

New formulation for forecasting streamflow: evolutionary polynomial regression vs. extreme learning machine

Mohammad Rezaie-Balf and Ozgur Kisi

ABSTRACT

Streamflow forecasting is crucial in hydrology and hydraulic engineering since it is capable of optimizing water resource systems or planning future expansion. This study investigated the performances of three different soft computing methods, multilayer perceptron neural network (MLPNN), optimally pruned extreme learning machine (OP-ELM), and evolutionary polynomial regression (EPR) in forecasting daily streamflow. Data from three different stations, Soleyman Tange, Perorich Abad, and Ali Abad located on the Tajan River of Iran were used to estimate the daily streamflow. MLPNN model was employed to determine the optimal input combinations of each station implementing evaluation criteria. In both training and testing stages in the three stations, the results of comparison indicated that the EPR technique would generally perform more efficiently than MLPNN and OP-ELM models. EPR model represented the best performance to simulate the peak flow compared to MLPNN and OP-ELM models while the MLPNN provided significantly under/overestimations. EPR models which include explicit mathematical formulations are recommended for daily streamflow forecasting which is necessary in watershed hydrology management.

Key words | evolutionary polynomial regression, optimally pruned extreme learning machine, river flow forecasting, Tajan catchment

Mohammad Rezaie-Balf
Department of Civil Engineering,
Graduate University of Advanced
Technology-Kerman,
P.O. Box 76315-116,
Kerman,
Iran

Ozgun Kisi (corresponding author)
Center for Interdisciplinary Research,
International Black Sea University,
Tbilisi,
Georgia
E-mail: okisi@ibsu.edu.ge

INTRODUCTION

Accurate forecasting of daily river flow is considered essential for many hydrological management decisions such as assessment of risk and control of floods and droughts, hydroelectric energy production, planning of navigation, and sustainable management of water resources. To manage water levels/discharges and to operate water structures more efficiently, new approaches that forecast river flow with high precision and performance are needed (Chen *et al.* 2015). The data-driven modeling depends on distilling and reutilizing information which is incorporated in the hydrological data lacking in straightforwardly clarifying the physical principles that result in river flow processes (Samsudin *et al.* 2011; Ghorbani *et al.* 2016).

In recent decades, soft computing data-driven techniques including artificial neural network (ANN), model

tree (MT), adaptive neuro fuzzy interference systems, gene expression programming, group method of data handling, and multivariate adaptive regression splines (MARS) have been utilized as suitable approaches to estimate complex nonlinear time series in hydrological processes and hydraulic engineering (Sharda *et al.* 2008; Cimen & Kisi 2009; Adamowski *et al.* 2012; Kisi *et al.* 2013; Azamathulla *et al.* 2014; Hamidi *et al.* 2015; Najafzadeh 2015; Seo *et al.* 2015, 2016; Najafzadeh *et al.* 2016a, 2016b; Gorgij *et al.* 2017; Kim *et al.* 2017; Li *et al.* 2017; Yin *et al.* 2017). Sharda *et al.* (2008) predicted the runoff of mid-Himalayan micro-watersheds with limited data using the MARS model. Adamowski *et al.* (2012) compared the accuracy of MARS and wavelet-ANNs models in forecasting runoff of Himalayan micro-watersheds with limited data. Kim *et al.* (2016)

assessed the aggregation and disaggregation of rainfall using different data-driven models and wavelet.

Many AI approach-based data driven models, such as optimally pruned extreme learning machine (OP-ELM) and evolutionary polynomial regression (EPR), can obtain a robust correlation (between predicted and observed value) to forecast daily river flow. ANNs, first presented in 1940s, are the most widely used machine learning methods in various areas of water-related research, e.g., rainfall-runoff modeling, precipitation forecasting, groundwater modeling, and discharge prediction (Bhattacharya & Solomatine 2003; Kisi & Cigizoglu 2007; Nourani *et al.* 2009; Taormina *et al.* 2012; Tapoglou *et al.* 2014). Bhattacharya & Solomatine (2003) used ANNs and MT for modeling the water level-discharge relationship of an Indian river. Kisi & Cigizoglu (2007) compared various data-driven ANN methods in river flow forecasting. Tapoglou *et al.* (2014) forecasted groundwater-levels under climate change scenarios utilizing ANNs with particle swarm optimization.

In recent years, a novel learning approach named extreme learning machine (ELM), which significantly reduces the time required for training single-hidden layer feed-forward neural networks (SLFNs) has been presented (Huang *et al.* 2004). ELM is a nonlinear technique through which all weights and biases related to hidden node parameters can be generated and fixed randomly without tuning. Thus, the weights at the output layer are analytically determined (Schmidt *et al.* 1992). This tends to be regarded as the advantage of the ELM approach, which is able to solve simple linear least squares optimization against the nonlinear optimization. This technique is used to find solutions to various problems in different areas of the hydrologic engineering field, such as predicting reference evapotranspiration (Abdullah *et al.* 2015; Kumar *et al.* 2016), rainfall-runoff modeling (Taormina & Chau 2015), and daily streamflow forecasting (Lima *et al.* 2016).

EPR is a gray-box approach which is generally applied as a predictive model to forecast daily flow of the Tajan used in this research. Giustolisi & Savic (2009) were the first researchers to propose the EPR method in hydroinformatic fields and environmental issues. The methodology of EPR follows the objectives simultaneously as follows: (i) maximizing model accuracy and (ii) minimizing both model coefficients and the number of variables. Thus,

these two criteria would be able to introduce the best model (Laucelli & Giustolisi 2011). Although this method has been used in civil engineering, including modeling failures in urban water systems (Berardi *et al.* 2005; Savic *et al.* 2006), predicting scour depth around piers (Najafzadeh *et al.* 2016a), and soil permeability modeling (Ahangar-Asr *et al.* 2011), it has rarely been used for hydrological issues (e.g., rainfall-runoff modeling, flood forecasting, reference evapotranspiration modeling, etc.). Successful applications of black-box data-driven models in water resources considerations have inspired the exploration of their ability to forecast daily river flow. Investigating preceding research showed that EPR and OP-ELM methods could be rarely employed in hydrological applications.

The main aim of this paper is to investigate the ability of multilayer perceptron neural network (MLPNN), OP-ELM, and EPR in forecasting daily river flow. The paper presents some important points regarding EPR, a case study and datasets further describing development of the proposed models in daily flow forecasting, terminating with forecasting results and conclusions of the research.

STUDY AREA AND DATA

In order to investigate the application of proposed models, the Tajan River basin, located in Mazandaran, Iran was selected as the case study site. Daily river flow estimation was verified for the Tajan catchment, located in the northern part of Iran. The Tajan catchment ($53^{\circ} 56' - 36^{\circ} 17'$ north latitude and $53^{\circ} 7' - 53^{\circ} 42'$ east longitude) is drained by the Tajan River, in Sari, the center of Mazandaran Province, with an area of 4,147.22 km². The Tajan watershed has experienced several floods with the most drastic in 1997. The catchment climate is between semi-humid and cold humid. The average slope of the area, discharge of the river, and annual rainfall are 85%, 20 m³/s, and 539 mm, respectively. The catchment elevation varies between 26 and 3,728 m. There exist nine active hydrometric stations in the Tajan watershed. In this study, datasets of three hydrometric stations, namely, Soleyman Tange, Perorich Abad, and Ali Abad were used for modeling purposes (see Figure 1). The characteristics of these stations are presented in Table 1.

