Flood Forecasting Using Radial Basis Function Neural Networks

Fi-John Chang, Jin-Ming Liang, and Yen-Chang Chen

Abstract—A radial basis function (RBF) neural network (NN) is proposed to develop a rainfall–runoff model for three-hour-ahead flood forecasting. For faster training speed, the RBF NN employs a hybrid two-stage learning scheme. During the first stage, unsupervised learning, fuzzy min–max clustering is introduced to determine the characteristics of the nonlinear RBFs. In the second stage, supervised learning, multivariate linear regression is used to determine the weights between the hidden and output layers. The rainfall–runoff relation can be considered as a linear combination of some nonlinear RBFs. Rainfall and runoff events of the Lanyoung River collected during typhoons are used to train, validate, and test the network. The results show that the RBF NN can be considered as a suitable technique for predicting flood flow.

Index Terms—Flood flow, hydrological processes, nonlinear, radial basis function neural network (RBF NN), rainfall–runoff.

I. INTRODUCTION

Increased public awareness has made flooding a prominent focus of hydrological studies; however, for decades we have not coped well with floods. Part of the reason lies in the complex nature of floods and the varied responses to them. Flood forecasting undoubtedly is the most challenging and important task of operational hydrology. Conventional methods for establishing the relationship between rainfall and runoff need to understand the behavior of hydrological cycles and processes; however, the relation is complex and notoriously nonlinear. Many hydrographic and physiographic factors associated with the changing of both space and time may affect the relationship of rainfall and runoff. Some sophisticated hydrological models, which usually include a great number of parameters and observed variables, have been developed to describe complicated hydrological processes. Unfortunately, such models are difficult to use and are impractical in Taiwan, which has a subropical climate, high mountains, and steep upstream channels of all watersheds on the island.

To provide an alternative approach for accurate flood forecasting, an artificial neural network (ANN), which is capable for modeling and control of nonlinear and complex systems, is presented. One of the important characteristics of ANNs is their adaptive nature: learning by examples (input–output pairs). The ability of ANNs to extract dependencies from measured data and complement the existing analytic knowledge of the underlining phenomena makes them a valuable tool in a wide range of applications [1], [2]. In a hydrological context, ANNs have recently been used for streamflow prediction [3]–[5], rainfall estimation [6], [7], and groundwater modeling [8]. In this study, a modified radial basis function (RBF) NN with fuzzy min–max clustering is proposed to construct a rainfall–runoff model for forecasting one-hour-ahead, two-hour-ahead, and three-hour-ahead flood flows during typhoon periods. Fuzzy min–max clustering is used to determine the center and the number of RBFs, and the method of least squares is used to determine the weights of the network. The results show that the modified RBF NN can be applied successfully to build a rainfall–runoff model and provide high accuracy of flood flow prediction.

II. DESCRIPTION OF THE STUDY CATCHMENT AND DATA

The RBF NN is applied to the Lanyoung River catchment, as shown in Fig. 1, for predicting flood flows during high flow periods. The Lanyoung River, the most important river in northeastern Taiwan, is a mountainous river with a steep slope and a catchment area covered by mature forests. The catchment area, which extends from the mountains (elevation 3535 m) to the sea level, drains an area of 821 km$^2$ to the Pacific Ocean. Within this area, the river is 67 km long and the average slope is 5.3%. Due to its subtropical location, the catchment receives an average of 3173 mm of rainfall each year; however, the rainfall is uneven. Typhoons accompanied by torrential rainfall, which usually occur in the summer, are principally responsible for the floods. Consequently, flood forecasting becomes one of the most important tasks of hydrologists and engineers.

Four rain and one stream gaging stations operated by Taiwan Water Conservancy Agency hourly record precipitation totals and stage. The streamflow is generated from the stage-discharge rating curve. Seventeen typhoon events with rainfall and streamflow data during typhoons are available for the period of 1980–1997. The data set is split into three independent subsets: the training, validation, and testing subsets. The training subset, the first 11 typhoon events with 371 set data, is used for parameter estimation and model development. The validation subset, which consists of the data of Typhoons Tim, Doug, and Yanni with 101 set data, is applied to choose the best model from the candidate ones. The testing subset, the latest three typhoon events with 103 set data, is devoted to show the performance of the selected model.

