ELSEVIER

Journal of Hydrology 270 (2003) 158–166

Journal of Hydrology

www.elsevier.com/locate/jhydrol

Estuary water-stage forecasting by using radial basis function neural network

Fi-John Chang*, Yen-Chang Chen

Department of Bioenvironmental System Engineering, National Taiwan University, Roosevelt Road, Taipei 10770, Taiwan, ROC

Received 31 October 2001; revised 1 September 2002; accepted 6 September 2002

Abstract

The Radial basis function neural network (RBFNN) has been successfully applied to many tasks due to its powerful properties in classification and functional approximation. This paper presents a novel RBFNN for water-stage forecasting in an estuary under high flood and tidal effects. The RBFNN adopts a hybrid two-stage learning scheme, unsupervised and supervised learning. In the first scheme, fuzzy min–max clustering is proposed for choosing best patterns for cluster representation in an efficient and automatic way. The second scheme uses supervised learning, which is a multivariate linear regression method to produce a weighted sum of the output from the hidden layer. Since this network has only one layer using a supervised learning algorithm, its training process is much faster than the error back propagation based multilayer perceptrons. Moreover, only one parameter, θ, must be determined manually. The other parameters used in this model can be adjusted automatically by model training. The water-stage data of the Tanshui River under tidal effect are used to construct a water-stage forecasting model that can also be used during flood. The results show that the RBFNN can be applied successfully and provide high accuracy and reliability of water-stage forecasting in an estuary.

Keywords: Estuary; Neural network; Nonlinear; Radial basis function; Tidal effect; Water-stage forecasting

1. Introduction

An estuary is a semi-enclosed coastal body of water which has free connection to the open sea and within which sea water is measurably diluted with fresh water derived from land drainage (Cameron and Pritchard, 1963). The hydrological systems in an estuary are unique and the water-stage is continually changing under the interaction of riverine and marine processes. The most obvious factors having a profound influence on the water-stage include shape of the estuary, astronomical tide, wind, salinity, temperature, sediment, river discharge, storm surge, and others, are too complex to model directly. Consequently, the hydrodynamic processes of estuaries are manifestly complex and notoriously nonlinear.

Water-stage forecasting in a river under tidal effects is among the most important outstanding problems of flood management. It is never an easy task, because in order to develop a water-stage forecasting model one must know the behavior of the physical processes. The river flow conditions under tidal effects are rarely steady or uniform.

* Corresponding author. Tel.: +886-2-2363-9461; fax: +886-2-2363-5854.

E-mail address: changfj@ccms.ntu.edu.tw. (F.J. Chang).

0022-1694/03/$ - see front matter © 2002 Published by Elsevier Science B.V.
PII: S0022-1694(02)00289-5

However, many of the concepts and principles derived from other watercourses can be applied. For example, the momentum equation is used to study surges and floodwaves, stochastic hydraulics is employed to estimate discharge (Chen and Chiu, 2002) and the diffusion equation with kinetics is applied to evaluate the pollutant movement in estuarine systems. Some empirical equations also have been developed to characterize the flow (Wright et al., 1973; Prandle, 1991; Van Dongeren, 1992). Many deterministic and stochastic models have been developed to forecast the water-stage (Lin and Lee, 1996; Perumal and Ranga Raju, 1998). One of the major disadvantages of using these models is that the parameters are usually difficult to determine from the observed data. Owing to the lack of practicality and difficulty in use, the application of such sophisticated models in Taiwan has so far been unsuccessful.

Even with all these difficulties and challenges, artificial neural networks (ANNs), a relatively new computational tool that has found wide acceptance in many disciplines, provide an alternative way to make important contributions to one step ahead understanding and/or managing these hydrological processes. The attractiveness of ANNs comes from their information processing characteristics, such as nonlinearity, parallelism, noise tolerance, and learning and generalization capabilities (Basheer and Hajmeer, 2000). Recently, ANNs have been successfully used for modeling hydrological processes (Hsu et al., 1995; Rogers et al., 1995; Kuligowski and Barros, 1998; Sajikumar and Thandaveswara, 1999; Govindaraju and Rao, 2000; Chang and Chen, 2001; Chang et al., 2002).

Most of the water-stage forecasting models built by ANNs are based on training the rainfall–runoff data set where the previous stages of rainfall and/or water-stage could dominate the current water-stage. In estuary problems, there are many effects that could influence the water-stage. Moreover, the poorly defined or even misunderstood riverine and marine interaction in the estuary makes it impossible to model its hydrological processes in the circumstance. To develop an estuary water-stage forecasting neural network, a great amount of relative information (data) would be involved, and thus massive network structure. In this study, we present a novel Radial basis function neural network (RBFNN) for solving this poorly defined and complex problem.

