ELSEVIER

Suspended sediment flux modeling with artificial neural network: An example of the Longchuanjiang River in the Upper Yangtze Catchment, China

Yun-Mei Zhu [a,b,*], X.X. Lu [a], Yue Zhou [b]

[a] *Department of Geography, National University of Singapore, 119260, Singapore*
[b] *Department of Environmental Science, Kunming University of Science and Technology, 650093, Yunnan, China*

Abstract

Artificial neural network (ANN) was used to model the monthly suspended sediment flux in the Longchuanjiang River, the Upper Yangtze Catchment, China. The suspended sediment flux was related to the average rainfall, temperature, rainfall intensity and water discharge. It is demonstrated that ANN is capable of modeling the monthly suspended sediment flux with fairly good accuracy when proper variables and their lag effect on the suspended sediment flux are used as inputs. Compared with multiple linear regression and power relation models, ANN can generate a better fit under the same data requirement. In addition, ANN can provide more reasonable predictions for extremely high or low values, because of the distributed information processing system and the nonlinear transformation involved. Compared with the ANNs that use the values of the dependent variable at previous time steps as inputs, the ANNs established in this research with only climate variables have an advantage because it can be used to assess hydrological responses to climate change.
© 2006 Elsevier B.V. All rights reserved.

Keywords: Artificial neural network; Suspended sediment flux; Climate variables; Upper Yangtze

1. Introduction

Suspended sediment flux in a river is an important parameter for the management of hydraulic projects, and an index for the status of soil erosion and ecological environment of a catchment. Many empirically- and physically-based models have been developed to model the suspended sediment flux of a catchment. Empirical models estimate suspended sediment flux by relating it to catchment characteristics such as drainage area, topography, land cover and climate (Flaxman, 1972; Walling, 1983; Verstraeten and Poesen, 2001; Zhou et al., 2002), as well as deposition rates in ponds or reservoirs (Verstraeten et al., 2003). They are widely used because of their relatively simple structure and mathematical methods involved, and their ability to work with limited input data. Although this type of model is not able to represent the spatial variability of hydrologic processes and catchment parameters that influence the suspended sediment flux in a river, Beven (2000) argued that as far as the response of an entire water system is concerned, "it would appear that, where calibration data are available, simple lumped parameter models can provide simulations as good as those from complex physically based models". However, conventional linear or nonlinear regression models can

---

* Corresponding author. Department of Geography, Faculty of Arts and Social Science, National University of Singapore, 1 Arts Link 119260, Singapore. Tel.: +65 65163831; fax: +65 67773091.
E-mail address: zhuyunmei@nus.edu.sg (Y.-M. Zhu).

only simulate the highly nonlinear suspended sediment flux with limited accuracy, due to their simple model structure and mathematical methods employed.

Physically-based models attempt to represent the spatial heterogeneity of variables by dividing the catchment into grids, and describe the processes of the sediment transport from grid to grid with simplified partial differential equations (Wicks and Bathurst, 1996; Morgan et al., 1998; Van Oost et al., 2000). Their distributed structure allows us to evaluate the influence of land management measures on soil erosion. However, some researchers argued that the inadequate scientific basis was a major constraint for the application of distributed hydrological models to suspended sediment flux modeling (Refsgaard and Abbott, 1996). The partial differential equations used for sediment transport were based on some "unrealistic simplifying assumptions for flow and empirical relationships for erosive effects of rainfall and flow" (Kisi, 2004). Although sophisticated distribution models can provide satisfactory simulation and prediction for small and heavily instrumented catchments (usually less than 100 km^2), their applications at regional and larger scales are unrealistic, because the quantity and quality of necessary input data are usually insufficient.

Artificial neural network (ANN), a massively parallel-distributed information processing system, is based on concepts derived from research on the nature of human brains (Müller et al., 1995). It has many distinct advantages for hydrological modeling. For example, it can approximate any arbitrary continuous functions, simulate a nonlinear system without a priori assumption of processes involved, and give a good solution even when input data are incomplete or ambiguous (ASCE, 2000a,b).

ANN was introduced into hydrological modeling in the 1990s (Singh and Woolhiser, 2002) and has been successfully used in rainfall–runoff modeling (Tokar and Markus, 2000; Rajurkar et al., 2004), stream flow forecasting (Campolo et al., 1999; Liong et al., 2000; Cigizoglu, 2003), water quality assessment (Clair and Ehrman, 1998), draught analysis (Shin and Salas, 2000) and reservoir operations modeling (Golob et al., 1998).

There are, however, not many reports on the application of ANN in sediment studies, though it may also offer a promising alternative for conventional empirical and physically-based models. The research conducted by Abrahart and White (2001), Jain (2001), Tayfur (2002), Kisi (2004) and Agarwal et al. (2005) may be deemed as pathfinder experiments in this area. Details of these studies, such as variables used and temporal scales employed, are summarized in Table 1. Abrahart and White (2001) applied ANN to predict sediment flux occurring under different types of agriculture and land use management. They predicted suspended sediment loads in four catchments from rainfall, runoff, runoff coefficient and fuzzy set membership variables representing land use status. The results indicated that ANN could provide a better fit to the data than the multiple linear regression method. Tayfur (2002) used ANN to simulate experimentally observed sediment fluxes from different slopes under various rainfall intensities. The study indicated that the performance of ANN with only slope and rainfall information as inputs could be as good as that of a physically-based model with much more variables such as flow velocity, infiltration rate, shear stress, stream power, and unit stream power. Jain (2001) applied ANN to establish an integrated stage–discharge–sediment concentration relation for two sites on the Mississippi River.

Table 1
Characteristics of previous work on modeling sediment discharge with ANN

Reference	Dependent variable	Independent variables	Drainage area	Temporal scale	Time span
Abrahart and White (2001)	Sediment load (kg)	Rainfall, peak 30 min rainfall intensity, runoff, runoff/rainfall ratio, land use	Four small catchments in Bvumbwe, Malawi	Daily	117 records related to 1981~1985 rain seasons
Agarwal et al. (2005)	Suspended sediment discharge (kg s^{-1})	Rainfall, runoff, sediment yield	7820 km^2	Daily, weekly, ten-daily, monthly	10 years (1984~1989, 1992~1995)
Bhattacharya et al. (2005)	Sediment transport rate (m^3 s^{-1})	Flow velocity, depth of flow, particle size, energy slope	55 flume and 24 field observations		Experimental data
Jain (2001)	Suspended sediment concentration (mg l^{-1})	River stage, water discharge, sediment concentration	1) 1,835,276 km^2 2) 1,847,190 km^2	Daily	1) 1985~1987 2) 1990~1991
Kisi (2004)	Suspended sediment concentration (mg l^{-1})	Runoff, sediment concentration	1) 10,521 km^2 2) 13,932 km^2	Daily	5 years (1977~1981)
Tayfur (2002)	Sheet sediment transport (kg m^{-1} s^{-1})	Slope, rainfall intensity	1.52 m × 4.58 m flume	Minute	Experimental data

Five combinations of inputs including water stage, water discharge and sediment concentration at current and previous time steps were evaluated. By comparing the performance of ANN with the conventional curve-fitting method, Jain (2001) concluded that ANN could provide an estimate closer to observed suspended sediment concentration. Kisi (2004) used ANN to simulate daily suspended sediment concentration at two stations on the Tongue River in Montana, USA. He tried different combinations of inputs to predict the sediment concentration, e.g. water discharges at both current and previous time steps, sediment concentrations at previous time steps at the station of interest, as well as data from the upstream station. These studies demonstrated that the modeling of sediment, including its concentration in a river or flux from a slope or a watershed, is possible through the use of ANN. The time scale of most of the aforementioned papers is daily, except that Agarwal et al. (2005) tried to relate suspended sediment flux to rainfall and runoff on a monthly scale with limited records of a 10-year span. A common approach adopted by the above papers is that water and suspended sediment flux at previous time steps were used as inputs. It may increase the accuracy of the simulation. However, ANNs established by this method are unable to explain the contribution from climatic variables. Also, they are insufficient to predict sediment flux if water and sediment data for previous time periods are not available.

