WILEY | Hindawi

Research Article

Prediction for Various Drought Classes Using Spatiotemporal Categorical Sequences

Rizwan Niaz [ID],[1] **Mohammed M. A. Almazah**,[2,3] **Xiang Zhang** [ID],[4,5] **Ijaz Hussain** [ID],[1] **and Muhammad Faisal**[6]

[1]*Department of Statistics, Quaid-i-Azam University, Islamabad, Pakistan*
[2]*Department of Mathematics, College of Sciences and Arts (Muhyil), King Khalid University, Muhyil 61421, Saudi Arabia*
[3]*Department of Mathematics and Computer, College of Sciences, Ibb University, Ibb 70270, Yemen*
[4]*National Engineering Research Center of Geographic Information System, School of Geography and Information Engineering, China University of Geosciences (Wuhan), Wuhan 430074, China*
[5]*State Key Laboratory of Information Engineering in Surveying, Mapping, and Remote Sensing (LIESMARS), Wuhan University, Wuhan 430079, China*
[6]*Faculty of Health Studies, University of Bradford, BD7 1DP, Bradford, UK*

Correspondence should be addressed to Rizwan Niaz; razwanniaz11@gmail.com and Xiang Zhang; zhangxiangsw@whu.edu.cn

Drought frequently spreads across large spatial and time scales and is more complicated than other natural disasters that can damage economic and other natural resources worldwide. However, improved drought monitoring and forecasting techniques can help to minimize the vulnerability of society to drought and its consequent influences. This emphasizes the need for improved drought monitoring tools and assessment techniques that provide information more precisely about drought occurrences. Therefore, this study developed a new method, Model-Based Clustering for Spatio-Temporal Categorical Sequences (MBCSTCS), that uses state selection procedures through finite mixture modeling and model-based clustering. The MBCSTCS uses the functional structure of first-order Markov model components for modeling each data group. In MBCSTCS, the suitable order K of the components is selected by Bayesian information criterion (BIC). In MBCSTCS, the estimated mixing proportions and the posterior probabilities are used to compute probability distribution associated with the future steps of transitions. Furthermore, MBCSTCS predicts drought occurrences in future time using spatiotemporal categorical sequences of various drought classes. The MBCSTCS is applied to the six meteorological stations in the northern area of Pakistan. Moreover, it is found that MBCSTCS provides expeditious information for the long-term spatiotemporal categorical sequences. These findings may be helpful to make plans for early warning systems, water resource management, and drought mitigation policies to decrease the severe effects of drought.

1. Introduction

Drought is relatively more volatile than other natural disasters, and traditional valuations or forecast procedures are failed to predict it. Its relatively unperceptive onset and the multifaceted impacts cause the new assessment methodologies [1–5]. Since last decades, it has become more prominent to distress the environment and economic sectors worldwide than other natural hazards [6–8]. Moreover, determining the onset and end times of the drought is still challenging for drought management. Structurally, the effects of droughts slowly add over a period, and they may linger for a long period [8–10]. However, it can be characterized by a precipitation deficiency, which has substantial impacts on the agriculture, hydrological systems, and on living standards of the people [11, 12]. Despite perceptible effects of drought, these effects acclimatize severity without appropriate measures and are sustained for the long term even after termination [9].

The advancements in drought assessing and monitoring procedures can lead to better drought preparation and decrease the susceptibility of society to drought and its forgoing influences [8, 10, 13]. Therefore, it is essential to find more suitable techniques and procedures to predict drought occurrences more instantaneously. The improved method can be helpful to make plans for the early warning system, drought mitigation policies, and water resource management and decrease the severe effects of drought. Furthermore, the occurrences and characteristics of drought trigger the discussion about the various methodologies and techniques. Generally based on the occurrences and characteristics of the drought, authors have been categorizing the drought into various groups, including "meteorological, hydrological, agricultural, and socioeconomic" [14]. Chang [15] and Eltahir [16] defined that meteorological drought can be occurred due to the shortage of precipitation over a region for some time. Several studies have considered precipitation data to analyze meteorological droughts [17, 18]. The streamflow data have been frequently used for analyzing hydrological drought [19–21]. Furthermore, the reduction in soil moisture usually causes agricultural drought. The reduction in soil moisture can be affected by meteorological and hydrological droughts. Socioeconomic drought is linked to the shortfall in water resource systems, and in this case, the water supply is unable to meet water demands.

In the past few decades, numerous drought indices have been proposed to assess the drought occurrences [22–26]. The drought indices are frequently used to characterize the drought. The indices are based on various parameters that describe the spatial and temporal extents. Obtaining accurate and precise information about drought occurrences using several drought indices is crucial for an early warning policy; however, consistent and eminent drought information plays a crucial part in preparing drought monitoring and mitigating policies. Numerous drought indices with their strengths and weaknesses exist in the literature and are used by decision-makers who build action plans for drought early warning systems and mitigation policies. For example, Palmer [27] developed a drought index named the Palmer Drought Severity Index (PDSI). The PDSI worked well especially for subhumid and semiarid regions. The PDSI provided weekly information related to abnormal evapotranspiration deficit for the various regions. Information obtained from PDSI can be helpful for the crops in the region. The moisture condition of the regions can be assessed. Gommes and Petrassi [28] have proposed the national rainfall index (NRI). The NRI was used to provide synthetic discussion in sub-Saharan countries in Africa. They used NRI to determine the pattern recognition of rainfall in various regions. The Surface-Water Supply Index (SWSI) was introduced by Shafer and Dezman [24]. The computation of the SWSI is based on two major sources of irrigation water supply, namely, spring-summer streamflow runoff and reservoir carryover. Both sources are accumulatively analyzed to determine the total availability of surface water supply in season. Van Rooy [29] developed the Rainfall Anomaly Index (RAI). The RAI helped to find

geographical anomalies of the rainfall pattern in varying regions. Weghorst [26] has introduced the Reclamation Drought Index. Palmer [22] has introduced crop moisture index (CMI). Bhalme and Mooley [23] has developed Bhalme and Mooly drought index (BMDI). The BMDI used precipitation data and provided both negative and positive values to measure drought intensities. McKee et al. [25] developed the Standardized Precipitation Index (SPI). The SPI considered the time series of a long-term record of precipitation in the climatic areas. The dynamic characteristic of SPI is that it can be studied for different time scales and used to compare varying climatic areas. Therefore, SPI is being used extensively for evaluating and recording drought characteristics [30–35]. Furthermore, the drought indices that are mentioned above have been used frequently for drought monitoring in the different studies, although having discrepancies among the indices, to gain consistent interpretation across several regimes and spatial climates. This study utilized SPI, which is often employed to assess and monitor meteorological drought and is recommended by the World Meteorological Organization [36].

