Monthly runoff forecasting via an improved extreme learning machine

Wen-jing Niu¹; Yu-bin Chen¹; Yao-wu Min¹; Yu-rong Li¹; Xiao Zhang¹; Zhong-kai Feng²,*

¹Bureau of Hydrology, ChangJiang Water Resources Commission, Wuhan 430010, China; Email: dgniuwenjing@163.com
²School of Hydropower and Information Engineering, Huazhong University of Science and Technology, Wuhan 430074, China; Email: myfellow@163.com

*Correspondence: myfellow@163.com; Tel.: +86-15842496712

Abstract: Generally, monthly runoff prediction is of great importance for effective water resource planning and management. Extreme learning machine (ELM) is a novel training tool for the famous single layer feed-forward neural network. Due to the satisfying performance, ELM is chosen for monthly runoff forecasting in this research. Nevertheless, it is unfortunately found that ELM easily falls into local optima in practice because the randomly-determined computational parameters remain unchanged during the learning process. Thus, this paper tries to develop an improved extreme learning algorithm (IELM) where the evolutionary algorithm is used to search for satisfying computational parameters while the Moore-Penrose generalized inverse method is used to determine the output weights. The IELM method is applied to forecast the monthly runoff of Hongjiadu reservoir in southwest China. The results show that the proposed method outperforms several traditional algorithms with respect to the performance indicators. Thus, this paper provides a new and effective artificial intelligence approach for the monthly runoff forecasting.

1. Introduction

Long-term runoff prediction is an important precondition for effective water resources planning, management and operation [1-3]. Therefore, in the past several decades, hydrologic time series prediction has received tremendous attention of researchers, and models from different perspectives have been proposed to improve the quality of forecasting accuracy [4-6]. In general, these models can be broadly classified into two different groups: physical based models and data based models [2, 3]. For physical based models, although they can help us understand the physical mechanism between rainfall and runoff, due to the complex structures and numerous parameters of models, the predictive ability of physical based models are limited, especially when it is lack of long-series data in monthly runoff prediction. On the other hand, data based models, having the ability of self-adaptation and self-study, can provide accurate prediction results with fewer data and have become increasingly popular in the long-term streamflow prediction among engineers and researchers [1, 7-9].

Since 1980s, plenty of data-based streamflow prediction methods were introduced into the field of hydrological prediction, such as artificial neural network models (ANN) [10, 11], support vector machine [12, 13] and hybrid models [9, 14]. Among them, as one of the most classical methods, ANN was widely used in the field of hydrological forecasting. However, ANN was usually trained by the conventional gradient-based learning algorithms represented by back-propagation (BP), leading to
relatively slow learning rate and easily be trapped in a local minimum [11]. Recently, a novel algorithm, called extreme learning machine (ELM), has been developed for feedforward neural networks with single hidden layer [15, 16]. In ELM, the weights between the input and hidden layers and hidden biases are first randomly, then the weights linking hidden layer to output layers are calculated by the Moore-Penrose (MP) generalized inverse [17, 18]. In this way, compared with other traditional algorithms, ELM not only avoids the calculating parameters like learning rate and stop criteria, but also has faster training speed and higher generalization ability [17]. Now, ELM is at the forefront of nonlinear regression and classification, and has been used for wind forecasting, system identification and machine learning. Therefore, in this paper, ELM is introduced for monthly hydrological series forecasting. However, it was found that, in some cases, ELM may be stuck in local optimum because the randomly preset parameters, such as the hidden neurons biases, may tend to contribute less in minimizing the cost function. In other words, there are certain promotion spaces to enhance the performance of ELM [16, 19, 20]. Evolutionary algorithms provide powerful tools for optimizing the neural network parameters [21, 22]. Thus, in this research, the particle swarm optimization (PSO), a popular population-based stochastic method [23, 24], was used to select the ELM network parameters. Moreover, to ensure better searching precision, chaos was employed to generate the initial population of PSO. Then, we propose a novel learning algorithm called improved extreme learning machine (IELM) where PSO and chaos were used in an effort to cooperate with ELM for monthly runoff prediction. Simulation results show that the proposed IELM method outperforms the ELM method.

The rest of this paper is organized as follows. The IELM method is first proposed after the brief introduction of ELM. Then, the study area, data sets and the measures of accuracy are introduced. In the next sections, the prediction results and discussions are given. Finally, the conclusions are presented followed by acknowledgments.

