

Evaluation of Psonns Model in Daily Flow Discharge Prediction (Case Study: Babolrood River, Iran)

V. Azimi^{1*}, A. Vakilifard¹, S. Khanmohamadi Fallah², S. Hosseinpour³

¹Department of Water Engineering, University of Tabriz, Tabriz, Iran

²Department of Water Engineering, Imam Khomeini International University, Qazvin, Iran

³Department of Water Engineering, Gorgan University of Agricultural Sciences and Natural Resources, Gorgan, Iran

*Corresponding Author Email: vahidazimi_irrigation@yahoo.com

Abstract

In recent years, artificial neural networks (ANNs) have been successfully used as a tool to model various non-linear relations, and the method is appropriate for modelling the complex nature of hydrological systems. Owing to the complexity of the hydrological process, Particle Swarm Optimization Neural Networks (PSONNs) is the superior model that is able to calibrate the daily flow discharge accurately by using only flow data. Therefore, a new evolutionary algorithm (EA) named particle swarm optimization (PSO) is proposed to train the feedforward neural networks. This particle swarm optimization feedforward Neural Networks (PSONNs) is applied to model the daily flow discharge for Babolrood river in Mazandaran a province in Iran. With the input data of antecedent flow discharge, the optimal configuration of PSONNs is able to simulate current flow discharge successfully with an accuracy of $R^2=0.683$, $MSE=0.0023$ and $MAE=0.0206$ for training and $R^2=0.736$, $MSE=0.0024$ and $MAE=0.0206$ for testing data set. The performance of the newly developed PSONNs demonstrated the success in modeling flow discharge for the Babolrood River.

Keywords: ANNs, PSO, PSONNs, Babolrood river, Daily flow prediction.

Introduction

McCulloch and Pitts (1943) introduced the initial idea about artificial neural networks (ANNs) by proposing the model of a neuron. Later, the ANNs got high recognition due to Rumelhart and McClelland (1986). They discovered the mathematically rigorous theoretical framework for the ANNs by presenting the generalized delta rule, or back-propagation algorithm (BPA) and demonstrated its capability, in training a multilayer ANNs. ANNs have been applied in different areas of hydrology such as flow discharge modeling, groundwater management, stream flow forecasting, precipitation forecasting, hydrologic time series and reservoir operations. Currently there are a few solutions proposed by (ANNs) researcher to overcome this slow convergence rate problem. many studies focused on streamflow predictions have proven that ANNs is superior to traditional regression techniques and time-series models including autoregressive (AR) and autoregressive moving average (ARMA) (Raman & Sunilkumar, 1995; Jain et al., 1999; Thirumalaiah & Deo, 2000; Abrahart & See, 2002; Castellano-Méndez et al., 2004; Kisi, 2003, 2008). Besides, ANN is also compared with non-linear prediction (NLP) method which is derived from the chaotic time series (Farmer & Sidorowich, 1987). Laio et al (2003) carried out a comparison of ANN and NLP for flood predictions and found that ANN performed slightly better at long forecast time while the situation was reversed for shorter time. Sivakumar et al (2002) found that ANN was worse than NLP in short-term river flow prediction.

Rogers and Dowla (1994) used the approach of combining the genetic algorithm (GA) with ANN was predicted the fitness measures of the generated pumping pattern of 20 wells. GA-ANN optimization results were compared with other approaches of combining the groundwater flow model with a non-linear programming with a quasi-Newton search method. The counter propagation network (CPN) and the generalized regression neural network (GRNN) algorithms have been used in a small number of studies to forecast streamflow (Aytek et al., 2008; Chang & Chen, 2001; Chang et al., 2001). The GRNN was found to outperform FFBP ANN methods during daily prediction (Cigizoglu, 2005a) and monthly streamflow (Cigizoglu, 2005b; Kisi, 2008). Kumar et al (2005) forecasted monthly river flows by using two different networks, named the feed forward network and the recurrent neural network. Hence recurrent neural network are recommend as a tool for river flow forecasting.

Mohammadi et al (2005) presented three different methods, artificial neural network (ANN), ARIMA time series and regression analysis between some hydroclimatological data and inflow. Akhtar et al (2009) explored the use of flow length and travel time as a pre-processing step for incorporating spatial precipitation information into Artificial Neural Network (ANN) models used for river flow

forecasting. Srinivasulu and Jain (2009) presented an integrated approach for river flow prediction in an attempt to achieve better forecast accuracy. The integrated model uses conceptual, ANN, genetic algorithm, data-decomposition, and model fusion techniques. The ANN is able to capture the dynamics of the flow series by using previous observed flow values as inputs during forecasting the daily flows from the flow data alone. De Vos and Rientjes (2005) suggested that an effective solution to forecast lag effect is to obtain new model inputs by moving average (MA) over the original discharge data. ANN consists of simple processing units in a parallel sequence. The connections between these units will specify the network function and performance. For doing of some tasks or functions ANN can be trained by adjusting the connection strength (weights) between units (Beal et al., 2010).

Artigue et al (2012) used ANN for flash flood forecasting on ungauged basins in southern France. Study demonstrates that efficient forecasting can be derived from a feedforward model using the available measured discharges. This paper examines the capability of the PSONNs based on simulation-optimization model where the training dataset for the ANNs model were generated by the PSO based on daily flow discharge model. This PSONNs is proposed to improve the convergence rate of NNs and avoid solutions being trapped at local minima.

Methodology

Study area

Babolrood River originates in the Alborz mountains and is one of the major rivers in Iran. The minimum and maximum of altitude is 100-800m respectively. According to classifying continental Domartan, this basin is very humid. This district is situated in 52° 41' 40" to 52° 44' 55" east longitude and 36° 16' 28" to 36° 19' 15" north latitude. As it is obvious, approximately the soil type in mentioned borehole and in most part of study area is clay. It's watershed is 1630 km² with 78 km length, and about 50–60 m width at the mouth to 100 m in the upstream. The average discharge of it, is 16 m³/s. The river receives abundant water from snowmelt and rainfall, particularly in winter and when substantial precipitation has fallen in the mountains, which makes it full-flowing river (Farrokhzad et al., 2012) (Figure 1).