Furthermore, many clustering techniques are considered in the literature [37–41]. The clustering techniques focus on grouping the data so that the data group with similar characteristics would be selected within the cluster, while distinct information can exist among other clusters. Various clustering techniques have been frequently considered in machine learning approaches, especially in statistics and computer science, due to the variety of their applications [41–45]. Among the various techniques, model-based clustering groups data and presumes that each data cluster can be perceived as a part of any probability distribution [46, 47]. In various data groups, numerous distributions are preferred, and finite mixture models are desired [48]. The performances of the model-based clustering are outstanding in spectrometry data, text classification, social networks, and distinct grouping objects. Model-based clustering is used for time series [49] and regression time series analyses [50]. Several studies related to model-based clustering are available in the literature; however, it has not yet received greater attention in drought analysis. Therefore, this study developed a new technique known as Model-Based Clustering for Spatio-Temporal Categorical Sequences (MBCSTCS) to precisely predict drought occurrences for spatiotemporal categorical sequences. The performance of the proposed technique is assessed by using six meteorological stations in the northern area of Pakistan.

2. Methods

2.1. Standardized Precipitation Index (SPI). The long-term record of precipitation in the climatic area observed in the time sequence can be used to compute SPI. The vital feature of SPI is that it can be considered for various time scales and is being widely used to calculate and record drought occurrences [34, 35, 51, 52]. The analysis with various time scales can provide varying information. For example, the moisture conditions in different seasons can be assessed using SPI at a three-month time scale. The SPI

can assess information related to the water deficiency at a twelve-month time scale. Furthermore, the use of SPI describes the best characteristics in forecasting and analyzing risks as a probabilistic approach [31, 35, 53]. The SPI has been frequently used for drought monitoring in several aspects, for example, spatiotemporal analysis, forecasting, frequency analysis, and climatic studies [33, 35, 51, 52]. As precipitation is only used to determine the climatic condition for a particular area, it offers spatially reliable interpretations across various climates [32, 34, 35]. Therefore, it can be advantageous for the areas where other parameters are available that are required to calculate other indices and of significantly great concern to the various environmental and temporal circumstances [54]. This study focuses on the new methodology developed for monitoring drought more precisely and comprehensively in a specific area. The SPI at various time scales (1, 3, 6, 9, 12, and 24) is used for the current analysis.

2.2. Model-Based Clustering for Spatio-Temporal Categorical Sequences (MBCSTCS). Model-based clustering has been used for time series [49] and regression time series analyses [50]. Various studies associated with model-based clustering are available in the literature; the technique has significant importance for many applications; however, it has not yet received greater attention in drought analysis. Furthermore, in drought classification, categorical sequences are required for obtaining reliable results for the drought characterization. In this perspective, this study proposed MBCSTCS to analyze the categorical drought sequences for various time scales and stations. The MBCSTCS provides more significant results by using a categorical grouping of sequences than traditional approaches that have been used for the prediction. The MBCSTCS reflects the steering behavior of drought classes on various time scales and stations. Moreover, the selected drought classes (states) ("(Extremely Dry (ED), Severely Dry (SD), Normal Dry (ND), Median dry (MD), Median Wet (MW), Severely Wet (SW), and extremely Wet (EW)") are considered for the region [55].

Moreover, the first-order Markov model has a rationale in statistical modeling. The MBCSTCS considers the functional shape of first-order Markov model components for each data group. Furthermore, in the MBCSTCS the data groups consist of various sequences of drought states. For example, we let observation $X = (x_1, \ldots, x_m)^T$ that specifies for an ordered sequence, where each of its elements x_J consists of a categorical value that is specified for varying drought states and coded by natural integers. Furthermore, it is assumed that the number of unique drought states equals p, i.e., $x_J \, \varepsilon \, \{1, 2, \ldots p\}$ for $j = 1, 2, \ldots m$. Moreover, using a joint probability expression it can be written as $P(X = x) = P(X_1 = x_1, \ldots, X_m = x_m)$. In this format, the first-order Markov model provides an interesting method to describe the transitions between varying states. The probability of transitions of drought states in the next step depends only on the present state and has no connection to the drought states that are observed in the past. The joint probability using the first-order Markov model is given in the following equation:

$$P(X = x) = P(X_1 = x_1, \ldots, X_m = x_m) = P(X_1 = x_1) \prod_{J=2}^{m} P(X_j = x_j | X_{j-1} = x_{j-1}). \tag{1}$$

Furthermore, to simplify the notations, we use β to denote initial state probability and γ to represent the transition probability. For example, β_{x_1} shows the probability that the initial state is x_1 and transition probability of x_{j-1} to x_j is represented by $\gamma_{x_{j-1} x_j}$. So, utilizing the given notations, we can write as there are p states in the Markov model, and in this case, the initial state probabilities can be represented as $\beta = (\beta_1, \ldots \beta_p)^T$ and the matrix of the transitions as $\Gamma = (\gamma_{jr})_{p \times p}$. $P(X = x) = \beta_{x_1} \prod_{j=1}^{m} \gamma_{x_{j-1} x_j}$. Moreover, for the specific component based on finite mixture modeling the β_{x_1} and $\gamma_{x_{j-1} x_j}$ are replaced by the $\beta_{k x_1}$ and $\gamma_{k x_{j-1} x_j}$ and the model can be written as follows:

$$f(x; \theta) = \sum_{k=1}^{K} \pi_k \beta_{k x_1} \prod_{j=1}^{m} \gamma_{k x_{j-1} x_j}. \tag{2}$$

The log-likelihood of equation (2) can be expressed as follows:

$$\log L(\theta; \{x_i\}_{i=1}^n) = \sum_{i=1}^{n} \log \left\{ \sum_{k=1}^{K} \pi_k \prod_{l=1}^{p} (\beta_{kl})^{I(=x_{i1}=l)} \prod_{j=1}^{m_i} \gamma_{k x_{i(j-1)} x_{ij}} \right\}. \tag{3}$$

In equation (3), the $I(.)$ is indicator function and m_i indicates the length of ith categorical sequence. Expectation-maximization (EM) algorithm is employed to estimate the parameters [56].

