



Drought modelling by standard precipitation index (SPI) in a semi-arid climate using deep learning method: long short-term memory

Alireza Docheshmeh Gorgij¹ · Meysam Alizamir² · Ozgur Kisi³ · Ahmed Elshafie^{4,5}

Received: 11 November 2020 / Accepted: 31 August 2021 / Published online: 30 September 2021
© The Author(s), under exclusive licence to Springer-Verlag London Ltd., part of Springer Nature 2021

Abstract

Drought modelling is an important issue because it is required for curbing or mitigating its effects, alerting the people to the its consequences, and water resources planning. This study investigates the capability of a deep learning method, long short-term memory (LSTM), in forecasting drought calculated from monthly rainfall data obtained from four stations of Iran. The outcomes of LSTM compared with extra-trees (ET), vector autoregressive approach (VAR) and multivariate adaptive regression spline (MARS) methods in forecasting four drought indices, SPI-3, SPI-6, SPI-9 and SPI-12, taking into account numerical criteria, root-mean-square errors (RMSE), Nash–Sutcliffe efficiency and correlation coefficient together with the visual methods, time variation graphs, scatter plots and Taylor diagrams. The overall results showed that the LSTM method performed superior to the ET, VAR and MARS in forecasting drought based on SPI-3, SPI-6, SPI-9 and SPI-12. The RMSE of ET, VAR and MARS was improved by about 17.1%, 12.8% and 9.6% for SPI-3, by 10.5%, 6.2% and 5% for SPI-6, by 7.3%, 4.1% and 6.2% for SPI-9 and by 22.2%, 27% and 10.6% for SPI-12 using LSTM. The MARS method was ranked as the second best, while the ET provided the worst results in forecasting drought based on SPI.

Keywords Drought forecasting · Standard precipitation index · Deep learning · LSTM · Extra-trees · VAR · MARS

1 Introduction

Drought as a meteorological phenomenon, relative to the expected amount, could be described as the lack of rainfall intensity over a certain period of time. Mathematically, when the real rainfall rate is equal to or less than the specified percentage of the predicted rainfall over the same time span, this particular duration is called a drought occurrence [41]. For example, if the rainfall rate is equal to 75% of the average long term for the given duration, the presence of the drought might be considered, whereas others may consider it to be about 60% or 50% [7]. In terms of societal influences, some scientists have indicated that drought should be established. They also argued that it typically adversely affects the developed economy of the region before a decrease in precipitation becomes a matter for society [45]. Agricultural and hydrological droughts can occur, as well as meteorological droughts, which can contribute to disagreements as to whether a drought has actually occurred [23, 22, 45].

✉ Meysam Alizamir
alizamir@iauh.ac.ir; meysamalizamir@gmail.com

✉ Ozgur Kisi
ozgur.kisi@iliauni.edu.ge
Alireza Docheshmeh Gorgij
gorgij.a.d@eng.usb.ac.ir
Ahmed Elshafie
elshafie@um.edu.my

¹ Faculty of Industry and Mining (Khash), University of Sistan and Baluchestan, Zahedan, Iran

² Department of Civil Engineering, Hamedan Branch, Islamic Azad University, Hamedan, Iran

³ Department of Civil Engineering, Ilia State University, 0162 Tbilisi, Georgia

⁴ Civil Engineering Department, Faculty of Engineering, University of Malaya, Kuala Lumpur, Malaysia

⁵ National Water Center, United Arab Emirates University, P.O. Box 15551, Al Ain, United Arab Emirates

Definitions of what causes a drought will vary based on the possibility of specific human activities requiring rainfall within a given area. One of the expected effects of climate change is the possible increase in both the frequency and intensity of extreme weather events, such as hurricanes, floods and droughts. The warming of the planet would improve ocean–atmosphere interactions that could increase the regularity and intensity of the frequency and status of tremendous weather and environment. Drought research is needed to find out the likelihood of a shortfall in the supply of water in future. It also encourages researchers to present a study that can curb drought, minimize the impact of drought, alert the public to the effects of drought and plan for the future. A practical estimate and forecast of potential drought events is also compelling.

Existing studies that have been reviewed to assess and forecast both meteorological and hydrological droughts in Semnan, Iran, are introduced in Sanusi et al. [46]. As a result from this study, artificial intelligence and data-driven models will be used to forecast future drought events. A few suitable drought indices such as the SPI (standardized precipitation index), RAI (rainfall anomaly index), SIAP (standard index of annual precipitation), SDI (streamflow drought index), SSFI (standardized streamflow index) and SWSI (standardized water storage index) for the Semnan, Iran, were estimated [17]. The data-driven models to be used are traditional stochastic models for time-series forecasting [10] and more recent machine learning models [9].

It was established that, as opposed to a satellite-based drought index, a data-driven drought index would be used to predict drought for this analysis. Satellite-based drought indices are, in reality, adaptive over a given spatial and temporal scale to changes in vegetative land cover. Because of the time it takes for the impact of a drought to be noticeable on vegetative surfaces, satellite-based drought indices are not as effective in detecting the emergence of droughts. Of the above data-driven drought indices discussed, the PDSI [27, 33, 34] and the SPI [32, 43] have found widespread application in the field of drought forecasting. The fact that they are standardized is the principal strength of these two drought indices. The PDSI assumes parameters such as soil characteristics are uniform over a climatic region in view of the very complex empirical derivations of the PDSI. The SPI, however, is geographically independent, is relatively easy to quantify and allows both short- and long-term droughts to be quantified. It was determined that the SPI is the most acceptable drought index for drought forecasting, considering these benefits [13].

The need to acquire fast and precise prediction is significant in operation and management of water resources. This is to make sure that an acceptable time is given to the

authorities for implementing their protection procedures. In general, by simulating a catchment's response to rainfall, watershed hydrology models are optimized for these goals. Simulating the answer using mathematical models, however, is not a straightforward job as the hydrological processes are complicated and the effect on them is not well known by geomorphological conditions and climatic influences [7]. Consequently, one of the most optimistic subjects of hydrology is now the pursuit of specific, accurate and scientifically realistic models [8, 26]. Since the earliest statistical simulation, the logical approach for peak discharge, many hydrologic models have been accounted for, was developed by Arvind et al. [6]. These hydrological models may be grouped as knowledge-driven or data-driven by their internal definition of the hydrological processes and, according to the spatial representation of managed watersheds, as lumped or dispersed [28, 35]. Physically based or processed-based, mechanistic models and conceptual models are created through knowledge-driven models.

The internal parameters of the model are only related to the model configuration, while the physical parameters of the hydrological routes (i.e. run-off generation) are not taken into account by model structure, making data-driven models appear simple and easy to create. Likewise, because of the adversity of data processing or the lack of hydrometeorological networks, the scarcity of hydrological data creates rational difficulties for the use of knowledge-driven models [44]. In this case, data-driven modelling appears to be the only other option where model inputs can only be easily obtained from previous documents. Conversely, in some real-world situations, the greatest difficulty is to make precise and timely estimates in certain regions. A rational data-driven model is enough to establish a direct plot between inputs and outputs. At present, the data-driven model is becoming increasingly popular in the worldwide hydrological society. Scientists also remain very uncertain about the intent of the acceptable modelling style and adequate model structure. No modelling method triumphs over others, as each modelling method demonstrates advantages and disadvantages. It was not possible for a global data-driven model to relate all hydrological conditions accurately. Hybrid models are then proposed in which, first of all, several classifications subdivide hydrological conditions and then isolated models are created for each of them [11, 36].

Using various and assemblages of historical rainfall events, the efficient drought index (EDI) and the normal precipitation index (SPI), they estimated quantitative values for drought indices [24, 25]. Direct multistep neural network (DMSNN) and linear stochastic models of recursive multistep neural network (RMSNN) were evaluated for drought forecasting by Mishra and Desai [32] and

Quilty et al. [37]. In the hydrological area, data-driven models have proven to be effective, but there is still a position available to develop better methods of forecasting [38]. By simply applying such models to an initial time series, the specifics of adjustment are overlooked, decreasing the accuracy of estimation [42]. The original time series decomposed with discrete wavelet transformation (DWT) enhances a forecasting model by providing useful information on various degrees of resolution [27]; however, there has not been much research on using a wavelet for drought prediction. Shabri [40] proposed a hybrid wavelet and adaptive neuro-fuzzy inference approach for SPI-based drought prediction in an analysis (WANFIS). The study showed that the SPI values are very sensitive to parametric distribution feature selection, and most drought indices DIs have been developed for particular areas, and some DIs are better suited for specific uses than others [13]. Therefore, owing to varying hydro-climatic conditions and many other considerations, any of the current DIs might not be specifically applicable to other areas. In fact, drought index tracking in geographic areas may be dependent on the accessibility of hydrometeorological information data and the aptitude of the drought index to accurately track the temporal and spatial pattern variations [1]. Various climate and water quality parameters are applied to describe the severity of a drought event. Although, in all cases, none of the key indices are inherently superior to the others, for some reasons, some indices are better suited than others.

As can be seen from the above, there are various forms of machine learning (ML) models that can be used for drought forecasting. However, the existing ML models that can memorize the pattern of inputs–outputs within the model structure are the most suitable. Recurrent neural networks are the most appropriate form of neural network (RNN) [39]. In reality, there are different RNNs from traditional neural feed-forward networks. This disparity in the integration of complexity comes with the promise of new habits that traditional methods cannot achieve. RNN also has an internal state which can reflect contextual awareness [18]. Furthermore, it retains information about previous inputs for a period of time that is not set a priori, instead depending on its weights and the input data. RNNs whose inputs are not fixed but represent an input sequence may be used to transform an input sequence into an output sequence if contextual information is flexibly taken into account first, in order to be able to retain information for an unspecified time in the method. Second, the model architecture should be noise immune (i.e. input variations that are unpredictable or unrelated to estimating a correct output). Third, the parameters of model architecture can be learned (in a short period of time).

RNNs must use context when making predictions, but the relevant context must be learned to some extent. RNNs include loops which feed network activations from a previous time point as inputs to the network to affect forecasts at the current time point. Internal network states, which are capable of maintaining long-term transient contextual information in principle, maintain these activations. RNNs may use this approach to control a contextual window that shifts over the course of the input sequence.

LSTM success is one of the first implementations to overcome technical problems and fulfil the promise of persistent neural networks in their statement that they can effectively memorize the historical pattern [16]. As a result, regular RNNs fail to remember lag-times that are between 5 and 10 s apart from the desired output. Because of the unseen fault problem, it is unclear whether standard RNNs can still outperform time-window-based feed-forward networks in terms of functionality [16]. Graves et al. [18] found that LSTM could adjust to a minimum lag-time of more than 10 time-steps. The two technical challenges LSTMs tackle are the absence of gradients and the bursting of gradients, both linked to how the network is trained. Unfortunately, the collection of qualitative information that traditional RNNs can access is, in fact, rather limited. Since it runs through the repeated interconnections of the network, the effect of the selective input patterns prearranged on the following phase within the hidden-layers and thus on the desired output from the model, either decays or exponentially enlarges [18]. In fact, long short-term memory (LSTM) is an RNN architecture mainly built to address the problem of the vanishing gradient, which is referred to as the vanishing gradient problem in the literature. The key to the LSTM approach to technical challenges was the simple internal configuration of the units used in the model, which is controlled by its ability to cope with disappearing and bursting gradients, the most common difficulty in the design and training of RNNs.

