

T.C.
AKDENİZ ÜNİVERSİTESİ



ZAMAN SERİSİ REGRESYONU'NUN MATEMATİKSEL OLARAK
İNCELENMESİ, MODELLERİN AÇIKLANMASI VE BUNUN ÜZERİNE BİR
UYGULAMA

Yağmur KARACA ÜLKÜTANIR

FEN BİLİMLERİ ENSTİTÜSÜ

MATEMATİK

ANABİLİM DALI

YÜKSEK LİSANS TEZİ

EYLÜL 2022

ANTALYA

T.C.
AKDENİZ ÜNİVERSİTESİ



**ZAMAN SERİSİ REGRESYONU'NUN MATEMATİKSEL OLARAK
İNCELENMESİ, MODELLERİN AÇIKLANMASI VE BUNUN ÜZERİNE BİR
UYGULAMA**

Yağmur KARACA ÜLKÜTANIR

FEN BİLİMLERİ ENSTİTÜSÜ

MATEMATİK

ANABİLİM DALI

YÜKSEK LİSANS TEZİ

EYLÜL 2022

ANTALYA

**T.C.
AKDENİZ ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ**

**ZAMAN SERİSİ REGRESYONU'NUN MATEMATİKSEL OLARAK
İNCELENMESİ, MODELLERİN AÇIKLANMASI VE BUNUN ÜZERİNE BİR
UYGULAMA**

**Yağmur KARACA ÜLKÜTANIR
MATEMATİK
ANABİLİM DALI
YÜKSEK LİSANS TEZİ**

**Bu tez T.C. Akdeniz Üniversitesi Bilimsel Araştırma Projeleri (BAP)
Koordinasyon Birimi tarafından FYL-2020-5379 nolu proje ile desteklenmiştir.**

EYLÜL 2022

T.C.
AKDENİZ ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

ZAMAN SERİSİ REGRESYONU'NUN MATEMATİKSEL OLARAK
İNCELENMESİ, MODELLERİN AÇIKLANMASI VE BUNUN ÜZERİNE BİR
UYGULAMA

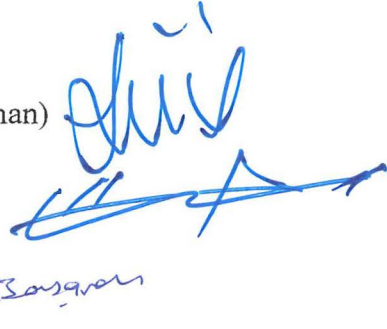
Yağmur KARACA ÜLKÜTANIR
MATEMATİK
ANABİLİM DALI
YÜKSEK LİSANS TEZİ

Bu tez 22/09/2022 tarihinde jüri tarafından Oybirliği / Oyçokluğu ile kabul edilmiştir.

Dr. Öğretim Üyesi Füsun YALÇIN (Danışman)

Prof. Dr. Özkan ÖCALAN

Prof. Dr. Murat Alper BAŞARAN



ÖZET

ZAMAN SERİSİ REGRESYONU'NUN MATEMATİKSEL OLARAK İNCELENMESİ, MODELLERİN AÇIKLANMASI VE BUNUN ÜZERİNE BİR UYGULAMA

Yağmur KARACA ÜLKÜTANIR

Yüksek Lisans Tezi Matematik Anabilim Dalı

Danışman: Dr. Öğr. Üyesi Füsun YALÇIN

Eylül 2022;68 sayfa

Zaman serileri, matematik, istatistik, fizik gibi pozitif bilimlerde artma, azalma, büyüme, küçülme zamana bağlı değişim gösteren veri setlerinde kullanıldığı gibi finans, bankacılık, ekonometri gibi sosyal bilimler alanlarında da nüfus, borsa gibi zaman ile ilişkili birçok alt başlık altında kullanılmaktadır. Bu çalışmada zaman serilerinin hangi alanlarda kullanıldığından bahsedilmiş, zaman serilerinde kullanılan tanımlar verilmiş, zaman serisi regresyonu ve zaman serilerinde ARIMA modelleri araştırılarak bir uygulama yapılmıştır. Zaman serisi regresyonu uygulaması için 1928-2020 yılları arasında Türkiye'deki ebe sayıları yıllık veri olarak TÜİK'ten alınmıştır. Zaman serisi regresyonu için 7 farklı model veriye uygulanmış ve en iyi model bulunmuştur. Bulunan bu model için 2021 ve 2022 yılları için öngörü değerleri elde edilmiştir.

2014 yılında günlük olarak alınan hava kalite indeksi verileri kullanılarak yapılan uygulamada ise ARIMA modelleri veriye uygulanmıştır. Veriler günlük olarak alındığı için tahmin değerleri 2015 yılının ocak ayı için hesaplanmıştır. Öncelikle serinin mevsimsellik barındırıp barındırmadığına ve durağan olup olmadığına bakılmıştır. Gerekli analizler yapılmış ve bunun sonucunda serinin mevsimsellik barındırmadığı ve serinin düzeyde durağan olduğu görülmüştür. Serinin mevsimsel bir seri olmaması ve düzeyde durağan olması nedeniyle veri setine en uygun ARIMA modelinin ARMA(p,q) modeli olduğuna karar verilerek modelleme yapılmıştır. Yapılan modeller arasından en iyi model seçilerek 2015 yılının ocak ayına ait öngörü değerleri bulunmuş ve bulunan öngörü değerleri EPA hava kalite indeksi tablosuna göre değerlendirilmiştir.

ANAHTAR KELİMELER: Hava Kalite İndeksi, Regresyon analizi, Zaman Serilerinde Regresyon Analizi, Zaman serilerinde modelleme, ARIMA

JÜRİ: Dr. Öğr. Üyesi Füsun YALÇIN

Prof. Dr. Özkan ÖCALAN

Prof. Dr. Murat Alper BAŞARAN

ABSTRACT

MATHEMATICAL EXAMINATION OF TIME SERIES REGRESSION, EXPLANATION OF MODELS AND A RELATED APPLICATION

Yağmur KARACA ÜLKÜTANIR

MSc Thesis in MATHEMATICS

Supervisor: Asst. Prof. Dr. Füsun YALÇIN

September 2022; 68 pages

Time series are used in positive sciences such as mathematics, statistics, physics, in data sets that show increase, decrease, growth, shrinkage depending on time, as well as in social sciences such as finance, banking and econometrics, they are used under many time-related subheadings such as population and stock market. In this study, it is mentioned in which areas time series are used, definitions used in time series are given, an application is made by searching time series regression and ARIMA models in time series. For the time series regression application, the number of midwives in Turkey between the years 1928-2020 was taken from TUIK as annual data. For time series regression, 7 different models were applied to the data and the best model was found. For this model, prediction values were obtained for the years 2021 and 2022.

In the application made by using the air quality index data taken daily in 2014, ARIMA models were applied to the data. Since the data are taken on a daily basis, the estimation values are calculated for January 2015. First of all, it was checked whether the series contains seasonality and whether it is stationary. Necessary analyzes were made and as a result, it was seen that the series did not contain seasonality and that the series was stationary at the level. Modeling was done by deciding that the most suitable ARIMA model for the data set is the ARMA(p,q) model, since the series is not a seasonal series and is stationary at the level. The best model was selected among the models, and the prediction values for January 2015 were found and the predicted values were evaluated according to the EPA air quality index table.

KEYWORDS: Air Quality Index, Regression analysis, Regression Analysis in Time Series, Modeling in time series, ARIMA models

COMMITTEE: Asst. Prof. Dr. Füsun Yalçın

Prof Dr. Özkan Öcalan

Prof. Dr. Murat Alper BAŞARAN

ÖNSÖZ

Zaman serileri, birçok bilim dalında kullanılmaktadır. Zamana bağlı olan, düzenli aralıklarla alınan her veri setine uygulanabilmektedir. Bu da zaman serileri için geniş bir çevre yaratmaktadır. Birçok bilim dalında kullanılabilir olması açısından çok yönlü bir konudur. Zaman serisi verisi ileriye yönelik tahminlerde bulunmamız için bir araç görevi görmektedir. Özellikle gelişmekte olan ülkelerde iktisat teorilerinin açıklanmasında mühendislikte korozyon sonuçlarının tahminini hesaplamada, jeolojide zeminde oluşan erozyonun gelecekteki doğurabileceği sonuçları tahmin etmede, biyolojide bakterilerin çoğalması ile ilgili durumlarda, tıpta tümördeki değişimlerin zamanla ölçülerek durumunun saptanması, ileriye yönelik tahminlerinin bulunması ve yeni tedavi yöntemlerine karar verilmesinde ve benzeri birçok konuda ve bilim dalında kullanılmaktadır. Zaman serisinin çok yönlü bir konu olması, elimizdeki verilerde gelecek tahmini yapmamıza olanak sağlaması bu çalışmayı yapmamın temel sebeplerinden bir tanesidir. Bu tez çalışmasında zaman serisi ile ilgili, mevsimsellik, durağanlık, trend, ana kütle, örneklem ve diğer bazı tanımlar verilmiştir. Regresyon kavramından bahsedilmiş zaman serisi regresyonu modelleri ve zaman serilerinde ARIMA modelleri incelenmiştir. Zaman serisi modellemelerinin denklemleri verilmiş ve zaman serisi analizi üzerine bir uygulama yapılmıştır.

Bu tez giriş, kaynak taraması, materyal ve metot, bulgular ve tartışma, sonuçlar olmak üzere 5 ayrı bölümden oluşmaktadır.

Giriş kısmında zaman serisi analizlerinden genel bahsedilmiş ve bu tezin yapılaş amacı verilmiştir.

Kaynak taraması kısmında zaman serisi ile ilgili tezde kullanılan tanımlar, verilmiştir. Bunun yanında zaman serisi analizi ile ilgili literatür taraması yapılmış ve bu çalışmaların içeriğinden bahsedilmiştir.

Materyal metot kısmında ise zaman serisi regresyonu, mevsimsellik, durağanlık, ve ARIMA modelleri verilmiştir.

Bulgular kısmında ise 1928-2020 yılları arasında Türkiye'deki ebe sayıları ile oluşturulan veri seti kullanılarak yapılan zaman serisi regresyon modellemelerinin

analiz sonuçları ve 2014 yılı için günlük alınan hava kalite indeksi verilerinden yapılan ARIMA modellerinin çıktıları çizelgeler ve şekiller ile sunulmuş ve bu çalışmadan elde edilen çıktılar açıklanmıştır.

Sonuçlar kısmında ise bulgular kısmındaki analizlerin sonuçları tartışılmıştır.

2018 yılında yüksek lisans serüvenime başladığımda yanımda ailem ve Sayın Danışmanım Dr. Öğr. Üyesi Füsun Yalçın vardı. Şimdi bu serüvenin sonuna geldiğimde bu yolda bana eşim ve oğlumda katıldı. Pandemi ile birlikte geçen zorlu bir süreç idi.

Bu süreçte tüm zorluklara rağmen yoluma devam etmem konusunda beni hep destekleyen, asla pes etmeyen ve asla pes etmeme müsaade etmeyen danışmanım Dr. Öğr. Üyesi Füsun YALÇIN'a sonsuz teşekkürlerimi ve minnet duygularımı sunuyorum.

Bu çalışmayı destekleyen Akdeniz Üniversitesi BAP Koordinasyon Birimi'ne teşekkürlerimi sunarım.

Maddi manevi yanımda olan eşim Metecan ÜLKÜTANIR 'a, bu yola başlarken beni her zaman destekleyen anne ve babama, sonsuz teşekkür ederim.

Yüksek Lisans tez sürecinde hayatıma katılan, asla çalışmama müsaade etmeyen ancak bu yolda hayatıma katılarak bana güç veren oğlum Ateş ÜLKÜTANIR' a sonsuz sevgilerimi sunuyorum.

Son olarak yüksek lisans tez sürecimin sonuna gelmiş bulunmaktayım ve şimdide hayatıma bir kız çocuğu daha katılacak. Manevi olarak gücüme güç kattığı için bu satırlarla ona da sonsuz sevgilerimi sunuyorum.

İÇİNDEKİLER

ÖZET.....	i
ABSTRACT.....	iii
ÖNSÖZ.....	v
AKADEMİK BEYAN.....	ix
SİMGELER VE KISALTMALAR.....	x
ŞEKİLLER DİZİNİ.....	xii
ÇİZELGELER DİZİNİ.....	xiii
1. GİRİŞ.....	1
2. KAYNAK TARAMASI.....	3
3. MATERYAL VE METOT.....	11
3.1. Regresyon.....	11
3.2. Durbin-Watson Testi.....	13
3.3. Zaman Serisi Regresyonu.....	15
3.3.1. Basit doğrusal regresyon modeli.....	15
3.3.2. Birinci farklar regresyon modeli.....	16
3.3.3. Üstel regresyon modeli.....	16
3.3.4. Karesel regresyon modeli.....	16
3.3.5. Lojistik regresyon modeli.....	17
3.3.6. Kübik regresyon modeli.....	17
3.3.7. Logaritmik regresyon modeli.....	17
3.3.8. Güç regresyon modeli.....	17
3.3.9. S regresyon modeli.....	17
3.3.10. Ters regresyon modeli.....	18
3.4. Trend ve Mevsimsellik.....	18
3.5. Durağanlık.....	20
3.6. Kesikli ve Sürekli Zaman Serileri.....	22
3.7. Arttırılmış Dickey-Fuller Testi (The Augmented Dickey-Fuller Test).....	23
3.8. Phillips Perron Testi.....	24

3.9. ARIMA Modelleri	25
3.9.1. AR(p) modeli	25
3.9.2. MA(q) modeli	27
3.9.3. ARMA(p,q) modeli.....	29
3.9.4. ARIMA(p,d,q) modeli	30
3.9.5. ARFIMA (p,d,q) modeli	31
4. BULGULAR VE TARTIŞMA	33
4.1 Zaman Serisi Regresyon Analizi	33
4.2 Arıma Modellerinin Tahmini.....	46
5. SONUÇLAR	64
6. KAYNAKLAR	67
ÖZGEÇMİŞ	

AKADEMİK BEYAN

Yüksek Lisans Tezi olarak sunduğum “Zaman Serisi Regresyonu’nun Matematiksel Olarak İncelenmesi, Modellerin Açıklanması Ve Bunun Üzerine Bir Uygulama” adlı bu çalışmanın, akademik kurallar ve etik değerlere uygun olarak yazıldığını belirtir, bu tez çalışmasında bana ait olmayan tüm bilgilerin kaynağını gösterdiğimi beyan ederim.

22/09/2022

Yağmur KARACA ÜLKÜTANIR



SİMGELER VE KISALTMALAR

Simgeler

y_t	: Zaman serisi
h_0	: Yokluk hipotezi
h_1	: Alternatif hipotez
$\gamma(h)$: Otokovaryans fonksiyonu
$E(y_t)$: Beklenen değer
Δy_t	: Zaman serisinin birinci farkı
$\ln(y_t)$: Zaman serisinin doğal logaritması
t_δ	: Test istatistiği
L	: Gecikme işlemcisi
ρ	: AR(1) Parametresi
θ	: MA(1) Parametresi

Kısaltmalar

AR	: Otoregresif seri
ARFIMA	: Otoregresif kesirli tümleşik hareketli ortalama
ARH	: Günlük ortalama bağıl nem
ARIMA	: Otoregresif entegre edilmiş hareketli ortalama
ARM	: Toplamsal regresyon modeli
ARMA	: Otoregresif hareketli seri
AT	: Günlük ortalama sıcaklık
Cov	: Kovaryans
EMA	: Üssel hareketli ortalama
EPA	: Amerika Birleşik Devletleri Çevre Koruma Ajansı
HKI	: Hava kalite indeksi

MA : Hareketli ortalama
MAE : Ortalama mutlak hata
MAPE : Ortalama mutlak yüzde hata
RMSE : Ortalama kare hatalarının karekökü
TÜİK : Türkiye İstatistik Kurumu

ŞEKİLLER DİZİNİ

Şekil 3.1. Standart normal dağılım tablosu.....	15
Şekil 4.1. 1928-2020 yılları arası Türkiye'deki ebe sayısı zaman yolu grafiği	34
Şekil 4.2. Türkiye'deki ebe sayıları için basit doğrusal regresyon modeline ait bazı istatistikler.....	35
Şekil 4.3. Türkiye'deki ebe sayısı tahmini (basit doğrusal regresyon modeli).....	35
Şekil 4.4. Türkiye'deki ebe sayıları birinci farklar regresyon modeli için bazı istatistikler.....	36
Şekil 4.5. Türkiye'deki ebe sayısı tahmini (birinci farklar regresyon modeli).....	37
Şekil 4.7. Türkiye'deki ebe sayısı tahmini (üstel regresyon modeli)	38
Şekil 4.8. Türkiye'deki ebe sayıları için karesel regresyon modeline ait bazı istatistikler	39
Şekil 4.9. Türkiye'deki ebe sayısı tahmini (karesel regresyon modeli).....	40
Şekil 4.10. Türkiye'deki ebe sayısı için lojistik regresyon modeline ait bazı istatistikler	41
Şekil 4.11. Türkiye'deki ebe sayısı tahmini (lojistik regresyon modeli).....	41
Şekil 4.13. Türkiye'deki ebe sayısı tahmini (kübik regresyon modeli).....	43
Şekil 4.15. Türkiye'deki ebe sayısı tahmini (logaritmik regresyon modeli)	44
Şekil 4.16. Hava kalite indeksinin günlük değişim zaman yolu grafiği	47
Şekil 4.17. Hava kalite indeksi korelogramı	48
Şekil 4.18. AR(1) modeli kalıntı korelogramı	50
Şekil 4.19. AR(2) modeli kalıntı korelogramı	52
Şekil 4.20. MA(1) modeli kalıntı korelogramı	53
Şekil 4.21. MA(2) model kalıntı korelogramı	55
Şekil 4.22. ARMA(1,1) model kalıntı korelogramı.....	56
Şekil 4.23. ARMA(2,1) model kalıntı korelogramı.....	58
Şekil 4.24. ARMA (1,2) model kalıntı korelogramı.....	60
Şekil 4.25. ARMA(2,2) model kalıntı korelogramı.....	62

ÇİZELGELER DİZİNİ

Çizelge 4.1. Türkiye’deki ebe sayıları için basit doğrusal regresyon modeli tahmini....	34
Çizelge 4.2. Türkiye’deki ebe sayıları için birinci farklar regresyon modeli tahmini....	36
Çizelge 4.3. Türkiye’deki ebe sayıları için üstel regresyon modeli tahmini	37
Çizelge 4.4. Türkiye’deki ebe sayıları için karesel regresyon modeli tahmini.....	39
Çizelge 4.5. Türkiye’deki ebe sayıları için lojistik regresyon modeli tahmini.....	40
Çizelge 4.6. Türkiye’deki ebe sayıları için kübik regresyon modeli tahmini.....	42
Çizelge 4.7. Türkiye’deki ebe sayısı için logaritmik regresyon modeli tahmini.....	43
Çizelge 4.8. Zaman serisi regresyon modelleri.....	45
Çizelge 4.9. Kübik regresyon modeline ait tahmin değerleri	46
Çizelge 4.10. Değişkenlerin tanımlayıcı istatistikleri	46
Çizelge 4.11. ADF birim kök testi çıktıları	48
Çizelge 4.12. PP birim kök testi (trendli ve sabitli model) çıktıları	49
Çizelge 4.13. AR(1) model tahmin analizleri	49
Çizelge 4.14. AR(2) model tahmin analizleri	51
Çizelge 4.15. MA(1) model tahmin analizleri	52
Çizelge 4.16. MA(1) model tahmin analizleri	54
Çizelge 4.17. ARMA(1,1) model tahmin analizleri.....	55
Çizelge 4.18. ARMA(2,1) model tahmin analizleri.....	57
Çizelge 4.19. ARMA(1,2) model tahmin analizleri.....	58
Çizelge 4.20. ARMA(2,2) model tahmin analizleri.....	60
Çizelge 4.21 ARMA(2,2) model kalıntıları için ADF testi	62
Çizelge 4.22. ARMA(2,2) tahmin değerleri	63
Çizelge 5.1 EPA hava kalite indeksi (Zencirci ve Işıklı 2017).....	65

1. GİRİŞ

İki olayın birbirini etkileyip etkilemediği korelasyon ve regresyon analizinde incelenebilir. Korelasyon analizi iki değişkenin birbirini etkileyip etkilemediğini açıklarken, regresyon analizi bu iki değişkenin birbirini ne kadar etkilediğini bize sayısal olarak açıklar. Regresyon analizinde kullanılan matematiksel model doğrusal ya da eğrisel olabilir. Verilerin saçılım grafiğine bakarak araştırmacı buna karar verebilir. Ayrıca regresyon analizinin varsayımları araştırmacı tarafından mutlaka incelenmelidir. Birçok disiplindeki çalışmalarda regresyon analizi sıkça karşımıza çıkmaktadır. Özellikle zaman serisi regresyonu araştırmacıların ilgisini çekmektedir. Zaman serisi regresyonu modelleri öngörü ve tahmin yapmak için sıkça kullanılır. Bu nedenle bu çalışmada zaman serisi regresyonu ve zaman serisi modellerine ağırlık verilmiştir.

Geçmişe dayalı veriyi kullanarak öngörü ve tahminlerde bulunmak bilim dünyası için önemli bir kavramdır. Birçok alanda geçmişe ait veriler kullanılarak ileriye dönük tahminler ve öngörüler yapılmaktadır. Düzenli zaman dilimlerinde ve ardışık zaman aralıklarında ölçülen verilere zaman serisi verisi denir. Zaman serisi istatistik, ekonometri, matematiksel finans alanlarında kullanılan veri noktalarının sıklığını ifade eder. Zaman serisi analizi, zamana bağlı veri setlerinin incelenmesi, analiz edilmesi, belirli zaman dilimlerinde alınan verilerin incelenerek geleceğe dönük tahminler yapılması açısından önemlidir.

Veri analizi yatay kesit verisi, panel veri ve zaman serisi olmak üzere 3 farklı başlıkta toplanabilir. Son yıllarda dördüncü bir veri grubu olarak büyük veri (big data) literatürde dikkat çekmektedir. Yatay kesit verisi, belli bir zaman anında farklı birimlerden toplanan veriler ile elde edilir. Yatay kesit verilerinin ise bir zamana yayılarak oluşturulması panel veri olarak adlandırılır. Zaman serisi ise farklı zaman dilimlerinde, birimlerden elde edilen veri setleri ile oluşmaktadır.

