Journal of Hydrology 265 (2002) 225–245

Journal of Hydrology

www.elsevier.com/locate/jhydrol

# River flow forecasting: use of phase-space reconstruction and artificial neural networks approaches

B. Sivakumar*, A.W. Jayawardena, T.M.K.G. Fernando

*Department of Civil Engineering, The University of Hong Kong, Pokfulam Road, Hong Kong, People's Republic of China*

Received 3 July 2001; revised 1 May 2002; accepted 7 May 2002

**Abstract**

The use of two non-linear black-box approaches, phase-space reconstruction (PSR) and artificial neural networks (ANN), for forecasting river flow dynamics is studied and a comparison of their performances is made. This is done by attempting 1-day and 7-day ahead forecasts of the daily river flow from the Nakhon Sawan station at the Chao Phraya River basin in Thailand. The results indicate a reasonably good performance of both approaches for both 1-day and 7-day ahead forecasts. However, the performance of the PSR approach is found to be consistently better than that of ANN. One reason for this could be that in the PSR approach the flow series in the phase-space is represented step by step in local neighborhoods, rather than a global approximation as is done in ANN. Another reason could be the use of the multi-layer perceptron (MLP) in ANN, since MLPs may not be most appropriate for forecasting at longer lead times. The selection of training set for the ANN may also contribute to such results. A comparison of the optimal number of variables for capturing the flow dynamics, as identified by the two approaches, indicates a large discrepancy in the case of 7-day ahead forecasts (1 and 7 variables, respectively), though for 1-day ahead forecasts it is found to be consistent (3 variables). A possible explanation for this could be the influence of noise in the data, an observation also made from the 1-day ahead forecast results using the PSR approach. The present results lead to observation on: (1) the use of other neural networks for runoff forecasting, particularly at longer lead times; (2) the influence of training set used in the ANN; and (3) the effect of noise on forecast accuracy, particularly in the PSR approach. 

*Keywords:* River flow; Forecasting; Phase-space reconstruction; Artificial neural networks; Local and global approximations; Number of variables; Noise

## 1. Introduction

Accurate forecasting of river flow dynamics is necessary for, among others: (1) optimal design of water storage and drainage networks; and (2) management of extreme events, such as floods and droughts. However, the problem is non-trivial, because: (1) the various physical mechanisms governing the river flow dynamics act on a wide range of temporal and spatial scales; and (2) almost all mechanisms involved in the river flow process present some degree of non-linearity.

During the past few decades, a great deal of research has been devoted to the modeling and forecasting of river flow dynamics. Such efforts have led to the formulation of a wide variety of approaches and the development of a large number of

* Corresponding author. Address: On leave from: Department of Land, Air and Water Resources, University of California, Davis, CA 95616, USA. Fax: +1-530-752-5262.

*E-mail address:* sbellie@ucdavis.edu (B. Sivakumar).


PII: S0022-1694(02)00112-9

models. The existing models for river flow forecasting may broadly be grouped under two main categories: (1) physically-based models; and (2) black-box models. The physically-based models are specifically designed to mathematically simulate or approximate (in some physically realistic manner) the general internal sub-processes and physical mechanisms that govern the river flow process, whereas the black-box models are designed to identify the connection between the inputs and the outputs, without going into the analysis of the internal structure of the physical process.

While the physically-based models are very useful to our understanding of the physical mechanisms involved in the river flow (or any other hydrological) process, unfortunately, they also possess great application difficulties, essentially for the following reasons: (1) they require a large number of parameters for modeling the complexity of river flow dynamics; and (2) extension of a particular model to even slightly different situations is very difficult. The black-box models, on the other hand, though may not necessarily lead to a better understanding of the river flow process (in a physically realistic manner), have an advantage in that they are easier to apply for even different conditions since the modeling and forecasting procedure is usually analogous. Furthermore, the analysis of the characteristic parameters of the black-box models can furnish useful information on the dynamics of the phenomenon.

In the absence of accurate information about the physical mechanisms underlying or the 'exact' equations involved in the dynamics of river flow at a particular location, the use of a black-box model seems to have an edge over the use of a physically-based model, since the former is capable of representing arbitrarily the complex non-linear river flow process, by relating the inputs and the outputs of the underlying system. In view of this, the present study investigates the use of two non-linear black-box approaches: (1) the phase-space reconstruction (PSR) approach; and (2) the artificial neural networks (ANN) approach, for forecasting the river flow dynamics, and also compares their performances. Even though both the approaches employ a non-linear function to model the input–output relationship, they are different in the way they subdivide the function domain (or phase-space) to perform the forecast. The PSR approach is a local approximation approach, where the domain is subdivided into many sub-domains or subsets, each of which identifies some approximations valid only in that subset. In this way, the system dynamics is represented step by step locally in the phase-space. The ANN approach, on the other hand, is a global approximation approach, in the sense that it uses all the values that were generated in the past as input for the forecast.

The use of these two approaches for studying the river flow (or any other hydrological) process is not new. Applications of these approaches to the river flow dynamics (either using the river flow series alone or using both rainfall and river flow series, i.e. rainfall-runoff modeling), in spite of their state of infancy and inherent limitations, have been on the rise since early last decade. Having said that, the ANN approach has been more popular, and the studies that have employed this approach (e.g. Karunanithi et al., 1994; Hsu et al., 1995; Minns and Hall, 1996; Fernando and Jayawardena, 1998; Jayawardena and Fernando, 1998; See and Openshaw, 1999; Zealand et al., 1999) have certainly outnumbered the ones that have employed the PSR approach (e.g. Jayawardena and Lai, 1994; Porporato and Ridolfi, 1997; Jayawardena and Gurung, 2000; Sivakumar et al., 2000, 2001). The outcomes of such studies are encouraging, as the two approaches have been found to be very useful in providing important information regarding the non-linear (dynamical) characteristics of the river flow and its predictability. The use and the validity of the PSR and other related concepts in hydrology are discussed in detail by Sivakumar (2000), whereas a review of the applications of the ANN approach in hydrology is made by Govindaraju (2000).

An important observation that can be made from the studies conducted thus far on the applications of the PSR approach and the ANN approach for river flow forecasting is that each of these studies has been limited either to the application of the PSR approach alone or to the application of the ANN approach alone. Also, to the authors' knowledge, none of the river flow data sets used in the above studies has been studied by both the PSR and ANN approaches. Consequently, a comparison of the performances of the two approaches for forecasting river flow dynamics at a particular location could not be made, which, in turn, eliminated the possibility of choosing the appropriate (or better) approach for the river flow dynamics under investigation. This forms the basis for the present study. The study employs the above two approaches for forecasting the same river flow dynamics. The river flow dynamics chosen for the analysis is the daily river flow observed at the Nakhon Sawan gaging station in the Chao Phraya River basin in Thailand. The forecasting is made using only a scalar time series, i.e. the river flow series itself. The performance of the two approaches is investigated by forecasting the river flow series for two different lead times, i.e. 1 day and 7 days.

The organization of this paper is as follows. Section 2 starts with a brief introduction about the concept of PSR followed by a description of the local approximation forecasting approach. In Section 3, the fundamentals of the ANN and the forecasting approach therein are presented. Section 4, starting with the details of the Chao Phraya River basin study area and the data used, presents the analysis of the data using the two forecasting approaches and the results obtained. A comparison of the results obtained using the two approaches is also presented in Section 4. Important conclusions drawn from the present study as well as the potential areas for further research are presented in Section 5.

## 2. Phase-space reconstruction forecasting approach

The concept of phase-space is a useful tool for characterizing dynamical systems, such as the river flow system. According to this concept, a dynamical system can be described by a phase-space diagram, which is essentially a coordinate system, whose coordinates are all the variables that enter the mathematical formulation of the system (i.e. the variables necessary to completely describe the state of the system at any moment). The trajectories of the phase-space diagram describe the evolution of the system from some initial state (assumed to be known) and hence represent the history of the system. A point in the phase-space represents the state of the system at a given time. Phase-space is a powerful concept because with a model and a set of appropriate variables, dynamics can represent a real-world system as the geometry of a single moving point.

A common problem encountered while dealing with real systems, such as the river flow system, is the absence of information about all the variables involved in the underlying system. Under such circumstances, one way to represent the dynamics of the system is through the PSR, i.e. reconstruction (or embedding) of a single-dimensional (or variable) time series in a multi-dimensional phase-space. The physics behind such a reconstruction is that a non-linear system is characterized by self-interaction, so that a time series of a single variable can carry the information about the dynamics of the entire multi-variable system.

Among a variety of methods available for reconstructing the phase-space, the method of delays (e.g. Takens, 1981) is the most popular one. The method is based on the concept that, using its past history and an appropriate delay time, a scalar (or single-variable) time series $X_i$, where $i = 1, 2, \ldots, N$, can be reconstructed in a multi-dimensional phase-space to represent the underlying dynamics, according to

$$\mathbf{Y}_j = (X_j, X_{j+\tau}, X_{j+2\tau}, \ldots, X_{j+(m-1)\tau}) \quad (1)$$

where $j = 1, 2, \ldots, N - (m-1)\tau$, $m$ is the dimension of the vector $\mathbf{Y}_j$, called as embedding dimension; and $\tau$ is a delay time (Packard et al., 1980; Takens, 1981). A (correct) PSR in a dimension $m$ allows one to interpret the underlying dynamics in the form of an $m$-dimensional map $f_T$, that is

$$\mathbf{Y}_{j+T} = f_T(\mathbf{Y}_j) \quad (2)$$

where $\mathbf{Y}_j$ and $\mathbf{Y}_{j+T}$ are vectors of dimension $m$, describing the state of the system at times $j$ (current state) and $j+T$ (future state), respectively (In real situations, however, the optimal embedding dimension for reconstruction is not known a priori. In such cases, a trial and error procedure has to be adopted, i.e. by increasing the embedding dimension and selecting the one that gives the best (forecast) results). The problem now is to find an appropriate expression for $f_T$ (e.g. $F_T$).

