Original Paper

Spatial statistics techniques for SPEI and NDVI drought indices: a case study of Khuzestan Province

M. Nejadrekabi1 · S. Eslamian1,2 · M. J. Zareian3

Received: 11 April 2021 / Revised: 8 October 2021 / Accepted: 13 December 2021 / Published online: 29 January 2022
© Islamic Azad University (IAU) 2022

Abstract
Drought is a major water resources management issue in Iran. Khuzestan Province is in a drought state due to water shortage. Therefore, identifying areas at high risk of drought and when drought occurs is essential for drought management. For this purpose, this study used precipitation and temperature data of 12 selected stations and MODIS sensor images from the United States Geological Survey database in 2000–2017. The Standardized Precipitation Evapotranspiration Index (SPEI) and the Standardized Normalized Difference Vegetation Index (NDVI) were calculated using the Hargreaves–Samani method and ENVI software. Moreover, different spatial statistics techniques were used in the ArcGIS environment to analyze the results. Also, time series diagrams were drawn, and the trend was evaluated using the Mann–Kendall test. Finally, the distribution of NDVI values was investigated using EasyFit software, and the amount of drought damage was determined using NDVI. The investigation of the cluster maps of the Anselin Local Moran’s Index along with hot and cold spots formed for both SPEI and NDVI showed that drought severity was higher at the southern stations than at the semi-northern and northwestern ones in the province. Moreover, the survey results using the EasyFit software showed that the southern stations, including the Ahvaz, Mahshahr, and Omidiyeh-Aghajari stations, were more at risk of drought than the other stations due to the drought threshold. Furthermore, the total damage caused by drought for the Ahvaz and Abadan stations showed a damage rate of 50%.

Keywords Drought · Standardized precipitation evapotranspiration index (SPEI) · Normalized difference vegetation index (NDVI) · Spatial autocorrelation · Hot spot analysis

Introduction
Drought is a natural event that occurs frequently or intermittently in almost any type of climate. It is generally the result of a natural decrease in precipitation over a period of time (Badaq Jamali et al. 2005). In Iran, Khuzestan Province is in a drought state due to reduced rainfall. In recent years, drought and water salinization have destroyed groves in the province. Due to the permanent presence of drought-induced dust, Khuzestan Province has many respiratory and lung patients. To manage dust, it is necessary to identify areas with a high risk of drought and determine the scope of its impact to reduce dust management costs. At present, Coronavirus Disease 2019 has increased urban water consumption in Khuzestan Province by 15–20%, and thus water shortage problems have increased in some areas in the province. In addition, essential products, such as rough rice, are produced in the province’s extensive agricultural lands. In recent years, water withdrawal from the network for rough rice cultivation has caused problems in supplying drinking water to people in the cities and villages of Abadan, Khorramshahr, and Shadegan in the province. Therefore, it is necessary to identify drought-prone areas to make more rational decisions regarding crop selection.

According to studies, the weight of rainfall in drought indices is more than other variables. Therefore, this variable alone describes more than 80% of changes in drought indices. This is why indices often used in many countries take this variable into account (Keyantash and Dracup 2002).
Serrano et al. (2011) investigated the effect of global warming on droughts and water sources in Spain using the two drought indices, the Standardized Precipitation Index (SPI) and the Standardized Precipitation Evapotranspiration Index (SPEI). They reported similar results for the two indices (Vicente-Serrano et al. 2011). SPEI results are more accurate than SPI results when evaporation and rainfall play a significant role in drought events (Wang et al. 2015). Serrano et al. (2010) argued that the Palmer index, although having significant benefits, is weak because it cannot provide short-term and long-term effects of drought by determining drought conditions at different time scales. They also discussed the relatively new Respiratory Disturbance Index (RDI). They believed that the precipitation to evapotranspiration ratio was not defined in RDI when the evapotranspiration rate was zero (in the winter months in different parts of the world) and the range of ratios was applied. The reason is that the ratio is significantly small and cannot reflect the role of temperature (Vicente-Serrano et al. 2010). SPEI has a higher utilization rate than other indices, and its main advantage is the ability to detect the effect of changes in evapotranspiration and temperature concerning global warming. However, the quality of a drought index can only be as good as the input data (Vicente-Serrano et al. 2010; Wang et al. 2015). Moreover, having the ability to incorporate both temperature and precipitation, SPEI may be a useful drought index (Gurrapu et al. 2014) (Vicente-Serrano et al. 2015). Over the past decades, researchers have developed various indices, including SPI, the Palmer Drought Severity Index, RDI, and the Standard Index of Annual Precipitation, to monitor the drought situation and evaluate its quantitative effects. In this study, two drought indices, namely SPEI and the Standardized Normalized Difference Vegetation Index (NDVI), were used to monitor the moisture period.

The effects of increased temperature are far greater than the effects of reduced precipitation on the increased drying rate in an area. Accordingly, SPEI can be used as an appropriate index for determining drought severity. Three variables of precipitation, temperature, and evapotranspiration potential are considered in this index. One of the limitations of SPI is the inability to consider water balance based on evapotranspiration. In contrast, SPEI considers water balance in calculating the drought index based on precipitation and potential evapotranspiration. In comparison to SPI, SPEI can also show the main effects of temperature rise on water demand (Saeidipour et al. 2019).

