Natural Hazards (2023) 117:681–701
https://doi.org/10.1007/s11069-023-05877-3

ORIGINAL PAPER

Streamflow forecasting using a hybrid LSTM-PSO approach: the case of Seyhan Basin

Bulent Haznedar[1] · Huseyin Cagan Kilinc[2] · Furkan Ozkan[3] · Adem Yurtsever[4,5]

Received: 7 October 2022 / Accepted: 7 February 2023 / Published online: 24 February 2023

Abstract
The conditions which affect the sustainability of water cause a number of serious environmental and hydrological problems. Effective and correct management of water resources constitutes an effective and important issue among scales. In this sense, a precise estimation of streamflow time series in rivers is one of the most important issues in optimal management of surface water resources. Therefore, a hybrid method combining particle swarm algorithm (PSO) and long short-term memory networks (LSTM) are proposed to predict flow with data obtained from different flow measurement stations. In this respect, the data gathered from three Flow Measurement Stations (FMS) from Zamanti and Eğlence rivers located on Seyhan Basin are utilized. Besides, the proposed LSTM-PSO method is compared to an adaptive neuro-fuzzy inference system (ANFIS) and the LSTM benchmark model to demonstrate the performance achievement of proposed method. The prediction performances of the developed hybrid model and the others are tested on the determined stations. The forecasting performances of the models are determined with RMSE, MAE, MAPE, SD, and R^2 metrics. The comparison results indicated that the LSTM-PSO method provides highest results with values of R^2 (≈ 0.9433), R^2 (≈ 0.6972), and R^2 (≈ 0.9273) for the Değirmenocağı, Eğribük, and Ergenusagi FMS data, respectively.

Keywords Forecasting · Streamflow · ANFIS · LSTM · PSO · Time Series

✉ Bulent Haznedar
haznedar@gaziantep.edu.tr

Huseyin Cagan Kilinc
huseyincagankilinc@aydin.edu.tr

Furkan Ozkan
furkan.ozkan@hku.edu.tr

Adem Yurtsever
adem.yurtsever@iuc.edu.tr

1 Department of Computer Engineering, Gaziantep University, Gaziantep, Turkey

2 Department of Civil Engineering, Istanbul Aydın University, Istanbul, Turkey

3 Department of Computer Engineering, Hasan Kalyoncu University, Gaziantep, Turkey

4 Department of Environmental Engineering, İstanbul University-Cerrahpaşa, Istanbul, Turkey

5 Environmental Research and Application Center, Hasan Kalyoncu University, Gaziantep, Turkey

1 Introduction

The continuity of the balance of all living species on earth with each other is sustained with the help of water. Civilizations that have lived throughout history have established their presence near water resources and utilized these resources in accordance with the conditions of the era. As long as water resources persist its existence, people will be able to continue their agricultural, industrial, social and industrial activities. In previous years, the rapidly increasing population, expansion of agriculture, climate change, energy and industrial activities conditions impacted the water resources. In addition, the circumstances caused a decrease in the amount of water (Karahan 2021). These conditions, which affect the sustainability of water, cause a number of serious environmental and hydrological problems. Effective and correct management of water resources constitutes an effective and important issue among scales.

Natural disasters such as floods, droughts and storms are among the most influential environmental events of recent years. The conditions impacting the environment are of great matter to researchers. A precise estimation of streamflow time series in rivers is one of the most critical issues in the optimal management of surface water resources; in particular, making appropriate decisions when coping with hydrological data such as floods and droughts (Kilinc and Haznedar 2022). Flow forecasting can be considered in two various groups. The first one which is short-term time series contains minutely, hourly, and daily time periods commonly used in flood management systems. The second one which is long-term forecasting generally includes weekly, monthly and annual flow data which can be widely applied in reservoir operation, and hydroelectric generation systems (Wegayehu and Muluneh 2022). In general, there are two distinct approaches for predicting streamflows; conceptual (physical) and machine learning (data-driven) models (Yaseen et al. 2019; Nourani et al. 2014).

Artificial intelligence methodologies such as fuzzy logic and neural network have begun to gain popularity recently. In studies, hydrological data were used as input parameters (flow, temperature, precipitation, evaporation, etc.) and predictions were made with multiple models. The solution sought to be achieved with this popularity is called data-driven since it is learned or driven directly from the data without assuming a predetermined equation as a model (Cao et al. 2016). Considering the river flow predictions, the data-driven solution rather than the model-driven solution turned out to be more applicable and caught our attention. Given its data-driven appetite and non-reliance on governing physics, machine learning becomes a natural choice for solving these problems. Data-driven modeling assumes the existence of a substantial and sufficient amount of data describing the underlying system. The data are mainly used to perform classification, pattern recognition, relational and predictive analysis tasks (Solomatine and Ostfeld 2008).

The increasing amount of data and, accordingly, the growth of computer systems that can process the data quickly and more accurately than humans allow significant advances in the field of artificial intelligence. The fact that it exhibits high cognitive functions or autonomous behaviors specific to human intelligence. For instance, perception, learning, connecting plural concepts, thinking, reasoning, problem-solving, communication, inference, and decision making have heightened the interest in artificial intelligence models.

Expert systems, fuzzy logic, artificial neural networks, machine learning, and genetic algorithms constitute artificial intelligence technologies. Artificial intelligence (AI) is directly related to machine learning (ML). Nowadays machine learning is a method that researchers frequently operate to explore patterns from datasets by allowing them to learn on their own. Moreover, with the development of large-scale computation techniques, data learning gets more straightforward. These data are very essential for the prediction performance of the model and to ensure that it can be generalized to unknown data. Deep learning models, on the other hand, represent a new form of learning in AI and ML (Yağın 2022; Türkşen and Akgün 2018; Chen et al. 2020).

Deep learning is an artificial intelligence method that employs multilayered artificial neural networks in the fields of image, sound, or text processing. Moreover, deep learning is a subclass of machine learning. It is a learning process of the resulting model obtained by passing the sample data taken as input through multilayer neural networks. It consists of some input, hidden, and output layers. Likewise, it is quite good at discovering complex structures in high-dimensional data and thus can be applied to numerous fields of science (Tüzün 2022; Dolezal et al. 2021).

When the studies in the literature are examined, it has been observed that compelling results are obtained in the field of hydrology (rainfall-runoff models, temperature models, humidity and evaporation forecasting). In addition, autoregressive integrated moving average (ARIMA), artificial neural network (ANN), convolutional neural network (CNN), recurrent neural network (RNN) models, which are classical prediction methods in the literature, and long short-term memory (LSTM), gated recurrent unit (GRU), bounded Boltzmann machine (RBM), deep belief network (DBN), genetic algorithm (GA) and fuzzy logic adaptive network, which are prevalent among deep learning models. There are numerous modeling studies involving based those models (Yamashita et al. 2018; Cheng et al. 2022; Ye et al. 2022).

RNN, which is a type of artificial neural network, depends on the result of calculating the sequential, time-series inputs of the output data of the neural network. In RNN training; backpropagation is required as in conventional neural networks. In this case, a prominent concern came out by the RNN arises. The gradient is the value that authorizes adjusting all the weights. Since the neural network data are utilized in all layers, the gradient in each output relies not only on the current layer but also on the previous layer. If the backpropagation process is constantly renewed in more than one-time interval, the outcome gets more smallish and the disappearing gradient problem arises. In the same situation, if the gradients are greater than one, the result becomes more extensive and the exploding gradient problem arises (Yu et al. 2022; Zhang et al. 2018; Feng et al. 2017).

On the other hand, the LSTM model is a special type of Recurrent Neural Networks (RNN) that can learn long-term associations (recurring values, sequence, etc.) between data. LSTMs also have chain-like structures as RNNs; however, the structure of the repeating module is different. LSTMs are remarkably convenient for classifying, predicting, and processing time series, regarding temporal delays. The relative insensitivity to its temporal length provides LSTM an advantage over alternative RNNs and other learning methods in different applications. LSTMs recall data for a long time and pursue the process through learning (Nazimi 2021).

The optimization of LSTM parameters greatly determines the ultimate effect of network. Gradient descent-based optimizers are commonly used to optimize parameters of the current LSTM network. However, there exists a problem in all above optimizers, that the iterative process may fall into a local optimum and cannot achieve the global optimum, which will affect the effect of model. So, different approaches are required for optimizing LSTM successfully (Jang 1993; İpek 2021; Durgut and Aydın 2021; Haznedar et al. 2021; Oyelade et al. 2022). PSO can be given as an example as the most frequently used swarm-based optimization algorithm in the field of water resources and hydrology. PSO algorithm is an intelligence optimization theory and one of the brightest optimization approaches because it has advantages over other optimization methods in terms of ease of implementation, sphere convergence ability and robustness. The algorithm was developed by observing the normal behavior of the presented fish groups and birds to reduce the optimization problems. In this technique, a particle represents each member in the swarm, and each particle has a velocity and position vector. These particles behave in the same way: Each particle is relocated to the optimum location in the swarm whose value is closest to the target. Recently, numerous model studies hybridizing with metaheuristic algorithms take part in the literature. The impacts of mentioned models on river flows and other hydrological parameters were studied by researchers.