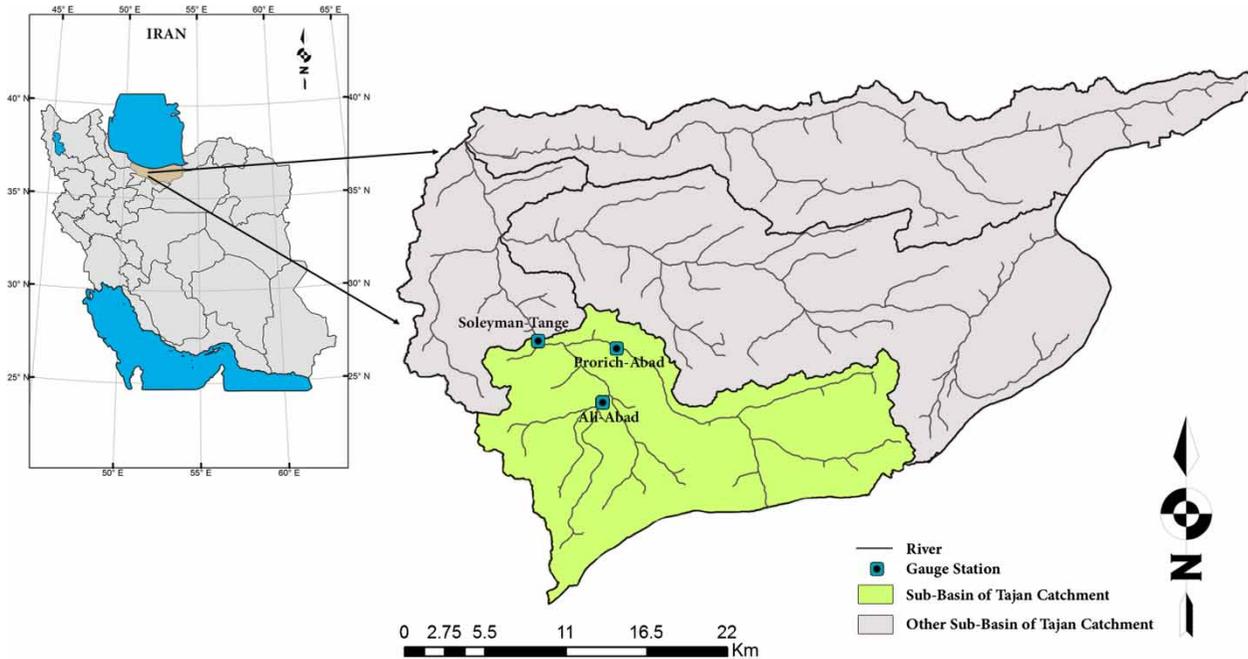


Figure 1 | Location of the three selected flow measuring stations in the Tajan basin of Iran.

The data were obtained from the Meteorological Organization of Mazandaran Province. For this study, daily discharge (m^3/s) data for 12 years (2002–2013) from three gauging stations (i.e., Soleyman Tange, Perorich Abad, and Ali Abad) were used for forecasting. The data were divided into two phases: approximately 75% (1,920 data points) of the datasets were used for the training phase while the remaining datasets (25% of the whole dataset, 637 data points) were kept for testing purposes. Training and test parts can be clearly seen in Figure 2 where time variation graph for the river flows of each station are shown. The other details of datasets and parameter statistics used for the proposed models are given in Tables 2 and 3. In these tables, Q_{\min} , Q_{\max} , Q_{mean} , and S_d indicate the minimum, maximum, mean, and standard deviation of the stream flow data. It is clear

from Table 3 that Soleyman Tange station has the highest standard deviation followed by Perorich Abad.

METHODS

Hydrological process forecasting is one of the important issues providing reliable and accurate applications in water resources activities. Development of three different soft computing techniques, including traditional ANN, OP-ELM, and EPR approaches which were applied to forecast river daily flow are described in this section.

MLPNN

The ANN computational method is inspired by the biological nervous system, which is based on the human brain. The most noteworthy benefit of this method over conventional hydrological models could be its ability to successfully identify both the linear and the nonlinear hydrologic relationship between input and output. Furthermore, ANN can adapt itself to changing circumstances leading to improved performance of the model, reduced calculation times and

Table 1 | Characteristics of the stations

Station name	Geographic coordinate		Height above the sea level (m)
	North latitude	East longitude	
Soleyman Tange	36° 15'	53° 13'	400
Perorich Abad	36° 14'	53° 19'	516
Ali Abad	36° 10'	53° 20'	670

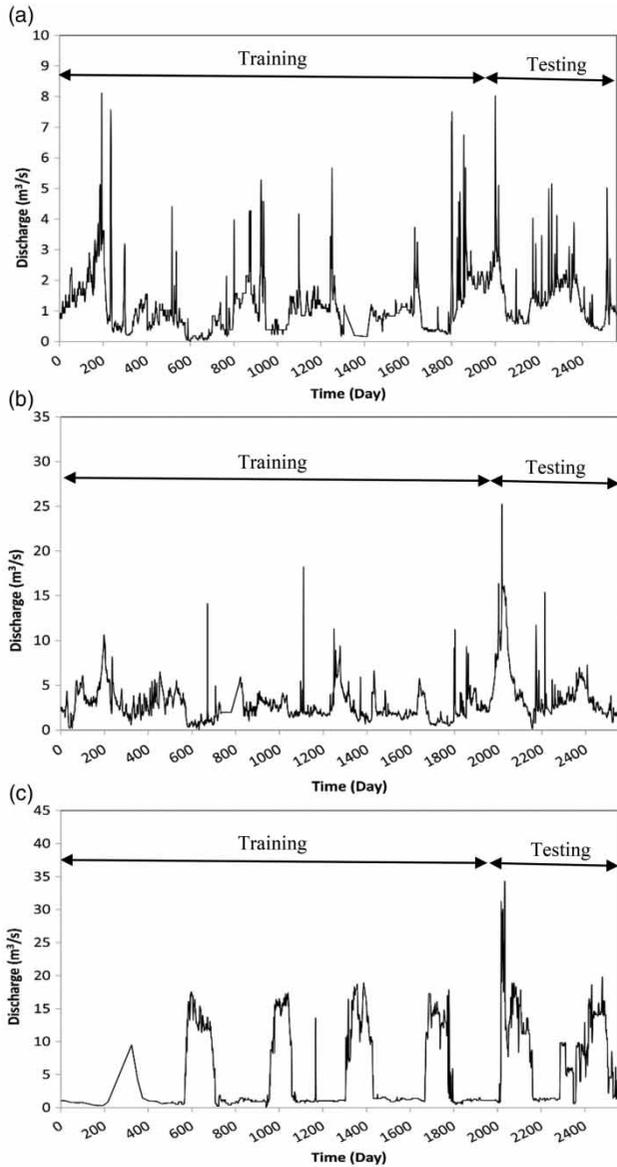


Figure 2 | Time variation for the observed river flows of (a) Ali Abad, (b) Pererich Abad, and (c) Soleyman Tange stations.

Table 2 | Statistical parameters of data and data periods used in training and testing

Stations	Data period	
	Training	Testing
Soleyman Tange	Sep/23/2002– Dec/25/2011	Dec/26/2011– Sep/22/2013
Pererich Abad	Sep/23/2002– Dec/25/2011	Dec/26/2011– Sep/22/2013
Ali Abad	Sep/23/2002– Dec/25/2011	Dec/26/2011– Sep/22/2013

accelerated model enhancement (Birikundavyi *et al.* 2002). Even though there are different ANN types, MLPNN is the common type of neural system most widely used in hydrologic problems (McGarry *et al.* 1999). The MLPNN can comprise one to many neurons and layers. Generally, the structure of a network consists of three layers: the input layer, through which data enter the network; the hidden layer or layers, where data are processed; and the output layer, which is responsible for producing an appropriate response to the given input.

To establish the structure of a neural network, the number of its layers and neurons in each layer, and excitation function of each neuron should be determined to minimize the error. Different schemes and methods are available for the above mentioned process. Some methods proposed to minimize errors in the training process are the Levenberg–Marquardt algorithm, steepest descent, conjugate gradient algorithm, Bayesian approach, and momentum that are in ascending order of calculation speed, all of which follow a back propagation approach. First, some random values are assigned to the weight and bias of each neuron. Subsequently, the initial training sample vector is fed into the network and the output is calculated and compared to the available observed data. The process is followed by modifying weights or parameters using an iterative algorithm (methods mentioned above) in order to make the error become smaller than its present value (Farfani *et al.* 2015). More information on ANN structures can be found in Haykin (1998). Furthermore, the architecture of the MLPNN approach is shown in Figure 3.

ELM

Huang *et al.* (2006) proposed an original algorithm to determine the weights of hidden neurons (see Figure 4) named ELM. It could decrease the computational time required in training and selecting a model structure by hundreds. Moreover, the algorithm is rather simplistic, which makes the implementation easy to process (Miche *et al.* 2008).

