T.C.
ÇUKUROVA ÜNİVERSİTESİ
SAĞLIK BİLİMLERİ ENSTİTÜSÜ
BİYOİSTATİSTİK ANABİLİM DALI

# SALGINLARDAKİ VAKA TAHMİNLERİNDE KULLANABİLECEK MODELLERİN KARŞILAŞTIRILMASI COVİD-19 VAKA SAYILARI TAHMİNİ

Zeynep BAHAR KILINÇ

BİYOİSTATİSTİK TEZLİ YÜKSEK LİSANS PROGRAMI
YÜKSEK LİSANS TEZİ

DANIŞMANI
Doç. Dr. İlker ÜNAL

ADANA-2024

T.C.
ÇUKUROVA ÜNİVERSİTESİ
SAĞLIK BİLİMLERİ ENSTİTÜSÜ
BİYOİSTATİSTİK ANABİLİM DALI

# SALGINLARDAKİ VAKA TAHMİNLERİNDE KULLANABİLECEK MODELLERİN KARŞILAŞTIRILMASI COVİD-19 VAKA SAYILARI TAHMİNİ

Zeynep BAHAR KILINÇ

BİYOİSTATİSTİK TEZLİ YÜKSEK LİSANS PROGRAMI
YÜKSEK LİSANS TEZİ

DANIŞMANI
Doç. Dr. İlker ÜNAL

ADANA-2024

## KABUL VE ONAY

**Biyoistatistik** Anabilim Dalı

**Tezli** Yüksek Lisans Programı Çerçevesinde yürütülmüş olan

"Salgınlardaki Vaka Tahminlerinde Kullanılacak Modellerin Karşılaştırılması COVİD-19 Vaka Sayıları Tahmini " adlı çalışma, aşağıdaki jüri tarafından **Yüksek Lisans Tezi** olarak kabul edilmiştir.

**Tarihi:** 28/ 06 / 2024

**TEZ SINAV JÜRİSİ**

Doç. Dr. İlker ÜNAL
Çukurova Üniversitesi
**Başkan**

Prof. Dr. Gülşah SEYDAOĞLU
Çukurova Üniversitesi
**Üye**

Prof. Dr. Gülhan TEMEL
Mersin Üniversitesi
**Üye**

**Üye**

**Üye**

Yukarıdaki Tez, Yönetim Kurulunun / / tarih ve sayılı kararı ile kabul edilmiştir.

Prof.Dr. Behice DURGUN
**Sağlık Bilimleri Enstitü Müdürü**

# T.C. ÇUKUROVA ÜNİVERSİTESİ SAĞLIK BİLİMLERİ ENSTİTÜSÜ DOKTORA TEZ ÇALIŞMASI

## ETİK BEYANI

Çukurova Üniversitesi Bilimsel Araştırma ve Yayın Etiği Yönergesini okuduğumu ve anladığımı ve Çukurova Üniversitesi Sağlık Bilimleri Enstitüsü Tez Yazım Kurallarına uygun olarak hazırladığım bu tez çalışmasında;

- Tez içinde sunduğum verileri, bilgileri ve dokümanları akademik ve etik kurallar çerçevesinde elde ettiğimi,
- Tüm bilgi, belge, değerlendirme ve sonuçları bilimsel etik kurallarına uygun olarak sunduğumu,
- Tez çalışmasında yararlandığım eserlerin tümüne uygun atıfta bulunarak kaynak gösterdiğimi,
- Kullanılan verilerde ve ortaya çıkan sonuçlarda herhangi bir değişiklik yapmadığımı,
- Tez olarak sunduğum bu çalışmanın özgün olduğunu,

bildirir, aksi bir durumda bu konuda hakkımda yapılacak tüm yasal işlemleri ve aleyhime doğabilecek tüm hak kayıplarını kabullendiğimi beyan ederim. 01/07/2024

Zeynep BAHAR KILINÇ

Kayıtlı olunan Program : Tezli Yüksek Lisans
Tezin Konusu : Salgınlardaki Vaka Tahminlerinde Kullanılacak Modellerin Karşılaştırılması COV
Sayıları Tahmini

Tezin Türü : Yüksek Lisans : [X] Dokto[ ]

Danışmanın Adı-Soyadı : Doç. Dr. İlker ÜNAL
Danışmanın İletişim Bilgileri
Telefon : 0 (322) 338 60 60 -
E-Posta : ilkerun@cu.edu.tr

Öğrencinin İletişim Bilgileri
Telefon : 0 (533) 242 91 59
E-Posta : Virtue38@hotmail.com
Adresi : Çukurova Üni. Tıp Fakültesi
Biyoistatistik AD.
Sarıçam/ADANA

**Bu belgenin Lisansüstü eğitim tezleri savunmaya alınmadan önce öğrenci tarafından doldurulup imzalanarak Enstitü Müdürlüğüne teslim edilmesi gerekmektedir.*

## TEŞEKKÜR

Tez çalışmamın her aşamasında tüm bilgi birikimi ve deneyimiyle beni aydınlatan, her detayı titizlikle inceleyerek beni yönlendiren ve desteğini esirgemeyip büyük katkılar sağlayan değerli tez danışmanı hocam Doç. Dr. İlker ÜNAL'a teşekkür ederim. Tezime önerileriyle sağladıkları değerli katkılardan dolayı Prof. Dr. Gülşah SEYDAOĞLUNA ve Prof. Dr. Gülhan OREKİCİ TEMEL'e, yüksek lisans eğitimimde ki katkılarından dolayı Doç. Dr. Yaşar SERTDEMİR'e, tez sürecimdeki fikir alışverişleri ve güler yüzlü tavırlarıyla destek veren Biyoistatistik Anabilim Dalı araştırma görevlisi hocalarıma teşekkür ederim. Üniversite öğrenimime başladığım ilk günden beri yol gösterici olan, yüksek lisans eğitimine başlamamı sağlayan ve bu yolda beni yüreklendiren kıymetli hocam Prof. Dr. Zeliha Nazan ALPARSLAN'a ayrıca teşekkür ederim. Yüksek Lisans eğitimim boyunca desteklerini esirgemeyen Çukurova Üniversitesi Sağlık Bilimleri Enstitüsüne teşekkür ederim. Süreç boyunca tezi bitirmem için, tüm sabırlı ve anlayışlı tavırlarıyla bana en büyük manevi desteği vererek motive eden ve her koşulda yanımda olarak yardımlarını esirgemeyen, değerli annem Ayten BAHAR'a ve sevgili eşim Burak KILINÇ'a teşekkür ederim. Her zaman en büyük mutluluk kaynağım olan canım oğlum Kıvanç KILINÇ'a teşekkür ederim.

Zeynep BAHAR KILINÇ

# İÇİNDEKİLER

KABUL VE ONAY ..... III
ETİK BEYANI ..... IV
TEŞEKKÜR ..... V
İÇİNDEKİLER ..... VI
ÇİZELGELER DİZİNİ ..... VIII
ŞEKİLLER DİZİNİ ..... IX
SİMGELER DİZİNİ ..... XII
KISALTMALAR DİZİNİ ..... XIII
ÖZET ..... XV
ABSTRACT ..... XVI
1. GİRİŞ ..... 1
2. GENEL BİLGİ ..... 5
2.1. Matematiksel Modeller ..... 5
2.2. Kompartıman Modeller ..... 6
2.2.1. SIR ..... 6
2.2.2. SIS ..... 8
2.2.3. SIRD ..... 10
2.2.4. SIRV ..... 11
2.2.5. MSIR ..... 12
2.2.6. MSEIR ..... 13
2.2.7. SIRS ..... 14
2.2.8. SEIR ..... 15
2.3. Zaman Serisi Analizleri ..... 17
2.4. Box-Jenkins Yöntemi ..... 17
2.4.1. Mevsimsel Olmayan Box-Jenkins Modelleri ..... 19
2.4.1.1. AR (Otoregresyon Modelleri) ..... 19
2.4.1.2. MA (Hareketli Ortalama Modeli) ..... 20
2.4.1.3. ARMA (Otoregresif Hareketli Ortalam Modeli) ..... 20
2.4.1.4. ARIMA (Otoregresif Hareketli Entegre Ortalamalar Modeli) ..... 21
2.4.2. Mevsimsel Box-Jenkins Modelleri ..... 21

2.4.2.1. Mevsimsel SAR (Mevsimsel Otoregresyon) Modeli ..... 21
2.4.2.2. Mevsimsel SMA (Mevsimsel Hareketli Ortalama) Modeli ..... 22
2.4.2.3. Mevsimsel SARMA (Mevsimsel Otoregresif Hareketli Ortalama) Modeli ..... 22
2.4.2.4. Mevsimsel SARIMA( Mevsimsel Otoregresif Entegre Hareketli Ortalama) Modeli ..... 22
3. GEREÇ ve YÖNTEM ..... 23
5. BULGULAR ..... 26
5.1. SIR Yaklaşımı ile Yapılan Tahminler ..... 26
5.1.1. Türkiye Verilerinin SIR Yaklaşımı ile Tahmini ..... 26
5.1.2. Almanya Verilerinin SIR Yaklaşımı ile Tahmini ..... 31
5.1.3. Amerika Verilerinin SIR Yaklaşımı ile Tahmini ..... 35
5.1.4. Çekya Verilerinin SIR Yaklaşımı ile Tahmini ..... 38
5.2. ARIMA Yaklaşımı ile Yapılan Tahminler ..... 42
5.2.1. Türkiye Verilerinin ARIMA Yaklaşımı ile Tahmini ..... 42
5.2.2. Almanya Verilerinin ARIMA Yaklaşımı ile Tahmini ..... 46
5.2.3. Amerika Verilerinin ARIMA Yaklaşımı ile Tahmini ..... 48
5.2.4. Çekya Verilerinin ARIMA Yaklaşımı ile Tahmini ..... 49
6. TARTIŞMA ..... 53
7. SONUÇ ve ÖNERİLER ..... 56
KAYNAKÇA ..... 57
EKLER ..... 60
ÖZGEÇMİŞ ..... 67

# ÇİZELGELER DİZİNİ

**Çizelge 4.1.** Türkiye verilerinin 3, 5 ve 7 gün önceki veriler kullanılarak yapılan günlük ve haftalık vaka tahminlerinin hataları .....30

**Çizelge 4.2**.Türkiye verilerinin farklı periyotlardaki 3, 5 ve 7 gün önceki veriler kullanılarak yapılan günlük ve haftalık vaka tahminlerinin hataları .....31

**Çizelge 4.3**.Almanya verilerinin 3, 5 ve 7 gün önceki veriler kullanılarak yapılan günlük ve haftalık vaka tahminlerinin hataları .....34

**Çizelge 4.4**.Amerika verilerinin 3, 5 ve 7 gün önceki veriler kullanılarak yapılan günlük ve haftalık vaka tahminlerinin hataları .....38

**Çizelge 4.5.** Çekya verilerinin 3, 5 ve 7 gün önceki veriler kullanılarak yapılan günlük ve haftalık vaka tahminlerinin hataları .....42

**Çizelge 4.6**. Türkiye verilerinin ilk 250 gün ve 275. gün sonrası eriler kullanılarak yapılan günlük ve haftalık vaka tahminlerinin hataları .....46

**Çizelge 4.7**. Almanya verilerinin günlük ve haftalık vaka tahminlerinin hataları .....47

**Çizelge 4 8.** Amerika verilerinin günlük ve haftalık vaka tahminlerinin hataları .....48

**Çizelge 4.9**. Çekya verilerinin günlük ve haftalık vaka tahminlerinin hataları .....50

# ŞEKİLLER DİZİNİ

**Şekil 2.1.** SIR modelinde bölmeler arası geçiş katsayıları ..... 7
**Şekil 2.2.** SIS modelinde bölmeler arası geçiş katsayıları..... 9
**Şekil 2.3.** SIRD modelinde bölmeler arası geçiş katsayıları ..... 10
**Şekil 2.4.** SIRV modelinde bölmeler arası geçiş katsayıları ..... 11
**Şekil 2.5**. MSIR modelinde bölmeler arası geçiş katsayıları..... 13
**Şekil 2.6.** MSEIR modelinde bölmeler arası geçiş katsayıları ..... 14
**Şekil 2.7.** SIRS modelinde bölmeler arası geçiş katsayıları ..... 15
**Şekil 2.8.** SEIR modelinde bölmeler arası geçiş katsayıları..... 16
**Şekil 4.1.** 3 gün önceki veriler ile oluşturulan SIR modeline ait Türkiye verilerinin günlük vaka tahmini ..... 27
**Şekil 4.2**. 3 gün önceki veriler ile oluşturulan SIR modeline ait Türkiye verilerinin haftalık vaka tahmini ..... 27
**Şekil 4.3.** 5 gün önceki veriler ile oluşturulan SIR modeline ait Türkiye verilerinin günlük vaka tahmini ..... 28
**Şekil 4.4**. 5 gün önceki veriler ile oluşturulan SIR modeline ait Türkiye verilerinin haftalık vaka tahmini ..... 28
**Şekil 4.5.** 7 gün önceki veriler ile oluşturulan SIR modeline ait Türkiye verilerinin günlük vaka tahmini ..... 29
**Şekil 4.6.** 7 gün önceki veriler ile oluşturulan SIR modeline ait Türkiye verilerinin haftalık vaka tahmini ..... 29
**Şekil 4.7.** 3 gün önceki veriler ile oluşturulan SIR modeline ait Almanya verilerinin günlük vaka tahmini ..... 31
**Şekil 4.8.** 3 gün önceki veriler ile oluşturulan SIR modeline ait Almanya verilerinin haftalık vaka tahmini ..... 32
**Şekil 4.9.** 5 gün önceki veriler ile oluşturulan SIR modeline ait Almanya verilerinin günlük vaka tahmini ..... 32
**Şekil 4.10.** 5 gün önceki veriler ile oluşturulan SIR modeline ait Almanya verilerinin haftalık vaka tahmini ..... 33
**Şekil 4.11.** 7 gün önceki veriler ile oluşturulan SIR modeline ait Almanya verilerinin günlük vaka tahmini ..... 33

**Şekil 4.12.** 7 gün önceki veriler ile oluşturulan SIR modeline ait Almanya verilerinin haftalık vaka tahmini ..... 34
**Şekil 4.13.** 3 gün önceki veriler ile oluşturulan SIR modeline ait Amerika verilerinin günlük vaka tahmini ..... 35
**Şekil 4.14.** 3 gün önceki veriler ile oluşturulan SIR modeline ait Amerika verilerinin haftalık vaka tahmini ..... 35
**Şekil 4.15.** 5 gün önceki veriler ile oluşturulan SIR modeline ait Amerika verilerinin günlük vaka tahmini ..... 36
**Şekil 4.16.** 5 gün önceki veriler ile oluşturulan SIR modeline ait Amerika verilerinin haftalık vaka tahmini ..... 36
**Şekil 4.17.** 7 gün önceki veriler ile oluşturulan SIR modeline ait Amerika verilerinin günlük vaka tahmini ..... 37
**Şekil 4.18.** 7 gün önceki veriler ile oluşturulan SIR modeline ait Amerika verilerinin haftalık vaka tahmini ..... 37
**Şekil 4.19.** 3 gün önceki veriler ile oluşturulan SIR modeline ait Çekya verilerinin günlük vaka tahmini ..... 39
**Şekil 4.20.** 3 gün önceki veriler ile oluşturulan SIR modeline ait Çekya verilerinin haftalık vaka tahmini ..... 39
**Şekil 4.21.** 5 gün önceki veriler ile oluşturulan SIR modeline ait Çekya verilerinin günlük vaka tahmini ..... 40
**Şekil 4.22.** 5 gün önceki veriler ile oluşturulan SIR modeline ait Çekya verilerinin haftalık vaka tahmini ..... 40
**Şekil 4.23.** 7 gün önceki veriler ile oluşturulan SIR modeline ait Çekya verilerinin günlük vaka tahmini ..... 41
**Şekil 4.24.** 7 gün önceki veriler ile oluşturulan SIR modeline ait Çekya verilerinin haftalık vaka tahmini ..... 41
**Şekil 4.25.** ARIMA modeli ile Türkiye verilerinin günlük vaka tahmini ..... 43
**Şekil 4.26.** ARIMA modeli ile Türkiye verilerinin haftalık vaka tahmini ..... 43
**Şekil 4.27.** ARIMA modeli ile ilk 250 günlük Türkiye verilerinin günlük vaka tahmini ..... 44
**Şekil 4.28.** ARIMA model ile ilk 250 günlük Türkiye verilerinin haftalık vaka tahmini ..... 44

**Şekil 4.29.** ARIMA modeli ile 275. gün sonrası Türkiye verilerinin günlük vaka tahmini ..... 45
**Şekil 4.30.** ARIMA modeli ile 275. gün sonrası Türkiye verilerinin haftalık vaka tahmini ..... 45
**Şekil 4.31.** ARIMA modeli ile Almanya verilerinin günlük vaka tahmini ..... 47
**Şekil 4.32.** ARIMA modeli ile Almanya verilerinin haftalık vaka tahmini ..... 48
**Şekil 4.33.** ARIMA modeli ile Amerika verilerinin günlük vaka tahmini ..... 49
**Şekil 4.34.** ARIMA modeli ile Amerika verilerinin haftalık vaka tahmini ..... 49
**Şekil 4.35.** ARIMA modeli ile Çekya verilerinin günlük vaka tahmini ..... 50
**Şekil 4.36.** ARIMA modeli ile Çekya verilerinin haftalık vaka tahmini ..... 51
**Şekil 4.37**. Ülkelere göre SIR ve ARIMA modellerinin RMSE ve göreceli RMSE değerleri ..... 51

## SİMGELER DİZİNİ

| | | |
|---|---|---|
| S | : | Hastalığa duyarlı birey sayısı |
| I | : | Enfekte olmuş birey sayısı |
| R | : | İyileşen birey sayısı |
| E | : | Hastalığa maruz kalan birey sayısı |
| M | : | Doğuştan bağışıklığı olan birey sayısı |
| I | : | Arıma fark alma katsayısı |
| $\beta$ | : | S kompartımanından ve I kompartımanına iletim katsayısı |
| $\gamma$ | : | Bulaşıcı bireyin iyileşme hızı |
| $\delta$ | : | Maruz kalan bireyin bulaşıcı hale gelmesi |
| $\mu$ | : | Ölen kişi sayısı |
| $\theta$ | : | Hareketli ortalama katsayısı |
| $\omega$ | : | Bağışıklık geçtikten sonra duyarlı hale gelme katsayısı |
| $\phi$ | : | Otoregresyon katsayısı |
| rv | : | Aşılanma katsayısı |
| $R_0$ | : | Temel üreme sayısı |
| p | : | Otoregresyon parametresi |
| q | : | Hareketli ortalama parametresi |
| d | : | Fark alma parametresi |
| P | : | Mevsimsel otoregresyon katsayısı |
| Q | : | Mevsimsel hareketli ortalama katsayısı |
| D | : | Mevsimsel fark alma katsayısı |
| S | : | Mevsim periyodu katsayısı |

## KISALTMALAR DİZİNİ

| | | |
|---|---|---|
| DSÖ (WHO) | : | Dünya Sağlık Örgütü (World Health Organization) |
| COVİD-19 | : | Korona Virüs Hastalığı |
| SIR | : | Duyarlı-Enfekte-İyileşmiş (Susceptible-Infected-Removed) |
| SIS | : | Duyarlı-Enfekte-Duyarlı(Susceptible-Infected-Susceptible) |
| SEIR | : | Duyarlı-Maruz kalan-Enfekte-İyileşmiş (Susceptible-Exposed-Infected-Removed) |
| SIRS | : | Duyarlı-Enfekte-İyileşmiş-Duyarlı (Susceptible-Infected-Removed-Susceptible) |
| SIRV | : | Duyarlı-Enfekte-İyileşmiş-Aşılanmış (Susceptible-Infected-Removed-Vaccinated) |
| SIRD | : | Duyarlı-Enfekte-İyileşmiş-Ölmüş (Susceptible-Infected-Removed-Death) |
| MSIR | : | Doğuştan bağışıklığı olan-Duyarlı-Enfekte-İyileşmiş (Maternally immunity-Susceptible-Infected-Removed) |
| MSEIR | : | Doğuştan bağışıklığı olan-Duyarlı-Maruz Kalan-Enfekte-İyileşmiş (Maternally immunity-Susceptible-Exposed-Infected-Removed) |
| AR | : | Otoregresyon (Auto Regressive) |
| MA | : | Hareketli Ortalama (Moving Average) |
| ARMA | : | Otoregresif Hareketli Ortalama (Auto Regressive Moving Average) |
| ARIMA | : | Otoregresif Entegre Hareketli Ortalama (Auto Regressive Integrated Moving Average) |
| SAR | : | Mevsimsel Otoregresyon (Seasonal Auto Regression) |
| SMA | : | Mevsimsel Hareketli Ortalama (Seasonal Moving Average) |
| SARMA | : | Mevsimsel Otoregresyon Hareketli Ortalama (Seasonal Auto Regression Moving Average) |
| SARIMA | : | Mevsimsel Otoregresyon Entegre Hareketli Ortalama (Seasonal Auto Regression Integrated Moving Average) |
| ACF | : | Otokorelasyon Fonksiyonu (Auto Correlation Function) |

| | | |
|---|---|---|
| PACF | : | Kısmi Otokorelasyon Fonksiyonu (Partial Auto Correlation Function) |
| HKOK | : | Hata Kareleri Ortalama Karekökü(Root Mean Squared Error) |
| GHKOK | : | Göreceli Hata Kareleri Ortalama Karekökü |

# ÖZET

## Salgınlardaki Vaka Tahminlerinde Kullanabilecek Modellerin Karşılaştırılması: Covid-19 Vaka Sayıları Tahmini

**Toplum üzerinde sağlık, ekonomik ve kültürel gibi yönlerden ciddi etkiler yaratan salgının kontrol altına alınabilmesi için, hastalık nedenlerinin belirlenmesi, bulaşın engellenmesi ve hastaların hızlı şekilde tedavi edilmesi gerekmektedir. Bu basamakların geçilmesinde özellikle hastalığın seyri ve vaka sayıları büyük önem taşımaktadır. Bu nedenle, mevcut vakalardan hareketle ilerleyen dönemlerde ortaya çıkabilecek yeni vakaların sayısını belirleyebilecek matematiksel modellere her zaman ihtiyaç duyulmuştur. 20.yy başlarında popülerliği artan bu modeller tahmin başarıları yönünden değişkenlik göstermektedir. Bu tez çalışmasında bu modellerden sıklıkla kullanılan SIR modeli ve ARIMA modeli Covid-19 pandemi verileri kullanılarak karşılaştırılmış ve üstünlük ve eksiklikleri belirlenmiştir.**

**Tez çalışmasında Almanya, Amerika, Türkiye ve Çekya ülkelerinin ilk vakasının görüldüğü günden başlayarak, 550 günlük verileri kullanılmıştır. SIR modelinin parametrelerini tahmin edebilmek için geriye yönelik 3, 5 ve 7 günlük değerler kullanılmıştır. Günlük vaka sayıları ve günler arası ötelenme görülebildiği için haftalık toplam vaka sayıları modeller kullanılarak tahmin edilmiştir. Bu nedenle ARIMA modelinde frekans değeri 365 ve 52 olarak alınmıştır. Tahmin başarıları hata değerleri ortalamasının karekökü (RMSE) ve seyir (trend) grafikleri ile incelemiştir.**

**Yapılan analizler sonucunda tüm ülkeler için geçerli olmak üzere SIR modelinde parametre tahmini için kullanılacak geriye yönelik gün sayısının etki ettiği gözlenmiştir. Buna göre genellikle 5 veya 7 günlük geçmiş bilginin tahmin başarısını arttırdığı gözlenmiştir. ARIMA modelinin günlük vaka tahminlerinde haftalık vaka tahminlerine göre daha az hata yaptığı görülmüştür.**

**Bu çalışmanın sonucunda modellerin birbirleri üzerinde belirgin bir üstünlüğü olmadığı ve SIR modelinde parametre tahmininde en az 5 günlük veriye ihtiyaç olduğu, ARIMA modellerinde ise haftalık vaka tahminleri için daha uzun veri periyotlarına ihtiyaç olduğu belirlenmiştir.**

**Anahtar Kelimeler: Salgın Hastalık, Matematiksel Modelleme, COVID-19, SIR Modeli, ARIMA Modeli, WHO**

## ABSTRACT

**A Comparison of Models That Can Be Used in Case Estimations in Epidemics: Estimating the Number of Covid-19 Cases**

**In order to control the epidemic, which may have serious effects on society in terms of health, economy and education, it is necessary to determine the causes of the disease, prevent transmission and treat patients quickly. In the course of implementing these steps, it is of great importance to monitor the disease's progression and the number of cases. Consequently, mathematical models that can estimate the number of new cases have always been needed. These models, which became popular in the early 20th century, vary in terms of their predictive success. In this thesis, the SIR model and the ARIMA model, which are frequently used among these models, were compared using Covid-19 pandemic data and their advantages and disadvantages were determined.**

**In this thesis, 550-day data were used, starting from the day the first case was seen in Germany, America, Turkey and the Czech Republic. In order to estimate the parameters of the SIR model, previous 3, 5 and 7-day values were employed. The daily case numbers and as well as the weekly total case numbers were estimated using the models. The frequency value in the ARIMA model was set at 365 and 52. The performance of the models were examined using root mean square error (RMSE) values and trend graphs.**

**The results of the analyses indicated that the number of days to be used for parameter estimation in the SIR model affected the performance of the SIR model. Furthermore, the use of 5 or 7 days generally increased the estimation success. Additionally, it was observed that the ARIMA model has lower RMSE values in daily case estimations than in weekly case estimations.**

**The findings of this study indicate that the models do not exhibit a significant degree of superiority over one another. Furthermore, the study revealed that at least five days of data are required for parameter estimation in the SIR model, while longer data periods are necessary for weekly case estimation in ARIMA models.**



**Key Words: Epidemics, Mathematical Modeling, COVID-19, SIR Model, ARIMA Model, WHO**

# 1. GİRİŞ

Salgın belirli bir popülasyonda belirli bir zamanda enfeksiyon hastalığına yakalanan bireylerin sayısında kontrolsüz artış demektir. Epidemiyolojide salgın, bir hastalıktaki vaka sayısının yer ve mevsim normallerine göre ani artış göstermesi olarak tanımlanır [(1)]. Hastalık bir alanda kalabilir veya daha geniş bir alana yayılabilir ve günler veya yıllar sürebilir. Belli bir bölgede yayılan salgınlara epidemi, dünya çapında yayılan salgınlara ise pandemi denir. Yüzyıllar boyunca salgın hastalıklar sonucu insanlık tarihinde birçok pandemi görülmüştür. 2019 yılının sonunda Çin'in Wuhan kentinde başlayıp tüm dünyayı etkisi altına alan Covid 19 pandemisinin etkileri hala devam etmektedir.

Salgınlar toplumları sağlık, ekonomi, eğitim, sosyal yaşam gibi birçok yönden etkilemektedir. Bu etkilerin büyüklükleri, salgının yayılım hızına ve vaka yoğunluğuna bağlıdır. Bu nedenle salgın ile ilgili istatistikler çok önemlidir. Salgının yayılımı, parametre tahmini ve vaka sayılarının önceden belirlenip tedbir alınabilmesi için matematiksel modellere ihtiyaç vardır. Matematiksel modeller sayesinde salgın hastalıkların yayılım oranı ve hızı önceden tespit edilerek önlemler alınır ve bunun sonucunda hastalık artışının engellenmesi hedeflenir. Matematiksel modeller bulaşıcı hastalıklara ilişkin parametreleri istatistiksel yöntemler kullanarak tahmin etmeye çalışır ve bunun sonucunda modellemeler hangi müdahale tekniklerinin kullanılacağına karar verir.