2. Extreme Learning Machine (ELM)

Suppose the number of nodes in input layer, hidden layer and output layer of the ELM networks are $M$, $L$ and $N$, respectively. The input and output layers in the ELM network employ the linear activation function, while the hidden layer uses the nonlinear activation function, such as sigmoid function and radial basis function. In this paper, the sigmoid function is chosen as the transfer function of the hidden layer because it behaves well in many different fields, which can be given as the following formula:

$$g(x) = \frac{1}{1 + e^{-x}}$$  \hspace{1cm} (1)

where $x$ and $g(x)$ denote the input value and output value of the hidden node, respectively.

The optimization objective of ELM is to minimize the total empirical errors of the training data sets with $J$ samples $(x_j, t_j)$ [15, 25], where $x_j = [x_{j,1}, x_{j,2}, \ldots, x_{j,M}] \in \mathbb{R}^M$, $t_j = [t_{j,1}, t_{j,2}, \ldots, t_{j,N}] \in \mathbb{R}^N$, and $j = 1, 2, \ldots, J$. The objective function can be expressed as

$$\min_{w, b} E(w, b) = \sum_{j=1}^{J} ||f_j - o_j||$$  \hspace{1cm} (2)

where $w$ is the connection weight matrix between the input layer and the hidden layer; $b$ is the bias of the hidden layer; $\beta$ is the connection weight matrix between the hidden layer and the output layer; $o_j$ is the output vector of the ELM network after given inputs $x_j$.

Once weights $w$ and bias $b$ are randomly assigned, the optimization problem in Eq. (2) is equivalent to obtain the least square solution of a linear system showed in Eq. (3). Using the Moore-Penrose generalized inverse matrix theory [17], the weight matrix $\beta$ between hidden layer and output layer can be estimated as $\beta = H^T$, where $H^T$ is the generalized inverse matrix of the hidden layer outputs matrix $H$.

$$H^T \beta = T$$  \hspace{1cm} (3)
\[
H = \begin{bmatrix}
H_1 \\
H_2 \\
\vdots \\
H_{j=J+L}
\end{bmatrix}
= \begin{bmatrix}
g(w_1 \cdot x_1 + b_j) & g(w_2 \cdot x_1 + b_j) & \cdots & g(w_{JL} \cdot x_1 + b_j) \\
g(w_1 \cdot x_2 + b_j) & g(w_2 \cdot x_2 + b_j) & \cdots & g(w_{JL} \cdot x_2 + b_j) \\
\vdots & \vdots & & \vdots \\
g(w_1 \cdot x_{JL} + b_j) & g(w_2 \cdot x_{JL} + b_j) & \cdots & g(w_{JL} \cdot x_{JL} + b_j)
\end{bmatrix}
\]
\quad(4)

where \( w_j \) is the connection weight vector between the input layer and the \( j \)th node in the hidden layer; \( w_i \cdot x_j \) denote the inner product of the weight vector \( w_i \) and input vector \( x_j; \) \( b_j \) is the bias of the \( j \)th node in the hidden layer; \( \beta^i_j = \left[ \beta_{i1}, \beta_{i2}, \ldots, \beta_{iN} \right] \in \mathbb{R}^N \) is the weight vector connecting the \( i \)th hidden node and all output nodes.

### 3. Improved Extreme Learning Machine (IELM)

Since ELM can directly obtain the weights between hidden and output layers by solving the linear system analytically, it can save much of time spent on tuning parameters. However, the output weights of ELM are based on the pre-decided input-hidden weights and biases of hidden layer, the model trained by ELM is random and may be easy to be run into local optimum [19, 22]. Therefore, in this section, a novel method called IELM is proposed, where PSO is employed for estimating optimal ELM parameters. Given a set of training data and \( L \) hidden nodes with an activation function \( g(x) \), the IELM algorithm can be summarized in the following steps:

#### Step 1. Preparation for optimization. Set maximize iterations and population size in PSO, and divide data into training and testing sets.

#### Step 2. Set \( k=1 \), and initialize all particles with chaos mapping.

A set of \( I \) individuals, where each one is composed of all input weights and hidden bias of a ELM network, are initialized using chaos mapping as the first population. The particle \( X_i^k \) can be expressed as follows:

\[
X_i^k = \left[ w_{i1,1}^{j,k}, w_{i1,2}^{j,k}, \ldots, w_{i1,l}^{j,k}, w_{i2,1}^{j,k}, \ldots, w_{i2,l}^{j,k}, \ldots, w_{iM,1}^{j,k}, \ldots, w_{iM,l}^{j,k}, b_{i1}^{j,k}, \ldots, b_{iN}^{j,k} \right]
\quad(6)

Where \( w_{im,l}^{j,k} \) in particle \( X_i^k \) represents the weight between the \( m \)th input node and the \( j \)th hidden node; \( b_{iN}^{j,k} \) in particle \( X_i^k \) represents the bias of the \( l \)th hidden node.

