Hydrological Sciences Journal

ISSN: 0262-6667 (Print) 2150-3435 (Online) Journal homepage: https://www.tandfonline.com/loi/thsj20

River flow time series prediction with a range-dependent neural network

T. S. HU , K. C. LAM & S. T. NG

River flow time series prediction with a range-dependent neural network

T. S. HU*

Department of Hydraulic Engineering, Wuhan University, Wuhan, Hubei Province, China

K. C. LAM

Department of Building and Construction, City University of Hong Kong, Tat Chee Avenue, Hong Kong

e-mail: bckclam@city.edu.hk

S. T. NG

Department of Civil Engineering, The University of Hong Kong, Pokfulam Road, Hong Kong

Abstract Artificial neural networks provide a promising alternative to hydrological time series modelling. However, there are still many fundamental problems requiring further analyses, such as structure identification, parameter estimation, generalization, performance improvement, etc. Based on a proposed clustering algorithm for the training pairs, a new neural network, namely the range-dependent neural network (RDNN) has been developed for better accuracy in hydrological time series prediction. The applicability and potentials of the RDNN in daily streamflow and annual reservoir inflow prediction are examined using data from two watersheds in China. Empirical comparisons of the predictive accuracy, in terms of the model efficiency R^2 and absolute relative errors (*ARE*), between the RDNN, back-propagation (BP) networks and the threshold auto-regressive (TAR) model are made. The case studies demonstrated that the RDNN network performed significantly better than the BP network, especially for reproducing low-flow events.

Key words river flow; prediction; hydrological time series; artificial neural networks; range-dependent neural network; threshold auto-regressive (TAR) model

Prévision de séries temporelles de débits en rivière par un réseau de neurones dépendant d'échelle

Résumé Les réseaux de neurones artificiels fournissent une alternative prometteuse pour la modélisation des séries temporelles en hydrologie. Cependant, plusieurs problèmes fondamentaux se posent encore, comme l'identification de la structure, l'estimation des paramètres, la généralisation de l'amélioration des performances, etc. Un nouveau type de réseau de neurones, basé sur un algorithme de classification des couples d'apprentissage et baptisé réseau de neurones dépendant d'échelle, a été développé pour améliorer la précision de la prévision des séries hydrologiques. Nous avons étudié l'applicabilité et le potentiel de ce modèle pour la prévision des débits journaliers et annuels de deux bassins versants chinois. Nous avons comparé empiriquement la qualité de la prévision, en termes d'efficacité du modèle R^2 et d'erreur relative absolue, entre le réseau de neurones dépendant d'échelle, le réseau à rétro-propagation et le modèle auto-régressif seuillé. Les études de cas ont montré que le réseau de neurones dépendant d'échelle donne des résultats significativement meilleurs que le réseau à rétro-propagation, en particulier pour la prévision des débits faibles.

Mots clefs écoulement en rivière; séries hydrologiques temporelles; réseau de neurones dépendant d'échelle; modèle auto-régressif seuillé

* *Also at:* Department of Building and Construction, City University of Hong Kong, Tat Chee Avenue, Hong Kong.

Open for discussion until 1 April 2002

INTRODUCTION

When studying river flow one deals with an extremely complex problem area. The flow process can be influenced by a wide variety of factors including the precipitation intensity and distribution, channel characteristics, watershed geology and topography, vegetation cover, human activities (e.g. land-use changes) and even, indirectly, the greenhouse gas releases. Attempts to improve the accuracy of river flow predictions have been a relentless pursuit by hydrologists and water resources engineers. Various mathematical models have been developed to tackle the problems. These models can be classified into two broad categories: physically-based hydrological (PBH) models and system identification (SI) models (Amorocho & Hart, 1964). Since PBH models require the extremely complex relationships between river flow and its influencing factors to be established, coupled with the fact that this kind of model places a very stringent demand for retrospective hydrometeorological data, many hydrologists favour the use of SI models. Examples of SI models include the simple linear (SL) model, the seasonally-based linear perturbation (LP) model, the auto-regressive and moving average (ARMA) model, and the constrained linear system (CLS) model. These models have been applied to many catchments worldwide with varying degrees of success.

Owing to the inherent nonlinear nature of river flow processes and the poor results as confirmed by the linear SI models, many researchers are becoming interested in the nonlinear SI models (Hughes & Murrell, 1986; Kachroo & Natale, 1992; Kember & Flower, 1993). More recently, chaos theory and artificial neural networks (ANNs) have been considered as alternatives to the more traditional nonlinear SI models for studying nonlinear hydrological processes (Wilcox *et al*, 1981; Rodriguez-Iturbe *et al*, 1989; See & Openshaw, 2000).

Many studies on ANN-based hydrological modelling have been reported over the last decade, especially since 1995 (e.g. French *et al.*, 1992; Lachtermacher & Fuller, 1995; Lorrai & Sechi, 1995; Raman & Sunilkumar, 1995; Smith & Eli, 1995; Mason *et al.*, 1996; Raman & Chandramouli, 1996; Shamseldin, 1997; Thirumalaiah & Deo, 1998; Atiya *et al.*, 1999; See & Openshaw, 1999; Sajikumar & Thandaveswara, 1999; Whitley, 1999; ASCE Task Committee on Artificial Neural Networks in Hydrology, 2000a,b; Coulibaly *et al.*, 2000; Liong *et al.*, 2000; Luk *et al.*, 2000; Maier & Dandy, 2000). The distinguishing feature of the ANN-based hydrological model is the use of the ANN's capacity in approximating any continuous function to a degree of accuracy, which is not otherwise available from other conventional hydrological modelling techniques. Additional benefits include data error tolerance and the characteristic of being data-driven, thereby providing a capacity to learn and generalize patterns in noisy and ambiguous input data without the need for prior knowledge.

Despite these features, problems associated with structure identification and parameter estimation, such as time-consuming training processes, the existence of numerous multi-local minima, and a lack of well-recognized structure identification algorithms applicable to the catchments, still exist in ANN-based hydrological modelling. Other factors complicating ANN-based hydrological modelling include learning algorithms, training sets, activation functions, an objective function for the training, sensitivity to input errors, input and output pre-processing, input window lengths, lengths and contents of calibration data, prior knowledge and the specific nature of the watershed under study.

Apart from these theoretical problems, two major problems were found in ANN-based hydrological modelling (Hu, 1996, 1997). First, ANNs have problems reproducing extreme hydrological events, such as overestimation of low flows and underestimation of peak flows. Second, under some circumstances, high calibration errors could not be reduced to a desired level when there were additional hidden nodes or additional hidden layers in a single-neural-network-based modelling, and this could give rise to high validation errors (see Karunanithi *et al.*, 1994; Minns & Hall, 1996; Campolo *et al.*, 1999; Dawson & Wilby, 1998; Jain *et al.*, 1999; Zealand *et al.*, 1999; Imrie *et al.*, 2000).

To alleviate these two problems, a range-dependent neural network (RDNN), which is based on a clustering algorithm for training pairs, is proposed herein. The objective of this paper is to examine the potential of the RDNN in estimating river flows by comparing the results with those generated by the commonly used back-propagation (BP) network and a threshold auto-regressive (TAR) model.

