Pure Appl. Geophys.
© 2020 Springer Nature Switzerland AG
https://doi.org/10.1007/s00024-020-02609-7

Pure and Applied Geophysics

A Novel Fuzzy Random Forest Model for Meteorological Drought Classification and Prediction in Ungauged Catchments

Ali Danandeh Mehr,[1] Rifat Tur,[2] Cafer Çalışkan,[3] and Erkin Tas[2]

Abstract—This paper presents a new tree-based model, namely Fuzzy Random Forest (FRF), for one month ahead Standardized Precipitation Evapotranspiration Index (SPEI) classification and prediction with a noteworthy application in ungauged catchments. The proposed FRF model uses global SPEI dataset as the meteorological drought quantifier and applies a fuzzy inference system to extract fuzzified and crisp SPEI values for an ungauged catchment. The evolved crisp series is then transformed into the polynomial label vector of extremely wet, wet, near normal, dry, and extremely dry categories. In the end, the state-of-the-art random forest algorithm was used to classify and predict the label vector using the lagged SPEI series of the selected global grid points. To demonstrate the development and verification process of the FRF model, the global SPEI-6 values for the period of 1961–2015 were retrieved from four global grid points around the Central Antalya Basin, Turkey. The new model was trained and validated using 70% and 30% of the data sets, respectively. The performance of the new model was examined in terms of total accuracy, misclassification, and Kappa statistics and cross-validated with the fuzzy decision tree model developed as the benchmark in this study. The results showed the promising performance of the FRF for drought classification and prediction with outstanding efficiency for extremely wet and dry events classification. According to the Kappa statistic, the proposed FRF model is 25% more accurate than the benchmark FDT model.

Keywords: Drought, SPEI, Classification models, Decision tree, Fuzzification, Random forest.

1. Introduction

Drought is a climatological extreme that undesirably affects human life through limited access to freshwater. It makes significant impacts on the quantity and quality of water resources at local and regional scales which might yield famine and socioeconomic crisis (Danandeh Mehr et al. 2020b). To measure drought severity, the corresponding indices are calculated using a set of certain meteorological, agricultural, hydrological, or even socioeconomic information. From a meteorological drought perspective, variables such as precipitation and temperature are commonly used to calculate the drought index. There are several meteorological drought indices, such as the standardized precipitation index (SPI; McKee et al. 1993), Palmer's drought severity index (PDSI; Palmer 1965), drought area index (DAI; Bhalme and Mooley 1980), and the most recently developed standardized precipitation evapotranspiration index (SPEI; Vicente-Serrano et al. 2010). Historical precipitation data is used to calculate SPI and ADI, whereas temperature is also required to calculate PDSI and SPEI. The time scale of an index is also important in the evaluation of drought impacts. For instance, the SPEI with a monthly scale (i.e., SPIE-1) is a criterion indicating the monthly variation of wet and dry spells with respect to the long-term mean precipitation and evapotranspiration in a region. The choice of the time scale, in practice, depends on the aim of the study. Despite being a meteorological index, use of short time scales (1-, 3- or 6-month) is suggested for meteorological and agriculture drought analysis, while the longer time scales, such as -9 or -12 can be used for hydrological drought monitoring (Caloiero and Veltri 2019; Alsafadi et al. 2020).

Because of the importance of sustainable water supply for human life, many studies have addressed both natural and anthropogenic climate change impacts on water resources (e.g., Maghrebi et al.

[1] Department of Civil Engineering, Antalya Bilim University, Antalya, Turkey. E-mail: ali.danandeh@antalya.edu.tr
[2] Department of Civil Engineering, Akdeniz University, Antalya, Turkey.
[3] Department of Computer Engineering, Antalya Bilim University, Antalya, Turkey.

Published online: 03 November 2020

2020; Danandeh Mehr et al. 2020a). Several studies have also attempted to improve the accuracy of drought prediction/forecasting models (e.g., Pongracz et al. 1999; Bogardi et al. 2004; Danandeh Mehr et al. 2020b). To this end, the capability of different machine learning (ML) techniques such as artificial neural networks (ANN; e.g., Mishra et al. 2007; Morid et al. 2007; Barua et al. 2012; Rezaeian-Zadeh and Tabari 2012; Nair et al. 2018), support vector machine (SVM; e.g., Belayneh and Adamowski 2012), fuzzy logic (FL; Pesti et al. 1996; Pongracz et al. 1999; Bogardi et al. 2004; Keskin et al. 2009; Ozger et al. 2011; Huang et al. 2015; Abdourahamane and Acar 2019) and genetic programming (Mehr et al. 2014; Abbasi et al. 2019; Danandeh Mehr et al. 2021) were extensively studied.

Focusing on the FL-based drought forecasting models, our review showed that a few applications exist in the relevant literature. For instance, in a seminal paper, Pesti et al. (1996) presented a fuzzy-based model to forecast regional drought in New Mexico using atmospheric pressure patterns of the Western United States. To create a fuzzy model, triangular function membership was applied, and 16 linguistic fuzzy rules were derived to make a relationship between drought characteristics and atmospheric pressure patterns. Pangracz et al. (1999) developed a fuzzy rule-based model to forecast regional drought in terms of Palmer Modified Drought Index (PMDI) for Nebraska. Different large-scale atmospheric circulation patterns were used as the PMDI predictors. A comparison was performed between the observed and FL-predicted values. The authors showed that the FL can regress observed and multivariate estimated PMDI time series for South-Central Nebraska. Huang et al. (2015) developed an Integrated Drought Index that combines the meteorological, hydrological, and agricultural factors by using fuzzy set theory. The study area includes the Yellow River, the second largest river in China. The results showed that the developed drought index with applying fuzzy set theory is reliable and can be used for integrated drought studies. More recently, Abdourahamane and Acar (2019) developed a fuzzy-rule based model to forecast 3-month SPI of Western Niger. The southern oscillation index, sea surface temperature, relative humidity, and sea level pressure were used on the model as the SPI predictors. Fuzzy membership functions and linguistic rules were derived considering expert knowledge and literature survey. Statistical analysis showed that the developed fuzzy-based model may provide accurate drought forecasts.

Despite the inclusion of human knowledge, inadequate accuracy of in FL-based models for long lead time forecast was also reported (Mehr et al. 2014). To overcome the problem and enhance forecasting accuracy, hybrid ML models that typically are attained through the combination of different ML techniques were suggested (e.g., Shirmohammadi et al. 2013; Deo et al. 2017; Kisi et al. 2019 among others). For example, Ozger et al. (2011) developed a wavelet-fuzzy logic (WFL) model to forecast drought in different climate zones of Texas using several meteorological parameters as predictors. The Palmer Drought Severity Index was used as the predictand on the developed model. Özger et al. (2012) presented three different drought forecasting models including WFL, ANN, and wavelet-ANN. ENSO and previous PMDI observations were used to forecast long lead drought in Texas. The comparisons of the developed model showed that the hybrid WFL model is superior to other models to forecast regional drought.