In this research, we applied ANN to simulate monthly suspended sediment flux from 1960 to 2001 in the Longchuanjiang River of the Upper Yangtze Catchment, China. Instead of using water and suspended sediment flux at previous time steps as inputs, we attempted to relate the suspended sediment flux to the original driving forces, i.e., climatic variables such as rainfall, temperature, and rainfall intensity, to establish an ANN model that can be used to explore the relationships between the climate inputs and sediment responses. The ANN constructed from climatic variables only will have a potential of filling missing data in a suspended sediment flux time series and predicting the influence of climatic change on suspended sediment flux. The advantages of ANN were also evaluated by comparing its performance with that of multiple linear regression (MLR) models and power relation (PR) models.

2. Study area and data used

2.1. Study area

The study area along the Longchuanjiang River is located at 24°45′N~26°15′N and 100°56′E~102°02′E,

Southwest China (Fig. 1). The Longchuanjiang is a tributary of the Upper Yangtze River. The length of the main channel is 231.2 km. The Huangguayuan gauging station is located at the lower reach of the river (Fig. 1), which has a drainage area of 5560 km[2]. The Longchuanjiang catchment has a subtropical monsoon climate characterized by dramatically different rainfall levels between the wet and dry seasons. The average annual rainfall is ca. 800 mm, with 86~94% of the total rainfall occurring in the wet season from May to October. Consequently, most of sediment transport in the river occurs during the wet season. There are six weather stations within or nearby the catchment (Fig. 1).

More than 50% of the area in the catchment is affected by soil erosion (Yunnan Bureau of Water Resources and Hydropower, Tianjin Survey and Design Institute, 1987). The erosion is attributed to both the natural environment and human activity. More than 60% of the catchment is covered by "purplish soil" (belongs to skeletol primitive soils in the Chinese classification by China National Soil Survey, 1992), which is very susceptible to water erosion. Furthermore, rainfall in the wet season has high intensity and frequency. In addition, there are human influences under the pressure from population growth and rapid economic development, such as deforestation, reforestation, intensified agriculture activity, reservoir building, stone excavation, and road construction. For example, forest cover in the catchment declined from 36.9% in 1949 to 24.7% in 1985, followed by a gradual recovery in the late 1990s; the total storage capacity of the reservoirs in the catchment increased from 0.28 m[3] in 1948 to 8.76 m[3] in 2002, which would contribute to sediment storage and a subsequent decrease of sediment in the river (Zhou et al., 2004; Lu, 2005) as happened in other places in the Upper Yangtze (Lu and Higgitt, 1998; Lu et al., 2003); and the road length drastically increased from 451 km in 1949 to 14801 km in 2002. All the human activities may contribute positively or negatively to the sediment concentration in the river.

The average annual water discharge and suspended sediment load at Huangguayuan are 817 million m[3] and 5.38 million tons, respectively. Both the water discharge and suspended sediment load tended to increase after 1990, but the rate of increase in the sediment load was faster (Fig. 2). This disproportional change of water and sediment could result from their nonlinear relationship and/or other factors such as land surface disturbance influencing the generation and transportation of sediment. If the effect of the non-linearity is dominant, ANN with climatic and water discharge data as inputs would be effective in modeling suspended sediment transport,
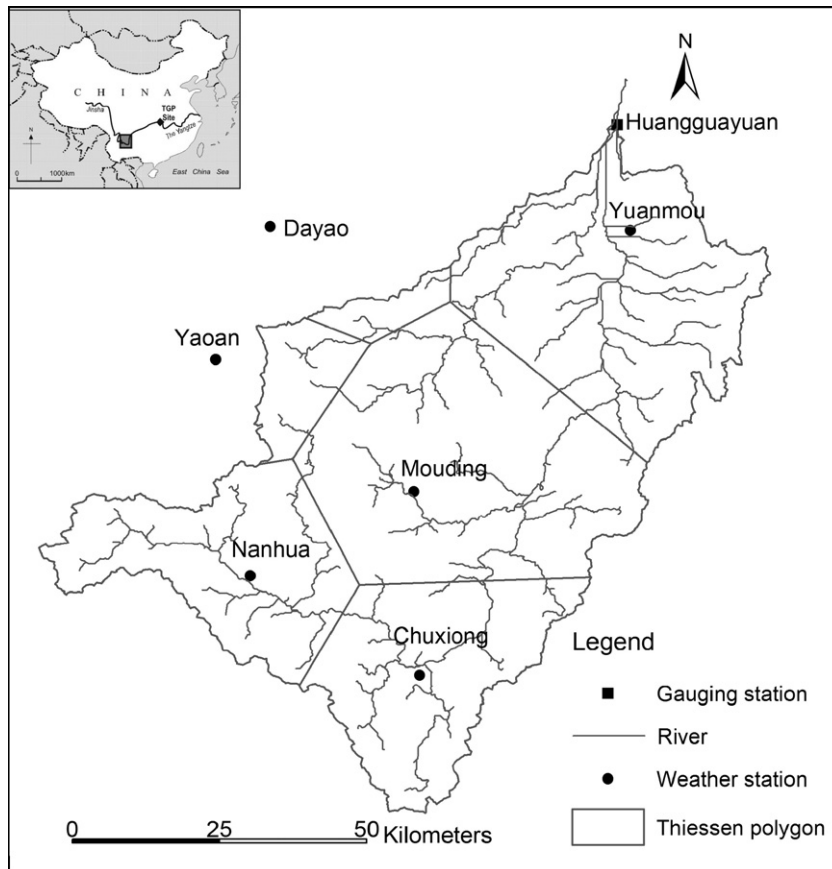
Fig. 1. Map of the Longchuanjiang catchment and the Thiessen polygons of weather stations.

but large error is expected if factors other than climate and water dominate.

2.2. Data

Daily suspended sediment flux and water discharge data at the Huangguayuan gauging station from 1960 to 2001 were collected. At the station, water level was recorded automatically 24 h a day. Suspended sediment concentration (SSC) and water discharge were measured manually once a day at 8:00 a.m., and more samples were taken during floods. Suspended sediment flux at each time point is calculated as SSC multiplied by water discharge (Chinese Ministry of Water Resources, 1999). Monthly average suspended sediment flux and water discharge were derived from the collected daily sediment and water discharge data.

The climatic inputs were selected based on physical relationships between the input and output variables and a pruning process. Rainfall is the direct driving force of sediment production and transportation. Temperature can influence the sediment generation and transportation

through several indirect ways, e.g. through its influence on the evapotranspiration, runoff and residue decomposition rate. Syvitski et al. (2003) used temperature as one of the variables, among basin area and relief, to predict the long-term sediment discharge. In the beginning of our study, four climatic variables, monthly average temperature (T), rainfall (R), evaporation (E) and humidity (H) were considered as inputs. Neural networks
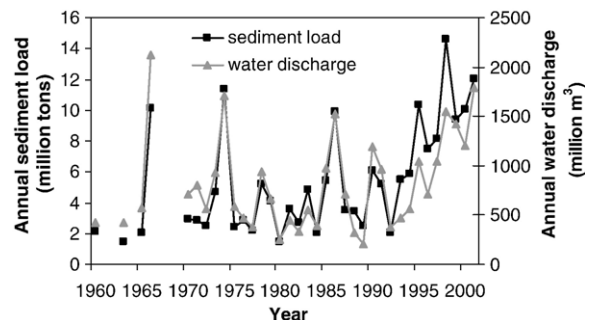

Fig. 2. Time series of annual suspended sediment load and water discharge at Huangguayuan.

