A novel hybrid neural network based on continuity equation and fuzzy pattern-recognition for downstream daily river discharge forecasting

Xiao-yun Chen, Kwok-wing Chau and Wen-chuan Wang

ABSTRACT

Forecasting of river discharge is crucial in hydrology and hydraulic engineering owing to its use in the design and management of water resource projects. The problem is customarily settled with data-driven models. In this research, a novel hybrid model which combines continuity equation and fuzzy pattern-recognition concept with artificial neural network (ANN), is presented for downstream river discharge forecasting in a river network. Time-varying water storage in a river station and fuzzy feature of river flow are considered accordingly. To verify the proposed model, traditional ANN model, fuzzy pattern-recognition neural network model, and hydrological modeling network model have been employed as the benchmark models. The root mean squared error, Nash–Sutcliffe efficiency coefficient and accuracy are adopted as evaluation criteria. The proposed hybrid model is applied to compute downstream river discharge in the Yellow River, Georgia, USA. Results indicate that the proposed hybrid model delivers better performance, which can effectively improve forecasting capability at the studied station. It is, therefore, proposed as a novel model for downstream river discharge forecasting because of its highly nonlinear, fuzzy and non-stationary properties.

Key words | artificial neural network, continuity equation, daily river discharge forecasting, fuzzy pattern-recognition, particle swarm optimization, storage reservoir

Xiao-yun Chen
Kwok-wing Chau (corresponding author)
Department of Civil and Environmental
    Engineering,
Hong Kong Polytechnic University,
Hung Hom,
Kowloon,
Hong Kong
E-mail: *cekwchau@polyu.edu.hk*

Wen-chuan Wang
Department of Water conservancy,
North China University of Water Resources and
    Electric Power,
Zhengzhou,
450011,
China

INTRODUCTION

The assessment of flow in a river system is of vital interest in hydraulic engineering for flood warning and/or evacuation measures. To control water levels/discharges and to operate water structures more efficiently, models that forecast river discharge are desired to be of high precision and certain degree of accuracy (ACC). Artificial neural network (ANN) models have been widely applied in this area because they do not require the a priori knowledge of the involving complex physical processes and are capable of addressing the non-linear nature of the system (Campolo *et al.* 2003). Hybrid models that combined ANN with other algorithms (e.g., genetic algorithm) in river systems have been undertaken in many studies (see Wu & Chau 2006; Firat & Güngör 2007; Pramanik *et al.* 2011; Santos & Silva 2013). However, limitations still exist.

Traditional ANN is a 'black box' which cannot reflect the physical relation between the input and output variables. It provides no insight into the natural phenomenon (Haykin 1994). An important feature of flow in a river system is that it varies temporally and spatially. Some studies were undertaken by integrating ANN and a conceptual model which could provide underlying physical processes. For instance, Song *et al.* (2011) demonstrated a feasible hybrid approach of combining back-propagation (BP) ANN with a semi-distributed Xinanjiang model in the Yanduhe watershed; Rezaeianzadeh *et al.* (2013) applied the standard conceptual HEC-HMS's soil moisture accounting algorithm and the multi-layer perceptron for forecasting daily outflows at the outlet of Khosrowshirin watershed in Iran. Other researchers have added physical

doi: 10.2166/hydro.2015.095

equations into neural networks to render data-driven models more understandable. A hydrological modeling network (HYMN) model was proposed first by Yang et al. (1998), which integrated continuity equation of flow in to an ANN model. The basic idea was that the nodes in the hidden and output layers were regarded as storage reservoirs, which could provide water mass conservation and accord with the non-linear nature. The model yielded good forecasting results in the preliminary study of Irwell River basin at Salford University. Li & Gu (2003) further extended the HYMN model to stream flow and sediment transport forecasting at the Jingjiang reach of the Yangtze River and Dongting Lake, China. Such HYMN model overcomes the drawback of a BP algorithm in traditional ANN since the parameters with physical meanings are time-varying. Piotrowski et al. (2007) showed that the forecasted peak concentrations of a transported pollutant in a river were more approximate when mass conservation was included. The number of studies which directly integrate fundamental physical principles into ANN structure is scanty, thus, a motivation of the present study is the necessity to advance the application of ANN hybrid models for river flow forecasting.

Combining fuzzy pattern-recognition ideas with a ANN model is another alternative to deal with the non-linear and fuzzy hydrological models. Qiu et al. (1998) introduced a model termed fuzzy pattern-recognition neural network (FPNN) to forecast annual runoff at Mayadu station in the Yili River of Xijiang. The fuzzy notions in this practical FPNN model were the high and low runoffs due to wet and dry seasons, respectively. The goal of pattern-recognition was to classify runoffs into a number of categories to reflect the non-linear character of the river system. Zhao & Chen (2008) applied this hybrid FPNN model in ungauged basins with consideration of the fuzziness in the concept of similar basins. Li & Chen (2010) proposed a classified method for basin floods based on the variable fuzzy sets theory; the simulation forecast was excellent as demonstrated in the verification. The combination of neural network and fuzzy theory has been conducted in many other studies, such as the adaptive neuro-fuzzy inference system (Talei et al. 2010) and fuzzy optimization neural network (Peng & Liang 2009). In addition, a hybrid model of ANN and fuzzy pattern has been applied for

modeling rainfall–runoff process (e.g., Nourani et al. 2011; Nourani & Komasi 2013). Such hybrid models have demonstrated their applicability and validity in the field of hydrology.

In the training process of ANN, usually with BP algorithm, slow convergence and easy entrapment in a local optimization is not uncommon. The parameters (such as weights and bias) are constant, making it unable to solve problems that vary with time. Besides, the traditional optimization methods used in ANN are incapable of dealing with non-differentiable objective functions (Piotrowski & Napiorkowski 2011). Recently, evolutionary algorithms, which exhibit their advantages in global optimization, have been developed as alternative methods for neural network. They are generally stochastic methods, inspired by natural selection and biological evolution. Particle swarm optimization (PSO), proposed by Kennedy & Eberhart (1995), is one common evolutionary algorithm. It has been employed and testified in hydraulic and hydrology fields for decades. Clerc & Kennedy (2002) demonstrated that PSO is an effective algorithm to find the global optimum with a large probability and high-convergence rate. In the same manner, Chau (2006) forecasted the flow stage of ShingMun River using PSO in the training algorithm. Gill et al. (2006) introduced PSO and multi-objective PSO into a conceptual rainfall–runoff model for parameter estimation. Wang et al. (2013) used PSO to determine parameters of support vector machine and forecasted the annual rainfall–runoff based on a decomposition model. The PSO also shows its distinct advantages with non-smooth objective functions and multiple minima to solve reservoir operation problems (Reddy & Nagesh Kumar 2007) and the design of water supply systems (Montalvo et al. 2010). It is thus adopted in this study as an efficient optimization technique for forecasting models.

