



A simple and efficient rainfall–runoff model based on supervised brain emotional learning

Sara Parvinizadeh¹ · Mohammad Zakermoshfegh¹ · Maryam Shakiba²

Received: 6 November 2020 / Accepted: 26 August 2021 / Published online: 9 September 2021
© The Author(s), under exclusive licence to Springer-Verlag London Ltd., part of Springer Nature 2021

Abstract

To achieve a robust data-driven flood forecasting model, features such as fast learning, appropriate training using insufficient data and reliable prediction of flood flows are of essential importance. These models also have notable vulnerabilities such as decreased accuracy in forecasting peak discharges, challenging simulation of rainy events, performance deterioration in confronting with inadequate training data and weakness due to reduced number of training epochs. In this paper, the supervised brain emotional learning (SBEL) neural network has been used in daily rainfall–runoff modeling of the Dez Dam watershed in the southwest of Iran, as its first application in the field of hydrology. SBEL is a supervised neurocomputing model inspired by the limbic system in the mammalian brain. To create the right responses, the SBEL models the processing of emotional stimuli and the inhibitory mechanism of incorrect responses to stimuli in the emotional brain. The performance of SBEL was compared to the well-known multilayer perceptron (MLP) with 15–8–1 architecture, through different perspectives. The SBEL outperforms MLP in peak flow prediction, limiting the training epochs, reducing the training samples and predictions of rainy events, while improving the mean relative error by 21%, 59.4%, 74.5% and 14.4%, respectively. By placing reduced training data in dry, normal and wet periods, it has been observed that SBEL has more generalization ability in all flow regimes. Overall, the use of this type of emotional intelligence-based model can be of particular interest in developing reliable early flood forecasting and warning systems.

Keywords Supervised brain emotional learning · Limited training samples · Precise peak discharge forecasting · Fast learning

1 Introduction

Over the past two decades, the data-driven watershed runoff models have played a key role in flood forecasting studies. Extensive efforts have been conducted to enhance the capabilities of artificial intelligence-based methods in the development of flood forecasting and warning systems

[1, 2]. However, due to the complex, nonlinear and dynamic nature of the rainfall–runoff process, accurate modeling of this process has always been challenging [3–5]. The models presented to simulate the hydrologic behavior of the watershed can be classified into three general categories, including physical, conceptual and data-driven models [6]. Where the calibration of conceptual models is challenging and time-consuming due to the multiple parameters and insufficient available datasets, data-driven models can be appropriate alternatives. The popularity of data-driven models is due to their ability to discover complex relationships and create a nonlinear mapping between input and output spaces in the presence of limited and noisy data sets [7–9]. Nowadays, artificial intelligence including machine learning approaches, fuzzy inference systems, evolutionary computations and complementary wavelet models has been found wide applications in hydrology [10–21]. In this regard, artificial neural

✉ Mohammad Zakermoshfegh
moshfegh@jsu.ac.ir

Sara Parvinizadeh
sparvinizadeh@jsu.ac.ir

Maryam Shakiba
shakiba@jsu.ac.ir

¹ Department of Civil Engineering, Jundi-Shapur University of Technology, Dezful, Iran

² Department of Electrical and Computer Engineering, Jundi-Shapur University of Technology, Dezful, Iran

networks (ANNs) as one of the most promising machine learning approaches have been widely used in rainfall–runoff modeling [22–26]. An artificial neural network is a parallel data processing system that is capable of storing information obtained through the learning process as well as generalizing the acquired knowledge to different events [27, 28].

So far, many studies have been conducted to compare different ANN models in river flow forecasting. Anctil et al. [29], using the multilayer perceptron (MLP) neural network in the single-step-ahead flow forecasting, showed that this model is affected by the length of the training dataset. Cigizoglu [30] used the generalized regression neural network (GRNN) to estimate daily mean flows and showed that this algorithm has fewer problems such as trapping at the local minima and has higher accuracy compared to the feed-forward neural network, regression and stochastic methods. Studying different ANN models in some cases has shown the superiority of the radial basis function neural network (RBFNN) model to other supervised methods, due to features such as smaller extrapolation errors, faster convergence and higher reliability [31–36]. However, RBFNN performance is deteriorated by inadequate training data [37]. Furthermore, to enhance the forecasting accuracy of a neural network, the model can be combined with other methods of artificial intelligence; including evolutionary algorithms and wavelet-based methods, which have shown good performance in hydrology science [38–41]. Nourani [42] employed the emotional backpropagation (EmBP)-based neural network for the rainfall–runoff modeling and obtained better results than the feed-forward neural network in multi-step-ahead forecasting. Yaseen et al. [43] used the extreme learning machine (ELM) neural network to the river flow forecasting and showed that the method has a simpler architecture and more learning speed than the MLP. Yaseen et al. [44] combined the ELM model with wavelet-based methods and achieved higher accuracy in river flow forecasting. Araghinejad et al. [45] used a modified neural network that is different in terms of activation function compared to the MLP. Moreover, to increase the forecasting accuracy of low and high flow values, initially, they performed the classification of these flows using a probabilistic neural network. Darbandi and Pourhosseini [46] used the MLP neural network based on the firefly optimization algorithm to river flow forecasting. The firefly optimization algorithm is based on moving the population toward the brightest firefly. Ni et al. [47] predicted the monthly river flow, by developing two hybrid models based on the long short-term memory network (LSTM). A comparison of models performance showed that the LSTM-based hybrid models, LSTM and MLP models illustrated the highest accuracy, respectively.

Some other data-driven methods such as support vector machines (SVMs) have also shown good performance in various prediction problems. YU et al. [48] proposed an algorithm to train SVMs on a bound vectors set that was extracted based on Fisher projection. They used linear and nonlinear problems to verify the proposed algorithm. For each case, they selected a certain ratio samples whose projections were adjacent to those of the other class as bound vectors. For the first time, a two-side cross-domain collaborate filtering model was proposed by YU et al. [49]. They assumed that there existed two auxiliary domains, i.e., user-side domain and item-side domain, where this two-side domain shared the same aligned users and items, respectively. YU et al. [50] proposed a cross-domain collaborative filtering algorithm to overcome some drawbacks including that most recommender systems only take advantage of information from a two-side cross-domain, i.e., user-side domain and item-side domain.

Despite the abilities of artificial neural networks for rainfall–runoff modeling, these methods also shown some challenges and problems in the previous studies. Over-training, the challenge of predicting extreme values, local minima, adjusting the architecture and tuning the training parameters are some of these problems [22, 43, 51–53]. Also, inadequate and limited training data sets reduce the accuracy of ANN prediction [54, 55]. On the other hand, in a low number of training epochs, the accuracy of the ANN will also be reduced. So far, various methods have been proposed to solve these problems, such as wavelet-based artificial neural networks [56–58]. Therefore, providing a model with the least complexity in terms of the number of parameters, simple structure and fast learning at the same accuracy is very important. Recently, the study of human psychological features, as an intelligent emotional entity, has led to the appearance of a new class of ANNs based on emotions [59]. Emotional ANNs can be classified into two general categories, including ANNs based on emotional backpropagation (EmBP) and ANNs based on brain emotional learning (BEL) [60, 61]. BEL network is based on the mammalian emotional brain, which is more like human emotional processes than EmBP [62]. In this paper, as the first application in hydrology, the supervised BEL (SBEL) neural network has been used in rainfall–runoff modeling. Biologically, when faced with a danger in which the logical mind does not have enough time to process, the emotional stimulus can react more quickly using short paths in the emotional brain [63, 64]. The operation of BEL is similar to the response of the brain to emotional stimuli. Rapid response to stimuli due to the existence of short paths in the emotional brain and the presence of a part that inhibits inappropriate responses in critical situations are some characteristics of this kind of neural network, which will increase the accuracy of the model [65]. Therefore, the

model is based on the interaction of emotion and cognition. On the other hand, it has a simple architecture and, due to the lack of hidden layers in its structure, does not encounter the usual challenge in designing the architecture of a MLP neural network. The special features of this class of ANNs motivate this study. The SBEL has been successful in different applications such as chaotic time series prediction [66–69], prediction of geological science events [59, 70, 71] and prediction of wind power [62]. Besides, it has been applied to controlling the heat and air conditions [72] and smart machines [73, 74].

This paper aims to develop a rainfall–runoff model using SBEL on a daily time scale and compare its performance with the MLP as the universal approximator. The biological basis of the emotional brain, structure and mathematical model of SBEL will be described. Then, the ability of SBEL to predict peak flows and also discharge at non-rainy days as well as rainy days (with at least one recorded rainfall event in all precipitation stations) will be investigated. Also, the network performance will be examined in conditions of limited training data by reducing the length of the training dataset in three different scenarios, including dry, normal and wet training periods. The learning speed and accuracy of the network will also be measured by reducing the number of training epochs. In all aforementioned analysis, the SBEL performance has been compared to the MLP as the most common ANN in hydrologic modeling.

2 Methodology

2.1 Multilayer perceptron artificial neural network

The MLP neural network with backpropagation training algorithm is a widely used type of ANNs, applied in various fields of hydrology to predict nonlinear systems [22, 75]. The MLP model, called the universal approximator, can create nonlinear mappings between input and output spaces, which can estimate any nonlinear function desirably [76]. The MLP structure includes one input layer, one or more hidden layers and one output layer [27]. The schematic diagram of an MLP neural network with one hidden layer is presented in Fig. 1. Due to the effect of topology and learning algorithm on the performance of the MLP, it is essential to precisely determine some factors such as the number of neurons in each hidden layer, the activation functions and, finally, the network weights [77]. The number of neurons in the input and output layers is determined based on the problem statement [78]. The number of hidden layers, as well as the number of neurons in each layer, depends on the complexity of the problem

and are usually determined by trial and error [79]. The small number of neurons in hidden layer can lead to reduced network learning ability, while a large number may cause problems such as overtraining [78]. The activation functions in the neurons are responsible for the non-linearization of the input signals [80], and the most commonly used functions are sigmoid and hyperbolic tangent. Adjusting network weights is also performed in the training phase based on the back propagation (BP) algorithm by propagating the error to the back layers to minimize the difference between the observed and calculated values [81].

2.2 Brain emotional learning-based artificial neural network

Emotional learning is caused by emotional stimuli such as rewards and punishments received from various real-life situations and associated with emotional states such as happiness, fear, etc. [82]. All of these processes, such as receiving rewards and punishments, processing emotional stimuli and creating emotional states, take place in the central area of the brain called the limbic system (LS), which plays an important role in the emotional process [83, 84]. Figure 2 shows the limbic system and its main components, including the thalamus, sensory cortex, amygdala, orbitofrontal cortex, hypothalamus and hippocampus [83–85]. There are two ways to get external stimuli by the amygdala. One path is short and fast, including a naïve content that comes directly from the thalamus, and the other is a long and slower path, including more veritable information coming from the sensory cortex [86]. Given the pathways mentioned above, the amygdala will give an imprecise but quick response to stimuli. Then, based on the interactive mechanism that exists in the emotional brain between the amygdala and the orbitofrontal cortex, imprecise responses of the amygdala to the stimuli will be inhibited by the orbitofrontal cortex [87, 88].

