

**ANKARA ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ**

YÜKSEK LİSANS TEZİ

**MEVSİMSSEL ZAMAN SERİSİ YÖNTEMLERİNİN ÇEVRE VERİLERİNE
UYGULAMASI: KOCAELİ ÖRNEĞİ**

Faruk ÇELİKKANAT

İSTATİSTİK ANABİLİM DALI

**ANKARA
2018**

Her hakkı saklıdır

TEZ ONAYI

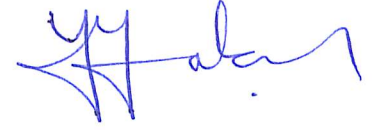
Faruk ÇELİKKANAT tarafından hazırlanan “**Mevsimsel Zaman Serisi Yöntemlerinin Çevre Verilerine Uygulaması: Kocaeli Örneği**” adlı tez çalışması 21/06/2018 tarihinde aşağıdaki jüri tarafından oy birliği ile Ankara Üniversitesi Fen Bilimleri Enstitüsü İstatistik Anabilim Dalı’nda **YÜKSEK LİSANS TEZİ** olarak kabul edilmiştir.

Danışman : Prof. Dr. Yılmaz AKDİ
Ankara Üniversitesi İstatistik Anabilim Dalı



Jüri Üyeleri:

Başkan: Doç. Dr. Yeliz YALÇIN
Gazi Üniversitesi Ekonometri Anabilim Dalı



Üye : Prof. Dr. Yılmaz AKDİ
Ankara Üniversitesi İstatistik Anabilim Dalı



Üye : Doç. Dr. Rukiye DAĞALP
Ankara Üniversitesi İstatistik Anabilim Dalı



Yukarıdaki sonucu onaylarım.

Prof. Dr. Atila YETİŞEMİYEN
Enstitü Müdürü

ETİK

Ankara Üniversitesi Fen Bilimleri Enstitüsü tez yazım kurallarına uygun olarak hazırladığım bu tez içindeki bütün bilgilerin doğru ve tam olduğunu, bilgilerin üretilmesi aşamasında bilimsel etiğe uygun davrandığımı, yararlandığım bütün kaynakları atıf yaparak belirttiğimi beyan ederim.

21/06/2018



Faruk ÇELİKKANAT

ÖZET

Yüksek Lisans Tezi

MEVSİMSSEL ZAMAN SERİSİ YÖNTEMLERİNİN ÇEVRE VERİLERİNE UYGULAMASI: KOCAELİ ÖRNEĞİ

Faruk ÇELİKKANAT

Ankara Üniversitesi
Fen Bilimleri Enstitüsü
İstatistik Anabilim Dalı

Danışman: Prof. Dr. Yılmaz AKDİ

Hava kirliliği alınan tüm önlemlere rağmen dünya genelinde önde gelen sorunlardan biridir. Bu kirliliğin sebeplerinden başlıca olanları insan kaynaklı sebeplerdir. Bu sebepler yok edilebilir ya da bazı önlemler alınarak etkileri azaltılabilir. Böylece, hava kirliliğinin öngörülmesi, ilgili tüm kurumlar için önemli bir konudur. Bu çalışmada, hava kirliliği oranlarının gelecekteki değerlerini öngörmek için iki farklı model yaklaşımı ele alınmıştır. Bunlardan biri Box-Jenkins modelleri olarak da bilinen ARIMA modelleri, diğeri ise klasik modellere alternatif olarak trigonometrik zaman serisi modelidir. Çalışmada ana hava kirliliği ölçüm endeksi olarak aylık ortalama Partikül Madde miktarı PM_{10} kullanılmıştır. Türk imalat sanayinin %13'üne sahip Türkiye 'nin en önemli sanayi kenti olan Kocaeli, temsili şehir olarak seçilmiştir. Bu nedenle, veriler 2005'den 2018'e kadar Kocaeli şehrindeki hava izleme istasyonlarından aylık olarak elde edilmiştir. Zaman serisi verilerin durağanlığı Augmented Dickey Fuller birim kök testi ile kontrol edilmiştir.

Haziran 2018, sayfa 80

Anahtar Kelimeler: Hava kirliliği, PM_{10} , Mevsimsellik, Durağanlık, Öngörü

ABSTRACT

Master Thesis

APPLICATIONS OF SEASONAL TIME SERIES METHODS TO ENVIRONMENTAL DATA: KOCAELI CASE

Faruk ÇELİKKANAT

Ankara University
Graduate School of Natural and Applied Sciences
Department of Statistics

Supervisor: Prof. Dr. Yılmaz AKDİ

Around the world air pollution is a leading social problem despite all the precautions. Of the reasons, human-induced ones rank the first. These can be prevented or at least reduced by some measures. Thus, forecasting air pollution is an important issue for all related agencies. In this study, in order to forecast future values of air pollution rates, two different model approaches were considered. The one of these approaches is ARIMA models also known as Box-Jenkins models and the another one is a trigonometric time series model as an alternative to the traditional models. As a main air pollution measurement index, monthly average amounts of Particulate Matter, PM_{10} was used. Being the most important industrial city of Turkey with the 13% share of Turkish manufacturing industry, Kocaeli was selected as a representative city. Hence, data was obtained monthly from the air monitoring stations in the city of Kocaeli between the period of 2005 and 2018. The stationarity of time series data was checked by Augmented Dickey Fuller unit root test.

June 2018, 80 pages

Key Words: Air Pollution, PM_{10} , Seasonality, Stationarity, Forecasting.

TEŞEKKÜR

Öncelikle bana desteğini esirgemeyen saygıdeğer danışman hocam Prof. Dr. Yılmaz AKDİ (Ankara Üniversitesi İstatistik Anabilim Dalı)'ye, tez çalışmam esnasında desteğini esirgemeyen Piyade Albay Cemalettin İLHAN'a ve diğer mesai arkadaşlarıma teşekkür ediyorum.

Faruk ÇELİKKANAT

Ankara, Haziran 2018

İÇİNDEKİLER

TEZ ONAY SAYFASI	
ETİK	i
ÖZET	ii
ABSTRACT	iii
TEŞEKKÜR	iv
KISALTMALAR DİZİNİ	vi
ŞEKİLLER DİZİNİ	vii
ÇİZELGELER DİZİNİ	viii
1. GİRİŞ	1
2. DURAĞAN ZAMAN SERİLERİ	8
2.1 Temel Kavramlar	8
2.1.1 Zaman serileri	8
2.1.2 Durağanlık	9
2.1.3 Otokovaryans ve otokorelasyon	10
2.1.4 Regresyon	12
2.1.5 Öngörü	14
2.2 Bazı Durağan Zaman Serisi Modelleri	15
2.2.1 Hareketli ortalama (MA) serileri	15
2.2.2 Otoregresif (AR) serileri	22
2.2.3 ARMA serileri	35
3. DURAĞAN OLMAYAN ZAMAN SERİLERİ	44
3.1 Birim Kök Testleri	44
3.1.1 Dickey-Fuller birim kök testi	44
3.1.2 Phillips-Perron birim kök testi	51
3.2 Mevsimsel Birim Kök Testleri	53
3.2.1 HEGY testi	55
3.3 Periodogramlar	57
4. BULGULAR ve TARTIŞMA	58
5. SONUÇ VE ÖNERİLER	76
KAYNAKLAR	78
ÖZGEÇMİŞ	80

KISALTMALAR DİZİNİ

PM ₁₀	Partiküler Madde
MA	Moving Average (Hareketli Ortalama)
AR	AutoRegressive (Oto regresif)
SAR	Seasonal AutoRegressive (Mevsimsel Oto regresif)
ARMA Ortalama)	AutoRegressive Moving Average (Oto regresif Hareketli
ARIMA	AutoRegressive Integrated Moving Average (Oto regresif Bütünleşik Hareketli Ortalama)
$E()$	Beklenen Değer
$Var()$	Varyans
$Cov()$	Kovaryans
$\gamma()$	Otokovaryans Fonksiyonu
$\rho()$	Otokorelasyon Fonksiyonu
$\phi()$	Kısmi otokorelasyon Fonksiyonu
$det()$	Matrisin Determinantı
ACF	Otokorelasyon Fonksiyonu
PACF	Kısmi Otokorelasyon Fonksiyonu
WN	White Noise (Beyaz Gürültü Serisi)
TÜİK	Türkiye İstatistik Kurumu
EKK	En Küçük Kareler
HEGY	Hylleberg,Engle,Granger ve Yoo
DHF	Dickey, Hasza ve Fuller
ADF	Adjusted Dickey-Fuller
PP	Phillips-Perron
MSE	Mean Square of Error (Hata Kareler Ortalaması)
$\hat{\cdot}$	Parametrenin Tahmini

ŞEKİLLER DİZİNİ

Şekil 2.1 $AR(1)$ zaman serisi modeline uygun kısmi otokorelasyon grafiği	11
Şekil 2.2 $MA(1)$ ve $MA(2)$ modellerine ilişkin otokovaryans ve otokorelasyon fonksiyonları.....	16
Şekil 2.3 $\theta = 0.3$ olan bir $MA(1)$ modelinin otokovaryans ve otokorelasyon fonksiyonları	17
Şekil 2.4 $\theta_1 = 0.4$ ve $\theta_2 = -0.2$ olacak şekilde türetilen bir $MA(2)$ modelinin otokovaryans ve otokorelasyon fonksiyonları	18
Şekil 2.5 Rasgele üretilen $MA(1)$ ve $MA(2)$ modellerine ait kısmi otokorelasyon fonksiyonları.....	19
Şekil 2.6 $X_t = 0.8e_{t-1} + e_t$ şeklinde verilen $MA(1)$ modeline ait otokorelasyon ve kısmi otokorelasyon fonksiyonları	20
Şekil 2.7 $X_t = 0.8e_{t-1} + e_t$ şeklinde verilen $MA(1)$ modeline uygun üretilen 100 birimlik veri setine ait otokorelasyon ve kısmi otokorelasyon fonksiyonları	21
Şekil 2.8 Farklı ρ değerleri için zaman serisi, ACF ve PACF grafikleri.....	24
Şekil 2.9 $X_t = 1.2X_{t-1} - 0.3X_{t-2} + e_t$ modelinin otokorelasyon grafiği.....	30
Şekil 2.10 Kısmi otokorelasyon grafiği	31
Şekil 2.11 α_1 ve α_2 'nin değişik değerleri için $AR(2)$ modeline ilişkin grafikler	32
Şekil 2.12 Farklı α değerleri için $AR(4)$ modeline ilişkin grafikler	34
Şekil 2.13 $X_t = 1.2X_{t-1} - 0.3X_{t-2} + e_t + 2e_{t-1}$ modeline ait otokorelasyon ve kısmi otokorelasyon grafikleri.....	38
Şekil 2.14 Üretilen $ARMA(2,1)$ modeline ait otokorelasyon ve kısmi otokorelasyon grafikleri	39
Şekil 2.15 $X_t = 1.7X_{t-4} - 0.72X_{t-8} + e_t$ modeline ait otokorelasyon ve kısmi otokorelasyon grafikleri	41
Şekil 2.16 $X_t = 1.7X_{t-4} - 0.72X_{t-8} + e_t$ modeline uygun üretilen veriye ait otokorelasyon ve kısmi otokorelasyon grafikleri.....	42
Şekil 2.17 $X_t = 0.8X_{t-12} + e_t$ modeline ait otokorelasyon ve kısmi otokorelasyon grafikleri	43
Şekil 2.18 $X_t = 0.8X_{t-12} + e_t$ modeline uygun üretilen veriye ait otokorelasyon ve kısmi otokorelasyon grafikleri.....	43
Şekil 3.1 A firmasının ay bazında 5 yıllık kar/zarar durumuna ait verilerinin zaman serisi grafikleri	50
Şekil 4.1 Kocaeli iline ait aylık PM_{10} değerleri (Ocak 2005-Aralık 2017).....	61
Şekil 4.2 Periodogramların sıklıklara karşı grafiği (Kocaeli, 2005:01-2017:12).....	68
Şekil 4.3 Öngörü değerleri ve gerçekleşen değerler	74

ÇİZELGELER DİZİNİ

Çizelge 2.1 $X_t = 0.8e_{t-1} + e_t$ modeline uygun üretilen veri setinin otokorelasyon ve kısmi otokorelasyon değerleri	21
Çizelge 2.2 Üretilen 50 veriye ait bilgiler	26
Çizelge 2.3 ARMA (2.1) modeline ait otokorelasyon ve kısmi otokorelasyon değerleri	38
Çizelge 2.4 $X_t = 1.2X_{t-1} - 0.3X_{t-2} + e_t + 2e_{t-1}$ modeline uygun üretilen veri setinin otokorelasyon ve kısmi otokorelasyon değerleri	38
Çizelge 2.5 $X_t = 1.7X_{t-4} - 0.72X_{t-8} + e_t$ modelinin otokorelasyon ve kısmi otokorelasyon değerleri	40
Çizelge 2.6 $X_t = 1.7X_{t-4} - 0.72X_{t-8} + e_t$ modeline uygun üretilen veri setine ait otokorelasyon ve kısmi otokorelasyon değerleri	41
Çizelge 3.1 A firmasının ay bazında 5 yıllık kar/zarar durumu (* Yüz Bin TL)	49
Çizelge 3.2 $(X_t - \mu) = \alpha_0 + \alpha X_{t-1} + e_t$ modeline ilişkin parametre tahminleri	50
Çizelge 3.3 ∇X_t 'nin X_{t-1} üzerine regresyonundan elde edilen istatistiki değerler	51
Çizelge 3.4 A firması kar/zarar verisi Phillips-Perron testi	52
Çizelge 3.5 Üretilen veri seti	53
Çizelge 3.6 $(X_t - \mu) = \alpha(X_{t-4} - \mu) + e_t$ modeline uygun üretilen veriye ait parametre tahminleri	54
Çizelge 4. 1 Aylık ortalama PM_{10} değerlerine ilişkin (1.1) de verilen model sonuçları	61
Çizelge 4.2 Değişik zaman serisi modellerine ilişkin AIC istatistiklerinin değerleri	62
Çizelge 4.3 ADF ve PP Birim Kök Test Sonuçları	63
Çizelge 4.4 Eşitlik (4.1) modeline ait parametre tahminleri PROC ARIMA	69
Çizelge 4.5 Eşitlik (4.1) modeline ait parametre tahminleri PROC REG	69
Çizelge 4.6 Bir yıllık öngörü değerleri (2018:01-2018:12)	70
Çizelge 4.7 Eşitlik (4.1) modeline ait en büyük periodogram değeri, sıklıkları, periyotları ve V istatistik değerleri	70
Çizelge 4.8 Eşitlik (4.6) modeline ait parametre tahminleri	72
Çizelge 4.9 Eşitlik (4.7) modeline ait parametre tahminleri	73
Çizelge 4.10 Öngörü değerleri ve gerçekleşen değerler	74

1. GİRİŞ

Herhangi bir değişkenin belirli zaman aralıklarında aldığı değerler, değişkenin daha sonra alacağı değerlerin öngörülmesinde kullanılır. Bir rasgele değişkenin birim zaman aralıklarında aldığı değerler kümesi bir zaman serisi olarak adlandırılır. Zaman serilerine ekonomi, tarım, hayvancılık, sağlık ve meteoroloji gibi bilimin değişik alanlarında örneklerine rastlanabileceği gibi bilimin hemen hemen her alanında uygulamalarına rastlanabilir. Bir şirketin aylık tüketim harcamaları, yıllık ihraç edilen gübre miktarı veya hava kirliliğinin aylık gözlenen değerleri zaman serilerine örnek olarak verilebilir. Zaman serileri genellikle istatistiğin bazen de ekonometrinin bir uygulama alanı olarak karşımıza çıkar.

Hayatın her alanında, bireysel ve kurumsal her seviyede geleceğe yönelik bilinmeyen parametreleri minimum hata ile tahmin ederek öngörüde bulunmak ve bu öngörülere dayanarak gelecek dönemler için planlama yapmak hayati öneme sahiptir. Bir ülkenin gelecekteki 5 yıllık işsizlik oranlarını öngörüp buna uygun yatırımlar yaparak bu oranı düşürmeye çalışması veya bir firmanın olabilecek satış miktarını kestirerek üretim kademelerinde gerekli tedbirleri alması hem kaynakların uygun kullanılması hem de o ülke ya da firmanın istikrarı ve devamlılığı için vazgeçilmezdir. Bu tür öngörülerin yapılabilmesi için genellikle zaman serileri teknikleri kullanılır.

Herhangi bir zaman serisi verileri kullanılarak istatistiki sonuç çıkarımı yapabilmek için önce temel varsayımların sağlanması gerekir. Eğer varsayımlar sağlanamıyorsa herhangi bir dönüşüm yada bir teknik ile varsayımlar sağlatılmaya çalışılır. Zaman serilerinde en belirgin varsayımlardan biri serinin durağanlığıdır. Bu varsayımın sağlanmadığı durumda yapılacak öngörüler hatalı politikaların belirlenmesine neden olabilir.

Sonuçlarının ciddiyeti ve önemi gereği hava kirliliği birçok devletin ve ilgili kuruluşlarının üzerinde çalıştığı bir konudur. Aslında hava kirliliği bir ülke yada kurum sorunu olmaktan ziyade bütün dünyanın bir sorunu olup özellikle son zamanlarda Birleşmiş Milletler ve Dünya Sağlık Örgütü gibi uluslararası kuruluşlar bu sorunun

üzerinde arařtırmalar yapmaktadır. Elde edilen bulgular düzenli bir řekilde bütün dünya kamuoyu ile paylaşılmaktadır.

Tüm dünyada olduđu gibi ölkemizin de en önemli sorunlarından biri hava kirliliđidir. Hava kirliliđi çeřitli sađlık sorunlarını da beraberinde getirmektedir. Bununla birlikte artan sađlık harcamaları, ölümler, eldeki iş gücünün etkin kullanılamaması ve üretim kaybı, sađlıksız bir nesil ve kirlı bir çevre gibi ciddi problemlere neden olabilmektedir.

Bu çalışmada, ölkemizin en büyük sanayi merkezlerinde Kocaeli’ ndeki hava kirliliđi problemi zaman serileri teknikleri ile incelenmeye çalışılacaktır. Hava kirliliđinin önemli göstergelerinden Partiküler Madde (PM_{10}) indeksi kullanılarak öncelikle bir öngörü modeli elde edilecek ve bu model yardımıyla gelecek dönemlere ait kestirimler yapılacaktır. Bu kestirimlerin yapılması genel anlamda ölkemizde ve özellikle sanayi merkezlerindeki kirliliđin azaltılmasına yönelik hazırlanacak proje ve alınacak tedbirlere yön verilmesine imkan sađlayacaktır.

Hava kirliliđi gibi birçok çevresel durum ve olaylarda mevsimsellik hemen göze çarpmaktadır. Yani yıl içinde bazı zamanlarda daha etkili olurken etkisini yitirdiđi periyotlara da rastlanmaktadır. Birim zaman aralıklarında gözlenen bu deđerler kullanılarak hava kirliliđinin ileriki dönemlerdeki boyutu hakkında bilgi sahibi olunur. Çalışmada mevsimselliđin belirgin olduđu hava kirliliđi problemi ele alınmıř, bu alanda yapılan çalışmalar da deđerlendirilerek zaman serileri yöntemleri ile birlikte “trigonometrik bir regresyon “ eřitliđinden de yararlanılmıřtır.

Türkiye’nin başkenti Ankara’nın bazı semtlerinde (Siteler, Yenimahalle, Tandođan ve Ařađı Ayrancı) SO_2 ve duman kirletici gözlenen deđerleri kullanılıp Box–Jenkins teknikleri ile modellenerek bir öngörü modeli elde edilmeye çalışılmıřtır. Duman serileri istasyonlara ve öteki kirleticiye göre farklılık gösterirken; SO_2 serilerinin birbirine benzer nitelikte oldukları görölmüřtür. Sonuçta istasyonlara ve diđer kirleticilere göre farklılık gösteren duman serileri için bir model önerilmesi zor iken, birbirine benzer nitelikte olduđu görölen SO_2 serileri için ARIMA (1,1,1) (0,1,1) modelinin uygun olduđu önerilmektedir (Yarcı 1991).

Pope vd. (1992), PM_{10} un semptomatik ve asemptomatik bir çocuk üzerindeki akut sađlık etkilerini iki farklı yaklaşımla lojistik regresyon yöntemini kullanarak araştırmışlardır. PM_{10} ve solunum yolu semptomları arasında pozitif bir ilişki gözlenmiştir.

Ostro vd (1999), PM_{10} ve günlük mortalite arasındaki ilişkiyi Poisson regresyonu ile incelemeye çalışmışlar, bu değişkenler arasında bir ilişkinin varlığı tespit edilmiştir.

Berktaş ve Bircan (2003), astım nedeniyle Atatürk Göğüs Hastalıkları ve Göğüs Cerrahisi Eğitim ve Araştırma Hastanesi acil servisine başvuran kişi sayısı ile PM_{10} , SO_2 ve meteorolojik değişkenlerin miktarı arasındaki ilişkiyi Pearson ve Spearman korelasyon ve Ki-kare testlerini kullanarak analiz etmiştir. Astım nedeniyle acil servise başvuranların sayısı ile SO_2 miktarı arasında istatistiksel olarak anlamlı düşük düzeyli bir ilişki tespit ederken, PM_{10} miktarı ile hastaneye kabul edilen kişi sayısı arasında istatistiksel olarak anlamlı bir ilişki elde edilmiştir.

Çiçek vd. (2004), Ankara Sıhhiye bölgesinde ölçülen 5 hava kirliliđi endeksini (SO_2 , PM_{10} , NO, NO_2 ve CO) incelemiş ve aşamalı regresyon yöntemini kullanarak hava kirliliđi ile iklim koşulları arasındaki düşük ve orta düzey bir ilişki tespit etmiştir.

Koutrakis vd (2005), Santiago, Şili'de $PM_{2.5}$, PM_{10} ve meteorolojik ölçümleri karma regresyon modeli ile kategorik verilere dönüştürerek incelemişler. Meteorolojik değişkenler ile partiküler madde miktarı arasında istatistiksel olarak anlamlı ilişkiler elde edilmiştir. Ayrıca havadaki partikül madde miktarı ile trafik yoğunluğu arasında önemli bir ilişki tespit edilmiştir.

Genç vd. (2010) hava kirliliđi endeksini öngörmek için Ankara'daki yerleşim yerlerinden toplanan PM_{10} ve SO_2 miktarları ile Hava Kirliliđi İndeksi (API) çoklu doğrusal regresyon yöntemi ile incelemeye çalışmışlardır.

Turner vd. (2011) ise cox nisbi bozulma regresyon modeli kullanarak uzun vadeli partiküler madde hava kirliliđi ile akciđer kanseri arasındaki ilişkiyi incelemişlerdir.

PM_{2,5} konsantrasyonlarındaki her 10 µg / m³’lük artışın akciğer kanseri tanısında % 15-27 artışa neden olduğu gözlenmiştir.

Son yıllarda doğalgaz kullanımında görülen artış özellikle şehir merkezlerinde hava kirliliğini önemli ölçüde etkilemesine rağmen sanayi şehirlerinde gözlenen değerler uluslararası standartların oldukça üzerindedir. Türk Toraks Derneği’nin yürüttüğü bir çalışmada Kasım 2016 – Ekim 2017 tarihleri arasında partiküler madde açısından sadece Rize’de Dünya Sağlık Örgütü’nün sağlık açısından izin verdiği sınırın aşılmadığını, Şırnak’ta yeterli ölçüm yapılmadığını, geri kalan 79 ilde ise hava kirliliğinin yaşandığı açıklanmıştır. Hava kirliliğine ait bu oranların büyük şehirlerde (İstanbul vb.) çok fazla olduğu belirtilmektedir.

Öztaş (2012), Erzurum ilindeki doğalgaz tüketimini Box – Jenkins ve “Bileşenlerine Ayırma Yöntemi” ile analiz etmiş ve bu oranlara ilişkin kestirim modelleri önermiştir.

Novikov vd. (2012) tarafından İsrail’de kamu hastanelerinin acil servislerindeki hasta sayılarını önceden kestirebilmek için Geçişli Poisson regresyonu ve zaman serileri analizini kullanarak bir modelleme çalışması yapılmış, elde edilen kestirim değerleri daha sonra yapılan gözlem değerlerine çok yakın olduğu gözlenmiştir. Bu çalışmada, günlük hasta sayıları, meteorolojik değerler, elektrik tüketimi ve hava kirliliği oranları kullanılmıştır.

Yağış miktarı verisinin zaman serileri ile analizi konusunda yapılan çalışmada doğrusal olmayan zaman serileri için yapay sinir ağları ve Box-Jenkins metotları kullanılarak iklim zaman serilerine ilişkin kestirimlerin elde edilmesi ve karşılaştırılması amaçlanmıştır. Ankara, İstanbul, Antalya, Erzurum ve Samsun illerine ait 1960-2012 yılları arasında aylık yağış miktarı verileri kullanılmış ve her iki metoda göre elde edilen sonuçlar 12 dönemlik kestirimlerine dayalı olarak karşılaştırılmıştır. Yapay sinir ağları metodunun Box-Jenkins yöntemine göre nispeten daha iyi kestirim değerlerine sahip olduğu değerlendirilmiştir (Kurt 2013).

Tomasetti vd. (2013), çalışmalarında birçok alerjinin ve solunum problemlerinin kaynağı olan hava partiküllerinin (polenler ve fungal sporlar) konsantrasyonlarının meteorolojik parametreler yoluyla tahmin etme imkanını araştırmışlardır. Bu kapsamda karmaşık bir yapay sinir ağı mimarisinin kullanımına dayanan Operasyonel Sınırlı Alan Meteorolojik modelinden faydalanılarak bir sinir ağı modeli kullanmışlar ve Orta İtalya'da geniş bir alanda bu partikül konsantrasyonlarının bir haritasını tatmin edici bir doğrulukla çıkarmışlardır.

Hava kirliliği ile ilgili bir diğer çalışmada Ivanov ve Gocheva- Ilieva çalışmalarında ortam hava kalitesi ile ilgili verilerin modellenmesi için mevsimsel otoregresif entegrasyonlu hareketli ortalama (SARIMA) ve yeni bir Generalized Path Seeker (GPS) Regularized Regresyon yöntemi olmak üzere iki tür tahmini modelleme tekniği kullanılmışlardır. Modeller, Shumen kasabesindeki birincil hava kirleticisi olan PM₁₀ partiküler madde için ölçülen veriler için üretilmiştir. Zaman serisi analizi, bir aylık periyot boyunca altı meteorolojik değişkene göre saatlik verilere dayanılarak gerçekleştirilmiştir. Oluşturulan modeller, kısa vadeli dört günlük öngörüler için kullanılmıştır. Elde edilen sonuçlar mevsimsel ARIMA stokastik modellenmesi ve uygulanabilirliği üzerine GPS yönteminin bazı avantajlarını göstermektedir. Bu, kentsel alanlarda olası kirlilik sorunlarını analiz etmek ve önlemek için yeni bir perspektif sunmaktadır (Ivanov ve Gocheva 2013).

Turgut ve Deniz (2015), Sıhhiye istasyonundan 1 Ocak 2010 ile 31 Ekim 2014 arasında alınan haftalık ölçümlerin ARIMA (3,0,0) yani AR (3) modeline uygun olduğunu ve Kasım 2014, Aralık 2014 ve Ocak 2015 ayı değerlerini tahmin ettiklerini belirtmişlerdir. Öngörüler sonucunda, PM₁₀ miktarının olası değerlerinin 83,21 µg / m³'e yakın olduğunu tespit ettiler. Buradan hareketle Ankara için gerekli önlemlerin alınmasını gerektiği sonucuna ulaşmışlardır.

Liu ve arkadaşları partikül hava kirliliğine maruz kalma ile tip 2 şeker hastalığı, glikoz ve glikozile hemoglobin düzeyleri arasındaki ilişki hakkında kesitsel bir araştırma yapmışlardır. Ülke çapında yapılan bu araştırma ile PM_{2.5}'e maruz kalınmasının tip 2 Diyabet riskini arttırdığı sonucuna ulaşmışlardır (Liu vd. 2016)

Shi vd. (2016), PM_{2.5}'in solunum yolu süpürasyonunun bir biyolojik işareti olan fraksiyonel ekshalasyon nitrik oksit (FeNO) üzerindeki etkilerini analiz etmek için Şangay, Çin'de bir panel veri anketi üzerinde çalışmıştır. Bu ilişkiyi değerlendirmek için doğrusal karışık-etki modeli kullandılar. PM_{2.5}'in FeNO'yu tetiklediğini sonucuna vardılar.