A recent review study conducted by Fung et al. (2019) showed that most of the existing drought forecasting studies have focused on either regression analysis between climate indicators and drought pattern or time series modeling of drought indices (Farokhnia et al. 2011; Özger et al. 2012; Danandeh Mehr et al. 2014; Belayneh and Adamowski 2012; Nguyen et al. 2015; Mokhtarzad et al. 2019; Abdourahamane and Acar 2019; Mehdizadeh et al. 2020; Danandeh Mehr et al. 2021). Only a few studies have attempted to model and predict drought classes. For example, Chiang and Tsai (2013) designed a two-stage SVM for reservoir drought prediction and showed that classification accuracy of the two-stage SVM is higher than standalone SVM, ANN, Maximum likelihood, and Bayes classifiers. Nourani and Molajou (2017) proposed a hybrid decision tree-association rules model to discover the relationship among SPI-based drought events in Tabriz City and Kermanshah City and sea surface temperature of the Black Sea, the Mediterranean Sea, and the Red Sea. The state-of-the-art random forest (RF) technique was applied for SPI time series forecasting by Chen et al. (2012). The authors also applied RF, for the first time, to forecast the number of dry days in four cities in China. The results demonstrated that the RF is more efficient than the classic autoregressive integrated moving average model for both short- and long-term drought classification. Park et al. (2016) applied three ML models including RF, boosted regression trees, and Cubist for SPI time series modeling based on remote sensing meteorological data. Park et al. (2018) also applied RF for short-term drought forecasting in the East Asia region using satellite-based climatic data. The study showed that the RF can capture the sudden change in drought conditions when Madden–Julian oscillations is used as a feature for drought prediction.

The aforementioned studies have proven the robustness of RF in drought forecasting using regression analysis. The promising role of RF for probabilistic nowcasting of low-visibility procedure states was also reported in a recent study by Dietz et al. (2019). Inspired by the high performance of RF for regression tasks, the main objectives of this study are (i) to explore the efficiency of RF for drought classification and (ii) for the first time, to develop and apply a novel hybrid fuzzy random forest (hereafter FRF) for draught classification, 1-month in-advance. As RF is from the classic decision tree (DT) family, a Fuzzy-DT (hereafter FDT) model was also developed in this study as the benchmark. To the best of the author's knowledge, the efficiency of FRF and FDF for SPEI classification/prediction has never been explored yet.

2. Study area and data exploration

Antalya province with a population of more than two million is located on the Mediterranean coast of south-west Turkey, between the Taurus Mountains and the Mediterranean Sea. The Mediterranean climate prevails in the province with hot and dry summers and rainy springs and winters. Its temperature ranges from − 4.3 °C to 43.4 °C during a year. It is very rare to cold down to 10° in winter. The average precipitation is 1070 mm per square meter. To obtain local monthly SPEI series around the Central Antalya Basin (CAB), the global SPEI database, which offers long-time SPEI series at a global scale with a 0.5 degrees spatial resolution was used in this study. The database has a multi-scale character, providing SPEI timescales between 1 and 48 months. The historical SPEI-6 data in the period between January 1961 and December 2015 from four global grid points (i.e., G1, G2, G3, G4) located in the study area were used (Fig. 1) in the present study. For details on the calculations of SPEI in different time scales are referred to Vicente-Serrano et al. (2010) but here, it is worth mentioning that SPEI-6 reflects rainfall and temperature conditions over the past 6-month period. Danandeh Mehr and Vaheddoost (2020) reported that the index also is suitable for monitor and analyze drought under climate change.

Figure 2 depicts the time distribution of SPEI-6 series at each grid point attained for the period of 1961–2015. The statistical characteristics of the series were summarized in Table 1. The fuzzy inference system (FIS) based on 81 fuzzy rules exclusively defined by experts (is explained later in Methodology Section) was applied using these SPEI-6 series to produce a reliable set of SPEI-6 for the CAB.

3. Methods

3.1. Overview of Decision Tree (DT)

DT, conceptually, is an if–then–else algorithm that can be used to predict a result based on historical experiments. DT is a supervised ML algorithm with a tree-like structure where the tree has branches, leaves, and the root node at the top. In general, the process of prediction (deciding) starts with exploring the existing attributes (input vectors) in order to split the data into subsets. In other words, the data is separated according to a series of questions relevant to the interrelations among the attributes and a given label. Each subset represents a certain class of data available in the associated attribute. For example, if outlook, humidity, and wind are considered as the attributes to predict the temperature of a day, the outlook vector can be divided into subsets of sunny, overcast, and rainy. The process of subdividing the

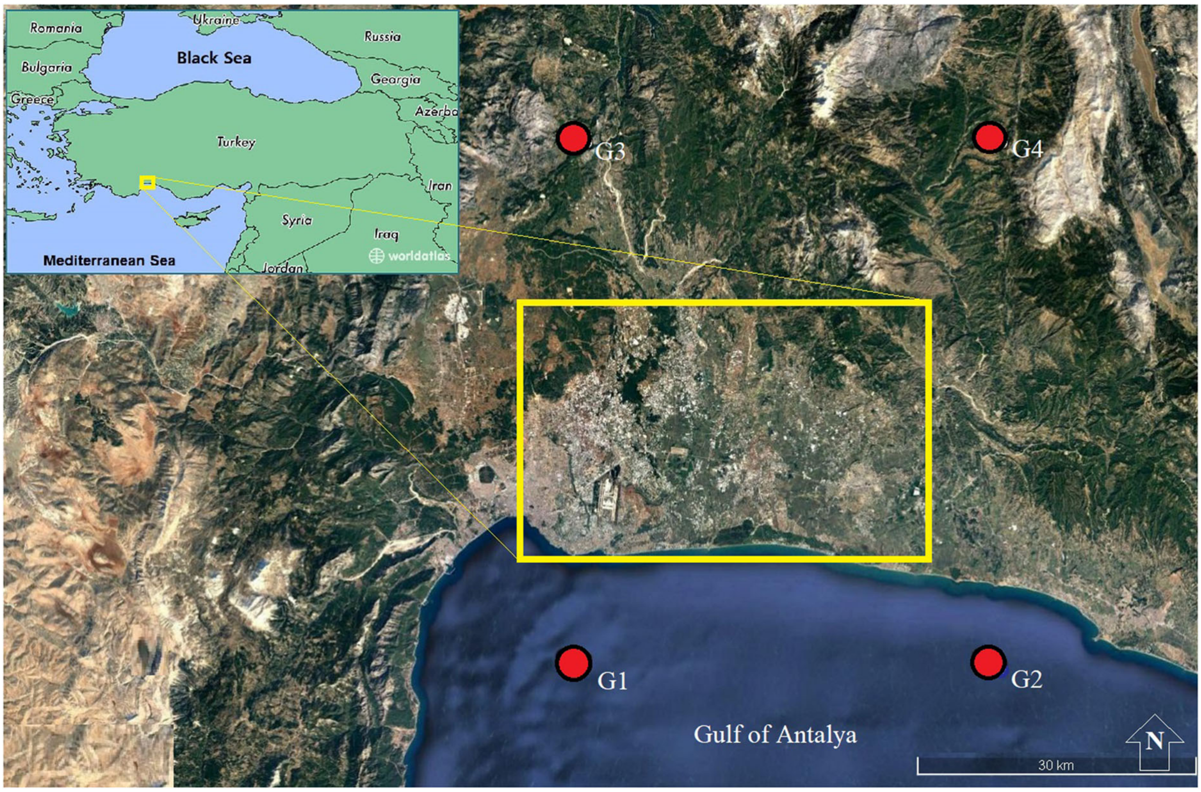

Figure 1
Location of the closest global grid points (*red points*) to the CAB (*yellow polygon*) in Turkey

dataset into smaller and smaller subsets of the same or related type is continued like the divide-and-conquer algorithm of sorts until a subset becomes pure which means the branch is simple enough to decide directly. The result is a multi-branched tree that is used to predict future experiments. To this end, the modeler looks at which subset that the new experiments fall into and then uses the dominant class in that subset.