There are several approaches for determining $F_T$, the most widely used one being the local approximation method proposed by Farmer and Sidorowich (1987). According to this method, the $f_T$ domain is subdivided into many subsets (neighborhoods), each of which identifies some approximations $F_T$, valid only in that subset and, hence, in this way, the system dynamics is represented step by step locally in the phase-space. The identification of the sets in which to subdivide the domain is done by fixing a metric $\| \|$ and, given the starting point $\mathbf{Y}_j$ from which the forecast is initiated, identifying neighbors $\mathbf{Y}_j^p$, $p = 1, 2, \ldots, k$, with $j^p < j$, nearest to $\mathbf{Y}_j$, which constitute the set corresponding to $\mathbf{Y}_j$. The local functions can then be built, which take each point in the neighborhood to the next neighborhood: $\mathbf{Y}_j^p$ to $\mathbf{Y}_{j+1}^p$. The local map $F_T$, which does this, is determined by a least squares fit minimizing

$$\sum_{p=1}^{k} \left\| \mathbf{Y}_{j+1}^p - F_T \mathbf{Y}_j^p \right\|^2 \tag{3}$$

In this study, the local maps are learned in the form of local polynomials (e.g. Abarbanel, 1996), and the forecasts are made forward from a new point $\mathbf{Z}_0$ using these local maps. For the new point $\mathbf{Z}_0$, the nearest neighbor in the learning or training set is found, which is denoted as $\mathbf{Y}_q$. Then the evolution of $\mathbf{Z}_0$ is found, which is denoted as $\mathbf{Z}_1$ and is given by

$$\mathbf{Z}_1 = F_q(\mathbf{Z}_0) \tag{4}$$

The nearest neighbor to $\mathbf{Z}_1$ is then found, and the procedure is repeated to forecast the subsequent values. The forecasting algorithm is implemented herein using the cspW software (Randle Inc., 1996).

## 3. Artificial neural networks forecasting approach

An ANN is a massively parallel-distributed information-processing system that has certain performance characteristics resembling biological neural networks of the human brain, where knowledge is acquired through a learning process and finding optimum weights for the different connections between the individual nerve cells (Haykin, 1994). The advantage of the ANN is that with no a priori knowledge of the actual physical process and, hence, the 'exact' relationship between sets of input and output data, if acknowledged to exist, the network can be 'trained' to 'learn' such a relationship. The ability to 'train' and 'learn' the output from a given input makes ANN capable of describing large scale arbitrarily complex non-linear problems.

A neural network is characterized by its architecture that represents the pattern of connection between nodes, its method of determining the connection weights, and the activation function (Fausett, 1994). A typical ANN consists of a number of nodes that are organized according to a particular arrangement. One way of characterizing ANNs is by the number of layers, as single-layer, bi-layer, and multi-layer. Another way of characterizing ANNs is based on the direction of information flow and processing, as feed-forward (where the information flows through the nodes from the input to the output side) and recurrent (where the information flows through the nodes in both directions). Among these combinations, the multi-layer feed-forward networks, also known as multi-layer perceptrons (MLPs), trained with a back-propagation learning algorithm have been found to provide the best performance with regard to input–output function approximation, such as forecasting. As the present study uses an MLP trained with a back-propagation algorithm for the purpose of river flow forecasting, the architecture of such a network is described here.

An MLP can have many layers. A typical MLP with one hidden layer is shown in Fig. 1. The first layer connects with the input variables and is called the input layer. The last layer connects to the output variables and is called the output layer. The layer in-between the input and output layers is called the hidden layer (there may be more than one hidden layer in an MLP). The processing elements in each layer are called nodes or units. Each of the nodes is connected to the nodes of neighboring layers. The parameters associated with each of these connections are called weights.

The architecture of a typical node (in the hidden or output layer) is also shown in Fig. 1. Each node $j$ receives incoming signals from every node $i$ in the previous layer. Associated with each incoming signal ($x_i$) is a weight ($w_{ji}$). The effective incoming signal ($s_j$) to node $j$ is the weighted sum of all the incoming signals

$$s_j = \sum_{i=0}^{n} w_{ji} x_i \tag{5}$$

The effective incoming signal, $s_j$, is passed through a

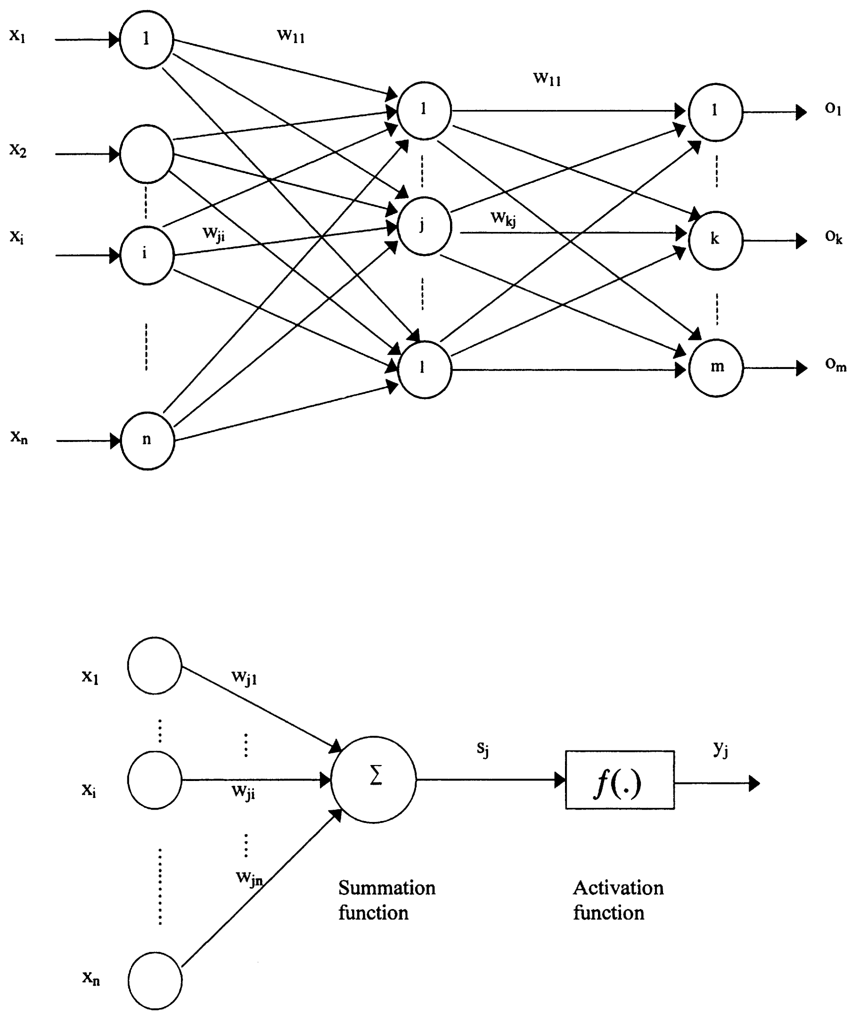


Fig. 1. Typical three-layer feedforward artificial neural network.

non-linear activation function (sometimes called a transfer function or threshold function) to produce the outgoing signal ($y_j$) of the node. The most commonly used function in an MLP trained with back-propagation algorithm is the sigmoid function. The characteristics of the sigmoid function are that it is bounded above and below, it is monotonically increasing, and it is continuous and differentiable everywhere (Hecht-Nielsen, 1990). The sigmoid function most often used for ANNs is the logistic function:

$$f(s_j) = \frac{1}{1 + \exp^{-s_j}} \quad (6)$$

in which $s_j$ can vary on the range $\pm\infty$, but $y_j$ is bounded between 0 and 1.

## 4. Analysis, results and discussion

### 4.1. Study area and data

In the present study, daily river flow series observed in the Chao Phraya River basin in Thailand is analyzed to study the use of the PSR and the ANN approaches for forecasting river flow dynamics. The Chao Phraya River basin, situated between 13.58–15.67°N and

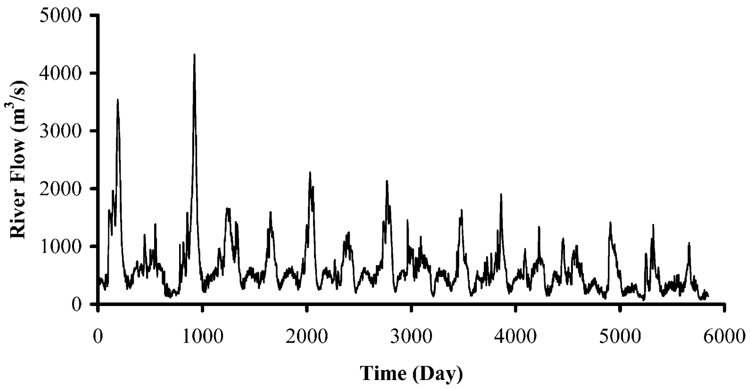


Fig. 2. Variation of daily river flow series at Chao Phraya River basin (day 1 corresponds to April 1, 1978).

100.10–101.00°E, is one of the most intensively monitored basins in Asia under the auspices of the global energy and water cycle experiment (GEWEX) and the GEWEX Asian monsoon experiment (GAME). Recent years have seen a large number of studies on the hydrology of this basin, notable among them are the studies by Oki et al. (1995), Wijesekera et al. (1995), and Raghunath et al. (2000).

Throughout the Chao Phraya River basin, flow is measured at a number of locations. In the present study, flow data observed at the Nakhon Sawan gaging station (Global River Flow Data Center station #2964100) is studied. This station is situated at 15.67°N and 100.12°E. The area of the basin at this station is 110,569 $km^2$ and, therefore, can be considered as a macro scale basin. For this station, daily river flow measurements are available from April 1978. The data set obtained, from the Global River Flow Data Center, Germany, for the present investigation contains the river flow data collected during the period from April 1978 to March 1994 (17 years), consisting of a total of 5844 data points. The time series plot of this river flow series is shown in Fig. 2. As can be seen in the figure, the river flow series exhibits significant variations, though an annual cycle seems to be evident. The analysis of this data series using the PSR approach and the ANN approach and the results obtained are presented below.

As both the PSR and the ANN approaches are black-box models, depending primarily on good training and learning of the data to establish relationships between the input and the output, it is imperative to select such a good training set from the available data series; this is particularly more important as far as the latter approach is concerned. The best way to achieve this goal seems to be to include all (or most of) the extreme events, such as very high and very low values (both qualitatively and quantitatively), in the training set. The inclusion of biased samples in the training set is also not recommended, as this will lead to a longer training time but not better results (and also may result in overtraining). A preliminary investigation of the river flow series (through a qualitative visual inspection of Fig. 2 as well as using quantitative basic statistical parameters) indicates that the most significant events (i.e. very high values and significant variations in the series) fall within the first quarter of the series. The latter part of the series (in particular the last one-third) seems to contain a lot of biased samples (i.e. with a large number of low flow values). The authors are not certain about the reasons for the significant change of trend (i.e. a decline) in the river flow values over time (i.e. very high flows and significant variations during the first few years and then somewhat consistently very low flows and less variations in the following years), but there may be two possible reasons for the above: (1) an increase in abstractions in the upstream locations for irrigation and other purposes; and (2) a change in basin regulations.

