GEOFIZIKA VOL. 37 2020

DOI: https://doi.org/10.15233/gfz.2020.37.8
Short communication

Defining rainfall intensity clusters in Turkey by using the fuzzy c-means algorithm

Utku Zeybekoğlu[1] *and Asli Ülke Keskin*[2]

[1] Boyabat Vocational School of Higher Education, Construction Department, Sinop University, Sinop, Turkey

[2] Ondokuz Mayis University, Engineering Faculty, Civil Engineering Department, Samsun, Turkey

Received 27 March 2020, in final form 2 September 2020

Turkey has seven traditionally accepted climatic zones that are defined primarily by maritime and topographic influences. Across these zones, the annual amount of rainfall, including its intensity and its seasonal distribution, vary considerably. These variations, which impact on both urban and rural communities, including the occurrence of water shortages and flash flooding events, are increasing in both frequency and magnitude due to global warming and climate change. Several types of climate occur in Turkey where climate zones have been defined with various methodologies. To better understand rainfall intensity patterns across Turkey, this study used the Fuzzy C-Means (FCM) algorithm to define their spatial distribution. In the first stage, the annual maximum rainfall intensity records for periods ranging from 30 to 78 years were obtained from 95 stations operated by the Turkish State Meteorological Service, and the longitude, latitude and altitude data for the stations were compiled for cluster analysis. Secondly, all rainfall intensities and geographical values were normalized, and in the third stage, the FCM algorithm was applied. The comparison of annual maximum rainfall intensities revealed five clusters. Four clusters were identified as discrete zones and one was identified as a transitional zone. Weather stations located in different geographical regions sometimes fell into the same clusters. In other words, rainfall events of similar intensity can occur in different climatic zones. This study, which brought a different perspective to clustering studies, showed that rainfall intensity values can be successfully analyzed at a national scale with the FCM technique.

Keywords: rainfall intensities, fuzzy c-means (FCM), cluster analysis, Turkey, climate

1. Introduction

Climate, which is the dominant weather conditions occurring over a very long period in a large region, including extreme weather events, also determines the character and vegetation of a region (Dönmez, 1984; Gürkan et al., 2016).

Topography, the distance to large water bodies or seas, altitude, air masses, precipitation and temperature, are the principal factors used to differentiate climate zones. The concept of climate classification is widely employed in climate and climate change research, geography, hydrology, the history of civilization, agriculture, ecology and education. The diversity of climate regimes makes climate classification necessary. The systematic classification of climate, which pools or separates areas surrounding individual weather stations on the basis of the characteristics of their data, is one of the topics most studied by climatologists (Erinç, 1996; Erlat, 2014; Öztürk et al., 2017), with large climatic zones created by merging similar types (Dönmez, 1984). In the various methods developed for climate classification, precipitation and temperature are the main parameters used to describe a particular climate type.

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). The climate classification system developed by Köppen (1918) was the first classification system developed and is still the most widely used. In the last century, several methodologies were developed, Köppen (1918, 1936), De Martonne (1942), Thornthwaite (1948), Köppen and Geiger (1954), Centroid Methods, Average Linkage Method (Sokal and Michener, 1958), K-Nearest Neighbours algorithm (KNN) (MacQuinn, 1967; Anderberg, 1973), Fuzzy C-Means (FCM) algorithm (Dunn, 1974; Bezdek, 1981), Agglomerative clustering (Murtagh, 1983), Self-organizing Feature Maps (SOM) methods (Kohonen, 1990) and Expectation Maximization (EM) algorithm (McLachlan and Krishnan, 1997), and are widely applied to identify and describe climate/rainfall zones. As can be seen in this list of methodologies, climate science is continually evolving to meet the needs of local, regional and even continental communities, which is becoming increasing important as the effects of climate change increase the variability of global weather patterns.

Turkey, which straddles continental Europe and Asia, is located between 36º and 42º north latitude and 26º and 45º east longitude and contains both moderate temperate and subtropical climatic zones. In addition, there are various sub-climatic types because the country is surrounded on three sides by the sea, therefore, it is strongly affected by maritime influences, and it also has a highly variable topography that includes high mountain range systems. Due to this diversity, there are substantial variations in local and regional climate regimes, including precipitation, and the need to identify zones with similar characteristics has emerged. Therefore, many studies have focused on the determination of Turkey's precipitation and climate zones. Turkey's traditionally accepted seven climatic zones are the Aegean, Black Sea, Central Anatolia, Eastern Anatolia, Marmara, Mediterranean and South-eastern Anatolia Regions (Fig. 1). These zones were defined by Erinç (1984) on the basis of similarities and differences in climatic data and topographic features and are still accepted as valid by most climatologists in Turkey.

Erinç (1949) used the Thornthwaite method in the first study on the climate of Turkey. In that study, four main climate types were described. They were the

