

Daily river flow forecasting using wavelet ANN hybrid models

Niranjan Pramanik, Rabindra K. Panda and Adarsh Singh

ABSTRACT

Advance time step stream flow forecasting is of paramount importance in controlling flood damage. During the past few decades, artificial neural network (ANN) techniques have been used extensively in stream flow forecasting and have proven to be a better technique than other forecasting methods such as multiple regression and general transfer function models. This study uses discrete wavelet transformation functions to preprocess the time series of the flow data into wavelet coefficients of different frequency bands. Effective wavelet coefficients are selected from the correlation analysis of the decomposed wavelet coefficients of all frequency bands with the observed flow data. Neural network models are proposed for 1-, 2- and 3-day flow forecasting at a site of Brahmani River, India. The effective wavelet coefficients are used as input to the neural network models. Both the wavelet and ANN techniques are employed to form a loose type of wavelet ANN hybrid model (NW). The hybrid models are trained using Levenberg–Marquart (LM) algorithm and the results are compared with simple ANN models. The results revealed that the predictabilities of NW models are significantly superior to conventional ANN models. The peak flow conditions are predicted with better accuracy using NW models than compared to ANN models.

Key words | artificial neural networks, discrete wavelet transformation, flow forecasting

Niranjan Pramanik (corresponding author)
Rabindra K. Panda
Adarsh Singh
Agricultural and Food Engineering Department,
Indian Institute of Technology,
Kharagpur 721302,
West Bengal,
India
E-mail: nrrjnp@iitkgp.ac.in

INTRODUCTION

The ability to compute stream flow quickly and accurately is of crucial importance in flood forecasting. Hydrodynamic models provide a sound physical basis for this purpose and have the capability to simulate a wide range of flow situations. However, these models require accurate stream geometry data, which are generally scanty in most of the developing countries such as India. It is also not possible to integrate the observed data directly at desired locations to improve the model results using physically based models. Moreover, physically based models require high computation time thereby making them incapable of simulating real-time forecasting required for flood warnings.

To overcome the above-mentioned difficulties associated with the application of physically based models, data-driven techniques are commonly adopted by various

researchers for modelling water resources systems. The data-driven methods such as artificial neural networks (ANNs), fuzzy theory and chaos theory have been widely applied in the domain of hydrology and water resources (Zealand *et al.* 1999; Dawson & Wilby 2001; Bray & Han 2004; Nayak *et al.* 2005, 2007; Shrestha *et al.* 2005; Han *et al.* 2006, 2007; Sahoo & Ray 2006; Pramanik & Panda 2009).

The application of ANNs has been extended to their use in remote sensing data for retrieving hydrological variables such as precipitation at different scales, needed for stream flow forecasting (Evora & Coulibaly 2009). A comprehensive review of the application of ANNs in hydrology is presented in the ASCE Task Committee report (2000a,b). Comparison of the performances of different optimization algorithms and types of ANN networks used in neural

doi: 10.2166/hydro.2010.040

network modelling have been made by various researchers (Kumar *et al.* 2005; Chau 2006; Kisi 2007, 2008a; Kisi & Cigizoglu 2007; Mukerji *et al.* 2009). Feed forward ANN architecture using Levenberg–Marquart (LM) back-propagation algorithm is reported to perform well in most of the investigations.

Tokar & Johnson (1999) reported that ANN models provide greater training and testing accuracy than that of the regression and simple conceptual models. In this respect, some researchers attempted coupling of neural networks with linear dynamic models such as autoregressive with exogenous input (ARX) and autoregressive moving average with exogenous inputs (ARMAX) to form neural network auto-regressive with exogenous input (NNARX) and neural network auto regressive moving average with exogenous input (NNARMAX) (Gautam *et al.* 2000; Kishor & Singh 2007). They found that NNARMAX models perform superior to simple ANN models.

In many hydrologic applications, fuzzy partitioning mechanisms are adopted to the input data for creating a rule base to generate the output. Jacquin & Shamseldin (2009) provided a detailed review on the application of fuzzy inference system (FIS) in river flow forecasting. They found that use of FIS is not widespread by the hydrologists compared to ANNs. In FIS, the back-propagation algorithm is used to optimize the fuzzy membership parameters for best input-output mapping. The integration of fuzzy inference system with the back-propagation algorithm leads to the development of adaptive neuro-fuzzy inference system (ANFIS). A good number of studies are reported on rainfall–runoff modelling and river flow forecasting using ANFIS together with its performance comparison with the ANN technique (Chau *et al.* 2005; Nayak *et al.* 2005; Aqil *et al.* 2007; Mukerji *et al.* 2009; Pramanik & Panda 2009). In most of the investigations, it is reported that ANFIS performs better than neural networks and fuzzy models.

Most of the studies dealing with the integration of ANN with other techniques such as fuzzy inference system, data normalization and principal component analysis fall under the phase of data pre-processing in neural network modelling. Data pre-processing is an important step in ANN modelling, which leads to a reduction in the forecasting error. Data pre-processing methods uses different statistical techniques to explore the hidden properties in

the datasets that help in efficient input-output mapping during model training. Normalization of the raw data within appropriate ranges and selection of important inputs are most vital in ANN modelling to enhance the forecasting ability of the models (Hsu *et al.* 1995; Gunn 1998; Dawson & Wilby 1999; Kisi 2004).

Partial auto-correlation analysis is employed to identify the suitable lag time in the time series data and inputs are selected for development of ANN models based on this (Sudheer *et al.* 2002). Principal component analysis is adopted to reduce the dimensions of input data matrix to build simpler ANN architecture to yield good results (Wilby *et al.* 2003). Another data-driven method referred as Evolutionary Polynomial Regression (EPR), developed by Giustolisi & Savic (2004, 2006), is used in hydrological data analysis and water resources modelling. EPR integrates the best features of numerical regression (Draper & Smith 1998) with genetic programming (Koza 1992). Recently, Giustolisi & Savic (2009) improved the EPR strategy using multi-objective genetic algorithm (EPR-MOGA) to select best models and successfully applied the method in the prediction of ground water fluctuation.

Some recent studies reported that integration of wavelet transformation technique with ANN yields superior results compared to simple ANN and regression models (Anctil & Tape 2004; Chou & Wang 2004; Zhou *et al.* 2006; Partal & Kisi 2007; Nourani *et al.* 2008; Kisi 2008b, 2009a,b; Remesan *et al.* 2009). This advanced pre-processing of raw data to capture the non-stationary behaviour of the time series data by decomposing the original series into wavelet coefficients of different frequency bands has been effectively applied by Wang & Ding (2003) for short term flood forecasting.

Anctil & Tape (2004) used a continuous wavelet transform technique through histograms and power spectra to identify the major wavelet bands based on which wavelet-ANN models are developed and trained for 1-day stream flow forecasting. Adamowski (2008) performed wavelet and cross wavelet analysis of flow and meteorological time series to exploit the non-stationarity in the time series data and developed wavelet-based constituent components to use them in the forecasting models for 1-, 2- and 6-day advance flood forecasting.

The superiority of wavelet-ANN hybrid models is explained by their ability to capture the useful information

in the time series on various frequency components (Wang & Ding 2003; Labat 2005; Liang & Shizhong 2006). Kisi (2008b) determined the effective wavelet components based upon the results of correlation analysis of the wavelet coefficients of different components with the observed stream flow data. He proposed a wavelet-ANN model for monthly stream flow forecasting. However, very few studies are reported on short-term stream flow forecasting using wavelet-ANN approach so far. The use of wavelet transformation theory in hydrology and its coupling with ANN is a recent attempt, which needs further investigation using datasets of different basins worldwide.

The present study aims to forecast short-term flow up to 3 days in advance using wavelet-ANN approach at a site of the Brahmani River, India. In the present study, the wavelet transformation technique was used to generate the wavelet coefficients of different frequency components. The effective components were selected based on the correlation analysis of the wavelet coefficients of different resolution scales. Three ANN models were proposed for 1-, 2- and 3-day advance stream flow forecasting at the Jenapur site of the River Brahmani. ANN models were trained and tested using the time series of the wavelet coefficients as inputs to the ANN. Finally, the results obtained from the wavelet ANN hybrid models were compared with simple ANN models.

