Journal of Agricultural Sciences (Tarim Bilimleri Dergisi) 2024, 30 (1) : 61 – 78 DOI: 10.15832/ankutbd.1067486

Journal of Agricultural Sciences
(Tarim Bilimleri Dergisi)

J Agr Sci-Tarim Bili
e-ISSN: 2148-9297
jas.ankara.edu.tr

JAS

Meteorological Drought Assessment and Prediction in Association with Combination of Atmospheric Circulations and Meteorological Parameters via Rule Based Models

Fatemeh Shaker SUREH[a], **Mohammad TAGHI SATTARI**[b,c*], **Hashem ROSTAMZADEH**[d], **Ercan KAHYA**[e]

[a]*Water Resources Engineering, Faculty of Agriculture, Tarbiat Modares University, IRAN*

[b]*Department of Water Engineering, Faculty of Agriculture, University of Tabriz, IRAN*

[c]*Department of Agricultural Engineering, Faculty of Agriculture, Ankara University, TURKEY*

[d]*Department of Climatology, Geography and planning Faculty, University of Tabriz, Tabriz, IRAN*

[e]*Water Resources, Hydrology Civil Engineering Department, Istanbul Technical University, Istanbul, TURKEY*

ARTICLE INFO

Research Article

Corresponding Author: Mohammad TAGHI SATTARI, E-mail: mtsattar@gmail.com

Received: 03 February 2022 / Revised: 13 July 2023 / Accepted: 22 July 2023 / Online: 09 January 2024

Cite this article

Sureh F S, Sattari Taghi M, Rostamzadeh H, Kahya E (2024). Meteorological Drought Assessment and Prediction in Association with Combination of Atmospheric Circulations and Meteorological Parameters via Rule Based Models. Journal of Agricultural Sciences (Tarim Bilimleri Dergisi), 30(1): 61-78. DOI: 10.15832/ankutbd.1067486

ABSTRACT

The development of data-driven models in conjunction with the advances in technologies regarded as remote sensing in generating recorded data from satellites has guided water management studies towards using these technologies, especially in the regions dealing with drought, like the Lake Urmia basin, Iran. In this basin, the agricultural sector has been exposed to dryness due to a decrease in rainfall and uncontrolled water consumption. In the last decade, many studies have tried to brighten this arena of water knowledge. However, the relationship between meteorological variables and atmospheric circulation with the meteorological drought of Lake Urmia had never been determined. The relationship between meteorological variables and atmospheric circulation with Lake Urmia's meteorological drought has been determined. This study calculated Standardized Precipitation Evapotranspiration Index (SPEI) values based on meteorological variables. Then a combination of the meteorological variables and atmospheric circulation values was considered a data mining model input for estimating the droughts. The series of the SPEI values for 3-, 6-, 9-, 12-, 24-, and 48-month time scales were obtained during 1988-2016. In this study, both the M5 Tree model and Associate Rules were used to predict and analyze the meteorological drought at six synoptic stations in the basin, considering both the atmospheric circulations (North Atlantic Oscillation (NAO), Southern Oscillation Index (SOI), Mediterranean Oscillation Index of Gibraltar-Israel (Mogi), Mediterranean Oscillation Index of Algiers-Cairo (MOac), Western Mediterranean Oscillation Index (WEMO), Mediterranean, Red, Black, Caspian, and Persian Gulf SSTs) and the meteorological variables (lagged relative humidity, evapotranspiration, average temperature, minimum-maximum temperature, and pressure). The results showed that using a combination of the atmospheric circulation indices and meteorological variables in the models increases the model's accuracy and improves the results in a long-term period. The best result in the study of drought in the Lake Urmia basin is related to SPEI48 (R = 0.85, RMSE = 0.08, MAE = 0.11), and in the association rules, the value of the lifting index of the best rule is 1.32. Although both approaches provided acceptable results, the M5 Tree model had a comparative advantage due to simple and practical linear relationships.

Keywords: SPEI, Associate Rules, M5 Tree Model, The Urmia Lake basin, Iran

1. Introduction

Drought, as a natural disaster, is one of the main aspects of the climate in Iran. There are a number of research studies involving atmospheric circulations in drought prediction potentials. The teleconnections between El-Niño Southern Oscillation (ENSO) and critical flood and drought events have been reported all over the world (i.e., Barlow et al. 2001; Kahya & Karabörk 2001; Lee & Julien 2017; Kuzay et al. 2022). The most frequently used drought indexes are Palmer Drought Severity Index (PDSI) (Palmer 1965), Standardized Precipitation Index (SPI) (McKee et al. 1993) and Standardized Precipitation Evapotranspiration Index (SPEI) (Stagge et al. 2015; Wang et al. 2015).

In fact, various analysis methods have been proposed to monitor drought in two main categories: (i) statistical analysis (e.g., non-linear regression (Masocha et al. 2017), Geostatistical approach (Buttafuoco et al. 2018), log-linear (Moreira 2016) and Markov chain (Alam et al. 2017)) and (ii) data mining, which were developed to make a predictive decision based on the obtained knowledge. Data mining methods facilitate access to a large data set (Indurkhya & Weiss 1998), and allow the extraction of rules through pattern recognition technologies (e.g., neural networks, machine learning, and genetic programming).

Dracup & Kahya (1994) have examined the four regions for a relationship between streamflow and the cold (La Nifia) phase of the Southern Oscillation (SO). They pointed out a strong and significant connection between streamflow and La Niña in southwestern US. The results of this study show significant midlatitude streamflow responses to the tropical SO phenomenon but major limitation to this study is the limited length of records, which shows only nine El Nifio and La Nifla episodes occurring in the 41 years of streamflow data. Tadesse et al. (2004), who used data based approaches in drought analysis in the Nebraska for improve understanding of the characteristics and relationships of atmospheric and oceanic parameters that cause drought, examined a group of large-scale circulation the Results showed that the three indices have a stronger relationship with drought in the study area. The study suggests that data mining techniques can help us to monitor drought using oceanic indices as a precursor of drought but they did not investigate meteorological variables. Le et al. (2016) aims examining the lagged climate signals to predict SPEI at Khanhhoa province, using artificial neural network. They surveyed atmospheric circulations to predict SPEI using an artificial neural network method. Their results showed that adding more atmospheric circulations in the analysis could lead to better forecasting. The developed model can benefit developing long-term policies for reservoir and irrigation regulation and plant alternation schemes in the context of drought hazard. Nourani et al. (2017) proposed a hybrid application of Decision Tree and Associate Rule in SPI data in Tabriz and Kermanshah synoptic stations in Iran and de-trend SST data in the Black Sea, the Mediterranean Sea, and the Red Sea. They used classification, selection, and extraction to provide decision rules to monitor drought. Confidence and Heidke Skill Score (HSS) were used to evaluate the performance of the hybrid data mining methods, and thus, high Confidence was demonstrated between the monthly SPI values and the de-trend SST. This has resulted from the existence of a relative correlation between the three seas' de-trend SSTs and drought in Tabriz but they only used two station in Iran and can be used other ocean-atmospheric climate phenomena as predictors to predict maximum MP of station through the proposed method. The purpose of the study by Sezen & Partal (2017) is determining the effects of NAO and NCP on the temperature and precipitation regime of Mediterranean region in Turkey. As a result, NAO and NCP have remarkable influences on temperature and precipitation regime of Mediterranean region of Turkey for either seasonal or annual. It can be concluded that when the value of aforementioned teleconnection rises, the temperature value reduces. Nam et al. (2018) developed a satellite-based hybrid drought index called the vegetation drought response index for South Korea (VegDRI-SKorea) that could improve the spatial resolution of agricultural drought monitoring on a national scale. Their results showed that the hybrid drought index put forwarded spatially more improved drought patterns as opposed to that of station-based drought indices. Meteorological droughts threaten human societies in many parts of the world that are struggling with the water crisis.

Sattari et al. (2020) successfully predicted monthly precipitations at six meteorological stations in the Urmia Lake basin situated in north western Iran. They proved increased accuracy rate of models when using atmospheric circulation indices together with meteorological variables. Uzun & Ustaoglu (2022) examining current climatic conditions and atmospheric indices together on the Mediterranean crop yield. Hence, extreme weather conditions affect the yield of crops with high economic value in the Mediterranean Basin. Hosseini et al. (2022) determine the relationship between the NAO index and the precipitation of the APHRODITE base in Iran, to besides investigating the effect of these two phenomena in the country, the zones with positive, negative and no correlation to be identified with accuracy. The results in an analysis of North Atlantic Oscillation and monthly precipitation in Iran during the Last Half Century show that significant correlation is different between these two variables in time and space scale. In general, in more than 90% of Iran's area, there is no direct correlation between rainfall and NAO. In 9%, the correlation is negative and in only 1% of the Iranian zone, the correlation is positive.

In this study, we aimed to explore the linkage between atmospheric circulations and meteorological variables in the Urmia Lake basin for the purpose of drought assessment and prediction. The mergences of meteorological variables and atmospheric circulations using the Association Rules and M5 model has important implications in analysing drought prediction in the study area. This study has, therefore, special relevance to the understanding of drought patterns within the Urmia Lake basin. For this purpose, it used two M5 tree and Association Rules models, one of them evaluates the data numerically and the other after the discrete data examines the relationship variables. Both of them identify effective variables on the SPEI of the area.

2. Material and Methods

2.1. Study area

According to the Ramsar Convention, in 1971, Urmia Lake was declared an important international wetland. Urmia Lake, once the second largest saline lake in the world, is on the verge of complete desiccation. It has been suggested that the desiccation is caused by intensified human activities, and prolonged droughts in the lake basin. Finding the factors affecting the drought in this region is very important. In this study, we used from atmospheric circulations and meteorological parameters and data mining methods to investigate drought of Urmia lake basin. The closed catchment basin of Urmia Lake, which is considered as the main basin in the watershed division in Iran, has an area of 51 801 km^2 (Figure 1). Urmia Lake, situated in northwestern Iran, is one of the largest hypersaline lakes all over the world. Mediterranean air mass provides the main source of rainfall to the Urmia Lake. In our study area, there are six meteorological stations (Table 1) having an observation period spanning from January 1988 to December 2016; namely, Urmia, Tabriz, Sarab, Takab, Mahabad, and Maragheh. In this study, we used the meteorological station's local meteorological data in study area. The maximum statistical period was 28 (1988-2016) years and was available only for these six stations, which were provided to us by the Weather bureau. Only this number of stations were available in the area for the parameters of calculating evapotranspiration.

