HYDROLOGICAL PROCESSES
Hydrol. Process. **26**, 281–296 (2012)
Published online 29 July 2011 in Wiley Online Library
(wileyonlinelibrary.com) DOI: 10.1002/hyp.8227

Simulating and predicting river discharge time series using a wavelet-neural network hybrid modelling approach

Shouke Wei,[1]* Jinxi Song[2] and Nasreen Islam Khan[1,3]

[1] *Eawag, Swiss Federal Institute of Aquatic Science and Technology, Ueberlandstrasse 133, CH-8600 Dübendorf*
[2] *College of Urban and Environmental Sciences, Northwest University, CN-710127 Xi'an, China*
[3] *Department of Geography and Environment, Dhaka University, Dhaka-1000, Bangladesh*

Abstract:

Accurate simulation and prediction of the dynamic behaviour of a river discharge over any time interval is essential for good watershed management. It is difficult to capture the high-frequency characteristics of a river discharge using traditional time series linear and nonlinear model approaches. Therefore, this study developed a wavelet-neural network (WNN) hybrid modelling approach for the predication of river discharge using monthly time series data. A discrete wavelet multiresolution method was employed to decompose the time series data of river discharge into sub-series with low (approximation) and high (details) frequency, and these sub-series were then used as input data for the artificial neural network (ANN). WNN models with different wavelet decomposition levels were employed to predict river discharge 48 months ahead of time. Comparison of results from the WNN models with those of the ANN models alone indicated that WNN models performed a more accurate prediction. Copyright © 2011 John Wiley & Sons, Ltd.

KEY WORDS river discharge; wavelet; artificial neural network; Weihe River

Received 1 December 2010; Accepted 5 July 2011

INTRODUCTION

Accurate simulation and prediction of water discharge of a river for both short and long periods (yearly, monthly, daily or hourly) is essential for various activities in hydrological and water resources management, such as limiting pollutant load, calculating sedimentation transportation, controlling flood and drought, determining environmental flow, reservoir operation, agricultural irrigation, water supply to industry and households, to name just as few. A variety of methods have been developed to simulate and predict river discharge, including time series methods (linear and nonlinear), conceptual models, physical models, artificial neural network (ANN) models, and integrated models.

Linear time series modelling approaches exploit the autocorrelation structure in a time series to model future events as a function of the previous time step (Srinivasulu and Jain, 2009). The autoregressive moving average (ARMA) approach is one of the most commonly applied time series techniques in the literature for hydrological estimation and prediction because it is comparatively easy to implement. Many studies have attempted to estimate the parameters of an ARMA model, such as Kalman filter-based methods (Gardner *et al*., 1980; Azrak and Melard, 1998), Goal programming (Mohammadi *et al*., 2006), Clustering (Wang *et al*., 2008), and Artificial Neural Networks (Chenoweth *et al*., 2000). These ARMA models are sufficient for time series analysis where there is no nonlinearity in the data (Abrahart and See, 2000; Hwarng, 2001; Mohammadi *et al*., 2006), data are stationary (Nourani *et al*., 2009), and the prediction period is short (Toth *et al*., 2000). However, data series in nature are usually non-stationary with transitory components (Salerno and Tartari, 2009). River discharge is affected by many different factors, primarily precipitation, temperature, evaporation and water use in the upper streams which supply the river and hence discharge usually has strong nonlinear and non-stationary characteristics. Therefore, these linear-related time series methods cannot accurately capture the variations of river discharge. Furthermore, when more detailed factors like periodic and stochastic variations are included in times series, they become even more complex and difficult for those methods to adequately describe.

To overcome some of these problems, conceptual models (CM) have been developed based on simplified physical laws. These models are generally nonlinear, time-invariant, and deterministic with parameters representing physical characteristics such as topology, vegetation and soil type (Kisi and Cigizoglu, 2007). A CM usually uses simplified forms to incorporate interconnected physical elements, which signify a hydrologic process of the rainfall-runoff transformation (Chen and Adams, 2006). For example, a set of conceptual water balance models were applied to estimate monthly river flow or monthly runoff based on temperature and precipitation (Xu *et al*., 1996; Müller-Wohlfeil *et al*., 2003), and a conceptual rainfall-runoff model was employed

* Correspondence to: Shouke Wei, Eawag, Swiss Federal Institute of Aquatic Science and Technology, Ueberlandstrasse 133, CH-8600 Dübendorf. E-mail: shouke.wei@eawag.ch

with important snowpack in wintertime (Vandewiele and Nir-Lar-Win, 1993). While CMs are reliable for the evaluation of hydrological processes based on relationships between model parameters representing measurable physical characteristics (O'Connor, 1997), they become more unreliable when they attempt to represent complex hydrological process involving a number of interconnected elements, especially when watershed geomorphological characteristics or spatial and temporal variations of model inputs are also considered (Chen and Adams, 2006).

Physically based models can describe the complicated hydrological processes in terms of physical laws expressed by partial differential equations. SWAT (Neitsch *et al*., 2005) and MIKE-SHE (DHI, 2007) are two widely used physical models. These models usually require a large quantity of data, enormous time and large computer resources (Kisi and Cigizoglu, 2007), and also require a good knowledge of the physical laws governing the complex hydrological process. The accuracy of model prediction is usually influenced by subjective components as parameter establishment is dependent on the user's expertise and experience (Duan *et al*., 1992; Tokar and Johnson, 1999; Chang and Chen, 2001; Pulido-Calvo and Portela, 2007). In addition, these modelling approaches will be unattractive and difficult to implement where the datasets are insufficient, especially where only one time series data set is available.

ANN models are a type of black-box model, involving a data-driven approach and a function approximator, which are well suited to complex problems and have a great ability to model and forecast nonlinear hydrologic time series (Kisi and Cigizoglu, 2007; Nourani *et al*., 2009). The advantages of ANN models are that they do not require information on whether the data is from a specific statistical distribution or stationary, nor what the exact relationships amongst the various input variables are (Maier and Dandy, 1996; Kisi and Cigizoglu, 2007). Even though the ANN modelling procedure does not require a detailed knowledge of hydrological characteristics, a well trained ANN model can easily be applied to water resources management issues (Srinivasulu and Jain, 2009). ANN also has the advantages of self-learning, self-organising, and self-adapting (Ertay and Çekyay, 2006; Feng and Hong, 2008), and the technique has been successfully applied to various different fields including hydrological flow estimation and prediction (Jain *et al*., 1999; Zealand *et al*., 1999; Salas *et al*., 2000; Sivakumar *et al*. 2002; Valença, *et al*., 2005; Nilsson *et al*., 2006; Kisi and Cigizoglu, 2007). Many comparison studies have confirmed that ANN models have very good forecasting performance (Tsai and Lee, 1999; Chu and Chen, 2000; Tsai *et al*., 2002; Valença, *et al*., 2005; Feng and Hong, 2008). However, when river discharge time series are characterized by high nonlinearity and are non-stationary with excessive noise, it will still be difficult for ANN to accurately capture river discharge behaviour.

Since model accuracy is a key performance indicator, many studies have employed integrated models to improve modelling accuracy by combining two or more approaches, which include, for example, river flow model combining CM, ANN, elitist real-coded genetic algorithm (ERGA), data-decomposition and model-fusion techniques (Srinivasulu and Jain, 2009), runoff models combining CA and ANN (Nilsson *et al*., 2006; Chen and Adams, 2006), daily runoff prediction models integrating wavelet decomposition and stepwise regression (Li *et al*., 2008), and a band-pass (frequency) filter and auto-regression (AR(p)) integrated model for annual reservoir inflow prediction (Wei, 2008).

More recently, WNN and ANN hybrid analyses have attracted increasing interest due to advantages in terms of good fitting and prediction accuracy compared to other models. Wavelet transformation analysis is therefore becoming a popular analysis technique due to its ability to reveal simultaneously both spectral and temporal information within one signal (Nourani *et al*., 2009). This method advances Fourier analysis, where the basic shortcoming was that the Fourier spectrum contained only globally average information. Wavelet transformation is able to decompose a time series into its sub-series and capture useful information at different resolution levels. A WNN hybrid model was first proposed by Aussem *et al*. (1998) to predict a financial time series. Zhang and Dong (2001) subsequently developed an adaptive neural-wavelet model for short-term load forecasting in the competitive market in Australia. In hydrology and water resource management, Kim and Valdes (2003) applied dyadic wavelet transforms and neural networks to forecast droughts in the Conchos River Basin in Mexico, and showed that the conjunction model significantly improved the ability of neural networks to forecast regional drought. Wang and Ding (2003) applied WNN to predict shallow groundwater levels in Beijing and daily discharge of the Yangtze River in China, with results that suggested the model could increase the forecast accuracy and extend the prediction time. Cannas *et al*. (2006) employed continuous and discrete wavelet transforms and data partitioning to investigate the effects of data processing for river flow forecasting using a neural network, with results that showed that networks trained with pre-processed data performed better than networks trained with an unrecompensed signal. Chen *et al*. (2007) analysed tide forecasting and supplement of tides around Taiwan and South China Sea using WNN, and found that the developed model accurately predicted tide for periods of 1–5 years with improved prediction quality. Partal and Cigizoglu (2008) estimated and forecasted daily suspended sediment data using a wavelet-neural network (WNN), and the forecasting results were very close to the real values. Kisi (2008) applied a neuro-wavelet model to forecast (one-month-ahead) stream flows of the Canakdere and Goksudere Rivers in Turkey and compared the results with those of single Multilayer Perceptron (MLP), multilinear regression (MLR) and auto-regression (AR) models, revealing that the combined WNN model could increase both the forecast accuracy and performance. WNN methods have

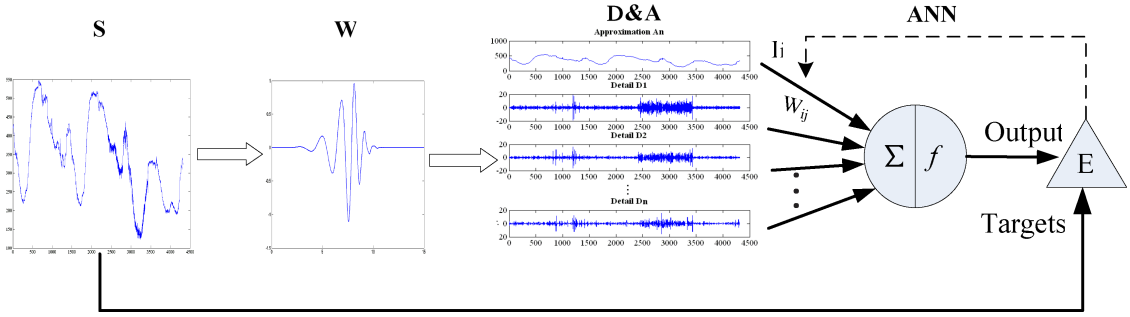

Figure 1. A sketch of the Wavelet-artificial neural network model process: a signal (S) is decomposed by a Wavelet (W) into details and approximation (D&A), which are regarded as the inputs of ANN

also been used in water quality monitoring (Kim *et al.*, 2006), precipitation (Partal and Kişi, 2007; Nourani *et al.*, 2009) monthly level fluctuation modelling (Kisi, 2009), evapotranspiration (Partal, 2009), and understanding the base-flow components in river discharge in karst environment (Salerno and Tartari, 2009).

