DOI:10.28974/idojaras.2025.1.6

IDŐJÁRÁS

Quarterly Journal of the HungaroMet Hungarian Meteorological Service
Vol. 129, No. 1, January – March, 2025, pp. 89–105

Clustering of the Black Sea Region meteorological stations of Türkiye with fuzzy c-means, k-means, and silhouette index analysis methods by precipitation, temperature and wind speed

Aslı Ulke Keskin[1], **Gurkan Kır**[1], and **Utku Zeybekoglu**[2,*]

[1]*Faculty of Engineering, Ondokuz May University*
55139, Atakum, Samsun, Türkiye

[2]*Boyabat Vocational School of Higher Education,*
Sinop University, 57200 Boyabat, Sinop, Türkiye

* *Correspondening Author E-mail*: *utkuz@sinop.edu.tr*

(Manuscript received in final form April 11, 2024)

Abstract— Recent years have seen a marked increase in the number of disasters caused by the effects of global climate change. In response, a range of studies have been conducted in Türkiye and worldwide with the aim of reducing the impact of climate change.The classification of regions affected by climate change into similar classes in terms of climate parameters is crucial for the application of consistent methods in studies conducted in these regions. Consequently, the formulation of effective strategies to mitigate the repercussions of climate change in these regions is contingent upon the accurate determination of the aforementioned strategy.The observation records evaluated within the scope of the study were obtained from 31 stations of the Turkish State Meteorological Service in the Black Sea Region, encompassing the period between 1982 and 2020, encompassing precipitation, temperature, and wind speed records.. The maximum number of clusters was determined as 5, the cluster analysis study was carried out by using fuzzy c-means and k-means methods for 2, 3, 4, and 5 cluster numbers according to these three data together form a matrix. The determination of the optimum cluster numbers was carried out by silhouette index analysis. For the data matrix where precipitation, temperature, and wind speed were evaluated together, the most appropriate classification was obtained by the k-means method by choosing the number of clusters as 4.

Key-words: cluster analysis, silhouette index analysis, wind speed, precipitation, temperature, fuzzy c-means, k-means, Türkiye

1. Introduction

Climate, defined as the extreme values of various meteorological parameters such as precipitation, temperature, and wind, is the collective state of the atmosphere for a specific location and time period (*Demircan et al.*, 2017).It was widely accepted that there was no change in the long-term averages of the parameters of this collective structure until the mid-1950s. However, as the 20th century progressed, the rapid advancements in industry led to unplanned consumption of natural resources, escalating environmental pollution levels relative to the population, and substantial increases in greenhouse gas emissions into the atmosphere.Consequently, the increasing presence of greenhouse gases, which are capable of trapping heat in the atmosphere, has initiated a series of alterations in climate parameters over time. These alterations in climate parameters are designated as global climate change (*Türkeş*, 2010; *Özkoca*, 2015).The phenomenon of climate change, in its global manifestation, has been demonstrated to manifest locally in the form of various disasters, including but not limited to floods, droughts, and storms. The escalating impact of global climate change, a subject that has been extensively documented in numerous articles in recent years, has been shown to exert a detrimental effect on human life in economic and social spheres.In this regard, studies undertaken to comprehend climate change and to devise measures in this context are also assuming increasing significance. The classification of regions exhibiting analogous climatic characteristics is believed to facilitate various studies, including those focused on combating climate change, safeguarding water resources, and strategising land use. Cluster analysis is a methodology that has been used in climatology research for at least 30 years (*Kalkstein et al.*, 1987; *Fovell* and *Fovell* 1993). *Erinç* (1949) classified precipitation and temperature data obtained from 53 meteorological stations in Türkiye for 4 different climate zones using the Thornthwaite method. With this study, the regional and detailed classification of Türkiye's geography with sufficient data was carried out for the first time. *Türkeş* (1996) classified the precipitation data of Türkiye with the help of the normalization procedure method proposed by *Kraus* in 1977. In the study, in which the aspirations of the 1930-1993 period were used, 7 different regions were determined. *Kulkarni* and *Kripalani* (1998), using the fuzzy c-means method, determined the similar classes of Indian precipitation data. Using the precipitation data for the 1871–1984 period, 306 meteorological stations were divided into 4 different clusters. *Unal et al.* (2003) determined the similar classes of temperature and precipitation data covering the period 1951–1998 in Türkiye with 5 different clustering methods. In the study where the single linkage, complete linkage, centroid, Ward’s minimum variance, and average distance methods were used, it was stated that the most effective method was the Ward's method. *Soltani* and *Modarres* (2006) divided the precipitation data of 28 stations in Iran into similar classes with the help of hierarchical and non-hierarchical clustering methods. In the study in which 8 different classes were determined, the Ward's method and the k-means algorithm were used. *Sönmez* and *Kömüşcü* (2008) used the k-means algorithm in their study in which they determined the precipitation regions of Türkiye. In the study, in which monthly total precipitation series covering the years 1977–2006 obtained from 148 stations were used, 6 different precipitation regions were determined. *Şahin* (2009) used monthly average temperature, monthly relative humidity, and monthly total precipitation data obtained from 150 stations to determine similar climate classes in Türkiye. Using the Ward's method, the Kohonen artificial neural network, and the fuzzy artificial neural network, 7 different regions were determined. *Dikbas et al.* (2012) determined 6 different precipitation regions by using the 1967-1998 records of 188 stations in Türkiye using the fuzzy c-means method. *Şahin* and *Cığızoğlu* (2012) determined the sub-climate and sub-precipitation regime classes of Türkiye by using the Ward's method and the fuzzy artificial neural network. Using the precipitation, temperature, and humidity data of 232 stations in the 1974–2002 period, 7 precipitation regime regions and 7 climate regions were determined. *Firat et al.* (2012) determined 7 different regions with similar characteristics by using the k-means method of the classes of annual total precipitation, which was measured at 188 precipitation observation stations in Türkiye covering the period of 1967–1998. In the study of *İyigün et al.* (2013), a clustering analysis study was carried out with precipitation, temperature, and relative humidity data using the Ward's method. It was obtained from 244 stations in Türkiye and its period covered the years 1970–2010. As a result of the study, 14 different clusters were identified. *Rau et al.* (2017) divided the precipitation data of the Peruvian Pacific slope and coast into regions with similar characteristics. Using the regional vector method and the k-means algorithm, 9 different precipitation regions were determined. *Zeybekoğlu* and *Ülke Keskin* (2020) realized clustering analysis by adding the latitudes, longitudes, and elevations of the stations to the precipitation intensity series using the fuzzy c-means algorithm. It has been determined that 95 meteorological observation stations in Türkiye form 5 different clusters.

