DOI: 10.1007/s00267-002-2862-9

Modeling Flow and Sediment Transport in a River System Using an Artificial Neural Network

LI YITIAN
The Key Laboratory of Water and Sediment Sciences of Ministry of Education of China
Wuhan University, Wuhan, China, 430072

ROY R. GU*
Department of Civil and Construction Engineering
Iowa State University, Ames, Iowa 50011, USA

ABSTRACT / A river system is a network of intertwining channels and tributaries, where interacting flow and sediment transport processes are complex and floods may frequently occur. In water resources management of a complex system of rivers, it is important that instream discharges and sediments being carried by streamflow are correctly predicted. In this study, a model for predicting flow and sediment transport in a river system is developed by incorporating flow and sediment mass conservation equations into an artificial neural network (ANN), using actual river network to design the ANN architecture, and expanding hydrological applications of the ANN modeling technique to sediment yield predictions. The ANN river system model is applied to modeling daily discharges and annual sediment discharges in the Jingjiang reach of the Yangtze River and Dongting Lake, China. By the comparison of calculated and observed data, it is demonstrated that the ANN technique is a powerful tool for real-time prediction of flow and sediment transport in a complex network of rivers. A significant advantage of applying the ANN technique to model flow and sediment phenomena is the minimum data requirements for topographical and morphometric information without significant loss of model accuracy. The methodology and results presented show that it is possible to integrate fundamental physical principles into a data-driven modeling technique and to use a natural system for ANN construction. This approach may increase model performance and interpretability while at the same time making the model more understandable to the engineering community.

A river system is a complex network of intertwining channels and tributaries, where floods may frequently occur, with interacting flow and sediment transport processes. The assessment of sediment volume being transported by streamflow is of vital interest in hydraulic engineering due to its importance in the design and management of water resources projects. Numerical simulations of water discharges, flow velocities, and sediment transport rates have been attempted to investigate water and sediment problems. Hydrodynamic models have been widely used for the analysis, prediction, design, and management of a wide range of water–sediment systems. However, the spatial heterogeneity of various physical and geomorphologic properties of a river system can not be easily represented, and the requirement for large amounts of data by sophisticated deterministic models must be considered. In the application of conventional models, the emerging issue is the requirement of detailed topographical, geophysical, and morphometric data. A river network covers a vast area consisting of many watersheds and subbasins, where a complete set of data may not be available. A practical user-friendly model is needed for quick simulations and predictions with minimum data requirement and without significantly compromising the model accuracy.

An artificial neural network (ANN) has the characteristics of parallel link, error correction, and nonlinear transfer and is an emerging technique for the flow and connection of information. It is constructed to obtain a prediction of system response without attempting to reach an understanding of or provide insight into the nature of the phenomena that are represented (Haykin 1994, Fausett 1994, Hornik 1989, Fahlamn 1989, Rumelhart and others 1986, Rogers and Lamarsh 1992). It is a tool for nonlinear input–output mapping, which usually consists of input, output, and layers of hidden units or elements called neurons. ANNs have been successfully applied to the field of pattern recognition (Bishop 1995) and are increasingly used in the areas of the aquatic environment (Aly and Peralta 1999, Dawson and Wilby 1998, French and others 1992, Chang and Tsang 1992, Hsu and others 1995, Zhang and Stanley 1997). Dibike and others (1999) conducted the encapsulation of numerical–hydraulic models in ANNs for a flow forecasting problem with encouraging results. Dibike and Abbott (1999) employed ANNs to simulate plane two-dimensional flow. Many studies have

KEY WORDS: Artificial neural networks; River system; Streamflow; Sediments; Water resources management

*Author to whom correspondence should be addressed; *email:* roygu@iastate.edu

used ANNs to investigate the rainfall–runoff relationship over watersheds (Zhu and others 1994, Tokar and Johnson 1999). Yang and others (1998) developed a hydrology-based ANN model using nonlinear reservoir theory. However, the ANN technique has not yet been applied to sediment yield predictions.

In view of growing applications of ANNs in the areas of water resources (Smith and Eli 1995, Raman and Sunilkumar 1995, Minns and Hall 1996, Carrieve and others 1996), a comprehensive review of their concept and applications was carried out by ASCE (2000a, 2000b). It was concluded that ANNs could perform as well as existing models. However, the physics of the underlying process is locked up in a set of ANN configuration and not revealed back to the user after training. Moreover, in previous studies using ANN models, investigators experienced difficulty in determining the appropriate ANN structure or optimum network architecture, especially hidden layers, due to a lack of fundamental physical principles and an understanding of internal conditions of the system being simulated. Previous investigators had to use a trial-and-error approach, i.e., going through extensive experiments and many trials, to determine the number of units in hidden layers. The trial-and-error approach is somewhat frustrating and time-consuming (Karunanithi and others 1994). The integration of ANN and hydrodynamics can avoid the disadvantages of a hydrodynamic model in real-time prediction, e.g., large data requirement and long computational time, while overcoming the difficulty in choosing the optimal ANN structure and internal functions. If a real river network is used as the ANN model structure, it is possible and necessary for the model to employ the physical functions and internal conditions of the river system.

This paper is concerned with the application of the artificial neural network to water-sediment problems in a complex river system. The objectives of this study are to develop an ANN river system model for simulating and predicting flow and sediment transport in a river system and to demonstrate the practical capability and usefulness of the ANN technique. The model utilizes the real river system being simulated as the ANN architecture and incorporates the physical behavior and internal conditions of the system. Water and sediment mass conservation equations are integrated into a multilayer feed-forward network with an error back propagation algorithm. The integration of physical functions into the model and the use of an actual river network for ANN construction makes it possible to have an appropriate or optimum architecture for ANN modeling and makes it easier for the engineering community to understand the technique and interpret results. The model is intended for real-time

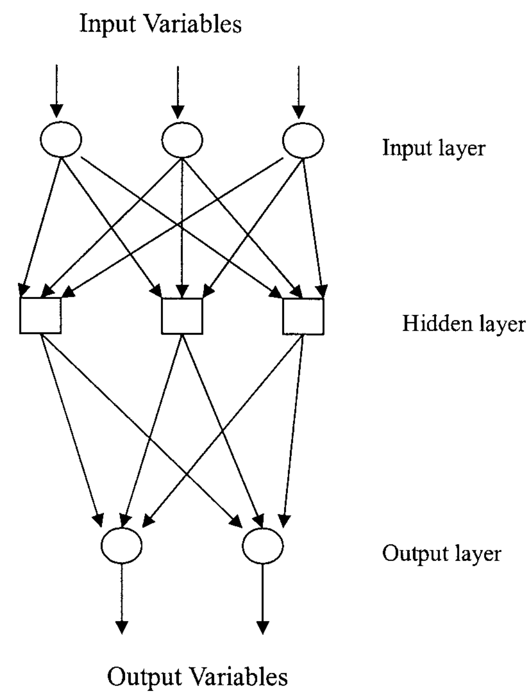

Figure 1. Typical three-layer feed-forward ANN.

prediction of flow and sediment transport in a complex waterway network with less data required for topographical and morphometric information than a conventional hydrodynamic model without compromising modeling accuracy. The model is calibrated and tested against observed water discharges and sediment transport quantities in the Jingjiang Reach of Yangtze River and Dongting Lake, China.