The existence of short paths and rapid responses to the learning process of the emotional brain, as well as the correction of inaccurate responses, has led to mathematical modeling of this type of learning process and the emergence of a new class of artificial neural networks based on emotional intelligence. Several models of BEL have been presented so far, all of which are based on the amygdala–orbitofrontal model [61, 87]. In this study, we used the supervised BEL (SBEL) as a universal approximator that was first proposed by Lotfi and Akbarzadeh-T [89] by modifying the amygdala–orbitofrontal model. In Fig. 3, SBEL is presented with n inputs, single output and one sensory cortex area.

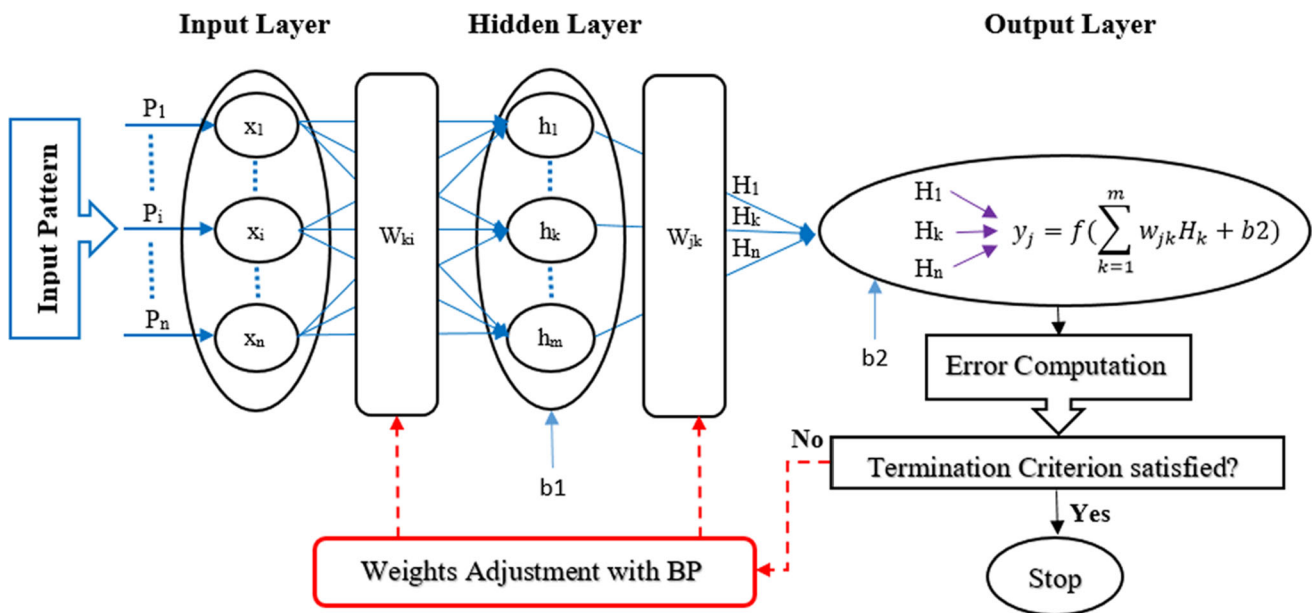


Fig. 1 Multilayer perceptron artificial neural network

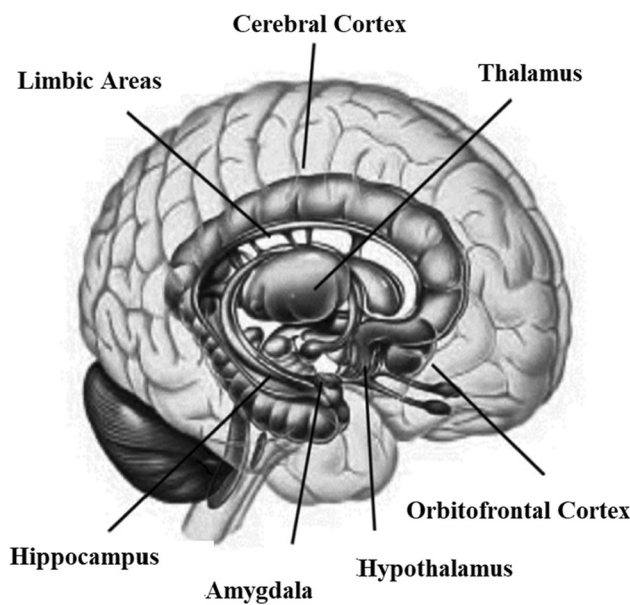


Fig. 2 Limbic system in the brain [82]

The model consists of four neural components in the emotional brain, including the thalamus, sensory cortex, amygdala and orbitofrontal cortex that interact with each other. Figure 3 shows the data flow as solid lines and the learning flow in the form of dashed lines. The number of nodes in each area equals the number of inputs except in the amygdala, which has an additional node.

First, inputs enter the thalamus, and the added imprecise input will be generated using a feature expansion function according to Eq. (1). The feature expansion function makes

the new features by the input characteristics and can be a sinusoidal, Gaussian or max function.

If $p = [p_1, p_2, \dots, p_n]$ be the input vector, then the expanded feature will be computed as:

$$p_{n+1} = \max_{j=1, n, n} (p_j) \tag{1}$$

where p_{n+1} is the expanded input to the amygdala. The sensory cortex receives the input signal p from the thalamus and releases it to the amygdala and orbitofrontal cortex. The orbitofrontal cortex receives n inputs from the sensory cortex. It does not receive any input from the thalamus, while the amygdala, in addition to receiving inputs from the sensory cortex, also directly receives the expanded feature of the thalamus as input. Finally, the SBEL output will be calculated according to Eq. (2).

$$E = E_a - E_o \tag{2}$$

where E_a and E_o are the outputs of the amygdala and orbitofrontal cortex, respectively, and are calculated as follows:

$$E_a = \sum_{j=1}^{n+1} (v_j t \times p_j) \tag{3}$$

$$E_o = \sum_{j=1}^n (w_j \times p_j) + b \tag{4}$$

In Eqs. (3) and (4), v_j is the learning weight of the amygdala, and w_j and b are the learning weight and the bias of orbitofrontal, respectively.

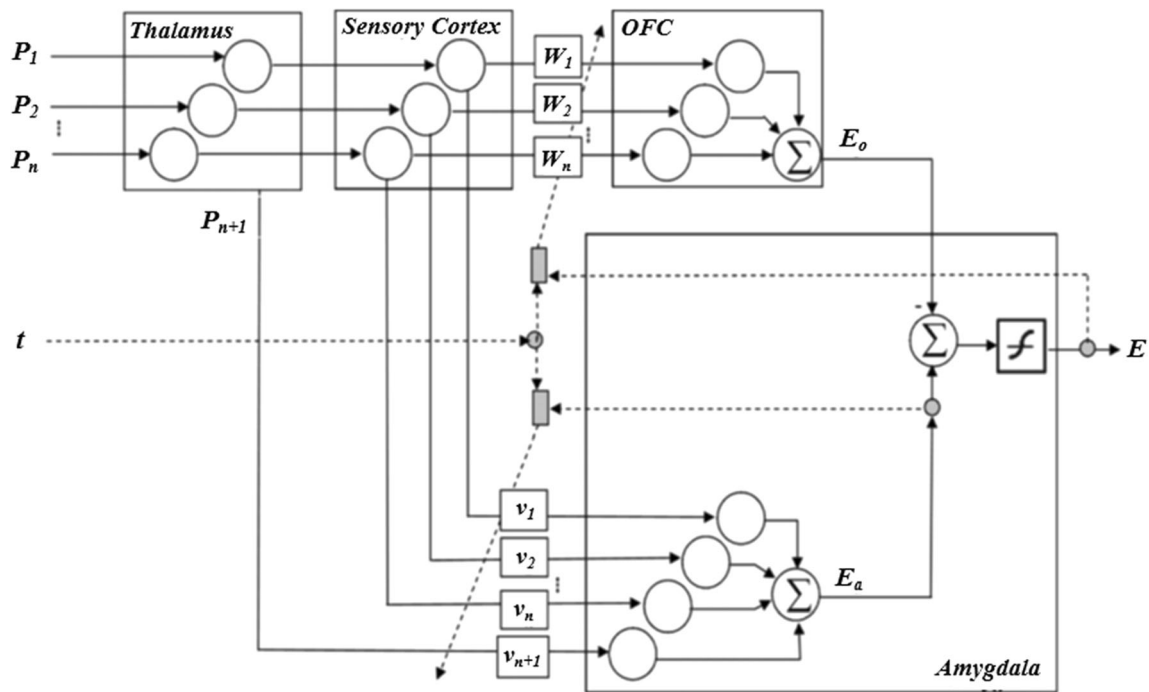


Fig. 3 Brain emotional learning-based artificial neural network [82]

After calculating the final output, it is necessary to correct the weights w , v and the bias so that the network output leads to the least error as follows:

$$v_j^{k+1} = (1 - \gamma) \times v_j^k + \alpha \times \max(T - E_a, 0) \times p_j^k \quad j = 1, \dots, n + 1. \quad (5)$$

$$w_j^{k+1} = w_j^k + \beta \times (E - T) \times p_j^k \quad j = 1, \dots, n \quad (6)$$

$$B^{k+1} = B^k + \beta \times (E - T) \quad (7)$$

where α and β are learning rates, γ is the proposed decay rate, T is the target value, and k is the learning step. The max operator in Eq. (5) will also result in monotonic learning [82]. In the learning process of the amygdala, γ controls the effects of target values as well as monotonic learning in the model and simulates the forgetting role of the amygdala [90]. In this way, controlling the monotonic learning will lead to model performance improvement and consistent decision making [89].

3 Case study and dataset

The Dez River, with upstream catchment area of 17,000 sq. km and an average annual discharge of 245 cms released to the Dez Dam reservoir, is one of the major rivers of Iran in terms of runoff volume. The presence of the 203 meters height Dez Dam and hydroelectric power plant upstream of Dezful city with the aim of flood control, power generation

and supplying water to the Dezful Plain has increased its importance.

With many destructive flood events generated annually in this catchment, it is necessary to develop flood management strategies, especially early flood forecasting and warning systems on the Dez River to optimal maneuvering of the dam gated spillways and reservoir operation to protect the downstream plane.

The Taleh-Zang hydrometric station has been constructed on the Dez River, upstream of the Dez Dam, that measures the inflow to the dam reservoir. The mean daily discharge data at this station besides precipitation data at twelve rain gauges for the period September 1982 to September 2014 (7304 events) were used in this study (Fig. 4).

To predict mean discharge of Taleh-Zang hydrometric station through rainfall–runoff modeling at a given day, fifteen independent variables have been used, including discharges of one and 2 days ago, daily rainfall of twelve rain-gauge stations and a weighted antecedent precipitation index (WAPI). The WAPI is a weighted sum of previous rainfall values used as a modified antecedent soil moisture content of the watershed.