Table 2
Statistical parameters of hydro-climatic data for the Longchuanjiang catchment

Statistics	T (°C)	R (mm day^{-1})	R_{25} (mm month^{-1})	R_{50} (mm month^{-1})	Q (m^3 s^{-1})	Q_s (kg s^{-1})
Max	24.3	11.5	249.5	117.3	288	2350
Min	7.4	0	0	0	0.26	0
Mean	18.5	3.1	23.21	3.87	38.08	259.81
S.D.	4.5	2.4	35.31	15.28	45.40	394.04
Coefficient of variation	0.24	0.75	1.52	3.94	1.19	1.52
Skewness	−0.91	0.60	2.12	4.11	2.16	2.40

were established with these inputs and it was found that rainfall and temperature were closely related with the sediment flux, while adding evaporation and humidity as inputs did not improve the performance of the networks. As a result, evaporation and humidity were deleted from the input list. In addition, given the influence of intensive rainfall on sediment, variables that represent the frequency of storm events (N_{25} and N_{50}: numbers of ≥25 mm and ≥50 mm rainfall days in a month, respectively) and the intensity of the events (R_{25} and R_{50}: cumulative rainfall in ≥25 mm and ≥50 mm rainfall days in a month) were also considered. However, the correlation coefficients between N_{25} and R_{25} as well as N_{50} and R_{50} are larger than 0.95. Thus, N_{25} and N_{50} were not used for analysis. Consequently, R, T, R_{25} and R_{50} were selected as input variables. The time series of daily temperature and rainfall from the six weather stations were collected, and the climatic variables were converted into the areal averaged value using the Thiessen method (Croley and Hartmann, 1985). The Thiessen polygons in the study area are shown in Fig. 1.

The hydrologic and climatic data for 36 years, 1960, 1963, 1965–1966 and 1970–2001, were used in this study. Because of the influence from the monsoon climate,

sediment fluxes from January to April were very low, accounting for less than 3% of the annual total sediment load. To study sediment fluxes in this period looks insignificant for both model evaluation and practical applications. Thus, only data in 8 months (May to December) were used for the modeling. The statistical parameters of the climatic and hydrologic variables, R, T, R_{25}, R_{50}, as well as water discharge (Q) and suspended sediment flux (Q_s) are listed in Table 2. The correlation coefficients between these variables of up to previous 3 months are given in Table 3. It can be seen that the suspended sediment flux has relatively higher linear correlations with water discharge, rainfall and R_{25}.

3. Methodology

3.1. Data processing

The original data were processed through two steps: data standardization and data set partition. The original input and output data consist of different parameters with different physical meaning and units, and thus their ranges are highly variable. To ensure that each variable is treated equally in a model, data are usually rescaled to

Table 3
Correlation coefficients (**r**) of the hydro-climatic data for the Longchuanjiang catchment

|	T	R	R_{25}	R_{50}	Q	T_{t-1}	R_{t-1}	Q_{t-1}	T_{t-2}	R_{t-2}	Q_{t-2}	T_{t-3}	R_{t-3}	Q_{t-3}	Q_s
T	1	0.594	0.357	0.179	0.253	0.882	0.261	0.000	0.267	0.191	*0.339*	*0.355*	*0.570*	*0.586*	0.352
R		1	0.732	0.384	0.707	0.698	0.467	0.128	0.477	0.024	*0.214*	0.049	*0.361*	*0.445*	0.760
R_{25}			1	0.609	0.564	0.406	0.258	0.066	0.256	*0.062*	*0.135*	*0.050*	*0.227*	*0.275*	0.672
R_{50}				1	0.238	0.184	0.000	*0.084*	0.057	*0.122*	*0.134*	*0.115*	*0.143*	*0.144*	0.320
Q					1	0.364	0.701	0.597	0.455	0.397	0.214	0.287	0.077	*0.080*	0.855
T_{t-1}						1	0.394	0.096	0.570	0.028	*0.187*	*0.014*	*0.357*	*0.470*	0.418
R_{t-1}							1	0.717	0.714	0.490	0.170	0.538	0.143	*0.101*	0.557
Q_{t-1}								1	0.387	0.712	0.615	0.466	0.458	0.288	0.326
T_{t-2}									1	0.418	0.126	0.708	0.189	*0.064*	0.406
R_{t-2}										1	0.737	0.707	0.587	0.284	0.148
Q_{t-2}											1	0.425	0.746	0.672	*0.037*
T_{t-3}												1	0.543	0.289	0.133
R_{t-3}													1	0.771	*0.135*
Q_{t-3}														1	*0.240*
Q_s															1

Italic numbers represent negative correlations.
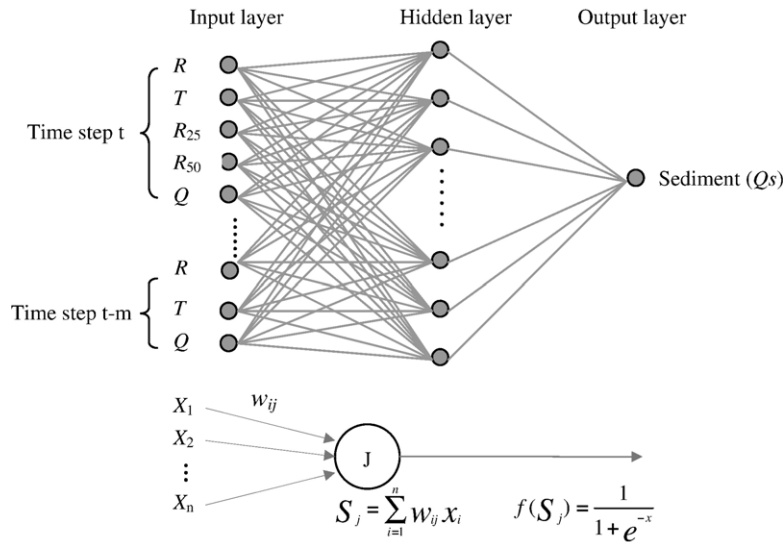
Fig. 3. Architecture of the MLP used and the schematic representation of a neuron.

a certain interval such as $[-1, 1]$, $[0.1, 0.9]$ or $[0, 1]$ (Dawson and Wilby, 2001). The advantage of using $[0.1, 0.9]$ is the accommodation of possible extreme values outside the range of the calibration data (Hsu et al., 1995; Imrie et al., 2000). The input and output variables for the present study were standardized into the interval $[0.1, 0.9]$:

$$X_{istd} = 0.1 + 0.8 \times \frac{X_i - X_{imin}}{X_{imax} - X_{imin}} \tag{1}$$

where x_{istd} is the standardized value of variable i, x_i is the original value, and x_{imax} and x_{imin} are the maximum and minimum of variable i, respectively.

For model calibration and validation, the input data are usually divided into a calibration set and a validation set. To avoid the problem of over-fitting, one more data set, a test set, was extracted and used for cross-validation of the network in this study (Anctil and Lauzon, 2004). The representativeness of the data sets used for calibration and test should be carefully considered. Networks trained with a data set that represents the characteristics of the hydrologic patterns will achieve higher generalization ability. Tokar and Johnson (1999) suggested that for rainfall–runoff modeling, the network trained with the dry-, average- and wet-year data could produce more accurate prediction than those trained with only dry- and average-year data. Given the close relationship among rainfall, runoff and sediment, this method may also be applied to data set partitioning for sediment modeling. Thus, in this study, the data were first classified into wet-, average- or dry-year data groups

according to the rainfall of each year. The number of the years in the wet group is 10, that in the average group is 12, and that in the dry group is 14. Then 15 years' data, with five from each group, were extracted randomly to form the validation set; 5 years' data, with one from the dry-year, two from the average-year and the other two from the wet-year data group, were extracted to form the test set; the rest 16 years' data were combined to form the calibration dataset.

3.2. Artificial neural networks (ANNs)

ANNs can be categorized into feed forward and recurrent networks according to the direction of the information flow and processing. Multilayer perception (MLP) is a feed forward network. Detailed information about MLP is found in literature (e.g. Müller et al., 1995; Schalkoff, 1997). Many studies have shown that MLP is a universal approximator. A MLP with one hidden layer is capable of approximating any finite nonlinear function with very high accuracy (Hornik et al., 1989; Schalkoff, 1997). The MLP in the present study consisted of an input layer, a hidden layer and an output layer. Each neuron in each layer was connected to all neurons in the adjacent layers but the information flowed only in one direction, from the input side to the output side. The architecture of the neural network used in this study and the schematic representation of a neuron are shown in Fig. 3.

