Erciyes Üniversitesi
Fen Bilimleri Enstitüsü Dergisi
Cilt 41, Sayı 1, 2025

Erciyes University
Journal of Institue Of Science and Technology
Volume 41, Issue 1, 2025

Estimation on River Water Flow Rate Using Long Short-term Memory (LSTM) Neural Network and Adaptive Neuro-fuzzy Inference Systems (ANFIS)

Akin Ilhan[*1]

[*1] Department of Energy Systems Engineering, Faculty of Engineering and Natural Sciences, Ankara Yildirim Beyazit University, Ankara

(Alınış / Received: 06.07.2024, Kabul / Accepted: 06.04.2025, Online Yayınlanma / Published Online: 30.04.2025)

Keywords
Adaptive Neuro Fuzzy Inference System (ANFIS), Fuzzy C-means (FCM), Subtractive Clustering (SC), Grid Partitioning (GP), Long Short-term Memory (LSTM)

Abstract: The forecasting of river water flow rate (RWFR) plays a prominent role in planning and constructing of new hydraulic dams, or running the ones that were formerly built. This study suggests algorithms of machine learning to predict future water flow rate values for river flow. Namely, estimation models were advanced according to the past time-series RWFR data to obtain the values of future RWRF. Accordingly, long short-term memory (LSTM) neural network, adaptive neuro-fuzzy inference systems (ANFIS) including fuzzy c-means (FCM), subtractive clustering (SC), and grid partitioning (GP) were advanced for the aim of RWFR predictions. A measurement station (MS), settled at the border of Türkiye and Bulgaria, named as Svilengrad MS was selected on the Maritsa River, as the study region. Accordingly, it was concluded that SC algorithm of ANFIS have generated better results compared with respect to other three methods. The comparisons of the data estimations according to the real observed water flow values were accomplished depending on the statistical error values including mean absolute error (MAE), root mean square error (RMSE), else the correlation coefficient (R). Eventually, it was concluded and shown that the superior model of SC has generated those statistical accuracy values, respectively to correspond 3.04 m^3/s MAE, 4.91 m^3/s RMSE, and 0.9979 R, among the total of 102 tested models using FCM, SC, GP, and LSTM.

Uzun Kısa Süreli Bellek (LSTM) Sinir Ağı ve Uyarlanabilir Nöro Bulanık Çıkarım Sistemleri (ANFIS) Kullanılarak Nehir Suyu Akış Hızının Tahmini

Anahtar Kelimeler
Uyarlanabilir Nöro Bulanık Çıkarım Sistemi (ANFIS), Bulanık C-ortalamaları (FCM), Çıkarım Kümeleme (SC), Izgara Bölümleme (GP), Uzun Kısa Süreli Bellek (LSTM)

Öz: Nehir suyu akış hızının (RWFR) tahmini, yeni su barajlarının planlanması ve inşa edilmesinde veya daha önce inşa edilmiş olanların işletilmesinde çok önemli bir rol oynar. Bu çalışma, nehir akışının gelecekteki su akış hızı değerlerini tahmin etmek için makine öğrenme algoritmaları önermiştir. Yani tahmin modelleri, gelecekteki RWRF değerlerini elde etmek için geçmiş zaman serisi RWFR verilerine dayalı olarak geliştirilmiştir. Buna göre, uzun kısa süreli bellek (LSTM) sinir ağı, bulanık c-ortalamaları (FCM), çıkarım kümelemeyi (SC), ızgara bölümlemeyi (GP) içeren uyarlanabilir nöro-bulanık çıkarım sistemleri (ANFIS), RWFR tahminleri amacıyla geliştirilmiştir. Çalışma bölgesi olarak Meriç Nehri üzerinde Türkiye-Bulgaristan sınırında yer alan ve Svilengrad MS olarak adlandırılan ölçüm istasyonu (MS) seçilmiştir. Buna göre, ANFIS'in SC algoritmasının diğer üç yönteme kıyasla daha iyi sonuçlar ürettiği sonucuna varılmıştır. Veri tahminlerinin gerçek gözlemlenen su akış değerlerine göre karşılaştırılması, ortalama mutlak hata (MAE), ortalama karekök hata (RMSE) ve korelasyon katsayısı (R) dahil olmak üzere istatistiksel hata değerlerine bağlı olarak gerçekleştirilmiştir. Neticede, SC'nin en iyi modelinin, FCM, SC, GP ve LSTM kullanan toplam 102 test modeli arasında, sırasıyla, 3,04 m^3/s MAE, 4,91 m^3/s RMSE ve 0,9979 R'ye karşılık gelen istatistiksel doğruluk değerlerini ürettiği sonucuna varılmıştır ve gösterilmiştir.

*İlgili Yazar, email: akinilhan@aybu.edu.tr

Nomenclature

Abbreviation	Definition
σ_c	State activation function
σ_g	Gate activation function
$\odot$	Hadamard product
a(j)	Anticipated instantaneous value
ACO	Ant colony optimization
ANFIS	Adaptive neuro fuzzy inference system
ANN	Artificial neural network
b	Bias
BP	Backpropagation
ELM	Extreme learning machine
f	Forget gate
FCM	Fuzzy c-means
FIS	Fuzzy inference system
g	Cell candidate
GA	Genetic algorithm
GP	Grid partitioning
GRNN	Generalized regression neural network
HD	Number of the historical data
HL	The number of the hidden layer
i	Input gate
IR	The value of the influence radius
LSTM	Long short-term memory
MAE	Mean absolute error
MFs	The number of the membership functions
o	Output gate
PSO	Particle swarm optimization
R_i	The input gate for the recurrent weight
R	Correlation coefficient
r(j)	Real measured instantaneous value
RES	Renewable energy source
RMSE	Root mean square error
RNN	Recurrent neural network
RVM	Relevance vector machine
SC	Subtractive clustering
SVM	Support vector machine
W	Input weight
WNN	Wavelet neural network

1. Introduction

There is a significant increase of the installations of renewable energy sources (RESs) in many countries of the World, especially in the recent years. This is due to the swift rise of greenhouse gas emissions, causing global warming as well as resulting in especially health problems in humans and even in other living organisms. Accordingly, RESs have a rapid spreading potential compared with respect to the conventional fossil based energy sources. For this reason, RESs extend all over the World so quickly and they are recently found in many geographical places of the earth globe. Among the renewable energy sources, most important types include wind, solar, and the hydraulic types of energy production. Although hydraulic energy facility installations have partially declined in the latest years with the rising interest in other types of renewable energy sources; the hydraulic energy is still a type of renewable energy generation method that is one of the most widely used and preferred power source in the World. Hydropower along with the solar energy having these significant features compete closely with each other both produce the most power worldwide, compared to the other renewable sources.

Figure 1 indicates the growth trend of most important types of renewable sources for the World, including hydraulic, solar, and wind; considering the year range between 2014 to 2024. On the other hand, their cumulative amounts including the other sources are presented with dark blue total column, shown in this figure. The other sources consist of renewable sources such as marine energy, biomass energy, energy captured from solid biofuels and renewable waste, energy obtained from bagasse, energy acquired from renewable municipal waste, and finally, energy achieved from other solid biofuels, liquid biofuels, biogas, plus the geothermal energy. Since the capacity of the other sources is so low in magnitude of the total installed power including these cited sources, for this reason, those are totally shown with a single column, and exhibited in this figure using yellow color. As displayed in Figure 1, the total capacity for the renewable energy has significantly augmented from 1,706.98 GW of installations to reach to 4,758.23 GW, respectively from 2014 to 2024. Namely, this corresponds to an increase of approximately 2.8 times, in the last decade of the installed global power analysis. Another striking point in this way is that while the total installed power of wind and solar power plants was quite low as of 2014, by the year of 2023 these two technologies have become competitive with the hydraulic technology, considering the worldwide situation. Another important issue is that the amount of solar installed power, which was quite less compared to the wind installed power as of 2014, has become almost equalized as of 2019 and 2020; and by 2024, the amount of solar installed power has greatly exceeded the total amount of wind installed power, i.e., this ratio increase of solar installations to wind installations corresponded to around 1.7 times more by 2024. Taking the last year of analysis into account demonstrates that total installations of wind, solar, hydro, and other power types respectively correspond to 1,160.20 GW, 1,962.97 GW, 1,605.66 GW, and 29.41 GW, constituting a cumulative installed worldwide power of 4,758.23 GW.

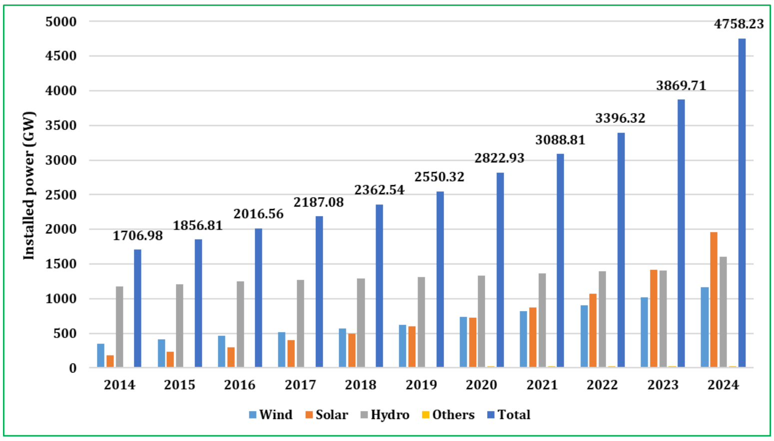

Figure 1. World installed renewables by sources

In the literature, the machine learning and artificial intelligence predictions are mainly implemented to wind and solar power estimations. Accordingly, there are a variety of different studies in obtaining output power anticipations of photovoltaic power plants. For instance, Lurwan et al. used a model of Hottel designed in Matlab software to obtain solar radiation forecast [1]. On the other hand, predictions on the electricity production and the related consumption of the solar power plant of Izmir Bakircay University were actualized, using the artificial neural network (ANN) and the conventional analysis of regressions [2]. Theocharides et al. performed a performance comparison of machine learning methods including regression trees (RT), support vector regression (SVR), and artificial neural networks (ANN) at a test center found in University of Cyprus. In their study, they have shown that the highest prediction capability was achieved during the utilization of the ANN method [3]. Scott et al. utilized different learning algorithms consisting of support vector machines (SVM), traditional linear regression (LR), random forest (RF), and the neural network (NN) in energy estimations of systems of PV. In their study, algorithm input parameters involved pressure of air in milli-bar, temperature in °C, wind speed given in mile per hour, the humidity provided in percentage and the amount of the cloudiness supplied in percentage. The root mean square error statistical results were reported in their study to choose the best prediction tool to be RF [4]. Furthermore, in a study conducted by Gaboitaolelwe et al., the prediction performances of the support vector machine (SVM), gradient boosting (GB), extreme gradient boosting (XGB), support vector regressor (SVR), random forest (RF), and the algorithm of multi-layer perceptron were checked against each other. In this study, the superior performances were obtained on the behalf of gradient boosting (GB) and the random forest (RF) [5].

On the other hand, Liu et al., in their study, proposed autoregressive integrated moving average (ARIMA) and the EMD (empirical mode decomposition). They considered this developed model to actualize forecasts implemented to wind speed in order to obtain a model for a warning system of strong wind for railways [6]. Kavasseri and Seetharaman suggested wind speed forecasts to obtain the magnitudes of wind blowing speed for the next day and two days later, using the ARIMA type of modelling, in which the model was exerted to the region North provinces of Dakota of USA [7]. In the literature, the ANNs are also utilized in acquiring forecasts for wind speed and power. As similar to ANNs, the long short-term memory (LSTM) is mainly exploited to obtain anticipations for power output of wind turbines [8].

