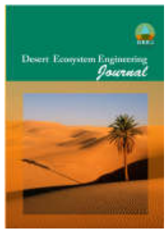
DEEJ (2023) 12 (9) 33-48

University of Kashan

Desert Ecosystem Engineering Journal

Journal homepage: http://deej.kashanu.ac.ir



Assessing Drought Hazard Via Combined Drought Index Using Machine Learning Techniques: A Case Study of Ilam Province

Zahedeh Heidarizadi[1], Majid Ownegh*[2], Chooghi Bairam Komaki[3]

Received: 28/05/2022 Accepted: 09/07/2023

Abstract

As one of the most important natural hazards worldwide, drought increases the vulnerability of the agricultural sector, raises economic loss, and threatens human life, making the characterization of drought and its hazard assessment to be of great significance. Therefore, this study used twelve various remotely sensed indices derived from Moderate Resolution Imaging Spectroradiometer (MODIS) and digital elevation model (DEM) to monitor drought throughout the 2000–2018 growing season. Moreover, the Standardized Precipitation Index (SPI) was used as reference data, with the relevant time scales ranging from 1 to 12 months. Finally, the correlation between thirteen indices and SPI in Ilam Province was modulated using three machine learning approaches, including random forest, boosted regression trees, and Cubist. The results indicated that among the three approaches mentioned above, random forest delivered the best performance (R^2 = 0.88) in terms of SPI prediction. It was also found that Land Surface Temperature (LST) and Evapotranspiration (ET) had higher relative significance in terms of short-term meteorological drought, whereas Normalized Difference Vegetation Index (NDVI) and Soil Adjusted Vegetation Index (SAVI) had higher relative significance in terms of long-term meteorological drought when treated by random forest approach. In the next step, relative soil moisture, Standardized Precipitation Evapotranspiration Index (SPEI), and crop yield data were used to validate the collected data. Finally, the Drought Hazard Index (DHI) was generated based on the probability occurrences of drought using the comprehensive drought model made in the previous step. Accordingly, the results of the DHI map indicated that 65% and 18% of the study area fell under the very high and high classes of drought hazard, respectively. Overall, the results of this study provide a comprehensive method for assessing regional drought.

Keywords: DHI, Ilam province, MODIS.

1. Ph.D. student of Combat Desertification, Gorgan University of Agricultural Science and Natural Resources
2. Professor, Department of Watershed and Arid Zone Management, Gorgan University of Agricultural Sciences and Natural Resources (GUASNR), Corresponding author; m.ownegh@yahoo.com
3. Assistant prof, Department of Watershed and Arid Zone Management, Gorgan University of Agricultural Sciences and Natural Resources (GUASNR
DOI: 10.22052/JDEE.2023.246432.1079

1. Introduction

As one of the most serious natural hazards worldwide, drought causes major threats to the ecological environment, socioeconomic development, and agricultural production (Dai., 2011), having a variety of types, including meteorological, agricultural, hydrologic, and socio-economic ones (Wilhite and Glantz., 1985; Wilhite et al., 2007). From among these types of droughts, meteorological drought is of greater significance, as it could cause all other types of droughts (Nasrollahi et al., 2018).

Droughts are common natural disasters in Iran, bringing about devastating consequences, especially in the country's arid and semi-arid provinces. Moreover, drought-induced crop failure and water shortages are among the frequent problems occurring almost annually in Iran (Nasrollahi, 2015).

Monitoring and predicting drought are considered by local, regional, and international policymakers as proper strategies to reduce the negative influence of drought events, commonly implemented via traditional meteorological monitoring and remote sensing methods (Dai et al., 2011; Wang, 2007). In general, it is necessary to use a reliable drought index to accurately identify drought events and investigate their spatial and temporal changes.

While drought indices are used to investigate drought, they have mainly been developed based on a relevant climate variable. For instance, most drought indices depend on climatic variables such as precipitation, runoff, or soil moisture (Waseem et al., 2015). However, applying a multiple-variable-based drought index (multiple drought index) could help predict drought events more accurately by taking a variety of variables into account. Moreover, a multiple drought index can usually reflect different types of droughts (meteorological, hydrological, agricultural, or socio-economic droughts). On the other hand, single-variable drought indicators are unable to reveal the complex relationship existing between different variables. Therefore, the application of a composite index is suggested as a way to overcome the aforementioned problem.

Combining indices started in the 1990s when many composite indices were produced. Similarly, drought indices that can effectively evaluate and control drought conditions in different climatic regions should be developed by combining satellite-based drought factors and parameters. However, when it comes to combining multiple variables, it is important to determine what weighting method to use. In this regard, equal weighting and linear combination approaches are commonly used due to their simplicity (Kogan, 1993; Rhee et al., 2010; Zhang and Jia, 2013). Nevertheless, considering the fact that the factors affecting drought are dependent on the region, time, and the type of drought, more advanced and consistent weighting methods should be explored to combine several factors so that drought can better be controlled.

Single-variable-based drought indices are unable to provide reliable early warnings on complex drought conditions due to their potential prediction limitations. For instance, SPI is merely based on precipitation, lacking any drought information generated by other variables or sources such as temperature or soil moisture. Therefore, predicting drought based on SPI may not provide sufficient information to use early preventive measures for operational drought management.

Furthermore, various variables such as precipitation, temperature, soil moisture, and relative humidity may contribute (or be related) to drought conditions through different mechanisms, all of which need to be considered in monitoring and forecasting drought events. Therefore, drought prediction can be improved by considering factors such as land surface, feedback among variables, and the correlation between climate indicators, or certain combinations of them (Mishra et al., 2015, Wood et al., 2015).

The composite index requires developing an index based on different drought indices (Rajsekhar, 2015) which can represent multiple-variable drought factors simultaneously. (Huang et al., 2015; Chang et al., 2016). In recent years, many studies have been conducted to integrate various factors to build a multiple-variable drought index/model (AghaKouchak et al., 2015; Park et al., 2016; Yin et al., 2018; Shen et al., 2019; Proodhan et al., 2021).

The traditional regression method is commonly used to construct a comprehensive model, requiring the determination of weights to do so. On the other hand, approaches comprising linear combinations with equal weighting are generally used because of their simplicity (Rhee et al., 2010; Zhang and Jia., 2013). However, more advanced and adaptive weighting approaches should be examined to be used for combining multiple factors so that the drought could be better monitored (Park et al., 2016). Seeking to assess drought hazard, this study, therefore, used machine learning approaches are one of those advanced approaches mentioned above.