Figure 1. Location map of the study area.

Streamflow data

Daily mean flow data from Babolrood River of Mazandaran province are used in this study. The flow data from Babolrood River were acquired at Babol hydrology station which is at the upper stream of Babolrood watershed. The data period covers a 13 years long duration (Mar. 1, 1997- Mar. 1, 2010). In the process of PSONNs modeling, the raw flow data is often partitioned into two parts training set and testing set. The training set serves the model training and the testing set is used to evaluate the performances of models. The same data partition was adopted in two daily flow series: the first 80% of the entire flow data as training and the remaining data as testing set.

Table 1 presents related information about some descriptive statistics of the original data and two data subsets, including mean (μ), standard deviation (S_x), coefficient of variation (C_v), skewness coefficient (C_s), minimum (X_{min}), and maximum (X_{max}). As shown in Table I, the training set cannot fully include the testing set.

Table 1. Related information for flow data of Babolrood River.

Data sets	No. of data	Data Type	Statistical parameters					
			μ	S_x	C_v	C_s	X_{min}	X_{max}
Whole	4744	Flow (m3/s)	15.72	22.24	1.41	5.07	0.006	332
Training	3795	Flow (m3/s)	16.16	23.62	1.46	5	0.006	332
Testing	949	Flow (m3/s)	13.9	15.41	1.11	3.68	0.008	151

Particle Swarm Optimization Neural Networks (PSONNs)

The PSO is made up of particles, where each particle has a specific position and velocity. The purpose of PSO in NNs is to get the best set of weights (or particle positions) where several particles (problem solution) are trying to achieve the best solution and this will avoid the solution trap at local minima (Eberhart & Shi, 2001). The advantage of the PSO over many of the other optimization algorithms is its relative simplicity (Van den Bergh, 2001). According to Jones (2005), the only two equations used in PSO are the movement (Equation 1) and the velocity update equation (Equation 2). The movement equation provides the actual movement of the particles using their specific vector velocity while the velocity update equation provides velocity vector adjustment given the two competing forces (gbest and pbest). The inertia weight (ω) was introduced by Eberhart and Shi (1998) to improve the convergence rate of the PSO algorithm.

$$presLocation = prevLocation + Vi\Delta t \tag{1}$$

$$Vi = \omega Vi-1 + c1*rund()*(pbest - presLocation) + c2 *rund()*(gbest - presLocation) \tag{2}$$

where Vi is the current velocity, Δt defines the discrete time interval over which the particle will move, ω is the inertia weight, $Vi-1$ is the previous velocity, $presLocation$ is the present location of the particle, $prevLocation$ is the previous location of the particle and $rand()$ is a random number between (0, 1), $c1$ and $c2$ are the acceleration constants for “gbest” and “pbest” respectively. Particle velocities are limited by a user-specified value, maximum velocity V_{max} , to prevent the particles from moving too far from a potential solution. The hybrid of PSO and NNs is applying PSO to train the feedforward NNs for enhancing the convergence rate and learning process. According to Al-kazemi and Mohan (2002), the position of each particle in a PSONNs represents a set of weights for the current iteration. The dimension of each particle is the number of weights associated with the network. The learning error of this network is computed using the Mean Squared Error (MSE) between the observed and simulated runoff. The learning process involves finding a set of weights that minimizes the learning error. Hence, the particle will move within the weight space attempting to minimize learning error.

The learning process of PSONNs is further illustrated in figure 2. The learning process of PSONNs is initialized with a group of random particles (step 1), which are assigned randomly PSO positions (weight and bias). The PSONNs is trained using the initial particles position (step 2). Then, the feedforward NNs in PSONNs will produce the learning error (particle fitness) based on an initial weight and bias (step 3). The training error at the current epoch or iteration will be reduced by changing the particles position, which will update the weight and bias of the network. The pbest value (each particle’s lowest learning error so far) and gbest value (lowest learning error found in the entire learning process so far) are applied to use in the velocity update equation (Equation 2) to produce a value for position adjustment for achieving the best solutions or targeted learning error (step 4). The new sets of positions (NNs weight and bias) are produced by adding the calculated velocity value to the current position value using the movement equation (Equation 1).

Then, the new sets of positions are used to produce new learning errors for the feedforward NNs (step 5). This process must be repeated until the stopping conditions, either minimum learning error or maximum numbers of iterations are met (step 6). The optimization output, which is the solution for the optimization problem, is based on the gbest position value.

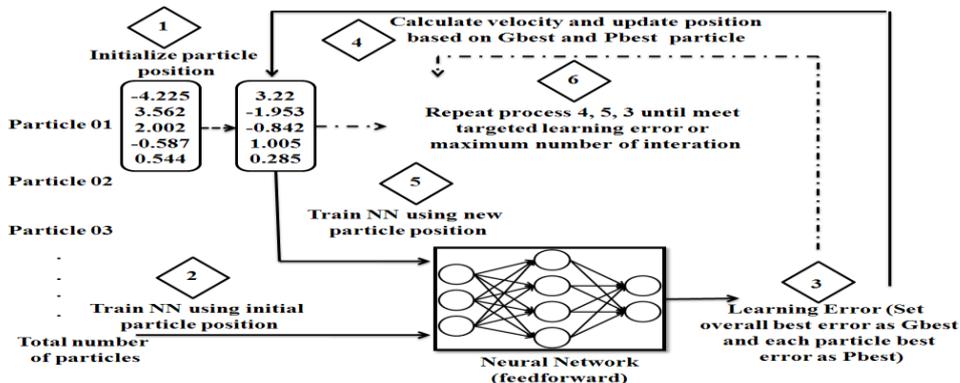


Figure 2. PSONNs learning process.