Besides, machine learning is also taken into account for to attain guesses on temperature and pressure of the weather. In this way future magnitudes of pressure and temperature of the air can be presumed. Information on the weather data projected for the future is significantly important, especially for the safe travel of airplanes and the ships. Namely, when the future conditions of the weather are known, it will guide all sea vehicles, air vehicles and land vehicles during their secure voyages. Weather predictions were executed depending on the time-series data obtained in historical data cluster, by Zaytar and Amrani. Deep neural network was considered in their study. They have shown the historical data estimations obtained in the form of time-series, carried out by the LSTM simulations have exhibited that the result accuracy was better than the accuracy of the results acquired by the conventional methods [9]. On the other hand, LSTM approximations were conducted to form the prediction data set of one-hour ahead, in the study of Prabha et al [10]. Three different machine learning approaches of ANFIS, comprising of FCM, SC, and GP methodologies, referring respectively to fuzzy c-means, subtractive clustering, and grid partitioning were together considered in the study of Benmouiza and Cheknane [11]. In the study, they have concluded FCM to give better results compared to SC and GP methodologies [11]. Besides, machine learning applications were even taken into account to obtain estimations for heat index map in Türkiye. In the study, it was demonstrated that the ANN tool outperformed the ANFIS [12]. A variety of different machine learning methods were considered in order to predict the aerodynamic coefficients, in the study of Tumse et al. [13]. The methods of FNN, ENN, and ANFIS were exhibited to accurately estimate the aerodynamic coefficients of the delta wind [13]. In the study of Tumse et al., output power of wind turbine was estimated by using a variety of soft computing models. Accordingly, it was reported that the ANFIS tool can be implemented in wind output power forecasts [14].

On the other hand, there are variety of studies dealing with the data predictions of the hydraulic phenomena. In this regards, Hewett et al. have considered a technique of machine learning in order to perform predictions for water flows in regional base for the lake of Okeechobee [15]. The lake has been exploited to purvey water to the south provinces of city of Florida and it has been chosen as the interested location for the study. On the other hand, the forecasts that were implemented for flood hydrography were also taken into consideration with the utilization of the ant colony optimization (ACO), by considering genetic algorithm (GA), through employing particle swarm optimization (PSO), as well as taking the artificial neural network (ANN) into account [16]. Besides, water flow predictions of rivers have been executed, utilizing the suggested models implementing RVM (relevance vector machine), in a study of Flake [17]. Also, in the estimation approaches, in addition to the RVM, the utilization of the learning machine of the probabilistic kernel-based has been actualized. Ultimately, Flake has concluded that a high precision can be acquired for obtaining estimations on the flows of river water, during the utilization of the RVM approaches [17]. In a study of Farhadi et al.; the support vector machine (SVM) has been proposed to be implemented safely as a technical method in obtaining anticipations of water flow in channels of straight composite structure [18]. In this study, they have shown that the statistical accuracy results obtained from this novel technique, used in comparisons were pretty ahead of the error outcomes of the conventional methods. Furthermore, in flow estimations for stream water, the ANN was also implemented. In such ANN applications, appropriate activation function as well as the correct hidden nodes were used to obtain a high accuracy flow forecasts [19]. They had demonstrated that during the implementation of the proper initial conditions and in the case of selecting correct adjustment parameters; on the forecasts of the flow of the river water, the performance of ANN predictions was significantly improved [19]. Additionally, in a model of flood hazard rating, the long short-term memory (LSTM) was suggested. The flood risk area, named as Samseong District was selected as the location of the LSTM study. Namely, in the study of Il Kim and Kim, appropriate learning algorithms for data predictions were aimed to be obtained for to identify the amount of the fall of the rain in the district that was reported to be risky for floods [20]. Another interested location was selected in Iraq, which is the Tigris river, in order to obtain predictions for the flow rates of the river in terms of the discharge water flow. In the study, a variety of diverse prediction models were compared in terms of the performances, namely, the SVM, the extreme learning machine (ELM), plus the generalized regression neural network (GRNN) were compared based on the statistical accuracy outcomes. Accordingly, Yaseen et al. have exhibited that forecasts belonging to the ELM model have demonstrated superior prediction outcomes than the results of the other estimation models obtained from SVM and GRNN [21]. The learning algorithms were also exerted in obtaining flood detections in the runoff forecasts. In this context, Xiao et al. have used the extreme learning machine (ELM), the backpropagation (BP) network, the wavelet neural network (WNN), plus the generalized regression neural network (GRNN), to perform forecasts in this area [22]. In this regards, the flow rate of the water else the water level both were anticipated for the interested river of Xijiang. The predictions, in this study, were conducted based on the data measured at the station of Wuzhou, found on this river of Xijiang. Xiao et al. have reported that the superior anticipation outcomes were acquired by the models of GRNN, in the case of the stream flow, whereas, best predictions were obtained by WNN models, in the case of the level of the water, in which both were evaluated for flood detections [22].

As a novelty, in this study, the water volumetric flow rate was predicted, by a variety of machine learning algorithms, using the historical data of water flow, considering the Maritsa river. On the other hand, the data that has been used in the predictions has been acquired from the Svilengrad measurement station, located on this Maritsa river. Accordingly, four machine learning methods have been followed. The forecasts were done initially using the fuzzy c-means (FCM) algorithm of the adaptive neuro fuzzy inference system (ANFIS), followed by the subtractive clustering (SC) and grid partitioning (GP) of ANFIS, and finally, the long short-term memory (LSTM) method.

Knowing the future values of river water flow rate based on past time series data; enables some preliminary plans to be made with the predictions to be obtained within the framework of flow rates that have not yet occurred but are likely to occur. For example, the operation of hydraulic dams depends on the stable water flow. In this context, the early detection of deviations in the stable hydraulic flow rate of water flowing upstream to the water turbines located in the dam body concrete gains importance. With the large amount of historical water flow rate (WFR) data at hand, the computer learns with machine learning techniques which times of the year the decrease of the water flow rate eventuates, and in this direction, the times when the water flow will decrease can be determined in advance. And, accordingly, it can be reported that the energy provided to the grid will decrease; and it can be mentioned officially that alternative energy supply channels should be turned to. On the other hand, since the periods when excessive increases in flow rate can occur will be learned by the computer via the machine learning as a result of long-term analyses; necessary precautions can be taken to discharge this excess water without damaging the surrounding cities, villages and the fields, at times when even the flow of excess water from the spillway will not be sufficient. In this way, the utilization of the machine learning algorithms in water flow rate predictions will assist in helping water management, dam operations, and the flood prevention.

2. Material and Method

Figure 2 summarizes the boundary conditions of the research. The aim of the study is to obtain an output of water flow rate (WFR) time-series function, shown as function g(t) in the figure described as input parameter of time-series, demonstrated in the same figure with the function of f(t). The learning task of the f(t) function regarding the logic of every instantaneous WFR data was accomplished by machine learning tools including fuzzy c-means (FCM), subtractive clustering (SC), and grid partitioning (GP) of adaptive neuro-fuzzy inference system (ANFIS), and besides by long short-term memory (LSTM) method, as indicated in Figure 2. This formed the methodology of the study.

On the other hand, the boundary conditions of these tools are explained in this figure. Namely, considering the FCM, the adjustments of $3 \leq HD \leq 10$ having $\Delta HD = 1$, and $2 \leq MFs \leq 10$ having $\Delta MFs = 2$; were implemented. Taking SC into account, the regulations of $3 \leq HD \leq 10$ having $\Delta HD = 1$, and $0.20 \leq IR \leq 0.28$ having $\Delta IR = 0.02$ as well as the case of $IR = 0.29$; were applied. When the GP algorithm was kept in view, the setting parameters included the cases of $3 \leq HD \leq 4$ having $\Delta HD = 1$, and $2 \leq MFs \leq 4$ having $\Delta MFs = 1$; were performed. On the other hand, when the LSTM was taken in cognizance of, the range for the number of the hidden layers (HL) was scaled to $5 \leq HL \leq 300$. More comprehensive explanations on these issues are provided in the following sub-sections. Using these setting parameters exerted to the input function of f(t), generated the output of g(t) function, and in this way, a total of 102 models were created. The error magnitudes between both functions were evaluated, checked and compared, using the error parameters including mean absolute error (MAE), root mean square error (RMSE), and the correlation coefficient (R), as also exhibited in Figure 2. And, detailed information on these error parameters are also given in the following sub-sections.

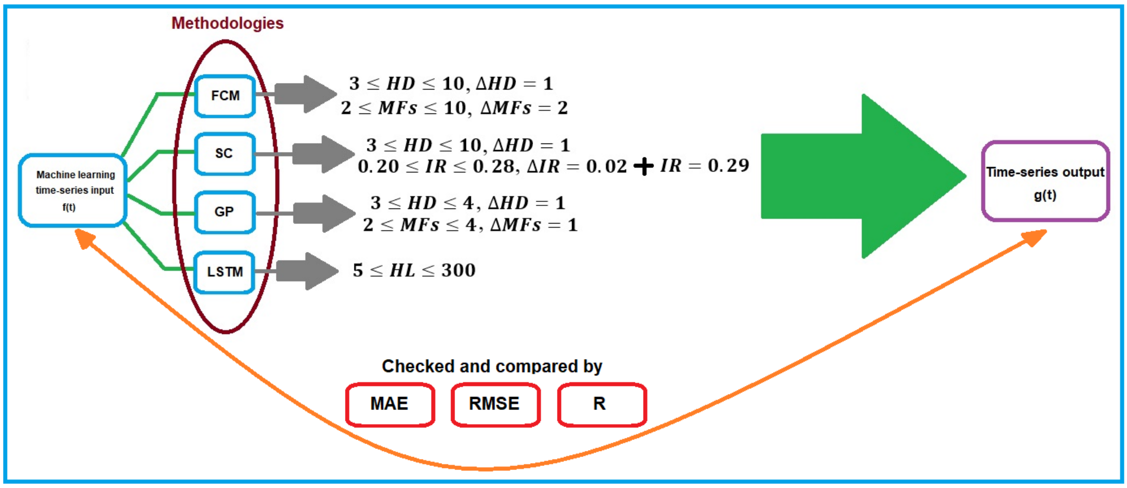

Figure 2. Boundary conditions of the research

2.1. Adaptive neuro fuzzy inference system (ANFIS)

The rapid computing algorithm that is used in all the World, namely, the adaptive neuro fuzzy inference system (ANFIS) is known to be an accomplished universal predictor. It provides an intense set-up, also allows any desired accuracy degree. On the other hand, it can be handled for any type of real observed function, given in continuous form of data cloud. Fundamentally, when the capabilities of the neural learning affixed to the fuzzy systems of Sugeno-type, ANFIS which is referred to be a network statement, will be structured. The rules provided in the form of fuzzy if-then will be created having proper MFs (membership functions), formed from input towards the output, during the employment of the learning algorithm of the neural network. The fuzzy inference system (FIS) development procedure which implements the frame of the adaptive neural network is named as ANFIS [23, 24]. The general structure of the ANFIS is presented either in detail or superficially in the literature [25, 26].

