

# A comparative study of artificial neural network (MLP, RBF) and support vector machine models for river flow prediction

Mohammad Ali Ghorbani<sup>1</sup> · Hojat Ahmad Zadeh<sup>1</sup> · Mohammad Isazadeh<sup>1</sup> · Ozlem Terzi<sup>2</sup>

Received: 18 April 2015 / Accepted: 11 December 2015 / Published online: 10 March 2016  
© Springer-Verlag Berlin Heidelberg 2016

**Abstract** This study investigates the applicability of multilayer perceptron (MLP), radial basis function (RBF) and support vector machine (SVM) models for prediction of river flow time series. Monthly river flow time series for period of 1989–2011 of Safakhaneh, Santeh and Polanian hydrometric stations from Zarrinehrud River located in north-western Iran were used. To obtain the best input–output mapping, different input combinations of antecedent monthly river flow and a time index were evaluated. The models results were compared using root mean square errors and the correlation coefficient. A comparison of models indicates that MLP and RBF models predicted better than SVM model for monthly river flow time series. Also the results showed that including a time index within the inputs of the models increases their performance significantly. In addition, the reliability of the models prediction was calculated by an uncertainty estimation. The results indicate that the uncertainty in the SVM model was less than those in the RBF and MLP models for predicting monthly river flow.

**Keywords** Multilayer perceptron · Radial basis function · Prediction · Support vector machine · Uncertainty · Zarrinehrud River

## Introduction

River flow predicting is important for planning and management of river basins, assessment of risk and control of floods and droughts, development of water resources, production of hydroelectric energy, navigation planning and allocation of water for agriculture (Bayazit 1988; Khatibi et al. 2012). The data-driven modeling is based on extracting and reusing information that is included in the hydrological data without directly taking into explanation the physical rules that cause the river flow processes (Samsudin et al. 2011). In recent years, artificial neural networks and support vector machines have been widely used in the modeling of various complex environmental problems (Yilmaz and Kaynar 2011).

A neural network is an adaptable system that learns relationships from the input and output data sets and then is able to predict a previously unseen data set of similar characteristics to the input set (Haykin 1999; ASCE Task Committee 2000). Multilayer perceptron (MLP) and radial basis function (RBF) are widely used neural network architecture in literature for regression problems (Cohen and Intrator 2002; Kenneth et al. 2001; Loh and Tim 2000). From the literature, several researchers have used the MLP and RBF techniques in river flow forecasting.

Jayawardena and Fernando (1995) used both MLP and RBF methods for flood forecasting and compared their performance with the statistical ARMAX model. The results showed the two data driven models are better than ARMAX model. Sudheer and Jain (2003) used MLP and RBF methods in daily river flow estimation. The results showed that the MLP results were similar to those of the RBF. Kisi (2005) used auto-regressive and artificial neural network models to estimate river flow for two rivers in the USA, namely the Blackwater River and the Gila River

---

✉ Mohammad Ali Ghorbani  
Ghorbani@tabrizu.ac.ir; m\_ali\_ghorbani@ymail.com

<sup>1</sup> Department of Water Engineering, University of Tabriz, Tabriz, Iran

<sup>2</sup> Suleyman Demirel University, 32260 Isparta, Turkey

the Filyos Stream in Turkey. He indicated that the artificial neural network model has better results than the auto-regressive model. Sahoo and Ray (2006) compared the accuracy of the RBF and MLP models in prediction of daily river flows of a Hawaiian river. The results showed that the MLP and RBF results were similar. Kisi (2008) investigated the capability of three different architectures of ANN techniques for the estimation of monthly river flow of two rivers in Turkey. He found that all methods can be used successively for river flow prediction. Terzi (2011) investigated multilinear regression, multilayer perceptron, radial basis function network, decision table, REP tree and KStar algorithms for monthly river flow for Kızılırmak River in Turkey. The results indicated that the multilinear regression model has better results than other models. Terzi and Ergin (2014) used autoregressive (AR), gene expression programming, radial basis function network, feed-forward neural networks, and adaptive neural-based fuzzy inference system (ANFIS) techniques to predict monthly river flow for Kızılırmak River in Turkey. The results indicated that AR(2) model gave the best performance among all data driven models. Awchi (2014) compared the accuracy of feed-forward neural networks (FFNN), generalized regression neural networks (GRNN), the radial basis function neural networks (RBF) and Multiple regression models for forecasting monthly river flow of Upper and Lower Zab rivers in Northern Iraq. The results showed that the ANNs performed better than the MLR and also the FFNN was almost better than other networks.

Support vector machine (SVM) proposed by Vapnik (1995) is a powerful tool for modeling of hydrologic processes, which are non-linear in nature (Bhagwat and Maity 2012). Many researchers have verified the SVM capability for the modelling of river flow prediction.

Wang et al. (2008) developed monthly river flow prediction model based on the genetic programming, ANFIS, SVM and ARMA. The results showed the three data driven models are better than ARMA model. Kisi and Cimen (2011) also applied discrete wavelet transform and SVM conjunction models for monthly river flow prediction. They concluded that the wavelet support vector machine model results were better than single SVM model. Kisi et al. (2012) compared the accuracy of ANFIS, ANNs and SVM for forecasting daily stream flows of two stations in north-western Turkey. The results showed that the ANN and ANFIS gave the best forecasts for the first and second stations respectively. Kalteh (2013) used support vector regression (SVR) and ANN methods coupled with wavelet transform in monthly river flow prediction. The results showed that both ANN and SVR models coupled with wavelet transform are better than the single ANN and SVR models.

All data-driven techniques such as MLP, RBF and SVM models are sensitive to input parameters and this lead to uncertainty in predicted values. So the reliability or uncertainty information about river flow prediction is very important for any application. Because information on predict uncertainty helps users to feel more confident about their decisions (Frick and Hegg 2011). Evaluation of predict uncertainty has been an active research subject in water related problems (Montanari and Grossi 2008; Coccia and Todini 2011; Noori et al. 2015).

The main objective of this study is to investigate the applicability of two different well-known types of ANNs including multilayer perceptron (MLP), radial basis neural networks (RBF) and comparison with support vector machine model to predict the monthly river flow in the Zarrinehrud River for three hydrometric stations in north-western Iran. The uncertainty involved in artificial intelligence (AI) techniques for river flow prediction has rarely been reported. The novelty of the study is that we have used the reliability of three AI-based techniques, including the MLP, RBF and SVM for predicting the monthly river flow in natural rivers.

## Methodology

### Artificial neural networks (ANNs)

ANNs are parallel information processing systems consisting of a set of neurons arranged in layers. These neurons provide suitable conversion functions for weighted inputs.

ANNs can be used to solve problems that are hard for conventional computers or human beings (Demuth and Beale 2001). In this study, MLP and RBF methods are used for the prediction of monthly river flow.

The MLP trained with the use of back propagation learning algorithm. Figure 1a represents a three-layer structure (MLP) that consists of (1) input layer, (2) hidden layer and (3) output layer. The input layer accepts the data and the hidden layer processes them and finally the output layer displays the resultant outputs of the model. For more information see Ghorbani et al. (2013).

The RBF is the most widely used architecture in ANN and simpler than MLP neural network. The RBF has an input, hidden and output layer. The hidden layer consists of RBF activation function or  $h(x)$  as networks neuron. There are different types of radial basis functions, but the most widely used type is the Gaussian function.

Figure 1b shows the structure of an RBF neural network. The basis functions in the hidden layer produce a localized response to the input. That is, each hidden neuron has a localized receptive field. For more information see (Yu et al. 2011; Terzi 2011).

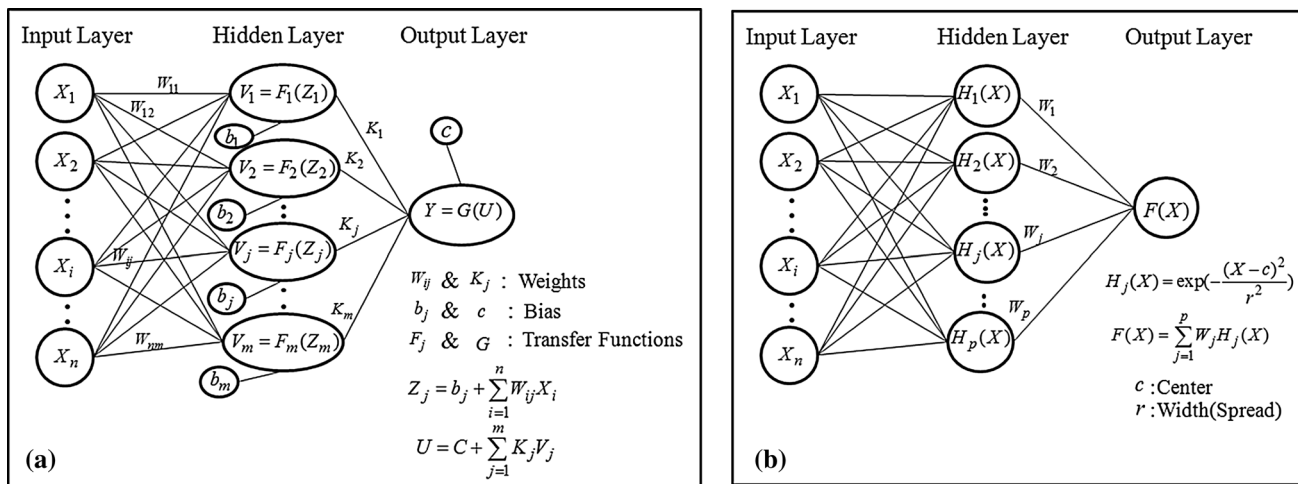


Fig. 1 Simple configuration of a MLP and; b RBF neural networks (Haykin 1999)

Support vector machines (SVM)

The Support Vector Machines (SVM) is a computer algorithm that learns by example to find the best function of classifier/hyperplane to separate the two classes in the input space. The SVM analyzed two kinds of data, i.e. linearly and non-linearly separable data. The example of linearly separated data is shown in Fig. 2. The best hyperplane between two classes can be found by measuring the hyperplane margin and finding out the maximum points. The margin is defined as the distance between hyperplane and the closest pattern of each class, which is called support vector (Vapnik 1998).