RANGE-DEPENDENT NEURAL NETWORK

This section describes the motivations and development processes of the RDNN. Instead of using one complex feed-forward neural network (FFNN) for river flow time series modelling, the RDNN employs several simple but adequate networks based on the threshold value of the TAR model developed for the time series, each of which learns to handle a subset of the complete set of training samples.

Fundamentals of the feed-forward neural network

Many different ANN models have been developed since the 1980s, the most popular for water resources research being the FFNNs, Hopfield networks, and Kohonen's self-organizing networks. Despite substantial developments in the use of ANNs for forecasting, the BP algorithm for training the FFNNs was not introduced until 1986 (Rumelhart *et al.*, 1986).

Consider a three-layer FNN for river flow prediction, in which the numbers of neurons in the input layer, the hidden layer and the output layer are denoted by N_1, N_2, N_3, respectively:

$$\begin{cases} OO_{pk} = f(\sum_{j=1}^{N_2} w_{kj}^{\mathrm{S}} OH_{pj} - \theta_k^{\mathrm{O}}) \\ OH_{pj} = f(\sum_{j=1}^{N_2} w_{ji}^{\mathrm{F}} OI_{pi} - \theta_j^{\mathrm{H}}) \\ OI_{pi} = I_{pi} \end{cases} \tag{1}$$

where $w_{kj}^{\mathrm{S}}, w_{ji}^{\mathrm{F}}$ are the connection strengths between neuron k of the output layer and neuron j of the hidden layer, and between neuron j of the hidden layer and neuron i of the input layer, respectively; $\theta_k^{\mathrm{O}}, \theta_j^{\mathrm{H}}$ are the bias of the unit k and j, OO_{pk}, OH_{pj}, OI_{pi} are the output of the unit k, j, i for the pth input pattern; I_{pi} is the ith input in the input layer for pth input pattern. The goal of BP network training is to minimize the sum of the squares of the errors between the modelled outputs (OO_{pk}) and the target outputs (T_{pk}), which can be expressed as the following optimization problem:

$$\mathrm{BP}\begin{cases} \min F = \sum\limits_{p}\sum\limits_{k} (T_{pk} - OO_{pk})^2 \\ \text{subject to } OO_{pk} - f(\sum w_{kj}^{\mathrm{S}} OH_{pj} - \theta_k^{\mathrm{O}}) = 0 \\ \qquad\qquad\quad OH_{pj} - f(\sum w_{ji}^{\mathrm{F}} OI_{pi} - \theta_j^{\mathrm{H}}) = 0 \end{cases} \tag{2}$$

Clustering algorithm for training pairs

If BP is used for training a single, multi-layer network to perform different subtasks on different occasions, there will be generally strong interference effects that may lead to slow learning and poor generalization (Jacobs *et al.*, 1991). For a better modelling performance, Hampshire & Waibel (1989) developed a system so that division into subtasks is known prior to training. Jacobs *et al.* (1991) presented a new supervised learning procedure involving the training of many separate networks using data that correspond to various classes within the available data set. A similar concept was adopted by Rodriguez & Serodes (1996) in their modelling of the adequacy of chlorine dosage for drinking water disinfection with training samples being divided into subsets representing winter and summer conditions.

In the context of rainfall–runoff modelling, a neural network with complex structure was frequently found unable to adapt to the complexity of rainfall–runoff transformation and other hydrological processes. Calibration errors were found to remain high with the inclusion of additional hidden nodes or even additional hidden layers (Hu, 1997). This may be partially attributed to the fact that the nonlinear dynamics governing the low, medium and high magnitudes of flows are quite different. Low-flow events are sustained by baseflow, while high-flow events are fed primarily by fast surface flow (Zhang & Govindaraju, 2000). It would be difficult for a neural network with satisfactory validation accuracy for high-flow events to simultaneously reproduce low-flow and medium-flow events very well.

Based on this argument, instead of using one complex network, several simple but adequate networks are proposed herein to model the hydrological time series. Each of the networks has its own training pairs obtained by the flowing clustering algorithm and serves different magnitudes of flow predictions. This clustering algorithm is based on the calibration results of the TAR model for a hydrological time series.

The TAR model was first proposed by Tong (1978) and further refined by Tong & Lim (1980). The basic idea of the TAR model is to describe a given stochastic process by a piecewise autoregressive model according to the value of an observable variable, termed the threshold variable. If this variable is a logged value of the time series, the model is called a self-exciting threshold auto-regressive (SETAR) model, and the SETAR is adopted in the case study below. Other extensions of TAR models include the smooth threshold auto-regressive (STAR) model, the time-varying STAR (TV-STAR) model and the multiple regime STAR (MRSTAR) model.

If one estimates the river flow based on the observed time series X:

$$X = (x_1, x_2, ..., x_t, x_{t+1}, ..., x_m) \tag{3}$$

$TAR(p_1, p_2, \ldots, p_L, r_1, r_2, \ldots, r_L, d)$ could be identified as follows:

$$\begin{cases} X_t = \varphi_0^l + \sum_{j=1}^{p_l} \varphi_j^l X_{t-j} + \varepsilon_t^l \\ X_{t-d} \in R_l \qquad l = 1, 2, ..., L \end{cases} \tag{4}$$

where $R_l = (r_{l-1}, r_l)$ is the threshold, d is the time lag, and $\varphi^l_j (j = 1, 2, \ldots, p_l)$ is the coefficient of the TAR model. Calibration of the $TAR(p_1, p_2, \ldots, p_L, r_1, r_2, \ldots, r_L, d)$ consists of three stages: identification, parameter estimation and diagnostic checking. For given values of delay parameter d and threshold region $R_l = (r_{l-1}, r_l)$, the coefficients in equation (4) can be estimated by fitting separate autoregressive models with the standard least squares to the appropriate subsets of the time series data. The threshold values can be chosen for which an information criterion such as the Normalized Akaike's Information Criterion (*NAIC*) attains its minimum when d remains fixed and selecting that value of d for which $NAIC(d)$ attains its minimum value by repeating the above processes. Readers can refer to Tong (1990) for a more detailed discussion in the TAR model. Finally, with the information obtained from this calibration process, the number of clusters, which is also the number of FFNNs eventually used, is designated as L in this study.

Network structure selection

Both validation and calibration performance of ANNs are highly dependent on the structure of the neural network. If the architecture is too large the network may not converge during the training, or it may over-fit the data and memorize the river flow history rather than generalizing it. On the other hand, if the network is too small it may not have sufficient degrees of freedom to correctly learn the underlying river flow process. Some guidelines and criteria (e.g. the Schwartz Information Criterion (SIC) and the Akaike's Information Criterion (AIC)) have been proposed in deciding the network structure. However, the authors are not aware of any well-recognized algorithm that can determine the optimal number of layers or nodes in the hidden layers under all circumstances.