In this study, PSONNs program has been developed based on PSO program for Sombrero function optimization. The particle position represents two-dimensional (2D) vector of x and y values in Sombrero function. The objective is to reach the value of 1 based on value of x and y in Sombrero equation (Equation 3) with the goal of maximizing the function.

$$z = 6 * \cos\left(\frac{\sqrt{x * x + y * y}}{x * x + y * y + 6}\right) \tag{3}$$

where x is the value in x-axis, y is the value in y-axis and z is the value in z-axis. Figure 3 shows that 5 particles move or fly to the solution in 2D problem to solve the Sombrero function. The fitness of the particle on the Sombrero is represented by z-axis.

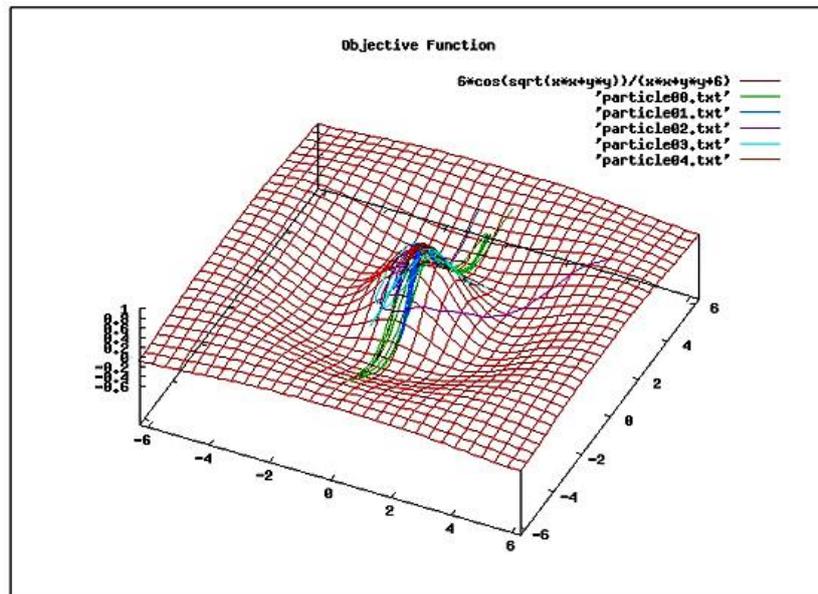


Figure 3. Particle movement in sombrero function optimization.

Model development

Various models were calibrated to determine to optimal configuration of PSONNs models. PSONNs architecture with well-selected parameter set can have good performance (Shi, 2004). Hence, 7 parameters were investigated named:

- a) Acceleration constants for gbest (c1)
- b) Acceleration constants for pbest (c2)
- c) The time interval (Δt)
- d) The number of particle (D).
- e) Number of maximum iteration (stopping condition)
- f) Number of hidden neurons
- g) Length of training and testing data
- h) Number of antecedent day.

Four basic parameters in PSO namely acceleration constants for gbest (c1), acceleration constants for pbest (c2), time interval (Δt), number of particles (D) are affecting the optimal configuration of PSONNs (Jones, 2005). Besides, other parameters including maximum iteration (stopping condition), number of hidden neurons in the hidden layer, length of input data and number of antecedent hours are also investigated for model calibration and optimization. Meanwhile, parameters such as particle dimension also affect the optimization results. Number of dimension in PSONNs is referring to number of weight and bias that is based on the dataset and PSONNs architecture. PSONNs dimension is calculated using Eq. 4:

$$\text{Dimension} = (\text{input} * \text{hidden input}) + (\text{hidden} * \text{output hidden}) + \text{hidden bias} + \text{output bias} \tag{4}$$

5 scenarios were developed to investigate the effect of daily antecedent to the configuration of PSONNs. The input data of the models consists of antecedent flow, Q (t-1), Q (t-2),...,Q (t-n). Where else, the output is the flow discharge at the current day, Q (t).

Where t=time in day, and Q=flow discharge in m3/s. These 5 scenarios are shown in table 2. As time is one of the important factors in the model, the inputs shall be arranged in sequent.

Table 2. Combinations of PSONNs model scenarios.

Scenario	Input combinations	Output
1	$Q_{(t-1)}$	Q_t
2	$Q_{(t-1)}, Q_{(t-2)}$	Q_t
3	$Q_{(t-1)}, Q_{(t-2)}, Q_{(t-3)}$	Q_t
4	$Q_{(t-1)}, Q_{(t-2)}, Q_{(t-3)}, Q_{(t-4)}$	Q_t
5	$Q_{(t-1)}, Q_{(t-2)}, Q_{(t-3)}, Q_{(t-4)}, Q_{(t-5)}$	Q_t

Data normalization

One advantage of data normalization is to avoid attributes in greater numeric ranges dominating those in smaller numeric ranges, and another advantage is to avoid numerical difficulties during the calculation. It is recommended to linearly scale each attribute to the range of [0.1, 0.9], [- 1, +1] or [0, 1]. In the modeling process, the data sets of maximum and minimum temperature, streamflow, and suspended sediment load were scaled to the range between 0 and 1 as follow:

$$N_i = \frac{x_i - x_{\min}}{x_{\max} - x_{\min}} \tag{5}$$

Where N_i is the normalized value, x_i is the original data and x_{\min} , x_{\max} are, respectively, the minimum and maximum of the original data.

Error functions

Measures of goodness of fit between observed and predicted datasets are based on the coefficient of determination, R^2 , as well as the three measures defined below:

-Mean squared error (MSE):

$$MSE = \frac{\sum_{i=1}^t (P_i - O_i)^2}{t} \tag{6}$$

-Mean absolute error (MAE):

$$MAE = \frac{\sum_{i=1}^t |P_i - O_i|}{t} \tag{7}$$

Where P_i and O_i are the simulated and observed values respectively, and t the sample size. R^2 varies between 0 and 1; the closer the values to 1, the better the goodness of fit is. The goodness of fit improves as the MSE approaches 0, and as the MAE decreases.

