

**T.C.
ONDOKUZ MAYIS ÜNİVERSİTESİ
LİSANSÜSTÜ EĞİTİM ENSTİTÜSÜ
İSTATİSTİK ANA BİLİM DALI**



**FARKLI ÜLKELERDE COVID 19 VAKA SAYILARININ
TAHMİNİNDE KULLANILAN ZAMAN SERİLERİNİN
KARŞILAŞTIRILMASI**

Yüksek Lisans Tezi

Mohammed Zaki NOORİ

Danışman

Prof.Dr. Mehmet Ali CENGİZ

SAMSUN
2022

TEZ KABUL VE ONAYI

Mohammed Zaki NOORİ tarafından, **Prof. Dr. Mehmet Ali CENGİZ** danışmanlığında hazırlanan “**Farklı Ülkelerde Covid 19 Vaka Sayılarının Tahmininde Kullanılan Zaman Serilerinin Karşılaştırılması**” başlıklı bu çalışma, jürimiz tarafından 16.2.2022 tarihinde yapılan sınav sonucunda oy birliği ile başarılı bulunarak Yüksek Lisans Tezi olarak kabul edilmiştir.

	Unvanı Adı Soyadı Üniversitesi Ana Bilim/Ana Sanat Dalı	İmza	Sonuç
Başkan	Prof.dr. Mehmet ali cengiz Ondokuz mayıs üniversitesi Fen-edebiya fakültesi/ istatistik bölümü	☒ ☐	Kabul Ret
Üye	Doç.dr. Erol terzi Ondokuz mayıs üniversitesi Fen-edebiyat fakültesi / istatistik bölümü	☒ ☐	Kabul Ret
Üye	Dr. Öğr. Üyesi Hamza ERDOĞDU Harran Üniversites İktisadi ve İdari Bilimler Fakültesi Ekonometri Bölümü	☒ ☐	Kabul Ret

Bu tez, Enstitü Yönetim Kurulunca belirlenen ve yukarıda adları yazılı jüri üyeleri tarafından uygun görülmüştür.

ONAY

... / ... / ...

Prof. Dr. Ali BOLAT
Enstitü Müdürü

BİLİMSEL ETİĞE UYGUNLUK BEYANI

Hazırladığım yüksek lisans tezinin bütün aşamalarında bilimsel etiğe ve akademik kurallara riayet ettiğimi, çalışmada doğrudan veya dolaylı olarak kullandığım her alıntıya kaynak gösterdiğimi ve yararlandığım eserlerin Kaynaklar’da gösterilenlerden oluştuğunu, her unsurun enstitü yazım kılavuzuna uygun yazıldığını ve TÜBİTAK Araştırma ve Yayın Etiği Kurulu Yönetmeliği’nin 3. bölüm 9. maddesinde belirtilen durumlara aykırı davranılmadığını taahhüt ve beyan ederim.

Etik Kurul Gerekli mi ?

Evet ☐ (Gerekli ise ekler kısmına ekleyiniz)

Hayır ☒

İmza

28 / 02 / 2022

Mohammed Zaki NOORİ

TEZ ÇALIŞMASI ÖZGÜNLÜK RAPORU BEYANI

Tez Başlığı: “Farklı Ülkelerde Covid 19 Vaka Sayılarının Tahmininde Kullanılan Zaman Serilerinin Karşılaştırılması”

Yukarıda başlığı belirtilen tez çalışması için şahsım tarafından 12.01.2022 tarihinde intihal tespit programından alınmış olan özgünlük raporu sonucunda;

Benzerlik oranı: % 6

Tek kaynak oranı: %1 çıkmıştır.

İmza

28 / 02 / 2022

Prof. Dr. MEHMET ALİ CENGİZ

ÖZET

FARKLI ÜLKELERDE COVID 19 VAKA SAYILARININ TAHMİNİNDE KULLANILAN ZAMAN SERİLERİNİN KARŞILAŞTIRILMASI

Mohammed Zaki NOORİ

Ondokuz Mayıs Üniversitesi

Lisansüstü Eğitim Enstitüsü

İstatistik Ana Bilim Dalı

Yüksek Lisans, Ocak/ 2022

Danışman: Prof.Dr. Mehmet Ali CENGİZ

Ülkelerdeki aktif COVID-19 vaka oranları ve dolayısıyla salgının ne zaman zirve yapacağını bilmek, hükümetlerin sosyal ve sağlık politikalarını değiştirmesine izin vereceği için pandeminin etkisini önemli ölçüde azaltma potansiyeline sahiptir. COVID-19 pandemisi, mevcut dünyamızda büyük bozulmalara neden oldu. Bu krizin etkisi ve büyüklüğü, dünya çapında çok fazla korku, belirsizlik ve endişeye neden oldu. Bu nedenle, gelecekte kayıpları önlemek için COVID-19 pandemisini ve toplumumuz üzerindeki etkilerini anlamak kritik öneme sahiptir.

Bu tez çalışmasında; on farklı ülke için (Türkiye, Brezilya, Rusya, Avustralya, İran, Amerika, ABD, Çin, Fransa ve Almanya) COVID-19 vakalarına ilişkin dört farklı veri seti (vakalar - kümülatif vakalar - yeni ölümler - kümülatif ölümler), zaman serisi kullanılarak incelenmiştir. Herbir ülke için dört farklı zaman serisi verileri ayrı ayrı modellenmiş ve birbirleriyle karşılaştırılmıştır. Ayrıca on farklı ülke için dört farklı zaman serisi modelleri, ülke bazlı olarak birbirleriyle de karşılaştırılmıştır.

Anahtar Sözcükler: COVID-19, zaman serisi, ARIMA, Vaka sayısı,

ABSTRACT

COMPARISON OF TIME SERIES USED TO ESTIMATE COVID 19 CASE NUMBERS IN DIFFERENT COUNTRIES

Mohammed Zaki NOORİ

Ondokuz Mayıs University

Institute of Graduate Studies

Department of Statistics

Master, January/2022

Supervisor: Prof. Dr. Mehmet Ali CENGİZ

Knowing the active COVID-19 case rates in countries and when the epidemic will peak, have the potential to significantly reduce the impact of the pandemic, as it will allow governments to change their social and health policies. The COVID-19 pandemic has caused major disruptions in our current world. The impact and magnitude of this crisis has caused a lot of fear, uncertainty and concern around the world. Therefore, it is critical to understand the COVID-19 pandemic and its effects on our society in order to prevent losses in the future.

In this thesis, in ten different countries (Turkey, Brazil, Russia, Australia, Iran, America, China, France and Germany) COVID-19 cases in relation to four different data sets (cases cumulative cases - deaths - cumulative deaths), has been investigated by using the time series. Four different time series data for each country were modelled separately and compared with each other. In addition, four different time series models for ten different countries were compared with each other on a country-by-country basis.

Keywords: Covid-19, Time series, ARIMA, the number of cases

TEŞEKKÜR

Tez çalışma sürecimde her daim yanımda olan, karşılaştığım sorunların çözümünde tüm ilgisiyle ve hoşgörüsüyle yol gösteren danışmanım Sayın Prof. Dr. Mehmet Ali CENGİZ'e sonsuz teşekkürlerimi sunarım.

Bu süreçte beni hep motive eden, sonsuz sabrıyla yanımda olan ve destekleyen, aileme çok teşekkür ederim.

Mohammed Zaki NOORİ

İÇİNDEKİLER

TEZ KABUL VE ONAYI	i
BİLİMSEL ETİĞE UYGUNLUK BEYANI	ii
TEZ ÇALIŞMASI ÖZGÜNLÜK RAPORU BEYANI	ii
ÖZET	iii
ABSTRACT	iv
TEŞEKKÜR	v
İÇİNDEKİLER	vi
KISALTMALAR	vii
TABLolar LİSTESİ.....	viii
ŞEKİLLER LİSTESİ.....	ix
1. GİRİŞ	1
1.1. COVID 19 Pandemisi	2
1.2. COVID 19 Pandemisine Yönelik Tahmin Modellemesi Yapan Çalışmalar	3
2. ZAMAN SERİLERİ	12
2.1. Zaman Serilerinin Tanılanması.....	12
2.2. Zaman Serisinin Bileşenleri.....	12
2.2.1. Trend (T)	12
2.2.2. Mevsimsel Dalgalanmalar (M).....	13
2.2.3. Konjonktürel Değişimler (K)	14
2.2.4. Düzensiz Hareketler (D).....	15
2.3. Durağanlık	16
2.3.1. Otokorelasyon Fonksiyonu (ACF)	16
2.3.2. Kısmi Otokorelasyon Fonksiyonu (PACF)	17
2.3.3. Beyaz Gürültü Süreci	18
2.3.4. Durağanlığın Belirlenmesi	18
2.3.4.1. Grafik Analizleri	18
2.3.4.2. Korelogramlar	19
2.3.4.3. Birim Kök Testleri	19
2.4. Zaman Serileri Analizleri	20
2.5. BOX-Jenkins Yöntemi	22
2.5.1. ARIMA (Auto Regressive Integrated Moving Average)	23
2.5.1.1. Model Belirleme	25
2.5.1.2. Parametrelerin Tahmin edilmesi	30
2.5.1.3. Teşhis ve Denetim Yapılması	31
2.5.1.4. Öngörü (Tahmin)	32
3. UYGULAMA.....	33
3.1. Genel Bilgi.....	33
3.2. Veri Seti.....	33
3.3. Bulgular	34
3.3.1. İran İçin Model Belirleme ve Tahmin.....	35
3.3.2. Almanya İçin Model Belirleme ve Tahmin.....	41
3.3.3. Rusya İçin Model Belirleme ve Tahmin	45
3.3.4. Suudi Arabistan İçin Model Belirleme ve Tahmin.....	48
3.3.5. Türkiye İçin Model Belirleme ve Tahmin.....	52
3.3.6. ABD İçin Model Belirleme ve Tahmin	56
3.3.7. Avustralya İçin Model Belirleme ve Tahmin.....	60
3.3.8. Brezilya İçin Model Belirleme Ve Tahmin	64
3.3.9. Çin İçin Model Belirleme ve Tahmin.....	67
3.3.10. Fransa İçin Model Belirleme ve Tahmin.....	71
4. SONUÇ	80
KAYNAKLAR	83
ÖZGEÇMİŞ.....	87

KISALTMALAR

ACF	: Otokorelasyon Fonksiyonu
ADF	: Geniřletilmiş Dickey-Fuller Birim K�k Testi
AR(p)	: Otoregresif S�re�
ARIMA	: Otoregresif Entegre Hareketli Ortalama S�reci
ARMA	: Otoregresif Hareketli Ortalama S�reci
MA (q)	: Hareketli Ortalama S�reci
MAE	: Ortalama Mutlak Hata
MAPE	: Ortalama Mutlak Hata Y�zdesi
MSE	: Hata Kareler Ortalaması
PACF	: Kısmi Oto Korelasyon Fonksiyonu
RMSE	: Hata Kareler Ortalamasının Kare K�k�

TABLÖLAR LİSTESİ

Tablo 3.1. Zaman serileri için çeşitli Box-Jenkins modelleri	35
Tablo 3.2. Olağan modelin parametreleri	35
Tablo 3.3. COVID-19 salgını gözlemlerinin tahmin değerlerinin, üst ve alt limitleri	40
Tablo 3.4. Zaman serileri için ilgili Box-Jenkins modelleri	41
Tablo 3.5. Seçilen modelin parametreleri	41
Tablo 3.6. Gözlemlerin tahmin değerlerinin üst ve alt sınırları	44
Tablo 3.7. Zaman serileri için ilgili Box-Jenkins modelleri	45
Tablo 3.8. Seçilen modelin parametreleri	45
Tablo 3.9. Gözlemlerin tahmin değerlerinin, üst ve alt sınırları	48
Tablo 3.10. Önerilen Box - Jenkins modelleri	48
Tablo 3.11. Seçilen modelin parametreleri	49
Tablo 3.12. Gözlemlerin tahmin değerlerinin, üst ve alt sınırları	51
Tablo 3.13. Önerilen Box - Jenkins modelleri	52
Tablo 3.14. Seçilen modelin parametreleri	52
Tablo 3.15. Gözlemlerin tahmin değerlerinin, üst ve alt sınırları	55
Tablo 3.16. Önerilen Box-Jenkins modelleri	56
Tablo 3.17. Seçilen modelin parametreleri	56
Tablo 3.18. Gözlemlerin tahmin değerlerinin, üst ve alt sınırları	59
Tablo 3.19. Önerilen Box-Jenkins modelleri	60
Tablo 3.20. Seçilen modelin parametreleri	60
Tablo 3.21. Gözlemlerin tahmin değerleri	63
Tablo 3.22. Önerilen Box Jenkins modelleri	64
Tablo 3.23. Seçilen modelin parametreleri	64
Tablo 3.24. Gözlemlerin tahmin değerleri	67
Tablo 3.25. Önerilen Box-Jenkins modelleri	68
Tablo 3.26. Seçilen modelin parametreleri	68
Tablo 3.27. Gözlemlerin tahmin değerlerinin, üst ve alt limiti	71
Tablo 3.28. önerilen Box-Jenkins modelleri	71
Tablo 3.29. Seçilen modelin parametreleri	72
Tablo 3.30. Gözlemlerin tahmin değerlerinin, üst ve alt limiti	74
Tablo 3.31. Yeni vakalar	76
Tablo 3.32. Kümülatif vakalar	77
Tablo 3.33. Yeni ölümler	78
Tablo 3.34. Kümülatif ölümler	79

ŞEKİLLER LİSTESİ

Şekil 2.1. Doğrusal ve doğrusal olmayan trend sergileyen zaman serisi örneği.....	13
Şekil 2.2. 1990:01 – 1994: 12 yıllarına ait çimento üretim miktarı.	14
Şekil 2.3. 1989 – 2008 yıllarına ait GSYİH verisi	15
Şekil 2.4. Düzensiz hareketler sergileyen zaman serisi örneği	15
Şekil 2.5. Türkiye'nin eğitim harcamaları ve Gayri Safi Milli Hasıla verilerine ait zaman serisi grafikleri	19
Şekil 2.6. ARIMA modeli temel işleyişi	24
Şekil 3.1. Otokorelasyon işlevi (ACF)	36
Şekil 3.2. Kısmi otokorelasyon (PACF)	36
Şekil 3.3. Otokorelasyon katsayıları	37
Şekil 3.4. Kısmi otokorelasyon katsayıları	38
Şekil 3.5. Bir serinin tahmin edilen değerleri ile orijinal değerleri arasındaki fark ...	39
Şekil 3.6. Otokorelasyon	42
Şekil 3.7. Kısmi otokorelasyon	42
Şekil 3.8. Otokorelasyon ve kısmi otokorelasyon katsayıları	43
Şekil 3.9. Bir serinin tahmin edilen değerleri ile orijinal değerleri arasındaki fark ...	44
Şekil 3.10. Otokorelasyon	46
Şekil 3.11. Kısmi otokorelasyon	46
Şekil 3.12. Seri için tahmin değerleri	47
Şekil 3.13. Bir serinin tahmin edilen değerleri ile orijinal değerleri arasındaki fark.	47
Şekil 3.14. Otokorelasyon	49
Şekil 3.15. Kısmi otokorelasyon	49
Şekil 3.16. Seri için tahmin değerleri	50
Şekil 3.17. Bir serinin tahmin edilen değerleri ile orijinal değerleri arasındaki fark.	51
Şekil 3.18. Otokorelasyon	53
Şekil 3.19. Kısmi otokorelasyon	53
Şekil 3.20. Seri için tahmin değerleri	54
Şekil 3.21. Bir serinin tahmin edilen değerleri ile orijinal değerleri arasındaki fark.	55
Şekil 3.22. Otokorelasyon	57
Şekil 3.23. Kısmi otokorelasyon	57
Şekil 3.24. Seri için tahmin değerleri	58
Şekil 3.25. Bir serinin tahmin edilen değerleri ile orijinal değerleri arasındaki fark.	59
Şekil 3.26. Otokorelasyon	61
Şekil 3.27. Kısmi otokorelasyon	61

Şekil 3.28. Seri için tahmin edilen değerler	62
Şekil 3.29. Bir serinin tahmin edilen değerleri ile orijinal değerleri arasındaki fark. 63	
Şekil 3.30. Otokorelasyon	65
Şekil 3.31. Kısmi otokorelasyon	65
Şekil 3.32. Seri için tahmin değerleri.....	66
Şekil 3.33. Bir serinin tahmin edilen değerleri ile orijinal değerleri arasındaki fark. 67	
Şekil 3.34. Otokorelasyon	69
Şekil 3.35. Kısmi otokorelasyon	69
Şekil 3.36. Seri için tahmin değerleri.....	70
Şekil 3.37. Bir serinin tahmin edilen değerleri ile orijinal değerleri arasındaki fark. 70	
Şekil 3.38. Otokorelasyon	72
Şekil 3.39. Kısmi otokorelasyon	73
Şekil 3.40. Seri için tahmin değerleri.....	73
Şekil 3.41. Bir serinin tahmin edilen değerleri ile orijinal değerleri arasındaki fark. 74	

1. GİRİŞ

Corona Virüs Hastalığı 2019 (COVID-19), 31 Aralık 2019'da Çin'in Hubei kentinde rapor edilmiş ve Dünya Sağlık Örgütü (DSÖ) tarafından bir ay sonra küresel bir pandemik hastalık olarak ilan edilmiştir. Enfeksiyon endişe verici bir oranda hızla tüm dünya ülkelerine yayılmıştır. COVID-19 ile enfekte olan kişilerde çok az semptom olabilmekte veya hiç olmayabilmektedir. Semptomlar hafif semptomlardan ciddi hastalıklara kadar değişebilmekte ve COVID-19'un kuluçka süresi 2 hafta veya daha uzun sürebilmektedir. Gizli enfeksiyon döneminde hastalık hala bulaşıcı olabilmekte ve virüs solunum damlacıkları ve kişiden kişiye yakın temas yoluyla yayılabilmektedir (Gebretensae ve Asmelash, 2021).

Pandemi ile mücadelede, salgının yayılma hızını tespit edebilmek için çok önemlidir. Herhangi bir zamanda yayılma seviyelerinin farkındalığı, hükümetlerin pandeminin sonuçlarıyla baş etmek için halk sağlığı politikalarını planlamasına ve formüle etmesine yardımcı olabilir. Yayılma ve zirve zamanlarını anlamının tek yolu, herhangi bir zamanda aktif vaka sayısını doğru bir şekilde tahmin edebilmektir. (Guo, vd., 2020).

Bu nedenlerle farklı tekniklerle, araştırmalar yapılarak salgınla ilgili tahmin çalışmaları gerçekleştirilmektedir. Matematiksel modellerin bulaşıcı hastalıkların kontrolünü ve yayılmasını analiz etmek için en iyi teknoloji olduğu ifade edilmektedir. Zaman serisi analizi, matematiksel modelin zaman içinde gözlemlenen tarihsel değerlerin düzenliliğine ve eğilimine dayalı olarak oluşturulduğu ve COVID-19'un yayılmasını tahmin etmede yaygın olarak kullanılan bir araçtır (Fanelli ve Piazza, 2020; Thompson, vd., 2020).

Model tabanlı tahmin, çeşitli yazarlar tarafından araştırılmıştır. Örneğin, COVID-19 verilerini zaman serisi olarak ele alan Tandon, vd. (2020) ve Bayyurt ve Bayyurt (2020) gelecekteki enfekte vakaları ve ölümü tahmin etmek için sırasıyla otoregresif entegre hareketli ortalama (ARIMA) modellerini uygulamışlardır. Petropoulos ve Makridakis (2020), kümülatif enfeksiyon sayısını modellemek için üstel yumuşatma yönteminden faydalanmışlardır. Sidner et al. (2020) Sosyal mesafenin ABD'deki COVID-19 salgınına yavaşlatmaya yardımcı olduğunu göstermek için zaman serisi yöntemlerini kullanmışlardır. Diğer taraftan, salgın modelleme de geniş çapta ele alınmaktadır. Örneğin, Fanelli ve Piazza (2020), Çin, İtalya ve Fransa'da COVID-19'un

yayılmasını analiz etmek ve tahmin etmek için duyarlı enfekte kaldırılmış (SIR) modelini kullanmışlardır. Wang vd. (2020) Çin'in Wuhan kentinde COVID-19'un patlak vermesine yönelik ilaç dışı müdahaleyi değerlendirmek için SIR modellerini genişletmişlerdir.

Hastalığı modellemek ve günlük olası vaka sayısına ilişkin geleceğe yönelik tahminler sağlamak, sağlık sisteminin yeni hastalara hazırlanmasına yardımcı olabilir. Bu nedenle, küresel pandemik tehditleri tahmin etmek ve izlemek için istatistiksel tahmin modelleri kullanılabilir. Bu nedenle, politika yapıcılara ve sağlık personeline yardımcı olmak için hem hesaplama açısından yetkin hem de pratik modeller oluşturmak son derece önemlidir (Fanelli ve Piazza, 2020; Thompson, vd., 2020).

1.1. COVID 19 Pandemisi

Coronavirüs 1965 yılından beri keşfedilen büyük bir virüs grubudur; bugüne kadar 7 türün insanları etkilediği rapor edilmiştir. Bu virüslerin 3 genotipi alfa, beta ve gamadır. Bu hastalıkların doğal rezervuarları memeliler ve kuşlardır ve bu nedenle zoonotik hastalıklar olarak kabul edilirler (Anderson, vd., 2004; Peeri, vd., 2020). Şiddetli akut solunum sendromuna (SARS), insanları, yarasaları ve diğer bazı memelileri enfekte eden ve 2002 ve 2003 yıllarında salgınlara yol açan bir koronavirüs türü neden olmuştur (Zhau ve Yan, 2003). Orta Doğu solunum sendromu (MERS), 2012 senesinde Suudi Arabistan'da salgına sebep olmuştur (Mobaraki and Ahmadzadeh, 2019).

SARS'ın ortaya çıkmasından sonra MERS, büyük bir küresel halk sağlığı krizine neden olan ikinci koronavirüs oldu. İlk kez 2012 yılında Suudi Arabistan'da 60 yaşındaki bir erkeğin şiddetli zatürree şikayeti ile ortaya çıkmıştır (Mackay ve Arden, 2015). Virüs salgını 2 yıl sonrasına kadar, 2014 yılında, tespit edilen toplam 662 vaka ve %32,97 vaka-ölüm oranı gerçekleşmiştir (Al-Omari, vd., 2019). 2014 yılından 2016' yılına kadar Suudi Arabistan'da 1364 ayrı vaka kaydedilmiştir. Avrupa, Asya, Orta Doğu ve Kuzey Amerika'yı kapsayan salgın sırasında MERS'ten toplam 27 ülke etkilenmiştir. Güney Kore'de bir süper yayıcı sonucu 186 kişinin enfekte olduğu salgın da dahil olmak üzere Orta Doğu dışında tespit edilen vakalar, daha önce Orta Doğu'da enfekte olmuş bireylere bulaşmıştır (Liu vd., 2020). 2012'den bu yana, 2494 laboratuvar tarafından doğrulanmış MERS vakası rapor edilmiş ve buna bağlı 858 ölüm meydana gelmiştir (%34,4 vaka-ölüm oranı). (WHO, 2019; Al-Omari, vd., 2019). Son olarak 2019 yılı Aralık ayında Çin'in Whan eyaletinde yeni tip koronavirüs olarak nitelendirilen ve daha sonra COVID-19 salgınına dönüşen virüs tespit edilmiştir.

COVID-19, soğuk algınlığı semptomları, solunum/ateş semptomları, öksürük, nefes darlığı, böbrek yetmezliği ve hatta ölüm dahil olmak üzere çok çeşitli semptomlarla solunum yolu hastalıklarının doğrudan ve dolaylı olarak bulaşmasına neden olur (Liu, vd., 2020). COVID-19'un kesin klinik belirtileri henüz bilinmemektedir ve bulaşma riski tam olarak anlaşılamamıştır (Adhikari ve ark., 2020). Ancak virüsün esas olarak temas, damlacıklar, solunum ve dışkı yoluyla bulaştığına inanılmaktadır. Genel olarak herkes bu viral hastalığa karşı hassastır. Raporlara göre, COVID-19'un ortalama kuluçka süresi 5,2 gün (4,1-7 gün) ve temel üreme sayısı (R_0) 2,2'dir (%95 GA: 1,4 ila 3,9) (Li ve diğerleri, 2020). Başka bir çalışmada, ortalama kuluçka süresi 0-24 gün, ortalama 6.4 gündür olarak belirtilmiştir. Tahmin modelinden bağımsız olarak, erken COVID-19'un R_0 'ı SARS ve MERS'den daha yüksektir ve çoğu hasta (%80,9) asemptomatik veya hafif pnömoni olarak kabul edilir (Lipsitch ve ark., 2020). Ölüm oranları sırasıyla %2 (12), %2.3 (8) ve %3.46'dır. Altta yatan hastalıkları olan yaşlı erkeklerin ölüm riski daha yüksektir (Wang ve ark., 2020). 29 Mart 2020 itibariyle, Dünya Sağlık Örgütü (WHO) 100'den fazla ülkede/bölgede (en yaygın olarak Amerika Birleşik Devletleri, İtalya, Çin, İspanya, Almanya, İran ve Fransa'da) bir COVID-19 pandemisi ilan etmiştir (Worldometer, 2020).

Türkiye'de ise ilk koronavirüs vakası 11 Mart 2020'de Sağlık bakanlığı tarafından bildirilmiştir. O günden itibaren inişli, çıkışlı bir seyir gösteren salgından korunmak için kısıtlamalar ve önlemler alınmıştır. Ancak yine de salgın zaman zaman düşüş seyrine geçse de yayılımı engellenememiştir. 29 Ekim 2021 tarihi itibariyle 7.601.626 vaka sayısının olduğu Türkiye'de aktif vaka sayısı 214.089, hasta sayısı 8.744 olup, salgının başlangıcından itibaren ölenlerin sayısı 67.225'dir (TC. Sağlık Bakanlığı, 2021).

1.2. COVID 19 Pandemisine Yönelik Tahmin Modellemesi Yapan Çalışmalar

Yeni koronavirüsün (COVID-19) patlak vermesi dünya çapında ilgi görmüştür. Başta sağlık sektörü olmak üzere küresel ekonomiler için önemli bir zorluk teşkil etmiştir. Sağlam bir sağlık sistemine sahip olsalar bile ülkeler COVID-19'un etkilerine hazırlıksız yakalanmışlardır. Enfeksiyonun yayılmasını analiz etmek için COVID-19 salgınının çeşitli istatistiksel, dinamik ve matematiksel modeller geliştirilmiştir. Dünyanın farklı ülkelerindeki araştırmacılar farklı veriler ve yöntemler kullanarak COVID-19 salgınına yönelik tahminler gerçekleştirmektedirler. Bu bölümde COVID-19 salgınına yönelik tahmin modellemesi yapan araştırmalardan bahsedilmektedir.

Maurya, vd. (2020), COVID-19 pandemik hastalığı için mevcut verilerden yola çıkarak yedi tahmin yöntemi kullanmıştır. Ancak en az hataya sahip dört yöntemi rapor etmiştir. Bu yöntemler: NAIVE, Holt'un lineer trend yöntemi, Holt'un kış mevsimi yöntemi ve ARIMA'dır. Çalışma, MSE değerine göre NAIVE modelinin en iyi model olduğunu, çünkü bu modelin hatasının diğer modellere göre çok düşük olduğunu bulmuştur. Diğer yöntemler, veri noktası çok az olduğu için salgının tahmini için daha az uygun görülmüştür.

Chandu (2020), Nepal'de teyit edilmiş COVID-19 vaka sayısını tahmin etmek için 14 günlük bir zaman serisi analizi yapmışlardır. 18 Nisan - 19 Mayıs 2020 tarihleri arasında Nepal'deki COVID-19 vakalarının sayısına ilişkin önceki veriler DSÖ COVID-19 durum raporlarından toplanmıştır. Nepal'de son bir ay içinde teyit edilmiş COVID-19 vakalarının sayısının zaman serisi verileri için en uygun modeli belirlemek için SPSS zaman serisi analizinde Expert Modeler özelliğini kullanmıştır. Modelleyici, ARIMA (0,1,0)'in eldeki zaman serisi verileri için en uygun model olmasını önermiştir. Araştırmaya göre, Nepal'de 2 Haziran 2020'ye kadar vaka sayısının nokta tahmininin 1266 (%95 GA: 529-2602) olacağını öngörmüştür.

Zhao (2020), araştırmasında yeni koronavirüsün ilk ortaya çıktığı Hubei Eyaleti ve salgının en şiddetli olduğu Amerika Birleşik Devletleri için bir zaman serisi modeli oluşturmak, yeni koronavirüsün yayılma eğilimini ve kısa vadeli tahminini analiz etmeyi amaçlamıştır. Kümülatif tanı sayısı için Holt modeli, kümülatif tedavi ve ölümler için ARIMA (1,2,0) modeli kullanılmıştır. Tahmin edilen değerler, aynı dönemin gerçek değerleri ile karşılaştırılmıştır. Hubei Eyaletindeki salgın durumunun temelde Mayıs ayından sonra sona erdiğini, ancak Amerika Birleşik Devletleri'ndeki salgın durumunun Mayıs'tan sonra daha şiddetli hale geldiği görülmüştür, bu nedenle Holt modeli ve ARIMA modeli, kısa vadede salgın durumunu tahmin etmede çok uygun olduğu sonucuna varılmıştır.

Gola vd., (2020), ulusal düzeyde karantinalar sırasında Hindistan'da günlük olarak teyit edilen COVID-19 vakalarıyla ilgili verileri dikkate alan bu çalışmada, regresyon modelleri, ARIMA modeli, çok katmanlı algılayıcı, vektör otoregresyonu, duyarlı maruz kalan enfekte olmuş iyileşen (SEIR), duyarlı enfekte iyileşen (SIR), tekrarlayan sinir ağları (RNN'ler), uzun kısa süreli bellek ağları (LSTM) ve toplam COVID-19 teyitli vakaların tahmininde üstel büyüme modeli gibi farklı modeller arasındaki doğruluğu karşılaştırılmıştır. Karşılaştırmalar ülke çapında karantinalar sırasında Hindistan'daki

gerçek COVID-19 oranlarıyla bu modellerin tahmin edilen oranları karşılaştırılarak yapılmıştır. Araştırma sonucunda; bu modellerden çok azı COVID-19 insidansını ve ölüm oranlarını altı hafta içinde doğru bir şekilde öngörmüştür, ancak bazıları yakın sonuçlar vermiştir. Gelişmiş uyarı, yaklaşmakta olan veya devam eden bir salgının hafifletilmesine ve hazırlanmasına yardımcı olabilirken, tahmin için uygun olmayan modellerin kullanılması önemli olumsuz sonuçlara yol açabileceği belirtilmiştir.

Keskin, Demir ve Demir (2020) araştırmalarında Türkiye’de yeni koronavirüsün yayılım seyrinin kısıtlama öncesi ve sonrasındaki günlerin istatistik modellemesinin yapılması amaçlanmıştır. Araştırmada 17 Mart 2020 ile 19 Mayıs 2020 arasındaki 64 günlük veriler kullanılarak, kısıtlamanın olmadığı ve olduğu günlerde yeni koronavirüs hastalığına yakalanmış birey sayısını belirlemeye yönelik modelleme yapılmıştır. Kısıtlamalı günler için Üstel (Exponential) fonksiyon, kısıtlama olmayan günler için Karesel (Quadratic) fonksiyon kullanılmıştır. Araştırmada belirleme katsayıları; Karesel fonksiyon için %78, Üstel fonksiyon için %98 olarak bulunmuştur. Her iki model de istatistik olarak anlamlı bulunmuştur.