Method

ANN Overview

Among many ANN structures, the most widely used one in the area of hydrology is the multilayer, feed-forward network (Rumelhart and others 1986). In a feed-forward network, data flows in one direction (Rumelhart and others 1986, White 1990, Gallant and White 1992). An ANN has input, output, and hidden middle layers (Figure 1). A neural network consists of a large number of simple processing elements or units, namely neurons or nodes. Each neuron is connected to other neurons by means of direct communication links, each with an associated weight. The neurons in one layer are not connected among themselves. The data passing through the connections from one neuron to another are multiplied by weights that control the strength of a passing signal. The input layer neurons receive the input vector and transmit the values to the next layer of processing elements across connections. This process is continued until the output layer is reached. The weights represent information being used by the net to solve a problem. The net usually has two or more layers of processing units where each unit in each layer is connected to all processing units in the adjacent layers. The desired output is achieved by adjusting the weights on the links between the neurons and calculating the value of error function for a particular input and then back-propagating the error from one layer to the previous one (Rumelhart and others 1986). Each neuron multiplies every input by its interconnection weight, sums the product, and then passes the sum through an activation (or transfer) function to produce its result. Activation functions typically used in previous studies (Imrie and others 2000, Jain and Chalisgaonkar 2000) included logistic sigmoid, bipolar sigmoid, linear, sigmoid plus linear, and cubic polynomial. The output for sigmoid function is always bounded between 0 and 1.

ANN Learning

The process of determining ANN weights is called learning or training and is similar to calibration of a mathematical model. The ANNs are trained with a training set of input and known output data. At the beginning of training, the weights are initialized either with a set of random values or based on some previous experience. Next, the weights are systematically changed by the learning algorithm such that for a given input the difference between the ANN output and actual output is small. The ANN learning process is terminated when this difference is less than a specified tolerance. In general, there are two additional termination procedures that are commonly used: (1) when the weights are updated a maximum number of times and (2) when the error calculated for a separate test dataset begins to increase, i.e., by cross-validation. At this stage, the ANN is considered trained. If regional information is not available or insufficient on important river reaches, a hydrodynamic model still can be used to generate training data for an ANN.

The most widely used learning rule for ANNs or multilayer perceptions is the error back-propagation (BP) algorithm developed by Rumelhart and others (1986) (Tchaban and others 1998, Bishop 1995). The BP algorithm is based upon the generalized delta rule. In BP processes, all nodes change their weights based on the accumulated derivatives of the error with respect to each weight. A set of inputs and outputs is selected from the training set, and the network calculates the output based on the inputs. This output is subtracted from the actual output to find the output layer error. The error is back-propagated through the network, and the weights are suitably adjusted. This process continues for the number of prescribed sweeps or until a prescribed error tolerance is reached. The mean square error over the training samples is the typical objective function to be minimized. After training is completed, the ANN performance is validated and implemented for its intended use. An ANN is better trained when a wider range of environmental scenarios is used, under which input data are collected. If the model is trained using a dataset that contains a limited range of values, it may perform poorly when encountering events having previously unobserved values. The failure to generalize may limit its use as a tool in applications where the data available for calibration is unlikely to cover all possible scenarios. Imrie and others (2000) presented a methodology for improving the generalization performance of an ANN model by adding a guidance system to a learning architecture and including a simple cross-validation procedure in ANN training. The method can produce models that generalize well on new data and extrapolate beyond the range of values included in the calibration range. Imrie and others (2000) scaled river flow data so that the training data values lay between 0.2 and 0.8 or between 0.1 and 0.9. However, reducing the range of the scaled values further may lead to loss of information and a poorer overall network performance (Imrie and others 2000).

A temporal back-propagation algorithm, i.e., a recurrent ANN, was presented in work by Haykin (1994) and Elman (1990). The recurrent ANN is a nonlinear system in which the outputs from the net at one time step become the inputs at the next time step, are inherently dynamic in nature, and are able to deal with time-varying information (Anmala and others 2000). This dynamic feature embedded in BP ANN architecture makes it suitable to modeling time-dependent flow and sediment transport process.

The determination of optimal network architecture is a part of the learning strategy (Fahlman and Lebiere 1990). The number of input, output, and hidden layer nodes depends on the problem being studied. There are no fixed rules about the number of nodes in the hidden layer. However, if the number of nodes in the hidden layer is too 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 overfit the data (Karunanithi and others 1994). In previous studies (Jain and Chalisgaonkar 2000, Anmala and others 2000) the number of neurons or units in the hidden layer was determined after many trials. The configuration that gave the minimum sum of the square of errors

(SSE) was selected as the ANN structure (Jain and Chalisgaonkar 2000). Such an approach is frustrating and time-consuming (Karunanithi and others 1994). To conquer this problem, a constructive algorithm (namely cascade correlation learning architecture) was developed by Fahlman and Lebiere (1990). It is a type of feed-forward ANN that constructively builds the network by adding one hidden unit at a time to avoid the need to manually locate the optimum network structure (Imrie and others 2000). However, the straightforward cross-validation procedure employed in error-back propagation training can not be applied to this algorithm (Imrie and others 2000). Several alternatives were suggested by Russell and Norvig (1995), Karnin (1990), and Hsu and others (1995). Various empirical guidelines based on the number of inputs or training patterns were also proposed (Weigend and others 1990). In this study, a river system is modeled by an ANN configuration, which is based on the structure of the river network, so that the difficult and time-consuming practice in finalizing the hidden layer structure can be avoided.

In most cases, network output depends not only on the current and previous input but also on the condition of the system. For instance, when significant degradation occurs in a river, sediment outflow is influenced by local scouring or erosion in the channel in addition to inflow. In this case, the input–output relation is very complex and difficult to simulate using a traditional ANN transfer function. A conventional BP ANN is a "black box" without physical meanings of internal parameters and physical relations between the parameters and output. It may give good training results, but poor implementation results when system conditions vary. Therefore, water and sediment mass conservation over the whole network and at all nodes needs to be considered when using BP algorithm for simulating water movement and sediment transport in a river system.

