Research Article

Prediction of Drought Severity Using Model-Based Clustering

Rizwan Niaz ¹, **Ijaz Hussain** ¹, **Xiang Zhang** ², **Zulfiqar Ali** ³,
Elsayed Elsherbini Elashkar^{4,5}, **Jameel Ahmad Khader**⁶, **Sadaf Shamshoddin Soudagar**⁷,
and **Alaa Mohamd Shoukry** ^{8,9}

¹Department of Statistics, Quaid-i-Azam University, Islamabad, Pakistan

²National Engineering Research Center of Geographic Information System, School of Geography and Information Engineering, China University of Geosciences (Wuhan), Wuhan 430074, China

³State Key Laboratory of Hydro-Science and Engineering and Department of Hydraulic Engineering, Tsinghua University, Beijing 100084, China

⁴Administrative Sciences Department, Arriyadh Community College, King Saud University, Riyadh, Saudi Arabia

⁵Applied Statistics Department, Faculty of Commerce, Mansoura University, Mansoura, Egypt

⁶College of Business Administration, King Saud University Muzahimiyah, Muzahimiyah, Saudi Arabia

⁷College of Business Administration, King Saud University Riyadh, Riyadh, Saudi Arabia

⁸Arriyadh Community College, King Saud University, Riyadh, Saudi Arabia

⁹KSA Workers University, Nsar, Egypt

Correspondence should be addressed to Xiang Zhang; zhangxiang76@cug.edu.cn and Alaa Mohamd Shoukry; aabdulhamid@ksu.edu.sa

Received 1 April 2021; Accepted 9 July 2021; Published 23 July 2021

Academic Editor: Bosheng Song

Copyright © 2021 Rizwan Niaz et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Drought is a common climatic extreme that frequently spreads across large spatial and time scales. It affects living standard of people throughout the globe more than any other climate extreme. Therefore, the present study proposed a new technique, known as model-based clustering of categorical drought states sequences (MBCCDSS), for monthly prediction of drought severity to timely inform decision-makers to anticipate reliable actions and plans to minimize the negative impacts of drought. The potential of the proposed technique is based on the expectation-maximization (EM) algorithm for finite mixtures with first-order Markov model components. Moreover, the proposed approach is validated on six meteorological stations in the northern area of Pakistan. The study outcomes provide the basis to explore and frame more essential assessments to mitigate drought impacts for the selected stations.

1. Introduction

Drought is a multifaceted and recurring event characterized by precipitation insufficiency, which has significant effects on hydrological systems, agriculture, and society [1, 2]. Drought lasts for a long time and brings extreme meteorological consequences, causing distress to crop yield and other plant reproduction [3]. In recent decades, drought has dramatically impacted the environment and economies worldwide [4, 5]. The determination of the incoming and termination times of the drought is still problematic for drought management. Structurally, the effects of drought slowly add over a period, and it may linger for an extended

period. Albeit having abstruse visual effects, these impacts of drought become severe without proper action and remain for a prolonged period even after termination [6–8].

According to drought occurrences and their characteristics, the well-known drought categories are meteorological, agricultural, hydrological, and socioeconomic [9, 10]. Among these categories of drought, a meteorological drought is a climatic event that is associated with a decrease in precipitation. In contrast, all other drought categories have more extensive human and social features [8, 11]. Moreover, the meteorological drought can lead to the other three types of drought; because of the intricacy and severity of drought, it becomes challenging to recognize and evaluate

drought characteristics. Therefore, in recent decades, many drought indices have been developed to assess and monitor drought events. Reliable and quality drought knowledge is essential for mitigation policies and preparation in disaster-stricken regions globally. Obtaining knowledge about drought occurrence is crucial for an early warning to lessen the adverse effects. Several drought indices are available in the literature and have been used by decision-makers to mitigate the negative impacts of drought.

There are different commonly known drought indices; for example, Palmer [12] has proposed a drought index called the Palmer Drought Severity Index (PDSI). This index incorporates soil moisture, precipitation, and temperature in a water balance model. Gibbs and Maher [13] have introduced a Decile Index (DI), Shafer and Dezman [14] proposed the Surface Water Source Index (SWSI), while the Standardized Precipitation Index (SPI) was introduced and has been used as a meteorological index by McKee et al. [15]. Albeit having a subtle discrepancy among the indices, the present analysis is accomplished using the SPI [15], which frequently has been used for drought monitoring policies and acquired endorsement from the World Meteorological Organization [16, 17]. It produces a consistent interpretation across various regimes and various spatial climates. Furthermore, it depicts ideal characteristics in forecasting and risk analyses as probabilistic approaches [18–20].

