Spatiotemporal Analysis of Droughts Over Different Climate Regions Using Hybrid Clustering Method

Kiyoumars Roushangar1,2 · Roghayeh Ghasempour1 · Vahid Nourani1,2,3

Received: 2 April 2021 / Accepted: 15 September 2021 / Published online: 23 September 2021
© The Author(s), under exclusive licence to Springer Nature B.V. 2021, corrected publication 2021

Abstract
Assessment of spatiotemporal variations of drought is an efficient method for implementing drought mitigation strategies and reducing its negative impacts. This study aimed to assess the spatiotemporal pattern of short- to long-term droughts for an area with different climates. Therefore, 31 stations located in Iran were selected and the Standardized Precipitation Index (SPI) series with timescales of 3, 6, and 12 months were computed during the 1951-2016 period. A hybrid methodology including Maximal Overlap Discrete Wavelet Transform (MODWT) and K-means methods was used for obtaining the SPIs time-frequency properties and multiscale zoning of the area. The energy amounts of the decomposed subseries via the MODWT were applied as inputs for K-means. Also, the statistics in drought features (i.e. drought duration, severity, and peak) were assessed and the results showed that shorter term droughts (i.e. SPI-3 and -6) were more frequent and severe in the northern parts where the lowest values were obtained for drought duration. It was observed that the regions with more droughts frequency had the highest energy values. For shorter term droughts a direct relationship was obtained between the energy values and the mean SPIs, drought severity, and drought peak, whereas an inverse relationship was obtained for longer term drought. It was found that increasing the degree of SPI led to more similarity between the stations of each cluster. Also, the homogeneity of stations for the SPI-12 was slightly higher than the SPI-3 and -6.

Keywords Different climate · Drought features · Homogeneity · MODWT · Spatiotemporal analysis

1 Introduction
Drought is considered as a natural disaster and its effects accumulate slowly over a long period (Mann and Gleick 2015; Marvel et al. 2019). This extreme event has adverse effects on water supply and quality, agricultural productivity, land degradation, public...
health, desertification, and famine (Nanzad et al. 2019). The widespread undesirable impacts of drought can be seen anywhere in the world (Zhou et al. 2020). Generally, due to the invidious nature of drought, its definition and understanding is difficult (Vicente-Serrano 2007; Van Loon 2015; Dadson et al. 2019). Monitoring the spatiotemporal variation of drought makes it possible to reduce vulnerability and establish adaptation strategies (Chanda et al. 2014). There are different techniques for quantitatively analysis of drought phenomenon and among them; drought indices have been widely applied for this aim (Kwak et al. 2016; Mukherjee et al. 2018) such as Evapotranspiration Deficit Index (EDI), Surface Water Supply Index (SWSI), Palmer Drought Severity Index (PDSI), Standardized Precipitation Index (SPI), Standardized Runoff Index (SRI), and Standardized Precipitation Evapotranspiration Index (SPEI). Also, Tigkas et al. (2020) introduced the Crop Reconnaissance Drought Index (CRDI) as a new drought index. This index is an adjustment of the Reconnaissance Drought Index (RDI), in which the crop evapotranspiration is replaced by reference evapotranspiration parameter. The SPI is an ideal tool in drought severity characterization (Li et al. 2015; Kwak et al. 2016). Nowadays, this index is known as one of the most common methods for drought monitoring. The SPI index has statistical consistency, and is capable to show short- and long-term drought impacts on precipitation anomalies (Tigkas et al. 2019).

According to Madadgar and Moradkhani (2013), the required information for drought preparedness can be provided by continuous drought monitoring. This information accuracy depends directly on the applied models’ efficiency. Numerous methods including regression, time series, artificial intelligence, and physical approaches have been developed to monitor drought events (Madadgar and Moradkhani 2013; Surendran et al. 2019; Sakizadeh et al. 2019). Although these methods showed promising results in modeling of drought; however, due to the undesirable effects of droughts, it is necessary to use more efficient methods. Also, since most of the hydrological signals have seasonal fluctuations or are non-stationary, the time series preprocessing approaches are commonly used for decomposition and exca-
vation of complex, irregular, and periodic hydrological signals (Roushangar and Alizadeh 2018). The Discrete Wavelet Transform (DWT) is a common approach which used by many researchers for time series decomposition (Percival and Walden 2000; Agarwal et al. 2016). The use of the DWT in decomposition of non-stationary signals provides beneficial information in both temporal and frequency domains (Roushangar et al. 2019. The modified version of DWT is Maximal Overlap Discrete Wavelet Transform (MODWT). In this approach, time series should be projected onto a set of non-orthogonal basis functions in order to generate the wavelet coefficients. In the current study, drought series were broken down via the MODWT.

