Dokuz Eylül Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dergisi,
Cilt:25, Sayı:1, Yıl:2010, ss.41-67.

Ege Bölgesi Yağış Verilerinin Fonksiyonel Veri Analizi İle İncelenmesi

İstem Köymen KESER*

Özet

Araştırmalarda incelenen gözlem noktası sayısı arttıkça bir diğer deyişle çalışma sahası genişledikçe gözlemlerin altta yatan reel bir fonksiyondan örneklendiği varsayılır. Bu tip verileri analiz etmek için geliştirilen istatistiksel metodlar Fonksiyonel Veri Analizi (FVA) terimi ile adlandırılır. Fonksiyonel Veri Analizinde ilk adım ayrık noktalardaki gözlemlerden oluşan şans örneğini reel fonksiyonlardan oluşan bir şans örneğine dönüştürmektir. Bunun için ilk olarak Baz Fonksiyon Yaklaşımı ve daha sonra da Pürüzlü Ceza Yaklaşımı kullanılmıştır.

Bu çalışmada 22 meteoroloji istasyonundan alınan aylık ortalama yağış verileri incelendiğinden ve yağış verileri doğal olarak periyodik bir yapı izlediğinden baz fonksiyon yaklaşımı olarak Fourier baz fonksiyonları ele alınmıştır. Daha sonra verilerdeki değişkenlik yapısını incelemek üzere Düzgünleştirilmiş Fonksiyonel Ana Bileşenler Analizi kullanılmış ve burada λ düzgünleştirme parametresinin değeri Genelleştirilmiş Çapraz Geçerlilik Metoduna göre belirlenmiştir.

Uygulama olarak 2000 ve 2005 yılları arasında Ege Bölgesinde bulunan 22 farklı meteoroloji istasyonundan 71 noktada alınan aylık ortalama yağış verileri incelenmiş ve öncelikle 22 farklı istasyon için Pürüzlü Ceza Yöntemine göre reel fonksiyonlar oluşturulmuştur. Daha sonra yağış verileri için ortalama fonksiyonu, kovaryans yüzeyi ve Düzgünleştirilmiş Fonksiyonel Ana Bileşenler Analizine göre belirlenen ana bileşen fonksiyonları ve birinci ana bileşen türev fonksiyonu oluşturulmuş ve yorumlanmıştır. Birinci ana bileşen fonksiyonuyla yağışlar arasındaki en yüksek değişkenliğin kış aylarında meydana geldiği tespit edilmiştir.

Anahtar Sözcükler: Fonksiyonel Veri Analizi, Düzgünleştirme, Fonksiyonel Ana Bileşenler Analizi

The Analyse Of Aegean Region Rainfall Data By Using Functional Data Analysis

Abstract

As the number of observation points increase, it is assumed that they are sampled from an underlying real function. Statistical methods that are developed for analyzing this kind of data are called Functional Data Analysis (FDA). The first step in FDA is

* Yrd.Doç.Dr., Dokuz Eylül Üniversitesi, İktisadi ve İdari Bilimler fakültesi, Ekonometri Bölümü,(istem.koymen@deu.edu.tr)

transforming the random sample, which consists of observations on separate points, into a random sample which consists of real functions. Therefore Basis Function Approach and then Roughness Penalty Approach is used in this study.

Since we examined average monthly rainfall and since rainfall data are naturally periodical, Fourier basis functions are utilized. Regularized Functional Principle Components Analysis is used in order to investigate the variation structure in the data. Then the value of the smoothing parameter, λ*, determined by using Generalized Cross-Validation Method.*

Monthly average rainfall data taken at 71 points from 22 meteorology stations in Aegean Region between 2000 and 2005 are investigated in this study. Firstly, real functions are formed for 22 different stations by using Roughness Penalty Approach. Then, average function, covariance surface, principle component functions, and first principle component derivative function are formed and interpreted.The first principle component function shows that the highest variation between rainfalls occurs in winter months.

Keywords: Functional Data Analysis, Smoothing, Functional Principal Components Analysis, Roughness PenaltyApproach

---

1-Giriş

Günümüz teknolojisi artık çok geniş hacimli örneklerle çalışılabilmeye ve bunların istatistiksel analizine imkan vermektedir. İncelenen çalışma sahası bir diğer deyişle, örneğe dahil edilen gözlem noktası sayısı arttıkça aslında ayrık noktalarda gözlenen bu verilerin altta yatan reel bir fonksiyondan örneklendiği varsayılır. Dolayısıyla bu gözlemler “Fonksiyonel Veriler” olarak adlandırılabilir. Fonksiyonel verileri analiz etmek için geliştirilen istatistiksel metodlar ilk olarak Ramsay ve Dalzell (1991) tarafından “Fonksiyonel Veri Analizi” terimi ile adlandırılmıştır.

Fonksiyonel Veri Analizinin ilk adımı ayrık noktalarda gözlenen verilerin sürekli fonksiyonlar haline dönüştürülmesidir. Bu adım interpolasyon (interpolation) veya düzgünleştirme (smoothing) ile sağlanır. Bu çalışmada buharlaşma v.b. nedenlerden dolayı ölçüm hatalarına müsait olan yağış verileri incelendiğinden düzgünleştirme süreci benimsenmiştir. 22 farklı birey (bu çalışmada meteoroloji istasyonu) ve 71 ayrık noktada (bu çalışmada aylar) gözlem yapıldığından incelenen ana kütle karmaşık bir hal almış ve verilerdeki bu karmaşık yapının anlaşılması amacıyla çalışmada Düzgünleştirilmiş Fonksiyonel Ana Bileşenler Analizi kullanılmıştır.

2- Fonksiyonel Verilere Dönüşüm

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. Eğer gözlemlenen değerlerin hatasız oldukları varsayılırsa, bu süreç interpolasyon yöntemi ile yapılır. Fakat verilerin elde edilmesi örneğin bir ölçüm prosesinin sonucu ise ve böylece verilerde ortadan kaldırılması gereken bazı ölçüm prosesinden kaynaklanabilecek hatalar mevcutsa, kesikli verilerden fonksiyonel verilere yapılan bu dönüşüm süreci, düzgünleştirme adını alır (Ramsay ve Silverman, 1997: 9). Bu çalışmada da, doğrudan doğruya interpolasyon uygulaması yerine, düzgünleştirme yöntemi olarak öncelikle **Baz Fonksiyon Yaklaşımı** ve ikinci aşama olarak da istatistiksel uygulamalarda çoğu kez tercih edilen **Pürüzlü Ceza Yaklaşımı** kullanılmaktadır.

2.1. Baz Fonksiyon Yaklaşımı

$x_i(t)$ fonksiyonunu oluştururken esnek yöntemlere ihtiyaç duyulmaktadır. Bunun için K tane baz fonksiyondan (basis function) oluşan bir sistem seçilmektedir. Oluşturulmak istenilen $x_i(t)$ fonksiyonu bu baz fonksiyonların ağırlıklandırılmış bir toplamı olarak şu şekilde yazılabilir:

$$x_i(t) = c_1\theta_1(t) + c_2\theta_2(t) + ... + c_k\theta_k(t) \qquad (2.1)$$

Bu ifadede yer alan $\theta_i(t)$ i.inci baz fonksiyon ve c_i ise bu baz fonksiyona karşılık gelen katsayıdır. Genel olarak en çok kullanılan baz fonksiyonlar Kuvvetler, Fourier baz ve B-Splayn baz şeklinde sıralanabilir. Bu çalışmada yağış verileri ile çalışıldığından ve yağışlar genel olarak periyodik bir yapı gösterdiğinden dolayı **Fourier baz** kullanılmıştır.

c_i, $i=1, 2, \dots, K$ katsayıları ise $x_i(t)$ fonksiyonunun şeklini ve biçimini belirleyen katsayılardır. Bir anlamda parametre olarak yorumlanabilirler. Bu çalışmada "**Pürüzlü Ceza**" yaklaşımı ile de c_i katsayılarının tahminlenmesi amaçlanmaktadır.

2.1.1. Fourier Baz

Trigonometrik Fonksiyonlar geniş ölçüde periyodik fonksiyonlara yaklaşım için kullanılırlar. Bir periyodik fonksiyon $x(t)$ sonlu veya sonsuz Fourier serisi cinsinden genel olarak ;

$x(t) = c_0 + c_1 \sin wt + c_2 \cos wt + c_3 \sin 2wt + c_4 \cos 2wt + \ldots t \in T[a,b]$
(2.2)

şeklinde ifade edilebilir. Burada bazlar periyodiktir ve w parametresi 2π /w periyodunu belirler. Eğer t_j değerleri ilgilenilen reel aralık olan T de eşit ölçeklenmişse ve periyod T aralığının uzunluğuna eşitse bu durumda baz ortogonaldir, ayrıca baz fonksiyonları uygun sabitlere bölerek bazlar ortonormal hale getirilebilir(Benko, 2004:13).