In this paper, based on multivariate correlations among time-series characteristics and streamflow information, the parameters of LSTM network are optimized using PSO algorithm so as to improve the prediction performance. The proposed LSTM-PSO model performance was compared with the performance of LSTM and ANFIS models to predict flow with data obtained from various flow observation stations. The prediction performances of the models were tested on the determined station. In this way, it was sought to obtain a long correlation in a short time. In other respects, the LSTM-PSO method was conducted limitedly in real-time river flow prediction. In addition, a limited number of studies are found for the prediction of river flow in the Seyhan basin in the literature that show the originality of this study. In the following section, the datasets are introduced and methods are explained. In the third and fourth sections, the results are presented and discussed in detail.

2 Literature review

Xu et al. (2022) generated a method to enhance the flood forecasting model. LSTM networks and particle swarm optimization (PSO) were hybridized to the estimation of the flood data. The hybrid model was compared with ANN, PSO-ANN, and the benchmark model (LSTM). The outcomes displayed that the PSO-LSTM hybrid model significantly improved the benchmark model. Kilinc and Yurtsever (2022) designed a hybrid model based on the GWO algorithm. The daily streamflow data were tested with GRU (Gated Recurrent Unit). The results revealed that GWO algorithms augmented the performance of the hybrid model. Siva Kumar et al. (2021) employed a hybrid model linked with a genetic algorithm (GA) and an adaptive neuro-fuzzy inference system (ANFIS). The results uncovered that the GA algorithm promisingly increased the performance of the ANFIS model. Likewise, Dalkiliç and Hashimi (2020) operated the artificial neural network (ANN), wavelet neural network (WNN), and adaptive neuro-fuzzy inference system (ANFIS) models to forecast the daily streamflow. The outcomes demonstrated that the WNN model indicated the most satisfactory performance compared with the ANN and ANFIS model. Zhang et al. (2014) built a hybrid model grounded on particle swarm optimization and support vector machine model for streamflow forecasting. The test results revealed that the hybrid model enhanced the single SVM model.

Kushwaha et al. (2022) reported a novel prediction model established on the ANFIS and GA. According to the analysis, the GA-ANFIS model depicted a greater prediction capability than the ANFIS model. Kayhomayoon et al. (2022) conducted adaptive neuro-fuzzy inference systems (ANFIS) to predict the groundwater level. Furthermore, several metaheuristic algorithms were operated to improve the model. The findings uncovered that metaheuristic algorithms enhanced the performance of the ANFIS model. When the studies in the literature were reviewed, prosperous results of metaheuristic algorithms hybridizing with deep learning models in river flows and other areas have been observed. Asaad et al. (2022) evaluated the performance of LSTM, multilayer perceptron (MLP) and ANFIS for long-term streamflow forecasting. The results revealed that the LSTM model had a better prediction performance, surpassing the MLP and ANFIS models. Adaryani et al. (2022) compared the performances of three machine and deep learning-based rainfall forecasting approaches including a hybrid optimized-by-PSO support vector regression (PSO-SVR), long short-term memory (LSTM), and convolutional neural network (CNN). The findings showed that hybrid PSO-SVR model was improved the accuracy of the forecasting. Du et al. (2022) combined LSTM and PSO algorithm in order to predict water demand. The results showed that PSO algorithm improved the performance of the LSTM model. Song et al. (2020) generated hybrid LSTM-PSO model to time-series prediction. The hybrid model was compared with traditional machine learning models and accuracy predictions of the models were observed. The findings revealed that proposed model outperformed other approaches.

Kim and Cho (2021) proposed a method of optimizing CNN-LSTM neural networks with PSO algorithm. PSO-based hybrid model was outperformed other deep learning and machine learning models. Pranolo et al. (2022) aimed to hybridize LSTM with PSO and Bifold-Attention mechanism. The created model was compared with classical machine learning models. The proposed model comparison was based on the accuracy of each model in forecasting multivariate time series. The findings revealed that the proposed model outperformed other approaches. This ground-breaking innovation was valuable for time-series analysis research, particularly the implementation of deep learning for time-series forecasting. Sheikh Khozani et al. (2022) introduced new hybrid model, namely ARIMA-LSTM neural network, to forecast the ground water level time series. In order to determine the hyperparameters of the LSTM algorithm, PSO algorithm and other three algorithm were coupled with the LSTM model. The results indicated that all hybrid models were increased the performance of the LSTM model. Wang et al. (2022) compared five popular machine learning methods, including PSO algorithm and LSTM models. Each model was hybridized with each other. The results showed that LSTM- and PSO-based hybrid models were achieved better single prediction accuracy. Lv et al. (2018) improved LSTM-based PSO algorithm, which was applied to predict time-series data. Compared with typical algorithms, the findings showed that LSTM has better performance in reliability and adaptability, and hybrid PSO-LSTM model had better accuracy.

3 Materials and methods

3.1 Location and characterization of the study area

The Seyhan Basin extends from Çukurova to the north, and its upper part is located in the Central Anatolian region, while its middle and lower parts are located in the Mediterranean Region. It is located in the north of Adana Province, which is located in the Eastern Mediterranean Region of Turkey, between 36° 30′ and 39° 15′ north latitudes and 34° 45′ and 37° 00′ east longitudes and has a surface area of 22.042 km^2 (Topaloğlu 1999). The most important streams of the Seyhan Basin are the Seyhan River, which names to the basin, and its tributaries. The Zamanti River is one of the important tributaries of Seyhan River. Zamanti River, which originates from the Kayseri highlands, has a length of 308 km and an average annual flow of 65 m^3/s. In this study, a total of three stations, two stations belonging to Zamanti River and one station belonging to the Eğlence River located in Adana Province, were studied. Eğlence River is approximately 29 km long and forms the Seyhan River with the Körkun Stream, which originates from the north of Karaisalı. As a result of climate change in Turkey, the decrease in precipitation and drought is a natural occurrence (Komuscu et al. 1998). The Seyhan Basin, which is one of the basins where the effects of this change are observed, has a structure whose hydrological form can deteriorate when it undergoes climate change with its climate type characteristics. For this reason, changes in the amount of precipitation in the basin also affect surface water resources (Daba and You 2020). Due to the large number of karst structures of the basin, underground drainage rather than surface drainage predominates in the hydrological basin. The predominance of underground drainage causes the basin to acquire an arid appearance. This makes it important to determine the hydrological status of the basin. Since these features play an important role in water resources management and planning, Seyhan Basin has been considered as a study area.

3.2 Datasets and pre-processing

In order to regulate the prediction potential of the compared models, three flow measurement stations (FMS) bearing various hydrological conditions of the Seyhan Basin were employed. In addition, long-term, 24-year stream flow data are obtained daily from Eğlence River-Eğribük (E18A025), Zamanti River-Ergenuşağı (E18A026) and Zamanti River-Değirmenocağı (E18A027) stations. The positions of the related stations in the Seyhan Basin on the Türkiye map are presented in Fig. 1. The twenty-four years of flow data at

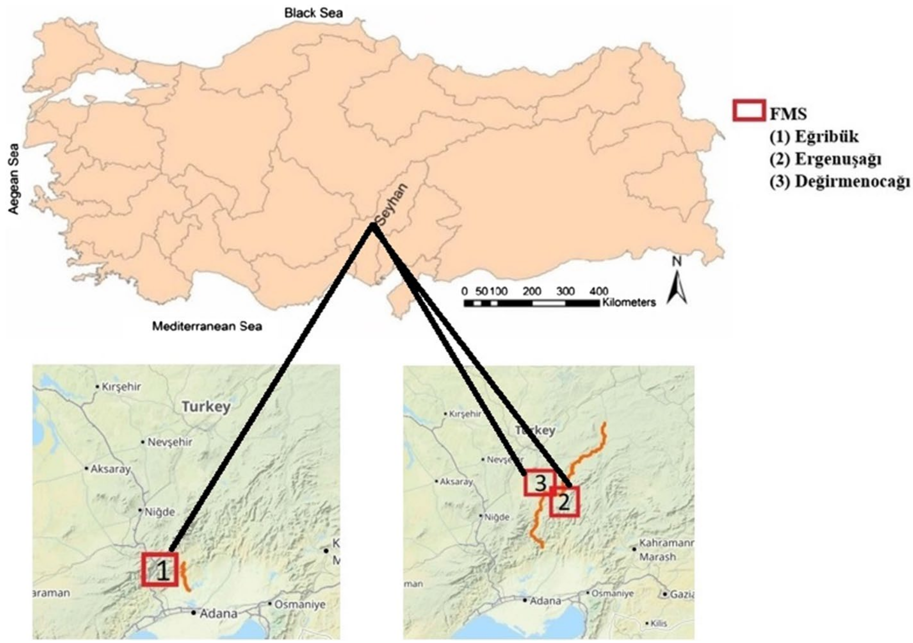

Fig. 1 Study sites in the Seyhan river basin

all stations between 1988 and 2011 are 8928 daily measured flow values. The locations of all stations are given in Table 1 with their geographical coordinates.