The main purpose of this paper was to develop a WNN hybrid modelling method, which could estimate and predict river discharge time series. Initially, the paper begins with an introduction of the wavelet and ANN integrated modelling method followed by a description of the study area and data. Subsequent sections present the results and discussion, including model comparison and evaluation, results, interpretation and model generalisation improvement, with the final section concluding the paper with a summary of the study findings.

METHODS

In this study, two modelling techniques, wavelet analysis (WA) and ANN, were combined to produce a WNN hybrid model which integrated the advantages of both wavelet transformation and ANN models. The integrated modelling process included two steps: (1) The river discharge time series, i.e. a signal, was decomposed with Discrete wavelet transformation (DWT) into an approximation (An) with low frequency, and details (D1, D2, …, Dn) with high frequency; and (2) These subseries were used as the input for the network, with the original series as the target (Figure 1). All the analyses were performed using the Wavelet Toolbox™ 4 and Neural Network Toolbox™ 6 in MATLAB.

Wavelet analysis

While Fourier analysis (FA) is a useful traditional tool to transform a stationary signal from a time domain into frequency-based constituent sinusoids, time information on local variations (like trends, breakdown points, discontinuities, and self-similarity) are lost during transformation to a non-stationary signal. Dennis Gabor's adaptation, the Short-Time Fourier Transform (STFT) (1946), was an effort to correct FA's deficiency by analysing only a small section of the signal at a time, called windowing the signal (Misiti *et al.*, 2008). The drawback of STFT was that the size of the time window was fixed for all frequencies, which made the analysis less precise. Wavelet analysis (WA) is a promising time-frequency technique for signal analysis, and has several advantages over the traditional Fourier analysis (FA) and the adapted 'Short-Time Fourier Transform'. WA is an improved version of the 'short time Fourier transform' which can elucidate time characteristics in data (Daubechies, 1990; Rioul and Vetterli, 1991). WA is a windowing technique with variable-sized regions, which allows the use of long time intervals for low-frequency information, and shorter regions for high-frequency information. In this sense, it is usually characterized as a *mathematical microscope* (Wang and Ding, 2003). There are two main types of wavelet transformation methods, *continuous wavelet transformation* and *discrete wavelet transformation*.

One-dimensional continuous wavelet transformation. The continuous wavelet transform is usually expressed by the following equation (Chen *et al.*, 2007; Kisi, 2008; Partal, 2009; Misiti *et al.*, 2010).

$$W_f(a,\tau) = < f(t), \psi(t) > = |a|^{-1/2} \int_{-\infty}^{+\infty} f(t)\psi \times \left(\frac{t-\tau}{a}\right) dt \quad a \in R,\ \tau \in R, a \neq 0 \quad (1)$$

$$\psi_{a,\tau}(t) = |a|^{-1/2}\psi\left(\frac{t-\tau}{a}\right) \quad (2)$$

where $W_f(a, \tau)$ are the wavelet coefficients, t is time interval, $f(t)$ presents the input signal, $\psi(t)$ is a base wavelet function also called the mother wavelet, * corresponds to the complex conjugate of $\psi(t)$, a is a scaling factor stretching or compressing the mother wavelet to the frequency of the signal, and τ is a translating factor shifting the mother wavelet to the time domain of the signal. Equation (1) expresses that the wavelet transformation is the decomposition of $f(t)$ under different resolution levels (scales). The essence of wavelet transformation is to filter a wave for $f(t)$ with different filters (Wang and Ding, 2003).

The property of the mother wavelet $\psi(t)$ is that the set of its continuous variants $[\psi_{a,\tau}(t)]$ forms an orthogonal basis in a two-dimensional space $L^2(R)$, and it has finite energy and fast decay. Its Fourier transformation ($\hat{\psi}(\omega)$

has to satisfy admissibility conditions.

$$C_\psi = \int_R \frac{|\hat{\psi}(\omega)|^2}{|\omega|} d\omega < \infty, \quad \int_R \psi(t)dt = \hat{\psi}(0) = 0 \tag{3}$$

Under those conditions, the input signal can be reconstructed using an inverse wavelet transform expressed by

$$f(t) = \frac{1}{C_\psi} \int_0^{+\infty} \frac{da}{a^2} \int_{-\infty}^{+\infty} W_f(a, \tau) a^{-1/2} \psi\left(\frac{t-\tau}{a}\right) d\tau \tag{4}$$

One-dimensional discrete wavelet transformation. Continuous wavelet transformation calculates wavelet coefficients at every possible scale, which is a time-consuming process producing excessive (often redundant) amounts of data. When the scaling factor a and shifting factor τ of the basic wavelet function ($\psi_{a,\tau}(t)$) are limited to only discrete values, the analysis is more efficient and accuracy is retained. This process is called *discrete wavelet transformation*, and has been recently used as a computational technique for extracting information about non-stationary signals (Daubechies, 1990; Kim *et al*., 2006). It can be expressed by

$$\text{Let } a = a_0^j, \ \tau = k a_0^j \tau_0, \ a_0 > 0, \ \tau_0 \in R, \ \forall j, \quad k = 0, 1, 2, 3, \ldots, m \in Z \tag{5}$$

where ($\psi_{a,\tau}(t)$) can be written as:

$$\psi_{j,k}(t) = a_0^{-j/2} \psi[a_0^{-j}(t - k a_0^j \tau_0)] = a_0^{-j/2} \psi(a_0^{-j} t - k\tau_0) \tag{6}$$

Therefore, the discrete wavelet transformation can correspondingly be expressed by

$$W_f(j,k) = a_0^{-j/2} \int_{-\infty}^{+\infty} f(t)\psi \times (a_0^{-j} t - k\tau_0) dt \tag{7}$$

The simplest and most efficient practice is that scales and position are selected based on the powers of two logarithms, called *dyadic* scales and positions (Mallat, 1989). That is, let $a_0 = 2$, $\tau_0 = 1$, then the discrete wavelet transformation becomes a binary one.

$$W_f(j,k) = 2^{-j/2} \int_{-\infty}^{+\infty} f(t)\psi \times (2^{-j} t - k) dt \tag{8}$$

$W_f\ (a, \tau)$ and $W_f\ (j, k)$ can reflect the characters of the original time series in frequency (a or j) and time domains (τ or k). When a or j is small, the frequency resolution is very low, but the time domain is very high. When a or j become large, the frequency resolution is high, but the time domain is low.

As a discrete time series $f(t)$, in which $f(t)$ occurs at discrete integer time steps t, the dyadic discrete wavelet transformation can be written as

$$W_f(j,k) = \sum_{j,k \in z} f(t) 2^{-j/2} \psi(2^{-j} t - k) \tag{9}$$

The input signal can be reconstructed using the equation

$$f(t) = \sum_{j,k \in z} W_f(j,k) \psi_{i,k}(t) \tag{10}$$

In this equation, wavelet coefficients $W_f(j,k)$ are divided into an approximation (or low frequency) coefficient (cA_n) at level n through a low pass filter $l(\psi_{i,k}(t))$, and detail (or high frequency) coefficients (cD1, cD2, cD3,..., cD_n) at different levels 1, 2, ..., n through a high-pass filter $h(\psi_{i,k}(t))$. cA_n provides background information on the original signal, while cD1, cD2, cD3, ..., cD_n contains the detail information on the original signal such as period, break and jump. Then the original signal can be expressed as

$$f(t) = cA_n l(\psi_{i,k}(t)) + \sum_{n=1} cD_n h(\psi_{i,k}(t)) \tag{11}$$

or simplified to the form

$$f(t) = A_n(t) + \sum D_n(t) \tag{12}$$

where $A_n(t)$ is the approximation of the original signal at level n, and $D_n(t)$ are the details of the original signal at different levels (1, 2, 3,.., n).

Daubechies wavelets are one of the most commonly used set of discrete wavelet transforms, which was formulated by the Belgian mathematician Ingrid Daubechies in 1988 (Mohamed and Atta, 2010). Daubechies wavelet family is usually written as 'dbN', where db is the 'surname', and N the order of the wavelet. After Ingrid Daubechies, the Daubechies family goes from 1 to 20. Daubechies No 1 (db1), the first and simplest one, represents the same Haar wavelet, whereas the complex one, for example Daubechies No 9 (db9), has about the same shape as Daubechies Wavelet No 5 (db5), but with many more oscillations in the tail. In this study, we used db5, one of the most commonly used mother wavelets (Goring, 2006; Abibullaev *et al*., 2009; Hu *et al*., 2010; Sasi *et al*., 2010; Badiei and Mohammadi, 2011), to decompose the series. Figure 2 displays the wavelet functions (psi) of db1, db5 and db9 of the Daubechies wavelet family numbers.