In view of the above observations, it is decided to use only the first 2550 data points for analysis in this study. Out of these 2550 points, the first 2150 points, which represent about 84% of the series, are selected as training set, whereas the remaining 400 points, accounting for about 16% of the series, are used for

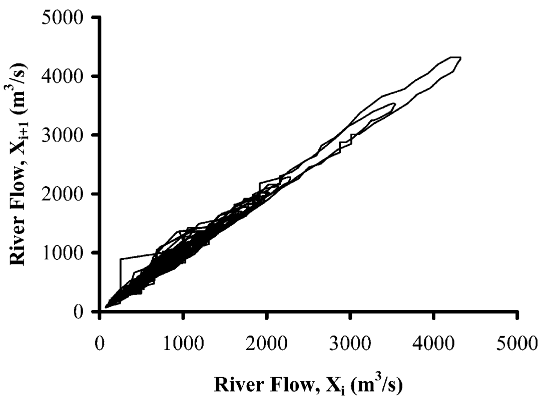


Fig. 3. Phase-space plot of daily river flow series at Chao Phraya River basin.

testing the forecasting performance of the PSR and ANN approaches. Also, in the case of ANN, the training set of 2150 points is further divided into two parts; training set and validation set. The first 1750 points are selected as the training set and the next 400 points are taken as the validation set. The choice of the length of the validation set is based on the recommendation to use about 10–20% of the training set (e.g. Kasabov, 1996). The performance of the PSR and the ANN approaches for forecasting the river flow series is tested by making forecasts for two different lead times, i.e. 1 day and 7 days.

Having said that, the consideration of (only) the first 1750 values of the river flow series for the purpose of training may raise serious questions for (at least) two reasons: (1) the 1750 values used for training happen to contain the highest recorded flow event and also exhibit significant variations; and (2) the testing set used (i.e. the latter part of the series) does not exhibit significant variations. The concern implied in these reasons is that the testing set is less variable and more predictable than the training set and, therefore, there may be a bias in the analysis (to show that the two approaches work well). Although, one cannot dismiss such a concern, the following points are also to be noted in order to understand why such a selection is made: (1) it is necessary to have a training set that could represent the overall structure of the flow series to capture the input–output relationship, which means that it is important to include the extreme events (this is particularly the case in the ANN approach); (2) since the objective is forecasting, it is important to capture the changes in the system with respect to time, which means that events from the first few years, for instance, should form the basis for the events that follow (this is particularly the case in the PSR approach); and (3) it may be necessary to have a reasonably long data set for training in order to sufficiently capture the dominant characteristics of the system under investigation (this seems to be the case particularly in the PSR approach, details of which are not reported herein, but can be found in, for instance, Sivakumar (2000)).

With respect to the above concern, however, unfortunately, the first few years of the river flow series used in this study happen to consist of the extreme events. Therefore, the selection of the data sets for training and testing is only logical, for the reasons also stated above. Having said that, it is also important to note that the compromise thus made in the selection of training and testing sets may also lead to inaccurate results, since the ANNs are not good extrapolators, and, therefore, one has to be careful in interpreting the results. However, the problem of extrapolation may be overcome to some extent by selecting a proper normalization range for the output variables, which is also done in the present study (see Section 4.3.1 for details).

It is important to note that the forecasting procedures in the PSR and the ANN approaches adopted herein with respect to the number of variables used as inputs are somewhat different. In the PSR approach, forecasting is made using lagged variables (or embedding dimensions) from 1 to 9. On the other hand, in the ANN approach, the analysis is started with 7 lagged variables (or input nodes), in order to reduce the computational time, and for each lead time, the optimal network structure is obtained by varying the number of nodes in the hidden layer. However, with the optimal network structure obtained in each case, a sensitivity analysis is performed to assess the importance of each input variable for forecasting.

The accuracy of forecasts is evaluated using a variety of (absolute and relative) error indicators, as follows: mean absolute error (MAE), mean square error (MSE), root mean square error (RMSE), maximum absolute error (MAXAE), minimum absolute error (MINAE), correlation coefficient ($r$), coefficient of determination ($R^2$), coefficient of

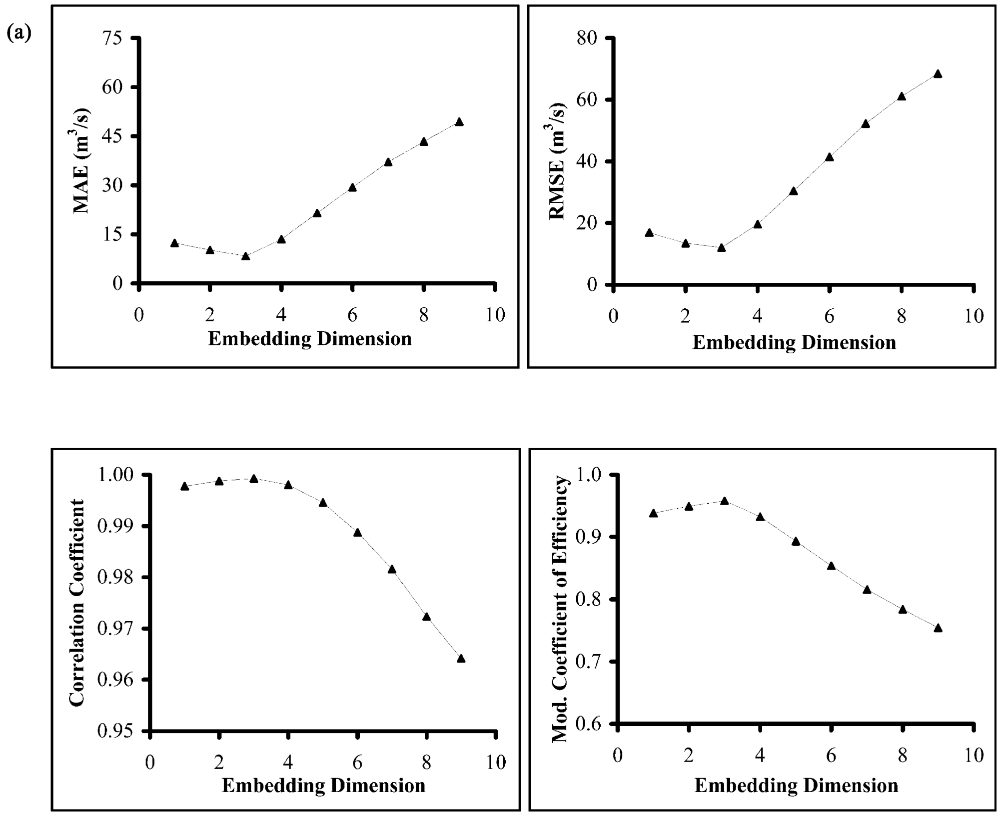


Fig. 4. PSR forecasting results for daily river flow series at Chao Phraya River basin: forecast accuracy versus embedding dimension for: (a) lead time = 1 day; and (b) lead time = 7 days.

efficiency ($E$), and modified coefficient of efficiency ($E_1$). Among these, the mean absolute error, the root mean square error, the correlation coefficient, and the modified coefficient of efficiency are considered the most important and, therefore, presented in all the forecasting cases that follow. These four error indicators are defined in Appendix A. The time series plots and the scatter diagrams are also used to choose the best forecasts among a large combination of results achieved with the different number of input variables (i.e. embedding dimensions in case of PSR approach and input nodes in case of ANN approach).

### 4.2. Phase-space reconstruction forecasting results

Fig. 3 illustrates how the scalar river flow series is reconstructed in a higher dimensional phase-space, according to Eq. (1), to represent the underlying dynamics. The figure presents the reconstruction of the series in a two-dimensional phase-space ($m = 2$), i.e. the projection of the attractor on the plane $\{X_i, X_{i+1}\}$, i.e. with $\tau = 1$. As can be seen, such a reconstruction yields a well-defined attractor for the river flow series, i.e. an attractor contained in a small region within the phase-space. The presence of such a well-defined attractor, even in just two dimensions, may be an indication of the possibility of obtaining reasonably good forecasts for the river flow series, especially with the use of local approximation methods.

The PSR approach with a local approximation forecasting method in the form of local polynomials, explained in Section 2, is now employed to forecast the river flow dynamics. For the data set mentioned above, embedding dimensions from 1 to 9 are used for the reconstruction purposes, and forecasts are made for 1-day and 7-day lead times.

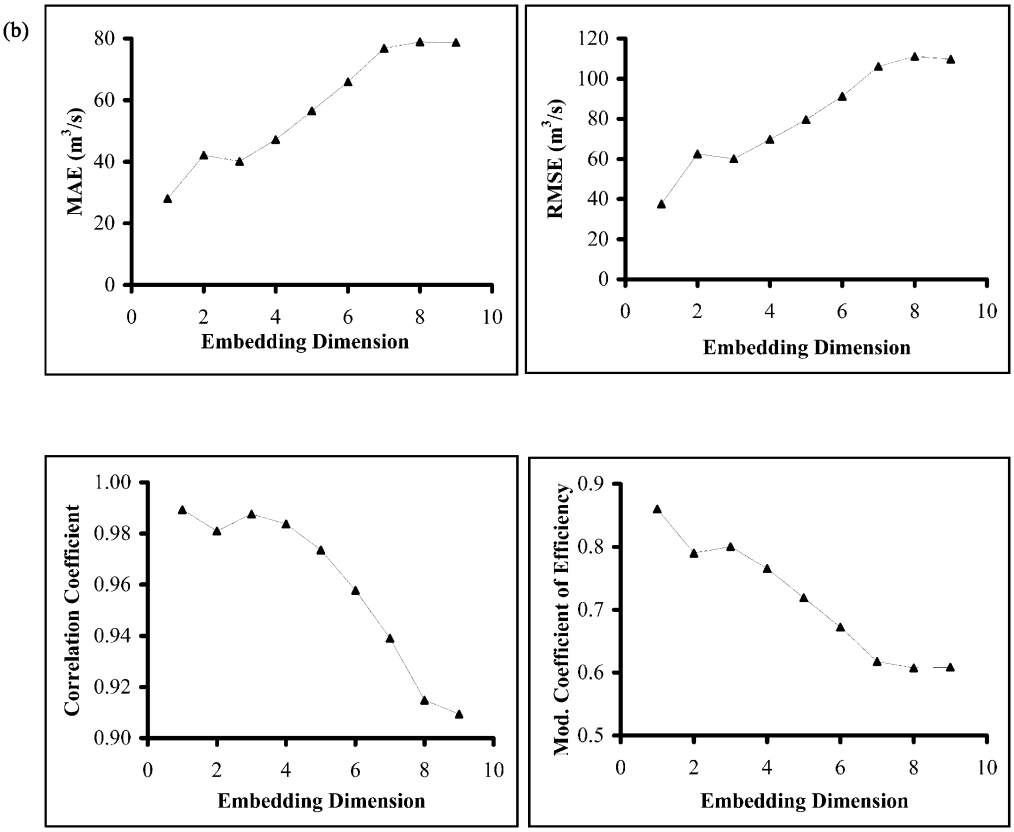


Fig. 4 (*continued*)

