PREDICTING WEEKLY DISCHARGE USING ARTIFICIAL NEURAL NETWORK (ANN) OPTIMIZED BY ARTIFICIAL BEE COLONY (ABC) ALGORITHM: A CASE STUDY

[1]Mahmood Rahsepar and [2]Hamid Mahmoodi

[1]Departement of Civil Engineering, Islamic Azad University of Darab, Iran
[2]Department of Computer Science and Engineering, Shiraz University, Shiraz, Iran

ABSTRACT

Water surface management in the field if hydrology has been considered as an important issue due to many recently drought years specially in warm and dry regions like south of iran. This paper is conducted to propose ANN-ABC mixture algorithm in order to forecast the future discharge of Tang-e Karzin hydrometric station located in sub domain of salman farsi dam. A Feed Forward Neural Network (FFNN) was utilized to forecast the future discharge of a case study station using the 36 past years discharge information. Moreover, Artificial Bee Colony (ABC) optimization algorithm was applied within the training phase of the ANN network to optimize the weights of the MLP network. Through substitution of parameters of hidden layer (kind of activation function and number of neurons) of the neural network, best combination of parameters found in the first phase of the algorithm. Next in the second phase, artificial bee colony algorithm was used to find the global solution of the objective function of the problem. Simulation results indicated that ABC algorithm could significantly improve the neural network results.

KEYWORDS

Multi-Layer Perceptron(MLP) Algorithm, Artificial Bee Colony(ABC) Optimization technique, River Discharge Prediction, Neural Network Weight Optimization.

1.INTRODUCTION

1.1.Descriptions and previous works

Many hydrological engineers eager to innovate precise relationships between rainfall and runoff system because of its importance and impacts on many people life. Such relations are essential in planning and managing in exploitation of water resources. Before forecasting models come in to use, physic and process oriented models were common methods to measure the several features. Main problem of physics based models is that the ways in which measure the spatial distribution of most variables such as precipitation are not practical as much as forecasting models in many regions of the world [1]. In spite of need to great amount of data and human effort to calibrate, validate and test the model, physics based methods are useful to understand the entire of underlying process [2]. In the other hand, good advantage of forecasting models is that they just need limited amount of data, but their drawback is that they seem like a black box in usages and also need lengthy parameterization. Multi Linear Regression (MLR), Fuzzy Logic Based System and Artificial Neural Networks are more commonly used of forecasting methods [3].

Neural networks (NN) have been well suited and successfully applied in variety fields of hydrology including water resources [4, 5, 6, 7, 8]. In hydrological context, artificial neural network has been shown as a sufficient alternative in rainfall-runoff models [4, 7]. A good advantage of ANNs over traditional methods is that they don't need to explicitly characterize the complex mathematical form of underlying process [9]. In the condition of having enough accurate data, ANN has shown satisfaction results in most reports [10]. This demonstrates that ANN model can well cope with complex mapping of nonlinear inputs to their corresponding outputs [11, 12, 13, 14, 15]. Emiroglu et al. applied ANN model to estimate the discharge coefficient of triangular labyrinth side-weir [11]. They investigated several structure of ANN model in terms of number of hidden layer's neurons and found the best parameter of the model. They compared the yield optimum model of ANN with the nonlinear regression model and concluded that ANN can act better than the regression models. Demirel et al. applied the soil and water assessment tool (SWAT) together with the ANN to evaluate the daily stream flow. They tested their model on daily flow of the Pacana basin in Portugal and showed that ANN can well predict the flow much better than the mathematical base methods [16]. Piotrowski et al. proposed ANN structure for rainfall-runoff prediction in the Annapolis river [17]. They trained the ANN using Levenberg–Marquardt learning algorithm and eight evolutionary computation methods. By varying the number of input neurons and also hidden nodes as variable parameters of the ANN, They finally found out that having high quality of data, smaller structures of ANN perform better in rainfall-runoff prediction models. Nguyen et al. applied Back-Propagation Neural Network (BPANN) model to predict the river discharge in Thailand and reached the quite reasonable overall performance [18]. Mrutyunjaya et al. also used BPANN to predict discharge in a compound open channel flow [19]. Riahi-Madvar et al. studied on the efficiency of ANNs to simulate alluvial regime channels [20].

In spite of many good news mentioned about the ANN, danger of getting stuck into local minima is an undeniable problem of ANN from which makes researchers to be concern about their results. Recently Swarm optimization algorithms have been recommended as one of best solution for this issue. Kisi in a suspended sediment estimation study, proposed a differential evolution model instead of back propagation algorithm to learn the ANN [21]. He utilized artificial bee colony (ABC) algorithm along with an ANN model to modeling the daily suspended sediment concentration. Kisi in one of his next studies [22], proposed a new ANN-ABC mixed algorithm for modeling daily discharge-suspended sediment relationship. He compared the yield results of the proposed model with the accuracy of his previous studies such as neural differential evolution [21], adaptive neuro fuzzy [23], neural networks [24] and sediment rating curve [25]. The results showed that generally ANN-ABC model performs better than the mentioned methods. He also used logarithm transformed data as input to the ANN-ABC model and found out that this manipulation in input data increase accuracy of the model. Ebrahimi et al. used the honey-bees mating optimization (HBMO) algorithm to optimize the parameters which governed on sediment concentration and discharge of the river. Results showed that the heuristic approaches like HBMO can be chosen as a superior alternative to regression techniques [26].