Artificial neural networks (ANN)

Basic concept. An ANN is defined as a structure comprised of a number of simple interconnected operating elements called neurons (units, cells or nodes), as inspired by the biological nervous system (Beale *et al*., 2010). A neural network can perform a particular function mapping between inputs and outputs, by adjusting the values of the connections (weights) between neurons. The feed-forward multilayer perceptron (FFML) network is the most widely used type of ANN in hydrological modelling (Wang *et al*., 2006). A typical FFML network includes a number of neurons connected together in layers. Its first and final layers are called the input and output layers, respectively, and all other layers are hidden layers. There are various ANN algorithms including Hebbian, Deklta,

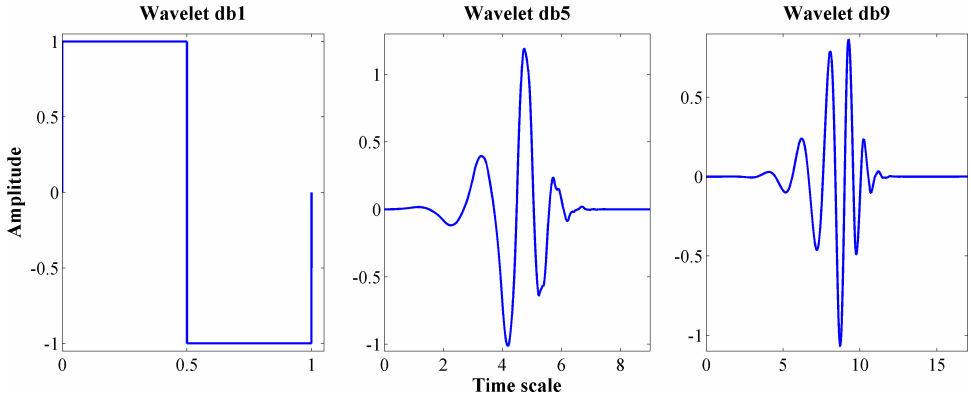

Figure 2. Wavelet function psi (ψ) of Daubechies Wavelet No. 1 (db1), Daubechies Wavelet No. 5 (db5) and Daubechies Wavelet No. 9 (db9) of the Daubechies wavelet family

Kohonen, and BP (Ertay and Çekyay, 2006; Feng and Hong, 2008). In recent years, the error back-propagation algorithm (BP) has become the most popular one in the water resources literature (Cigizoğlu, 2004). The main working process of the BP network are the input, hidden and output units which are completely connected in a feed-forward way, and an input is propagated forward to output units through hidden layers. When the forwarding reaches the output layer, comparison with the target (observations) allows calculation of the mean sum of squares of the network error (*MSE*) using Equation (13).

$$MSE = \frac{1}{N}\sum_{h=1}^{N}\sum_{n=1}^{m}(T_{hn} - O_{hn})^2 \tag{13}$$

where N is the training data number, n is the output node number, T_{hn} is the target value, (i.e. observed data), and O_{hn} is the output of the neural network and, h is the pattern of the network. Each error signal can be back-propagated from the output to input layer to adjust the weight. These processes can continue iteratively until the network output matches the target within a specified level.

Model process. The important decisions necessary to build a neural network model include neural network type, network structure, methods of pre- and post-processing of input/output data, choice of training algorithms as well as training stop criteria (Wang *et al.*, 2006). Determining the numbers of hidden layers and neurons (i.e. network structure) is the most important task in building the network. While a trial and error procedure to determine the number of neurons is still the most commonly used methods some alternative algorithms have also been proposed. In this study, Feed-forward multilayer perceptron (FFML) ANN was used as the model structure, and a back-propagation algorithm (BP) was used for the network training. The Levenberg-Marquardt (LM) algorithm was employed for BP training due to its advantages of faster convergence compared to the traditional widely used gradient descent algorithm. Preprocessing of the dataset for network training, called data normalszation, is also an important step in ANN modelling, in which all the data are transformed into a small intervals usually [−1, 1] or [0,1] depending on the transfer (activation) function of neurons. This process is important to ensure that all variables receive equal attention and that the efficiency of the training network is improved (Dawson and Wilby, 2001; Wang *et al.*, 2006). Various methods can be used for data normalisation (Wang *et al.*, 2006). In this study, the data series of river discharge (Q_t) was normalized into the range [0,1] using Equation (14).

$$Q_t^{'} = \frac{Q_t - Q_{\min}}{Q_{\max} - Q_{\min}} \tag{14}$$

where $Q_{\max}$ and $Q_{\min}$ are the maximum and minimum values from the river discharge series, respectively.

The transfer function associated to the neurons in the hidden layer was a tangent sigmoid function (Equation (15)), while to output neurons, a logarithm sigmoid function (Equation (16)) was employed considering the output value of this function was in the range of [0,1].

$$tansig(n) = \frac{2}{(1+e^{-2n})} - 1 \tag{15}$$

$$logsig(n) = \frac{1}{(1+e^{-n})} \tag{16}$$

Improving generalisation. Overfitting is one of the common problems that usually occurs during neural network training, in which the error on the training set is driven to a very small value, but when new data is presented to the network the error is large, indicating that the network is able to memorize the training examples, but will not generalize well to new situations. Two methods are usually used for improving the network generalisation: early-stopping and regularisation. For the early-stopping method, data is normally divided into three subsets: one for training, one for validation and one for testing. This method was applied to stop the training process when the optimum result occurred, i.e. the performance of network failed to improve. Another method for improving generalisation is called regularisation, which involves modifying the performance function, normally the mean sum of squares of the network error (*MSE*)

(Equation (13)). The generalisation will be improved if the performance function (*MSE*) is modified by adding a term consisting of the sum of squares of the network weights and biases (Beale *et al*., 2010). This regularisation method is usually expressed by

$$msereg = \gamma mse + (1-\gamma)msw \tag{17}$$

$$msw = \frac{1}{n}\sum_{i=1}^{n} w_j^2 \tag{18}$$

where *msereg* means the regularisation *MSE*, γ is the performance ratio, *w*, *msw* is the weights of network *MSE*, and their average.

The problem with regularisation is that it is difficult to determine the optimum value for the performance ratio parameter. The network will overfit the training data if this parameter is too large. In contract, the network does not adequately fit the training data if the ratio is too small (Beale *et al*., 2010). One approach to overcome this problem is the Bayesian regularisation (BR) (MacKay 1992), which enables the determination of the optimal regularisation parameters in an automated fashion.

Model evaluation

Since the developed models are applied for management and planning appropriate model evaluation methods are essential. Many measures can be applied to assess the accuracy of developed models. In this study, besides the mean squared error (*MSE*) expressed by Equation (13), the following error statistics were also used to evaluate the models.

Correlation coefficient (R)

$$R = \frac{\sum_{t=1}^{n}(Q_t - \overline{Q}_t)(\hat{Q}_t - \overline{\hat{Q}}_t)}{\sqrt{\sum_{t=1}^{n}(Q_t - \overline{Q}_t)^2 \sum_{t=1}^{n}(\hat{Q}_t - \overline{\hat{Q}}_t)^2}} \tag{19}$$

where t is time unit, $\hat{Q}_t$ is modelled river discharge, Q_t is observed discharge, $\overline{Q}_t$ and $\overline{\hat{Q}}_t$ are the mean ones.

Mean absolute error (MAE)

$$MAE = \frac{1}{n}\sum_{t=1}^{n}\left|\hat{Q}_t - Q_t\right| \tag{20}$$

Mean absolute relative error (MARE)

Mean absolute relative errors (*MARE*) are calculated based on Equation (21).

$$MARE = \frac{1}{n}\sum_{t=1}^{n}\left|\frac{\hat{Q}_t - Q_t}{Q_t}\right| \tag{21}$$

Absolute relative errors of maximum (RE_{mxd}) and *minimum discharge* (RE_{mnd}) are used to measure the model ability to predict the peak and lowest discharge.

$$RE_{\mathrm{mxd}} = \left|\frac{\hat{Q}_{\max} - Q_{\max}}{Q_{\max}}\right| \tag{22}$$

$$RE_{\mathrm{mnd}} = \left|\frac{\hat{Q}_{\min} - Q_{\min}}{Q_{\min}}\right| \tag{23}$$

EXPERIMENTAL AREA AND MATERIALS

The dataset used in this modelling experiment was a time series of average monthly river discharge during the period 1956–1995, which was collected from the Weijiabao gauging hydrological station in the Weihe River Basin. The Weihe river originates from the Wushu Mountain in Weiyuan County of the Gansu province in west-central China, and runs across the Shaanxi province and joins the Yellow River at Tonguan in Shaanxi. The river is 818 km long with a watershed area of 135 000 km^2, and is the largest tributary of the Yellow River. It is called the Mother River of the Guanzhong region in the Shaanxi province, and plays a large role in developing West China and maintaining ecosystem health of the Yellow River. However, since the 1980s, the river has suffered many problems, including greatly reduced annual river runoff, geologic hazards caused by water projects, high concentrations of sediment and consequential heavy flooding, as well as heavily water pollution (Song *et al*., 2007). Since the late 1990s, many parts of the river have lost ecosystem functionality and such problems have restricted the sustainable development of the region (Zhao, 2003; Wang *et al*., 2004). In this context, it is of significant interest to have an advanced technique to model the river's discharge with improved prediction accuracy. Since the Weihe River becomes flat midstream, where high sedimentation occurs, the Weijiabo gauging station is one of more important midstream monitoring stations along the Weihe River (Figure 3).

The average monthly river discharge of the Weijiabo station was 103·4 m^3/s with a maximum discharge of 705 m^3/s in October 1975 and a minimum discharge of 0·98 m^3/s in February 1969. Since records were kept, discharge has decreased from an average monthly discharge of 140·4m^3/s during the first 10 year period (1956–1965), to 94·2 m^3/s during second 10 year period (1966–1975), to 107·8 m^3/s during the third 10 year period (1976–1985), and finally to 89·7 m^3/s during the last 10 years (Figure 4(a)). The river discharge also has clear seasonal character, and the highest discharge occurs from July to October and the lowest discharge from December to March (Figure 4(b)).