Xiaoqiang et al. (2017) investigated the effect of drought in 2009–2010 on vegetation using NDVI, the Enhanced Vegetation Index (EVI), the Normalized Difference Water Index (NDWI), and Land Surface Temperature (LST). They reported that NDVI, EVI, and NDWI levels decreased while LST increased in southwest China. Tien Le Thuy Du et al. (2018) examined the six most widely used vegetation indices (VIs) to evaluate droughts based on their sensitivity to rainfall, soil moisture, and their relationship with crop production and forest fires. The six VIs representing four main groups, including VI₅ based on greenery (the Vegetation Condition Index), VI₆ based on water (the Normalized Difference Water Index and the Land Surface Water Index), VI₇ based on temperature (the Temperature Condition Index), and hybrid VI₈ (the Vegetation Health Index and the Normalized Difference Drought Index), were tested using MODIS data from January 2001 to December 2016 through the Google Earth Engine platform. The results showed that droughts occurred almost every year but with varying severity. Afshar et al. (2021) calculated meteorological droughts using SPI in more than 24,000 pixels with a spatial resolution of 0.25 × 0.25 located in Central Europe. They compared the heterogeneous mean values of VIs in dry and wet periods resulting from meteorological droughts statistically to evaluate the ability of VIs in drought conditions. Jie Zhang et al. (2016) used the Drought Severity Index (DSI) to monitor agricultural drought and analyze its impact in northern China using the MODIS DSI. Orimoloye et al. (2019) evaluated drought severity in the Cape Town region of South Africa using the Geographic Information System (GIS) and remote sensing data obtained from the United States Geological Survey (USGS) database between 2014–2018. They examined vegetation health and drought severity in the study area using VIs and the Normal Difference Drought Index. Fei Xie, Hui Fan (2021) examined the possibility of deriving drought indices from the MODIS VIs (EVI/NDVI) and the LST indices. They investigated three widely used reconstruction methods, including the Savitzky–Golay filter, the Harmonic Analysis of Time Series Algorithm, and Whittaker smoother, in the Lancang-Mekong River Basin using MOD13Q1 and MOD11A2 products of the MODIS sensor and the Google Earth Engine cloud-based platform for the years 2001–2018. Sandeep et al. (2020) used the Normalized Vegetation Supply Water Index to monitor agricultural drought in the semi-arid ecosystem of the Indian subcontinent through indices derived from the CHIRPS and MODIS series data sets between 2000–2018. Abebe Senamaw et al. (2021) used the Vegetation Condition Index eMODIS NDVI and SPI between 2000–2016 to show the spatial and temporal variations of agricultural drought and meteorology in Ethiopia. Qian Wang et al. (2019) assessed the spatial aggregation of the agricultural drought crisis in China between 1978–2016 using the Global Moran’s Index (GMI) and the Local Indicators of Spatial Association. They reported that the direction of the drought crisis was toward the northwest. They cited regional differences in rainfall and temperature, rational irrigation, and environmental changes as factors changing the focus of drought. Gevorg Tepanosyana et al. (2019) used the Local Moran’s Index to identify spatial clusters and hot pot points of lead, Mo, and Ti in urban soils of
Yerevan. In the case of lead, they concluded that high-value spatial clusters were constructed in urban areas and were spatially correlated with the prevailing northeast winds. Tongtiegang Zhao et al. (2019) presented significant spatial patterns of GCM seasonal global precipitation predictions using GMI based on the spatial map of the anomaly correlation between the mean of the forecast and the observation groups. Luković Jelena et al. (2015) examined the spatial pattern of the effect of North Atlantic fluctuations on rainfall in Serbia. Moran’s I statistic calculation showed a statistically significant correlation (p < 0.01) for all months, seasons, and years to confirm the spatial clustering of the effect of NAO fluctuations on the amount of rainfall during a year. In addition, both positive and negative Gi Z scores showed clustered simultaneous flood hazards over the Po River Basin in Italy. For this purpose, they proposed a pair of agglomerative hierarchical clustering algorithms, a vital tool in the earth sciences to analyze dependency information. Furthermore, by considering the interaction between the influential variables, they identified spatial sub-regions with similar behaviors during floods. Melita Perćec Tadić et al. (2015) used the Krigeing geostatistical method, which is part of the study of spatial statistics, to present a map of maximum snow load values with a 50-year return period in Croatia. Chen Shaoqian et al. (2013) in their study used the Moran’s Index to study the spatial distribution characteristics of traffic accidents. They reported that the temporal and spatial distribution of traffic phenomena had a clustering behavior. They also observed that the temporal and spatial changes of traffic phenomena were continuous. Agossou Gadedjisso-Tossou et al. (2021) analyzed rainfall and temperature trends and their significance on dry cereal yield in northern Togo using the Mann–Kendall test Fooladia et al. (2021) investigated the Shahpour River in southern Iran to identify human effects and climate diversity at the river basin scale. They used Mann–Kendall tests, kpss statistics, and homogeneity test at the 5% confidence level. Wenyun Huang et al. (2021) used the Mann–Kendall test to determine the spatio-temporal variation of drought in northwestern China during 1950–2012 to determine the increasing or decreasing drought trend. In the present study, SPEI, NDVI, the Mann–Kendall test, and spatial statistics were used for drought analysis, and areas with high drought risk were identified.