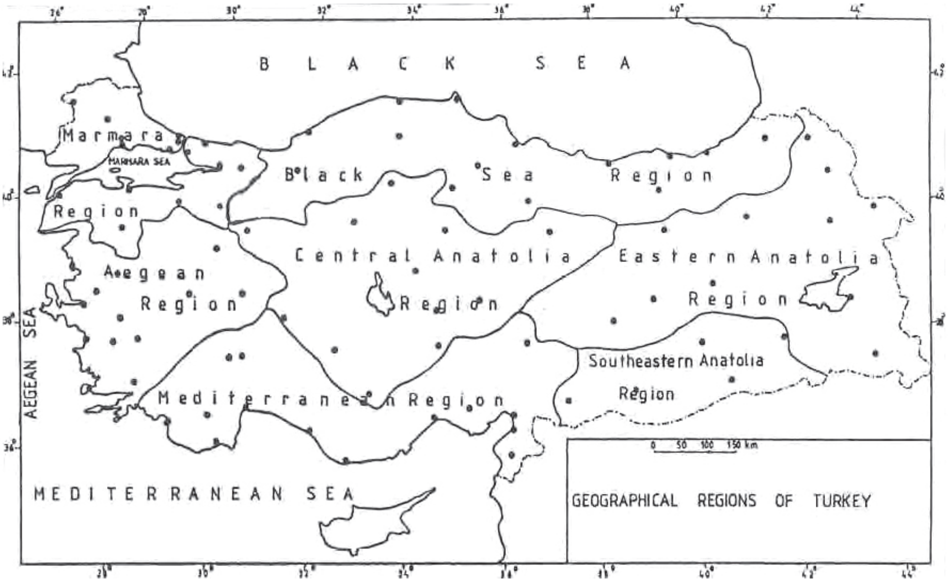

Figure 1. Geographical and climatic zones of Turkey (Erinç, 1984; Ünal et al., 2003).

Mediterranean type with hot, dry summers and mild, wet winters; the Pontic type with warm summers, mild winters and sufficient precipitation in all seasons; the Sub-Continental type with fairly warm summers but very cold winters and sufficient precipitation in all seasons; and the Semiarid type with cold winters and hot, dry summers. However, the same author emphasized the necessity of a more detailed investigation into Turkey's climate zones.

Turkes (1996) examined the precipitation records of 91 weather stations in Turkey and defined seven regions with distinct rainfall regimes, namely the Black Sea, Continental Central Anatolia, Continental Eastern Anatolia, Continental Mediterranean, Marmara Transition, Mediterranean and Mediterranean to Central Anatolia Transition. Turkes (1996) also stated that a similar precipitation regime was observed in the Mediterranean and Aegean geographical regions. In addition, the same author defined the region between the Central Anatolian and Aegean Regions as the Mediterranean Transition region, and the Marmara region was defined as a transition region between the Black Sea and Mediterranean precipitation regions.

Ünal et al. (2003) applied cluster analysis to the maximum and minimum temperatures, monthly mean temperature and monthly precipitation totals from 113 weather stations in Turkey to define homogeneous climate zones. Several hierarchical clustering procedures, namely single linkage, complete linkage, average distance within clusters, average distance between clusters and Ward's method were applied, with the last mentioned generating the best results. The same authors identified seven climate zones when rainfall and temperature values were combined, and for temperature values alone. In contrast, the use of only precipitation records generated six climate zones.

Evrendilek and Berberoğlu (2008) investigated the spatial distribution of bioclimatic zones by using 12 climatic variables, 11 bioclimatic indexes and four location descriptors from 272 meteorological stations with discriminant analysis (DA), hierarchical and non-hierarchical cluster analysis (CA), principal components analysis (PCA) and multiple linear regression (MLR) modeling. The analyses allocated the meteorological stations to heterogeneous clusters that equated to seven climatic zones. Sönmez and Kömüşcü (2008) defined six rainfall zones in Turkey by employing the K-Means clustering algorithm and the total monthly rainfall records from 148 meteorological stations.

In another study, Sariş et al. (2010) investigated the precipitation patterns across Turkey by using the multivariate methodology to analyze all of the monthly precipitation data from 107 stations. They identified two distinct coastal precipitation zones, two transitional zones and three inland zones. In a later study, Sönmez and Kömüşcü (2011) examined Turkey's rainfall zones with the 'k means' methodology. Six rainfall clusters were defined for the monthly rainfall records from 148 stations covering the period 1977 to 2006. Their results characterized the Aegean–Marmara and the Eastern Anatolia–Central Anatolia geographic regions as a single rainfall cluster, in contrast to the conventionally understood geographical regions.

In the same year, Türkes and Tatli (2011) generated eight clusters of precipitation from data from 96 stations in Turkey through the use of the spectral clustering technique. The clusters represented seven zones, namely the Black Sea, Continental eastern and south-eastern Anatolia, Eastern Continental Central Anatolia, Mediterranean, North-west Turkey, Southern Aegean and Western Mediterranean, and Western Continental Central Anatolia. In the following year, Dikbas et al. (2012) classified a Turkish precipitation series and identified six homogeneous groups by employing the fuzzy cluster method. In addition, they checked the homogeneity status of groups with the 'l-moments-based' regional homogeneity test. Their testing demonstrated that the fuzzy cluster method is useful for the classification of precipitation series and for identifying hydrologically homogenous regions.

In the same year, Firat et al. (2012) applied the k means and Ward clustering method to analyze the annual precipitation records and the longitude, latitude and altitude data of 88 stations operated by the Turkish State Meteorological Service (TSMS) and reported 7 distinct clusters. They also did regional homogeneity testing for the clusters and found that one of the clusters generated with k means and two of the clusters determined with the Ward's method were not homogenous. Iyigun et al. (2013) applied the Ward method to air temperature, precipitation total and relative humidity series from 244 meteorological stations across Turkey and reported that 14 clusters represented the climate of Turkey more realistically.

Three years later, Yilmaz and Cicek (2016) applied the Thorntwaite climate classification system to monthly average precipitation and temperature data for Turkey and generated eight different precipitation effectiveness index classes, eight different temperature effect index classes, six different drought and moistness index classes and eight different evaporation index classes.

