

**T.C.  
ONDOKUZ MAYIS ÜNİVERSİTESİ  
FEN BİLİMLERİ ENSTİTÜSÜ**

**AYRIK NOKTALARDA GÖZLENEN VERİLERİN ANALİZİ VE BİR  
UYGULAMA**

**YÜKSEK LİSANS TEZİ**

**Çağlar SÖZEN**

**İstatistik Anabilim Dalı**

**AĞUSTOS 2014  
SAMSUN**





**T.C.  
ONDOKUZ MAYIS ÜNİVERSİTESİ  
FEN BİLİMLERİ ENSTİTÜSÜ**



**İSTATİSTİK ANABİLİM DALI**

**AYRIK NOKTALARDA GÖZLENEN VERİLERİN ANALİZİ VE BİR  
UYGULAMA**

**YÜKSEK LİSANS TEZİ**

**Çağlar SÖZEN  
(12210124)**

**Tezin Savuma Tarihi : 18 Ağustos 2014**

**Tez Danışmanı : Doç. Dr. Yüksel ÖNER**



**Ondokuz Mayıs Üniversitesi Fen Bilimleri Enstitüsü**

**İstatistik Anabilim Dalında**

**Çağlar SÖZEN Tarafından Hazırlanan**

**AYRIK NOKTALARDA GÖZLENEN VERİLERİN ANALİZİ VE BİR  
UYGULAMA**

**başlıklı bu çalışma jürimiz tarafından 18/08/2014 tarihinde yapılan sınav ile  
YÜKSEK LİSANS tezi olarak kabul edilmiştir.**

**Başkan : Doç. Dr. Mehmet GÜRCAN** .....  
Fırat Üniversitesi

**Jüri Üyeleri : Doç. Dr. Vedat SAĞLAM** .....  
Ondokuz Mayıs Üniversitesi

**Doç. Dr. Yüksel ÖNER** .....  
Ondokuz Mayıs Üniversitesi

**....../....../2014**

**Prof. Dr. Hüseyin DEMİR**

Enstitü Müdürü



*Aileme,*



## **ÖNSÖZ**

Ondokuz Mayıs Üniversitesi Fen Bilimleri Enstitüsü İstatistik Anabilim Dalında hazırlamış olduğum yüksek lisans tez çalışmamda bilimsel desteğini esirgemeyen danışman hocam, Sayın Doç. Dr. Yüksel ÖNER' e, tezimin başından sonuna kadar göstermiş olduğu ilgiden dolayı saygılarımı sunarım.

Ağustos 2014

Çağlar SÖZEN



## İÇİNDEKİLER

### Sayfa

|                                                                            |      |
|----------------------------------------------------------------------------|------|
| ÖNSÖZ.....                                                                 | vii  |
| İÇİNDEKİLER .....                                                          | ix   |
| ÇİZELGELER LİSTESİ.....                                                    | xi   |
| ŞEKİLLER LİSTESİ.....                                                      | xiii |
| KISALTMALAR .....                                                          | xv   |
| AYRIK NOKTALARDA GÖZLENEN VERİLERİN ANALİZİ VE BİR                         |      |
| UYGULAMA.....                                                              | xvii |
| ÖZET.....                                                                  | xvii |
| DATA OBSERVED AT SEPERATED POINTS OF ANALYSES AND AN                       |      |
| APPLİCATION.....                                                           | xix  |
| ABSTRACT .....                                                             | xix  |
| 1. GİRİŞ .....                                                             | 1    |
| 2. GENEL BİLGİLER.....                                                     | 3    |
| 2.1 Fonksiyonel Verilerle Çok Değişkenli Verilerin Karşılaştırılması ..... | 3    |
| 2.2 Baz Fonksiyon Yaklaşımı.....                                           | 4    |
| 2.2.1 Fourier baz fonksiyon yaklaşımı .....                                | 5    |
| 2.2.2 B-splayn baz fonksiyon yaklaşımı.....                                | 6    |
| 2.2.3 Pürüzlü ceza yaklaşımı .....                                         | 7    |
| 2.3 Düzgünleştirme Parametresinin Belirlenmesi .....                       | 8    |
| 3. MATERYAL VE YÖNTEM.....                                                 | 11   |
| 3.1 Ana Bileşenler Analizi.....                                            | 11   |
| 3.2 Fonksiyonel Ana Bileşenler Analizi .....                               | 12   |
| 3.3 Düzgünleştirilmiş Fonksiyonel Ana Bileşenler Analizi .....             | 14   |
| 4. BULGULAR VE TARTIŞMA.....                                               | 17   |
| 5. SONUÇ VE ÖNERİLER.....                                                  | 33   |
| KAYNAKLAR .....                                                            | 35   |
| ÖZGEÇMİŞ.....                                                              | 37   |



## ÇİZELGELER LİSTESİ

|                                                              | <b><u>Sayfa</u></b> |
|--------------------------------------------------------------|---------------------|
| <b>Çizelge 4.1.</b> Sıcaklık için ana bileşen skorları ..... | 25                  |
| <b>Çizelge 4.2.</b> Yağış için ana bileşen skorları .....    | 29                  |



## ŞEKİLLER LİSTESİ

### Sayfa

|                                                                                                           |    |
|-----------------------------------------------------------------------------------------------------------|----|
| Şekil 4.1. 2012 yılı Karadeniz Bölgesi günlük ortalama sıcaklık verileri .....                            | 18 |
| Şekil 4.2. 2012 yılı Karadeniz Bölgesi günlük ortalama yağış verileri .....                               | 18 |
| Şekil 4.3. Sıcaklık verileri için 18 bireysel fonksiyon ve ortalama fonksiyonu .....                      | 19 |
| Şekil 4.4. Yağış verileri için 18 bireysel fonksiyon ve ortalama fonksiyonu.....                          | 20 |
| Şekil 4.5. Sıcaklığa ait $\psi$ değerleri karşısındaki GCV değerleri .....                                | 21 |
| Şekil 4.6. Yağışa ait $\psi$ değerleri karşısındaki GCV değerleri .....                                   | 22 |
| Şekil 4.7. Sıcaklık için birinci ana bileşen fonksiyonu ve ortalama fonksiyonu<br>karşılaştırılması ..... | 23 |
| Şekil 4.8. Sıcaklık için ikinci ana bileşen fonksiyonu ve ortalama fonksiyonu<br>karşılaştırılması .....  | 24 |
| Şekil 4.9. Sıcaklık için ana bileşen skorlarının dağılımı.....                                            | 26 |
| Şekil 4.10. Yağış için birinci ana bileşen fonksiyonu ve ortalama fonksiyonu<br>karşılaştırılması .....   | 27 |
| Şekil 4.11. Yağış için ikinci ana bileşen fonksiyonu ve ortalama fonksiyonu<br>karşılaştırılması.....     | 28 |
| Şekil 4.12. Yağış için ana bileşen skorlarının dağılımı.....                                              | 30 |



## KISALTMALAR

|                                         |                                                                 |
|-----------------------------------------|-----------------------------------------------------------------|
| <b>N</b>                                | : Örneklem Hacmi                                                |
| <b><math>n_i</math></b>                 | : i.inci örnek bireyi için yapılan ölçüm sayısı                 |
| <b>sd</b>                               | : Serbestlik Derecesi                                           |
| <b><math>t_{ij}</math></b>              | : i.örnek bireyi için j.inci ölçümlemenin alındığı nokta        |
| <b><math>y_{ij}</math></b>              | : y değişkeninin $t_{ij}$ noktasındaki değeri                   |
| <b><math>\bar{x}(t)</math></b>          | : Ortalama Fonksiyonu                                           |
| <b><math>Var_x(t)</math></b>            | : Varyans Fonksiyonu                                            |
| <b><math>Cov_x(t)</math></b>            | : Kovaryans Fonksiyonu                                          |
| <b><math>Cor_x(t)</math></b>            | : Korelasyon Fonksiyonu                                         |
| <b><math>\theta_i(t)</math></b>         | : i.inci baz fonksiyon                                          |
| <b><math>c_i</math></b>                 | : i.inci baz fonksiyonuna karşılık gelen katsayı                |
| <b>B</b>                                | : B-Splayn Baz Fonksiyonları                                    |
| <b>t</b>                                | : Bağımsız Değişken                                             |
| <b><math>\underline{t}</math></b>       | : Polinomun Derecesi                                            |
| <b><math>\{B_{i,d}\}_{i=1}^K</math></b> | : B-Splaynları oluşturmada kullanılan düğüm vektörü             |
| <b><math>CKT_\psi</math></b>            | : Cezalı Kareler Toplamı                                        |
| <b>R</b>                                | : Pürüzlü Ceza Matrisi                                          |
| <b><math>\underline{y}</math></b>       | : Gözlem Vektörü                                                |
| <b><math>\underline{c}</math></b>       | : k x 1 boyutlu katsayılar vektörü                              |
| <b><math>\phi</math></b>                | : $\theta_j(t)$ baz fonksiyonlarından oluşan bir set            |
| <b>Lx</b>                               | : Harmonik Hız Operatorü                                        |
| <b><math>PEN_L(x)</math></b>            | : Ceza Yaklaşımı                                                |
| <b>P</b>                                | : Ortogonal Matris                                              |
| <b>tr</b>                               | : Matrisin İzi                                                  |
| <b>SSE</b>                              | : Hata Kareler Toplamı                                          |
| <b>GCV</b>                              | : Genelleştirilmiş Çapraz Geçerlilik Yöntemi                    |
| <b><math>\Psi</math></b>                | : Düzgünleştirme Parametresi                                    |
| <b><math>\Sigma</math></b>              | : Varyans-Kovaryans Matrisi                                     |
| <b><math>\lambda</math></b>             | : Özvektörler Matrisi                                           |
| <b><math>\hat{c}</math></b>             | : Baz Fonksiyon Yaklaşımına Göre Pürüzlü Ceza Tahminleri        |
| <b><math>\gamma_{jk}</math></b>         | : j.inci ana bileşen için k.ıncı değişkenin önemi               |
| <b><math>\lambda_j</math></b>           | : Her bir bileşene yüklenebilir varyans miktarı                 |
| <b><math>x_i(t)</math></b>              | : Her bir birimden ortalama fonksiyon değerinin çıkarılmış hali |
| <b><math>V(Y_j)</math></b>              | : j.inci ana bileşen için varyans                               |
| <b><math>\underline{Y}_j</math></b>     | : P ortogonal matrisinin j.inci sütunu                          |
| <b>CABV</b>                             | : Cezalı Ana Bileşenler Varyansı                                |



# AYRIK NOKTALARDA GÖZLENEN VERİLERİN ANALİZİ VE BİR UYGULAMA

## ÖZET

İnceleme yapılan çalışma sahası genişledikçe ya da örneğe eklenen gözlem noktası sayısı arttıkça aslında ayrik noktalarda gözlenen verilerin altta yatan reel bir fonksiyondan örneklendiği varsayılmaktadır. Bu verileri analiz etmek için geliştirilen metot Fonksiyonel Veri Analizi ismiyle adlandırılmıştır. Fonksiyonel Veri Analizinin ilk adımı ayrik noktalardaki verileri reel fonksiyonlardan oluşan bir veri örneğine dönüştürmektir. Bunun için Baz Fonksiyon ve Pürüzlü Ceza Yaklaşımları kullanılmaktadır.

Bu çalışmada Karadeniz Bölgesinde bulunan 18 şehre ait günlük ortalama sıcaklık ve yağış verileri 65 ayrik noktada incelenmeye çalışılmıştır. Sıcaklık ve yağış verileri periyodik bir yapı izlediklerinden baz fonksiyon yaklaşımı olarak Fourier baz fonksiyonları ele alınmıştır. Daha sonra 65 ayrik noktada ele alınan günlük ortalama yağış ve sıcaklık verileri Baz fonksiyon ve Pürüzlü Ceza Yaklaşımıyla sürekli bir fonksiyon haline getirilmiştir. Her bir değişkene ait düzgünleştirme parametresinin belirlenmesinde Genelleştirilmiş Çapraz Geçerlilik Yöntemi kullanılmıştır. Pürüzlü Ceza Yöntemiyle elde edilen katsayılara Düzgünleştirilmiş Fonksiyonel Ana Bileşenler Analizi uygulanmıştır. Tüm fonksiyonlar ile ele alınıp incelenmesi zor gözüken yağış ve sıcaklık fonksiyonları arasındaki değişim ana bileşen fonksiyonu için oluşturulan ortalama fonksiyonuna uygun bir çarpanla çarpılmış, ana bileşen fonksiyonu eklenerek ve çıkarılarak elde edilen fonksiyonlarla ortalama fonksiyonu aynı grafik üzerinde çizdirilerek ortaya çıkartılmıştır.

**Anahtar Kelimeler:** Fonksiyonel Veri Analizi; Fonksiyonel Ana Bileşenler Analizi; Düzgünleştirilmiş Fonksiyonel Ana Bileşenler Analizi; Ayrik Noktalarda Gözlenen Veriler.



# DATA OBSERVED AT SEPERATED POINTS OF ANALYSES AND AN APPLICATION

## ABSTRACT

As the examined study field extends or the number of observation points added to the sample increases, it is actually assumed that the data observed at separated points are sampled from an underlying real function. The method developed to analyse this data is named as Functional Data Analysis. The first step of the Functional Data Analysis is to transform the data at seperated points into a data sample composed of real functions. To achieve this, Base Function and Roughness Penalty Approaches are used.

In this study, the average daily temperature and rain data of 18 cities in Black Sea Region at 65 separated points are analysed. Since temperature and rain data follow a periodical structure, Fourier base functions are used as the base function approach. Later, the average daily temperature and rain data at 65 separated points are transformed into a continuous function by Base Function and Roughness Penalty Approaches. Generalized Cross Validity Method is used to determine the smoothing parameter for each variable. Smoothed Functional Principle Components Analysis is carried out to the coefficients obtained by the Roughness Penalty Approach. Since it is difficult to deal with all functions, the change in the rain and temperature functions are multiplied by a multiplier appropriate to the average function developed for the Principle Components Function, then average function and the functions obtained from adding and subtracting Principle Components Function is generated by being drawn on the same graphic.

