

## Karadeniz Bölgesi İstasyonlarının K-Ortalamlar, Bulanık C-Ortalamlar ve Siluet İndeks Analizi Yöntemleri Kullanılarak Kümelenmesi

Gürkan KIR<sup>1</sup>, Aslı ÜLKE KESKİN<sup>1</sup>, Utku ZEYBEKOĞLU<sup>2</sup>  
ORCID: 0000-0001-8156-9431, 0000-0002-9676-8377, 0000-0001-5307-8563

### Öz

Son yıllarda küresel iklim değişikliği etkilerinden kaynaklanan afet sayılarında hissedilebilir bir artış görülmektedir. Bu kapsamda iklim değişikliği etkilerini azaltmak amacıyla ülkemizde ve dünyada çeşitli çalışmalar yapılmaktadır. İklim değişikliğinden etkilenen bölgelerin iklim parametreleri bakımından benzer sınıflara ayrılması bu bölgelerde yapılacak olan çalışmalarda benzer yöntemlerin uygulanması açısından önemlidir. Böylece iklim değişikliğinin etkilerini azaltmak amacıyla yapılacak olan çalışmalarda doğru bir stratejinin belirlenmesi sağlanacaktır. Çalışma kapsamında Meteoroloji Genel Müdürlüğüne ait Karadeniz Bölgesinde yer alan 31 istasyonun 1982-2020 periyoduna ait yıllık sıcaklık kayıtları değerlendirilmiştir. Bulanık C-Ortalamlar ve K-Ortalamlar yöntemleri kullanılarak kümeleme analiz çalışması gerçekleştirilmiştir. Çalışma sonucunda optimum küme sayısı Siluet indeksi ile tespit edilmiştir. Sıcaklık gözlemler için en uygun sınıflandırma, küme sayısı 5 seçilerek K-Ortalamlar yöntemi ile elde edilmiştir.

**Anahtar Kelimeler:** Kümeleme Analizi, Sıcaklık, Silüet Analizi, Afet, Küresel İklim Değişikliği.

## Clustering of Black Sea Region Stations Using K-Means, Fuzzy C-Means and Silhouette Index Analysis Methods

### Abstract

In recent years, there has been a noticeable increase in the number of disasters caused by the effects of global climate change. In this context, various studies are carried out in our country and in the world in order to reduce the effects of climate change. The classification of regions affected by climate change into similar classes in terms of climate parameters is important in terms of applying similar methods in studies to be carried out in these regions. Thus, a correct strategy will be determined in the studies to be carried out in order to reduce the effects of climate change. Within the scope of the study, annual temperature records of 31 stations in the Black Sea Region belonging to the Turkish State Meteorological Service for the period 1982-2020 were evaluated. Cluster analysis was carried out using the Fuzzy C-Means and K-Means methods. As a result of the study, the optimum number of clusters was determined by Silhouette index analysis. The most suitable classification for the temperature observations was obtained by the K-Means method by choosing the number of clusters as 5.

**Keywords:** Clustering Analyses, Temperature, Silhouette Analysis, Disaster, Global Climate Change

<sup>1</sup>İnşaat Müh. Bölümü, Ondokuz Mayıs Üniversitesi, Samsun, Türkiye

<sup>2</sup>İnşaat Bölümü, Sinop Üniversitesi, Sinop, Türkiye

\* İlgili yazar/Corresponding author: utkuz@sinop.edu.tr

Gönderim Tarihi / Received Date: 02.08.2023

Kabul Tarihi / Accepted Date: 26.11.2023

Bu makaleye atf yapmak için- To cite this article

Kır, G., Ülke Keskin, A., Zeybeoğlu, U. (2023). Karadeniz Bölgesi İstasyonlarının K-Ortalamlar, Bulanık C-Ortalamlar ve Siluet İndeks Analizi Yöntemleri Kullanılarak Kümelenmesi. Resilience, 325-337

## 1. Giriş

Yağış, sıcaklık ve rüzgâr gibi çeşitli meteorolojik parametrelerin ekstrem değerleri olarak tanımlanan iklim, belirli bir zaman aralığında belirli bir konum için atmosferin kolektif durumudur (Demircan vd., 2017). Bu kolektif yapıya ait parametrelerin uzun dönem ortalamalarında 1950'li yılların ortalarına kadar değişim olmadığı kabul edilmekteydi. 20. yüzyıla doğru ilerlerken sanayi alanında yaşanan hızlı gelişmeler; doğal kaynakların plansızca tüketilmesine, çevre kirliliğinin insan sayısı ile orantılı olarak artmasına ve atmosfere yoğun miktarda sera gazı salınımına neden olmuştur (Türkeş vd., 2000). Bu doğrultuda atmosferde ısı tutma özelliğine sahip sera gazı, zamanla iklim parametrelerinde değişiklikler meydana getirmeye başlamıştır. İklim parametrelerinde meydana gelen bu değişiklikler küresel iklim değişikliği olarak adlandırılmaktadır (Türkeş, 2010; Özkoca, 2015).