In this paper, WAPI is calculated by Eq. (8) based on the weighted summation of daily precipitation values of 12 precipitation stations in the past 5 days.

$$API = \frac{1(\bar{p}_{t-1}) + 0.9(\bar{p}_{t-2}) + 0.8(\bar{p}_{t-3}) + 0.7(\bar{p}_{t-4}) + 0.6(\bar{p}_{t-5})}{1 + 0.9 + 0.8 + 0.7 + 0.6} \quad (8)$$

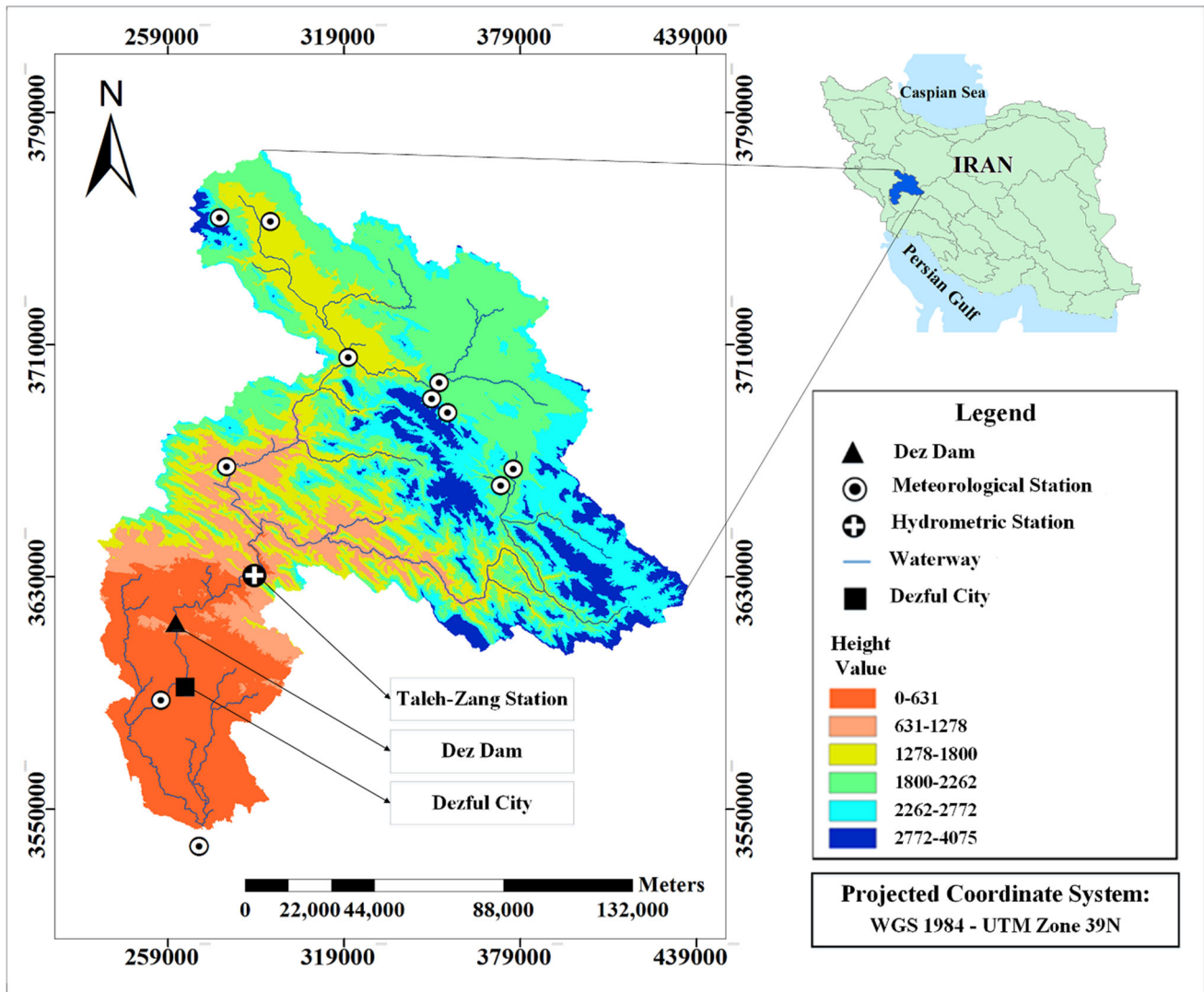


Fig. 4 The location map of the study area

In Eq. (8), WAPI is the weighted antecedent precipitation index in millimeters and \bar{P}_{t-1} , \bar{P}_{t-2} , \bar{P}_{t-3} , \bar{P}_{t-4} and \bar{P}_{t-5} are the antecedent mean daily precipitation of whole 12 precipitation stations during one to five previous days. The weight given to each rainfall average is such that the recent days will receive greater weights, linearly.

4 Results and discussion

4.1 Experimental design

In this paper, the data set was scaled in order to balance the data range, prevent early saturation of the neurons and avoid variables with large numerical ranges dominating the role of the smaller numerical ranges.

The following equation is used for data normalization:

$$x_n = 0.05 + 0.9 \times \frac{(x - x_{min})}{(x_{max} - x_{min})} \tag{9}$$

where x , x_n , x_{min} and x_{max} are raw, normalized, minimum and maximum values of observations, respectively. Accordingly, the data are converted to the range [0.05, 0.95]. Using this domain instead of the range [0, 1] is due to add more flexibility to the model in simulating possible future events. This option allows values less than the minimum as well as larger than maximum of the available data to be better processed in probable future circumstances.

In this paper, statistical indicators such as coefficient of determination (R^2), root mean square error (RMSE), mean relative error (MRE), Taylor diagram and Violin plot have been used to evaluate the performance of MLP and SBEL models in rainfall–runoff simulation.

$$R^2 = \left[\frac{\sum_{i=1}^n (Q_{oi} - \overline{Q_o})(Q_{ci} - \overline{Q_c})}{\sqrt{\sum_{i=1}^n (Q_{oi} - \overline{Q_o})^2 \sum_{i=1}^n (Q_{ci} - \overline{Q_c})^2}} \right]^2 \tag{10}$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (Q_{oi} - Q_{ci})^2}{n}} \tag{11}$$

$$MRE (\%) = \frac{1}{n} \sum_{i=1}^n \frac{|Q_{oi} - Q_{ci}|}{Q_{oi}} \times 100 \tag{12}$$

where R^2 , $RMSE$ and MRE are the coefficient of determination, root mean square error and mean relative error percentage, respectively. Also Q_{oi} , Q_{ci} , $\overline{Q_o}$, $\overline{Q_c}$ and n are the observed discharge, computed discharge, mean observed discharge, mean computed discharge and number of observations, respectively.

The Taylor diagram is a graph that simplifies the comparison and evaluation of different models. It has recently been frequently used in weather studies [91–94]. This diagram was first proposed by Taylor [95], which is based on the geometric relationship between correlation coefficient (R), the standard deviation of time series and root mean square difference ($RMSD$). The latter is calculated as follows [91, 95]:

$$RMSD^2 = \sigma_c^2 + \sigma_o^2 - 2\sigma_c\sigma_oR. \tag{13}$$

$$\sigma_c^2 = \frac{1}{n} \sum_{i=1}^n (Q_{ci} - \overline{Q_c})^2 \tag{14}$$

where σ_c^2 and σ_o^2 are the variance of computed and observed data, respectively.

The Taylor diagram is presented in the form of a half-circle (to show positive and negative correlations) or a quarter circle (to show positive correlations). In either case, values of the correlation coefficient in the form of the radius of the circle, the standard deviation values in the form of co-centric circles around the origin of coordinates and the values of $RMSD$ as co-centric circles around the reference point are plotted. The reference point represents the location of the observed data. To evaluate and compare the performance of models, the location of the models based on the three above-mentioned indices will be plotted on the graph, and any model whose position on the chart is closer to the reference point will have more precision in the forecasting.

The Violin graph is a combination of box and density graphs representing the distribution of data and their probability density. The superiority of this plot than the box diagrams is the form of the data distribution that is drawn as prominence on the left and right of the box plot. It will also be able to identify data clusters, as well as minimum and maximum values [96]. A wider section of the graph in a given value indicates the greater the density of

data at that value, and the smaller width, the less likely the sample will take that value.

4.2 Overall performance of the models

In this section, the ability of the MLP and SBEL models in simulating the daily rainfall–runoff process of Dez Dam watershed is compared. The proposed SBEL and MLP models have been written and evaluated on MATLAB R2014b. In the MLP model, determining the number of hidden layers and the number of neurons in each layer, as well as the number of training epochs, is considered to be an important issue. In this paper, through trial and error, the MLP neural network with one hidden layer containing eight neurons and 176 training epochs has been used. In the hidden and output layers, tangent sigmoid and linear activation functions have been used, respectively. Also, by trial and error, the momentum coefficient and the learning rate of MLP are set as 0.0011 and 0.0017, respectively. On the other hand, due to the absence of the hidden layer in SBEL structure, the challenge of determining the number of hidden neurons does not exist in this model. The performance of SBEL was evaluated using the sinusoidal, Gaussian and max operators as feature expansion functions. No significant difference in the SBEL performance was observed using these functions. Therefore, since the previous studies used the max operator to feature expansion, in this study the max function was selected as well [55, 69, 82, 89]. All three parameters of the SBEL model can also take values in the range [0, 1], which are adjusted to 0.01, 0.01 and 0 by trial and error for α , β and γ , respectively. These values are consistent with the results reported by Lotfi and Akbarzadeh [90]. The linear activation function in SBEL is adopted for both the amygdala output and the final output. In order to evaluate the performance of the MLP and SBEL models, the watershed data have been divided into 70, 15 and 15 percentages for training, cross-validation and validation, respectively.

The distribution of the observed and forecasted river flow by MLP and SBEL models is presented in the Violin plot, as shown in Fig. 5. This figure shows that the most river discharges are in the range from 0 to 200 cms, and the river flow distribution predicted by SBEL is closer to the observed flow distribution. Considering the higher correlation between the observed and predicted data by SBEL, as shown in the scatter plots of Fig. 6, it can be concluded that the overall performance and accuracy of the SBEL in the river flow prediction is higher than the MLP. Also, SBEL can estimate the peak discharges closer to their observations and has higher accuracy than MLP. In table 1, the lower values of $RMSE$ and MRE and higher correlation between observed and predicted discharge values in all three training, cross-validation and validation phases

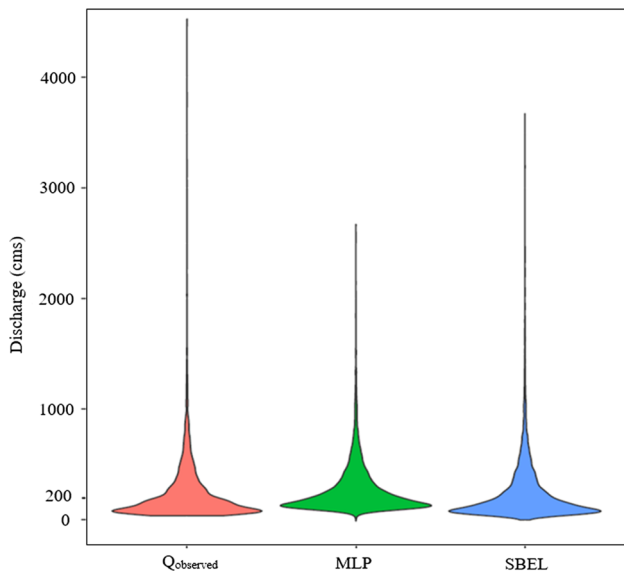


Fig. 5 Violin plot of the observed and predicted runoff time series using MLP and SBEL models (all datasets)

denote more accuracy in SBEL compared to MLP. Also, SBEL has been able to improve the MRE percentage of MLP up to 59.9%, 56.2%, 57.5% and 69.8% for predicting total data, training, cross-validation and validation data, respectively.