Drought hazard assessment can provide solutions to mitigate the adverse consequences of drought. In this regard, many studies conducted recently have considered the significance of drought hazards and risks (Shahid and Behrawan., 2008; Dabanli., 2018; Nasrollahi et al., 2018; Khoshnazar et al., 2021; Lin et al., 2021; Sahana et al., 2021).

According to what was already discussed, it could be argued that assessing and monitoring drought based on a multiple-variable index provides more reliable results than the single-variable indices do. To develop a multiple-variable drought index, many indices should be selected and modified based on the drought type, study area, data availability, and climatic regimes.

Located in western Iran, Ilam province is often suffering from drought due to its arid and semi-arid conditions. Moreover, most parts of Iran lack a sufficient number of meteorological stations to cover a suitable statistical period. The problem is more critical in western Iran, where intense topography has led to the creation of several microclimatic areas with different characteristics in terms of soil, vegetation, and climatic conditions.

In addition, most of the studies conducted on drought in Ilam province have merely used single-factor indices. Considering the fact that few studies have considered a comprehensive drought index comprising the province's health condition of vegetation, temperature, and geographical condition, this study sought to evaluate drought in a specific period using comprehensive indices made up of different variables, attempting to prepare a drought hazard map based on the developed index.

2. Materials and Methods

2.1. Case study

Located in the western Zagros Mountain between 32° 03' to 34° 02' north latitude and 45° 40' to 48° 03' southwest of Iran (Figure 1), Ilam province covers 20106 square kilometers of Iran, which is about 1.2% of the country's total area. The mountainous areas in the north and northwest of the province are relatively cold with long winters, where the minimum temperature in winter reaches -15 degrees, and the rainfall rate is more than 500mm. The western and southwestern regions of the province are hot, with their maximum rainfall rate being about 200mm. Moreover, while the temperature is up to -15 degrees in the middle regions of the province, the maximum rate of temperature goes up to more than 40 degrees in the southern parts during the summer. The amount of precipitation in these areas is estimated as the average rate of precipitation in the two regions mentioned above, varying from 250 to 400mm (Statistical yearbook of Ilam province in 1998).

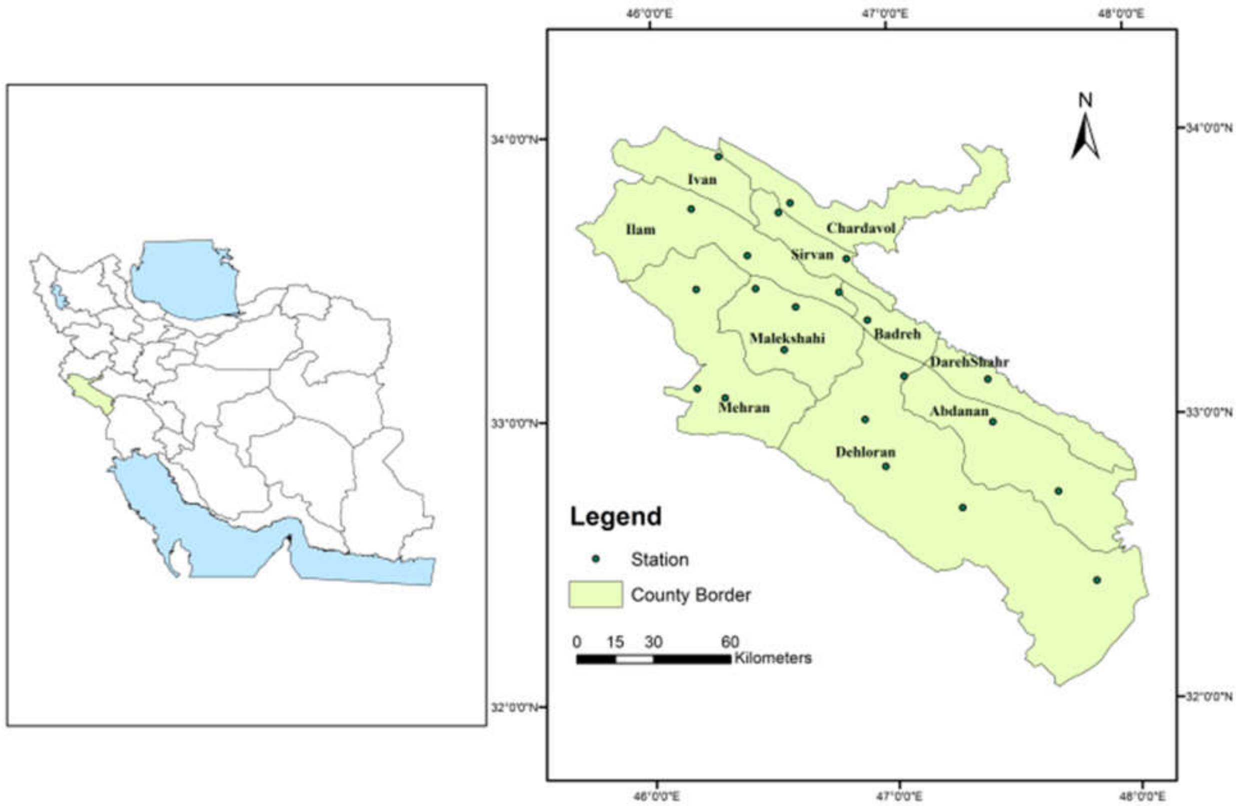


Figure (1): Geographical location of Ilam Province with rainfall gauging stations

2.2. Data

2.2.1. Remote sensing data

This study used 12 drought-based satellite indices that were obtained from atmospherically corrected MODIS on the Terra platform from 2000 to 2018. The MOD16A2 ET, the MOD11A2 LST, and MOD13A3 NDVI data were available with 1 km spatial resolution and the MOD09A1 surface reflectance data were available with 500 m spatial resolution (modis.gsfc.nasa.gov), all of which were used for the growing season (March to June). The MOD11A2 and the MOD09A1 surface reflectance provide 8-day composite data, which were converted into monthly data to fit the purposes of this study.