Zaman serisi analizi yaparken bazı adımlara dikkat etmek gerekmektedir. Öncelikle elimizdeki veri setinin belirli zaman dilimlerinde alınmış, zamana bağlı bir veri seti olması gereklidir. Veri setinin uzunluğu, verinin ani bir şok karşısında gösterdiği direnç açısından önemlidir. Literatürde bununla ilgili net bir fikir birliği bulunmasa da bazı kaynaklar 30 veri seti önermektedir. Bunun yanı sıra mevsimselliğe

bakılmaktadır. Yıllık verilerden daha çok aylık haftalık verilerde daha çok mevsimselliğe rastlanmaktadır. Bu durumda veriler mevsimsellikten arındırılarak analize devam edilmelidir. Mevsimselliğin hangi yapıda olduğu bu noktada önem arz etmektedir. Mevsimsellik deterministik ya da skotastik yapıda olabilir. Deterministik yapıda olan mevsimselliği arındırmak daha güvenli olacaktır. Zaman serisi analizlerinde mevsimselliğin yapısının belirlenmesi için mevsimsel birim kök testleri kullanılabilir. Eğer zaman serisinde mevsimsel birim kök bulunuyorsa SARIMA model tahmini yapılabilir. Mevsimselliğin yapısı deterministik ise arındırma işlemi uygulanarak sıradan birim kök testleri ile analize devam edilir (Mert ve Çağlar 2019).

Bu tezin amacı zaman serisi analizini hem teorik hem uygulamalı olarak açıklayarak bir derleme oluşturmaktır. Tez çalışmasının uygulamasında iki farklı veri seti kullanılmıştır. Zaman serisi regresyonu uygulaması yapılırken 1928-2020 yılları arasında Türkiye'deki ebe sayısı ile oluşturulan yıllık veri seti kullanılmıştır. Bu veri setine 7 farklı regresyon modeli uygulanmış ve aralarından veriyi en iyi modelleyen regresyon modeli seçilmiştir. Bu regresyon modeli için 2021 ve 2022 yıllarına ait olan Türkiye'deki ebe sayıları tahmin edilmiştir.

Zaman serisi modellerinden ARIMA modellerinin uygulaması yapılırken ise 2014 yılına ait günlük meteorolojik hava tahmin verileri kullanılarak 2015 yılının ocak ayı hava kalite indeksi değerlerinin tahmini üzerine bir uygulama yapılmıştır.

Zaman serisi analizleri mühendislik, sağlık, istatistik, temel bilim dalları olmak üzere hemen hemen her alanda kullanılmaktadır. Zaman serisi için bahsedebileceğimiz en önemli beyinlerden biri George Box'tur. 1913-2013 yılına kadar Bayes çıkarımı, zaman serisi analizi ve diğer alanlarda çalışmalar yapmıştır. 20. Yüzyılın en büyük istatistik beyinlerinden biri olarak anılmıştır. İngiltere doğumlu olan George Box üniversitede kimya okumaya başlamıştır ve daha sonrasında askere çağrılmıştır. Dünya savaşı sırasında ise İngiliz ordusu için küçük hayvanları zehirli bir gaza maruz bırakan deneyler yapmıştır. Bunların analizlerini yaparken ise metinlerden okuyarak istatistik öğrenmiştir. Savaştan sonra ise matematik ve istatistik alanında lisans derecesi almıştır. 1953 yılında ise bu başarısını doktora derecesi ile taçlandırmıştır. Bir süre istatistikçi olarak çalışmıştır. Ancak daha sonrasında 1960 yılında Box, istatistik bölümü oluşturmak için Wisconsin-Madison üniversitesine taşındı (Anonim 1).

2. KAYNAK TARAMASI

Bu bölümde zaman serisi regresyonu ile ilgili literatürde yapılan çalışmalardan bazıları ve bu tezde kullanılacak olan zaman serisi regresyonu ile ilgili bazı tanım ve teoremler verilecektir.

Zaman serisi analizleri mühendislik, sağlık, istatistik, temel bilim dalları olmak üzere hemen hemen her alanda kullanılmaktadır. Zaman serisi veri noktalarının sıklığını ifade eder ve düzenli zaman aralıklarında ve ardışık zaman aralıklarında ölçülür. Zaman serisi analizi yapmak için birkaç yöntem bulunmaktadır. Bunlardan biride zaman serisi tahminidir. Zaman serisi tahmini daha önceden bilinen olaylar referans alınarak gelecek olayları tahmin etmenin kavramsal modelidir. Zaman serisi analizinde kullanılan formüllerden biri;

$$X = \{X_1, X_2, \dots\}$$

Şeklinde ifade edilmiştir. Bu zaman serisini temsil eden en temel formüldür. Burada X doğal sayılar dizisidir (Akdi 2012).

Diğer kullanılan yaygın formül ise şu şekildedir;

$$Y = \{Y_t; t \in T\}$$

Zaman serisi analizleri bahsettiğimiz gibi birçok farklı disiplinde kullanılmaktadır. Özellikle ülkemizde de hala izlerini gördüğümüz bir Covid-19 pandemisi geçirdik. Tüm dünya da yaşadığımız bu pandemi sürecinde Covid-19 üzerine birçok çalışma yapılmıştır. Zaman serileri kullanılarak yapılan bu çalışmalardan bazıları aşağıda verilmiştir.

“COVID-19 transmission in Mainland China is associated with temperature and humidity: A time series analysis” adlı makalede dünyadaki herkesin etkilendiği COVID-19 virüsünün sıcaklık ve nem ile bir ilişkisinin bulunup bulunmadığı araştırılmıştır. Bu çalışma 30 Çin ilinde incelenmiştir. Çin’deki tüm illerden laboratuvar onaylı vakalar Çin Halk Cumhuriyeti Ulusal Sağlık Komisyonu’nun resmi raporlarından toplanmıştır. Her ilin günlük ortalama sıcaklığı ve günlük ortalama bağıl nemi dahil olmak üzere meteorolojik veriler analiz edilmiştir.

Her il başkentinin günlük ortalama sıcaklığı (AT) ve günlük ortalama bağıl nemi (ARH) dahil olmak üzere meteorolojik veriler Weather Underground'dan alınmıştır (Anonim 2).

Meteorolojik değişkenler ile covid-19 vakaları arasında illere özgü ilişkileri ölçmek için geliştirilmiş bir katkı modeli oluşturulmuştur. Günlük vaka sayılarının varyanslarının, ortalamalarından daha büyük olduğu göz önüne alındığında covid-19 vakalarının dağılımının negatif iki terimli olduğu varsayılmıştır. Covid-19'un kuluçka dönemi göz önünde bulundurularak günlük ortalama sıcaklık, günlük ortalama bağıl nem ve diğer ortak değişkenlerin etkilerini, potansiyel gecikme etkisini hesaba katmak için 14 günlük üstel ortalama (EMA) ile modellendi.

Bu çalışmanın sonucunda hem günlük sıcaklık hem de bağıl nemin Çin'in Hubei eyaletinde ve bazı illerde Covid-19 oluşumunu etkilediği gözlemlenmiştir. Ancak iller arasında Covid-19 ile günlük ortalama sıcaklık (AT) ve günlük ortalama bağıl nem arasındaki ilişkiler tutarlı olmamıştır. Zaman serisinin uzunluğunun Hubei eyaletinde diğer illerden daha fazla olmasından dolayı ve zaman serisinin uzunluğu ne kadar fazla ise o kadar tutarlı olacağından Günlük sıcaklık ve bağıl nem 'in Covid-19'un bulaşması üzerinde olumlu etkileri bulunmuştur. Ancak çalışma diğer illerde meteorolojik faktörler baz alındığında tutarsızlık ile sonuçlanmıştır.

Ünsal, yaptığı bu çalışmada, regresyon ve varyans analizi yöntemi ile Türkiye'ye 1985-1995 yılları arasında, aylık ve üç aylık dönemlerde gelen turist sayısının oluşturduğu zaman serisinde, mevsimsel bir etkinin olup olmadığı araştırmıştır. Mevsimsel bir etkinin olup olmadığını araştırırken ise varyans analizini kullanmıştır.

Doğrusal olmayan zaman serileri için parametrik ve parametrik olmayan yöntemlerin kullanıldığı bilinmektedir. Bu çalışmada, parametrik yöntemlerden otoregresif (AR) ve kendinden eşik değerli (SETAR) modelleri, parametrik olmayan yöntemlerden ise toplamsal regresyon modeli (ARM) kullanılmıştır. Parametrik olmayan regresyon teknikleri hatalardaki otokorelasyonun varlığına genellikle duyarlıdır. Bu duyarlılığın pratik sonuçları düzeltme parametresinin uygun seçimiyle açıklanır. Bu amaçla, Türkiye'deki ihracat birim değer endeks verisi, AR, SETAR ve ARM modelleri ile tahmin edilerek uygun model belirlenmeye çalışılmıştır. Bu

çalışmanın uygulama kısmında 2002:12-2012:10 yılları arasındaki ihracat birim değeri endeksi verileri kullanılmıştır. Bu uygulamanın verileri ise Türkiye İstatistik Kurumu'nun internet sitesinden alınmıştır (Anonim 3).

Ünsal, zaman serilerinde regresyon ve varyans analizi yöntemleri ile mevsimsel dalgalanmaların araştırılması ve bir uygulama adlı çalışmasında regresyon ve varyans analizi metodu ile 1985-1995 yılları arasında Türkiye'yi aylık ve üç aylık dönemlerde ziyarete gelen turistlerin oluşturduğu zaman serisinde mevsimler bir etkinin var olup olmadığını araştırmıştır. Zaman serilerinde mevsimsel etkinin varlığı ya da yokluğu için kullanılan bazı yöntemler vardır. Bu yöntemlerden bir tanesi de varyans analizi ve regresyondur. Ünsal, bu çalışmasında mevsimselliği araştırmak için varyans analizi ile mevsimsellik araştırması yaparken, regresyon modeli ile de mevsimselliği incelemiştir. Eğer aylık veri setinde mevsimsellik incelemesi yapılıyor ise varyans analizi modeli aşağıdaki şekilde kurulur (Ünsal 1997).

$$Y_{ij} = \mu + a_i + \varepsilon_{ij} \quad i: 1,2, \dots, 12 \quad j: 1,2, \dots, n \quad (2.1)$$

(2.1) Numaralı model kullanılarak mevsimsel etkinin varlığı araştırılırken hipotez;

$$H_0: a_1 = a_2 = \dots = a_{12}$$

H_1 : en az iki a_i birbirinden farklıdır.

mevsimsel etkinin varlığından bahsetmek için H_0 hipotezinin reddedilmesi gerekmektedir.

Ünsal'ın 1997'de yaptığı bu çalışmada mevsimsel etki 5 farklı model ile açıklanmıştır. Aylık veriler ile oluşturulan regresyon modelinin parametre tahmini için aralık ayı sabit terim olarak tanımlanmıştır ve modelin dışına çıkartılmıştır. Bu durumda parametreler tahmin edilebilir hale getirilmiştir. Bu modelin sonucunda $F = 11.141$ değerinin büyük olması sebebiyle $\alpha = 0.05$ anlamlılık düzeyinde anlamlı bulunmuştur. Parametrelerin tahmin değerleri ise $Q_1, Q_2, Q_3, Q_4, Q_{11}$ aylar dışında t-istatistiği ile $\alpha = 0.05$ anlamlılık düzeyinde anlamlı bulunmuştur. Bağımlı değişkendeki (turist sayısı) varyansın bağımsız değişkenler (aylar) ile açıklanan kısmı $R^2=0.505$ bulunmuştur. Varyans analizi modeli ve regresyon modelinin sonuçlarının klasik mevsim verileri ile yapılan analizlerden daha iyi sonuçlar verdiği gözlemlenmiştir (Ünsal 1997).

Son zamanlarda etkinliğini artıran dijital para birimleri uluslararası piyasalarda çokça talep görmeye başlamıştır. Bu dijital para birimleri arasında ise hacmi ve getiri oranları dikkate alındığında en çok göz önünde olan para birimi ise bitcoindir. Bitcoin herhangi bir merkezi otoriteye bağlı değildir *ve P2P³* teknolojisini kullanılır. Karşılığında altın gibi herhangi bir fiziksel maden olmamasından dolayı tamamen dijitaldir. Sınırlı bir kullanım alanı mevcuttur ve sigortalanamaz. Bitcoin, 2008 yılının kasım ayında Satoshi Nakamoto tarafından kapalı bir mail grubuna yollanan bir e-mail ile tanıtılmıştır. İlk Bitcoin yazılımı ise 2009 yılında yayınlanmıştır (Aslantaş Ateş, 2016). Bitcoinin, piyasa değerinin giderek yükselmesi ile insanların çokça ilgisine maruz kalması, bazı ödeme yöntemlerinde bitcoin para biriminin geçerli kılınması bitcoinin bu kadar göz önünde olmasına sebep olmuştur. Bunun sonucunda dijital paralar ile ilgili bilimsel çalışmalarda yoğunluk kazanmıştır.

2018 yılında Shaily Roy, Samiha Nanjiba, Amitabha Chakrabarty tarafından yapılan zaman serisi çalışmasında öncelikte tutarlı bir veri seti oluşturup önceki bitcoin verilerine dayanarak gelecekteki Bitcoin kapanış fiyatı tahmin edilmeye çalışılmıştır. Veri seti 2013 Temmuz ayından 2017 Ağustos ayına kadar olan günlük piyasa sermayesi, işlem hacmi, bitcoin'in USD cinsinden açılış ve kapanış fiyatını kapsamaktadır. Veri seti coindex'ten alınmıştır. Durağanlık kontrol edilirken Augmented Dickey-Fuller testi kullanılmıştır. Artırılmış Dicket-Fuller testi, birim kök testi adı verilen bir tür olgusal testtir ve daha büyük ve karmaşık zaman serileri için kullanılması iyidir. Çok sayıda birim kök testi vardır bunların içinde en genel olarak Augmented Dickey-Fuller öne çıkmaktadır. Dickey-Fuller testi kullanılarak p değeri ölçülmeye çalışılmış ve bu p değeri 0.05'ten büyük olduğundan verilerin durağan olmadığına karar verilmiştir. Bunun sonucunda verilerin logaritması alınıp farkı alınarak bir dönüşüm uygulanmıştır. Logaritma ile dönüşüm yapıldığında veriler normal ya da normale yakın bir dağılım göstermektedir (Roy vd. 2018).

Otoregresif entegra hareketli ortalama modeli (ARIMA), Otoregresif model (AR) ve Hareketli ortalama (MA) zaman serisi modelleri işlenmiş veri setine uygulanmıştır ve elde edilen grafik çizilmiştir. Modellerin doğruluğuna dayanarak ve verilerin içine daha iyi uyum sağlamasından yola çıkılarak Bitcoin fiyatını tahmin etmek için ARIMA modeli kullanılmıştır. Analizler yapıldıktan sonra MA modelinin

%87.58 ile en düşük doğruluğa sahip olduğunu ve ARIMA modelinin %90.31 ile kesinlikle daha iyi olduğu sonucuna ulaşılmıştır. Bu çalışmada 10 günlük bitcoin fiyatları tahmin edilmiştir. Gelecekte ise aynı modeller doğrultusunda rastgele bir tarih için bitcoin tahmini yapılmak hedeflenmiştir (Roy vd. 2018).

Su kalitesi insan yaşamını temelden etkileyen dünya çapında bir sorundur. Su kalitesinin bozulması sonucunda su kıtlığı çoğalmaktadır. Bu sorunun kaynağı olarak; nüfus artışı, ekonomik gelişmeler ve su kirliliği sebep olarak gösterilebilir. Taheri Tizro ve diğerlerinin 2014 yılında yaptıkları çalışmada gelecekte yaşanması muhtemel su kalitesi bozulmalarını önlemek amacı ile su kalitesi araştırılmış ve zaman serisi modellemesi kullanılarak tahminler incelenmiştir. Hor Rood nehrinin kalitesi, zaman serisi analizi kullanılarak Kakareza istasyonunda incelenmiştir. TDS , EC , HCO_3^- , SO_4^{2-} , Mg^{2+} , Ca^{2+} , Na gibi 9 adet su parametresi incelenmiştir. Ph ve Sar olmadığında tüm parametreler için artan bir eğilim olduğu gözlenmiştir. Nehir suyunun kalitesini, oluşturmak ve tahmin etmek için ARIMA modeli uygun bulunmuştur. Üretim ve tahmin sonuçlarının değerlendirilmesinde AIC , R^2 , $RMSE$, % kriterleri kullanılmıştır. Sonuçlar su kalitesinin tahmininde zaman serisi modellemelerinin oldukça yetenekli olduğunu göstermiştir. Üretilen ve tahmin edilen bütün parametreler için R^2 değeri SO_4^{2-} hariç 0.66 üstünde bulunmuştur. R^2 nin değeri üretilen SO_4^{2-} için 0.48, tahmin değeri ise 0.43 olarak hesaplanmıştır. Bu çalışmanın sonucunda parametrelerin çoğu için artan bir eğilime bağlı olarak su kalitesinin kötüleştiği ve ciddi anlamda yönetimsel eylemlerin gerekli olduğu kanısına varılmıştır (Tizro vd. 2014).

Roussel ve diğerlerinin yaptığı bu çalışmada, Pakistan örneği kullanılmış ve 1980 yılından 2019 yılına kadar olan para talebini etkileyen faktör araştırılmıştır. Bağımlı değişken olarak geniş para seçilirken, bağımsız değişken olarak ise hane halkı tüketimi, devlet tüketimi, faiz oranı, tüketici fiyat endeksi, nüfus artışı ve işçi dövizleri seçilmiştir. Değişkenlerin durağanlığını test etmek için ADF ve PP birim kök testleri kullanılmıştır. Modelin değişkenleri arasındaki eşbütünlümü bulmak için ARDL yöntemi ve nedensellik ilişkisini incelemek içinse Granger nedensellik testi kullanılmıştır. Çalışmanın tahmin edilen sonuçları, Pakistan'daki para talebinin belirlenmesinde sosyo-ekonomik faktörlerin önemli bir rol oynadığını göstermiştir.

Bu tezde kullanılacak olan zaman serisi regresyonu ile ilgili tanım ve teoremlerden bazıları aşağıda verilmiştir.

Tanım 2.1

i) $P(A) \geq 0$, bütün $A \in U$

ii) $P(\Omega) = 1$

iii) A_n ler U da ayrık ($A_k \cap A_j = \emptyset, k \neq j$)

olayların bir dizisi ise

$$(\bigcup_{n=1}^{\infty} A_n) = \sum_{n=1}^{\infty} P(A_n) \text{ dir (Akdi 2012).}$$

Tanım 2.2 (Ω, U, P) bir olasılık uzayı olsun. Ω örnek uzayından \mathbb{R} kümesine tanımlı

$$X: \Omega \rightarrow \mathbb{R} \quad w \rightarrow X(w)$$

X fonksiyonu, $\forall a \in \mathbb{R}$ için $\{w: X(w) \leq a\} \in U$ özelliğini sağlıyorsa X fonksiyonuna bir rasgele değişken denir. Benzer şekilde

$$X: \Omega \rightarrow \mathbb{R}^k$$

$$w \rightarrow X(w) = (X_1(w), \dots, X_k(w))$$

fonksiyonu her i ve her $a_i \in \mathbb{R}$ için

$\{w: X_{i(w)} \leq a_i\} \in U$ koşulunu sağlıyorsa X 'e k - boyutlu rasgele vektör denir (Akdi 2012).

Tanım 2.3 (Ω, U, P) bir olasılık uzayı, T de bir indis kümesi olsun. Bir zaman serisi $\Omega \times T$ çarpım uzayından reel sayılara giden bir fonksiyondur. Yani bir zaman serisi:

$$X(.,.): \Omega \times T \rightarrow \mathbb{R}$$

$$(w, t) \rightarrow X(w, t)$$

şeklinde tanımlanan fonksiyondur (Akdi 2012).

Tanım 2.4 $\{X_t: t \in T\}$ bir zaman serisi

i) $E(X_t) = \mu$ (beklenen değerin zamana göre değişmediğini gösterir.

ii) $Cov(X_t, X_s)$ sadece $|t-s|$ nin bir fonksiyonudur.

koşullarını sağlıyorsa $\{X_t: t \in T\}$ zaman serisine durağandır denir (Akdi 2012).

Tanım 2.5 Tanım 2.4 (ii) koşuluna göre $Cov(X_{\{t\}}, X_{\{t+h\}})$ kovaryansı sadece h nin bir fonksiyonudur. Bu fonksiyona serinin otokovaryans fonksiyonu denir (Akdi 2012).

$\gamma(h)$ ile gösterilir. μ serinin beklenen değeri olmak üzere otokovaryans fonksiyonu daha açık olarak

$$\gamma(h) = Cov(X_t, X_{t+h}) = E[(X_t - \mu)(X_{t+h} - \mu)]$$

şeklinde yazılır.

Tanım 2.6 Beklenen değeri 0, otokovaryans fonksiyonu

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

şeklinde olan bir zaman serisine beyaz gürültü (White Noise) denir (Akdi 2012).

Tanım 2.7 $\{X_t: t \in T\}$ zaman serisi verildiğinde X_t nin $X_{t-1}, X_{t-2}, X_{t-3}, \dots, X_{t-h}$ üzerine regresyonunda X_{t-h} nin kat sayısına serinin h . Kısmi otokorelasyonudur denir (Akdi 2012).

Tanım 2.8 X_1, X_2, \dots, X_n rasgele değişkenleri verildiğinde X_{n+1} için bir öngörü

$$\hat{X}_{n+1} = E(X_{n+1} | X_1, X_2, \dots, X_n)$$

koşullu beklenen değeridir (Akdi 2012).

Tanım 2.9 Bir zaman serisi analizinin uzun dönemli bileşeni olup, uzun bir zaman dilimindeki büyümeyi veya küçülmeyi ifade eden kavrama trend denir (Akmüt vd. 1999).

Tanım 2.10 Bir veri dizisinin orta konumunu tek bir sayı ile ifade eden betimsel istatistik ölçüsüne ortalama denir (Anonim 4).

Yığın ortalaması $\mu = \frac{\sum x}{N}$ şeklinde ifade edilir. Burada

$\sum x$: yığında yer alan gözlem değerlerinin toplamı

N : Yığında yer alan gözlem sayısı, şeklinde ifade edilir (Akmüt vd. 1999).

Tanım 2.11 Değişken değerlerinin ortalama etrafındaki dağılımını ifade eden kavrama varyans denir (Akmüt vd. 1999).

$$Var(x) = \sigma_x^2 = E(x - \mu)^2$$

şeklinde ifade edilir. Buradan yığının varyansı;

$$\sigma_x^2 = \frac{\sum E(x - \mu)^2}{N} = \frac{\sum x^2 - \frac{(\sum x)^2}{N}}{N} \text{ dir.}$$

Tanım 2.12 Örneklem veya yığında yer alan değerlerin örnek veya yığın ortalamasından farkını ölçen parametreye standart sapma adı verilir. Varyansın pozitif karekökü alınarak hesaplanır ve σ ile gösterilir (Akmüt vd.1999).

$$\sigma = \sqrt{\frac{\sum E(x - \mu)^2}{N}} = \sqrt{\frac{\sum x^2 - \frac{(\sum x)^2}{N}}{N}} \text{ dir.}$$

Tanım 2.13 Bir sayı dizisinin ortasındaki sayıya medyan denir. Sayı dizisi sıralandıktan sonra, varsa ortadaki gerçek sayıdır, yoksa ortadaki iki sayının ortalamasıdır (Nicholas 1990).

Tanım 2.14 Bir sayı dizisinde en çok tekrar eden sayıya mod denir (Nicholas 1990).