For a given training data with  $N$  number of samples, represented by  $(x_1, y_1), \dots, (x_N, y_N)$ , where  $x$  is an input vector and  $y$  is a corresponding output value, SVM estimator ( $f$ ) on regression can be represented by:

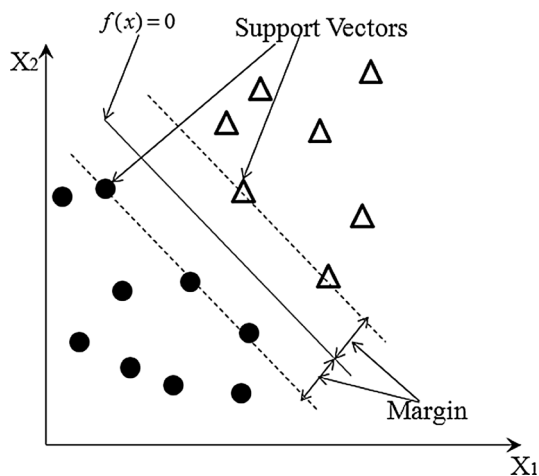


Fig. 2 A schematic figure of linearly separated data

$$f(x) = w \cdot \phi(x) + b \tag{1}$$

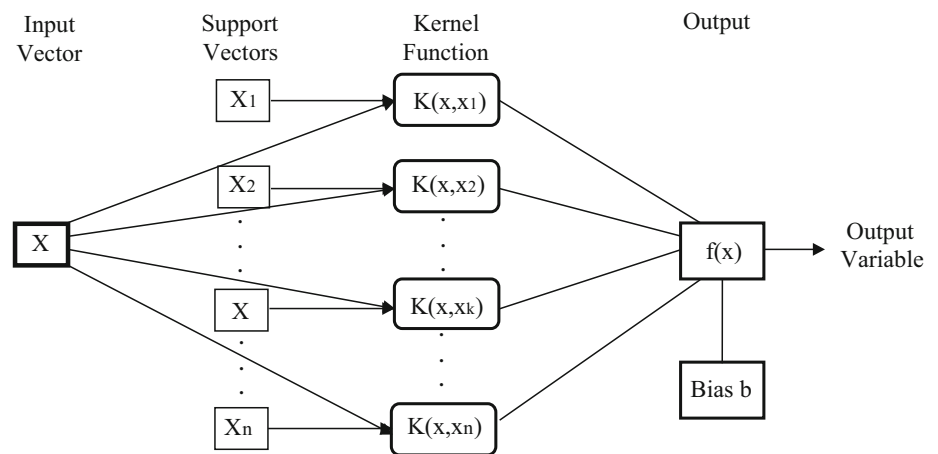
where  $w$  is a weight vector,  $b$  is a bias, “.” denotes the dot product and  $\phi$  is a non-linear mapping function. A smaller value of  $w$  indicates the flatness of equation (x), which can be obtained using minimizing the Euclidean norm as defined by  $w^2$ . Vapnik (1995) introduced the following convex optimization problem with an  $\epsilon$ -insensitivity loss function to obtain the solution to Eq. (2):

$$\begin{aligned} &\text{minimize } \frac{1}{2}w^2 + C \sum_{k=1}^N (\xi_k^+ + \xi_k^-) \\ &\text{subject to } \begin{cases} y_k - (w \cdot \phi(x_k) + b) \leq \epsilon + \xi_k^+ \\ -y_k + (w \cdot \phi(x_k) + b) \leq \epsilon + \xi_k^- & k = 1, 2, \dots, N \\ \xi_k^+, \xi_k^- \geq 0 \end{cases} \end{aligned} \tag{2}$$

where  $C$  is a positive trade off parameter that determines the degree of the empirical error in the optimization problem and determines the trade-off between the flatness of the function and the amount to which deviations larger than  $\epsilon$  are tolerated. Also  $\xi_k^-, \xi_k^+$  are slack variables representing upper and lower constraints on the output system over the error tolerance  $\epsilon$  (Misra et al. 2009). Lagrangian multipliers and imposing the Karush–Kuhn–Tucker (KKT) method used to solve the optimization of Eq. (2) in a dual form. The inequality constraint converts into an equation by the KKT method by adding or subtracting slack variables. Support vectors are the input vectors that have nonzero Lagrangian multipliers under the KKT condition (Yoon et al. 2011). Figure 3 shows a schematic diagram of the SVM used in this study.

In natural processes almost the predictor variables (input space) are non-linearly related to the predicted variable. This limits a linear formulation of the problem as shown in

**Fig. 3** A schematic structure of SVM model (Yoon et al. 2011)



Eq. (2). This limitation is solved by mapping the input space on to some higher dimensional space (feature space) using a kernel function. The kernel function enables us to implicitly work in a higher dimensional feature space. Then Eq. (1) becomes the explicit function of Lagrangian multipliers or  $\alpha_i$  and  $\alpha_i^*$  as follows:

$$f(x, \alpha_i, \alpha_i^*) = \sum_{i=1}^N (\alpha_i - \alpha_i^*) \cdot K(x, x_i) + b \tag{3}$$

where  $K(x, x_i)$  is kernel function. The used kernel functions including the linear, polynomial, radial basis and sigmoid kernel functions. Because linear kernel is only appropriate for linear problems, polynomial kernel has computational difficulties and Sigmoid kernel function is not widely used (Chang and Lin 2005), in this study the most common kernel or RBF function is used as follows:

$$\text{RBF: } K(x, x_i) = \exp\left(-\frac{1}{2\gamma^2} \|x - x_i\|^2\right) \tag{4}$$

In (4)  $\gamma$  is the kernel parameter. The program of SVM was constructed by using MATLAB (The Math Works Inc 2012).

**Uncertainty analysis**

In this study we used a method proposed by Abbaspour et al.(2007) to uncertainty analyze the predicted river flow. In this method, the percentage of measured data bracketed by 95 percent of predicted uncertainties (95PPU) calculated by 2.5th ( $X_L$ ) and 97.5th ( $X_U$ ) percentiles of normal distribution function obtained from  $n$  (for example, 1000) times of the simulation results as follows:

$$\text{Bracketed by 95PPU} = \frac{1}{k} \text{Count} \left( j | X_L^l \leq X_{\text{reg}}^l \leq X_U^l \right) \times 100 \tag{5}$$

where 95PPU represents 95 % prediction uncertainty;  $k$  is the number of observed data points;  $l$  indicates the current month, taking values from 1 to  $k$  ( $l = \overline{1, k}$ );  $X_L^l$  and  $X_U^l$  represent the lower and the upper limit of the uncertainty defined as 2.5 % and 97.5 % levels of the cumulative distribution of output variables;  $X_{\text{reg}}^l$  is the registered values in the current month  $l$ ;  $j$  is increasing with one unit each time when the registered variable for the current month  $l$  is located between the uncertainty limits; the maximum value of  $j$  is obtained when  $l = k$ . If all registered values are within these limits, the maximum value of “Bracketed by 95PPU” is 100.