2.1.1. Fuzzy c-means of adaptive neuro fuzzy inference system (ANFIS-FCM)

Among different types of ANFIS, FCM is filing approach of ANFIS, permitting every data point to possess poly clusters and belong to different degrees of MFs. The essence of FCM consists of diminishing the target function. Information on FCM, similarly can be obtained from the literature [25, 26]. The FCM method is based on the minimization of the objective function, as shown in Eq. (1) [23, 24]:

$$J_m = \sum_{i=1}^{D}\sum_{j=1}^{N} \mu_{ij}^m \left\| x_i - c_j \right\|^2 \tag{1}$$

Here, in the above equation, the denotations of D and N, stand respectively for the amount of the data points else the clusters. On the other hand, the exponent of the fuzzy partition matrix is represented by the designation of m. The value of the m is greater than 1.

During forming the clusters, the initializing of the values of the cluster membership, μ_{ij} are performed randomly. Besides, the centers of the clusters are computed as indicated with the following equation:

$$c_j = \left(\sum_{i=1}^{D} \mu_{ij}^m x_i\right) / \left(\sum_{i=1}^{D} \mu_{ij}^m\right) \tag{2}$$

The following equation is considered for the update of μ_{ij}:

$$\mu_{ij} = \frac{1}{\left[\sum_{i=1}^{N} \left(\frac{\|x_i - c_j\|}{\|x_i - c_k\|}\right)^{\frac{2}{m-1}}\right]} \tag{3}$$

2.1.2. Subtractive clustering of adaptive neuro fuzzy inference system (ANFIS-SC)

The method of SC takes each data point as a center of candidate cluster into account. Additionally, the calculation of the potential for every data point is actualized by density measuring of the data point surrounding the center of the cluster [11].

The index of the density, shown by the abbreviation of D_i, that corresponds to the data x_i, is determined utilizing below equation,

$$D_i = \sum_{j=1}^{n} \exp\left(-\frac{\|x_i - x_j\|}{\left(r_a/2\right)^2}\right)^2 \quad (4)$$

The measurements of the density (x_i) for every data point are computed using equation,

$$D_i' = D_i - D_{c1} \exp\left(-\frac{\|x_i - x_{c1}\|}{\left(r_b/2\right)}\right)^2 \quad (5)$$

2.1.3. Grid partitioning of adaptive neuro fuzzy inference system (ANFIS-GP)

The method of GP allocates the space of the input data to rectangular subspace. In this division, a partition of axis-paralleled is used. The partition of each input is actualized by membership functions of identical shape. The amount of the rules of fuzzy if-then is identical to the value of M^n, here the n stands for the input dimension and the amount of the subsets for fuzzy partitions at each input variable corresponds to the value of M [11].

2.2. Long short-term memory (LSTM)

The method of LSTM was presented to the literature initially by Liu and Liu [27]. It is a type of recurrent neural network (RNN) which proposes solutions to problems through the addition of the cell states or the memory cells. However, this addition is executed with constant errors, in this way, the errors can be regenerated without having any disappear of the gradients. Furthermore, an LSTM consists of a total of three gates which are distinct. Among these three gates, the input gate is responsible of learning to protect the flow of constant error inside the memory cell, from inputs of unrelated type. Whereas, the output gate is responsible of learning to protect other units from storing of irrelevant memory content inside the memory cell. Terminally, learning how to control for the time duration of storing a value inside the memory cell, is accomplished by the forget gate [27].

The general structure of an LSTM is exposed schematically in Figure 3. This figure denotes X time-series flow having S-length with C properties, i.e., channels given along the layer of LSTM. Besides, in this figure, the output which is also referred to be the hidden state is abbreviated by the symbol of h_t, whereas, the status of the cell at a time step of t is curtailed by the symbol of c_t. In order to compute for the initial output as well as to update the status of the cell, the initial part of the LSTM is used to be considered for the initial status of the network as well as the former time-step of the series. At a time-step of t, this block of the LSTM utilizes the present status of the network, namely, c_{t-1} and h_{t-1} are used, as well as the following time step of the data set will be considered in order to compute for the output plus the updated status of the cell, i.e., the c_t will be reckoned according to knowledge received from former steps.

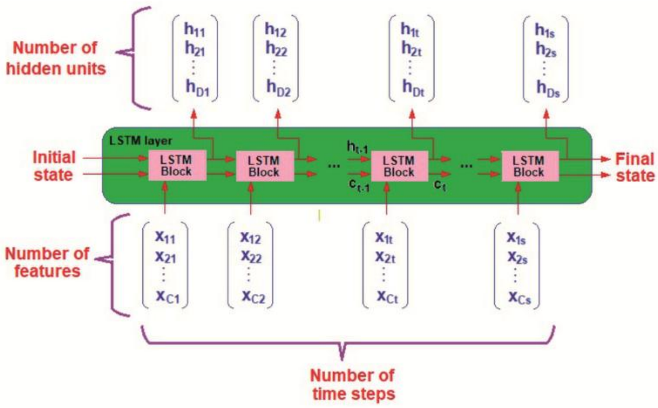

Figure 3. The schematic structure of layer of LSTM [28]

The layer state encloses the hidden state, which is also referred as the output state; else the status of the cell. At the time step of t of the hidden state, the output obtained from the LSTM layer at this time step will be included. The information extracted from the former time steps will be covered by the cell status. For every time step, the LSTM layer exerts the addition task of information on the cell status or carries out the duty of removing task of information from the cell status. The checking of such updates are handled by these introduced gates including the input gate, output gate, and the forget gate. In controlling of the cell status plus the hidden status of the layer, some constituents of the LSTM layer are utilized. Such as, in order to control the level for update of the cell status, the input gate, i operates, on the other hand, in order to check the level of cell status incorporated to the hidden status, the output gate, o performs the related task. Additionally, the grade of the reset of the cell status, that is referred as the forget of the cell is accomplished by the forget gate, f. Another gate is needed to be acquainted at this step. In this context, the instruction will be adjoined to the cell status by the cell candidate, g.

Furthermore, the data flow at a time interval of t is represented in Figure 4. How such gates including the forget gate, the update gate including the cell candidate and the input, as well as the output gate perform their tasks on the cell as well as on the hidden status, are indicated in this figure.

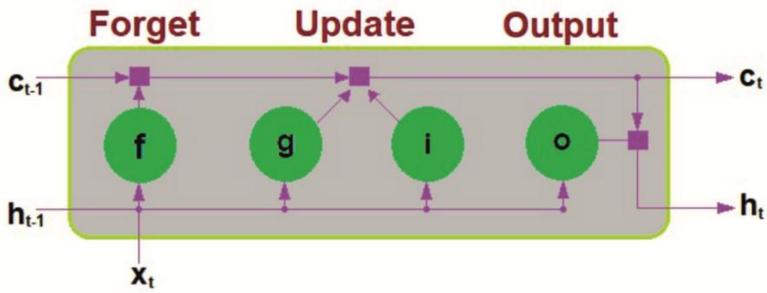

Figure 4. The data influx at a time interval of t

After LSTM gates are introduced, it is proper to introduce the weights assigned to these gates. The weights belonging to the input, shown by W, belonging to the recurrent, demonstrated by R, as well as the bias, indicated by b are referred to be the learnable weights for the layer of LSTM. Therefore, the concatenation of the matrices for those weight parameters of W, R, and b are exhibited in Eq. (6):

$$W = \begin{bmatrix} W_i \\ W_f \\ W_g \\ W_o \end{bmatrix}, R = \begin{bmatrix} R_i \\ R_f \\ R_g \\ R_o \end{bmatrix}, b = \begin{bmatrix} b_i \\ b_f \\ b_g \\ b_o \end{bmatrix} \tag{6}$$

In this equation, the input gate for the input weight, recurrent weight, and the bias is indicated by the subscript denotation of i, attached respectively to form W_i, R_i, and b_i. Whereas, the forget gate for the same weights and the bias is revealed by the subscript abbreviation of f, affixed respectively to constitute W_f, R_f, and b_f. The cell candidate given for the same weights and the bias is exhibited by subscript abridgement of g, annexed respectively to compose W_g, R_g, and b_g. Finally, the output gate which is called by the subscript symbol of o is used for the same weights as well as the bias, joined respectively to carve out W_o, R_o, and b_o. On the other hand, the status of the cell at the time step interval of t, is specified by Eq. (7);

$$c_t = f_t \odot c_{t-1} + i_t \odot g_t \tag{7}$$

Here, in Eq. (7), the product of the vectors provided in the form of element wise is indicated by the abridgment of $\odot$. Similarly, the status of the hidden layer, i.e., the h_t, at the time step interval of t is specified by Eq. (8),

$$h_t = o_t \odot \sigma_c(c_t) \tag{8}$$

In Eq. (8), the function for the state activation is indicated by the abbreviation of σ_c. Considering the management of the LSTM layer, by default, the state activation function is computed through the utilization of the function of the hyperbolic tangent. In Eq. (11), the relation of the cell candidate gate (g_t) at the time step of t is identified by the state activation function, σ_c. Following four equations describe the components at a time step of t, given for the formerly defined four gates [27]:

$$i_t = \sigma_g(W_i x_t + R_i h_{t-1} + b_i) \tag{9}$$

$$f_t = \sigma_g\left(W_f x_t + R_f h_{t-1} + b_f\right) \tag{10}$$

$$g_t = \sigma_c\left(W_g x_t + R_g h_{t-1} + b_g\right) \tag{11}$$

$$o_t = \sigma_g\left(W_o x_t + R_o h_{t-1} + b_o\right) \tag{12}$$

Taking Eqs. (9), (10), and (12) into account, the compendium of σ_g used in these equations, exhibits the function of the gate activation. These equations respectively characterize the input gate (i_t), the forget gate (f_t), and the output gate (o_t), at a time step of *t*. Besides, the LSTM layer utilizes the function of the sigmoid in order to compute the function of the gate activation. In this context, it is expressed by Eq. (13);

$$\sigma(x) = (1 + e^{-x})^{-1} \tag{13}$$

2.3. Statistical accuracy analysis

In the current study, the evaluation of the statistical accuracy was executed by the comparisons of the anticipation data with respect to the real observed counterparts and was achieved utilizing the mean absolute error (MAE), root mean square error (RMSE), plus the coefficient of the correlation (R). Therefore, the precisions of these error criteria were controlled depending on the comparisons of the anticipations crosschecked against the real measured variables. Accordingly, Eqs. (14), (15), and (16) are considered for this purpose of statistical accuracy evaluation, corresponding respectively to MAE, RMSE, and R:

Mean absolute error (MAE):

$$MAE = \frac{1}{M}\sum_{j=1}^{M} |a(j) - r(j)| \tag{14}$$

Root mean square error (RMSE):

$$RMSE = \sqrt{\frac{1}{M}\sum_{j=1}^{M} [a(j) - r(j)]^2} \tag{15}$$

Correlation coefficient (R):

$$R = \left(\sum_{j=1}^{M} [a(j) - \bar{a}][r(j) - \bar{r}]\right) \Big/ \left(\sqrt{\sum_{j=1}^{M} [a(j) - \bar{a}]^2}\sqrt{\sum_{j=1}^{M} [r(j) - \bar{r}]^2}\right) \tag{16}$$

Here, *a(j)* and *r(j)*, exploited in three of the equations above, stand for the abbreviations of the anticipated instantaneous value as well as the real measured instantaneous value. On the other hand, the dashed abbreviations of those symbols respectively correspond to average value of the anticipated data cloud and the mean value of the real data cluster, given respectively as $\bar{a}$ and $\bar{r}$, and those were included only in Eq. (16). On the other hand, in order to call any instant data sample of both anticipated and real data clusters, the *j* indicator is utilized in three of the above equations, having a value analysis in the data set range of $1 \leq j \leq M$. Therefore, the summation signs employed in three of the above equations are already compatible with this data set range of $1 \leq j \leq M$. Finally, the cumulative amount of the data in these three equations is demonstrated by the curtailment of M.