Bir diğer çalışmada Yurtseven ve Serengil (2017), İstanbul iline ait mevsimsel meteorolojik verileri kullanıp ileriye yönelik yağış tahminlerinde bulunarak bölgedeki ekosistemlerin sağlığı ve geleceği açısından gereken tedbirleri almayı ve bu konuda yapılması gerekenleri ortaya koymaya çalışmışlardır. Bu mevsimsel serinin ileriye dönük tahminini yaparken Mann-Kendall trend analizini kullanmışlardır. Bu analizin sonuçlarına göre Florya, Kireçburnu ve Kumköy bölgelerinde sonbahar mevsimi yağışlarında, Göztepe bölgesinde ise yaz mevsimi yağışlarında istatistiksel anlamda artan bir trend bulunmuştur.

Çalışma beş bölümden oluşmaktadır. Birinci bölüm giriş bölümüdür. Bu bölümde öncelikle zaman serisi mantığından ve istatistikten ekonomiye sağlıktan tarım ve hayvancılığa geniş bir yelpazede kullanım alanına sahip olduğundan bahsedilmiştir. Ayrıca hava kirliliğinin ve sonuçlarının hem ulusal hem de uluslararası önemine vurgu yapılmıştır.

Çalışmanın ikinci bölümünde zaman serileri ile ilgili temel kavramlara yönelik genel bilgiler verilerek durağan zaman serileri modellerinden ve bu serilere ait öngörü tahminlerine değinilmiştir.

Üçüncü bölümde durağan olmayan zaman serilerine ve öngörü tahminlerine değinilmiş, birim kök testleri, periodogramlar ve periyodik fonksiyonlardan bahsedilmiştir.

Dördüncü bölümde Kocaeli ilimize ait hava kirliliği indekslerinden PM₁₀ verilerini çeşitli zaman serileri modellerinde uygulayarak öngörüler elde edilmiş ve bu öngörülerini farklı

bir yol izleyerek elde etmeye çalıştığımız trigonometrik model ile ulaşılan sonuçlarla karşılaştırma yapılarak modeller arası benzerlikler ve farklılıklar görülmeye çalışılmıştır.

Beşinci bölüm ise çalışmanın sonuç ve tartışma bölümüdür. Bu bölüm elde edilen sonuçlara dayanarak hava kirliliğinin Kocaeli için ne yönde seyredeceği ve uygulanması gereken tedbirler şeklinde tavsiye niteliğinde bilgilendirme kısmıdır.

2. DURAĞAN ZAMAN SERİLERİ

Zaman serileri ortalamadan ve varyansdan sapma durumuna göre durağan ve durağan olmayan seriler diye iki ana başlık altında incelenir. Bir zaman serisinin ortalaması, varyansı ve kovaryansı zamana göre değişmiyorsa veya seri periyodik dalgalanmalardan arınmışsa bu seri durağan zaman serisi olarak adlandırılır (Fuller 1976). Durağan zaman serileri hareketli ortalama (Moving Average, MA) modeller, otoregresif (AutoRegressive, AR) modeller ve otoregresif hareketli ortalama (AutoRegressive Moving Average, ARMA) modellerden oluşmaktadır.

2.1 Temel Kavramlar

Bu kısımda zaman serileri ile ilgili temel kavramlar özetlenecektir. Hemen hemen bütün zaman serileri ile ilgili kaynaklarda bulunabilen bu kavramların ayrıntılarına girilmeyecektir.

2.1.1 Zaman serileri

Bir zaman serisi, en basit anlamda, birim zaman aralıklarında yapılan gözlemlerin bir topluluğudur. Aşağıda bir zaman serisi açıklanmaya çalışılmıştır.

Ω boş olmayan bir küme, U ise Ω nın bazı alt kümelerinden oluşan bir sınıf olmak üzere;

$$i) \Omega \in U$$

$$ii) \text{ Her } A \in U \text{ için } A^c \in U \text{ (} A^c \text{ A'nın tümleyenidir)}$$

$$iii) A_n \in U \text{ olmak üzere, } \bigcup_{n=1}^{\infty} A_n \in U, n=1,2,3,\dots$$

özelliklerini sağlıyorsa U 'ya Ω üzerinde bir sigma cebir, (Ω, U) ikilisine de ölçülebilir bir uzay denir. U 'nun her bir elemanı bir olayı ifade eder.

(Ω, U) ölçülebilir bir uzay ve U üzerinde tanımlanan

$$P: U \longrightarrow [0,1]$$

$$A \longrightarrow P(A)$$

P küme fonksiyonu

i) $P(A) \geq 0$, her $A \in U$

ii) $P(\Omega) = 1$

iii) A_n ' ler U da ayırık olayların dizisi ($A_i \cap A_j = \emptyset$, $i \neq j$) olmak üzere

$$P\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} P(A_n)$$

özelliklerini sağlıyorsa P ye bir olasılık ölçüsü, $P(A)$ değerine A olayının olasılığı ve (Ω, U, P) üçlüsüne de olasılık uzayı denir.

T herhangi bir indis kümesi ve (Ω, U, P) de bir olasılık uzayı olmak üzere bir zaman serisi $\Omega \times T$ çarpım uzayından reel sayılara giden bir fonksiyondur. Yani, bir zaman serisi

$$\begin{aligned} X(w, t): \Omega \times T &\rightarrow \mathbb{R} \\ (w, t) &\rightarrow X(w, t) \end{aligned}$$

olarak ifade edilebilir. Zaman serisi $X_t(w)$ veya kısaca X_t şeklinde gösterilir (Akdi 2012). Buna göre bir zaman serisi $\{X_t: t \in T\}$ şeklinde ifade edilebilir. $\{X_t: t \in T\}$ zaman serisi için $w = w_0$ sabit tutulduğunda zaman serisi reel değerli bir fonksiyon olarak ele alınabilir. Bu fonksiyon zaman serisinin bir yörüngesini oluşturur. Aslında internette, televizyonlarda gördüğümüz zaman serisi grafikleri birer yörüngedir. $t=t_0$ sabit tutulduğunda ise $X_t(w)$ bir rasgele değişkendir.

2.1.2 Durağanlık

Zaman serilerinde durağanlık istatistiki sonuç çıkarımı için çok önemlidir. Zaman serileri ortalamadan sapma durumuna göre durağan ve durağan olmayan zaman serileri olarak iki grupta toplanabilir. Durağan olmayan bir seri ile öngörü yapılamaz. Bu tür serilerde öngörü yapabilmek için serinin durağanlaştırılması gerekir. Herhangi bir zaman serisi verildiğinde, öngörü yapabilmek için önce durağanlığın sınanması gerekir. Eğer seri durağan değilse önce durağanlaştırılmalı daha sonra öngörü yapılmalıdır. Ayrıca, öngörü yapabilmek için önce öngörülere uygun bir modelin belirlenmesi gerekir.

$\{X_t: t \in T\}$ gibi herhangi bir zaman serisi verildiğinde, eğer

- i) $E(X_t) = \mu, \forall t$
- ii) $Var(X_t) = \sigma^2, \forall t$
- iii) $Cov(X_t, X_{t+h}) = \gamma(h), \forall t, h$

özellikleri sağlamıyorsa yani ortalaması ve kovaryansı zaman göre değişmiyorsa bu zaman serisi durağandır. Daha genel olarak $t + h = s$ dersek $Cov(X_t, X_s)$ kovaryansı sadece $|t-s|$ nin bir fonksiyonudur. Buradaki $Cov(X_t, X_{t+h})$ kovaryansı zaman serisinin otokovaryans fonksiyonudur. Bu fonksiyon zaman serilerinin modellenmesinde ve sezgisel de olsa durağanlığının tespitinde önemli bir araçtır. Aşağıda otokovaryans fonksiyonu herhangi bir ayrıntıya girmeden kısaca özetlenmiştir.

2.1.3 Otokovaryans ve otokorelasyon

$\{X_t: t \in T\}$ gibi verilen herhangi bir zaman serisinin otokovaryans fonksiyonu

$$\gamma(h) = Cov(X_t, X_{t+h})$$

şeklinde verilir. Kolayca görüleceği gibi serinin varyansı $\gamma(0) = Cov(X_t, X_t) = Var(X_t)$ dir. $\gamma(h)$ fonksiyonu simetrik ($\gamma(-h) = \gamma(h)$) olup Cauchy – Schwartz eşitsizliğinden bütün h 'ler için $|\gamma(h)| \leq \gamma(0)$ olduğu kolayca görülür. Diğer taraftan $\gamma(h)$ fonksiyonu negatif olmayan tanımlıdır. Ayrıca serinin otokorelasyon fonksiyonu da $\rho(h) = \gamma(h)/\gamma(0)$ şeklinde tanımlanır ve bu fonksiyon da benzer özelliklere sahiptir.

Özellikle otoregresif zaman serilerinin modellenmesinde kullanılan kısmi otokorelasyon fonksiyonu da otokorelasyon fonksiyonu yardımı ile hesaplanabilir. Aslında serinin kısmi otokorelasyon fonksiyonu $\phi(h)$, X_t nin $X_{t-1}, X_{t-2}, \dots, X_{t-h}$ üzerine regresyonu yapıldığında X_{t-h} nin katsayısıdır. Fonksiyonun bu şekilde hesaplanması biraz zahmetli olabilir. Bunun yerine,

$$P_h = \begin{pmatrix} 1 & \rho(1) & \rho(2) & \dots & \rho(h-2) & \rho(h-1) \\ \rho(1) & 1 & \rho(1) & \dots & \rho(h-3) & \rho(h-2) \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ \rho(h-1) & \rho(h-2) & \rho(h-3) & \dots & \dots & 1 \end{pmatrix}$$

olmak üzere bu matrisin son kolonunun $[\rho(1), \rho(2), \dots, \rho(h)]'$ ile değiştirilmesi ile elde edilen

$$P_h' = \begin{pmatrix} 1 & \rho(1) & \rho(2) & \dots & \rho(h-2) & \rho(1) \\ \rho(1) & 1 & \rho(1) & \dots & \rho(h-3) & \rho(2) \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ \rho(h-1) & \rho(h-2) & \rho(h-3) & \dots & \dots & \rho(h) \end{pmatrix}$$

matrisleri yazıldığında serinin otokorelasyon fonksiyonu

$$\phi(h) = \frac{\det(P_h')}{\det(P_h)}$$

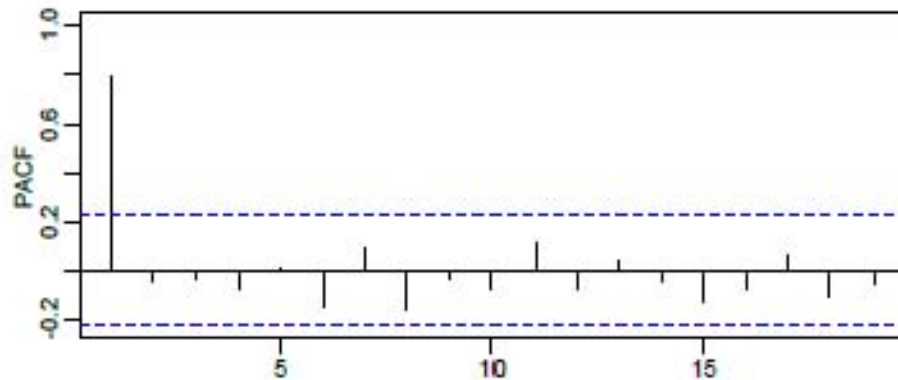
şeklinde bulunur.

Ayrıca h inci kısmi otokorelasyon değeri için

$$\begin{aligned} \rho(h) &= \frac{y(h) - \alpha_1 y(h-1) - \alpha_2 y(h-2) - \dots - \alpha_{h-1} y(1)}{y(h) - \alpha_1 y(1) - \alpha_2 y(2) - \dots - \alpha_{h-1} y(h-1)} \\ &= \frac{\rho(h) - \alpha_1 \rho(h-1) - \alpha_2 \rho(h-2) - \dots - \alpha_{h-1} \rho(1)}{1 - \alpha_1 \rho(1) - \alpha_2 \rho(2) - \dots - \alpha_{h-1} \rho(h-1)} \end{aligned}$$

formülü de yazılabilir. Buradan Cramer kuralına göre α_i katsayıları $\alpha_i = \det(P_i') / \det(P_i)$ formülü ile hesaplanır (Wei 2006).

Birinci dereceden otoregresif ($AR(1)$) zaman serisi modeline uygun olarak rasgele üretilen 100 birimlik bir zaman serisinin kısmi otokorelasyon fonksiyonunun grafiği aşağıda verilmiştir.



Şekil 2.1 AR(1) zaman serisi modeline uygun kısmi otokorelasyon grafiği

Bu grafik incelendiğinde kısmi otokorelasyonlar birinci gecikmeden sonra sifıra yakın değerler almaktadır. Bu da verilerin $AR(1)$ olarak modellenmesine işaret etmektedir.

2.1.4 Regresyon

Zaman serisi modelleri genelde regresyon modellerine benzer. (X,Y) ortak olasılık (veya olasılık yoğunluk) fonksiyonu $f(x,y)$ olan iki boyutlu bir rasgele vektör olmak üzere Y nin X üzerine regresyonu $E(Y|X=x)$ koşullu beklenen değeri Y nin X üzerine regresyonudur. $(Y, X_1, X_2, \dots, X_p)$ ortak olasılık (veya olasılık yoğunluk) fonksiyonunu $f_{(Y, X_1, X_2, \dots, X_p)}(y, x_1, \dots, x_p)$ olan $(p \times 1)$ boyutlu bir rasgele vektör olmak üzere $E(Y|X_1=x_1, X_2=x_2, \dots, X_p=x_p)$ koşullu beklenen değeri de Y nin $X_1, X_2, \dots, X_p$ üzerine çoklu regresyonudur. Eğer bu koşullu beklenen değer x_i 'lerin lineer bir fonksiyonu ise lineer regresyon aksı halde lineer olmayan regresyon adını alır. Basit doğrusal regresyon modeli

$$Y_t = \beta_0 + \beta_1 x_t + e_t, t = 1, 2, \dots, n$$

şeklinde verildiğinde e_t ' ler birbirinden bağımsız ($e_t \sim iid(0, \sigma^2)$) beklenen değeri 0 varyansı σ^2 olan bağımsız rasgele değişkenler, β_0 ve β_1 model parametreleri (β_0 kesim noktası, β_1 eğim), x_i ' ler açıklayıcı değişkenler (rasgele olmayan) olup Y_i ' ler bağımlı değişkenleri ifade etmektedir. Buradaki bağımlılık Y_i lerin x_i lere bağımlılığı anlamında olup aslında beklenen değeri $E(Y_i) = \beta_0 + \beta_1 x_i$, varyansı σ^2 olan bağımsız rasgele değişkenlerdir.

Kurulan regresyon modelinde bağımsız değişkenlerin (x_i) bilinen yani rasgele olmayan değişkenler olması, e_t hata terimlerinin birbirinden bağımsız ve aynı dağılıma sahip olmaları gerekir. İstatistiki sonuç çıkarımı için hataların normal dağılması gerekir. Bu varsayımlar sağlanmadan kurulacak modeller ve doğal olarak yapılacak analizler ve çıkarımlar hatalı olacaktır. Regresyonda amaç bilinmeyen model parametrelerinin ($\beta_i, i = 0, 1, 2, \dots, p$ ve σ^2) tahmin edilmesi ve uygun modelin tespit edilmesidir.

Bahsedilen bu varsayımların sağlanıp sağlanmadığını kontrol etmek için birçok yöntem bulunmaktadır. Örneğin hata terimlerinin ilişkisiz olup olmadığını tespit etmek için

Durbin-Watson istatistiği kullanılabilir. $\hat{e}_t = Y - \hat{Y}$ şeklinde hesaplandıktan sonra Durbin-Watson test istatistiğinin değeri ,

$$d = \frac{\sum_{t=2}^n (e_t - e_{t-1})^2}{\sum_{t=1}^n e_t^2}$$

formülü ile hesaplanır. Hata terimleri normal dağılımlı ve ilişkisiz ise paydaki toplam yaklaşık olarak $2(n-1)\sigma^2$ dir. Paydadaki toplamın beklenen değeri ise $n\sigma^2$ dir. Bu iki değeri oranlarsak yaklaşık 2 değerini buluruz. Yani hata terimleri otokorelasyonlu değil ise d' nin alacağı değer yaklaşık 2 olacaktır. d değeri 0 ile 4 arasında değer alır. $d > 2$ ise negatif korelasyon tersi durumda ise pozitif korelasyon vardır (Brockle ve Dickey 1986).

Hataların normalliği Kolmogorov – Simirnov testi, Shapiro – Wilk testi gibi testlerle sınanabilir. Hata terimlerinin normal dağılıma sahip olup olmadığını belirlemek için çeşitli görsel teknikler de vardır. Artıkların histogramı, kutu çizgi grafikleri ve normal olasılık grafikleri bunlardan bazılarıdır.

Regresyon analizinde hataların bağımsız olması ve normal dağılması kadar bunların sabit varyanslı olması yani varyansların zaman içinde değişmemesi da önemli varsayımlardan birisidir. Bu durumun tespiti için de değişik testler (White testi gibi) vardır. $Y = X\beta + e$ şeklindeki bir regresyon modeli için $Var(e) = \sigma^2 I$ matrisi olması gerekirken $Var(e) = \sigma^2 V$ olursa değişen varyans sorunu var demektir. Dönüşüm yapıldıktan sonra analiz ve sonuç çıkarım aşamasına geçilmelidir. V bir singüler matris olmak üzere;

$$Y^* = V^{-1/2} Y, \quad X^* = V^{-1/2} X, \quad e^* = V^{-1/2} e$$

dönüşümleri yapılır. Yeni modelimiz $Y^* = X^*\beta + e^*$ halini alır. Buradan,

$Var(e^*) = Var(V^{-1/2} e) = V^{-1/2} Var(e) V^{-1/2} = \sigma^2 I$ olur ve varsayım sağlanmış olur. Bu modele göre,

$$\hat{\beta}^* = (X^{*'} X^*)^{-1} X^{*'} Y^* = (X' V^{-1} X)^{-1} X' V^{-1} Y$$

olup, β^* nin genelleştirilmiş en küçük kareler tahmin edicisidir.

Regresyon modelimiz için diğ er bir varsayımımız ise oluřturulan modelin ve modelde var olan katsayıların istatistiksel a ıdan anlamlı olmasıdır. Eđer modelimiz anlamsız  ıkarsa modelimizi deđiřtirmeli veya bařka metotları kullanmalıyız. Katsayılarımızdan anlamsız olanları mutlaka modelimizden  ıkarmalı ve yeniden modelimizi kurmalıyız. Aksi halde elde edeceđimiz sonu lar ve  ıkarımlar hatalı olacaktır.

2.1.5  ng r 

Zaman serilerinde ilgili analizler yapılp model kurulduktan sonra analizini yaptığımız bu rasgele deđiřkenin gelecek d nemlerde alacađı deđerleri  ng rmek ř phesiz  nemli bir yer tutmaktadır. İstatistikte “ tahmin (estimation)”, “ kestirim (prediction)” ve “  ng r  (forecasting)” kavramları birbirine yakın anlamlar tařımalarına rađmen farklı řeyleri ifade ederler.

Tahmin  zerinde  alıřılan  rneklemin geldiđi kitleye ait bir parametre i in  nerilen tahmin edicinin deđeridir. Yani kitleye ait ve bizim ger ekte hi bir zaman bilemeyeceğimiz parametreler i in  eřitli y ntemleri kullanarak tahminde bulunmaktadır.  rneđin, $X_1, X_2, \dots, X_n$ beklenen deđer μ varyansı σ^2 olan bir kitleden alınmış  rnekleme olsun. μ parametresi i in $\bar{X}_n$ tahmin edicisi  nerilebilir. Bu tahmin edicinin alacađı deđer μ i in bir tahmindir. Rasgele deđiřkenin deđerleri

5,8	8,7	9,6	3,4	11,4	6,6	3,1	8,3	7,5	4,9
-----	-----	-----	-----	------	-----	-----	-----	-----	-----

olarak g zlenmiş olsun. μ i in tahminimiz $\hat{X}_n = 6.93$ olur.

Kestirim bir rasgele deđiřken i in elde edilen bir modelin parametrelerinin tahminleri modelde yerine konulduđunda elde edilen deđerdir.  rneđin, $Y_t = \beta_0 + \beta_1 x_t + e_t$, $t = 1, 2, \dots, 10$ řeklinde bir modeli elimizdeki rasgele deđiřken i in kurduđumuzu d ř nelim. X ve Y deđerlerimiz ř  řekilde olsun:

X	1	2	3	4	5	6	7	8	9	10
Y	7.5	6.4	9.8	8.1	8.9	10.8	9.6	6.6	11.4	7.7

Modelimizin parametreleri olan β_0 ve β_1 için en küçük kareler tahmin edicileri $\beta_0=7.7733$ ve $\beta_1=0.1648$ olarak bulunmuştur. Kestirim denklemi $\hat{Y} = 7.7733 + 0.1648 x_t$ olur. $Y_1=7.5$ iken modelimize göre $\hat{Y}_1=7.66$ olarak kestirilmiştir.

Öngörü ise bir modelde parametrelerin tahmin edilmesi ile rasgele değişkenin gelecek dönemlerde alabileceği değerleri kestirmektir. Örneğin yukarıdaki regresyon modelimizin 10 aylık buğday üretimi olduğunu ve modelimizin de bu veri seti için uygun olduğunu düşünelim. 11'inci ayda üretilebilecek buğday miktarını öngörmek istediğimizde;

$\hat{Y} = 7.7733 + 0.1648 (11) = 9.58$ buluruz. Yani önümüzdeki ay 9.58 ton buğday üreteceğimizi öngörebiliriz.

2.2 Bazı Durağan Zaman Serisi Modelleri

Bu kısımda, uygulamada çok karşılaşılan bazı durağan zaman serisi modelleri kısaca gözden geçirilecektir. Bu modellerden başlıcaları Hareketli ortalama (Moving Average, MA) modelleri, Aotoregresif (Autiregrasive, AR) zaman serisi modeller, Mevsimsel zaman serisi modelleri, Ootoregresive Hareketli Ortalama (Autoregressive Moving Average, ARMA) modelleridir. Bu modeller ilişkin temel özellikler kısaca özetlenecektir.

2.2.1 Hareketli ortalama (MA) serileri

$\{X_t : t \in T\}$ t dönemindeki değeri, aynı dönemdeki hata terimi (e_t) ve onun geçmiş dönemlerine ait gecikmeli değerlerinin bir fonksiyonu şeklinde ifade edilebiliyorsa $\{X_t : t \in T\}$ zaman serisi hareketli ortalama süreci (Moving Average, MA) olarak tanımlanır ve

$$X_t = \mu + e_t + \sum_{j=1}^q \beta_j e_{t-j}$$

şeklinde gösterilir. Buradaki q değeri MA modelinin derecesini μ de serinin varyansını göstermektedir. Ayrıca, e_t ' ler beklenen değeri 0 varyansı σ^2 olan ilişkisiz rasgele değişkenlerdir ($E(e_t) = 0, Var(e_t) = \sigma^2, Cov(e_t, e_s) = 0, s \neq t$). Hareketli ortalama serilerine

en belirgin örnek beyaz gürültü serisi verilebilir. $\gamma(h)$ serinin otokovaryans fonksiyonunu göstermek üzere, $\{e_t = t \in T\}$ zaman serisi

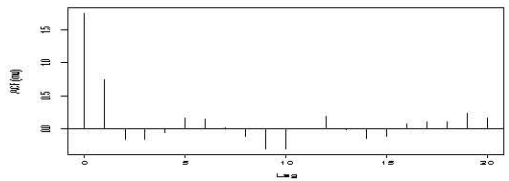
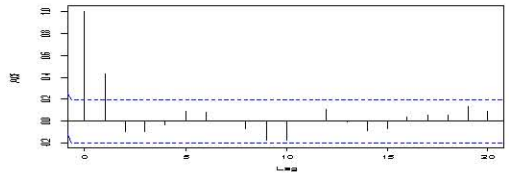
$$\gamma(h) = \begin{cases} \sigma^2, & h = 0 \\ 0, & d.d. \end{cases}$$

şeklinde bir otokovaryans fonksiyonuna sahipse $\{e_t = t \in T\}$ zaman serisine bir beyaz gürültü serisi denir ve $e_t \sim WN(0, \sigma^2)$ ile ifade edilir.

$X_t \sim MA(q)$ olmak üzere, X_t nin beklenen değeri μ olduğu ($E(X_t) = \mu$) açık olup otokovaryans fonksiyonu $h > 0$ için $\gamma(-h) = \gamma(h)$

$$\gamma(h) = Cov(X_t, X_{t+h}) = \sigma^2 \sum_{j=0}^{q-h} \beta_j \beta_{j+h}$$

dir. Buna göre serinin beklenen değeri ve otokovaryansları zamana bağlı değildir. Yani, $MA(q)$ modeli her zaman durağandır. Özel olarak $q = 1$ ve $q = 2$ için $MA(1)$ ve $MA(2)$ modellerine ilişkin otokovaryans ve otokorelasyon fonksiyonları sırası ile şekil 2.2'de gösterilmiştir.

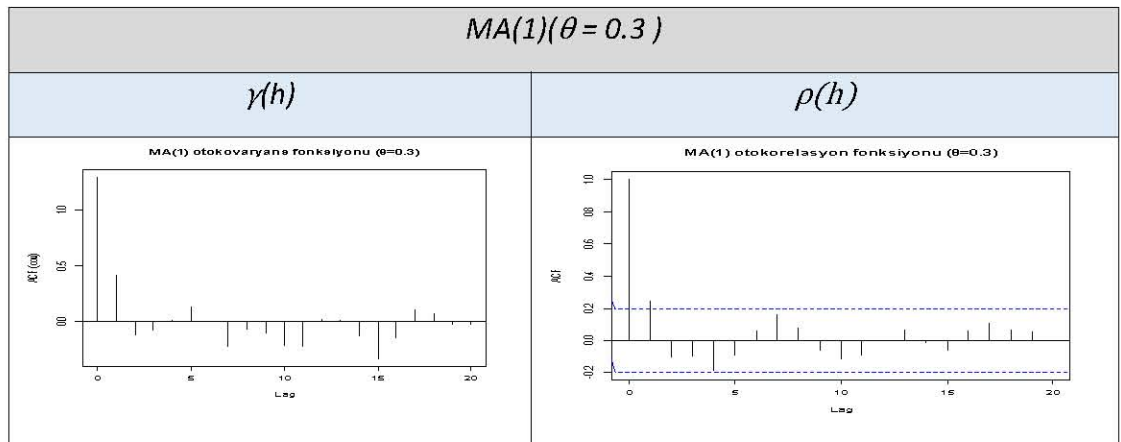
MA(1)	
$\gamma(h)$	$\gamma(h)$ fonksiyon grafiği
$\gamma(h) = \begin{cases} \sigma^2(1 + \theta^2), & h = 0 \\ \sigma^2\theta, & h = \pm 1 \\ 0, & h \geq 2 \end{cases}$	
$\rho(h)$	$\rho(h)$ fonksiyon grafiği
$\rho(h) = \begin{cases} 1, & h = 0 \\ \frac{\theta}{1 + \theta^2}, & h = \pm 1 \\ 0, & h \geq 2 \end{cases}$	

Şekil 2.2 $MA(1)$ ve $MA(2)$ modellerine ilişkin otokovaryans ve otokorelasyon fonksiyonları

MA(2)	
$\gamma(h)$	$\gamma(h)$ fonksiyon grafiği
$\gamma(h) = \begin{cases} \sigma^2(1 + \theta_1^2 + \theta_2^2), h = 0 \\ \sigma^2(\theta_1 + \theta_1\theta_2), h = \pm 1 \\ \sigma^2\theta_2, h = \pm 2 \\ 0, h \geq 3 \end{cases}$	
$\rho(h)$	$\rho(h)$ fonksiyon grafiği
$\rho(h) = \begin{cases} 1, h = 0 \\ \frac{\theta_1(1 + \theta_2)}{1 + \theta_1^2 + \theta_2^2}, h = \pm 1 \\ \frac{\theta_2}{1 + \theta_1^2 + \theta_2^2}, h = \pm 2 \\ 0, h \geq 3 \end{cases}$	

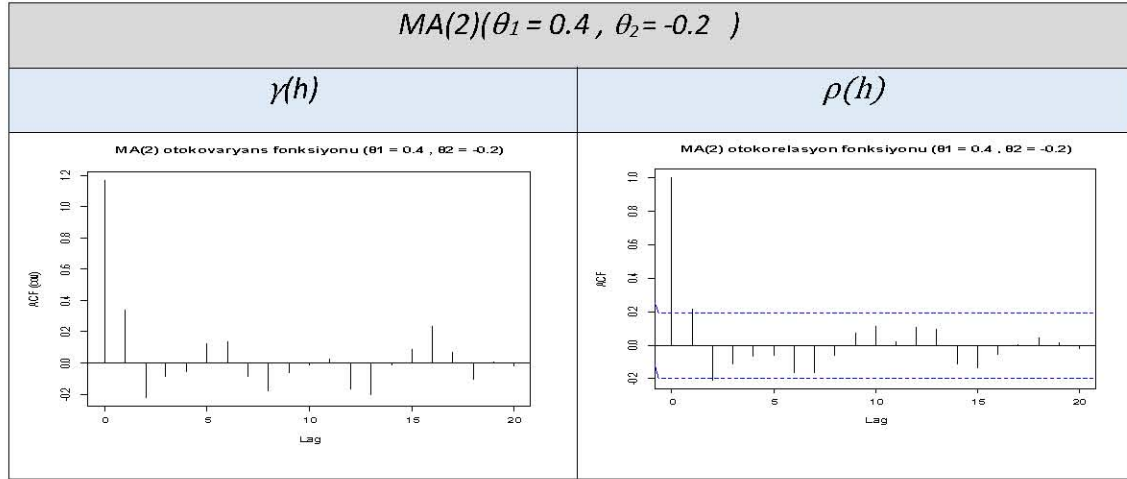
Şekil 2.2 MA(1) ve MA(2) modellerine ilişkin otokovaryans ve otokorelasyon fonksiyonları (devam)

Bu modellere uygun olarak rasgele üretilen 100 birimlik verilere ilişkin hesaplanan otokovaryansların grafikleri aşağıda verilmiştir. Bu fonksiyonların grafikleri özel olarak seçilen θ katsayılarına göre otokovaryans ve otokorelasyon fonksiyonları değişim göstermektedir. Buna göre $\theta = 0.3$ olan bir MA(1) modelinin otokovaryans ve otokorelasyon fonksiyonları şöyledir:



Şekil 2.3 $\theta = 0.3$ olan bir MA(1) modelinin otokovaryans ve otokorelasyon fonksiyonları

Ayrıca $\theta_1 = 0.4$ ve $\theta_2 = -0.2$ olacak şekilde türetilen bir $MA(2)$ modelinin otokovaryans ve otokorelasyon fonksiyonları şekil 2.4'te gösterilmiştir.