RESULTS AND DISCUSSION

The 480-month river discharge data were divided into two parts, data of the first 432 months (36 years) for network

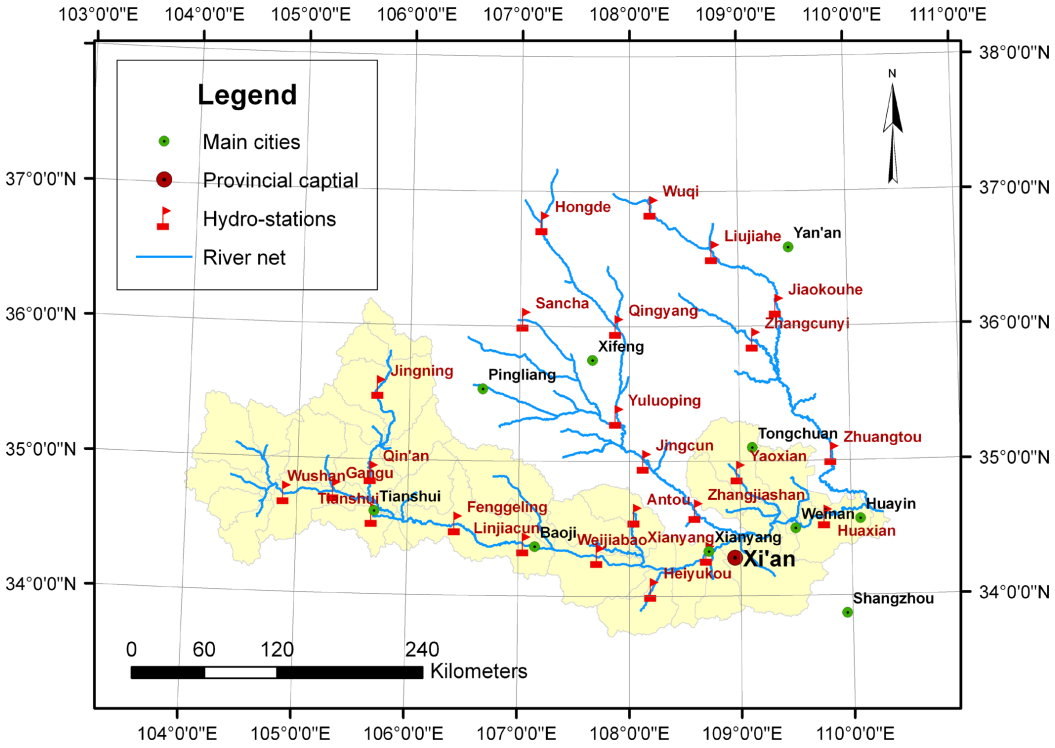

Figure 3. A sketch of the water system of the Weihe River Basin, China

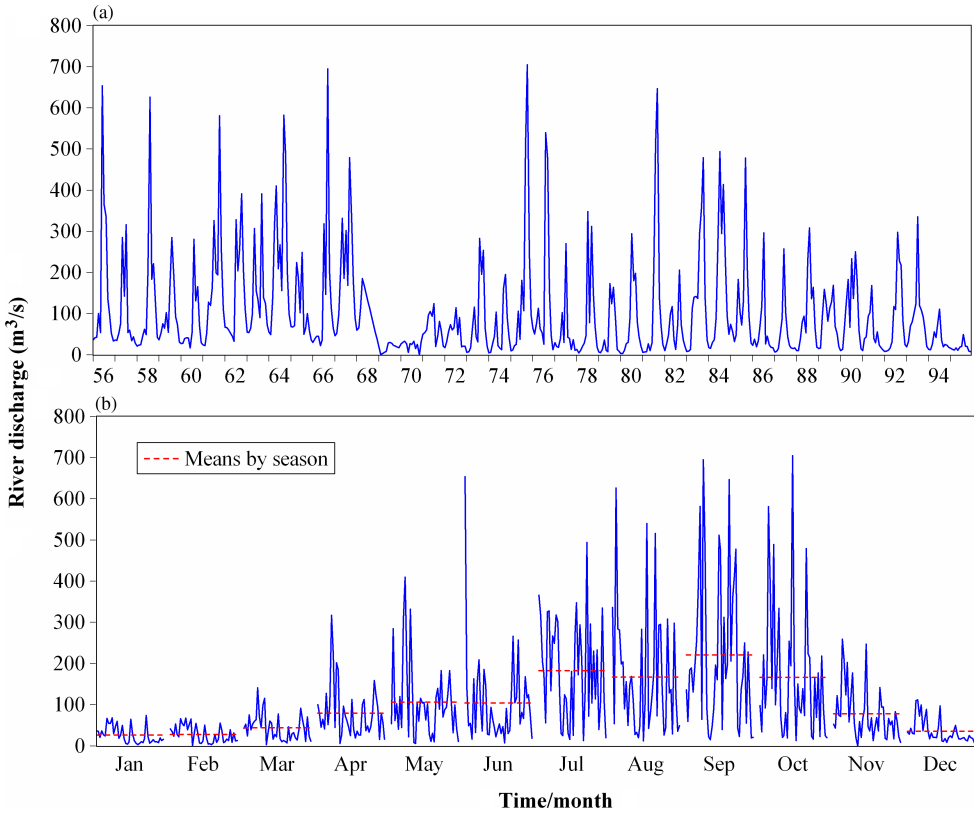

Figure 4. Monthly mean discharge at the Weijiabao gauging station in the Weihe River, China

training and the remaining 48-month data (4 year) for model prediction testing. Since water decision makers would make an annual plan at the beginning of a year based on the situation in the past year(s) it was essential to establish a model capable of predicting 12-month river discharge for the new year using the 12-month data in the previous year. In this context, we reconstructed the data set using the previous 12-month data as inputs and the following 12-month data as the targets for network training (Table I). The results of wavelet decomposition,

Table I. Data division for model training and testing. P = Input vectors; T = Target vectors; S = Signal, i.e. observed river discharge dataset

	Input	Target
Training	P1 = [S1, S2, S3, ..., S12]; P2 = [S13, S14, S15, ..., S24]; P3 = [S25, S26, S27, ..., S36]; ..... P35 = [S409, S410, S411, ..., S420]	T1 = [S13, S14, S15, ..., S24]; T2 = [S25, S26, S27, ..., S36]; T3 = [S37, S38, S39, ..., S48]; ..... T35 = [S421, S422, S423, ..., S432];
Testing	P36 = [S421, S422, S423, ..., S432]; P37 = [S433, S434, S435, ..., S444]; P38 = [S445, S446, S447, ..., S456]; P39 = [S457, S458, S459, ..., S468];	T36 = [S433, S434, S435, ..., S444]; T37 = [S445, S446, S447, ..., S456]; T38 = [S457, S458, S459, ..., S468]; T39 = [S469, S470, S471, ..., S480];

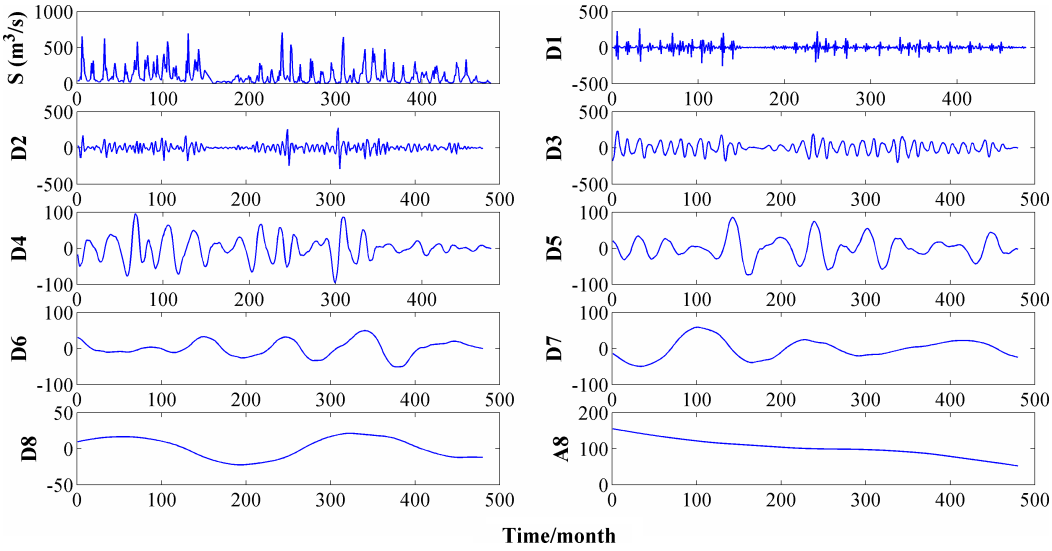

Figure 5. Wavelet decomposition of the observed monthly mean river discharge series (i.e. original signal (S)) at Weijiabao gauging station using Daubechies Wavelet No. 5 (db5) with 8 levels, D1, D2, D3, ..., D8 are the details and A8 is the approximation with wavelet coefficients as the vertical axis

model evaluation for training and prediction performance, as well as results of the improved models are analysed in the following sections.

Wavelet decomposition

The time series dataset of river discharge, i.e. the signal (S) in the experimental study area was decomposed into various details (Ds) at different resolution levels using discrete wavelet transformation (DWT). The new decomposed sub-series of wavelet transformation coefficients represent variations of the original times series on different periods. A decomposition at eight resolution levels (2-4-8-16-32-64-128-256) using db5 wavelet is illustrated as an example in Figure 5, which shows the time series of 2-month mode (D1), 4-month mode (D2), 8-month mode (D3), 16-month mode (D4), 32-month mode (D5), 64-month model (D6), 123-month mode (D7), 256-month mode (D8), and the approximation mode (A8). Each mode displays the variations of the original river discharge time series over a different period. For instance, the approximation mode (A8) illustrated that river discharge had a clear decreasing trend, and the 256-month mode (D8) showed that there were two major cycles during the estimated period.

Results of model training

A three layer BP network structure (p, m, n) was used for training both WNN and ANN models, in which p is the input numbers, m is the neurons in the hidden layers and n is the output numbers. For network training, the maximum training epoch, performance goal (MSE) and learning rate were set at 1000, 0·001 and 0·01, respectively. The training process started using 1 neuron (i.e. $m = 1$), and the neuron numbers were increased progressively until the performance goal ($MSE = 0{\cdot}001$) was met. The results of model performances of training are presented in Table II. In general, the results showed that WNN hybrid models of db5L2 to db5L8 met the performance goal quickly after 6 neurons were used in the hidden layer, and the db5L1 model also reached the performance goal after 7 hidden neurons were used. However, the training results reveal that the ANN model alone did not meet the goal until 8 neurons were used in the hidden layers with 451 iterations. The training results of the ANN model with 8 neurons (ANN3) showed that this model had a very good performance for

Table II. Evaluation of the WNN and ANN model training performances

Model	Model Structure[a]	Epochs	R[b]	MSE[c] $((m^3/s))^2$	MAE[d] (m^3/s)	$MARE$[e]
db5L1-1[f]	(24,06,12)	1000	0·908	0·0058	0·035	0·503
db5L1-2[f]	(24,07,12)	111	0·984	0·0010	0·022	0·396
db5L2[f]	(36,06,12)	217	0·983	0·0010	0·023	0·488
db5L3[f]	(48,06,12)	59	0·984	0·0010	0·022	0·477
db5L4[f]	(60,06,12)	35	0·983	0·0010	0·022	0·477
db5L5[f]	(72,06,12)	61	0·986	0·0009	0·021	0·377
db5L6[f]	(84,06,12)	90	0·984	0·0010	0·022	0·430
db5L7[f]	(96,06,12)	23	0·984	0·0010	0·024	0·497
db5L8[f]	(108,06,12)	15	0·986	0·0009	0·023	0·451
ANN1[g]	(12,06,12)	1000	0·846	0·0092	0·038	0·452
ANN2[g]	(12,07,12)	1000	0·970	0·0014	0·026	0·527
ANN3[g]	(12,08,12)	451	0·987	0·0008	0·020	0·407

[a] Structure (m, n, p): The first number (m), the second one (n) and the third one (p) signify elements of input vectors, neurons in hidden layers and the output numbers, respectively.
[b] R = Correlation coefficient.
[c] MSE = Mean squared error.
[d] MAE = Mean absolute error.
[e] $MARE$ = Mean absolute relative error.
[f] db5LN = Daubechies Wavelet 5 at N (1, 2, 3, . . .,8) level.
[g] ANN1-3 = three models of ANN alone.

fitting historical data due to high correlation coefficient ($R = 0{\cdot}98$) and smaller errors ($MSE = 0{\cdot}0008$ and $MAE = 0{\cdot}020$ and $MARE = 0{\cdot}407$). Comparison of the training results of WNN with ANN models suggested that db5L1-2, db5L2 to db5L8 and ANN3 all had very good simulated performances in terms of R, *MSE, MAE* as well as *MARE*.