Due to the nature of the estimation algorithms, it is not possible to obtain a perfect estimation on the actual observed values. Because, algorithms learn how the data shows a sequence by starting from the past data set and produce their own values by interpreting on this. While the data sequence is made regarding the produced estimation values, a mathematical relationship is created in between. As in the traditional linear and curvilinear regressions, in this context, it is not possible to produce estimation instantaneous values that perfectly overlap on the real instantaneous values. However, the aim here is to minimize the error values. In other words, bringing the statistical error values such as mean absolute error (MAE) and the root mean square error (RMSE) closer to zero, and increasing the values such as the correlation coefficient (R) indicating the fit of the data closer to one, increases the quality of the estimation functions, in this frame. These improvements in error values can be achieved by increasing the total number of the historical data samples that is used in the data cluster for training of the algorithms, by selecting the correct estimation algorithm, and by ensuring that the tuning parameters of the estimation algorithms are set to their correct values. Besides, if a prediction function is to be produced depending on the predicted output values defined according to certain physical input parameters; choosing the correct and sufficient amount of physical input parameters and the most effective ones that influence the correct prediction results of the output parameter, will also help to minimize the statistical error values. In such functions defined according to certain input physical parameters, while the situation regarding the wrong or excessive selection of physical inputs is sometimes expected to increase the quality of the prediction, on the contrary, it deteriorates the quality of the predictions.

3. Results

3.1. Interpretation of data and interested area location

The study was performed taking the hydraulic data of Maritsa river into account. In Turkish, this river name is provided with the word of Meriç, shown on Figure 5. Maritsa river is found in the northwest region of Türkiye and this river emerges in Bulgaria, gets in Türkiye via Bulgaria, then flows in the direction of Edirne province of the country and continues its route towards the Aegean Sea to be eventually poured to this sea. To provide further elaboration on this river route; Maritsa initially gets mixed with the Arda river, in the north of Edirne inside the purlieu of Türkiye. This new mixed river including the waters of both rivers is referred to be Maritsa, again. Later, at the south of Edirne, Tunca river joins with Maritsa. The new formed river including the waters received from three of them together is again named as a single river called Maritsa river. Finally, the Ergene river that is flowing from upstream north east direction to the downstream south west direction and Maritsa river formed previously from the constitution of the water flows of three rivers, are mixed in the rural of Adasarhanlı, that has a distance of thirty kilometers referenced to the İpsala city center. Before this final shuffle of Ergene river and the Maritsa river, Maritsa flows through downstream south direction by drawing the Turkish-Greek border for a long time. On the other hand, the volumetric water data utilized in this study has been acquired from the internet site of the 11th regional directorate of the institution of DSI. The DSI is a government agency of Türkiye that is occupied with most of the hydraulic issues of the country. Accordingly, the water flow data of the Maritsa river, used in the current study, has been acquired from the measurements taken from the Svilengrad measurement station, found on Maritsa river. This measurement station is situated in the Svilengrad town of Haskovo city of Bulgaria. It is situated at the border of Bulgaria, Türkiye, and the Greece. In this regards, Figure 4 presents the exact location of this interested area, taken into account for the RWFR simulations. In this figure, the Maritsa river is provided on the map and the Svilengrad measurement station seated on this river is demonstrated with a large pink circle. Besides, the other three rivers consisting of Arda, Tunca, and Ergene; all merging ultimately with the Maritsa, are clearly shown from their upstream paths flowing towards on their downstream paths.

All instantaneous data readings were read manually and daily, one by one, from the website of the Turkish DSI institution and those were transferred to the Excel program [29]. Thus, the real observed and measured water flow rate (WFR) data folder was created in this way, to which machine learning techniques were later implemented. After the application of the machine learning methods, the simulated estimated values were later added to this real data folder, and the real value function and the corresponding estimated functions were plotted on this data folder, and as a result, and ultimately, the compatibility and overlap between the predicted curves and their real counterparts were obtained using statistical error methods and the outputs of the current study were revealed. Namely, in this machine learning study, a cumulative of 1,465 daily measurements have been taken into account for the 102 models of the machine learning computations. On the other hand, the historical data covers the instantaneous time series distribution of past 4 years, considered as inputs to the machine learning methods. The data is related on the volumetric flow rate of the river water flow. In this measurement station, the measurement of the rate of the river water flow was usually performed twice in one day, that was executed at 08.00 as well as 16.00 o'clock.

In the daily measurements carried out twice a day, it was observed that there were not many sudden changes in the water flow values and therefore it is reported that the measurements during the day deviated from the average RWFR values by a small amount at most. On the strength of this situation, the study was started by taking arithmetic average of flow rate of the river, performed for each day; therefore, a new data set was created. Namely, a reconstituted data cloud for four years, including an average water flow data at each day was formed. On the other hand, the volumetric measurement data physical unit was m^3/s, as provided by this institution of DSI. Accordingly, the unit studied in the machine learning simulations was also materialized in this direction. Besides, a correct allocation of the data separated for training and testing of the algorithms was aimed to be compatible with the literature studies. In this frame, among the total of 1,465 daily average data, belonging to the new data cluster; while 80% of all was utilized to obtain the training of the ANFIS-FCM, SC, GP, and LSTM algorithms, the rest of 20% of the cumulative data set was used for testing of these same ANFIS and LSTM algorithms.

Figure 5. Svilengrad measurement station in the border of Türkiye and Bulgaria, situated on the Maritsa river

The whole data cluster considered in the study that includes the total of 1,465 data samples and which was allocated to two sub-clusters consisting of training data cluster and the testing data cluster parts, is demonstrated in Figure 6. In this regards, while the training data cluster is exhibited in this figure by yellow color, the testing data cluster is shown in the same figure by the black color.

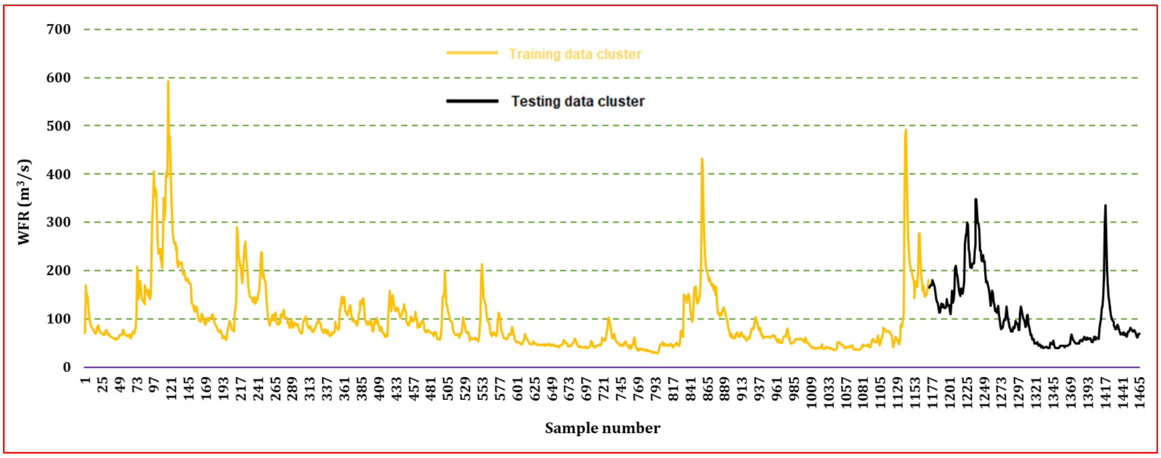

Figure 6. Training and testing allocations of WFR data

Table 1. Statistical information of the data cluster

Type of data	Information response
Smallest value of the cluster (m³/s)	28.50
Greatest value of the cluster (m³/s)	592.00
Arithmetic average value of the cluster (m³/s)	98.50
Standard deviation (m³/s)	67.83
Total number of data in the cluster (entire set-100%)	1,465
Total number of training data in the cluster (80% of the set)	1,172
Total number of testing data in the cluster (20% of the set)	293

Table 1 points the statistical information on the utilized data of all data cluster formed by the measurements performed in Svilengrad WFR measurement station. The amount of data samples considered in training of the machine learning algorithms correspond to 1,172, while the amount of data samples taken into account for testing of the machine learning algorithms equal to 293. On the other hand, the value range of all data samples is described by the relation of 28.50 $m^3/s \leq WFR \leq$ 592.00 m^3/s. The arithmetic average value of the data cluster, regarding 1,465 total observations correspond to the value of 98.50 m³/s. The standard deviation which describes the total deviation of the values of the data samples from the mean normalized for a single sample, namely which indicates the mean deviation of the whole data cluster from the average value is also indicated in Table 1. In this context, it is concluded that the mean deviation of the whole data cluster from the average value is reported to be 67.83 m³/s.

3.2. The applied forecasting methods

3.2.1. Fuzzy c-means (FCM) of ANFIS

Considering fuzzy c-means (FCM) analysis of ANFIS, the historical data (*HD*) considered during training of the tested models were altered in between 3 and 10, with 1 increment increase, applied to the former value to find for the trailing value of *HD*. Besides, 5 different number of MFs was tried for every value of the historical data (*HD*), and those were changed between 2 and 10, with 2 increment increase, implemented to the former value to find for the trailing value of MFs. Therefore, a cumulative of 40 different models were forged for the FCM algorithm of ANFIS, however the maximum epoch number was set to the same value of 100 for these 40 different FCM models. The parameters considered for set-up of the models including *HD*, MFs, and maximum epoch number as well as the output statistical accuracy results including MAE, RMSE, and *R* are indicated in Table 2.

The best outcome of every number of historical data (*HD*) has been represented utilizing bold format in Table 2, that is generally exhibiting the lowest values on the MAE as well as RMSE, whereas demonstrating the highest values on the *R*, i.e., the best result for every *HD* group is displayed by bold color. These cases of MAE, RMSE, and *R* for the best model of every *HD* group stand for the highest harmony of real and the forecasted data of the *HD* group in question, and offers the best correlation of the instantaneous scatter for the same *HD* group. On the other hand, the leading result of all bests for every historical data number utilized for training has been presented by simultaneous patterns of bold and italic color in this table. Besides, during the adjustment of *HD* to 3 as well as setting MFs at 8 have turned the best outcome for this number of *HD* value. Namely, MAE of 3.21 m^3/s, RMSE of 5.44 m^3/s, and *R* of 0.9967 was computed, at 3 *HD* and 8 MFs. On the other hand, when the number of the *HD* was set at 4, with the simultaneous adjustment of the value of the MFs to 2, given as the first rank of this *HD* group; has turned with the statistical error results of MAE, RMSE, plus *R* to correspond, respectively to 3.25 m^3/s, 4.88 m^3/s, and 0.9977, to be best reported for this *HD*=4. In the case of the *HD* to be arranged to correspond to 5 as well as when MFS was arranged to be 2; 3.13 m^3/s, 4.90 m^3/s, as well as 0.9978 values, were respectively computed for MAE, RMSE, plus *R*, to be reported as superior for the number of *HD*=5. Likewise, the *HD*=6 and MFs=2 have yielded 3.16 m^3/s for MAE, 4.97 m^3/s for RMSE, and 0.9975 for *R*, to be concluded as the best outcome for the historical data collimation of *HD*=6. Additionally, among five different MFs trials of *HD*=7; the best forecast outcome was acquired at MFs=2; resulting 3.17 m^3/s on MAE, 5.11 m^3/s on RMSE, and 0.9973 on *R*. As in the case of previous *HD* numbers except the regulation of *HD*=3, at the number of the historical data corresponding to *HD*=8, again the best statistical accuracy result was handled at MFs=2; namely, providing the statistical accuracy parameters of 3.25 m^3/s at MAE, 5.18 m^3/s at RMSE, and 0.9972 at *R*. Finally, the set of MFs=2 twice, in order for *HD*=9 and *HD*=10; have effectuated the superior outcomes for those historical data values. Accordingly, initially at *HD*=9, 3.38 m^3/s at MAE, 5.29 m^3/s for RMSE, and 0.9971 on *R* were obtained. Followed by, 3.48 m^3/s upon MAE, 5.65 m^3/s upon RMSE, and 0.9967 upon *R* have been achieved at *HD*=10. However, among these 8 different trials of *HD* and 5 different trials of MFs for every considered *HD* value; namely, among considered 8x5=40 models in the total, the best model in terms of the statistical errors were reported to correspond to the model arranged at the adjustments including 5 *HD* and 2 MFs, that have turned with the computing results to correspond to 3.13 m^3/s MAE, 4.90 m^3/s RMSE, and 0.9978 *R*. Under the light of these outcomes, while the best results corresponding to a certain value of MFs for each *HD* studied in between $3 \leq HD \leq 10$, except *HD*=5 were indicated with bold patterns in Table 2; however, since, *HD*=5 has generated the superior result among the other *HD* numbers, the statistical accuracy results for this *HD* value were exposed using the bold pattern plus the italic representation in this table to be distinguished from the others. Accordingly, in this wise, the best model for each *HD* was distinguished from the other four models in the *HD* group that is in question. On the other hand, the best model of Table 2 was also distinguished from the remaining 39 models.