Literatürde, salgın hastalıklar için formül geliştirmede ilk adım 1765 yılında çiçek hastalığı için yaşam tablolarının yapılması amacıyla Bernoulli tarafından atılmıştır [(2)]. Formül geliştirme konusunda bir başka çalışma ise Sir Ronald Ross tarafından yapılan, sıtmanın böcek vektörleri ve insan konakçılar arasında yayılmasına ilişkin matematiksel model geliştirme çalışmasıdır [(3)].

Yirminci yüzyılın başlarında salgın için matematiksel model geliştirme çalışmaları yapılırken bir yandan da salgın hastalıkların neden tüm duyarlı kişilere bulaşmadan bittiği konusunda çalışmalar yapılmaktaydı. Bu konuda bir ilk olarak, Hamer [(4)] yaptığı çalışma ile bulaşıcı hastalıkların yayılmasının dinamiklerinin doğrusal olmadığını ve salgını durdurabilecek tek şeyin duyarlı kişilerin yoğunluğunun azalması olduğunu belirtmiştir.

Salgın hastalıkların yayılımı için matematiksel model geliştirme ise ilk olarak 1927 yılında yayımlanan "A Contribution to the Mathematical Theory of Epidemics" isimli makaleleri ile A. G. McKendrick ve W. O. Kermack tarafından yapılmıştır [5]. Kermack ve McKendrick çalışmalarında Kermack ve McKendrick teorisi olarak geçen ve SIR modeli olarak bilinen bir epidemiyolojik model önermiştir. SIR modeli salgın hastalıkların evrimi için en çok kullanılan modellerden biridir. Model, üç niceliğin gelişimini gösteren birleştirilmiş diferansiyel denklemler sistemidir. Bu üç nicelik; enfekte olabilen ancak henüz bir hastalık tarafından enfekte edilmemiş bireylere karşılık gelen bir popülasyonun "duyarlı" S(t) sayısı; enfekte kişi sayısı I(t) ve iyileşen kişi sayısı R(t) dir [6]. SIR modeli daha da geliştirilerek doğum ve ölüm oranları da sisteme katılmıştır [7].

Hethcote [8] salgın hastalıklar için üç temel epidemiyolojik model türünü tanımlamıştır: SIS modeli, vital bulgusu olmayan SIR modeli ve vital bulgulu SIR modeli. Bu modellerde temel üreme oranı (basic reproductive rate) $R_0$ ve temas edilen kişi sayısı $\sigma$ gibi değerleri kullanmıştır. Kuznetsov ve Piccardi sinüzoidal olarak değişen temas hızıyla SEIR ve SIR salgın modellerinin periyodik çözümlerinin değişimlerini araştırmıştır [9]. Peng ve arkadaşları latent dönemde enfeksiyonun yayılmasını hesaba katan değiştirilmiş SEIR modeli uygulamıştır [10]. Wallinga ve Teunis SARS virüsünün patojen olarak ortaya çıkmasından sonra müdahale tedbirlerinin etkinliğini değerlendirmek amacıyla salgın hastalık verilerini gerçek zamanlı analiz etmek için matematiksel modelleme kullanmıştır [11]. Bir başka tahmin modeli olarak zaman serisi tahminlerinden en popüler olanı otomatik regresyonlu entegre hareketli ortalama (ARIMA) modelidir. Sato, zaman serileri ile hastalık yönetimi teknikleri için ARIMA modelini kullanmıştır [12]. Bu çalışmada sağlık uygulamalarındaki farklılıkları tespit etme ve takip etme seçeneklerine önem verilmesi gerektiği vurgulanmaktadır.

Çin'in Wuhan kentinde 2019'da yeni bir koronavirus olan SARS-CoV-2 virüsü ortaya çıkmış ve kısa sürede tüm şehri etkilemiştir. Sonrasında tüm dünyaya kontrolsüz bir şekilde yayılmış ve dünyayı etkisi altına almıştır. Virüs insanlarda ölümcül solunum sendromuna neden olmaktadır ve çok bulaşıcı olduğu için insandan insana hızlı bir şekilde aktarılmaktadır. Bunun sonucunda ortaya çıkan salgın, dünya çapında bir sürü kişinin ölümüne yol açmış ve Mart 2020'de Dünya Sağlık Örgütü (WHO) tarafından Covid-19 pandemisi olarak ilan edilmiştir [13]. Ülkeler salgınla mücadele konusunda

karantina, sosyal mesafe ve seyahat kısıtlamaları gibi birçok yol izlemelerine rağmen bu kısıtlamalar salgını durdurmada yeterli olmamıştır. Salgın ülkelerin ekonomilerini, sağlık sistemlerini ve sosyal hayatını olumsuz yönde etkilemiştir. Birçok farklı alandan bilim insanı Covid-19 salgınını farklı açılardan araştırmıştır. Salgının gelişimini analiz etmek ve sonuçlarını tahmin etmek için birçok çalışma yapılmış ve yeni tahmin modelleri önerilmiştir. [14-23]

Wuhan'daki (Hubei Eyaleti, Çin) COVID-19 salgınının boyutunu değerlendirmek için Imai ve ark [24], insan bulaşmasına odaklanan potansiyel bir salgının hesaplama modellerini önermiştir. Bulguları, müdahalenin salgınları önlemek için SARS-CoV-2 bulaşlarının %60'ını engelleyebileceğini öne sürmüştür. Gao ve ark. [25], SARS-CoV-2'nin bulaşıcılığını değerlendirmek ve vaka sayılarını tahmin etmek için bir algoritma geliştirmiştir. SARS-CoV-2'nin konakçılarının yarasalar ve vizonlar olabileceğini öne sürmüşlerdir. Tang ve ark. [26], klinik hastalığın ilerlemesini, bireysel epidemiyolojik durumu ve katılımcı davranışını içeren deterministik bir parçalı model önermiştir. Karantina eşliğinde müdahale yöntemlerinin (örneğin, enfekte olmuş kişilerin dışlanması) SARS-CoV-2'nin bulaşma ve üreme olasılığını azaltabileceğini belirtmişlerdir. Prasse ve ark [27] Hubei eyaletindeki salgını tanımlamak için şehirler ve trafik akışına göre oluşturulan ağ tabanlı bir model tasarlamıştır.

Wu ve ark. [28] , Wuhan'daki enfeksiyonların sayısını çıkarmak için Wuhan'dan uluslararası olarak ihraç edilen vaka sayısına ilişkin 31 Aralık 2019'dan 28 Ocak 2020'ye kadar olan verileri (25 Aralık 2019'dan 19 Ocak 2020'ye kadar bilinen semptomların başlangıç günleri) kullanarak Duyarlı-Maruz Kalmış-Bulaşıcı-İyileşen (SEIR) modeli ile analiz yapmıştır. Read ve ark. [29] Poisson günlük zaman artışları varsayımı ile veri dayalı olarak SEIR modeli için üreme sayısını (RN) 3.1 olarak elde etmiştir.

Pandemideki vaka tahmininde, basitliği ve sistematik yapısı nedeniyle tercih edilen ARIMA zaman serisi modeli, Ceylan tarafından İtalya, İspanya ve Fransa'da Kovid-19 salgınının yayılma boyutunu araştırmak amacıyla kullanılmıştır. [30]

Alzahrani ve arkadaşları [31], Suudi Arabistan'da ARIMA tahmin modelini kullanarak Covid-19 salgının yayılmasını incelemiştir. Yazarlar, ARIMA modelinin parametrelerinin "Akaike bilgi kriteri (AIC)" değerine göre seçildiği "doğrusal parametrik model tahmin yaklaşımını" kullanmışlardır.

Mohammad ve ark. Covid-19 salgını için doğrusal olmayan denklemlerin türevlerine dayanarak bir model oluşturmuş, bu modelle parametrelerin matematiksel tahminlerini yaparak grafiklerle mevcut toplam vaka sayısını elde etmişlerdir [32].

Güzey alınan önlemlerin Covid-19 virüsünün yayılmasını nasıl etkilediği ve daha sonra alınacak önlemlerin pandeminin seyri üzerindeki etkisine dair model oluşturulmuştur. Bu model için 91 günlük Türkiye verisi kullanılmış ve MATLAB programı ile parametre kestirimi yapılmıştır [33].

Literatürde salgındaki vaka tahminleri için farklı yaklaşımlar kullanılarak birçok çalışma yapıldığı görülmektedir. Ancak bu çalışmalarda yöntemlerin birbirine olan üstünlük ve eksiklikleri yeterince tartışılmamıştır. Bu noktadan hareketle bu çalışmada covid-19 salgınında vakaların önceden tahmin edilmesi için SIR ve ARIMA modellerinin uygulanması ve bu modellerin etkinliğinin karşılaştırılması amaçlanmıştır

# 2. GENEL BİLGİ

## 2.1. Matematiksel Modeller

Epidemiyoloji toplumda hastalık ve salgınların nedenlerini, sıklıklarını ve dağılımlarını inceleyen bilim dalıdır. Bu bilgilere göre, salgın ve hastalıkları azaltmayı amaçlar ve yöntemler belirler. Epidemiyolojik çalışmalar halk sağlığını hedef alır ve temel amaçları sağlıkla ilgili durumları tanımlamak, ölçmek ve nedenlerini çözümlemeye çalışmaktır. Epidemiyolojide salgın hastalıkların nedenini ve etkilerini araştırırken matematiksel modeller kullanılır.

Salgının yayılımı, ölüm ve iyileşme sayıları gibi parametrelerin kestirimi salgın yönetimi açısından büyük önem taşımaktadır. Bu kestirimleri yapılmasında matematiksel modellemeler büyük rol oynar [34]. Matematiksel modellerin amacı gerçek dünyadaki olayları matematiksel bir dille anlatmaktır. Matematiksel modellemeler sayesinde salgın hastalıkların yayılım oranı ve hızı önceden tespit edilerek önlemler alınır ve bunun sonucunda hastalık artışının engellenmesi hedeflenir. Modellemeler ile bir salgının olası sonucu belirlenerek hastalığın nasıl ilerlediği tahmin edebilir ve yapılacak halk sağlığı müdahaleleri ve bunların olası sonuçları belirlenebilir.

Bir bulaşıcı hastalığın seyrini tahmin eden ve ilk defa Kermack ve McKendrick tarafından önerilen matematiksel modelleme epidemiyoloji alanında yapılan ilk çalışmadır. Bu öneride model, üç niceliğin gelişimini gösteren birleştirilmiş diferansiyel denklemler sistemidir. Bu üç nicelik; enfekte olabilen ancak henüz bir hastalık tarafından enfekte edilmemiş bireylere karşılık gelen bir popülasyonun "duyarlı" S(t) sayısı; enfekte kişi sayısı I(t) ve iyileşen kişi sayısı R(t) dir [35].

Bulaşıcılık hastalık seyirleri bir zaman serisindeki seyir olarak da kabul edilebilir. Bu durumda, zaman serilerini tahmin etmeye yönelik ARIMA modeli salgının seyri için kullanılabilir. Bu model aynı zamanda Box-Jenkins yöntemi olarak bilinir, George E. Box ve Gwilym M. Jenkins tarafından önerilmiştir.

Epidemiyolojide literatürde önerilmiş birçok matematiksel model tekniği vardır. Bu tez çalışmasında kompartıman modellerden SIR ve BOX-JENKSİN metodundan ARIMA yöntemi detaylı olarak incelenmiştir.

### 2.2. Kompartıman Modeller

Kompartıman modeller, hastalık bulaşma dinamiklerini anlamak için uzun yıllardır kullanılmakta olan çok genel bir modelleme tekniğidir. Adını verinin kompartımanlara (bölümlere) ayrılmasında alan bu modellerde, toplam nüfus belirli bölümlere (kompartımanlara) atanır ve bölümler arasında geçiş olduğu varsayılır. Ayrıca modelde bölümlerin sıralı olduğu ve bölümler arası geçişlerin belirli sıraya yapıldığı olduğu varsayılır.

Kompartıman modeller ilk olarak 20. Yüzyılın başlarında önerilmiştir. Bu yöntemin kullanıldığı ilk çalışma, 1916 yılında Ronald Ross tarafından sıtmanın bulaşması üzerine yaptığı çalışmadır. Daha sonrasında 1927 yılında Kermack ve McKendrick zaman içerisinde popülasyonda bulaşıcı hastalık vakalarının sayısını ve dağılımını tahmin eden hipotezlerinde bu model tekniklerini kullanmıştır.

Kompartıman modellerde bölümler arası geçişler ve vaka tahminleri basit diferansiyel denklemler (ODE) kullanılarak yapılır. Bu tahminler sonucunda enfekte olan kişi sayısı, salgının yayılması ve alınan önlemlerin salgının seyrini nasıl etkileyeceği konusunda fikir vermektedir.

Bu modellerden en basit olanı SIR modelidir. Modelin salgın hastalıkların seyrine ve gerekliliklerine göre değişmesi ve çeşitli değişkenlerin eklenmesiyle gelişmiş farklı varyasyonları da mevcuttur. Bu modeller salgınlardaki birçok ayrıntıyı barındırır ve gerçeğe daha yakın sonuçlar elde edilerek, salgının seyrini daha başarılı şekilde açıklar. SIS, SIRD, SIRV, SIRS, MSIR, MSEIR ve SEIR modeli bunlardan bazılarıdır. Bir sonraki bölümde bu modellerin özelliklerine değinilecektir. Bu çalışmada kompartıman modellerden SIR modeli ele alınmıştır.

#### 2.2.1. SIR

Salgın hastalıkların yayılımı için matematiksel modeller 1927'de Kermack ve McKendrick tarafından önerilmiştir. Bu modellerde toplumdaki bireyler hastalığa ilişkin durumlarına göre belli gruplara ayrılıp bunlar arasındaki geçişler matematiksel modeller aracılığı ile belirlenir. Bu çerçevede geliştirilen bölmeli modellerin en temeli Duyarlı-Enfekte-İyileşen (Susceptible-Infected-Recovered – SIR) modelidir.

Bu yöntemde popülasyon 3 bölmeye ayrılır:

- S (Duyarlı – Susceptible): Popülasyondaki duyarlı bireyler, henüz hastalığa maruz kalmamış fakat hastalıkla karşılaştığında enfekte olma olasılığı olan kişilerdir.
- I (Enfekte – Infected): Virüs bulaşmış ve sonucunda enfekte olmuş ve hastalığı bulaştırma riski olan kişilerdir.
- R (İyileşen – Recovered): Virüs bulaşmış ve sonrasında iyileşmiş kişilerdir.

Bölmeler arası tek yönlü geçişler vardır;

**S** Zamanla azalır, bölmeye geri akış yoktur. Duyarlı bireyler enfekte olduktan sonra I (enfekte) bölmesine geçerler.

**I** Bu bölmede hem giriş hem çıkış vardır. Duyarlı bireyler bölmeye artı olarak gelebilir. İyileşen ve ölen bireyler bölmeden çıkar.

**R** Bölmeye giriş vardır. Ancak ölüm olursa bölmeden çıkış olur. İyileşen kişiler bağışıklık kazanmış demektir.




**Şekil 2.1.** SIR modelinde bölmeler arası geçiş katsayıları

Model parametreleri;

$\beta$ : İletim katsayısı, bulaşıcı bir bireyin hastalığı bulaştırma hızı.

$\gamma$ : Bulaşıcı bireylerin iyileşme hızı

Modelin varsayımları;

- Model popülasyonu sabit kabul eder.
- Popülasyonun izole olduğunu varsayar.
- Bireylerin genellikle bağışıklıkla doğmadığını kabul eder.
- R bölmesinden ancak ölüm olursa çıkış olur.

- Hastalığı atlatan bireyler tam bağışıklık kazanır.
- Hastalığın ömrü sınırlıdır.
- Virüsün kuluçka süresi yoktur, bireyler hastalığı alır almaz başkalarını hasta edebilir
- Model yaş, cinsiyet, göç, doğum ve ölüm gibi değişkenleri hesaba katmaz.

Model denklemleri;

Model t zamanındaki geçişleri, basit diferansiyel denklem (ODE) kullanarak belirler.

$$N = S(t) + I(t) + R(t)$$

$$S'(t) = -\frac{\beta S(t) I(t)}{N}$$

$$I'(t) = \frac{\beta S(t) I(t)}{N} - \gamma I(t)$$

$$R'(t) = \gamma \mathrm{I(t)}$$

SIR modeli için $R_0$ değeri;

$$R_0 = \frac{\beta}{\gamma}$$

Temel üreme katsayısı $R_0$, bir hastalığın ne kadar transfer edilebildiğinin ölçüsüdür. Tek bir bulaşıcı kişinin enfekte edebileceği insan sayısını gösterir. Salgının durumunun anlaşılmasını sağlar: $R_0 > 1$ ise salgın artmakta, $R_0 < 1$ ise salgın kontrol altına alınmakta anlamına gelmektedir.

### 2.2.2. SIS

Matematiksel modellerin içerisinde kompartıman modellerin en temelleri SIR ve SIS modelleridir. SIR modelinin aksine, SIS modelinde kişiler enfeksiyondan kurtulduktan sonra tamamen bağışıklık kazanmaz ve hastalık ile karşılaştığı zaman tekrardan enfekte olma olasılığı vardır. Bu nedenle tekrar duyarlı bölmesine geçiş yapar. Bunun sonucunda SIS modelinde iki tane bölme vardır. Tekrarlamalı hesaplama yöntemi kullanılarak hesaplanır.

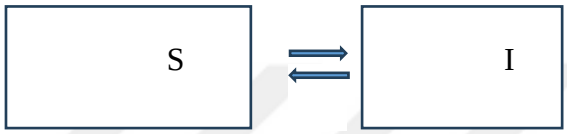
- S (Susceptible):Popülasyondaki duyarlı bireyler.
- I (Infected): Virüs bulaşmış ve sonucunda enfekte olmuş kişiler.

Bu modelde iki bölme arasında geçişler vardır. S bölmesindeki bireyler (duyarlı), enfekte olup I bölmesine geçtikten sonra tekrar S bölmesine döneceklerdir.

Modelde parametreleri ;

$\beta$ : Hastalığın bulaşma hızı

$\gamma$ : İyileşme oranı.




**Şekil 2.2.** SIS modelinde bölmeler arası geçiş katsayıları

Modelin varsayımları;

- Model popülasyonu sabit kabul eder.
- Hastalıklı kişilerin iyileştikten sonra tekrar enfekte olmadığı kabul edilir.

Bu model, soğuk algınlığı, grip, bel soğukluğu veya klamidya gibi cinsel yolla bulaşan hastalıklar gibi sıklıkla tekrarlanan enfeksiyonlara sahip hastalıklar için uygundur. Bu tarz enfeksiyonlar geçici direnç sağlarlar fakat hastalık geçtikten sonra kalıcı bağışıklık oluşmaz bu nedenle bireyler tekrar duyarlı konuma geçerler.

Model denklemleri;

SIS modeli analitik olarak ODE diferansiyel denklemi kullanılarak hesaplanır.

$$N = S(t) + I(t)$$

$$S'(t) = \frac{-\beta I(t)S(t)}{N} + \gamma I(t)$$

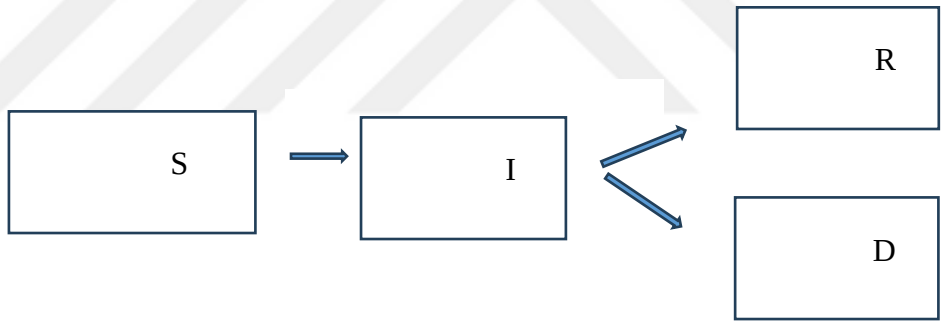
$$I'(t) = \frac{-\beta I(t)S(t)}{N} - \gamma I(t)$$

### 2.2.3. SIRD

Salgın hastalıkların araştırılması için kullanılan bir diğer matematiksel model SIRD modelidir. Bu model, hastalıktan iyileşen kişiler ve hastalıktan ölen kişiler arasındaki ayrımı gösterir. Modelde 4 tane bölme vardır.

- S (Susceptible):Popülasyondaki enfeksiyona yakalanma riski olan duyarlı bireyler.
- I (Infected): Virüs bulaşmış ve sonucunda enfekte olmuş kişiler.
- R (Recovered): Hastalıktan iyileşen kişi sayısı.
- D (Deceased): Hastalıktan ölen kişi sayısı.

Model parametreleri ;
β : Hastalığın bulaşma hızı
γ : İyileşme oranı
μ : Ölüm oranı.




**Şekil 2.3**. SIRD modelinde bölmeler arası geçiş katsayıları

Modelin varsayımları ;

- İyileşen kişiler artık enfeksiyona duyarlı değildir.
- Araştırılan enfeksiyon hastalığından bağımsız ölümlerin sayısı dikkate alınmaz.
- İncelenen bölgenin diğer bölgelerden izole olduğu kabul edilir.
- Popülasyon sabittir.

Model denklemleri ;
SIRD modeli diferansiyel denklemler kullanılarak hesaplanır.

$$N = S(t) + I(t) + R(t) + D(t)$$

$$S'(t) = \frac{-\beta I(t)S(t)}{N}$$

$$I'(t) = \frac{-\beta I(t)S(t)}{N}\gamma I(t) - \mu I(t)$$

$$R'(t) = \gamma I(t)$$

$$D'(t) = \mu I(t)$$

### 2.2.4. SIRV

Aşılama, duyarlı sayısını azalttığı için salgın hastalıkların yayılımını kontrol altına almak için kullanılan yöntemler arasındadır. Son zamanlarda yapılan matematiksel çalışmalarda aşılamayı hesaba katan salgın modelleri geliştirilmiştir. SIR modelinde duyarlı kişilerin aşılanması hesaba katılarak SIRV modeli geliştirilmiştir.

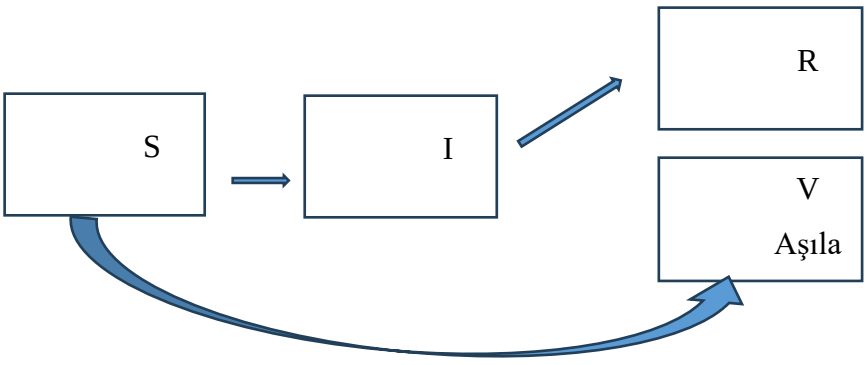
- S (Susceptible):Popülasyondaki enfeksiyona yakalanma riski olan duyarlı bireyler.
- I (Infected): Virüs bulaşmış ve sonucunda enfekte olmuş kişiler.
- R (Recovered): Hastalıktan iyileşen kişi sayısı.
- V (Vaccinated): Aşılanmış duyarlı birey sayısı.

Model parametreleri ;

β : Hastalığın bulaşma hızı

γ : İyileşme oranı

v : Duyarlı kişilerin aşılanma oranı.




**Şekil 2.4.** SIRV modelinde bölmeler arası geçiş katsayıları

Modelin varsayımları;

- Popülasyon sabittir.
- Aşılamadan önce tüm bireyler S bölmesine aittir, veri eksikliği nedeniyle kendi kendine aşı yapan bireyler modele dahil edilmez.
- Model aşılanan kişilerin direk R bölmesine geçtiğini ve tamamen bağışıklık kazandığını varsayar.
- Model varyantları hesaba katmaz.

Model denklemleri;

SIRV modeli diferansiyel denklemler kullanılarak matematiksel olarak hesaplanır.

$$N = S(t) + I(t) + R(t) + V(t)$$

$$S'(t) = \frac{-\beta S(t) I(t)}{N} - V(t) S(t)$$

$$I'(t) = \frac{-\beta S(t) I(t)}{N} - \gamma I(t)$$

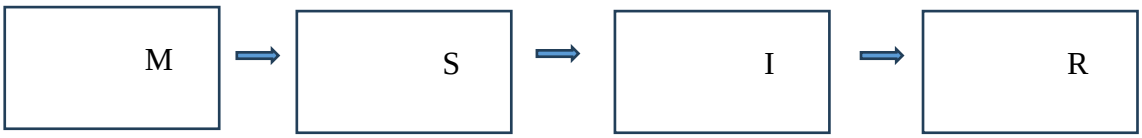
$$R'(t) = \gamma I(t)$$

### 2.2.5. MSIR

SIR modeliyle geliştirilen bir diğer matematiksel metot da MSIR modelidir. Bu modelde bebeklerde bazı hastalıklar için doğuştan gelen bağışıklık olduğu düşünülerek modele M (maternally-derived immunity) bölmesi eklenmiştir. Modelde 4 bölme vardır.

- M (Maternally-derived immunity): Pasif immun yenidoğan.
- S (Susceptible):Popülasyonda ki enfeksiyona yakalanma riski olan duyarlı bireyler.
- I (Infected): Virüs bulaşmış ve sonucunda enfekte olmuş kişiler.
- R (Recovered): Hastalıktan iyileşen kişi sayısı.

Yenidoğanlar annelerinden aldığı antikorlar sayesinde bazı hastalıklara karşı pasif bağışıklığa sahip olarak doğarlar, bu sebepten duyarlı bölmesinde değerlendirilmezler. Bu bağışıklık yaşam boyunca devam etmez ilk birkaç ay içindir, sonrasında antikorlar uzaklaştığında birey duyarlı gruba dahil edilir.