Fig. 4(a) presents a summary of the PSR forecasting results achieved for the river flow series when 1-day ahead forecasts are made, whereas the results of the 7-day ahead forecasts are summarized in Fig. 4(b). As the figures indicate, overall, reasonably good forecasts are achieved for the river flow series for both 1-day and 7-day lead times for all the embedding dimensions used for the PSR. However, the figures also reveal that the best forecasts for 1-day lead time are achieved (consistently in terms of the evaluation statistics) when the embedding dimension is three, i.e. $m_{\rm opt} = 3$ (with MAE = 8.379, RMSE = 12.058, $r = 0.99924$, $E_1 = 0.95832$) and for 7-day lead time are achieved when the embedding dimension is one, i.e. $m_{\rm opt} = 1$ (with MAE = 28.095, RMSE = 37.532, $r = 0.98392$, $E_1 = 0.86026$). In regards to $m_{opt}$, a comparison, using time series plots and scatter diagrams, of the observed values with the forecasted ones (obtained for the nine embedding dimensions) reveals that the best forecasts are indeed achieved when the embedding dimensions are three and one for 1-day and 7-day ahead forecasts, respectively, thus supporting the above observation.

Fig. 5(a) and (b) compare, using time series plots, the forecasted river flow values with the observed ones for 1-day and 7-day lead times, respectively (the ANN forecasted values are also presented therein, details of which will be presented later). The plots shown correspond to the results achieved with $m = 3$ and $m = 1$, respectively, i.e. the best forecasts achieved among the nine embedding dimensions used for the PSR. As can be seen, the forecasted values are, in general, in good agreement with the observed ones. A closer look at the two (i.e. observed and forecasted) time series (for both 1-day and 7-day lead times) reveals that the local approximation method with local polynomials not only very well captures the major trends in the river flow series but also reasonably

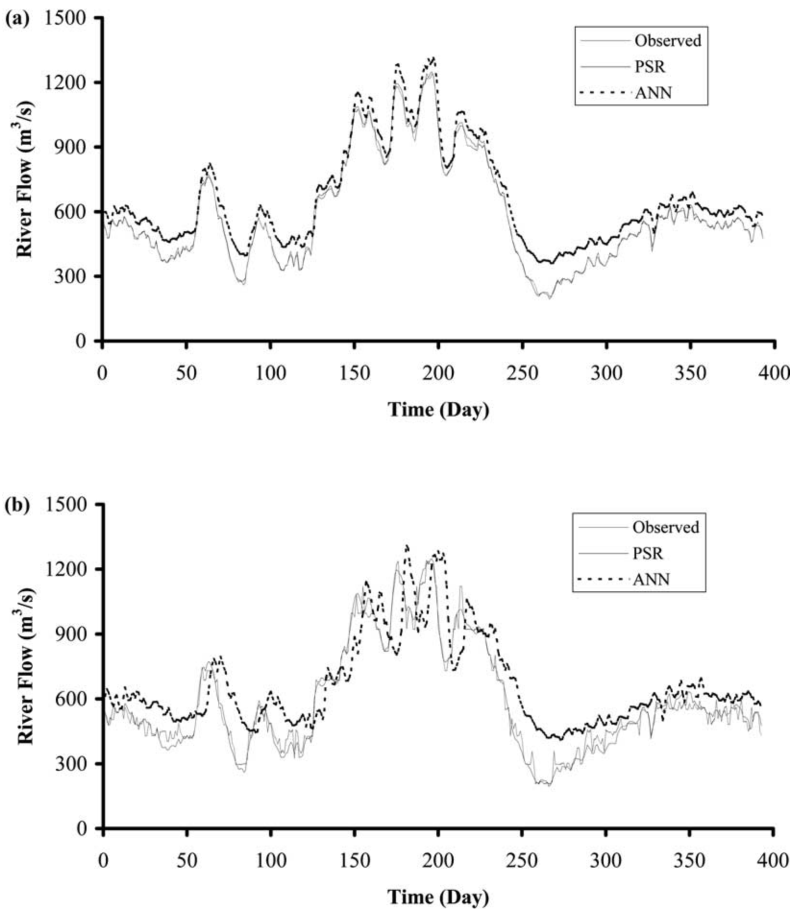


Fig. 5. Time series comparison of observed and PSR and ANN forecasted daily river flow series at Chao Phraya River basin for: (a) lead time = 1 day; and (b) lead time = 7 days.

preserves the minor (noisy) fluctuations. As can be seen, even extreme (i.e. very high) values are very well forecasted. The good agreement between the observed and forecasted series can also be revealed by plotting the scatter diagrams, shown in Fig. 6(a) and (b) for the two lead times, where the solid 1:1 (diagonal) line is plotted for reference. The time series plots, the scatter diagrams, the low MAE and RMSE values and the high $r$ and $E_1$ values clearly indicate the suitability of the PSR forecasting approach with the local approximation method for forecasting the river flow dynamics. The ability of the local approximation procedure in forecasting the river flow dynamics lies essentially in representing the dynamics captured in the phase-space step by step in local neighborhoods.

### 4.3. Artificial neural networks forecasting results

#### 4.3.1. Training and testing details

For the data set considered in the present study, the input variables as well as the target variables are first normalized linearly in the range of 0.1–0.9. This range is selected because of the use of the logistic function (which is bounded between 0.0 and 1.0) as the activation function for the output layer, i.e. Eq. (6). The normalization is done using the following equation

$$X_{\mathrm{norm}} = 0.1 + 0.8(X - X_{\min})/(X_{\max} - X_{\min}) \qquad (7)$$

where $X_{\min}$ and $X_{\max}$ are the minimum and maximum values in the data set, respectively. The synaptic weights of the networks are initialized with normally

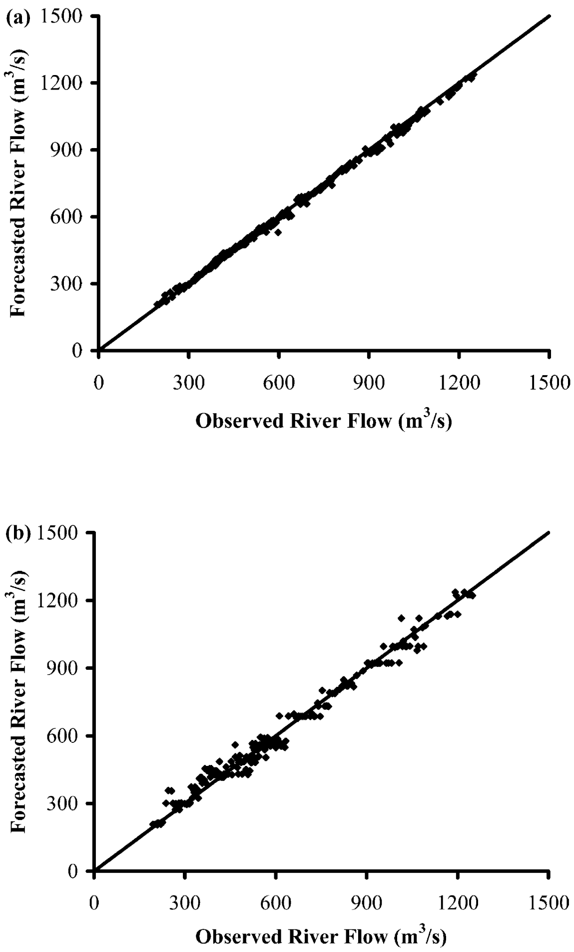


Fig. 6. Scatter plot of observed and PSR forecasted daily river flow series at Chao Phraya River basin for: (a) lead time = 1 day; and (b) lead time = 7 days.

distributed random numbers in the range of $-1$ to 1. The same initial weights are adopted for all the simulations in one set of simulations in order to make a direct comparison. The training is carried out in a pattern mode and the order of presenting the training samples to the network is also randomized from iteration to iteration. A learning rate of 0.1 and a momentum term of 0.3 are used for training, since these values have been found to yield better performance for this particular river flow series (see Jayawardena et al. (2000), Jayawardena and Fernando (2001) for details). Two stopping criteria, the cross validation and the fixed number of iterations, are adopted. The maximum number of iterations is set at 1000 and the training is continued for another 250 iterations after reaching a minimum in the validation set. The error reduction curves for the training and the testing sets are used to assess the convergence speed of the networks.

*4.3.2. Results of 1-day lead time forecasts*

For the 1-day lead time forecasts, the input layer consists of 7 nodes representing the daily river flow values at times $t$, $t-1,\ldots,t-6$ and the output layer consists of a single node representing the river flow value at $t+1$. The optimal number of hidden nodes is chosen based on a trial and error procedure, by varying the hidden nodes from 2 to 10 in steps of 2. With respect to the effect of number of hidden nodes, a significant increase in the learning speed is observed only when the number of hidden nodes is increased from 2 to 4, but not for further increase in nodes (i.e. 6, 8, and 10), as can be seen in Fig. 7. The mean square error of the training set (shown in Fig. 7) as well as the validation set (Figure not shown) is found to decrease gradually during training as expected and all the simulations are terminated after 1000 iterations. However, the rate of change of mean square error is found to be very small at the latter part of the training irrespective of the number of nodes in the hidden layer. The error measures for the training set and the testing set for these simulations (presented in Fig. 8, see below for details) demonstrate that over-fitting or over-training would not be a problem for a network with enough training samples, as suggested by Amari et al. (1997).

Fig. 8(a) presents a summary of the ANN forecasting results achieved for the river flow series when 1-day ahead forecasts are made. As the results indicate, overall, reasonably good forecasts are achieved for both the training and the testing sets for all the five different combinations of number of hidden nodes (the variations between the best case and the others are only marginal, i.e. less than 6%), even though the best results are achieved for the network with 4 hidden nodes.

