Indian Journal of Fundamental and Applied Life Sciences ISSN: 2231– 6345 (Online)
An Open Access, Online International Journal Available at www.cibtech.org/sp.ed/jls/2015/04/jls.htm
2015 Vol. 5 (S4), pp. 322-329/Malekzadeh and Fereydooni
Research Article

EVALUATION OF MLP NEURAL NETWORK IN FLOW DISCHARGE PREDICTION IN TANGAB DAM BASINFIROUZABAD RIVER

Mohsen Malek Zade and *Mehrdad Fereydooni
Department of Civil Engineering, Larestan Branch, Islamic Azad University, Larestan, Iran
**Author for Correspondence*

ABSTRACT

Accurate prediction of river flow for Safe and economic design of facilities and structures as well as establishing appropriate procedures and guidelines during years of drought and rain, and also proper management and utilization of water resources for sustainable development have been always taken into consideration by experts. In the present research, the application of MLP artificial neural network in management and simulation of Tangab dam basin (Firouzabad River) in Fars province was studied. After processing the data (discharge, evaporation, rainfall, temperature) the dependence of variables was determined and, using Corrcoef order, the correlation coefficient of different variables could be identified and the input patterns to the neural network could be determined. For neural network training, the data is first normalized and then it is classified so that 70% of the data is considered as training input and output, and 30% is considered as validation input and output. After determining input patterns, the best neural network structure for river discharge (BR.LM) and (FNN) in the output range of (0 and 1) based on the assessment model including root mean square error (RMSE) and correlation coefficient (R^2) was compared with different structures. In that study, the sixth input pattern with 7 neurons in the first layer and 7 neurons in the middle layer and RMSE equal to 0/062 and correlation coefficient equal to 0/85 predicted the best modeling result. Furthermore, in order to compare the previously observed data and the ones calculated from neural network as well as seeing neural network function, both their diagrams are drawn.

Keywords: *Discharge Prediction, Neural Network, Tangab Dam, Firouzabad River, Correlation Coefficient*

INTRODUCTION

Using neural network, Zhang & Govindaraju (2000) predicted the amount of monthly runoff in three areas of medium size in Kansas, the United States. Results of the study indicated that the proposed network for predicting minimum and average flows provides acceptable results but it does not show its high efficiency in maximum flow forecast. Using artificial neural network, Lorrai and Sechi (1995) modeled the monthly rainfall-runoff process in several different areas. Results indicated that the neural network method can accurately predict the monthly runoff. Atiya *et al.,* (1999) used artificial neural network to predict the average discharge of the Nile River. In this study, the discharge data of the previous period was used as the network input and results showed that in order to do long-term predictions such as annual prediction it is preferable to provide a model for annual prediction and do the prediction as incorporated. Using Box Jenkins Model and artificial neural network, MariaKaslanvMendez *et al.,* (2004) did the monthly and daily discharge predictions in Kalas River in North West of Spain. Findings indicated that the neural network made a more accurate prediction. Nilsson *et al.,* (2005) used conceptual models and neural networks as well as a combination of them to simulate the monthly runoff of two auriferous areas in Norway and concluded that the neural network method and its combination with conceptual models provide a better estimation of monthly runoff for both areas. KaramSigizoghlou (2004) made a comparison between multilayer perceptron and AR model to predict daily suspended sediment. In his study, too, the superiority of the neural network's prediction ability has been proved compared to that of regression methods. Nayebi (2005) predicted the river flow of Kor River's basin using artificial neural network model in the form of four various models. The inputs of different recorded models of rain gauge, hydrometric and evaporation measurement stations of the area were considered