The travel times of the flow from the rain gage stations to the stream gaging station are less than 3 h because of the steep
slope and small watershed. Since the lag time of the drainage catchment relates to the antecedent rainfalls and streamflows, the inputs of the model include the rainfall and the streamflow information between adjacent values for up to 3 hours before the study. The outputs of the model are the one-hour-ahead, two-hour-ahead, and three-hour-ahead streamflows. The inputs, outputs, and structure of the rainfall–runoff model using the RBF NN are shown in Fig. 2 and the established model can be expressed as

$$Q_t = f(Q(t - 1), Q(t - 2), Q(t - 3), P_N(t - 1), P_N(t - 2), P_L(t - 3), P_L(t - 1), P_L(t - 2), P_F(t - 1), P_F(t - 2), P_F(t - 3))$$

where $Q_t$ is the predicted streamflow of the Lanyoung River at the Lanyoung Bridge at time $t$ and $t$ will be 0, 1, and 2. $Q(t)$ is the observed streamflow at time $t$; $P_N(t - 1), P_L(t - 2), P_F(t - 3)$ are the rainfall of Nanshan, Liumaoan, Tuchang, and Fanfan at time $t - 1, t - 2, t - 3$, respectively.

### III. Radial Basis Function Neural Network

The RBF NN, which is capable of universal approximation, may be traced back to an earlier work of Hardy [9]. The present RBF NN, which always consists of three layers, is motivated by Moody and Darken [10], Hush and Horne [11], Pogg and Girosi [12], and others. An important property of the RBF NN is that multidimensional space nonlinearity, such as the relation of rainfall and runoff, can be taken to be a linear combination of the nonlinear RBF [2], [13], [14]. For the purpose of faster training speed, the RBF NN with hybrid learning scheme suggested by Moody and Darken [10] is applied herein. A common feature of the RBF NN is its fast training as compared with the backpropagation networks [13]. Fig. 2 illustrates the architecture of an RBF NN. Three layers are involved in RBF NN architecture and each layer is made up of several nodes. The input layer introduces the outside information into the network. The only hidden layer, which has $J$ nodes, processes the input information with a nonlinear transformation. The transformation associated with each node of hidden layer implemented herein is a Gaussian function defined as

$$Z_j(x) = \exp \left( -\frac{d_j^2}{2\sigma_j^2} \right)$$
$$d_j = ||x - \mu_j||$$

where $Z_j(x)$ is the output of the $j$th hidden node and bias node; $Z_0(x)$, is fixed at 1; $x$ is the input with $n$-dimension; $\sigma_j$ is the width of the receptive field of the $j$th hidden node; $\mu_j$ is the center of the $j$th hidden node; and $d_j$ is the Euclidean distance between $x$ and $\mu_j$. The third layer, the output layer with $L$ nodes, fully interconnects to each hidden node. The output of the network is a linear combination of the nonlinear radial basis functions

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

where $y_l$ is the $l$th output node and $w_{lj}$ is the weight of the connection between the $j$th hidden node and the $l$th output node.

Fig. 3 illustrates the nonlinear relation of rainfall and runoff, constructed by two-dimensional (2-D) inputs (rainfall $P_1$ and $P_2$) and one-dimensional (1-D) output [runoff $Q(P_1, P_2)$]. The cones are the Gaussian functions for some hidden nodes. Many RBFs can be chosen to cover the entire region of the rainfall and runoff data. The irregular surface $Q(P_1, P_2)$ that is a nonlinear relation of rainfall and runoff, is taken to be the sum of linear
weighted nonlinear RBFs $Z_{ij}(P_1, P_2)$, and is a linear combination of finite RBFs. Thus, (1) can then be written as

$$Q_\tau = \sum_{j=0}^{J} w_{j\tau} Z_{j\tau}(Q(t), Q(t-1), Q(t-2), P_N(t))$$

$$P_N(t-1), P_N(t-2), P_L(t)$$

$$P_L(t-1), P_L(t-2), P_T(t)$$

$$P_T(t-1), P_T(t-2), P_T(t)$$

$$P_T(t-1), P_T(t-2).$$  \( (5) \)