ANN Representation of a River System

A river system is a network of intertwining channels, tributaries, lakes, and other waterbodies connected to each other. It may be nonlinear and multivariate, and the variables involved may have complex interrelationships. Such problems can be efficiently solved using ANNs as there are many similarities between a neural network and a river system. The processes that involve several parameters in a complex system such as a river system are amenable to neurocomputing. Therefore, it is feasible to use an ANN model to simulate flow and sediment transport processes in a complex river system.

A river system can be conceptualized according to research needs and the connection and interaction between waterbodies. To meet the requirement for developing an ANN model for flow and sediment transport, some considerations or assumptions are made. The river network is represented by a system of interconnected nonlinear reservoirs. Upstream inflows and sediment loads to various waterbodies are used as the model input and downstream discharges and sediment transport rates at a downstream station as the model output. It is assumed that there is no interaction, i.e., water and sediment exchanges, between reservoirs in the same layer of the network. Interaction between reservoirs in the adjacent layers is represented by the weight, to which a zero is assigned if there is no exchange between two layers. This treatment makes the use of the similarity between the ANN and the river network. The nodes in the first and last layers of the network serve as input and output, respectively, and do not have storage capability. The nodes of the internal layers have storage functions through which there are water and sediment exchanges. Reservoir storage-outflow function is nonlinear for both water and sediment. Water and sediment mass conservation principles apply to all reservoirs in the river system.

With the above-mentioned assumptions and considerations, the simplified representation of a river network (Figure 2) consists of three components: water and sediment inflows (sources) as input, internal reservoirs in parallel or series, and water and sediment outflows as output. A nonlinear relationship exists between the input to the first layer and the output from the last node. The water continuity and sediment transport equations are used for satisfying water and sediment mass conservation over the whole river system as well as at all nodes.

Mass Conservation–Transfer Function

Mass conservation is used as the transfer (or activation) function in this ANN modeling of a river system. In the representation of a river system, i.e., a conceptual model characterizing the river net, mass conservation is satisfied at all nodes and over the entire network by the water continuity equation,

$$\frac{\partial V_{w,j}^{k+1}}{\partial T} = \sum_{i=1}^{N_k} \omega_{i,j}^{w,k+1} Q_{w,i}^{k} - Q_{w,j}^{k+1} \tag{1}$$

and the sediment continuity equation,

$$\frac{\partial V_{s,j}^{k+1}}{\partial T} = \sum_{i=1}^{N_k} \omega_{i,j}^{s,k+1} Q_{s,i}^{k} - Q_{s,j}^{k+1} \tag{2}$$

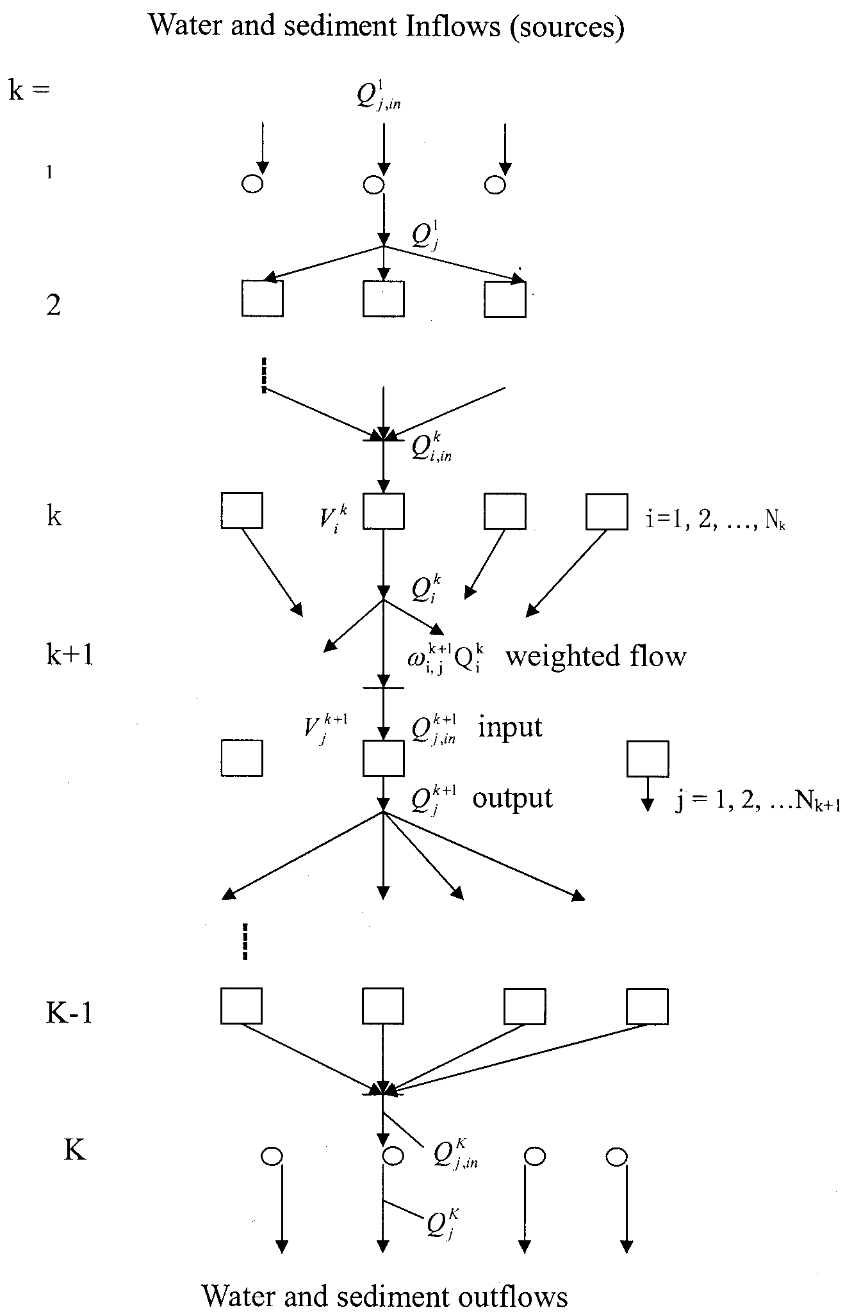

Figure 2. ANN representation of a river network.

where subscripts and superscripts w and s denote water and sediment, respectively; index k is for the layers of reservoirs; index i is for reservoirs in a previous layer ($i = 1, 2, \ldots, N_k$); index j is for reservoirs in the current layer ($j = 1, 2, \ldots, N_{k+1}$); V is the water storage or sediment deposition; Q is water flow-rate (discharge) or sediment transport rate; T is time; ω is the fraction of water or sediment from a reservoir in the immediate upstream layer entering a reservoir in the current layer, i.e., corresponding weight; and N_k is the total number of reservoirs in the previous layer (k) entering the current layer ($k + 1$). The sediment transport rate, Q_s, is equal to the product of discharge, Q, and sediment concentration, C, i.e. $Q_s = QC$.