**Şekil 2.5**. MSIR modelinde bölmeler arası geçiş katsayıları

Model denklemleri;

MSIR modeli diferansiyel denklemler kullanılarak matematiksel olarak hesaplanır.

$$N = S(t) + I(t) + R(t) + M(t)$$

$$M'(t) = \varphi - \delta M(t) - \mu M(t)$$

$$S'(t) = \delta M(t) - \frac{\beta S(t) I(t)}{N} - \mu S(t)$$

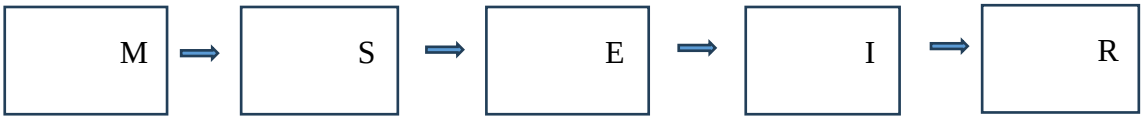
$$I'(t) = \frac{\beta S(t) I(t)}{N} - \gamma I(t) - \mu I(t)$$

$$R'(t) = \gamma I(t) - \mu R(t)$$

### 2.2.6. MSEIR

Bu modelde anneden gelen antikorlar sayesinde pasif bağışıklığa sahip olanlar için M bölmesi ve virüsü aldıktan sonra enfekte olana kadar geçen süre (latent periyot) için E bölmesi eklenmiştir. Hastalık anneye ulaşmışsa doğacak olan bebek pasif bağışıklıkla doğar, fakat antikorlar vücuttan uzaklaştığında tekrar S (duyarlı) bölmeye geçiş yapar. Duyarlı olan bireyler hasta olan kişilerle temas ettiğinde enfekte olana kadar geçen bir süre vardır, bu süre boyunca kişiler E (maruz kalmış) bölmesinde kalırlar. Modelde 5 tane bölme vardır.

- M (Maternally-derived immunity): Pasif immun yenidoğan.
- S (Susceptible):Popülasyondaki enfeksiyona yakalanma riski olan duyarlı bireyler.
- E (Exposed): Enfekte olmuş fakat henüz bulaşıcı olmayan ve semptomları olmayan kişiler.
- I (Infected): Virüs bulaşmış ve sonucunda enfekte olmuş kişiler.
- R (Recovered): Hastalıktan iyileşen kişi sayısı.




**Şekil 2.6.** MSEIR modelinde bölmeler arası geçiş katsayıları

Model denklemleri;

MSEIR modeli diferansiyel denklemler kullanılarak matematiksel olarak hesaplanır.

$$M'(t) = \varphi - \delta M(t) - \mu M(t)$$

$$S'(t) = \delta M(t) - \frac{\beta S(t) I(t)}{N} - \mu S(t)$$

$$E'(t) = \frac{\beta S(t) I(t)}{N} - (\varepsilon + \mu) E(t)$$

$$I'(t) = \varepsilon E(t) - (\gamma + \mu) I(t)$$

$$R'(t) = \gamma I(t) - \mu R(t)$$

### 2.2.7. SIRS

SIRS modeli SIR modelinden farklı olarak iyileşen bireylerin bağışıklığını kaybedeceğini ve hastalığa tekrar yakalanma durumu olduğundan, duyarlı bölmeye tekrar geçebileceğini varsaymaktadır. Bu model kızamık, suçiçeği, hepatit, grip, kabakulak ve AIDS gibi hastalıklar için uygundur.

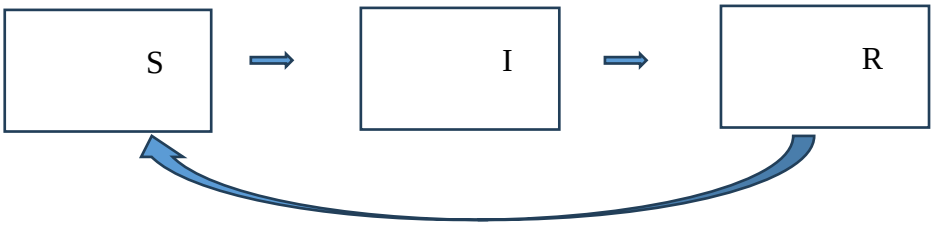
- S (Susceptible):Popülasyondaki enfeksiyona yakalanma riski olan duyarlı bireyler.
- I (Infected): Virüs bulaşmış ve sonucunda enfekte olmuş kişiler.
- R (Recovered): Hastalıktan iyileşen kişi sayısı.

Model parametreleri ;

β : Hastalık bulaşma katsayısı,

γ : İyileşme katsayısı

δ : İyileşen bireylerin tekrar duyarlı bölmeye geçme katsayısıdır.




**Şekil 2.7.** SIRS modelinde bölmeler arası geçiş katsayıları

Modelin Varsayımları;

- Bireyin bağışıklığı zamanla azabilir,
- İyileşme ömür boyu bağışıklık sağlamaz
- İyileşen bireyler hastalığa tekrar yakalanabilir.

Model Denklemleri;

SIRS modeli analitik olarak ODE diferansiyel denklemi kullanılarak hesaplanır.

$$N = S + I + R$$

$$S'(t) = -\beta IS + \delta R$$

$$I'(t) = \beta IS - \gamma I$$

$$R'(t) = \gamma I - \delta R$$

### 2.2.8. SEIR

SEIR modeli, Kermack ve McKendrick tarafından 1927'de sunulan klasik SIR modelinin bir uzantısıdır. Çoğu hastalıkta, ilk enfeksiyondan sonra konak, bulaşıcı hale gelmeden önce latent bir aşamada kalır ve bu süre, enfeksiyon dönemiyle karşılaştırıldığında enfekte olabilir. Böylece kuluçka (hastalığın ortaya çıkma evresi) süreside hesaba katılarak modele Exposed (maruz kalan) bölmesi eklenmiştir. Model 4 bölmeden oluşur.

- S (Susceptible): Popülasyonda ki duyarlı bireyler, henüz hastalığa maruz kalmamış fakat hastalıkla karşılaştığında enfekte olma olasılığı olan kişiler.
- E (Exposed): Enfekte olmuş fakat henüz bulaşıcı olmayan ve semptomları olmayan kişiler.

- I (Infected): Virüs bulaşmış ve sonucunda enfekte olmuş ve hastalığı bulaştırma riski olan kişiler.
- R (Recovered): Virüs bulaşmış ve sonrasında iyileşmiş kişiler.

Bölmeler arasında geçişler vardır.

**S** Duyarlı bölmesi zamanla azalır. Bireyler E (maruz kalan) bölmesine geçiş yapar. Bölmeye geri akış yoktur.

**E** Bireyler bu bölmeye giriş-çıkış yapabilirler. Bireyler I bölmesine geçerek E bölmesini terk edebilir.

**I** Bu bölmeye hem giriş hem çıkış vardır. Bireyler bölmeyi iyileşerek veya ölerek terk eder.

**R** Bölmeye giriş vardır. Hiçbir birey bölmeden dışarı çıkamaz.




**Şekil 2.8.** SEIR modelinde bölmeler arası geçiş katsayıları

Model Parametreleri;

β : Popülasyondaki bulaşıcı bir bireyin birim zamanda yaptığı ortalama bulaşıcı temas sayısı. Yüksek β değeri virüsün yayılmak için daha fazla fırsatı olduğunu gösterir.

δ : Maruz kalan bireyin bulaşıcı hale gelme hızı, bulaşıcı duruma gelene kadar geçen süre.

γ : Bulaşıcı bireylerin iyileşme hızı.

Model Varsayımları;

- Popülasyon sabittir.
- Büyük popülasyonları tek tip kabul eder, bireylerin birbirleri ile etkileşime girme olasılığını hesaba katmaz.
- Model popülasyonun izole olduğunu varsayar.
- Bireylerin genellikle bağışıklıkla doğmadığını kabul eder.

Model Denklemleri;

$$N = S + E + I + R$$

$$S'(t) = -\frac{\beta SI}{N}$$

$$E'(t) = \frac{\beta SI}{N} - \delta E$$

$$I'(t) = \delta E - \gamma I$$

$$R'(t) = \gamma I$$

SEIR modeli için $R_0$ değeri;

$$R_0 = \frac{\beta\delta}{(\gamma + \mu)(\delta + \mu)}$$

### 2.3. Zaman Serisi Analizleri

Zaman serileri, zaman içerisinde gözlemlenen değerlerden elde edilen veri setledir. Verinin zaman serisi olarak adlandırılması için 4 farklı unsuru birlikte veya ayrı ayrı göstermesi gerekir. Bunlar trend, konjonktür dalgalanmalar, mevsimlik dalgalanmalar ve tesadüfi hareketlerdir. Veride gözlemlenen değerler zamana bağlıdır ve t zamanında gözlenen bir değer serinin önceki değerleri kullanılarak tahmin edilir. Zaman serisi verileri günlük, haftalık, aylık ve yıllık olabilir. Bu analizlerde, verilerin gelecekteki değerlerinin geçmiş değerlerinden bağımsız olamadığı ve birbirleriyle korelasyon içerisinde olduğu varsayılır. Zaman olarak birbirlerine yakın gözlemler, uzak olan gözlemlere göre daha benzer özellik gösterirler. Genel olarak zaman serisi analizlerinde geçmiş gözlem değerleri ve şimdiki gözlem değerleri kullanılarak, gelecekteki değerler tahmin edilmeye çalışılır. Bir zaman serisinin geleceğe yönelik tahmin yapabilmesi için trend, mevsim ve konjonktür dalgalanmalarının etkisinden arındırılmış olması gerekir. Bu etkilerden arındırılmış serilere durağan zaman serisi denir ve bu seriler geleceğe yönelik tahmin modellerinde kullanılabilir.

### 2.4. Box-Jenkins Yöntemi

Bu yöntem, zaman serisinin periyodik hareketlerinin zaman içerisinde tekrar etmesinden yararlanarak, geçmiş dönem gözlem değerleri ile gelecek dönem tahmini yapılmasını sağlar. Herhangi bir kısıtlama olmaksızın karmaşık zaman serileri içinde

uygulanmaktadır. Veride otokorelasyon var ise kullanılabilir. Box Jenkins metodu seri için en uygun modeli seçer ve seçilen modelin uygunluğunu kontrol eder.

Model Aşamaları;

- Uygun modelin belirlenmesi
- Modelin tahmini
- Denetim
- Öngörü şeklindedir.

Uygun modelin belirlenebilmesi için verinin durağan olması gerekir, yani serinin özelliklerinin zamandan etkilenmemesi ve bağımsız olması demektir. Seri trend ve mevsimsellik etkisinden arınmış olmalıdır. Trend, verideki uzun süreli iniş ya da çıkışlardır. Seri bunları barındırıyorsa durağan değil demektir. Bu durum serideki gözlemleri etkiler ve zamana bağlı olarak değişmesine neden olur. En uygun modeli belirlemek için ACF (otokorelasyon fonksiyonu) ve PACF (kısmi otokorelasyon fonksiyonu) grafiklerinden faydalanılır. ACF grafiğinde ilişki miktarları yavaş yavaş azalıyor ve PACF grafiğinde bu azalma hızlı oluyorsa AR modeli (otoregresyon), PACF grafiğinde ilişki miktarları yavaş yavaş azalıyor ve ACF grafiğinde bu azalma hızlı oluyorsa MA modeli (hareketli ortalama), hem ACF hem PACF grafiklerindeki ilişki miktarının azalışı yavaş yavaş oluyorsa ARMA (otoregresif hareketli ortalama) modelidir.

Modelin tahmini için, modeldeki katsayıların önemlilik testi sonuçlarına bakılması gerekir. Karar verilen model katsayılarından herhangi birinin test sonucu önemsiz çıkarsa bu model seriye uygun değildir.

Modelin denetlenmesi için tahmin ve gerçek değerlerin grafikleri çizilir ve birbirine olan uyumuna bakılır. Tüm bu şartlar sağlandıktan sonra en uygun model üzerinden öngörü değeri hesaplanır.

Box-Jenkins modelleri aşağıdaki gibidir.

1. Mevsimsel Olmayan Box-Jenkins Modelleri
   - Otoregresyon Modelleri (AR)
   - Hareketli Ortalama Modelleri (MA)

- Otoregresif Hareketli Ortalama Modelleri (ARMA/ARIMA)

2. Mevsimsel Box-Jenkins Modelleri
   - Mevsimsel Otoregresyon Modelleri (SAR)
   - Mevsimsel Hareketli Ortalama Modelleri (SMA)
   - Mevsimsel Otoregresif Hareketli Ortalama Modelleri (SARMA/SARIMA)

Modelin varsayımları,

- Modelin uygulanması için gerekli gözlem sayısı en az 50 civarında olmalıdır.
- Gecikmeler arasında otokorelasyon olmamalıdır, eğer var ise modelin kapsamadığı bilgiler var demektir.
- Gecikme ortalaması sıfır olmalıdır, eğer bu koşul sağlanmazsa tahminler yanlı demektir.
- Gecikmelerin varyansı sabit olmalıdır.
- Gecikmeler normal dağılmalıdır.
- İncelenen bağımlı değişken, bu değişken üzerindeki etkili olan diğer değişken veya değişkenler bağımsız kabul edilir.

### 2.4.1. Mevsimsel Olmayan Box-Jenkins Modelleri

Mevsimsel olmayan zaman serisi için p,d,q parametreleri belirlenmelidir. Seri durağan değilse fark alma işlemi yapılarak d parametresi belirlenir, ACF ve PACF grafiklerinin seyrine bakılarak da p ve q parametrelerinin dereceleri belirlenir. ACF (otokorelasyon fonksiyonu) MA (hareketli ortalama) kısmının derecesini yani q değerini, PACF (kısmı otokorelasyon fonksiyonu) AR (otoregresyon) kısmının derecesini yani p değerini gösterir.

#### 2.4.1.1. AR (Otoregresyon Modelleri)

AR(p) modeli için ACF grafiğindeki ilişkiler yavaş yavaş azalırken, PACF grafiğindeki ilişki bir anda azalmaktadır. PACF grafiğindeki ilk gecikmeye ait ilişkiler önemlidir. Böylece p'inci gecikmeden sonraki gecikmeler önemsizdir.

Model denklemi;

$$AR(p) \rightarrow \ yt = c + \emptyset 1 y_{t-1} + \emptyset 2 y_{t-2} + \cdots + \emptyset p y_{t-p} + \varepsilon t$$

$$\varepsilon t: Hata\ terimi$$

p=1 ise AR(1) birinci dereceden otoregreyon modeli, yani PACF grafiğindeki ilk gecikmeye ait ilişki miktarı istatistiksel olarak önemlidir.

p=2 ise AR(2) ikinci dereceden otoregreyon modeli, yani PACF grafiğindeki ilk iki gecikmeye ait ilişki miktarı istatistiksel olarak önemlidir.

#### 2.4.1.2. MA (Hareketli Ortalama Modeli)

MA(q) modeli için PACF grafiğindeki ilişkiler yavaş yavaş azalırken, ACF grafiğindeki ilişki bir anda azalmaktadır. ACF grafiğindeki ilk gecikmeye ait ilişkiler önemlidir. Böylece q'uncu gecikmeden sonraki gecikmeler önemsizdir.

Model denklemi;

$$MA(q) \rightarrow \ yt = c + \varepsilon t + \theta 1 \varepsilon_{t-1} + \theta 2 \varepsilon_{t-2} + \cdots + \theta q \varepsilon_{t-q}$$

q=1 ise MA(1) birinci dereceden hareketli ortalama modeli, yani ACF grafiğindeki ilk gecikmeye ait ilişki miktarı istatistiksel olarak önemlidir.

q=2 ise MA(2) ikinci dereceden hareketli ortalama modeli, yani ACF grafiğindeki ilk iki gecikmeye ait ilişki miktarı istatistiksel olarak önemlidir

#### 2.4.1.3. ARMA (Otoregresif Hareketli Ortalam Modeli)

ARMA(p,q) modeli için hem ACF hem de PACF grafiğindeki ilişkiler gecikme sayısı arttıkça yavaş yavaş azalmaktadır. Modelde p değeri PACF grafiğindeki gecikmelerden, q değeri ise ACF grafiğindeki gecikmelere ait önemli ilişkilerden belirlenir.

Model denklemi;

$$\left(yt - \emptyset 1 y_{t-1} - \cdots \emptyset p y_{t-p}\right) = (\varepsilon t - \theta 1 y_{t-1} - \cdots - \theta q y_{t-q})$$

p=1 ve q=1 olduğunda ARMA(1,1) modeli, hem ACF hem PACF grafiği için ilk gecikmeler önemlidir.

#### 2.4.1.4. ARIMA (Otoregresif Hareketli Entegre Ortalamalar Modeli)

ARIMA modeli zaman serilerini tahmin etmeye yönelik matematiksel yaklaşımlardan biridir. Box-Jenkins yöntemi olarak bilinir, George E. Box ve Gwilym M. Jenkins tarafından önerilmiştir. Eşit zaman aralıklarında gözlenen zaman serisi verilerine uygulanır. Verilerdeki korelasyonu modelleyerek bir zaman serisini en iyi temsil eden tahmin modeli ARIMA'dır.

Mevsimsel Olmayan ARIMA modeli;

- AR (Auto Regressive): Otoregresyon, p parametresi denklemdeki değişkenlerin gecikmelerini ifade eder.
- I (Integrated): Entegre, d parametresi modeli durağan yapmak için alınan fark işleminin derecesini gösterir.
- MA (Moving Average): Hareketli ortalama, q parametresi tahmin hata terimlerindeki gecikmeleri ifade eder.

Model denklemi;

$$ARIMA(p,d,q) \rightarrow yt' = c + \varphi 1 y'_{t-1} + \cdots + \varphi p y'_{t-p} + \theta 1 \varepsilon_{t-1} + \cdots + \theta q \varepsilon_{t-q} + \varepsilon$$

Mevsimsel zaman serisi aynı dönemlere ait gözlemlerin değerlerindeki artış veya azalış değişimlerini ifade etmektedir. Yılların, mevsimlerin ve ayların belirli zamanlarında aynı dalgalanmayı düzenli şekilde gösteren serilerde mevsim etkisi var demektir. Örneğin ayın belirli bir günü, yılın belirli bir ayı veya her yıl aynı zaman diliminde yapılan kutlama buna bir örnek olarak sayılabilir. Uygun ARIMA modeli için mevsimsel etkinin ortadan kaldırılması gerekir ve fark alma işlemi uygulanarak D parametresi belirlenir. Bu işlemden sonra yine ACF ve PACF grafiklerine bakılarak P ve Q parametrelerinin dereceleri belirlenir. Bu işlemlerden sonra modelin doğruluğunun kontrolü için istatistik önemlilik test sonuçlarına bakmak gerekir.

### 2.4.2. Mevsimsel Box-Jenkins Modelleri

#### 2.4.2.1. Mevsimsel SAR (Mevsimsel Otoregresyon) Modeli

Derecesi P, periyodu s olan SAR(P)s modelidir. PACF grafiğindeki önemli ilişkiler, ilk gecikmede değil, ilk periyoda ait gecikmelerdir. 1.dereceden farkı alınmış periyodu 12 olan otoregreson modeli ;

$$ARIMA(1,1,0)(1,1,0)_{12}$$

#### 2.4.2.2. Mevsimsel SMA (Mevsimsel Hareketli Ortalama) Modeli

Derecesi Q, periyodu s olan SMA(Q)s modelidir. ACF grafiğindeki önemli ilişkiler, ilk gecikmede değil, ilk periyoda ait gecikmelerdir. 1.dereceden farkı alınmış periyodu 12 olan otoregreson modeli ;

$$ARIMA(0,1,1)(0,1,1)_{12}$$

#### 2.4.2.3. Mevsimsel SARMA (Mevsimsel Otoregresif Hareketli Ortalama) Modeli

Hem ACF hem PACF grafiklerindeki gecikme ilişkileri yavaş yavaş azalır ve ACF grafiğinde periyodun yanındaki ilişkilerin yönünden Q parametresi, PACF grafiğinde periyodun yanındaki gecikmeden P parametresi belirlenir. 1.dereceden farkı alınmış periyodu 12 olan otoregresif hareketli ortalama modeli

$$ARIMA(1,0,1)(1,0,1)_{12}$$

#### 2.4.2.4. Mevsimsel SARIMA( Mevsimsel Otoregresif Entegre Hareketli Ortalama) Modeli

Periyodun yanındaki gecikmelere ait ilişkiler aynı büyüklükte ve aynı işaretli olmalıdır. Mevsimsel SARIMA modeli;

- SAR (Seasonal Auto Regressive): Mevsimsel otoregresyon, P parametresi denklemdeki değişkenlerin gecikmelerini ifade eder.
- I (Integrated): Entegre, D parametresi modeli durağan yapmak için alınan mevsimsel fark işleminin derecesini gösterir.
- SMA (Seasonal Moving Average): Mevsimsel hareketli ortalama, Q parametresi tahmin hata terimlerindeki gecikmeleri ifade eder.
- s mevsim periyodunu ifade eder.

1.dereceden farkı alınmış periyodu 12 olan otoregresif hareketli ortalama modeli;

$$ARIMA(1,1,1)(1,1,1)_{12}$$

# 3. GEREÇ ve YÖNTEM

Bu çalışmada SIR ve ARIMA modelleri kullanılarak Türkiye, Almanya, Amerika ve Çekya Covid-19 verileri analiz edilmiş ve modellerin vaka tahmini başarıları karşılaştırılmıştır. Her ülkenin araştırma için seçilme kriterleri farklıdır. Almanya'nın seçilme nedeni nüfusunun Türkiye'ye yakın olması ve bir Avrupa ülkesi olmasıdır. Amerika vaka sayısının fazlalığı ve ölüm oranının yüksek olması sebebiyle tercih edilmiştir. Çekya yapılan Benford[38] yasası uyumu araştırması sonucu en doğru vaka sayısı bildirimi yapan ülke olmasından dolayı seçilmiştir. Veri setleri ourworldindata sitesinde github.com/owid/covid-19 dan alınmıştır. Seçilen 4 ülke için paylaşılan Covid-19 veri setleri MS Excel programında düzenlenmiştir. Düzenlemelerde ülkelerin paylaştığı rakamlar doğrultusunda popülasyon, toplam vaka sayısı ve ölen kişi sayısı kullanılarak S,I ve R sayıları oluşturulmuştur. S ülke nüfusundan toplam vaka sayısı çıkarılarak elde edilmiştir. I toplam vaka sayısı olarak alınmıştır. R ise enfekte olduktan sonra 14.günde toplam vaka sayısından ölenlerin sayısı çıkarılarak elde edilmiştir. Bu hesaplamada T.C. Sağlık Bakanlığı Covid-19 rehberi referans alınmıştır [36]. Bunun sonucunda analiz için kullanılacak veri setleri elde edilmiştir. Covid-19 salgınında verilen bilgiler doğrultusunda SIR modeli için gerekli olan duyarlı, enfekte ve iyileşen kişi sayıları elde edilebilirken diğer kompartıman modeller için gerekli bilgiler doğru bir şekilde elde edilememektedir. Bu nedenle, çalışmada en temel model olan SIR modeli ile araştırma yapılmıştır.

Analizler için R programı ve bu program için yazılmış olan *forecast, ggplot2* gibi paketler kullanılmıştır. SIR modeli için R programında hazır bir paket bulunmaması nedeniyle bu analiz için birçok fonksiyon ve kodlar yazılmıştır (Ek1).

Sonuçların karşılaştırılabilir olması amacıyla, her ülke için salgının başlangıcından itibaren 550 gün boyunca sunulan veriler kullanılmıştır. Bu tezde SIR modeli, günlük yeni vaka değerlerini tahmin edebilmesi yönünde modifiye edilmiştir. Bu amaçla, veri düzenleme işlemi sonrasında I değeri belirtilen gündeki toplam vaka sayısını gösterdiği için belirtilen günde yeni enfekte olan vaka sayısı ardışık günlerdeki değerler arasındaki fark alınarak günlük yeni vaka sayısı olarak hesaplanmıştır. Analizlerde günlük vaka tahminlerinde kaymalar ve ötelenmeler (bugünün vaka sayısının olduğundan daha düşük ertesi gün vaka sayısını olduğundan daha yüksek tahmin etme) olabildiği

görüldüğünden haftalık toplam vakalar elde edilerek SIR modelleri hem günlük vaka hem de haftalık vaka tahminleri yönünden karşılaştırılmıştır.

SIR modeli için, beta ve gamma parametreleri farklı senaryolar baz alınarak günlük olarak hesaplattırılmıştır. Beta ve gamma değerleri belirli bir periyottaki veriler kullanılarak tahmin edilmektedir. Belirlenen periyot tahmin başarını etkileyebilmektedir. Bu nedenle, beta ve gamma tahminlerini daha az hatalı yapabilmek için ilgilenilen günden 3,5 ve 7 gün öncesindeki veriler kullanılmıştır.

Yapılan tahminler için değerlendirme ölçütü olarak seyir grafikleri (belirlenen tarih aralığındaki vaka sayıları ve tahminlerinin grafiği) ve artık (rezidü) değerleri kullanılmıştır. Artık değerlerinden hareketle hata terimleri kareler ortalamasının karekökü (RMSE) elde edilmiştir. Bunun için kullanılan formül şu şekildedir:

$$HKOK(RMSE) = \sqrt{\frac{\sum\left(GV_i - \widehat{GV}_i\right)^2}{n-p}}$$

Burada GV günlük vaka sayını, n tahmin yapılan gün sayısını, p ise tahmin işleminde kullanılan parametre sayısını (2 olarak alınmıştır) göstermektedir. Formül benzer şekilde haftalık vaka tahmini için de kullanılmıştır.

Tahmin edilmeye çalışılan değerlerin çok yüksek değerler olması nedeniyle bu değerlerdeki tahmin hataları da yüksek elde edilmektedir. Bu da toplam hatanın çok büyük olduğunu düşündürmektedir. Bu problemi aşabilmek için artık değerleri tahmin edilmeye çalışılan değere göre göreceli olarak hesaplanmış ve bu değerlerden hareketle RMSE değerleri tekrar elde edilmiştir. Göreceli artık değeri şu formülle elde edilmektedir.

$$\hat{e}_i^r = \frac{GV_i - \widehat{GV}_i}{GV_i}$$

$$GHKOK = \sqrt{\frac{\sum(\hat{e}_i^r)^2}{n-p}}$$

ARIMA modellerinde SIR modelinde olduğu gibi ülkelerin ilk 550 günlük değerleri kullanılmıştır. Bu modellemede modelin belirlenmesi, tahmin ve öngörülerin (forecasting) yapılabilmesi için verinin ilk 500 günlük değerleri kullanılmıştır. İlk 500 gün için tahminler ve sonraki 50 gün için öngörüler, günlük ve haftalık vaka değerleri için frekans değerleri 365 ve 52 alınarak yapılmıştır. Modellerde önsel bir bilginin

olmaması nedeniyle *auto.arima* fonksiyonu ile veriye uyan en iyi modelin belirlenmesi bilgisayara bırakılmıştır. Gerçek değerler ve tahmin değerlerinin olan uyumu için grafiklerden faydalanmıştır. Tahmin başarılarını kıyaslamak için yukarıda formülü verilen RMSE ve göreceli artıklar için RMSE değerlerine kullanılmıştır.

Tüm analizlerde elde edilen tahminler ve gerçek değerler seyir (trend) grafikleri olarak sunulmuştur. Bu grafiklerde vaka sayılarının pik ve dip yaptığı günler görülmektedir. Yapılan tahmin işlemlerinde bu değerlerde yöntemlerin tahmin başarıları şekillerin birbirine olan yakınlıkları üzerinden tartışılmıştır.

Türkiye verilerinde 2020 Aralık ayı içinde vaka tanımı değiştiği için (önceden semptomlu vakalar bildirilirken daha sonra test sonucu pozitif çıkan tüm vakalar bildirildi) günlük toplam vaka sayısında ciddi bir değişim olmuştur. Bu da modellerin tahmin başarısını etkilemektedir. Bu problemi aşabilmek için Türkiye verileri 3 parçaya ayrılmıştır: İlk 250 günkü veriler, 250-274 gün arası değerler ve 276-550 arası değerler. Burada vaka sayısı tanımının değişim gösterdiği gün 275.güne (bir önceki gündeki yeni vaka sayısı 31712, 275.günde yeni vaka sayısı 823325 ve bir sonraki gündeki yeni vaka sayısı 32106) denk geldiği için bu ayrım yapılmıştır. Analizlerde bu gündeki vaka sayısı dikkate alınmamıştır. Üç parçadan ayrı ayrı elde edilen toplam hata kareleri ve göreceli hata kareleri değerleri toplandıktan sonra tüm Türkiye verisi için hata kareleri ortalamaları hesaplanmıştır.