Küresel ölçekte gerçekleşen iklim değişikliği yerelde etkilerini frekans ve şiddeti artan taşkın, sel, kuraklık ve fırtına gibi farklı afetler şeklinde göstermektedir (Çelik vd., 2018). Son yıllarda birçok makaleye konu olan küresel iklim değişikliğinin gün geçtikçe etkisini artırması insan hayatını ekonomik ve sosyal alanda olumsuz etkilemektedir. Bu doğrultuda iklim değişikliğini anlamak ve bu kapsamda önlemler almak için gerçekleştirilen çalışmalar da giderek önem kazanmaktadır. İklim parametreleri bakımından benzerlik gösteren bölgelerin sınıflara ayrılması; iklim değişikliği ile mücadele, su kaynaklarının korunması ve arazi kullanımının planlanması gibi farklı çalışmalara katkı sağlayacağı düşünülmektedir. Literatürde iklim sınıflarının belirlenmesi üzerine yurt içinde ve yurt dışında pek çok çalışma gerçekleştirilmiştir. Bu çalışmalar incelendiğinde iklim parametresi olarak daha çok yağış ve sıcaklık parametreleri kullanılmaktadır. Erinç (1949), Thorntwaite yöntemiyle Türkiye'de bulunan 53 meteoroloji istasyonundan alınan yağış ve sıcaklık verilerini 4 farklı iklim bölgesi için sınıflandırmıştır. Bu çalışma ile Türkiye coğrafyasının yeterli veri ile bölgesel ve ayrıntılı sınıflandırılması ilk kez gerçekleştirilmiştir. Türkeş (1996), Kraus tarafından 1977 yılında önerilen Normalizasyon Prosedürü yöntemi yardımıyla Türkiye'nin yağış verilerini sınıflandırmıştır. 1930-1993 periyodundaki özelemlerin kullanıldığı çalışmada 7 farklı bölge tespit edilmiştir. Kulkarni ve Kripalani (1998), Bulanık C Ortalamalar yöntemi ile Hindistan yağış verilerinin benzerlik gösteren sınıflarını tespit etmişlerdir. 1871-1984 periyodundaki yağış verileri kullanılarak 306 meteoroloji gözlem istasyonu 4 farklı kümeye ayrılmıştır. Ünal vd. (2003), 5 farklı kümeleme yöntemi ile Türkiye'de yer alan ve periyodu 1951-1998 yılları arasında kapsayan sıcaklık ve yağış verilerinin benzerlik gösteren sınıflarını belirlemişlerdir. Tek Bağlantı, Tam Bağlantı, Merkez, Ward'ın Minimum Varyansı ve Ortalama Mesafe yöntemlerinin kullanıldığı çalışmada en etkili yöntemin Ward metodu olduğu belirtilmiştir. Soltani ve Modarres (2006), İran'da bulunan 28 adet meteoroloji istasyonuna ait yağış verilerini hiyerarşik ve hiyerarşik olmayan kümeleme yöntemleri yardımıyla benzerlik gösteren sınıflara ayırmışlardır. 8 farklı sınıfın belirlendiği çalışmada Ward metodu ve K-Ortalamlar algoritması kullanılmıştır. Sönmez ve Kömüştü (2008), Türkiye'nin yağış bölgelerini belirledikleri çalışmalarında K-Ortalamları algoritmasından yararlanmışlardır. 148 meteoroloji istasyonundan temin edilen ve 1977-2006 yılları arasında kapsayan aylık toplam yağış serilerini kullanıldığı çalışmada 6 farklı yağış bölgesi tespit edilmiştir. Şahin (2009), Türkiye'nin benzer iklim sınıflarını belirlemek amacıyla 150 meteoroloji istasyonundan temin edilen aylık ortalama sıcaklık, aylık bağıl nem ve aylık toplam yağış verilerini kullanmıştır. Ward, Kohonen Yapay Sinir Ağı ve Bulanık Yapay Sinir Ağı yöntemlerinden kullanılarak 7 farklı bölge tespit edilmiştir. Dikbaş vd. (2011), Bulanık C-Ortalamlar yöntemiyle Türkiye'deki 188 istasyona ait 1967-1998 kayıtlarını kullanarak 6 farklı yağış bölgesi belirlemişlerdir. Şahin ve Cıgızoğlu (2012), Ward metodu ve Bulanık Yapay Sinir Ağı yöntemlerini kullanarak Türkiye'nin alt iklim ve alt yağış rejimi sınıflarını belirlemişlerdir. 232 meteoroloji istasyonunun 1974-2002 periyodundaki yağış, sıcaklık ve nem verileri kullanılarak 7 yağış rejimi bölgesi ve 7 iklim bölgesi tespit edilmiştir. Fırat vd. (2012), Türkiye'de bulunan 188 yağış gözlem istasyonunda ölçülmüş olan ve periyodu 1967-1998

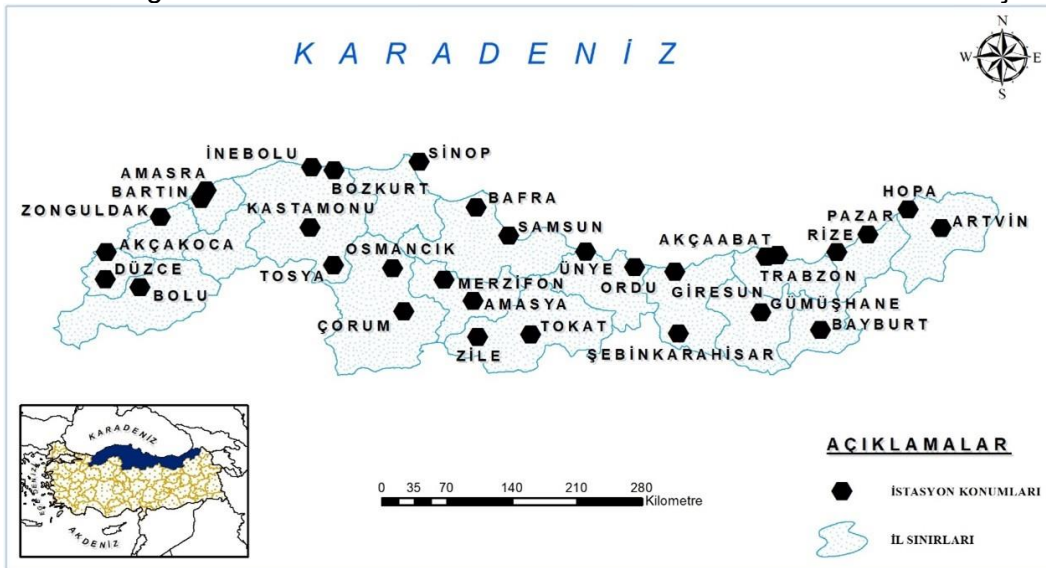
yılları arası kapsayan yıllık toplam yağışların benzer özellik gösteren sınıflarını K-Ortalamları yöntemini kullanarak 7 farklı benzer özellik gösteren bölge belirlemiştirler. İyigün vd. (2013), Ward metodu kullanarak yağış, sıcaklık ve bağıl nem verileri ile kümeleme analiz çalışması gerçekleştirmişlerdir. Türkiye’de bulunan 244 adet meteoroloji istasyonundan temin edilmiş olup periyodu 1970-2010 yılları arasını kapsamaktadır. Çalışma sonucunda 14 farklı küme tespit edilmiştir. Rau vd. (2017), Peru Pasifik yamacı ve kıyısına ait yağış verilerini benzer özellik gösteren bölgelere ayırmışlardır. Bölgesel Vektör Yöntemi ve K-Ortalamları algoritması kullanılarak 9 farklı yağış bölgesi tespit edilmiştir. Zeybekoğlu ve Ülke Keskin (2020), Bulanık C-Ortalamlar algoritmasını kullanarak yağış şiddeti serilerini gözlem istasyonlarına ait enlem, boylam ve yükseklik değerlerini de ekleyerek kümeleme analizi gerçekleştirmiştir. Türkiye’de bulunan 95 adet meteoroloji gözlem istasyonununun 5 farklı küme oluşturduğu tespit edilmiştir.