Abotalep (2020), araştırmasında, çok sayıda teyit edilmiş vakanın olduğu ülkelerde günlük durumların doğrulanması için en iyi tahmin modellerini bulmayı amaçlamıştır. Çalışma, enfeksiyon oranının en yüksek olduğu ülkeler olan Çin, İtalya ve Amerika Birleşik Devletleri üzerinde gerçekleştirilen çalışmada, ARIMA ve üstel büyüme modeli kullanılarak istatistiksel tahmin modelleri aracılığıyla tahminler yapılmıştır. Sonuçlar, üstel büyüme modelinin COVID-19 vakalarını tahmin etmek için ARIMA modellerinden daha iyi olduğunu göstermiştir.

Ahmadi, vd., (2020) yaptıkları çalışmada; İran'da COVID-19'un salgın eğilimini ve tahminini belirlemeyi amaçlamışlardır. İran Sağlık ve Tıp Eğitimi Bakanlığı tarafından yayınlanan kesin COVID-19 hastalarının günlük raporlarının kullanıldığı analizde; Gompertz, von Bertalanffy'nin salgın projeksiyon modelleri ve en küçük kare hatası (LSE) ile yüzde hatası, 3 Nisan 2020'den 13 Mayıs 2020'ye kadar hastaneye yatış vakalarının sayısını tahmin etmek için kullanılmıştır. Araştırma sonucunda; Von Bertalanffy, Gompertz ve LSE tarafından 3 Nisan 2020'deki hasta sayısının %95 güven aralığı (CI) ile tahmini 44 200 (39 208-53 809), 47 500 (38 907) olarak tahmin edilmiştir. Von Bertalanffy modeline göre ölen COVID-19 hastalarının sayısı da 3100 (2633-3717), Gompertz modeline göre 3700 (2900-4310), LSE'ye göre ise 3850 olarak tahmin edilmektedir (3200-4580). Gompertz modeline göre düz salgın eğrisinin ve hasta

sayısının 67 000 (61 500-87 000) vakayı yansıtacağı tahmin edilmektedir. Gompertz ve von modellerine göre 13 Mayıs - 1 Haziran 2020 tarihleri arasında sırasıyla 7900 (6200-9300) ve 4620 (3930-5550) ölüm meydana gelecek ve ardından eğrinin düzleşeceği saptanmıştır.

Yusuf, vd., (2020) araştırmalarında; Pakistan'daki COVID-19 teyit edilmiş vakaların yanı sıra ölüm ve iyileşme sayısı tahmini üretmişlerdir. Ortalama Model (ARIMA) uygun tahmin modelleri, Pakistan'daki teyit edilmiş vaka, ölüm ve iyileşme sayısında yüksek üssel büyüme olduğunu ortaya koymuştur. Model tahminine dayanarak, teyit edilen vaka sayısının 2,7 kat artacağı, bu durumun Pakistan açısından endişe verici olduğu belirtilmiştir.

Amar, vd. (2020), Mısır'daki enfekte hasta sayısını tahmin etmek için yedi regresyon modeli (yani, üstel, logit büyüme, ikinci dereceden, üçüncü derece, dördüncü derece, beşinci derece ve altıncı derece) uygulamıştır. Modellerin geliştirilmesi için 15 Şubat - 15 Haziran 2020 tarihleri arasındaki resmi veri tabanı kullanılmıştır. Bu modeller, gelecek 15 gün, bir ay boyunca COVID-19 salgını ve son salgın boyutu ve zamanını öngörmüştür. Sonuçlar üstel, dördüncü derece, beşinci derece ve altıncı derece modellerin sonraki 15 gün için en iyi tahmin sonuçlarını verdiğini doğrulamıştır. Bu arada, dördüncü derece model, bir ay boyunca tahmin yeteneğini kanıtlamıştır. Logit büyüme regresyon modeli kullanılarak salgının zirve ve bitiş zamanının sırasıyla 22 Haziran 2020 ve 8 Eylül 2020 olduğu tahmin edilmiştir. Ayrıca toplam salgın boyutunun 166 olacağı 760 vaka tahmin edilmiştir. Ancak, rapor edilen sonuçlar sosyal ve doğal (iklimsel) koşullardan kaynaklanan belirsizliğe maruz kalmıştır.

Saba ve Elsheikh (2020), Mısır'da COVID-19'un yayılmasını tahmin etmek için doğrusal olmayan otoregresif yapay sinir ağlarını (NARANN) ve otoregresif entegre hareketli ortalamayı (ARIMA) kullanmışlardır. Model, 1 Mart ile 1 Mayıs 2020 arasında rapor edilen veri seti kullanılarak geliştirilmiştir. Modeli değerlendirmek için belirleme katsayısı, ortalama mutlak hata, ortalama karekök hata, sapma oranı ve artık kalite katsayısını kullanılmıştır. İstatistiksel göstergeler, NARANN modelinin ARIMA modeline kıyasla daha iyi tahmin yeteneklerine sahip olduğunu doğrulamaktadır.

Moftakhar, Seif ve Güvenli (2020) araştırmada; İran'da gelecekteki otuz gün boyunca COVID-19 ile günlük yeni enfekte vakaların sayısını tahmin etmek için Yapay Sinir Ağları (YSA) ve Otomatik Regresif Entegre Hareketli Ortalama (ARIMA) modelleri

tahmin için uygulanmıştır. Veriler, İran Sağlık Bakanlığı'nın günlük raporlarından ve JOHN Hopkins tarafından sağlanan açık veri setlerinden hazırlanmıştır. Araştırma sonucunda; Her iki algoritma da yeni enfekte olmuş hasta sayısında üstel bir artış öngörmüştür. ARIMA tahmininin YSA'dan daha doğru olduğu doğrulanmıştır

Gebretensae ve Asmelash (2021), çalışmalarında, Etiyopya'da COVID-19'un eğilimini analiz etmek ve yayılmayı tahmin etmek ve uygun bir istatistiksel tahmin modeli geliştirmeyi amaçlamışlardır. ARIMA modelinin geliştirilmesi için 13 Mart 2020 ile 31 Ağustos 2020 arasındaki günlük yayılmaya ilişkin veriler toplanmıştır. Durağanlık testi, parametre testi ve model teşhisi yapılmıştır. Ayrıca, otokorelasyon fonksiyonu (ACF) ve kısmi otokorelasyon fonksiyonları (PACF) kullanılarak aday modeller elde edilmiştir. Son olarak, ARIMA modellerinin uydurma, seçim ve tahmin doğruluğu, RMSE ve MAPE model seçim kriterleri kullanılarak değerlendirilmiştir. Etiyopya'da 13 Mart - 31 Ağustos 2020 tarihleri arasında toplam 51.910 onaylanmış COVID-19 vakası rapor edilmiştir. 31 Ağustos 2020 itibariyle toplam iyileşen ve ölüm oranları sırasıyla %37,2 ve %1,57 olup, Ağustos ortasından sonra yüksek bir artış düzeyine ulaşmıştır. Bu çalışmada, ARIMA (0, 1, 5) ve ARIMA (2, 1, 3), en düşük RMSE, MAPE ve BIC'ye dayalı olarak sırasıyla doğrulanmış ve iyileşmiş COVID-19 vakaları için en uygun model olduğu doğrulanmıştır.

Elsheikh, vd. (2021) çalışmalarında; Suudi Arabistan'daki toplam teyit edilmiş vaka sayısını, toplam iyileşen vakaları ve toplam ölümleri tahmin etmek için sağlam bir derin öğrenme modeli olarak uzun kısa süreli bellek (LSTM) ağı önermişlerdir. Model, resmi olarak bildirilen veriler kullanılarak geliştirilmiştir. Modelin tahmin doğruluğunu maksimize eden parametrelerinin optimal değerleri belirlenmiştir. Modelin tahmin doğruluğu yedi istatistiksel değerlendirme kriteri, yani ortalama karekök hatası (RMSE), belirleme katsayısı (R), ortalama mutlak hata (MAE), verimlilik katsayısı (EC), genel indeks (OI), varyasyon katsayısı (COV) ve artık kütle katsayısı (CRM) kullanılarak değerlendirilmiştir. Önerilen modelin tahmin doğruluğu diğer iki modelle karşılaştırılmıştır. Birincisi, otoregresif entegre hareketli ortalama (ARIMA) adı verilen istatistiksel tabanlı bir modeldir. İkincisi, doğrusal olmayan otoregresif yapay sinir ağı (NARANN) adı verilen yapay zeka tabanlı bir modeldir. Son olarak, önerilen LSTM modeli, altı farklı ülkede teyit edilen vakaların yanı sıra ölümlerin toplam sayısını tahmin etmek için uygulanmıştır; Brezilya, Hindistan, Suudi Arabistan, Güney Afrika, İspanya

ve ABD. Bu ülkeler, farklı politikalar uyguladıkları ve farklı yaş yapısı, hava durumu ve kültüre sahip oldukları için farklı salgın eğilimlerine sahiptirler.

Gnanvi, vd., (2021), COVID-19 için kullanılan modelleme teknikleri eğilimleri özetlemek için küresel bir sistematik literatür taraması yapılmıştır. 1 Ocak 2020 ile 30 Kasım 2020 tarihleri arasında kümülatif vakalar ve ölümler için öngörülen ve gözlemlenen değerlerin yanı sıra bu tahminlerin belirsizliklerini karşılaştırarak tahminlerin doğruluğunu ve kesinliği incelenmiştir. Tanımlanmış anahtar kelimelerle bulunan 4311 hakemli makale ve ön baskıdan 242'si tamamen analiz edilmiştir. Çalışmaların çoğu Asya (%78.93) ve Avrupa (%59.09) ülkeleri üzerinde yapılmıştır. Çoğu bölümsel modeller (yani SIR ve SEIR) (%46,1) ve istatistiksel modeller (büyüme modelleri ve zaman serileri) (%31,8) kullanırken, az bir kısmı yapay zeka (%6,7), Bayes yaklaşımı (%4,7), Ağ modelleri (%2.3) ve Etmen tabanlı modeller (%1.3) kullanılmıştır. Kümülatif vaka sayısı için, tahmin edilenin gözlemlenen değerlere oranı ve tahminlerin güven aralığı (CI) veya güvenilirlik aralığının (CrI) genişliğinin oranı ve merkezi değer, yanlış ve kesin olmayan tahminler ve büyük varyasyon durumlarını gösteren ortalama tahminleri 1'den büyük olduğu saptanmıştır. Bu iki oran için kullanılan modeller arasında önemli bir fark yoktur. %75'lik bir CI veya CrI tahmini sağlamada gözlemlenen değer, öngörülen kümülatif vakanın %95 CI veya CrI'si dahilindedir. Çalışmaların sadece %3,7'si kümülatif ölüm sayısını tahmin etmiştir. Tahminlerin %70'i için, tahmin edilenlerin gözlenen kümülatif ölümlere oranı 1'e yakın veya daha az olduğu belirlenmiştir. Ayrıca, Bayes modeli, klasik istatistiksel modellere göre gerçeğe daha yakın tahminler ortaya koymuştur, ancak bu farklılıklar, veri kümesindeki az sayıda tahmin nedeniyle (toplamda 9) yalnızca fikir vericidir. Ek olarak, anlamlı bir negatif korelasyon bulunmuştur ($\rho = -0,56, p = 0.021$) bu oran ile modellemenin kapsadığı sürenin uzunluğu (gün olarak) arasındadır; bu, modelin kapsadığı süre ne kadar uzun olursa tahminlerin o kadar doğru olma eğiliminde olduğunu gösterir. Bulgular, farklı modeller tarafından yapılan tahminlerin pandemi seyrini anlamak ve politika oluşturmaya rehberlik etmek için yararlı olsa da, bazılarının nispeten doğru ve kesin olduğunu, bazılarının ise olmadığını göstermiştir.

Geçili, Ziady ve Szczesniak (2021) makalelerinde; R'deki 'tahmin' paketinden kolayca erişilebilen dört farklı istatistiksel/zaman serisi modelini ele almışlardır. Bu modellerle enfekte vakaların sayısını (doğrulanmış vakalar ve benzer şekilde ölüm ve iyileşme sayısını tahmin ettik)) noktasal tahminler etrafındaki belirsizliği tahmin etmek

için karşılık gelen %90 tahmin aralığı ile yeni uygulamalar gerçekleştirmişlerdir. Bu dört model, hem ABD hem de İtalya için COVID-19 salgınının kamuya açık verilerine uygulanmıştır. Tüm modellerin, COVID-19'un gelecekteki teyit edilmiş vaka, ölüm ve iyileşme sayılarını makul bir şekilde tahmin ettiği gözlenmiştir. Bununla birlikte, analizlerin çoğu için, otoregresif entegre hareketli ortalamalı (ARIMA) zaman serisi modeli ve kübik düzleştirme spline modellerinin her ikisi de, Box ile Holt ve Trigonometrik Üstel düzleştirme durum uzayı modeline kıyasla daha küçük tahmin hatalarına ve daha dar tahmin aralıklarına sahip olduğu belirlenmiştir.

Sandoval, vd. (2021) araştırmalarında; Şili'deki doğrulanmış COVID-19 vakalarının ve ölümlerin sayısını tahmin etmek için farklı zaman serisi metodolojilerini karşılaştırmışlardır. Bu araştırmada kullanılan metodoloji, zamana bağlı sayım verileri için ARIMA modelleri, Üstel Düzeltme teknikleri ve Poisson modelleri kullanılarak Şili'de COVID-19 kaynaklı hem doğrulanmış tanıların hem de ölümlerin modellenmesinden oluşmuştur. Ek olarak, bir eğitim seti ve bir test seti kullanarak tahminlerin doğruluğunu değerlendirilmiştir. Araştırma sonucunda; kullanılan veri seti, doğrulanmış Şili'deki COVID-19 vakalarının sayısını tahmin etmek için en uygun modelin ARIMA zaman serisi modeli olduğunu göstermiştir.

Kumar ve Kumar (2021) araştırmalarında; bulanık C-arac kümeleme tekniğini kullanarak Hindistan'da yaklaşmakta olan COVID-19 ile enfekte vakaların ve ölümlerin tahmini için hibrit bir bulanık zaman serisi modeli önermişlerdir. Önerilen modelin iki aşaması vardır. Aşama-I'de, modifiye bulanık C-ortalamar kümeleme tekniği, kümelerin centroid yardımıyla temel aralıkları oluşturmak için kullanılırken, Aşama-II'de bu aralıklar, alt aralıklar oluşturacak şekilde yükseltilmektedir. Önerilen model, ortalama kare hata, ortalama karekök hata ve ortalama tahmin hata oranına dayalı performansının ölçümü için mevcut COVID-19 verilerine karşı test edilmiştir. Önerilen modelin yeniliği, gelecek 31 gün için COVID-19 ile enfekte vakaların ve ölümlerin tahmininde yatmaktadır. Bunun yanı sıra ihtiyaç duyulan yaklaşık izolasyon yatağı ve yoğun bakım ünitesi sayısı tahmini yapılmıştır.

Singh , Chowdhury ,Kumar Panja & Neogy (2021), araştırmalarında Hindistan için 12 Mart 2020'den 2 Temmuz 2020'ye kadar teyit edilmiş vakaların, iyileşen vakaların ve ölüm vakalarının verilerini kullanarak, tahmine dayalı model oluşturmayı amaçlamışlardır. Çalışmada pozitif vaka sayısı, yapılan test sayısı ve ortalama pozitiflik oranı dikkate alınarak ARIMA tabanlı bir tahmin modeli geliştirilmiştir. Analiz,

karantinadan sonra artan günlük vaka sayısındaki artışın anormal bir eğilim olmadığını göstermiştir. Tecrit verilerine dayanan tahmini grafik, COVID vakalarının sayısında önemli bir artış göstermiştir. Geliştirilen model ve tahmin edilen sonuçlar, gerçek vaka sayısı ile çok yakından uyumludur.

Tetali vd., (2021) büyük ölçekli ilaç dışı müdahalelerin COVID-19 pandemisi üzerindeki etkisini incelemeyi amaçlamışlardır. Tecrit öncesi, karantina ve açılış döneminde COVID-19 pozitifliğini karşılaştırarak Hindistan'ın dört eyaletinde müdahalelerin etkisini değerlendirmek için yarı deneysel bir model olan kesintili zaman serisi analizi (ITS) kullanılmıştır. Araştırma sonucunda; dört eyalette de pozitiflik karantina sırasında düşmüş ve eyaletler açıldıkça karantina önlemleri gevşetildikten hemen sonra eğilimler tersine dönmüştür. Pozitifliğin azalma oranı, eyaletler arasında önemli ölçüde farklılık göstermiştir. Kapanma ve açılış dönemi arasında, eyaletler arasında önemli farklılıklar olan tüm eyaletlerde pozitiflikte bir artış kaydedilmiştir. Analiz, karantina önlemlerinin COVID-19'un yükünü azaltmada olumlu bir etkisi olduğuna dair kesin kanıtlar sunmuştur.

Chu (2021), çalışmasında Avrupa'da en çok etkilenen iki ülke olan İtalya ve İspanya'daki yeni Coronavirüs (COVID-19) salgınının basit bir istatistiksel analizini yapmıştır. Her ilgili ülkede ilk vakaların doğrulanmasından sonraki yaklaşık ilk ay boyunca günlük ve kümülatif insidans verilerini kullanarak eğilimleri analiz etmiş ve insidans modelleyerek, epidemiyolojide iki yaygın yaklaşımı kullanarak temel üreme değerini tahmin etmiştir: SIR modeli ve log-lineer model. SIR modelinden elde edilen sonuçlar, salgının ilk aşamalarında İspanya ve en çok etkilenen bölgelerinin kümülatif insidansına yeterli bir uyum göstermiştir, ancak İtalya ve en çok etkilenen bölgelerinde tahminin önemli ölçüde altında olduğunu göstermiştir. Modelden salgının erken evresindeki temel üreme sayısının tahminlerinin her durumda birden fazla olduğu ve bu da COVID-19'un artan bulaşıcılığına işaret etmiştir. Günlük insidans oranına log-lineer regresyon modeli uygulandığında, İtalya ve İspanya'daki salgının büyüme aşamasının sonuçları, İtalya'ya (ve en çok etkilenen bölgelerine) kıyasla İspanya'nın daha yüksek bir büyüme oranına sahip olduğunu göstermiştir.

Ma vd., (2021), Bu çalışma, geleneksel LSTM modellerinin verilerin sapmalarını öngörmesi açısından Markov modeli (LSTM-Markov) ile birleştirilmiş bir LSTM modeli önerilmektedir. İlk olarak, model dört ülkeden doğrulanmış vaka verileriyle denenmiştir: ABD, İngiltere, Brezilya ve Rusya. Ardından, LSTM modelini düzelten Markov modelini

kullanarak her ÷lkede 20 Şubat 2021'den önceki teyit edilen vaka sayısı tahmin edilmiştir. Şubat 2021'e kadar İngiltere'de vaka sayısının istikrara kavuşacağı ve salgının kontrol altına alınacağı, ABD, Brezilya ve Rusya'da ise vaka sayısının artmaya devam edeceği öngör÷lmüştür. Sonuçlar, önerilen LSTM-Markov modelinin tahmin eğrisinin gerçek salgın eğrisine daha yakın olduğunu, ortalama RMSE'nin LSTM modelinin sadece %40'ı olduğunu göstermiştir. Bu nedenle, LSTM-Markov'un tahmin doğruluğı, LSTM modelinden çok daha yüksek olduğu saptanmıştır.



2. ZAMAN SERİLERİ

2.1. Zaman Serilerinin Tanılanması

Zaman serisi bir olgunun değişkenlerinin belirli periyotlarla gözlemlenen ardışık değerlerinden oluşur. Zaman serisi verileri kendi içlerinde bağımlıdır. Seri içerisinde birbirine yakın olan değerler genellikle benzerdir. Birbirine yakın olan veriler arasında korelasyon güçlü iken uzak olan veriler arasındaki korelasyon azdır (Erdoğan, 2006).

Zaman serisinde değerler arasındaki süre eşit olmalıdır. İki ardışık değer arasında saniyelik, dakikalık, saatlik fark olabileceği gibi günlük, haftalık, aylık, mevsimlik ve yıllık farklar da olabilir. Ölçülen değerler süreklilik gösteren bir olguya ait olsa bile, gözlemler belirli zaman aralıkları ile kaydedildiği için zaman serisinin verileri kesikli değerlerden meydana gelir.

Bağımlı değişkenin zaman olduğu durumlarda geleceğe yönelik tahmin için zaman serisi analizi kullanılabilir. Bu nedenle zaman serilerinin uygulama alanları oldukça geniştir. Ekonomi alanında faiz ve işsizlik oranları, satış ve hisse senedi fiyatları, sosyolojide suç ve boşanma oranları, meteorolojide yağış miktarı, sıcaklık, rüzgar hızı, ekolojide çevre ve su kirliliği tahmininde zaman serisine dayalı tahmin teknikleri kullanılmaktadır (Erdoğan, 2006).

Zaman serisi verileri, doğaları gereği karmaşık bir yapıya sahiptir ve bunların ayıklanması ve yorumlanması oldukça zordur. Bir zaman serisinin zaman içerisindeki davranışını anlamamanın en kolay yolu, bu serinin "zaman serisi grafiği" adı verilen grafiğini çizmektir (Goggin, 2007). Zaman serisi analizi, anlamlı istatistiksel sonuçlar elde etmek ve verilerin diğer özelliklerini ifade etmek için zaman serisi verilerini analiz etme yöntemlerini içerir (Lin vd, 1995).

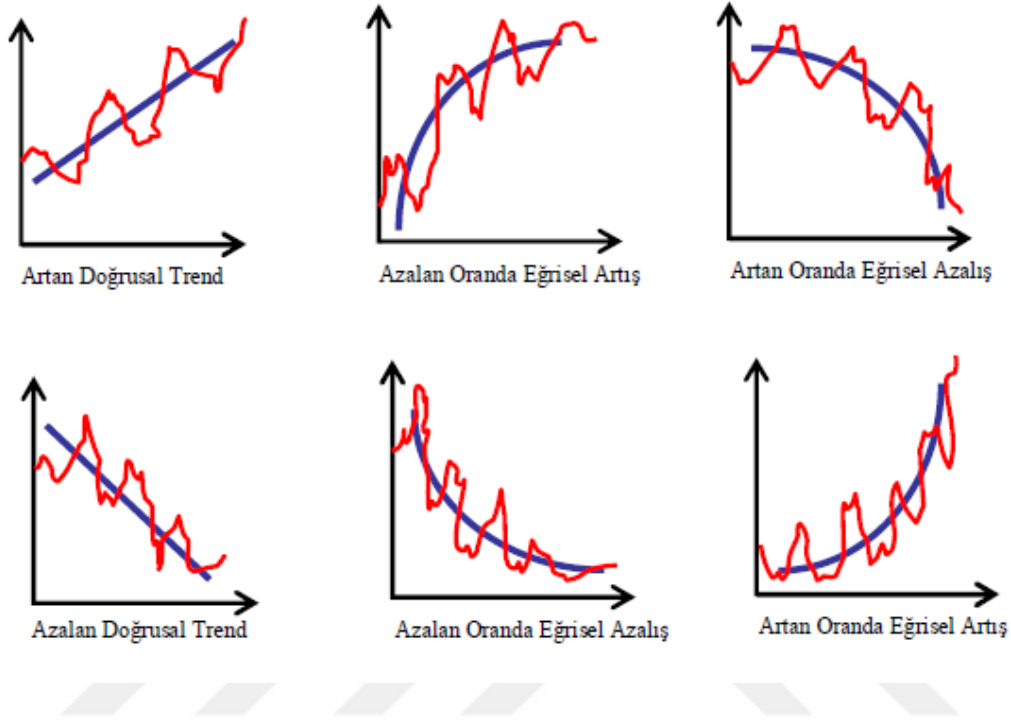
2.2. Zaman Serisinin Bileşenleri

Bir zaman serisindeki gözlemin değerlerini etkileyen çeşitli nedenlere zaman serisinin bileşenleri denir. Bu bileşenler genel olarak Eğilim (Trend), Mevsimsel, Döngüsel ve Düzensiz değişimler olmak üzere dört ana bileşene ayrılır.

2.2.1. Trend (T)

Bir zaman serisinin uzun bir süre belirli bir yönde gösterdiği ana eğilime trend adı verilir (Kırçıl, 2013). Trend, serinin temelini teşkil eden uzun vadeli davranış veya örüntüdür. Eğilim, bir zaman serisindeki uzun dönemli yükseliş veya düşüşü gösterebilir.

Trend; nüfus artışı, fiyat enflasyonu ve genel ekonomik değişiklikler gibi etkilerin sonucu olabilir. Eğilim iki şekilde gözlemlenebilir: Doğrusal Eğilim ve Doğrusal Olmayan Eğilim (Verbesselt vd, 2010). Bazı eğilim örnekleri şunlardır: Nüfus, tarımsal üretim, endüstriyel üretim, doğum ve ölüm sayıları.



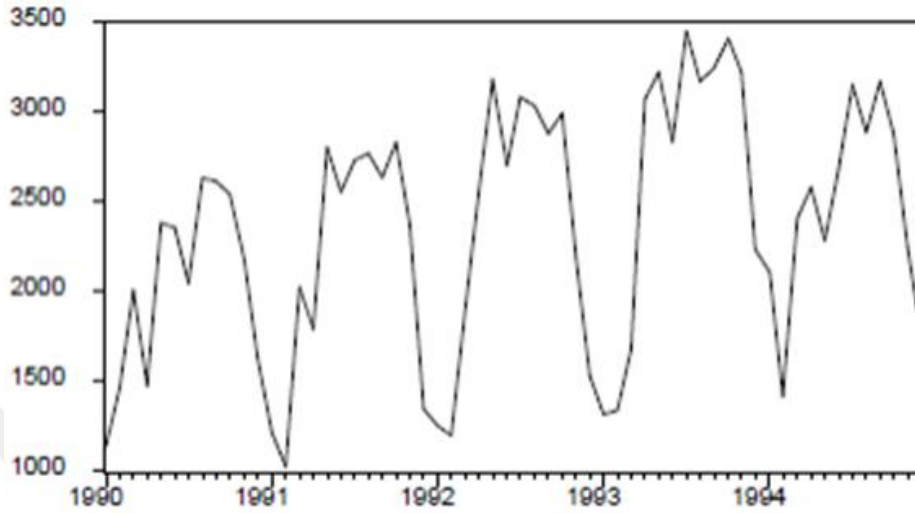
Şekil 2.1. Doğrusal ve doğrusal olmayan trend sergileyen zaman serisi örneği (Özek, 2010).

2.2.2. Mevsimsel Dalgalanmalar (M)

Zaman serisi içerisindeki düzenli dalgalanmalara mevsimsellik adı verilir. Bu dalgalanmalar yalnızca aylık veriler olarak düşünülmemelidir. Örneğin saatlik verilerin bulunduğu zaman serilerinde trafik yoğunluğu ve hava kirliliği verilerinin göstermiş olduğu dalgalanmalarda genelde mevsimsellik tanımıyla örtüşmektedir. Mevsimsel dalgalanmalar periyodu belli olan döngülerdir. Dalgalanmaların başlangıç ve bitiş noktaları minimum ve maksimum olduğu noktalar genel olarak her periyotta aynı zaman dilimine tekabül eder.

Mevsimsel olarak etkilenen değişkenler yılın bazı dönemlerinde diğer dönemlere göre daha yüksek veya daha düşük değerlere ulaşmaktadır. Bu davranış örüntüsü mevsimsellik olarak bilinir ve birçok farklı şekilde ortaya çıkabilir. Örneğin belirli mevsimlerde, belirli aylarda, bir yılın belirli haftalarında, bir çeyreğin belirli bir ayında, belirli bir haftada veya belirli bir günde meydana gelebilir (Miyake, vd., 2009).

Mevsimsel deęişimlere ilişkin bazı örnekler şunlardır: Kış ve yaz aylarında soęuk iecek talebinin artması veya azalması, yazın su tüketiminin artması, Ramazan ayında alışverişin artması ve benzeri. Mevsimsellik; altı ay, üç ay, bir ay, bir hafta, bir gün ve hatta bir saat gibi dilimleri kapsayabilir.



Şekil 2.2. 1990:01 – 1994: 12 yıllarına ait imento üretim miktarı

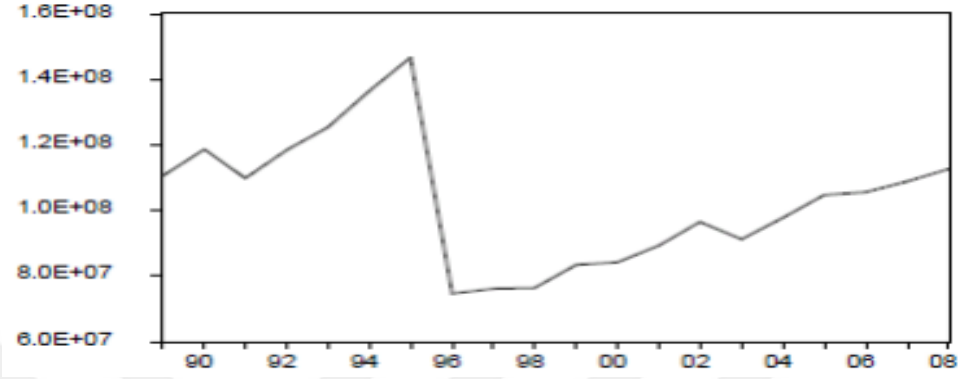
Mevsimsellik serilerin varyansını önemli derecede etkilemektedir. Mevsimsellik dikkate alınmadan analizler ve tahmin alışmaları gerçekleştirilecek olursa sonuçlar hatalı olacaktır (Tıraşoęlu, 2012). Bu tip hatalarla karşılaşılmaması iin Üstel Düzeltme, X-11 metodu, Census II yöntemi gibi teknikler kullanılarak serinin mevsimsellikten arındırılmasıyla analizler yapılabilir (Kıril, 2013).

2.2.3. Konjonktürel Deęişimler (K)

Bir trend doğrusu veya eğrisi etrafında oluşan uzun dönem dalgalanmalara konjonktürel dalgalanmalar denilmektedir. Bu hareketler mevsimsel dalgalanmalara benzetilseler de dalgaların uzunluğu ve sürelerinin periyodik olmamaları yönüyle mevsimsel dalgalardan ayrılmaktadır. Konjonktürel hareketlerin, genellikle ekonomik şartlara baęlı olduğu ve savaş, afet, kriz gibi olaęandışı dönemleri yansıttıkları ifade edilebilir (Kıril, 2013).

Konjonktürel Deęişimler, verilerde bir yıldan uzun süredir var olan kısa vadeli dalgalanmalardır (yükselme ve düşmeler). Özellikle, veri sabit bir sıklığa sahip olmayan yükselme ve düşme gösterdiğinde döngü oluşur. Genellikle, enflasyon veya hisse senedi fiyatındaki keskin yükseliş gibi iş döngüsü ile ilişkilendirilen beklenmeyen veya öngörülemeyen olaylardan kaynaklanır.

Mevsimsel ve konjonktürel değişim arasındaki temel fark, evvelkinin bir yıl içerisinde sabit bir uzunlukta olması, sonrakinin ise bir yıldan daha uzun bir süre çeşitli uzunluklara sahip olmasıdır ve ayrıca dalgalanmalar esnek ve değişken döngüseldir, ancak sıklık (frekans) değişmiyorsa ve birtakım takvimsel özellikler gösteriyorsa, örüntü mevsimseldir (Chen vd, 2016).