To characterize impurity in a data set, DT applies entropy theory and fits a classification tree for the data given by the primary information of each class. Considering a binary classification problem with label L which consists merely of yes and no samples, the entropy of L is expressed as below.

$$Entropy(L) = -p_y(log_2 p_y) - p_n(log_2 p_n) \quad (1)$$

where p_y and p_n are the proportions of yes and no samples, respectively.

The DT corresponding to the SPEI classification described in Sect. 2 was illustrated in Fig. 3. The tree starts with the root node question of splitting SPEI into positive (yes) and negative (no) classes. Then, the yes answers go the one way (right hand) and the no goes to the other way (left hand). The yes/no questions are continued until we reach the leaves of the tree that figures out a pure meteorological drought condition.

3.2. Overview of Random Forest (RF)

RF (Breiman 2001) method is an ensemble learning algorithm that aims, firstly, to build small decision trees with few features each of which is computationally feasible and, secondly, to combine these constructed trees to form a strong learning process. In this learning process, for a given input vector, computation is carried over each of the trees

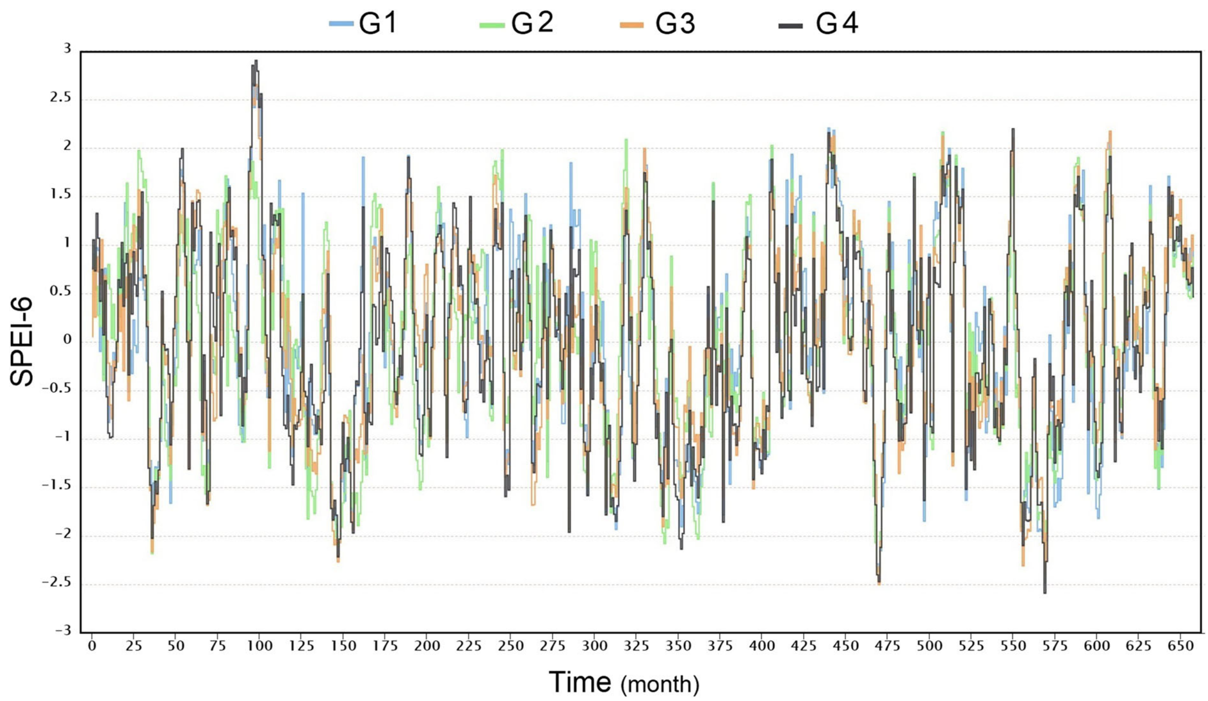

Figure 2
SPEI-6 series (classification features) at the grid points nearby the study area

Table 1

Statistical characteristics of SPEI time series used in the study

Grid point	Latitude	Longitude	Min	Mean	Max	Standard deviation	Skewness
G1	36.75	30.75	− 2.483	0.03	2.655	1.05	0.025
G2	36.75	31.25	− 2.473	0.01	2.165	1.03	− 0.085
G3	37.25	30.75	− 2.502	0.037	2.657	1.00	− 0.008
G4	37.25	31.25	− 2.591	− 0.017	2.901	1.02	0.065

in the forest. Once the classification model is obtained from each tree, the method chooses the most common class outcome. The error rate basically depends on the correlation between any two trees as well as the strength of individual trees in the forest. If the forest includes individual trees with a low error rate, it will decrease the overall error rate of the forest. Also, less correlation among the trees gives rise to a lower error rate for the overall process.

RF method is relatively a fast method compared to other classifier methods. In the forest, each tree is constructed using a different bootstrap sample. For the construction of a tree T, some portion of the original data is not taken into consideration and this left-out data part (so-called out-of-band (OOB) data) is used for testing purposes for this particular tree T. RF performance mainly depends on the number of ancillary data in each random tree and the number of trees in the forest (Ghorbani et al. 2020). The increase in the number of trees does not create overfitting problems. On the other side, using too many trees may not contribute to the classification accuracy. These parameters are typically optimized by an iterative method. As a result, the authors use 100 many trees when running the random forest method. Moreover, the maximal tree depth is kept as 15. For

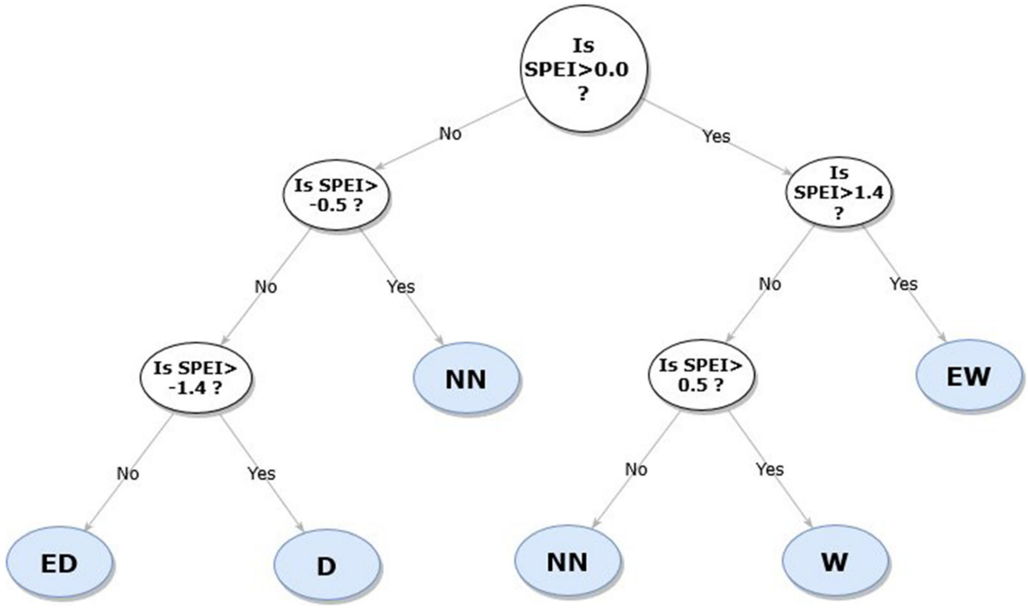

Figure 3
The perfect DT classifier for meteorological drought classification based on SPEI series at a single station

details about the fundamentals of RF and its applications in water engineering, the interested reader is referred to the comprehensive review paper that has recently been provided by Tyralis et al. (2019).