Tanım 2.15 Birbirini izleyen yılların, mevsimlerin, ayların günlerin aynı zaman noktalarında zaman serisi gözlem değerlerinde bir artma ve bir azalma şeklinde görülen değişimlerdir (Özmen ve Poyraz 1999)

3. MATERYAL VE METOT

Bu bölümde zaman serisi analizi yapılırken çizilmesi gereken harita ve uygulanacak olan metotlardan bahsedilmiştir.

3.1. Regresyon

İki değişken arasındaki istatistiksel ilişkiye regresyon adı verilir (Akdi 2012). Y bağımlı değişken olmak üzere; x ile isimlendirilen bir veya birden fazla bağımsız değişken ile Y bağımlı değişkeni arasındaki ilişkiyi modelleme yaklaşımına basit (doğrusal) regresyon denir. Açıklayıcı yani x bağımsız değişkeninin birden fazla olması durumunda bu regresyon modeline çoklu doğru regresyon modeli adı verilir (Kılıç 2013).

X ve Y gibi iki değişken arasında $Y = f(X)$ şeklinde bir ilişki, f fonksiyonu biliniyor ise deterministiktir (Akdi 2012).

Gerçek hayatta değişkenlerden bir tanesi sabit tutularak deney birden fazla tekrarlandığında diğer değişkenin değerleri gözlemlenir. Aynı X değeri için birden fazla Y değeri bulunabilir. Yani aslında Y değerleri belli bir hata ile gözlenir. Bu durumda Y ile X arasında

$$Y = f(X) + e$$

gibi bilinmeyen bir f fonksiyonuna bağlı bir ilişkiden söz edilmektedir. Böyle bir ilişkiye stokastik ilişki denir.

Regresyonda önemli olan bazı koşullar altında bilinmeyen f fonksiyonunun belirlenmesidir. Bu koşullar ise genellikle e hata terimi üzerindedir (Akdi 2012).

$X_1, X_2, X_3, \dots, X_p$ değişkenlerinin değerlerinin bilinmesi halinde X_i ler ile Y arasında

$$Y = f(X_1, X_2, \dots, X_p) + e$$

Gibi istatistiksel bir ilişki göz önüne alınırsa Y bağımlı değişkeni rasgele olmak üzere, X_i ler açıklayıcı değişkenlerdir yani reel sayılardır. Bu durumda

$$Y = f(x_1, x_2, \dots, x_p) + e$$

regresyon eşitliğinden bahsedilmektedir (Akdi 2012).

Burada Y bağımlı değişken, x_1, x_2, \dots, x_p ' ler açıklayıcı yani bağımsız değişkenler, e hata terimi, f de bilinmeyen parametre içeren ve regresyonun yapısını gösteren bir fonksiyondur (Akdi 2012).

Regresyonun amacı hataların kareleri en küçük olacak şekilde f fonkiyonunu tahmin etmektir.

x_1, x_2, \dots, x_p ler, X_1, X_2, \dots, X_p değişkenlerinin aldığı değerler olmak üzere, Y ' nin X_1, X_2, \dots, X_p üzerine regresyonu,

$$E(Y|X_1, \dots, X_p = x_p)$$

şeklinde koşullu beklenen değerdir. Bu koşullu beklenen değer bazen doğrusal bazense doğrusal değildir, buda regresyon denklemlerini doğrusal ve doğrusal olmayan regresyon olarak ikiye ayırır (Akdi 2012).

Regresyonda asıl amaç bilinmeyen parametre içeren f fonkiyonunun tahminidir. Bu fonksiyon doğrusal ya da doğrusal olmayan bir fonksiyon olarak karşımıza çıkabilir. Genele bakılırsa genelde doğrusal bir fonksiyon olarak karşımıza çıkmaktadır.

Doğrusal regresyon modeli $x_{i,t}$ ($t = 1,2,3, \dots, n$) ve ($i = 1,2,3, \dots, p$) açıklayıcı değişkenleri, Y_t bağımlı değişkeni göstermek üzere,

$$Y_t = \beta_0 + \beta_1 x_{1,t} + \beta_2 x_{2,t} + \dots + \beta_p x_{p,t} + e_t, t = 1,2, \dots, n$$

olarak verilir. Bu durumda $E(e_t) = 0$, $Var(e_t) = \sigma^2$ ve $t \neq s$ için $Cov(e_t, e_s) = 0$ ise bu denkleme doğrusal regresyon denklemi adı verilir. Bu denklemin basit doğrusal regresyon denklemi olabilmesi için $p = 1$ olması gerekmektedir. Bu durumda böyle bir eşitliğin regresyon denklemi olabilmesi için x açıklayıcı değişkenlerinin bilinen, e_t hata terimlerinin bağımsız aynı dağılıma sahip olan değişkenler olması gerekir. İstatistiki sonuçların çıkarımları bakımından e_t hata terimlerinin normal dağılımlı olduğu varsayılmaktadır. Veriler analiz edilmeden önce bu varsayımların geçerliliği test

edilmeli, varsayımların dışında gelişen durumlarda dönüşümler yapılarak varsayımların geçerliliği sağlanmalıdır. Buradaki amaç model parametrelerinin ($\beta_i, i = 0,1,2, \dots, p$ ve σ^2) tahmin edilmesi ve eldeki verilere uygun modelin belirlenmesidir (Akdi 2012).

Varsayımların geçerliliğini test etmek için farklı yöntemler mevcuttur. Örneğin, hata terimleri kendi aralarında otokorelasyonlu olabilir. Yani ;

$$e_t = \rho e_{t-1} + e_t, |\rho| < 1$$

şeklinde bir ilişki olabilir. Hata terimlerinin otokorelasyonlu olup olmadığı anlamak için Durbin-Watson testi kullanılabilir (Akdi 2012).

3.2. Durbin-Watson Testi

Durbin-Watson testi diğer yöntemlere göre daha çok tercih edilmektedir. Bunun nedeni ise Durbin-Watson'ın ifade ettiği gibi "hatalar bilenemeyeceğine göre test kalıntılara dayanmaktadır." (Genceli 1973).

Durbin-Watson Testi, bir regresyon modelinin tahmini yapıldıktan sonra hata terimlerinin korelasyon halinde olup olmadığını kontrol eder. Durbin-Watson test istatistiğinin 2 civarında çıkması otokorelasyonun olmadığını göstermektedir (Anonim 4).

$\hat{\beta}_i$ 'ler β_i parametrelerinin en küçük kareler tahmin edicilerini göstermek üzere, kestirimler $\hat{Y}_t = \hat{\beta}_0 + \hat{\beta}_1 x_{1,t} + \dots + \hat{\beta}_p x_{p,t}$ ve artıklar da $\hat{e}_t = Y_t - \hat{Y}_t$ şeklinde hesaplanır (Akdi 2012).

Bu durumda Durbin-Watson test istatistiğinin değeri;

$$d = \left[\sum_{t=1}^n \hat{e}_t^2 \right]^{-1} \sum_{t=2}^n (\hat{e}_t - \hat{e}_{t-1})^2$$

formülü ile hesaplanır (Akdi 2012).

Hata terimleri normal dağılımlı ve $t \neq s$ için $Cov(e_t, e_s) = 0$ ise d 'nin payındaki toplamın beklenen değeri yaklaşık olarak $2(n-1)\sigma^2$ dir. Paydadaki toplamın beklenen değeri ise $n\sigma^2$ dir. Sonuç olarak hata terimleri arasında bir

korelasyon ilişkisi yok ise d 'nin beklenen değeri yaklaşık olarak 2 olmalıdır. d 'nin alabileceği en küçük değer ise 0'dır. Durbin-Watson test istatistiği için tablolar düzenlenmiştir. Bu tablo değerlerini kullanarak hata terimlerinin otokorelasyonlu olup olmadığına bakılabilir (Akdi 2012).

Hata terimlerinin normal dağılım gösterip göstermediğini, hata terimlerinin histogramını, kutu çizgilerini ve normal olasılık grafiğini çizerek görebiliriz. Normal olasılık grafiği çizmek için önce regresyondan elde edilen hata terimleri küçükten büyüğe doğru olacak şekilde sıralanmalıdır. Bunlar $\hat{e}_{(t)}$ ile gösterildiğinde normal dağılım varsayımı altında, $E(\hat{e}_{(t)})$ yaklaşık olarak $Z((t - 0.375)/(n + 0.25))$ dir (Neter, Wasserman ve Kutner, 1985 s.118-119). Bu değerler normal dağılım tablosu incelenerek $\hat{e}_{(t)}$ değerlerinin

$$Z((t - 0.375)/(n + 0.25))$$

Tablo değerlerine karşı grafiği çizilir. Grafik çizildikten sonra doğrusallık gözlemleniyor ise hata terimlerinin normal dağılıma uygun olduğundan bahsedilebilir (Akdi 2012).

z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.0000	0.0040	0.0080	0.0120	0.0160	0.0199	0.0239	0.0279	0.0319	0.0359
0.1	0.0398	0.0438	0.0478	0.0517	0.0557	0.0596	0.0636	0.0675	0.0714	0.0753
0.2	0.0793	0.0832	0.0871	0.0910	0.0948	0.0987	0.1026	0.1064	0.1103	0.1141
0.3	0.1179	0.1217	0.1255	0.1293	0.1331	0.1368	0.1406	0.1443	0.1480	0.1517
0.4	0.1554	0.1591	0.1628	0.1664	0.1700	0.1736	0.1772	0.1808	0.1844	0.1879
0.5	0.1915	0.1950	0.1985	0.2019	0.2054	0.2088	0.2123	0.2157	0.2190	0.2224
0.6	0.2257	0.2291	0.2324	0.2357	0.2389	0.2422	0.2454	0.2486	0.2517	0.2549
0.7	0.2580	0.2611	0.2642	0.2673	0.2704	0.2734	0.2764	0.2794	0.2823	0.2852
0.8	0.2881	0.2910	0.2939	0.2967	0.2995	0.3023	0.3051	0.3078	0.3106	0.3133
0.9	0.3159	0.3186	0.3212	0.3238	0.3264	0.3289	0.3315	0.3340	0.3365	0.3389
1.0	0.3413	0.3438	0.3461	0.3485	0.3508	0.3531	0.3554	0.3577	0.3599	0.3621
1.1	0.3643	0.3665	0.3686	0.3708	0.3729	0.3749	0.3770	0.3790	0.3810	0.3830
1.2	0.3849	0.3869	0.3888	0.3907	0.3925	0.3944	0.3962	0.3980	0.3997	0.4015
1.3	0.4032	0.4049	0.4066	0.4082	0.4099	0.4115	0.4131	0.4147	0.4162	0.4177
1.4	0.4192	0.4207	0.4222	0.4236	0.4251	0.4265	0.4279	0.4292	0.4306	0.4319
1.5	0.4332	0.4345	0.4357	0.4370	0.4382	0.4394	0.4406	0.4418	0.4429	0.4441
1.6	0.4452	0.4463	0.4474	0.4484	0.4495	0.4505	0.4515	0.4525	0.4535	0.4545
1.7	0.4554	0.4564	0.4573	0.4582	0.4591	0.4599	0.4608	0.4616	0.4625	0.4633
1.8	0.4641	0.4649	0.4656	0.4664	0.4671	0.4678	0.4686	0.4693	0.4699	0.4706
1.9	0.4713	0.4719	0.4726	0.4732	0.4738	0.4744	0.4750	0.4756	0.4761	0.4767
2.0	0.4772	0.4778	0.4783	0.4788	0.4793	0.4798	0.4803	0.4808	0.4812	0.4817
2.1	0.4821	0.4826	0.4830	0.4834	0.4838	0.4842	0.4846	0.4850	0.4854	0.4857
2.2	0.4861	0.4864	0.4868	0.4871	0.4875	0.4878	0.4881	0.4884	0.4887	0.4890
2.3	0.4893	0.4896	0.4898	0.4901	0.4904	0.4906	0.4909	0.4911	0.4913	0.4916
2.4	0.4918	0.4920	0.4922	0.4925	0.4927	0.4929	0.4931	0.4932	0.4934	0.4936
2.5	0.4938	0.4940	0.4941	0.4943	0.4945	0.4946	0.4948	0.4949	0.4951	0.4952
2.6	0.4953	0.4955	0.4956	0.4957	0.4959	0.4960	0.4961	0.4962	0.4963	0.4964
2.7	0.4965	0.4966	0.4967	0.4968	0.4969	0.4970	0.4971	0.4972	0.4973	0.4974
2.8	0.4974	0.4975	0.4976	0.4977	0.4977	0.4978	0.4979	0.4979	0.4980	0.4981
2.9	0.4981	0.4982	0.4982	0.4983	0.4984	0.4984	0.4985	0.4985	0.4986	0.4986
3.0	0.4987	0.4987	0.4987	0.4988	0.4988	0.4989	0.4989	0.4989	0.4990	0.4990

Şekil 3.1. Standart normal dağılım tablosu

3.3. Zaman Serisi Regresyonu

Zaman serilerinde trend karşımıza deterministik bir yapıda çıkabilir. Bu durumda zaman serisinde öngörü yapmak için, zaman serisinin trendine uygun olan bir matematiksel form bulunması gerekir. Bu matematiksel formlara karar verebilmek için hataya dayalı bazı istatistiklere başvurulmaktadır. Bu başlıkta deterministik trend yapılarının yardımı ile mevsimsel olmayan zaman serisi regresyonunda eğri uydurma ve öngörü metotlarından bahsedilecektir (Mert ve Çağlar 2019).

3.3.1. Basit doğrusal regresyon modeli

Doğrusal bir trende sahip olan zaman serileri için uygun olan regresyon

modelidir (Mert ve Çağlar 2019).

$$y_t = a + bt$$

denklemini kullanılmaktadır.

Eğer denklem doğrusal bir trende sahip ise $y_t = a + bt$ denklemi en küçük kareler yöntemi ile tahmin edilir ve öngörü değerleri elde edilir (Mert ve Çağlar 2019).

3.3.2. Birinci farklar regresyon modeli

Eğer zaman serisinin trendi doğrusal olmayan bir biçimde ise birinci farklar regresyon modeli kullanılabilir. Bu modelde tahmin edilecek olan denklem

$$\Delta y_t = a + bt$$

şeklindedir ve bu denklemde bağımlı değişken zaman serisinin birinci farkı olmaktadır ($\Delta y_t = y_t - y_{t-1}$) (Mert ve Çağlar 2019).

3.3.3. Üstel regresyon modeli

Üstel regresyon modeli, birinci farklar regresyon modelinin bir alternatifi olarak karşımıza çıkmaktadır, Büyüme modeli olarakta bilinmektedir. Modelin denklemi

$$\ln(y_t) = a + bt$$

şeklindedir ve bu denklemde bağımlı değişken zaman serisinin doğal logaritmasıdır (Mert ve Çağlar 2019).

3.3.4. Karesel regresyon modeli

Birinci farklar regresyon modeli ve üstel regresyon modelinin bir benzeride karesel regresyon modelidir. Karesel regresyon modelinde trendin artış hızı üstel modele göre daha yavaştır. Modelin denklemi

$$y_t = a + bt + ct^2$$

şeklindedir (Mert ve Çağlar 2018).

3.3.5. Lojistik regresyon modeli

Lojistik regresyon modeli, trendi doğrusal olmayan formlar için kullanılabilir.

Model denklemi

$$\ln\left(\frac{L}{y_t} - 1\right) = a + bt$$

şeklindedir ve bu denklemde L değeri serideki en büyük gözlem değerinden daha büyük olan keyfi bir değerdir (Mert ve Çağlar 2019).

3.3.6. Kübik regresyon modeli

Kübik regresyon modeli trend olarak lojistik regresyon modeline benzemektedir ancak serinin son gözlem değerlerinde trend aşağıya doğru yönelme eğilimine girmektedir. Kübik regresyon modeli

$$y_t = a + bt + ct^2 + dt^3$$

şeklindedir (Mert ve Çağlar 2019).

3.3.7. Logaritmik regresyon modeli

Logaritmik regresyon denklemi

$$y_t = a + b \cdot \ln(t)$$

şeklindedir ve bu denklemde bağımsız değişken, t değişkeninin doğal logaritmasıdır.

3.3.8. Güç regresyon modeli

Güç regresyon modeli, birinci farklar, üstel ve karesel regresyon modellerinin bir alternatifi olarak kullanılmaktadır. Güç regresyon modelinin denklemi

$$\ln(y_t) = a + b \cdot \ln(t)$$

şeklindedir.

3.3.9. S regresyon modeli

S regresyon modeli, kübik ve lojistik regresyon modellerinin bir alternatifidir. S

regresyon modeli

$$\ln(y_t) = a + b \left(\frac{1}{t} \right)$$

şeklindedir ve bu denklemde zaman serisinin doğal logaritmik değerleri, bağımlı değişken olarak kullanılmaktadır (Mert ve Çağlar 2019).

3.3.10. Ters regresyon modeli

Ters regresyon modeli denklemi

$$y_t = a + b \left(\frac{1}{t} \right)$$

şeklindedir ve logaritmik regresyon modelinin alternatifidir.

Yukarıdaki modeller seriye uygulanarak en uygun model kullanılabilir (Mert ve Çağlar 2019).

3.4. Trend ve Mevsimsellik

Zaman serisi analizi, zamana bağlı olayları ölçmek adına önemli bir buluştur. Zaman serisi analizi yapmak için en önemli şey veri setidir. Veri setinin ise zaman serisi analizine uygun olması beklenir.

Verilerin elde edilmesinin bir yolu gözlemdir. Gözlem aşamasında bir ölçme işlemi gerçekleşir. Bu ölçme işleminde ana kütle ya tam ölçülür ya da ana kütlede bir örneklem seçilerek ölçme işlemine devam edilir. Genelde ölçme işlemleri örneklem seçilerek yapılmaktadır. Bunun nedeni ise ana kütle ile ölçüm yapmak hem vakit almasından dolayı hem de zor bir süreç olmasından dolayı kaynaklanır (Sevüktekin ve Çınar 2014).

Zaman serisi verilerinde en önemli durum veri setinin zamana bağlı olmasıdır. Verinin düzenli ve art arda gelen zaman aralıklarında alınması gerekmemeyle birlikte düzenli ve art arda gelen zaman aralıklarından alınması daha doğru sonuçlar verecektir. Ayrıca bu zaman aralığı günlük, haftalık, aylık, yıllık seçilebilir (Mert 2018).

Zaman serisinde trend önemli bir yer kaplamaktadır. Zaman serilerinde trend doğrusal ve ya doğrusal olmayan şekilde karşımıza çıkabilir.

Zaman serisi analizi yapılırken öncelikle mevsimsellik incelemesi yapılır. Mevsimsellik incelemesi yıllık verilerden daha çok aylık ve haftalık verilerde karşımıza çıkmaktadır (Mert 2018).

Mevsimsel zaman serilerinde, uygun fark ve dönüşümün belirlenmesi önceliklidir. Örnek gösterirsek yaz aylarında dondurma tüketimi artış göstermektedir. Bu sebeple ağustos ayında tüketilen dondurma ile bir yıl önceki ağustos ve temmuz aylarında tüketilen dondurma miktarı birbiri ile ilişkili olabilir (Akdi 2012).

$\varepsilon_t \sim WN(0, \sigma^2)$ olmak üzere

$$X_t - \mu = \alpha(X_{t-12} - \mu) + e_t, \quad t = 1, 2, 3 \dots, n \quad (3.1)$$

Zaman serisi modelini inceleyelim. (3.1) de ki model aylık olarak alınan veri setine uygulanabilir.

$\mu = 100$ ve $\alpha = 0.8$ için en son aralık ayına ait gözlenen değer $X = 120$ ise Aralık ayının öngörülen değeri

$$100 + 0,8(120 - 100) = 116$$

şeklinde alınabilir.

Bir sonraki aralık ayı için öngörü tahmini ise

$$100 + 0,8^2(120 - 100) = 112,8 \text{ (Akdi 2012).}$$

şeklinde olur.

Bu öngörülerin benzeri, diğer aylar içinde uygulanabilir.

Bu model üzerinden incelersek, k.ncı yılın aralık ayı için öngörü tahminini

$$100 + 0,8^k(120 - 100) \text{ (Akdi 2012).}$$

şeklinde hesaplayabiliriz.

Bu model dikkate alındığında ise geçen yılın aralık ayına ait öngörünün bu sene ki aralık ayını etkilediği gözlemlenmiştir.

(3.1) Numaralı modelde $\sigma = 1$ seçilirse, serinin μ parametresi yani beklenen değeri modelden düşecektir ve model

$$X_t = X_{t-12} + e_t \quad (3.2)$$

şeklinde olur (Akdi 2012).

(3.1) Numaralı modele karşılık gelen karakteristik denklem

$$m^{12} - 1 = 0$$

şeklinde ve bu karakteristik denklemin kökleri mutlak değerce 1'dir.

Bu durumda

$$Z_t = (1 - B^{12})X_t \quad (3.3)$$

dönüşümü ile seri durağan hale getirilir (Akdi 2012).

Zaman serisi analizlerinde seriyi oluşturan bileşenlerin elde edilmesi için serinin bileşenlerinden ayrılması durumu ortaya çıkabilir. Özellikle mevsimselliğin görüldüğü serilerde, serinin mevsimsellikten arındırılarak analize devam edilmesi önemlidir (Mert 2018).

3.5. Durağanlık

Zaman serisi analizi yapılırken incelenmesi gereken bir diğer kavram ise durağanlık kavramıdır. Durağanlık kavramı veri setinin geçirdiği bir şok karşısında gösterdiği dirençtir. Yani veri seti olağan dışı bir durum ile karşılaşılırsa ve bu durum uzun süreli olan veri setinin varyansını ve ortalamasını etkilemiyorsa bu veri seti için durağandır diyebiliriz. Karşılaşılan olağanüstü durumlarda veri seti direnç göstermiyorsa, uzun süreli bakıldığında ortalamasında ve varyansın da bir değişiklik

meydana geliyor ise bu veri seti için durağan değildir diyebiliriz. Başka bir tabirle ise veri setinde birim kök bulunurda diyebiliriz. Özetle bir serinin durağan olması o serinin birim kök içermediğini gösterirken, durağan dışı olması ise seride birim kökün varlığını gösterir (Mert 2018).