STUDY AREA

The wavelet ANN hybrid approach was applied to the Brahmani River located in the eastern part of India. The lower reach of the Brahmani River is the main cause of flood in the downstream delta area of the basin. During the rainy season (June–September), flood is a common problem in the deltaic region of the basin due to lowland topography of the floodplain and low carrying capacity of the river. At the upstream location of the lower reach of Brahmani, daily water level and discharge are measured and flood forecasting is conducted to save lives and properties during the period of heavy rain. The water level and discharge at that location (Jenapur) has been continuously monitored by the Central Water Commission (CWC) of the Government of India. The average annual rainfall of the basin is about 1,305 mm. The maximum discharge and water level at Jenapur site of

the river has been recorded as 24,246 m³/s and 24.78 m respectively on 20 August 1975 (Jain *et al.* 2007). A location map of the Brahmani basin depicting the Jenapur gauging site is shown in Figure 1. The time series data pertaining to daily water level and discharge at the site for seven years (1999–2005) were collected from the CWC, Bhubaneswar, India and were used in this study.

THEORETICAL BACKGROUND

Artificial neural networks (ANNs)

An ANN is a computing system made up of interconnected information processing units which are analogous to biological neurons. The neurons collect inputs and produce output in accordance with certain transfer functions. The modelling using ANN is performed by optimizing the connection weights between the nodes in order to produce output close to the observed values.

The fundamental principle of neural computing is due to its decomposition property, which generates series of linearly separable steps for input-output relationships using hidden layers (Haykin 1994). There are three basic layers or levels of data processing units in ANN, namely the input layer, the hidden layer and the output layer (Figure 2).

Each of these layers consists of processing units called neurons. The connection between the processing neurons is weighted by scalar weight (w), which is adapted during model training, and a bias (b). The number of input neurons, output neurons and the neurons in the hidden layer depend upon the problem being studied. If the number of neurons in the hidden layer is small, the network may not have sufficient degrees of freedom to learn the process correctly. If the number is too high, the training will take a long time and the network may sometimes over-fit the data (Karunanithi *et al.* 1994).

During the training process, the connection weights are updated. At the beginning of the training, the initial values of the weights can be assigned randomly or based on experience. In this process, the learning algorithm systematically changes the weights to correctly perform a desired input-output relationship. The process of training is said to be finished when the difference between the ANN output

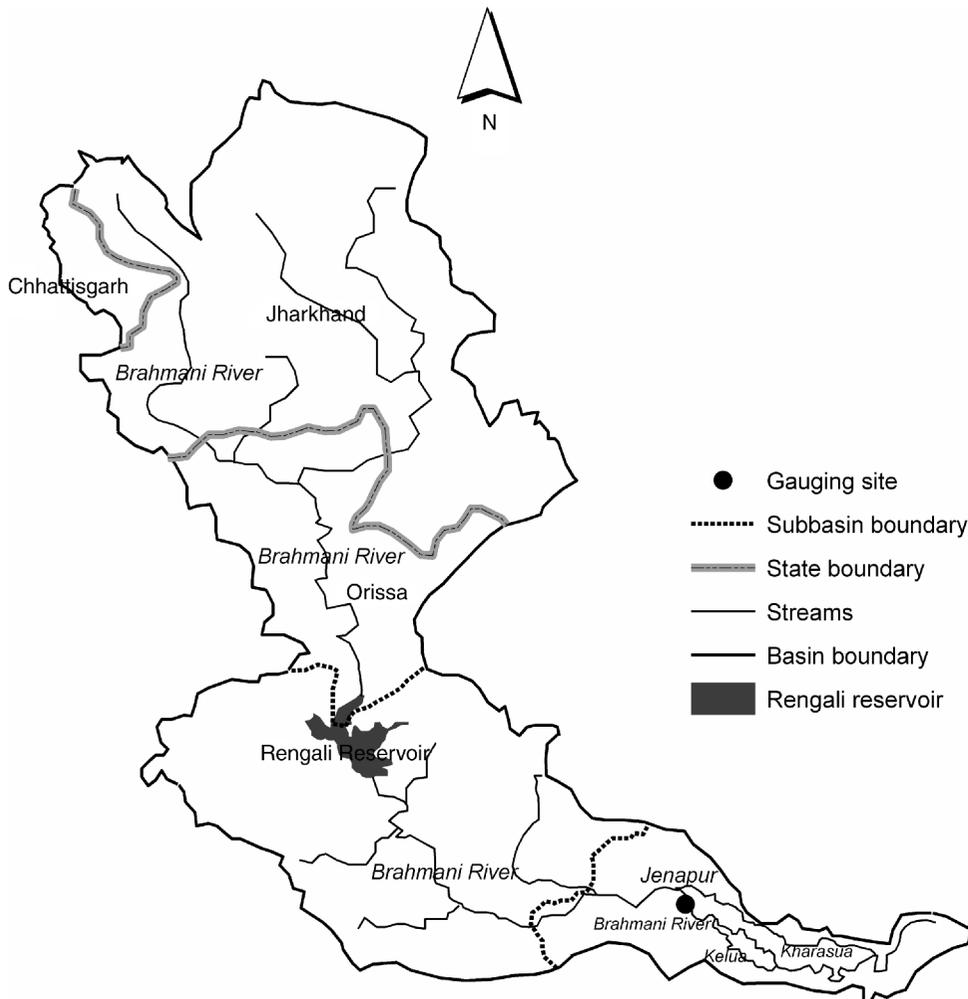


Figure 1 | Index map of Brahmani basin showing streams and Jenapur gauging site.

and the actual output becomes closer to the preset error goal. During the testing phase, the performance of the trained ANN model is evaluated using unseen datasets.

There are four distinct steps followed in the development of an ANN-based solution to a certain problem such as flood forecasting. The first step is data transformation, scaling or normalization. A large variation in the input data can slow down or even prevent the training of the network. To overcome this problem, the data are usually scaled using statistical, min-max, sigmoid or principal component transformations (Priddy & Keller 2005). It is also important that absolute input values are scaled to avoid asymptotic issues (Haykin 1994). The second step is the definition of network architecture, in which the number of hidden layers, the number of neurons in each layer and the connectivity

between the neurons are set. Obtaining an optimum number of hidden neurons in an ANN architecture is a trial and error process, which depends on the quality of data and the type of the problem. In the third step, a learning algorithm is used to train the network to respond correctly to a given set of inputs. In the last step, the trained architecture (having the optimized values of the connection weights) is tested using the unseen datasets to evaluate the model performance.

Wavelet analysis

In wavelet analysis, the signals are analyzed both in the time and frequency domain by decomposing the original signals in different frequency bands using wavelet functions. This is

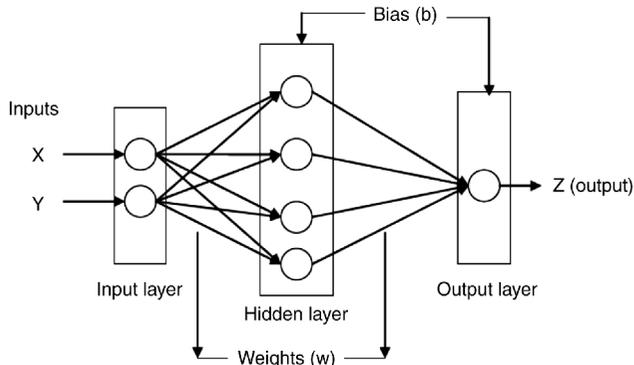


Figure 2 | ANN architecture with one hidden layer (Pramanik & Panda 2009).

different from Fourier analysis, in which signals are analyzed using sine and cosine functions. The wavelet transform (WT) uses the scalable windowing technique for analyzing local variation in the time series (Torrence & Compo 1998). Wavelet transforms provide useful decompositions of original time series, so that wavelet-transformed data improve the ability of a forecasting model by capturing useful information on various resolution levels (Adamowski 2008).

The time series data are pre-processed using wavelet transformation techniques to obtain decomposed wavelet coefficients that are used as inputs in the forecasting models (Aussem et al. 1998; Zheng et al. 2000; Zhang & Dong 2001). Several researchers have applied wavelet transformation techniques to different areas of hydrology, e.g. rainfall-runoff modelling, flood forecasting and stream flow prediction (Smith et al. 1998; Labat et al. 1999; Coulibaly et al. 2000; Saco & Kumar 2000; Kim & Valdes 2003; Partal & Kisi 2007; Remesan et al. 2009). They observed that use of wavelet techniques, to pre-process the hydrologic time series data into series of wavelet coefficients of different resolutions, produced significantly better results than that of the linearly scaled inputs (simple ANN models) when used as inputs.