Şekil 2.4 $\theta_1 = 0.4$ ve $\theta_2 = -0.2$ olacak şekilde türetilen bir $MA(2)$ modelinin otokovaryans ve otokorelasyon fonksiyonları

Şekil 2.3-2.4'ten de görüldüğü gibi MA modellerinde otokorelasyonlar belli bir yerden sonra sıfırdır ($MA(1)$ için birinci, $MA(2)$ için ikinci gecikmeden sonra sıfırdır). Buna göre, serinin otokovaryansları belli bir yerden sonra sıfır oluyorsa (örneğin ikinci gecikmeden sonra), bu serinin $MA(2)$ olarak modellenmesi uygundur.

Verilen herhangi bir zaman serisinin $\{x_1, x_2, \dots, x_n\}$ gözlem değerlerine karşılık otokovaryansların tahmin değerlerinin

$$\hat{\gamma}_n(h) = \frac{1}{n-h} \sum_{t=1}^{n-h} (x_t - \bar{x}_n)(x_{t+h} - \bar{x}_n)$$

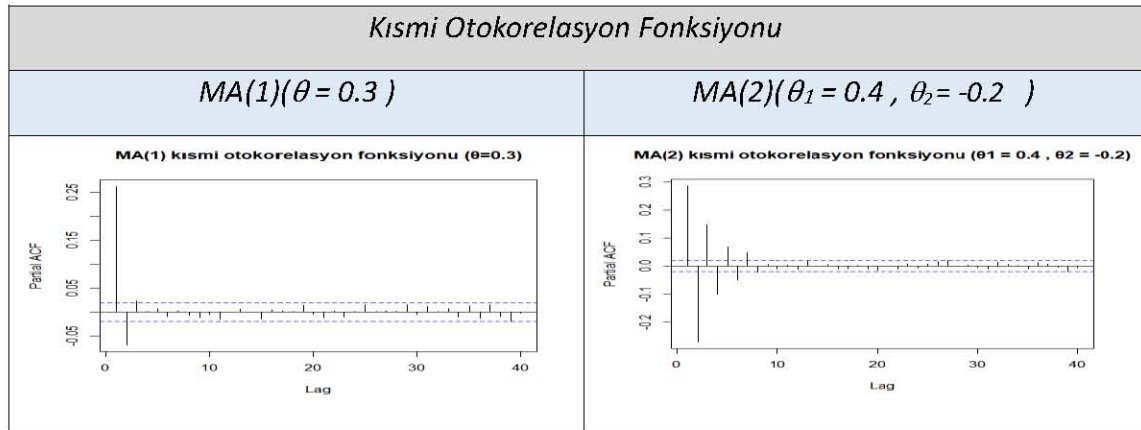
şeklinde hesaplandığını biliyoruz. Ayrıca $\gamma(0) = \text{Var}(X_t)$ olup $\{X_t : t \in T\}$ otokorelasyon fonksiyonu da

$$\rho_n(h) = \frac{\gamma_n(h)}{\gamma_n(0)}$$

olarak hesaplanır. Verilen herhangi bir $\{X_t : t \in T\}$ zaman serisinin $\{x_1, x_2, \dots, x_n\}$ gözlem değerlerine dayalı hesaplanan $\rho_n(h)$ otokorelasyonlarının h gecikmelerine ait grafikleri çizildiğinde bu grafiklerde belli bir gecikmeden sonra sıfır değerleri gözleniyorsa bu tür serilerin MA olarak modellenmesi uygundur. Bu “belli bir yer” ise MA modelinin

derecesidir. Yukarıda $MA(1)$ ve $MA(2)$ modellerine uygun rasgele üretilmiş $n = 100$ birimlik serilere ilişkin otokorelasyon fonksiyonlarının grafiklerine baktığımızda bu durum görülebilmektedir.

Diğer taraftan zaman serilerinin sezgisel de olsa modellenmesinde kullanılan diğer bir fonksiyon ise kısmi otokorelasyon fonksiyonudur. Yukarıda rasgele üretilen verilere ilişkin kısmi otokorelasyon fonksiyonlarının grafikleri aşağıdadır.



Şekil 2.5 Rasgele üretilen $MA(1)$ ve $MA(2)$ modellerine ait kısmi otokorelasyon fonksiyonları

Bu grafiklerden, $\rho_n(h)$ lar belli bir yerden sonra sıfır olmasına rağmen $\phi_n(h)$ kısmi otokorelasyonlarında benzer bir davranış göze çarpmamaktadır. Kısaca, eğer serinin $\rho_n(h)$ otokorelasyonları belli bir yerden sonra sıfır ve $\phi_n(h)$ kısmi otokorelasyonları bir azalış eğimi göstermesine rağmen sıfır olmuyorsa bu tür seriler MA olarak modellenmelidir.

$e_t = WN(0, \sigma^2)$ olmak üzere $MA(1)$ zaman serisi modeli

$$X_t = 0.8e_{t-1} + e_t$$

şeklinde verilmiş olsun. Bu modelin otokorelasyon ve kısmi otokorelasyon değerlerini hesaplayalım. Otokorelasyon değerleri,

$$\rho(0) = 1$$

$$\rho(1) = \theta/1+\theta^2 = 0.8/1+(0.8)^2 = 0.48$$

$$\rho(2) = \rho(3) = \dots = \rho(10) = 0$$

olup kısmi otokorelasyon değerlerinden ilk beş tanesi,

$$\phi(h) = \frac{(-1)^{h+1} \theta^h (1 - \theta^2)}{1 - \theta^{2(h+1)}}$$

formülünden

$$\Phi(1) = 0.48$$

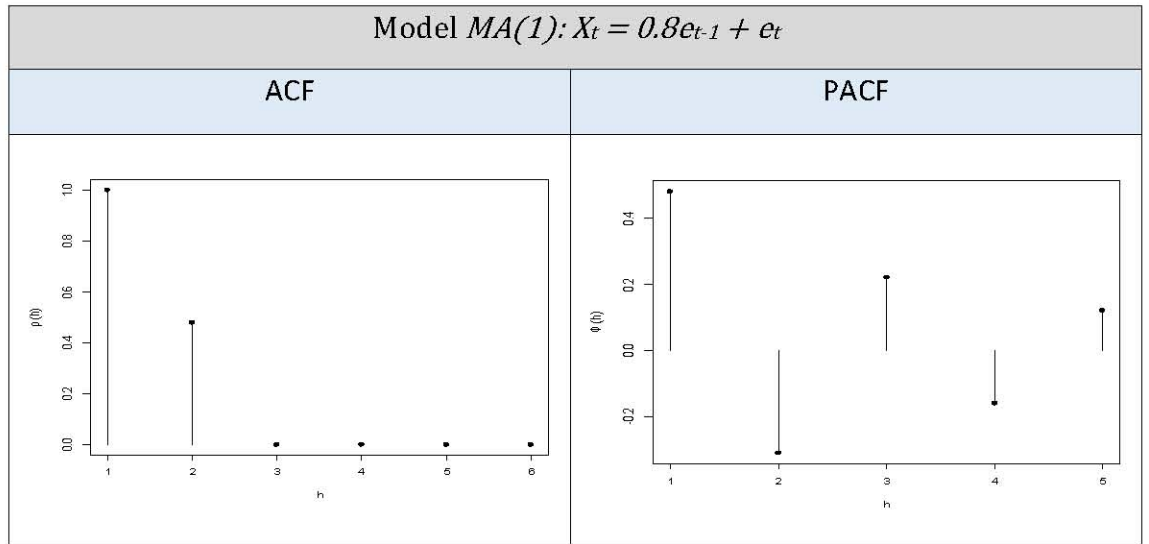
$$\Phi(2) = -0.31$$

$$\Phi(3) = 0.22$$

$$\Phi(4) = -0.16$$

$$\Phi(5) = 0.12$$

olarak bulunur. Bu değerler grafik olarak şekil 2.6'da verilmiştir.

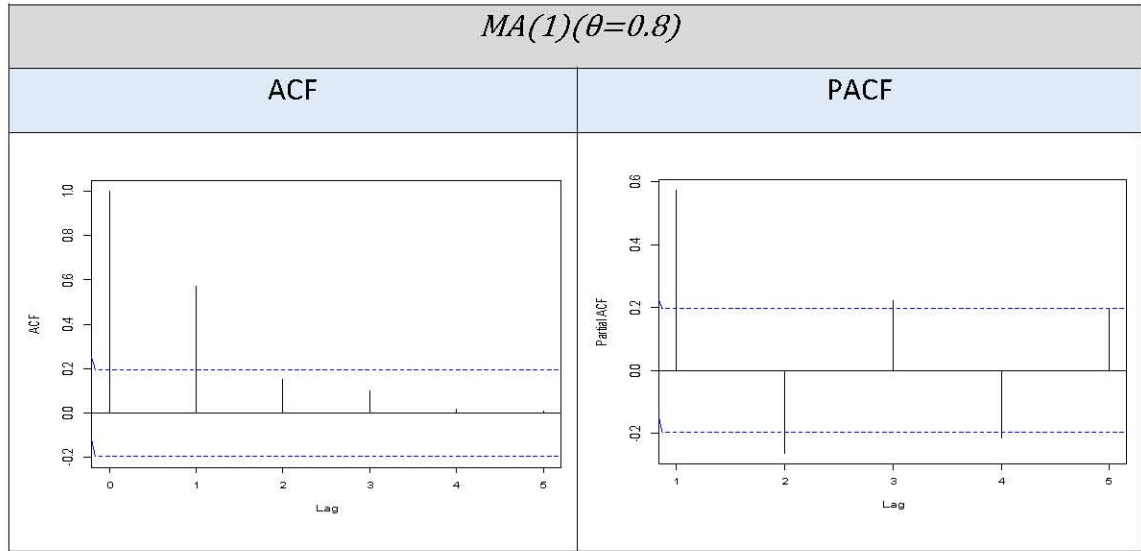


Şekil 2.6 $X_t = 0.8e_{t-1} + e_t$ şeklinde verilen MA(1) modeline ait otokorelasyon ve kısmi otokorelasyon fonksiyonları

Şimdi de bu modelimize uygun 100 birimlik veri üretelim ve bu ürettiğimiz veri setinin aynı şekilde otokorelasyon ve kısmi otokorelasyon değerlerini bulalım ve grafik olarak gösterelim.

Çizelge 2.1 $X_t = 0.8e_{t-1} + e_t$ modeline uygun üretilen veri setinin otokorelasyon ve kısmi otokorelasyon değerleri

h	0	1	2	3	4	5
$\rho(h)$	1.00	0.573	0.152	0.099	0.014	0.008
$\Phi(h)$		0.573	- 0.262	0.223	- 0.213	0.195



Şekil 2.7 $X_t = 0.8e_{t-1} + e_t$ şeklinde verilen $MA(1)$ modeline uygun üretilen 100 birimlik veri setine ait otokorelasyon ve kısmi otokorelasyon fonksiyonları

Her iki örnekte otokorelasyon grafiği 1'inci dereceden sonra sıfır olmakta kısmi otokorelasyon grafikleri ise mutlak değerce üstel olarak azalmakta ancak sıfır olmamaktadır.

Bölüm 2.2.1'de de belirtildiği gibi $MA(q)$ modelleri q gecikme sayısı sonlu olduğu sürece durağandır. Eğer $MA(q)$ serisi $e_t \sim WN(0, \sigma^2)$ olmak üzere $q = +\infty$ için

$$X_t = \mu + \sum_{j=0}^{\infty} \beta_j e_{t-j}$$

verilmiş olursa X_t serisinin beklenen değeri $\sum_{j=0}^{\infty} \beta_j$ serisinin yakınsaklığına bağlıdır. Eğer

bu seri yakınsak değilse serinin beklenen değeri tanımsızdır.

2.2.2 Otoregresif (AR) seriler

Bir önceki kısımda $MA(q)$ modellerinin q sonlu olduğu sürece her zaman durağan olduğu görüldü. Eğer $q = \infty$ ise $MA(q)$ zaman serisi modeli $e_t = WN(0, \sigma^2)$ olmak üzere

$$X_t = \mu + \sum_{j=0}^{\infty} \beta_j e_{t-j} \quad (2.1)$$

şeklinde yazılabilir. Burada $E(X_t)$ nin ve kovaryansların hesaplanabilmesi β_j

katsayılarının durumuna bağlıdır. Eğer $\sum_{j=0}^{\infty} \beta_j^2 < \infty$ veya $\sum_{j=0}^{\infty} |\beta_j| < \infty$ ise beklenen değer

ile sonsuz toplam yer değiştirebilir ve $E(e_t) = 0$ olduğundan

$$E(X_t) = \mu, \gamma(0) = Var(X_t) = \sigma^2 \sum_{j=0}^{\infty} \beta_j^2$$

ve

$$h > 0 \text{ için } \gamma(h) = \sigma^2 \sum_{j=0}^{\infty} \beta_j \beta_{j+h}$$

olarak yazılabilir. Eğer $\sum_{j=0}^{\infty} |\beta_j| < \infty$ (veya $\sum_{j=0}^{\infty} \beta_j^2 < \infty$) varsayımı sağlanmıyor ise modelin momentleri hesaplanamaz. Dolayısı ile modelin durağan olduğu söylenemez.

Yukarıda eşitlik (2.1) de verilen $MA(q)$ zaman serisi modeli için $|\rho| < \infty$ olmak üzere, $\beta_j = \rho^j$ alalım. Bu durumda

$$\sum_{j=0}^{\infty} \beta_j^2 = \frac{1}{1-\rho^2} < \infty$$

olduğundan $E(X_t) = \mu$, $Var(X_t) = \frac{\sigma^2}{1-\rho^2}$ ve $h > 0$ için $\gamma(h) = \rho^h (\frac{\sigma^2}{1-\rho^2})$ dir. Yani model durağandır. Yine $|\rho| < 1$ ve $\beta_j = \rho^j$ olmak üzere $MA(\infty)$ modelini ($\mu = 0$ alınsın),

$$X_t = \mu + \sum_{j=0}^{\infty} \rho^j e_{t-j}$$

şeklinde yazalım. Buradan serinin terimleri açık olarak

$$X_t = e_t + \rho e_{t-1} + \rho^2 e_{t-2} + \rho^3 e_{t-3} + \dots$$

şeklinde yazılabilir. Burada ρX_{t-1} serisi ise

$$\rho X_{t-1} = \rho e_{t-1} + \rho^2 e_{t-2} + \rho^3 e_{t-3} + \dots$$

olacaktır. Bu iki seri taraf tarafa çıkartıldığında $X_t - \rho X_{t-1} = e_t$ serisi elde edilir. Yani

$$\begin{array}{rcl} X_t & = & e_t + \rho e_{t-1} + \rho^2 e_{t-2} + \rho^3 e_{t-3} + \dots \\ - \rho X_{t-1} & = & \rho e_{t-1} + \rho^2 e_{t-2} + \rho^3 e_{t-3} + \dots \\ \hline X_t - \rho X_{t-1} & = & e_t \end{array}$$

dir. Buradan da

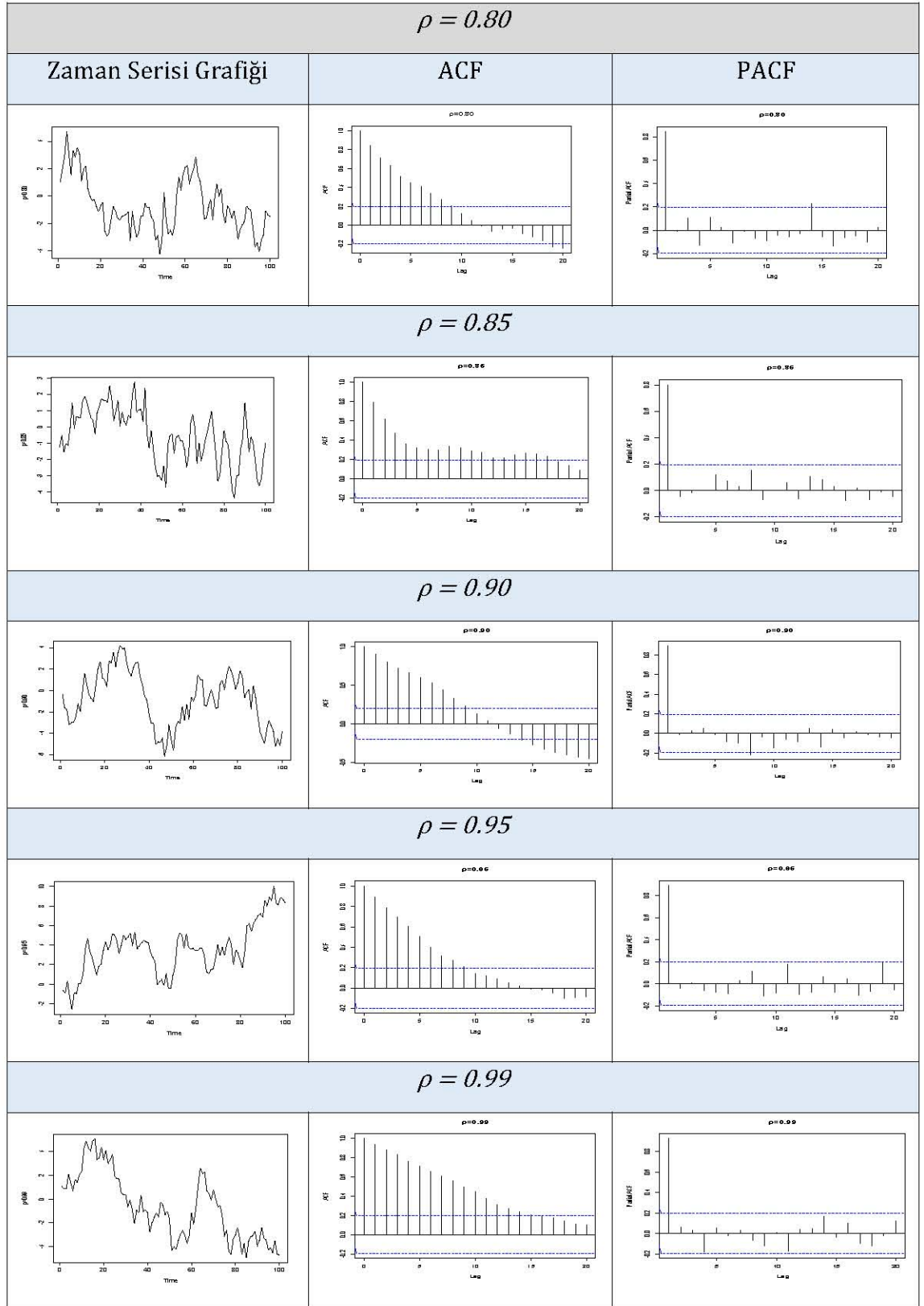
$$X_t = \rho X_{t-1} + e_t$$

modeli elde edilir. Bu da birinci dereceden otoregresif zaman serisi olup $AR(1)$ şeklinde gösterilir. Yani $AR(1)$ zaman serisi modeli $e_t = WN(0, \sigma^2)$ olmak üzere $X_t = \rho X_{t-1} + e_t$ şeklinde gösterilir.

Yukarıda da sözü edildiği gibi $AR(1)$ modelinin durağanlığı ρ katsayısının durumuna

bağlıdır. Eğer $|\rho| < 1$ ise model durağandır ve $E(X_t) = 0$, $Var(X_t) = \frac{\sigma^2}{1 - \rho^2}$, $\gamma(h) = \rho^h \gamma(0)$

dir. Aşağıda ρ nun farklı değerleri için rasgele üretilen 100 birimlik zaman serisine ilişkin grafikler bulunmaktadır.



Şekil 2.8 Farklı ρ değerleri için zaman serisi, ACF ve PACF grafikleri

Şekil 2.8’de de görüldüğü gibi serinin otokorelasyonlarındaki azalma hızı ρ ’nun değerleri 1’e yaklaştıkça yavaşlamaktadır. Kısmi otokorelasyonlar ise birinci gecikmeden sonra sıfır (yada sıfıra çok yakın) değerlerini almaktadır.

$AR(1)$ zaman serisi modeli $e_t = WN(0, \sigma^2)$ olmak üzere

$$X_t = \rho X_{t-1} + e_t, t=1,2,\dots,n$$

şeklinde verilmiş olsun. Eğer model durağan ($|\rho| < 1$) ise serinin beklenen değeri sabittir ($E(X_t) = E(X_{t-1}) = E(X_{t-2}) = \dots = E(X_{t-500}) = \mu$) ve $E(X_t) = 0$ dır. [$E(X_t) = E(\sum \beta_j e_{t-j}) = \sum \beta_j E(e_{t-j}) = 0$]

Ayrıca,

$$X_t X_{t-h} = \rho X_{t-1} X_{t-h} + e_t X_{t-h}$$

olup

$$\gamma(h) = E(X_t X_{t-h}) = E(\rho X_{t-1} X_{t-h}) + E(e_t X_{t-h})$$

$$\Rightarrow \gamma(h) = \rho \gamma(h-1) + \sigma^2 I(h=0)$$

yazılır. Burada $I(h=0)$ değeri $h=0$ için 1, diğer durumlarda 0 değeri alan gösterge fonksiyonudur. Buna göre beklenen değeri sıfır olan ($\mu=0$) $AR(1)$ zaman serisi modeli için

$$\gamma(h) = \rho \gamma(h-1) + \sigma^2 I(h=0)$$

denklemleri yazılır. Yani,

$$\gamma(0) = \rho^2 \gamma(0) + \sigma^2$$

ve $h > 0$ için

$$\gamma(h) = \rho \gamma(h-1) \tag{2.2}$$

Yule-Walker denklemleri elde edilir.

$AR(1)$ durağan zaman serisi modeli

$$(X_t - \mu) = \rho(X_{t-1} - \mu) + e_t, t=1,2, \dots, n$$

şeklinde de verilebilir. Bu model

$$X_t = \mu(1-\rho) + \rho X_{t-1} + e_t, t=1,2, \dots, n$$

veya $\alpha_0 = \mu(1-\rho)$ ve $\alpha_1 = \rho$ olmak üzere

$$X_t = \alpha_0 + \alpha_1 X_{t-1} + e_t, t=1,2, \dots, n$$

şeklinde de yazılabilir.

Bu model görünüşte X_t 'nin X_{t-1} üzerine regresyonu gibi görünmektedir. Bir an için bu modelin regresyon modeli olduğunu varsayalım ve regresyon parametrelerinin (α_0 , α_1) en küçük kareler tahmin edicileri

$$\hat{\alpha}_1 = \frac{\left[\sum (X_t - \bar{X})(X_{t-1} - \bar{X}_{(-1)}) \right]}{\sum (X_{t-1} - \bar{X}_{(-1)})^2}$$

ve

$$\hat{\alpha}_0 = \bar{X}_t - \hat{\alpha}_1 \bar{X}_{(-1)}$$

dir.

Diğer taraftan, herhangi bir $\{x_1, x_2, \dots, x_n\}$ zaman serisi verileri için otokovaryanslar bir model varsayımına gerek duyulmadan

$$\hat{\gamma}_n(h) = \frac{1}{n-h} \sum_{t=1}^{n-h} (X_t - \bar{X}_n)(X_{t+h} - \bar{X}_n)$$

şeklinde hesaplanır. Buradan, eşitlik (2.2)'deki Yule-Walker denklemlerinde $\gamma(h) = \rho\gamma(h-1)$ olmak üzere ρ nun başka bir tahmin edicisi

$$\hat{\rho}_h = \frac{\hat{\gamma}(h)}{\hat{\gamma}_n(h-1)}$$

ile tahmin edilebilir. Buna ρ ' nun Yule-Walker tahmin edicisi denir.

Şimdi rasgele 50 veri üretelim. Bu ürettiğimiz veri seti ve bu veri setine ait otokovaryans, otokorelasyon ve kısmi otokorelasyon değerleri ile güven aralıkları çizelge 2.2'de verilmiştir.

