Modeling and analysis of effects of precipitation and vegetation coverage on runoff and sediment yield in Jinsha River Basin

Jun DU*, 1, 2, Chang-xing SHI2, Chen-di ZHANG1

1. State Key Laboratory of Hydroscience and Engineering, Tsinghua University, Beijing 100084, P. R. China
2. Key Laboratory of Water Cycle and Related Land Surface Processes, Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences, Beijing 100101, P. R. China

Abstract: This paper focuses on the effects of precipitation and vegetation coverage on runoff and sediment yield in the Jinsha River Basin. Results of regression analysis were taken as input variables to investigate the applicability of the adaptive network-based fuzzy inference system (ANFIS) to simulating annual runoff and sediment yield. Correlation analysis indicates that runoff and sediment yield are positively correlated with the precipitation indices, while negatively correlated with the vegetation indices. Furthermore, the results of stepwise regression show that annual precipitation is the most important factor influencing the variation of runoff, followed by forest coverage, and their contributions to the variation of runoff are 69.8% and 17.3%, respectively. For sediment yield, rainfall erosivity is the most important factor, followed by forest coverage, and their contributions to the variation of sediment yield are 49.3% and 24.2%, respectively. The ANFIS model is of high precision in runoff forecasting, with a relative error of less than 5%, but of poor precision in sediment yield forecasting, indicating that precipitation and vegetation coverage can explain only part of the variation of sediment yield, and that other impact factors, such as human activities, should be sufficiently considered as well.

Key words: precipitation; vegetation coverage; runoff; sediment yield; adaptive network-based fuzzy inference system (ANFIS); Jinsha River Basin

1 Introduction

The issue of runoff and sediment yield and their factors is one of the hot spots in hydrological science. Previous studies show that runoff and sediment yield in a river basin are mainly affected by local natural conditions, such as precipitation, vegetation coverage, terrain, lithology, and soil structure, as well as human activities (Zingg 1940; Wischmeier and Smith 1978; Oztas et al. 2003; Xu 2006). Generally, terrain, lithology, and soil structure are relatively...
stable, while precipitation and vegetation coverage are subject to climate change and human activities. Therefore, research on variation of precipitation and vegetation coverage is of great significance to the variation of regional runoff and sediment yield over a relatively short period.

Vegetation coverage affects runoff and sediment yield in the following three ways: (1) vegetation above ground can reduce runoff and splash erosion effectively by weakening the gravitational energy of raindrops and retaining a portion of rainfall (Dou 1975); (2) surficial litter can improve soil structure, increase surface roughness, and intercept some rainfall (Yu et al. 1997); and (3) the root system of plants can reduce runoff and improve soil stability by enhancing the porosity and infiltration of soil (Sun 1989; Li et al. 1998; Gyssels and Poesen 2003).

Precipitation is the source for runoff generation in most river basins. Studies show that the total amount, intensity, energy, and duration of rainfall are all directly correlated with runoff (Guan 1996; Wu 1996). In addition, precipitation is the source power for sediment yield. The basic sediment processes, such as soil erosion and sediment transport, cannot occur or proceed without rainfall.

Although there is a relatively clear understanding of the correlation at the micro-scale between precipitation, vegetation coverage, runoff, and sediment yield, there is no widely accepted conclusion from the perspective of river basins. Most scholars point out that runoff and sediment yield increase with the amount of rainfall, and decrease with the increase of regional vegetation coverage (Liu and Zhong 1978; Singer and Bissonnais 1998; Casermeeiro et al. 2004; Al-Seikh 2006; Merzer 2007; Yu et al. 2006; Xin et al. 2009; Mohammad and Adam 2010). Others argue that runoff is positively correlated with vegetation coverage, especially in semi-humid and humid regions of China (Ma 1987; Cheng 1999). As for the upper reaches of the Yangtze River, most studies indicate that vegetation can significantly decrease the peak discharge of short-term floods but increase annual runoff (Jin 1989; Ma 1987; Chen 1999; Xu 2000), and that there is no significant correlation between sediment yield and vegetation coverage (Shi and Du 2009). As the main stream of the upper Yangtze River, the Jinsha River flows through the transitional area from semi-arid to humid zones. The relationships between vegetation coverage, runoff, and sediment yield in this basin may come against the common viewpoint. We tried to discuss the effects of precipitation and vegetation coverage on runoff and sediment yield in the Jinsha River Basin in this study.