In the present paper, a novel hybrid forecasting model which takes the effect of both river storage and flow season into consideration, motivated by the notions of HYMN and FPNN model, has been formulated. The objective is basically to forecast downstream daily river discharge based on upstream river discharges and precipitation, by incorporating continuity equation and fuzzy pattern-recognition concept into a neural network. The model is trained and tested against observed discharges in the Yellow River

system of Georgia, USA. This paper is outlined thus: the development of forecasting models based on neural network is presented; the novel hybrid model is then described in detail; the study area and available data are depicted, followed by the expression of three evaluation criteria; the computed results are discussed and the conclusions are finally drawn.

DESCRIPTION OF MODEL DEVELOPMENT

Modeling hydrological processes is of profound importance in providing reliable and accurate applications in water resource projects. This section introduces a traditional ANN model which would be applied to forecast river discharge. Then, a FPNN model which could reflect fuzzy and non-linear features of the river system is described. The algorithm of a HYMN model is further demonstrated, aimed at satisfying water mass conservation in the river network.

ANN model

ANN, being a data-driven model, is powerful in real-time forecasting. It imitates the function of the human brain and nervous system, acting as an information process system which is composed of layers and nodes. A three-layer feed-forward neural network is the most commonly used ANN in practical applications. It consists of the input, hidden, and output layers. The input layer $\{p_1, p_2, \ldots, p_k\}$ has k nodes, representing the data introduced to the network. The weighted sum of inputs and bias is passed with a predetermined activation function $f(.)$ to the nodes in the hidden layer (Thirumalaiah & Deo 1998)

$$t_i = f\left(\sum_{j=1}^{k} p_j w_{ji} + b_i\right) \tag{1}$$

where t_i ($i = 1, 2, \ldots, s$) represent nodes in the hidden layer and p_j ($j = 1, 2, \ldots, k$) represent nodes in the input layer. The weight parameter from the input layer to the hidden layer is denoted by w_{ji}, and b_i is the bias value. The

computed nodes in the output layer are obtained by similar forward pass from nodes in the hidden layers (Thirumalaiah & Deo 1998)

$$a_h = F\left(\sum_{i=1}^{s} t_i \bar{w}_{ih} + \bar{b}_h\right) \tag{2}$$

where a_h ($h = 1, 2, \ldots, r$) represent nodes in the output layer, and the weight parameter from the hidden layer to the output layer and bias are denoted by $\bar{w}_{ih}$ and $\bar{b}_h$, respectively. For traditional ANN models, the activation functions from the input layer to the hidden layer and from the hidden layer to the output layer are usually non-linear functions (e.g., radial basis function) and linear functions, respectively. They can reveal the relation of nodes between two layers, although having no physical meaning. Based on the error between target and computed outputs, the value of weights and bias are adjusted. That is, an optimal set of parameters $\{w_{11}, \ldots, w_{ks}, b_1, \ldots, b_s, \bar{w}_{11}, \ldots, \bar{w}_{sr}, \bar{b}_1, \ldots, \bar{b}_r\}$ will be determined in the learning process. Therein, gradient descent algorithm, Gauss–Newton algorithm and Levenberg–Marquardt algorithm are common optimization methods in conventional ANNs. The major drawbacks of the above methods are slow convergence rate and incapability to solve non-differential problems.

FPNN model

The fuzzy pattern-recognition idea can be combined with neural network by introducing a conceptual activation function. For this so-called FPNN model, the activation function from the input layer to the hidden layer is demonstrated as follows (Qiu *et al.* 1998)

$$Q_i = \frac{1}{\sum_{l=1}^{C} \sum_{j=1}^{k} \left[w_{ji}\left(Q_j^{in} - M_i\right)\right]^2 / \sum_{j=1}^{k} \left[w_{ji}\left(Q_j^{in} - M_l\right)\right]^2} \tag{3}$$

where Q_i ($i = 1, 2, \ldots, s$) represent nodes in the hidden layer and Q_j^{in} ($j = 1, 2, \ldots, k$) represent nodes in the input layer. The weight parameter from the input layer to the hidden layer is denoted by w_{ji}. A model vector is

denoted by $M = [M_i] = [M_l]$, that contains a number of patterns in the hidden layer. The introduction of model vector can demonstrate fuzzy pattern-recognition idea in the hidden layer, since the inputs are classified into a number of categories in terms of different patterns. The parameter C refers to the number of elements in the model vector as well as the number of nodes in the hidden layer (i.e., $C = s$). Generally, a higher value of C generates a higher precision for the forecasting result, since it implies that there are more categories in the hidden layer and represents a higher degree of non-linearity. The values $C = 5$ and $M = (1.0, 0.75, 0.50, 0.25, 0)$ were adopted in a previous study (Qiu *et al.* 1998), which meant that the degree of membership is 1.0 for 'wet' model in wet season and 0 for 'dry' model in dry season. Practically, the degrees of membership should not be linear only. However, a model vector with enough elements could circumvent this limitation. We further give a general expression for the vector M: if the number of the nodes in the hidden layer equals to C (≥ 2), then $M = (1.0, (C-2)/(C-1), (C-3)/(C-1), \ldots, 1/(C-1), 0)$. This would fully cover the models ranging from 'wet' to 'dry' season. The model vector M (including the value of C) will be determined based on the forecasting performances, which would be a trial and error process akin to the determination of the number of nodes in the hidden layer for classic ANNs. Meanwhile, the activation function of

FPNN model from the hidden layer to the output layer, is given by Qiu *et al.* (1998) as follows:

$$Q_h^{\text{out}} = \frac{1}{1 + \left[\left(\sum_{i=1}^s \bar{w}_{ih} Q_i \right)^{-1} - 1 \right]^2} \tag{4}$$

where Q_h^{out} ($h = 1, 2, \ldots, r$) represent nodes in the output layer and $\bar{w}_{ih}$ denotes the weight parameter from the hidden layer to the output layer. Thus, in the FPNN model, the parameters to be optimized are w_{ji} and $\bar{w}_{ih}$. The framework of the FPNN model is demonstrated in Figure 1. The structure of three-layer feed-forward neural network is kept. However, the activation functions are replaced by Equations (3) and (4). When compared with traditional ANN model, the FPNN model includes a pattern-recognition concept in the algorithm, which makes the recognition effects more efficient and reveals highly non-linear features. It is a viable model for fuzzy and non-linear systems, such as river flow with distinct seasonal features. However, in previous studies, the focus of FPNN models have been quite limited to simple implementation in stationary forecasting problems, which could be extended or combined with other algorithms.

