New Approaches to Hydrological Prediction in Data-sparse Regions (Proc. of Symposium HS.2 at the Joint IAHS & IAH Convention, Hyderabad, India, September 2009). IAHS Publ. 333, 2009.

Modelling hydrological time series data using wavelet neural network analysis

Y. R. SATYAJI RAO & B. KRISHNA

Deltaic Regional centre, National Institute of Hydrology, Siddartha Nagar, Kakinada-3, Andhra Pradesh, India
yrs_rao@yahoo.com

Abstract Time series analysis requires mapping complex relationships between input(s) and output(s), because the forecasted values are mapped as a function of patterns observed in the past. In order to improve the precision of the forecasts, a Wavelet Neural Network (WNN) model, based on a combination of wavelet analysis and Artificial Neural Network (ANN), has been proposed. The WNN and ANN models have been applied to daily streamflow and monthly groundwater levels series where there is a scarcity of other hydrological time series data. The calibration and validation performance of the models is evaluated with appropriate statistical indices. The results of daily streamflow and monthly groundwater level series modelling indicated that the performances of WNN models are more effective than the ANN models. This paper also highlights the capability of WNN models in estimating low and high values in the hydrological time series data.

Key words time series; ANN; WNN; streamflow; groundwater levels; India

INTRODUCTION

Time series modelling of hydrological variables is of utmost important in the planning and management of water resources. Most time series modelling procedures fall within the framework of multivariate Auto Regressive Moving Average (ARMA) models. Traditionally, ARMA models have been widely used for modelling water resources time-series modelling (Maier *et al*., 1997). Time-series models are more practical than conceptual models because it is not essential to understand the internal structure of the physical processes that are taking place in the system being modelled. However, these models do not attempt to represent the nonlinear dynamics inherent in the hydrological process, and may not always perform well (Tokar *et al.*, 1999). Presently nonlinear models such as neural networks are widely used for time series modelling. Artificial Neural Networks (ANN) models were widely used to overcome many difficulties in time series modelling of hydrological variables (ASCE Task Committee, 2000a,b). However, it is also reported that ANN models are not very satisfactory in terms of precision because they consider only a few aspects of the behaviour of the time series (Wensheng, 2003). To raise the precision, a wavelet analysis has been used along with ANN. The advantage of the wavelet technique is that it provides a mathematical process for decomposing a signal into multiple levels of details and analysis. Wavelet analysis can effectively decompose signals into the main frequency components while also extracting local information of the time series. In recent years, wavelet theory has been introduced in the field of hydrology (Smith *et al*., 1998; Labat *et al*., 2000). Due to the similarity between wavelet decomposition and the single-hidden layer neural network, the idea of combining both wavelet and neural networks has resulted recently in formulation of the wavelet neural network, which has been used in various fields (Xiao *et al*., 2005).

In this paper, a Wavelet Neural Network (WNN) model, which is the combination of wavelet analysis and ANN, has been proposed for time series modelling of four west-flowing rivers in India: the Kollur, Seethanadi, Varahi and Gowrihole rivers, and for simulating groundwater levels in three observation wells in the Central Godavari Delta, east coast of India, where the availability of other hydrological time series data is limited.

WAVELET ANALYSIS

Wavelet analysis involves the decomposition of a signal into shifted and scaled versions of the original (or mother) wavelet. In wavelet analysis, the use of a fully scalable modulated window solves the signal-cutting problem. The window is shifted along the signal and for every position the spectrum is calculated. Then this process is repeated many times with a slightly shorter (or longer) window for every new cycle. In the end, the result will be a collection of time-frequency representations of the signal, all with different resolutions. Because of this collection of representations, it can be called a multi-resolution analysis. By decomposing a time series into time–frequency space, one is able to determine both the dominant modes of variability and how those modes vary in time. Wavelets have proven to be a powerful tool for the analysis and synthesis of data from long memory processes. Wavelets are strongly connected to such processes in that the same shapes repeat at different orders of magnitude. The ability of the wavelets to simultaneously localize a process in a time and scale domain results in representation of many dense matrices in a sparse form.

The discrete wavelet transform of a time series $f(t)$ is defined as:

$$f(a,b) = \frac{1}{\sqrt{a}} \int_{-\infty}^{\infty} f(t)\psi\left(\frac{t-b}{a}\right)dt \quad (1)$$

where $\psi(t)$ is the basic wavelet with effective length (t) that is usually much shorter than the target time series $f(t)$; a is the scale or dilation factor that determines the characteristic frequency so that its variation gives rise to a "spectrum"; and b is the translation in time so that its variation represents the "sliding" of the wavelet over $f(t)$. The wavelet spectrum is thus customarily displayed in the time-frequency domain. For low scales, i.e. when $|a| << 1$, the wavelet function is very concentrated (shrunk, compressed) with frequency content mostly in the higher frequency bands. Inversely, when $|a| >> 1$, the wavelet is stretched and contains mostly low frequencies. For small scales we obtain thus a more detailed view of the signal (also known as "higher resolution"), whereas for larger scales we obtain a more general view of the signal structure.