Discrete wavelet transform

The basic objective of wavelet transforms is to achieve a complete timescale representation of localized and transient phenomena occurring at different timescales (Labat et al. 2000). The continuous wavelet transform is defined as

the sum over all time of the signal multiplied by scale and shifted versions of wavelet function ψ :

$$W(a, b) = \frac{1}{\sqrt{a}} \int_{-\infty}^{+\infty} f(x) \psi * \left(\frac{x-b}{a} \right) dx \quad (1)$$

where a is a scale parameter; b is a position parameter; and $*$ corresponds to the complex conjugate.

Several families of wavelets (ψ) that have proven to be useful for various applications are described in related references (Mallat 1998; Rao & Bopardikar 1998). The coefficient plots of the continuous wavelet transform are precisely the timescale view of the signal. However, calculating wavelet coefficients at every possible scale is time consuming and generates large amount of information. Thus, the use of the continuous wavelet transform for forecasting is not practically possible.

Discrete wavelet transformation (DWT) is therefore preferred in most of the forecasting problems of water resource systems because of its simplicity and ability to compute with less time (Cannas et al. 2005, 2006). The DWT involves choosing scales and positions based on powers of two, so-called dyadic scales and translations. The mother wavelet is rescaled or dilated by powers of 2 and translated by integers. The DWT algorithm is capable of producing coefficients of fine scales for capturing high-frequency information and coefficients of coarse scales for capturing low-frequency information. The DWT with respect to a mother wavelet is defined:

$$f(t) = \sum c_{j_0, k} \phi_{j_0, k}(t) + \sum_{j>j_0} \sum w_{j, k} 2^{j/2} \psi(2^j t - k) \quad (2)$$

where j is the dilation or level index, k is the translation or scaling index and $\phi_{j_0, k}$ is a scaling function of coarse scale coefficients. $c_{j_0, k}$, $w_{j, k}$ is the scaling function of detail (fine scale) coefficients and all functions of $\psi(2^j t - k)$ are orthonormal.

High-pass and low-pass filters of different cut-off frequencies are used to separate the signal at different scales. The time series is decomposed into one containing its trend (the approximation) and one containing the high frequencies and the fast events (the detail). The scale is changed by upsampling and downsampling operations. The DWT has many advantages in compressing a wide range of

signals. With the DWT technique, a very large proportion of the coefficients of the transform can be set to zero without appreciable loss of information. In addition, if more properties other than the stationary properties of a signal are desired, DWT would certainly be a better choice than that of the traditional Fourier transformation technique. The limitation of the application of DWT in time-series analysis is that it suffers from a lack of translation invariance.

METHODOLOGY

Data preparation for ANN models

A remarkable property of ANNs is its ability to handle nonlinear, noisy and non-stationary data. However, with suitable data preparation beforehand, it is possible to improve the modelling performance (Maier & Dandy 2000; Bray & Han 2004). Data preparation involves a number of processes such as data collection, data division and data pre-processing. In the present study, daily flow as well as water level data at Jenapur gauging site were collected for a period of 7 years (1999–2005) from CWC, Bhubaneswar, India. The total dataset was divided into training and testing samples. Many researchers have discussed the process of data division in ANN modelling (ASCE 2000a). In this connection, the appropriate size of datasets used to train the ANN models in order to obtain accurate result is also reported by the ASCE task committee (2000a,b). Normally a major portion of the datasets is used for training and a relatively smaller portion is used for testing the ANN models.

In the present study, data for the period 1999–2003 and 2004–2005 were used for ANN training and testing, respectively. The descriptive statistics of the flow datasets used in the study are presented in Table 1. The minimum, maximum and average water level at Jenapur site during the period 1999–2005 were observed to be 17.27 m, 23.36 m and 18.41 m, respectively. In ANN application, the time series of flow data are normalized between suitable ranges to obtain better training (Shanker et al. 1996; Gunn 1998; Singh et al. 2009). The appropriate data normalization range for daily flow forecasting using ANN is reported in

Table 1 | Basic statistics of the training and testing datasets

Basic statistics	Training (1999–2003)	Testing (2004–2005)
Sample size	1,826	731
Minimum flow (m ³ /s)	37.0	56.535
Maximum flow (m ³ /s)	10,077.0	10,314.0
Flow range (m ³ /s)	10,040.0	10,257.0
Mean flow (m ³ /s)	568.63	463.37
Std. deviation (m ³ /s)	957.45	748.48
Std. error (m ³ /s)	25.049	27.683
Skewness	4.5754	6.658
Kurtosis	28.416	67.981

Singh et al. (2009). They observed that the normalization range of 0.2 to 0.8 produced better results as compared to other normalization ranges such as –1.0 to 1.0, 0 to 1 and 0.1 to 0.9 considered in their study. In the present study, the data were normalized in the range of 0.2 to 0.8 and were used in the training ANN architectures.

Input selection for ANN models

Partial auto-correlation analysis of the time series of daily discharge data was carried out to study the effect of preceding flow values on succeeding flow values. The partial autocorrelation statistics at 95% confidence bands from zero time lag to 20-day time lag were estimated and presented in Figure 3. It was observed that there is significant correlation up to 3-day time lags in the time

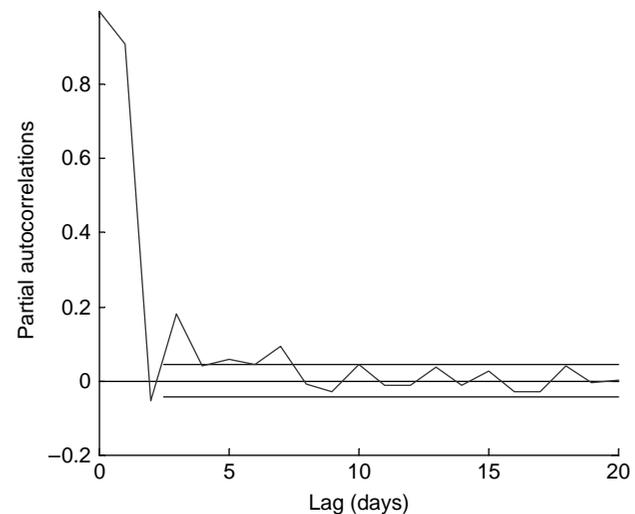


Figure 3 | Partial autocorrelation plot of the daily flow series.

series of the flow data, which indicates the antecedent flow values up to 3-day lag should be considered as input in the models for advance time step forecasting. Therefore, three antecedent flow values were selected for developing ANN models.

Three models were proposed for 1-, 2- and 3-day flood forecasting and are presented in Table 2. In addition to the consideration of flow values in the input datasets, time series of water level data was also included as input in the models. This is because of the water level values corresponding to 1-day previous time step has considerable influence on the flood situation at current time step. In the case of all three models, antecedent flow values up to 3-day time lags i.e. $Q_{(t-1)}$, $Q_{(t-2)}$ and $Q_{(t-3)}$ along with Q_t and H_t were included in the input data matrix. Q_t and H_t are the current (at time t) flow and stage values at Jenapur gauging site, respectively. The Levenberg–Marquart (LM) algorithm was used to train the ANN models because of its wide application and superiority over other training algorithms such as conjugate gradient (CG) and gradient descent with momentum (GD) as reported in previous studies (Pramanik & Panda 2009). Moreover, the LM algorithm is faster than other algorithms and is the standard method for minimization of the mean square error (MSE) criterion, due to its rapid convergence properties and robustness.

Wavelet decomposition of the flow time series

Deciding the optimal decomposition level of the time series data in wavelet analysis plays an important role in preserving the information and reducing the distortion of the datasets. Information content in the wavelet coefficients at every level contains a certain entropy value that serves as a measure of roughness present in the time series data. In reality, decomposition can proceed until the individual details consist of a single sample. The number of decomposition levels controls the flow approximation in the data. The decomposition of the time series to a higher level than

optimum can disturb the inherent properties of the datasets. Therefore, the maximum allowable decomposition level is avoided and a smaller value is considered.

In the present study, the `wmaxlev` library function of Matlab was used to determine the maximum level decomposition of the flow data. It was found that flow time series could be decomposed up to a maximum of 12 resolution levels. However, in the present study, 10 levels were adopted, based on the above discussions, to decompose the time series data for onward use in the ANN models.

The observed daily time series of the flow data were decomposed using the DWT technique. Depending on the selected resolution levels, the time series of flow data were decomposed into a number of wavelet components. The observed time series of discharge flow data was decomposed at 10 decomposition levels (2–4–8–16–32–64–128–256–512–1,024 day). The time series of the 2-day mode (DW_1), 4-day mode (DW_2), 8-day mode (DW_3), 16-day mode (DW_4), 32-day mode (DW_5), 64-day mode (DW_6), 128-day mode (DW_7), 256-day mode (DW_8), 512-day mode (DW_9) and 1,024-day mode (DW_{10}) are presented in Figure 4. The approximate and the original stream flow series are presented in Figure 5. Matlab codes were developed using its library functions to perform wavelet decomposition of the time series data and the required computation in the whole ANN modelling process.