© *Copyright 2014 | Centre for Info Bio Technology (CIBTech)*

individually or as combined. Results of the study showed that the neural network model with the input of daily data from rain gauge stations had the best performance in terms of predicting the area's next day output. Using the neural network model, Nazarnezhad & Ghorbani (2005) studied the probability of predicting maximum monthly discharge of Nazlou River. The amounts of discharges and rainfalls in the previous month as well as the rainfalls of the current month were the inputs of the proposed model in that study. Findings showed that the amounts of predicted discharges had high accuracy and were in good agreement with the amounts of observed discharges. Karamouz *et al.,* (2003) investigated the use of artificial neural network in hydro-ecological modeling of the prediction of inflow to the reservoir. Two different approaches have been suggested to do long-term prediction of river flow. In the first approach, a model is used to make a relationship between hydro-climate variables and the river flow so that the variables rainfall, air temperature, runoff and snow budget observed currently are used as the model's inputs. In the second approach, a combination of two weather forecasting prediction models were used to predict the rainfall and air temperature, and a hydrologic model and back-propagation networks as well as comparative Adeline were used to model the prediction of Zayanderood monthly flow where the river entered Zayanderood dam.

Results obtained from the two approaches above indicated the artificial neural network had a better performance compared to that of the first approach (Karamouz and Tabari, 2003). In their investigations, Fereidouni *et al.,* (2011) used an artificial neural network to model and predict each balance parameter (rainfall, evaporation, groundwater level and runoff) in Mand and Bakhtegan basins (Jahrom, GhareAghach and Gavdar saline rivers) in three regions of different climate in Fars province. Their findings indicated proper functioning of the artificial neural network to simulate the runoff evaporation and groundwater level of the selected regions (Fereidouni *et al.,* 2011).

MATERIALS AND METHODS

Methodology

The aim of this study is to model Tangab River discharge using the statistics and information obtained from Hanifian station through applying MLP artificial neural network with the use of discharge, rainfall, temperature, and evaporation amounts measured monthly in the past.

The Research Process

1. Preprocessing of the available data and determining the dependence of effective parameters on discharge of Firouzabad River using Corrcoef Function in Matlab software
2. Determining various input patterns
3. Modeling Firouzabad River discharge using MLP artificial neural network

The Software Used in This Research

1. MATLAB R2009a software
2. Minitab software

Model Evaluation Criteria

To evaluate the efficiency of the modelings done in this research, two statistical parameters including Root Mean Square Error (RMSE) and/or the correlation coefficient (R^2) were used. The Root Mean Square Error is stated as equation 1-4:

$$RMSE = \sqrt{\frac{\sum_{i=1}^{N}(o_i - t_i)^2}{N}} \qquad (1\text{-}4)$$

In which N is the number of educational pairs, Oi is the observed discharge and ti is the estimated discharge. The large amounts of RMSE show that the modeling has not been done with high accuracy. The nearer the amount to zero, the higher the accuracy of the model is. The correlation coefficient is shown as equation 2-4:

[2-4]

In whichN is the number of educational pairs, Oi is the observed discharge, ti is the estimated discharge (model outputs) and Oi is the average discharges observed. The moles with the highest R^2 coefficient are selected as the optimal models.

© *Copyright 2014 | Centre for Info Bio Technology (CIBTech)*

Indian Journal of Fundamental and Applied Life Sciences ISSN: 2231– 6345 (Online)
An Open Access, Online International Journal Available at www.cibtech.org/sp.ed/jls/2015/04/jls.htm
2015 Vol. 5 (S4), pp. 322-329/Malekzadeh and Fereydooni

MLP Model

To determine the input patterns for modeling with the use of artificial neural networks, the sensitivity analysis has to be done on the used data. For modeling a river discharge or other things, it should first be understood which variables have an important effect on that phenomenon so that the intended variables could be considered as the model inputs. Furthermore, as any phenomenon in dynamic systems depends on its past, the extent of this dependence has to be identified. For example, when we have a second order differential equation, it means that this system's output depends on its past to the second order. Therefore, the higher the degree of differential equation, the more this dependence on further past will be. For systems whose differential equation with the extent of dependence on the past is not much clear, the dependence analysis methods are used to identify it. In neural network models, selecting the variables might also have a theoretical basis. In fact, some researchers have shown that if selecting the input variables is done based on theoretical opinions, the prediction result will be more accurate. Therefore, some statistical analyses or data dependency methods might be used so that the extent of dependency between different variables gets clear. For instance, the Corrcoef order in Matlab software environment provides the correlation coefficient of different variables which are in the form of measured quantities. Hence, using the results obtained from the Corrcoef order, we can determine the input patterns for neural networks.