#### Step 3. Calculate the output weights and fitness values of all individuals.

Decode the individual vector to an ELM network, and calculate the ELM network output weight matrix \( \beta^i_j \) with the following equations.

\[
\beta^i_j = \left( H_i^k \right)^T T
\quad(7)

Where \( H_i^k \) is the hidden layer outputs matrix of the \( i \)th particle in the \( k \)th generation.

Then the fitness value \( f(x) \) of each individual is calculated using the following equations.

\[
f \left( X_i^k \right) = \sqrt{\frac{1}{N \times J} \sum_{j=1}^{J} \sum_{i=1}^{I} \left( \sum_{l=1}^{L} \beta_l g \left( w_i \cdot x_j + b_l \right) - t_j^k \right)^2}
\quad(8)

#### Step 4. Update the best-known position of each particle and the best known position of the whole population according to the following two formulas:

\[
PBest_i^k = \begin{cases} X_i^k & \text{if } (k = 1) \\ \arg \min \left\{ f \left( PBest_i^k \right), f \left( X_i^k \right) \right\} & \text{otherwise} \end{cases}
\quad(9)

3
Step 5. Calculate the inertia factor, and update current position of each particle.
Step 6. Set $k = k + 1$, and check whether the calculation can be stopped or not. Output the optimal ELM network if the maximum iterations are reached; otherwise, go back to Step 3.

### 4. Measures of Accuracy

The performances of forecasting models can be evaluated by different statistical measure criteria. In this paper, four commonly used metrics, respectively defined from Eqs. (11)-(14), are selected to evaluate the forecast results: the root mean squared error (RMSE), mean absolute percentage error (MAPE), Nash-Sutcliffe efficiency coefficient ($E$), and coefficient of correlation ($R$). The RMSE measures the precision of the model by comparing the deviation between the observed and forecasted data; the MAPE evaluates the absolute percentage error of the prediction; the $E$ reveals the capability of the model in predicting values away from the mean; the $R$ shows the linear relationship between the forecasted and measured data.

\[
RMSE = \sqrt{\frac{1}{n} \sum_{t=1}^{n} (y_t - \bar{y}_t)^2} \tag{11}
\]

\[
MAPE = \frac{1}{n} \sum_{t=1}^{n} \left| \frac{y_t - \bar{y}_t}{y_t} \right| \times 100 \tag{12}
\]

\[
R = \frac{\sum_{t=1}^{n} (y_t - \bar{y}_t)(\bar{y}_t - \bar{y}_{avg})}{\sqrt{\sum_{t=1}^{n} (y_t - \bar{y}_t)^2}(\bar{y}_t - \bar{y}_{avg})^2} \tag{13}
\]

\[
E = 1 - \frac{\sum_{t=1}^{n} (y_t - \bar{y}_t)^2}{\sum_{t=1}^{n} (y_t - \bar{y}_{avg})^2} \tag{14}
\]

where $y_t$ and $\bar{y}_t$ are the observed value and model output value of $i$th data, respectively. $y_{avg}$ and $\bar{y}_{avg}$ represent the mean value of observed and modeled values, respectively. $n$ is the total number of observed data values.

### 5. Study Area and Data Set

In this study, the monthly streamflow data from Hongjiadu reservoir located in Wu River was used to test the effectiveness of the proposed method. The monthly streamflow data sets from January 1951 to December 2015 are studied. The first 54 years from 1951 to 2005 were used for training while the rest were used for validation. To avoid the numerical problems caused by larger attribute values dominating smaller ones, the data sets are usually normalized before applied for prediction. In ELM and IELM modeling process, the runoff data sets were scaled to the range between 0 and 1 by Eq. (15).