Model testing evaluation

The general testing evaluation results for model prediction ability of WNN db5L1-2, db5L2 to db5L8, and ANN3 using the observed discharge data from the last 48 months are summarized in Table III. In terms of correlation coefficients (R values), the WNN hybrid models of db5L2, db5L3, db5L4 and dbL6, and ANN3 exhibited a good correlation between observed data and simulated results. However, models of db5L2 and dbL6 had weaker prediction ability due to larger *MSE, MAE* and *MARE*. WNN hybrid models of db5L3, db5L4 and db5L8 had much better prediction results than any other WNN models and the ANN3 model in terms of *MSE, MAE* and *MARE*. For further comparison of these better models, the maximum absolute relative error (*MxARE*), minimum absolute relative error (*MnARE*), and absolute relative error of highest river discharge (RE_{mxd}) and lowest river discharge (RE_{mnd}), called *extreme error indices* were considered. The largest *MxARE* (28·333), RE_{mxd} (1·034) and RE_{mnd} (0·877) of ANN3 proved that this model had weaker prediction ability than the WNN hybrid models, especially for the highest and lowest discharge events (Table IV).

The visualized comparison of the simulated and observed discharge series between the selected three WNN hybrid models and the ANN3 model is illustrated in Figure 6. This figure clearly illustrates that compared to the three WNN models, even though ANN had a very good performance for fitting historical data during the training, ANN3 greatly overestimated the highest discharge and underestimated the lowest discharge for the 48-month testing period.

Model generalisation improvement

The previous analyses had shown that WNN hybrid models of db5L3, db5L4 and db5L8 had the best performance compared to all other WNN models and to the ANN models alone. In the meantime, the error indexes also showed that the error on the training set was much smaller while the error on the new (testing) data was larger. This revealed that the network was able to memorize the training examples, but had not learnt to generalize the new situations. In this sense, overfitting problems occurred during network training of the WNN models. For the ANN3 model, overfitting was a much more serious issue due to the much smaller errors during training but exhibiting the largest errors during testing (Figure 6). While two methods, early-stopping and regularisation, could be used to solve the overfitting problem as both these methods can ensure network generalisation when applied properly; for a small dataset, however, Bayesian regularisation provides a much better generalisation performance than early-stopping. This is because Bayesian regularisation (BR) does not need a validation dataset separated from the training dataset, and consequently can utilize all available data (Beale *et al*., 2010). In Matlab, BR was implemented in the training function *trainbr*, via the Matlab Neural Network Toolbox User's Guide (Beale *et al*., 2010).

The evaluation results of the three selected WNN models (db5L3, db5L4 and db5L8) and ANN3 model after employing Bayesian regularisation are displayed in Tables V and VI. The evaluation indices for the model testing showed that the prediction abilities of the three WNN models and the ANN3 model had been significantly improved, while retaining the good fitting

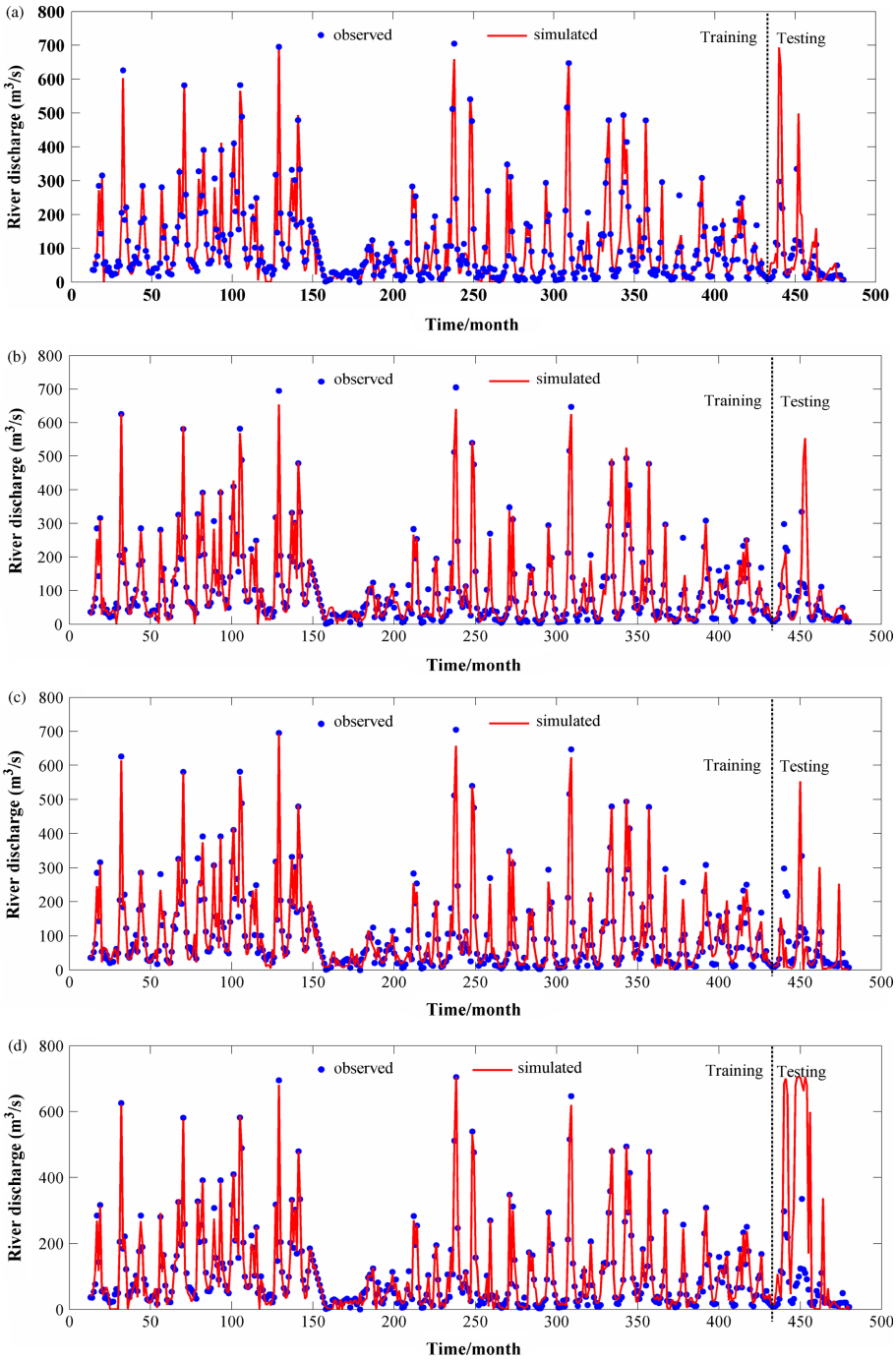

Figure 6. Comparison of simulated results and observed data of: (a) WNN model db5L3, (b) WNN model db5L4, (c) WNN model db5L8, and (d) model ANN3

abilities for the historical data as evidenced by the training evaluation indices (Table V). The lower extreme error indices also confirmed that the four improved models had better prediction performance for the highest and lowest river discharge than the unimproved models. Comparison of the three WNN hybrid models with ANN3 model, via model evaluation indices, suggested that the three hybrid WNN models significantly improved the general estimation of intermediate and moderate discharge (Table V), while comparison of extreme error indices indicated that the three hybrid models also had abilities to accurately forecast the highest and lowest river discharges (Table VI). The prediction abilities of the three hybrid models were also superior to ANN3 as indicated by direct comparison of simulated results and observed data (Figure 7).

Comparing the evaluation indices of the four models, in terms of smallest errors ($MSE = 0{\cdot}004$, $MAE = 0{\cdot}041$, $MnARE = 0{\cdot}002$, $RE_{\mathrm{mxd}} = 0{\cdot}068$) and highest R value ($R = 0{\cdot}82$), revealed that the WNN model of db5L4 had the best performance to estimate general river and peak discharge. The good performance of the best model (WNN db5L4) and its comparison to ANN3 model are displayed using regression plots (Figure 8), in which

Table III. Statistic evaluation measures of the WNN and ANN model prediction performances

Model	Model Structure[a]	R[b]	MSE[c] $((m^3/s))^2$	MAE[d] (m^3/s)	$MARE$[e]
db5L1-2[f]	(24,07,12)	0·484	0·138	0·221	4·190
db5L2[f]	(36,06,12)	0·664	0·075	0·150	2·148
db5L3[f]	(48,06,12)	0·713	0·025	0·081	0·980
db5L4[f]	(60,06,12)	0·512	0·020	0·066	0·747
db5L5[f]	(72,06,12)	0·488	0·019	0·077	1·015
db5L6[f]	(84,06,12)	0·551	0·162	0·246	3·763
db5L7[f]	(96,06,12)	0·085	0·036	0·118	1·788
db5L8[f]	(108,06,12)	0·295	0·022	0·081	0·988
ANN3[g]	(12,08,12)	0·695	0·157	0·243	2·916

[a] Structure (m, n, p): The first number (m), the second one (n) and the third one (p) signify elements of input vectors, neurons in hidden layers and the output numbers, respectively.
[b] R = Correlation coefficient.
[c] MSE = Mean squared error.
[d] MAE = Mean absolute error.
[e] $MARE$ = Mean absolute relative error.
[f] db5LN = Daubechies Wavelet 5 at N (1, 2, 3, . . .,8) level.
[g] ANN3 = model 3 of ANN alone.