As shown in Fig. 2, the daily flow potentials of all three stations were discussed during the flow analysis of the river and the lowest and highest flows were 0.95 m3/s and 434 m3/s, respectively. When the three stations are analyzed concurrently and the stream flow is controlled, the lowest stream flow was observed at the Eğribük station in August 2008 with 0.95 m3/s. The highest flow rate was noted at Ergenuşağı station in December 2002 with 434 m3/s, and there is a difference of 106 m3/s from the closest value. The highest flow rates were detected at all three stations in the first years of the data set, in the 1998–2002 period.

All values in the data set contain daily flow rate values. The daily flow measurement station data operated in the study were gathered from DSİ (Hydraulic State Works). 7143 daily flow data, which is 80% percent of the data set of the stations between 01.01.1988 and 03.11.2007, were employed in the training phase of the models, and 1785 daily flow data, while the 20% percent of data set between 03.12.2007 and 12.31.2011, were used in the testing phase. The data operated in the training phase were maintained to examine the indicators in the models and the test data were applied for performance comparisons of all models. In addition, some rivers may have mislaid data from past dates, with concerns such as interruptions in flow monitoring stations and the inability to make measurements. In the analysis of the data, the length and uninterruptedness of the series are significantly influential in terms of the sensitivity and reliability of the results to be achieved as a result of the study. As indicated in Table 1, the time span of data from all stations was from January 1988 to December 2011.

Elements such as the Seyhan Basin being considered an important food production area for Turkey and Europe, extensive stockbreeding and meadow-pasture in the upper basin, covering agricultural functions such as very strategic products such as wheat, and the entire basin being within national borders are the water sources that are negatively affected by climate change today. Planning, development and management of resources have made the studies very important. Dry farming takes place in the upper and middle parts of the basin, while irrigated agriculture is practiced in the lower parts (Özfidaner et al. 2018). Prominent agricultural products such as wheat, maize, fruit, and vegetables are cultivated in both areas. Two sub-basins have been determined within the scope of reducing the adverse impacts of water scarcity on production resources and socioeconomic life as ensuring the rational and sustainable use of limited water resources in the basin. These sub-basins are listed as the Zamanti River sub-basin and Seyhan Dam confluence sub-basin. The Eğribük station, which is located on the important river branches of the basin due to the surface water potential of the Seyhan Dam confluence sub-basin, and the Ergenuşağı station, which is located close to the discharge point of the Zamanti River sub-basin with the surface

Table 1 Three flow measurement stations located along the Seyhan Basin

FMS	River-FMS	Coordinates		Catchment Area (km^2)	Elevation(m)	Observation (Year)
		East (° ′ ″)	North (° ′ ″)			
E18A025	Eğribük	35 11 35	37 21 50	544.5	222	1988–2011
E18A026	Ergenuşağı	35 34 47	37 39 55	8698.1	360	1988–2011
E18A027	Değirmenocağı	35 29 10	37 51 18	7718	740	1988–2011

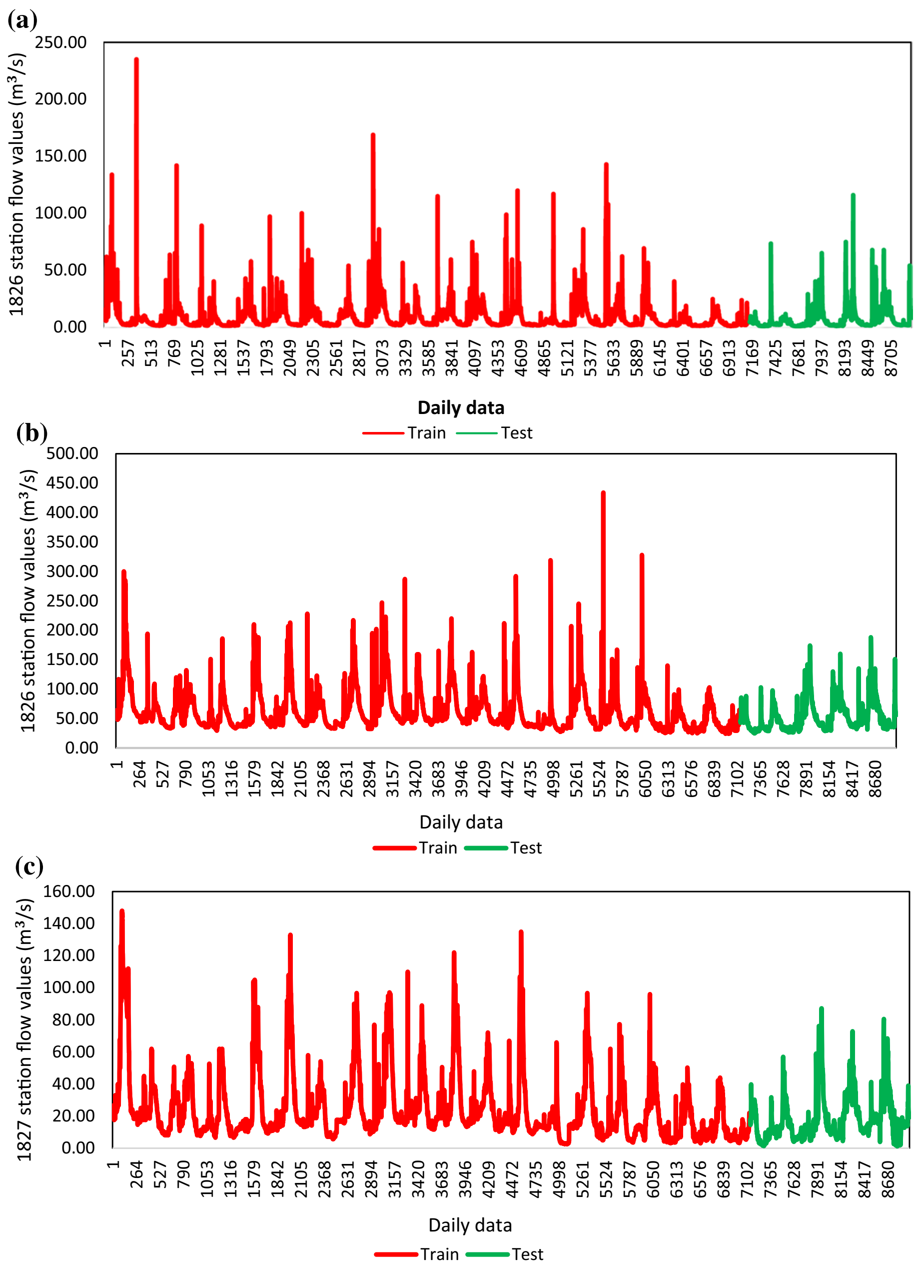

Fig. 2 Daily streamflow for **a** Eğribük and **b** Ergenuşağı **c** Değirmenocağı stations

water potential of the basin, are located in the sub-basins used in this study (Barbaros et al. 2021). Finally, Değirmenocağı station, which is located similarly in the Zamanti River sub-basin by utilizing the feature of being in different branches of the same river and selected for the estimation of prospective flow data, is the additional studied station.

3.3 Adaptive neuro-fuzzy inference system (ANFIS)

ANFIS has developed as an intelligent system with the combination of fuzzy logic and artificial neural networks and is applied for solving numerous engineering issues (Karaboga and Kaya 2019). Its concept was initially suggested by Jang (1993) and as in other artificial neural networks, it is utilized to estimate related problems by comprehending using training data sets. In this study, five-layered ANFIS structure with two inputs (x_1 and x_2) and one output (f) is applied and it is demonstrated in Fig. 3. The layers of ANFIS structure have described as below.