The original signal $X(n)$ passes through two complementary filters (low pass and high pass filters) and emerges as two signals: approximations (A) and details (D). The approximations are the high-scale, low frequency components of the signal. The details are the low-scale, high frequency components. Normally, the low frequency content of the signal (approximation, A) is the most important. It demonstrates the signal identity. The high-frequency component (detail, D) is nuance. The decomposition process can be iterated, with successive approximations being decomposed in turn, so that one signal is broken down into many lower resolution components (Fig. 1).

ARTIFICIAL NEURAL NETWORKS

An ANN can be defined as a system or mathematical model consisting of many nonlinear artificial neurons running in parallel, which can be generated as single or multiple layered. Although the concept of artificial neurons was first introduced by McCulloch & Pitts (1943), the major applications of ANNs have arisen only since the development of the back-propagation (BP) method of training by Rumelhart *et al*. (1986). Following this development, ANN research has resulted in the successful solution of some complicated problems not easily solved by traditional modelling methods when the quality/quantity of data is very limited. ANN models are "black box" models with particular properties, which are greatly suited to dynamic nonlinear system modelling. The main advantage of this approach over traditional methods is that it does not require the complex nature of the underlying process under consideration to be explicitly described in mathematical form. ANN applications in hydrology vary, from real-time to event-based modelling.

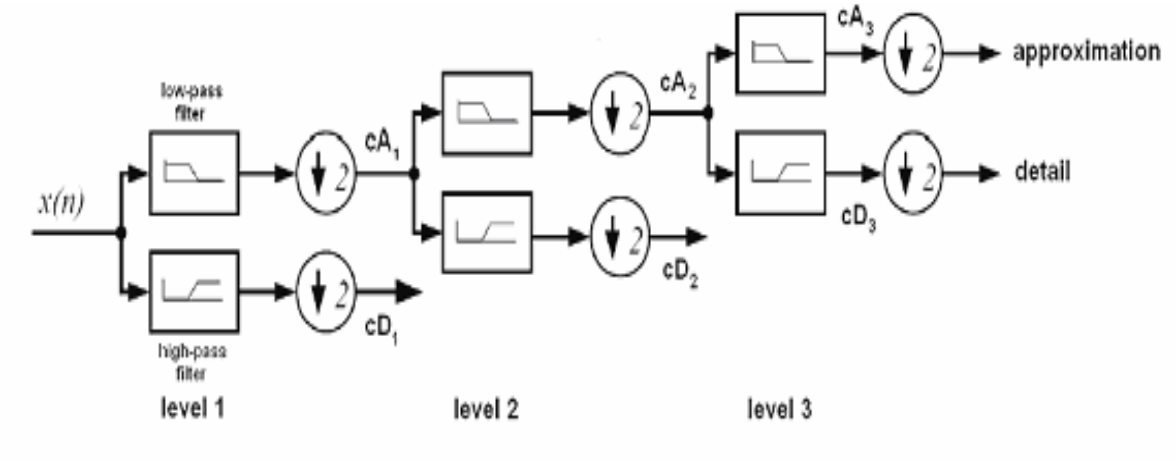


Fig. 1 Diagram of multi-resolution analysis of signal.

Neural network structure

The most popular ANN architecture in hydrological modelling is the multilayer perceptron (MLP) trained with a BP algorithm (ASCE, 2000a,b). A multilayer perceptron network consists of an input layer, one or more hidden layers of computation nodes, and an output layer. The number of input and output nodes is determined by the nature of the actual input and output variables. The number of hidden nodes, however, depends on the complexity of the mathematical nature of the problem, and is determined by the modeller, often by trial and error. The input signal propagates through the network in a forward direction, layer by layer. Each hidden and output node processes its input by multiplying each of its input values by a weight, summing the product and then passing the sum through a nonlinear transfer function to produce a result. For the training process, where weights are selected, the neural network uses the gradient descent method to modify the randomly selected weights of the nodes in response to the errors between the actual output values and the target values. This process is referred to as training or learning. It stops when the errors are minimized or another stopping criterion is met. The BPNN can be expressed as:

$$Y = f(\Sigma W X - \theta) \tag{2}$$

where X is the input or hidden node value; Y is the output value of the hidden or output node; $f()$ is the transfer function; W is weights connecting the input to hidden, or hidden to output nodes; and θ is the bias (or threshold) for each node.

Methods of network training

In the current investigation, the Levenberg-Marquardt (LM) method was used for training the given network. LM is a modification of the classic Newton algorithm for finding an optimum solution to a minimization problem. The algorithm uses the second-order derivatives of the function so that better convergence behaviour is observed. In the ordinary gradient descent method, only the first-order derivatives are evaluated and the parameter change information is contained solely in the direction along which the cost is minimized. In practice, LM is faster and finds better optima for a variety of problems than most other methods (Hagan & Menhaj, 1994). The method also takes advantage of the internal recurrence to dynamically incorporate past experience in the training process (Coulibaly *et al*., 2000).

The Levenberg-Marquardt algorithm is given by:

$$X_{k+1} = X_k - (J^T J + \mu I)^{-1} J^T e \tag{3}$$

where, X contains the weights of the neural network, J is the Jacobian matrix of the performance criteria to be minimized, μ is a learning rate that controls the learning process and e is residual error vector.