One method of determining time dependency between variables is the use of Corrcoef function in Matlab software environment. This calculates the function of dependency coefficient between variables. The inputs of this function are the discharge, rainfall, temperature and evaporation in their previous delays while its output is a matrix whose elements' sizes imply the relationship between the input and output. The closer the matrix elements are to one, the more relationships between variables exist, and the closer they are to zero, the less relationship between them can be seen. For example, to find out the dependency of monthly discharge of the flow on its previous delays, the river discharge and the previous delays of the discharge are given to the function as input variables so that a matrix is obtained as can be seen in table (1-4). This matrix shows that the discharge at the time of t+1is dependent on the three delays before it which are the discharges at the times of t, t-1 and t-2. Besides, to find out the dependency of daily discharge and monthly rainfall, the flow discharge as well as the rainfalls and the delays before rainfalls are given to the function as the input. Therefore, a matrix as in table (2-4) is obtained which shows that the discharge at the time of t+1 dependence on the rainfalls at the times of t and t-1. This process is done similarly to find out the dependency of monthly discharge on the temperature and evaporation. Results can be seen in tables (3-4) and (4-4)

Teaching Artificial Neural Networks

The following procedures have to be done in order to do modeling through the use of MLP artificial neural network or any other artificial neural networks:

Data Normalization

Before entering the data, whether the input data or the output ones, they have to be normalized. Data normalization has the following advantages:

1. It causes the network to reach the intended or ideal result more quickly.
2. Normalization also shows the extent of data efficacy.

There are several ways to normalize the data used. In the present study, equation (3-4) has been used to normalize the data. This function normalizes the input and output data within 0.05 and 0.95.

$$x_{normalize} = 0.05 + 0.95\left(\frac{x - x_{min}}{x_{max} - x_{min}}\right) \tag{3-4}$$

In this equation, Xmin and Xmax are the minimum and maximum amounts while Xnormalized is the normalized data.

Data Classification

In the next step, the data must be classified. In other words, we should specify the training and validation data sets. We also specify the input as well as output data in the present study. As the used data include

© Copyright 2014 | Centre for Info Bio Technology (CIBTech)

the monthly statistics of discharge, rainfall and temperature of Hanifian station on Firouzabad River in south of Fars province from 1985 to 1986, the classification is as follows:

1. The class relates to the network training. In this thesis, 70% of the data (209 pieces of data) have been regarded as the training inputs and outputs.
2. The class relates to network validation. In this study 30% of the entire data (89 pieces of data) have been regarded as the validation inputs and outputs.

Determining Input Patterns for Modeling the Monthly River Discharge by MLP Artificial Neural Network

In this stage, according to the preprocessing done by Coorcoef order, different input patterns are determined in order to reach our final goal which is the modeling of monthly discharge by MLP artificial neural network. For all input patterns, the target vector or network output is the Shapoor River discharge at the time of t+1. The input patterns used in this study are s follows:

Input pattern 1: rainfall at the time of (t), temperature at the time of (t), evaporation at the time of (t) and discharge at the time of (t)

Input pattern 2: rainfall at the times of (t) and (t-1), temperature at the time of (t), evaporation at the time of (t) and discharge at the time of (t)

Input pattern 3: rainfall at the time of (t), temperature at the time of (t), evaporation at the time of (t) and discharge at the times of (t) and (t-1)

Input pattern 4: rainfall at the times of (t) and (t-1), temperature at the time of (t), evaporation at the time of (t) and discharge at the times of (t) and (t-1)