3.3. *Overview of Fuzzy Inference System*

Fuzzy logic (Zadeh 1965) is a soft computing technique that attempts to mathematically emulate human reasoning to solve complex problems. It allows modelers to employ human knowledge and decision-making rules to resolve the systems where an explicit analytical-process model is not available (Carr and Shearer 2007). The fuzzy logic system comprises three main elements to transform descriptive variables into a scalar response variable. These are input/output membership functions that can take any value in the range between 0 and 1, fuzzy rules, and an inference engine (Abdourahamane and Acar 2019). In a conventional implementation approach, first, a degree of membership (characteristic) is defined for each variable and then a set of fuzzy rules is created by an expert(s) to define the existing relationship between individual input and output variables. In the last step, the resulting fuzzy output set is transformed into a numerical (crisp) set throughout a defuzzification technique. In the following, the fuzzy rule-based methodology will be used to the present case is described in a step-by-step manner. For more details about fuzzy systems and its applications in water resources engineering, the interested reader is referred to Tayfur (2014).

3.4. *The Proposed Fuzzy Random Forest (FRF) Model*

As illustrated in Fig. 4, the FRF modeling process for meteorological drought classification and prediction is commenced by the collection of a set of input vectors from the global SPEI repository. We selected SPEI-6 from four grid points (G1, G2, G3, and G4) so that they are closest to the study area. In the next step, the target SPEI-6 for the study area (label) is created through the FIS illustrated in Fig. 6b. Three fuzzy linguistic terms including *low*, *medium,* and *high* were defined for all the inputs. Regarding the output, five fuzzy linguistic terms, namely extremely wet (EW), extremely dry (ED), dry (D), near normal (NN), and wet (W), were also defined in our implementation. Inasmuch as no sharp distinction

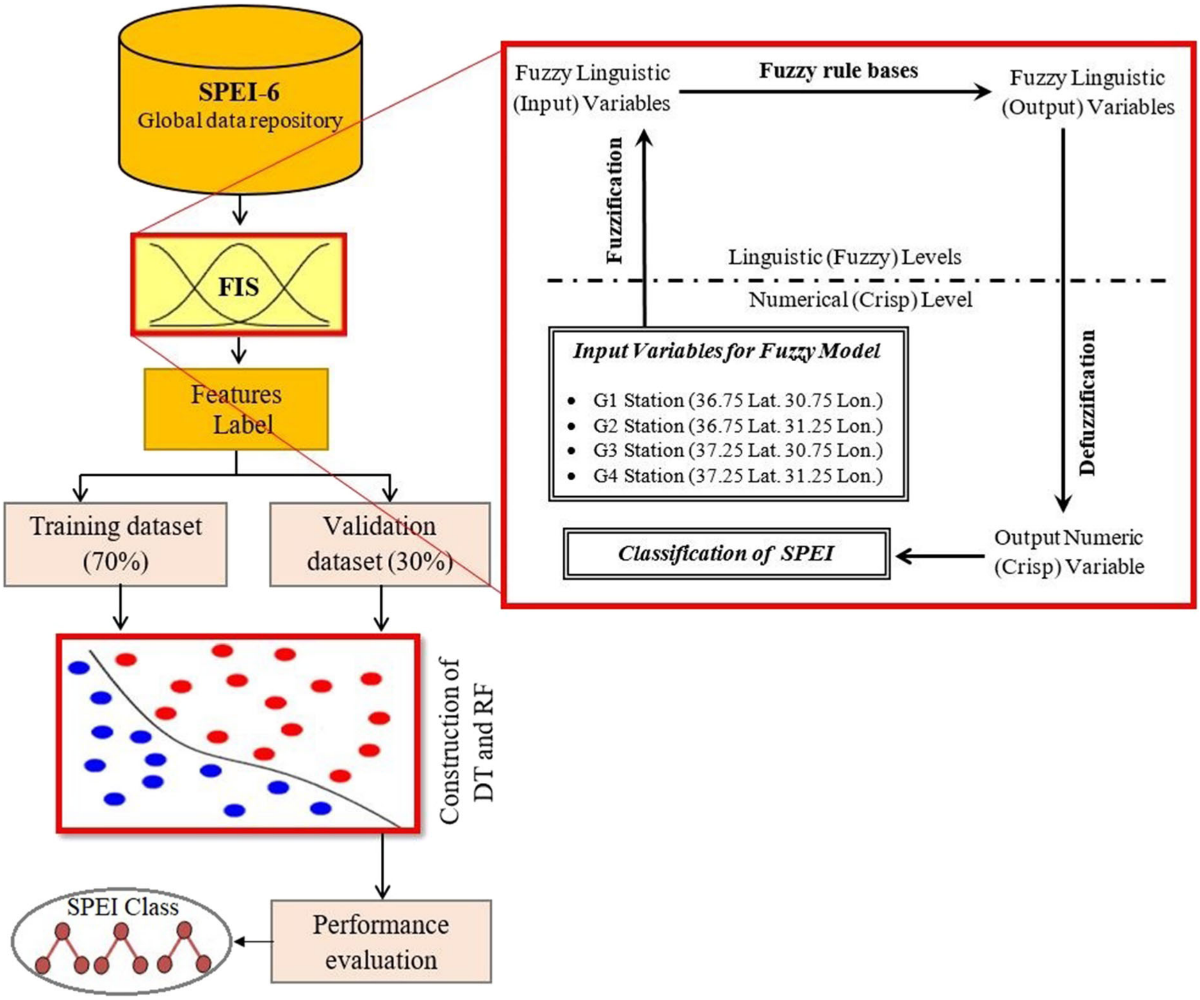

Figure 4
Methodology flowchart of the construction of the FRF model

usually exists between low, medium, and high drought, Gaussian membership function was used to assign a degree of membership for each sample of the input and output SPEI vectors.

Figure 5 compares the classical SPEI sets (Fig. 5-left) of low (SPEI < -1), medium ($-1 \leq$ SPEI ≤ 1), and high (SPEI > 1) with the fuzzified SPEI set used in the present study. If one conventionally classifies (evaluates the degree of membership), for example, the SPEI value 0.5, they get degree 1 for set medium and 0 for sets low and high. However, the corresponding classes (degrees) can be to some extent low (0.04), medium (0.23), and high (0.55) (Fig. 5right) in our FIS which is closer to expert reasoning and thus more realistic.

Since the FIS used in our use case has four input vectors and each of which has three fuzzy linguistic terms, 3^4 (or 81) fuzzy rules (complete rule bases) have been defined by the authors. The rules having the form of “IF condition THEN conclusion” inserted among the series of global SPEI vectors and a fuzzy set was created with respect to their degree of membership. Examples of the rules describing the relation between input and output linguistic variables are given below:

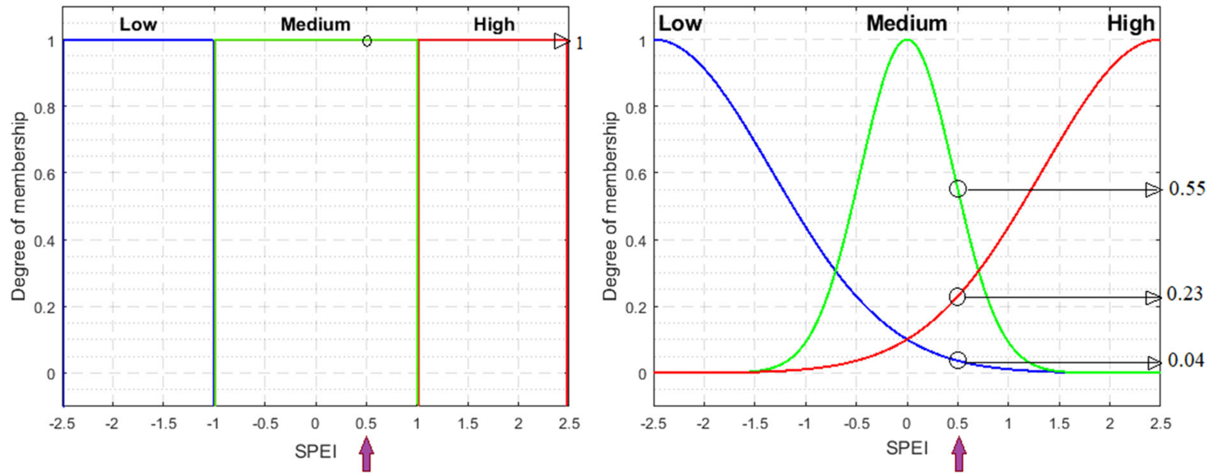

Figure 5
The classic (left) and fuzzy sets (right)

- **If** G1 is high & G2 is high & G3 is high & G4 is high, **then** SPEI in CBA is EW.
- **If** G1 is low & G2 is low & G3 is low & G4 is low, **then** SPEI in CBA is ED.
- **If** G1 is high & G2 is high & G3 is low & G4 is low, **then** SPEI in CBA is NN.
- **If** G1 is low & G2 is low & G3 is high & G4 is high, **then** SPEI in CBA is NN.
- **If** G1 is high & G2 is high & G3 is high & G4 is medium, **then** SPEI in CBA is EW.
- **If** G1 is low & G2 is low & G3 is low & G4 is medium, **then** SPEI in CBA is ED.
- **If** G1 is high & G2 is high & G3 is high & G4 is low, **then** SPEI in CBA is W.
- **If** G1 is low & G2 is low & G3 is low & G4 is high, **then** SPEI in CBA is D.

The last steps in the FIS is the defuzzification of our output fuzzy set (Fig. 6) followed by categorization of the attained crisp vector. During defuzzification, the output fuzzy sets for each rule

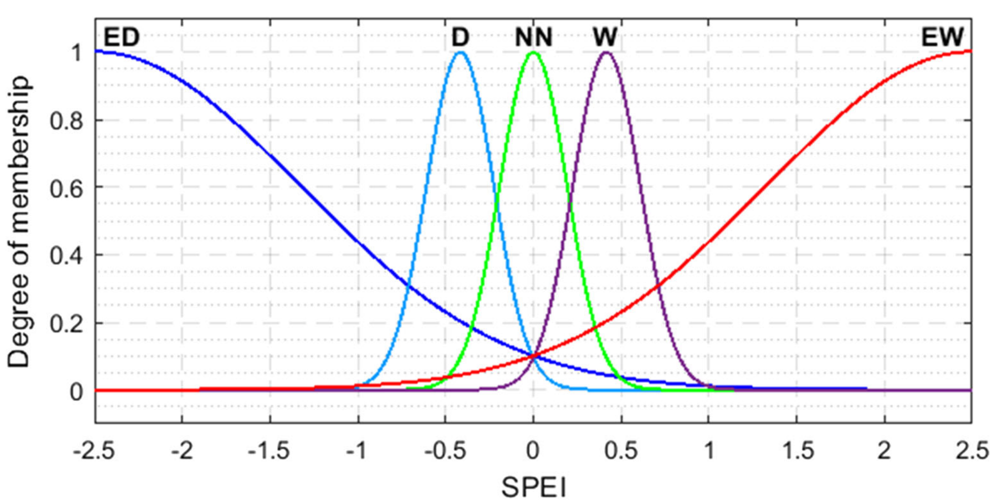

Figure 6
The Gaussian membership functions defined for output fuzzy sets attained based on the running of 81 rules in FIS

Table 2

Absolute count and fractions of meteorological drought across the Central Antalya Basin

Drought class	Acronym	Absolute count	Fraction
Wet	W	146	22.2
Near normal	NN	151	22.9
Dry	D	143	21.7
Extremely wet	EW	116	17.6
Extremely dry	ED	103	15.6

are aggregated to a single output fuzzy set and then each fuzzy linguistic output variable is converted to a crisp value. To this end, we used the most prevalent and physically appealing centroid (Center of Area-CoA) method (Ross 2004).

The resulting series is a vector of 659 samples of scalar SPEI-6 values that varies in the range [-2.03, 1.97]. Following the SPEI thresholds proposed by Danandeh Mehr et al. (2020a), we categorized the resulting series in five classes as tabulated in Table 2. The absolute count and fractions of each class across the CAB were also presented in the table.

Like other ML models, features and corresponding label are split into training and validations sets in our FRF model. The first 70% of the data were used to train the models. As previously mentioned, the classification/prediction algorithms used in this study is RF. However, a baseline DT is used for cross-validation task. Various statistical measures are used to evaluate the performance of the evolved FRF and FDT models.

3.5. Performance Evaluation Criteria

As the evolved models were developed for polynomial classification and forecasting tasks, their performance is assessed using confusion matrix and different Boolean statistics namely, total accuracy (TA), Kappa (KA; Landis and Koch 1977), recall (RE), and classification error (CE) in this study. The TA (Eq. 2) is the average number of correctly classified events. The kappa (Eq. 3) is a metric that compares an observed accuracy with an expected accuracy (random chance). It generally thought to be a more robust measure than simple percentage correct prediction calculation since it considers the correct prediction occurring by chance. The CE (Eq. 4) is the misclassification rate which equals 0.0 for a perfect classifier.

$$TA = \frac{\mathrm{TP} + \mathrm{TN}}{M} \times 100 \qquad (2)$$

$$KA = \frac{\mathrm{O}_{AG} + \mathrm{E}_{AG}}{1 - \mathrm{E}_{AG}} \qquad (3)$$

$$CE = 100 - \mathrm{AC} \qquad (4)$$

$$RE = \frac{\mathrm{TP}}{TP + \mathrm{FN}} \times 100 \qquad (5)$$

where TP (true positive predictions) and TN (true negative predictions) are the drought conditions classified correctly as classified in Table 2. FN is false negative predictions, and M is the total number of samples. The O_{AG} and E_{AG} are observed agreement and expected agreement, respectively. The former is defined as the number of samples that were classified correctly throughout the entire confusion matrix. The latter is the accuracy that any random classifier would be expected to achieve based on the confusion matrix.

4. Results

In this study, meteorological drought events in the CAB (see Table 2) is predicted over the lead time of 1-month which will be beneficial for early warning and taking preventive measures. The corresponding tree-based decision models are created based on the spatiotemporal relationship between SPEI-6 values of the nearby global grid points and the drought class in the CAB attained through FIS. Like other ML techniques in the first step, the entire input/output datasets at each scenario was divided into two subsets of training ($\sim$ 70%) and validation ($\sim$ 30%) periods. Then, DT and RF models were trained and validated using the corresponding data sets.

The evolved benchmark FDT model was illustrated in Fig. 7. In this tree, the most important variable to split on is G2. Inasmuch as the tree is created by the process of recursive partitioning, G2 followed by G3 and G5 are the dominant predictors that bubbled at the top nodes. It is seen that the samples were split regarding the threshold of -0.161.

