



Application of artificial intelligence hybrid models for meteorological drought prediction

Seyed Mohammad Ehsan Azimi¹ · Seyed Javad Sadatinejad¹ · Arash Malekian² · Mohammad Hossein Jahangir¹

Received: 6 October 2021 / Accepted: 11 December 2022
© The Author(s), under exclusive licence to Springer Nature B.V. 2022

Abstract

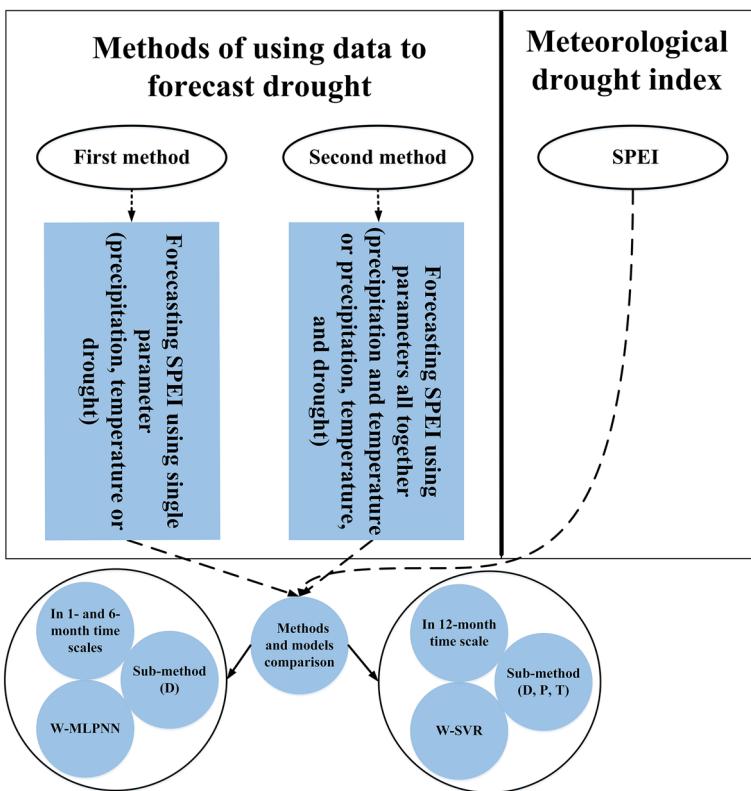
Drought is a prolonged dry period that has a serious impact on health, agriculture, economies, energy, and the environment. Thus, there have been numerous attempts to make this phenomenon more predictable for preventing the aforementioned effects. The present study aims to determine the best combination of input data sets and predict the Standardized Precipitation Evapotranspiration Index (SPEI) in 1, 6, and 12-month time scales using Artificial Intelligence (AI) models (Multilayer Perceptron Neural Network (MLPNN), Support Vector Regression (SVR), Adaptive Neuro-Fuzzy Inference System (ANFIS), and Ensemble Decision Tree (EDT)), which all models are hybridized with a wavelet transformation at three synoptic stations named Ardebil, Khalkhal, and Moghan. To this end, monthly lags of precipitation, temperature, and SPEI were used in northwestern Iran from 1987 to 2018. The methods were classified into single parameter and multiparameter, and each sub-method was designed based on a combination of the parameters. Moreover, Autocorrelation Function (ACF), Partial Autocorrelation Function (PACF), Effective Factor Elimination Technique (EFET), and Feature Scaling (FS) were used to determine the best lags of parameters. In this regard, Root Mean Square Error (RMSE), Correlation Coefficient (CC), and Nash–Sutcliffe Efficiency Index (NSE) were used as the statistical criteria to assess AI models, methods, and sub-methods. The results revealed that the sub-method (D) with W-MLPNN in the 1-month and 6-month time scales and the sub-method (D, P, T) with W-SVR in the 12-month time scale were the best models and sub-methods of this area, respectively. Moreover, based on the results, the efficiency of the AI was enhanced in longer time scales (more than the 6th month) and the longer the time scale, the more the number of lags (under the 6th month) in input data is decreased.

✉ Seyed Mohammad Ehsan Azimi
ehsan.azimi.92@ut.ac.ir

¹ Faculty of New Sciences and Technologies, University of Tehran, Tehran, Iran

² Faculty of Natural Resources, University of Tehran, Karaj 31585-3314, Iran

Graphical Abstract



Keywords Artificial intelligence · Drought forecasting · Hybrid models · Northwestern iran · Time series

1 Introduction

The total damage of natural disasters to the global economy is approximately 33 percent (Keshavarz et al. 2013). Natural disasters caused approximately \$ 27 billion in damage to developing countries' food security annually from 2005 to 2016, of which five percent was due to drought (FAO 2018; UNISDR 2015). Because of this infliction, approximately 55 million people in the world suffer from drought annually (WHO 2020). Finding the most effective ways for recognizing and predicting it to reduce its effects is critical (Mokhtarzad et al. 2017), to this end, it is necessary identifying appropriate drought indices in the first step. Notably, due to the ease of implementation and simplicity in the equation, almost most researchers (e.g., Bacanli et al. 2009; Durdu 2010; Belayneh et al. 2014; Abdourahmane and Acar 2019) have used single parameter indices like standard precipitation index (SPI). Hence, the lack of using multiparameter indices, including the Standardized Precipitation Evapotranspiration Index (SPEI) is obvious in drought studies because of employing the temperature and precipitation balance equation (Vicente-Serrano et al. 2010a, b)

besides all features of single parameter indices (simplicity of use, calculation in different time scales, and probability) (Belayneh and Adamowski 2012).

In the second step, to be aware of drought status for future planning, it is crucial to have a suitable drought prediction model. These models are generally divided into dynamics and statistical classes (Djibo et al. 2015), and researchers have strongly considered statistical models because of developing and designing (Belayneh and Adamowski 2012; Adamowski 2008). Time series of drought are almost nonlinear (Liu et al. 2018), thus water and climate researchers should employ the most significant precise prediction models (Mishra and Desai 2006). Although the efficiency of machine learning models has been proved for water and climate science (Le et al. 2017; Bai et al. 2016; Borji et al. 2016; Choubin et al. 2014), they are not completely applicable to dealing with non-stationary water and climate data (Belayneh and Adamowski 2012). Therefore the tool for preprocessing of data called Wavelet Transform (WT) is utilized to confront the mentioned limitations (Zhang et al. 2017). Kim and Valdés (2003) analyzed the application of WT along the neural network for drought prediction in Mexico and concluded that WT enhances the efficiency of the neural network. Belayneh et al. (2014) compared Multilayer Perceptron Neural Network (MLPNN), Support Vector Regression (SVR), WT-MLPNN, and WT-SVR in the Awash river basin of Ethiopia and found that W-MLPNN is more efficient than the basic models. In another study, Shirmohammadi et al. (2013) found that the W- Adaptive Neuro-Fuzzy Interference System (ANFIS) has the highest efficiency compared to WT-MLPNN, and MLPNN for forecasting SPEI in Azerbaijan Province, Iran.

In the third step, due to many challenges in selecting the best data sets as a method of prediction, many studies have employed various variables for drought prediction. For instance, Dehghani et al. (2017) employed lags of Standardized Hydrological Drought Index (SHDI) for drought forecasting in the Black River basin of the United States. They concluded that using the drought index as one of the variables of input data sets significantly improves the efficiency of models. Also, Bacanli et al. (2009) found that hydroclimatic data and drought indices all together in the input data sets in most cases could enhance the efficiency of AI models for drought forecasting in central Anatolia, Turkey. Poornima and Pushpalatha (2019) also utilized precipitation, temperature, humidity, wind speed, sunshine duration, and drought indices (SPI and SPEI) for drought forecasting in the Hyderabad region. In summary, previous studies have found that hydroclimate variables have positive effects on the results of drought prediction and using preprocessors models become more efficient. The present study aims to (1. Determine the best combination of input data sets among lags of synoptic parameters and drought index, 2. Run and compare the efficiency of hybrid AI models, and 3. Predict SPEI) in 1, 6, and 12-month time scales in Ardebil province for the first time.

2 Methodology

2.1 Study area and data

Ardebil Province is one of the most important provinces located in northwestern Iran ($37^{\circ}45'$ N to $39^{\circ}42'$ N and $47^{\circ}30'$ E to $48^{\circ}55'$ E)) Fig. 1 and covers an area of approximately $17,900 \text{ km}^2$. The minimum and maximum elevation of the area are 20 and 4780 m above sea level, respectively. Based on the Köppen-Geiger climate classification, the climate of this province is Mediterranean ('Csa') (Farajzadeh and Matzarakis 2009). Because

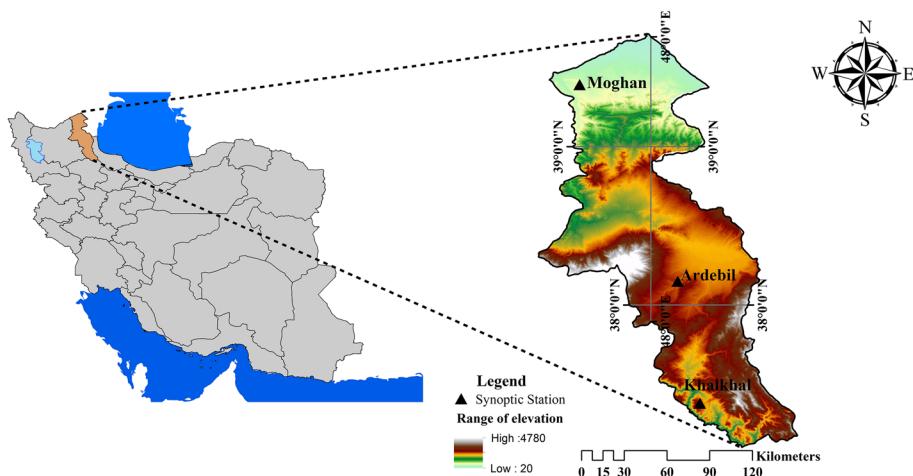


Fig. 1 The geographical location of the study and its synoptic stations

of this climate, winters are often cold with intensive snowfall, springs and autumns are relatively moderate, and summers are hot and arid. It should be noted that a large proportion of the local people's income comes from natural ecosystems, thus drought could play important role in their income. In this regard, it is crucial to have a vision of the upcoming drought status in this region.

In this study, monthly precipitation and temperature data for 31 years (1987–2017) are derived from Iran Meteorological Organization for Ardebil, Khalkhal, and Moghan synoptic stations. It should be noted that although there are lots of synoptic stations in this area, these three stations have the most complete time series. Table 1 presents the specifications of the synoptic stations and the seasonal average of precipitation and temperature.