A plethora of studies have been conducted on the determination of climate classes in the literature. A close examination of these studies reveals that they predominantly emphasise precipitation and temperature data as climate parameters.Furthermore, the utilisation of silhouette index analysis for the evaluation of clusters determined by using fuzzy c-means and k-means methods is not a prevalent practice in climate studies, to the best of the authors' knowledge (*Kır*, 2021). The Black Sea Region, selected as the study area, has been experiencing the impacts of global climate change, manifesting in the form of disasters such as floods, droughts, and severe storms.The objective of this study is to identify similar clusters in the wind speed series of the Black Sea Region by employing various clustering algorithms.The analyses conducted for different cluster numbers using the fuzzy c-means and k-means methods were used to determine the most suitable cluster number by means of the silhouette index analysis. This analysis, conducted using a matrix encompassing wind speed parameters alongside precipitation and temperature metrics, represents a pioneering endeavour within the context of the Black Sea Region, particularly within the context of Türkiye.

2. Materials and methods

2.1. Materials

The present study utilised observations of annual precipitation, annual temperature, and annual wind speed, which were recorded over a period of 39 years (1982–2020) at 31 observation stations operated by the Turkish State Meteorological Service in the Black Sea Region of Türkiye. In order to ensure sufficient statistical validity, it was imperative to consider a minimum record length of 30 years (*Kite*, 1991).The observation stations utilised in this study are situated in 17 distinct provinces across the Black Sea Region. Eleven of the stations are located in the western Black Sea region, including Duzce, Akçakoca, Bolu, Zonguldak, Bartın, Amasra, Kastamonu, İnebolu, Bozkurt, Tosya, Sinop. The remaining ten stations are situated in the central Black Sea region, Samsun, Bafra, Çorum, Osmancık, Amasya, Merzifon, Tokat, Zile, Ordu and Ünye. The final ten stations are located in the eastern Black Sea Region, including Giresun, Şebinkarahisar, Trabzon, Akçaabat, Gümüşhane, Bayburt, Rize, Pazar, Artvin and Hopa. Comprehensive details concerning the stations are delineated in *Table* 1, while the geographical distribution of the stations is exhibited in *Fig. 1*.

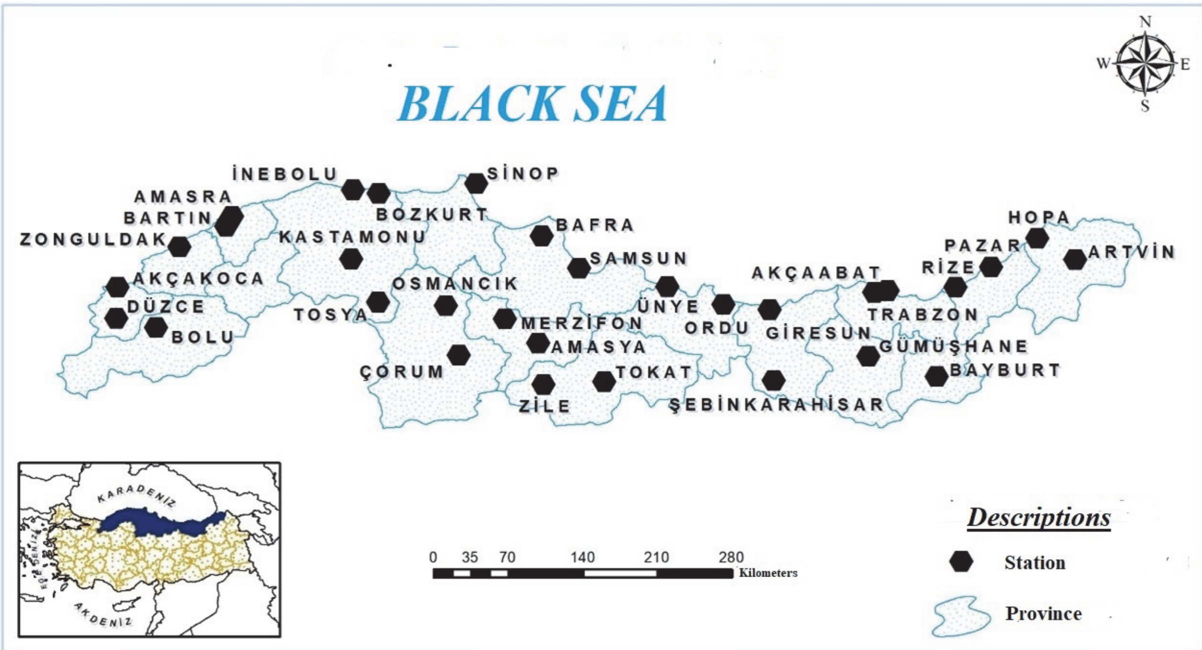

Fig. 1. Geographical distribution of meteorological stations

Table 1. List of meteorological stations and geographical details

Station Name	Station No	Latitude (N)	Longitude (E)	Elevation (m)	Period
Düzce	17072	40°50'37.3"	31°08'55.7"	146	1982-2020
Akçakoca	17015	41°05'22.2"	31°08'14.6"	10	1982-2020
Bolu	17070	40°43'58.4"	31°36'07.9"	743	1982-2020
Zonguldak	17022	41°26'57.3"	31°46'40.5"	135	1982-2020
Bartın	17020	41°37'29.3"	32°21'24.8"	33	1982-2020
Amasra	17602	41°45'09.4"	32°22'57.7"	73	1982-2020
Kastamonu	17074	41°22'15.6"	33°46'32.2"	800	1982-2020
İnebolu	17024	41°58'44.0"	33°45'49.0"	64	1982-2020
Kastamonu/Bozkurt	17606	41°57'34.9"	34°00'13.3"	167	1982-2020
Tosya	17650	41°00'47.5"	34°02'12.1"	870	1982-2020
Çorum	17084	40°32'46.0"	34°56'10.3"	776	1982-2020
Osmancık	17652	40°58'43.3"	34°48'04.0"	419	1982-2020
Sinop	17026	42°01'47.6"	35°09'16.2"	32	1982-2020
Amasya	17085	40°40'00.5"	35°50'07.1"	409	1982-2020
Merzifon	17083	40°52'45.5"	35°27'30.6"	754	1982-2020
Samsun Bölge	17030	41°20'39.0"	36°15'23.0"	4	1982-2020
Bafra	17622	41°33'05.4"	35°55'28.9"	103	1982-2020
Tokat	17086	40°19'52.3"	36°33'27.7"	611	1982-2020
Zile	17681	40°17'45.6"	35°53'25.8"	719	1982-2020
Ordu	17033	40°59'01.7"	37°53'08.9"	5	1982-2020
Ünye	17624	41°08'34.8"	37°17'34.8"	16	1982-2020
Giresun	17034	40°55'21.7"	38°23'16.1"	38	1982-2020
Şebinkarahisar	17682	40°17'13.9"	38°25'09.5"	1364	1982-2020
Gümüşhane	17088	40°27'35.3"	39°27'55.1"	1216	1982-2020
Trabzon Bölge	17037	40°59'54.6"	39°45'53.6"	25	1982-2020
Akçaabat	17626	41°01'57.0"	39°33'41.4"	3	1982-2020
Bayburt	17089	40°15'16.9"	40°13'14.5"	1584	1982-2020
Rize	17040	41°02'24.0"	40°30'04.7"	3	1982-2020
Rize/Pazar	17628	41°10'39.7"	40°53'57.5"	78	1982-2020
Artvin	17045	41°10'30.7"	41°49'07.3"	613	1982-2020
Hopa	17042	41°24'23.4"	41°25'58.8"	33	1982-2020

2.2. k-means algorithm

K-means, one of the oldest clustering algorithms, was developed in 1967 by MacQueen (*MacQueen*, 1967). The assignment mechanism of k-means, one of the most widely used unsupervised learning methods, allows each data to belong to only one cluster. Therefore, it is a sharp clustering algorithm. It is a method based on the main idea that the central point represents the cluster (*Han* and *Kamber*, 2006). It tends to find globular clusters of equal size (*Isık* and *Camurcu,* 2007).