2. Radial basis function neural network

RBFNNs rank among the most popular tools for function approximation and have currently been widely applied in many areas, such as nonlinear control, speech processing, and pattern recognition (Gorinevsky and Vukovich, 1997; Pedrycz, 1998). An important property of RBFNNs is that a high-dimensional space nonlinearity problem can be easily broken down through a set of linear combination of radial basis functions. Another important feature of an RBFNN is the ability to be quickly trained. For the purpose of faster training speed, RBFNNs with the hybrid learning scheme applied herein, which is suggested by Moody and Darken (1989), have a feedforward structure that involves three layers. The input layer is composed of n input nodes. The only hidden layer consists of J locally tuned units and each unit has a radial basis function acting like a hidden node. The hidden node output $z_j(x)$ calculates the closeness of the input and projects the distance to an activation function. The activation function of the jth hidden node used in this study is the Gaussian function given by

$$z_j(x) = \exp\left(-\frac{\|x-\mu_j\|^2}{2\sigma_j^2}\right) \tag{1}$$

where x is the n-dimensional input vector; μ_j is the center of the radial basis function for hidden node j; and $\|x-\mu_j\|$ denotes the Euclidean distance between the center of the radial basis function and input; σ_j is a parameter for controlling the smoothness properties of the radial basis functions. The third layer of the network is the output layer with L nodes that are fully interconnected to each hidden node. The output of the network is the sum of linear weighted $z_j(x)$

$$y_l = \sum_{j=0}^{J} w_{lj} z_j(x) \tag{2}$$

$$z_0(x) = 1 \tag{3}$$

where y_l is the lth component of the output layer; w_{lj} is the synaptic weight between the jth node of hidden layer and the lth node of output layer. Eq. (3) denotes the constant, w_{l0}, in the regression Eq. (2).

3. Network training

Training an RBFNN for a specific problem involves selecting the type of basis functions with associated center location $\bar{u}$and width σ, the number of functions K and the weights. Apparently, this is a non-convex optimization problem. A great number of studies have been done to explore the efficient ways. For batch mode training, RBFNNs with localized basis functions offer a very attractive way that, in practice, the estimation of parameters can be decoupled into a two-stage procedure: (1) determine the centers and widths, and (2) based on the results obtained in step (1), determine the weights to the output units.

The two-stage training scheme is shown in Fig. 1. In the first stage only the input values are used for determining the centers and the widths of the radial basis functions. Thus, learning is unsupervised. Once the function parameters are fixed, supervised training (i.e. with target information) can be used for determining the second layer weights. More detail of the two-stage scheme is as follows.

3.1. Unsupervised training—fuzzy min–max clustering

The key to determining the locations and widths of the localized basis functions is to view them as representing the input data density. An efficient approach is to cluster the input vectors and then locate the basis functions at the centers. A variety of clustering techniques can be used to cluster the data with simultaneous water-stage conditions into the same hidden node. The K-means clustering algorithm, which minimizes the sum of squares error between the inputs and hidden node centers, is commonly used to locate the centers of the radial basis functions (Oukhellou and Aknin, 1999); however, this algorithm must be given or input the number of radial basis functions J, and the widths σ_j of every radial basis function usually have to be the same. One has to guess this number J and σ_j before starting training the neural

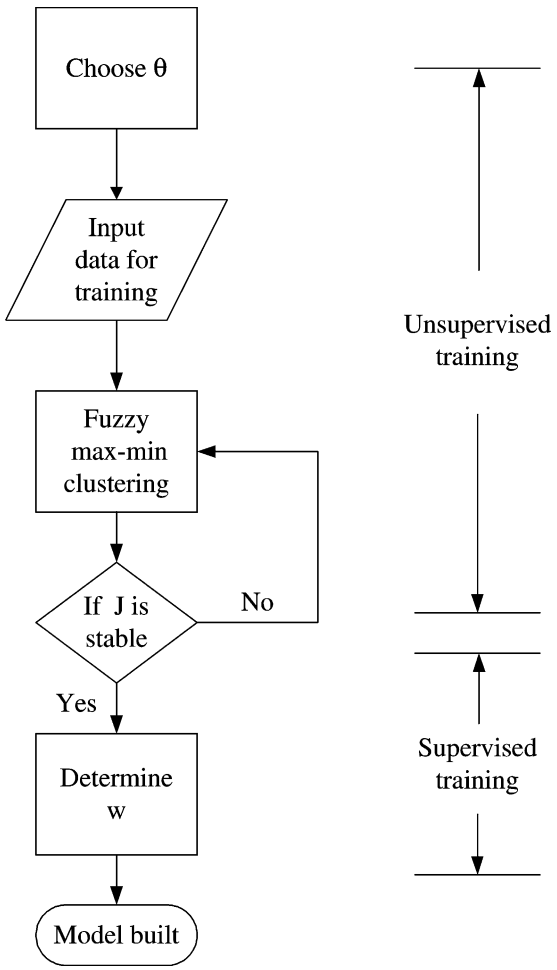

Fig. 1. General flowchart of algorithm for constructing a rainfall–runoff model using RBFNN.

network. Instead of the K-means clustering method, the fuzzy min–max clustering method (Simpson, 1993) is employed in this study. The advantage of the fuzzy min–max clustering method is that the number, centers and σ_j of radial basis functions can all be determined systematically and automatically.