2.3. Prediction of Future Drought Occurrences for Spatial-Temporal Categorical Sequences. The setting of transition probability matrices can be represented by $\Gamma_1, \Gamma_1, \ldots, \Gamma_k$ and a probability distribution $\pi_1, \pi_2, \ldots, \pi_K$ connected with mixture components, and the M-step transition probability matrix can be created by

$$\Gamma^M = \sum_{k=1}^{K} \alpha_k \Gamma_k^M, \tag{4}$$

where Γ_k^M indicates the matrix Γ^M raised to the power M. The choice of the appropriate distribution $\pi_1, \pi_2, \ldots, \pi_K$ is linked with the application. However, the $(\widehat{\alpha_1}, \widehat{\alpha_2}, \ldots, \widehat{\alpha_K})$ and (i.e., $\widehat{z_{i1}}, \widehat{z_{i2}}, \ldots, \widehat{z_{iK}}$), which are the mixing proportion estimated vector and the posterior probability estimated vector, respectively, associated with a particular sequence, can hold significant influence for the computation of probability distribution for future drought occurrences.

3. Application

The choice of the region is based on its structural impacts and other climatic characteristics that affect the other parts of the country. The outcomes of the study are obtained from the six selected stations with time-series data from January 1971 to December 2017 of the northern area of Pakistan (Figure 1) using SPI at various time scales. The selected stations have significant importance for the selected region and other regions of the country. For example, the reservoir system and agriculture sector are highly associated with the selected region; therefore, the climatic discrepancy of the region is significant for the other parts of the country [57, 58]. Furthermore, the fluctuation of the weather pattern in other regions within the country also contributed to their impacts on socioeconomic and environmental sectors. Most of the parts of the country have been facing the highest temperature, and these parts are being highly influenced by global warming [58, 59]. Undoubtedly, extreme climate events, including high temperatures, rainstorms, and droughts, are frequently associated with global climate warming. Climate warming significantly affects the universe, which usually causes a high temperature and water deficiency. These issues are associated with drought occurrences that damage the environment, natural resources, and lives of the people distinctly more than any other natural hazard. Furthermore, it produces convoluted consequences for society and the economic sectors of the country. Therefore, it is vital to recognize the drought occurrences more instantly by emerging comprehensive and efficient frameworks and techniques. In this regard, a new technique is applied to the selected stations that will influentially expand the capability of detecting drought occurrences and improve the competencies for drought evaluation and its assessment.

3.1. Results. The findings of this study are obtained by using long time series data collected from six climatological stations in the northern area of Pakistan. The selected stations

are observed to show homogenous results for the specific indices when calculated for varying stations with a single time scale [55]. However, on the varying time scales, the observations of the indices may vary. Furthermore, the inconsistency in their observations and varying generating processes of the drought states causes for developing a new method (i.e., MBCSTCS). The MBCSTCS considers the various time scales for a particular station as sequences with inconsistency in their sizes and varying data generating processes to analyze the spatiotemporal behavior of the drought states. It means that the observations of the SPI at scale-1 (SPI-1) for Astore station are considered as sequence-1, sequence-2 takes all observations of Astore station in SPI at scale-3 (i.e., SPI-3), and these sequences are continued to the last scale (SPI-24). Accordingly, these sequences can be assigned for other stations and time scales. Furthermore, the observations of each sequence assume that they come from the specific components that are selected appropriately for the data. The selected states are observed corresponding to every calculated value of SPI. These selected states are further distributed categorically for the computation of this study.

Moreover, Niaz et al. [55] proposed a new technique for monthly forecasting drought intensities using model-based clustering of categorical drought state sequences. The mentioned study is performed on various stations based on a single time scale. However, in this study, the various time scales are accumulatively considered for the monthly prediction of drought severity in a region. The outcomes of the current analysis are more appropriate, especially for the selected stations, and help the policymakers to make better policies related to various kinds of droughts including meteorological, hydrological, and socioeconomic. Furthermore, the current analysis is performed by using the R package ClickClust [45] that handles the case of coming observations from various probability distributions (K-components). The package is based on finite mixtures with Markov model components and is used to find the specific outcomes related to the specific sequence. The appropriate order K of the components (i.e., the mixture model) is identified by minimizing the Bayesian information criterion (BIC) [60]. Moreover, for a specific sequence, the mixing proportions estimated vector and the associated posterior probability estimated vector were used to calculate probability distribution associated with future steps of transitions from the last state of these sequences. Furthermore, climatological statistics on the given data of various stations are provided in Figure 2. To accomplish the analysis, the R package named *propagate* is used to provide appropriate findings and permit the specific analysis. In the mentioned package, various distributions are considered; among the given distributions, the appropriate choice of the distribution is based on the BIC values. This selection criterion is helpful to find the best fitting for the time scale and stations specified for the analysis.