After simulation, the normalized output data should be rescaled to the original forecasted data following the contrary procedure of Eq. (15).

\[
q_t = \frac{q_t - q_{min}}{q_{max} - q_{min}} \tag{15}
\]

where $q_t$ and $q_i$ are the scaled and original flow value at period $t$, respectively; $q_{max}$ and $q_{min}$ are the maximum and minimum values of original streamflow data series, respectively.
6. Prediction Modeling and Input Selection
In general, forecasting models that predict future runoff value based on past records are built. The object using antecedent values to forecast inflow values can be generalized in the following form

\[
q = \text{Model}(x),
\]

which denotes the nonlinear relationship between input vector \( x \) and the target value \( q \). In this research, \( q \) is the flow value in the next period, while, for the sake of simplicity, \( x \) is usually the antecedent flow values with various time lags.

![Autocorrelation and Partial Autocorrelation](image)

Figure 1. (left) ACF and (right) PACF value of Hongjiadu monthly inflow series.

Input variables occupy an important place in model training phase and implementation phase because it can reflect the basic information about the system being studied. Therefore, the input variables should be selected carefully to enhance the precision of the model. In this study, the autocorrelation function (ACF) and the partial autocorrelation function (PACF) are two classic indicators used to determine the number of input vectors [26]. Figure 1 shows ACF and PACF values for lag 1-24 in Hongjiadu, respectively. It can be found that, at lag 12, ACF exhibits peak value while PACF shows a significant correlation. Hence, on the basis of the above-mentioned analysis, the variables \( q_{t-1}, q_{t-2}, \cdots, q_{t-12} \) possess abundant information for future flow \( q_t \) and are selected as input vector for monthly forecasting model.

7. Results and Discussion
In order to verify the effectiveness of the proposed method, a typical three-layer feed-forward ANN model using traditional back-propagation algorithm is selected as the benchmark method. For the sake of comparison, the same training and validation data sets, respectively, are used for three forecasting models, and the four quantitative indexes are employed to evaluate the performances of all the developed models. For ANN model optimized by back-propagation algorithm, the optimal number of ANN hidden nodes was identified with a trial-error procedure by varying the number of hidden nodes. Similarly, the number of hidden neurons in ELM and IELM models is also determined by searching for the optimal performance. For IELM, the maximal iteration is 200 and the number of individuals is 50.

Table 1 shows the statistics results of Hongjiadu reservoir in terms of four performance indicators. It can be seen from Table 1 that three models developed have satisfactory performances during training and validation periods, and the IELM method are superior to those in ANN and ELM. In the training period, the IELM model can improve the ANN model with 6.25% and 12.63% reduction in RMSE and MAPE, while the improvements regarding \( R \) and \( E \) were 6.51% and 3.21%, respectively. In the validation period, when compared with ELM model, the \( R \) and \( E \) values of the IELM method increase by 2.85% and 1.59% respectively, while the values of the forecasting results regarding RMSE and MAPE decrease by 4.05% and 0.60%. Thus, it can be concluded that the IELM model can obtain better results than ANN and ELM models with obvious improvements in four statistical indexes for monthly streamflow prediction.

| Models | Training Period | Validation Period |
|--------|-----------------|-------------------|
|        | RMSE | MAPE(%) | \( E \) | \( R \) | RMSE | MAPE(%) | \( E \) | \( R \) |
|        |      |        |      |      |      |        |      |      |      |
Table 2 shows the observed and forecasted peak flow values using three models. In term of the maximum peak flow in the validation period, the forecasted value of ANN, ELM and IELM are 413.1 m$^3$/s, 404.4 m$^3$/s, and 429.4 m$^3$/s, corresponding to about 40.1%, 41.3% and 37.7% underestimation, respectively. The main reason was that, from 2011 to 2012, it was dry years where runoff was mainly recharged by base flow, and this kind of situation was a fairly rare occurrence in the history of Hongjiadu reservoir, making the performances of all models poorer than that in other years. On the whole, the absolute average relative error of IELM is 25.7%, while ANN has 26.2% error and ELM predicts peak with 20.3% error. Thus, for peak flow estimation, the IELM model has better nonlinear mapping ability than ELM and ANN.