Bulanık C-Ortalamlar (BCO) ve K-Ortalamlar (KO) yöntemlerinin birlikte kullanılarak elde edilen sonuçların Siluet analizi ile değerlendirilmesine iklim çalışmalarında çok fazla rastlanmamaktadır (Kır, 2021). Bu doğrultuda bu çalışmanın amacını Karadeniz Bölgesi’ne ait sıcaklık gözlem kayıtlarının farklı kümeleme algoritmaları aracılığı ile benzer özellik gösterenlerinin kümelenmesi oluşturmaktır. BCO ve KO algoritmaları ile farklı küme sayılları için gerçekleştirilen analizlerde en uygun küme sayısı ile kümeleme algoritması Siluet indeks analizi ile belirlenmiştir.

## 2. Materyal ve Metot

### 2.1. Materyal

Bu çalışmada Karadeniz Bölgesinde, Meteoroloji Genel Müdürlüğü (MGM) tarafından işletilen 31 gözlem istasyonunda 1982-2020 (39 yıl) yılları arasında kaydedilen rüzgar hızı gözlemleri kullanılmıştır. Temin edilen verilerin istatistiki açıdan yeterli olabilmesi için en az 30 yıl kayıt uzunluğuna sahip olmasına dikkat edilmiştir (Kite, 1991). Çalışmada kullanılan gözlem istasyonları Karadeniz Bölgesindeki 17 farklı ilde yer almaktadır. İstasyonlardan 11 tanesi Batı Karadeniz’de (Düzce, Akçakoca, Bolu, Zonguldak, Bartın, Amasra, Kastamonu, İnebolu, Bozkurt, Tosya, Sinop), 10 tanesi Orta Karadeniz’de (Samsun, Bafra, Çorum, Osmancık, Amasya, Merzifon, Tokat, Zile, Ordu, Ünye), geriye kalan 10 tanesi ise Doğu Karadeniz’de yer almaktadır (Giresun, Şebinkarahisar, Trabzon, Akçaabat, Gümüşhane, Bayburt, Rize, Pazar, Artvin, Hopa). İstasyonların coğrafyadaki dağılımı Şekil 1’de gösterilmiştir. İstasyonların coğrafi konum bilgileri ile verilere ait temel istatistikleri Tablo 1 ve Tablo 2’de verilmiştir.



Şekil 1. Verisi kullanılan istasyonların coğrafyadaki dağılımı (Kır, 2021).

Tablo 1. Verisi kullanılan istasyonların coğrafi konum bilgileri.

İstasyon	Enlem (K)	Boylam(D)	Rakım (m)
Düzce	40°50'37.3"	31°08'55.7"	146
Akçakoca	41°05'22.2"	31°08'14.6"	10
Bolu	40°43'58.4"	31°36'07.9"	743
Zonguldak	41°26'57.3"	31°46'40.5"	135
Bartın	41°37'29.3"	32°21'24.8"	33
Amasra	41°45'09.4"	32°22'57.7"	73
Kastamonu	41°22'15.6"	33°46'32.2"	800
İnebolu	41°58'44.0"	33°45'49.0"	64
Bozkurt	41°57'34.9"	34°00'13.3"	167
Tosya	41°00'47.5"	34°02'12.1"	870
Çorum	40°32'46.0"	34°56'10.3"	776
Osmancık	40°58'43.3"	34°48'04.0"	419
Sinop	42°01'47.6"	35°09'16.2"	32
Amasya	40°40'00.5"	35°50'07.1"	409
Merzifon	40°52'45.5"	35°27'30.6"	754
Samsun	41°20'39.0"	36°15'23.0"	4
Bafra	41°33'05.4"	35°55'28.9"	103
Tokat	40°19'52.3"	36°33'27.7"	611
Zile	40°17'45.6"	35°53'25.8"	719
Ordu	40°59'01.7"	37°53'08.9"	5
Ünye	41°08'34.8"	37°17'34.8"	16
Giresun	40°55'21.7"	38°23'16.1"	38
Şebinkarahisar	40°17'13.9"	38°25'09.5"	1364
Gümüşhane	40°27'35.3"	39°27'55.1"	1216
Trabzon	40°59'54.6"	39°45'53.6"	25
Akçaabat	41°01'57.0"	39°33'41.4"	3
Bayburt	40°15'16.9"	40°13'14.5"	1584
Rize	41°02'24.0"	40°30'04.7"	3
Pazar	41°10'39.7"	40°53'57.5"	78
Artvin	41°10'30.7"	41°49'07.3"	613
Hopa	41°24'23.4"	41°25'58.8"	33

Tablo 2. İstasyonların sıcaklık gözlemlerine ait temel istatistikler.

İstasyon	Ortalama	S. Sapma	Min.	Mak.	Varyasyon	Çarpıklık
Düzce	13.38	0.95	11.45	15.35	0.07	0.28
Akçakoca	13.21	0.93	11.59	15.26	0.07	0.34
Bolu	10.68	0.87	8.70	12.78	0.08	0.15
Zonguldak	13.82	0.84	12.59	15.66	0.06	0.49
Bartın	12.88	0.78	11.36	14.59	0.06	0.30
Amasra	13.98	0.85	12.74	15.73	0.06	0.21
Kastamonu	9.85	0.75	8.33	11.47	0.08	0.20
İnebolu	13.50	1.04	11.98	16.89	0.08	1.19
Kastamonu/Bozkurt	13.39	0.83	12.14	15.46	0.06	0.37
Tosya	11.37	0.89	9.57	13.35	0.08	0.34
Çorum	10.79	0.98	9.00	13.17	0.09	0.63
Osmancık	13.64	0.80	12.17	15.50	0.06	0.50
Sinop	14.46	0.92	13.02	16.70	0.06	0.55
Amasya	13.63	0.95	11.78	15.85	0.07	0.45
Merzifon	11.58	0.86	9.76	13.45	0.07	0.20