In the following year, Özturk et al. (2017) generated Köppen-Geiger climate zones for Turkey by analyzing the data collected from 512 meteorological stations. Under their classification system, the largest area had a temperate climate, the smallest area had an arid climate, and the highlands of the Central Taurus and Eastern Anatolia regions had a continental climate.

Building on the body of earlier work in Turkey, the aim of this study was to identify the spatial distribution of rainfall intensity in Turkey based on rainfall intensities from 95 meteorological stations across the country and three sets of location data (latitude, longitude, altitude). For that purpose, the fuzzy c-means (FCM) method was used to analyze the data sets. To the knowledge of the authors, the use of rainfall intensity to better understand weather patterns differs from all earlier research in Turkey.

2. Materials and methods

2.1. Materials

The dataset used in this study included annual maximum rainfall intensity (mm/min) data recorded at 95 stations by the Turkish State Meteorological Service (TSMS) during the period 1938–2015 that ranged from 30 to 78 years. The

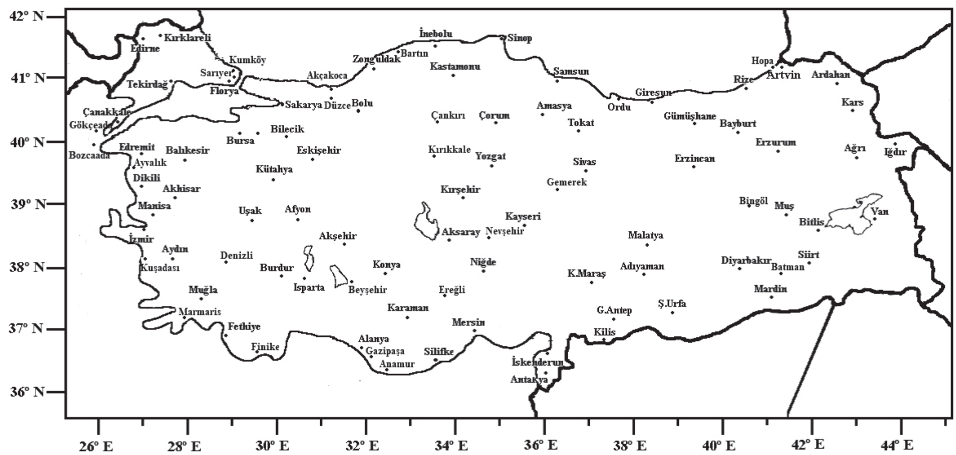

Figure 2. Geographical distribution of meteorological stations used to determine rainfall intensity patterns in Turkey.

Table 1. List of meteorological stations and geographical details used to determine rainfall intensity patterns in Turkey.

Name of Station	Period (to 2015)	Longitude (ºN)	Latitude (ºE)	Altitude (m)	Name of Station	Period (to 2015)	Longitude (ºN)	Latitude (ºE)	Altitude (m)
Adiyaman	1965	37.7553	38.2775	672	Gumushane	1966	40.4598	39.4653	1,216
Afyonkara-hisar	1957	38.738	30.5604	1,034	Hopa	1965	41.4065	41.433	33
Agri	1967	39.7253	43.0522	1,646	Igdir	1966	39.9227	44.0523	856
Akcakoca	1968	41.0895	31.1374	10	Isparta	1957	37.7848	30.5679	997
Akhisar	1965	38.9118	27.8233	92	Inebolu	1959	41.9789	33.7636	64
Aksaray	1965	38.3705	33.9987	970	Iskenderun	1965	36.5924	36.1582	4
Aksehir	1964	38.3688	31.4297	1,002	Izmir	1938	38.3949	27.0819	29
Alanya	1964	36.5507	31.9803	6	Kahraman-maras	1966	37.576	36.915	572
Amasya	1965	40.6668	35.8353	40	Karaman	1965	37.1932	33.2202	1,018
Anamur	1965	36.0686	32.8649	2	Kars	1965	40.6042	43.1073	1,777
Antakya	1957	36.2048	36.1513	104	Kastamonu	1948	41.371	33.7756	800
Ardahan	1967	41.1061	42.7055	1,827	Kayseri	1950	38.687	35.5	1,094
Artvin	1965	41.1752	41.8187	613	Kirikkale	1967	39.8433	33.5181	751
Aydin	1959	37.8402	27.8379	56	Kirklareli	1966	41.7382	27.2178	232
Ayvalik	1967	39.3113	26.6861	4	Kirsehir	1942	39.1639	34.1561	1,007
Balikesir	1957	39.6326	27.9201	102	Kilis	1966	36.7085	37.1123	640
Bartin	1966	41.6248	32.3569	33	Kocaeli	1945	40.7663	29.9173	74
Batman	1969	37.8636	41.1562	610	Konya	1950	37.8687	32.4713	1029
Bayburt	1966	40.2547	40.2207	1,584	Kumkoy	1965	41.2505	29.0384	38
Beysehir	1965	37.6777	31.7463	1,141	Kusadasi	1966	37.8597	27.2652	25
Bilecik	1960	40.1414	29.9772	539	Kutahya	1941	39.4171	29.9891	969
Bingol	1966	38.8847	40.5007	1,139	Malatya	1958	38.3367	38.2173	950
Bitlis	1966	38.475	42.1625	1,785	Manisa	1958	38.6153	27.4049	71
Bolu	1949	40.7329	31.6022	7,43	Mardin	1966	37.3103	40.7284	1,040
Bozcaada	1970	39.8326	26.0728	30	Marmaris	1966	36.8395	28.2452	16
Burdur	1964	37.722	30.294	957	Mersin	1958	36.7808	34.6031	7
Bursa	1951	40.2308	29.0133	100	Mugla	1944	37.2095	28.3668	646
Canakkale	1958	40.141	26.3993	6	Mus	1966	38.7509	41.5023	1,322
Cankiri	1959	40.6082	33.6102	755	Nevsehir	1965	38.6163	34.7025	1,260
Corum	1958	40.5461	34.9362	776	Nigde	1959	37.9587	34.6795	1,211
Denizli	1959	37.762	29.0921	425	Ordu	1965	40.9838	37.8858	5
Dikili	1959	39.0737	26.888	3	Rize	1940	41.04	40.5013	3
Diyarbakir	1940	37.9094	40.2133	680	Sakarya	1962	40.7676	30.3934	30
Duzce	1965	40.8437	31.1488	146	Samsun	1957	41.3435	36.2553	4
Edirne	1949	41.6767	26.5508	51	Sariyer	1955	41.1464	29.0502	59
Edremit	1965	39.5592	27.0253	19	Siirt	1959	37.9319	41.9354	895
Eregli	1970	37.5255	34.0485	1,046	Silifke	1964	36.3824	33.9373	10
Erzincan	1957	39.7523	39.4868	1,216	Sinop	1965	42.0299	35.1545	32
Erzurum	1956	39.9058	41.2544	1,860	Sivas	1958	39.7437	37.002	1,294
Eskisehir	1940	39.7656	30.5502	801	Sanlıurfa	1959	37.1608	38.7863	550
Fethiye	1960	36.6266	29.1238	3	Tekirdag	1963	40.9585	27.4965	4
Finike	1966	36.3024	30.1458	2	Tokat	1966	40.3312	36.5577	611
Florya	1938	40.9758	28.7865	37	Usak	1941	38.6712	29.404	919
Gaziantep	1957	37.0585	37.351	854	Van	1956	38.4693	43.346	1,675
Gazipasa	1983	36.2715	32.3045	21	Yenisehir	1986	40.2552	29.5624	238
Gemerek	1966	39.185	36.0805	1,182	Yozgat	1960	39.8243	34.8159	1,301
Giresun	1966	40.9227	38.3878	38	Zonguldak	1945	41.4492	31.7779	135
Gokceada	1970	40.191	25.9075	79					