2.2.2. Reference data

The current study used the SPI as reference data to monitor drought. The World Meteorological Organization has recommended SPI as a standard for describing meteorological droughts (McKee, 1993; Dutra et al., 2013). To calculate the SPI index, there must be no missing data in the time series, and the statistical period should be at least 30 years (Svoboda, 2000). In this regard, this study used total monthly precipitation data collected from 22 stations (Table 1, as provided by the Meteorological Department of Ilam Province) for the 1988-2018 period to calculate SPI using DrinC software with different time scales (e.g., 1-, 3-, 6-, 9- and 12-month SPI data) to consider the lag time effect between precipitation and drought (Park et al., 2016). It should be noted that the SPI values were considered as reference data.

Table (1): Geographical location of the stations whose data were used in this study

Station	Longitude	latitude	Height above sea level (meters)
Arkavaz	648794.6689	3696776.245	1423
Ema	633191.9299	3703942.424	1106
Ilam	629919.843	3716836.456	1326
Ivan	621330.2657	3661276.769	1169
Abdanan	725824.5816	3652032.893	916
Bishederaz	684077.1853	3634523.828	358
Pahle	675945.8749	3652862.22	768
Takhtan	691194.0158	3669787.021	1080
Chenarbashi	665751.0716	3702597.913	1100
Chenan	608056.3701	3735048.286	946
Chaharmele	618609.0572	3755505.731	1338
Dareshahr	723886.732	3668632.344	630
Dashteabbas	766452.4025	3590155.067	175
Dehloran	714081.985	3618486.369	232
Sarableh	646622.9019	3737411.13	1052
Salehabad	609959.2904	3703649.821	623
Karezan	6442047.2325	3733644.817	1262
Kolm	676798.6435	3691700.984	926
Gonbad	644391.6254	3680073.544	903
Lumar	668623.1034	3715588.884	789
Mormori	751455.3486	3624910.228	524
Mehran	610441.234	3661147.936	145

2.2.3. Validation data

Crop yield data**:** The present study used the total crop yield information as the reference data to monitor the drought. The data were collected from the Jihad Agricultural Organization of Ilam province for the 2005-2018 period, including the data on wheat and barley.

Soil moisture data**:** Several studies previously conducted in the same study area have provided data concerning the soil relative moisture at a depth of 10 cm for the 2000-2018 period. Therefore, this study used monthly relative moisture data (at the depth of 10 cm) as another reference data to validate the drought model.

Standardized Precipitation apotranspiration Index (SPEI): a multivariate index in which precipitation and temperature data are combined. The 12-month SPEI data (for the 2000-2018 period) were used to validate the constructed model.

2.3. Methodology

The Drought Hazard Index (DHI) was calculated by combining multi-sensor and geographical indices using learning machine approaches. Figure 2 illustrates the overall framework adopted in this study.

Remote sensing-based drought factors used in this study (Table 2) included Land Surface Temperature (LST), Normalized Difference Vegetation Index (NDVI), Normalized Difference Drought Index (NDDI), Normalized Difference Water Index (NDWI), Normalized Multi-band Drought Index (NMDI), Soil Adjusted Vegetation Index (SAVI), and actual ET (collected from MODIS). Moreover, the digital elevation map (DEM) was as an auxiliary layer. The data collected for each month from 2000 to 2018 were scaled from 0 to 1 using max-min scaling (Kogan, 1993). Then, the dataset was divided into two groups: training (70%) and validation datasets (30%). Table 2 shows the remote sensing indices and their formula used in this study.

Table 2: Remote sensing indices used in this study and their formula

Index	Formula	Reference
NDVI	$\frac{\mathrm{NIR}-\mathrm{RED}}{\mathrm{NIR}+\mathrm{RED}}$	Rouse, 1974
NMDI	(NIR − (SWIR2 − SWIR3))/(NIR + (SWIR2 − SWIR3))	Wang and Qu., 2007
EVI	$G\times\frac{P_{NIR}-P_{RED}}{P_{NIR}+C_1\times P_{RED}-C_2\times P_{BLUE}+L}$	Huet., 1984
SAVI	$\frac{(\mathrm{NIR}-\mathrm{RED})(1+L)}{(\mathrm{NIR}+\mathrm{RED}+L)}$	Huet., 1984
NDWI5, 6 & 7	(ρband2 − ρband5(or 6 or 7))/(ρband2 + ρband5 (or 6 or 7)	Gao., 1996
NDDI5, 6 & 7	(NDVI − NDWI)/(NDVI + NDWI)	Gu et al., 2007

2.3.1. Drought monitoring and drought model construction

As for the theoretical basis of combining indices to monitor drought, it could be argued that a multiple-variable drought index includes a variety of variables that are not only related to precipitation but also to factors that influence drought, with each of them representing the drought from a different perspective. On the other hand, elevation significantly affects regional drought.

To model the drought, this study used a total of 13 variables as input variables, all of which were derived from previous studies as relevant important factors (Park et al., 2016; Shen et al., 2019). Moreover, the data regarding each month (for the 2000-2018 period) were scaled from 0 to 1 using max-min scaling. Three machine learning approaches, including BRT, RF, and Cubist, were also used to examine the relationship between indices and drought conditions. It should be noted that 70% of the samples were used randomly as the training data and the remaining 30% were used for validation.

2.3.2. Machine learning approaches

Random Forest (RF): a technique for developing a regression model between a set of inputs and the desired output. The RF model also provides the relative significance of each variable (Kim et al., 2013; Park et al., 2016) through the application of sensitivity analysis using the mean decrease accuracy (MDA). In this study, the RF model calculations were performed in the R4.1.0 software environment using the “Random Forest” package (Liaw & Wiener., 2002).

Boosted Regression Tree (BRT): The method uses the capabilities of two algorithms, including regression trees and an adaptive method to combine a large number of simple models for delivering an appropriate performance (Elith et al., 2008). In the current study, the calculations of the RF model were performed in the R4.1.0 software environment using the “GBM” package.

Cubist: Aa a commercial rule-based machine, Cubist is a software learning tool that draws a modified regression tree (Rule Quest, 2012). The tool has so far been used for various drought studies (Brown et al., 2008). In the present study, the calculations of the RF model were performed in the R4.1.0 software environment using the “model data” package.

Generally, a total of thirteen factors were used as independent variables, and one-month, three-month, six-month, nine-month, and twelve-month SPIs were used as dependent variables in the machine learning models. Moreover, the performance of each machine learning model was evaluated using the coefficient of determination (R^2) and Root-Mean-Square Error (RMSE).