Based on the "lag component" of the TAR model of river flow time series, the numbers of neurons in the input and output layers of the lth network are set as $p_l (l = 1, 2, \ldots, L)$ and 1, respectively. A trial-and-error procedure was used in this study for determining the optimal number of hidden nodes, which varies within the range set by a widely applied heuristic as suggested by Weigend *et al.* (1992) (also see Lachtermacher & Fuller, 1995):

$$1.1NP(l) \le 10NH(l)[NI(l)+1] \le 3NP(l) \tag{5}$$

where $NI(l)$, $NH(l)$ represent, respectively, the number of neurons in the input layer and hidden layer of the neural network l, and $NP(l)$ is the number of training samples of the lth neural network.

Based on the calibration results of the TAR model, the training pairs (TP) for the lth network can be obtained from the observed time series, which can be represented by the following expression:

$$\mathrm{TP}_l = (x_{t-p_l}, x_{t-p_l+1}, ..., x_{t-1}; x_t. \quad t \in (1, T)) \tag{6}$$

where $X_{t-p_l}, X_{t-p_l+1}, ..., X_{t-1}$ represent the inputs to the lth neural network, and X_t is the desired output of the tth training samples.

Modification of the training objective function

Minimization of the sum of squares of errors (*SSE*) between the estimated and observed flows is generally used as an objective function for training neural networks, with both of the flows being initially scaled to lie within the range [0–1] during the process of data normalization. The advantages of the SSE are that (a) it can be easily calculated, (b) it has a simple closed-form solution for a linear system which penalizes large errors, and (c) it lies close to the body of normal distribution (Dawson & Wilby, 1998). However, the generalization capacity for the test patterns when river flow is small was often found to be less accurate than that of other test patterns (Hu, 1996). Hsu *et al.* (1995) observed that the low flows were consistently overestimated in their daily rainfall–runoff (RR) modelling. Similar results were also reported in the ANN applications to the RR process modelling. This may be partially attributed to the objective function that induces overemphasized high-flow events, higher nonlinearity inherent in extreme flow events, and stringent requirements, such as normality and independence about the errors.

To improve training and/or generalization performance, a number of alternative accuracy measures have been investigated including the mean squared error (*MSE*), the mean absolute percentage error (*MAPE*), the median absolute percentage error (*MdAPE*), the geometric mean relative absolute error (*GMRAE*), and the AIC and BIC (Drucker & Cun, 1992; Liano, 1996; Ooyen & Nienhuis, 1992). In the context of RR modelling, Dawson & Wilby (1998) used the mean squared relative error (*MSRE*) and the root mean squared error (*RMSE*) as performance criteria in the application of ANN to flow forecasting in two flood-prone UK catchments. Hsu *et al.* (1995) suggested the fitting criterion used for calibrating should be based on matching the logs of flows.

In addition, when errors do not follow a Gaussian distribution, or when little is known about the system, the use of the sum of absolute errors between the estimated and observed flows as the objective function is considered a more appropriate choice (Bos, 1989; Cottrell *et al.*, 1995). In this connection, the weighted sum of the absolute deviations between outputs of the network and the desired outputs is adopted as the objective function in order to strike a balance between high- and low-flow events in this study. Thus, the calibration of the RDNN model can be represented by the following nonlinear programming (NLP) problem:

$$
\text{NLP}\begin{cases}\min F=\sum_{p}\sum_{k}\left|T_{pk}-OO_{pk}\right| \\ \text{subject to } OO_{pk}-f(\Sigma w_{kj}^{\mathrm{S}}OH_{pj}-\theta_k^{\mathrm{O}})=0 \\ \qquad\qquad OH_{pj}-f(\Sigma w_{ji}^{\mathrm{F}}OI_{pi}-\theta_j^{\mathrm{H}})=0\end{cases} \tag{7}
$$

Learning algorithm for the RDNN

The steepest gradient descent algorithm is generally used for calibrating the FFNNs as specified in equation (2), which gave rise to the well-known BP algorithm. The gradient descent algorithm is not the best option for calibrating FFNNs and has been widely criticized for slow convergence, inefficiency, and lack of robustness. Furthermore, it can be very sensitive to the choice of learning rate. A wide variety of alternative learning algorithms have been developed with varying degrees of success (Falman, 1988). Taking advantage of the derivative of the sigmoid function being less than or equal to 0.25, successive linear programming (SLP) was employed for parameter estimation of the RDNN in this study. The SLP methods solve a sequence of linear programming approximations to a nonlinear programming problem. Griffith & Stewart (1961) developed one of the earliest approaches, which is the method of approximation programming (MAP). This algorithm has been widely applied in nonlinear optimization practice (Baker & Lasdon, 1985). The popularity of the SLP is due to its ability to utilize existing LP codes and to solve large problems as well as a history of successful performance. Giving initial values for the weights and thresholds of the neural networks, a linear programming is solved. If the new point provides an improvement, it becomes the incumbent and the process is repeated. If the new point does not yield an improvement, step bounds assigned for the weights and thresholds may need to be reduced or one may be close enough to an optimum for stopping. This process is repeated until a pre-specified value for the objective function is attained, and that depends on the number of training samples involved in the case study. Readers are referred to Hu (1997) for details of the calibration procedure.

APPLICATION OF THE RANGE-DEPENDENT NEURAL NETWORK

The RDNN was applied to two basins in China to evaluate its ability to predict the river flow at the flow gauging stations of two different rivers, Danjiangkou station on the Hanjiang River in Hubei Province and Pingshan station on the Jingsha River in Sichun Province. For comparison purposes, the BP neural network and TAR models were also applied to the river flow prediction at the two gauging stations. The prediction accuracy was quantified by the following two measures; the first is the model efficiency criterion R^2:

$$
R^2=1-\frac{\sum_{p=1}^{P}\left(T_{pk}-OO_{pk}\right)^2}{\sum_{p=1}^{P}\left(T_{pk}-\overline{T}_{pk}\right)^2}\qquad k=1 \tag{8}
$$

The higher the value of R^2, the better is the forecasting performance. If R^2 is equal to 1, it implies that the forecast replicates observation 100% of the time. If all forecast values are equal to the long-term observed mean, R^2 would assume the value of 0. If R^2 is less than 0, however, it implies that the forecast is worse than the long-term observed mean. The second measure is the absolute relative error (*ARE*):

$$ARE_p = \left| \frac{(T_{pk} - OO_{pk})}{T_{pk}} \right| \times 100 \qquad k = 1;\ p = 1, 2, \ldots, P \tag{9}$$

Here, a value of 0 for the *ARE* implies a perfect forecast. Therefore the lower the values of *ARE*, the better is the forecasting accuracy.

Annual river flow estimation at Danjiangkou power station

Whether China is able to feed herself in future depends largely on the ability of the country's water resources to cope with the rapid rate of development. Unfortunately, the geographical variability between regions is striking, with water availability per capita as high as 32 000 m^3 $year^{-1}$ in the southeastern coastal region of China and at a low of 225 m^3 $year^{-1}$ in the North China Plain. As a result, China is now considering a large-scale, inter-basin, water transfer plan—the South-to-North Scheme. This scheme is a strategic means to solve the water resources shortage in north China, to enable advancements in productivity in areas now frequently plagued by droughts. The water transfer scheme includes three projects, the West, Middle and East Routes, each of which serves a different area.