The generalized form of equations 1 and 2 is written as

$$\frac{\partial V_j^{k+1}}{\partial T} = \sum_{i=1}^{N_k} \omega_{i,j}^{k+1} Q_i^k - Q_j^{k+1} \tag{3}$$

After discretization, the difference form of equation 3 is

$$\frac{\Delta V_j^{k+1}}{\Delta T} = \sum_{i=1}^{N_k} \omega_{i,j}^{k+1} Q_i^k - Q_j^{k+1} \tag{4}$$

The state variable for jth reservoir in $(k + 1)$th layer at time step $T + \Delta T$ is determined by

$$V_j^{k+1}|_{T+\Delta T} = V_j^{k+1}|_T + \left(\sum_{i=1}^{N_k} \omega_{i,j}^{k+1} Q_i^k - Q_j^{k+1} \right)\Big|_T \Delta T \tag{5}$$

In general, changes to external conditions influence flow pattern and sediment transport process in a reservoir or river channel, which may cause imbalanced sediment transport and scouring or deposition. Riverbed aggradation or degradation, in turn, affects flow and sediment transport. Sediments serve as a "bridge" in the variations of flow and sediment transport and between balanced and imbalanced sediment transport. Therefore, riverbed aggradation/degradation need to be considered in computations of reservoir water and sediment outflows. If accumulated reservoir deposition is used to express reservoir morphometry, reservoir outflow function is

$$Q_w|_T = f_w(V_w|_T,\ V_s|_T \ldots) \tag{6}$$

Similarly, sediment out flow at each of the nodes (reservoir) is a function of sediment inflow and reservoir geometry,

$$Q_s|_T = f_s(Q'|_T,\ V_s|_T \ldots) \tag{7}$$

in which $Q' = \sum_{i=1}^{N_k} \omega_{i,j}^{k+1} Q_i^k$.

When reservoir aggradation/degradation is insignificant, reservoir outflow may be considered as a nonlinear function of reservoir storage, i.e., the relation between outflow and storage at time step T becomes:

$$Q_w|_T = f_w(V_w|_T) \tag{8}$$

Similarly, sediment outflow function can be simplified as

$$Q_s|_T = f_s(Q_s'|_T) \tag{9}$$

The water and sediment functions vary from watershed to watershed, depending on the physical and geometrical characteristics of the catchment and the inherent hydrological and sediment properties of the river basin/system. Empirical relations or regressions to measured data can be used to determine the functions, considering the requirements in the mathematical schemes or algorithms of the ANN model. In this study, the functions are represented in the ANN model by empirical expressions (Yang and others 1998) in the form of $1/[1 + \exp(-Ax)]$, where A is the calibration coefficient and x stands for V_w, V_s, Q', and Q_s'. The coefficient A in the empirical water and sediment equations was determined by measured flow and sediment data, respectively.

ANN Model for a River System

In the ANN river system model, the mass balance for inflow and outflow relations at a node (j) in the first layer (Figure 2) is expressed as

$$Q_j^1 = Q_{j,in}^1 \tag{10a}$$

$$\sum_{i=1}^{N_1} \omega_{i,j}^2 = 1 \tag{10b}$$

where $Q_{j,in}^1$ is inflow (i.e., water or sediment source) to node j in layer 1 and Q_j^1 is outflow from node j in layer 1.

For the internal nodes or reservoirs in the ANN model (Figure 2), the input–output relations are

$$Q_{j,in}^{k+1} = \sum_{i=1}^{N_k} \omega_{i,j}^{k+1} Q_i^k \tag{11a}$$

$$Q_j^{k+1} = f(V,\ Q') \tag{11b}$$

$$\sum_{i=1}^{N_{k+1}} \omega_{i,j}^{k+2} = 1 \tag{11c}$$

For the output nodes (Figure 1), the relations are

$$Q_{j,in}^K = \sum_{i=1}^{N_{k-1}} \omega_{i,j}^K Q_i^{K-1} \tag{12a}$$

$$Q_j^K = Q_{j,in}^K \tag{12b}$$

where K is the total number of layers.

The ANN model for water and sediment transport simulations is constructed with the integration of mass conservation principles into a BP ANN, including water and sediment continuity equations 1 and 2, outflow-storage functions, equations 8 and 9, and the input–output relations, equations 10–12. Although the algorithm in this model is similar to that in traditional ANN, the incorporation of the water and sediment conservation principles gives the model some new features. The interaction between water sediment and reservoir aggradation/degradation has an impact on water and sediment outflow–storage functions. The intertwining river system is reasonably represented by the ANN model for their similarity in structures of both networks. Connections between adjacent layers are described by parameters that have physical meaning. Mass conservation is satisfied at all nodes and over the whole river system. In this flow- and sediment-based ANN river system model, the impact of upstream and downstream sediment deposition conditions on reservoir storage and deposition is simulated; the traditional BP ANN can not deal with this. The model may be simplified to the traditional BP model if the storage variation term, $\delta V_j^{k+1}/\delta T$, in the mass conservation equation 3 is dropped.

Application

Study Site

The study site is the middle (Jingjiang) reach of Yangtze River and Dongting Lake in the center of Hunan province (Figure 3). The plain area of the middle Yangtze River is a heavily populated region with rapidly developing transportation and economy. Dongting Lake is located on the south side of Jingjiang reach of the Yangtze River and used as a detention basin for attenuating flood peaks in Yangtze River. The Yangtze River flow enters Dongting Lake through three diversion works at Songzi, Taiping, and Ouchi. The lake has four tributaries from the southwest, Xiang, Zi, Yuan, and Li. It is divided into three regions (west, east, and south) with flood pathways connecting them. According to historical records (1951–1988 hydrological data), long-term annual inflow is 3.018×10^{11} m^3 including 1.12×10^{11} m^3 from three flood diversion inlets (37.1% of the total runoff), 1.647×10^{11} m^3 from the tributaries (54.6%), and 0.251×10^{11} m^3 of surface runoff or overland flow from the surrounding area of the lake (8.3%). Three cutoffs (1967–1972) along the Jingjiang reach have changed the flow and sediment diversions at the three inlets and influenced flow and sediment transport processes in and exchange between Jingjiang reach and the lake.

Some of the Yangtze River water (Yichang plus Qingjiang) stays in its main channel, Jingjiang reach. The rest enters the lake through the diversion works, meeting water from the tributaries in the lake, which returns to the main channel of the Yangtze River at Chenglingji after routing through flood pathways at the west and south regions of the lake. The Dongting basin is a radiating river network with the lake at its center. Morphological and meteorological conditions vary significantly from watershed to watershed. During wet seasons, storm centers move around in an area of 1.0×10^6 km^2, resulting in water rises in the upstream rivers or tributaries and elevating flood peaks in the lake due to superposition of arriving floods. The flood season lasts for almost eight months (March–October) with four months of high flood flows (May–August). The complexity and intertwining nature of the river system causes frequent flooding. To identify the impact of water and sediments in the tributaries on the flow and sediment transport processes in the lake, the tributaries need to be treated as independent elements in the river network and each has its individual role in the network. For simplicity and convenience, overland runoff is treated as part of the flow in the tributaries and distributed according to the ratio of tributary flow to the total flow.