With the assistance of data-driven techniques, the study aims primarily to develop robust artificial intelligence forecasting models that can easily and reliably predict drought in the Semnan, Iran. This includes the development and testing of a modern drought forecasting approach based on time-series methods such as multivariate adaptive regression spline (MARS) and vector autoregression (VAR) beside the traditional ML models such as extra-trees (ET) method. The importance of this analysis and the overall findings will be discussed in this study. As most of the current drought indices (DIs) have been developed for the use in a particular area, the suitability of these DIs for Semnan province has not been examined, although similar studies have been undertaken in other parts of the world; however, these studies examined different types of models. Specifically, the objectives of the study are to estimate the

standardized precipitation index (SPI) from the raw monthly precipitation data for characterization of meteorological droughts. In addition, develop and evaluate the performance of deep learning method, LSTM in forecasting four SPI indices, SPI-3, SPI-6, SPI-9 and SPI-12. Furthermore, compare the outcomes of LSTM with the three machine learning methods, ET, VAR and MARS.

2 Materials and methods

2.1 Study area

Due to specific hydrologic and climatic parameters in semi-arid regions, it is necessary to have a wise knowledge about them. Semnan is one of the Iran’s provinces, located in a semi-arid area at 34° 17' to 37° 00' and 51° 58' to 57° 58' north latitude and east longitude, respectively. Its approximate area is about 97,491 km² and average high about

1630 m. Figure 1 depicts a general view of the study area and the location of stations, which their statistics have been evaluated in the present study.

Considering the differences between the annual precipitation amount, 120 mm, and the annual average evaporation, 220.9, in the study area, the necessity of precise assessment is obvious. The brief statistics of the time series of four SPI indices are provided in Table 1.

2.2 Standard precipitation index (SPI)

As it mentioned before, the hydrologic parameters analysis is a complex procedure in arid and semi-arid areas, due to vast temporal and spatial diversity. Among the various indices for drought study in these regions, SPI is one of the most appropriate ones [29]. It has several time scales as a meteorological drought index, e.g. 1, 3, 6, 9, 12 and 24 months. Each one of above-mentioned scales is applicable for a distinct purpose [47].

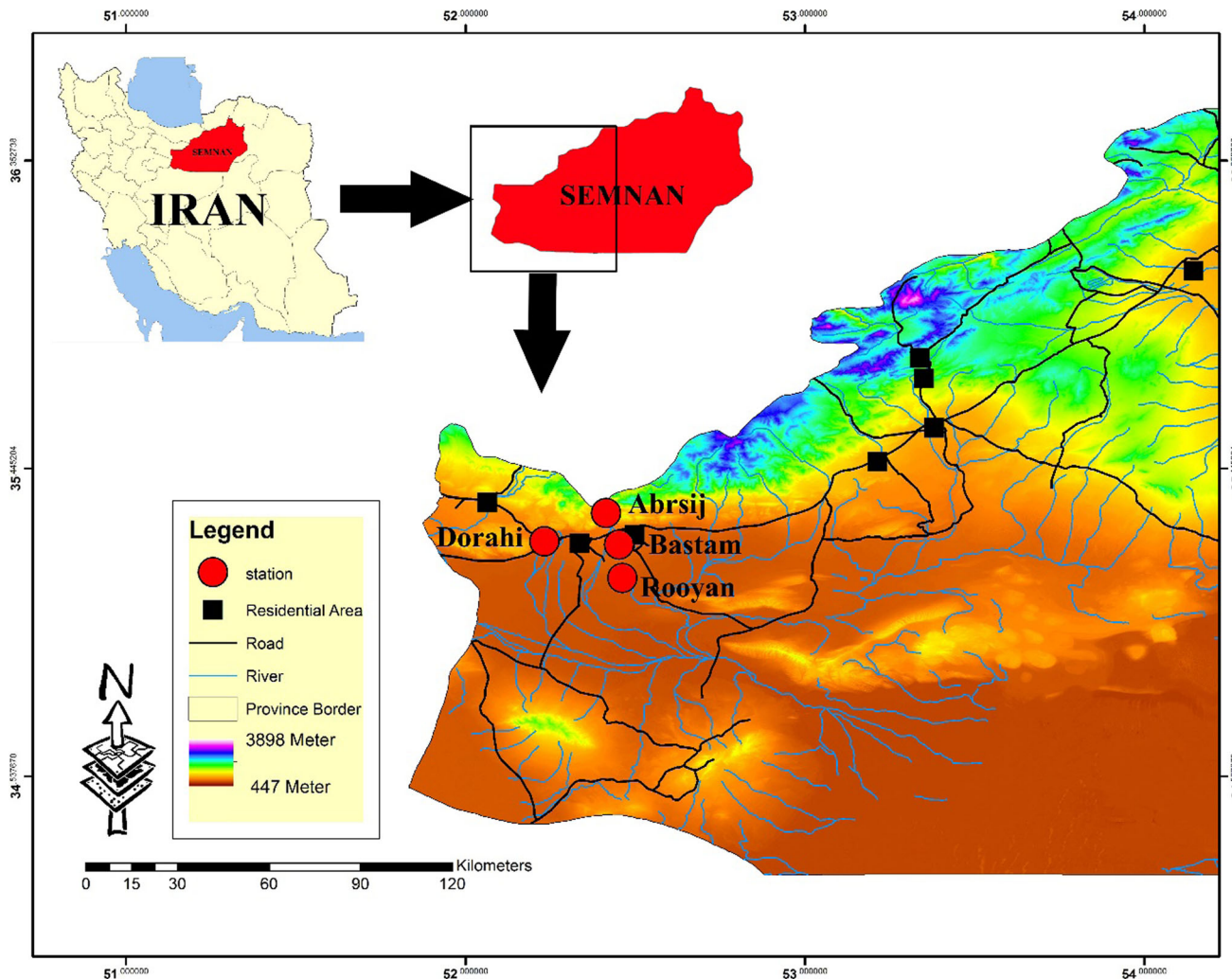


Fig. 1 General view of the study area

Table 1 Statistical properties of data for each station

Station	Data set	Average	Min.	Max.	St. Dev.	Variation coefficient
Abrsij	SPI-3	0.029	− 1.487	3.195	0.945	33
	SPI-6	0.001	− 2.939	2.771	1.001	1001
	SPI-9	− 0.001	− 2.409	2.866	1.001	− 1001
	SPI-12	0.0	− 2.572	2.407	1.001	None
Bastam	SPI-3	0.016	− 2.065	2.245	0.979	61
	SPI-6	0.004	− 3.093	2.242	1.002	251
	SPI-9	0.002	− 3.585	2.073	1.001	501
	SPI-12	0.001	− 3.615	2.054	1.001	1001
Dorahi	SPI-3	0.032	− 1.583	2.675	0.942	29
	SPI-6	0.022	− 1.884	2.381	0.957	44
	SPI-9	0.023	− 2.005	2.311	0.945	41
	SPI-12	0.025	− 2.167	1.984	0.931	37
Rooyan	SPI-3	0.095	− 0.922	2.904	0.856	9
	SPI-6	0.021	− 1.787	2.108	0.963	46
	SPI-9	0.018	− 2.410	1.992	0.968	54
	SPI-12	0.008	− 4.582	2.104	1.003	125

Since the precipitation shows a gamma probability distribution, the SPI is based on it. The gamma function can be expressed as below based on Kisi et al. [29]:

$$g(x) = \frac{1}{\beta^\alpha \cdot \Gamma(\alpha)} x^{\alpha-1} \cdot e^{-\frac{x}{\beta}} \quad \text{For } x > 0 \tag{1}$$

where $\Gamma(\alpha) = \int_0^{+\infty} x^{\alpha-1} e^{-x} dx$, $\hat{\alpha} = \frac{1}{4A} (1 + \sqrt{1 + \frac{4A}{3}})$, $A = \ln(\bar{x}) - \frac{\sum \ln(x)}{n}$ and $\hat{\beta} = \frac{\bar{x}}{\hat{\alpha}}$.

$G(x)$ is a cumulative probability and can be calculated through below equation:

$$G(x) = \int_0^x g(x) \tag{2}$$

In arid and semi-arid regions, the precipitation value is zero frequently, and the cumulative probability equation, hence, is obtained as below:

$$H(x) = q + (1 - q)G(x) \tag{3}$$

where $q = \frac{m}{n}$, while the m and n are the amount of zero values in precipitation, and observation number, correspondingly. Lloyd Hughes and Saunders [30] proposed the SPI calculation equation as below, considering the before mentioned terms.

$$SPI = - \left(t - \frac{c_0 + c_1 t + c_2 t^2}{1 + d_1 t + d_2 t^2 + d_3 t^3} \right), \quad 0 < H(x) \leq 0.5 \tag{4}$$

$$SPI = + \left(t - \frac{c_0 + c_1 t + c_2 t^2}{1 + d_1 t + d_2 t^2 + d_3 t^3} \right), \quad 0.5 < H(x) \leq 1.0 \tag{5}$$

where $c_0 = 2.515517$, $c_1 = 0.802853$, $c_2 = 0.010328$,

$d_1 = 1.432788$, $d_2 = 0.189269$ and $d_3 = 0.001308$ (McKee et al., [31]). In addition, t obtained from next mentioned equations:

$$t = \sqrt{\ln \frac{1}{(H(x))^2}}, \quad 0 < H(x) \leq 0.5 \tag{6}$$

$$t = \sqrt{\ln \frac{1}{(1 - H(x))^2}}, \quad 0.5 < H(x) \leq 1.0 \tag{7}$$

2.3 Long short-term memory (LSTM)

Among the various types of recurrent neural networks (RNNs), LSTM is recognized as an advanced form of it, which can cover the flaws of general RNN structure through the long-term dependency leaning. Hochreiter and Schmidhuber proposed LSTM for the first time in 1997, although it has improved and generalized progressively by numerous scholars [12].

Cell state (C_t) is the main concept of LSTM and passes the information through the gates in unaffected form. The above-mentioned gates, which are three, control the cell state to let information flow arbitrary.

Forget gate is known as the first gate. It selects that which cell state vector (C_{t-1}) should be forgotten and explained as below:

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f) \tag{8}$$

where f_t is the output vector, while the W_f and b_f are the parameters that can be trained for the first gate. f_t ranges from 0 to 1 and reveals the degree of forget.