Results and discussion

Learning mechanism

Three layers PSONNs were used in this study to calibrate the model. Number of input neurons depends on the model used (number of antecedent data). There is only one neuron in the output layer. PSONNs were trained and tested to find the best configuration using various dataset as shown below:

- a) 1.4 to 2.3 of c_1 and c_2 values.
- b) Δt of 0.005, 0.010, 0.015, 0.020 and 0.025
- c) 16, 17, 18, 19, 20, 21 and 22 numbers of particles.
- d) 50, 75, 100, 125 and 150 number of hidden neurons in hidden layer.
- e) 400, 450 and 500 number of maximum iteration.
- f) 117 months of training data (3795 day) and 39 months of testing data (949 day) constantly.
- g) 1, 2, 3, 4 and 5 number of antecedent data.

Discussions in PSONNs model

The optimal model required 100 numbers of hidden neurons. The best maximum iteration obtained, is 450. Length of training and testing data in scenarios is constant. The effect of number of antecedent days is to determine the best time series to produce the best network. Scenario 4 performs the best with the highest coefficient of correlation compared with other PSONNs scenarios as shown in Table 3 and Figure 4. The result showed that Scenario 4 is the optimal model for simulating daily flow discharge in Babolrood River.

Table 3. Combinations of PSONNs model scenarios.

Scenario	Training			Testing		
	R ²	MSE	MAE	R ²	MSE	MAE
1	0.686	0.0024	0.021	0.630	0.0025	0.020
2	0.682	0.0025	0.019	0.741	0.0029	0.021
3	0.577	0.0032	0.026	0.501	0.0043	0.027
4	0.683	0.0023	0.021	0.736	0.0024	0.021
5	0.686	0.0024	0.020	0.689	0.0024	0.023

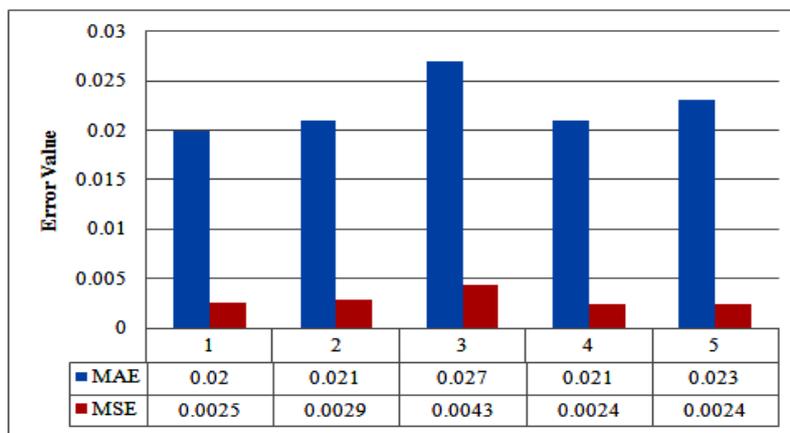


Figure 4. MAE and MAE error values in scenarios.

The acceleration constant c_1 and c_2 represent the stochastic acceleration that pulls each particle toward “pbest” and “gbest” position (Eberhart and Shi, 2001). The c_1 constant affects the influence of the global best solution over the particle, whereas the c_2 constant affects how much influence of personal best solution has over the particle. The best c_1 and c_2 values required for optimal configuration of scenario 4 is 1.8. The Δt parameters define the time interval over which movement takes place in the solution space. Decreasing these parameters provides higher granularity movement within the solution space and higher Δt value, performs lower granularity movement (greater distance achieved in less time). It was concluded the best Δt obtained in this study is 0.020. In conclusion the number of particles in PSO determines the amount of space that is covered in the problem. Generally, this study indicates that the best number of particles is 19. The optimization period was getting longer by increasing number of particles. This is because the more particles presented, the greater amount of space is covered in the problem. The optimal PSONNs model determined above will be used to generate daily flow discharge. The results of training and testing step (first 200 day) in 200 first step of this study are presented in Figure 5 to 14 for all scenarios.

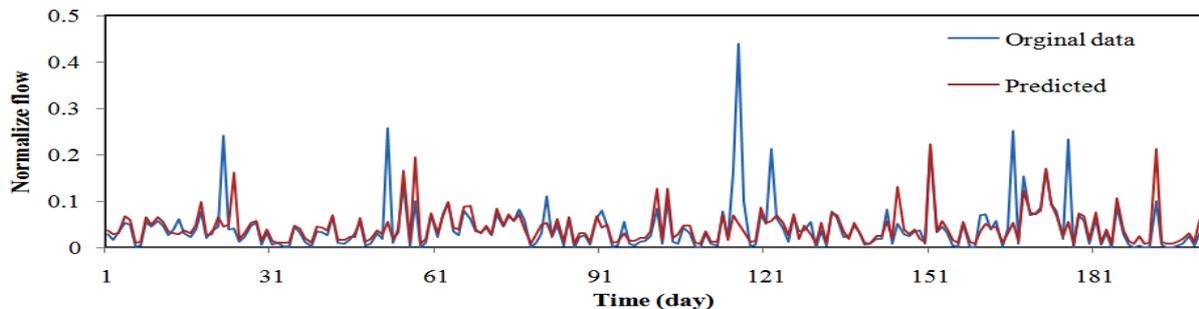


Figure 5. Performance of PSONNs for training data set in scenario 1.