Bir diğer gösterimle, K bir çift tamsayı olmak üzere , Fourier baz,

$$\theta_o(t) = \frac{1}{\sqrt{T}}$$

$$\theta_{2r\text{-}1}(t) = \frac{1}{\sqrt{T/2}} \sin rwt$$

$$\theta_{2r}(t) = \frac{1}{\sqrt{T/2}} \cos rwt \qquad r = 1, \ldots, K/2 \text{ için} \qquad (2.3)$$

şeklinde tanımlanabilir. Ayrıca katsayıların uygun sabitler olarak adlandırılan $\frac{1}{\sqrt{T}}$ ve $\frac{1}{\sqrt{T/2}}$ ye bölünmesiyle baz fonksiyonlar ortonormal hale gelir, bir anlamda,

$$\int_T \theta_{k_1}(t)\theta_{k_2}(t)dt = \begin{cases} 0 & k_1 \neq k_2 \\ 1 & k_1 = k_2 \end{cases} \qquad (2.4)$$

halini alır (Ulbricht, 2004: 19).

2.1.2 Pürüzlü Ceza Yaklaşımı

$\theta_i(t)$ şeklinde gösterilen Fourier baz fonksiyonları belirlendikten sonra bireysel fonksiyonları elde etmek için ikinci adım olarak c_i katsayılarının tahminlenmesine geçilir. Bu amaçla da Pürüzlü Ceza Yaklaşımı kullanılır. Fonksiyonel veri analizinde verilere bir eğrinin uyumunu sağlarken tek amaç yalnızca iyi bir uyum yapmak değil, aynı zamanda bu amaçla aslında çatışan diğer bir amaç da çok fazla iniş çıkış göstermeyen bir eğri tahmini elde etmektir. Fonksiyonel veri analizinde fonksiyonları düzgünleştirirken yaygın olarak kullanılan Pürüzlü Ceza Yaklaşımının temel amacı eğrinin pürüzlülüğünü (roughness) ölçmek ve verilerin eğriye uyumu ve eğrinin pürüzlülüğü arasında bir uzlaşma sağlamaktır.

Bu iki çatışan amaç bir anlamda istatistiğin temel prensibinin iki elemanına karşılık gelebilir. Bilindiği gibi, Ortalama Karesel Hata, Sapmanın karesi ile Örnekleme Varyansının toplamına eşittir. Örnekleme varyansını azaltmak için sapmadan biraz taviz verilebilir, bu da tahminlenen eğriye düzgünleştirme yüklenmesinin temel nedenidir(Ramsay, 2005: 85).

Pürüzlü Ceza Yaklaşımında, pürüzlü ceza tahminlerini elde ederken, $x(t)$ fonksiyonu $T=[a,b]$ aralığında tanımlı türevi alınabilen bir fonksiyon ve $\lambda>0$ düzgünleştirme parametresi olsun. Cezalı kareler toplamı (CKT_λ),

$$CKT_\lambda = \sum_j \left(y_j - x(t_j)\right)^2 + \lambda \|Lx\|^2 \qquad (2.5)$$

şeklinde ve baz fonksiyon yaklaşımına göre vektör bazında,

$$CKT_\lambda = [\underline{\mathbf{y}} - \mathbf{\Phi}\, \underline{\mathbf{c}}]^T [\underline{\mathbf{y}} - \mathbf{\Phi}\, \underline{\mathbf{c}}] + \lambda\, \underline{\mathbf{c}}^{\mathbf{T}}\, \mathbf{R}\, \underline{\mathbf{c}} \qquad (2.6)$$

şeklinde ifade edilebilir. Burada **R** Pürüzlü Ceza Matrisi , $\underline{y}$ gözlem vektörü, $\mathbf{\Phi}$, $\theta_j(t)$ baz fonksiyonlarından oluşan bir sete sahip olunduğu varsayıldığında, $\theta_i(t_j)$, i=1,2, … ,K; j=1,2, … n elemanlarına sahip (nxK) boyutlu bir matris ve $\underline{\mathbf{c}}$ (Kx1) boyutlu katsayılar vektörüdür.

Periyodik verileri analiz ederken bir eğrinin pürüzlülüğünün belirlenmesinde *x* eğrileri ideal olarak belirli bir diferansiyel denklemi sağlamalıdır ve bu denklemden sapmalar cezalandırılmak istenebilir. Eğer yağış ve sıcaklık gibi periyodik veriler analiz ediliyorsa,

$$Lx = D^3x + \omega^2 Dx \qquad (2.7)$$

şeklinde harmonik hız (acceleration) operatörünü eğrinin pürüzlülüğünün değerlendirilmesinde kullanmak doğaldır. Burada sıfır pürüzlülük $x(t)$ eğrisinin,

$$x(t) = c_1 + c_2 \sin \omega t + c_3 \cos \omega t \qquad (2.8)$$

formunda olduğunu gösterir ve w burada periyodu belirtmektedir. Bir diğer deyişle *x* eğrisi gerçekten sinüsodial ise bu durumda L*x* sıfıra eşit olur.

Böylece periyodik veriler için pürüzlülüğü ölçmenin evrensel olarak kabul edilen bir yolu,

$$PEN_L(x) = \int [(Lx)^2](t)dt = \|Lx\|^2 \qquad (2.9)$$

şeklinde Lx'in integralinin karesi olarak tanımlanır(Ramsay, 2005: 93). Bir anlamda sinüsodial fonksiyondan sapmalar cezalandırılır.

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

λ şeklindeki düzgünleştirme parametresi de, Artık Kareler Toplamı ile ölçülen 'verilerin eğriye uyumu' ve $\|Lx\|^2$ ile ölçülen '*x* fonksiyonunun pürüzlülüğü' arasındaki değişim oranını ölçer. Eğer λ çok büyük ise bu durumda doğrusal olmayan fonksiyonlar CKT_λ da büyük bir pürüzlülük cezası içerir ve $\lambda \to 0$ *x* eğrisi $f(t_j) = y_j$ (j =1,2, … , n) şeklinde verileri interpole etmeye yaklaşır. Bir diğer deyişle, Cezalı Kareler Toplamındaki temel katkı Artık Kareler Toplamı ile olur. Uyumlaştırılan eğri verileri daha fazla değişkenlik olsa bile daha yakın izler. Burada pürüzlülük üzerine daha az ceza konduğundan dolayı eğri daha değişken olur. Bu limit durumunda bile interpole edilen eğri keyfi değişken değildir, bunun yerine, bu eğri tüm türev alınabilen eğriler içinde verilere uyum gösteren en düzgün eğridir (Ramsay, 2005: 83). Bir diğer deyişle, $\hat{x}_i(t)$ şeklindeki Cezalı Kareler Toplamını minimize eden eğri tahmini düzgünleştirme ve uyum iyiliği arasındaki en iyi uzlaşmadır.

2.1.3.Düzgünleştirme Parametresinin Belirlenmesi

Bireysel fonksiyonları elde ederken son aşama olarak λ ile sembolize edilen düzgünleştirme parametresinin belirlenmesine kısaca değinilecektir. Düzgünleştirme parametresinin belirlenmesinde Green ve Silverman (1994: 29) iki farklı yaklaşımdan bahsetmişlerdir. Bunlardan bir tanesi subjektif bir diğeri ise otomatik seçimdir. Burada otomatik bir yaklaşım olan ve fonksiyonel veri analizinde sıklıkla kullanılan Genelleştirilmiş Çapraz Geçerlilik Yöntemi kullanılmıştır.

Craven ve Wahba (1979) tarafından geliştirilen Genelleştirilmiş Çapraz Geçerlilik Yöntemi Çapraz Geçerlilik prosedürünün daha basit bir versiyonu olarak geliştirilmiştir. Bu kriter genellikle,

$$GCV = \frac{n^{-1}SSE}{\left[n^{-1}trace(I - S_{\varphi,\lambda})\right]^2} \tag{2.10}$$

olarak gösterilir. Burada SSE hata kareler toplamı ve $S_{\varphi,\lambda}$,

$$S_{\varphi,\lambda} = \mathbf{\Phi}\ (\mathbf{\Phi}^{\mathbf{T}}\ \mathbf{\Phi})^{\mathbf{-1}}\ \mathbf{\Phi} \tag{2.11}$$

şeklindeki düzgünleştirme matrisidir. Ve ayrıca df(λ) = trace ($S_{\varphi,\lambda}$) olmak üzere GCV,

$$GCV(\lambda) = \left(\frac{n}{n - df(\lambda)}\right)\left(\frac{SSE}{n - df(\lambda)}\right) \tag{2.12}$$

şeklinde de ifade edilebilir. Amaç λ 'ye göre GCV'nin minimizasyonudur(Ramsay, 2005: 97). Trace matrisin izini ve df de serbestlik derecesini belirtmektedir.

