International Journal of Advanced Science and Computer Applications (2025); 4(1):
DOI: 10.47679/ijasca.v4i1.63
https://ijasca.org/index.php/ijasca

RESEARCH ARTICLE

A New Agricultural Drought Index to Characterize Agricultural Drought Using Data Mining Techniques

Shubhangi Wankhede [1*)]; Leisa Armstrong [2]; Amiya Kumar Tripathy [3]

Published online: 21 March 2025

Abstract

Drought monitoring is a critical task as its occurrence and extent vary according to many factors like drought type, risk, agricultural losses, and impact. Monitoring drought is important because the footprint of this hazard is larger than that of other natural hazards. Many drought indices are developed to monitor complex drought conditions. The intensity and severity of drought in a particular region and at a particular time can be tracked by the drought indicator. In this research, a new agricultural drought index, Yield-Evapotranspiration Drought Index (YEDI) is developed using crop yield, potential, and reference crop evapotranspiration. Data mining and Neural Network techniques have been used to model the drought index. The agricultural and climatic data used is selected from the year 1983 to 2015 (33 years) from the period of June to October (Kharif period) for Maharashtra state in India. The drought index generates the positive values which are further divided into a range of high, medium, and low intensities of drought. SPI and SPEI indices are used for validation against YEDI. Results show that there is a correlation between YEDI and SPEI whereas a low correlation is between YEDI and SPI. YEDI proves to be useful for agricultural drought monitoring.

Keyword: Agricultural drought; drought index; monitoring; data mining; neural network; prediction

Introduction

The drought condition occurs due to deficiency of moisture, and it harms vegetation, soil, animals, and human beings over an area [1]. Drought is considered both sometimes hazard and sometimes a disaster per the situation. It is called a hazard because it is an unpredictable natural occurrence and a disaster because it is a condition causing interference in the water supply to all natural, agricultural and human activities [2]. Previous studies mention that certain human activities and the effect of global warming cause drought events to occur [3]. Drought characterization is a complex process that uses a wide range of available indices for its characterization. It can be characterized by its severity (intensity), duration and geographic extent. Indices are quantitative measures to characterize drought with the help of data from one or many variables called indicators (indices) [4]. The drought index is an important factor for drought monitoring and quantitative evaluation of drought in terms of intensity, spatial extent and frequency [5].

In order to assess droughts and create methods to mitigate them, it is necessary to comprehend the associated drought indices for specific kinds [6]. Considering the features and characteristics of the type of drought, the appropriate drought index needs to be chosen [5]. According to [6], indicators provide a clear picture of the occurrence of drought in a certain area at a specific moment. According to [7], the indices are used to monitor the onset, duration, and regional and temporal patterns of drought.

The kind of water shortage in the specific geographic area that will be examined will be taken into account while developing an index [6]. Precipitation, or rainfall, in the study area is a significant variable for meteorological droughts; stream flow and reservoir data are significant variables for hydrological droughts. Crop output and soil moisture are important factors in agricultural drought [5]. An index must meet the following requirements: (i) have an accurate temporal range; (ii) be a quantitative indication of conditions of drought; and (iv) have an authentic historical record that can be computed or presented, (iii) An index ought to be appropriately applied to the subject under investigation, (v) An index ought to be computed in close.

A more thorough evaluation of drought conditions at the appropriate time and location can be obtained by combining the several drought indices rather than relying just on one [7]. Some characteristics of drought indices are as follows: (i)Time: It could take a few weeks or a few years. (ii) A build-up of water deficits below the threshold level occurs during droughts. (iii) Intensity is the drought's length divided by its severity (iv). The degree of precipitation absence is measured by severity (v). Coverage by Region: This report on the drought is aerial (vi). The return phase of the drought is defined as the average interval between drought incidents, or frequency [4].

The complexity of drought occurrences, along with their variability in terms of location, meteorological circumstances,

Department of Computer Science University of Mumbai and CKT ACS College, New Panvel (Autonomous) Mumbai, India

*) corresponding author

Manna Dey

Email: shubhangi.ak4@gmail.com

Received 21 November 2024 | Received in revised form 21 Januari 2025 | Accepted 16 February 2025
Available online 21 March 2025

and data quality, means that index performance would likely fluctuate [8]. A drought indicator can be made up of a single element or a combination of various different kinds of data. It requires doing some computations [9]. There are about 150 documented drought indices, and more than 50 of these are widely used in numerous nations [4, 10, 11]. The application of drought indices to assess the severity of drought has been the subject of numerous studies.

For example, the drought indices SPI, SSFI, and STI, which have input parameters including precipitation, thermometric data, and watershed observation records, were employed in a study [12]. The results showed that the Mellah watershed (in northeastern Algeria) has years of gaps, and that the intensity, duration, and frequency of the dry spells provide a more precise location for them.

The finding provide a valuable resources for assessing the risks associated with climatic variability. SPI and CDI were the indicators utilized in another study conducted by [13]. Crop water shortage and potential crop yield were the input parameters. The study investigated the effect of meteorological dryness on crop water deficits and agricultural yield reduction in different agro-climatic zones of Poland. [14] employed a modified composite drought index (MCDI) including temperature, evaporation, precipitation, and surface water content as input factors. China's Hubei Province's drought conditions were measured using the multivariate drought index, or MCDI. There was a high correlation discovered when comparing the observations of drought occurrences with the meteorological drought composite index (CI).

[15] studied input factors such temperature, precipitation, GIMMS AVHRR NDVI, soil moisture, and VCI were compared with the drought conditions identified by the various drought indices utilized (PCI, TCI, SMCI, VCI, CMDI, and CVDI). The conditions found and a comparison of several indices are displayed in the results. Crop production, precipitation, maximum and minimum temperatures, and water field capacity were the variables used in another study by [16] that employed the index SPEI. The findings showed a significant agricultural yield response to the duration of the drought. The research findings indicate that fluctuations in the response of agricultural yield to drought are governed by meteorological factors.

[17] showed in another study that alternative drought indicators, like CZI, MCZI, and Statistical Z Score, can take the place of SPI, which depends on rainfall as an input parameter.

[18] compares the MSDI to the data from the U.S. Drought Monitor (USDM) and the standardized indices for drought monitoring in another study. The results showed that MSDI possesses characteristics such as an increased probability of detecting drought. Comparative analysis between SPI and PHDI were conducted in [19] paper. The results demonstrated a reduction in the frequency and duration of drought episodes upon switching from 3- and 6-month SPI to PHDI, with the majority of hydrological droughts following meteorological droughts.