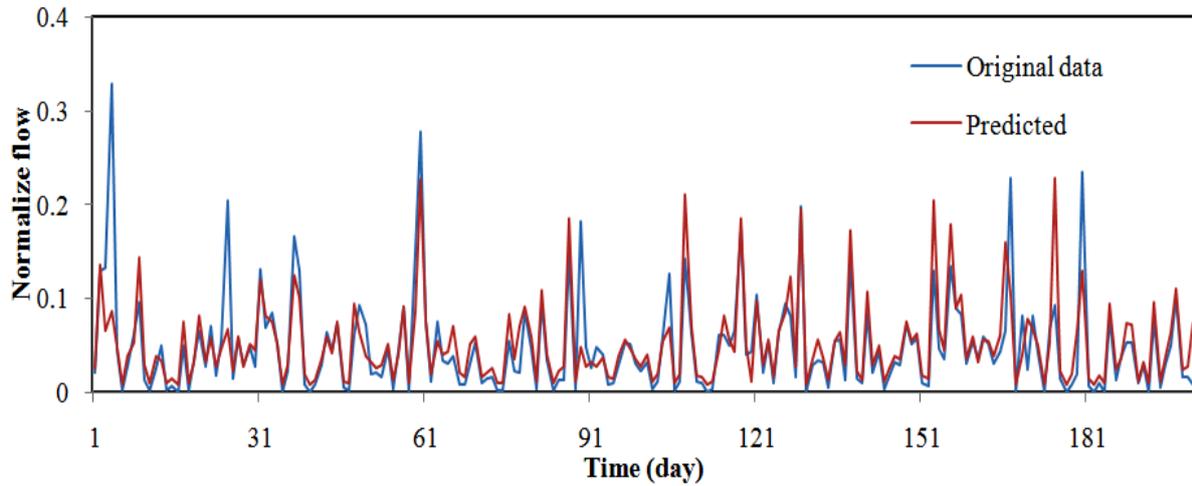


Figure 6. Performance of PSONNs for testing data set in scenario 1.

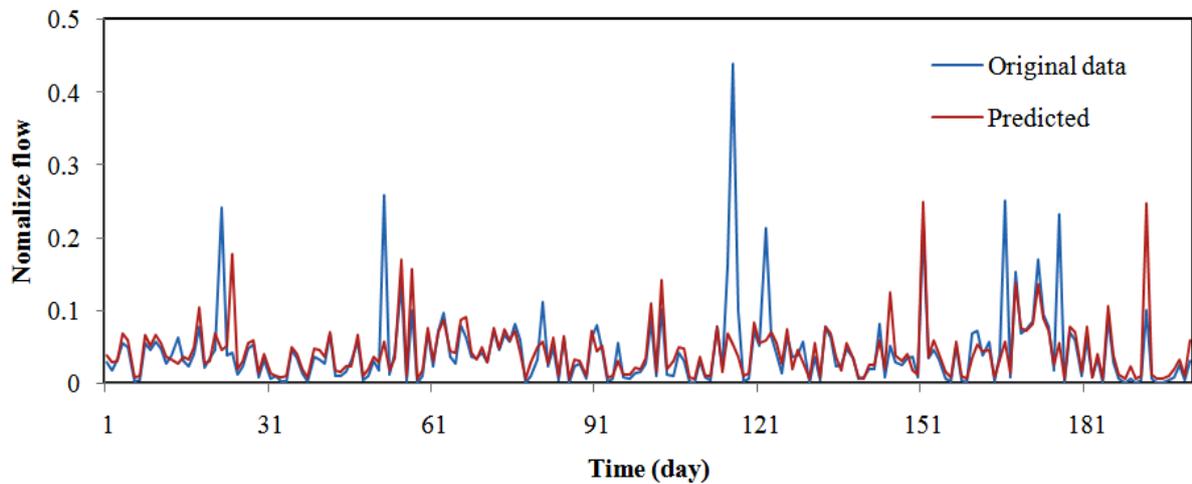


Figure 7. Performance of PSONNs for training data set in scenario 2.

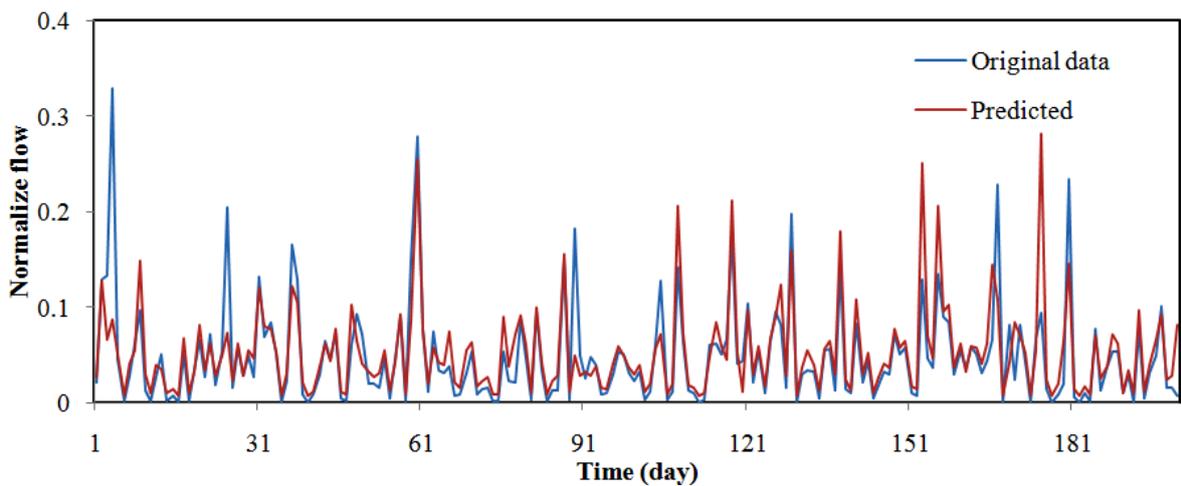


Figure 8. Performance of PSONNs for testing data set in scenario 2.