In layer one, the outputs of the nodes in this layer $\left(O_{1i}\right)$ are indicated in Eqs. (1) and (2).

$$O_{1i} = \mu A_i(x) \quad i = 1,\ 2 \tag{1}$$

$$O_{1i} = \mu B_{i-2}(x) \quad i = 3,\ 4 \tag{2}$$

A_i and B_i represent any membership function and μA_i and μB_i illustrate the membership degree. In this study, it was obtained the Gaussian membership functions (MFs) by using input data according to the following Eq. (3).

$$\mu A_i = \exp\left[-\left(\frac{x - c_i}{a_i}\right)^2\right] \tag{3}$$

where x is the input data to i node and a_i, c_i specifies membership function parameters of this set and these parameters represent to premise parameters.

Later on, in layer two, fuzzy sets are gathered with the help of operating the membership functions calculated in layer 1. The membership values are multiplied to obtain the membership degree in the second layer which is called the rule layer. The output which was acquired in this layer is a product of the input.

$$O_{2i} = w_i = \mu A_i(x).\mu B_i(y) \quad i = 1,\ 2 \tag{4}$$

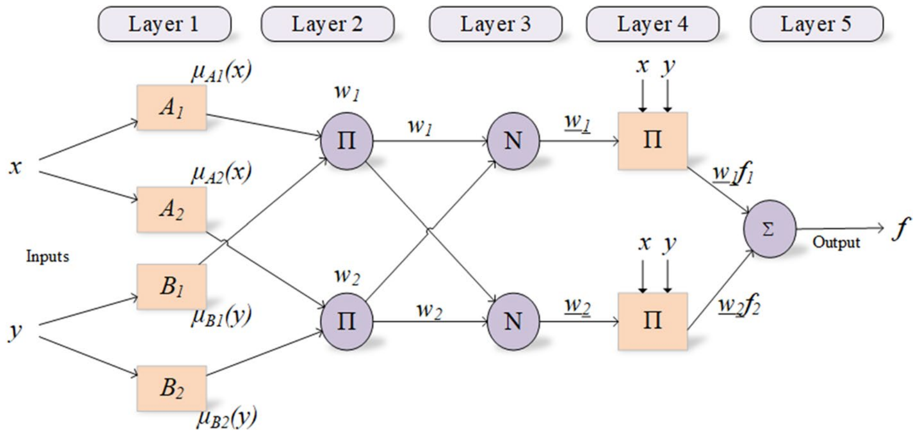

Fig. 3 Basic structure of ANFIS

where w_i refers the membership degree and $\mu A_i(x)$ and $\mu B_i(y)$ are represented membership degree of x in A_i and y in B_i, respectively.

The third layer is the normalization layer, and the membership degree normalization is performed in this layer.

$$O_{3i} = \overline{w_i} = \frac{w_i}{w_1 + w_2} \quad i = 1, 2 \tag{5}$$

where w_i illustrates the normalized membership degree.

The fourth layer is the layer where the outputs of all rules are computed. The calculated outputs are called as consequence parameters which are p_i, q_i, r_i.

$$O_{4i} = \overline{w_i}.f_i = \overline{w_i}.(p_i x + q_i y + r_i) \quad i = 1, 2 \tag{6}$$

The ultimate outputs of each rules are calculated in the fifth layer.

$$O_{5i} = f = \sum \overline{w_i}.f_i = \frac{\sum w_i.f_i}{\sum w_i} \quad i = 1, 2 \tag{7}$$

3.4 Long short-term memory network (LSTM)

The recurrent neural network (RNN) is widely conducted for processing sequential data. LSTM is a RNN-like system, and unlike RNN, it is operated to create long-term memory with short-term dependencies by using nodes (Dong et al. 2020). The LSTM network expands the learning capacity by utilizing long-term memory and is highly efficacious to predict at higher accuracy with multivariate data.

As demonstrated in Fig. 4, the LSTM algorithm consists of three gates named the input gate (i_t), the output gate (o_t), and the forget gate (f_t), which rule the inside-outside information flow in the cell state. These gates ensure to forget the unrelated information, which can initially be calculated by Eq. (8), to transfer the necessary information from the previous loop to the next state Eq. (9) and to produce an output Eq. (10), respectively.

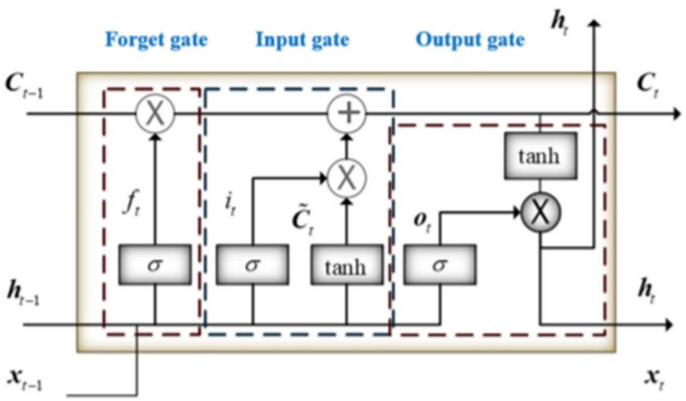

Fig. 4 Schematic of LSTM structure

$$f_t = \sigma\left(W_{f,x} * X_t + W_{f,h} * h_{t-1} + b_f\right) \tag{8}$$

$$i_t = \sigma\left(W_{i,x} * X_t + W_{i,h} * h_{t-1} + b_i\right) \tag{9}$$

$$o_t = \sigma\left(W_{o,x} * X_t + W_{o,h} * h_{t-1} + b_o\right) \tag{10}$$

The input modulate gate ($\tilde{C}_t$) is computed using the current input (X_t) and the cell output in the previous hidden state (h_{t-1}) as in Eq. (11). The current state cell (C_t) and the hidden state (h_t) are calculated using Eqs. (12) and (13), respectively.

$$\tilde{C}_t = \tanh\left(W_{c,x} * X_t + W_{c,h} * h_{t-1} + b_c\right) \tag{11}$$

$$C_t = C_{t-1} * f_t + i_t * \tilde{C}_t \tag{12}$$

$$h_t = o_t * \tanh\left(C_t\right) \tag{13}$$

3.5 Particle swarm optimization algorithm (PSO)

PSO is one of the popular metaheuristic algorithms and was initially presented by Kennedy and Eberhart (1995) inspired by social acts such as fish and bird flocking. In PSO, in order to solve the optimization issues, each particle named the potential solution in the group perceives its position and adjusts its subsequent position and velocity in the process of search according to the performances and fitness of the other members of the group. In searching for the global optimum solution, the higher-performing particle is determined at each iteration and iteratively obtains the best solution which is *pbest*. Moreover, if the calculated final value is more promising than the current optimal fitness value of the swarm (*gbest*), *gbest* is set to the current value of the particle then the position and velocity of that particle are recalculated according to the equations given below. The primary principle of PSO is to adjust the velocity of each swarm towards the *pbest* and *gbest* positions in each iteration.

$$\vec{x}(t+1) = \vec{x}(t) + \vec{v}(t+1) \tag{14}$$

$$\vec{v}(t+1) = w\vec{v}(t) + \phi_1 r\,\mathrm{and}\,(0,1)\left(\vec{p}(t) - \vec{x}(t)\right) + \phi_2 r\,\mathrm{and}\,(0,\ 1)\left(\vec{g}(t) - \vec{x}(t)\right) \tag{15}$$

where x_i^t and v_i^t show the current position and velocity; $x_i^{(t+1)}$ and $v_i^{(t+1)}$ refer the position and velocity at the following time, respectively; r_1 and r_2 are two random numbers generated from uniformly distributed in [0,1]; c_1 and c_2 are acceleration coefficients. The whole process described above is repeated according to the flow presented in Fig. 5 until the most satisfactory result is obtained.