A sensitivity analysis is now performed to evaluate the importance of each input parameter in training the network for making 1-day ahead forecasts. For this purpose, the trained weights of the network with structure 7-4-1 (input-hidden-output nodes) that yielded the best results (as discussed above) are

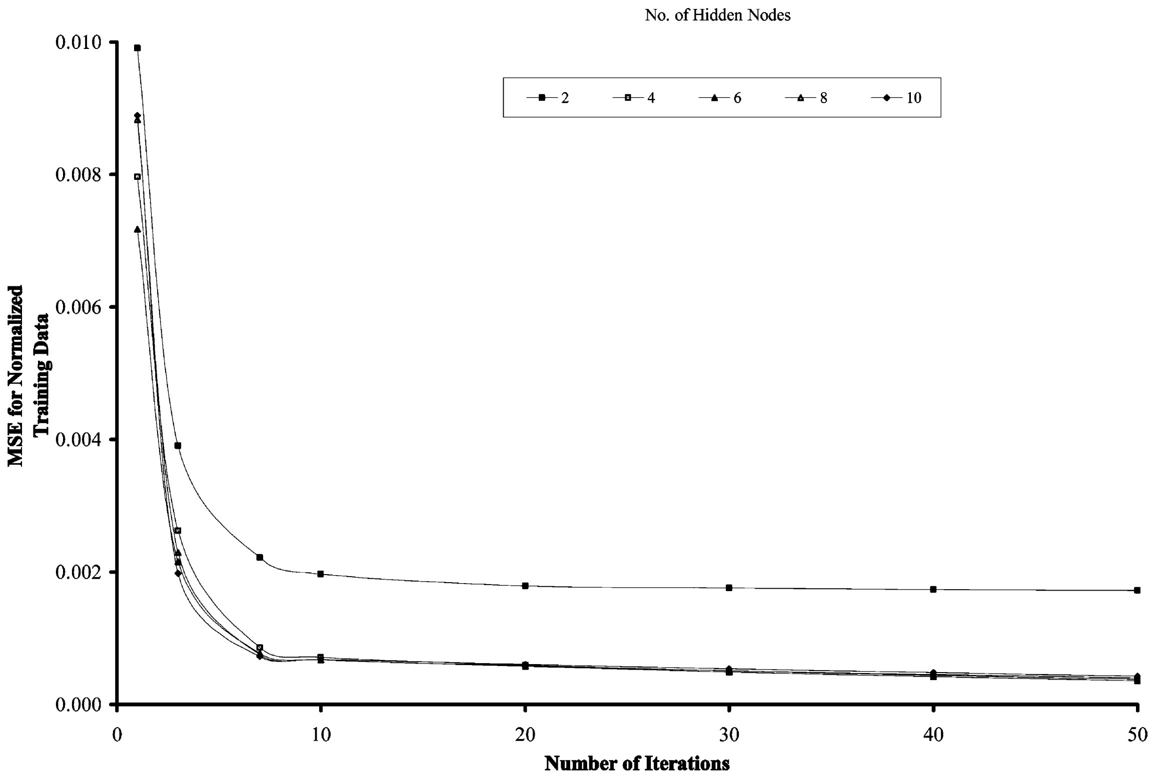


Fig. 7. Variation of mean square error of normalized training set during ANN training of the daily river flow series at Chao Phraya River basin with different number of nodes in the hidden layer: lead time = 1 day; input variables = 7.

used. The sensitivity analysis is carried out as follows. First, each input parameter is increased by 10 percent at a time. This gives 7 new input data series. These input series are used to make the corresponding forecasts using the trained network with original inputs. The forecasts obtained with the increased input variables are compared with the forecasts obtained with the original inputs. The sensitivity of the input variables is evaluated by computing the sensitivity indicators, i.e. sensitivity level (SL) and the absolute percentage variation of mean square error (APMSE). The results indicate that only the river flow value observed at $t$ has a significant impact on $t+1$ forecasts (with SL = 73.3, APMSE = 156.6). For the remaining input variables (i.e. river flow values observed at $t-1$ to $t-6$), the SL and APMSE are found to be considerably less, with the maximum SL of 7.9 and the maximum APMSE of 10.4 observed for the river flow value at $t-6$. These results seem to suggest that the use of lagged value at time $t$ might be sufficient for making forecasts at time $t+1$. Therefore, the next simulations are conducted with only one input variable, which is the river flow observed at time $t$.

For the one input case, the simulations are started with 10 hidden nodes and then gradually increased up to 60 hidden nodes. The simulation with 16 hidden nodes has an identical number of synaptic weights with the networks 7-4-1, which yielded the best results with 7 input variables, as presented earlier. All these simulations are terminated after 1000 iterations similar to the termination in the case of 7 input nodes. The error indicators for the training, validation and testing sets are presented in Fig. 8(b). As can be seen, in general, the accuracy of the forecasts increases with an increase in the hidden layer nodes up to 30, and then decreases with further increase in the hidden nodes. The network with 30 hidden layer nodes consistently results in the best forecasts. A

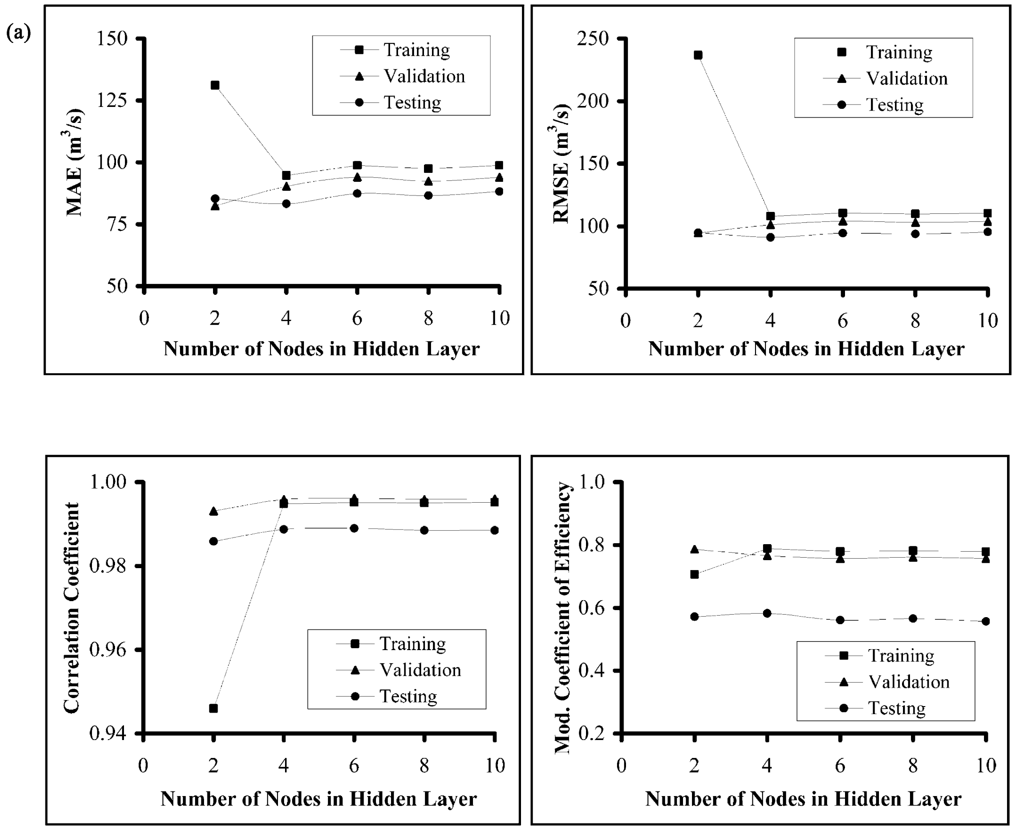


Fig. 8. Variation of error indicators for ANN training, validation, and testing of daily river flow series at Chao Phraya River basin for different number of nodes in the hidden layer: (a) lead time = 1 day, input variables = 7; (b) lead time = 1 day, input variables = 1; and (c) lead time = 1 day, input variables = 3.

relevant point to be noted here is that even this best result with the single input variable is worse than the worst result obtained with 7 input variables (Fig. 8(a)), clearly indicating that the use of only the river flow at time $t$ is not sufficient to make reliable forecasts for 1-day ahead, in spite of the fact that it is the only input variable that has a remarkable impact.

In view of the above, it is decided to assess the effect of the use of 3 input variables in forecasting, i.e. using river flow values at times $t$, $t-1$ and $t-2$ for forecasting the value at time $t+1$. In this case, the simulations are carried out with $5, 8, 10, 20, 40$, and 60 hidden layer nodes. All these simulations are terminated using the stopping criterion of maximum number of iterations. A summary of the error indicators is shown in Fig. 8(c). As can be seen, the accuracy of the forecasts increases with the increase in the number of hidden layer nodes up to 40 and then decreases with further increase in the number of hidden nodes. When the number of synaptic weights is identical with that of the network with 7 inputs, the results seem to be more or less the same. The results are further improved by increasing the hidden layer nodes and the best results are achieved with the network structure 3-40-1. These best results are a significant improvement over all the results achieved using different network structures. A time series comparison of these forecasted river flow values and the observed values is presented in Fig. 5(a), whereas Fig. 9 shows a scatter plot comparison of the same.