Input pattern 5: rainfall at the time of (t), temperature at the time of (t), evaporation at the time of (t) and discharge at the times of (t), (t-1) and (t-2)

Input pattern 6: rainfall at the times of (t) and (t-1), temperature at the time of (t), evaporation at the time of (t) and discharge at the times of (t), (t-1) and (t-2)

Modeling the Monthly Discharge of Firouzabad River through MLP Artificial Neural Network

After determining the input patterns, the best artificial neural network structure for the river discharge could be selected in order to be used in the next stages. This evaluation was done using two existing artificial neural networks composed of a combination of two algorithms and one leading network. In the present study, the algorithms and neural networks used included (*LM*, *BR*) and (FNN). In addition, the stimulus function used for the neurons in middle and outer layers must be continuous and differentiable. For approximation of functions usually two types of Sigmoid functions are used in the middle layers: Hyperbolic TangentSigmoid and Log-Sigmoid.

Due to having high flexibility and output range of (0 & 1), the Log-Sigmoid stimulus function has more applications than the other one. Hence, the Log-Sigmoid function and the linear function have been used in the middle layers and the outer layers, respectively, to model the monthly discharge of the river under investigation.

In order to determine the number of middle nodes, the try & error method was used so that for each input pattern, different neural networks were created and trained with the changes in the number of the hidden layer's neurons. Based on the model evaluation criteria including root meansquare error (RMSE) and correlation coefficient (R^2), the different structures were compared.

Now we are going to deal with monthly discharge modeling through MLP neural networks. The best results of input patterns at training and validation stages are provided in the following tables.

Determining the Best Pattern and Model for Modeling and Evaluating the Monthly Discharge of the River through MLP Artificial Neural Network

In this part of the thesis and according to the results obtained from modeling by MLP artificial neural network and the used algorithm as well as comparing the results, we are going to determine the best network and input structure as well as the best input pattern. It was found that the input pattern 6 with an FNN-LM structure and 7 neurons in the first layer, 7 neurons in the hidden layer, RMSE coefficient equal to 0.062, and correlation coefficient R^2 equal to 0.85 at validation stage provides the best result for modeling and predicting the monthly discharge.

© Copyright 2014 | Centre for Info Bio Technology (CIBTech)

RESULTS AND DISCUSSION

Table 1-4: Results of river discharge preprocessing with past data

	Q(t)	Q(t-1)	Q(t-2)	Q(t-3)	Q(t-4)
Q(t)	1	0.39	0.27	0.12	0.009
Q(t-1)	0.39	1	0.39	0.27	0.12
Q(t-2)	0.27	0.39	1	0.39	0.27
Q(t-3)	0.12	0.27	0.39	1	0.27
Q(t-4)	0.009	0.12	0.27	0.39	1

Table 2-4: Results of river discharge preprocessing with rainfall past data

	P(t)	P(t-1)	P(t-2)	P(t-3)	P(t-4)
Q(t)	1	0.72	0.40	0.29	0.11
Q(t-1)	0.72	1	0.34	0.099	-0.26
Q(t-2)	0.40	0.34	1	0.34	0.099
Q(t-3)	0.29	0.099	0.34	1	0.34
Q(t-4)	0.11	-0.267	0.099	0.34	1

Table 3-4: Results of river discharge preprocessing with temperature past data

	T(t)	T(t-1)	T(t-2)	T(t-3)	T(t-4)
Q(t)	1	-0.41	-0.36	-0.24	-0.054
Q(t-1)	-0.41	1	0.81	0.47	0.010
Q(t-2)	-0.36	0.81	1	0.81	0.47
Q(t-3)	-0.24	0.47	0.81	1	0.81
Q(t)	-0.054	0.010	0.47	0.81	1

