

TÜRKİYE İKLİM BÖLGELERİNİN HİYERARŞİK KÜMELEME YÖNTEMİ İLE BELİRLENMESİ

Elçin KARTAL*
Fidan M. FAHMI****
Vilda PURUTÇUOĞLU*****
Gülser KÖKSAL*****

Cem İYİGÜN**
Ceylan YOZGATLIGİL****
İnci BATMAZ*****
Murat TÜRKEŞ*****

ÖZET

Hükümetlerarası İklim Değişikliği Paneli (IPCC)'ne göre yerküre ısınmaktadır. Bunun nedeni atmosferdeki CO₂ ve diğer sera gazı birikimindeki artıştır. Küresel ısınmaya bağlı iklim değişikliğinin etkisi deniz seviyesinin yükselmesi, buzulların erimesi, iklim bölgelerinin sınırlarının değişmesi, meteorolojik olaylar ve buna bağlı doğal afetlerin artması gibi bölgesel ve zamansal olarak değişik biçimlerde ortaya çıkmaktadır (Türkeş vd., 1999a; Türkeş, 2008). İklim değişikliğini klimatolojik ve meteorolojik verinin analizi ile incelemek mümkün olabilmektedir. Bu çalışmada tüm dünyada görülen iklim değişikliğinin, Türkiye'nin daha önce belirlenen iklim bölgelerinin sınırlarına olan etkisini incelemek amaçlanmıştır. Türkiye Devlet Meteoroloji İşleri Genel Müdürlüğü'nün farklı istasyonlarda 1950-2006 yılları arasında kaydettiği yağış ve çeşitli sıcaklık (ortalama, minimum, maksimum vb.) verilerine bir dizi ön işlem yapıldıktan sonra 'hiyerarşik kümeleme analizi' uygulanmıştır. Sıcaklık ve yağış değişkenlerinin farklı niteliklere sahip olması nedeniyle veriler standartlaştırılarak çok değişkenli kümeleme analizi ele alınmıştır. Çalışmada uygulanan hiyerarşik kümeleme yöntemleri, dört farklı uzaklık tanımını içeren tek bağlantı, ortalama bağlantı, tam bağlantı ve Ward yöntemleridir. Bu yöntemlerin kullanılmasıyla elde edilen kümelerin kararlılıklarının test edilebilmesi için yöntemler 1950-2006 ve 1970-2006 yılları arasında olmak üzere iki farklı zaman dilimi için uygulanmıştır. Yöntemlerden Ward yönteminin daha kararlı ve anlamlı kümeler oluşturduğu gözlemlenmiştir. Ward yöntemi sonucuna göre sıcaklık ve yağış özellikleri bakımından Türkiye'de yedi iklim bölgesi bulunduğu söylenebilir. Ancak bu iklim bölgeleri önceden bilinen bölgelerden farklıdır.

Anahtar Kelimeler: Hiyerarşik kümeleme yöntemi, Öznel değerlendirme, Türkiye iklim bölgeleri.

*Ar. Gör., Orta Doğu Teknik Üniversitesi, İstatistik Bölümü, e-posta: ekartal@metu.edu.tr

**Yrd. Doç. Dr., Orta Doğu Teknik Üniversitesi, Endüstri Mühendisliği Bölümü, e-posta: iyigun@ie.metu.edu.tr

***Doktora Öğrencisi, Orta Doğu Teknik Üniversitesi, İstatistik Bölümü, e-posta: fidantelaferli@yahoo.com

****Yrd. Doç. Dr., Orta Doğu Teknik Üniversitesi, İstatistik Bölümü, e-posta: ceylan@metu.edu.tr

*****Yrd. Doç. Dr., Orta Doğu Teknik Üniversitesi, İstatistik Bölümü, e-posta: vpurutcu@metu.edu.tr

*****Doç. Dr., Orta Doğu Teknik Üniversitesi, İstatistik Bölümü, e-posta: ibatmaz@metu.edu.tr

*****Prof. Dr., Orta Doğu Teknik Üniversitesi, Endüstri Mühendisliği Bölümü, e-posta: koksai@ie.metu.edu.tr

*****Prof. Dr., Çanakkale Onsekiz Mart Üniversitesi, Coğrafya Bölümü, e-posta: murat.turkes@comu.edu.tr

Bu çalışma, Orta Doğu Teknik Üniversitesi tarafından BAP-2008-01-09-02 nolu bilimsel araştırma projesi olarak desteklenmektedir.

Makale, 28-30 Haziran 2010 tarihlerinde ODTÜ tarafından düzenlenen olan 7. İstatistik Günleri Sempozyumu'nda sunulmuş olup, herhangi bir yerde yayımlanmamıştır.

1. GİRİŞ

İklim değişiminin etkisi gün geçtikçe artmaktadır. Bu etki bölgesel ve zamansal olarak farklı biçimlerde ortaya çıkmaktadır (Türkeş vd., 1999a; Türkeş, 2008). İklim bölgelerindeki olası değişimleri klimatolojik ve meteorolojik verinin analizi ile incelemek mümkündür. Bu amaçla kullanılan en yaygın ampirik yöntemler hava durumu etkilerini göz önüne alan Koeppen ve Thorntwaite sınıflandırmasıdır (Türkeş, 2010). Bu nicel yöntemde kullanılan kurallar özeldir ve iklim kuşakları geçişler yerine kesin sınırlarla çizilmektedir. Diğer yandan öznel sınıflandırmaların yerine, nesnel bir yöntem olan tanımlayıcı veri madenciliği gibi istatistiksel kuralların gelişmesiyle birlikte, benzer iklim özelliklerine sahip bölgeleri gruplamak da mümkün olabilmektedir. Bu amaçla kullanılan en önemli yöntemlerden biri **kümeleme** analizidir. Örneğin Mexico City'nin iklim kuşaklarının belirlenmesine ait bir çalışmada çok değişkenli analizlerden biri olarak hiyerarşik kümeleme yöntemi kullanılmıştır (Estrada, 2009). Benzer şekilde Amerika Birleşik Devletlerinin iklim bölgeleri, sıcaklık ve yağış verilerine uygulanan hiyerarşik kümeleme yöntemi ile belirlenmeye çalışılmıştır (Fovell ve Fovell, 1993). Türkiye'ye ait bir çalışmada ise iklim kuşakları, 1951-1998 yılları arasındaki minimum, maksimum, ortalama sıcaklık ve toplam yağış değişkenlerine hiyerarşik kümeleme yöntemi uygulanarak incelenmiş ve analiz sonucunda yedi farklı iklim kuşağı elde edilmiştir (Ünal vd., 2003). Bu çalışmada elde edilen bölgeler daha önceden bilinen yedi iklim bölgesi (Erinç, 1984) ile karşılaştırılmış ve sınırlarda farklılaşmalar olduğu ortaya konmuştur. Türkiye'ye ait bir diğer çalışmada da iklim bölgeleri, merkeze dayalı kümeleme yöntemleri ile analiz edilmiş ve sonuçlar karşılaştırılmıştır (Fahmi vd., 2010).