4.3 Predicting peak discharges

The accurate prediction of peak values is usually a challenge in data-driven modeling, including rainfall–runoff simulation. Therefore, due to the importance of peak flows in flood management, the peak discharge in each water year has been extracted and the prediction accuracy of peak flow values is investigated using MLP and SBEL models. According to the results shown in Table 2, SBEL has a lower RMSE than MLP and has improved the MRE

percentage than the MLP model up to 21%. On the other hand, considering the scatter plots presented in Fig. 7 and the hydrograph illustrated in Fig. 8, SBEL is more capable of predicting peak flows. If peak discharges can be interpreted as emotional stimuli, it can be deduced that the brain’s emotional learning features mentioned in Sect. 2.2, including the existence of the emotional processor and also imprecise response modifier, have improved the SBEL performance. Therefore, SBEL can be useful in developing flood warning systems with an emphasis on forecasting peak discharges.

4.4 Limiting the training data

In hydrological time series, high flows usually have much less abundance than low flows. Therefore, most data-driven methods require a large and varied set of training data to predict high flow values accurately. As training data are not always available sufficiently in many cases, this is a serious limitation for data-driven models. In ANNs, inadequate training data also lead to a decrease in generalization capability and an increase in forecasting error. In this section of the study, the performance of the MLP and SBEL models has been evaluated by limiting the number of training samples from 70% to 10% of all samples (equivalent to two water years). The observed and predicted runoff hydrographs are presented in Fig. 9. Results show that the SBEL has been more effectively trained and also has forecasted the peak discharge values with more accuracy even under the condition of a limited training dataset. On the other hand, considering Fig. 10, the RMSE and MRE of the SBEL are significantly lower than the MLP for the total data, training, cross-validation and validation data, e.g., the SBEL has been able to improve the MRE by 74.5%. According to Table 3, the predicted runoff volume by the SBEL is closer to the observed runoff volume.

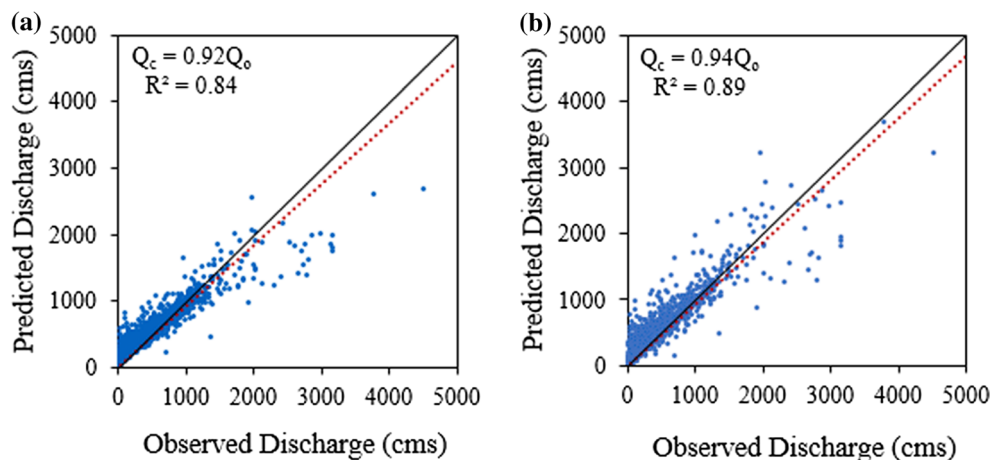


Fig. 6 The scatter plots of the predicted versus the observed runoff time series using **a** MLP and **b** SBEL models (all datasets)

Table 1 Overall performance indicators of MLP and SBEL models

Error criterion	Data							
	All data		Training		Cross-validation		Validation	
	SBEL	MLP	SBEL	MLP	SBEL	MLP	SBEL	MLP
Q_c	0.92 Q_o	0.94 Q_o	0.9 Q_o	0.93 Q_o	Q_o	0.98 Q_o	1.22 Q_o	1.02 Q_o
R^2	0.84	0.89	0.86	0.89	0.77	0.85	0.63	0.72
RMSE	98.1	95	102.6	101.4	93.1	86.5	80.2	68.2
MRE (%)	41.4	16.6	33.3	14.6	49.7	21.1	71	21.4
RMSE improvement by SBEL (%)	+ 3.2		+ 1.2		+ 7.1		+ 15	
MRE improvement by SBEL (%)	+ 59.9		+ 56.2		+ 57.5		+ 69.8	

Table 2 Performance of MLP and SBEL models in peak discharges prediction of each water year

Model	Q_c	R^2	RMSE	MRE (%)	RMSE Improvement by SBEL (%)	MRE Improvement by SBEL (%)
MLP	0.62 Q_o	0.87	844.4	30.9	+ 15.6	+ 21
SBEL	0.71 Q_o	0.73	712.7	24.4		

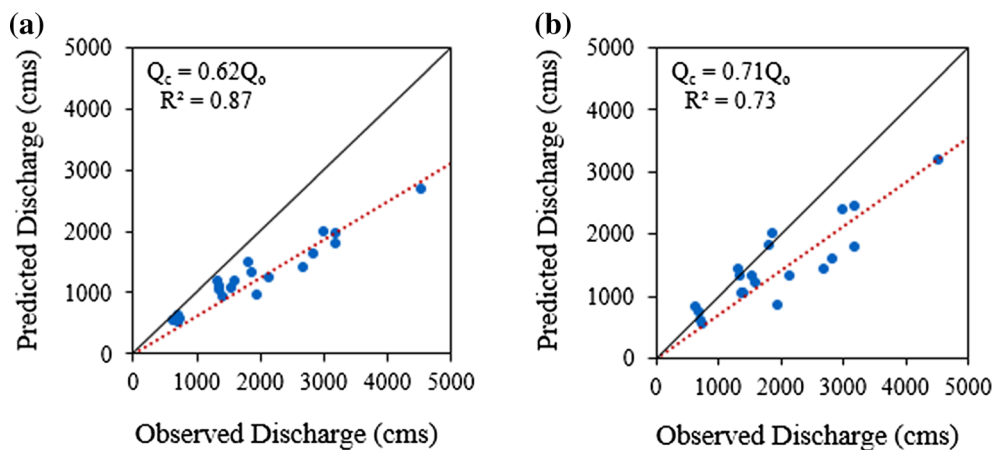


Fig. 7 The scatter plots of the predicted versus the observed peak discharges using **a** MLP and **b** SBEL models

Considering the satisfactory agreement between the observed and predicted runoff resulted by SBEL in the cross-validation and validation phases, it can be concluded that the SBEL model has a higher generalization ability even under conditions of limited data availability.

Moreover, to further compare the performance of both ANNs trained with 10% of the total data, where the training data are not sufficiently varied, three different training scenarios, including training data located in dry, normal and wet periods, separately, are considered. The results presented in Table 4 show that by reducing training data from 70% to 10% and restricting them to a specific scenario, the MLP network could not adequately be trained, and its generalization capability in cross-validation

and validation is reduced compared to SBEL. In contrast, SBEL has been better trained and is able to generalize its knowledge with acceptable performance under different conditions.

The less RMSE and MRE in the SBEL indicate the superiority of this model in the condition of limiting the training data. Also, SBEL has been able to improve the MRE percentage of MLP in predicting total data, training, cross-validation and validation data, respectively, up to 68.2%, 54.2%, 70.8% and 67.7% for the scenario of dry water years, 68.7%, 64.7%, 66.5% and 70.8% for the scenario of normal water years and 76.6%, 65.1%, 72.9% and 75.2% for the scenario of wet water years.

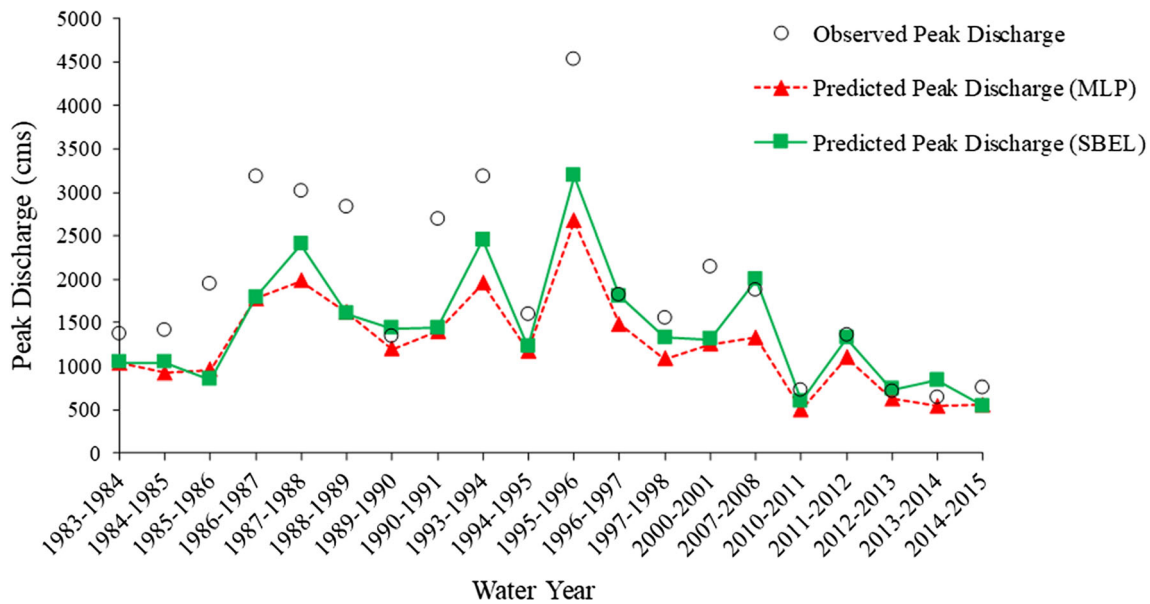


Fig. 8 Observed and predicted peak discharge time series

According to the Taylor diagram in Fig. 11, it can be seen that the performance of both models decreases during the training phase in the normal period than the wet period, as well as the training in the dry period compared to the normal period. However, SBEL is more accurate than MLP in all three scenarios. The results show a higher performance of SBEL versus MLP in the case of limited training data.

4.5 Discharge forecasting in non-rainy events

To evaluate the performance of the models at daily events in which no rainfall has been observed, the days with zero precipitation records have been extracted. Then, the accuracy of the models has been evaluated in terms of forecasting discharge on these events. According to Fig. 12 and Table 5, the correlation between observed and predicted flows in the SBEL is somewhat more than the MLP. Also, the RMSE in the SBEL model is amended, and the MRE is improved by 78.3% compared to the MLP.

4.6 Discharge forecasting in rainy events

River flow forecasting when the basin response is due to rainfall excitation is more complex and challenging than when no precipitation enters the hydrological system. Hence, the performance of models has been evaluated in discharge forecasting of days which include recorded rainfall events in at least one meteorological station. The results in Fig. 13 and Table 6 indicate that the SBEL improved the coefficient of determination and MRE

percentage and was more accurate in discharge forecasting of rainy events.

4.7 Training speed

In most early flood forecasting and warning systems, the training (or calibration) speed of the forecasting model is of paramount importance in system reliability. Through this perspective, a model that is capable of achieving appropriate learning in fewer training epochs is more efficient. Figure 14 illustrates the trend of training error reduction in MLP and SBEL models versus the number of iterations. Considering the vertical dashed lines drawn in the diagram, the error will reach an acceptable value (3% of the previous step error value) after 26 epochs in SBEL and after 49 epochs in MLP models. Therefore, the SBEL model has achieved faster training with fewer iterations.