The BIC values are given in Table 1 for the selected probability distributions fitting appropriately to the several time scales and stations. For example, at Astore station for scasle−1, the BIC value (−1036.5) of three-parameter (3P) Weibull distribution is found minimum among other distributions. Therefore, the (3P) Weibull distribution is
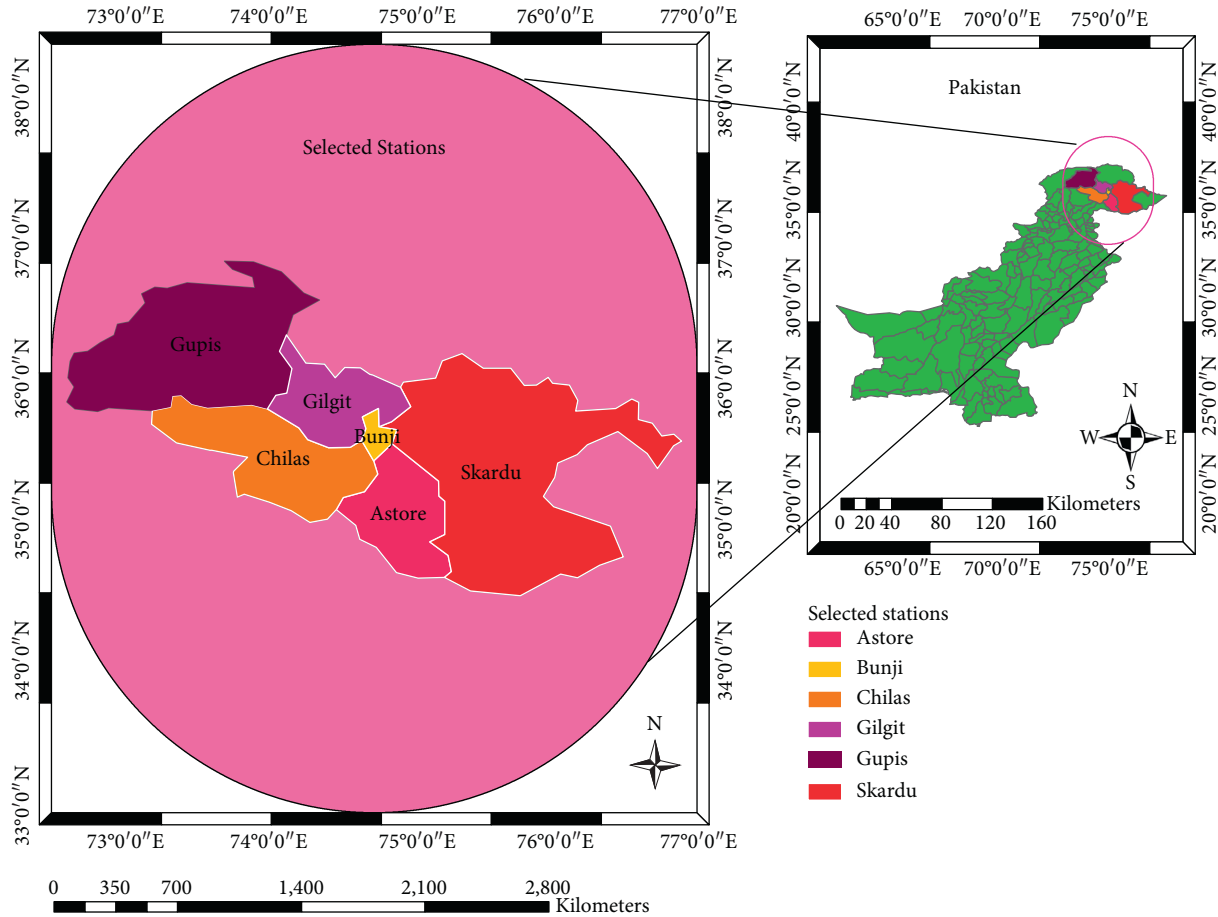
FIGURE 1: The study area consisting of six required stations in the northern area of Pakistan is presented.

considered as best fitted distribution for the Astore station at a scale−1. Furthermore, in Astore station for scale−3, the Gamma distribution is selected with the minimum BIC value (−1279.1). Moreover, in scale−6 and scale−9, it is also found that the Gamma distribution is best fitted in Astore station with minimum BIC values −892.8 and −896.1, respectively. Furthermore, Cosine and Skewed- normal distributions are considered for scale−12 and scale−24, respectively, in the Astore station. In Bunji station at scale−1, the (3P) Weibull is showing the minimum value of BIC (i.e., 1,031.0) and specified for the computation. At scale−3, scale−6, scale−9, scale−12, and scale−24 at Bunji station, the Gamma, Skewed-normal, Normal, Laplace, and Laplace are selected with BIC values −824.9, −1162.2, −649.1, −688.1, and −843.7, respectively. In Gupis station 4p beta has a minimum value of BIC (−788.7) for scale−1. In other scales (3, 6, 9, 12, and 24), the gamma, Gumbel, Johnson SU, and scaled/shifted t have minimum values of BIC −1264.9, −1305.4, 1519.0, −937.6, and 1408.0, respectively.

Accordingly, various distributions are selected for various time scales for the other three stations (Chilas, Gilgit, and Skardu) (1, 3, 6, 9, 12, and 24). After standardization with a selected probability distribution, the next step is the classification of the SPI for various drought states (Table 2). In Figure 3, the temporal behavior of the SPI at scale-1 is presented for various stations. However, the behavior of SPI