**Materials and methods**

**Study area**

Khuzestan Province, with an area of about 64,236 square kilometers, is located between 47° 41′ to 51° 39′ east longitude and 29° 58′ to 33° 4′ north latitude in southwestern Iran. The cities of Izeh and Hindijan in the province have the highest and lowest altitudes with 764 m and 2 m above sea level, respectively. The summer season in this province starts in May and continues until October. The mean temperature in this period is 31.2 °C, and the average temperature in winter is 14.9 °C. The mean rainfall in Khuzestan is 265.5 mm per year. Climatic conditions vary widely in the province. Accordingly, the western, southwestern, and southern regions are affected by the hot, dry climate, and the central, southeastern, and northwestern regions are affected by the hot desert climate; these two climates cover most of the province. Moreover, hot semi-arid and semi-arid climates affect the eastern and northern regions along with large parts of the northeastern region in the province. Mediterranean and semi-humid climates also affect the eastern and northeastern highlands in the province.

Some of the major consequences of drought and water shortage in the province include a decrease in the development of agricultural projects, the destruction of palmetum, and, most significantly, the dust crisis. The occurrence of drought and the decreased vegetation density in recent years have caused the haze phenomenon and other factors, which have increased many respiratory diseases in the province. Respiratory diseases in the area have caused many patients to die from the coronavirus in recent months.

In this study, statistical data of 12 synoptic stations of the Meteorological Organization and the Water Resources Management Company of the Ministry of Energy was used based on the homogeneity test and the common statistical period (Fig. 1). There are 10 stations in the study area. For better overlap of the stations, the Eydanak and Paypol stations were considered outside the study area. The datasets of the stations include total precipitation during rainy months between 2000–2017. The UTM geographical coordinates of the stations are given in Table 1.

**Method**

In this study, the rainfall and temperature data of the meteorological stations at the Khuzestan Meteorological Organization and the Khuzestan Water and Electricity Organization and MODIS sensor images were used between 2010–2017. After sorting the precipitation and temperature data using SPSS software, outlier data were identified and discarded. Then, the normality and homogeneity of the data were examined. SPEI was calculated by the Hargreaves–Samani method in Excel software monthly and annually, and drought severity was determined. SPEI drought values were entered into ArcGIS 10.2. After interpolation using the IDW
Geostatistical method, hot and cold spots and drought clusters were determined with the help of spatial statistics extension. Then, the drought behavior was identified in the study area. Monthly and annual time series diagrams were drawn, and the drought trend of the stations was determined using the Mann–Kendall method. ENVI 5.3 software was used to determine values of the standardized normalized difference vegetation index (NDVI). NDVI values were entered into ArcGIS 10.2, and hot and cold spots of NDVI were determined with the help of spatial statistics. Finally, the drought behavior of the study area was studied by the vegetation change index.

The best statistical distribution matching with the NDVI values of the stations was identified using EasyFit software. The average monthly NDVI values of the stations were obtained, and by identifying poor coverage areas for each station, the percentage of data with open coating, i.e., being in a drought condition, was determined. For validating the work carried out in EasyFit, artificial data was generated, and the open coating values were evaluated and compared with the original data. Figure 2 presents a schematic diagram of the steps performed in the study.

**The standardized precipitation evapotranspiration index (SPEI) calculation**

Since data were limited on the study area, the Hargreaves-Samani method was used to calculate SPEI. This method
Investigating the total drought damage

Assessing the risk of drought onset

Investigating the probability of drought onset

Investigating the statistical distribution of the stations

Determining hot and cold spots using GETIS-ORD G*

Determining the clustering or dispersion of SPEI values using Moran's correlation coefficient

Determining clusters and outliers using the Anselin Local Moran's Index

Using ENVI 5.3 software

Using Excel software

Using SPSS software

Using SPSS software

Using Excel and Hargreaves-Samani equation

Using GIS 10.2 IDW interpolation

Using GIS spatial statistics extension

Fig. 2 Schematic of the steps
requires calculating the average air temperature (T) and the difference between the average maximum and minimum temperature (TR) in the desired period.

Monthly and annual SPEI was calculated by formulation in the Excel environment using the Hargreaves-Samani method, and monthly and annual drought severity was determined (Alizadeh 2001).

\[
ET_0 = 0.0023Ra(T + 17.8)\sqrt{TR} \tag{1}
\]

\[
TR = T_{\max} - T_{\min} \tag{2}
\]

In this formula, ET0 is the amount of evapotranspiration with the same unit of Ra, T is the average air temperature in degrees Celsius, TR is the mean difference between the maximum and minimum temperatures in degrees Celsius, and Ra is the extraterrestrial radiation in millimeters of water evaporating per day for different latitudes.

The defined ranges to determine SPEI severity did not contribute to detecting drought severity in the region. For this reason, to define drought severity ranges for each station, the average value and standard deviation were determined. For SPEI classification, the formulation was performed in the Excel environment, and the drought index was classified into four groups (Mohammadyari et al. 2015):

1) Values smaller than average—standard deviation.
2) (Mean—Standard Deviation) to mean.
3) (Mean + Standard Deviation) to mean.
4) Values greater than (mean + standard deviation).

Spatial autocorrelation

Spatial autocorrelation is related to the relationship between residual values along the regression line. Strong autocorrelation occurs when residual values are strongly correlated; that is, changes in them occur systematically.