Samsun Bölge	14.73	0.90	13.27	16.58	0.06	0.31
Bafra	13.96	0.84	12.47	15.75	0.06	0.20
Tokat	12.60	1.01	10.66	14.90	0.08	0.40
Zile	11.90	1.00	9.96	13.93	0.08	0.13
Ordu	14.73	0.87	13.23	16.63	0.06	0.20
Ünye	14.55	0.88	12.99	16.47	0.06	0.25
Giresun	14.83	0.91	13.31	16.85	0.06	0.47
Şebinkarahisar	9.38	1.07	6.93	12.17	0.11	0.32
Gümüşhane	9.60	1.01	7.33	12.05	0.10	0.65
Trabzon Bölge	15.01	0.91	13.55	16.95	0.06	0.39
Akçaabat	14.65	0.86	13.21	16.57	0.06	0.31
Bayburt	7.23	1.16	4.65	10.08	0.16	0.45
Rize	14.69	0.89	13.14	16.83	0.06	0.30
Rize/Pazar	13.79	0.89	12.02	15.92	0.06	0.46
Artvin	12.24	0.99	10.00	14.53	0.08	0.36
Hopa	14.77	0.98	13.15	16.96	0.07	0.41

## 2.2. Bulanık C-Ortalamlar algoritması (BCO)

Kümeleme algoritmaları arasından yaygın olarak kullanılan BCO tekniği Dunn (1973) tarafından geliştirilip, Bezdek (Bezdek 1980; Bezdek vd., 1984) tarafından iyileştirilmiştir. Kümeleme işlemi sırasında kümelere ait bulanıklık derecesini belirleyen [1, 2] arasında değişen bulanıklaştırma parametresi (m) kullanılmaktadır (Bezdek 1980; Bezdek vd., 1984; Vani vd., 2019). Algoritmanın işlem adımları:

Adım 1: Küme merkezleri rastgele oluşturulur.

$$J_{KM}(X;V)=\sum_{i=1}^c \sum_{j=1}^n D_{ij}^2 \quad (1)$$

Adım 2: Denklem (1) ile Öklid mesafesi kullanılarak bir veri noktasından kümelerin merkezlerine olan uzaklık matrisi oluşturulur.

$$V_i = \frac{\sum_{j=1}^n u_{ij}^m x_j}{\sum_{j=1}^n u_{ij}^m}; 1 \leq i \leq c \quad (2)$$

Adım 3: Denklem (3) ile bulanıklaştırma parametresi aracılığıyla üyelik matrisi hesaplanır.

$$u_{ij} = \left( \sum_{k=1}^c \left( \frac{D_{ijA}}{D_{kjA}} \right)^{\frac{2}{m-1}} \right)^{-1}; 1 \leq i \leq c, 1 \leq j \leq n \quad (3)$$

$$J_{KM}(U,\lambda;X)=\sum_{i=1}^c \sum_{t=1}^T u_{it}^m d_{it}^2 \sum_{j=1}^n D_{ij}^2 \quad (4)$$

Adım 4:  $U_{ij}$  matrisinin değerleri şuna eşit veya küçük olmalıdır: ( $U_{ij} \leq 1$ )

Adım 5: Yeni ağırlık merkezi hesaplanır.

Adım 6: Yeni merkezler oluşturularak küme merkezleri optimize edilir.

Adım 7: Veri noktaları için kümeler belirlenir.

Girdi:  $x_1$ , veri vektörü;  $V_i$ , bulanık kümelerin merkez noktaları;  $c$ , bulanık küme sayısı;  $m$ , bulanıklaştırma parametresi;  $U$ , her örneğe, bir veri örneğinden  $n$ 'inci kümeye üyelik değerini gösteren Bulanık üyelik değeri atar;  $\varepsilon$  - durdurma kriteri;  $D_{ij}$ , mesafe ölçüsü ve  $n$ , veri noktası sayısı.

Çıktı: Veri noktaları uygun kümelere atanır.

Algoritmanın Avantajları (Vani vd., 2019):

- Çakışan veri kümesi için BCO, KO'ya göre daha iyi sonuçlar verir.
- Her bir veri noktası, her küme merkezine bir üyelik değeri ile atanır, sonuç olarak, veri noktası birden fazla küme merkezine ait olabilir.

Algoritmanın Dezavantajları (Vani vd., 2019):

- BCO'da, küme sayısının önceden belirtilmesini gerektirir.
- Bulanıklaştırıcı katsayı 'm'nin daha düşük değeriyle bile daha fazla yineleme alır.

### 2.3. K-Ortalamlar algoritması (KO)

KO kümeleme algoritması (Xin vd, 2011), en basit denetimsiz ve zor kümeleme algoritmalarından biridir. Bu yöntem, belirli bir veri setini çeşitli kümelere sınıflandırmak için kullanılır (Vani vd., 2019). Algoritmanın işlem adımları:

Adım 1: Küme merkezleri rastgele merkezleri belirlenir.

Adım 2: Ağırlık merkezi ve veri noktaları arasındaki mesafeler hesaplanır.

Adım 3: k-ortalamlar, minimum Öklid mesafe ölçüsünü kullanarak verileri kümeler.

$$J_{KM}(X;V)=\sum_{i=1}^c \sum_{j=1}^n D_{ij}^2 \quad (5)$$

Adım 4: Yeni ağırlık merkezi hesaplanır.

$$V_i = \sum_{j=1}^{n_i} \frac{D_{ij}}{n_i}; 1 \leq i \leq c \quad (6)$$

Adım 5: Yeni ağırlık merkezinin eski ağırlık merkezine eşit olup olmadığı kontrol edilir.

Adım 6: Yeni ağırlık merkezi ile eski ağırlık merkezi eşitse algoritma tamamlanır, aksi durumda algoritma 2 adımdan başlanarak tekrarlanır.

Girdi: V merkez numarası; x ve y, ağırlık merkezi ile veri noktaları arasındaki uzaklık merkezi değerleri;  $x_1$  ve  $y_1$ , veri noktasının değerleri;  $x_{11}$  ve  $y_{11}$ , düzeylerle birlikte yeni merkezin değerleridir;  $D_{ij}$ , her veri noktası arasındaki Öklid mesafesi ve merkezler ve n, yineleme sayısı.

Çıktı: Küme sayısı.

Algoritmanın Avantajları (Vani vd., 2019):

- KO algoritmasının anlaşılması daha kolay ve uygulaması basittir.

Algoritmanın Dezavantajları (Vani vd., 2019):

- Çakışan kümeler için etkili değildir
- Heterojen verileri etkili bir şekilde kümeleyemez.
- Karesi alınmış hata fonksiyonunun yerel optimumunu sağlar.
- Küme merkezini rastgele seçmek iyi sonuçlar vermeyebilir.