According to the indices presented in Table 7, it is clear that by reducing the number of training epochs, the performance of both models deteriorates. However, the SBEL is more robust than MLP, confronting the reduction of training epochs. At the same time, the SBEL was able to perform lower RMSE and MRE values than the MLP model in different numbers of iterations. Also, the SBEL was appropriately trained, even with a very low number of iterations, i.e., 10 epochs, and due to its performance in cross-validation and validation periods, it has shown acceptable generalization ability. In the MLP and SBEL models, after 176 and 107 training epochs, respectively, the error reached its minimum value and then remained constant. Therefore, in Table 7, the number of iterations of 176, which is related to the lowest error rate in MLP, is

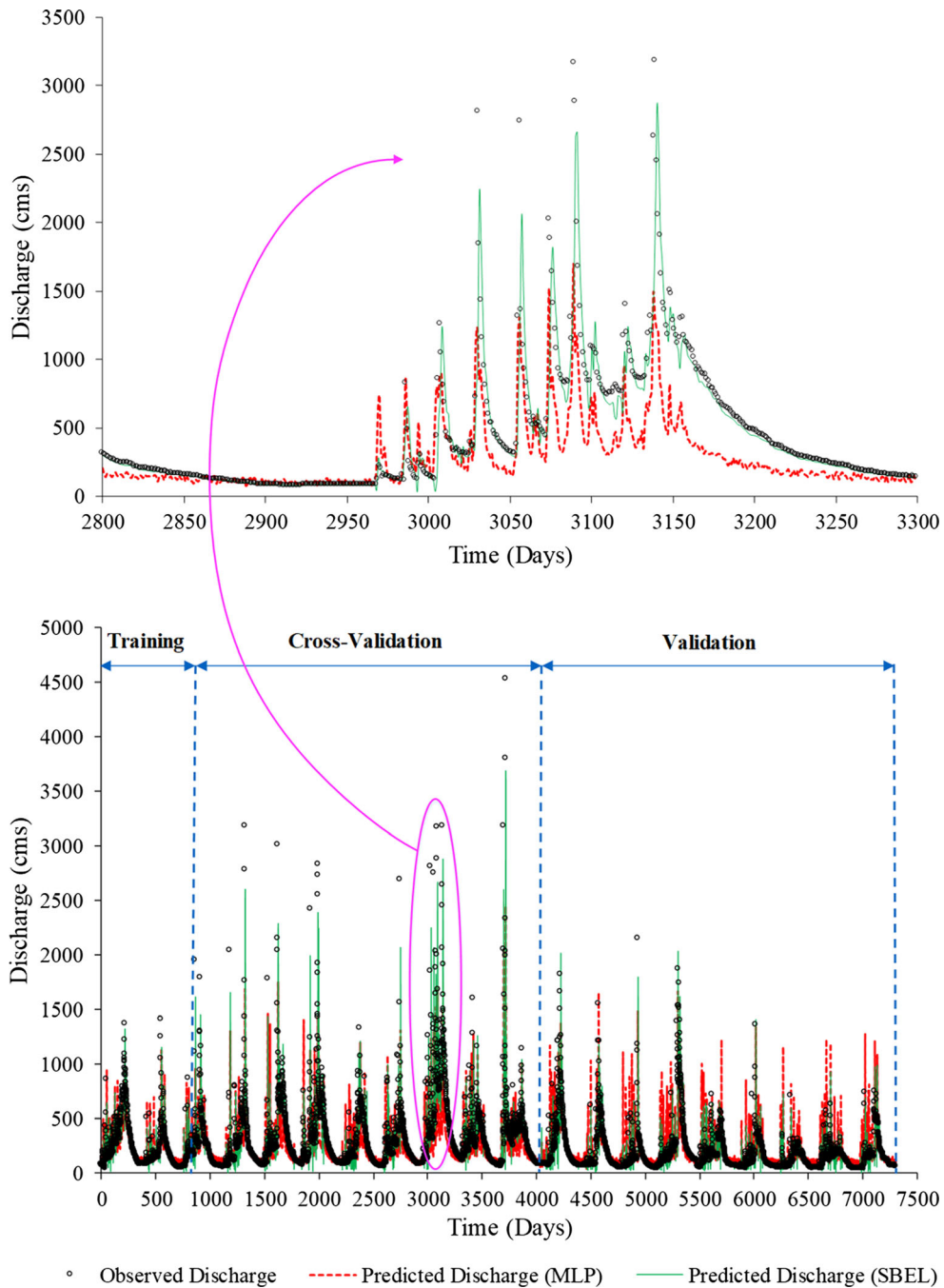


Fig. 9 Observed and predicted hydrographs (trained models by a limited dataset)

considered to start comparing the performance of the models while reducing the training epochs.

Figure 15 demonstrates the RMSE value of both models in four different training epochs. According to this figure, the SBEL performance is nearly constant with decreasing the number of training epochs from 75 to 35, while in the MLP, the error is continuously increasing. It can be concluded that SBEL can learn faster with less number of

iterations, while capable of generalizing its knowledge achieved in the training phase to different conditions with acceptable accuracy. Therefore, the application of SBEL in flood forecasting and warning systems can be of particular interest.

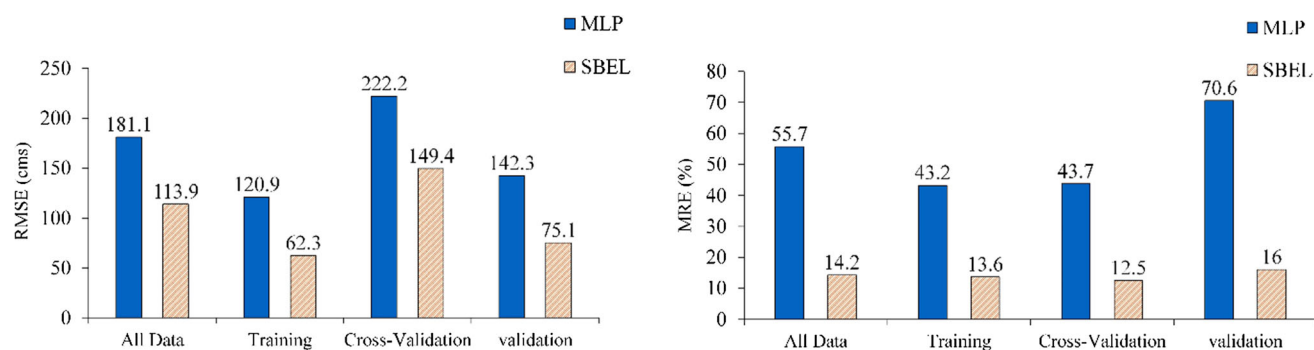


Fig. 10 Performance of MLP and SBEL models in case of the limited training data

Table 3 The observed and predicted runoff volume using the MLP and SBEL models with limited training data

Data	Observed	Predicted	
		MLP	SBEL
Runoff volume (billion m ³)			
All data	156.8	144	153.4
Training	15.6	15.3	15.5
Cross-validation	91.2	71.8	87.6
Validation	50	56.9	50.3

5 Conclusions

A reliable early flood forecasting and warning system (FFWS) is an essential part of non-structural integrated flood management strategies that can significantly reduce flood damages. To improve the performance of the FFWS, features such as prediction accuracy, fast learning, appropriate training even in the absence of sufficient data, reliable prediction of peak flows and accurate forecasts in rainy events are of essential importance. The cognitive model of brain emotional learning taken from the emotional brain has been developed as a center for receiving, processing and creating feelings in the human brain. SBEL has a simple structure and, due to the short paths in its structure, it can make quick responses to input impulses. Determining the number of hidden layers and the number of neurons in each hidden layer is challenging in the topology design and training of the MLP, which is not an issue in the SBEL due to the lack of the hidden layer in its structure.

In this paper, as the first application in hydrology, the SBEL neural network has been applied to the rainfall–runoff simulation of the Dez Dam watershed and has been compared with the MLP. The results show that the overall performance of the SBEL is superior to the MLP and improves the RMSE and MRE of MLP by 3.2% and 59.9%, respectively. Due to the importance of accurate

peak flow prediction in flood forecasting systems, the estimation accuracy of these values has been investigated. The results of peak discharge estimation in each water year show that the SBEL model reduced the RMSE and MRE in MLP by 15.6% and 21%, respectively. Lack of sufficient recorded data to train data-driven models usually results in performance deterioration, especially in high flow predictions. To evaluate the robustness of models encountered with a limited training data, by reducing the number of training samples from 70% to 10% of total data and limiting them to a specific scenario, including placing reduced training data in dry, normal and wet periods, it has been observed that SBEL can generalize its knowledge with acceptable accuracy. In training scenarios located in the dry, normal and wet periods, SBEL improved the RMSE values of the MLP by 38.2%, 29.1% and 25.7%, respectively, and also improved the MRE values of the MLP by 68.2%, 68.7% and 76.6%, respectively. The prediction accuracy of both models is more desirable when training data are placed in the wet water years, than other scenarios. The river flow forecasting is more challenging in rainy events than days without precipitation records. Along with the precise performance of SBEL in predicting river flow in rainy days, this model can reduce the RMSE and MRE of MLP by 0.7% and 14.4%, respectively. On the other hand, fast learning is another essential feature needed in flood forecasting models. By reducing the number of training epochs, the results from networks performance indicated that SBEL could be trained faster with fewer training epochs, without significant changes in model performance. Also, compared to MLP, the SBEL generalization capability is reasonably maintained. For example, at very low iterations, i.e., 10 training epochs, the percentage of RMSE and MRE improvements of the SBEL compared to MLP were 46.6% and 59.4%, respectively.

Notable features of the SBEL neural network used in the present study are the simple structure, fast learning and considerable generalization capability. Also, due to some advantages presented in this case study, such as acceptable learning ability in case of insufficient recorded

Table 4 Performance indicators of the MLP and SBEL models, trained with limited training data in dry, normal and wet scenarios

Training scenario (limited data)	Error criterion	Data							
		MLP	SBEL	MLP	SBEL	MLP	SBEL	MLP	SBEL
Dry water years	Q_c	0.31 Q_o	0.75 Q_o	0.72 Q_o	0.94 Q_o	0.28 Q_o	0.76 Q_o	0.35 Q_o	0.74 Q_o
	R^2	0.43	0.62	0.42	0.53	0.48	0.65	0.45	0.54
	RMSE	279	172.4	85.8	67.1	337.1	191.4	240.3	168.5
	MRE (%)	77.3	24.6	55.9	25.6	77.5	22.6	81.8	26.4
	RMSE Improvement by SBEL (%)	+ 38.2		+ 21.8		+ 43.2		+ 29.9	
	MRE Improvement by SBEL (%)	+ 68.2		+ 54.2		+ 70.8		+ 67.7	
Normal water years	Q_c	0.69 Q_o	0.89 Q_o	0.81 Q_o	0.95 Q_o	0.63 Q_o	0.87 Q_o	0.84 Q_o	0.95 Q_o
	R^2	0.48	0.78	0.54	0.88	0.59	0.75	0.35	0.8
	RMSE	182.7	129.6	123	71.3	224.2	169.8	143.3	85.9
	MRE (%)	54	16.9	43	15.2	43.3	14.5	67.1	19.6
	RMSE Improvement by SBEL (%)	+ 29.1		+ 42		+ 24.3		+ 40.1	
	MRE Improvement by SBEL (%)	+ 68.7		+ 64.7		+ 66.5		+ 70.8	
Wet water years	Q_c	0.98 Q_o	0.98 Q_o	0.9 Q_o	0.95 Q_o	0.94 Q_o	0.97 Q_o	1.16 Q_o	1.05 Q_o
	R^2	0.58	0.84	0.85	0.9	0.59	0.83	0.4	0.79
	RMSE	157.7	117.1	147.3	132	158.6	126.3	159.1	103.2
	MRE (%)	87.8	20.5	37.3	13	66.5	18	99.3	24.6
	RMSE Improvement by SBEL (%)	+ 25.7		+ 10.4		+ 20.4		+ 35.1	
	MRE Improvement by SBEL (%)	+ 76.6		+ 65.1		+ 72.9		+ 75.2	