Moreover, multiple techniques have been developed in various studies to evaluate and predict drought occurrences [21–24]. However, drought is considered a complicated dynamic; therefore, much more fundamental work needs to be done to clarify the critical issues and demonstrate the effectiveness in enhancing both the monitoring and prediction of droughts. Hence, it is important to handle a drought process as a predictable dynamic system that helps to reduce the critical effects [5, 23, 25, 26]. Therefore, the current study proposes a new technique, known as model-based clustering of categorical drought states sequences (MBCCDSS) for grouping the categorical drought state sequences to predict the drought severity in the selected stations. The MBCCDSS may accurately and timely inform decision-makers to anticipate reliable actions and plans to mitigate negative drought impacts.

2. Methods

2.1. Standardized Precipitation Index (SPI). The SPI is commonly used for computing and recording drought occurrences [15]. It can be calculated for different periods based on monthly precipitation data. It provides a spatially reliable interpretation across several climates [27, 28]; Guttman 1998; [20]. Furthermore, the use of SPI is significantly high in geographical and temporal circumstances. The simplicity of calculation and availability of the SPI make it the most familiar worldwide. Usually, SPI-1 and SPI-3 consider meteorological drought, and SPI-6 and SPI-9 envisage agriculture drought. Moreover, hydrological drought is usually envisaged by SPI-12 and SPI-24 [29, 30]. However, the present study considers SPI-1 for quantifying drought occurrences from the data ranging from January 1971 to December 2017.

2.2. Model-Based Clustering of Categorical Drought State Sequences (MBCCDSS). The primary focus of the clustering technique is to group the data based on similar information. In contrast, specific information can be available from one another. It is prevalent in statistics and computer science due to its great variety of applications. There are numerous clustering techniques contemplated in the literature. Among them, there are various hierarchical clustering algorithms [31, 32], well-known k -means [33], and k -medoids [34] clustering algorithms. Moreover, model-based clustering is a technique that groups the objects of the data and assumes that each object of the cluster can be observed as a sample from some probability distribution [35, 36]. In case there are numerous data groups, various distributions are desired, and finite mixture models are needed [37]. Model-based clustering performance is outstanding in distinctly grouping objects [38]. Multiple challenging applications can be addressed by this technique, including mass spectrometry data [38, 39], text classification [40], and social networks [41]. Some works related to model-based clustering have been done in time series [42] and regression time series [43]. A high number of applications can be handled more reliably by using categorical grouping of sequences [39, 41–43]; however, in drought analysis, it has not established greater attention yet. In drought classification, the analysis of categorical sequences is important to obtain consistent results. Therefore, the present study proposed MBCCDSS that considers the transition pattern of the drought states and provides the basis for using model-based clustering to substantiate more reliable results about drought occurrences. The MBCCDSS is based on finite mixture modeling. The mathematical form for the finite mixtures can be written as

$$f(x|\theta) = \sum_{k=1}^K \alpha_k f_k(x|\theta_k), \quad (1)$$

where K is representing the total number of component distributions $f_k(.|\theta_k)$ with corresponding parameter vectors θ_k and $\alpha_1, \alpha_2, \dots, \alpha_K$ showing the mixing proportions, subject to $\alpha_k > 0$ and $\sum_{k=1}^K \alpha_k = 1$. $\theta = (\alpha_1, \alpha_2, \dots, \alpha_{K-1}, \theta_1^T, \theta_1^T, \dots, \theta_K^T)^T$ showing the entire parameter vector that has to be estimated. Moreover, the MBCCDSS models each data group by using a functional form of the first-order Markov model components. Furthermore, MBCCDSS used various sequences of drought states. These sequences reflect the steering behavior of drought states and reflect the importance of this on the application site. The drought states (extremely dry (ED), severely dry (SD), normal dry (ND), median dry (MD), median wet (MW), severely wet (SW), and extremely wet (EW)) are classified according to [44]. Now, let $X_i = (X_{i1}, X_{i2}, \dots, X_{iS})^T$ show the i -th categorical drought state sequence of length S_i following the first-order Markov model with p unique states. Then, we can write

$$P(X_i = x_i) = P(X_{i1} = x_{i1}) \prod_{s=2}^{S_i} (X_{is} = x_{is} | X_{i(s-1)} = x_{i(s-1)}), \quad (2)$$