Table IV. Extreme statistic errors for evaluation of model prediction

Model	$MxARE$[a]	$MnARE$[b]	RE_{mxd}[c]	RE_{mnd}[d]
db5L2-2[e]	20·209	0·018	0·050	0·195
db5L3[e]	4·0781	0·056	0·750	0·263
db5L4[e]	4·0365	0·053	0·443	0·435
db5L6[e]	34·951	0·149	0·389	0·649
db5L8[e]	12·377	0·007	0·457	0·733
ANN3[f]	28·333	0·003	1·034	0·877

[a] $MxARE$ = Maximum absolute relative error.
[b] $MnARE$ = Minimum absolute relative error.
[c] RE_{mxd} = Absolute relative errors of maximum river discharge.
[d] RE_{mnd} = Absolute relative errors of minimum discharge.
[e] db5LN = Daubechies Wavelet 5 at N (2, 3, 4, 6, 8) level.
[f] ANN3 = model 3 of ANN alone.

the network outputs are plotted with respect to targets. For a perfect fit, the data should fall along a 45° line, where the network outputs are equal to the targets. For WNN db5L4, R values of 0·978 and 0·820 in training and testing indicated that the fit was reasonably good. In contrast, ANN3 exhibited a less precise performance for river discharge prediction in terms of a lower R value ($R = 0{\cdot}667$) in testing even though the higher R value of 0·974 in training indicated that ANN3 was also able to fit historical data well. A detailed comparison of the results of 12-month forecasts, 48 months ahead (i.e. 4-year period) of the WNN db5L4 and ANN3 models and their evaluations are displayed in Table VII. In general, these results revealed that the 48-month-ahead prediction of WNN db5L4 was more accurate than that of the ANN3 model based on both maximum, minimum and mean absolute errors (AEs), and the AEs of peak and lowest discharge, though ANNs had a comparatively smaller maximum AE (128·1) in the first 12-month-ahead forecast (annual) and smaller AEs for peak and lowest discharge (19·5 and 6·1) in the third annual forecasting period than WNN db5L4.

Table VI. Extreme statistic errors for model predictions evaluation after generalisation improvement

Model	Model structure[a]	$MxARE$[b]	$MnARE$[c]	RE_{mxd}[d]	RE_{mnd}[e]
Db5L3[f]	(48,06,12)	2·598	0·022	0·259	0·931
Db5L4[f]	(60,06,12)	4·369	0·002	0·068	0·809
Db5L8[f]	(108,06,12)	2·917	0·037	0·653	0·266
ANN3[g]	(12,08,12)	7·275	0·069	0·716	1·942

[a] Structure (m, n, p): The first number (m), the second one (n) and the third one (p) signify elements of input vectors, neurons in hidden layers and the output numbers, respectively.
[b] $MxARE$ = Maximum absolute relative error.
[c] $MnARE$ = Minimum absolute relative error.
[d] RE_{mxd} = Absolute relative errors of maximum river discharge.
[e] RE_{mnd} = Absolute relative errors of minimum river discharge.
[f] db5LN = Daubechies Wavelet 5 at N (3, 4, 8) level.
[g] ANN3 = the third model of ANN alone.

CONCLUSIONS

This study improved the model prediction performance of river discharge time series using a Wavelet Neural network (WNN) hybrid model approach. The WNN hybrid models were successfully applied to the Weijiabao

Table V. Evaluation of the WNN and ANN model training and testing performances

Model	Model structure[a]	Training				Testing			
		R[b]	MSE $((m^3/s))^2$[c]	MAE (m^3/s)[d]	$MARE$[e]	R[b]	MSE $((m^3/s))^2$[c]	MAE (m^3/s)[d]	$MARE$[e]
db5L3[f]	(48,06,12)	0·979	0·001	0·025	0·559	0·783	0·013	0·053	0·735
db5L4[f]	(60,06,12)	0·978	0·001	0·026	0·550	0·820	0·004	0·041	0·684
db5L8[f]	(108,06,12)	0·989	0·001	0·020	0·481	0·661	0·006	0·048	0·615
ANN3[g]	(12,06,12)	0·974	0·002	0·029	0·789	0·667	0·009	0·060	1·216

[a.] Structure (m, n, p): The first number (m), the second one (n) and the third one (p) signify elements of input vectors, neurons in hidden layers and the output numbers, respectively.
[b.] R = Correlation coefficient.
[c.] MSE = Mean squared error.
[d.] MAE = Mean absolute error.
[e.] $MARE$ = Mean absolute relative error.
[f.] db5LN = Daubechies Wavelet 5 at N (3, 4 and 8) level.
[g.] ANN3 = the third model of ANN alone.

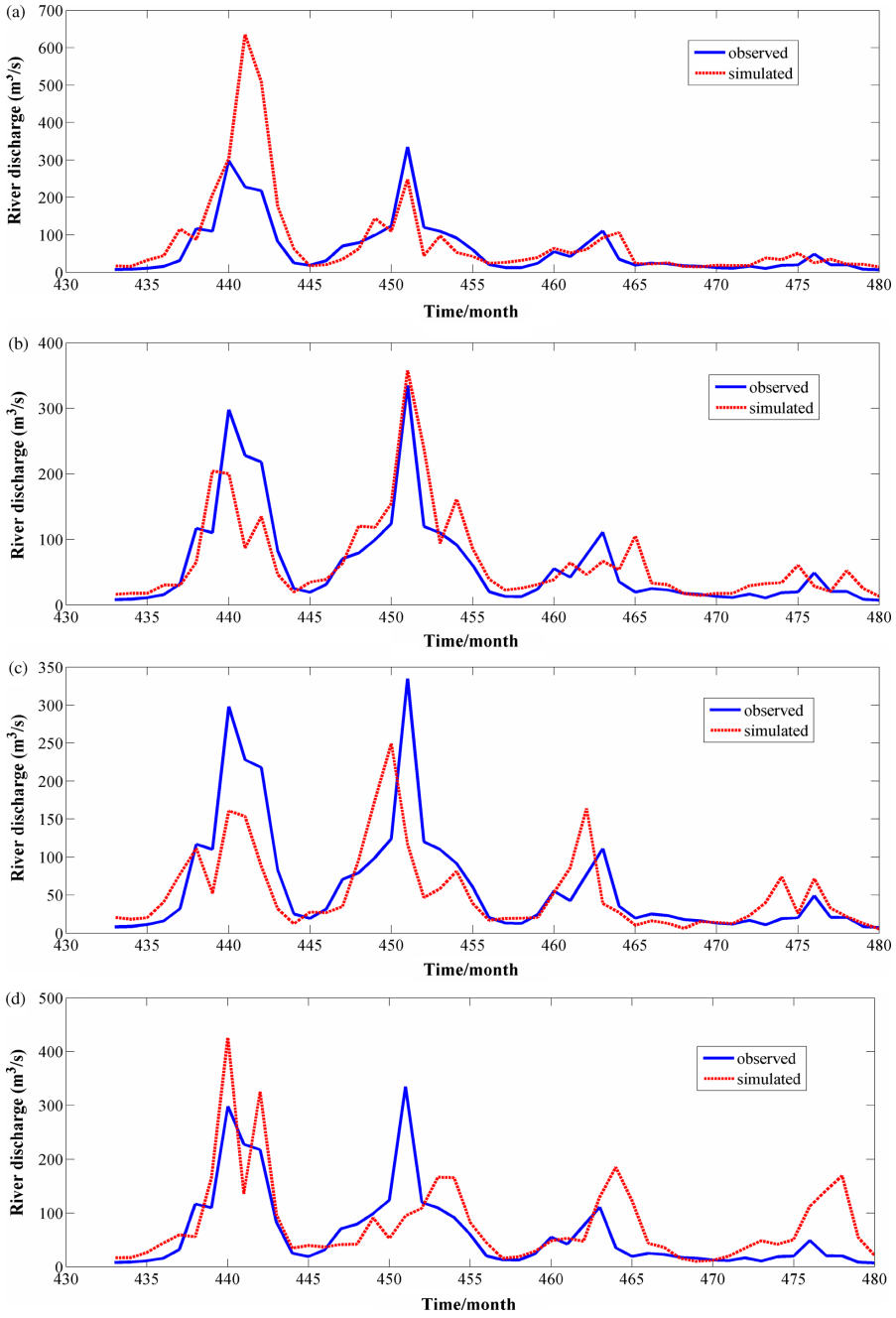

Figure 7. Comparison of forecasting results and observed data after employing Bayesian regularisation in network training: (a) WNN model db5L3, (b) WNN model db5L4, (c) WNN model db5L8, and (d) model ANN3

hydrological gauging station data in the Weihe River, China to predict river discharge time series 48-month ahead.

The discharge time series data was initially normalized into range of [0, 1] at the beginning of the model process, mainly to enable fast training convergence of the ANN model. The pre-processed time series data was decomposed into 8 different sub-series at various levels using DWT of the wavelet function of Daubenchies 5 (db5). The sub-series at these 8 different levels were input into the designed network, which formed 8 different WNN hybrid models. The results from the 8 WNN models were compared with those from the ANN models alone in two steps. In the first step, a performance goal ($MSE = 0{\cdot}001$) was included in the training process. Three hybrid models and one ANN model with good performances were selected through evaluating of the models' fitting and forecasting abilities. In the second step, the selected models were improved by employing Bayesian regularisation (BR) in the training function.