2.2 Drought index

2.2.1 Standardized precipitation evapotranspiration index

Standardized Precipitation Evapotranspiration Index (SPEI) is developed by Vicente-Serrano et al. (2010a, b) based on the evapotranspiration system to assess the drought status in different time scales for most latitudes and longitudes (Zhang et al. 2019). SPEI is designed by multi equations which Potential Evapotranspiration (PET) is one of the most important

Table 1 General synoptic stations' specifications

| Station | Longitude | Latitude | Elevation (m) | Spring | | Summer | | Autumn | | Winter | |
|----------|-----------|----------|---------------|--------|----|--------|----|--------|----|--------|----|
| | | | | T | P | T | P | T | P | T | P |
| Ardebil | 48°17' E | 35°18' N | 1332 | 6 | 30 | 10 | 8 | 1 | 27 | -4 | 26 |
| Khalkhal | 48°31' E | 38°37' N | 1796 | 4 | 45 | 12 | 9 | -1 | 35 | -6 | 34 |
| Moghan | 47°55' E | 39°39' N | 31 | 13 | 26 | 19 | 14 | 6 | 28 | 1 | 22 |

The temperature (T) and precipitation (P) unities are degrees celsius and millimeter, respectively

fundamental computational equations of this index. Numerous equations have been developed to calculate PET, which Thornthwaite (1948) is used in this study. The distribution function is another one of those equations that make this index more flexible and led it to be used in different time scales and locations and should be determined by trial and error to make the index more practical (Jahangir et al. 2023). Based on Behrang Manesh et al. (2019), the Log-Logistic distribution function is suitable for this index in northwestern Iran. Additionally, Table 2 presents SPEI classifications and interested researchers can refer to Vicente-Serrano et al. (2010a, b) to know how it works.

2.3 Artificial intelligence

2.3.1 Wavelet transformation

Wavelet Transformation (WT) is a mathematical tool that reduces non-stationary effects in time series (Poornima and Pushpalatha 2019). In most cases, WT can better discover data trends such as breakdown points, discontinuities, and local minimums or maximums rather than other signal analysis techniques (Kim and Valdés 2003). Considering the advantages of WT, this tool can help enhance the efficiency of AI models in water and climate science, of which Continuous Wavelet Transform (CWT) and Discrete Wavelet Transform (DWT) are generally known as two types of this tool. The basic WT equation is defined as:

$$W(\tau, s) = \frac{1}{\sqrt{|s|}} \int_{-\infty}^{\infty} x(t) \psi^* \left(\frac{t - \tau}{s} \right) dt \quad (1)$$

where (s) is the scale parameter, (τ) is the change, and (*) is related to the complex conjugate (Kim and Valdes, 2003). Due to the simplicity of calculations and efficiency in dealing with a small dataset, DWT (Cannas et al. 2005) is usually preferred to CWT (Zhang et al. 2017).

2.3.2 Multilayer perceptron neural network

The concept of Artificial Neural Network (ANN) is inspired by the operation of the neuron in the organisms' neural network to tackle problems that might not be solved by traditional mathematical equations. This technique can provide unique methods to model complex systems by establishing relations between input and target data (Sulaiman et al. 2018),

Table 2 SPEI classifications (Vicente-Serrano et al. 2010a, b)

| Drought class | SPEI value | Probability of cumulative distribution function |
|------------------|---------------|---|
| Extremely wet | ≥ 2 | ≥ 0.935 |
| Severely wet | 1.5 to 1.99 | 0.934 |
| Moderately wet | 1 to 1.49 | 0.842 |
| Normal | -0.99 to 0.99 | 0.5 |
| Moderate drought | -1 to -1.49 | 0.158 |
| Severe drought | -1.5 to -1.99 | 0.066 |
| Extreme drought | ≤ -2 | ≤ 0.022 |

of which one of the most well-known ANN models is the Multilayer Perceptron Neural Network (MLPNN). This model has three layers (input, hidden, and output), of which the input layer involves input and target data that are always single; the hidden layer involves at least one layer, of which each layer has at least one neuron that is responsible for developing a relationship between input and target data, and the output layer involves a single presents the output data) Fig. 2).

In the MLPNN structure, the net is a developed model, b is a neuron's bias, w is a neuron's weight, and p symbolizes the input data. It should be noted that all neurons and layers are given a weight that is specified by training algorithms.

2.3.3 Adaptive neuro-fuzzy inference system

Adaptive Neuro-Fuzzy Inference System (ANFIS) refers to a technique using the training algorithms of neural networks for modeling (Brown and Harris 1994) that was developed by Jang (1993). The structure of ANFIS involves many nodes and layers that are connected to directional links, of which each node has constant and adjustable parameters. Utilizing linguistic knowledge and neural network learning algorithms in this model simultaneously leads to a model with wide applications. In this way, Fullér (2000) presented a set of similarities and differences between ANFIS and ANN.

According to the architecture of ANFIS)Fig. 3(, a structure of ANFIS is composed of fuzzification, rules and arguments, normalization, defuzzification, and cumulative layers. In this architecture, the first layer classifies data according to user-defined fuzzy rules, inputs are multiplied in each node to yield the weight of the rules in the second layer, the relative rule weights are calculated in the third layer, and each node calculates the contribution of rules to the output layer in the fourth layer (Mohammadi et al. 2014). The last layer is eventually the output layer minimizes the discrepancy between the acquired output and the actual output (Mokhtarzad et al. 2017). For more information about the algorithms and ANFIS Jang (1993) is suggested.

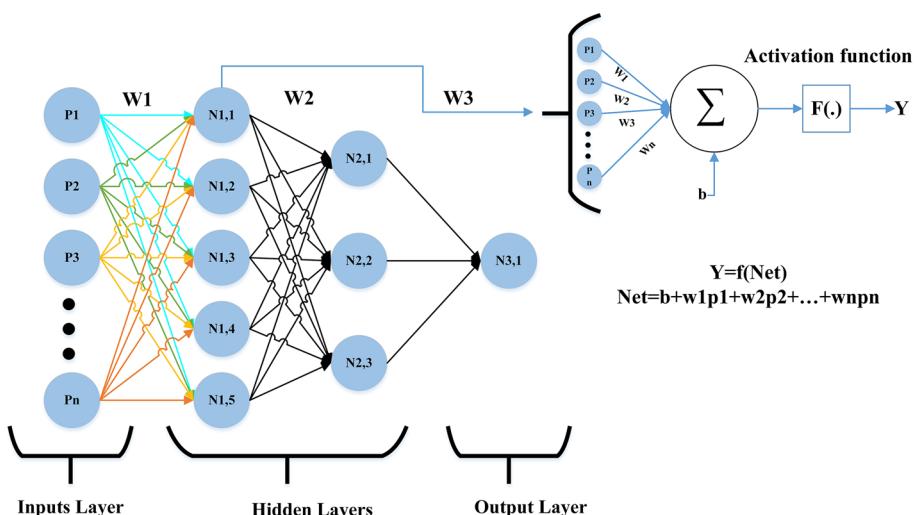


Fig. 2 A schematic of MLPNN structure

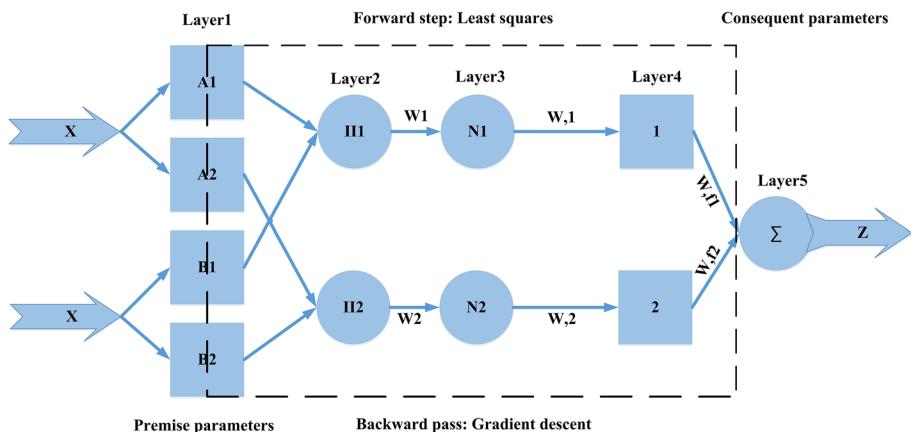


Fig. 3 The general architecture of ANFIS

2.3.4 Support vector regression (SVR)

Support Vector Regression (SVR) was proposed by Sain and Vapnik (1996) to provide a model based on structural error reduction that models data by estimating the correlation function between training and target data. One of the critical components of SVR that is responsible for dividing the space between the input and target data and identifying the best space based on the maximum margin with training data is the Kernel function (KF) (Gunn 1998). SVR consists of three main parameters, namely gamma (γ), cost (C), and epsilon (ϵ) are responsible for controlling complexity, limiting model dimensions, controlling capacity, and calculating the cost function, respectively (Kisi and Cimen 2011). Interested researchers can refer to Sain and Vapnik (1996) to know about its learning algorithms and computational details.

Based on SVR architecture (Fig. 4), Support Vectors (SVs) are derived from the input data, the model is run by KFs, training algorithms define their parameters, and the process of modeling finally ends by adding biases and weights to the output.

2.3.5 Ensemble decision tree (EDT)

Ensemble Decision Tree (EDT) is one of the most popular models that has been used by water and climate researchers (Chen et al. 2012). EDT refers to an ensemble of decision trees that are designed to improve AI models' resistance to overfitting, which offers more reliable predictions than a single decision tree (Rhee and Im 2017). To be aware of EDT, it is necessary to know single decision tree concepts. A decision tree architecture involves a set of nodes, branches, and leaves, of which each one has a specific role (Breiman et al. 1996). In this architecture, internal nodes and leaves specify input parameters and target data, respectively. In the process of modeling, all training data are first defined and processed in a first root node, then sub-root nodes divide training data as long as the parameters are optimized (each branch leads to a leaf). In this respect, EDT refers to a model that utilizes several decision trees simultaneously to be more applicable.

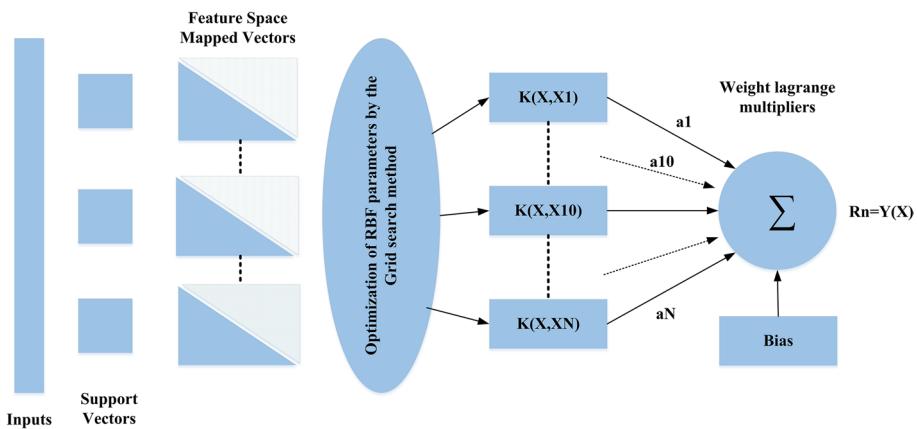


Fig. 4 The general architecture of SVR

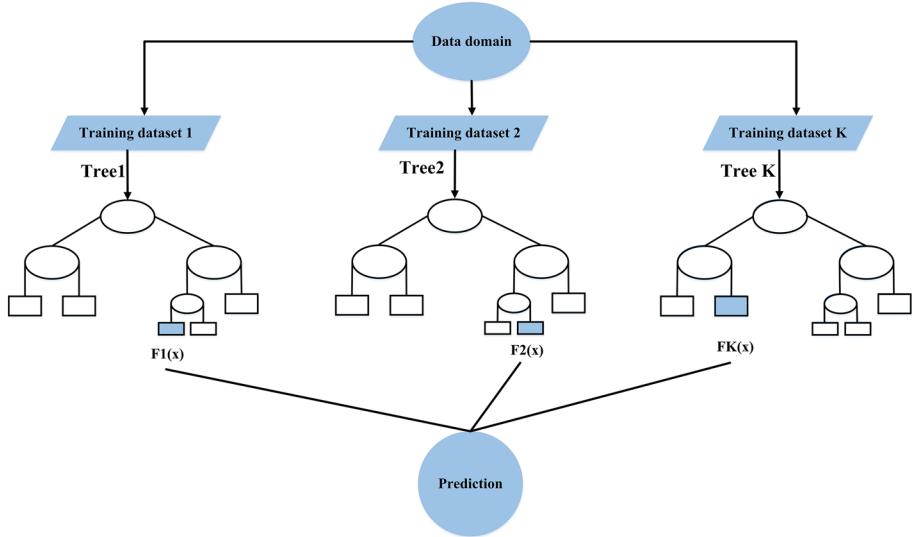


Fig. 5 The architecture of EDT

The architecture of an EDT (Fig. 5) depicts that the input data are divided into K classes, then each class is fed into a decision tree, the training algorithm is applied for each decision tree (model), and the final result is eventually obtained by calculating a weighted average of each output of the decision trees.