Model Structure

The study area is divided into 17 interconnected regions and flood waterways in parallel or series, forming a network. Table 1 lists nodes used in the ANN model and corresponding river regions in the river system. In the ANN model interactions exist between adjacent layers of nodes. However, adjacent regions in the river network may not physically have any connections. Therefore, in this paper, the weights are set at zero for adjacent regions that have no interactions. For instance, region 3 has no water exchange with regions in layer 3 except region 11. In addition, the three diversion flows are combined to one for simplicity. The interaction and exchanges between the Jingjiang River and Dongting Lake are characterized by a network as shown in Figure 4. Based on the representative river network, the ANN network for water and sediment transport simulations has six layers of nodes and reservoirs. There are five nodes in the first layer, Yichang and Qingjiang inflow and four tributaries. The second to fifth layers have six, seven, three, and two reservoirs, respectively. The sixth layer has one node, i.e., outflow at Luoshan Station. As indicated in the structure of the ANN river system model (Figure 4), the number of input nodes is the number of factors influencing output variables, which in turn equals to the number of output nodes.

The simplified pattern of the river network can be considered as the appropriate ANN architecture. The reasons are that it is a physical representative of the river net and that mass conservation equations and flow–storage–sediment functions are incorporated into the model as the transfer functions, which describe the behavior and conditions of the system. It is convenient and reasonable to construct the ANN network (Figure 4) by considering the structural pattern of the river system due to the similarity between them. The use of a river system for the ANN structure makes it unnecessary to carry out many experiments and time-consuming trials, as performed in previous investigations.

The number of hidden units in the internal layers, which can be adjusted according to the river system

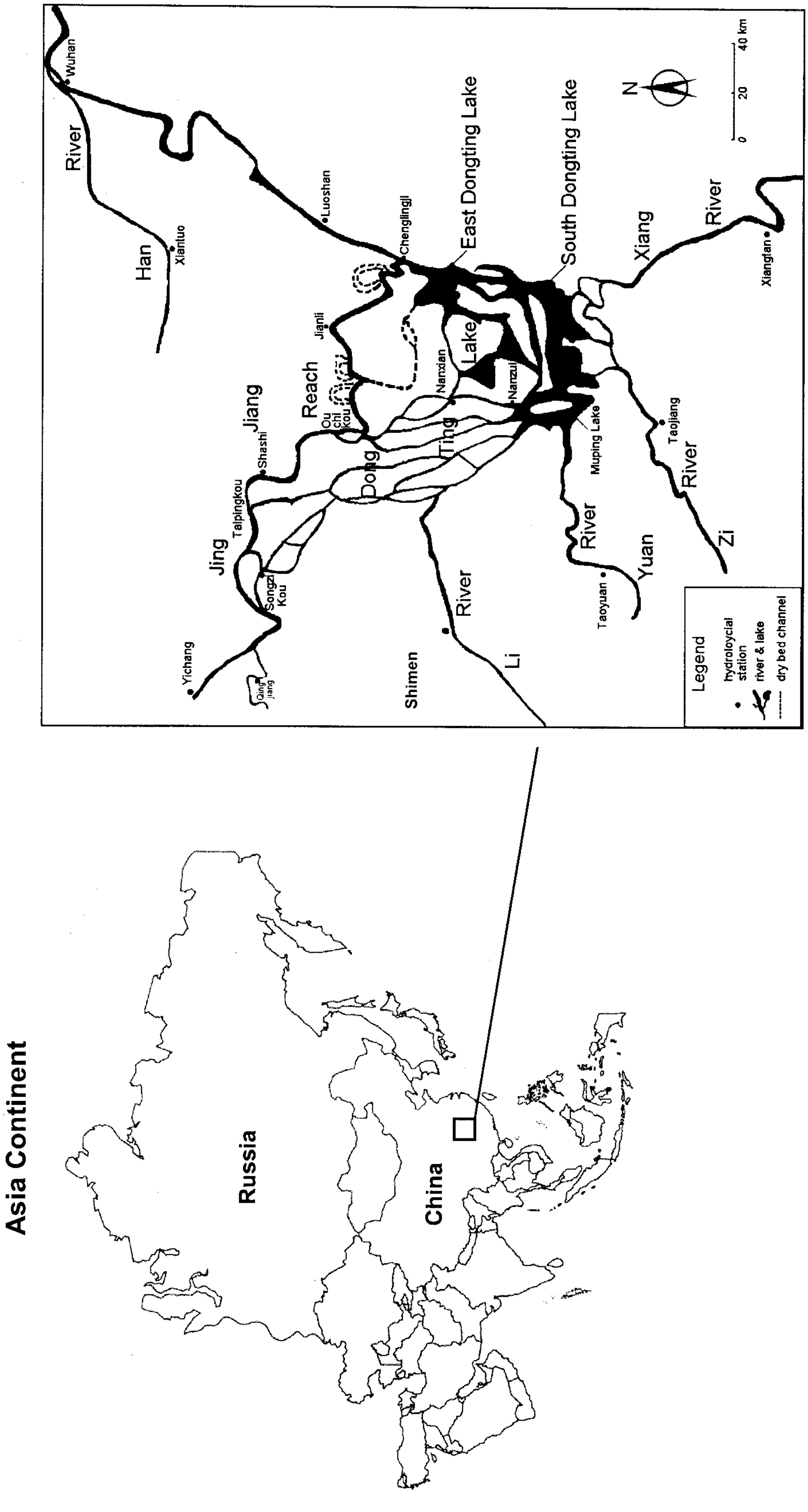

Figure 3. Study site—middle reach of the Yangtze River, China.