The fuzzy min–max clustering algorithm involves three phases: expansion of a hyperbox, overlap test, and contraction of a hyperbox. During the network training process, a large number of n-dimensional hyperboxes that range from 0 to 1 along each dimension will be generated. Each hyperbox is

viewed as a hidden unit. The boundary of a hyperbox is defined by the max and min points of the form

$$v_j = (v_{j1}, v_{j2}, \ldots, v_{jn}) \tag{4}$$

$$u_j = (u_{j1}, u_{j2}, \ldots, u_{jn}) \tag{5}$$

where v_j and u_j are the min and max points for the jth hyperbox; v_{jn} and u_{jn} are the min and max value for the nth dimension. The membership function for measuring the degree of the hth input x_h falling within the hyperbox j is defined as

$$H_j(x_h, v_j, u_j) = \frac{1}{n}\sum_{i=1}^{n}[1 - f(x_{hi} - u_{ji}) - f(v_{ji} - x_{hi})] \tag{6}$$

$$f(\xi) = \begin{cases} 1 & \xi > 1 \\ \xi & \text{if } 0 \le \xi \le 1 \\ 0 & \xi < 0 \end{cases} \tag{7}$$

where x_{hi} is the ith node of the hth input; $H_j(\cdot)$ is the membership value setting to the unit interval [0,1]; ξ is either $x_{hi} - u_{ji}$ or $v_{ji} - x_{hi}$. The membership values are used to determine which hyperbox needs to be expanded.

At the beginning, the max and min points of the first hyperbox are set to be the first input data. In the subsequent, the degree of membership values will be calculated for every new input, and the hyperbox with the highest degree of membership is tested for possible hyperbox expansion if

$$\sum_{i=1}^{n}(\max(u_{ji}, x_{hi}) - \min(v_{ji}, x_{hi})) \le n\theta \tag{8}$$

in which θ is a user-defined value and $0 \le \theta \le 1$. A small θ means more hyperboxes will be created. If the hyperbox is expanded, the min and max points of the hyperbox will be adjusted as

$$v_{ji}^{\text{new}} = \min(v_{ji}^{\text{old}}, x_{hi}) \tag{9}$$

$$u_{ji}^{\text{new}} = \max(u_{ji}^{\text{old}}, x_{hi}) \tag{10}$$

If no hyperbox can be expanded

$$\left(\sum_{i=1}^{n}(\max(u_{ji}, x_{hi}) - \min(v_{ji}, x_{hi})) > n\theta\right),$$

a new hyperbox will be generated.

$$v_{ji}^{\text{new}} = u_{ji}^{\text{new}} = x_{hi} \tag{11}$$

After the hyperbox is expanded, the hyperbox overlap test will be used to determine whether hyperboxes overlap or not. If overlapping is found between hyperboxes, the max and/or min points of each dimension of hyperbox j could be contained within another hyperbox k. Thus, the hyperboxes will be contracted with the minimal disturbance principle, and only the one dimension that has the minimum overlap is adjusted.

The entire training data will be presented for clustering again and again until no hyperbox needs to be adjusted. The parameter σ_j is determined by half of the width of the hyperbox, as given:

$$\sigma_j = \frac{\sqrt{(v_{j1} - u_{j1})^2 + (v_{j2} - u_{j2})^2 + \ldots + (v_{jn} - u_{jn})^2}}{2} \tag{12}$$

3.2. Supervised learning—multivariate linear regression

After the clustering is completed, the radial basis functions are fixed. In the second scheme, the weights w_{lj} are determined to let the output of network y_{hl} approximate to the target y_{hl}^*. The supervised training algorithm aims to minimize the following sum of squares error,

$$\text{SSE} = \frac{1}{2}\sum_{h=1}^{N}\sum_{l=1}^{L}(y_{hl}^* - y_{hl})^2 \tag{13}$$

where y_{hl} and $y_{hl}^*$are the lth node of the hth set output and target, respectively. Since the outputs of the network are linear combinations of the outputs of the hidden layer, a multivariate linear regression model, given as

$$y^* = zw + \epsilon \tag{14}$$

can be used to determine the weights. In Eq. (14) y^* is the target that is an $N \times L$ matrix; z is the output of the hidden layer that is an $N \times (J + 1)$ matrix; w is the weight between hidden and output layers that is a $(J + 1) \times L$ matrix; and ϵ is independent noise with zero mean. The method of least squares selecting