list of stations is given in Tab. 1, together with their longitude, latitude and altitude. The geographical distribution of stations is shown in Fig. 2.

Cluster analysis was applied to the annual maximum rainfall intensities (mm/min) for 14 standard durations (5 to 1440 min), and latitude (º), longitude (º) and altitude (m) data from the 95 stations. The data were normalized with the appropriate transformation functions (Eqs. 1–3) because variables with different units can adversely influence clustering results (Lin and Chen 2006; Cannarozzo et al., 2009; Lim and Voeller, 2009; Dikbas et al., 2012; Firat et al., 2012):

$$I_{nti} = \frac{I_{ti} - I_{tmin}}{I_{tmax} - I_{tmin}}, \quad (1)$$

$$X_{ni} = \frac{X_i - X_{min}}{X_{max} - X_{min}}, \quad (2)$$

$$Z_{ni} = \frac{Z_i}{Z_{max}}. \quad (3)$$

In the above equations, I_{ti} is the rainfall intensity of duration t at station i; I_{nti} is the normalized rainfall intensity of duration t at the station i; I_{tmax} is the maximum rainfall intensity of duration t; I_{tmin} is the minimum rainfall intensity of duration t, X_i is the latitude or longitude of the station, i; X_{ni} is the normalized latitude or longitude of the station i; X_{max} is the maximum latitude or longitude; X_{min} is the minimum latitude or longitude; Z_i is the altitude of the station, i; Z_{ni} is the normalized altitude of station i, and Z_{max} is the maximum altitude of any station.

2.2. Fuzzy c-means clustering algorithm

The fuzzy c-means (FCM) clustering algorithm was proposed by Dunn (1974) and developed and extended by Bezdek (1981). The algorithm is based on the Fuzzy Logic method introduced by Zadeh (1965) (Bezdek et al., 1984; Kulkarni and Kripalani, 1998). Fuzzy c-means is a clustering method that allows each data point to belong to more than one cluster and with varying degrees of membership. In this method, clustering is performed by minimizing a defined objective function:

$$J_m = \sum_{i=1}^{n} \sum_{j=1}^{k} u_{ij}^m \left\| x_i - c_j \right\|^2, 1 \le m \le \infty, \quad (4)$$

where m is a weighting component controlling the degree of fuzzification, n is number of data sets to be clustered, k is the number of clusters determined by researchers, u_{ij} is the degree of membership of x_i in the cluster j, x_i is the i^{th} of d-dimensional measured data, c_j is the d-dimension center of the cluster, and ‖*‖ is any norm expressing the similarity between any measured data and the center of the cluster. Fuzzy partitioning is carried out through an iterative optimization of the objective function shown above, with the updating of membership, u_{ij}, and the cluster centers, c_j, with the equation:

$$u_{ij} = \left[\sum_{k=1}^{c} \left(\frac{\left\| x_i - c_j \right\|}{\left\| x_i - v_k \right\|} \right)^{\frac{2}{m-1}} \right]^{-1}, \quad c_j = \frac{\sum_{i=1}^{n} u_{ij}^{m} x_i}{\sum_{i=1}^{n} u_{ij}^{m}}. \tag{5}$$

This iteration process ends when the $max\left\{ \left| u_{ij}^{(l+1)} - u_{ij}^{(l)} \right| \right\} < 7$, where 7 is a termination criterion between 0 and 1, and l is the number of iteration steps. This procedure converges J_m to a local minimum or a saddle point (Zhang et al., 2007).