The adaptive network-based fuzzy inference system (ANFIS) is the combination of artificial neural networks (ANNs) and the fuzzy inference system (Takagi and Sugeno 1985; Jang 1993). Compared with the black box system of ANNs, ANFIS needs less data training and performs better in forecasting with its rule-based inference mechanism. The Jinsha River Basin is too large and complicated for a normal physical model, and the data used in this study are not sufficient to establish a physical model. Thus, the grey box ANFIS model may be a proper
method for simulation and forecasting in this area. Considering that previous studies on hydrological modeling by ANFIS were based only on a single data sequence of precipitation or runoff (Chang and Chang 2006; Mahmut and Mahmud 2008; Ma and Hu 2008; Talei et al. 2010). The other purpose of this study was to investigate the capability of ANFIS of simulating annual runoff and sediment yield based on precipitation and vegetation coverage.

2 Study area

Being the uppermost part of the Yangtze River, the Jinsha River has a length of 3486 km, accounting for 77% of the length of the upper Yangtze River, and a drainage area of $4.8 \times 10^5$ km$^2$ (upstream of the Pingshan Hydrological Station), accounting for 50% of the area of the upper Yangtze River. The location of the Jinsha River and its drainage area is shown in Fig. 1. The north of the river basin is a part of the Qinghai-Tibet Plateau, and the south is the west margin of the Sichuan Basin. Since it crosses these two terrain steps of China, the Jinsha River has a height difference of up to 5142 m, which accounts for 95% of that of the whole Yangtze River Basin. Gravitational erosion, such as collapses and landslides, occurs frequently in the region with dense faults. The average annual sediment yield in the Jinsha River is $2.56 \times 10^8$ t, accounting for 48.8% of that in the upper Yangtze River. Therefore, the Jinsha River is the main sediment source area in the upper Yangtze River.

![Fig. 1 Location of Jinsha River Basin and its drainage area](image)

The climate of the river basin shows typical spatial heterogeneity. The northern area on the Qinghai-Tibet Plateau is characterized by a typical continental climate, while large parts of
the southern area are affected by monsoons. Moreover, a vertical climate difference is also observable. The average annual precipitation of the river basin is 750 mm, 90% of which occurs from May to October. However, the annual precipitation of the dry and hot valley area in the downstream area of the Jinsha River is no more than 400 mm. Limited by the unfavorable climate condition, vegetation in this area generally shows poor growth, especially in the dry and hot valley area. Most of the forest areas are distributed in the southern monsoon region. The average annual runoff of the Jinsha River Basin is 152.2 km³, accounting for 34.7% of that of the upper reaches of the Yangtze River.

3 Data processing

In this study, the data sequences of annual runoff, sediment yield, vegetation coverage, and precipitation from 1981 to 2006 were adopted for quantitative analysis and simulation, and the types of land use were also considered.

3.1 Precipitation data

The total amount and intensity of rainfall are the main factors influencing runoff and sediment yield. Meanwhile, sediment yield is closely correlated with rainfall erosivity, which is the product of rainfall intensity and rainfall kinetic energy according to the work of Wischmeier and Smith (1958). We selected annual precipitation and rainfall erosivity as two indices to study the influence of precipitation on the runoff and sediment yield of the Jinsha River Basin. The simplified formula for rainfall erosivity $R$ is as follows (Arnoldus 1977):

$$R = \frac{12}{1.735 \times 10^{1.5 \log \frac{p^2}{p_i} - 0.818}}$$

where $p$ and $p_i$ are the average annual precipitation and monthly precipitation, respectively. The data for annual precipitation and rainfall erosivity of the Jinsha River Basin are derived from the interpolation of precipitation data from 154 weather stations in the upper reaches of the Yangtze River, collected from the database of the China Meteorological Administration.