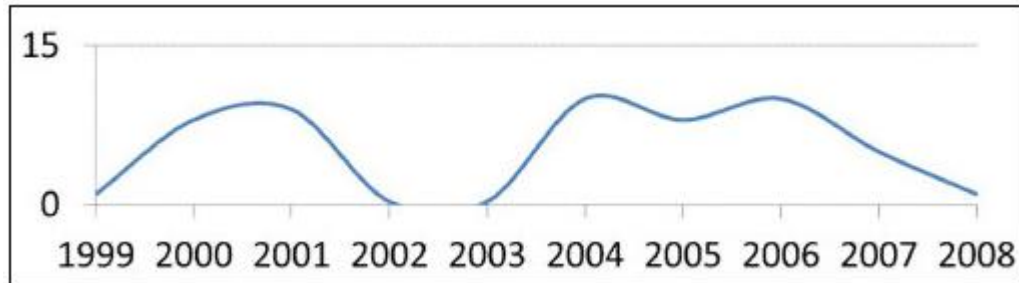


Şekil 2.3. 1989 – 2008 yıllarına ait GSYİH verisi

2.2.4. Düzensiz Hareketler (D)

Zaman serilerinde belirli bir sistematığa uyum sağlamadan gerçekleşen rastgele değişimlere düzensiz hareketler veya rassallık adı verilir. Rassallığın oluşumundaki faktörlerin kestirimi mümkün değildir (Kırçıl, 2013).

Düzensiz hareketler, zaman serilerindeki düzensiz değişimlerdir ve gözlemlenemezler. Düzenli değişimler değildir ve bütünüyle rastgele veya düzensiz bir doğaya sahiplerdir. Zaman serisinin diğer tüm bileşenler çıkarıldıktan sonra geriye kalan açıklanamayan kısımdır. Bu bileşenin tanımlanabilir bir şekli yoktur (Yasue, vd., 1993).



Şekil 2.4. Düzensiz hareketler sergileyen zaman serisi örneği

2.3.Durağanlık

Bir zaman serisinin ortalama, varyans, kovaryans gibi özellikleri zamanla değişmeyip sabit kalıyorsa seri durağandır denir. Bu tip seriler dalgalanma içermez (Duru, 2007). Fakat böylesi zaman serilerine rastlamak pek mümkün değildir. Çünkü zaman serilerinin nerdeyse tamamı bazı dalgalanmalar veya eğilimler içerir. Bu durum serinin durağanlığını kaybetmesine sebep olur.

Durağan olmayan zaman serilerinde ise değerler belli bir ortalama etrafında dağılmaz ve varyans zamanla değişir. Ayrıca serinin değerleri arasında uzaklık arttıkça benzerliğin azalacağından bahsedilemez (Sevüktekin ve Nargeleçekenler 2005). Zaman serilerinin çoğunluğu durağan olmayan zaman serileridir. Bu nedenle tahmin yaparken başarılı sonuçlar ortaya çıkarmak oldukça zordur. Tahminlerden başarılı sonuçlar ortaya çıkarabilmek için durağan olmayan zaman serilerine farklı dönüşümler uygulayarak seriler durağan hale getirilebilir.

2.3.1. Otokorelasyon Fonksiyonu (ACF)

Zaman serisi verilerinde, otokorelasyon, aynı veri setinin zaman içinde farklı noktalarda gözlemleri arasındaki korelasyondur. Farklı zaman serisi modellerine duyulan ihtiyaç, kısmen zaman serisi verilerinde bulunan otokorelasyondan kaynaklanmaktadır (Eric, 2019).

Otokorelasyon katsayısı, zaman serilerinin farklı dönemlerdeki gözlemleri arasındaki ilişkiyi gösterir. Zaman serilerinin en önemli göstergesi katsayılarıdır. Katsayılar; zaman serisi gözlemleri arasındaki bağımlılığın derecesini ortaya çıkarır. Yüksek otokorelasyon katsayısı zaman serisinin geçmiş değerlere bağlı olduğunu gösterir, düşük olması ise zaman serisinin rassal olduğunu göstermektedir. Teorik olarak, k gecikmesinde otokorelasyon katsayısı (Kıran, 2015); $P_k = \frac{\gamma_k}{\gamma_0}$ olarak hesaplanmaktadır. Denklem 2.1'deki γ_k ; k gecikmede kovaryansı ifade ederken, γ_0 ise varyansı ifade etmektedir. Açık şekilde ifade edilirse;

$$\gamma_k = \frac{E[(Y_t - \mu)(Y_{t-k} - \mu)]}{\sqrt{E(Y_t - \mu)^2 E(Y_{t-k} - \mu)^2}} = \frac{KOV(Y_t, Y_{t-k})}{\sigma_{y_t} \sigma_{y_{t-k}}} = \frac{\gamma_k}{\sigma_y^2} = \frac{\gamma_k}{\gamma_0} \quad (2.1)$$

olmaktadır. $k = 0$ durumu gecikmenin 0 olduğunu göstermektedir ve bu gecikmede otokorelasyon katsayısı; $\rho_0 = \frac{\gamma_0}{\gamma_0} = 1$ 'dir. Uygulamada sürecin tamamı değil, bir kısmı gözlemlendiği için tahmini değerler olan örnek otokorelasyon katsayıları hesaplanmaktadır. Dolayısıyla ρ_k otokorelasyon katsayısının bir tahmini olan $\hat{\rho}_k$ 'nin hesaplanabilmesi için öncelikle k gecikmede örneklem kovaryansının $\hat{\gamma}_0$ aşağıdaki gibi elde edilmesi gerekmektedir.

$$\hat{\gamma}_k = \frac{\sum (Y_t - \bar{Y})(Y_{t-k} - \bar{Y})}{n} \quad (2.2)$$

$$\hat{\gamma}_0 = \frac{\sum (Y - \bar{Y})^2}{n}$$

Buradan hareketle örnek otokorelasyon kats $\hat{\rho}_k = \frac{\hat{\gamma}_k}{\hat{\gamma}_0}$ şeklinde hesaplanmaktadır.

2.3.2. Kısmi Otokorelasyon Fonksiyonu (PACF)

Regresyon analizlerinde bağımlı değişken Y_t , X_1 ve X_2 gibi bağımsız değişkenler üzerine regres edildiğinde örnek olarak X_2 nin kısmi etkisinin dışlanması yada sabit tutulması durumunda X_1 'in kısmi etkisinin ne olacağı hususunda bir sonuç belirlenebilmektedir. Y_t 'nin X ile regresyonundan kalıntılar elde edilir ve X_1 ile ilgili kalıntılarla korelasyonu bulunur.

Zaman serisi analizlerinde de benzer bir kavram bulunmaktadır.

Kısmi otokorelasyon katsayıları; diğer gecikmelerin etkisi sabit olduğunda Y_t değerleri ile herhangi bir k gecikme ile oluşturulmuş Y_{t-k} 'lar arasındaki bağımlılığın derecesini gösteren katsayılardır. Y_t ile Y_{t-k} arasındaki bu katsayının diğer tüm gecikmelerin etkisinin sabit tutularak elde edildiğini not etmek gerekmektedir.

Kısmi otokorelasyon katsayılarını otokorelasyon katsayılarından aşağıdaki gibi elde etmek mümkündür.

$$\begin{aligned} \rho_{11} &= \rho_1 \\ \rho_{22} &= (\rho_2 - \rho_1^2) / (1 - \rho_1^2) \end{aligned} \quad (2.3)$$

ve ilave gecikmeler için

$$\rho_{kk} = \left(\rho_k - \sum_{j=1}^{k-1} \rho_{k-1,j} \rho_{k-j} \right) / \left(1 - \sum_{j=1}^{k-1} \rho_{k-1,j} \rho_j \right) \quad (2.4)$$

olmaktadır.

2.3.3. Beyaz Gürültü Süreci

Ekonometri ve zaman serilerinde önemli olan durağan süreç türleri; "beyaz gürültü" adı verilen olasılık süreci. Bu sürecin özelliği, ortalamanın sıfır olması, varyansın sabit olması ve otokorelasyonun olmamasıdır. Yani. Bu tür süreçler aynı zamanda bağımsız, özdeş ve normal dağılımlıdır. Buna "Gauss beyaz gürültü" süreci denir (Kıran, 2015);

Beyaz gürültü modeli çoğu zaman serisi modelinin yapı taşıdır ve şu şekilde tanımlanır (Guerrier vd., 2019): $W_t \stackrel{iid}{\sim} N(0, \sigma_w^2)$ Bu tanım şu anlama gelir:

$$E[W_t | \Omega_t] = 0 \quad E[W_t | \Omega_t] = 0 \text{ hepsi için } t,$$

$$\text{cov}(W_t, W_{t-h}) = 0 \quad \text{cov}(W_t, W_{t-h}) = 0 \text{ hepsi için } t, h$$

Daha spesifik olarak, $H \in \mathbb{N} + h \in \mathbb{N} + \text{gecikmeli}$ değişkenler arasındaki zaman farkıdır. Bu nedenle, bu süreçte zamansal (veya seri) bir korelasyon yoktur ve homoskedastiktir (yani sabit bir varyansa sahiptir). Daha fazla ayrıntıya girerek, beyaz gürültü iki tür sürece ayrılabilir: zayıf ve güçlü. Süreç (W_t) Bir olan zayıf beyaz gürültü varsa; $E[W_t] = 0 \quad E[W_t] = 0$ hepsi için t , $\text{var}(W_t) = \sigma^2 \quad \text{var}(W_t) = \sigma^2$ hepsi için t , $\text{cov}(W_t, W_{t-h}) = 0 \quad \text{cov}(W_t, W_{t-h}) = 0$ hepsi için t ve herkes için $h \neq 0$.

Bununla birlikte, bağımsızlık kavramı, güçlü bir beyaz gürültüyü şu şekilde tanımlamak için kullanılır: $E[W_t] = 0$ ve $\text{var}(W_t) = \sigma^2 < \infty \quad \text{var}(W_t) = \sigma^2 < \infty$, hepsi için t ,

$F(W_t) = F(W_{t-h})$ için t, h (nerede $F(W_t)$ $F(W_t)$ marjinal dağılımını gösterir W_t), W_t ve W_{t-h} herkes için bağımsız t ve herkes için $h \neq 0$.

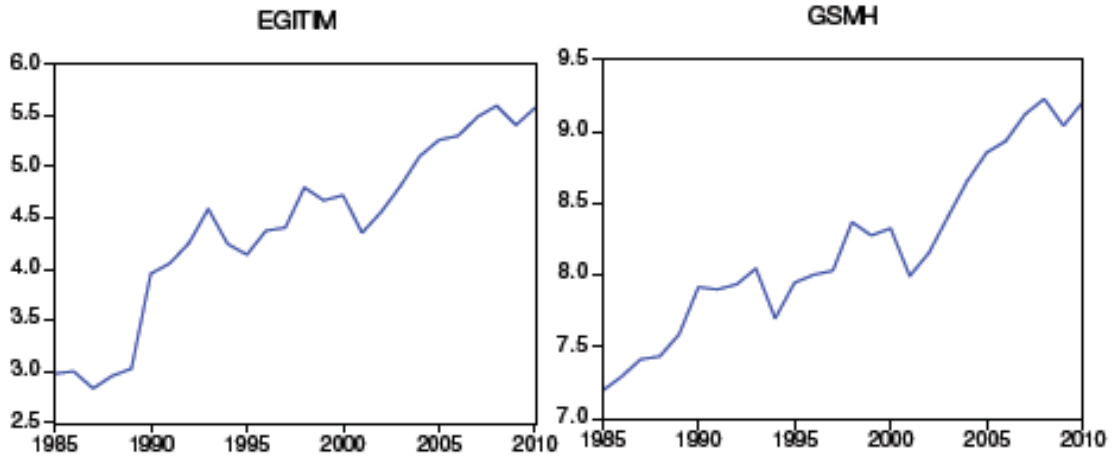
2.3.4. Durağanlığın Belirlenmesi

Bir zaman serisinin durağanlığını incelemenin temelde üç yolu vardır: Grafik analizleri, korelogramlar (serisel korelasyon grafikleri) ve birim kök testleridir.

2.3.4.1. Grafik Analizleri

Durağanlığı test etmenin en temel yöntemi zaman serisi verilerinin zaman serisi grafiğini çizmektir. Genellikle, böyle bir enformel analiz, bir zaman serisinin durağan olup olmadığına dair ilk ipuçlarını sunabilir. Zaman serisi grafiği; artan veya azalan bir eğilim, mevsimsel veya döngüsel değişim, değişen ortalama ve varyans gösteriyorsa, verilerin durağan olmadığı söylenebilir (Gujarati ve Porter, 2003).

Grafik Analizleri durağanlığı tespit etmek için kesinlikle güvenilir değildir ve genellikle verilere ilişkin kesin iddialarda bulunmaktan ziyade ilk izlenim elde etmek için kullanılmaktadır.



Şekil 2.5. Türkiye'nin eğitim harcamaları ve Gayri Safi Milli Hasıla verilerine ait zaman serisi grafikleri

2.3.4.2. Korelogramlar

Zaman serilerinin durağanlığını test etmenin bir başka yolu da korelogramlar çizmektir. Bir korelogram, zaman içerisinde değişen verilerdeki (diğer bir deyişle zaman serisi verilerindeki) serisel korelasyonu gösteren bir grafikdir. Verilerin Otokorelasyon Fonksiyonu (ACF) ve Kısmi Otokorelasyon Fonksiyonu (PACF) grafiklerini bir panel grafik şeklinde sunar. Dolayısıyla, çeşitli gecikmelerdeki otokorelasyon ve kısmi otokorelasyon katsayıları sıfır civarında ise, zaman serilerinin muhtemelen durağan olduğunu söyleyebiliriz. Ancak gecikmelerdeki korelasyonların ve kısmi korelasyonların istatistiksel olarak anlamlı olmaları serinin durağan olmadığını göstermektedir (Gujarati ve Porter, 2003).

2.3.4.3. Birim Kök Testleri

Bir zaman serisinin durağan olup olmadığını veya durağansa hangi mertebeden durağan olduğunu belirlemek için bir takım testler geliştirilmiştir. Durağanlığın sınanmasında genel olarak birim kök testleri kullanılmaktadır. Birim kök testi genel olarak $y_t = \rho y_{t-1} + \epsilon_t$ olarak ifade edilir.

Burada ϵ_t hata terimi pür rassaldır. ve y_{t-1} ' in katsayısı ρ , 1'e eşitse bu seri birim kök içeriyor denir. Katsayının bire eşitliğini sınanan test t testidir. Bu durumda durağanlık testi için hipotez testi aşağıdaki gibi geliştirilmiştir (Gujarati ve Porter, 2009).

$H_0: \rho = 1$ seri durağan değildir. Serinin birim kökü vardır.

$H_a: \rho < 1$ seri durağandır. Serinin birim kökü yoktur. Hipotezlere ilişkin test istatistiği $t = \frac{\hat{\rho}-1}{sp}$ ile hesaplanır.

Herhangi bir y_t serisinin “t” istatistiğinin hesaplanabilmesi için zaman serisinin durağan olması gerekmektedir. Durağan olmayan zaman seriler için hesaplanan istatistiği t dağılımına uymaz ve sapmalı olur. Bu durumda H_0 hipotezinin, “t” istatistiğinin t dağılımı ile test edilmesi gerçekliliğini yitirir. Yani (2.9)’da yer alan y_{t-1} varlığından dolayı etkinlik kaybı olur ve $\hat{\rho}$ aşağı yönlü sapmalı olur. Bu sebeplerden dolayı standart hata büyür ve hipotez testiyle yanlış karar verilir. Bu durumun üstesinden gelmek için çeşitli birim kök testleri geliştirilmiştir (Gujarati ve Porter, 2009).

Zaman serisi verilerinin durağanlığının tespitinde en fazla kullanılan testler: Dickey-Fuller testi, Artırılmış Dickey-Fuller (ADF) testi ve Phillips-Perron (PP) testi gibi istatistiksel testlerdir.

2.4. Zaman Serileri Analizleri

Zaman serisi analizi, belirli bir zaman diliminde toplanan bir dizi veri noktasını analiz etmenin özel bir yöntemidir. Zaman serisi analizinde, veri noktalarını aralıklı veya rastgele kaydetmek yerine, veri noktaları belirli bir süre boyunca tutarlı aralıklarla kaydedilir. Ancak bu tür bir analiz, yalnızca zaman içinde veri toplama eylemi değildir. Zaman serisi verilerini diğer verilerden ayıran şey, analizin değişkenlerin zaman içinde nasıl değiştiğini gösterebilmesidir. Diğer bir deyişle, zaman çok önemli bir değişkendir çünkü verilerin veri noktalarında nasıl ayarlanacağını ve nihai sonucu gösterir. Ek bir bilgi kaynağı ve veriler arasında belirli bir bağımlılık sırası sağlar. Zaman serisi analizi, tutarlılık ve güvenilirlik sağlamak için tipik olarak çok sayıda veri noktası gerektirir. Kapsamlı bir veri seti, temsili bir örneklem büyüklüğüne sahip olmayı ve bu analizin gürültülü verileri kesebilmesini sağlar. Ayrıca, keşfedilen eğilimlerin veya modellerin aykırı değerler olmamasını ve mevsimsel değişiklikleri hesaba katabilmesini sağlar. Ek olarak, zaman serisi verileri, geçmiş verilere dayalı olarak gelecekteki verileri tahmin etmek için tahmin için kullanılabilir (<https://www.tableau.com/>) .

Zaman serisi analizi yeni bir çalışma olmasa da teknoloji onu daha erişilebilir kılmaktadır. Onlarca yıldır, konunun temel teorisini ve pratiğini öğreten önerilen metinlerin çoğu hazırdır. Ancak yöntemin kendisi bundan daha da eskidir. Zaman serisi

analizi binlerce yıldır kullanılmaktadır ve eski gezegensel hareket ve navigasyonla izlenebilir. Zaman serisi analizi, durağan olmayan veriler için kullanılır, zaman içinde sürekli dalgalanan veya zamandan etkilenen şeylerde, örneğin; finans, perakende ve ekonomi gibi sektörler, para birimi ve satışlar her zaman değiştiği için zaman serisi analizi sıklıkla kullanılır. Borsa analizi, özellikle otomatik ticaret algoritmaları olmak üzere zaman serisi analizine iyi bir örnektir. Benzer şekilde, zaman serisi analizi de hava değişikliklerini tahmin etmek için idealdir. Meteorologların yarının hava durumu raporundan gelecek yıllardaki iklim değişikliğine kadar her şeyi tahmin etmelerine yardımcı olur. Çalışan zaman serisi analizi örnekleri şunları içerir:

- Hava durumu verileri
- Yağış ölçümü
- Sıcaklık okuma
- Kalp atış hızı izleme (EKG)
- Beyin İzleme (EEG)
- Üç aylık satışlar
- Hisse senedi fiyatı
- Otomatik hisse senedi ticareti
- Sektör tahmini
- Faiz oranı

Zaman serisi tahmini, tahmin yapmak ve stratejik kararların alınmasında bilgi sağlamak için zaman serisi verilerini analiz etmek için istatistik ve modelleme kullanma sürecidir. Her zaman kesin bir tahmin değildir ve tahminlerin olasılığı, özellikle zaman serisi verilerinde yaygın olarak dalgalanan değişkenlerle ve kontrol dışındaki faktörlerle uğraşırken farklı bir biçimde değişebilir. Bununla birlikte, hangi sonuçların ortaya çıkma olasılığının diğer potansiyel sonuçlara göre daha olası veya daha az olduğu hakkında öngörü yürütülmesine yardımcı olur. Çoğu zaman, eldeki veriler ne kadar kapsamlı olursa, tahminler o kadar doğru olabilir. Tahmin ve "öngörü" genellikle aynı şeyi ifade etse de, dikkate değer bir ayrım vardır. Bazı endüstrilerde tahmin, zaman içinde belirli bir gelecek noktasındaki verilere atıfta bulunabilirken, öngörü genel olarak gelecekteki verilere atıfta bulunur. Zaman serisi analizi, altta yatan nedenleri anlamak için verileri

anlamak için modeller geliştirmeyi içerir. Analiz, görülen sonuçların arkasındaki “neden”i açıklayabilir. Daha sonra tahmin, bu bilgiyle ne yapılacağına dair bir sonraki adımı ve gelecekte ne olabileceğine dair tahmin edilebilir öngörülerini ortaya çıkarır. (<https://www.tableau.com/learn/articles/time-series-forecasting>).

2.5. BOX-Jenkins Yöntemi

Box-Jenkins modeli, belirli bir zaman serisinin girdisine dayalı olarak veri aralığını tahmin etmek için tasarlanmış matematiksel bir modeldir. Box-Jenkins modeli, tahmin için çok sayıda farklı zaman serisi verisini analiz edebilir. Yöntemi, sonucu belirlemek için veri noktaları arasındaki farkı kullanır. Bu yöntem, modelin tahminler oluşturmak için eğilimleri belirlemek için otoregresyon, hareketli ortalamalar ve mevsimsel değişiklikleri kullanmasına izin verir.

Box-Jenkins Modeli iki matematikçi olan George Box ve Gwilym Jenkins tarafından oluşturulmuştur. Yazarlar, 1970 yılında yayınladıkları "Zaman Serisi Analizi: Tahmin ve Kontrol" adlı eserlerinde bu modeli ortaya koymuşlardır.

Box-Jenkins Modelinin parametrelerinin tahminleri çok karmaşık olabilir. Bu nedenle, diğer zaman serisi regresyon modellerine benzer şekilde, en iyi sonuçlar tipik olarak programlanabilir yazılımların kullanılmasıyla elde edilebilmektedir. Box-Jenkins Modeli ayrıca genellikle 18 ay veya daha kısa süreli kısa vadeli tahminler için uygundur.

Box-Jenkins modeli, tahmincilerin programlama tahmin yazılımı kullanırken karşılaştıkları birkaç zaman serisi analiz modelinden biri olabilir. Çoğu durumda yazılım, tahmin edilecek zaman serisi verilerine dayalı olarak en uygun tahmin yöntemini otomatik olarak kullanacak şekilde programlanacaktır. Box-Jenkins'in çoğunlukla kararlı ve düşük oynaklığa sahip veri kümeleri için en iyi seçim olduğu belirtilmektedir.

Box-Jenkins'in zaman serisi modelleme yöntemi, bir dizi yineleme yoluyla gözlemlenen verilerden öngörülebilir hareketlerin (veya kalıpların) çıkarılmasını içerir. Tek değişkenli Box-Jenkins yöntemi tamamen bir tahmin aracıdır; regresör tipi değişkenler olmadığı için hiçbir açıklama yapılmamıştır. Box-Jenkins yaklaşımı üç aşamalı bir prosedürü takip eder (Aljandali, 2017):

Model tanımlama: Box-Jenkins (BJ) modelinin belirli bir kategorisi, geçmiş verilerin analizinden hesaplanan çeşitli istatistikler kullanılarak tanımlanır.

Model tahmini ve doğrulama: Bir kez tanımlandıktan sonra, "en iyi model", uydurulan değerlerin gerçek veriler tarafından sergilenen modeli yakalamaya mümkün olduğunca yakın olacağı şekilde tahmin edilir.

Tahmin: Nihai model, zaman serisini tahmin etmek ve tahminle ilişkili belirsizliği ölçmek için kullanılan bir güven aralığı geliştirmek için kullanılır.

Box-Jenkins modeli, verileri tahmin etmek için üç ilke kullanır: otoregresif, fark ve hareketli ortalama. Bu üç ilke p , d ve q olarak adlandırılır. Her ilke Box-Jenkins analizinde kullanılır; birlikte ARIMA (p , d , q) olarak adlandırılırlar.

Otoregresif (p) süreç testi veri durağanlığı. Kullanılan veriler sabitse, tahmin süreci basitleştirilebilir. Kullanılan veriler durağan değilse, türevlenmesi gerekir (d). Veriler üzerinde bir hareketli ortalama uyum testi de gerçekleştirilmiştir (analiz sürecinin q bölümünde tamamlanmıştır). Genellikle, verilerin ilk analizi parametreleri (p , d ve q) ayarlar ve daha sonra bu parametreleri tahmin yapmak ve tahmine hazırlanmak için uygular (Scott, 2021).

2.5.1. ARIMA (Auto Regressive Integrated Moving Average)

Zaman serilerinin tahmini için en iyi bilinen ve en başarılı doğrusal istatistiksel modellerden biri ARIMA modelidir. 1970'lerin başında Box ve Jenkins tarafından önerilmiştir. Tam adı Otoregresif Entegre Hareketli Ortalama modelidir (Yüksel, 2015). ARIMA, bir değişkenin gelecekteki değerini bulmak için kendi geçmiş değerlerini kullanarak tahmin etmede kullanılan bir zaman serisi tahmin yaklaşımıdır (Alabdulrazzaq, vd.,2021).

Zaman serisi verileri, durağan ve durağan olmayan veriler olmak üzere iki kategoriye ayrılmaktadırlar. Durağan olan zaman serisi verisinin zamana göre kalıpları yoktur, oysa durağan olmayan zaman serisi verisinin mevsimsellik olarak da bilinen kalıpları vardır. Bu nedenden dolayı, durağan olmayan verilerin ortalaması ve varyansı zaman içinde sabit değildir. Durağan olmayan zaman serileri verilerini, iki ardışık gözlem arasındaki farkları hesaplanarak durağan hale getirilmektedir. Bu tekniğe fark alma işlemi denmektedir. Fark alma işlemi zaman serisi düzeyindeki değişiklikleri ortadan kaldırarak trendleri ve mevsimselliği ortadan kaldırmakta daha sonra elde ettiği verileri tahmin için kullanmaktadır (Kumar, vd., 2021).

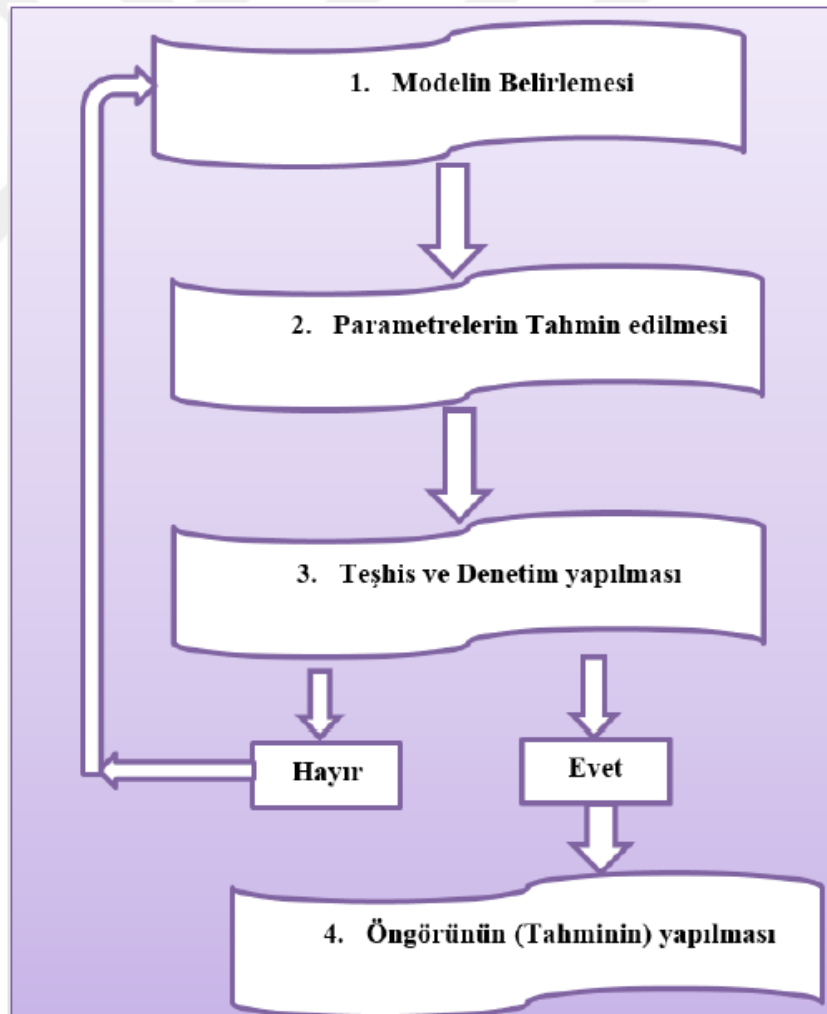
Durağanlaştırma işlemi için serilerin farkı alınır. Literatürdeki çalışmalarda serilerin en çok iki kez farkının alındığına rastlanılmıştır.

Eşitlik 2.1’de ARIMA (p,d,q) modelinin genel ifadesi verilmektedir; $(1-\phi L-\phi L^2-\dots-\phi_p L^p)\Delta^d Y_t = \delta + e_t + \theta e_{t-1} + \dots + \theta_g e_{t-g}$

ARIMA zaman serisi verilerine dayalı istatistiksel modeller geliştirmeden önce bir takım hipotezler uygulanmaktadır. Bunların başlıcaları (Hasmida 2009):

1. Verilerin zamana göre durağanlığının olması,
2. Verilerin normal dağılımının olması,
3. Veriler içinde herhangi bir anormallik veya aykırılık olmaması,
4. Verilerin herhangi bir eksikliğinin olmaması

Şekil 2.6’da ARIMA zaman serisi modelinin 4 temel aşamalı işleyişi gösterilmiştir.



Şekil 2.6. ARIMA modeli temel işleyişi (Lee and Ko 2011; Muhammad 2012; Kohvakka 2017)

2.5.1.1. Model Belirleme

ARIMA modellerinin en önemli varsayımlarından biri elde edilen gözlemlerin eşit zaman aralıklarında kesikli ve durağan verilerden oluşmasıdır. Seride durağanlık yok ise seri durağan hale getirilmektedir. Seriler durağanlaştırılarak periyodik dalgalanmaların önüne geçilmesini sağlamaktadır (Bircan ve Karagöz 2003). Bu modellerdeki temel amaç en az sayıda parametre içeren en uygun modellerin elde edilmesidir (Temur and Yıldız 2021). Uygun olan ARIMA zaman serisi modelinin belirlenmesinde durağanlık durumunda fark alma işlemi ve AR ve MA gecikmelerinin tespitini yapmak yani (p, d, q) değerlerini belirlemek gerekmektedir (Önen 2020).

Bir zaman serisinin oto-korelasyon (ACF), ters korelasyon (IACF) ve kısmi korelasyon (PACF) gözlemleri incelenerek zaman serilerinin durağanlık durumu incelenmektedir (SAS 2014). Bir ARIMA modeli, ARIMA (p, d, q) olarak tanımlanmakta her bir model değişkeni aşağıdaki gibi ifade edilmektedir (Temur, vd., 2019).

p: Otoregresif model derecesi (AR)

d: Fark derecesi

q: Hareketli ortalama modelinin derecesi (MA)

AR modeli zaman serisi içindeki serinin son değerinin kendinden önceki geçmiş zaman verilerine bağımlı olduğunu ifade etmektedir (Attanayake and Perera 2021). Basitçe bir AR modeli denklemi aşağıda formül’de verilmiştir (Çelik 2013).

$y_t = a_0 + a_1 y_{t-1} + \varepsilon_t$ verilen model denklemi birinci dereceden AR(1) modeli olarak tanımlanmakta ve Markov süreci olarak ta bilinmektedir (Akgül, 2003).

y_t : Değişkeninin bugünkü değerini,

a_0 : Sabit sayıyı,

$a_1 y_{t-1}$: a_1 Gecikme değerleriyle bir dönem önceki y_{t-1} var olan değerlerinin çarpılmasıyla elde edilen değerleri ve ε_t : Öngörülemez hata (beyaz gürültü) değerini göstermektedir.

Bu modelin genişletilmesinin otomatik regresyon AR(p) olarak ifade edilmesi aşağıdaki gibi yazılabilir (Akgül, 2003; SAS 2014);

$$y_t = a_0 + a_1 y_{t-1} + a_2 y_{t-2} + \dots + a_p y_{t-p} + \varepsilon_t$$

$y_t = t$ zamanında öngörüsü yapılacak veri,

$y_{(t-p)}$: farklı gecikme	}	y_{t-1} = t zamanından bir önce oluşmuş veriler,
dönemlerindeki değişken		y_{t-2} = t zamanından iki önce oluşmuş veriler,
değerini göstermektedir.		y_{t-3} = t zamanından üç önce oluşmuş veriler,

$a_0, a_1, \dots, a_p \rightarrow$ AR modeli katsayıları,

$\varepsilon_t \rightarrow$ Hata değerlerinin toplamını göstermektedir (Uzundumlu, vd., 2019).