Table 1. Nodes or reservoirs in ANN model and regions in the river network

Node or reservoir	River region	Node or reservoir	River region
1	Zhijiang-Ouchikou	10	Songzi River (east and west tributaries), Dahu River, Zizhiju River, Guanwan River, Hudu River
2	Songzi River, MituoTemple-Zhonghe mouth	11	Lower reach of Xiang River
3	Upper reach of Xiang River	12	Lower reach of Zi River
4	Upper reach of Zi River	13	Lower reach of Yuan River
5	Upper reach of Yuan River	14	Taoxuan-Jianli
6	Upper reach of Li River	15	Zhuzi
7	Ouchi-Taoxuan	16	Region West Dongting Lake and tributaries and south Dongting Lake
8	West and middle tributaries of Ouchi River	17	Jianli-Luoshan
9	East tributaries of Ouchi River	18	East Dongting Lake to Chenglingji

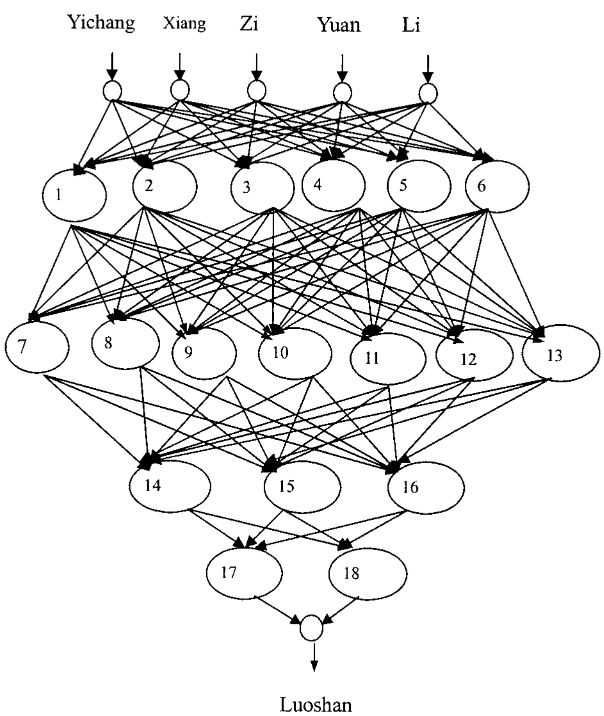

Figure 4. Network schematic of the middle reach of the Yangtze River.

being studied, determines the network's suitability and capability. They should be selected with care. If the number of hidden reservoirs is too small, the model's applicability to a wide range of problems may be compromised. If the number of hidden nodes or reservoirs is too high, it may result in "overtraining" or "overfit" of the model. The disadvantage of overtraining is that it tries to memorize the relationship accurately in the training phase, resulting in good performance during training, but poorer generalization capacities in the implementation phase. A large number of internal units may lead to excessively long model learning time and not necessarily give better accuracy or fewer errors. After a certain number of hidden nodes, the increase in the number of nodes may not contribute to better predictive capabilities of the model. Beyond a certain point, the addition of hidden nodes is governed by the law of diminishing returns for such problems (Anmala and others 2000). In previous studies, the selection of the appropriate number of hidden units in the internal layers, depending on project need and the number of input and output nodes, was made from experience and experiments because a useful analytical method was lacking.

Although the use of an actual river network in the selection of hidden nodes in internal layers eliminates the trial-and-error approaches used in the previous ANN applications, its appropriateness should be examined and verified by other approaches. An empirical formula was suggested by Yang and others (1998), as a reference, for computing the appropriate range of hidden units in an internal layer, $n_h = \sqrt{n_i + n_o} + a$, in which n_i is the number of input nodes, n_o is number of output nodes, and a is a constant in the range of 1–10. For the Jingjiang–Dongting water system considered in this study, the selected numbers of hidden units in the internal layers were based on the pattern of the actual river system as presented in Table 1 and Figure 3. The selected hidden unit numbers as indicated in Figure 4 fall within the range of values (3–12) computed by the empirical formula suggested by previous investigators (Yang and others 1998, Jain and Chalisgaonkar 2000). This comparison verifies that the selected numbers of internal layer nodes, by considering the actual river network, are appropriate.

Results and Discussion

A statistical parameter needs to be selected to measure agreement or disagreement between actual obser-

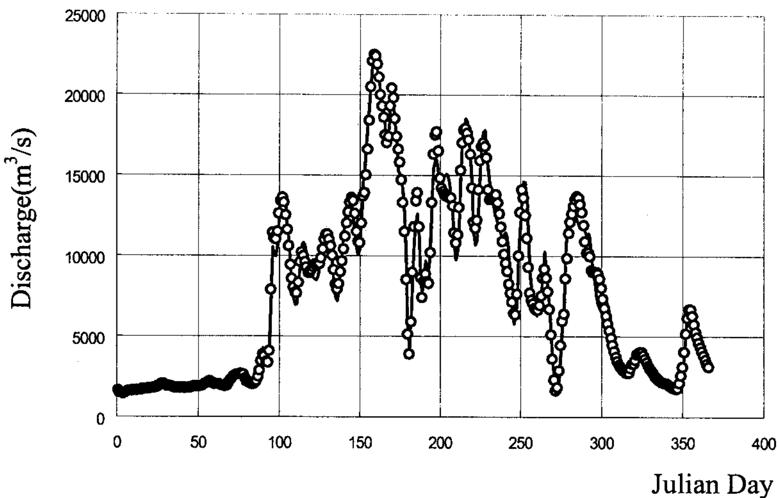

Figure 5. Measured (circles) and simulated (solid line) discharges at Chenglingji station (1984), $R^2 = 0.985$.

vation and model simulations. The goodness of fit between observed data and model results was primarily measured by a regression that is constrained through the origin. The overall performance of the ANN model was judged with respect to observed (actual) flow and sediment data on the basis of the coefficient of efficiency or regression coefficient, R^2, calculated as follows:

$$R^2 = 1 - \frac{s_e^2}{\sigma^2} \tag{13}$$

where s_e^2 is standard error of model predictions (sum of the square of errors) and σ^2 is standard deviation of the mean of the observed data, given as

$$s_e^2 = \frac{\sum_{m=1}^{M} (p_m - d_m)^2}{M} \tag{14}$$

and

$$\sigma^2 = \frac{\sum_{m=1}^{M} (d_m - \bar{d})^2}{M} \tag{15}$$

where d_m and p_m are observed data (target values) and model predictions respectively, M is the number of data points (p_m, d_m) and $\bar{d}$ is the mean of the observed values. A perfect match between observed data and model simulations is obtained when R^2 approaches 1.0 and s_e^2 goes to 0.0, which are the objective values for model training (or calibration) and validation.

The river system ANN model for hydrodynamics and sediment transport simulations was trained and tested separately. Decoupled water flow and sediment transport simulations are performed using different sets of data due to data availability for flows and sediments. Instream flow (daily discharge) and nonpoint source surface runoff data at Dongting Lake, Yichang station, four tributaries, and Luoshan station (outlet) are employed to train and test the ANN model developed in this study for flow simulations. Luoshan station is the closest downstream station to the outlet (near Chenglingji), which collects water and sediments from the entire study area. The flood travel time from Yichang to Luoshan is about 3 days and from the four tributaries about 2 days. In model training and validation, the inflow data (input variables) were observations on a different day from outflow data (output variables). The lead time of the forecast period (lag) is 2–3 days. The 1981–1983 flow data set is used for model training or calibration and the 1984 set for model verification. To avoid the problem of exceeding the upper and lower limits of the value range of the model units, the flow data were scaled so that the values of dimensionless variables lay within the range of 0.1–0.9 (Imrie and others 2000). With the weights obtained in the training phase (1981–1983), the performance of the ANN model was tested using the validation period data (1984).