Second gate is called input gate and chooses those values, which should be updated as below:

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \tag{9}$$

where the output variable is assigned to i_t . It varies from zero to one. W_i and b_i are parameters which are trained, and x_t and h_{t-1} are the current input and last hidden state, respectively. In the next step, a potential vector is calculated for cell state:

$$\tilde{C} = \tanh(W_c \cdot [h_{t-1}, x_t] + b_c) \tag{10}$$

where vector \tilde{C} ranges from zero to one, while \tanh is hyperbolic. Moreover, W_c and b_c are parameters that can be trained. Old cell state, then, can be updated into a new one called C_t :

$$C_t = f_t * C_{t-1} + i_t * \tilde{C} \tag{11}$$

The third gate is output gate and selects the output, using a sigmoid layer as below:

$$o_t = \sigma(W_o[h_{t-1}, x_t] + b_o) \tag{12}$$

where o_t is an output vector and varies from zero to one. W_o and b_o are trainable parameters. The new hidden state h_t is calculated afterwards as the final step:

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

Figure 2 shows a simple architecture of a LSTM cell. Also, Fig. 3 shows different steps of LSTM model. In this study, during the training phase, one of the most widely optimization techniques for deep learning approaches, Adam algorithm, has been applied to optimize the weight of the LSTM model, which is one of the extensions of stochastic gradient descent (SGD). Also, it can be mentioned that rectified linear units (ReLU) activation function has been used. Since the learning rate parameter has a crucial role in training process, minimizing the loss function, and convergence of the model, it was fixed to 0.001.

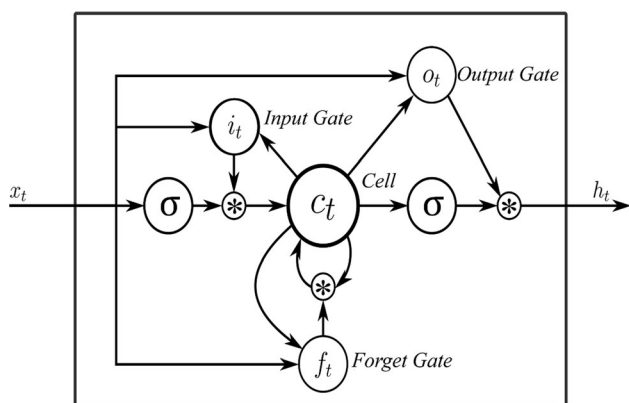


Fig. 2 Schematic architecture of LSTM cell

2.4 Extra-trees method

In 2006, an extension of random forest model suggested by Geurts et al. [19] called extra-trees. It is a tree-based learning method for decision-making and performs as a classifier and regression finder [15]. Its ability is that it can accept high-dimensional input and outputs [3]. It generates a set of unpruned decisions or regression trees upon to a classic top-down method. However, it has two major differences with other ensemble methods: at first, it separates nodes randomly via cut-point approach and second is that it applied all samples for learning to grow the tree in spite of bootstrap model.

The extra-trees, indeed, is more resourceful and extremely randomized extended than its basic model, i.e. random forest method. It has two spectacular distinctions, comparing to random forest method. At first, it does not apply the tree-bagging step in subset making. Secondly, it chooses the finest feature along with the corresponding value, in node splitting process. Above-mentioned contrasts lead to overfitting prevention and high performance in extra-trees utilization.

2.5 Multivariate adaptive regression spline (MARS)

MARS is a kind of nonparametric models, but works similar to a systematic linear regression modelling. Friedman proposed MARS in 1991. The main domain of MARS application is the improving and recognizing the complicated phenomena amongst objective and predictor variable [2]. A simple stepwise procedure of MARS model generation can be as below:

1. Data gathering.
2. Utilizing the basis mapping with knots in order to assess the candidate mapping procedure.
3. Selecting the number of terms and interaction levels to set limitations.
4. Identifying the obtained maps with minimum error in training step.
5. Eliminating the terms of model, which have the least effects.
6. Validating the model via choosing the best number of terms.

Figure 4 reveals the flowchart of MARS method in a schematic view. It should be noticed that MARS has this ability to find the pattern of hidden data and their interactions in a proficient way. MARS does that through the selecting of variables amalgamation, represented as basis function. MARS model can be explained as follows:

Fig. 3 LSTM flowchart

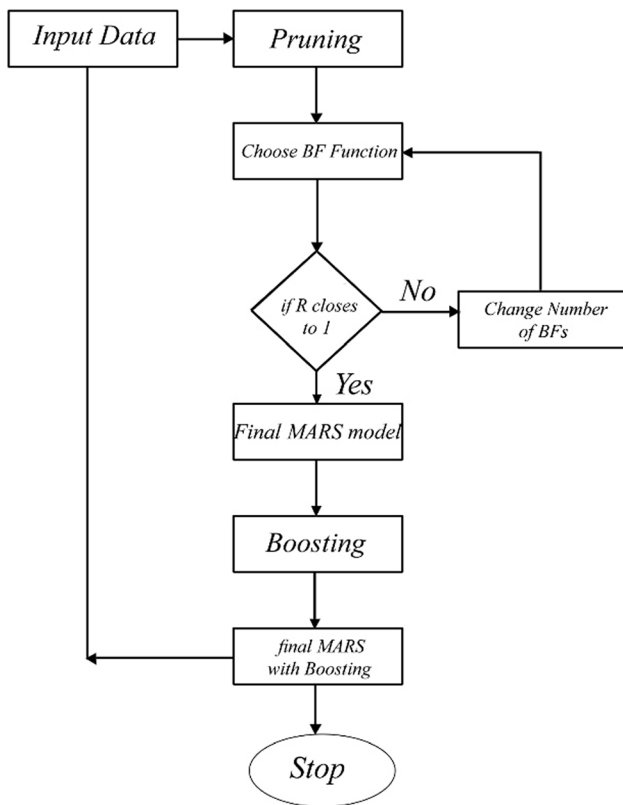
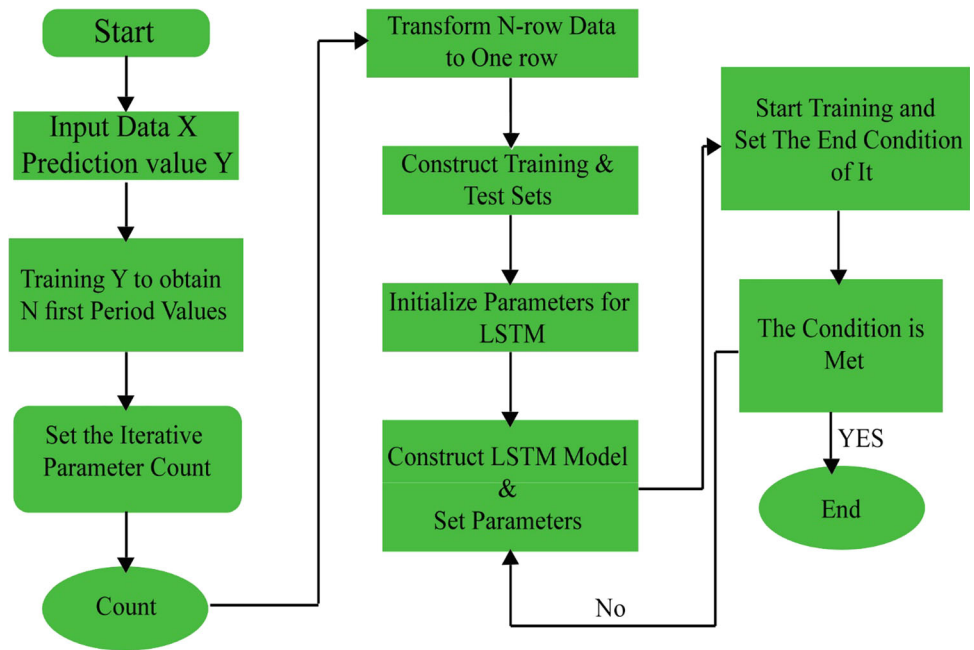


Fig. 4 MARS schematic flowchart

$$f(X) = \alpha_0 + \sum_{n=1}^N \alpha_n h_n(X) \tag{14}$$

where α is obtained from the minimized residual error calculation, and $h_n(X)$ is a function from the candidate ones

and shows the importance of variable [2]. The main excellence of MARS model is that it can evaluate the cut points (knots). Before mentioned advantages enable the MARS to be efficient in a vast area of sciences [4, 5].

2.6 Vector Autoregressive approach (VAR)

VAR is one of the most applicable approaches in multiple time-series prediction. It searches the normalized time-series data and explains their dependency and interdependency.

A VAR approach with order p , VAR (p), can be stated simply as below [14]:

$$z_{1,t} = \phi_0 + \sum_{i=1}^p \phi_i z_{t-1} + a_t \tag{15}$$

where z_t is the multiple time series, ϕ_0 is a vector which is

constant, ϕ_i is a matrix for $i > 0$, and a_t is an order of random vectors.

A simple description of VAR can be as follows:

1. Choosing the study variables
2. Generating an arbitrary model of order p
3. Defining the p value
4. Assessing the model parameters
5. Checking of diagnosis
6. Examining that are the assessed residuals that are in intended range or not.
7. Structure analysing
8. Final forecasting and evaluating the results.

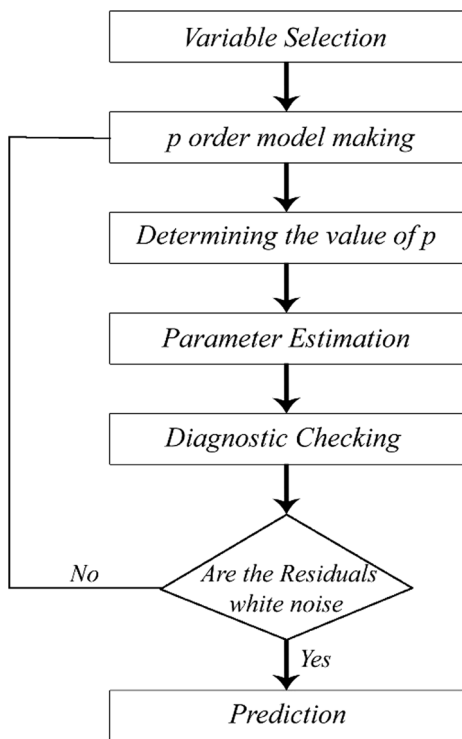


Fig. 5 VAR schematic flowchart

Figure 5 depicts the VAR steps in a simple way. Ljung–Box test is often used to check the VAR model sufficiency and is calculated as below:

$$Q_k(m) = T^2 \sum_{k=1}^m \frac{1}{T-k} t_r(\hat{R}_k^T \hat{R}_0^{-1} \hat{R}_k \hat{R}_0^{-1}) \quad (16)$$

where T is the sample size, m is the number of cross-correlation matrices of the multiple residuals, \hat{R}_0 and \hat{R}_k are the lag zero and k sample of cross-correlation. The prime sign, on the other hand, means the matrix transpose. The model adequacy depends on $Q_k(m)$, if it was higher than a considered value.