In addition, clustering methods are used to identify homogenous drought regions. In fact, clustering is the classification of a large number of data into a smaller number of groups so that data belonging to the same cluster are as similar as possible while, they have heterogeneity with the data belonging to other groups (Jain et al. 1999; Barton et al. 2016). The clustering approaches regard the spatiotemporal domains of components; so, the obtained outcomes from them indicate the local characteristics and temporary variations. The K-means is among the most popular clustering techniques.

Drought is one of the challenging natural disasters in distressing the lives of people directly or indirectly. Monitoring the spatiotemporal variation of drought makes it possible to reduce vulnerability and establish adaptation strategies. Therefore, in this study, a multi-spatiotemporal approach based on the short- to long-term SPIs time series was proposed to analysis the drought characteristics in Iran. Moreover, due to non-stationary properties of the drought series, the statistics in drought features including severity, peak,
Spatiotemporal Analysis of Droughts Over Different Climate…

and duration were calculated to assess the spatiotemporal variations in drought features. Iran is a country in Asia that has regions with different climates. However, arid and semi-arid climates are predominant in most regions and therefore, drought is a major concern across these regions. The applied procedure was based on the SPIs series decomposition via the MODWT technique. First, the SPI with 3, 6 and 12 months timescales (i.e. the SPI-3, SPI-6, and SPI-12 for the short, mid, and long term analyses, respectively) were calculated using precipitation data of 31 stations located in different climates of Iran during the period of 1951-2016. Then, the statistics in drought features including severity, peak, and duration were calculated to assess the spatiotemporal variations in drought features. In the next step, the SPIs series were broken down to various periodicity scales using the MODWT approach. The energy values of subseries were computed and the subseries with less energy variations were used for spatial categorization of the selected area via the K-means method. Also, it was tried to determine a relationship between the energy values of subseries and characteristics of drought. The results of this study can help to better understanding the drought behavior. Further, the proposed method provides valuable references for drought assessment and management, specifically at the regional level, and gives a better quantitative way of analyzing drought with varying drought intensities.

2 Materials and Methods

2.1 Study Area

In the current study, Iran which is a country in west of Asia was selected as a case study. The Caspian Sea is located in the north of Iran, and the Persian Gulf and sea of Oman are located in the south of country. Iran’s climate is known as arid or semi-arid and the annual mean precipitation is approximately 250 mm. Generally, due to large annual rainfall and temperature variations across the country, its climate is very diverse (Dinpashoh 2006). The higher altitudes in the mountains have the most precipitation where the temperature is the lowest. The amount of precipitation gradually reduces from the northwest to the southeast, and vice versa, the amount of temperature increases. In this study, monthly precipitation data from 31 sites obtained from Iranian Meteorological Organization were used for drought modeling in regions with different climates. The data at first were checked for accuracy and homogeneity. Table 1 and Fig. 1 show the selected sites geographical locations and their relevant information.

2.2 Proposed Methodology

In the current study, the spatiotemporal variations of short- to long-term droughts were assessed for Iran. For this aim, as shown in Fig. 2, in the first step after data collections and calculating the SPI-3, SPI-6 and SPI-12 series, their standard deviation (STD), mean values, and the number of drought events were computed for different parts of study area. Also, drought features including peak, duration, and severity were computed. In the second step, the MODWT was applied in order to decompose the drought series. The subseries energy values were computed and used for categorization of the area. Also, the relationships between the characteristics of drought and energy values were investigated for the SPIs with different degrees. Finally, the energy values of the selected series were considered as inputs of the K-means and the area clustering was
### Table 1