The sum of squared errors criterion (SSE) is most commonly used in the evaluation of the k-means clustering method. The clustering result with the lowest

SSE value gives the best results. The sum of the squares of the distances of the objects from the center points of the cluster they are located in is calculated with the following formula (*Pang-Ning et al.,* 2006; *Isık* and *Camurcu*, 2007):

$$SSE = \sum_{i=1}^{K} \sum_{x \in C_i} dist^2(m_i, x). \tag{1}$$

As a result of this criterion, it is aimed to result in k clusters as dense and separate from each other as possible. The algorithm tries to identify *k* pieces that will reduce the quadrature-error function. The k-means algorithm divides the data set consisting of *n* data into k clusters with the *k* parameter given to the algorithm by the user. Cluster similarity is measured by the mean value of the objects in the cluster, which is the center of gravity of the cluster (*Xu* and *Wunsch*, 2005; *Isık* and *Camurcu*, 2007).

2.3. Fuzzy c-means (FCM) algorithm

Fuzzy c-means (FCM) algorithm is the best known and widely used method among fuzzy division clustering techniques. The fuzzy c-means algorithm was introduced by *Dunn* (1974) and developed by *Bezdek* (1981) (*Höppner et al.,* 2000). The fuzzy c-means algorithm is also an objective function-based method. The FCM allows objects to belong to two or more clusters. According to the fuzzy logic principle, each data belongs to each of the clusters with a membership value varying between [0,1]. The sum of the membership values of a data to all classes must be “1”. Whichever cluster center the object is close to, the membership of that cluster will be larger than the membership of other clusters. The clustering process is completed when the objective function converges to the determined minimum progress value (*Isık* and *Camurcu*, 2007).

The algorithm works to minimize the following objective function, which is a generalization of the least squares method (*Höppner et al.,* 2000; *Isık* and *Camurcu*, 2007):

$$J_m = \sum_{i=1}^{N} \sum_{j=1}^{C} u_{ij}^m \left\| x_i - c_j \right\|^2; 1 \leq m < \infty . \tag{2}$$

The algorithm is started by randomly assigning the U membership matrix. In the second step, the center vectors are calculated. The centers are calculated with the following formula(*Höppner et al.,* 2000; *Isık* and *Camurcu*, 2007):

$$c_j = \frac{\sum_{i=1}^{N} u_{ij}^m x_i}{\sum_{i=1}^{N} u_{ij}^m} . \tag{3}$$

According to the calculated cluster centers, the U matrix is recalculated using the following formula:

$$u_{ij} = \frac{1}{\sum_{k=1}^{C}\left(\frac{\|x_i - c_j\|}{\|x_i - c_k\|}\right)^{2/(m-1)}} . \tag{4}$$

The old U matrix and the new U matrix are compared, and the process continues until the difference is less than ε (*Moertini*, 2002; *Isık* and *Camurcu,* 2007). As a result of the clustering operation, the U membership matrix containing fuzzy values reflects the result of the clustering. If desired, these values can be rounded to 0 and 1 by clarification (*Isık* and *Camurcu*, 2007).

2.4. Silhouette index analysis

In this method developed by *Rousseeuw* (1987), the suitability of each element in the data set to the cluster to which it is assigned is defined by the silhouette index value obtained between [-1, +1]. A positive silhouette index value indicates that the element is assigned to the correct cluster, and a negative value indicates that the element is assigned to the wrong cluster. The amount of the silhouette index value indicates the degree of membership in the cluster to which the element is assigned. The silhouette index value is calculated by the following formula (*Günay Atbaş*, 2008; *Sönmez* and *Kömüşcü* 2008):

$$S(i) = \frac{min\{b(i,m) - a(i)\}}{max\{a(i), min(b(i,m))\}} , \tag{5}$$

where $a(i)$ is the average distance between point i and all other points in the same cluster, $b(i,m)$ represents the average distance between the ith point and all the points in the mth cluster.

3. Results and discussion

In this study, the k-means and FCM methods were employed to ascertain clusters exhibiting analogous characteristics. This was achieved by utilising a matrix comprising annual total precipitation, annual average temperature, and annual average wind speed observations, encompassing the period between 1982 and 2020, for a total of 31 stations.The analyses were conducted within the MATLAB R2016a software framework. The number of clusters was determined to be five, ensuring that the total number of clusters was less than the square root of the number of stations (*Pal* and *Bezdek*, 1995; *Zhang et al.,* 2008; *Karahan*, 2011, 2019). Since precipitation, temperature, and wind speed are variables with different scales, these data were standardized before classification (*Ünal et al.,* 2003; *Lin* and *Chen*, 2006; *Cannarozzo et al.,* 2009; *Lim* and *Voeller*, 2009; *Dikbas et al.,* 2012; *Firat et al.,* 2012) as follows:

$$z = \frac{x_i - \bar{x}}{s}, \tag{6}$$

where x_i denotes the value to be standardized, $\bar{x}$ is the mean of data, s denotes the standard deviation, and z is the standardized hydrometeorological data.

The classification results obtained when the number of clusters was 2 using the k-means method of the series consisting of precipitation, temperature, and wind speed variables are shown in *Fig.2*. When the results are examined, cluster A consists of 16 stations located in the western, central, and eastern Black Sea coastal areas. Cluster B consists of 15 stations located in the western, central, and eastern Black Sea Regions. The maximum, minimum, and mean values of precipitation, temperature, and wind speed of the determined clusters are presented in *Table 2*.