2.7 Model development

The ability of a deep learning method, LSTM, was investigated in forecasting drought using four SPI indices, SPI-3, SPI-6, SPI-9 and SPI-12, calculated from monthly rainfall data obtained from four stations, Ahrsij, Bastam, Dorahi and Rooyan. The outcomes of the LSTM method were compared with three machine learning methods, extra-trees, VAR and MARS, which were also applied to the same data sets. Data were separated into two sets, 80% for training and 20% for testing. The criteria used for assessments are root-mean-square error (RMSE), Nash–Sutcliffe efficiency (NSE) and correlation coefficient (R):

$$RMSE = \sqrt{\frac{\sum_{i=1}^N (SPI_{p,i} - SPI_{o,i})^2}{N}} \quad [0, +\infty) \quad (17)$$

$$NSE = 1 - \frac{\sum_{i=1}^N (SPI_{p,i} - SPI_{o,i})^2}{\sum_{i=1}^N (SPI_{o,i} - \overline{SPI}_o)^2} \quad (-\infty, 1] \quad (18)$$

$$R = \frac{\sum_{i=1}^N (SPI_{o,i} - \overline{SPI}_o)(SPI_{p,i} - \overline{SPI}_p)}{\sqrt{\sum_{i=1}^N (SPI_{o,i} - \overline{SPI}_o)^2 \sum_{i=1}^N (SPI_{p,i} - \overline{SPI}_p)^2}} \quad [0, 1] \quad (19)$$

where $SPI_{o,i}$ and $SPI_{p,i}$ are the observed and predicted SPI values, while \overline{SPI}_o and \overline{SPI}_p are the mean observed and predicted SPI values.

Before applying LSTM, ET, VAR and MARS methods, the optimal lagged inputs were decided according to the correlation analysis (partial auto-correlation functions). Considering 95% confidence level, the inputs provided in Table 2 were obtained for forecasting the SPI-3, SPI-6, SPI-9 and SPI-12 indices. For example, for the SPI-3 of Ahrsij station, 5 lags including SPI3t-4, SPI3t-3, SPI3t-2, SPI3t-1 and SPI3t were selected as input to the methods to forecast SPI3t + 1. It should be noted here that the Dorahi and Rooyan stations have same inputs for the four indices. A brief illustration of the model development is shown in Fig. 6.

3 Application and results

Test results of the implemented methods are summed up in Table 3 for the SPI-3 index of four stations. As observed from the table, the LSTM has performed better than the other three methods with the lowest RMSE ranging from 0.415 to 0.511 and the highest NSE ranging from 0.697 to 0.682. Compared to ET, VAR and MARS, the deep learning method has, respectively, improved the forecasting accuracy in terms of RMSE by 22.8%, 18.9% and 13.7% for Ahrsij station, by 13.9%, 8.5% and 3.5% for Bastam station, by 22.4%, 15.2% and 13.1% for Dorahi station, by 9.4%, 8.6% and 8.2% for Rooyan station. The MARS method has been ranked as the second best, while the ET has provided the worst accuracy in forecasting SPI-3 index in all stations.

Table 4 summarizes the test performances of the LSTM, ET, VAR and MARS methods in forecasting SPI-6 of four stations with respect to RMSE, NSE and R statistics. The LSTM performs superior to the other methods. Its RMSE, NSE and R value has the ranges 0.341–0.390, 0.780–0.815 and 0.889–0.909, while the corresponding values of the second best method have the ranges 0.362–0.411, 0.748–0.805 and 0.865–0.900, respectively. The RMSE accuracies of the ET, VAR and MARS methods were

Table 2 Input and output parameters used for forecasting SPI

Station	Output	Inputs
Abrsij	SPI-3	SPI3t-4, SPI3t-3, SPI3t-2, SPI3t-1, SPI3t
	SPI-6	SPI6t-2, SPI6t-1, SPI6t
	SPI-9	SPI9t-2, SPI9t-1, SPI9t
	SPI-12	SPI12t-1, SPI12t
Bastam	SPI-3	SPI3t-5, SPI3t-4, SPI3t-3, SPI3t-2, SPI3t-1, SPI3t
	SPI-6	SPI6t-3, SPI6t-2, SPI6t-1, SPI6t
	SPI-9	SPI9t-2, SPI9t-1, SPI9t
	SPI-12	SPI12t
Dorahi	SPI-3	SPI3t-6, SPI3t-5, SPI3t-4, SPI3t-3, SPI3t-2, SPI3t-1, SPI3t
	SPI-6	SPI6t-2, SPI6t-1, SPI6t
	SPI-9	SPI9t-2, SPI9t-1, SPI9t
	SPI-12	SPI12t
Rooyan	SPI-3	SPI3t-6, SPI3t-5, SPI3t-4, SPI3t-3, SPI3t-2, SPI3t-1, SPI3t
	SPI-6	SPI6t-2, SPI6t-1, SPI6t
	SPI-9	SPI9t-2, SPI9t-1, SPI9t
	SPI-12	SPI12t

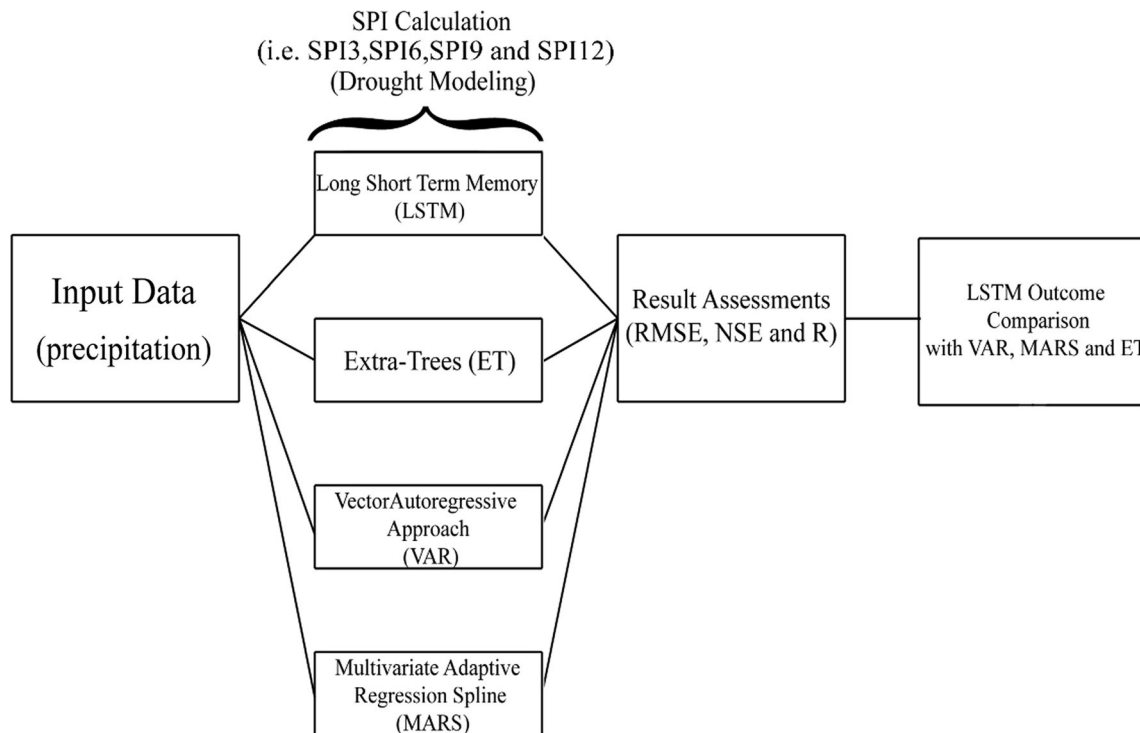


Fig. 6 Flowchart of the model development

improved by 5.1%, 5.1% and 4.9% for Abrsij station, by 7%, 1.3% and 2.8% for Bastam station, by 13.7%, 8.6% and 5.8% for Dorahi station, by 16.1%, 9.9% and 6.5% for Rooyan station using the LSTM method. In this station also, ET had the worst forecasts whereas it caught the same accuracy with VAR in Abrsij station. For the Bastam station, the VAR method performed slightly better than the MARS in forecasting SPI-6.

Test results of the LSTM, ET, VAR and MARS methods in forecasting SPI-9 of four stations can be seen from Table 5. Here also the LSTM having the ranges of RMSE, NSE and R as 0.322–0.405, 0.777–0.848 and 0.893–914, respectively, indicated a superior accuracy compared to other three methods. The LSTM has improved the forecasting performance of the ET, VAR and MARS in respect of RMSE by 4.1%, 4.1% and 3.9% for Abrsij station, by

Table 3 Testing results in forecasting SPI-3

Station	Model	Testing phase		
		RMSE	NSE	R
Abrsij	LSTM	0.511	0.690	0.856
	Extra-tree	0.662	0.479	0.749
	VAR	0.630	0.528	0.792
	MARS	0.592	0.584	0.803
Bastam	LSTM	0.495	0.697	0.836
	Extra-tree	0.575	0.592	0.769
	VAR	0.541	0.639	0.807
	MARS	0.513	0.674	0.821
Dorahi	LSTM	0.451	0.682	0.846
	Extra-tree	0.581	0.473	0.770
	VAR	0.532	0.558	0.791
	MARS	0.519	0.580	0.805
Rooyan	LSTM	0.415	0.696	0.839
	Extra tree	0.458	0.631	0.796
	VAR	0.454	0.636	0.825
	MARS	0.452	0.640	0.808

Table 5 Testing results in forecasting SPI-9

Station	Model	Testing phase		
		RMSE	NSE	R
Abrsij	LSTM	0.371	0.848	0.893
	Extra-tree	0.387	0.762	0.889
	VAR	0.387	0.762	0.887
	MARS	0.386	0.762	0.884
Bastam	LSTM	0.405	0.834	0.914
	Extra-tree	0.410	0.829	0.911
	VAR	0.411	0.829	0.911
	MARS	0.410	0.829	0.911
Dorahi	LSTM	0.331	0.777	0.891
	Extra-tree	0.376	0.712	0.860
	VAR	0.358	0.739	0.872
	MARS	0.394	0.684	0.837
Rooyan	LSTM	0.322	0.819	0.905
	Extra-tree	0.366	0.766	0.875
	VAR	0.333	0.806	0.898
	MARS	0.334	0.806	0.898

Table 4 Testing results in forecasting SPI-6

Station	Model	Testing phase		
		RMSE	NSE	R
Abrsij	LSTM	0.390	0.806	0.909
	Extra-tree	0.411	0.784	0.898
	VAR	0.411	0.784	0.900
	MARS	0.410	0.785	0.900
Bastam	LSTM	0.386	0.815	0.903
	Extra-tree	0.415	0.787	0.888
	VAR	0.391	0.811	0.900
	MARS	0.397	0.805	0.898
Dorahi	LSTM	0.341	0.790	0.906
	Extra-tree	0.395	0.719	0.884
	VAR	0.373	0.750	0.903
	MARS	0.362	0.763	0.896
Rooyan	LSTM	0.375	0.780	0.889
	Extra-tree	0.447	0.687	0.841
	VAR	0.416	0.728	0.860
	MARS	0.401	0.748	0.865

1.2%, 1.5% and 1.2% for Bastam station, by 12%, 7.5% and 16% for Dorahi station, by 12%, 3.3% and 3.6% for Rooyan station, respectively. In two stations (Abrsij and Bastam), the ET, VAR and MARS have almost the same accuracy. In Rooyan station, the VAR and MARS have

similar test results and they are superior to the ET while VAR forecasts SPI-9 better than the MARS and ET in Dorahi station.