# 5. BULGULAR

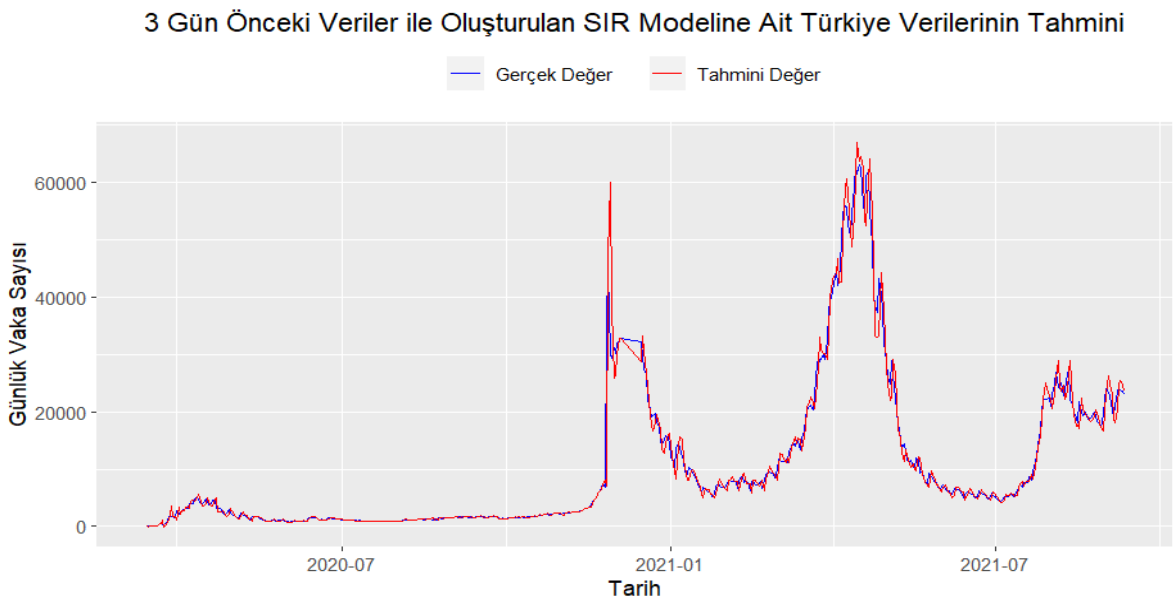
Tez çalışmasında SIR ve ARIMA modelleri, Our World in data (http://ourworldindata.org) web sitesinden elde edilen Türkiye, Almanya, Amerika Birleşik Devletleri ve Çekya Covid19 verileri kullanılarak karşılaştırılmıştır. Bulgular, SIR modellerinin ülkelere göre sonuçları, ARIMA modellerinin ülkelere sonuçları ve iki yaklaşımın genel karşılaştırması olarak 3 ana bölümden oluşmaktadır.

## 5.1. SIR Yaklaşımı ile Yapılan Tahminler

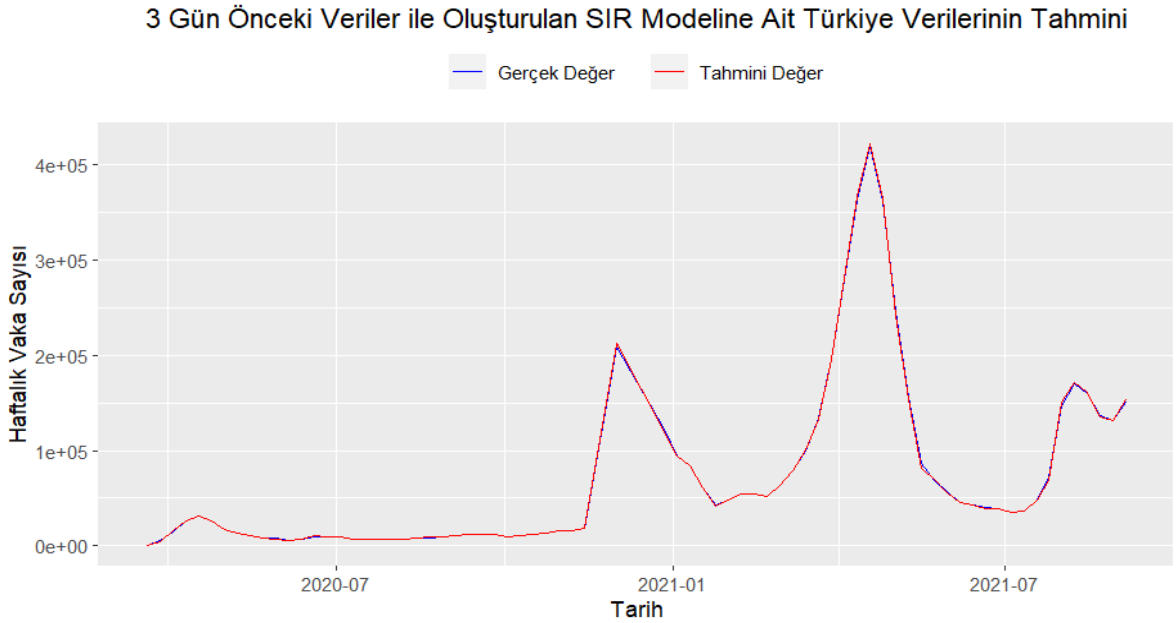
SIR yaklaşımında bir sonraki güne ait toplam risk altındaki kişi sayısı (S), toplam enfekte kişi sayısı (I) ve toplam iyileşen kişi sayısı (R) hesaplanmaktadır. Bu çalışmada karşılaştırılabilir olması adına iki ardışık günün toplam enfekte kişi sayıları arasındaki fark alınarak günlük enfekte kişi sayısı ve bu farkın bir hafta arayla iki toplam arasında hesaplanmasıyla haftalık toplam enfekte kişi sayısı hesaplanmıştır. SIR modelinin tahmin performansını etkileyebileceği düşüncesiyle tahmin edilmeye çalışılan günden 3, 5 ve 7 gün önceki değerlerin kullanarak tahminler yapılmıştır. Sonuçlar ülkeler içinde ayrı ayrı sunulmuştur.

### 5.1.1. Türkiye Verilerinin SIR Yaklaşımı ile Tahmini

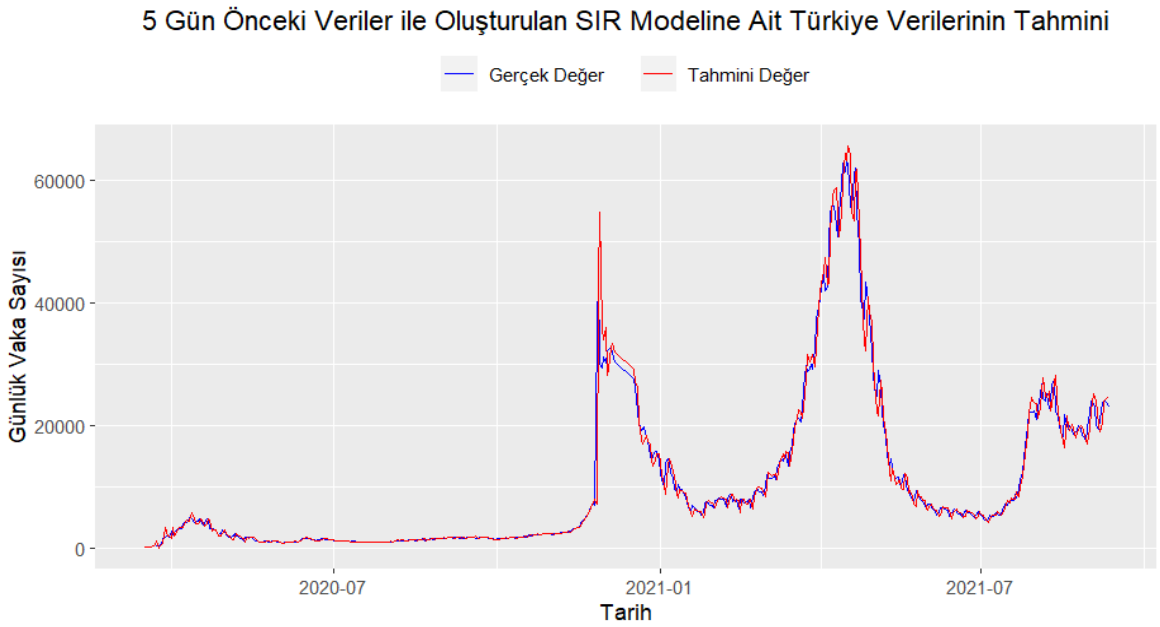
Türkiye için salgının başlangıcından itibaren ilk vakanın görüldüğü andan başlayarak ilk 550 gün için 3, 5 ve 7 gün önceki değerler kullanılarak yapılan günlük vaka tahminleri ve haftalık vaka tahminleri Şekil 4.1 – Şekil 4.6 arasında sırasıyla sunulmuştur. Görüldüğü üzere yapılan tahminler gerçek değere çok yakındır.



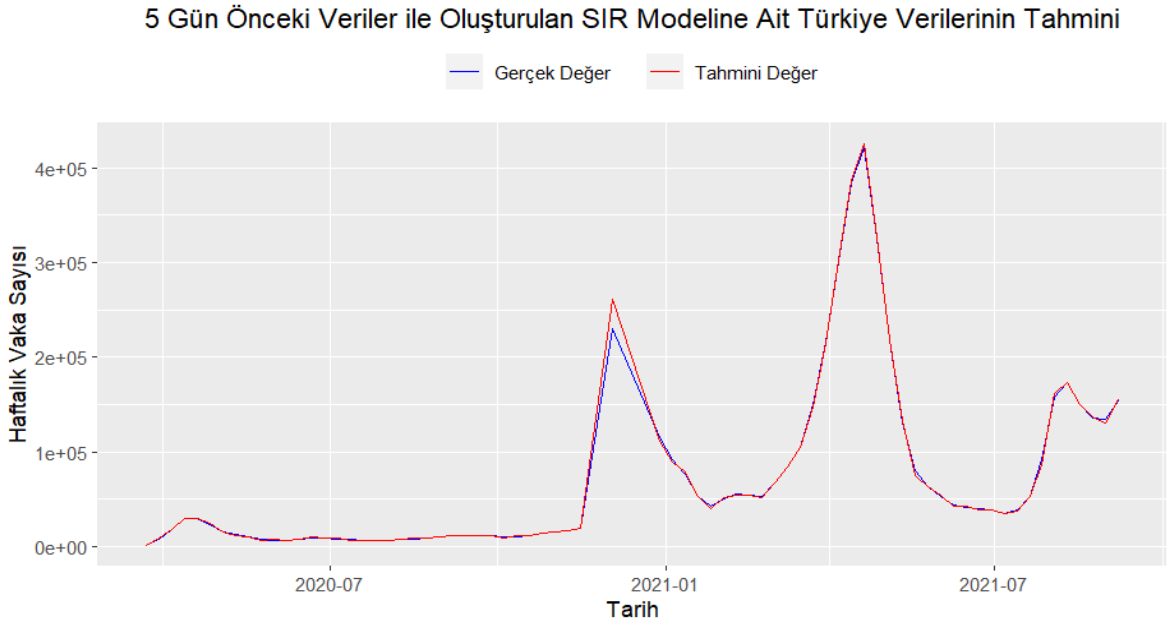

**Şekil 4.1.** 3 gün önceki veriler ile oluşturulan SIR modeline ait Türkiye verilerinin günlük vaka tahmini



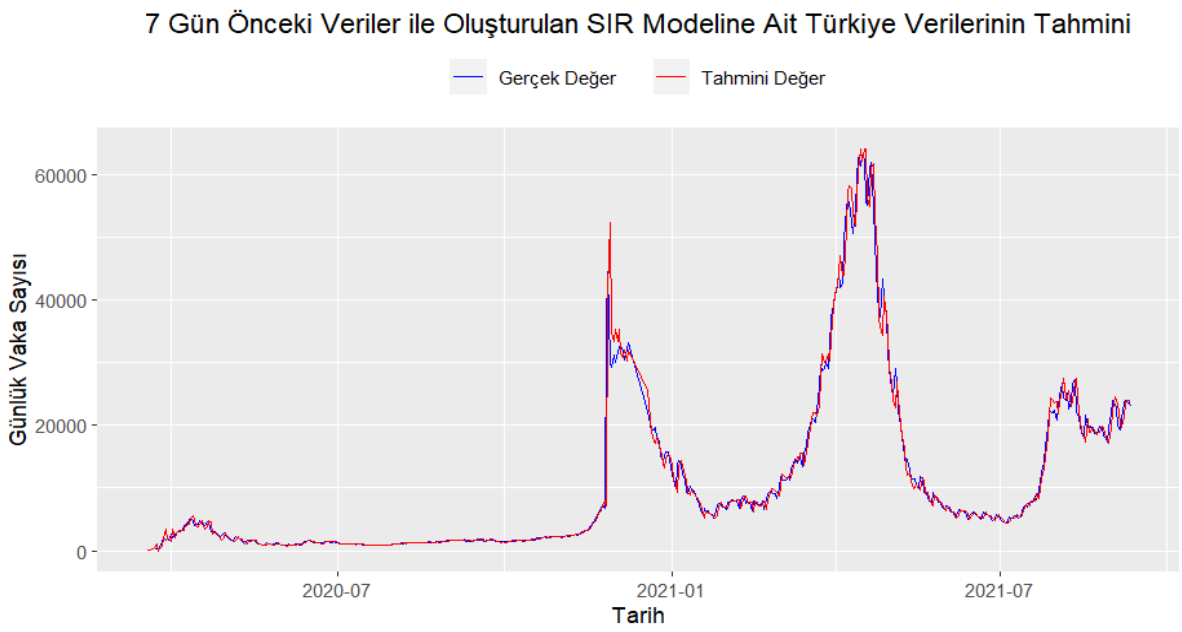

**Şekil 4.2.** 3 gün önceki veriler ile oluşturulan SIR modeline ait Türkiye verilerinin haftalık vaka tahmini



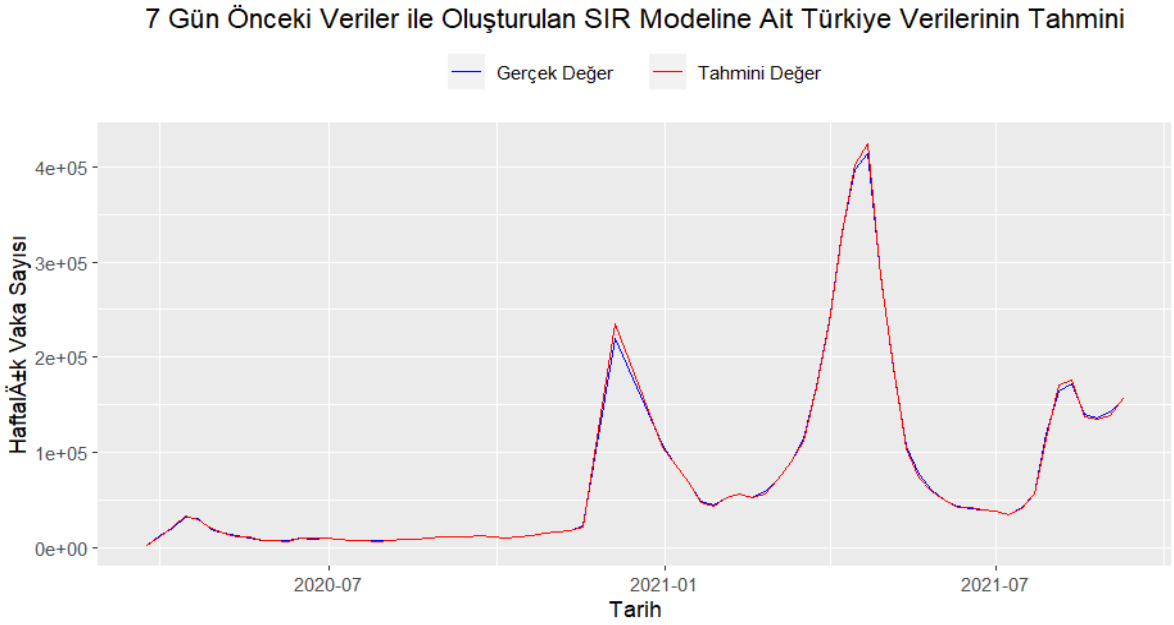

**Şekil 4.3.** 5 gün önceki veriler ile oluşturulan SIR modeline ait Türkiye verilerinin günlük vaka tahmini




**Şekil 4.4**. 5 gün önceki veriler ile oluşturulan SIR modeline ait Türkiye verilerinin haftalık vaka tahmini




**Şekil 4.5.** 7 gün önceki veriler ile oluşturulan SIR modeline ait Türkiye verilerinin günlük vaka tahmini




**Şekil 4.6.** 7 gün önceki veriler ile oluşturulan SIR modeline ait Türkiye verilerinin haftalık vaka tahmini

Çizelge 4.1. Türkiye verilerinin 3, 5 ve 7 gün önceki veriler kullanılarak yapılan günlük ve haftalık vaka tahminlerinin hataları

| | Günlük vaka tahminleri | | Haftalık vaka tahminleri | |
|---|---|---|---|---|
| | Hata Kareleri Ortalaması (RMSE) | Göreceli Hata Kareleri Ortalaması | Hata Kareleri Ortalaması (RMSE) | Göreceli Hata Kareleri Ortalaması |
| 3 gün öncesi | 2442.0 | 48.8 | 1451.5 | 2.34 |
| 5 gün öncesi | 2266.0 | 18.0 | 5176.0 | 5.68 |
| 7 gün öncesi | 2082.9 | 17.1 | 2880.3 | 3.06 |

Çizelge 4.1.'de verilen SIR yöntemi ile 3, 5 ve 7 gün önceki verileri kullanarak yapılan tahminlerin hata payları incelendiğinde hata kareleri ortalaması yönünden en iyi tahminlerin; günlük vaka tahminlerinde 7 gün önceki değerlerin kullanılmasının, haftalık vaka tahminlerinde ise 3 gün önceki değerlerin kullanılmasının gerektiği anlaşılmaktadır. Öte yandan hata kareleri hatanın yapıldığı değerin çok yüksek olması halinde yüksek değerler olabilmekte dolayısıyla bu değerler hata karelerin ortalamasını yükselmektedir. Bu sapmayı azaltmak için göreceli hata kareleri elde edilmiştir. Bu sonuçlar incelendiğinde de aynı sonuçların elde edildiği görülmüştür.

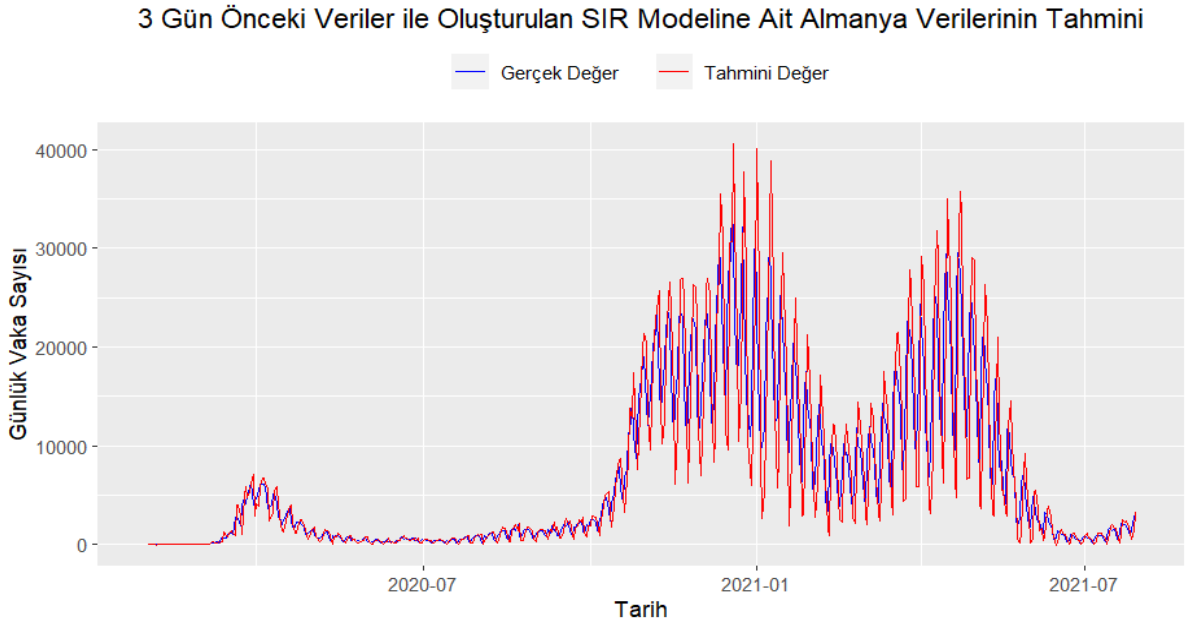
Gereç ve yöntemde de belirtildiği üzere Türkiye verilerinde 275.günde gözlenen aşırı yükseliş nedeniyle üç periyoda ayrılan Türkiye verisi, periyotlar içinde de ayrı ayrı incelenmiş ve elde edilen RMSE değerleri Çizelge 4.2'de sunulmuştur. Bu sonuçlara göre, hata kareleri ortalaması yönünden en iyi tahminlerin; günlük vaka tahminlerinde ilk 250 gün ve 276 gün sonrası günler için tahminlerde 5 gün öncesi değerlerin kullanılması halinde hatanın azaldığı görülmektedir. Benzer şekilde, 251-274 arası günler için ise 7 gün önceki değerlerin daha az hata yaptığı görülmüştür. Haftalık vaka tahminlerinde ise RMSE değerlerine göre tüm periyotlarda 3 gün önceki değerlere bağlı tahminlerin hatayı azalttığı görülmektedir. Göreceli RMSE değerlerine bakıldığında ise, günlük vaka tahminlerinde tüm aralıklarda 7 gün önceki değerlerin kullanılması, haftalık vaka tahminlerinde ise tüm aralıklarda 3 gün önceki değerlerin kullanılması gerektiği görülmektedir.

Çizelge 4.2.Türkiye verilerinin farklı periyotlardaki 3, 5 ve 7 gün önceki veriler kullanılarak yapılan günlük ve haftalık vaka tahminlerinin hataları

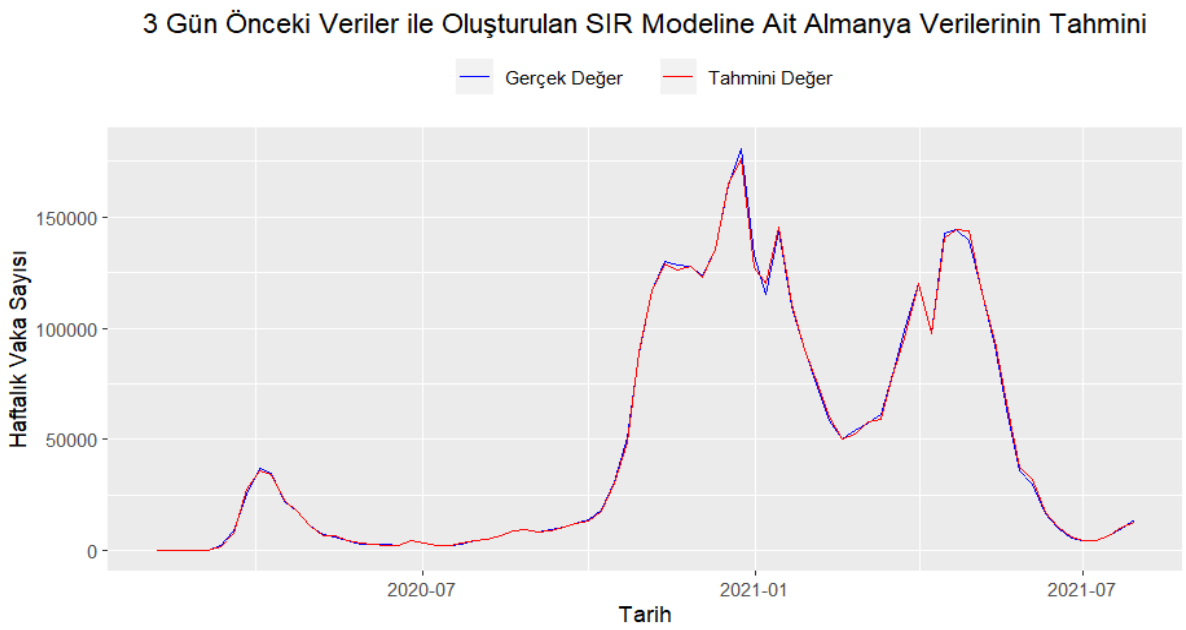
| | | Hata Kareleri Ortalaması (RMSE) | | | Göreceli Hata Kareleri Ortalaması | | |
|---|---|---|---|---|---|---|---|
| | | 3 gün öncesi | 5 gün öncesi | 7 gün öncesi | 3 gün öncesi | 5 gün öncesi | 7 gün öncesi |
| İlk 250 gün | Günlük vaka tahminleri | 312.8 | 274.5 | 277.9 | 70.7 | 23.1 | 22.3 |
| | Haftalık vaka tahminleri | 283.8 | 295.5 | 446.2 | 3.05 | 3.21 | 3.69 |
| 251-274 gün arası | Günlük vaka tahminleri | 11136.9 | 10197.8 | 9747.0 | 34.7 | 31.3 | 29.6 |
| | Haftalık vaka tahminleri | 3068.4 | 30006.8 | 12337.5 | 1.70 | 30.6 | 6.71 |
| 276 gün sonrası | Günlük vaka tahminleri | 2211.8 | 2091.5 | 1814.3 | 10.5 | 10.2 | 8.83 |
| | Haftalık vaka tahminleri | 1905.9 | 2198.6 | 2861.2 | 1.58 | 2.17 | 2.04 |

### 5.1.2. Almanya Verilerinin SIR Yaklaşımı ile Tahmini

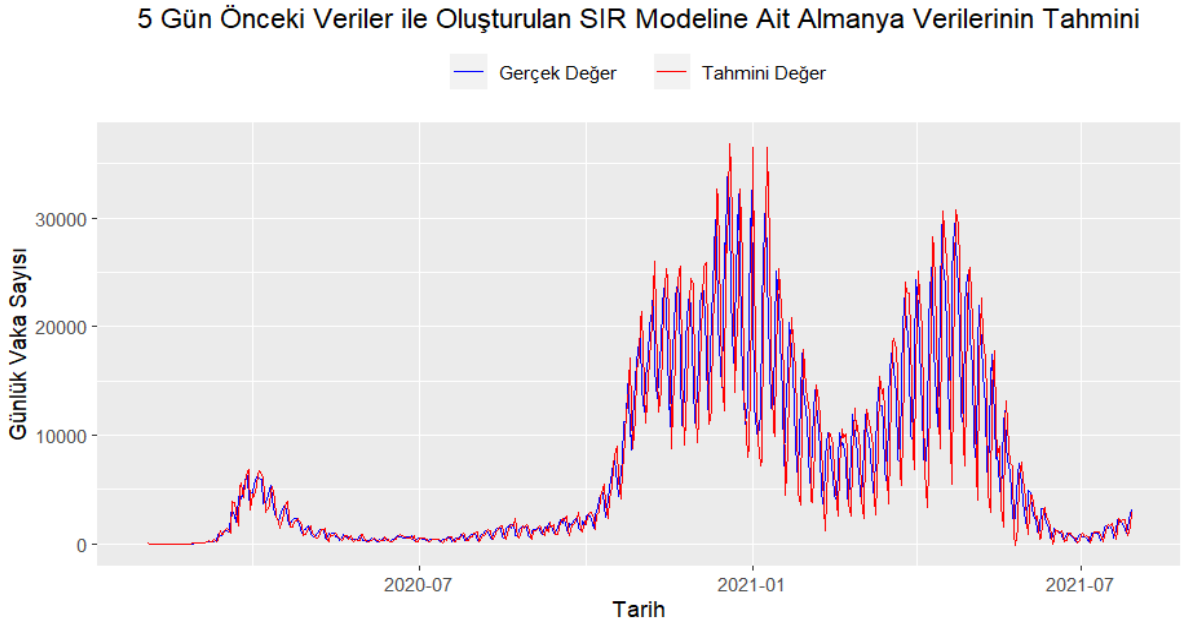
Almanya için salgının başlangıcından itibaren ilk vakanın görüldüğü andan başlayarak ilk 550 gün için 3, 5 ve 7 gün önceki değerler kullanılarak yapılan günlük vaka tahminleri ve haftalık vaka tahminleri Şekil 4.7 – Şekil 4.12 arasında sırasıyla sunulmuştur. Görüldüğü üzere yapılan tahminlerde sapmalar olsa bile özellikle haftalık tahminlerde gerçek değere çok yakındır.