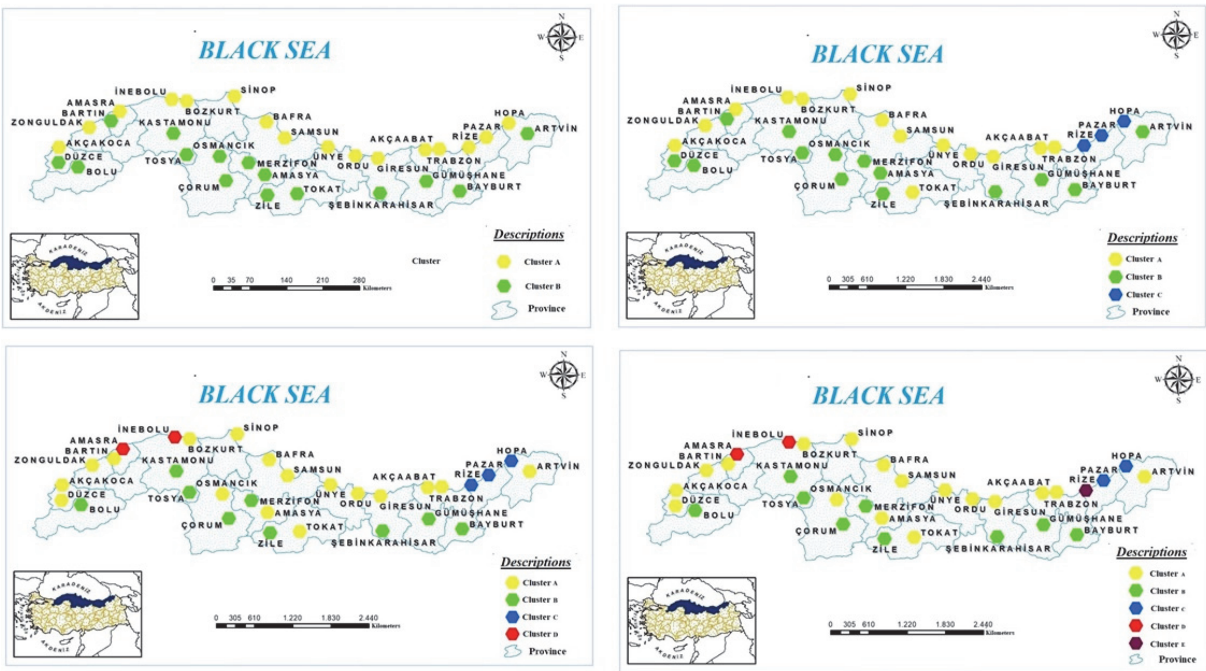

Fig. 2. Geographical distribution of stations for k-means

*Table 2***.** Minimum, maximum, and mean values, and standard deviations for meteorological records from 2 clusters solution for k-means

Cluster	Precipitation (mm)				Temperature (°C)				Wind speed (m s^{-1})			
	Min.	Max.	Mean	Std. Dev.	Min.	Max.	Mean	Std. Dev.	Min.	Max.	Mean	Std. Dev.
A	716.5	2329.7	1225.8	540.3	13.2	15.0	14.2	0.6	1.2	4.8	2.2	0.9
B	423.4	1051.1	555.2	177.1	7.2	13.6	11.4	1.8	0.5	2.2	1.4	0.5

The classification results obtained when the number of clusters is selected as 3 using the k-means method is shown in *Fig.2*. When the results are examined, it is seen that cluster C is separated as a subset of cluster A in the previous distribution. Here, clusters A and B consist of 14 stations located in the western, central, and eastern Black Sea Regions. Cluster C consists of 3 stations located in the eastern Black Sea coast. The maximum, minimum, and mean values of the observations in the determined clusters are presented in *Table 3*.

Table 3. Minimum, maximum, and mean values, and standard deviations for meteorological records from 3 clusters solution for k-means

Cluster	**Precipitation (mm)**				**Temperature (°C)**				**Wind speed (m s^{-1})**			
	Min.	**Max.**	**Mean**	**Std. Dev.**	**Min.**	**Max.**	**Mean**	**Std. Dev.**	**Min.**	**Max.**	**Mean**	**Std. Dev.**
A	444.3	1308.1	952.7	253.1	12.6	15.0	14.1	0.7	1.2	4.8	2.3	0.9
B	423.4	1051.1	563.2	181.0	7.2	13.6	11.3	1.9	0.5	1.9	1.3	0.4
C	2105.4	2329.7	2239.8	118.6	13.8	14.8	14.4	0.5	1.2	2.3	1.8	0.6

The classification results obtained when the number of clusters is selected as 4 using the k-means method is shown in *Fig. 2*. When the results are examined, it is seen that cluster D is separated as a subset of cluster A in the previous distribution. In addition, it is seen that cluster C maintains its integrity in the previous distribution. Thus, cluster A consists of 17 stations located in the western, central, and eastern Black Sea Regions. Cluster B consists of 9 stations located in the inner parts of the western, central, and eastern Black Sea Regions. Clusters C and D, on the other hand, consist of 2 and 3 stations located in the western and eastern Black Sea coastal areas, respectively. The maximum, minimum, and average values of the observations in the determined clusters are presented in *Table 4*.

Table 4. Minimum, maximum, and mean values, and standard deviations for meteorological records from 4 clusters solution for k-means

Cluster	**Precipitation (mm)**				**Temperature (°C)**				**Wind speed (m s^{-1})**			
	Min.	**Max.**	**Mean**	**Std. Dev.**	**Min.**	**Max.**	**Mean**	**Std. Dev.**	**Min.**	**Max.**	**Mean**	**Std. Dev.**
A	423.4	1308.1	869.5	285.3	12.2	15.0	13.9	0.8	1.0	2.8	1.8	0.5
B	444.4	568.6	489.4	49.1	7.2	11.9	10.3	1.4	0.5	1.9	1.3	0.5
C	2105.4	2329.7	2239.8	118.6	13.8	14.8	14.4	0.5	1.2	2.3	1.8	0.6
D	981.6	1053.8	1017.7	51.0	13.5	14.0	13.7	0.3	3.6	4.8	4.2	0.9

The classification results obtained when the number of clusters is selected as 5 using the k-means method is shown in *Fig.2.* When the results are examined, it is seen that cluster E is separated as a subset of cluster C in the previous distribution. Thus, cluster A consists of 17 stations located in the western, central, and eastern Black Sea Regions. Cluster B consists of 9 stations located in the inner parts of the western, central, and eastern Black Sea Regions. Clusters C and E consist of 2 and 1 stations, respectively, located in the Eastern Black Sea coast. Cluster D consist of 2 stations located in the western Black Sea coast. The maximum, minimum, and mean values of the observations in the determined clusters are presented in *Table 5*.