3.6 Optimizing LSTM network using the PSO algorithm

The LSTM network is operated by researchers for predicting streamflow owing to its excellent nonlinear prediction capacity and the capability to receive long-term correlations

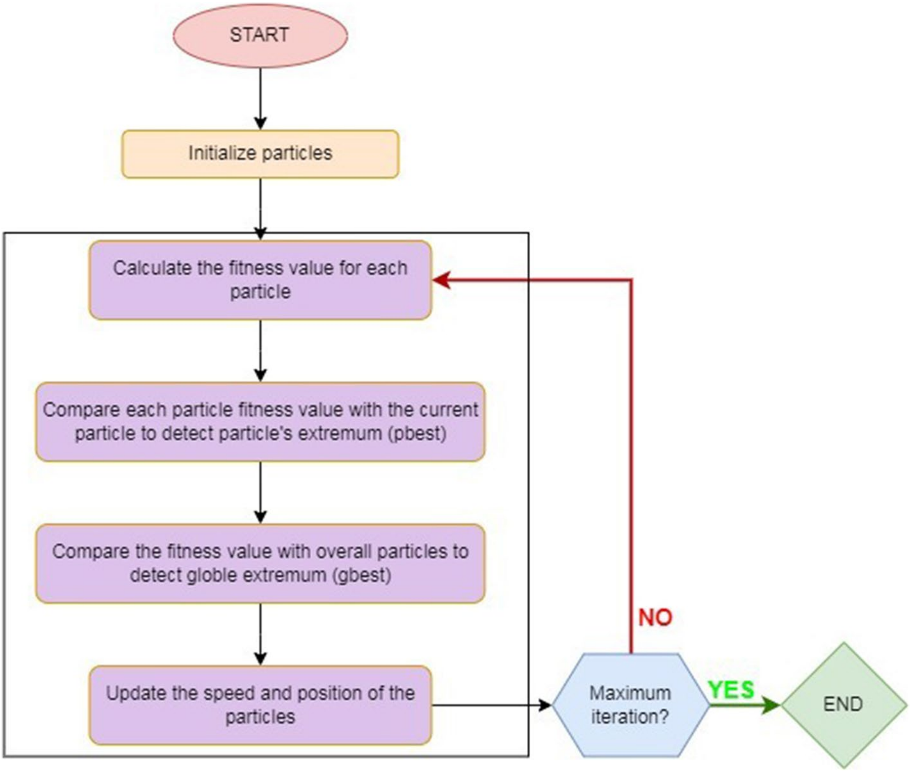

Fig. 5 Flowchart of PSO

using time series. Not long ago, hybrid models have been conducted to enhance the LSTM's internal parameters such as hidden layer nodes and the learning rate. In the present study, PSO, which is the intelligent algorithm, was utilized to uncover the optimum value of these parameters of the LSTM. In the study, the LSTM network was primarily operated to receive the best forecast results as a reference. Then, the dataset was organized into two sets for testing and training. 80% of the dataset was operated for training the network while the remaining data (20%) was used for testing. Hence, the PSO algorithm was performed to explore the global optimum solution in the LSTM network utilizing the fitness function. The fitness value was computed at each iteration to discover the lowest value for the optimum solution. RMSE, MAE, MAPE, and SD criteria were employed to determine the model forecasting performance. Furthermore, LSTM-PSO hybrid model was also explained meticulously, and the model flowchart is demonstrated in Fig. 6.

4 Results and discussion

In this section, the performance of the ANFIS, LSTM-PSO, and LSTM benchmark model was investigated. Test data of streamflow from long-term annually are plotted in Fig. 2. The model's performance was analyzed with 1785 test data for Değirmenocağı, Eğribük, and Ergenusağı FMS. The performance of the hybrid model against ANFIS appears to be

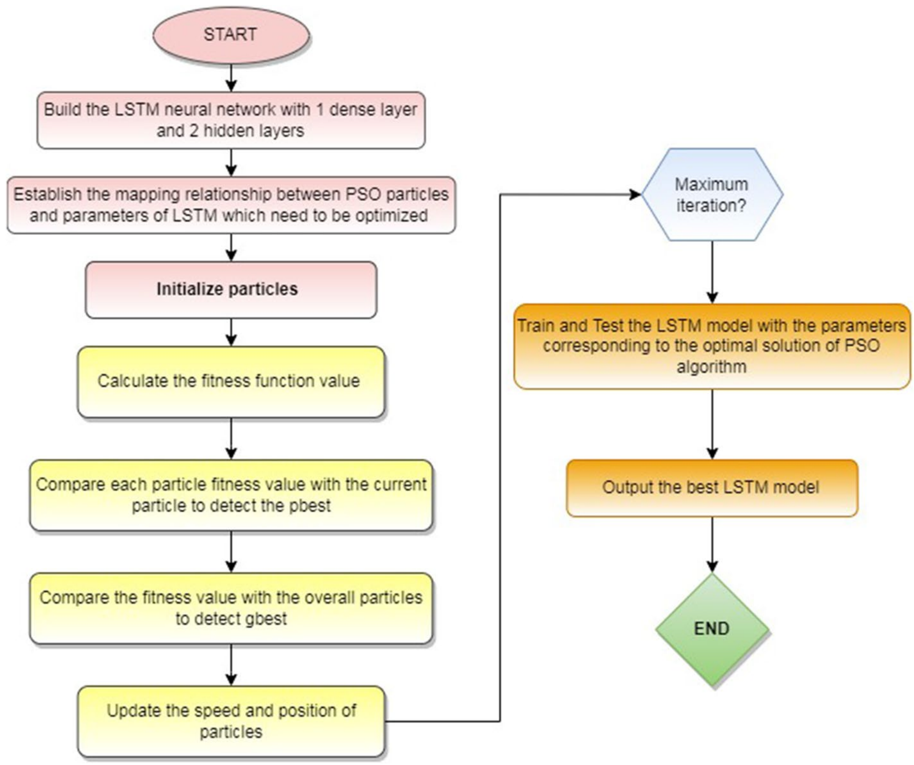

Fig. 6 The flowchart of hybrid LSTM-PSO model

relatively successful when the statistical metrics. Additionally, the purpose is to support the outcomes of the statistical measurements of the hybrid and benchmark models comprised in the study. Numerous attempts were made to determine the PSO algorithm's initial parameter values. As a consequence of the attempts, it was uncovered that the number of iterations was 1000, the population size was 50, the crossover rate was 0.8, and the mutation rate was 0.01.

In the statistical measurement performances of the study, mean square error (MSE), root mean square error (RMSE), mean absolute error (MAE), mean absolute percentage error (MAPE), R^2, and standard deviation (STD) were utilized. These evaluation methods have been widely engaged in various works and are provided as measurement tools for estimating daily flow values and determining the effectiveness of the model (Kilinc and Haznedar 2022). Table 2 illustrates the statistical measure of the model results. Likewise, Table 2 indicates the proposed model perpetrates more pleasing with bold when the error measures are concerned. The comparison estimations of these three models for Değirmenocağı, Eğribük, and Ergenuşağı FMS, respectively, are presented in Table 2. While the estimated MAE values of the benchmark LSTM and ANFIS were 1.7768, 1.3023, 3.0511 and 1.8082, 1.2954, 3.1524, respectively, the estimated MAE values of the proposed LSTM-PSO are 1.7587, 1.3023, and 3.0124, respectively. Concerning RMSE, the proposed LSTM-PSO model is detected to be 3.2762, 4.8109, and 6.6825 while benchmark LSTM is 3.3215, 4.8109, and 6.7725 and ANFIS is 3.3790,

Table 2 Performance measures (All values are in m^3/s)

FMS	Model	MSE	RMSE	MAPE	MAE	R^2	StD
Değirmenocağı	ANFIS	11.4178	3.3790	16.7569	1.8082	0.9390	0.2020
	LSTM	11.0322	3.3215	15.7780	1.7768	0.9408	0.2004
	LSTM-PSO	**10.7337**	**3.2762**	**15.2126**	**1.7587**	**0.9433**	**0.2293**
Egribuk	ANFIS	24.3237	4.9319	13.0760	1.2954	0.6854	0.7819
	LSTM	24.1444	4.8409	13.0419	1.3123	0.6830	0.7859
	LSTM-PSO	**23.1444**	**4.8109**	**13.0319**	**1.3023**	**0.6972**	**0.7849**
Ergenuşağı	ANFIS	47.0879	6.8621	5.1373	3.1524	0.9248	0.1085
	LSTM	45.8666	6.7725	5.1609	3.0511	0.9256	0.1073
	LSTM-PSO	**44.6553**	**6.6825**	**4.9781**	**3.0124**	**0.9273**	**0.1073**

4.9319, and 6.8621. The best MAPE values are discovered as 15.2126, 13.0319, and 4.9781 for the proposed LSTM-PSO model. Ultimately, according to the coefficient of determination (R^2), the proposed LSTM-PSO model is found to be 0.9433, 0.6972, and 0.9273 while benchmark LSTM is 0.9408, 0.6972, 0.9256 and ANFIS is 0.9390, 0.6854, and 0.9248, respectively.

A residual which is also referred to as ‘error value’ can be described as the distinction between the actual data point and the predicted data point. It is a measure of well a line fits for the given regression line and important for showing model performance. In this context, it is analyzed by its magnitude and whether it form a pattern when determining the quality of a model. The proposed LSTM-PSO method and the other residual performance are displayed in Figs. 7, 8, 9. It is apparent that residual values are too undersized and formed a group.