Also  $d$ -factor is used to evaluate the average width of confidence interval band as Eq. (6) :

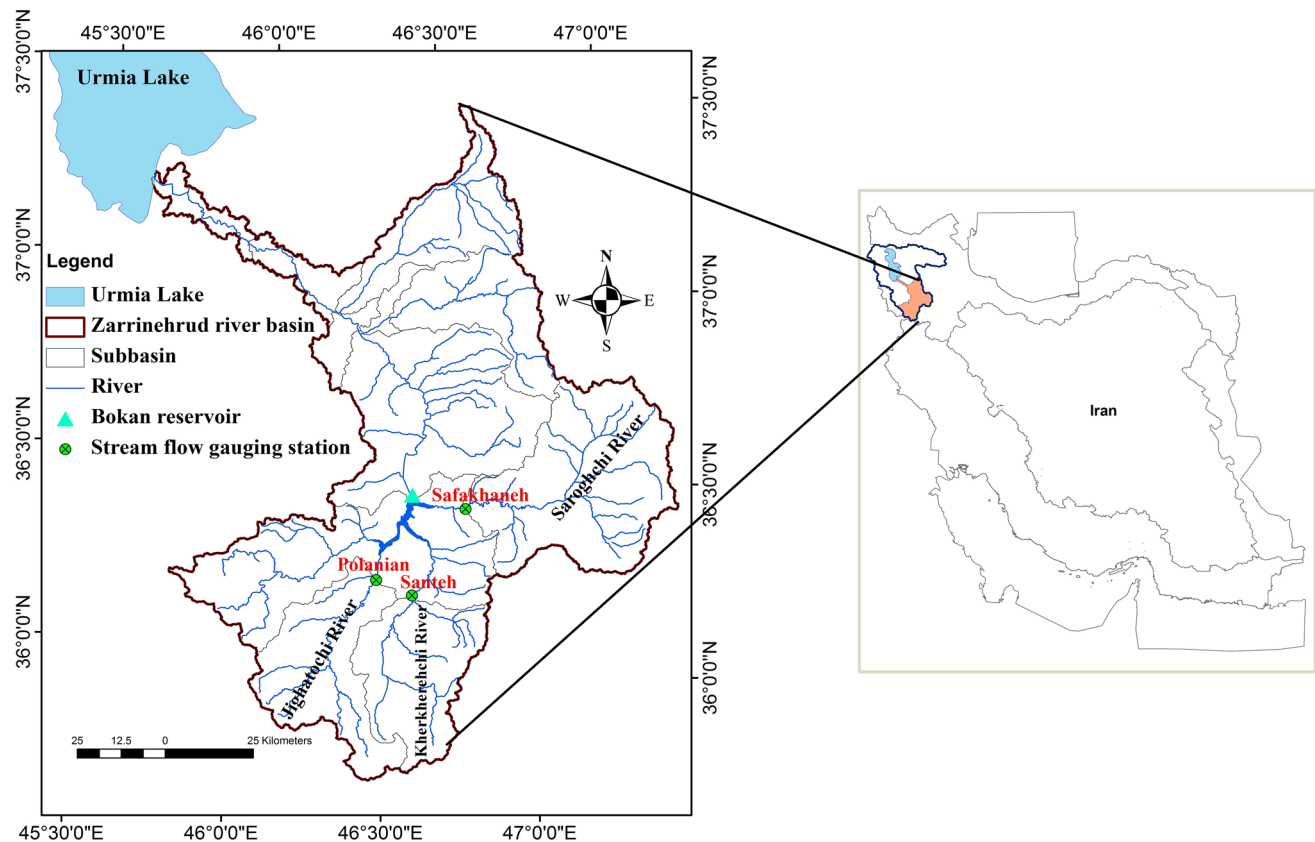
$$d\text{-factor} = \frac{\overline{d_X}}{\sigma_X} \tag{6}$$

$$\overline{d_X} = \frac{1}{k} \sum_{l=1}^k (X_U - X_L)_l \tag{7}$$

where  $\overline{d_X}$  is the average distance between the upper and lower bands and  $\sigma_X$  is the standard deviation of the observed data. Ideally, we would like to bracket most of the measured data (plus their uncertainties) within the 95PPU band, while having the narrowest band amplitude ( $d$ -factor  $\rightarrow 0$ ) (Azimi et al. 2013).

**Study area, the used data and performance criteria**

The monthly river flow time series between 1989 and 2011 years for Safakhaneh, Santeh and Polanian stations of Zarrinehrud River were used in this study. Zarrinehrud River is one of the largest rivers in north-western Iran and drains to Urmia Lake. The drainage area and length of the river are 12,025 km<sup>2</sup> and 300 km. The location of the Zarrinehrud River and three sites is shown in Fig. 4. The



**Fig. 4** The location of study sites in the Zarrinehrud River basin

measured data from January 1, 1989 to October 31, 2004 (190 records or 70 % of the whole data set) are used for training and the data from November 1, 2004 to August 30, 2011 (82 records or 30 % of the whole data) are chosen for testing (Fig. 5).

The statistical parameters of the Zarrinehrud River flow data at three sites are given in Table 1. In Table 1, the  $x_{\text{mean}}$ ,  $S_x$ ,  $C_v$ ,  $C_{sx}$ ,  $x_{\text{max}}$ , and  $x_{\text{min}}$  denote the mean, standard deviation, variation coefficient, skewness, maximum and minimum, respectively.

Two performance criteria are used in this study to assess the goodness of fit of the models, which are correlation coefficient (CC) and root mean square error (RMSE) (further discussed by Ghorbani et al. 2013).

## Results and discussion

### Input selection

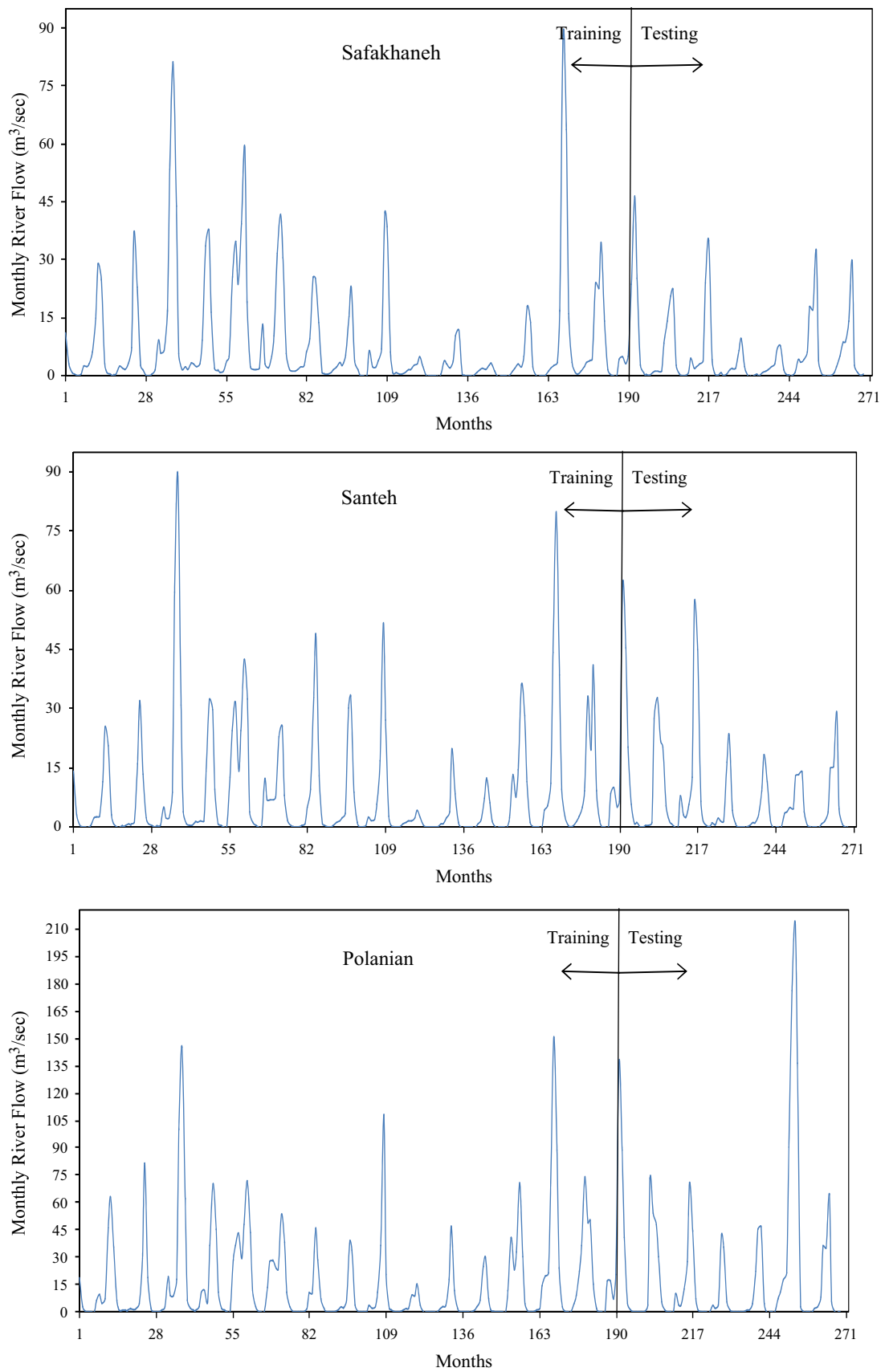
Selection of the input variables is one of the most important problems when developing prediction models. Hence, cross correlations analysis was performed between input and output variables in order to determine the suitable time lag

and the number of input variables (Table 2) (Luk et al. 2000).