Çizelge 2.2 Üretilen 50 veriye ait bilgiler

<i>Veri seti</i>				
-0.56	0.56	-1.59	0.28	-0.53
-1.61	0.69	-2.31	-0.77	-1.67
-1.89	0.18	-2.51	-0.97	-0.90
-1.89	0.36	-3.00	-1.61	-2.57
-0.47	0.72	-1.71	-2.31	-2.35
-0.64	0.53	-1.76	-2.10	-2.29
-0.21	-0.72	-1.27	-1.43	-3.46
-0.68	-1.43	-0.92	-3.32	-2.26
0.61	-1.59	-0.70	-2.72	-1.13
-0.63	-2.34	-1.32	-0.45	0.23

Çizelge 2.2 Üretilen 50 veriye ait bilgiler (devam)

Farklı α düzeylerinde güven aralıkları		
	Güven Aralığı	
%90	-1,46964	-0,94755
%95	-1,52149	-0,8957
%99	-1,62587	-0,79132

Otokovaryans, otokorelasyon ve kısmi otokorelasyon değerleri											
h	0	1	2	3	4	5	6	7	8	9	10
γ	1.187	0.762	0.552	0.392	0.245	0.039	0.002	-0.15	-0.12	-0.03	-0.06
ρ	1.00	0.642	0.465	0.330	0.207	0.033	0.002	-0.12	-0.10	-0.03	-0.05
Φ		0.642	0.09	0.001	-0.04	-0.17	0.061	-0.17	0.108	0.121	-0.11

Durağan $AR(1)$ modelinin $MA(\infty)$ şeklinde yazılabildiği bir önceki kısımda gösterildi. Benzer şekilde durağan $AR(2)$ modeli de $MA(\infty)$ şeklinde yazılabilir. Yani, $e_t = WN(0, \sigma^2)$ olmak üzere,

$$\hat{\rho}_h = \frac{\hat{\gamma}(h)}{\hat{\gamma}_n(h-1)} \quad (2.3)$$

şeklinde yazılabilir. w_j katsayıları $w_0 = 1$, $w_1 = \alpha_1$ ve $w_j = \alpha_1 w_{j-1} + \alpha_2 w_{j-2}$ olacak şekilde seçildiğinde Eşitlik (2.3)'de verilen $MA(\infty)$ modeli

$$X_t = \alpha_1 X_{t-1} + \alpha_2 X_{t-2} + e_t$$

şeklinde yazılabilir. Burada, $\alpha_0 = \mu(1-\alpha_1-\alpha_2)$ olup μ modelin beklenen değerini göstermektedir. Bu modelin durağanlığı $m^2 - \alpha_1 m - \alpha_2 = 0$ karakteristik denklemin köklerine bağlıdır. Eğer denklemin her iki kökü de mutlak değerce 1'den küçükse model durağandır. Eğer köklerden en az biri mutlak değerce 1 (ya da 1'den büyük) ise model durağan değildir. Ayrıca, $\alpha_1 + \alpha_2 = 1$ ise model durağan değildir. Bununla birlikte $\alpha_1 + \alpha_2 = 1$ ise modelin beklenen değeri μ modelden düşer. $AR(2)$ modelinin durağanlığı;

$$i) \alpha_1 + \alpha_2 < 1$$

$$ii) \alpha_2 - \alpha_1 < 1$$

$$iii) -1 < \alpha_2 < 1$$

koşullarının sağlanmasına da denktir (Wei, 2006). $AR(2)$ modeli $e_t = WN(0, \sigma^2)$ olmak üzere ,

$$X_t = \alpha_1 X_{t-1} + \alpha_2 X_{t-2} + e_t$$

şeklinde verildiğinde, basit hesaplamalardan sonra otokovaryans fonksiyonunun $h \geq 1$ için,

$$\gamma(h) = \alpha_1 \gamma(h-1) + \alpha_2 \gamma(h-2)$$

olduğu görülür. Özel olarak $h = 0, 1$ ve 2 için

$$\gamma(0) = \alpha_1 \gamma(1) + \alpha_2 \gamma(2) + \sigma^2$$

$$\gamma(1) = \alpha_1 \gamma(0) + \alpha_2 \gamma(1)$$

$$\gamma(2) = \alpha_1 \gamma(1) + \alpha_2 \gamma(0)$$

eşitlikleri yazılır. Buradan

$$\begin{bmatrix} \gamma(0) & \gamma(1) \\ \gamma(1) & \gamma(0) \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \alpha_2 \end{bmatrix} = \begin{bmatrix} \gamma(1) \\ \gamma(2) \end{bmatrix}$$

denklemleri elde edilir. Bu denklemlerinden α_1 ve α_2 model parametreleri elde edilebilir. Herhangi bir X_t zaman serisinin $\{x_1, x_2, \dots, x_n\}$ gözlem değerleri için otokovaryansların tahmin değerleri

$$\hat{\gamma}_n(h) = \frac{1}{n-h} \sum_{t=1}^{n-h} (X_t - \bar{X}_n)(X_{t+h} - \bar{X}_n)$$

ile hesaplanır. Buradan $\hat{\gamma}_n(0), \hat{\gamma}_n(1)$ ve $\hat{\gamma}_n(2)$ değerleri hesaplandığında α_1 ve α_2 parametrelerinin Yule-Walker tahmin edicileri bulunur. Ayrıca α_1 ve α_2 parametreleri bilindiğinde otokovaryanslar

$$\gamma(0) = \frac{(1-\alpha_2)\sigma^2}{(1+\alpha_2)[(1-\alpha_2)^2 - \alpha_1^2]}$$

$$\gamma(1) = \frac{\alpha_1 \sigma^2}{(1 + \alpha_2)[(1 - \alpha_2)^2 - \alpha_1^2]}$$

$$\gamma(2) = \frac{[\alpha_1^2 + \alpha_2(1 - \alpha_2)]\sigma^2}{(1 + \alpha_2)[(1 - \alpha_2)^2 - \alpha_1^2]}$$

şeklinde hesaplanır. Serinin bu üç otokovaryans değeri bilindiğinde diğer otokovaryanslar

$$\gamma(h) = \alpha_1 \gamma(h-1) + \alpha_2 \gamma(h-2)$$

Yule – Walker denklemlerinden ardışık olarak hesaplanır. Bununla birlikte serinin

otokorelasyon fonksiyonunun $\rho(h) = \frac{\gamma(h)}{\gamma(0)}$ şeklinde olduğunu biliyoruz. $\rho(0) = 1$ olduğu

hatırlandığında serinin ilk iki otokorelasyon değeri

$$\rho(1) = \frac{\alpha_1}{1 - \alpha_2}$$

$$\rho(2) = \frac{\alpha_1^2 + \alpha_2(1 - \alpha_2)}{(1 - \alpha_2)}$$

şeklinde elde edilir.

Örnek 1: $e_t = WN(0, \sigma^2)$ olmak üzere $AR(2)$ zaman serisi modeli

$$X_t = 1.2X_{t-1} - 0.3X_{t-2} + e_t$$

şeklinde verilmiş olsun. Burada, $\alpha_1 = 1.2$ ve $\alpha_2 = -0.3$ olup $\alpha_1 + \alpha_2 = 0.9 < 1$, $\alpha_2 - \alpha_1 = -1.5 < 1$

ve $-1 < \alpha_2 < 1$ olduğundan model durağandır. Bu modelin otokorelasyonlarından bazıları

$$\rho(0) = 1$$

$$\rho(1) = \alpha_1 / (1 - \alpha_2) = 1.2 / 1 - (-0.3) = 0.923$$

$$\rho(2) = \alpha_1^2 + \alpha_2(1 - \alpha_2) / (1 - \alpha_2) = (1.2)^2 - 0.3(1 - (-0.3)) / 1 - (-0.3) = 0.807$$

$$\rho(3) = 1.2\rho(2) - 0.3\rho(1) = 1.2(0.807) - 0.3(0.923) = 0.691$$

$$\rho(4) = 1.2\rho(3) - 0.3\rho(2) = 1.2(0.691) - 0.3(0.807) = 0.581$$

$$\rho(5) = 1.2\rho(4) - 0.3\rho(3) = 1.2(0.581) - 0.3(0.691) = 0.490$$

$$\rho(6) = 1.2\rho(5) - 0.3\rho(4) = 1.2(0.490) - 0.3(0.581) = 0.413$$

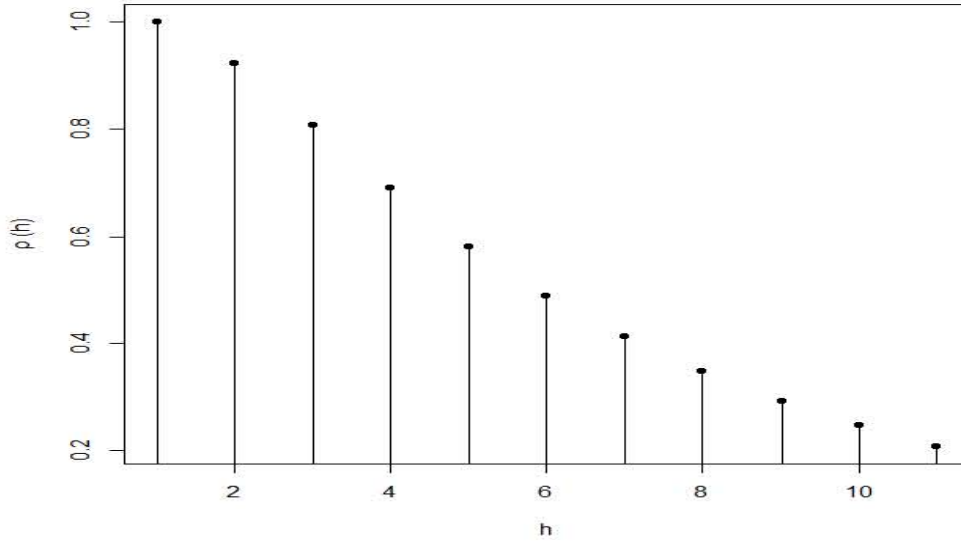
$$\rho(7) = 1.2\rho(6) - 0.3\rho(5) = 1.2(0.413) - 0.3(0.490) = 0.348$$

$$\rho(8) = 1.2\rho(7) - 0.3\rho(6) = 1.2(0.348) - 0.3(0.413) = 0.293$$

$$\rho(9) = 1.2\rho(8) - 0.3\rho(7) = 1.2(0.293) - 0.3(0.348) = 0.247$$

$$\rho(10) = 1.2\rho(9) - 0.3\rho(8) = 1.2(0.247) - 0.3(0.293) = 0.208$$

olarak bulunur. Otokorelasyonların grafiği şekil 2.9'da görüldüğü gibidir.



Şekil 2.9 $X_t = 1.2X_{t-1} - 0.3X_{t-2} + e_t$ modelinin otokorelasyon grafiği

Grafikten de görüldüğü gibi otokorelasyonlar mutlak değerce bir azalma eğilimi göstermesine rağmen hiçbir zaman sıfır olmamaktadır.

Modelin kısmi otokorelasyonlarından bazıları da aşağıda hesaplanmıştır.

$$\phi(1) = \rho(1) = 0.923$$

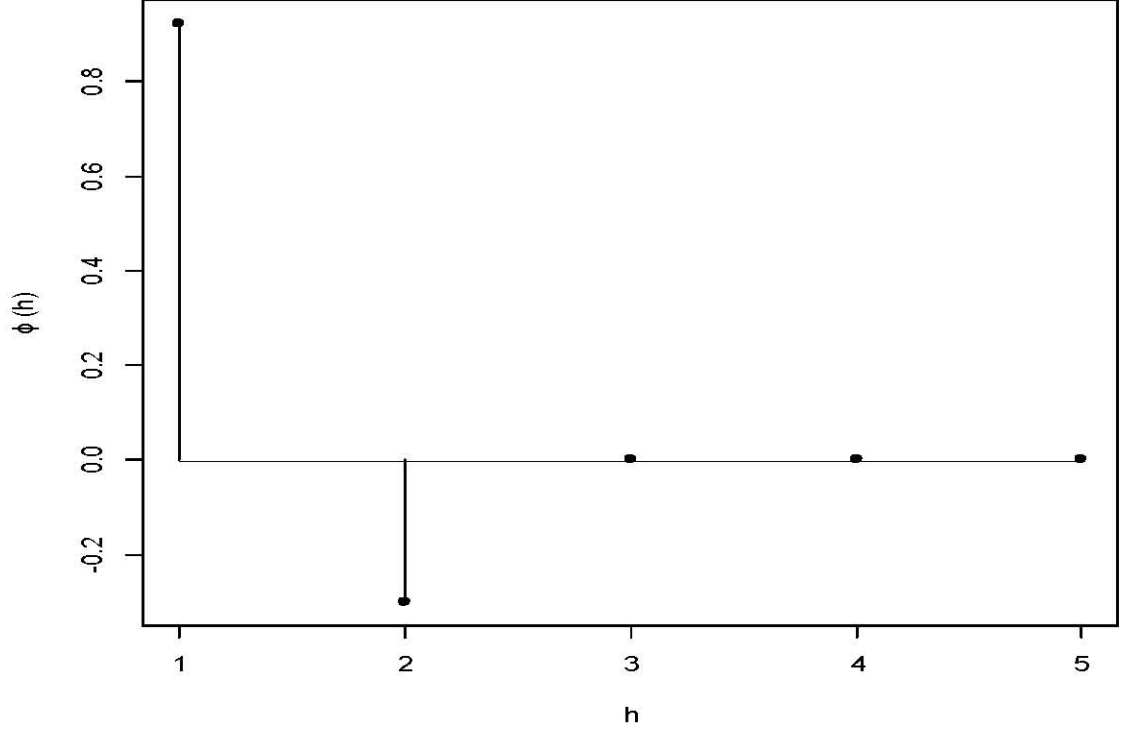
$$\phi(2) = \frac{\begin{vmatrix} 1 & \rho(1) \\ \rho(1) & \rho(2) \end{vmatrix}}{\begin{vmatrix} 1 & \rho(1) \\ \rho(1) & 1 \end{vmatrix}} = \alpha_2 = -0.3$$

$$\phi(3) = \frac{\begin{vmatrix} 1 & \rho(1) & \rho(2) \\ \rho(1) & 1 & \rho(2) \\ \rho(2) & \rho(1) & \rho(3) \end{vmatrix}}{\begin{vmatrix} 1 & \rho(1) & \rho(2) \\ \rho(1) & 1 & \rho(1) \\ \rho(2) & \rho(1) & 1 \end{vmatrix}} \approx 0$$

$$\phi(4) \approx 0$$

$$\phi(5) \approx 0$$

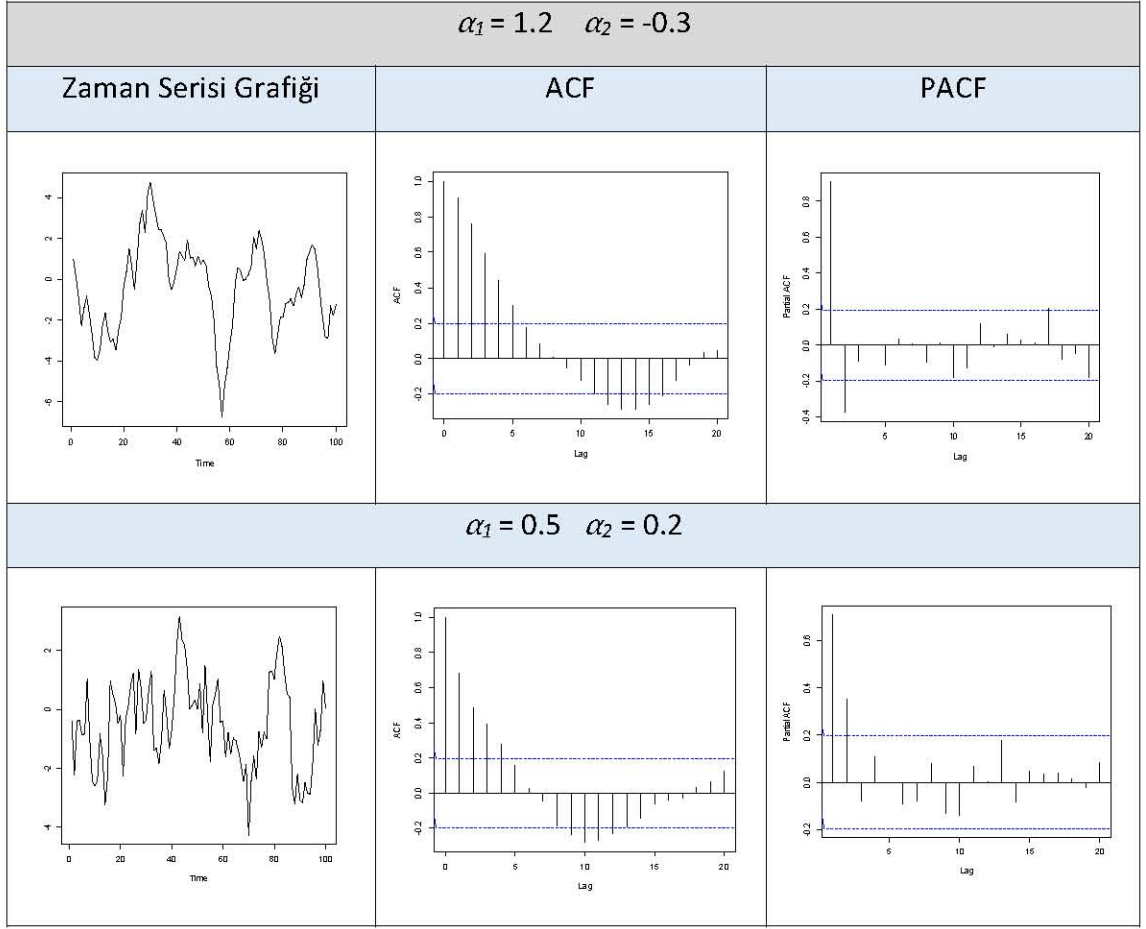
dir. Modelin kısmi otokorelasyonlarının grafiđi ařađıda verildiđi gibidir.



řekil 2.10 Kısmi otokorelasyon grafiđi

Grafikten de g r ld đ  gibi, kısmi otokorelasyonlar ikinci gecikmeden sonra sıfırdır. Eđer herhangi bir zaman serisinin $\rho(h)$ ve $\phi(h)$ grafikleri buna benzer ise serinin $AR(2)$ olarak modellenebileceđini sezgisel olarak s yleyebiliriz.

Ařađıda α_1 ve α_2 'nin deđiřik deđerleri i in $AR(2)$ modeline uygun rasgele  retilen 100 birimlik verilere iliřkin hesaplanan otokorelasyon ve kısmi otokorelasyonların grafikleri bulunmaktadır.



Şekil 2.11 α_1 ve α_2 'nin değişik değerleri için $AR(2)$ modeline ilişkin grafikler

Yukarıdaki örnekte görüldüğü gibi $AR(2)$ modelinin durağanlığı α_1 ve α_2 parametrelerinin değerlerine bağlıdır. $AR(2)$ zaman serisi modeli

$$X_t = \alpha_1 X_{t-1} + \alpha_2 X_{t-2} + e_t$$

şeklinde verildiğinde bu modelin karakteristik denklemi

$$m^2 - \alpha_1 m - \alpha_2 = 0$$

şeklinde dir. Eğer bu polinomun her iki kökü de mutlak değerce 1'den küçük ise model durağandır. Yukarıdaki örneğin karakteristik denklemi

$$m^2 - 1.2m + 0.3 = 0$$

olup denklemin kökleri $m_1 = 0.84$ ve $m_2 = 0.35$ bulunur. Bu kökler mutlak değerce 1'den küçük olduğu için model durağandır.

Daha genel hali ile p'inci dereceden durağan otoregresif zaman serisi modeli $e_t = WN(0, \sigma^2)$ olmak üzere,

$$(X_t - \mu) = \sum_{i=1}^p \alpha_i (X_{t-i} - \mu) + e_{t-i}$$

ya da $\alpha_0 = \mu(1 - \alpha_1 - \alpha_2 - \dots - \alpha_p)$ olmak üzere

$$X_t = \alpha_0 + \alpha_1 X_{t-1} + \alpha_2 X_{t-2} + \dots + \alpha_p X_{t-p} + e_t$$

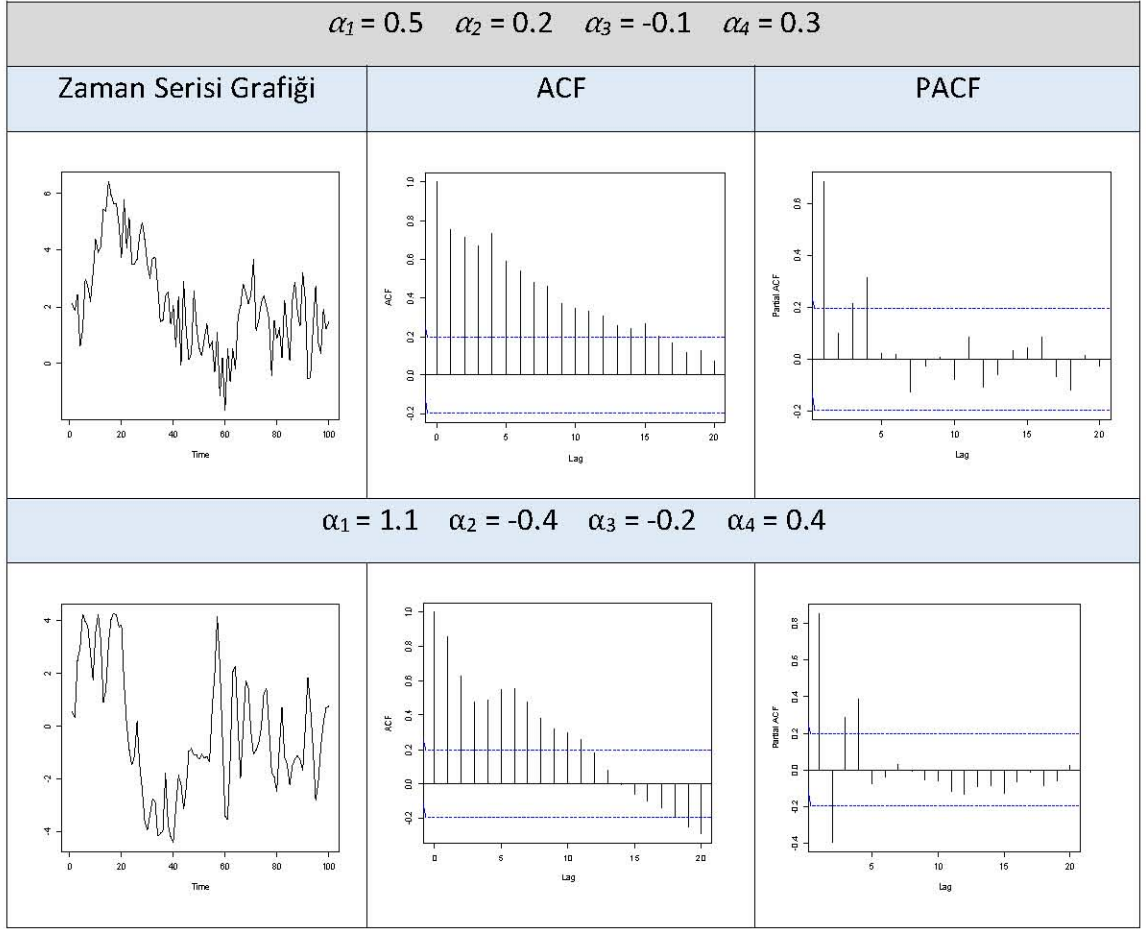
şeklinde verilir. $AR(p)$ modelinin durağanlığı

$$m^p - \sum_{i=1}^p \alpha_i m^{p-i} = 0$$

karakteristik denkleminin köklerine bağlıdır. Eğer bütün kökler mutlak değerce 1'den küçükse model durağandır. Eğer $\{X_t : t \in T\}$ durağan $AR(p)$ modeline uygun ise serinin otokorelasyonları

$$\rho(h) = \alpha_1 \rho(h-1) + \alpha_2 \rho(h-2) + \dots + \alpha_p \rho(h-p)$$

Yule-Walker eşitliğinden ardışık olarak hesaplanır. Yine, $AR(2)$ modelinde olduğu gibi otokorelasyonlar mutlak değerce üstel olarak azalmasına rağmen hiçbir zaman sıfır değildir. Oysa kısmi otokorelasyonlar p'inci gecikmeden sonra sıfırdır. Aşağıda $AR(4)$ modeline uygun rasgele üretilen 100 birimlik verilere ilişkin zaman serisi grafikleri bulunmaktadır.



Şekil 2.12 Farklı α değerleri için $AR(4)$ modeline ilişkin grafikler

Verilen ilk modelimiz,

$$X_t = 0.5X_{t-1} + 0.2X_{t-2} - 0.1X_{t-3} + 0.3X_{t-4}$$

şeklinde olup karakteristik denklemi

$$m^4 - 0.5m^3 - 0.2m^2 + 0.1m - 0.3 = 0$$

şeklinde. Bu denklemin kökleri $m_1 = -0.73$, $m_2 = 0.94$, $m_3 = 0.14 - 0.64i$ ve $m_4 = 0.14 + 0.64i$ bulunur. Bu kökler mutlak değerce 1'den küçük olduğu için model durağandır.

Ürettiğimiz diğer veri setimize uygun modelimiz,

$$X_t = 1.1X_{t-1} - 0.4X_{t-2} - 0.2X_{t-3} + 0.4X_{t-4}$$

şeklinde olup karakteristik denklemi

$$m^4 - 1.1m^3 + 0.4m^2 + 0.2m - 0.4 = 0$$

şeklindedir. Bu denklemin kökleri $m_1 = -0.60$, $m_2 = 0.93$, $m_3 = 0.385-0.749i$ ve $m_4 = 0.385+0.749i$ bulunur. Bu kökler mutlak değerce 1'den küçük olduğu için bu model de durağandır.

2.2.3 ARMA Serileri

Yukarıda, *MA* modellerinde otokorelasyonlar belli bir yerden sonra sıfır olup *AR* modellerinde ise otokorelasyonlar üstel olarak mutlak değerce azalmasına rağmen kısmi otokorelasyonlar belli bir yerden sonra sıfır olur. Bazen bu fonksiyonların her ikisi de sıfır olmayabilir. Bu durumda verileri *MA* ve *AR* modellerinin karışımı olan *ARMA* olarak modellenir. Liteartürde bu modeller bazen karışık (mixed) modeller olarak da bilinir. Model dereceleri p ve q olan *ARMA*(p,q) zaman serisi modeli $e_t = WN(0, \sigma^2)$ olmak üzere,

$$X_t = \mu + \alpha_0 + \sum_{i=1}^p \alpha_i X_{t-i}$$

şeklinde ifade edilir. Burada $\alpha_0 = \mu(1 - \sum \alpha_i)$ olup μ modelin beklenen değeridir. *ARMA* modellerinin durağanlığı da modelin *AR* kısmının durağanlığına bağlıdır. Yani, *AR* kısmına karşılık gelen

$$m^p - \sum_{i=1}^p \alpha_i m^{p-i} = 0$$

karakteristik denklemin bütün kökleri mutlak değerce 1'den küçük ise model durağandır. Ayrıca, *MA* kısmına karşılık gelen

$$m^q - \sum_{j=1}^q \beta_j m^{q-j} = 0$$

karakteristik denklemin bütün kökleri mutlak değerce 1'den küçük ise modele tersinirdir (invertible) denir.

Bilindiği gibi herhangi bir X_t zaman serisi durağan ise bu seri *MA*(∞) serisi olarak yazılabilir. Yani, $\sum |\beta_j| < \infty$ olmak üzere X_t serisi

$$X_t = \sum_{j=0}^{\infty} w_j e_{t-j}$$

olarak ifade edilebilir. Eğer model tersinir ise, $\sum |\pi| < \infty$ olmak üzere

$$e_t = \sum_{j=0}^{\infty} \pi_j X_{t-j}$$

şeklinde ifade edilebilir.

Örnek 2: $e_t = WN(0, \sigma^2)$ olmak üzere $ARMA(2,1)$ zaman serisi modeli

$$X_t = 1.2X_{t-1} - 0.3X_{t-2} + e_t + 2e_{t-1}$$

olarak verilmiş olsun.

Modelin durağanlığı tespit etmek için AR kısmına karşılık gelen karakteristik denklemin köklerine bakılmalıdır. Bu denklem,

$$m^2 - 1.2m + 0.3 = 0$$

olup kökleri $m_1 = 0.85$ ve $m_2 = 0.35$ olduğundan ve mutlak değerce 1'den küçük olduğundan model durağandır. Dolayısı ile verilen $ARMA(2,1)$ modeli $MA(\infty)$ olarak yazılabilir. Yani, bu model

$$X_t = \sum_{i=0}^{\infty} w_i e_{t-i}$$

şeklinde ifade edilebilir. w_j katsayıları,

$$w_0 = \beta_0 = 1, w_1 = \beta_1 + \alpha_1 = 1 + 1.2 = 2.2$$

başlangıç koşulları ile $i \geq 2$ için $w_i = 1.2 w_{i-1} - 0.3 w_{i-2}$ homojen fark denkleminin çözülmesi ile bulunur. $m_1 = 0.85$ ve $m_2 = 0.35$ kökleri ile genel çözüm

$$w_i = c_1 m_1^i + c_2 m_2^i$$

dir. $w_0 = 1$ ve $w_1 = 2.2$ başlangıç koşulları kullanılarak c_1 ve c_2 katsayıları için $w_0 = 1$ ise $c_1 + c_2 = 1$ ve $w_1 = 2.2$ ise $0.85c_1 + 0.35c_2 = 2.2$ denklemleri yazılabilir. Bu denklem sisteminin çözümünden $c_1 = 3.7$ ve $c_2 = -2.7$ bulunur ve genel çözüm

$$w_i = 3.7(0.85)^i - 2.7(0.35)^i$$

olarak elde edilir. O halde, verilen $ARMA(2,1)$ zaman serisi modeli

$$X_t = \sum_{i=0}^{\infty} [3.7(0.85)^i - 2.7(0.35)^i] e_{t-i}$$

şeklinde ifade edilir. Bu modelin başlangıçta bu şekilde verildiğini düşünelim. Böyle verilen $MA(\infty)$ zaman serisi modelinin $ARMA(2,1)$ şeklinde yazılabileceğini görelim. Bunun için X_{t-1} ve X_{t-2} zaman serileri

$$X_{t-1} = \sum_{i=0}^{\infty} [3.7(0.85)^i - 2.7(0.35)^i] e_{t-1-i} \quad \text{ve} \quad X_{t-2} = \sum_{i=0}^{\infty} [3.7(0.85)^i - 2.7(0.35)^i] e_{t-2-i}$$

şeklinde yazıldığında, eşitliğin sol tarafı $X_t - 1.2X_{t-1} + 0.3X_{t-2}$ olacak şekilde düzenlenirse,

$$\begin{aligned} X_t &= e_t + 2.2e_{t-1} + \sum_{k=2}^{\infty} [3.7(0.85)^k - 2.7(0.35)^k] e_{t-k} \\ -1.2X_{t-1} &= -1.2e_{t-1} + \sum_{k=2}^{\infty} [-5.23(0.85)^k + 9.31(0.35)^k] e_{t-k} \\ 0.3X_{t-2} &= \sum_{k=2}^{\infty} [1.53(0.85)^k - 6.61(0.35)^k] e_{t-k} \end{aligned}$$

eşitlikleri elde edilir. Bu ifadeler taraf tarafa toplandığında,

$X_t - 1.2X_{t-1} + 0.3X_{t-2} = e_t + e_{t-1}$ elde edilir. Yani $ARMA(2,1)$ olarak verilen X_t zaman serisi modeli

$$X_t = \sum_{i=0}^{\infty} w_i e_{t-i}$$

olarak yazılabilir. w_i katsayıları belirlenirse otokovaryans fonksiyonu

$$\gamma(h) = \sigma^2 \sum_{i=0}^{\infty} w_i w_{i+h}$$

şeklinde olur. Burada otokovaryans fonksiyonunun hesaplanabilmesi için (sonsuz toplam ile beklenen değerin yer değiştirebilir)

$$\sum_{i=0}^{\infty} w_i < \infty$$

olması gerekir. Otokovaryans fonksiyonu w_i katsayılarının yerine konulması ile,

$$\gamma(h) = \sigma^2 \sum_{i=0}^{\infty} w_i w_{i+h} = \sigma^2 \sum_{i=0}^{\infty} [3.7(0.85)^i - 2.7(0.35)^i] [3.7(0.85)^{i+h} - 2.7(0.35)^{i+h}]$$

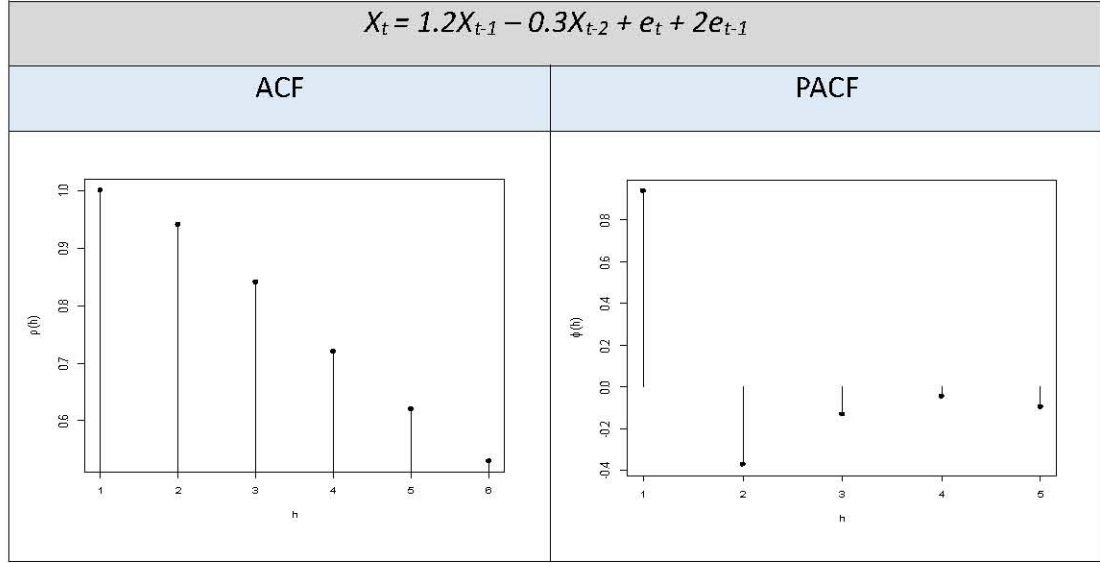
$\gamma(h) = \sigma^2 [35.26(0.85)^h - 5.77(0.35)^h]$ olarak bulunmuştur. Buradan otokorelasyon fonksiyonu da,

$$\rho(h) = (0.034)[35.26(0.85)^h - 5.77(0.35)^h]$$

şeklinde olur. Otokorelasyon ve kısmi otokorelasyonların bazıları hesaplanarak aşağıdaki tabloda verilmiş ve grafik olarak sunulmuştur.

Çizelge 2.3 $ARMA(2,1)$ modeline ait otokorelasyon ve kısmi otokorelasyon değerleri

h	0	1	2	3	4	5
ρ	1.00	0.94	0.84	0.72	0.62	0.53
ϕ		0.94	- 0.37	-0.13	- 0.045	-0.096



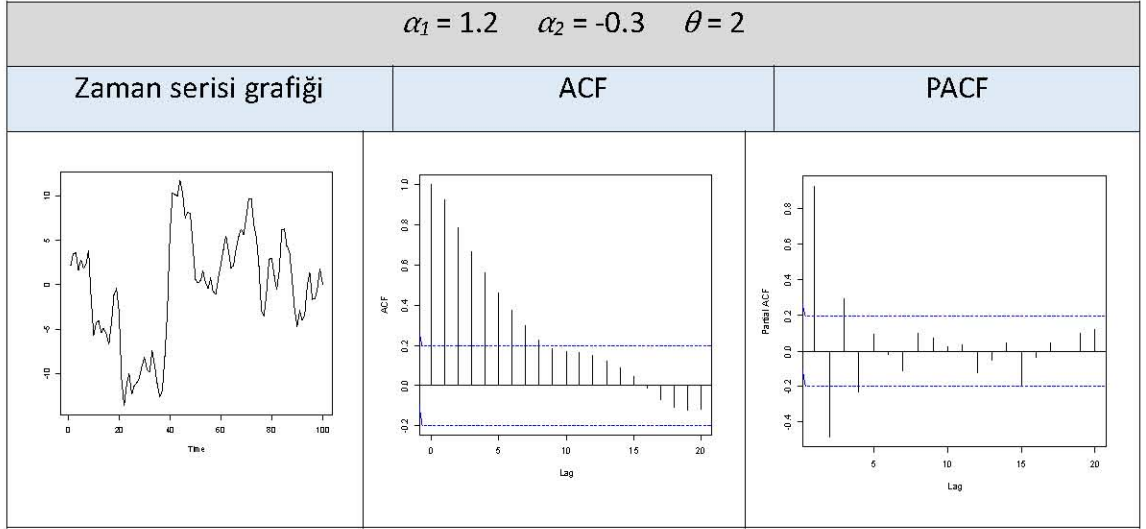
Şekil 2.13 $X_t = 1.2X_{t-1} - 0.3X_{t-2} + e_t + 2e_{t-1}$ modeline ait otokorelasyon ve kısmi otokorelasyon grafikleri

Her iki fonksiyon da mutlak değerce azalmasına rağmen, ne otokorelasyonlar ne de kısmi otokorelasyonlar belli bir noktadan sonra sıfır değildir.

Şimdi yukarıda verilmiş olan modelimize uygun rasgele 100 birimlik bir veri üretelim. Modelimiz ile ilgili otokorelasyon ve kısmi otokorelasyon değerleri ve ilgili grafikler çizelge 2.4'te verilmiştir.

Çizelge 2.4 $X_t = 1.2X_{t-1} - 0.3X_{t-2} + e_t + 2e_{t-1}$ modeline uygun üretilen veri setinin otokorelasyon ve kısmi otokorelasyon değerleri

h	0	1	2	3	4	5	6	7	8	9	10
ρ	1.00	0.925	0.786	0.664	0.558	0.459	0.376	0.298	0.228	0.185	0.169
ϕ		0.925	- 0.48	0.294	- 0.23	0.092	- 0.02	- 0.11	0.097	0.071	0.023



Şekil 2.14 Üretilen ARMA (2,1) modeline ait otokorelasyonve kısmi otokorelasyon grafikleri

Grafiklere bakıldığında serinin ne *MA* gibi ne de *AR* gibi davranmadığı görülür. Otokorelasyon grafiğinde *MA* kısmının derecesinden sonra sıfır olmadığı gibi kısmi otokorelasyon grafiği de *AR* kısmının derecesinden sonra sıfır olmamaktadır. Böyle bir durumda her iki modeli içine alacak *ARMA* modelinin kullanılması uygun olacağı söylenebilir.

Zaman serilerinde veriler birim zaman aralıklarında toplanır. Bu “birim zaman aralıkları” bazen gün, hafta, ay, yıl olabilir. Verilerin toplanış yapısına göre bütün seriler mevsimsel zaman serileri modellerine uygundur. Örneğin TÜİK tarafından yılda 4 defa (her üç ayda bir) açıklanan sanayi üretim indeksleri için mevsimsel zaman serisi modeli gözönüne alınabilir. Böyle bir veri için,

$$(X_t - \mu) = \alpha(X_{t-4} - \mu) + e_t$$

şeklinde ya da $\mu = 0$ olması halinde

$$X_t = \alpha X_{t-4} + e_t$$

şeklinde ifade edilebilir. Bu modelde dördüncü dereceden otoregresif zaman serisi modeline benzemektedir. Dolayısı ile modelin durağanlığı

$$m^4 - \alpha = 0$$

denkleminin köklerine bağlıdır. Bu denklemin bütün kökleri mutlak değerce 1'den küçükse model durağandır. Eğer $m^4 = z$ denirse bu denklem $AR(1)$ modelinin karakteristik denklemine benzer ($z - \alpha = 0$). Dolayısı ile model $SAR_4(1)$ olarak ifade edilir. Benzer şekilde, X_t zaman serisi modeli

$$X_t = \alpha_0 + \alpha_4 X_{t-4} + \alpha_8 X_{t-8} + e_t$$

şeklinde verilmiş olsaydı bu modele karşılık gelen karakteristik denklem

$$m^8 - \alpha_4 m^4 - \alpha_8 = 0$$

olacaktı ve $m^4 = z$ denirse denklem $z^2 - \alpha_4 z - \alpha_8$ şeklinde $AR(2)$ modelinin karakteristik denklemine dönüşür. Bu nedenle model $X_t \sim SAR_4(2)$ olarak ifade edilir.

Örnek 3: $e_t = WN(0, \sigma^2)$ olmak üzere X_t zaman serisinin

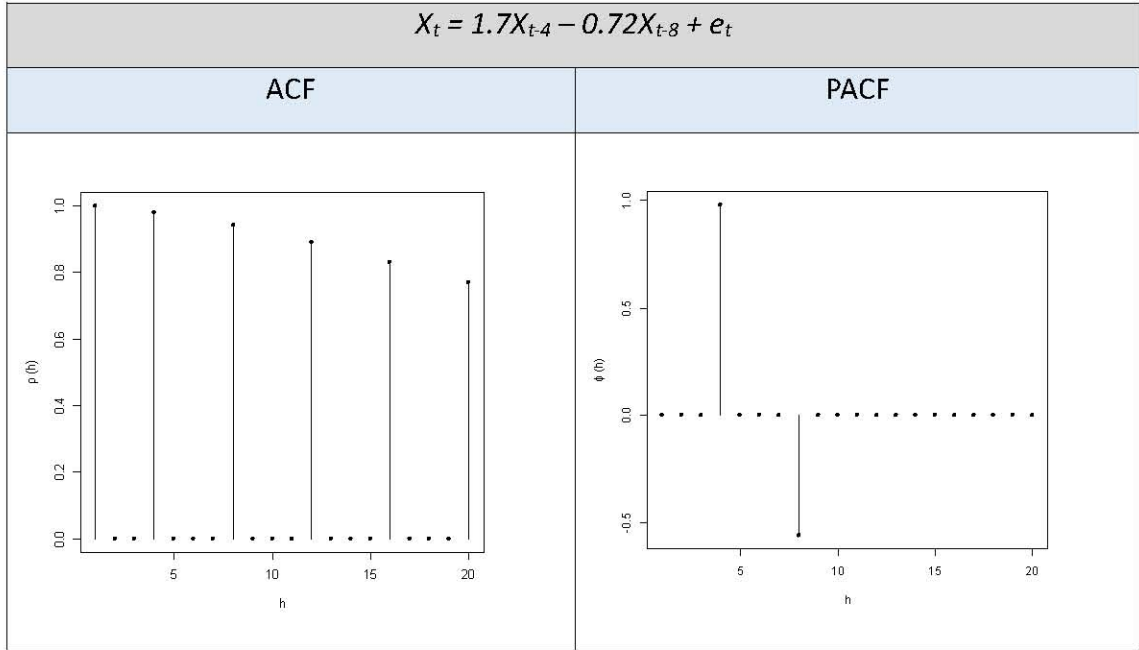
$$X_t = 1.7X_{t-4} - 0.72X_{t-8} + e_t$$

modeline uygun olduğunu düşünelim. $m^8 - 1.7m^4 + 0.72 = 0$ denkleminin bütün kökleri mutlak değerce 1'den küçük olduğundan model durağandır. $\alpha_4 = 1.7$ ve $\alpha_8 = -0.72$, bu iki değer bilmesi ile bütün otokorelasyon ve kısmi otokorelasyonlar hesaplanabilir. Bazı değerlere karşılık gelen otokorelasyon ve kısmi otokorelasyonlar hesaplanmış ve grafikleri çizdirilerek aşağıda sunulmuştur.

Çizelge 2.5 $X_t = 1.7X_{t-4} - 0.72X_{t-8} + e_t$ modelinin otokorelasyon ve kısmi otokorelasyon değerleri

Otokorelasyon ve kısmi otokorelasyon değerleri											
h	0	1	2	3	4	5	6	7	8	9	10
ρ	1.00	0	0	0	0.98	0	0	0	0.94	0	0
ϕ		0	0	0	0.98	0	0	0	0.56	0	0
h	11	12	13	14	15	16	17	18	19	20	
ρ	0	0.89	0	0	0	0.83	0	0	0	0.77	
ϕ	0	0	0	0	0	0	0	0	0	0	

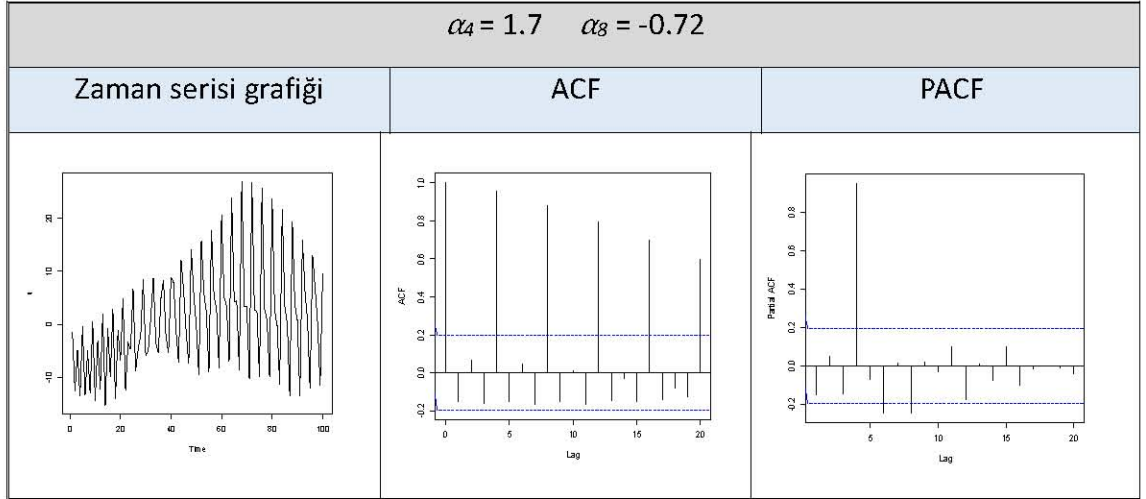
Şekil 2.15 $X_t = 1.7X_{t-4} - 0.72X_{t-8} + e_t$ modeline ait otokorelasyon ve kısmi otokorelasyon grafikleri



Şimdi yukarıda verilmiş olan modelimize uygun rasgele 100 birimlik bir veri üretelim. Modelimiz ile ilgili otokorelasyon ve kısmi otokorelasyon değerleri ve ilgili grafikler aşağıdadır.

Çizelge 2.6 $X_t = 1.7X_{t-4} - 0.72X_{t-8} + e_t$ modeline uygun üretilen veri setine ait otokorelasyon ve kısmi otokorelasyon değerleri

Otokorelasyon ve kısmi otokorelasyon değerleri											
h	0	1	2	3	4	5	6	7	8	9	10
ρ	1.00	-0.15	0.069	-0.16	0.953	-0.15	0.046	-0.16	0.88	-0.15	0.013
ϕ		-0.15	0.047	-0.14	0.951	-0.06	-0.24	0.014	-0.24	0.019	-0.03
h	11	12	13	14	15	16	17	18	19	20	
ρ	-0.165	0.794	-0.148	-0.031	-0.149	0.697	-0.14	-0.08	-0.127	0.598	
ϕ	0.1	-0.176	0.006	-0.078	0.102	-0.101	-0.018	-0.001	-0.012	-0.04	



Şekil 2.16 $X_t = 1.7X_{t-4} - 0.72X_{t-8} + e_t$ modeline uygun üretilen veriye ait otokorelasyon ve kısmi otokorelasyon grafikleri

Grafiklere bakıldığında verideki dalgalanmanın belirli periyotlarda tekrar ettiğini görebiliriz. Modelde 4'üncü ve 8'inci birim zamanlarında tekrar eden bir dalgalanma söz konusudur. Böyle durumlarda mevsimsellikten söz edilebilir ve bu veriler mevsimsel zaman serisi modelleri ile modellenebilirler.

Aylık veriler için kullanılan $SAR_{12}(1)$ modeli $e_t \sim WN(0, \sigma^2)$ olmak üzere,

$$X_t = \alpha_0 + \alpha_1 X_{t-12} + e_t, \quad t = 1, 2, \dots, n$$

şeklinde (veya $(X_t - \mu) = \alpha (X_{t-12} - \mu) + e_t$) verilmiş olsun. Modelin otokovaryansları

$$\gamma(h) = \alpha \gamma(h-12) + \sigma^2 I(h=0)$$

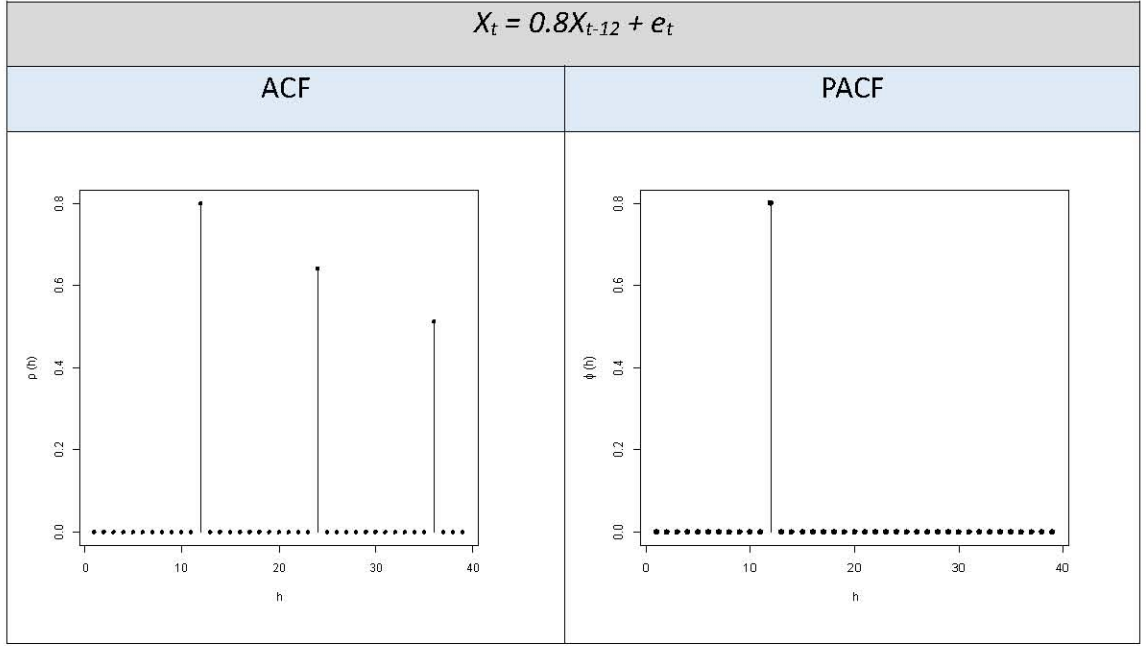
şeklinde olup

$$\gamma(h) = \begin{cases} \alpha^{\frac{h}{12}} \frac{\sigma^2}{1-\alpha^2}, & h = 12s, s = 0, 1, 2, \dots \\ 0, & d.d. \end{cases}$$

olarak yazabiliriz. Otokorelasyonlar ise

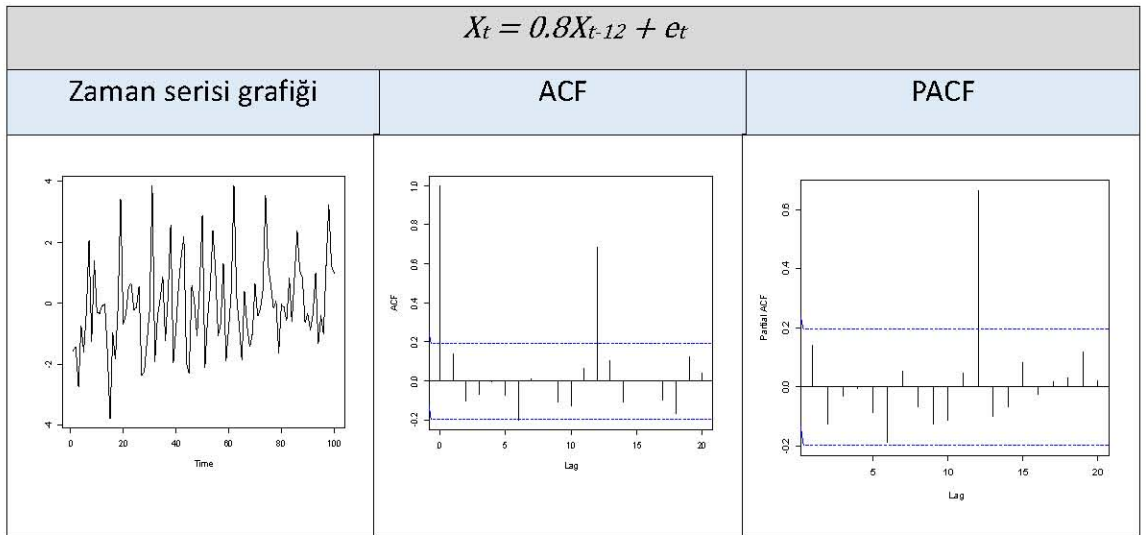
$$\rho(h) = \begin{cases} \alpha^{\frac{h}{12}}, & h = 12s, s = 0, 1, 2, \dots \\ 0, & d.d. \end{cases}$$

şeklinde olur. $\alpha = 0.8$ için otokorelasyonların ve kısmi otokorelasyonların grafiği şekil.2.17'de verilmiştir.



Şekil 2.17 $X_t = 0.8X_{t-12} + e_t$ modeline ait otokorelasyon ve kısmi otokorelasyon grafikleri

Ayrıca bu modele uygun rasgele üretilen 100 birimlik verilere ilişkin grafikler şekil 2.18'de verilmiştir.



Şekil 2.18 $X_t = 0.8X_{t-12} + e_t$ modeline uygun üretilen veriye ait otokorelasyon ve kısmi otokorelasyon grafikleri

3. DURAĞAN OLMAYAN ZAMAN SERİLERİ

Bir önceki bölümde bazı durağan zaman serisi modelleri incelendi. Orada da görüldüğü gibi $MA(q)$ modelleri her zaman durağandır. Ancak AR veya AR bileşeni içeren modeller durağan olmayabilir.

Zaman serilerinde en önemli amaçlarından biri serinin gelecekteki değerlerinin öngörülmesidir. Öngörü yapabilmek için de en önemli varsayım serinin durağanlığıdır. Birçok iktisadi zaman serisi durağan değildir. Durağan olmayan seriler denildiğinde ilk akla gelen birim köklü zaman serileridir. Zaman serileri analizinde durağanlığın öngörü yapmak için bir şart olduğu gerçeğinden hareketle durağanlığından şüphelenilen veri setlerinde yapılması gereken ilk şey bu veri setinin durağan olup olmadığının test edilmesidir.

3.1 Birim Kök Testleri

Parametrelerin tahmin edicilerinin farklı özelliklerine göre değişen birim kök testleri olmakla beraber bunların arasında en yaygın olan ve parametrelerin EKK tahmin edicisinin dağılımına bağlı olarak geliştirilen Dickey-Fuller birim kök testi ile bu yöntemin modife edilmiş hali olan Phillips-Perron birim kök testinden bahsedilecektir.

3.1.1 Dickey-Fuller birim kök testi

Herhangi bir $AR(1)$ zaman serisi modeli $e_t \sim WN(0, \sigma^2)$ olmak üzere

$$(Y_t - \mu) = \rho(Y_{t-1} - \mu) + e_t$$

veya $X_t = Y_t - \mu$ denildiğinde

$$X_t = \rho X_{t-1} + e_t$$

şeklinde verilmiş olsun. Burada, $|\rho| < 1$ ise model durağandır ve ρ' nun en küçük kareler tahmin edicisi

$$\hat{\rho}_n = \frac{\sum X_t X_{t-1}}{\sum X_{t-1}^2}$$

olmak üzere,

$$\hat{\rho}_n \xrightarrow{P} \rho, n \rightarrow \infty$$

ve

$$\sqrt{n}(\hat{\rho}_n - \rho) \xrightarrow{D} N(0, 1 - \rho^2) \text{ dir (Fuller, 1996).}$$

Eğer $\rho = 1$ ise model $X_t = X_{t-1} + e_t$ şeklinde olup

$$X_t = \sum_{j=1}^{X_0+t} e_j$$

şeklinde yazılabilir. Buradan $X_0=0$ (veya $Y_0 = \mu$) varsayımı altında ρ 'nun en küçük kareler tahmin edicisi,

$$\hat{\rho}_n = \frac{\sum_{t=1}^n X_t X_{t-1}}{\sum_{t=1}^n X_{t-1}^2} = \rho + \frac{\sum_{t=1}^n e_t X_{t-1}}{\sum_{t=1}^n X_{t-1}^2}$$

eşitliğinden $\rho = 1$ için

$$\hat{\rho}_n - 1 = \frac{\sum_{t=1}^n e_t X_{t-1}}{\sum_{t=1}^n X_{t-1}^2}$$

şeklinde yazılır. Ayrıca,

$$\sum_{t=1}^n e_t X_{t-1} = O_p(n) \text{ ve } \sum_{t=1}^n X_{t-1}^2 = O_p(n^2)$$

olduğundan

$$n(\hat{\rho}_n - 1) = \frac{\frac{1}{n} \sum_{t=1}^n e_t X_{t-1}}{\frac{1}{n^2} \sum_{t=1}^n X_{t-1}^2}$$

olup

$$n(\hat{\rho}_n - 1) = O_p(1) \text{ dir (Fuller 1996).}$$

Diğer taraftan,

$$\frac{1}{n} \sum_{t=1}^n e_t X_{t-1} = \frac{1}{n} \sum_{t=1}^n e_t \left(\sum_{j=1}^{t-1} e_j \right) = \frac{1}{n} \left[\left(\sum_{t=1}^n e_t \right)^2 - \left(\sum_{t=1}^n e_t^2 \right) \right]$$

olduğundan

$$\frac{1}{n} \sum_{t=1}^n e_t^2 X_{t-1} \xrightarrow{D} \frac{\sigma^2}{2} (\chi_{t-1}^2)$$

olup

$$\hat{\rho}_n - 1 = \frac{\sum_{t=1}^n e_t X_{t-1}}{\sum_{t=1}^n X_{t-1}^2}$$

$$n(\hat{\rho}_n - 1) = \frac{\frac{1}{n} \sum_{t=1}^n e_t X_{t-1}}{\frac{1}{n^2} \sum_{t=1}^n X_{t-1}^2} \xrightarrow{P} \frac{\chi_1^2 - 1}{2 \int_0^1 w^2(t) dt}$$

elde edilir (Fuller 1996).

Bu dağılım Dickey–Fuller dağılımlarından biridir. Benzer şekilde t–türü istatistik ise

$$S_n^2 \frac{1}{n-2} \sum_{t=1}^n (X_t - \hat{\rho}_n X_{t-1})^2 = \frac{1}{n-2} \sum_{t=1}^n \hat{e}_t^2$$

olmak üzere

$$\hat{\tau}_n = \frac{\hat{\rho}_n - 1}{S(\hat{\rho}_n)} = \frac{\hat{\rho}_n - 1}{\sqrt{S_n^2 \left(\sum_{t=1}^n X_{t-1} \right)^{-\frac{1}{2}}}}$$

olup bu dağılım doğrudan standart t–dağılımı değildir. Literatürde bu dağılım da Dickey–Fuller dağılımı olarak bilinir. Bunlar sırası ile $n(\hat{\rho}_n - 1)$, $n(\hat{\rho}_{n,\mu} - 1)$, $n(\hat{\rho}_{n,\tau} - 1)$ ve t–türü istatistikler de $\hat{\tau}$, $\hat{\tau}_\mu$, $\hat{\tau}_\tau$ olarak isimlendirilmiştir. Bu istatistiklerin dağılımları ve kritik değerleri hemen hemen bütün zaman serisi kitaplarında bulunabileceği gibi birçok paket program bu istatistiklerin değerlerini ve kritik değerlerini doğrudan hesaplar (Eviews gibi).

Yukarıda $AR(1)$ modeli

$$(Y_t - \mu) = \rho(Y_{t-1} - \mu) + e_t$$

veya

$$X_t = \rho X_{t-1} + e_t$$

şeklinde verildiğinde $|\rho| < 1$ olması halinde modelin durağan olduğunu biliyoruz. $\rho = 1$ olması halinde

$$X_t = \sum_{j=1}^t e_j$$

olup

$$Var(X_t) = t\sigma^2 \text{ ve } Cov(X_t, X_{t+h}) = \sigma^2 \min\{t, t+h\}$$

olduğundan model durağan değildir. Ayrıca,

$$(Y_t - \mu) = \rho(Y_{t-1} - \mu) + e_t$$

modeli $\alpha_0 = \mu(1-\rho)$ olmak üzere

$$Y_t = \alpha_0 + \rho Y_{t-1} + e_t$$

şeklinde de yazılabilir. Eğer $\rho = 1$ ise μ parametresi modelden düşer. Ayrıca ρ ' nun en küçük kareler tahmin edicilerinin dağılımları da bilinen dağılımlardan farklıdır. Bu dağılımların kritik değerleri kullanılarak $H_0: \rho = 1$ yokluk hipotezi $H_a: |\rho| < 1$ alternatif hipotezine karşı test edilir. Eğer test istatistiğinin hesaplanan değeri kritik değerden küçük ise H_0 hipotezi reddedilir (Yani serinin durağan olduğu sonucuna varılır). Genellikle t-türü istatistikler kullanılır. Yukarıda $AR(1)$ modeli için açıklanan en küçük kareler tahmin ediciler ve bunların $H_0: \rho = 1$ altındaki dağılımları daha yüksek dereceden modeller için de uygulanabilir. Örneğin $AR(p)$ zaman serisi modeli $e_t \sim WN(0, \sigma^2)$ olmak üzere,

$$X_t = \alpha_0 + \sum_{i=1}^p \alpha_i X_{t-i} + e_t \quad (3.1)$$

şeklinde verildiğinde, bu modelin durağanlığının

$$m^p - \sum_{i=1}^p \alpha_i m^{p-i} = 0$$

karakteristik denkleminin köklerine bağlı olduğunu biliyoruz. Eğer $m=1$ bu denklemin bir

kökü ise $\sum_{i=1}^p \alpha_i = 1$ ($m=1 \Leftrightarrow \sum_{i=1}^p \alpha_i = 1$) dir. Diğer taraftan (3.1) de verilen $AR(p)$ modeli

$$\theta = \sum_{i=1}^p \alpha_i, \theta_i = -\sum_{j=1}^p \alpha_j \quad i=2,3,\dots,p$$

olmak üzere

$$X_t = \theta_1 X_{t-1} + \sum_{i=2}^p \theta_i (X_{t-i} - X_{t-i+1}) + e_t, \quad t = 1, 2, \dots, n \quad (3.2)$$

şeklinde de yazılabilir. Bu nedenle H_0 : seri birim köklüdür hipotezi $H_0: \theta = 1$ hipotezine denktir. Dolayısı ile θ_1 parametresinin en küçük kareler tahmin edicisinin dağılımı da yukarıda açıklanan Dickey–Fuller dağılımlarından biridir. O halde $AR(p)$ modelleri içinde serinin durağanlığının testi için aynı dağılımlar (yani kritik değerler) kullanılabilir.

Bir serinin Dickey–Fuller testi ile birim köklü olup olmadığı sınaması aşağıda anlatılmıştır:

A) Model: $(X_t - \mu) = \alpha(X_{t-1} - \mu) + e_t, t = 1, 2, \dots, n$

i) X_t : t zamanındaki gözlem

ii) e_t : hata terimi (normal olduğu varsayılmaktadır)

iii) μ : beklenen değer

iv) α : otoregresif model parametresi

B) Problem: $H_0: \alpha = 1$ yokluk hipotezinin test edilmesi

i) $\alpha = 1$ için μ modelden düşer ve öngörüler ortalamaya yaklaşmaz

ii) Test istatistiğinin değerini hesaplamak için modelin her iki tarafından X_{t-1} çıkartılarak model $X_t - X_{t-1} = (\alpha - 1)X_{t-1} + e_t$ şeklinde yazılır. Buradan, $D_t = X_t - X_{t-1}$ ler hesaplanır.

iii) D_t lerin X_{t-1} üzerine regresyonu yapılır. Bu regresyon modeline göre X_{t-1} in katsayısının sıfır olup olmadığı yukarıda belirtilen test istatistiklerinin $(n(\hat{\rho}_n - 1), n(\hat{\rho}_{n,\mu} - 1), n(\hat{\rho}_{n,\tau} - 1))$ veya bu istatistiklerinin yerine kullanılabilecek olan t-türü istatistiklerinin $(\hat{\tau}, \hat{\tau}_\mu, \hat{\tau}_\tau)$ değerleri ile kritik değerler karşılaştırılır. Hesaplanan değer, kritik değerden küçük ise $H_0: \alpha = 1$ yokluk hipotezi red edilir.

iv) Bu yöntem yüksek dereceden modeller için de geçerlidir.

C) Sonraki adım trendden arındırma safhasıdır. Şu şekilde yapılır:

i) $D_t = X_t - X_{t-1}$ değişkeni X_{t-1} üzerine regresyonu yapılır (α_0 yok)

ii) $D_t = X_t - X_{t-1}$ değişkeni X_{t-1} üzerine regresyonu yapılır (α_0 var)

iii) $D_t = X_t - X_{t-1}$ değişkeni X_{t-1} ve t üzerine regresyonu yapılır (α_0 yok)

iv) Her üç regresyon için test istatistiğinin değerleri modelin kesim noktası ve trend olup olmadığına göre hesaplanır ve kritik (tablo) değeri ile karşılaştırılır.

D) Yüksek dereceden modellerde uygulanması şöyledir:

i) Model : $(X_t - \mu) = \sum_{i=1}^p \alpha_i (X_{t-i} - \mu) + e_t, t = 1, 2, 3, \dots, n$

ii) $D_t = X_t - X_{t-1}$ olmak üzere, $D_{t-1}, D_{t-2}, \dots, D_{t-p}$ farkları hesaplanır.

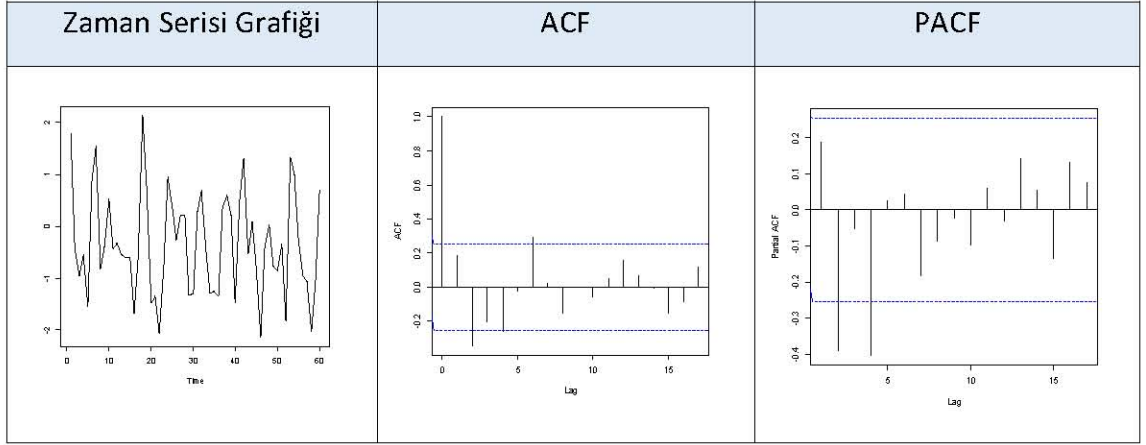
iii) D_t nin $X_{t-1}, D_{t-1}, D_{t-2}, \dots, D_{t-p}$ üzerine regresyonu yapılır. Bu regresyona göre X_{t-1} in katsayısının $(1 - \alpha_1 - \alpha_2 - \dots - \alpha_p)$ sıfır olup olmadığını test etmek için B şikkında belirtilen test istatistiklerinin (veya t-türü istatistikler) değerleri hesaplanır ve tablo değeri ile karşılaştırılır.

Örneğin yeni kurulan bir A silah firmasının ay bazında 5 yıllık kar/zarar durumu çizelge 3.1'de verilmiştir.

Çizelge 3.1 A firmasının ay bazında 5 yıllık kar/zarar durumu (* Yüz Bin TL)

1.78	-0.43	-0.95	-0.54	-1.54	0.89	1.54	-0.81	-0.39	0.53	-0.43	-0.32
-0.54	-0.60	-0.60	-1.67	-0.59	2.14	0.75	-1.47	-1.35	-2.06	-0.78	0.95
0.38	-0.26	0.21	0.22	-1.33	-1.30	0.28	0.70	-0.46	-1.29	-1.25	-1.35
0.38	0.60	0.22	-1.47	0.51	1.31	-0.53	0.09	-0.78	-2.14	-0.38	0.04
-0.77	-0.85	-0.34	-1.82	1.33	1.00	-0.25	-0.94	-1.07	-2.02	-1.06	0.70

Bu veri setinin zaman serisi grafikleri şekil 3.1'de verilmiştir.



Şekil 3.1 A firmasının ay bazında 5 yıllık kar/zarar durumuna ait verilerinin zaman serisi grafikleri

Veri setinin grafikleri incelendiğinde $AR(1)$ olarak modellenmesinin uygun olacağı değerlendirilmektedir. Buradan verilere,

$$(X_t - \mu) = \alpha_0 + \alpha X_{t-1} + e_t, t = 1, 2, 3, \dots, 60$$

şeklinde $AR(1)$ modelinin uygun olduğunu düşünelim. Bu model dikkate alınarak elde edilen parametre tahminleri ve bazı istatistikî sonuçlar çizelge 3.2'de verilmiştir.

Çizelge 3.2 $(X_t - \mu) = \alpha_0 + \alpha X_{t-1} + e_t$ modeline ilişkin parametre tahminleri

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	1	2.04979	2.04979	2.23	0.1405
Error	57	52.29862	0.91752		
Corrected Total	58	54.34841			
Parameter Estimates					
Variable	DF	Parameter Estimate	Standard Error	t Value	Pr > t
Intercept	1	-0.30556	0.13243	-2.31	0.0247
x1	1	0.18830	0.12598	1.49	0.1405

Model,

$$X_t = \alpha_0 + \alpha X_{t-1} + e_t, t = 1, 2, 3, \dots, 60$$

şeklinde yazılarak, ∇X_t 'nin X_{t-1} üzerine regresyonundan elde edilen istatistiki değerler aşağıdadır. t-türü istatistiğinin değeri $\hat{\tau}_{n,\mu} = -6.44$ olarak bulunmuştur. Bu değer % 5 anlam düzeyindeki kritik değerden küçük ise $H_0 : \alpha = 1$ yokluk hipotezi red edilir.

Çizelge 3.3 ∇X_t 'nin X_{t-1} üzerine regresyonundan elde edilen istatistiki değerler

Parameter Estimates					
Variable	DF	Parameter Estimate	Standard Error	t Value	Pr > t
Intercept	1	-0.30556	0.13243	-2.31	0.0247
x1	1	-0.81170	0.12598	-6.44	<.0001

$$\hat{\tau}_{n,\mu} = -6.44 < -2.91 = \hat{\tau}_{\mu}$$

olduğundan $H_0 : \alpha = 1$ hipotezi red edilir. Buna göre, A firmasının yıllık durumuna ait veri birim köklü değildir.

3.1.2 Phillips – Perron birim kök testi

Dickey-Fuller birim kök testi hata terimlerinin istatistiksel olarak bağımsız olduğu ve sabit varyansa sahip olduğu varsayımına dayanır. Phillips – Perron hata terimlerinin otokorelasyonlu olduğu durumlarda DF testinin hatalı olabileceği düşüncesinden hareketle DF test istatistiğine bir düzeltme faktörü ekleyerek hataların otokorelasyon problemine çözüm getirmeye çalışmıştır (Akdi 2012). Yani Dickey–Fuller'in hata terimleri ile ilgili varsayımlarını genişletmiştir. Diğer testlerin hata terimlerini,

$$\psi(B) = 1 + \beta_1 B + \beta_2 B^2 + \dots + \beta_q B^q$$

olarak tanımlanan bir düzeltme faktörü ile modife ederek $u_t = \psi(B)e_t$ şekline çevrilmiş olsun. Durağanlığı test etmek için Dickey–fuller test istatistiklerine bir düzeltme faktörü eklenir ve aynı şekilde modelin sabit terim ve trend içerip içermemesine göre test istatistikleri belirlenir. Örneğin,

$$(X_t - \mu) = \alpha(X_{t-1} - \mu) + u_t, t = 1, 2, \dots, n$$

$AR(1)$ modelini gözönüne alalım. Bu model,

$$X_t = \alpha_0 + \alpha X_{t-1} + u_t, t = 1, 2, 3, \dots, n$$

olarak yazılabilir. Burada, $H_0: \alpha = 1$ yokluk hipotezini test etmek için Dickey–Fuller istatistiklerine belirtilen düzeltme faktörü eklendikten sonra X_t nin X_{t-1} üzerine regresyonundan (daha yüksek dereceden modeller için de benzerdir) elde edilen artıklar $\hat{u}_t$ olsun. Bu artıklar serisinin otokovaryanslarını $\hat{\gamma}(h)$ ile gösterelim. Ayrıca,

$$\hat{\lambda}^2 = \hat{\gamma}(0) + 2 \sum_{j=1}^q [1 - j / (q+1)] \hat{\gamma}(j)$$

olsun. X_t nin X_{t-1} üzerine regresyonundan hata kareler ortalaması,

$$S_n^2 = \frac{1}{n-2} \sum_{t=1}^n \hat{u}_t^2$$

dir. Buna göre Phillips–Perron birim kök test istatistiği,

$$Z_\mu = \left[\sqrt{\frac{\hat{\gamma}(0)}{\hat{\lambda}^2}} \right] \hat{\tau}_\mu - \left\{ \frac{1}{2} (\hat{\lambda}^2 - \hat{\gamma}(0)) (n S_{\hat{\alpha}} / S_n) \right\} \hat{\lambda}^{-1}$$

olarak bulunur. $S_{\hat{\alpha}}$ değeri, $\hat{\alpha}$ nın standart hatasıdır.

Yukarıda verilmiş olan A silah firmasının kar/zarar verilerine bu yöntemi uyguladığımızda, yani $H_0: \alpha = 1$ yokluk hipotezini %5 anlam düzeyinde test ettiğimizde

$$Z_{n,\mu} = -6.264 < -2.91 = \hat{\tau}_\mu (tablo)$$

olarak bulunur. Buradan $H_0: \alpha = 1$ yokluk hipotezi %5 anlam düzeyinde reddedilir. Yani, A firmasının kar/zarar verileri birim köklü değildir.

Çizelge 3.4 A firması kar/zarar verisi Phillips-Perron testi

		Adj. T-Stat	Prob.
Phillips-Perron test statistic		-6.264	0.01
Test critical values	1% level	-3.54	
	5% level	-2.91	
	10% level	-2.59	

3.2 Mevsimsel Birim Kök Testleri

Durağan mevsimsel zaman serisi önceki bölümlerde incelenmişti. Bu kısımda, verilen bir zaman serisinin mevsimsel birim köklü olup olmadığının nasıl sınanacağı incelenecektir. Mevsimsel durağan zaman serilerinde otokorelasyonlar periyodik bir şekilde azalır. Bu azalma bazen yavaş olabilir. Böyle durumlarda serinin durağanlığından şüphelenilir. Bununla beraber, mevsimsel zaman serisi modeline karşılık gelen karakteristik denklemin en az bir tane birim kökü mutlak değerce 1 ise, bu birim köklerin tekrar ettiğini biliyoruz. Bu durumda, birim kök sayısının tespiti için ardışık birim kök testlerinin uygulanması doğru olmayabilir. Bunu görebilmek için $SAR_4(1)$ modeli

$$X_t = \alpha X_{t-4} + e_t, t = 1, 2, \dots, n$$

şeklinde verilmiş olsun. Modelin karakteristik denklemi $m^4 - \alpha = 0$ dır. Bu dördüncü dereceden polinomun 4 tane kökü vardır. $|\alpha| < 1$ ise model durağandır. $\alpha = 1$ için model durağan değildir. $m^4 - 1 = 0$ polinomunun kökleri, $m_{1,2} = \pm 1$ ve $i = \sqrt{-1}$ olmak üzere $m_{3,4} = \pm i$ şeklindedir. Bütün kökler mutlak değerce 1' dir. Buradan serinin 4 tane birim kökü vardır. Başka bir ifade ile, bu seriyi durağan hale getirebilmek için serinin 4 defa farkının alınması gerekir. Oysa, bu hem doğru olmayabilir hem de pratik değildir. Bu seri 4 birim köklü ise, $\nabla^4 X_t = (1-B)^4 X_t$ modelinin durağan olması beklenir. Böyle bir dönüşüm modeli çok karmaşık hale getirebilir. Serinin dördüncü dereceden farkı

$$(1-B)^4 X_t = (1-4B+6B^2-4B^3+B^4)X_t$$

şeklindedir. Yukarıdaki basit regresyon modeli çok karmaşık hale dönüşür. Bunun yerine, $\nabla_4 X_t = (1-B^4)X_t = X_t - X_{t-4}$ olup, model $\nabla_4 X_t = e_t$ haline gelir. Bu model durağandır.

Örnek 4: Elimizde aşağıdaki gibi bir veri seti olsun.

Çizelge 3.5 Üretilen veri seti

22.9	23.9	22.9	22.6	20.1	21.3	21.9	23.1	19.9	19.9
21.3	22.5	19.7	19.7	22.3	21.5	18.2	17.8	24.0	21.4
19.0	16.6	23.7	19.7	19.2	15.3	23.1	19.8	18.5	15.0
23.4	19.5	17.7	14.6	22.1	18.5	16.5	13.2	24.1	19.0
17.6	14.3	24.0	19.5	15.5	14.3	23.6	17.4	14.4	15.7
22.0	15.2	15.3	15.2	21.2	14.5	16.6	15.4	21.0	14.4

Bu verilere bir model uydurmak için AIC ve SBC değerleri dikkate alınarak bu veri setine

$$(X_t - \mu) = \alpha(X_{t-4} - \mu) + e_t, t = 1, 2, 3, \dots, 60$$

şeklinde bir modelin uygun olacağı değerlendirilmiştir. Model parametrelerine ait tahmin değerleri ve bazı istatistiki değerler aşağıda özetlenmiştir. $\hat{\alpha} = 1$ olduğundan seri durağan değildir.

Çizelge 3.6 $(X_t - \mu) = \alpha(X_{t-4} - \mu) + e_t$ modeline uygun üretilen veriye ait parametre tahminleri

Conditional Least Squares Estimation					
Parameter	Estimate	Standard Error	t Value	Pr > t	Lag
MU	22.49358	0.57607	39.05	<.0001	0
AR1,1	1.00000	0.03453	28.96	<.0001	4

Karakteristik denklem $m^4 - 1 = 0$ olduğuna göre seri 4 tane birim köke sahiptir. $W_t = (1 - B)^4 X_t$ ve $Z_t = (1 - B^4) X_t$ dönüşümleri altında serinin grafikleri incelendiğinde W_t dönüşümü ile hem serideki hem de otokorelasyonlardaki periyodiklik devam etmektedir. Yani, bu dönüşüm verinin modellenmesi ve istatistiki sonuç çıkarımı açısından yetersiz kalmış ve bir iyileştirme sağlamamıştır. Z_t dönüşümü ile hem serideki hem de otokorelasyonlardaki periyodiklik ortadan kalkmakta ve bu sayede modelleme ve sonuç çıkarımı yapılabilecek duruma gelmektedir. Buradan, $\nabla_4 X_t = e_t$ şeklinde bir model uygun görülmektedir.

Bu örnekte görüldüğü gibi, birim kök sayısı kadar fark almak seriyi durağan hale getirmeyebilir. Mevsimsel zaman serilerinin durağan olup olmadığını sınamak için daha önce belirtilen birim kök testleri anlamlı olmayabilir. Literatürde, mevsimsel birim kök testleri ile ilgili birçok yöntem bulunmasına rağmen, Hylleberg, Engle, Granger ve Yoo (1990) tarafından önerilen ve HEGY testi olarak bilinen en pratik test yöntemidir. Bunun yanında Dickey, Hasza ve Fuller (1984) tarafından önerilen ve simetrik en küçük kareler yöntemi (DHF) de kullanılmaktadır.

Mevsimsel seriler haftalık, aylık, 3 aylık, 6 aylık gibi verilerin toplanma şekline göre değişiklik gösterir. Mevsimsel birim kök testlerinin çoğu 3 aylık veriler üzerine kurulan modeller için geliştirilmiştir. Aşağıda, mevsimsel birim kök testlerinden HEGY testi özetlenmiştir.

3.2.1 HEGY testi

Hylleberg, Engle, Granger ve Yoo (1990) tarafından önerilen ve HEGY testi olarak bilinen bu yöntem mevsimsel birim kök testlerinden en çok kullanılan yöntemlerden biridir. Çeyrek periyoda sahip veriler için $SAR_4(1)$ modeli,

$$(X_t - \mu) = \alpha(X_{t-4} - \mu) + e_t, t = 1, 2, 3, \dots, n$$

şeklinde verilmiş olsun. Modelin karakteristik denklemi $m^4 - \alpha_4 = 0$ olup $\alpha_4 = 1$ için model durağan değildir ve beklenen değer (μ) modelden düşer. Denklemin kökleri $m_{1,2} = \pm 1$ ve $i = \sqrt{-1}$ olmak üzere, $m_{3,4} = \pm i$ olup bütün kökler mutlak değerce 1'dir. $\alpha_4 = 1$ için $\alpha(B) = (1-B^4)$ olmak üzere, model $\alpha(B)X_t = e_t$ olarak yazıldığında,

$$1-B^4 = (1-B)(1+B)(1-iB)(1+iB) = (1-B)(1+B)(1+B^2)$$

eşitliğinden yukarıdaki birim kökler sırası ile , sıfır sıklık, yılda 2 devir ve yılda 1 devir olarak adlandırılır. Buradaki m_3 ve m_4 birim kökleri birbirlerinden ayrılamayıp yıllık devir olarak yorumlanmaktadır. HEGY yönteminde yardımcı bir regresyondan yararlanılır. Bu yardımcı regresyon modeli,

$$Y_{1,t} = (1+B+B^2+B^3)X_t$$

$$Y_{2,t} = -(1-B+B^2-B^3)X_t$$

$$Y_{3,t} = -(1-B^2)X_t$$

$$Y_{4,t} = (1-B^4)X_t = \nabla_4 X_t$$

olmak üzere,

$$\alpha^*(B)Y_{4,t} = \pi_1 Y_{1,t-1} + \pi_2 Y_{2,t-1} + \pi_3 Y_{3,t-2} + \pi_4 Y_{3,t-1} + e_t$$

şeklinde kurulmuştur. Ayrıca modele sabit veya mevsimsel lineer trend eklenebilir. Bu regresyon modeli $\alpha^*(B) = 1$ olduğu göz önüne alınarak $\pi_2 = \pi_3 = \pi_4 = 0$ varsayımı altında $H_0 : \pi_1 = 0$ hipotezinin test edilmesi standart Dickey-Fuller testidir. Bu varsayım altında model

$$Y_{1,t} = (1 + \pi_1)Y_{1,t-1} + e_t$$

şeklinde yazılabildiğinden iki hipotezin aynı olduğu açıktır. Benzer şekilde, $\pi_1 = \pi_3 = \pi_4 = 0$ varsayımı altında model

$$Y_{2,t} = -(1 + \pi_2)Y_{2,t-1} + e_t$$

şeklinde yazılabilir. $H_0 : \pi_2 = 0$ hipotezi karakteristik denklemin kökünün -1 olduğunu göstermektedir.

Çeyreklik mevsimsel serilerin birim kök testi için HEGY yöntemi aşağıda açıklanan şekilde uygulanır.

Adım 1: Aşağıda belirtilen dönüşümler uygulanır.

$$Y_{1,t-1} = (1+B+B^2+B^3)X_{t-1} = X_{t-1} + X_{t-2} + X_{t-3} + X_{t-4}$$

$$Y_{2,t-1} = -(1-B+B^2-B^3)X_{t-1} = -X_{t-1} + X_{t-2} - X_{t-3} + X_{t-4}$$

$$Y_{3,t-1} = -(1-B^2)X_{t-1} = -X_{t-1} + X_{t-3}$$

$$Y_{3,t-2} = -X_{t-2} + X_{t-4}$$

$$Y_{4,t} = (1-B^4)X_t = \nabla_4 X_t$$

Adım 2: Kurulan yardımcı regresyon modeline göre parametreler tahmin edilir.

$$Y_{4,t} = \pi_1 Y_{1,t-1} + \pi_2 Y_{2,t-1} + \pi_3 Y_{3,t-2} + \pi_4 Y_{3,t-1} + e_t$$

Bu regresyon modelinin parametreleri tahmin edilirken, $(1-B^4)X_{t-i}$ değişkenleri de modele ilave edilebilir. Bu aşamada, e_t hata terimlerinin beyaz gürültü serisi olduğu kontrol edilmelidir.

Adım 3: $H_0: \pi_1 = 0$ hipotezini test etmek için t-istatistiğinin hesaplanan değeri kritik değer ile karşılaştırılır. $H_0: \pi_1 = 0$ hipotezi red edilemez ise, seri birim köklüdür. Daha sonra, $H_0: \pi_2 = 0$ hipotezi için t-istatistiğinin değeri hesaplanarak aynı şekilde kritik değer ile karşılaştırılır. Bu yokluk hipotezi red edilemez ise, seri 2 devirli (6 aylık dönemler) birim köke sahiptir. $H_0: \pi_3 = \pi_4 = 0$ hipotezi F istatistiğinin bulunan değeri kritik değer ile karşılaştırılır. F-istatistiğinin değeri, kritik değerden küçük ise $H_0: \pi_3 = \pi_4 = 0$ yokluk hipotezi red edilemez ve seri mevsimsel birim köklü olduğu sonucuna ulaşılır. Bu adımda bahsedilen kritik değerler Hylleberg vd. (1990) çalışmasında mevcuttur.