All evaluation criteria confirm that the hybrid LSTM-PSO model is quite successful. The performance of the ANFIS model remains inadequate compared to other models. Due to the distinctions in the instantaneous variability of the currents in the Eğribük River, the all models showed the lowest performance. Following the trend line at Değirmenocağı station, the current values make the prediction capacity higher than

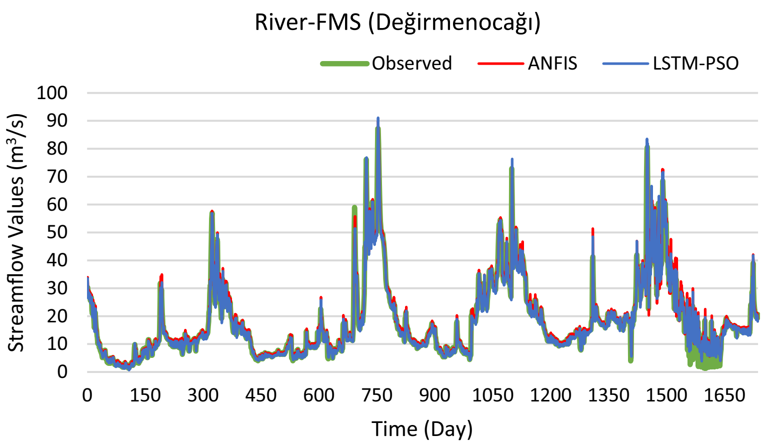

Fig. 7 Comparative plots of the observed and predicted flow of the models for Değirmenocağı FMS

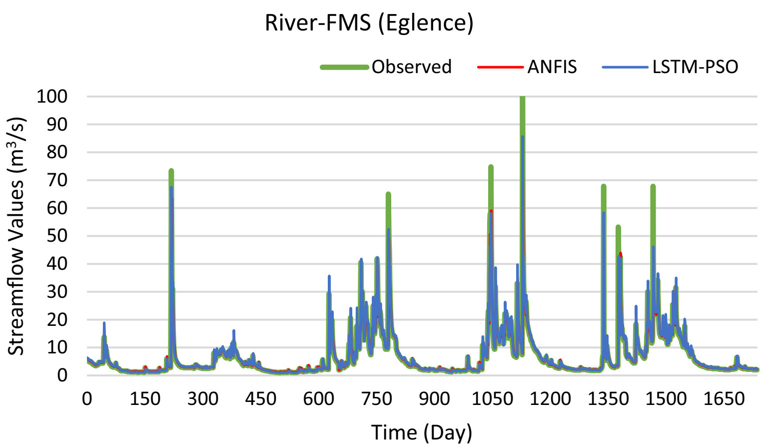

Fig. 8 Comparative plots of the observed and predicted flow of the models for Eğribük FMS

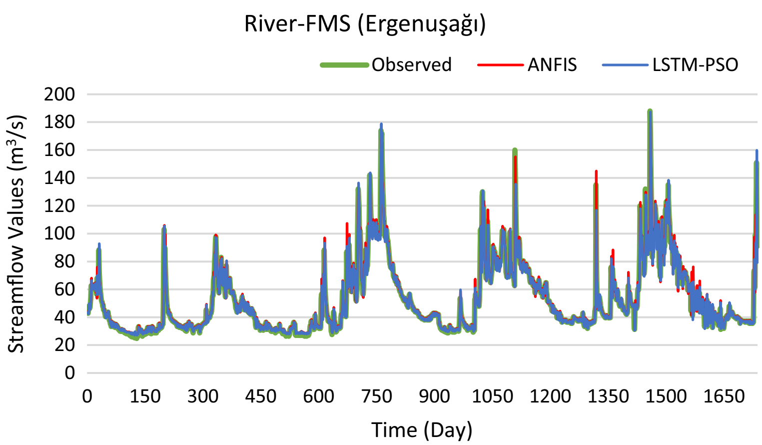

Fig. 9 Comparative plots of the observed and predicted flow of the models for Ergenuşağı FMS

anticipated. Even though ANFIS is approaching LSTM at this station, the success of the hybrid LSTM-PSO model remains at the forefront.

The scatter plots for the LSTM-PSO, LSTM, and ANFIS model for the test data are indicated in Figs. 10, 11, 12 to investigate the coefficient of determination between actual and predicted streamflow data. The proposed LSTM-PSO method results are closer to actual streamflow data even if the Değirmenocağı is commonly further upstream. Thus, LSTM-PSO carries a high success of prediction with an $R^2$of 0.9433.

The LSTM network is an essential function in determining the performance of prediction models. Combining models such as PSO and LSTM appear to provide benefits in the time-series prediction issues such as river flow prediction. The benefit becomes evident when the optimization and generation of other units bring to completion. The proposed hybrid model comprehended to discover the optimal level of river flows and was able to anticipate the following day's flow value. This situation is demonstrated by its significant performance compared to the benchmark model.

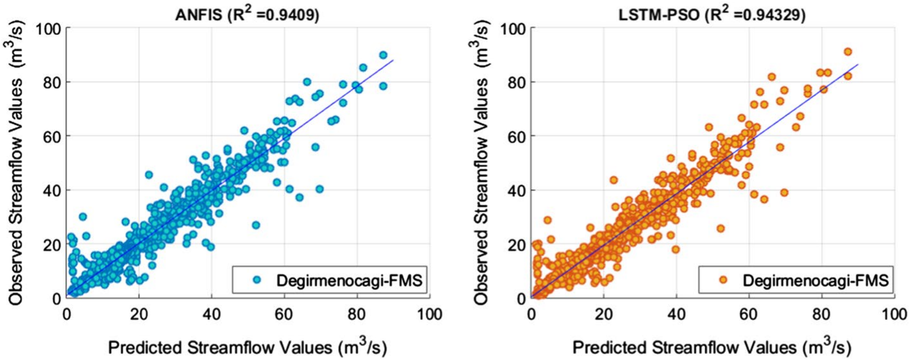

Fig. 10 Scatter plots of the ANFIS and LSTM-PSO models belonging to Değirmenocağı-FMS

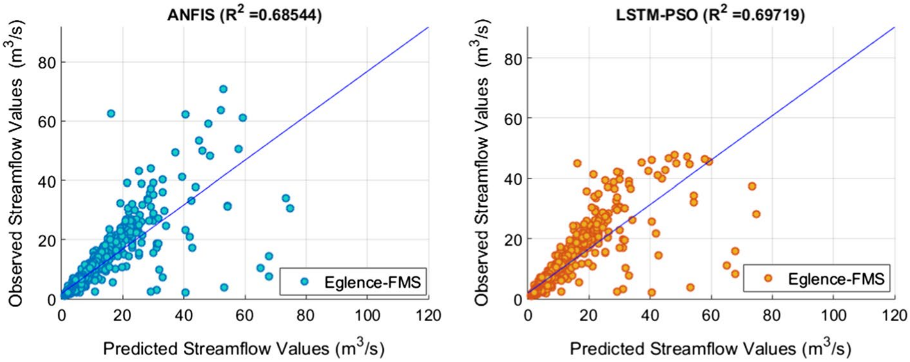

Fig. 11 Scatter plots of the ANFIS and LSTM-PSO models belonging to Eğribük-FMS

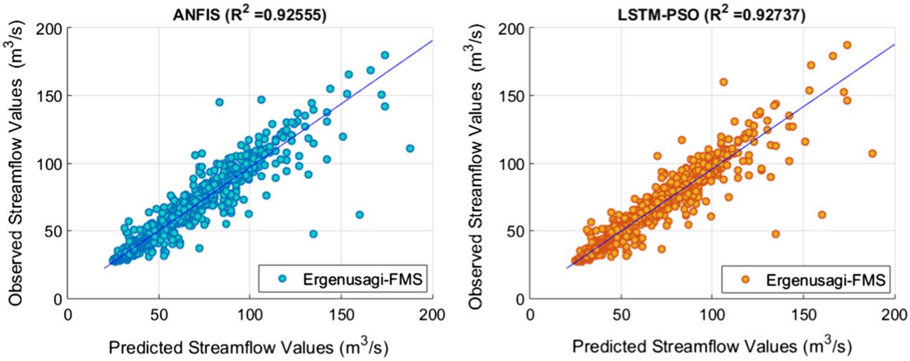

Fig. 12 Scatter plots of the ANFIS and LSTM-PSO models belonging to Ergenuşağı-FMS

The developed hybrid model showed that the method used achieves an optimum result in river flow estimations. Research results and statistical methods supported this situation. In addition, the superiority of the hybrid model over the comparison models, ANFIS and LSTM, was observed at all three stations. Recently, various hybrid models have been developed and used to predict stream flows. When the literature is examined in detail, the contribution of many models to the accuracy and prediction performance is undeniably good. Zhang (2022) used hybrid PSO-SVM model for a comprehensive evaluation. Empirical results research showed that the PSO-SVM algorithm has certain applicability in forecasting accuracy. Mohammed et al. (2022) proposed a novel hybrid machine learning model based on an artificial neural network (ANN) and the Marine Predators algorithm (MPA) for modeling monthly water levels. MPA-ANN algorithm's performance was compared with recent constriction coefficient-based particle swarm optimization and chaotic gravitational search algorithm and slime mold algorithm. The results showed that the proposed model was offered good results. Abdul Kareem et al. (2022) methodology that involved data pre-processing and an artificial neural network (ANN) optimized with the coefficient-based particle swarm optimization and chaotic gravitational search algorithm (CPSOCGSA-ANN) to forecast the monthly water streamflow. The hybrid model outperformed other comparison models. Also, the suggested methodology offered accurate results. According to the results observed from the literature research, it has been shown that the hybrid model has better performance than the comparative models and increases the prediction precision.