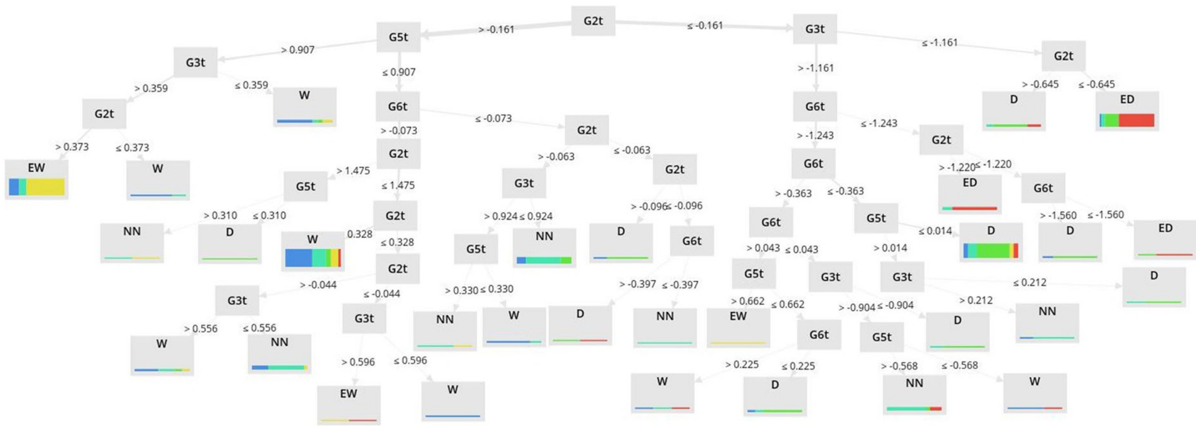

Figure 7
The DT model evolved for 1-month ahead drought classification in the CAB, Turkey

Table 3

Prediction results of FDT model

	True W	True NN	True D	True EW	True ED
Training samples (M = 468)					
Pred. W	61	26	9	18	5
Pred. NN	14	40	12	3	5
Pred. D	10	23	63	2	9
Pred. EW	19	11	0	59	1
Pred. ED	0	7	18	0	53
	True W	True NN	True D	True EW	True ED
Validation samples (M = 191)					
Pred. W	18	12	7	4	2
Pred. NN	6	20	2	0	1
Pred. D	5	6	27	4	7
Pred. EW	10	5	0	26	1
Pred. ED	3	1	5	0	19

Tables 3 presented the confusion matrix, i.e., the matrix of outputs, of the FDT model at training and validation sets. Each row in the table shows which drought event was predicted by the model. The columns are the ground truth indicating the number of events in each class. The values lying across the main diagonal shows the success of the model at each class. The summation of these values creates the numerator of Eq. (2) needed for AC calculations. The total number of events in the training and validation periods, M, were also given in the table that helps the reader to calculate all the performance criteria described in Sect. 3.5. The overall accuracy in the training and validation periods are 58.33% and 57.59%, respectively which were made up from accuracies of each drought class.

From the environmental management perspective, extreme weather conditions are more important to be precisely predicted. Indeed, EW and ED events may influence both biotic and abiotic factors more seriously. Hence, using the associated outcomes in Table 3 (columns 5 and 6), we calculated the classes recall recognizing how often the FDT correctly predicts the extreme events. Considering EW and ED outcomes in the training period, the model is 71.95% and 72.60% precise, respectively.

$$RE = \begin{cases} \frac{59}{18+3+2+59+0} \times 100 = 71.95 for EW \\ \frac{53}{5+5+9+1+53} \times 100 = 72.60 for ED \end{cases} \quad (6)$$

Looking at the validation period, the corresponding RE is 76.47% and 63.33%, respectively, which indicates the evolved FDT is not a highly reliable classifier, particularly for ED events.

To increase classification and forecasting accuracy, several RF models were developed, and the accuracy results of the best classifier were presented in Table 4. It's worth mentioning that the required modeling parameters were slightly altered to find their optimum values. The number of random trees was initialized by 100 and then increased to 1000. The maximal depth was limited to 15 to avoid overtraining. The classification matrices of the evolved FRF were tabulated in Table 4. Returning to

Table 4

Prediction results of FRF model

	True W	True NN	True D	True EW	True ED
Training samples (M = 468)					
Pred. W	66	13	7	12	0
Pred. NN	8	60	10	1	1
Pred. D	9	17	67	3	10
Pred. EW	20	11	0	66	0
Pred. ED	1	6	18	0	62
	True W	True NN	True D	True EW	True ED
Validation samples (M = 191)					
Pred. W	29	8	4	3	0
Pred. NN	5	25	3	1	1
Pred. D	3	7	26	0	2
Pred. EW	5	3	0	30	0
Pred. ED	0	1	8	0	27

the FDT results, it must be noted that a significant drop in the model performance was observed when the maximum depth of the tree could be larger than 10 which indicates an overtrading situation in FDT. However, in FRF, the model efficiency in both training and validation periods progressively increased up to 15 levels. The overall accuracy in the training and validation periods is 68.59% and 71.73%, respectively that shows significant improvement over the FDT model.

Like FDT evaluation, the recall index for ED and EW classes were calculated to recognize how often the FRF correctly predicts the extreme events. The results exhibited that the model is 80.49% and 84.93% precise, respectively. Looking at the recall in the validation period, the model is 88.23% and 90.00% precise, respectively, which indicates the evolved FRF is a highly reliable classifier for both ED and EW classes. It is seen from the table that from 73 ED events in the training period, the RF truly predicted 62 events. False predictions include 10 D and one NN so that no ED event was predicted as W or EW events. Regarding the validation outcomes, the model can predict 27 ED events out of 30. As the remaining three events were classified as the D or NN cases, the robust performance of the model for ED classification is concluded.

To get a sense of how well the evolved FRF works in practice, cross-validation with the evolved FDT

Table 5

Performance results of the FDT and FRF models in the training and validation periods

Model	Training			Validation		
	TA (%)	CE (%)	KA	TA (%)	CE (%)	KA
FDT	58.97	41.03	0.49	57.59	42.41	0.47
FRF	68.59	31.41	0.61	71.73	28.27	0.65

model was accomplished in terms of overall model accuracy, misclassification rate, and Kappa statistics (Table 5). It is seen from the table that the RF is superior to the FDT in terms of all the performance measures. In both training and validation datasets, the misclassification rate of the RF is approximately less than 30% that approximately shows a 25% and 30% reduction in the misclassification rate of the FDT. According to KA statistic in the training dataset, which allows comparing the classifiers based on the number of drought events along with the number of instances that correctly labeled at each class, the FRF satisfactorily classified the events 25% more accurate than FDT. It is also 34% more precise in the validation period. Following Landis and Koch (1997) who considers KA in the range [0.21–0.40] as fair, [0.41–0.60] as moderate, and [0.61–0.80] as substantial classification, the FDT and FRF models are respectively interpreted as moderate and substantially satisfactory classifiers in this study.