where x_{is} , $s = 1, 2, \dots, S_i$, takes values in $\{1, 2, \dots, p\}$ and shows the drought state observed in the s -th position of x_i . Furthermore, to specify the notation, we denote the initial state probability as $\beta_j = P(X_{i1} = j)$ and transition probability as $\gamma_{jj'} = P(X_{is} = j' | X_{i(s-1)} = j)$, subject to the restrictions $\sum_{j=1}^p \beta_j = 1$ and $\sum_{j=1}^p \gamma_{jj'} = 1$ for $j = 1, 2, \dots, p$. Now it can be written as

$$\begin{aligned} P(X_i = x_i) &= P(X_{i1} = x_{i1}) \prod_{s=2}^{S_i} (X_{is} = x_{is} | X_{i(s-1)} = x_{i(s-1)}) \\ &= x_{i(s-1)} = \prod_{j=1}^p \beta_j^{I(x_{i1}=j)} \prod_{j=1}^p \prod_{j'=1}^p \gamma_{jj'}^{y_{ijj'}}, \end{aligned} \quad (3)$$

where $I(\cdot)$ is considered as an indicator function and $y_{ijj'}$ shows the frequency of the transitions from state j to state j' within the i -th sequence and assume that each categorical sequence originates from one of the K components. K

represents the total number of components. The order of these components is detected by minimizing the Bayesian information criterion (BIC) [45]. Using the notations and the final form of the finite mixture model with the first-order Markov model with $p \times p$ matrix Y_i having elements $y_{ijj'}$, equation (1) can be written as

$$f(x_{i1}, Y_i | \theta) = \sum_{k=1}^K \alpha_k \prod_{j=1}^p \beta_j^{I(x_{i1}=j)} \prod_{j=1}^p \prod_{j'=1}^p \gamma_{jj'}^{y_{ijj'}}, \quad (4)$$

where information of the i -th sequence is summarized in terms of the first state observed and the transition frequency matrix, which is considered as a minimal sufficient statistic, i.e., a pair (x_{i1}, Y_i) for estimating parameters of the model given in equation (2). The estimation of the parameters is done by the expectation-maximization (EM) algorithm [46]. The EM algorithm consists of two steps: expectation (E step) and maximization (M step). In the E step, the EM algorithm finds the conditional expectation of the complete-data log-likelihood function given observed data, and θ is used to maximize the conditional expectation in the M step. In the expectation step of the EM algorithm, posterior probabilities are calculated at the l -th iteration as

$$z_{ik}^l = \alpha^{(l-1)} \prod_{j=1}^p (\beta_{kj}^{(l-1)})^{I(x_{i1}=j)} \prod_{j=1}^p \frac{\prod_{j'=1}^p (\gamma_{jj'}^{(l-1)})^{x_{ijj'}}}{\sum_{k'=1}^K \alpha_{k'}^{(l-1)} \prod_{j=1}^p (\beta_{k'j}^{(l-1)})^{I(x_{i1}=j)} \prod_{j=1}^p \prod_{j'=1}^p (\gamma_{jj'}^{(l-1)})^{x_{ijj'}}}, \quad (5)$$

and the maximization step involves updating the parameter estimates by the following equations:

$$\alpha_{(k)}^l = \frac{1}{n} \sum_{i=1}^n z_{ik}^{(l)}, \quad (6)$$

$$\beta_{kj}^{(l)} = \frac{\sum_{i=1}^n z_{ik}^{(l)} I(x_{i1} = j)}{\sum_{i=1}^n z_{ik}^{(l)}}, \quad (7)$$

$$\gamma_{jj'}^{(l)} = \frac{\sum_{i=1}^n z_{ik}^{(l)} x_{ijj'}}{\sum_{i=1}^n z_{ik}^{(l)} \sum_{j'=1}^p x_{ijj'}}. \quad (8)$$

2.3. Prediction of Drought States. Using the set of transition probability matrices $\Gamma_1, \Gamma_1, \dots, \Gamma_K$ and a probability distribution $\pi_1, \pi_2, \dots, \pi_K$ linked with mixture components, the L -step transition probability matrix can be found by

$$\Gamma^L = \sum_{k=1}^K \pi_k \Gamma_k^L, \quad (9)$$

where Γ_k^L shows the matrix Γ_k raised to the power L . For instance, $\Gamma_k^4 = \Gamma_k \Gamma_k \Gamma_k \Gamma_k$. The choice of the distribution $\pi_1, \pi_2, \dots, \pi_K$ depends on the specific application. However, in the current scenario, the mixing proportion

estimated vector (i.e., $\widehat{\alpha_1}, \widehat{\alpha_2}, \dots, \widehat{\alpha_K}$) and the posterior probability estimated vector (i.e., $\widehat{z_{i1}}, \widehat{z_{i2}}, \dots, \widehat{z_{iK}}$) are associated with a specific sequence used to calculate the probability distribution for future drought state occurrences.