Determination of effective wavelet components

The effective wavelet components were decided based on the higher correlation coefficients obtained from the correlation analysis between the observed flow data and the wavelet coefficients of different decomposition levels (Partal & Kisi 2007). The correlations between each wavelet component of 1-day previous daily stream flow data and the observed daily stream flow data were performed and are presented in Table 3.

It is observed that the correlation between the wavelet component DW_8 (256-day periodical component) of the daily stream flow and the observed daily stream flow data has the highest correlation value ($R = 0.4572$). In addition, the DW_1 , DW_2 , DW_3 , DW_4 , DW_5 , DW_6 and DW_7 components show significantly higher correlations compared to the remaining DW components. From Table 3, it is

Table 2 | ANN models for 1-, 2- and 3-day flood forecasting

Models	Output	Inputs
ANN-1	$Q_{(t+1)}$	$f(Q_t, Q_{(t-1)}, Q_{(t-2)}, Q_{(t-3)}, H_t)$
ANN-2	$Q_{(t+2)}$	$f(Q_t, Q_{(t-1)}, Q_{(t-2)}, Q_{(t-3)}, H_t)$
ANN-3	$Q_{(t+3)}$	$f(Q_t, Q_{(t-1)}, Q_{(t-2)}, Q_{(t-3)}, H_t)$

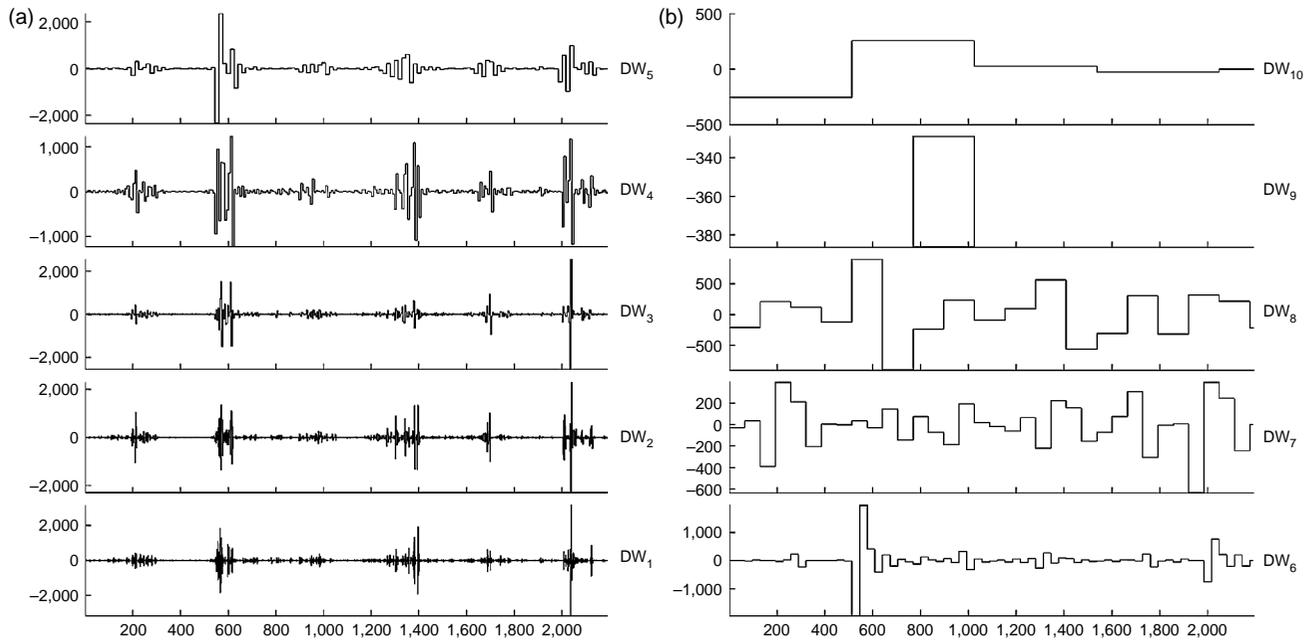


Figure 4 | Decomposed wavelet sub time series components (a) DW₁, DW₂, DW₃, DW₄, DW₅ and (b) DW₆, DW₇, DW₈, DW₉, DW₁₀ of stream flow data.

observed that the correlation between the DW₈ component of 1-day previous time step with the observed daily stream flow values has the highest correlation value ($R = 0.4595$). Also, DW₂, DW₃, DW₄, DW₅, DW₆ and DW₇ show high correlation coefficients (Table 3).

These results show that the periodic components, DW₃, DW₄, DW₅, DW₆, DW₇ and DW₈ (decomposition levels varying between 8 and 256 days) of daily stream flow and past stream flow data have notably high correlations with the observed daily stream flow data. The correlation

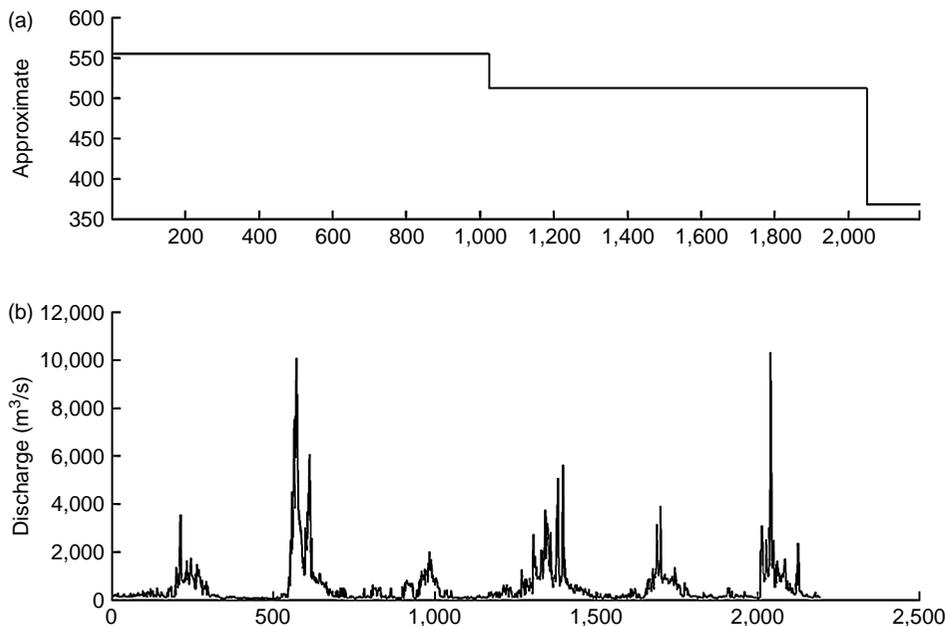


Figure 5 | (a) Approximate wavelet sub-series and (b) original time series of the flow data.

Table 3 | The correlation coefficients between the discrete wavelet components and the observed discharge data

Discrete wavelet components	Correlation between Q_t and DW_{it}	Correlation between the Q_t and $DW_{i(t-1)}$	Correlation between the Q_t and $DW_{i(t-2)}$
DW ₁	0.0021	0.00686	-0.12303
DW ₂	0.0237	0.05443	-0.21349
DW ₃	0.2645	0.28879	0.21547
DW ₄	0.2838	0.26737	0.18431
DW ₅	0.4050	0.38770	0.26123
DW ₆	0.4321	0.43224	0.27653
DW ₇	0.4463	0.43625	0.26134
DW ₈	0.4572	0.45957	0.32196
DW ₉	0.2803	0.28065	0.27085
DW ₁₀	0.1968	0.19654	0.18693
Approximation	0.0093	0.04234	0.03278

analysis between the wavelet components of 2-day previous daily stream flow data with the observed daily stream flow data was also performed and is presented in Table 3. A notable correlation was observed between the observed daily stream flow data with the periodic component, DW₆ and DW₈ of two-day lag. The correlation analysis between the DW values of one day lag and Q_{t+1} , Q_{t+2} and Q_{t+3} flow series was performed. It was found that the correlation coefficients decreased from Q_{t+1} to Q_{t+3} . The results of the correlation analysis showed that DW₄, DW₅, DW₆, DW₇ and DW₈ are the most effective components to be considered for forecasting.