Table 6 sums up the test accuracies of the implemented methods in forecasting SPI-12 of four stations. As clearly seen from Table 6, the LSTM has provided the best accuracy with the lowest RMSE (0.233–0.309) and the

Table 6 Testing results in forecasting SPI-12

Station	Model	Testing phase		
		RMSE	NSE	R
Abrsij	LSTM	0.246	0.865	0.932
	Extra-tree	0.269	0.838	0.921
	VAR	0.355	0.718	0.859
	MARS	0.274	0.832	0.926
Bastam	LSTM	0.309	0.904	0.951
	Extra-tree	0.393	0.846	0.923
	VAR	0.336	0.887	0.950
	MARS	0.325	0.894	0.946
Dorahi	LSTM	0.233	0.852	0.925
	Extra-tree	0.330	0.704	0.858
	VAR	0.371	0.626	0.831
	MARS	0.265	0.808	0.922
Rooyan	LSTM	0.238	0.887	0.942
	Extra-tree	0.338	0.773	0.885
	VAR	0.350	0.756	0.939
	MARS	0.281	0.843	0.921

highest NSE (0.852–0.904) and R (0.925–0.951). The RMSE of the ET, VAR and MARS was improved by 8.6%, 30.7% and 10.2% for Abrsij station, by 21.4%, 8% and 4.9% for Bastam station, by 29.4%, 37.2% and 12.1% for Dorahi station, by 29.6%, 32% and 15.3% for Rooyan

station applying LSTM, respectively. The MARS method has the second rank in forecasting SPI-12.

The accuracies of the applied methods in forecasting SPI-12 are graphically compared in Figs. 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18 and 19. Time variation graphs of the

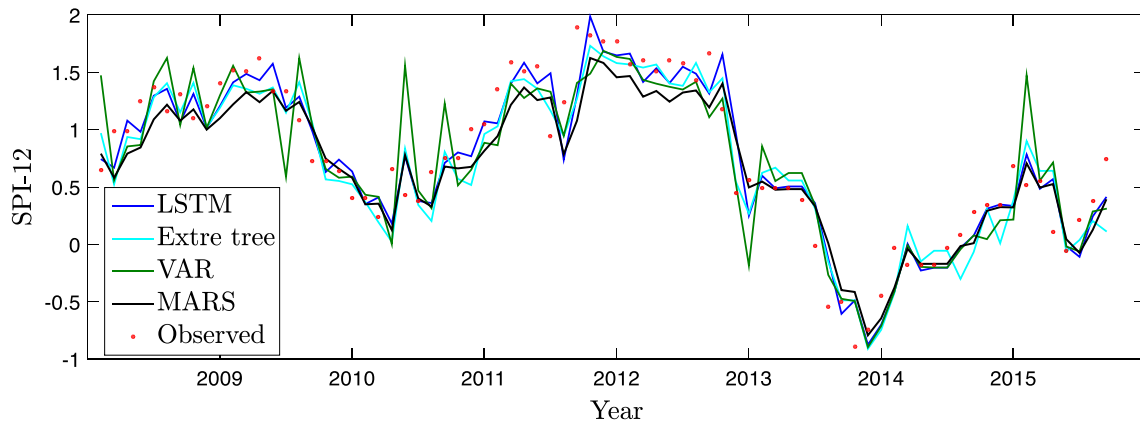
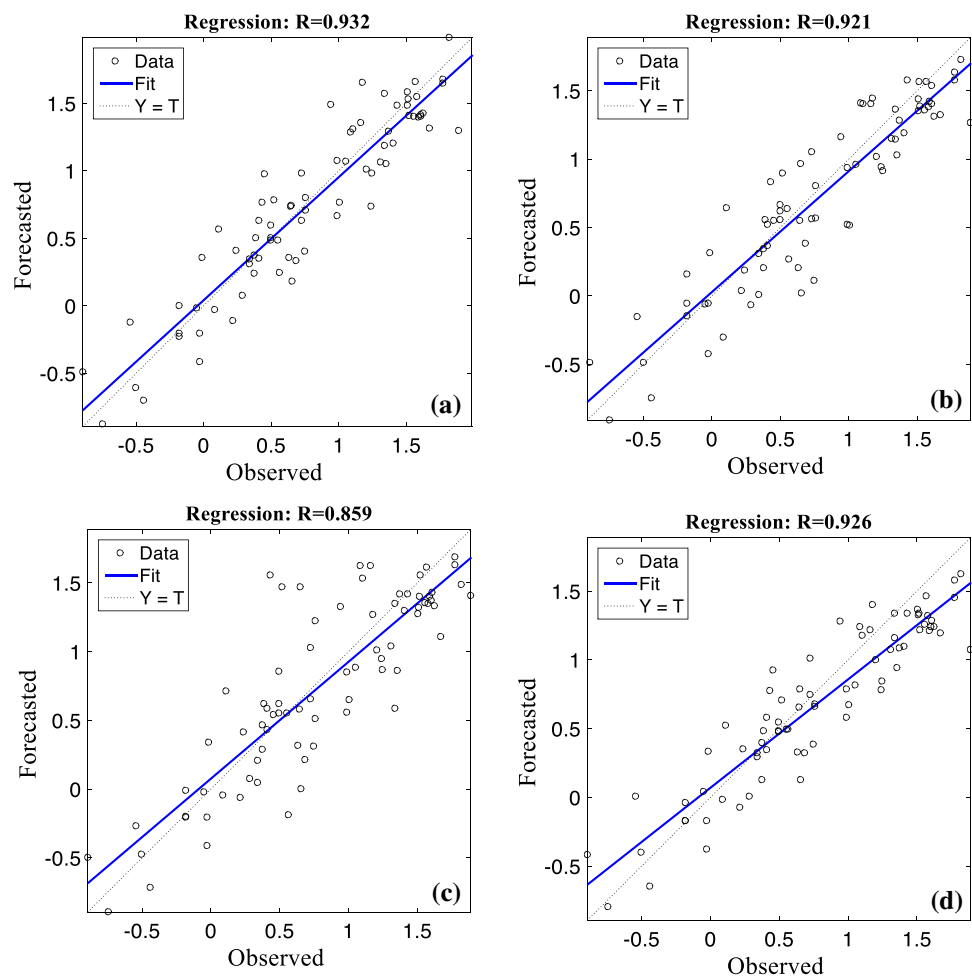


Fig. 7 Observed and forecasted time variation graphs of different models in forecasting SPI-12 at Abrsij station

Fig. 8 Scatter plots of forecasted and observed SPI-12 using: a LSTM, b extra-tree, c VAR and d MARS for Abrsij station



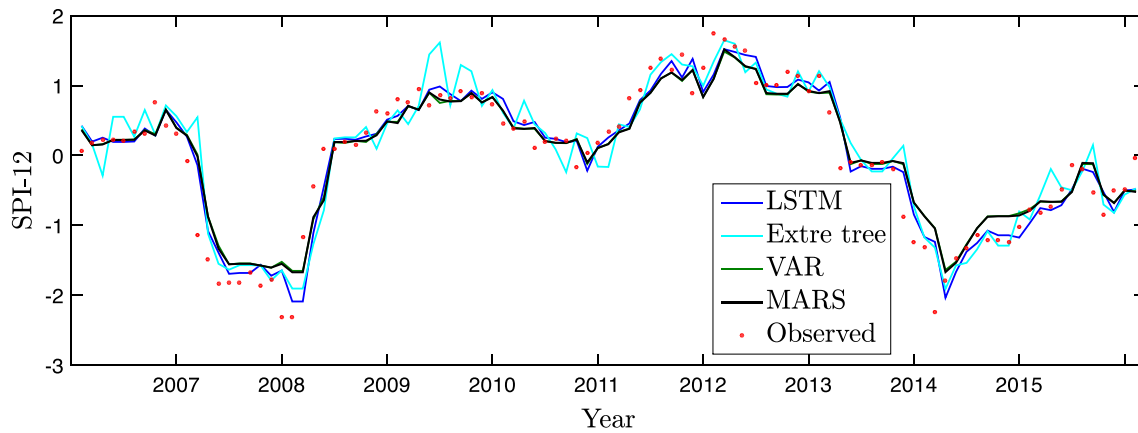
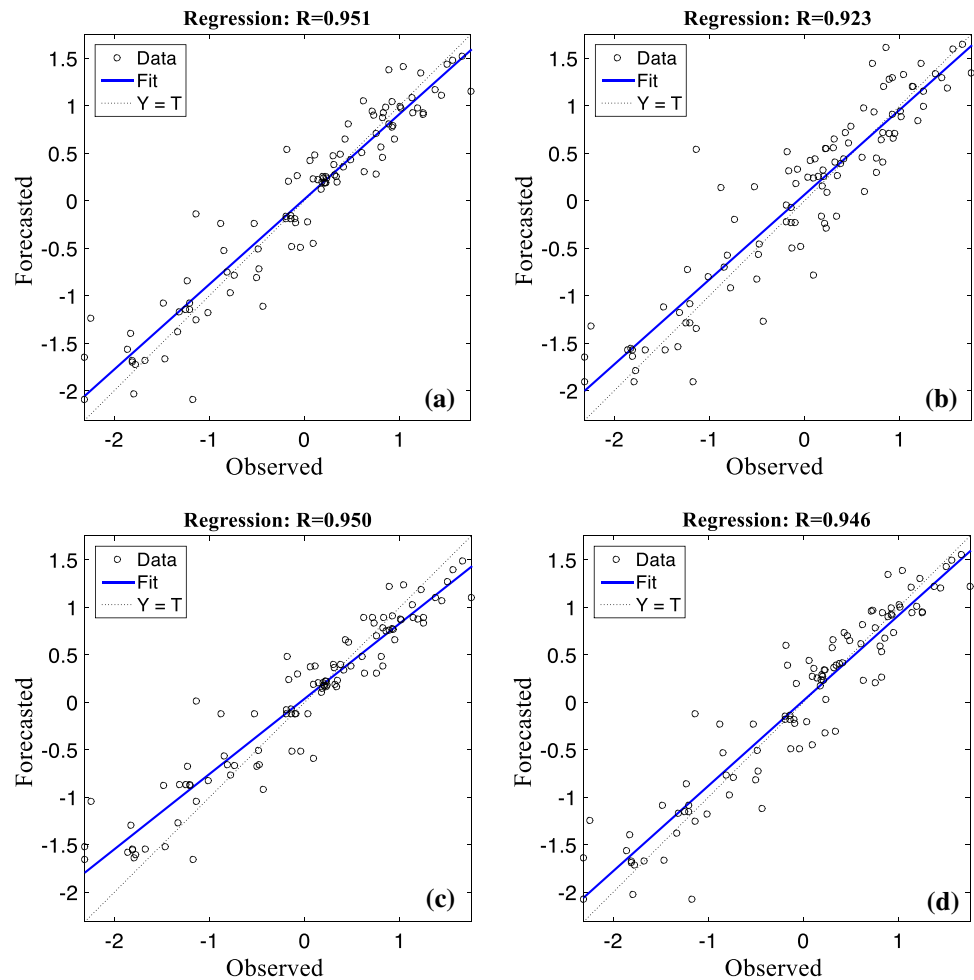