The research published in [8] evaluates how well the six drought indices performed in identifying previous instances of drought in the Upper Blue Nile Basin. According to the results, SPI and SPEI reported drought situations at an early stage. The ETDI, SMDI, and SRI-3 indicate that two of the droughts had early onsets, whereas one had a late start. A study by ADI found that one drought started early and two had late onset. An integrated strategy that considered crop phenology and drought development was necessary for an efficient drought assessment, as [20] discovered in another study. This led to the development of a framework for evolution process-based multi-sensor collaboration (EPMC). The findings demonstrated a strong link between ADI and wheat yield loss in an integrated model.

The MASH drought indicator was utilized in [21] study. Using thirteen different drought indicators, it tried to construct a generic drought index that considers the effects of drought on meteorological, agricultural, hydrological, and stream health. In a different study, [22] analyzed nine drought indicators (RDI, Normal SPI, Gamma SPI, Log SPI, CZI, MCZI, RAI, PN, and DI) and assessed how well they performed in terms of forecasting and tracking drought over the Middle Euphrates region using rainfall and temperature as input factors. The effectiveness of commonly used drought indicators and susceptibility criteria in predicting yearly drought in European macro regions was also investigated in a study by [23]. It was discovered that the best indicators of potential dangers were the 12-month SPEI as well as the sector- and macro region-specific sensitivity of drought indicators.

By analyzing the relationship between soil moisture and drought, [24] shows how to predict agricultural dryness in the Xiangjiang River basin in another study. The findings demonstrated that SPEI-6 significantly more correctly captures soil moisture than SPEI-3 and SPEI-1. Three parameters are analyzed in [25] research to determine the extent of the drought in the Cilacap Regency. There is no difference in the methods used to estimate the severity of drought, as the comparison between SPI and SPEI shows. In [26], stream flow measurements were utilized to analyze the drought in Turkey's Küçük Menderes Basin and served as an input parameter for SSI and SDI indices.

[27] evaluated the temporal and spatial characteristics of the meteorological drought in the Upper Tana Watershed using the SPI and SPEI indices. Remotely gathered data was used. The study produced new data that can be used for drought forecasting, early warning systems, and the planning and management of water resources within a watershed. [28] used the two indices (SDI and SRI) to investigate multi-time step hydrological dryness. Between the chosen years, both indices' results indicated a mild drought.

[29] conducted a follow-up study to evaluate the suitability of the selected drought indices (SPI, EDI, Statistical Z-Score, CZI, RD, and RDDI) in the central Indian Ken River Basin. [30] conducted an evaluation of drought features in the Bundelkhand region of India using a regression discontinuity approach. It was found that RDI-based drought studies work best in semi-arid areas. [31] conducted a study that assessed the state of drought in the Luanhe River Basin by utilizing the SPI and SPEI indices. The findings indicate a relationship between the drought chronology and the output of different drought indices. Additionally, it showed that when evaporation and precipitation are taken into account, SPEI outperforms SPI in accuracy.

Various studies that use indices are outlined in this section. Materials and techniques utilized in this study are described in the following section. The study's results, insights, and recommendations are presented in the ensuing parts.

Method

Study Area and Dataset

India's Maharashtra state serves as the study area. Situated in the western peninsula of the country, it covers a land area of 3,07,713 sq km, which accounts for 9.36% of the entire size of the country. The state is situated between latitudes

15°35'N and 22°02'N and longitudes 72°36'E to 80°54'E. There is an annual temperature range of 25°C to 27°C and an annual rainfall range of 400 mm to 6,000 mm, [32] Thirty-five districts comprise the state. According to [33], around 70% of the sown land is vulnerable to drought. Every year, the rainy season starts in June and lasts until October. Maharashtra's main occupation is agriculture [34].

The datasets used for this study contained three parameters: reference crop evapotranspiration (RCE), crop yield of jowar (Sorghum spp.), and potential evapotranspiration (PE). These datasets came from a range of sources, such as official and public datasets. The Maharashtra states' jowar crop output statistics were provided by the Commissioner of Agriculture, Government of Maharashtra, Pune; potential and reference crop evapotranspiration data were gathered from publicly available web sources.

The chosen parameters were collected during a 33-year period. Depending on the availability, the years ranged from 1983 to 2015. 25 district data were gathered. For the Kharif season (June to October) for each year, the jowar crop production statistics (tonnes per hectare), potential evapotranspiration (mm), and reference crop evapotranspiration (mm) were calculated and saved in a numeric format. The data is preprocessed in accordance with the intended usage after data collection.

Methodology

Parameters for the development of YEDI: Agricultural drought is a slow natural hazard that begins with a precipitation deficit. Other factors such as temperature, soil moisture, potential and reference crop evapotranspiration also can influence the drought situation [35]. Precipitation is considered the major parameter for drought occurrence conditions in most situations. The other two important variables are potential evapotranspiration- an amount of evaporation that would occur in the presence of sufficient water availability and reference crop evapotranspiration- which represents the evapotranspiration rate of green vegetation. It is a combination of normal evaporation and transpiration on the earth's surface. These are crucial elements because they demonstrate how mass and energy are exchanged within the soil, water, and vegetation system [36]. These two parameters are associated with agricultural drought conditions and hence are used in this study to develop a simple equation that defines the drought condition.

This study attempts to track the drought situation in Maharashtra, where the Jowar crop is one of the most important Kharif crops [34]. The jowar crop production, which was expressed in tonnes per hectare, was the third factor taken into account when creating a drought index. To more accurately anticipate and track drought conditions, a new agricultural drought index has been created using the combination of these three characteristics.

Bayesian Network: The Bayesian theorem is the foundation of Bayesian classifiers. According to [39], it forecasts class membership probabilities. It determines the likelihood that the provided data instance belongs to the given class [39]. It is a method for supervised classification [40]. Using a probabilistic method, it is used to simulate a prediction problem [39]. According to this method [39, 40], attributes' probabilities are determined by presuming that each attribute's likelihood on a particular class value is unrelated to the other attributes. The classifier gets superior outcomes as a result of this presumption. Conditional independence is what is meant by this [39, 40]. The Naïve Bayes theorem is represented as:

$$\Pi(\chi|\xi)=\Pi(\xi|\chi)\Pi(\chi)/\Pi(\xi) \quad (1)$$

P(c|x) is the posterior probability, which is determined using P(C), the class prior probability, P(x), the predictor prior probability, and P(x|c), the likelihood according to [40]. On the prepared datasets, Bayes net and Naive Bayes classifiers were employed. Weka, a tool that can conduct preprocessing, classification, visualization, and feature selection, was employed for this purpose.