Moran’s spatial autocorrelation analysis can describe the spatial properties of a variable over an entire region. It can be used to identify the mean spatial difference between all spatial features (pixels) and their adjacent features (pixels). (Poudyal 2013; Peters and Matsui 2015).

In general, if the value of the Moran’s Index is close to +1, the data have a spatial autocorrelation and a cluster pattern. However, if the value of the Moran’s Index is close to -1, the data are discrete and scattered. In the case of this tool, the null hypothesis (H0) is that there is no spatial clustering between drought index values. H0 can be rejected when the P-value is too small, and the calculated Z-value (its absolute value) is too large (outside the confidence range).

Evaluation of clusters and outliers with the Anselin local Moran’s index

Cluster and outlier analysis, also known as the Anselin Local Moran’s Index, shows the statistical distribution of phenomena in space. Based on weighted geographic features, this tool shows where large and small amounts of these phenomena are distributed as a cluster in space and which features have very different values from the surrounding features. If the value of I is positive, it means that similar features surround the desired feature (here drought). Therefore, the desired features are part of that cluster. If the value of I is negative, it means that the desired feature is surrounded by features that are not similar to it at all. This type of feature is called an outlier. The Anselin Local Moran’s Index can only be interpreted in terms of calculated standard score and P-Value. The output layer created by this tool includes high-value clusters (HH), low-value clusters (LL), the outlier in which a high value is surrounded by low values (LH), and outliers in which a low-value feature is surrounded by high values (LH). (Anselin 2000, 2013).

The General Statistics Index (G)

The Moran’s Index is designed to describe the global spatial correlation. However, it is not efficient to identify different classification types of spatial patterns. These patterns are sometimes referred to as hot and cold spots. If high values are close to each other, the Moran’s Index and Geary’s coefficient indicate relatively high positive spatial autocorrelation. This cluster of high values may be referred to as a hot spot. However, the high positive spatial autocorrelation shown by the Moran’s Index and Geary’s coefficient may be due to low values adjacent to each other.

This type of cluster is called a cold spot. The Moran’s Index cannot distinguish between these two types of spatial autocorrelation. Therefore, the General Statistics Index, G, is preferable to the Moran’s Index in determining positive (hot) and negative (cold) spots in the study area. These hot and cold spots can be considered as spatial concentrations (Balyani and H. 2014).

Getis-Ord-Gi statistics are used when there is evidence of a clustering pattern at data collection points. However, the aim is to know whether high values lead to a clustering pattern or low cluster values (Boots and Kanaroglou 1988).

The hypothesis of a high or low clustering method is defined as follows:

Null hypothesis (H0): There is no spatial clustering in the desired variable values for the features in the study area.
The alternative hypothesis (H1): There is a high (positive) or low (negative) spatial clustering pattern in the desired variable values for the features in the study area.

The purpose of using the General Statistics Index, \( G \), is to detect the presence or absence of high or low clustering in spatial data (Burnham and Anderson 2002).

This index can be calculated using the following equation:

\[
G(d) = \frac{\sum_{i=1}^{m} \sum_{j=1}^{m} \omega_{ij}(d)y_i y_j}{\sum_{i=1}^{m} \sum_{j=1}^{m} y_i y_j}, \quad i \neq j
\]

where \( m \) is the number of zones and \( \omega_{ij}(d) \) is the weight based on the spatial vicinity between the regions \( i \) and \( j \). This index measures spatial dependence using all pairs of positional values \((y_i, y_j)\) so that the positions \( i \) and \( j \) are at a distance \( d \) from each other (Ord and Getis 2010).

The standard score of the General Statistics Index, \( G \), is measured as follows:

\[
Z_G = \frac{G - E[G]}{\sqrt{V(G)}}
\]

where

\[
E[G] = \frac{\sum_{i=1}^{n} \sum_{j=1}^{n} W_{ij} x_i}{n(n-1)}, \quad \forall j \neq i
\]

\[
V(G) = E[G^2] - E[G]^2
\]

In a spatial pattern where the geographical location of high values is close to other high values, the \( Z_G \) statistic is positive. If clustering is based on low values, the \( Z_G \) statistic is negative. (Getis et al. 1995, 1992).

**Gi * index or hot spots**

Another spatial autocorrelation index is the Gi * index. This index shows the distribution points of location-based data according to the type of data and spatial correlation analysis as a standard score number. This index, also known as GetisOrd \( G^* \) or hot spots, can be evaluated as the following equation:

\[
G^*_i = \frac{\sum_{j=1}^{n} W_{ij} X_j - \bar{X} \sum_{j=1}^{n} W_{ij}}{\sqrt{\frac{\sum_{j=1}^{n} X_j^2 - (\sum_{j=1}^{n} W_{ij} \bar{X})^2}{n-1}}}
\]

where \( X_j \) is the attribute value for the feature \( j \), \( W(i, j) \) is the spatial weight between the features \( I \) and \( j \), and \( n \) is the total number of features.

Since \( Gi^* \) is a kind of Z-score, no further calculation is needed.

**The standardized precipitation evapotranspiration index (SPEI) analysis with Getis Ord G * statistics**

The regulated data in the Excel environment was transferred to ArcGIS software, and monthly drought zoning maps were determined using the IDW interpolation method. Then, spatial statistics of the ArcGIS environment and the Moran’s Index were used to determine the formation or non-formation of clusters.