In this study, several input combinations of river flow values were used. Here  $Q_t$ ,  $Q_{t-1}$ ,  $Q_{t-2}$ ,  $Q_{t-3}$  represents the river flow values at time  $(t)$ ,  $(t - 1)$ ,  $(t - 2)$  and  $(t - 3)$ . The  $\tau$  is the month number between 1 and 12.

### ANNs models

In this study, all data were first normalized and then divided into training and testing data sets. A typical MLP model with a back-propagation algorithm is constructed for predicting monthly river flow. The back-propagation training algorithm is a supervised training mechanism and is normally adopted in most of the engineering application. The neurons in the input layer have no transfer function. The logarithmic sigmoid transfer function was used in the hidden layer and linear transfer function was employed from the hidden layer to the output layer as an activation function, because the linear function is known to be robust for a continuous output variable. The network was trained in 1000 epochs using the Levenberg–Marquardt learning algorithm with a learning rate of 0.001 and a momentum coefficient of 0.9. The optimal number of neuron in the



**Fig. 5** Time series plot for the river flow data

**Table 1** The monthly statistical parameters of river flow data

| Site       | Data set | Number of data | $x_{\text{mean}}$ | $S_x$ | $C_v$ | $C_{\text{sx}}$ | $x_{\text{max}}$ | $x_{\text{min}}$ |
|------------|----------|----------------|-------------------|-------|-------|-----------------|------------------|------------------|
| Safakhaneh | Training | 190            | 8.7               | 14.4  | 1.7   | 2.8             | 88.1             | 0.0              |
|            | Testing  | 82             | 6.0               | 9.3   | 1.6   | 2.3             | 46.6             | 0.0              |
|            | Total    | 272            | 7.8               | 13.1  | 1.7   | 2.9             | 88.1             | 0.0              |
| Santeh     | Training | 190            | 8.6               | 14.3  | 1.7   | 2.7             | 89.4             | 0.0              |
|            | Testing  | 82             | 7.9               | 12.7  | 1.6   | 2.5             | 61.1             | 0.0              |
|            | Total    | 272            | 8.4               | 13.8  | 1.6   | 2.6             | 89.4             | 0.0              |
| Polanian   | Training | 190            | 16.1              | 25.3  | 1.6   | 2.6             | 150.7            | 0.0              |
|            | Testing  | 82             | 21.6              | 40.0  | 1.9   | 2.9             | 213.2            | 0.0              |
|            | Total    | 272            | 17.7              | 30.5  | 1.7   | 3.0             | 213.2            | 0.0              |

**Table 2** Cross correlations analysis between input and output variables

| Station           | $Q(t - 1)$ | $Q(t - 2)$ | $Q(t - 3)$ |
|-------------------|------------|------------|------------|
| $Q(t)$ Safakhaneh | 0.638      | 0.170      | -0.053     |
| Santeh            | 0.614      | 0.131      | -0.101     |
| Polanian          | 0.663      | 0.218      | -0.041     |

hidden layer was identified using a trial and error procedure by varying the number of hidden neurons from 1 to 20, and the structures of best networks are shown in Table 3. As shown in Table 3, the performance criteria show that ANN(3,5,1), ANN(4,2,1) and ANN(2,8,1) models perform relatively better than the others combinations for the

Safakhaneh, Santeh and Polanian stations, respectively. Here, the structure of the ANN(3,5,1) model denotes an ANN model comprising three input, five hidden and one output nodes. The results also showed that the models including a time index increases their performance significantly. The decrease between RMSE values of the models without a time index and the models including a time index for testing set ranged from 15.6 to 17.4 %, 7 to 18.7 %, 7.3 to 14.6 % for Safakhaneh, Santeh and Polanian, respectively.

Figure 6 shows the comparative plots between the results obtained from the MLP model and the the observed monthly river flow for testing data set. From Fig. 6, it can be observed that the values of the monthly river flow are generally predicted close to the observed value. Table 3

**Table 3** Performance of MLP model for monthly river flow prediction

| Station    | Input combination                                  | Output | Structure      | Training    |                          | Testing     |                          |
|------------|--|--------|----------------|-------------|--------------------------|-------------|--------------------------|
|            |  |        |                | CC          | RMSE (m <sup>3</sup> /s) | CC          | RMSE (m <sup>3</sup> /s) |
| Safakhaneh | (1) $Q_{t-1}$                                      | $Q_t$  | (1,6,1)        | 0.68        | 10.54                    | 0.67        | 7.11                     |
|            | (2) $Q_{t-1}, \tau$                                | $Q_t$  | (2,4,1)        | 0.85        | 7.48                     | 0.83        | 6.00                     |
|            | (3) $Q_{t-1}, Q_{t-2}$                             | $Q_t$  | (2,11,1)       | 0.81        | 8.33                     | 0.80        | 6.55                     |
|            | <sup>a</sup> (4) $Q_{t-1}, Q_{t-2}, \tau$          | $Q_t$  | <b>(3,5,1)</b> | <b>0.87</b> | <b>7.30</b>              | <b>0.83</b> | <b>5.41</b>              |
|            | (5) $Q_{t-1}, Q_{t-2}, Q_{t-3}$                    | $Q_t$  | (3,5,1)        | 0.78        | 8.91                     | 0.74        | 6.70                     |
|            | (6) $Q_{t-1}, Q_{t-2}, Q_{t-3}, \tau$              | $Q_t$  | (4,4,1)        | 0.87        | 7.05                     | 0.85        | 5.58                     |
| Santeh     | (1) $Q_{t-1}$                                      | $Q_t$  | (1,3,1)        | 0.68        | 10.46                    | 0.64        | 10.25                    |
|            | (2) $Q_{t-1}, \tau$                                | $Q_t$  | (2,4,1)        | 0.84        | 7.96                     | 0.76        | 8.33                     |
|            | (3) $Q_{t-1}, Q_{t-2}$                             | $Q_t$  | (2,3,1)        | 0.77        | 9.02                     | 0.68        | 9.59                     |
|            | (4) $Q_{t-1}, Q_{t-2}, \tau$                       | $Q_t$  | (3,20,1)       | 0.85        | 7.48                     | 0.75        | 8.92                     |
|            | (5) $Q_{t-1}, Q_{t-2}, Q_{t-3}$                    | $Q_t$  | (3,5,1)        | 0.80        | 8.98                     | 0.70        | 9.25                     |
|            | <sup>a</sup> (6) $Q_{t-1}, Q_{t-2}, Q_{t-3}, \tau$ | $Q_t$  | <b>(4,2,1)</b> | <b>0.81</b> | <b>8.34</b>              | <b>0.77</b> | <b>8.25</b>              |
| Polanian   | (1) $Q_{t-1}$                                      | $Q_t$  | (1,11,1)       | 0.62        | 20.73                    | 0.78        | 25.97                    |
|            | <sup>a</sup> (2) $Q_{t-1}, \tau$                   | $Q_t$  | <b>(2,8,1)</b> | <b>0.85</b> | <b>13.33</b>             | <b>0.84</b> | <b>22.18</b>             |
|            | (3) $Q_{t-1}, Q_{t-2}$                             | $Q_t$  | (2,9,1)        | 0.79        | 15.31                    | 0.81        | 24.21                    |
|            | (4) $Q_{t-1}, Q_{t-2}, \tau$                       | $Q_t$  | (3,4,1)        | 0.74        | 16.83                    | 0.80        | 25.41                    |
|            | (5) $Q_{t-1}, Q_{t-2}, Q_{t-3}$                    | $Q_t$  | (3,6,1)        | 0.61        | 20.40                    | 0.73        | 28.33                    |
|            | (6) $Q_{t-1}, Q_{t-2}, Q_{t-3}, \tau$              | $Q_t$  | (4,17,1)       | 0.88        | 12.27                    | 0.78        | 26.25                    |