for other selected time scales can be presented accordingly. After calculating values to quantify SPI at various time scales, we first categorized SPI for its magnitude. The behavior of several drought classes for SPI at a one-month time scale in selected stations is provided in Table 3, where the observations are taken as an example for various months of the year 2017. Accordingly, the behavior of several drought classes for other years for different time scales is calculated. These observed drought classes are further used to find the probability distribution associated with the three-step transition from the last state in the various sequences. The posterior vector related to these sequences specifies the parameter values (briefly described in Section 2.2). The obtained results show that the most likely state to visit in three steps is ND, which means the probability associated with ND is higher than the other selected states in varying sequences (Table 4). For example, for the Astore station, in sequence-1, the value indicates that the probability of ND occurrence is 0.6668, which is higher than other selected states. This probability of occurrence for ND can be observed from other sequences. Further in sequence-2, the probability of ND occurrence after three months is 0.6729. Moreover, the probabilities of ND in sequence-3, sequence-4, sequence-5, and sequence-6 are 0.6611, 0.6221, 0.6450, and 0.6729, respectively. It means that the policymakers should make their plans accordingly for ND. Other information can
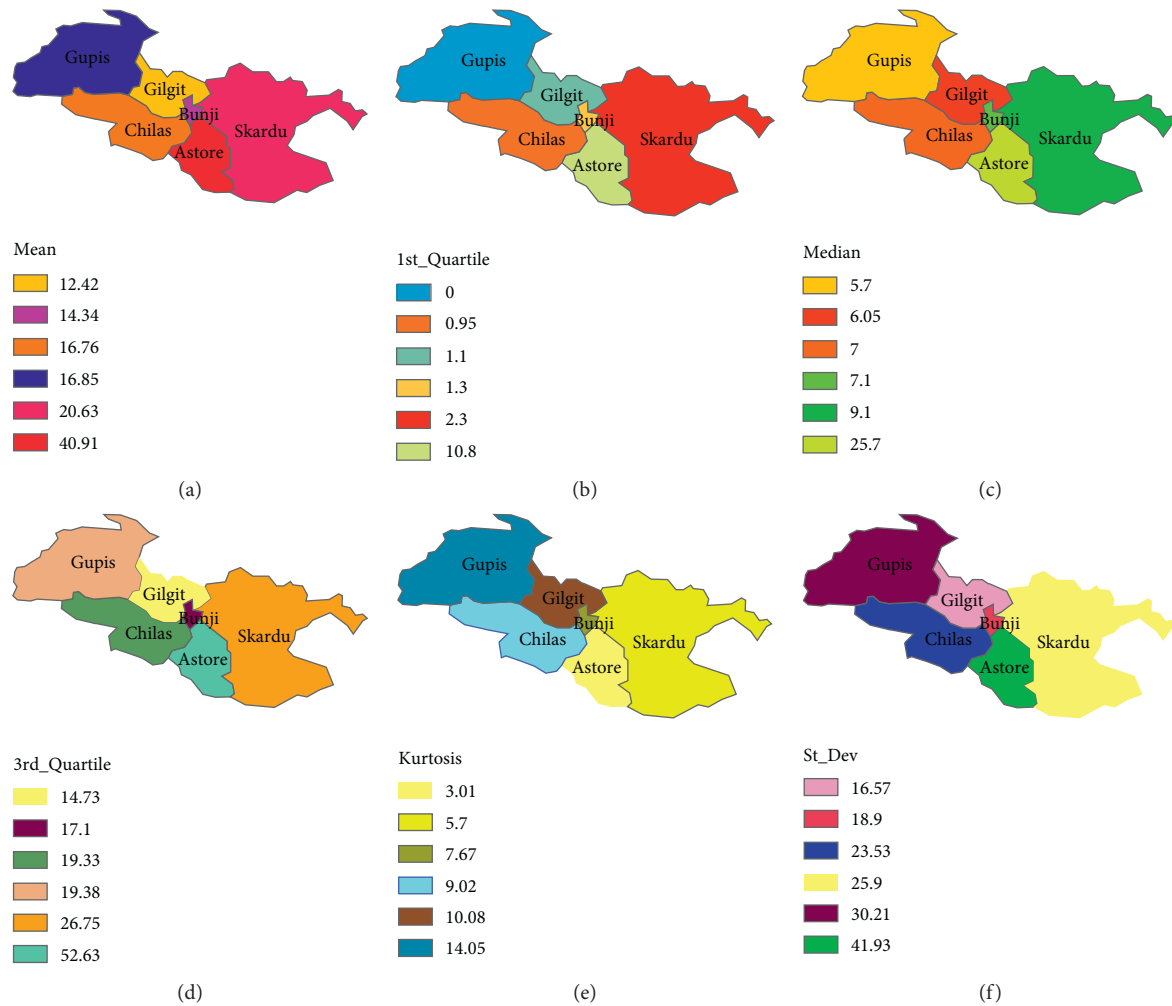
FIGURE 2: The climatological characteristics of the precipitation to describe the period from January 1971 to December 2017 of specified stations. The (a) mean, (b) 1st-quartile, (c) median, (d) 3rd_quartile, (e) kurtosis, and (f) st_dev.

be observed from the various sequences for different time scales. However, the ND is prevailing in all time scales in the selected region. So, the policymakers should work to mitigate negative impacts for this specific drought state (ND).

3.2. Discussion. The time series data were collected from six meteorological stations in the northern area of Pakistan. The drought index SPI is used for the analysis with various time scales for selected stations. The reliable and efficient outcomes of the analysis provide strong indications about the drought occurrences that can significantly help for an early warning system [31, 53, 58, 59, 61]. Therefore, a new MBCSTCS method is developed for the drought monitoring and mitigation policies that explicitly envisage spatiotemporal information. The proposed technique uses the long-run behavior of drought states (categorical sequences) from various time scales and stations in the selected region. If a time scale changes, then the categorical sequence sizes are affected. Therefore, in past studies, researchers have not been studied various time scales accumulatively due to inconsistency in their sizes and the phenomenon that has

generated the observations for varying stations. However, these issues are being resolved effectively by the current technique. Furthermore, the outcomes associated with the present technique help to accomplish the current objective and provide more substantial outcomes for the selected drought states based on varying time scales and stations. MBCSTCS uses state selection procedures through finite mixture modeling and model-based clustering. Niaz et al. [55] developed a new model-based clustering technique that predicts probabilities for various drought classes. They computed categorical drought state (classes) sequences for selected drought classes and predicted their probabilities for the future. The mentioned study used a single time scale on various stations. However, in this study, the varying time scales are accumulatively measured for the monthly prediction of drought severity in selected stations. Therefore, it is a novel method for predicting drought severity using spatiotemporal categorical sequences. MBCSTCS is applied to six meteorological stations in the northern area of Pakistan. It is found that MBCSTCS provides expeditious information for the long-term spatiotemporal categorical sequences. The present analysis results are more suitable,

TABLE 1: The BIC of selected probability distributions for SPI at various timescales (1, 3, 6, 9, 12, and 24) for the selected stations

Index	Scale	Astore Distribution	Astore BIC	Bunji Distribution	Bunji BIC	Gupis Distribution	Gupis BIC	Chilas Distribution	Chilas BIC	Gilgit Distribution	Gilgit BIC	Skardu Distribution	Skardu BIC
SPI	1	3P Weibull	−1036.5	3p Weibull	−1031.0	4p beta	−788.7	4P beta	−805.6	3P Weibull	−1097.4	3P Weibull	−735.1
	3	Gamma	−1279.1	Gamma	−824.9	Gamma	−1264.9	3P Weibull	−810.2	Exponential	−824.2	Generalized extreme value	−653.9
	6	Gamma	−892.8	Skewed-normal	−1162.2	Gumbel	−1305.4	Triangular	−1190.6	Gamma	−1204.8	Log–normal	−1018.5
	9	Gamma	−896.1	Normal	−649.1	Johnson SU	−1519.0	Normal	−1266.7	Laplace	−1206.3	Log–normal	−1109.6
	12	Cosine	−913.3	Laplace	−688.1	Johnson SU	−937.6	Laplace	−1325.8	Inverse gamma	−1266.4	Gamma	−1025.7
	24	Skewed-normal	−1294.4	Laplace	−843.7	Scaled/shifted t	−1408.0	Laplace	−826.8	Trapezoidal	−681.7	Inverse Gaussian	−926.8