Before launching such an expensive water-transfer scheme, it is crucial to know both how much water is currently available to be diverted from the Danjiangkou reservoir and the potential change in water availability under future weather con-

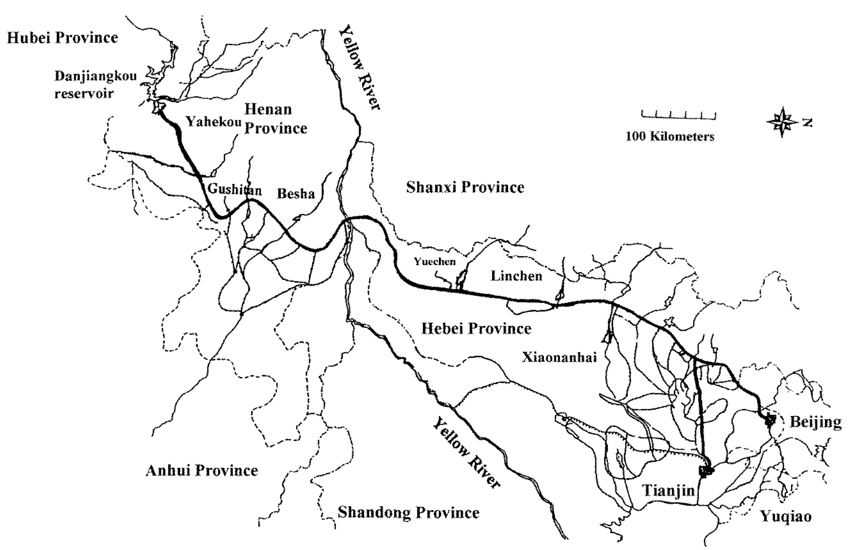

Fig. 1 Relative positions of the MRP and Danjiangkou reservoir in China.

ditions due to global warming. Therefore, an annual river flow prediction is extremely important to the planning and proper management of such a large water resources system across a range of different climatic conditions and watershed area magnitudes. Figure 1 presents the relative positions of the Middle Route Project (MRP) in China.

The Danjiangkou reservoir has a catchment area of 95 200 km^2 and a gross storage capability of 30×10^9 m^3. The installed capacity of the hydropower plant is 900 MW. The annual reservoir inflows to the Danjiangkou reservoir for a 35-year time period (1956–1990) and a 60-year time period (1930–1990) are 40.85 and 40.66×10^9 m^3, respectively. The objectives of this project include flood control, power generation and irrigation at its present scale. The MRP plans to serve an area of 1 550 000 km^2 with a population of 106 810 000 based on an average annual rainfall of 658.3 mm.

Figure 2 shows the historical time series of annual river flow at Danjiangkou station. Also shown is the average annual river flow (dashed line). Annual reservoir inflow data are available from 1931 to 1995. The three models, i.e. the TAR, BP and RDNN, were calibrated using data extending from 1931 to 1985, and then tested for the period 1986–1995. This test period was selected to examine the capacity of the neural network in estimating low-flow events because most of the annual inflows then were less than the average inflow. The following TAR model was calibrated for the Danjiangkou reservoir inflow time series:

$$x_t = \begin{cases} 1614.14 - 1.071x_{t-1} + 0.8501x_{t-2} + 0.2554x_{t-3} & x_{t-1} > 1360.67 \\ 1378.22 + 0.044x_{t-1} - 0.3405x_{t-2} & x_{t-1} \le 1360.67 \end{cases} \tag{10}$$

Based on this calibration result and the trial-and-error procedure outlined above, two separate neural networks, RDNN1 (2,3,1) and RDNN2 (3,3,1), have been identified to model the Danjiangkou reservoir inflow when it is greater than 1360.67 m^3 s^{-1} or lower than 1360.67 m^3 s^{-1}, respectively. The numbers inside the parentheses denote the number of nodes in the input, hidden and output layers, respectively.

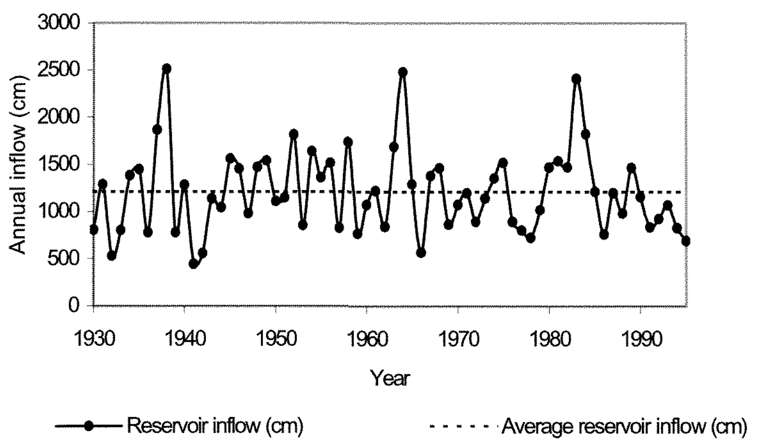

Fig. 2 Annual reservoir inflow for the Danjiangkou station, Hubei province, China.

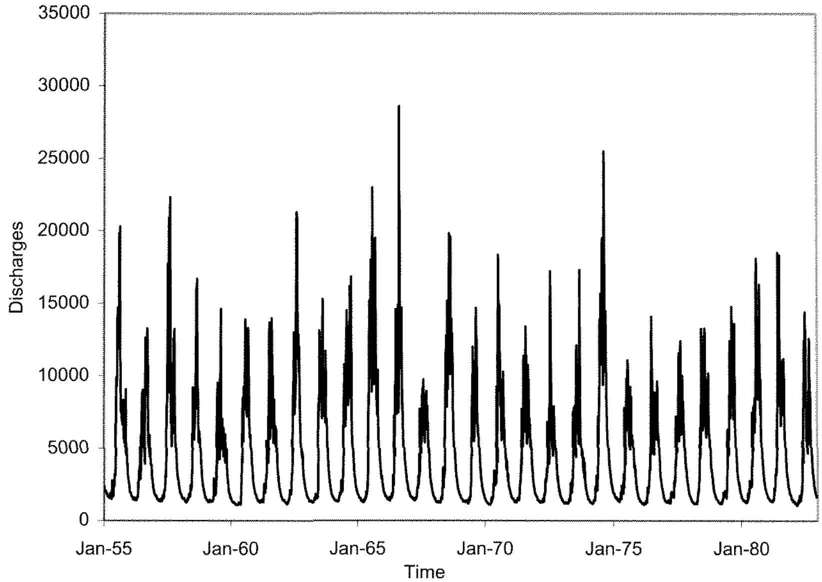

Fig. 3 The daily time series at Pingshan gauging station and the threshold values of the corresponding TAR model.