ABC is a sort of strong nature inspired and population based search algorithm which is recently proposed by Karaboga in 2005 [27]. One of ABC specifications is in which both exploration and exploitation are well balanced. This is the key point of ABC that why it doesn't get stuck into local minima and why it is used in this study. Applying a swarm optimization algorithm like ABC along with ANN makes a powerful and flex model which improve the global convergence of the model and eliminates the concerning about the results don't come from a local minima state.

Although some kind of application of ABC in few hydrological fields is proposed, there wasn't any published study indicating the usage of ABC in discharge estimation model.

In this paper and in the first phase multilayer perceptron (MLP) neural network was applied to predict the average weekly discharge rate of the underlying hydrometric station (*Tang-e Karzin*). For this purpose the underlying station and either it's two upside hydrometric station (*Band-e Bahman* and *Ali Abad*) discharge data was used. After finding the best parameters (kind of activation function in hidden layer and number of hidden layer neurons) of the ANN model, in the second phase of the algorithm along with the training of the network in each iteration, artificial bee colony algorithm was utilized to optimize the weights of the network. Also, the parameters of the hidden layer such as number of neurons and kind of activation function were alternatively changed to evaluate the different results. Finally, the optimum model was chosen as best prediction model.

1.2. Artificial Neural Network

Artificial neural networks are known as the sort of dynamic, complex and intelligent systems. The importance of the neural networks is in simulation of processes in which usually there isn't clear and comprehension definition for them. Artificial neural networks has been applied in prediction tool of various field of science specially in hydrology science and water management affaires [4-20,22-25,33]. MLP neural network is considered as the class of feed-forward back-propagation network which includes of several layers of computational units. Each units is a representation of neuron in which consists of a linear or nonlinear activation function. The structure of the network is a directed and multiple layer graph and neurons in each layer is fully connected to neurons of the next layer. General structure of a feed-forward neural network known as MLP is depicted in Figure 1.

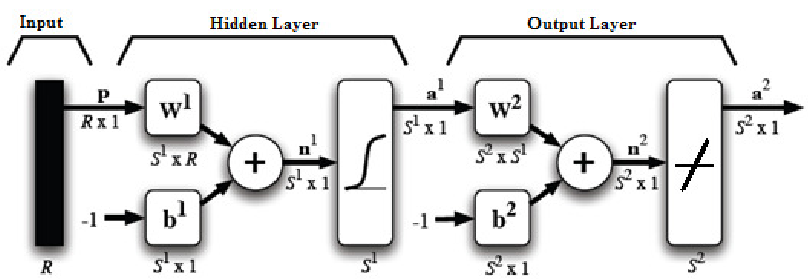


Figure 1. General structure of a MLP network

The weights of the network are usually initialize randomly and are progressively changed in each iteration during the training process. Training in such networks means that the network has to learn a target function. To do so, each input together with its corresponding output is presented to the network. Learning algorithm tries to adjust the weights in all layers in a way that the error between computed output and correct output become small. The most known and popular learning algorithm is back-propagation learning algorithm. This learning algorithm tries to minimize the overall error of the network based on optimization method called gradient descent. After injection of any input to first layer of the network, the error is calculated base on some pre-defined error function in the output layer. Then the error is fed back to previous layer (hidden layer) to adjust those weights which connect the hidden layer to output layer. These chained adjustments continue until the adjustment of weights of the first layer. This process is done for all records of training data in each epoch. Training is done in sufficient large number of epochs where network converges to a steady state in which the overall error becomes very small. A detailed explanation of FFBP can be found in [36].

The most known activation functions used in multilayer perceptron networks is as follows [36]:

- Purelin activation function: It is known as a pure linear function which outputs any inputs without any change. It is often used in input and sometimes output layer of perceptron networks.

- Hardlim activation function: This simple linear function just outs two values 0 or 1. If the input of the function be greater than or equal to 0 then the output would be 1 and adversely if the input be less than zero then the output would be 0. This function is usually applied in output layer of the perceptron networks.