Properties of the used stations

| Station         | Label | Longitude | Latitude | Mean rainfall | Station         | Label | Longitude | Latitude | Mean rainfall |
|-----------------|-------|-----------|----------|---------------|-----------------|-------|-----------|----------|---------------|
| Khoi            | 1     | 44.908    | 38.617   | 24.144        | Kermanshah      | 17    | 46.645    | 34.425   | 37.282        |
| Urmia           | 2     | 44.908    | 37.546   | 27.524        | Arak            | 18    | 49.188    | 34.145   | 27.859        |
| Tabriz          | 3     | 46.526    | 37.784   | 23.533        | Khorraram abad  | 19    | 48.51     | 33.586   | 41.654        |
| Zanjan          | 4     | 46.468    | 36.55    | 25.198        | Dezful          | 20    | 48.353    | 32.838   | 32.712        |
| Rasht           | 5     | 50.096    | 37.261   | 110.966       | Esfahan         | 21    | 52.694    | 33.181   | 10.041        |
| Ramsar          | 6     | 50.833    | 36.785   | 100.547       | Shahre kord    | 22    | 50.452    | 32.41    | 27.418        |
| Babolsar        | 7     | 52.537    | 36.68    | 74.538        | Yazd           | 23    | 55.549    | 32.224   | 4.767         |
| Gorgan          | 8     | 54.26     | 36.956   | 47.972        | Birjand         | 24    | 59.576    | 32.673   | 13.828        |
| Ghazvin         | 9     | 49.843    | 36.1     | 26.248        | Ahwaz           | 25    | 49.053    | 31.353   | 18.833        |
| Tehran          | 10    | 51.66     | 35.787   | 19.444        | Abadan          | 26    | 48.411    | 30.282   | 12.862        |
| Mashhad         | 11    | 59.146    | 36.568   | 20.836        | Shiraz          | 27    | 52.18     | 29.897   | 20.474        |
| Shahrood        | 12    | 55.836    | 35.775   | 13.135        | Kerman          | 28    | 56.58     | 30.15    | 11.695        |
| Sabzevar        | 13    | 58.01     | 35.51    | 15.653        | Zahedan         | 29    | 60.831    | 29.597   | 6.785         |
| Torbat Heydarie | 14    | 59.466    | 35.196   | 22.313        | Bushehr         | 30    | 50.952    | 28.94    | 20.261        |
| Sanandaj        | 15    | 47.178    | 35.738   | 37.614        | Bandarabas      | 31    | 56.42     | 27.213   | 14.159        |
| Hamedan         | 16    | 48.492    | 34.786   | 26.970        |                 |       |           |          |               |
done for identifying the regions with more similarities. In the following parts of the study the used materials were explained.

2.3 The Standardized Precipitation Index

The Standardized Precipitation Index (SPI) has been commonly used by researchers for drought monitoring. The McKee et al. (1993) defined the SPI to show the abnormal wetness and dryness. In fact, they designed the SPI to investigate the rainfall deficiency in multiple timescales. SPI is capable of showing rainfall anomalies in short- to long-term timescales. According to Mirabbasi et al. (2013), via the SPI different timescales the effects of rainfall deficit on different water resources components can be investigated. For
instance, soil moisture conditions indicate the short-term rainfall anomalies, while groundwater, reservoir storage, and streamflow respond to rainfall anomalies on a relatively long scale (Mirabbasi et al. 2013). Positive or negative amounts can be obtained for the SPI series, in which negative values indicate the occurrence of drought event and positive values show that the drought ends. Table 2 shows the descriptive features of the SPI. In this study the SPIs values less than zero were considered for drought events detecting. Severity, peak, and duration are three significant characteristics of droughts. The characteristics of drought event are shown in Fig. 3. Drought duration is the period when the SPI is continuously negative. Also, drought severity and peak can be calculated as:

\[
\text{Peak} = \text{Min}(\text{SPI}_D) \tag{1}
\]

\[
\text{Severity} = -\sum_{i=1}^{D} \text{SPI}_i \tag{2}
\]

where \(D\) is drought duration.