If the scalar μ is very large, the above expression approximates gradient descent with a small step size; while if it is very small; the above expression becomes the Gauss-Newton method using the approximate Hessian matrix. The Gauss-Newton method is faster and more accurate near an error minimum. Hence we decrease μ after each successful step and increase it only when a step increases the error. Levenberg-Marquardt has large computational and memory requirements, and thus it can only be used in small networks (Maier & Dandy, 1998). However, it is faster and less easily trapped in local minima than other optimization algorithms (Coulibaly, 2001a,b,c; Toth *et al*., 2000).

Selection of network architecture

Based on a physical knowledge of the problem and statistical analysis, different combinations of antecedent values of the time series were considered as input nodes. The output node is the time series data to be predicted in one step ahead. Time series data was standardized for zero mean and unit variation, and then normalized into the range [0 to 1]. The activation functions used for the hidden and output layer were logarithmic sigmoidal and pure linear function, respectively. For deciding the optimal hidden neurons, a trial and error procedure was started with two hidden neurons initially, and the number of hidden neurons was increased up to 10 with a step size of 1 in each trial. For each set of hidden neurons, the network was trained in batch mode to minimize the mean square error at the output layer. To check for any over-fitting during training, a cross-validation was performed by keeping track of the efficiency of the fitted model. The training was stopped when there was no significant improvement in the efficiency, and the model was then tested for its generalization properties. Figure 2 shows the multilayer perceptron neural network architecture when the original signal is used as the input to the neural network architecture.

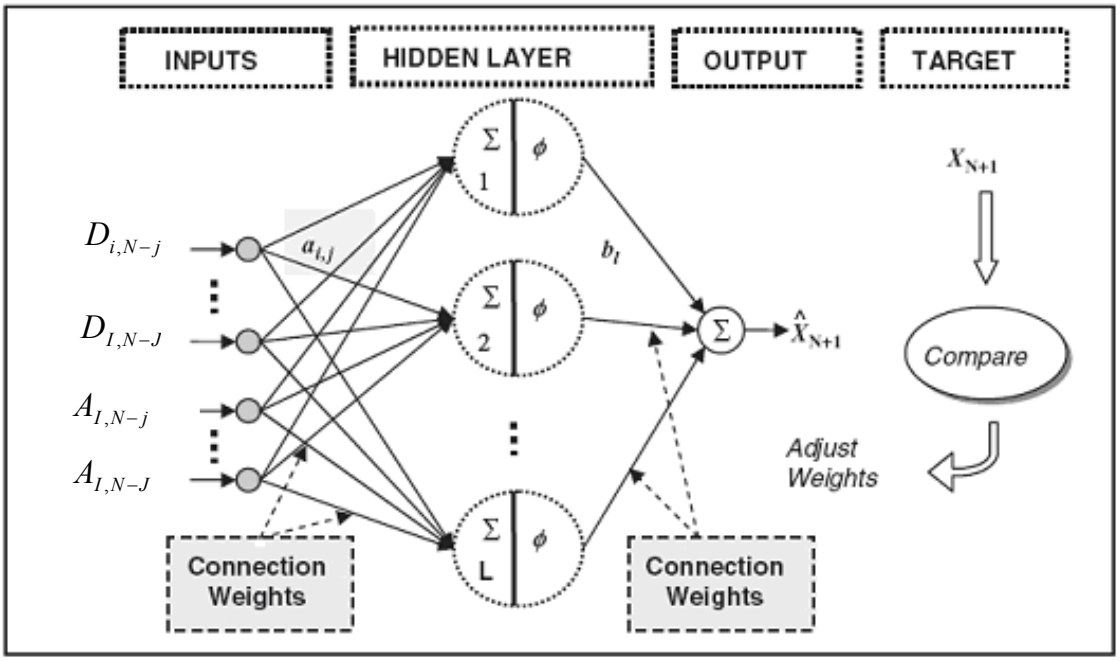


Fig. 2 Scatter plots between observed and modelled stream flow in different basins.

METHOD OF COMBINING THE WAVELET ANALYSIS WITH ANN

The decomposed details (D) and approximation (A) are used as inputs to the neural network structure as shown in Fig. 1. Here, *i* is the level of decomposition varying from 1 to *I*, *j* is the number of antecedent values varying from 0 to *J*, and *N* is the length of the time series. To obtain

the optimal weights (parameters) of the neural network structure, the LM algorithm is used to train the network. The output node represents the original value at one step ahead.

PERFORMANCE CRITERIA

The performance of various models (WNN and ANN) during calibration and validation were evaluated by using statistical indices: the root mean squared error (RMSE), correlation coefficient (R), and the Nash and Sutcliffe coefficient of efficiency (COE). RMSE measures the residual variance; which indicates a quantitative measure of the model error in units of the variable; the optimal value is 0. The correlation coefficient (R) measures the linear correlation between the observed and modelled values; the optimal value is 1.0. The COE, or efficiency (E%) compares the modelled and observed values and evaluates how far the network is able to explain total variance of data; the optimal value of COE is 1.0.