TABLE 2: The classification of the SPI for various drought states

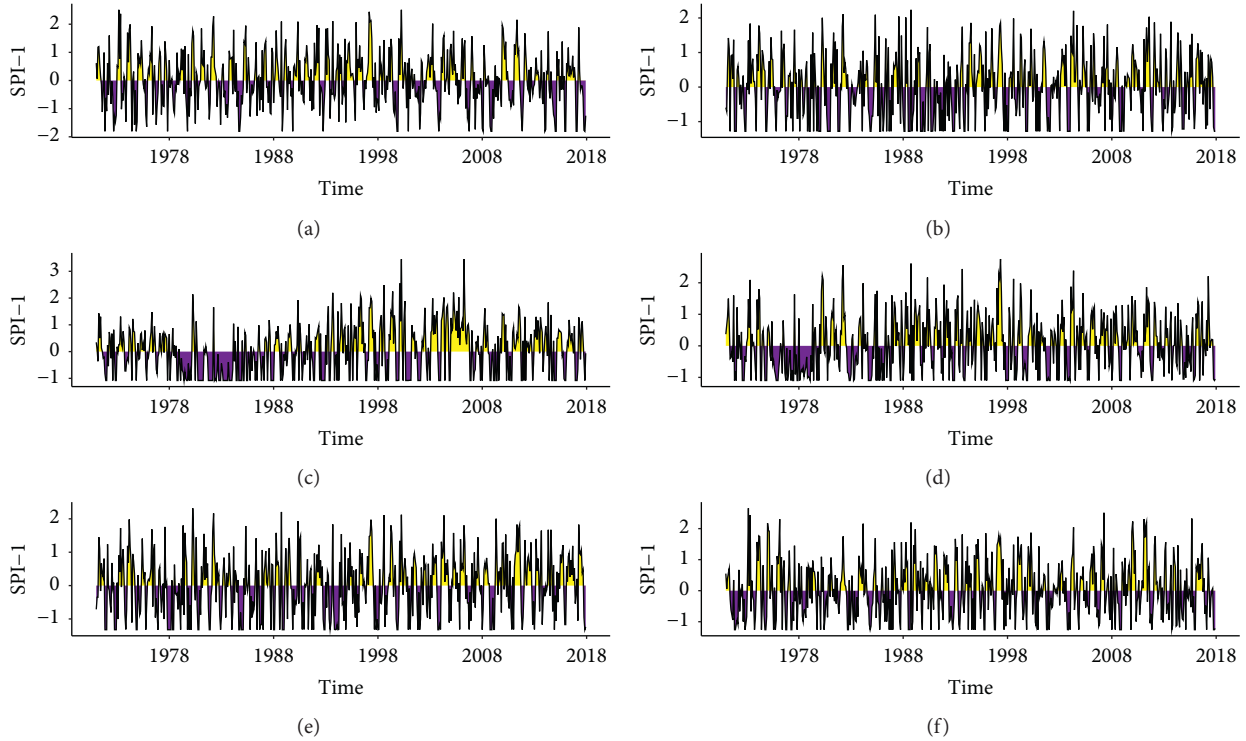| SPI	Observed drought states
SPI ≤ −2	Extremely dry (ED)
SPI > −2 and SPI ≤ −1.5	Severely dry (SD)
SPI > −1.5 and SPI ≤ −1	Median dry (MD)|
SPI > −1 and SPI ≤ 1	Normal dry (ND)
SPI > 1 and SPI ≤ 1.5	Median wet (MW)
SPI > 1.5 and SPI ≤ 2	Severely wet (SW)
SPI > 2	Extremely wet (EW)

(a)

(b)

(c)

(d)

(e)

(f)

FIGURE 3: The temporal presentation of SPI at scale-1 for varying stations. (a) Astore, (b) Bunji, (c) Gupis, (d) Chilas, (e) Gilgit, and (f) Skardu.

TABLE 3: The behavior of various drought classes for SPI at scale-1 in selected stations.

Month	Astore		Bunji		Gupis		Chilas		Gilgit		Skardu	
	SPI	Class	SPI	Class	SPI	Class	SPI	Class	SPI	Class	SPI	Class
Jan	−1.8044	SD	−1.2152	MD	−1.0805	MD	−0.3254	ND	−0.230	ND	−0.4827	ND
Feb	−1.8044	SD	0.2629	ND	−1.0805	MD	0.8075	ND	0.201	ND	−1.2581	ND
Mar	−0.4745	ND	−1.1059	MD	−0.6655	ND	−0.1535	ND	−1.052	MD	−0.3096	ND
Apr	1.8879	SW	0.9377	ND	1.2811	MW	2.2094	EW	1.828	SW	1.0615	EW
May	0.3314	ND	0.6975	ND	0.9813	ND	0.8297	ND	0.972	ND	0.0796	ND
Jun	−0.3302	ND	−0.6380	ND	−0.2883	ND	−0.0637	ND	0.931	ND	−0.8574	ND
Jul	−0.1857	ND	0.8192	ND	0.7872	ND	0.1366	ND	1.057	MW	−0.2401	ND
Aug	−0.1967	ND	0.7181	ND	1.1317	MW	0.2058	ND	0.386	ND	0.0086	ND
Sep	−0.5938	ND	0.2824	ND	0.4029	ND	0.1937	ND	0.476	ND	0.0739	ND
Oct	−1.7685	SD	−1.1576	MD	−0.2883	ND	−1.0539	MD	−0.957	ND	−1.2581	MD
Nov	−1.8044	SD	−1.2806	MD	−1.0805	MD	−1.0995	MD	−1.323	MD	−1.2581	MD
Dec	−1.2499	MD	−1.2806	MD	−0.0621	ND	−1.0539	MD	−1.240	MD	−1.2168	MD

Note. The observations are considered for varying months of the year 2017. The specific year is selected as an example. Accordingly, the behavior of various drought classes for other years for various time scales is calculated.