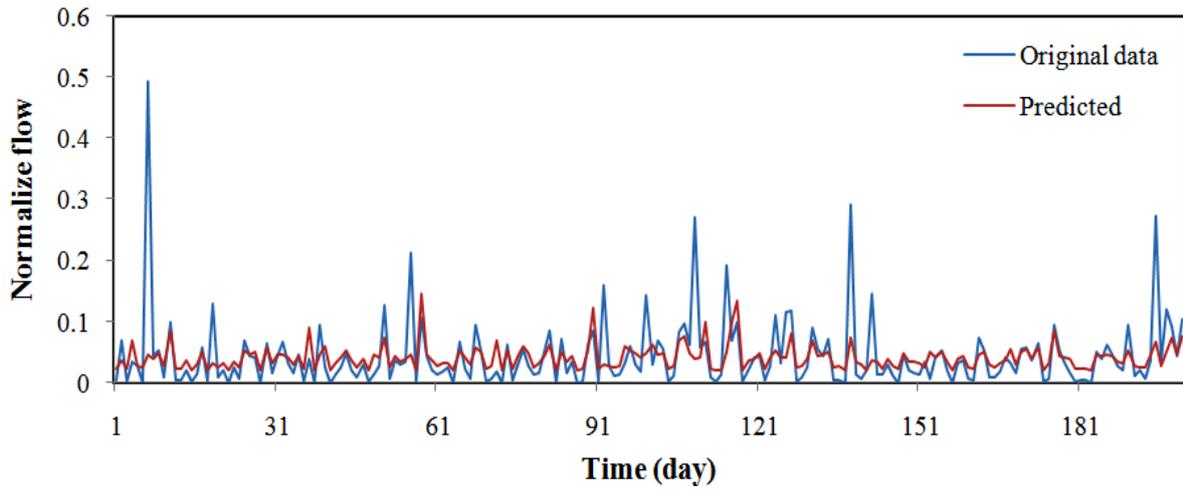


Figure 9. Performance of PSONNs for training data set in scenario 3.

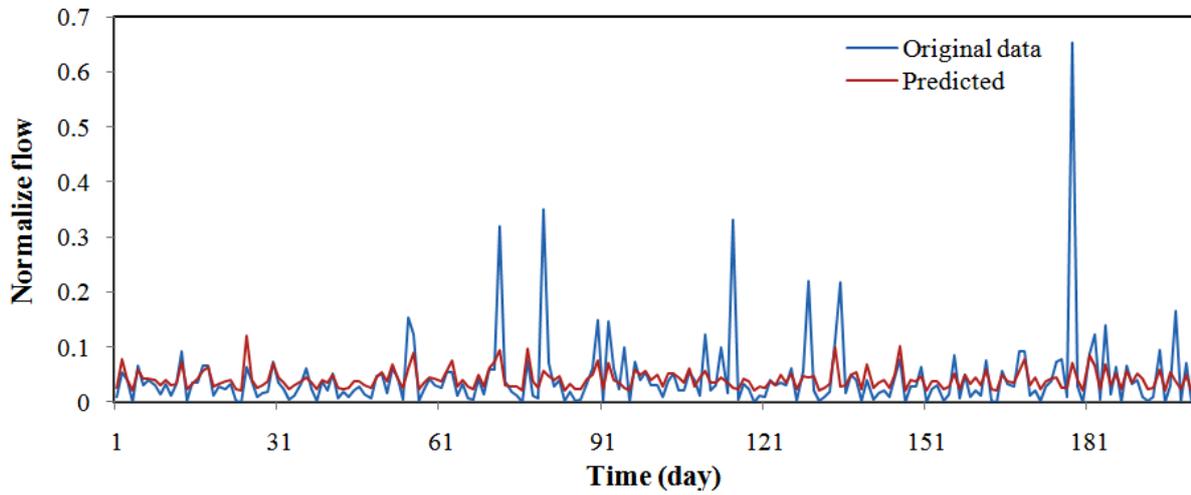


Figure 10. Performance of PSONNs for testing data set in scenario 3.

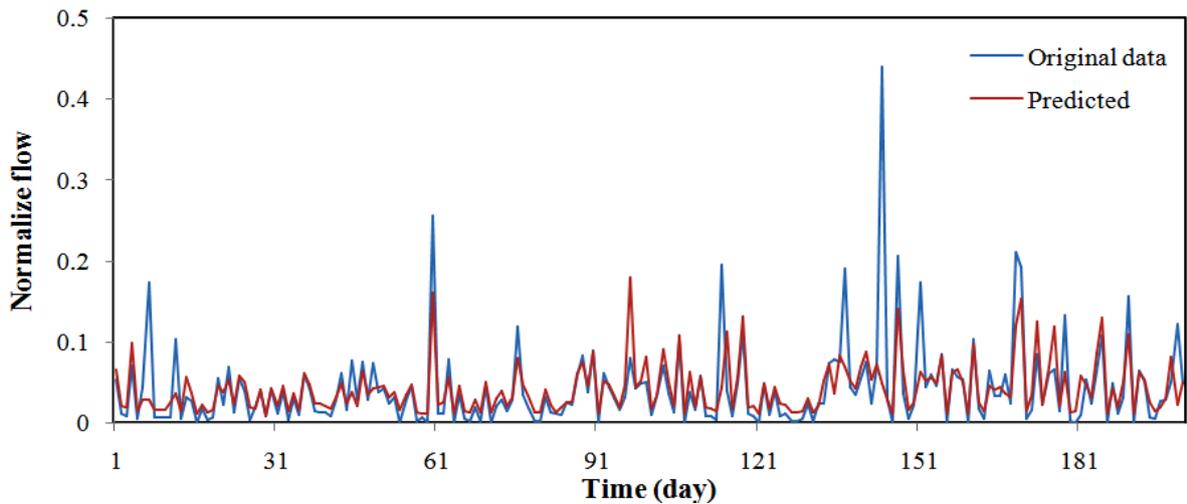


Figure 11. Performance of PSONNs for training data set in scenario 4.