2.3.3 Drought Hazard Index (DHI)

Drought hazard is defined as the product of drought characteristics, including frequency and magnitude. Several studies have fully described the procedure for calculating drought events, (i.e., Kim et al., 2013; Dabanli., 2018; Nasrollahi et al., 2018; Khoshnazar et al., 2021) which could briefly be summarized as follows: the drought's vulnerable area is calculated based on the percentage of the drought occurrence probability using the ratio between occurrence in time and the total occurrence.

In this study, first, the model with the best performance was selected, followed by the construction of a comprehensive drought index (i.e., considering the pixel of remote sensing images as a meteorological station) for each year and the preparation of the value of the blended-drought index. Then, the occurrence probability for each pixel and each county was obtained, and after that, the frequency of drought classes was calculated. Finally, the occurrence probability was calculated by dividing the frequency of drought occurrence in each drought class, considering all possible cases of drought. Using a weighting system, the drought events are classified into four classes: low (L), moderate (M), High (H), and Very High (VH) (Table 3).

As the occurrence of different drought severity does not have equal value in determining the risk of an area, each drought class was weighed to generate the DHI map. Weight (W) and rating (R) scores were assigned to each category based on severity and occurrence probability. The probability of occurrence was then classified into four ratings using the Jenks' natural break method (Poortaheri et al., 2013; Kim et al., 2013). Severity and probability of occurrence are weight (Dw) and rating (Dr) for each case (Table 3).

Table 3: Weights and rates assigned to the drought category

Drought index value	category	Weight (DW)	Occurrence Probability	Rating (Dr)
0.75-1	Very High (VH)	4	Very high	4
			High	3
			Low	2
			Very low	1
0.50-0.75	High (H)	3	Very high	4
			High	3
			Low	2
			Very low	1
0.25-0.50	Moderate (M)	2	Very high	4
			High	3
			Low	2
			Very low	1
0-0.25	Low (L)	1	Very high	4
			High	3
			Low	2
			Very low	1

Finally, DHI was calculated through equation 1 as follows:

$$DHI = LDr \times LDw + MDr \times MDw + HDr \times HDw + VHDr \times VHDw \qquad \text{(Eq. 1)}$$

Where LDr, MDr, HDr, and VHDr stand for the low, moderate, high, and very high rates assigned to the drought category respectively. On the other hand, LDw, MDw, HDw, and VHDw represent the weights designated for the low, moderate, high, and very high drought category, respectively. Moreover, the DHI values were normalized through the min-max normalization approach, converting all values to rates ranging from 0 to 1.

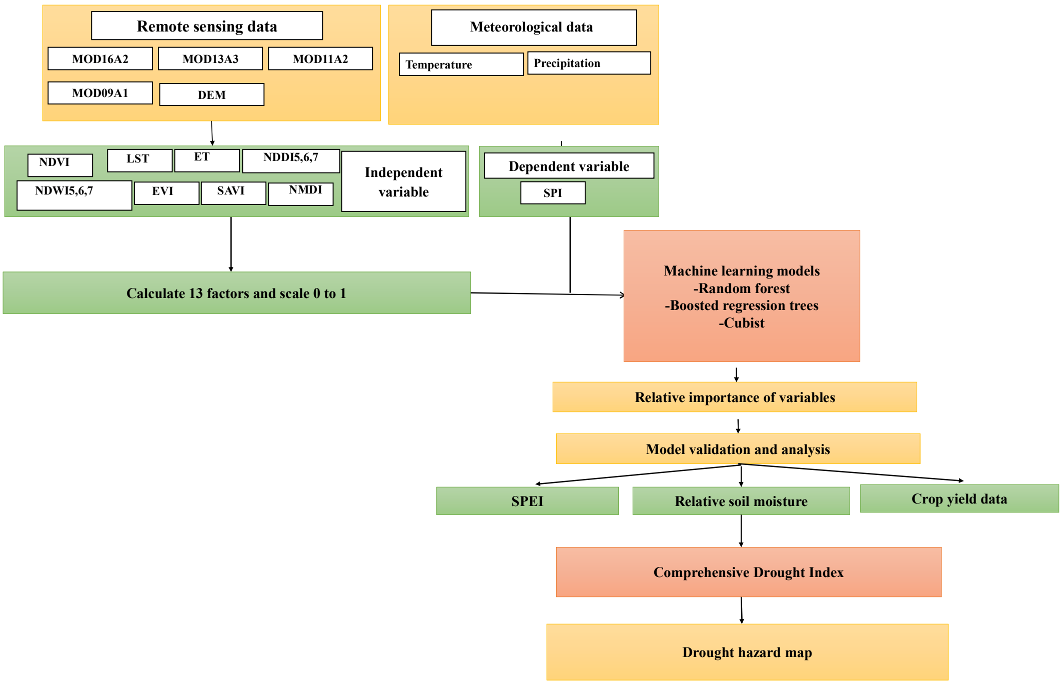


Figure (2): The Schema of the methodology used for calculating the Drought Hazard Index (DHI)

3. Results

Remote sensing data:

As mentioned earlier, twelve satellite indicators were introduced to run the models. Figure 3 shows the indices used for 2018 as an example. here only the map of 2018 is given as an example.

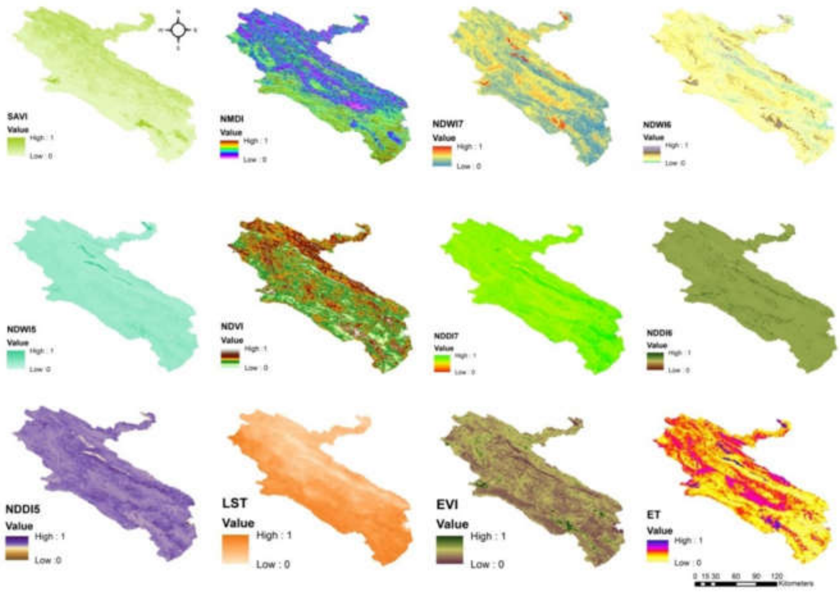


Figure (3): The indices used in drought modeling in 2018

3.1. Monitoring the meteorological drought and constructing its model:

BRT Model

Table 4 shows the results of BRT performance. Accordingly, the value of R^2 is generally higher for SPIs within a longer period, with BRT predicting the 6-month SPI with good accuracy. However, the model does not deliver a relatively good performance for SPIs during different periods. The R^2 for BRT was found to vary between 0.3 and 0.6.