The neuron in the output layer represents suspended sediment flux (Fig. 3). The number of neurons in the hidden layers was decided by a trial-and-error method.

Table 4
Performances of ANNs, MLR and PR models

Model	Group		Inputs	Calibration [a]		Testing [b]		Validation [c]	
				RMSE [d]	r^2	RMSE [d]	r^2	RMSE [d]	r^2
ANN	I	1	$(T,R)_t$	260.31	0.666	184.23	0.541	216.28	0.663
		2	$(T,R)_t, (T,R)_{t-1}$	216.18	0.767	141.92	0.729	195.39	0.765
		3	$(T,R)_t, (T,R)_{t-1}, (T,R)_{t-2}$	214.86	0.769	145.98	0.713	206.10	0.746
		4	$(T,R)_t, (T,R)_{t-1}, (T,R)_{t-2}, (T,R)_{t-3}$	214.55	0.770	145.21	0.717	202.99	0.755
	II	5	$(T,R,R_{25},R_{50})_t$	255.57	0.680	187.84	0.524	201.01	0.706
		6	$(T,R,R_{25},R_{50})_t, (T,R)_{t-1}$	202.76	0.790	141.87	0.732	182.56	0.800
		7	$(T,R,R_{25},R_{50})_t, (T,R)_{t-1}, (T,R)_{t-2}$	200.11	0.793	143.59	0.725	201.44	0.778
		8	$(T,R,R_{25},R_{50})_t, (T,R)_{t-1}, (T,R)_{t-2}, (T,R)_{t-3}$	202.66	0.793	145.05	0.720	181.82	0.784
	III	9	$(T,R,Q)_t$	186.91	0.817	120.99	0.817	164.30	0.893
		10	$(T,R,Q)_t, (T,R,Q)_{t-1}$	179.41	0.834	121.22	0.808	168.23	0.888
		11	$(T,R,Q)_t, (T,R,Q)_{t-1}, (T,R,Q)_{t-2}$	177.95	0.838	119.62	0.818	172.90	0.871
		12	$T,R,Q)_t, (T,R,Q)_{t-1}, (T,R,Q)_{t-2}, (T,R,Q)_{t-3}$	170.09	0.848	120.82	0.813	184.88	0.864
	IV	13	$(Q)_t$	239.97	0.696	145.13	0.716	200.13	0.836
		14	$(Q)_t, (Q)_{t-1}$	212.85	0.767	147.85	0.714	166.65	0.882
		15	$(Q)_t, (Q)_{t-1}, (Q)_{t-2}$	194.89	0.803	128.54	0.804	182.00	0.857
		16	$(Q)_t, (Q)_{t-1}, (Q)_{t-2} (Q)_{t-3}$	193.78	0.806	126.46	0.806	173.67	0.865
MLR	A		$(T,R)_t, (T,R)_{t-1}$	232.34	0.668			227.08	0.690
	B		$(T,R,R_{25},R_{50})_t, (T,R)_{t-1}$	224.10	0.691			215.64	0.721
	C		$(T,R,Q)_t$	196.75	0.762			184.69	0.862
PR			$(Q)_t$	535.83	0.569			1135.58	0.631

[a] Number of records is 128 for ANNs and 168 for MLR and RC models.
[b] Number of records is 40.
[c] Number of records is 120.
[d] In kg s^{-1}.

Neurons in the input layer represent input variables. In this study, sixteen input combinations, which fell in four groups, were used (Table 4). The networks in different groups were designed to compare the performances of different sets of causal variables; while those in the same group were designed to examine the degree of lag-effect between the inputs and the outputs. The networks in Group I (Table 4), ANN_1 to ANN_4, used rainfall (R) and temperature (T) only as causal variables. In ANN_1, R and T of the current month were used as inputs. R and T in previous 1, 2, and 3 months were added as extra inputs to form ANN_2, _3, _4, respectively, to consider the lag effect between the causal variables and the sediment output. In Group II, in view of the influence of intensive rainfall on the sediment, R_{25} and R_{50} at the current month were added as extra inputs. Networks in Group III used both climatic and hydrologic variables as inputs. They were constructed by adding water discharge as additional causal variables to the networks in Group I. Networks in Group IV tried to predict sediment from water discharge only.

The networks were trained with a Back-Propagation (BP) algorithm. BP training involves information processing in two directions, the feed forward of the input information and the back-propagation of the error. The

input information is processed in the neurons of the input layer and is passed down to the next layer through the links. Each neuron calculates its net input (S_j) as the weighted sum of all inputs, as defined in Fig. 3. Before the information is passed to the next layer, a further transformation is conducted to amplify or inhibit neuron's net input through a transformation function, $f(S_j)$, associated with each neuron. A logistic sigmoid function, which was widely used in current hydrological modeling, was chosen as the transformation function (Fig. 3). When the information processes reach the last layer and the final output is produced, an error indicating the difference between the predicted and observed outputs is computed. This error is back-propagated through the network to each neuron, and correspondingly the connection weights are adjusted based on the steepest gradient descent of the error function, with the direction vector being set equal to the negative of the gradient vector:

$$\Delta W_{ij}(n) = -\varepsilon \frac{\partial E}{\partial W_{ij}} + \alpha \Delta W_{ij}(n-1) \qquad (2)$$

where $\Delta W_{ij}(n)$ and $\Delta W_{ij}(n-1)$ are weight increments between neuron i and j during the n-th and $(n-1)$-th
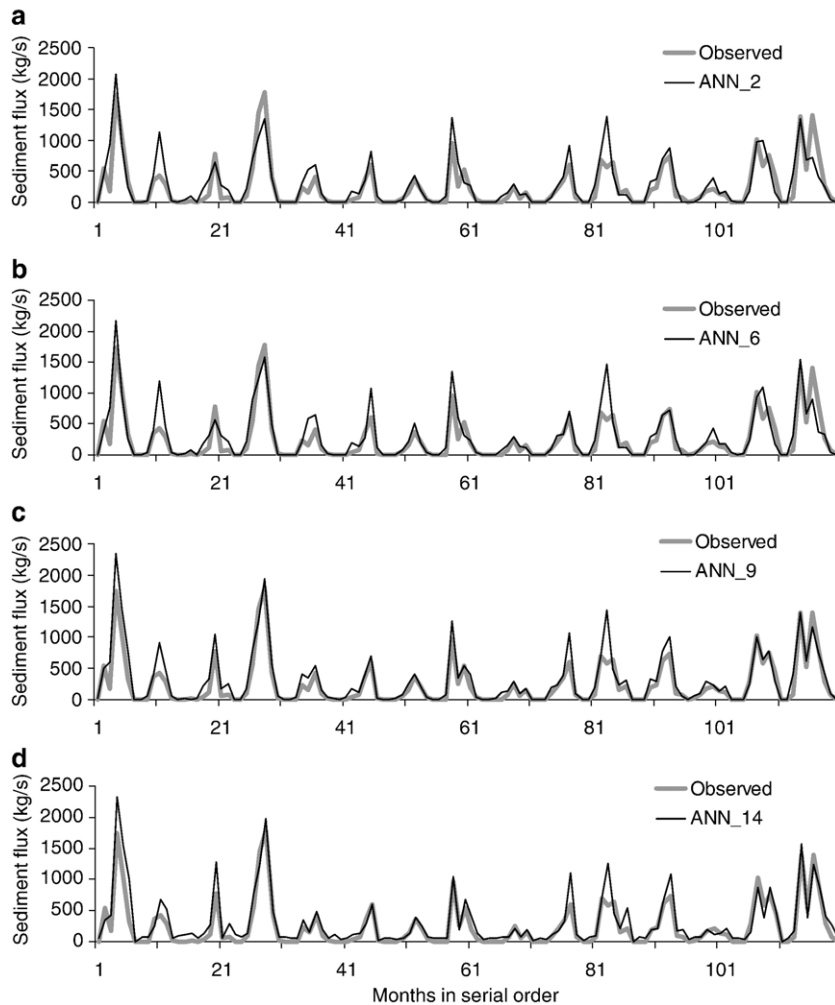
Fig. 4. Comparisons between the observed and predicted sediment fluxes based on validation data: (a) ANN_2, (b) ANN_6, (c) ANN_9 and (d) ANN_14.

epoch (or round), ε is the learning rate, and α is the momentum. In this study, the training was done with the software Neuroshell 2. The training process was stopped when the average error of the network started to rise. The network was recorded on the best test result.