For a network with *p* input units, *q* hidden neurons, and *c* outputs, the *i*-th output at time step *t*, is given by:

$$o_i(t) = m_i^T h(t) \tag{1}$$

Table 3 | Data statistics for calibration, validation, and whole of dataset for each station

Stations	Calibration dataset				Validation dataset				Whole dataset			
	Q _{max}	Q _{min}	Q _{mean}	Sd	Q _{max}	Q _{min}	Q _{mean}	Sd	Q _{max}	Q _{min}	Q _{mean}	Sd
Soleyman Tange	18.9	0.07	4.35	5.47	34.3	0.62	7.44	6.40	34.3	0.07	5.12	5.86
Perorich Abad	18.2	0.08	2.69	1.66	25.2	0.17	4.10	3.13	25.2	0.08	3.05	2.21
Ali Abad	8.12	0.02	1.07	0.91	8.02	0.37	1.49	0.86	8.12	0.02	1.17	0.92

Unit of discharge is cubic meters per second.

where $m_i \in \mathbb{R}^q, \forall i \in \{1, \dots, c\}$, is weight vector which links hidden neurons to the i -th output neuron, and $h(t) \in \mathbb{R}^q$ is hidden neuron vector output for a given input pattern $X(t) \in \mathbb{R}^q$ from a dataset $\{X(t)\}_{t=1}^n$. Vector $h(t)$ itself is

$$h(t) = [f(W_1^T X(t) + b_1), f(W_2^T X(t) + b_2), \dots, f(W_q^T X(t) + b_q)]^T \tag{2}$$

where b_l is the bias of the l -th hidden neuron, $W_l \in \mathbb{R}^q$ is the weight vector of the l -th hidden neuron and $f(\cdot)$ is a sigmoidal activation function. Weight vectors w_l are randomly sampled from a normal or even from a uniform distribution. On the other hand, $H = [h(1) h(2) \dots h(n)]$ is a $q \times n$ matrix whose t -th column is the hidden-layer output vector $h(t) \in \mathbb{R}^q$. Similarly, $D = [d(1) d(2) \dots d(n)]$ is a $c \times n$ matrix whose t -th column is the desired (target) vector $d(t) \in \mathbb{R}^c$ associated with the input pattern $x(t), t = 1, \dots, N$. Finally, $M = [m(1) m(2) \dots m(c)]$ is a $q \times c$ matrix, whose i -th column

is the weight vector $m_i \in \mathbb{R}^q, i = 1, \dots, c$. These three matrices are related by linear mapping:

$$D = M^T H \tag{3}$$

where matrices D and H are known and built by the data, while the weight matrix M is unknown. Nonetheless, the weight matrix M can be easily computed by means of the Moore–Penrose pseudo-inverse method as follows:

$$M = (HH^T)^{-1} HD^T \tag{4}$$

Assuming that the number of output neurons is equal to the number of classes, class index i^* for a new input pattern can be calculated by the following decision rule:

$$i^* = \arg \max_{i=1, \dots, c} \{O_i\} \tag{5}$$

where o_i is computed by Equation (1).

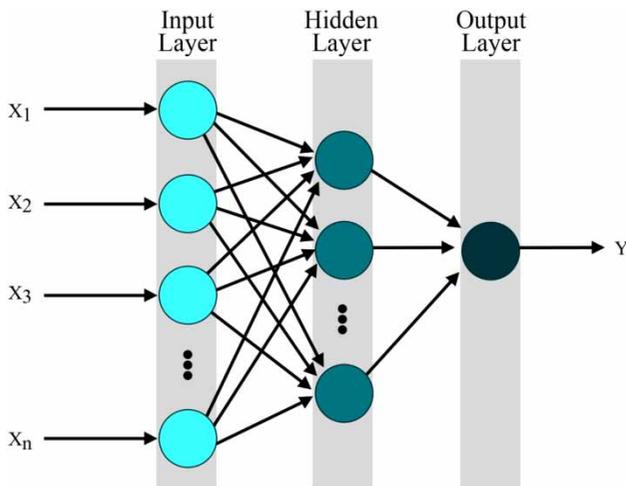


Figure 3 | The MLPNN architecture (shown with a single output variable).

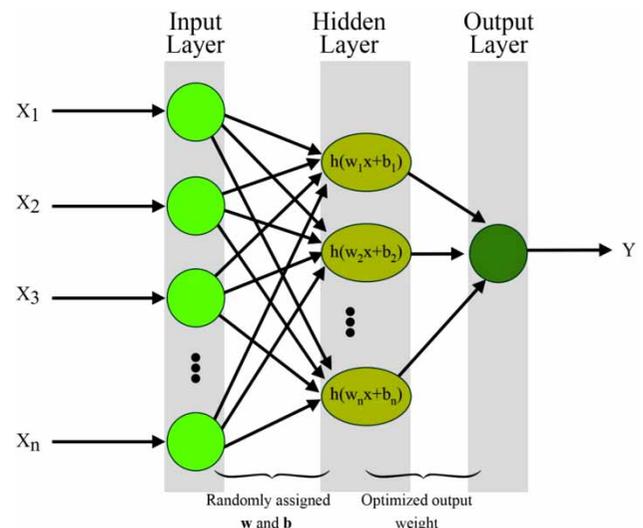


Figure 4 | The OP-ELM architecture (shown with a single output variable).

OP-ELM

The OP-ELM contains three main stages, which are shown in Figure 5.

First, a SLFN is created by plenty of hidden layer neurons. To enhance robustness, OP-ELM methodology tends to use a combination of three different types of kernels instead of sigmoid kernel. Linear, sigmoid, and Gaussian are regarded as possible kernel types. The second stage consists of constructing a multi-response sparse regression (MRSR) responsible for mapping outputs of the hidden layer to targets. MRSR can be considered as an extension of least angle regression algorithm, hence, actually a variable ranking technique rather than a selected one (Alencar *et al.* 2016). In the third stage, the decision on the optimal number of neurons is made by a leave-one-out (LOO) validation method after ranking the neurons. However, the dataset includes a high number of samples, in which case the LOO method can be really time-consuming. To overcome this

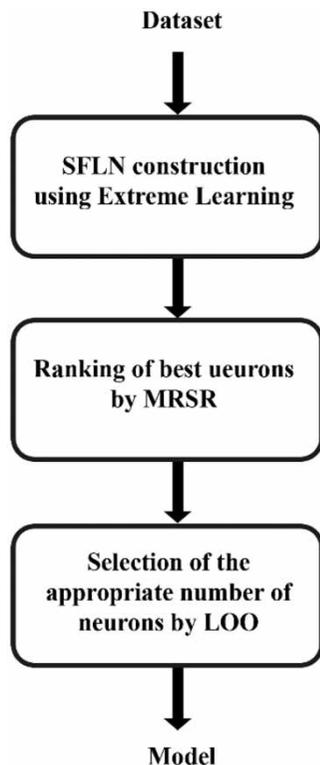


Figure 5 | The three steps of the OP-ELM algorithm.

disadvantage, the prediction sum of squares statistics, can be used giving a direct and exact formula to calculate LOO error in linear models. More details in this respect can be found in Mische *et al.* (2008).

EPR

EPR is a method founded on evolutionary computation programmed to seek mathematical expressions of polynomial structures which illustrate a true physical system. In EPR, at first, symbolic structures made by a genetic algorithm (GA) are searched and the next stage starts by estimating the constant values by solving a least square (LS) linear problem (Giustolisi & Savic 2006, 2009). This technique is a non-linear global stepwise regression, providing symbolic formulas of models.

EPR combines evolutionary algorithm power with numerical regression in order to improve its efficiency. Polynomial models are capable of combining the independent variables simultaneously with user-defined function (Lauccelli *et al.* 2005):

$$y = \sum_{j=1}^m F(X, f(X), a_j) + a_0 \quad (6)$$

where F is a function produced by the process, y is the estimated output of the system, X is the matrix of input variables, a_j is a constant value, f is a function defined by the user, and m is the number of terms of the expression excluding the term a_0 . To enhance the performance of the response (dependent) variable, user-defined function can be included. To construct symbolic models in EPR there are two basic stages: (i) identification of structure and (ii) estimation of parameters, in which a GA simple search method is used to search in the space of the model structure. Similar to other numerical regression, EPR applies the LS method to estimate the parameters of the selected model structure based on the performed GA search. Crossover and mutation are two standard operators of GA where search progresses with them. This is not an exhaustive search since it is impractical to conduct such a search in an infinite search space (Lauccelli *et al.* 2005). EPR permits pseudo-polynomial

expressions which belong to Equation (6):

$$\begin{aligned} \hat{y} &= a_0 + \sum_{j=1}^m a_j (X_1)^{ES(j,1)} \dots (X_K)^{ES(j,K)} f((X_1)^{ES(j,K+1)}) \\ &\quad \dots f((X_K)^{ES(j,2K)}) \text{ case 1} \\ \hat{y} &= a_0 + \sum_{j=1}^m a_j f((X_1)^{ES(j,1)}) \dots (X_K)^{ES(j,K)} \text{ case 2} \\ \hat{y} &= a_0 + \sum_{j=1}^m a_j (X_1)^{ES(j,1)} \dots (X_K)^{ES(j,K)} f((X_1)^{ES(j,K+1)}) \\ &\quad \dots (X_K)^{ES(j,2K)} \text{ case 3} \\ \hat{y} &= g \left(a_0 + \sum_{j=1}^m a_j (X_1)^{ES(j,1)} \dots (X_K)^{ES(j,K)} \text{ case 4} \right) \end{aligned} \quad (7)$$

where \hat{y} is model prediction vector, ES is the matrix of exponents (elements of matrix can assume values within user-defined bounds), and K is the number of variables (inputs) which are independent predictors, whereas each element is the candidate exponent used for each single input. User-specified functions f shown in Equation (7) might be tangent hyperbolic, natural logarithmic, exponential, and so on. It should be noted that the last structure in Equation (7) needs to be assumed as an invertible function g , due to the following stage of estimating parameters. Using the term 'pseudo-polynomial expressions' here is due to the fact that the parameters in any expressions of Equation (7) can be computed as a linear problem and/or true polynomial expression. A typical flow diagram for the EPR procedure is presented in Figure 6. For more detailed information on the EPR procedure the reader can refer to Giustolisi & Savic (2006) and Laucelli *et al.* (2005).