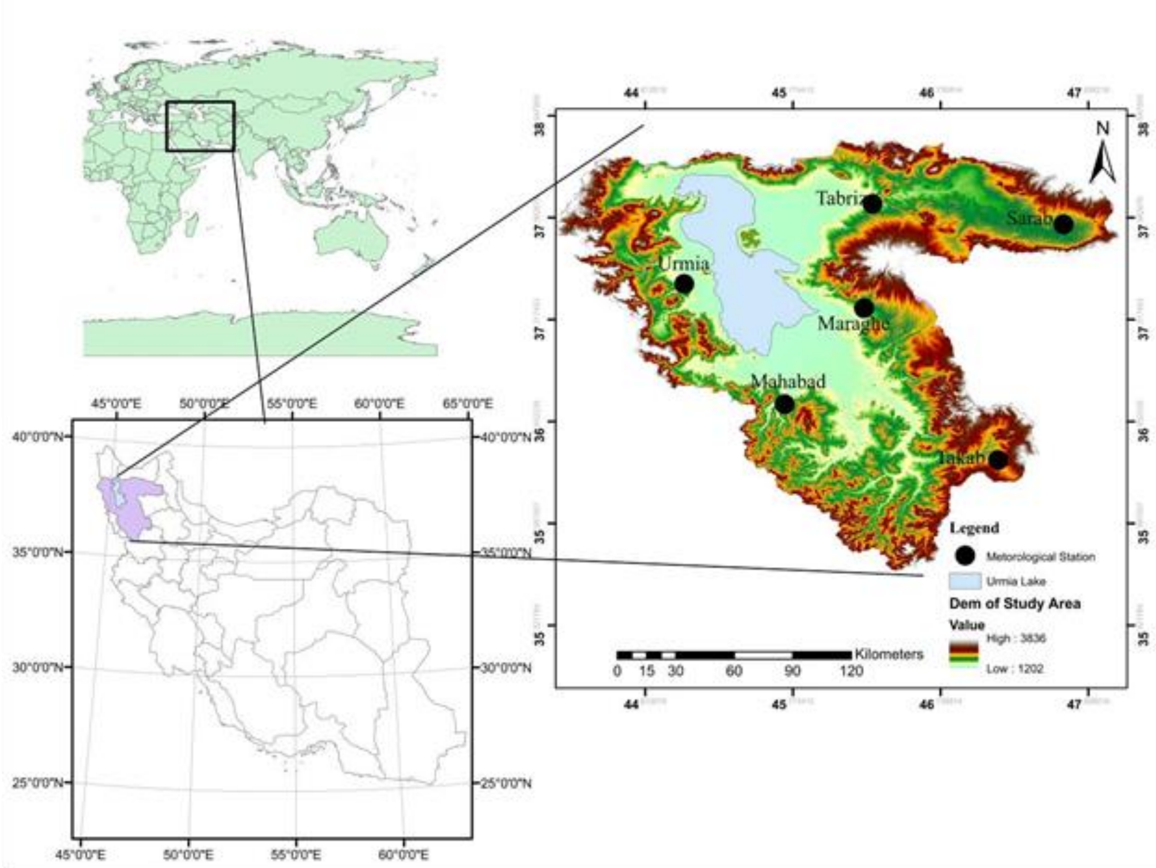


Figure 1- The geographic location of The Urmia Lake basin

Table 1- The statistical characteristics of the meteorological variables of six meteorological stations used in the Urmia Lake basin

Station		*Mahabad*	*Maragheh*	*Sarab*	*Tabriz*	*Takab*	*Urmia*
R (mm)	min	0.00	0.00	0.00	0.00	0.00	0.00
	max	155.92	114.81	102.24	114.84	165.72	147.54
	mean	33.19	23.67	20.29	20.57	26.85	25.54
	SD	34.06	25.89	19.10	19.95	27.77	27.48
ET (mm)	min	19.50	17.91	-30.00	21.35	19.51	16.21
	max	235.14	271.08	10.40	272.30	276.52	233.87
	mean	118.56	123.12	-5.67	126.05	117.15	109.77
	SD	61.93	71.61	9.97	72.25	69.55	62.536

In this research, the input variables in the models including meteorological variables and atmospheric circulation indices are described in Table 2.

Table 2- The Input variables used

Atmospheric circulations	*Meteorological variables*
Mediterranean sea surface temperature (MSST)	Minimum Temperature (Tmin)
Black Sea surface temperature (BSST)	Maximum Temperature (Tmax)
Red Sea surface temperature (RSST)	Rainfall (R)
Persian Gulf surface temperature (PSST)	Average Temperature (Tmean)
SOI	Average Relative Humidity (RH)
NAO	Average Wind speed (W)
MOac	Evapotranspiration (ET)
MOgi	------
WEMO	-------

In recent decades, drought has been affecting the Urmia Lake basin which can be seen in the average annual rainfall distribution (Figure 2). The Urmia Lake, which is the largest territorial lake in Iran and the second largest lake in the world, drying at an alarming rate. Therefore, various drought factors were studied to determine the causes of drought in the region in this study.

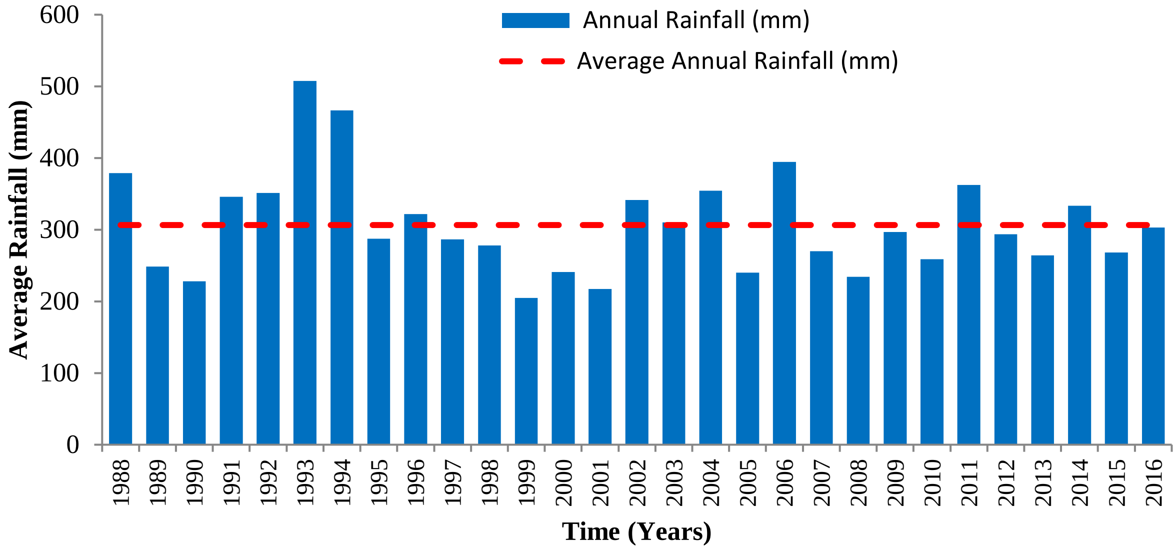


Figure 2- Annual and average rainfall distribution (mm) in The Urmia Lake basin

2.2. Meteorological drought index

Due to the vague and complex nature of droughts and the lack of identical definition of it, researchers often use drought indicators derived from hydrological variables to assess droughts and decide on drought management based on these indicators. So far, many indicators have been developed by researchers for drought monitoring. In this study, the SPEI was used, which is the multi-scale index comprising of precipitation and evapotranspiration variables, to analyse the danger of drought. Vicente-Serrano et al. (2010) proposed using the SPEI involving specifically the values of precipitation differences and evapotranspiration as expressed in equation 1, representing water balance value. In this equation, with a value for $ETo i$, the difference between the precipitation (P) and reference evapotranspiration (ETO) will be measured for i-th month.

$$Di = Pi - ETo i \tag{1}$$

Different methods have been proposed to calculate $ETo i$. Some studies have compared different ET calculation methods (Sheffield et al. 2012) and it has been shown that the Penman-Monteith method gives more accurate results because the formula is more based on the atmospheric evaporation demand. In 1998, in its publication No. 56, the FAO introduced the Penman-Monteith method as the standard method for estimating the evapotranspiration of the reference plant (Allen et al. 1998):

$$ET_0 = \frac{0.408\Delta(R_n - G) + \gamma\left[\frac{890}{T+273}\right]U_2(e_a - e_d)}{\Delta + \gamma(1 + 0.34U_2)} \tag{2}$$

ET0 Evapotranspiration - Reference plant (mm / day), Rn :Net radiation at the vegetation level (mega joules per square meter per day), T: average air temperature (°C),: U_2 wind speed at a height of 2 meters above the ground (meters per second), ea ed Lack of steam pressure at a height of 2 meters (kPa), Δ: Slope of vapour pressure curve (kPa), γ: The coefficient of psychometric (kilopascals per degree °C), G: is the heat flux into the soil (mega joules per square meter-day).

Since the study area is semi-arid, SPEI index was preferred for drought analysis. SPEI also shows the effect of climate change on drought by considering precipitation and evaporation together. The advantage of SPEI over other multi-scalar drought indices (i.e., SPI) highlights the role of temperature on drought through potential evapotranspiration. In this case, SPEI is a reliable tool to assess global warming influences on drought events. Many studies have been done using the SPEI index, among the most recent of which Mehdizadeh et al. (2020), Inoubli et al. (2020), Danandeh Mehr & Fathollahzadeh Attar. (2021)'s articles can be mentioned.

2.3. Selecting inputs

Vicente-Serrano et al. (2011) found that drought prediction might be improved considering the lag impacts of ENSO events. In addition, other large scale atmospheric circulations (Sattari et al. 2020), namely SOI, NAO, the Mediterranean Oscillation Index of Algiers-Cairo (MOac), the Mediterranean Oscillation Index of Gibraltar-Israel (MOgi), and finally the Western Mediterranean Oscillation Index (WEMO). The last three indexes are all based on standardized pressure differences between the two preselected locations in the Mediterranean Sea. Following Sattari et al. (2020), we here included the following average sea surface temperature data observed in (i) Mediterranean Sea, (ii) Red Sea, (iii) Caspian Sea, (iii) Black Sea, and (iv) Persian Gulf. The Pearson correlation method was adopted to determine the lag times of each input variable in relation to SPEI index and a probable predictor was chosen if it has the highest correlation with SPEI among various lag values. The combined effects of meteorological variables and large-scale atmospheric circulation indices on drought in the Urmia Lake basin are of primary interest in this research. Meteorological variables include lagged observations of SPEI index and Rainfall (R), Wind speed (W), mean temperature (Tmean), Evapotranspiration (ET), Average Relative Humidity (RH), minimum temperature (Tmin) and maximum temperature (Tmax).

2.4. Feature selection by RELIEF algorithm

A number of algorithms exists in the efforts of feature selection, along with processing speed problems. Due to the advent of fast computers and large storage capacities, a large data series of new issues have led to the continuing invention of the fast algorithm. The RELIEF method uses statistical solutions to select a feature as well as a weight-based method inspired by sample-based algorithms (Sattari et al. 2020). The RELIEF algorithm is known to be an effective tool for a dataset with a small number of training samples. Therefore, we preferred to adapt in identifying effective parameters in rainfall predictions due to the short span of our data.