Table (4): Performance of BRT model for SPIs of one, three, six, nine, and twelve months

Index	RSME	MAE	R^2
SPI1	0.19	0.11	0.36
SPI3	0.12	0.14	0.54
SPI6	0.21	0.10	0.60
SPI9	0.16	0.15	0.59
SPI12	0.28	0.11	0.57

The Cubist Model

Table 5 shows the results of cubist performance. Accordingly, the highest R^2 value belonged to the one-month SPI. However, the results obtained from the evaluation of the Cubist model revealed that contrary to expectations, the model does not deliver a good performance in terms of modeling, with the model's R^2 values varying from 0.3 to 0.5.

Table (5): Performance of BRT model for SPIs of one, three, six, nine, and twelve months

Index	RSME	MAE	R^2
SPI1	0.17	0.13	0.31
SPI3	0.15	0.11	0.44
SPI6	0.18	0.12	0.48
SPI9	0.17	0.15	0.52
SPI12	0.19	0.14	0.50

RF model:

Figure 4 shows the results of the RF model. Accordingly, the random forest model indicated acceptable accuracy in simulating the SPI at different time scales except for the 1-month SPI, clearly underestimating SPI-1 values.

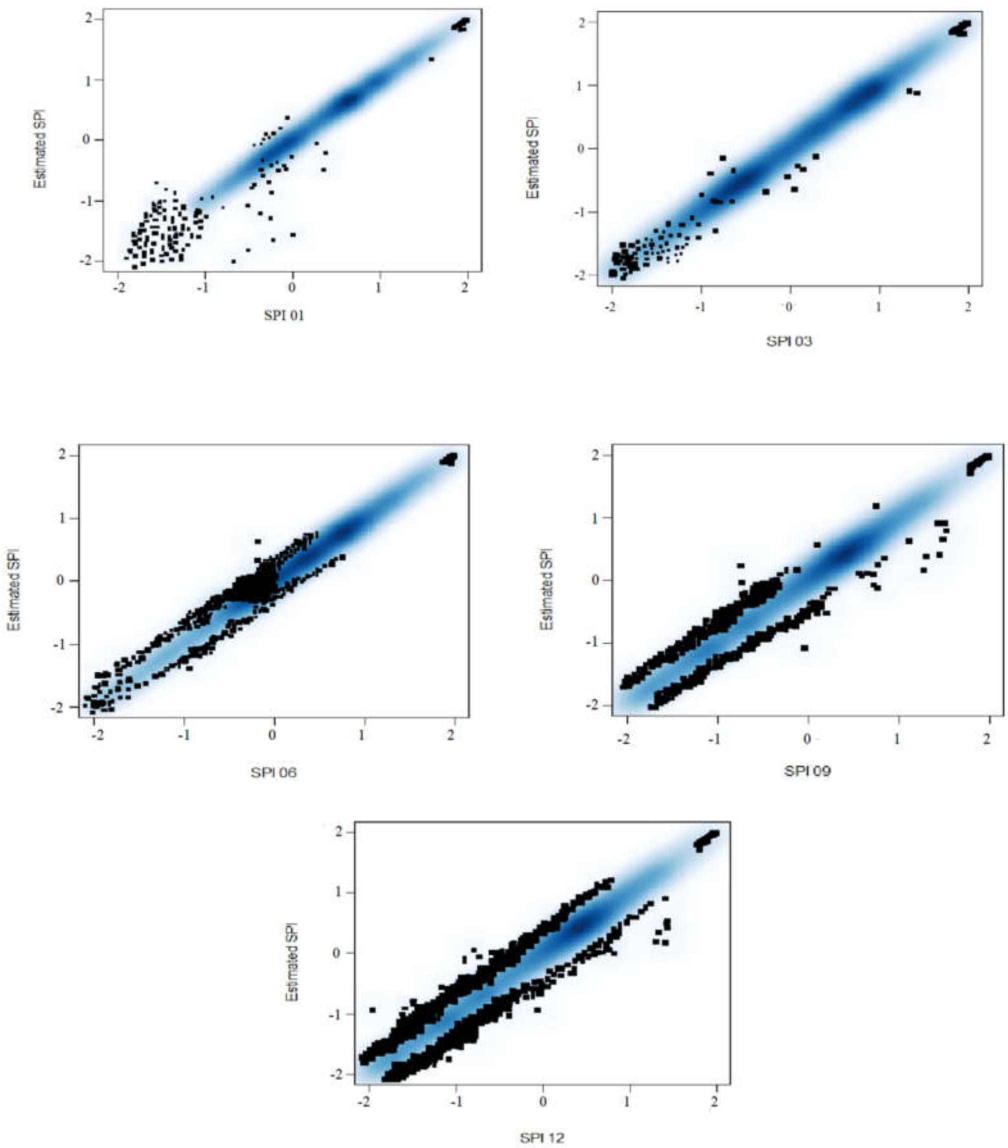


Figure (4): Scatter plots found for the estimated SPIs at periods of one, three, six, nine, and twelve months using the RF model

Table 6 shows the results of RF performance, according to which, the highest R^2 value belonged to the 6-month SPI (0.88). Generally, it could be argued that R^2 values were higher for SPIs with a longer period.

Table (6): Performance of random forest model for 1-, 3-, 6-, 9-, and 12-month SPIs

Index	RSME	MAE	R^2
SPI1	0.29	0.11	0.66
SPI3	0.32	0.09	0.74
SPI6	0.41	0.009	0.88
SPI9	0.36	0.01	0.79
SPI12	0.38	0.08	0.83

As shown in Table 6, from among different SPIs, the 6-month SPI delivered the best performance, followed by the 12-month one. In general, it could be said that the model delivered a good performance for SPIs at different periods. For instance, the range of the predicted 6-month SPI obtained according to the RF is similar to the range of the actual 6-month SPI. These results are related to the accuracy of the model in the training phase.