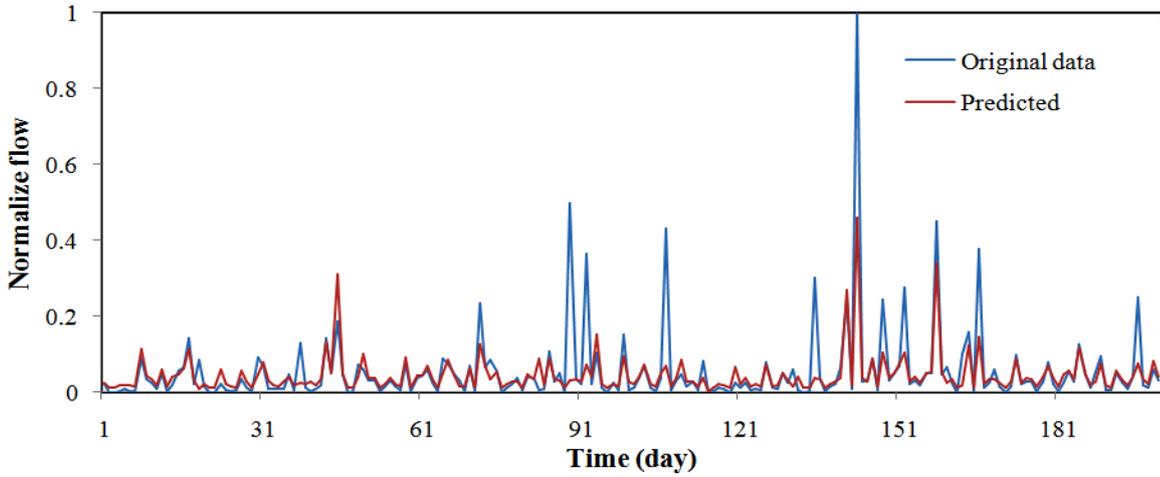


Figure 12. Performance of PSONNs for testing data set in scenario 4.

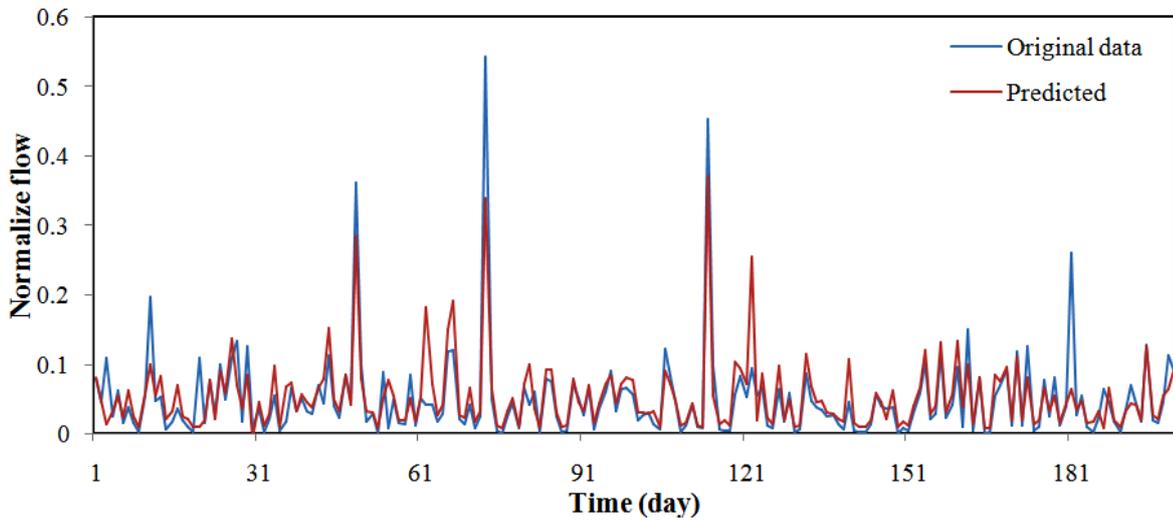


Figure 13. Performance of PSONNs for training data set in scenario 5.

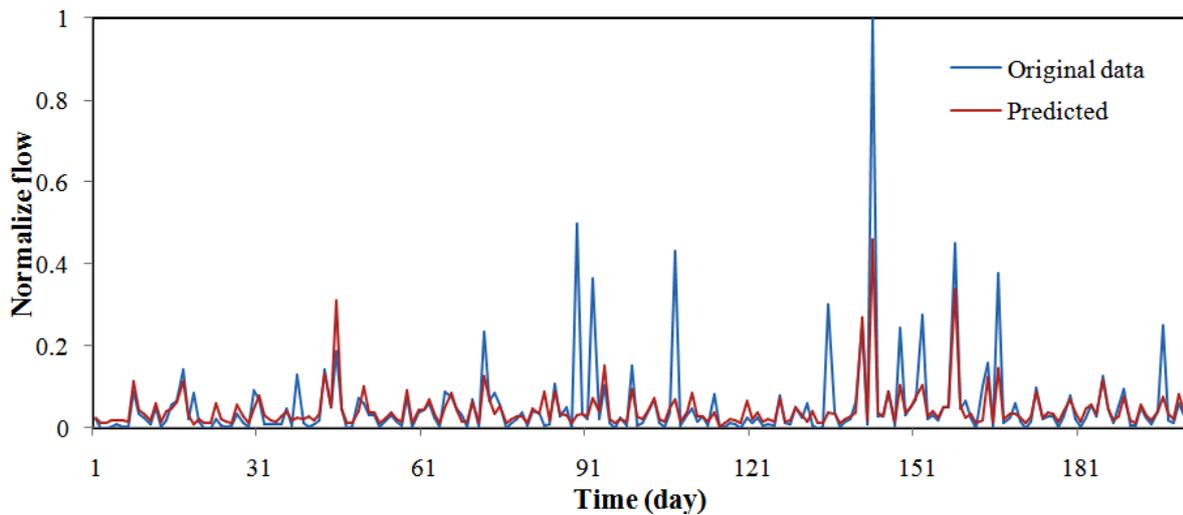


Figure 14. Performance of PSONNs for testing data set in scenario 5.