ARIMA zaman serisi modelinin verilerinin işlenmesine başlanmadan önce zaman serileri verilerini Microsoft Excel programında bir sütun içinde t, t_1, t_2 ve t_3 serileri verilerini t baz alınarak üç dönemlik gecikme verileri aktararak analizlere tabii tutulmaktadır.

Denklem 216'da $p=1$ diğer ifadesiyle AR (1) 1'ci derece otoregresif modeller ve $p=2$ diğer ifadesiyle AR (2)'de 2'nci derece otoregresif modeller olarak adlandırılırlar. AR (p) diğer ifadeyle p'nci dereceden otoregresif serilerin oto ortak varyansları fonksiyonunu aşağıda verilmiştir (Yeşilyayla 2013; Mech 2017);

$$\gamma_h = \theta_0 + \theta_1 \gamma_{h-1} + \theta_2 \gamma_{h-2} + \dots + \theta_p \gamma_{h-p} + \varepsilon_t \text{ ve } h > 0 \quad (2.5)$$

AR (p) yani p'nci dereceden otoregresif serilerin otokorelasyon fonksiyonunu aşağıda verilmiştir (Yeşilyayla, 2013; Mech, 2017);

$$\rho_h = \theta_0 + \theta_1 \rho_{h-1} + \theta_2 \rho_{h-2} + \dots + \theta_p \rho_{h-p} + \varepsilon_t \text{ ve } h > 0 \quad (2.6)$$

ARIMA modelinde model tespiti için ACF ve PACF yapıları incelenerek anlamlı katsayılar sahip olmayan seriler durağan olmadığı tespit edilmektedir. Uygun gecikme sayıları ile serilerin durağan hale getirilmesi gerekmektedir (Kasap ve Güney 2010). ARIMA (p, d, q) sürecinin teorik PACF'si genellikle ilk p gecikmelerinde sıfırdan farklı PACF gösterir, geri kalan gecikmeler ise sıfır PACF'ye sahiptir. Teorik ACF içinse ilk q gecikmeleri sıfırdan farklı ACF bildirirken, geri kalan gecikmeler sıfır ACF'ye sahiptir.

Böylelikle, q ve p değerlerini, sırasıyla ACF ve PACF sınırlarının dışında kalan katsayılar olan %5 seviyesindeki anlamlı gecikmelerin toplam sıklığı (frekansı) ile belirleyebiliriz (Gujarati, 2011). P, d ve q değerleri yanlış seçilirse, türetilen modeller yetersiz olur ve bu nedenle öngörü için kullanılamaz.

Tablo 3.1. ACF ve PACF'nin teorik davranışları (Sevüktekin ve Çınar, 2017)

Model	Otokorelasyon Fonksiyonu	Kısmi Otokorelasyon Fonksiyonu
AR(p)	Üstel/geometrik olarak azalır	q -gecikmeden sonra kesilir
MA(q)	p-gecikmeden sonra Kesilir	Üstel/geometrik olarak azalır
ARMA(p, q)	Üstel/geometrik olarak azalır	Üstel/geometrik olarak azalır

ARIMA modelinin sağlıklı sonuç verebilmesi için öncelikle veri setinin durağanlığını test etmek için Augmented Dickey-Fuller ve Phillips-Pheron birim kök testleri yapılmaktadır. Test sonuçlarına göre veri setinin birim kök içerdiği, dolayısıyla durağan olmadığı zamanlarda verilerden birinci, ikinci veya üçüncü dereceden verilerin farkları alınarak veriler durağan yapılabilmektedir (Cançelik, 2021). Bu testin açıklanması için uygulanan modelin denklemi aşağıda verildiği gibi olduğu varsayılırsa;

$$y_t = \rho y_{t-1} + u_t$$

u_t = rastgele hata terimi.

Modelin eşitliğinin her iki tarafından y_{t-1} çıkarılmasından,

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

$(\rho - 1) = \gamma$ oluşmak üzere formül aşağıda verildiği gibidir.

$$\Delta y_t = \gamma y_{t-1} + u_t$$

$$H_0: \rho = 1 \text{ ve } H_1: \rho < 1$$

$(\rho - 1) = 0$ ve ya $\gamma = 0$ olduğu durumda y_t serisi bir birim kökü içerdiği

anlaşılmaktadır. Ancak $|\rho| < 1$ olduğu durumlarda serinin durağanlığı anlaşılmaktadır. Bu denklemde mutlak olarak hesaplanmış “t” değerinin, mutlak olarak hesaplanması Dickey-Fuller veya McKinnon Dickey-Fuller kritik değerlerinden küçük olması durumunda $H_0: \rho = 1$ fonksiyonu reddedilmektedir ve $H_1: \rho < 1$ fonksiyonu kabul edilmesiyle serinin birim kökü olmadığı ve bu sebeple serinin durağan olduğu ortaya çıkmaktadır (Hanedar, vd. 2006).

Ayrıca Anderson (2003)’te aynı durumları aşağıda verilen denklemlerle açıklamıştır. $H_0: \rho = 1$, durumu için durağan değildir; $y_t = \alpha_1 + \beta T + \varepsilon_t$ sabit eğilimde, T bir zaman eğiliminde $y_t = \alpha_2 + \rho y_{t-1} + \varepsilon_t$ rastgele yürüyüş, T bir zaman eğiliminde denklem durağan değildir. Denklem de sabit bir artış ve sonraki denklem’de ise serinin rastlantısal şekilde bir önceki konumuna dönmediği anlamına gelmekte ve rastlantısal bir

şekilde dolanmaktadır. Bu durumun rastgele yürüyüş modeli olarak tanımlandığı bilinmektedir. $H_0: \rho < 1$, durumu durağandır; $y_t = \alpha + \beta T + \rho y_{t-1} + \varepsilon_t$ değişkenin bir eğilimi, t bir zaman eğiliminde, durağanlığın testi için otokorelasyonun olup olmadığının belirlenmesini sağlayan testlerden biriside Durbin-Watson Testi'dir. Modelde olası otokorelasyon problemini inceleyen Durbin-Watson istatistiğinin de 2 civarında olması hata terimleri arasındaki korelasyonun düşük olduğunu göstermektedir (Cançelik, 2021).

$$H_0: \rho = 0 \text{ ve } H_1: \rho \neq 0$$

$$d = \frac{\sum (u_t - u_{t-1})^2}{\sum u_t^2} \quad (2.7)$$

Eğer Durbin-Watson Testi $d=2$ ise durağanlığın olduğunu göstermektedir.

Ayrıca,

$$E(Y_t) = \mu \text{ (bütün } t \text{ 'ler için)}$$

$$\text{Var}(Y_t) = E(Y_t - \mu)^2 = \sigma^2 \text{ (bütün } t \text{ 'ler için)}$$

$$\text{Cov}(Y_t, Y_{t+k}) = \gamma_k \text{ sabit (bütün } t \text{ 'ler için tüm } k \neq 0 \text{ için)}$$

Bir serinin yukarıdaki 3 şartı sağlayarak yani serilerin veri ortalaması, varyansı ve otovaryansı sabit kalması ile seri durağan olabilmektedir (Cançelik, 2021).

Trend ve Fark Durağan Süreç serileri durağanlaştırma yöntemleri olup seriyi trendden arındırmak için, trend modele eklenmekte ve modele EKK uygulanmaktadır. Burada artıklar trendden arındırılmış olmaktadır. Fark Durağan Süreçte ise bu serinin farkı alınarak seri durağanlaştırılmaktadır (Uğurlu 2019). Durağanlık testinin ilk aşamasında EKK yöntemi uygulanarak hata teriminin tahmini yapılmakta ve ikinci aşamada da bu hata terimleri çekilip verilerde birim kök araştırması yapılabilmektedir (Ata ve Yücel 2003).

Zaman serilerinde gecikmeli bulunan hata terimi son bulunan hata terimine etki edebiliyorsa burada MA Modelleri dikkate alınmalıdır (Yeşilyayla, 2013). Bu modeller zaman serisinin belirli bir dönemindeki gözlem ve hata teriminin belirli sayıda geçmiş dönem hata terimleri ile olan doğrusal ilişkisi olarak ifade edilmektedir. MA modelleri geçmiş dönemlerdeki hata terimleri sayılarına göre bir, iki, üçüncü hatta genelde q 'nuncu dereceden MA modelleri olarak isimlendirilmektedirler (Karakaş, 2019).

Basit bir MA (q) model; $y_t = \varepsilon_t - \theta \varepsilon_{t-1}$ verilen denklem birinci dereceden MA(1) modelini tanımlamakta olup formülün değişkenleri şu şekilde ifade edilmektedir (Kaya, 2019).

y_t : Bugünkü verileri,

ε_t : Bugünkü dönemin hata terimlerini,

ε_{t-1} : Bir önceki dönemin hata terimlerini ve

θ : Bilinmeyen parametreleri ifade etmektedir.

MA (q) modelinin y_t ortalaması sıfır, $h > q$ için $\gamma(q) > 0$ ve γ_h (Oto ortak Varyans fonksiyonu) durağan olduğunda q'nuncu seviyede hareketli ortalama serisinin denklemi aşağıda verilmiştir (Çelik, 2013; Yeşilyayla, 2013).

$$y_t = \mu + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \dots - \theta_q \varepsilon_{t-q}$$

Buradaki ε_t , $WN(0, \sigma^2)$ denkleminde verilen beyaz gürültü serileridir.

MA (1) ve MA (2) aşağıdaki şekilde ifade edilirler (Akgül 2003).

$$MA(1) y_t = \mu + \varepsilon_t - \theta_1 \varepsilon_{t-1}$$

$$MA(2) y_t = \mu + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2}$$

Zaman serisinin durağan olması durumunda ekstra bir işleme tabi tutulmadan ARMA (p, q) modelinin uygulanmasına geçilmektedir. Ancak serinin durağan olmaması durumunda seri durağanlaştırılmak için geciktirme işlemi yapılmasıyla seri durağanlaştırılmaktadır (Yeşilyayla 2013). Geciktirme işlemleri serinin durağan olabilmesi için işlem sürdürülmeye devam ettirilir (Uzundumlu, vd., 2018). Zaman serilerinde 2 yıl için uygulanabilen geciktirme denklemi aşağıda verilmektedir.

y_t denkleminde AR”, “I” ve “MA”, burada “AR” otoregresif anlamına gelir, “I” fark anlamına gelir ve “MA” Hareketli ortalamaya eşittir (Yang ve diğerleri, 2021). ARMA modeli için uygulanacak denklem aşağıda verilmiştir.

$$y_t = a_0 + a_1 y_{t-1} + \dots + a_p y_{t-p} + \varepsilon_t - a_1 \varepsilon_{t-1} - a_2 \varepsilon_{t-2} - \dots - a_q \varepsilon_{t-q}$$

ARMA modeli durağan olması durumunda ARIMA modeline dönüşmektedir (Çelik 2013). $y_t (a_0 + a_1 y_{t-1} + \dots + a_p y_{t-p}) * (1 - B)^d + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \dots - \theta_q \varepsilon_{t-q}$ Ayrıca Kadılar (2009) ARIMA modelinin denklemini (2.14)'da ki gibi ifade etmektedir.

$$(1 - a_1 B - a_2 B^2 - \dots - a_p B^p) * (1 - B)^d y_t = (1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q) \varepsilon_t$$

$(1 - B)^d$ terimini d'ninci dereceden fark alma işlemi olduğu, $(1 - B)^d y_t$ ifadesi ise $d=1$ için $B y_t = y_{t-1}$ şeklinde yazılabilmektedir. Ayrıca $d=2$ için $B^2 y_t = y_{t-2}$ veya $B y_{t-1} = y_{t-2}$ şeklinde yazılabilmektedirler.

$p=0, d=1$ ve $q=1$ olması durumundaki modeli ARIMA (0, 1, 1) yani $(1 - B)^{d=1} y_t = (1 - \theta B)\varepsilon_t$ olduğu için $y_t - B y_t = \varepsilon_t - \theta B \varepsilon_t$ ise $y_t - y_{t-1} = \varepsilon_t - \varepsilon_{t-1}$ ve $y_t = y_{t-1} + \varepsilon_t - \theta \varepsilon_{t-1}$ ve $-1 < \theta < 1$

$p=1, q=0$ ve $d=1$ olması durumundaki modeli ARIMA (1, 1, 0) yani $(1 - a_1 B)(1 - B)^{d=1} y_t = \varepsilon_t$ olduğu için $(1 - B)^{d=1} y_t = y_t - B y_t$ 'ye eşit olmaktadır. İşlemin biraz daha açılmasıyla $(1 - B)^1 y_t = y_t - y_{t-1}$ eşitliğini oluşturmaktadır. $(1 - a_1 B)(y_t - y_{t-1}) = \varepsilon_t$ açılımının yapılması $y_t - y_{t-1} + (-a_1 B y_t) + (a_1 B y_{t-1}) = \varepsilon_t$ ve $y_t - y_{t-1} + a_1(-y_{t-1} + y_{t-2}) = \varepsilon_t$ $y_t = y_{t-1} - a_1(-y_{t-1} + y_{t-2}) + \varepsilon_t$ veya $y_t = y_{t-1} + a_1(y_{t-1} - y_{t-2}) + \varepsilon_t$ olmaktadır.

$p=1, q=1$ ve $d=1$ olması durumundaki modeli ARIMA (1, 1, 1) yani $(1 - a_1 B)(1 - \theta B)(1 - B)^{d=1} y_t = \varepsilon_t$

$(1 - B)^{d=1} y_t = (1 - \theta B)\varepsilon_t$ olduğu için $(1 - B)^{d=1} y_t = y_t - B y_t$ 'ye eşit olmakta, işlemin açılımı ile $(1 - B)^1 y_t = y_t - y_{t-1}$ olmaktadır. $(1 - a_1 B)(y_t - y_{t-1}) = \varepsilon_t - \theta \varepsilon_{t-1}$ olmakta ve $y_t - y_{t-1} + (-a_1 B y_t) + (a_1 B y_{t-1}) = \varepsilon_t - \theta \varepsilon_{t-1}$

$y_t - y_{t-1} + a_1(-y_{t-1} + y_{t-2}) = \varepsilon_t - \theta \varepsilon_{t-1}$ eşitliğini sağlamak ve

$y_t = y_{t-1} - a_1(-y_{t-1} + y_{t-2}) + \varepsilon_t - \theta \varepsilon_{t-1}$ veya

$y_t = y_{t-1} + a_1(y_{t-1} - y_{t-2}) + \varepsilon_t - \theta \varepsilon_{t-1}$ olmaktadır.

2.5.1.2. Parametrelerin Tahmin edilmesi

Model belirleme işlemi tamamlanınca modelin parametre tahminleri yapılmaktadır. Model tahmin yöntemleri klasik (EKK ve En Çok Olabilirlik) ve sağlamcı tahmin yöntemleri olmak üzere ikiye ayrılmaktadır (Dönmez 2014). AR modeli parametre tahmininde, EKK, MA modeli için en çok olabilirlik veya EKK yöntemi ve ARMA modeli içinse doğrusal olmayan optimizasyon yöntemi parametre tahmininde kullanılmaktadır (Griffiths 1993; Erdoğan 2006).

Bu çalışmada EKK yöntemi kullanılarak seçenekler içerisinde en uygun parametre tahminleri yapmak arzulanmıştır. Model için kullanılmış parametrelerden modelin önemini açıklayan ve kimi verilerin model için gerekliliğini belirleyen "t" istatistik değeridir (SAS, 2014). En iyi model için değerleri için genelde en küçük olan AIC ve

BIC değerlerine bakılmakta olup karşılaştırılan modellerde en küçük AIC değeri tercih edilerek model seçilmelidir. Ancak bilimsel çalışmalarda BIC modelinin AIC modelinde nazaran daha çok kullanıldığı bunun nedenlerinden biri AIC modelinin sahip olduğu büyük sayılar nedeniyle büyük hesaplamalar gerektirmesidir (Ekmekci 2016). Bu çalışmada AIC, BIC, DW, HQC, MAE, MAPE, SSE, RMSE, SBC, MSE ve R2 sonuçlarına göre en uygun modeller belirlenmiştir. Genelde DW'ın 2,00'a ve R2'değerinde 1,00'e yakın olması istenmekte AIC, BIC, HQC, MAE, MAPE, SSE, RMSE, SBC ve MSE'de ise en küçük değerler en iyi modeli oluşturmaktadır. Bu 11 kriterin en fazlasını sağlayan model en uygun model olarak kabul edilmiştir.

AIC hesaplaması için kullanılan fonksiyon aşağıda verilmektedir (Aho ve diğerleri 2014). $AIC = n \ln \hat{\sigma}^2 + 2P$ denklemde, P model için parametrelerin sayısı olduğu, $P = p+q+1$ işlemiyle elde ederler. n gözlenen veri sayısını ve $\hat{\sigma}^2$ minimum kalıntı-varyans kriterleridir (Ucal, 2006; Aho, vd., 2014). BIC hesaplaması için gerekli fonksiyon aşağıda verilmektedir (Aho, vd., 2014). $BIC (SIC) = n \ln \hat{\sigma}^2 + P \ln n$

2.5.1.3. Teşhis ve Denetim Yapılması

Elde edilen parametrelerle geçici modelin en uygun model olduğu uygunluk testleri ile belirlenmektedir. İlk başta tahmin hataları serisi oluşturulmakta ve sonra bu hata serileri için otokorelasyon katsayıları hesaplanmaktadır. Eğer tahminlerin hataları otokorelasyon katsayılarının seyri standart hata limitleri ile karşılaştırılarak fazla bir fark olmadığı anlaşılırsa geçici model nihai en uygun model olduğuna karar verilmektedir. Eğer fark büyük ise geçici model en uygun model olmayıp süreç bir önceki adıma giderek yeniden model belirlenmektedir (Duru, 2007).

Ayrıca modelin yeterliliğinin uyumuna ve kalıntı analiz verilerine de bakılabilmektedir.

Box-Pierce tarafından geliştirilen Q istatistiği ile Otokorelasyon katsayılarını bireysel incelemektense belirli sayıdaki hataların otokorelasyonunu bir arada toplu incelemek modelin uygun olup olmadığını daha açık ortaya koyabilmektedir.

$$Q = n \sum_{k=1}^K r_k^2 \quad \text{ve } k=1,2,3,\dots \quad (2.8)$$

r_k^2 : örneklem tahmin hatalarının, çeşitli gecikmelerdeki otokorelasyon katsayısını,

N : Örneklem hacmini,

d : Fark alma derecesini,

n: N-d'yi ve

K : hesaplanan otokorelasyon sayılarını göstermektedir.

Q istatistiki yaklaşık olarak X^2 dağılımını göstermekte ve serbestliklerinin derecelerini mevsimsel olmayan modellerde (K - p - q) olarak verilmektedir. Bu testler sonucunda hataların otokorelasyon katsayılarının sıfırdan farklı olup olmadığına yani hatalar serilerinin rassal seriler olup olmadıklarına karar vererek, modellerin uygunluğu test edilmektedir. Hesaplanan Q değeri (K- p - q) serbestlik derecelerinde ve verilen anlam seviyelerindeki X^2 tablo değerinden küçükse hatalar serilerinin rassal olduğunu ve modelin uygun olduğunu göstermektedir (Duru, 2007).

2.5.1.4. Öngörü (Tahmin)

En uygun ARIMA modeline karar verdikten sonra zaman serisi analizlerinin en temel ikinci amacı olan gelecek değerlere ilişkin öngörüler elde edilebilir. Özetle, zaman serilerinin gelecekteki değerleri öngörülebilir ve öngörülerin güven aralıkları oluşturulabilir. ARIMA modellemesinin popülerliğinin arkasındaki nedenlerden biri öngörüdeki başarısıdır. Birçok durumda, bu teknikle elde edilen öngörüler, özellikle de kısa vadeli öngörüler, geleneksel ekonometrik modellemeden elde edilenlere nazaran daha güvenilirdir (Gujarati ve Porter, 2003).

Uygunluk testleri ile uygunluğu sağlanan p ve q değerleri modelde kullanılarak en uygun model belirlenmiş olmaktadır. En uygun model için yazılan p ve q değerleriyle sabit ortalama ve varyansa sahip olan tahmin modeli oluşturulmaktadır. Tahmin yapmada belli dönemlerdeki gözlenmiş değerlerle veya gözlenecek değerlerin gerçek değerlere yakın öngörüleme yapılması amaçlanmaktadır (Erdoğan, 2006).

3. UYGULAMA

3.1. Genel Bilgi

Bu çalışmada ARIMA modeli kullanılarak 10 farklı ülkeye yönelik COVID-19 aktif vaka oranları tahmin edilmeye çalışılmıştır. Bunun için literatür kısmında belirtilen 4 aşamalı bir süreç takip edilmiştir. İlk aşamada model belirlenmiş, ikinci aşamada parametreler tahmin edilmiştir. Üçüncü aşamada teşhis ve denetim yapılarak son aşamada öngörü tahmin çalışması gerçekleştirilmiştir.

31 Aralık 2019'da Çin'in Wuhan kentinde ilk tespit edildiğinden bu yana, bu patojen hızla yayıldı ve Mart 2020 itibarıyla 300.000'den fazla ölümle dünya çapında 4 milyondan fazla insanı enfekte etti. Bu virüs, 2.5 ila 3.5 arasında tahmin edilen bir temel üreme sayısı (R_0) ile oldukça bulaşıcıdır (Gan, et al., 2020). Bu hastalık, eşlik eden hastalıkları %2 ile %18 arasında değişen yaşlı nüfus arasında özellikle yüksek ölüm oranlarına sahiptir (Onder, et al, 2020). COVID-19 için özel bir tedavi, tedavi veya aşı bulunmadığından, dünya çapında sağlık sistemlerinin bu pandemiye yönetmek için karşılaştığı zorluklar çok büyük. Salgın yönetimi ve kaynak planlaması açısından sağlık sistemlerinin hazırlıklı olması için COVID-19 durum eğilimlerinin doğru tahmin edilmesi esastır. Bulaşıcı hastalıkların matematiksel ve istatistiksel modellenmesi, otoregresif hareketli ortalama (ARIMA) ve otoregresif mevsimsel hareketli ortalama (SARIMA) gibi sağlık modellerinin bulaşıcı hastalık eğilimlerinin istatistiksel modellenmesi ve tahmini için büyük ölçekte yapılmasını sağlayacak güçlü araçlardır (Allard, 1998). ARIMA modelleri, akut enfeksiyon hastalıklarının kısa vadeli etkilerine yönelik araştırmalara uygun oldukları ve birçok yola uymaya uygun esnek bir model sınıfı olduğu için bu bağlamda tercih edilmektedir ve literatürde iyi bir şekilde belgelenmiştir (Allard, 1998; Imai and Hashizume, 2015). Ayrıca mevcut durumda, ARIMA modelleri, veri noktalarının bir yıldan daha kısa bir süreyi kapsaması ve herhangi bir mevsimsel değişimin etkisini ortadan kaldırması nedeniyle SARIMA muadillerine göre tercih edilmektedir. ARIMA modelleri, COVID-19 salgınlarındaki eğilimleri tahmin etmek için İtalya, Brezilya, Çin, İran, Tayland ve Güney Kore'deki çeşitli çalışmalarda kullanılmıştır (Sing, et al, 2020).

3.2. Veri Seti

Çalışmada ARIMA modeli; 3 Ocak 2020 ile 7 Aralık 2021 tarihleri arasında günlük olarak teyit edilmiş yeni COVID-19 vakalarını tahmin etmek için kullanılmıştır.

Çalışmaya 10 ülke dahil edilmiştir. Bu ülkeler; Türkiye, Brezilya, Rusya, Avustralya, İran, Amerika, Çin, Fransa, Almanya ve Suudi Arabistan. Veriler Dünya Sağlık Örgütü'nün resmi web sitesinden elde edilmiştir (<https://covid19.who.int/>).

ARIMA modelleri, gözlemlenen zaman serisi verilerinin üzerine uygulanmıştır. ARIMA modelini doğrulamak ve modelin uygunluğunu ve tahminlerin doğruluğunu belirlemek için eğitim testi çerçevesi kullanılmıştır. Bu çerçevede, tahmin doğruluğunun güvenilir olduğunu ve ARIMA modelinin genel özelliklerinin aşağıdaki gibi olduğunu bulunmuştur: (p, d, q) burada p, otoregresif bileşenin (AR) sırasını temsil eder, d farkın derecesidir. . Orijinal zaman serisinde d, orijinal zaman serisindeki farkın derecesidir ve q, hareketli ortalama bileşenin sırasındır.

Hareketli ortalama bileşenin (MA) sırası. Uygun bir ARIMA modelinin belirlenmesi üç aşamayı içeriyordu - model tanımlama, parametre tahmini ve model doğrulama.

Kararlılığı doğrulamak ve p, d ve q parametrelerinin koşulları için spesifikasyonlar sağlamak için orijinal zaman serisi verilerinden otokorelasyon fonksiyonu (ACF) ve kısmi otokorelasyon fonksiyonu (PACF) çizimleri alınmıştır. Parametre tahmini adımı, ARIMA modelinin AR ve MA bileşenlerine karşılık gelen bilinmeyen katsayılar çok değişkenli regresyon analizi ile tahmin edilmiştir. Yukarıdaki prosedürler ARIMA modelleme metodolojisinde standarttır ve tüm adımları gerçekleştirmek için SPSS 24.0 istatistiksel yazılımı kullanılmıştır.

Günlük COVID-19 vakalarının, varyans tahmini ve normallik testi ile belirlendiği üzere yüksek derecede bir varyans (gürültü) ve çarpıklık sergilediği bulunmuştur. Gürültü bileşenini kaldırmak için, verilerin daha yumuşak MA noktası kullanılarak hareketli bir pencere üzerinden ortalaması alınmıştır. Tahmin modelinin geliştirilmesi için dikkate alınan ortak değişkenler (1) yeni COVID-19 vakaları, (2) kümülatif vakalar ve (3) yeni ölümler (4) kümülatif ölümler arasından seçilmiştir.

3.3. Bulgular

COVID-19 salgını ile ilgili araştırma verileri, 10 ülkenin 1/4/2020 ile 7/12/2021 arasındaki zaman serilerinden oluşmuştur. Dünya Sağlık Örgütü'nün resmi web sitesinden ulaşılan veriler, SPSS 24.0 istatistik programı ile analiz edilmiştir.

Jenkins Box metodolojisinin ilk aşaması olan veri toplama sürecinden sonra, veri serisinin davranışını tespit etmek için 1/4/2020 ile 7/12/2021 tarihleri arasında ülke grubu için COVID-19 salgını veri serisinin çizimine geçilmiştir.

3.3.1. İran İçin Model Belirleme ve Tahmin

Modelin sıralamasını daha kesin olarak belirlemek için dört özellik modeli kontrol edilmiş ve farklılaştırma kriterlerine göre en iyisi seçilmiştir: MAPE, RMSE, R - kare, R _ durağan kare, BIC, MAE. Değerleri Tablo 3.1’de verilmiştir.

Tablo 3.1. Zaman serileri için çeşitli Box-Jenkins modelleri

MODELER	ARIMA	R_ stationary squared	R - squared	RMSE	MAPE	MAE	BIC
1	(2,1,2)	0,362	0,976	1407,83	8,216	765,45	14,55
2	(1,2,1)	0,219	1,000	1481,71	0,42	804,023	14,63
3	Wintes.	0,570	0,970	25,263	10,863	18,382	6,49
4	(0,2,6)	0,311	1,000	25,919	0,069	18,207	6,52

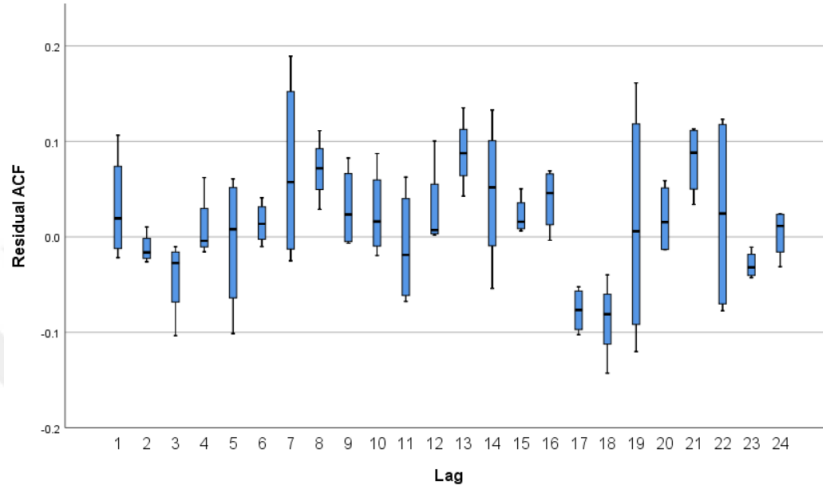
Farklılaştırma kriterlerinin sonuçları ve seçilen modellerin parametrelerinin önemi sayesinde, Jenkins Box modellerinden en iyi modelin ARIMA (0.2.6) modeli olduğu sonucuna varılabilir, çünkü bu model için standart değerler, modeller arasında en az olası olanlardır.

Tablo 3.2. Olağan modelin parametreleri

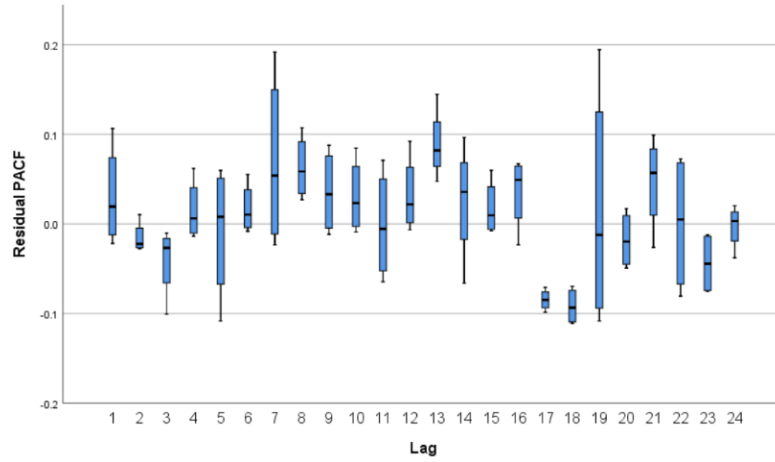
Estimate	SE	t	sig
0,508	0,035	14,560	0,00
-0,111	0,035	-3,173	0,002
-0,101	0,038	-2,663	0,008
0,894	0,035	25,877	0,000
0,689	0,059	11,630	0,000

Ve p değerinin değerine dayalı t-testi yoluyla, p değerinin değeri (0.05)'ten küçük olduğu için (0.2.6) formunun anlamlı olduğu fark edilmektedir.

Tahmin edilen modelin artıklarının otokorelasyon ve kısmi otokorelasyon fonksiyonları çıkarılıp çizilmiştir. Böylece Şekil 3.1, 3.2, 3.3 ve 3.4'ten otokorelasyon katsayılarının ve kısmi otokorelasyon değerlerinin değerlerinin olduğu sonucuna varılmaktadır. Önerilen modeldeki artıkların otokorelasyonu güven sınırları içindedir (%95). Bu, artık serisinin rastgele değişkenleri temsil ettiği anlamına gelmekte, bu da tahmin edilen modelin iyi, verimli bir tahmin için uygun olduğunu göstermektedir.