2.4 Input data selection and model development

Selecting appropriate input data sets is a time-consuming process, especially in nonlinear processes and there are not lots of ways to determine input data sets for drought prediction. The present study aims to employ synoptic and drought data in two methods (inputs of

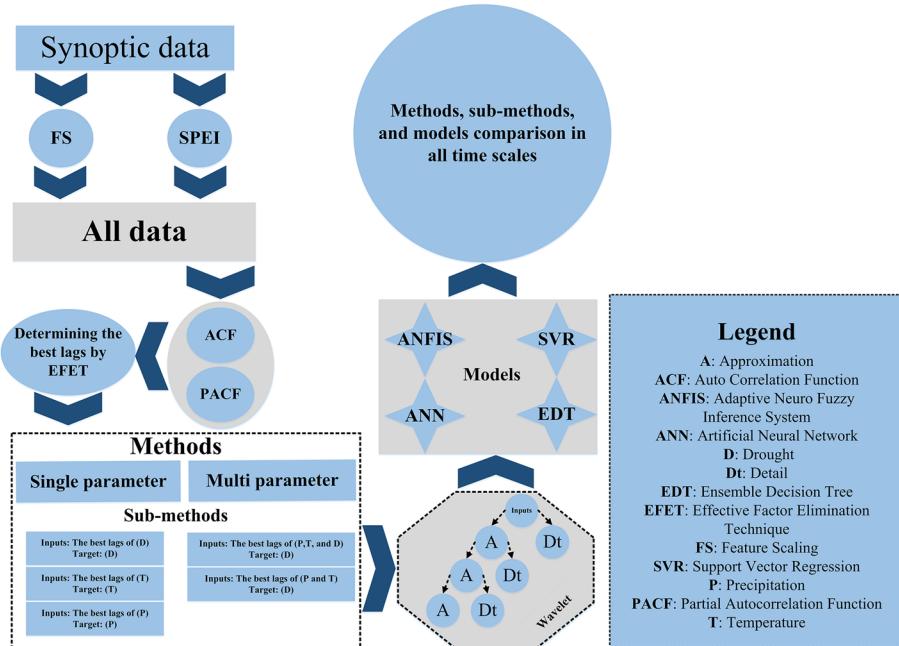


Fig. 6 Flowchart of the present research

models with single or multiple data sets) and five sub-methods (combination of the parameters in the input) to predict drought using hybrid AI models (Fig. 6). For instance, a single parameter method might just use a drought index as input and output, while a multiparameter method would utilize several parameters (e.g., rainfall, temperature, evaporation, etc.) as input and a drought index as output. In the first step, the SPEI 1-, 6-, and 12-months are calculated at 3 synoptic stations in Ardebil province. In the second step, the numerical scales of temperature and precipitation are transformed into the numerical scale of drought using by Scaling (FS) that is used to normalize the numerical range variables. In data processing, it is also known as data normalization and is generally performed during the data preprocessing step, which the formula for normalization is given as:

$$X' = a + \frac{(x - \min(x))(b - a)}{\max(x) - \min(x)} \quad (2)$$

where (a) is the minimum value of drought, (b) is the maximum value of drought, (x) is the value of Temperature or Precipitation (TOP), and $\min(x)$ and $\max(x)$ are the minimum and maximum of the value array of the TOP, respectively. One of the most critical aspects of determining the best combination of input data sets in modeling is ascertaining the seasonality of data. Hence in the third step, Auto Correlation Function (ACF) and Partial Autocorrelation Function (PACF) are used to measure the seasonality or serial correlation between the variables and show correlation between targets and inputs within the lags, respectively (Malik et al. 2019). In this regard, both (ACF and PACF) are utilized to determine how many lags have significant correlations with the target within the past. In the fourth step, the Effective Factor Elimination Technique (EFET) is used to define

the best combination of lags by eliminating all parameters one by one to find a significant correlation between each lag of rescaled parameters (temperature and precipitation) and drought through R^2 . It should be noted that in the process of determining the input data if the R^2 between the lag and target is more than 0.75, it is considered as one of the variables of an input data set.

After determining the best combination of data based on sub-methods, WT is used to decompose the data into two levels as the end of the matter of pre-modeling. And now the efficiency of methods, sub-methods, and models is analyzed by statistical criteria as the last step. This process is based on two methods and five sub-methods, of which the first method involves three sub-methods (each sub-method uses the lags of a single parameter) and the second method involves two sub-methods (each sub-method uses the lags of multiparameter).

2.5 Performance criteria

In this study uses the following three metrics, namely Root Mean Square Error (RMSE), Correlation Coefficient (CC), and Nash–Sutcliffe Efficiency Index (NSEI) as statistical criteria to evaluate the performance of the methods, sub-methods, and models:

$$\text{RMSE} = \sqrt{\frac{1}{N} \sum_{i=1}^N (\text{SPEI}_{\text{Model}} - \text{SPEI}_{\text{Target}})^2} \quad (0 < \text{RMSE} < \infty) \quad (3)$$

$$\text{NSEI} = 1 - \left[\frac{\sum_{i=1}^N (\text{SPEI}_{\text{Model}} - \text{SPEI}_{\text{Target}})^2}{\sum_{i=1}^N (\text{SPEI}_{\text{Model}} - \overline{\text{SPEI}}_{\text{Target}})^2} \right] \quad (-\infty < \text{NSEI} < 1) \quad (4)$$

$$\text{CC} = \frac{\sum_{i=1}^N (\text{SPEI}_{\text{Model}} - \overline{\text{SPEI}}_{\text{Target}})(\text{SPEI}_{\text{Model}} - \overline{\text{SPEI}}_{\text{Target}})}{\sqrt{\sum_{i=1}^N (\text{SPEI}_{\text{Model}} - \overline{\text{SPEI}}_{\text{Target}})^2 \sum_{i=1}^N (\text{SPEI}_{\text{Model}} - \overline{\text{SPEI}}_{\text{Target}})^2}} \quad (-1 < \text{CC} < 1) \quad (5)$$

3 Results and discussion

3.1 Results

3.1.1 Input data set

Determining the best input data sets for AI models to predict drought was one of the main aims of this study. To this end, the lags of precipitation, temperature, and drought index to 13 months ago based on the results of PACF and ACF in three-time scales (due to space limitation, only the graph of one station in one-time scale is shown in (Fig. 7) as an instance) were considered as input. According to Table 3, in all sub-methods, the target had a direct relationship to the abundance of the lags, so the longer the time scale, the more the number of lags (under the 6th month) was decreased in the input data sets. The

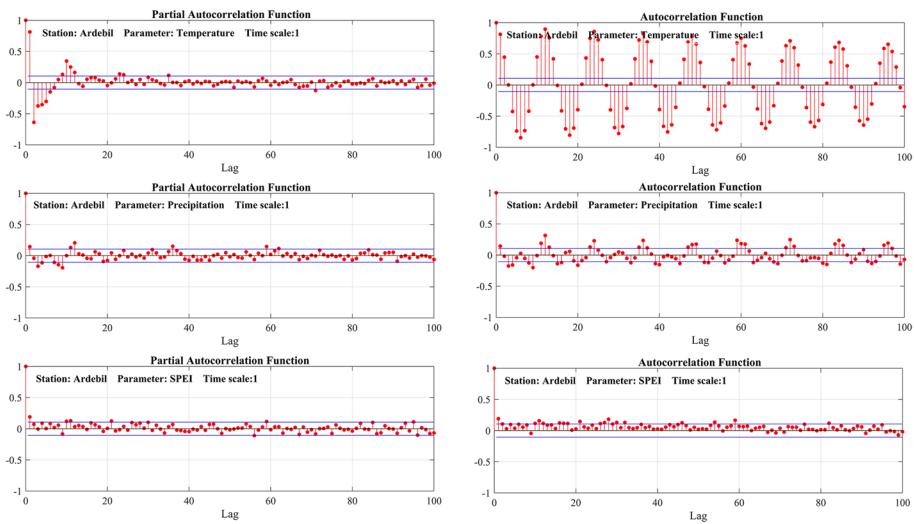


Fig. 7 Correlation of PACF and ACF of precipitation, temperature, and SPEI in 1-month time scale at Ardebil station

combination of the input data sets showed that the abundance of T was more than P in the sub-method (P, T) and D had the most in the sub-method (D, P, T) in all time scales.

3.1.2 Statistical assessment

3.1.2.1 Ardebil station As Table 4 showed, the sub-methods (D) and (D, P, T) provided the best results in all time scales. Based on the average results of all the models in the 1-month time scale, RMSE, NSE, and CC for the sub-method (D) were 0.75, 0.486, and 0.74, respectively, and RMSE, NSEI, and CC for the sub-method (D, P, T) were 0.788, 0.321, and 0.673, respectively, which the sub-method (D) outperformed all. In addition, W-SVR and W-MLPNN exhibited better efficiency than other models, which W-MLPNN in the sub-method (D) had the best results with RMSE = 0.579, NSEI = 0.690, and CC = 0.841. Based on the average results of all the models in the 6-month time scale, RMSE, NSEI, and CC for the sub-method (D) were 0.213, 0.626, and 0.896, respectively, and RMSE, NSEI, and CC for the sub-method (D, P, T) were 0.260, 0.817, and 0.887, respectively, which the sub-method (D) outperformed like the 1-month time scale. Moreover, W-MLPNN in the sub-method (D) had the best results among all models with RMSE = 0.129, NSEI = 0.936, and CC = 0.964. In the 12-month time scale eventually, based on the average results of all the models, RMSE, NSEI, and CC for the sub-method (D) were 0.08, 0.824, and 0.913, respectively, and RMSE, NSEI, and CC for the sub-method (D, P, T) were 0.236, 0.714, and 0.846, respectively. Similar to the previous time scales, the sub-method (D) outperformed the sub-method (D, P, T), but the best results in this time scale belonged to W-SVR in the sub-method (D, P, T) with RMSE = 0.063, NSEI = 0.927, and CC = 0.973. Figure 8 shows the graphical comparison between the time series of the target and the best model at Ardebil station.