Zaman serisi analizlerinde ise durağanlık, zayıf durağanlık ve kovaryans durağanlık olarak karşımıza çıkmaktadır. Bir y_t zaman serisine durağandır diyebilme için 3 koşulun sağlanması gerekmektedir. Bu koşullar aşağıdaki şekilde sıralanmıştır (Mert 2018).

$$E(y_t) = \mu \quad (3.4)$$

$$Var(y_t) = \sigma^2 \quad (3.5)$$

$$Cov(y_t, y_{t-k}) = \rho_k \quad (3.6)$$

(3.4) eşitliğinde zaman serisinin, zaman süresince ortalamasının değişkenlik göstermediğini, (3.5) eşitliğinde zaman serisinin varyansının, zaman süresince sabit kaldığını ve (3.6) eşitliğinde ise serinin kendi gecikmeleriyle olan kovaryansın zamandan bağımsız hareket edip gecikme uzunluğuna bağlı olarak değiştiğini göstermektedir. (3.7) denkleminde ki model $\phi = 1$ için sabitsiz ve trendsiz bir rassal yürüyüş süreci olarak tanımlanır (Mert 2018).

$$y_t = \phi_{t-1} + \varepsilon_t \quad (3.7)$$

(3.7) eşitliğinde $\phi = 1$ için rassal yürüyüş sürecinin durağan dışı bir süreç olduğu görülebilir. Her bir t zamanı için süreç aşağıdaki gibi olacaktır (Mert 2018).

$$y_1 = \varepsilon_1$$

$$y_2 = y_1 + \varepsilon_2 = \varepsilon_1 + \varepsilon_2$$

$$y_3 = y_2 + \varepsilon_3 = \varepsilon_1 + \varepsilon_2 + \varepsilon_3$$

$$y_t = y_{t-1} + \varepsilon_t = \varepsilon_1 + \varepsilon_2 + \varepsilon_3 + \dots + \varepsilon_t$$

Dolayısı ile $E(y_t) = E(\varepsilon_1 + \varepsilon_2 + \varepsilon_3 + \dots + \varepsilon_t) = E(\varepsilon_1) + \dots + E(\varepsilon_t) = 0$

eşitliği elde edilir.

Böylece (3.4) eşitliğinde gösterilen durağanlığın ilk koşulu sağlanmış olur.

(3.5) eşitliğinde;

$$Var(y_t) = Var(\varepsilon_1 + \varepsilon_2 + \varepsilon_3 + \dots + \varepsilon_t) = Var(e_1) + \dots + Var(e_t) = t\sigma^2$$

eşitliği elde edilir. Bu eşitlikte $t\sigma^2$ varyansı t zamandan bağımsız olmadığından dolayı (3.5) eşitliğinde verilen ikinci koşulun sağlanamadığı görülmektedir. Yani rassal yürüyüş sürecinin durağan dışı olduğundan bahsedebiliriz. Bu durumda (3.4) denkleminde $\phi = 1$ olması yani birim kökün varlığı rassal yürüyüş sürecinin durağan dışı olduğunu gösterecektir. Bu durumda $\phi = 1$ olması durumunda bu denklemde durağanlık sağlanacaktır. Özetle birim kök testlerindeki genel uygulama serinin rassal yürüyüş süreci ile karşılaştırılarak $\phi = 1$ yokluk hipotezinin (birim kökün varlığı), $\phi < 1$ alternatif hipotezine (birim kökün yokluğu) karşı dayanmaktadır (Mert, 2018).

3.6. Kesikli ve Sürekli Zaman Serileri

Eğer zaman serilerinde incelenen gözlem değerleri devamlı olarak elde edilebiliyor ise ortaya çıkan seri için sürekli zaman serisidir diyebiliriz. Sürekli zaman serisi genel olarak zamanda eşit olmayan aralıklarda alınan gözlem değerlerinden meydana gelir. Bundan farklı olarak eğer gözlem değerleri belli zaman aralıklarında elde edilmişse bu tür zaman serilerine ise kesikli zaman serileri denir.

İncelenen zaman serilerinin gözlem değerleri zaman içinde devamlı olarak elde ediliyorsa, meydana gelen seri sürekli zaman serisidir. Bu tür seriler genellikle zaman içinde eşit olmayan aralıklarla elde edilen gözlem değerlerinden oluşur. Eğer gözlem sadece belirli zaman aralıkları ile yapılıyorsa, böyle serilere kesikli zaman serileri denir. Kesikli zaman serileri genellikle eşit zaman aralıklarıyla yapılan gözlem değerlerinden oluşur. Uygulamada en çok üzerinde çalışılan zaman serileri kesikli zaman serileridir (Duru 2007).

3.7. Arttırılmış Dickey-Fuller Testi (The Augmented Dickey-Fuller Test)

Arttırılmış Dickey-Fuller testi durağanlığın tespiti için kullanılabilir. Büyük ve karmaşık veri setleri için arttırılmış Dickey-Fuller testi kullanıma uygundur. Bir çok birim kök testi olmasına karşın Arttırılmış Dickey-Fuller testi genel anlamda tercih edilenler arasında öne çıkmaktadır (Roy vd. 2018).

Dickey-Fuller testi (3.7) de verilen eşitlikte rassal yürüyüş sürecinin birinci farkından yararlanılır ve (3.7) eşitliğindeki denklemin birinci farkı için eşitliğin her iki tarafından da y_{t-1} çıkartılır ve aşağıdaki eşitlik elde edilir (Mert ve Çağlar 2019).

$$y_t - y_{t-1} = \phi y_{t-1} - y_{t-1} + \varepsilon_t$$

$$\Delta y_t = (\phi - 1)y_{t-1} + \varepsilon_t$$

olmak üzere eşitlik aşağıdaki gibi elde edilir.

$$\Delta y_t = \delta y_{t-1} + \varepsilon_t \quad (3.8)$$

(3.8) eşitliğinden elde edilen model sabitsiz ve trendi olmayan bir modeldir. Sabitli model

$$\Delta y_t = \mu + \delta y_{t-1} + \varepsilon_t \quad (3.9)$$

şeklinde verilmiştir.

Trendli ve sabitli model ise;

$$\Delta y_t = \mu + \beta_t + \delta y_{t-1} + \varepsilon_t \quad (3.10)$$

şeklinde verilir.

Bu durumda birim kökün varlığından bahsedilmek için $\delta = 1 - \phi = 0$ olmalıdır. Aynı zamanda durağanlıktan söz edebilmek içinse $\delta = \phi - 1 < 0$ olur. (Mert ve Çağlar 2019).

Standart Dickey-Fuller testi için hipotezler aşağıdaki gibi kurulabilir. (Mert ve Çağlar 2019).

$H_0: \delta = 0$ (seri durağan değildir, birim kök vardır)

$H_1: \delta < 0$ (seri durağandır ve birim kök yoktur)

kurulan hipotezde test istatistiği $t_\delta = \hat{\delta}/S_{\hat{\delta}}$ olmaktadır.

Standart Dickey-Fuller testi (3.8), (3.9) ve (3.10) denklemlerinde görüldüğü üzere AR(1) sürecinden yararlanmaktadır. Eğer ki seride yüksek dereceden bir korelasyon mevcutsa ε_t temiz dizi olma özelliğini yitirecektir. Arttırılmış Dickey-Fuller testi ise bu sorunu halledilmek için AR(p) sürecinden faydalanmıştır ve denkleme p gecikmeli fark terimleri eklemiştir. Bunun sonucunda sabitsiz ve trendli olmayan, sabitli, trendli ve sabitli, arttırılmış Dickey-Fuller denklemleri aşağıdaki gibi verilmiştir (Mert ve Çağlar 2019).

$$\Delta y_t = \delta y_{t-1} + \sum_{i=1}^p \beta_i \Delta y_{t-i} + \varepsilon_t \quad (3.11)$$

$$\delta y_{t-1} + \sum_{i=1}^p \beta_i \Delta y_{t-i} + \varepsilon_t \quad (3.12)$$

$$\Delta y_t = \mu + \beta_t + \delta y_{t-1} + \sum_{i=1}^p \beta_i \Delta y_{t-i} + \varepsilon_t \quad (3.13)$$

(3.11), (3.12) ve (3.13) denklemlerinde birim kökün varlığını araştıran hipotezler aynıdır ve t_δ istatistiğinin asimptotik dağılımı denkleme eklenen gecikmeli fark terimlerinden bağımsızdır. Test istatistiği ilgili kritik değerden eğer daha küçük ise H_0 hipotezi reddedilecektir (Mert ve Çağlar 2019).

3.8. Phillips Perron Testi

Phillips-Perron testi Dickey-Fuller testinden farklı olarak hata terimindeki ardışık ilişkiyi de varsaymak için gecikmeli fark değerlerini eklemeyen parametrik olmayan istatistiksel yöntemleri kullanır. Dickey-Fuller testi ise hata terimlerinin arasında otokorelasyonun olmadığını varsaymaktadır. PP testinde zaman serisinin hata terimleri arasında parametrik olmayan düzeltme kullanılır ve bu düzeltme sonucunda otokorelasyon sorunu ortadan kalkmaktadır. Bu durumda özellikle belli bir trende sahip

olan veri setlerinde PP testi DF testine göre daha güçlüdür. PP testinde kullanılan başlıca denklemleri şu şekilde verebiliriz:

$$\text{Sabit terimsiz model: } \Delta Y_t = \beta(t - \frac{1}{2} T) + \rho Y_{t-1} + \varepsilon_t$$

$$\text{Sabit terimli model: } \Delta Y_t = \beta_0 + \beta(t - \frac{1}{2} T) + \rho Y_{t-1} + \varepsilon_t$$

$$\text{Sabit terimli ve trendli model: } \Delta Y_t = \beta_0 + \beta_t + \beta(t - \frac{1}{2} T) + \rho Y_{t-1} + \varepsilon_t$$

Bu modellerde T gözlem değerini göstermektedir ve hata terimlerinin beklenen değeri ise sifıra eşittir. Hipotezler şu şekilde kurulabilir;

$$H_0: \rho = 1 \text{ ise seri durağan değildir.}$$

$$H_1: \rho < 1 \text{ ise seri durağandır.}$$

DF ve ADF testi için kullanılan Mackinnon kritik tablo değerleri, PP test istatistiği ile karşılaştırılarak hipotezler kabul ya da reddedilebilir (Sevüktekin ve Nargeçekenler 2007; Şenesen ve Şenesen 2012; Kaya 2018).

3.9. ARIMA Modelleri

Doğrusal zaman serisi analizlerinden olan ARIMA metodolojisi araştırmacılar tarafından sıklıkla kullanılan öngörü yöntemlerinden biridir. Analiz edilen serideki birim kökün varlığı ya da yokluğu bu modellerde çok fazla önem kazanmaktadır. Dolayısı ile ARIMA analizinin ilk aşaması birim kök testidir. ARIMA modellerinin uygulanabilmesi için veri setinin tutarlı olması gerekmektedir (Roy vd. 2018).

3.9.1. AR(p) modeli

Bir seri kendi gecikmeli değerlerinden etkileniyorsa otoregresif bir seridir denir. Eğer durağan bir $y_{\{t\}}$ zaman serisi, sadece bir dönem gecikmesinden etkileniyorsa birinci dereceden otoregresiftir. ($p = 1$) ve AR(1) modelinin kapalı formu

$$(1 - \rho L)(y_t - \mu) = \varepsilon_t \quad (3.14)$$

şeklindedir.

Burada;

L : gecikme işlemcisi

ρ : AR(1) parametresi

μ : sabit terim

$\varepsilon_t \sim N(0, \sigma^2)$ olan temiz dizidir. (3.14)'deki denklemin açık hali aşağıdaki gibi yazılabilir.

$$y_t - \mu - \rho Ly_t + \rho\mu = \varepsilon_t$$

$$y_t = \mu(1 - \rho) + \rho y_{t-1} + \varepsilon_t$$

$c = \mu(1 - \rho)$ sabit olmak üzere AR(1) denkleminin açık formuna (3.15)'deki gibi ulaşılır.

$$y_t = c + \rho y_{t-1} + \varepsilon_t \quad (3.15)$$

Eğer durağan zaman serisi hem bir dönem hem de iki dönem gecikmeli değerlerinden etkileniyorsa ($p = 2$) ikinci dereceden otoregresiftir denir. Bu durumda AR(2) modelinin kapalı formu;

$$(1 - \rho^1 L - \rho^2 L^2)(y_t - \mu) = \varepsilon_t \quad (3.16)$$

şeklindedir.

(3.16) 'da gösterilen AR(2) modelinin kapalı formu aşağıdaki gibi açılabilir;

$$y_t - \mu - \rho_1 Ly_t + \rho_1 \mu - \rho_2 L^2 y_t + \rho_2 \mu = \varepsilon_t$$

$$y_t = \mu - \rho_1 \mu - \rho_2 \mu + \rho_1 y_{t-1} + \rho_2 y_{t-2} + \varepsilon_t$$

$$c = \mu(1 - \rho^1 - \rho^2)$$

sabiti olmak üzere AR(2) modelinin açık formuna (3.17)'teki gibi ulaşılır.

$$y_{\{t\}} = c + \rho_1 y_{\{t-1\}} + \rho_2 y_{\{t-2\}} + \varepsilon_{\{t\}} \quad (3.17)$$

(4)'te ρ_1 ve ρ_2 sırası ile AR(1) ve AR(2) katsayıları olmaktadır. Buradan hareketle genelleştirilirse; durağan bir zaman serisi için p. dereceden otoregresif AR(p) model denkleminin kapalı formu (3.18)'teki gibi olacaktır.

$$\left(1 - \sum_{\{j=1\}}^p p_j L^j\right) (y_t - \mu) = \varepsilon_t \quad (3.18)$$

(3.18) denklemi açılır ise AR(p) denklemi

$$(1 - \sum_{j=1}^p p_j L^j) (y_t - \mu) = \varepsilon_t \quad (3.19)$$

$c = \mu - \sum_{j=1}^p p_j \mu$ sabiti olmak üzere

$$y_t = c + \rho_1 y_{t-1} + \dots + \rho_p y_{t-p} \quad (3.20)$$

Bu denklemde modelin açık formu görülmektedir ve $\rho_p, p = 1, 2, \dots, p$ AR katsayıları olmaktadır (Mert ve Çağlar 2019).

3.9.2. MA(q) modeli

Bir zaman serisi, otoregresif süreçler gibi kendi geçmiş değerlerinden değil de rassal şoklardan etkilenebilir. Böyle serilere hareketli ortalama serileri denir. Eğer durağan y_t zaman serisi (3.21)'de görüldüğü gibi, bir dönem önceki rassal şoklardan etkileniyorsa birinci dereceden hareketli ortalama serisidir, ($q = 1$) ve MA(1) ile gösterilmektedir.

$$y_t - \mu = (1 + \theta L) \varepsilon_t \quad (3.21)$$

(3.16)'da verilen MA(1) kapalı formu aşağıdaki gibi açılabilir.

$$y_t - \mu = \varepsilon_t + \theta L \varepsilon_t$$

$$y_t = \mu + \theta \varepsilon_{t-1} + \varepsilon_t \quad (3.22)$$

(3.21) ve (3.22)'de μ :sabit terim , θ : $MA(1)$ katsayısı ve $\varepsilon_t \sim N(0, \sigma^2)$ olan temiz dizi özelliğinde hata terimi olmaktadır.

Eğer durağan bir zaman serisi hem bir dönem hemde iki dönem önceki rassal şoklardan etkileniyorsa ($q = 2$) ikinci dereceden hareketli ortalama serisidir denir. Bu durumda $MA(2)$ modelinin kapalı denklemi (3.23)'deki gibi elde edilir.

$$y_t - \mu = (1 + \theta_1 L + \theta_2 L^2) \varepsilon_t \quad (3.23)$$

(3.23)'de verilen $MA(2)$ model denklemi aşağıdaki gibi açılacaktır.

$$y_t - \mu = \varepsilon_t + \theta_1 L \varepsilon_t + \theta_2 L^2 \varepsilon_t$$

$$y_{\{t\}} = \mu + \varepsilon_t + \theta_1 L \varepsilon_t + \theta_2 L^2 \varepsilon_t \quad (3.24)$$

(3.21)'de θ_1 ve θ_2 sırasıyla $MA(1)$ ve $MA(2)$ katsayıları olmaktadır. Buradan hareketle genelleştirilirse; durağan bir zaman serisinin q . dereceden hareketli ortalama $MA(q)$ model denklemi kapalı formda (3.25)'deki gibi olacaktır (Mert ve Çağlar 2019).

$$y_t - \mu = \left(1 + \sum_{\{j=1\}}^q \theta_j L^j \right) \varepsilon_{\{t\}} \quad (3.25)$$

(3.25) eşitliğinin kapalı formunun açılımı;

$$y_t - \mu = \varepsilon_t + \sum_{j=1}^q \theta_j L^j \varepsilon_t$$

$$y_t - \mu = \varepsilon_t + \theta_1 L \varepsilon_t + \dots + \theta_q L^q \varepsilon_t$$

$$y_t = \mu + \theta_1 \varepsilon_{t-1} + \dots + \theta_q \varepsilon_{t-q} + \varepsilon_t$$

eşitliğinde $\theta_q, q = 1, 2, \dots, q$ MA katsayıları olmaktadır.

3.9.3. ARMA(p,q) modeli

Bir zaman serisi hem otoregresif hem de hareketli ortalama serisi yapısında olabilir. Böyle serilere otoregresif hareketli ortalama serileri denilmektedir. Bu durumda "p" AR parametrelerinin, "q" ise MA parametrelerinin denklemlerle sayılarını verecektir. Örneğin durağan y_t zaman serisi ARMA(1,1) sürecine sahipse model denklemleri kapalı formda (3.26)'deki gibidir.

$$(1 - \rho L)(y_t - \mu) = (1 + \theta L)\varepsilon_t \quad (3.26)$$

μ :sabit terim

ρ : AR(1) parametresi,

θ :MA(1) parametresi ve

$\varepsilon_t \sim N(0, \sigma^2)$ temiz dizi özelliğinde hata terimi olmaktadır.

(3.26)'daki kapalı denklemin açılımı aşağıdaki gibi yapılmaktadır:

$$(y_t - \mu - \rho L y_t + \rho \mu) = \varepsilon_t + \theta L \varepsilon_t$$

$$y_t = \mu - \rho \mu + \rho y_{t-1} + \theta \varepsilon_{t-1} + \varepsilon_t$$

$$c = \mu - \rho \mu$$

olmak üzere, ARMA(1,1) modelinin açık hali (3.27)'de gösterilmektedir.

$$y_t = c + \rho y_{t-1} + \theta \varepsilon_{t-1} + \varepsilon_t \quad (3.27)$$

Örneğin durağan y_t zaman serisi ARMA(2,1) sürecine sahipse model denklemleri kapalı formda (3.28)'teki gibi olacaktır.

$$(1 - \rho^1 L - \rho^2 L^2)(y_t - \mu) = (1 + \theta L)\varepsilon_t \quad (3.28)$$

(3.28)'in açılımı aşağıdaki gibi olacaktır:

$$y_t - \mu - \rho_1 L y_t + \rho_1 \mu - \rho_2 L^2 y_t + \rho_2 \mu = \varepsilon_t + \theta L \varepsilon_t$$

$$y_t = \mu - \rho_1 \mu - \rho_2 \mu + \rho_1 y_{t-1} + \rho_2 y_{t-2} + \theta \varepsilon_{t-1} + \varepsilon_t$$

$$c = \mu - \rho_1 \mu - \rho_2 \mu$$

olmak üzere $ARMA(2,1)$ modelinin açık formu (3.29)'da ki gibidir.

$$y_t = c + \rho_1 y_{t-1} + \rho_2 y_{t-2} + \theta \varepsilon_{t-1} + \varepsilon_t \quad (3.29)$$

Bu örneklerden hareketle genelleştirme yapılırsa $ARMA(p,q)$ için model denkleminin kapalı formu (3.30)'teki gibi olacaktır.

$$\left(1 - \sum_{\{j=1\}}^{\rho} \rho_j L^j\right) (y_{\{t\}} - \mu) = \left(1 + \sum_{\{j=1\}}^q \theta_j L^j\right) \varepsilon_t \quad (3.30)$$

(3.30) de verilen deklemin kapalı formu ise;

$$y_t - \sum_{\{j=1\}}^{\rho} \rho_j L^j y_t - \mu + \sum_{\{j=1\}}^{\rho} \rho_j \mu = \varepsilon_t + \sum_{j=1}^q \theta_j L^j \varepsilon_t$$

$$y_t = c + \rho_1 y_{t-1} + \dots + \rho_p y_{t-p} + \theta_1 \varepsilon_{t-1} + \dots + \theta_q \varepsilon_{t-q} + \varepsilon_t \quad (3.31)$$

(3.31)'de verilen denklemde $c = \mu - \sum_{j=1}^p \rho_j \mu$ sabiti, ρ_p , $p=1,2,\dots,p$ AR katsayılarını ve θ_q , $q = 1,2,\dots,q$ MA katsayılarını göstermektedir (Mert ve Çağlar 2019).

3.9.4. ARIMA(p,d,q) modeli

ARMA modellerinde zaman serisi $\{y_t\}$ 'nin düzeyde durağan olma zorunluluğu vardır. Başka bir deyişle zaman serisi $I(0)$ olmalıdır. Ancak zaman serisi düzeyde durağan değil d. Dereceden entegre (tümleşik) bir seri ise (yani $I(d)$ ise) kullanılacak model $ARIMA(p,d,q)$ modeli olacaktır. Dolayısı ile d tümleşme derecesini, p modeldeki AR gecikme sayısını, q ise MA gecikme sayısını gösterecektir. Dolayısı ile

seriyi modellemeden önce tümleşme derecesine karar vermek için birim kök testi yapılmalıdır. Birim kök testi sonucuna göre seri düzeyde durağan ise uygun model $ARMA(p, q) = ARIMA(p, 0, q)$ modeli olacak ancak serinin tümleşme derecesi d ise uygun model $ARIMA(p, d, q)$ modeli olacaktır. Sonuç olarak; seri birinci dereceden entegre ise (yani $I(1)$ ise) serinin düzey değerleri değil, birinci fark değerleri, ikinci dereceden entegre ise yani ($I(2)$ ise) ikinci fark değerleri ve sonuçta d . dereceden entegre ise (yani $I(d)$ ise) d . fark değerleri denklemlere alınacaktır. En genel haliyle $ARIMA(p, d, q)$ modelinin kapalı formu (3.33)'da gösterilmiştir (Mert ve Çağlar 2019).

$$(1 - L)^d \left(1 - \sum_{j=1}^p \rho_j L^j\right) (y_t - \mu) = \left(1 + \sum_{j=1}^q \theta_j L^j\right) \quad (3.32)$$

(3.30)'da $(1 - L)^d$: fark polinomudur. Örneğin seri düzeyde durağan ise bu polinom $d = 0$ olacağından 1'e eşit olacaktır. Bu durumda ise (3.32)'deki $ARIMA$ modeli (3.31)'te verilen

$$ARIMA(p, 0, q) = ARMA(p, q)$$

modeline dönüşecektir.

3.9.5. ARFIMA (p,d,q) modeli

Durağanlık kavramı, bir zaman serisinin karşılaştığı şok karşısında gösterdiği direnç ile açıklanmıştır. Uygulanan birim kök testleri $I(0)$ ise seri karşılaştığı şoklara karşı dirençli yani durağan, $I(1)$ ise karşılaştığı şoklar karşısında dirençsiz yani durağan olmayan bir seridir. Bu yaklaşımda; serinin şoklara karşı göstermiş olduğu direnç tümleşme derecesinin 0 ya da 1 olmasına göre ölçülmektedir. Ancak bazı zaman serilerinde özellikle de finansal veri setlerinde, veri setinin karşılaştığı şok uzun süreli olmasa bile verinin tekrardan ortalamasına dönmesi uzun sürmektedir. Bu durumda kesirli tümleşme yaklaşımıyla, tümleşme derecesinin 0 ve 1 değerleri arasında, tam sayı dışında da değerler almasına olanak verilerek durağanlık kavramı dışında serinin uzun hafızalı olup olmadığı incelenir. Bu inceleme ARFIMA modelleri ile tümleşme değerleri tahmin edilerek gerçekleşir. Serinin uzun hafızalı olması, veri setinin çok eski gözlemlerden bile hala etkileniyor olması demektir. Bu durumdaki veri setlerinin iyi

modellenebilmesi uzun süreli tahminler yapabilmeye olanak sağlar (Mert ve Çağlar 2019).