- Logsig activation function: This function is known as the most used nonlinear activation function within [-∞ +∞] interval (Equation (1)). The output stands in [0 1] interval. Most important features of the sigmoid functions are continuity and differentiability.

$$\mathrm{Logsig(n)} = \frac{1}{1+\exp(-n)} \tag{1}$$

- Tansig activation function: Tansig is another sigmoid function used in hidden layer of perceptron networks. It is symmetric of Logsig function (Equation (2)) and the output stands in [-1 1] interval.

$$\mathrm{Tansig(n)} = \frac{2}{1+\exp(-2n)} - 1 \tag{2}$$

- Radbas activation function: Radbas is the acronym of radial basis whose output depends on the distance from center of Gaussian function [37]. It is typically used in approximate and regression functions (Equation (3)).

$$\mathrm{Radbas(n)} = \exp\left(\frac{-n^2}{\sqrt{2}\sigma}\right) \tag{3}$$

1.3. Artificial Bee Colony Algorithm (ABC)

ABC [27] is known as the kind of optimization techniques. It is mostly used in solving multidimensional and multarget optimization problems [28]. There are some unsupervised interacting individuals in the algorithm which intelligently interact each other to share and transmit their information to whole swarm in order to maximize the nectar amount loaded to the hive. There are also some food sources positions and colony consists of three groups of self-organized and intelligent bees: employed bees, onlooker and scout bees [29]. Self-organization of the individuals (bees) is an essential requirement of dynamic nature of the system. Self-organization means that any decide of bees to change in their position or state directly depend on the state of the swarm or environment [22].

The process is as follows. Each employed bee evaluates the nectar of already discovered food source and then comeback to hive. They start to do a special kind of dance in the hive to attract the more onlooker bees. The dance carries the distance, direction and the amount of nectar information of the food source. Each onlooker chooses and follows one of the employed bees based on its evaluation of the chosen employed bee. The ignored bees change to be scout bees and start to do global search in order to discover new food sources. Their duties are important for the sake of making a balance between exploitation and exploration. This is why mostly is claimed that swarm optimization searches don't get stuck into a local minima and have capability to escape to reach the global minima.

The main difference between onlooker and scout bees relate to what information provide for them in foraging task. Onlooker bees use information shared by employed bees while scout bee use external or random signals to find new undiscovered sources. Memorization of the best food source makes the employed bees as experience individuals about their surrounding environment. So their process is called local search (exploitation) due to the experience and their followers in searching of the food sources [27]. As described, all individuals are foraging in their special ways. The foraging is the essential task to be ensured that colony would be survived.
To make a relationship between underlying ANN with the ABC algorithm, each randomly initialized perceptron neural network was considered to be a source food. To reach a global optimum solution, after training phase of the network at each iteration, the parameters (i.e. weights) of the neural network were changed iteratively based on the ABC algorithm. All process of proposed ANN-ABC algorithm is described in problem formulation section.

In order to know much more about the ABC algorithm, the reference [30] is recommended. In respect to other proposed optimization problems like Genetic Algorithm (GA) and Differential Evolution (DE), ABC algorithm is claimed more simple, robust and flexible [31, 32].

2. Area under consideration

Ghare Aghaj River is taken as the case study of the present study. The river basin is located in the center of Fars Province of Iran. It originate from Bond Rood heights in skirt of Task mountains of Zangane village (30 km of northeast of Kazeroon city). Ghare Aghaj River passes through the Band-e Bahman and Ali Abad-e Khafr hydrometric stations. After joining to Shoor River in Firooz Abad City, Ghare Aghaj crosses the Tang-e Karzin hydrometric station and finally enters the Persian Gulf through the Mond River. The area of the basin is 13055 Km^2 and geographical location of consideration area is from eastern longitude 51º47´ to 54º14´, from northern latitude 28º22´ to 29º54´. Yearly average discharge of *Ghareh-Aghaj* river is reported 18 m^3/s in *Tang-e Karzin* station. The Lowest and most discharge in this station is reported 3.5 and 43 m^3/s, respectively. The location of the area is shown in Figure 2.

The importance of *Salman-e Farsi* dam is for providing the drinking water and agricultural land surround the dam. The type of dam is gravity dams made from concrete. With respect to severe drought years in south region in iran specially in recent decade, dewatering of the dam has vital role in life of many people who lives in cities near the dam. This dam can reserve 43 million cubic meters yearly can provide 13 megawatt electricity impress on life of jahram, ghir, karzin, gerash and larestan cities.

The daily real value discharge dataset used in the present study were collected in 36 years of daily observations (01.01.1971-31.12.2007) of three stations (*Tang-e Karzin, Band-e Bahman and Ali abad-e Khafr).* Altogether 13149 daily records relate to 36 years with respect to have 9 leap years was gathered. The data just contained daily discharge values. All data was provided by Fars Regional Water Authority[1] located in Shiraz (Center of fars province). Unfortunately, no more information (such as mean rainfall, runoff characteristic and …) was provided by the authority and wasn't found in any other source.

3. Problem Formulation

First, 13149 records of daily discharge values transformed to weekly information and then normalized each station as follow:

$$Norm(x) = \frac{x - min_i}{max_i - min_i} \tag{4}$$

In which min_i and max_i are maximum and minimum discharge of i'th station, respectively.
The objective of the first phase of the current study is to construct an ANN prediction structure to predict the weekly discharge of Tang-e Karzin station. In addition to Tang-e Karzin information, it's two topside stations *Band-e Bahman* and *Ali abad-e Khafr* discharge information was also used in the constructed model.