### 2.4 Maximal Overlap Discrete Wavelet Transform (MODWT)

Pre-processing approaches emphasize on original time series breaking down into stationary and more regular subseries in which are more explicit to analyze by omitting the redundant and irrelevant properties of the signals (Cornish et al. 2006; Roushangar et al. 2019). This process will enhance the quality of the time series. Maximal Overlap Discrete Wavelet
Transform (MODWT) is one of time series pre-processing techniques. This method is the modified version of the DWT and can be applied for original time series decomposition into an approximation component (scaling coefficients) and many detail components (wavelet coefficients). The scaling coefficients contain the low-frequency information which is the most crucial section in providing the signal identity, while the wavelet coefficients reveal flavor of the signal (Sujjaviriyasup 2017). Via the MODWT, a Multi Resolution Analysis (MRA), which allows for the decomposition of the time series into a sum of simpler time series, can be developed (Sujjaviriyasup 2017). The formula of MODWT is as bellow:

\[ X = \sum_{j=1}^{L} W_j + T_N \]  

in which \( X \) is time series, \( W_j \) is wavelet details, \( T_N \) is the overall trend of original signal, and \( L \) is the number of decomposition level. According to Dghais and Ismail (2013), the MODWT is non-orthonormal and shift-invariant and can be used for time series with different sizes. For more details on MODWT see Percival and Walden (2000). The subseries decomposed via the MODWT method (i.e. scaling and wavelet coefficients) were used as inputs to categorize the selected stations. However, in order to decrease the number of inputs, the energy amounts of these subseries were computed using Eq. (4). Finally, the subseries energy values \( E \) were applied as clustering method basis.

\[ E = \frac{\sum (SP_m^2)}{\left( \sum SP_m \right)^2} \quad m = 1, 2 \ldots \]  

where \( SP \) shows the SPIs amount decomposed via the MODWT and \( m \) is the SPIs month.

### 2.5 K-means Approach

The aim of clustering techniques is to divide a set of patterns into separate clusters. In this process, the patterns with the similar features are located in the same cluster, while different patterns are assigned to other clusters. Among different clustering techniques, the k-means generally applied for data mining cluster analysis and vector quantization (Pham et al. 2005). Using the k-means, \( n \) observations are classified into \( k \) clusters and in this process, each entity (observation) is assigned to the cluster with the nearest mean (Pham et al. 2005; Yang and Deng 2010). Based on Likas et al. (2003), the following steps should be performed in data clustering through the k-means method:

1. Selecting an initial number of clusters and define the centers;
2. Assigning each observation to a cluster with the closest center considering its Euclidean distance;
3. Reordering the centers’ positions after assigning all observation to one cluster;
4. Repeating steps 2 and 3 until the memberships of cluster do not vary.

Due to the k-means ability in large datasets clustering as well as its implementation in a short time, it is one of the popular clustering methods.

The outcomes of spatial clustering were evaluated based on the Silhouette coefficient (SCi) validation metric. The SCi amounts indicate the degree of similarity for samples
within the cluster. According to Eq. (5), the range of SCi varies between -1 and 1 and higher values indicate that the samples are well-matched to their own clusters (Hsu and Li 2010).

$$\text{SCi} = \frac{b(i) - a(i)}{\text{max}[a(i), b(i)]}$$  \hspace{1cm} (5)

In Eq. (5), $a(i)$ is the average dissimilarity of $i$ sample with other samples of the same cluster and $b(i)$ is the lowest average dissimilarity of $i$ sample with other clusters.

### 3 Results and Discussion

#### 3.1 Spatiotemporal Pattern of the SPIs Time Series and Drought Features

Study of monthly rainfall time series in the considered stations showed that the average rainfall during the 65-year statistical period differs from the lowest value with 4.6 mm to the highest value with 143 mm. Comparison of rainfall data with the mean amounts showed that frequent droughts had occurred in the considered stations. The SPIs with 3, 6 and 12 months timescales were calculated for all selected stations located in various parts of Iran. In Fig. 4a, the SPIs series are shown for the Abadan station. The temporal drought characteristics of the SPIs time series showed that different levels of drought occurred in different parts of the selected area. It was find that the most extreme drought period lasted 48 months from 1999-2002.