2.5. Feature selection based on correlation (CFS)

Selection of feature based on correlation introduced by Hall (1999) is a commonly used method for selecting input variables and reducing problems in dimensions. The correlation method shows the subsets that have the properties with the highest correlation coefficient with the sample class, and the variables with the highest scores are considered as the main variables. This algorithm has a high ability to quickly detect irrelevant, additional, and error-free data, which generally results in the removal of half the data. This feature increases the production of the models by reducing the dimensions of the problem.

2.6. M5 tree predictions method

This model relies on a binary decision tree in which there are linear regression functions in the terminal node, revealing a relationship between independent and dependent variables. Tree-based models, which can be used for qualitative and quantitative data, resemble to a divide-and-conquer method in constructing a relationship between independent and dependent variables. The use of M5 tree model combined with neural networks gives more accurate result than decision trees like the CART method. The superiority of M5 tree model against regression trees is being smaller than regression trees. In comparison to neural networks, it results in easily understandable rules (Etemad-Shahidi & Mahjoobi 2009; Sattari et al. 2020). The advantage of M5 model tree over other previous linear models is that model trees are generally much smaller than regression trees and have proven more accurate in the tasks investigated. (Nourani et al. 2019). The M5 model tree can learn efficiently and can control tasks with very high dimensionality. This ability has developed the popularity of the M5 model tree and caused more usage in different fields of engineering. (Nourani et al. 2019). The fundamentals of the tree models are based on decision-making and problem solving. The structure of decision tree resembles to a tree whose body includes a system of leave, node, branch and root. Drawing the tree starts from top to bottom in which the root is placed as the first node on top and successively the chain of branches and nodes with the leaf. In the M5 tree model, the split criterion maximizes in reducing standard deviation in the child node (Sattari et al. 2020). In the case of reducing the standard deviation of the child node data impossible, the parent node do not branch out, instead reaching to final node or leaf. The first step in constructing a model tree is to calculate the standard deviation, which is expressed as in equation 3 (Quinlan. 1992).

$$\mathrm{SDR} = \mathrm{SD(T)} - \sum_{i=1}^{N} \frac{|T_i|}{T}\,\mathrm{SD(Ti)} \tag{3}$$

Where; SD(T) is the standard deviation of the N input data points expressed as in equation 4.

$$\mathrm{SD(T)} = \sqrt{\frac{1}{N}\left(\left(\sum_{i=1}^{N} Y_i^2\right) - \frac{1}{N}\left(\sum_{i=1}^{N} Y_i\right)^2\right)} \tag{4}$$

In these relations, T denotes a set of samples corresponding to each node as Ti does a subset of the subsample representing the sample. Our aim in this study is to evaluate the success of data mining methods which different from classical statistical methods. The M5 tree model offers practical, easier and more understandable solutions by generating linear relationships and if-then rules between target attribute and others. One of the major advantages of M5 model trees over normal regression trees is that normal regression trees can never predict the values outside the range of the trained model but the same is not the case with M5 model trees as they can extrapolate (Fayaz et al. 2022).

2.7. Association rules

Over the past years, among the methods of data mining, there has been a special interest in the algorithms of the discovery of patterns. As the name of these algorithms is known, we look for patterns existing in the dataset. In the meantime, the obtained algorithms of the repeated set of items have been discussed, which ultimately lead to the creation of Association Rules whose purpose is to find the number of abundances in the series or a database, in which the events preceding and occurring take place together. Support (S) is a fraction of transactions that contains all of the objects in a set of specific objects (Eq. 5).

$$\text{Support } (X \rightarrow Y) = P(X \cap Y) \tag{5}$$

Confidence (C) is the fraction of transactions containing all the objects of the conditional branch of the Association Rules that shows the accuracy of the rule. Unlike backup, an example for measuring the Confidence of a set of objects cannot be provided, because this criterion is only meaningful for associate rules (Eq. 6).

$$\text{Confidence } (X \rightarrow Y) = \frac{P(X \cap Y)}{P(X)} \tag{6}$$

Rules, which have a low limit of both Support and Confidence, are called strong rules of the Association Rules, and all algorithms are aimed to be associated with such rules.

LIFT is a criterion, which shows the degree of independence between the objects A and B, which can be a numerical value from zero to infinity. Combining this benchmark with Support and Confidence is one of the best ways to explore associate rules. The LIFT criterion is obtained in accordance with equation 7.

$$\text{Lift } (X \rightarrow Y) = \frac{P(X \cup Y)}{P(X)P(Y)} \tag{7}$$

All analytical steps in this study were processed using Weka, which is an open source tool collection of machine learning algorithms. Figure 3. illustrates analytical steps in our analysis.

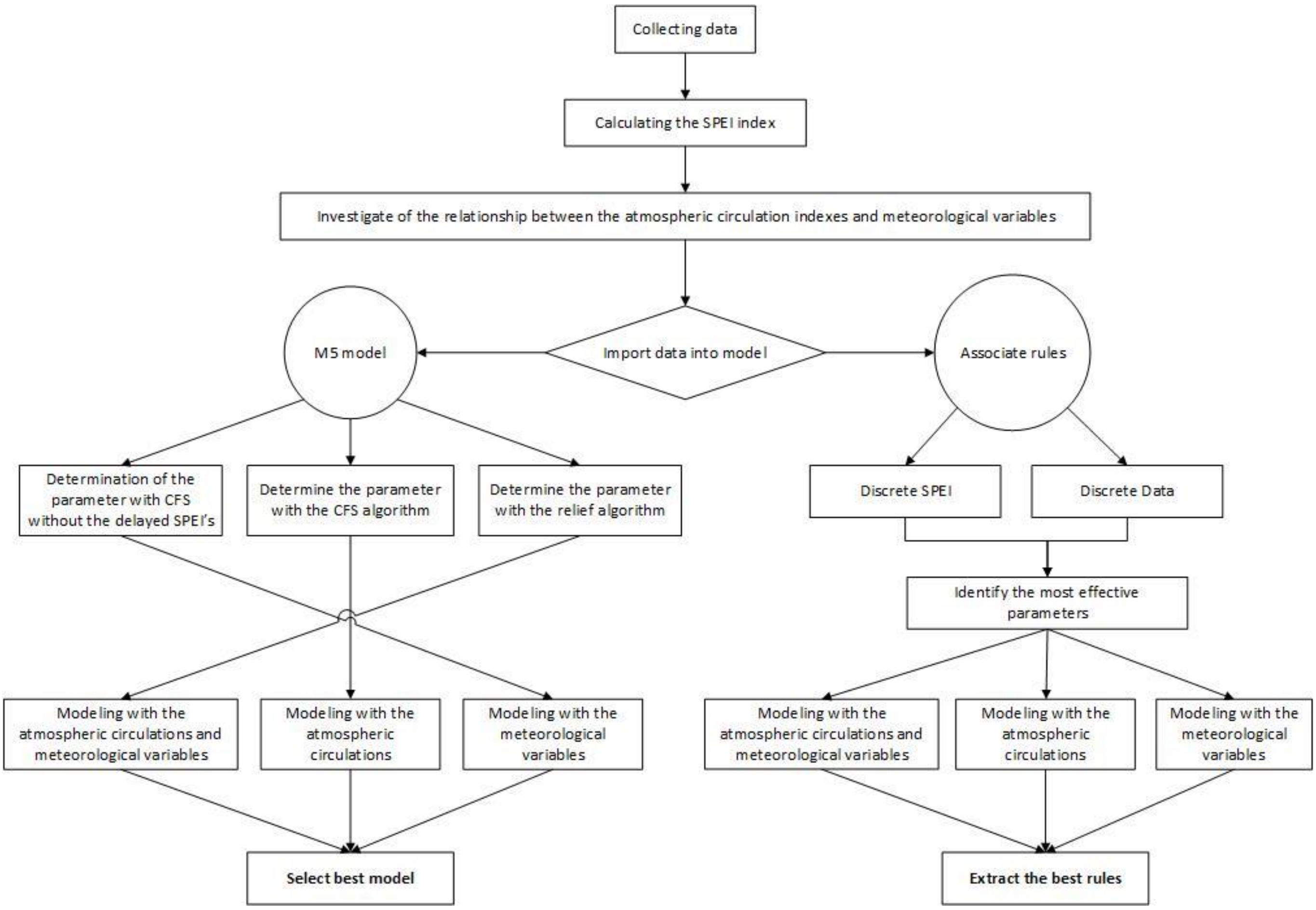


Figure 3- Analytical steps in drought predictions and forecasting

2.8. Model evaluation criteria

This criteria measures error rate and determines patterns and structures, exhibiting the least rainfall forecasting error. In this study, we used three model evaluation criteria, namely (i) correlation coefficient (R) (eq. 8), (ii) root mean square error (RMSE) (eq. 9) and (iii) mean absolute error (MAE) (eq. 10).

$$R = \frac{\sum_{i=1}^{n}(x_i-\bar{x})(y_i-\bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i-\bar{x})^2 \sum_{i=1}^{n}(y_i-\bar{y})^2}} \tag{8}$$

$$RMSE = \sqrt{\frac{\sum_{i=0}^{n}(y_i-x_i)^2}{N}} \tag{9}$$

$$MAE = \frac{1}{n}\sum_{I=0}^{N}|X_i - y_i| \tag{10}$$

In these relations, xi (yi) are observational (computational) values. The higher (lower) the correlation coefficient (RMSE and MAE) values are the more accurate the model is. In addition, we used Taylor diagram, which is useful in evaluating complex models to study geophysical phenomena, as a model evaluation tool in this study. It provides a concise statistical summary of how well patterns match each other in terms of correlation, RMSE, and the ratio of variances. Readers are referred to Taylor (2001) for further information about theoretical basis for the diagram.

3. Results and Discussion

In this study, an index, which includes both the information of precipitation and evaporation was used to study the phenomenon of meteorological droughts in the Urmia Lake basin. The SPEI index is a preferable tool to analyse droughts in warm and relatively dry regions such as Iran due to evaporation. For this purpose, predictions procedures using new data mining methods such as a decision tree and Association Rules were carried out. We examined droughts in the Urmia Lake basin using the SPEI index, which was calculated in the following sex steps: 3-, 6-, 9-, 12-, 24-, and 48-month (Figure 4). It is evident that drought events dominated after 1999, but much more after 2008 throughout all of the time scales. As the time scale gets higher, the spells of drought tend to unify in terms of the length of time but decrease in magnitude.