For the training and testing phases of the data analysis, the data were divided into training sets and test sets. The Bayesian network was built using training data, which also involved building the network topology and probabilities connecting network nodes [41]. The impact class was divided into high, medium, and low categories based on ranges produced by drought index estimates. On the train data set, the Bayes Net and Naive Bayes classifiers were run, and the model they produced was saved and later used on the test data set. To determine the best classifier for the current data set, the performances of the two classifiers were compared.

Artificial Neural Network: Neurons are found in neural networks at various layers. Every neuron of one layer is linked to every other layer's neuron by synapses. The neural network is made up of a collection of connected input and output units, and each connection has a corresponding weight. In order to predict the correct class label for a particular data instance, the network's weights can be changed throughout the learning phase [42]. The data values produced by the drought index in this study were modeled using Multi-Layer Perceptron (MLP), Sequential Minimal Optimization (SMO), and Radial Basis Function (RBF).

Decision Trees: Data mining uses classification algorithms like decision trees. The decision tree creates a top-down tree structure based on the characteristics of the supplied data sets. According to the previously created model, predictive modelling is utilized to classify the given data items [43]. Models for decision tree-based classifiers take into account a number of factors, including computational overheads used, highlights, productivity, and exactness, and then produce results [44]. Root nodes, internal nodes, and terminal nodes are all parts of the tree structure. The Root is a top node that is not singular and has one or more outward edges but no incoming edges. One incoming edge and one or more outgoing edges make up an internal node. Every internal node indicates a test of an attribute, and every edge shows the test's outcome. The final proposed or projected attribute (label) of a data object is represented by the terminal or leaf node [45]. Tree building and pruning are the two stages of the decision tree categorization process. Building trees is done from the top down. The tree is recursively partitioned at this point until all of the data elements have the same class label. As the training data set must be constantly reprocessed, this stage requires numerous computation cycles. Tree pruning is carried out from the ground up. By reducing over fitting, tree pruning reduces over fitting and increases prediction and classification accuracy outcomes of the algorithm [43].

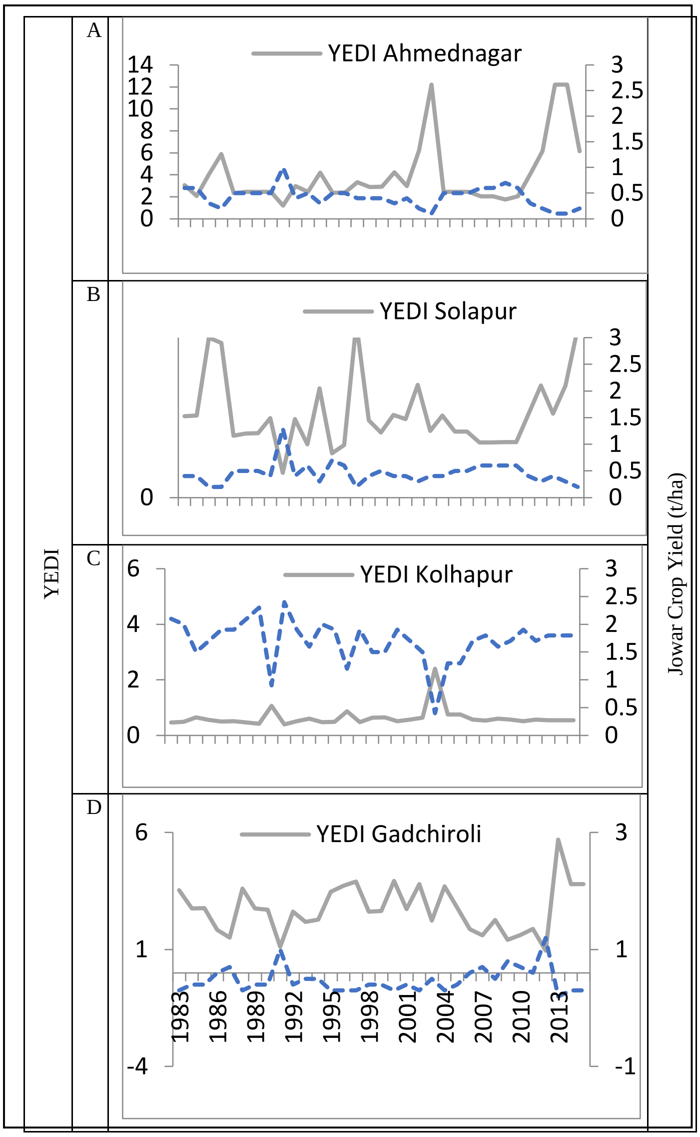

Fig. 1. Relation between YEDI and JCY for Dry Districts A)Ahmednagar, B) Solapur and Wet Districts C) Kolhapur and D) Gadchiroli

Validation of YEDI: Validation by Using Crop Yield Data-In any given region, agricultural dryness results in low crop productivity. Its impact might be long-lasting and persistent, contributing to lower crop harvest yields in the years that follow. Depending on how severe and intense the initial drought was, this affect varies [36]. The YEDI was validated using the association between the Jowar crop yield and the YEDI. The relationship between YEDI and Jowar crop yield for both wet and dry districts and years is depicted in Fig. 1.

Validation by comparison with Standard Precipitation Index (SPI) and Standard Precipitation Evapotranspiration Index (SPEI): Using historical long-term precipitation data, SPI calculates the variation in precipitation for a given time period. It is feasible to identify the type of drought due to its different time scales (1, 3, 6, etc.) [36]. SPI may be computed during the necessary period of interest at various time scales and is straightforward and adaptable. Since rainfall is the primary parameter in this study and its conditions determine when a drought occurs, SPI was employed. Using precipitation data for the chosen years and districts in this study, the 12-month SPI was computed.

SPEI and SPI are comparable, but SPEI takes into account the variation in water supply—rainfall—and demand—potential evapotranspiration. The influence of other related elements including humidity, air temperature, solar radiation, and wind speed is also considered because potential evapotranspiration is taken into account [37].