IV. LEARNING IN RBF NEURAL NETWORK

Training a RBF NN occurs in two stages. In the first stage, the unsupervised training scheme is employed to determine the parameters of the RBFs ($\mu_j$ and $\sigma_j$). The method for finding the characteristics of the RBFs is always of concern. The $k$-means clustering algorithm, which minimizes the sum of squares error (SSE) between the inputs and hidden node centers, is commonly used to locate a set of $k$ RBF centers. However, there is no formal method for specifying the number of hidden nodes. The numbers $J$ and $\sigma_j$ have to be determined in advance. The fuzzy min-max clustering algorithm, which can determine the number of nodes dynamically and automatically, is used to effectively locate $\mu_j$ and $\sigma_j$ of every RBF. Compared to $k$-means clustering that minimizes an objective function resulting in the mean for each of the $k$ clusters, fuzzy min-max clustering does not minimize any objective function [15]. Thus, the greatest advantage of fuzzy min-max clustering is that it is not necessary to predetermine the number and characteristics of the RBF. The entire training work, including unsupervised and supervised training, has to be done repeatedly when $k$-mean clustering is used. By using fuzzy min-max clustering, however, the iterrant training work will be executed during the unsupervised training stage and the work or supervised training is executed only once. Clearly, fuzzy min-max clustering can save training work and time. During network training, many $n$-dimensional hyperboxes, which can be viewed as hidden nodes, will be generated. The boundaries of a hyperbox are defined by the maximum and minimum points. The degree of membership, which measures the degree of the input falling within the hyperbox, will be calculated when an input is present. The membership function is defined by

$$H_j(x_i, v_{ji}, u_{ji}) = \frac{1}{n} \sum_{i=1}^{n} \left[1 - f(x_i - u_{ji}) - f(v_{ji} - x_i)\right]$$  \( (6) \)

$$f(\xi) = \begin{cases} 1, & \xi > 1 \\ \xi, & 0 \leq \xi \leq 1 \\ 0, & \xi < 0 \end{cases}$$  \( (7) \)

where

- $H_j$: degree of membership setting to [0,1];
- $x_i$: $i$th dimension of input;
- $u_{ji}$ and $v_{ji}$: $i$th dimension of maximum and minimum points of the $j$th hyperbox, respectively;
- $\xi$: either $x_i - u_{ji}$ or $v_{ji} - x_i$.

The fuzzy min-max clustering algorithm involves three phases: expansion of a hyperbox, overlap test, and contraction of a hyperbox. At the beginning, the maximum and minimum points of the first hyperbox are set to be the first input data. The degree of membership values will be calculated for every new input. During expansion of the hyperbox, the hyperbox with the highest degree of membership will be tested for hyperbox expansion according to

$$\sum_{i=1}^{n} (\max(u_{ji}, x_i) - \min(v_{ji}, x_i)) \leq n\theta$$  \( (8) \)

in which $\theta$ is the only one user-defined value and $0 \leq \theta \leq 1$. A small $\theta$ means more hyperboxes will be created. If the hyperbox can be expanded, the old minimum and maximum points of the hyperbox will be replaced with the new minimum and maximum values. If no hyperbox can be expanded, a new hyperbox containing the new input data will be generated. 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 maximum and/or minimum points of each dimension of the hyperbox could be contained within another hyperbox. Thus, the hyperboxes will be contracted with the minimal disturbance principle, and only 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 $\sigma_j$ is set to be half the distance between the maximum and minimum points of the $j$th hidden node.

Once the receptive field widths and centers of hidden nodes are found, the RBFs are kept fixed. During the second stage, supervised training is used to determine the weights between hidden and output layers to let the output of the network approximate to the target. The multivariate linear regression method is used to optimize the weights by minimization of SSE. This consumes less time for model training when compared with that of the other three-layer NNs, e.g., the backpropagation NNs.
CHANG et al.: FLOOD FORECASTING USING RADIAL BASIS FUNCTION NNs 533

Fig. 4. Relation of $\theta$ to NRMSE with hidden nodes is 145.

Only one parameter, $\theta$, is needed to be set in this study. $\theta$ determines the number and characteristics of RBFs, and the performance of the model. The other parameters such as $\omega_j$, $\mu_j$, and $\sigma_j$ will be automatically adjusted during the training processes. Choosing a $\theta$ properly becomes the key to constructing a satisfactory flood-forecasting model.