3.1.2.2 Khalkhal station According to Table 4, based on the average of all the models in the 1-month time scale, RMSE, NSE, and CC for the sub-method (D) were 0.922, 0.398, and

Table 3 The selected inputs in each time scale in the stations based on the sub-methods

| Station | Method | Sub-method | Inputs | Target |
|---------|-----------------------------|------------|---|---|
| Ardebil | First (single parameter) | D | SPEI _{t-1} , SPEI _{t-2} -1, SPEI _{t-3} -1, SPEI _{t-4} -1, SPEI _{t-5} -1, SPEI _{t-6} -1, SPEI _{t-8} -1, SPEI _{t-9} -1, SPEI _{t-11} -1 SPEI _{t-6} , SPEI _{t-8} -6 SPEI _{t-12} , SPEI _{t-12} | SPEI _{t+1} -1 SPEI _{t+1} -6 SPEI _{t+1} -12 |
| P | | | P _{t-1} , P _{t-1} , P _{t-2} -1, P _{t-3} -1, P _{t-4} -1, P _{t-6} -1, P _{t-7} -1, P _{t-8} -1, P _{t-9} -1, P _{t-10} -1, P _{t-11} -1 P _{t-6} , P _{t-6} , P _{t-2} -6, P _{t-3} -6, P _{t-4} -6, P _{t-5} -6, P _{t-6} -6, P _{t-7} -6, P _{t-8} -6, P _{t-9} -6, P _{t-10} -6, P _{t-13} -6 P _{t-12} , P _{t-12} , P _{t-2} -12, P _{t-3} -12, P _{t-8} -12, P _{t-9} -12, P _{t-11} -12, P _{t-13} -12 T _{t-1} , T _{t-1} , T _{t-2} -1, T _{t-2} -1, T _{t-3} -1, T _{t-4} -1, T _{t-5} -1, T _{t-6} -1, T _{t-7} -1, T _{t-8} -1, T _{t-9} -1, T _{t-10} -1, T _{t-11} -1, T _{t-12} -1 T _{t-6} , T _{t-6} , T _{t-2} -6, T _{t-3} -6, T _{t-4} -6, T _{t-5} -6, T _{t-6} -6, T _{t-7} -6, T _{t-9} -6, T _{t-11} -6, T _{t-12} -6, T _{t-13} -6 T _{t-12} , T _{t-12} , T _{t-3} -12, T _{t-5} -12, T _{t-9} -12, T _{t-10} -12, T _{t-11} -12, T _{t-12} -12, T _{t-13} -12 P _{t-1} , P _{t-1} , P _{t-2} -1, P _{t-3} -1, P _{t-4} -1, P _{t-6} -1, P _{t-8} -1, P _{t-10} -1 T _{t-1} , T _{t-1} , T _{t-2} -1, T _{t-2} -1, T _{t-4} -1, T _{t-5} -1, T _{t-12} -1 P _{t-6} , P _{t-6} , P _{t-3} -6, P _{t-4} -6, P _{t-5} -6, P _{t-6} -6, P _{t-7} -6, P _{t-9} -6, P _{t-13} -6, T _{t-6} , T _{t-6} , T _{t-2} -6, T _{t-3} -6, T _{t-4} -6, T _{t-5} -6, T _{t-6} -6, T _{t-7} -6, T _{t-9} -6, T _{t-11} -6, T _{t-12} -6, T _{t-13} -6 P _{t-12} , P _{t-12} , P _{t-2} -12, P _{t-3} -12, P _{t-8} -12, P _{t-9} -12, P _{t-12} , P _{t-3} -12 T _{t-12} , T _{t-12} , T _{t-3} -12, T _{t-5} -12, T _{t-9} -12, T _{t-10} -12, T _{t-11} -12, T _{t-12} -12 SPEI _{t-1} , SPEI _{t-1} , SPEI _{t-2} -1, SPEI _{t-4} -1, SPEI _{t-8} -1, SPEI _{t-9} -1 P _{t-1} , P _{t-1} , P _{t-6} -1, P _{t-8} -1 T _{t-1} , T _{t-1} , T _{t-12} -1 SPEI _{t-6} , SPEI _{t-6} , SPEI _{t-3} -6, SPEI _{t-5} -6, P _{t-6} , P _{t-6} , P _{t-3} -6, P _{t-4} -6, P _{t-6} -6, P _{t-9} -6, T _{t-6} , T _{t-6} , T _{t-2} -6, T _{t-3} -6, T _{t-4} -6, T _{t-5} -6, T _{t-6} -6, T _{t-9} -6, T _{t-11} -6, T _{t-12} -6, T _{t-13} -6 SPEI _{t-12} , SPEI _{t-12} , SPEI _{t-12} , SPEI _{t-9} -12, SPEI _{t-10} -12, SPEI _{t-11} -12, SPEI _{t-12} -12 P _{t-12} , P _{t-12} , P _{t-3} -12, P _{t-8} -12 T _{t-12} , T _{t-12} , T _{t-5} -12 | SPEI _{t+1} -1 SPEI _{t+1} -6 SPEI _{t+1} -12 |
| T | | P, T | | SPEI _{t+1} -1 SPEI _{t+1} -6 SPEI _{t+1} -12 |
| | Second (multiparameter) | D, P, T | | SPEI _{t+1} -1 |

Table 3 (continued)

| Station | Method | Sub-method | Inputs | Target |
|----------------------------|-----------------------------|------------|--|-------------------------|
| Khalkhal | First (single parameter) | D | SPEI _{t-1} , SPEI _{t-2} -1, SPEI _{t-3} -1, SPEI _{t-4} -1, SPEI _{t-5} -1, SPEI _{t-6} -1, SPEI _{t-9} -1, SPEI _{t-10} -1, SPEI _{t-11} -1 | SPEI _{t+1} -1 |
| | | | SPEI _{t-6} , SPEI _{t-6} | SPEI _{t+1} -6 |
| | | | SPEI _{t-12} , SPEI _{t-12} | SPEI _{t+1} -12 |
| P | | | P _{t-1} , P _{t-1} -1, P _{t-2} -1, P _{t-3} -1, P _{t-4} -1, P _{t-5} -1, P _{t-6} -1 | SPEI _{t+1} -1 |
| | | | P _{t-6} , P _{t-7} , P _{t-8} , P _{t-9} , P _{t-10} , P _{t-11} , P _{t-12} , P _{t-13} | SPEI _{t+1} -6 |
| | | | P _{t-12} , P _{t-12} | SPEI _{t+1} -12 |
| T | | | T _{t-1} , T _{t-1} -1, T _{t-2} -1, T _{t-3} -1, T _{t-4} -1, T _{t-5} -1, T _{t-6} -1, T _{t-7} -1, T _{t-8} -1, T _{t-9} -1, T _{t-10} -1, T _{t-11} -1, T _{t-12} -1, T _{t-13} -1 | SPEI _{t+1} -1 |
| | | | T _{t-6} , T _{t-6} | SPEI _{t+1} -6 |
| | | | T _{t-12} , T _{t-12} | SPEI _{t+1} -12 |
| Second (multiparameter) | P, T | | P _{t-1} , P _{t-1} , P _{t-2} -1, P _{t-3} -1, P _{t-4} -1, P _{t-5} -1, P _{t-6} -1 | SPEI _{t+1} -1 |
| | | | T _{t-1} , T _{t-1} -1, T _{t-2} -1, T _{t-3} -1, T _{t-4} -1, T _{t-5} -1, T _{t-6} -1, T _{t-7} -1, T _{t-8} -1, T _{t-9} -1, T _{t-10} -1, T _{t-11} -1, T _{t-12} -1, T _{t-13} -1 | |
| | | | P _{t-6} , P _{t-7} , P _{t-8} , P _{t-9} , P _{t-10} | SPEI _{t+1} -6 |
| | | | T _{t-6} , T _{t-6} | |
| | | | P _{t-12} , P _{t-12} , P _{t-2} -12, P _{t-4} -12, P _{t-6} -12, P _{t-8} -12, P _{t-10} -12, P _{t-12} -12 | SPEI _{t+1} -12 |
| D, P, T | | | T _{t-12} , T _{t-12} , T _{t-12} , T _{t-12} , T _{t-7} -12, T _{t-8} -12, T _{t-9} -12, T _{t-13} -12 | |
| | | | SPEI _{t-1} , SPEI _{t-1} , SPEI _{t-2} -1, SPEI _{t-3} -1, SPEI _{t-5} -1, SPEI _{t-9} -1 | |
| | | | P _{t-1} , P _{t-1} , P _{t-2} -1, P _{t-3} -1, P _{t-4} -1, P _{t-5} -1, P _{t-6} -1 | |
| | | | T _{t-1} , T _{t-1} -1, T _{t-2} -1, T _{t-3} -1, T _{t-4} -1, T _{t-5} -1, T _{t-6} -1, T _{t-8} -1, T _{t-11} -1, T _{t-13} -1, | |
| | | | SPEI _{t-6} , SPEI _{t-6} | |
| | | | P _{t-6} , P _{t-6} | |
| | | | T _{t-6} , T _{t-6} | |
| | | | SPEI _{t-12} , SPEI _{t-12} , SPEI _{t-6} -12, SPEI _{t-6} -12, SPEI _{t-11} -12, SPEI _{t-12} -12, | |
| | | | P _{t-12} , P _{t-12} , P _{t-2} -12, P _{t-7} -12, P _{t-12} -12 | |
| | | | T _{t-12} , T _{t-12} , T _{t-3} -12, T _{t-6} -12, T _{t-7} -12, T _{t-8} -12, T _{t-9} -12 | |