Conclusion

PSO is successfully developed for daily flow discharge simulation for Babolrood River. The result shows that this PSO based on algorithm can train the NNs as other existing techniques. The Statistic indicators to help to evaluate performance of different models are R^2 , MSE and MAE. According to value of R^2 , MSE and MAE the comparative advantage of models is obviously understandable. The best PSO models in this model for daily flow discharge prediction is scenario 4 by yielding $R^2=0.683$, $MSE=0.0023$ and $MAE=0.0206$ for training and $R^2=0.736$, $MSE=0.0024$ and $MAE=0.0206$ for testing data set. This models with best configuration to model daily flow discharge is using 100 number of hidden neurons, 450 number of max iteration, 117 months of input data for training and 39 months of testing data (constant), c1 and c2 values of 1.8, 4 antecedent day, 19 number of particles and time interval of 0.020. It was also found that input data including antecedent flow was sufficient for PSO models to predict current flow accurately. Besides, Scenario 5, Scenario 1, Scenario 2 and Scenario 3 are the next priorities, respectively. The last conclusion introduces designed PSO models to estimate the Babolrood River daily flow discharge using multiple inputs.

References

- Abrahart RJ, See L, 2002. Multi-model data fusion for river flow forecasting: an evaluation of six alternative methods based on two contrasting catchments. *Hydrology and Earth System Sciences*. 6(4): 655-670.
- Akhtar MK, Corzo GA, Andel SJ, Jonoski A, 2009. River flow forecasting with artificial neural network using satellite observed precipitation pre-processed with flow length and travel time information: case study of the Ganges river basin. *Journal of Hydrology and Earth System Sciences*. 6: 3385-3416.
- Al-kazemi B, Mohan CK, 2002. Training feedforward neural network using multi-phase particle swarm optimization. *Proceeding of the 9th International Conference on Neural Information Processing*, New York.
- Artigue G, Johannet A, Borrell V, Pistre S, 2012. Flash flood forecasting in poorly gauged basins using neural networks: case study of the Gardon deMialet basin (southern France). *Nat Hazards Earth Syst Sci*. 12: 3307-3324.
- Aytek A, Asce M, Alp M, 2008. An application of artificial intelligence for rainfall-runoff modeling. *Journal of Earth System Science*. 117(2): 145-155.
- Beal MH, Hagan MT, Demuth HB, 2010. *Neural Network Toolbox™ 7*.
- Castellano-Méndez M, Gonzalez-Manteigaa W, Febrero-Bande M, Manuel Prada-Sancheza J, Lozano Calderon R, 2004. Modeling of the monthly and daily behavior of the runoff of the Xallas river using Box-Jenkins and neural networks methods. *Journal of Hydrology*. 296: 38-58.
- Chang FJ, Chen YC, 2001. A counter propagation fuzzy-neural network modeling approach to real time stream flow prediction. *Journal of Hydrology*. 245: 153-164.
- Chang FJ, Hu HF, Chen YC, 2001. Counter propagation fuzzy-neural network for streamflow reconstruction. *Hydrological Processes*. 15(2): 219-232.
- Cigizoglu HK, 2005a. Application of generalized regression neural networks to intermittent flow forecasting and estimation. *Journal of Hydrologic Engineering*. 10(4): 336-341.
- Cigizoglu HK, 2005b. Generalized regression neural network in monthly flow forecasting. *Civil Engineering and Environmental Systems*. 22(2): 71-84.
- De Vos NJ, Rientjes THM, 2005. Constraints of artificial neural networks for rainfall-runoff modeling: trade-offs in hydrological state representation and model evaluation. *Hydrology and Earth System Sciences*. 9: 111-126.
- Eberhart R, Shi Y, 2001. Particle swarm optimization: developments, applications and resources. *IEEE Int Conf on Evolutionary Computation*. 2001: 81-86.
- Farmer JD, Sidorowich JJ, 1987. Predicting chaotic time series. *Phys Rev Lett*. 59(4): 845-848.
- Farrokhzad F, Choobbasti A J, Barari A, 2012. Liquefaction microzonation of Babol city using artificial neural network. *Journal of King Saud University-Science*. 24(1): 89-100.
- Jain SK, Das A, Drivastava DK, 1999. Application of ANN for reservoir inflow prediction and operation. *J Water Resour Plann Manage*. 125(5): 263-271.
- Jones MT, 2005. *AI application programming*. 2nd edn. Hingham: Massachusetts.
- Kisi O, 2003. River flow modeling using artificial neural networks. *Journal of Hydrologic Engineering*. 9(1): 60-63.
- Kisi O, 2008. River flow forecasting and estimation using different artificial neural network techniques. *Hydrology Research*. 39(1): 27-40.
- Kumar DN, Srinivasa Raju K, Sathish T, 2005. River flow forecasting using recurrent neural networks. *Journal of Water Resources Management*. 18: 143-161.
- Laio F, Porporato A, Revelli R, Ridolfi L, 2003. A comparison of nonlinear flood forecasting methods. *Water Resour Res*. 39(5): 1129.
- McCulloch WS, Pitts W, 1943. A logical calculus of the ideas imminent in nervous activity. *Bull Math Biophys*. 5: 115-133.
- Mohammadi K, Eslami HR, Dayyani Dardashti Sh, 2005. Comparison of regression ARIMA and ANN models for reservoir inflow forecasting using snowmelt equivalent (a case study of Karaj). *Journal of Agricultural Science and Technology*. 17: 17-30.

- Raman H, Sunilkumar N, 1995. Multivariate modeling of water resources time series using artificial neural networks. *Hydrological Sciences Journal*. 40(2): 145-163.
- Rogers LL, Dowla FU, 1994. Optimization of groundwater remediation using artificial neural networks with parallel solute transport modeling. *Water Resour Res*. 30(2): 457-481.
- Shi Y, 2004. Particle swarm optimization. *IEEE Neural Network Society*. 8-13.
- Sivakumar B, Jayawardena AW, Fernando TMK, 2002. River flow forecasting: use of phase-space reconstruction and artificial neural networks approaches. *Journal of Hydrology*. 265(1): 225-245.
- Srinivasulu S, Jain A, 2009. River flow prediction using an integrated approach. *Journal of Hydrologic Engineering, ASCE*. 14(1): 75-83.
- Thirumalaiah K, Deo MC, 2000. Hydrological forecasting using neural networks. *Journal of Hydrologic Engineering*. 5(2): 180-189.
- Van den Bergh M, 2001. An analysis of particle swarm optimizers. PhD Thesis, University of Pretoria.