5. *Discussion*

Figure 8 displays the count of ground truth (observed) and predicted drought classes in the training and validation periods. The figure reveals that the CAB has experienced 104 (42) W, 108 (42) NN, 102 (41) D, 83 (33) EW, and 71(32) ED months during the training (validation) period. The applied FIS, naturally classified most of the observed samples as NN, followed by W and D events in both training and validation periods. Regarding the extreme event, the basin has faced more EW events than ED (e.g., 83 vs. 71 in the training period). However, the number of ED months is so close to that of EW (32 vs. 33) in the

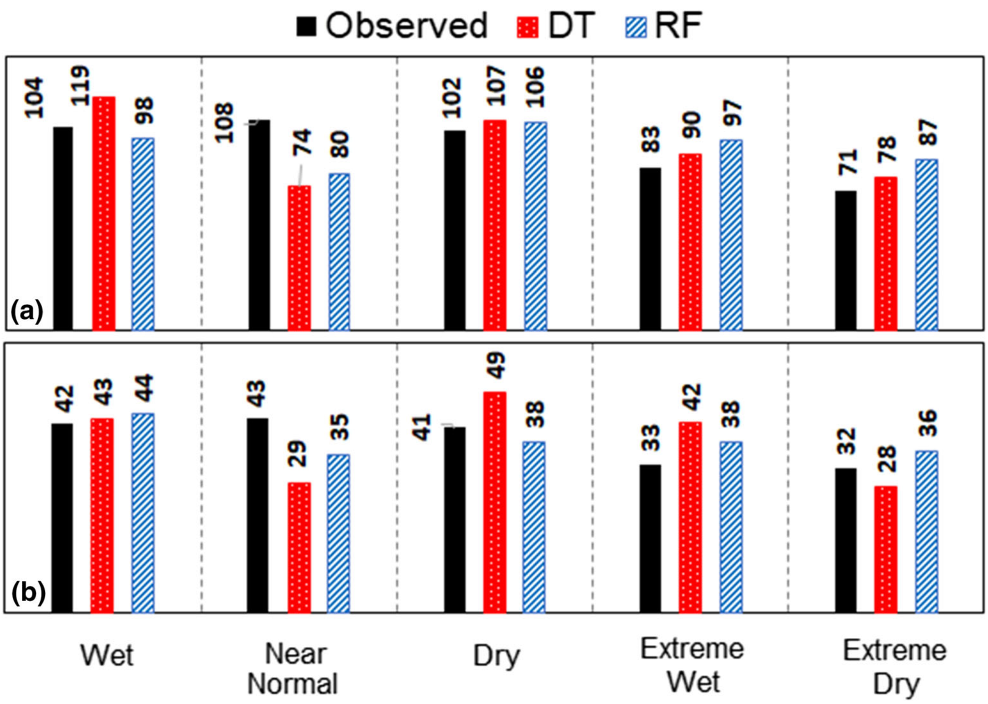

Figure 8
The observed and classified drought conditions in **a** training and **b** validation period

validation period that indicates an increasing trend for ED events in the most recent years which could be considered as the signal for climate variability in the basin. Both FDT and FRF overclassified extreme events in the training period with a slight superior performance of the FDT over FRF. However, the FRF shows better performance in the validation period. By contrast, both the models underestimate the NN conditions, but the RF is still superior to its counterpart. The FDT predictions are very poor in both training and validating periods. The FRF exhibits a greater class precision. Despite the higher total accuracy of the FRF, the W events were perfectly classified by both FRF and FDT in the validation period.

6. Conclusion

Drought prediction is known as a challenging task owing to the extremely stochastic feature of drought classification indices. While most of the preceding studies attempted to model the temporal variation of drought indices using station-based historical indicators, this study for the first time, presented a new and robust classifier, called FRF, with of paramount importance application in ungauged catchments. The proposed FRF model is a supervised classifier that uses the global drought information, prepared based on multiple satellite-, and model-based meteorological data sets, for a 1-month ahead classification and prediction of SPEI drought in the CAB, Turkey. The overall performance of the new model was assessed in terms of TA, KA, and CE statistics, and its class precision distilled using the matrix of the outputs. In addition, the classification accuracy of the FRF was cross validated comparing with a baseline model, FDT. The results showed that FRF can perfectly handle the polynomial target vector with specific superiority over FDT in the classification of ED events. Despite dealing with the highly stochastic features, the maximum overall misclassification on the validation set was approximately 28% with a KA value of 0.65 that indicates the FRF as a substantially desirable classifier.

From the modeling perspective, when a nonlinear relationship is a predominant factor among the data sets, the results showed that FDT fits the data inefficiently. The ensemble FRF algorithm combines multiple trees through the bagging, i.e., bootstrapped aggregation, to increase the accuracy of FDT. It is worth mentioning that the trees themselves are independent. The present study was limited to a 1-month ahead classification of the SPEI-6 from a single global drought repository. The design of drought classification models with higher lead times would be beneficial for planning mitigation strategies. To diminish the inherent uncertainties available in global data, the use of drought indices from other global drought repositories such as the Global Integrated Drought Monitoring and Prediction System (Hao et al. 2014) and Climate of the Carpathian region project-CARPATCLIM (2019) would be instrumental in reducing drought impacts, particularly in ungauged catchments.

Compliance with Ethical Standards

Conflict of interest None.

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

References

Abbasi, A., Khalili, K., Behmanesh, J., & Shirzad, A. (2019). Drought monitoring and prediction using SPEI index and gene expression programming model in the west of Urmia Lake. *Theoretical and Applied Climatology, 138*(1–2), 553–567.

Abdourahamane, Z. S., & Acar, R. (2019). Fuzzy rule-based forecast of meteorological drought in Western Niger. *Theoretical and Applied Climatology, 135,* 157–168.

Alsafadi, K., Mohammed, S. A., Ayugi, B., Sharaf, M., & Harsányi, E. (2020). Spatial–temporal evolution of drought characteristics over hungary between 1961 and 2010. *Pure and Applied Geophysics, 177*, 3961–3978.

Belayneh, A., & Adamowski, J. (2012). Standard precipitation index drought forecasting using neural networks, wavelet neural networks, and support vector regression. *Applied computational intelligence and soft computing, 2012,* 6.

Bhalme, H. N., & Mooley, D. A. (1980). Large-scale droughts/floods and monsoon circulation. *Monthly Weather Review, 108*(8), 1197–1211.

Barua, S., Ng, A. W. M., & Perera, B. J. C. (2012). Artificial neural network–based drought forecasting using a nonlinear aggregated drought index. *Journal of Hydrologic Engineering, 17*(12), 1408–1413.

Bogardi, I., Bardossy, A., Duckstein, L., & Pongracz, R. (2004). Fuzzy logic in hydrology and water resources. In *Fuzzy logic in geology* (pp. 153-VIII). Academic Press.

Breiman, L. (2001). Random forests. *Machine learning, 45*(1), 5–32.

CARPATCLIM. (2019). Climate of the Carpathian region project. https://www.carpatclim-eu.org/pages/about/. Accessed May 2020

Caloiero, T., & Veltri, S. (2019). Drought assessment in the Sardinia Region (Italy) during 1922–2011 using the standardized precipitation index. *Pure and Applied Geophysics, 176*(2), 925–935.

Carr, D., & Shearer, J. (2007). *Nonlinear Control and Decision Making Using Fuzzy Logic in Logix*. Rockwell Automation, Inc.

Chen, J., Li, M., & Wang, W. (2012). Statistical uncertainty estimation using random forests and its application to drought forecast. *Mathematical Problems in Engineering*, 2012.

Chiang, J. L., & Tsai, Y. S. (2013). Reservoir drought prediction using two-stage SVM. In *Applied Mechanics and Materials* (Vol. 284, pp. 1473–1477). Trans Tech Publications Ltd.

Danandeh Mehr, A., & Vaheddoost, B. (2020). Identification of the trends associated with the SPI and SPEI indices across Ankara, Turkey. *Theoretical and Applied Climatology, 139*(3–4), 1531–1542.

Danandeh Mehr, A., Sorman, A. U., Kahya, E., & Hesami Afshar, M. (2020a). Climate change impacts on meteorological drought using SPI and SPEI: Case study of Ankara, Turkey. *Hydrological Sciences Journal, 65*(2), 254–268.