Figures 2-4 show the scatter plots of observed data versus forecasted data using various models developed for Hongjiadu reservoir during training and validation periods. The observed and calculated monthly streamflow by three models in the training period and validation period are shown in Figure 5. It can be clearly seen that all three models can give close approximations for observed runoff data, and the IELM model mimics flow data better than that by ELM and ANN models. It means that the IELM can overcome drawbacks of both ELM and ANN models, and effectively capture the dynamic changing processes of Hongjiadu reservoir.

| Peak NO. | Date  | Observed peak | Forecast Peak | Relative error (%) |
|----------|-------|---------------|---------------|--------------------|
|          |       |               | ANN | ELM | IELM | ANN | ELM | IELM |
| 1        | 2006/7| 173.9         | 241.57 | 278.70 | 202.82 | 38.9 | 60.3 | 16.6 |
| 2        | 2007/7| 329.7         | 261.97 | 275.84 | 203.91 | -20.5 | -16.3 | -38.2 |
| 3        | 2008/6| 422.7         | 368.8 | 284.7 | 323.6 | -12.8 | -32.7 | -23.4 |
| 4        | 2009/7| 201.8         | 190.6 | 224.6 | 200.0 | -5.5 | 11.3 | -0.9 |
| 5        | 2010/7| 335.0         | 318.7 | 273.7 | 309.2 | -4.9 | 18.3 | -7.7 |
| 6        | 2011/6| 111.6         | 52.0 | 66.6 | 78.5 | -53.4 | 40.3 | 29.7 |
| 7        | 2012/7| 689.1         | 413.1 | 404.4 | 429.4 | -40.1 | 41.3 | -37.7 |
| 8        | 2013/6| 288.4         | 263.9 | 267.9 | 287.2 | -8.5 | 7.1 | -0.4 |
| 9        | 2014/7| 509.2         | 199.3 | 373.0 | 387.4 | -60.9 | 26.7 | -23.9 |
| 10       | 2015/6| 206.5         | 229.7 | 222.0 | 256.6 | 11.2 | 7.5 | 24.2 |
| Average (absolute) | | | 25.7 | 26.2 | 20.3 | | | |
Figure 2. The scatter plots of observed vs. forecasted value using ANN model during the (left) training and (right) validation periods.

Figure 3. The scatter plots of observed vs. forecasted value using ELM model during the (left) training and (right) validation periods.

Figure 4. The scatter plots of observed vs. forecasted value using IELM model during the (left) training and (right) validation periods.

Figure 5. Observed and forecasted values in training and validation phases by various models.

To summary, the abovementioned results show that, in the monthly streamflow prediction modeling, the IELM model using population-based algorithms for parameter selection can enhance the generalization ability of the ELM method in some degrees. In addition, due to the complexity of
hydrological system, the models for prediction should be developed according to the hydrological characteristics of the river basins being studied.

8. Conclusions

In the present study, the accuracy of extreme learning machine (ELM), one method with growing popularity for regression and classification, has been investigated for forecasting reservoir monthly streamflow. In order to enhance the performance of ELM, the improved extreme learning machine (IELM) using particle swarm optimization and chaos to select the parameters of ELM network is proposed. The results in Hongjiadu reservoir indicate that the IELM method outperforms the ANN and ELM models in terms of four different statistics measures, increasing the $R$ and $E$ by 2.85% and 1.59% respectively compared to ANN. Thus, with the measurable improvements on the precision of monthly runoff prediction, the IELM model can act as a superior tool to ANN and ELM, and can help reservoir operate in a more sustainable manner. Moreover, the IELM model considering the effects of other hydrological variables in different temporal scales needs to be studied in the future.

Acknowledgments

This study is supported by the National Key R&D Program of China (2018YFC1508002), National Natural Science Foundation of China (51709119), Natural Science Foundation of Hubei Province (2018CFB573), and Fundamental Research Funds for the Central Universities (HUST: 2017KFYXJ193).