ARFIMA modeli için kesirli tümleşme derecesi Granger ve Joyeux (1980), Hosking (1981) tarafından (3.34) numaralı denklem yardımıyla elde edilmektedir (Mert ve Çağlar 2019).

$$\nabla^d = (1 - L)^d = \sum_{k=0}^{\infty} \binom{d}{k} (-1)^k L^k = \sum_{k=0}^{\infty} \frac{\Gamma(-d+k)}{\Gamma(-d)\Gamma(k+1)} L^k \quad (3.33)$$

(3.33) numaralı eşitlikte $-\frac{1}{2} \leq d < \frac{1}{2}$ olup Γ gamma fonksiyonudur.

Tümleşme derecesinin (d) değeri büyüdükçe, karşılaştığı şok karşısında ortalamaya dönmek isteyen serinin gösterdiği direnç yüksek olmaktadır. Tümleşme derecesinin (d) 0 olması durumu serinin kısa hafızalı olduğunu gösterir. Bunun sonucunda ise seriye ARIMA modelleri uygulanabilir. Tümleşme derecesinin 1' den küçük olması durumu ise şokların etkisinin yok olduğunu yani serinin ortalamaya döndüğünü göstermektedir. $d \geq 1$ durumu serinin ortalamasına dönmeyeceğini yani şokların etkisinin kalıcı olduğunu gösterirken, $0 < d < 0.5$ olması durumu serinin uzun hafızalı olduğunu gösterir (Mert ve Çağlar 2019).

4. BULGULAR VE TARTIŞMA

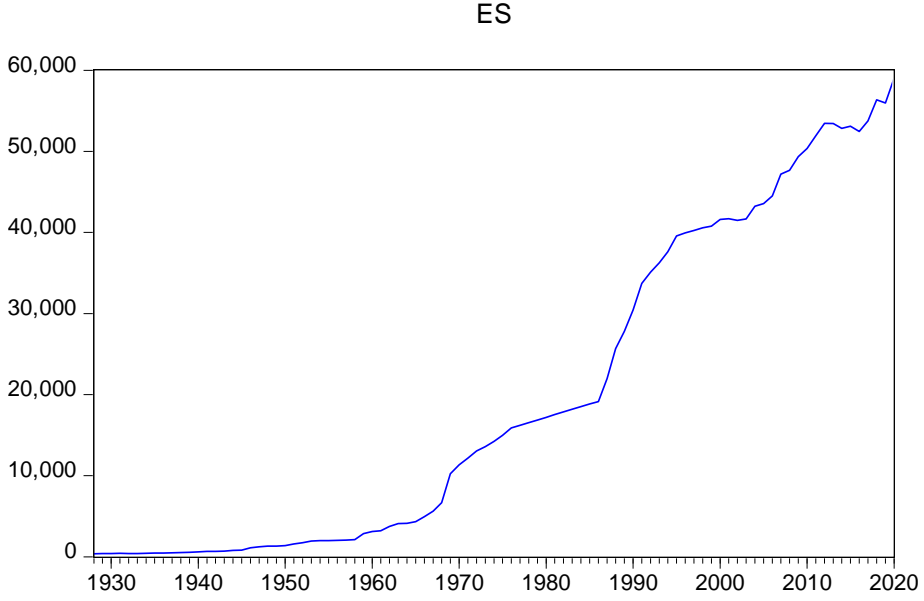
Araştırmanın ikinci kısmında iki farklı veri türü kullanılarak zaman serisi regresyonu ve ARIMA modelleri uygulamaları yapılmıştır. Zaman serisi regresyonu uygulaması için veri seti TÜİK'ten elde edilmiştir. 1928-2020 yılları arasında Türkiye'deki ebe sayılarından veri seti yıllık olarak oluşturulmuştur. Oluşturulan veri seti üzerinde 7 farklı regresyon modeli denenmiş ve en uygun model seçilmiştir. Seçilen en uygun model için 2021 ve 2022 yıllarına ait ebe sayıları tahmin edilmiştir.

Zaman serisi modellerinden olan ARIMA modelleri uygulaması için Antalya iline ait 2014 yılı günlük hava kalite indeksi verileri alınmıştır (Yalçın 2018; Tepe 2018; Doğan 2018; Çizmeci 2018). Meteorolojiden alınmış olan hava tahminleri ile oluşturulan hava kalite indeksi (HKİ) değerleri ARIMA modelleri ile incelenmiş ve 2015 yılı ocak ayı için hava kalite indeksi öngörü değerleri elde edilmiştir.

Veri ile ilgili bu tür bir çalışmaya literatürde rastlanmamıştır. Çalışmanın analizleri için SPSS ve Eviews programları kullanılmıştır.

4.1 Zaman Serisi Regresyon Analizi

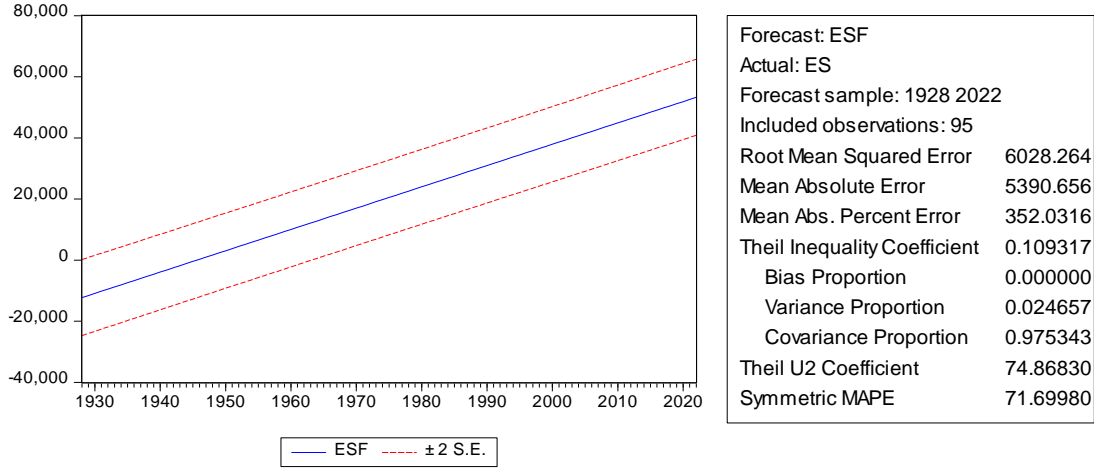
Bu bölümde zaman serisi regresyonunda bahsedilen denklemlerin bilgisayar uygulamaları verilmiştir. Bu bölümde yapılan zaman serisi regresyonu uygulamasının veri seti 1928-2020 yılları arasında Türkiye'deki ebe sayılarına ait yıllık olarak TÜİK'ten alınmış veriler ile oluşturulmuştur. 7 farklı zaman serisi regresyon modeli veri setine uygulanmış ve aralarından en iyi model seçimi yapılmıştır. Seçilen model için 2021 ve 2022 yıllarına ait ebe sayıları tahmin değerleri elde edilmiştir.



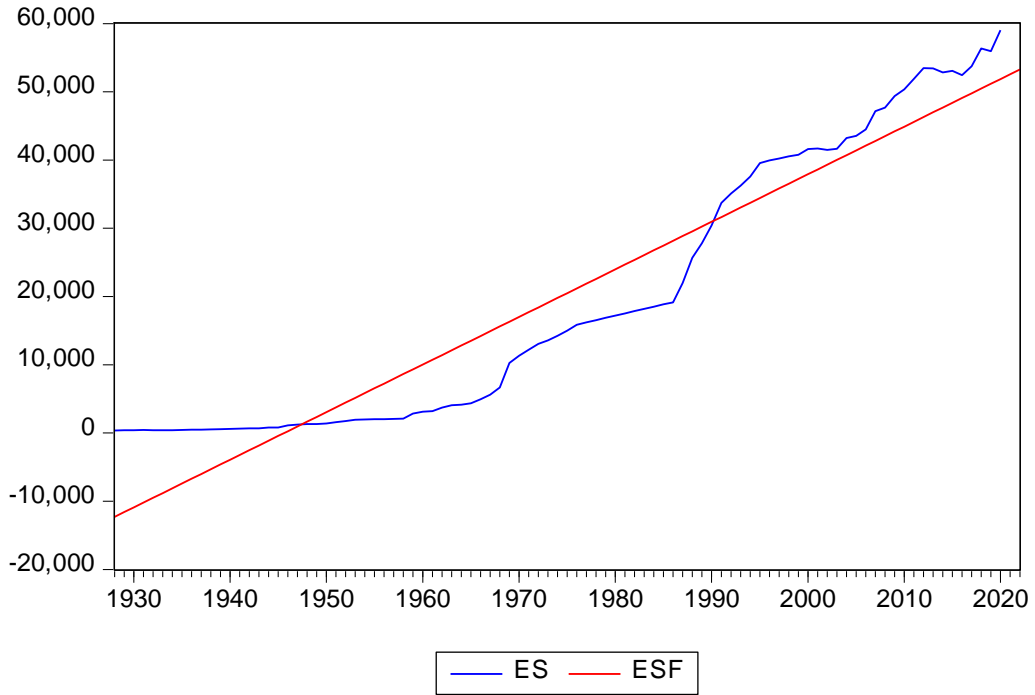
Şekil 4.1. 1928-2020 yılları arası Türkiye’deki ebe sayısı zaman yolu grafiği

Çizelge 4.1. Türkiye’deki ebe sayıları için basit doğrusal regresyon modeli tahmini

Değişken	Katsayı	Standart hata	t-istatistiği	Prob.
C	-12290.04	1253.743	-9.802681	0.0000
@TREND	697.4046	23.53989	29.62651	0.0000
R^2	0.906063	Bağımlı değişken ortalaması		19790.57
AIC	20.28932			
SIC	20.34378			
HQ	20.31131			
F- istatistiği	877.7300	DW test istatistiği		0.022850
Prob(F-istatistiği)	0.000000			



Şekil 4.2. Türkiye'deki ebe sayıları için basit doğrusal regresyon modeline ait bazı istatistikler

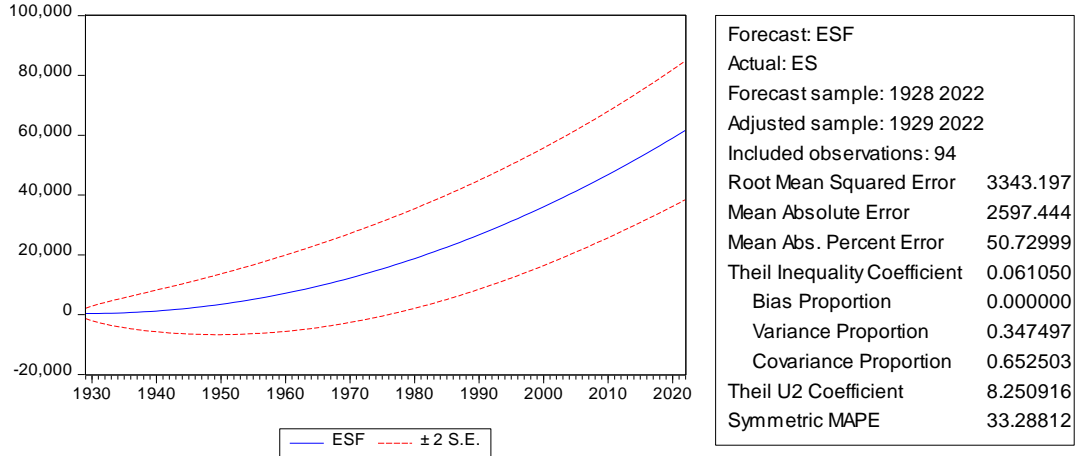


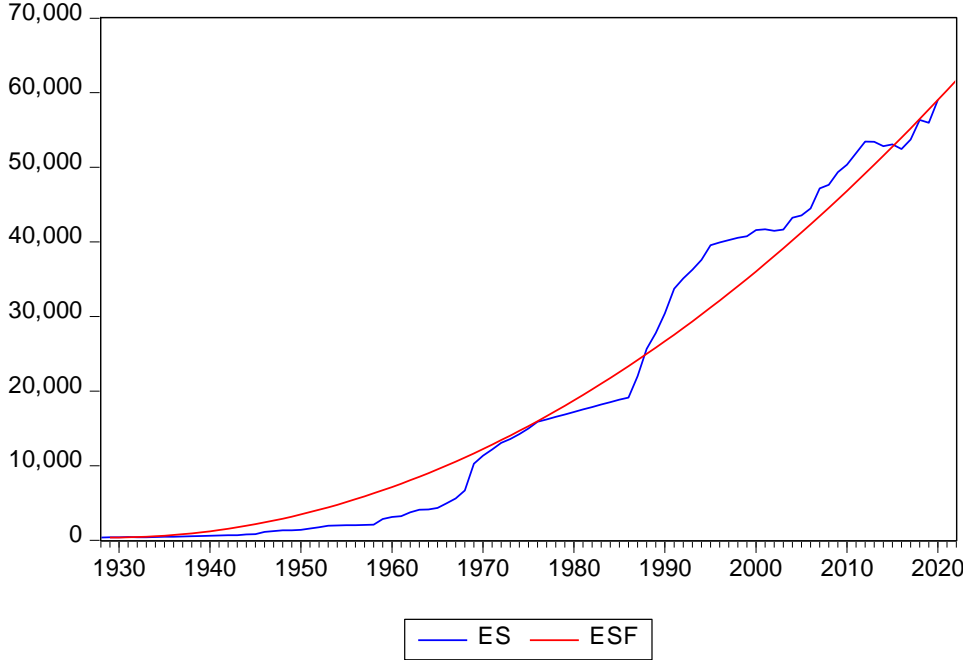
Şekil 4.3. Türkiye'deki ebe sayısı tahmini (basit doğrusal regresyon modeli)

Çizelge 4.1'de basit doğrusal regresyon modeli için tahmin sonuçları görülmektedir. Bu sonuçlara göre $a=-12290.04$ $b=697.4046$ şeklinde bulunmuştur. Ve olasılık değerleri $p<0.01$ olduğundan .01 düzeyinde istatistiksel olarak anlamlı bulunmuştur.

Çizelge 4.2. Türkiye’deki ebe sayıları için birinci farklar regresyon modeli tahmini

Değişken	Katsayı	Standart hata	t-istatistiği	Prob.
C	-23.39608	176.9785	-0.132197	0.8951
@TREND	697.4046	23.53989	29.62651	0.0000
R^2	0.170519	Bağımlı değişken ortalaması		637.6413
AIC	16.33058			
SIC	16.38540			
HQ	16.35270			
F- istatistiği	18.50156	DW test istatistiği		1.214632
Prob(F-istatistiği)	0.000043			

**Şekil 4.4.** Türkiye’deki ebe sayıları birinci farklar regresyon modeli için bazı istatistikler



Şekil 4.5. Türkiye’deki ebe sayısı tahmini (birinci farklar regresyon modeli)

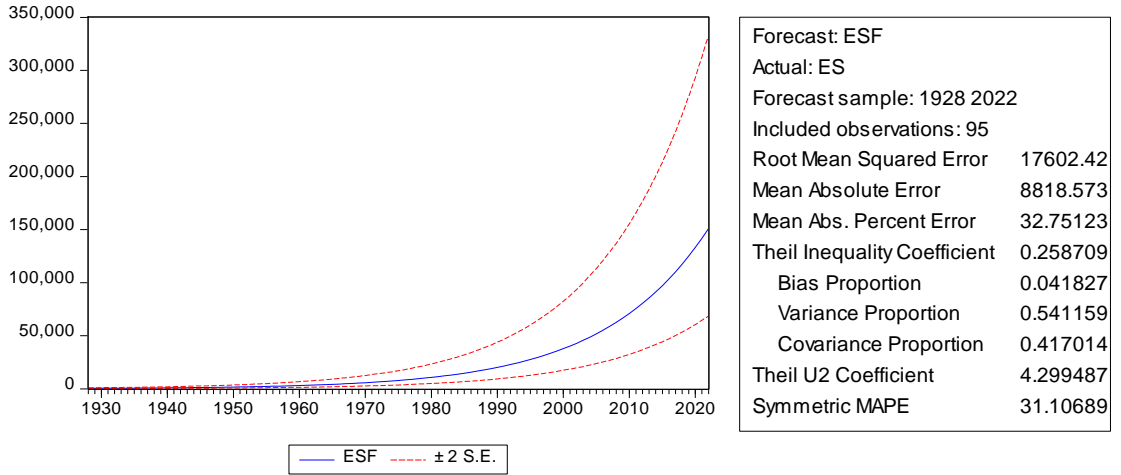
Çizelge 4.2’de birinci farklar regresyon modeline ait tahmin değerleri görülmektedir. Bu tahmin değerlerine göre $a=-23.39608$ $b=697.4046$ bulunmuştur. Sabit katsayı istatistiksel olarak anlamlı bulunmamıştır. Regresyon katsayısı ise olasılık değeri $p<0.01$ olduğundan .01 düzeyinde anlamlı bulunmuştur.

Çizelge 4.3. Türkiye’deki ebe sayıları için üstel regresyon modeli tahmini

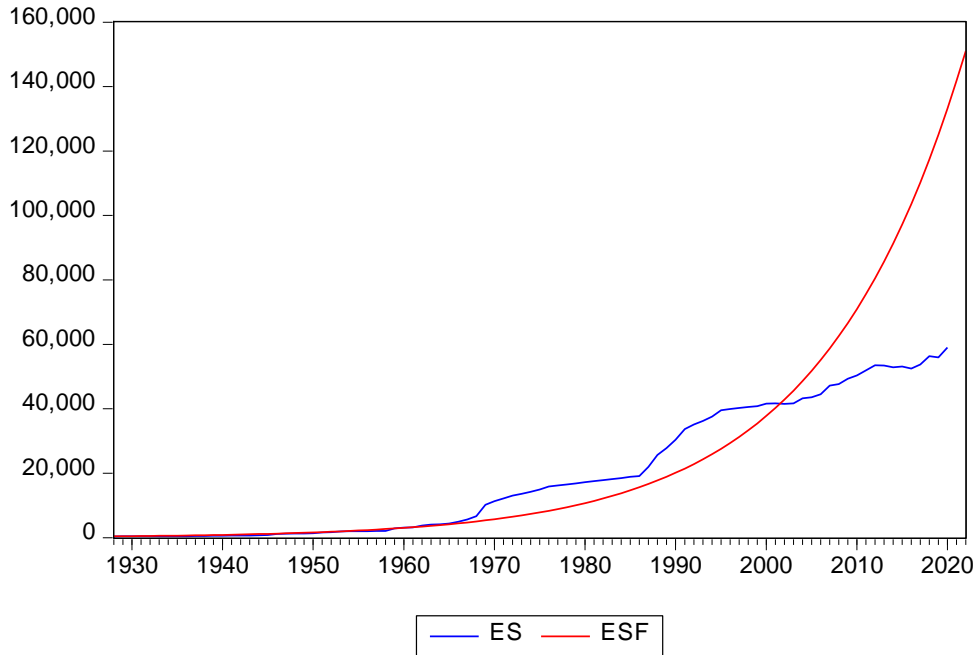
Değişken	Katsayı	Standart hata	t-istatistiği	Prob.
C	6.004849	0.079467	75.56398	0.0000
@TREND	0.062984	0.001492	42.21284	0.0000
R^2	0.951413	Bağımlı değişken ortalaması		8.902096
AIC	0.956714			
SIC	1.011179			
HQ	0.978706			

Çizelge 4.3'ün devamı

F- istatistiği	1781.924	DW test istatistiği	0.033709
Prob(F-istatistiği)	0.000000		



Şekil 4.6. Türkiye'deki ebe sayısı için üstel regresyon modeline ait bazı istatistikler

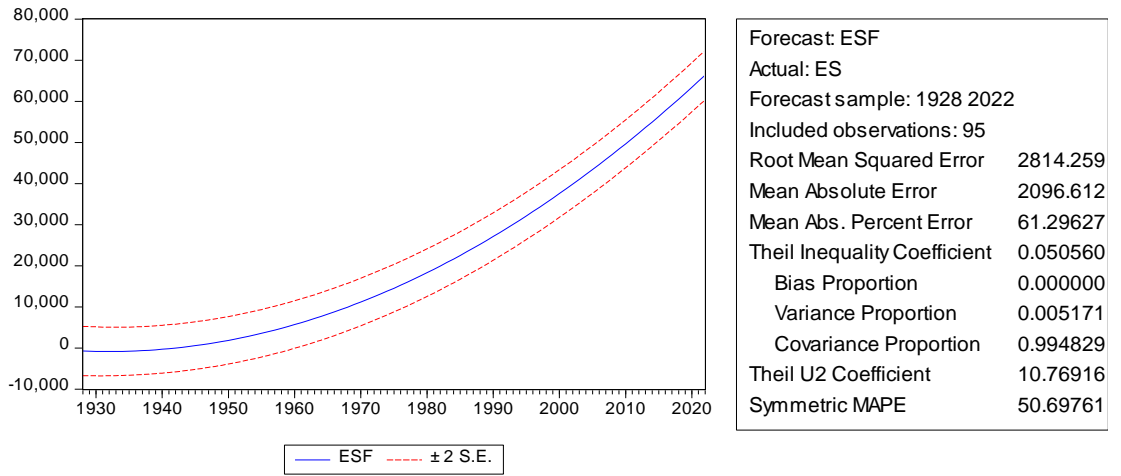


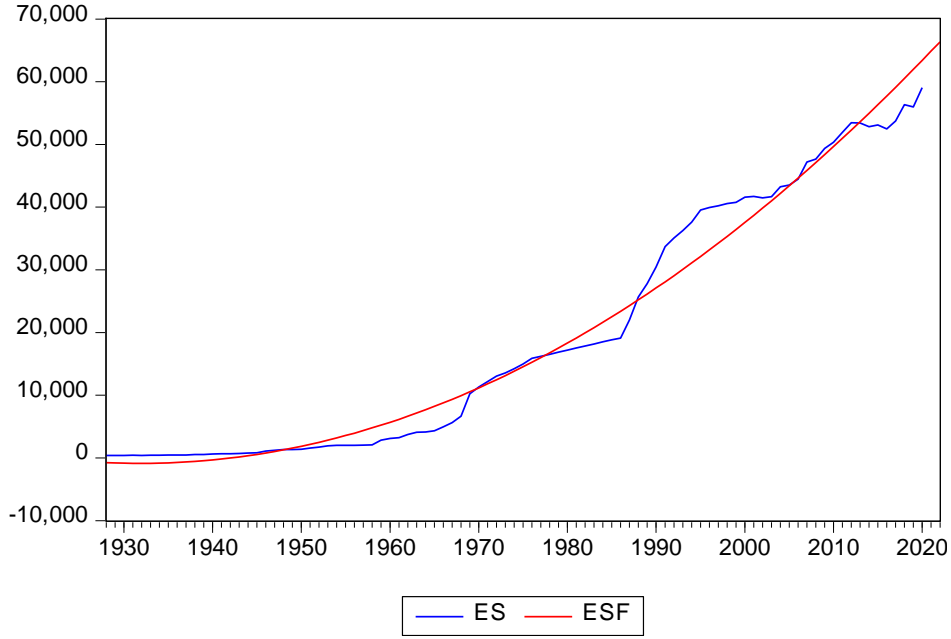
Şekil 4.7. Türkiye'deki ebe sayısı tahmini (üstel regresyon modeli)

Çizelge 4.3'te üstel regresyon modeli için tahmin sonuçları verilmiştir. Bu tahmin sonuçlarına göre $a=6.004849$ $b=0.062984$ şeklinde bulunmuştur. Bu katsayılar olasılık değeri $p<0.01$ olduğundan .01 düzeyinde anlamlıdır.