V. RESULTS

The performance of the hydrological models is usually evaluated by the correlation coefficient, given by

$$\rho = \frac{\sum (Q_o - \bar{Q}_o)(Q_e - \bar{Q}_e)}{\sqrt{\sum (Q_o - \bar{Q}_o)^2 \sum (Q_e - \bar{Q}_e)^2}}$$

which measures how well the predicted streamflows correlate with the observed streamflows. $N$ is the number of data sets and $Q_o$, $Q_e$, $\bar{Q}_o$, and $\bar{Q}_e$ denote the observed, predicted, mean of observed, and mean of predicted streamflows, respectively.

A simple test, which is a widely used measure of forecast error, is the error rate of peak flow, having the form

$$EQ_P = \frac{Q_{P_e} - Q_{P_o}}{Q_{P_o}}$$

in which $Q_{P_e}$ and $Q_{P_o}$ are the forecasted and observed peak flows, respectively. $EQ_P$ is a measure of peak flow error, that is, the difference between the peak flow forecasted and the value that actually occurs. It provides an indication of the performance of the model.

The criterion to choose the best model from the validation subset data is normalized root-mean-square error (NRMSE), given by

$$NRMSE = \frac{1}{\sigma} \left[ \frac{\sum (Q_o - Q_e)^2}{N} \right]^{1/2}$$

in which $\sigma$ standard deviation of measured flood flow. NRMSE indicates the closeness of the prediction to the observation and is used to measure the forecast error. A perfect forecast that all forecasted values are identical to the observations exists only if $NRMSE = 0$. Since the Lanyoung River is a mountain river with a steep channel, floods resulting from typhoons dropping large amounts of rain within a brief period move at very fast speed. The time of concentration, which is the time of flow from the farthest point in the watershed to the outlet of the watershed, is very short. The peak flows usually arrive the gaging station within hours, whereas the accuracy of flood forecasting decreases when forecasting time increases. One-hour-ahead flood forecasting hence becomes the key to constructing the flood-forecasting model.

One-hour, two-hour, and three-hour flood forecasting are weighted by 0.6, 0.3, and 0.1. Fig. 4 shows the relation between $\theta$ and NRMSE. The minimum weighted NRMSE is obtained when $\theta = 0.2$ and the number of hidden nodes is 145. The performance of training from the 11 training subset data with $\theta = 0.2$ is shown in Fig. 5. In all cases, the runoff hydrographs are well simulated. Fig. 6 demonstrates the best performance of validation subset data. Fig. 7 illustrates the performance of the flood flow forecasting of the three testing typhoon events with $\theta = 0.2$. The hyetographs,
VI. CONCLUSIONS

A simple but reliable rainfall–runoff model is established to predict flood flows during typhoons. Particularly pertinent to this paper are results that show the modified RBF NN is capable of providing arbitrarily good prediction of flood flow up to three hours ahead. Despite the fact that the nonlinear relation between rainfall and runoff is extremely difficult to explore, this relation, which could not be well clarified and simulated by conventional hydrological models, can be taken to be a linear combination of some nonlinear RBFs. The modified RBF NN, which is a model-free estimator, can be used successfully for constructing rainfall–runoff models and provide valuable flood flow predictions. The modified RBF NN employs a hybrid two-stage learning scheme, unsupervised and supervised training. During the unsupervised training, fuzzy \( \text{min–max} \) clustering is introduced to determine the width and center of the receptive fields and the number of the RBFs. The advantage of using fuzzy \( \text{min–max} \) clustering is that the characteristics of...
RBFs are determined dynamically and automatically and only one parameter, $\theta$, is needed when the rainfall–runoff model is constructed. During the supervised training, multivariate linear regression, which consumes less training time, is used to determine the weights between layers. The optimal rainfall–runoff model using the validation subset data is chosen from the candidate model with minimum weighted NRMSE. The rainfall and runoff data of the Lanyoung River collected during typhoons are used to construct the rainfall–runoff model. The results show that the modified RBF NNs can be applied successfully to building rainfall–runoff models and offer high accuracy of flood forecasting.

**ACKNOWLEDGMENT**

The authors wish to thank Prof. Frank for helping re-edit, and the reviewers for their valuable suggestions to improve the paper.