The number of iterations (i.e. 1000) used to carry out the above simulations seems to be sufficient and reliable as the rate of change of MSE of the normalized training set is found to be very small at

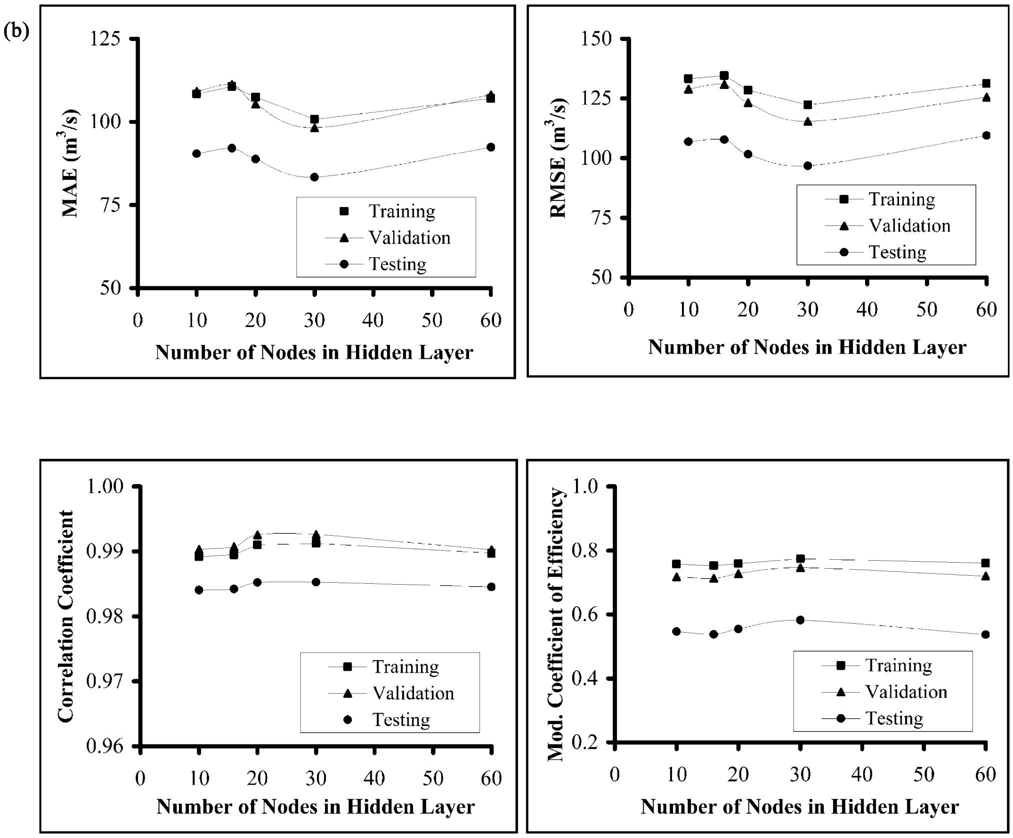


Fig. 8 (*continued*)

the latter part of the training irrespective of the number of input variables and the hidden layer nodes. For instance, in the case of 7 inputs, the rate of change in the normalized MSE per iteration is about $1.4 \times 10^{-3}$ and $1.3 \times 10^{-7}$ at the beginning of the training and at the latter part of the training, respectively. However, the learning curves indicate a slow reduction in the error until the termination. Therefore, in order to verify whether further improvement in the results could be achieved by increasing the number of iterations, two more simulations are conducted with networks 7-4-1 and 3-40-1 by setting the maximum number of iterations to 5000. The results do not indicate any improvement in the forecast accuracy with the increase in the number of iterations; rather there seems to be a decrease in the accuracy. The effective learning rate is found to be significantly smaller at the latter part of the training and both networks seem to converge to a minimum after about 1000 iterations. All these results suggest that 1000 iterations are reasonably sufficient to train the network to establish the input–output relationship of the river flow series studied and hence its forecasting for 1-day ahead.

#### *4.3.3. Results of 7-day lead time forecasts*

For the 7-day lead time forecasts, the input layer consists of 7 nodes representing the daily river flow values at times $t$, $t-1, \ldots, t-6$ and the output layer consists of a single node representing the river flow value at $t+7$. The number of nodes in the hidden layer is started at 2 and then gradually increased up to 10 to find the optimal number. The learning speeds of the networks do not seem to vary significantly with the increase in the number of hidden nodes (figure not shown). However, the use of only 2 hidden nodes does not seem to be sufficient to capture the underlying

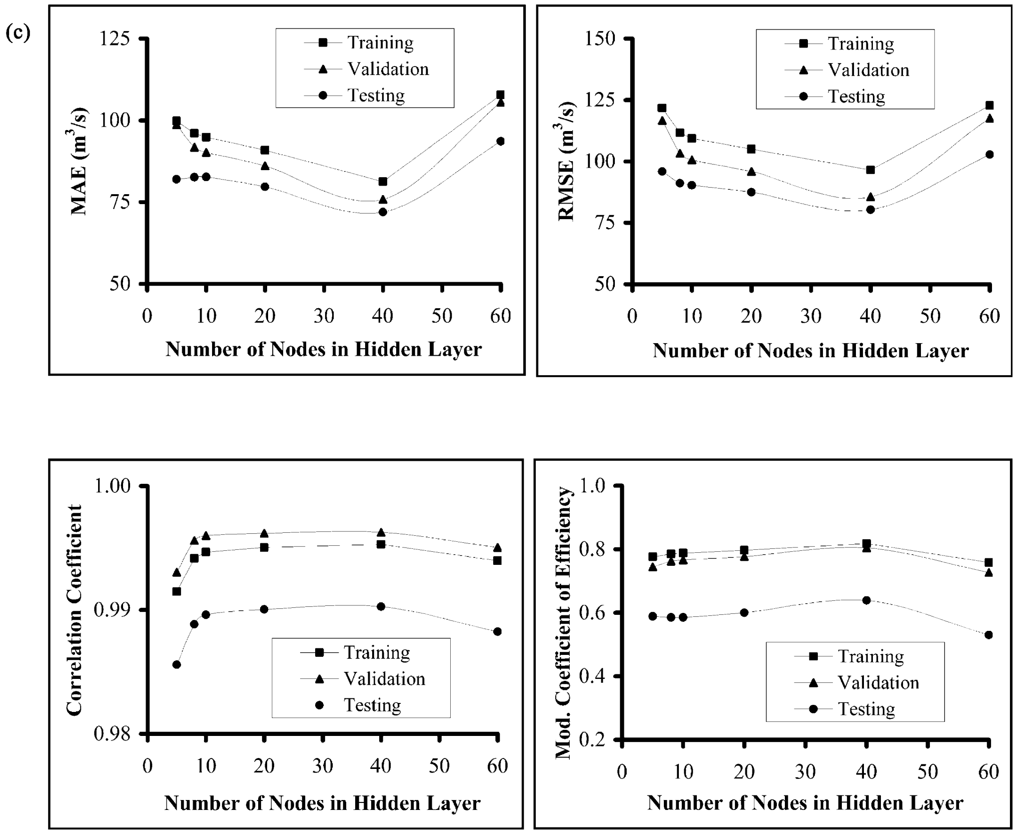


Fig. 8 (*continued*)

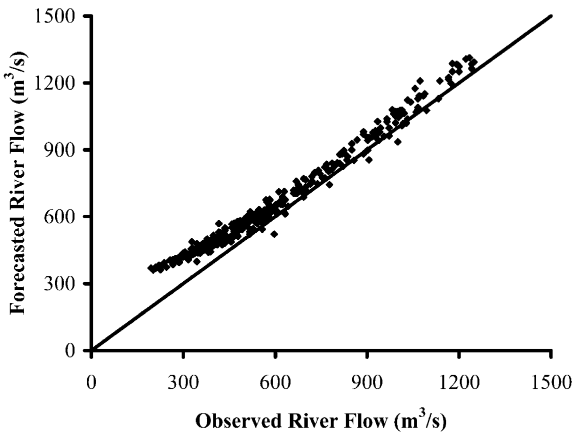


Fig. 9. Scatter plot of observed and ANN forecasted daily river flow series at Chao Phraya River basin for lead time = 1 day.

behavior of the data set. Unlike with the 1-day lead-time forecasts, the simulations in this case are terminated using the cross validation criterion except for the network 7-2-1. Fig. 10(a) presents a summary of the ANN forecasting results achieved for the river flow series when the forecasts are made 7 days ahead. As can be seen, the best results obtained with the network 7-6-1 are slightly better than those obtained with the other networks tested. The time series comparison of the 7-day ahead forecasted river flow values from such a network and the observed values is presented in Fig. 5(b), whereas Fig. 10(b) shows a scatter plot comparison of the same. As can be seen, the forecasts are in reasonably good agreement with the observed values; however, they are not as good as, and sometimes much worse than, the ones obtained using the PSR approach (see below for further details).

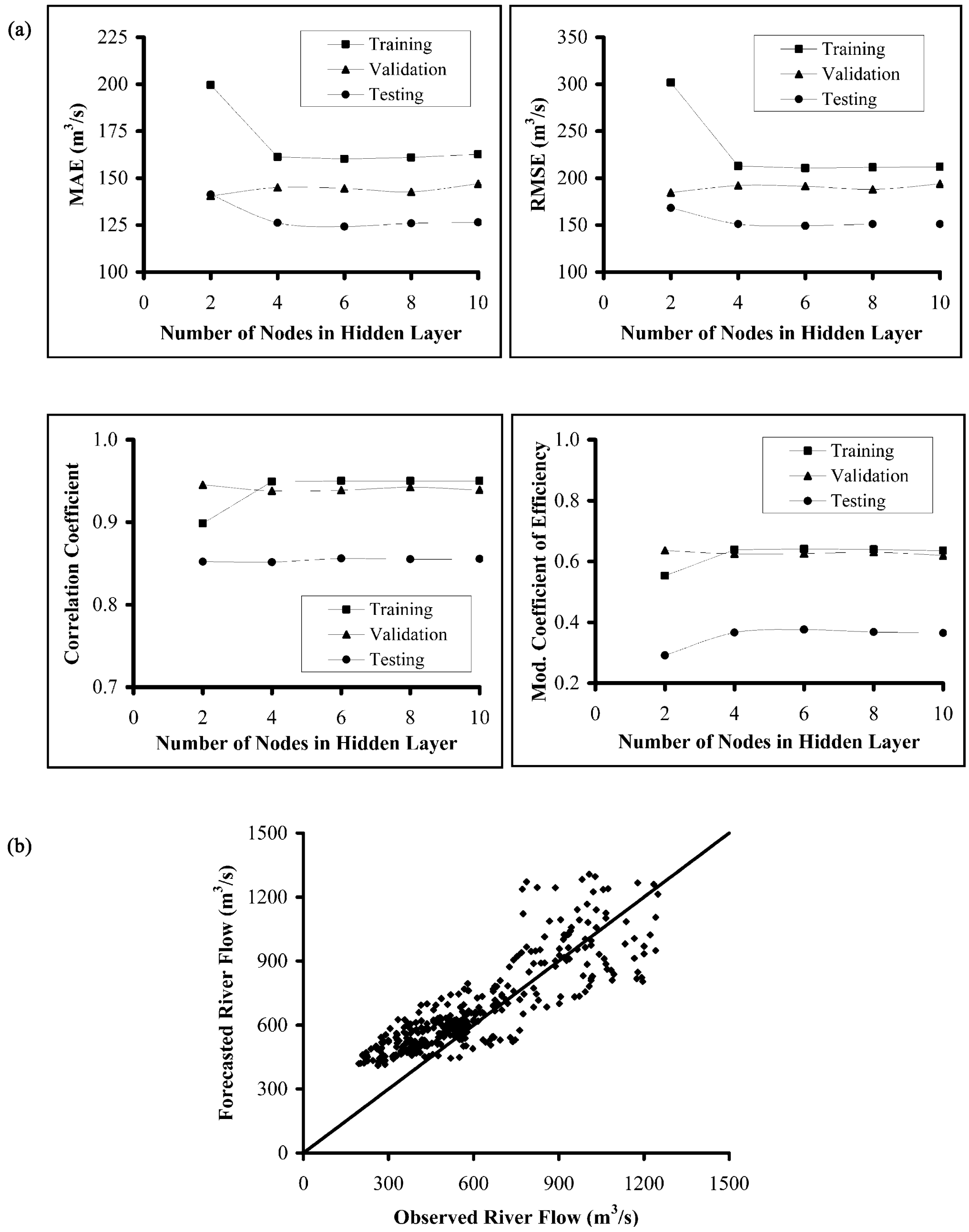


Fig. 10. (a) Variation of error indicators for ANN training, validation, and testing of daily river flow series at Chao Phraya River basin for different number of nodes in the hidden layer: lead time = 7 days, input variables = 7; and (b) scatter plot of observed and ANN forecasted daily river flow series at Chao Phraya River basin: lead time = 7 days.