Based on the correlation coefficients, DW₃, DW₄, DW₅, DW₆, DW₇ and DW₈ were selected as the dominant wavelet components. A new time series was composed by adding the dominant wavelet components. Addition of all dominant wavelet components obtained from correlation analysis into a single sum series (SW) was performed to take into account the integrated effect of all the inherent properties of the dominant components to capture the non-linearity in the input-output mapping. Summing up all the dominant components into a single series would represent the transformed form of the observed time series having more relevant properties that would help to portray the non-linearity of the flow data. Moreover, use of each dominant wavelet component as separate inputs would have resulted in a larger input data matrix, which would

Table 4 | Wavelet ANN models for 1-, 2- and 3-day flood forecasting

Models	Output	Inputs
NW-1	SW _(t+1)	$f(SW_t, SW_{(t-1)}, SW_{(t-2)})$
NW-2	SW _(t+2)	$f(SW_t, SW_{(t-1)}, SW_{(t-2)})$
NW-3	SW _(t+3)	$f(SW_t, SW_{(t-1)}, SW_{(t-2)})$

have made the ANN architecture more complex. The new series was used as inputs in the NW models for 1-, 2- and 3-day advance flow forecasting. Accordingly, three wavelet-ANN models were proposed using the values of the new time series corresponding to time t , $t-1$ and $t-2$ (Table 4).

Performance evaluation of the models

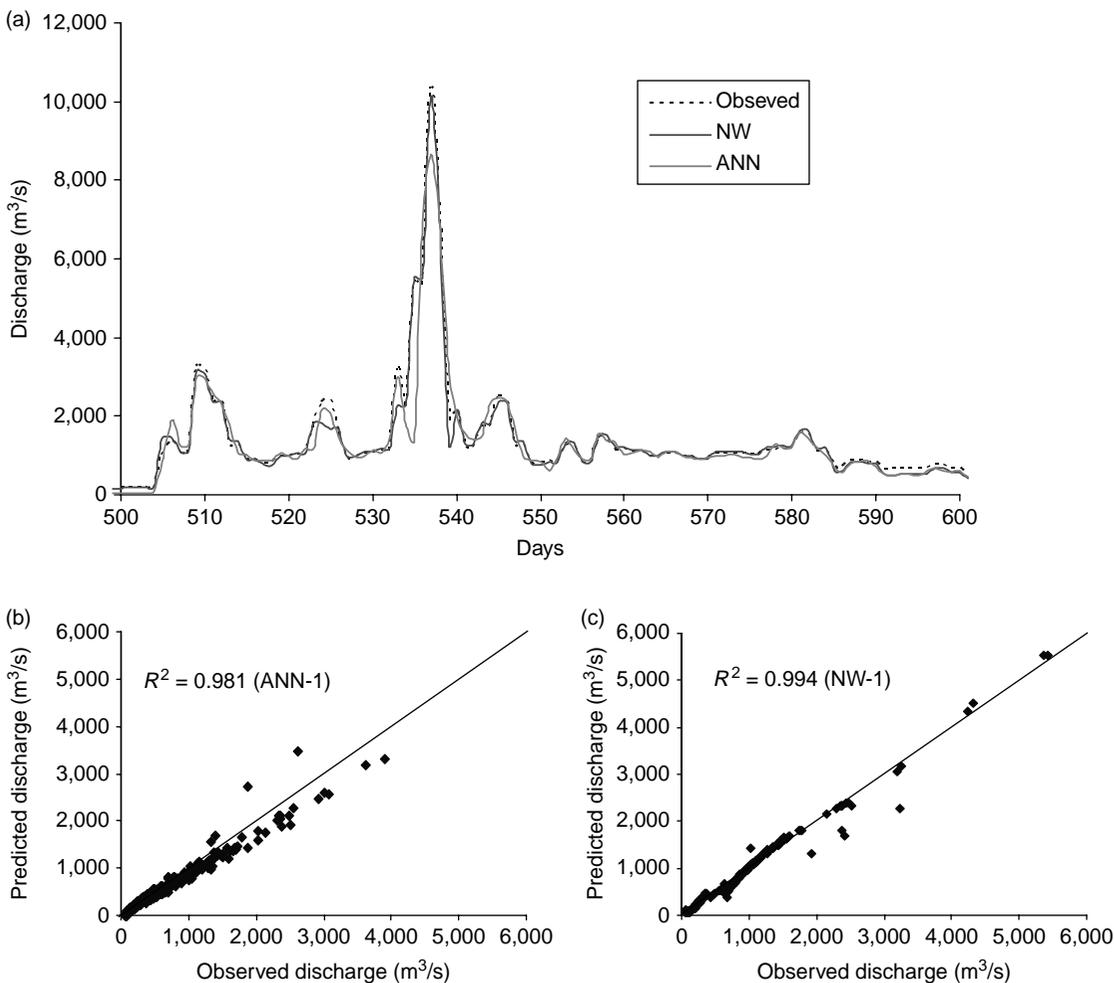
The performance of the trained network was assessed using the standard goodness of fit criteria such as modelling efficiency (E), coefficient of determination (R^2), index of agreement (IOA) and root mean square error (RMSE). These coefficients are independent of the scale of data used and are useful in assessing the goodness of fit of the model (Dawson & Wilby 1999). The peak flow conditions in the flow time series data need to be forecasted accurately because it is the main component of the flood hydrograph that causes damage. Therefore, another goodness of fit index called difference in peak flow (DP) was determined to evaluate the model performances. The mathematical expressions of the goodness of fit indices used in the study are presented in Table 5.

Table 5 | Mathematical expressions of goodness of fit indices (GFI)

Goodness of fit indices	Abbreviation	Mathematical expression
Index of agreement	IOA	$1 - \frac{\sum_{i=1}^n (O_i - P_i)^2}{\sum_{i=1}^n (P_i - \bar{O}_i + O_i - \bar{O}_i)^2}$
Coefficient of determination	R^2	$\left(\frac{\sum_{i=1}^n (O_i - \bar{O}_i)(P_i - \bar{P}_i)}{\sqrt{\sum_{i=1}^n (O_i - \bar{O}_i)^2 \sum_{i=1}^n (P_i - \bar{P}_i)^2}} \right)^2$
Modeling efficiency	E	$1 - \frac{\sum_{i=1}^n (O_i - \bar{O}_i)^2}{\sum_{i=1}^n (O_i - \bar{O}_i)^2}$
Root mean square error	RMSE	$\sqrt{\frac{1}{n} \sum_{i=1}^n (O_i - \bar{P}_i)^2}$
Difference in peak	DP	$D_p = \max(O_i) - \max(P_i)$

Table 6 | Performance indices of ANN and WN models

ANN and wavelet-ANN models	ANN structure	RMSE (m ³ /s)	Coefficient of determination (R ²)	Nash-sutcliffe coefficient (E)	Index of agreement (IOA)	Difference in peak (DP) m ³ /s
ANN-1	5-2-1	214.34	0.981	0.9281	0.9757	3,652
NW-1	3-2-1	69.76	0.994	0.9922	0.9981	149
ANN-2	5-2-1	490.95	0.768	0.5896	0.8505	5,229
NW-2	3-2-1	326.69	0.875	0.7597	0.9326	419
ANN-3	5-2-1	590.21	0.644	0.4071	0.7610	6,670
NW-3	3-2-1	398.74	0.761	0.5484	0.8635	471

**Figure 6** | (a) Comparison of the observed with the ANN and NW predicted stream flow for 1-day advance flow forecasting during the testing period and (b, c) scatter plot of the predicted stream flow values by ANN and NW models and observed values along with the best fit line.

RESULTS AND DISCUSSION

Results obtained from the ANN and NW models for 1-, 2- and 3-day advance flow forecasting at Jenapur site of

Brahmani basin are presented in this section. The goodness of fit indices were estimated for both ANN and NW models and were used to compare modelling performance of the models.

Comparison of ANN and NW models

The normalized datasets within the range of 0.2 and 0.8 were used for ANN training and testing for 1-, 2- and 3-day advance flow forecasting. One hidden layer was chosen to construct the optimized ANN architecture by changing the number of neurons in the hidden layer in a trial and error manner. Two hidden neurons were found to yield the best results in the case of all three models. The prediction accuracy of the models was found to decrease from 1-day forecast to 3-day forecasts in the case of both ANN and NW models.