3. Application

The proposed technique is validated on six meteorological stations of the region, Northern area, Pakistan (Figure 1). The selection of the region is based on its structural importance and significant climatological characteristics [47]. The appearance in the atmosphere of the selected region adds significant effects on other parts of the country. Moreover, several changes have been observed in the country due to fluctuating weather patterns in the season in various regions. However, the highest temperature has been observed in larger parts of the country, and these parts were highly affected by global warming [48, 49]. Furthermore, global warming has not been affecting the Pakistan atmosphere alone but also the world. Its impact can be observed on temperature and water that cause high temperature and water deficiency. Although future climate changes can be problematic, these changes substantially impact rural livelihoods and their coping accomplishments. Furthermore, drought occurrences can damage several vital sectors of the country; for example, these occurrences can negatively affect the economy, agriculture, and natural resources. Therefore,

FIGURE 1: Geographical locations of the six selected stations of the region of Pakistan.

it is important to understand the drought occurrences more instantaneously by developing inclusive and efficient techniques. In these perspectives, the present study proposed a new technique that meaningfully improves the competency of observing drought occurrences in the selected area. These findings may enhance the capabilities of drought monitoring and mitigation policies.

3.1. Results. The inadequacy of precipitation and anarchy to an expected precipitation pattern cause drought events. The summary statistics of precipitation for selected stations is presented in Table 1. The monthly occurrence of precipitation in various months of the Chilas station is presented in

Figure 2, and the precipitation occurrence over the selected period for Gilgit is presented in Figure 3. We took these two stations to present precipitation occurrence; however, the precipitation occurrence for other selected stations can be presented accordingly. The theoretical versus empirical histograms of SPI-1 for selected stations are presented in Figure 4. The presented results in histograms can be envisaged as the discrepancy among stations; this divergence can be arisen due to the natural enactment of the data. In the recent past, many researchers have been working on modeling such discrepancy recitals in the data. Moreover, new procedures were proposed for the standardization based on nonparametric functions and mixture distribution functions [50], but still handling the discrepancy is under

TABLE 1: The statistics to address the period 1971–2017 of six selected stations.

Variable	Station	Mean	Median	St. dev
Precipitation	Astore	40.912	25.708	41.934
	Bunji	14.340	7.109	18.901
	Gupis	16.852	5.707	30.214
	Chilas	16.763	7.006	23.535
	Gilgit	12.420	6.056	16.576
	Skardu	20.631	9.107	25.907

The mean, median, and standard deviation of precipitation at different stations are provided.

FIGURE 2: The monthly precipitation occurrence at Chilas station. The colors in the figure show different levels of monthly precipitation that occurs. The horizontal axis indicates the selected years and the vertical axis shows the varying months.

FIGURE 3: The monthly precipitation behavior for Gilgit station. The colors in the figure show different levels of monthly precipitation that occurs in the selected years.

contemplation. Furthermore, the temporal behavior of the SPI-1 at various stations can be envisaged in Figure 5. Furthermore, the selected stations show more similar behavior in data over the region for a specific drought index [44]. However, varying distributions can be observed in selected stations (Table 2) to generate categorical values. The BIC are used to select appropriate distributions among the fitted distributions for the selected stations. The BIC value that is -1036.513 is considered the minimum for three parameters (3P) Weibull distribution in Astore station.

Furthermore, the three parameters (3P) Weibull distribution has minimum BIC values among all BIC values of other distributions for SPI-1 in particular stations such as Astore, Bunji, Gilgit, and Skardu with values of -1030.985 , -1097.487 , and -735.125 , respectively. The BIC values calculated from some distributions for Gupis station, 4P-Beta with BIC values that are -788.076 and Chilas with BIC value -805.614 , were considered the minimum among other specified distributions for these two stations. Usually, the Weibull distribution has applications in hydrology and

FIGURE 4: Theoretical versus empirical histograms of the selected distribution for six stations.

FIGURE 5: Continued.

FIGURE 5: Temporal behavior in various plots of selected stations for the SPI at scale-1.

TABLE 2: BIC of selected probability distributions at SPI-1 in selected stations.

Distribution	Astore SPI	Bunji SPI	Gupis SPI	Chilas SPI	Gilgit SPI	Skardu SPI
3P Weibull	-1036.513	-1030.985	-735.144	-800.233	-1097.487	-735.125
4P-Beta	-1031.384	-1020.690	-788.076	-805.614	-1085.123	-700.236

BIC values given in bold font represent the most likely candidate distributions used for the standardization of the SPI-1 in the selected stations under the current climatological characteristics.

associated disciplines [51] and has more significant candidacy features for standardization.