3. Results and discussion

In the present study, the annual maximum rainfall intensity series and 3 sets of geographical location data for 95 Turkish weather stations were analyzed with the use of the FCM algorithm and MATLAB2016a software. In cluster analysis, the most appropriate cluster number is decided by trial and error method (Karahan, 2019). The number of clusters may differ according to the methods used and the researcher's approach. A consensus has not yet been reached among researchers in determining the most appropriate cluster number (Zhang et al., 2008; Karahan, 2019).

The cluster analysis process was repeated for various cluster numbers, and the most suitable cluster number was determined as five, taking into account the geographical and climatic characteristics. The geographical distribution of the stations, which was defined by five clusters, is shown in Fig. 3, and the stations in each cluster are listed in Tab. 2.

According to Fig. 3 and Tab. 2, cluster A (44 stations) covered all of the coasts of Turkey and all of the Marmara region, and cluster B (11 stations) covered a transitional region between the Aegean, Marmara and Black Sea Regions and the Central Anatolia Region. Surprisingly, the Igdir station, which is at low altitude with respect to other stations in the Eastern Anatolia region, was allocated to cluster B.

Cluster C (18 stations) covered the western and southern parts of the Central Anatolia region, the Lakes Region, the vicinity of the Tuz Lake, and the transition zone between the South-eastern Anatolia and the Eastern Anatolia regions. cluster D (15 stations) covered most of the stations in the Eastern Anatolia region, except Igdir, and some stations in the Central Anatolia region, and cluster E (7 stations) includes stations from South-eastern Anatolia and the Artvin station located in the Black Sea region. That means that the Artvin station, which is located inland and at high altitude with respect to the most of the

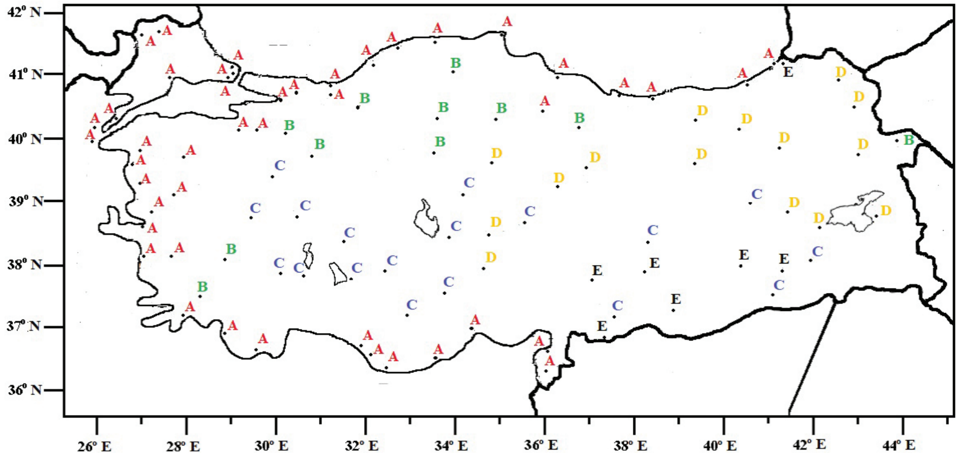

Figure 3. Geographical distribution of stations for five clusters solution.

Table 2. Clustering of weather stations in Turkey on the basis of rainfall intensity.

Cluster A		Cluster B	Cluster C	Cluster D	Cluster E
Akcakoca	Gokceada	Bilecik	Afyon	Ağri	Adiyaman
Akhisar	Hopa	Bolu	Aksaray	Ardahan	Artvin
Alanya	Izmir	Cankiri	Aksehir	Bayburt	Batman
Amasya	Inebolu	Corum	Beysehir	Bitlis	Diyarbakir
Anamur	Iskenderun	Denizli	Bingol	Erzincan	Kahramanmaras
Antakya	Kirklareli	Eskisehir	Burdur	Erzurum	Kilis
Aydın	Kocaeli	Igdir	Eregli	Gemerek	Sanliurfa
Ayvalik	Kumkoy	Kastamonu	Gaziantep	Gumushane	
Balikesir	Kusadasi	Kirikkale	Isparta	Kars	
Bartin	Manisa	Mugla	Karaman	Mus	
Bozcaada	Marmaris	Tokat	Kayseri	Nevsehir	
Bursa	Mersin		Kirsehir	Nigde	
Canakkale	Ordu		Konya	Sivas	
Dikili	Rize		Kutahya	Van	
Duzce	Sakarya		Malatya	Yozgat	
Edirne	Samsun		Mardin		
Edremit	Sariyer		Siirt		
Fethiye	Silifke		Usak		
Finike	Sinop				
Florya	Tekirdag				
Gazipasa	Yenisehir				
Giresun	Zonguldak				

other stations in the Black Sea region, is subject to different weather conditions to the other stations in the region.

Table 3 shows that cluster A had the lowest average elevation and cluster D, which includes stations in the Eastern Anatolia region which has mountain-

Table 3. Number of meteorological stations and the minimum, maximum and average elevation for each rainfall intensity cluster in Turkey.