After the networks had been calibrated and validated, their performances were assessed with two statistics, root mean square error (RMSE) and coefficient of determination (r^2). The best performing networks of each group were identified and their input combinations were used for multiple linear regression models.

3.3. Multiple linear regression (MLR) and power relation (PR) models

Multiple linear regression (MLR) and power relation (PR) models were employed to simulate the relationship between the input variables and the suspended sediment flux. The MLR model is denoted as:

$$\hat{Q}_{si} = b_0 + b_1 x_{i1} + b_2 x_{i2} + \cdots + b_k x_{ik} + \varepsilon_i \tag{3}$$

where x_{ij} is the j-th independent variable for the i-th pattern; ε_i is the value of random fluctuation or error for the i-th pattern; b_0 is the regression intercept and b_j is the coefficient of the j-th independent variable. Three MLR models with the same input combinations as the best performing networks from ANN Group I, Group II and Group III, respectively, were established.

The PR model in the present study took the form of the conventional rating curve model but worked on monthly scale:

$$Q_s = aQ^b \tag{4}$$

where a and b are the constants. Because the PR model stemmed from the rating curve, which is designed to study the direct relationship between water discharge and sediment flux, we used the original data for the modeling (sediment flux in kg s^{-1} and water discharge in m^3 s^{-1}), instead of the standardized ones. The performances of the MLR and PR models were also evaluated using RMSE and r^2, and were compared with those of ANNs.

4. Application and results

4.1. ANN

The RMSE and r^2 values of the sixteen ANNs during calibration, testing and validation periods are given in Table 4. Our focus of the discussion below is on the values for the validation period, because the generalization ability of the networks is of interest for the application.

The networks in Group I based only on the average rainfall and temperature show relatively poor performances. The r^2 of ANN_1, ANN_2, ANN_3 and ANN_4 during the validation period are 0.663, 0.765, 0.746 and 0.755, respectively. ANN_2, with the lowest RMSE and highest r^2, is identified as the best performing network of this group.

In Group II, the RMSE and r^2 of ANN_5 during the validation period are 201.01 kg s^{-1} and 0.706, respectively. ANN_6 shows a significant improvement, due to the inclusion of T and R for the previous month, with an RMSE of 182.56 kg s^{-1} and an r^2 of 0.800. The performances of ANN_7 and ANN_8 are not as good as the performance of ANN_6, although more information from previous months is used. ANN_6 is selected as the best performing network of Group 2.

The performances of the networks in Group III are dramatically improved, compared with those in Groups I and II. This is mainly due to the close relationship between the suspended sediment flux and water discharge. The r^2 values of ANN_9, _10, _11 and _12 during the validation period are 0.893, 0.888, 0.871 and 0.864, respectively. ANN_9, with T, P and Q only at the current month provided the best simulation.

ANN_14 is the best performing network in Group IV, with an RMSE of 166.65 kg s^{-1} and an r^2 of 0.836 in the validation period. The comparison of the observed and predicted suspended sediment fluxes for the four best performing networks, ANN_2, _6, _9 and _14, are plotted in Fig. 4. It can be observed that the pattern of sediment flux including peaks is well predicted. All the four networks provide better simulation results for years with lower suspended sediment flux.

Table 4 shows that in each group, as more inputs are used, the performances of the networks tend to increase during the calibration and testing periods, especially in Groups III and IV. However, as noted, their performances in the validation period do not show the same trend. Their performance may be improved in the first few steps, but usually drops after a certain point. For example, the best performing networks in the four groups are ANN_2, _6, _9 and _14, instead of those with larger number of inputs, suggesting that the generalization ability of a network may decrease if too many inputs are used, due to the increased complexity of the network. As long as the closely related information representing the causal variables and their lag effect is provided, the network can generate satisfactory results without using many variables.

4.2. MLR and PR models

Three MLR models, MLR_A, MLR_B and MLR_C, with the same inputs as ANN_2, ANN_6 and ANN_9, respectively, and one PR model were established (Table 5). The data sets used for calibration and testing in the ANNs were used for calibration in the MLR and PR models, and the same validation set was used for all the models.

The RMSE and r^2 of the MLR and PR models are given in Table 4. The r^2 values of MLR_A and MLR_B in the validation period are 0.690 and 0.721, respectively. MLR_C, with Q as an extra input, has the lowest RMSE (184.69 g s^{-1}) and highest r^2 (0.862).

The PR model generates a poorer prediction than the MLR models. Its r^2 is 0.631 in the validation period and its RMSE is 1135.58 kg s^{-1}. The observed and predicted suspended sediment fluxes by MLR_A, MLR_B, MLR_C and PR are plotted in Fig. 5. When the suspended sediment flux is very low, the MLR models predict negative values. The PR model overestimates the peaks.

5. Discussion

The performances of the ANN, MLR and PR models were evaluated in terms of goodness-of-fit, model

Table 5
Estimated MLR and PR models

Model	Equation
MLR_A	$Q_s = 0.22 + 0.21 * T_t + 0.725 * R_t - 0.645 * T_{t-1} + 0.254 * R_{t-1}$
MLR_B	$Q_s = 0.183 + 0.22 * T_t + 0.551 * R_t + 0.28 * R_{25} - 0.001 * R_{50}$ $- 0.608 * T_{t-1} + 0.278 * R_{t-1}$
MLR_C	$Q_s = -0.059 + 0.021 * R_t + 0.26 * T_t + 0.785 * Q_t$
PR	$Q_s = 0.031 * Q_t^{2.265}$

simplicity and data requirement. The goodness-of-fit was examined by comparing the statistics of the results (RMSE and r^2; Table 4), the scatter plots of the observed and estimated suspended sediment fluxes (Fig. 6), and the plots of observed and estimated cumulative suspended sediment loads (Fig. 7). When the performance of each neural network group was compared with the corresponding the MLR and PR models, the best performing network in each group was considered.

Group I networks have the lowest data requirement and relatively simple structure among the ANNs. However, their goodness-of-fit is lower than the others. Group II networks perform better than their counterparts in Group I, but they require daily climatic information to compile R_{25} and R_{50}. Because the Groups I and II networks use only climate data as inputs, they can be used to predict the sediment flux directly from the climate inputs and to predict the influence of climate change on the sediment flux, if other situations, e.g. land use status, can be considered as constant. The Groups III and IV networks have higher goodness-of-fit than the Groups I and II networks. However, because water discharge is used as input, they cannot be used to predict sediment flux directly from the climate input. Therefore, in ungauged catchments, one extra step is required to predict the water discharge before the sediment flux can be predicted. The error may become cumulative in such a case and final accuracy may be lower than those in Table 4.