**Key Words:** Functional Data Analysis; Functional Principle Components Analysis; Smoothed Functional Principle Components Analysis; Data Observed at Seperated Points.



## 1. GİRİŞ

Her geçen gün teknolojinin gelişmesine paralel olarak artan veri toplama kapasiteleri sayesinde verileri klasik tek değişkenli ve çok değişkenli yöntemlerle analiz etmek ve yorumlamaya çalışmak gittikçe zorlaşmaktadır. İncelenen birey yada obje için inceleme yapılan değişken sayısı arttıkça bu gözlemleri birbiri ardına gelen sayı noktaları halinde incelemektense incelenen her birey yada obje için bireysel girişler olarak ele almak daha avantajlıdır. Bireysel girişler olarak ele almanın en kolay yolu ise her bireyin ya da objenin bir fonksiyonel gözlem olarak değerlendirilmesinden geçmektedir. Örneğin 10 birey için 100 farklı noktada inceleme yapıldığı varsayılırsa, elimizde 1000 farklı noktanın ya da her birey için 100 farklı noktanın incelenmesi problemini fonksiyonel gözlemler sayesinde 10 farklı fonksiyonun incelenmesine bırakır (Ramsay ve Silverman, 2005).

Günümüzde geniş hacimli örneklerle çalışabilmek ve bunların istatistiksel yöntemlerle analizini yapmak mümkündür. İncelenen çalışma sahası genişledikçe ya da başka bir ifadeyle örneğe dâhil edilen gözlem noktası sayısı arttıkça aslında ayırık noktalarda gözlenen bu verilerin altta yatan reel bir fonksiyondan örneklediği varsayılmaktadır. İşte bu yüzden bu gözlemler "Fonksiyonel Veriler" ismini almaktadır. Fonksiyonel verileri analiz etmek için geliştirilen istatistiksel metotlar ise ilk defa Ramsay ve Dalzell (1991) tarafından "Fonksiyonel Veri Analizi" terimiyle ifade edilmiştir.

Ayırık noktalarda gözlenen verileri fonksiyonel verilere döndürmenin bilinen en pratik yolları interpolasyon (interpolation) ya da düzgünleştirme (smoothing) yöntemleridir. Eğer gözlenen değerlerin hatasız oldukları kabul ediliyorsa fonksiyonları oluşturma süreci interpolasyon yöntemiyle yapılır. Eğer verilerde ortadan kaldırılması gereken bazı gözlemsel hatalar varsa, örneğin verilerin elde edilmesi deneysel bir sürece dayanıyorsa kesikli verilerden fonksiyonel verilere yapılan bu dönüşüm süreci düzgünleştirme yöntemi olarak adlandırılmıştır (Ramsay ve Silverman, 2005, s.11).

Fonksiyonel veri analizinin ilk adımı ve amacı;  $y_{ij}$  gözlem değerlerini herhangi bir  $t$  değeri için hesaplanması mümkün olan bir  $x_i(t)$  reel sürekli fonksiyonuna dönüştürmektir. Fonksiyonel veri analizinde, veri matrisinde yer alan  $i$ .inci gözlem aslında  $x_i(t)$  ( $i=1,2, \dots, N$ ) biçiminde reel sürekli bir fonksiyon olmasına rağmen verilerin ayrık noktalarda gözlemlendiği varsayılır. Fonksiyonel Veri Analizinde kullanılan genel notasyonunu şu şekilde özetleyebiliriz.

$N$ : Örneklem Hacmi

$n_i$ :  $i$ .inci örnek bireyi için yapılan ölçüm sayısı

$t_{ij}$ :  $i$ .örnek bireyi için  $j$ .inci ölçümlemenin alındığı nokta

$y_{ij}$ :  $y$  değişkeninin  $t_{ij}$  noktasındaki değeri

şeklinde verilebilir.

Fonksiyonel Veri Analizinin amaçları istatistiğin diğer dallarıyla benzerlik gösterir ve amaçlarını şu şekilde sıralayabiliriz.

- 1.İleri analizlere yardım edecek şekilde verileri sunmak.
- 2.Çeşitli özellikleri öne çıkaracak biçimde verileri sergilemek.
- 3.Veriler arasındaki önemli modelleri ve noktaları tanımlamak.
- 4.Bağımlı değişkendeki ilişkiyi bağımsız değişken yapısını kullanarak açıklamak.

Özellikle Ramsay ve Silverman ile beraber ivme kazanan Fonksiyonel Veri Analizinin önemi her geçen gün artmaktadır. Bunun temel nedeni interpolasyon yöntemi sayesinde kayıp verilerle baş edilebilmesi ve oluşturulan fonksiyonların istenilen dereceden türevlerinin incelenebilmesidir.

Fonksiyonel Veri Analizi ile ilgili daha detaylı bilgi için Ramsay ve Silverman (2005), Ulbricht (2004), Ramsay ve Dalzell (1991) çalışmaları incelenebilir.

Fonksiyonel Veri Analizi ile ilgili Türkiye’de de yapılan çalışmalarda mevcuttur. Keser (2010) tarafından Ege Bölgesine ait yağış verilerinin Fonksiyonel Veri Analiziyle incelenmesi, Keser ve Ertaş (2007) tarafından yapılanan İMKB verilerinin Düzgünleştirilmiş Fonksiyonel Ana Bileşenler Analiziyle incelenmesi, Keser (2008) tarafından yapılanan GSYİH verilerinin Fonksiyonel Ana Bileşenler Analiziyle incelenmesi bu çalışmalara örnek olarak verilebilir.

## 2. GENEL BİLGİLER

### 2.1 Fonksiyonel Verilerle Çok Değişkenli Verilerin Karşılaştırılması

Fonksiyonel Veri Analizinde i.inci birey için yapılan bu gözlemler,  $y_{i1}, y_{i2}, \dots, y_{in_i}$  şeklinde verilebilir. N hacimlik rastgele örneği oluşturan veri seti ise;

$$y_{ij}, i=1,2, \dots, N; j=1,2, \dots, n_i$$

şeklinde özetlenmesi mümkündür. Burada her bir gözlem için inceleme yapılan nokta sayısı eşit ve aynı noktalarda alınmak zorunluluğu olmadığından bu değişebilirliği göstermek üzere  $n_i$  notasyonu kullanılmıştır. Klasik çok değişkenli kavramlarda ise üzerinde çalışılan ölçümler klasik olarak şu şekilde ifade edilebilirler.

$$Y, [y_{ij}], i=1,2, \dots, N; j=1,2, \dots, p$$

biçiminde belirtilen veri matrisidir (Ramsay ve Silverman, 1997).

Her bir fonksiyonel gözlem aynı noktalarda gözlemlendiğinde çok değişkenli veriye benzese de fonksiyonel veri çok değişkenli veriden farklılık göstermektedir.

1. Bir fonksiyonel gözlemin boyutu  $n_i$  o kadar büyüktür ki bazen fonksiyonel gözlem sürekli bir fonksiyon olarak dikkate alınabilir. Buna ek olarak boyut genellikle örnek hacminden büyük olur. Başka bir ifadeyle incelenen birey sayısı gözlem sayısından büyük olur. Bu durumda  $n_i > N$  şeklinde ifade edilecektir. Böylece ilgili kovaryans matrisi tam ranklı olmadığından dolayı standart çok değişkenli yöntemler başarısızlığa uğrayabilir (Keser, 2008).

2. Bir fonksiyonel veri analizinde,  $(y_{ij}, t_{ij})$  biçiminde gözlenen veri temelde yer alan düzgün bir fonksiyondan örneklenmiştir. Çok değişkenli veri setinin gözlenen vektörü için böyle bir yapıdan söz edilemez (Keser, 2008).

3. Gözlem noktaları  $t_{ij}$  ( $i=1,2, \dots, n_i$ ) bir fonksiyonel gözlemden diğerine değişebileceğinden dolayı bu durumda doğrudan çok değişkenli yöntemlerin uygulanması mümkün olmayabilir. Buna ek olarak Fonksiyonel Veri Analizi sıklıkla zamana ait verilerle ilgilenmesine rağmen faaliyet alanı ve amaçları zaman serisi analizlerinden farklılık gösterir. Zaman serisi daha çok verilerin modellenmesi ve

gelecek gözlemlerin tahmin edilmesi için kullanılır. Fonksiyonel Veri Analizi tekniklerinde ise önemli olan şekillerdir (Costanzo, 2005).

Gözlemlerin fonksiyonlar olarak ele alınmasındaki en büyük etken fonksiyonların istenilen dereceden türevlerinin incelenebilmesidir. Ayrıca interpolasyon yoluyla kayıp verilerle de baş edilmektedir.

Buna ek olarak Fonksiyonel Veri Analizinde  $x_1(t)$ ,  $x_2(t)$ , ... ,  $x_N(t)$  şeklinde belirtilen N tane fonksiyon olduğu kabul edildiğinde klasik istatistikler benzer biçimde fonksiyonel verilere de uygulanabilmektedir. Özetlemek gerekirse;

Ortalama Fonksiyonu:

$$\bar{x}(t) = \frac{1}{N} \sum_{i=1}^N x_i(t_j), \quad (2.1)$$

Varyans Fonksiyonu:

$$Var_X(t) = \frac{1}{N-1} \sum_{i=1}^N (x_i(t) - \bar{x}(t))^2, \quad (2.2)$$

Kovaryans Fonksiyonu:

$$Cov_X(s, t) = \frac{1}{N-1} \sum_{i=1}^N \{x_i(s) - \bar{x}(s)\} \{x_i(t) - \bar{x}(t)\}, \quad (2.3)$$

Korelasyon Fonksiyonu:

$$Cor_X(s, t) = \frac{Cov_X(s, t)}{\sqrt{Var_X(s)Var_X(t)}}, \quad s, t \in [a, b] \quad (2.4)$$

## 2.2 Baz Fonksiyon Yaklaşımı

Fonksiyonel Veri Analizinde  $x_i(t)$  fonksiyonunu oluşturabilmek için esnek yöntemlere ihtiyaç duyulur. Bunun içinde k tane baz fonksiyonundan oluşan bir sistem seçilmelidir (Keser, 2010). Oluşturulmak istenen  $x_i(t)$  fonksiyonunu bu baz fonksiyonlarının ağırlıklandırılmış bir toplamı olarak yazmak olup şu şekilde ifade edilebilir.

$$x_i(t) = c_1\theta_1(t) + c_2\theta_2(t) + \dots + c_k\theta_k(t) \quad (2.5)$$

Bu ifadede yer alan;

$\theta_i(t)$  : i.inci baz fonksiyon

$c_i$  : Baz fonksiyonuna karşılık gelen katsayı

olarak ifade edilebilir.

Burada  $c_i$  ,  $i=1,2, \dots ,k$  katsayıları  $x_i(t)$  fonksiyonunun biçimini belirleyen katsayılar oldukları için bir nevi parametre olarak kabul edilirler.

Özetleyecek olursak;

$k$  herhangi bir pozitif tamsayı ve  $c_i \in \mathbb{R}$  olmak üzere  $x_{i(t)}$  fonksiyonunun,

$$x_i(t)=\sum_{i=1}^k c_i \theta_{i(t)} \quad (2.6)$$

biçiminde  $k$  tane baz fonksiyonunun ağırlıklandırılmış bir toplamı olarak tahmin edilmesini amaçlayan yaklaşım "Baz Fonksiyon Yaklaşımı" olarak ifade edilmektedir. Genel olarak en çok kullanılan baz fonksiyon yaklaşımları Fourier Baz ve B-Splayn Baz Fonksiyon yaklaşımlarıdır (Keser ve Ertaş, 2010).

Ne kadar çok baz fonksiyon içerilirse uyumu sağlayan fonksiyonda o kadar karmaşık bir hal alır. İdeal olan nispeten küçük bir  $k$  değeri kullanılarak mükemmel bir yaklaşım veren baz fonksiyonunun seçimidir (Ramsay ve Silverman, 1997, s.46).

Ramsay ve Silverman 20–30 arası baz fonksiyonunun genelde göze çarpan görünümleri özetlemek amacıyla yeterli olduğunu söylemiştir. Fonksiyon tahmini için iyi çalışan bir baz nispeten zayıf türev tahminleri verebilir. Bu nedenle bir baz için seçim kriterlerinden biriside yaklaşımın bir veya daha fazla türevinin hareket edip etmediğidir (Ramsay ve Silverman, 1997, s.46).

### 2.2.1 Fourier baz fonksiyon yaklaşımı

Trigonometrik fonksiyonlar genelde periyodik fonksiyonlara yaklaşım için kullanılmaktadır. Bir periyodik fonksiyon  $x(t)$  sonlu veya sonsuz Fourier Serisi cinsinden notasyonu şu şekilde ifade edilebilir.

$$x(t)= c_0+c_1 \sin wt+c_2 \cos wt+ c_3 \sin 2wt+ c_4 \cos 2wt+\dots , t \in T[a, b] \quad (2.7)$$

Burada bazlar periyodiktir. Eğer  $t_j$  değerleri ilgilenilen reel aralık olan  $T'$  de eşit ölçeklenmişse ve periyod  $T$  aralığının uzunluğuna eşitse bu durumda baz ortogonaldır. Ayrıca baz fonksiyonlarını uygun sabitlere bölerek bazlar ortonormal hale getirilebilir (Benko, 2004, s.13).

Bir diğer gösterim şeklide;

$K$  bir çift tam sayı olmak üzere, Fourier baz şu şekilde ifade edilebilir.

$$\begin{aligned}\theta_o(t) &= \frac{1}{\sqrt{T}} \\ \theta_{2r-1}(t) &= \frac{1}{\sqrt{\frac{T}{2}}} \sin rwt, r=1,2, \dots, \frac{K}{2} \\ \theta_{2r}(t) &= \frac{1}{\sqrt{\frac{T}{2}}} \cos rwt\end{aligned}\tag{2.8}$$

Ayrıca burada katsayılar uygun sabitler olarak bilinen  $\frac{1}{\sqrt{T}}$  ve  $\frac{1}{\sqrt{\frac{T}{2}}}$ 'ye bölünmesiyle

baz fonksiyonları ortonormal hale gelecektir. Başka bir ifadeyle,

$$\int_a^b \theta_{k_1}(t) \theta_{k_2}(t) dt = f(x) = \begin{cases} 0, & k_1 \neq k_2 \\ 1, & k_1 = k_2 \end{cases}\tag{2.9}$$

halini alacaktır (Ulbricht, 2004, s.19).