3.2 Vegetation data

This study adopted the normalized difference vegetation index (NDVI) as the level indicator to detect the growing condition of plants. NDVI is calculated from the red and near-infrared bands of remote sensing data (Walsh et al. 2001), and is widely used to analyze the regional status of vegetation coverage. The NDVI data in this study were taken from NOAA/AVHRR, and updated every 15 days, with its original image resolution being 8 km. Since the regional runoff and sediment yield mainly occur in the flood season (from June to October), we selected the annual average NDVI data from June to October as the basic NDVI database. Moreover, the land use types, especially those accounting for a large proportion of the study area, could also influence runoff and sediment yield. The annual NDVI data of forest
(N_f) and grassland (N_g) in the Jinsha River Basin were extracted from the basic NDVI database using ArcGIS technology according to the regional land use type in 1980, 1995, and 2000. The land use data were collected from the Data Center for Resources and Environmental Sciences of the Chinese Academy of Sciences.

3.3 Annual runoff and sediment yield

The annual runoff and sediment yield data from the Pingshan Hydrological Station, the control hydrological station of the Jinsha River Basin, used in this study were collected from the sediment bulletins, which are published by the Changjiang Water Resources Commission every year.

4 Methods

Regarding the river basin as a system of hydrological and erosional processes with annual precipitation (P and R) and vegetation coverage (N_f and N_g) conditions as the input variables and annual runoff (R_F) and sediment yield (S_Y) as the output variables, we attempted to establish a quantitative relationship between the output and input variables based on Pearson correlation analysis and stepwise regression analysis. Furthermore, taking the precipitation and vegetation indices selected from the analysis above as the input variables, we tried to simulate annual runoff and sediment yield using the ANFIS model. Considering that the Pearson correlation analysis is a commonly used statistical method, the introduction herein is only given for the less commonly used methods: the stepwise regression and ANFIS model.

4.1 Stepwise regression

Stepwise regression, a classical statistical technique, is a method for adding independent variables to and deleting them from a multiple linear model based on their statistical significance in a regression analysis. The method starts with an initial model, which includes a subset of all candidate independent variables. Then, the candidate variables are added to the model step by step if the P-value of the added variable is less than an entrance tolerance. If there are some variables in the model with P-values greater than an exit tolerance, the one with the largest P-value is removed, and the procedure turns back to the previous step. The procedure terminates when the model cannot be improved or when a specified maximum number of steps has been reached. The advantage of stepwise regression is that it can obtain the best combination of independent variables to predict dependent variables without causing an impact on inter-correlations among independent variables.

4.2 ANFIS model

ANFIS has proven to be useful in modeling many hydrological processes, especially in rainfall-runoff time series modeling. It is the combination of ANNs and the fuzzy inference system. The fuzzy inference system can be treated as a rule-based grey box system, in which
several inputs are fuzzified through membership functions at first, and then processed by the if-then rule operator, and an output is obtained after the defuzzification process in the end. In the conventional fuzzy inference system, the membership function parameters are defined based on artificial choices, and sometimes the process is arbitrary and complicated. ANFIS incorporates the learning ability of ANNs, either with the back propagation algorithm or hybrid algorithm, to adjust the membership function parameters automatically according to the given input and output data. Compared with the black box system of ANNs, ANFIS needs less training data and performs better in forecasting with its rule-based inference mechanism.

There are two types of fuzzy inference methods: the Takagi-Sugeno inference method and the Mamdani inference method, which are different in their definitions of consequences. The outputs of the Takagi-Sugeno method are either linear or constant (Takagi and Sugeno 1985). Though the Mamdani method is commonly used, only the Takagi-Sugeno method works in ANFIS.