Furthermore, the concept of varying probability distributions selected for the varying stations advocates finite mixture modeling. Therefore, MBCCDSS is proposed for the prediction of various categorical drought states using a mixture of first-order Markov models. The use of Markov models reflects the dynamics of the drought occurrences. The MBCCDSS assumes the first-order Markov models in this analysis; however, higher-order Markov models can be included [38]. Furthermore, the MBCCDSS uses the categorical values corresponding to each drought state. These categorical values are specified for the various drought states that are classified according to Niaz et al. [44]. Moreover, it assumes that each categorical sequence of the selected states instigates from one of the K components. The mixture model order K is detected by minimizing the BIC [45]. The

mixture model with two components ($K = 2$) based on BIC values is selected for the analysis. The performance of the model is detected by including initial state probabilities and without initial state probabilities. It can be observed from Table 3 that, for the first case, the maximized log-likelihood (LogL) value is equal to -3467.271 , while in the second case, it is equal to -3477.991 . Expectedly, the inclusion of the initial state probabilities in the model yields a higher LogL value as it slightly better fits the data. The variability is rather marginal, and based on the BIC value, the model with initial state probabilities with BIC 7061.757 is preferred over the model without initial state probabilities with BIC equal to 7065.28. However, the BIC have superiority over other competitors in finite mixture modeling, which is used for model selection and its performance.

Moreover, the mixing proportions and the posterior probabilities associated with a specific sequence are used to

TABLE 3: The performance of the MBCCDSS is detected by using initial state probabilities and without initial state probabilities.

Methods		$K = 1$	$K = 2$	$K = 3$	$K = 4$
EM algorithm (with β)	LogL	-3539.739	-3467.271	-3447.74	-3435.261
	BIC	7142.191	7061.757	7087.199	7126.744
EM algorithm (without β)	LogL	-3550.46	-3477.991	-3458.461	-3445.982
	BIC	7154.673	7065.28	7081.764	7112.35

The corresponding maximized LogL and BIC values are given for each number of components $K = 1, 2, 3$, and 4. BIC value provided in bold font shows the best model detected, which is $K = 2$ compared to $K = 1, 3$, and 4.

calculate the drought state prediction. The six sequences are used in MBCCDSS, which means that each sequence contains the monthly categorical observations of the specific station. For example, varying states in the Astor station are considered in sequence one (sequence-1); the second sequence (sequence-2) considers all various states of the Bunji station, states of Gupis's station are considered in the third sequence (sequence-3), and so forth. The prediction at one month of varying drought states is given in Table 4. However, MBCCDSS can predict the probabilities of the selected states in other months. The obtained results from sequence-1 show that the most likely state to visit in one month is ND; the probability (0.6247) associated with this prediction is higher than other drought states. In sequence-2, the ND state prevails among other drought states. The probabilities of the selected states at one month in other sequences can be observed accordingly. Furthermore, the appearance of NA values shows that the specific state is not available in six categorical drought state sequences. Therefore, MBCCDSS cannot predict any value for this drought state.

3.2. Discussion. The present study uses a mixture of first-order Markov models to develop a new technique for clustering categorical sequences of drought states. The proposed technique is applied to the six meteorological stations of the northern areas. The calculation of the MBCCDSS is based on the categorical sequences of the drought states. These categorical sequences of the drought states are calculated by SPI-1 and used in MBCCDSS to predict drought severity (i.e., ED, SD, MD, ND, MW, SW, and EW) in the selected stations. Furthermore, the outcomes of the MBCCDSS provide information about drought occurrences more plainly and accurately and can be used to support the mitigation strategies. Moreover, the probabilities obtained from MBCCDSS may be used to compare various drought indices, get more precise results about the drought occurrences for various drought states, find several propagations, and calculate various thresholds for different drought intensities in the selected region. Moreover, in MBCCDSS, the initial state and the transition probabilities are considered constant. The MBCCDSS assumes that the observations are time-homogeneous; however, these probabilities can be constructed by considering time as a function. The inclusion of temporal characteristics will improve the efficiency of MBCCDSS for drought monitoring. Furthermore, the results obtained in

TABLE 4: The one-month prediction from the last state in six sequences.