Daily river flow prediction at Pingshan gauging station

A set of 28 years of data (1955–1982, Fig. 3) was used for calibration and validation of the TAR, BP and RDNN models in this study. The first 10 water years of data were used for the three model calibrations, while the remaining 18 water years (1965–1982) were used for the model validations. Likewise, the following model $\mathrm{TAR}(3, 5, 3, 2030, 7180, 1)$ was obtained based on the observed daily flow from 1955 to 1964 at Pingshan gauging station:

$$x_t = \begin{cases} 1.36x_{t-1} - 0.14x_{t-2} - 0.12x_{t-3} & x_{t-1} < 2030 \\ 1.52x_{t-1} - 0.39x_{t-2} - 0.05x_{t-3} - 0.10x_{t-4} - 0.04x_{t-5} & 2030 \le x_{t-1} \le 7180 \\ 1.63x_{t-1} - 0.82x_{t-2} + 0.13x_{t-3} & x_{t-1} > 7180 \end{cases} \tag{11}$$

where 2030 and 7180 are the two threshold values. As a result, three RDNN neural models, RDNN1(3, 1, 1), RDNN2(5, 2, 1), RDNN3(3, 2, 1), were identified to model the low-flow, medium-flow and high-flow events, respectively. After training with its own training set, one of the three RDNNs was selected every time in order to generate a 1-day-ahead prediction based on the values of X_{t-1}. Comparisons of the 1-day lead-time forecasts of daily flow for the BP network with different architectures, RDNN and TAR models 1 in the calibration and verification periods are given in Table 1.

Analysis of results

In the case of annual reservoir inflow estimation, Fig. 4 shows the forecasting results of the BP, RDNN and TAR models for annual inflows in terms of the *ARE*. It can be

Table 1 One-day ahead prediction using different models for Pingshan station.

Models	Calibration period:		Verification period:	
	R^2	Maximum *ARE*	R^2	Maximum *ARE*
BP(3, 2, 1)	73.36	95.73	65.64	239.76
BP(3, 3, 1)	71.35	85.37	69.36	221.46
BP(3, 5, 1)	75.57	62.34	62.35	279.27
BP(3, 6, 1)	72.14	76.64	66.63	247.77
RDNN	97.36	47.89	91.27	89.74
TAR	77.42	69.74	68.68	241.37

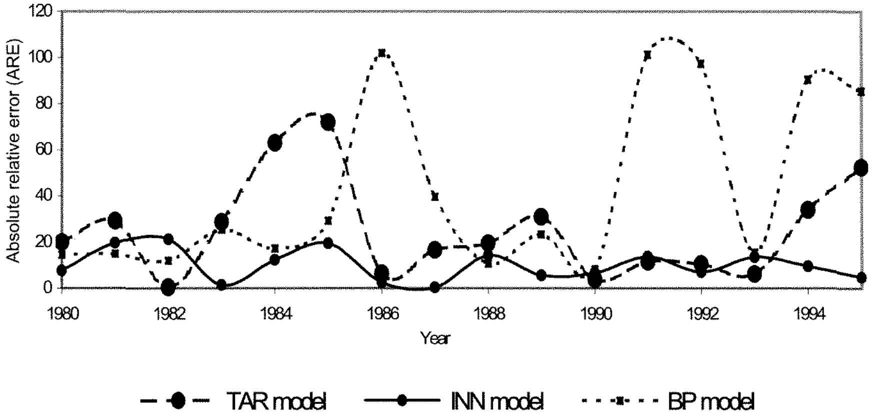

Fig. 4 Comparison of the absolute relative errors of the BP, INN and TAR models.

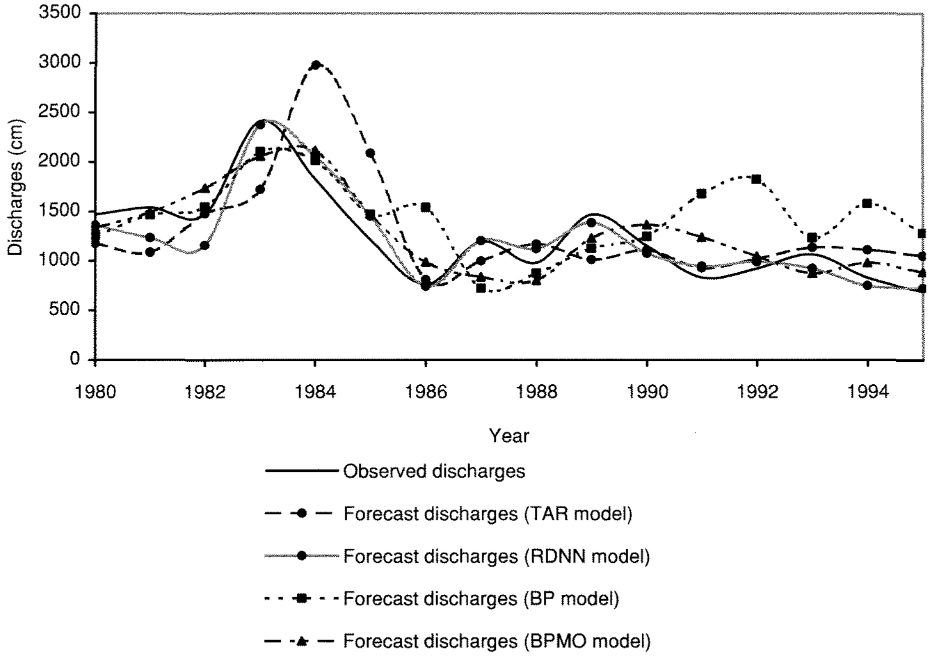

Fig. 5 Comparisons of forecast discharges at Danjiangkou by different models.

seen from Fig. 4 that the absolute relative errors of the RDNN model were the lowest in 12 out of 16 years of annual reservoir inflow estimation, indicating that the RDNN model performs best on average in terms of the *ARE* statistics. On average, the TAR model compared favourably with results obtained using the BP model.

Furthermore, examination of the results of Figs 4 and 5 indicates that the performance of the BP deteriorates as the magnitude of annual inflow decreases, with the two worst *ARE* values of the BP occurring in the two driest years, 1986 and 1991. In order to have a better insight into model generalization performance for low, medium and high annual inflows, comparisons were made between the performance of the BP network and those of the BP model with the modified objective function (BPMOF) as shown in equation (7) and the RDNN model. Figure 6 presents the scatter plots of the observed *vs* 1-year ahead forecast of annual inflows from 1981 to 1995 using the BP, the BPMOF and the RDNN. Figure 6(a) shows that the scatter plot was well spread over the 45° line when the observed inflows were large enough, i.e. greater than the average annual reservoir inflow of 1208.63 $m^3 s^{-1}$, while forecasts for the low and medium inflows tended to dramatically diverge from the 45° line. In other words, the BP model seemed to be accurate for forecasting inflows that are higher than approximately 1360.67 $m^3 s^{-1}$, which is the threshold value of the TAR model for the inflow time series of the Danjiangkou reservoir, while its performance deteriorated when the inflows were lower than 1360.67 $m^3 s^{-1}$.