Düzgünleştirme parametresinin belirlenmesi ile ilgili ayrıntılı bilgi için Craven ve Wahba (1979), Eubank (1985), Hutchinson ve Hoog (1985), Raz vd. (1989), Hardle (1997;147-187), Hurvich vd. (1997), Wei (2005) çalışmalarına başvurulabilir.

Son olarak tüm gerekli hesaplamalar yapıldıktan sonra CKT_{λ}'yı minimize eden baz fonksiyon yaklaşımına göre Pürüzlü Ceza Tahminleri,

$$\underline{\hat{\mathbf{c}}} = (\mathbf{\Phi}^{\mathbf{T}}\ \mathbf{\Phi} + \lambda\, \mathbf{R})^{-1}\, \mathbf{\Phi}^{\mathbf{T}}\ \underline{y} \tag{2.13}$$

şeklinde tahminlenir.

Özetle bu çalışmada da ilk adımda Fourier Baz Fonksiyon yaklaşımı ve daha sonra da Pürüzlü Ceza Yaklaşımına göre bir diğer deyişle Cezalı Kareler Toplamını minimize ederek ve Genelleştirilmiş Çapraz Geçerlilik Yöntemine göre düzgünleştirme parametresini belirleyerek 22 farklı meteoroloji istasyonu için bireysel yağış fonksiyonları ve ortalama fonksiyonu Şekil (4.2) deki biçimde elde edilmiştir.

3. Düzgünleştirilmiş Fonksiyonel Ana Bileşenler Analizi

Fonksiyonel Ana Bileşenler Analizinde de amaç klasik Ana Bileşenler Analizinde olduğu gibi incelenen çalışma sahası genişledikçe oluşabilecek veriler arasındaki karmaşıklığın çözümlenmesini sağlamaktır. Fonksiyonel açıdan ele alındığında bireysel eğriler arasındaki değişimin önemli modlarını tanımlamak için Ana Bileşenler Analizinin (ABA) kullanımı güçlü bir araçtır. Ana Bileşenler Analizi sistemde olması beklenen ve aynı zamanda da önceden fark edilmeyen ilişkileri ortaya çıkarır. Bu nedenlerden dolayı Ana Bileşenler Analizi fonksiyonel veri analizinde ele alınan anahtar tekniktir**.**

Fonksiyonel veriler için uygulanan Fonksiyonel Ana Bileşenler Analizinde asıl amaç çok değişkenli veriler için uygulanan Ana Bileşenler Analizi ile aynı olup verilerdeki değişimi etkili bir biçimde tanımlayan bu sefer birkaç ortogonal vektör değil birkaç **ortogonal fonksiyon** elde etmektir. γ_j biçiminde belirtilen ortogonal fonksiyonlar olan ana bileşen ağırlıkları (bunlar genelde harmonik olarak da adlandırılır) şimdi zamanın veya ilgili başka bir değişkenin fonksiyonlarıdır.

Fonksiyonel kavramda her bir ana bileşen bir fonksiyonel veri ile aynı T aralığında tanımlı, verilerin temel "Değişim Modlarını" tanımlayan bir ana bileşen ağırlık fonksiyonu ($\gamma(t)$) ile belirtilir ve doğrusal kombinasyon aşağıdaki biçimde tanımlanır:

$$Y_j = \left\langle \gamma_j , x - E(x) \right\rangle = \int \gamma_j(t) \left\{ x(t) - E\,x(t) \right\} dt \qquad (3.1)$$

Bundan sonra artık γ_j ile belirtilen ağırlıklar $\gamma(t)$ değerlerine sahip bir ağırlık fonksiyonu halini alır. Burada $Y_{j,}$ her bir $x(t)$ için γ_j üzerine $\left\{ x(t) - E\,x(t) \right\}$nin izdüşüm (projection) miktarıdır(Castro v.d., 1986).

Fonksiyonel Ana Bileşenlerin ilk adımında ağırlık fonksiyonu (ana bileşen fonksiyonu veya harmonik fonksiyonu) γ_1 ,

$$\|\gamma_1\|^2 = \int \gamma_1(t)^2\, dt = 1 \qquad (3.2)$$

kısıtı altında, doğrusal bileşenin varyansı olan,

$$\mathrm{Var}(Y_j) = \mathrm{Var}\left\langle \gamma_j , x - E(x) \right\rangle = \iint \gamma_j(s)\, \mathrm{Cov}(s,t)\, \gamma_j(t)\, ds\, dt \qquad (3.3)$$

ifadesini maksimum yapacak biçimde belirlenir.

İkinci ağırlık fonksiyonunun hesaplanması için Klasik Ana Bileşenler Analizinde olduğu gibi Fonksiyonel Ana Bileşenler Analizinde de ağırlık fonksiyonunun,

$$\langle \gamma_j, \gamma_m \rangle = \int \gamma_j(t)\ \gamma_m(t)\ dt = 0 \ (j \neq m) \qquad (3.4)$$

şeklindeki ilave kısıt olan ortogonallik koşullarını da sağlaması gerekir. Her bir ağırlık fonksiyonunun eğrilerdeki değişimin en önemli modunu tanımlama görevi vardır ve burada her bir modun önceki adımlarda tanımlanan modlara ortogonal olması gerekir(Ramsay & Silverman, 1997; 88). Bunun sonucu olarak ağırlık fonksiyonları her aşamada maksimum değişimi açıklayabilecek biçimde oluşturulan ortogonal baz fonksiyonlar setidir.

Doğrusal bileşenin varyansının maksimum yapılması problemi Klasik Ana Bileşenler Analizinde olduğu gibi fonksiyonel veriler içinde özdeğer-özfonksiyon ayrışımı ile çözümlenebilir.

Bu amaç için gerekli fonksiyonel özdenklem aşağıdaki biçimdedir:

$$\int Cov(s,t)\ \gamma(t)\ dt = \psi\ \gamma(s) \qquad s,t \in T\ [a,b] \qquad (3.5)$$

Bu ifadede,

$$Cov(s,t) = N^{-1}\{\sum_{i=1}^{N} x_i(s)\ x_i(t)\} \qquad s,t \in T\ [a,b] \qquad (3.6)$$

şeklindedir. Bu kovaryans ifadesinde $x_i(t)$ her bir birimden ortalama fonksiyon değerinin çıkarılmış halidir. Fonksiyonel Ana Bileşenler Analizinde bu özdenklemi sağlayan farklı **özdeğer - özvektör** değil **özdeğer - özfonksiyon** çiftleri vardır. $\gamma_1, \gamma_2, \ldots$ biçimindeki özfonksiyonlar karşılık geldikleri $\psi_1 \geq \psi_2 \geq \ldots$ özdeğerlerine göre sıralanırlar.

Fonksiyonel Ana Bileşenler Analizi ile elde edilen özfonksiyonlar bir diğer deyişle ana bileşen ağırlıkları da pürüzlü olabilir. Bu pürüzlülük örnekleme varyansından veya gözlem gürültüsünden (observation noise) ve kullanılan fonksiyonel bazın esnekliğinden kaynaklanabilir. Daha durağan ve daha yorumlanabilir sonuçlara sahip olmak için özfonksiyonlar düzgünleştirilebilir.

Fonksiyonel Ana Bileşenler Analizinin değeri düzgünleştirmenin Ana Bileşenler Analizine dahil edilmesi ile biraz daha artar. Fonksiyonel Ana

Bileşenler Analizini düzgünleştirme sadece klasik Ana Bileşenler Analizi ile elde edilen bileşenleri düzgünleştirmek değildir. Düzgünleştirme Ana Bileşenlerin orijinal tanımının içine dahil edilir. Düzgünleştirilmiş Fonksiyonel Ana Bileşenler Analizinde klasik ortonormallik kısıtları fonksiyonların pürüzlülüğünü de hesaba katan bir ortonormallikle yer değiştirir.

Düzgünleştirilmiş Fonksiyonel Ana Bileşenler Analizinde, $\left\|\gamma_j\right\|^2 = \int \gamma_j(t)^2\, dt = 1$ şeklindeki kısıt γ 'nın pürüzlülüğünü de dikkate alan,

$$\int \gamma_j(t)^2 dt + \lambda \mathrm{PEN}_L(\gamma) = 1 \qquad (3.7)$$

kısıtı ile yer değiştirir.

Bu durumda Var (Yj),

$$\int \gamma_j(t)^2 dt + \lambda \mathrm{PEN}_L(\gamma) = 1$$

kısıtına bölünerek Cezalı Ana Bileşen Varyansı (CABV),

$$\mathrm{CABV} = \mathrm{Var}\,(Y_j) = \frac{\iint \gamma_j(s) Cov(s,t) \gamma_j(t) dsdt}{\int \gamma_j(t)^2 dt + \lambda PEN_L(\gamma)} \qquad (3.8)$$

şeklinde elde edilir(Silverman, 1996). Bu analizde GCV'ye göre düzgünleştirme parametresi (λ) 4 olarak belirlenmiştir.