3.1.2. Determining the relative importance of the drought-based factors

As the RF model showed the best performance in modeling the drought (out of three machine learning models investigated in this study), it was selected for this purpose. Table 7 shows the most appropriate environmental variables for modeling derived from the analysis of the variance inflation factor. Accordingly, six variables were selected as the most optimal parameters (out of thirteen variables), whose relative importance was identified using the data collected from a nineteen-year period (2000–2018).

Table (7): The Relative importance of the most significant indices found for 1-, 3-, 6-, 9-, and 12-month SPIs using the RF Model

Rank	SPI 1	SPI 3	SPI 6	SPI 9	SPI 12
1	ET(100)	SAVI (100)	NDWI7 (96)	NDWI6 (98)	NDWI6 (88)
2	LST (96)	NDWI6 (.98)	NDVI (95.2)	NDVI (88)	NMDI (81)
3	DEM (61)	NDVI (93)	NDWI6 (87.4)	DEM (77)	NDVI (76)
4	NDWI7 (43.2)	NDDI6 (91)	SAVI (86.2)	ET (66)	DEM (68)
5	NDVI (11.2)	EVI (74)	ET (72.7)	NMDI (69)	ET (65)
6	NDDI6 (3.5)	DEM (63)	DEM (64)	NDDI6 (12)	SAVI (11)

According to the results presented in Table 7, the importance of surface conditions-related indices such as surface temperature and evapotranspiration is greater in short drought periods than in long ones. Moreover, drought occurrence and the lack of precipitation are instantaneously connected to the elevation of the study area. The SAVI is also more important in short drought periods.

3.1.3. Analysis of the model's monitoring results

According to Table 8, the correlation between constructed model and the SPEI was found to be 0.88. Furthermore, the correlation between the crop yield data and the constructed model was reported as 0.68. In general, the results of correlation analysis revealed that the highest correlation existed between the constructed model and the reference data.

Table (8): Correlation between the constructed model and the reference data

Index	Constructed model	SPI	SPEI	Crop yield	Soil moisture
Constructed model	1				
SPI	0.85	1			
SPEI	0.88	0.78	1		
Crop yield	0.77	0.64	0.69	1	
Soil moisture	0.68	0.52	0.59	0.72	1

3.2. Mapping drought hazard

The DHI indicates the drought hazard. On the other hand, the Ilam province could be classified into four classes based on the DHI, whose map shows more drought severity in the south and southwest counties of the province than other parts of the province, considering the fact that these parts possess lower altitudes and this feature makes these Counties warmer. According to Table 9, 65% of the study area falls within the very high class of drought hazard (Figure 5).

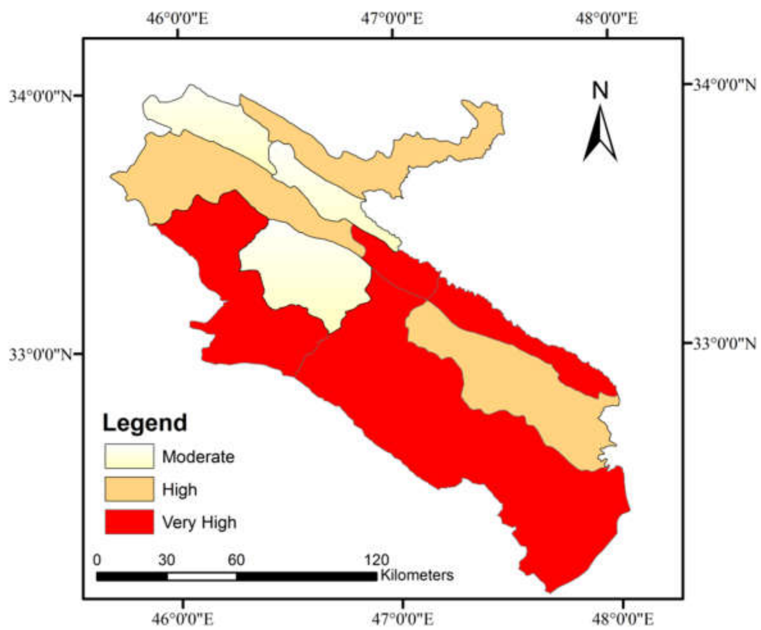


Figure 5: DHI map of Ilam province

Table (9): Area related to drought hazard classes

classes	Area	Percent (%)
Moderate	3238/46	16/09
High	3665/61	18/23
Very High	13202/3	65/66

4. Discussion

Considering the fact that drought is regarded as a very complex phenomenon, it is necessary to use a reliable drought index to accurately identify drought events and investigate their spatial and temporal changes. On the other hand, as several drought indices have so far been proposed, a suitable model needs to be developed by selecting and integrating important and highly influential variables in a specific region.

Located in western Iran, Ilam province usually suffers from drought due to its dry and semi-arid climate. Therefore, this study sought to evaluate and monitor drought in the province during a certain period (by considering various variables), trying to develop a drought hazard map. To do so, twelve satellite-extracted drought indices and one environmental indicator (which were introduced as efficiency indicators according to previous studies), were used and analyzed using machine learning approaches.

Moreover, BRT, RF, and Cubist approaches were used to simulate the SPI and determine the relative weight and importance of the indices. The results indicated that the BRT model achieved a higher efficiency than the Cubist one, whose efficiency proved to be the lowest among the three approaches mentioned above.

The acceptable performance of the RF model could partly be attributed to the self-evaluation and self-correction capability of the model during the construction of multiple trees (Kouranjadi and Porqashmi, 2018). Therefore, the present study used the random forest model as the most appropriate and efficient method for the study area. In this regard, the weight and relative importance of each index were determined using RF, followed by the removal of less important indices from the model to acquire six final indices.

The results of the model validation suggested that the correlation between the constructed model, soil moisture, and crop yield was higher than other variables. In other words, the constructed model for drought showed more realistic conditions of the drought occurrence in the region, which were consistent with the results found by Shen et al (2019). Moreover, the results are in agreement with the findings reported by Mizzell (2008), Wardlow et al. (2012), Park et al. (2016), Luetkemeier et al. (2017) and Sahana et al. (2021) who argued that the combination of indices achieved a higher efficiency in evaluating and monitoring drought than single indices.