The comparisons between ANN_2 and MLR_A, ANN_6 and MLR_B as well as ANN_9 and MLR_C based on Figs. 4 and 5 indicate that under the same data requirement or the input combination, the neural networks generate better estimation than the multiple regression models do. Both ANN_14 and PR use only
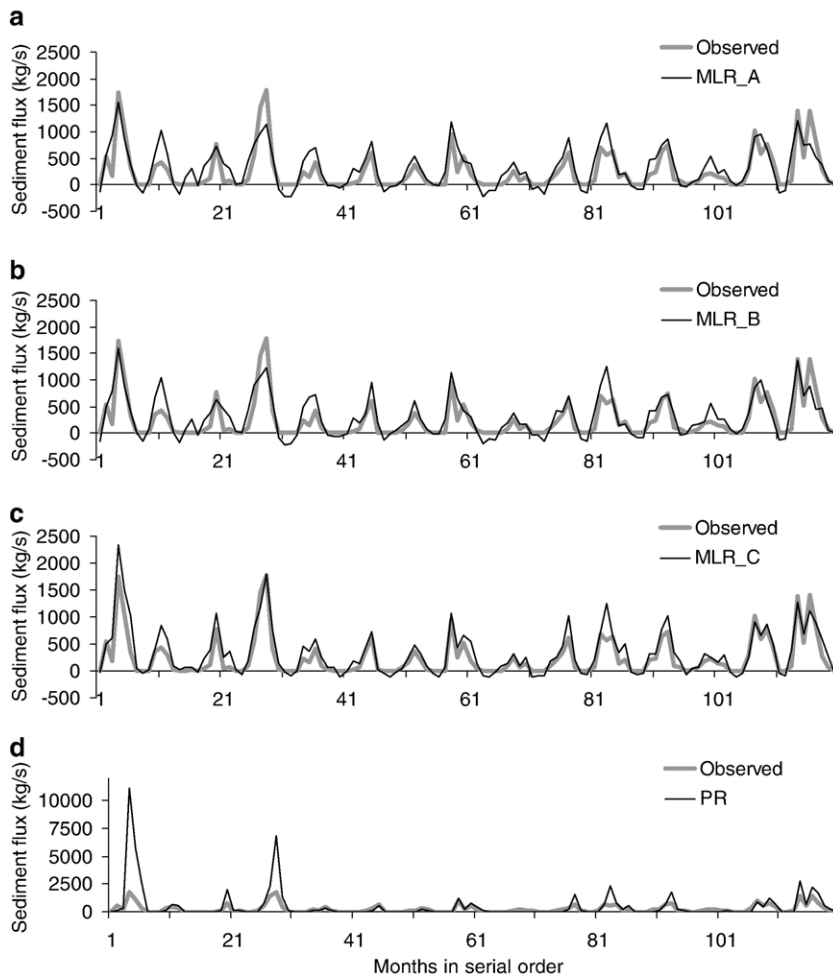

Fig. 5. Comparisons between the observed and predicted sediment fluxes based on validation data: (a) MLR_A, (b) MLR_B, (c) MLR_C, and (d) PR.
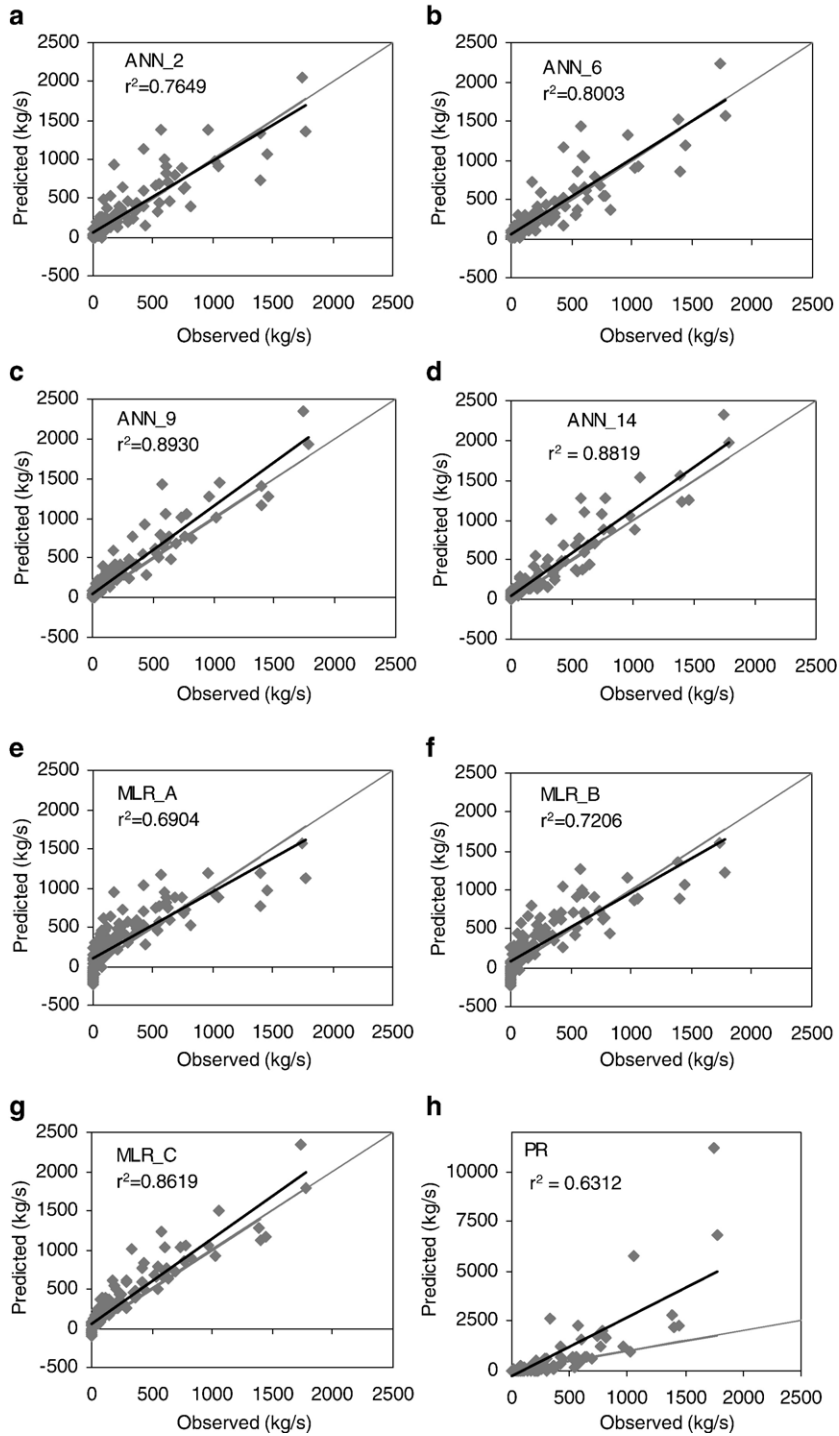
Fig. 6. Scatter plots of the observed and predicted sediment fluxes by the best performing ANN in each group and the MLR/PR models, based on validation data: (a) ANN_2, (b) ANN_6, (c) ANN_9 and (d) ANN_14, (e) MLR_A, (f) MLR_B, (g) MLR_C, and (h) PR. Gray line shows observed=predicted. Black line shows linear regression line.

water discharge to predict the sediment flux based on a nonlinear estimation. The comparison between them indicates that ANN_14 provides much better estimation.

Furthermore, the neural networks generate more reasonable predictions for the extreme values of the output variable. The MLR models may generate
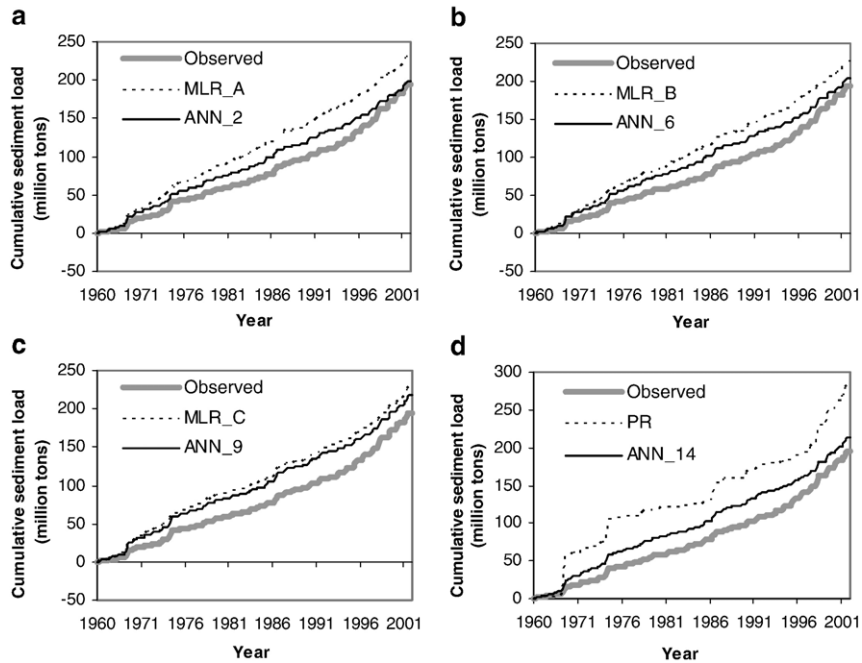
Fig. 7. Comparison between observed and predicted cumulative suspended sediment load values by the best performing ANN of each group and MLR/PR models based on validation data: (a) MLR_A and ANN_2, (b) MLR_B and ANN_6, (c) MLR_C and ANN_9, (d) PR and ANN_14.

negative values of sediment flux at the points where fluxes are zero or close to zero; whereas the neural networks can generate reasonable estimation due to the nonlinear transformation process involved. The PR model overestimates most of the peaks, and very large overestimations are found at some points. For example, the estimated value for the first peak in Fig. 5d is 6.4 times as large as the observed one. The cumulative suspended sediment load predicted by the PR model is much higher than that by ANN_14 (Fig. 7d).