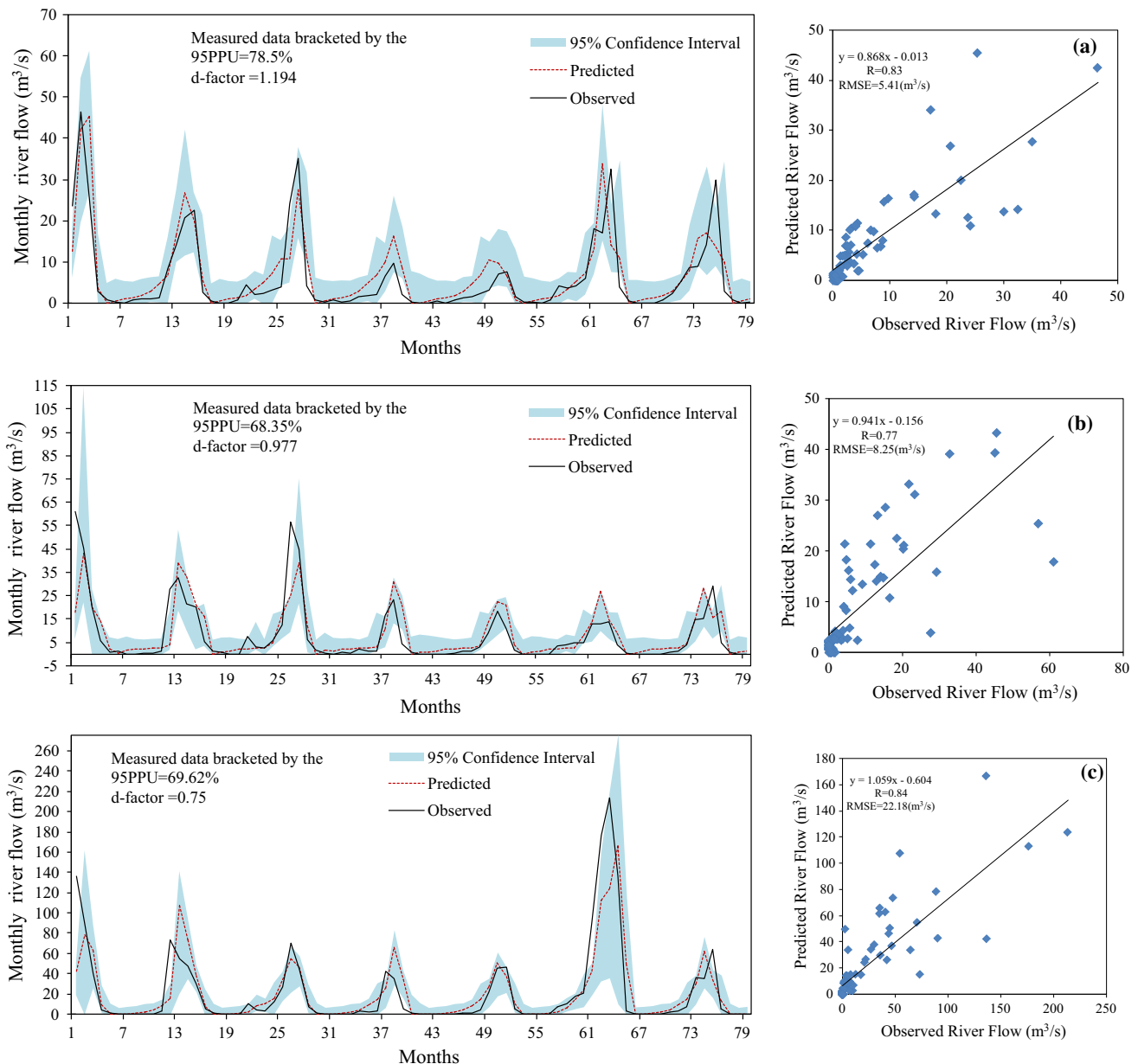
<sup>a</sup> The results in bold show the selected model

and Fig. 6 prove that the MLP model is able to predict monthly river flow, perfectly. The values bracketed by 95PPU indicate that about 79, 68 and 70 % of the data were bracketed by the 95PPU for Safakhaneh, Santeh and Polanian, respectively. Also the *d-factor* had a value of 1.194, 0.977 and 0.75 for Safakhaneh, Santeh and Polanian station, respectively. The values bracketed by 95PPU and *d-factor* values for all station indicating small prediction uncertainties.

For the all stations, to find the best results of RBF, trial and error procedure is used to approve a suitable hidden

neurons number (HN) and spread value (*r*) is selected by normalization method for each combination. The training and testing of RBF network were done using the same data sets applied in the MLP network. The best HN ranged from 3 to 40 and best spread value (*r*) is from 4 to 83 for the all combinations. The effect of changing the number of hidden neurons and spread value on the CC and RMSE for each combination is presented in Table 4.

The performance criteria show that RBF(HN:6, *r*:19), RBF(HN:36, *r*:22) and RBF(HN:14, *r*:83) structure performs relatively better than the other combinations for the



**Fig. 6** Comparative plots of observed and predicted monthly river flow by MLP model for testing period (2004–2011): **a** Safakhaneh, **b** Santeh, **c** Polanian



Safakhaneh, Santeh and Polanian stations, respectively. Here, the RBF(HN:6,  $r$ :19) structure denotes an RBF model comprising six hidden neurons number and spread value equal to 19.

Table 4 shows that the decrease in RMSE values of testing set due to including a time index within the inputs of the models is ranged from 4.4 to 17.2 %, 2.6 to 18.6 % and 2.3 to 18.2 % for Safakhaneh, Santeh and Polanian,

respectively. In other words, adding a time index within the inputs of the RBF model increases their performance significantly.

Figure 7 shows the comparative plots of the results of RBF model and the observed monthly river flow for testing data set. From Fig. 7 and Table 4, it can be observed that predicted and observed river flow was reasonably good.

**Table 4** Performance of RBF model for monthly river flow prediction

| Station    | Input combination                     | Output | Structure                          | Training    |                          | Testing     |                          |
|------------|---------------------------------------|--------|------------------------------------|-------------|--------------------------|-------------|--------------------------|
|            |                                       |        |                                    | CC          | RMSE (m <sup>3</sup> /s) | CC          | RMSE (m <sup>3</sup> /s) |
| Safakhaneh | (1) $Q_{t-1}$                         | $Q_t$  | HN: 3; $r$ : 9<br>BF: Gaussian     | 0.65        | 10.83                    | 0.72        | 6.69                     |
|            | <sup>a</sup> (2) $Q_{t-1}, \tau$      | $Q_t$  | HN: 6; $r$ : 19<br>BF: Gaussian    | <b>0.76</b> | <b>9.28</b>              | <b>0.82</b> | <b>5.54</b>              |
|            | (3) $Q_{t-1}, Q_{t-2}$                | $Q_t$  | HN: 9; $r$ : 23<br>BF: Gaussian    | 0.78        | 9.03                     | 0.75        | 6.76                     |
|            | (4) $Q_{t-1}, Q_{t-2}, \tau$          | $Q_t$  | HN: 16; $r$ : 42.2<br>BF: Gaussian | 0.83        | 8.05                     | 0.82        | 6.01                     |
|            | (5) $Q_{t-1}, Q_{t-2}, Q_{t-3}$       | $Q_t$  | HN: 15; $r$ : 7<br>BF: Gaussian    | 0.88        | 6.85                     | 0.75        | 6.30                     |
|            | (6) $Q_{t-1}, Q_{t-2}, Q_{t-3}, \tau$ | $Q_t$  | HN: 20; $r$ : 51<br>BF: Gaussian   | 0.83        | 7.91                     | 0.81        | 6.02                     |
| Santeh     | (1) $Q_{t-1}$                         | $Q_t$  | HN: 3; $r$ : 50.8<br>BF: Gaussian  | 0.69        | 10.32                    | 0.64        | 10.06                    |
|            | <sup>a</sup> (2) $Q_{t-1}, \tau$      | $Q_t$  | HN: 36; $r$ : 22<br>BF: Gaussian   | <b>0.92</b> | <b>5.50</b>              | <b>0.80</b> | <b>8.19</b>              |
|            | (3) $Q_{t-1}, Q_{t-2}$                | $Q_t$  | HN: 16; $r$ : 4<br>BF: Gaussian    | 0.91        | 5.76                     | 0.68        | 9.98                     |
|            | (4) $Q_{t-1}, Q_{t-2}, \tau$          | $Q_t$  | HN: 5; $r$ : 27.8<br>BF: Gaussian  | 0.77        | 8.99                     | 0.70        | 9.26                     |
|            | (5) $Q_{t-1}, Q_{t-2}, Q_{t-3}$       | $Q_t$  | HN: 11; $r$ : 6<br>BF: Gaussian    | 0.88        | 6.63                     | 0.70        | 9.62                     |
|            | (6) $Q_{t-1}, Q_{t-2}, Q_{t-3}, \tau$ | $Q_t$  | HN: 8; $r$ : 72.6<br>BF: Gaussian  | 0.78        | 8.92                     | 0.71        | 9.37                     |
| Polanian   | (1) $Q_{t-1}$                         | $Q_t$  | HN: 13; $r$ : 34<br>BF: Gaussian   | 0.67        | 18.70                    | 0.80        | 24.77                    |
|            | <sup>a</sup> (2) $Q_{t-1}, \tau$      | $Q_t$  | HN: 14; $r$ : 83<br>BF: Gaussian   | <b>0.74</b> | <b>16.92</b>             | <b>0.87</b> | <b>20.27</b>             |
|            | (3) $Q_{t-1}, Q_{t-2}$                | $Q_t$  | HN: 15; $r$ : 57<br>BF: Gaussian   | 0.81        | 14.68                    | 0.80        | 24.70                    |
|            | (4) $Q_{t-1}, Q_{t-2}, \tau$          | $Q_t$  | HN: 14; $r$ : 75.8<br>BF: Gaussian | 0.81        | 14.84                    | 0.81        | 24.12                    |
|            | (5) $Q_{t-1}, Q_{t-2}, Q_{t-3}$       | $Q_t$  | HN: 35; $r$ : 56<br>BF: Gaussian   | 0.86        | 12.57                    | 0.82        | 23.62                    |
|            | (6) $Q_{t-1}, Q_{t-2}, Q_{t-3}, \tau$ | $Q_t$  | HN: 40; $r$ : 53<br>BF: Gaussian   | 0.89        | 11.44                    | 0.83        | 22.99                    |