Sequence	One-month prediction						
	ED	SD	MD	ND	MW	SW	EW
Sequence-1	NA	0.1374	0.1569	0.6247	0.0595	0.0010	0.0205
Sequence-2	NA	0.0010	0.3880	0.5170	0.0572	0.0308	0.0060
Sequence-3	NA	0.0280	0.1260	0.6899	0.0929	0.0460	0.0173
Sequence-4	NA	0.0010	0.3880	0.5170	0.0572	0.0308	0.0060
Sequence-5	NA	0.0237	0.3495	0.5350	0.0576	0.0258	0.0084
Sequence-6	NA	0.0392	0.1296	0.5418	0.0844	0.0552	0.1498

The obtained results from sequence-1 show that the most likely state to visit in next month (i.e., January 2018) is ND; the probability associated with this prediction is slightly higher than 0.624. In other sequences, the ND is also prevailing among other states. This shows that policymakers should make their plans according to this drought state (category) to mitigate its negative impacts.

this study are significant for the existing conditions of the application site as the forthcoming promising climate circumstances can be unsuitable for the extrapolations based on the present analysis.

4. Conclusion

Drought is a slowly emerging issue, and the determination of its occurrence is still an issue to be solved. Structurally, the consequences of drought gradually accumulate over a period, and they may last for a long period. Drought distresses the lives of the people directly more than any other natural hazard and causes maleficent results for the society and the economy of the country. Therefore, it is necessary to handle drought occurrences as a predictable dynamic system, which used a particular memory and helps to minimize the critical effects. A new technique, known as MBCCDSS, is proposed for the monthly prediction of drought severity using model-based clustering. The MBCCDSS employed an EM algorithm for finite mixtures with first-order Markov model components. The MBCCDSS provides future probabilities for each of the drought states in selected stations. Moreover, the outcomes of the study may accurately and timely inform decision-makers to anticipate reliable policies and plans to mitigate the adverse effects of drought.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that there are no conflicts of interest.

Acknowledgments

This work was supported by grants from the National Natural Science Foundation of China program (41801339). The authors are also thankful to the Deanship of Scientific Research at King Saud University, through research group no. RG-1437-027.