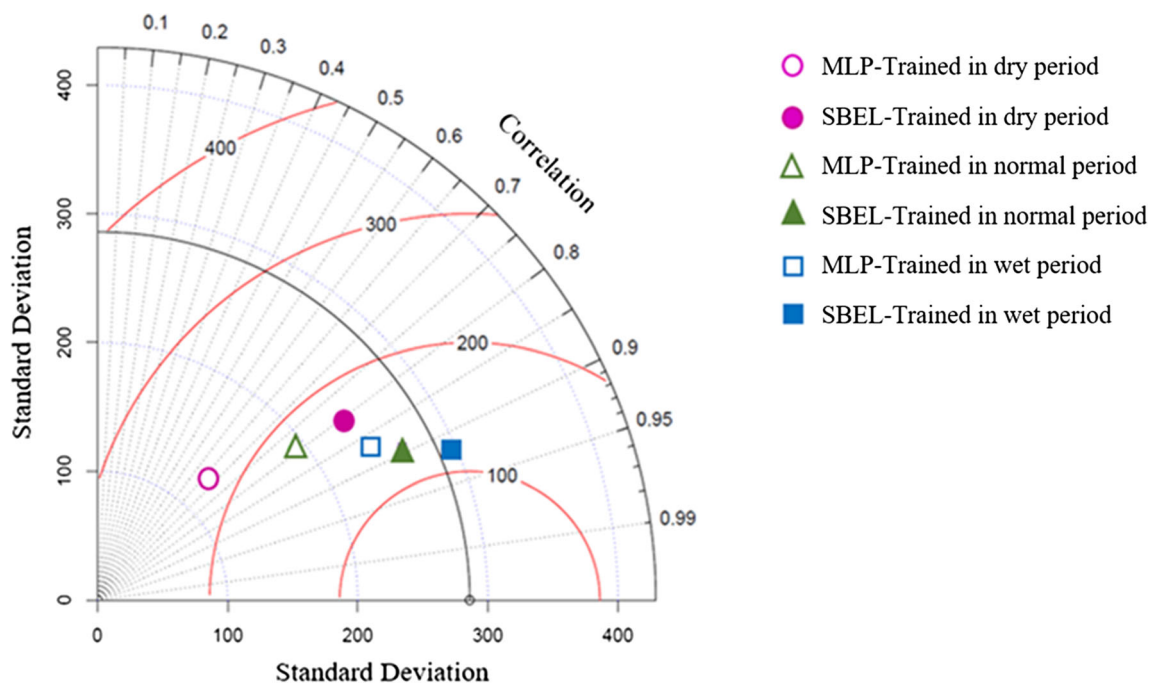


Fig. 11 The Taylor diagram of the MLP and SBEL models trained with limited training data, in dry, normal and wet scenarios

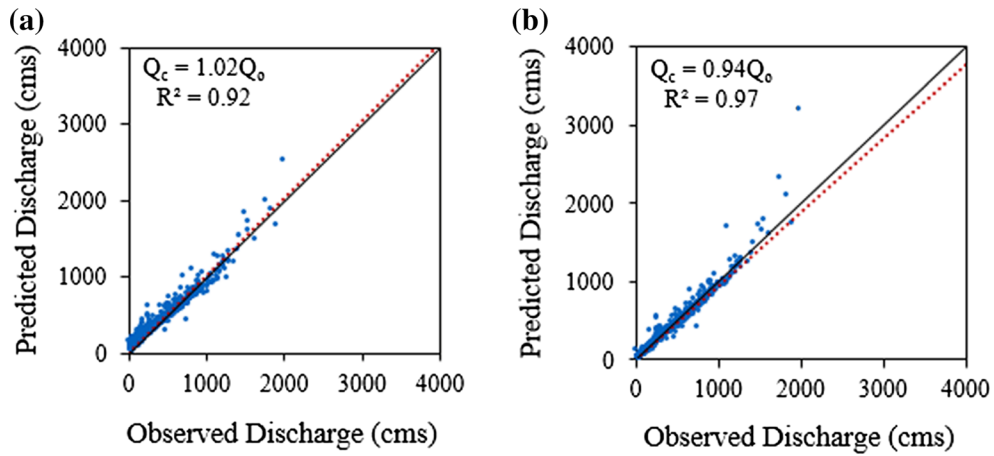


Fig. 12 Scatter plots of predicted versus observed runoff using **a** MLP and **b** SBEL models (non-rainy events)

Table 5 Performance criteria of MLP and SBEL (non-rainy events)

Event	Model	Q_c	R^2	RMSE	MRE (%)	RMSE improvement by SBEL (%)	MRE improvement by SBEL (%)
No rainfall	MLP	$1.02Q_o$	0.92	48.3	42.8	+ 27.7	+ 78.3
	SBEL	$0.94Q_o$	0.97	34.9	9.3		

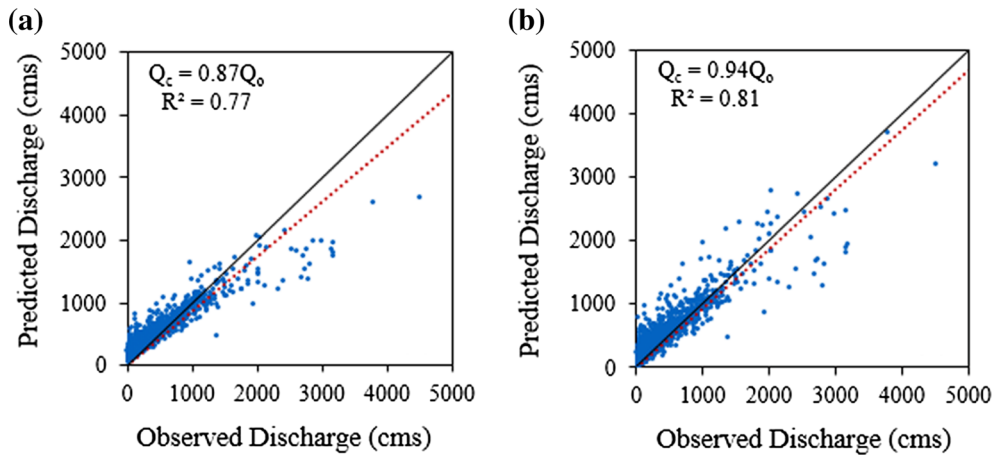


Fig. 13 Scatter plots of the predicted versus the observed runoff using **a** MLP and **b** SBEL models (rainy events)

Table 6 Performance indicators of the MLP and SBEL models in the discharge prediction (rainy days)

Event	Model	Q_c	R^2	RMSE	MRE (%)	RMSE Improvement by SBEL (%)	MRE improvement by SBEL (%)
Rainy days	MLP	$0.87Q_o$	0.77	158.4	50.8	+ 0.7	+ 14.4
	SBEL	$0.94Q_o$	0.81	157.2	43.5		

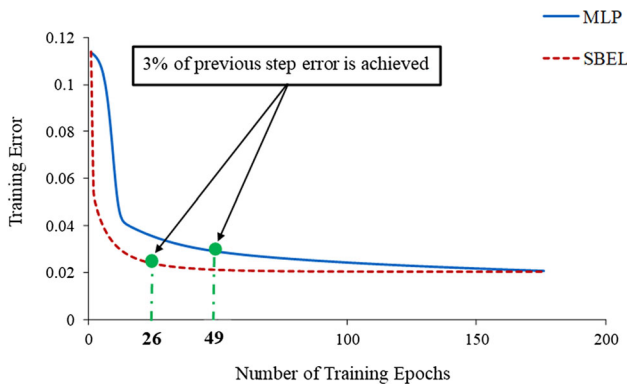


Fig. 14 Training error versus the number of training epochs

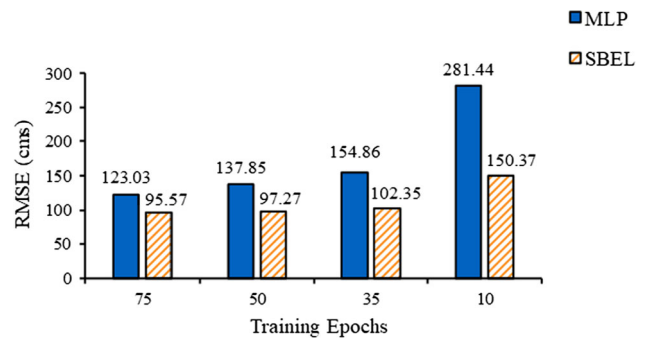


Fig. 15 The performance of the MLP and SBEL models by reducing the number of training epochs (all data)

Table 7 Results of the MLP and SBEL models in different training epochs

Number of epochs	Error criterion	Data							
		All data		Training		Cross-Validation		Validation	
		MLP	SBEL	MLP	SBEL	MLP	SBEL	MLP	SBEL
176	Q_c	0.92 Q_o	0.94 Q_o	0.9 Q_o	0.93 Q_o	Q_o	0.98 Q_o	1.22 Q_o	1.02 Q_o
	R^2	0.84	0.89	0.86	0.89	0.77	0.85	0.63	0.72
	RMSE	98.1	95	102.6	101.4	93.1	86.5	80.2	68.2
	MRE (%)	41.4	16.6	33.3	14.6	49.7	21.1	71	21.4
	RMSE Improvement by SBEL (%)	+ 3.2		+ 1.2		+ 7.1		+ 15	
	MRE improvement by SBEL (%)	+ 59.9		+ 56.1		+ 57.5		+ 69.8	
50	Q_c	0.87 Q_o	0.94 Q_o	0.84 Q_o	0.93 Q_o	0.98 Q_o	0.99 Q_o	1.34 Q_o	1.04 Q_o
	R^2	0.64	0.88	0.7	0.89	0.51	0.84	0.42	0.69
	RMSE	137.9	97.3	143.4	102.9	129.2	90.5	118.6	74
	MRE (%)	62.1	17.5	50.9	15.3	72.7	22.2	91.9	23.1
	RMSE improvement by SBEL (%)	+ 29.4		+ 28.2		+ 30		+ 37.6	
	MRE Improvement by SBEL (%)	+ 71.8		+ 69.9		+ 69.4		+ 74.8	
10	Q_c	0.31 Q_o	0.88 Q_o	0.27 Q_o	0.87 Q_o	0.5 Q_o	0.95 Q_o	0.72 Q_o	1.03 Q_o
	R^2	0.39	0.73	0.55	0.75	0.29	0.67	0.28	0.5
	RMSE	281.4	150.4	320.1	160.7	190.2	137.3	117.6	107
	MRE (%)	77.3	31.4	86.1	29.4	54.7	34.8	58.9	37.1
	RMSE improvement by SBEL (%)	+ 46.6		+ 49.8		+ 27.8		+ 9	
	MRE improvement by SBEL (%)	+ 59.4		+ 65.8		+ 36.4		+ 37	

training data as well as high accuracy in predicting peak flows, the presented brain emotional learning-based neural network can be of particular interest for researchers to model hydrological processes in the future studies.