Performance evaluation criteria

To compare the rainfall-runoff simulation performance of the developed models, different statistical indices (Equations (8)–(10)) were used. The indices include correlation coefficient (R), root mean squared error (RMSE), and relative absolute error (RAE):

$$R = \frac{\sum_{i=1}^M (O_i - \bar{O}) \cdot (P_i - \bar{P})}{\sqrt{\sum_{i=1}^M (O_i - \bar{O})^2 \sum_{i=1}^M (P_i - \bar{P})^2}} \quad (8)$$

$$RMSE = \frac{\sum_{i=1}^M (P_i - O_i)^2}{M} \quad (9)$$

$$RAE = \frac{\sum_{i=1}^M |P_i - O_i|}{\sum_{i=1}^M |O_i - \bar{O}|} \quad (10)$$

where O and P signifies the observed runoff and predicted runoff by the model, respectively. \bar{O} is the mean of the observed values and \bar{P} represents the mean of the predicted values. M shows the total number of dataset samples. The R measures how well considered independent variables account for the measured dependent variable. The RMSE is used to measure estimating accuracy, which produces a positive value by squaring the errors. RMSE would increase from zero for perfect estimates to large positive values as the differences between simulations and observations grow. A small value of RMSE and high value of R (up to 1) show high efficiency of the model. RAE is the ratio of the absolute error of the measurement to the accepted measurement. Lower value of the RAE illustrates high performance of the proposed model (Kisi *et al.* 2013; Najafzadeh *et al.* 2014; Kisi & Parmar 2016).

RESULTS AND DISCUSSION

Analysis of results for proposed models

In this study, the MLPNN, OP-ELM, and EPR approaches were investigated for present daily discharge forecasting in Tajan River. As there is no procedure to select the relevant inputs to forecast daily river flow, several input combinations were applied using MLPNN in the three stations studied. The inputs of the MLPNN models present the previously recorded daily discharge values (t-1 to t-6), whereas the output corresponds to the daily flow value at the current time (t). Thus, the combinations of input data of daily flow values are defined in six scenarios that are illustrated in Table 4. Thereafter, with calculating the statistical measures that are presented in Table 5, the optimal input combinations were selected for forecasting the daily flow of Tajan River.

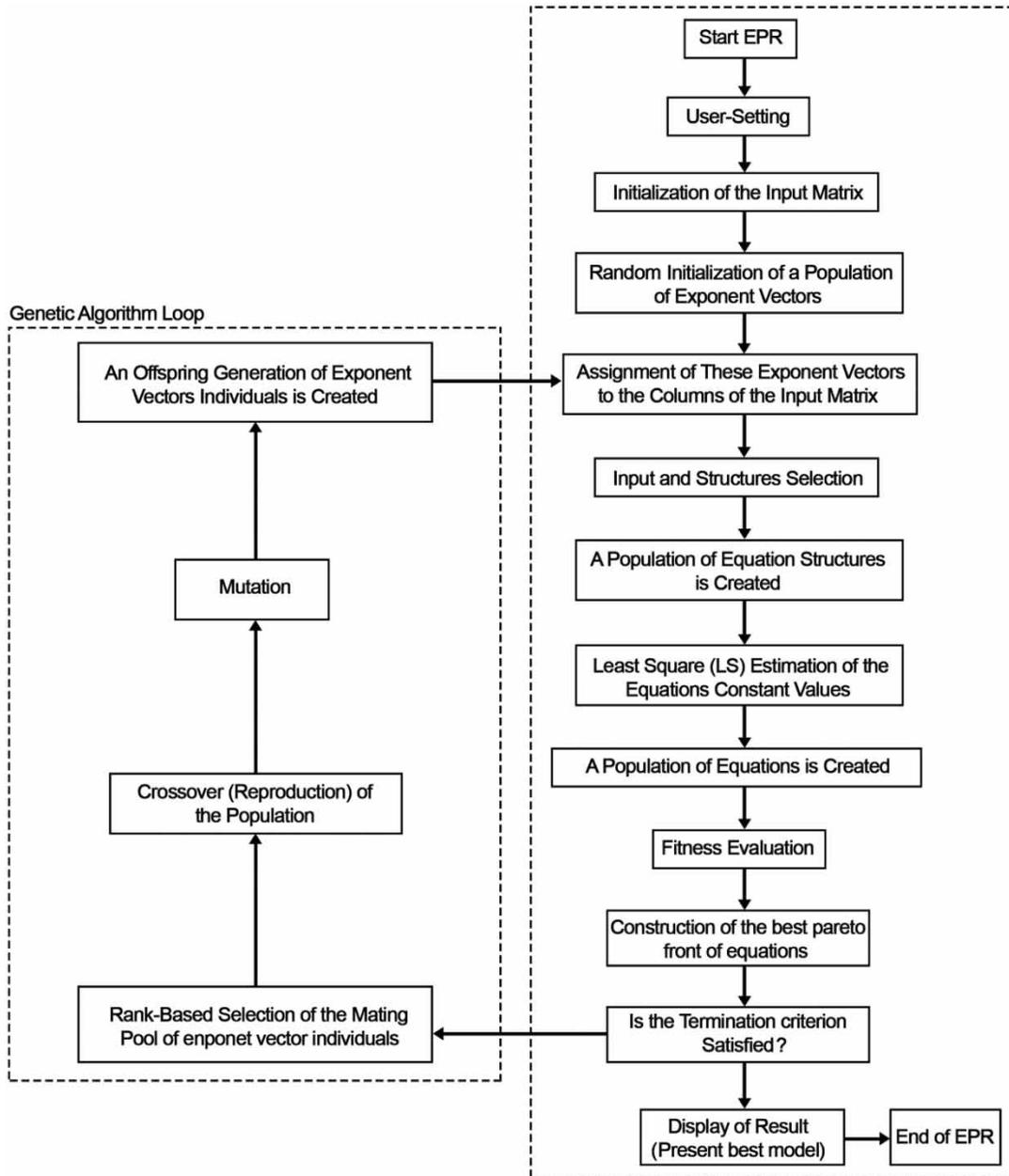


Figure 6 | Flow diagram for the EPR procedure.