3.3 Periodogramlar

Zaman serileri analizinde ele alınan seriler zamanın bir fonksiyonu olarak değerlendirilir ve bu serilerin otokorelasyon fonksiyonları serilerin yapısı hakkında bilgi verir. Yani seriler zaman boyutu esas alınarak incelenir ve analiz edilirler. Aynı zaman serileri çeşitli frekans bileşenlerine ayrıştırılarak frekansın bir fonksiyonu şeklinde ifade edilebilmektedir. Bu durumda frekans boyutu öne çıkmaktadır. Frekans boyutlu analizlerde otokorelasyon fonksiyonu yerine spektral yoğunluk fonksiyonu incelenerek serideki periyodik yapı ortaya çıkarılmaya çalışılır. Bu geçişi sağlamak için bazı fonksiyonlar ve matematiksel araçlar geliştirilmiştir. “Fourier Dönüşümü” ve “Spektral Analiz” yöntemleri bunlardan ikisidir. Periodogramlarla ilgili bilgiler dördüncü bölümde detaylı olarak verilmiştir.

4. BULGULAR VE TARTIŞMA

Kocaeli iline ait hava kirliliği indekslerinden biri olan PM_{10} maddesine ait 2005:1 – 2017:12 aralığına ait 156 aylık veri seti kullanılmıştır. Ocak 2007 – Ekim 2011 aylarına ait aylık veriler Türkiye İstatistik Kurumu(TÜİK) internet sitesinden diğer aylık veriler Çevre ve Şehircilik Bakanlığı'nın Hava Kalitesi İzleme İstasyonları Web Sitesinden elde edilmiştir. Bu istasyonlardan alınan değerlerin ortalamaları alınarak veriler elde edilmiştir.

Partiküler maddeler (PM), hava içerisinde asılı olarak bulunan katı ve sıvı parçacıkların karışımından oluşan, yaygın hava kirletici olarak bilinmektedir. Partiküller çapları 10 mikrogram/metreküp veya 2,5 mikrogram/metreküp'ten daha küçük olan partiküller olarak aerodinamik çapları uyarınca tanımlanırlar. Partiküler maddelere hem kısa, hem de uzun dönemde maruz kalmak; solunum, kalp damar hastalıkları, astım atakları, solunumla ilgili rahatsızlıklar ve akciğer kanserinden ölümlerde artışa yol açtığı bilinmektedir. Dünya Sağlık Örgütü, ' PM_{10} ' için yıllık sınır değeri 20 mikrogram/metreküp, 24 saatlik sınır değeri ise 50 mikrogram/metreküp olarak belirmiş iken bu değer ' $PM_{2,5}$ ' için yıllık değer 10 mikrogram/metreküp, 24 saatlik değer ise 25 mikrogram/metreküp'tür.

Sonuçlarının ciddiyeti ve önemi gereği hava kirliliği birçok devletin ve ilgili kuruluşlarının üzerinde çalıştığı bir konudur. Aslında hava kirliliği bir ülke yada kurum sorunu olmaktan ziyade bütün dünyanın bir sorunu olup özellikle son zamanlarda Birleşmiş Milletler ve Dünya Sağlık Örgütü gibi uluslararası kuruluşlar bu sorunun üzerinde araştırmalar yapmaktadır. Elde edilen bulgular düzenli bir şekilde bütün dünya kamuoyu ile paylaşılmaktadır. Dünya Sağlık Örgütü'ne göre günlük PM_{10} konsantrasyonlarında her 10 mikrogram/metreküp artış ölümlerde yüzde 0.2 ile 0.6 artışa yol açarken, uzun dönem PM_{10} konsantrasyonlarında her 10 mikrogram/metreküp artış ise kalp-akciğer hastalıklarına bağlı ölümlerde % 6 ile % 13 arasında artışa sebep olmaktadır.

Tüm dünyada olduğu gibi ülkemizin de en önemli sorunlarından biri hava kirliliğidir. Hava kirliliği çeşitli sağlık sorunlarını da beraberinde getirmektedir. Bununla birlikte artan sağlık harcamaları, ölümler, eldeki iş gücünün etkin kullanılamaması ve üretim kaybı, sağlıksız bir nesil ve kirli bir çevre gibi ciddi problemlere neden olabilmektedir. Bu sebeplerle dünyada ve ülkemizde hava kirliliği ile alakalı birçok araştırma yapılmış ve halen yapılmaktadır. Giriş bölümünde bahsedilen benzer çalışmalara ek olarak yapılan çalışmalardan bazılarını aşağıda yer verilmiştir.

Femke de Zwart ve diğerleri Hollanda'da bulunan yaşlı yetişkinlere yönelik yaptıkları çalışmalarında performans dayalı verileri (yürüme hızı, bir sandalyeden yükselme kabiliyeti, hırka giyme ve soyunma ve denge testi) 1762 katılımcıya uygulayarak elde etmişler ve hava kirliliğine maruz kalan yetişkin insanların fiziksel performansları ciddi anlamda olumsuz olarak etkilendiği sonucuna varmışlardır (Zwart vd. 2018).

Hooper vd. (2018) tarihinde yaptıkları çalışmada uzun süre partiküler madde ($PM_{2.5}$ ve PM_{10}) ve nitrojen dioksit (NO_2) veya her ikisine birden uzun süre maruz kalınmasının kronik bronşit ile ilişkisini araştırmışlar ve güçlü bir ilişki tespit etmişlerdir.

Çatık vd. (2016) yaptıkları çalışmada Türkiye için Çevresel Kuznets Eğrisinin (ÇKE) varlığını, bölgeler arasındaki kirlilik yayılımı dikkate alınarak 1990-2001 yıllarını kapsayan yıllık verilerle incelemişlerdir. Kirliliğin yayılımının test edilmesinde mekansal ekonometrik yöntemlerden yararlanmışlardır. Sürdürülebilir kalkınmanın elde edilmesi için çevresel politikaların bölgesel ve ulusal düzeydeki politikalarla desteklenmesi gerektiği sonucuna ulaşmışlardır.

Demirarslan (2016) sanayi ve teknolojinin gelişmesi ve hızlı kentleşme ile birlikte Doğu Karadeniz Bölgesi sınırları içinde olan Artvin, Bayburt, Giresun, Gümüşhane, Ordu, Rize, Trabzon illerinin merkezlerindeki partikül maddelerden meydana gelen hava kirliliği hakkında araştırma yapılmış ve PM_{10} kirleticisi değerlendirilmiştir. Yapılan bu çalışma ile Doğu Karadeniz Bölgesi içerisinde PM_{10} bakımından hava kalitesinin en yüksek olduğu iller arasında Giresun ve Artvin, en kötü olduğu iller arasında ise Trabzon ve Gümüşhane'nin yer aldığı belirlenmiştir.

Rabia Ece Omay ve Esra Canpolat (2013) birlikte yaptıkları çalışmalarında Türkiye için kişi başına düşen kükürt dioksit (SO_2) ve partiküler madde (PM_{10}) ile kişi başına düşen gayri safi yurtiçi hasıla (GSYH) arasındaki ilişki modellemişlerdir. Türkiye için GSYH ile SO_2 ve PM_{10} arasındaki ilişkinin modellenmesinde, parametrik olmayan modellerin iyi sonuçlar verdiğini gözlemlemişlerdir.

Tomasetti vd. (2013) çalışmalarında birçok alerjinin ve solunum problemlerinin kaynağı olan hava partiküllerinin (polenler ve fungal sporlar) konsantrasyonlarının meteorolojik parametreler yoluyla tahmin etme imkanını araştırmışlardır. Bu kapsamda karmaşık bir yapay sinir ağı mimarisinin kullanımına dayanan Operasyonel Sınırlı Alan Meteorolojik modelinden faydalanılarak bir sinir ağı modeli kullanmışlar ve Orta İtalya’da geniş bir alanda bu partikül konsantrasyonlarının bir haritasını tatmin edici bir doğrulukla çıkarmışlardır.

Onur Günay ve Cavit Işık Yavuz (2009) çalışmalarında hava kirliliği ölçüm parametrelerinden kükürtdioksit (SO_2) ve partiküler madde (PM) düzeyi ile myokard enfarktüsü (MI) nedeni hastane yatışları arasında bir ilişki olup olmadığını araştırmışlardır. Araştırma kapsamında 2004 yılı içerisinde Kocaeli’de koroner yoğun bakım ünitesi olan üç hastanenin kayıtlarından 428 MI yatış verisine ulaşılmıştır. Bu kayıtların 417 bireye ait olduğu saptanmıştır. Ulaşılan 428 MI kaydının uygunluk ölçütleri açısından 318’i ÇVK analizine dahil edilmiştir. Sonuç olarak, Kocaeli’de MI nedeni hastaneye yatışlar ile ölçülmüş SO_2 ve PM_{10} düzeyleri arasındaki ilişki incelenerek, SO_2 düzeylerindeki $10 \mu g/m^3$ bir artışın MI geçirme riskini soğuk sezonda (Ekim-Mart) anlamlı olarak arttırdığı tespit edilmiştir.

Hava kirliliği yaşanabilir bir ortam için en önemli göstergelerden biridir. Hava kirliliği için değişik indeksler olmakla birlikte, bunlar arasında SO_2 ve PM_{10} indeksleri öne çıkanlardır. Bu çalışmada, Türkiye’nin önemli sanayi merkezlerinden olan ve Türkiye’nin nüfus olarak en kalabalık (15 milyon üzeri) şehirlerinden olan İstanbul’a en yakın şehir olan Kocaeli için aylık olarak ölçülen PM_{10} değerleri 2005:01-2017:12 dönemi dikkate alınarak analizleri yapılmıştır.

Hava kirliliğini mevsimsel olarak değişebileceği gibi aylara göre de değişiklik göstermesi beklenir. Ayrıca, son zamanlarda özellikle doğalgaz kullanımının yaygınlaşması nedeni ile hava kirliliğinde bir azalmanın olması beklenmektedir. Diğer taraftan, doğalgaz kullanımının yanında nüfus ve sanayileşmedeki artış dikkate alındığında yıllara göre değişim olmayabilir. Ancak, aylar arasında bir farklılığın olması beklenmektedir. Aylar arasında bir farklılığın olup olmadığını görebilmek için verilere

$$Y_{ij} = \mu + \alpha_i + e_{ij}, \quad i = 1, 2, \dots, 12; j = 1, 2, \dots, 13 \quad (4.1)$$

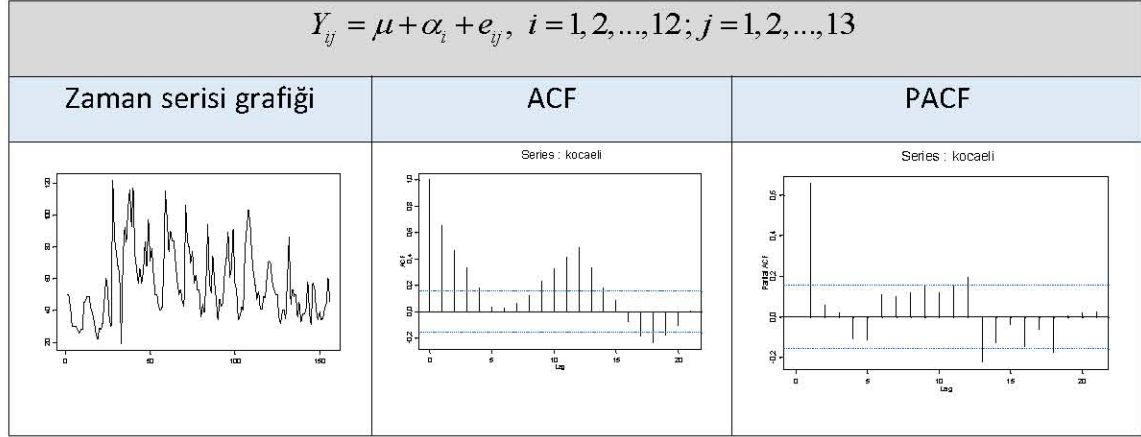
şeklinde bir modelin uygun olduğunu varsayalım. Bu modele ilişkin SAS da PROG GLM çıktıları aşağıda çizelge 4.1'de verilmiştir.

Çizelge 4.1 Aylık ortalama PM₁₀ değerlerine ilişkin Eşitlik (1.1) de verilen model sonuçları

Dependent Variable: pm10					
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	11	22423.84615	2038.53147	5.71	<.0001
Error	144	51412.92308	357.03419		
Corrected Total	155	73836.76923			

mon	N	Mean	Std Dev
1	13	66.7692308 (2)	22.5468174
2	13	64.6153846	20.9783160
3	13	63.9230769	20.4835769
4	13	63.2307692	28.1103594
5	13	51.6923077	17.1385755
6	13	46.0769231	14.8293714
7	13	42.0000000	12.7540843
8	13	42.3846154	11.6586053
9	13	39.8461538	9.9485858
10	13	50.3846154	16.5052439
11	13	73.1538462 (2)	22.3340031
12	13	74.3846154 (1)	21.0655266

Yukarıda çizelge 4.1’ de verilen sonuçlar incelendiğinde, PM₁₀ değerleri aylara göre değişiklik gösterdiği söylenebilir (hesaplanan F istatistiğine ilişkin p-değeri<0.0001 dir).



Şekil 4.1 Kocaeli iline ait aylık PM₁₀ değerleri (Ocak 2005-Aralık 2017)

Verilere uygun bir model belirleyebilmek için şekil 4.1’de verilen zaman serisi grafikleri de dikkate alınarak değişik zaman serisi modelleri için AIC değerleri hesaplanmıştır. Bunlar arasından en küçük AIC değerini veren model verilere en uygun modeldir.

Çizelge 4.2 Değişik zaman serisi modellerine ilişkin AIC istatistiklerinin değerleri

P	1	2	3	4	12	13	(4)	(12)	(1,12)	(1,12,13)
AIC	1319	1321	1323	1323	1312	1307	1402	1361	1302	1298

Yukarıda verilen AIC değerleri göz önüne alındığında, verilere en uygun modelin

$$Y_t = \alpha_0 + \alpha_1 Y_{t-1} + \alpha_{12} Y_{t-12} + \alpha_{13} Y_{t-13} + e_t, t = 1, 2, 3, \dots, n \quad (4.2)$$

ARIMA modelinin olduğu görülmektedir (en küçük AIC değeri dikkate alınmıştır).

Daha önce belirtildiği gibi zaman serilerinde en önemli varsayımlardan biri serinin durağanlığıdır. Eğer seri durağan ise öngörülebilirdir. Bilindiği gibi, MA modelleri her zaman durağan olup, AR modelleri ya da AR bileşeni içeren (ARMA, SAR ya da SARMA türü) modeller durağan olmayabilir. Öizelge 4.2’deki sonuçlara göre, verilere

en uygun modelin (4.2) de verilen model olduğu sonucu elde edilmişti. Herhangi bir $AR(p)$ zaman serisi modeli,

$$Y_t = \alpha_0 + \alpha_1 Y_{t-1} + \alpha_{12} Y_{t-12} + \alpha_{13} Y_{t-13} + e_t, t = 1, 2, 3, \dots, n \quad (4.3)$$

şeklinde verildiğinde, bu modele karşılık gelen karakteristik denklem

$$m^p - \sum_{i=1}^p \alpha_i m^{p-i} = 0 \quad (4.4)$$

olup, bu denklemin bütün kökleri mutlak değerce 1' den küçük ise model durağandır. Eğer köklerden en az biri mutlak değerce 1 ise seri birim köklü yani durağan değildir (köklerin 1' den büyük olması uygulama açısından anlamlı değildir). Kolayca görüleceği

gibi, köklerden en az birinin mutlak değerce 1 olması için gerek ve yeter koşul $\sum_{i=1}^p \alpha_i = 1$

olmasıdır. Veri setine en uygun modelin (4.2) eşitliğinde verilen model olduğuna göre, parametre tahminleri SAS da PROC ARIMA ve PROG REG ile hesaplanarak aşağıda Çizelge 4.3a ve 4.3b' de verilmiştir. Bu değerler PROC ARIMA da , $\hat{\alpha}_1 + \hat{\alpha}_{12} + \hat{\alpha}_{13} = 0.79181$ ve PROG REG de, $\hat{\alpha}_1 + \hat{\alpha}_{12} + \hat{\alpha}_{13} = 0.76001$ olarak gözlenmiştir. Buna göre, serinin durağan olduğu söylenebilir. Serinin durağanlığını test etmek için literatürde oldukça çok yöntem bulunmakla birlikte, bunlar arasında parametrelerin en küçük kareler tahmin edicisinin dağılımına dayalı Dickey-Fuller yöntemi ile bu yöntemin biraz düzenlenmiş hali olarak bilinen Phillips-Perron yöntemi öne çıkmaktadır. Bu iki yöntemle göre maksimum gecikme uzunluğu 13 olduğu varsayımına göre birim kök testleri yapılmış elde edilen sonuçlar çizelge 4.3' de verilmiştir. Her iki yöntemle göre de aylık PM₁₀ verileri durağandır.

Çizelge 4.3 ADF ve PP Birim Kök Test Sonuçları

	t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic	-5.622680	0.0000
Phillips-Perron test statistic	-5.717368	0.0000
Test critical values: 1% level	-3.472813	
5% level	-2.880088	
10% level	-2.576739	

*MacKinnon (1996) one-sided p-values.

Bununla birlikte veride periyodik bir bileşenin varlığından şüpheleniliyorsa, periodogramlara dayalı birim kök testinden de faydalanılabilir. Periodogramlar genellikle veride muhtemel bir periyodikliğin araştırılmasında kullanıldığı gibi, serinin durağanlığının test edilmesinde de kullanılmaktadır. Bir serinin durağanlığının sınanmasına ilişkin oldukça çok metot vardır. Bu kısımda, periodogramlara dayalı birim kök test yöntemi kısaca açıklandıktan sonra serinin durağanlığı test edilecektir.

Zaman serilerinde istatistiki sonuç çıkarımlar için serinin durağanlığı en önemli varsayımdır. Ayrıca zaman serilerinde en önemli amaç, serinin gözlenen değerlerini kullanarak gelecek zamanlarda serinin alacağı değerlerin öngörülmesidir. Serinin öngörülebilir olması için de durağanlık en önemli varsayımdır. Daha önce belirtildiği gibi durağanlığın testi başta Dickey-Fuller ve Phillips-Perron testleri olmak üzere değişik yöntemlerle yapılmaktadır. Zaman serilerinde MA türü seriler her zaman durağan olup, AR veya AR bileşeni içeren modeller durağan olmayabilir. İktisadi serilerin çoğu durağan olmayan zaman serileridir. Bununla birlikte, mevsimsel serilerin durağanlığının sınanmasında ise HEGY türü testler kullanılır. Eğer herhangi bir periyodiklikten şüphelenilmiyorsa, serinin durağanlığı için yukarıda bahsedilen testlerden biri kullanılabilir.

Periodogramlar genellikle seride bulunan gizli periyodikliklerin ortaya çıkarılmasında kullanılır (Fuller 1976, Wei 2006, Brockwell ve Davis 1987). Bununla birlikte Akdi ve Dickey (1998) periodogramlara dayalı bir test önermişlerdir. Periodogramlar ile ilgili genel açıklamalar aşağıdadır.

Periyodik fonksiyonlar denildiğinde ilk trigonometrik fonksiyonlar aklımıza gelir. O halde, serinin periyodik bir bileşen içerip içermediğini sınamak için $\{Y_1, Y_2, \dots, Y_n\}$ gibi verilen herhangi bir zaman serisini

$$Y_t = \mu + R \cos(\omega t + \phi) + e_t, \quad t = 1, 2, \dots, n \quad (4.4)$$

şeklinde yazalım. Buradaki μ , R , ϕ ve ω terimleri sırası ile beklenen değer, genişlik (genlik/amplitude), faz (phase) ve sıklıkları (frequency) ifade etmektedir. Bu parametrelerin tahmin edilmesi gerekir. Eğer $\omega_k = 2\pi k / n$ olarak seçildiğinde ω_k

' ler Fourier sıklıkları olarak bilinir. Kosinüs fonksiyonunun özelliklerinden $\alpha = R \cos(\phi)$ ve $\beta = R \sin(\phi)$ olmak üzere bu model

$$Y_t = \mu + \alpha \cos(w_k t) + \beta \sin(w_k t) + e_t, t = 1, 2, \dots, n \quad (4.5)$$

şeklinde yazılabilir. Bu modele göre, $H_0 : \alpha = \beta = 0$ yokluk hipotezi ret edilirse veride periyodik bileşen vardır. Bu hipotezin testi için standart F testi kullanılabilir. Ancak w_k sıklıkları bilinmediği durumda F istatistiğinin kullanımı anlamlı değildir (Wei 2006). Bu modele göre μ , α ve β parametrelerinin en küçük kareler tahmin edicileri sırası ile,

$$\hat{\mu} = \bar{Y}_n, \quad a_k = \frac{2}{n} \sum_{t=1}^n (Y_t - \bar{Y}_n) \cos(w_k t) \quad \text{ve} \quad b_k = \frac{2}{n} \sum_{t=1}^n (Y_t - \bar{Y}_n) \sin(w_k t)$$

dir. Buradaki hesaplanan a_k ve b_k değerlerine Fourier katsayıları denir. Yine trigonometrik fonksiyonların özelliklerinden

$$\sum_{t=1}^n \cos(w_k t) = \sum_{t=1}^n \sin(w_k t) = 0$$

olduğundan Fourier sıklıkları ortalamaya göre değişmezdir (invariant). Bu Fourier katsayıları yardımı ile zaman serisinin w_k sıklığındaki Periodogram (periodogram ordinate),

$$I_n(w_k) = \frac{n}{2} (a_k^2 + b_k^2)$$

şeklinde hesaplanır.

Zaman serileri genellikle iki grupta incelenir. Biri zaman boyutu (time-domain), diğer sıklık boyutu (frequency domain). Zaman boyutunda en önemli araç serinin otokorelasyon fonksiyonudur. Sıklık boyutunda ise spectral yoğunluk fonksiyonudur. Ayrıca, serinin otokorelasyon fonksiyonu ile spectral yoğunluk fonksiyonu arasında da bir geçiş vardır (Herlogtz Teoremi, Brockwell ve Davis 1987). Eğer $f(w_k)$ durağan zaman serisinin spectral yoğunluk fonksiyonu ise $I_n(w_k) / f(w_k)$ istatistiğinin asimptotik dağılımı serbestlik derecesi 2 olan ki-kare olup, asimptotik dağılımın beklenen değeri 1 olan üsteldir. Yani, normalleştirilmiş periodogramların asimptotik dağılımının olasılık yoğunluk fonksiyonu

$$f(x) = \begin{cases} e^{-x} & , \quad x > 0 \\ 0 & , \quad d.y. \end{cases}$$

dir. Dolayısı ile, periodogramlar spectral yoğunluk fonksiyonunun bir tahmin edicisi olarak alınabilir.

Periodogramlar veride muhtemel periyodikliklerin araştırılmasında (yani yukarıdaki H_0 hipotezinin test edilmesinde) da kullanılmaktadır. Durağan herhangi bir zaman serisi için her bir k sıklığındaki $I_n(w_k)$ periodogram değerleri hesaplanır. $I_n(w_{(1)})$ en büyük periodogram değerini m de $n/2$ nin tam kısmı ($m = [n/2]$) olmak üzere,

$$V = I_n(w_{(1)}) \left[\sum_{k=1}^m I_n(w_k) \right]^{-1}$$

istatistiğini tanımlayalım. Herhangi bir periyodik bileşen yoksa ($H_0 : \alpha = \beta = 0$ altında) V istatistiği için

$$P(V > c_\alpha) = \alpha \cong m(1 - c_\alpha)^{m-1}$$

dir (Wei 2006, page 294-295). Buradan, seçilen herhangi bir α anlam düzeyi için c_α kritik değeri

$$c_\alpha = 1 - (\alpha / m)^{1/(m-1)}$$

olarak hesaplanır. Eğer $V > c_\alpha$ ise $H_0 : \alpha = \beta = 0$ hipotezi ret edilir ve seride periyodik bileşen olduğu sonucuna varılır (Fisher testi).

Yukarıda da bahsedildiği gibi, verilen zaman serisi durağan ise, normalleştirilmiş periodogramın asimptotik dağılımı serbestlik derecesi 2 olan ki-karedir. $f(w_k)$ durağan zaman serisinin spectral yoğunluk fonksiyonu olmak üzere,

$$I_n(w_k) / f(w_k) \xrightarrow{D} \chi_2^2, \quad n \rightarrow \infty$$

dir (Fuller 1976, Brockwell ve Davis 1987, Wei 2006). Serinin durağan olmadığı (birim köklü) varsayımı altında her bir sabit w_k için,

$$T_n(w_k) = \frac{2(1 - \cos(w_k))}{\hat{\sigma}_n^2} I_n(w_k) \xrightarrow{D} Z_1^2 + 3Z_2^2, \quad n \rightarrow \infty$$

dir (Akdi 1995, Akdi ve Dickey 1998). Burada, Z_1 ve Z_2 bağımsız standart normal dağılımlı rasgele değişkenler olup $\hat{\sigma}_n^2$ hata terimin varyansının bir tahmin edicisidir. Kısaca asimptotik dağılım her sabit k için

$$T_n(w_k) \xrightarrow{D} \chi_1^2 + 3\chi_1^2, n \rightarrow \infty$$

dir. Yine Akdi ve Dickey (1999) yöntemin mevsimsel zaman serileri için de uygulanabilir olduğunu göstermişlerdir. Yani, $T_n(w_k)$ istatistiği serinin durağanlığını (birim köklü) sınamak için kullanılabilir. Asimptotik dağılım her bir sabit w_k için geçerli olmakla birlikte hipotez testlerinde genellikle w_1 sıklığı kullanılmaktadır. Dağılımın kritik değerleri ilgili çalışmada verilmiştir. Seride özellikle periyodik bir bileşen varsa, bu yöntemin diğer yöntemlere göre avantajlı olduğu söylenebilir. Bununla birlikte,

i) Yöntemde herhangi bir model varsayımı bulunmamaktadır. Yani, metod modele göre değişmezdir (invariant). Yani Periodogramlar herhangi bir model varsayımına bağlı olmadan hesaplanabilir.

ii) Bilindiği gibi, diğer birim kök testleri zaman serisi modeline bağlı olduğundan, öncelikle model parametrelerinin tahminine ihtiyaç duyulur. Burada ise, beyaz gürültü serisinin varyansı haricinde herhangi bir parametre tahminine ihtiyaç duyulmaz.

iii) $T_n(w_k)$ istatistiğinin dağılımı H_0 yokluk hipotezi ve H_a alternatif hipotez altında belli olduğundan, test için analitik güç fonksiyonu mevcuttur.

iv) Test istatistiğinin kritik değerleri örneklem hacmine bağlı değildir.

v) Periodogramlar verinin trigonometrik dönüşümleri ile hesaplandığı için eğer veride herhangi bir periyodik bileşen varsa, bu periyodikliği dikkate aldığından dolayı daha anlamlı olacağı açıktır.

Bu yöntemin yukarıda bahsedilen özellikleri dikkate alındığında, özellikle veride herhangi bir periyodik bileşenin bulunması durumunda diğer yöntemlere göre daha avantajlı olduğu düşünülebilir.