The structure of neural network used in this paper, had three layers as input, output and hidden layer. Main idea was alternatively changing both number of neurons within hidden layer and also kind of activation function in this layer. For the sake of using three station in prediction task, number of neurons existed in input layer was set to three. Also one neuron was assigned for output layer due to discharge prediction of just one station. Furthermore number of neurons within hidden layer of the network altered from 5 to 30 ascending by step 5.

Evaluation of the network was performed for multiple values of two mentioned parameters of MLP network. Five activation functions (two linear and three nonlinear) was separately employed in hidden layer. Having 6 different values for number of neurons (5, 10,15,20,25 and 30) and 5 activation functions, 30 different states of the MLP network was tested.

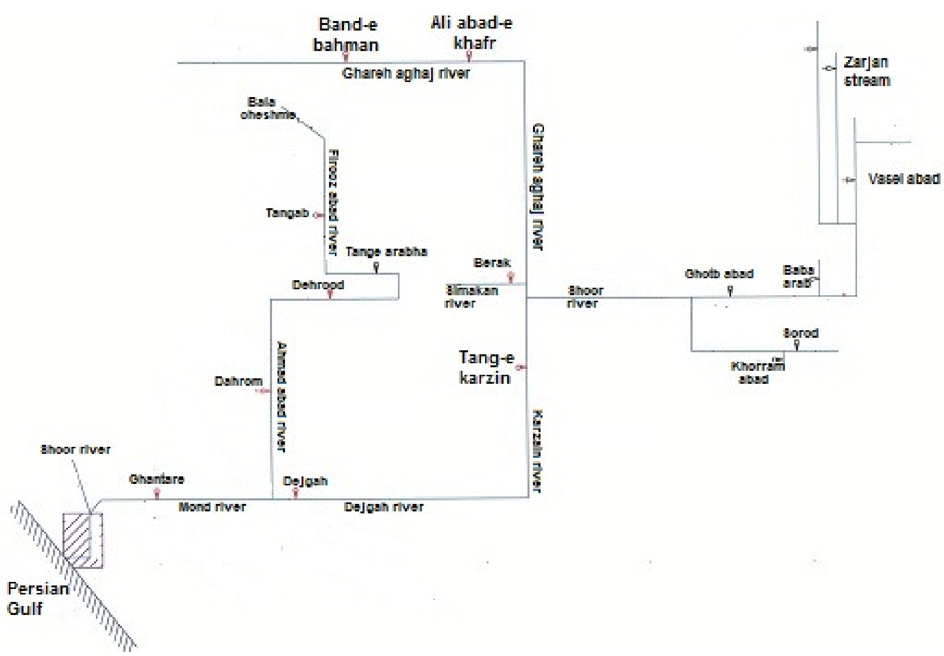


Figure 2. The schematic situation of underlying stations location

For more confidence and integration in localization of the algorithm, more searches were done surrounding the super neuron (i.e. one of six values of neurons number). To do this, three minus and three plus of super neuron value along with best discovered activation function was tested. Although the results of MLP network was satisfying in discharge prediction, it had the worth to try use an optimization algorithm in order to more adjust the structure parameters of the network (i.e. weights). To achieve this purpose, in the second phase of the algorithm, artificial bee colony idea was used along with the training of the network at each iteration. This algorithm during its process tried to adjust the hidden and output weights of the hidden layer and simultaneously implemented by the train step of perceptron neural network. More description on applying the ABC algorithm in MLP networks is as follows.

At first, number of n MLP neural network was created and weights of each network were randomly initialized. Each network is corresponds to one food source in bee colony algorithm. Assigning different values of weights in n networks, leads to have different amounts of nectars and also different positions of n food sources. In the next step, each network was trained using the same train data and then evaluated them based on the test data. The network with the minimum error was chosen as best food among n other foods. All weights of the first layer of network structure depend on number of input and hidden neuron. Since network is fully connected, multiplying the number of input and hidden neurons gives the number of parameters which have to be considered in ABC algorithm. Due to linearity in output layer, the weights in the second layer were ignored. The employed bee corresponding to best food, in order to improve and even maximize its performance starts to explore its neighbors. Chosen neighbor equivalent to change the value of weights. The changing equation is shown in equation (6) and (7).

$$v_{ij} = x_{ij} + \emptyset_{ij}\left(x_{ij} - x_{kj}\right) \;\; i \neq k, k\epsilon\{1,2,\dots,N\}, \emptyset\epsilon[0\ 1] \tag{6}$$

where v_{ij} defines the neighbor of chosen employed bee, x_{ij} defines the coordinates of the best food position and also $\emptyset_{ij}$ is a random and control function to prevent violation of scope of the search space and k is a random value alters within $1,\dots,N$ range. Simpler form of equation (6) is shown in equation (7).

$$New_{x_i^j} = x_i^j + rand[0,1](x_i^j - neighbor(x^j)) \tag{7}$$

where $x_i{}^j$ stands for i'th weight, which was chosen randomly in j'th MLP structure. The process was iteratively performed on each food source in predetermined cycles. In each iteration, evaluation was done on new food using test data. If the fitness of evaluated new food (i.e. food whose weights are changed) was better than previous food (i.e. before changing the weights), then new food would replace with the older one.