The performance evaluation of the MLPNN technique through the R, RMSE, and RAE statistical indexes is made in Table 5. As can be seen from the table, six different MLPNN models with different configurations were applied for three stations. 1-2-1 in the third column of the table indicates an ANN model having 1 input, 2 hidden, and 1 output nodes, respectively. As seen from Table 5, the RMSE of the

MLPNN model's range is 2.743–11.32, 3.874–7.414 and 0.720–2.107 for the Soleyman Tange, Perorich Abad and Ali Abad stations, respectively. It should be noted that these ranges are parallel to the data ranges given in Table 3, in which the streamflow data of Soleyman Tange has the highest range (0.071–34.3 m³/s) and standard deviation (5.86 m³/s). Based on the performance evaluation

Table 4 | Model structures for forecasting daily flow of case study stations

Scenario	Model input
S1	$Q_{(t+1)} = f(Q_{(t-1)})$
S2	$Q_{(t+1)} = f(Q_{(t-1)}, Q_{(t-2)})$
S3	$Q_{(t+1)} = f(Q_{(t-1)}, Q_{(t-2)}, Q_{(t-3)})$
S4	$Q_{(t+1)} = f(Q_{(t-1)}, Q_{(t-2)}, Q_{(t-3)}, Q_{(t-4)})$
S5	$Q_{(t+1)} = f(Q_{(t-1)}, Q_{(t-2)}, Q_{(t-3)}, Q_{(t-4)}, Q_{(t-5)})$
S6	$Q_{(t+1)} = f(Q_{(t-1)}, Q_{(t-2)}, Q_{(t-3)}, Q_{(t-4)}, Q_{(t-5)}, Q_{(t-6)})$

results, it is clear that Scenario 3 (S3) for Soleyman Tange station among the other scenarios (i.e., S1, S2, S4, S5 and S6) has the highest correlation (R = 0.936) and the lowest errors in terms of RMSE (2.742) and RAE (1.666). Also, for Perorich Abad and Ali Abad stations, S3 and S2 were

Table 5 | Test performance measures of the MLPNN model in daily flow forecasting

Scenario	Stations	Best structure of the ANN model	RMSE	R	RAE
S1	Soleyman Tange	1-2-1-1	6.362	0.931	68.01
	Perorich Abad	1-1-1-1	3.874	0.863	6.365
	Ali Abad	1-2-1-1	0.720	0.850	2.556
S2	Soleyman Tange	2-3-1-1	7.632	0.934	21.00
	Perorich Abad	2-2-1-1	4.060	0.865	6.867
	Ali Abad	2-2-1-1	0.547	0.865	2.250
S3	Soleyman Tange	3-2-1-1	2.742	0.936	1.666
	Perorich Abad	3-2-1-1	5.009	0.938	12.98
	Ali Abad	3-2-1-1	1.469	0.862	2.077
S4	Soleyman Tange	4-2-1-1	9.944	0.92	15.11
	Perorich Abad	4-2-1-1	5.416	0.932	13.70
	Ali Abad	4-3-1-1	1.678	0.847	2.162
S5	Soleyman Tange	5-3-1-1	6.528	0.910	26.32
	Perorich Abad	5-3-1-1	7.414	0.924	10.62
	Ali Abad	5-3-1-1	2.107	0.806	2.427
S6	Soleyman Tange	6-3-1-1	11.32	0.820	14.56
	Perorich Abad	6-3-1-1	7.201	0.891	9.441
	Ali Abad	6-4-1-1	2.099	0.783	2.409

Table 6 | Results of performances for proposed models – Soleyman Tange station

	MLPNN			OP-ELM			EPR		
	RMSE	R	RAE	RMSE	R	RAE	RMSE	R	RAE
Training stage	1.181	0.977	0.414	0.928	0.985	0.263	0.841	0.991	0.219
Testing stage	2.741	0.931	1.666	2.136	0.942	0.764	1.836	0.962	0.567

selected as optimal structures for daily flow forecasting, respectively (Table 5).

Comparison of MLPNN, OP-ELM, and EPR methods for the data of Soleyman Tange station

Three different AI techniques were developed for forecasting of daily flow at Soleyman Tange station. The performance of the tested approaches was analyzed by computing the statistical error functions for daily river flow using the MLPNN, OP-ELM, and EPR, as presented in Table 6. Flow values of three previous periods that were identified in the previous application were used as inputs to the compared models. From the training stage (Table 6), it can be seen that the EPR method estimated the daily flow of Soleyman Tange station with a higher correlation (R = 0.991) and lower error statistics of RMSE (0.841) and RAE (0.219) compared to the OP-ELM (R = 977, RMSE = 0.928, and RAE = 0.263) and MLPNN (R = 977, RMSE = 1.181, and RAE = 0.414) models.

It should be noted that one of the key aspects in using EPR is the capability of this method in giving mathematical expression between the input-output variables, which is not the case in MLPNN and OP-ELM models. The corresponding explicit expression of the present EPR model is:

$$\begin{aligned}
 QS_{(t)} = & 1.1307QS_{(t-1)} \times \exp(-0.5QS_{(t-3)} - 2QS_{(t-2)} - 0.5QS_{(t-1)}) \\
 & + 2.478QS_{(t-3)}^{0.5} QS_{(t-1)}^{1.5} \times \exp(-0.5QS_{(t-3)} - 0.5QS_{(t-1)})
 \end{aligned}
 \tag{11}$$

where $QS_{(t)}$ is daily discharge at present time (t) at Soleyman Tange station. In the testing stage (Table 6), it can be noted that Equation (11), given by the EPR model, estimated the daily flow with a better accuracy than those using OP-ELM and MLPNN techniques similar to the training stage. Figure 7 shows the scatterplots and time

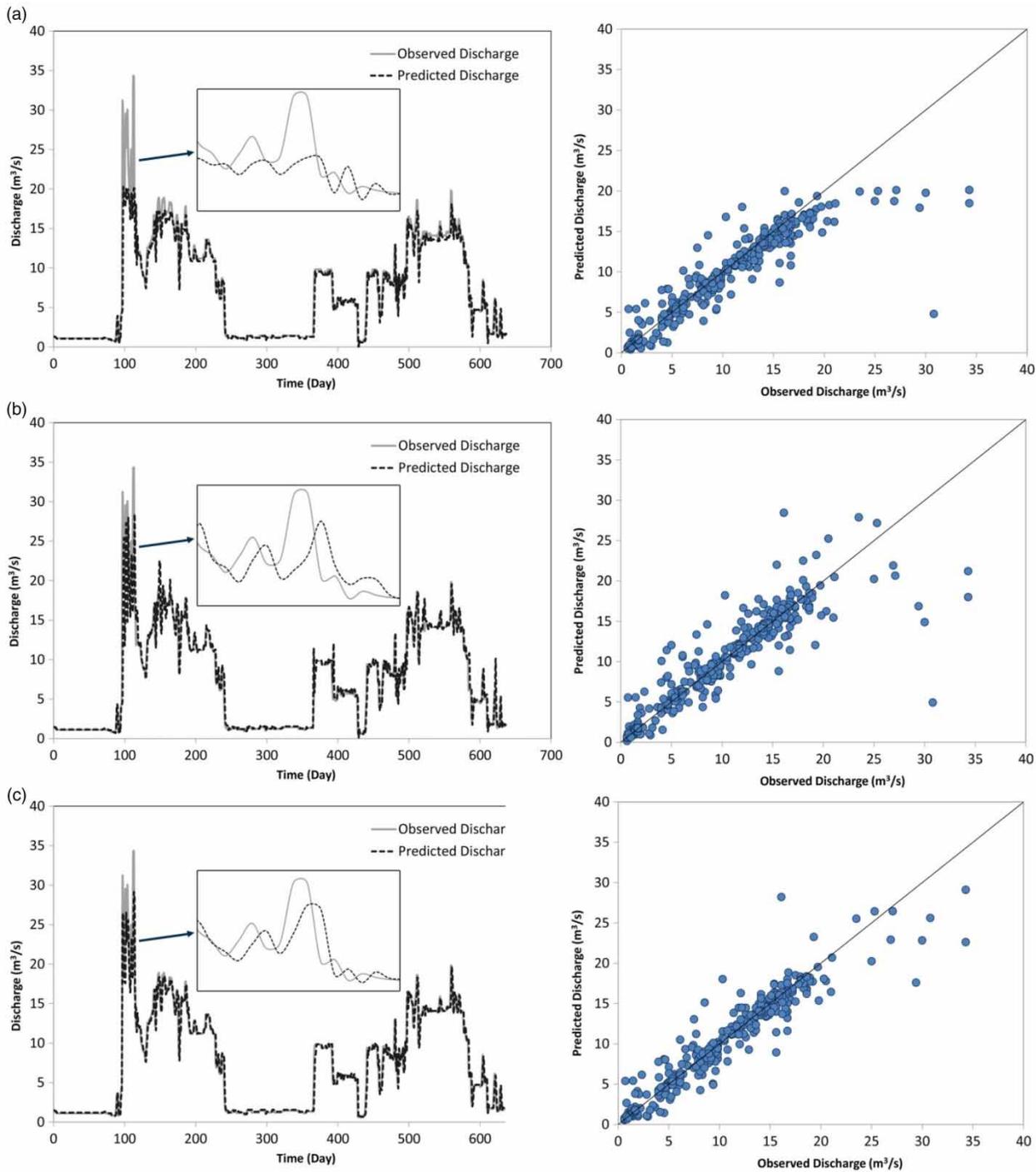


Figure 7 | Hydrographs and scatterplots of observed and predicted river flows for proposed models at Soleyman Tange station: (a) MLPNN model, (b) OP-ELM model, and (c) EPR model.

series graphs of the observed and predicted river flow forecasts using MLPNN, OP-ELM, and EPR models for testing datasets.