Table 5. Minimum, maximum, and mean values, and standard deviations for meteorological records from 5 clusters solution for k-means

Cluster	Precipitation (mm)				Temperature (°C)				Wind speed (m s^{-1})			
	Min.	**Max.**	**Mean**	**Std. Dev.**	**Min.**	**Max.**	**Mean**	**Std. Dev.**	**Min.**	**Max.**	**Mean**	**Std. Dev.**
A	423.4	1308.1	869.5	285.3	12.2	15.0	13.9	0.8	1.0	2.8	1.8	0.5
B	444.4	568.6	489.4	49.1	7.2	11.9	10.3	1.4	0.5	1.9	1.3	0.5
C	2105.4	2329.7	2217.6	158.6	13.8	14.8	14.3	0.7	1.8	2.3	2.1	0.3
D	981.6	1053.8	1017.7	51.0	13.5	14.0	13.7	0.3	3.6	4.8	4.2	0.9
E	2284.4	2284.4	2284.4	-	14.7	14.7	14.7	-	1.2	1.2	1.2	-

The classification results obtained when the number of clusters is selected as 2 using the FCM is shown in *Fig.3*. When the results are examined, cluster A consists of 17 stations located in the western, central, and eastern Black Sea coastal areas. Cluster B consists of 14 stations located in the inner parts of the western, central, and eastern Black Sea Regions. The maximum, minimum, and mean values of the observations in the determined clusters are presented in *Table 6.*

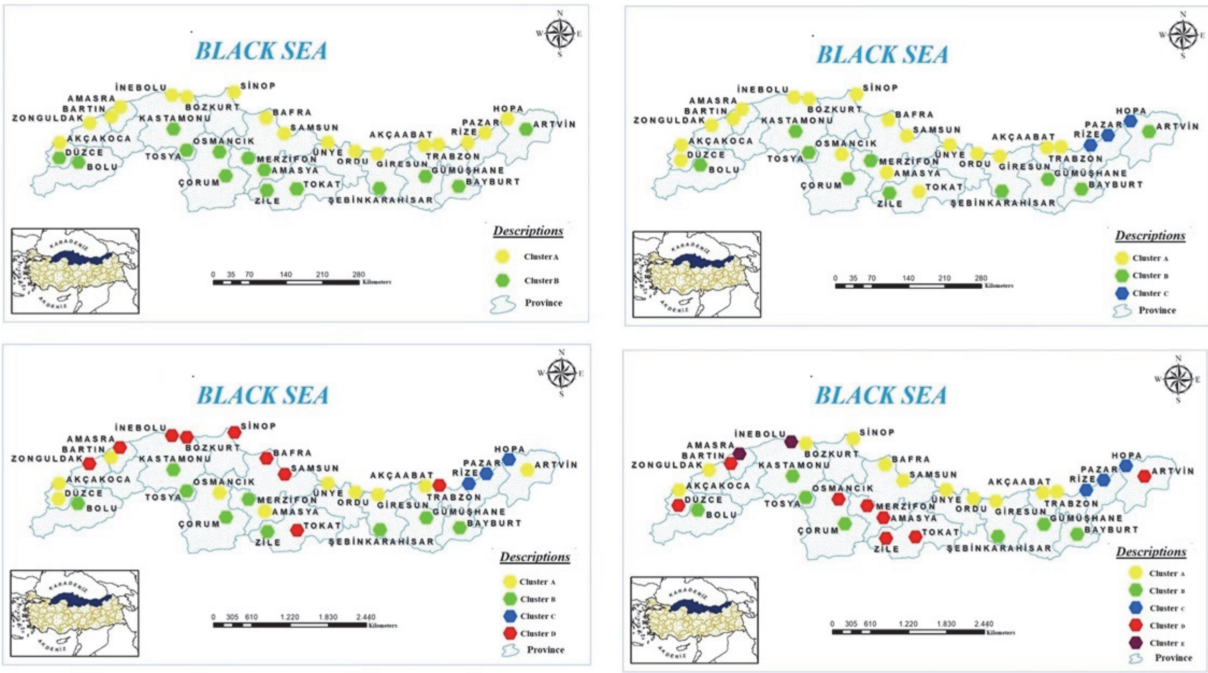

Fig. 3. Geographical distribution of stations for FCM

Table 6. Minimum, maximum, and mean values, and standard deviations for meteorological records from 2 clusters solution for FCM

Cluster	Precipitation (mm)				Temperature (°C)				Wind speed (m s^{-1})			
	Min.	**Max.**	**Mean**	**Std. Dev.**	**Min.**	**Max.**	**Mean**	**Std. Dev.**	**Min.**	**Max.**	**Mean**	**Std. Dev.**
A	716.5	2329.7	1215.5	524.9	12.9	15.0	14.2	0.7	1.2	4.8	2.2	0.9
B	423.4	818.4	519.8	116.2	7.2	13.6	11.3	1.8	0.5	2.2	1.4	0.5

The classification results obtained when the number of clusters 3 is selected using the FCM is shown in *Fig.3*. When the results are examined, it is seen that cluster C is separated as a subset of cluster A in the previous distribution. Thus, cluster A consists of 18 stations located in the western, central, and eastern Black Sea Regions. Cluster B consists of 10 stations located in the inner parts of the western, central, and eastern Black Sea Regions. Cluster C consists of 3 stations located in the eastern Black Sea coast. The maximum, minimum, and mean values of the observations in the determined clusters are presented in *Table 7*.

Table 7. Minimum, maximum, and mean values, and standard deviations for meteorological records from 3 clusters solution for FCM

Cluster	**Precipitation (mm)**				**Temperature (°C)**				**Wind speed(m s^{-1})**			
	Min.	**Max.**	**Mean**	**Std. Dev.**	**Min.**	**Max.**	**Mean**	**Std. Dev.**	**Min.**	**Max.**	**Mean**	**Std. Dev.**
A	423.4	1308.1	894.2	278.3	12.6	15.0	13.9	0.7	1.0	4.8	2.1	0.9
B	444.4	721.4	512.6	86.8	7.2	12.2	10.5	1.5	0.5	1.9	1.4	0.5
C	2105.4	2329.7	2239.8	2239.8	13.8	14.8	14.4	14.4	1.2	2.3	1.8	1.8

The classification results obtained when the number of clusters is selected as 4 using the FCM is shown in *Fig.3*. When the results are examined, it is seen that Rize, Pazar, and Hopa stations maintain their integrity. It is seen that cluster B is separated as a subset of cluster A in the previous distribution. Thus, clusters A and D consist of 10 and 9 stations, respectively, located in the western, central, and eastern Black Sea Regions. Cluster B consists of 9 stations located in the inner parts of the western, central and eastern Black Sea Regions. Cluster C consists of 3 stations located in the eastern Black Sea coast. The maximum, minimum, and mean values of the observations in the determined clusters are presented in *Table 8*.