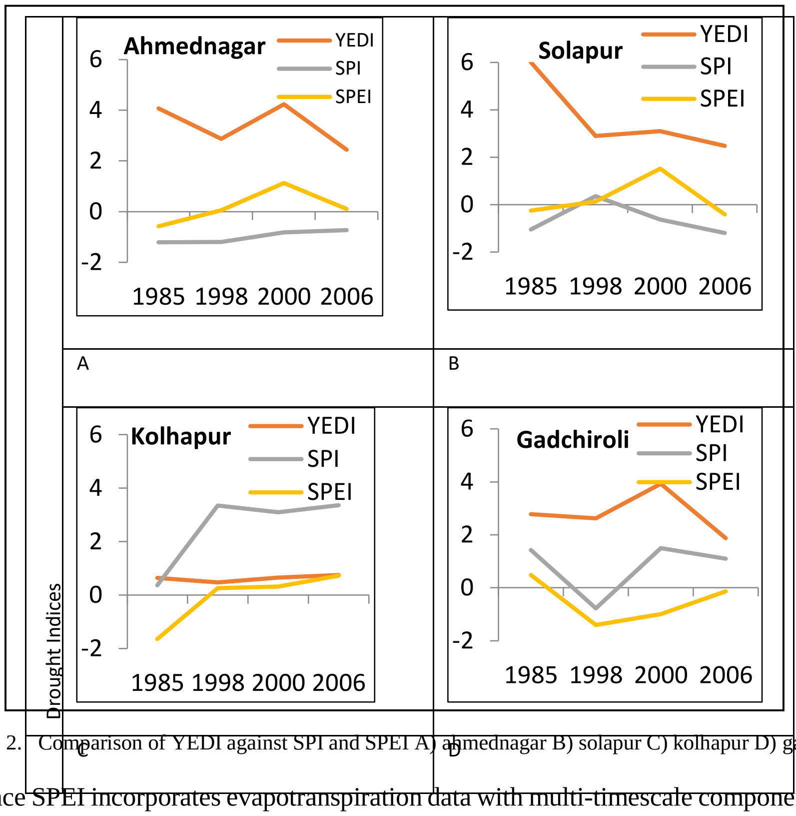

Fig. 2. Comparison of YEDI against SPI and SPEI A) ahmednagar B) solapur C) kolhapur D) gadchiroli

Since SPEI incorporates evapotranspiration data with multi-timescale components of SPI, it is thought to be a more realistic drought index [38]. SPEI is computed using temperature, evapotranspiration, and precipitation data over a 12-month period. Next, the computed outcomes of SPI and SPEI were contrasted with YEDI. The comparison of YEDI with SPI and SPEI is displayed in Fig. 2.

YEDI Calculation using the selected parameters: The RCE and PE parameters along with the JCY were chosen for the development of a new agricultural drought indicator. Initially, during pre-processing, the data gaps were filled by averaging the closest values. The JCY production and area data were available from which tonnes per hectare were calculated for each of the 25 districts. For June through October, the averages of PE and RCE were calculated. Graphs were used to visualize the general characteristics of the data sets and identify the trends. To determine the yield behavior, the JCY production data was employed. The values produced by the drought equation developed through this research were further calculated using the final data sheet that had been collected. The positive numbers produced by this new drought formula were then divided into ranges to determine the severity of the drought.

The intensity levels were calculated by developing an equation via a hybrid methodology. The resulting values were all positive and ranged from 0 to 12. Three intervals were used to categorize the range's three levels of drought intensity: high, medium, and low. The drought formula's ranges and these positive values were combined with the other chosen factors. Future effect values were characterized and predicted using data mining and neural network approaches. The columns for PE, RCE, JCY, drought index, and impact were included in the final Excel file for analysis. Weka tool was used to do this modeling for Bayes Net, Naïve Bayes, J48, LADTree, Random Tree, MLP, RBF and SMO techniques.

Results and Discussion

Graph-Based Visualisation Results

The findings published in the previous study served as the foundation for the findings reported in this paper. The dry and wet years from 1983 to 2015 as well as the dry and wet districts chosen from 25 districts served as the basis for these findings. Districts and two of each dry and wet year are selected for additional research. The findings indicate that 1988 and 2000 are the dry years whereas 1998 and 2006 are the wet years. Kolhapur and Gadchiroli were considered to be the wet districts,and Ahmednagar and Solapur to be the dry districts.

Data on rainfall and Jowar crop yield were analyzed and correlated to make these selections. These years and districts are taken into consideration for additional study.

The impact of PE and RCE on JCY in the Maharashtra divisions is displayed in Fig. 3. RCE is between 3 and 6 mm, whereas PE is between 4 and 7 mm. It was observed that when both PE and RCE were high, the JCY was high for the divisions of Aurangabad and Nasik. The JCY was high in the divisions of Pune, Kolhapur, Latur, Amravati, and Nagpur when the PE and RCE were low.