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

$$\int \gamma_j(t)\gamma_m(t)dt + \lambda \int (L\gamma_j(t))(L\gamma_m(t))dt = 0 \qquad j \neq m \qquad (3.9)$$

ilave kısıtı ile maksimize eder.

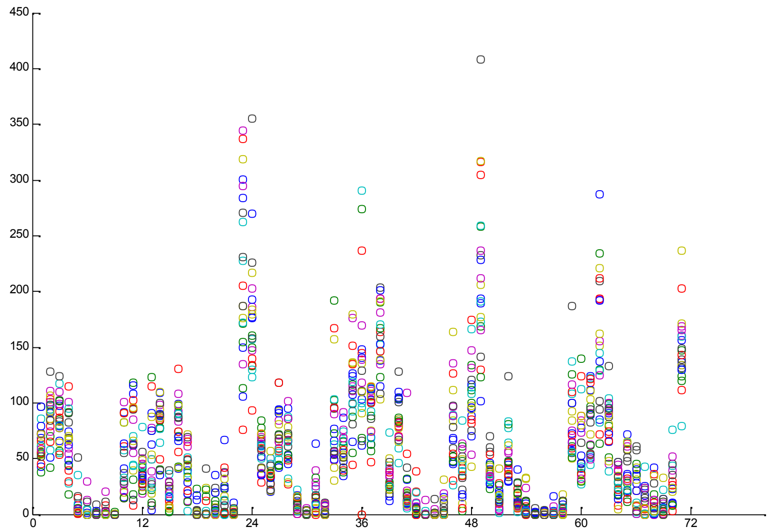
Ana Bileşenler Analizine düzgünleştirme yüklense bile ana bileşen ağırlık fonksiyonlarının açık bir biçimde yorumlanması her zaman mümkün olmayabilir. Bu durumda sonuçların yorumlanmasına yardımcı olacak ilk yaklaşım ana bileşen skorlarının işaretlenmesidir ve bir diğer ikinci yaklaşımda ağırlık fonksiyonunun sabit bir çarpanıyla ortalama fonksiyonun karşılaştırılması olabilir(Silverman, 1995). Burada ortalama fonksiyona ilgilenilen, uygun bir çarpanla çarpılmış, ana bileşen fonksiyonu eklenerek ve çıkarılarak elde edilen fonksiyonlarla ortalama fonksiyon aynı grafik üzerinde çizdirilir ve karşılaştırmalar yapılarak verilerin yapısı ile ilgili çeşitli yorumlamalarda bulunulabilir.

4- Uygulama

Bu çalışmada Ege Bölgesinde bulunan 22 meteoroloji istasyonundan alınan 2000 ve 2005 yılları arasındaki aylık ortalama yağış verileri incelenmiştir. Bu istasyonlar sırasıyla Soma, Köprübaşı, Turgutlu, Alaşehir, Kuyucak, Dikili, Akhisar, Manisa, İzmir, Çeşme, Kuşadası, Didim, Aydın, Bergama, Demirci, Bornova, Salihli, Seferihisar, Ödemiş, Sultan, Selçuk ve Nazilli şeklinde sıralanabilir.

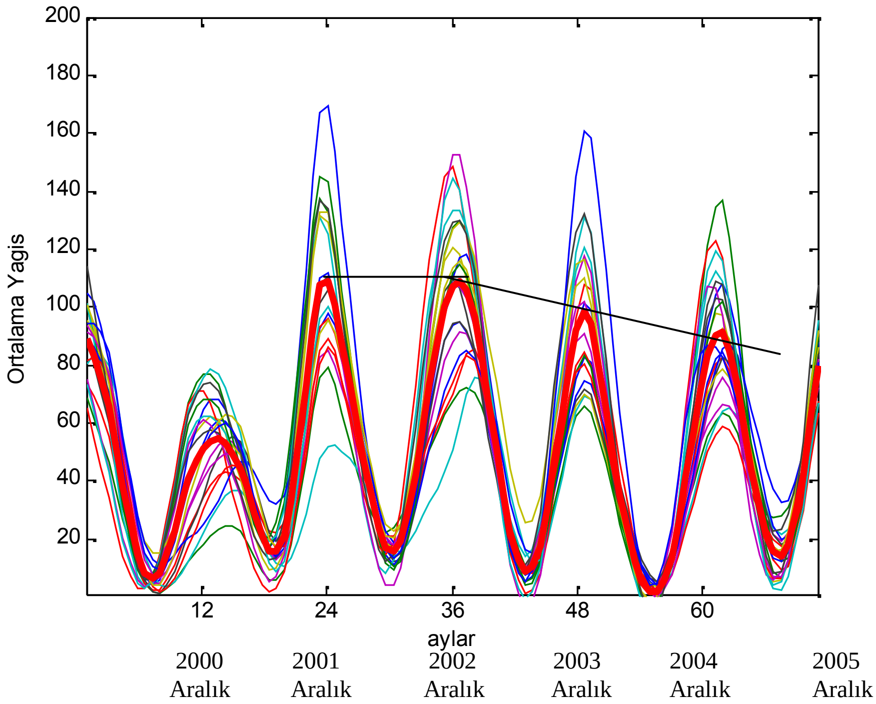
Burada öncelikle yukarıda belirtilen 22 istasyondan 71 ayrık noktada ölçülen ortalama yağış verileri Baz fonksiyon ve Pürüzlü Ceza Yaklaşımlarıyla sürekli birer fonksiyon haline dönüştürülmüş ve öncelikle oluşturulan bu 22 bireysel fonksiyon ve ortalama fonksiyonu incelenmiştir. Düzgünleştirme parametresinin belirlenmesinde Genelleştirilmiş Çapraz Geçerlilik Yöntemi kullanılmıştır. Daha sonra 71 tane değişkene (aya) ait kovaryans yüzeyleri oluşturulmuş ve Pürüzlü Ceza Yaklaşımı ile tahminlenen katsayılara Düzgünleştirilmiş Fonksiyonel Ana Bileşenler Analizi uygulanıp tüm fonksiyonlar birlikte ele alındığında gözlenmesi güç olan bireysel yağış fonksiyonları arasındaki değişim, bir diğer deyişle meteoroloji istasyonları arasındaki yağış miktarları açısından değişim ana bileşen fonksiyonu (özfonksiyon) yardımıyla ortaya konulmaya çalışılmıştır. Bu çalışmada Matlab 7.0 paket programından faydalanılmıştır.

Şekil(4.1) 22 farklı istasyonun her biri için 71 ayrık noktada gözlenen verilerin 22 bireysel fonksiyona dönüştürülmeden önceki halini vermektedir.



Şekil (4.1) : 2000-2005 Ege Bölgesi Aylık Ortalama Yağış Verileri

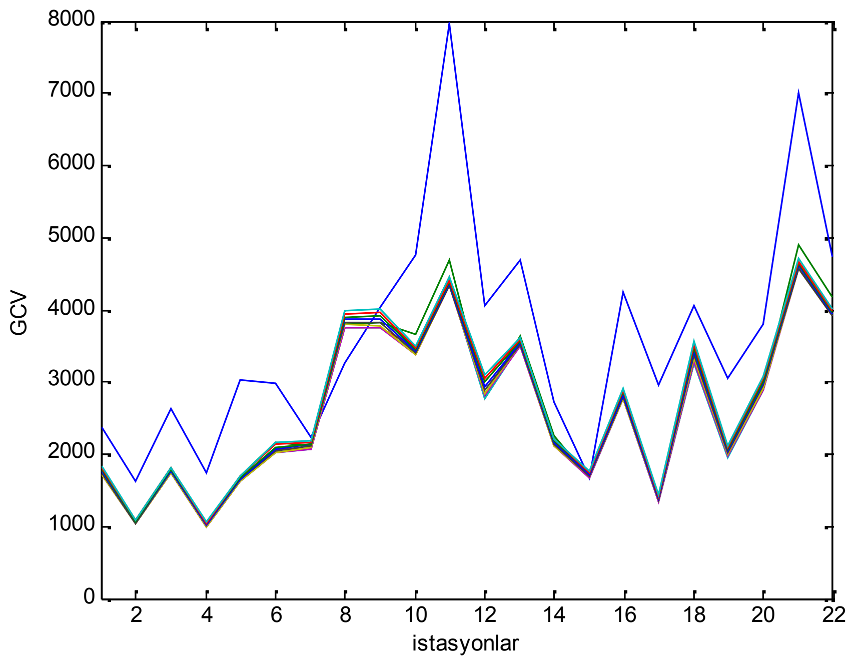
Burada tam (22x71=1562) 1562 tane gözlem noktası bulunmaktadır. Şekilden Ege Bölgesi için aylık ortalama yağışların genel veya meteoroloji istasyonları bazında bireysel seyirlerini çıkarmak Şekil (4.2)ye göre daha güç görünmektedir.