Fig. 9 Observed and forecasted time variation graphs of different models in forecasting SPI-12 at Bastam station

Fig. 10 Scatter plots of forecasted and observed SPI-12 using **a** LSTM, **b** extra-tree, **c** VAR and **d** MARS for Bastam station



observed and forecasted SPI-12 values are illustrated in Figs. 7, 9, 11 and 13, while Figs. 8, 10, 12 and 14 show the scatter diagrams. It is clear from the time variation graphs that the LSTM forecasts are closer to the corresponding

observed SPI-12 values compared to other three methods in all stations. After the LSTM, the MARS takes the second place, while the ET and VAR seem to be insufficient to catch SPI-12 values. Scatterplots tell us that the LSTM has

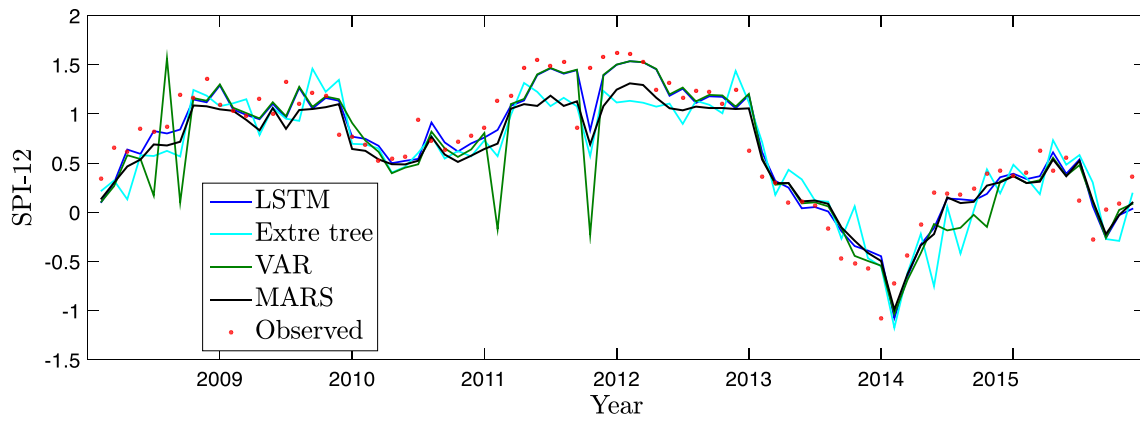
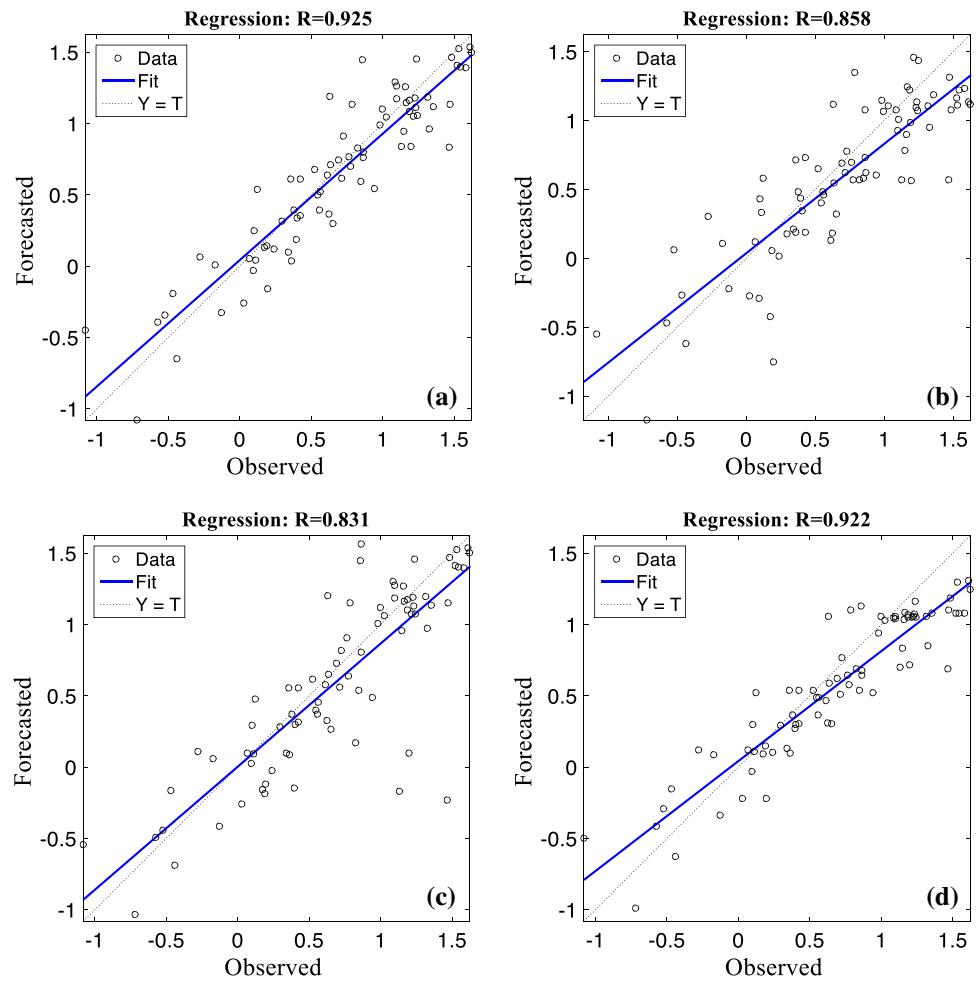


Fig. 11 Observed and forecasted time variation graphs of different models in forecasting SPI-12 at Dorahi station

Fig. 12 Scatter plots of forecasted and observed SPI-12 using **a** LSTM, **b** extra-tree, **c** VAR and **d** MARS for Dorahi station



less scattered forecasts than the other alternative methods. The RMSE and NSE values of the implemented methods are also compared in Fig. 15. As clearly observed, the LSTM has the lowest RMSE and the highest NSE in all

stations and its accuracy is followed by the MARS method. Taylor diagrams in Figs. 16, 17, 18 and 19 also compare test results of the four applied methods. As evident from these graphs, the LSTM has the highest correlation and its

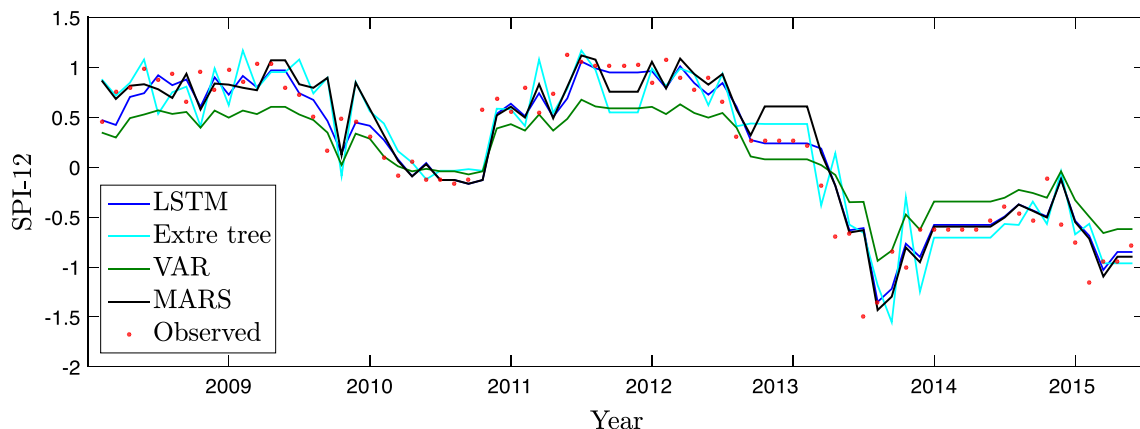
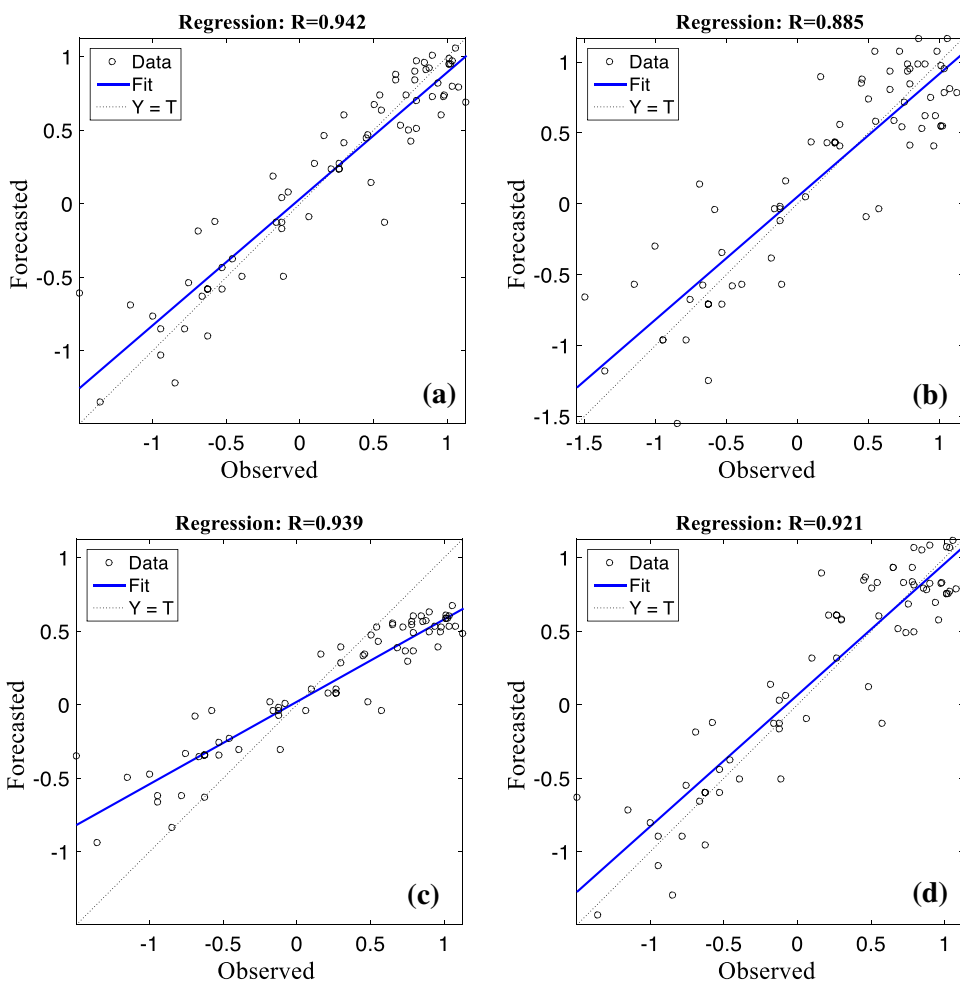


Fig. 13 Observed and forecasted time variation graphs of different models in forecasting SPI-12 at Rooyan station

Fig. 14 Scatter plots of forecasted and observed SPI-12 using a LSTM, b extra-tree, c VAR and d MARS for Rooyan station



standard deviation is closer to the observed one in all four stations. The less accuracy of ET and VAR compared to MARS is clearly observed from the Taylor diagrams.