Çizelge 4.4. Türkiye’deki ebe sayıları için karesel regresyon modeli tahmini

Değişken	Katsayı	Standart hata	t-istatistiği	Prob.
C	-747.9312	871.1455	-0.858561	0.3929
@TREND	-63.61367	43.76541	-1.453515	0.1496
@TREND^2	8.271938	0.460298	17.97084	0.0000
R^2	0.979527	Bağımlı değişken ortalaması		19790.57
AIC	18.78730			
SIC	18.86900			
HQ	18.82029			
F- istatistiği	2153.012	DW test istatistiği		0.087520
Prob(F-istatistiği)	0.000000			

**Şekil 4.8.** Türkiye’deki ebe sayıları için karesel regresyon modeline ait bazı istatistikler



Şekil 4.9. Türkiye’deki ebe sayısı tahmini (karesel regresyon modeli)

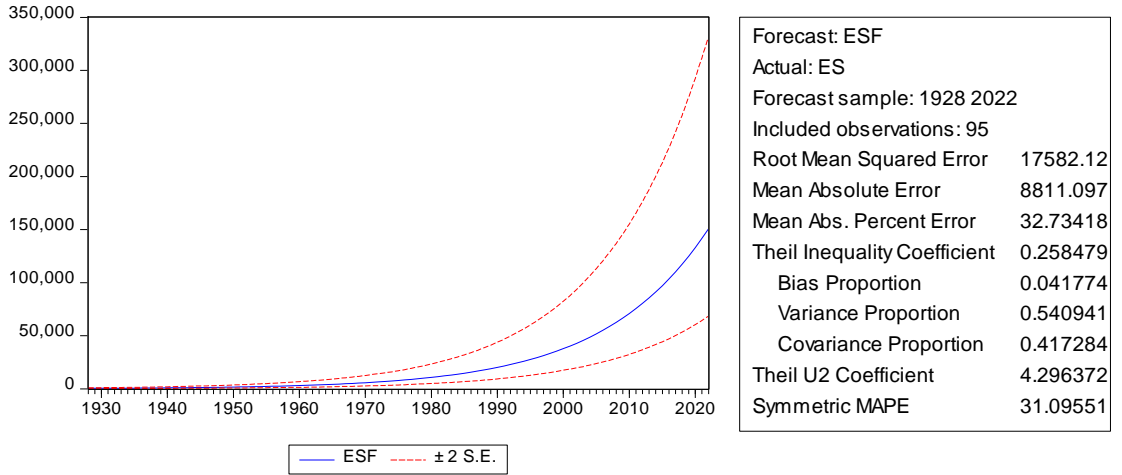
Çizelge 4.4’te karesel regresyon modeline ait tahmin sonuçları bulunmaktadır. Bu tahmin sonuçlarına göre $a=-747.9312$ $b=-63.61367$ $c=8.271938$ şeklinde hesaplanmıştır. Bu modelde a ve b katsayıları istatistiksel olarak anlamlı bulunamazken c katsayısı olasılık değeri $p<0.01$ olduğundan .01 düzeyinde istatistiksel olarak anlamlı bulunmuştur.

Çizelge 4.5. Türkiye’deki ebe sayıları için lojistik regresyon modeli tahmini

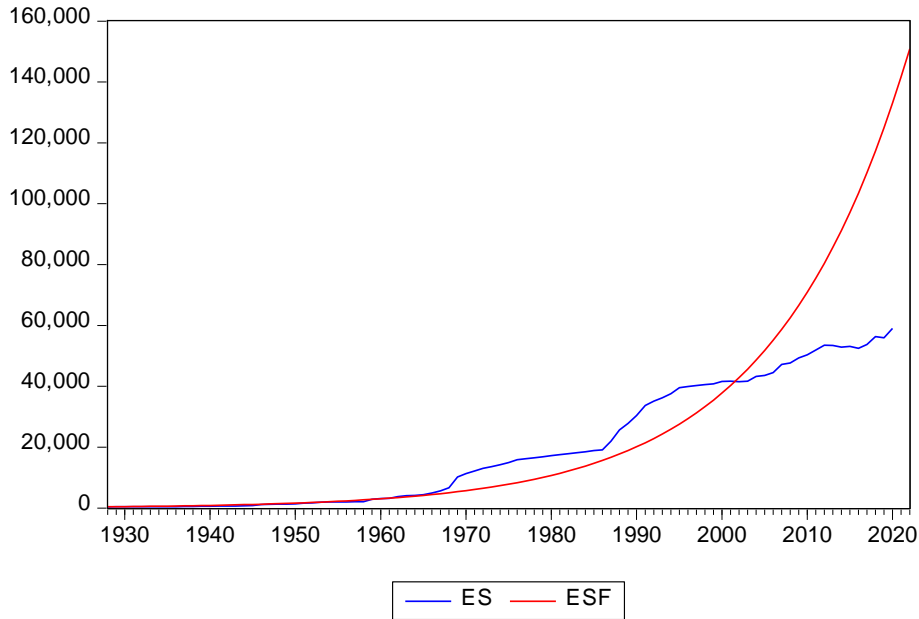
Değişken	Katsayı	Standart hata	t-istatistiği	Prob.
C	12.41595	0.079460	156.2547	0.0000
@TREND	-0.062991	0.001492	-42.22142	0.0000
R^2	0.951432	Bağımlı değişken ortalaması		9.518387
AIC	0.956530			
SIC	1.010994			
HQ	0.978521			

Çizelge 4.5'in devamı

F- istatistiği	1782.648	DW test istatistiği	0.033719
Prob(F-istatistiği)	0.000000		



Şekil 4.10. Türkiye'deki ebe sayısı için lojistik regresyon modeline ait bazı istatistikler



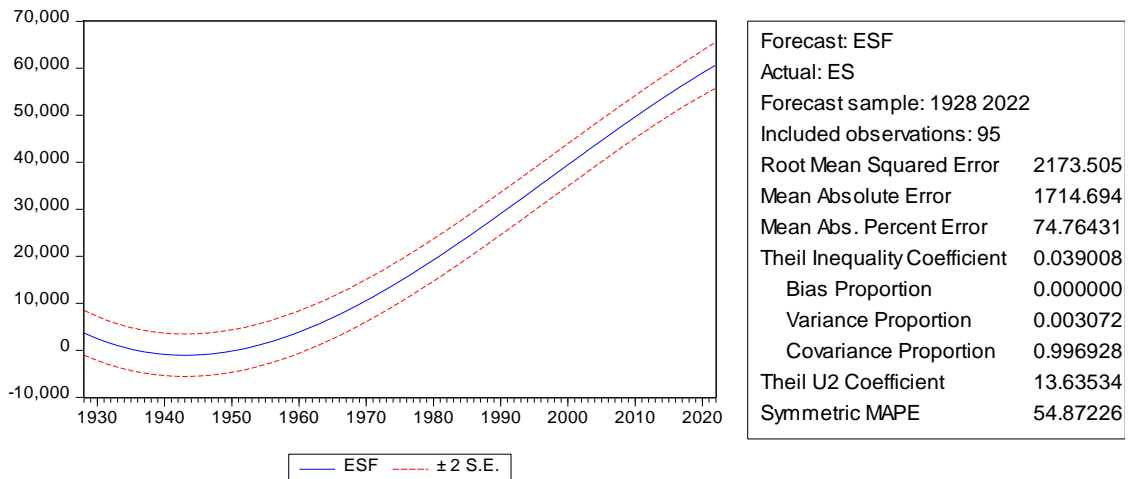
Şekil 4.11. Türkiye'deki ebe sayısı tahmini (lojistik regresyon modeli)

Çizelge 4.5'te lojistik regresyon modeline ait tahmin sonuçları verilmiştir. Bu tahmin sonuçlarına göre $a=12.41595$ $b=-0.062991$ şeklinde bulunmuştur. a ve b

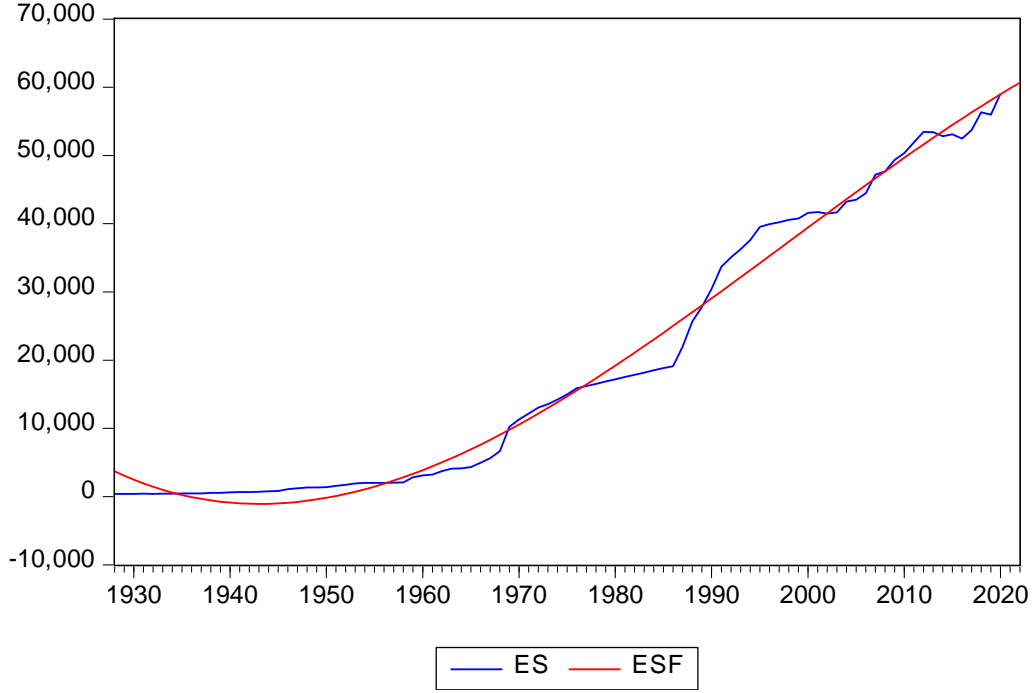
katsayıları olasılık değeri $p < 0.01$ olduğundan .01 düzeyinde istatistiksel olarak anlamlıdır.

Çizelge 4.6. Türkiye’deki ebe sayıları için kübik regresyon modeli tahmini

Değişken	Katsayı	Standart hata	t-istatistiği	Prob.
C	3686.340	885.6185	4.162447	0.0001
@TREND	-658.1219	83.81820	-7.851778	0.0000
@TREND^2	24.51469	2.123583	11.54402	0.0000
@TREND^3	-0.117701			
R^2	0.987788	Bağımlı değişken ortalaması		19790.57
AIC	18.29209			
SIC	18.40102			
HQ	18.33607			
F- istatistiği	2399.701	DW test istatistiği		0.143770
Prob(F-istatistiği)	0.000000			



Şekil 4.12. Türkiye’deki ebe sayısı için kübik regresyon modeline ait bazı istatistikler

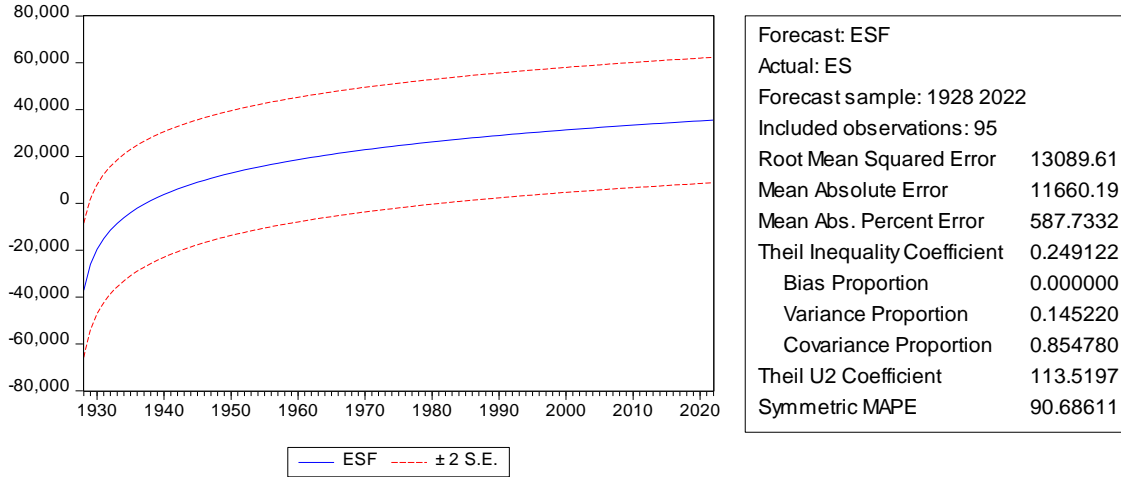


Şekil 4.13. Türkiye’deki ebe sayısı tahmini (kübik regresyon modeli)

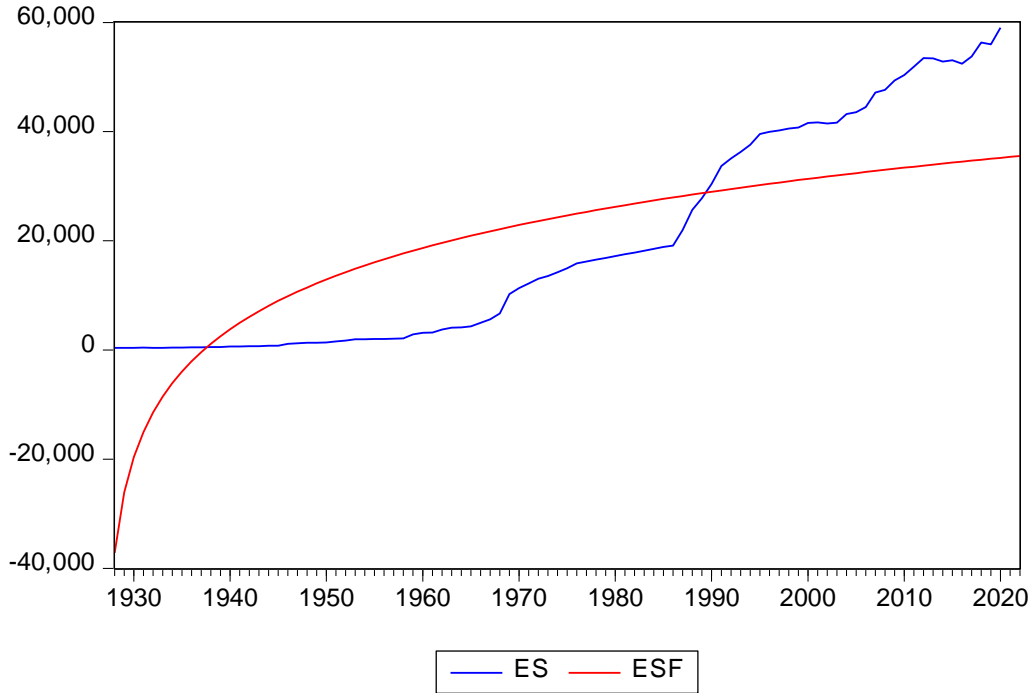
Çizelge 4.6’da kübik regresyon modeline ait tahmin değerleri görülmektedir. Bu tahmin değerlerine göre $a=3686.340$ $b=-658.1219$ $c=24.51469$ $d=0.117701$ şeklinde bulunmuştur ve bu katsayılar olasılık değerleri $p<0.01$ olduğundan .01 düzeyinde istatistiksel olarak anlamlı bulunmuştur.

Çizelge 4.7. Türkiye’deki ebe sayısı için logaritmik regresyon modeli tahmini

Değişken	Katsayı	Standart hata	t-istatistiği	Prob.
C	-37147.83	5496.019	-6.759042	0.0000
LOG(T)	15963.18	1492.061	10.69874	0.0000
R^2	0.557098	Bağımlı değişken ortalaması		19790.57
AIC	21.84004			
SIC	21.89450			
HQ	21.86203			
F- istatistiği	114.4630	DW test istatistiği		0.020486
Prob(F-istatistiği)	0.000000			



Şekil 4.14. Türkiye’deki ebe sayısı için logaritmik regresyon modeline ait bazı istatistikler



Şekil 4.15. Türkiye’deki ebe sayısı tahmini (logaritmik regresyon modeli)

Çizelge 4.7’de logaritmik regresyon modeli için tahmin sonuçları hesaplanmıştır. Bu tahmin değerlerine göre $a=-37147.83$ $b=15963.18$ şeklinde bulunmuştur ve bulunan katsayılar olasılık değerleri $p<0.01$ olduğundan .01 düzeyinde istatistiksel olarak anlamlıdır.

Şekil 4.2, şekil 4.4, şekil 4.6, şekil 4.8, şekil 4.10, şekil 4.12, şekil 4.14'te regresyon modellerine ait olan RMSE, MAE, MAPE, U1, U2, AIC, SIC, HQ istatistikleri çizelge 4.8'de toplanmış ve değerlendirilmiştir.

Çizelge 4.8. Zaman serisi regresyon modelleri

model	RMSE	MAE	MAPE	U1	U2	AIC	SIC	HQ
Basit doğrusal regresyon modeli	6028.264	5390.656	352.0316	0.109317	74.86830	20.28932	20.34378	20.31131
Birinci farklar regresyon modeli	3343.197	2597.444	50.72999	0.061050	8.250916	16.33058	16.38540	16.35270
Üstel regresyon modeli	17602.42	8818.573	32.75123	0.258709	4.299487	0.956714	1.011179	0.978706
Karesel regresyon modeli	2814.259	2096.612	61.29627	0.050560	10.766916	18.78730	18.86900	18.82029
Lojistik regresyon modeli	17582.12	8811.097	32.73418	0.258479	4.296372	0.956530	1.010994	0.978521
Kübik regresyon modeli	2173,505	1714,694	74.76431	0.039008	13.63534	18.29209	18.40102	18.33607
Logaritmik regresyon modeli	13089.61	11660.19	587.7332	0.249122	113.5197	21.84004	21.89450	21.86203

Yukarıdaki zaman serisi regresyonunda 1928-2020 yılları arası Türkiye'de ki ebe sayısı yıllık verileri kullanılmıştır. Öngörü dönemi 2 dönem genişletilmiş ve 2021-2022 yılları için tahmin yapmak amacıyla en iyi model seçimi yapılmıştır.

Çizelge 4.8'e baktığımızda 7 modelin analizlerinden elde edilen istatistiksel değerleri görüyoruz. Burada AIC, SIC ve HQ değerleri karşılaştırılırken bağımlı değişkenlerin aynı formda olması gerektiğine dikkat edilmiştir. Örnek verecek olursak burada üstel regresyon modelimizde bağımlı değişken logaritmik bir formdadır ve bu sebeple üstel regresyon modeli ile basit doğrusal regresyon modeli AIC, SIC ve HQ değerlerine bakılarak karşılaştırılmaz. Bu nedenle en iyi model tercihi yaparken RMSE, MAE, MAPE, U1, U2 değerlerinden yola çıkılmıştır. Çünkü bu değerler serinin

düzey değerleri ve tahmini düzey değerlerinden elde edildiğinden bağımlı değişkenin formu farklı olsa da bir önem arz etmez. Model seçiminde hataların karelerinin en küçük olması istendiğinden RMSE değeri en küçük olan model kübik regresyon modelidir. Bu da en uygun modelin kübik regresyon modeli olduğunu göstermektedir. Çizelge 4.9’da kübik regresyon modeli için ulaşılan tahmin değerleri verilmiştir.

Çizelge 4.9. Kübik regresyon modeline ait tahmin değerleri

2021	59834.8469587145
2022	59834.8469587145

4.2 Arıma Modellerinin Tahmini

Arıma modeli tahmin sürecinde serinin korelogram grafiğine bakılmış, birim kök testi uygulanmıştır. Serinin mevsimsel olmadığına ve durağan bir yapıda olduğuna çıkan analiz sonuçları ile birlikte karar verilmiştir. Bu tahmin sürecinde iyi bir model elde edebilmek için tüm AR ve MA kat sayılarının anlamlı olması gerekmektedir. Bunun yanı sıra elde edilen kalıntılarda otokorelasyon olmamalıdır. Durbin-Watson testine göre $1.7 < D.W < 2.3$ arasında seyreden bir test istatistiği için otokorelasyon yoktur denilebilir. Son olarak ise kalıntılar eğer temiz dizi olma özelliği gösterirse iyi bir model tahmini yapılmış olacaktır. Temiz dizi olma özelliğine ise korelogram grafiği incelenerek bakılabilir (Mert ve Çağlar).

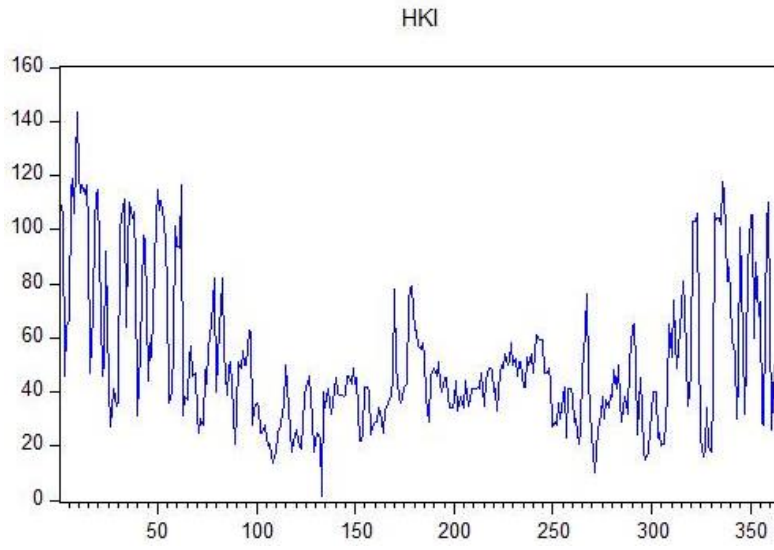
Bizim veri setimiz şekil 4.17’de verilen HKI verisi korelogram grafiğine göre mevsimsellik barındırmamaktadır. Aynı zamanda çıktılar çizelge 4.11’de ve çizelge 4.12’de verilen ADF VE PP testlerine göre düzeyde durağan bulunmuştur. Yani $I(0)$ olduğundan $d=0$ olmalıdır. Bu nedenle uygun olan model $ARMA(p,q)$ modeli olarak alınmıştır.