Şekil(4.2) : 22 Bireysel Fonksiyon ve Ortalama Fonksiyonu

Pürüzlü Ceza Yöntemine göre 22 bireysel yağış fonksiyonunu elde ettiğimizde Şekil (4.2)den tüm fonksiyonların bireysel davranışları ve genel seyirleri nispeten rahatlıkla gözlemlenebilmektedir. Şekilden genel olarak fonksiyonların sinüsodial bir yapı gösterdikleri rahatlıkla söylenebilir. Ancak incelenen birey (istasyon) sayısı arttıkça fonksiyonların genel seyirleri hakkında bilgi edinilmesi de kompleks bir hal alacak ve zorlaşacaktır. Bu durum periyodik olmayan veriler için daha karmaşık bir hal alabilir. Burada kalın çizgi ile ortalama fonksiyon çizdirilmiştir. Şekilden 2000 ve 2005 yılları arasındaki yağış yoğunluğunun özellikle kış aylarına denk geldiği rahatlıkla gözlemlenebilmektedir. Yaz aylarında neredeyse tüm yıllarda özellikle 54-56 ıncı aylar arasına denk gelen 2004 yazında hiç yağış gözlemlenememiştir. Genel anlamda bakıldığında 12. aya karşılık gelen 2000 yılı sonu 2001 kışı tüm yıllara nisbeten yağışlar en düşük seviyede seyretmektedir. Ayrıca 36.aydan sonrada ortalama fonksiyonda zirve noktalarını bir doğru ile birleştirdiğimizde yağışların azalma eğiliminde olduğu rahatlıkla gözlemlenebilmektedir ki buralarda bile ortalama fonksiyonu yukarıya doğru çeken pozitif yönde uç bireysel meteoroloji istasyonları bulunmaktadır. Özellikle bu uç bireysel fonksiyonlara Aralık 2001- Ocak 2002 döneminde rastlanmaktadır.

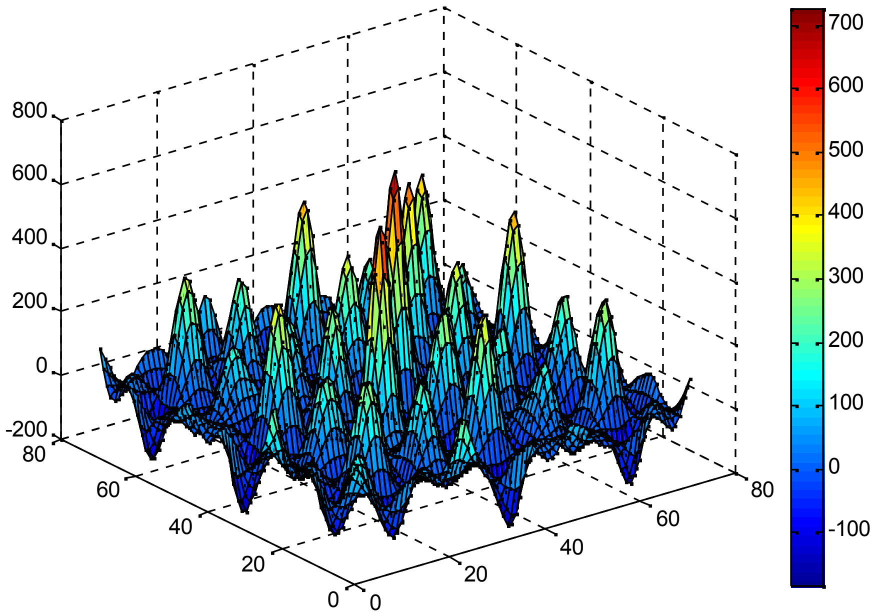
Pürüzlü Ceza Yaklaşımında λ şeklindeki düzgünleştirme parametresi belirlenirken Genelleştirilmiş Çapraz Geçerlilik (GCV) Metodu kullanılmıştır. Farklı Lamda değerleri için simülasyon yapılmış ve tüm istasyonlar açısından minimum GCV değerini veren Lamda değeri düzgünleştirme parametresi olarak belirlenmiştir. Şekil (4.3)de yapılan simülasyon sonuçları yer almaktadır.



Şekil(4.3): Farklı Lambda değerleri karşısında GCV değerleri

Bu çalışmada λ şeklindeki düzgünleştirme parametresine 10^(-4) ile 10 arasında 11 farklı değer verilmiştir. Şekilde en altta görülen, her bir istasyon için minimum GCV değerini veren grafik λ=4 değerine karşılık gelmektedir. Başlangıç noktası olarak 10^(-4) değerinin alınmasının nedeni uygulamalarla ilgili kapsamlı bir araştırma sonucu λ ile belirtilen düzgünleştirme parametresi için 10^{-4}, 10^{-3} ve 10^{-2} değerlerinin iyi çalıştığının gözlenmiş olmasıdır(Ramsay & Li;1998).

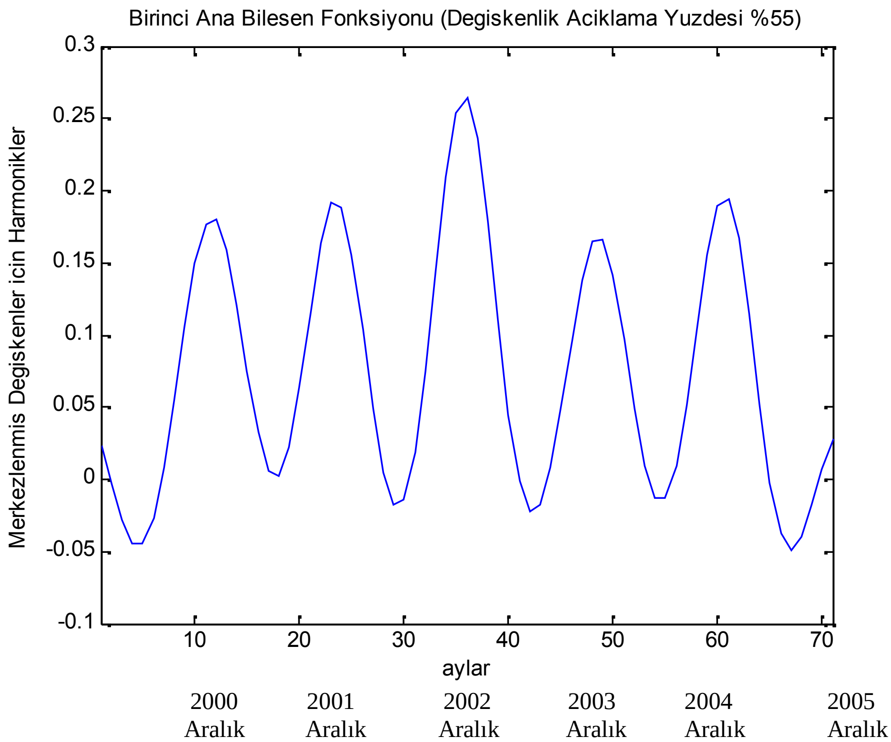
Şekil(4.4) de 71 farklı değişken için kovaryans yüzeyi oluşturulmuştur ve bu yüzeyin yüksekliği zamanın (veya ilgili değişkenin) her bir noktasında eğrilerin değişkenliğini ve birlikte değişiminin ölçüsünü vermektedir.



Şekil (4.4): Kovaryans Yüzeyi

Şekil (4.4)de de kovaryans yüzeyi 71 ayrık noktada 22 istasyon için inceleme yapıldığından dolayı oldukça karmaşıklaşmış ve farklı t zamanlarında gözlemlerin birlikte değişimleri çok zor incelenebilir hale gelmiştir. Hatta değişken sayısı arttıkça bu durum daha da karmaşıklaşmakta ve yüzey üzerindeki yükseklikler tanımlanamaz hale gelebilmektedir. Kovaryans yüzeyinde dikey köşegen üzerinde bulunan belli bölgelerde bir diğer ifadeyle zaman noktaları arasında değişkenlikte bir artış olduğu görülebilmekte ancak bu bölgeler net olarak çok zor tespit edilebilmektedir.