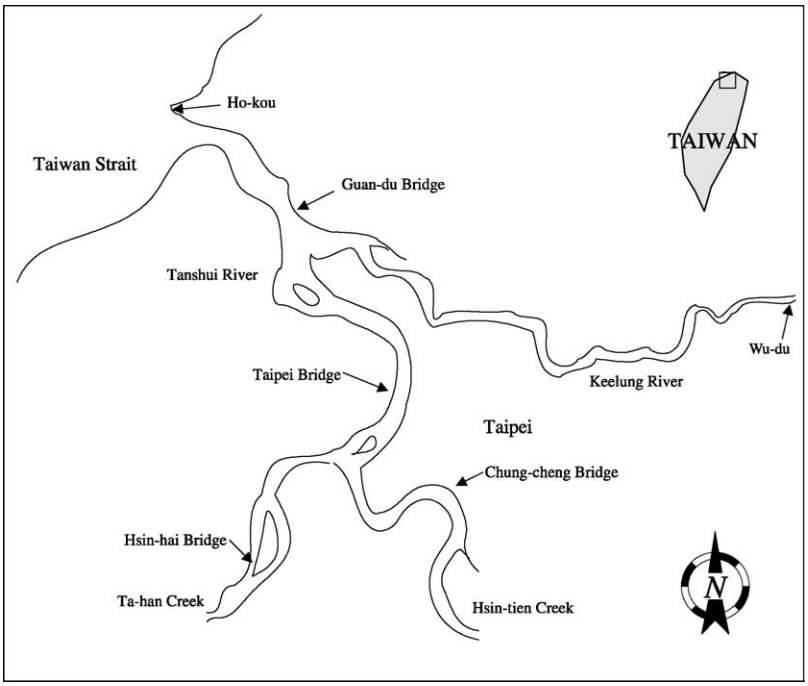

Fig. 2. Locations of the study watershed.

weight matrix w^* to minimize SSE is given by

$$w^* = (z^{\mathrm{T}}z)^{-1}z^{\mathrm{T}}y^* \tag{15}$$

4. Study watersheds and data

To illustrate the foregoing with practical applications, the Tanshui River will be considered. The Tanshui River, composed of three major tributaries, the Ta-han Creek, the Hsin-tien Creek, and the Keelung River, as shown in Fig. 2, is situated near the city of Taipei. The third largest river in Taiwan, it drains 1183 km^2 into the Taiwan Strait and its downstream section for approximately 25 km from the river mouth is under tidal effects. The water-stage in this section is needed as the boundary condition for flood forecasting after heavy rainfalls brought by tropical storms every year. Recently, a flood control project was undertaken on the Tanshui River, including forecasting models, costly levees and bank protection works that can reduce the frequency of inundation of Taipei, mitigate losses by flood damages, and safeguard lives and properties. Water-stage forecasting, especially during a flood, is one of the most crucial parts of the project. Unfortunately, because of the complex hydrological processes in these highly unsteady flows, the traditional physics-based models and time series models cannot be used. The RBFNN can be then used to test its applicability. Six gaging stations equipped with automatic water-stage recorders are located in the study watershed. The available water-stage data collected by Taiwan Water Conservancy Agency are the hourly data measured at the gaging stations.

5. Practical applications

The travel times of flow between the Taipei Bridge Station and the other gaging stations are less than 3 h because of the small watershed. For water-stage forecasting, the inputs of the model include the lunar month, lunar day, time and the water-stages of the six gaging stations for up to 3 h before the study. The only output is the 1-h-ahead water-stage of the

Taipei Bridge Station. In order to employ fuzzy min–max clustering, the different units and scales of the input data must be scaled to have values be between 0 and 1. The model using the RBFNN can be represented by

$$G_t(t+1) = \sum_{j=0}^{J} w_j z_j [T_m(t), T_d(t), T_h(t), G_t(t), G_t(t-1), G_t(t-2), G_h(t), G_h(t-1), G_h(t-2), G_c(t), G_c(t-1), G_c(t-2), G_s(t), G_s(t-1), G_s(t-2), G_w(t), G_w(t-1), G_w(t-2), G_g(t), G_g(t-1), G_g(t-2)] \quad (16)$$

where $G(t+1)$ is the water-stage at time $t+1$; the subscripts t, h, c, s, w, and g represent gaging stations of the Taipei Bridge, Ho-kou, the Chung-cheng Bridge, the Hsin-hai Bridge, Wu-du, and the Guan-du Bridge, respectively; $T_m(t)$, $T_d(t)$, and $T_h(t)$ are lunar month, lunar day, and time. The water-stages of Ho-kou and the Guan-du Bridge are used to describe storm surge, ebb and flood flows. The water-stages of the Chung-cheng Bridge, the Hsin-hai Bridge and Wu-du represent the inflows of the Hsin-tien Creek, the Ta-han Creek and the Keelung River, respectively. The effect of astronomical tide is related to lunar month, lunar day and time.