It is assumed that the classic Takagi-Sugeno inference system includes two inputs, $x$ and $y$, and one output, $z$. For the inference system with linear consequence, its if-then rules can be expressed as follows: rule 1: if $x$ is at grade $A_1$, and $y$ is at grade $B_1$, then $z_1 = p_1 x + q_1 y + r_1$; and rule 2: if $x$ is at grade $A_2$, and $y$ is at grade $B_2$, then $z_2 = p_2 x + q_2 y + r_2$. $A_i$ and $B_i$ are the membership grades characterized by corresponding membership functions $\mu_{A_i}$ and $\mu_{B_i}$; and $p_i$, $q_i$, and $r_i$ ($i = 1, 2$) are the parameters. For the inference system with constant consequence, the output $z$ is constant (i.e., $p_i = q_i = 0$). The structure of ANFIS consists of five layers:

Layer 1 (input nodes): Each node in this layer generates the membership grade for a corresponding input variable. The output of the four nodes is defined as

$$
O_{li} = \mu_{A_i}(x), \quad O_{l(i+2)} = \mu_{B_i}(y) \quad i = 1, 2
$$

Layer 2 (rule nodes): The firing strength of the $i$th rule is obtained by multiplying the incoming signals from the previous layer, and the output ($w_i$) of node $i$ is defined as

$$
O_{2i} = w_i = \mu_{A_i}(x) \mu_{B_i}(y) \quad i = 1, 2
$$

Layer 3 (average nodes): In this layer, the ratio of each rule’s firing strength is calculated against the sum of all rules’ firing strength. The output of node $i$ is defined as

$$
O_{3i} = \frac{w_i}{\sum_i w_i} \quad i = 1, 2
$$

Layer 4 (consequence nodes): The contribution of the $i$th rule to the output function is calculated for node $i$ in this layer, and the function is defined as
Layer 5 (output node): In the last layer, the overall output of the single node is calculated by summing all the incoming signals generated at the previous layer. The output of the system is defined as

\[ O_i = \sum w_i z_i = w_i \left( p_i x + q_i y + r_i \right) \quad i = 1, 2 \tag{5} \]

In this study, the model was implemented in the ANFIS editor GUI of the fuzzy logic toolbox in MATLAB.

5 Results and discussion

5.1 Description of dynamically changing trends

The annual variation curves of \( N_f, N_g, P, R, S_Y, \) and \( R_F \) in the Jinsha River Basin are presented in Fig. 2. Most of the curves, such as \( N_f, N_g, R, \) and \( S_Y, \) show decreasing trends, while the rest show increasing trends according to their correlation coefficients with corresponding years (Table 1), but all these trends are statistically insignificant. The analysis of the coefficients of variation of different variables indicates that the fluctuations of all variables, especially \( N_f, N_g, \) and \( P, \) are very small except \( S_Y \) (Table 1).

![Fig. 2 Annual data of different variables in Jinsha River Basin from 1981 to 2006](image)

<table>
<thead>
<tr>
<th>Variable</th>
<th>Coefficient of correlation</th>
<th>Coefficient of variation</th>
<th>Variable</th>
<th>Coefficient of correlation</th>
<th>Coefficient of variation</th>
</tr>
</thead>
<tbody>
<tr>
<td>( N_f )</td>
<td>-0.277</td>
<td>0.062</td>
<td>( R )</td>
<td>-0.143</td>
<td>0.157</td>
</tr>
<tr>
<td>( N_g )</td>
<td>-0.067</td>
<td>0.048</td>
<td>( S_Y )</td>
<td>-0.250</td>
<td>0.337</td>
</tr>
<tr>
<td>( P )</td>
<td>0.151</td>
<td>0.072</td>
<td>( R_F )</td>
<td>0.285</td>
<td>0.159</td>
</tr>
</tbody>
</table>

The annual sediment yield of the Jinsha River Basin was comparatively stable before 1997, but the curve has dropped sharply since 1998. This variation is consistent with that of
precipitation (Fig. 2). However, human activities, such as the operation of the Ertan Reservoir and road construction, also play an important role in the variation of regional sediment yield apart from the impact of climate (Zhang and Wen 2002; Du et al. 2010). Compared with the curve of sediment yield, the fluctuation of the runoff curve is much smaller, and the total amount of annual runoff is not affected significantly by human activities in the upper reaches of the Yangtze River (Zou et al. 2007).