Table 4-4: Results of river discharge preprocessing with evaporation past data

	E(t)	E(t-1)	E(t-2)	E(t-3)	E(t-4)
Q(t)	1	-0.41-	-0.38	-0.25	-0.086
Q(t-1)	-0.41	1	0.848	0.49	0.036
Q(t-2)	-0.38	0.848	1	0.848	0.49
Q(t-3)	-0.25	0.49	0.84	1	0.84
Q(t-4)	-.08	0.36	0.49	0.849	1

Table 5-4: Results of monthly discharge modeling by the training network with MLP algorithm at LM stage

Input pattern	Network	Network structure	R^2	RMSE
1	**FNN-LM**	4-6-2	0.87	0.08
2	**FNN-LM**	5-7-1	0.88	0.07
3	**FNN-LM**	5-5-1	0.85	0.074
4	**FNN-LM**	6-3-1	0.85	0.08
5	**FNN-LM**	6-20-1	0.84	0.082
6	**FNN-LM**	7-7-1	0.89	0.072

© Copyright 2014 | Centre for Info Bio Technology (CIBTech)

Table 5-4: Results of monthly discharge modeling by the validation network with MLP algorithm at LM stage

Input pattern	Network	Network structure	R^2	RMSE
1	**FNN-LM**	4-6-1	0.77	0.081
2	**FNN-LM**	5-7-1	0.80	0.075
3	**FNN-LM**	5-5-1	0.75	0.084
4	**FNN-LM**	6-3-1	0.79	0.085
5	**FNN-LM**	6-20-1	0.78	0.083
6	**FNN-LM**	7-7-1	0.85	0.062

Table 5-4: Results of monthly discharge modeling by the MLP network with BR algorithm at training stage

Input pattern	Network	Network structure	R^2	RMSE
1	**FNN-BR**	4-6-1	0.84	0.082
2	**FNN-BR**	5-5-1	0.83	0.079
3	**FNN-BR**	5-5-1	0.79	0.08
4	**FNN-BR**	6-15-1	0.82	0.075
5	**FNN-BR**	6-10-1	0.82	0.079
6	**FNN-BR**	7-10-1	0.86	0.078

Table 5-4: Results of monthly discharge modeling by the MLP network with BR algorithm at validation stage

Input pattern	Network	Network structure	R^2	RMSE
1	**FNN-BR**	4-6-1	0.78	0.084
2	**FNN-BR**	5-5-1	0.73	0.082
3	**FNN-BR**	5-5-1	0.76	0.082
4	**FNN-BR**	6-15-1	0.80	0.08
5	**FNN-BR**	6-10-1	0.81	0.073
6	**FNN-BR**	7-10-1	0.80	0.075