Table 3 (continued)

| | Station | Method | Sub-method | Inputs | Target |
|---------|-----------------------------|--------|---|---|--|
| Moghani | First (single parameter) | D | SPEI _{t-1} , SPEI _{t-2} -1, SPEI _{t-3} -1, SPEI _{t-4} -1, SPEI _{t-5} -1, SPEI _{t-6} -1, SPEI _{t-7} -6, SPEI _{t-8} -6, SPEI _{t-9} -6, SPEI _{t-10} -6, SPEI _{t-11} -6, SPEI _{t-12} -6 | SPEI _{t-6} , SPEI _{t-1} -6, SPEI _{t-2} -6, SPEI _{t-3} -6, SPEI _{t-4} -6, SPEI _{t-5} -6, SPEI _{t-6} -6, SPEI _{t-7} -6, SPEI _{t-8} -6, SPEI _{t-9} -6, SPEI _{t-10} -6, SPEI _{t-11} -6, SPEI _{t-12} -6 | SPEI _{t+1} -1, SPEI _{t+1} -6 |
| P | Second (multiparameter) | T | P _{t-1} , P _{t-1} -1, P _{t-2} -1, P _{t-3} -1, P _{t-4} -1, P _{t-5} -1, P _{t-6} -1, P _{t-8} -1 P _{t-6} , P _{t-1} -6, P _{t-2} -6, P _{t-3} -6, P _{t-4} -6, P _{t-5} -6, P _{t-6} -6, P _{t-7} -6, P _{t-8} -6, P _{t-9} -6, P _{t-10} -6, P _{t-13} -6 P _{t-12} , P _{t-1} -12, P _{t-8} -12, P _{t-11} -12, P _{t-12} -12 T _{t-1} , T _{t-1} -1, T _{t-2} -1, T _{t-4} -1, T _{t-6} -1, T _{t-10} -1, T _{t-11} -1 T _{t-6} , T _{t-1} -6, T _{t-3} -6, T _{t-5} -6, T _{t-11} -6, T _{t-12} -6, T _{t-13} -6 T _{t-12} , T _{t-1} -12, T _{t-2} -12, T _{t-4} -12, T _{t-5} -12, T _{t-6} -12, T _{t-7} -12, T _{t-8} -12, T _{t-9} -12, T _{t-10} -12, T _{t-11} -12, T _{t-12} -12, T _{t-13} -12 | P _{t-1} , P _{t-1} -1, P _{t-2} -1, P _{t-3} -1, P _{t-4} -1, P _{t-6} -1, T _{t-1} , T _{t-1} -1, T _{t-2} -1, T _{t-6} -1, T _{t-10} -1 P _{t-6} , P _{t-1} -6, P _{t-3} -6, P _{t-4} -6, P _{t-5} -6, P _{t-7} -6, P _{t-9} -6, P _{t-12} -6 T _{t-6} , T _{t-1} -6, T _{t-3} -6, T _{t-5} -6, T _{t-11} -6, T _{t-12} -6, T _{t-13} -6 P _{t-12} , P _{t-1} -12, P _{t-8} -12, P _{t-12} -12 T _{t-12} , T _{t-1} -12, T _{t-2} -12, T _{t-5} -12, T _{t-7} -12, T _{t-12} -12 | SPEI _{t+1} -12, SPEI _{t+1} -1, SPEI _{t-3} -12, SPEI _{t-4} -12, SPEI _{t-5} -12, SPEI _{t-6} -12, SPEI _{t-7} -6, SPEI _{t-8} -6, SPEI _{t-9} -6, SPEI _{t-10} -6, SPEI _{t-11} -6, SPEI _{t-12} -6 |
| D, P, T | | | SPEI _{t-1} , SPEI _{t-1} -1, SPEI _{t-2} -1, SPEI _{t-3} -1, SPEI _{t-4} -1, SPEI _{t-6} -1, SPEI _{t-11} -1 P _{t-1} , P _{t-1} -1, P _{t-2} -1, P _{t-3} -1, P _{t-4} -1, P _{t-6} -1 T _{t-1} , T _{t-1} -1, T _{t-2} -1, T _{t-6} -1, T _{t-10} -1 SPEI _{t-6} , SPEI _{t-1} -6, SPEI _{t-2} -6, SPEI _{t-6} -6, SPEI _{t-8} -6, SPEI _{t-11} -6, SPEI _{t-12} -6 P _{t-6} , P _{t-1} -6, P _{t-3} -6, P _{t-4} -6, P _{t-5} -6, P _{t-7} -6, P _{t-12} -6 T _{t-6} , T _{t-1} -6, T _{t-3} -6, T _{t-5} -6, T _{t-11} -6 P _{t-12} , P _{t-1} -12, SPEI _{t-1} -12, SPEI _{t-2} -12, SPEI _{t-3} -12, SPEI _{t-11} -12, SPEI _{t-12} -12 T _{t-12} , T _{t-1} -12, T _{t-2} -12, T _{t-5} -12, T _{t-12} -12 | P _{t-1} , P _{t-1} -1, P _{t-2} -1, P _{t-3} -1, P _{t-4} -1, P _{t-6} -1 T _{t-1} , T _{t-1} -1, T _{t-2} -1, T _{t-6} -1, T _{t-10} -1 SPEI _{t+1} -12, SPEI _{t+1} -6 | |

Table 4 Statistical results of the models at the stations based on the sub-methods in all time scales

| Station | Method | Sub-method | Time scale | W-MLPNN | | | W-ANFIS | | | W-SVR | | | W-EDT | | | Average of all models | | | |
|-------------------------------|------------------------------|------------|------------|---------|-------|-------|---------|-------|-------|-------|-------|-------|-------|-------|-------|-----------------------|-------|-------|--|
| | | | | RMSE | NSEI | CC | RMSE | NSEI | CC | RMSE | NSEI | CC | RMSE | NSEI | CC | RMSE | NSEI | CC | |
| Ardebil (single parameter) | First | D | 1 | 0.579 | 0.690 | 0.841 | 0.913 | 0.194 | 0.753 | 0.614 | 0.621 | 0.8 | 0.925 | 0.442 | 0.566 | 0.757 | 0.486 | 0.74 | |
| | | | 6 | 0.129 | 0.936 | 0.964 | 0.308 | 0.588 | 0.836 | 0.140 | 0.914 | 0.956 | 0.277 | 0.667 | 0.830 | 0.213 | 0.776 | 0.896 | |
| | | | 12 | 0.06 | 0.9 | 0.958 | 0.088 | 0.857 | 0.928 | 0.065 | 0.922 | 0.960 | 0.144 | 0.619 | 0.807 | 0.089 | 0.824 | 0.913 | |
| | P | 1 | 1 | 0.26 | 0.522 | 1.02 | 0.192 | 0.28 | 0.871 | 0.240 | 0.572 | 1.016 | 0.31 | 0.329 | 0.976 | 0.25 | 0.425 | | |
| | | | 6 | 0.451 | 0.327 | 0.640 | 0.599 | 0.391 | 0.331 | 0.409 | 0.277 | 0.574 | 0.431 | 0.398 | 0.505 | 0.472 | 0.348 | 0.512 | |
| | | | 12 | 0.240 | 0.362 | 0.837 | 0.262 | 0.412 | 0.568 | 0.201 | 0.293 | 0.615 | 0.257 | 0.420 | 0.509 | 0.24 | 0.37 | 0.632 | |
| T | T | 1 | 1.08 | 0.16 | 0.29 | 1.17 | 0.156 | 0.244 | 0.951 | 0.094 | 0.441 | 1.195 | 0.180 | 0.218 | 1.09 | 0.147 | 0.298 | | |
| | | | 6 | 0.571 | 0.310 | 0.38 | 0.56 | 0.251 | 0.457 | 0.504 | 0.29 | 0.464 | 0.580 | 0.369 | 0.410 | 0.553 | 0.305 | 0.427 | |
| | | | 12 | 0.299 | 0.380 | 0.513 | 0.541 | 0.379 | 0.508 | 0.280 | 0.4 | 0.525 | 0.363 | 0.422 | 0.540 | 0.37 | 0.395 | 0.521 | |
| | Second (multiparameter) P, T | 1 | 0.863 | 0.312 | 0.601 | 1.01 | 0.21 | 0.360 | 0.705 | 0.432 | 0.667 | 0.901 | 0.322 | 0.490 | 0.869 | 0.319 | 0.529 | | |
| | | | 6 | 0.375 | 0.464 | 0.762 | 0.648 | 0.26 | 0.409 | 0.312 | 0.578 | 0.762 | 0.449 | 0.328 | 0.527 | 0.446 | 0.407 | 0.615 | |
| | | | 12 | 0.299 | 0.48 | 0.789 | 0.412 | 0.31 | 0.586 | 0.182 | 0.590 | 0.805 | 0.238 | 0.491 | 0.531 | 0.282 | 0.467 | 0.677 | |
| D, P, T | D, P, T | 1 | 0.709 | 0.535 | 0.744 | 0.892 | 0.202 | 0.618 | 0.656 | 0.568 | 0.774 | 0.986 | 0.42 | 0.557 | 0.81 | 0.431 | 0.673 | | |
| | | | 6 | 0.155 | 0.907 | 0.956 | 0.495 | 0.65 | 0.768 | 0.146 | 0.906 | 0.953 | 0.246 | 0.738 | 0.873 | 0.26 | 0.817 | 0.887 | |
| | | | 12 | 0.3 | 0.631 | 0.870 | 0.446 | 0.640 | 0.725 | 0.063 | 0.927 | 0.973 | 0.136 | 0.659 | 0.816 | 0.236 | 0.714 | 0.846 | |

Table 4 (continued)

| Station | Method | Sub-method | Time scale | W-MLPNN | | | W-ANFIS | | | W-SVR | | | W-EDT | | | Average of all models | | |
|----------|------------------------------|------------|------------|---------|-------|-------|---------|-------|-------|-------|-------|-------|-------|-------|-------|-----------------------|-------|-------|
| | | | | RMSE | NSEI | CC | RMSE | NSEI | CC | RMSE | NSEI | CC | RMSE | NSEI | CC | RMSE | NSEI | CC |
| Khalkhal | First (single parameter) | D | 1 | 0.87 | 0.552 | 0.749 | 1 | 0.35 | 0.640 | 0.821 | 0.488 | 0.737 | 0.997 | 0.245 | 0.523 | 0.922 | 0.408 | 0.662 |
| | | | 6 | 0.194 | 0.872 | 0.946 | 0.348 | 0.494 | 0.803 | 0.144 | 0.922 | 0.962 | 0.291 | 0.646 | 0.835 | 0.244 | 0.733 | 0.886 |
| | | | 12 | 0.06 | 0.805 | 0.916 | 0.07 | 0.921 | 0.961 | 0.066 | 0.934 | 0.967 | 0.153 | 0.647 | 0.849 | 0.08 | 0.826 | 0.923 |
| | P | 1 | 1 | 0.227 | 0.501 | 1.03 | 0.195 | 0.278 | 0.981 | 0.270 | 0.573 | 1.032 | 0.132 | 0.254 | 1.01 | 0.206 | 0.401 | |
| | | | 6 | 0.52 | 0.58 | 0.506 | 0.533 | 0.220 | 0.419 | 0.418 | 0.272 | 0.587 | 0.543 | 0.228 | 0.403 | 0.503 | 0.325 | 0.478 |
| | | | 12 | 0.236 | 0.652 | 0.756 | 0.209 | 0.342 | 0.731 | 0.203 | 0.385 | 0.796 | 0.778 | 0.316 | 0.450 | 0.356 | 0.423 | 0.683 |
| T | T | 1 | 1.11 | 0.201 | 0.345 | 1.19 | 0.18 | 0.129 | 1.08 | 0.204 | 0.25 | 1.24 | 0.176 | 0.13 | 1.15 | 0.190 | 0.213 | |
| | | | 6 | 0.629 | 0.334 | 0.425 | 0.929 | 0.21 | 0.34 | 0.532 | 0.218 | 0.451 | 0.407 | 0.223 | 0.408 | 0.624 | 0.246 | 0.406 |
| | | | 12 | 0.385 | 0.366 | 0.521 | 0.433 | 0.29 | 0.516 | 0.393 | 0.33 | 0.522 | 0.607 | 0.303 | 0.5 | 0.454 | 0.322 | 0.514 |
| | Second (multiparameter) P, T | 1 | 0.96 | 0.229 | 0.368 | 1.13 | 0.197 | 0.261 | 1.02 | 0.207 | 0.526 | 1.07 | 0.233 | 0.432 | 1.045 | 0.216 | 0.396 | |
| | | | 6 | 0.377 | 0.519 | 0.811 | 0.603 | 0.517 | 0.679 | 0.326 | 0.557 | 0.762 | 0.492 | 0.41 | 0.660 | 0.449 | 0.5 | 0.728 |
| | | | 12 | 0.15 | 0.56 | 0.817 | 0.29 | 0.526 | 0.798 | 0.159 | 0.620 | 0.828 | 0.249 | 0.47 | 0.777 | 0.212 | 0.544 | 0.805 |
| D, P, T | D, P, T | 1 | 0.89 | 0.249 | 0.499 | 1.06 | 0.446 | 0.604 | 0.907 | 0.376 | 0.653 | 0.992 | 0.253 | 0.575 | 0.962 | 0.331 | 0.582 | |
| | | | 6 | 0.15 | 0.923 | 0.967 | 0.327 | 0.554 | 0.824 | 0.144 | 0.912 | 0.956 | 0.275 | 0.684 | 0.839 | 0.224 | 0.768 | 0.896 |
| | | | 12 | 0.076 | 0.735 | 0.882 | 0.114 | 0.804 | 0.909 | 0.053 | 0.929 | 0.969 | 0.155 | 0.641 | 0.861 | 0.09 | 0.77 | 0.905 |