5.2 Correlation analysis of runoff and sediment yield in Jinsha River Basin

Considering the increasing influence of human activities on regional sediment yield in recent years (Du et al. 2010), the annual data of sediment yield from 2001 to 2006 was not considered in the correlation and regression analysis in order to better identify the quantitative relationships between precipitation, vegetation coverage, runoff, and sediment yield.

Correlation analysis was applied to the variables of $S_Y$, $R_f$, $P$, $R$, $N_f$, and $N_g$, and the results are shown in Table 2. It can be seen that runoff and sediment yield are significantly correlated with the precipitation and vegetation indices in the Jinsha River Basin, and that runoff and sediment yield are positively correlated with the precipitation indices, while negatively correlated with the vegetation indices. The correlations between precipitation indices, runoff, and sediment yield are all statistically significant, with significance levels less than 0.01. The index $P$ is more closely related to $R_f$, while the index $R$ is more closely related to $S_Y$. By contrast, only the correlations of $N_f$ with $R_f$ and $S_Y$ are statistically significant in the vegetation indices.

**Table 2** Correlation coefficients between different indices in Jinsha River Basin

<table>
<thead>
<tr>
<th>Variable</th>
<th>$S_Y$</th>
<th>$R_f$</th>
<th>$P$</th>
<th>$R$</th>
<th>$N_f$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$R_f$</td>
<td>0.759**</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$P$</td>
<td>0.771**</td>
<td>0.918**</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$R$</td>
<td>0.811**</td>
<td>0.677**</td>
<td>0.718**</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$N_f$</td>
<td>−0.522*</td>
<td>−0.525**</td>
<td>−0.430</td>
<td>−0.382</td>
<td></td>
</tr>
<tr>
<td>$N_g$</td>
<td>−0.320</td>
<td>−0.253</td>
<td>−0.237</td>
<td>−0.139</td>
<td>0.692**</td>
</tr>
</tbody>
</table>

Note: ** means that the significance level is less than 0.01, and * means that the significance level is less than 0.05.

In addition, the precipitation indices are also negatively correlated with the vegetation indices, and the correlation between $P$ and $N_f$ is statistically significant, with a significance level less than 0.05. It seems that the larger the amount of rainfall is, the worse the vegetation growth is. Some scholars have shown that the large amount of rainfall in this area increases the amount of clouds, which reduces the solar radiation and decreases the temperature. Thus, the vegetation growth is inhibited to some extent (Zhang et al. 2009).

5.3 Stepwise regression analysis of runoff and sediment yield

Correlation analysis can determine the basic quantitative relationships between
precipitation, vegetation coverage, runoff, and sediment yield in the Jinsha River Basin, but the multi-collinearity among different variables is also evident, which may impact the accuracy of results. In order to eliminate this kind of influence, the stepwise multiple linear regression method is used to relate $R_F$ and $S_Y$, respectively, with the indices $P$, $R$, $N_f$, and $N_g$ as the independent variables.