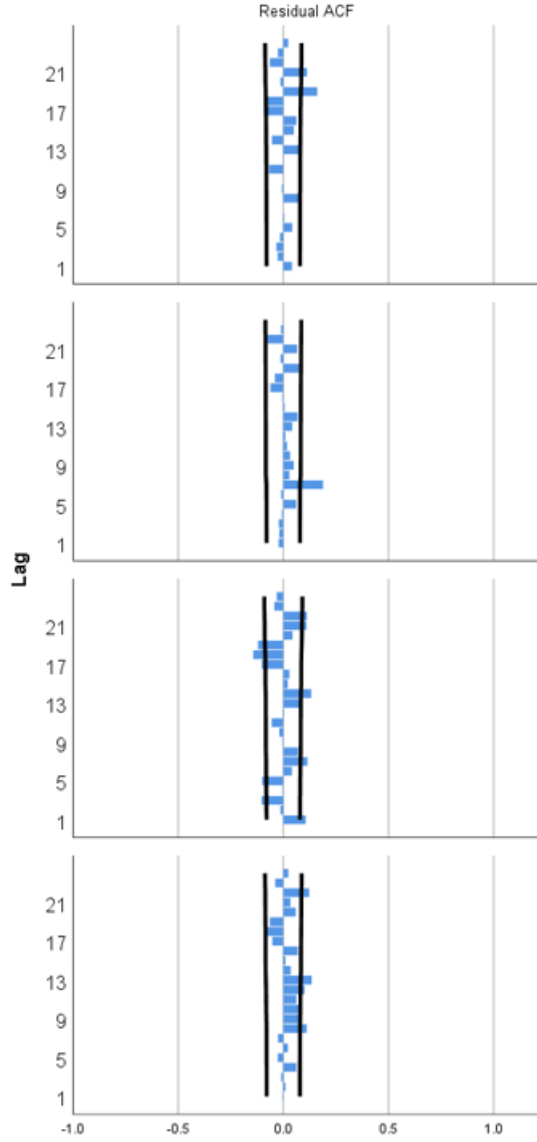


Şekil 3.1. Otokorelasyon işlevi (ACF)



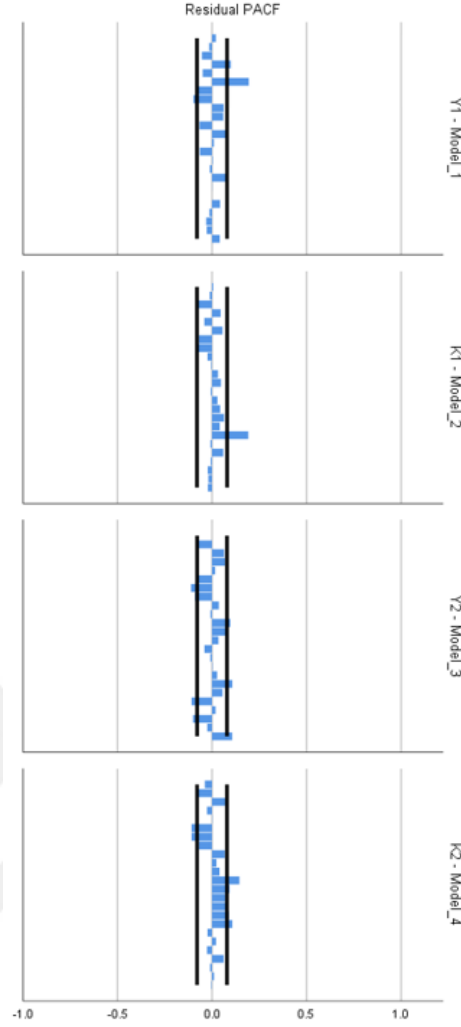
Şekil 3.2. Kısmi otokorelasyon (PACF)

Şekil 3.1. ve 3.2 aracılığıyla, verilerin genel bir eğilim aldığı için ortalama olarak zaman serisinin kararsızlığı ile dalgalandığını fark ettik ve daha fazla doğruluk için sırasıyla otokorelasyon fonksiyonunu (ACF) ve kısmi otokorelasyonu (PACF) çizilmiştir.



Şekil 3.3. Otokorelasyon katsayıları

Şekil 3.3'te görüldüğü üzere otokorelasyon katsayılarının (ACF) üstel olarak azalmaya başladığını ve anlamlılık düzeyinde güven sınırlarının dışına çıktığı görülmüştür (%95).

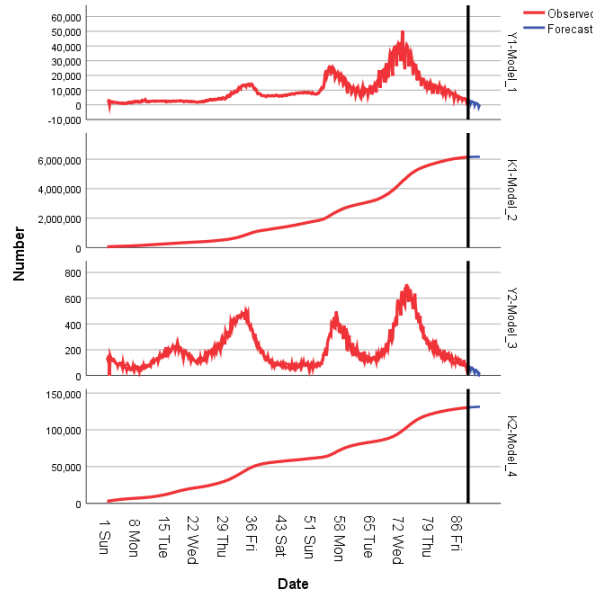


Şekil 3.4. Kısmi otokorelasyon katsayıları

Kısmi otokorelasyon katsayılarının (PACF) davranışı hakkında Şekil (3.4)'ten ayrıca yer değiştirmenin (12) kısmi otokorelasyon katsayıları için güven sınırlarının dışında olduğunu ve bu durumun zincirde kararlılık eksikliğini gösterdiğini not ediyoruz.

Serinin ortalamadaki kararsızlığından kurtulmak için farklarını alıyoruz ve eğer birinci ve ikinci fark alındıktan sonra durağanlık sağlanıyorsa ortaya çıkan serinin resimli hali aşağıdaki şekildeki gibi oluyor (3.4) (3.3) Serinin durağan hale geldiği görülüyorsa, otokorelasyon fonksiyonu ve ceza otokorelasyonu çizildi ve katsayılarının güven sınırları (± 0.2) içinde olduğu farklar alındıktan sonra %95 güven düzeyi ile yani veriler Şekil 3.3 ve 3.4'teki gibi sabittir.

Modeli tanı testleri analiz ettikten sonra, COVID-19 salgını serisinin tahmin değerlerini temsil eden Şekil 3.5'teki modeli tahmin için kullanmak mümkün hale gelmektedir.



Şekil 3.5. Bir serinin tahmin edilen değerleri ile orijinal değerleri arasındaki fark

Kırmızı rengin serinin orijinal değerlerini, mavi renklerin ise güven düzeyi (%95) ile öngörülen değerleri gösterdiği ve Corona salgını katsayılarının pozitif değerler aldığı Şekil (5)'ten fark edilmektedir. 12/08/2021 ile (12/2021/29) arasındaki tahmin döneminin uzunluğu için Genel olarak, bir serinin tahmin edilen değerleri ile orijinal değerleri arasındaki fark yakındı, bu da doğruluğu gösterir. tahminin. Aşağıdaki tabloda Corona salgını, üst limit ve alt limit gözlemlerinin tahmin değerleri gösterilmektedir. Bu tahmin değerleri aylar arasındaki değişkenlerin yer aldığı ARIMA (0,2,6) modeli kullanılarak elde edilmiştir (Tablo 3.3).

Tablo 3.3. COVID-19 salgını gözlemlerinin tahmin değerlerinin, üst ve alt limitleri

	Tahmin değerleri	Üst limit	Alt limit
89 Sun	130363	130414	130313
89 Mon	130435	130525	130344
89 Tue	130435	130639	130369
89 Wed	130565	130749	130381
89 Thu	130623	130864	130383
89 Fri	130688	130992	130384
89 Sat	130751	131127	130375
90 Sun	130824	131285	130363
90 Mon	130886	131440	130332
90 Tue	130945	131599	130290
90 Wed	130997	131759	130236
90 Thu	131047	131922	130172
90 Fri	131102	132096	130108
90 Sat	131155	132096	130108
91 Sun	131320	132473	129963
91 Mon	131270	132668	129873
91 Tue	131320	132866	129774
91 Wed	131364	133065	129663
91 Thu	131406	133268	129544
91 Fri	131452	133481	129424
91 Sat	131497	13398	129295

3.3.2. Almanya İçin Model Belirleme ve Tahmin

Almanya'daki Corona salgını verileri için en iyi modelin teşhis ve tahminine gelince, dört işaret modeli de test edildi ve farklılaştırma kriterlerine göre en iyi model test edildi.

Tablo 3.4. Zaman serileri için ilgili Box-Jenkins modelleri

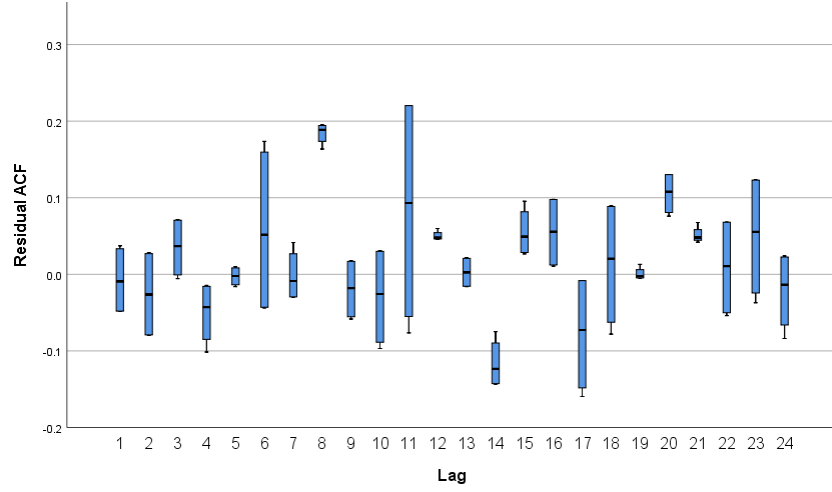
MODELER	ARIMA	R_ stationary squared	R - squared	RMSE	MAPE	MAE	BIC
1	(0,1,3)	0,282	0,961	2458,932	20,250	1285,13	15,647
2	(0,2,3)	0,283	1,000	2478,806	0,093	1320,46	15,663
3	(0,0,6)	0,354	0,931	62,271	50,073	32,803	8,316
4	(0,1,6)	0,352	1,000	62,288	32,764	32,76	8,316

Farklılaştırma kriterlerinin sonuçları ve seçilen modellerin parametrelerinin önemi sayesinde, bu modelin standart değerleri mümkün olan en az olduğu için en iyi modelin ARIMA (0,1,6) modeli olduğu sonucuna varılabilir. Model parametreleri aşağıdaki Tablo (5)'te gösterildiği gibi tahmin edilmiştir.

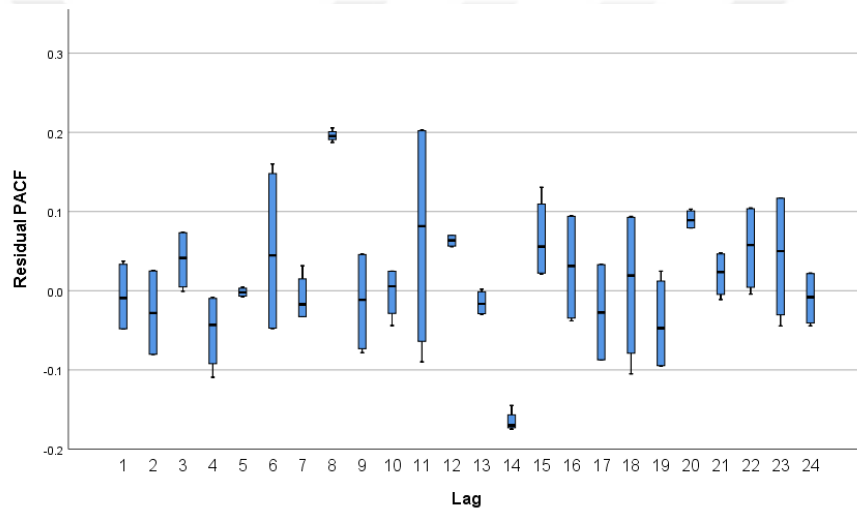
Tablo 3.5. Seçilen modelin parametreleri

Estimate	SE	t	sig
0,678	0,039	-17,244	0,000
-0,247	0,047	-5,262	0,000
-0,084	0,040	-2,118	0,035
-0,133	0,039	-3,395	0,001
-0,274	0,039	-7,115	0,000

P değerine dayalı t-testi ile modelin (0,1,6) anlamlı olduğunu çünkü p değerinin (0.05) den küçük olduğunu not ediyoruz. Otokorelasyon katsayılarının değerleri ve artıkların kısmi otokorelasyonu olarak Önerilen modelde güven sınırları (%95) dahilindedir, bu, artık serisinin rastgele değişkenleri temsil ettiği anlamına gelir ve bu, tahmin edilen modelin aşağıdaki şekiller (6) ve (7)'deki gibi iyi olduğunu gösterir.

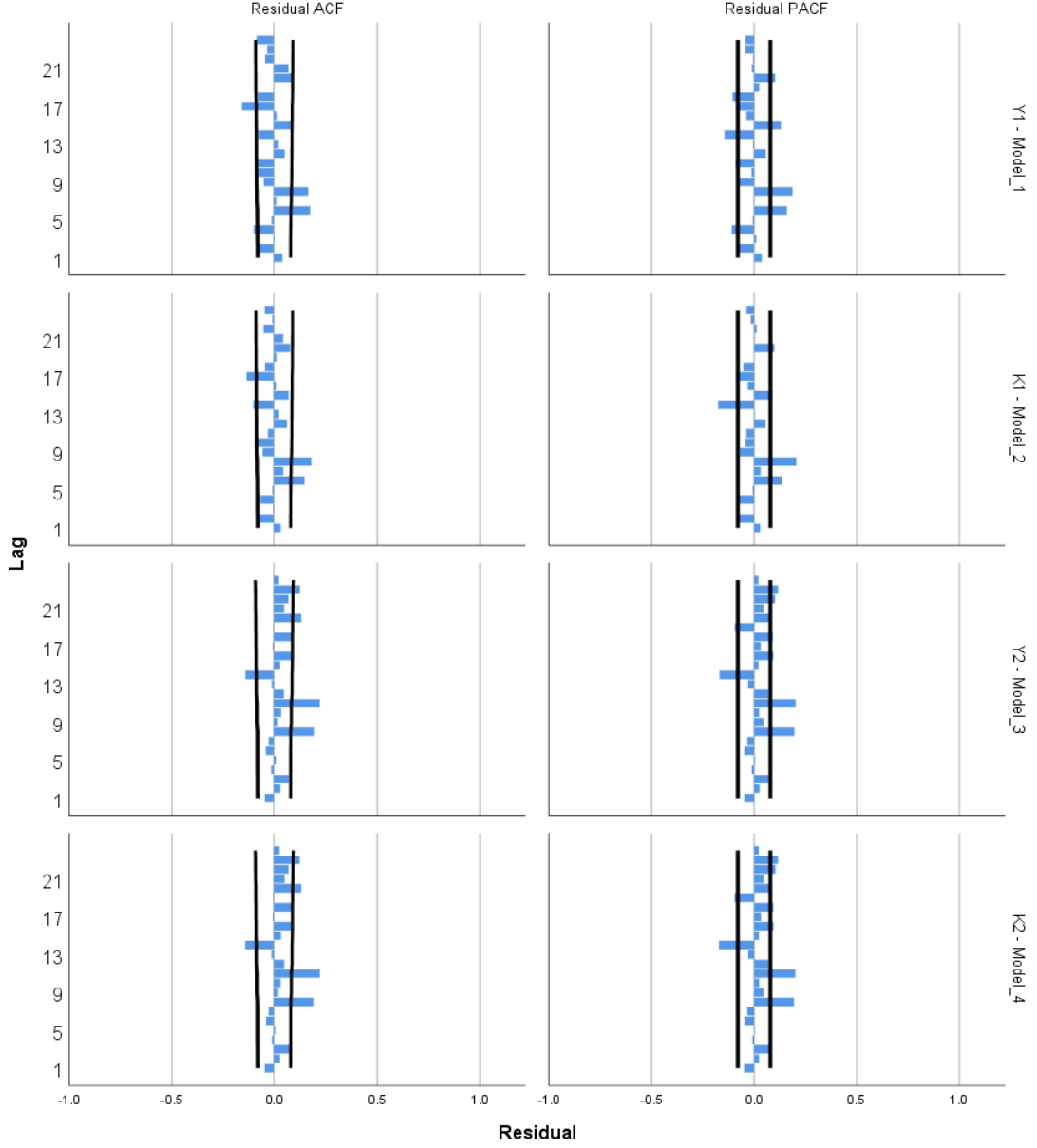


Şekil 3.6. Otokorelasyon



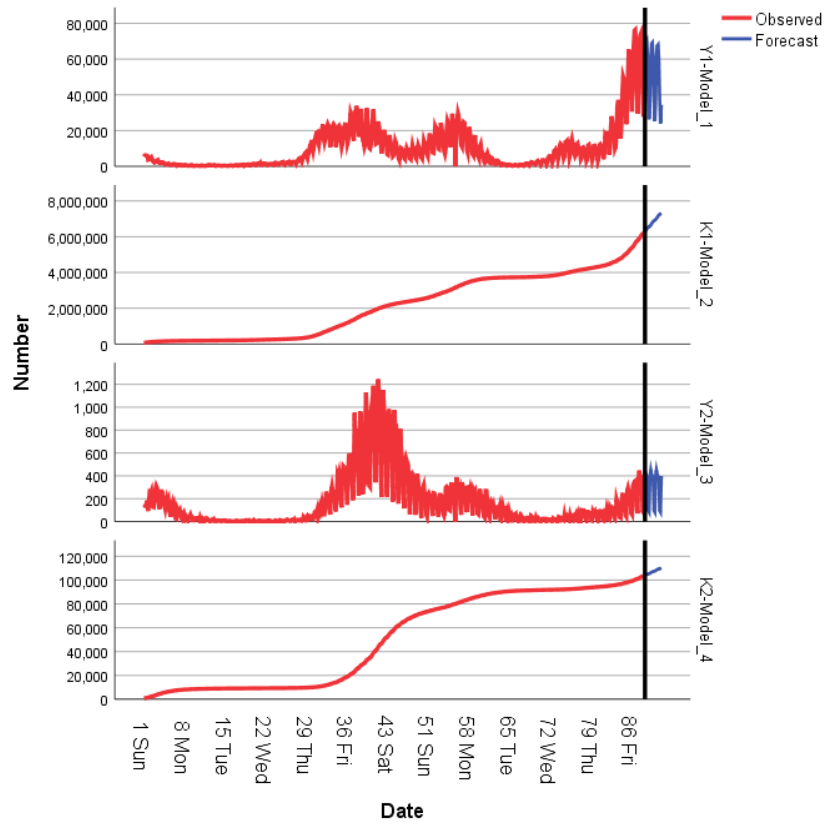
Şekil 3.7. Kısmi otokorelasyon

Modelin test edilmesinden sonra testler, bir serinin tahmin değerlerinin çizilmesini temsil eden Şekil 3.8'deki gibi tahmin için kullanılmıştır ve model kullanılarak (0,1.6), haftalar arasındaki değişimler gösterildiği gibi not edilmiştir (Şekil 3.8).



Şekil 3.8. Otokorelasyon ve kısmi otokorelasyon katsayıları

Şekil 3.9'dan kırmızı renklerin serinin orijinal değerlerini, mavi renklerin ise güven düzeyi (%95) ile öngörülen değerleri gösterdiğini ve COVID-19 salgın katsayılarının tahmin döneminin uzunluğu için pozitif değerler aldığı fark edilmiştir.



Şekil 3.9. Bir serinin tahmin edilen değerleri ile orijinal değerleri arasındaki fark

Tablo 3.6. Gözlemlerin tahmin değerlerinin üst ve alt sınırları

	Tahmin değerleri	Üst limit	Alt limit
89 Sun	103969	104091	103847
89 Mon	104378	104617	104139
89 Tue	104775	105110	104440
89 Wed	105138	105553	104722
89 Thu	105235	105718	104753
89 Fri	105319	105718	104753
89 Sat	105718	106342	105095
90 Sun	106167	106917	105417
90 Mon	106576	107478	105675
90 Tue	106973	108019	105927
90 Wed	107336	108514	106159
90 Thu	107434	108729	106138
90 Fri	107518	108928	106107
90 Sat	107917	109445	106388
91 Sun	108365	110052	106679
91 Mon	108775	110639	106910
91 Tue	109172	111211	107132
91 Wed	109534	111738	107330
91 Thu	109632	111989	107275
91 Fri	109716	112222	107209
91 Sat	110115	112773	107457

3.3.3. Rusya İçin Model Belirleme ve Tahmin

Rusya'daki en iyi görüş modelleri de erdemli kriterlere göre teşhis edildi ve dağılık modeller şu şekildeydi:

Tablo 3.7. Zaman serileri için ilgili Box-Jenkins modelleri

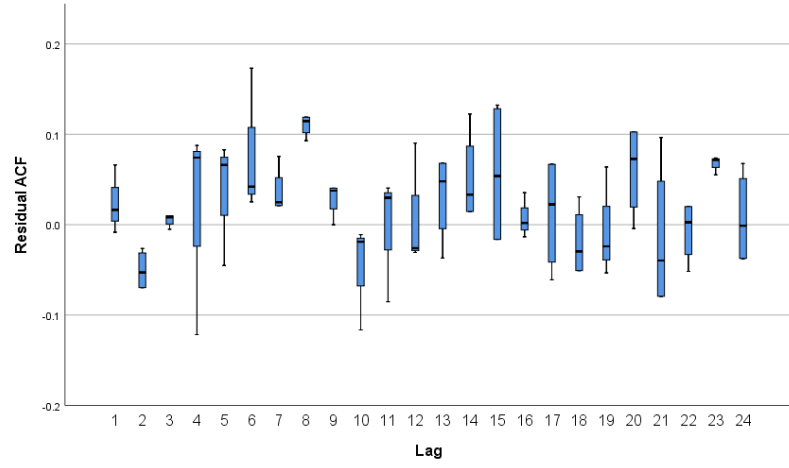
MODELER	R-stat	R-squared	RMSE	MAPE	MAE	BIC
WINTER	0.484	0.966	594.08	3.517	427.358	12.805
3,2,6	0.478	1.000	598.66	0.101	430.574	12.832
_0,1,1	0.497	0.994	24.355	7.194	17.880	6.407
0,2,1	0.496	1.000	27.374	0.158	17.885	6.408

Sonuçlar ve seçilen modellerin parametrelerinin önemi sayesinde, modelin modelde en iyisi olduğu sonucuna varıyoruz (ARIMA (0,1,1), çünkü bu modelin standart değerleri en az olasıdır ve dolayısıyla model parametreleri aşağıdaki tablodaki gibi tahmin edilmiştir (Tablo 3.8)

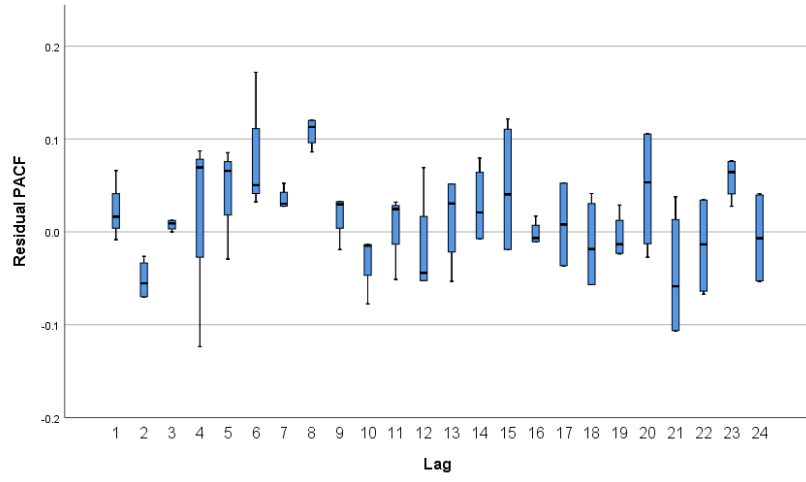
Tablo 3.8. Seçilen modelin parametreleri

ESTI	S,E	t	sig
0,481	0,036	13,400	0,00
0,759	0,027	27,691	0,000

T-testi yoluyla ve p value değerine bağlı olarak, p value değerinin 0,05'ten küçük olması nedeniyle ARIMA (0,1,1) modelinin anlamlı olduğunu not ediyoruz, bu nedenle otokorelasyon değerlerinin ve Önerilen modeldeki artıkların kısmi katsayıları güven sınırları içindedir ve bu, tahmin edilen modelin aşağıdaki şekillerde (10)(11) olduğu gibi iyi ve verimli olduğunu gösterir.

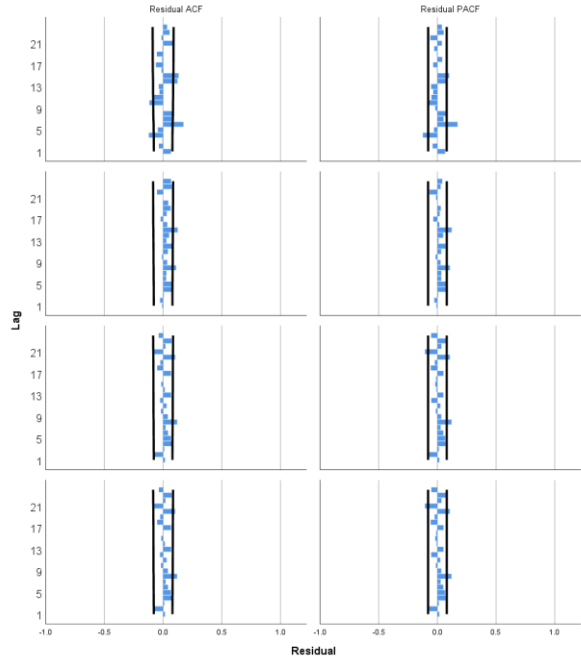


Şekil 3.10. Otokorelasyon



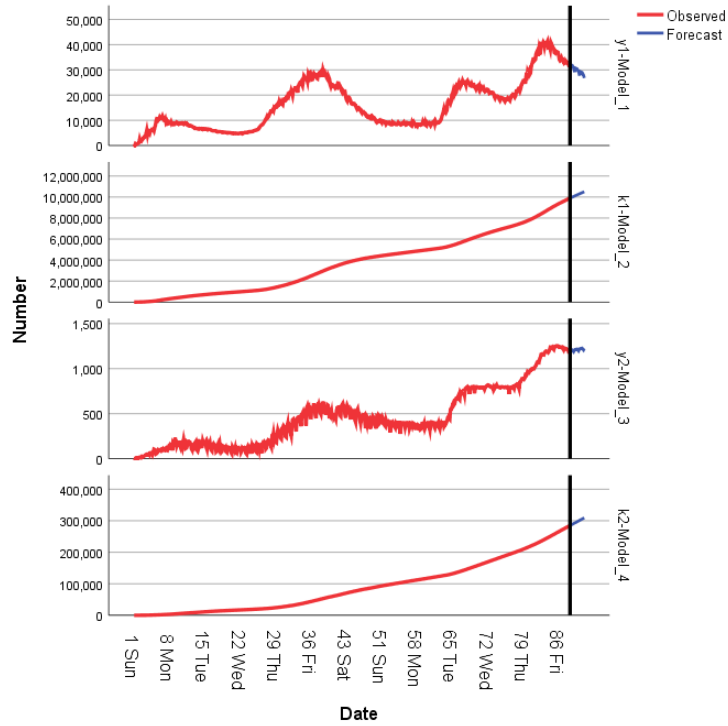
Şekil 3.11. Kısmi otokorelasyon

Model, bir serinin tahmin değerlerinin çizimini temsil eden sonuçlardan aktarılan Şekil 3.2'deki gibi tahmin için kullanılmıştır.



Şekil 3.12. Seri için tahmin değerleri

Haftalar arasındaki değişiklikleri fark edersek, aşağıdaki şekilde gösterilmiştir (Şekil 3.13).



Şekil 3.13. Bir serinin tahmin edilen değerleri ile orijinal değerleri arasındaki fark

Yukarıdaki Şekil 3.13 %95 güven seviyesi ile kırmızı renklerin serinin orijinal değerlerini, mavi renklerin ise tahmin edilen değerleri gösterdiğini göstermektedir.

Tablo 3.9. Gözlemlerin tahmin değerlerinin, üst ve alt sınırları

	Tahmin değerleri	Üst limit	Alt limit
89 Sun	1203	1251	1156
89 Mon	1206	1259	1152
89 Tue	1207	1266	1147
89 Wed	1209	1273	1144
89 Thu	1199	1268	1130
89 Fri	1188	1261	1114
89 Sat	1203	1280	1125
90 Sun	1213	1299	1128
90 Mon	1215	1306	1124
90 Tue	1216	1312	1121
90 Wed	1218	1319	1118
90 Thu	1209	1314	1104
90 Fri	1197	1307	1088
90 Sat	1213	1327	1099
91 Sun	1223	1345	1101
91 Mon	1225	1352	1098
91 Tue	1226	1358	1094
91 Wed	1228	1365	1091
91 Thu	1219	1361	1077
91 Fri	1207	1354	1060
91 Sat	1222	1374	1071

3.3.4. Suudi Arabistan İçin Model Belirleme ve Tahmin

Suudi Arabistan eyaletindeki en iyi görüş modelleri de erdemli kriterlere göre teşhis edildi ve dağılık modeller şu şekildeydi:

Yer işaretleri için modeller seçilseydi ve eklem kriterlerine göre en iyi model seçilmiş olsaydı ve bahsi geçen modeller aşağıdaki gibi olsaydı

Tablo 3.10. Önerilen Box - Jenkins modelleri

MODELAR	R-stat	R-SQARED	RMSE	MAPE	MAE	BIC
3,1,0	0.324	0.954	2.584	16.956	1.721	1.930

Modellerin farklılaşma kriterlerinin sonuçları üzerinden model (3,1,0) en iyi model olarak seçilmiştir çünkü bu model için standart değerler mümkün olduğunca düşük çıkmıştır.

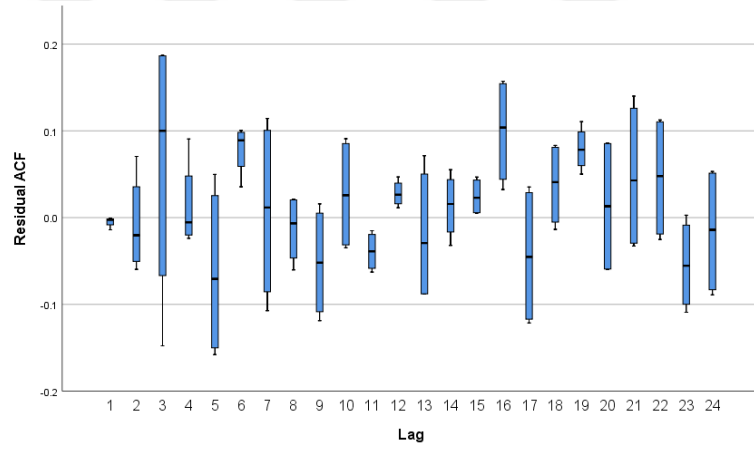
Modelin parametreleri aşağıdaki tabloda gösterildiği gibi tahmin edilebilir.