Table 4 (continued)

| Station | Method | Sub-method | Time scale | W-MLPNN | | | W-ANFIS | | | W-SVR | | | W-EDT | | | Average of all models | | |
|---------|-----------------------------|------------|------------|---------|-------|-------|---------|-------|-------|-------|-------|-------|-------|-------|-------|-----------------------|-------|-------|
| | | | | RMSE | NSEI | CC | RMSE | NSEI | CC | RMSE | NSEI | CC | RMSE | NSEI | CC | RMSE | NSEI | CC |
| Moghan | First (single parameter) | D | 1 | 0.865 | 0.33 | 0.614 | 0.982 | 0.230 | 0.685 | 0.840 | 0.436 | 0.710 | 0.998 | 0.205 | 0.560 | 0.921 | 0.3 | 0.642 |
| | | | 6 | 0.147 | 0.919 | 0.955 | 0.467 | 0.271 | 0.699 | 0.196 | 0.871 | 0.934 | 0.369 | 0.545 | 0.789 | 0.294 | 0.651 | 0.844 |
| | | | 12 | 0.081 | 0.882 | 0.940 | 0.18 | 0.667 | 0.876 | 0.092 | 0.911 | 0.956 | 0.182 | 0.659 | 0.834 | 0.133 | 0.779 | 0.901 |
| | | P | 1 | 1.03 | 0.342 | 0.420 | 1.2 | 0.126 | 0.236 | 0.933 | 0.305 | 0.566 | 1.14 | 0.113 | 0.339 | 1.07 | 0.22 | 0.39 |
| | | | 6 | 0.341 | 0.524 | 0.837 | 0.876 | 0.321 | 0.443 | 0.320 | 0.657 | 0.823 | 0.469 | 0.265 | 0.609 | 0.5 | 0.441 | 0.678 |
| | | | 12 | 0.338 | 0.55 | 0.63 | 0.192 | 0.619 | 0.816 | 0.186 | 0.644 | 0.855 | 0.245 | 0.483 | 0.763 | 0.24 | 0.574 | 0.766 |
| T | Second(multiparameter) | T | 1 | 1.187 | 0.382 | 0.522 | 1.26 | 0.114 | 0.229 | 1.08 | 0.395 | 0.308 | 1.14 | 0.1 | 0.32 | 1.16 | 0.247 | 0.344 |
| | | | 6 | 0.45 | 0.41 | 0.53 | 0.96 | 0.36 | 0.414 | 0.551 | 0.49 | 0.579 | 0.742 | 0.21 | 0.501 | 0.675 | 0.367 | 0.506 |
| | | | 12 | 0.331 | 0.49 | 0.58 | 0.61 | 0.4 | 0.43 | 0.351 | -0.51 | 0.59 | 0.573 | 0.41 | 0.6 | 0.466 | 0.197 | 0.55 |
| | | P, T | 1 | 0.848 | 0.356 | 0.613 | 1.09 | 0.28 | 0.524 | 0.918 | 0.377 | 0.589 | 1.11 | 0.21 | 0.395 | 0.99 | 0.305 | 0.53 |
| | | | 6 | 0.353 | 0.491 | 0.771 | 0.661 | 0.458 | 0.584 | 0.296 | 0.406 | 0.747 | 0.507 | 0.240 | 0.576 | 0.454 | 0.398 | 0.669 |
| | | | 12 | 0.22 | 0.544 | 0.848 | 0.335 | 0.515 | 0.685 | 0.206 | 0.564 | 0.778 | 0.258 | 0.313 | 0.697 | 0.254 | 0.484 | 0.752 |
| D, P, T | | D, P, T | 1 | 0.895 | 0.283 | 0.536 | 1.2 | 0.28 | 0.659 | 0.790 | 0.497 | 0.725 | 0.975 | 0.241 | 0.549 | 0.965 | 0.325 | 0.617 |
| | | | 6 | 0.177 | 0.871 | 0.933 | 0.594 | 0.37 | 0.713 | 0.178 | 0.893 | 0.946 | 0.352 | 0.585 | 0.820 | 0.325 | 0.679 | 0.853 |
| | | | 12 | 0.087 | 0.866 | 0.931 | 0.152 | 0.760 | 0.889 | 0.088 | 0.919 | 0.959 | 0.187 | 0.640 | 0.825 | 0.128 | 0.796 | 0.901 |

Table 4 (continued)

| Station | Method | Sub-method | Time scale | W-MLPNN | | | W-ANFIS | | | W-SVR | | | W-EDT | | | Average of all models | | |
|-------------------------|--------------------------|------------|------------|---------|-------|-------|---------|-------|-------|-------|-------|-------|-------|-------|-------|-----------------------|-------|-------|
| | | | | RMSE | NSEI | CC | RMSE | NSEI | CC | RMSE | NSEI | CC | RMSE | NSEI | CC | RMSE | NSEI | CC |
| Average of all stations | First (single parameter) | D | 1 | 0.771 | 0.524 | 0.734 | 0.965 | 0.528 | 0.692 | 0.758 | 0.515 | 0.749 | 0.973 | 0.297 | 0.549 | 0.866 | 0.466 | 0.681 |
| | | | 6 | 0.156 | 0.909 | 0.955 | 0.374 | 0.451 | 0.779 | 0.16 | 0.902 | 0.95 | 0.312 | 0.619 | 0.818 | 0.25 | 0.72 | 0.875 |
| | | | 12 | 0.067 | 0.862 | 0.938 | 0.112 | 0.815 | 0.921 | 0.07 | 0.922 | 0.961 | 0.159 | 0.641 | 0.83 | 0.102 | 0.81 | 0.912 |
| | P | 1 | 1.01 | 0.276 | 0.481 | 1.08 | 0.171 | 0.264 | 0.928 | 0.271 | 0.57 | 1.06 | 0.185 | 0.307 | 1.01 | 0.225 | 0.405 | |
| | | | 6 | 0.437 | 0.477 | 0.661 | 0.669 | 0.31 | 0.397 | 0.382 | 0.402 | 0.661 | 0.481 | 0.297 | 0.505 | 0.492 | 0.371 | 0.556 |
| | | | 12 | 0.271 | 0.521 | 0.741 | 0.221 | 0.457 | 0.705 | 0.196 | 0.44 | 0.755 | 0.426 | 0.406 | 0.574 | 0.278 | 0.456 | 0.693 |
| Second(multiparameter) | T | 1 | 1.125 | 0.247 | 0.385 | 1.2 | 0.15 | 0.2 | 1.03 | 0.231 | 0.333 | 1.19 | 0.152 | 0.222 | 1.136 | 0.195 | 0.285 | |
| | | | 6 | 0.55 | 0.351 | 0.445 | 0.816 | 0.273 | 0.403 | 0.529 | 0.332 | 0.498 | 0.576 | 0.267 | 0.439 | 0.617 | 0.305 | 0.446 |
| | | | 12 | 0.338 | 0.412 | 0.538 | 0.528 | 0.356 | 0.484 | 0.341 | 0.07 | 0.545 | 0.514 | 0.378 | 0.546 | 0.430 | 0.304 | 0.528 |
| | D, P, T | 1 | 0.890 | 0.299 | 0.527 | 1.07 | 0.229 | 0.381 | 0.881 | 0.338 | 0.594 | 1.02 | 0.255 | 0.43 | 0.965 | 0.28 | 0.483 | |
| | | | 6 | 0.368 | 0.491 | 0.781 | 0.637 | 0.411 | 0.557 | 0.311 | 0.513 | 0.757 | 0.482 | 0.326 | 0.587 | 0.449 | 0.435 | 0.67 |
| | | | 12 | 0.223 | 0.528 | 0.818 | 0.345 | 0.450 | 0.689 | 0.182 | 0.591 | 0.803 | 0.248 | 0.424 | 0.668 | 0.249 | 0.498 | 0.744 |

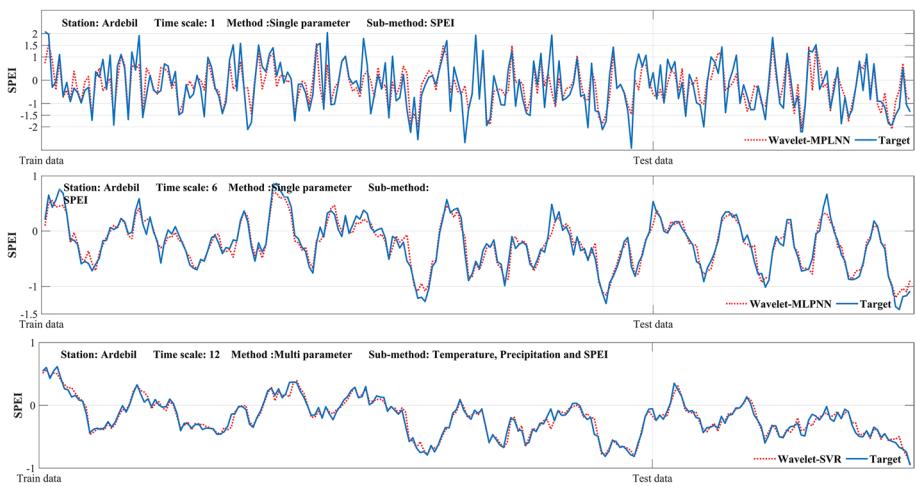


Fig. 8 Time series graphical comparison between the target and the best model at Ardebil station

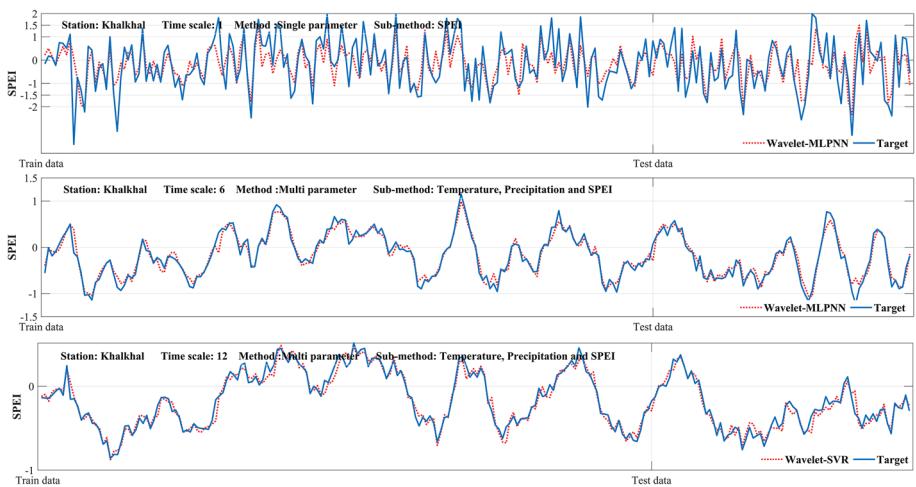


Fig. 9 Time series graphical comparison between the target and the best model at Khalkhal station