The selection of the input variables also plays an important role in determining the accuracy of the network. Rainfall and temperature were selected from the originally considered climatic variables including evaporation and humidity, because they were found closely correlated with the sediment flux and could be used to represent the influence of climate. The close relationship between temperature and sediment flux may be a result of two mechanisms. First, in the study area, the average annual potential evaporation is three times as much as that of the annual rainfall, and temperature, to some degree, affects the potential evaporation. Second, temperature, especially in the previous month, is an index of soil moisture, which significantly affects soil erosion processes and resultant sediment supply to rivers. In addition to rainfall and temperature, adding variables that represent the intensity

of the storm event into the network would improve its performance, because it is directly related to erosion. The inclusion of water discharge has further improved the performance of the network.

The lag effect of the input variables is another issue that should be taken into account in constructing the network. The best performing networks in Groups I and II are those with the information of the current month and the previous 1 month, suggesting that a 1-month lag-effect exits between the climate inputs and sediment flux. In Group III, when T, R and Q are used as inputs, the current month's information is good enough for the estimation of the sediment flux, reflecting the strong influence of Q without lag time. The degree to which the information from previous time steps should be involved can be decided from the physical relationship between the inputs and the output. This research suggests that, if the inputs are directly or closely related to the output, like water discharge to sediment, no or only shorter lag effect should be considered. In other cases a longer lag effect may be required, as for rainfall and temperature.

Most previous studies that employed ANN to model water discharge or sediment flux used the values of the dependant variable at previous time steps as inputs, in addition to climate variables (Tokar and Markus, 2000; Raid and Mania, 2004), or used them as the only type of input to the network (Cigizoglu, 2003; Kisi, 2004).
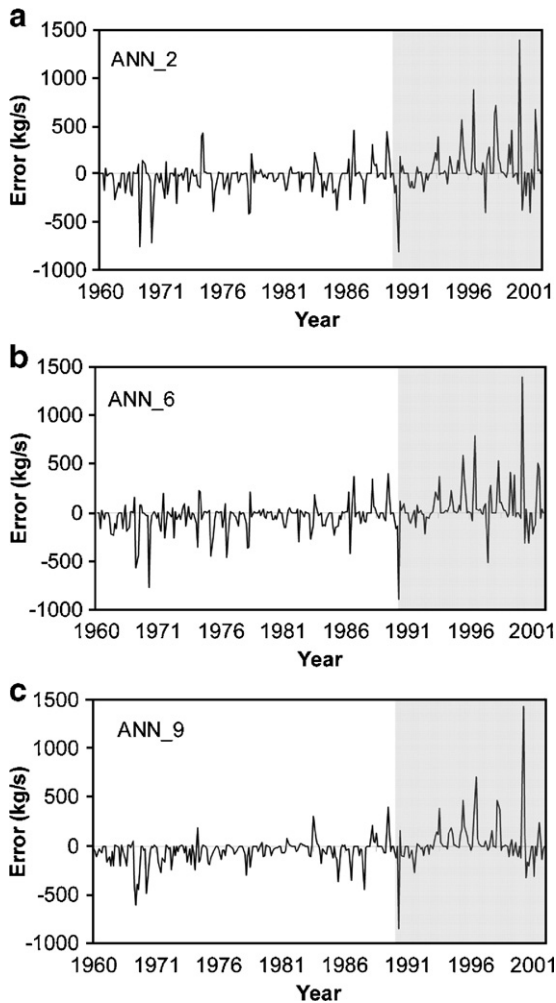
Fig. 8. Difference between observed and predicted sediment fluxes by selected ANNs: (a) ANN_2, (b) ANN_6, (c) ANN_9.

Although it has been proved that the performance of the network can be improved by including the dependent variable at previous time steps, the application of such a network is limited due to the same variable employed. This type of network actually estimates the difference between the values of the dependent variable at the current and previous time steps. More accurate result generated by this type of network may be of great interest for applications focusing on the status of water or sediment themselves, but it does not provide any information about the contribution of the driving force such as climate. Furthermore, the application of this type of network is limited to research on a short temporal scale. It is usually carried out on an hourly, daily or weekly scale, and the accuracy of its future prediction decreases sharply when the prediction time horizon increases (Campolo et al., 1999), due to limited relations

between water and sediment at a longer temporal interval.

The ANNs established in this research also provide information about the influence from factors other than climate and water discharge on the suspended sediment flux in the catchment. The records at the Huangguayuan gauging station showed that the annual suspended sediment flux increased rapidly after 1990 (Fig. 2). The annual average suspended sediment load from 1960 to 1989 was 4.02 million tons, while it increased up to 8.07 million tons for the period from 1990 to 2001. A report from the local hydraulic government attributed this increase to the increased frequency of intensive rainfall, as well as change in land use/cover (Chuxiong Hydraulic Bureau, 1997). If the increase of sediment mainly resulted from the higher rainfall amount and intensity, it can be predicted by the ANNs with average rainfall and rainfall intensity as inputs. Fig. 8 shows the errors of ANN_2, ANN_6 and ANN_9, which are computed as the differences between the observed and estimated suspended sediment fluxes. All the three networks generate consistent lower estimation for most of the months after 1990 (shaded area in Fig. 8). This under-estimation is also detectable from Fig. 7. The observed and predicted cumulative sediment fluxes by the ANNs tend to diverge with time in the beginning, but converge at the end because the cumulative overestimation in the earlier period is offset by the underestimation in the later period. This indicates that the ANNs with average climate status, rainfall intensity and water discharge only as inputs cannot fully predict or explain the increase in sediment after 1990. In other words, the relationships of the suspended sediment flux with climate and water discharge have changed since the 1990s, due to the factors which are not included as inputs in the networks. Since the early 1990s, the influence of climate variation seems to be more limited than that of human activity. Zhou et al. (2004) and Lu (2005) denoted that human activity related to land surface disturbance, such as deforestation, afforestation, intensified agriculture activity, and road construction played an important role in the recent increase of suspended sediment flux in the Longchuanjiang catchment.

6. Conclusions

Artificial neural network (ANN) was applied to predict the monthly suspended sediment flux in the Longchuanjiang catchment, by relating it to average rainfall, temperature, rainfall intensity and water discharge. It is demonstrated that ANN is capable of modeling the monthly suspended sediment flux with fairly good accuracy when proper input variables and

their lag effect on suspended sediment flux are included. ANN can generate a better fit to the observed suspended sediment flux than the multiple regression models and the power relation model under the same data requirement. The most prominent feature of ANN is that it can provide more reasonable predictions for extremely high or low values, because of the distributed information processing system and the nonlinear transformation involved. Compared with the previous ANN models which use the values of the dependent variable at previous time steps as inputs, the models established in this research with only climate variables as inputs may have lower goodness-of-fit, but they permit the assessment of hydrological responses to climate change, which is an important issue in recent years especially in relation to global warming.

The ANN models in this research were constructed under the assumption that land use/cover in the catchment has remained unchanged during the study period. However, land use/cover is an important factor for the production and transportation of sediment. A more accurate prediction of suspended sediment flux may be achieved by adding variables representing the land use/cover status into the neural network. We plan to perform such analyses for the Longchuanjiang catchment in the near future.