The data of eight typhoon events during 1994 were obtained, and the continuous hourly water-stage data were measured from July 1996 to December 2000 by Taiwan Water Conservancy Agency. Totally, 22 864 sets data are used in this study. The data are split into three independent subsets: the training, evaluation, and testing subsets, respectively. The training subset includes 13 269 sets of data, the evaluation subset has 3469 sets, while the test subset has the remaining 6126 sets. The training subset is used for model development and parameter estimation. In this phase, a great number of models (networks) are created due to different values of the predetermined parameter θ. In the second phase, the best model from the above candidates is chosen by simulating the performance of evaluation subset data. Neither networks' structure nor their parameters can be adjusted at this stage. As the best model is determined in the previous phase, the testing subset data which is never part of training and validation subsets is then devoted to access the performance of the selected model without any modification.

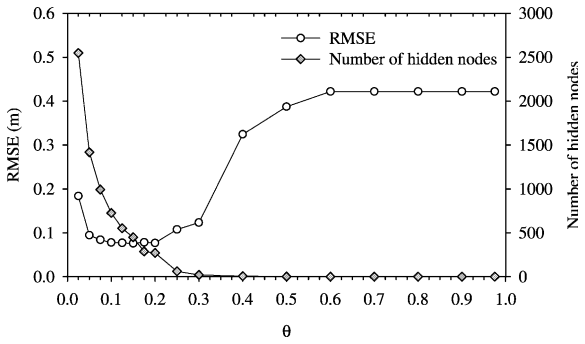

Fig. 3. θ for RMSE and number of hidden nodes.

Correlation coefficient and root-mean-square error (RMSE) are used to evaluate the performance of the networks. The correlation coefficient indicates the strength of relationships between observed and estimated water-stages. The RMSE evaluates the residual between observed and forecasted water stages.

Fig. 3 shows the effect of θ on RMSE and number of hidden nodes. It appears that θ controls the number of hidden nodes (or radial basis functions) where the number of hidden nodes decreases as θ increases. According to the evaluation subset data, the best model, in terms of minimum RMSE, is $\theta = 0.15$ and

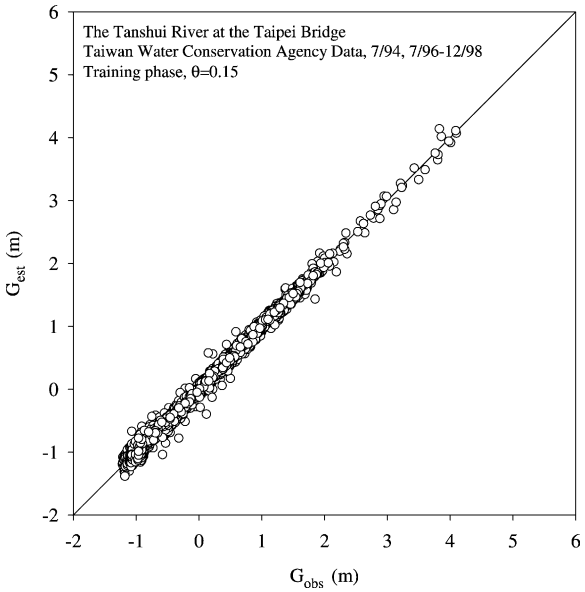

Fig. 4. Observed and forecasted water-stages of the model at training phase.