STUDY AREA AND DATA

Streamflows

The Western Ghats form a range of mountains in peninsular India running approximately parallel to the west coast and are home to the largest tracts of moist tropical forest in India. The study area, is between latitudes 12°N to 14°N and longitude 74°E to 76°E. The basins and gauging stations considered for this study are: Kollur at Jadkal (108 km^2), Sitanadi at Kokkarne (343 km^2), Varahi at Dasanakatte (135 km^2) and Gowrihole at Sarve Bridge (126 km^2). These streamgauges are maintained by the Water Resources Development Organization (WRDO), Govt of Karnataka. The annual rainfall is as high as 4000 mm; the basins are characterized by steep gradients and good forest cover. The bulk of the rainfall occurs during the monsoon season (June–September), which accounts for about 80% of the annual total. The daily discharge data for Kollur, Sitanadi, Varahi and Gowrihole basins for period of 1981–2002 (22 years), 1973–1998 (26 years), 1978–2003 (26 years), and 1979–2003 (25 years), respectively, are considered for the study. Due to geographical conditions the measurement of daily rainfall and other climatic variables is difficult in densely forested basins and the availability of historical data is very limited.

Groundwater levels

The study area forms a part of the River Godavari delta system in the East Godavari District of Andhra Pradesh, India. Geographically the study area, Central Godavari Delta, is located between 16°25′N to 16°55′N latitude and 81°44′E to 82°15′E longitude with its hydrological boundaries the River Gowthami Godavari in the east, the River Vasistha Godavari in the west and the Bay of Bengal to the south. The total geographical area is 825 km^2. The study area receives more than half of its annual rainfall during the southwest monsoon (i.e. June–September), while a large portion of the rest occurs in the month of October. In the study area the groundwater is under water table conditions and with dense canal network systems. The groundwater utilization for irrigation purposes is very limited. The information on historical monthly groundwater draft data and monthly canal seepage data into groundwater is very limited. The representative groundwater levels are from three observation wells in the Central Godavari Delta: Kattunga, Munganda and Cheyyeru, which have long records of groundwater level data (1977–1994).

RESULTS AND DISCUSSION

Streamflows

The daily discharge data of four west-flowing river basins, namely Kollur, Sithanadi, Varahi and Gowrihole for the total period of 1981–2002 (22 years), 1973–1998 (26 years), 1978–2003

Table 1 Model inputs for WNN model.

Model I	Q(t)=f (x[t–1])
Model II	Q(t)=f (x[t–1], x[t–2])
Model III	Q(t)=f (x[t–1], x[t–2], x[t–3])
Model IV	Q(t)=f (x[t–1], x[t–2], x[t–3], x[t–4])
Model V	Q(t)=f (x[t–1], x[t–2], x[t–3], x[t–4], x[t–5])

Q(t) is stream flow and groundwater level; f(x( )) is decomposed series of streamflow and groundwater level.

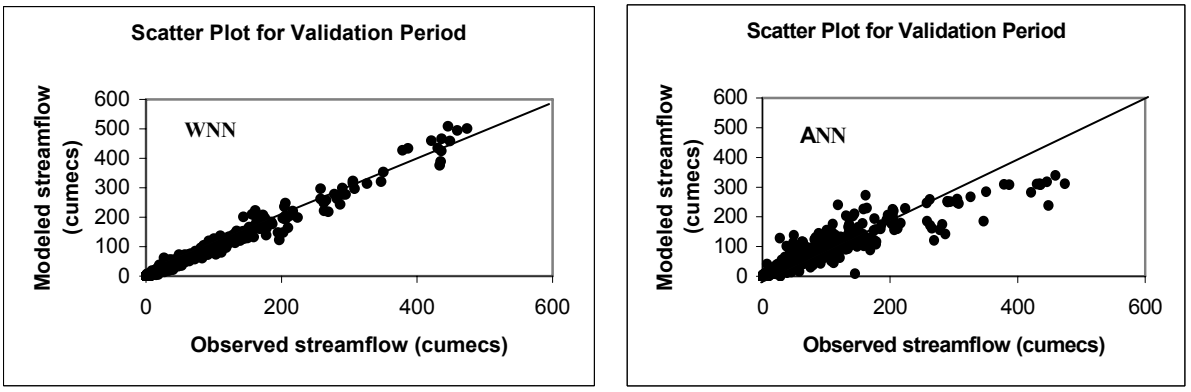
Table 2 Model inputs for ANN model.