High/low clustering statistics were used to determine the SPEI values of the formed clusters.

**The standardized normalized difference vegetation index (NDVI)**

Ground weather stations provide a relatively good source of information for drought monitoring. However, they are often scattered and insufficient due to the limited distribution and coverage of gauging stations. Particularly in developing countries where the stations are scarcely available, drought monitoring is more arduous. A growing number of Earth observation satellites provide useful data sources to monitor the changing dynamics of soil, water, and vegetation in the terrestrial surface (Tatem et al. 2008). Remote sensing-based VIs, which combine different spectral bands, have been widely used to monitor vegetation responses to weather conditions, and thus, assess drought conditions (AghaKouchak et al. 2015).

Based on the positive significant correlation between NDVI and rainfall anomalies, several studies concluded that NDVI as the most common form of VIs could be used effectively in the drought early warning system (Martiny et al. 2006; Murthy et al. 2009; Quiring and Ganesh 2010).

Rouse (1974) presented the following formula for calculating this index:

\[
NDVI = \frac{NIR - VIS}{NIR + VIS}
\]
In this regard, NIR is the near-infrared band, and VIS is the visible red band.

The index has normal values, i.e., it has an absolute value equal to or less than one, and its range is between (-1 to +1). Therefore, it can be easily examined in charts and histograms. Further, it is easily possible to interpret images representing this index and stretch its values visually.

Since plant indices are designed based on the spectral properties of plants, values obtained from them reflect the density of vegetation in the range of the desired pixel. Regarding NDVI, values are between 0.1 and 0.05 for poor coverage areas, 0.6 and 0.1 for normal coverage areas, and 0.7 and 0.6 for density coverage areas. Water, snow, and ice have negative NDVI values, soils have NDVI values less than 0.05, and clouds generally have zero NDVI values (Rouse et al. 1974).

The standardized normalized difference vegetation index (NDVI) calculation and analysis with spatial statistics

Based on 16-day NDVI MODIS images with a resolution of 250 m, monthly NDVI stations for the study area were downloaded directly from the USGS database. Monthly NDVI values were determined for each station using ENVI 5.3 software. The obtained values were arranged and categorized in Excel. Similar to SPEI, areas determining the NDVI drought severity did not contribute to the detection of drought severity in the region. For this reason, in this section, a new drought severity classification was determined for each station based on the average NDVI value and NDVI standard deviation.

Then, the NDVI values were entered into ArcGIS 10.2, interpolation was performed using the IDW method, and an NDVI zoning map was prepared. Next, the spatial statistics section of ArcGIS 10.2 was used, and the hot and cold spots of NDVI were determined on the map using the Getis Ord G* statistic.

Examination of time series charts and trend changes for both the Standardized Normalized Difference Vegetation Index (NDVI) and the Standardized Precipitation Evapotranspiration Index (SPEI).

The Mann–Kendall test

The existence of trends in drought severity can be indicative of how it changes. There are different methods for examining changes and trends in the time series of drought indices that are divided into two categories of parametric and non-parametric statistics. The linear regression method belongs to the parametric category, and the most common non-parametric method is the I-Kendall method. The Mann–Kendall method was first proposed by Mann (1945) and then developed by Kendall (1970). The null hypothesis of the Mann–Kendall test indicates that the trend in the data series is random, and the acceptance of the hypothesis one (rejection of the null hypothesis) indicates the existence of a trend in the data series (Ghasaabe Feyz and Eslami 2017).

The Mann–Kendall test is recommended by the World Meteorological Organization (168, 2009.MO, NO) to find the trend. In the Mann–Kendall test for the time series X (x1, x2, x3,…, xn) with the length n, based on the comparison of data in a pair, the statistic S is constructed in Eq. (9), where sgn(xk-xi) has values of 1, 0, and -1 when xk > xi and xk < xi, respectively (Kumar et al., 2009).

The number of sign statements is equal to \( \frac{n(n-1)}{2} \). Based on H0 stating that the data are independent and evenly distributed, the mean statistic S is expressed as E (S) = 0, and the modified variance of V (S) according to Hamed (2008) is expressed in Eq. (10).

\[
S = \sum_{i=1}^{n-1} \sum_{k=i+1}^{n} \text{sgn}(x_k - x_i) \quad (9)
\]

\[
V(S) = \frac{n(n-1)(2n+5)}{18} - \sum_{j=1}^{m} \frac{t_j(t_j - 1)(2t_j + 5)}{18} \quad (10)
\]

where

n: the length (or number of observations in) of time series x.

\( t_j \): the frequency of data with the same value.

m: the number of series with at least one duplicate data.

Mann (1945) and Kendall (1975) showed that with increasing the number of observations (i.e., larger n), the statistical distribution of s tends to be normalized. By assuming n > 10, the significance of the trend can be tested by the standard variable Zmk in Eq. (11) compared to the standard normal variable Zc at α significant level. From the statistical science perspective, if α is selected equal to 0.05 by the standard two-way variable test method, \( Z(1_{-\frac{1}{2}}) = Z_{0.975} \) will be equal to 1.96 in the statistical table (Behroozi et al. 2017) (Alijani et al. 2012).

\[
Z_{MK} = \begin{cases} 
\frac{S-1}{\sqrt{V(S)}} & S > 0 \\
0 & S = 0 \\
\frac{(S-1)}{\sqrt{V(S)}} & S < 0 
\end{cases} \quad (11)
\]
Fig. 3 Spatial correlation between the Standardized Precipitation Evapotranspiration Index (SPEI) and the Moran’s Index in the studied months.
The average annual SPEI was used to draw the SPEI time series diagram, and the Mann–Kendall test was used to determine the SPEI trend.