In order to implement the scout bee's role in algorithm, randomizing task was entered within the search process based on fitness which earned through each employed bees. Based on equation (8) whatever the fitness was higher, the probability of corresponding employed bee becomes less to become scout bee and vise versa.

$$\mathrm{prob(i)} = 1 - \frac{\mathrm{fitness(i)}}{\sum_{\mathrm{k=1}}^{\mathrm{N}} \mathrm{fitness(k)}} \tag{8}$$

If the calculated probability of each employed bee was less than a generated random value in [0 1] range then it would change to scout bee. Scout bees play the globalization role in the algorithm. For each scout bee, new MLP network was created and randomly initialized. Then the created network was trained by train data and evaluated the fitness under test data. Finally, the network (food) corresponding to scout bee was replaced with the best food which had ever been found if the fitness of the scout bee was higher.

4. Experimental Results

4.1. Evaluation Measures

The accuracy of predictions is evaluated using the variety of error indicators as follows:
Mean Square Error (MSE)[33]:

$$\mathrm{MSE} = \frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{y}_i)^2$$

(9)

where N is the number of samples, y_i is the true value of output of i'th sample and $\hat{y}_i$ is corresponding forecasted output value.

The scatter index (SI) is a normalized measure of error and standard metric for model inter-comparisons. It is simply defined as ratio of standard deviation of difference to mean of measurements. Lower value of SI indicates better prediction [34]. Relation of SI is as follows:

$$\mathrm{SI} = \frac{\sqrt{\frac{1}{N}\sum_{i=1}^{N}((y_i-\bar{y})(\hat{y}_i-\bar{\hat{y}}))^2}}{\bar{\hat{y}}} \tag{10}$$

Pearson correlation coefficient represents the linear dependence between observation and their corresponding predictions [35].

$$r = \frac{\sum_{i=1}^{N}(y_i-\bar{y})(\hat{y}_i-\bar{\hat{y}})}{\left[\sum_{i=1}^{N}(y_i-\bar{y})^2\right]^{.5}\left[\sum_{i=1}^{N}(\hat{y}_i-\bar{\hat{y}})^2\right]^{.5}} \tag{11}$$

where $\bar{y}$ and $\bar{\hat{y}}$ are the mean of the true and predicted values in the dataset, respectively. Correlation coefficient (r) varies from 0 to 1 and whatever its value be more then it denotes better agreement between observed and predicted values.

4.2. Results

This paper tries to simulate a prediction structure to predict the average weekly discharge of *Tang-e Karzin* station. Prediction structure is a mixture of MLP network and Artificial Bee Colony (ABC) algorithm. At the first phase, MLP network was employed without any optimization algorithm. As described both number of neurons and the kind the activation function within the hidden layer was alternatively changed to gain the best values of these two arguments of MLP network. Activation function of the only output neuron was set to purelin. The activation function of hidden layer was varied among five mention activation functions and also number of neurons in this layer altered from 5 to 30 with steps 5. The criterion from which used to compare the results was correlation coefficient. When the most sufficient number of neurons was found, in order to regard the globalization part three minus and three plus number of best found neuron was also considered. Two following tables, results of super neuron for any five activation function have been shown for train and test data.

Table 1. Results of train data on MLP

Hardlim	**Purelin**	**Radbas**	**Tansig**	**Logsig**	**Activation Function**	
5	5	4	5	**3**	Best number of hidden layer neurons	
0.112	0.092	0.082	0.085	**0.071**	MSE	Evaluation Measures
0.669	0.624	0.572	0.583	**0.532**	SI	
0.891	0.911	0.913	0.915	**0.921**	r	

Table 2. Results of test data on MLP

Hardlim	**Purelin**	**Radbas**	**Tansig**	**Logsig**	**Activation Function**	
5	5	4	5	**8**	Best number of hidden layer neurons	
0.181	0.112	0.111	0.109	**0.098**	MSE	Evaluation Measures
0.841	0.734	0.730	0.728	**0.695**	SI	
0.849	0.876	0.881	0.886	**0.891**	r	

Table 3. Results of train data on MLP along with bee colony algorithm

Hardlim	**Purelin**	**Radbas**	**Tansig**	**Logsig**	**Activation Function**	
15	15	13	15	**11**	Best number of hidden layer neurons	
0.109	0.089	0.08	0.081	**0.062**	MSE	Evaluation Measures
0.634	0.596	0.523	0.517	**0.458**	SI	
0.89	0.913	0.92	0.921	**0.932**	r	

Table 4. Results of test data on MLP along with bee colony algorithm

Hardlim	**Purelin**	**Radbas**	**Tansig**	**Logsig**	**Activation Function**	
14	14	11	14	**18**	Best number of hidden layer neurons	
0.176	0.106	0.101	0.096	**0.083**	MSE	Evaluation Measures
0.813	0.726	0.708	0.691	**0.643**	SI	
0.854	0.885	0.890	0.898	**0.913**	r	

Tables 1 and 2 demonstrate that Logsig activation function with three and eight neurons in the hidden layer gives the best results for train and data respectively. Also Tansig and Radbas functions are the next sufficient activate functions. Purelin and Hardlim activate functions give lowest results because of their linear structure. The hidden layer is added to solve the complex and nonlinear problems. So this layer needs a nonlinear function and linear functions like Purelin and Hardlim are not sufficient.