Tablo 3.11. Seçilen modelin parametreleri

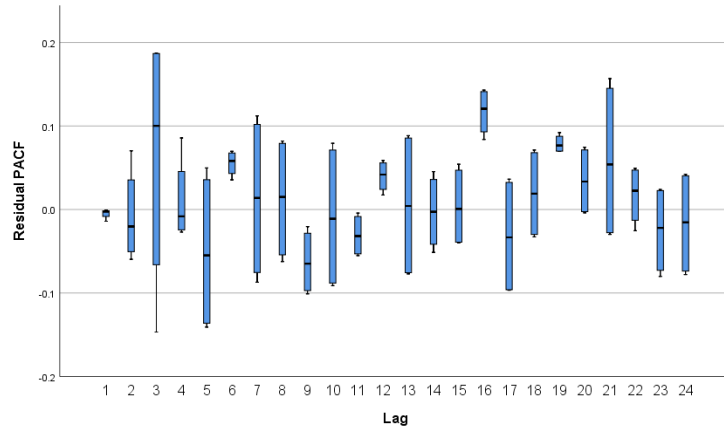
ESTI	S,E	t	sıg
-0,595	0.040	-14.86	0.000
-0,43	0.046	-3.078	0.002
-0,143	0.040	-3.560	0.000

Ve t'yi seçerek ve p değerine bağlı olarak, p değeri 0.05'ten küçük olduğu için 3,1,0 modelinin anlamlı olduğunu gözlenmektedir.

Önerilen modelde artıkların otokorelasyon ve kısmi otokorelasyon katsayıları %95 güven sınırları içerisinde, bu da artık serilerin rastgele değişkenleri temsil ettiği ve tahmin edici modelin Şekil 3.14 ve 3.15'deki gibi iyi olduğu anlamına gelmektedir.

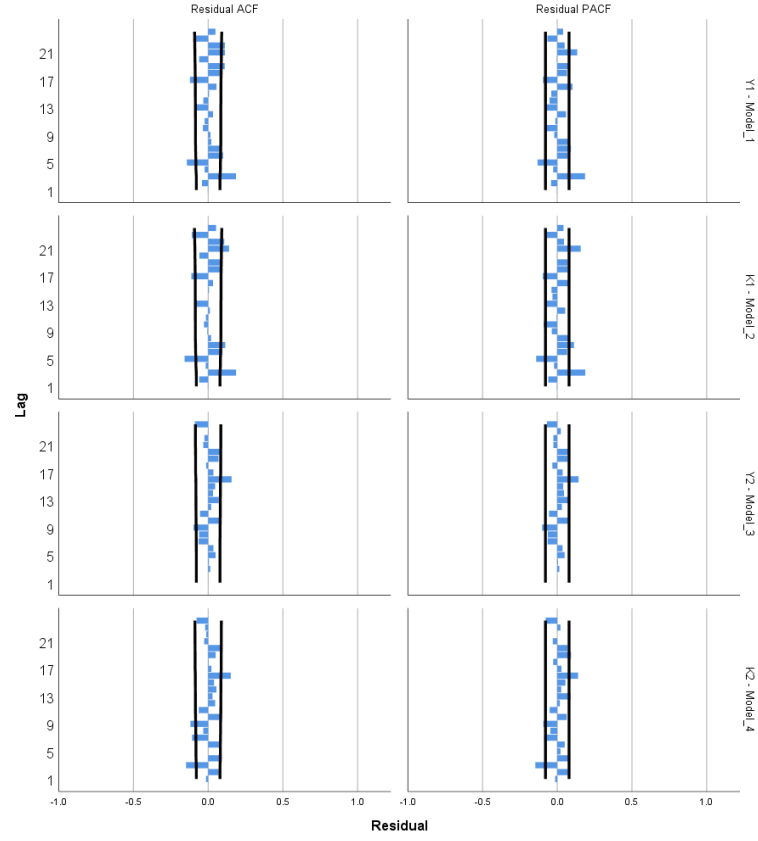


Şekil 3.14. Otokorelasyon



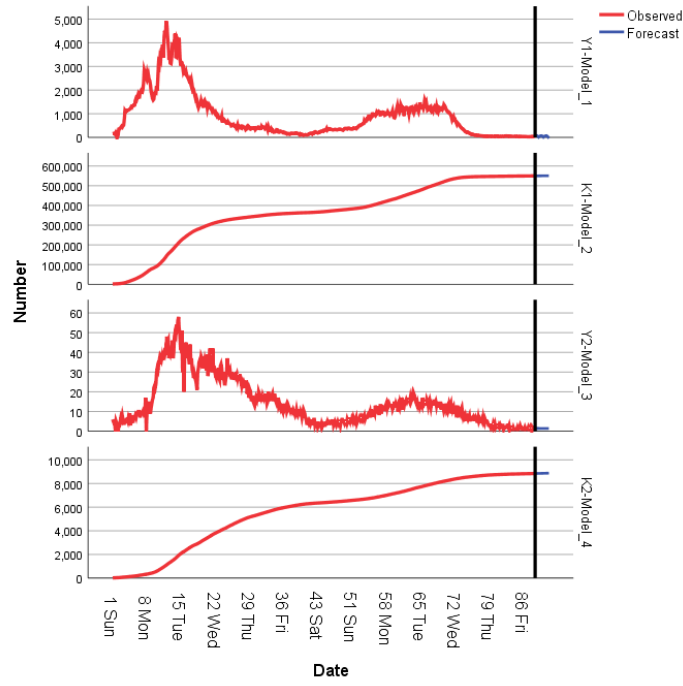
Şekil 3.15. Kısmi otokorelasyon

Şekil 3.16'daki gibi önerilen model kullanılarak tahmin edilmiştir.



Şekil 3.16. Seri için tahmin değerleri

Bir serinin tahmin değerlerinin çizilmesini temsil eden ve model (3,1,0) kullanılarak haftalar arasındaki değişiklikleri fark edersek sonuçlardan aktarılır ve aşağıdaki şekilde 3.17 gösterilmiştir.



Şekil 3.17. Bir serinin tahmin edilen değerleri ile orijinal değerleri arasındaki fark

Tablo 3.12. Gözlemlerin tahmin değerlerinin, üst ve alt sınırları

	Tahmin değerleri	Üst limit	Alt limit
89 Sun	1	7	-4
89 Mon	1	7	-4
89 Tue	1	8	-5
89 Wed	1	8	-5
89 Thu	1	9	-6
89 Fri	1	9	-6
89 Sat	1	10	-7
90 Sun	1	10	-7
90 Mon	1	10	-8
90 Tue	1	10	-8
90 Wed	1	10	-8
90 Thu	1	11	-9
90 Fri	1	11	-9
90 Sat	1	12	-10
91 Sun	1	13	-10
91 Mon	1	13	-10
91 Tue	1	13	-10
91 Wed	1	14	-11
91 Thu	1	14	-11
91 Fri	1	14	-11
91 Sat	1	14	-12

3.3.5. Türkiye İçin Model Belirleme ve Tahmin

Türkiye'deki en iyi veri modelleri teşhis edildi ve farklılaştırma kriterlerine göre en iyi modeli seçmek için seçilen modeller aşağıdaki tablodaki gibidir (13)

Tablo 3.13. Önerilen Box - Jenkins modelleri

MODELAR	R-stat	R-SQARED	RMSE	MAPE	MAE	BIC
0,1,6	0.358	0.989	1318.162	5.565	693.800	14.410
0,2,6	0.324	1.000	1317.32	0.066	709.868	14.409

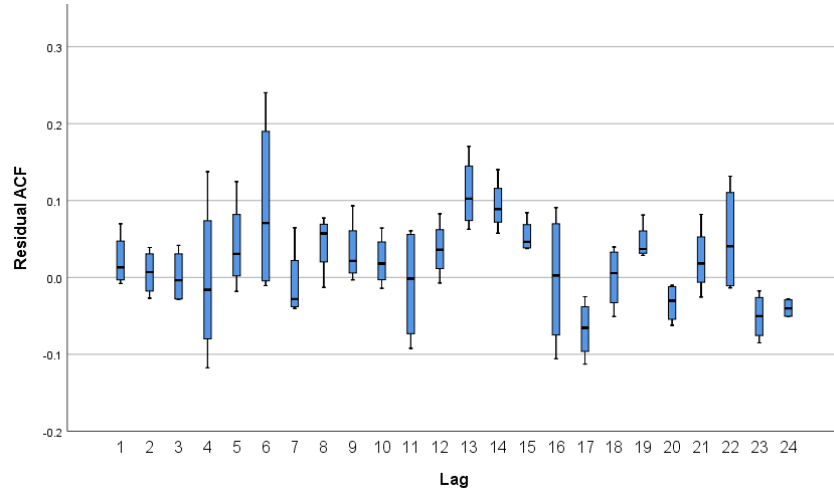
Sonuçlar aracılığıyla en iyi modelin (0,1,6) ARIMA olduğu bulunmuştur çünkü bu modelin standart değerleri olabildiğince az olup, bu nedenle aşağıdaki tablodaki gibi model parametrelerinin tahmin edilmesi mümkündür (Tablo 3.14).

Tablo 3.14. Seçilen modelin parametreleri

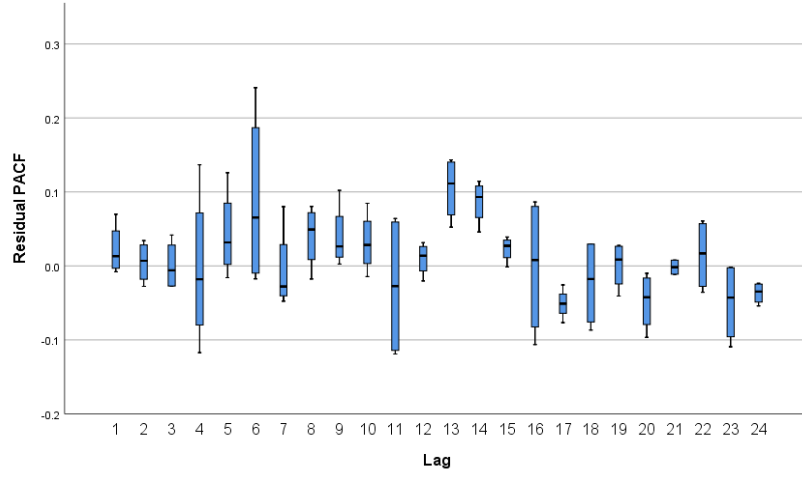
ESTI	S,E	t	sig
-0.136	0.040	-3.381	0.001
-0.225	0.040	-5.639	0.000
-0.322	0.46	-6.960	0.000
0.935	0.020	46.883	0.000

Ve p değerine göre t'yi seçerek anlamlı bir (0,1,6) model fark ederiz çünkü p değeri 0,05'ten küçüktür,

Önerilen modeldeki artıklar için kısmi otokorelasyon katsayılarının %95 güven sınırları içinde olması tahmin edilen modelin iyi olduğunu göstermekte ve aşağıdaki şekillerde 3.18 ve 3.19 sonuçlardan bir not aktarmaktadır.

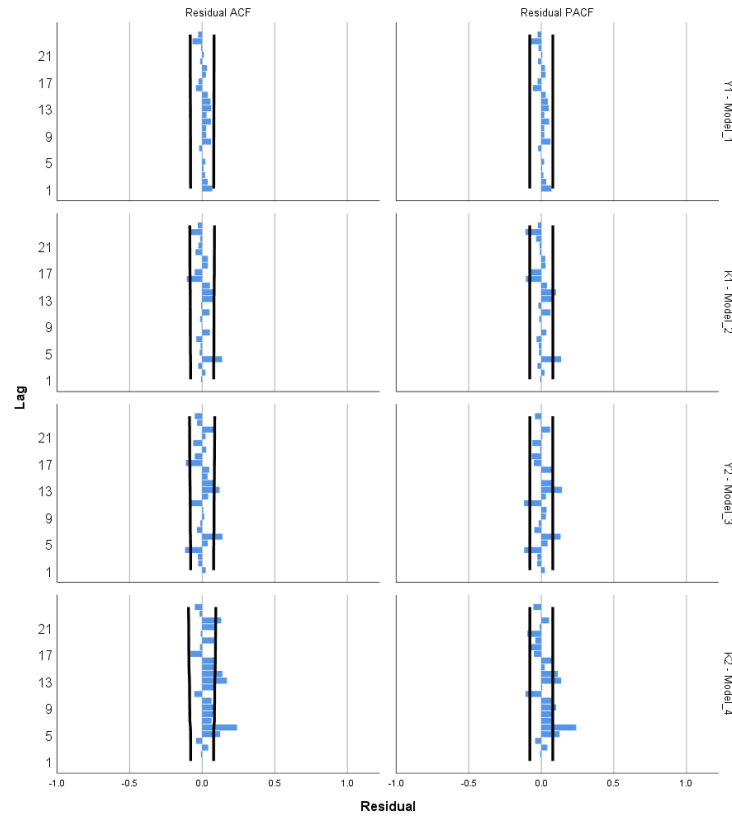


Şekil 3.18. Otokorelasyon



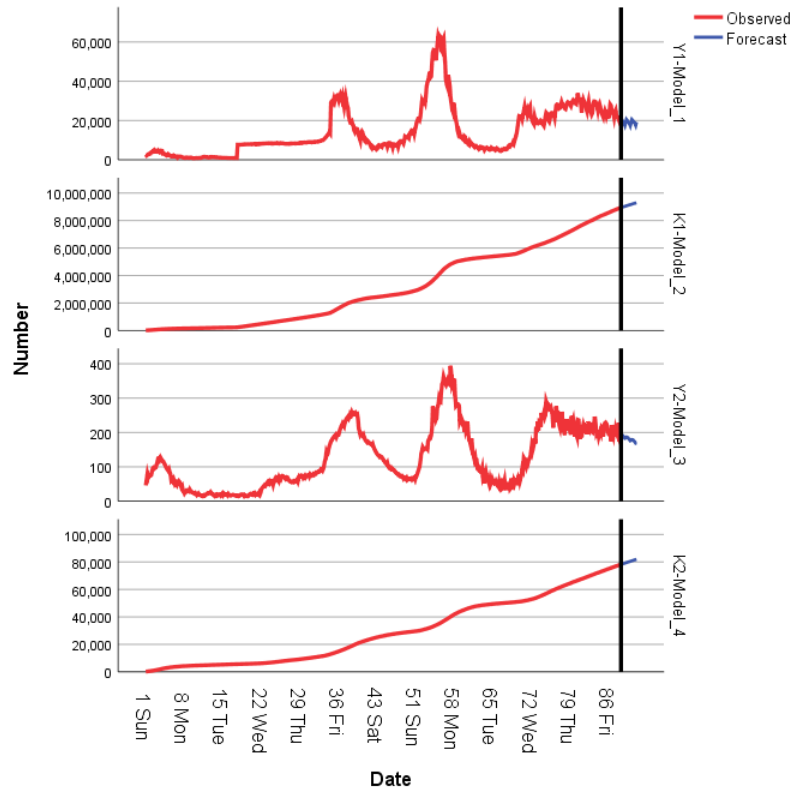
Şekil 3.19. Kısmi otokorelasyon

Modeli seçtikten sonra, Şekil 3.20 'deki gibi tahmin için seçenekler kullanıldı, şekil bir serinin tahmin değerlerinin çizimini temsil eden sonuçlardan aktarıldı.



Şekil 3.20. Seri için tahmin değerleri

Model (0.1.6) aracılığıyla haftalar arasındaki değişimler gözlemlenmekte ve aşağıdaki şekilde gösterilmektedir (Şekil 3.21)



Şekil 3.21. Bir serinin tahmin edilen değerleri ile orijinal değerleri arasındaki fark

Tablo 3.15. Gözlemlerin tahmin değerlerinin, üst ve alt sınırları

	Tahmin değerleri	Üst limit	Alt limit
89 Sun	21359	24235	18620
89 Mon	20572	24602	18915
89 Tue	19763	24639	15297
89 Wed	19355	24963	14295
89 Thu	18610	24932	13002
89 Fri	17473	24323	11502
89 Sat	18966	26909	12103
90 Sun	20580	29974	12565
90 Mon	20100	30413	11465
90 Tue	19286	30030	10520
90 Wed	19014	30847	9457
90 Thu	18319	30780	8444
90 Fri	17311	30219	7298
90 Sat	18959	33207	7927
91 Sun	20628	39349	8498
91 Mon	20215	36655	7740
91 Tue	19412	36366	6798
91 Wed	19174	36815	6249
91 Thu	18500	36609	5484
91 Fri	17527	35914	4615
91 Sat	19205	39096	5175

3.3.6. ABD İçin Model Belirleme ve Tahmin

ABD’de için en iyi özel veri modelinin teşhis ve tahmininde ise farklılaştırma kriterlerine göre en iyi model seçilmiş ve modeller aşağıdaki gibidir.

Tablo 3.16. Önerilen Box-Jenkins modelleri

MODELER	R-stat	R-SQARED	RMSE	MAPE	MAE	BIC
0,2,1	0.438	1.000	14674.467	0.068	8266.467	19.209
0,2,2	0.508	1.000	303.380	0.089	184.055	11.462

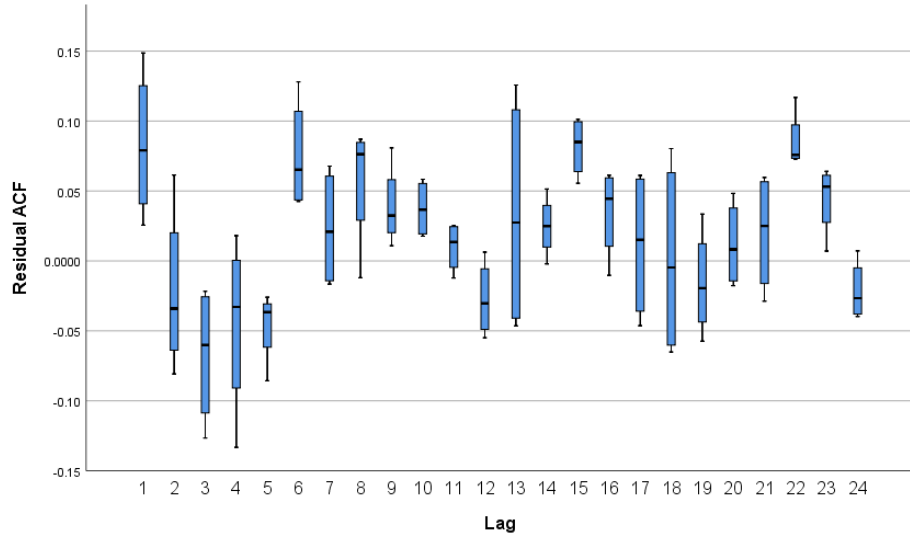
Farklılaştırma kriterlerinin sonuçları ve seçilen modellerin parametrelerinin önemi ile, en iyi modelin (0,2,2) ARIMA olduğu sonucuna varıyoruz, çünkü bu model için kriter değerleri mümkün olduğunca az ve model parametreleri aşağıdaki tabloda (3.17) tahmin edilmiştir.

Tablo 3.17. Seçilen modelin parametreleri

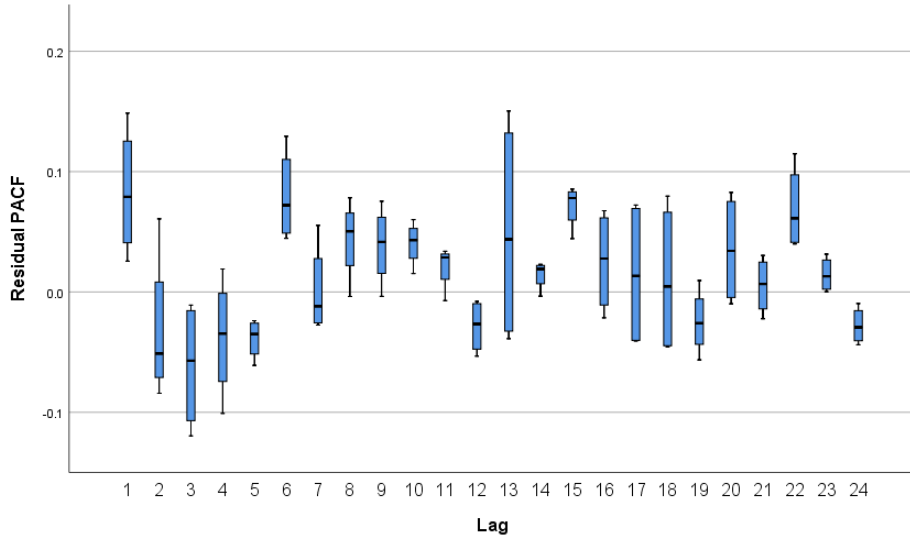
ESTI	S,E	t	sıg
0.529	0.041	13.057	0.00
0.201	0.041	4.943	0.00
0.712	0.030	23.502	0.00

p değerine göre t'yi seçerek, p değeri 0,05'ten küçük olduğu için anlamlı bir (0,2,2) model fark edilmektedir.

Önerilen modeldeki artıkların otokorelasyon ve kısmi katsayı değerlerinin %95 güven sınırları içinde olması, tahmin edilen modelin aşağıdaki şekillerde gösterildiği gibi olduğunu gösterir (Şekil 3.22 ve Şekil 3.23).

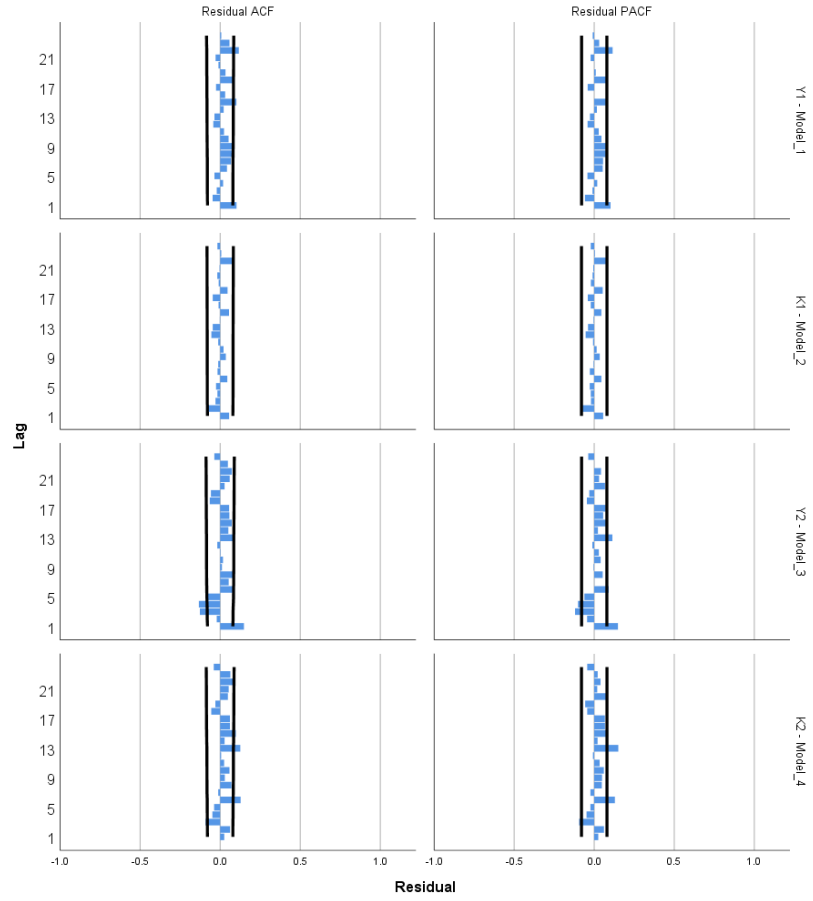


Şekil 3.22. Otokorelasyon



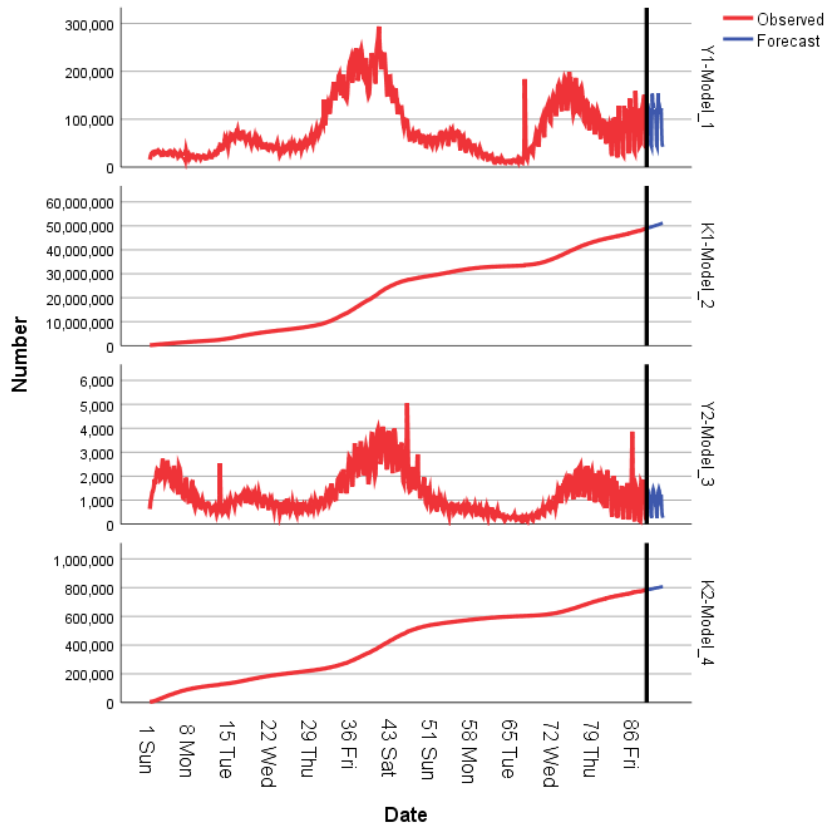
Şekil 3.23. Kısmi otokorelasyon

Model seçildikten sonra, Şekil 3.24'deki gibi tahmin için seçenekler kullanılmıştır.



Şekil 3.24. Seri için tahmin değerleri

Aşağıdaki şekilde gösterildiği gibi, zaman serisinin tahmin değerlerinin çizilmesini temsil eden ve aşağıdaki şekilde gösterildiği gibi haftalar arasındaki değişiklikleri fark ettiğimiz modeli (0,2,2) kullanarak



Şekil 3.25. Bir serinin tahmin edilen değerleri ile orijinal değerleri arasındaki fark

Tablo 3.18. Gözlemlerin tahmin değerlerinin, üst ve alt sınırları

	Tahmin değerleri	Üst limit	Alt limit
89 Sun	783344	783937	782752
89 Mon	784929	785982	783875
89 Tue	786487	787961	785012
89 Wed	787935	798831	786040
89 Thu	789446	791773	787118
89 Fri	789850	792625	787074
89 Sat	790060	793300	786821
90 Sun	791378	795186	787569
90 Mon	792955	797374	788536
90 Tue	794506	799558	789453
90 Wed	795947	801654	790240
90 Thu	797450	803833	791066
90 Fri	797847	804928	790765
90 Sat	798050	805851	790249
91 Sun	799360	807973	790748
91 Mon	800930	810399	791461
91 Tue	802473	812927	792120
91 Wed	803907	815173	792642
91 Thu	805403	817606	793200
91 Fri	805792	818959	792626
91 Sat	805988	820143	791834

3.3.7. Avustralya İçin Model Belirleme ve Tahmin

Avustralya ile ilgili olarak en iyi model teşhis edilmiş ve değerlendirilmiştir. Modeller farklılaştırma kriterlerine göre en iyi model olan nirengi noktaları için seçilmiştir. Seçilen modeller aşağıdaki tabloda gösterildiği gibidir (Tablo 3.19).

Tablo 3.19. Önerilen Box-Jenkins modelleri

MODELER	R-stat	R-SQARED	RMSE	MAPE	MAE	BIC
0,2,6	0.396	1.000	27.297	0.122	41.061	9.000
0,1,4	0.477	0.676	3.394	52.630	1.446	2.463
3,2,1	0.451	1.000	3.446	0.404	1.483	2.51

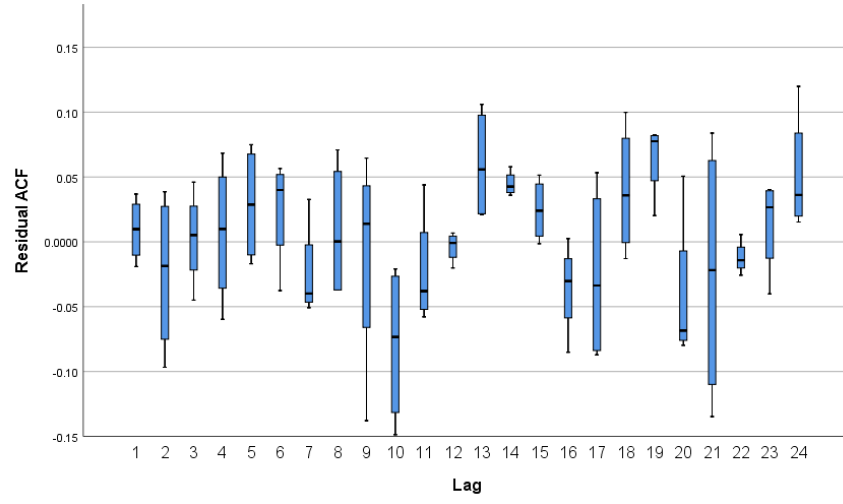
Elde edilen sonuçlardan en iyi modelin (3,2,1) olduğu sonucuna varılabilir çünkü bu model için kriter değerleri mümkün olduğunca azdır. Modele ait parametreler tablo (20)'deki gibi tahmin edilmiştir.

Tablo 3.20. Seçilen modelin parametreleri

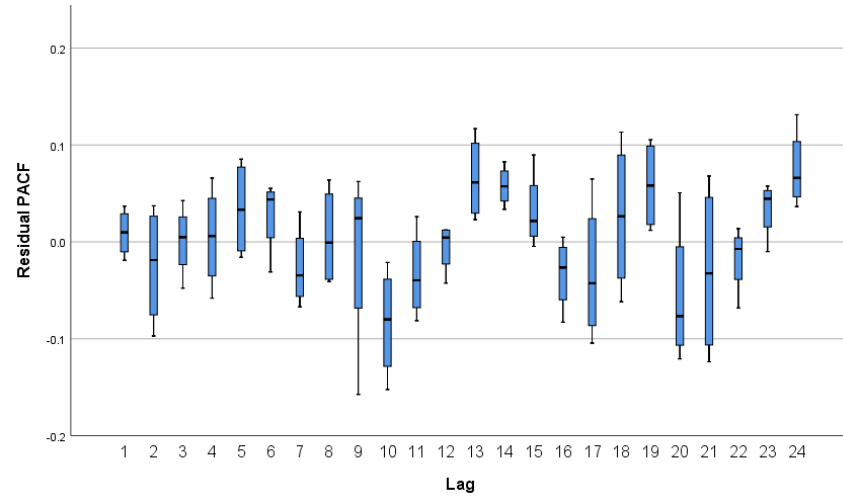
ESTI	S,E	t	sig
-0.386	0.068	-5.721	0.000
-0.357	0.056	-6.324	0.000
-0.308	0.049	-6.302	0.000
0.497	0.066	7.526	0.000

p değerine göre t'yi seçerek, p değeri 0,05'ten küçük olduğu için anlamlı bir (3,2,1) model fark ederiz.

Önerilen modelde artıkların otokorelasyon ve kısmi katsayı değerlerinin %95 güven sınırları içinde olması, aşağıdaki şekillerde gösterildiği gibi tahmin edilen modelin iyi olduğunu gösterir.