The results also indicated that the LST and ET indices were more significant in shorter periods of drought, which was consistent with the findings of Zhang & Jia (2013). On the other hand, vegetation-related variables such as NDWI gained more significance with an increase in the drought period, which was consistent with the results reported by Gessner et al. (2013), Piao et al. (2003), and Park et al. (2016). Furthermore, Farrokhzadeh et al (2017) argued that remote sensing data, NDVI, and EVI can be used in areas with insufficient rain gauge data where meteorological stations are inappropriately distributed to be resorted to for monitoring drought. The SAVI and NDVI were also found to be of great significance for monitoring drought in Ilam Province, which was consistent with the results of a study carried out by Faramarzi et al (2018) in the province.

Moreover, assessing the drought hazard throughout the study's statistical period (2000-2018) revealed that Ilam was exposed to moderate to very high drought hazard risks and that there was a very high drought hazard risk in some of the province's cities, including Mehran, Dehlran, Dehrashahr, and Badreh. Also, it was found that 65% of the study area fell in a very high class of drought hazard risk, which could be considered as a threat to the region.

5. Conclusions

This study used machine learning approaches to construct a multivariate drought index using remote sensing data and geographical information. Considering the limited and insufficient number of meteorological stations in Iran, especially in Ilam province, and their inappropriate distribution in the region, the model proposed in the current study could be considered as a novel and effective method for monitoring regional drought. In this regard, the study found that satellite imagery with high accuracy can play a very important role in monitoring, evaluating, and modeling drought.

In this study, the pixel size of satellite images was used as a meteorological station that can significantly increase the accuracy of drought assessment since the number of meteorological stations in the province is limited due to its specific geographical conditions. While this study used machine learning methods, other data mining methods such as deep learning, support vector models, and other efficient tools in the field of drought monitoring should be used to prove the constructed model.

References

1. AghaKouchak, A., Farahmand, A., Melton, F.S., Teixeira, J., Anderson, M.C., Wardlow,

B.D. and Hain, C.R., 2015. Remote sensing of drought: progress, challenges and opportunities, Journal of Reviews of Geophysics. 53, 452–480.

2. Brown, J.F., Wardlow, B.D., Tadesse, T., Hayes, M.J. and Reed, B.C. 2008. The Vegetation Drought Response Index (VegDRI): a new integrated approach for monitoring drought stress in vegetation. *GIScience Remote Sens*. 45 (1), 16–46.
3. Chang, J., Li, Y., Wang, Y., Yuan, M., 2016. Copula-based drought risk assessment combined with an integrated index in the Wei River Basin. China J. Hydrol. 540, 824–834.
4. Chang, J., Li, Y., Wang, Y. and Yuan, M., 2016. Copula-based drought risk assessment combined with an integrated index in the Wei River Basin. China J. Hydrol. 540, 824–834.
5. Dabanli, I, 2018. Drought Risk Assessment by Using Drought Hazard and Vulnerability Indexes, Natural Hazards and Earth System. Sciences Discuss. 1–15.
6. Dai, A, 2011. Erratum: drought under global warming: a review, Wiley Interdisciplinary Reviews: Climate Change, 2 (1), 45–65.
7. Dutra, E., Giuseppe, F. D., Wetterhall, F. and Pappenberger, F, 2013. Seasonal forecasts of droughts in African basins using the Standardized Precipitation Index, Hydrology and Earth System Sciences, 17(6), 2359-2373. doi:10.5194/hess-17-2359-2013.
8. Elith, J., Leathwick, J. R., & Hastie, T. (2008). A working guide to boosted regression trees. Journal of animal ecology, 77(4), 802-813. doi: 10.1111/j.1365-2656.2008.01390.x
9. Faramarzi M, Heidarizadi Z, Mohamadi A, Heydari M. Detection of Vegetation Changes in Relation to Normalized Difference Vegetation Index (NDVI) in Semi-Arid Rangeland in Western Iran. JAST. 20 (1) :51-60.
10. Farrokhzadeh, B, Mansouri, Sh, and Sepehari, A, 2016, determination of the correlation between vegetation indices NDVI and EVI with meteorological drought index SPI (case study: plain pastures of Golestan province), Agricultural Meteorology Journal, Volume 5, No. 2, pp. 65-56.
11. Gao, B.-C, 1996. NDWI—a normalized difference water index for remote sensing of vegetation liquid water from space, Remote Sensing Environment (RSE). 58 (3), 257–266.
12. Gessner, U., Naeimi, V., Klein, I., Kuenzer, C., Klein, D., and Dech, S. (2013). The relationship between precipitation anomalies and satellite-derived vegetation activity in Central Asia. Global and Planetary Change, 110, 74-87. https://doi.org/10.1016/j.gloplacha.2012.09.007.
13. Gu, Y., Brown, J.F., Verdin, J.P. and Wardlow, B, 2007. A five-year analysis of MODIS NDVI and NDWI for grassland drought assessment over the central Great Plains of the United States, Geophysical Research Letters, 34 (6).
14. 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. J. Hydrol. 527, 608–618.
15. Huete, A. R., Post, D. F., and Jackson, R. D, 1984. Soil Spectral effects and 4-space vegetation discrimination, Journal of Remote sensing of Environment, 15:155-165.
16. Khoshnazar, A., Corzo Perez, G.A., and Diaz, V, 2021. Spatiotemporal Drought Risk Assessment Considering Resilience and Heterogeneous Vulnerability Factors:

Lempa Transboundary River Basin in The Central American Dry Corridor, Journal of Marine Science and Engineering, 9(4):386. https://doi.org/10.3390/jmse9040386