HN hidden neurons,  $r$  spread value, BF basic function

<sup>a</sup> The results in bold show the selected model

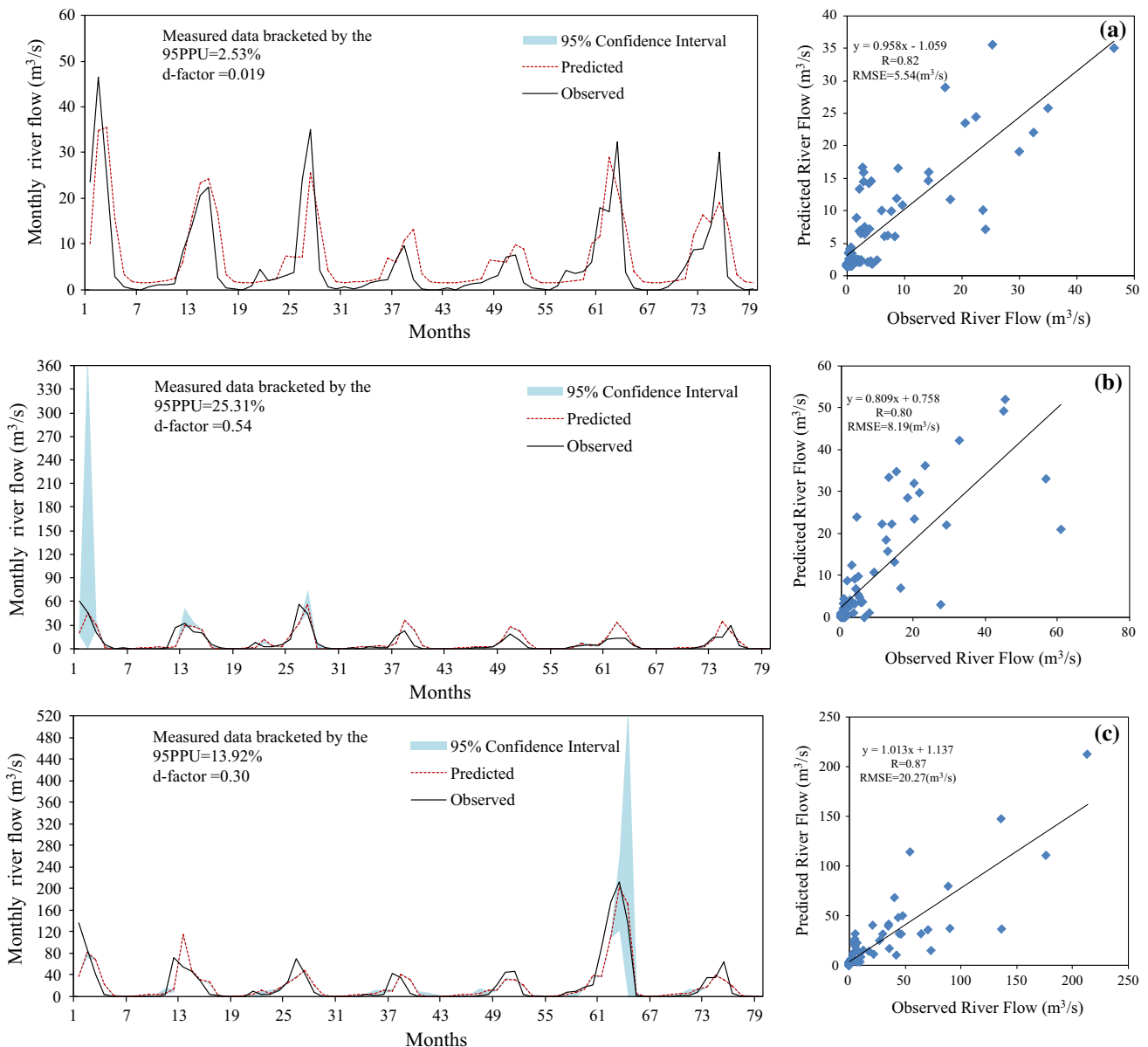
The values bracketed by 95PPU indicate that about 3, 25 and 14 % of the data were bracketed by the 95PPU for Safakhaneh, Santeh and Polanian, respectively. Also the *d*-factor had a value of 0.019, 0.54 and 0.3 for Safakhaneh, Santeh and Polanian station, respectively.

**SVM model**

There are two main steps for developing a SVM: (a) the selection of the kernel function, and (b) the identification of the specific parameters of the kernel function, i.e. *C* and  $\epsilon$ . Many studies on the use of SVM in hydrological modeling have demonstrated the favorable performance of the radial

basis function (Khan and Coulibaly 2006; Lin et al. 2006; Liong and Sivapragasam 2002; Yu et al. 2006). Therefore, in this study the RBF kernel with parameters (*C*,  $\epsilon$ ,  $\gamma$ ) is used as the kernel function for river flow modeling and the accuracy of a SVM model is dependent to identify the parameters. To obtain a suitable value of these parameters (*C*,  $\epsilon$ ,  $\gamma$ ), the RMSE was used to optimize parameters. The results of the RBF kernel based for each data combination in terms of *CC* and RMSE are given in Table 5. For each combination of inputs, the values of kernel parameters (*C*,  $\epsilon$ ,  $\gamma$ ) that provide the minimum RMSE were used.

Of the developed models, the model of input combination (6) with kernel parameters (13.14, 0.1, 7.31) for the



**Fig. 7** Comparative plots of observed and predicted monthly river flow by RBF model for testing period (2004–2011): **a** Safakhaneh, **b** Santeh, **c** Polanian

Safakhaneh station, the model of input combination (2) with kernel parameters (3.17, 0.2, 2.37) for the Santeh station and the model of input combination (4) with kernel parameters (10.56, 1.3, 6.49) shows better performance than the other combinations.

Table 5 shows that the decrease in RMSE values for testing duration due to including a time index within the inputs of the models ranged from 3.8 to 21.8 %, 6 to 15.5 % and 7.8 to 27.6 % for Safakhaneh, Santeh and Polanian, respectively.

Figure 8 shows the comparative plots of the results obtained from the SVM model the observed monthly river flow for testing data set. From Fig. 8 and Table 5, it can be observed a relatively good match between the observed and predicted river flow especially for low values. The values bracketed by 95PPU indicate that about 4, 3 and 6 % of the data were bracketed by the 95PPU for Safakhaneh, Santeh and Polanian, respectively. Also the *d-factor* had a value of 0.02, 0.013 and 0.027 for Safakhaneh, Santeh and Polanian station, respectively.

The results have in general smaller prediction uncertainties as indicated by smaller *d-factors*.

### Comparison of ANNs (MLP, RBF) and SVM models

The performances of modelling river flow time series are compared using the three techniques of MLP, RBF and SVM. The values of performance measures are given in Table 6, which indicates that the performance of each model on river flow prediction is acceptable and these approaches are applicable for modeling river flow time series data.