Overall, the numerical and visual assessment of the application outcomes indicated that the LSTM is superior to the ET, VAR and MARS methods in forecasting

droughts based on SPI index. It was observed that the accuracy of the methods increases by the increment in the window of SPI (from SPI-3 to SPI-9). LSTM is a deep learning method and having some advantages to classical methods as also mentioned before. The most important one is the memory of its cells which is able to keep information

Fig. 15 **a** RMSE and **b** NSE values of different models in forecasting SPI-12

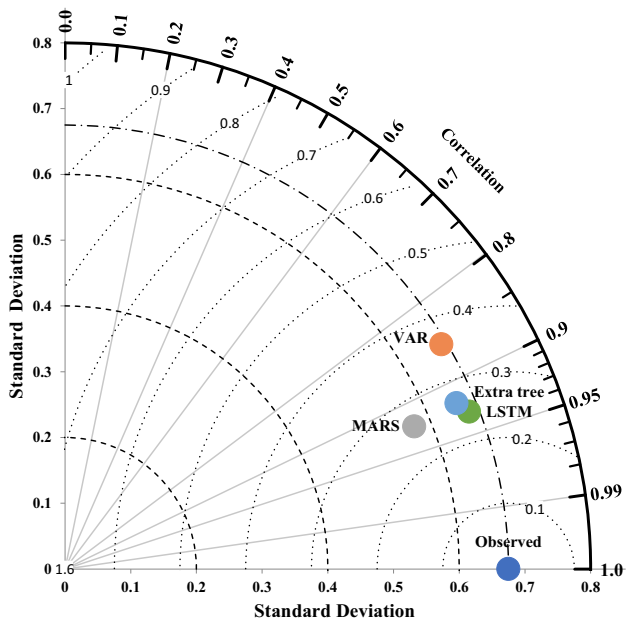
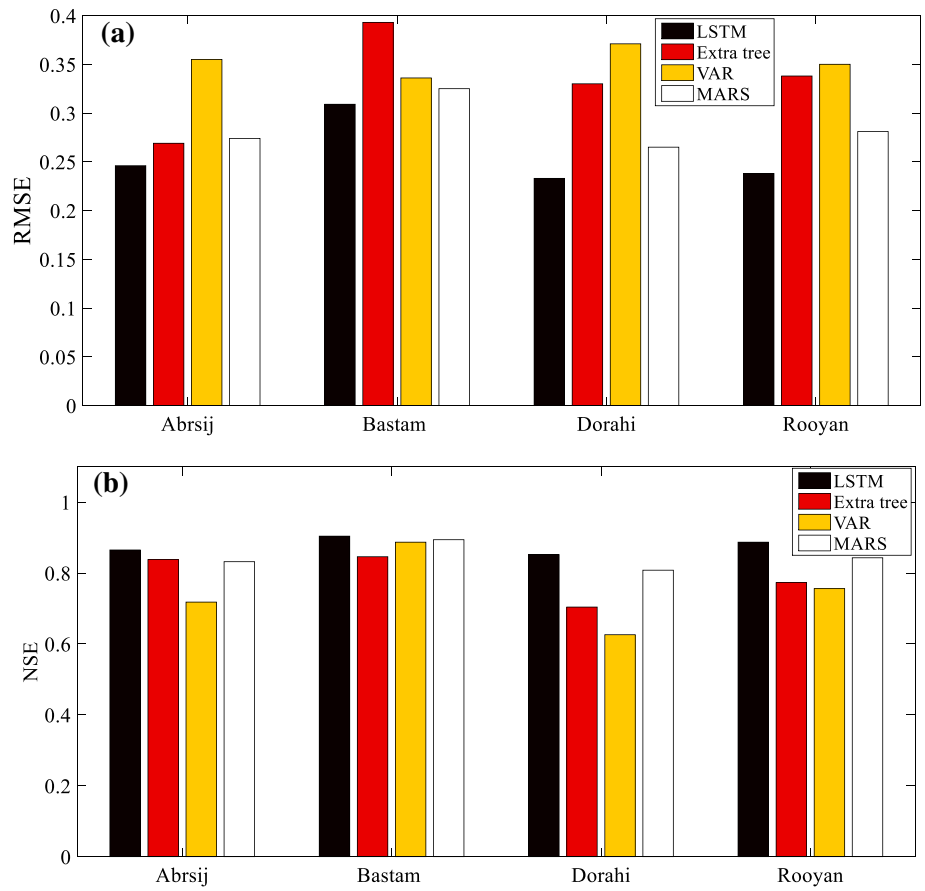


Fig. 16 Taylor diagram of different models in forecasting SPI-12 at Abrsij station

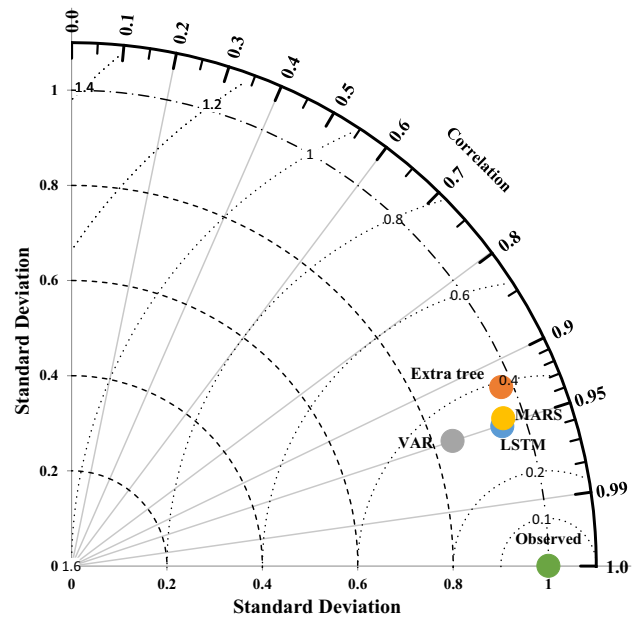


Fig. 17 Taylor diagram of different models in forecasting SPI-12 at Bastam station

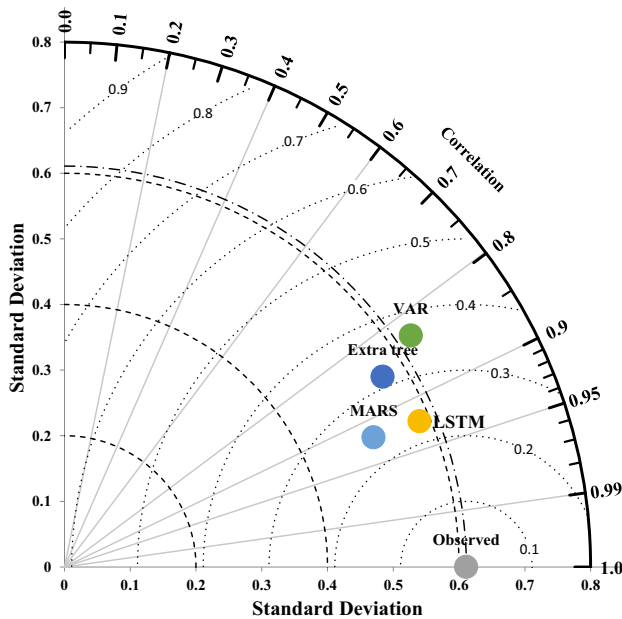


Fig. 18 Taylor diagram of different models in forecasting SPI-12 at Dorahi station

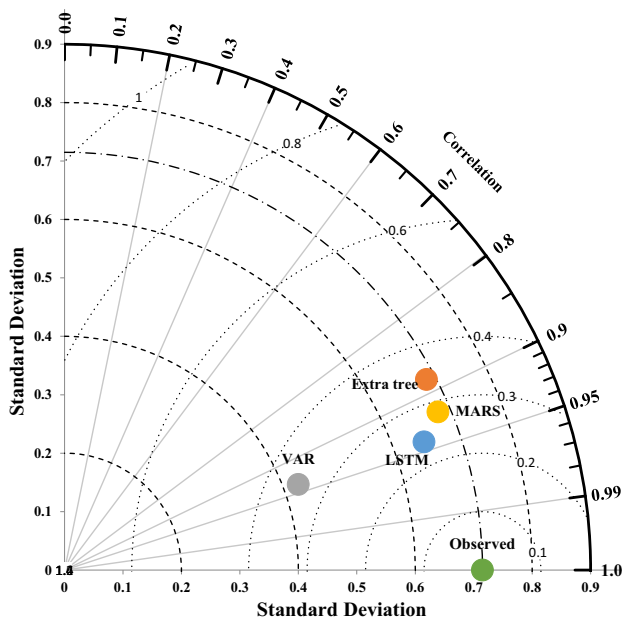


Fig. 19 Taylor diagram of different models in forecasting SPI-12 at Rooyan station

of previous time step, and this causes more effective learning process.

The presented study (LSTM model) is compared with the existing literature in Table 7 with respect to correlation

coefficient (R). Belayneh and Adamowski [7] used wavelet neural networks (WNN) in estimating SPI-3 and SPI-12 and obtained R of 0.883 and 0.953 for the best models, respectively. Jalalkamali et al. [21] applied MLP ANN and ANFIS methods in modelling SPI-3, SPI-6 and SPI-9 and the optimal models provided the R as 0.895, 0.880 and 0.840, respectively. Kisi et al. [29] compared the ANFIS and metaheuristic algorithms in estimation of different SPI indices, and they found that the optimal models gave R of 0.872, 0.922, 0.878 and 0.912 for the SPI-3, SPI-6, SPI-9 and SPI-12, respectively. The R ranges of the best model (LSTM) of the present study obtained through the four stations clearly indicate that the proposed model successfully estimates four SPI indices.

4 Conclusion

The study investigated the ability of LSTM deep learning method in forecasting four SPI indices obtained from four stations of Iran. The outcomes produced by this method were compared with extra-trees, VAR and MARS methods considering numerical criteria, RMSE, NSE and R and visual methods, time variation graphs, scatter plots and Taylor diagrams. The overall outcomes tell us that the LSTM method performed superior to the other methods in forecasting droughts based on SPI-3, SPI-6, SPI-9 and SPI12. The proposed method considerably improved the forecasting accuracy; improvements in RMSE accuracy of the extra-trees, VAR and MARS are by about 17.1%, 12.8% and 9.6% for SPI-3, by 10.5%, 6.2% and 5% for SPI-6, by 7.3%, 4.1% and 6.2% for SPI-9 and by 22.2%, 27% and 10.6% for SPI-12. The accuracy rank of the methods applied is LSTM > MARS > VAR > ET in forecasting drought based on standard precipitation index. The evaluation outcomes recommend the use of deep learning method (LSTM) in forecasting SPI-based droughts. This can be useful for the manager of decision-makers in planning water resources and management. The main limitation of the presented method is having more complex structure and training duration compared to MARS, VAR and ET methods. By the advancement in the computer science (advanced computers), however, this problem can be easily coped. The generalization of the present study can be improved by using different benchmarks from different regions of the globe.