Table 2. FCM algorithm results of ANFIS methodology

HD	MFs	Max epoch	MAE (m^3/s)	RMSE (m^3/s)	R
3	2	100	3.24	5.54	0.9969
	4	100	3.43	5.58	0.9966
	6	100	3.24	6.05	0.9959
	8	**100**	**3.21**	**5.44**	**0.9967**
	10	100	3.83	6.96	0.9947
4	**2**	**100**	**3.25**	**4.88**	**0.9977**
	4	100	3.25	5.65	0.9965
	6	100	3.48	5.70	0.9964
	8	100	3.96	7.44	0.9949
	10	100	4.40	7.70	0.9936
	2	***100***	***3.13***	***4.90***	***0.9978***
	4	100	3.22	5.88	0.9962

5	6	100	3.62	6.35	0.9955
	8	100	4.29	8.13	0.9938
	10	100	4.73	8.84	0.9917
6	**2**	**100**	**3.16**	**4.97**	**0.9975**
	4	100	3.34	6.13	0.9959
	6	100	5.58	9.88	0.9901
	8	100	4.53	8.47	0.9936
	10	100	5.54	9.84	0.9904
7	**2**	**100**	**3.17**	**5.11**	**0.9973**
	4	100	3.62	6.37	0.9955
	6	100	3.87	7.14	0.9944
	8	100	4.39	9.20	0.9923
	10	100	5.97	10.82	0.9894
8	**2**	**100**	**3.25**	**5.18**	**0.9972**
	4	100	4.04	7.00	0.9946
	6	100	3.95	7.28	0.9941
	8	100	5.17	10.16	0.9907
	10	100	5.81	11.73	0.9880
9	**2**	**100**	**3.38**	**5.29**	**0.9971**
	4	100	4.36	7.40	0.9939
	6	100	4.22	7.44	0.9939
	8	100	5.60	10.72	0.9895
	10	100	6.96	13.13	0.9849
10	**2**	**100**	**3.48**	**5.65**	**0.9967**
	4	100	4.25	7.16	0.9943
	6	100	4.49	7.88	0.9931
	8	100	5.48	10.62	0.9889
	10	100	7.11	13.87	0.9818

Figures 7 and 8 show the data distributions of the real data cloud simultaneously demonstrated with its counterpart of FCM anticipated data function and the correlation results of the instantaneous data readings indicated against the predictions, respectively. Furthermore, the average value of the FCM correlation outcome of the best result obtained among different values of *HD* and MFs trials, namely, corresponding to the statistical error results captured at 5 *HD* and 2 MFs, has been also exhibited in Figure 8, corresponding to the value of 0.9978 *R*. Both figures indicate the high compatibility of the real data versus its counterpart WFR anticipations. Namely, the anticipation function presented in Figure 7 with dashed black colored curve, overlaps on the real data cloud of continuous yellow colored curve, sufficiently and with a high quality. Besides, this situation is also observed from Figure 8, revealing the correlation scatter diagram of both actual and guess functions. Namely, all instantaneous data samples having two roots, one corresponding to the real data cloud, the other stands for the anticipated data cluster, are observed to be scattered all in the vicinity of the regression line, that is exhibited by dotted blue colored line. As well as, the regression line almost resembles the symmetry line of the real and predicted data axes.

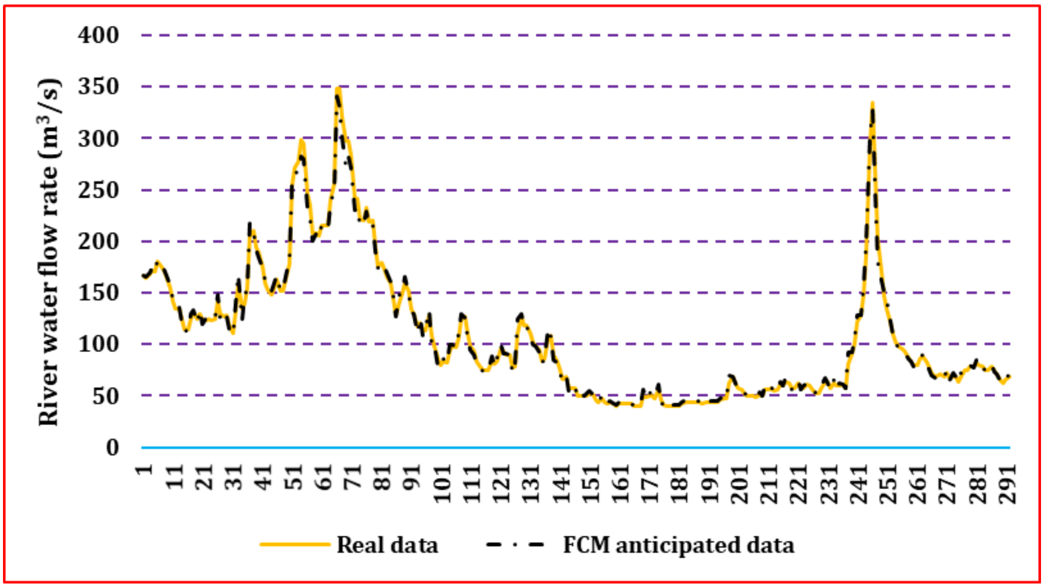

Figure 7. Real flow data (WFR) given according to FCM anticipations

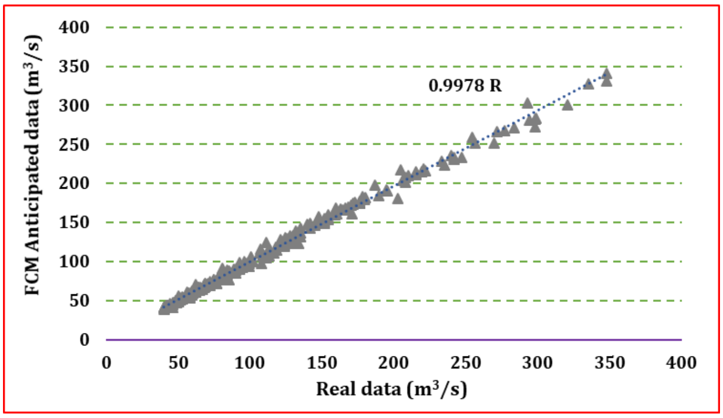

Figure 8. The FCM correlation scatter for water flow time-series

3.2.2. Subtractive clustering (SC) of ANFIS

The ANFIS subtractive clustering simulation results are brought in Table 3, in terms of the mean absolute error (MAE), root mean square error (RMSE), and the correlation coefficient. Selected number of historical data (*HD*) at training of the models was considered in the value range of 3 and 10, again by 1 increment realized to the previous value to call for the following selected number of *HD*. Additionally, six different influence radius (*IR*) values were studied for every selected number of historical data in SC algorithm. The tested influence radius (*IR*) values for every number of *HD* were chosen to fall within the range of 0.20 and 0.29, and to describe the ensuing *IR* value, the relation $IR_{next} = IR + 0.02$ was used, except for the relation between last two ones described by $IR_{next} = IR + 0.01$. In this way, a cumulative of 8x6=48 models were created, and the epoch number was bound to same value of 100. The algorithm adjustment parameters including the selected number of historical data (*HD*) used for training of the algorithm, influence radius (*IR*) values, and the maximum epoch number, as well as the statistical accuracy error results including MAE, RMSE, and *R* are all shown for different models in Table 3.

In Table 3, the best outcome of every *HD* number of SC method is denoted by only bold color, whereas, the best result of the bests is depicted by simultaneous usage of bold color and italic pattern in this table. In this context, the superior result of all trials was produced during the adjustment of *HD*=4 and *IR*=0.20. The model generated at these adjustments as well as at 100 epoch number turned the statistical accuracy error results of MAE, RMSE, and *R*, corresponding to 3.04 m^3/s, 4.91 m^3/s, and 0.9979, respectively. On the other hand, the other good results one step below the best result, that is, the best results for each *HD* number, were obtained as follows, and those are indicated in Table 3 using bold color: Namely, at *HD*=3 and *IR*=0.20, statistical accuracy outcomes of 3.05 m^3/s MAE, 4.94 m^3/s RMSE, and 0.9979 *R* were achieved, as the best output of the selected number of historical data (*HD*) used for training corresponding to 3. Besides, 3.04 m^3/s MAE, 5.06 m^3/s RMSE, and 0.9978 *R* were acquired at 5 *HD* and 0.22 *IR*, and this model was reported to be the best model of *HD*=5 models. Considering six different models of *HD*=6 model series, MAE, RMSE, and *R* corresponding to 3.11 m^3/s, 5.07 m^3/s, and 0.9978 were attained at 0.26 *IR*, and this model was concluded to be the best model of these *HD*=6 models. Statistical error results of 3.21 m^3/s, 5.07 m^3/s, and 0.9979 *R* were reached as best model at *IR*=0.20, respectively for MAE, RMSE, and *R* error values, at *HD*=7 model series. Taking six different models of *HD*=8 into account, it was concluded that the best model was obtained during the regulation of the influence radius either to *IR*=0.20 and *IR*=0.22, that eventuated 3.22 m^3/s MAE, 5.08 m^3/s RMSE, and 0.9981 *R*. The number of *HD* and *IR*, respectively set to 9 and 0.24, caused the best model having 3.33 m^3/s MAE, 5.32 m^3/s RMSE, and 0.9978 *R*. Last group of SC models including six models formed at 10 *HD*, resulted the best model for this group at 0.22 *IR*, generating 3.36 m^3/s MAE, 5.38 m^3/s RMSE, and 0.9977 *R*.