Table 8. Minimum, maximum, and mean values, and standard deviations for meteorological records from 4 clusters solution for FCM

Cluster	**Precipitation (mm)**				**Temperature (°C)**				**Wind speed (m s^{-1})**			
	Min.	**Max.**	**Mean**	**Std. Dev.**	**Min.**	**Max.**	**Mean**	**Std. Dev.**	**Min.**	**Max.**	**Mean**	**Std. Dev.**
A	423.4	1308.1	888.1	304.2	12.2	14.8	13.8	0.9	1.0	1.9	1.5	0.3
B	444.4	568.6	489.4	49.1	7.2	11.9	10.3	1.4	0.5	1.9	1.3	0.5
C	2105.4	2329.7	2239.8	118.6	13.8	14.8	14.4	0.5	1.2	2.3	1.8	0.6
D	444.3	1226.7	881.8	252.9	12.6	15.0	13.9	0.7	2.1	4.8	2.7	0.9

The classification results obtained when the number of clusters is selected as 5 using the FCM is shown in *Fig.3*. When the results are examined, it is seen that only cluster C preserves its integrity. Here, clusters A and D consist of 11 and 8 stations located in the western, central, and eastern Black Sea coastal areas, respectively. Cluster B consists of 7 stations located in the inner parts of the western, central, and eastern Black Sea Regions. Cluster C consists of 3 stations located in the Eastern Black Sea coast. Cluster E consists of 2 stations located in the western Black Sea Region. The maximum, minimum, and mean values of the observations in the determined clusters are presented in *Table 9*.

Table 9. Minimum, maximum, and mean values, and standard deviations for meteorological records from 5 clusters solution for FCM

Cluster	Precipitation (mm)				Temperature (°C)				Wind speed (m s^{-1})			
	Min.	Max.	Mean	Std. Dev.	Min.	Max.	Mean	Std. Dev.	Min.	Max.	Mean	Std. Dev.
A	716.5	1308.1	987.1	234.5	13.2	15.0	14.3	0.6	1.2	2.8	2.0	0.5
B	450.2	568.6	502.2	48.5	7.2	11.4	9.8	1.4	0.9	1.9	1.5	0.3
C	2105.4	2329.7	2239.8	118.6	13.8	14.8	14.4	0.5	1.2	2.3	1.8	0.6
D	423.4	1051.1	601.6	235.4	11.6	13.6	12.7	0.8	0.5	2.2	1.3	0.5
E	981.6	1053.8	1017.7	51.0	13.5	14.0	13.7	0.3	3.6	4.8	4.2	0.9

Silhouette index values of stations in clusters for each number of clusters are presented in *Figs. 4* and *5* for k-means and FCM, respectively. Average silhouette index values and negative silhouette index numbers for each cluster determined by k-means and FCM from the clusters 2 to 5 are presented in *Tables10* and *11*, respectively.