HYMN model

The river network, in which the upstream river carries water flow into the downstream river, has to satisfy the following
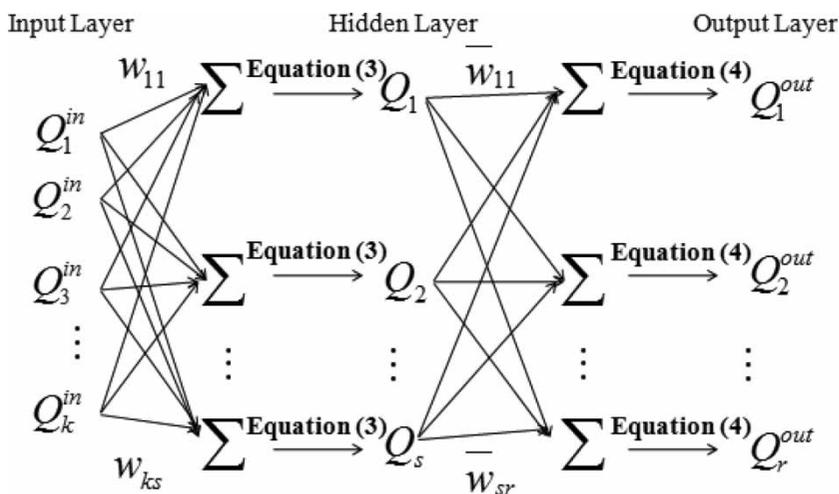

Figure 1 | Architecture of a three-layer feed-forward FPNN model.

continuity equation (Yang *et al.* 1998)

$$\frac{\partial S_h}{\partial T} = \sum_{i=1}^{s} w_{ih} Q_i - Q_h \tag{5}$$

where S is water storage, Q is water discharge, and T is time. Meanwhile, i $(1, 2, …, s)$ refers to the reservoir in a previous layer and h $(1, 2, …, r)$ refers to the reservoir in a current layer. That is, stations in the upstream river reach are represented by reservoirs in a previous layer and stations in the downstream river reach by reservoirs in a current layer. The fraction of water from a reservoir in the previous layer entering into a reservoir in the current layer is denoted by w_{ih}, which has the same meaning of weight parameter in the ANN structure. This equation implies that the rate of change of storage in the river section is determined by the difference with the source river discharge at the upstream river reach. The continuity equation is used to denote water mass conservation over the entire river system.

The discretized form of Equation (5) is

$$\frac{\Delta S_h}{\Delta T} = \sum_{i=1}^{2} w_{ih} Q_i - Q_h \tag{6}$$

where ΔT is the time step between two layers. The water storage S in the current layer at time $T + \Delta T$ is determined by the following equation

$$S_{h(T+\Delta T)} = S_{h(t)} + \left(\sum_{i=1}^{s} w_{ih} Q_{i(T)} - Q_{h(T)} \right) \times \Delta T \tag{7}$$

By setting $P_{h(t)} = \sum_{i=1}^{s} w_{ih} Q_{i(T)} \times \Delta T$, Equation (7) in its simplified form is given by

$$S_{h(T+\Delta T)} = \lambda_{h(T)} \times \left(S_{h(T)} + P_{h(T)} \right) \tag{8}$$

where $\lambda_{h(T)} = 1 - ((Q_{h(t)} \times \Delta T)/(S_{h(t)} + P_{h(t)}))$. Here, λ could be regarded as a recession coefficient, which is assumed to be independent of time (Yang *et al.* 1998). The recession coefficient represents the capability of a reservoir to absorb and store water. A higher value of recession coefficient indicates that the reservoir can store more water. Once the storage at time $T + \Delta T$ is obtained from Equation (8), the

discharge in the current layer $Q_{h(T+\Delta T)}$ is evaluated as a non-linear function of storage as follows:

$$Q_{h(T+\Delta T)} = \frac{1}{1 + \exp\left[-\left(S_{h(T+\Delta T)} + P_{h(T+\Delta T)}\right)\right]} \tag{9}$$

The non-linear feature of the reservoir lies in the non-linear relation between the discharge and the storage of the reservoir, which can be represented by an empirical expression. The above derivation of Equations (6)–(9) can be found in Yang *et al.* (1998).

The above process can be applied in the neural network to forecast downstream river discharge, which is termed the HYMN model. In the HYMN model, the river network is viewed as having the same architecture as a feed-forward neural network. The nodes in the input layer export river discharges directly. The nodes in the hidden and output layers are generalized as non-linear reservoirs with storage capacity. They receive the outflows from the previous layer, and generate discharges after computing the entire water storage. The reservoirs in the same layer do not exchange water discharges. Thus, there are two time step parameters: ΔT_1 representing the time of flow from the input layer to the hidden layer, and ΔT_2 representing that from the hidden layer to the output layer. The water storage varies with time, denoting that the previous storage in this river will affect the discharge in the next time step. Consequently, large storage in wet season will result in high discharge, exhibiting the physical phenomenon of flow in a river basin. The initial water storage $S_{h(T=0)}$ of each reservoir is prescribed before the computation. Then, storage $S_{h_{(T+\Delta T)}}$ is obtained from the initial one by every time step ΔT, which is a time-varying parameter in the model. If the storage variation term $\partial S_h / \partial T$ in Equation (5) is neglected, the HYMN model may be simplified to the traditional ANN model. Three kinds of parameters are to be optimized: weight parameters w_{ji} and $\bar{w}_{ih}$, whose definitions are the same as those in ANN model; recession coefficients λ_i and $\bar{\lambda}_h$ for the nodes in the hidden and output layers, respectively, initial storages $S_{i(T=0)}$ and $\bar{S}_{h(T=0)}$ for the nodes in the hidden and output layers, respectively. Each of them has specific meanings related to the non-linearity and storage capacity of the reservoir. Accordingly, the HYMN model could be applied

to a river system in which observed river stations could be regarded as storage and non-linear reservoirs.