References

[1] Wang, W.C.; Chau, K.W.; Cheng, C.T.; Qiu, L., A comparison of performance of several artificial intelligence methods for forecasting monthly discharge time series. *J Hydrol* 2009, 374, (3-4), 294-306.
[2] Valipour, M.; Banihabib, M.E.; Behbahani, S.M.R., Comparison of the ARMA, ARIMA, and the autoregressive artificial neural network models in forecasting the monthly inflow of Dez dam reservoir. *J Hydrol* 2013, 476, (2), 433-441.
[3] Chau, K.W.; Wu, C.L.; Li, Y.S., Comparison of Several Flood Forecasting Models in Yangtze River. *J Hydrol Eng* 2005, 10, (6), 485-491.
[4] Zhao, T.; Zhao, J., Joint and respective effects of long and short-term forecast uncertainties On reservoir operations. *J Hydrol* 2014, 517, (2), 83-94.
[5] Wu, C.L.; Chau, K.W.; Li, Y.S., Methods to improve neural network performance in daily flows prediction. *J Hydrol* 2009, 372, (1-4), 80-93.
[6] Huang, S.; Chang, J.; Huang, Q.; Chen, Y., Monthly streamflow prediction using modified EMD-based support vector machine. *J Hydrol* 2014, 511, (7), 764-775.
[7] Liu, P.; Lin, K.; Wei, X., A two-stage method of quantitative flood risk analysis for reservoir real-time operation using ensemble-based hydrologic forecasts. *Stochastic Environmental Research & Risk Assessment* 2015, 29, (3), 803-813.
[8] Lund, J.R., Flood Management in California. *Water-Sui* 2012, 4, (1), 157-169.
[9] Cheng, C.T.; Feng, Z.K.; Niu, W.J.; Liao, S.L., Heuristic Methods for Reservoir Monthly Inflow Forecasting: A Case Study of Xinfengjiang Reservoir in Pearl River, China. *Water-Sui* 2015, 7, (8), 4477-4495.
[10] Wang, W.C.; Chau, K.W.; Lin, Q.; Chen, Y.B., Improving forecasting accuracy of medium and long-term runoff using artificial neural network based on EEMD decomposition. *Environ Res* 2015, 139, 46-54.
[11] Chen, Y.H.; Chang, F.J., Evolutionary artificial neural networks for hydrological systems forecasting. *J Hydrol* 2009, 367, (1-2), 125-137.
[12] Lin, G.F.; Chen, G.R.; Huang, P.Y.; Chou, Y.C., Support vector machine-based models for hourly reservoir inflow forecasting during typhoon-warning periods. *J Hydrol* 2009, 372, (1-4), 17-29.
[13] Lin, G.F.; Chou, Y.C.; Wu, M.C., Typhoon flood forecasting using integrated two-stage Support Vector Machine approach. *J Hydrol* **2013**, *486*, 334-342.

[14] Wu, C.L.; Chau, K.W.; Li, Y.S., Predicting Monthly Streamflow Using Data-Driven Models Coupled with Data-Preprocessing Techniques. *Water Resour Res* **2009**, *45*(8), 2263-2289.

[15] Huang, G.B.; Zhou, H.; Ding, X.; Zhang, R., Extreme learning machine for regression and multiclass classification. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)* **2012**, *42*(2), 513-29.

[16] Huang, G.B.; Zhu, Q.Y.; Siew, C.K., Extreme learning machine: Theory and applications. *Neurocomputing* **2006**, *70*(1-3), 489-501.

[17] Huang, G.B.; Chen, L.; Siew, C.K., Universal approximation using incremental constructive feedforward networks with random hidden nodes. *IEEE Transactions on Neural Networks* **2006**, *17*(4), 879-92.

[18] Yin, Y.; Zhao, Y.; Li, C.; Zhang, B., Improving Multi-Instance Multi-Label Learning by Extreme Learning Machine. *Applied Sciences* **2016**, *6*(6), 160.

[19] Zhu, Q.Y.; Qin, A.K.; Suganthan, P.N.; Huang, G.B., Evolutionary extreme learning machine. *Pattern Recogn* **2005**, *38*(10), 1759-1763.

[20] Cao, J.; Lin, Z.; Huang, G.B., Self-Adaptive Evolutionary Extreme Learning Machine. *Neural Process Lett* **2012**, *36*(3), 285-305.

[21] Ch, S.; Anand, N.; Panigrahi, B.K.; Mathur, S., Streamflow forecasting by SVM with quantum behaved particle swarm optimization. *Neurocomputing* **2013**, *101*(3), 18-23.

[22] Nahvi, B.; Habibi, J.; Mohammadi, K.; Shamshirband, S.; Razgan, O.S.A., Using self-adaptive evolutionary algorithm to improve the performance of an extreme learning machine for estimating soil temperature. *Computers & Electronics in Agriculture* **2016**, *124*, 150-160.

[23] Chau, K.W., Particle swarm optimization training algorithm for ANNs in stage prediction of Shing Mun River. *J Hydrol* **2006**, *329*(3-4), 363-367.

[24] Kennedy, J.; Eberhart, R., In *Particle swarm optimization*, IEEE International Conference on Neural Networks, 1995. Proceedings, 1995; 1995; pp. 1942-1948 vol.4.

[25] Huang, G.; Huang, G.B.; Song, S.; You, K., Trends in extreme learning machines: A review. *Neural Networks* **2015**, *61*, 32-48.

[26] Lin, J.; Chuntian, C.; Chau, K., Using support vector machines for long-term discharge prediction. *Hydrological Sciences Journal* **2010**, *51*(4), 599-612.