On the basis of this figure, the observed and predicted values of daily discharge from EPR showed better agreement than the MLPNN and OP-ELM, especially for the peak

flows. Significantly over- and underestimations are seen for the MLPNN model.

Comparison of MLPNN, OP-ELM, and EPR methods for the data of Perorich Abad station

Similar to the methodology used for Soleyman Tange station, three proposed models (i.e., MLPNN, OP-ELM, and EPR) were constructed for the purpose of daily flow estimation in the Perorich Abad station. The evaluation criteria RMSE, R, and RAE during training and testing stages for proposed techniques are illustrated in Table 7. Statistics given for the training stage clearly show that the EPR model whose inputs are the discharge of 3 previous days better approximates the corresponding observed flow values than the other models with a higher R (0.913) and lower RMSE (0.384) and RAE (1.125). The proposed equation for daily flow of Tajan River at Perorich Abad from the EPR model is:

$$\begin{aligned}
 QP(t) = & 1.2155QP_{(t-1)} \times \exp(-2QP_{(t-3)} - 2QP_{(t-2)} - 2QP_{(t-1)}) \\
 & + 4.7655QP_{(t-3)}^{0.5} QP_{(t-1)}^{1.5} \\
 & \times \exp(-0.5QP_{(t-3)} + 0.5QP_{(t-2)} - 2QP_{(t-1)})
 \end{aligned}
 \tag{12}$$

where $QP(t)$ is daily discharge at present time (t) at Perorich Abad station. Also, it can be observed that Equation (12) from EPR is able to attain acceptable forecasting results as the R, RMSE, and RAE values during the testing period are 0.932, 1.301, and 0.883, respectively. Figure 8 indicates the observed and modeled runoff values with the optimal MLPNN, OP-ELM, and EPR models during the test datasets through the scatterplots and time series graphs. MLPNN gives significant overestimates for the daily river flows of Perorich-Abad station. The peak discharge forecast by the

MLPNN model is about 4.68 m³/s higher when compared with the observed value. This over-forecasting for peak discharge is intolerable for flood warning. According to the scatterplots, the OP-ELM significantly overestimates high discharges (> 10 m³/s) while the EPR significantly underestimates them. The EPR technique provides a better forecast of the peak discharge, and performs better than the MLPNN and OP-ELM techniques.

Comparison of MLPNN, OP-ELM, and EPR methods for the data of Ali Abad station

Similar to the previous two stations (i.e., Soleyman Tange and Perorich Abad), the MLPNN, OP-ELM, and EPR were applied to forecast daily river flow at Tajan basin. The training and testing performance statistics of the proposed techniques are given in Table 8. The EPR model again was the best model in terms of RMSE (0.384), RAE (1.125), and R (0.913) in the training stage compared with the MLPNN and OP-ELM models. The RAE (2.250) and RMSE (0.865) of the MLPNN approach in the testing stage indicate poor performance of estimation of daily flow at Ali Abad station. Figure 9 illustrates the scatterplots and time series graphs of the predicted and observed daily flow forecasts using applied techniques in training and testing stages at Ali Abad station.

As can be seen from Figure 9, the developed OP-ELM model at Ali Abad station is not as good as that constructed for Soleyman Tange and Perorich Abad stations. The OP-ELM and EPR methods have approximately similar performances in low flow forecasting, but in peak flow forecasting the EPR is the more robust method with a higher accuracy than the MLPNN and OP-ELM methods. Also, the peak discharge forecasts by the MLPNN model are underestimated when compared with the observed value. After all these comparisons, the selected equation

Table 7 | Results of performances for proposed models – Perorich Abad station

	MLPNN			OP-ELM			EPR		
	RMSE	R	RAE	RMSE	R	RAE	RMSE	R	RAE
Training stage	0.715	0.902	1.373	0.724	0.903	1.713	0.691	0.912	1.627
Testing stage	5.009	0.937	12.98	1.650	0.924	1.39	1.301	0.932	0.883

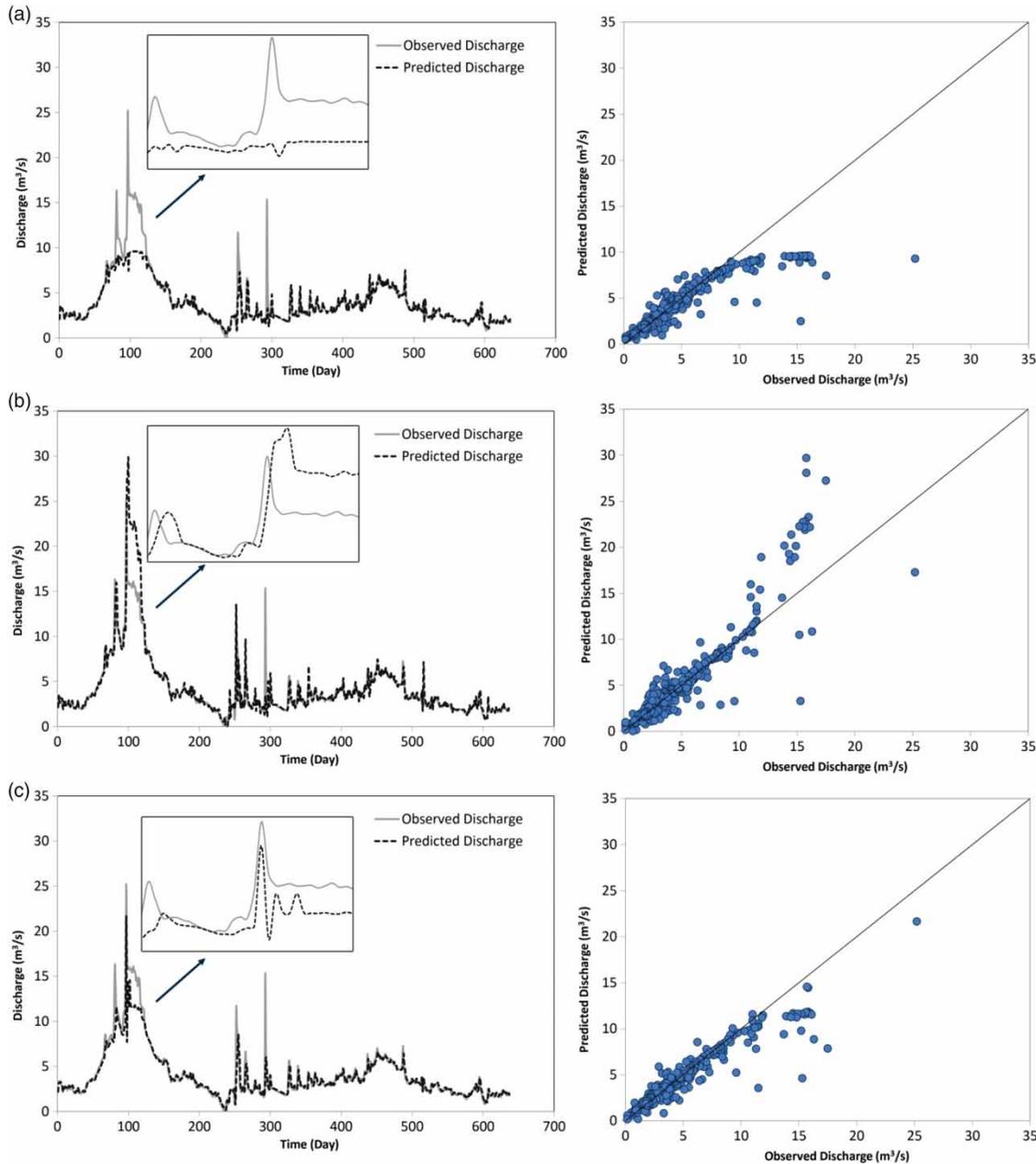


Figure 8 | Hydrographs and scatterplots of observed and predicted river flows for proposed models at Perorch Abad station: (a) MLPNN model, (b) OP-ELM model, and (c) EPR model.

for daily discharge as a trade-off between accuracy and parsimony at Ali Abad station is:

$$\begin{aligned}
 QA_{(t)} = & 1.025QA_{(t-1)} \times \exp(-2QA_{(t-3)} - 1.5QA_{(t-1)}) \\
 & + 3.837QA_{(t-1)}^2 \times \exp(-2QA_{(t-1)}) \\
 & + 0.0114QA_{(t-2)}^2 QA_{(t-1)}^2 \times \exp(1.5QA_{(t-2)} + 1.5QA_{(t-1)})
 \end{aligned}
 \quad (13)$$

where $QA_{(t)}$ is daily discharge at present time (t) at Ali Abad station.