We obtained correlation coefficients between the SPEI and large-scale atmospheric circulation indices and climate variables at different points in time. Numerical values denote simultaneous and lagged correlations from 1- to 6-month at the 5% of significance level. In predicting, SPEI3 exhibits significant correlations with SOI, NAO, WeMO, MOgi, and MOac in the short term time scale. The SOI and WeMO have significant correlations with all the SPEI timescales, with the highest SOI correlation for SPEI9 with 5 steps of delay (R = -0.440) and the highest WeMO correlation for SPEI6 without a delay step (R = 0.239). Among the climate variables, the maximum temperature, mean humidity, mean rainfall and mean wind speed appeared to be significantly linked to the SPEI in short time lags. The remaining variables did not show a meaningful relationship with SPEI. Sattari et al. (2016) pointed out that drought index prediction works only in the short-term (i.e., time-scale 1) and the model accuracy decreases as the prediction time increases. Therefore, the SPEI index was predicted only for a single point in time in this research. In forecasting, SPEI3 with SOI, NAO, and WeMO, there is a significant correlation in the short-term scale. SOI and WeMO have a significant correlation with all SPEI time steps, with the highest correlation between SOI in SPEI9 with 4 steps of delay (R = -0.440) and the highest WeMO correlation in SPEI9 without a delay step (R = 0.199). Among the meteorological variables in the short-term scale, the maximum Temperature and average Relative Humidity and mean Rainfall and mean Wind speed are significantly correlated with 5% significance level.

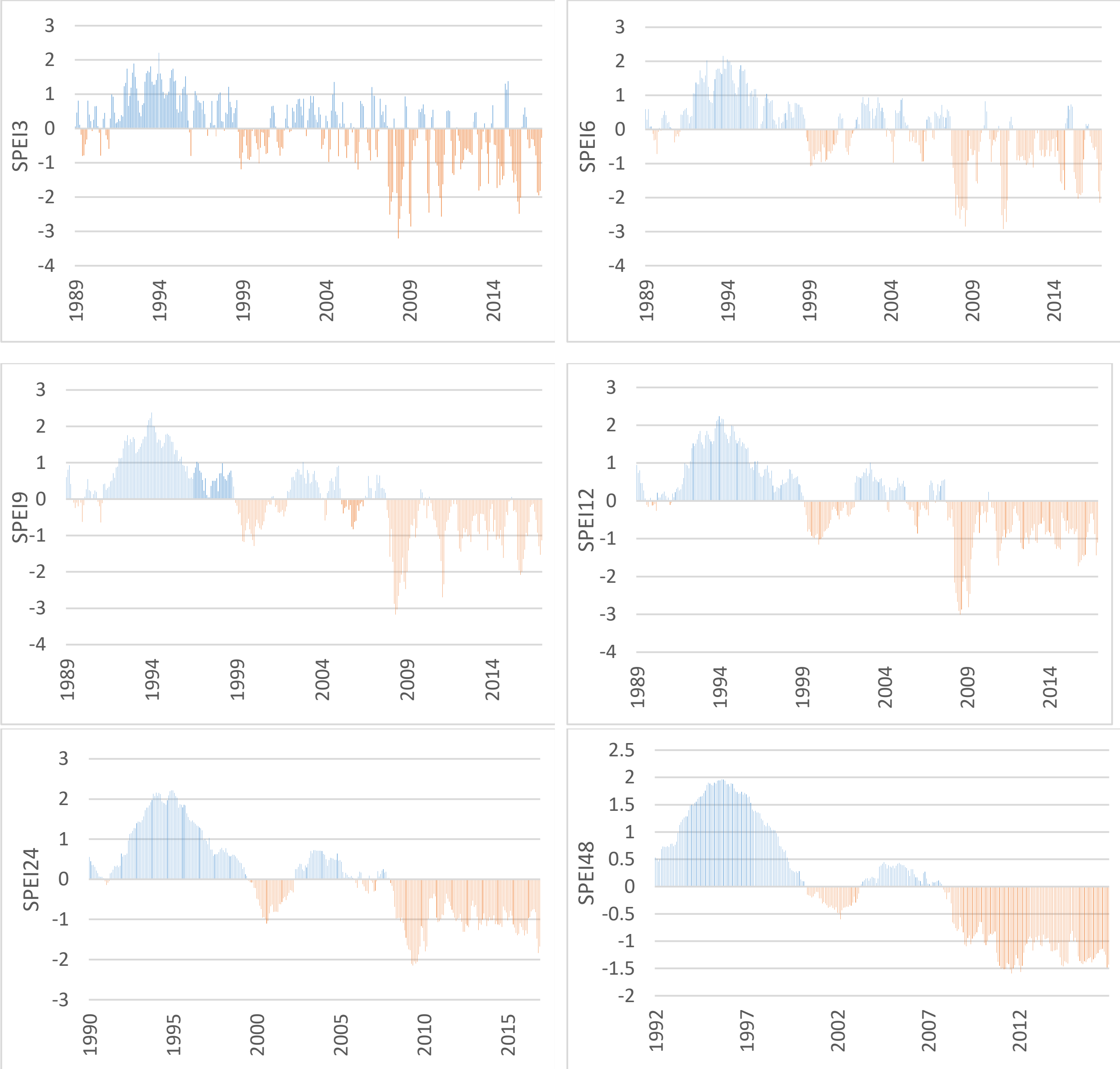


Figure 4- The SPEI index time series at time scales: 3, 6, 9, 12, 24, and 48 months in the Urmia Lake basin

In this research, after identifying the optimal combination by the RELIEF and correlation test, three input structures were determined to form an M5 tree model of possible combinations. In the first structure, the meteorological variables were used in addition to the SPEI with specific delays. In the second structure, atmospheric circulations were adopted in the model with the SPEI delays. In the third structure, a combination of these two groups of factors with the SPEI delays were considered.

In addition, each structure was given as three combinations in the model. In the first combination, the optimal combination of RELIEF tests with the SPEI delays were considered as the combination of the model. In the second combination, the correlation test for input parameters was used with considering the SPEI delays. In the third combination, the correlation test for input parameters was used without considering the SPEI delays. The optimal final composition of the three structures and three compounds was obtained according to these structures and combinations. The results of optimal combinations and structures are presented in Table 3. The optimal structures were entered into the M5 tree model. In Table 4, the results of the evaluation of the M5 tree model for assessing meteorological drought at the time scale of 3, 6, 9, 12, 24 and 48 months are presented in three structures with optimal combinations. In most situations, the third structure appears to be the best model performance when the input set is a combination of large-scale atmospheric circulation indices and climate variables. In general, the climate variables are more effective in the short-term analysis, whereas the atmospheric circulation indices are more effective in the long-term analysis (Miley et al. 2020; Dehghani et al. 2020).

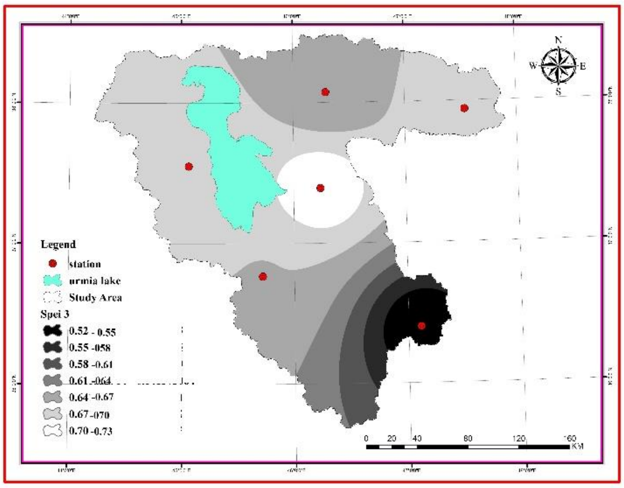
Table 3- Best input structure for M5 tree model for drought predictions and forecasting

		Structure	*Input*
SPEI3	Predicting	Combinatorial	SPEI3t-1 , SOIt-2 , MOgit , Wt-1, Rt-2
	Forecasting	Combinatorial	SPEI3t-3, SPEI3t-2, SPEI3t-1, RHt, MOgit, Rt-1,ETt, Wt
SPEI6	Predicting	Meteorological variables	Tmaxt, Tmint, Tmeant, SPEI6t-1, SPEI6t-2, SPEI6t-3
	Forecasting	Atmospheric circulations	SPEI6t-3, SPEI6t-2, SPEI6t-1,WEMOt, SOIt-4, MOgit
SPEI9	Predicting	Meteorological variables	ETt, Tmeant, Tmint, SPEI9t-1, SPEI9t-2, SPEI9t-3
	Forecasting	Meteorological variables	RHt,Wt, SPEI9t-3, SPEI9t-2, SPEI9t-1
SPEI12	Predicting	Atmospheric circulations	WEMOt-4, MOact, MOgit , SPEI12t-1, SPEI12t-2, SPEI12t-3
	Forecasting	Combinatorial	Wt , SOIt-4, SPEI12t-1, SPEI12t-2, SPEI12t-3
SPEI24	Predicting	Combinatorial	SOIt-6,Wt-6, Rt-1, MOgi, MOac, Tmax, SPEI24t-1, SPEI24t-2, SPEI24t-3
	Forecasting	Meteorological variables	Wt , RHt , SPEI24t-1, SPEI24t-2, SPEI24t-3
SPEI48	Predicting	Atmospheric circulations	MOact , WEMOt , SPEI48 t-1, SPEI48t-2, SPEI48t-3
	Forecasting	Combinatorial	W t-6, Rt, RHt, Tmaxt, BSSTt , MOact , WEMO t-6, SPEI48 t-1, SPEI48t-2, SPEI48t-3

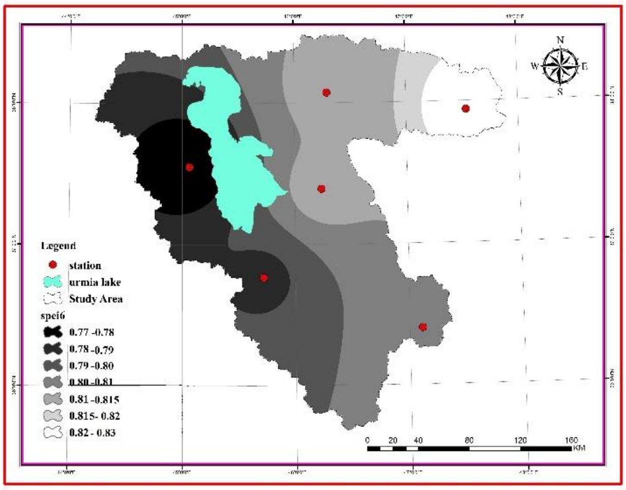
Except the SPEI3 predictions (RELIEF + SPEI delays), in the rest, the second combination (SPEI delays+cfs) was best input structure.

