td>
</tr>
<tr>
<td>DDW 4</td>
<td>0.375</td>
<td>0.186</td>
<td>0.330</td>
<td>0.270</td>
</tr>
<tr>
<td>DDW 5</td>
<td>0.246</td>
<td>0.236</td>
<td>0.387</td>
<td>0.387</td>
</tr>
<tr>
<td>DDW 6</td>
<td>0.202</td>
<td>0.282</td>
<td>0.373</td>
<td>0.250</td>
</tr>
<tr>
<td>DDW 7</td>
<td>0.219</td>
<td>0.221</td>
<td>0.329</td>
<td>0.249</td>
</tr>
<tr>
<td>DDW 8</td>
<td>0.1136</td>
<td>0.1124</td>
<td>0.121</td>
<td>0.121</td>
</tr>
<tr>
<td>DDW 9</td>
<td>0.0855</td>
<td>0.0832</td>
<td>0.105</td>
<td>0.105</td>
</tr>
<tr>
<td>DDW App</td>
<td>0.069</td>
<td>0.0296</td>
<td>0.059</td>
<td>0.058</td>
</tr>
</tbody>
</table>

Since majority of the decomposed time series are of lower variability we prepare for the increase of the prediction precision.

According to the Table 2, DDW 1, DDW 2,
DDW 9, DDW 10 and DDW App. show low correlation with original SSL, therefore DDW 3, DDW 4, DDW 5, DDW 6, DDW 7 and DDW 8, were selected as effective wavelet components and were summed together to obtain total discharge discrete wavelet (TDDW). In this manner, SDW 1, SDW 2, SDW 10 and SDW App. show low correlation with original SSL, thus SDW 3, SDW 4, SDW 5, SDW 6, DDW 7, DDW 8 and SDW 9 which have high correlation, were chosen as forceful wavelet components and summed together to provide total sediment discrete wavelet (TSDW). In Figure 3 the TDDW and TSDW are shown.

According to Table 3, correlation coefficient between SSL ($S_t$) and $Q_{t-1}$ is 0.616, while it is 0.654 for $TDDW_{t-1}$ and correlation coefficient between SSL and $Q_{t-2}$ is 0.424, but it is 0.537 for $TDDW_{t-2}$. Also, SSL lag 1 day autocorrelation coefficient is 0.712 (Table 1), while it is 0.8 for $TSDW_{t-1}$ and SSL lag 2 days autocorrelation coefficient is 0.48, but it is 0.648 for $TSDW_{t-2}$. It is obvious that the wavelet analysis is extremely useful when employed before correlation coefficient assessment to extract effective sub-time series in SSL prediction. In Figure 4 the construction of the proposed WNF model is shown.

Wavelet transforms provide useful decompositions of main time series, so that wavelet-transformed data improve the ability of a predicting model by capturing useful information on various resolution levels. Hence a hybrid WNF model which uses summed multi-scale signals as input data may present more probable prediction rather than a single pattern input.

### 4 Comparison result of three models

Figure 5 shows measured and estimated cumulative SSL for verification period, and Figure 6 shows, WNF and SRC performances in hysteresis phenomenon modeling for verification period.
5 CONCLUSION

In this study a new model by conjunction of wavelet analysis and NF approach was applied to daily suspended sediment load prediction in a gauging station in the USA. In the provided WNF model, at first the observed time series of river discharge and suspended sediment load were decomposed to some sub-time series at different scales by discrete wavelet analysis. Then, effective sub-time series were summed together to obtain useful river discharge and SSL time series in prediction. In the WNF model, selection of appropriate decomposed time series is important in model performance. Afterwards, these total time series were imposed as inputs to the NF model for SSL prediction in one day ahead.

The paper presents a comparative study on convenient classic and new generation hybrid intelligence approaches in SSL modeling. Therefore it will be of particular utility to researchers that require time-series of SSL and but do not have the resources to support sampling or turbidity monitoring and are deciding between various models that predict the needed data from discharge values. This research prepares prediction benchmarks for SSL prediction in the type of numerical and visual contrast between NF, WNF and SRC models. Results indicate that the WNF model is suitable in predictions and improve the NF and SRC performances. Also, the WNF model can satisfactory estimates cumulative suspended sediment load and predicted high SSL values by this model show better fitting to the observed data than the other models. The NF approach is in reasonable agreement with the actual SSL data than the SRC technique but the errors demonstrate that some contributions of the physics are disguised. WNF model goes someway towards including this unknown physics and so improve the prediction accuracy. Non-stationary time series wavelet decomposition into various sub-series prepares an interpretation of the signal construction and extracts important knowledge about its history with the employ of just a few coefficients. The WNF model considers periodic and stochastic characteristics of suspended sediment phenomenon and may provide suitable constructions not clearly seen in the suspended sediment event.

Data pre-processing technique warrants further investigation. In fact it should be noted that in general, and in Pecos River basin in particular, river discharge and SSL time series are characterized by high non-linearity and non-stationarity. NF models may simply be unable to cope with these two different features if preprocessing of the input and/or output data is not performed. Tests undertaken on data preprocessed using a wavelet transformation
showed that the best results were obtained with the WNF model. Overall these results provide evidence of the promising role of combining data clustering and discrete wavelet transforms in SSL prediction.

The results of this paper illustrated the advantage of WNF model to NF approach in simulation of suspended sediment time series.

References:


