

## Bulanık c ortalamalar yöntemi ile Karadeniz bölgesi standart süreli yağış şiddetlerinin kümelenmesi

*Clustering of standard duration rainfall intensity of the Black Sea region by fuzzy c means method*

Utku ZEYBEKOĞLU\*<sup>1,a</sup>

<sup>1</sup>Sinop Üniversitesi, Boyabat Meslek Yüksekokulu, İnşaat Bölümü, 57200, Boyabat, Sinop

• Geliş tarihi / Received: 19.01.2021 • Düzeltilerek geliş tarihi / Received in revised form: 26.02.2021 • Kabul tarihi / Accepted: 28.02.2021

### Düzeltilme

Gümüşhane Üniversitesi Fen Bilimleri Enstitüsü Dergisi'nin sekizinci cildinin birinci sayısında yer alan "Zeybekoğlu, U. (2018). Bulanık C Ortalamalar Yöntemi ile Karadeniz Bölgesi Standart Süreli Yağış Şiddetlerinin Kümelenmesi. Gümüşhane Üniversitesi Fen Bilimleri Enstitüsü Dergisi, 8(1), 147-154. doi: 10.17714/gumusfenbil.334693" referanslı makalede yazar tarafından kaynakçada ve metin içerisinde yer almasına rağmen makale içerisinde bazı yerlerde bir referansın makale hazırlanma aşamasındayken sehven unutulduğu fark edilmiştir. Yapılan bu hatadan dolayı okuyucularımızdan özür dileriz. Bu düzeltme yazısında ilgili makalenin metin içerisinde sehven unutulmuş atıfları kapsayan yerler sunulmuştur.

### Erratum

In the article with "Zeybekoğlu, U. (2018). Clustering of Standard Duration Rainfall Intensity of the Black Sea Region by Fuzzy c Means Method. Gümüşhane University Journal of the Science and Technology Institute, 8(1), 147-154. doi: 10.17714 / gumusfenbil.334693" citation information which was published in the first issue of the eighth volume of the Gümüşhane University Journal of Science and Technology Institute, the author unintentionally forgotten a reference in some parts of the article while the article was being prepared, although it is included in the references and in the text. We apologize to the readers for the mistake. In this erratum, an unintentionally forgotten citation is presented in some parts of the text of the related article.

**Orijinal Makalenin DOI numarası:** 10.17714/gumusfenbil.334693

\*<sup>a</sup> Utku ZEYBEKOĞLU; utkuz@sinop.edu.tr, Tel: (0368) 315 01 01, orcid.org/0000-0001-5307-8563

## 1. Giriş

Yağışın süresi, şiddeti, alansal ve zamansal değişimi vb. özelliklerinin bilinmesi; su kaynakları, tarım, kentleşme, drenaj, taşkın kontrolü ve ulaşım gibi farklı sektörler için planlama çalışmaları ile bu sistemler üzerindeki çeşitli mühendislik yapılarının güvenli ve ekonomik olarak boyutlandırılması ve işletilmesi için gereklidir. Hidro-meteorolojik verilerin (yağış, akış, buharlaşma, sıcaklık...) eksiksiz ve düzenli olarak kayıt altına alınması birçok sektör açısından önem arz etmektedir. Bu kayıtlar ile ilgili yapılmış basılı son çalışma DSİ ve DMİ tarafından 1990 yılında yapılmıştır ve güncel verileri içermediğinden sağlıklı çalışmalar gerçekleştirilememektedir (Karahana, 2011).

Kümeleme analizinde genel olarak, aynı özelliklere sahip benzer verilerin bir araya toplanması ve ortak bir grup altında sınıflandırılması esas alınmaktadır. Her bir küme bir küme merkezi ile ifade edilmekte ve küme merkezinin yakınındaki veri noktaları gruplandırılarak o kümeye dahil edilmektedir. Bu işlem, veri noktaları ile küme merkezleri arasında hesaplanan uzaklıkların minimize, küme merkezleri arasındaki uzaklıkların ise maksimize edilmesi esasına dayanmaktadır (Karahana, 2011).

Hidrolojik kümeleme çalışmalarında, hidro-meteorolojik veriler yerine coğrafi özellikler kullanıldığından, bu çalışmalar sonucu elde edilen kümeler hidrolojik açıdan homojen olmazlar. Hidro-meteorolojik verilerin coğrafi konum ile birlikte kullanılması hidrolojik kümeleme çalışmalarının daha iyi sonuç vermesini sağlayabilmektedir.

Ülkemizde ve dünyada hidrolojik kümelerin belirlenmesi üzerine birçok araştırmacı farklı hidro-meteorolojik veriler üzerine farklı kümeleme yöntemleri kullanarak homojen kümeleri belirlemişlerdir. Bu çalışmalardan başlıcaları Mosley (1981), Acreeman ve Sinclair (1986), Burn (1989, 1997, 2000), Lecce (2000), Burn vd. (1997), Guttman, (1993), Soltani ve Modarres, (2006), Demirel (2004), Turan (2005), Kâhya vd. (2007), Demirel vd. (2007), Isik ve Singh (2009), Karahana (2011), Fırat vd. (2012; 2013), Dikbaş vd. (2012), Alemaw ve Chaoka (2016) gerçekleştirmiş oldukları çalışmalarıdır.

Ancak yazarın literatür araştırması sonucunda yağış şiddetleri verileri kullanarak yapılmış kümeleme çalışmalarının sınırlı sayıda bulunduğu görülmüştür.

Önerilen bu çalışmada Bulanık C Ortalamalar kümeleme yöntemi kullanılarak Karadeniz bölgesindeki istasyonların Standart Süreli Yağış Şiddeti (SSYŞ) verilerine göre homojen alt bölgelere ayrılması amaçlanmaktadır. Bunun için Karadeniz bölgesinde yer alan 15 ile ait 18 meteoroloji gözlem istasyonuna ait veriler kullanılmış olup kümeleme analizi gerçekleştirilmiştir.

## 2.2. Yöntem

### 2.2.1. Bulanık c ortalamalar

Bu çalışmada ortak özellik taşıyan veri setlerinin kümeleme işlemi, literatürde yaygın olarak kullanılan yöntemlerden biri olan Bulanık C-Ortalamalar (Fuzzy C-Means) yöntemi kullanılarak yapılmıştır (Bezdek, 1981; Karahana, 2011).

Fuzzy c-means algoritmasının en önemli özelliği olan üyelik matrisinin kümeleme üzerinde olumlu etkileri vardır. Bu matris belirsiz durumların tanımlanmasını kolaylaştırır (Azem, 2013). Ayrıca üyelik dereceleri düşük olduğundan sıra dışı verilerin etkisi azdır. Esnek bir yapıya sahiptir. Örtüşen kümeleri bulma kabiliyeti diğer bölümlü algoritmalarla göre daha fazladır.

Bulanık c-ortalama yönteminde, kümeleme işlemi tanımlanan bir amaç fonksiyonunun minimize edilmesiyle yapılmaktadır. Yöntem, aşağıdaki hesap adımlarına göre kümeleme işlemini yapmaktadır (Karahana, 2011).

$X = \{x_j\}_{j=1}^n \subseteq R^d$  kümelenecek n adet veri setini

ve c küme sayısını göstermek üzere, X veri setinin c adet kümeye bölünebilmesi için  $c \times n$  boyutunda ve aşağıdaki koşulları (Denklemler 1-3) sağlayan üyelik fonksiyonu matrisinin,  $U = [u_{ij}]$

başlangıçta rastgele olarak oluşturulmaktadır (Karahana, 2011).