	No. of stations	Altitude (m)		
		Min	Max	Average
Cluster A	44	2	238	45.4
Cluster B	11	425	856	695.5
Cluster C	18	854	1,141	1,003.4
Cluster D	15	1,182	1,860	1,477.1
Cluster E	7	550	680	619.6

ous topography, had the highest average elevation. Anamur and Finike stations were at the lowest elevations in cluster A and Yenisehir station at the highest elevation. In cluster B, Denizli and Igdir were at the lowest and highest altitudes, respectively. In clusters C, D and E, Beysehir, Erzurum and Diyarbakir, respectively, were the stations at the highest altitude, and Gaziantep, Gemerek and Sanliurfa, respectively, were at the lowest altitude.

The minimum (I_{min}), maximum (I_{max}), average (I_{ave}) and standard deviation (I_{sd}) of the rainfall intensity values for the five clusters are presented in Tabs. 4–6.

In Tabs. 4–6, it can be seen that cluster A had the highest rainfall intensity values. This indicates that the more intense rainfall events occur in coastal areas. In terms of rainfall intensity, cluster A was followed by cluster B and then clusters C, E and D in succession. Based on the range of precipitation intensity and altitude values included in cluster B, and its location between the coastal cluster A and cluster C in the interior, cluster B was determined toed be a transitional cluster.

Table 4. Minimums, maximums, averages and standard deviations for rainfall intensities of clusters A and B in Turkey.

Duration	Cluster A Rainfall intensities (I; mm/min)				Cluster B Rainfall intensities (I; mm/min)			
	I_{min}	I_{max}	I_{ave}	I_{sd}	I_{min}	I_{max}	I_{ave}	I_{sd}
5’	0.2000	10.1000	1.6006	0.6793	0.1000	4.8000	1.2706	0.6520
10’	0.1100	6.0600	1.1882	0.5237	0.0800	3.1000	0.9185	0.4816
15’	0.1467	4.7133	0.9942	0.4560	0.0800	2.8067	0.7521	0.3983
30’	0.0800	3.0300	0.6953	0.3553	0.0600	1.7633	0.4908	0.2806
60’	0.0583	2.0833	0.4553	0.2552	0.0367	1.1650	0.2930	0.1748
120’	0.0500	1.4367	0.2814	0.1648	0.0392	1.0167	0.1706	0.1017
180’	0.0339	1.2828	0.2094	0.1250	0.0300	0.7594	0.1234	0.0729
240’	0.0254	1.0667	0.1705	0.1016	0.0229	0.6192	0.0977	0.0577
300’	0.0203	0.8600	0.1454	0.0862	0.0200	0.5023	0.0817	0.0479
360’	0.0189	0.7544	0.1279	0.0754	0.0175	0.4208	0.0705	0.0410
480’	0.0152	0.5927	0.1034	0.0611	0.0131	0.3398	0.0561	0.0326
720’	0.0101	0.4408	0.0765	0.0454	0.0087	0.2300	0.0408	0.0238
1080’	0.0068	0.4304	0.0566	0.0340	0.0058	0.1565	0.0296	0.0176
1440’	0.0101	0.3238	0.0500	0.0272	0.0065	0.1185	0.0278	0.0170

Table 5. Minimums, maximums, averages and standard deviations for rainfall intensities of clusters C and D in Turkey.

Duration	Cluster C Rainfall intensities (I; mm/min)				Cluster D Rainfall intensities (I; mm/min)			
	I_{min}	I_{max}	I_{ave}	I_{sd}	I_{min}	I_{max}	I_{ave}	I_{sd}
5’	0.1400	5.4800	1.0401	0.5762	0.1200	3.7200	0.9802	0.5143
10’	0.1000	3.6700	0.7551	0.4066	0.100	2.3900	0.7057	0.3645
15’	0.1000	3.0000	0.6143	0.3324	0.0867	2.0733	0.5688	0.3019
30’	0.0600	1.9967	0.4007	0.2213	0.0567	1.5567	0.3662	0.2013
60’	0.0533	1.2300	0.2440	0.1310	0.0500	0.9633	0.2205	0.1210
120’	0.0308	0.6358	0.1457	0.0718	0.0317	0.5450	0.1300	0.0641
180’	0.0228	0.4239	0.1065	0.0493	0.0267	0.3933	0.0944	0.0434
240’	0.0171	0.3179	0.0856	0.0390	0.0204	0.3025	0.0753	0.0328
300’	0.0137	0.2543	0.0719	0.0324	0.0183	0.2443	0.0634	0.0268
360’	0.0117	0.2119	0.0624	0.0279	0.0153	0.2094	0.0552	0.0227
480’	0.0088	0.1698	0.0497	0.0223	0.0115	0.1573	0.0440	0.0176
720’	0.0058	0.1367	0.0360	0.0166	0.0079	0.1050	0.0322	0.0132
1080’	0.0039	0.1131	0.0263	0.0127	0.0053	0.0725	0.0238	0.0105
1440’	0.0077	0.1124	0.0260	0.0117	0.0075	0.0849	0.0235	0.0103

Table 6. Minimums, maximums, averages and standard deviations for rainfall intensities of cluster E in Turkey.