HYBRID NEURAL NETWORK MODEL

This paper aims at forecasting downstream river discharge based on precipitation and upstream river discharges. The ANN, FPNN, and HYMN models have been proved to be efficient forecasting models in previous studies. However, there are some limitations with these models. For example, in the HYMN model, the nodes in the hidden layer are assumed as storage reservoirs, which apart from being impractical, render the model complicated. First, it is difficult to determine the number of reservoirs between the input and output river stations (i.e., the number of nodes in the hidden layer). Second, it is not realistic to define time step of river flow from the input layer to a virtual reservoir in the hidden layer. Finally, the recession coefficients for reservoirs in the hidden layer are physically meaningless. Above all, it is unrealistic to regard the nodes in the hidden layer as storage reservoirs.

A tentative practice is to apply Equation (3) as the activation function from the nodes in the input layer to those in the hidden layer, and to consider the nodes in the output layer as storage reservoirs. This could also satisfy the mass conservation principle and represent the time-varying feature of the river network. The framework within such a hybrid neural network (HNN) model is as follows: (i) obtain the nodes in the hidden layer (i.e., determine the value of C by trial and error) by classifying the input variables into a number of categories in terms of different seasons from Equation (3); (ii) the sorted flows reach reservoirs in the output layer with a time step; (iii) compute the storage of reservoirs; and (iv) output the discharges. In this way, the fuzzy characteristics of river flow and the time-dependent storage capacity of the observed station are well considered. This novel model is much easier since storage reservoirs in the hidden layer are excluded. Similarly, the parameters to be optimized for the proposed HNN model are weight parameters w_{ji} and $\bar{w}_{ih}$; recession coefficients $\bar{\lambda}_h$; and initial storages $\bar{S}_{h(T=0)}$ for the reservoirs in the output layer. In the present study, the objective is to forecast the downstream river discharge. As a result, there is only one node to be considered as the reservoir in the output layer. The set of parameters to be optimized is correspondingly $\{w_{11}, \ldots w_{ks}, \bar{w}_{11}, \ldots, \bar{w}_{s1}, \bar{\lambda}, \bar{S}_{(T=0)}\}$.

The PSO algorithm is employed to optimize the parameters in the learning process of HNN forecasting model, as outlined in Figure 2. A detailed description of the PSO algorithm is given by Clerc & Kennedy (2002). The focus of this paper is on the application of PSO
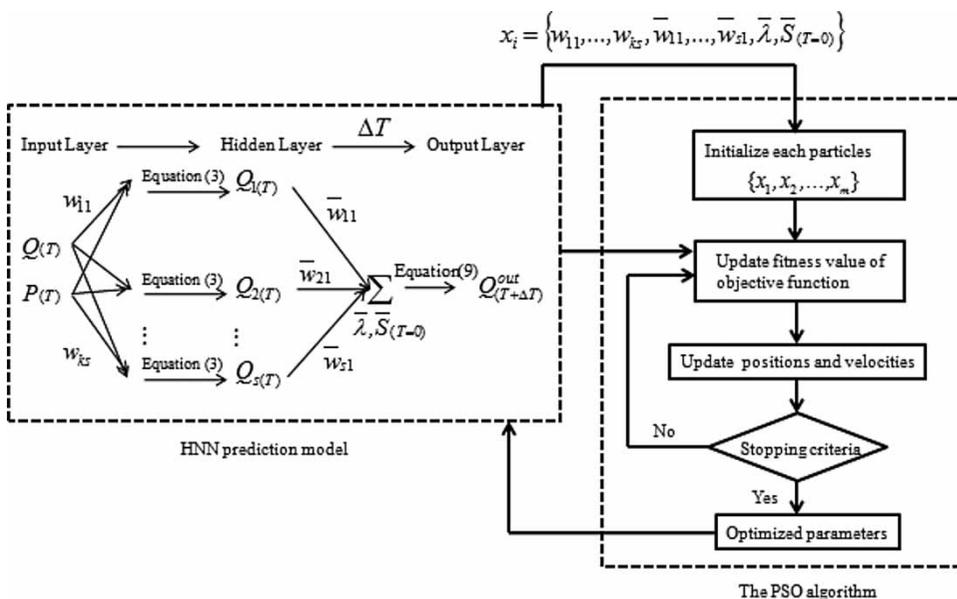

Figure 2 | Outline of the PSO algorithm for the calibration of HNN forecasting model.

algorithm in the parameter optimization for the HNN model. As shown in the left part of Figure 2 for the framework of HNN model, the output discharge at time $T + \Delta T$ is computed from inputs (discharge Q and precipitation P) at time T with two activation functions Equations (3) and (9). Particularly, values of storage vary with time and are obtained from previous time step values. The set of parameters to be optimized in the HNN model include particle x_i and D dimensions $(D = k \times s + s \times 1 + 1 + 1)$, where D is the number of parameters in the set, i.e., the sum of number of w_{ji}, $\bar{w}_{ih}$, $\bar{\lambda}$ and $\bar{S}_{(T=0)}$. The process of optimizing the parameters with PSO is established in the right-hand side of Figure 2 and described as follows: (i) a population with randomly initialized positions and velocities within the range of parameters is generated for the PSO algorithm; (ii) define the objective function of the HNN model, which would be used to evaluate the fitness of each particle; (iii) train the proposed HNN model with parameters corresponding to the current particle and obtain the fitness value of the objective function; (iv) the velocity and position of each particle are updated in each iteration until a stopping criterion is satisfied. Consequently, an optimal set of parameters is obtained with respect to the fitness value of the objective function for the forecasting model. In this way, a HNN forecasting model based on continuity equation and fuzzy pattern-recognition combining with PSO algorithm is developed to forecast downstream river discharge.

CASE STUDY

Study area and data

To verify the application of suggested models, the Altamaha River basin located in Georgia, USA is selected as the case study site. It is a large river basin on the Atlantic coast whose drainage basin is about 36,000 km² in size. The Yellow River flows generally southward for 122 km, as a tributary of the Ocmulgee River. The flow of the Yellow River exhibits a seasonal behavior, which is low in dry season and high during wet months. The river flow varies depending on various impacts such as climate and human activities. An accurate forecasting of river discharge is vital

since it allows engineers to make efficient decisions in water management and to prevent flooding. The daily time series of discharge and precipitation in this river basin were downloaded from the USGS web server (http://water-data.usgs.gov/ga/nwis/rt).