Acknowledgements

References

Abrahart, R.J., White, S.M., 2001. Modelling sediment transfer in Malawe: comparing backpropagation neural network solution against a multiple linear regression benchmark using small data sets. Physics and Chemistry of the Earth (B) 26, 19–24.

Agarwal, A., Singh, R.D., Mishra, S.K., Bhunya, P.K., 2005. ANN-based sediment yield models for Vamsadhara river basin (India). Water S.A. 31, 95–100.

Anctil, F., Lauzon, N., 2004. Generalisation for neural networks through data sampling and training procedures, with applications to stream-flow predictions. Hydrology and Earth System Sciences 8, 940–958.

ASCE, 2000a. Artificial neural networks in hydrology: 1. Preliminary concepts. Journal of Hydrologic Engineering 5, 115–123.

ASCE, 2000b. Artificial neural networks in hydrology: 2. Hydrology applications. Journal of Hydrologic Engineering 5, 124–136.

Beven, K.J., 2000. Rainfall–Runoff Modelling: the Primer. John Wiley, Chichester, UK, pp. 85–114.

Bhattacharya, B., Price, R.K., Solomatine, D.P., 2005. Data-driven modelling in the context of sediment transport. Physics and Chemistry of the Earth 30, 297–302.

Campolo, M., Andreussi, P., Soldati, A., 1999. River flood forecasting with a neural network model. Water Resources Research 35, 1191–1197.

China National Soil Survey, 1992. Chinese Soil Taxonomy. China Agriculture Press, Beijng. (in Chinese).

Chinese Ministry of Water Resources, 1999. Code for Hydrologic Data Compilation. Water Resources and Hydropower Publication, Beijing, pp. 47–52 (in Chinese).

Chuxiong Hydraulic Bureau, 1997. Effect of the "Changzhi" project on sediment control and reduction. Technical Report. Chuxiong (in Chinese).

Cigizoglu, H.K., 2003. Estimation, forecasting and extrapolation of river flows by artificial neural networks. Hydrological Sciences Journal 48, 349–361.

Clair, T.A., Ehrman, J.M., 1998. Using neural networks to assess the influence of changing seasonal climates in modifying discharge, dissolved organic carbon, and nitrogen export in eastern Canadian rivers. Water Resources Research 34, 447–455.

Croley, T.E., Hartmann, H.C., 1985. Resolving Thiessen polygons. Journal of Hydrology 76, 363–379.

Dawson, C.W., Wilby, R.L., 2001. Hydrological modelling using artificial neural networks. Progress in Physical Geography 25, 80–108.

Flaxman, E.M., 1972. Predicting sediment yield in western United States. Proceedings of the American Society of Civil Engineers Journal of the Hydraulics Division, vol. 98(HY12). ASCE, pp. 2073–2085.

Golob, R., Stokelj, T., Grgic, D., 1998. Neural-network-based water inflow forecasting. Control Engineering Practice 6, 593–600.

Hornik, K., Stinchcombe, M., White, H., 1989. Multilayer feedforward networks are universal approximators. Neural Networks 2, 359–366.

Hsu, K., Gupta, H.V., Sorooshian, S., 1995. Artificial neural network modeling of the rainfall–runoff process. Water Resources Research 31, 2517–2530.

Imrie, C.E., Durucan, S., Korre, A., 2000. River flow prediction using artificial neural networks: generalisation beyond the calibration range. Journal of Hydrology 233, 138–153.

Jain, S.K., 2001. Development of integrated sediment rating curves using ANNs. Journal of Hydraulic Engineering 127, 30–37.

Kisi, Ö., 2004. Multi-layer perceptions with Levenberg–Marquardt training algorithm for suspended sediment concentration prediction and estimation. Hydrological Sciences Journal 49, 1025–1040.

Liong, S.Y., Lim, W.H., Paudyal, G.N., 2000. River stage forecasting in Bangladesh: neural network approach. Journal of Computing in Civil Engineering 14, 1–8.

Lu, X.X., 2005. Spatial variability and temporal changes of water discharge and sediment flux in the lower Jinsha tributary: impact of environmental changes. River Research and Applications 21, 229–243.

Lu, X.X., Higgitt, D.L., 1998. Recent changes of sediment yield in the Upper Yangtze, China. Environmental Management 22, 697–709.

Lu, X.X., Ashmore, P., Wang, J., 2003. Seasonal water discharge and sediment flux changes in the Upper Yangtze, China. Mountain Research and Development 23, 56–64.

Morgan, R.P.C., Quinton, J.N., Smith, R.E., Govers, G., Poesen, J.W.A., Auerswald, K., Chisci, G., Torri, D., Styczen, M.E., 1998. The European Soil Erosion Model (EUROSEM): a dynamic approach

from predicting sediment transport from fields and small catchments. Earth Surface Processes and Landforms 23, 527–544.

Müller, B., Reinhardt, J., Strickland, M.T., 1995. Neural Networks: an Introduction. Springer-Verlag, New York, pp. 52–62.

Raid, S., Mania, J., 2004. Rainfall–runoff model using an artificial neural network approach. Mathematical and Computer Modelling 40, 839–846.

Rajurkar, M.P., Kothyari, U.C., Chaube, U.C., 2004. Modeling of the daily rainfall–runoff relationship with artificial neural network. Journal of Hydrology 85, 96–113.

Refsgaard, J.C., Abbott, M.B., 1996. The role of distributed hydrological modelling in water resources management. In: Abbott, M.B., Refsgaard, J.C. (Eds.), Distributed Hydrological Modeling. Kluwer Academic, Boston, pp. 1–16.

Schalkoff, R.J., 1997. Artificial Neural Networks. Mcgraw-Hill, New York, pp. 146–188.

Shin, H.S., Salas, J.D., 2000. Regional drought analysis based on neural networks. Journal of Hydraulic Engineering 5, 145–155.

Singh, V.P., Woolhiser, D.A., 2002. Mathematical modeling of watershed hydrology. Journal of Hydrologic Engineering 7, 270–292.

Syvitski, J.P.M., Peckham, S.D., Hilberman, R., Mulder, T., 2003. Predicting the terrestrial flux of sediment to the global ocean: a planetary perspective. Sedimentary Geology 162, 5–24.

Tayfur, G., 2002. Artificial neural networks for sheet sediment transport. Hydrological Sciences Journal 47, 879–892.

Tokar, A.S., Johnson, P.A., 1999. Rainfall–runoff modeling using artificial neural networks. Journal of Hydrologic Engineering 4, 232–239.

Tokar, A.S., Markus, M., 2000. Precipitation–runoff modeling using artificial neural networks and conceptual models. Journal of Hydrologic Engineering 5, 156–161.

Van Oost, K., Govers, G., Desmet, P., 2000. Evaluating the effects of changes in landscape structure on soil erosion by water and tillage. Landscape Ecology 15, 577–589.

Verstraeten, G., Poesen, J., 2001. Factors controlling sediment yield from small intensively cultivated catchments in a temperate humid climate. Geomorphology 40, 123–144.

Verstraeten, G., Poesen, J., de Vente, J., Koninckx, X., 2003. Sediment yield variability in Spain: a quantitative and semiqualitative analysis using reservoir sedimentation rates. Geomorphology 50, 327–348.

Walling, D.E., 1983. The sediment delivery problem. Journal of Hydrology 65, 209–237.

Wicks, J.M., Bathurst, J.C., 1996. SHESED: a physically based, distributed erosion and sediment yield component for the SHE hydrological modeling system. Journal of Hydrology 175, 213–238.

Yunnan Bureau of Water Resources and Hydropower, Tianjin Survey and Design Institute, 1987. Report on Soil Erosion in Chuxiong Prefecture. Project Report, Chuxiong (in Chinese).

Zhou, G., Goel, N.K., Bhatt, V.K., 2002. Stochastic modeling of the sediment flux of the Upper Yangtze River (China). Hydrological Sciences Journal 47, S93–S105.

Zhou, Y., Lu, X.X., Huang, Y., Zhu, Y.M., 2004. Anthropogenic impact on the sediment flux in the dry-hot valleys of Southwest China – an example of the Longchuan River. Journal of Mountain Science 1, 239–249.