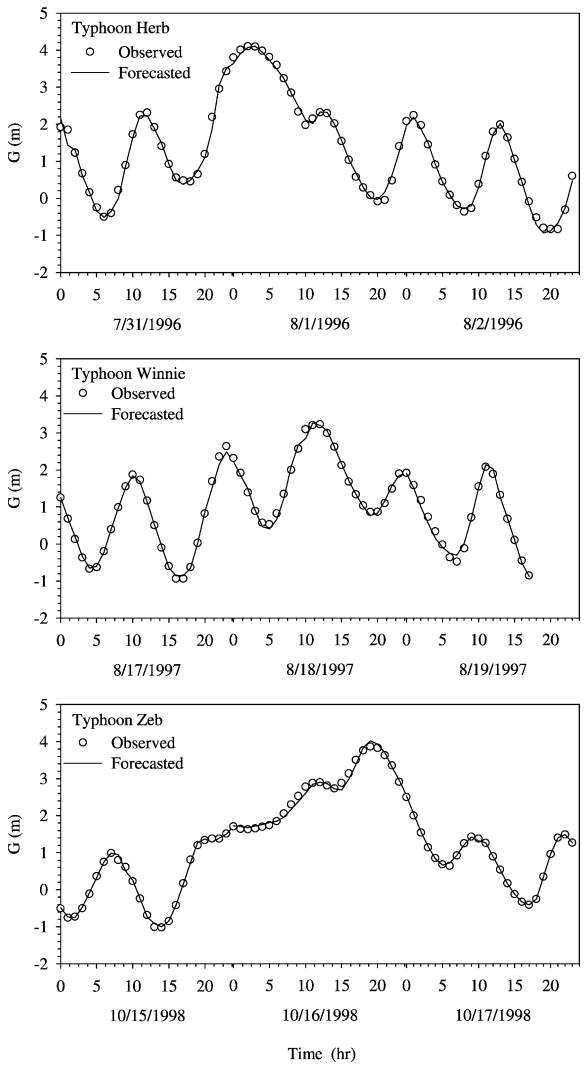

Fig. 5. Accuracy of water-stage forecasting during typhoons at training phase.

450 hidden nodes. That means, on average, each node might include 332 data sets. Compared with the training data sets (13 269), 450 is relative small and can be acceptable. Fig. 4 shows the observed and forecasted water-stages at training phase with $\theta = 0.15$. G_{obs} and G_{est} denote the observed and forecasted water-stages, respectively. All the data points nicely fall onto the line of agreement. Three typhoon events at the training phase are shown in Fig. 5. Fig. 6 is similar to Fig. 4 and shows the accuracy of the water-stages at the evaluation phase

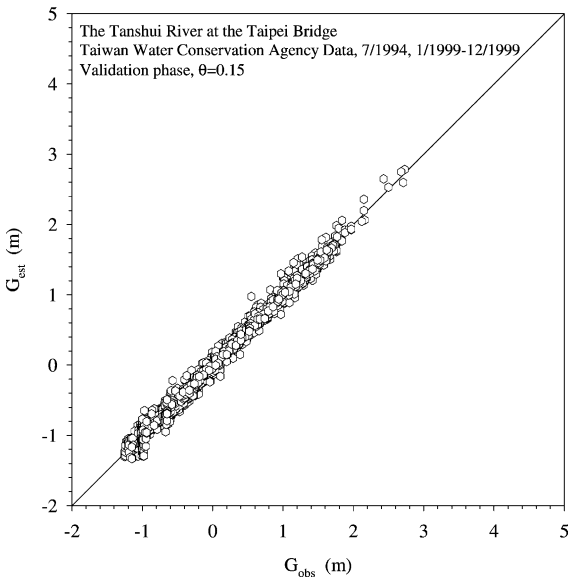

Fig. 6. Accuracy of 1 h ahead water-stage forecasting in the Tanshui River at evaluation phase.

forecasted by the RBFNN. Fig. 7 shows the performance of one of the typhoon events at the evaluation phase that the hydrograph is well predicted. Finally, in the testing phase, the best model identified in previous phase is directly implemented without any adjustment to its structure or parameters. The forecasted and observed water-stages at the testing phase are compared in Fig. 8. Again, it shows that the performance of the RBFNN is very good. The water-stage forecasting model is also exhibited in Fig. 9, which includes four typhoon events at the

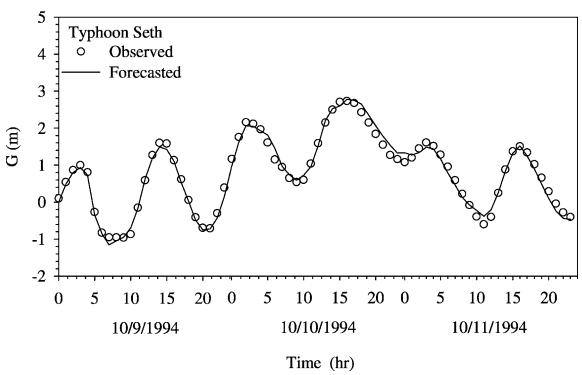

Fig. 7. Accuracy of water-stage forecasting phase during typhoons at evaluation phase.