### 2.2.2 B-splayn baz fonksiyon yaklaşımı

B-Splayn Baz Fonksiyonları periyodik olmayan durumlarda tercih edilmektedir. B-Splaynların esnek ve nümerik olarak sağlam olması ve değerlerinin bant matris yapısında olması B-Splaynların yaygın kullanımını arttırmaktadır.

Polinomun derecesi herhangi bir  $d \in \mathbb{Z}^+$ ,  $j = 1, 2, \dots, n$  için  $t = t_j$ , B-Splaynları oluşturmada kullanılan  $\{B_{i,d}\}_{i=1}^K$ , B-Splaynları göstermek üzere Splaynlar B-Splaynların doğrusal kombinasyonu olarak şu şekilde ifade edilmektedir.

$$x_{i(t)} = \sum_{i=1}^K c_i B_{i,d}\tag{2.10}$$

Tanımlamalara bağlı kalarak, d'inci dereceden m'inci mertebeye sahip bir j.inci B-Splayn aşağıdaki tekrarlı ilişkiyle tanımlanabilir.

$$B_{j,d}(t) = \frac{t-t_j}{t_{j+d}-t_j} B_{j,d-1}(t) + \frac{t_{j+1+d}-t}{t_{j+1+d}-t_{j+1}} B_{j+1,d-1}(t)\tag{2.11}$$

B-Splaynların mertebeleri genellikle m ile gösterilmekte olup derecelerinin bir fazlasıdır. Burada t bağımsız değişken olarak ele alınmış olup tüm reel sayılar için tanımlıdır.

Düzgünleştirmenin kontrolü Splayn Teorisinin temel amacıdır. Tekrarlı ilişkiler sayesinde verilen bir düzgünleştirme seviyesinde ki B-Splayn baz fonksiyonları kullanılarak bir mertebe daha yüksek düzgünleştirme seviyesindeki yeni bir B-Splayn elde etmektir (Keser ve Ertaş, 2008).

Bir B-Splayn oluşturabilmek için düğümlerin sayısının en az  $d+2$  tane olması gerekir. Eğer düğüm sayısı tekrarlı değilse oluşturulacak baz sayısı fonksiyonel veri analizinde yaklaşımın doğruluğunu garantiye alacak büyüklükte olması bakımından aşağıda verilen eşitlik yardımıyla bulunur.

$$\text{Baz Fonksiyon Sayısı} = \text{Düğüm Sayısı} + \text{Mertebe} - 2 \quad (2.12)$$

B-Splayn baz fonksiyonlarını elde ettikten sonra  $x_i(t)$  fonksiyonlarını elde edebilmek için ikinci adım  $c_i$  katsayılarının Pürüzlü Ceza Yaklaşımı ile tahmin edilmesi olacaktır (Keser ve Ertaş, 2008).

### 2.2.3 Pürüzlü ceza yaklaşımı

Fonksiyonel Veri Analizinde verilere bir eğrinin uyumunu sağlarken tek amaç iyi bir uyum yakalamak değil aynı zamanda çok fazla iniş çıkış göstermeyen bir eğri tahmini yapmaktır. Fonksiyonel Veri Analizinde fonksiyonları düzgünleştirmek için kullanılan Pürüzlü Ceza Yaklaşımında temel amaç eğrinin pürüzlülüğünü ölçmektir. Aynı zamanda verilerin eğriye uyumu ve eğrinin pürüzlülüğü arasında bir uzlaşma sağlamaktır (Keser ve Ertaş, 2010).

Bu amaç istatistiğin temel prensiplerinden birisine karşılık gelir. Ortalama Karesel Hata, Sapmanın karesiyle Örneklem Varyansının toplamına eşittir. Örneklem Varyansını azaltabilmek için sapmadan biraz taviz verilebilir, buda eğriye düzgünleştirme yüklenmesinin temel nedenidir (Ramsay ve Silverman, 2005).

Pürüzlü Ceza Yaklaşımında, pürüzlü ceza tahminlerini elde ederken  $x(t)$  fonksiyonu  $T=[a,b]$  aralığında tanımlı türevi alınabilen bir fonksiyon olsun ve  $\psi > 0$  düzgünleştirme parametresi olarak kabul edelim. O halde Cezalı Kareler Toplamı ( $CKT_\psi$ ),

$$CKT_\psi = \sum_j (y_j - x(t_j))^2 + \psi \|Lx\|^2 \quad (2.13)$$

şeklinde ifade edilir. Baz fonksiyon yaklaşımına göre vektör bazında ise;

$$CKT_\psi = [\underline{y} - \underline{\phi c}]^T [\underline{y} - \underline{\phi c}] + \psi \underline{c}^T R \underline{c} \quad (2.14)$$

şeklinde ifade edilir.

Burada;

R: Pürüzlü Ceza Matrisi

$\underline{y}$ : Gözlem Vektörü

$\underline{c}$ :  $k \times 1$  boyutlu katsayılar vektörü

$\phi$ :  $\theta_j(t)$  baz fonksiyonlarından oluşan bir set

Periyodik verileri analiz etmek için bir eğrinin pürüzlülüğünün belirlenmesi  $x$  eğrilerinin ideal olarak belirli bir diferansiyel denklemiyle sağlanır ve bu denklemden sapmalar cezalandırılmak istenir. Eğer periyodik veriler analiz ediliyorsa,

$$Lx = D^3(x) + W^2 Dx \quad (2.15)$$

şeklinde harmonik hız operatörünü eğrinin pürüzlülüğünün değerlendirilmesinde kullanılması son derece normaldir. Böylelikle sıfır pürüzlülük  $x(t)$  eğrisinin,

$$x(t) = c_1 + c_2 \sin wt + c_3 \cos wt \quad (2.16)$$

formunda olduğunu kanıtlar. Başka bir ifadeyle  $x$  eğrisi gerçekten sinüsoidal ise  $Lx$  sıfıra eşit olacaktır.

Böylelikle periyodik veriler için pürüzlülüğü ölçmenin evrensel kabul edilen bir yolu,

$$PEN_L(x) = \int [(Lx)^2](t) dt = \|Lx\|^2 \quad (2.17)$$

şeklinde  $Lx$  'in integralinin karesi olarak tanımlanır (Ramsay ve Silverman, 2005). Bir başka deyişle sinüsoidal fonksiyondan sapmalar cezalandırılmak istenir.

$\psi \|Lx\|^2$  şeklinde gösterilen pürüzlü ceza terimi belirli bir eğrinin cezalı en küçük karelerinin sadece  $\sum_j (y_j - x(t_j))^2$  şeklinde Artık Kareler Toplamı ile ölçülen verilere uyum iyiliği ile değil aynı zamanda  $\|Lx\|^2$  şeklinde pürüzlülüğüne de bakarak karar verilmesini garanti altına alacaktır (Green ve Silverman, 1994, s.5).

$\psi$  şeklinde ki düzgünleştirme parametresi de, Artık Kareler Toplamıyla ölçülen verilerin eğriye uyumunu ölçer.  $\|Lx\|^2$  ise ölçülen  $x$  fonksiyonunun pürüzlülüğü arasındaki değişim oranını ölçer. Özetle  $\hat{x}_i(t)$  şeklindeki Cezalı Kareler Toplamını minimize eden eğri tahmini düzgünleştirme ve uyum iyiliği arasındaki en iyi uzlaşmadır (Ramsay ve Silverman, 2005).

### 2.3 Düzgünleştirme Parametresinin Belirlenmesi

$\psi$  düzgünleştirme parametresinin belirlenmesi ile ilgili iki farklı yaklaşım söz konusudur. Sübjektif yaklaşım düzgünleştirme parametresinin belirlenmesinde serbest seçim prosedürünün avantajlı bir özelliğini ele almaktadır. Düzgünleştirme

parametresini değiştirerek farklı ölçeklerde ortaya çıkan veri özellikleri incelenebilmektedir. Eğer tek bir tahmine ihtiyaç varsa sübjektif bir seçimle en iyi görünümü veren parametre değeri seçilebilir.  $\psi$  'nün seçimi araştırmacının ihtiyaçlarından ortaya çıkmaktadır ve aynı verilere farklı değerler uygulanan analizleri görmekten memnun olurlar (Green ve Silverman, 1994).

Buna ilaveten Ramsay ve Li'nin "Curve Registration (1998)" isimli çalışmalarında uygulamalarla ilgili kapsamlı bir araştırma sonucu düzgünleştirme parametresi  $\psi$  için  $10^{-2}, 10^{-3}$  ve  $10^{-4}$  değerlerinin iyi sonuçlar verdiklerini keşfetmişlerdir.

Bu alanda Ulbricht (2004) tarafından yapılan bir tez çalışmasında ise  $\psi$  düzgünleştirme parametresi, baz fonksiyon sayısının gözlem sayısından büyük olduğu durumlarda;

$$\psi = 10^{-4} [\text{tr}(B^T B) / \text{tr} R] \quad (2.18)$$

şeklinde ifade edilmiştir.

Otomatik bir yaklaşım olan Genelleştirilmiş Çapraz Geçerlilik Yöntemi de Fonksiyonel Veri Analizi çalışmalarında sıklıkla ele alınmıştır.

Craven ve Wahba (1979) tarafından geliştirilen Genelleştirilmiş Çapraz Geçerlilik Yöntemi Çapraz Geçerlilik prosedürünün daha basit bir versiyonudur.

$$GCV = \frac{n^{-1} SSE}{[n^{-1} \text{tr}(I - S_{\phi, \psi})]^2} \quad (2.19)$$

olarak ifade edilebilir.

Burada;

SSE: Hata Kareler Toplamı

$$S_{\phi, \psi} = \Phi(\Phi^T \Phi)^{-1} \Phi \quad (2.20)$$

şeklindeki düzgünleştirme matrisidir.

$df(\Psi) = \text{tr}(S_{\phi, \psi})$  olmak üzere Genelleştirilmiş Çapraz Geçerlilik Yöntemi,

$$GCV_{\psi} = \left( \frac{n}{n - sd(\Psi)} \right) \left( \frac{SSE}{n - sd(\Psi)} \right) \quad (2.21)$$

şeklinde de ifade edilebilmektedir. Amaç  $\Psi$  'ye göre Genelleştirilmiş Çapraz Geçerlilik Yönteminin minimizasyonudur (Ramsay ve Silverman, 2005). Burada tr matrisin izini ve sd de serbestlik derecesini belirtmektedir.

Son olarak tüm gerekli hesaplamalar yapıldıktan sonra  $CKT_\psi$  'yi minimize eden baz fonksiyon yaklaşımına göre Pürüzlü Ceza Tahminleri,

$$\underline{\hat{c}} = (\Phi^T \Phi + \Psi R)^{-1} \Phi^T \underline{y} \quad (2.22)$$

şeklinde tanımlanır.

### 3. MATERYAL VE YÖNTEM

Karadeniz Bölgesinde bulunan 18 şehrin 2012 yılına ait günlük ortalama sıcaklık ve günlük ortalama yağış verileri Samsun Meteoroloji 10. Bölge Müdürlüğünden temin edilmiştir. Bu iller sırasıyla Samsun, Ordu, Giresun, Çorum, Amasya, Tokat, Trabzon, Rize, Artvin, Gümüşhane, Bayburt, Bartın, Zonguldak, Sinop, Bolu, Düzce, Kastamonu ve Karabük şeklinde sıralanabilir. Bu verilerin analiz edilmesinde yararlanılan ve kullanılan yöntemler Ana Bileşenler Analizi, Fonksiyonel Ana Bileşenler Analizi ve Düzgünleştirilmiş Fonksiyonel Ana Bileşenler Analizidir.

#### 3.1 Ana Bileşenler Analizi

Bir diğer adı Karhunen - Loeve metodu olan Ana Bileşenler Analizinde amaç bireylerin ve ölçümde alınan noktaların sayısı arttıkça oluşan kompleks yapıların anlaşılmasıyla ilgili önemli bir probleme etkili bir çözüm sağlamasıdır. Veri Karmaşıklığını gidermek üzere bireyler arasındaki değişimin önemli noktalarını tanımlamak için bir boyut indirgeme yöntemi olan Ana Bileşenler Analizinin kullanılması önerilir. Ana Bileşenler Analizi genel anlamda sistemde olması beklenen ve önceden fark edilmeyen ilişkilerin ortaya çıkmasını sağlar. Bu nedenden dolayı Ana Bileşenler Analizi Fonksiyonel Veri Analizi için ele alınan önemli bir tekniktir (Keser, 2010, s.48).

Ana Bileşenler Analizi varyans - kovaryans yapısına daha aydınlatıcı bir şekilde bakılmasını sağlar (Ramsay ve Silverman, 1997, s.85). Sahip olduğu ana fikir varyans - kovaryans yapısını, orijinal değişkenlerin maksimum varyansa sahip doğrusal kombinasyonu aracılığıyla açıklamaya çalışarak veri setinin boyutunun indirgenmesini sağlamaktır (Mardia, 1989, s.213). Bu yöntemde karşılıklı bağımlılık yapısını gösteren, ölçüm sayısı  $n$  olan  $p$  adet değişken; doğrusal, dikey ve birbirinden bağımsız olma özelliklerini taşıyan  $k$  tane yeni değişkene dönüştürülmektedir. Yüksek boyutlu verilerdeki genel özellikleri bularak boyut sayısının azaltılmasını sağlar. Boyut azaltılmasıyla bazı özelliklerin kaybedileceği kesindir fakat amaçlanan bu kaybolan özelliklerin veri seti hakkında çok az bilgi içeriyor olmasıdır (Yılmaz vd., 2010, s.4).