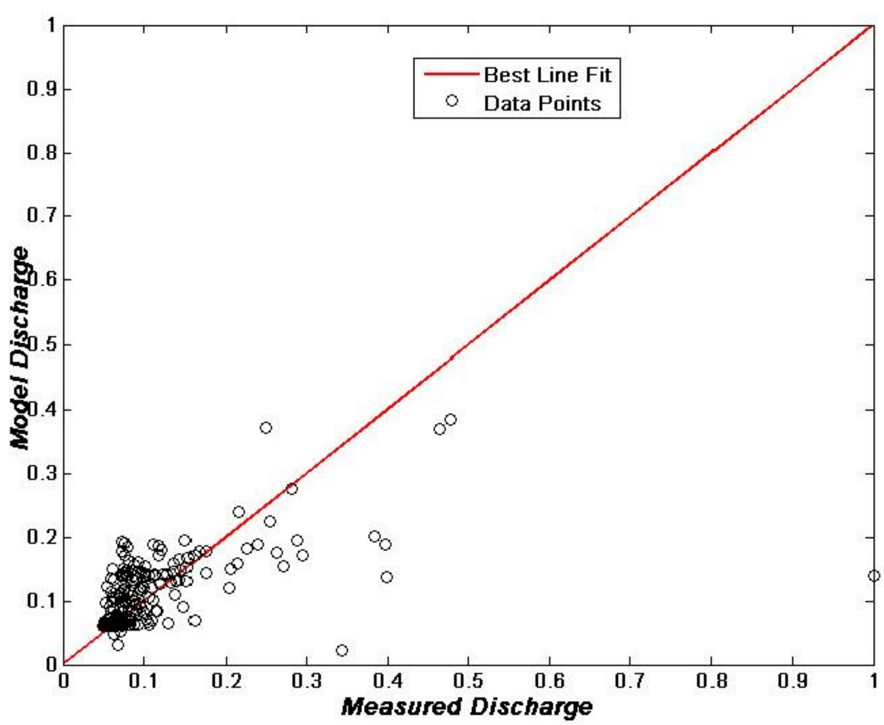

Figure 1-4: The best fitted line across the points at the training stage

© Copyright 2014 | Centre for Info Bio Technology (CIBTech)