### 2.4. Silüet indeks analizi

Rousseeuw (1987) tarafından geliştirilen bu yöntemde veri setindeki her bir elemanın atandığı kümeye uygunluğu [-1. +1] arasında elde edilen silüet indeks değeri ile tanımlanmaktadır. Silüet indeks değerinin pozitif olması elemanın doğru kümeye atandığını, negatif olması ise elemanın yanlış kümeye atandığını ifade etmektedir. Silüet indeks değerinin miktarı ise elemanın atandığı kümeye olan üyelik derecesini göstermektedir (Örneğin tespit edilen Silüet indeks değeri +1 ise elemanın kesinlikle doğru kümeye atandığı. -1 ise elemanın kesinlikle yanlış kümeye atandığı anlaşılmaktadır). Silüet indeks değeri Denklem 7 ile hesaplanmaktadır (Günay Atbaş, 2008; Sönmez ve Kömüscü, 2008).

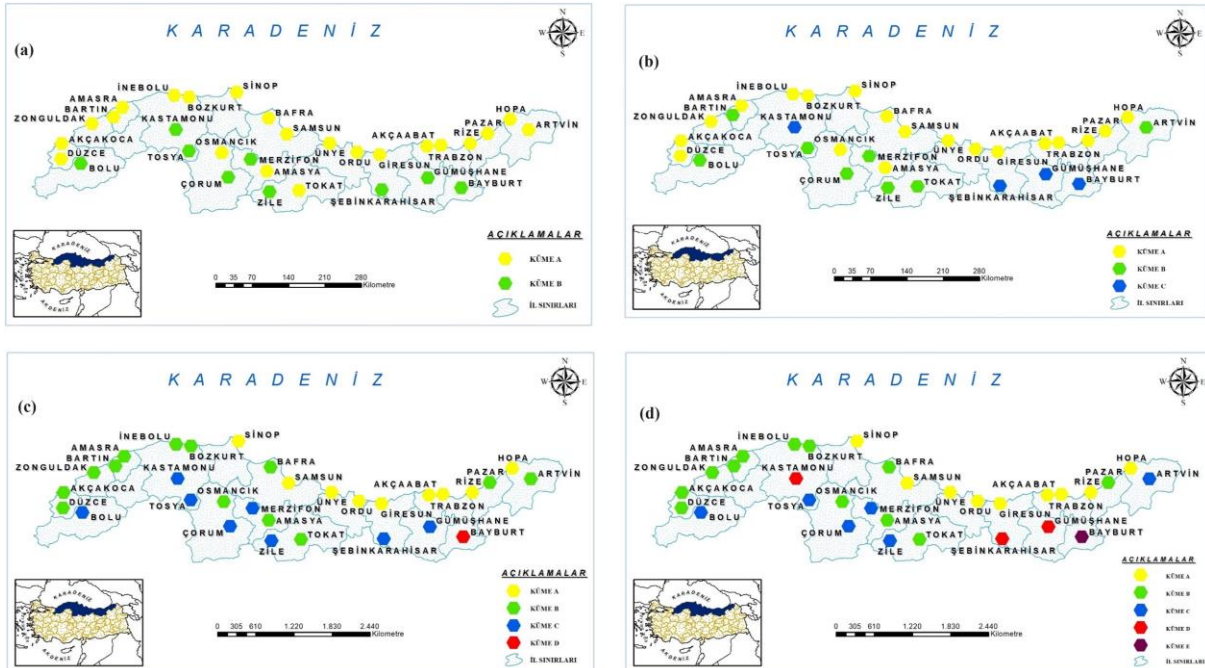
$$S(i) = \frac{\min\{b(i.m)-a(i)\}}{\max\{a(i). \min(b(i.m))\}} \quad (7)$$

Burada,  $a(i)$ ;  $i$ . nokta ile aynı kümede yer alan diğer tüm noktalar arasındaki ortalama mesafeyi,  $b(i.m)$ ;  $i$ . nokta ile  $m$ . kümede yer alan tüm noktalar arasındaki ortalama mesafeyi ifade etmektedir.