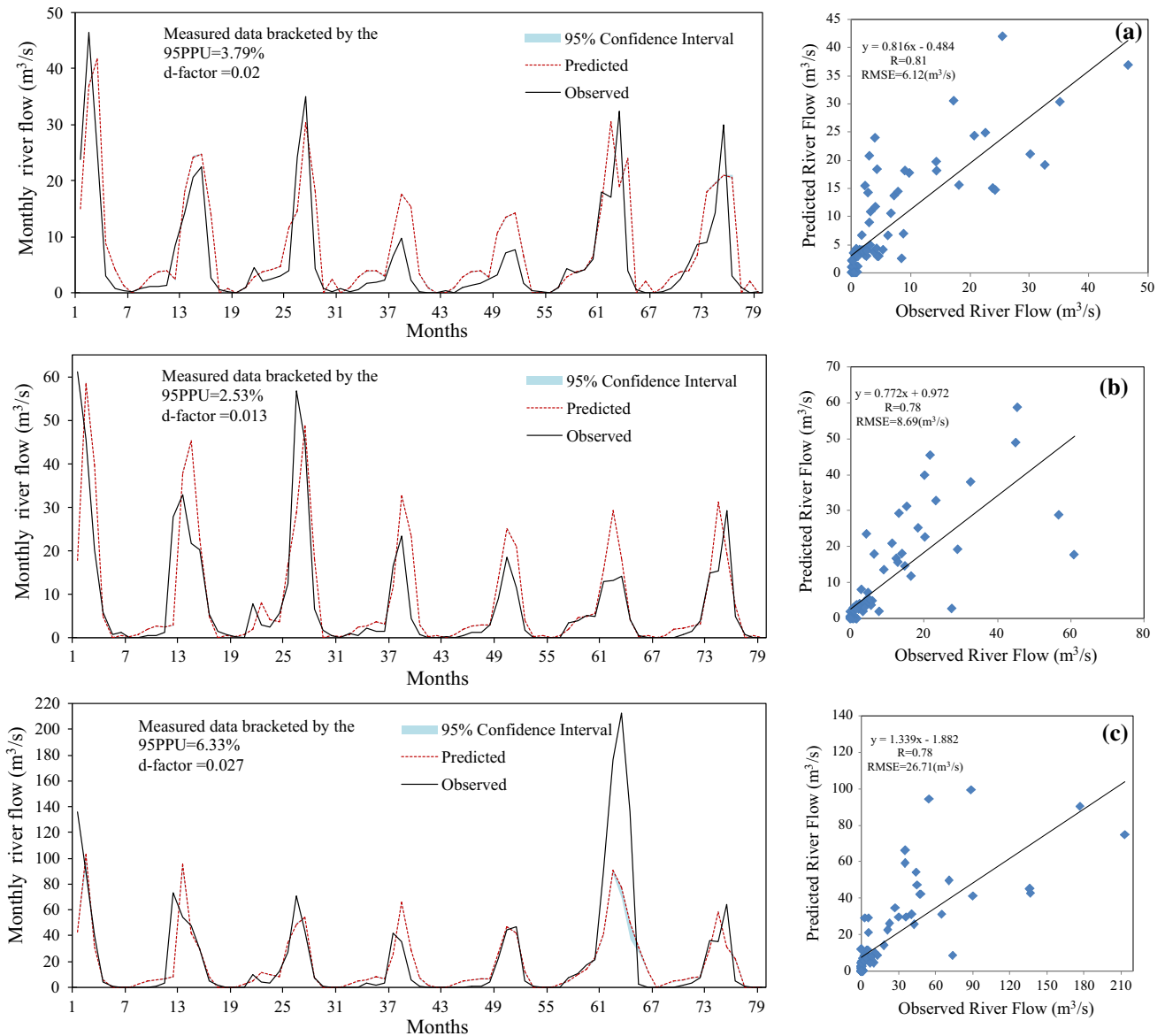
The arithmetic mean of CC values of three stations for MLP, RBF and SVM models are 0.813, 0.830 and 0.790, respectively. Also, MLP, RBF and SVM models of three stations have an arithmetic mean of RMSE values of 11.947, 11.333, 13.840 m<sup>3</sup>/s, respectively. As a result of the comparison of CC and RMSE values of the monthly river flow models, it was obtained that the results of the RBF and MLP models are better than SVM model. Also, RBF model is slightly better than MLP model in prediction of monthly river flow. This study shows that the way forward is to understand the performance of the individual models and use as many models as possible to gather more evidence for the selected models.

**Table 5** Performance of SVM model for monthly river flow prediction

| Station    | Input combination                                  | Output | Structure*   | Training    |                          | Testing     |                          |
|------------|--|--------|--|-------------|--------------------------|-------------|--------------------------|
|            |  |        |  | CC          | RMSE (m <sup>3</sup> /s) | CC          | RMSE (m <sup>3</sup> /s) |
| Safakhaneh | (1) $Q_{t-1}$                                      | $Q_t$  | C: 3710.37, $\epsilon$ : 0.1, $\gamma$ : 21,795.36 | 0.95        | 4.97                     | 0.62        | 8.90                     |
|            | (2) $Q_{t-1}, \tau$                                | $Q_t$  | C: 1.67, $\epsilon$ : 0.3, $\gamma$ : 36.21        | 0.93        | 5.42                     | 0.79        | 7.72                     |
|            | (3) $Q_{t-1}, Q_{t-2}$                             | $Q_t$  | C: 16.47, $\epsilon$ : 1.5, $\gamma$ : 16.38       | 0.84        | 7.73                     | 0.71        | 8.07                     |
|            | (4) $Q_{t-1}, Q_{t-2}, \tau$                       | $Q_t$  | C: 1.93, $\epsilon$ : 0.1, $\gamma$ : 8.41         | 0.93        | 5.41                     | 0.79        | 7.76                     |
|            | (5) $Q_{t-1}, Q_{t-2}, Q_{t-3}$                    | $Q_t$  | C: 76.47, $\epsilon$ : 1.3, $\gamma$ : 24.27       | 0.76        | 9.28                     | 0.63        | 7.83                     |
|            | <sup>a</sup> (6) $Q_{t-1}, Q_{t-2}, Q_{t-3}, \tau$ | $Q_t$  | C: 13.14, $\epsilon$ : 0.1, $\gamma$ : 7.31        | <b>0.85</b> | <b>7.50</b>              | <b>0.81</b> | <b>6.12</b>              |
| Santeh     | (1) $Q_{t-1}$                                      | $Q_t$  | C: 2.43, $\epsilon$ : 0.3, $\gamma$ : 3.72         | 0.71        | 10.05                    | 0.63        | 10.29                    |
|            | <sup>a</sup> (2) $Q_{t-1}, \tau$                   | $Q_t$  | C: 3.17, $\epsilon$ : 0.2, $\gamma$ : 2.37         | <b>0.87</b> | <b>6.96</b>              | <b>0.78</b> | <b>8.69</b>              |
|            | (3) $Q_{t-1}, Q_{t-2}$                             | $Q_t$  | C: 5.71, $\epsilon$ : 0.4, $\gamma$ : 53           | 0.83        | 7.91                     | 0.67        | 11.63                    |
|            | (4) $Q_{t-1}, Q_{t-2}, \tau$                       | $Q_t$  | C: 12.13, $\epsilon$ : 2, $\gamma$ : 123.36        | 0.89        | 6.30                     | 0.76        | 10.35                    |
|            | (5) $Q_{t-1}, Q_{t-2}, Q_{t-3}$                    | $Q_t$  | C: 7.52, $\epsilon$ : 0.2, $\gamma$ : 46.87        | 0.83        | 7.84                     | 0.67        | 10.86                    |
|            | (6) $Q_{t-1}, Q_{t-2}, Q_{t-3}, \tau$              | $Q_t$  | C: 39.26, $\epsilon$ : 0.1, $\gamma$ : 1160.8      | 0.89        | 6.50                     | 0.76        | 10.21                    |
| Polanian   | (1) $Q_{t-1}$                                      | $Q_t$  | C: 955.10, $\epsilon$ : 1.2, $\gamma$ : 5002.7     | 0.64        | 19.14                    | 0.66        | 31.40                    |
|            | (2) $Q_{t-1}, \tau$                                | $Q_t$  | C: 6.41, $\epsilon$ : 0.8, $\gamma$ : 6.03         | 0.84        | 13.57                    | 0.77        | 26.88                    |
|            | (3) $Q_{t-1}, Q_{t-2}$                             | $Q_t$  | C: 100.02, $\epsilon$ : 1.7, $\gamma$ : 5461.86    | 0.77        | 15.95                    | 0.63        | 36.90                    |
|            | <sup>a</sup> (4) $Q_{t-1}, Q_{t-2}, \tau$          | $Q_t$  | C: 10.56, $\epsilon$ : 1.3, $\gamma$ : 6.49        | <b>0.84</b> | <b>13.74</b>             | <b>0.78</b> | <b>26.71</b>             |
|            | (5) $Q_{t-1}, Q_{t-2}, Q_{t-3}$                    | $Q_t$  | C: 36.56, $\epsilon$ : 2, $\gamma$ : 363.71        | 0.79        | 15.38                    | 0.71        | 29.19                    |
|            | (6) $Q_{t-1}, Q_{t-2}, Q_{t-3}, \tau$              | $Q_t$  | C: 11.19, $\epsilon$ : 1.0, $\gamma$ : 6.94        | 0.85        | 13.43                    | 0.78        | 26.89                    |

\* C: Magnitude of penalty term,  $\epsilon$ : width/deviation of the error margin,  $\gamma$ : Gaussian radial basis function parameter

<sup>a</sup> The results in bold show the selected model



**Fig. 8** Comparative plots of observed and predicted monthly river flow by SVM model for testing period (2004–2011): **a** Safakhaneh, **b** Santeh, **c** Polanian

**Table 6** Comparison of ANNs (MLP, RBF) and SVM models for monthly river flow prediction for testing period

| Model | Safakhaneh |                          | Santeh |                          | Polanian |                          | Arithmetic mean of performance criteria |                          |
|-------|------------|--------------------------|--------|--------------------------|----------|--------------------------|---|--------------------------|
|       | CC         | RMSE (m <sup>3</sup> /s) | CC     | RMSE (m <sup>3</sup> /s) | CC       | RMSE (m <sup>3</sup> /s) | CC                                      | RMSE (m <sup>3</sup> /s) |
| MLP   | 0.83       | 5.41                     | 0.77   | 8.25                     | 0.84     | 22.18                    | 0.813                                   | 11.947                   |
| RBF   | 0.82       | 5.54                     | 0.80   | 8.19                     | 0.87     | 20.27                    | 0.830                                   | 11.333                   |
| SVM   | 0.81       | 6.12                     | 0.78   | 8.69                     | 0.78     | 26.71                    | 0.790                                   | 13.840                   |

## Conclusion

The performances of three modelling techniques are reported by this study to provide evidence for suitable techniques for predicting monthly river flow values. The studied techniques comprise: Artificial Neural Networks (MLP, RBF) and Support Vector Machines (SVM).