In Figure 3, five hydrological stations along the Yellow River are marked, demonstrating their locations on the map. The objective is to forecast the discharge in station 02208000 based on the data from the nearest upstream station 02207335, whose ID, name, latitude, and longitude are summarized in Table 1. As shown in Figure 4, the daily discharge varies quite irregularly and the peak value could be as high as 199.35 m³/s. The distance from station 02207335 to 02208000 is roughly 19.3 km, and the travel time of flow between the two stations is estimated to be 16 hours based on the mean velocity during flood flow condition (http://waterdata.usgs.gov/ga/nwis/rt). Daily data from 1 January 2010 to 31 December 2013 were selected, and separated into training and testing sets. The datasets for the training stage are from 1 January 2010 to 31 December 2011, taking around 50% of all data. The data of year 2012 are utilized for validation. During the calibration process, the training will be strictly terminated at the point where the error in the validation set begins to rise. Validation is a necessary and crucial procedure to avoid overfitting the training data (Faber & Rajko 2007). The remaining data from 1 January 2013 to 31 December 2013 are used to assess the efficiency of the calibration and evaluate the performance of the forecasting models.

The downstream station 02208000 receives its flows mainly from two sources: discharge from the upstream station and precipitation at the current station. Thus, two input variables are selected in the forecasting model, which are discharge at station 02207335 and precipitation at station 02208000, as shown in Figure 5. The maximum daily precipitation is as high as 0.078 m, and heavy rainfall is presented in this area. A preliminary study is carried out to examine the contribution of precipitation to the forecasting of downstream flow by ANN model, and the computed discharges are demonstrated in Figure 6. As can be seen, the model with precipitation input outperforms the other one with evidence of intensively distributed dots along the ideal line. Thus, precipitation is adopted as an indispensable input variable for this study site.
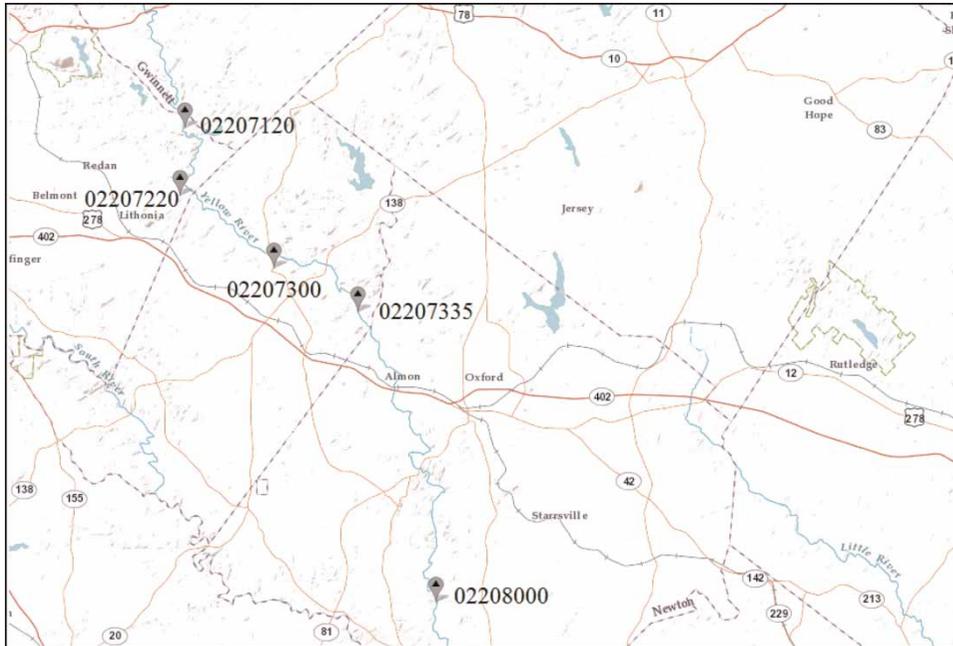
Figure 3 │ Location of two stations along the Yellow River in Altamaha River basin, Georgia (http://waterdata.usgs.gov/ga/nwis/current).

Table 1 │ Stations' ID, name, and location in Altamaha River Basin, Georgia

Station ID	Station name	Latitude	Longitude
02207335	Yellow River at Gees Mill road	33°40′01″N	83°56′17″W
02208000	Yellow River at Rocky Plains road	33°29′59.5″N	83°53′03″W
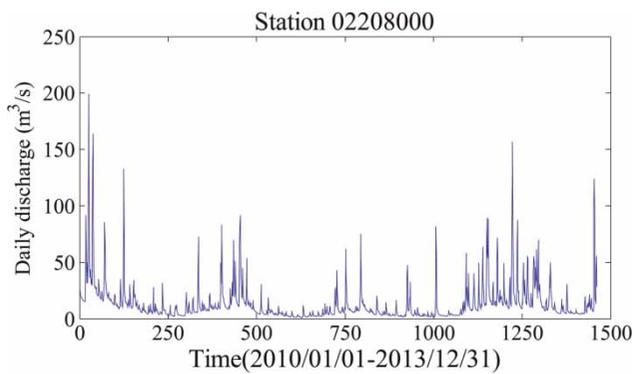
Figure 4 │ Daily discharges at station 02208000 as output for forecasting models.

Figure 5 │ Discharge at station 02207335 (a), and precipitation at station 02208000 (b), as inputs.