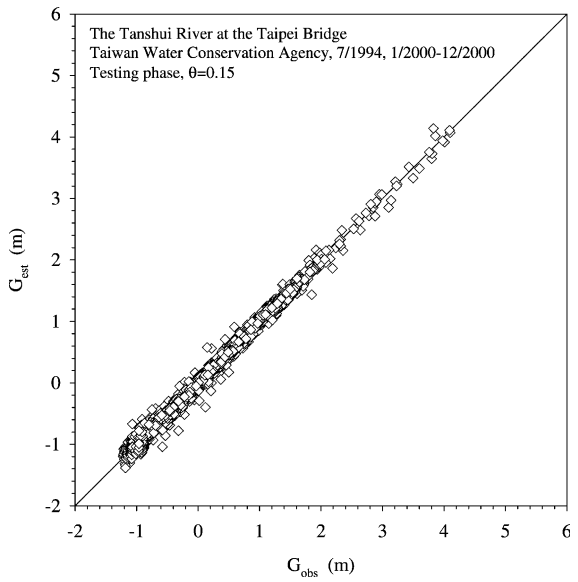

Fig. 8. One hour ahead water-stage forecasting in open channel under tidal effect.

testing phase. The forecasts are very satisfactory and accurate; in particular, the peaks of the hydrographs are captured by the model. These figures indicate that the water-stages forecasted by RBFNN agree quite well with the observed water-stages. The RMSEs of the training, evaluation and testing subsets are 0.059, 0.07 and 0.059, respectively. Moreover, the correlation coefficients of the training, evaluation and testing are 0.998, 0.997 and 0.998. They results show that the model performance is accurate and consistent in these different subsets. All correlation coefficients are very close to unity, and all RMSEs are relatively smaller. It demonstrates that the RBFNN can be successfully applied to establish the model and provide accurate and reliable one-step-ahead water-stage forecasting.

6. Summary and conclusions

The hydrological processes in an estuary are nonlinear and extremely complicated, and it is difficult to quote physically model. The RBFNN, an intelligently adaptive model, has been successfully used for many tasks. In this study, we proposed a novel RBFNN, which employs hybrid unsupervised

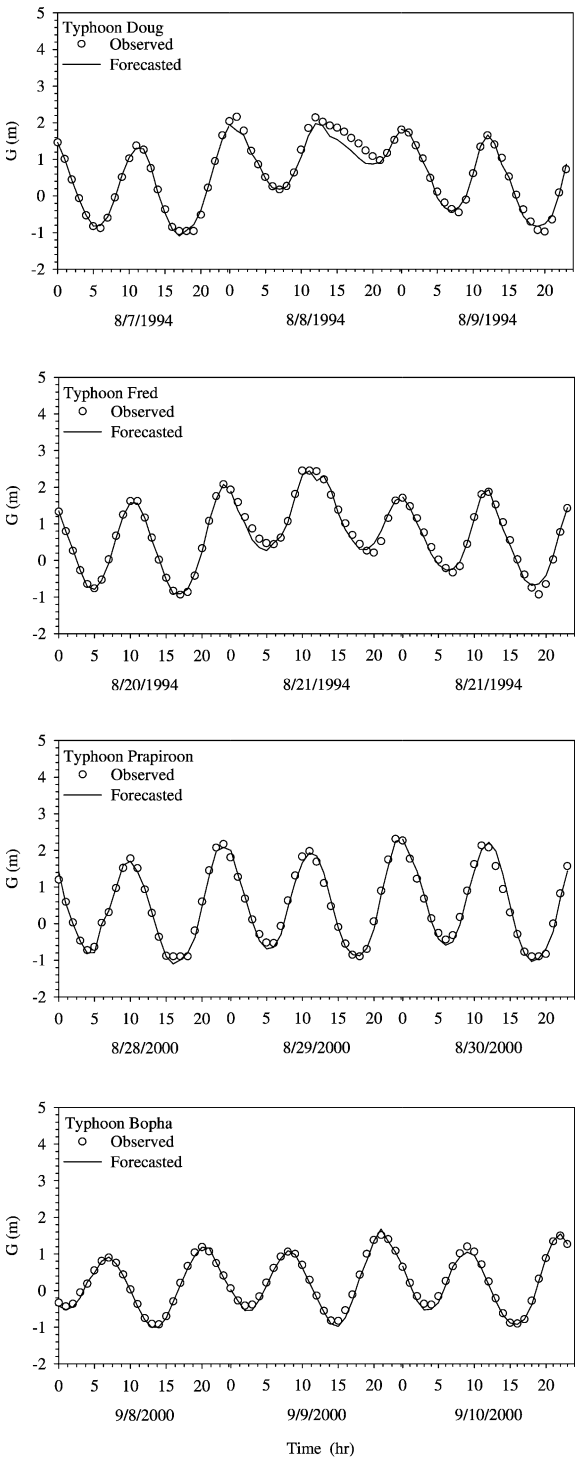

Fig. 9. Accuracy and reliability of RBFNN during typhoons at testing phase.