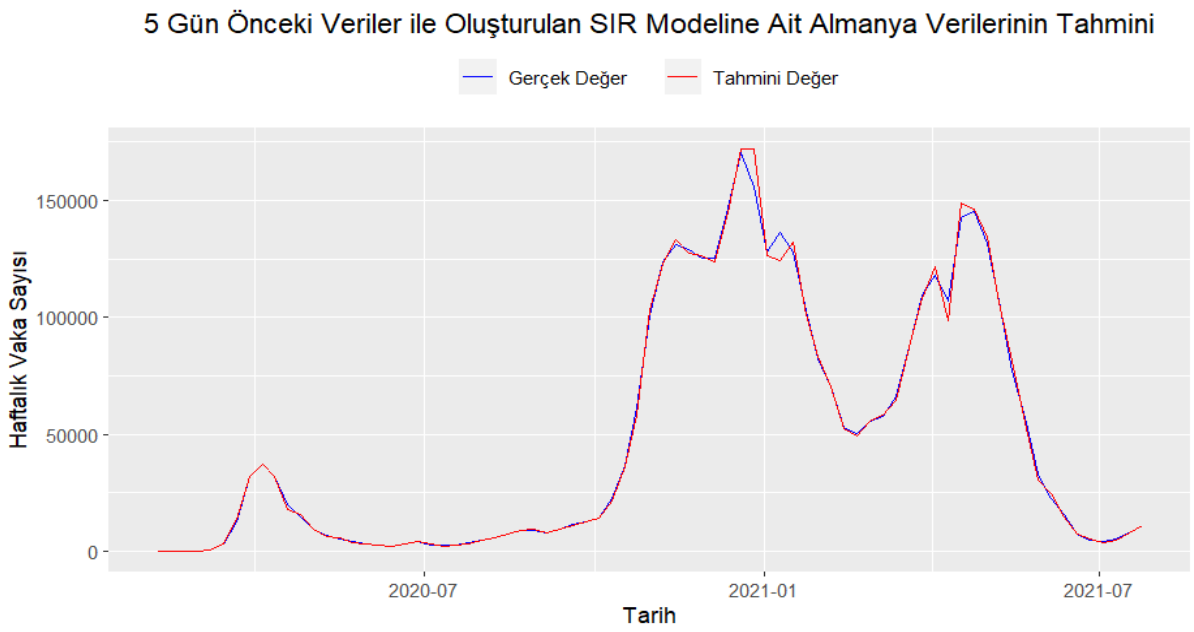

**Şekil 4.7.** 3 gün önceki veriler ile oluşturulan SIR modeline ait Almanya verilerinin günlük vaka tahmini



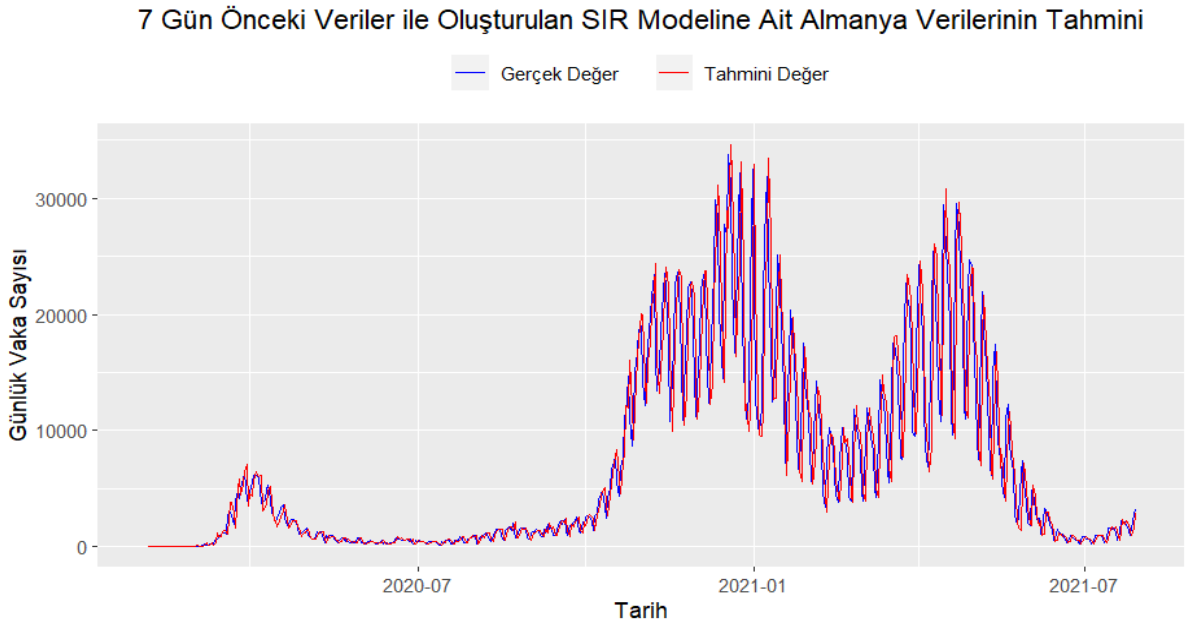

**Şekil 4.8.** 3 gün önceki veriler ile oluşturulan SIR modeline ait Almanya verilerinin haftalık vaka tahmini



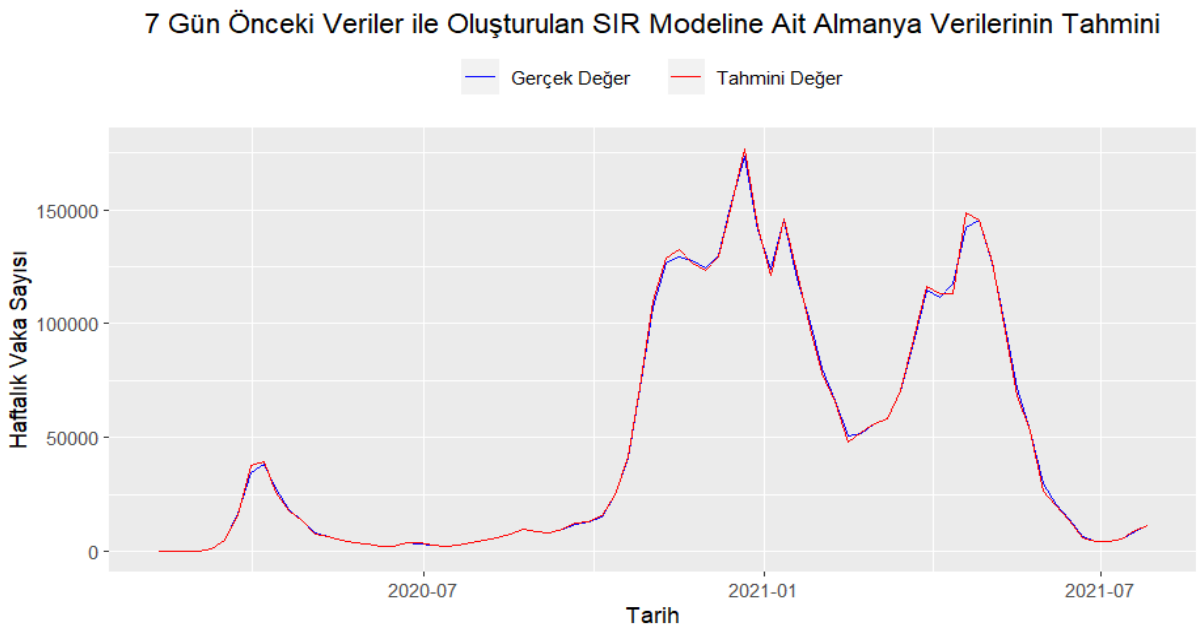

**Şekil 4.9.** 5 gün önceki veriler ile oluşturulan SIR modeline ait Almanya verilerinin günlük vaka tahmini




**Şekil 4.10.** 5 gün önceki veriler ile oluşturulan SIR modeline ait Almanya verilerinin haftalık vaka tahmini




**Şekil 4.11.** 7 gün önceki veriler ile oluşturulan SIR modeline ait Almanya verilerinin günlük vaka tahmini




**Şekil 4.12.** 7 gün önceki veriler ile oluşturulan SIR modeline ait Almanya verilerinin haftalık vaka tahmini

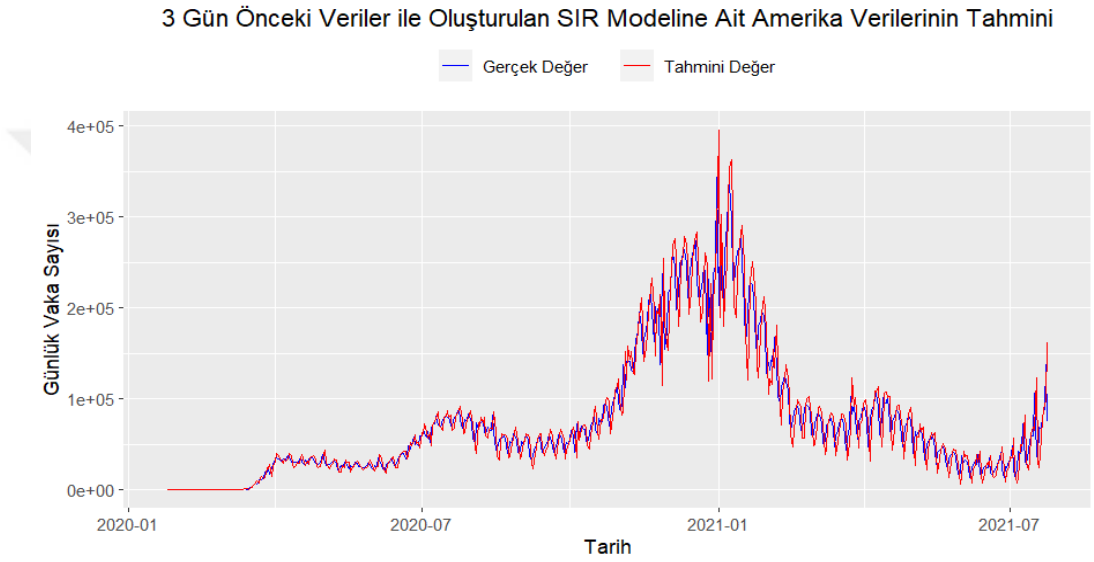
Çizelge 4.3.Almanya verilerinin 3, 5 ve 7 gün önceki veriler kullanılarak yapılan günlük ve haftalık vaka tahminlerinin hataları

| | Günlük vaka tahminleri | | Haftalık vaka tahminleri | |
|---|---|---|---|---|
| | Hata Kareleri Ortalaması (RMSE) | Göreceli Hata Kareleri Toplamı | Hata Kareleri Ortalaması (RMSE) | Göreceli Hata Kareleri Toplamı |
| 3 gün öncesi | 3984.8 | 53.9 | 1514.6 | 5.79 |
| 5 gün öncesi | 3831.6 | 55.4 | 2948.9 | 5.39 |
| 7 gün öncesi | 3068.4 | 47.3 | 1534.8 | 14.1 |

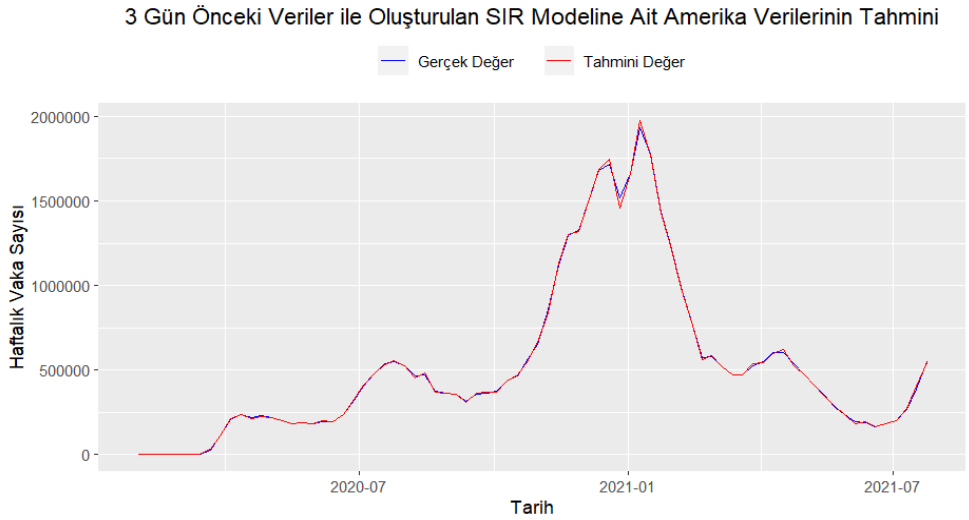
Çizelge 4.3.'de verilen SIR yöntemi ile 3, 5 ve 7 gün önceki verileri kullanarak yapılan tahminlerin hata payları incelendiğinde hata kareleri ortalaması yönünden en iyi tahminlerin; günlük vaka tahminlerinde 7 gün önceki değerlerin kullanılmasının, haftalık vaka tahminlerinde ise 3 gün önceki değerlerin kullanılmasının gerektiği anlaşılmaktadır. Ancak göreceli hata kareleri incelendiğinde ise, en iyi tahminlerin; günlük vaka tahminlerinde yine 7 gün önceki değerlerin kullanılmasının, haftalık vaka tahminlerinde ise bu sefer 5 gün önceki değerlerin kullanılmasının gerektiği görülmüştür.

### 5.1.3. Amerika Verilerinin SIR Yaklaşımı ile Tahmini

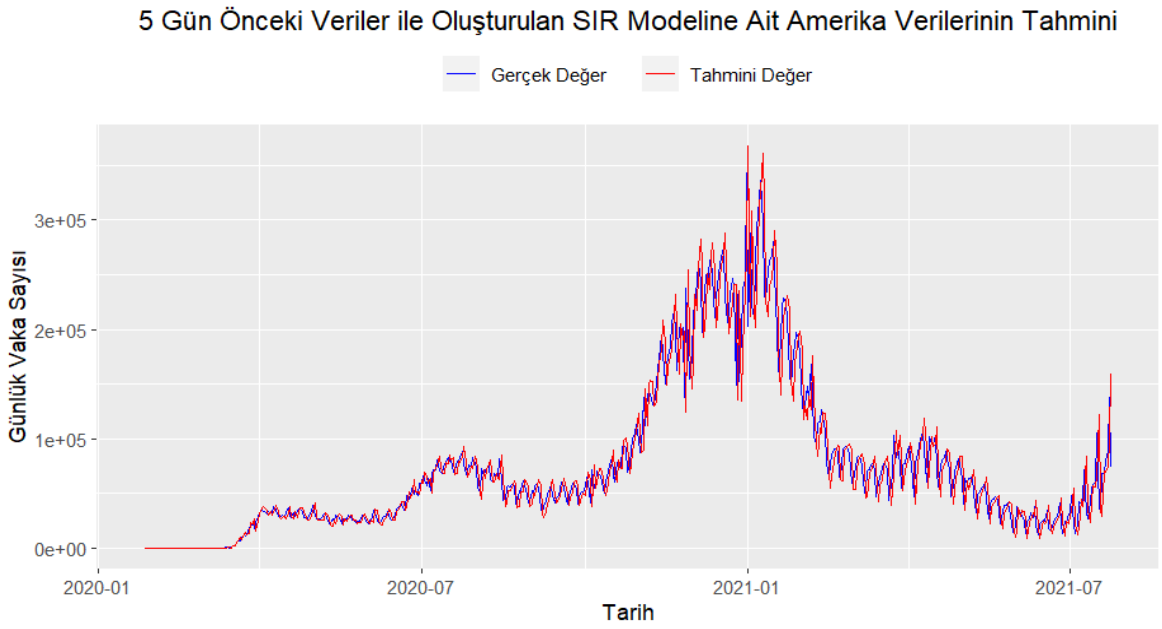
Amerika için salgının başlangıcından itibaren ilk vakanın görüldüğü andan başlayarak ilk 550 gün için 3, 5 ve 7 gün önceki değerler kullanılarak yapılan günlük vaka tahminleri ve haftalık vaka tahminleri Şekil 4.13 – Şekil 4.18 arasında sırasıyla sunulmuştur. Görüldüğü üzere günlük vaka tahminlerinde sapmalar özellikle 3 gün önceki verilerde daha yüksek olsa da 7 gün önceki veriler ile yapılan tahminler gerçek değerlere oldukça yakındır.



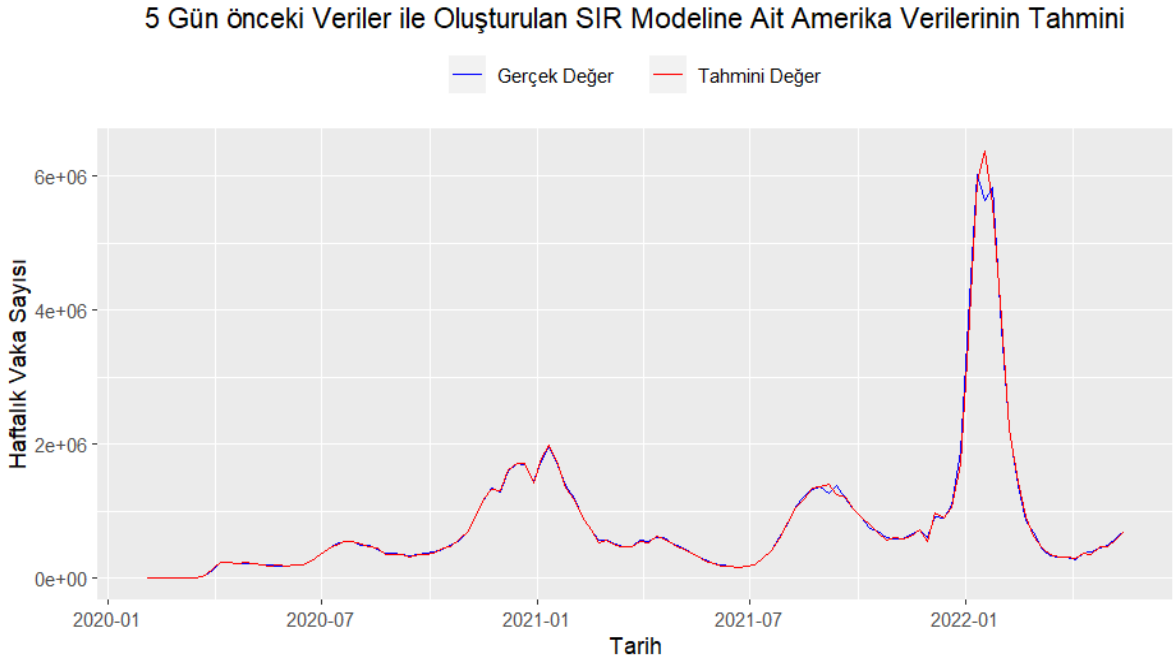

**Şekil 4.13.** 3 gün önceki veriler ile oluşturulan SIR modeline ait Amerika verilerinin günlük vaka tahmini



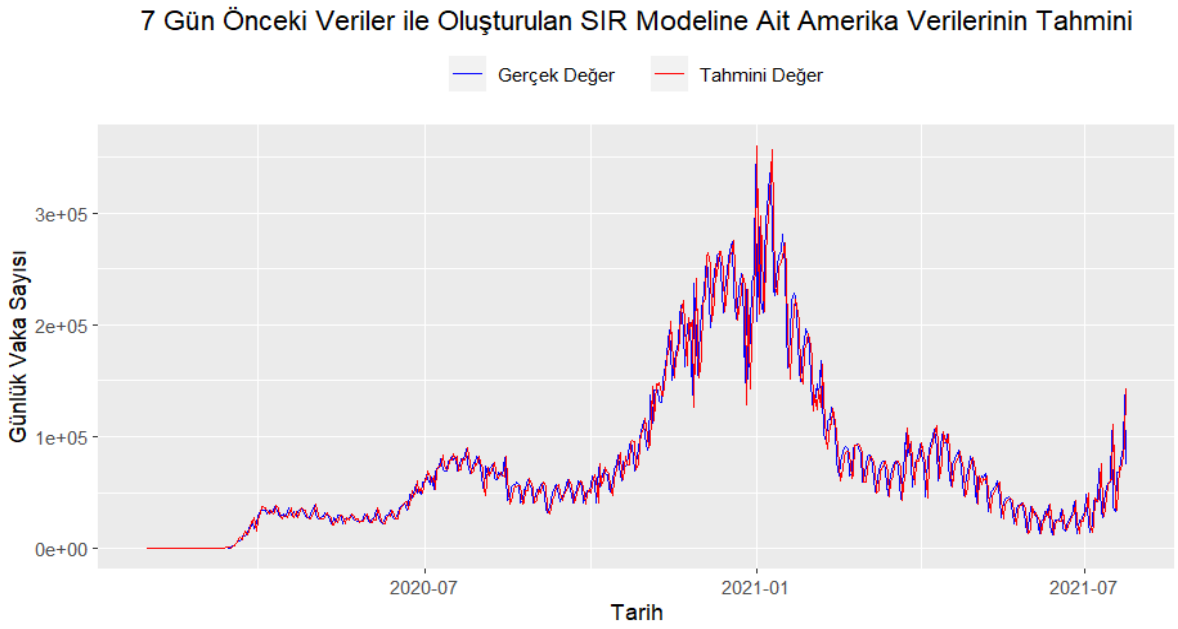

**Şekil 4.14.** 3 gün önceki veriler ile oluşturulan SIR modeline ait Amerika verilerinin haftalık vaka tahmini



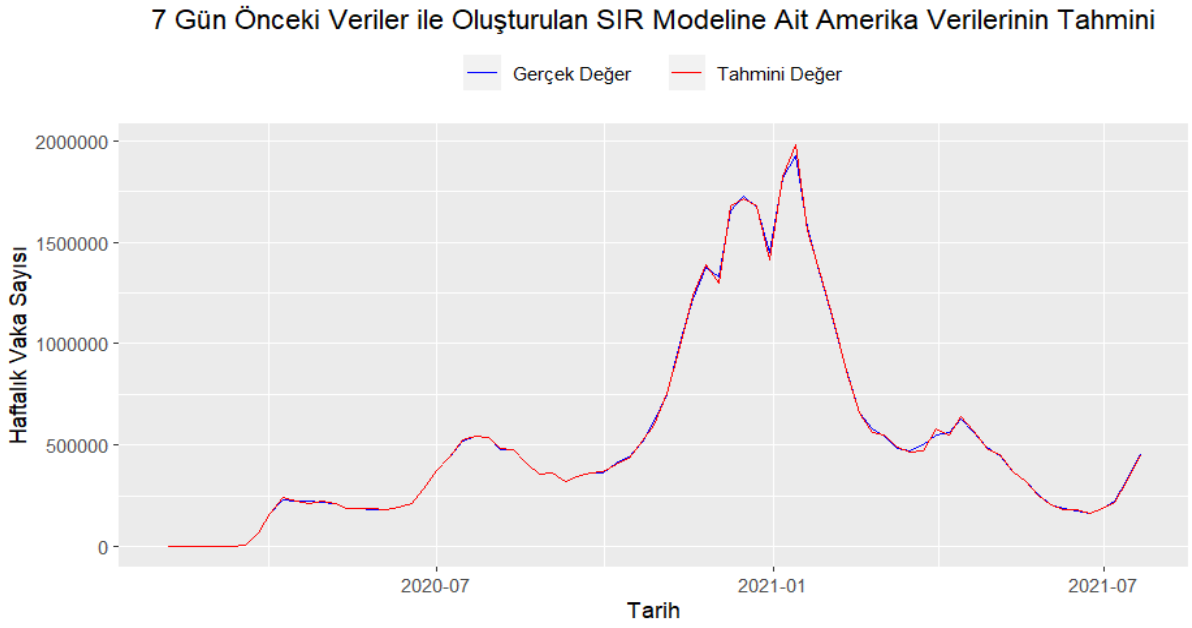

**Şekil 4.15.** 5 gün önceki veriler ile oluşturulan SIR modeline ait Amerika verilerinin günlük vaka tahmini




**Şekil 4.16.** 5 gün önceki veriler ile oluşturulan SIR modeline ait Amerika verilerinin haftalık vaka tahmini




**Şekil 4.17.** 7 gün önceki veriler ile oluşturulan SIR modeline ait Amerika verilerinin günlük vaka tahmini




**Şekil 4.18.** 7 gün önceki veriler ile oluşturulan SIR modeline ait Amerika verilerinin haftalık vaka tahmini

**Çizelge 4.4.**Amerika verilerinin 3, 5 ve 7 gün önceki veriler kullanılarak yapılan günlük ve haftalık vaka tahminlerinin hataları