**REFERENCES**


**CHANG et al.: FLOOD FORECASTING USING RADIAL BASIS FUNCTION NNs**

**TABLE I**

**SUMMARY RESULTS OF THE RBF NEURAL NETWORK APPLIED DURING THE STUDY**

<table>
<thead>
<tr>
<th>Subset</th>
<th>Typhoon event</th>
<th>Peak flow m$^3$/s</th>
<th>$E_Q_t$</th>
<th>$\rho$</th>
<th>NRMSE</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>$t+1$</td>
<td>$t+2$</td>
<td>$t+3$</td>
</tr>
<tr>
<td>Training</td>
<td>Norris</td>
<td>1280</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>Ike</td>
<td>317</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>June</td>
<td>1330</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>Andy</td>
<td>2916</td>
<td>0.91</td>
<td>5.35</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>Nelson</td>
<td>409</td>
<td>0</td>
<td>20.65</td>
<td>0</td>
</tr>
<tr>
<td>Validation</td>
<td>Brenda</td>
<td>1230</td>
<td>0</td>
<td>7.21</td>
<td>-0.081</td>
</tr>
<tr>
<td></td>
<td>Alex</td>
<td>393</td>
<td>1.13</td>
<td>-4.63</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>Gerald</td>
<td>1350</td>
<td>6.60</td>
<td>-8.65</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>Lyon</td>
<td>3060</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>Sarah</td>
<td>3090</td>
<td>-2.73</td>
<td>1.87</td>
<td>-0.77</td>
</tr>
<tr>
<td></td>
<td>Dot</td>
<td>3640</td>
<td>-3.50</td>
<td>0</td>
<td>-2.87</td>
</tr>
<tr>
<td>Testing</td>
<td>Tim</td>
<td>3350</td>
<td>-3.47</td>
<td>-3.36</td>
<td>0.02</td>
</tr>
<tr>
<td></td>
<td>Doug</td>
<td>1030</td>
<td>1.87</td>
<td>17.13</td>
<td>33.28</td>
</tr>
<tr>
<td></td>
<td>Yanni</td>
<td>1880</td>
<td>-3.34</td>
<td>2.47</td>
<td>2.56</td>
</tr>
<tr>
<td></td>
<td>Gladys</td>
<td>1470</td>
<td>-0.65</td>
<td>0.17</td>
<td>10.50</td>
</tr>
<tr>
<td></td>
<td>Seth</td>
<td>1090</td>
<td>0.92</td>
<td>0.85</td>
<td>-2.86</td>
</tr>
<tr>
<td></td>
<td>Herb</td>
<td>3020</td>
<td>-0.73</td>
<td>-3.88</td>
<td>9.29</td>
</tr>
</tbody>
</table>

**Fi-John Chang** received the B.S. and M.S. degrees in agricultural engineering from National Taiwan University (NTU), Taipei, Taiwan, R.O.C., in 1980 and 1984, respectively, and the Ph.D. degree in civil engineering from Purdue University, West Lafayette, IN, in 1988. From 1994 to 1997, he was Director of the Hydrotech Research Institute and from 1997 to 2000, Chairman of the Department of Agricultural Engineering (renamed as the Department of Bioenvironmental Systems Engineering) at NTU. He has also been Project Leader in a number of research projects funded by the Ministry of Economic Affairs and Council of Agriculture, Taiwan. He is currently a Professor in the Department of Bioenvironmental Systems Engineering at NTU and is involved in research and development of stochastic hydrology, hydrological information systems, and ANNs.

**Jin-Ming Liang** received the B.S. degree in hydraulic engineering from Feng Chia University, Taiwan, R.O.C., in 1991 and the M.S. degree in agricultural engineering from National Taiwan University (NTU), Taipei, Taiwan, in 1993. He is currently pursuing the Ph.D. degree at NTU, where he is working with Prof. F.-J. Chang on the prediction and analysis of hydrological systems.

**Yen-Chang Chen** received the B.S. and M.S. degrees in agricultural engineering from National Taiwan University (NTU), Taipei, Taiwan, R.O.C., in 1989 and 1991, respectively, and the Ph.D. degree in civil and environmental engineering from the University of Pittsburgh, Pittsburgh, PA, in 1998. He is currently a Postdoctoral Research Associate with Prof. F.-J. Chang in the Department of Bioenvironmental Systems Engineering, NTU, and is currently involved in research and development of hydrological information systems.