The interaction and exchange between the Jingjiang reach of the Yangtze River and Dongting Lake are bridged by flow and sediment transport at the diversion openings. Therefore, to satisfactorily catch the developing and evolutionary process of the river reach, the variation patterns of flow and sediment transport at the diversion openings must be correctly reflected during ANN model training. An R^2 value of 0.995, the measure of model performance, was obtained for the training period 1981–1983 and 0.985 for the validation period of 1984. The verification result is presented in Figure 5,

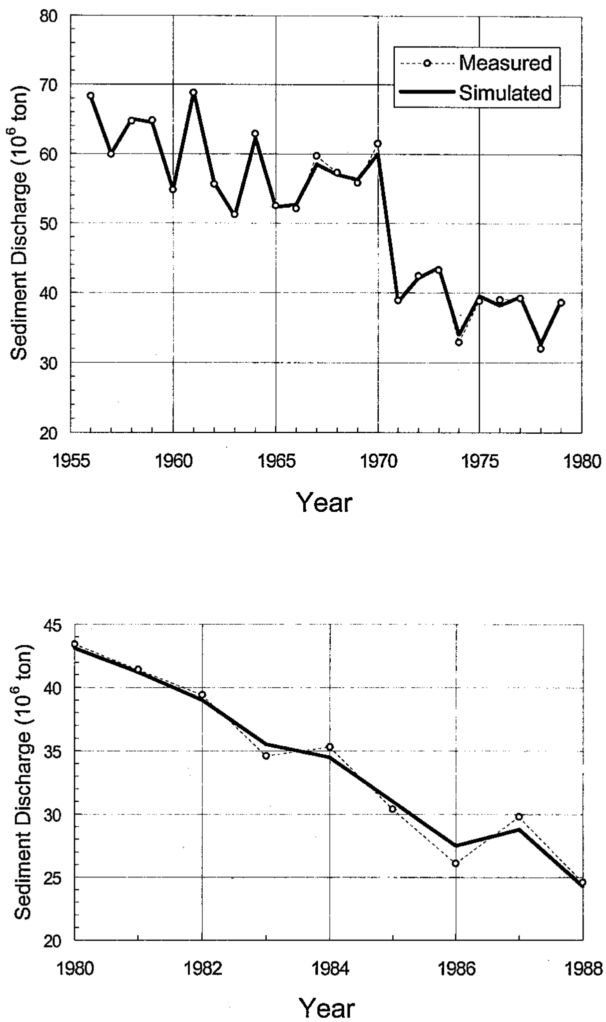

Figure 6. Training (top, R^2 = 0.997) and verification (bottom, R^2 = 0.986) results of the ANN river system model—measured and simulated sediment discharges at Chenglingji Station.

containing a plot of observed and computed daily discharges at Chenglingji station during the year of 1984. It is seen that the simulated hydrograph, using the ANN model developed in this study and trained with 1981–1983 flow data, agrees well with the field measurements. The plot shows a very good match, as indicated by an R^2 performance value of 0.985, except for some deviations near flow peaks during Julian days 180–270. The average relative error, i.e., the average of ratio of difference between measured and calculated to the measured, is less than 3%. The maximum relative error is less than 6%. It is demonstrated that the ANN model is capable of reproducing the flow process in a river network and can be effectively used for mapping flow variations.

Annual sediment discharges at Dongting Lake, the tributaries, and Chenglingji station for the period 1956–1979 are used to train the model for sediment transport simulations by calibrating the weights (ω) in the ANN model. Model verification is conducted with the 1980–1988 annual sediment discharge data. The results of model training and verification for sediment transport are displayed in Figure 6. The simulated sediment discharges at the Chenglingji station during the calibration period of 1956–1979 are generated after the calibration of weights in the model. The model duplicates the measured variation of annual sediment discharge very well. The performance measure, R^2 value, is 0.997 for the calibration period (1956–1979) and 0.986 for the validation period (1980–1988). The relative errors in the result from model testing or validation with the 1980–1988 data set are within a satisfactory or required range of accuracy, demonstrating the model's usefulness and adequacy in simulating and analyzing flow and sediment transport processes in the Dongting area of the Yangtze River. Based on the results, it is clear that the ANN technique is capable of modeling the flow and sediment transport behavior of a river system.

Conclusions

In the simulations of a complex river network, the successful employment of deterministic models or hydrodynamic models is restricted by a need for a great amount of site-specific data. It is impractical to use a traditional hydrodynamic model for quick and simple simulations and real-time predictions, especially when a complete set of data is not available. Considering the similarity in structure and in the input–output relation between a neural network and a river system, the use of an ANN approach for simulating water and sediment motions in a river network is adequate and practical. The ANN technique has been used to establish an integrated river system simulation model. A significant advantage of using the ANN technique is that it can successfully model the unsteady flow behavior and sediment transport in a complex system of rivers.

An ANN river system model has been developed by integrating the mass continuity equations and storage functions into a back-propagation ANN. The encapsulation of equations describing physical processes in the ANN model as well as the similarity in structures of a river net and an ANN configuration overcomes the difficulty in selecting the appropriate network architecture. The use of the actual river net for configuring the ANN avoids the trial-and-error procedure for choosing the optimal number of units in each hidden layer. Connections between adjacent layers are described by parameters that have physical meaning. Incorporating the mass conservation as the activation function makes the ANN model no longer a "black box." The method and results presented demonstrated that it is practical and possible to integrate fundamental physical principles into a data-driven modeling technique and to use a natural system for ANN construction.

The model has been applied to the middle reach of the Yangtze River, i.e., Jingjiang reach and Dongting Lake. Model results and field measurements agree well. It is concluded that the ANN river system model is capable of describing flow and sediment transport processes in a river system of interconnected channels and lakes. Given upstream inflows, the physically based ANN river system model, which was trained with a wide range of data, can be used to simulate and predict downstream discharges and sediments. The simplicity of the model makes it possible and practical to conduct quick, real-time predictions during floods. Nevertheless, it should be noted that the ANN technique is most suitable for climatically, hydrologically, and morphologically stationary systems. Significant deviations from conditions the ANN is trained for, such as changes in rainfall–runoff relations or stream avulsions, may lead to inaccurate predictions with ANN parameters trained on historic data.

Acknowledgments

This paper is based on the results of a key research project (No. 59890200) funded by the National Natural Science Foundation of China and completed at the Key Laboratory of Water and Sediment Sciences of Ministry of Education of China (Wuhan University) in cooperation with the visiting scholar program sponsored and supported by the Ministry of Education of China.