In addition, it is recommended to normalize each attribute in order to avoid larger data dominating smaller data. In this paper, the datasets of discharge are scaled linearly
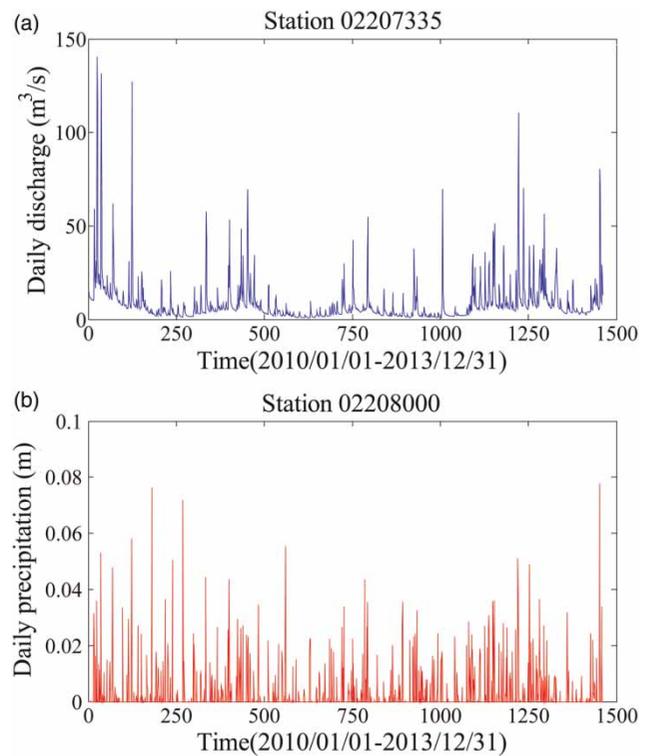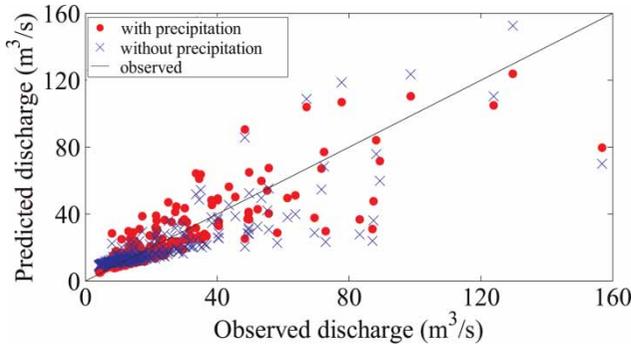
Figure 6 | Observed and computed discharges by ANN model with different input variables.

to the range between 0.1 and 0.9 as follows (Campolo et al. 1999)

$$Q_i' = 0.1 + 0.8 \times \frac{Q_i - Q_{\min}}{Q_{\max} - Q_{\min}} \tag{10}$$

where Q_i' is the scaled value, Q_i is the original discharge value, and $Q_{\min}$, $Q_{\max}$ are the minimum and maximum of the flow series, respectively.

Performance evaluation

To evaluate the performance of forecasting models, three statistical indices were used as evaluation criteria. First, the root mean squared error (RMSE) is a commonly used error index statistic which is defined as follows (Legates & McCabe 1999)

$$\text{RMSE} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(Q_i - \hat{Q}_i)^2} \tag{11}$$

It is used as the objective function in the calibration period in this study as well. Second, the Nash–Sutcliffe efficiency coefficient (NSEC) recommended by Nash & Sutcliffe (1970), is a normalized efficiency coefficient to assess forecasting results. It is formulated in Equation (12), exhibiting the relative magnitude of the residual variance compared to the measured data variance. The statistic NSEC scales the mean squared error as well as RMSE, therefore could reflect the performance on high values. When NSEC = 1, it is a perfect fit between the forecasted discharge and the observed data; and NSEC = 0

indicates that the model reaches ACC when the mean of the observed data are forecasted

$$\text{NSEC} = 1 - \frac{\sum_{i=1}^{N}(Q_i - \hat{Q}_i)^2}{\sum_{i=1}^{N}(Q_i - \bar{Q})^2} \tag{12}$$

The lastly employed statistical index is known as ACC, which is obtained from the mean relative error (De Vos & Rientjes 2008). Its mathematical expression is presented in Equation (13). Obviously, a higher value of ACC reveals a better forecasting performance of the model.

$$\text{ACC} = 1 - \frac{1}{N}\sum_{i=1}^{N}\frac{\left|Q_i - \hat{Q}_i\right|}{Q_i} \tag{13}$$

In the above equations, Q_i and $\hat{Q}_i$ are, respectively, observed and forecasted discharges, $\bar{Q}$ is the mean of observed data, and N is the number of data. In the following, RMSE, NSEC, and ACC are adopted to evaluate the performances of the various models developed.

RESULTS AND DISCUSSION

First, we determined the model vector M and its corresponding value C for the proposed model by comparing the fitness values during the training period. As illustrated in Figure 7, the fitness value varies with the number of nodes in the hidden layer (value C). It decreases with the fluctuation while C increases, and attains a minimum value when $C = 11$. The corresponding vector is $M = (1.0, 0.9, 0.8, \ldots, 0.1, 0)$, which is large
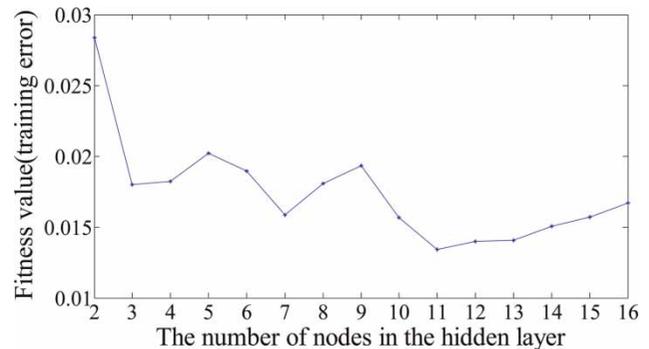

Figure 7 | Fitness value for different numbers of nodes in the hidden layer for HNN model.

enough to perform the non-linear property. Similar tests were conducted to determine the number of nodes in the hidden layer for the other three forecasting models.

Since the travel time of flow from station 02207335 to 02208000 is estimated as 16 hours, the lead time for the ANN and FPNN model is selected as 1 day, which is considered useful and necessary for practical purposes in this case study. That is, the output discharge at time T is computed from upstream discharge at time $T–1$ and precipitation at time $T–1$. Then, time step ΔT is a vital parameter for both HYMN and HNN models. As demonstrated in Figure 2 for the HNN model, $Q^{out}_{(T+\Delta T)}$ is computed from $\{Q_{(T)}, P_{(T)}\}$, which means that the downstream flow has ΔT days delay from the inputs. That is, the inputs at day T strongly influence downstream discharge at day $T + \Delta T$. We try different time steps ΔT (1, 2, 3, 4, 5 day) for the proposed model as shown in Table 2 for comparison. As a similar concept with forecasting lead time, ΔT is selected as 1 day for HYMN and HNN models. Accordingly, the models provide a 1 day lead time forecast for the flow in station 02208000.