Model I	Q(t)=f (Q[t–1])
Model II	Q(t)=f (Q[t–1], Q[t–2])
Model III	Q(t)=f (Q[t–1], Q[t–2], Q[t–3])
Model IV	Q(t)=f (Q[t–1], Q[t–2], Q[t–3], Q[t–4])
Model V	Q(t)=f (Q[t–1], Q[t–2], Q[t–3], Q[t–4], Q[t–5])

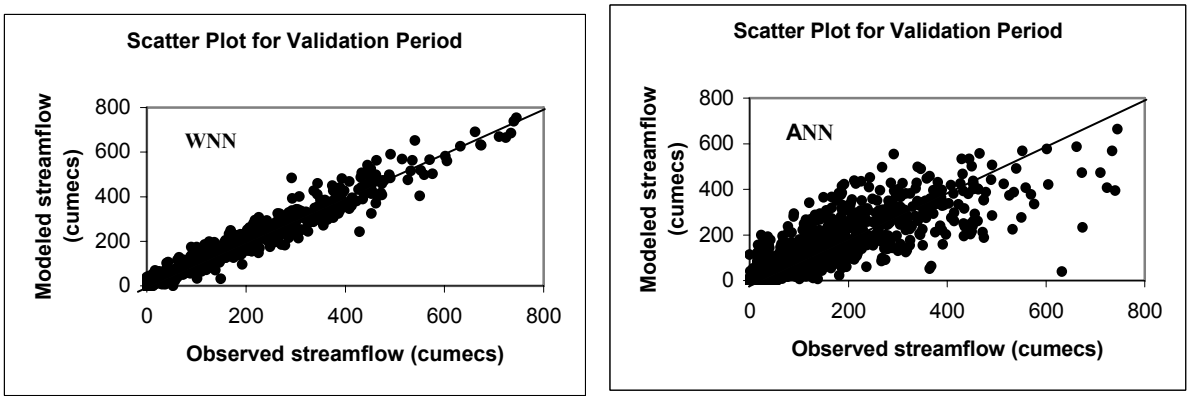
Q(t) is streamflow and groundwater level.

Table 3 Performance of WNN and ANN models during calibration and validation.

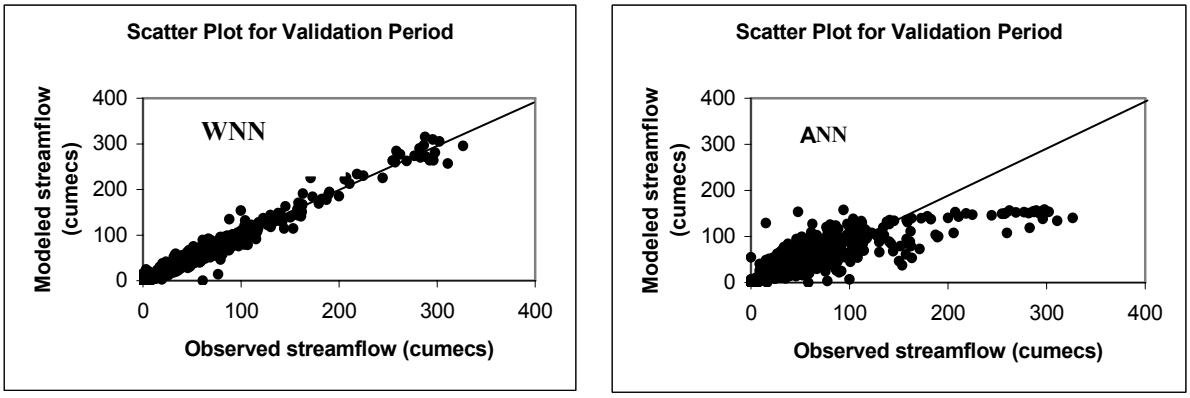
Station	Model	Calibration:			Validation:		
		RMSE	R	E (%)	RMSE	R	E (%)
Jadkal	WNN I	11.36	0.9527	90.76	12.08	0.9703	94.14
	II	5.69	0.9883	97.68	7.01	0.9907	98.03
	III	5.07	0.9907	98.16	6.80	0.9908	98.14
	IV	**4.37**	**0.9931**	**98.63**	**6.12**	**0.9925**	**98.49**
	V	3.82	0.9947	98.95	8.48	0.9859	97.11
Jadkal	ANN I	14.39	0.9230	85.20	30.70	0.7893	62.12
	II	13.93	0.9280	86.13	16.69	0.9492	88.80
	III	**13.75**	**0.9300**	**86.49**	**15.90**	**0.9522**	**89.83**
	IV	13.37	0.9339	87.22	16.71	0.9446	88.78
Kokkarne	WNN I	20.16	0.9804	96.12	21.57	0.9794	95.91
	II	17.71	0.9849	97.00	19.65	0.9830	96.61
	III	13.27	0.9915	98.32	17.09	0.9873	97.43
	IV	**12.20**	**0.9928**	**98.57**	**16.75**	**0.9876**	**97.53**
Kokkarne	ANN I	43.67	0.9044	81.80	46.25	0.9015	81.19
	II	44.06	0.9026	81.48	46.25	0.9012	81.19
	III	**44.22**	**0.9019**	**81.34**	**46.36**	**0.9009**	**81.10**
	IV	41.54	0.9140	83.54	48.48	0.8915	79.33
Dasanakatte	WNN I	11.67	0.9510	90.45	7.30	0.9655	93.22
	II	6.39	0.9857	97.17	3.33	0.9933	98.59
	III	5.47	0.9896	97.93	3.45	0.9924	98.48
	IV	**4.72**	**0.9922**	**98.45**	**2.80**	**0.9950**	**98.99**
	V	4.47	0.9930	98.61	3.06	0.9914	98.81
Dasanakatte	ANN I	16.16	0.9054	81.98	7.72	0.9614	92.40
	II	15.21	0.9167	84.04	7.00	0.9684	93.76
	III	14.96	0.9195	84.56	7.07	0.9680	93.64
	IV	**14.93**	**0.9199**	**84.62**	**7.02**	**0.9682**	**93.73**
	V	14.98	0.9193	84.52	7.07	0.9678	93.64
Sarve Bridge	WNN I	10.04	0.9285	86.22	11.10	0.9536	90.94
	II	5.55	0.9787	95.78	6.84	0.9833	96.55
	III	**4.55**	**0.9857**	**97.17**	**5.05**	**0.9905**	**98.12**
	IV	4.06	0.9886	97.74	6.29	0.9859	97.09
Sarve Bridge	ANN I	13.02	0.8767	76.86	18.61	0.8658	74.50
	II	12.92	0.8786	77.20	17.88	0.8860	76.48
	III	**13.02**	**0.8766**	**76.84**	**16.48**	**0.9031**	**80.00**
	IV	13.13	0.8744	76.46	19.10	0.8623	73.14