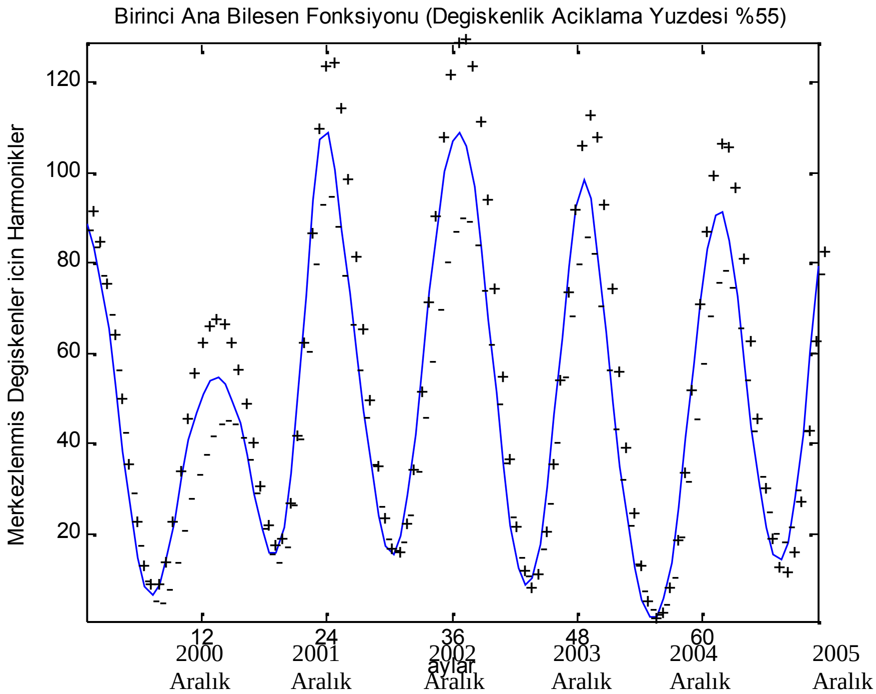
Kovaryans yüzeyinin yorumlanmasının güç olduğu durumlarda, ki bu durum çalışma sahası genişledikçe daha çok ortaya çıkmaktadır, ana bileşen fonksiyonundan yararlanılır. Şekil (4.5)de bireysel fonksiyonlar arasındaki değişkenliği açıklamaya yönelik olarak birinci ana bileşen fonksiyonu verilmiştir.



Şekil (4.5): Birinci Ana Bileşen Fonksiyonu

Ana bileşen fonksiyonunda zirve noktalarının her yılın kış aylarına denk geldiği rahatlıkla gözlemlenebilmektedir. Kış aylarında özellikle 2002-2003 (34-38 ayları arası) kışında meteoroloji istasyonları arasında yağışların değişkenliğinde bir artış olduğu rahatlıkla gözlemlenebilmektedir. Verilen ilgili kovaryans yüzeyi ile kıyaslandığında ana bileşen fonksiyonu çok daha rahatlıkla yorumlanabilmektedir. Ana bileşen fonksiyonuna bakar bakmaz %55 değişkenlik açıklama gücü ile yağış açısından eğriler arasındaki birinci temel değişkenliğin kış aylarından kaynaklandığı görülmektedir. Bir diğer deyişle veriler arasındaki değişkenliğin birinci modu kış ayları değişkenliğidir.

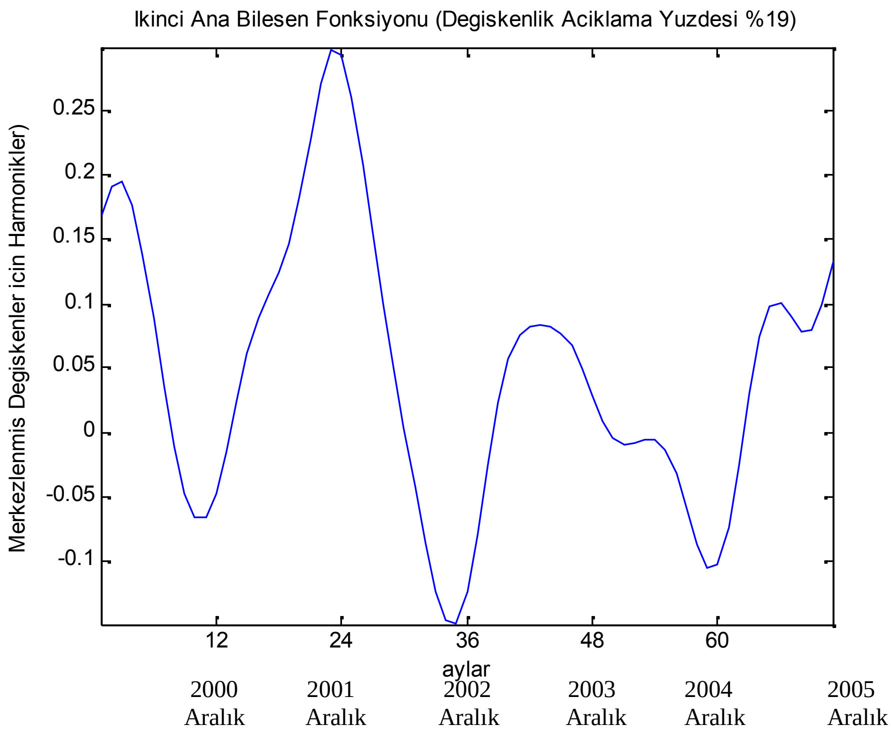
Ana bileşen fonksiyonu incelendikten sonra özellikle ana bileşen fonksiyonunun yorumlanmasının güç olduğu durumlarda alternatif yardımcı yöntemler olarak kullanılan ortalama fonksiyona ana bileşen fonksiyonlarının bir sabitle çarpanının eklenmesi ve çıkarılmasının etkileri Şekil (4.6)da verilmektedir.



Şekil (4.6): Ana Bileşen Fonksiyonu ve Ortalama Fonksiyonu Karşılaştırılması

\+ ve – noktalar ortalama fonksiyona ana bileşen fonksiyonunun belirli bir sabitle çarpanının eklenmesinin ve çıkarılmasının etkilerini göstermektedir. Düz çizgi ile verilen fonksiyon ise ortalama fonksiyonudur. + ve – noktalar ortalama fonksiyonundan ne kadar uzaksa ortalamadan sapmaların o kadar yüksek olduğu bu grafik yardımıyla da gözlemlenebilir. Ana bileşen fonksiyonuna benzer şekilde bu şekilden de özellikle yağışların fazla olduğu kış aylarında sapma çok net bir biçimde görülebilmektedir. Özellikle 2002-2003 (34-38 ayları arası) kışında sapma maksimuma ulaşmaktadır. Bu grafik de

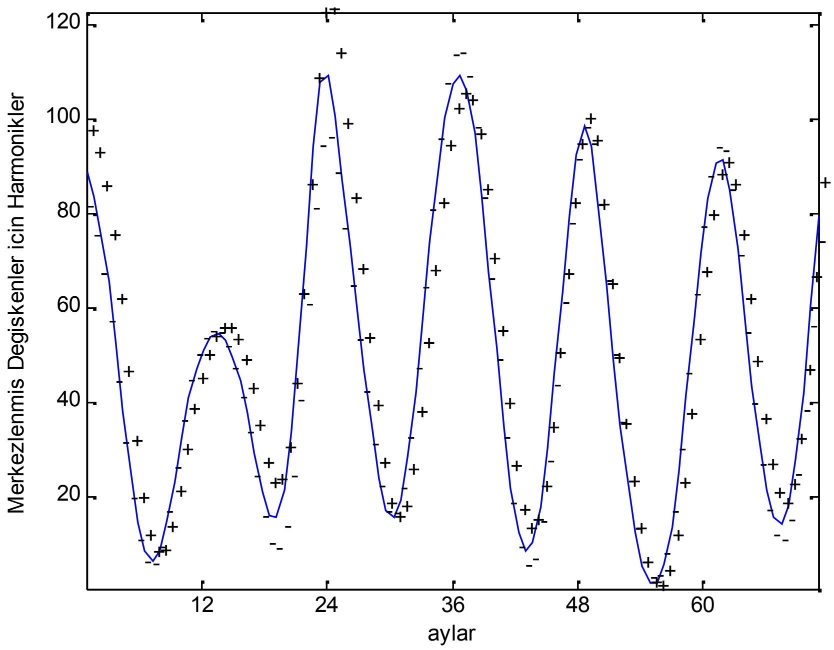
Şekil(4.5) de verilen birinci ana bileşen fonksiyonuyla ilgili yapılan yorumlamaları doğrulamaktadır.



Şekil (4.7): İkinci Ana Bileşen Fonksiyonu

Birinci ana bileşene ortogonal olan ikinci ana bileşen fonksiyonu Şekil (4.7)de verilmektedir. İkinci ana bileşen ile ilgili olarak birinci ana bileşen fonksiyonunda olduğu gibi çok rahatlıkla mevsimsel bir yorumlama yapılamamaktadır. Burada Aralık 2001- Ocak 2002 gibi fonksiyonda bir zirve noktası oluşmuştur. Ama doğrudan tüm kış veya yaz aylarında değişkenliğin fazla olduğu söylenemez. Yorumlamayı kolaylaştırmak için ana bileşen fonksiyonu ortalama fonksiyonu ile karşılaştırılacak ve ana bileşen skorları incelenecektir.