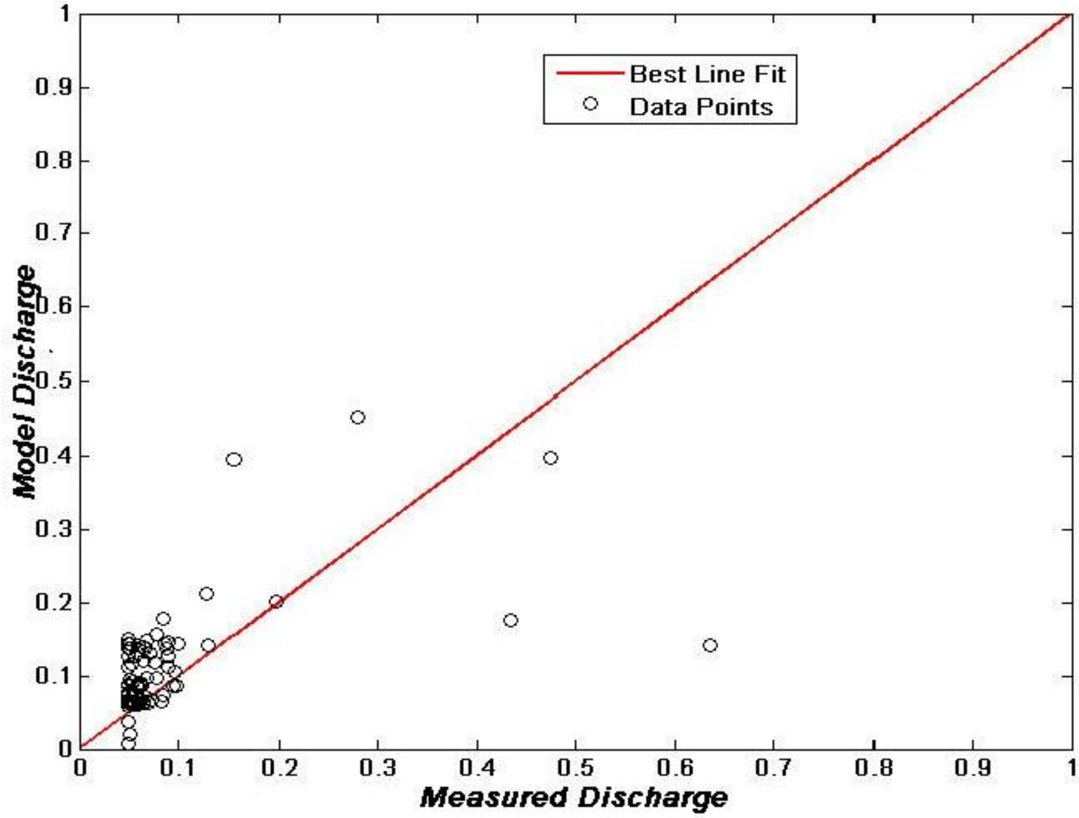

Figure 2-4: The best fitted line across the points at the validation stage

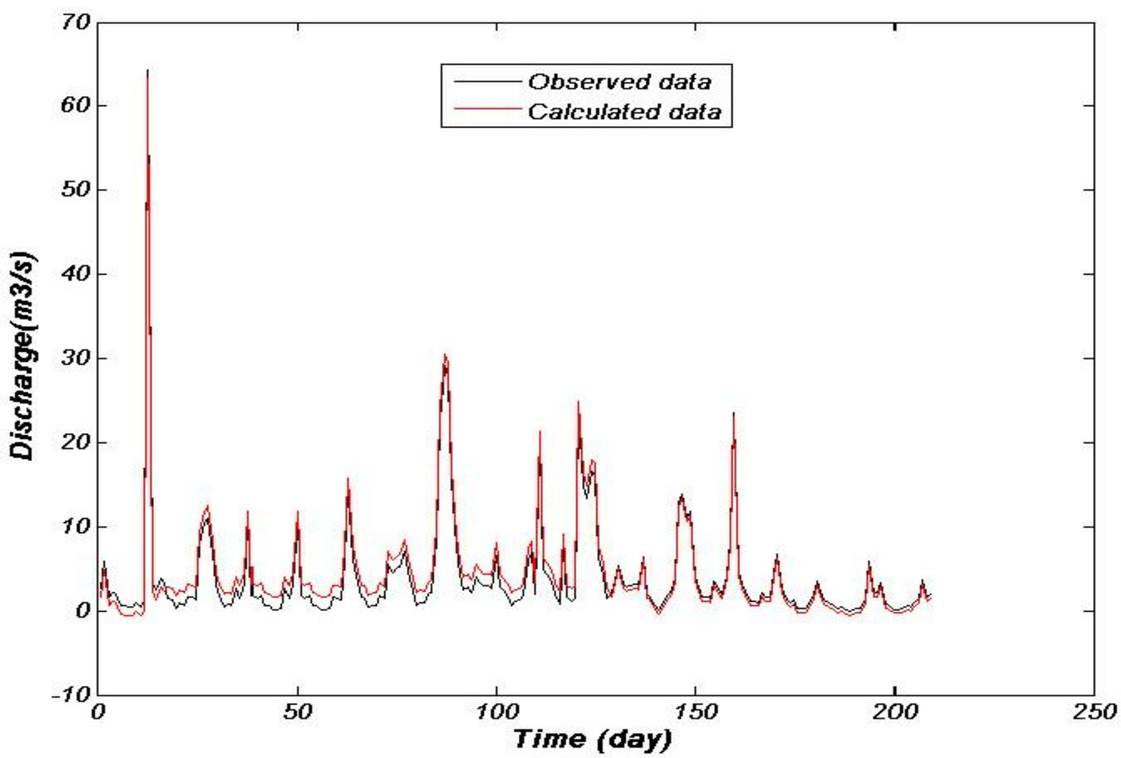

Figure 3-4: Diagram of the comparison between calculated data of the selected model and the observed data of monthly discharge at training stage

© Copyright 2014 | Centre for Info Bio Technology (CIBTech)