It is observed that ANN predicted peak flow values were significantly under-predicted by an amount of $3,652 \text{ m}^3/\text{s}$; the under-prediction is very low ($149 \text{ m}^3/\text{s}$) in

the case of the NW model for 1-day forecasting. This promising result of the NW model may be due to the ability of DWT functions, which exploit the hidden properties of the flow data making it easier to capture the non-linearity in input-output datasets. The forecasting accuracy of NW models up to 3 days was found to be much better than that of corresponding ANN models. The superiority of the NW models can be attributed to the fact that NW used the highly conceivable wavelet coefficients as inputs in the ANN models, which ensured very good training and thereby predicting the peak flow conditions with greater accuracy. The performance of ANN-2 and ANN-3 are considerably poorer than that of NW-2 and NW-3, respectively, as highlighted in Table 6.

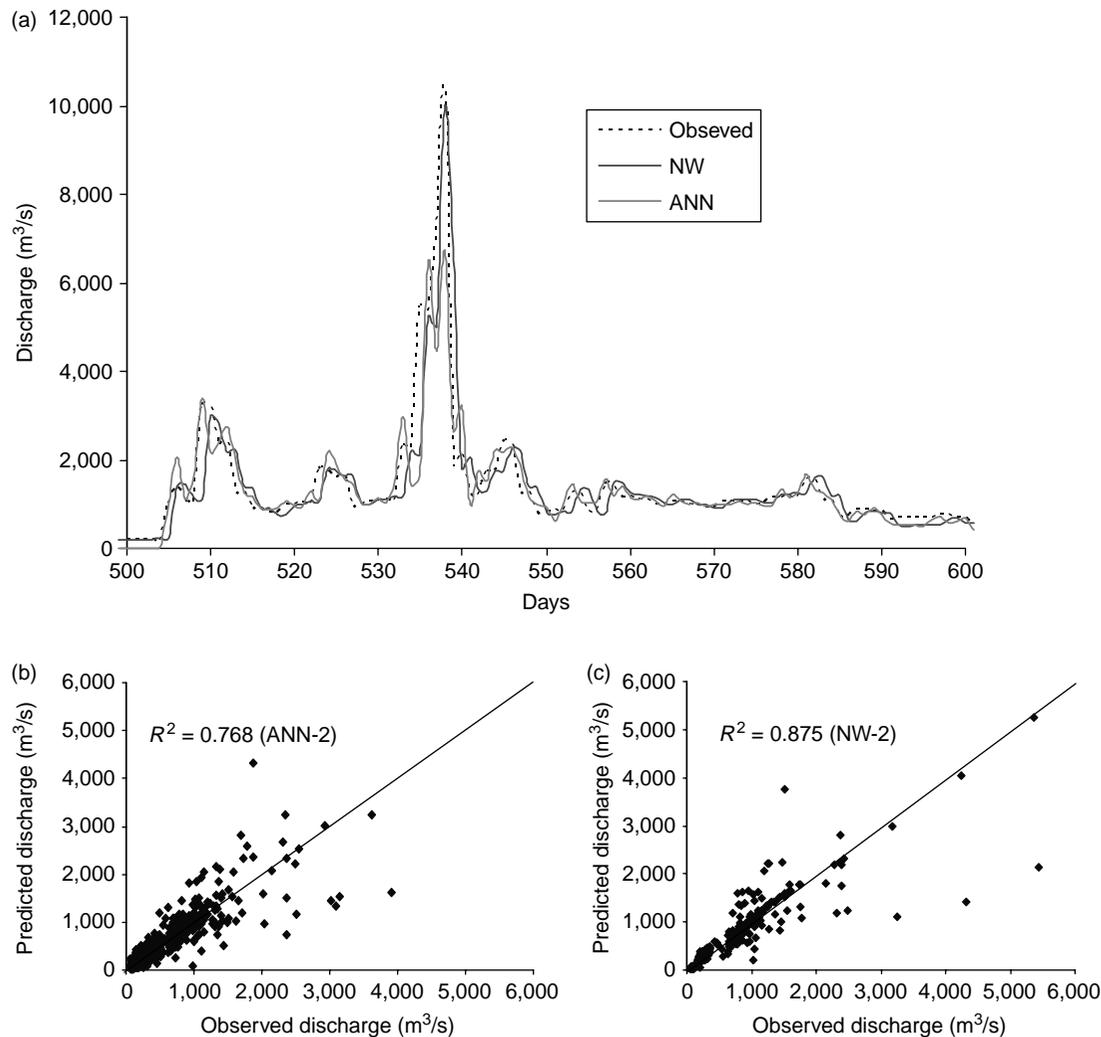


Figure 7 | (a) Comparison of the observed with the ANN and NW predicted stream flow for 2-day advance flow forecasting during the testing period and (b, c) scatter plot of the predicted stream flow values by ANN and NW models and observed values along with the best fit line.

The values of RMSE obtained from ANN-2 and ANN-3 are $490.95 \text{ m}^3/\text{s}$ and $590.21 \text{ m}^3/\text{s}$, respectively. The corresponding values in the case of NW-2 and NW-3 are $326.69 \text{ m}^3/\text{s}$ and $398.74 \text{ m}^3/\text{s}$, respectively. The estimated values of the RMSE, R^2 , E, IOA and DP obtained from ANN and NW models are listed in Table 6. The optimum numbers of hidden nodes were found to be two for both the types of models. It is observed that the result obtained from the hybrid model (NW-1) for 1-day forecasting is better than that of the other two hybrid models (NW-2 and NW-3) used for 2-day and 3-day forecasting. However, the performances of NW-2 and NW-3 models were found

satisfactory in terms of their abilities to yield lower DP than that of their respective ANN models.

The 1-day ahead forecasted stream flow hydrographs by ANN-1 and NW-1 models were compared with the observed stream flow values and are presented in Figure 6(a) during the testing period. It is observed from the hydrographs (Figure 6(a)) that the NW-1 model predicted flow values much closer to the corresponding observed flow values as compared to ANN-1 model. The observed data points, along with the ANN-1 and NW-1 predicted stream flow values, are depicted in Figure 6(b,c) in the form of scatter diagrams. It is observed from Figure 6(c) that the

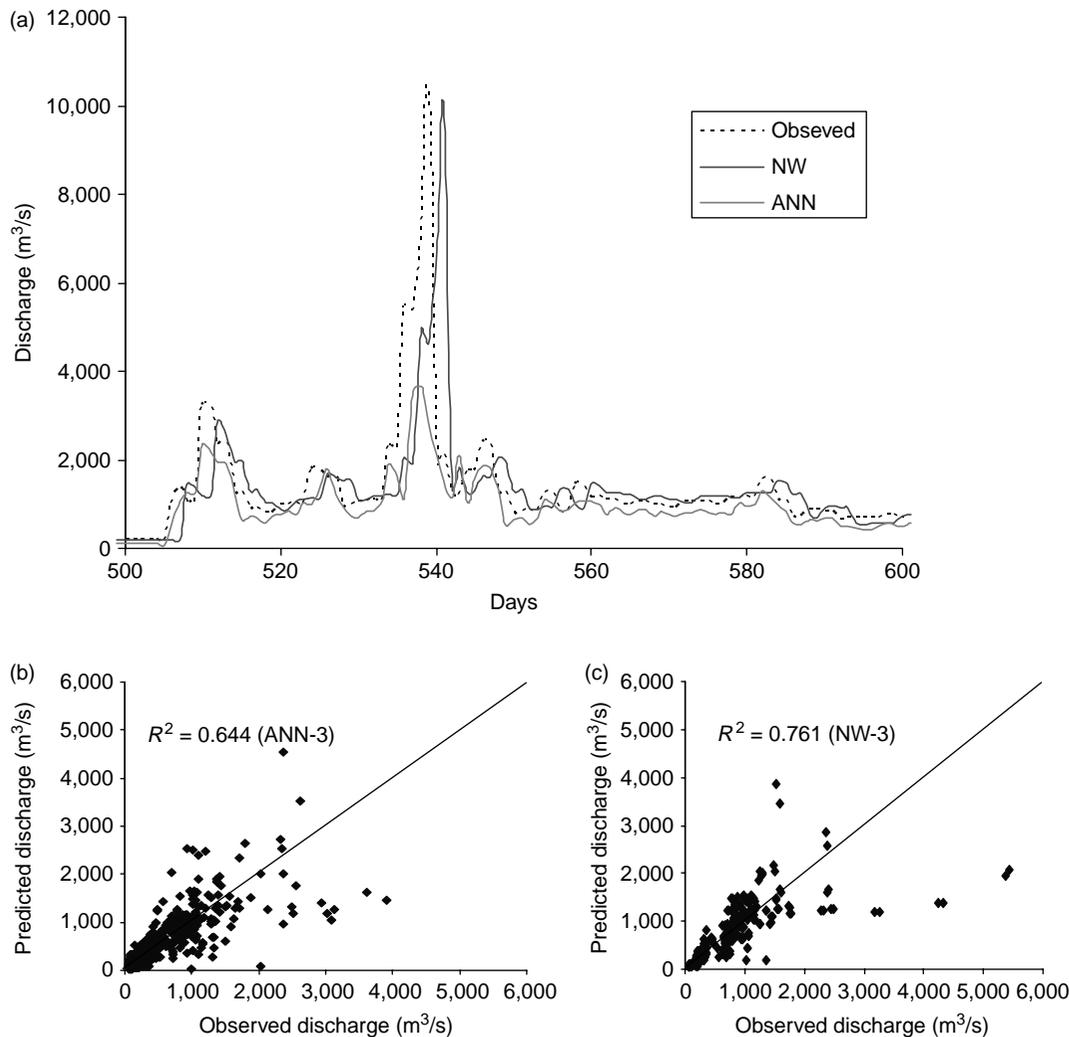


Figure 8 | (a) Comparison of the observed with the ANN and NW predicted stream flow for 3-day advance flow forecasting during the testing period and (b, c) scatter plot of the predicted stream flow values by ANN and NW models and observed values along with the best fit line.