5 Conclusions

In this study, a hybrid method building PSO and LSTM is proposed to estimate flow data. The performance of the proposed model has been tested on the data of the Ergenuşağı, Eğribük, and Değirmenocağı streamflow measurement stations. Basically, although the LSTM neural network shows a good learning ability for time-series predictions, its performance falls short due to the ineffectiveness of some hyperparameters. For example, LSTM models still have problems such as slow convergence rate and low learning capacity, resulting in too long training time or even poor fit. For this reason, the effective powers of the PSO parameters were used. PSO was used to search for suitable values of LSTM parameters. In addition, the performance of the models was compared with ANFIS, which is one of the classical estimation methods. The training and application areas of these models are related to the size and accuracy of datasets. Therefore, statistical measurement criteria such as RMSE, MAE, MAPE, SD, and R^2 are used especially for estimation measurements and basic criteria for observing the performance of the model. These criteria were used to evaluate the performance of the proposed method. The results showed that with the proposed LSTM-PSO approach, the estimation error of the flow data was reduced more successfully than the benchmark model. As a result, it was determined that the approach had low metrics and these results became meaningful with the evaluation criteria. This shows that the hybrid model approach can improve the benchmark model.

As mentioned in the previous section, it was determined that the prediction performance of LSTM-PSO was higher when compared to the ANFIS and LSTM results at all three measurement stations. In all three measurement stations, it was obtained better results for LSTM-PSO on average by 2–10% in all evaluation criteria, and it was achieved 10.14% improvement for MAPE, especially at the Değirmenocağı measurement station. This improvement value shows that LSTM-PSO provides 10% better quality and more accurate streamflow forecasting. LSTM-PSO can ensure the long-term correlation of time series with high accuracy and it can have said that it has a stronger nonlinear forecast capability to improve the reliability of streamflow prediction. Additionally, the results obtained in Değirmenocağı FMS are very close to the observed values given in Fig. 7, and the R^2 value obtained in this FMS is the highest with 0.9433. In Ergenuşağı FMS, on the other hand, the R^2 value was obtained as 0.9273, and the observed and predicted flow values in this FMS are very close to each other (Fig. 9). In addition, in the study, generally lower error rates were obtained in the LSTM model compared to the ANFIS model. As a result, it was determined that the approach had low metrics and these results became meaningful with the evaluation criteria. This shows that the hybrid model approach can improve the benchmark model. Thus, the LSTM-PSO model proposed in this study performs better in three different hydrological stations compared to other models, which shows that this model has universal applicability and that more accurate results can be obtained in flow forecasting for different hydrological stations. Therefore, it can be considered as a promising alternative to improve long-term daily streamflow prediction quality.

Nevertheless, the study has some limitations. The most significant aspect impacting the success of heuristic optimization algorithms is the optimization of their parameters. Every optimization algorithm has parameters. The number of iterations is the most widely accustomed and known parameter. In addition, the speed of convergence is one of the criteria that best measures the performance of heuristic optimization techniques. If the number of iterations is low, the algorithm cannot converge, and if it is too high, it causes disproportionate use of solution time and computer resources. The detriments of the Particle Swarm Optimization (PSO) algorithm, which is one of the heuristic algorithms operated in the study, are that it is straightforward to descend to the local optimum in high-dimensional space and has a low convergence rate in the iterative process. Hybridizing the PSO algorithm with LSTM augments its performance. As shown in Table 2, the hybrid model successfully improved the optimum values of the benchmark model. On the other hand, it has been seen that the LSTM model, which boosts the performance of the PSO algorithm and is frequently used in time prediction series, delivers successful results. Yet, LSTM network performance occasionally presents unsatisfactory results owing to the random selection of initialization parameters. Thus, hybrid modeling studies persist to attract increasing attention from researchers to acquire better performance results. Therefore, the random selection of the initial parameters and optimization of the hyper parameters are extremely crucial in order to improve the performance measures of the LSTM model. Despite all its advantages and its enduring structure, there are also handicaps of the ANFIS model. In this method, models suitable for certain dataset structures are created, models suitable for every common dataset are not very common. In addition, many attempts should be made in the process of revealing the appropriate model for the data set at hand. The selection of the sample for the data, the size of the data set, the learning method and the type of membership function to be operated are the factors determining the quality of the output values and these values are determined by experimentation. As a result, ANFIS, which is incorporated in the study as a classical comparison model, can be hybridized in forthcoming studies to observe success performance with other hybrid models. Furthermore, the current data are nonlinear time series. Studies with other parameters with these features will play an active role in the management of river regimes. The inclusion of data on a minute, hourly or monthly basis can also be regarded as a criterion in performance analysis. As an alternative to the PSO algorithm, which was successful by hybridizing in the study, the contribution of the hybrid models to be created with the popular metaheuristic algorithms of recent times to the prediction accuracy can be examined.

Author contribution BH contributed to conceptualization, methodology, writing—original draft preparation, and software. HCK contributed to supervision, methodology, validation, and software. FO contributed to writing—reviewing and editing. AY contributed to writing—reviewing and editing. The authors give their full consent for the publication of this manuscript.

Funding The authors declare that no funds, grants, or other support were received during the preparation of this manuscript.

Data availability The data will be available upon reasonable request.

Declarations

Conflict of interests The authors declare no conflict of interests.

Ethical approval The authors undertake that this article has not been published in any other journal and that no plagiarism has occurred.

Consent to participate The authors agree to participate in the journal.

Consent to publish The authors agree to publish in the journal.

References

Abdul Kareem B, Zubaidi SL, Ridha HM et al (2022) Applicability of ANN model and CPSOCGSA algorithm for multi-time step ahead river streamflow forecasting. Hydrology 9:171. https://doi.org/10.3390/hydrology9100171

Adaryani FR, Jamshid Mousavi S, Jafari F (2022) Short-term rainfall forecasting using machine learning-based approaches of PSO-SVR, LSTM and CNN. J Hydrol 614:128463. https://doi.org/10.1016/j.jhydrol.2022.128463

Asaad MN, Eryürük Ş, Eryürük K (2022) Forecasting of streamflow and comparison of artificial intelligence methods: a case study for Meram stream in Konya Turkey. Sustainability 14:6319. https://doi.org/10.3390/su14106319

Barbaros F, Onuşluel Gül G, Boyacioğlu H (2021) Evaluation of seasonality in water quality with non-parametric statistical methods in the sample of Küçük Menderes Basin. J Suleyman Demirel Univ Insti Sci Nat Appl Sci 25(2):195–207. https://doi.org/10.19113/sdufenbed.790331

Cao Q, Banerjee R, Gupta S, et al (2016) Data driven production forecasting using machine learning. In: Day 2 Thu, June 02, 2016. SPE, Buenos Aires, pp D021S006R001

Chen Y, Song L, Liu Y et al (2020) A review of the artificial neural network models for water quality prediction. Appl Sci 10:5776. https://doi.org/10.3390/app10175776

Cheng Y, Wang C, Wu J et al (2022) Multi-dimensional recurrent neural network for remaining useful life prediction under variable operating conditions and multiple fault modes. Appl Soft Comput 118:108507. https://doi.org/10.1016/j.asoc.2022.108507

Daba MH, You S (2020) Assessment of climate change impacts on river flow regimes in the upstream of Awash Basin, ethiopia: based on IPCC fifth assessment report (AR5) climate change scenarios. Hydrology 7:98. https://doi.org/10.3390/hydrology7040098

Dalkiliç HY, Hashimi SA (2020) Prediction of daily streamflow using artificial neural networks (ANNs), wavelet neural networks (WNNs), and adaptive neuro-fuzzy inference system (ANFIS) models. Water Supply 20:1396–1408. https://doi.org/10.2166/ws.2020.062

Dolezal JM, Trzcinska A, Liao C-Y et al (2021) Deep learning prediction of BRAF-RAS gene expression signature identifies noninvasive follicular thyroid neoplasms with papillary-like nuclear features. Mod Pathol 34:862–874. https://doi.org/10.1038/s41379-020-00724-3