Table 1
Comparison of forecasting results between PSR and ANN approaches for the daily river flow series from Chao Phraya River basin

| Error indicator | Phase-space | | Neural networks | |
|---|---|---|---|---|
| | Lead time = 1 | Lead time = 7 | Lead time = 1 | Lead time = 7 |
| Mean absolute error ($m^3/s$) | 8.379 | 28.095 | 71.754 | 122.958 |
| Mean square error ($m^3/s$) | 145.40 | 1408.68 | 6416.93 | 22298.80 |
| Root mean square error ($m^3/s$) | 12.058 | 37.532 | 80.106 | 149.328 |
| Maximum absolute error ($m^3/s$) | 67.136 | 113.120 | 174.607 | 485.563 |
| Minimum absolute error ($m^3/s$) | 0.033 | 0.277 | 0.405 | 1.294 |
| Correlation coefficient | 0.9992 | 0.9893 | 0.9904 | 0.8556 |
| Coefficient of determination | 0.9985 | 0.9788 | 0.9809 | 0.7321 |
| Coefficient of efficiency | 0.9977 | 0.9777 | 0.8983 | 0.6466 |
| Mod. coefficient of efficiency | 0.9583 | 0.8603 | 0.6431 | 0.3834 |

A sensitivity analysis is now carried out to evaluate the significance of each input parameter in making 7-day ahead forecasts using 7 input variables. Similar to the procedure adopted in the 1-day ahead forecast case, the trained weights of the optimal network (7-6-1) are used for this analysis. Each input variable is increased by 10 percent at a time and the forecasts obtained with the increased input are compared with the original forecasts. The results, i.e. sensitivity indicators (SL and APMSE), indicate that only the river flow value observed at the previous day (i.e. $t$) has a significant impact on the 7-day ahead ($t+7$) forecasts (with $\mathrm{SL}=92$, $\mathrm{APMSE}=75$), similar to the observation in the 1-day ahead forecast case. However, unlike in the 1-day ahead forecasts, the other input variables also seem to have a relatively significant impact in making 7-day ahead forecasts. The river flow value observed at time $t-2$ has the second highest significance on the 7-day lead time forecasts (with $\mathrm{SL}=35$, $\mathrm{APMSE}=11$), followed by the ones observed at $t-1$, $t-5$ and $t-6$.

In an attempt to obtain the optimal network structure for the 7-day ahead forecasts, simulations are also carried out with river flow values observed at times $t$, $t-1$, and $t$-2 as input parameters, and with $10, 12, 30, 40, 60, 100,$ and 150 hidden nodes. The results indicate that none of these networks yields satisfactory forecasts when compared to the best results obtained with 7 input nodes and, therefore, suggest that just 3-day lagged variables are not sufficient for making 7-day ahead forecasts of the daily river flow series under study.

*4.4. Comparison between PSR and ANN forecasting results*

*4.4.1. Forecast accuracy*

Table 1 presents, using a variety of absolute and relative error indicators, a comparison of the performance of the PSR approach and the ANN approach in forecasting the daily river flow dynamics in the Chao Phraya River basin at time scales of 1-day and 7 days ahead. As can be seen, in general, both the approaches yield reasonably good forecasts for both the 1-day and 7-day lead times, though the forecasts are better for the 1-day ahead case. However, the forecasts obtained using the PSR approach are significantly better than those using the ANN approach irrespective of the lead time. This can be better seen through, for instance, a direct time series comparison of these forecasted values with the observed values, as shown in Fig. 5(a) and (b) for the 1-day and 7-day lead times, respectively, and also using scatter diagrams, as presented in Figs. 6(a) and 9 for the 1-day lead time and Figs. 6(b) and 10(b) for the 7-day lead time. As can be seen, from Fig. 5(a) and (b), the PSR approach not only very well captures the major trends in the river flow series but also preserves very well the minor (noisy) fluctuations, as the forecasted values are extremely closer to and almost indistinguishable from the observed values, particularly for the 1-day lead time. As seen in Figs. 5(a),(b) and 6(a),(b), even extreme (i.e. very high) values are very well forecasted using the PSR approach. On the other hand, the ANN approach provides sometimes a significant overestimation of the river flow values, as obtained for both the 1-day and 7-day lead time cases (Figs. 5(a),(b), 9, and 10(b)), and at other times a significant underestimation of the flow values, as obtained for the 7-day lead time case (Figs. 5(b) and 10(b)).

It would be interesting to know what makes the PSR approach a better forecaster than the ANN approach (at least) for the river flow series studied herein. The importance of this question lies in the facts that: (1) both the PSR approach and the ANN approach are black-box (or data-driven) approaches, based on training and learning to establish possible connections between the inputs and the outputs; and (2) the ANN approach has been shown to yield very good forecasts for a large number of river flow series studied (e.g. Karunanithi et al., 1994; Fernando and Jayawardena, 1998). It is difficult at this stage to provide strong reasons for the above question, but the following could be some of the possible reasons for such results.

First, the PSR approach captures the most important features of the river flow dynamics, in a better way than the ANN approach does, essentially due to the local approximation method that represents the dynamics captured in the phase-space step by step in local neighborhoods. Such a local approximation method is certainly capable of better capturing the dynamics of the system when compared to a global approximation method, as used in the ANN approach, when the system under investigation exhibits low-dimensional chaotic dynamical behavior. The fact that the river flow series analyzed in the present study has been identified to exhibit low-dimensional chaotic behavior with a correlation dimension of about 2.90 (e.g. Jayawardena and Gurung, 2000) only supports the above interpretation.

Second, ANNs perform generally very well for one step ahead forecasts, but not for longer lead times. This is particularly the case with MLPs trained with back-propagation learning algorithm, as is the case in the present study, since such networks may not be the most appropriate networks for forecasts of two steps ahead or more. The results achieved in the present study only seem to support the above as: (1) the 1-day ahead forecasts achieved for the Chao Phraya flow series using the ANN approach are in good agreement with the observed values and are highly comparable with the forecasts achieved using the PSR approach; and (2) the 7-day ahead forecasts using the ANN approach are not in very good agreement with the observed values (yielding significant overestimation in general) and also are significantly different and less accurate when compared to those using the PSR approach.

In addition to the above, the selection of the training set in the ANN approach may also contribute to such results. In other words, for instance, very high values in the training set may drive the forecasted solution toward higher values, and very low values may drive the forecasted solution toward lower values. As far as the present study is concerned, the selection of extreme events (i.e. very high values) in the training set may have resulted in an overestimation of the forecasted values, which is also supported by the forecasts obtained for the 1-day ahead (Figs. 5(a) and 9) and 7-day ahead (Figs. 5(b) and 10(b)), particularly for the latter. The clear time shifts between the forecasted and the observed values for the 1-day and 7-day ahead, shown in Fig. 5(a) and (b), also seem to support the concern regarding the selection of the training set and the use of MLPs, in particular for the 7-day ahead forecasts.

In regards to the above problems with the use of MLPs and the selection of training set, it is necessary: (1) to study the possibility of improving the forecasts (particularly at longer lead times) using other types of networks; and (2) to study the influence of (the selection of) training set on the forecast results. Research in these directions are being carried out, details of which will be reported elsewhere.

#### *4.4.2. Optimal number of variables*

Another important observation that can be made from the PSR and the ANN forecast results is with reference to the optimal number of variables (i.e. the number of variables that yields the best forecasts), as identified by the two approaches, for capturing the dynamics of the river flow series. In regards to this, there is a complete agreement between the PSR and the ANN approaches as far as 1-day lead-time forecasts are concerned, as both the approaches identify the optimal number of variables as 3 (i.e. embedding dimension equal to 3 in the PSR approach, and number of inputs equal to 3 in the ANN approach). Appropriately, this result is also consistent with the optimal number of variables identified based on the correlation dimension of this river flow series, that is the nearest integer above the correlation dimension value (of about 2.90) obtained (e.g. Jayawardena and Gurung, 2000). However, a significant discrepancy is observed in the optimal number of variables, as identified by the PSR and the ANN approaches, when the forecasts are made for 7 days ahead. In this case, the optimal number of variables identified by the PSR approach is 1, whereas that identified by the ANN approach is 7. It is not known what causes such a discrepancy, but one possible reason could be the effect of the noise (e.g. measurement error) present in the river flow series on the forecasts in the case of the PSR approach.

In regards to the above, the effect of noise generally increases with the increase in the embedding dimension used for the PSR (e.g. Sugihara and May, 1990; Sivakumar, 2000). Therefore, the river flow forecast accuracy achieved for higher embedding dimensions might be significantly different and worse than the actual forecast accuracy that could be achieved if there were no noise. One possible way to verify this is to reduce (or remove) the noise in the river flow series and to make forecasts for the noise-reduced series. Investigations in this direction are underway, details of which are beyond the scope of the present study. It is relevant, however, to note at this point that the effect of noise on the river flow forecasts using the PSR approach is also recognized when 1-day ahead forecasts are made, as a decrease in the forecast accuracy is observed with an increase in the embedding dimension beyond 3 (Fig. 4(a)). On the other hand, the ANN approach is less susceptible to the presence of noise in the data series and, therefore, it is believed that the forecast accuracy obtained for the river flow is not affected much due to the presence of noise. All these observations seem to imply that: (1) the optimal number of variables for the 1-day ahead forecasts is 3, as identified by both the PSR and the ANN approaches; (2) the optimal number of variables for the 7-day ahead forecasts is expected to be greater than 3 and, therefore, could be closer to 7, as identified by the ANN approach (and is not 1, as identified by the PSR approach); and (3) the significantly worse performance of the ANN approach for the 7-day ahead forecasts (and to some extent the 1-day ahead forecasts as well) may perhaps be due to the inadequacy of the training procedure or the training data set adopted to learn the input–output relationship (rather than the insufficiency of the number of variables used as inputs) and/or due to the inability of the MLP (as discussed earlier).