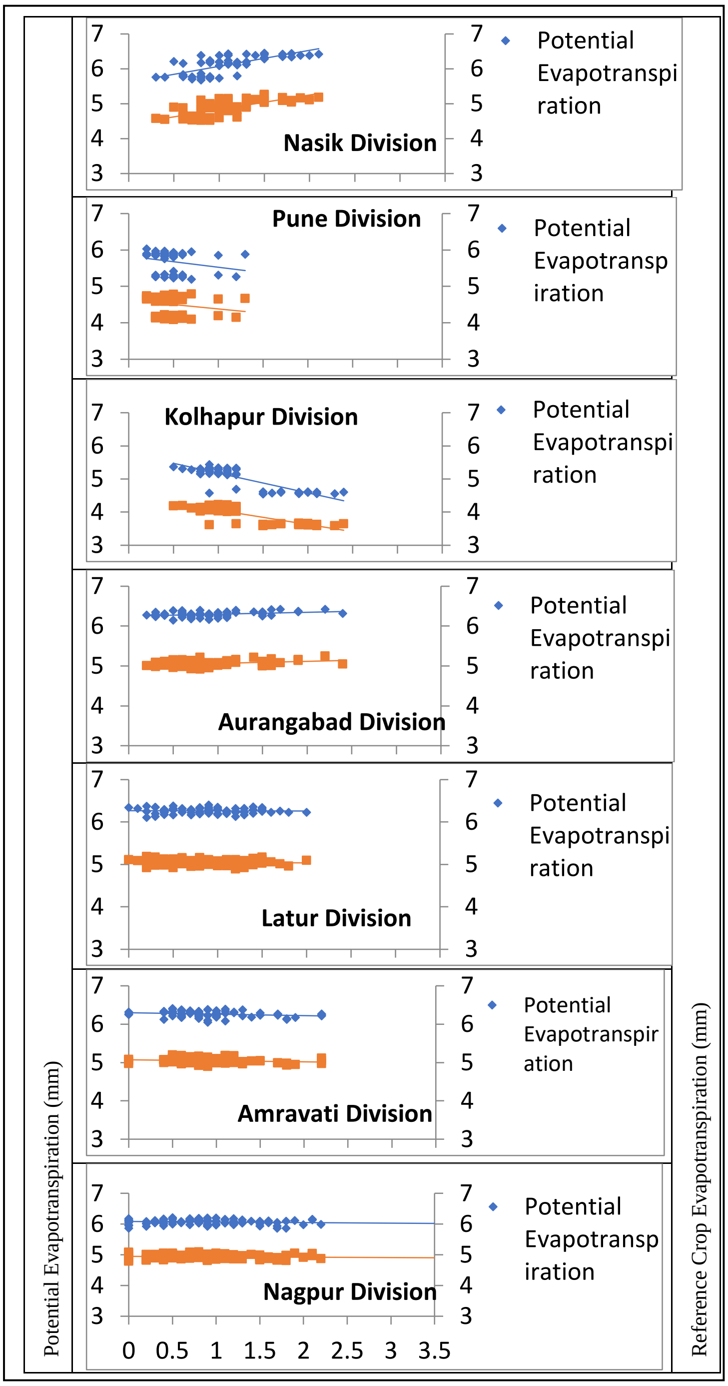

Fig. 3. Effect of PE and RCE on JYC in divisions of Maharashtra state, India

Development of Yield Evapotranspiration Drought Index

The agricultural data, which included JCY and climatic characteristics like PE and RCE, were used to create a novel formula for the agricultural drought index. Potential evapotranspiration is the amount of water that would evaporate if there was enough water available. The rate at which green vegetation evapotranspirates is represented by reference crop evapotranspiration. It is a result of both typical evaporation and transpiration from trees on the surface of the earth. The annual jowar yield, expressed in tons per hectare, is known as the jowar crop yield. The newly developed agricultural drought index is given as equation (2).

$$YEI=(PETi-RCETi)/Yi \qquad (2)$$

Where, YEI=Yield Evapotranspiration Index

PETi=Potential evapotranspiration for selected kharif season of one year

RCETi=Reference crop evapotranspiration for selected kharif season of one year

Yi=Jowar crop yield for the selected kharif season of one year

A range of high, medium, and low drought intensities are produced by this indicator, which also yields positive values. Table 1 specifies the range.

Table 1. Range and drought intensities for the generated drought index.

Sr. No.	Range	Intensity
1	1.8 and above	High
2	1 to 1.8	Medium
3	0 to 0.9	Low

In order to estimate the impact of the drought using the recently developed drought formula, these index-generated values are modeled using a variety of data mining and neural network techniques. The sections that follow provide more information about these.

Agricultural drought Prediction by YEDI

Bayesian Network Results: To simulate the impact of drought severity on jowar yield, two Bayesian network techniques were applied. Using the selected parameters that were available, Bayes Net and Naïve Bayes algorithms were applied. Climate parameters (PE and RCE) and agricultural parameters (JCY) made up the data set. The information was stored in a CSV file. The first Bayesian technique was used with Weka to analyze the dataset and forecast how drought will affect jowar crop productivity. The accuracy of the BayesNet method was 98.09%. The precision was 98 and the sensitivity was 98.71%.

The resulting values for the mean absolute error (MAE), relative absolute error (RAE), root mean square error (RMSE), and R2 were 0.0146, 0.1082, 3.8457, and 24.85%, respectively. Naive Bayes, another Bayesian technique, has an accuracy of 93.21%. Both the precision and sensitivity were 93.21%. Using this method, the MAE was 0.0583, the RMSE was 0.1892, the RAE was 15.3302, and the R2 was 43.47. Table 2 displays a comparison of the two classifiers operating under the Bayesian network. Table 3 presents a comparison of the Bayesian network algorithms' performance metrics.

Table 2. Comparison of Bayesian approach techniques for new drought index

Sr. No.	Classifier Name	Accuracy	Sensitivity	Precision
1	BayesNet	98%	98%	98
2	NaiveBayes	93%	93%	93

Table 3. Comparison of Bayesian approach performance measures for new drought index

Sr. No.	Classifier Name	MAE	RMSE	RAE	RRSE
1	BayesNet	0.0146	0.1082	3.8457%	24.85%
2	NaiveBayes	0.0583	0.1892	15.3302%	43.47%

Neural Network Results: Three neural network techniques were applied to the available data including MLP, RBF and SMO. MLP algorithm was found to have an accuracy of 98.03 with a sensitivity of 98.03% and a precision of 98%. The MAE generated was 0.0352, RMSE was 0.1154, RAE was 9.239 and the R2 was 26.52%. Another NN approach was the RBF network provided an accuracy of 96.23% with a sensitivity and precision was 96.3% and 96 respectively. The MAE generated with this technique was 0.0312, RMSE was 0.162, RAE was 8.2116 and the R2 was 37.22. A third NN approach SMO was also tested and produced an accuracy of 92% and sensitivity and precision was 92.6%. The MAE generated with this technique was 0.2444, RMSE was 0.3103, RAE was 64.24 and the root relative squared error was 71.31. A comparison of the three classifiers under NN is shown in Table 4. A comparison of performance measures of neural network algorithms is shown in Table 5.

Table 4. Comparison of NN for new drought index

Sr. No.	Classifier Name	Accuracy	Sensitivity	Precision
1	MLP	98.03%	98.03%	98
2	RBF	96.23%	96.3%	96
3	SMO	92.0%	92.6%	92

Table 5. Comparison of NN approach performance measures for new drought index

Sr. No.	Classifier Name	MAE	RMSE	RAE	R2 Error
1	MLP	0.0352	0.1154	9.239%	26.52%
2	RBF	0.0312	0.162	8.2116%	37.22%
3	SMO	0.2444	0.3103	64.24%	71.31%

Decision Tree Results: Three Decision tree techniques. J48, LADTree and RandomTree classifiers were used in this study. The J48 classifier produced an accuracy of 99% and a sensitivity of 99%. It also showed an RAE of 0.0132, RMSE was 0.0825, RAE was 3.4661%, and R2 of 18.96%. The accuracy and sensitivity of the LADTree were both 98.08%. It also showed an MAE of 0.0076, RMSE was 0.0817, RAE was 2.0016% and R2 of 18.76%. The last Decision tree assessed was the RandomTree classifier which achieved an accuracy of 98.2%and sensitivity is 98.1%. It also showed an MAE of 0.0133, RMSE was 0.1155, RAE was 3.5041% and R2 of 26.53%. Among decision tree classifiers applied to the dataset, it was found that J48 tree reported a high accuracy of 99%. A comparison of three decision tree algorithms is shown in Table 6. The comparison of performance measures of decision tree algorithms is shown in Table 7.