Danandeh Mehr, A. Vaheddoost, B., & Mohammadi B. (2020b). A neuro-annealing model for multi-station drought prediction. *Computers and Geosciences,* 104622. https://doi.org/10.1016/j.cageo.2020.104622.

Danandeh Mehr, A., Safari, M. J. S., & Nourani, V. (2021). Wavelet Packet-Genetic Programming: A New Model for Meteorological Drought Hindcasting. *Teknik Dergi*, 32(4).

Dietz, S. J., Kneringer, P., Mayr, G. J., & Zeileis, A. (2019). Correction to: Forecasting Low-Visibility Procedure States with Tree-Based Statistical Methods. *Pure and Applied Geophysics, 176*(6), 2645–2658.

Deo, R. C., Tiwari, M. K., Adamowski, J. F., & Quilty, J. M. (2017). Forecasting effective drought index using a wavelet extreme learning machine (W-ELM) model. *Stochastic Environmental Research and Risk Assessment, 31*(5), 1211–1240.

Farokhnia, A., Morid, S., & Byun, H. R. (2011). Application of global SST and SLP data for drought forecasting on Tehran plain using data mining and ANFIS techniques. *Theoretical and Applied Climatology, 104,* 71–81.

Friedman, J. H. (2001). Greedy function approximation: a gradient boosting machine. *Annals of statistics*, 1189–1232.

Fung, K. F., Huang, Y. F., Koo, C. H., & Soh, Y. W. (2019). Drought forecasting: A review of modelling approaches 2007–2017. *Journal of Water and Climate Change*. jwc2019236

Ghorbani, M. A., Deo, R. C., Kim, S., Kashani, M. H., Karimi, V., & Izadkhah, M. (2020). Development and evaluation of the cascade correlation neural network and the random forest models for river stage and river flow prediction in Australia. *Soft Computing, 24,* 12079–12090.

Hao, Z., AghaKouchak, A., Nakhjiri, N., & Farahmand, A. (2014). Global integrated drought monitoring and prediction system. *Scientific Data, 1,* 140001.

Huang, S., Chang, J., Leng, G., & Huang, Q. (2015). Integrated index for drought assessment based on variable fuzzy set theory: a case study in the Yellow River Basin, China. *Journal of Hydrology, 527,* 608–618.

Keskin, M. E., Terzi, O., Taylan, E. D., & Küçükyaman, D. (2009). Meteorological drought analysis using data-driven models for the Lakes District, Turkey. *Hydrological sciences journal, 54*(6), 1114–1124.

Kisi, O., Gorgij, A. D., Zounemat-Kermani, M., Mahdavi-Meymand, A., & Kim, S. (2019). Drought forecasting using novel heuristic methods in a semi-arid environment. *Journal of Hydrology, 578,* 124053.

Landis, J. R., & Koch, G. G. (1977). An application of hierarchical kappa-type statistics in the assessment of majority agreement among multiple observers. *Biometrics*, 363–374.

Maghrebi, M., Noori, R., Bhattarai, R., Mundher Yaseen, Z., Tang, Q., Al-Ansari, N., & Torabi Haghighi, A. (2020). Iran's Agriculture in the Anthropocene. *Earth's Future*, *8*(9), e2020EF001547.

McKee, T.B., Doesken, N.J., & Kleist, J. (1993). The relationship of drought frequency and duration to time scales. In *Proceedings of the International 8th Conference on Applied Climatology*. American Meteorological Society, Anaheim, CA, USA, 17–22 January. pp. 179–184.

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

Mehr, A. D., Kahya, E., & Özger, M. (2014). A gene–wavelet model for long lead time drought forecasting. *Journal of Hydrology, 517,* 691–699.

Mishra, A. K., Desai, V. R., & Singh, V. P. (2007). Drought forecasting using a hybrid stochastic and neural network model. *Journal of Hydrologic Engineering, 12*(6), 626–638.

Mokhtarzad, M., Eskandari, F., Vanjani, N. J., & Arabasadi, A. (2017). Drought forecasting by ANN, ANFIS, and SVM and comparison of the models. *Environmental Earth Sciences, 76*(21), 729.

Morid, S., Smakhtin, V., & Bagherzadeh, K. (2007). Drought forecasting using artificial neural networks and time series of drought indices. *International Journal of Climatology: A Journal of the Royal Meteorological Society, 27*(15), 2103–2111.

Nguyen, L. B., Li, Q. F., Ngoc, T. A., & Hiramatsu, K. (2015). Adaptive neuro-fuzzy inference system for drought forecasting in the Cai River basin in Vietnam. *Journal of the Faculty of Agriculture, Kyushu University, 60,* 405–415.

Nair, A., Singh, G., & Mohanty, U. C. (2018). Prediction of monthly summer monsoon rainfall using global climate models through artificial neural network technique. *Pure and Applied Geophysics, 175*(1), 403–419.

Nourani, V., & Molajou, A. (2017). Application of a hybrid association rules/decision tree model for drought monitoring. *Global and Planetary Change, 159,* 37–45.

Ozger, M., Mishra, A. K., & Singh, V. P. (2011). Estimating palmer drought severity index using a wavelet fuzzy logic model based on meteorological variables. *International Journal of Climatology, 31*(13), 2021–2032.

Özger, M., Mishra, A. K., & Singh, V. P. (2012). Long lead time drought forecasting using a wavelet and fuzzy logic combination model: a case study in Texas. *Journal of Hydrometeorology, 13*(1), 284–297.

Palmer, W.C. (1965). Meteorological Drought, Weather Bureau Research Paper No. 45, U.S. Department of Commerce, Washington, D.C.

Park, S., Im, J., Jang, E., & Rhee, J. (2016). Drought assessment and monitoring through blending of multi-sensor indices using machine learning approaches for different climate regions. *Agricultural and Forest Meteorology*, *216*, 157–169.

Park, S., Seo, E., Kang, D., Im, J., & Lee, M. I. (2018). Prediction of drought on pentad scale using remote sensing data and MJO index through random forest over East Asia. *Remote Sensing, 10*(11), 1811.

Pesti, G., Shrestha, B. P., Duckstein, L., & Bogardi, I. (1996). A fuzzy rule-based approach to drought assessment. *Water Resourches Research, 32*(6), 1741–1747.

Pongracz, R., Bogardi, I., & Duckstein, L. (1999). Application of fuzzy rule-based modelling technique to regional drought. *Journal of Hydrology, 224*(3–4), 100–114.

Rezaeian-Zadeh, M., & Tabari, H. (2012). MLP-based drought forecasting in different climatic regions. *Theoretical and Applied Climatology, 109*(3–4), 407–414.

Ross, T. J. (2004). *Fuzzy logic with engineering applications*. New York: Wiley.

Shirmohammadi, B., Moradi, H., Moosavi, V., Semiromi, M. T., & 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. *Natural Hazards, 69,* 389–402.

Tayfur, G. (2014). *Soft computing in water resources engineering: Artificial neural networks, fuzzy logic and genetic algorithms*. New York: WIT Press.

Tyralis, H., Papacharalampous, G., & Langousis, A. (2019). A brief review of random forests for water scientists and practitioners and their recent history in water resources. *Water, 11*(5), 910.

Vicente-Serrano, S. M., Beguería, S., & López-Moreno, J. I. (2010). A multiscalar drought index sensitive to global warming: The standardized precipitation evapotranspiration index. *Journal of Climate, 23*(7), 1696–1718.

Zadeh, L. A. (1965). Fuzzy sets. *Information and Control, 8*(3), 338–353.

(Received May 21, 2020, revised October 9, 2020, accepted October 16, 2020)