and supervised training schemes, for water-stage forecasting in an estuary. During the first scheme, the commonly used K-means clustering method is replaced by the fuzzy min–max clustering for determining the characteristics of the radial basis functions. The advantage of using fuzzy min–max clustering is that the number, centers and σ of the radial basis function can be determined systematically and automatically. It is not necessary to estimate many parameters. Only one parameter, θ, needs to be pre-decided. A small θ, which will generate more radial basis functions and might overfit the data at the training phase, does not necessary guarantee better performance in forecasting. The weights between the hidden layer and output layer are obtained by multivariate linear regression method. The output of the RBFNN is simply the sum of the weighted output of the hidden layer. Several candidate models are built by using the training subset data. The model, which applies to the validation subset data and has minimum RMSE, is chosen from those candidate models. The chosen model is then verified, without any further change, through the test subset data to evaluate its applicability and suitability. The water-stage data of the Tanshui River under tidal effect is used for developing the water-stage forecasting model. The results demonstrate that RBFNN networks can successfully model nonlinear hydrological systems and accurately forecast 1-h-ahead water-stage in rivers under tidal and typhoon effects.

Acknowledgements

This paper is based on partial work supported by the Water Resources Bureau and National Science Council of Republic of China (Grant no. NSC 89-2313-13-002-041). The first author thanks Professor Sorooshian, S and Dr. Hsu, K.-L. for their kindly support and valuable advise during their visit of University of Arizona (fall, 2001).

References

Basheer, I.A., Hajmeer, M., 2000. Artificial neural networks—fundamentals, computing, design, and application. J. Microbiol. Meth. 43 (1), 3–31.

Cameron, W.M., Pritchard, D.W., 1963. Estuaries. In: Hill, M.N., (Ed.), The Sea, vol. 2. Wiley, New York.

Chang, F.-J., Chen, Y.-C., 2001. A counterpropagation fuzzy-neural network modeling approach to real time streamflow prediction. J. Hydrol. 245 (1–4), 153–164.

Chang, F.-J., Chang, L.-C., Huang, H.-L., 2002. Real-time recurrent learning neural network for stream-flow forecasting. J. Hydrol. Processes 16, 2577–2588.

Chen, Y.-C., Chiu, C.-L., 2002. An efficient method of discharge measurement in tidal streams. J. Hydrol. 265 (1–4), 212–224.

Gorinevsky, D., Vukovich, G., 1997. Control of flexible spacecraft using nonlinear approximation of input shape dependence on reorientation maneuver parameters. Control Engng Practice 5 (12), 1661–1671.

Govindaraju, R.S., Rao, A.R., 2000. Artificial Neural Networks in Hydrology, Kluwer, The Netherlands.

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

Kuligowski, R.J., Barros, A.P., 1998. Experiments in short-term precipitation forecasting using artificial neural networks. Monthly Weather Rev. 126 (2), 470–482.

Lin, S.-C., Lee, T.-Y., 1996. Water level forecasting of self-combined models in tidal river segment. J. Civil Hydrol. Engng 22 (4), 3–16.in Chinese.

Moody, J., Darken, C., 1989. Fast learning in networks of locally-tuned processing units. Neural Comput. 1, 281–294.

Oukhellou, L., Aknin, P., 1999. Hybrid training of radial basis function networks in a partitioning context of classification. Neurocomputing 28, 165–175.

Pedrycz, W., 1998. Conditional fuzzy clustering in the design of radial basis function neural networks. IEEE Trans. Neural Networks 9 (4), 601–612.

Perumal, M., Ranga Raju, K.G., 1998. Variable-parameter stage-hydrograph routing method. I: Theory. J. Hydrol. Engng 3 (2), 109–114.

Prandle, D., 1991. Tides in estuaries and embayments. In: Parker, B.B., (Ed.), Tidal Hydrodynamics, Wiley, New York, pp. 125–152.

Rogers, L.L., Dowla, F.U., Johnson, V.M., 1995. Optimal fieldscale gorundater remediation using neural networks and the genetic algorithm. Environ. Sci. Technol. 29 (5), 1145–1155.

Sajikumar, N., Thandaveswara, B.S., 1999. A non-linear reinfall–runoff model using an artificial network. J. Hydrol. 216 (1), 32–55.

Simpson, P.K., 1993. Fuzzy min–max neural networks. Part 2: Clustering. IEEE Trans. Fuzzy Syst. 1 (1), 32–45.

Van Dongeren, A., 1992. A model of the morphological behaviour and stability of channels and flats in tidal basins. Delft Hydraul. Rep., H824.55.

Wright, L.D., Coleman, J.M., Thom, B.G., 1973. Processes of channel development in a high-tide-range environment: Gambridge Gulf-Ord River delta, Western Australia. J. Geol. 81, 15–41.