Table 6. Comparison of decision tree algorithms for new drought index

Sr. No.	Classifier Name	Accuracy	Sensitivity
1	J48	99%	99%
2	LADTree	98%	98%
3	RandomTree	98%	98.1%

Table 7. Comparison of decision tree algorithms performance measures for new drought index

Sr. No.	Classifier Name	MAE	RMSE	RAE	RRSE
1	J48	0.0132	0.0825	3.4661%	18.96%
2	LADTree	0.0076	0.0817	2.0016%	18.76%
3	RandomTree	0.0133	0.1155	3.5041%	26.53%

Validation of Results: Validation by comparison with jowar crop yield- To assess the accuracy and compatibility of YEDI, validation was done using district-wise and year-wise jowar crop yield data. Correlation analysis was done for annual jowar crop yield (t/ha) and the YEDI values were calculated for each year and district. Fig 4. shows the correlation of YEDI and jowar crop yield for all the 25 selected districts in this research. The years for which the comparison was done are the selected two dry years (1985 and 2000) and two wet years (1998 and 2006). Fig 6. shows the correlation of YEDI and jowar crop yield from 1983 to 2015 (33 years) for the selected dry districts (Ahmednagar and Solapur) and the wet districts (Kolhapur and Gadchiroli).

Fig 1. indicates that when the crop yield is more the YEDI value is low and when the crop yield is low the YEDI value is high. Table 1 shows that high values of YEDI indicate high drought severity and low values indicate low severity. Based on this information, it can be reported that YEDI has given significant observations for high and low crop production.

Fig 5. Displays Scatter Plots and R2 and P values between JCY and YEDI for dry and wet years. Statistically significant correlations were observed for the dry years 1985 (R2=0.6, P value=0.0008) and 2000 (R2=0.8, P value=0.0001). For wet years the observed correlations are 1998 (R2=0.8, P value=0.0002) and 2006 (R2=0.7, P value=0.0002). The results showed that YEDI and JCY are highly correlated for the dry and wet years. Fig. 6 shows Scatter Plots and R2 and P values between JCY and YEDI for Dry and wet districts. These results also show significant correlations for dry districts Ahmednagar (R2=0.6, P value=0.0001) and Solapur (R2=0.6, P value=0.0002) and wet districts Kolhapur (R2=0.7, P value=0.0004) and Gadchiroli (R2=0.7, P value=0.0005).

The significant correlation values indicate that JCY and YEDI are highly correlated. This evaluation proved that YEDI reflects drought's influence on agricultural growth. The YEDI values also can be used to categorize the drought severity (high, medium, and low), suggesting that YEDI provides a good monitor for agricultural drought conditions in Maharashtra, India.

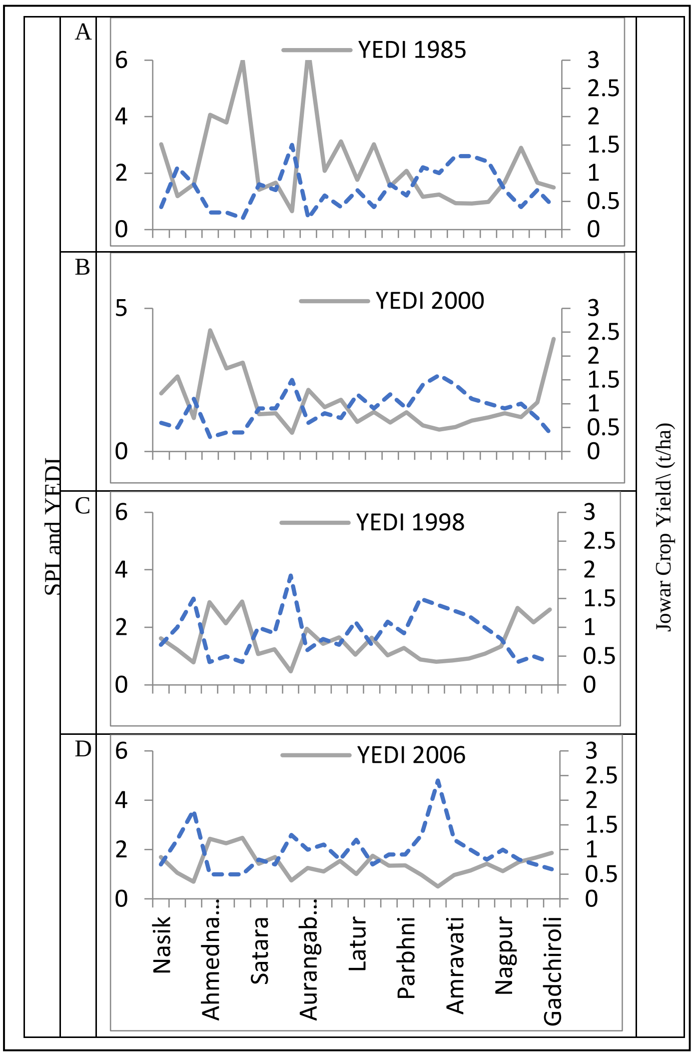

Fig. 4. Relation between YEDI and JYC for Dry Years (a)1985 and (b) 2000 and Wet Years (c) 1998 and (d) 2006

Validation by comparison with SPI and SPEI: An evaluation of YEDI with respect to SPI and SPEI was made using two dry years and districts as well as two wet years and districts as shown in Fig 2. For Ahmednagar district, the correlation of YEDI with SPI is (R2=0.07) and with SPEI it is (R2=0.2). For Solapur district the correlation for SPI is (R2=0.09) and for SPEI (R2=0.1). For Kolhapur district, the correlation with SPI is (R2=0.02) and with SPEI is (R2=0.1) and for Gadchiroli district the SPI correlation is (R2=0.07) and SPEI is R2=0.2. This shows a very low correlation between YEDI and SPI and YEDI and SPEI. The P values also showed statistical insignificance (P>0.05). There is a low correlation between YEDI and SPEI whereas very low correlation between YEDI and SPI. This states that YEDI is not capable of monitoring the meteorological drought (as compared to SPI). As compared to SPEI, YEDI proves to be having very low correlation and it cannot be used to much extent. These overall results state that YEDI along with JCY data can be used to monitor and categorize agricultural drought.