$$0 \leq u_{ij} \leq 1; i = 1,2,3,\dots,c; j = 1,2,3,\dots,n \quad (1)$$

$$\sum_{i=1}^c u_{ij}; j = 1,2,3,\dots,n \quad (2)$$

$$\sum_{j=1}^n u_{ij} > 0; i = 1,2,3,\dots,c \quad (3)$$

Denklemler 1-3'te,  $u_{ij}$ , i-nolu küme merkezinden j-nolu veri noktasına olan üyelik değerini göstermektedir. Hesaplanan üyelik değerleri

kullanılarak aşağıdaki amaç fonksiyonunun (Denklem 4) minimize edilmesiyle kümeleme işlemi yapılmaktadır (Karahan, 2011).

$$J_m(U, V) = \sum_{i=1}^c \sum_{j=1}^n u_{ij}^m \|x_j - v_i\|^2 \quad (4)$$

Denklem 4'te m üyelik değerlerinin ağırlıklarını ayarlamakta kullanılan bulanıklık terimini ( $1 \leq m \leq 2$ ),  $\|\bullet\|$ , normu ve  $V = \{v_i\}_{i=1}^c \subset R^d$  i

se her bir küme merkezini içeren çözüm vektörünü göstermektedir. Bu çalışmada; Öklid normu kullanılmıştır (Karahan, 2011). Küme merkezleri, Denklem 4'ün V'ye göre kısmi türevinin alınıp sıfıra eşitlenmesiyle Denklem 5'teki gibi hesaplanmaktadır (Karahan, 2011).

$$V_i = \frac{\sum_{j=1}^n u_{ij}^m x_j}{\sum_{j=1}^n u_{ij}^m} \quad i = 1, 2, 3, \dots, c \quad (5)$$

Bu aşamadan sonra, hesaplanan küme merkezleri için yeni uygunluk değerleri, Denklem 4'ün U'ya göre kısmi türevinin alınarak sıfıra eşitlenmesi sonucu Denklem 6 kullanılarak hesaplanmaktadır (Karahan, 2011).

$$u_{ij} = \left( \sum_{k=1}^c \left( \frac{\|x_j - v_k\|}{\|x_j - v_i\|} \right)^{\frac{2}{m-1}} \right)^{-1} \quad i = 1, 2, 3, \dots, c; j = 1, 2, 3, \dots, n \quad (6)$$

Denklem 6 kullanılarak güncellenen üyelik değerleri ile Denklem 4 ile 6 arasındaki işlemler tekrar yapılmakta ve aşağıda tanımlanan durma

koşulu sağlanıncaya kadar devam etmektedir (Karahan, 2011).

$$\|V^{(l+1)} - V^{(l)}\| < \zeta \quad (7)$$

Denklem 7'de,  $V^{(l+1)}$  ve  $V^{(l)}$  değerleri sırasıyla  $l+1$  ve  $l$  -inci iterasyonlardaki küme merkezlerini,  $\zeta$  ise iki iterasyon arasında sağlanması gereken hoşgörü değerini ifade etmektedir (Karahan, 2011).

Bulanık C Ortalamalar yöntemi ile literatürde kullanılan diğer kümeleme yöntemlerinin tümünde karşılaşılan en önemli sorun, hesap başlangıcındaki küme sayısının seçilmesidir. Ancak, gerçekte küme sayısının ne olması gerektiği de bilinmeyen olup, veri setine bağlı olarak değişim göstermektedir (Karahan, 2011).

Bu nedenle, küme sayısının değerini belirlemek için çeşitli indeksleme yaklaşımları geliştirilmiş ve küme sayıları bu indeks değerlerine göre belirlenmiştir (Pal vd., 1995; Velthuzien vd., 1997; Sugar ve James, 2003; Shen vd., 2005; Zhang vd., 2008; Karahan, 2011). Bu indeksleme yöntemlerinde maksimum küme sayısının önceden seçilmesi gerekmektedir. Ancak; maksimum küme sayısının belirlenmesinde araştırmacılar arasında genel bir uzlaşma bulunmamasına rağmen, araştırmacıların çoğunluğunun maksimum küme sayısının istasyon sayısının karekökünden küçük olması gerektiği konusunda uzlaştıkları Zhang vd. (2008) tarafından ifade edilmektedir (Karahan, 2011).