Comparison of the results showed that while ANN might meet the training goals provided sufficient neurons were included in the hidden layers, in general, ANN alone failed to accurately predict the peak and lowest river discharge. In contrast, WNN hybrid models displayed much better performance both in the training and testing periods than ANN alone. WNN models significantly improved

Table VII. Comparison of model results of models WNN db5L4 and ANN3 ($\times m^3/s$)

Time/month	First annual period					Second annual period					Third annual period					Fourth annual period				
	O^a	WNN db5L4		ANN3		O^a	WNN db5L4		ANN3		O^a	WNN db5L4		ANN3		O^a	WNN db5L4		ANN3	
		S^b	AE^c	S^b	AE^c		S^b	AE^c	S^b	AE^c		S^b	AE^c	S^b	AE^c		S^b	AE^c	S^b	AE^c
1	8·2	16·2	7·9	17·0	8·8	19·3	34·8	15·5	39·8	20·5	13·0	22·7	9·7	15·8	2·8	16·2	14·8	1·4	10·1	6·1
2	8·9	17·9	9·0	17·2	8·3	31·7	39·1	7·4	36·9	5·2	12·7	25·6	12·9	18·8	6·1	12·9	17·6	4·7	12·0	0·9
3	11·2	17·9	6·7	26·6	15·4	70·6	63·3	7·3	41·7	28·9	24·4	31·3	6·9	29·0	4·6	11·6	17·6	6·0	20·1	8·5
4	15·9	30·5	14·6	44·1	28·2	79·4	120·3	40·9	42·2	37·2	55·4	38·7	16·7	48·3	7·1	16·7	29·3	12·6	34·9	18·2
5	32·2	30·0	2·2	59·5	27·3	99·5	118·5	19·0	91·3	8·2	42·4	64·5	22·1	53·3	10·9	10·7	32·5	21·8	48·7	38·0
6	117·0	65·6	51·4	55·8	61·2	124·0	154·8	30·8	52·5	71·5	76·7	46·2	30·5	47·7	29·0	18·9	34·2	15·3	41·8	22·9
7	110·0	204·2	94·2	166·1	56·1	335·0	357·8	22·8	95·1	239·9	111·0	66·8	44·2	130·5	19·5	20·0	60·7	40·7	50·8	30·8
8	298·0	200·0	98·0	426·1	128·1	120·0	238·8	118·8	109·0	11·0	35·4	53·7	18·3	185·5	150·1	49·1	28·6	20·5	112·7	63·6
9	228·0	86·9	141·1	135·8	92·2	110·0	93·9	16·1	166·8	56·8	19·6	105·2	85·6	122·4	102·8	20·6	20·6	0·0	142·1	121·5
10	218·0	135·0	83·0	325·3	107·3	91·8	161·5	69·7	166·0	74·2	25·0	33·2	8·2	43·7	18·7	20·5	52·3	31·8	169·6	149·1
11	83·2	46·9	36·3	96·7	13·5	60·2	85·7	25·5	82·6	22·4	23·0	30·9	7·9	36·0	13·0	8·7	25·6	16·9	54·4	45·7
12	25·2	19·6	5·6	34·7	9·5	20·4	39·5	19·1	45·0	24·6	17·7	17·6	0·1	15·7	2·0	7·3	13·2	5·9	21·5	14·2
Max.[d]	298·0	204·2	141·1	426·1	128·1	335·0	357·8	118·8	166·8	239·9	111·0	105·2	85·6	185·5	150·1	49·1	60·7	40·7	169·6	149·1
Min.[e]	8·2	16·2	2·2	17·0	8·3	19·3	34·8	7·3	36·9	5·2	12·7	17·6	0·1	15·7	2·0	7·3	13·2	0·0	10·1	0·9
Mean	96·3	72·5	45·8	117·1	46·3	96·8	125·7	32·8	80·7	50·0	38·0	44·7	21·9	62·2	30·5	17·8	28·9	14·8	59·9	43·3
Maxd[f]	298·0	200·0	98·0	426·1	128·1	335·0	357·8	22·8	95·1	239·9	111·0	66·8	44·2	130·5	19·5	49·1	28·6	20·5	112·7	63·6
Mind[g]	8·2	16·2	7·9	17·0	8·8	19·3	34·8	15·5	39·8	20·5	12·7	25·6	12·9	18·8	6·1	7·3	13·2	5·9	21·5	14·2

[a] Observed river discharge data.
[b] Simulated (or predicted) river discharge.
[c] Absolute error, i.e. differences between observed data and predictions.
[d] Maximum value.
[e] Minimum value.
[f] Maximum river discharge and related AE indices.
[g] Minimum river discharge and related AE indices.

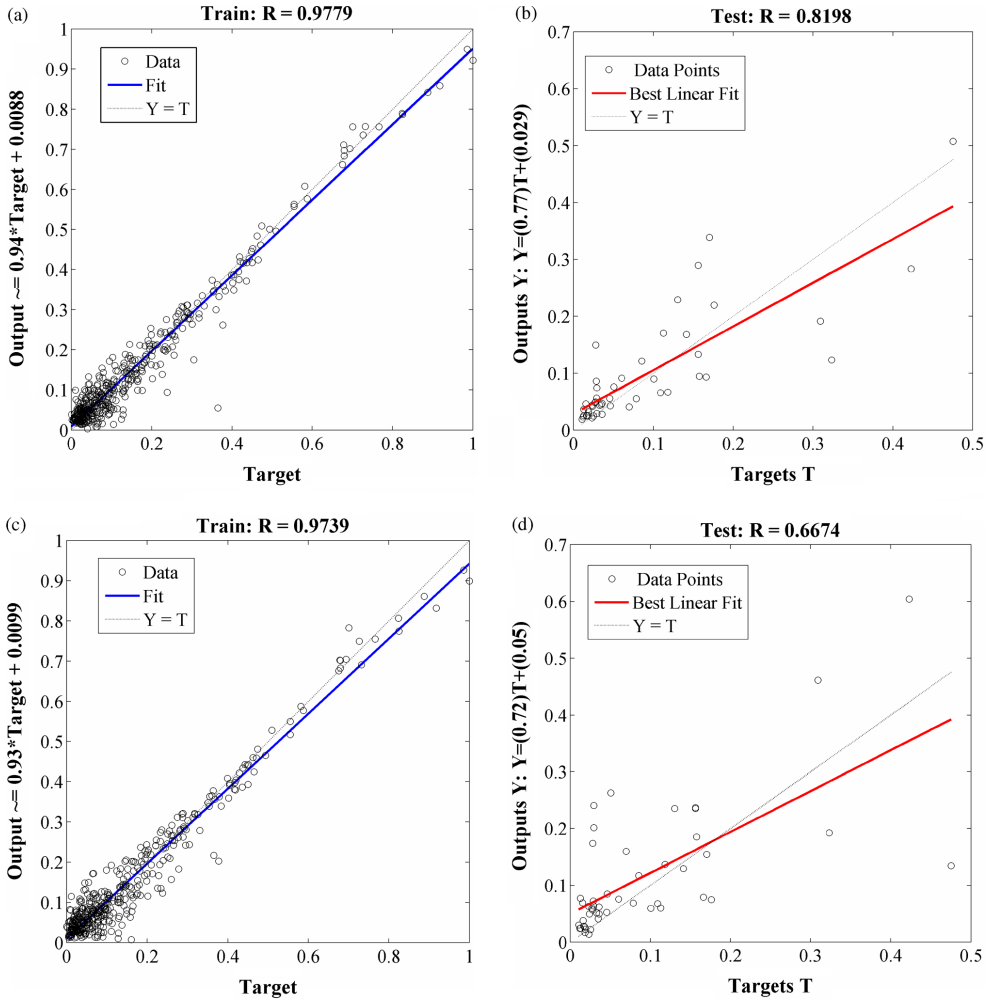

Figure 8. Regression plots displaying the network outputs with respect to targets after employing Bayesian regularisation in network training: (a) training result of WNN model db5L4, (b) testing result of model db5L4, (c) training results of model ANN alone, and (d) testing results of model ANN alone

the estimation of intermediate and moderate discharge in general, as well as the peak and lowest discharges. These results are very significant for flood prevention and river base discharge assessment. The final results of the study also confirmed that the WNN model at level 4 (db5L4) had the best simulation performance for modelling the river discharge. However, since only one river discharge time series was used for the WNN simulation and prediction analysis, more river discharge time series are required to conclusively prove the advantages of the proposed WNN modelling approach.

ACKNOWLEDGEMENTS

This work is part of the Project 'Determination of Environmental Flow Requirement and Its Safeguard Measures in the Wei River in China (2009DFA22980)' supported by the Sino-Swiss Science and Technology Cooperation Program, Switzerland and the Ministry of Science and Technology, China. We also would like to express our sincere thanks to Dr Gary Owens from the Centre for Environmental Risk Assessment and Remediation (CERAR) at the University of South Australia for reading the research paper and providing many useful scientific and English language corrections which we certainly believe helped to improve the quality of the final manuscript.

REFERENCES

Abibullaev B, Kim MS, Seo HD. 2009. Seizure detection in temporal lobe epileptic EEGs using the best basis wavelet functions. *Journal of Medical Systems* **34**(4): 755–765.

Abrahart RJ, See L. 2000. Comparing neural network and autoregressive moving average techniques for the provision of continuous river flow forecasts in two contrasting catchments. *Hydrological Processes* **14**: 2157–2172.

Aussem A, Campbell J, Murtagh F. 1998. Wavelet-based feature extraction and decomposition strategies for financial forecasting. *Journal of Computational Intelligence in Finance* **6**(2): 5–12.

Azrak R, Melard G. 1998. The exact quasi-likelihood of time dependent ARMA models. *Journal of Statistical Planning and Inference* **68**: 31–45.

Badiei E, Mohammadi S. 2011. An empirical analysis to study the cyclical trends on stock exchange using wavelet methods. *Management Science Letters* **1**: 57–64.

Beale MH, Hagan MT, Demuth MH. 2010. *Neural Network Toolbox™ 7 User's Guide*. MathWorks, Inc.: Natick, MA.

Cannas B, Fanni A, See L, Sias G. 2006. Data preprocessing for river flow forecasting using neural networks: wavelet transforms and data partitioning. *Physics and Chemistry of the Earth* **31**(18): 1164–1171.

Chang FJ, Chen YC. 2001. A counterpropagation fuzzy-neural network modeling approach to real time streamflow prediction. *Journal of Hydrology* **245**: 153–164.

Chen BF, Wang HD, Chu CC. 2007. Wavelet and artificial neural network analyses of tide forecasting and supplement of tides around Taiwan and South China Sea. *Ocean Engineering* **34**: 2157–2175.

Chen J, Adams BJ. 2006. Integration of artificial neural networks with conceptual models in rainfall-runoff modeling. *Journal of Hydrology* **318**: 232–249.

Chenoweth T, Hubata R, St. Louis RD. 2000. Automatic ARMA identification using neural networks and the extended samples autocorrelation function: a reevaluation. *Journal of Decision Support Systems* **29**: 21–30.

Chu CC, Chen BF. 2000. An application of artificial networks—tidal predication and supplement around Taiwan waters. In *The 22nd Ocean Engineering Conference*, Kaohsiung, Taiwan.