A comparison of the forecasting results of the BPMOF model (Fig. 6(b)) with those of the BP model (Fig. 6(a)) indicates that the modification of objective function improves the predictive accuracy for low-flow events, especially for inflows being lower than 1000 $m^3 s^{-1}$. This may be due to the fact that the combined effect of the SSE objective function and output data normalization resulted in an over-emphasized high-flow event at the expense of the relatively poor low-flow accuracy. Further, as shown in Fig. 6(c), inflows across the entire flow range tend to fall closer to the 45° line, indicating an improved generalization performance of the RDNN model. Moreover, it has also indicated that the RDNN model improved its *ARE* performance for medium and low inflows without diminishing its performance for high inflows. This may be attributed to the combined effect of training pairs clustering and objective function modification.

Referring to the daily streamflow prediction in Pingshan gauging station, it was found that the BP(3, 3, 1) network gave rise to both the highest R^2 coefficient and the lowest maximum *ARE* among 50 combinations of network architectures examined, with input and hidden nodes varying from 1 to 5 and 1 to 10, respectively. Table 1 presents the best four BP networks. It was noted that the high sum of square errors between the observed and simulated inflows during network training could not be reduced to a desired level by increasing the number of hidden nodes to more than seven when the number of input nodes was three. This problem in turn has resulted in a difficulty in the generalization improvement. In addition to this 3-7-1 case, similar results were also observed with different network architectures such as 1-12-1, 2-9-1 and 4-8-1.

Statistics shown in Table 1 suggest that the RDNN model performed significantly better than the best BP network, in both the calibration and verification periods. It can be seen from Table 1 that the combined effect of objective function modification and training pairs clustering increased the R^2 from 0.6936 of the BP(3, 3, 1) network to

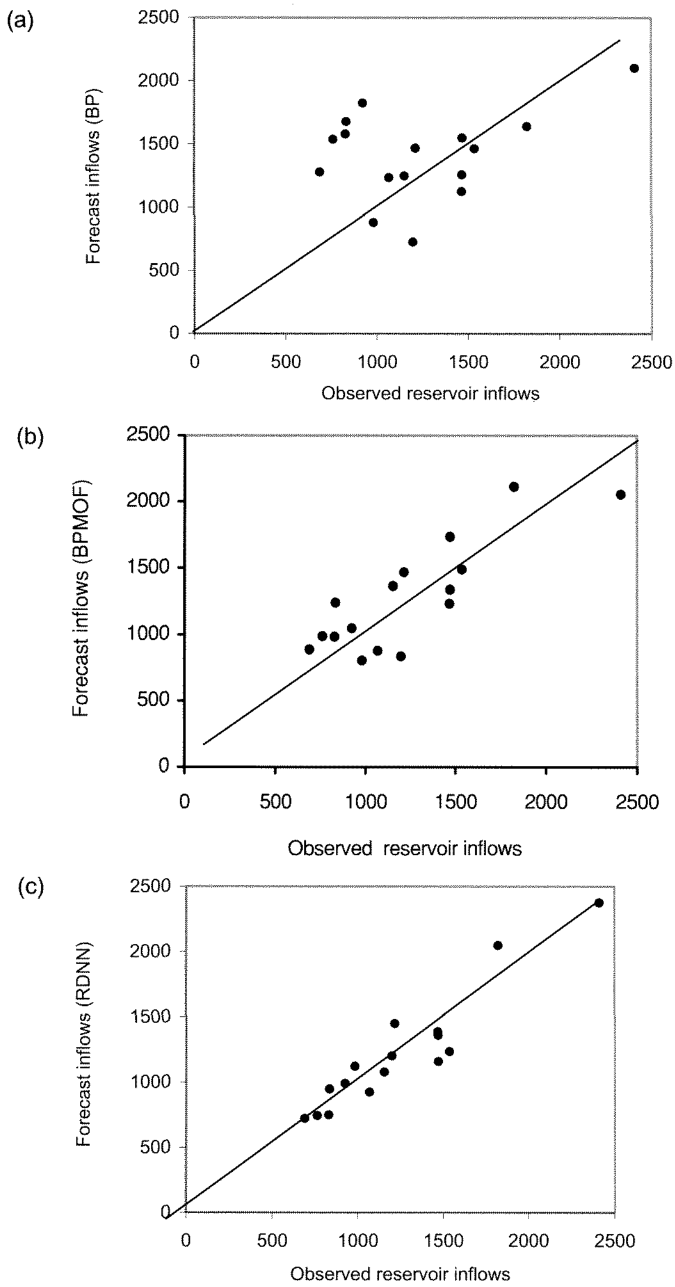

Fig. 6 Scatterplots of forecast inflows against observed inflows.

0.9127 of the RDNN network and decreased the maximum *ARE* from 221.46% of the BP(3, 3, 1) to 89.74% of the RDNN network in the verification period. Furthermore, a significant improvement of the R^2 values from 0.7135 of the BP(3, 3, 1) to 0.9736 of the RDNN network in the calibration period suggests that the RDNN network was effective in solving the aforementioned high calibration error problem, which was experienced by the BP network in the case of the Pingshan station.