Results and discussion

Results of the Moran’s index

According to Fig. 3, the Moran’s Index values are close to one or more in all the months except October. A very small P-value and a large Z-value indicate the rejection of H0 and the cluster pattern of SPEI values in these months. This clustering pattern of SPEI values shows that the values have a spatial correlation (Askari 2011).

Results of the cluster and outlier study with the anselin local Moran’s index

According to Fig. 4 and Table 2, the study analyzed by the Anselin Local Moran’s Index showed that no clusters were formed in any of the stations. However, the formed clusters for the eastern stations of Dez Dam and Izeh Dam and the Abadan Meteorological station located in the southwest of the region should be considered. For the Izeh Dam and Shahid Abbaspour Dam stations, LL clusters formed in October, April, and May indicated lower SPEI and drought severity at a 95% confidence level. The Abadan meteorological station, the formed LL clusters showed a lower SPEI value and drought severity from November to March at a 95% confidence level.

Results of the standardized precipitation evapotranspiration index (SPEI) analysis using Getis Ord G * statistics

The study of the type of clusters formed and drought hot and cold spots determined using Getis Ord G * statistics are shown in Figs. 4 and 5 and Table 3.

Getis Ord G * statistics showed that the Abadan, Ahvaz, and Mahshahr stations were identified as drought hot and cold spots among the southern stations. Drought hot spots were created for these stations in October, November, April, and May. Among the northern stations, drought cold spots were created for the Dez Dam station in October, April, and May. Drought hot and cold spots were better identified at the eastern stations than at the other stations. The Dez Dam and Izeh Dam stations had similar spatial behavior in terms of drought. Accordingly, drought cold spots were identified in October, November, April, and May, while drought hot spots were identified in December, January, February, and March at these stations. No significant spots were formed for the rest of the stations.

Results of drought hot and cold spots based on the standardized normalized difference vegetation index (NDVI)

According to the results presented in Table 4, during the study period, among the southern stations, hot spots with high vegetation density were identified at the Shohada Dam station, and cold spots with open covers were identified at the Aghajari and Ahvaz stations.

Among the northern stations, hot spots with high vegetation density were observed at the Dezful and Gotvand Regulatory Dam stations, in the months during which spatial behavior was identified, and cold spots with low vegetation density were identified at the Dez Dam station.

For the eastern stations of Izeh Dam and Shahid Abbaspour Dam, cold spots were identified in October, November, and December, and hot spots with dense vegetation were identified in February, March, and April.

Results of the Mann–Kendall test

The time series diagram of the stations is shown in Fig. 6. The results of the Mann–Kendall test at a 95% confidence level are according to Table 5. Time series charts for the Dez Dam, Izeh Dam, Shahid Abbaspour Dam, Gotvand Dam, and Dezful Regulatory Dam stations showed a decreasing SPEI trend toward negative numbers. The SPEI levels suddenly decreased at all the mentioned stations from 2012, indicating an increase in drought severity since 2012.

As can be observed in the table, the SPEI time series diagram of the Dez Dam, Shahid Abbaspour Dam, Izeh Dam, Gotvand Dam, and Dezful Regulatory Dam stations had a trend.

The mean annual NDVI values and the Mann–Kendall test were used to determine the NDVI trend. The time series diagram is shown in Fig. 7. The MANN–Kendall test results at a 95% confidence level are according to Table 6.

The table shows that based on the absolute value of the Mann–Kendall test statistics, the NDVI time series diagram of the Shahid Abbaspour Dam, Gotvand Dam, Dezful Regulatory Dam, and Ahvaz Meteorological stations had a trend.

An increase in vegetation density was observed at the Gotvand Dam station from 2010, the Izeh Dam station from 2014, and the Shahid Abbaspour Dam station from 2015.
Fig. 4 Cluster and outlier pattern formation in the studied months using the Anselin Local Moran’s Index
Fig. 4 (continued)
However, a decrease in vegetation density was detected at the Dezful Regulatory Dam and Ahvaz stations from 2014 and 2015, respectively.

**Results of the statistical distribution of the Standardized Normalized Difference Vegetation Index (NDVI) values using Easy-Fit software**

The best statistical distribution for each station was determined with Kolmogorov–Smirnov, Anderson Darling, and Chi-square tests using Easy-Fit software based on NDVI values of each station. By introducing the NDVI value of the drought onset threshold to the software, the probability of the drought onset, the amount of drought values, the risk of drought occurrence, and the total damage caused by drought were obtained. The results are shown in Table 7.

After introducing the stations, the results of the best statistical distribution are given. In the fifth column, the drought onset threshold is between 0.12 and 0.18 based on the value indicating the NDVI thinned vegetation of each station. Accordingly, the probability of drought onset for each station is given in the sixth column. The two southern and adjacent stations, Omidiyeh-Aghajari and Mahshahr, are more likely to start drought. Column 7 shows what percentage of the NDVI values for each station is less than the drought threshold introduced. According to the results of the seventh column, the Ahvaz, Mahshahr, Omidiyeh-Aghajari stations, and the Dez dam have the highest risk of drought. Drought values set for the Ahvaz, Abadan, and Shahid Abba-spour stations estimate the degree of damage to be more than 50% of the total damage, the results of which are given in the last column.