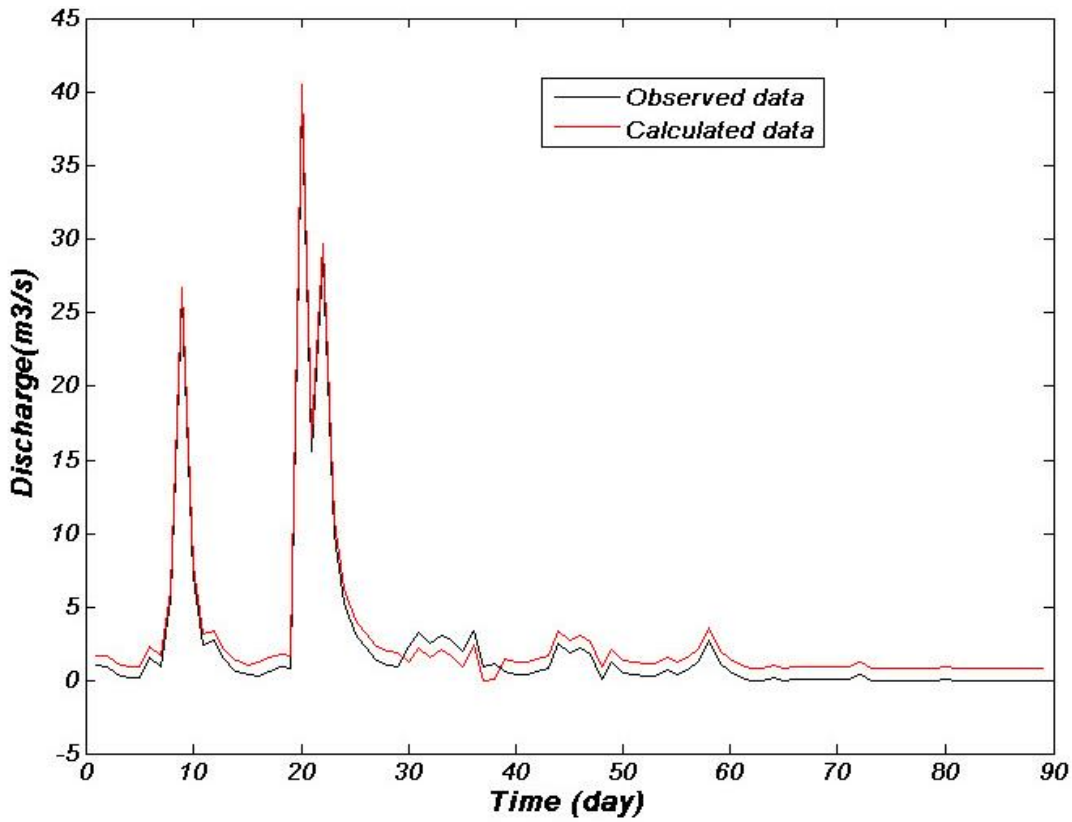

Figure 4-4: Diagram of the comparison between calculated data of the selected model and the observed data of daily discharge at validation stage

REFERENCES

Atiya AM, El-Shora Suzan M, Shaheen Samir I and El-Sherif Mohamed S (1999). A comparison between neural network forecasting techniques- Case study: River flow forecasting. *IEEE Transaction on Nural Network.*

Castellano-Mendez M, Gonzalez- Manteiga W, Febrero-Bande M, Prada-Sanchez JM and Lozano-Calderon R (2004). Modelling of monthly and daily behavior of the run off the Xallas river using Box-Jenkins and Neural networks methods. *Journal of Hydrology* 296.

Cigizoglu HK (2004). Suspended sediment estimation and forecasting using artificial neural networks.

Fereidouni *et al.*, (2011). Used an artificial neural network to model and predict each balance parameter (rainfall, evaporation, groundwater level and runoff) in Mand and Bakhtegan basins (Jahrom, GhareAghach and Gavdar saline rivers) in three regions of different climate in Fars province. *National Conference on Applied Research in Science and Engineering*.

Karamouz M and Tabari MRR (2003). *The First Annual Conference on Water Resources Management, Iran, Tehran*.

Lorrai M and Sechi HM (1995). Neural nets for modeling rainfall-runoff transformations. *Water Resources Management*.

Nayebi (2005). Predicting Short-term River Flow Using Artificial Neural Network, Shiraz University, Master Thesis.

Nazarnezhad H and Ghorbani MA (2005). Application of Artificial Neural Network to Predict the Maximum Daily River Discharge. *The Second National Conference on Watershed and Water Resources Management*.

Zhang B and Govindaruja RS (2000). Prediction of watershed runoff using Bayesian concepts and modular neural networks. *Water Resources Research.*

© *Copyright 2014 | Centre for Info Bio Technology (CIBTech)*