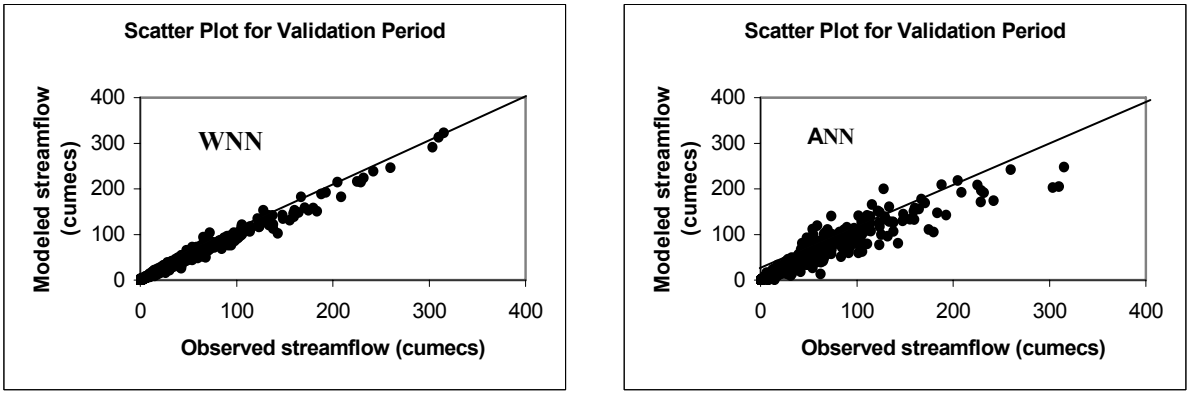
(A) Jadkal Station



(B) Kokkarne Station



(C) Dasanakatte Station



(D) Sarve Bridge Station

Fig. 3 Scatter plots between observed and modelled daily streamflows.

(26 years), and 1979–2003 (25 years), respectively, is used for time series modelling. The daily discharge data of these four basins for the period of 1981–1995 (15 years), 1973–1990 (18 years), 1978–1995 (18 years), and 1979–1995 (17 years) are used for calibration and for periods of 1996–2002 (7 years), 1991–1998 (8 years), 1996–2003 (8 years) and 1996–2003 (8 years) are used for validation period, respectively. The model inputs for each WNN model are indicated in Table 1 and the model inputs for each ANN model given in Table 2.

A total of five sub models were developed for each basin and these models were calibrated and tested for daily discharge series data modelling. The performance of these models in terms of global statistical tests (RMSE, R and %E) are given in Table 3. Similarly, five ANN models have been developed for four basins for the same period in which WNN models are developed. The performance of the ANN models in terms of global statistics are shown in Table 3.

Table 3 reveals that the RMSE is much better (2.8 to 21.57 m^3/s) for WNN models as compared to ANN (7.00 to 48.48 m^3/s) models in all four basins during the validation period. From Table 3, it was also observed that among different antecedent values of the time series models (WNN), the model marked with bold shows lowest RMSE (2.8 to 16.75 m^3/s), high correlation coefficient (0.9876 to 0.9950) and highest efficiency (97.53 to 98.99%) during the validation period in all four basins. Therefore, the WNN model (bold in Table 3) is selected as the best-fit model as compared to the ANN model to forecast the streamflow for the rivers. Figure 3 shows the scatter plot between the observed and modelled values of WNN and ANN for four stations and shows that the forecast daily flows from WNN models were close to the measured values.

Groundwater levels

The monthly groundwater level data from 1977 to 1994 were taken for three wells, namely Kattunga, Munganda and Cheyyeru in the Central Godavari Delta. The period 1977–1989 (13 years) was taken for calibration and 1990–1994 (5 years) was taken for validation. The model inputs for each WNN model are indicated in Table 1 and the model input for each ANN model is indicated in Table 2. The performance of the WNN and ANN models are shown in Table 4 which

Table 4 Performance of WNN and ANN models during calibration and validation.