Overall, the EPR seems to more adequate for forecasting daily streamflow when compared to MLPNN and OP-ELM methods. The flexible structure of EPR may have provided it with a better ability to learn investigated phenomena. MLPNN generally underestimated the peak discharge

Table 8 | Results of performances for proposed models – Ali Abad station

	MLPNN			OP-ELM			EPR		
	RMSE	R	RAE	RMSE	R	RAE	RMSE	R	RAE
Training stage	0.392	0.905	2.469	0.390	0.904	1.345	0.384	0.913	1.125
Testing stage	0.547	0.865	2.250	0.432	0.871	0.311	0.408	0.905	0.288

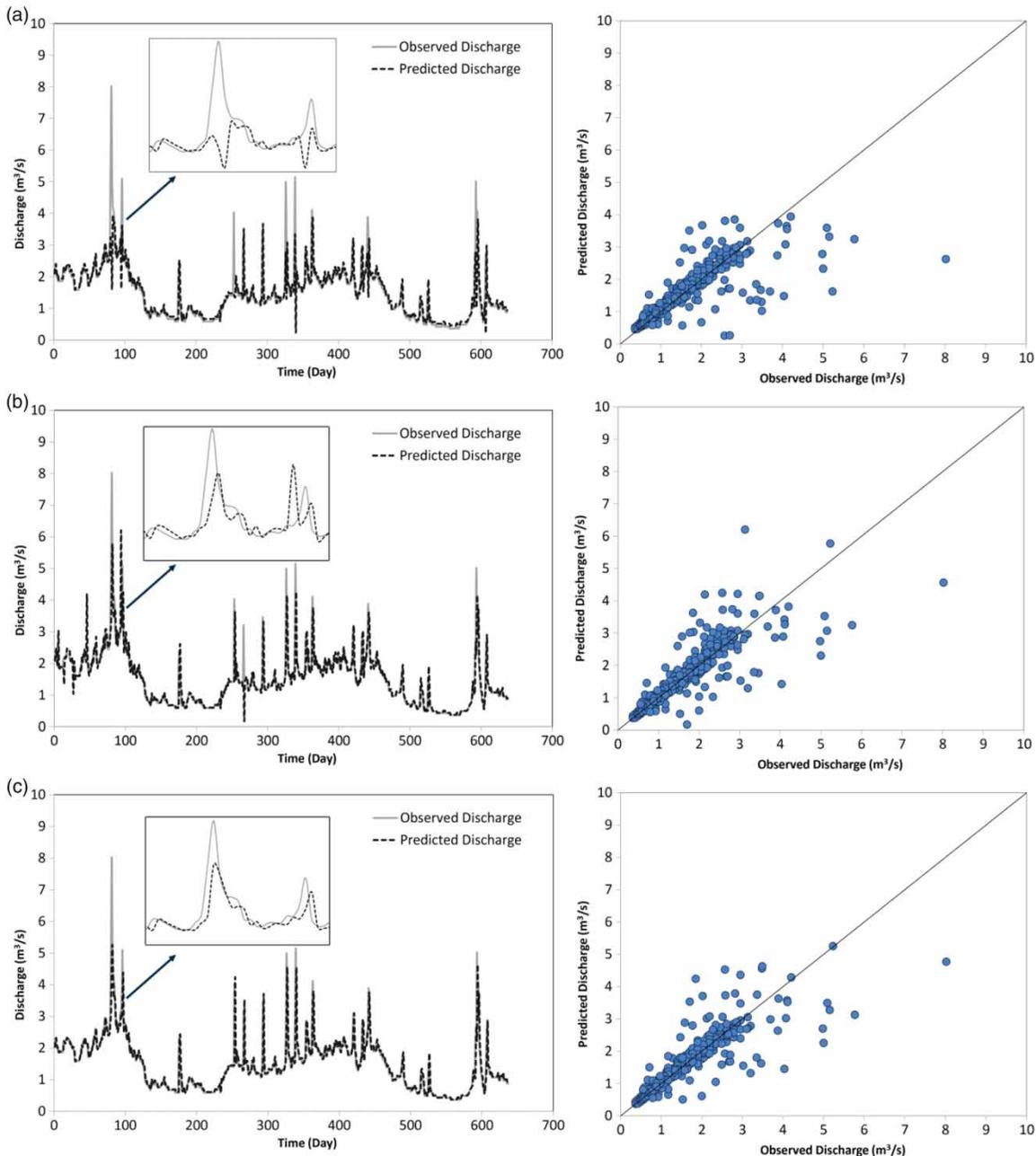


Figure 9 | Hydrographs and scatterplots of observed and predicted river flows for proposed models at Ali Abad station: (a) MLPNN model, (b) OP-ELM model, and (c) EPR model.

values at all three stations. The main reason for this may be the lesser number of data of extreme values. The main disadvantage of the MLPNN method is that it has a high number of data for adequately learning.

CONCLUSION

In this study, MLPNN, OP-ELM, and EPR models were employed for forecasting of daily river flow using inputs of previous daily flows of Soleyman Tange, Perorich Abad and Ali Abad stations that are located along the Tajan River, Mazandaran region of Iran. The MLPNN model was tested by applying it to different input combinations of daily flow data. After selecting the optimal input combination by MLPNN, the performances of the three methods were evaluated based on three different criteria, R, RMSE, and RAE. The obtained results indicated that the EPR model was superior to the MLPNN and OP-ELM models for forecasting daily flows of the studied river. MLPNN model could not simulate the daily discharge values well and the accuracy of this predictive model was generally found to be low. On the other hand, the EPR approach provided a better forecast of the peak flow than the MLPNN and OP-ELM techniques. The main advantage of the EPR model is its explicit mathematical formulations. It can be simply used in practical applications. Instead, the MLPNN and OP-ELM models are black-box models whose formulations are closed.

The proposed techniques may also be compared in other hydrological applications (e.g., short-term wind speed predictions, sea water level forecasting, prediction of daily evapotranspiration). The efficiency and accuracy of the proposed EPR models may be compared with other AI methods in future studies.

REFERENCES

- Abdullah, S. S., Malek, M. A., Abdullah, N. S., Kisi, O. & Yap, K. S. 2015 Extreme learning machines: a new approach for prediction of reference evapotranspiration. *Journal of Hydrology* **527**, 184–195.
- Adamowski, J., Chan, H. F., Prasher, S. O. & Sharda, V. N. 2012 Comparison of multivariate adaptive regression splines with coupled wavelet transform artificial neural networks for runoff forecasting in Himalayan micro-watersheds with limited data. *Journal of Hydroinformatics* **14** (3), 731–744.
- Ahangar-Asr, A., Faramarzi, A., Mottaghifard, N. & Javadi, A. A. 2011 Modeling of permeability and compaction characteristics of soils using evolutionary polynomial regression. *Computers & Geosciences* **37** (11), 1860–1869.
- Alencar, A. S., Neto, A. R. R. & Gomes, J. P. P. 2016 A new pruning method for extreme learning machines via genetic algorithms. *Applied Soft Computing* **44**, 101–107.
- Azamathulla, H. M., Yusoff, M. A. M. & Hasan, Z. A. 2014 Scour below submerged skewed pipeline. *Journal of Hydrology* **509**, 615–620.
- Berardi, L., Savic, D. A. & Giustolisi, O. 2005 Investigation of burst-prediction formulas for water distribution systems by evolutionary computing. *Computer and Control in Water Industry (CCWI)* **2**, 275–280.
- Bhattacharya, B. & Solomatine, D. P. 2005 Neural networks and M5 model trees in modeling water level-discharge relationship for an Indian river. In: *European Symposium on Artificial Neural Networks*, Bruges, Belgium, 23–25 April, pp. 407–412.
- Birikundavyi, S., Labib, R., Trung, H. T. & Rousselle, J. 2002 Performance of neural networks in daily streamflow forecasting. *Journal of Hydrologic Engineering* **7** (5), 392–398.
- Chen, X. Y., Chau, K. W. & Wang, W. C. 2015 A novel hybrid neural network based on continuity equation and fuzzy pattern-recognition for downstream daily river discharge forecasting. *Journal of Hydroinformatics* **17** (5), 733–744.
- Cimen, M. & Kisi, O. 2009 Comparison of two different data-driven techniques in modeling lake level fluctuations in Turkey. *Journal of Hydrologic Engineering* **378**, 253–262.
- Farfani, H. A., Behnamfar, F. & Fathollahi, A. 2015 Dynamic analysis of soil-structure interaction using the neural networks and the support vector machines. *Expert Systems with Applications* **42** (22), 8971–8981.
- Ghorbani, M. A., Zadeh, H. A., Isazadeh, M. & Terzi, O. 2016 A comparative study of artificial neural network (MLP, RBF) and support vector machine models for river flow prediction. *Environmental Earth Sciences* **75**, 476.
- Giustolisi, O. & Savic, D. 2006 A symbolic data-driven technique based on evolutionary polynomial regression. *Journal of Hydroinformatics* **8** (3), 207–222.
- Giustolisi, O. & Savic, D. A. 2009 Advances in data-driven analyses and modelling using EPR-MOGA. *Journal of Hydroinformatics* **11** (3–4), 225–236.
- Gorgij, A. D., Kisi, O. & Moghaddam, A. A. 2017 Groundwater budget forecasting, using hybrid wavelet-ANN-GP modelling: a case study of Azarshahr Plain, East Azerbaijan, Iran. *Hydrology Research* **48** (2), 455–467. doi:10.2166/nh.2016.202.
- Hamidi, O., Poorolajal, J., Sadeghifar, M., Abbasi, H., Maryanaji, Z., Faridi, H. R. & Tapak, L. 2015 A comparative study of support vector machines and artificial neural networks for predicting precipitation in Iran. *Theoretical and Applied Climatology* **119** (3–4), 723–731.
- Haykin, S. 1998 *Neural Networks – A Comprehensive Foundation*, 2nd edn. Prentice-Hall, Upper Saddle River, NJ, USA.
- Huang, G. B., Zhu, Q. Y. & Siew, C. K. 2004 Extreme learning machine: a new learning scheme of feedforward neural