Klasik Ana Bileşenler Analizine  $Cov(\underline{x}) = \Sigma$ ,  $P^T \Sigma P = \Lambda$  olmak üzere,  $\underline{x}$  rastgele vektörünün j.inci ana bileşeni,

$$Y_j = \underline{\gamma}_j(\underline{x} - \underline{\mu}), \quad j=1,2, \dots, p \quad (3.1)$$

olarak tanımlanmaktadır. Burada  $\underline{\gamma}_j$ :  $p \times 1$  boyutlu ve sütunları  $\Sigma$  matrisinin  $\lambda_j$  özdeğerlerine karşılık gelen özvektör olup,  $P$  ortogonal matrisinin j.inci sütunudur. Ayrıca j,  $k=1,2, \dots, p$  olmak üzere;  $\gamma_{jk} \in \underline{\gamma}_j$  olup j.inci ana bileşen için k.inci değişkenin önemini belirten katsayıdır.

Ana Bileşenler Analizinde j.inci doğrusal bileşenin varyansı  $V(Y_j)$ ,

$$V(Y_j) = \underline{\gamma}_j^T \Sigma \underline{\gamma}_j = \lambda_j \quad (3.2)$$

şeklinde ve

$$\begin{aligned} 1) & \langle \underline{\gamma}_j, \underline{\gamma}_m \rangle = \underline{\gamma}_j^T \underline{\gamma}_m, & 1 & \quad j = m \text{ ise}; j, m = \overline{1, p} \\ 2) & \langle \underline{\gamma}_j, \underline{\gamma}_m \rangle = \underline{\gamma}_j^T \underline{\gamma}_m, & 0 & \quad j \neq m \text{ ise}; j, m = \overline{1, p} \end{aligned} \quad (3.3)$$

kısıtları altında maksimum varyansa sahip olup  $Cov(\underline{\gamma}_j, \underline{\gamma}_m) = \underline{\gamma}_j^T \Sigma \underline{\gamma}_m = 0$  olmakla beraber ilk adımda hesaplanıp maksimize edilen  $V(Y_j)$  değişimin en güçlü ve en önemli noktasıdır.

Doğrusal bileşenin varyansının maksimum yapılması problemi Ana Bileşenler Analizinde

$$\Sigma \underline{\gamma}_j = \lambda_j \underline{\gamma}_j \quad (3.4)$$

şeklindeki özdeğer - özvektör ayrışımını gerektirmektedir.

### 3.2 Fonksiyonel Ana Bileşenler Analizi

Fonksiyonel veriler için kullanılan Fonksiyonel Ana Bileşenler Analizinde temel amaç çok değişkenli veriler için uygulanan Ana Bileşenler Analizi ile benzer olup verilerdeki değişimleri etkili bir biçimde tanımlayacak birkaç ortogonal fonksiyon elde etmektir. Aralarındaki fark  $\gamma_j$  biçiminde belirtilen ortogonal fonksiyonlar olan ana bileşen ağırlıkları zamanın ya da ilgilenilen başka bir değişkenin fonksiyonları olmasıdır (Keser, 2007).

Fonksiyonel Ana Bileşen Analizinde çok değişkenli analizde yer alan değişken değerleri kesikli indeks j sürekli indeks t ile  $\gamma_{jk}$  biçimindeki ağırlıklar  $\gamma_j(t)$  fonksiyon

değerleri ile ve vektör uzayındaki skaler çarpım fonksiyon uzayındaki skaler çarpımlarla yer değiştirir. Toplam indisi de integrale dönüşür (Keser, 2008, s.919).

Fonksiyonel Ana Bileşenler Analizinde ana bileşen fonksiyonu;

$$\|\gamma_j\|^2 = \int \gamma_j(t)^2 dt = 1 \quad (3.5)$$

kısıtı altında doğrusal bileşenin varyansı  $V(Y_j)$ ,

$$\text{Var}(Y_j) = \text{Var}\langle \gamma_j, x - E(x) \rangle = \iint \gamma_j(s) \text{cov}(s,t) \gamma_j(t) ds dt = \lambda_j \quad (3.6)$$

şeklinde;

$$\begin{aligned} 1) \langle \gamma_j, \gamma_m \rangle &= \int \gamma_j(t) \gamma_m(t) dt, & 1 & \quad j = m \text{ ise}; j, m = \overline{1, n} \\ 2) \langle \gamma_j, \gamma_m \rangle &= \int \gamma_j(t) \gamma_m(t) dt, & 0 & \quad j = m \text{ ise}; j, m = \overline{1, n} \end{aligned} \quad (3.7)$$

kısıtları altında maksimum varyansa sahip olup  $\lambda_j$  biçimindeki ağırlıklar her bir bileşene yüklenebilir varyans miktarını gösterir.

$\text{Var}(Y_j)$  ifadesinde yer alan  $\text{cov}(s, t)$ ,

$$\text{cov}(s, t) = N^{-1} \{x_i(s) x_i(t)\} \quad (3.8)$$

şeklinde dir.

$x_i(t)$  ise her bir birimden ortalama fonksiyon değerinin çıkarılmış hali olarak ifade edilmektedir.

Buradaki amaç kovaryans fonksiyonunun ortogonal ayrışımını sağlayarak fonksiyonel değişimi dominant bileşenlerine ayırmaktır.

Fonksiyonel Ana Bileşenler Analizi kendi ortalamasından ayrılan fonksiyonel veri setlerini açıklar. Bu nedenle uygulamadan önce her bir değişken için karşılık gelen ortalamalar çıkartılmaktadır. Bu yapıldığında ana bileşen skorunun ortalama karesini maksimize etmek kendi varyansını maksimize etmeye karşılık gelecektir (Keser ve Ertaş, 2008, s.920).

Fonksiyonel Ana Bileşen Analizinde doğrusal bileşenin varyansının maksimum yapılması;

$$\int \text{cov}(s, t) \gamma(t) = \lambda \gamma(s) \quad (3.9)$$

şeklindeki özdeğer - özfonksiyon eşitliğinin çözümünü gerektirir. Artık fonksiyonel olarak ele alındığında öz denklemi sağlayan özdeğer - özvektör değil özdeğer - özfonksiyon çiftleridir.  $\gamma_1, \gamma_2, \dots, \gamma_{N-1}$  şeklindeki özfonksiyonlar karşılık geldikleri

$\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_{N-1}$  özdeğerlerine göre sıralanırlar. Klasik çok değişkenli analizdeki felsefeye paralel olarak j.inci bileşenin varyansı j.inci özdeğere eşit olmaktadır (Keser, 2010, s.49).

### 3.3 Düzgünleştirilmiş Fonksiyonel Ana Bileşenler Analizi

Fonksiyonel Ana Bileşenler Analiziyle elde edilen ana bileşen ağırlıkları pürüzlü olabilir. Bu pürüzlülük örneklem varyansından ya da bazın esnekliğinden kaynaklanabilir. Daha iyi yorumlanabilir sonuçlar için özfonksiyonların düzgünleştirilmesi gerekir (Keser, 2010, s.49).

Fonksiyonel Ana Bileşenler Analizini düzgünleştirmek sadece klasik ana bileşenler analiziyle elde edilen bileşenleri düzgünleştirme anlamına gelmez. Düzgünleştirme Ana Bileşenleri orijinal tanımın içine dahil edilmektedir.

Düzgünleştirilmiş Fonksiyonel Ana Bileşenler Analizinde,

$$\|\gamma_j\|^2 = \int \gamma_j(t)^2 dt = 1$$

şeklindeki kısıt  $\gamma$  ' nün pürüzlülüğünü de dikkate alan,

$$\int \gamma_j(t)^2 dt + \psi PEN_L(\gamma) = 1 \quad (3.10)$$

kısıtı ile yer değiştirmektedir.

Böylece  $Var(Y_j)$  ,

$$\int \gamma_j(t)^2 dt + \psi PEN_L(\gamma) = 1$$

kısıtına bölünerek Cezalı Ana Bileşenler Varyansı (CABV),

$$CABV = Var(Y_j) = \frac{\iint \gamma_j(s) cov(s,t) \gamma_j(t) ds dt}{\int \gamma_j(t)^2 dt + \psi PEN_L(\gamma)} \quad (3.11)$$

şeklinde elde edilecektir (Silverman, 1996).

Pürüzlü Ceza ikinci, üçüncü ve daha yüksek dereceli düzgünleştirilmiş ana bileşenlere ekstra kısıtlar ekler. j.inci ana bileşen fonksiyonu,

$$\int \underline{\gamma}_j(t) \underline{\gamma}_m(t) dt + \Psi \int (L \gamma_j(t)) (L \gamma_m(t)) dt = 0 \quad j \neq m \quad (3.12)$$

ilave kısıtı ile maksimize edilmesi sağlanır.

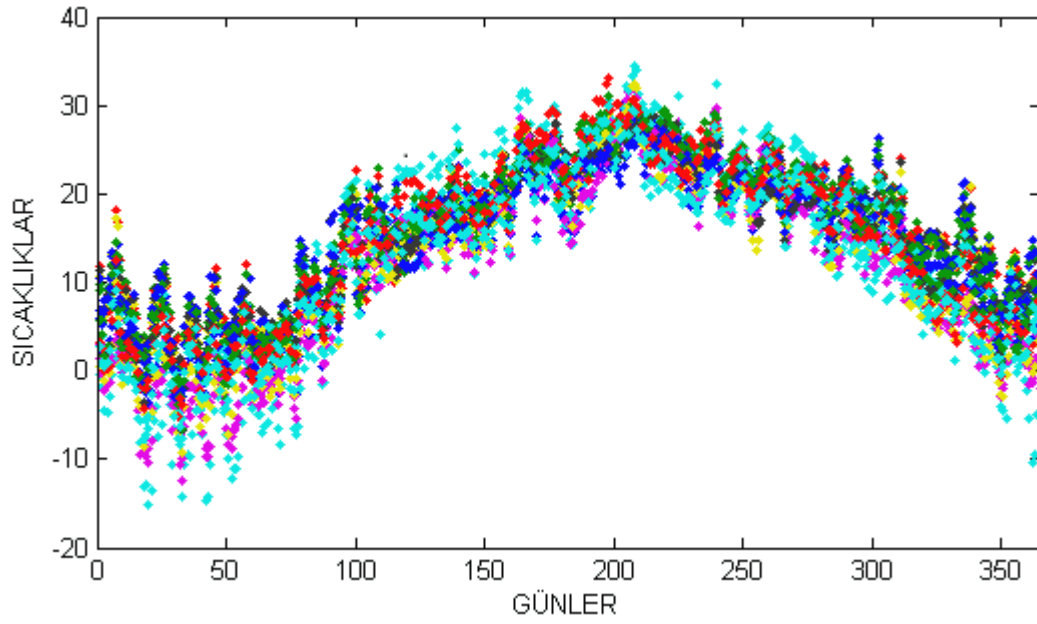
Ana Bileşenler Analizine düzgünleştirme yüklense bile ana bileşen ağırlık fonksiyonlarının net bir şekilde yorumlanması her zaman mümkün olmayabilir. Bu durumlarda sonuçların yorumlanmasına yardımcı olacak ilk yaklaşım ana bileşen skorlarının işaretlenmesidir. Diğer ikinci yaklaşımda ise ağırlık fonksiyonunun sabit bir çarpanıyla ortalama fonksiyonunun karşılaştırılması gerekir (Silverman, 1995). Burada ortalama fonksiyona uygun bir çarpanla çarpılmış, ana bileşen fonksiyonu eklenerek veya çıkarılarak elde edilen fonksiyonlarla ortalama fonksiyonu aynı grafik üzerinde çizdirilerek verilerin yapısıyla ilgili çeşitli yorumlar yapılabilir (Keser, 2010, s.50).



#### 4. BULGULAR VE TARTIŞMA

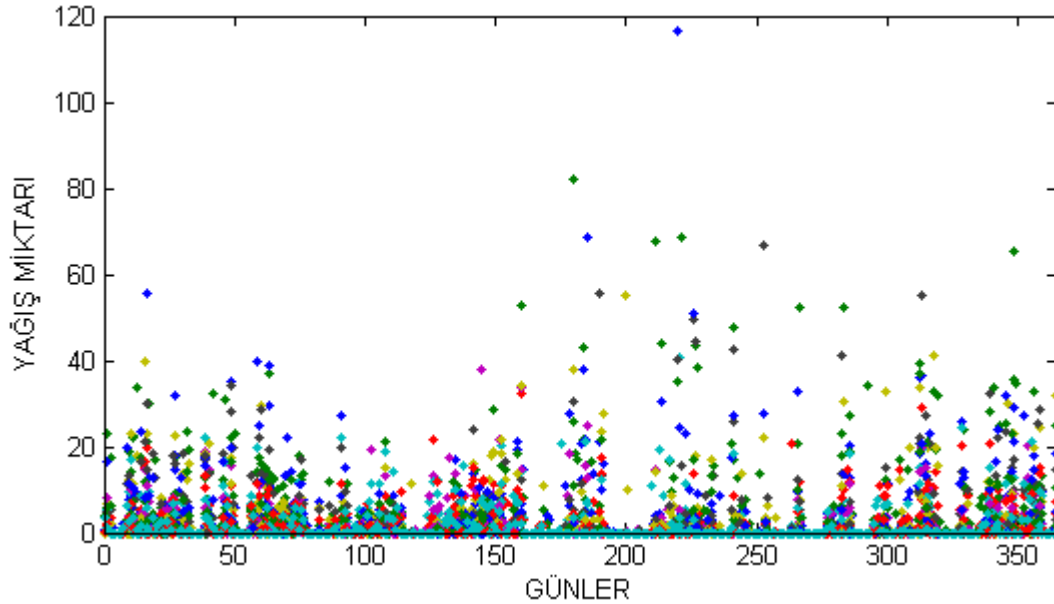
Öncelikle günlük ortalama sıcaklık ve yağış verilerini incelememiz için yeterli olan baz fonksiyon sayısı belirlenmiştir. Veriler hakkında gerekli bilgilere sahip olabilmemiz için 65 baz fonksiyonunun yeterli olabileceği düşünülmüştür. 18 şehre ait 65 ayık noktada ölçülen ortalama yağış ve sıcaklık verileri Baz fonksiyon ve Pürüzlü Ceza yaklaşımlarıyla sürekli bir fonksiyon haline getirilmiştir. Her bir değişken için ayrı ayrı oluşturulan 18 bireysel fonksiyon ve ortalama fonksiyonları incelenmiştir. Her bir değişkene ait Düzgünleştirme Parametrelerinin belirlenmesinde Genelleştirilmiş Çapraz Geçerlilik Yöntemi kullanılmıştır. Pürüzlü Ceza Yaklaşımıyla elde edilen katsayılara Düzgünleştirilmiş Fonksiyonel Ana Bileşenler Analizi uygulanmıştır. Tüm fonksiyonlar ile ele alınıp incelenmesi çok zor görünen bireysel yağış ve sıcaklık fonksiyonları arasındaki değişim ana bileşen fonksiyonu için oluşturulan ortalama fonksiyonuna uygun bir çarpanla çarpılmış, ana bileşen fonksiyonu eklenerek ve çıkarılarak elde edilen fonksiyonlarla ortalama fonksiyonu aynı grafik üzerinde çizdirilerek ortaya çıkarılmaya çalışılmıştır. Buna ek olarak ana bileşen skorları da incelenerek yapılan yorumların doğru olduğu kanıtlanmaya çalışılmıştır. Bu uygulamada Matlab R2012a paket programı kullanılmıştır.

Şekil 4.1 ve Şekil 4.2 18 şehre ait günlük ortalama sıcaklık ve yağış verilerinin fonksiyonel verilere dönüştürülmeden önceki halini ifade etmektedir.



**Şekil 4.1.** 2012 yılı Karadeniz Bölgesi günlük ortalama sıcaklık verileri

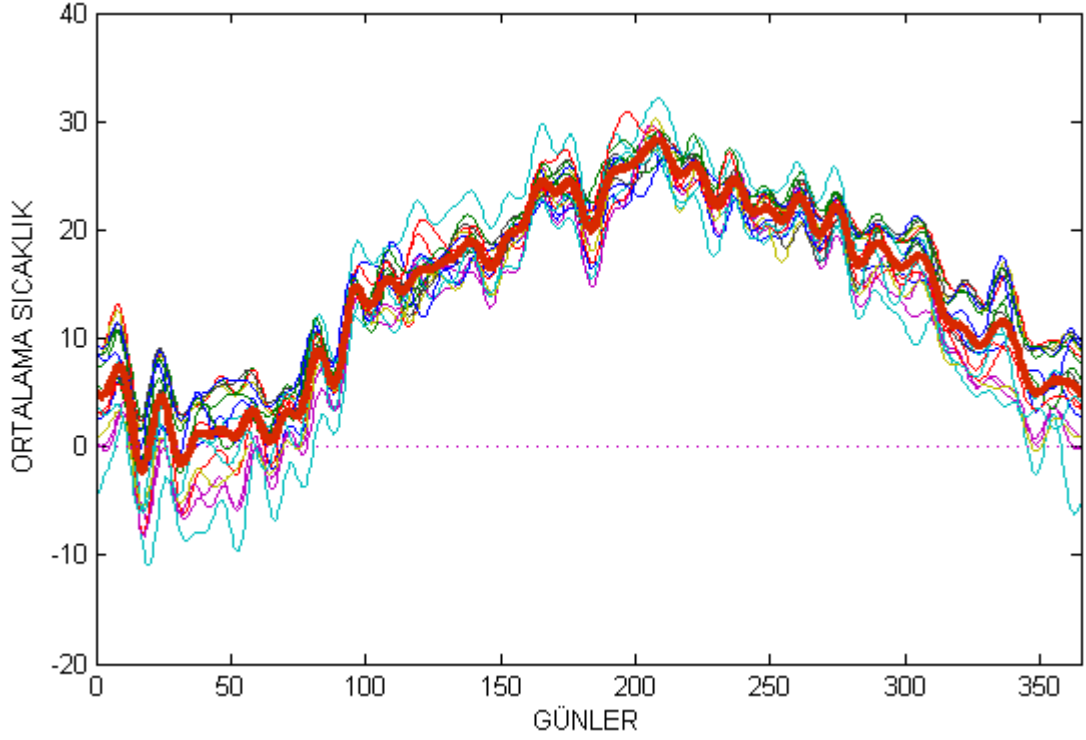
Burada  $(365 \times 18 = 6570)$  gözlem noktası bulunmaktadır. Şekilde Karadeniz Bölgesi için günlük ortalama sıcaklıkların bireysel seyirlerini çıkarmak Şekil 4.3 e göre daha zordur. İnceleme yapılan gözlem noktası sayısı arttıkça bu zorluk seviyesinin daha fazla artacağı rahatlıkla söylenebilir.



**Şekil 4.2.** 2012 yılı Karadeniz Bölgesi günlük ortalama yağış verileri

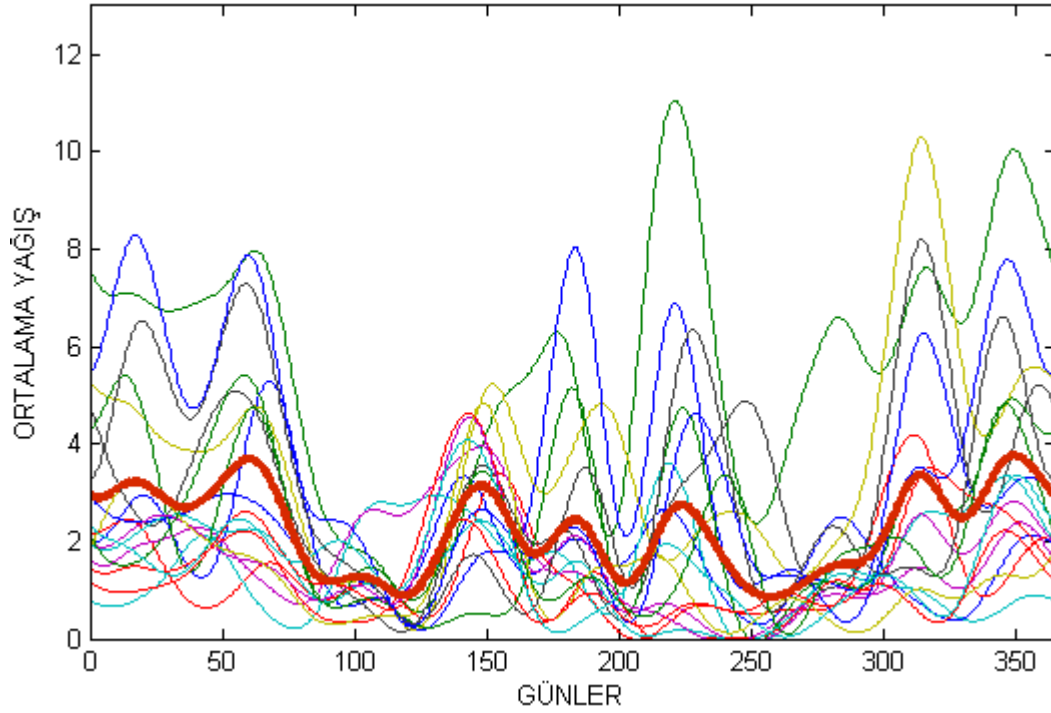
Burada  $(365 \times 18 = 6570)$  gözlem bulunmaktadır. Şekilde Karadeniz Bölgesi için günlük ortalama yağışların bireysel seyirlerini çıkarmak Şekil 4.4 e göre daha güçtür.

Örneğe dâhil edilen gözlem noktası sayısı arttıkça verilerin analiz edilmesinin daha güç olacaktır.



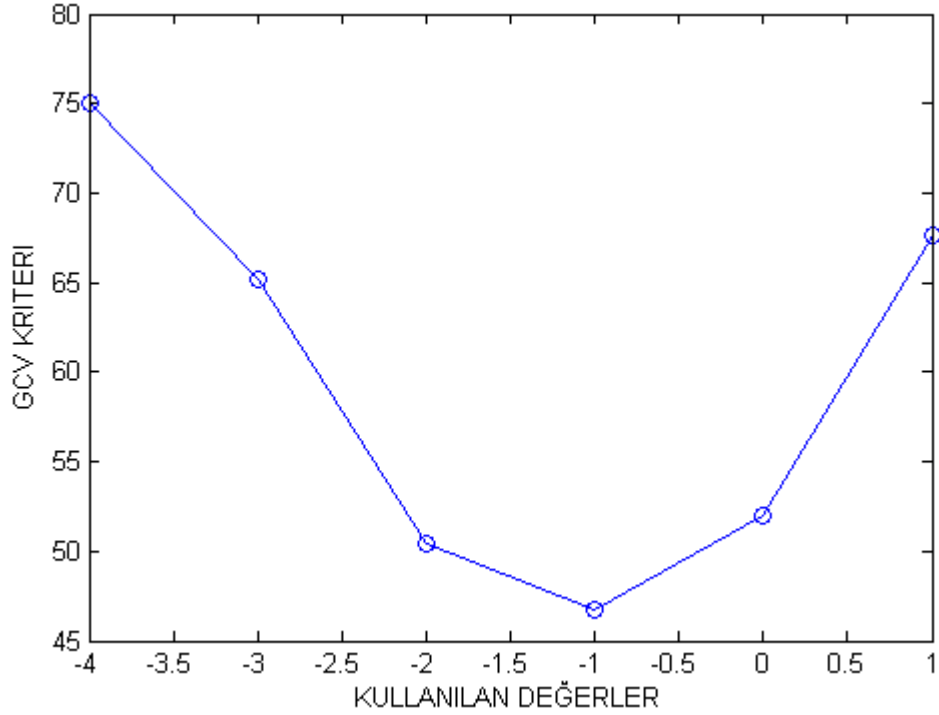
**Şekil 4.3.** Sıcaklık verileri için 18 bireysel fonksiyon ve ortalama fonksiyonu

Pürüzlü Ceza Yöntemine göre 18 bireysel sıcaklık fonksiyonunu elde ettiğimizde tüm fonksiyonların bireysel davranışları ve genel seyirleri nispeten rahatlıkla gözlenebilmektedir. Şekilden genel olarak fonksiyonların sinüsoidal bir yapı gösterdikleri söylenebilir. Ancak incelenen birey(şehir) sayısı arttıkça fonksiyonların genel seyirleri hakkında bilgi edinilmesi de karmaşık bir yapı alacak ve zorlaşacaktır. Bu durum periyodik olmayan veriler için daha karmaşık bir yapı olabilir. Şekilde kalın çizgiyle ortalama fonksiyonu çizdirilmiştir. Şekilden 2012 yılındaki sıcaklık yoğunluğunun özellikle yaz aylarına denk geldiği görülmektedir. 2012 yılındaki en düşük sıcaklıkların kış aylarına geldiği rahatlıkla görülebilmektedir. Bunlara ek olarak 200. günden sonra sıcaklıkların bir miktar daha arttığı, sıcaklıklar tepe noktasına ulaştıktan sonra sıcaklıkların azalma eğiliminde olduğu görülmektedir. Buralar da bile ortalama fonksiyonunu yukarı doğru çeken pozitif yönde uç bireysel sıcaklık değerleri mevcuttur.



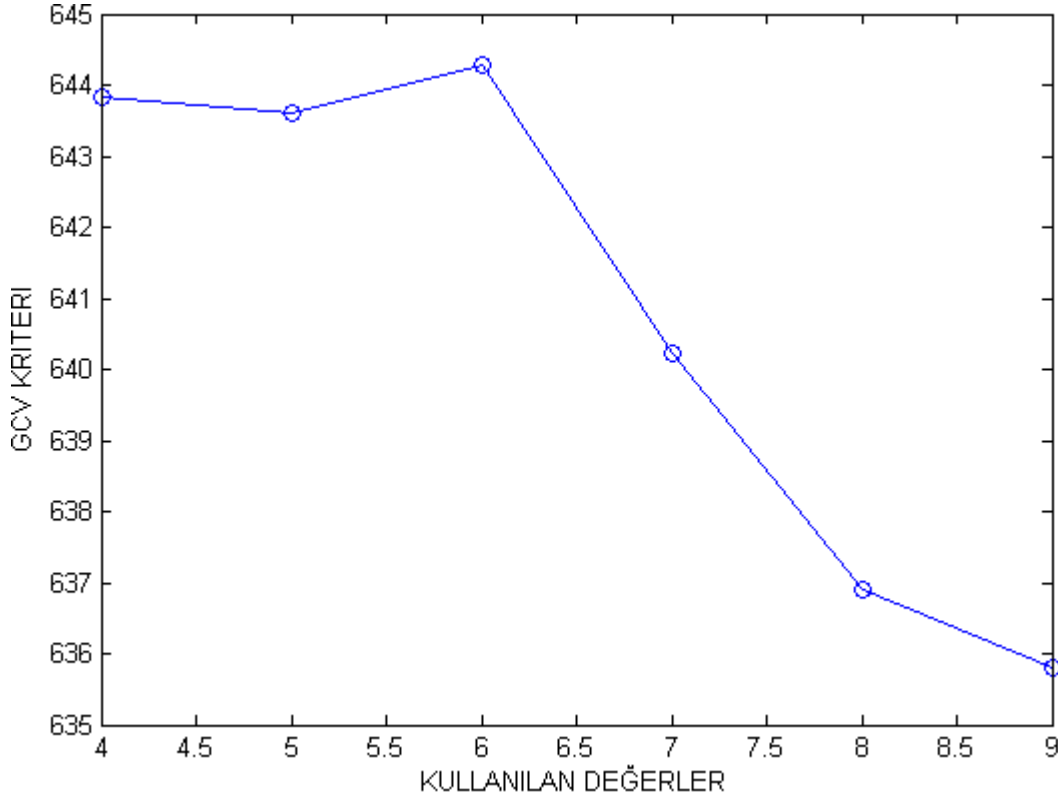
**ŞEKİL 4.4.** Yağış verileri için 18 bireysel fonksiyon ve ortalama fonksiyonu

Pürüzlü Ceza Yöntemine göre 18 bireysel yağış fonksiyonunu elde ettiğimizde tüm fonksiyonların bireysel davranışları ve genel seyirleri daha rahat gözlenmektedir. Şekilden genel olarak fonksiyonların sinusoidal bir yapı gösterdikleri söylenebilir. Ancak inceleme yapılan gözlem noktası sayısı arttıkça fonksiyonların genel seyirleri hakkında bilgi edinilmesi gittikçe zorlaşacaktır. Bu durum periyodik olmayan veriler için daha karmaşık bir hal alabilir. Şekilde kalın çizgiyle ortalama fonksiyonu çizdirilmiştir. Ortalama fonksiyonu sayesinde fonksiyonların genel seyirleri hakkında daha fazla bilgi sahibi olmamız mümkündür. Şekilden 2012 yılı için yağışların önemli bir bölümünün kış aylarına rastladığı görülmektedir. Kış aylarından sonra yağışların azalma eğiliminde olduğu anlaşılmaktadır. Hatta kış aylarında bile ortalama fonksiyonunu aşağıya çeken negatif yönde uç bireysel yağış değerleri tespit edilmiştir. 150. ve 250. günler arası yağışların azalma eğiliminde olduğu anlaşılmaktadır. Hatta buralarda bile ortalama fonksiyonunu yukarı çeken pozitif yönde uç bireysel yağış değerleri mevcuttur. Genellikle bu uç bireysel fonksiyonlara 150. ve 250. günler arası rastlanılmıştır.



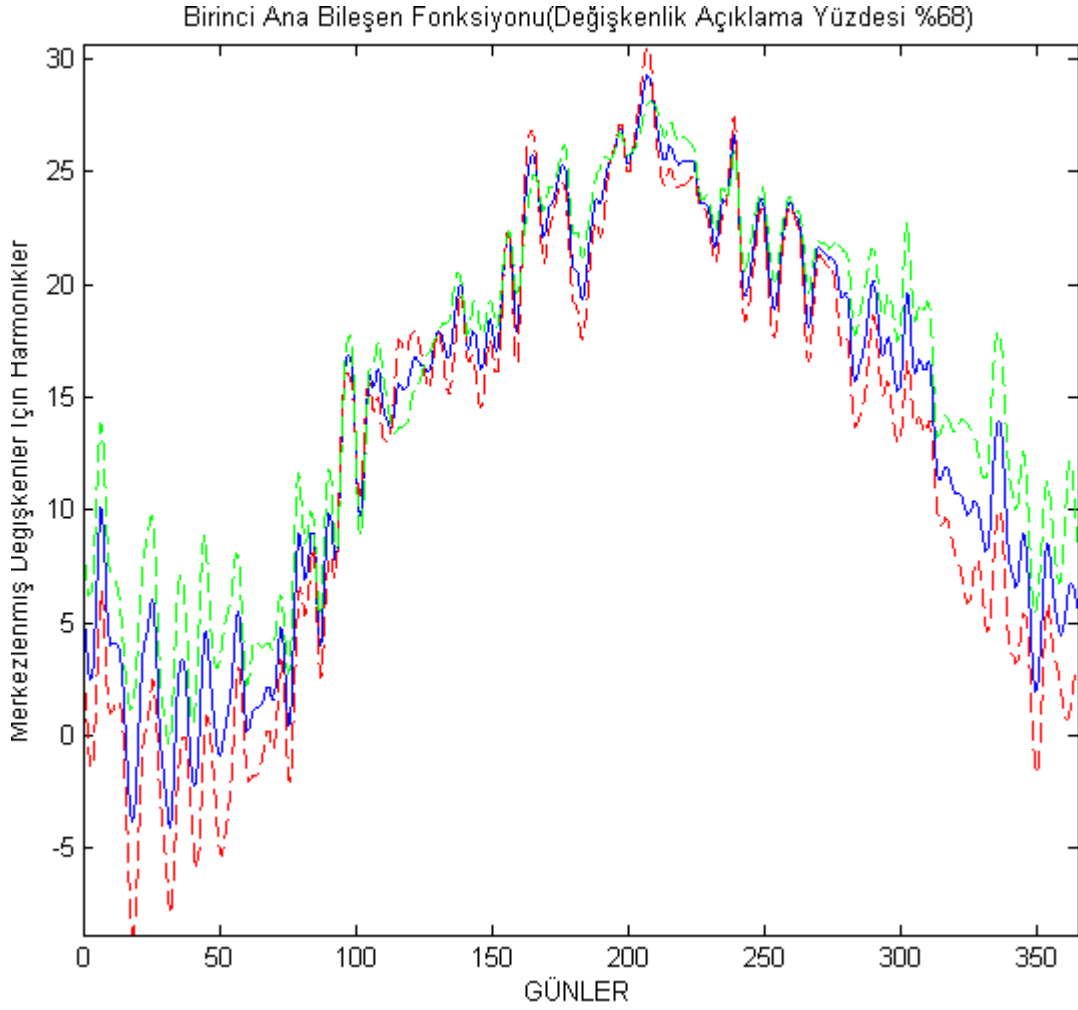
**Şekil 4.5.** Sıcaklığa ait farklı  $\psi$  değerleri karşısındaki GCV değerleri

Sıcaklık verileri için Pürüzlü Ceza Yaklaşımında  $\psi$  düzgünleştirme parametresi belirlenirken Genelleştirilmiş Çapraz Geçerlilik yöntemi tercih edilmiştir. Genelleştirilmiş Çapraz Geçerlilik Yöntemine göre düzgünleştirme parametresinin bulunmasında 365 baz fonksiyonu kullanılmıştır. Farklı düzgünleştirme parametresi değerleri için simülasyonlar yapılmış ve minimum GCV değerini veren  $\psi$  değeri düzgünleştirme parametresi olarak belirlenmiştir. Şekil 4.5 de yapılan simülasyon sonuçları gösterilmektedir. Bu çalışmada  $\psi$  şeklindeki düzgünleştirme parametresine  $10^{-4}$  ile  $10^1$  arasında farklı değerler verilmiştir. Minimum GCV değeri -1 değerine karşılık gelmektedir. Başlangıç noktası olarak  $10^{-4}$  değerinin alınmasının nedeni yapılan araştırmalar sonucu bu değer iyi sonuçlar verdiğinin tespit edilmiş olmasıdır.



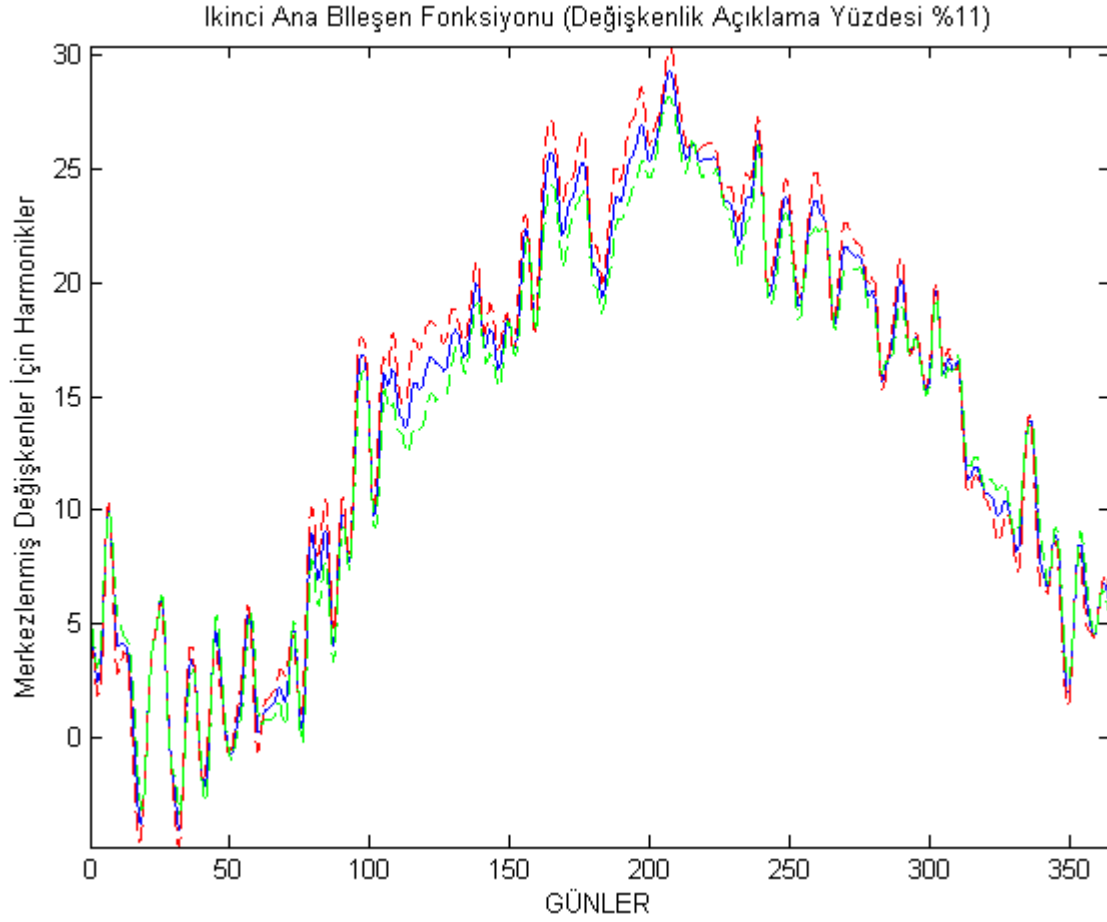
**Şekil 4.6.** Yağışa ait farklı  $\psi$  değerleri karşısındaki GCV değerleri

Yağış verileri için Pürüzlü Ceza Yaklaşımında  $\psi$  düzgünleştirme parametresini belirlenirken sıcaklık verilerinde olduğu gibi Genelleştirilmiş Çapraz Geçerlilik Yöntemi tercih edilmiştir. Genelleştirilmiş Çapraz Geçerlilik Yöntemine göre düzgünleştirme parametresinin bulunmasında 365 baz fonksiyonu dikkate alınmıştır. Farklı düzgünleştirme parametresi değerleri için simülasyonlar yapılmış ve minimum GCV değerini veren  $\psi$  değeri düzgünleştirme parametresi olarak belirlenmiştir. Şekil 4.6 da yapılan simülasyon sonuçları gösterilmiştir. Bu çalışmada  $\psi$  şeklindeki düzgünleştirme parametresine  $10^4$  ile  $10^9$  arasında farklı değerler verilmiştir. Minimum GCV değeri 9 değerine karşılık gelmektedir. Yapılan çalışmalar  $10^{-4}$ ,  $10^{-3}$  ve  $10^{-2}$  değerlerinin düzgünleştirme parametresi olarak iyi sonuçlar verdiklerini gösterse de bu çalışmada olduğu gibi her zaman en iyi değeri vermeleri gerekir diye de bir genelleme yapmak yanlış olacaktır.



**Şekil 4.7.** Sıcaklık için birinci ana bileşen fonksiyonu ve ortalama fonksiyonu karşılaştırılması

Kesikli kırmızı ve yeşil çizgiler ana bileşen için oluşturulan ortalama fonksiyonunun belirli bir sabitle çarpanının eklenmesinin ve çıkarılmasının etkilerini göstermektedir. Mavi çizgiyse birinci ana bileşen fonksiyonu için ortalama fonksiyonunu ifade etmektedir. Şekil 4.7 de ortalama fonksiyonuna uygun bir çarpanla çarpılmış, ana bileşen fonksiyonu eklenerek ve çıkarılarak elde edilen fonksiyonlarla ortalama fonksiyonu aynı grafik üzerinde çizdirilmiştir. Bundaki amaç ana bileşen fonksiyonunun yorumlanmasını kolaylaştırmaktır. Kesikli kırmızı ve yeşil çizgiler ortalama fonksiyonundan ne kadar uzaksa ortalamadan sapmaların o kadar yüksek olduğu görülmektedir. Bu sapmalar özellikle sıcaklıkların düşük seyrettiği kış aylarında daha rahat ayırt edilebilmektedir. Bu nedenle %68 değişkenlik açıklama gücüyle sıcaklık açısından eğriler arasındaki birinci temel değişkenliğin kış aylarından kaynaklandığı görülmektedir. Diğer bir ifadeyle veriler arasındaki değişkenliğin birinci sebebinin kış ayları olduğu anlaşılmaktadır.