0.662 and RMSE, NSEI, and CC for the sub-method (D, P, T) were 0.926, 0.33, and 0.582, respectively, of which the sub-method (D) outperforms all. W-SVR and W-MLPNN now were more efficient than other models same as the Ardebil station in this time scale, of which the W-MLPNN in the sub-method (D) with RMSE=0.87, NSEI=0.552, and CC=0.749, was the best. In the 6-month time scale, based on the average results of the models, the sub-method (D) exhibited RMSE=0.244, NSEI=0.733, and CC=0.886 and the sub-method (D, P, T) exhibited RMSE=0.224, NSEI=0.768, and CC=0.896, which the sub-method (D, P, T) had the highest accuracy. Moreover, W-MLPNN in the sub-method (D, P, T) with RMSE=0.15, NSEI=0.923, and CC=0.967 showed the best result. Eventually, based on the average results of all the models in the last time scale, RMSE=0.08, NSEI=0.826, and CC=0.923 were for the sub-method (D) and RMSE=0.09, NSEI=0.777, and CC=0.905 were for the sub-method (D, P, T), of which the sub-method (D) slightly outperformed the

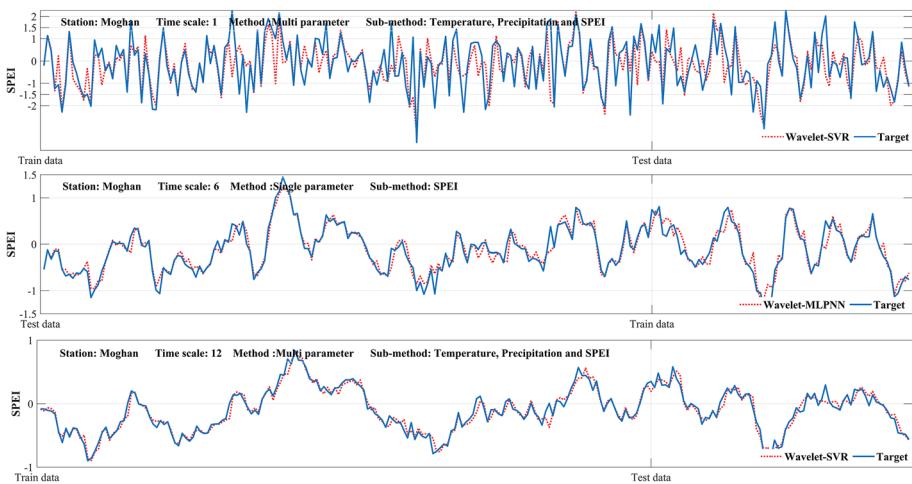


Fig. 10 Time series graphical comparison between the target and the best model at Moghan station

(D, P, T) sub-method. On the contrary, the best results in this time scale belonged to W-SVR in the sub-method (D, P, T) with RMSE = 0.053, NSEI = 0.929, and CC = 0.969. Figure 9 shows the graphical comparison between the time series of the target and the best model at Khalkhal station.

3.1.2.3 Moghan station Based on the results that are given in Table 4, the best results belonged to the sub-methods (D) and (D, P, T), of which based on the average results of all the models in the 1-month time scale, RMSE = 0.921, NSEI = 0.3, and CC = 0.642 and RMSE = 0.963, NSEI = 0.325, and CC = 0.617 were for the sub-method (D) and (D, P, T), respectively, of which the sub-method (D) slightly outperformed all. Same as other stations, W-SVR and W-MLPNN exhibited higher efficiency than ANFIS and EDT in this time scale, which W-SVR in the sub-method (D, P, T) with RMSE = 0.725, NSEI = 0.497, and CC = 0.790 showed better results. In the 6-month time scale, based on the average results of all the models, RMSE, NSEI, and CC for the sub-method (D) were 0.294, 0.651, and 0.844 and RMSE, NSEI, and CC for the sub-method (D, P, T) were 0.325, 0.679, and 0.8853, respectively, of which the sub-method (D, P, T) obtained better results. Furthermore, W-MLPNN in the sub-method (D) with RMSE = 0.147, NSEI = 0.919, and CC = 0.955 had the best result. In the 12-month time scale eventually, RMSE = 0.133, NSEI = 0.779, and CC = 0.903 were for the sub-method (D) and RMSE = 0.128, NSEI = 0.796, and CC = 0.901 were for the sub-method (D, P, T), of which the sub-method (D) slightly outperformed other sub-methods. Also, W-SVR in the sub-method (D, P, T) with RMSE = 0.088, NSEI = 0.919, and CC = 0.959 was the best model in the 12-month time scale. Figure 10 shows the graphical comparison between the time series of the target and the best model at Moghan station.

3.1.3 Overall assessment

Table 5 summarizes the results of the best sub-methods, the best models, and sub-methods of the best models at all stations in two parts. According to the upper part of the table, the sub-methods (D) and (D, P, T) had better efficiency in the (1-month and 6-month) and 12-month time scales, respectively. Also, the best models and their sub-methods were

Table 5 Summary of the best sub-methods, models, and sub-methods of the best models in all time scales

| Criteria | Time Scale | Synoptic stations | | Ardeabil | |
|--|------------|-------------------|------------|------------|------------|
| | | Moghān | Khalkhal | | |
| The best sub-method based on the average of models | 1 | D | D | D | D |
| | 6 | (D, P, T) | (D, P, T) | D | D |
| | 12 | (D, P, T) | D | D | D |
| The best model and its sub-method | 1 | Model | Sub-method | Sub-method | Sub-method |
| | 6 | W-SVR | (D, P, T) | W-MLPNN | D |
| | 12 | W-MLPNN | D | W-MLPNN | W-MLPNN |
| | 1 | W-SVR | (D, P, T) | W-SVR | (D, P, T) |
| | 6 | W-SVR | (D, P, T) | W-SVR | (D, P, T) |
| | 12 | W-SVR | (D, P, T) | W-SVR | (D, P, T) |

presented in the lower part of Table 5, of which W-MLPNN in the 1-month and 6-month time scales and W-SVR in the 12-month time scale acquired the best results. Moreover, based on their sub-methods, the sub-method (D) in the 1-month and 6-month time scales and sub-method (D, P, T) in the 12-month time scale stood superior. Hence, based on the abundance of the best sub-method in all sections of Table 5, the sub-method (D) with W-MLPNN in the 1-month and 6-month time scales and the sub-method (D, P, T) with W-SVR in the 12-month time scale were determined as the best models and sub-methods in this area.

3.2 Discussion

To achieve the best outcomes, all aspects of modeling that have substantial effects on the results including the type of models, time scales, architectures of the models, optimizers (e.g., preprocessors), climates and sub-climates, and variables must be considered, run, and analyzed. To this end, many studies have explored the best hybrid models with preprocessors for more accurate drought forecasting in different climates and time scales (Zhang et al. 2017; Liu et al. 2018). Although numerous studies have compared the efficiency of popular AI models (e.g., EDT, MLP, SVR, ANFIS, etc.) for drought prediction (Tan et al. 2022; Alawsy et al. 2022), the genuine competition is between MLP and SVR in most cases. The current study found that W-MLPNN and W-SVR are more capable than W-ANFIS and W-EDT for drought prediction. Shirmohammadi et al. (2013) concluded that ANFIS is better than MLP, and using WT in AI models significantly improves the accuracy of results. Soh et al. (2018) compared W-MLP and W-ANN and concluded that W-MLPNN is more capable than W-ANFIS in drought modeling. It should be noted that one of the reasons for the failure of ANFIS is fuzzy rules; in fact, data type and their fluctuations range can negatively affect the ANFIS flexibility. In addition, because the fuzzy rule-based is developed based on linguistic knowledge, the ANFIS training algorithm categorizes the data based on the specific adjustments, thus the model could be less applicable.

Using the lags of all precipitation, temperature, and drought in the combination of input data sets simultaneously and analyzing their efficiency in the models in 1, 6, and 12-month time scales for drought prediction by trial and error is one of the properties of this study. In this regard, the results indicated that the combination of precipitation, temperature, and drought all together in the input data sets positively affected the accuracy of the models in longer time scales (more than 6-month time scales). Jalalkamali et al. (2015) used precipitation, minimum temperature, maximum temperature, drought index, and their lags to develop four methods of setting input data sets for ANFIS, SVM, MLP, and ARMAX. Based on their findings, the method that used all parameters had better output almost in all cases. The current study determined the best lags as an input data set and found that increasing the time scale affects the abundance of longer lags in the input data sets. Studies such as Hosseini-Moghari et al. (2017) and Nguyen et al. (2017) concluded that short-term lags have a close correlation with target data. It should be noted that some studies conducted that just lags of drought have a positive effect on the results, but this study concluded that the combination of synoptic parameters (e.g., temperature and precipitation) with drought data as input data sets can improve the efficiency of models in longer time scales (more than 6-month). Moreover, this study attempted to determine the best combination of input data by EFET, which is a strength for enhancing the efficiency of the models.

Eventually, according to the best methods and models, increasing the time scale improves the accuracy of models almost in all cases. Hosseini-Moghari et al. (2017) in this

regard concluded that the prolongation in the time scale improves the efficiency of models and significant fluctuations (especially under 6-month time scales) can increase the regression model error. In another study, Nguyen et al. (2017) concluded that occurring recurrent dry and wet periods in the short-term time scales have a negative effect on models' flexibility. Also, Mishra and Desai (2005) maintained that the most important cause of decreasing the efficiency of models is white noise in short-term time scales.

4 Conclusion

This study aimed to predict drought in 1-, 6-, and 12-month time scales based on the two methods and five sub-methods of using four hybrid AI models. Based on the results, the abundance of small lags (under the 6th month) in the input data sets for the 1-month and 6-month time scales was more than others and vice versa. In addition, the abundance of the temperature lags in the sub-method (P,T) and drought in the sub-method (D, P, T) were more than other variables in the combination of input data sets in all sub-methods. Based on the abundance of the best sub-methods and sub-methods of the best models in all stations, the sub-method (D) in most 1- and 6-month time scales and sub-methods (D, P, T) in most 12-month time scales were more impressive. Also, the best results of the models belonged to MLPNN in the 1- and 6-month time scales and W-SVR in the 12-month time scale, respectively. It should be noted that the best results belonged to W-SVR at Khalkhal station in the 12-month time scale. Due to the potential and abilities of the W-MLP and W-SVR in the aforementioned time scales, it is possible to apply such models for the forecasting of drought indices which are based on synoptic parameters. Moreover, utilizing the lags of precipitation, temperature, and drought all together in the input data sets could improve the efficiency of models, so these points must be considered by hydrologists, water managers, and policymakers to deal with the consequences ahead of the drought and make a clear vision of this phenomena in future of this region.