References

- [1] T. D. Delbiso, J. M. Rodriguez-Llanes, A.-F. Donneau, N. Speybroeck, and D. Guha-Sapir, "Drought, conflict and children's undernutrition in Ethiopia 2000-2013: a meta-analysis," *Bulletin of the World Health Organization*, vol. 95, no. 2, pp. 94–102, 2017.
- [2] N. Puletti, W. Mattioli, F. Bussotti, and M. Pollastrini, "Monitoring the effects of extreme drought events on forest health by sentinel-2 imagery," *Journal of Applied Remote Sensing*, vol. 13, no. 2, Article ID 020501, 2019.
- [3] C. Rosenzweig, A. Iglesias, X. B. Yang, P. R. Epstein, and E. Chivian, "Climate change and extreme weather events-implications for food production, plant diseases, and pests," *Global Change and Human Health*, vol. 2, no. 2, pp. 90–104, 2001.
- [4] A. Dai, "Drought under global warming: a review," *WIREs Climate Change*, vol. 2, no. 1, pp. 45–65, 2011.
- [5] M. S. Hosseini, D. Samsampour, M. Ebrahimi, J. Abadia, and M. Khanahmadi, "Effect of drought stress on growth parameters, osmolyte contents, antioxidant enzymes and glycyrrhizin synthesis in licorice (*Glycyrrhiza glabra* L.) grown in the field," *Phytochemistry*, vol. 156, pp. 124–134, 2018.
- [6] M. A.-A. Hoque, B. Pradhan, and N. Ahmed, "Assessing drought vulnerability using geospatial techniques in north-western part of Bangladesh," *Science of the Total Environment*, vol. 705, Article ID 135957, 2020.
- [7] L. Sharafi, K. Zarafshani, M. Keshavarz, H. Azadi, and S. Van Passel, "Drought risk assessment: towards drought early warning system and sustainable environment in western Iran," *Ecological Indicators*, vol. 114, Article ID 106276, 2020.
- [8] D. A. Wilhite and M. Buchanan-Smith, "Drought as hazard: understanding the natural and social context," *Drought and Water Crises: Science, Technology, and Management*, pp. 3–29, CRC Press, Boca Raton, FL, USA, 2005.
- [9] C. Salvador, R. Nieto, C. Linares, J. Díaz, and L. Gimeno, "Effects of droughts on health: diagnosis, repercussion, and adaptation in vulnerable regions under climate change. Challenges for future research," *Science of the Total Environment*, vol. 703, Article ID 134912, 2020.
- [10] D. A. Wilhite, "Drought as a natural hazard: concepts and definitions," 2000.
- [11] O. V. Wilhelmi and D. A. Wilhite, "Assessing vulnerability to agricultural drought: a Nebraska case study," *Natural Hazards*, vol. 25, no. 1, pp. 37–58, 2002.
- [12] W. C. Palmer, *Meteorological Drought*, Vol. 30, US Department of Commerce Weather Bureau, Washington DC, USA, 1965.
- [13] W. J. Gibbs and J. V. Maher, *Rainfall Deciles As Drought Indicators*, Commonwealth of Australia, Melbourne, Australia, 1967.
- [14] B. A. Shafer and L. E. Dezman, "Development of surface water supply index (SWSI) to assess the severity of drought condition in snowpack runoff areas," in *Proceedings of the Western Snow Conference*, Reno, Nevada, January 1982.
- [15] T. B. McKee, N. J. Doesken, and J. Kleist, "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, American Meteorological Society, Boston, MA, USA, January 1993.
- [16] M. J. Hayes, M. D. Svoboda, B. D. Wardlaw, M. C. Anderson, and F. Kogan, "Drought monitoring: historical and current perspectives," 2012.
- [17] W.-H. Nam, M. J. Hayes, M. D. Svoboda, T. Tadesse, and D. A. Wilhite, "Drought hazard assessment in the context of climate change for South Korea," *Agricultural Water Management*, vol. 160, pp. 106–117, 2015.
- [18] N. B. Guttman, "Comparing the palmer drought index and the standardized precipitation index," *Journal of the American Water Resources Association*, vol. 34, no. 1, pp. 113–121, 1998.
- [19] A. K. Mishra and V. P. Singh, "A review of drought concepts," *Journal of Hydrology*, vol. 391, no. 1-2, pp. 202–216, 2010.
- [20] A. Zargar, R. Sadiq, B. Naser, and F. I. Khan, "A review of drought indices," *Environmental Reviews*, vol. 19, pp. 333–349, 2011.
- [21] P. Aghelpour, H. Bahrani-Pichaghchi, and O. Kisi, "Comparison of three different bio-inspired algorithms to improve ability of neuro fuzzy approach in prediction of agricultural drought, based on three different indexes," *Computers and Electronics in Agriculture*, vol. 170, Article ID 105279, 2020.
- [22] E. Arnone, M. Cucchi, S. D. Gesso, M. Pettita, and S. Calmanti, "Droughts prediction: a methodology based on climate seasonal forecasts," *Water Resources Management*, vol. 34, no. 14, pp. 4313–4328, 2020.
- [23] M. D. P. Jiménez-Donaire, A. Tarquis, and J. V. Giráldez, "Evaluation of a combined drought indicator and its potential for agricultural drought prediction in southern Spain," *Natural Hazards and Earth System Sciences*, vol. 20, no. 1, pp. 21–33, 2020.
- [24] A. Malik and A. Kumar, "Meteorological drought prediction using heuristic approaches based on effective drought index: a case study in Uttarakhand," *Arabian Journal of Geosciences*, vol. 13, no. 6, pp. 1–17, 2020.
- [25] C. Bhuiyan, R. P. Singh, and F. N. Kogan, "Monitoring drought dynamics in the Aravalli region (India) using different indices based on ground and remote sensing data," *International Journal of Applied Earth Observation and Geoinformation*, vol. 8, no. 4, pp. 289–302, 2006.
- [26] H. West, N. Quinn, and M. Horswell, "Remote sensing for drought monitoring & impact assessment: progress, past challenges and future opportunities," *Remote Sensing of Environment*, vol. 232, Article ID 111291, 2019.
- [27] R. V. Cumbie-Ward and R. P. Boyles, "Evaluation of a high-resolution SPI for monitoring local drought severity," *Journal of Applied Meteorology and Climatology*, vol. 55, no. 10, pp. 2247–2262, 2016.
- [28] D. H. Mlenga, A. J. Jordaan, and B. Mandebvu, "Monitoring droughts in Eswatini: a spatiotemporal variability analysis using the standard precipitation index," *Jamba*