According to Fig. 4b, the considered region was zoned using the mean SPIs series and standard deviation (STD) values. It was found that the mean SPIs highest values for the SPI-3 and -6 occurred in the northern part and in the northwest, northern, and central parts, respectively. With increasing the SPI degree to 12 month, the amount of drought index increased in the whole country. However, it was observed that during the selected statistical period, in the central and southeast districts the arid regions increased for mid- and long-term droughts. Also, the northwest to northeast parts had the highest STD values for short term SPI, while for mid- and long-term SPIs the STD values were approximately the same (see Fig. 4b). Since the variability of the annual rainfall is high in the northern and northwest parts (districts with higher rainfall); therefore, droughts are more frequent in these parts. In addition, Alborz mountains which extends from northwest to northeast and Zagros mountains which extends from northwest to south affect the drought values. In the mentioned regions, high elevation is a significant driving element of precipitation and temperature, due to its negative impacts on the humidity flow from the country western and northern regions to the central parts. Also, from Fig. 4b, it could be seen that for the SPI-3 the number of drought events was highest in the northwest to northeast parts, where droughts were more frequent. With increasing the SPI degree the highest values were obtained for southeast part of the study area. In the next section, the SPIs series were analyzed to assess drought severity, peak, and duration features during the period of 1951-2016.

The mean values of drought features including severity, peak, and duration values were calculated and the results were illustrated in Fig. 4c. Based on Fig. 4c, it could be seen that the mean values of drought severity ranged from 1 for SPI-3 to 16 for SPI-12. For all selected regions, the severity values increased with increasing the SPI degree. For the shorter terms of SPI, the higher mean amounts of severity were observed in the northern
and northwest parts, whereas the higher values of this parameter for the SPI-12 were observed in the southern part. Also, the highest mean drought peak for the shorter terms of SPI occurred in the northern part, while for the SPI-12 the highest peak value occurred in southern and southwest parts. The mean values of drought duration varied between 1 to 22 months. It could be seen that with increasing the SPI degree, an increase in duration term values was observed. The lowest drought duration was obtained for northern part and
the highest drought duration was obtained for southeast part. Generally, and based on the results, it could be deduced that short-term and long-term droughts were more probable in the northern and southern parts of the selected region, respectively.

3.2 Decomposition of the SPIs and Computing the Coefficients’ Energy Values

For further investigating the SPIs series spatiotemporal variations in the selected region, the MODWT method was applied to decompose the SPIs series into several subseries. For this regard, according to Montaseri et al. (2018), daubechies (db2 and db4) mother wavelets were trained. Also, the optimum level of decomposition was determined via a trial and error process. The Root Mean Square Error (RMSE) criterion was used to select the optimum value. From the results, the db4 mother wavelet and decomposition level of 3 and 4 led to better outcomes.

The MODWT requires an infinite SPI_t signal, where t = …, -2, -1, 0, 1, …, N-2, N-1, N. However, in most cases the datasets are recorded over a specified period of time at discrete times. Therefore, in order to use this approach, the signal extension is required to specify unobserved amounts, SPI_0, SPI_1, …, SPI_{N+1}, SPI_{N+2}, prior to pre-processing. The right end of the SPIs time series (i.e. SPI_{N+1}, SPI_{N+2}, …) must be properly expanded. In this research, the data extension is performed using the following equation:

\[
L_j - 1 = (2^j - 1) \times 3
\]  

where j is decomposition level and L_j represents the number of removed data. Considering above equation and using the db4 wavelet, the first 45 data were removed. Figure 5a shows the results of the SPI-3 decomposition with and without data omitting for the Abadan station. Based on this figure, the lower wavelet levels expressed the higher frequencies component, while the higher levels showed the component with lower frequencies (including trend of the signal). In fact, the W subseries indicate specific time periods. For instance, the W1, W2, W3, and W4 show periods of 2, 4, 8, and 16 months for monthly data, respectively. Also, the V4 shows the trend of time series.

According to Fig. 5b, after signal decomposition, the subseries energy values were computed. From Fig. 5b, W1 and V3 in the SPI-3 and W2 and V4 in the SPI-6 and -12 showed the least variations. In order to investigate the relationship between energy values and the SPIs series, the energy values of the subseries were used to zone the selected region. Due to less variation of the W1 and V3 subseries in the SPI-3 and the W2 and V4 subseries in the SPI-6 and -12, these coefficients were used for this aim. Based on Fig. 5c, it could be seen that the SPI degree increasing led to reduction in the amount of energy. The highest energy values were obtained for SPI-3 and for the northern and northwest parts, where droughts were more frequent. The energy values of the central and southern parts were the lowest, where there were more arid regions. For shorter term droughts, the northern parts, which had the highest values of mean SPI, drought peak, and drought severity and the lowest value of drought duration, had the highest amount of energy. For long term drought (SPI-12), the southern parts, which had the highest values of mean SPI, drought peak, drought severity, and drought duration, had the lowest amount of energy. Therefore, it seems that in the shorter term droughts the energy had a direct relationship with the mean SPI, drought severity, and drought peak, while in the long term drought the energy had an inverse relationship with these quantities.
3.3 Results of the SPIs Regionalization Using the Decomposed Subseries Energy Values and K-means Method