Çizelge 4.10. Değişkenlerin tanımlayıcı istatistikleri

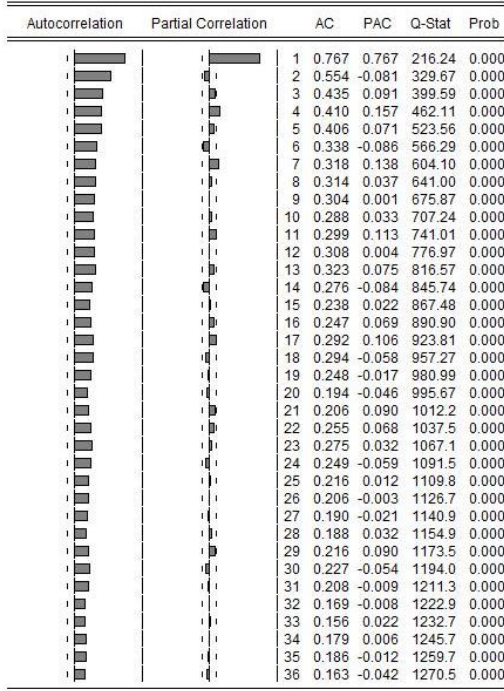
Tanımlayıcı istatistikler	HKI
N	728
Ortalama	48,8857916
Medyan	42
Mod	33

Çizelge 4.10'un devamı

Standart Sapma	25,39556742
Basıklık	1,121
Çarpıklık	0,731
Değişim Aralığı	142,56875
Minimum	1
Maksimum	143,56875



Şekil 4.16. Hava kalite indeksinin günlük değişim zaman yolu grafiği



Şekil 4.17. Hava kalite indeksi korelogramı

Çizelge 4.11. ADF birim kök testi çıktıları

			t istatistiği	Prob.*
Augmented Dickey-Fuller test istatistiği			-7.001044	0.0000
Kritik değerler	1% level		-3.448111	
	5% level		-2.869263	
	10% level		-2.570952	

Şekil 4.17’de 2014 yılında alınan hava kalite indeksi günlük verilerinin korelogram grafiği görülmektedir. 1, 2, 3 şeklinde devam eden sayılar gecikmeleri, AC değerleri otokorelasyon katsayılarını, PAC değerleri ise kısmi otokorelasyon katsayılarını göstermektedir. Otokorelasyon katsayılarına baktığımızda %95 güven aralığının dışında kaldığını görüyoruz. Bu ise 2014 HKI günlük verilerinde otokorelasyonun karşımıza çıkabileceğini, dolayısı ile serinin durağan olmayabileceğini göstermektedir. Mevsimsel serilerde otokorelasyon grafiği sarmal biçimde devam etmektedir. Ancak HKI serisinin otokorelasyon grafiği sarmal biçimde değildir. Sonuç olarak veri setimiz için mevsimselliğin bulunmadığını söyleyebiliriz. İlk gözlem olarak

durağanlığın olmamasından bahsedilebilmekle beraber ADF ve PP birim kök testi yapılmış, durağanlığın olup olmadığına kesin bir şekilde bakılmıştır.

Çizelge 4.11’de verilen ADF testi için yokluk hipotezi HKI verisi birim kök içerir şeklinde kurulmuştur ve sabitli model üzerinden ADF denkleminin tahmin edildiği görülmektedir. Çizelge 4.11’deki çıktıdan test istatistiği $t_{\delta} = -7.001044$ olarak bulunmuştur. Test istatistiğinin yanı sıra 0.01, 0.05, 0.10 yanılma düzeyleri için MacKinnon (1996) tarafından elde edilen kritik değerler verilmiştir. Hesaplanan test istatistiği tüm değerlerden küçük olduğu için kurulan yokluk hipotezi reddedilecektir. Yani HKI verisi birim kök içermez alternatif hipotezi kabul edilir. Dolayısı ile veri seti durağandır. Eğer seri durağan dışı bir seri olsaydı serinin farkı alınarak durağanlaştırılması gerekirdi ancak verimizde böyle bir işlem yapılmasına gerek kalmamıştır. Serimiz düzeyde durağan bulunduğu için I(0) denir. Birim kök testleri arasından PP testinin de bir uygulaması yapılmış ve çıktıları çizelge 4.12’de verilmiştir.

Çizelge 4.12. PP birim kök testi (trendli ve sabitli model) çıktıları

			Adj. t-Stat	Prob.*
Phillips-Perron test istatistiği			-6.944023	0.0000
Kritik test değerleri:	1% level		-3.983471	
	5% level		-3.422218	
	10% level		-3.133955	

Çizelge 4.14’de görüldüğü üzere test istatistiği $t_{\delta} = -6.944023$ bulunmuştur. Test istatistiği MacKinnon (1996) tarafından üretilen 0.05, 0.01, 0.1 yanılma düzeyleri için kritik değerlerden küçük olduğu için yokluk hipotezi reddedilir ve serinin birim kök içermediğini ve sonuç olarak düzeyde durağan bir seri olduğunu söyleyebiliriz.

Çizelge 4.13. AR(1) model tahmin analizleri

Değişken	Katsayı	Standart hata	t istatistiği	Prob.
C	51.58025	4.814524	10.71347	0.0000
AR(1)	0.775137	0.036779	21.07581	0.0000
SIGMASQ	287.9412	16.76589	17.17423	0.0000
R^2	0.595853	Bağımlı değişken ortalaması		51.20724

Çizelge 4.13'ün devamı

Akaike kriteri	8.519588		
Schwarz kriteri	8.551642		
Hannan-Quinn kriteri	8.532327		
F-istatistiği	266.8565	Durbin-Watson istatistiği	1.881202
Prob(F istatistiği)	0.000000		

Date: 09/13/22 Time: 12:19
Sample: 1/01/2014 1/31/2015
Included observations: 365

Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob
1	0.053	0.053	1.0354	0.309	
2	-0.098	-0.101	4.5919	0.101	
3	-0.129	-0.119	10.748	0.013	
4	0.009	0.013	10.779	0.029	
5	0.168	0.147	21.285	0.001	
6	-0.049	-0.081	22.181	0.001	
7	0.006	0.044	22.193	0.002	
8	0.030	0.057	22.522	0.004	
9	0.052	0.036	23.552	0.005	
10	-0.012	-0.031	23.603	0.009	
11	0.035	0.081	24.072	0.012	
12	0.027	0.017	24.342	0.018	
13	0.150	0.149	32.861	0.002	
14	0.023	0.015	33.068	0.003	
15	-0.053	-0.014	34.133	0.003	
16	-0.047	-0.033	34.999	0.004	
17	0.108	0.122	39.468	0.002	
18	0.124	0.053	45.386	0.000	
19	0.043	0.058	46.085	0.000	
20	-0.089	-0.063	49.160	0.000	
21	-0.050	-0.013	50.134	0.000	
22	0.085	0.036	52.937	0.000	
23	0.118	0.097	58.349	0.000	
24	0.054	0.023	59.482	0.000	
25	-0.015	0.028	59.566	0.000	
26	0.040	0.037	60.203	0.000	
27	-0.007	-0.023	60.222	0.000	
28	-0.031	-0.049	60.595	0.000	
29	0.058	0.088	61.938	0.000	
30	0.071	0.017	63.956	0.000	
31	0.058	0.012	65.301	0.000	
32	-0.032	-0.024	65.717	0.000	
33	-0.045	0.004	66.535	0.000	
34	0.037	0.010	67.099	0.001	
35	0.075	0.032	69.367	0.000	
36	0.002	-0.068	69.368	0.001	

Şekil 4.18. AR(1) modeli kalıntı korelogramı

Çizelge 4.13'te verilen AR(1) model tahmin sonuçlarına göre AR(1) katsayısı 0.775137 şeklinde bulunmuştur ve sabit katsayı 51.58025 olarak hesaplanmıştır. AR(1) katsayısı istatistiksel olarak anlamlı bulunmuştur. Çünkü $p=0.000 < 0.10$ bulunmuştur. Durbin-Watson test istatistiği ise 1.88 çıktığından kalıntılarda otokorelasyon mevcut değildir. Son olarak kalıntıların korelogram grafiğine bakılmış ve şekil 4.18'de çıkan sonuçlarla birlikte temiz dizi olmadığına karar verilmiştir. Çünkü AC değerleri anlamlı bulunmuştur. Bu aşamada bir diğer model olan AR(2) modeline geçilmiştir.

Çizelge 4.14. AR(2) model tahmin analizleri

Değişken	Katsayı	Standart hata	t istatistiği	Prob.
C	51.51916	4.462169	11.54577	0.0000
AR(1)	0.834555	0.049304	16.92657	0.0000
AR(2)	-0.077573	0.043233	-1.794293	0.0736
SIGMASQ	286.2186	16.91972	16.91628	0.0000
R^2	0.598270	Bağımlı değişken ortalaması		51.20724
Akaike kriteri	8.519093			
Schwarz kriteri	8.561832			
Hannan-Quinn kriteri	8.536078			
F istatistiği	179.2049	Durbin-Watson istatistiği		1.988200
Prob(F istatistiği)	0.000000			

Date: 09/13/22 Time: 13:51
 Sample: 1/01/2014 1/31/2015
 Included observations: 365

Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob
		1 -0.000	-0.000	8.E-05	0.993
		2 -0.057	-0.057	1.2023	0.548
		3 -0.092	-0.093	4.3475	0.226
		4 0.027	0.023	4.6175	0.329
		5 0.186	0.178	17.465	0.004
		6 -0.048	-0.054	18.318	0.005
		7 0.024	0.048	18.530	0.010
		8 0.038	0.068	19.073	0.014
		9 0.064	0.051	20.608	0.015
		10 -0.005	-0.026	20.619	0.024
		11 0.047	0.083	21.464	0.029
		12 0.022	0.015	21.655	0.042
		13 0.155	0.148	30.756	0.004
		14 0.026	0.025	31.009	0.006
		15 -0.036	-0.011	31.494	0.008
		16 -0.038	-0.043	32.053	0.010
		17 0.111	0.114	36.764	0.004
		18 0.119	0.055	42.198	0.001
		19 0.049	0.063	43.132	0.001
		20 -0.075	-0.060	45.322	0.001
		21 -0.037	-0.021	45.844	0.001
		22 0.088	0.028	48.893	0.001
		23 0.112	0.096	53.816	0.000
		24 0.054	0.026	54.964	0.000
		25 -0.012	0.023	55.021	0.000
		26 0.051	0.035	56.068	0.001
		27 0.002	-0.019	56.070	0.001
		28 -0.025	-0.059	56.317	0.001
		29 0.063	0.081	57.920	0.001
		30 0.068	0.020	59.750	0.001
		31 0.062	0.014	61.286	0.001
		32 -0.025	-0.025	61.529	0.001
		33 -0.036	0.001	62.040	0.002
		34 0.042	0.006	62.755	0.002
		35 0.077	0.038	65.167	0.001
		36 -0.001	-0.069	65.167	0.002

Şekil 4.19. AR(2) modeli kalıntı korelogramı

AR(2) tahmin modeli için analiz yapılmıştır. Çizelge 4.14'te görüldüğü üzere AR(1) ve AR(2) katsayıları anlamlı bulunmuştur. Durbin-Watson test istatistiği de bize kalıntılarda bir otokorelasyon bulunmadığını göstermiştir. Ancak şekil 4.19'da görüldüğü üzere yine AR(2) Modeli için kalıntıların korelogram grafiği kalıntıların temiz dizi olmadığını göstermiştir. Dolayısı ile model tam anlamıyla iyi bir model olmamıştır.

Çizelge 4.15. MA(1) model tahmin analizleri

Değişken	Katsayı	Standart hata	t istatistiği	Prob.
C	51.20156	2.139237	23.93450	0.0000
MA(1)	0.672525	0.033969	19.79812	0.0000
SIGMASQ	389.6598	30.71296	12.68715	0.0000

Çizelge 4.15'in devamı

R^2	0.453083	Bağımlı değişken ortalaması	51.20724
Akaike info kriteri	8.821239		
Schwarz kriteri	8.853293		
Hannan-Quinn kriteri	8.833977		
F istatistiği	149.9460	Durbin-Watson istatistiği	1.427598
Prob(F istatistiği)	0.000000		

Date: 09/13/22 Time: 14:00
Sample: 1/01/2014 1/31/2015
Included observations: 365

Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob
■	■	1 0.277	0.277	28.245	0.000
■	■	2 0.485	0.443	115.20	0.000
■	■	3 0.186	-0.015	128.04	0.000
■	■	4 0.297	0.074	160.80	0.000
■	■	5 0.290	0.213	192.09	0.000
■	■	6 0.181	-0.060	204.35	0.000
■	■	7 0.233	0.021	224.60	0.000
■	■	8 0.172	0.086	235.67	0.000
■	■	9 0.236	0.062	256.67	0.000
■	■	10 0.145	-0.041	264.55	0.000
■	■	11 0.235	0.112	285.47	0.000
■	■	12 0.153	0.025	294.41	0.000
■	■	13 0.265	0.087	321.14	0.000
■	■	14 0.141	-0.020	328.75	0.000
■	■	15 0.175	-0.027	340.44	0.000
■	■	16 0.124	-0.006	346.32	0.000
■	■	17 0.223	0.133	365.44	0.000
■	■	18 0.183	0.025	378.39	0.000
■	■	19 0.178	-0.012	390.64	0.000
■	■	20 0.102	-0.064	394.66	0.000
■	■	21 0.127	0.007	400.96	0.000
■	■	22 0.177	0.079	413.23	0.000
■	■	23 0.185	0.084	426.57	0.000
■	■	24 0.169	-0.019	437.84	0.000
■	■	25 0.126	-0.023	444.06	0.000
■	■	26 0.146	0.001	452.45	0.000
■	■	27 0.119	-0.003	458.02	0.000
■	■	28 0.109	-0.041	462.71	0.000
■	■	29 0.154	0.098	472.15	0.000
■	■	30 0.141	0.018	480.05	0.000
■	■	31 0.149	-0.026	488.90	0.000
■	■	32 0.093	-0.028	492.39	0.000
■	■	33 0.098	0.015	496.27	0.000
■	■	34 0.112	0.004	501.39	0.000
■	■	35 0.136	0.028	508.91	0.000
■	■	36 0.097	-0.049	512.76	0.000

Şekil 4.20. MA(1) modeli kalıntı korelogramı

Çizelge 4.15'te görüldüğü gibi MA(1) model tahmini yapılmıştır ancak MA(1) katsayısı anlamlı olsa bile Durbin-Watson test istatistiğine göre kalıntılarda otokorelasyon olduğu görülmüştür. Ayrıca şekil 4.20'de kalıntıların korelogramına bakıldığında yine temiz dizi olmadığını anlaşılmıştır. Dolayısı ile AR(1), AR(2) modellerinin yanında MA(1) modeli daha başarısız olduğuna karar verilmiştir.

Çizelge 4.16. MA(1) model tahmin analizleri

Değişken	Katsayı	Standart hata	t istatistiği	Prob.
C	51.35835	2.417601	21.24351	0.0000
MA(1)	0.772785	0.041311	18.70645	0.0000
MA(2)	0.488432	0.042486	11.49619	0.0000
SIGMASQ	314.6644	19.61945	16.03838	0.0000
R^2	0.558345	Bağımlı değişken ortalaması		51.20724
Akaike kriteri	8.613655			
Schwarz kriteri	8.656394			
Hannan-Quinn kriteri	8.630640			
F istatistiği	152.1265	Durbin-Watson istatistiği		1.743149
Prob(F istatistiği)	0.000000			

Date: 09/13/22 Time: 14:23
 Sample: 1/01/2014 1/31/2015
 Included observations: 365

Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob	
		1	0.121	0.121	5.3448	0.021
		2	0.137	0.125	12.302	0.002
		3	0.291	0.269	43.544	0.000
		4	0.090	0.026	46.531	0.000
		5	0.267	0.215	72.976	0.000
		6	0.112	-0.008	77.650	0.000
		7	0.098	0.032	81.266	0.000
		8	0.185	0.051	94.040	0.000
		9	0.125	0.069	99.933	0.000
		10	0.100	-0.010	103.70	0.000
		11	0.156	0.075	112.95	0.000
		12	0.096	0.006	116.44	0.000
		13	0.194	0.125	130.81	0.000
		14	0.125	0.009	136.80	0.000
		15	0.066	-0.005	138.47	0.000
		16	0.090	-0.058	141.59	0.000
		17	0.159	0.103	151.34	0.000
		18	0.149	0.051	159.96	0.000
		19	0.117	0.054	165.27	0.000
		20	0.044	-0.081	166.01	0.000
		21	0.073	-0.013	168.07	0.000
		22	0.134	0.015	175.09	0.000
		23	0.151	0.121	184.01	0.000
		24	0.104	0.014	188.30	0.000
		25	0.060	-0.018	189.70	0.000
		26	0.133	0.011	196.70	0.000
		27	0.072	-0.016	198.77	0.000
		28	0.045	-0.047	199.58	0.000
		29	0.130	0.073	206.32	0.000
		30	0.100	0.023	210.36	0.000
		31	0.102	0.023	214.54	0.000
		32	0.074	-0.031	216.74	0.000
		33	0.022	-0.022	216.94	0.000
		34	0.102	0.009	221.18	0.000
		35	0.123	0.058	227.32	0.000
		36	0.019	-0.079	227.48	0.000

Şekil 4.21. MA(2) model kalıntı korelogramı

Seri için MA(2) model tahmini yapılmış ve çizelge 4.16’da sonuçları verilmiştir. Yapılan tahminde MA(1) ve MA(2) katsayıları anlamlı Durbin-Watson test istatistiği 1.74 olarak bulunmuştur. Durbin-Watson test istatistiğine göre ise kalıntıların otokorelasyon içermediği söylenebilir. Şekil 4.21’de verilen kalıntıların korelogramına bakılırsa yine temiz dizi olma özelliği taşımadığı görülmüştür. Ancak MA(1) modeline göre daha iyi sonuç vermiştir.

Çizelge 4.17. ARMA(1,1) model tahmin analizleri

Değişken	Katsayı	Standart hata	t istatistiği	Prob.
AR(1)	0.957168	0.012859	74.43372	0.0000
MA(1)	-0.037060	0.048379	-0.766042	0.4442
SIGMASQ	316.3099	14.24465	22.20553	0.0000

Çizelge 4.17'nin devamı

R^2	0.556035	Bağımlı değişken ortalaması	51.20724
Akaike kriteri	8.617636		
Schwarz kriteri	8.649689		
Hannan-Quinn kriteri	8.630374		
Durbin-Watson istatistiği	1.989782		

Date: 09/13/22 Time: 14:58
Sample: 1/01/2014 1/31/2015
Included observations: 365

Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob
█	█	1 -0.014	-0.014	0.0706	0.790
█	█	2 -0.195	-0.195	14.134	0.001
█	█	3 -0.216	-0.231	31.475	0.000
█	█	4 -0.050	-0.116	32.403	0.000
█	█	5 0.131	0.033	38.752	0.000
█	█	6 -0.095	-0.185	42.112	0.000
█	█	7 -0.037	-0.069	42.634	0.000
█	█	8 -0.009	-0.051	42.664	0.000
█	█	9 0.019	-0.063	42.798	0.000
█	█	10 -0.049	-0.136	43.695	0.000
█	█	11 -0.000	-0.029	43.695	0.000
█	█	12 -0.004	-0.083	43.700	0.000
█	█	13 0.128	0.071	49.977	0.000
█	█	14 -0.005	-0.039	49.988	0.000
█	█	15 -0.093	-0.072	53.271	0.000
█	█	16 -0.088	-0.108	56.252	0.000
█	█	17 0.083	0.053	58.936	0.000
█	█	18 0.106	0.008	63.244	0.000
█	█	19 0.017	0.031	63.353	0.000
█	█	20 -0.126	-0.091	69.531	0.000
█	█	21 -0.087	-0.062	72.478	0.000
█	█	22 0.061	-0.019	73.923	0.000
█	█	23 0.100	0.058	77.813	0.000
█	█	24 0.031	0.003	78.199	0.000
█	█	25 -0.044	0.015	78.968	0.000
█	█	26 0.014	0.035	79.049	0.000
█	█	27 -0.035	-0.026	79.547	0.000
█	█	28 -0.059	-0.065	80.939	0.000
█	█	29 0.037	0.072	81.483	0.000
█	█	30 0.053	0.016	82.599	0.000
█	█	31 0.041	0.018	83.263	0.000
█	█	32 -0.057	-0.020	84.569	0.000
█	█	33 -0.070	0.004	86.543	0.000
█	█	34 0.018	0.009	86.667	0.000
█	█	35 0.061	0.038	88.153	0.000
█	█	36 -0.016	-0.068	88.253	0.000

Şekil 4.22. ARMA(1,1) model kalıntı korelogramı

Çizelge 4.17'de verilen Arma(1,1) Model tahmin sonuçlarına bakılırsa AR(1) katsayısı anlamlı ancak MA(1) katsayısı anlamlı bulunmamıştır. Kalıntılarda otokorelasyona rastlanmamıştır. Ancak şekil 4.22'de ARMA(1,1) modelinin kalıntı korelogramına bakıldığında temiz dizi olmadığı açıkça görülmüştür. Çizelge 4.2.9'da ARMA(2,1) modelinin tahmin sonuçları ve şekil 4.2.8'de ARMA(2,1) kalıntı korelogramı verilmiştir.

Çizelge 4.18. ARMA(2,1) model tahmin analizleri

Değişken	Katsayı	Standart hata	t istatistiği	Prob.
C	51.51517	4.516898	11.40499	0.0000
AR(1)	0.634354	0.505345	1.255288	0.2102
AR(2)	0.075234	0.394205	0.190849	0.8488
MA(1)	0.205726	0.504818	0.407526	0.6839
SIGMASQ	285.8827	16.92735	16.88880	0.0000
R^2	0.598742	Bağımlı değişken ortalaması		51.20724
Akaïke kriteri	8.523404			
Schwarz kriteri	8.576828			
Hannan-Quinn kriteri	8.544635			
F istatistiği	134.2946	Durbin-Watson istatistiği		2.003480
Prob(F istatistiği)	0.000000			

Date: 09/13/22 Time: 15:06
 Sample: 1/01/2014 1/31/2015
 Included observations: 365

Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob
		1 -0.008	-0.008	0.0241	0.877
		2 -0.042	-0.042	0.6857	0.710
		3 -0.096	-0.097	4.1213	0.249
		4 0.021	0.018	4.2924	0.368
		5 0.182	0.176	16.654	0.005
		6 -0.050	-0.056	17.589	0.007
		7 0.026	0.043	17.847	0.013
		8 0.034	0.067	18.274	0.019
		9 0.063	0.051	19.783	0.019
		10 -0.007	-0.029	19.803	0.031
		11 0.049	0.084	20.724	0.036
		12 0.020	0.015	20.883	0.052
		13 0.154	0.147	29.891	0.005
		14 0.023	0.025	30.088	0.007
		15 -0.033	-0.011	30.509	0.010
		16 -0.038	-0.042	31.074	0.013
		17 0.110	0.116	35.698	0.005
		18 0.115	0.055	40.832	0.002
		19 0.049	0.061	41.751	0.002
		20 -0.074	-0.060	43.882	0.002
		21 -0.036	-0.021	44.394	0.002
		22 0.087	0.030	47.328	0.001
		23 0.110	0.096	52.060	0.000
		24 0.055	0.027	53.260	0.001
		25 -0.012	0.021	53.316	0.001
		26 0.050	0.035	54.311	0.001
		27 0.001	-0.017	54.312	0.001
		28 -0.025	-0.059	54.556	0.002
		29 0.063	0.082	56.161	0.002
		30 0.065	0.021	57.876	0.002
		31 0.061	0.013	59.391	0.002
		32 -0.025	-0.025	59.636	0.002
		33 -0.035	0.001	60.119	0.003
		34 0.040	0.005	60.771	0.003
		35 0.076	0.039	63.127	0.002
		36 -0.000	-0.068	63.127	0.003

Şekil 4.23. ARMA(2,1) model kalıntı korelogramı

Çizelge 4.18’de ARMA(2,1) modelinin tahmin sonuçları verilmiştir. Sonuçlarda AR(1), AR(2), MA(1) katsayıları anlamlı çıkmamıştır. Durbin-Watson test istatistiğine bakılırsa kalıntılarda otokorelasyon mevcut olmadığı anlaşılmıştır. Şekil 4.23’te ise ARMA(2,1) modeli için kalıntıların korelogram grafiği çizilmiştir ve buradan yine temiz dizi olmadığı açıkça görülmüştür. Sonuç olarak ARMA(2,1) modeli de iyi bir model olmamıştır. ARMA(1,2) modeli tahmin sonuçları çizelge 4.19’da verilmiştir.

Çizelge 4.19. ARMA(1,2) model tahmin analizleri

Değişken	Katsayı	Standart hata	t istatistiği	Prob.
C	51.58019	5.384501	9.579381	0.0000

Çizelge 4.19'un devamı

AR(1)	0.832417	0.049681	16.75526	0.0000
MA(1)	-0.030597	0.068136	-0.449052	0.6537
MA(2)	-0.148353	0.059658	-2.486709	0.0133
SIGMASQ	284.8866	16.85308	16.90413	0.0000
R^2	0.600140	Bağımlı değişken ortalaması		51.20724
Akaike kriteri	8.519974			
Schwarz kriteri	8.573397			
Hannan-Quinn kriteri	8.541205			
F istatistiği	135.0788	Durbin-Watson istatistiği		1.960211
Prob(F istatistiği)	0.000000			

Date: 09/13/22 Time: 15:13
 Sample: 1/01/2014 1/31/2015
 Included observations: 365

Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob	
		1	0.013	0.013	0.0667	0.796
		2	-0.002	-0.003	0.0689	0.966
		3	-0.145	-0.145	7.8419	0.049
		4	-0.027	-0.024	8.1136	0.088
		5	0.132	0.135	14.572	0.012
		6	-0.074	-0.102	16.589	0.011
		7	0.012	0.006	16.639	0.020
		8	-0.001	0.042	16.639	0.034
		9	0.045	0.025	17.390	0.043
		10	-0.026	-0.050	17.638	0.061
		11	0.046	0.081	18.430	0.072
		12	0.012	0.013	18.484	0.102
		13	0.141	0.132	26.090	0.017
		14	0.004	0.007	26.097	0.025
		15	-0.037	-0.019	26.633	0.032
		16	-0.049	-0.031	27.543	0.036
		17	0.098	0.128	31.242	0.019
		18	0.099	0.051	35.006	0.009
		19	0.038	0.045	35.558	0.012
		20	-0.082	-0.061	38.173	0.008
		21	-0.047	-0.009	39.029	0.010
		22	0.073	0.048	41.118	0.008
		23	0.102	0.102	45.161	0.004
		24	0.058	0.026	46.492	0.004
		25	-0.014	0.019	46.573	0.006
		26	0.033	0.033	47.000	0.007
		27	-0.013	-0.010	47.071	0.010
		28	-0.031	-0.049	47.459	0.012
		29	0.056	0.094	48.690	0.012
		30	0.056	0.024	49.924	0.013
		31	0.053	0.007	51.065	0.013
		32	-0.033	-0.022	51.497	0.016
		33	-0.039	0.010	52.126	0.018
		34	0.025	0.006	52.382	0.023
		35	0.067	0.039	54.186	0.020
		36	-0.000	-0.062	54.186	0.026

Şekil 4.24. ARMA (1,2) model kalıntı korelogramı

ARMA(1,2) tahmin sonuçlarının verildiği çizelge 4.19 incelendiğinde MA(1) katsayısı anlamsız bulunmuştur. Durbin Watson test istatistiğine göre ise kalıntılarda otokorelasyona rastlanmamıştır. Şekil 4.24'de ARMA(1,2) modeli için oluşturulan kalıntı korelogramında kalıntıların temiz dizi olmadığı görülmüştür. Sonuç olarak ARMA(1,2) modeli de iyi bir tahmin olmamıştır. Son olarak ise 2014 HKI verilerine ARMA(2,2) modeli uygulanmıştır.

Çizelge 4.20. ARMA(2,2) model tahmin analizleri

Değişken	Katsayı	Standart hata	t istatistiği	Prob.
C	57.95337	12.03346	4.816019	0.0000
AR(1)	1.497855	0.080793	18.53944	0.0000

Çizelge 4.20'nin devamı

AR(2)	-0.503441	0.078150	-6.441997	0.0000
MA(1)	-0.716079	0.082030	-8.729514	0.0000
MA(2)	-0.197235	0.061253	-3.220016	0.0014
SIGMASQ	273.0207	15.60551	17.49515	0.0000
R^2	0.616795	Bağımlı değişken ortalaması		51.20724
Akaike kriteri	8.484307			
Schwarz kriteri	8.548415			
Hannan-Quinn kriteri	8.509784			
F istatistiği	115.5669	Durbin-Watson istatistiği		1.979954
Prob(F istatistiği)	0.000000			

Date: 09/13/22 Time: 15:23
 Sample: 1/01/2014 1/31/2015
 Included observations: 365

Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob	
		1	0.003	0.003	0.0032	0.955
		2	0.011	0.011	0.0482	0.976
		3	-0.095	-0.095	3.4001	0.334
		4	-0.013	-0.012	3.4610	0.484
		5	0.129	0.132	9.6086	0.087
		6	-0.102	-0.115	13.502	0.036
		7	-0.019	-0.025	13.641	0.058
		8	-0.023	0.008	13.835	0.086
		9	0.013	-0.005	13.901	0.126
		10	-0.057	-0.085	15.135	0.127
		11	0.008	0.038	15.160	0.175
		12	-0.028	-0.031	15.452	0.218
		13	0.108	0.094	19.852	0.099
		14	-0.021	-0.027	20.024	0.129
		15	-0.061	-0.053	21.450	0.123
		16	-0.073	-0.074	23.471	0.102
		17	0.068	0.088	25.229	0.090
		18	0.070	0.023	27.101	0.077
		19	0.014	0.022	27.175	0.101
		20	-0.099	-0.089	30.998	0.055
		21	-0.062	-0.037	32.511	0.052
		22	0.051	0.020	33.526	0.055
		23	0.071	0.079	35.488	0.046
		24	0.028	0.006	35.806	0.057
		25	-0.040	-0.008	36.438	0.065
		26	0.018	0.006	36.568	0.082
		27	-0.029	-0.031	36.900	0.097
		28	-0.054	-0.069	38.040	0.098
		29	0.032	0.074	38.448	0.113
		30	0.030	0.011	38.801	0.130
		31	0.034	-0.001	39.276	0.146
		32	-0.045	-0.034	40.098	0.154
		33	-0.051	0.000	41.167	0.155
		34	0.015	0.004	41.262	0.183
		35	0.053	0.039	42.391	0.182
		36	-0.020	-0.067	42.558	0.210

Şekil 4.25. ARMA(2,2) model kalıntı korelogramı

		t istatistiği	Prob.*
Augmented Dickey-Fuller test istatistiği		-6.922925	0.0000
Kritik değerler	1% level	-3.448161	
	5% level	-2.869285	
	10% level	-2.570963	

Çizelge 4.21 ARMA(2,2) model kalıntıları için ADF testi

Çizelge 4.20'de analiz sonuçlarına bakılırsa AR(1), AR(2), MA(1) ve MA(2) katsayıları anlamlı bulunmuştur ve Durbin-Watson test istatistiği ise 2 değerine çok yakın çıkmıştır buda bize kalıntılar arasında otokorelasyon olmadığını göstermiştir. Bu durumda iyi bir model olma koşullarından ilk ikisini sağlamıştır. Şekil 4.25'te verilen ARMA(2,2) modeli için kalıntıların korelogramını incelediğimizde değerler 0.01 değerinden büyük bulunmuştur. Bunun yanında kalıntıların ADF testide yapılmış ve

çizelge 4.21’de verilmiştir. Buradan kalıntıların durağan olduğu da görülmüştür. Dolayısı ile ARMA(2,2) modeli temiz dizi olma özelliği göstermiştir. Sonuç olarak ARMA(2,2) modeli kurduğumuz en iyi model olmuştur ve ARMA (2,2)’nin tahmin değerleri Çizelge 4.21’de verilmiştir.

Çizelge 4.22. ARMA(2,2) tahmin değerleri

43.72580925188715
50.14675609129683
53.42292691475138
55.09758996727097
55.95663475160637
56.40026578492558
56.63228266408107
56.75646840273442
56.82567390060654
56.8668135518499
56.89359392379099
56.91299566612601
56.92857433642579
56.94214130139443
56.95461971341027
56.96648040520654
56.97796386349241
56.98919326613983
57.00023204494972
57.01111319881169
57.02185422112925
57.03246470150138
57.04295019686622
57.05331420291746
57.06355915810891
57.07368695501661
57.08369920082853
57.09359735007958
57.10338277232865
57.11305678670451
57.12262067957937

5. SONUÇLAR

'Hava Kalitesi İndeksi''(HKİ) kirleticilerin belli bir bölgedeki hava kalitesinin karakterize edilmesi için ülkelerin kendi sınır değerlerine göre dönüştürdükleri ve kirlilik sınıflandırılmasının yapıldığı bir indekstir (Zencirci ve Işıklı 2017). Bulunulan bölgedeki hava kirliliğinin ne seviyede olduğunu, bu seviyede ortamda bulunan canlıların sağlığını etkileyebilecek düzeyde olup olmadığı hakkında bilgi sahibi olmamıza yardımcı olur. Hava kalite indeksi 0-500 arasında düzenlenmiş bir skaladır ve değerler yukarıya doğru çıktıkça havanın kalitesinin azaldığı düşünülmektedir. Bu durumda hava kalite indeksi ile havanın kalitesi arasında ters bir orantı olduğunu açıkça söyleyebiliriz (Zencirci ve Işıklı 2017). Çizelge 5.1'de EPA hava kalite indeksi kategorileri, sağlık endişe seviyeleri renklendirilmiş bir biçimde gösterilmiştir.

Bu yüksek lisans tez çalışmasının birinci amacı zaman serileri regresyonunun ve zaman serilerinde ARIMA modellerinin incelenmesi ve zaman serisine ait olan veriler ile uygulama yapılarak bu verilere en uygun modelin belirlenmesidir. Öncelikle zaman serisi regresyonu için 1928-2020 yılları arasında TÜİK'ten alınan Türkiye'deki ebe sayıları ile veri seti oluşturulmuştur. Bu veri setinin zaman yolu grafiği incelenmiş ve deterministik yapıda artan bir trendinin olduğu görülmüştür. Sonrasında ise veri seti 7 farklı regresyon modeli ile modellenmiştir ve aralarında veri setine en uygun model belirlenmiştir. Belirlenen en uygun model için 2021 ve 2022 yılları için ebe sayıları tahmin değerleri bulunmuştur.

ARIMA modellerinin uygulaması için 2014 yılına ait günlük olarak alınmış hava kalite indeksi verileri kullanılmıştır. Verilerin günlük alınmış olmasından dolayı tahmin değerleri 2015 yılının ocak ayı için hesaplanmıştır. Modelimizin öncelikle mevsimsel olup olmamasına bakılmış ve mevsimsel olmadığı anlaşılmıştır. Mevsimselliğe karar verildikten sonra ise veri setine uygulanan Philips-Perron ve Augmented Dickey-Fuller birim kök testleri sonucunda düzeyde durağan olduğu sonucuna ulaşılmıştır. Düzeyde durağan olması ise bizi ARIMA modellerinden ARMA(p,q) modeline yaklaştırmıştır. ARMA(p,q) modelleri ise sırası ile yapılmış sonuçları çizelge ve şekiller ile bulgular kısmında açıklanmıştır ve en uygun model ARMA(2,2) modeli olmuştur. Tüm katsayılar anlamlı bulunmuş, kalıntılar arasında otokorelasyon tespit edilmemiş ve temiz dizi olma özelliği sağlamıştır. Modelin

uygunluđuna karar verildikten sonra ise öngörü deđerleri hesaplanmıřtır ve bu deđerler çizelge 4.21’de verilmiřtir. Bulunan sonuçlar ise çizelge 5.1’de verilen EPA hava kalite indeksi tablosuna göre deđerlendirilmiřtir.

Çizelge 5.1 EPA hava kalite indeksi (Zencirci ve Iřıklı 2017)

Hava kalite indeksi	Sađlık endiře seviyesi	Anlamı
0-50	İyi	Hava kalitesi yeterli kabul edilir ve hava kirliliđi az riskli ya da risk oluřturmuyor.
51-100	Orta	Hava kalitesi kabul edilebilir düzeyde; ancak bazı kirleticiler için, hava kirliliđine çok hassas olan az sayıda kiři için orta düzeyde sađlık endiřesi olabilir
101-150	Hassas gruplar için sađlıksız	Hassas gruplarda sađlık etkileri oluřabilir. Nüfusun tamamının etkilenmesi olası deđildir.
151-200	Sađlıksız	Herkes sađlık etkileri yařamaya bařlayabilir, hassas gruplarda ise ciddi sađlık etkileri yařanabilir.
201-300	Çok sađlıksız	Sađlık alarmı: Herkes daha ciddi sađlık etkileri yařayabilir.
301-500	Tehlikeli	Sađlıkla ilgili acil durum: Nüfusun tamamının etkilenmesi olasıdır.

Bulgular kısmında verilen tanımlayıcı istatistiklerde HKI verilerinin ortalaması 48 olarak hesaplanmıřtır. 2015 yılının ocak ayına ait kurulan ARMA(2,2) modelinin öngörü deđerleri ise 43 ile 57 arasında seyretmiřtir. Dolayısı ile 2015 ocak ayı için tahmin edilen HKI verilerinin ortalamannın biraz üstünde seyrettiđini görebiliriz. Çizelge

5.1 de verilen EPA hava kalite indeksi tablosuna göre hava kalitesinin ortalamasının 2014 yılı için iyi bir düzeyde olduğunu söyleyebiliriz ki bu da bize çizelge 5.1’den hava iyi düzeyde olduğunu söylemiştir. Ancak 2015 yılının ocak ayı tahminleri ise yine çizelge 5.1 EPA hava kalite indeksi değerlerine göre artık hava kalitesinin orta seviyeye düştüğünü, hava kalitesinin kabul edilebilir düzeyde olmasına karşın çok hassas olan kişiler için orta seviyede de olsa bir sağlık endişesi oluşturduğunu bize açıkça ifade etmiştir.

Zaman serileri farklı disiplinlerde oldukça yaygın olarak kullanılmaktadır. Matematiksel modelleme üzerine yapılan bu inceleme literatüre fayda sağlayacak niteliktedir. Bu çalışmanın uygulaması zaman serisi regresyonu için yıllık olarak alınmış Türkiye’deki ebe sayıları ile veri seti oluşturularak, ARIMA modelleri uygulaması ise günlük olarak alınmış hava kalite indeksi verileri ile yapılmıştır. Zaman serisi regresyonu verileri yıllık olarak alındığından tahmin değerleri 2 yıllık olarak hesaplanmıştır. 2 yıl, tahmin yapmak için önemli bir zaman dilimidir. Bu çalışmanın farklı veri türleri kullanılarak yapılması mümkündür. Bu açıdan bir kaynak niteliği taşımaktadır.

ARIMA modelleri uygulamasında ise veri seti günlük olarak alındığından tahmin değerleri 1 aylık olarak hesaplanmıştır. Hava kalite indeksi alınarak yapılan bu çalışma genişletilerek yıllık veri toplandığında yıllık bir zaman serisi uygulaması ve yine yıllık tahminler ortaya çıkacaktır. Buda bize ilerleyen süreçlerde hava kalitesi hakkında oldukça önemli öngörü değerleri sunacaktır. Yapılan tahminler sonucunda hava kalite indeksinin risk oluşturduğu düzeylerde yetkili birimlerin hava kalitesini artıracak nitelikte önlemler almasına fayda sağlayacaktır. Modelleme için yapılan çalışmada kullanılan veri daha öncesinde tez danışmanın da dahil olduğu proje ve projeden üretilmiş makaleden alınmıştır. Bu veri üzerine bir zaman serisi modellemesi çalışılmamıştır.

6. KAYNAKLAR

- Akdi, Y. (2012). Zaman serileri analizi:(birim kökler ve kointegrasyon). Gazi Kitabevi.
- Akmut, Ö., Aktaş, R., & Binay, H. S. (1999). Öngörü teknikleri ve finans uygulamaları: Ankara: Ankara Üniversitesi Siyasal Bilgiler Fakültesi.
- Anonim1:https://en.wikipedia.org/wiki/George_E._P._Box [Son erişim tarihi: 26.05.2022]
- Anonymous 1: <https://www.wunderground.com/> [Son erişim tarihi: 26.05.2022].
- Anonim 3: <https://www.tuik.gov.tr/> [Son erişim tarihi: 26.05.2022].
- Anonim 4: <https://tr.wikipedia.org/wiki/Ortalama> [Son erişim tarihi: 06.09.2022].
- Anonim5: <https://sim.csb.gov.tr/Home/HKI> [Son erişim tarihi: 16.09.2022].
- Duru, Ö. (2007). Zaman serileri analizinde ARIMA modelleri ve bir uygulama. Yüksek Lisans Tezi, İstanbul Üniversitesi, İstanbul, 103 s.
- Durbin, J. (1960). Estimation of parameters in time-series regression models. Journal of the royal statistical society: Series B (Methodological), 22(1), 139-153.
- Genceli, M. (1973). İki değişkenli doğrusal regresyonda zaman faktörü. İstanbul Üniversitesi İktisat Fakültesi Mecmuası, 33(1-4).
- Granger, C. W., & Joyeux, R. (1980). An introduction to long-memory time series models and fractional differencing. Journal of time series analysis, 1(1), 15-29.
- Gujarati, D. N., Porter, D. C., Şenesen, Ü., & Günlük-Şenesen, G. (2012). Temel ekonometri. Literatür Yayıncılık.
- Güneş, M. (2010). Zaman serilerinde müdahale analiz teknikleri üzerine bir çalışma (Master's thesis, Namık Kemal Üniversitesi).
- Hosking, J. R. M. (1981). Fractional differencingi Biometrika 68.
- Kaya, E. (2019). Zaman serileri analizinde box-jenkins yöntemi ile savunma sanayi verileri üzerine bir uygulama (Master's thesis, Sosyal Bilimler Enstitüsü)
- Kiliç, S. (2013). Linear regression analysis. Psychiatry and Behavioral Sciences, 3(2), 90.
- Mert, M., & Çağlar, A. E. (2019). Eviews ve Gauss uygulamalı zaman serileri analizi. Ankara: Detay Yayıncılık.
- Nicholas, J. (1990). Tanımlayıcı istatistiklere giriş .Matematik Öğrenme Merkezi, Sidney Üniversitesi.
- Özmen, A., & Poyraz, K. (1999). Zaman serilerini mevsimlik etkiden arındırmada uygun tekniğin belirlenmesine ilişkin bir yaklaşım. Dumlupınar Üniversitesi Sosyal Bilimler Fakültesi, 3, 167-186.
- Öznur, İ. (2012). Doğrusal Olmayan Otoresif Zaman Serileri Modellerinin Kestirimi. Muğla Üniversitesi Sosyal Bilimler Enstitüsü Dergisi, (28), 205-218
- Roussel, Y., Ali, A., & Audi, M. (2021). Measuring the money demand in Pakistan: a time series analysis. Bulletin of Business and Economics (BBE), 10(1), 27-41.

- Roy, S., Nanjiba, S., & Chakrabarty, A. (2018, December). Bitcoin price forecasting using time series analysis. In 2018 21st International Conference of Computer and Information Technology (ICIT) (pp. 1-5). IEEE.
- Sevüktekin, M., & Nargeleçekenler, M. (2007). Ekonometrik zaman serileri analizi (2. Baskı). Ankara: Nobel Yayın Dağıtım.
- Taheri Tizro, A., Ghashghaie, M., Georgiou, P., & Voudouris, K. (2014). Su kalitesi parametrelerinin zaman serisi analizi. Su ve Atıksuda Uygulamalı Araştırmalar Dergisi , 1 (1), 40-50.
- Ünsal, A. (1997). Zaman serilerinde regresyon ve varyans analizi yöntemleri ile mevsimsel dalgalanmaların araştırılması ve bir uygulama. Ekonomik Yaklaşım, 8(26), 119-130.
- Yalçın, F., Tepe, AM, Doğan, G., & Çizmeci, N. (2021). Türkiye'deki Şehirlerde AQI Değerlerinin Açıklanmasında Mekansal Regresyon Modelleri. Kocaeli Fen Bilimleri ve Mühendislik Dergisi,4(1), 1-15.
- Zencirci SA, Işıklı B. Hava Kirliliği. Türk Dünyası Uygulama Ve Araştırma Merkezi Halk Sağlığı Dergisi. 2017; 2(2),24-36.
- Qi, H., Xiao, S., Shi, R., Ward, M. P., Chen, Y., Tu, W., ... & Zhang, Z. (2020). COVID-19 transmission in Mainland China is associated with temperature and humidity: A time-series analysis. Science of the total environment,728, 138778

ÖZGEÇMİŞ

Yağmur KARACA ÜLKÜTANIR

krcyagmur26@gmail.com



ÖĞRENİM BİLGİLERİ

Yüksek Lisans 2018-2022	Akdeniz Üniversitesi Fen Bilimleri Enstitüsü, Matematik A.B.D, Antalya
Lisans 2013-2018	Akdeniz Üniversitesi Fen Fakültesi, Matematik Bölümü, Antalya

ESERLER

Ulusal bilimsel toplantılarda sunulan ve bildiri kitaplarında basılan bildiriler

1- Yalcin F., Ülkütanir Y., Özer Ö. (2019). İstatistiksel Analizde Bazı Parametrik Testler İçin Ampirik Bir Uygulama: Akademik Başarı Değerlendirmesi International Symposium on Advanced Engineering Technologies. ISADET, Kahramanmaraş, Türkiye.