Table 3. SC algorithm results of ANFIS methodology

HD	IR	Max epoch	MAE (m^3/s)	RMSE (m^3/s)	R
3	**0.20**	**100**	**3.05**	**4.94**	**0.9979**
	0.22	100	3.08	4.97	0.9978
	0.24	100	3.10	4.98	0.9977
	0.26	100	3.10	4.99	0.9977
	0.28	100	3.09	4.98	0.9977
	0.29	100	3.11	5.00	0.9977
4	***0.20***	***100***	***3.04***	***4.91***	***0.9979***
	0.22	100	3.04	4.92	0.9979
	0.24	100	3.03	4.92	0.9979

	0.26	100	3.04	4.93	0.9979
	0.28	100	3.07	4.95	0.9978
	0.29	100	3.05	4.94	0.9978
5	0.20	100	3.06	*5.06*	*0.9978*
	0.22	**100**	**3.04**	**5.06**	**0.9978**
	0.24	100	3.04	5.07	0.9977
	0.26	100	3.03	5.07	0.9977
	0.28	100	3.02	5.08	0.9977
	0.29	100	3.02	5.08	0.9977
6	0.20	100	3.14	5.07	0.9978
	0.22	100	3.13	5.08	0.9978
	0.24	100	3.12	5.08	0.9978
	0.26	**100**	**3.11**	**5.07**	**0.9978**
	0.28	100	3.15	5.12	0.9977
	0.29	100	3.14	5.10	0.9977
7	**0.20**	**100**	**3.21**	**5.07**	**0.9979**
	0.22	100	3.24	5.06	0.9979
	0.24	100	3.19	5.11	0.9978
	0.26	100	3.20	5.11	0.9978
	0.28	100	3.22	5.10	0.9978
	0.29	100	3.22	5.10	0.9978
8	**0.20**	**100**	**3.22**	**5.08**	**0.9981**
	0.22	**100**	**3.22**	**5.08**	**0.9981**
	0.24	100	3.24	5.11	0.9980
	0.26	100	3.26	5.15	0.9979
	0.28	100	3.28	5.25	0.9976
	0.29	100	3.29	5.25	0.9976
9	0.20	100	3.41	5.47	0.9973
	0.22	100	3.41	5.46	0.9973
	0.24	**100**	**3.33**	**5.32**	**0.9978**
	0.26	100	3.34	5.35	0.9977
	0.28	100	3.40	5.45	0.9974
	0.29	100	3.40	5.47	0.9973
10	0.20	100	3.32	5.39	0.9976
	0.22	**100**	**3.36**	**5.38**	**0.9977**
	0.24	100	3.44	5.48	0.9974
	0.26	100	3.45	5.50	0.9974
	0.28	100	3.37	5.38	0.9976
	0.29	100	3.42	5.45	0.9974

The mutual data distributions which consist of the actual data function given with respect to the counterpart SC anticipations are reported in Figure 9. On the other hand, the correlation scatter of the real data versus the predictions is demonstrated in Figure 10 for the forecasts obtained using the SC algorithm. The simulation results with reference to these computations, namely, the best SC outcome that was obtained as soon as the SC adjustment parameters were set to 4 for *HD*, 0.20 for *IR*, and maximum epoch of 100, and that has resulted the statistical accuracy results of 3.04 m^3/s MAE, 4.91 m^3/s RMSE, and 0.9979 R, are shown in these two figures. On the other hand, the mean value of the correlation coefficient (R) is also pointed in Figure 10, to correspond to the value of 0.9979 R. In both Figures of 9 and 10, the best SC prediction concordance of WFR on the factual observed WFR counterparts are observed. The function of the forecast shown in this Figure 9, utilizing the curve of red color and dashed type, face coincidently the actual measured WFR function, which is shown in the same figure, using the curve of purple color and continuous type. Additionally, this case is also monitored in Figure 10, in which the figure stands for the scatter data distributions of actual and forecast instantaneous data. These instantaneous data distributions which each represent the actual reading on the x-axis versus the predicted virtual reading on y-axis are scattered so close to the regression line, that is demonstrated by the dotted burgundy colored line. Similarly, regression line also serves as the symmetric line of the data axes including the real and the forecast readings.

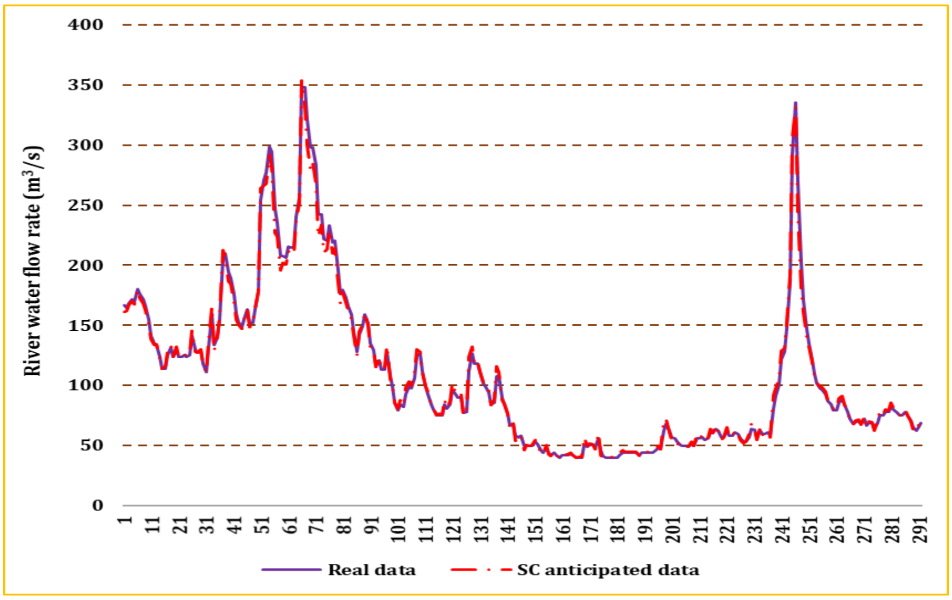

Figure 9. Real flow data (WFR) given according to SC anticipations

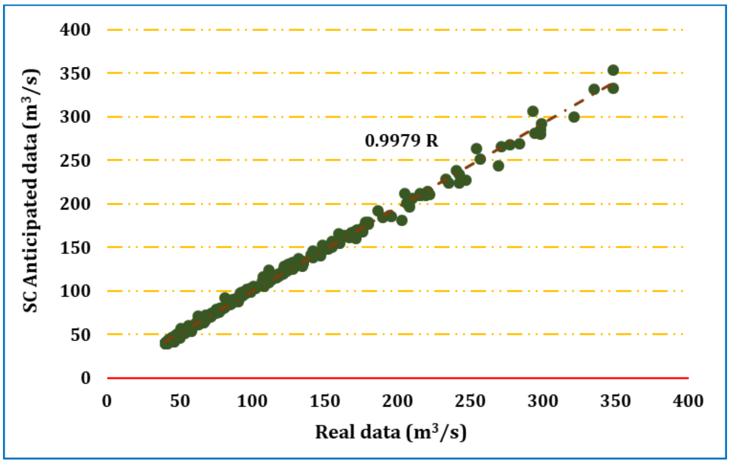

Figure 10. The SC correlation scatter for water flow time-series

3.2.3. Grid partitioning (GP) of ANFIS

The simulation outcomes of grid partitioning (GP) tool of ANFIS are given in Table 4. Similarly, in this table, the statistical error displays are given in terms of the parameters including the mean absolute error (MAE), root mean square error (RMSE), and the correlation coefficient. However, during the GP training stages of the models, the number of the historical data (*HD*) was selected at a narrow range, such as it was considered to be in the value range of $3 \leq HD \leq 4$. Accordingly, two *HD* values were studied in this value range of *HD*, including *HD*=3 as well as *HD*=4. Intercalary, the number of the membership functions (MFs) was studied in the value range of $2 \leq MFs \leq 4$, for *HD*=3; whereas, the number of the membership functions (MFs) was taken into account in the value range of $2 \leq MFs \leq 3$, for *HD*=4. In this way, three models for *HD*=3, plus two models for *HD*=4 were constituted, making a total of five models for GP of ANFIS. In order to reduce the computation time of the models, the epoch number was uniformed to the constant value of 30 for all models. On the other hand, the algorithm tuning parameters including the chosen number of historical data (*HD*) considered for algorithm training, the amount of the membership functions (MFs), and also maximum epoch numbers, as well as the error results for statistical accuracy parameters involving MAE, RMSE, and *R* are demonstrated in total in Table 4.

The best result of each *HD* number of GP method is indicated with the utilization of the bold pattern, besides, the best result of the bests is signified by together usage of bold and italic patterns simultaneously in Table 4. In this regards, the upper result of five computation trials was generated at the arrangement of the number of historical data (*HD*) and the number of the membership functions (MFs), respectively to correspond to 3 and 2. The model that was achieved at these arrangements and at the tuning of 30 epoch number accomplished the error results of the statistical parameters encapsulating 3.53 m^3/s, 5.79 m^3/s, and 0.9964 *R*. Likewise, the other good outcome being one step below the superior result, namely, the best result obtained at the number of *HD*=4 that is presented in Table 4 using bold color, has given statistical accuracy results of 7.70 m^3/s, 21.48 m^3/s, and 0.9501 *R*.

Table 4. GP algorithm results of ANFIS methodology

HD	MFs	Max epoch	MAE (m^3/s)	RMSE (m^3/s)	R
3	***2***	***30***	***3.53***	***5.79***	***0.9964***
	3	30	15.15	79.75	0.5919
	4	30	76.84	502.00	0.1246
4	**2**	**30**	**7.70**	**21.48**	**0.9501**
	3	30	98.30	618.61	0.1499

Visual monitor of the GP results given in Table 4 is exposed in Figure 11, for the function of the WFR forecasted data versus the counterpart function of the real measured WFR data. Moreover, the scatter of correlations of the actual data function given with respect to the forecasts is manifested in Figure 12, regarding the GP outlooks. The best outcome of the simulations considering these computations, i.e., the topnotch GP result that was acquired during the GP arrangement parameters were set to 3 *HD*, 2 MFs and 30 maximum epoch, and which has eventuated the statistical errors of 3.53 m^3/s MAE, 5.79 m^3/s RMSE, and 0.9964 *R*, are disclosed in these Figures of 11 and 12. Extra, the average value of the correlation coefficient (*R*) for this best model can withal be monitored in Figure 12, namely, the figure presents the value of 0.9964 *R* attached on it. In these two figures, the best GP harmony of WFR forecasts given on the authentic measured WFR counterparts are visualized.

The forecast function that is depicted in Figure 11 uses the color of purple being in the style of dashed format, whereas the real data function that is shown in the same figure utilizes the color of yellow being in the style of continuous form. Although the success of tracking capability of the forecast function on the real data function is not as successful as the former models of FCM and SC of ANFIS, it is still concluded that the GP forecast function tracks the real data function with a high success and performance. In addition to this situation, the similar conclusion can also be derived from Figure 12 which demonstrates the data disperse of the real observation versus the forecasted instantaneous WFR data. The WFR data scatters given in instantaneous form, represent the *x*-axis reading to correspond to real data, whereas, the *y*-axis reading to correspond to predictions, are generally close to the regression line, which is exhibited by the dotted red colored line. But, due to relatively lower value of the mean correlation coefficient corresponding to 0.9964 *R*; especially considering the sample number increase over the value of 200 has caused somewhat more deviations from the regression line.