Station	Model	Calibration: RMSE	R	E (%)	Validation: RMSE	R	E (%)
Kattunga	WNN						
	I	0.789	0.894	79.93	0.975	0.875	76.43
	II	**0.318**	**0.983**	**96.73**	**0.536**	**0.964**	**92.86**
	III	0.261	0.988	97.80	0.596	0.955	91.19
Kattunga	ANN						
	I	1.07	0.790	62.50	1.62	0.587	34.43
	II	**0.952**	**0.841**	**70.80**	**1.36**	**0.739**	**54.10**
	III	0.782	0.895	80.27	1.37	0.736	53.14
Munganda	WNN						
	I	0.341	0.861	74.16	0.454	0.870	74.91
	II	0.165	0.969	93.92	0.325	0.935	87.10
	III	**0.131**	**0.980**	**96.18**	**0.324**	**0.934**	**87.21**
Munganda	ANN						
	I	0.447	0.745	55.59	0.646	0.732	49.22
	II	**0.367**	**0.836**	**70.04**	**0.668**	**0.687**	**45.73**
	III	0.333	0.868	75.38	0.672	0.694	44.93
Cheyyeru	WNN						
	I	0.203	0.891	79.40	0.346	0.826	67.56
	II	**0.107**	**0.970**	**94.25**	**0.188**	**0.954**	**90.44**
	III	0.082	0.982	96.57	0.222	0.931	86.64
Cheyyeru	ANN						
	I	0.341	0.648	42.02	0.442	0.695	47.10
	II	**0.330**	**0.676**	**45.74**	**0.430**	**0.731**	**49.93**
	III	0.278	0.784	61.53	0.468	0.642	40.60

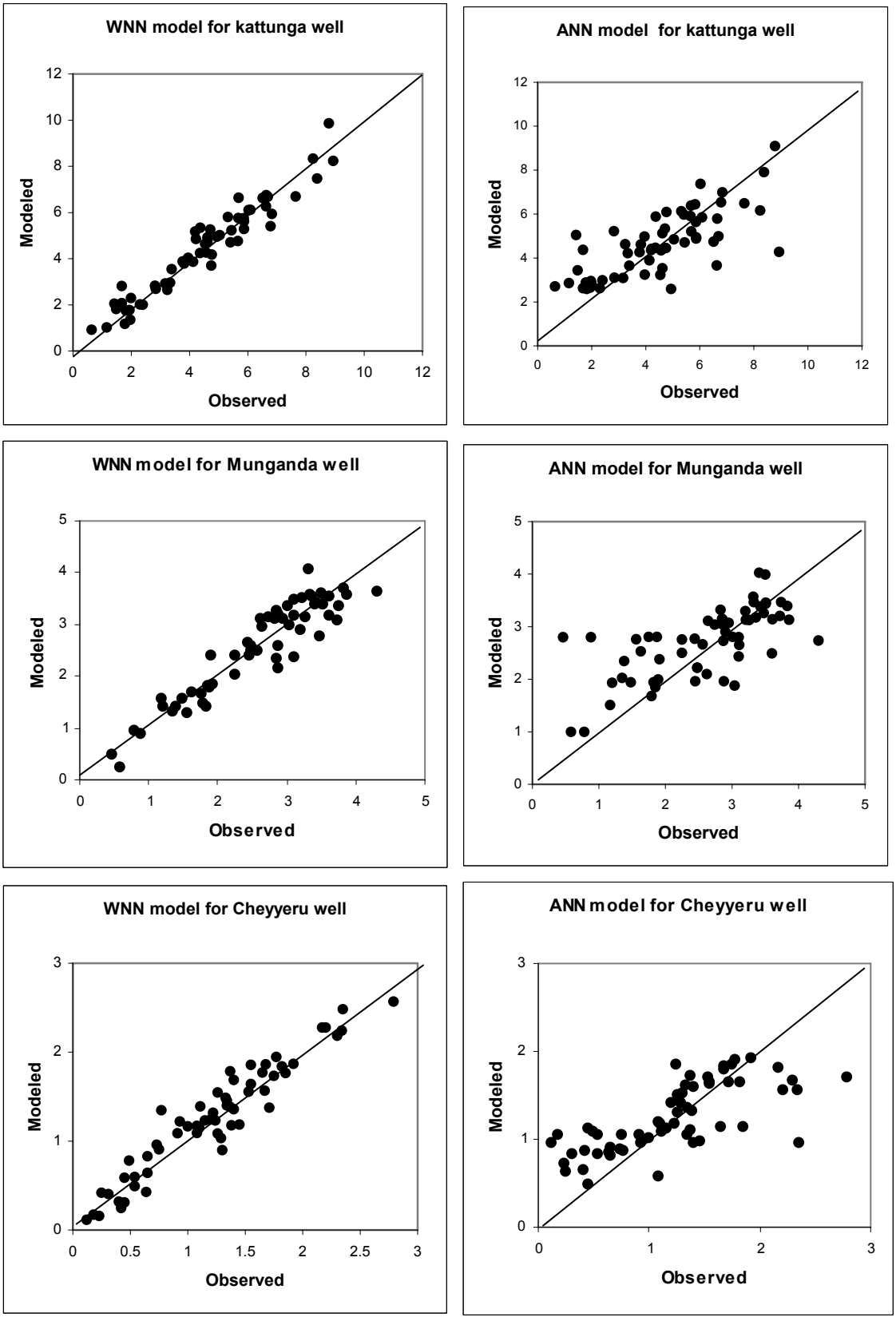


Fig. 4 Scatter plot between observed and modelled monthly groundwater levels.