Funding There was no funding for this study.

Declarations

Conflict of interest None.

References

- Abdourahmane ZS, Acar R (2019) Fuzzy rule-based forecast of meteorological drought in Western Niger. *Theor Appl Climatol* 135:157–168. <https://doi.org/10.1007/s00704-017-2365-5>
- Adamowski JF (2008) Development of a short-term river flood forecasting method for snowmelt driven floods based on wavelet and cross-wavelet analysis. *J Hydrol* 353:247–266. <https://doi.org/10.1016/j.jhydrol.2008.02.013>
- Alawsi MA, Zubaidi SL, Al-Bdairi NS, Al-Ansari N, Hashim K (2022) Drought forecasting: a review and assessment of the hybrid techniques and data pre-processing. *Hydrology* 9(7):115. <https://doi.org/10.3390/hydrology9070115>
- Bacanli UG, Firat M, Dikbas F (2009) Adaptive neuro-fuzzy inference system for drought forecasting. *Stoch Environ Res Risk Assess* 23:1143–1154. <https://doi.org/10.1007/s00477-008-0288-5>

- Bai Y, Chen Z, Xie J, Li C (2016) Daily reservoir inflow forecasting using multiscale deep feature learning with hybrid models. *J Hydrol* 532:193–206. <https://doi.org/10.1016/j.jhydrol.2015.11.011>
- Behrang Manesh M, Khosravi H, Heydari Alamdarloo E, Saadi Alekasir M, Gholami A, Singh VP (2019) Linkage of agricultural drought with meteorological drought in different climates of Iran. *Theor Appl Climatol* 138(1):1025–1033. <https://doi.org/10.1007/s00704-019-02878-w>
- Belayneh A, Adamowski J (2012) Standard precipitation index drought forecasting using neural networks, wavelet neural networks, and support vector regression. *Appl Comput Intell Soft Comput* 2012:1–13. <https://doi.org/10.1155/2012/794061>
- Belayneh A, Adamowski J, Khalil B, Ozga-Zielinski B (2014) Long-term SPI drought forecasting in the Awash River Basin in Ethiopia using wavelet neural networks and wavelet support vector regression models. *J Hydrol* 508:418–429. <https://doi.org/10.1016/j.jhydrol.2013.10.052>
- Borji M, Malekian A, Salajegheh A, Ghadimi M (2016) Multi-time-scale analysis of hydrological drought forecasting using support vector regression (SVR) and artificial neural networks (ANN). *Arab J Geosci* 9(19):1. <https://doi.org/10.1007/s12517-016-2750-x>
- Breiman L (1996) Bagging predictors. *Mach Learn* 24:123–140. <https://doi.org/10.1007/bf00058655>
- Brown M, Harris CJ (1994) Adaptive neurofuzzy systems for difficult modelling and control problems. IEE colloquium on advances in neural networks for control and systems. IEEE, Berlin, Germany, pp 15–21
- Cannas B, Fanni A, Sias G, Tronci S, Zedda MK (2005) River flow forecasting using neural networks and wavelet analysis. *Geophys Res Abstr* 7:08651
- Chen J, Li M, Wang W (2012) Statistical uncertainty estimation using random forests and its application to drought forecast. *Math Probl Eng* 1:2012. <https://doi.org/10.1155/2012/915053>
- Choubin B, Khalighi-Sigaroodi S, Malekian A, Ahmad S, Attarod P (2014) Drought forecasting in a semi-arid watershed using climate signals: a neuro-fuzzy modeling approach. *J Mt Sci* 11(6):1593–1605. <https://doi.org/10.1007/s11629-014-3020-6>
- Dehghani M, Saghafian B, Rivaz F, Khodadadi A (2017) Evaluation of dynamic regression and artificial neural networks models for real-time hydrological drought forecasting. *Arab J Geosci* 10(12):1–3. <https://doi.org/10.1007/s12517-017-2990-4>
- Djibo AG, Karambiri H, Seidou O, Sittichok K, Philippon N, Paturel JE, Moussa Saley H (2015) Linear and non-linear approaches for statistical seasonal rainfall forecast in the Sirba watershed region (SAHEL). *Climate* 3(3):727–752. <https://doi.org/10.3390/cli3030727>
- Durdur ÖF (2010) Application of linear stochastic models for drought forecasting in the Büyük Menderes River Basin, Western Turkey. *Stoch Environ Res Risk Assess* 24:1145–1162. <https://doi.org/10.1007/s00477-010-0366-3>
- Farajzadeh H, Matzarakis A (2009) Quantification of climate for tourism in the northwest of Iran. *Meteorol Appl* 16:545–555. <https://doi.org/10.1002/met.155>
- Fullér R (2000) Introduction to neuro-fuzzy systems. Springer, Cham
- Gunn SR (1998) Support vector machines for classification and regression. *ISIS Tech Report* 14(1):5–16
- Hosseini-Moghari SM, Araghinejad S, Azarnivand A (2017) Drought forecasting using data-driven methods and an evolutionary algorithm. *Model Earth Syst Environ* 3:1675–1689. <https://doi.org/10.1007/s40808-017-0385-x>
- Jahangir MH, Azimi SME, Arast M (2023) Determining the most appropriate probability distribution function for meteorological drought indices in Urmia Lake Basin. *Iran Environ Monit Assess* 195:2. <https://doi.org/10.1007/s10661-022-10639-y>
- Jalalkamali A, Moradi M, Moradi N (2015) Application of several artificial intelligence models and ARI-MAX model for forecasting drought using the standardized precipitation index. *Int J Environ Sci Technol* 12:1201–1210. <https://doi.org/10.1007/s13762-014-0717-6>
- Jang JS (1993) ANFIS: adaptive-network-based fuzzy inference system. *IEEE Trans Syst Man Cybern* 23(3):665–685
- Keshavarz M, Karami E, Vanclay F (2013) The social experience of drought in rural Iran. *Land Use Policy* 30:120–129. <https://doi.org/10.1016/j.landusepol.2012.03.003>
- Kim TW, Valdés JB (2003) Nonlinear model for drought forecasting based on a conjunction of wavelet transforms and neural networks. *J Hydrol Eng* 8:319–328. [https://doi.org/10.1061/\(ASCE\)1084-0699\(2003\)8:6\(319\)](https://doi.org/10.1061/(ASCE)1084-0699(2003)8:6(319))
- Kisi O, Cimen M (2011) A wavelet-support vector machine conjunction model for monthly streamflow forecasting. *J Hydrol* 399:132–140. <https://doi.org/10.1016/j.jhydrol.2010.12.041>
- Le JA, El-Akmary HM, Allali M, Struppa DC (2017) Application of recurrent neural networks for drought projections in California. *Atmos Res* 188:100–106. <https://doi.org/10.1016/j.atmosres.2017.01.002>
- Liu ZN, Li QF, Nguyen LB, Xu GH (2018) Comparing machine-learning models for drought forecasting in vietnam's cao river basin. *Polish J Environ Stud* 27:2633–2646

- Malik A, Kumar A, Singh RP (2019) Application of heuristic approaches for prediction of hydrological drought using multi-scalar streamflow drought index. *Water Resour Manag* 33(11):3985–4006. <https://doi.org/10.1007/s11269-019-02350-4>
- Mishra AK, Desai VR (2005) Drought forecasting using stochastic models. *Stoch Environ Res Risk Assess* 19:326–339. <https://doi.org/10.1007/s00477-005-0238-4>
- Mishra AK, Desai VR (2006) Drought forecasting using feed-forward recursive neural network. *Ecol Model* 198:127–138. <https://doi.org/10.1016/j.ecolmodel.2006.04.017>
- Mohammadi A, Tavakoli A, Ebrahimi A (2014) Predicting product life cycle using fuzzy neural network. *Manag Sci Lett* 4:2057–2064. <https://doi.org/10.5267/j.msl.2014.8.016>
- Mokhtarzad M, Eskandari F, Jamshidi Vanjani N, Arabasadi A (2017) Drought forecasting by ANN, ANFIS, and SVM and comparison of the models. *Environ Earth Sci* 76(21):1. <https://doi.org/10.1007/s12665-017-7064-0>
- Nguyen VH, Li QF, Nguyen LB (2017) Drought forecasting using ANFIS- a case study in drought prone area of Vietnam. *Paddy Water Environ* 15(3):605–616. <https://doi.org/10.1007/s10333-017-0579-x>
- Poornima S, Pushpalatha M (2019) Drought prediction based on SPI and SPEI with varying timescales using LSTM recurrent neural network. *Soft Comput* 23(18):8399–8412. <https://doi.org/10.1007/s00500-019-04120-1>
- Rhee J, Im J (2017) Meteorological drought forecasting for ungauged areas based on machine learning: using long-range climate forecast and remote sensing data. *Agric for Meteorol* 237:105–122. <https://doi.org/10.1016/j.agrformet.2017.02.011>
- Sain SR, Vapnik VN (1996) The nature of statistical learning theory. *Technometrics* 38:409. <https://doi.org/10.2307/1271324>
- Shirmohammadi B, Moradi H, Moosavi V, Semiroomi MT, Zeinali A (2013) Forecasting of meteorological drought using wavelet-ANFIS hybrid model for different time steps (case study: Southeastern part of east Azerbaijan province, Iran). *Nat Hazard* 69(1):389–402. <https://doi.org/10.1007/s11069-013-0716-9>
- Soh YW, Koo CH, Huang YF, Fung KF (2018) Application of artificial intelligence models for the prediction of standardized precipitation evapotranspiration index (SPEI) at Langat River Basin, Malaysia. *Comput Electron Agric* 144:164–173. <https://doi.org/10.1016/j.compag.2017.12.002>
- Sulaiman SO, Shiri J, Shiralizadeh H, Kisi O, Yaseen ZM (2018) Precipitation pattern modeling using cross-station perception: regional investigation. *Environ Earth Sci* 77(19):1–1. <https://doi.org/10.1007/s12665-018-7898-0>
- Tan YX, Ng JL, Huang YF (2022) A review on drought index forecasting and their modelling approaches. *Arch Comput Method Eng*. <https://doi.org/10.1007/s11831-022-09828-2>
- Thornthwaite CW (1948) An approach toward a rational classification of climate. *Geogr Rev* 38:55. <https://doi.org/10.2307/210739>
- Vicente-Serrano SM, Beguería S, López-Moreno JI (2010a) A multiscale drought index sensitive to global warming: the standardized precipitation evapotranspiration index. *J Clim* 23:1696–1718. <https://doi.org/10.1175/2009JCLI2909.1>
- Vicente-Serrano SM, Lasanta T, Gracia C (2010b) Aridification determines changes in forest growth in *Pinus halepensis* forests under semiarid mediterranean climate conditions. *Agric Meteorol* 150:614–628. <https://doi.org/10.1016/j.agrformet.2010.02.002>
- 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(10):2471–2481. <https://doi.org/10.1007/s00477-017-1437-5>
- Zhang R, Chen ZY, Xu LJ, Ou CQ (2019) Meteorological drought forecasting based on a statistical model with machine learning techniques in Shaanxi province, China. *Sci Total Environ* 665:338–346. <https://doi.org/10.1016/j.scitotenv.2019.01.431>

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

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.