estimated peak flows in the case of the NW-1 model are closely placed around the best fit line.

Figure 7 presents the hydrographs and the scatter plot between the observed and modelled flow values for 2-day advance forecasting during the testing period. It is observed that the NW-2 forecasted flow values were found to be in close agreement with the observed values. The performance of the NW-2 model was found to be satisfactory as compared to the ANN-2 model, which is revealed from its estimated peak flow values presented in Table 6.

Similarly, the ANN and NW predicted hydrographs and the scatter plot for 3-day advance forecasting were computed and are presented in Figure 8 for the testing period. Although the performances of the NW-3 were found to be lower than for the NW-1 and NW-2 models, the prediction accuracy was found to be superior to the corresponding ANN models.

In the case of all the models (ANNs and NWs), the peaks as well as some high-magnitude flow values were found to be considerably under-estimated whereas low flow conditions were estimated fairly accurate. This under-estimation is found more in the case of ANN models than NW models. The reason for under-prediction could be attributed to the fact that there are less numbers of high magnitude flow values in the training datasets, which biased the training process and forced the architecture to map the low flow situations accurately. However, the degree of under-prediction is low in the case of NW models, which ensured a close agreement between the observed and modelled outputs. These performances of NW models are due the non-linear property of the DWT technique, which scaled the original datasets into a series of wavelet coefficients of varying magnitude in a regular fashion. This property of the DWT technique avoids biased training, thereby enhancing good prediction for all flow situations.

CONCLUSIONS

This paper provides a thorough review of the application of different data-driven methods applied for modelling wide range of problems in the hydrological system, including rainfall-runoff modelling and stream flow forecasting. The study concludes that the forecasting abilities of the ANN

models are found to be improved when the wavelet transformation technique is adopted for data pre-processing. The decomposed periodic components obtained from the DWT technique are found to be most effective in yielding accurate forecast when used as inputs in the ANN models.

It is concluded that the wavelet-ANN hybrid models are capable of approximating the nonlinear relationship between the inputs and output much better than simple ANN models. The peaks in the time series of the flow data can be estimated more accurately using the wavelet-ANN approach than that of simple ANN models, making the DWT technique very efficient and useful in hydrologic forecasting. The approach used in this study may be similar to the few applications made in the past, but the study is original in terms of study area, data used and the results obtained. The results of the study will be useful in controlling flood impact through flood warning.

The forecasting accuracy could be further enhanced if additional variables such as river stage and rainfall on the basin could be considered for wavelet analysis and the respective effective components would be used as inputs to the NW models.

REFERENCES

- Adamowski, J. F. 2008 Development of a short-term stream flood forecasting method for snowmelt driven floods based on wavelet and cross-wavelet analysis. *J. Hydrol.* **353**, 247–266.
- Anctil, F. & Tape, D. G. 2004 An exploration of artificial neural network rainfall-runoff forecasting combined with wavelet decomposition. *J. Environ. Eng. Sci.* **3**, 21–128.
- Aqil, M., Kita, I., Yano, A. & Nishiyama, S. 2007 A comparative study of artificial neural networks and neuro-fuzzy in continuous modeling of the daily and hourly behaviour of runoff. *J. Hydrol.* **337**, 22–34.
- ASCE Task Committee 2000a Artificial neural networks in hydrology I: preliminary concepts. *J. Hydrol. Eng.* **5** (2), 115–123.
- ASCE Task Committee 2000b Artificial neural networks in hydrology II: hydrologic application. *J. Hydrol. Eng.* **5** (2), 124–137.
- Aussem, A., Campbell, J. & Murtagh, F. 1998 Wavelet-based feature extraction and decomposition strategies for financial forecasting. *J. Comput. Intell. Finance* **6**, 5–12.
- Bray, M. & Han, D. 2004 Identification of support vector machines for runoff modeling. *J. Hydroinformatics* **6** (4), 265–280.
- Cannas, B., Fanni, A., Sias, G., Tronci, S. & Zedda, M. K. 2005 Stream flow forecasting using neural networks and wavelet analysis. *Eur. Geosci. Union* **7**, 45–51.