After decomposition of the SPIs series to $W_i$ ($i=1, \ldots, 4$) and $V_4$ coefficients, since some coefficients have a weak correlation with the signal; therefore, it was tried to optimize the

![Fig. 5](attachment:image.png)

**Fig. 5** (a): The SPI-3 series decomposition using db(4, 3) for the Abadan station, (b): the energy values of decomposed subseries and their variation ranges, and (c): zoning map of the study area via the energy amounts of the decomposed subseries

3.3 Results of the SPIs Regionalization Using the Decomposed Subseries Energy Values and K-means Method

After decomposition of the SPIs series to $W_i$ ($i=1, \ldots, 4$) and $V_4$ coefficients, since some coefficients have a weak correlation with the signal; therefore, it was tried to optimize the
inputs of k-means. In this regard, different combinations of mentioned coefficients [i.e. V4, W1+V4, W2+V4, W3+V4, W4+V4, \( \sum_{i=1}^{4} W_i + V_4 \)] were considered. The SCI criterion was applied to validate the outcome of the clustering method. In the use of k-means clustering approach, at first the number of clusters should be selected. Therefore, the number of clusters was changed between 2 and 7 and clustering operations were done to determine the optimal number of clusters. Based on Fig. 6a, combining the V4 and W2 with considering

![Figure 6a](image)

**Fig. 6 (a):** The SCI validity index for the SPI-6 via MODWT-k-means clustering model, (b): comparison of the clustering results obtained for the proposed method and SPIs historical data, (c): the spatial clustering of the study area, and (d): the SCI validity index for MODWT-k-means clustering model
the cluster number of five led to better clustering result in the case of SPI-6. Between developed combinations, in the SPI-3 the V3+W1 combination with clustering number of six, in the SPI-6 the V4+W2 combination with clustering number of five, and in the SPI-12 the V4+W2 combination with clustering number of six were selected as appropriate coefficients to spatially identify the homogenous drought districts. For assessing the capability of the proposed methodology the 65 years of historical SPIs in the period of 1951-2016 was inserted to the k-means and the obtained results were compared with the proposed

Fig. 7 The subseries energy values of the clusters for the SPI-3 (a), SPI-6 (b), and SPI-12 (c)
hybrid methodology. According to Fig. 6b, the SCi and Mean Squared Error (MSE) criteria were used for this assessment. From this figure, it could be stated that clustering beads on the historical data was not led to desirable accuracy and the hybrid MODWT-k-means was more successful in determining the homogenous areas.

The spatial position of stations based on clustering by k-means method is illustrated in Fig. 6c, d. It could be clearly seen that the stations located in a specific cluster were spread in different areas and the proximity of the stations to each other was not the basis of clustering. For example, stations located near the Caspian Sea (stations 5, 6, 7, and 8), which had higher amounts of rainfall and located geographically side by side, were located in different clusters because the amounts and trends of energy variations were different. However, homogeneity in terms of hydrological characteristics at the SPI 12-month scale was slightly higher than the SPIs 3- and 6-month scales. Also, the SCi values were slightly higher for the SPI-12. In general, according to the results, it seems that the used method for zoning the study area had desirable performance. For example, the stations located in the central and southern parts (places with more arid regions) were in the same clusters.

In order to find more similarities among stations located in each cluster, the trend of energy variations in clusters were further assessed. Based on Fig. 7, the energy values showed wide variations, while in terms of within-cluster, they had the most similarities (homogeneity). It could be seen that as the degree of drought index increased, the similarity within the clusters also increased. It could be induced that the energy amounts of subseries can be used as an efficient method in zoning a region with different climates. For example, the trend of energy changes in cluster 2 of SPI-6 for all stations followed a unique trend while it was different from other clusters.