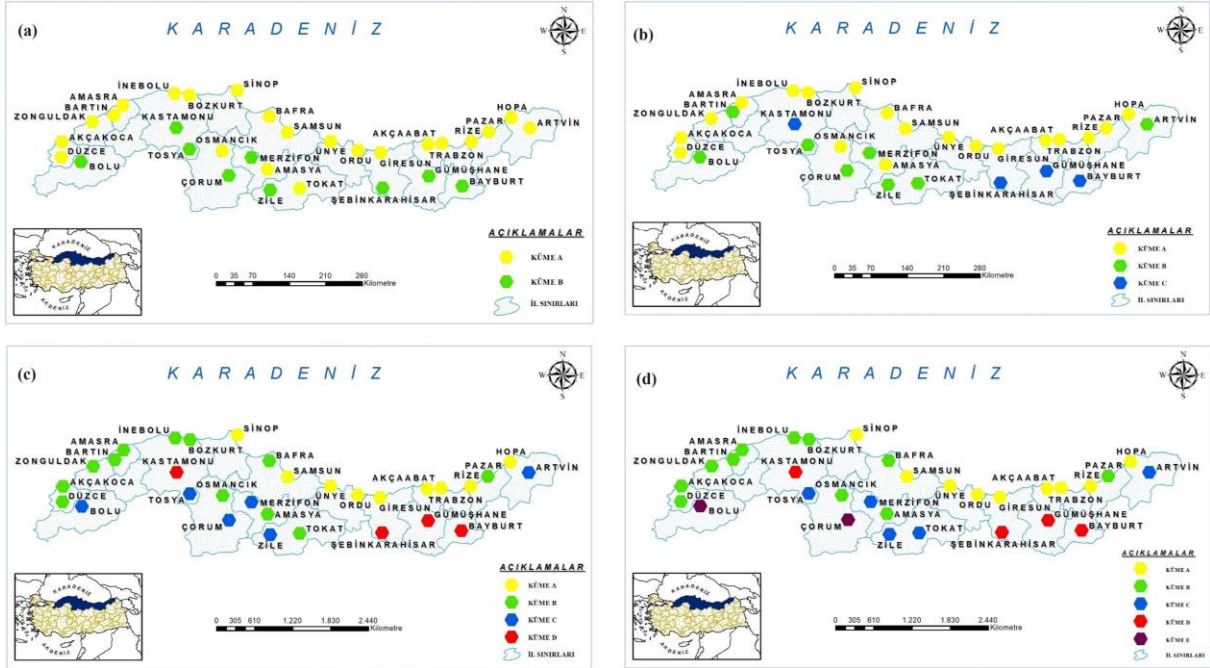
### 3. Bulgular

Bu çalışmada Karadeniz Bölgesindeki 31 istasyona ait periyodu 1982-2020 yılları arasında kapsayan yıllık ortalama sıcaklık gözlemleri kullanılarak benzer özellik gösteren kümeleri belirlemek amacıyla KO ve BCO algoritmalarından yararlanılmıştır. Analizler MATLAB R2016a kullanılarak gerçekleştirilmiştir. Çalışmada maksimum küme sayısı istasyon sayısının karekökünden küçük olacak şekilde 5 seçilmiştir (Pal ve Bezdek, 1995; Zhang vd., 2008; Karahan, 2011, 2019). Gözlem kayıtları kümeleme analizi yapılmadan önce bu veriler  $z = \frac{x_i - \bar{x}}{s}$  formülü kullanılarak standardize edilmiştir. Burada,  $x_i$   $i$ . sıradaki standardize edilen gözlem değeri;  $\bar{x}$  veri setinin ortalaması,  $s$  veri setinin standart sapması;  $z$  standardize değer olarak tanımlanmaktadır (Ünal vd., 2003; Dikbas vd., 2011).

2'den başlayarak maksimum küme sayısı olarak belirlenen 5'e kadar her bir küme sayısı için yapılan sınıflandırmada Karadeniz Bölgesi rüzgâr hızı serilerinin küme sayısı 2, 3, 4 ve 5 seçilerek KO ve BCO ile oluşturulan kümeler sırasıyla Şekil 2-3'te, kümelere ait özet bilgiler ise Tablo 3-8'de sunulmaktadır.



Şekil 2. KO ile oluşturulan kümeler.



Şekil 3. BCO ile oluşturulan kümeler.

KO ve BCO algoritmaları kullanılarak küme sayısı 2 seçildiğinde elde edilen sınıflandırma sonuçları her iki yöntem için de aynı olup Şekil 2a ve Şekil 3a'da gösterilmektedir. Sonuçlar incelendiğinde A kümesi Batı, Orta ve Doğu Karadeniz'de yer alan 22 istasyondan oluşmaktadır. B kümesi ise Batı, Orta ve Doğu Karadeniz iç kesimlerinde yer alan 9 istasyondan oluşmaktadır. Belirlenen kümelere ait yıllık maksimum, minimum, ortalama ve standart sapma değerleri Tablo 3'te sunulmaktadır.

Tablo 3. Küme sayısı 2 için KO ve BCO ile oluşturulan kümelerin özeti (°C).

Küme	Minimum	Maksimum	Ortalama	Standart Sapma
A	12.24	15.01	13.93	0.79
B	7.23	11.90	10.27	1.44

KO ve BCO kullanılarak küme sayısı 3 seçildiğinde elde edilen sınıflandırma sonuçları her iki yöntem için de aynı olup Şekil 2b ve Şekil 3b'de gösterilmektedir. Sonuçlar incelendiğinde B ve C kümelerinin bir önceki dağılımda yer alan B kümesinin iki alt kümesi olarak ayrıldığı görülmektedir. Böylece A kümesi Batı, Orta ve Doğu Karadeniz'de yer alan 19 istasyondan oluşmaktadır. B kümesi Batı, Orta ve Doğu Karadeniz'de yer alan 8 istasyondan oluşmaktadır. C kümesi ise Batı ve Doğu Karadeniz'de yer alan 4 istasyondan oluşmaktadır. Belirlenen kümelere ait maksimum, minimum, ortalama ve standart sapma değerleri Tablo 4'te sunulmaktadır.

Tablo 4. Küme sayısı 3 için KO ve BCO ile oluşturulan kümelerin özeti (°C).

Küme	Minimum	Maksimum	Ortalama	Standart Sapma
A	13.21	15.01	14.14	0.60
B	10.68	12.88	11.76	0.80
C	7.23	9.85	9.02	1.21

KO algoritması kullanılarak küme sayısı 4 seçildiğinde elde edilen sınıflandırma sonuçları Şekil 2c'de gösterilmektedir. Sonuçlar incelendiğinde D kümesi bir önceki dağılımda yer alan C kümesinin bir alt kümesi olarak ayrıldığı görülmektedir. Burada A kümesi Batı, Orta ve Doğu



Karadeniz kıyı kesiminde yer alan 9 istasyondan oluşmaktadır. B kümesi Batı, Orta ve Doğu Karadeniz’de yer alan 13 istasyondan oluşmaktadır. C kümesi Batı, Orta ve Doğu Karadeniz iç kesimlerinde yer alan 8 istasyondan oluşmaktadır. D kümesi ise Doğu Karadeniz iç kesiminde yer alan 1 istasyondan oluşmaktadır. Belirlenen kümelere ait maksimum, minimum, ortalama ve standart sapma değerleri Tablo 5’te sunulmaktadır.

Tablo 5. Küme sayısı 4 için KO ile oluşturulan kümelerin özeti (°C).

Küme	Minimum	Maksimum	Ortalama	Standart Sapma
A	14.46	15.01	14.71	0.16
B	12.24	13.98	13.39	0.53
C	9.38	11.90	10.64	0.95
D	7.23	7.23	7.23	-

KO kullanılarak küme sayısı 5 seçildiğinde elde edilen sınıflandırma sonuçları Şekil 2d ile gösterilmektedir. Sonuçlar incelendiğinde C ve D kümeleri bir önceki dağılımda yer alan C kümesinin iki alt kümesi olarak ayrıldığı görülmektedir. Ayrıca A ve E kümelerinin bir önceki dağılımdaki bütünlüğünü koruduğu, Bayburt istasyonunun D kümesinden ayrılıp E kümesine atandığı görülmektedir. Böylece A kümesi Batı, Orta ve Doğu Karadeniz kıyı kesimlerinde yer alan 9 istasyondan oluşmaktadır. B kümesi Batı, Orta ve Doğu Karadeniz’de yer alan 12 istasyondan oluşmaktadır. C kümesi Batı, Orta ve Doğu Karadeniz iç kesimlerinde yer alan 6 istasyondan oluşmaktadır. D kümesi Batı ve Doğu Karadeniz iç kesimlerinde yer alan 3 istasyondan oluşmaktadır. E kümesi ise Doğu Karadeniz iç kesiminde yer alan 1 istasyondan oluşmaktadır. Belirlenen kümelere ait maksimum, minimum, ortalama ve standart sapma değerleri Tablo 6’de sunulmaktadır.