The results of stepwise regression analysis are consistent with those of correlation analysis, but the quantitative relationships of precipitation and vegetation coverage with runoff and sediment yield are more definite (Table 3). For the runoff of the river basin, $P$ is the most important factor, followed by $N_f$. The index $P$ has a positive impact on runoff, while the impact of $N_f$ is negative according to the standardized coefficients. The adjusted multiple correlation coefficient ($R^2$) of the regression model is 0.871, revealing that the independent variables could explain 87.1% of the variation of the dependent variable. Based on this value, the contributions of $P$ and $N_f$ to the variation of runoff are 69.8% and 17.3%, respectively. As for the sediment yield of the river basin, $R$ is the most important factor, and $N_f$ takes the second place. The impact of $R$ on sediment yield is positive, while that of $N_f$ is negative. The independent variables could explain 73.5% of the variation of sediment yield, and the contributions of $R$ and $N_f$ are 49.3% and 24.2%, respectively.

<table>
<thead>
<tr>
<th>Dependent variable</th>
<th>Independent variable</th>
<th>Standardized coefficient</th>
<th>t-value</th>
<th>Significance level</th>
<th>Number of data</th>
<th>Adjusted $R^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$R_F$</td>
<td>$P$</td>
<td>0.832</td>
<td>10.235</td>
<td>0.000</td>
<td>26</td>
<td>0.871</td>
</tr>
<tr>
<td></td>
<td>$N_f$</td>
<td>-0.206</td>
<td>-2.532</td>
<td>0.019</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$S_Y$</td>
<td>$R$</td>
<td>0.669</td>
<td>4.546</td>
<td>0.000</td>
<td>20</td>
<td>0.735</td>
</tr>
<tr>
<td></td>
<td>$N_f$</td>
<td>-0.328</td>
<td>-2.233</td>
<td>0.042</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Though the indices $R$ and $P$ are all significantly correlated with $R_F$, the result of stepwise regression analysis indicates that $P$ in climatic indices is the factor that really explains the variation of runoff. This result may relate to local runoff yield mechanisms. Runoff can only be generated under the saturated underlying conditions because of the regional thin soil layer and low permeability of bedrock (Xu 2000). Thus, the amount of rainfall plays a more important role in runoff generation than rainfall erosivity, which can reflect rainfall intensity to some extent. As for sediment yield, the index $R$ has a substantial impact on its variation. This is because the processes of soil erosion and sediment yield are more closely related to the energy and intensity of rainfall.

The correlation and regression analyses show that the index $N_g$ has little impact on runoff and sediment yield in the Jinsha River Basin. This is understandable when the regional land use pattern in 2000 is presented (Fig. 3). It was found that grasslands were mainly distributed in the northern part of the river basin (upstream of the Batang Hydrological Station), with an area of
0.19 × 10^6 km², accounting for 37.5% of the whole river basin, but output only 17.9% and 5.7% of the runoff and sediment yield of the whole river basin, respectively (Pan 1999). This means that the northern river basin dominated by grasslands was not the main source of runoff and sediment yield. Therefore, the runoff and sediment yield of the whole river basin cannot be significantly related to \( N_g \). However, the situation of forest lands in the Jinsha River Basin was different. The southern part of the river basin, where large amounts of runoff and sediment were generated, was mainly covered by different kinds of forests, most of which were natural forests and plantations, with a coverage density of more than 30%. This makes it possible for the forest growth to impact the runoff and sediment yield of the whole river basin. Besides, in the dry and hot valley of the southern river basin, the amount of annual evaporation was at least three times higher than that of rainfall, which meant that potential evaporation was great, and that local trees could encourage regional evaporation easily. Thus, runoff could be significantly reduced under the impact of forests (Liu and Zhong 1978). Forests reduce sediment yield not only by suppressing runoff, but also by their other effects on soil erosion. Therefore, the impact of forests on sediment yield is larger than that on runoff.