Hava kirliliği günümüzün en önemli sağlık, çevre ve sosyal sorunlarından biridir. Hava kirliliğinin gelişen teknoloji ve artan nüfus ile birlikte giderek arttığı da bir gerçektir. Bu

çalışmada, 2005:01-2017:12 dönemine ait Kocaeli ili için ölçülen PM_{10} verilerinin durağanlığı incelendikten sonra eğer seri durağan ise muhtemel periyotlar elde edilmeye çalışılacaktır. Verilerden, hata terimlerinin varyansı (Eşitlik (4.1) de verilen model) için tahmin değeri Çizelge 4.3'de verildiği gibi $\hat{\sigma}_n^2 = 234.6583$ ve SAS/PROC SPECTRA'da $I_n(w_1) = 15373.31$ olarak elde edilmiştir. Buna göre, test istatistiğinin değeri de $T_n(w_1) = 0.10626$ dir. Test istatistiğinin kritik değeri $\alpha = 0.05$ için $c_\alpha = 0.178$ ve $\alpha = 0.01$ için $c_\alpha = 0.0348$ dir (Akdi ve Dickey, 1998). Buna göre

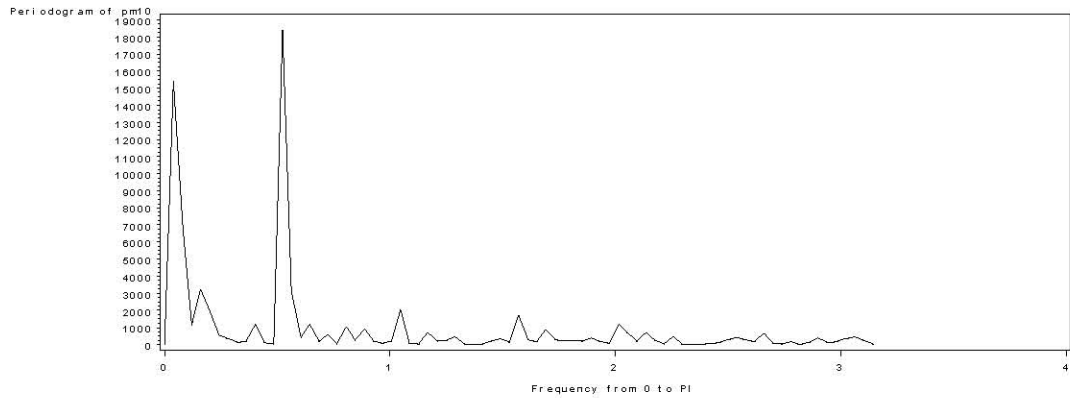
$$T_n(w_1) = 0.10626 < 0.178 = c_{0.05}$$

olduğundan seri $\alpha = 0.05$ anlam düzeyinde durağan ve

$$T_n(w_1) = 0.10626 > 0.0348 = c_{0.01}$$

olduğundan $\alpha = 0.01$ anlam düzeyinde durağan değildir. Dolayısı ile Eşitlik (4.2)'de verilen model $\alpha = 0.05$ anlam düzeyinde durağandır.

Şimdi serinin periyodik bileşen içerip içermediğini sınamak isteyelim. Eğer seri periyodik bileşen içeriyorsa, bu periyodları tespit edelim. Bunun için de serinin periodogram değerlerini kullanalım.



Şekil 4.2 Periodogamların sıklıklara karşı grafiği (Kocaeli, 2005:01-2017:12)

Çizelge 4.4 Eşitlik (4.1) modeline ait parametre tahminleri PROC ARIMA

Conditional Least Squares Estimation					
Parameter	Estimate	Standard		Approx	
		Error	t Value	Pr > t	Lag
MU	53.04557	5.20131	10.20	<.0001	0
AR1,1	0.59666	0.06616	9.02	<.0001	1
AR1,2	0.40159	0.07674	5.23	<.0001	12
AR1,3	-0.20644	0.08092	-2.55	0.0117	13

Constant Estimate			11.04392		
Variance Estimate			234.6583		
Std Error Estimate			15.31856		
AIC			1298.125		
SBC			1310.324		
Number of Residuals			156		

Çizelge 4.5 Eşitlik (4.1) modeline ait parametre tahminleri PROC REG

Parameter Estimates					
Variable	DF	Parameter		t Value	Pr > t
		Estimate	Error		
Intercept	1	14.11374	4.39815	3.21	0.0017
pm101	1	0.55597	0.07152	7.77	<.0001
pm1012	1	0.40140	0.07878	5.10	<.0001
pm1013	1	-0.19736	0.08328	-2.37	0.0192

Çizelge 4.4 – 4.5’ teki değerler incelendiğinde model parametreleri anlamlıdır (p-değerleri).

Bu modele göre hesaplanan bir yıllık (2018 yılı Ocak-Aralık) öngörü değerleri çizelge 4.6’ da verilmiştir.

Çizelge 4.6 Bir yıllık öngörü değerleri (2018:01-2018:12)

Forecasts for variable pm10				
Obs	Forecast	Std Error	95% Confidence Limits	
157	42.8882	15.3186	12.8643	72.9120
158	53.1243	17.8381	18.1623	88.0863
159	52.6594	18.6531	16.1001	89.2188
160	47.7808	18.9347	10.6694	84.8922
161	44.7426	19.0340	7.4366	82.0486
162	46.3828	19.0692	9.0078	83.7578
163	44.3213	19.0818	6.9218	81.7209
164	44.5251	19.0862	7.1169	81.9334
165	47.2514	19.0878	9.8400	84.6628
166	48.6377	19.0884	11.2252	86.0502
167	54.8693	19.0886	17.4564	92.2822
168	49.8686	19.0886	12.4556	87.2817

En büyük 5 periodogram değeri ve bunlara karşılık gelen sıklık ve periyotlar ile V istatistiğinin değerleri aşağıdaki çizelgede verilmiştir.

Çizelge 4.7 Eşitlik (4.1) modeline ait en büyük periodogram değeri,sıklıklar, periyotlar ve V istatistik değerleri

Sıra	$I_n(w_k)$	Sıklık (w_k)	Periyod	V
$I_n(w_{(1)})$	18378.85	0.52360	12	0.24886
$I_n(w_{(2)})$	15373.31	0.04028	156	0.27713
$I_n(w_{(3)})$	6808.18	0.08055	78	0.16978
$I_n(w_{(4)})$	3209.40	0.16111	39	0.09642
$I_n(w_{(5)})$	3116.22	0.56388	11.143	0.10359

Diğer taraftan,

$$T = \sum_{k=1}^m I_n(w_k) = 73851.5385 \text{ olup}$$

$$V_1 = \frac{I_n(w_{(1)})}{T} \cong 0.24886$$

$$V_2 = \frac{I_n(w_{(2)})}{T - I_n(w_{(1)})} \cong 0.27713$$

$$V_3 = \frac{I_n(w_{(3)})}{T - I_n(w_{(1)}) - I_n(w_{(2)})} \cong 0.16978$$

$$V_4 = \frac{I_n(w_{(4)})}{T - I_n(w_{(1)}) - I_n(w_{(2)}) - I_n(w_{(3)})} \cong 0.09642$$

$$V_5 = \frac{I_n(w_{(5)})}{T - I_n(w_{(1)}) - I_n(w_{(2)}) - I_n(w_{(3)}) - I_n(w_{(4)})} \cong 0.10359$$

olarak bulunur.

Kritik değerler $c_\alpha = 1 - (\alpha / m)^{1/(m-1)}$ formülünden

$$c_{0.01} = 0.10987, \quad c_{0.05} = 0.09107 \text{ ve } c_{0.10} = 0.08285$$

olarak hesaplanmıştır. Burada, $n = 156$ örneklem hacmi olup $m = 78$ dir.

Veri setimize zaman serileri açısından yaklaştığımızda elde ettiğimiz ARIMA modeli ve bu modele ait sonuçlar yukarıda tartışılmıştı. Aynı veri daha önce bahsedilen trigonometrik fonksiyon olarak ele aldığında iki farklı modele ulaşılır. Birincisi

$$\begin{aligned} Y_t = & \mu + A_1 \cos\left(\frac{2\pi t}{12}\right) + B_1 \sin\left(\frac{2\pi t}{12}\right) + A_2 \cos\left(\frac{2\pi t}{156}\right) + B_2 \sin\left(\frac{2\pi t}{156}\right) \\ & + A_3 \cos\left(\frac{2\pi t}{78}\right) + B_3 \sin\left(\frac{2\pi t}{78}\right) + e_t, \quad t = 1, 2, \dots, 300 \end{aligned} \quad (4.6)$$

şeklinde olacaktır. Bu modele ait parametre tahminleri aşağıda çizelge 4.8'de gösterilmiştir.

Çizelge 4.8 Eşitlik (4.6) modeline ait parametre tahminleri

Parameter Estimates					
		Parameter	Standard		
Variable	DF	Estimate	Error	t Value	Pr > t
Intercept	1	56.53846	1.19650	47.25	<.0001
c1	1	12.17284	1.69211	7.19	<.0001
s1	1	9.35138	1.69211	5.53	<.0001
c2	1	-12.65189	1.69211	-7.48	<.0001
s2	1	6.08468	1.69211	3.60	0.0004
c3	1	-8.89416	1.69211	-5.26	<.0001
s3	1	-2.85977	1.69211	-1.69	0.0931

Çizelge 4.8'de görüleceği üzere s3 olarak tanımlanmış B_3 katsayısı hariç diğer katsayılar istatistiksel olarak anlamlı iken B_3 için sınırda bir anlamlılıktan bahsedilebilir. Bu yüzden modelden çıkarılmamıştır. Sonuç olarak;

$$Y_t = 56.53 + 12.17 \cos\left(\frac{2\pi t}{12}\right) + 9.35 \sin\left(\frac{2\pi t}{12}\right) - 12.65 \cos\left(\frac{2\pi t}{156}\right) + 6.08 \sin\left(\frac{2\pi t}{156}\right) - 8.89 \cos\left(\frac{2\pi t}{78}\right) - 2.85 \sin\left(\frac{2\pi t}{78}\right)$$

şeklindeki ilk modele (Model 1) ulaşılmıştır.

Verideki gizli periyodiklikler dikkate alındığında Model 1'deki yaklaşımla aşağıdaki trigonometrik modele (Model 2) ulaşılır.

$$Y_t = \mu + A_1 \cos\left(\frac{2\pi t}{12}\right) + B_1 \sin\left(\frac{2\pi t}{12}\right) + A_2 \cos\left(\frac{2\pi t}{156}\right) + B_2 \sin\left(\frac{2\pi t}{156}\right) + A_3 \cos\left(\frac{2\pi t}{78}\right) + B_3 \sin\left(\frac{2\pi t}{78}\right) + A_4 \cos\left(\frac{2\pi t}{39}\right) + B_4 \sin\left(\frac{2\pi t}{39}\right) + A_5 \cos\left(\frac{2\pi t}{11.143}\right) + B_5 \sin\left(\frac{2\pi t}{11.143}\right) + e_t, t=1,2,\dots,300 \quad (4.7)$$

Bu modele ait parametre tahminleri çizelge 4.9'de verilmiştir.

Çizelge 4.9 Eşitlik (4.7) modeline ait parametre tahminleri

Parameter Estimates					
		Parameter	Standard		
Variable	DF	Estimate	Error	t Value	Pr > t
Intercept	1	56.53838	1.09154	51.80	<.0001
c1	1	12.17185	1.54367	7.89	<.0001
s1	1	9.35075	1.54367	6.06	<.0001
c2	1	-12.65205	1.54367	-8.20	<.0001
s2	1	6.08467	1.54367	3.94	0.0001
c3	1	-8.89432	1.54367	-5.76	<.0001
s3	1	-2.85979	1.54367	-1.85	0.0660
c4	1	3.87639	1.54367	2.51	0.0131
s4	1	-5.11065	1.54367	-3.31	0.0012
c5	1	-5.18554	1.54368	-3.36	0.0010
s5	1	-3.61427	1.54366	-2.34	0.0206

Çizelge 4.9'de görüleceği üzere s3 olarak tanımlanmış B_3 katsayısı hariç diğer katsayılar istatistiksel olarak anlamlı iken B_3 için sınırda bir anlamlılıktan bahsedilebilir. Bu yüzden modelden çıkarılmamıştır. Sonuç olarak;

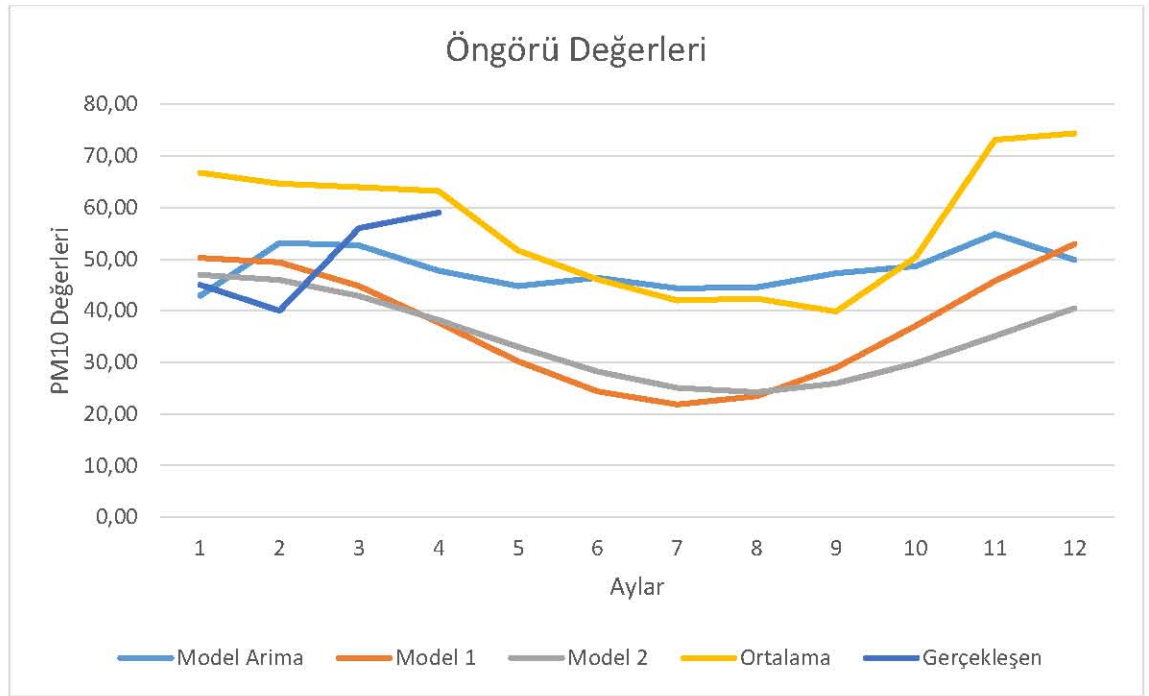
$$Y_t = 56.53 + 12.17 \cos\left(\frac{2\pi t}{12}\right) + 9.35 \sin\left(\frac{2\pi t}{12}\right) - 12.65 \cos\left(\frac{2\pi t}{156}\right) + 6.08 \sin\left(\frac{2\pi t}{156}\right) - 8.89 \cos\left(\frac{2\pi t}{78}\right) - 2.85 \sin\left(\frac{2\pi t}{78}\right) + 3.87 \cos\left(\frac{2\pi t}{39}\right) - 5.11 \sin\left(\frac{2\pi t}{39}\right) - 5.18 \cos\left(\frac{2\pi t}{11.143}\right) - 3.61 \sin\left(\frac{2\pi t}{11.143}\right) + e_t, t = 1, 2, \dots, 300$$

şeklindeki ikinci modele (Model 2) ulaşılır.

Yukarıda bahsedilen modellere ait öngörü değerleri ve gerçekleşen değerler Çizelge 4.10 ve Şekil 4.3' te özetlenmiştir.

Çizelge 4.10 Öngörü değerleri ve gerçekleşen değerler

Model Arima	Model 1	Model 2	Ortalama	Gerçekleşen
42,89	50,26	46,95	66,76	45
53,12	49,36	45,93	64,61	40
52,66	44,74	42,84	63,92	56
47,78	37,70	38,20	63,23	59
44,74	30,19	32,96	51,69	
46,38	24,33	28,23	46,07	
44,32	21,82	25,06	42,00	
44,53	23,45	24,20	42,38	
47,25	28,96	25,90	39,84	
48,64	37,05	29,81	50,38	
54,87	45,74	35,08	73,15	
49,87	52,93	40,50	74,38	



Şekil 4.3 Öngörü değerleri ve gerçekleşen değerler

Çizelge 4.10 ve şekil 4.3 model öngörülleri ile gerçekleşen değerler arasında uyumu görsel olarak sunmakla birlikte modeller arasında seçim yapmak için her bir modelin Hata

Kareler Ortalama (MSE) değerlerini elimizde bulunan 2018 yılının ilk 4 aylık değerlerine göre bulalım. Öncelikle,

Y_i : Gerçekleşen değer

$\hat{Y}_{i,a}$: ARIMA öngörülleri

$\hat{Y}_{i,m1}$: Model 1 öngörülleri

$\hat{Y}_{i,m2}$: Model 2 öngörülleri

olsun. Her bir modele ait MSE değerleri aşağıdadır.

$$MSE(ARIMA) = \sum_{i=1}^3 (Y_i - \hat{Y}_{i,a})^2 = (45 - 42.88)^2 + (40 - 53.12)^2 + (56 - 52.65)^2 + (59 - 47.87)^2 = 311.72$$

$$MSE(\text{Model 1}) = \sum_{i=1}^3 (Y_i - \hat{Y}_{i,m1})^2 = (45 - 50.26)^2 + (40 - 49.36)^2 + (56 - 44.74)^2 + (59 - 37.70)^2 = 695.75$$

$$MSE(\text{Model 2}) = \sum_{i=1}^3 (Y_i - \hat{Y}_{i,m2})^2 = (45 - 46.95)^2 + (40 - 45.93)^2 + (56 - 42.83)^2 + (59 - 38.20)^2 = 645.05$$

Buradan da görüldüğü gibi ilk 4 aylık verilere göre hesaplanan MSE değerlerine bakıldığında en düşük MSE değerine sahip olan ARIMA modelinin en iyi model olduğu söylenebilir.

5. SONUÇ VE ÖNERİLER

Hava kirliliği tehlikeli bir sorundur ve giderek artan nüfus, kentleşme, artan sanayileşme ve karbon içeren yakıtların artan kullanımı nedeniyle artmaktadır. İnsan kaynaklı nedenlerin önlenmesi veya en azından belli ölçüde azaltılabileceği kabul edilmektedir. Hava kirliliğinin ana nedenlerinden birinin organize sanayi bölgelerinin varlığı olduğu açıktır. Böylece hava kirliliğinin öngörülmesi önlem almakla görevli ilgili tüm yetkililer için önemli bir konudur.

Bu çalışmada, hava kirliliğinin ana bileşenlerinden PM₁₀ değerleri zaman serisi boyutunda incelenmiştir. Verilerin durağanlığı Augmented Dickey Fuller birim kök testi ile test edilmiş ve durağanlık varsayımının sağlandığı görülmüştür. Bu veriyi modellemek için Box-Jenkins modeli (AR, p = (1,12,13)) ve iki trigonometrik regresyon modeli düşünülmüş ve karşılaştırılmıştır. 2005:01 - 2017:12 verileri kullanılarak elde edilen üç farklı model ile 2018 yılı için aylık öngörü değerleri elde edilmiş ve bu değerler gerçek gözlem değerleri ile karşılaştırılmıştır. Aylık PM₁₀ gözlenen değerlerin son yıllarda özellikle yaz aylarında azalma eğilimi gösterdiği ve bu azalmanın trigonometrik modellerce daha iyi tespit edildiği görülmüştür. Buna rağmen modellerin hata kareler ortalamasına (MSE) baktığımızda ARIMA modelinin istatistiksel olarak daha iyi sonuçlar verdiği söylenebilir. Bu nedenle, bir sanayi ilimiz olan Kocaeli için PM₁₀ 'un gelecek değerlerini öngörmek için

$$Y_t = \alpha_0 + \alpha_1 Y_{t-1} + \alpha_{12} Y_{t-12} + \alpha_{13} Y_{t-13} + e_t, \quad t = 1, 2, 3, \dots, n$$

modeli tercih edilebilir.

Dünya Sağlık Örgütü (WHO) raporlarına göre yıllık ortalama PM₁₀ limit değeri 20 µg / m³ olmalı, ancak bahsedilen modellerle elde edilen öngörü değerleri bu limit değerlerin çok üzerindedir. Doğal gazın ülke genelinde yaygın olarak kullanılmasına rağmen, PM₁₀ değerleri hala yüksektir. Özellikle Kocaeli ilinde aylık ortalamalar Avrupa Birliği tarafından belirlenen standartların çok üzerindedir. Bu değerlerin standartların çok üzerinde olması sağlıktan ekonomiye bir çok alanda hem o ile hem de ülkemize külfet getirdiği aşıkardır. Gerekli tedbirlerin alınmaması durumunda bu değerlerin artmaya

devam edeceđi beklenmektedir. Bu noktada ilgili birimlerin konuya dikkati çekilmeli ve gereken tedbirlerin alınması hususunda kendileri bilgilendirilmelidir.

KAYNAKLAR

- Akdi, Y. 1995. Periodogram Analysis for Unit Roots. Ph.D. Thesis. North Carolina State University.
- Akdi, Y. 2012. Zaman Serileri Analizi. Ankara Üniversitesi. Gazi kitabevi, 377 s., Ankara.
- Akdi, Y. and Dickey, D.A. 1998. Periodograms of unit root time series: distributions and tests. *Communications in Statistics Theory and Methods*, 27:1, 69-87.
- Akdi, Y. and Dickey, D. A. 1999. "Periodograms for Seasonal Time Series With a Unit Root", *İstatistik, Journal of the Turkish Statistical Association*, 2, 3, 153-164.
- Berktaş, B.M. and Bircan, A. 2003. Effects of atmospheric sulphur dioxide and particulate matter concentrations on emergency room admissions due to asthma in Ankara. *Journal of Tuberculosis and Thorax* , 51(3):231-238.
- Brockle, J.C. and Dickey, D.A. 1986. SAS System for Forecasting Time Series. SAS Institute, Cary, NC.
- Brockwell P.J. and Davis, R.A. 1987. Time Series: Theory and Methods. Springer-Verlag.
- Çatık, A.N., Karaçuka, M. and Huyugüzel-Kışla, G. 2016. Air pollution and income relationship in Turkish provinces: A spatial approach. *Marmara Üniversitesi Açık Arşiv Sistemi*.
- Cicek, I., Turkoglu, N. ve Gorgen, G. 2004. Ankara'da hava kirliliğinin istatistiksel analizi. *Firat University Journal of Social Science* Volume:14, Issue:2, 1-18.
- Demirarslan, K.O. ve Akıncı, H. 2016. Doğu Karadeniz Bölgesindeki Partikül Madde Dağılımlarının Coğrafi Bilgi Sistemleri Yardımıyla Belirlenmesi. *Doğal Afetler ve Çevre Dergisi* Cilt 2 Sayı 1.
- Dickey, D. A., Hazsa, D. P. and Fuller, W. A. 1984. Testing for Unit Roots in Seasonal Time Series. *Journal of American Statistical Association*.
- Fuller, W.A. 1976. Introduction to Statistical Time Series, John Wiley & Sons.
- Genc, D.D., Yesilyurt, C. and Tuncel, G. 2010. Air pollution forecasting in Ankara, Turkey using air pollution index and its relation to assimilative capacity of the atmosphere. *Environ Monit Assess*, 166:11-27.
- Günay, O. ve Yavuz, C.I. 2009. Bir Case Crossover (Çapraz Vaka Kontrol) Çalışması: Bir Yıllık Myokard Enfarktüsü Nedenli Yatışlar ve Hava Kirliliği İlişkisinin İncelenmesi. *Kocaeli Örneği. TAF Preventive Medicine Bulletin* . Sep/Oct2009, Vol. 8 Issue 5, p381-388. 8p. 5 Charts.
- Hooper, G.L. and Kaufman, J.D. 2018. Ambient Air Pollution and Clinical Implications for Susceptible Populations. *ATS Journals* Vol. 15.
- HYLLEBERG, S., ENGLE, R.F., GRANGER, C.W.J. and YOO, B.S. 1990. "Seasonal Integration and Cointegration", *Journal of Econometrics*.
- Ivanov, A. and Gocheva-Ilieva, S. 2013. Short-time particulate matter PM10 forecasts using predictive modeling techniques. *AIP Conference Proceedings*. Oct2013, Vol. 1561 Issue 1, p209-218. 10p. 5 Charts, 5 Graphs.
- Koutrakis, P., Sax, S.N., Sarnat, J.A., Coull, B., Demokritou, P., Oyola, P., Garcia, J. and Gramsch, E. 2005. Analysis of PM₁₀, PM_{2.5}, and PM_{2.5-10} concentrations in Santiago, Chile from 1989 to 2001. *Journal of the Air & Waste Management Association*, 55:3, 342-351.

- Kurt, S. 2013. Bazı illere ilişkin yağış miktarı verisinin doğrusal ve doğrusal olmayan zaman serileri analizi. Yüksek lisans tezi. Gazi Üniversitesi.
- Liu, C., Yang, C., Zhao, Y., Ma, Z., Bi, J., Liu, Y., Meng, X., Wang, Y., Cai, J., Kan, H. and Chen, R. 2016. Associations between long-term exposure to ambient particulate air pollution and type 2 diabetes prevalence, blood glucose and glycosylated hemoglobin levels in China. *Environment International*, 92-93, 416-421
- Novikov, I., Kalter-Leibovici, O., Chetrit, A., Stav, N. and Epstein, Y. 2012. Weather conditions and visits to the medical wing of emergency rooms in a metropolitan area during the warm season in Israel: a predictive model. *INTERNATIONAL JOURNAL OF BIOMETEOROLOGY*; JAN, 2012, 56 1, p121-p127, 7p.
- Omay, R. E. ve Canpolat, E. 2013. Parametrik Olmayan Sabit Etkiler Panel Veri Modelleri: Türkiye için Hava Kirliliği ve Gelir İlişkisi. *Anadolu Üniversitesi Bilim ve Teknoloji Dergisi Cilt 14 Sayı1*
- Ostro, B.D., Hurley, S. and Lipsett, M.J. 1999. Air pollution and daily mortality in the Coachella Valley, California: a study of PM₁₀ dominated by coarse particles. *Environmental Research Section A* 81, 231-238.
- Öztaş, S. 2012. Erzurum ilindeki doğalgaz tüketiminin zaman serileri analizi ile incelenmesi ve bir uygulama. Yüksek lisans tezi. Atatürk Üniversitesi.
- Pope III, C.A. and Dockery, D.W. 1992. Acute health effects of PM₁₀ pollution on symptomatic and asymptomatic children. *Am Rev Respir Dis*, 145:1123–1128.
- Shi, J., Chen, R., Yang, C., Lin, Z., Cai, J., Xia, Y., Wang, C., Li, H., Johnson, N., Xu, X. and Zhao, Z. 2016. Association between fine particulate matter chemical constituents and airway inflammation: A panel study among healthy adults in China. *Environmental Research*, 150, 264-268.
- Tomassetti, B., Lombardi, A., Cerasani, E., Di Sabatino, A., Pace, L., Ammazalorso, D. and Verdecchia, M. 2013. Mapping of Alternaria and Pleospora concentrations in Central Italy using meteorological forecast and neural network estimator. *AEROBIOLOGIA*; MAR, 2013, 29 1, p55-p70, 16p.
- Turgut, D. and Temiz, I. 2015. Time series analysis and forecasting for air pollution in Ankara: A Box-Jenkins approach. *Alphanumeric Journal the Journal of Operations Research, Statistics, Econometrics and Management Information Systems*, Volume 3, Issue 2, 131-138.
- Turner, M., Krewski, D., Pope III, C.A., Chen, Y., Gapstur, S. and Thun, M. 2011. Long-term Ambient Fine Particulate Matter Air Pollution and Lung Cancer in a Large Cohort of Never-Smokers. *American Journal of Respiratory and Critical Care Medicine*, Vol. 184, No.12.
- Wei, W.W.S. 2006. *Time Series Analysis: Univariate and Multivariate Methods*. Pearson-Addison Wesley, 2nd ed.
- Yarç, F. 1991. Hava kirliliği verilerinin zaman serileri analizinde Box-Jenkins tekniklerinin uygulanması üzerine bir araştırma. Doktora tezi. Uludağ Üniversitesi.
- Yurtseven, İ. and Serengil, Y. 2017. Changes and trends of seasonal total rainfall in the province of Istanbul, Turkey. *İstanbul Üniversitesi Orman Fakültesi Dergisi*, Vol 67, Iss 1, Pp 1-12 (2017).
- Zwart F., Brunekreef B., Timmermans E., Deeg D. and Gehring U. 2018. Air Pollution and Performance-Based Physical Functioning in Dutch Older Adults. *Environmental Health Perspectives*.

ÖZGEÇMİŞ

Adı Soyadı : Faruk ÇELİKKANAT
Doğum Yeri : ANKARA
Doğum Tarihi : 04.07.1983
Medeni Hali : Bekar
Yabancı Dili : İngilizce

Eğitim Durumu

Lise : Kuleli Askeri Lisesi (2001)
Lisans : Kara Harp Okulu Sistem Mühendisliği (2005)
Yüksek Lisans : Ankara Üniversitesi Fen Bilimleri Enstitüsü İstatistik Anabilim Dalı
(Eylül 2015 – Temmuz 2018)

Çalıştığı Kurum/Kurumlar ve Yıl

Türk Silahlı Kuvvetleri (2005 -)