Bu çalışmada da Türkiye'nin iklim bölgeleri 1950-2006 yılları arasındaki yedi sıcaklık ve bir yağış değişkenlerine uygulanan hiyerarşik kümeleme yaklaşımıyla incelenmiştir. Hiyerarşik kümeleme yöntemi olarak tek bağlantı (single linkage), ortalama bağlantı (average linkage), tam bağlantı (complete linkage) ve Ward yöntemi olmak üzere dört yöntem kullanılmıştır. Yöntemler, oluşturdukları kümelerin kararlılık ve anlamlılıklarına göre karşılaştırılmıştır. Karşılaştırma sonucunda en kararlı ve anlamlı kümelemenin Ward yöntemiyle elde edildiği gözlenmiş ve analizlerin devamında bu yöntem esas alınarak Türkiye'deki iklim bölgeleri incelenmiştir. Analizler sonucunda Türkiye'de yedi iklim bölgesi olduğu ortaya konmuş ve bu bölgeler önceki çalışma sonuçlarıyla karşılaştırılmıştır. Makalenin içeriği ise şu şekilde sunulmuştur: Veriler hakkında ayrıntılı bilgi makalenin 2. Bölümü'nde yer almaktadır. Bölüm 3'de hiyerarşik kümeleme yönteminin esasları ve farklı uzaklık tanımlarına sahip dört hiyerarşik kümeleme yöntemlerinden söz edilmektedir. Bölüm 4, yöntemlerin uygulandığı ve bulguların grafikler yardımıyla sunulduğu bölümdür. En son olarak da Bölüm 5'te sonuç ve gelecekteki çalışmalara yer verilmektedir.

2. VERİLER

Türkiye'deki iklim bölgelerinin kümeleme yöntemi ile incelenmesi için gerekli olan klimatolojik ve meteorolojik veri Devlet Meteoroloji İşleri Genel Müdürlüğü (DMİGM)'nden alınmıştır. Çalışmada kullanılan değişkenler, iklim üzerinde etkileri bilinen sıcaklık ve yağış değişkenleridir. Sıcaklık verisi, ortalama, minimum, maksimum, minimum sıcaklıkların ortalaması, maksimum sıcaklıkların ortalaması, ortalama sıcaklıkların maksimumu ve ortalama sıcaklıkların minimumu olmak üzere yedi sıcaklık değişkenini içermektedir. Yağış verisi olarak da ortalama yağış değişkeni

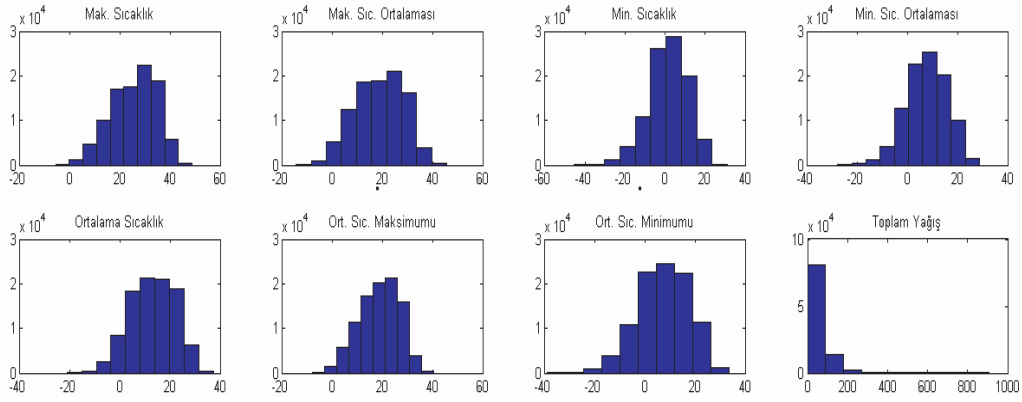
ele alınmıştır. Veriler, 277 istasyonda, 1950-2006 yılları arasında aylık olarak kaydedilmiş gözlemlerden oluşmaktadır. Ancak birçok istasyonda kayıp gözlemler olması nedeniyle veriler önışlemlerden geçirilmiştir. Bu amaçla en az iki yıl üst üste kayıp gözlem içeren istasyonlar çalışmada gözardı edilirken, yıl içindeki eksik gözlemler ortalama eksik doldurma yöntemi ile tamamlanmıştır. Aslan vd. (2010a, 2010b) çalışmalarında aynı klimatolojik ve meteorolojik veri üzerinde Beklenti En Büyükleme (EM) Markov Zinciri Monte Carlo (MCMC) yönteminin, kayıp verileri doldurmada en iyi yöntem olduğunu göstermişlerdir. Nitekim kayıp verilerin önerilen bu yöntemle doldurulması, halen devam eden bir çalışmadır. Bu sebeple bu çalışmada, Türkiye'deki iklim bilimcilerin kullandığı ortalama alma yöntemi kullanılmıştır. Ayrıca sonuçlarda oluşabilecek yanlılığı azaltmak amacıyla, en az iki yıl üst üste kayıp gözlem içeren istasyonlar çalışmada kullanılmamıştır. Sonuç olarak 1950-2006 yıllarında bu koşulu sağlayan 62 istasyon bulunmuş ve bunların tamamı da analizlerde kullanılmıştır. En çok istasyon bilgisinin yer aldığı zaman aralığı ise 1970-2006 yılları olarak gözlemlenmiştir.

1970-2006 zaman aralığı sıcaklık ve yağış değişkenlerine ilişkin tanımlayıcı istatistikler Tablo 1'de verilmiştir. Değişkenlerin basıklık ve çarpıklık istatistikleri ve histogramları incelendiğinde (Şekil 1), sıcaklık değişkenlerinin az da olsa sola çarpık bir dağılıma sahip olduğu söylenebilir. Yağış değişkeninin dağılımı ise sağa çarpık bir dağılımdır. Ayrıca sıcaklık değişkenlerine ait basıklık katsayıları üç değerine yakın olduğu için dağılımların normal dağılımdan çok uzaklaşmadığı söylenebilir. Oysa ki 14.26 basıklık değeriyle yağış değişkeni normal dağılımdan farklılaşarak sağa çarpık bir dağılım sergilemektedir. Nitekim Dünya Meteoroloji Örgütü'nün raporunda yıllık yağış verilerinin gamma dağılıma sahip olabileceği vurgulanmıştır (WMO, 1983).

Tablo 1. Sıcaklık ve yağış değişkenlerine ilişkin tanımlayıcı istatistikler

İstatistikler	Mak.	Mak.Sic. Ort.	Min.	Min. Sic. Ort.	Ort.	Ort. Sic. Mak.	Ort. Sic. Min.	Toplam Yağış
Ortalama	25,50	18,76	1,56	7,64	13,00	18,72	7,10	52,22
Standard Sapma	8,96	9,57	9,75	8,00	8,82	7,93	10,18	57,33
Minimum	-5,8	-14,5	-45,6	-27,9	-21,2	-8,2	-39,3	0,0
Maksimum	48,8	45,6	31,1	28,7	37,3	40,5	33,5	907,2
Basıklık (Kurtosis)	2,42	2,29	3,26	2,88	2,46	2,46	2,79	14,26
Çarpıklık (Skewness)	-0,33	-0,18	-0,48	-0,26	-0,22	-0,29	-0,31	2,52

(Verilerde Mak., Min., Sic. ve Ort. kısaltmaları, sırasıyla, maksimum, minimum, sıcaklık ve ortalamaı ifade etmektedir.)



Şekil 1. Sıcaklık ve yağış değişkenlerine ilişkin histogramlar
(Grafiklerdeki Mak., Min. ve Sic. kısaltmaları, sırasıyla, maksimum, minimum ve sıcaklığı ifade etmektedir.)

3. YÖNTEM

Atmosferik araştırmalarda yaygın olarak kullanılan yöntemlerden birisi kümeleme analizidir (Kalkstein vd., 1987). Kümeleme analizi, aynı grup içerisindeki nesnelerin birbirine benzer veya ilişkili olması; farklı gruptakilerin ise birbirinden farklı olması ya da ilişkilerinin bulunmamasını göz önüne alarak nesneleri gruplandırır. Kümeleme işleminde nesneler arasındaki uzaklığı (benzerliği) belirlemek için Öklit uzaklığı, Manhattan, Mahalanobis gibi değişik uzaklık ölçüleri kullanılmaktadır (Tan vd., 2006). Öklit uzaklığı birçok çalışmada olduğu gibi atmosferik çalışmalarda da en sık kullanılan uzaklık ölçüsüdür.

Kümeleme analizinde birçok çok değişkenli istatistik analizinde önemli olan, verilerin normalliği varsayımı, doğrusallık, ortak varyansa sahip olma gibi varsayımlar önem taşımamaktadır. Fakat bu varsayımlar yerine, örneklemin kitleyi temsil etme gücü dikkate alınır (Grimm ve Yarnold, 2000). Çünkü bu analizde kümeler, kitlenin yapısını göstermek için türetilirler. Çalışmada, Türkiye'nin çeşitli şehir ve ilçelerine yerleştirilmiş olan istasyonlardan elde edilen klimatolojik ve meteorolojik veri kullanıldığı için örneklemin kitleyi temsil gücünün yüksek olduğu düşünülmektedir.

Ayrıca iklim kuşaklarının belirlenmesine dair çoğu çalışmada hiyerarşik kümeleme analizi kullanılmıştır (Fovell ve Fovell, 1993; Ünal vd., 2003; Estrada vd., 2009). Hiyerarşik kümeleme yönteminde, verideki her bir gözlemin bir küme olduğu kabul edilerek analize başlanır ve her aşamada aralarında en yüksek benzerlik (en yakın) ölçümüne sahip olan kümeler birleştirilir. Kümelendirme işlemi tüm elemanların bir kümede birleşmesiyle son bulur. Kümeleme sırasındaki birleştirmeler uzaklık veya benzerlik matrisi sonuçlarını taşıyan hiyerarşik ağaç veya dal grafiği (dendogram) ile gösterilebilir (Tan vd., 2006) ve bu dendogramlar yardımıyla uygun **küme sayısı** belirlenebilir.

Hiyerarşik yöntemler, nesneler arasındaki uzaklıkları hesaplama yöntemlerine göre farklılıklar göstermektedirler. Tek bağlantı, tam bağlantı, ortalama bağlantı ve Ward yöntemi bunlardan bazılarıdır. Tek bağlantı yöntemi, her küme için o kümede yer alan gözlemlerin birbirlerine olan Öklit uzaklıklarını karşılaştırır. Kümelemeler yerel olarak

yapılır. İki kümenin birbirine en yakın olduğu alan dikkate alınarak kümenin uzak bölümleri ve kümelerin genel yapısı gözardı edilir. Bu nedenle tek bağlantı yönteminde gözlemlerin kümelerine sıralı biçimde bağlanabilmesinden kaynaklanan “zincirleme” problemi sıkça ortaya çıkmaktadır. Tam bağlantı yöntemi, zincirleme problemini ortadan kaldırmaya yönelik olarak geliştirilen bir yöntemdir. Bu yöntem, tek bağlantı yöntemine göre veriyi daha anlamlı kümelerle ayırmasına rağmen, aykırı gözlemlerden çok fazla etkilenmektedir. Ward yöntemi ise iklim çalışmasında sıkça kullanılan ve iyi sonuçlar verdiği ortaya konan bir yöntemdir (Ünal vd., 2003; Kalkstein vd., 1987). Bu yöntemde küme-içi hata kareler toplamındaki artış küçültülmeye çalışılmaktadır. Bu nedenle, Ward yöntemindeki yaklaşım, olabildiğince küçük ve eşit sayıda gözlem içeren kümelerin oluşmasını sağlamaktadır. İklim çalışmalarında eşit sayıda gözlem içeren kümelerin elde edilmesi gibi bir kısıt olmadığından dolayı Ward yöntemi ile elde edilen sonuçların anlamlılığı dikkatle incelenmelidir. Ortalama bağlantı, diğer yöntemlerin eksikliklerini ortadan kaldıran bir yöntemdir. Buna bağlı olarak tek bağlantı yönteminde gözlenen zincirleme problemi, bu yöntemde çok nadir olarak gözlemlenmektedir. Ward yönteminden farklı olarak ise küme-içi varyansın küçülmesi ile birlikte kümeler arası varyansın büyümesi de dikkate alınır. Bu nedenle, iklim çalışmalarında diğer yöntemlere göre daha anlamlı sonuçlar verebildiği söylenebilir (Kalkstein vd., 1987).

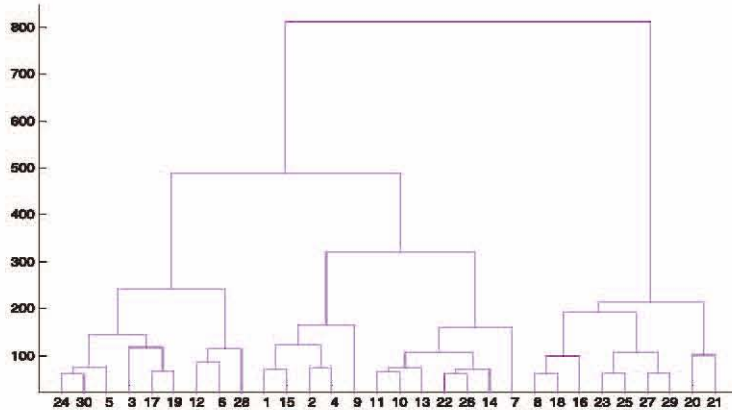
4. YÖNTEMLERİN UYGULANMASI VE BULGULAR

Türkiye’deki iklim bölgelerini yeniden incelemek amacıyla Türkiye DMİGM’ye ait 277 istasyonda kaydedilen yedi sıcaklık değişkeni (minimum, maksimum, ortalama, vb.) ile toplam yağış değişkeni ele alınmıştır. Hiyerarşik kümeleme yöntemi olarak farklı uzaklık hesaplama teknikleri içeren tek bağlantı, ortalama bağlantı, tam bağlantı ve Ward yöntemleri kullanılmıştır. Bu yöntemlerde Öklid uzaklık ölçüsü kullanıldığından dolayı değişkenlerin aynı ölçekte olmaları önem kazanmaktadır. Bu nedenle kümeleme analizinde eşit olarak ağırlıklandırılabilmesi için farklı ölçeklerde olan sıcaklık ve yağış değişkenleri ortalamaları sıfır ve varyansları bir olacak şekilde standartlaştırılmıştır. Ayrıca Bölüm 2’de vurgulandığı gibi en az iki yıl üst üste kayıp gözlem içeren istasyonlar çalışmada gözardı edilmiştir. Bu nedenle 1950-2006 yıllarında analizlerde kullanılabilen istasyon sayısı sadece 62’dir.

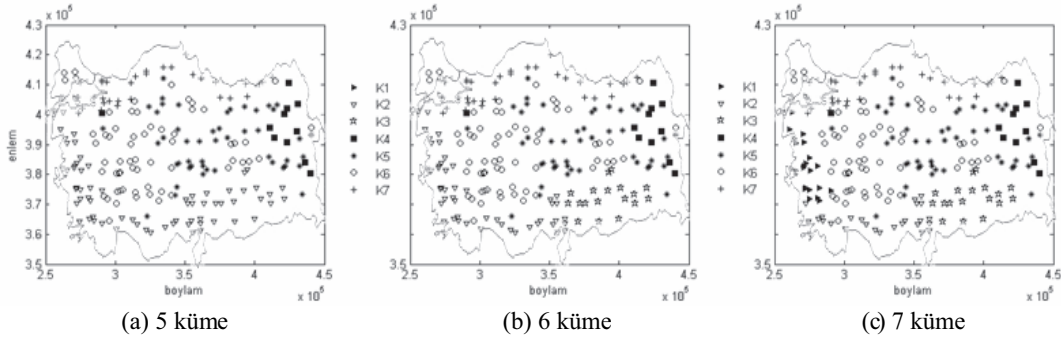
Buna ek olarak analizlerde, kümeleme analizinin daha güvenilir sonuçlar üretebilmesini sağlamak amacıyla daha fazla (220) istasyona ait klimatolojik ve meteorolojik verinin yer aldığı 1970-2006 zaman aralığında kullanılmıştır. Bu analizler sonucunda farklı yöntemlerle elde edilen kümelerin kararlılıklarını test edebilmek için 1950-2006 zaman aralığında klimatolojik/meteorolojik verisine sahip olan 62 istasyon için kümeleme analizi yeniden uygulanmıştır. Her iki veri kümesi için benzer gruplamalar oluşturularak kararlı küme yapıları sunan yöntemler, oluşturdukları kümelerin anlamlılıklarına göre de değerlendirilmişlerdir.

Bilindiği gibi birçok kümeleme yöntemindeki en önemli problem küme sayısının bilinmemesidir. Diğer yöntemlerden farklı olarak hiyerarşik kümelemede, oluşturulan kümeleri dendogram aracılığıyla geometriksel veya şekilsel olarak tanımlamak mümkündür. Dendogramda birbirine benzer istasyonlar, dallarla birbirine bağlanır (Şekil 2) ve bu dallar kümeler arası uzaklıkları ifade etmektedir. Dalların uzunluğu arttıkça kümeler arası uzaklık da artmaktadır. Bu özellikten faydalanarak anlamlı kümeler oluşturulabilir. Bu çalışmada, yöntemlerden elde edilen dendogramlar incelenmiş ve en uygun küme sayısının, tek bağlantı yöntemi için dört ve ortalama, tam

bağlantı ve Ward yöntemi içinse yedi olduğu sonucuna varılmıştır. Bu sonucun gerçekten anlamlı kümeler oluşturup oluşturulmadığını görebilmek amacıyla da yöntemler, değişik küme sayıları için uygulanmış ve sonuçların geçerliliği doğrulanmıştır. Bahsedilen analizlere ilişkin bir örnek Şekil 3’de verilmiştir. Bu şekilde, Ward yönteminin beş, altı ve yedi küme sayısı için oluşturduğu kümeler yer almaktadır. Bu kümeler Karadeniz, Marmara, Ege, Akdeniz, İç Anadolu, Doğu Anadolu ve Güneydoğu Anadolu olarak bilinen yedi bölgenin (Erinç, 1984) ismi kullanılarak değerlendirilmiştir. Şekil 3(a)’da K2, K4, K5, K6 ve K7 olmak üzere beş küme yer almaktadır. K2 ile ifade edilen küme Ege, Akdeniz ve Güneydoğu Anadolu bölgesindeki istasyonları kapsamaktadır. Bu küme, Şekil 3(b)’de bölünerek K2 ve K3 kümelerini oluşturmuştur. Burada K2 kümesi, Ege ve Akdeniz bölgesindeki istasyonları kapsarken, K3 ile ifade edilen küme, Güneydoğu Anadolu bölgesindeki istasyonları içermektedir. Şekil 3(c)’de ise Şekil 3(a)’da yer alan K2 kümesi K1, K2 ve K3 olmak üzere üç kümeye ayrılmıştır ve sırasıyla Ege, Akdeniz ve Güneydoğu Anadolu bölgelerindeki istasyonları gruplamaktadır. Dolayısıyla Şekil 3(c)’deki kümelendirmelerin Türkiye iklim özellikleri açısından anlamlı kümeleri oluşturduğu düşünülerek, Ward yöntemi sonucu bulunan anlamlı küme sayısı, yedi olarak belirlenmiştir. Benzer analizin tek bağlantı yöntemi ile tekrarlanması sonucunda ise uygun küme sayısının dört, ortalama bağlantı ve tam bağlantı yöntemleri ile belirlenmesiyle de yedi olduğu gözlenmiştir. Tek bağlantı yönteminin diğer yöntemlerden farklı sonuçlar koyma sebebinin, yöntemde sıkça görülen zincirleme probleminden kaynaklandığı belirlenmiştir.



Şekil 2. 1970-2006 zaman aralığındaki değişkenler için Ward yöntemine ait dendrogram
(İstasyon sayısı 30’dan fazla olduğu için alt dallar birleştirilmiştir. Bu nedenle en alttaki dallar birden fazla istasyonu içerebilmektedir.)



Şekil 3. Ward yönteminin 1970-2006 verilerine uygulanmasıyla elde edilen kümeler

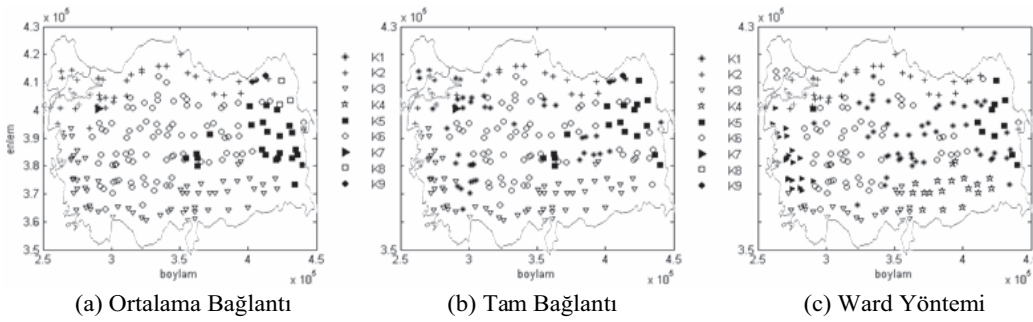
Diğer yandan, yöntemlerin uygulanmasıyla elde edilen kümelerin veriye uygunluğunu test edebilmek amacıyla Cophenetic İlişki Katsayısı (CPCC) kullanılmıştır (Sokal ve Rohlf, 1962). Değerlendirmede, bu katsayı bir'e yaklaştıkça elde edilen kümelemenin veriye uygunluğu artmaktadır. Dört farklı yöntemle oluşturulan kümelerin klimatolojik ve meteorolojik veriye uygunluğunu ölçen CPCC katsayıları Tablo 2'de verilmiştir.

Tablo 2. Cophenetic İlişki Katsayıları

Veri	Tek Bağlantı	Ortalama Bağlantı	Tam Bağlantı	Ward Yöntemi
1950-2006	0.601	0.782	0.659	0.635
1970-2006	0.508	0.722	0.701	0.634

Tablo 2'deki sonuçlar incelendiğinde, klimatolojik ve meteorolojik veriye en uygun kümelemenin ortalama bağlantı yöntemi ile oluşturulduğu söylenebilir. Ayrıca tek bağlantı tekniği ile elde edilen kümelerin veriye uygunluğunun, tam bağlantı, ortalama bağlantı ve Ward yöntemi ile elde edilen kümelerin veriye uygunluğundan daha kötü olduğu görülmektedir. Bu nedenle analize sadece ortalama bağlantı, tam bağlantı ve Ward yöntemleri ile devam edilmiştir. Şekil 4'de üç yöntemin yedi küme sayısı için edilen kümelendirme sonuçları gösterilmektedir.

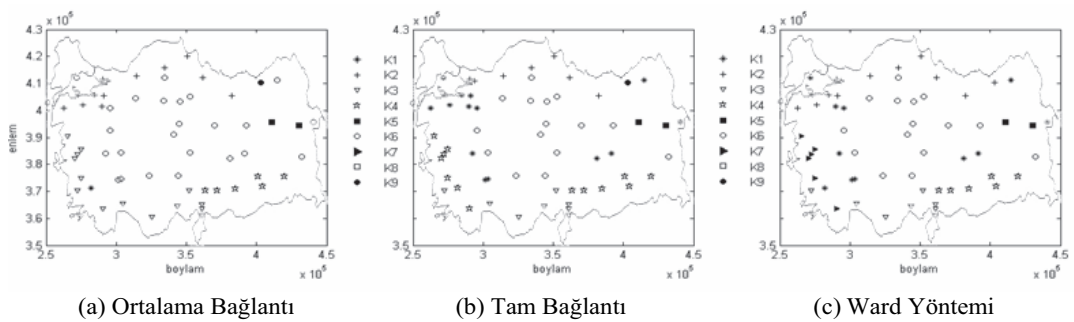
Şekil 4(a)'da ortalama bağlantı yöntemi ile elde edilen yedi kümede K7 kümesinin sadece Uludağ (Bursa) istasyonundan oluştuğu görülmektedir. K8 ve K9 kümeleri ise kuzeydoğu Karadeniz bölgesindeki ve kuzeydoğu Anadolu bölgesindeki üç istasyonu içermektedir. Bunlarla birlikte Ege, Akdeniz ve Güneydoğu Anadolu bölgelerindeki istasyonlar tek bir kümede yer almaktadır. Diğer üç kümeyi ise Karadeniz bölgesi, İç Anadolu Bölgesi ve Doğu Anadolu bölgesindeki istasyonlar oluşturmaktadır. Diğer yandan Şekil 4(b)'de yer alan tam bağlantı yönteminin sonuçları incelendiğinde Uludağ (Bursa) istasyonunun yine tek başına bir küme oluşturduğu ve Şekil 4(a)'daki gibi Doğu Karadeniz bölgesindeki üç istasyonun tek bir kümede yer aldığı bulunmuştur. Fakat Şekil 4(a)'dan farklı olarak, iç kısımlardaki istasyonlar K1 ve K6 olmak üzere iki farklı kümede yer almaktadır. Ayrıca Şekil 4(a)'da Karadeniz bölgesindeki istasyonlarla aynı kümede (K2) yer alan Marmara bölgesindeki istasyonlar, tam bağlantı yönteminde, İç Anadolu bölgesindeki istasyonlarla aynı kümeyi (K1) oluşturmaktadırlar.



Şekil 4. 1970-2006 zaman aralığındaki veriler için elde edilen kümeleme sonuçları

Son olarak Şekil 4(c)'de sunulan kümelerde, Ward yönteminin diğer iki yöntemden biraz daha farklı sonuçlar ortaya koyduğu bulunmuştur. Örneğin, K3 ile ifade edilen Ege, Akdeniz ve Güneydoğu Anadolu bölgesindeki istasyonları kapsayan küme, Ege bölgesindeki istasyonların yer aldığı K7, Akdeniz ve Güneydoğu Anadolu'daki istasyonların yer aldığı K3 ve K4 kümelerine ayrılmıştır. Ayrıca Marmara bölgesindeki istasyonlardan sadece kuzeybatı Marmara (Trakya)'da yer alan istasyonlar İç Anadolu bölgesindeki istasyonlarla benzer özellik göstererek aynı kümede yer almaktadırlar. İç bölgelerdeki kümelenme ise ortalama bağlantı yöntemindeki gibi bütünsel değildir. Yani buradaki istasyonları, sınırları belirgin iki kümede toplamak mümkün olmamaktadır. Bunun dışında diğer iki yöntemden farklı olarak, Bursa istasyonunun Doğu Anadolu bölgesindeki istasyonlarla aynı kümede yer aldığı-görülmektedir. Sonuç olarak, Türkiye'nin iklim özelliği göz önüne alındığında Ward yöntemi ile elde edilen kümelendirmenin diğer iki yöntemle elde edilen kümelendirmelerden daha gerçekçi olduğu söylenebilir.

Son olarak, ortalama, tam bağlantı ve Ward yöntemleri ile elde edilen kümelerin kararlılıklarını incelemek amacıyla, sadece 62 istasyon bilgisinin yer aldığı 1950-2006 zaman dilimindeki klimatolojik ve meteorolojik veri ele alınmıştır. Bu verilere uygulanan hiyerarşik kümeleme yöntemlerinin sonuçları Şekil 5'de yer almaktadır.



Şekil 5. 1950-2006 zaman aralığındaki veriler için elde edilen kümeleme sonuçları

Kümelendirmede, Şekil 5(a) ve (b)'de ortalama ve tam bağlantı yöntemleri ile elde edilen kümelerin, 1970-2006 yılları için elde edilen kümelerden daha farklı olduğu görülmektedir. Nitekim Şekil 4(a) ve (b)'de tek bir kümede (K3) yer alan Ege, Akdeniz ve Güneydoğu Anadolu istasyonları, Şekil 5(a)'da yer alan ortalama bağlantı yöntemi ile iki bölgeye ayrılmıştır. Bu bölgeler, Ege ve Akdeniz bölgesindeki istasyonları içeren K3 kümesi ile Güneydoğu Anadolu bölgesindeki istasyonları içeren K4 kümesidir.

Benzer şekilde Şekil 4(a) ve (b)'de tek bir kümede (K3) yer alan Ege, Akdeniz ve Güneydoğu Anadolu istasyonları, Şekil 5(b)'de yer alan tam bağlantı yöntemi ile Ege ve Güneydoğu Anadolu bölgesindeki istasyonları içeren K4 kümesi ile Akdeniz bölgesindeki istasyonları içeren K3 kümesine bölünmektedir. Buna karşın Ward yöntemi sonuçlarına göre Şekil 4(c) ile Şekil 5(c) de yer alan kümelendirmeler birbirine benzemektedir. Zira her iki veri için Ege, Akdeniz ve Güneydoğu bölgelerindeki istasyonlar K7, K3 ve K4 kümelerinde yer almaktadır. Ayrıca Kuzeybatı Marmara (Trakya) bölgesindeki istasyonların, İç Anadolu bölgesindeki istasyonlarla aynı kümede yer aldığı bilgisi de geçerliliğini korumaktadır. Dolayısıyla tüm bu sonuçlar dikkate alındığında, Ward yönteminin diğer iki yöntemle göre daha kararlı bir yapı sergilediği söylenebilir. Bu nedenle çalışmada Ward yöntem ile elde edilen kümelendirmeler ele alınmış ve analiz sonuçları, bu yöntemle göre değerlendirilmiştir. Belirlenen yeni bölgeler ve bu bölgelerdeki istasyonlar EK'de sunulmuştur.

5. TARTIŞMA VE SONUÇ

Bu çalışmada Türkiye'deki iklim bölgeleri sıcaklık ve yağış özelliklerine göre hiyerarşik kümeleme yöntemi ile incelenmiştir. Dört farklı hiyerarşik kümeleme yönteminin uygulanması ile elde edilen kümeler Türkiye iklim özellikleri açısından anlamlılıklarına ve kararlı yapıda olup olmamalarına göre karşılaştırılmışlardır. Analiz sonucunda Ward yöntemi ile elde edilen sonuçlar daha anlamlı bulunmuştur.

Bu yöntemle elde edilen sonuçlara göre Türkiye'de yedi iklim bölgesi olduğu söylenebilir. Fakat bulunan bu yedi bölge, önceden bilinen yedi iklim bölgesinden (Erinç, 1984) farklıdır. Örneğin, Kuzey Marmara (Trakya) bölgesi, İç Anadolu bölgesi ile benzer özellik göstermektedir. Ege Bölgesindeki istasyonlara bakıldığında ise orta ve doğu Ege'deki istasyonlarla İç Anadolu bölgesindeki istasyonlar aynı grupta yer almaktadır. Bunun dışında Güney Ege ile Akdeniz bölgesindeki istasyonlar birleşerek tek bir grup oluşturmuşlardır. Marmara bölgesi ile Karadeniz bölgesi birleşmiştir. Doğu Anadolu'nun Batı bölgesindeki istasyonların İç Anadolu iklim özelliği göstermesinden dolayı Doğu Anadolu bölgesinin sınırları daralarak, sadece en Doğu kesimlerdeki istasyonlar tek bir grupta toplanmışlardır. Bununla birlikte Doğu Beyazıt (Ağrı) ve Iğdır istasyonları da İç Anadolu özelliği taşımaktadır (Bkz., EK). Sonuç olarak İç Anadolu iklim yapısı Türkiye genelinde yayılırken kıyı bölgelerindeki iklim kümeleri daralmaktadır. Bu da kuraklığın yaygınlaşması olarak değerlendirilebilir.

Türkiye'nin iklim bölgelerinin incelendiği en son çalışmalardan Fahmi vd. (2010) tarafından elde edilen sonuçlar incelendiğinde bu çalışmada bulunan sonuçların desteklendiği söylenebilir. Fahmi vd. (2010) da sadece sıcaklık değişkenleri ele alınmış ve çalışma sonucunda beş bölge elde edilmiş olmasına rağmen Ege bölgesindeki istasyonların sıcaklık özelliği bakımından Akdeniz bölgesinde yer alan istasyonlarla benzer özellikler göstermesi ve Trakya'daki istasyonlar haricindeki Marmara bölgesi istasyonlarının Karadeniz bölgesindeki istasyonlarla aynı grupta yer alması iki çalışmanın ortak sonucudur.

Ünal vd. (2003)'nin çalışmasında da Ward yöntemi ile yedi iklim bölgesi elde edilmiştir. Çalışmada Ege ve Batı Akdeniz bölgeleri birleşmiş ve Doğu Akdeniz olarak başka bir bölge tanımlanmıştır. Bu çalışmada ise Güney Ege ile Akdeniz bölgesindeki istasyonlar birleşerek tek bir grup oluşturmuşlardır. Yani, Akdeniz bölgesinin sınırları batıya doğru biraz daha genişlemiştir. Ancak Ünal vd. (2003) bütün sıcaklık

değişkenlerini kullanmak yerine sadece üç sıcaklık verisi kullandığı için değişkenlerin bütün niteliklerini tam kapsayamamaktadır. Ayrıca küresel ısınma etkilerinin açık bir şekilde gözlemlendiği 1998 sonrası verileri içermemektedir. Bu nedenle sunduğumuz çalışma daha geniş kapsamlı olup bilinmekte olan eski iklim bölgelerinin (Erinç, 1984) halen var olduğunu, ancak sınırlarda değişmeler olduğu, Orta Anadolu bölgesi kurak ikliminin daha yaygınlaşıp kıyı şeridi ikliminde daralma olduğunu göstermektedir. Bunun iklim değişikliğinin ve Türkiye’de olası kuraklığın göstergelerinden biri olarak değerlendirilebileceği düşünülmektedir.

6. KAYNAKLAR

Aslan, S., Yozgatlıgil, C., İyigün, C., Batmaz, İ., Türkeş, M., Tatlı, H., 2010a. Comparison of missing value imputation methods for Turkish monthly total precipitation data. 9th International Conference of the Computer Data Analysis and Modeling: Complex Stochastic Data and Systems, Tam Metin Kitabı, Minsk, Belarus State University, 2, 137-140.

Aslan, S., Yozgatlıgil, C., İyigün, C., Batmaz, İ., Tatlı, H., 2010b. Meteorolojik zaman serilerinde kayıp veri tahmin yöntemlerinin başarımının korelasyon boyutu analiziyle karşılaştırılması. 7. İstatistik Günleri Sempozyumu, Özetler Kitabı ODTÜ, Ankara, Türkiye. 100-102.

Erinç, S. 1984. Climatology and its methods., 3rd edition. İstanbul, Gür-ay Press Inc. Estrada, F., Martinez-Arroyo A., Fernandez-Equiarte A., Luyando E., Gay C., 2009. Defining climate zones in Mexico City using multivariate analysis. Atmosfera, 22, 175-193.

Fahmi, F. Kartal, E., İyigün, C., Türkeş, M., Yozgatlıgil, C., Purutçuoğlu, V., Batmaz, İ., Köksal, G., 2010. Determining the climate zones of Turkey by center-based clustering methods. Nonlinear Dynamics of Complex Systems: Applications in Physical, Biological and Financial Systems, (Eds.) J.A. Tenreiro Machado, Dumitru Baleanu ve Albert Luo. Berlin: Springer (Basımda).

Fovell, G. R., Fovell, M., 1993. Climate zones of conterminous United States defined using cluster analysis. Journal of Climate, 6, 2103-2120.

Grimm, L. G., Yarnold, P. R., 2000. Reading and understanding more multivariate statistics. American Psychological Association, Washington, DC.

Kalkstein, L. S., Tan, G., Skindlov, J. A., 1987. An evaluation of three clustering procedures for use in synoptic climatological classification. Journal of Climate and Applied Meteorology, 26, 717-730.

Sokal, R. R., Rohlf F. J., 1962. The comparison of dendograms by objective methods. Taxon, 11, 33-40.

Tan, P. N., Steinbach, M., Kumar, V., 2006. Introduction to Data Mining. Pearson Addison Wesley, Inc., Boston.

Türkeş, M., 2008. Küresel iklim değişikliği nedir? Temel kavramlar, nedenleri, gözlenen ve öngörülen değişiklikler. İklim Değişikliği ve Çevre, 1, 45-64.

Türkeş, M., Sümer, U. M., Çetiner, G., 1999a. 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.

Türkeş, M., 2010. Klimatoloji ve Meteoroloji, Kriter Yayınevi, İstanbul.

Ünal, Y., Kindap, T., Karaca, M., 2003. Redefining the climate zones of Turkey using cluster analysis. International Journal of Climatology, 23, 1045-1055.

WMO (World Meteorological Organization), 1983. Guide to Climatological Practices, 2nd Ed., Secretariat of World Meteorological Organization, Geneva, Switzerland, No: 100.

IDENTIFYING CLIMATE ZONES OF TURKEY BY HIERARCHICAL CLUSTERING METHOD

ABSTRACT

According to the Intergovernmental Panel on Climate Change (IPCC), the globe has been warming due to the increase in the amount of CO₂ and other greenhouse gases accumulated in the atmosphere. The effect of climate change result from global warming has been observed in different spatio-temporal forms such as rise in the sea level, melting glaciers, the change in the climate regions' boundaries, meteorological events and increase in disasters (Türkeş et al., 1999a; Türkeş, 2008). It is possible to examine the change in climate by analyzing atmospheric data. In this study, it is aimed to investigate the effect of climate change observed globally, on the boundaries of climate regions of Turkey. To accomplish this aim, after preprocessing, hierarchical cluster analysis is employed on the temperature (e.g. mean, minimum, maximum) and precipitation variables recorded in different stations of Turkey State Meteorology Service General Directorate in the period of 1950-2006. Since precipitation and temperature variables have different characteristics, multivariate cluster analysis is applied after all variables are standardized. In the study, four different hierarchical clustering methods with different distance measures are used. These are single linkage, average linkage, complete linkage and Ward method. To be able to test the stabilities of the clusters obtained using these methods, they are applied to the same variables obtained in two different time span: 1950-2006 and 1970-2006. It has been observed that Ward method creates more stable and meaningful clusters. Depending on the results of Ward method, it can be concluded that Turkey has seven climate regions with respect to temperature and precipitation characteristics. However, these regions are different than the ones known previously.

Keywords: Hierarchical clustering method, Subjective evaluation, Climate zones of Turkey.

EK: Ward Yöntemi ile Elde Edilen Yedi Bölgedeki İstasyonların Listesi

Ege	Akdeniz	Güneydoğu Anadolu	Doğu Anadolu	Doğu-İç Anadolu	Batı-İç Anadolu	Karadeniz
Göçeada	Bodrum	Siirt	Ardahan	Kastamonu	Artvin	Bartın
Bozcaada	Dalaman	K.Maraş	Erzurum	Çankırı	Edirne	Zonguldak
Çanakale	Fethiye	G.Antep	Kars	Çorum	Kırklareli	İnebolu
Edremit	Datca	Kilis	Ağrı	Gümüşhane	Bolu	Sinop
Ayvalık	Marmaris	Adıyaman	Uludağ	Bayburt	Merzifon	Samsun
Dikili	Antalya	Urfa	Horasan	Sivas	Amasya	Ordu
Akhisar	Alanya	Mardin	Sarıkamuş	Erzincan	Tokat	Giresun
Manisa	Anamur	Diyarbakır	Hınıs	Esenboğa	Iğdır	Trabzon
İzmir	Silifke	Batman	Malazgirt	Yozgat	Bilecik	Rize
Çeşme	Mersin	Ergeril	Özalp	Gemerek	Ankara	Hopa
Kuşadası	Adana	Cermik	Başkale	Van	Kırıkkale	Çorlu
Aydın	İskenderun	Siverek		Nevşehir	Kütahya	Tekirdağ
Denizli	Antakya	Nusaybin		Kayseri	Kırşehir	Kumköy
Muğla	Finike	Cizre		Muş	Tunceli	Sarıyer
Bergama	Kozan	Islahiye		Tatvan	Uşak	Göztepe
Bornova	Köyceğiz	Birecik		Niğde	Afyon	Kocaeli
Salihli	Karaisalı	Ceylanpınar		Hakkari	Cihanbeyli	Sakarya
Seferihisar	Manavgat	Akçakale		Ilgaz	Aksaray	Düzce
Ödemiş	Erdemli			Kızılcahamam	Malatya	Bandırma
Selçuk	Ceyhan			İspir	Elazığ	Osmangazi
Nazilli	Dört Yol			Oltu	Bingöl	Yalova
Milas	Gazipaşa			Ş.Karahisar	Burdur	Amastra
Yatağan	Yumurtalık			Suşehri	Akşehir	Bozkurt
	Karataş			Tortum	Isparta	Şile
	Samandağ			Keleş	Konya	Akçakoca
				Zara	Karaman	Bahçeköy
				Tercan	Ereğli	Bafra
				D.Beyazıt	Uzunköprü	Ünye
				Divriği	Lüleburgaz	Akçaabat
				Kulu	İpsala	Pazar
				Kaman	Tosya	Florya
				Boğazlıyan	Nallıhan	Geyve
				Kangal	Beypazarı	Gönen
				Solhan	Zile	
				Erciş	Dursunbeyli	
				Muradiye	Bozüyük	
				Pınarbaşı	Tavşanlı	
				Ahlat	Sivrihisar	
				Ürgüp	Polatlı	
				Develi	Çiçekdağı	
				Tomarza	Simav	
				Sarız	Emirdağ	

EK: Ward Yöntemi ile Elde Edilen Yedi Bölgedeki İstasyonların Listesi (Devam)

Ege	Akdeniz	Güneydoğu Anadolu	Doğu Anadolu	Doğu-İç Anadolu	Batı-İç Anadolu	Karadeniz
				Bitlis	Arapkir	
				Göksun	Çemişkezek	
				Afşin	Bolvadin	
				Elbistan	Yunak	
				Ulukışla	Keban	
				Hadım	Palu	
					Güney	
					Senirkent	
					Ilgın	
					Dinar	
					Uluborlu	
					Eğirdir	
					Acıpayam	
					Tefenni	
					Beyşehir	
					Seydişehir	
					Çumra	
					Karapınar	
					Korkuteli	
					Elmalı	