For investigating the reliability of the results, 2000 NDVI artificial data were generated for each station. Then, the best distribution diagram was determined for the data. Afterward, by introducing the drought onset threshold (drought threshold), the results of the artificial data were again compared with those of real NDVI data. In all the cases, the results obtained from the artificial data were very close to those obtained from the real data, and in many of the cases, the results of both the data groups were the same.

**Conclusion**

In this study, the rainfall and temperature data of 12 stations were used to identify drought behavior in Khuzestan Province. The study findings showed that the use of SPEI alone could not be sufficient to identify drought behavior. We attempted to achieve annual and monthly drought trends using NDVI and SPEI as well as satellite images. However, we observed no significant trend in time series charts in many of the cases. Spatial statistics, especially identifying hot and cold spots, along with studies with two indices, provide more comprehensive information on drought behavior.

Figures 4 and 5 and Tables 3 and 4 are presented using the Getis-Ord G* statistics results, which can be used as forecast maps for drought warning. The figures and tables mentioned above show drought prone places referred to as drought hot spots and the range of drought impact. The reliability degree of these hot spots was identified, and their high- or low-risk percentage was determined between 90–99% based on the confidence level. Hot spots detected based on SPEI, and NDVI values (indicating thinned vegetation) can be used to identify drought onset and the focus of dust.

As mentioned above, the fixed ranges defined for the drought severity classification of SPEI and NDVI values did not contribute to the detection of drought severity in

| Station                      | October | November | December | January | February | March | April | May |
|------------------------------|---------|----------|----------|---------|----------|-------|-------|-----|
| Shohada Dam                  | H–H     | –        | H–H      | –       | –        | –     | –     | –   |
| Dez Dam                      | L–L     | –        | –        | –       | –        | –     | –     | L–L |
| Izeh Dam                     | L–L     | H–H      | H–H      | L–L     | H–H      | H–H   | L–L   | L–L |
| Abbaspour                    | L–L     | H–H      | H–H      | L–L     | H–H      | H–H   | L–L   | L–L |
| Gotvand Dam                  | L–L     | –        | –        | L–L     | –        | –     | –     | L–L |
| Dezful Diversion Dam         | L–L     | –        | –        | –       | –        | –     | –     | L–L |
| Abadan Station               | H–H     | L–L      | L–L      | H–H     | L–L      | L–L   | H–H   | H–H |
| Ahwaz Station                | –       | –        | L–L      | –       | –        | –     | –     | –   |
| Mahshahr Station             | H–H     | –        | –        | –       | –        | –     | H–H   | –   |
| Omidiyeh–Aghajari Station    | H–H     | –        | –        | –       | –        | –     | –     | –   |
Fig. 5 Determination of hot and cold drought spots of the Standardized Precipitation Evapotranspiration Index (SPEI) using Getis Ord G * statistics in the studied months
Fig. 5 (continued)
the region. Therefore, the average and standard deviation were used to determine drought severity for each station. According to the study results, it is better to determine SPEI and NDVI in each region according to mean values and standard deviation in that region. It means that drought severity specific to each region should be classified.
Fig. 6  Time series diagram of the Standardized Precipitation Evapotranspiration Index (SPEI) of the stations
according to values obtained for the drought index in that region. It should be noted that in a study with spatial statistics, results of these classifications are not needed, and the software introduces the type of drought behavior or clustering based on similar drought values. Only during evaluation with spatial statistics and preparation of a zoning map using the IDW method, the mentioned classifications can provide a better view of the drought situation in the region. The investigation of SPEI with spatial statistics showed that the highest drought value with hot spots was formed at the southern stations in October, November, April, and May.

Moreover, it was observed that drought hot spots were formed at the eastern stations from November to March. However, no significant behavior was detected for the northern stations. The investigation of NDVI with spatial statistics showed that the formed spots had low vegetation density at the southern stations.

At the eastern stations, both hot and cold spots were formed. The cold spots revealed low vegetation density in autumn from October to December, and the hot spots showed high vegetation density from winter (January) to early spring (May). Most of the hot spots were created for the northern stations.

Comparing the spatial analysis results of both SPEI and NDVI showed that drought severity was higher at the southern stations than at the northern and northwestern stations in Khuzestan Province. The southeastern stations were also somewhat exposed to drought.

The decreasing SPEI trend in the time series diagram of the Dez Dam, Izeh Dam, Shahid Abbaspour Dam, Gotvand Dam, and Dezful Regulatory Dam stations showed the increased drought severity from 2012 onward. The Mann–Kendall test showed that the Dez Dam, Shahid Abbaspour Dam, Izeh Dam, Gotvand Dam, and Dezful Regulatory Dam stations had a trend. Moreover, the Mann–Kendall test for the NDVI time series showed that the Shahid Abbaspour Dam, Gotvand Dam, and Dezful Regulatory Dam stations had a trend. A decrease in vegetation density was observed at the Dezful Regulatory Dam station from 2014 and the Ahvaz station from 2015.