Cigizoğlu HK. 2004. Estimation and forecasting of daily suspended sediment data by multilayer perceptrons. *Advances in Water Resources* **27**: 185–195.

Daubechies I. 1990. The wavelet transform, time–frequency localization and signal analysis. *IEEE Transactions on Information Theory* **36**(5): 961–1005.

Dawson CW, Wilby RL. 2001. Hydrological modelling using artificial neural networks. *Progress in Physical Geography* **25**(1): 80–108.

DHI (Danish Hydraulic Institute). 2007. *Mike Zero: The common DHI user interface for project oriented water modeling*. DHI Water & Environment: Denmark.

Duan Q, Sorooshian S, Gupta VK. 1992. Effective and efficient global optimization for conceptual rainfall-runoff models. *Water Resources Research* **28**(4): 1015–1031.

Ertay T, Çekyay B. 2006. Integrated clustering modeling with backpropagation neutral network for efficient customer relationship management. In *Intelligent Data Mining: Techniques and Applications* Ruan D, Chen G, Kerre EE, Weta G (eds). Springer: Berlin; 355–373.

Feng LH, Hong WH. 2008. On hydrologic calculation using artificial neural networks. *Applied Mathematics Letters* **21**: 453–458.

Gardner G, Harvey AC, Phillips GDA. 1980. Algorithm AS 154: an algorithm for exact maximum likelihood estimation of autoregressive moving average models by means of Kalman filtering. *Journal of the Royal Statistical Society Series C—Applied Statistics* **29**: 311–322.

Goring D. 2006. Orthogonal wavelet decomposition. Available via Mulgor Consulting Ltd, web http://www.tideman.co.nz/Salalah/OrthWave Decomp.html. Accessed on 20 March 2011).

Hu NF, Wang AZ, Guan DX, Yuan FH, Jin CJ, Wu JB, Wang JJ. 2010. Multiple time scale analysis of precipitation series in Changbai Mountain Region in 1959–2006. *Journal of Applied Ecology* **21**(3): 549–56 (in Chinese).

Hwarng HB. 2001. Insights into neural network forecasting of time series corresponding to ARMA (p,q) structures. *Journal of Omega* **29**: 273–289.

Jain SK, Das D, Srivastava DK. 1999. Application of ANN for reservoir inflow prediction and operation. *Journal of Water Resources Planning and Management* **125**(5): 263–271.

Kim CK, Kwak IS, Cha EY, Chon TS. 2006. Implementation of wavelets and artificial neural networks to detection of toxic response behavior of chironomids (Chironomidae: Diptera) for water quality monitoring. *Ecological Modelling* **195**: 61–71.

Kim T, Valdes JB. 2003. Nonlinear model for drought forecasting based on a conjunction of wavelet transforms and neural networks. *Journal of Hydrologic Engineering* **8**(6): 319–328.

Kisi Ö. 2008. Stream flow forecasting using neuro-wavelet technique. *Hydrological Processes* **22**(20): 4142–4152.

Kisi Ö. 2009. Neural network and wavelet conjunction model for modelling monthly level fluctuations in Turkey. *Hydrological Processes* **23**: 2081–2092.

Kisi O, Cigizoglu HK. 2007. Comparison of different ANN techniques in river flow prediction. *Civil Engineering and Environmental Systems* **24**(3): 211–231.

Li H, Lian JJ, Wang XJ. 2008. Stepwise regression model for daily runoff prediction based on wavelet decomposition. *Journal of Hydraulic Engineering* **39**(12): 1334–1339 (in Chinese).

MacKay DJC. 1992. Bayesian interpolation. *Neural Computation* **4**(3): 415–447.

Maier HR, Dandy GC. 1996. The use of neural networks for the prediction of water quality parameters. *Water Resources Research* **32**(4): 1013–1022.

Mallat S. 1989. Multiresolution approximation and wavelet orthonormal bases of L2(R). *Transactions of the American Mathematical Society* **315**: 69–87.

Misiti M, Misiti Y, Oppenheim G, Poggi JM. 2008. *Wavelet Toolbox™ 4 User's Guid*. The MathWorks, Inc.: Natick, MA.

Misiti M, Misiti Y, Oppenheim G, Poggi JM. 2010. *Wavelet Toolbox™ 4 Getting Started Guide*. The MathWorks, Inc.: Natick, MA.

Mohamed MA, Atta MM. 2010. Automated classification of Galaxies using transformed domain features. *International Journal of Computer Science and Network Security* **10**(2): 86–91.

Mohammadi K, Eslami HR, Kahawita R. 2006. Parameter estimation of an ARMA model for river flow forecasting using goal programming. *Journal of Hydrology* **331**: 293–299.

Müller-Wohlfeil DI, Xu CY, Iversen HL. 2003. Estimation of monthly river discharge from Danish catchments. *Nordic Hydrology* **34**(4): 295–320.

Neitsch SL, Arnold JG, Kiniry JR, Williams JR. 2005. Soil and water assessment tool: Theoretical documentation (version 2005). Available via the Official SWAT WebSite http://swatmodel.tamu.edu/documentation. Accessed on 26 December 2009.

Nilsson P, Uvo CB, Berndtsson R. 2006. Monthly runoff simulation: Comparing and combining conceptual and neural network models. *Journal of Hydrology* **321**: 344–363.

Nourani V, Alami MT, Aminfar MH. 2009. A combined neural-wavelet model for prediction of Ligvanchai watershed precipitation. *Engineering Applications of Artificial Intelligence* **22**: 466–472.

O'Connor KM. 1997. Applied hydrology I-deterministic. Unpublished Lecture Notes. Department of Engineering Hydrology, National University of Ireland, Galway.

Partal T. 2009. Modelling evapotranspiration using discrete wavelet transform and neural networks. *Hydrological Processes* **23**: 3545–3555.

Partal T, Cigizoglu HK. 2008. Estimation and forecasting of daily suspended sediment data using wavelet–neural networks. *Journal of Hydrology* **358**: 317–331.

Partal T, Kişi Ö. 2007. Wavelet and neuro-fuzzy conjunction model for precipitation forecasting. *Journal of Hydrology* **342**: 199–212.

Pulido-Calvo I, Portela MM. 2007. Application of neural approaches to one-step daily flow forecasting in Portuguese watersheds. *Journal of Hydrology* **332**: 1–15.

Rioul O, Vetterli M. 1991. Wavelets and signal processing. *IEEE Signal Process* **8**: 14–38.

Salas JD, Markus M, Tokar AS. 2000. Streamflow forecasting based on artificial neural networks. In *Artificial Neural Networks in Hydrology,* Govindaraju RS, Rao AR (eds). Kluwer: Berlin.

Salerno F, Tartari G. 2009. A coupled approach of surface hydrological modelling and Wavelet Analysis for understanding the baseflow components of river discharge in karst environments. *Journal of Hydrology* **376**: 295–306.

Sasi B, Rao BPC, Jayakumar T, Raj B. 2010. Wavelet Transform-Based Denoising Method for Processing Eddy Current Signals. *Research in Nondestructive Evaluation* **21**(3): 157–170.

Sivakumar B, Jayawardena AW, Fernando TMKG. 2002. River flow forecasting: use of phase space reconstruction and artificial neural networks approaches. *Journal of Hydrology* **265**: 225–245.

Song JX, Xu ZX, Liu CM, Li HE. 2007. Ecological and environmental instream flow requirements for the Wei River—the largest tributary of the Yellow River. *Hydrological Processes* **21**: 1066–1073.

Srinivasulu S, Jain A. 2009. River flow prediction using an integrated approach. *Journal of Hydrologic Engineering* **14**(1): 75–83.

Tokar AS, Johnson PA. 1999. Rainfall-runoff modeling using artificial neural networks. *Journal of Hydrologic Engineering* **4**(3): 232–239.

Toth E, Brath A, Montanari A. 2000. Comparison of short-term rainfall prediction models for real-time flood forecasting. *Journal of Hydrology* **239**: 132–147.

Tsai CP, Lee TL. 1999. Back-propagation neural network in tidal-level forecasting. *Journal of Waterway, Port, Coastal, and Ocean Engineering* **125**(4): 195–202.

Tsai CP, Lin C, Shen JN. 2002. Neural network for wave forecasting among multi-stations. *Ocean Engineering* **29**(13): 1683–1695.

Valença M, Ludermir T, Valença A. 2005. River flow forecasting with constructive Neural Network. In *AI 2005: Advances in Artificial Intelligence*, Zhang S, Jarvis R. (eds). Springer: Berlin; 1031–1036.

Vandewiele GL, Nir-Lar-Win. 1993. Monthly water and snow balance models on basin scale. In *Runoff and Sediment Yield Modelling,* Banasik K, Zbikowski A (eds). Warsaw Agricultural University: Warsaw; 83–88.

Wang HR, Kang J, Lin X, Qian LX. 2008. Problems existing in ARIMA model of hydrologic series and some improvement suggestions. *System Engineering- Theory & Practice* **10**: 166–176 (in Chinese).

Wang W, Ding J. 2003. Wavelet network model and its application to the predication of hydrology. *Nature and Science* **1**(1): 67–71.

Wang W, Van Gelder PHAJM, Vrijling J.K, Ma J. 2006. Forecasting daily streamflow using hybrid ANN models. *Journal of Hydrology* **324**: 383–399.

Wang WK, Kong JL, Duan L, Wang YL, Ma XD. 2004. Research on the conversion relationships between the river and groundwater in the Yellow River drainage area. *Science in China Series E-Engineering & Materials Sciences* **14**(1): 25–41.

Wei S. 2008. On the use of game theoretic Models for water resources management, PhD. dissertation. The Faculty of Environmental Sciences and Process Engineering, the Brandenburg University of Technology (BTU) Cottbus, Germany.

Xu CY, Seibert J, Halldin S. 1996. Regional water balance modelling in the NOPEX area: development and application of monthly water balance model. *Journal of Hydrology* **180**: 211–236.

Zealand CM, Burn DH, Simonovic SP. 1999. Short term streamflow forecasting using artificial neural networks. *Journal of Hydrology* **214**: 32–48.

Zhang BL, Dong ZY. 2001. An adaptive neural wavelet model for short-term load forecasting. *Electric Power Systems Research* **59**: 121–129.

Zhao H. 2003. Evaluation of water quality and countermeasures of pollution prevention for the Wei River. *Northwest Water Resource and Water Engineering* **14**(1): 28–31 (in Chinese).