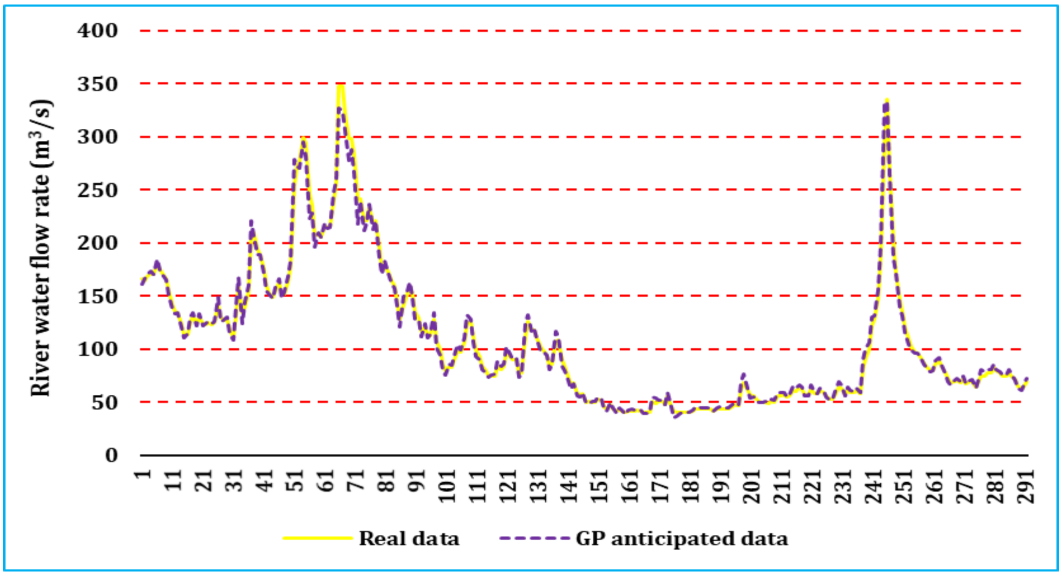

Figure 11. Real flow data (WFR) given according to GP anticipations

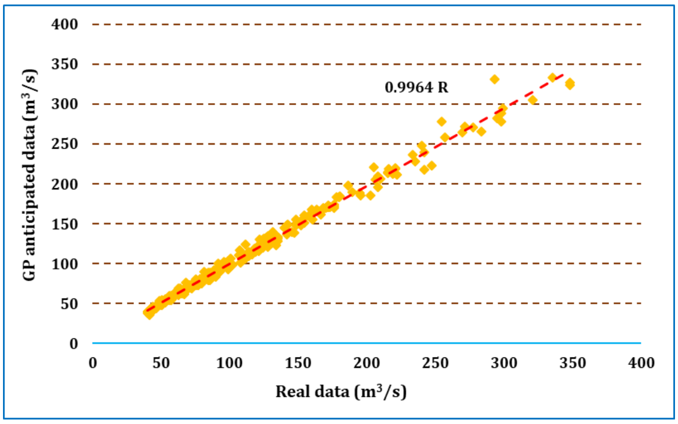

Figure 12. The GP correlation scatter for water flow time-series

3.2.4. Long short-term memory (LSTM)

Furthermore, the statistical accuracy outcomes of LSTM algorithm have been signified in Table 5. In this regards, the hidden layer (HL) was investigated in the range of $5 \leq HL \leq 300$. Since, the computation time of LSTM models was shorter and therefore LSTM turned the forecast results of the examined models so rapidly, the maximum epoch number was arranged to a constant of 300, as revealed in Table 5, in order to obtain more converging results. Due to this situation, the maximum epoch number of LSTM was chosen to be higher than the maximum epoch number of ANFIS algorithms. As pointed out in Table 5, a total of 9 different LSTM models were constituted. Interesting results were captured from LSTM models, in terms of the statistical error values. The minimum, namely the best MAE result was obtained when the HL was arranged to 10, which resulted 3.29 m^3/s MAE. Besides, again the minimum, i.e., the superior RMSE outcome was acquired during the adjustment of HL to 5, that generated 5.44 m^3/s RMSE. And, ultimately, the maximum, thus the strongest correlation coefficient (R) was produced as long as the number of the HL was set to 75. This value of HL calibration has propagated the correlation coefficient at 0.9990 R. Namely, in a total of 3 different cited models, every model has turned with only one best statistical accuracy parameter, among 3 different statistical accuracy parameters including MAE, RMSE, and R, taken into account in model comparisons. For this reason, among these three models including HL=5, HL=10, and HL=75, that could not be decided in terms of the superiority, two among three were chosen randomly and continuously, and a decision was made. Following this approach initially, the comparison was conducted regarding the models that were formed with regards to the number of HL=5 and HL=10. While MAE result of HL=10 model, which is 3.29 m^3/s was concluded to be smaller, so better than the MAE of HL=5 model, that is 3.38 m^3/s; whereas, the RMSE and R results of HL=5 model were reported to be superior than HL=10 model. Namely, RMSE of HL=5 model was computed as 5.44 m^3/s, being smaller and better than RMSE outcome of HL=10 model, being 6.01 m^3/s, as well as the correlation coefficient (R) of HL=5 model was determined as 0.9968, being bigger and so superior than the correlation coefficient (R) result of HL=10 model, being 0.9961. All in all, since two statistical accuracy parameters among three were observed to be better on the behalf of model having the number of HL=5, during the comparison stage of models of HL=5 and HL=10; it was concluded that model of HL=5 was better than model of HL=10. In the following situation and in a similar approach, model of HL=5 and model of HL=75 were compared. Accordingly, in the second stage of comparison for these two models, MAE and RMSE outcomes of HL=5 model was computed and thus reported to be smaller and better than MAE and RMSE results of HL=75 model. Although correlation coefficient (R) result of HL=75 model is bigger and better than the correlation coefficient (R) outcome of HL=5 model, it was reported that HL=5 model has satisfied better accuracy in terms of two statistical error values among total of three, compared with HL=75 model, therefore reported to be better than HL=75 model. Ultimately, because of determining and reporting that model of HL=5 better than models of HL=10 and HL=75; no comparison was needed thus conducted between the number of HL=10 and the number of HL=75 models, and eventually, the model generated during the adjustment of the number of hidden layer at HL=5 has been concluded to be the superior result of 9 trials, executed during LSTM computations. In this regards, the best LSTM result of HL=5 generating MAE, RMSE, and R, respectively to 3.38 m^3/s, 5.44 m^3/s, and 0.9968 R, has been denoted by the utilization of bold pattern in Table 5.

Table 5. LSTM algorithm results

HL	Epoch number	MAE (m^3/s)	RMSE (m^3/s)	R
5	**300**	**3.38**	**5.44**	**0.9968**
10	300	3.29	6.01	0.9961
25	300	6.83	13.26	0.9872
50	300	8.24	19.10	0.9757
75	300	6.09	11.00	0.9990

100	300	8.73	15.13	0.9777
125	300	7.49	12.84	0.9824
150	300	9.53	19.53	0.9642
300	300	9.63	18.65	0.9687

As in the case of former computations, Figures 13 and 14, respectively denote the data dispersions of the actual data cloud simultaneously shown against the counterpart of LSTM prediction data function and the correlation outcomes of the instantaneous data observations revealed with respect to the estimations. Furthermore, the mean value of the LSTM correlation result (R) of the superior outcome captured among diverse values of the number of HL, corresponding to the statistical accuracy outcomes generated at 5 HL, has been withal exposed in Figure 14, corresponding to the value of 0.9968 R. These two figures indicate the high rapport of the actual data given according to its counterpart guesses. In this regards, the prediction function displayed in dashed black colored curve presented in Figure 13, is sufficiently coincident on the real data cloud signified in continuous yellow colored curve, therefore, this situation points a high quality of data estimations. This case is also seen from Figure 14, which demonstrates the correlation scatter scene of both data series. Namely, all instantaneous data samples that have two roots, one corresponding to the actual data cluster, the other pointing for the WFR data estimations, are observed to be scattered whole close to the regression line. Also, this regression line drawn in the form of dotted blue colored line almost behaves again as a symmetry line of the measured and anticipated data axes.

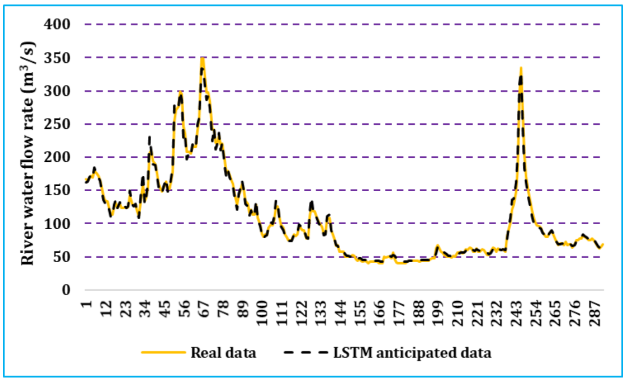

Figure 13. Real flow data (WFR) given according to LSTM anticipations

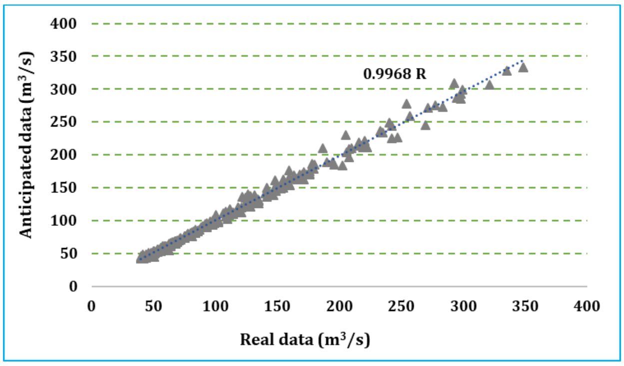

Figure 14. The LSTM correlation scatter for water flow time-series

3.2.5. The comparison of the applied methods

These implemented methods are not based on the experimental observations and historical data files including certain physical parameters that were created using observations obtained by past experimental methods are used in simulations. Namely, the data input considered in the simulations for future estimations are formed as past historical time-series data of the considered physical parameter. For this reason, these methods provide estimates for the very near future and cannot predict very accurately the distant future from the current time. In order to estimate more distant times with these methods, a second estimate will need to be derived from the near future estimates that are obtained by the simulations, which will increase the added error rate.

In the final stage, a comparison was subjected between the prediction capability of FCM, SC, GP of ANFIS and LSTM. In this context, the model generated during the calibration of the number of HD=4 and IR=0.20 for SC algorithm has been compared with respect to the models formed as soon as the number of HD and MFs were set respectively to 5 and 2 considering FCM algorithm, the number of HD and MFs were adjusted respectively to 3 and 2 keeping in view the GP algorithm, and the number of HL was arranged to 5 regarding LSTM methodology. Eventually, it was concluded that 3.04 m^3/s MAE, 4.91 m^3/s RMSE, 0.9979 R outcomes of SC were ahead in quality of 3.13 m^3/s, 4.90 m^3/s, 0.9978 R of FCM, 3.53 m^3/s, 5.79 m^3/s, 0.9964 R of GP, and 3.38 m^3/s, 5.44 m^3/s, 0.9968 R of LSTM. Namely, it was concluded and demonstrated that the statistical accuracy outcomes of MAE and RMSE on the behalf of SC were smaller and better than these statistical error results of FCM, GP, and LSTM, as well as, similarly, has been exhibited that the correlation result of R on the behalf of SC was bigger and superior than this correlation outcome of R of FCM, GP, and LSTM. This scene has been better observed and therefore concluded from the comparison of Figure 9 of SC with Figures 7 of FCM, 11 of GP, and 13 of LSTM.