The NDVI values representing drought onset in Easy Fit software showed that the probability of drought was higher on the southern stations of Mahshahr and Omidiyeh–Aghajari. In addition, the Abadan, Ahvaz, and Omidiyeh–Aghajari stations, the Dez dam, and the Shahid Abbaspour dam showed a higher percentage of drought-induced damage.

Spatial statistics techniques, such as Moran, Getis-Ord G *, and Anselin, can be used in water issues studies, such as those on flood warning forecasts, groundwater behavior changes, precipitation behavior detection, water quality, and the environment. In addition, the spatial statistics results can be used in various fields of civil engineering research, such as earthquake risk studies, urban traffic behavior studies, and land subsidence studies.

The present study revealed that primary data from other drought indices was the basis for using spatial statistics to study drought. The calculation of drought indices requires different parameters. The difference between these parameters can affect the result accuracy for each index. This study used the Hargreaves-Samani method to calculate SPEI due to the data limitations. The lack of required rainfall data from the studied stations in the western part of the region was another limitation of this study. The idea of using satellite imagery for drought studies that do not require the collection of various parameters to calculate drought and are not dependent on the meteorological station data can increase the accuracy of spatial statistics results.

| Station                  | Statistics $z$ | Null hypothesis | Hypothesis one | Trend |
|--------------------------|----------------|----------------|----------------|-------|
| Shohada Dam              | -0.19          |                |                |       |
| Dez Dam                  | -3.64          |                |                |       |
| Izeh Dam                 | -2.12          |                |                |       |
| Abbaspour                | -2.58          |                |                |       |
| Gotvand Dam              | -2.2           |                |                |       |
| Dezful Diversion Dam     | -2.58          |                |                |       |
| Abadan Station           | 0.04           |                |                |       |
| Ahwaz Station            | -0.38          |                |                |       |
| Mahshahr Station         | 0.53           |                |                |       |
| Omidiyeh–Aghajari Station| 1.36           |                |                |       |

Table 5: The Mann–Kendall test results at the 95% level for trend evaluation based on the Standardized Precipitation Evapotranspiration Index (SPEI)
Fig. 7 Time series diagram of the Standardized Normalized Difference Vegetation Index (NDVI) of the stations
Acknowledgements The authors would like to thank the Khuzestan Regional Water Company and the Meteorological Organization of Iran for providing information for this study.

Funding No funding was received for conducting this study.

Availability of data and material All data and models that support the findings of this study are available from the corresponding author upon reasonable request.

Declarations

Conflicts of interest The authors declare that there is no conflict of interest.

References

Afshar MH, Al-Yaari A, Yilmaz MT (2021) Comparative evaluation of microwave L-band VOD and optical NDVI for agriculture drought detection over central Europe. Remote Sens 13:1251. https://doi.org/10.3390/rs13071251

AghaKouchak A, Farahmand A, Melton FS et al (2015) Remote sensing of drought: progress, challenges and opportunities. Rev Geophys 53:452–480. https://doi.org/10.1002/2014RG000456

Aljani B, Mahmoudi P, Chogan AJ (2012) Investigation of the trend of changes in annual and seasonal annual precipitation in Iran using non-parametric method (estimation of Sense slope). J Climatol Res 112(3):553–564

Alizadeh A (2001) Principles of applied hydrology, 20th edn. Astan Quds Razavi, Mashhad, Iran

Anselin L (2000) spatial econometrics and social science research. J Geogr Syst 2:11–15
Vicente-Serrano SM, López-Moreno JI, Drumond A et al (2011) Effects of warming processes on droughts and water resources in the NW Iberian Peninsula (1930–2006). Clim Res 48:203–212
Vicente-Serrano SM, Van der Schrier G, Beguería S et al (2015) Contribution of precipitation and reference evapotranspiration to drought indices under different climates. J Hydrol 526:42–54. https://doi.org/10.1016/j.jhydrol.2014.11.025
Wang K, Li Q, Yang Y et al (2015) Analysis of spatio-temporal evolution of droughts in Luanhe River Basin using different drought indices. Water Sci Eng 8:282–290. https://doi.org/10.1016/j.wse.2015.11.004
Wang Q, Liu Y, Zhang Y et al (2019) Assessment of spatial agglomeration of agricultural drought disaster in China from 1978 to 2016. Sci Rep 9:14393. https://doi.org/10.1038/s41598-019-51042-x
Xie F, Fan H (2021) Deriving drought indices from MODIS vegetation indices (NDVI/EVI) and land surface temperature (LST): Is data reconstruction necessary? Int J Appl Earth Obs Geoinf 101:102352. https://doi.org/10.1016/j.jag.2021.102352
Zhang J, Mu Q, Huang J (2016) Assessing the remotely sensed drought severity index for agricultural drought monitoring and impact analysis in North China. Ecol Indic 63:296–309. https://doi.org/10.1016/j.ecolind.2015.11.062
Zhang X, Yamaguchi Y, Li F et al (2017) Assessing the Impacts of the 2009/2010 drought on vegetation indices, normalized difference water index, and land surface temperature in Southwestern China. Adv Meteorol 2017:1–9. https://doi.org/10.1155/2017/6837493
Zhao T, Zhang W, Zhang Y et al (2020) Significant spatial patterns from the GCM seasonal forecasts of global precipitation. Hydrol Earth Syst Sci 24:1–16. https://doi.org/10.5194/hess-24-1-2020