In the second phase of the algorithm, Artificial Bee Colony (ABC) algorithm was employed to optimize the weights of the MLP network while the training of the network is performed. Although applying the ABC had small increase in overall running time order, Tables 3 and 4 demonstrated the considerable improve in results.

As can be seen in Tables 3 and 4, the MLP algorithm always outperform the results of Tables 1 and 2, when the ABC optimization algorithm come into use. Again, Logsig activation function performed better among other activation function with 11 and 18 neurons in hidden layer for training and test, respectively. Closer look to four tables reveals that the number of neuron have been increased after applying the ABC algorithm. This issue may be accused by increase in more globalization in search process of the algorithm. Due to both exploration and exploitation nature of optimization algorithms, it leads to make a strong global search around the promising area.
In order to have better comparison scatter plots of three nonlinear activation functions has depicted for the test data. These comparisons were done both when MLP network trained independently (right column of Figure 3) and either when it used ABC optimization algorithm (left column of Figure 3). The scatter plots of left column of Figure 3demonstrate that logsig predictions are closer to true values relative to other two activation functions. With a closer look to right column of Figure 3 and compare each one to corresponding plot in left column, can observe that the sample points are more compact around the best fitted line. These plots clarify that applying ABC within the MLP network has a good effect in discharge prediction.

5. CONCLUSIONS

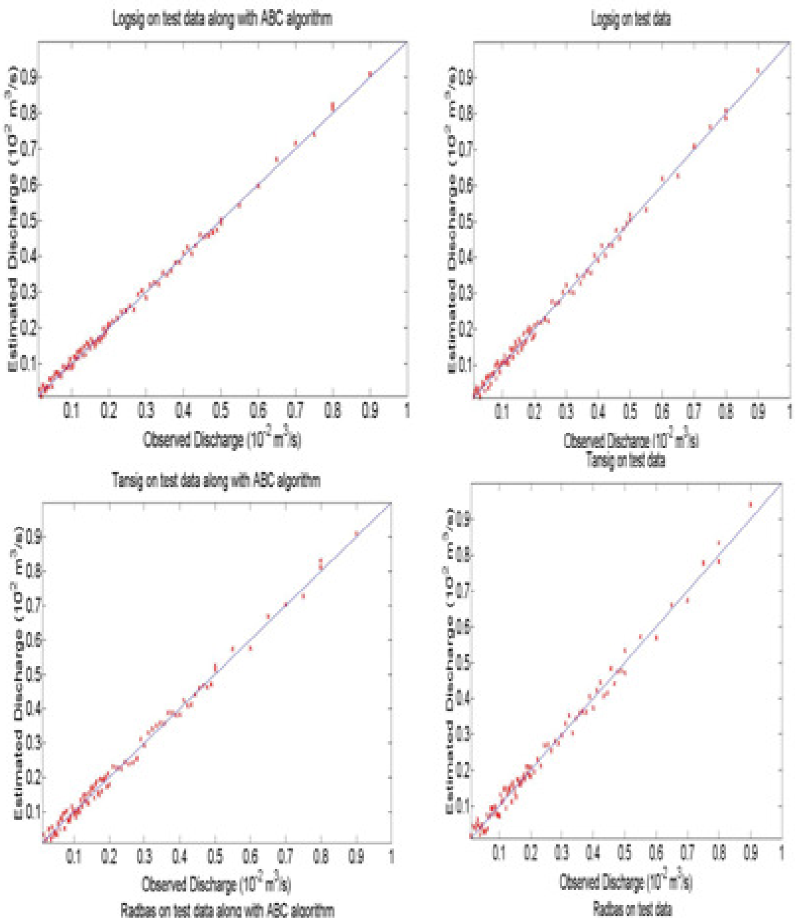
The case study was employed to investigate the capability of artificial neural network (ANN) along with artificial bee colony (ABC) optimization algorithm in modeling of the weekly discharge estimation of *Tang-e Karzin* station located on *Ghareh-Aghaj* river. The accuracy of mixture model was compared with neural network to evaluate the effect of ABC algorithm on ANN. As mentioned in the text, main drawback of ANN was its inability to get out of local minima. The reason is that there is no exploration plan in ANN learning structure to escape from local minima.