17. Kim H, Park J, Yoo J, Kim TW, 2013, Assessment of drought hazard, vulnerability, and risk: a case study for administrative districts in South Korea. J Hydro Environ Res 9(1):28–35 doi:https://digitalcommons.unl.edu/droughtnetnews/80
18. Kogan, F.N. 1993.United States droughts of late 1980's as seen by NOAA polar orbiting satellites. International Geoscience and Remote Sensing Symposium, 1:197-199
19. Kouranjadi, A, Pourqasmi, H. 2018. Landslide susceptibility assessment using data mining models, case study: Chelchai watershed. Watershed Engineering and Management, 11(1), 28-42.
20. Liaw A, Wiener M, 2002. “Classification and Regression by randomForest.” R News, 2(3), 18-22. https://CRAN.R-project.org/doc/Rnews/.
21. Lin, Y.-C.; Kuo, E.-D.; Chi, W.-J. Analysis of Meteorological Drought Resilience and Risk Assessment of Groundwater Using Signal Analysis Method. Water Resour. Manag. 2021, 35, 179–197. [CrossRef]
22. Luetkemeier, R.; Stein, L.; Drees, L.; Liehr, S.2017, Blended Drought Index: Integrated Drought Hazard Assessment in the Cuvelai-Basin. Climate, 5, 51. https://doi.org/10.3390/cli5030051
23. 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 Climatology (Vol. 17, No. 22, pp. 179-183).
24. Mishra, A.K., Sivakumar, B., Singh, V.P., 2015. Drought processes, modeling, and mitigation. J. Hydrol. 526, 1–2.
25. Mizzell, H., 2008. Improving Drought Detection in the Carolinas: Evaluation of Local, State, and Federal Drought Indicators. University of South Carolina.
26. Nasrollahi, M, 2015. Assessment of drought hazard, vulnerability and risk (case study: Semnan province). M.Sc. thesis. Faculty of Natural Resources, University of Tehran. 104p
27. Nasrollahi, M., Khosravi, H., Moghaddamnia, A. 2018. Assessment of drought risk index using drought hazard and vulnerability indices. Arab J Geosci 11, 606 .https://doi.org/10.1007/s12517-018-3971-y
28. 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. Doi:https://doi.org/10.1016/j.agrformet.2015.10.011.
29. Piao, S., Fang, J., Zhou, L., Guo, Q., Henderson, M., Ji, W. 2003. Interannual variations of monthly and seasonal normalized difference vegetation index (NDVI) in China from 1982 to 1999. Journal of Geophysical Research: Atmospheres, 108(D14). Doi:https://doi.org/10.1029/2002JD002848
30. Poortaheri M, Eftekhari A, Kazemi N, 2013. The role of drought risk management approach in reducing social—economic vulnerability of farmers and rural regions case study: Sulduz Rural District, Azerbaijan Gharbi. Rural Res 4(1):1–12
31. Proodhan, Foyez A., Jiahua Zhang, Fengmei Yao, Lamei Shi, Til P. Pangali Sharma, Da Zhang, Dan Cao, Minxuan Zheng, Naveed Ahmed, and Hasiba P. Mohana. 2021. "Deep Learning for Monitoring Agricultural Drought in South Asia Using Remote Sensing Data" Remote Sensing 13, no. 9: 1715. https://doi.org/10.3390/rs13091715.

32. R Core Team, 2021. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.URL https://www.R-project.org/.
33. Rajsekhar, D., V. P. Singh, and A. K. Mishra, 2015, Integrated drought causality, hazard, and vulnerability assessment 20 for future socioeconomic scenarios: An information theory perspective. J. Geophys. Res. Atmos., 120, 6346–6378. http://doi.org/10.1002/2014JD022670 .
34. Rhee, J., Im, J., Carbone, G.J., 2010. Monitoring agricultural drought for arid and humid regions using multi-sensor remote sensing data. Remote Sens. Environ. 114 (12), 2875–2887.
35. Rouse, J.W., 1974. Monitoring the vernal advancement of retro gradation of natural vegetation. NASA/GSFC, Type III, Final Report, Greenbelt, MD, pp. 371.
36. RuleQuest, 2012. http://www.rulequest.com/.
37. Sahana, V, Mondal, A, Sreekumar, P. 2021, Drought vulnerability and risk assessment in India: Sensitivity analysis and comparison of aggregation techniques, Journal of Environmental Management, 299(113689), pages 1-10. https://doi.org/10.1016/j.jenvman.2021.113689
38. Shahid, S., Behrawan, H., 2008. Drought risk assessment in the western part of Bangladesh. Nat. Hazards 46, 391e413.
39. Shen, R, Huang, A, Li, B, Guo, J., 2019, Construction of a drought monitoring model using deep learning based on multi-source remote sensing data,International Journal of Applied Earth Observation and Geoinformation,Volume 79,2019,Pages 48-57,
40. Svoboda, M., 2000. An introduction to the drought monitor. Drought Network News(1994-2001).80.
41. United Nation Development Program (2004) Reducing disaster risk, A challenge for development. United Nation Development Program/ Bureau for Crisis Prevention and Recovery, New York. http://www. undp.org/bcpr/disred/rdr.htm.
42. Wang, L., Qu, J.J., 2007. NMDI: A normalized multi-band drought index for monitoring soil and vegetation moisture with satellite remote sensing. Geophys. Res. Lett. 34 (20).
43. Wardlow, B. D., Anderson, M. C., & Verdin, J. P. (Eds.). (2012). Remote sensing of drought: Innovative monitoring approaches. CRC Press
44. Waseem, M., Ajmal, M., & Kim, T. W. 2015. Development of a new composite drought index for multivariate drought assessment. Journal of Hydrology, 527, 30-37. Doi:https://doi.org/10.1016/j.jhydrol.2015.04.044.
45. Wilhite, D.A., Glantz, M.H., 1985. Understanding: the drought phenomenon: the role of definitions. Water Int. 10 (3), 111–120.
46. Wilhite, D.A., Svoboda, M.D., Hayes, M.J., 2007. Understanding the complex impacts of drought: a key to enhancing drought mitigation and preparedness. Water Resour. Manage. 21 (5), 763–774.
47. Wood, E. F., S. D. Schubert, A. W. Wood, C. D. Peters-Lidard, K. C. Mo, A. Mariotti, and R. S. Pulwarty, 2015: Prospects for Advancing Drought Understanding, Monitoring, and Prediction. *J. Hydrometeor.*, 16, 1636–1657.
48. Yin, J., Zhan, X., Hain, C.R., Liu, J., Anderson, M.C., 2018. A method for objectively integrating soil moisture satellite observations and model simulations toward a blended drought index. Water Resour. Res.

54. https://doi.org/10.1029/ 2017WR021959.

49. Zhang, A., Jia, G., 2013. Monitoring meteorological drought in semiarid regions using multi-sensor microwave remote sensing data. Remote Sens. Environ. 134, 12–23. Doi:https://doi.org/10.1016/j.rse.2013.02.023