Tablo 6. Küme sayısı 5 için KO ile oluşturulan kümelerin özeti (°C).

Küme	Minimum	Maksimum	Ortalama	Standart Sapma
A	14.46	15.01	14.71	0.16
B	12.60	13.98	13.48	0.42
C	10.68	12.24	11.43	0.61
D	9.38	9.85	9.61	0.24
E	7.23	7.23	7.23	-

BCO algoritması kullanılarak küme sayısı 4 seçildiğinde elde edilen sınıflandırma sonuçları Şekil 3c’de gösterilmektedir. Sonuçlar incelendiğinde Kastamonu, Şebinkarahisar, Gümüşhane ve Bayburt istasyonlarının bir önceki dağılımdaki bütünlüğünün bozulmadığı ve D kümesine atandığı görülmektedir. Burada A kümesi Batı, Orta ve Doğu Karadeniz kıyı kesiminde yer alan 9 istasyondan oluşmaktadır. B ve C kümeleri Batı, Orta ve Doğu Karadeniz’de yer alan sırasıyla 12 ve 6 istasyondan oluşmaktadır. E kümesi ise Batı ve Doğu Karadeniz’de yer alan 4 istasyondan oluşmaktadır. Belirlenen kümelere ait maksimum, minimum, ortalama ve standart sapma değerleri Tablo 7’de sunulmaktadır.

Tablo 7. Küme sayısı 4 için BCO ile oluşturulan kümelerin özeti (°C).

Küme	Minimum	Maksimum	Ortalama	Standart Sapma
A	14.46	15.01	14.71	0.16
B	12.60	13.98	13.48	0.42
C	10.68	12.24	11.43	0.61
D	7.23	9.85	9.02	1.21

BCO yöntemi kullanılarak küme sayısı 5 seçildiğinde elde edilen sınıflandırma sonuçları Şekil 3d ile gösterilmektedir. Sonuçlar incelendiğinde A ve D kümelerinin bir önceki dağılımdaki bütünlüğünü koruduğu, E kümesinin ise bir önceki dağılımdaki C kümesinin bir alt kümesi olarak ayrıldığı görülmektedir. Böylece A kümesi Batı, Orta ve Doğu Karadeniz kıyı

kesimlerinde yer alan 9 istasyondan oluşmaktadır. B kümesi Batı, Orta ve Doğu Karadeniz’de yer alan 11 istasyondan oluşmaktadır. C kümesi Batı, Orta ve Doğu Karadeniz iç kesimlerinde yer alan 5 istasyondan oluşmaktadır. D kümesi Batı ve Doğu Karadeniz iç kesimlerinde yer alan 4 istasyondan oluşmaktadır. E kümesi ise Batı ve Orta Karadeniz iç kesimlerinde yer alan 2 istasyondan oluşmaktadır. Belirlenen kümelere ait maksimum, minimum, ortalama ve standart sapma değerleri Tablo 8’de sunulmaktadır.

Tablo 8. Küme sayısı 5 için BCO ile oluşturulan kümelerin özeti (°C).

Küme	Minimum	Maksimum	Ortalama	Standart Sapma
A	14.46	15.01	14.71	0.16
B	12.88	13.98	13.56	0.33
C	11.37	12.60	11.94	0.49
D	7.23	9.85	9.02	1.21
E	10.68	10.79	10.73	0.08

2’den başlayıp maksimum küme sayısı olarak belirlenen 5’e kadar her bir küme sayısı için KO ve BCO algortimaları kullanılarak kümeler belirlenmiştir. Oluşturulan kümelerin doğruluğunu analiz etmek ve optimum küme sayısını belirlemek için Silüet indeks analizinden yararlanılmıştır. Öncelikle her bir küme sayısı için elde edilen sonuçların Silüet indeks değerleri hesaplanmıştır. Daha sonra her bir kümeye ait ortalama silüet indeks değerleri ve negatif Silüet indeks sayıları belirlenmiştir. Optimum küme sayısı, ortalama silüet indeks değerinin maksimum olması ve negatif Silüet indeks değerinin bulunmaması koşuluna göre tespit edilmiştir (Sönmez ve Kömüşçü, 2008).

BCO ve KO ile 2 küme sayısından 5 küme sayısına kadar belirlenen her bir küme için ortalama Silüet indeks değerleri ve negatif silüet indeks sayıları sırası ile Tablo 9-10’da verilmiştir.

Tablo 9. KO ile belirlenen kümelerin Silüet İndeks Analizi sonuçları.

Küme sayısı	2	3	4	5
Ortalama Silüet İndeks Değeri	0.818	0.768	0.705	<b>0.778</b>
Negatif Silüet İndeks Sayısı	1	1	3	-

Tablo 10. BCO ile belirlenen kümelerin Silüet İndeks Analizi sonuçları.

Küme sayısı	2	3	4	5
Ortalama Silüet İndeks Değeri	0.818	0.768	<b>0.749</b>	0.665
Negatif Silüet İndeks Sayısı	1	1	-	3

Tablo 9-10 verilen Silüet indeks analizi yöntemi sonuçlarına göre sıcaklık gözlemleri kullanılarak KO ve BCO ile belirlenen kümeler arasından en uygun sınıflandırma ortalama silüet indeks değeri maksimum ve negatif silüet indeks sayısına sahip olmayan KO algoritması kullanılarak belirlenen 5 kümelili dağılım olmuştur. Buna karşılık BCO kullanılarak belirlenen kümeler arasından en uygun küme sayısı 4 olmaktadır.

#### 4. Sonuç

Bu çalışmada KO ve BCO kümeleme algoritmaları kullanılarak Karadeniz Bölgesindeki istasyonlara ait sıcaklık gözlem değerleri bakımından benzer özellik gösteren kümeler belirlenmiştir. Kümeleme analizi 2’den 5’e kadar 4 farklı küme sayısı için gerçekleştirilmiş olup optimum küme sayısı Silüet indeks analizi yöntemi kullanılarak belirlenmiştir. Yapılan analizler sonucunda BCO yöntemiyle 4 küme, KO yöntemine göre 5 küme belirlenmiştir. En uygun kümeler ise Silüet indeks analizi sonucunda göre KO yöntemine ve 5 küme olarak belirlenmiştir. Gerek KO gerekse de BCO kullanılarak elde edilen ve önerilen kümelerdeki

temel farklılığın yöntemlerden kaynaklandığı düşünülmektedir. Her iki kümeleme algoritması sonucu önerilen kümelerde, küme bazında bütünlük genel olarak korunmuş olup, benzer sıcaklık özelliklerine sahip istasyonların bir arada olduğu belirlenmiştir. Bu çalışmanın devamı olarak

- Sıcaklık gözlemlerine ek olarak farklı hidro-meteorolojik iklim parametreleri ile coğrafi konum bilgilerin kullanıldığı kombinasyonlar ile benzer özellik gösteren kümelerin belirlenmesi önerilmektedir.
- Hiyerarşik olmayan yöntemlerin dışında hiyerarşik yöntemler ile hibrit yöntemlerin kullanıldığı kümeleme çalışmalarının yapılması önerilmektedir.
- Kümeleme analizi çalışmasının Türkiye coğrafyasında yer alan diğer bölgeler ile tüm ülke coğrafyasını kapsayacak şekilde gerçekleştirilmesi önerilmektedir.