İkinci Ana Bileşen Fonksiyonu (Değişkenlik Açıklama Yüzdesi %19)



2000 Aralık | 2001 Aralık | 2002 Aralık | 2003 Aralık | 2004 Aralık | 2005 Aralık

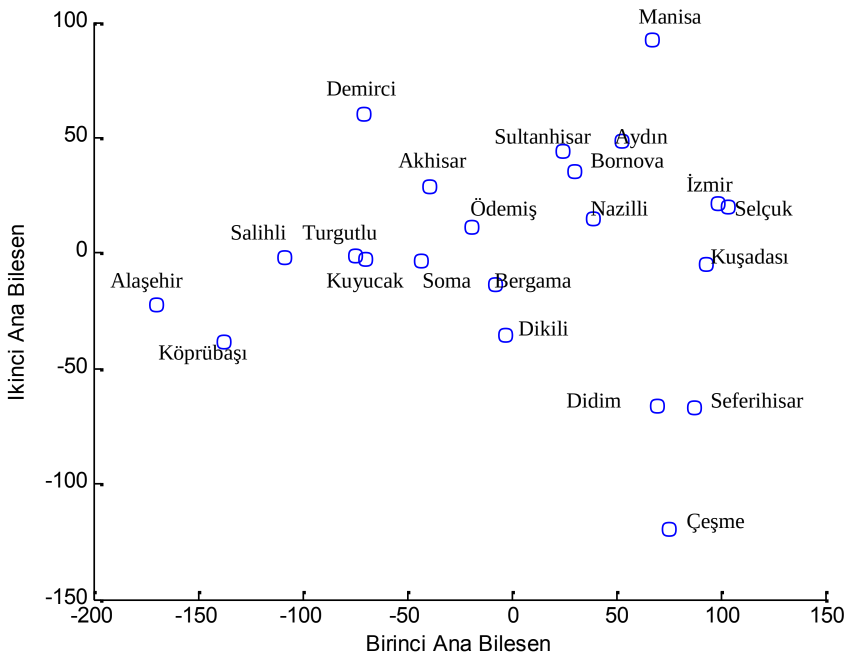
Şekil (4.8): Ana Bileşen Fonksiyonu ve Ortalama Fonksiyonu Karşılaştırılması

Şekil(4.8)de ortalama fonksiyonu ikinci ana bileşenin bir sabitle çarpanı ile karşılaştırıldığında yine Aralık 2001-Ocak 2002 arasında ortalama fonksiyondan sapmalarda artma olmuştur diğer bölgelerde ise + ve – ler neredeyse ortalama fonksiyonunun üzerinde seyretmektedir. İkinci ana bileşende değişkenlikteki en etkili artışın Aralık 2001-Ocak 2002 kışında olduğu görülmektedir. Ana Bileşen skorlarının yorumlanmasıyla burada en çok hangi bireysel fonksiyonun etkili olduğu görülebilecektir.

Ana bileşenlerin yorumlanmasına yardımcı olacak ikinci yaklaşım ana bileşen skorlarının incelenmesidir. Ana bileşen skor değerleri Tablo(4.1)de ve skor değerlerinin grafiksel gösterimi Şekil(4.9)da verilmektedir. Burada ilk iki ana bileşen dikkate alınmıştır.

Tablo (4.1): Ana Bileşen Skorları

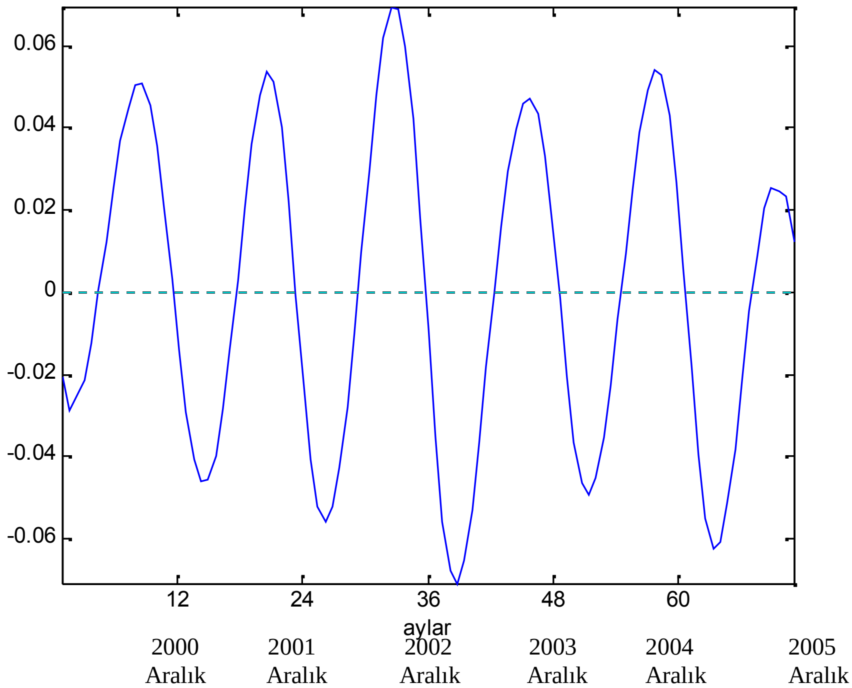
İstasyonlar	Birinci Ana Bileşen	İkinci Ana Bileşen
SOMA	-42.75	-3.38
KÖPRÜBAŞI	-137.66.	-38.61
TURGUTLU	-74.57	-1.00
ALAŞEHİR	-170.16	-22.31
KUYUCAK	-69.36	-2.77
DİKİLİ	-2.45	-35.67
AKHİSAR	-39.45	28.90
MANİSA	67.48	92.17
İZMİR	99.18	21.19
ÇEŞME	75.67	-119.75
KUŞADASI	92.78	-4.63
DİDİM	69.75	-66.49
AYDIN	54.43	48.83
BERGAMA	-7.74	-13.72
DEMİRCİ	-70.91	59.87
BORNOVA	29.98	35.61
SALİHLİ	-108.25	-1.95
SEFERİHİSAR	87.77	-67.09
ÖDEMİŞ	-18.92	11.48
SULTANHİSAR	24.44	44.42
SELÇUK	103.36	20.07
NAZİLLİ	36.37	14.82



Şekil (4.9) : Ana Bileşen Skorlarının Dağılımı

Ana bileşen skorları özellikle %55 değişkenlik açıklayabilme gücüne sahip birinci ana bileşen açısından incelendiğinde şekilden öncelikle Çeşme, Seferihisar, Kuşadası, Selçuk, İzmir ve Manisa'nın birinci ana bileşen üzerinde oldukça etkili olduğu, yüksek bir ana bileşen skoru değeri taşıdığı görülmektedir. Bunun tam tersi olarak da Alaşehir ve Köprübaşı da en yüksek negatif skor değerlerine sahiptirler. Gerçekten de en yüksek değişkenliğin kış döneminde meydana geldiğini vurgulayan birinci ana bileşen için, veriler incelendiğinde 2000-2005 yılları arasında genel olarak en yüksek ortalama yağışların Çeşme, Seferihisar, Kuşadası, Selçuk, İzmir ve Manisa'da ve en düşük ortalama yağışların ise Alaşehir ve Köprübaşı'nda meydana geldiği gözlenmiştir. 2005-2008 yılları arasındaki veriler incelenirse büyük olasılıkla İzmir ve yakın çevresi yüksek ortalama yağışları artık gösteremeyecektir. İkinci ana bileşen açısından bakıldığında en yüksek değişkenliğin bir diğer deyişle ortalamadan sapmaların Şekil (4.7) ve (4.8)den de görüldüğü üzere Aralık 2001-Ocak 2002 döneminde olduğu görülmektedir. Burada da ikinci ana bileşen üzerinde en etkili ilin Manisa olduğu gözlemlenmektedir. Veriler incelendiğinde bu dönemde gerçekten en yüksek ortalama yağışlar Manisa'da görülmüştür. Ayrıca Şekil (4.2) de Aralık 2001- Ocak 2002 dönemi için sapan bireysel fonksiyon olarak değerlendirilebilecek en üstte seyreden fonksiyon Manisa ilinin bireysel ortalama aylık yağış fonksiyonudur.

Fonksiyonel Veri Analizinin çok önemli bir avantajı da elde edilen bireysel fonksiyonların, ortalama fonksiyonlarının ve ana bileşen fonksiyonlarının sürekli arzu edilen dereceden türevi alınabilir fonksiyonlar olmasıdır. Bu çalışmada da olduğu gibi periyodik fonksiyonların bir avantajı da bunlar sonsuz türevi alınabilir fonksiyonlardır. Böylece fonksiyonlardaki büyüme hızı ve ivme gibi kavramlar rahatlıkla incelenebilmekte ve bu durumda analizciye çeşitli avantajlar sağlamaktadır. Bu çalışmada da birinci ana bileşen türev fonksiyonu incelenmiş ve Şekil (4.10)da verilmiştir. Çalışmada ana bileşen fonksiyonunun Şekil (4.10)da verilen aylara göre birinci türev fonksiyonu incelendiğinde değişkenlikteki pozitif ya da negatif değişimin hızı daha rahat bir şekilde görülebilmektedir.