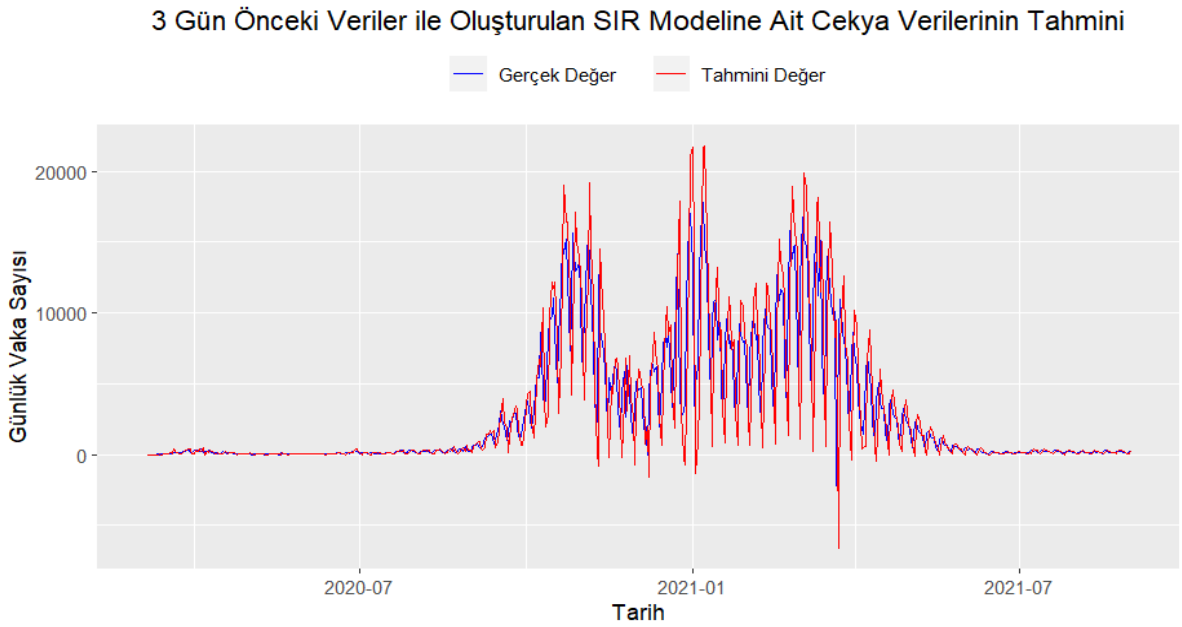
| | Günlük vaka tahminleri | | Haftalık vaka tahminleri | |
|---|---|---|---|---|
| | Hata Kareleri Ortalaması (RMSE) | Göreceli Hata Kareleri Ortalaması | Hata Kareleri Ortalaması (RMSE) | Göreceli Hata Kareleri Ortalaması |
| 3 gün öncesi | 22899.2 | 44.7 | 11320.5 | 11.4 |
| 5 gün öncesi | 21693.5 | 41.3 | 11883.9 | 8.91 |
| 7 gün öncesi | 19486.5 | 38.2 | 11862.2 | 5.24 |

Çizelge 4.4.'de verilen SIR yöntemi ile 3, 5 ve 7 gün önceki verileri kullanarak yapılan tahminlerin hata payları incelendiğinde hata kareleri ortalaması yönünden en iyi tahminlerin; günlük vaka tahminlerinde 7 gün önceki değerlerin kullanılmasının, haftalık vaka tahminlerinde ise 3 gün önceki değerlerin kullanılmasının gerektiği anlaşılmaktadır. Ancak göreceli hata kareleri incelendiğinde, en iyi tahminlerin; günlük vaka tahminlerinde yine 7 gün önceki değerlerin kullanılmasının, haftalık vaka tahminlerinde de 7 gün önceki değerlerin kullanılmasının gerektiği görülmüştür

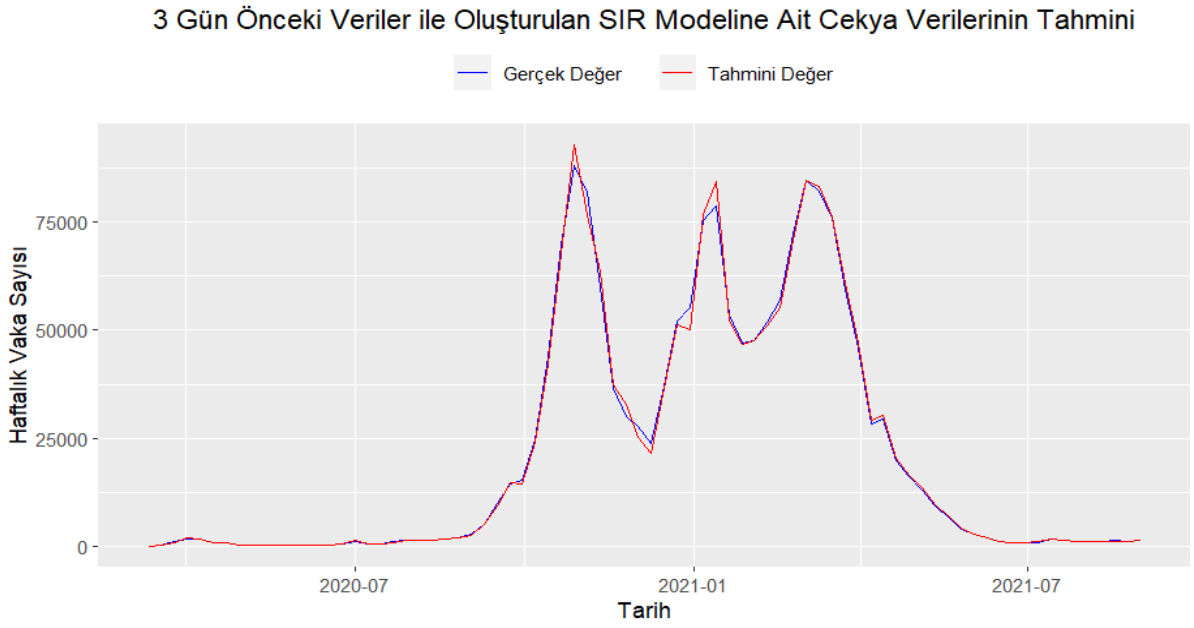
### 5.1.4. Çekya Verilerinin SIR Yaklaşımı ile Tahmini

Çekya için salgının başlangıcından itibaren ilk vakanın görüldüğü andan başlayarak ilk 550 gün için 3, 5 ve 7 gün önceki değerler kullanılarak yapılan günlük vaka tahminleri ve haftalık vaka tahminleri sırasıyla Şekil 4.19 – Şekil 4.24'te sunulmuştur. Görüldüğü üzere yapılan tahminler gerçek değerlere çok yakındır.

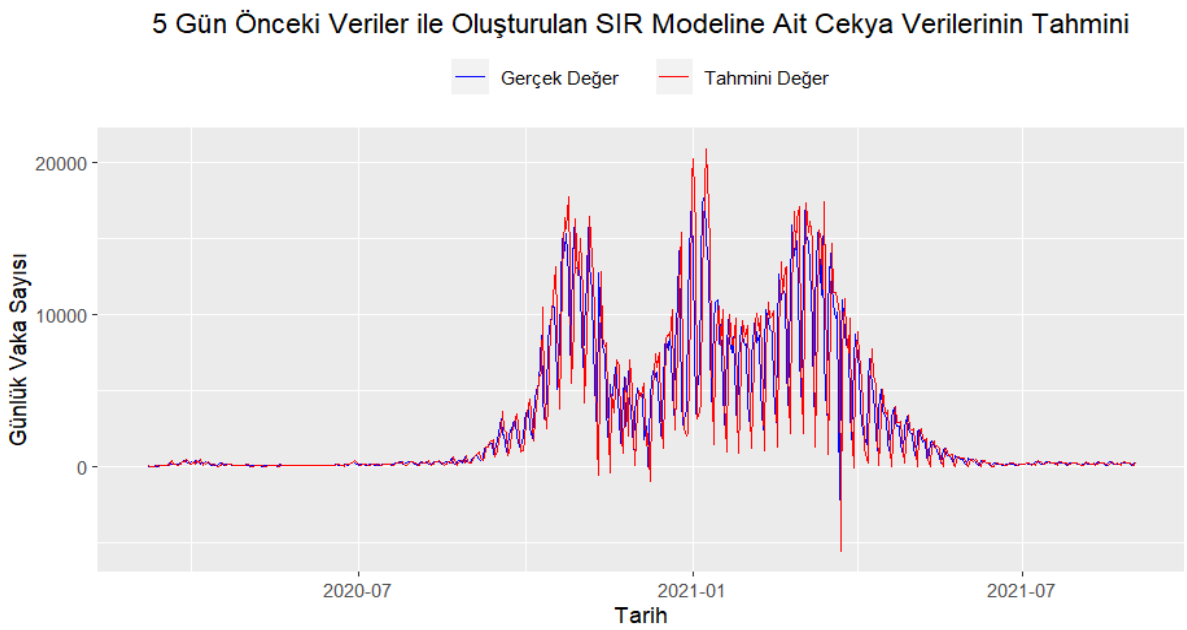
Çekya verilerinde 22 Mart 2021 tarihinde ardışık iki gün farkı ile hesaplanan günlük vaka değeri negatif bir değer (-2214) olarak elde edilmiştir. Bu değer veriden dışlanmadan tahminler yapıldığı için Mart 2021 tarihinde gerçek değer ve tahmin değerinin negatif değerler olarak grafikte gösterilmiştir. SIR modeli, günlük vaka grafiklerinden de görüldüğü üzere Mart 2021 tarihindeki değerin negatif olmasına rağmen bu değeri negatif olarak tahmin edebilmiştir.



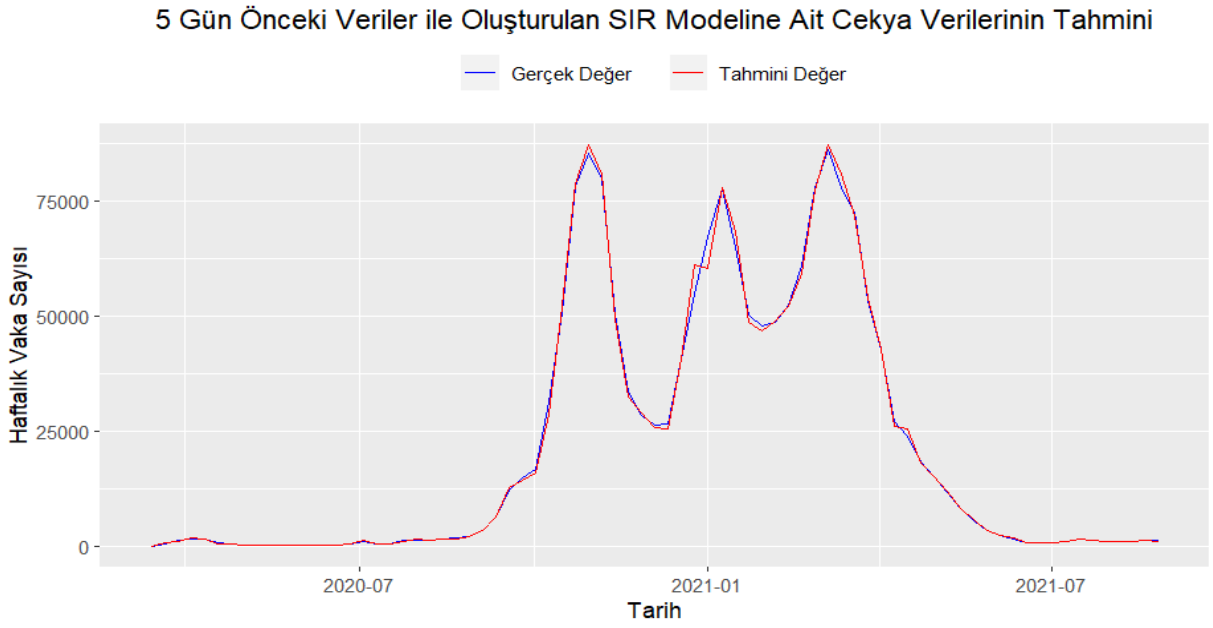

**Şekil 4.19.** 3 gün önceki veriler ile oluşturulan SIR modeline ait Çekya verilerinin günlük vaka tahmini



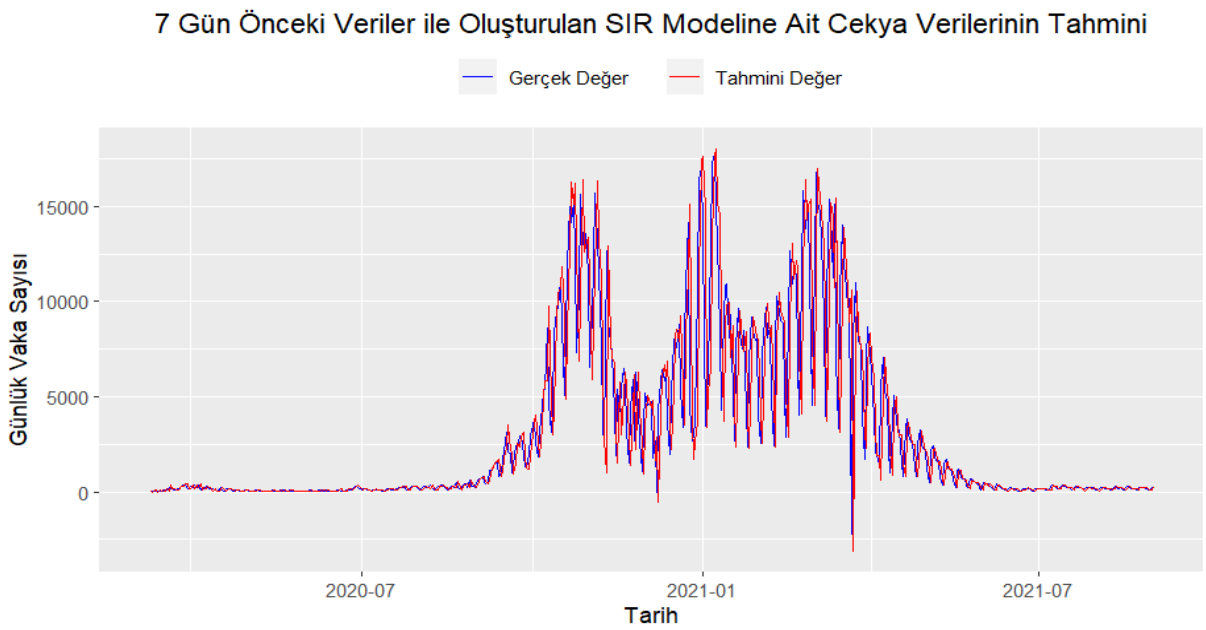

**Şekil 4.20.** 3 gün önceki veriler ile oluşturulan SIR modeline ait Çekya verilerinin haftalık vaka tahmini



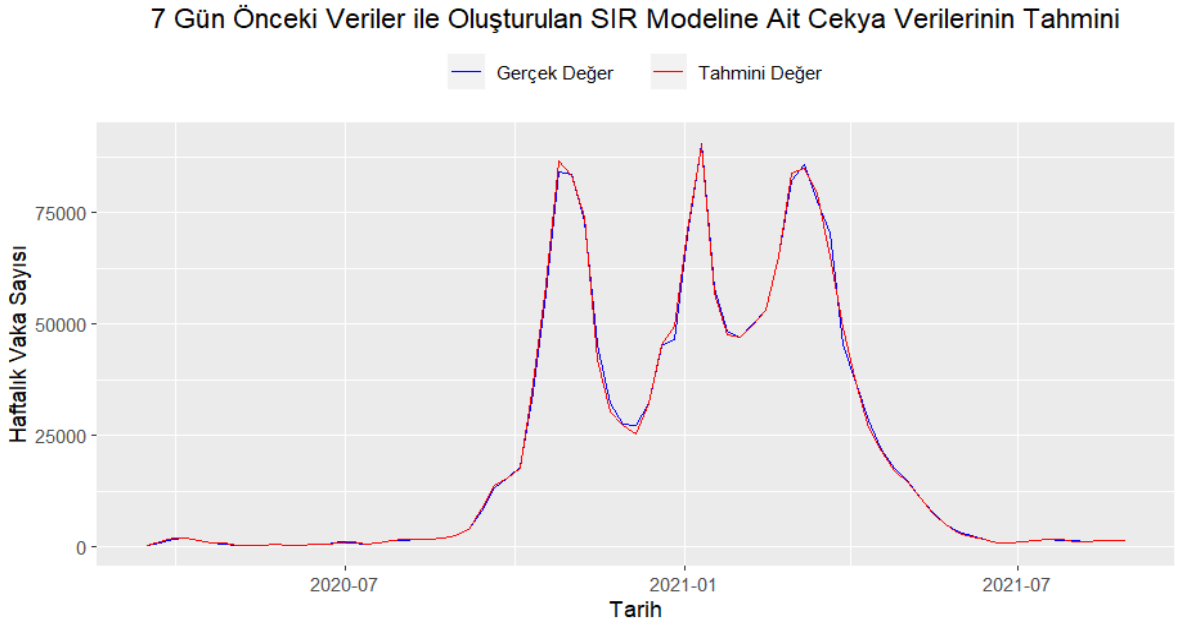

**Şekil 4.21.** 5 gün önceki veriler ile oluşturulan SIR modeline ait Çekya verilerinin günlük vaka tahmini




**Şekil 4.22.** 5 gün önceki veriler ile oluşturulan SIR modeline ait Çekya verilerinin haftalık vaka tahmini




**Şekil 4.23.** 7 gün önceki veriler ile oluşturulan SIR modeline ait Çekya verilerinin günlük vaka tahmini




**Şekil 4.24.** 7 gün önceki veriler ile oluşturulan SIR modeline ait Çekya verilerinin haftalık vaka tahmini

Çizelge 4.5. Çekya verilerinin 3, 5 ve 7 gün önceki veriler kullanılarak yapılan günlük ve haftalık vaka tahminlerinin hataları

| | Günlük vaka tahminleri | | Haftalık vaka tahminleri | |
|---|---|---|---|---|
| | Hata Kareleri Ortalaması (RMSE) | Göreceli Hata Kareleri Ortalaması | Hata Kareleri Ortalaması (RMSE) | Göreceli Hata Kareleri Ortalaması |
| 3 gün öncesi | 2619.2 | 157.5 | 1485.2 | 8.94 |
| 5 gün öncesi | 2537.7 | 149.8 | 1341.9 | 6.28 |
| 7 gün öncesi | 2142.9 | 146.0 | 1121.3 | 6.41 |

Çizelge 4.5.'de verilen SIR yöntemi ile 3, 5 ve 7 gün önceki verileri kullanarak yapılan tahminlerin hata payları incelendiğinde hata kareleri ortalaması yönünden en iyi tahminlerin; hem günlük vaka tahminlerinde hem de haftalık vaka tahminlerinde 7 gün önceki değerlerin kullanılmasının gerektiği anlaşılmaktadır. Göreceli hata kareleri incelendiğinde ise, en iyi tahminlerin; günlük vaka tahminlerinde 7 gün önceki haftalık vaka tahminlerinde 5 gün önceki değerlerin kullanılmasının gerektiği görülmüştür

## 5.2. ARIMA Yaklaşımı ile Yapılan Tahminler

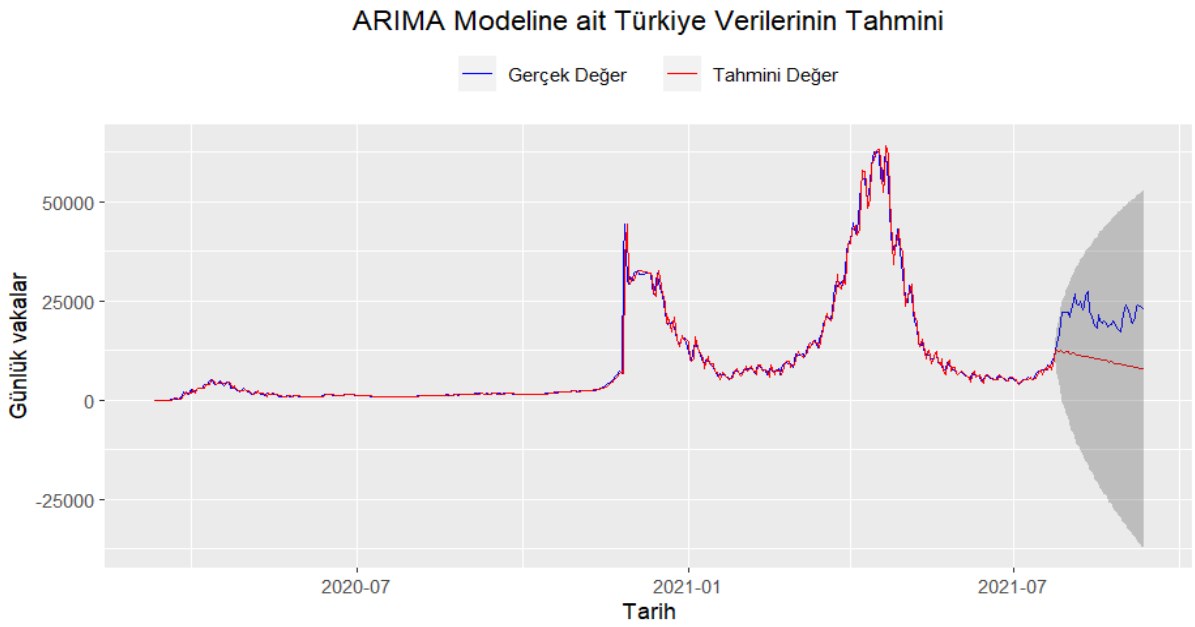
ARIMA yaklaşımında veri salgının başlangıç tarihi, bitiş tarihi, kaç günlük tahmin yapılacağı ve frekansı belirlenerek zaman serisi fonksiyonuna dönüştürülmüş ve Oto ARIMA (Auto ARIMA) kullanılarak en iyi model belirlenmiştir. ARIMA modelinde veri günlük vaka 365 ve haftalık vaka 52 günlük frekanslar kullanarak tahmin edilmiştir. Sonuçlar ülkeler içinde ayrı ayrı sunulmuştur.

### 5.2.1. Türkiye Verilerinin ARIMA Yaklaşımı ile Tahmini

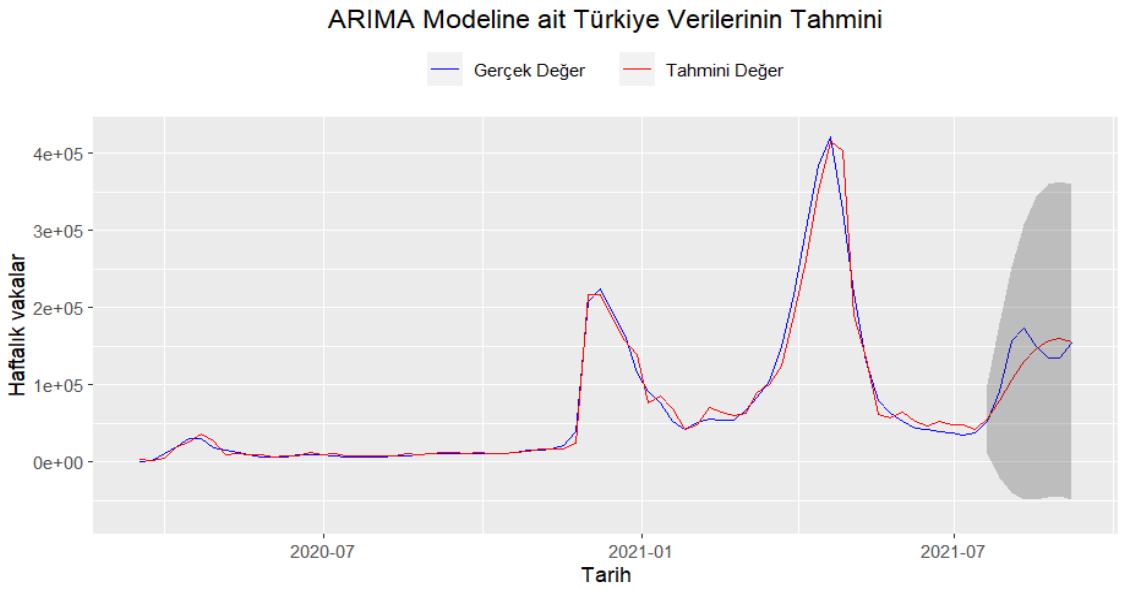
Tüm Türkiye verileri (275. Gün değeri çıkarılarak) ile yapılan tahmin ve öngörü değerlerinin seyir grafikleri Şekil 4.25 ve Şekil 4.26'da sunulmuştur. Görüldüğü üzere yapılan tahminlerde, gerçek değerler verinin başlangıcından itibaren tahmin değerleri ile yakın olsa da, son 50 günlük tahmin değerleri gerçek değerlerden uzaktır.

SIR modelinde uygulanan Türkiye verilerini bölerek analiz etme işlemi, ARIMA modeli için 3 yerine 2 bölüme ayrılarak (ilk 250 gün ve 276 gün ve sonrası) uygulanmıştır. Bunun nedeni 250-274 gün aralığında ARIMA modelinin kurulması ve tahmin yapılabilmesi için yeterli bir veri serisinin olmamasıdır. Burada belirtilen periyotlardaki ARIMA modelinin tahmin ve öngörü değerleri günlük ve haftalık olarak Şekil 4.27 –

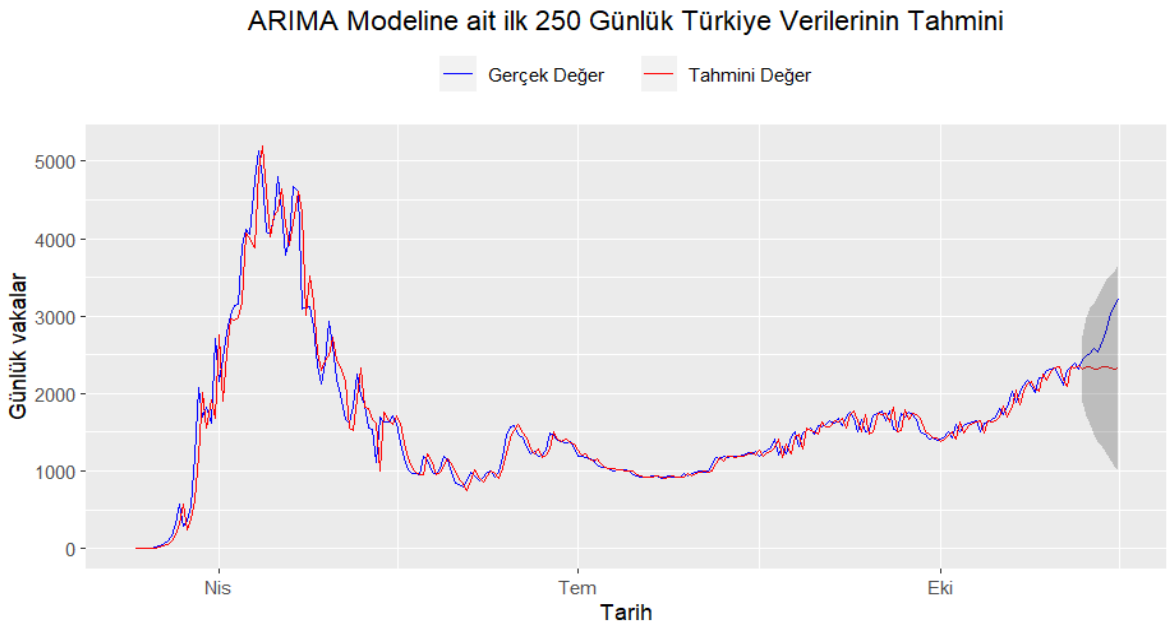
Şekil 4.30 arasında sunulmuş ve her biri için hata kareleri ortalamasının karekökü (RMSE) Çizelge 4.6'da gösterilmiştir.