## Kaynaklar

Bezdek, J.C. (1980). A Convergence Theorem for the Fuzzy ISODATA Clustering Algorithms. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2(1), 1-8.

Bezdek, J.C., Ehrlich, R., & Full, W. (1984). FCM: The Fuzzy C-Means Clustering Algorithm. *Computers and Geosciences*, 10(2-3), 191-203.

Çelik, M.A., Bayram, H., & Özüpekçe, S. (2018). An Assessment on Climatological, Meteorological and Hydrological Disasters That Occurred in Turkey in The Last 30 Years (1987-2017). *International Journal of Geography and Geography Education*, 38, 295-310.

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

Dikbas, F., Firat, M., Koc, A.C., & Gungor, M. (2011). Classification of precipitation series using fuzzy cluster method. *International Journal of Climatology*, 32(10), 1596-1603.

Dunn, J.C. (1973). A fuzzy relative of the ISODATA process and its use in detecting compact well-separated clusters. *Journal of Cybernetics*, 3(3), 32-57.

Eriç, S. (1949). The climates of Turkey according to Thornthwaite's classifications. *Ann. Assoc. Am. Geograp.*, 39, 26-46.

Firat, M., Dikbaş, F., Koç, A.C., & Güngör, M. (2012). Classification of annual precipitations and identification of homogeneous regions using k-means Method. *Technical Journal*, 23(113), 6037-6050.

İyigün, C., Türkeş, M., Batmaz, İ., Yozgatlıgil, C., Purutçuoğlu, V., Kartal Koç, E., & Öztürk, M.Z. (2013). Clustering current climate regions of Turkey by using a multivariate statistical method. *Theor Appl Climatol*, 114, 95-106.

Karahan, H. (2011). *Bölgesel Yağış-Şiddet-Süre-Frekans Bağıntılarının Diferansiyel Gelişim Algoritması Kullanılarak Elde Edilmesi*. TÜBİTAK (108Y299). Sonuç Raporu. 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 Univ Muh Bilim Derg*, 25(8), 998-1013

Kır, G. (2021). *Kümeleme Analiz Yöntemleri İle Karadeniz Bölgesi Meteorolojik Verilerinin Değerlendirilmesi*. Yüksek Lisans Tezi. Ondokuz Mayıs Üniversitesi, Lisansüstü Eğitim Enstitüsü, Samsun, 112.

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

Kulkarni, A., & Kripalani, R. (1998). Rainfall patterns over India: Classification with fuzzy c-means method. *Theor. Appl. Climatol.*, 59, 137–146.

Özkoca, T. (2015). *Orta Karadeniz kıyı illerinin hidrometeorolojik verilerinin trend analizi*. Yüksek Lisans Tezi, Ondokuz Mayıs Üniversitesi Fen Bilimleri Enstitüsü, Samsun, 89.

Pal, N.R., & Bezdek, J.C. (1995). On Cluster validity for the fuzzy c-means model. *IEEE Transactions On Fuzzy Systems*, 3, 370-379.

Rau, P., Bourrel, L., Labat, D., Melo, P., Dewitte, B, Frappart, F., Lavado, W., & Felipe, O. (2017). Regionalization of rainfall over the Peruvian Pacific slope and coast. *International Journal of Climatology*, 37(1), 143-158.

Rousseuw, P.J. (1987). Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. *Journal of Computational and Applied Mathematics*, 20, 53-65.

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

Sönmez, İ., & Kömüscü, A. (2008). Redefinition rainfall regions using k-means clustering methodology and changes of sub period. *İklim Değişikliği ve Çevre*, 1, 38–49.

Şahin, S. (2009). *Yapay sınır ağlarının iklim bölgelerinin belirlenmesinde kullanılması ve Ward metodu ile karşılaştırılması*. Doktora Lisans Tezi, İstanbul Teknik Üniversitesi Fen Bilimleri Enstitüsü, İstanbul, 344.

Şahin, S., & Cıgızoğlu, H.K. (2012). The sub-climate regions and the sub-precipitation regime regions in Turkey. *Journal of Hydrology*, 450-451, 180-189.

Türkeş, M. (1996). Spatial and temporal analysis of annual rainfall variations in Turkey. *Int. J. Climatol*, 16(9), 1057-1076.

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.

Türkeş, M., Sümer, U.M., & Çetiner, G. (2000). Küresel iklim değişikliği ve olası etkileri, Çevre Bakanlığı, Birleşmiş Milletler İklim Değişikliği Çerçeve Sözleşmesi Seminer Notları (13 Nisan 2000, İstanbul Sanayi Odası), 7-24, ÇKÖK Gn. Md., Ankara

Ünal, Y., Kındap, T., & Karaca, M. (2003). Redefining the climate zones of Turkey using cluster analysis. *Int. J. Climatol*, 23, 1045–1055.

Vani, H.Y., Anusuya, M.A., & Chayadevi, M.L. (2019). Fuzzy Clustering Algorithms-Comparative Studies For Noisy Speech Signals. *ICTACT Journal On Soft Computing*, 9(3), 1920-1926.

Xin, G.L., Min, F.Y., & Wen, T.H. (2011). Speech Recognition based on K-Means Clustering and Neural Network Ensembles. *Proceedings of 7th International Conference on Natural Computation*.

Zeybekoğlu, U., & Ülke Keskin, A. (2020). Defining rainfall intensity clusters in Turkey by using the fuzzy c-means algorithm. *Geofizika*, 37(2), 181–195.

Zhang, Y., Wang, W., Zhang, X., & Li, Y. (2008). A cluster validity index for fuzzy clustering. *Information Sciences*, 178, 1205–1218.