TABLE 4: Predicting future three-step transitions from the last state for various drought sequences at Astore station

	Three-month prediction in Astore						
Sequence	ED	SD	MD	ND	MW	SW	EW
1	0.0014	0.0766	0.0965	0.6668	0.0950	0.0409	0.0228
2	0.0010	0.0691	0.0927	0.6729	0.0962	0.0436	0.0241
3	0.0017	0.0829	0.0999	0.6611	0.0941	0.0388	0.0215
4	0.0065	0.0127	0.2198	0.6221	0.0815	0.0437	0.0137
5	0.0015	0.0014	0.1987	0.6450	0.0878	0.0506	0.0150
6	0.0014	0.0691	0.0927	0.6729	0.0962	0.0436	0.0241

especially for the selected region, and help the policymakers make better policies related to various kinds of drought, including meteorological, hydrological, and socioeconomic. The MBCSTCS may help to make plans for early warning systems, water resource management, and drought mitigation policies to reduce the severe effects of drought.

4. Conclusions

The outcomes of MBCSTCS provide the future probabilities corresponding to each of the drought states in varying stations and time scales. The obtained outcomes show that the most likely state to visit is ND, which means the probability associated with ND is higher than the other selected states in varying sequences. For instance, in sequence-1, the value shows that the probability of ND is 0.6668, which is higher than other selected states. Further in sequence-2, the probability of ND after three months is 0.6729. This probability of ND also prevails in other sequences. Furthermore, in sequence-6, the ND has a higher probability (0.6729) of occurrence in the future. Therefore, policymakers should work to reduce the negative impacts of this drought state (ND). In conclusion, this study suggests a more appropriate technique that emphasizes evaluating drought occurrences more instantaneously. The MBCSTCS helps the policymakers to make better policies related to various kinds of drought, including meteorological, hydrological, and socioeconomic. Furthermore, the analysis provides the basis to bring more attention to early warning systems. Moreover, the outcomes of the current analysis are only capable of transmitting in the present circumstances of the application site, as the circumstance of the selected stations will change the influence of the outcome for the extrapolations. Furthermore, the study can find some propagations and compute several thresholds for different drought severities for the region. Moreover, other drought indices can be incorporated to envisage the drought occurrences effectively.

Data Availability

The data used for the preparation of the manuscript are available from the corresponding author and can be provided upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

References

[1] E. Gidey, O. Dikinya, R. Sebego, E. Segosebe, and A. Zenebe, "Predictions of future meteorological drought hazard (~2070) under the representative concentration path (RCP) 4.5 climate change scenarios in Raya, Northern Ethiopia," *Modeling Earth Systems and Environment*, vol. 4, no. 2, pp. 475–488, 2018.

[2] M. Trnka, P. Hlavinka, M. Možný et al., "Czech drought monitor system for monitoring and forecasting agricultural drought and drought impacts," *International Journal of Climatology*, vol. 40, no. 14, pp. 5941–5958, 2020.

[3] M. Jehanzaib, J. Yoo, H.-H. Kwon, and T.-W. Kim, "Reassessing the frequency and severity of meteorological drought considering non-stationarity and copula-based bivariate probability," *Journal of Hydrology*, vol. 603, Article ID 126948, 2021.

[4] R. Salam, A. R. M. Towfiqul Islam, B. K. Shill et al., "Nexus between vulnerability and adaptive capacity of drought-prone rural households in northern Bangladesh," *Natural Hazards*, vol. 106, no. 1, pp. 509–527, 2021.

[5] P. Kalura, A. Pandey, V. M. Chowdary, and P. V. Raju, "Assessment of hydrological drought vulnerability using geospatial techniques in the tons river basin, India," *Journal of the Indian Society of Remote Sensing*, vol. 49, pp. 1–15, 2021.

[6] A. H. Payab and U. Türker, "Comparison of standardized meteorological indices for drought monitoring at northern part of Cyprus," *Environmental Earth Sciences*, vol. 78, no. 10, pp. 1–19, 2019.

[7] O. O. Ayantobo and J. Wei, "Appraising regional multi-category and multi-scalar drought monitoring using standardized moisture anomaly index (SZI): a water-energy balance approach," *Journal of Hydrology*, vol. 579, Article ID 124139, 2019.

[8] H. Salimi, E. Asadi, and S. Darbandi, "Meteorological and hydrological drought monitoring using several drought indices," *Applied Water Science*, vol. 11, no. 2, pp. 1–10, 2021.

[9] O. V. Wilhelmi, K. G. Hubbard, and D. A. Wilhite, "Spatial representation of agroclimatology in a study of agricultural drought," *International Journal of Climatology*, vol. 22, no. 11, pp. 1399–1414, 2002.

[10] S. Jiang, L. Wei, L. Ren et al., "Utility of integrated IMERG precipitation and GLEAM potential evapotranspiration products for drought monitoring over mainland China," *Atmospheric Research*, vol. 247, Article ID 105141, 2021.

[11] F. Hui-Mean, Z. Yusop, and F. Yusof, "Drought analysis and water resource availability using standardised precipitation

evapotranspiration index," *Atmospheric Research*, vol. 201, pp. 102–115, 2018.

[12] Y. Lin, F. Liu, and P. Xu, "Effects of drought on infant mortality in China," *Health Economics*, vol. 30, no. 2, pp. 248–269, 2021.

[13] 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.

[14] D. A. Wilhite, *Drought as a Natural hazard: Concepts and Definitions*, Springer, Berlin, Germany, 2000.

[15] T. J. Chang, "Investigation of precipitation droughts by use of kriging method," *Journal of Irrigation and Drainage Engineering*, vol. 117, no. 6, pp. 935–943, 1991.

[16] E. A. B. Eltahir, "Drought frequency analysis of annual rainfall series in central and western Sudan," *Hydrological Sciences Journal*, vol. 37, no. 3, pp. 185–199, 1992.

[17] T. J. Chang and X. A. Kleopa, "A proposed method for drought monitoring," *Journal of the American Water Resources Association*, vol. 27, no. 2, pp. 275–281, 1991.

[18] T. Estrela and E. Vargas, "Drought management plans in the European Union. The case of Spain," *Water Resources Management*, vol. 26, no. 6, pp. 1537–1553, 2012.

[19] Y. B. Zecharias and W. Brutsaert, "The influence of basin morphology on groundwater outflow," *Water Resources Research*, vol. 24, no. 10, pp. 1645–1650, 1988.

[20] S. Mohan and N. C. V. Rangacharya, "A modified method for drought identification," *Hydrological Sciences Journal*, vol. 36, no. 1, pp. 11–21, 1991.