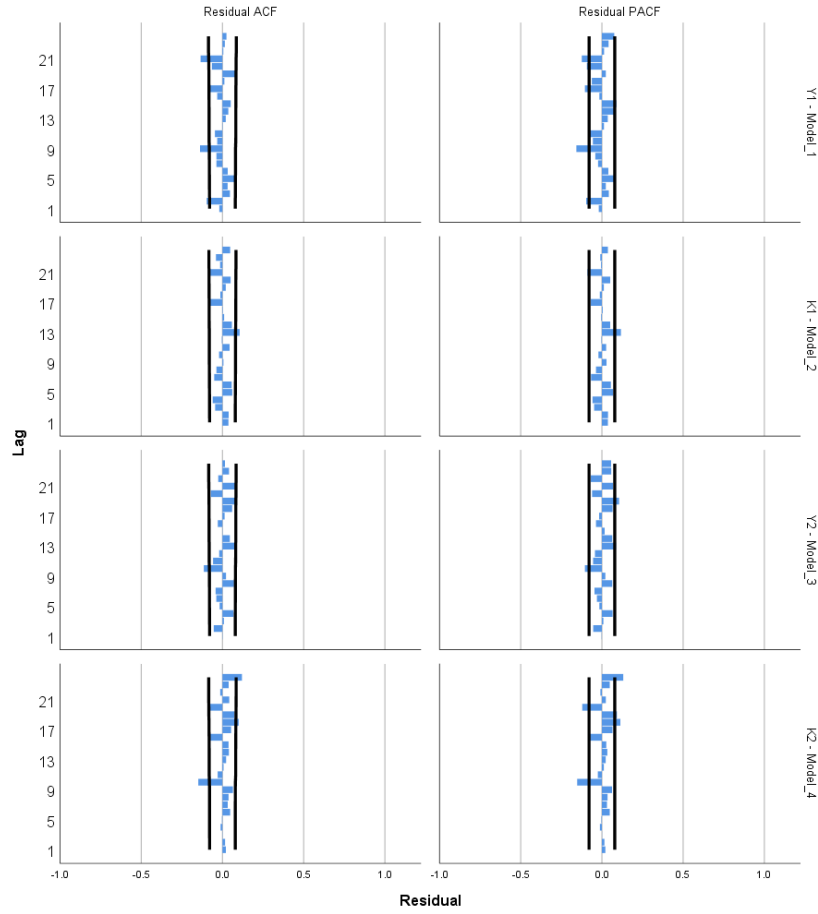


Şekil 3.26. Otokorelasyon



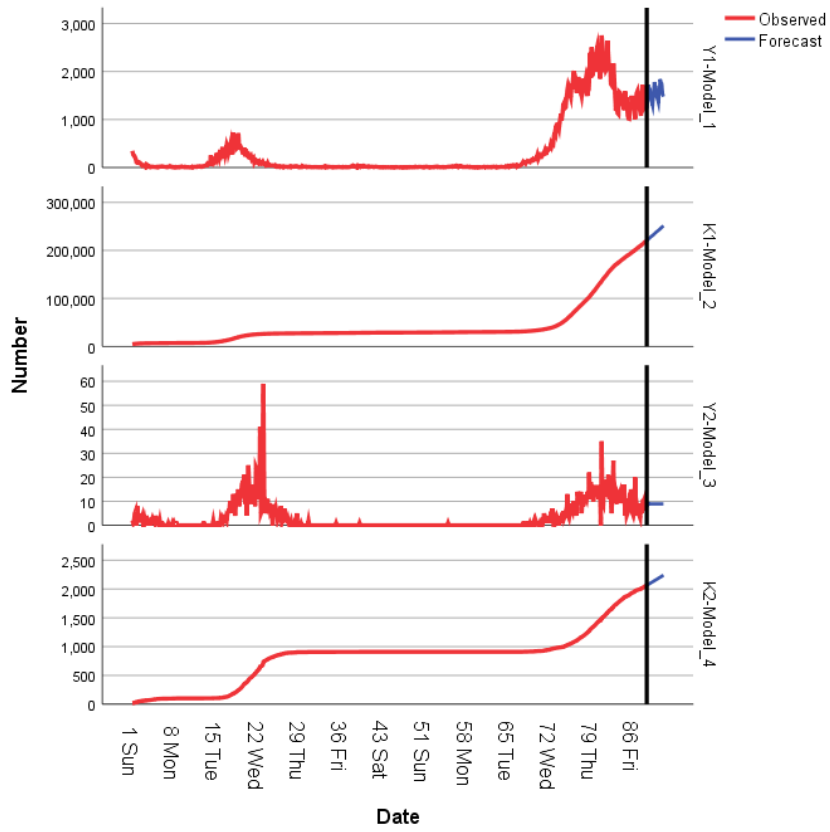
Şekil 3.27. Kısmi otokorelasyon

Model seçildikten sonra Şekil (28)'deki gibi tahmin için seçenekler kullanılmıştır.



Şekil 3.28. Seri için tahmin edilen değerler

Hangi zaman serisinin tahmin değerlerinin çizilmesini temsil eder ve yukarıdaki modeli kullanarak aşağıdaki şekilde gösterildiği gibi haftalar arasındaki değişiklikleri not ederiz.



Şekil 3.29. Bir serinin tahmin edilen değerleri ile orijinal değerleri arasındaki fark

Tablo 3.21. Gözlemlerin tahmin değerleri

	Tahmin değerleri	Üst limit	Alt limit
89 Sun	2065	2078	2052
89 Mon	2074	2093	2055
89 Tue	2083	2108	2058
89 Wed	2091	2121	2061
89 Thu	2099	2137	2062
89 Fri	2108	2154	2063
89 Sat	2117	2171	2063
90 Sun	2125	2188	2063
90 Mon	2134	2206	2063
90 Tue	2134	2225	2062
90 Wed	2151	2244	2061
90 Thu	2160	2263	2059
90 Fri	2169	2283	2057
90 Sat	2178	2303	2055
91 Sun	2187	2324	2052
91 Mon	2195	2345	2049
91 Tue	2204	2367	2046
91 Wed	2213	2389	2042
91 Thu	2222	2421	2038
91 Fri	2231	2435	2034
91 Sat	2240	2458	2029

3.3.8. Brezilya İçin Model Belirleme ve Tahmin

Brezilya'nın durumuna gelince, veriler için en iyi model teşhis edilmiş ve tahmin edilmiştir. Noktalar için modeller, farklılaştırma kriterlerine göre en iyi model seçilmiş ve seçilen modeller aşağıdaki tabloda gösterildiği gibidir.

Tablo 3.22. Önerilen Box Jenkins modelleri

MODELER	R-stat	R-SQARED	RMSE	MAPE	MAE	BIC
0,2,1	0.601	1.000	10740.626	0.227	6398.324	21.290
0,2,2	0.436	1.000	213.461	0.179	141.544	26.35

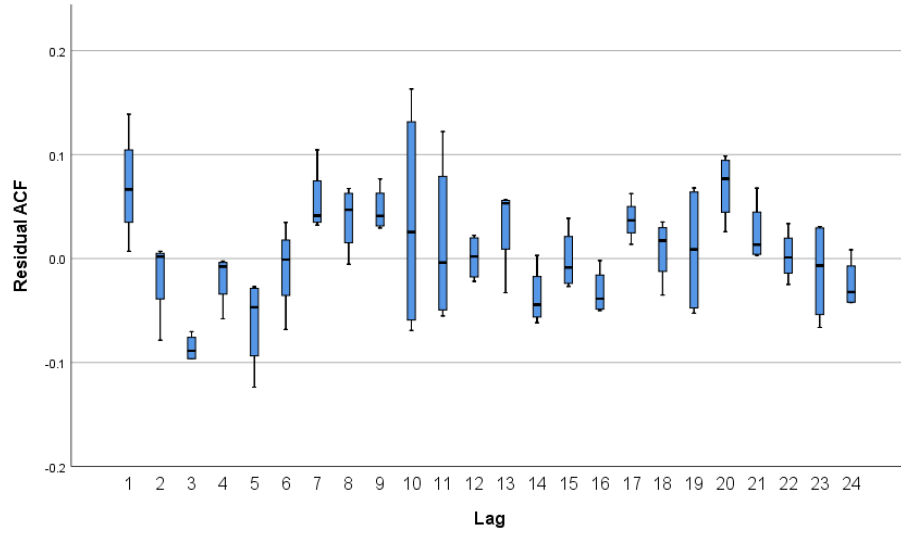
Karşılaştırma sonuçları ve seçilen modellerin parametrelerinin anlamlılığı sayesinde standart değerler en az olası olduğu için en iyi model (0,2,2) bulunmuştur. Model parametreleri buna göre tahmin edilmiştir (Tablo 3.23).

Tablo 3.23. Seçilen modelin parametreleri

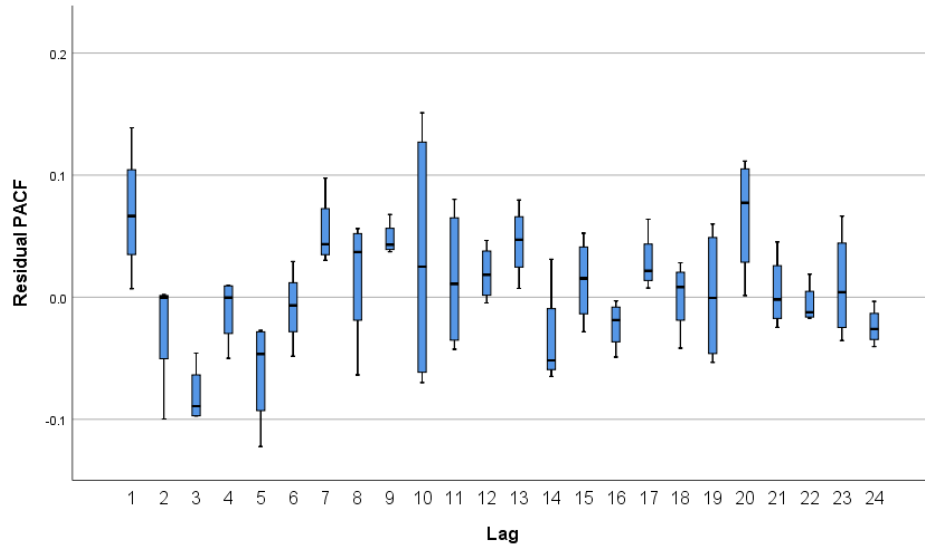
ESTI	S,E	t	sig
0.538	0.040	13.565	0.000
0.248	0.039	6.288	0.000
0.576	0.034	16.848	0.000

p değerine göre t'yi seçerek, p değeri 0,05'ten küçük olduğu için anlamlı bir (0,2,2) model fark ederiz.

Önerilen modelde artıkların otokorelasyon ve kısmi katsayı değerlerinin %95 güven sınırları içinde olması, aşağıdaki şekillerde gösterildiği gibi tahmin edilen modelin iyi olduğunu gösterir.

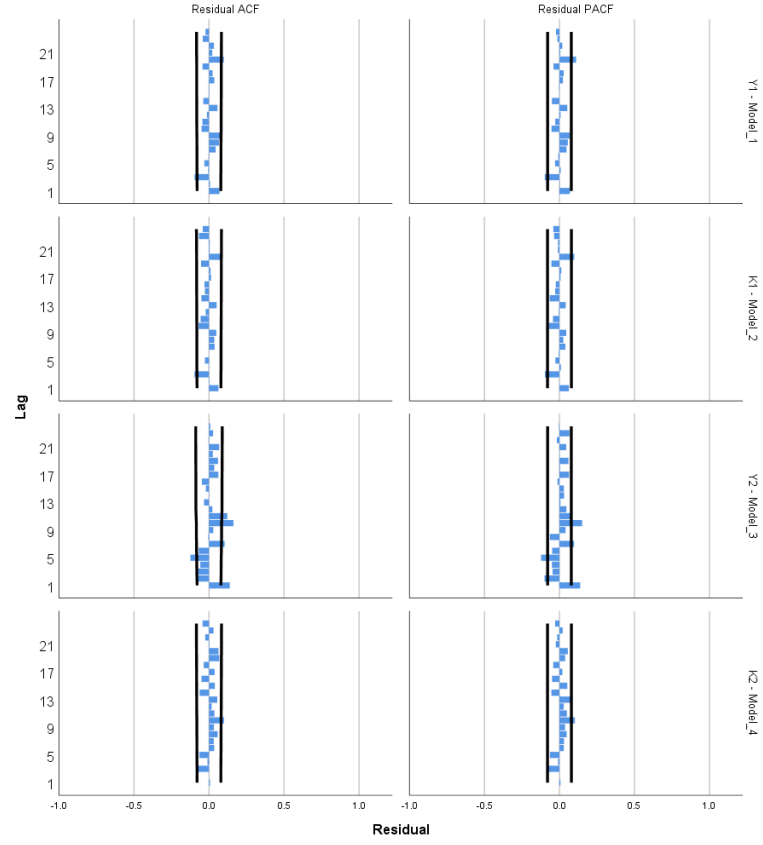


Şekil 3.30. Otokorelasyon



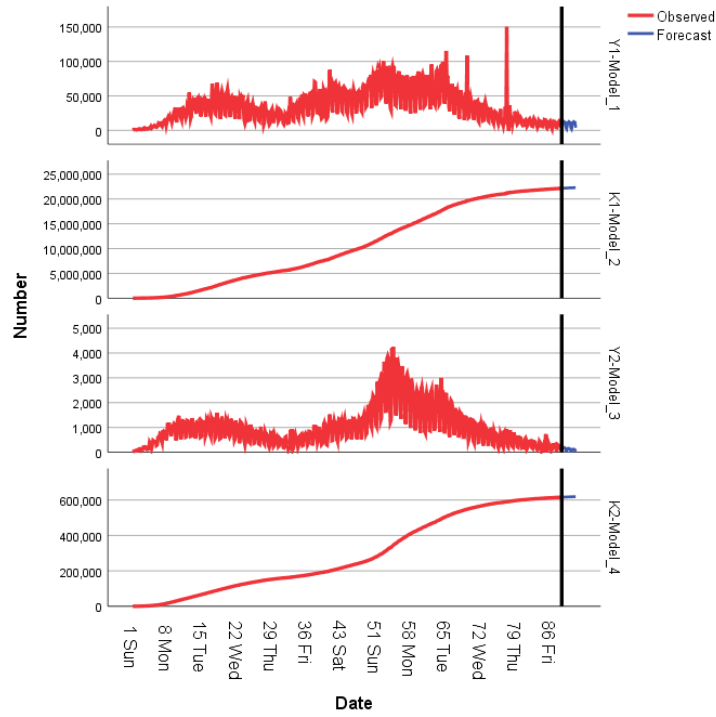
Şekil 3.31. Kısmi otokorelasyon

Model seçildikten sonra, Şekil 3.32'deki gibi tahmin için seçenekler kullanıldı. Şekil, bir serinin tahmin değerlerinin çizimini temsil eden sonuçlardan aktarılır.



Şekil 3.32. Seri için tahmin değerleri

Hangi zaman serisinin tahmin değerlerinin çizilmesini temsil eder ve yukarıdaki modeli kullanarak aşağıdaki şekilde gösterildiği gibi haftalar arasındaki değişiklikleri not ederiz.



Şekil 3.33. Bir serinin tahmin edilen değerleri ile orijinal değerleri arasındaki fark

Tablo 3.24. Gözlemlerin tahmin değerleri

	Tahmin değerleri	Üst limit	Alt limit
89 Sun	615709	616128	615290
89 Mon	615925	616668	615183
89 Tue	161171	617194	615149
89 Wed	161392	617686	615098
89 Thu	161614	618180	615048
89 Fri	161818	618661	614975
89 Sat	616846	618973	614720
90 Sun	616888	619396	614380
90 Mon	617084	620006	614161
90 Tue	617308	62654	163962
90 Wed	617507	621287	613728
90 Thu	61708	621931	613485
90 Fri	617891	622569	916213
90 Sat	617898	623042	612754
91 Sun	617918	623614	612223
91 Mon	618092	624377	611808
91 Tue	618295	625184	611407
91 Wed	618473	625982	610965
91 Thu	618652	626796	610509
91 Fri	618841	627608	610020
91 Sat	618800	628259	609341

3.3.9. Çin İçin Model Belirleme ve Tahmin

Çin’de en iyi veri modelini teşhis etmek mümkündür. Yer işaretleri için dört model seçilmiş ve kriterlere göre en iyi model seçilmiş ve modeller aşağıdaki gibidir.

Tablo 3.25. Önerilen Box-Jenkins modelleri

MODELER	R-stat	R-SQARED	RMSE	MAPE	MAE	BIC
0,1,6	0.196	0.833	9.628	42.151	17.900	7.390
0,2,6	0.196	1.000	39.662	0.018	17.922	7.392
0,0,0	2.2	2.243	52.113	163.036	6.014	7.917
0,1,0	1.6	0.989	52.155	0.124	6.015	7.919

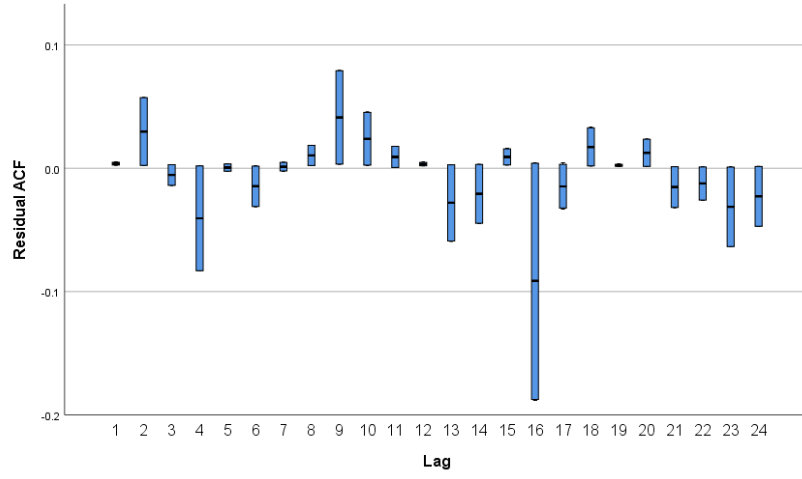
Farklılaştırma kriterlerinin sonuçları ve seçilen modellerin parametrelerinin önemi sayesinde, bu modelin standart değerleri mümkün olan en az olduğu için en iyi modelin (0,2,6) ARIMA olduğu sonucuna varılabilir. Modelin parametreleri aşağıdaki tabloda gösterildiği gibi tahmin edilmiştir.

Tablo 3.26. Seçilen modelin parametreleri

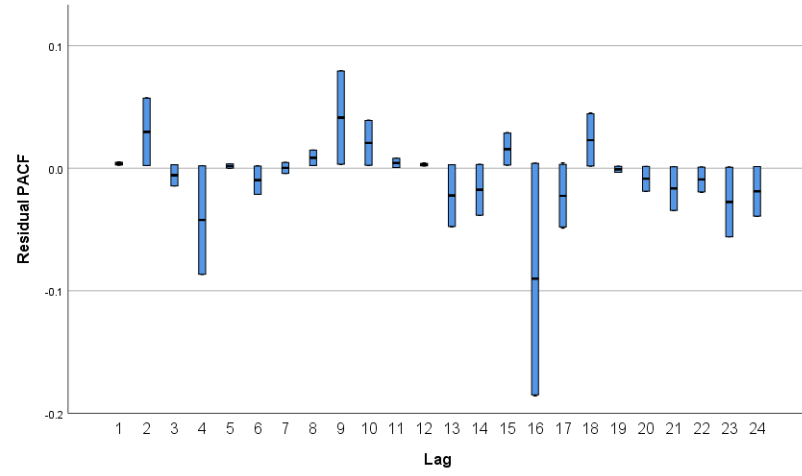
ESTI	S,E	t	Sıg
0.538	0.040	13.565	0.000
0.248	0.039	6.288	0.000
0.576	0.034	16.848	0.000

Ve p değerine dayalı olarak t'yi seçerek, p değeri 0,05'ten küçük olduğu için anlamlı bir (0,2,6) model fark ederiz.

Önerilen modelde artıkların otokorelasyon ve kısmi katsayı değerlerinin %95 güven sınırları içinde olması, aşağıdaki şekillerde gösterildiği gibi tahmin edilen modelin iyi olduğunu gösterir.

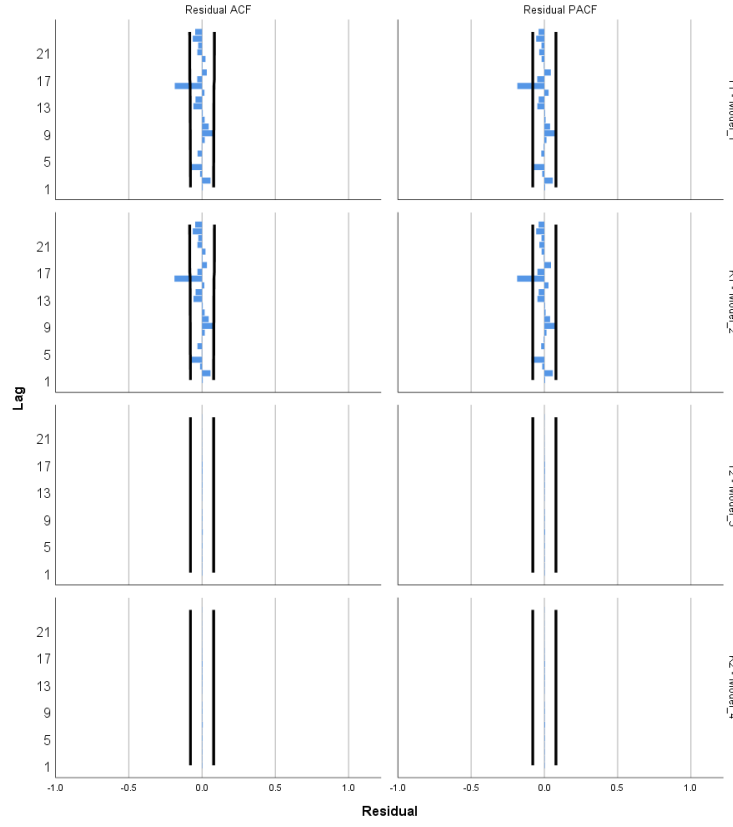


Şekil 3.34. Otokorelasyon



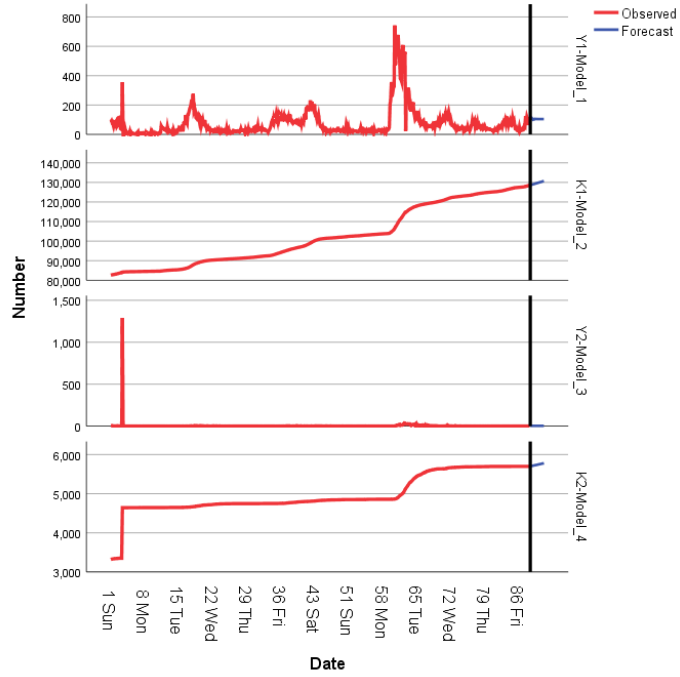
Şekil 3.35. Kısmi otokorelasyon

Model seçildikten sonra, Şekil 3.36'daki gibi tahmin için seçenekler kullanılmıştır.



Şekil 3.36. Seri için tahmin değerleri

Hangi zaman serisinin tahmin değerlerinin grafiğini temsil eder ve modelde not ederiz ((0,2,6 haftalar arasındaki değişiklikleri aşağıdaki şekilde olduğu gibi)



Şekil 3.37. Bir serinin tahmin edilen değerleri ile orijinal değerleri arasındaki fark

Tablo 3.27. Gözlemlerin tahmin değerlerinin, üst ve alt limiti

	Tahmin değerleri	Üst limit	Alt limit
89 Sun	128610	128688	128532
89 Mon	128720	128862	128557
89 Tue	128832	129046	125618
89 Wed	128938	129231	128644
89 Thu	129037	129418	128657
89 Fri	129139	129641	128665
89 Sat	129245	129827	128663
90 Sun	129350	130055	128644
90 Mon	129455	130295	128615
90 Tue	129560	130545	128575
90 Wed	129665	130804	128626
90 Thu	129769	131070	128468
90 Fri	129874	131345	128402
90 Sat	129979	131628	128329
91 Sun	130083	131919	128247
91 Mon	130188	132217	128159
91 Tue	130293	132522	128063
91 Wed	130397	132833	127961
91 Thu	130502	133151	127853
91 Fri	130606	133475	127738
91 Sat	130711	133805	127617

3.3.10. Fransa İçin Model Belirleme ve Tahmin

Fransa için en iyi modelin teşhisinin yanı sıra, nirengi noktaları için dört model seçilmiştir. Farklılaştırma kriterlerine göre en iyi model seçilmiştir. Model aşağıdaki gibidir.

Tablo 3.28. önerilen Box-Jenkins modelleri

MODELER	R-stat	R-SQARED	RMSE	MAPE	MAE	BIC
0,1,1	0.447	0.888	4403.82	33.916	.2255	16.812
0,2,6	0.419	1.000	4629.786	0.186	2489	16.923
Winter	0.623	0.790	111.025	53.389	34.45	9.451
0,1,6	0.136	1.000	108.688	0.135	48.49	9.430

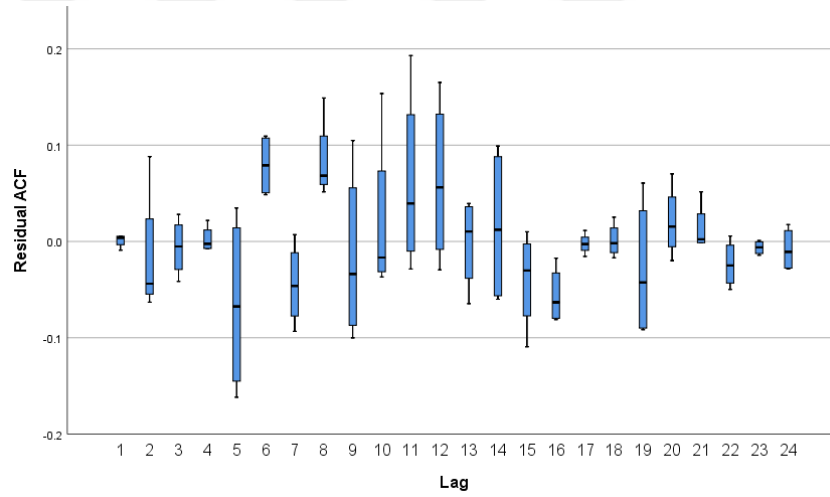
Seçilen modellerin erdemli kriterleri ve morali sonuçlarından en iyi modelin (0,1,6) olduğu sonucuna varılabilir çünkü bu modelin standart değerleri mümkün olduğunca az ve modelin parametreleri aşağıdaki tablodaki gibi tahmin edilmiştir.

Tablo 3.29. Seçilen modelin parametreleri

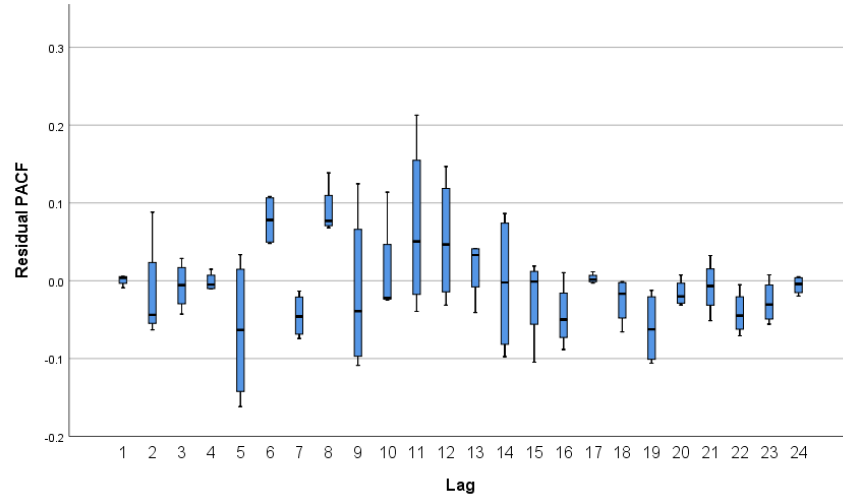
ESTI	S,E	t	sıg
-0.274	0.08	-7.309	0.000
-0.085	0.039	-2.164	0.031
-0.296	0.039	-7.500	0.000
-0.97	0.035	-27.750	0.000
-0.966	0.072	-13.758	0.000

p değerine dayalı t testinin sonuçlarıyla, p değeri 0,05'ten küçük olduğu için anlamlı bir (0,2.6) model fark ederiz.

Önerilen modelde artıkların otokorelasyon ve kısmi katsayı değerleri aşağıdaki şekillerde olduğu gibi %95 güven sınırları içinde olduğu için

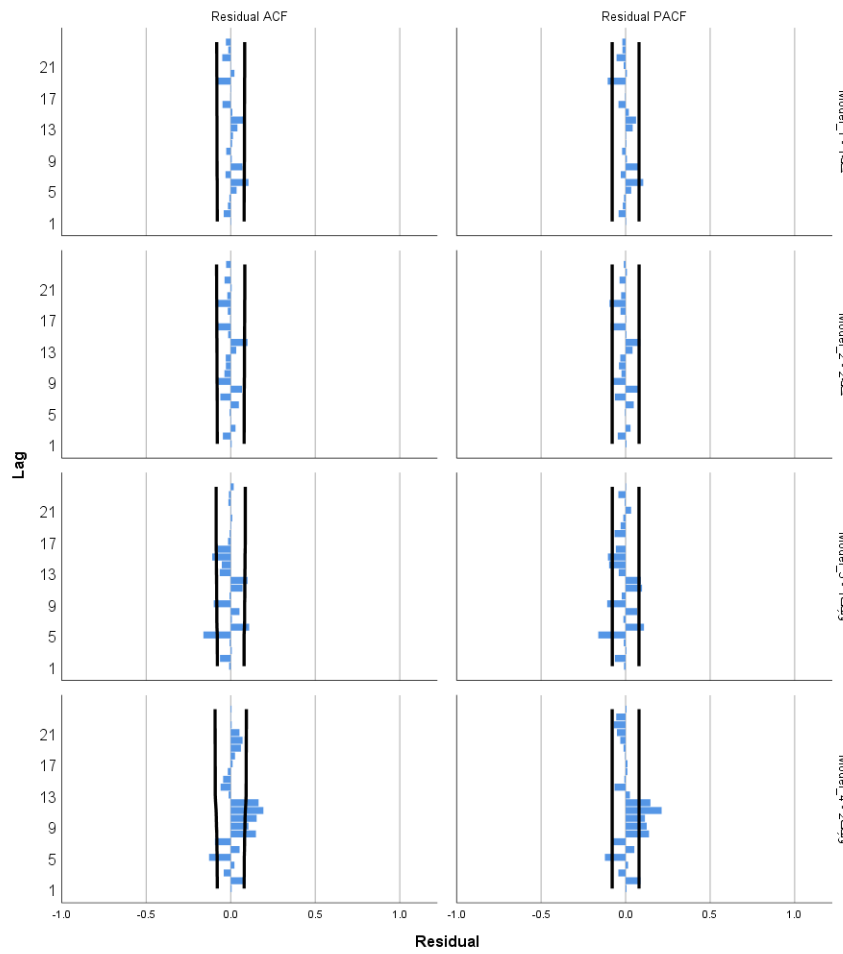


Şekil 3.38. Otokorelasyon



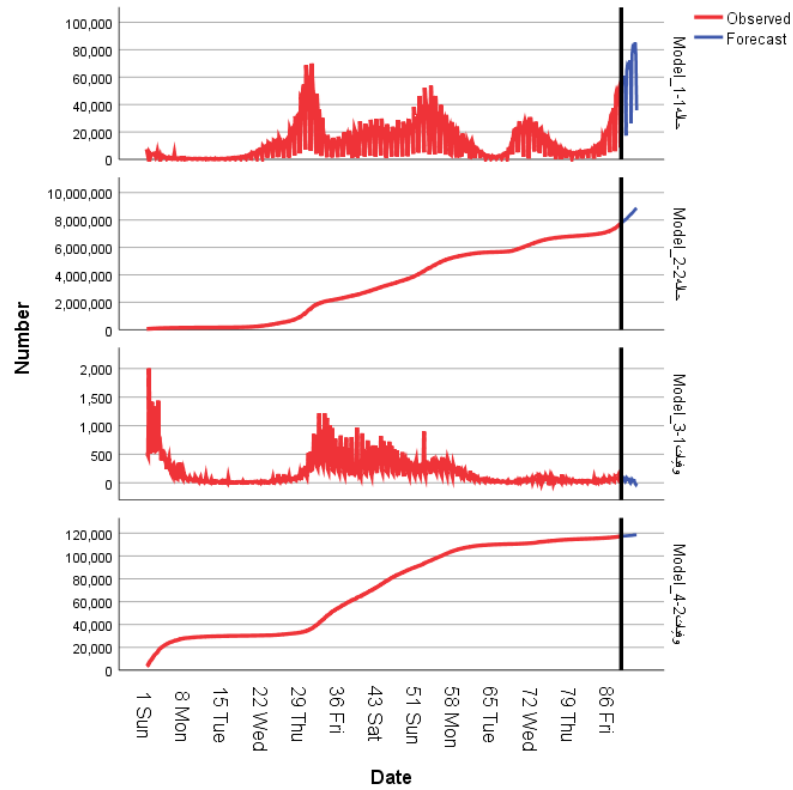
Şekil 3.39. Kısmi otokorelasyon

Model seçildikten sonra, Şekil 40'daki gibi tahmin için seçenekler kullanıldı. Şekil, bir serinin tahmin değerlerinin çizilmesini temsil eden sonuçlardan aktarılır.