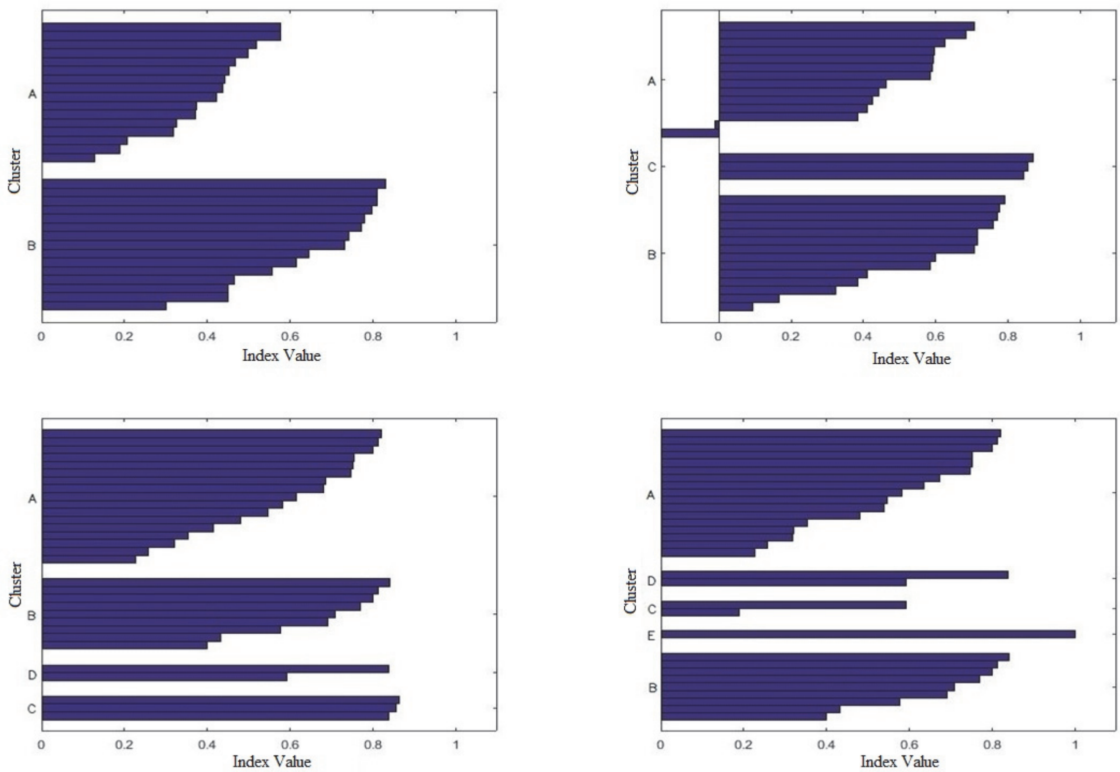

Fig. 4. Silhouette index analysis results of clusters for k-means

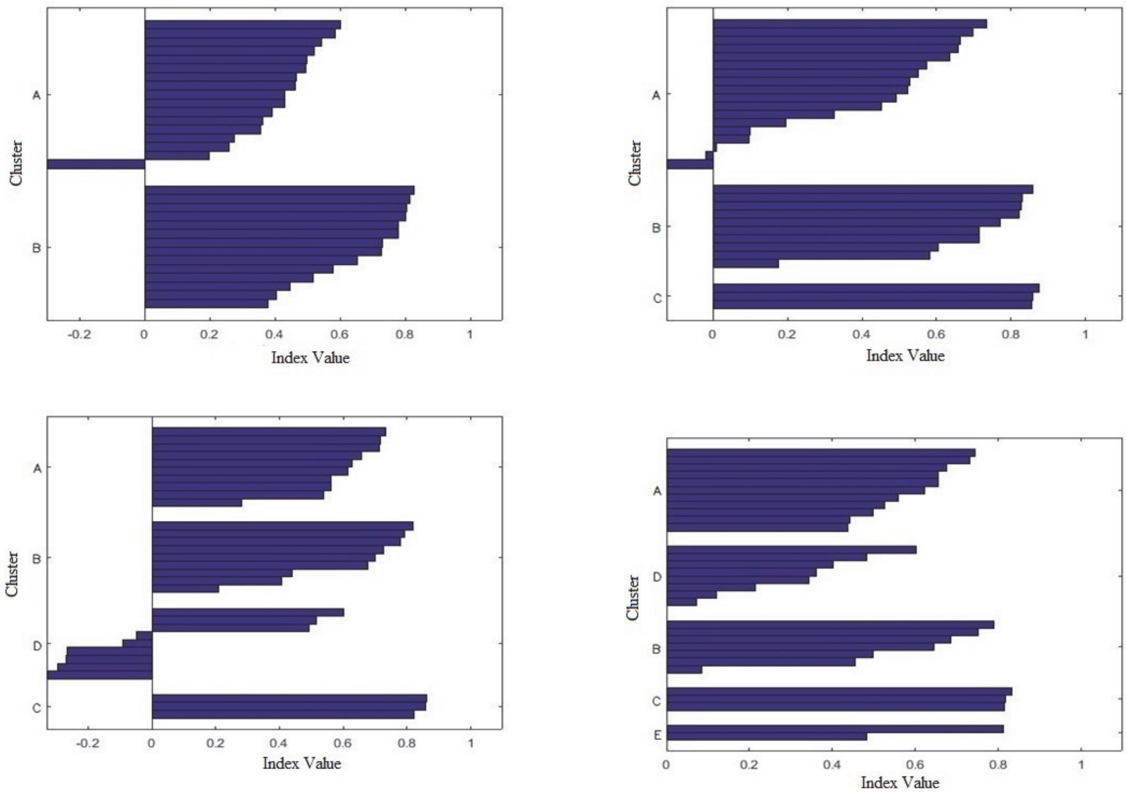

Fig. 5. Silhouette index analysis results of clusters for FCM

Table 10. Silhouette index analysis results for k-means

Cluster number	2	3	4	5
Mean silhouette index value	0.518	0.539	**<u>0.641</u>**	0.608
Number of negative silhouette indexes	-	2	-	-

Table 11. Silhouette index analysis results for FCM

Cluster number	2	3	4	5
Mean silhouette index value	0.510	0.535	0.464	**0.542**
Number of negative silhouette indexes	1	2	6	-

As a result of the silhouette analysis (*Figs. 4* and *5*, *Tables 10* and *11*) among the clusters determined using k-means and FCM, it was determined that the most suitable clusters were 4 clusters formed by the k-means method.

As a result of the analysis, the Black Sea Region stations were determined as 4 similar clusters in terms of precipitation, temperature, and wind speed characteristics. When compared with the studies covering the Black Sea Region (*Turkes*, 1996; *Unal et al.,* 2003; *Iyigun et al.,* 2013; *Ozturk et al.,* 2017; *Zeybekoglu* and *Ulke Keskin*, 2020), the methods used in the clustering analysis of the different clusters resulted in hydrometeorological parameters, different observation periods, sea effect, and the parallelism of the mountains to the coast, mountainous and rugged. It is thought to be caused by regional geographical features.

4. Conclusion

- In this study, the Black Sea Region of Türkiye was selected as the subject of analysis. The precipitation, temperature, and wind speed parameters of meteorological observation stations in this region were analysed to determine clusters with similar characteristics. The K-means and fuzzy c-means algorithms were utilised to identify the clusters. Cluster analysis was conducted for four different cluster numbers, ranging from two to five, and the optimal number of clusters was determined through the implementation of the silhouette index analysis. The findings of this study, incorporating both cluster analysis and silhouette index analysis, indicate that the most appropriate classification result is achieved through the delineation of four clusters employing the k-means method.While the outcomes derived from the k-means and FCM methods appear to be analogous, it is evident that the k-means method yields more favourable results.The subsequent phase of this study will entail:It is recommended to determine climate classes with different and new combinations that are not used much in the literature, by including hydro-meteorological parameters such as flow, humidity, and evaporation, as well as precipitation, temperature, and wind speed parameters.
- Apart from non-hierarchical clustering algorithms, clustering analyses containing different hierarchical clustering algorithms such as Ward's method should be performed in the literature, and comparative studies should be conducted.
- The cluster analysis study should also be applied to other regions in Türkiye's geography.

In the modern understanding of disaster management (*Gunduz,* 2022; *Usta,* 2023), activities should be carried out to identify risks and hazards to take all necessary precautions, to take responsibility for disasters, and to raise awareness of all individuals who make up the society, in order to reduce or prevent the damages of disasters.

Data availability: The meteorological data used in this manuscript were obtained from the Turkish State Meteorological Service (TSMS) for the master's thesis titled "Evaluation of the meteorological data of the Black Sea Region using clustering analysis methods" written by Gurkan Kir under the supervision of Asli Ulke Keskin.

Acknowledgements: The authors thank the Turkish State Meteorological Service for providing the meteorological data. The authors also thank the reviewers for their constructive criticisms which have considerably improved this manuscript.

References

Bezdek, J.C., 1981*: Pattern recognition with fuzzy objective function algorithms.* Plenum Press, USA.

Cannarozzo, M., Noto, L.V., Viola, F., and *La Loggia, G.,* 2009: Annual run-off regional frequency analysis in Sicily. *Phys. Chem. Earth 34,* 679–687. https://doi.org/10.1016/j.pce.2009.05.001

Demircan, M., Arabacı, H., Coşkun, M., Türkoğlu, N., and *Çiçek, İ.,* 2017: İklim değişikliği ve halk takvimi: Maksimum sıcaklık desenleri ve değişimi. IV. Türkiye İklim Değişikliği Kongresi İstanbul Medipol Üniversitesi, İstanbul. (In Türkish)

Dikbas, F., Firat, M., Koc, A.C., and *Gungor M.,* 2012: Classication of precipitation series using fuzzy cluster method. *Int. J. Climatol. 32(*10)*,* 1596–1603. https://doi.org/10.1002/joc.2350

Dunn, J.C., 1974: A fuzzy relative of the ISODATA process and its use in detecting compact, wellseparated clusters. *J. Cybernetics 3,* 32–57. https://doi.org/10.1080/01969727308546046

Erinç, S., 1949: The climates of Turkey according to Thornthwaite's classifications. *Ann. Assoc. Am. Geograp. 39,* 26–46. https://doi.org/10.1080/00045604909351994

Fırat, M., Dikbaş, F., Koç, A.C., and *Güngör, M.,* 2012: Classification of annual precipitations and identification of homogeneous regions using k-means method. *Techn. J. 23*(113), 6037-6050.

Fovell, B., and *Fovell, M.,* 1993: Climate zones of the conterminous United States defined using cluster analysis. *J. Climate 6,* 2103–2135. https://doi.org/10.1175/1520-0442(1993)006<2103:CZOTCU>2.0.CO;2

Gundüz, F., 2022: Lessons Learned from the Perspective of Women and Gender in Disasters, The Case of Haiti, and Japan Earthquake. *IBAD J. Social Scie. 12,* 440–460. https://doi.org.tr/10.21733/ibad.1039215

Günay Atbaş, A.C., 2008: A study on determining the number of clusters in cluster analysis. MSc Thesis, Ankara University, Ankara, Turkey.

Han, J., and *Kamber, M.,* 2006: Data Mining Concepts and Techniques*.* Morgan Kauffmann Publishers Inc.