- networks. In: *Neural Networks, 2004. Proceedings. 2004 IEEE International Joint Conference on*, Vol. 2, pp. 985–990.
- Huang, G. B., Zhu, Q. Y. & Siew, C. K. 2006 **Extreme learning machine: theory and applications**. *Neurocomputing* **70** (1), 489–501.
- Kim, S. W., Kisi, O., Seo, Y. M., Singh, V. P. & Lee, C.-J. 2017 **Assessment of rainfall aggregation and disaggregation using data-driven models and wavelet decomposition**. *Hydrology Research* **48** (1), 99–116. doi:10.2166/nh.2016.314.
- Kisi, O. & Cigizoglu, H. K. 2007 **Comparison of different ANN techniques in river flow prediction**. *Civil Engineering and Environmental Systems* **24**, 211–231.
- Kisi, O. & Parmar, K. S. 2016 **Application of least square support vector machine and multivariate adaptive regression spline models in long term prediction of river water pollution**. *Journal of Hydrology* **534**, 104–112.
- Kisi, O., Shiri, J. & Tombul, M. 2013 **Modeling rainfall-runoff process using soft computing techniques**. *Computers & Geosciences* **51**, 108–117.
- Kumar, D., Adamowski, J., Suresh, R. & Ozga-Zielinski, B. 2016 **Estimating evapotranspiration using an extreme learning machine model: case study in north Bihar, India**. *Journal of Irrigation and Drainage Engineering* **142** (9), 04016032.
- Laucelli, D. & Giustolisi, O. 2011 **Scour depth modelling by a multi-objective evolutionary paradigm**. *Environmental Modelling & Software* **26** (4), 498–509.
- Laucelli, D., Berardi, L. & Doglioni, A. 2005 **Evolutionary polynomial regression toolbox: version 1.SA**. Department of Civil and Environmental Engineering, Technical University of Bari, Bari, Italy. Available from: <http://www.hydroinformatics.it> (accessed March 2017).
- Li, X., Sha, J. & Wang, Z.-L. 2017 **A comparative study of multiple linear regression, artificial neural network and support vector machine for the prediction of dissolved oxygen**. *Hydrology Research* **48** (5), 1214–1225. doi:10.2166/nh.2016.149.
- Lima, A. R., Cannon, A. J. & Hsieh, W. W. 2016 **Forecasting daily streamflow using online sequential extreme learning machines**. *Journal of Hydrology* **537**, 431–443.
- McGarry, K. J., Wermter, S. & MacIntyre, J. 1999 **Knowledge extraction from radial basis function networks and multilayer perceptrons**. In: *Neural Networks, 1999. IJCNN'99. International Joint Conference on*, Vol. 4, pp. 2494–2497.
- Miche, Y., Sorjamaa, A. & Lendasse, A. 2008 **OP-ELM: theory, experiments and a toolbox**. In: *Proceedings of the 18th International Conference on Artificial Neural Networks*, Prague, Czech Republic, 3–6 September, pp. 145–154.
- Najafzadeh, M. 2015 **Neuro-fuzzy GMDH systems based evolutionary algorithms to predict scour pile groups in clear water conditions**. *Ocean Engineering* **99**, 85–94.
- Najafzadeh, M., Barani, G. A. & Azamathulla, H. M. 2014 **Prediction of pipeline scour depth in clear-water and live-bed conditions using group method of data handling**. *Neural Computing and Applications* **24** (3–4), 629–635.
- Najafzadeh, M., Balf, M. R. & Rashedi, E. 2016a **Prediction of maximum scour depth around piers with debris accumulation using EPR, MT, and GEP models**. *Journal of Hydroinformatics* **18** (5), 867–884.
- Najafzadeh, M., Etemad-Shahidi, A. & Lim, S. Y. 2016b **Scour prediction in long contractions using ANFIS and SVM**. *Ocean Engineering* **111**, 128–135.
- Nourani, V., Alami, M. T. & Aminfar, M. H. 2009 **A combined neural-wavelet model for prediction of Ligvanchai watershed precipitation**. *Engineering Applications of Artificial Intelligence* **22** (3), 466–472.
- Samsudin, R., Saad, P. & Shabri, A. 2011 **River flow time series using least squares support vector machines**. *Hydrology and Earth System Sciences* **15**, 1835–1852.
- Savic, D. A., Giustolisi, O., Berardi, L., Shepherd, W., Djordjevic, S. & Saul, A. 2006 **Modelling sewers failure using evolutionary computing**. *Proceedings of the Institute of Civil Engineers, Water Management* **159** (2), 111–118.
- Schmidt, W. F., Kraaijveld, M. & Duin, R. P. 1992 **Feedforward neural networks with random weights**. In: *11th IAPR International Conference on Pattern Recognition, 1992. Conference B: Pattern Recognition Methodology and Systems, Proceedings*, vol. II. IEEE, pp. 1–4.
- Seo, Y., Kim, S., Kisi, O. & Singh, V. P. 2015 **Daily water level forecasting using wavelet decomposition and artificial intelligence techniques**. *Journal of Hydrology* **520**, 224–243.
- Seo, Y., Kim, S., Kisi, O., Singh, V. P. & Parasuraman, K. 2016 **River stage forecasting using wavelet packet decomposition and machine learning models**. *Water Resources Management* **30** (11), 4011–4035.
- Sharda, V. N., Prasher, S. O., Patel, R. M., Ojasvi, P. R. & Prakash, C. 2008 **Performance of multivariate adaptive regression splines (MARS) in predicting runoff in mid-Himalayan microwatersheds with limited data**. *Hydrological Sciences Journal* **53**, 1165–1175.
- Taormina, R. & Chau, K. W. 2015 **Data-driven input variable selection for rainfall-runoff modeling using binary-coded particle swarm optimization and Extreme Learning Machines**. *Journal of Hydrology* **529**, 1617–1632.
- Taormina, R., Chau, K. W. & Sethi, R. 2012 **Artificial neural network simulation of hourly groundwater levels in a coastal aquifer system of the Venice lagoon**. *Engineering Applications of Artificial Intelligence* **25** (8), 1670–1676.
- Tapoglou, E., Trichakis, I. C., Dokou, Z., Nikolos, I. K. & Karatzas, G. P. 2014 **Groundwater-level forecasting under climate change scenarios using an artificial neural network trained with particle swarm optimization**. *Hydrological Sciences Journal* **59** (6), 1225–1239.
- Yin, Z. L., Wen, X. L., Feng, Q., He, Z. B., Zou, S. B. & Yang, L. S. 2017 **Integrating genetic algorithm and support vector machine for modeling daily reference evapotranspiration in a semi-arid mountain area**. *Hydrology Research* **48** (5), 1177–1191. doi:10.2166/nh.2016.205.