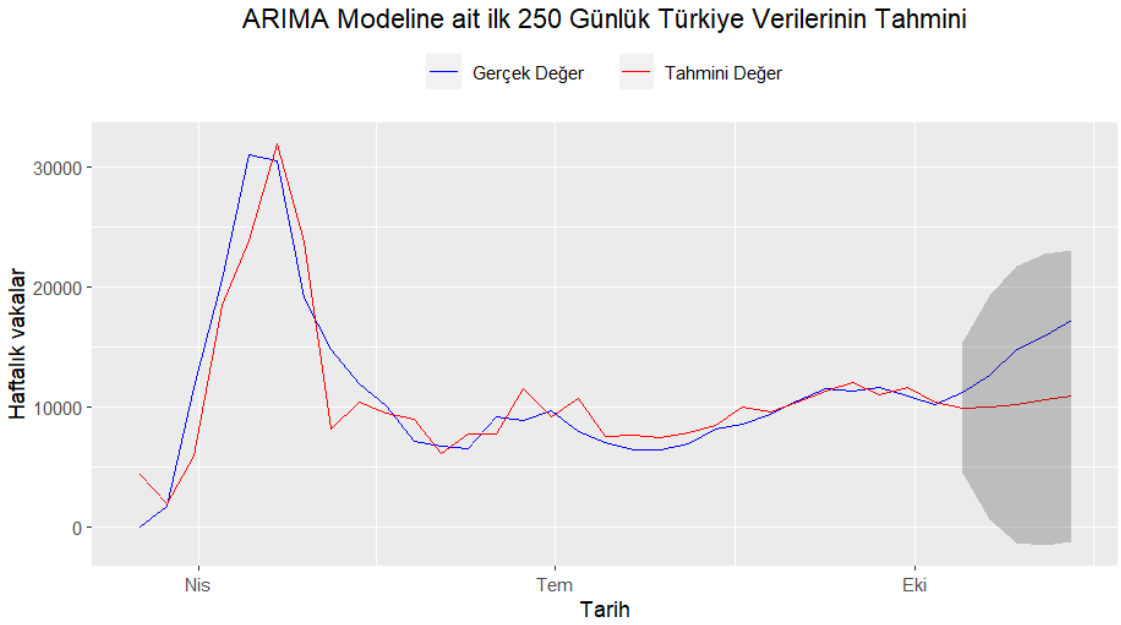

**Şekil 4.25.** ARIMA modeli ile Türkiye verilerinin günlük vaka tahmini



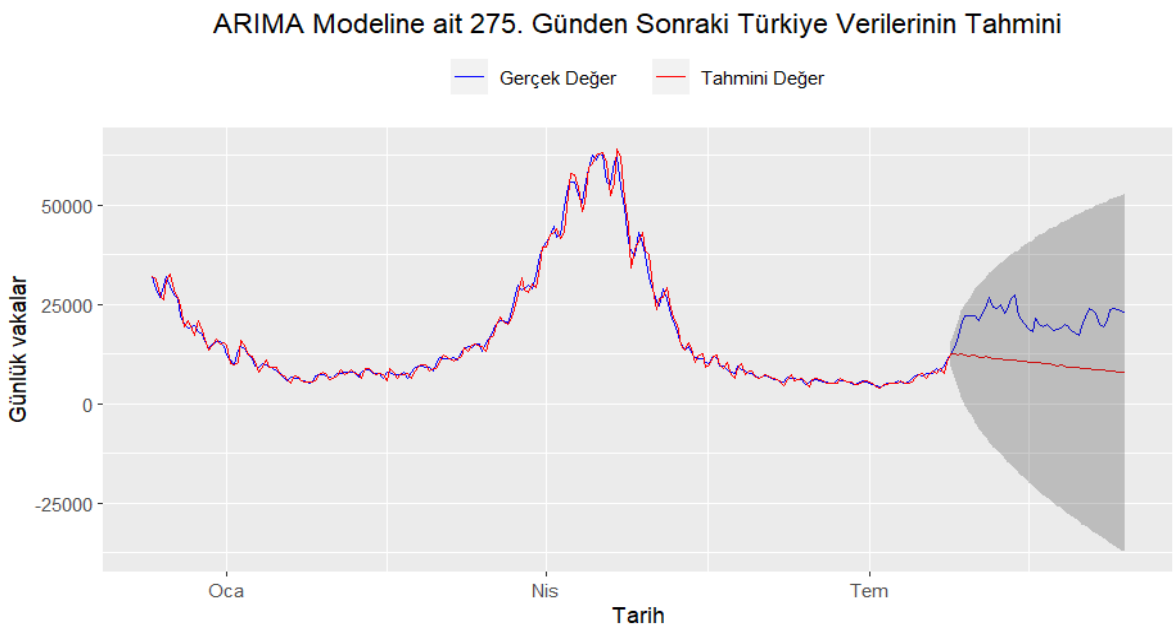

**Şekil 4.26.** ARIMA modeli ile Türkiye verilerinin haftalık vaka tahmini



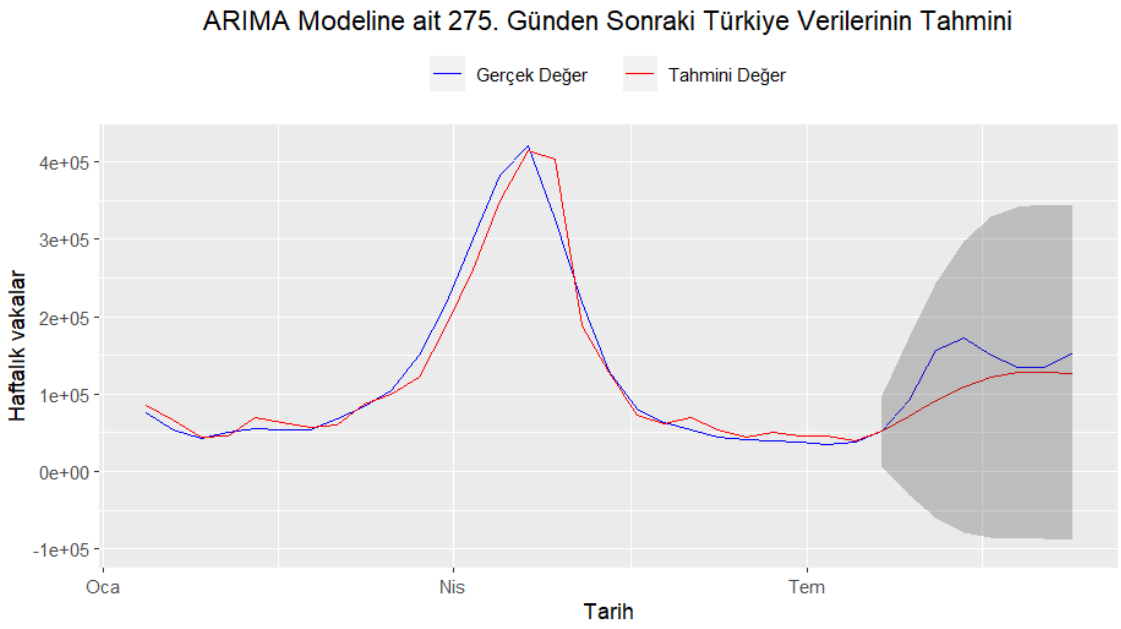

**Şekil 4.27.** ARIMA modeli ile ilk 250 günlük Türkiye verilerinin günlük vaka tahmini




**Şekil 4.28.** ARIMA model ile ilk 250 günlük Türkiye verilerinin haftalık vaka tahmini




**Şekil 4.29.** ARIMA modeli ile 275. gün sonrası Türkiye verilerinin günlük vaka tahmini




**Şekil 4.30.** ARIMA modeli ile 275. gün sonrası Türkiye verilerinin haftalık vaka tahmini

Çizelge 4.6. Türkiye verilerinin ilk 250 gün ve 275. gün sonrası eriler kullanılarak yapılan günlük ve haftalık vaka tahminlerinin hataları

| | Günlük Vaka Tahmini | | Günlük Vaka Öngörüsü | | Toplam | |
|---|---|---|---|---|---|---|
| | RMSE | Göreceli RMSE | RMSE | Göreceli RMSE | RMSE | Göreceli RMSE |
| İlk 250 gün | 217.7 | 21.3 | 462.2 | 15.7 | 228.9 | 21.0 |
| 275. günden sonrası | 1569.6 | 8.8 | 11505.5 | 52.0 | 5037.0 | 23.3 |
| Toplam | 1858.8 | 16.5 | 11706.4 | 52.9 | 3905.2 | 22.2 |
| | Haftalık Vaka Tahmini | | Haftalık Vaka Öngörüsü | | Toplam | |
| | RMSE | Göreceli RMSE | RMSE | Göreceli RMSE | RMSE | Göreceli RMSE |
| İlk 250 gün | 2771.7 | 1818.0 | 4359.9 | 29.5 | 2874.7 | 1670.0 |
| 275. günden sonrası | 22942.6 | 17.6 | 39640.0 | 25.2 | 27226.4 | 19.2 |
| Toplam | 14565.0 | 1100.1 | 33441.5 | 20.5 | 17343.0 | 1026.2 |

Çizelge 4.6.'da ARIMA modeli ile elde edilen modelde kullanılan verilerin tahmin değerleri, modelden hareketle yapılan öngörü değerleri ve bu iki grubun birleşimden oluşan toplam değerler için RMSE ve göreceli RMSE değerleri verilmiştir. Bu değerler önce ilk 250 gün için, sonra 276 gün ve sonrası için bir de bu iki grubun birleşimden hareketle toplam veri için elde edilmiştir.

Günlük vaka tahminleri, öngörü değerleri ve bu iki değerin bir arada alınması halinde RMSE değerlerinde ilk 250 günde daha az hata yapıldığı görülmektedir. Göreceli RMSE için günlük vaka tahminlerinde 276 gün ve sonrasında daha az hata yapılmış olsa da öngörü değerlerinde yapılan yüksek hatalı değerler toplamda 276 gün ve sonrasında tahminlerinde hatayı arttırmıştır.

Haftalık vaka tahminleri, öngörü değerleri ve bu iki değerin bir arada alınması halinde de RMSE değerlerine göre ilk 250 günde daha az hata yapılmıştır. Ancak göreceli RMSE değerlerinde 276 ve gün sonrasında hatalar düşmüş, bu da bu iki periyodun toplamından oluşan tüm veri tahminlerde hata oranını azaltmıştır.

### 5.2.2. Almanya Verilerinin ARIMA Yaklaşımı ile Tahmini

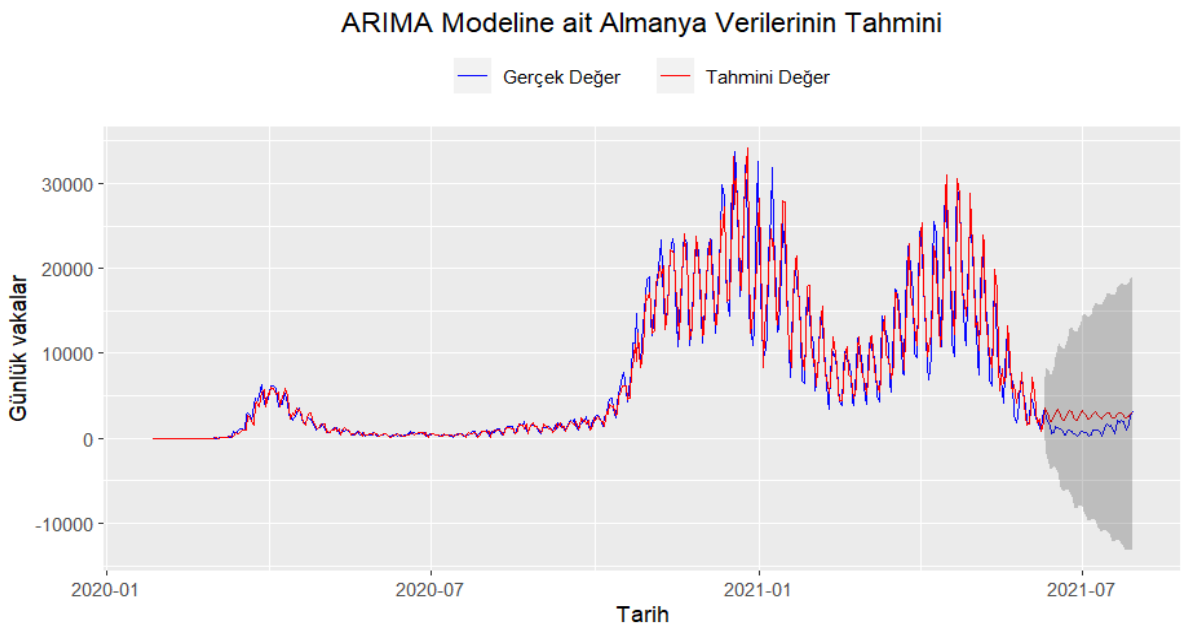
Almanya verileri ile yapılan tahmin ve öngörü değerlerinin seyir grafikleri Şekil 4.31 ve Şekil 4.32'de, modelin yaptığı tahmin işlemlerinin hata kareleri ortalamasının

karekökü (RMSE) Çizelge 4.7'de sunulmuştur. Çizelgeden de görüldüğü üzere günlük vaka öngörüsünde daha düşük RMSE değerleri olmasına rağmen göreceli RMSE değerlerine göre günlük vaka tahminlerinde hatanın daha düşük olduğu saptanmıştır. Haftalık vaka tahminlerinde hem RMSE hem de göreceli RMSE değerleri daha düşük elde edilmiştir.

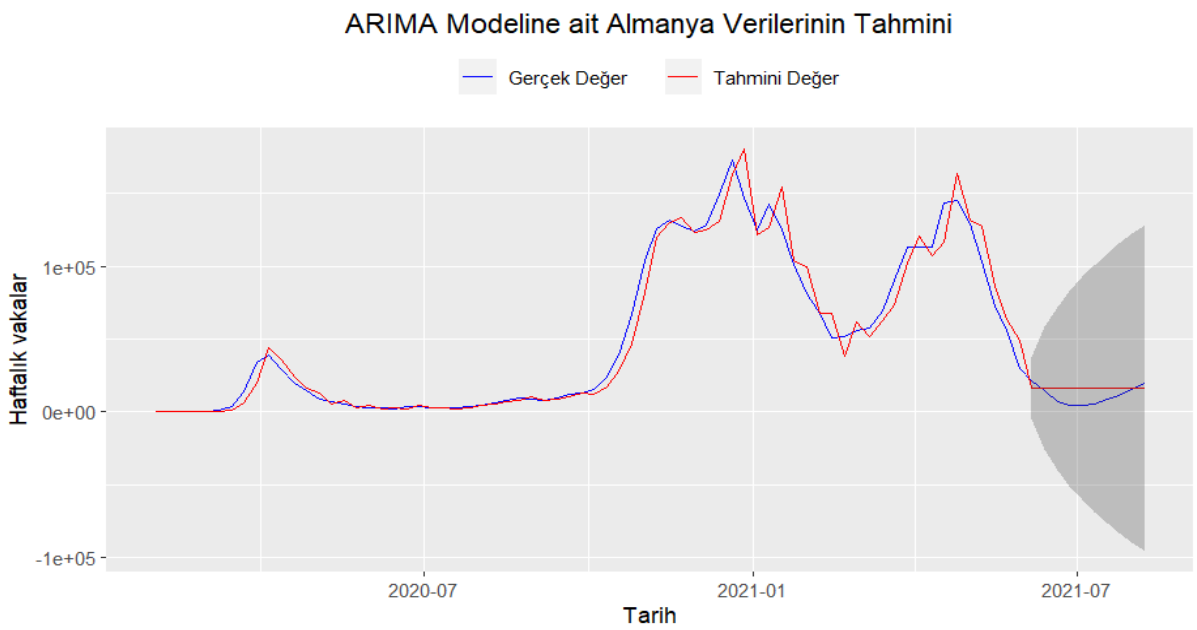
Çizelge 4.7. Almanya verilerinin günlük ve haftalık vaka tahminlerinin hataları

| | Günlük | | Haftalık | |
|---|---|---|---|---|
| | RMSE | Göreceli RMSE | RMSE | Göreceli RMSE |
| Vaka Tahmini | 1900.2 | 34.0 | 10673.0 | 30.3 |
| Vaka Öngörüsü | 1711.0 | 313.0 | 10962.4 | 163.8 |
| Toplam | 1880.8 | 98.2 | 10564.3 | 59.9 |

Çizelge 4.7'de sunulan bulgular Şekil 4.31 ve 4.32'den de teyit edilmektedir. Şekillerden de görüldüğü üzere gerçek değerler ile öngörü değerleri arasındaki fark gerçek değerler ile tahmin değerleri arasındaki farktan daha geniştir.




**Şekil 4.31.** ARIMA modeli ile Almanya verilerinin günlük vaka tahmini



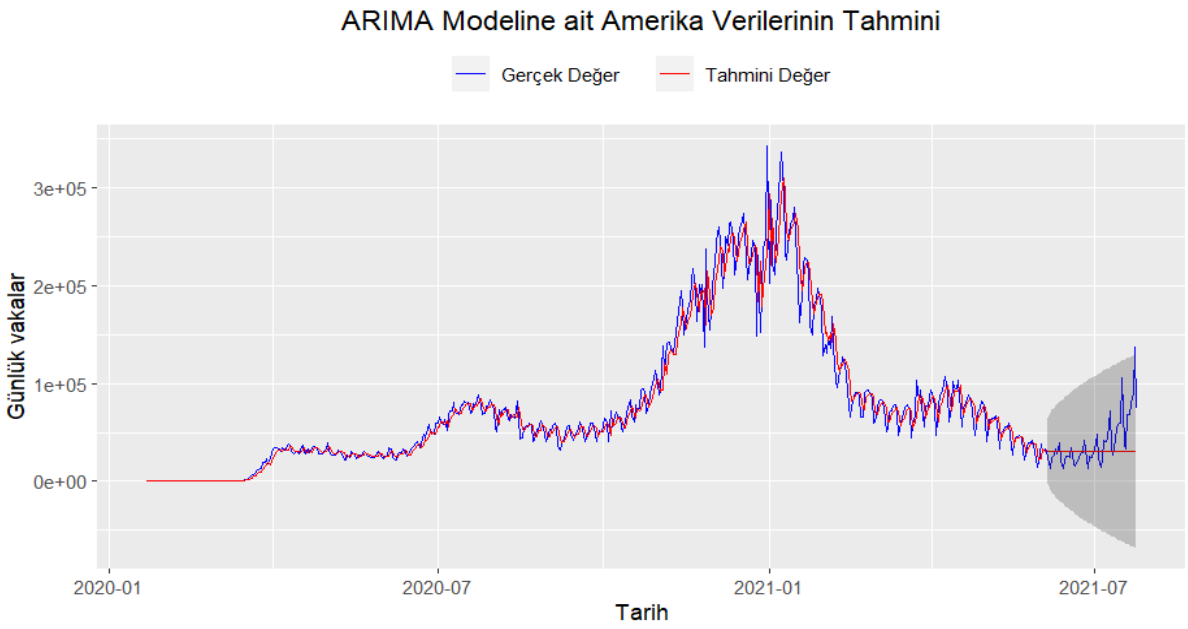

**Şekil 4.32.** ARIMA modeli ile Almanya verilerinin haftalık vaka tahmini

### 5.2.3. Amerika Verilerinin ARIMA Yaklaşımı ile Tahmini

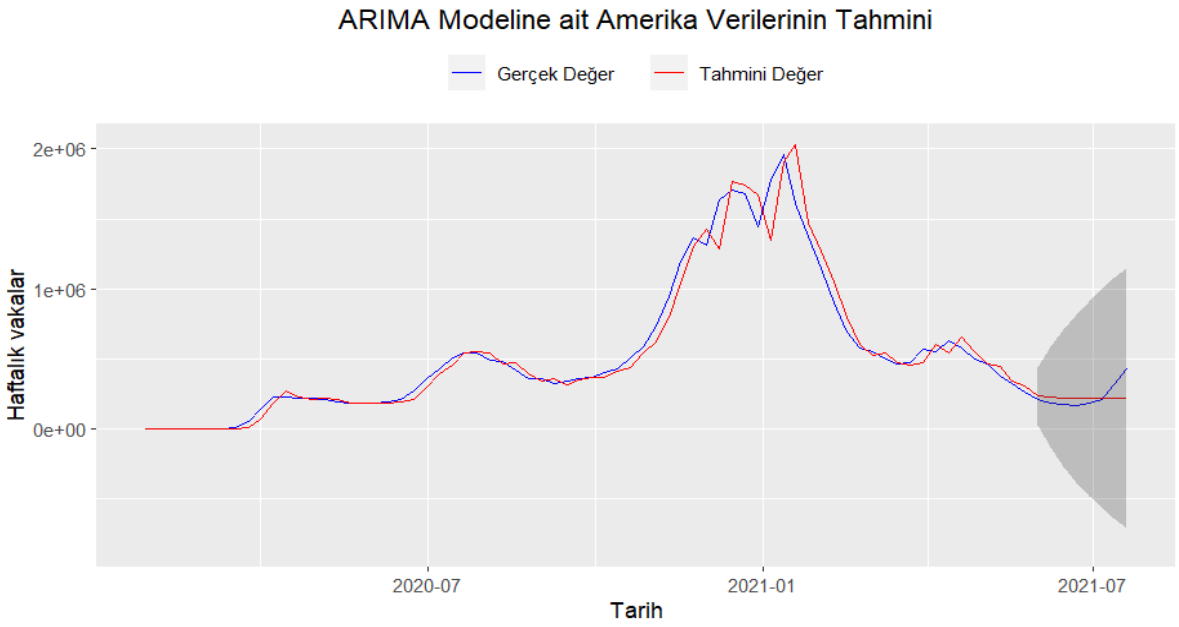
Amerika verileri ile yapılan tahmin ve öngörü değerlerinin seyir grafikleri Şekil 4.33 ve Şekil 4.34'te, modelin yaptığı tahmin işlemlerinin hata kareleri ortalamasının karekökü (RMSE) Çizelge 4.8'de sunulmuştur. Çizelgeden de görüldüğü üzere günlük vaka tahminlerinde RMSE ve göreceli RMSE değerlerinin daha düşük olduğu saptanmıştır. Haftalık vaka öngörüsünde hem RMSE hem de göreceli RMSE değerleri daha düşük elde edilmiştir..

Çizelge 4 8. Amerika verilerinin günlük ve haftalık vaka tahminlerinin hataları

| | Günlük | | Haftalık | |
|---|---|---|---|---|
| | RMSE | Göreceli RMSE | RMSE | Göreceli RMSE |
| Vaka Tahmini | 16970.3 | 26.2 | 105954.7 | 34.5 |
| Vaka Öngörüsü | 26307.5 | 53.5 | 90407.1 | 27.9 |
| Toplam | 17955.4 | 29.6 | 103041.8 | 33.4 |




**Şekil 4.33.** ARIMA modeli ile Amerika verilerinin günlük vaka tahmini




**Şekil 4.34.** ARIMA modeli ile Amerika verilerinin haftalık vaka tahmini

### 5.2.4. Çekya Verilerinin ARIMA Yaklaşımı ile Tahmini

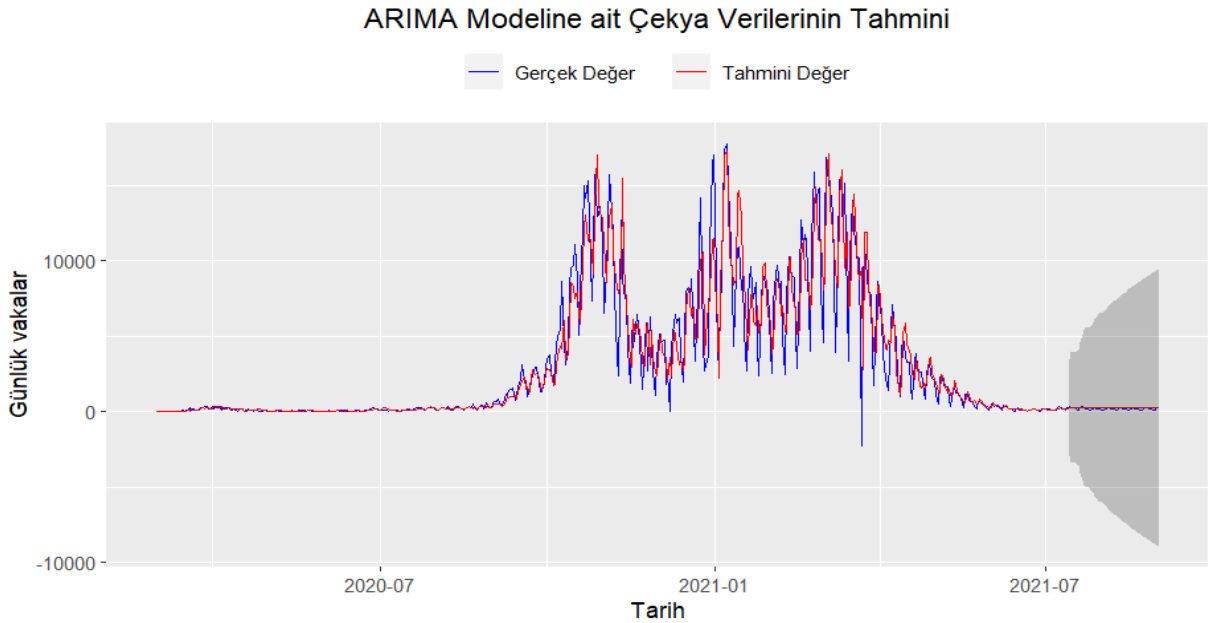
Çekya verileri ile yapılan tahmin ve öngörü değerlerinin seyir grafikleri Şekil 4.35 ve Şekil 4.36'da, modelin yaptığı tahmin işlemlerinin hata kareleri ortalamasının karekökü (RMSE) Çizelge 4.9'da sunulmuştur. Çizelgeden de görüldüğü üzere hem

günlük vaka öngörüsünde hem de haftalık vaka öngörüsünde daha düşük RMSE değerleri olmasına rağmen göreceli RMSE değerlerine göre günlük ve haftalık vaka tahminlerinde hatanın daha düşük olduğu saptanmıştır. Bunun sonucunda da toplamdaki RMSE değerleri daha düşük, göreceli RMSE değerleri de vaka tahminlerindeki RMSE değerlerine oldukça yakındır.