Höppner, F., Klawonn, F., Kruse, R., and *Runkler, T.,* 2000: Fuzzy Cluster Analysis*.* John Wiley&Sons, Chichester.

Işik, M., and *Çamurcu, A.Y.,* 2007: Applied Performance Determination Of K-Means, K-Medoids And Fuzzy C-Means Algorithms. *İstanbul Commerce Univ. J. Sci. 6*(11), 31–45.

Iyigun, C., Türkeş, M., Batmaz, İ., Yozgatligil, C., Purutçuoğlu, V., Kartal Koç, E., and *Öztürk, M.Z.,* 2013: Clustering current climate regions of Turkey by using a multivariate statistical method. *Theor. Appl. Climatol. 114,* 95–106. https://doi.org/10.1007/s00704-012-0823-7

Kalkstein, L.S., Tan, G., and *Skindlov, J.A.,* 1987: An evaluation of three clustering procedures for use in synoptic climatological classification. *J. Climate. Appl. Meteor. 26,* 717–730. https://doi.org/10.1175/1520-0450(1987)026<0717:AEOTCP>2.0.CO;2

Karahan, H., 2011: Determination of Rainfall Intensity-Duration-Frequency Relationships Using Differential Evolution Algorithm. 108Y299, Research Project, The Scientific and Technological Research Council of Turkey (TUBITAK), Final Report, Ankara, Türkiye.

Karahan, H., 2019: Determination of homogeneous sub-regions by using intensity-duration-frequency relationships and cluster analysis: an application for the Aegean Region. *Pamukkale University J. Eng. Sci. 25*(8), 998–1013. https://doi.org/10.5505/pajes.2019.09365

Kır, G., 2021: Evaluation of the meteorological data of the Black Sea Regıon using clustering analysis methods. MSc Thesis, Ondokuz May University, Samsun, Turkey.

Kite, G., 1991: Looking for Evidence of Climatic Change in Hydrometeorological Time Series. Western Snow Conference. Washington, Alaska.

Kulkarni, A. and *Kripalani, R.,* 1998: Rainfall patterns over India: Classification with fuzzy c-means method. *Theor. Appl. Climatol. 59,* 137–146. https://doi.org/10.1007/s007040050019

Lim, Y.H., and *Voeller, D.L.,* 2009: Regional flood estimations in red river using L-moment-based index-flood and bulletin 17B procedures. *J. Hydrol. Eng. 14,* 1002–1016. https://doi.org/10.1061/(ASCE)HE.1943-5584.0000102

Lin, G.F., and *Chen, L.H.,* 2006: Identification of homogeneous regions for regional frequency analysis using the self-organizing map. *J. Hydrol. 324,* 1–9. https://doi.org/10.1016/j.jhydrol.2005.09.009

Macqueen, J.B., 1967: Some Methods for Classification and Analysis of Multivariate Observations. Proc. Symp. Math. Statist. and Probability (5th), 281–297.

Moertini, V.S., 2002: Introduction to Five Clustering Algorithms. *Integral 7(2),* 87-96.

Özkoca, T., 2015: Trend analysis of hydrometeorological parameters at middle blacksea region coast band. MSc Thesis, Ondokuz May University, Samsun, Turkey.

Öztürk, M.Z., Çetinkaya, G., and *Aydın, S.,* 2017: Climate Types of Turkey According to Köppen-Geiger Climate Classification. *J. Geography 35,* 17–27. https://doi.org/10.26650/JGEOG295515

Pal, N.R., and *Bezdek, J.C.,* 1995: On Cluster validity fort he fuzzy c-means model. *IEEE Transactions On Fuzzy Syst. 3,* 370–379. https://doi.org/10.1109/91.413225

Pang-Ning Tan, P.N., Steinbach, M., and *Kumar, V.,* 2006: Introduction to Data Mining. Addison Wesley.

Rau, P., Bourrel, L., Labat, D., Melo, P., Dewitte, B., Frappart, F., Lavado, W., and *Felipe, O.,* 2017: Regionalization of rainfall over the Peruvian Pacific slope and coast. *Int. J. Climatol. 37*(1), 143–158. https://doi.org/10.1002/joc.4693

Rousseuw, P.J., 1987: Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. *J. Comput. Appl.Math. 20,* 53–65. https://doi.org/10.1016/0377-0427(87)90125-7

Soltani, S., and *Modarres, R.,* 2006: Classification of spatio temporal pattern of rainfall in Iran using a hierarchical and divisive cluster analysis. *J. Spatial Hydrol. 6(2),* 1-12.

Sönmez, İ., and *Kömüşcü, A.,* 2008: Redefinition rainfall regions using k-means clustering methodology and changes of sub period. *Iklim Değişikliği ve Çevre 1*, 38–49.

Şahin, S., 2009: *Applying artificial neural networks on determining climate zones and comparison with the Ward?s method*. PhD Thesis, Istanbul Technical University, Istanbul, Turkey.

Şahin, S., and *Cığızoğlu, H.K.,* 2012: The sub-climate regions and the sub-precipitation regime regions in Turkey. *J. Hydrol. 450–451,* 180–189. https://doi.org/10.1016/j.jhydrol.2012.04.062

Türkeş, M., 1996: Spatial and temporal analysis of annual rainfall variations in Turkey. *Int. J. Climatol. 16,* 1057–1076. https://doi.org/10.1002/(SICI)1097-0088(199609)16:9<1057::AIDJOC75>3.0.CO;2-D

Türkeş, M., 2010: Küresel iklim değişikliği: Başlıca Nedenleri, gözlenen ve öngörülen değişiklikler ve etkileri. Uluslararası Katılımlı 1. Meteoroloji Sempozyumu, Devlet Meteoroloji İşleri Genel Müdürlüğü, Ankara. (In Türkish)

Usta, G., 2023: Statistical Analysis of Disasters in the World (1900-2022). *Gümüşhane University Journal of Social Sciences Institute 14*(1), 172–186.

Ünal Y., Kındap, T., and *Karaca, M.,* 2003: Redefining the climate zones of Turkey using cluster analysis. *Int. J. Climatol. 23,* 1045–1055. https://doi.org/10.1002/joc.910

Xu, R., and *Wunsch, II D.,* 2005: Survey of Clustering Algorithms. *IEEE Transactions On Neural Networks 16(3),* 645-678. https://doi.org/10.1109/TNN.2005.845141

Zeybekoğlu, U., and *Ülke Keskin, A.,* 2020: Defining rainfall intensity clusters in Turkey by using the fuzzy c-means algorithm. *Geofizika 37(2),* 181–195. https://doi.org/10.1016/j.ins.2007.10.004

Zhang, Y., Wang, W., Zhang, X., and *Li, Y.,* 2008: A cluster validity index for fuzzy clustering. *Information Sciences 178,* 1205–1218. https://doi.org/10.1016/j.ins.2007.10.004