shows that RMSE is better (0.188 to 0.975 m) for WNN models as compared to ANN (0.43 to 1.62 m) models in all three-observation wells during validation period. From Table 4, it was also observed that the models of WNN having different antecedent values of the time series (marked with bold), estimated minimum RMSE (0.188 to 0.54 m), high correlation coefficient (0.934 to 0.964) and highest efficiency (87.21 to 92.86 %) during the validation period in all three observation wells. Therefore, WNN model is selected as the best-fit model as compared to ANN models to forecast the groundwater levels in Godavari delta. Figure 4 shows the scatter plot between the observed and modelled groundwater levels by WNN and ANN models. It was observed that the values forecasted from WNN models were close to the observed values.

CONCLUSIONS

This paper has reported on an application of wavelet neural network (WNN) models for time series modelling of daily discharge of four westward flowing rivers in India and monthly groundwater levels in the Godavri Delta, Andhra Pradesh, India. Further, ANN models were also developed and the results are compared with WNN models. The comparison revealed that the WNN model exhibits better performance in modelling daily discharge and monthly groundwater levels. This is mainly due to the capability of wavelets to decompose the time series into multi-levels of approximation and detail. The models developed for streamflow and groundwater levels would be useful for water resources planning in the Western Ghats and for groundwater management in the coastal aquifers, respectively.

Acknowledgements The authors are thankful to Mr B. Venkatesh who has provided the data of west flowing rivers for present analysis. They are also grateful to Dr Bhishmkumar, Scientist "F" and Coordinator, DRC, Kakinada and Dr R. D. Singh, Director, NIH, Roorkee, for their encouragement to submit the paper.

REFERENCES

ASCE Task Committee (2000a) Artificial neural networks in hydrology-I: Preliminary concepts. *J. Hydrologic Engng ASCE* **5**(2), 115–123.

ASCE Task Committee (2000b) Artificial neural networks in hydrology-II: Hydrologic applications. *J. Hydrologic Engng ASCE* **5**(2), 124–137.

Coulibaly, P., Anctil, F., Aravena, R. & Bobee, B. (2001a) Artificial neural network modelling of water table depth fluctuations. *Water Resour. Res.* **37**(4), 885–896.

Coulibaly, P., Anctil, F. & Bobee, B. (2001b) Multivariate reservoir inflow forecasting using temporal neural network. *J. Hydrologic Engng ASCE* 6(5), 367–376.

Coulibaly, P., Anctil, F., Rasmussen, P. & Bobee, B. (2000) A recurrent neural networks approach using indices of low-frequency climatic variability to forecast regional annual runoff. *Hydrol. Processes* **14**, 2755–2777.

Coulibaly, P., Bobee, B. & Anctil, F. (2001c) Improving extreme hydrologic events forecasting using a new criterion for ANN selection. *Hydrol. Processes* **15**, 1533–1536.

Hagan, M. T. & Menhaj, M. B. (1994) Training feed forward networks with Marquardt algorithm. *IEEE Trans. Neural Networks* **5**, 989–993.

Labat, D., Ababou, R. & Mangin, A. (2000) Rainfall–runoff relations for karstic springs: Part II. Continuous wavelet and discrete orthogonal multiresolution analyses. *J. Hydrol.* **238**, 149–178.

Maier, H. R. & Dandy, G. C.(1997) Determining inputs for neural network models of multivariate time series. *Microcomputers in Civil Engineering* **12**, 353–368.

Maier, H. R. & Dandy, G. C. (1998). Understanding the behavior and optimizing the performance of back-propagation neural networks: an empirical study. *Environ. Modelling and Software* **13**, 179–191.

McCulloch, W. & Pitts, W. (1943) A logical calculus of the ideas immanent in nervous activity. *Bull. Mathematical Biophysics* **5**, 115–133.

Rumelhart, D. E., Hinton, G. E. & Williams, R. J. (1986) Learning representations by back-propagating errors. *Nature* **323**, 533–536.

Smith, L. C., Turcotte, D. & Isacks, B. L.(1998) Streamflow characterization and feature detection using a discrete wavelet transform. *Hydrol. Processes* **12**, 233–249.

Tokar, A. S. & Johnson, P. A. (1999) Rainfall runoff modelling using artificial neural network. *J. Hydrologic Engng ASCE* **4**(3), 232–239.

Toth, E., Brath, A. & Montanari, A. (2000) Comparison of short-term rainfall prediction models for real-time flood forecasting. *J. Hydrol.* **239**, 132–147.

Wensheng, W. & Jing, D. (2003) Wavelet network model and its application to the prediction of hydrology. *Nature & Science* **1**(1), 67–71.

Xiao, F., Gao, X., Cao, C. & Zhang, J. (2005) Short-term prediction on parameter-varying systems by multiwavelets neural network. *Lecture Notes in Computer Science (LNCS)* no. 3611, 139–146. Springer-Verlag, Berlin, Germany.