Duration	Cluster E Rainfall intensities (I; mm/min)			
	I_{min}	I_{max}	I_{ave}	I_{sd}
5’	0.1600	3.8200	1.0149	0.5385
10’	0.1300	3.5600	0.7415	0.3925
15’	0.1333	2.7400	0.6038	0.3199
30’	0.0833	1.5367	0.3898	0.2038
60’	0.0500	0.9500	0.2372	0.1193
120’	0.0408	0.4950	0.1428	0.0653
180’	0.0300	0.3339	0.1060	0.0478
240’	0.0242	0.2504	0.0867	0.0382
300’	0.0197	0.2263	0.0740	0.0330
360’	0.0172	0.2025	0.0654	0.0290
480’	0.0133	0.1735	0.0532	0.0239
720’	0.0089	0.1607	0.0401	0.0187
1080’	0.0059	0.1084	0.0296	0.0141
1440’	0.0065	0.0835	0.0311	0.0115

When the rainfall distribution, rainfall intensity and altitude values for the clusters are viewed collectively, rainfall intensity decreased from the coastal regions to the interior and from west to east. The reduction in rainfall intensity from the coastal areas to the interior probably reflects a strong maritime influence manifesting as higher humidity and rainfall intensity in coastal areas. In addition, the general reduction in rainfall intensity from west to east probably indicates that the rain bearing systems move in that direction and that their moisture content is gradually depleted.

4. Conclusions

In this study, as distinct from earlier studies, annual maximum rainfall intensity values and location parameters were utilized to better understand weather patterns in Turkey. The use of a non-hierarchical clustering method known as the fuzzy c-means (FCM) algorithm produced five clusters from the data sets for 95 meteorological stations operated by the Turkish State Meteorological Service. Four clusters were identified as main rainfall zones and the other one was identified as a transitional zone. Especially the aggregation of stations near the sea in a single cluster (Cluster A) is understood to be the result of maritime influences.

Traditionally, Turkey has been divided into seven climate zones. The use of rainfall intensity over periods ranging from 5 minutes to 24 hours produced five clusters that are different from the traditionally accepted climatic zones. This result inherently represents the use of a different data set and provides a different perspective on weather and climate in Turkey.

Given that the decreasing and increasing intensity of rainfall events associated with climate change is linked to the increased probability of drought and flood events, respectively, the information generated by this study is potentially useful in the regional planning, design, construction and operation works of different sectors such as water resources, agriculture, urbanization, drainage, flood control and transportation. In terms of urbanism, it is thought to shed light on risky places in terms of natural disasters such as floods that may occur as a result of global climate change. In particular, through their incorporation in regional and local planning, the results of this study may help reduce the number of deaths and injuries and the damage to infrastructure and property caused by the flash flooding associated with the extreme rainfall events that are increasing in frequency across Turkey.

These kinds of researches, which are of great importance in determining regional differences, become a necessity for adaptation studies against climate change, which show their effects intensely.

Acknowledgements – Authors would like to thank the referees of our article for the time spent and their constructive criticism for the analysis of this article. Also, the authors specially thank Dr. Gregory T. Sullivan for helpful comments and the proofreading of earlier versions of this manuscript.

Conflict of interest statement – The authors have no conflicts of interest in relation to this study.

References

Anderberg, M. R. (1973): *Cluster analysis for applications*. Academic Press, Cambridge, 376 pp.

Bezdek, J. C. (1981): *Pattern recognition with fuzzy objective function algorithms*. New York, USA, Plenum Press, 256 pp, https://doi.org/10.1007/978-1-4757-0450-1.

Bezdek, J. C., Ehrlich, R. and Full W. (1984): FCM: The fuzzy c-means clustering algorithm, *Comput. Geosci.*, **10**, 191–203, https://doi.org/10.1016/0098-3004(84)90020-7.

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.

De Martonne, E. (1942): Nouvelle carte mondial de l'indice d'aridité, *Annales de Géographie*, **288**, 241–250, https://doi.org/10.3406/geo.1942.12050.

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

Dönmez, Y. (1984): *Umumî Klimatoloji ve İklim Çalışmaları*. Güryay Press Inc., Istanbul, 431 pp (in Turkish).

Dunn J. C. (1974): A fuzzy relative of the ISODATA process and its use in detecting compact, well-separated 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.

Erinç, S. (1984): *Climatology and its methods*. 3rd edition. Güryay Pres Inc., Istanbul, 539 pp (in Turkish).

Erinç, S. (1996): *Climatology and its methods*. 4th edition. Alfa Press Inc., Istanbul, 538 pp (in Turkish).

Erlat, E. (2014): *Dünya Iklimleri*. Ege University Faculty of Letters Publications, Izmir, 309 pp (in Turkish).

Evrendilek, F. and Berberoğlu, S. (2008): Quantifying spatial patterns of bioclimatic zones and controls in Turkey, *Theor. Appl. Climatol.*, **91**, 35–50, https://doi.org/10.1007/s00704-006-0294-9.

Fırat, M., Dikbaş, F., Koç, A. C. and Güngör, M. (2012): Classification of annual precipitations and identification of homogeneous regions using k-means method, *Teknik Dergi*, **23**, 1609–1622.

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.

Gürkan, H., Arabaci, H., Demircan, M., Eskioğlu, O., Şensoy, S. and Yazici B. (2016): Temperature and precipitation projectio ns based on GFDL-ESM2M using RCP4.5 and RCP8.5 scenarios for Turkey, *Turkish J. Geogr. Sci.*, **14**(2), 77–88, https://doi.org/10.1501/Cogbil_0000000174.

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. (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**, 998–1013, https://doi.org/10.5505/pajes.2019.09365.

Kohonen, T. (1990): The self-organizing map, *Proceedings of the IEEE*, **78**, 1464–1480, https://doi.org/10.1109/5.58325.

Köppen, W. (1918): Klassifikation der Klimate nach Temperatur, Niederschlag und Jahresablauf (Classification of climates according to temperature, precipitation and seasonal cycle), *Petermann. Geogr. Mitt.*, **64**, 193–203.

Köppen, W. (1936): *Das geographische System der Klimate (The geographic system of climates).* Borntraeger Science Publishers, Berlin, 44 pp.