![Fig. 3 Land use patterns of forests and grasslands in Jinsha River Basin in 2000](image)

5.4 Application of ANFIS model to simulation of runoff and sediment yield

According to the results described above, the indices of \( P \) and \( N_f \) and the indices of \( R \) and \( N_f \) were selected as the input variables of ANFIS, respectively, for simulating the recent runoff and sediment yield of the river basin. The sediment yield from 2001 to 2006 can hardly be simulated without consideration of human impacts. To achieve this object, the annual amount of sediment trapped by the Ertan Reservoir during this period was considered according to the work of Feng et al. (2008), which provided the calculated results of the annual amount of sediment trapped from 2001 to 2004, and the data from 2005 to 2006 were obtained by the ratio of the average annual amount of sediment trapped to the total sediment yield of the river basin.
ANFIS consists of a rule-based system and several parameters that are related to the construction and operation of rules, such as the type and number of membership functions, the number of training epochs, and the algorithm of learning, which need to be set before modeling. The relationships between precipitation, vegetation coverage, runoff, and sediment yield are complicated, and the rules can hardly be constructed artificially. Therefore, the method of trial and error was used to find the correct set of different parameters. The method of trial and error requires that the data be divided into three categories: training, calibration, and validation. Considering that there were only 26 groups of recorded data of each input and output variables (from 1981 to 2006), the former 20 groups of the data (1981 to 2000) were used for training, and three groups of the data were selected randomly from the remaining six groups for calibration; the other three groups were used for validation.

The training process started from the normalization of the input and output variables, and the parameters of ANFIS were calibrated according to the root mean square error ($E_{RMS}$) of the estimated result as compared with the measured data in the calibration process. The training process of ANFIS was terminated when the smallest value of $E_{RMS}$ was found, and the set of parameters giving the smallest value of $E_{RMS}$ were used for simulating runoff and sediment yield.

The detailed parameter set is given in Table 4. Gaussmf and Gauss2mf were chosen as the input membership functions in models 1 and 2, respectively. The basic formulas of these two functions are the same, and the only difference between them is that Gaussmf is a fundamental Gaussian curve, while Gauss2mf is the combination of two Gaussian curves. With the excellent smoothness, symmetry, and resolution, the functions are usually used to approximate general non-fuzzy number sets. Besides, the number of membership was three and four for model 1, and three and two for model 2. Taking model 1 as an example, the value ranges of initial input variables, including $P$ and $N_f$, were divided into three and four grades, respectively, and 12 kinds of permutations and combinations (rules) of the two input variables were used for training. The more rules for training there were, the more complicated the structure of ANFIS was, with more training epochs and time consumption. The hybrid algorithm, which combines the back propagation algorithm with the method of least squares, was chosen for learning because its performance was better than that of the back propagation algorithm in most tests.

<table>
<thead>
<tr>
<th>Model</th>
<th>Input</th>
<th>Output</th>
<th>Membership function</th>
<th>Number of training epoch</th>
<th>Algorithm of learning</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>$P$</td>
<td>$R_f$</td>
<td>Gaussmf</td>
<td>3</td>
<td>180</td>
</tr>
<tr>
<td></td>
<td>$N_f$</td>
<td></td>
<td>Constant</td>
<td></td>
<td>Hybrid algorithm</td>
</tr>
<tr>
<td>2</td>
<td>$R$</td>
<td>$S_y$</td>
<td>Gauss2mf</td>
<td>3</td>
<td>140</td>
</tr>
<tr>
<td></td>
<td>$N_f$</td>
<td></td>
<td>Constant</td>
<td></td>
<td>Hybrid algorithm</td>
</tr>
</tbody>
</table>
In order to identify the reliability of the two models, the absolute error ($E_A$), relative error ($E_R$), and root mean square error ($E_{RMS}$) were used to evaluate the modeling accuracy. The mathematical expression of each index is formulated as follows:

$$E_A = |Q_e - Q_o|$$  \hspace{1cm} (7)

$$E_R = \frac{|Q_e - Q_o|}{Q_o} \times 100\%$$  \hspace{1cm} (8)

$$E_{RMS} = \left[ \frac{1}{n} \sum_{i=1}^{n} (Q_e - Q_o)^2 \right]^{0.5}$$  \hspace{1cm} (9)

where $Q_e$ and $Q_o$ represent the estimates and observations, respectively, and $n$ is the number of data points. The results simulated by the stepwise regression model are also presented for comparison.