4 Conclusions

Drought as one of the severe and frequently occurring natural hazards has negative impacts on human life. In the current study, the spatial-temporal patterns of short- to long-term droughts were investigated for a region with different climates. The SPI-3, -6, and -12 series were selected for this aim and the statistics in drought features including severity, peak, and duration were assessed. For the short-term SPI, the areas which extended from the northwest to the northeast showed the highest STD, while for the mid- and long-term SPIs, the values of STD were approximately the same for all areas. The obtained results showed that the amount of drought severity and duration increased when the degree of SPI increased. The highest values for peak and severity features and the lowest values for duration feature were observed in the northern parts. Then, the energy values of the scaling and wavelet coefficients extracted from the MODWT was computed. The coefficients with less variation were selected to be used for zoning the area and finding a relationship between the SPIs series and energy amounts. From the results, the highest energy amounts were observed in the northern and northwest parts in which droughts were more frequent there, while the lowest amounts were observed in the central and southern parts (i.e. more arid regions). Generally, in the shorter term droughts, the energy had a direct relationship with the mean SPI and peak and severity of drought, and in the long term drought the energy had an inverse relationship with these properties. The MODWT-based input data were used for area clustering via the K-means to spatially determine the homogenous drought areas. The results showed that the stations located in a specific cluster were spread in different areas and the proximity of the stations to each other was not the basis of clustering. The
Homogeneity of stations was slightly higher for the SPI-12 compared to the SPI-3 and -6. The energy values of stations located at different clusters had wide variations, while in the within-cluster mode the similarities between stations were more. According to the results, increasing the degree of drought index led to more similarity between the members of a specific cluster. In general, the proposed model yielded appropriate results and it could be a useful method for drought investigating in areas with different climates due to having desirable values for the SCi index and having lower input data.

Authors Contributions Kiyoumars Roushangar: Conceptualization, Supervision, Methodology, Review & Editing. Roghayeh Ghasempour: Project administration, Investigation, Data Curation, Methodology, Writing. Vahid Nourani: Conceptualization, Methodology, Formal analysis, Review & Editing.

Funding The authors did not receive support from any organization for the submitted work.

Availability of Data and Material Monthly precipitation data of 31 sites of Iran are obtained from Iranian Meteorological Organization.

Declarations

Ethics Approval Not applicable.

Consent to Participate Not applicable.

Consent for Publication Not applicable.

Conflicts of Interest / Competing Interests The authors declare no conflicts of interest / competing interests.