Table 7 Comparison with the existing literature in forecasting SPI

Study	Applied models	Modelled index	Location	R
Belayneh and Adamowski [7]	WNN	SPI-3	Ethiopia	0.883
Belayneh and Adamowski [7]	WNN	SPI-12	Ethiopia	0.953
Jalalkamali et al. [21]	MLP ANN	SPI-3	Iran	0.895
Jalalkamali et al. [21]	ANFIS	SPI-6	Iran	0.880
Jalalkamali et al. [21]	MLP ANN	SPI-9	Iran	0.840
Hosseini-Moghari et al. [20]	ICA-RSVR	SPI-3	Iran	0.228
Hosseini-Moghari et al. [20]	ICA-RSVR	SPI-6	Iran	0.409
Hosseini-Moghari et al. [20]	ICA-RSVR	SPI-9	Iran	0.679
Hosseini-Moghari et al. [20]	ICA-RSVR	SPI-12	Iran	0.790
Kisi et al. [29]	ANFIS-PSO	SPI-3	Iran	0.872
Kisi et al. [29]	ANFIS-PSO	SPI-6	Iran	0.922
Kisi et al. [29]	ANFIS-BOA	SPI-9	Iran	0.878
Kisi et al. [29]	ANFIS-PSO	SPI-12	Iran	0.912
Present study	LSTM	SPI-3	Iran	0.836–0.856
Present study	LSTM	SPI-6	Iran	0.889–0.909
Present study	LSTM	SPI-9	Iran	0.891–0.914
Present study	LSTM	SPI-12	Iran	0.925–0.951

WNN wavelet neural networks, MLP ANN multi-layer perceptron ANN, ICA-RSVR imperialist competitive algorithm and recursive support vector regression, ANFIS-PSO ANFIS with particle swarm optimization, ANFIS-BOA ANFIS with butterfly optimization algorithm

Declarations

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

References

- Adamowski J, FungChan H, Prasher SO, Ozga-Zielinski B, Sliusarieva A (2012) Comparison of multiple linear and nonlinear regression, autoregressive integrated moving average, artificial neural network, and wavelet artificial neural network methods for urban water demand forecasting in Montreal, Canada. *Water Resour Res* 48:W01528. <https://doi.org/10.1029/2010WR009945>
- Adnan RM, Liang Z, Parmar KS, Soni K, Kisi O (2020) Modeling monthly streamflow in mountainous basin by MARS, GMDHNN and DENFIS using hydro-climatic data. *Neural Comput Appl*. <https://doi.org/10.1007/s00521-020-05164>
- Alizamir M, Kim S, Kisi O, Zounemat-Kermani M (2020) Deep echo state network: a novel machine learning approach to model dew point temperature using meteorological variables. *Hydrol Sci J*. <https://doi.org/10.1080/02626667.2020.1735639>
- Alizamir M, Kim S, Kisi O, and Zounemat-Kermani, M. 2020b. A comparative study of several machine learning based non-linear regression methods in estimating solar radiation: Case studies of the USA and Turkey regions. *Energy*: 117239.
- Alizamir M, Kim S et al (2020) Kernel extreme learning machine: an efficient model for estimating daily dew point temperature using weather data. *Water* 12(9):2600
- Arvind Singh T, Bhawana N, Lokendra S (2015) Drought spells identification with indices for Almora district of Uttarakhand, India. *Am Int J Res Sci Technol Eng Math* 12(1):2328–3491
- Belayneh A, Adamowski J (2013) Drought forecasting using new machine learning methods. *J Water Land Dev* 18:3–12
- Belayneh A, Adamowski J, Khalil B (2016) Short-term SPI drought forecasting in the Awash River Basin in Ethiopia using wavelet transforms and machine learning methods. *Sustain Water Resour Manag* 2:87. <https://doi.org/10.1007/s40899-015-0040-5>
- Bishop CM (1995) *Neural networks for pattern recognition*. Oxford University Press, Oxford, p 482
- Box GEP, Jenkins GM, Reinsel GC (1994) *Time series analysis, forecasting and control*. Prentice Hall, Englewood Cliffs, NJ, USA
- El-Shafie A, Noureldin A, Taha M, Hussain A (2011) Dynamic versus static neural network model for rainfall forecasting at Klang River Basin. *Malaysia Hydrol Earth Syst Sci Discuss* 8:6489–6532
- Fan H, Jiang M, Xu L, Zhu H, Cheng J, Jiang J (2020) Comparison of long short term memory networks and the hydrological model in runoff simulation. *Water* 12:175. <https://doi.org/10.3390/w12010175>
- Farahmand A, AghaKouchak A (2015) A generalized framework for deriving nonparametric standardized drought indicators. *Adv Water Resour* 76:140–145
- Fathian F, Fakheri-Fard A, Ouarda TBMJ, Dinpashoh YS, Mousavi Nadoushani S (2019) Multiple streamflow time series modeling using VAR–MGARCH approach. *Stoch Env Res Risk Assess*. <https://doi.org/10.1007/s00477-019-01651-9>
- Galelli S, Castelletti A (2013) Assessing the predictive capability of randomized tree-based ensembles in streamflow modelling. *Hydrol Earth Syst Sci* 17:2669–2684
- Gers FA, Schmidhuber J, Cummins F (2000) Learning to forget: continual prediction with LSTM. *Neural Comput* 12(10):2451–2471
- Gourabi BR (2010) The recognition of drought with DRI and SIAP method and its effects on rice yield and water surface in Shaft, Gilan, south Western of Caspian Sea. *Aust J Basic Appl Sci* 4(9):4374–4378
- Graves A, Liwicki M, Fernández S, Bertolami R, Bunke H, Schmidhuber J (2008) A novel connectionist system for

- unconstrained handwriting recognition. *IEEE Trans Pattern Anal Mach Intell* 31(5):855–868
19. Geurts P, Ernst D, Wehenkel L (2006) Extremely randomized trees. *Mach Learn* 63:3–42. <https://doi.org/10.1007/s10994-006-6226-1>
 20. Hosseini-Moghari SM, Araghinejad S, Azarnivand A (2017) Drought forecasting using data-driven methods and an evolutionary algorithm. *Model Earth Syst Environ* 3(4):1675–1689
 21. Jalalkamali A, Moradi M, Moradi N (2015) Application of several artificial intelligence models and ARIMAX model for forecasting drought using the Standardized Precipitation Index. *Int J Environ Sci Technol* 12(4):1201–1210
 22. Karavitis CA, Alexandris S, Tsesmelis DE, Athanasopoulos G (2011) Application of the standardized precipitation index (SPI) in Greece. *Water* 3:787–805
 23. Keyantash JA, Dracup JA (2004) An aggregate drought index: assessing drought severity based on fluctuations in the hydrologic cycle and surface water storage. *Water Resour Res* 40(9)
 24. Khadr M (2016) Forecasting of meteorological drought using Hidden Markov Model (case study: The upper Blue Nile river basin, Ethiopia). *Ain Shams Eng J* 7:47–56
 25. Khan MMH, Muhammad NS, El-Shafie A (2018) Wavelet-ANN versus ANN-based model for hydrometeorological drought forecasting. *Water* 10:998
 26. Khashei M, Bijari M (2011) A novel hybridization of artificial neural networks and ARIMA models for time series forecasting. *Appl Soft Comput* 11(2):2664–2675. <https://doi.org/10.1016/j.asoc.2010.10.015>
 27. Kim TW, Valdes JB (2003) Nonlinear model for drought forecasting based on a conjunction of wavelet transforms and neural networks. *J Hydrol Eng* 6:319–328
 28. Kisi O (2011) Wavelet regression model as an alternative to neural networks for river stage forecasting. *Water Resour Manag* 25:579–600
 29. Kisi O, Docheshmeh Gorgij A, Zounemat-Kermani M, Mahdavi-Meymand A, Kim S (2019) Drought forecasting using novel heuristic methods in a semi-arid environment. *J Hydrol* 578:124053
 30. Lloyd-Hughes B, Saunders MA (2002) A drought climatology for Europe. *Int J Climatol* 22:1571–1592
 31. McKee TB, Doesken NJ, Kleist J (1993) The relationship of drought frequency and duration to time scales. In: *Proceedings of the 8th conference on applied climatology*, Vol 17, No. 22, pp. 179–183
 32. Mishra AK, Desai VR (2006) Drought forecasting using stochastic models. *Stoch Env Res Risk Assess* 19(5):326–339
 33. Morid S, Smakhtin V, Moghaddasi M (2006) Comparison of seven meteorological indices for drought monitoring in Iran. *Int J Climatol* 26:971–985
 34. Morid S, Smakhtin V, Moghaddasi M (2006) Comparison of seven meteorological indices for drought monitoring in Iran. *Int J Climatol* 26(7):971–985
 35. 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>
 36. Peng T, Zhou J, Zhang C, Fu W (2017) Streamflow forecasting using empirical wavelet transform and artificial neural networks. *Water* 9:406
 37. Quilty J, Adamowski J, Boucher M-A (2019) A stochastic data-driven ensemble forecasting framework for water resources: a case study using ensemble members derived from a database of deterministic wavelet-based models. *Water Resour Res* 55:175–202. <https://doi.org/10.1029/2018WR023205>
 38. Remesan R, Shamim MA, Han D, Mathew J (2009) Runoff prediction using an integrated hybrid modelling scheme. *J Hydrol* 372:48–60
 39. Sak H, Senior A, and Beaufays F (2014) Long short-term memory based recurrent neural network architectures for large vocabulary speech recognition. arXiv preprint [arXiv:1402.1128](https://arxiv.org/abs/1402.1128)
 40. Shabri A (2015) A hybrid model for stream flow forecasting using wavelet and least Squares support vector machines. *Jurnal Teknologi* 73:89–96
 41. Singh VP (1994) *Elementary hydrology*. Prentice Hall of India, New Delhi, India
 42. Tayyab M, Zhoua J, Adnana R, Zenga X (2017) Application of artificial intelligence method coupled with discrete wavelet transform method. *Procedia Comput Sci* 107:212–217. <https://doi.org/10.1016/j.procs.2017.03.081>
 43. Tsakiris G, Vangelis H (2004) Towards a drought watch system based on spatial SPI. *Water Resour Manag* 18(1):1–12
 44. Van Loon AF, Van Lanen HAJ (2013) Making the distinction between water scarcity and drought using an observation-modeling framework. *Water Resour Res* 49:1483–1502. <https://doi.org/10.1002/wrcr.20147>
 45. Vicente-Serrano SM, Beguería S, López-Moreno JI (2010) A multiscalar drought index sensitive to global warming: the standardized precipitation evapotranspiration index. *J Clim* 23(7):1696–1718
 46. Sanusi W, Jemain AA, Zin WZW, Zahari M (2014) The drought characteristics using the first-order homogeneous Markov chain of monthly rainfall data in Peninsular Malaysia. *Water Resour Manag* 29:1523–1539
 47. Zhang Y, Li W, Chen Q, Pu X, Xiang L (2017) Multi-models for SPI drought forecasting in the north of Haihe River Basin, China. *Stoch Environ Res Risk Assess* 31:2471–2181

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