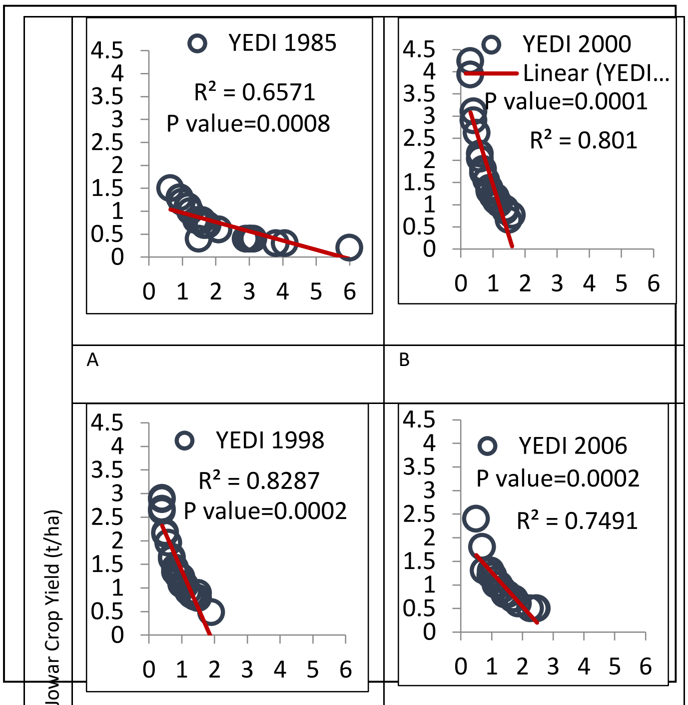

Fig. 5. Relation between YEDI and JYC for Dry Years (a)1985 and (b) 2000 and Wet Years (c) 1998 and (d) 2006

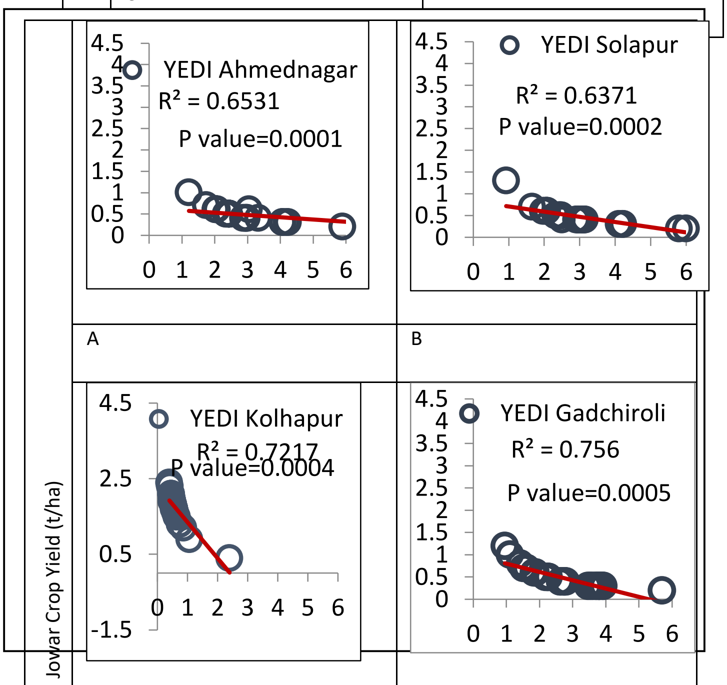

Fig. 6. Relation between YEDI and JCY for Dry Districts A)Ahmednagar and B) Solapur and Wet Districts C) Kolhapur and D) Gadchiroli

Discussion

According to the study's findings, YEDI is a useful indicator for tracking agricultural drought. YEDI's connection with SPI and SPEI was shown to be very low and low, respectively, based on a comparison of the two measures. In Maharashtra, YEDI is doing well in terms of crop productivity. All methods, have advantages and disadvantages. Since the annual drought monitoring method was the primary focus of this study, the time interval needs to be shortened from annually to monthly. The key benefit of YEDI is that it is easy to calculate and produce values that represent the intensity of the drought. It has shown to be highly helpful for Maharashtra in monitoring the drought, and it can be used to other regions in India. Decision-makers in the agriculture industry may find it helpful to utilize this indicator to determine the severity.

Conclusions and Recommendations

The Maharashtra state in India has faced regular drought events. Severe drought occurrences have led to many farmer suicides due to losses in crop productivity. Drought monitoring and early warning systems may help farmers and other stakeholders to improve their cropping activities in drought conditions. Drought indices play an important role in drought prediction. This study presents the development of a new agricultural index YEDI (Yield Evapotranspiration Drought Index). The newly generated formula is easy to apply and the forecasting done by using this index will prove to be useful for planning all agricultural activities. A comparison of YEDI with SPI and SPEI indicated that for the parameters selected for this study, YEDI proves to be more relatable. The positive values generated by YEDI are modeled using various data mining and neural network techniques. The results reported in this study have shown how data mining and neural network techniques can be used to model drought indices. These outcomes can be applied to drought decision support systems which will help in a better farming process. Also, it should not be ignored that agricultural drought is influenced by various factors and no single indicator can detect perfectly how drought will impact the affected area. The methods used in this research can be incorporated in other regions as well to monitor and detect agricultural drought conditions.

References

[1] Jiwan, J. (2012). Converging Information, Communication and Space Technologies in Drought Management in India. International Journal of Social Tomorrow, 1(5).

[2] Salee, W.; Pongsapukdee, V. (2013). Odds Prediction of Drought Category Using Log Linear Models Based on SPI in Northeast of Thailand. Silpakorn U Science and TechJ, 7(1), 32-40.

[3] Kaur, A., Sood, S. (2020) Cloud-Centric IoT-Based Green Framework for Smart Drought Prediction, IEEE IoT Journa. 7(2), 1111-1121.

[4] Zargar, A. (2011). A review of drought indices, Environ. Rev , 19, 333–349.

[5] Mishra, A.K., Singh, V.P. (2011). Drought Modeling-A Review, Journal of Hydrology, 3 (49), 157-175.

[6] Agwata, J. F. (2014). A Review of Some Indices used for Drought Studies, Civil and Environmental Research, 6(2), 14-21.

[7] Sun, L.; Mitcheel, S.; Davidson, A. (2011). Multiple Drought Indices for Agricultural Drought Risk Assessment on Canadian Prairies. Int. Journal of Climatology, Royal Meteorological Society. 32(11), 1628-1639.

[8] Bayissa, Y. (2018). Comparison of the Performance of Six Drought Indices in Characterizing Historical Drought for the Upper Blue Nile Basin, Ethiopia. Geosciences. , 8 (81), 1-26.

[9] Bachmair. (2016). Drought indicators revisited: the need for a wider consideration of environment and society. WIREs Water, 3(4), 516-536.