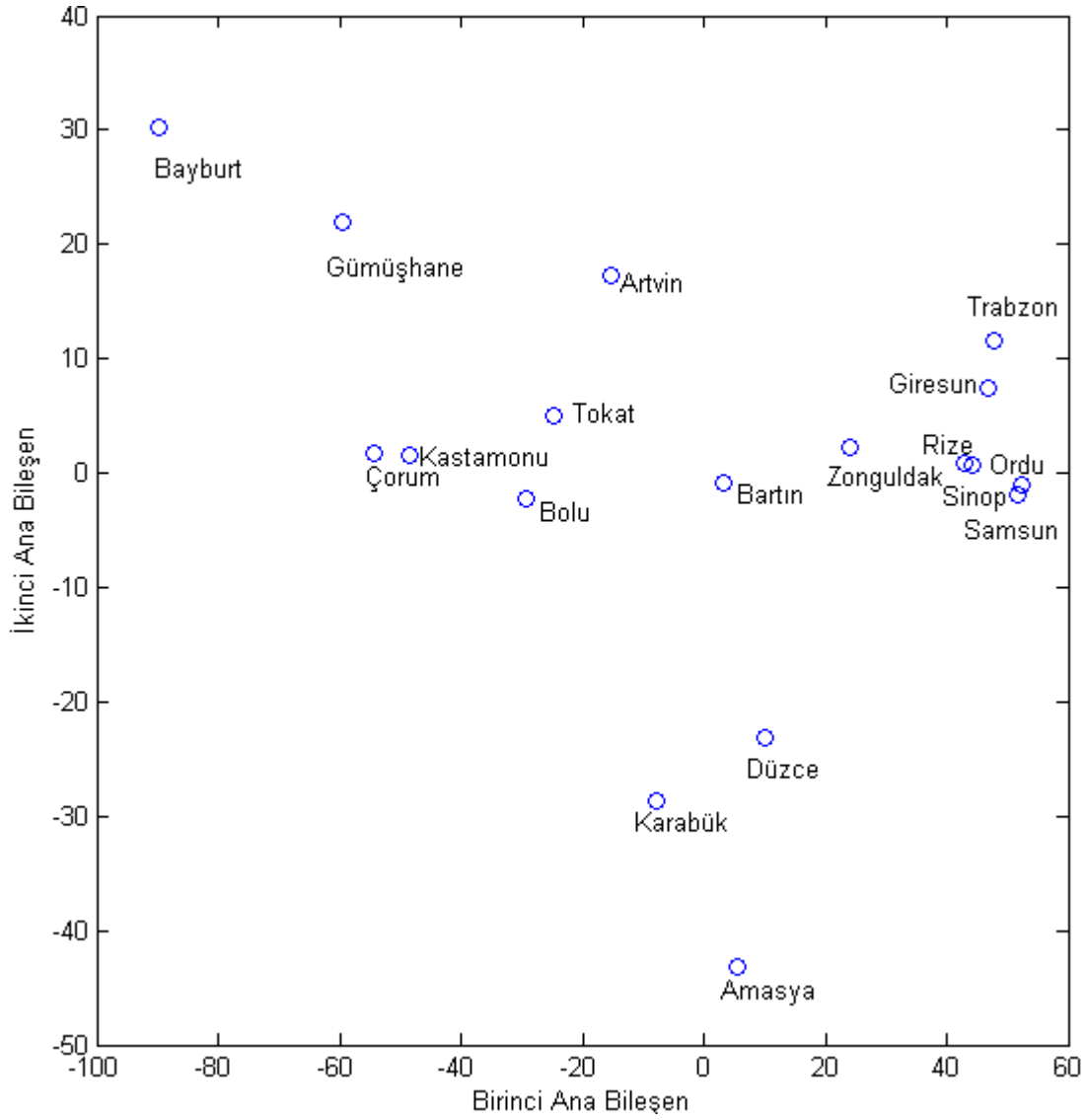


**Şekil 4.8.** Sıcaklık için ikinci ana bileşeni fonksiyonu ve ortalama fonksiyonunu karşılaştırılması

Kesikli kırmızı ve yeşil çizgiler ana bileşen için oluşturulan ortalama fonksiyonunun belirli bir sabitle çarpımının eklenmesinin ve çıkarılmasının etkilerini göstermektedir. Mavi çizgiyse ana bileşen fonksiyonu için ortalama fonksiyondur. Şekil 4.8 de ortalama fonksiyonuna uygun bir çarpanla çarpılmış ana bileşen fonksiyonu eklenerek ve çıkarılarak elde edilen fonksiyonlarla ortalama fonksiyonu aynı grafik üzerinde çizdirilerek ana bileşen fonksiyonunun yorumlanmasının kolaylaştırılması amaçlanmıştır. Kesikli kırmızı ve yeşil çizgiler ortalama fonksiyonundan ne kadar uzaksa ortalamadan sapmaların o kadar yüksek olduğu görülür. Bu sapmalar genellikle 100. ve 150. günler arasında görülmüştür. Fakat doğrudan kış, bahar ya da yaz günlerinde değişkenliğin fazla olduğu söylenemez. Ana bileşen skorlarının yorumlanmasıyla burada en çok hangi bireysel fonksiyonun etkili olduğu görülebilecektir.

**Çizelge 4.1.** Sıcaklık için ana bileşen skorları

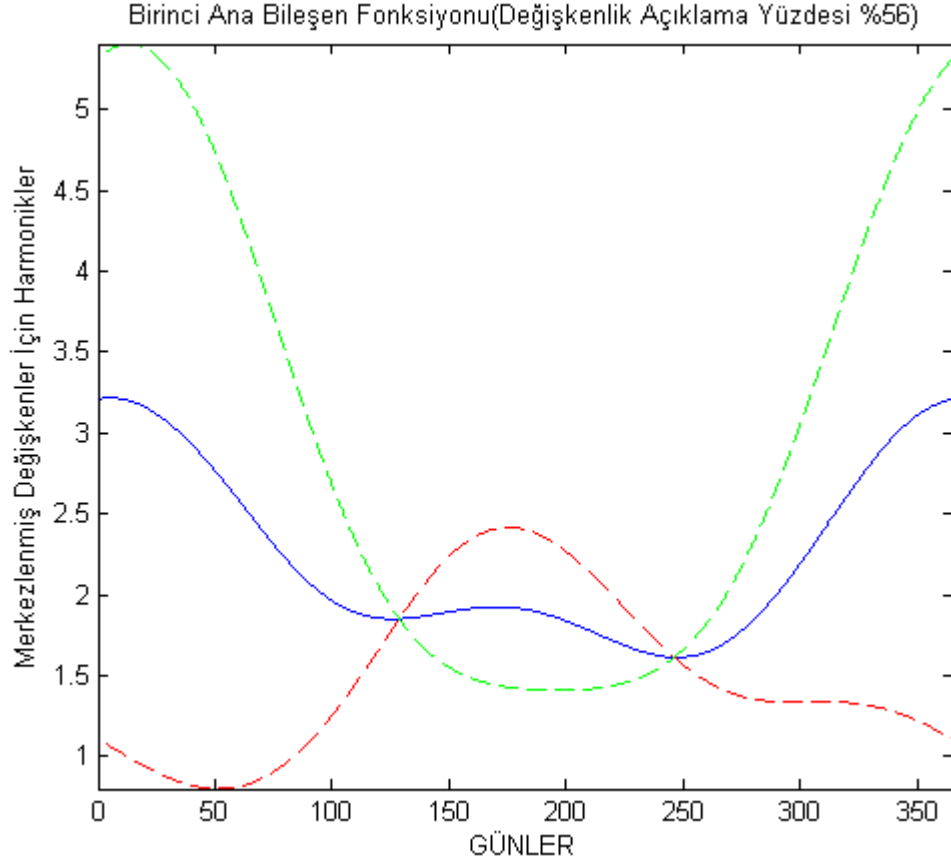
| Şehirler  | Birinci Ana Bileşen Skorları | İkinci Ana Bileşen Skorları |
|-----------|------------------------------|-----------------------------|
| Artvin    | -15,35                       | 13,99                       |
| Rize      | 44,17                        | 0,7327                      |
| Trabzon   | 47,83                        | 11,67                       |
| Bayburt   | -89,82                       | 30,28                       |
| Gümüşhane | -59,47                       | 21,99                       |
| Giresun   | 46,89                        | 7,365                       |
| Ordu      | 52,25                        | -0,9954                     |
| Samsun    | 51,82                        | -1,818                      |
| Sinop     | 43,04                        | 0,7975                      |
| Tokat     | -24,74                       | 5,101                       |
| Amasya    | 5,588                        | -43,24                      |
| Çorum     | -54,16                       | 1,805                       |
| Kastamonu | -48,43                       | 1,605                       |
| Bartın    | 3,176                        | -0,8542                     |
| Zonguldak | 24,15                        | 2,295                       |
| Düzce     | 10,13                        | -23,11                      |
| Karabük   | -7,862                       | -28,6                       |
| Bolu      | -29,22                       | -2,263                      |



**Şekil 4.9.** Sıcaklık için ana bileşen skorlarının dağılımı

Ana bileşen skorları özellikle %68 değişkenlik açıklayabilme gücüne sahip birinci ana bileşen açısından incelendiğinde şekilden öncelikle Rize, Ordu, Samsun, Sinop, Giresun, Trabzon'un birinci ana bileşen üzerinde oldukça etkili olduğu yüksek bir ana bileşen skoru değeri taşıdıkları görülmektedir. Bunun tam tersi olarak Bayburt, Gümüşhane ve Çorum en yüksek negatif skor değerine sahiptir. Gerçekten de en yüksek değişkenliğin kış aylarında meydana geldiğini vurgulayan birinci ana bileşen için verileri incelediğimizde en yüksek ortalama sıcaklıkların Rize, Ordu, Samsun, Sinop, Giresun ve Trabzon'da olduğu görülmektedir. En düşük sıcaklıklarında Bayburt, Gümüşhane ve Çorum'da olduğu görülmektedir. İkinci ana bileşen açısından bakıldığında en yüksek değişkenliğin başka bir ifadeyle ortalamadan sapmaların Şekil 4.8 de 100. ve 150. günler arası meydana geldiği

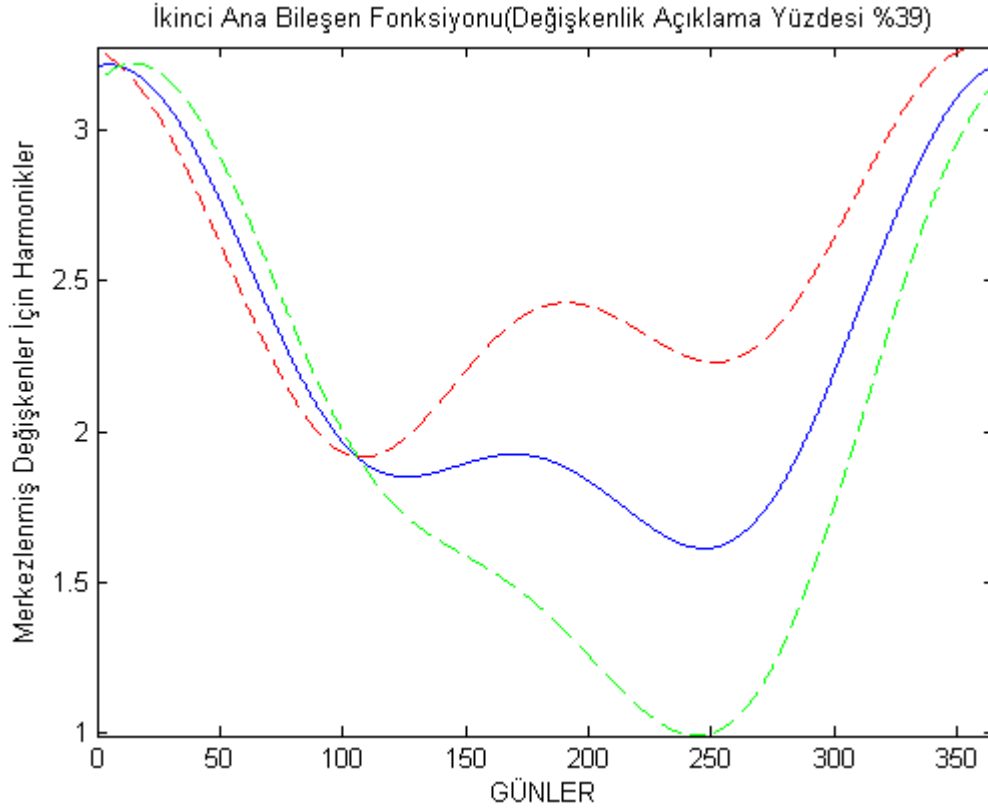
bilinmektedir. Burada ikinci ana bileşen üzerinde en etkili şehirlerin Bayburt, Gümüşhane olduğu açık bir şekilde görülmektedir. Veriler incelendiğinde en düşük sıcaklıkların bu dönemde bu illere ait oldukları görülecektir. Ayrıca Şekil 4.3 de 100. ve 150. günler arasında en düşük seyreden ortalama fonksiyonların bu illere ait oldukları görülebilmektedir.



**Şekil 4.10.** Yağış için birinci ana bileşen fonksiyonu ve ortalama fonksiyonu karşılaştırılması

Kesikli kırmızı ve yeşil çizgiler ana bileşen için oluşturulan ortalama fonksiyonunun belirli bir sabitle çarpanının eklenmesinin ve çıkarılmasının etkilerini göstermektedir. Mavi çizgiyse ana bileşen fonksiyonu için ortalama fonksiyonudur. Şekil 4.10 da ortalama fonksiyonuna uygun bir çarpanla çarpılmış ana bileşen fonksiyonu eklenerek ve çıkarılarak elde edilen fonksiyonlarla ortalama fonksiyonu aynı grafik üzerinde çizdirilerek yorumlamayı kolaylaştırmak amaçlanmıştır. Kesikli kırmızı ve yeşil çizgiler ortalama fonksiyonundan ne kadar uzaksa ortalamadan sapmaların o derece yüksek olduğu görülmektedir. Yağışların fazla olduğu kış aylarında bu sapmalar oldukça rahat görülmektedir. Bu nedenle %56 değişkenlik açıklama gücüyle yağış açısından eğriler arasındaki birinci temel değişkenliğin kış

aylarından kaynaklandığı görülmektedir. Diğer bir ifadeyle veriler arasındaki değişkenliğin birinci sebebi kış aylarıdır.