To validate the proposed model, forecasting results were compared to three benchmark models. The PSO algorithm was employed as the optimization method for the same basis of comparison. The evaluation criteria RMSE, NSEC, and ACC during training and testing stages for different models are provided in Table 3. It can be observed that the traditional ANN model is able to attain acceptable forecasting results as the NSEC value during the testing period is 0.7607. However, the peak discharge computed by ANN model is about 20 m^3/s less when compared with the observed value. This under-forecasting for peak discharge is intolerable for flood warning. The FPNN model provides a better forecast of the peak discharge, and performs better than the ANN model in terms of the ACC value. The improvement is yet not distinct. The HYMN model performs excellently in terms of both RMSE and NSEC values. However, the ACC value is only 0.6841 in the testing period. This indicates that the HYMN model is better in computing high values of discharges than the ANN model but, nevertheless, cannot ensure good ACC of all values. On the contrary, the present HNN model attains the best results for all evaluation criteria. For the training stage, there is a 56.79% reduction in RMSE value and 18.06% increase in NSEC value by the proposed model compared with the ANN model. The RMSE and NSEC values of the HNN model are 8.3465 and 0.8210 m^3/s, respectively, in the testing period, with a 36.10% reduction in RMSE and 46.40% improvement in NSEC value when compared with the FPNN model. The improvements regarding the NSEC and ACC value are 3.83 and 16.49%, respectively, over the HYMN model in the testing stage. This reveals that the proposed model outperforms the other three benchmark models with best generalization and forecasting ability.

To further illustrate these results, Figure 8 exhibits the observed and computed discharges during the testing period for all four models. It can apparently be perceived

Table 2 | Training and testing performances of different ΔT values for the HNN model

ΔT (day)	Training			Testing		
	RMSE (m^3/s)	NSEC	ACC	RMSE (m^3/s)	NSEC	ACC
1	3.4326	0.9661	0.7727	8.3465	0.8210	0.7969
2	5.5129	0.9126	0.6733	10.5426	0.7152	0.7758
3	6.4944	0.8789	0.6609	11.9661	0.6329	0.7451
4	6.7369	0.8698	0.6537	12.3854	0.6078	0.7505
5	6.8306	0.8663	0.6889	17.2968	0.2370	0.6832

Table 3 | Training and testing performances for different models

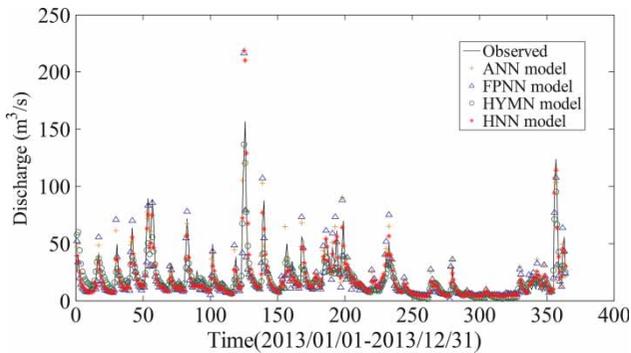| Model	Number of nodes	Training			Testing			Peak discharge (m^3/s)
		RMSE (m^3/s)	NSEC	ACC	RMSE (m^3/s)	NSEC	ACC	
ANN model	9	7.9437	0.8183	0.7259	9.6406	0.7607	0.7289	119.83
FPNN model	11	6.6239	0.8737	0.6496	13.0611	0.5608	0.7526	215.66
HYMN model	6	6.0746	0.8938	0.5903	9.0272	0.7907	0.6841	136.91
HNN model	11	3.4326	0.9661	0.7727	8.3465	0.8210	0.7969	218.72

Figure 8 │ Observed and computed discharges during testing period for different models.

that the peak value is over-forecasted by FPNN and HNN models, with triangular and asterisk symbols. For the high values of computed discharges, most results from ANN and FPNN models are larger than the observed values. These two models can forecast the main data of the discharges satisfactorily, yet the computation of high values is relatively poor. Apparently, HYMN and HNN models match the observed values much better than their counterpart. In addition, we obtain the recession coefficient for the reservoir in the output layer for HYMN and HNN models, which were 0.4369 and 0.4553, respectively. The coefficient computed from the two models does not vary significantly, and is thus acceptable. The reliability of the recession coefficient for reservoirs in the output layer can be assured. On the contrary, the recession coefficients for reservoirs in the hidden layer for the HYMN model vary from 0.0239 to 0.9790, revealing their uncertainty and impractical application. This offers concrete evidence for the assumption that it is incorrect to regard the nodes in the hidden layer as storage reservoirs in the HYMN model. Accordingly, a conclusion can be drawn that the proposed HNN model is an improvement over the others and capable of producing good and approximate results. The forecasting performance is enhanced significantly as it can reflect the physical processes of the hydrological cycle within the river system more accurately with factual supports.

CONCLUSIONS

In this paper, a novel hybrid model reflecting fuzzy features of river flows and non-linear storage reservoir is proposed. This

model incorporates continuity equation and fuzzy pattern-recognition concept into a neural network for downstream river discharge forecasting. The integration of physical equations and data-driven models in the river system renders the forecasting more meaningful physically. The PSO algorithm is applied as a feasible optimization technique in model calibration. The capability and applicability of the proposed model has been illustrated using the Yellow River in Georgia, USA as a case study. The performances of models in training and testing periods are analyzed using three evaluation criteria (RMSE, NSEC, and ACC). It can be concluded that the hybrid model is a better and improved forecasting model, since it outperforms other benchmark models with lower RMSE, higher NSEC, and ACC values. A meaningful value 0.4553, namely the recession coefficient of the downstream river section, is obtained when regarded as a storage reservoir. The hybrid model to forecast river discharge in this paper is as yet a pilot study. Since the uncertainty analysis is important for new proposed approaches (Khan & Coulibaly 2006; Boucher et al. 2009; Alvisi & Franchini 2012; Alvisi et al. 2012), it will be undertaken as a future study. In addition, the efficiency and adaptability of the proposed model with other input variables and optimization algorithms can be further investigated.

REFERENCES

Alvisi, S. & Franchini, M. 2012 Grey neural networks for river stage forecasting with uncertainty. *Phys. Chem. Earth* **42–44**, 108–118.

Alvisi, S., Creaco, E. & Franchini, M. 2012 Crisp discharge forecasts and grey uncertainty bands using data-driven models. *Hydrol. Res.* **43** (5), 589–602.

Boucher, M. A., Perreault, L. & Anctil, F. 2009 Tools for the assessment of hydrological ensemble forecasts obtained by neural networks. *J. Hydroinform.* **11** (3–4), 297–307.

Campolo, M., Andreussi, P. & Soldati, A. 1999 River flood forecasting with a neural network model. *Water Resour. Res.* **35** (4), 1191–1197.

Campolo, M., Soldati, A. & Andreussi, P. 2003 Artificial neural network approach to flood forecasting in the River Arno. *Hydrol. Sci. J.* **48** (3), 381–398.