Dong L, Fang D, Wang X et al (2020) Prediction of streamflow based on dynamic sliding window LSTM. Water 12:3032. https://doi.org/10.3390/w12113032

Du B, Huang S, Guo J et al (2022) Interval forecasting for urban water demand using PSO optimized KDE distribution and LSTM neural networks. Appl Soft Comput 122:108875. https://doi.org/10.1016/j.asoc.2022.108875

Durgut R, Aydïn M, (2021) Adaptive binary artificial bee colony for multi-dimensional knapsack problem. J Gazi Univ Fac Eng Archit 36(4):2333–2348. https://doi.org/10.17341/gazimmfd.804858
Feng W, Guan N, Li Y, et al (2017) Audio visual speech recognition with multimodal recurrent neural networks. In: 2017 International joint conference on neural networks (IJCNN). IEEE, Anchorage, pp 681–688
Haznedar B, Arslan MT, Kalinli A (2021) Optimizing ANFIS using simulated annealing algorithm for classification of microarray gene expression cancer data. Med Biol Eng Comput 59:497–509. https://doi.org/10.1007/s11517-021-02331-z
İpek B (2021) Comparison of ANFIS and ARIMA models via Industry 4.0 Data. Msc. Thesis, Karadeniz Technical University, Trabzon, Turkey
Jang J-SR (1993) ANFIS: adaptive-network-based fuzzy inference system. IEEE Trans Syst Man Cybern 23:665–685. https://doi.org/10.1109/21.256541
Karaboga D, Kaya E (2019) Adaptive network based fuzzy inference system (ANFIS) training approaches: a comprehensive survey. Artif Intell Rev 52:2263–2293. https://doi.org/10.1007/s10462-017-9610-2
Karahan SM (2021) Simulation of water resources of Tahtali-Seferihisar Sub-Basin based on weap model. Msc. Thesis, İzmir Institute of Technology University, İzmir, Turkey
Kayhomayoon Z, Babaeian F, Ghordoyee Milan S et al (2022) A combination of metaheuristic optimization algorithms and machine learning methods improves the prediction of groundwater level. Water 14:751. https://doi.org/10.3390/w14050751
Kennedy J, Eberhart R (1995) Particle swarm optimization. In: Proceedings of ICNN'95-international conference on neural networks. IEEE, Perth, WA, Australia, pp 1942–1948
Kilinc HC, Haznedar B (2022) A hybrid model for streamflow forecasting in the Basin of euphrates. Water 14:80. https://doi.org/10.3390/w14010080
Kilinc HC, Yurtsever A (2022) Short-term streamflow forecasting using hybrid deep learning model based on grey wolf algorithm for hydrological time series. Sustainability 14:1–20. https://doi.org/10.3390/su14063352
Kim T-Y, Cho S-B (2021) Optimizing CNN-LSTM neural networks with PSO for anomalous query access control. Neurocomputing 456:666–677. https://doi.org/10.1016/j.neucom.2020.07.154
Komuscu AU, Erkan A, Oz S (1998) Possible impacts of climate change on soil moisture availability in the southeast Anatolia development project region (GAP): an analysis from an agricultural drought perspective. Clim Change 40:519–545. https://doi.org/10.1023/A:1005349408201
Kushwaha OS, Uthayakumar H, Kumaresan K (2022) Modeling of carbon dioxide fixation by microalgae using hybrid artificial intelligence (AI) and fuzzy logic (FL) methods and optimization by genetic algorithm (GA). Environ Sci Pollut Res. https://doi.org/10.1007/s11356-022-19683-0
Lv L, Kong W, Qi J, Zhang J (2018) An improved long short-term memory neural network for stock forecast. In: MATEC web conferences, vol 232, pp 01024https://doi.org/10.1051/matecconf/201823201024
Mohammed SJ, Zubaidi SL, Al-Ansari N et al (2022) Hybrid technique to improve the river water level forecasting using artificial neural network-based marine predators algorithm. Adv Civ Eng 2022:1–14. https://doi.org/10.1155/2022/6955271
Nazimi N (2021) Monthly streamflow prediction of euphrates Basin by using FFNN, ANFIS and LSTM models. Msc. Thesis, Erzincan Binali Yıldırım University, Erzincan, Turkey
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
Oyelade ON, Ezugwu AE-S, Mohamed TIA, Abualigah L (2022) Ebola optimization search algorithm: a new nature-inspired metaheuristic optimization algorithm. IEEE Access 10:16150–16177. https://doi.org/10.1109/ACCESS.2022.3147821
Özfidaner M, Şapolyo D, Topaloğlu F (2018) Hydrological drought analysis of Seyhan basin flow data. Soil Water J. https://doi.org/10.21657/topraksu.410140
Pranolo A, Mao Y, Wibawa AP et al (2022) Robust LSTM With tuned-PSO and bifold-attention mechanism for analyzing multivariate time-series. IEEE Access 10:78423–78434. https://doi.org/10.1109/ACCESS.2022.3193643
Sheikh Khozani Z, Barzegari Banadkooki F, Ehteram M et al (2022) Combining autoregressive integrated moving average with long short-term memory neural network and optimisation algorithms for predicting ground water level. J Clean Prod 348:131224. https://doi.org/10.1016/j.jclepro.2022.131224

Siva Kumar M, Rajamani D, Abouel Nasr E et al (2021) A hybrid approach of ANFIS—artificial bee colony algorithm for intelligent modeling and optimization of plasma arc cutting on MonelTM 400 alloy. Materials 14:6373. https://doi.org/10.3390/ma14216373
Solomatine DP, Ostfeld A (2008) Data-driven modelling: some past experiences and new approaches. J Hydroinf 10:3–22. https://doi.org/10.2166/hydro.2008.015
Song X, Liu Y, Xue L et al (2020) Time-series well performance prediction based on long short-term memory (LSTM) neural network model. J Pet Sci Eng 186:106682. https://doi.org/10.1016/j.petrol.2019.106682
Topaloğlu F (1999) Investigation of a suitable method for the estimation of the magnitude and frequency of floods in the streams of the Seyhan basin. Phd. Thesis, Cukurova University, Adana, Turkey
Türkşen Ö, Akgün F (2018) Point estimation of nonlinear regression model parameters with the genetic-simplex hybrid algorithm. J Stat Stat Actuar Sci 2:81–92
Tüzün, BN (2022) Classification using deep learning models for brain tumors and some tumors affecting the brain. Msc. Thesis, Kütahya Dumlupınar University, Kütahya, Turkey
Wang Y, Tang H, Huang J et al (2022) A comparative study of different machine learning methods for reservoir landslide displacement prediction. Eng Geol 298:106544. https://doi.org/10.1016/j.enggeo.2022.106544
Wegayehu EB, Muluneh FB (2022) Short-term daily univariate streamflow forecasting using deep learning models. Adv Meteorol 2022:1–21. https://doi.org/10.1155/2022/1860460
Xu Y, Hu C, Wu Q et al (2022) Research on particle swarm optimization in LSTM neural networks for rainfall-runoff simulation. J Hydrol 608:127553. https://doi.org/10.1016/j.jhydrol.2022.127553
Yağın B (2022) COVID-19 Prediction with artificial intelligence based image processing methods. Msc. Thesis, İnönü University, Malatya, Turkey
Yamashita R, Nishio M, Do RKG, Togashi K (2018) Convolutional neural networks: an overview and application in radiology. Insights Imaging 9:611–629. https://doi.org/10.1007/s13244-018-0639-9
Yaseen ZM, Mohtar WHMW, Ameen AMS et al (2019) Implementation of univariate paradigm for streamflow simulation using hybrid data-driven model: case study in tropical region. IEEE Access 7:74471–74481. https://doi.org/10.1109/ACCESS.2019.2920916
Ye M, Wang J, Si X et al (2022) Analysis on dynamic evolution of the cost risk of prefabricated building based on DBN. Sustainability 14:1864. https://doi.org/10.3390/su14031864
Yu J, de Antonio A, Villalba-Mora E (2022) Deep learning (CNN, RNN) applications for smart homes: a systematic review. Computers 11:26. https://doi.org/10.3390/computers11020026
Zhang Z (2022) The application of PSO-SVM algorithm in the evaluation system of sports competition events. Secur Commun Netw 2022:1–8. https://doi.org/10.1155/2022/6865425
Zhang F, Dai H, Tang D (2014) A conjunction method of wavelet transform-particle swarm optimization-support vector machine for streamflow forecasting. J Appl Math 2014:1–10. https://doi.org/10.1155/2014/910196
Zhang Q, Yang LT, Chen Z, Li P (2018) A survey on deep learning for big data. Inf Fusion 42:146–157. https://doi.org/10.1016/j.inffus.2017.10.006