Table 4- Best results in the evaluation criteria for the M5 tree model for drought investigation and forecasting

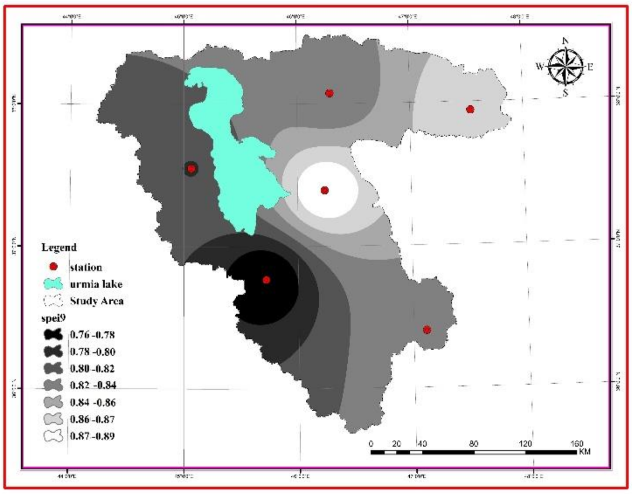
		SPEI3	*SPEI6*	*SPEI9*	*SPEI12*	*SPEI24*	*SPEI48*
Predicting	Structure	Combinatorial	Meteorological variables	Meteorological variables	Atmospheric circulations	Combinatorial	Atmospheric circulations
	R	0.6745	0.7866	0.7953	0.7574	0.8084	0.85
	RMSE	0.6901	0.3485	0.3777	0.2931	0.1926	0.0843
	MAE	0.5421	0.5221	0.2809	0.2153	0.1427	0.118
Forecasting	Structure	Combinatorial	Atmospheric circulations	Meteorological variables	Combinatorial	Meteorological variables	Combinatorial
	R	0.7086	0.7856	0.7917	0.7561	0.8049	0.8538
	RMSE	0.6654	0.528	0.2809	0.2144	0.1985	0.1168
	MAE	0.4921	0.3667	0.3794	0.2896	0.1503	0.0859



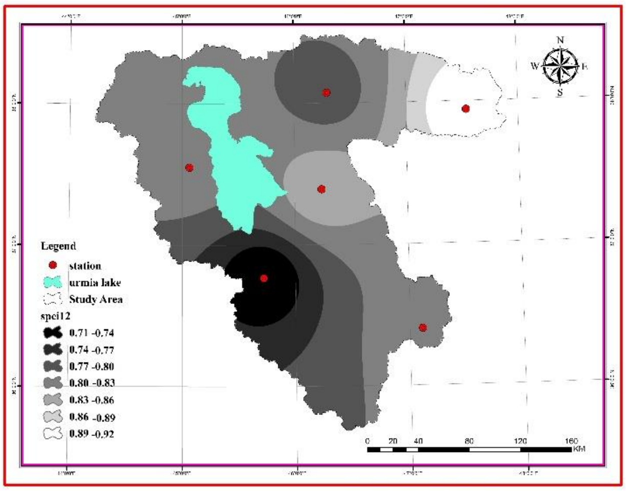
a) SPEI3



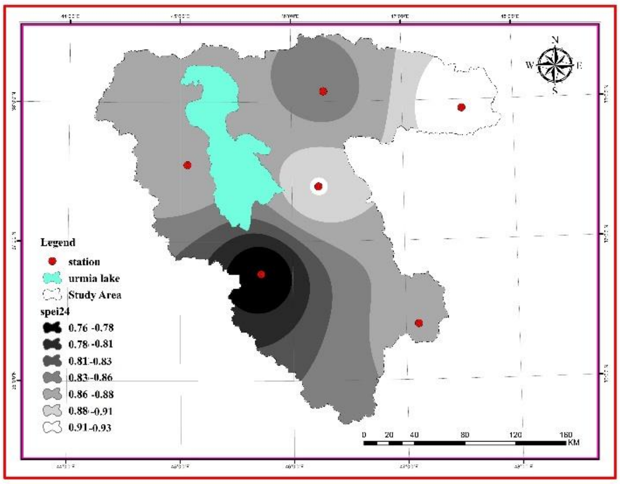
b) SPEI6



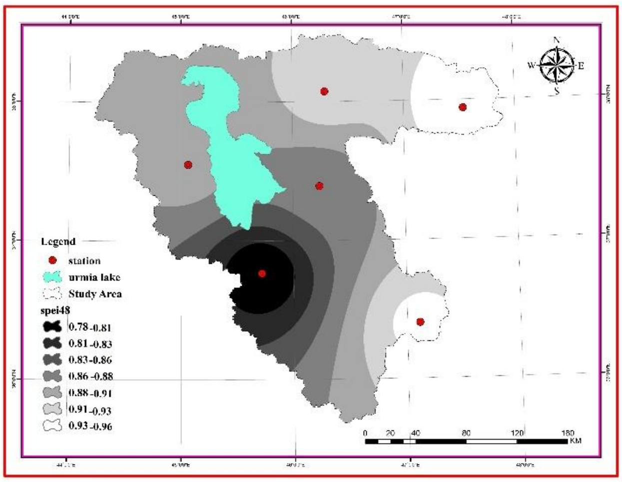
c) SPEI9



d) SPEI12



e) SPEI24



f) SPEI48

Figure 5- The correlation coefficients of M5 model for six stations between the years 1988–2016

The correlation coefficients are one of the criteria found in the evaluation in the M5 model and were calculated for six stations between the years 1988–2016. Figure 5 shows the results derived from the SPEI predictions for each station. If the model input is not the same for each time step of SPEI, comparisons cannot be made. The lighter colours in Figure 5 show higher correlation coefficients whereas the darker colours indicate lower correlation coefficients. Our results demonstrated that better information in most cases could be obtained from the combination of atmospheric circulations and meteorological variables. Moreover, the distribution charts display the best model for each SPEI time scale (Figures 6 and 7). As the time scale approaches the highest value (48-month), the accuracy of the resulting model increases. In other words, the SPEI48 model outperformed over the remaining models.

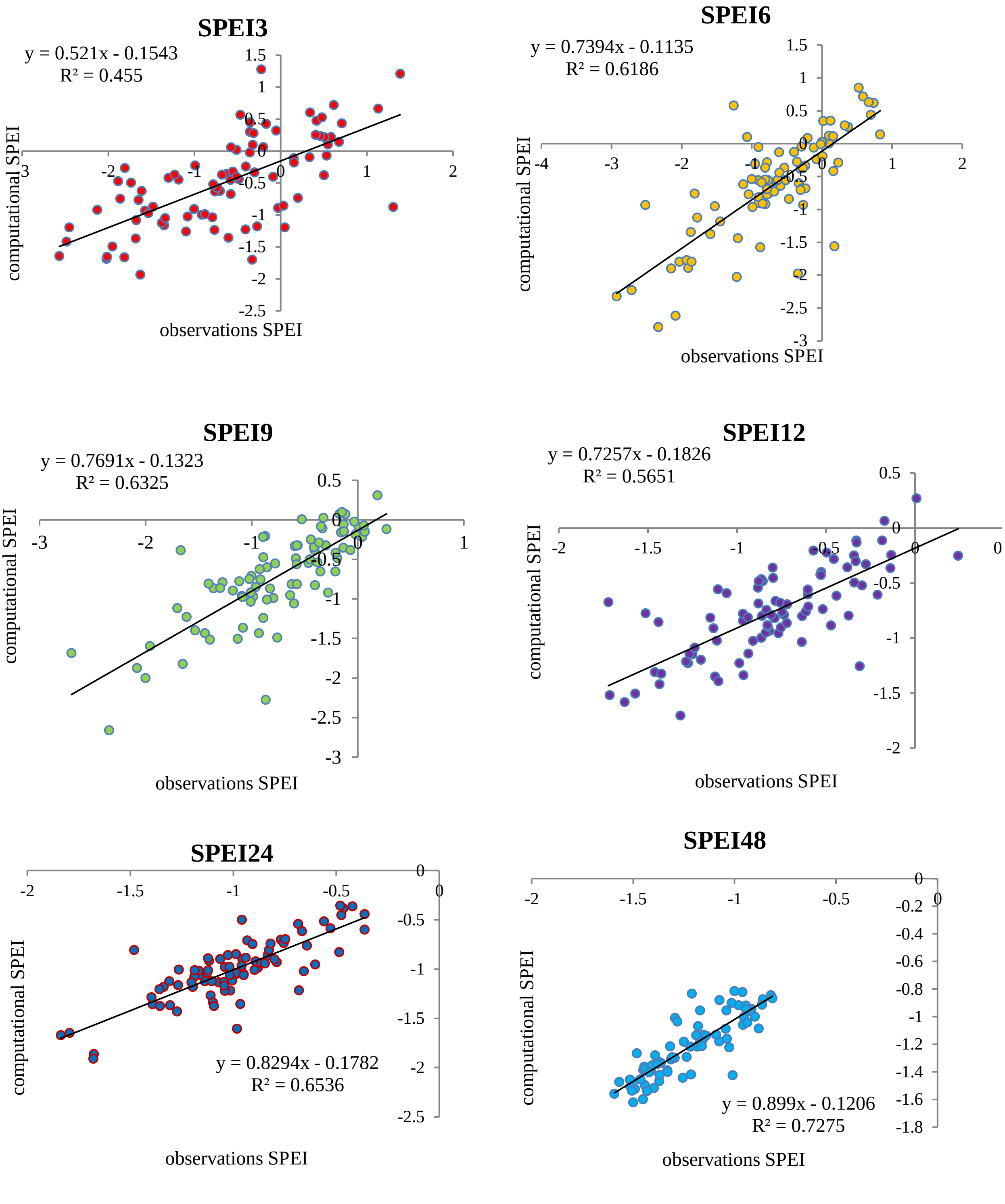


Figure 6- Distribution diagrams of actual and predicted SPEI for predictions of drought