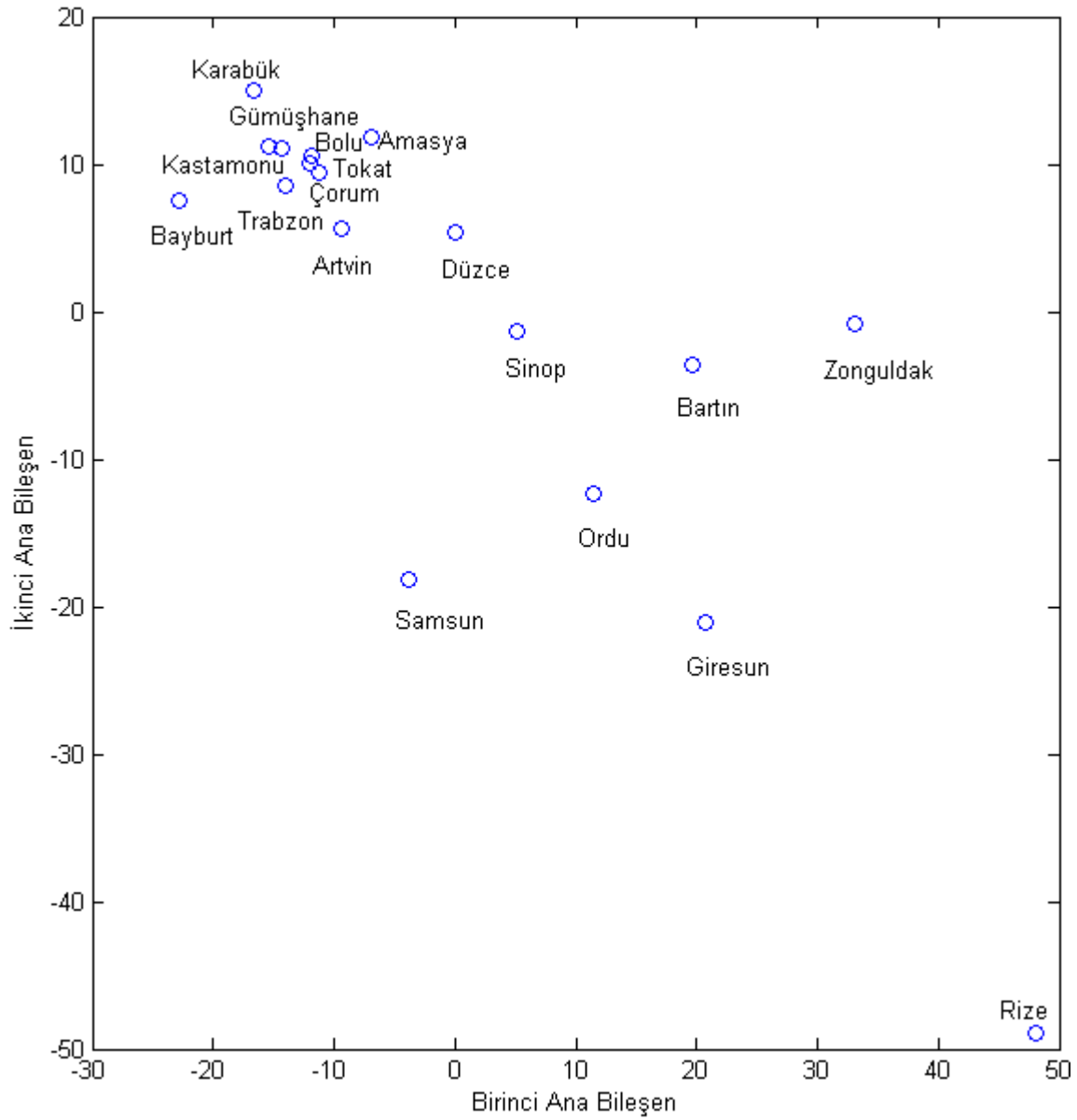


**Şekil 4.11.** Yağış için ikinci ana bileşen fonksiyonu ve ortalama fonksiyonu karşılaştırılması

Kesikli kırmızı ve yeşil çizgiler ana bileşen için oluşturulan ortalama fonksiyonunun belirli bir sabitle çarpanının eklenmesinin ve çıkarılmasının etkilerini göstermektedir. Mavi çizgiyse ana bileşen fonksiyonu için ortalama fonksiyonudur. Şekil 4.11 de ortalama fonksiyonuna uygun bir çarpanla çarpılmış ana bileşen fonksiyonu eklenerek ve çıkarılarak elde edilen fonksiyonlarla ortalama fonksiyonu aynı grafik üzerinde çizdirilerek yorumlamayı kolaylaştırmak amaçlanmıştır. Kesikli kırmızı ve yeşil çizgiler ortalama fonksiyonundan ne kadar uzaksa ortalamadan sapmaların o derece yüksek olduğu görülmektedir. İkinci ana bileşende değişkenlikteki en etkili artışın 150. ve 250. günler arasında olduğu ve ortalama fonksiyonunun azalma eğiliminde olduğu göze çarpmaktadır. Bu nedenle %39 değişkenlik açıklama gücüyle yağış açısından eğriler arasındaki ikinci temel değişkenlik bu dönemlerde meydana gelen düşük yağışlardan kaynaklandığı anlaşılmaktadır. Ana bileşen skorları incelenip yapılan yorumların doğru olduğu kanıtlanmaya çalışılacaktır.

**Çizelge 4.2.** Yağış için ana bileşen skorları

| Şehirler  | Birinci Ana Bileşen Skorları | İkinci Ana Bileşen Skorları |
|-----------|------------------------------|-----------------------------|
| Artvin    | -9,332                       | 5,67                        |
| Rize      | 48,08                        | -48,89                      |
| Trabzon   | -14                          | 8,51                        |
| Bayburt   | -22,84                       | 7,554                       |
| Gümüşhane | -14,34                       | 11,01                       |
| Giresun   | 20,72                        | -21,03                      |
| Ordu      | 11,47                        | -12,38                      |
| Samsun    | -3,804                       | -18,2                       |
| Sinop     | 5,167                        | -1,347                      |
| Tokat     | -12,07                       | 10,04                       |
| Amasya    | -6,858                       | 12,12                       |
| Çorum     | -11,24                       | 9,454                       |
| Kastamonu | -15,38                       | 11,15                       |
| Bartın    | 19,73                        | -3,563                      |
| Zonguldak | 33,07                        | -0,8414                     |
| Düzce     | 0,1093                       | 5,398                       |
| Karabük   | -16,66                       | 14,97                       |
| Bolu      | -11,83                       | 10,63                       |



**Şekil 4.12.** Yağış için ana bileşen skorları

Ana bileşen skorları özellikle %56 değişkenlik açıklayabilme gücüne sahip birinci ana bileşen açısından değerlendirildiğinde şekilden öncelikle Rize ve Zonguldak'ın birinci ana bileşen üzerinde oldukça etkili olduğu, yüksek bir ana bileşen değeri taşıdıkları görülmektedir. En yüksek değişkenliğin kış döneminde meydana geldiğini gösteren birinci ana bileşen için, veriler incelendiğinde en yüksek ortalama yağışların Rize ve Zonguldak illerinde olduğu görülecektir. İkinci ana bileşen açısından değerlendirildiğinde en yüksek değişkenliğin bir başka ifadeyle ortalamadan sapmalar 150. ve 250. günler arası görülmektedir. %39 değişkenlik açıklayabilme gücüne sahip ikinci ana bileşen açısından değerlendirildiğinde şekilden Karabük, Gümüşhane, Tokat, Kastamonu, Bolu ve Amasya'nın

ikinci ana bileşen üzerinde oldukça etkili olduğu ve yüksek bir ana bileşen değeri taşıdıkları görülmektedir. En düşük yağış miktarının bu dönemlerde meydana geldiğini gösteren ikinci ana bileşen için, veriler incelendiğinde en düşük ortalama yağışların bu illerde oldukları görülecektir. Ayrıca Şekil 4.4 de 150. ve 250. günler arası en düşük seyreden bireysel fonksiyonlar bu illere aittir.



## 5. SONUÇ VE ÖNERİLER

Teknolojik gelişmelerin daha büyük örneklemlemlerle çalışabilmeye imkân tanınması verilerin analiz edilmesinde klasik tek değişkenli ve çok değişkenli yöntemlerin kullanılmasını zorlaştırmaktadır. Fonksiyonel Veri Analizinin öneminin her geçen gün artmasının temel nedeni interpolasyon yöntemi sayesinde kayıp verilerle baş edilebilirken oluşturulan bireysel fonksiyonların istenilen dereceden türevlerinin de incelenmesine olanak sağlamasıdır. Fonksiyonel Veri Analizi sistemde olması beklenen ve önceden fark edilemeyen ilişkileri çok rahat bir biçimde ortaya çıkarabilmektedir. Bu ilişkileri ortaya çıkartırken verilere ait bireysel fonksiyonlardan, ortalama fonksiyonundan, kovaryans yüzeylerinden, ana bileşen fonksiyonlarından, ana bileşen fonksiyonu için oluşturulan ortalama fonksiyonundan yararlanılması mümkündür.

Bu uygulama da Karadeniz Bölgesi'nde bulunan 18 şehrin 2012 yılındaki günlük ortalama sıcaklık ve yağış verileri Baz Fonksiyon ve Pürüzlü Ceza Yaklaşımlarıyla 18 bireysel fonksiyona dönüştürülmüştür. Veriler arasındaki değişim yapısını ortaya koymak için Düzgünleştirilmiş Fonksiyonel Ana Bileşenler Analizinden yararlanılmıştır. Bu uygulamada örneklem hacmi değişken sayısından küçük olduğundan veriler arasındaki değişkenlik yapısını ortaya koymak için Klasik Çok Değişkenli Ana Bileşenler Analizini kullanmak doğru olmaz.

Düzgünleştirilmiş Fonksiyonel Ana Bileşenler Analizi sayesinde tüm fonksiyonlar ile ele alınıp incelenmesi zor gözükken yağış ve sıcaklık fonksiyonları arasındaki değişim ana bileşen fonksiyonu için oluşturulan ortalama fonksiyonuna uygun bir çarpanla çarpılmış, ana bileşen fonksiyonu eklenerek ve çıkarılarak elde edilen fonksiyonlarla ortalama fonksiyonu aynı grafik üzerinde çizdirilerek ortaya çıkartılmıştır. %68 değişkenlik açıklama gücüyle sıcaklık açısından eğriler arasındaki birinci temel değişkenliğin kış aylarından kaynaklandığı görülmüştür. Diğer bir ifadeyle sıcaklık için değişkenliğin birinci sebebinin kış ayları olduğu anlaşılmıştır. Sıcaklık için ana bileşen skorları incelenerek ikinci ana bileşen üzerinde Bayburt ve Gümüşhane illerinin etkili olduğu saptanmıştır. %56 değişkenlik açıklama gücüyle

yağıř aısından eđriler arasındaki birinci temel deđiřkenliđin kiř aylarından kaynaklandıđı saptanmıřtır. Diđer bir ifadeyle yağıř iin deđiřkenliđin birinci sebebinin kiř ayları olduđu anlařılmıřtır. %39 deđiřkenlik aıklama gcyle yağıř aısından eđriler arasındaki ikinci temel deđiřkenlik 150. ve 250. gnler arası meydana gelen dřk yağıřlardan kaynaklandıđı anlařılmıřtır.

## KAYNAKLAR

- Benko M., 2004. Functional Principal Components analysis, Implementation and Applications, A Master Thesis, Humboldt University Center of Applied Statistics and Economics, Berlin.
- Craven P., Wahba G., 1979. Smoothing Noisy Data with Spline Functions, *Numerische Mathematik*, 31.
- Costanzo G.D., 2005. *Functional Principal Component Analysis of Financial Time Series*, Cnam-Paris.
- Green.P.J., Silverman B.W., 1994. *Nonparametric Regression and Generalized Linear Models: A Roughness Penalty Approach*, Chapman & Hall, London.
- Kadir E., Keser İ., 2007. Düzgünleştirilmiş Fonksiyonel Ana Bileşenler Analizi ile İMKB Verilerinin İncelenmesi, *İstanbul Üniversitesi İktisat Fakültesi Ekonometri ve İstatistik Dergisi*, 8, 1-32.
- Keser İ., 2010. Ege Bölgesi Yağış Verilerinin Fonksiyonel Veri Analizi İle İncelenmesi, *Dokuz Eylül Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dergisi*, 1, 41-67.
- Keser İ., 2008. Karşılaştırmalı Olarak Fonksiyonel Ana Bileşenler Analizi ve GSYİH Verilerinin İncelenmesi, *Ege Akademik Bakış*, 8, 915-928.
- Mardia K.V., Kent J.T., Biby J.M., 1989. *Multivariate Analysis*, Academic Press, London.
- Ramsay J.O., Dalzell C., 1991. Some Tools For Functional Data Analysis, *Journal of the Royal Statistical Society: Series B*, 53(3), 539-572.
- Ramsay J.O., Li X., 1998. Curve Registration, *Journal of the Royal Statistical Society, Series B*, 60(2), 351-363.
- Ramsay J.O., Hooker G., 2010. fda: Functional Data Analysis, Matlab Package Version 2.2.6., <http://cran.r-project.org/src/contrib/Archive/fda/> (Ziyaret Tarihi 15 Mayıs 2014).
- Ramsay J.O., Silverman B.W., 2005. *Functional data analysis*, 2nd ed., Springer Series in Statistics, Springer, New York.
- Ramsay J.O., Silverman B.W., 1997. *Functional Data Analysis*, Springer Series in Statistics, Springer – Verlag, New York.
- Silverman B.W., 1995. Incorporating Parametric Effects into Functional Principal Components Analysis, *Journal of the Royal Statistical Society, Series B*, 57, 673-689.
- Silverman B.W., 1996. Smoothed Functional Principal Component Analysis By Choise Of Norm, *The Annals of Statistics*, 24(1).

- Ulbricht J., 2004. Representing Functional Data as Smooth Functions, A Master Thesis, Institute of Statistics and Econometrics, Humboldt University, Berlin.
- Yılmaz K., Çamurcu Y., Doğan B., Veri Madenciliğinde Temel Bileşenler Analizi ve Negatifsiz Matris Çarpanlarına Ayırma Tekniklerinin Karşılaştırmalı Analizi, <http://ab.org.tr/ab10/bildiri/34.doc/> (Ziyaret Tarihi 12 Mayıs 2014).
- Günlük Ortalama Sıcaklık ve Yağış Verileri, 2012. Samsun Meteoroloji 10. Bölge Müdürlüğü.

## **ÖZGEÇMİŞ**

**Adı Soyadı** : Çağlar SÖZEN  
**Doğum Yeri ve Tarihi** : Üsküdar/İSTANBUL, 1988  
**E-Posta** : caglarsozen@gmail.com  
**Lisans** : 2007-2011 Ondokuz Mayıs Üniversitesi Fen Edebiyat  
Fakültesi İstatistik Bölümü