Monthly river flow time series of three stations, i.e. Safakhaneh, Santeh and Polanian from Zarrinehrud River over a period of 23 years (1989–2011) were used in this study. The artificial intelligence models have advantages of being able to analyze and generalize relationships between parameters into the model, but it can be difficult or impossible using all parameters by developed models for the rivers having different climatical/topographical characteristics. Therefore, cross correlations analysis for various monthly river flow lags was used to find out the number of past observations to provide effective inputs to the models.

The study shows that the MLP, RBF and SVM models can be successfully applied to the tasks of prediction of monthly river flow. Although it is shown that the MLP, RBF models performs better than SVM and including a time index within the inputs increases their performance significantly.

In addition, the reliability of the models prediction was calculated by an uncertainty estimation. The results indicate that the uncertainty in the SVM model was less than those in the MLP and RBF models for predicting monthly river flow.

**Acknowledgments** The author is grateful for editor and anonymous reviewers for their helpful and constructive comments which greatly improved the quality of this paper.

## References

- Abbaspour KC, Yang J, Maximov I, Siber R, Bogner K, Mieleitner J, Zobrist J, Srinivasan R (2007) Modeling hydrology and water quality in the pre-alpine/alpine Thur watershed using SWAT. *J Hydrol* 333(2–4):413–430
- ASCE Task Committee on Application of Artificial Neural Networks in Hydrology (2000) Artificial neural networks in hydrology I: preliminary concepts. *J Hydrol Eng* 5:115–123
- Awchi TA (2014) River discharges forecasting in northern Iraq using different ANN techniques. *Water Resour Manag* 28:801–814
- Azimi M, Heshmati GhA, Farahpour M, Faramarzi M, Abbaspour KC (2013) Modeling the impact of rangeland management on forage production of sagebrush species in arid and semi-arid regions of Iran. *Ecol Model* 250:1–14
- Bayazit M (1988) Hidrolojik modeller. İ.T.Ü. Rektörlüğü, İstanbul
- Bhagwat PP, Maity R (2012) Multistep-ahead river flow prediction using LS-SVR at daily scale. *J Water Resour Prot* 4:528–539
- Chang CC, Lin CJ (2005) LIBSVM—a library for support vector machines, version 2.8. software available at <http://www.csie.ntu.edu.tw/~cjlin/libsvm>
- Coccia G, Todini E (2011) Recent developments in predictive uncertainty assessment based on the model conditional processor approach. *Hydrol Earth Syst Sci* 15:3253–3274
- Cohen S, Intrator N (2002) Automatic model selection in a hybrid perceptron/radial network. *Inf Fusion Special Issue Mult Experts* 3(4):259–266
- Demuth H, Beale M (2001) Neural network toolbox user's guide—version 4. The MathworksInc, Natick, p 840
- Frick J, Hegg C (2011) Can end-users' flood management decision making be improved by information about forecast uncertainty? *Atmos Res* 100:296–303
- Ghorbani MA, Khatibi R, Hosseini B, Bilgili M (2013) Relative importance of parameters affecting wind speed prediction using artificial neural networks. *Theor Appl Climatol* 114(1):107–114
- Haykin S (1999) Neural networks: a comprehensive foundation. Macmillan, New York
- Jayawardena AW, Fernando DAK (1995) Artificial neural networks in hydrometeorological modelling. In: Proceedings of the Fourth International conference on the application of artificial intelligence to civil and structural engineering—developments in neural networks and evolutionary computing in civil and structural engineering, Cambridge, UK
- Kalteh AM (2013) Monthly river flow forecasting using artificial neural network and support vector regression models coupled with wavelet transform. *Comput Geosci* 54:1–8
- Kenneth J, Wernter S, MacInyre J (2001) Knowledge extraction from radial basis function networks and multilayer perceptrons. *Int J Comput Intell Appl* 1(3):369–382
- Khan MS, Coulibaly P (2006) Application of support vector machine in lake water level prediction. *J Hydrol Eng* 11:199–205
- Khatibi R, Sivakumar B, Ghorbani MA, Kisi O, Kocak K, FarsadiZadeh D (2012) Investigating chaos in river stage and discharge time series. *J Hydrol* 414–415:108–117
- Kisi O (2005) Daily river flow forecasting using artificial neural networks and auto-regressive models. *J Environ Sci Eng* 29:9–20
- Kisi O (2008) River flow forecasting and estimation using different artificial neural network technique. *Hydrol Res* 39(1):27–40
- Kisi O, Cimen M (2011) A wavelet-support vector machine conjunction model for monthly streamflow forecasting. *J Hydrol* 399:132–140
- Kisi O, Moghaddam Nia A, GhafariGosheh M, JamalizadehTajabadi MR, Ahmadi A (2012) Intermittent streamflow forecasting by using several data driven techniques. *Water Resour Manag* 26:457–474
- Lin JY, Cheng CT, Chau KW (2006) Using support vector machines for long-term discharge prediction. *Hydrol Sci J* 51:599–612
- Liong SY, Sivapragasam C (2002) Flood stage Forecasting with support vector machines. *J Am Water Resour* 38:173–186
- Loh W, Tim L (2000) A comparison of prediction accuracy, complexity, and training time of thirty three old and new classification algorithm. *Mach Learn* 40(3):203–238
- Luk KC, Ball JE, Sharma A (2000) A study of optimal model lag and spatial inputs to artificial neural network for rainfall forecasting. *J Hydrol* 227:56–65
- Misra D, Oommen T, Agarwal A, Mishra SK, Thompson AM (2009) Application and analysis of support vector machine based simulation for runoff and sediment yield. *Biosyst Eng* 103:527–535
- Montanari A, Grossi G (2008) Estimating the uncertainty of hydrological forecasts: a statistical approach. *Water Resour Res* 44(12):8–17
- Noori R, Deng Z, Kiaghadi A, Kachoosangi F (2015) How reliable are ANN, ANFIS, and SVM techniques for predicting longitudinal dispersion coefficient in natural rivers? *J Hydraul Eng*. doi:10.1061/(ASCE)HY.1943-7900.0001062.04015039

- Sahoo GB, Ray C (2006) Flow forecasting for a Hawaii stream using rating curves and neural networks. *J Hydrol* 317:63–80
- Samsudin R, Saad P, Shabri A (2011) River flow time series using least squares support vector machines. *Hydrol Earth Syst Sci* 15:1835–1852
- Sudheer KP, Jain SK (2003) Radial basis function neural network for modeling rating curves. *J Hydrol Eng* 8:161–164
- Terzi O (2011) Monthly river flow forecasting by data mining process, knowledge-oriented applications in data mining, prof. kimitofunatsu (Ed.), ISBN: 978-953-307-154-1, InTech, Available from <http://www.intechopen.com/books/knowledge-oriented-applications-in-data-mining/monthly-river-flowforecasting-by-data-mining-process>
- Terzi O, Ergin G (2014) Forecasting of monthly river flow with autoregressive modeling and data-driven techniques. *Neural Comput Appl* 25:179–188
- The MathWorks Inc. (2012) Matlab the language of technical computing. Retrieved 4 September, from <http://www.mathworks.nl/products/matlab>
- Vapnik VN (1995) The nature of statistical learning theory. Springer, New York, p 314
- Vapnik VN (1998) Statistical learning theory. Wiley, New York, p 736
- Wang W, Men C, Lu W (2008) Online prediction model based on support vector machine. *Neurocomputing* 71:550–558
- Yilmaz I, Kaynar O (2011) Multiple regression, ANN (RBF, MLP) and ANFIS models for prediction of swell potential of clayey soils. *Expert Syst Appl* 38:5958–5966
- Yoon H, Jun SC, Hyun Y, Bae GO, Lee KK (2011) A comparative study of artificial neural networks and support vector machines for predicting groundwater levels in a coastal aquifer. *J Hydrol* 396:128–138
- Yu PS, Chen ST, Chang IF (2006) Support vector regression for real-time flood stage forecasting. *J Hydrol* 328:704–716
- Yu H, Xie T, Paszezynski S, Wilamowski BM (2011) Advantages of radial basis function networks for dynamics system design. *IEEE Trans Ind Electron* 58(12):5438–5450