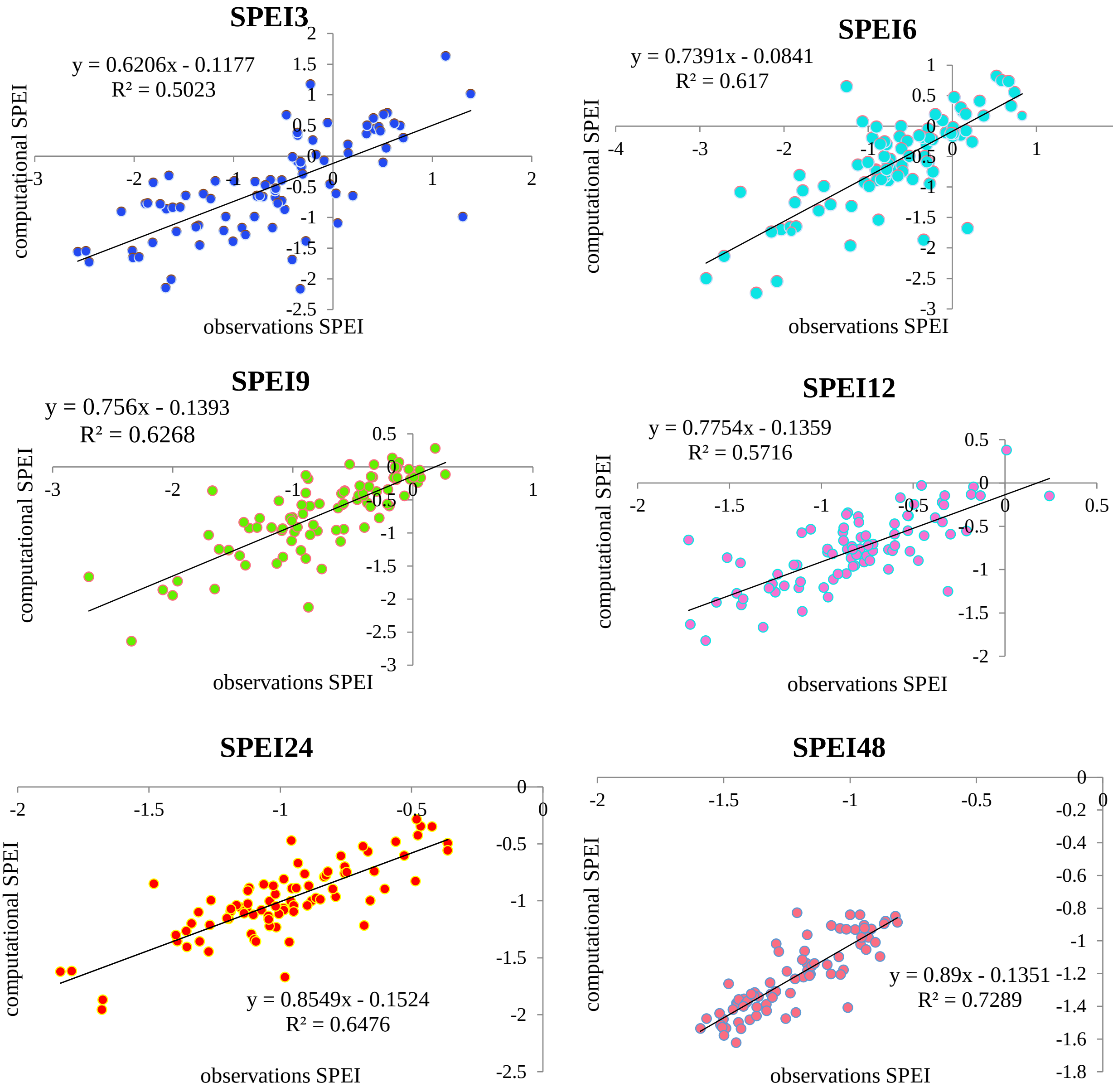


Figure 7- Distribution diagrams of actual and predicted SPEI for prediction of drought

Taylor diagrams are often applied to measure the predictions performance in hydrology and climatology. Information was plotted on Taylor's charts in identifying the best predictions and forecasting results for SPEI48 (Figure 8).

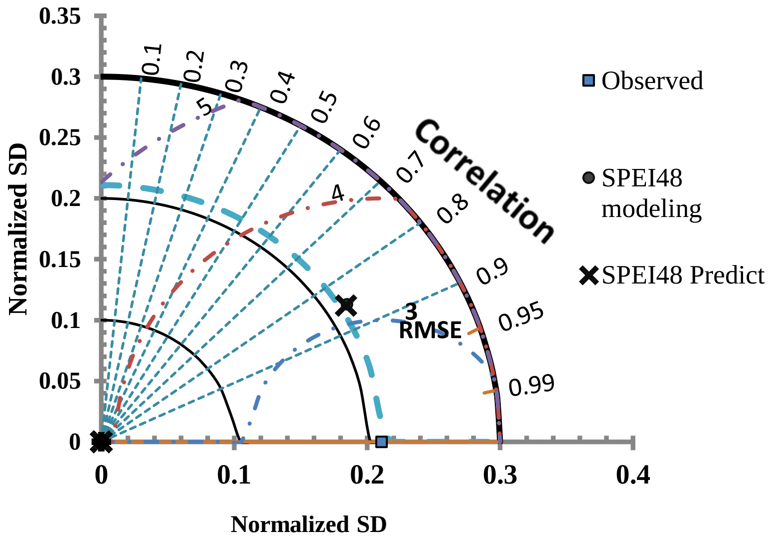


Figure 8- Taylor diagrams for SPEI48 predictions and prediction

Following to definition of drought index classification groups by the National Drought Reduction Center (Hayes 2003) we classified the monthly values of large-scale atmospheric circulation indices, climate variables and the drought index into seven discrete categories; namely, extremely dry (ED), severely dry (SD), moderately dry (MD), near normal (NN), moderately wet (MD), extremely wet (EW), and severely wet (SW). The classification threshold for drought indicators is shown in Figure 9. There is a lot of similarity between the SPI and SPEI indices. That is why many studies use the classification illustrated in Figure 9.

Drought Category	Extremely dry (ED)	Severely dry (SD)	Moderately dry (MD)	Near normal (NN)	Moderately wet (MW)	Severely wet (SW)	Extremely wet (EW)
SPEI	≤-2	-1.5≤SPEI≤-1.99	-1≤SPEI≤-1.49	0.99≤SPEI≤-0.99	1.49≤SPEI≤1	1.99≤SPEI≤1.5	≥2

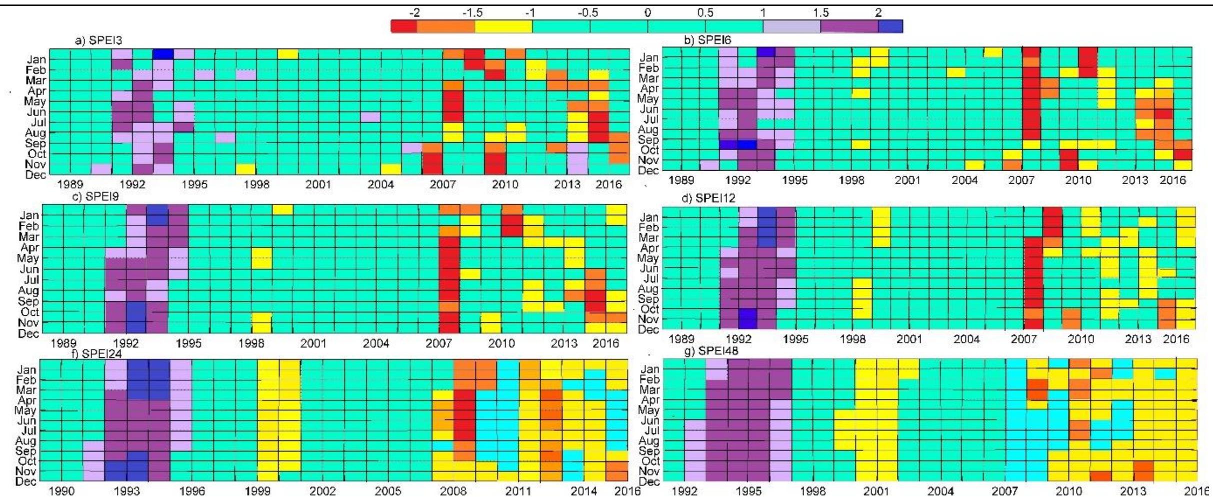


Figure 9- Drought index results in The Urmia Lake basin

Atmospheric circulation indices were further divided into seven categories with thresholds based on their frequency distribution. The data split is also shown in Figure 9, which presents the drought in recent years, from 2007 to 2016. It also indicates wetness from 1992 to 1997. This was ruled out in the case study. Assuming a normal distribution fit in to the period of 30-year observations, each circulation index and meteorological variable are divided with respect to magnitudes of 0.5, 1, and 1.5 times the standard deviation (SD). Table 5 shows the division of the data based on SD for the classification of each variable. For example, the SOI data were divided into seven groups as follows: SOI1 (group 1) corresponds to the SOI values more than 1.5 times its SD; SOI2 (group 2) for the range of 1 and 1.5 times its SD; SOI3 (group 3) for the range of 0.5 and 1 times its SD;

SOI4 (group 4) with internal values is 0.5 and -0.5 SD; SOI5 (group 5) for the range of -0.5 and -1 times its SD; SOI6 (group 6) for the range of -1 and -1.5 times its SD; and finally SOI7 (group 7) for the range of less than -1.5 times its SD. The similar rules apply for the classification of all meteorological variables and atmospheric circulation indices and this method was used by Tadesse et al. (2004).

Table 5- Classification of meteorological variables and atmospheric circulation indices

Variables	*Extremely high (EH)*	*Very high (VH)*	*High (H)*	*Near Normal (NN)*	*Low (L)*	*Very low (VH)*	*Extremely low (EL)*
Meteorological variables and sea surface temperature	1.5≤	1.5≤X≤1	1≤X≤0.5	0.5≤X≤-0.5	-1≤X≤-0.5	-1≤X≤-1.5	-1.5≥
NAO	3≤	3≤X≤2	2≤X≤1	1≤X≤-1	-1≤X≤-2	-2≤X≤-3	-3≥
SOI	2≤-	2≤-X≤1-	1≤-X≤0.5-	0.5≤X≤-0.5	1≤X≤0.5	1≤X≤2	2≥
MOac, MOgi	0.6≤	0.6≤X≤0.4	0.4≤X≤0.2	0.2≤X≤-0.2	-0.4≤X≤-0.2	-0.6≤X≤-0.4	-0.6≥

After determining these rules, keeping the definition of three different scenarios in mind, the Association Rules between the SPEI and large-scale atmospheric circulation indices and climate variables were extracted in the Urmia Lake basin in the next step. At the first scenario, only the effects of meteorological variables on effective SPEI groups were examined to determine which meteorological variables will mostly affect the SPEI in the basin. In the second scenario for the effective SPEI, the influences of atmospheric circulations indexes were examined; and finally, the combined effects of atmospheric circulations indexes on all the SPEI timescales were analysed in the third scenario. The Association Rules algorithm was implemented for various SPEIs. In each case, an average of 100 community rules was obtained, including prefixes of one to five members of the indexes. The occurrence of drought could be verified based on the meteorological variables and atmospheric circulation indices. As stated earlier in the formation of the Association Rules, only sufficiently repeated rules are desired, that is to have the minimum amount of support (in this research 10%) and a high degree of certainty. In other words, the achieved rules, at least in 34 cases occurred during the research period. That is to say, at least 10 percent of support is based on trial and error, implying that an increase in the amount of support will reduce not only the number of rules produced but also the increase in the rules so that their analysis becomes difficult to achieve.

The selected patterns were then determined for the SPEI3, SPEI6, SPEI9, SPEI12, SPEI24, and SPEI48 time scales. An effective group for the SPEI was determined as NN by the Association Rules. Most influential variables in the predictions of SPEI timescales are specified in Table 6, which shows that the most effective clustering of meteorological variables (Structure 1) for the SPEI3, SPEI6, SPEI9, SPEI12, SPEI24, SPEI48 are NN, VL, H, H, H, and VH, respectively. The most effective clustering of atmospheric circulations indices (Structure 2) is the H, NN, H, H, VL, and H, respectively. The most effective classification combination of the atmospheric circulation indices. And meteorological variables (Structure 3) on the SPEI time scales is NN, VH, VH, H, VL, and VH, respectively.