Chau, K. W. 2006 Particle swarm optimization training algorithm for ANNs in stage prediction of Shing Mun River. *J. Hydrol.* **329** (3), 363–367.

Clerc, M. & Kennedy, J. 2002 The particle swarm-explosion, stability, and convergence in a multidimensional complex space. *Evolut. Comput. IEEE Trans.* **6** (1), 58–73.

De Vos, N. J. & Rientjes, T. H. M. 2008 Multiobjective training of artificial neural networks for rainfall-runoff modeling. *Water Resour. Res.* **44** (8), W08434.

Faber, N. M. & Rajko, R. 2007 How to avoid over-fitting in multivariate calibration – the conventional validation approach and an alternative. *Analyt. Chim. Acta* **595** (1), 98–106.

Firat, M. & Güngör, M. 2007 River flow estimation using adaptive neuro fuzzy inference system. *Math. Comput. Simul.* **75** (3), 87–96.

Gill, M. K., Kaheil, Y. H., Khalil, A., McKee, M. & Bastidas, L. 2006 Multiobjective particle swarm optimization for parameter estimation in hydrology. *Water Resour. Res.* **42** (7), W07417.

Haykin, S. 1994 *Neural Networks: A Comprehensive Foundation.* Prentice Hall, Upper Saddle River, NJ.

Kennedy, J. & Eberhart, R. 1995 Particle swarm optimization. *1995 IEEE International Conference on Neural Networks Proceedings, Vol.* 1–6, pp. 1942–1948.

Khan, M. S. & Coulibaly, P. 2006 Bayesian neural network for rainfall-runoff modeling. *Water Resour. Res.* **42** (7), W07409.

Legates, D. R. & McCabe, G. J. 1999 Evaluating the use of 'goodness-of-fit' measures in hydrologic and hydroclimatic model validation. *Water Resour. Res.* **35** (1), 233–241.

Li, M. & Chen, S. Y. 2010 Fuzzy variable classified method and its application in basin floods. *Adv. Intell. Soft Comput.* **78**, 627–636.

Li, Y. T. & Gu, R. R. 2003 Modeling flow and sediment transport in a river system using an artificial neural network. *Environ. Manage.* **31** (1), 122–134.

Montalvo, I., Izquierdo, J., Pérez-García, R. & Herrera, M. 2010 Improved performance of PSO with self-adaptive parameters for computing the optimal design of water supply systems. *Eng. Appl. Artif. Intell.* **23** (5), 727–735.

Nash, J. E. & Sutcliffe, J. V. 1970 River flow forecasting through conceptual models part I – A discussion of principles. *J. Hydrol.* **10** (3), 282–290.

Nourani, V. & Komasi, M. 2013 A geomorphology-based ANFIS model for multi-station modeling of rainfall–runoff process. *J. Hydrol.* **490**, 41–55.

Nourani, V., Kisi, Ö. & Komasi, M. 2011 Two hybrid artificial intelligence approaches for modeling rainfall–runoff process. *J. Hydrol.* **402** (1), 41–59.

Peng, Y. & Liang, G. H. 2009 A fuzzy optimization neural network model using second order information. In: *Proceedings of Fuzzy Systems and Knowledge Discovery, 2009. FSKD'09. Sixth International Conference on. IEEE, 14–16 August, Tianjin, China,* pp. 221–227.

Piotrowski, A. P. & Napiorkowski, J. J. 2011 Optimizing neural networks for river flow forecasting–evolutionary computation methods versus the Levenberg-Marquardt approach. *J. Hydrol.* **407** (1), 12–27.

Piotrowski, A. P., Wallis, S. G., Napiorkowski, J. J. & Rowinski, P. M. 2007 Evaluation of 1-D tracer concentration profile in a small river by means of multi-layer perceptron neural networks. *Hydrol. Earth Syst. Sci.* **11**, 1883–1896.

Pramanik, N., Panda, R. & Singh, A. 2011 Daily river flow forecasting using wavelet ANN hybrid models. *J. Hydroinform.* **13** (1), 49–63.

Qiu, L., Chen, S. Y. & Nie, X. T. 1998 A forecast model of fuzzy recognition neural networks and its application. *Adv. Water Sci.* **9** (3), 258–264.

Reddy, M. J. & Nagesh Kumar, D. 2007 Multi-objective particle swarm optimization for generating optimal trade-offs in reservoir operation. *Hydrol. Process.* **21** (21), 2897–2909.

Rezaeianzadeh, M., Stein, A., Tabari, H., Abghari, H., Jalalkamali, N., Hosseinipour, E. Z. & Singh, V. P. 2013 Assessment of a conceptual hydrological model and artificial neural networks for daily outflows forecasting. *Int. J. Environ. Sci. Technol.* **10** (6), 1181–1192.

Santos, C. A. G. & Silva, G. B. L. 2013 Daily streamflow forecasting using a wavelet transform and artificial neural network hybrid models. *Hydrol. Sci. J.* **59** (2), 312–324.

Song, X. M., Kong, F. Z., Zhan, C. S. & Han, J. W. 2011 Hybrid optimization rainfall-runoff simulation based on Xinanjiang model and artificial neural network. *J. Hydrol. Eng.* **17** (9), 1033–1041.

Talei, A., Chua, L. H. C. & Quek, C. 2010 A novel application of a neuro-fuzzy computational technique in event-based rainfall–runoff modeling. *Exp. Syst. Applicat.* **37** (12), 7456–7468.

Thirumalaiah, K. & Deo, M. C. 1998 River stage forecasting using artificial neural networks. *J. Hydrol. Eng.* **3** (1), 26–32.

Wang, W. C., Xu, D. M., Chau, K. W. & Chen, S. Y. 2013 Improved annual rainfall-runoff forecasting using PSO–SVM model based on EEMD. *J. Hydroinform.* **15** (4), 1377–1390.

Wu, C. L. & Chau, K. W. 2006 A flood forecasting neural network model with genetic algorithm. *Int. J. Environ. Pollut.* **28** (3), 261–273.

Yang, R. F., Ding, J. & Liu, G. D. 1998 Preliminary study of hydrology-based artificial neural network. *J. Hydraul.* **8**, 23–27.

Zhao, R. H. & Chen, S. Y. 2008 A hybrid fuzzy and neural network model for hydrological forecasting in ungauged basins. In: *Proceedings of the International Association of Hydrological Sciences and the International Water Resources Association Conference, Guangzhou, China, 8–10 June 2006.*