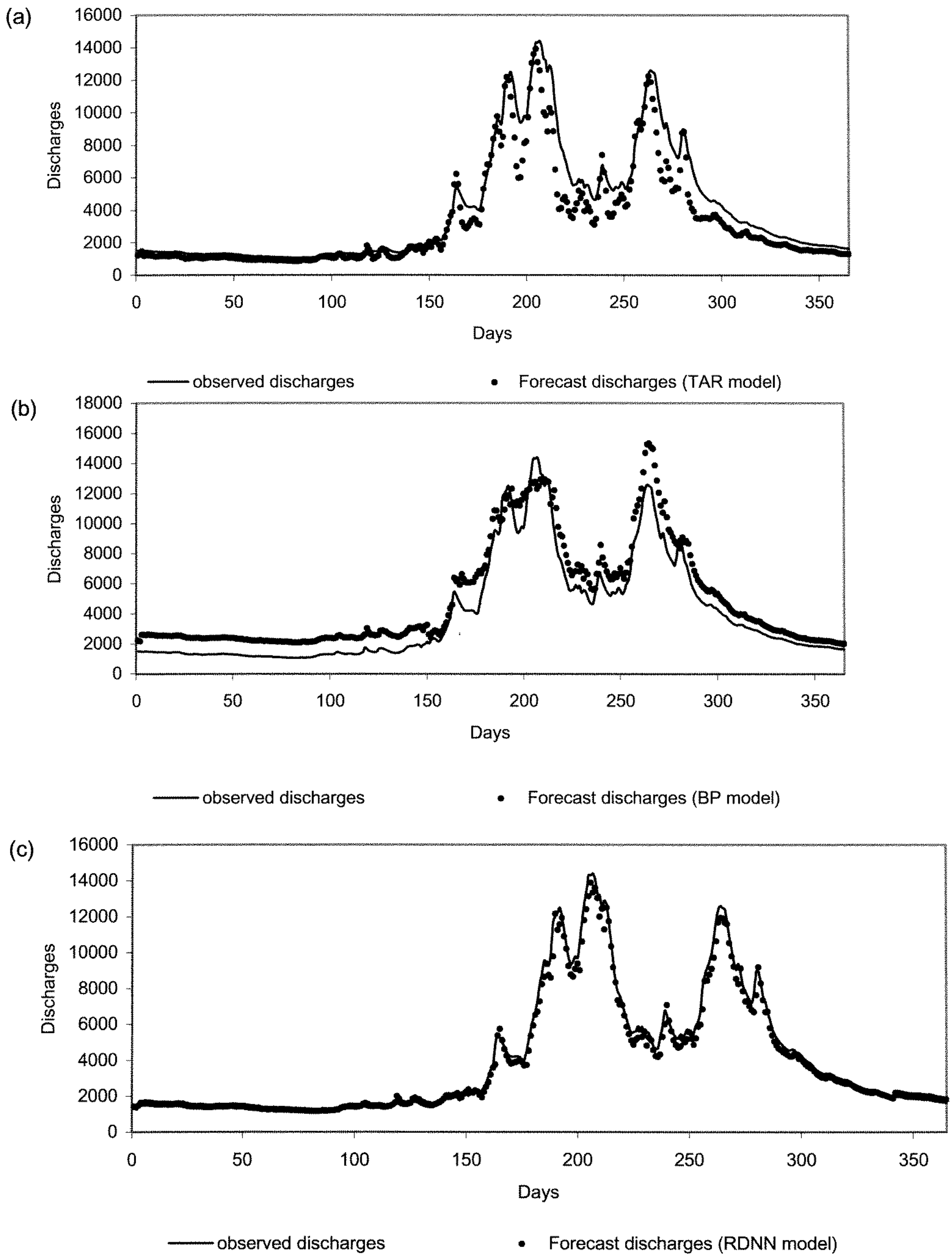

Fig. 7 Comparisons of the observed and estimated discharge hydrographs of (a) the TAR model, (b) the BP model and (c) the RDNN model.

Referring still to the Pingshan case, the BP model was able to predict a higher R^2 than the TAR model, which is different from its performance ranking status in the case of Danjiangkou reservoir inflow forecast. Nevertheless, the R^2 value of 0.6868 for the TAR model was just marginally lower than that of 0.6936 for the BP(3, 3, 1) network. Furthermore, the TAR model had a better calibration performance than the BP(3, 3, 1) network.

In order to indicate the differences in model performance, Fig. 7 presents a comparison of the observed and estimated discharge hydrographs of Pingshan station in 1982 using the TAR, BP and RDNN models. The TAR model tended to slightly underestimate the flows, while the BP(3, 3, 1) network consistently overestimated low flow in this study. The RDNN network predicted the entire range of flows very well with a slight tendency to underestimate the high flows. Similar observations of the three individual models for other validation years were found and were not included in Fig. 7 for brevity. The more even error distributions across the entire flow range, as shown in both Figs 6(c) and 7(c) suggest that the RDNN network is likely to be more effective in forecasting low-flow events as compared to the conventionally used BP network.

DISCUSSION AND CONCLUSIONS

A new neural network was proposed for better hydrological time series modelling, which is based on clustering calibration data into low-, medium- and high-flow sets and modification of the training objective function. The applicability and potential of the improved network in daily streamflow and annual reservoir inflow time series prediction were examined using data from two basins in China.

Daily river flow and annual reservoir inflow prediction in the two basins demonstrated that the RDNN significantly enhanced the network generalization performance, especially for low-flow events, as compared to the BP network. The high sum of square errors between the observed and simulated flows was found irreducible to the desired error level by increasing complexity of the BP network structure in the Pingshan case study. However, the RDNN network was shown to produce better calibration performance in the Pingshan station, thereby resulting in better generalization performance in this case. This finding indicates that several simple but adequate networks, which had been adopted in the RDNN structure, are likely to provide a better model performance in hydrological modelling than a single complex BP network.

More accurate prediction from a neural network for hydrological variables depends on its structure, learning algorithm, activations, and sensitivity and on the nature of specific problem being studied. These factors, coupled with the uncertainty and inaccuracy of hydrological data, make generalization improvement a very complex task. This study is by no means conclusive in proving that the RDNN network is the optimal form of neural network for river flow time series prediction. It merely provides a first step towards an improvement in generalization performance of neural networks.

The implication of this investigation is that more attention should be paid to enhancing the generalization performance of neural networks for hydrological prediction rather than simply focusing on improving learning performance of the FFNNs.

Acknowledgements The authors gratefully acknowledge the helpful comments and suggestions from the two anonymous reviewers, which are reflected in the present paper. This research was funded by both the He Yindong Foundation (Project no. 71071) and the Natural Science Foundation of China (Project no. 59609004). The authors would like to thank Professor Xia Jun and Professor Guo Shenglian for their helpful suggestions.

REFERENCES

ASCE Task Committee on Artificial Neural Networks in Hydrology (2000a) Artificial neural networks in hydrology. I. Preliminary concepts. *J. Hydrol. Engng* **5**(2), 115–123.

ASCE Task Committee on Artificial Neural Networks in Hydrology (2000b) Artificial neural networks in hydrology. II. Hydrologic applications. *J. Hydrol. Engng* **5**(2), 124–137.

Amorocho, J. & Hart, W. E. (1964) A critique of current methods in hydrologic systems investigation. *Trans. Am. Geophys. Union* **45**, 307–321.

Atiya, A. F., El-Shoura, S. M., Shaheen, S. I & El-Sherif, M. S. (1999) A comparison between neural-network forecasting techniques—case study: river flow forecasting, *IEEE Trans. Neural Networks* **10**(2), 402–409.

Baker, T. E. & Lasdon, L. S. (1985) Successive linear programming at Exxon. *Manage. Sci.* **31**(3), 264–274.

Bos, A. V. D. (1989) Nonlinear least-absolute values and minimal model fitting. *Automatica* **24**, 803–809.

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

Cottrell, M., Girard, B., Girard, Y., Mangeas, M. & Muller, C. (1995) Neural modeling for time series: a statistical stepwise method for weight elimination. *IEEE Trans. on Neural Networks* **6**(6), 1355–1364.

Coulibaly, P., Anctil, A. & Bobée, B. (2000) Daily reservoir inflow forecasting using ANNs with stopped training approach. *J. Hydrol.* **230**, 245–257.

Dawson, C. W. & Wilby, R. (1998) An artificial neural network approach to rainfall–runoff modelling. *Hydrol. Sci. J.* **43**(1), 47–66.

Drucker, H. & Cun, Y. L. (1992) Improving generalization using double backpropagation. *IEEE Trans. Neural Networks* **3**(6), 991–997.

Falman, S. E. (1988) An empirical study of learning speed in back-propagation networks. Tech. Report CMU-CS-88-162, Carnegie Mellon University, Pittsburgh, Pennsylvania, USA.