## 5. Conclusions

An attempt was made in this study to investigate the use of two non-linear black-box approaches: (1) the PSR, a local approximation approach; and (2) the ANN, a global approximation approach, for forecasting river flow dynamics, and also to compare their performances. For this purpose, the daily river flow series observed at the Nakhon Sawan gaging station in the Chao Phraya River basin in Thailand was analyzed, and forecasts were made for 1-day and 7-day lead times. The number of variables in the PSR approach (i.e. embedding dimension in the PSR) was varied between 1 and 9, whereas in the ANN approach the number of variables (i.e. the number of input nodes) used was 7. In the latter case, however, a sensitivity analysis was also carried out (with 1 and 3 input nodes, respectively) to find the optimal number of variables.

The results indicated that, in general, both the PSR and the ANN approaches yielded reasonably good forecasts for both the 1-day and the 7-day lead times, even though the 1-day ahead forecasts were found to be much better. However, a comparison between the PSR forecasts and the ANN forecasts clearly indicated a much better performance of the former. The forecasted river flow values using the PSR approach were found to be extremely closer to the observed ones, with not only the major trends (including extreme conditions) very well captured but also the minor noisy fluctuations well preserved. In the case of the ANN approach, the forecasts were sometimes found to be a significant overestimation and/or underestimation of the observed river flow values. It is believed that the significantly better performance of the PSR approach is due to the representation of the river flow dynamics in the phase-space step by step in local neighborhoods, as the river flow series exhibits low-dimensional chaotic behavior (e.g. Jayawardena and Gurung, 2000). On the other hand, the multi-layer perceptrons (MLPs) with a back-propagation learning algorithm, used in the present study, may have limitations, particularly when forecasts at longer lead times are attempted. Another possible reason could be the selection of the training set in the ANN approach.

The results from the present study raise two important questions. The first question is concerned with the (much worse) performance of the ANN approach when compared to the PSR approach, since the ANN approach has been reported to yield near-accurate forecasts for a number of river flow series. The second question is concerned with the (discrepancy in the) optimal number of variables identified by the PSR approach and the ANN approach for capturing the important dynamical features of the river flow dynamics while making 7-day ahead forecasts. The noise (e.g. measurement error) present in the river flow series seems to play an important role in this regard, which is also partially supported by the 1-day lead time forecasts using the PSR approach with respect to the (lower and higher) embedding dimensions used in the PSR. In view of the above, the immediate tasks are to: (1) use other, and possibly better, types of ANNs than the MLP used herein for river flow predictions, particularly at longer lead times; (2) study the effect of the selection of training set in the ANN approach; and (3) study the influence of noise on forecast accuracy, particularly in the PSR approach. Studies along these directions are underway and the details will be reported in subsequent articles.

## Acknowledgments

The work presented in this article, partially supported by the Hong Kong Research Grants Council Grant No. 7003/97E, was carried out when the first author was visiting the Department of Civil Engineering, The University of Hong Kong. The authors would like to thank the two reviewers for their valuable suggestions, which resulted in a much more accurate and complete presentation of the work.

## Appendix A. Definitions of error indicators

Definitions:

Mean absolute error

$$\mathrm{MAE} = \frac{1}{N}\sum_{i=1}^{N} |(O_i - P_i)| \qquad \text{(A1)}$$

Root mean square error

$$\mathrm{RMSE} = \sqrt{\frac{1}{N}\sum_{i=1}^{N} (O_i - P_i)^2} \qquad \text{(A2)}$$

Correlation coefficient

$$r = \left\{ \frac{\sum_{i=1}^{N} (O_i - \bar{O})(P_i - \bar{P})}{\left[\sum_{i=1}^{N} (O_i - \bar{O})^2\right]^{1/2} \left[\sum (P_i - \bar{P})^2\right]^{1/2}} \right\} \qquad \text{(A3)}$$

Modified coefficient of efficiency

$$E_1 = 1 - \frac{\sum_{i=1}^{N} |O_i - P_i|}{\sum_{i=1}^{N} |O_i - \overline{O}|} \qquad \text{(A4)}$$

where $N$ is the number of values in the evaluation set, $O$ and $P$ are the observed and the forecasted values in the evaluation set, respectively, and $\overline{O}$ is the mean of the observed values in the evaluation set.

Notes:

In general, RMSE $\geq$ MAE, and the degree to which RMSE exceeds MAE is an indicator of the extent to which outliers (variance in the differences between the observed and the forecasted values) exist in the evaluation set.

The correlation coefficient ($r$) ranges from 0 to 1 with higher values indicating better agreement between the observed and the forecasted values, whereas the modified coefficient of efficiency ($E_1$) ranges from $-\infty$ to 1, with the higher positive values indicating better agreement.

## References

Abarbanel, H.D.I., 1996. Analysis of Observed Chaotic Data, Springer, New York, USA.

Amari, S., Murata, N., Müller, K.B., Finke, M., Yang, H.H., 1997. Asymptotic statistical theory of overtraining and cross-validation. IEEE Trans. Neural Networks 8 (5), 985–996.

Farmer, D.J., Sidorowich, J.J., 1987. Predicting chaotic time series. Phys. Rev. Lett. 59, 845–848.

Fausett, L., 1994. Fundamentals of Neural Networks, Prentice Hall, Englewood, NJ.

Fernando, D.A.K., Jayawardena, A.W., 1998. Runoff forecasting using RBF networks with OLS algorithm. J. Hydrol. Engng, ASCE 3 (3), 203–209.

Govindaraju, R.S., 2000. Artificial neural networks in hydrology. II: Hydrological applications, by the ASCE task Committee on application of artificial neural networks in hydrology. J. Hydrol. Engng 5, 124–137.

Haykin, S., 1994. Neural Networks: A Comprehensive Foundation, MacMillan, New York, USA.

Hecht-Nielsen, R., 1990. Neurocomputing, Addison-Wesley, Reading, Massachusetts, USA.

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

Jayawardena, A.W., Fernando, D.A.K., 1998. Use of radial basis function type artificial neural networks for runoff simulation. Comput.-Aided Civ. Infrastruct. Engng 13, 91–99.

Jayawardena, A.W., Fernando, T.M.K.G., 2001. River flow prediction: an artificial neural network approach. In: Schumann, A.H., Acreman, M.C., Davis, R., Marino, M.A., Rosbjerg, D., Jun, X. (Eds.), Regional Management of Water Resources, Proceedings of the Sixth IAHS Scientific Assembly, 18–27 July, Maastricht, Netherlands, IAHS Publication No. 268, pp. 239–245.

Jayawardena, A.W., Gurung, A.B., 2000. Noise reduction and prediction of hydrometeorological time series: dynamical systems approach vs. stochastic approach. J. Hydrol. 228, 242–264.

Jayawardena, A.W., Lai, F., 1994. Analysis and prediction of chaos in rainfall and stream flow time series. J. Hydrol. 153, 23–52.

Jayawardena, A.W., Fernando, T.M.K.G., Chan, C.W., Chan, W.C., 2000. Comparison of ANN, dynamical systems and support vector approaches for river discharge prediction. The Nineteenth Chinese Control Conference, Proceedings of CCC, 2000, December, Hong Kong, vol. 2., pp. 504–508.

Karunanithi, N., Grenney, W.J., Whitley, D., Bovee, K., 1994. Neural networks for river flow prediction. J. Comp. Civ. Engng, ASCE 8 (2), 201–220.

Kasabov, N.K., 1996. Foundations of Neural Networks, Fuzzy Systems, and Knowledge Engineering, MIT Press, Cambridge, Massachusetts, USA.

Minns, A.W., Hall, M.J., 1996. Artificial neural networks as rainfall-runoff models. Hydrol. Sci. J. 41 (3), 399–417.

Oki, T., Musiake, K., Matsuyama, H., Masuda, K., 1995. Global atmospheric water-balance and runoff from large river basins. Hydrol. Processes 9 (5–6), 655–678.

Packard, N.H., Crutchfield, J.P., Farmer, J.D., Shaw, R.S., 1980. Geometry from a time series. Phys. Rev. Lett. 45 (9), 712–716.

Porporato, A., Ridolfi, L., 1997. Nonlinear analysis of river flow time sequences. Water Resour. Res. 33 (6), 1353–1367.

Raghunath, J., Herath, S., Musiake, K., 2000. River network solution for a distributed hydrological model and applications. Hydrol. Processes 14 (3), 575–592.

Randle Inc., 1996. cspW, Tools for Dynamics, Applied Nonlinear Sciences, LLC, Randle Inc, Great Falls, Virginia, USA.

See, L., Openshaw, S., 1999. Applying soft computing approaches to river level forecasting. Hydrol. Sci. J. 44 (5), 763–778.

Sivakumar, B., 2000. Chaos theory in hydrology: important issues and interpretations. J. Hydrol. 227 (1–4), 1–20.

Sivakumar, B., Berndtsson, R., Olsson, J., Jinno, K., Kawamura, A., 2000. Dynamics of monthly rainfall-runoff process at the Göta basin: a search for chaos. Hydrol. Earth Syst. Sci. 4 (3), 407–417.

Sivakumar, B., Berndtsson, R., Persson, M., 2001. Monthly runoff prediction using phase-space reconstruction. Hydrol. Sci. J. 46 (3), 377–388.

Sugihara, G., May, R.M., 1990. Nonlinear forecasting as a way of distinguishing chaos from measurement error in time series. Nature 344, 734–741.

Takens, F., 1981. Detecting strange attractors in turbulence. In: Rand, D.A., Young, L.S. (Eds.), Dynamical Systems and Turbulence, Lecture Notes in Mathematics 898, Springer, Berlin, pp. 366–381.

Wijesekera, N.T.S., Musiake, K., Herath, A.S., 1995. Water balance in Chao Phraya sub-catchments through watershed modelling. Proceedings of The Second International Study Conference on GEWEX in Asia and GAME, Pattaya, Thailand, pp. 268–271.

Zealand, C.M., Burn, D.H., Simonovic, S.P., 1999. Short term streamflow forecasting using artificial neural networks. J. Hydrol. 214, 32–48.