Köppen, W. and Geiger, R. (1954): *Klima der Erde (Climate of the Earth)*. Wall Map (1:16 Mill.), Klett-Perthes, Gotha.

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.

McLachlan, G. and Krishnan, T. (1997): *The EM algorithm and extensions*. Wiley, New York, 997 pp.

MacQueen, J. B. (1967): Some methods for classification and analysis of multivariate observations, in: *Proceedings of the 5th Berkeley Symposium on Mathematical Statistics and Probability*, 281–297.

Murtagh, F. (1983): A survey of recent advances in hierarchical clustering algorithms, *Comput. J.*, **26**(44), 354–359, https://doi.org/10.1093/comjnl/26.4.354.

Öztürk, M. Z., Çetinkaya, G. and Aydın, S. (2017): Climate types of Turkey according to Köppen-Geiger climate classification. *Istanbul University J. Geogr.*, **35**, 17–27, https://doi.org/10.26650/JGEOG330955.

Sariş, F., Hannah, D. and Eastwood, W. (2010): Spatial variability of precipitation regimes over Turkey, *Hydrolog. Sci. J.*, **55**, 234–249, https://doi.org/10.1080/02626660903546142.

Sokal, R. R. and Michener, C. D. (1958): A statistical methods for evaluating relationships, *University of Kansas Sci. Bull.*, **38**, 1409–1448.

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 (in Turkish).

Sönmez, I. and Kömüşcü, A. U. (2011): Reclassification of rainfall regions of Turkey by k-means methodology and their temporal variability in relation to North Atlantic Oscillation (NAO), *Theor. Appl. Climatol.*, **106**, 499–510, https://doi.org/10.1007/s00704-011-0449-1.

Thornthwaite, C. W. (1948): An approach toward a rational classification of climate, *Geogr. Rev.*, **38**, 55–94, https://doi.org/10.2307/210739.

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::AID-JOC75>3.0.CO;2-D.

Türkeş, M. and Tatli, H. (2011): Use of the spectral clustering to determine coherent precipitation regions in Turkey for the period 1929– 2007, *Int. J. Climatol.,* **31**, 2055–2067, https://doi.org/10.1002/joc.2212.

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.

Yılmaz, E. and Çiçek, İ. (2016): Thornthwaite climate classification of Turkey, *J. Hum. Sci.*, **13**, 3973-3994, https://doi.org/10.14687/jhs.v13i3.3994.

Zadeh L. A. (1965): Fuzzy sets, *Inform. Control*, **8**, 338–353, https://doi.org/10.1016/S0019-9958(65)90241-X.

Zhang, S., Wang, R. S. and Zhang, X. S. (2007): Identification of overlapping community structure in complex networks using fuzzy c-means clustering, *Physica A*, **374**, 483–490, http://doi.org/10.1016/j.physa.2006.07.023.

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

SAŽETAK

Definiranje klastera intenziteta oborine u Turskoj korištenjem algoritma neizrazitih c-srednjaka

Utku Zeybekoğlu i Asli Ülke Keskin

Turska ima sedam tradicionalno prihvaćenih klimatskih zona koje su definirane prvenstveno maritimnim i topografskim utjecajima. Diljem tih zona godišnja količina oborine, uključujući njezin intenzitet i sezonsku razdiobu, znatno se razlikuje. Te varijacije, koje utječu i na urbane i na ruralne ljudske zajednice, uključujući pojavu nestašice vode i poplave, povećavaju se i u učestalosti i magnitudi zbog globalnog zatopljenja i klimatskih promjena. U Turskoj se pojavljuje nekoliko tipova klime, pri čemu su klimatske zone definirane različitim metodologijama. Kako bi bolje razumjeli obrasce intenziteta oborine diljem Turske, u ovoj studiji korišten je algoritam „neizrazitih klasterskih srednjaka" („fuzzy c-means" – FCM) s ciljem definiranja njihove prostorne razdiobe. U prvom koraku, korišteni su zapisi godišnjih maksimalnih intenziteta oborine za razdoblja u rasponu od 30 do 78 godina s 95 postaja kojima upravlja Turska državna meteorološka služba, a podaci o zemljopisnoj dužini, zemljopisnoj širini i nadmorskoj visini postaja dodani su radi analize klastera. U drugom koraku normalizirani su svi intenziteti oborine i zemljopisni podaci, a u trećem je primijenjen FCM algoritam. Usporedba godišnjih maksimalnih intenziteta oborine definirala je postoojanje pet klastera. Četiri klastera identificirana su kao diskretne zone, a jedan je identificiran kao prijelazna zona. Meteorološke postaje koje pripadaju različitim zemljopisnim područjima ponekad pripadaju istom klasteru. Drugim riječima, oborinski događaji sličnog intenziteta mogu se pojaviti u različitim klimatskim zonama. Ova studija pokazala je da se vrijednosti intenziteta oborine na nacionalnoj razini mogu uspješno analizirati FCM tehnikom, doprinoseći drukčijem pogledu na studije koje koriste analizu klastera.

Ključne riječi: intenziteti oborine, neizraziti klasterski srednjaci (FCM), analiza klastera, Turska, klima

Corresponding author's address: Utku Zeybekoğlu, Boyabat Vocational School of Higher Education, Construction Department, Sinop University, Sinop, Turkey; tel: +90 368 315 0101 (internal: 5743), fax: +90 368 315 1441; e-mail: utkuz@sinop.edu.tr (ORCID: 0000-0001-5307-8563).