Şekil (4.10): Birinci Ana Bileşen Birinci Türev Fonksiyonu

Özellikle aylar itibariyle değişkenlikteki ani iniş ve çıkışların gücü kıyaslanmak istendiğinde orjine göre çok daha rahat bir kıyaslama yapılabilmektedir. Bu durum özellikle iniş çıkışların daha hareketli olabildiği periyodik olmayan veriler için çok kullanışlıdır. Bilindiği üzere orjinin üzerindeki bölgelerde fonksiyonlar artma eğiliminde orijin noktasında maksimum veya minimum değerlere sahip olmakta ve negatif bölgede de fonksiyonlar azalma eğilimindedir.

5.Sonuç

Gözlemler ilgilenilen veri aralığı genişledikçe bir sayı dizisi olarak görülmektense, zamanın veya yakın ilgili değişkenin bir fonksiyonu olarak görülmeye başlanır. Özellikle Ramsay ve Silverman (1997)dan sonra ivme kazanan Fonksiyonel Veri Analizinin önemi her geçen gün giderek artmaktadır. Bunun nedeni gittikçe ilerleyen teknolojiyle birlikte elde edilen verilerin analizi için klasik istatistiksel yöntemlerin yetersiz kalması, Fonksiyonel Veri Analizinin interpolasyon aracılığıyla düzensiz örneklenen fonksiyonlarla ve kayıp verilerle uğraşılmasına imkan vermesi ve düzgünleştirme ile oluşturulan fonksiyonların türevlerinin de incelenebilmesi gibi görsel olarak da çoğu konuda araştırmacılara yardımcı olması ve bu açıdan veri analizine yeni bir bakış açısı getirmesidir. Fonksiyonel Veri Analizinde özellikle Ana Bileşenler Analizinde esas olan görselliktir. Fonksiyonel Ana Bileşenler Analizi ve Düzgünleştirilmiş Fonksiyonel Ana Bileşenler Analizi Klasik Ana Bileşenler Analizi ile karşılaştırıldığında, Klasik Ana Bileşenler Analizinin veri yapısı nedeniyle uygulanmasının uygun olmadığı durumlar haricinde bu iki yöntem arasındaki temel fark, bir diğer deyişle verileri fonksiyonel açıdan ele almanın sağladığı avantaj, bireylere ait fonksiyonlar, ortalama fonksiyonu, kovaryans yüzeyleri, ana bileşen fonksiyonları ve elde edilen fonksiyonların türevlerinin de incelenebilmesi ve bir çok görünmeyeni ortaya çıkarmadaki yeteneğidir.

Bu çalışmada da Ege Bölgesi'nde bulunan 22 farklı meteoroloji istasyonu için 2000-2005 yılları arası aylık ortalama yağış verileri öncelikle Baz Fonksiyon ve Pürüzlü Ceza Yaklaşımlarına göre 22 bireysel fonksiyona dönüştürülmüş ve çalışma sahası oldukça geniş olduğundan ve veriler arasındaki değişim yapısını ortaya koymak üzere Düzgünleştirilmiş Fonksiyonel Ana Bileşenler Analizinden yararlanılmıştır. Bu çalışmada aynı zamanda incelenen gözlem noktası sayısı ele alınan meteoroloji istasyonu sayısından büyük olduğu, bir diğer deyişle, örnek hacmi değişken sayısından küçük olduğu için veriler arasındaki değişkenlik yapısını ortaya koymak üzere Klasik Çok Değişkenli Ana Bileşenler Analizi kullanmak zaten uygun değildir.

Düzgünleştirilmiş Fonksiyonel Ana Bileşenler Analizi ile kovaryans yüzeyinden ve bireysel fonksiyonları tek tek inceleyerek ortaya çıkarılması çok güç olan veriler arasındaki değişkenlik yapısı incelenen ana bileşen fonksiyonları yardımıyla çok daha rahat gözlemlenebilmiştir. Burada %55 değişkenlik açıklama gücüne sahip birinci ana bileşen fonksiyonu yardımıyla özellikle kış aylarında istasyonlar arasındaki yağışların değişkenliğinin yüksek olduğu ilgili şekillerden rahatlıkla gözlemlenebilir. Böylece veriler arasındaki birinci temel değişim modunun kış ayları değişkenliği olduğu söylenebilir. İkinci temel değişim modunda ise Aralık 2001-Ocak 2002 dönemi çok etkili olmuştur. Bu değişim modu üzerindeki en etkili ilin ise Manisa olduğu gözlemlenmiştir.

Fonksiyonel Veri Analizinde verilerin bireysel fonksiyonlar olarak ele alınması ve belirli bir matematiksel formda verilebilmesi hem fonksiyonların türev fonksiyonlarının incelenebilmesine ve hem de yeni görünmeyenleri ortaya çıkarmaya faydalı olabilmektedir. Örneğin Ramsay ve Silverman'ın (2005: 2) çocuklarla ilgili büyüme verileri ile ilgili bir araştırmalarında büyüme verileri incelendiğinde elde edilen ikinci türev fonksiyonlarından bireysel fonksiyonları inceleyerek çıkarılması güç olan ergenlik büyüme atakları rahatlıkla ortaya konulmuştur. Bu çalışmada da birinci ana bileşenin türev fonksiyonunun incelenmesi ile bireysel fonksiyonlarda gözlenmesi oldukça güç olan çok küçük iniş çıkışların bile rahatlıkla yakalanabildiği, türevlerin negatif ve pozitif olduğu bölgeler ile türev fonksiyonunun sıfır değerini aldığı asıl fonksiyonun maksimum ve minimum noktaları görsel olarak rahatlıkla gözlemlenebilmektedir.

KAYNAKÇA

BENKO M. (2004). *Functional Principal Components analysis, Implementation and Applications.* A Master Thesis. Humboldt University Center of Applied Statistics and Economics, Berlin.

CASTRO P. E, LAWTON W. H., & SYLVESTRE E. A. (1986) “Principal Modes Of Variation for Processes with Continuous Sample Curves”, *Technometrics,* 28(4).

CRAVEN P. & WAHBA G. (1979). “Smoothing Noisy Data with Spline Functions”, *Numerishe Mathematik.* 31.

EUBANK R.L (1985). “Diagnostic for Smoothing Splines”, *Journal of the Royal Statistical Society: Series B.,* 47(2).

GREEN.P.J., & SILVERMAN B.W. (1994). *Nonparametric Regression and Generalized Linear Models: A Roughness Penalty Approach*. Chapman & Hall:London.

HARDLE W. (1997). *Applied Nonparametric Regression*, Cambridge University Press:USA

HURVİCH C.M, SİMONOFF J.S & TSAIC-L. (1997). “Smoothing Parameter Selection in Nonparametric Regression Using an Improved Akaike Information Criterion”, *Journal of the Royal Statistical Society*: *Series B.* 60(2).

HUTCHİNSON M.F, & de HOOG F.R. (1985). “Smoothing Noisy Data with Spline Functions”, *Numerishe Mathematik.* 47.

RAMSAY J. O. , LI X. (1998). “Curve Registration”, *Journal of the Royal Statistical Society: Series B*, 60(2).

RAMSAY, J. O., & DALZELL C. (1991). “Some Tools For Functonal Data Analysis”, *Journal of the Royal Statistical Society*: *Series B.,*53 (3)

RAMSAY J.O., & SILVERMAN B.W. (1997). *Functional Data Analysis*. Springer – Verlag: New York.

RAMSAY J.O., & SILVERMAN B.W. (2005). *Functional Data Analysis*. *Second Edition.* Springer : USA

RAZ J., TURESTY B., & FEIN G. (1989). "Selecting the Smoothing Parameter for Estimation of Slowly Changing Evoked Potential Signals", *Biometrics*, 45.

SILVERMAN B. W. (1995). "Incorporating Parametric Effects into Functional Principal Components Analysis", *Journal of the Royal Statistical Society*: *Series B.*,57(4).

SILVERMAN B.W.(1996). "Smoothed Functional Principal Component Analysis By Choice Of Norm", *The Annals of Statistics*, 24(1)

ULBRICHT J.(2004*). Representing Functional Data as Smooth Functions*. A Master Thesis, Humboldt University Institute of Statistics and Econometrics, Berlin.

WEI H.W. (2005). "The Smoothing Parameter, Confidence Interval and Robustness for Smoothing Splines", *Nonparametric Statistics,* 00(0).