[10] Svoboda, M.; Fuchs, B.A. (2016). Handbook of Drought Indicators and Indices, World Meteorological Organization (WMO) and Global Water Partnership (GWP). Integrated Drought Management Programme (IDMP), Integrated Drought Management Tools and Guidelines Series, Geneva.

[11] Wang, Y. (2020). Evaluation of the suitability of six drought indices in naturally growing, transitional vegetation zones in Inner Mongolia (China). PLoS ONE, 15(5), 1-15.

[12] Bendjema, L. (2019). Characterization of the climatic drought indices application to the Mellah catchment, North-East of Algeria, Journal of Water and Land Development. 2019(43), 28-40.

[13] Labedzki, L.; Bak, B. (2017). Impact of meteorological drought on crop water deficit and crop yield reduction in Polish agriculture. Journal of Water and Land Development, 2017(34), 181–190.

[14] Chen, S. (2020). Comprehensive Drought Assessment Using a Modified Composite Drought index: A case study in Hubei Province, China. Water, 12(462), 1-12.

[15] Wang, K. (2019). Exploring Drought Conditions in the Three River Headwaters Region from 2002 to 2011 Using Multiple Drought Indices. Water, 11(190), 1-20.

[16] Gallardo, M. (2019). Response of Crop Yield to Different Time scales of Drought in United States: Spatio-Temporal Patterns and Climatic and Environmental Drivers, Agricultural and Forest Meteorology, 264, 40-55.

[17] Mayoor, M. (2018). Comparison of Four Precipitation Based Drought Indices in Marathwada Region of Maharashtra India. International Journal of Advance and Innovative Research, 5(4), 60-70.

[18] Hao, Z.; Aghakouchak, A. (2014). A Nonparametric Multivariate Multi-Index Drought Monitoring Framework. Journal of Hydrometeorology, 15, 89-101.

[19] Shin, J. (2018). Investigation of drought propagation in South Korea using drought index and conditional probability. Terr. Atmos. Ocean. Sci, 29(2), 231-241.

[20] Zhang, X. (2017). Multi-sensor integrated framework and index for agricultural drought monitoring. Remote Sensing of Environment, 188, 141–163.

[21] Esfahanian, E. (2015). Development and evaluation of a comprehensive drought index, Journal of Environmental Management, 185, 31-43.

[22] Alwan, I. (2018). Comparison of Nine Meteorological Drought Indices Over Middle Euphrates Region During Period from 1988 To 2017. International Journal of Engineering & Technology, 7(4), 602-607.

[23] Blauhut, V. (2016). Estimating drought risk across Europe from reported drought impacts, drought indices, and vulnerability factors, Hydrol. Earth Syst. Sci., 20, 2779–2800.

[24] Tian, Y. (2018). Agricultural drought prediction using climate indices based on support vector regression in Xiangjiang river basin, Science of the total environment, 622-623, 710- 720.

[25] Balbo, F. (2019). The evaluation of drought indices: Standard Precipitation Index, Standard Precipitation Evapotranspiration Index, and Palmer Drought Severity Index in Cilacap-Central Java. IOP Conf. Series: Earth and Environmental Science, 303, 1-10.

[26] Kerman C. and Gul G., Comparing Two Streamflow-Based Drought Indices, In Gastescu, Water resources and wetlands, 4th International Conference Water resources and wetlands, Tulcea (Romania), 5-9 September 2018.

[27] Okal, H. (2020). Spatio-temporal characterisation of droughts using selected indices in Upper Tana River watershed, Kenya. Scientific African, 7, 1-12.

[28] Pathak,A.; Dodamani, C. (2016). Comparison of two hydrological drought indices. Perspectives in Science, 8, 626-628.

[29] Jain, V. (2016). Comparison of drought indices for appraisal of drought characteristics in the Ken River Basin, Weather and Climate Extremes, 8, 1–11.

[30] Thomas, T. (2016). Reconnaissance drought index based evaluation of meteorological drought characteristics in Bundelkhand. Procedia Technology, 24, 23- 30.

[31] Wang, K. (2015). Analysis of spatio-temporal evolution of droughts in Luanhe River Basin using different drought indices. Water Science and Engineering, 8.

[32] Report: India State of Forest Report, 2019, 152-162.

[33] Choudhary, K., Tahlani, P., Bisen, P., Saxena, R., Ray, S. (2017) Assessment of drought indicators. Technical report, New Delhi.

[34] NIDM: Maharashtra. In: National Disaster Risk Reduction Portal, Maharashtra, 2012, 1-26.

[35] Liu, X. (2020). A remote sensing and artificial neural network-based integrated agricultural drought index: Index development and applications, Catena, 2020(186), 1-8.

[36] Mamane Barkawi Mansour Badamassi, M. B. M.; El-Aboudi, A.; Gbetkom, P. G. (2020). A New Index to Better Detect and Monitor Agricultural Drought in Niger Using Multisensor Remote Sensing Data. The Professional Geographer, DOI: 10.1080/00330124.2020.1730197, 1730197, 1-12.

[37] Wang, H.; Rogers, J.C.; Munroe, D.K. (2015). Commonly used drought indices as indicators of soil moisture in China. J. Hydrometeorol.

[38] Tian, T. (2018). Evaluation of six indices for monitoring agricultural drought in the southcentral United States, Agricultural and Forest Meteorology, 249(2018), 107-119.

[39] Gayathri, A. (2016). A survey on weather forecasting by data mining. IJARCCE, 5(2), 298–300.

[40] Netti, K.; Radhika, Y. (2016). Minimizing loss of accuracy for seismic hazard prediction using Naive Bayes classifier. IRJET, 3(4), 75–77.

[41] Heaton, J. (2013). Bayesian networks for predictive modeling. Forecast, 6–10.

[42] Ali, Z. (2017). Forecasting Drought Using Multilayer Perceptron Artificial Neural Network Model. Advances in Meteorology, 2017, 1-9.

[43] Hamoud, A. (2018). Predicting Student Performance in Higher Education Institutions Using Decision Tree Analysis, International Journal of Interactive Multimedia and Artificial Intelligence, 5(2), 26-31.

[44] Reddy, A.; Pachouri, S. (2016). Comprehensive Study on Efficient Diabetes Disease Prediction with Using Various Advance Decision Tree Models Algorithms. International Journal of Computer Techniques, 3(5), 57-65.

[45] Amjad, A. S. (2016). Educational Data Mining & Students' Performance Prediction. International Journal of Advanced Computer Science and Applications, 7(2016), 212-220,