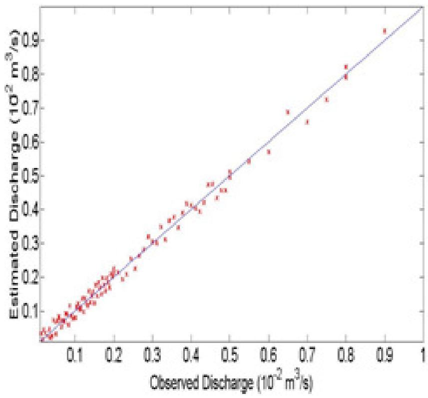
Exploration in ABC helps to algorithm have opportunity to search more globally and investigate undiscovered area in the problem space. The hope was that the combination of ANN and ABC solve the issue. Mean square error, scatter index and pearson correlation coefficient criteria were used as comparison criteria.

Providing correct and appropriate enough number of data, the kind of activation function and also number of neurons in hidden layer of MLP neural network, the comparison result showed that the ANN-ABC algorithm significantly increase the model accuracy respect to when ANN utilized solely.

Of all nonlinear activation function used in this study, Logsig activation function performed better than other activation function in ANN model and also in ANN-ABC model in terms of MSE, SI and r criterions. In the case of ANN-ABC, Logsig needed 11 neurons in training and 19 neurons in testing phase for hidden layer, while when ANN used solely it needed 3 and 8 neurons respectively for train and test phases. The increase in demand for more neurons in hidden layer in ANN-ABC case was probably due to increase in searches especially when both local and global search was taken into account in ABC algorithm.

In this study, ABC algorithm was used to optimize the learned weights of MLP neural network. Some other optimization techniques such as genetic algorithm, Ant colony and swarm optimization algorithms are suggested in order to train or optimize the weights of the ANN network.





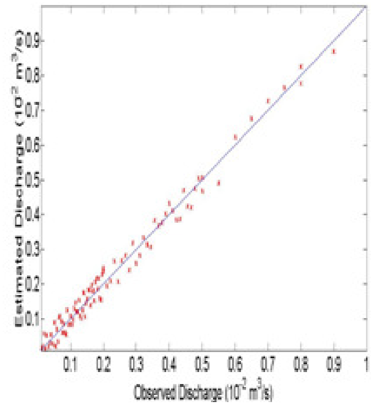


Figure 3. Scatter plot of MLP with and without using of ABC algorithm for Logsig,Tansig and Radbas activation functions, respectively

REFERENCES

[1] Guldal, V. & Muftuoglu, R.F.,(2001). "2D unit sediment graph theory", J. Hydrol. Eng. 6 (2), 132–140.

[2] VanderKwaak, J.E. & Loague, K., (2001). "Hydrologic-response simulations for the R-5 catchment with a comprehensive physics-based model", Water Resources Research 37 (4), 999.

[3] Wang, Y.-C., Chen, S.-T., Yu, P.-S. & Yang, T.C., (2008). "Storm-even rainfall–runoff modeling approach for ungauged sites in Taiwan". Hydrological Processes 22, 4322–4330.

[4] Lohani, A.K., Goel, N.K. & Bhatia, K.K.S., (2011). "Comparative study of neural network, fuzzy logic and linear transfer function techniques in daily rainfall–runoff modeling under different input domains". Hydrol. Process. 25, 175–193.

[5] Huo, Z., Feng, F., Kang, S., Mao, X. & Wang, F., (2010). "Numerically modeling groundwater in an arid area with ANN-generated dynamic boundary conditions". Hydrol. Process. 25, 705–713.

[6] Chen, D., Lu, J. & Shen, Y., (2010)."Artificial neural network modeling of concentrations of nitrogen, phosphorus and dissolved oxygen in a non-point source polluted river in Zhejiang Province, southeast China", Hydrol. Process. 24, 290–299.

[7] Agarwal, A., Mishra, S.K., Ram, S. & Singh, J.K., (2006). "Simulation of runoff and sediment yield using artificial neural networks". Biosystems Eng. 94 (4), 597–613.

[8] Panagoulia, S., (2006). "Artificial neural networks and high and low flows in various climate regimes", J. Hydrol. Sci. 51 (4), 563–587.

[9] Nayak, P.C., Sudheer, K.P. & Ramasastri, K.S., (2005). "Fuzzy computing based rainfall runoff model for real time flood forecasting". Hydrol. Process. 19 (4).

[10] Rabindra, K. Panda., Niranjan, Pramanik. & Biplab, Bala. (2010). "Simulation of river stage using artificial neural network and MIKE 11 hydrodynamic model", Journal of Computers & Geosciences, Volume 36, Issue 6, Pages 735-745.

[11] Emiroglu, ME., Bilhan, O. & Kisi, O. (2011). "Neural networks for estimation of discharge capacity of triangular labyrinth side-weir located on a straight channel", International Journal of Expert Systems 38(1):867-874.

[12] Xu, X.Z. & Li, J.Y. (2002). "Short-term inflow forecasting using an artificial neural network model", Journal of Hydrological Processes, 16, 2423–2439.

[13] Bruen, M. & Yang , J. (2005). "Functional networks in real time flood forecasting- a novel application", Journal of Advances in Water Resources, 28, 899–909.

[14] Kisi, O. (2005). “Daily river flow forecasting using artificial neural network and autoregressive models”. Turkish Journal of Engineering and Environmental Sciences, 29, 9–20.