4. Discussion and Conclusion

ANFIS and LSTM models generally have a structure that has a good learning capability and they easily perceive the harmony between instantaneous measurement data samples. When the number of historical data used is sufficient and in the case when there is no discontinuity in the real observed data function, the prediction data function usually predicts the real data function with a very high success. The situations when the prediction function may fail can be summarized as the case involving of very sudden fluctuations in the data values and the case of frequent very high data magnitude changes occurring between the ensuing data values. River flow velocities also generally show smooth and soft changes throughout the year, except for flash floods. In winter, in cases of snow and rain, there is an increase in the river flow rate depending on the total amount of the water quantity available in the river bed, but this increase usually does not occur so suddenly and the flow rate increases in a gradual form. Similarly, the decrease in the water flow rate, which decreases in the summer, also occurs gradually. Considering these reasons, the mentioned prediction algorithms can generally be easily applied to different rivers, and different and various hydrological conditions.

During training of the models, the considered four-year dataset was reported to be pretty sufficient in model training, since the magnitudes of the MAE and RMSE were taken down below the values of 3.10 m^3/s and 4.95 m^3/s, respectively for those error parameters. Namely, taking a water flow rate (WFR) dataset found in the range of $28.50\ m^3/s \leq WFR \leq 592.00\ m^3/s$ into account, these error rates are reported to be quite low. Similarly, when the correlation coefficient (R) result of the best model was considered, namely the value of R=0.9979 is reported to be rather high, approaching so close to the maximum value of R=1. On the other hand, the case of selecting or using a longer dataset can bring significant improvements in the training of the model, thus the accuracy of the implemented models may increase. In other words, examining the river flow regime over a longer period of time allows the computer to better learn the logic of how the river water flow rate behaves and the algorithm established by the computer, in this direction, can be better formed. As a result, learning the future water flow rates that have not yet occurred, using a larger number of past historical data can lead to more precise predictions of future water flow rates. However, the past historical dataset, which includes four years of water flow rate data of the current study, is still capable of providing a high success in estimating the future water flow rates, and thus offers a qualified forecast model.

In this study, as a novelty, the water flow rate was predicted, which is seldom seen in the literature, with four different machine learning algorithms, including methods of ANFIS-FCM, ANFIS-SC, ANFIS-GP, and LSTM. Accordingly, a total of 102 different models were created in large number by changing the tuning parameters of the relevant algorithms. In the generated models, the parameters that have a direct impact on the forecast performance, such as the number of the historical data used in algorithm training (HD), membership functions (MFs), radius of influence (IR) and the number of the hidden layers (HL), were altered and the best prediction model was obtained by the evaluation of the estimations due to the statistical error parameters, and thus, the results were reported. In short, applying these four algorithms together to water flow rate (WFR) estimation, considering these mentioned tuning parameters, will fill this gap in the literature.

In this present study, the volumetric rate of water flow (WFR) predictions of Maritsa River was performed, using the past time-series historical data of the water flow measurements of this river; considering fuzzy c-means (FCM), subtractive clustering (SC), grid partitioning (GP) algorithms of adaptive neuro fuzzy inference system (ANFIS), plus the neural network of long short-term memory (LSTM). In this regards, in the context of FCM computations, a total of 40 different models at different number of historical data (HD) utilized for training values as well as different MFs values were tried, considering SC simulations, a cumulative of 48 distinct models at different number of HD and influence radius (IR) were tested, taking GP modelling into account, an aggregate of 5 discrete models at different HD and MFs were studied, finally, a whole of 9 distinct models at different HL numbers were operated for LSTM algorithm computations. Based on these 102 different model computations, although it was obtained and reported that the results indicate more or less similar outcomes in terms of the statistical accuracy errors; on the other hand, SC results have indicated slightly better and more quality prediction data distributions compared to the actual observed volumetric water flow rate (WFR) data of these measurements captured from Svilengrad water measurement station, that is seated on the Maritsa river. Besides, the statistical accuracy parameters that were taken into account to obtain model comparisons were mean absolute error (MAE), root mean square error (RMSE), additionally, the correlation coefficient (R).

The mentioned best model of SC has procreated statistical accuracy results of 3.04 m^3/s on MAE, 4.91 m^3/s at RMSE, and 0.9979 for R, that is reported to be the leader among the total tested 102 models generated at different selected number of historical data considered for training values (HD), different membership function (MFs) values, discrete influence radius (IR) values, and distinct hidden layer (HL) values. Furthermore, the correlation result of this value reaching to 0.9979 R, and therefore quite converging to the unity value 1, implies that the functional distribution of the prediction function of this model almost perfectly resembles the functional distribution of the real measured past time-series of river water flow rate (WFR) data. In this regards, it was concluded that SC algorithm of ANFIS can be safely and quickly implemented on obtaining accurate estimations on future water flow rate (WFR) values of the river water flow.

In predictions of volumetric flow rate of rivers, different machine learning and artificial intelligence techniques are aimed to be studied, as future studies. In this way, the performance and the estimation qualities of the applied methods can be compared in a wider range, and the one having the highest prediction capability can be suggested to the literature as a novel technique.

References

[1] Lurwan, S. M., Mariun, N., Hizam, H., Radzi, M. A. M., Zakaria, A. 2014. Predicting Power Output of Photovoltaic Systems with Solar Radiation Model. 2014 IEEE International Conference on Power and Energy (PECon), 1-3 December, Kuching, Malaysia, 304-308.

[2] Oruç, E. N. 2021. Yapay sinir ağları kullanılarak kısa süreli güneş enerjisi tahmini. İstanbul Teknik Üniversitesi, Lisansüstü Eğitim Enstitüsü, Yüksek Lisans Tezi, İstanbul.

[3] Theocharides, S., Makrides, G., Georghiou, G. E., Kyprianou, A. 2018. Machine Learning Algorithms for Photovoltaic System Power Output Prediction. IEEE International Energy Conference (ENERGYCON), 3-7 June, Limassol, Cyprus, 1-6.

[4] Scott, C., Ahsan, M., Albarbar, A. 2023. Machine Learning for Forecasting a Photovoltaic (PV) Generation System. Energy, 278, 127807.

[5] Gaboitaolelwe, J., Zungeru, A. M., Yahya, A., Lebekwe, C. K., Vinod, D. N., Salau, A. O. 2023. Machine Learning Based Solar Photovoltaic Power Forecasting: A Review and Comparison. IEEE Access, 11, 40820-40845.

[6] Liu, H., Tian, H., Li, Y. 2015. An EMD-recursive ARIMA Method to Predict Wind Speed for Railway Strong Wind Warning System. Journal of Wind Engineering and Industrial Aerodynamics, 141, 27-38.

[7] Kavasseri, R. G., Seetharaman, K. 2009. Day-ahead Wind Speed Forecasting Using f-ARIMA Models. Renewable Energy, 34(5), 1388-1393.

[8] Zhang, J., Yan, J., Infield, D., Liu, Y., Lien, F. 2019. Short-term Forecasting and Uncertainty Analysis of Wind Turbine Power Based on Long Short-term Memory Network and Gaussian Mixture Model. Applied Energy, 241, 229-244.

[9] Zaytar, M. A., Amrani, C. E. 2016. Sequence to Sequence Weather Forecasting with Long Short-term Memory Recurrent Neural Networks. International Journal of Computer Applications, 143(11), 7-11.

[10] Prabha, P. P., Vanitha, V., Resmi, R. 2019. Wind Speed Forecasting Using Long Short Term Memory Networks. 2nd International Conference on Intelligent Computing, Instrumentation and Control Technologies (ICICICT), 5-6 July, Kannur, Kerala, India, 1310-1314.

[11] Benmouiza, K., Cheknane, A. 2019. Clustered ANFIS Network Using Fuzzy C-means, Subtractive Clustering, and Grid Partitioning for Hourly Solar Radiation Forecasting. Theoretical and Applied Climatology, 137, 31-43.

[12] Tumse, S., Bilgili, M., Sekertekin, A., Unal, S., Sahin, B. 2023. Comparison and Evaluation of Machine Learning Approaches for Estimating Heat Index Map in Türkiye. Neural Computing and Applications, 35(21), 15721-15742.

[13] Tumse, S., Bilgili, M., Sahin, B. 2022. Estimation of Aerodynamic Coefficients of a Non-slender Delta Wing Under Ground Effect Using Artificial Intelligence Techniques. Neural Computing and Applications, 34(13), 10823-10844.

[14] Tumse, S., Ilhan, A., Bilgili, M. 2022. Estimation of Wind Turbine Output Power Using Soft Computing Models. Energy Sources, Part A: Recovery, Utilization, and Environmental Effects, 44(2), 3757-3786.

[15] Hewett, R., Leuchner, J., Carvalho, M. 2001. From Climate History to Prediction of Regional Water Flows with Machine Learning. Institute of Electrical and Electronics Engineers (IEEE), 292-297.

[16] Tayfur, G., Singh, V. P., Moramarco, T., Barbetta, S. 2018. Flood Hydrograph Prediction Using Machine Learning Methods. Water (Multidisciplinary Digital Publishing Institute), 986, 1-13.

[17] Flake, J. T. 2007. Application of the relevance vector machine to canal flow prediction in the Sevier River basin. Utah State University, Electrical and Computer Engineering, Master of Science Thesis, Utah.

[18] Farhadi, H., Zahiri, A., Hashemi, M. R., Esmaili, K. 2019. Incorporating a Machine Learning Technique to Improve Open-channel Flow Computations. Neural Computing and Applications, 31, 909-921.

[19] Sahraei, A., Chamorro, A., Kraft, P., Breuer, L. 2021. Application of Machine Learning Models to Predict Maximum Event Water Fractions in Streamflow. Front Water, 3, 1-21.

[20] Il Kim, H., Kim, B. H. 2020. Flood Hazard Rating Prediction for Urban Areas Using Random Forest and LSTM. Water Resources and Hydrologic Engineering, 24, 3884-3896.

[21] Yaseen, Z. M., Jaafar, O., Deo, R. C., Kisi, O., Adamowski, J., Quilty, J., El-Shafie, A. 2016. Stream-flow Forecasting Using Extreme Learning Machines: A Case Study in a Semi-arid Region in Iraq. Journal of Hydrology, 542, 603-614.

[22] Xiao, L., Zhong, M., Zha, D. 2022. Runoff Forecasting Using Machine-learning Methods: Case Study in the Middle Reaches of Xijiang River. Front Big Data, 4, 1-11.

[23] Mathworks, 2020a. Fuzzy C-means Clustering. https://www.mathworks.com/help/deeplearning/ug/long-short-term-memory-networks.html (Access Date: 05.05.2024).

[24] Mathworks, 2020b. Long Short-term Memory Networks. https://www.mathworks.com/help/deeplearning/ug/long-short-term-memory-networks.html (Access Date: 07.06.2024).

[25] Jang, J. R. 1993. ANFIS: Adaptive-network-based Fuzzy Inference System. IEEE Transactions on Systems, Man, and Cybernetics, 23(3), 665-685.

[26] Chandy, K. M., Taylor, S. 1992. An Introduction to Parallel Programming. MA: Jones and Bartlett, Boston.

[27] Liu, R., Liu, L. 2019. Predicting Housing Price in China Based on Long Short-term Memory Incorporating Modified Genetic Algorithm. Soft Computing, 23, 11829-11838.

[28] Ilhan, A. 2023. Forecasting of Volumetric Flow Rate of Ergene River Using Machine Learning. Engineering Applications of Artificial Intelligence, 105983.

[29] Devlet Su Isleri (DSI). 2024. 11th regional directorate of DSI of Türkiye. http://edirnenehir.dsi.gov.tr/ (Access Date: 25.09.2024).