Funding No funding was received for conducting this study.

Declarations

Conflict of interest The authors declare that they have no conflict of interest.

References

- Hadid B, Duviella E, Lecoeuche S (2020) Data-driven modeling for river flood forecasting based on a piecewise linear ARX system identification. *J Process Control* 86:44–56. <https://doi.org/10.1016/j.jprocont.2019.12.007>
- Pandhiani SM, Sihag P, Shabri AB et al (2020) Time-series prediction of streamflows of Malaysian rivers using data-driven techniques. *J Irrig Drain Eng* 146:4020013. [https://doi.org/10.1061/\(ASCE\)IR.1943-4774.0001463](https://doi.org/10.1061/(ASCE)IR.1943-4774.0001463)
- Maier HR, Kapelan Z, Kasprzyk J et al (2014) Evolutionary algorithms and other metaheuristics in water resources: current status, research challenges and future directions. *Environ Model Softw* 62:271–299. <https://doi.org/10.1016/j.envsoft.2014.09.013>
- Maier HR, Dandy GC (2000) Neural networks for the prediction and forecasting of water resources variables: a review of

- modelling issues and applications. *Environ Model Softw* 15:101–124. [https://doi.org/10.1016/S1364-8152\(99\)00007-9](https://doi.org/10.1016/S1364-8152(99)00007-9)
5. Lipiwattanakarn S, Sriwongsitanon N, Saengsawang S (2004) Improving neural network model performance in runoff estimation by using an antecedent precipitation index. *J Hydrosoci Hydraul Eng* 22:141–154
 6. Clarke RT (1994) *Statistical modelling in hydrology*. Wiley, Chichester, UK
 7. Liu T, Wei H, Zhang C, Zhang K (2017) Time series forecasting based on wavelet decomposition and feature extraction. *Neural Comput Appl* 28:183–195. <https://doi.org/10.1007/s00521-016-2306-8>
 8. Mirabbasi R, Kisi O, Sanikhani H, Meshram SG (2019) Monthly long-term rainfall estimation in Central India using M5Tree, MARS, LSSVR, ANN and GEP models. *Neural Comput Appl* 31:6843–6862. <https://doi.org/10.1007/s00521-018-3519-9>
 9. Yadav B, Ch S, Mathur S, Adamowski J (2016) Discharge forecasting using an online sequential extreme learning machine (OS-ELM) model: a case study in Neckar River, Germany. *Measurement* 92:433–445. <https://doi.org/10.1016/j.measurement.2016.06.042>
 10. Liang Z, Li Y, Hu Y et al (2018) A data-driven SVR model for long-term runoff prediction and uncertainty analysis based on the Bayesian framework. *Theor Appl Climatol* 133:137–149. <https://doi.org/10.1007/s00704-017-2186-6>
 11. Koycegiz C, Buyukyildiz M (2019) Calibration of SWAT and two data-driven models for a data-scarce mountainous headwater in semi-arid Konya closed basin. *Water* 11:147. <https://doi.org/10.3390/w11010147>
 12. He Z, Wen X, Liu H, Du J (2014) A comparative study of artificial neural network, adaptive neuro fuzzy inference system and support vector machine for forecasting river flow in the semiarid mountain region. *J Hydrol* 509:379–386. <https://doi.org/10.1016/j.jhydrol.2013.11.054>
 13. Adamowski J, Sun K (2010) Development of a coupled wavelet transform and neural network method for flow forecasting of non-perennial rivers in semi-arid watersheds. *J Hydrol* 390:85–91. <https://doi.org/10.1016/j.jhydrol.2010.06.033>
 14. Wagena MB, Goering D, Collick AS et al (2020) Comparison of short-term streamflow forecasting using stochastic time series, neural networks, process-based, and Bayesian models. *Environ Model Softw* 126:104669. <https://doi.org/10.1016/j.envsoft.2020.104669>
 15. Chang TK, Talei A, Quek C, Pauwels VRN (2018) Rainfall-runoff modelling using a self-reliant fuzzy inference network with flexible structure. *J Hydrol* 564:1179–1193. <https://doi.org/10.1016/j.jhydrol.2018.07.074>
 16. Yolcu OC, Bas E, Egriglu E, Yolcu U (2020) A new intuitionistic fuzzy functions approach based on hesitation margin for time-series prediction. *Soft Comput* 24:8211–8222. <https://doi.org/10.1007/s00500-019-04432-2>
 17. Danandeh Mehr A, Nourani V, Kahya E et al (2018) Genetic programming in water resources engineering: a state-of-the-art review. *J Hydrol* 566:643–667. <https://doi.org/10.1016/j.jhydrol.2018.09.043>
 18. Feng Z, k., Niu W j., Tang Z y, et al (2020) Monthly runoff time series prediction by variational mode decomposition and support vector machine based on quantum-behaved particle swarm optimization. *J Hydrol* 583:124627. <https://doi.org/10.1016/j.jhydrol.2020.124627>
 19. Kisi O, Cimen M (2011) A wavelet-support vector machine conjunction model for monthly streamflow forecasting. *J Hydrol* 399:132–140. <https://doi.org/10.1016/j.jhydrol.2010.12.041>
 20. Nourani V, Kisi O, Komasi M (2011) Two hybrid artificial intelligence approaches for modeling rainfall–runoff process. *J Hydrol* 402:41–59. <https://doi.org/10.1016/j.jhydrol.2011.03.002>
 21. Sahay RR, Srivastava A (2014) Predicting monsoon floods in rivers embedding wavelet transform, genetic algorithm and neural network. *Water Resour Manag* 28:301–317. <https://doi.org/10.1007/s11269-013-0446-5>
 22. ASCE (2000) Artificial neural networks in hydrology. II: hydrologic applications. *J Hydrol Eng* 5:124
 23. Muñoz P, Orellana-Alvear J, Willems P, Céleri R (2018) Flash-flood forecasting in an Andean mountain catchment—development of a step-wise methodology based on the random forest algorithm. *Water* 10:1519. <https://doi.org/10.3390/w10111519>
 24. Pradhan P, Tingsanchali T, Shrestha S (2020) Evaluation of soil and water assessment tool and artificial neural network models for hydrologic simulation in different climatic regions of Asia. *Sci Total Environ* 701:134308. <https://doi.org/10.1016/j.scitotenv.2019.134308>
 25. Choong CE, Ibrahim S, El-Shafie A (2020) Artificial Neural Network (ANN) model development for predicting just suspension speed in solid-liquid mixing system. *Flow Meas Instrum* 71:101689. <https://doi.org/10.1016/j.flowmeasinst.2019.101689>
 26. Abraham RJ, Anctil F, Coulibaly P et al (2012) Two decades of anarchy? Emerging themes and outstanding challenges for neural network river forecasting. *Prog Phys Geogr* 36:480–513. <https://doi.org/10.1177/0309133312444943>
 27. Haykin S (1999) *Neural networks: a comprehensive foundation*. Prentice Hall
 28. Page GF, Gomm JB, Williams D (1993) *Application of neural networks to modelling and control*. Chapman & Hall New York, NY
 29. Anctil F, Perrin C, Andréassian V (2004) Impact of the length of observed records on the performance of ANN and of conceptual parsimonious rainfall-runoff forecasting models. *Environ Model Softw* 19:357–368. [https://doi.org/10.1016/S1364-8152\(03\)0135-X](https://doi.org/10.1016/S1364-8152(03)0135-X)
 30. Cigizoglu HK (2005) Application of generalized regression neural networks to intermittent flow forecasting and estimation. *J Hydrol Eng* 10:336–341. [https://doi.org/10.1061/\(ASCE\)1084-0699\(2005\)10:4\(336\)](https://doi.org/10.1061/(ASCE)1084-0699(2005)10:4(336))
 31. Danandeh Mehr A, Kahya E, Şahin A, Nazemosadat MJ (2015) Successive-station monthly streamflow prediction using different artificial neural network algorithms. *Int J Environ Sci Technol* 12:2191–2200. <https://doi.org/10.1007/s13762-014-0613-0>
 32. Fernando DAK, Jayawardena AW (1998) Runoff forecasting using RBF networks with OLS algorithm. *J Hydrol Eng* 3:203–209. [https://doi.org/10.1061/\(asce\)1084-0699\(1998\)3:3\(203\)](https://doi.org/10.1061/(asce)1084-0699(1998)3:3(203))
 33. Kagoda PA, Ndiritu J, Ntuli C, Mwaka B (2010) Application of radial basis function neural networks to short-term streamflow forecasting. *Phys Chem Earth, Parts A/B/C* 35:571–581. <https://doi.org/10.1016/j.pce.2010.07.021>
 34. Moradkhani H, Hsu KL, Gupta HV, Sorooshian S (2004) Improved streamflow forecasting using self-organizing radial basis function artificial neural networks. *J Hydrol* 295:246–262. <https://doi.org/10.1016/j.jhydrol.2004.03.027>
 35. Mutlu E, Chaubey I, Hexmoor H, Bajwa SG (2008) Comparison of artificial neural network models for hydrologic predictions at multiple gauging stations in an agricultural watershed. *Hydrol Process An Int J* 22:5097–5106. <https://doi.org/10.1002/hyp.7136>
 36. Zakermoshfegh M, Ghodsian M, Salehi Neishabouri SAA, Shakiba M (2008) River flow forecasting using neural networks and auto-calibrated NAM model with shuffled complex evolution. *J Appl Sci* 8:1487–1494. <https://doi.org/10.3923/jas.2008.1487.1494>
 37. Yaseen ZM, El-Shafie A, Jaafar O et al (2015) Artificial intelligence based models for stream-flow forecasting: 2000–2015.