[21] B. Clausen and C. P. Pearson, "Regional frequency analysis of annual maximum streamflow drought," *Journal of Hydrology*, vol. 173, no. 1-4, pp. 111–130, 1995.

[22] W. C. Palmer, "Keeping track of crop moisture conditions, nationwide: the new Crop Moisture Index," *Weatherwise*, vol. 21, no. 4, pp. 156–161, 1968.

[23] H. N. Bhalme and D. A. Mooley, "Large-scale droughts/floods and monsoon circulation," *Monthly Weather Review*, vol. 108, no. 8, pp. 1197–1211, 1980.

[24] 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.

[25] 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*, pp. 179–183, Boston, MA, USA, 1993 January.

[26] K. M. Weghorst, *The Reclamation Drought Index: Guidelines and Practical Applications*, Bureau of Reclamation, Denver, CO, USA, 1996.

[27] W. C. Palmer, *Meteorological Drought*, US Department of Commerce, Washington, DC, USA, 1965.

[28] R. Gommes and F. Petrassi, *Rainfall Variability and Drought in Sub-saharan Africa. SD Dimensions*, FAO, Rome, Italy, 1996.

[29] M. P. Van Rooy, *A Rainfall Anomaly Index Independent Of Time And Space*, NOTOS, Kifisia, Greece, 1965.

[30] J. González and J. B. Valdés, "New drought frequency index: definition and comparative performance analysis," *Water Resources Research*, vol. 42, no. 11, 2006.

[31] A. K. Mishra and V. P. Singh, "A review of drought concepts," *Journal of Hydrology*, vol. 391, no. 1-2, pp. 202–216, 2010.

[32] A. Zargar, R. Sadiq, B. Naser, and F. I. Khan, "A review of drought indices," *Environmental Reviews*, vol. 19, pp. 333–349, 2011.

[33] A. K. Mishra and V. P. Singh, "Drought modeling–A review," *Journal of Hydrology*, vol. 403, no. 1-2, pp. 157–175, 2011.

[34] B. Amirataee and M. Montaseri, "The performance of SPI and PNPI in analyzing the spatial and temporal trend of dry and wet periods over Iran," *Natural Hazards*, vol. 86, no. 1, pp. 89–106, 2017.

[35] M. M. Moghimi and A. R. Zarei, "Evaluating performance and applicability of several drought indices in arid regions," *Asia-Pacific Journal of Atmospheric Sciences*, vol. 57, no. 3, pp. 645–661, 2021.

[36] 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.

[37] 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.

[38] R. Xu and D. Wunsch, "Survey of clustering algorithms," *IEEE Transactions on Neural Networks*, vol. 16, no. 3, pp. 645–678, 2005.

[39] Y. W. Teh, M. I. Jordan, M. J. Beal, and D. M. Blei, "Hierarchical dirichlet processes," *Journal of the American Statistical Association*, vol. 101, no. 476, pp. 1566–1581, 2006.

[40] U. Von Luxburg, "A tutorial on spectral clustering," *Statistics and Computing*, vol. 17, no. 4, pp. 395–416, 2007.

[41] S. B. Dalirsefat, 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*, vol. 9, no. 1, pp. 1–8, 2009.

[42] J. H. Ward, "Hierarchical grouping to optimize an objective function," *Journal of the American Statistical Association*, vol. 58, no. 301, pp. 236–244, 1963.

[43] 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.

[44] V. Melnykov, "Challenges in model-based clustering," *Wiley Interdisciplinary Reviews: Computational Statistics*, vol. 5, no. 2, pp. 135–148, 2013.

[45] 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.

[46] M. S. Handcock, A. E. Raftery, and J. M. Tantrum, "Model-based clustering for social networks," *Journal of the Royal Statistical Society: Series A*, vol. 170, no. 2, pp. 301–354, 2007.

[47] J. D. Banfield and A. E. Raftery, "Model-based Gaussian and non-Gaussian clustering," *Biometrics*, vol. 49, no. 3, pp. 803–821, 1993.

[48] G. McLachlan and D. Peel, "Mixtures of factor analyzers," in *Proceedings of the 7th International Conference on Machine Learning*, Bellevue, WA, USA, June 2000.

[49] T. Warren Liao, "Clustering of time series data-a survey," *Pattern Recognition*, vol. 38, no. 11, pp. 1857–1874, 2005.

[50] W.-C. Chen and R. Maitra, "Model-based clustering of regression time series data via APECM-an AECM 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.

[51] A. K. Mishra, V. R. Desai, and V. P. Singh, "Drought forecasting using a hybrid stochastic and neural network model," *Journal of Hydrologic Engineering*, vol. 12, no. 6, pp. 626–638, 2007.

[52] V. Mishra, A. D. Tiwari, S. Aadhar et al., "Drought and f," *Geophysical Research Letters*, vol. 46, no. 4, pp. 2075–2083, 2019.

[53] R. R. Heim, "A review of twentieth-century drought indices used in the United States," *Bulletin of the American Meteorological Society*, vol. 83, no. 8, pp. 1149–1166, 2002.

[54] B. Lloyd-Hughes and M. A. Saunders, "A drought climatology for Europe," *International Journal of Climatology: A Journal of the Royal Meteorological Society*, vol. 22, no. 13, pp. 1571–1592, 2002.

[55] R. Niaz, I. M. Hussain, Z. Ali et al., "A novel spatially weighted accumulative procedure for regional drought monotoring," *Tellus A: Dynamic Meteorology and Oceanography*, vol. 72, 2020.

[56] A. P. Dempster, N. M. Laird, and D. B. Rubin, "Maximum likelihood from incomplete data via theEMAlgorithm," *Journal of the Royal Statistical Society: Series B*, vol. 39, no. 1, pp. 1–22, 1977.

[57] 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.

[58] 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.

[59] 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.

[60] G. Schwarz, "Estimating the dimension of a model," *Annals of Statistics*, vol. 6, no. 2, pp. 461–464, 1978.

[61] S. M. Vicente-Serrano, S. Beguería, and J. I. López-Moreno, "A multiscalar drought index sensitive to global warming: the standardized precipitation evapotranspiration index," *Journal of Climate*, vol. 23, no. 7, pp. 1696–1718, 2010.