French, M. N., Krajewski, W. F. & Cuykendall, R. R. (1992) Rainfall forecasting in space and time using a neural network. *J. Hydrol.* **137**,1–31.

Griffith, R. E. & Stewart, R. A. (1961) A nonlinear programming technique for the optimization of continuous processing systems. *Manage. Sci.* **7**, 379–392.

Hampshire, J. & Waibel, A. (1989) The meta-pi network: building distributed knowledge representation for robust pattern recognition. Tech. Report CMU-CS-89-166, Carnegie Mellon University, Pittsburgh, Pennsylvania, USA.

Hsu, K.-L., Gupta, H. V. & Sorooshian, S. (1995) Artificial neural network modelling of the rainfall–runoff process. *Wat. Resour. Res.* **31**(10), 2517–2530.

Hu, T. S. (1996) Hydrological time series prediction using radial basis function network. *J. Wuhan Univ. Hydraul. Electr. Engng* **29**(6), 1–6 (in Chinese).

Hu, T. S. (1997) *Neural Optimization and Prediction* (in Chinese). Dalian Maritime University Press, Dalian, China.

Hughes, D. A. & Murrell, H. C. (1986) Nonlinear runoff routing—a comparison of solution methods. *J. Hydrol.* **85**, 339–347.

Imrie, C. E., Durucan, S. & Korre, A. (2000) River flow prediction using artificial neural networks: generalization beyond the calibration range. *J. Hydrol.* **233**, 138–153.

Jacobs, R. A., Jordan, M. I., Nowlan, S. J. & Hilton, G. E. (1991) Adaptive mixtures of local experts. *Neural Comput.* **3**, 79–87.

Jain, S. K., Das., A. & Srivastava, D. K. (1999) Application of ANN for reservoir inflow prediction and operation. *J. Wat. Resour. Plan. Manage. ASCE* **125**(5), 263–271.

Kachroo, R. K. & Natale, L. (1992) Nonlinear modelling of the rainfall–runoff models. *J. Hydrol.* **135**, 341–369.

Karunanithi, N., Grenney, W. J., Whitley, D. & Bovee, K. (1994) Neural networks for river flow prediction. *J. Comput. Civil Engng* **8**(2), 201–219.

Kember, G. & Flower, A. C. (1993) Forecasting river flow using nonlinear dynamics. *Stochast. Hydrol. Hydraul.* **7**, 205–212.

Lachtermacher, G. & Fuller, J. D. (1995) Backpropagation in time-series forecasting. *J. Forecasting* **14**, 381–393.

Liano, K. (1996) Robust error measure fir supervised neural network learning with outliners. *IEEE Trans. Neural Networks* **7**(1), 246–250.

See, L. & Openshaw, S. (1999) Applying soft computing approaches to river level forecasting. *Hydrol. Sci. J.* **44**(5), 763-778.

See, L. & Openshaw, S. (2000) A hybrid multi-model approach to rover level forecasting. *Hydrol. Sci. J.* **45**(4), 523-536.

Liong. S.-Y., Lim, W. H. & Paudyal, G. N. (2000) River stage forecasting in Bangladesh: neural network approach. *J. Comput. Civil Engng* **14**(1), 1–8.

Lorrai, M. & Sechi, G. M. (1995) Neural nets for modelling rainfall–runoff transformation. *Wat. Resour. Manage.* **9**(4), 299–313.

Luk, K. C., Ball, J. E. & Sharma, A. (2000) A study of optimal model lag and spatial inputs to artificial neural network for rainfall forecasting. *J. Hydrol.* **227**, 56–65.

Maier, H. R. & Dandy, G. C. (2000) Neural networks for the prediction and forecasting of water resources variables: a review of modeling issues and applications. *Environ. Modeling & Software* **15**, 101–124.

Mason, J. C., Tem'me, A. & Price, R. K. (1996) A neural network model of rainfall–runoff using radial basis function. *J. Hydraul. Res.* **34**, 537–548.

Minns, A. W. & Hall, M. J. (1996) Artificial neural networks as rainfall–runoff models. *Hydrol. Sci. J.* **41**(3), 399–417.

Ooyen, A. V. & Nienhuis, B. (1992) Improving the convergence of back propagation problem. *Neural Networks* **5**, 465–471.

Raman, H. & Chandramouli, H. (1996) Deriving a general operating policy for reservoir using neural networks. *J. Wat. Resour. Plan. Manage.* **122**(5), 342–347.

Raman, H. & Sunilkumar, N. (1995) Multivariate modelling of water resources time series using artificial neural networks. *Hydrol. Sci. J.* **40**(2), 145–163.

Rodrigues, M. J. & Serodes, J. B. (1996) Neural network-based modeling of the adequate chlorine dosage for drinking water disinfection. *Can. J. Civil Engng* **23**, 621–631.

Rodriguez-Iturbe, I., Febres de Power, B., Sharifi, M. B. & Georgakakos, K. P. (1989) Chaos in rainfall. *Wat. Resour. Res.* **25**(7), 1667–1675.

Rumelhart, D. E., Hilton, G. E. & Williams, R. J. (1986) Learning representations by backpropagating errors. *Nature* **323** (6188), 533–536.

Sajikumar, N. & Thandaveswara, B. S. (1999) A nonlinear rainfall–runoff model using an artificial neural network. *J. Hydrol.* **216**, 32–55.

Shamseldin, A. Y. (1997) Application of a neural network technique to rainfall–runoff modelling. *J. Hydrol.* **199**, 272–294.

Smith, J. & Eli, R. N. (1995) Neural network models of rainfall–runoff process. *J. Wat. Resour. Plan. Manage.* **121**(6), 499–508.

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

Tong, H. (1978) On a threshold model. In: *Pattern Recognition and Signal Processing* (ed. by C. H. Chen). Chapter 5. Sijthoff & Noordhoff, Amsterdam, The Netherlands.

Tong, H. (1990) *Nonlinear Times Series: A Dynamical System Approach.* Oxford University Press, Oxford, UK.

Tong, H. & Lim, K. S. (1980) Threshold autoregression, limit cycles and cyclical data (with discussion). *J. Roy. Statist. Soc.* Ser. B **42**, 245–292.

Weigend, A. S., Rumelhart, D. E. & Huberman, B. A. (1992) Predicting the future: a connectionist approach. *Int. J. Neural Systems* **1**(3), 193–209.

Whitley. R. (1999) Approximate confidence intervals for design floods for a single site using a neural network. *Wat. Resour. Res.* **35**(1), 203–209.

Wilcox, B. P., Seyfried, M. S. & Martison, T. H. (1981) Searching for chaotic dynamics in snowmelt runoff. *Wat. Resour. Res.* **27**(6), 1005–1010.

Zealand, C. M., Burn., D. H. & Simonovic, S. P. (1999) Short term stream-flow forecasting using artificial neural network. *J. Hydrol.* **214**, 32–48.

Zhang, B. & Govindaraju, R. S. (2000) Prediction of watershed runoff using Bayesian concepts and modular neural networks. *Wat. Resour. Res.* **36**(3), 753–763.

Received 31 October 2000; accepted 21 May 2001