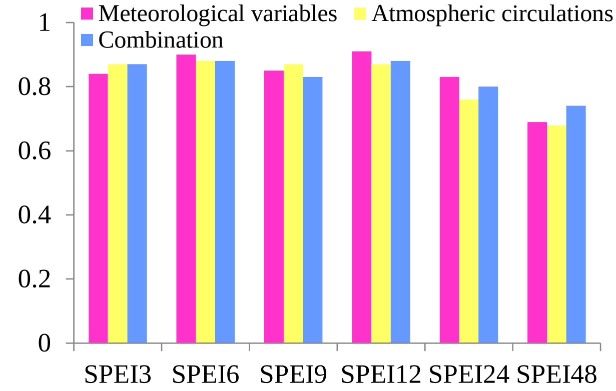
The most influential variables for the forecasting of SPEI timescales are specified in Table 6, which shows that the most effective classification of meteorological variables (Structure 1) for the SPEI3, SPEI6, SPEI9, SPEI12, SPEI24, SPEI48 are NN, VL, VH, VH, H, and VH, respectively. The most effective Classification of atmospheric circulation indices (Structure 2) is the (H, H, L, EL, VL, H), respectively. The most effective classification combination of atmospheric circulation indices and climate variables (Structure 3) on the SPEI time scales is NN, VH, VH, VL, NN, VH, respectively.

Table 6- Best rule of parameters with SPEI in Association Rules

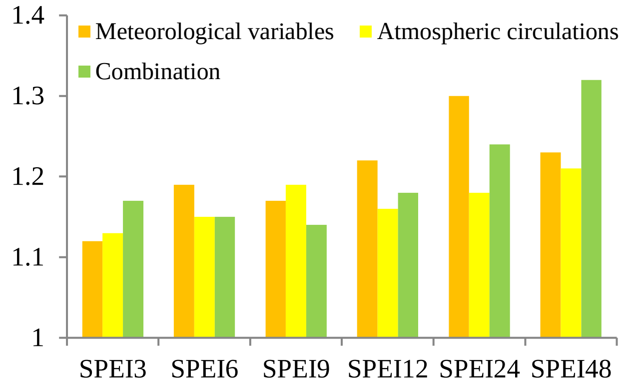
SPEI	*Structure*	*Combination*
SPEI3t	Meteorological variables	TMIN=NN, RH=NN ==> SPEI3=near normal
	Atmospheric circulations	MOgi=H ==> SPEI3=near normal
	Combinatorial	W=NN, RSST=VH ==> SPEI3=near normal
SPEI6t	Meteorological variables	TMAX= VL, ET= VL ==> SPEI6=near normal
	Atmospheric circulations	NAO=NN , SOI =NN ==> SPEI6=near normal
	Combinatorial	TMAX=VH, CSST=VH, MOac=EH ==> SPEI6=near normal
SPEI9t	Meteorological variables	W=H ==> SPEI9=near normal
	Atmospheric circulations	SOI= L ==> SPEI9=near normal
	Combinatorial	TMAX=VH ,MOac=EH 46 ==> SPEI9=near normal
SPEI12t	Meteorological variables	W=H ==> SPEI12=near normal
	Atmospheric circulations	SOI= L ==> SPEI12=near normal
	Combinatorial	TMAX=VH ,CSST=VH, MOac=EH 40 ==> SPEI12=near normal
SPEI24t	Meteorological variables	W=H ==> SPEI24=near normal
	Atmospheric circulations	RSST= VL, MSST= VL ==> SPEI24=near normal
	Combinatorial	R=NN, BSST= VL ==> SPEI24=near normal
SPEI48t	Meteorological variables	Tmean=VH, ET=VH==> SPEI48=near normal
	Atmospheric circulations	MOac=H==> SPEI48=near normal
	Combinatorial	Tmax=VH, PSST=VH,RSST=VH==> SPEI48=near normal
SPEI3t+1	Meteorological variables	Tmax=NN,RH=NN, ET=NN==> SPEI3=near normal
	Atmospheric circulations	MOgi=H==> SPEI3=near normal
	Combinatorial	RH=NN,MOgi=NN==> SPEI3=near normal
SPEI6t+1	Meteorological variables	Tmax = VL, ET= VL==> SPEI6=near normal
	Atmospheric circulations	MOgi=H==> SPEI6=near normal
	Combinatorial	Tmax =VH, CSST=VH,MOac=EH==> SPEI6=near normal
SPEI9t+1	Meteorological variables	Tmean=VH, ET=VH==> SPEI9=near normal
	Atmospheric circulations	SOI= L ==> SPEI9=near normal
	Combinatorial	ET=VH, MOac=EH==> SPEI9=near normal
SPEI12t+1	Meteorological variables	Tmax =VH, RH= VL==> SPEI12=near normal
	Atmospheric circulations	WEMO=EL==> SPEI12=near normal
	Combinatorial	RH= VL, CSST=VH==> SPEI12=near normal
SPEI24t+1	Meteorological variables	RH=H==> SPEI24=near normal
	Atmospheric circulations	BSST= VL==> SPEI24=near normal
	Combinatorial	R=NN, MSST= VL==> SPEI24=near normal
SPEI48t+1	Meteorological variables	Tmean=VH, ET=VH==> SPEI48=near normal
	Atmospheric circulations	SOI=H==> SPEI48=near normal
	Combinatorial	Tmax =VH, PSST=VH==> SPEI48=near normal

For each of the three structures, the values of confidence and Lift indexes computed with respect to the Association Rules are given in Figure 10. The confidence indicator for a fixed limit on the threshold is used to compare the SPEI time scale results. As evident in Figure. 10, with an increase in the SPEI time increment, the index of ascending Association Rules increases between the third scenario and SPEI, which indicates the high efficiency of the combined method in this study for the drought assessment in terms of SPEI. For example, the value of Lift index for Association Rules based on the SPEI48 and combination of large-scale atmospheric circulation indices and climate variables. Came out to be 1.32 and 1.24 for the SPEI24 and 1.1 for the SPEI12, indicating that the first combination is more relevant than other two alternatives.

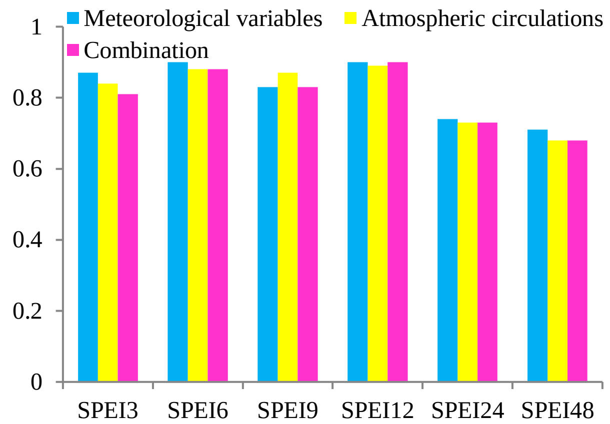
Atmospheric circulations can be also used solely as predictors without using meteorological variables – in longer horizons of SPEI (e.g.48-month). Due to Lake Urmia is close to the seas, it has been determined that the meteorological drought in this basin is more dependent on SOI and SST variables than other variables (Alizade Govarchin Ghale et al. 2018; Alizadeh-Choobari et al. 2016; Arkian et al. 2018). The Urmia Lake basin is one of the most important basins in Iran, facing many problems due to poor water management and rainfall reduction. Under current circumstances, it becomes critical to have an advanced understanding of rainfall patterns in the basin, setting the motivation of this study.



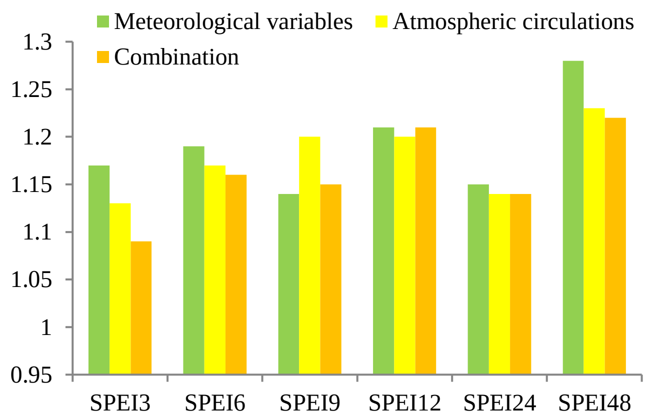
(a) Confidence Index values for SPEI time steps (present)



(b) LIFT index values for SPEI time steps (present)



(c) Confidence Index values for SPEI time steps (predict)



(d) LIFT index values for SPEI time steps (predict)

Figure 10- Confidence and Lift indicators of best rules

4. Conclusions

The drying of Lake Urmia and its socioeconomic side effects is one of Iran's major emerging environmental problems. This tragic phenomenon attracts the attention of many researchers in water resources and agricultural scientific communities. Many studies have used meteorological models to predict droughts in the Lake Urmia basin. However, the relationship between meteorological variables and atmospheric circulation and Lake Urmia's meteorological drought had never been determined. Therefore, this study focused on examining the droughts in this basin to develop predictive models in conjunction with large-scale atmospheric circulation indices and climate variables. The data mining algorithms were used to identify drought episodes associated with large-scale atmospheric circulation indices and climate variables. This study considered the meteorological drought in the Lake Urmia basin based on the SPEI index. SPEI index considers the effects of evapotranspiration with precipitation in drought. For the first time, in this study, a combination of meteorological variables, atmospheric circulations, and sea surface temperatures in the surrounding seas using data mining prediction methods provided a more comprehensive understanding of droughts. However, due to free access to the atmospheric circulations index data, the study area is limited to the Urmia basin. Our evaluation depends on three scenarios: the first included only the climate variables, the second with the atmospheric circulation indices, and finally, the third with a combination of both set variables. In the M5 tree model, forecasting was done for the next time step. The results showed that both methods are highly reliable and can be used for drought index modeling.

According to the results, atmospheric signals effectively forecast long-term drought. However, it was determined that the meteorological variables are more effective in forecasting short-term drought. Previous studies showed that atmospheric signals effectively forecast long-term precipitation and drought (Tadesse et al. 2004; Nikzad et al. 2013). Confidence and Lift indicators were computed to evaluate the rules extracted via the association rules method. A relationship was found between large-scale atmospheric circulation indices and climate variables and a combination of both with the SPEI index at different time steps. However, this relationship and efficiency vary regarding the delay time. The results revealed that regional atmospheric circulations have long-term impacts on weather patterns, and meteorological variables also play a significant role in short-term impacts. Therefore, it is plausible to use both sets of variables to reach acceptable results. Considering the recent droughts in the catchment area of the Lake Urmia and the other lakes with similar climatic characteristics, the availability of data on regional atmospheric circulations via satellite imagery and the use of introduced methods, including Association Rules and M5 tree models facilitate providing simple and understandable rules and relationships; therefore, it may be suggested as a new approach in forecasting droughts with relatively high precision. The approach outlined in this paper can be used in any earth region with a relatively similar climatic characteristic to the Lake Urmia basin. It is suggested that other indicators be used to study drought in the region. Considering that in this research, the Thiessen averaging of the region has been used to study the drought, it is suggested that other methods be used in the research for averaging the stations in the region (such as is pluvial line) and the results obtained with the results. It is suggested that other data mining methods, such as support vector machines or gene expression, be used to study and predict drought indicators. The effect of meteorological data on rainfall and drought of each station can be examined one by one. Also, it is suggested to use more stations in similar studies.