[15] Cigizoglu, H.K. (2003). “Incorporation of ARMA models into flow forecasting by artificial neural networks”. Journal of Environmetrics, 14, 417–427.

[16] Demirel, M.C., Venancio, A. & Kahya, E. (2009). “Flow forecast by SWAT model and ANN in Pracana basin, Portugal”. Journal of Advances in Engineering Software 40 (7), 467–473.

[17] Adam P. Piotrowski. & Jarosław J. Napiorkowski. (2011). “Optimizing neural networks for river flow forecasting– Evolutionary Computation methods versus the Levenberg–Marquardt approach”. Journal of Hydrology, Volume 407, Issues 1–4, 15, Pages 12–27.

[18] Nguyen, T.D., Phien, H.N., & Gupta A.D., (1999), “Neural Network Model for river flow forecasting”, water SA Vol.25 (1) pp 33-39.

[19] Mrutyunjaya, S., Khatua, KK. & Mahapatra, SS., (2011). “A neural network approach for prediction of discharge in straight compound open channel flow”. Flow Measurement and Instrumentation 22(5): 438-446.

[20] Riahi-Madvar H, Ayyoubzadeh SA & Atani MG (2011). “Developing an expert system for predicting alluvial channel geometry using ANN”. Expert Systems with Applications 38 (1):215–222.

[21] Kisi, O., (2010). “River suspended sediment concentration modeling using a neural differential evolution approach”. J. Hydrol. 389 (1–2), 227–235

[22] Kisi, Ö., Ozkan, C. & Akay, B., (2012). “Modeling discharge-sediment relationship using neural networks with artificial bee colony algorithm”. J. Hydrol. 428–429, 94–103.

[23] Kisi, O., Haktanir, T., Ardiclioglu, M., Ozturk, O., Yalcin, E. & Uludag, S., (2009). “Adaptive neuro-fuzzy computing technique for suspended sediment estimation”. Adv.Eng. Softw. 40 (6), 438–444.

[24] Kisi, O., (2008). “Constructing neural network sediment estimation models using a data-driven algorithm”. Math. Comput. Simul. 79 (1), 94–103.

[25] Tayfur, G., (2002). “Artificial neural networks for sheet sediment transport”. Hydrol. Sci. J. 47 (6), 879–892.

[26] Ebrahimi, H., Jabbari, E. & Ghasemi, M., (2013), “Application of Honey-Bees Mating Optimization algorithm on Estimation of Suspended Sediment Concentration”, World Applied Sciences Journal . 2013, Vol. 22 Issue 11, p1630-1638.

[27] Karaboga, D. (2005). “An Idea Based On Honey Bee Swarm for Numerical Optimization”, Technical Report-TR06,Erciyes University, Engineering Faculty, Computer Engineering Department.

[28] Pham, D., Ghanbarzadeh, A., Koc, E., Otri, S., Rahim, S., & Zaidi, M. (2005). “The Bees Algorithm”. Manufacturing Engineering Centre, Cardiff University, UK.

[29] Karaboga, D. & Basturk, B. (2007). “On the performance of artificial bee colony (ABC) algorithm”.Elsevier Applied Soft Computing, vol. 8, pp.687-697.

[30] Karaboga, D., Gorkemli, B., Ozturk, C. & Karaboga, N., (2011), “A comprehensive survey: artificial bee colony (ABC) algorithm and applications”, Artif Intell Rev, DOI 10.1007/s10462-012-9328-0.

[31] Akay, B. & Karaboga, D. (2012). “A modified Artificial Bee Colony algorithm for real-parameter optimization”, Journal of Information Sciences, vol.In Press, Corrected Proof.

[32] Aytek, A. & O. Kishi, (2007). ”A genetic programming approach to suspended sediment modeling”, Journal of Hydrology, 351: 288-298.

[33] Karunanithi, N., Grenney, W.J., Whitley D., & Bovee, K., (1994). “Neural networks for river flow prediction”, Journal of Computing in Civil Engineering, ASCE., 8(2): 201-220.

[34] Durrant, T. H., Greenslade, D. J. M. & Warren, G.R., (2009), “Assessing Upgrades to the Bureau of Meteorology's Wave Forecasting System Using Satellite Altimeter Data”, Centre for Australian Weather and Climate Research, Issue 12 of CAWCR technical report.

[35] Neto, A.M., Rittner, L., Leite, N., Zampieri, D. E., Lotufo, R. & Mendeleck, A., (2007), “Pearson's correlation coefficient for discarding redundant information in real time autonomous navigation system”, 16th IEEE International Conference on Control Applications, Part of IEEE Multi-conference on Systems and Control, Singapore.

[36] Haykin, S., (1998). Neural Networks – A Comprehensive Foundation, 2nd ed.Prentice-Hall, Upper Saddle River (NJ), pp. 26–32.

[37] Buhmann, Martin D. (2003), Radial Basis Functions: Theory and Implementations, Cambridge University Press, ISBN 978-0-521-63338-3.