References

Agarwal A, Maheswaran R, Sehgal V, Khos R, Sivakumar B, Bernhofer C (2016) Hydrologic regionalization using wavelet-based multiscale entropy method. J Hydrol 538:22–32
Barton Y, Giannakaki P, Von Waldow H, Chevalier C, Pfahl S, Martius O (2016) Clustering of regional-scale extreme precipitation events in southern Switzerland. Mon Weather Rev 144(1):347–69
Chanda K, Maity R, Sharma A, Mehrotra R (2014) Spatiotemporal variation of long-term drought propensity through reliability-resilience-vulnerability based drought management index. Water Resour Res 50(10):7662–76
Cornish CR, Bretherton CS, Percival DB (2006) Maximal overlap wavelet statistical analysis with application to atmospheric turbulence. Bound Layer Meteorol 119(2):339–74
Dadson SJ, Lopez HP, Peng J, Vora S (2019) Hydroclimatic extremes and climate change. Water Science, Policy, and Management: A Global Challenge, Wiley 11–28
Dghais AAA, Ismail MT (2013) A comparative study between discrete wavelet transform and maximal overlap discrete wavelet transform for testing stationarity. Int J Math Comput Sci Eng 12:1677–1681
Dinpashoh Y (2006) Study of reference crop evapotranspiration in IR of Iran. Agric Water Manag 84(1–2):123–129
Hsu KC, Li ST (2010) Clustering spatial–temporal precipitation data using wavelet transform and self-organizing map neural network. Adv Water Resour 33(2):190–200
Jain AK, Murty MN, Flynn PJ (1999) Data clustering: a review. ACM Comput Surv 31(3):264–323
Kwak J, Kim S, Jung J, Singh VP, Lee DR, Kim HS (2016) Assessment of meteorological drought in Korea under climate change. Adv Meteorol. https://doi.org/10.1155/2016/1879024
Li X, He B, Quan X, Liao Z, Bai X (2015) Use of the standardized precipitation evapotranspiration index (SPEI) to characterize the drying trend in southwest China from 1982–2012. Remote Sens 7(8):10917–10937
Likas A, Vlassis N, Verbeek JJ (2003) The global k-means clustering algorithm. Pattern Recognit 36(2):451–461
Madadgar S, Moradkhani H (2013) A Bayesian framework for probabilistic seasonal drought forecasting. Hydrometeorol 14(6):1685–1705
Mann ME, Gleick PH (2015) Climate change and California drought in the 21st century. Proc Natl Acad Sci U S A 112:3858–3859
Marvel K, Cook BJ, Bonfils CJ, Durack PJ, Smerdon JE, Williams AP (2019) Twentieth-century hydro-climate changes consistent with human influence. Nature 569:59–65
McKee TB, Doesken NJ, Kleist J (1993) The relationship of drought frequency and duration to time scales. In Proceedings of the 8th Conference on Applied Climatology 179–184
Mirabbasi R, Anagnostou EN, Fakheri-Fard A, Dinpanshoh Y, Eslamian S (2013) Analysis of meteorological drought in northwest Iran using the Joint Deficit Index. J Hydrol 7(492):35–48
Montaseri M, Ghavide SZZ, Sanikhani H (2018) Water quality variations in different climates of Iran: toward modeling total dissolved solid using soft computing techniques. Stoch Environ Res Risk Assess 32(8):2253–2273
Mukherjee S, Mishra A, Trenberth KE (2018) Climate change and drought: a perspective on drought indices.Curr Clim Chang Rep 4:145–163
Nanzad L, Zhang J, Tuvedendorj B, Nabi M, Zhang S, Bai Y (2019) NDVI anomaly for drought monitoring and its correlation with climate factors over Mongolia from 2000 to 2016. Arid Environ 164:69–77
Percival DB, Walden AT (2000) Wavelet Methods for Time Series Analysis, Cambridge Series in Statistical and Probabilistic Mathematics, 1st edn. Cambridge University Press, Cambridge, UK
Pham DT, Dimov SS, Nguyen CD (2005) Selection of K in K-means clustering. Proc Inst Mech Eng C J Mech Eng Sci 219(1):103–119
Roushangar K, Alizadeh F (2018) Entropy-based analysis and regionalization of annual precipitation variation in Iran during 1960–2010 using ensemble empirical mode decomposition. J Hydroinform 20(2):468–485
Roushangar K, Nourani V, Alizadeh F (2019) Corrigendum to a multiscale time-space approach to analyze and categorize the precipitation fluctuation based on the wavelet transform and information theory concept. Hydrolog Res 49(3):724–743
Sakizadeh M, Mohamed MM, Klammler H (2019) Trend analysis and spatial prediction of groundwater levels using time series forecasting and a novel spatio-temporal method. Water Resour Manag 33(4):1425–37
Sujjaviriyasup T (2017) A new class of MODWT-SVM-DE hybrid model emphasizing on simplification structure in data pre-processing: A case study of annual electricity consumptions. Appl Soft Comput 54:150–163
Surendran U, Anagha B, Raja P, Kumar V, Rajan K, Jayakumar M (2019) Analysis of drought from humid, semi-arid and arid regions of India using DrinC model with different drought indices. Water Resour Manag 33(4):1521–1540
Tigkas D, Vangelis H, Tsakiris G (2019) Drought characterisation based on an agriculture-oriented standardised precipitation index. Theor Appl Climatol 135(3):1435–1447
Tigkas D, Vangelis H, Tsakiris G (2020) Implementing crop evapotranspiration in RDI for farm-level drought evaluation and adaptation under climate change conditions. Water Resour Manag 34(14):4329–4343
Van Loon AF (2015) Hydrological drought explained. Wiley Interdisciplinary Reviews. Water 2:359–392
Vicente-Serrano SM (2007) Evaluating the impact of drought using remote sensing in a Mediterranean, semi-arid region. Nat Hazards 40(1):173–208
Yang L, Deng M (2010) Based on k-means and fuzzy k-means algorithm classification of Precipitation. In 2010 International Symposium on Computational Intelligence and Design 1:218–221. IEEE
Zhou X, Wang P, Tansey K, Zhang S, Li H, Wang L (2020) Developing a fused vegetation temperature condition index for drought monitoring at field scales using Sentinel-2 and MODIS imagery. Comput Electron Agric 168:10727–10741

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