- J Hydrol 530:829–844. <https://doi.org/10.1016/j.jhydrol.2015.10.038>
38. Aytek A, Alp M (2008) An application of artificial intelligence for rainfall-runoff modeling. *J Earth Syst Sci* 117:145–155. <https://doi.org/10.1007/s12040-008-0005-2>
 39. Partal T (2009) River flow forecasting using different artificial neural network algorithms and wavelet transform. *Can J Civ Eng* 36:26–38. <https://doi.org/10.1139/L08-090>
 40. Quilty J, Adamowski J (2018) Addressing the incorrect usage of wavelet-based hydrological and water resources forecasting models for real-world applications with best practices and a new forecasting framework. *J Hydrol* 563:336–353. <https://doi.org/10.1016/j.jhydrol.2018.05.003>
 41. Quilty J, Adamowski J, Boucher M (2019) A stochastic data-driven ensemble forecasting framework for water resources: a case study using ensemble members derived from a database of deterministic wavelet-based models. *Water Resour Res* 55:175–202. <https://doi.org/10.1029/2018WR023205>
 42. Nourani V (2017) An emotional ANN (EANN) approach to modeling rainfall-runoff process. *J Hydrol* 544:267–277. <https://doi.org/10.1016/j.jhydrol.2016.11.033>
 43. Yaseen ZM, Allawi MF, Yousif AA et al (2018) Non-tuned machine learning approach for hydrological time series forecasting. *Neural Comput Appl* 30:1479–1491. <https://doi.org/10.1007/s00521-016-2763-0>
 44. Yaseen ZM, Awadh SM, Sharafati A, Shahid S (2018) Complementary data-intelligence model for river flow simulation. *J Hydrol* 567:180–190. <https://doi.org/10.1016/j.jhydrol.2018.10.020>
 45. Araghinejad S, Fayaz N, Hosseini-Moghari SM (2018) Development of a hybrid data driven model for hydrological estimation. *Water Resour Manag* 32:3737–3750. <https://doi.org/10.1007/s11269-018-2016-3>
 46. Darbandi S, Pourhosseini FA (2018) River flow simulation using a multilayer perceptron-firefly algorithm model. *Appl Water Sci* 8:85. <https://doi.org/10.1007/s13201-018-0713-y>
 47. Ni L, Wang D, Singh VP et al (2019) Streamflow and rainfall forecasting by two long short-term memory-based models. *J Hydrol* 124296
 48. Yu X, Yang J, Xie Z (2014) Training SVMs on a bound vectors set based on Fisher projection. *Front Comput Sci* 8:793–806
 49. Yu X, Chu Y, Jiang F et al (2018) SVMs classification based two-side cross domain collaborative filtering by inferring intrinsic user and item features. *Knowledge-Based Syst* 141:80–91
 50. Yu X, Jiang F, Du J, Gong D (2019) A cross-domain collaborative filtering algorithm with expanding user and item features via the latent factor space of auxiliary domains. *Pattern Recognit* 94:96–109
 51. Brent RP (1991) Fast training algorithms for multilayer neural nets. *IEEE Trans Neural Networks* 2:346–354. <https://doi.org/10.1109/72.97911>
 52. Gori M, Tesi A (1992) On the problem of local minima in backpropagation. *IEEE Trans Pattern Anal Mach Intell* 14:76–86. <https://doi.org/10.1109/34.107014>
 53. Sudheer KP, Nayak PC, Ramasastri KS (2003) Improving peak flow estimates in artificial neural network river flow models. *Hydrol Process* 17:677–686. <https://doi.org/10.1002/hyp.5103>
 54. Bossley KM (1997) Neurofuzzy modelling approaches in system identification. University of Southampton
 55. Parsapoor M, Bilstrup U (2013) Chaotic time series prediction using brain emotional learning based recurrent fuzzy system (BELRFS). *Int J Reason Intell Syst* 5:113–126. <https://doi.org/10.1504/IJRIS.2013.057273>
 56. Kalth AM (2013) Monthly river flow forecasting using artificial neural network and support vector regression models coupled with wavelet transform. *Comput Geosci* 54:1–8. <https://doi.org/10.1016/j.cageo.2012.11.015>
 57. Nourani V, Baghanam AH, Adamowski J, Kisi O (2014) Applications of hybrid wavelet-artificial intelligence models in hydrology: a review. *J Hydrol* 514:358–377. <https://doi.org/10.1016/j.jhydrol.2014.03.057>
 58. Shoaib M, Shamseldin AY, Melville BW, Khan MM (2016) A comparison between wavelet based static and dynamic neural network approaches for runoff prediction. *J Hydrol* 535:211–225. <https://doi.org/10.1016/j.jhydrol.2016.01.076>
 59. Babaie T, Karimizandi R, Lucas C (2008) Learning based brain emotional intelligence as a new aspect for development of an alarm system. *Soft Comput* 12:857–873. <https://doi.org/10.1007/s00500-007-0258-8>
 60. Khashman A (2008) A modified backpropagation learning algorithm with added emotional coefficients. *IEEE Trans Neural Netw* 19:1896–1909. <https://doi.org/10.1109/TNN.2008.2002913>
 61. Morén J (2002) Emotion and learning: A computational model of the amygdala. Ph.D. Thesis, Lund Univ Lund, Sweden
 62. Lotfi E, Khosravi A, Akbarzadeh-T MR, Nahavandi S (2014) Wind power forecasting using emotional neural networks. In: 2014 IEEE international conference on systems, man, and cybernetics (SMC). IEEE, pp 311–316
 63. Goleman D (1996) Emotional intelligence: why it can matter more than IQ. Bloomsbury Publishing
 64. LeDoux JE (1991) Emotion and the limbic system concept. *Concepts Neurosci* 2:169–199
 65. Abdi J, Moshiri B, Abdulhai B, Sedigh AK (2012) Forecasting of short-term traffic-flow based on improved neurofuzzy models via emotional temporal difference learning algorithm. *Eng Appl Artif Intell* 25:1022–1042
 66. Lucas C, Abbaspour A, Gholipour A et al (2003) Enhancing the performance of neurofuzzy predictors by emotional learning algorithm. *INFORMATICA-LJUBLJANA* 27:137–146
 67. Parsapoor M, Lucas C, Setayeshi S (2008) Reinforcement_recurrent fuzzy rule based system based on brain emotional learning structure to predict the complexity dynamic system. In: 2008 third international conference on digital information management. IEEE, pp 25–32
 68. Parsapoor M, Bilstrup U (2012) Brain emotional learning based fuzzy inference system (BELFIS) for solar activity forecasting. In: 2012 IEEE 24th international conference on tools with artificial intelligence. IEEE, pp 532–539
 69. Parsapoor M, Bilstrup U (2012) Neuro-fuzzy models, BELRFS and LoLiMoT, for prediction of chaotic time series. In: 2012 international symposium on innovations in intelligent systems and applications. IEEE, pp 1–5
 70. Pavlos GP, Iliopoulos AC, Tsoutsouras VG et al (2011) First and second order non-equilibrium phase transition and evidence for non-extensive Tsallis statistics in Earth's magnetosphere. *Phys A Stat Mech its Appl* 390:2819–2839. <https://doi.org/10.1016/j.physa.2011.03.005>
 71. Gholipour A, Lucas C, Shahmirzadi D (2004) Purposeful prediction of space weather phenomena by simulated emotional learning. *Int J Model Simul* 24:65–72. <https://doi.org/10.1080/02286203.2004.11442288>
 72. Sheikholeslami N, Shahmirzadi D, Semsar E et al (2006) Applying brain emotional learning algorithm for multivariable control of HVAC systems. *J Intell Fuzzy Syst* 17:35–46
 73. Mehrabian AR, Lucas C, Roshanian J (2006) Aerospace launch vehicle control: an intelligent adaptive approach. *Aerosp Sci Technol* 10:149–155. <https://doi.org/10.1016/j.ast.2005.11.002>
 74. Milasi RM, Lucas C, Araabi BN (2005) Intelligent modeling and control of washing machine using LLNF modeling and modified BELBIC. In: 2005 international conference on control and automation. IEEE, pp 812–817
 75. Liu L, Moayedi H, Rashid ASA et al (2020) Optimizing an ANN model with genetic algorithm (GA) predicting load-settlement

- behaviours of eco-friendly raft-pile foundation (ERP) system. *Eng Comput* 36:421–433. <https://doi.org/10.1007/s00366-019-00767-4>
76. Hornik K, Stinchcombe M, White H (1989) Multilayer feedforward networks are universal approximators. *Neural Netw* 2:359–366. [https://doi.org/10.1016/0893-6080\(89\)90020-8](https://doi.org/10.1016/0893-6080(89)90020-8)
 77. Beheshti Z, Firouzi M, Shamsuddin SM et al (2016) A new rainfall forecasting model using the CAPSO algorithm and an artificial neural network. *Neural Comput Appl* 27:2551–2565. <https://doi.org/10.1007/s00521-015-2024-7>
 78. Badrzadeh H, Sarukkalige R, Jayawardena AW (2015) Hourly runoff forecasting for flood risk management: Application of various computational intelligence models. *J Hydrol* 529:1633–1643. <https://doi.org/10.1016/j.jhydrol.2015.07.057>
 79. Yaseen ZM, El-Shafie A, Afan HA et al (2016) RBFNN versus FFNN for daily river flow forecasting at Johor River, Malaysia. *Neural Comput Appl* 27:1533–1542. <https://doi.org/10.1007/s00521-015-1952-6>
 80. Brown M, Harris CJ (1995) A perspective and critique of adaptive neurofuzzy systems used for modelling and control applications. *Int J Neural Syst* 6:197–220. <https://doi.org/10.1142/S0129065795000159>
 81. Rumelhart DE, Hinton GE, Williams RJ (1986) Learning representations by back-propagating errors. *Nature* 323:533–536. <https://doi.org/10.1016/j.measurement.2017.09.025>
 82. Lotfi E, Akbarzadeh-T MR (2013) Brain emotional learning-based pattern recognizer. *Cybern Syst* 44:402–421. <https://doi.org/10.1080/01969722.2013.789652>
 83. LeDoux JE (1996) The emotional brain. Simon & Schuster
 84. Morén J, Balkenius C (2000) A computational model of emotional learning in the amygdala. *From Anim to Animat* 6:115–124
 85. LeDoux JE (2000) Emotion circuits in the brain. *Annu Rev Neurosci* 23:155–184. https://doi.org/10.1007/978-981-13-3320-0_10
 86. Phelps EA, LeDoux JE (2005) Contributions of the amygdala to emotion processing: from animal models to human behavior. *Neuron* 48:175–187. <https://doi.org/10.1016/j.neuron.2005.09.025>
 87. Balkenius C, Morén J (2001) Emotional learning: a computational model of the amygdala. *Cybern Syst* 32:611–636. <https://doi.org/10.1080/01969720118947>
 88. Elliott R, Dolan RJ, Frith CD (2000) Dissociable functions in the medial and lateral orbitofrontal cortex: evidence from human neuroimaging studies. *Cereb cortex* 10:308–317. <https://doi.org/10.1093/cercor/10.3.308>
 89. Lotfi E, Akbarzadeh TMR (2012) Supervised brain emotional learning. In: The 2012 international joint conference on neural networks (IJCNN). IEEE, pp 1–6
 90. Lotfi E, Akbarzadeh-T MR (2014) Practical emotional neural networks. *Neural Netw* 59:61–72. <https://doi.org/10.1016/j.neunet.2014.06.012>
 91. Lo Conti F, Hsu K-L, Noto LV, Sorooshian S (2014) Evaluation and comparison of satellite precipitation estimates with reference to a local area in the Mediterranean Sea. *Atmos Res* 138:189–204. <https://doi.org/10.1016/j.atmosres.2013.11.011>
 92. Griggs DJ, Noguer M (2002) Climate change 2001: the scientific basis. Contribution of working group I to the third assessment report of the intergovernmental panel on climate change. *Weather* 57:267–269. <https://doi.org/10.1256/004316502320517344>
 93. Pincus R, Batstone CP, Hofmann RJP et al (2008) Evaluating the present-day simulation of clouds, precipitation, and radiation in climate models. *J Geophys Res Atmos* 113:1–10. <https://doi.org/10.1029/2007JD009334>
 94. Wehner MF (2013) Very extreme seasonal precipitation in the NARCCAP ensemble: model performance and projections. *Clim Dyn* 40:59–80. <https://doi.org/10.1007/s00382-012-1393-1>
 95. Taylor KE (2001) Summarizing multiple aspects of model performance in a single diagram. *J Geophys Res Atmos* 106:7183–7192. <https://doi.org/10.1029/2000JD900719>
 96. Hintze JL, Nelson RD (1998) Violin plots: a box plot-density trace synergism. *Am Stat* 52:181–184. <https://doi.org/10.1080/00031305.1998.10480559>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.