Literature Cited

Aly, A. H., and R. C. Peralta. 1999. Optimal design of aquifer cleanup systems under uncertainty using a neural network and a genetic algorithm. *Water Resources Research* 35(8): 2523–2532.

Anmala, J., B. Zhang, and R. S. Govindaraju. 2000. Comparison of ANNs and empirical approaches for predicting watershed runoff. *Journal of Water Resource Planning and Management* 126(3):156–166.

ASCE (Task Committee on Application of Artificial Neural Networks in Hydrology). 2000a. Artificial neural networks in hydrology. I: Preliminary concepts. *Journal of Hydrologic Engineering* 5(2):115–123.

ASCE (Task Committee on Application of Artificial Neural Networks in Hydrology). 2000b. Artificial neural networks in hydrology. II: Hydrologic applications. *Journal of Hydrologic Engineering* 5(2):124–137.

Bishop, C. M. 1995. Neural networks for pattern recognition. Clarendon Press, Oxford.

Carrieve, P., S. Mohaghegh, and R. Gaskari. 1996. Performance of a virtual runoff hydrograph system. *Journal of Water Resource Planning and Management* 122(6):421–427.

Chang, A. T., and I. Tsang. 1992. A neural network approach to inversion of snow water equivalent from passive microwave measurements. *Nordic Hydrology* 23:173–182.

Dawson, C. W., and R. Wilby. 1998. An artificial neural network approach to rainfall-runoff modeling. *Hydrological Science Journal* 31(1):47–66.

Dibike, Y. B., and M. B. Abbott. 1999. Application of artificial neural networks to the simulation of two dimension flow. *Journal of Hydraulic Research* 37(4):435–446.

Dibike, Y. B., D. Solomatine, and M. B. Abbott. 1999. On the encapsulation of numerical-hydraulic models in artificial neural network. *Journal of Hydraulic Research* 37(2):147–161.

Elman, J. L. 1990. Finding structure in time. *Cognitive Science* 14:179–211.

Fahlman, S. E. 1989. Fast-learning variations on back-propagation: an empirical study. Pages 38–51 *in* D. Touretzky, G. Hinton, and T. Sejnowski (eds.), Proceedings, 1988 connectionist models summer school. Morgan Kaufmann Publishers, San Mateo, CA.

Fahlman, S. E., and C. Lebiere. 1990. The cascade-correlation learning architecture. Technical report CMU-CS-90–100, School of Computer Science Carnegic-Mellon University, Pittsburgh.

Fausett, L. 1994. Fundamental of neural networks. Prentice-Hall, Englewood Cliffs, New Jersey.

French, M. N., W. F. Krajewski, and R. R. Cuykendall. 1992. Rainfall forecasting in space and time using a neural network. *Journal of Hydrology* 137(1–4):1–31.

Gallant, A. R., and H. White. 1992. On learning the derivatives of an unknown mapping with multiplayer feedforward networks. *Neural Networks* 5:129–138.

Haykin, S. 1994. Neural networks: a comprehensive foundation. Macmillan, New York.

Hornik, K. 1989. Multilayer feedforward networks are universal approximators. *Neural Networks* 2:359–366.

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

Imrie, C. E., S. Durucan, and A. Korre. 2000. River flow prediction using artificial neural networks: generalization beyond the calibration range. *Journal of Hydrology* 233(1–4): 138–153.

Jain, S. K., and D. Chalisgaonkar. 2000. On setting up stage-discharge relations using ANN. *Journal of Hydrologic Engineering* 5(4):428–433.

Karnin, E. D. 1990. A simple procedure for pruning back-propagation trained neural networks. *IEEE Transactions, Neural Networks* 1:239–242.

Karunanithi, N., W. J. Grenney, D. Whitley, and K. Bovee.

1994. Neural networks for river flow prediction. *Journal of Computing in Civil Engineering* 8(2):201–220.

Minns, A. W., and M. J. Hall. 1996. Artificial neural networks as rainfall runoff models. *Hydrological Science Journal* 41(3): 399–418.

Raman, H., and N. Sunilkumar. 1995. Multivariate modeling of water resources time-series using artificial neural networks. *Hydrological Science Journal* 40(2):145–163.

Rogers, J. L., and W. J. Lamarsh. 1992. Application of a neural network to simulate analysis in an optimization process. *Proceedings, Artificial Intelligence in Design '92,* Kluwer Academic, Boston, pp. 739–754.

Rumelhart, D. E., J. L. McLelland, and PDP Research Group. 1986. Parallel distributed processing, explorations in the micro structure of cognition, Vol. I: Foundations, MIT Press, Cambridge, Massachusetts.

Russell, S. J., and P. Norvig. 1995. Artificial intelligence: a modern approach. Prentice-Hall, Englewood Cliffs, New Jersey.

Smith, J., and R. N. Eli. 1995. Neural network models of rainfall-runoff process. *Journal Water Resource Planning and Management.* 121(6):499–508.

Tchaban, T., M. J. Taylor, and J. P. Griffin. 1998. Establishing impacts of the inputs in a feedforward neural network. *Neural Computing Application* 7:309–317.

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

Weigend, A. S., D. E. Rumelhart, and B. A. Huberman. 1990. Predicting the future: a connectionist approach. *International Journal of Neural Systems* 1:193–203.

White, H. 1990. Connectionist nonparametric regression: multiplayer feed-forward networks can learn arbitrary mappings. *Neural Networks* 3:535–549.

Yang, R., J. Ding, and G. Liu. 1998. Preliminary study of hydrology-based artificial neural network. *Journal of Hydraulics* 8:23–27.

Zhang, Q., and S. J. Stanley. 1997. Forecasting raw-water quality parameters for the North Saskatchewan River by neural network modeling. *Water Research* 31:2340–2350.

Zhu, M., M. Fujita, and N. Hashimoto. 1994. Application of neural networks to runoff prediction. Pages 205–216 *in* K. W. Hipel and others (eds.) Stochastic and statistical method in hydrology and environmental engineering, Vol. 3. Kluwer Academic, Dordrechy, The Netherlands.

Appendix 1. Notation

A = calibration coefficient
C = sediment concentration
d = observed data
f = function
n = total number of data points
N = total number of reservoirs in a layer
P = model prediction (simulated data)
Q = water discharge or sediment transport rate
R^2 = coefficient measuring model performance, i.e., regression coefficient
S_2^e = standard error of model predictions
T = time
V = water storage or sediment deposition
x = V_w, V_s, Q', or Q'_s
ω = weight
σ^2 = standard deviation of the mean of the observed data

Subscripts

i = index for reservoirs in a previous layer
j = index for reservoirs in the current layer
m = index for data points
s = sediment
w = water
in = inflow

Superscripts

k = index for layers of reservoirs