The results of validation are satisfactory in general (Table 5). All the values estimated by the ANFIS model are more precise than those from the regression model, and $E_R$ of the ANFIS model for runoff is no more than 5%. However, the simulation accuracy for sediment yield of the ANFIS model is still very low, although the impact of the Ertan Reservoir is taken into consideration. This indicates that the mechanism of regional sediment yield is very complicated, and other impact factors, especially recent human activities, should be sufficiently considered in future studies.

Table 5 Results of validation from regression and ANFIS models in Jinsha River Basin

<table>
<thead>
<tr>
<th>$S_Y$</th>
<th>Measured $(t)$</th>
<th>Results of regression analysis</th>
<th>Results of ANFIS model</th>
</tr>
</thead>
<tbody>
<tr>
<td>Calculated $(t)$</td>
<td>$E_A (t)$</td>
<td>$E_R (%)$</td>
<td>$E_{RMS} (t)$</td>
</tr>
<tr>
<td>30 230.00</td>
<td>26 189.49</td>
<td>4 040.51</td>
<td>13.37</td>
</tr>
<tr>
<td>11 017.00</td>
<td>15 028.87</td>
<td>4 011.87</td>
<td>36.42</td>
</tr>
<tr>
<td>22 936.00</td>
<td>30 032.06</td>
<td>7 096.06</td>
<td>30.94</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$R_F$</th>
<th>Measured $(\text{km}^3)$</th>
<th>Results of regression analysis</th>
<th>Results of ANFIS model</th>
</tr>
</thead>
<tbody>
<tr>
<td>Calculated $(\text{km}^3)$</td>
<td>$E_A (\text{km}^3)$</td>
<td>$E_R (%)$</td>
<td>$E_{RMS} (\text{km}^3)$</td>
</tr>
<tr>
<td>174.200</td>
<td>161.399</td>
<td>12.801</td>
<td>7.35</td>
</tr>
<tr>
<td>154.700</td>
<td>144.971</td>
<td>9.729</td>
<td>6.29</td>
</tr>
<tr>
<td>164.800</td>
<td>155.105</td>
<td>9.695</td>
<td>5.88</td>
</tr>
</tbody>
</table>

6 Conclusions

In the Jinsha River Basin, neither the amount of annual precipitation and vegetation coverage nor the runoff and sediment yield presented remarkable tendencies from 1981 to 2006. However, the sediment yield has decreased sharply in recent years. Human activities, such as the construction of reservoirs and roads, may play an important role in this change.
Correlation analysis indicates that runoff and sediment yield are significantly correlated with the precipitation and vegetation indices, and that runoff and sediment yield are positively correlated with the precipitation indices, while the correlations of runoff and sediment yield with vegetation coverage are negative. Furthermore, the results of stepwise regression show that the annual precipitation is the most important factor in the change of runoff, followed by the forest index, and that their contributions are 69.8% and 17.3%, respectively. For sediment yield, rainfall erosivity is the most important factor, followed by the forest index, and their contributions to the variation of sediment yield are 49.3% and 24.2%, respectively.

Grasslands have little impact on the variation of runoff and sediment yield since areas covered with grasslands are not the main sources of runoff and sediment. The results of this study suggest that forests in the upper reaches of the Yangtze River reduce annual runoff to a certain extent due to the dry and hot climate in the lower Jinsha River Basin.

The performance of the ANFIS model in runoff simulation is excellent, with a relative error less than 5%, and the precision of the model in sediment yield simulation is also better than that of the conventional model, although the model for sediment yield simulation still does not meet the qualification for practice. The unsuccessful modeling for sediment yield indicates that regional precipitation and vegetation coverage can only explain part of the variation of sediment yield, and other factors, especially human activities, should be considered sufficiently in further research.

References


Xu, J. X. 2006. Effect of the changing rural socio-economic factors on sediment yield of the Jialinjiang River

(Edited by Ye SHI)