Şekil 3.40. Seri için tahmin değerleri

Hangi zaman serisinin tahmin değerlerinin grafiğini temsil eder ve modelde not ederiz (0,2,6 haftalar arasındaki değişiklikleri aşağıdaki şekilde olduğu gibi)



Şekil 3.41. Bir serinin tahmin edilen değerleri ile orijinal değerleri arasındaki fark

Tablo 3.30. Gözlemlerin tahmin değerlerinin, üst ve alt limiti

	Tahmin değerleri	Üst limit	Alt limit
89 Sun	117170	117379	116962
89 Mon	117266	117561	116972
89 Tue	117344	117705	116983
89 Wed	117473	117921	117025
89 Thu	117520	118050	116989
89 Fri	117500	117101	116899
89 Sat	117502	118195	116810
90 Sun	117625	118514	116735
90 Mon	117725	118775	116675
90 Tue	117803	118992	116614
90 Wed	117921	119260	116582
90 Thu	117965	119446	116483
90 Fri	117951	119563	116339
90 Sat	117951	119707	116195
91 Sun	118073	120049	116097
91 Mon	118169	120343	115995
91 Tue	118246	120601	115891
91 Wed	118375	120919	115831
91 Thu	118422	1211448	115695
91 Fri	118402	121300	115505
91 Sat	117405	121484	115325

Testleri yaptıktan ve dört vakadan (yeni vakalar, kümülatif vakalar, yeni ölümler ve kümülatif ölümler) her ülke için en iyi modeli teşhis ettikten sonra, aşağıdaki tabloda gösterildiği gibi yaralanmaları bilmek için en iyi modelleri elde etmek mümkündür.



Tablo 3.31. Yeni vakalar

Numara	Ülkeler	Modeldescription/AR IMA	Stationary R_squared	R squared	RMSE	MSE	Exponential smoothing model parameters/ sig	ARIMA Model parameters /sig MA	ARIMA Model parameters /sig AR	ARIMA Model parameters /sig MA seasonal
1	Türkiye	0,1,6	0,358	0,989	1318,162	1737551,05	yok	0,013	0,000	0,000
2	Almanya	0,1,3	0,282	0,961	2458,937	6046371,169	yok	0,000	yok	0,000
3	İran	2,1,2	0,362	0,976	1407,834	1981996,57	yok	0,000	0,000	0,000
4	ABD	Winters multiplicative	0,455	0,942	14360,953	206236971,068	0,802	0,000	yok	0,000
5	Rusya	Winters Additive	0,484	0,966	594,083	352934,610	0,000	0,000	0,001	0,000
6	Avustralya	Winters additive	0,553	0,981	86,349	7456,149	0,000	0,009	0,000	0,000
7	Suudi Arabistan	Simple seasonal	0,467	0,981	130,676	17076,216	0,000	yok	yok	yok
8	Brezilya	Simple seasonal	0,602	0,814	10653,223	113491160,28	0,000	yok	yok	yok
9	Çin	0,1,6	0,196	0,833	39,628	1570,378	yok	0,000	yok	0,034
10	Fransa	0,1,1	0,447	0,888	4403,821	19393639,4	yok	0,000	0,000	0,000

Tablo 3.32. Kümülatif vakalar

Numara	Ülkeler	Modeldescription/ARIMA	Stationary R_squared	R squared	RMSE	MSE	Exponential smoothing model parameters/sig	ARIMA parameters MA	Model /sig	ARIMA Model parameters /sig AR	ARIMA Model parameters /sig MA seasonal
1	Türkiye	0,2,6	0,324	1,000	1317,324	1735342,52	yok	0,001		0,000	0,000
2	Almanya	0,2,3	0,283	1,000	2478,806	6144479,185	yok	0,000		yok	0,000
3	İran	1,2,1	0,219	1,000	1481,710	2195464,52	yok	0,000		0,029	0,000
4	ABD	0,2,1	0,438	1,000	14674,467	215339981,73	yok	yok		0,000	0,000
5	Rusya	3.2.6	0,478	1.000	598.660	358393.79	yok	0,000		0,001	0,000
6	Avustralya	0,2,6	0,396	1,000	87,247	7612,039	yok	0,008		0,000	0,000
7	Suudi Arabistan	holt	0,011	1,000	132,675	17602,65	0,000	yok		yok	yok
8	Brezilya	0,2,1	0,601	1,000	10740,626	115361046,87	yok	0,000		yok	0,000
9	Çin	0,2,6	0,196	1,000	39,662	1573,074	yok	0,000		yok	0,034
10	Fransa	0,2,6	0,419	1,000	4629,786	21434918,40	yok	0,000		0,000	0,000

Tablo 3.33. Yeni ölümler

Numara	Ülkeler	Modeldescription/ ARIMA	Stationary R_squared	R squared	RMSE	MSE	Exponential smoothing model parameters/sig	ARIMA Model parameters /sig MA	ARIMA Model parameters /sig AR	ARIMA Model parameters /sig MA seasonal
1	Türkiye	Winters Additive	0,562	0,983	11,803	139,310	0,935	yok	yok	yok
2	Almanya	0.0.6	0,354	0.931	62.271	3877.677	yok	0,001	yok	yok
3	İran	Winters Additive	0.570	0.970	25.263	638.219	0.885	yok	yok	yok
4	ABD	Winters Multiplicative	0,544	0,885	293,570	86183,34	0,848	yok	yok	yok
5	Rusya	0.1.1	0,497	0.994	24.355	593.166	yok	0,000	0,001	0,000
6	Avustralya	0.1.4	0,477	0.676	3.394	11.5192	yok	0,000	yok	yok
7	Suudi Arabistan	3,1,0	0,324	0,954	2,584	6,677	yok	yok	0,002	yok
8	Brezilya	Winters Multiplicative	0,520	0.939	195.501	38220.64	0.121	yok	yok	yok
9	Çin	0,0,0	2,243E-14	2,243E-14	52,113	2715,764	yok	yok	yok	yok
10	Fransa	Winters Additive	0,623	0.790	111.025	12326.550	0.001	yok	yok	yok

Tablo 3.34. Kümülatif ölümler

Numara	Ülkeler	Model description/ARIMA	Stationary R_squared	R_squared	RMSE	MSE	Exponential smoothing model parameters/sig	ARIMA Model parameters /sig MA	ARIMA Model parameters /sig AR	ARIMA Model parameters /sig MA seasonal
1	Türkiye	Holt	0,067	1,000	12,242	149,866	0,000	yok	yok	yok
2	Almanya	0,1,6	0,352	1,000	62,288	3879,79	yok	0,001	yok	yok
3	İran	0,2,6	0,311	1,000	25,919	671,794	0.885	0,002	0,000	0,000
4	ABD	0,2,2	0,508	1,000	303,380	92039,42	yok	0,000	yok	0,000
5	Rusya	0,2,1	0,496	1,000	24.374	594,091	yok	0,000	yok	0,000
6	Avustralya	3,2,1	0,451	1,000	3,446	11,874	yok	0,000	0,000	yok
7	Suudi Arabistan	Holt	0,302	1,000	2,621	6,869	0,000	yok	yok	yok
8	Brezilya	0,2,2	0,436	1,000	213,461	45565,59	yok	0,000	yok	0,000
9	Çin	0,1,0	1,643E-14	0,989	52,155	2720,144	yok	yok	yok	yok
10	Fransa	0,1,6	0,136	1,000	108,688	11813,08	yok	0,000	0,000	0,000

4. SONUÇ

Ülkelerdeki aktif COVID-19 vaka oranları ve dolayısıyla salgının ne zaman zirve yapacağını bilmek, hükümetlerin sosyal ve sağlık politikalarını değiştirmesine izin vereceği için pandeminin etkisini önemli ölçüde azaltma potansiyeline sahiptir. COVID-19 pandemisi, mevcut dünyamızda büyük bozulmalara neden oldu. Bu krizin etkisi ve büyüklüğü, dünya çapında çok fazla korku, belirsizlik ve endişeye neden oldu. Bu nedenle, gelecekte kayıpları önlemek için COVID-19 pandemisini ve toplumumuz üzerindeki etkilerini anlamak kritik öneme sahiptir.

Çalışmada; COVID-19 vakalarının kullanıldığı veriye dayalı bir yaklaşım, dört vaka (ölümler - kümülatif vakalar - yeni ölümler - kümülatif ölümler) ve on ülke için (Türkiye, Brezilya, Rusya, Avustralya, İran, Amerika, Çin) Fransa Almanya) zaman serisi veri tabanı ve tahmin modelleri geliştirmeye yönelik ARIMA istatistik modellerinin COVID-19'i daha iyi anlamak için iyi çalıştığı görülmektedir.

Bu çalışmada, COVID-19 mortalite ve vakalarının ARIMA modellerini oluşturmak için, p (otomatik regresyon model sırası), d (fark sırası), model parametrelerinin çeşitli önerilerinin bulunduğu ESACF (Genişletilmiş Örnek Otokorelasyon Fonksiyonu) tabloları kullanılmıştır. Bu analizi yapmak için ARIMA modelleri geliştirilmiş, zaman periyodu ve tahmin edilen değerler, ARIMA modellerinin performansını değerlendirmek için COVID-19 ölümlerinin ve vakalarının gerçek değerleriyle karşılaştırılmıştır.

Bu çalışma için Ocak 2020 ile 7 Aralık 2021 tarihleri arasında veriler kullanılmıştır. 12/7/2021'den itibaren haftalar doğrulanmış ve aktif COVID-19 vakaları Dünya Sağlık Örgütü'nün resmi web sitesinden elde edilmiştir

COVID-19 veri kümesini kullanarak ARIMA modelini seçme ve oluşturma süreci, COVID-19 veri kümesi üzerinde farklı araştırma çalışmalarında kullanılan diğer modeller arasında doğruluğu test etmeye, karşılaştırmaya ve optimize etmeye yardımcı olur. Önceki yayınlar uzun süredir esas olarak COVID-19 vakalarını ve ölümlerini tahmin etmeye odaklanmış olsa da, bu modellerin doğruluk oranları az miktarda veri kullanılması nedeniyle tatmin edici değildir. Ayrıca devam eden bir pandemi olduğu için birçok değişken COVID-19 ölüm ve vaka sayılarını değiştirmeye gelebilir. Bu nedenle, uzun vadeli COVID-19 ölümlerini ve vakaları yalnızca COVID-19 mortalitesi ve vaka değişkenlerine dayalı olarak tahmin etmek yeterli değildir. Bununla birlikte, bu

çalışmada, 1/4/2020 ila 7/12/2021 dönemi için COVID-19 ölümlerini ve vakalarını ihmal edilebilir hatalarla tahmin etmek için modeller geliştirmek için büyük miktarda veri kullanılmıştır.

Elde edilen sonuçlar, ARIMA modelinin COVID-19 mortalite ve vakalarında iyi performans gösterdiğini göstermektedir. Bu nedenle, bu tür bir model, kısa vadeli tahminler için herhangi bir ilde veya dünya çapında salgının ilerleyişi ile ilgili sınırlı zaman ve malzemeye ihtiyaç duyulan bir durumda faydalı olabilir. Öte yandan, bu kısa vadeli tahmin, seyahat acenteleri, küçük işletmeler, eğitim sektörleri, hastaneler vb. için kısa vadeli planlama yapmak için faydalı olabilir. Ancak bu tahmin modeli geliştirilirken yaş ve cinsiyet dikkate alındığında COVID-19 verilerini analiz etmek için daha sağlam ve kullanışlı olabilir. Ayrıca, bu virüsün tatillerde (hafta sonları dahil), partilerde ve diğer sosyal etkinliklerde yayılma olasılığı vardır. p , d ve q parametrelerinin diğer başlangıç değerleri, COVID-19 veri analizi ve modelleme yaklaşımları üzerine gelecekteki araştırmalar için düşünülebilir.

Sonuçlar, farklılaşma kriteri sonuçları arasında en iyi modelin ARIMA modeli olduğunu ve modeldeki artıklar için otokorelasyon katsayıları ve kısmi otokorelasyon katsayılarının %95 güven sınırları içinde olduğunu yani beklenen modelin iyi olduğunu göstermektedir.

Son olarak, modelin tahmin doğruluğunu geliştirmek için gelecekteki değişiklikler, diğer faktörleri hesaba katan çok değişkenli zaman serisi modellemenin benimsenmesinin yanı sıra, genel olarak azaltmak için dünyanın en iyilerini birleştirecek ARMA modelinin kullanımını içerir. Hata, salgının yayılmasıyla doğrudan veya dolaylı olarak ilişkilidir.

Bu çalışmanın sonuçları, sağlık sektöründeki politika yapıcılar, planlayıcılar ve yöneticiler için olduğu kadar, bu kanıtlara dayalı olarak gelecekteki modellemeyi tasarlamada epidemiyologlara yardımcı olabilir. Spesifik modeller yeni veriler geldikçe güncellenebilmekte, bu tahmin düzenli olarak Türkiye'de ve dünyada ilgili departmanlarda karar verme süreçlerinde doğru bilgi sağlamak için kullanılabilir. Gelecekteki araştırmalar için öneriler, zaman serisi modellemesini ve günlük yeni enfeksiyonların, ölümlerin ve iyileşmelerin tahminini içerir, ancak bunun önceden bilinmesi ve belirlenmesi kolay bir iş olmayabilir. Farklı modellere (ARIMA, makine öğrenmesi algoritması vb.) dayalı bir karşılaştırma, Türkiye'de COVID-19'u modellemek ve tahmin etmek için en iyi modeli sağlayabilir. Bu çalışmanın sonuçlarına dayanarak,

Türkiye Hükümeti'ne ülke genelinde sıkı önlemler alarak virüsün daha fazla yayılmasını kontrol altına alması ve günlük yeni enfeksiyonları/vakaları azaltması tavsiye edilmektedir.



KAYNAKLAR

- Adhikari, S. P., S. Meng, Y. J. Wu, Y. P. Mao, R. X. Ye, Q. Z. Wang et al. (2020). Epidemiology, causes, clinical manifestation and diagnosis, prevention and control of coronavirus disease (COVID-19) during the early outbreak period: a scoping review. *Infect Dis Poverty*. 9(1).29.
- Ahmadi, Ali et al. (2020). Modeling and forecasting trend of COVID-19 epidemic in Iran until May 13, 2020. *Medical journal of the Islamic Republic of Iran*. 34. 27.
- Aho, K., Derryberry, D. and Peterson, T. (2014). Model selection for ecologists: the Worldviews of AIC and BIC. *Ecology*, 95(3).631-636.
- Akgül, I., 2003. Zaman Serilerinin Analizi ve ARIMA Modelleri, Der Yayınları, İstanbul.
- Alabdulrazzaq, H., Alenezi, M. N., Rawajfih, Y., Alghannam, B. A., Al-Hassan, A. A., & Al-Anzi, F. S. (2021). On the accuracy of ARIMA based prediction of COVID-19 spread. *Results in physics*, 27, 104509.
- Aljandali A. (2017) The Box-Jenkins methodology. In: Multivariate methods and Forecasting with IBM® SPSS® Statistics. *Statistics and Econometrics for Finance*. Springer, Cham.
- Al-Omari, A., A. A. Rabaan, S. Salih, J. A. Al-Tawfiq, Z. A. Memish (2019). MERS coronavirus outbreak: implications for emerging viral infections. *Diagn Microbiol Infect Dis*. 93. 265–85.
- Amar, A. A. Taha and M. Y. Mohamed (2020). Prediction of the final size for Covid-19 epidemic using machine learning: A Case study of Egypt.
- Anderson, R. M., C. Fraser, A. C. Ghani, C. A. Donnelly, S. Riley, N. M. Ferguson et al. (2004). Epidemiology, transmission dynamics and control of SARS: the 2002-2003 epidemic. *Philos Trans R Soc Lond B, Biol Sci*. 359(1447).1091–105.
- Anderson, S. (2003). *Practical business forecasting*. Oxford, U.K: Blackwell Publishing,
- Ata, A.Y. ve Yücel, F. (2003). Eş-bütünleşme ve nedensellik testleri altında ikiz açıklar hipotezi: Türkiye uygulaması. *Çukurova Üniversitesi Sosyal Bilimler Enstitüsü Sosyal Bilimler Dergisi*, 12(12).97-110.
- Attanayake, A. M. C. H., & Perera, S. S. N. (2021). 2 Time Series Analysis for. *Recent Advances in Time Series Forecasting*, 11.
- Barriá-Sandoval, C., G. Ferreira, K. Benz-Parra, P. López-Flores (2021). Prediction of confirmed cases of and deaths caused by COVID-19 in Chile through time series techniques: A comparative study. *PloS one*. 16(4).
- Bayyurt, L. and B. Bayyurt (2020). *Forecasting of COVID-19 cases and deaths using ARIMA models*. Erişim: 12 Aralık, 2021.
- Bircan, H. ve Karagöz, Y. (2003). Box-jenkins modelleri ile aylık döviz kuru tahmini üzerine bir uygulama. *Kocaeli Üniversitesi Sosyal Bilimler Dergisi*, (6) , 49-62.
- Chu, J. (2021). A statistical analysis of the novel coronavirus (COVID-19) in Italy and Spain. *PLoS ONE*. 16(3). Erişim: 12 Aralık, 2021.
- Duru, Ö. (2007). *Zaman serileri analizinde ARIMA modelleri ve bir uygulama*. Yüksek Lisans Tezi. İstanbul Üniversitesi Sosyal Bilimler Enstitüsü, İstanbul.
- Erdoğan, E. (2006). *Zaman serilerinde ARIMA modelleri*. Yüksek Lisans Tezi. Muğla Üniversitesi Fen Bilimleri Enstitüsü, Muğla.

- Eric, (2019). *Introduction to the fundamentals of time series data and analysis*. Eriřim: 12 Aralık, 2021, <https://www.aptech.com/blog/introduction-to-the-fundamentals-of-time-series-data-and-analysis/>
- Fanalli, D. and F. Piazza (2020). Analysis and forecast of Covid-19 spreading in China, Italy and France. *Chaos, Solitons and Fractals*. 134. 109761.
- Gebretensae, Y. A. And D. Asmelash (2021). Trend analysis and forecasting the spread of COVID-19 Pandemic in Ethiopia using box-jenkins modeling procedure. *Int J Gen Med*. 14. 1485-1498. Eriřim: 15 Aralık, 2021.
- Gecili, E., A. Ziady, R. D. Szczesniak (2021). Forecasting COVID-19 confirmed cases, deaths and recoveries: Revisiting established time series modeling through novel applications for the USA and Italy. *Plos One*. 16(1). Eriřim: 15 Aralık, 2021.
- Goggin, D. E. (2007). *Method and apparatus for time series graph display: Google Patents*. Washington, DC: U.S. Patent and Trademark Office.
- Griffiths, W.E. (1993). *Learning and practicing econometrics*. New York: John Wiley and Sons Inc.
- Guerrier, S., R. Molinari, H. Xu, Y. Zhang (2019). *Applied Time Series Analysis with R*. Eriřim: 12 Aralık, 2021.
- Gujarati, D. N. and D. C. Porter (2009). *Basic econometrics*. 4th Edition. New York: McGraw-Hill Education.
- Guo, Y. R., Q. D. Cao, Z. S. Hong et al. (2020). The origin, transmission and clinical therapies on coronavirus disease 2019 (Covid-19) outbreak - an update on the status. *Mil Med Res*. 7(1). 11.
- Hanedar, A.Ö., Akkaya, O. ve Bizim, Ç., 2006. Duraęanlık Analizi, Birim Kk Testleri ve Trend. Zaman Serileri ve Uygulamaları Ders Sunu Planı. <http://debis.deu.edu.tr/userweb//onder.hanedar/dosyalar/Metin.pdf>. Eriřim: 12 Aralık, 2021.
- Karakař, E. (2019). Çocuk yoęun bakım nitesine olan talebin zaman serisi yntemleri ile tahmin edilmesi. *Avrupa Bilim ve Teknoloji Dergisi*, (17). 454-462.
- Kaya, E. (2019). Zaman serileri analizinde box-jenkins yntemi ile savunma sanayi verileri zerine bir uygulama. Yksek Lisans Tezi. Karamanoęlu Mehmetbey niversitesi Sosyal Bilimler Enstits, Karaman.
- Kırıl M. (2013). Box-Jenkins yntemi ile konut doęalgaz talebinin tahminlenmesi: İzmır ili rneęi. Yksek Lisans Tezi. Sleyman Demirel niversitesi Sosyal Bilimler Enstits, Isparta.
- Kohvakka, S. (2017). Forecasting Univariate Time Series-Comparison of Statistical Methods and Software Resources Available to Undergraduate Students. MS Thesis.
- Kumar, N. and H. Kumar (2021). A novel hybrid fuzzy time series model for prediction of Covid-19 infected cases and deaths in India. *ISA Trans*.
- Kumar, R., Jain, A., Tripathi, A.K. and Tyagi, S. (2021). COVID-19 Outbreak: An Epidemic Analysis using Time Series Prediction Model. *2021 11th International Conference on Cloud Computing, Data Science & Engineering (Confluence)*, pp. 1090-1094.
- Lee, C.M. and Ko, C.N., 2011. Short-term load forecasting using lifting scheme and ARIMA models. *Expert Systems with Applications*, 38(5).5902-5911.
- Li, Q., X. Guan, P. Wu, X. Wang, L. Zhou, Y. Tong et al. (2020). Early transmission dynamics in Wuhan, China, of novel coronavirus-infected pneumonia. *N Engl J Med*. 382. 1199-1207.

- Lin, R., A. King-Ip and H. S. S. K. Shim (1995, Eylül). "Fast similarity search in the presence of noise, scaling, and translation in time-series databases". *Proceeding of the 21th International Conference on Very Large Data Bases*, Citeseer. 490-501.
- Lipsitch, M., D. L. Swerdlow, L. Finelli (2020). Defining the Epidemiology of COVID-19 - Studies Needed. *N Engl J Med*. 382(13). 1194–1196.
- Liu, Y., L. M. Yan, L. Wan, T. X. Xiang, A. Le, J. M. Liu (2020). Viral dynamics in mild and severe cases of COVID-19. *Lancet Infect Dis*. 1473-3099.
- Ma, R., X. Zheng, P. Wang et al. (2021). The prediction and analysis of COVID-19 epidemic trend by combining LSTM and Markov method. *Scientific Reports*. 11(1). 1-14. Erişim: 12 Aralık, 2021.
- Mackay, I. M. and K. E. Arden (2015). MERS coronavirus: diagnostics, epidemiology and transmission. *Virol Journal*. 12(1). 1-21.
- Maurya, S. and S. Singh (2020). "Time Series Analysis of the COVID-19 Datasets". *2020 IEEE International Conference for Innovation in Technology (INOCON)*.
- Mech, A. (2017). Status of tea production in assam: Past trends and its future projections. *Journal of Open Learning and Research Communication*, (3). 45-56.
- Miyake, K., Miyake, N., Kondo, S., Tabe, Y., Ohsaka, A., & Miida, T. (2009). Seasonal variation in liver function tests: a time-series analysis of outpatient data. *Annals of Clinical Biochemistry*, 46(5). 377–384.
- Mobaraki, K. and J. Ahmadzadeh (2019). Current epidemiological status of Middle East respiratory syndrome coronavirus in the World From 112017 To 1712018: A Cross-Sectional Study. *BMC Infect dis*. 19(1). 1-5.
- Muhammad, M.K.I. (2012). Time series modeling using markov and ARIMA models. MS Thesis, Faculty of Hydraulic and Hydrology, University Technology, Malaysia.
- Özek, T. (2010). Zaman serisi modelleri üzerine bir simülasyon çalışması. Yüksek Lisans Tezi. Selçuk Üniversitesi Fen Bilimleri Enstitüsü, Konya.
- Peeri, N. C., N. Shrestha, M. S. Rahman, R. Zaki, Z. Tan, S. Bibi, et al. (2020). The SARS, MERS and novel coronavirus (COVID-19) epidemics, the newest and biggest global health threats: what lessons have we learned? *Int J Epidemiol*. 49(3). 717-726.
- Petropoulos, F. and S. Makridakis (2020). Forecasting the novel coronavirus COVID-19. *PLOS ONE*, 15(3). Erişim: 16 Aralık, 2021.
- Saba, A. I. and A. H. Elsheikh (2020). Forecasting the prevalence of COVID-19 outbreak in Egypt using nonlinear autoregressive artificial neural networks *Process Saf. Environ*. 141. 1-8.
- SAS, 2014. SAS 13.2 User's Guide The ARIMA Procedure. SAS Institute Inc., Cary, NC, USA. <https://support.sas.com/documentation/onlinedoc/ets/132/ARIMA.pdf>. Erişim Tarihi: 24.05.2019.
- Scott, G. (2021). Box-Jenkins Model. <https://www.investopedia.com/terms/b/box-jenkins-model.asp>. Erişim: 12 Aralık, 2021.
- Siedner, M. J., G. Harling, Z. Reynolds, R. F. Gilbert, A. Venkataramani, and A. C. Tsai (2020). Social distancing to slow the U.S. COVID-19 epidemic: an interrupted time-series analysis. Erişim: 12 Aralık, 2021.
- Singh, S., C. Chowdhury, A. K. Panja et al. (2021). Time Series Analysis of COVID-19 Data to Study the Effect of Lockdown and Unlock in India. *Journal of The Institution of Engineers (India): Series B*. 102, 1275–1281. Erişim: 15 Aralık, 2021.

- Soy Temür, A. & Yıldız, Ş. (2021). Comparison of forecasting performance of ARIMA LSTM and HYBRID models for the sales volume budget of a manufacturing enterprise. *Istanbul Business Research*, 50 (1). 15-46
- Soy Temür, A., Akgün, M and Temür, G. (2019). Predicting housing sales in Turkey using ARIMA, LSTM and hybrid models. *Journal of Business Economics and Management*, 20(5). 920-938.
- Tandon, H., P. Ranjan, T. Chakraborty, V. Suhag (2020). *Coronavirus (COVID-19): ARIMA based time-series analysis to forecast near future*. arXiv:2004.07859
- TC. Sağlık Bakanlığı (2021). Erişim: 13 Aralık, 2021.
- Tetali, S., G. R. Jammy, E. S. Asirvatham, B. R. Kumar and L. P. Choudhury (2021). An Interrupted Time Series Analysis of COVID-19 Positivity before, during and after Lockdown in Four States of India. *Open Journal of Epidemiology*, 11(1). 47-55. Erişim: 14 Aralık, 2021.
- Thompson, R. N., T. D. Hollingsworth, V. Isham, et al. (2020). Key questions for modelling COVID-19 exit strategies. *Proceedings of the Royal Society B*. 287(1932).
- Time Series Forecasting: Definition, Applications, and Examples. <https://www.tableau.com/learn/articles/time-series-forecasting> Erişim: 12 Aralık, 2021.
- Ucal, M.Ş. (2006). Ekonometrik model seçim kriterleri üzerine kısa bir inceleme. *Cumhuriyet Üniversitesi İktisadi ve İdari Bilimler Dergisi*, 7(2). 41-57.
- Verbesselt, J., Hyndman, R., Zeileis, A., & Culvenor, D. (2010). Phenological change detection while accounting for abrupt and gradual trends in satellite image time series. *Remote Sensing of Environment*, 114(12). 2970-2980.
- Viswa, C. Chandu (2021). Department of Public Health Dentistry, SIBAR Institute of Dental Sciences, Takkellapadu, Guntur, Andhra Pradesh, India. Erişim: 15 Aralık, 2021.
- Wang, C., L. Liu, X. Hao, H. Guo, Q. Wang, J. Huang et al. (2020). Evolving epidemiology and impact of non-pharmaceutical interventions on the outbreak of Coronavirus disease 2019 in Wuhan, China. *Journal of American Medical Association*. 323. 1915–1923.
- Wang, Y., Y. Wang, Y. Chen, Q. Qin (2020). Unique epidemiological and clinical features of the emerging 2019 novel coronavirus pneumonia (COVID-19) implicate special control measures. *J Med Virol*. 92(6). 568-576.
- WHO, (2019). MERS-CoV. Erişim: 12 Aralık, 2021.
- Worldometer (2021). Covid-19 Coronavirus Outbreak. Erişim; 13 Aralık, 2021.
- Yasue, S., Munakata, K., Kato, M. and Mori, S. (1996). Calculation of the power law index for a time series by means of its fractal dimension. *Fractals*, 4(03). 265-271.
- Yeşilyayla, H. (2013). X-12 ARIMA metoduyla sosyo-ekonomik verilerin analizi. Yüksek Lisans Tezi. Pamukkale Üniversitesi Fen Bilimleri Enstitüsü, Denizli.
- Yüksel, D. (2015). Arap Baharı'ndan etkilenen yakın ve orta doğu ülkeleri ile Türkiye arasındaki ithalat ve ihracat miktarlarının ARIMA modelleri ile incelenmesi. Yüksek Lisans Tezi Dokuz Eylül Üniversitesi. Sosyal Bilimler Enstitüsü, İzmir.
- Zhao, B., Z. Wang, Z. Yu, C. Tian, J. Cao (2020). Time series analysis of the novel coronavirus (COVID-19). *J Immuno Allerg*. 1(3). 1-13.
- Zhou, G. and G. Yan (2003). Severe acute respiratory syndrome epidemic in Asia. *Emerging Infect Dis*. 9(12). 1608–1610.

ÖZ GEÇMİŞ

Mohammed Zakı Noori Lise öğrenimini Kerkük’te tamamladıktan sonra Kerkük Üniversitesi İktisadi ve Fen Bilimler Fakültesi’nden 2019 yılında mezun oldu. 2020 yılında OMÜ LEE İstatistik Yüksek Lisans programına başladı. 2022 yılında OMÜ, LEE, İstatistik Yüksek Lisans programını bitirdi. İyi derecede Arapça bilmektedir. (14.01.2022).

İletişim Bilgileri

Öğrenci no :19211885

ORCID ID : <https://orcid.org/0000-0002-7389-3669>

Yayınlar:

1. Noori, M.Z, Al-Khaiat, SH.S.J, CENGİZ, M.A. (2021). Application of the Regression Analysis in Python, Spss and Microsoft Excell Programs. 4. *International Conferannse on Pure Sciences and Agrivultural Sciences (ICPAS-2021)*. 26-29 May 2021, Antalya