- (Potchefstroom, South Africa), vol. 11, no. 1, Article ID 712, pp. 1–9, 2019.
- [29] J. P. Bloomfield and B. P. Marchant, “Analysis of groundwater drought building on the standardised precipitation index approach,” *Hydrology and Earth System Sciences*, vol. 17, no. 12, pp. 4769–4787, 2013.
 - [30] I. Nalbantis and G. Tsakiris, “Assessment of hydrological drought revisited,” *Water Resources Management*, vol. 23, no. 5, pp. 881–897, 2009.
 - [31] S. B. Dalirfesar, A. da Silva Meyer, and S. Z. Mirhoseini, “Comparison of similarity coefficients used for cluster analysis with amplified fragment length polymorphism markers in the silkworm, *Bombyx mori*,” *Journal of Insect Science (Online)*, vol. 9, no. 1, pp. 1–8, 2009.
 - [32] J. H. Ward, “Hierarchical grouping to optimize an objective function,” *Journal of the American Statistical Association*, vol. 58, no. 301, pp. 236–244, 1963.
 - [33] J. MacQueen, “Some methods for classification and analysis of multivariate observations,” in *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability*, vol. 1, pp. 281–297, no. 14, University of California Press, Berkeley, CA, USA, 1967.
 - [34] H.-S. Park and C.-H. Jun, “A simple and fast algorithm for K -medoids clustering,” *Expert Systems with Applications*, vol. 36, no. 2, pp. 3336–3341, 2009.
 - [35] J. D. Banfield and A. E. Raftery, “Model-based Gaussian and non-Gaussian clustering,” *Biometrics*, vol. 49, no. 3, pp. 803–821, 1993.
 - [36] C. Fraley and A. E. Raftery, “Model-based clustering, discriminant analysis, and density estimation,” *Journal of the American Statistical Association*, vol. 97, no. 458, pp. 611–631, 2002.
 - [37] D. Peel and G. J. McLachlan, “Robust mixture modelling using the t distribution,” *Statistics and Computing*, vol. 10, no. 4, pp. 339–348, 2000.
 - [38] V. Melnykov, “ClickClust: an R package for model-based clustering of categorical sequences,” *Journal of Statistical Software*, vol. 74, no. 9, pp. 1–34, 2016.
 - [39] V. Melnykov, “Challenges in model-based clustering,” *Wiley Interdisciplinary Reviews: Computational Statistics*, vol. 5, no. 2, pp. 135–148, 2013.
 - [40] K. Nigam, A. K. McCallum, S. Thrun, and T. Mitchell, “Text classification from labeled and unlabeled documents using EM,” *Machine Learning*, vol. 39, no. 2, pp. 103–134, 2000.
 - [41] M. S. Handcock, A. E. Raftery, and J. M. Tantrum, “Model-based clustering for social networks,” *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, vol. 170, no. 2, pp. 301–354, 2007.
 - [42] T. Warren Liao, “Clustering of time series data—a survey,” *Pattern Recognition*, vol. 38, no. 11, pp. 1857–1874, 2005.
 - [43] W.-C. Chen and R. Maitra, “Model-based clustering of regression time series data via APECM—an AEEM algorithm sung to an even faster beat,” *Statistical Analysis and Data Mining: The ASA Data Science Journal*, vol. 4, no. 6, pp. 567–578, 2011.
 - [44] R. Niaz, I. M. Hussain, Z. Ali et al., “A novel spatially weighted accumulative procedure for regional drought monitoring,” *Tellus A: Dynamic Meteorology and Oceanography*, vol. 72, 2020.
 - [45] G. Schwarz, “Cumulative index-volumes I–XVI,” *Progress in Optics*, vol. 6, no. 2, pp. 461–464, 1978.
 - [46] A. P. Dempster, N. M. Laird, and D. B. Rubin, “Maximum likelihood from incomplete data via the EM algorithm,” *Journal of the Royal Statistical Society: Series B (Methodological)*, vol. 39, no. 1, pp. 1–22, 1977.
 - [47] S. A. Awan, “The climate and flood risk potential of northern areas of Pakistan,” *Science Vision*, vol. 7, no. 3–4, pp. 100–109, 2002.
 - [48] S. Adnan and K. Ullah, “Development of drought hazard index for vulnerability assessment in Pakistan,” *Natural Hazards*, vol. 103, no. 3, pp. 2989–3010, 2020.
 - [49] K. Ahmed, S. Shahid, and N. Nawaz, “Impacts of climate variability and change on seasonal drought characteristics of Pakistan,” *Atmospheric Research*, vol. 214, pp. 364–374, 2018.
 - [50] S. Huang, Q. Huang, G. Leng, and S. Liu, “A nonparametric multivariate standardized drought index for characterizing socioeconomic drought: a case study in the Heihe River Basin,” *Journal of Hydrology*, vol. 542, pp. 875–883, 2016.
 - [51] E. S. Martins and J. R. Stedinger, “Generalized maximum-likelihood generalized extreme-value quantile estimators for hydrologic data,” *Water Resources Research*, vol. 36, no. 3, pp. 737–744, 2000.