Acknowledgments

We appreciate the kind efforts of Burhan Gürtunca and Bediz Çeliker, who are English Language lecturers in the School of Foreign Languages of ITU North Cyprus, for improving the language quality of the manuscript.

Conflict of Interest: The authors have no conflicts of interest to declare that are relevant to the content of this article.

Funding Source: This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Data availability: Data for this research are not publicly available due to governmental policies but available on request from the corresponding author.

References

Alam N M, Sharma G C, Moreira E, Jana C, Mishra P K, Sharma N K & Mandal D (2017). Evaluation of drought using SPEI drought class transitions and log-linear models for different agro-ecological regions of India. Physics and Chemistry of the Earth Parts A/B/C 100: 31-43. https://doi.org/10.1016/j.pce.2017.02.008

Alizade Govarchin Ghale Y, Altunkaynak A & Unal A (2018). Investigation anthropogenic impacts and climate factors on drying up of Urmia Lake using water budget and drought analysis. Water resources management 32: 325-337. https://doi.org/10.1007/s11269-017-1812-5

Alizadeh-Choobari O, Ahmadi-Givi F, Mirzaei N & Owlad E (2016). Climate change and anthropogenic impacts on the rapid shrinkage of Lake Urmia. *International Journal of Climatology* 36(13): 4276-4286. https://doi.org/10.1002/joc.4630

Arkian F, Nicholson S E & Ziaie B (2018). Meteorological factors affecting the sudden decline in Lake Urmia’s water level. Theoretical and Applied Climatology 131: 641-651. https://doi.org/10.1007/s00704-016-1992-6

Barlow M, Nigam S & Berbery E H (2001). ENSO Pacific decadal variability and US summertime precipitation drought and streamflow. *J Clim* 14: 2105–2128. https://doi.org/10.1175/1520-0442(2001)014<2105:EPDVAU>2.0.CO;2

Buttafuoco G, Caloiero T, Ricca N & Guagliardi I (2018). Assessment of drought and its uncertainty in a southern Italy area (Calabria region). Measurement 113: 205-210. https://doi.org/10.1016/j.measurement.2017.08.007

Danandeh Mehr A & Fathollahzadeh Attar N (2021). A gradient boosting tree approach for SPEI classification and prediction in Turkey. *Hydrological Sciences Journal* 66(11): 1653-1663. https://doi.org/10.1080/02626667.2021.1962884

Dehghani M, Salehi S, Mosavi A, Nabipour N, Shamshirband S & Ghamisi P (2020). Spatial analysis of seasonal precipitation over Iran: Co-variation with climate indices. *ISPRS International Journal of Geo-Information* 9(2): 73. https://doi.org/10.3390/ijgi9020073

Dracup J A & Kahya E (1994). the relationships between US streamflow and La Niña events. Water Resources Research 30(7): 2133-2141. https://doi.org/10.1029/94WR00751

Etemad-Shahidi A & Mahjoobi J (2009). Comparison between M5 model tree and neural networks for prediction of significant wave height in Lake Superior. Ocean Engineering 36(15-16): 1175-1181. https://doi.org/10.1016/j.oceaneng.2009.08.008

Fayaz S A, Zaman M & Butt M A (2022). Numerical and Experimental Investigation of Meteorological Data Using Adaptive Linear M5 Model Tree for the Prediction of Rainfall. Review of Computer Engineering Research 13: 9(1):1-2

Hall M A (1999). Correlation-based Feature Selection for Machine Learning, Ph.D. thesis University of Waikato.

Hayes M (2003). Drought indices. Available online at http://www.drought.unl.edu/whatis/indices.html

Hosseini S M, Kashki A & Karami M (2022). Analysis of the North Atlantic Oscillation Index and rainfall in Iran. Modeling Earth Systems and Environment 1-10. https://doi.org/10.1007/s40808-021-01309-y

Indurkhya N & Weiss S M (1998). Estimating performance gains for voted decision trees. Intelligent Data Analysis 2(4):303-310. https://doi.org/10.1016/s1088-467x(98)00028-6

Inoubli R, Abbes A B, Farah I R, Singh V, Tadesse T & Sattari M T (2020). A review of drought monitoring using remote sensing and data mining methods. In 2020 5th International Conference on Advanced Technologies for Signal and Image Processing (ATSIP) (pp. 1-6). IEEE. https://doi.org/10.1109/atsip49331.2020.9231697

Kahya E & Karabörk M C (2001). The Analysis of El Nino and La Nina Signals in stream flows of Turkey. International Journal of Climatology 21:1231-1250. https://doi.org/10.1002/joc.663

Kuzay M, Tuna M & Tombul M (2022). Determining the Relationship of Evapotranspiration with Precipitation and Temperature Over Turkey. *Journal of Agricultural Sciences* 28(3): 525-534. https://doi.org/10.15832/ankutbd.952845

Le M H, Perez G C, Solomatine D & Nguyen L B (2016). Meteorological drought forecasting based on climate signals using artificial neural network–a case study in Khanhhoa Province Vietnam. Procedia Engineering 154:1169-1175. https://doi.org/10.1016/j.proeng.2016.07.528

Lee J H & Julien P Y (2017). Influence of the El Niño/Southern Oscillation on South Korean Streamflow Variability: El Niño/Southern Oscillation on South Korean Streamflow Variability. Hydrological Processes 12:2162-2178. https://doi.org/10.1002/hyp.11168

Masocha M, Murwira A, Magadza C H, Hirji R & Dube T (2017). Remote sensing of surface water quality in relation to catchment condition in Zimbabwe. Physics and Chemistry of the Earth Parts A/B/C 100:13-18. https://doi.org/10.1016/j.pce.2017.02.013

McKee T B, Doesken N J & Kleist J (1993). The relationship of drought frequency and duration to time scales. In Proceedings of the 8th Conference on Applied ClimatologyAMS Boston MA 17(22):179–184

Mehdizadeh S, Ahmadi F, Mehr A D & Safari M J S (2020). Drought modelling using classic time series and hybrid wavelet-gene expression programming models. *Journal of Hydrology* 587: 125017

Miley K M, Downs J, Beeman S P & Unnasch T R (2020). Impact of the Southern Oscillation Index, temperature, and precipitation on Eastern equine encephalitis virus activity in Florida. *Journal of Medical Entomology* 57(5): 1604-1613. https://doi.org/10.1093/jme/tjaa084

Moreira E E (2016). SPI drought class prediction using log-linear models applied to wet and dry seasons. Physics and Chemistry of the Earth Parts A/B/C 94:136-145. https://doi.org/10.1016/j.pce.2015.10.019

Nam WH, Tadesse T, Wardlow BD, Hayes M J, Svoboda M D, Hong E M & Jang M W (2018). Developing the vegetation drought response index for South Korea (VegDRI-SKorea) to assess the vegetation condition during drought events. *International Journal of Remote Sensing* 39(5): 1548-1574. https://doi.org/10.1080/01431161.2017.1407047.

Nikzad M, Behbahani M & Rahimi Good A (2013). Detection of dependencies between oceanic-wet and climatic parameters for drought monitoring by data mining method Case study Khuzestan province. *Iranian Journal of Water Research* (13)7: 183-175

Nourani V, Davanlou Tajbakhsh A, Molajou A & Gokcekus H (2019). Hybrid wavelet-M5 model tree for rainfall-runoff modeling. *Journal of Hydrologic Engineering* 24(5): 04019012. https://doi.org/10.1061/(asce)he.1943-5584.0001777

Nourani V, Sattari M T & Molajou A (2017). Threshold-based hybrid data mining method for long-term maximum precipitation forecasting. Water Resources Management 31(9): 2645-2658. https://doi.org/10.1007/s11269-017-1649-y

Nourani V, Tajbakhsh A D & Molajou A (2019). Data mining based on wavelet and decision tree for rainfall-runoff simulation. Hydrology Research 50(1): 75-84. https://doi.org/10.2166/nh.2018.049

Palmer W C (1965). Meteorological Drought. Research paper, n 45, U. S. Department of Commerce Weather Bureau, Washington, D. C: 58

Quinlan J R (1992). Learning with Continuous Classes. Proceedings of AI'92 World Scientific 92: 343–348

Sattari M, Mirabbai Najafabadi R, Alimohammadi, M (2016). Application of M5 Tree Model in Forecasting Drought (Case Study, Maragheh, Iran). Hydrogeomorphology 3(8): 73-92. https://doi.org/10.1134/s0097807813030123

Sattari M T, Shaker Sureh F & Kahya E (2020). Monthly precipitation assessments in association with atmospheric circulation indices by using tree-based models. Natural Hazards 102: 1077-94. https://doi.org/10.1007/s11069-020-03946-5

Sezen C & Partal T (2017). The relation of North Atlantic oscillation (NAO) and North Sea Caspian pattern (NCP) with climate variables in Mediterranean region of Turkey. The Eurasia Proceedings of Science Technology Engineering and Mathematics (1): 366-371. https://doi.org/10.1007/s00703-019-00665-w

Sheffield J, Wood E F & Roderick M L (2012). Little change in global drought over the past 60 years. Nature 491: 435–438

Stagge J H, Tallaksen L M, Gudmundsson L, Van Loon A F & Stahl K (2015). Candidate distributions for climatological drought indices (SPI and SPEI). *Int J Climatol* 35: 4027–4040. https://doi.org/10.1002/joc.4267

Tadesse T, Wilhite D A, Harms S K, Hayes M J & Goddard S (2004). Drought Monitoring Using Data Mining Techniques: A Case Study for Nebraska USA. Natural Hazards 33(1): 137- 159. https://doi.org/10.1023/B:NHAZ.0000035020.76733.0b

Taylor K E (2001) Summarizing multiple aspects of model performance in a single diagram. *J Geophys Res* 106:7183–7192. https://doi.org/10.1029/2000JD900719

Uzun A & Ustaoğlu B (2022). The effects of atmospheric oscillations on crop (olive, grape and cotton) yield in the eastern part of the Mediterranean region, Turkey. *International Journal of Environment and Geoinformatics* 9(1): 147-161. https://doi.org/10.30897/ijegeo.1010181

Vicente-Serrano S M, Beguería S & Juan I A (2011). A multiscalar global evaluation of the impact of ENSO on droughts. *J Geophys Res Atm* 116 pp. https://doi.org/10.1029/2011JD016039

Vicente-Serrano S M, Beguería S & López-Moreno J I (2010). A multiscale drought index sensitive to global warming: the standardized precipitation evapotranspiration index. *Journal of Climate* 23(7): 1696-1718. https://doi.org/10.1175/2009JCLI2909.1

Wang W, Zhu Y, Xu R & Liu J (2015). Drought severity change in China during 1961–2012 indicated by SPI and SPEI. Nat Hazards 75: 2437–2451. https://doi.org/10.1007/s11069-014-1436-5