- Cannas, B., Fanni, A., See, L. & Sias, G. 2006 Data pre-processing for river flow forecasting using neural networks: wavelet transforms and data partitioning. *Phys. Chem. Earth* **31** (18), 1164–1171.
- Chau, K. W. 2006 Particle swarm optimization training algorithm for ANNs in stage prediction of Shing Mun River. *J. Hydrol.* **329** (3–4), 363–367.
- Chau, K. W., Wu, C. L. & Li, Y. S. 2005 Comparison of several flood forecasting models in Yangtze river. *J. Hydrol. Eng.* **10**, 485–491.
- Chou, C.-M. & Wang, R.-Y. 2004 Application of wavelet-based multi-model Kalman filters to real-time flood forecasting. *Hydrol. Process.* **18**, 987–1008.
- Coulibaly, P., Anctil, F. & Bobee, B. 2000 Daily reservoir inflow forecasting using artificial neural networks with stopped training approach. *J. Hydrol.* **230**, 244–257.
- Dawson, C. W. & Wilby, R. L. 1999 A comparison of artificial neural networks used for river forecasting. *Hydrol. Earth Sys. Sci.* **3** (4), 529–540.
- Dawson, C. W. & Wilby, R. L. 2001 Hydrological modelling using artificial neural networks. *Prog. Phys. Geogr.* **25**, 80–108.
- Draper, N. R. & Smith, H. 1998 *Applied Regression Analysis*. John Wiley and Sons, New York, USA.
- Evora, N. D. & Coulibaly, P. 2009 Recent advances in data-driven modeling of remote sensing applications in hydrology. *J. Hydroinformatics* **11** (3–4), 194–201.
- Gautam, M. R., Watanabe, K. & Saegusa, H. 2000 Runoff analysis in humid forest catchments with artificial neural networks. *J. Hydrol.* **235**, 117–136.
- Giustolisi, O. & Savic, D. A. 2004 A novel genetic programming strategy: evolutionary polynomial regression. In *Proceedings of Hydroinformatics 2004* (ed. S. Y. Liong, X. Phoon & V. Babovic), Vol. 1, pp. 787–794. World Scientific Publishing Company, Singapore.
- Giustolisi, O. & Savic, D. A. 2006 A symbolic data-driven technique based on evolutionary polynomial regression. *J. Hydroinformatics* **8** (3), 207–222.
- Giustolisi, O. & Savic, D. A. 2009 Advances in data-driven analyses and modeling using EPR-MOGA. *J. Hydroinformatics* **11** (3–4), 225–236.
- Gunn, S. R. 1998 Support vector machines for classification and regression. ISIS Tech. Report. ISIS-1-98, Image Speech Intelligent System Research Group, University of Southampton, Southampton, UK.
- Han, D., Kwong, T. & Li, S. 2006 Uncertainties in real-time flood forecasting with neural networks. *Hydrol. Process.* **21** (2), 223–228.
- Han, D., Chan, L. & Zhu, N. 2007 Flood forecasting using support vector machines. *J. Hydroinformatics* **9** (4), 267–276.
- Haykin, S. 1994 *Neural Networks: A Comprehensive Foundation*. Macmillan, New York, USA.
- Hsu, K. L., Gupta, H. V. & Sorooshian, S. 1995 Artificial neural network modeling of the rainfall-runoff process. *Water Resour. Res.* **31**, 2517–2530.
- Jacquín, A. P. & Shamseldin, A. Y. 2009 Review of the application of fuzzy inference systems in river flow forecasting. *J. Hydroinformatics* **11** (3–4), 202–210.
- Jain, S. K., Agarwal, P. K. & Singh, V. P. 2007 *Hydrology and Water Resources of India*. Springer, The Netherlands.
- Karunanithi, N., Grenney, W. J., Whitley, D. & Bovee, K. 1994 Neural networks for stream flow prediction. *J. Comput. Civ. Eng.* **8** (2), 201–220.
- Kim, T. W. & Valdes, J. B. 2003 Nonlinear model for drought forecasting based on a conjunction of wavelet transforms and neural networks. *J. Hydrol. Eng.* **8** (6), 319–328.
- Kishor, N. & Singh, S. P. 2007 Nonlinear predictive control for a NNARX hydro plant model. *Neural Comput. Appl.* **16** (2), 101–108.
- Kisi, O. 2004 River flow modeling using artificial neural networks. *J. Hydrol. Eng.* **9** (1), 60–63.
- Kisi, O. 2007 Streamflow forecasting using different artificial neural network algorithms. *J. Hydrol. Eng.* **12** (5), 532–539.
- Kisi, O. 2008a River flow forecasting and estimation using different artificial neural network techniques. *Hydrol. Res.* **39** (1), 27–40.
- Kisi, O. 2008b Stream flow forecasting using neuro-wavelet technique. *Hydrol. Process.* **22** (20), 4142–4152.
- Kisi, O. 2009a Neural network and wavelet conjunction model for modeling monthly level fluctuations of Van Lake in Turkey. *Hydrol. Process.* **23** (14), 2081–2092.
- Kisi, O. 2009b Neural networks and wavelet conjunction model for intermittent streamflow forecasting. *J. Hydrol. Eng.* **14** (8), 773–782.
- Kisi, O. & Cigizoglu, H. K. 2007 Comparison of different ANN techniques in river flow prediction. *Civ. Eng. Environ. Syst.* **24** (3), 211–231.
- Koza, J. R. 1992 *Genetic Programming: On the Programming of Computers by Means of Natural Selection*. MIT Press, Cambridge, MA, USA.
- Kumar, A. R. S., Sudheer, K. P., Jain, S. K. & Agarwal, P. K. 2005 Rainfall-runoff modeling using artificial neural networks: comparison of network types. *Hydrol. Process.* **19**, 1277–1291.
- Labat, D. 2005 Recent advances in wavelet analyses: part 1. A review of concepts. *J. Hydrol.* **314** (1–4), 275–288.
- Labat, D., Ababou, R. & Mangin, A. 1999 Wavelet analysis in Karstic hydrology 2nd part: rainfall-runoff cross-wavelet analysis. *C. R. Acad. Sci. IIA Earth Planet. Sci.* **329**, 881–887.
- Labat, D., Ababou, R. & Mangin, A. 2000 Rainfall-runoff relationships for karstic springs. part II: continuous wavelet and discrete orthogonal multiresolution analyses. *J. Hydrol.* **238**, 149–178.
- Liong, C. & Shizhong, J. 2006 Wavelet characteristics of hydrologic time series in upstream of the Longchuan stream. *J. Am. Sci.* **2** (2), 82–87.
- Maier, H. R. & Dandy, G. C. 2000 Neural networks for the prediction and forecasting of water resources variables: a review of modelling issues and applications. *Environ. Model. Softw.* **15**, 101–124.
- Mallat, S. G. 1998 *A Wavelet Tour of Signal Processing*, p. 577. Academic, San Diego.

- Mukerji, A., Chatterjee, C. & Raghuwanshi, N. S. 2009 Flood forecasting using ANN, Neuro-Fuzzy, and Neuro-GA Models. *J. Hydrol. Eng.* **14** (6), 647–652.
- Nayak, P. C., Sudheer, K. P. & Ramasastri, K. S. 2005 Fuzzy computing based rainfall-runoff model for real time flood forecasting. *Hydrol. Process.* **19**, 955–958.
- Nayak, P. C., Sudheer, K. P. & Jain, S. K. 2007 Rainfall-runoff modeling through hybrid intelligent system. *Water Resour. Res.* **43**, 1–17.
- Nourani, V., Alami, M. T. & Aminfar, M. H. 2008 A combined neural-wavelet model for prediction of Ligvanchai watershed precipitation. *Eng. Appl. Artif. Intell.* **22** (3), 466–472.
- Partal, T. & Kisi, O. 2007 Wavelet and Neuro-fuzzy conjunction model for precipitation forecasting. *J. Hydrol.* **342** (1–2), 199–212.
- Pramanik, N. & Panda, R. K. 2009 Application of neural network and adaptive neuro-fuzzy inference systems for stream flow prediction. *Hydrol. Sci. J.* **54** (2), 247–260.
- Priddy, L. K. & Keller, E. P. 2005 *Artificial Neural Networks, an Introduction*. SPIE Press, Bellingham, Washington.
- Rao, R. M. & Bopardikar, A. S. 1998 *Wavelet Transforms: Introduction to Theory and Applications*, p. 310. Addison-Wesley, MA.
- Remesan, R., Shamim, M. A., Han, D. & Mathew, J. 2009 Runoff prediction using an integrated hybrid modelling scheme. *J. Hydrol.* **372** (1–4), 48–60.
- Saco, P. & Kumar, P. 2000 Coherent modes in multiscale variability of streamflow over the united states. *Water Resour. Res.* **36**, 1049–1068.
- Sahoo, G. B. & Ray, C. 2006 Flow forecasting for a Hawaii stream using rating curves and neural networks. *J. Hydrol.* **317**, 63–80.
- Shanker, M., Hu, M. Y. & Hung, M. S. 1996 Effect of data normalization on neural network training. *Int. J. Manage. Sci.* **24**, 385–397.
- Shrestha, R. R., Theobald, S. & Nestmann, F. 2005 Simulation of flood flow in a stream system using artificial neural networks. *Hydrol. Earth Syst. Sci.* **9** (4), 313–321.
- Singh, A., Panda, R. K. & Pramanik, N. 2009 Appropriate data normalization range for daily river flow forecasting using an artificial neural network. In *Hydroinformatics in Hydrology, Hydrogeology and Water Resources 2009* (ed. I. D. Cluckie, Y. Chen, V. Babovic, L. Konikow, A. Mynett, S. Demuth & D. Savic), Vol. 331, pp. 51–57. IAHS Publication, UK.
- Smith, L. C., Turcotte, D. L. & Isacks, B. L. 1998 Stream flow characterization and feature detection using a discrete wavelet transform. *Hydrol. Process.* **12**, 233–249.
- Sudheer, K. P., Gosain, A. K. & Ramasastri, K. S. 2002 A data-driven algorithm for constructing artificial neural network rainfall-runoff models. *Hydrol. Process.* **16**, 1325–1330.
- Tokar, A. S. & Johnson, P. A. 1999 Rainfall-runoff modeling using artificial neural networks. *J. Hydrol. Eng.* **4** (3), 232–239.
- Torrence, C. & Compo, G. 1998 A practical guide to wavelet analysis. *Bull. Am. Meteorol. Soc.* **79** (1), 61–78.
- Wang, D. & Ding, J. 2003 Wavelet network model and its application to the prediction of hydrology. *Nat. Sci.* **1**, 67–71.
- Wilby, R. L., Abrahart, R. J. & Dawson, C. W. 2003 Detection of conceptual model rainfall-runoff processes inside an artificial neural network. *Hydrol. Sci. J.* **48** (2), 163–181.
- Zealand, C. M., Burn, D. H. & Simonovic, S. P. 1999 Short term streamflow forecasting using artificial neural networks. *J. Hydrol.* **214**, 32–48.
- Zhang, B. L. & Dong, Z. Y. 2001 An adaptive neural-wavelet model for short term load forecasting. *Electric Power Syst. Res.* **59** (2), 121–129.
- Zheng, T., Girgis, A. A. & Makram, E. B. 2000 A hybrid wavelet-Kalman filter method for load forecasting. *Electric Power Syst. Res.* **54**, 11–17.
- Zhou, H., Wu, L. & Guo, Y. 2006 Mid and long term hydrologic forecasting for drainage are based on NNNW and FRM. In *Sixth International Conference on Intelligent Systems Design and Applications*, pp. 7–12. ISDA.

First received 24 June 2009; accepted in revised form 25 September 2009. Available online 9 April 2010