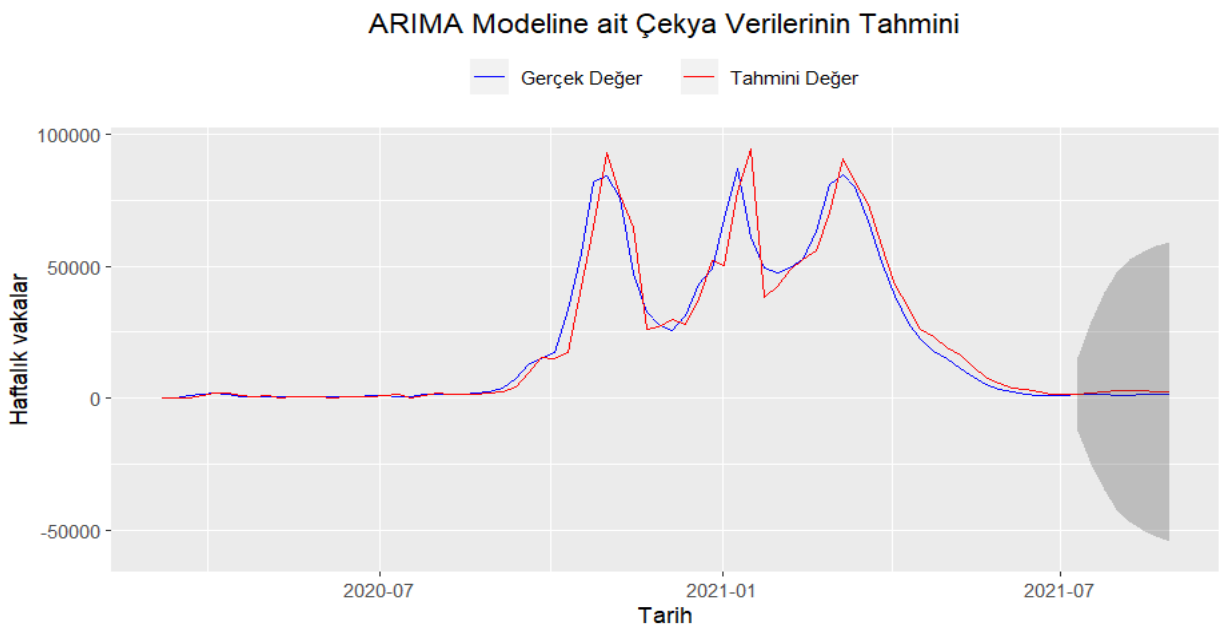
Çizelge 4.9. Çekya verilerinin günlük ve haftalık vaka tahminlerinin hataları

| | Günlük | | Haftalık | |
|---|---|---|---|---|
| | RMSE | Göreceli RMSE | RMSE | Göreceli RMSE |
| Vaka Tahmini | 1606.2 | 77.0 | 7126.3 | 40.8 |
| Vaka Öngörüsü | 98.8 | 94.2 | 1141.3 | 90.3 |
| Toplam | 1531.3 | 78.5 | 6657.6 | 47.9 |

Şekil 4.35 ve 4.36'da vaka öngörü değerlerinin gerçek değerlere çok yakın olduğu (neredeyse üst üste olduğu) görülmektedir. Öngörü değerlerinde gerçek değerlerin sıfıra yakın değerler olması ve buradaki tahmin değerlerinde rakamsal olarak küçük hata yapılsa da oransal olarak bu hataların yüksek olması nedeniyle göreceli RMSE değerleri öngörü tahminlerinde yükselmiştir.




**Şekil 4.35.** ARIMA modeli ile Çekya verilerinin günlük vaka tahmini




**Şekil 4.36.** ARIMA modeli ile Çekya verilerinin haftalık vaka tahmini

Şekil 4.37. Ülkelere göre SIR ve ARIMA modellerinin RMSE ve göreceli RMSE değerleri

| **Ülkeler** | **Modeller** | **Günlük vaka tahminleri** | | **Haftalık vaka tahminleri** | |
|---|---|---|---|---|---|
| | | **Hata Kareleri Ortalaması (RMSE)** | **Göreceli Hata Kareleri Ortalaması** | **Hata Kareleri Ortalaması (RMSE)** | **Göreceli Hata Kareleri Ortalaması** |
| Türkiye | SIR Modeli 3 gün öncesi | 2442.0 | 48.8 | 1451.5 | 2.34 |
| | 5 gün öncesi | 2266.0 | 18.0 | 5176.0 | 5.68 |
| | 7 gün öncesi | 2082.9 | 17.1 | 2880.3 | 3.06 |
| | ARIMA Modeli | 3905.2 | 22.2 | 17343.0 | 1026.2 |
| Almanya | SIR Modeli 3 gün öncesi | 3984.8 | 53.9 | 1514.6 | 5.79 |
| | 5 gün öncesi | 3831.6 | 55.4 | 2948.9 | 5.39 |
| | 7 gün öncesi | 3068.4 | 47.3 | 1534.8 | 14.1 |

| | | | | | |
|---|---|---|---|---|---|
| | ARIMA Modeli | 1880.8 | 98.2 | 10564.3 | 59.9 |
| Amerika | SIR Modeli | | | | |
| | 3 gün öncesi | 22899.2 | 44.7 | 11320.5 | 11.4 |
| | 5 gün öncesi | 21693.5 | 41.3 | 11883.9 | 8.91 |
| | 7 gün öncesi | 19486.5 | 38.2 | 11862.2 | 5.24 |
| | ARIMA Modeli | 17955.4 | 29.6 | 103041.8 | 33.4 |
| Çekya | SIR Modeli | | | | |
| | 3 gün öncesi | 2619.2 | 157.5 | 1485.2 | 8.94 |
| | 5 gün öncesi | 2537.7 | 149.8 | 1341.9 | 6.28 |
| | 7 gün öncesi | 2142.9 | 146.0 | 1121.3 | 6.41 |
| | ARIMA Modeli | 1531.3 | 78.5 | 6657.6 | 47.9 |

SIR ve ARIMA modelleri ile yapılan tahmin başarılarının tüm ülkelere göre RMSE ve göreceli RMSE değerleri Çizelge 4.10'da sunulmuştur. Bu çizelgeden de görüldüğü üzere

- günlük RMSE değerlerinden hareketle Türkiye verilerinde SIR modelinin diğer ülkelerde ise ARIMA modelinin daha az hatalı tahminler yaptığı
- günlük göreceli RMSE değerlerinden hareketle Türkiye (7 gün öncesi) ve Almanya verilerinde SIR modelinin, Çekya ve Amerika verilerinde ise ARIMA modelinin daha az hatalı tahminler yaptığı,
- hem haftalık RMSE hem de haftalık göreceli RMSE değerlerinden hareketle tüm ülkelerde SIR modelinin daha az hatalı tahminler yaptığı

görülmüştür.

# 6. TARTIŞMA

Yıllar boyunca salgın hastalıklar araştırma konusu olmuştur. SIR ve ARIMA modeli salgın hastalıkların analizi için geliştirilen matematiksel modellerdendir. Bu modeller salgın hastalıkların gidişatını matematiksel olarak ortaya koymakta ve yayılımının azalması için uygulanacak tedbirler konusunda araştırmacılara bilgi vermektedir. Son olarak yaşanılan Covid-19 salgını için literatürde bu modeller kullanılarak yapılan birçok çalışma vardır. SIR ve ARIMA metodu bu analizlerin yapılması için kullanılan önemli modellerdendir. Bu tezde yapılan çalışmalara ek olarak bu iki modelin vaka tahmin başarıları kıyaslanmıştır.

Literatürde modellerin birbirlerine olan üstünlükleri ve eksiklikleri yönünden tartışıldığı az sayıda çalışma mevcuttur. Bu çalışmalardan bir tanesi (bu teze başladıktan sonra yayınlanmıştır) olan Abolmaali ve Shirzaei'nin [37] çalışmasında Covid-19 pandemisi için SIR ve ARIMA modelleri ilk 200 günlük veri kullanılarak karşılaştırılmış ve karşılaştırılan modellerin veri elde edilme zamanı ve veri boyutundan etkilendiği gösterilmiştir. Çalışmada vaka sayısının en yüksek olduğu 4 farklı ülke için 4 farklı model (SIR, ARIMA, Lojistik Regresyon, …) kullanılmıştır. Modeller her ülke için farklı sonuçlar sergilemiştir. Yapılan analiz sonuçlarında SIR modelinin salgının ilk periyodunda iyi tahmin yapamadığını, ancak salgının ikinci periyodunda daha başarılı tahmin yaptığını göstermişlerdir. ARIMA modeli ile yapılan analizlerde tüm periyotlarda diğer modellerden daha başarılı olduğu belirtilmiştir. Bu çalışmanın sonucunda analiz edilen sürenin uzatılması ile sonuçların değişebileceğini belirtmişlerdir. Bu çalışmada SIR modeline, yapılan tahmin hataların eklenmesi sonucunda hatanın azaltılabileceği söylenmiştir.

Bu çalışmada Abolmaali ve Shirzaei'nin çalışmasında olduğu gibi farklı ülkelerde farklı yöntemlerin daha az hatalı tahminler yaptığı gösterilmiştir. Analizlere göre Türkiye verisinde günlük vaka sayısı tahmininde SIR modeli daha iyi sonuçlar verirken, diğer ülkelerin günlük vaka sayısı tahminlerinde ARIMA modelinin daha başarılı olduğu görülmüştür. Abolmaali ve Shirzaei'nin çalışmasındaki ülke verilerinin 2 periyoda ayrılmasındaki gibi Türkiye verileri de ilk 250 gün, 250 -275 gün ve 275 gün sonrası olarak 3 bölüme ayrılmıştır. Bu üç periyot için SIR modeli uygulanmıştır. Ancak 2.periyodun ARIMA modeli oluşturulabilmesi için çok kısa olması nedeniyle ilk ve son

periyot kullanılarak ARIMA modeli elde edilmiştir. Abolmaali ve Shirzaei'nin çalışmasından farklı olarak bu çalışmada Türkiye verilerinde SIR modelinin ilk periyottaki tahmin başarısının son periyottaki tahmin başarısından iyi olduğu görülmüştür. Bunun muhtemel nedeninin Türkiye verilerinde vaka tanımının değişmesi olduğu düşünülmektedir.

Bu tez çalışmasında da ARIMA modeli SIR modeline göre bazı ülkelerde daha başarılı bulunmuştur. Ancak bu başarı günlük vaka tahminlerinde gözlenirken, haftalık vaka tahminlerinde aynı düzeyde başarı elde edilememiştir. Haftalık vaka tahminlerinde başarının düşme nedeni ARIMA modelinde tahmin için kullanılan verinin yeterli büyüklükte olmamasıdır. Çalışmada 500 günlük verinin kullanılması sonucunda bu vakaların haftalık toplamlarının alınması nedeniyle toplam 70 haftalık veri ile ARIMA modeli kurulmuştur. Bu büyüklükteki verinin ARIMA modelinin tahmin başarısı negatif yönde etkilediği düşünülmektedir.

SIR modeli günlük vaka tahminlerinde haftalık vaka tahminlerine göre daha az başarılı bulunmuştur. SIR modelinde modeldeki her bir bileşen seçilen gündeki toplam gözlem (toplam risk altındaki kişi sayısı, toplam enfekte olan kişi sayısı, toplam iyileşen kişi sayısı) sayılarını göstermektedir. Bu çalışmadaki SIR modeli modifiye edilerek günlük yeni vaka sayıları üzerinden analiz edilmiştir. Bunun için ardışık iki gün arasındaki fark alınarak günlük yeni eklenen vaka sayıları kullanılmıştır. Haftalık vaka sayıları hesabında ise günlük yeni vaka sayılarının 7 günlük toplamları alınmıştır. SIR modeli toplam gözlem sayıları ile çalıştığından bu tezde elde edilen haftalık vaka tahminlerinde daha başarılı tahminler yapma bulgusu mantıklıdır.

Bu tezde hastalık seyri, yapılan müdahaleler ve virüsün öngörülemez gidişatı nedeniyle farklı zaman periyotlarında SIR modeli için parametre tahmininde kullanılacak veri boyutunun (geriye yönelik gün sayısı) sonucu etkilediği belirlenmiştir. Buna göre ülkelere göre değişmekle birlikte geriye yönelik gün sayısının 5 veya 7 olmasının tahmin başarısını olumlu yönde arttırdığı gösterilmiştir.

Güzey[33] tarafından 2021 yılında yapılan çalışmada,Covid-19 sürecinin başından itibaren 3 aylık Türkiye verisi ile, alınan tedbirler göz önüne alınarak SIR modeliyle salgının seyri incelenmiştir. Beta ve gamma parametreleri tedbir alınan 4 tarih aralığında farklı uzunluktaki aralık süreleri (5 gün, 26 gün, 12 gün ve 38 gün) için sabit bir değer olarak alınmıştır. Çalışmanın sonucunda yapılan tahmin başarılarının düşük olduğu ve

beta ve gamma parametrelerinin dinamik olarak hesaplanması gerektiğini belirtmiştir. Bu tezde ise, çalışmanın aksine beta ve gamma parametreleri her bir tarih için tek tek hesap edilmiş ve sabit bir değer alınmamıştır. Bunun sonucunda daha gerçekçi vaka tahminleri yapılmıştır.

Ceylan [30], İtalya, İspanya ve Fransa gibi bazı Avrupa ülkelerinde Covid-19 yayılımını incelemek amacıyla yaptığı çalışmada, ARIMA modelinin durağan zaman serilerinde kısa dönemde anlamlı sonuçlar verdiğini ve sonuçların verilerin zamana bağlı davranışlarına göre değiştiğini göstermiştir. Bu tezde de Almanya ve Çekya gibi Avrupa ülkelerinde ARIMA modelinin daha başarılı tahminler verdiği gösterilmiştir.

Bu tez çalışmasında SIR ve ARIMA modellerinin tahmin başarılarını kıyaslamak için RMSE ve göreceli RMSE yöntemi kullanılmıştır. RMSE yöntemine ilave olarak göreceli RMSE değerinin kullanılma amacı yapılan tahmin değerlerindeki yüksek hata değerlerinin RMSE değerini aşırı yükseltmesidir. Bu nedenle göreceli RMSE değerinde her nokta yapılan tahmindeki yüzdesel hata değeri ortalaması elde edilmiştir.

Cooper ve ark. [15] , çalışmalarında farklı ülkelerin (Çin, Güney Kore, Hindistan, Avustralya, İtalya ve ABD'nin Teksas eyaleti) Covid-19 salgınının zaman içerisindeki gelişimlerini, SIR modeli ile araştırmışlardır. Analizlerde, hükümetlerin uyguladığı tedbirlerin öncesi ve sonrası vaka sayısı değişimini incelenmiştir. Bu çalışmada SIR modelinin basit bir model olmasından dolayı, tahminlerdeki tutarlılığının tartışmalı olduğu ancak verilerin güvenilir olması halinde SIR modelinin virüsün yayılımı ile daha iyi öngördüğü belirtilmiştir.

Bu tezde genel bilgilerde de verilen, salgındaki vaka sayısını tahmin eden birçok matematiksel model bulunmaktadır. Ancak bu tezde, bahsi geçen birçok model, model için gerekli olan bilgilere ulaşılamaması nedeniyle, uygulanamamıştır. Örneğin bazı modellerin tahmin yapabilmesi için ihtiyaç duyduğu bazı bilgilere (maruz kalan kişi sayısı, doğuştan bağışıklığı olan kişi sayısı gibi) ulaşmak bu salgın için mümkün olmamıştır. Ayrıca, yine bu salgında geçerli olmak üzere aşılanmış kişi sayısı gibi aşı ile ilgili bilgilerin net olmaması (kaç doz uygulama tam bağışıklık kabul edilecek veya hangi tür aşılar yapılmış kişiler aşılanmış kişi olarak kabul edilecek gibi) nedeninden dolayı bazı modeller uygulanamamıştır.

# 7. SONUÇ ve ÖNERİLER

Bu tez çalışmasında COVID-19 pandemisindeki vaka sayılarının tahmininde kullanılmak üzere iki model, SIR ve ARIMA modelleri, 4 farklı ülkeden elde edilen veriler kullanılarak karşılaştırılmıştır. Çalışmanın sonucunda, COVID-19 vaka sayısı modellemesi için elde edilen sonuçlar ve öneriler şöyledir;

- SIR modeli yapısı gereği günlük yeni vaka tahmininden daha çok haftalık yeni vaka sayısı tahmininde kullanılmalıdır.
- SIR modelindeki beta ve gamma değerlerinin hesaplanması dinamik ve kısa süreli aralıklar için yapılmalıdır.
- SIR modeli yerine tahmin hata değerlerini de dikkate alan geliştirilmiş yeni bir SIR modeli tahmin başarısını artırabilir.
- ARIMA modeli günlük yeni vaka tahmininde kullanılabilir.
- ARIMA modelinin oluşturulabilmesi için veri boyutunun yüksek olması gerekmektedir.
- ARIMA modeli Çekya verilerinde gerçek değerlere çok yakın sonuçlar vermektedir, bu da vaka sayılarının daha güvenilir olduğu ülkeler için modelin daha başarılı tahmin yaptığını göstermektedir.

# KAYNAKÇA


1. **Green MS, Swartz T, Mayshar E, Lev B, Leventhal A, Slater PE, Shemer J (2002)**. "When is an epidemic an epidemic?" (PDF). Israel Medical Association Journal. 4 (1): 3–6. PMID 11802306.

2. **Dietz K, Heesterbeek JA.** Bernoulli was ahead of modern epidemiology. Nature. **2000**;408(6812):513–514. doi: 10.1038/35046270

3. **Ross R. Murray**; The prevention of malaria, **1911,** John Murray**,** London

4. **Hamer WH**. Epidemic disease in England – the evidence of variability and persistency of type. Lancet. **1906**;1:733–739

5. **Kermack O.W. and McKendrick A.G.** , A contribution to the mathematical theory of epidemics, Proc. R. Soc. Lond. A. **1927**; 115; 700-721.

6. **Kermack O.W. and McKendrick A.G.** , A contribution to the mathematical theory of epidemics, Proc. R. Soc. Lond. A. **1927**; 115; 700-721

7. **Kermack O.W. and McKendrick A.G** ,Contributions to the mathematical theory of epidemics. II The problem of endemicity **1932** . doi:10.1098/19320171

8. **Hethcote H.W**, Three basic epidemiological models : Applied mathematical ecology, Springer, Berlin, Heidelberg **1989,** 119-144

9. **Kuznetsov Y.A, Piccardi C**. : Bifurcation analysis of periodic seir and sir epidemic models. J. Math. Biol. 32, **1994**, 109–121

10. **Peng L, Yang W, Zhang D, Zhuge C, Hong L**. Epidemic analysis of covid-19 in china by dynamical modeling, arXiv preprint arXiv:**2002**.06563

11. **Wallinga J, Teunis P.** Different epidemic curves for severe acute respiratory syndrome reveal similar impacts of control measures. Am J Epidemiol. **2004**;160(6):509–516.

12. **Sato R.C** ,Disease management with ARIMA model in time series Einstein, 11 (1) **2013**, p.128

13. **WHO Director-General's** opening remarks at the media briefing on COVID19 -March **2020**

14. **Anastassopoulou C., Russo L., Tsakris A., Siettos C.** Data-based analysis, modelling and forecasting of the covid-19 outbreak. PloS One**. 2020;**15(3):1–21

15. **Cooper I., Mondal A., Antonopoulos C.G. A** SIR model assumption for the spread of COVID-19 in different communities. Chaos, Solitons & Fractals**. 2020;**139:110057**.**

16. **Giordano G., Blanchini F., Bruno R., Colaneri P., Di Filippo A., Di Matteo A., Colaneri M.** Modelling the covid-19 epidemic and implementation of population-wide interventions in Italy. Nature Medicine. **2020**

17. **Kucharski A., Russell T., Diamond C., Liu Y., CMMID nCoV working group. Edmunds J., Funk S., Eggo R. ,**Analysis and projections of transmission dynamics of nCoV in Wuhan**,2020**

18. **Ndairou F., Area I., Nieto J.J., Torres D.F.** Mathematical modeling of COVID-19 transmission dynamics with a case study of Wuhan. Chaos, Solitons & Fractals. **2020**

19. **Rafiq D., Suhail S.A., Bazaz M.A.,** Evaluation and prediction of COVID-19 in India: A case study of worst hit states. Chaos, Solitons & Fractals**. 2020**

20. **Russo L., Anastassopoulou C., Tsakris A., Bifulco G., Campana E., Toraldo G., Siettos C.** Tracing day-zero and forecasting the covid-19 outbreak in lombardy, Italy: A compartmental modeling and numerical optimization approach.,**2020**

21. **Silva P.C., Batista P.V., Lima H.S., Alves M.A., Guimarães F.G., Silva R.C. ,**COVID-ABS: An agent-based model of COVID-19 epidemic to simulate health and economic effects of social distancing interventions. Chaos, Solitons & Fractals**, 2020**

22. **Wu J., Leung K., Leung G. ,**Nowcasting and forecasting the potential domestic and international spread of the 2019-ncov outbreak originating in Wuhan, China: A modelling study. Lancet. 689–697,**2020.**

23. **Zhou F., Yu T., Du R., Fan G., Liu Y, Liu Z.…Guan L.** Clinical course and risk factors for mortality of adult inpatients with covid-19 in Wuhan, China: A retrospective cohort study. Lancet.1054–1062**,2020**

24. **Imai N. , Cori A. , Dorigatti I. , Baguelin M. , Donnelly CA. , Riley S.ve Ferguson NM**., Rapor 3: Transmissibility of 2019-nCoV, Imperial College London , **2020**

25. **Gao W. , Veeresha P. , Prakasha DG. , Baskonus HM. ve Yel G. , Chaos,** Solitons Fractals 134,109696 **2020.** doi: 10.1016/2020.109696

26. **Tang B. , Wang X. , Li Q. , Bragazzi NL. , Tang S. , Xiao Y. ve Wu J. , Clin Med**. **J**. ,9 (2), 462 **2020**. doi: 10.3390/9020462

27. **Prasse B. , Achterberg M.A. , Ma L.** et al. : Network-based prediction of the 2019-ncov epidemic outbreak in the chinese province hubei **2020**. arXiv:2002.04482

28. **Wu JT. , Leung K. ve Leung GM**. , Lancet; **2020**; 395 (10225), 689– 697

29. **Read JM. , Bridgen JR. , Cummings DA. , Ho A. ve Jewell CP**. , " Yeni koronavirüs 2019-nCoV: Epidemiyolojik parametrelerin ve salgın tahminlerin erken tahmini ", medRxiv **2020**. doi:10.1101/2020.01.23.20018549)

30. **Ceylan Z.** Estimation of COVID-19 prevalence in Italy, Spain, and France The Science of the Total Environment **2020**, p. 138817

31. **Alzahrani S.I , Aljamaan I.A, Al-Fakih E.A**. Forecasting the spread of the COVID-19 pandemic in Saudi Arabia using ARIMA prediction model under current public health interventions Journal of infection and public health, **2020**, pp. 914-919

32. **Mohammad M, Trounev A, Cattani C.** The dynamics of covid-19 in the uae based on fractional derivative modeling using riesz wavelets simulation ,115 **2021**. doi:10.1186/s13662-021-03262-7

33. **Nurbanu Guzey**, Türkiye'deki COVID-19 yayılımının SIR temelli modellenmesinde RSS yöntemi ile parametre kestirimi, GÜFBED/GUSTIJ **2021** 11 (3): 956-963. doi: 10.17714/gumusfenbil.757291

34. **Karthick K. , Algehyne E. A. , Sekar K**. , Analysis of 'earlyR' epidemic model and time series model for prediction of COVID-19 registered cases. Materials Today: Proceedings. **2020**

35. **Kermack O.W. and McKendrick A.G.** , A contribution to the mathematical theory of epidemics, Proc. R. Soc. Lond. A. **1927**; 115; 700-721.

36. T.C. Sağlık Bakanlığı Covid-19 Rehberi, Erişim linki: https://covid19.saglik.gov.tr/TR-66301/covid-19-rehberi.html, Son erişim tarihi: 01.06.2024

37. **Abolmaali S. and Shirzaei S.** , A comparative study of SIR Model, Linear Regression, Logistic Function and ARIMA Model for forecasting COVID-19 cases **2021**, doi: 10.3934/2021048

38. **Koylu F**. , Pandemi Sürecinde Ülkelere Göre Yayınlanan Covid-19 Vaka ve Ölüm Verilerinin Benford Kanununa Uyumu, **2024**

## EKLER

```
# SIR model parametrelerin belirlenmesi

betagamma_tahmin<-function(St,It,Rt,N) {
      gamma1<-0
      gamma2<-0
      beta1<-0
      beta2<-0
      gun<-length(St)-1
      for (i in 1:gun) {
         gamma1<-gamma1+(Rt[i+1]-Rt[i])*It[i]
         gamma2<-gamma2+It[i]^2
      }
      gamma<-gamma1/gamma2
      for (i in 1:gun) {
         beta1<-beta1+(It[i+1]-It[i]+gamma*It[i])*St[i]*It[i]
         beta2<-beta2+(St[i]*It[i])^2
      }
      beta<-N*beta1/beta2
      S<-St[gun]-beta/N*St[gun]*It[gun]
      I<-It[gun]+beta/N*St[gun]*It[gun]-gamma*It[gun]
      R<-Rt[gun]+gamma*It[gun]
      return(c(gamma,beta,S,I,R))
}

tahmin<-function(veri,t,g,b,kacgunluk) {
    # t. gunden itibaren verilen gamma (g) beta (b) degerlerine gore
    # belirtilen gun kadar (kacgunluk) tahmin sonucunu verir
    require(tidyverse)
    S<-veri$St[t]
    I<-veri$It[t]
    R<-veri$Rt[t]
    N<-veri$N[1]
    tahminler<-matrix(c(0),nrow=kacgunluk,ncol=3)
    for (i in 1:kacgunluk) {
       tahminler[i,1]<-S-b/N*S*I
       tahminler[i,2]<-I+b/N*S*I-g*I
       tahminler[i,3]<-R+g*I
       S<-tahminler[i,1]
       I<-tahminler[i,2]
       R<-tahminler[i,3]
    }
    yveri<-veri[(t+1):(t+kacgunluk),] %>%
mutate(S_tahmin =tahminler[,1],
I_tahmin=tahminler[,2],
R_tahmin=tahminler[,3]
```

```
        )
    return(yveri)
}


tumtahmin<-function(data,bas,oncesi,kacgunluk=(nrow(data)-oncesi)) {
    # Baslangic (bas) ve baslangic + oncesi gun arasindaki verilerle
    # gamma ve beta degerlerini hesaplar ve bas+oncesi gun sonrasi
    # belirtilen gun (kacgunluk) kadar tahmin verir
    # kacgunluk argumanina deger verilmediginde tum veriyi tahmin eder
    # tum veriyi tahmin etmek icin toplam veriden (nrow(data))
    # oncesi gun sayisini (oncesi) cikarip tahmin yapabiliriz

    tahminler<-NULL
    i=0
    bas1<-bas
    if (bas1 > 1) {
     bg<-betagamma_tahmin(data$St[(bas1-1):(bas+oncesi-1)],
                  data$It[(bas1-1):(bas+oncesi-1)],
                  data$Rt[(bas1-1):(bas+oncesi-1)],
                  data$N[1])
     aa<-tahmin(data,bas+oncesi-1,bg[1],bg[2],1)
     tahminler<-rbind(tahminler,aa)
    }
    while (i<kacgunluk) {
     bg<-betagamma_tahmin(data$St[(bas):(bas+oncesi)],
                  data$It[(bas):(bas+oncesi)],
                  data$Rt[(bas):(bas+oncesi)],
                  data$N[1])
     aa<-tahmin(data,bas+oncesi,bg[1],bg[2],1)
     tahminler<-rbind(tahminler,aa)
     bas<-bas+1
     i<-i+1
    }

    return(tahminler)
}

gunlukvaka<-function(veri)
{
    require(tidyverse)
    satir<-nrow(veri)
    oncekigunIt<-c(0,veri$It[1:satir-1])
    veri<-veri %>%
     as_tibble() %>%
     mutate(vakalar = It - oncekigunIt)
    if ("I_tahmin" %in% names(veri)) {
     oncekigunIt_tah<-c(0,veri$I_tahmin[1:satir-1])
```

```
  veri<-veri %>%
    mutate(vakalar_tahmin = I_tahmin - oncekigunIt_tah)
  }
  return(veri)
}



haftalikvaka<-function(veri,gunsayisi=7) {
  require(tidyverse)
  hafta<-floor(nrow(veri)/gunsayisi)
  kalan<-nrow(veri) - gunsayisi*hafta
  eklenen<-seq_along(vector("double",kalan))

  veri2<-veri %>%
   mutate(wk=c(rep(c(1:gunsayisi),hafta),eklenen)) %>%
   filter(wk==gunsayisi) %>%
   select(!wk) %>%
   gunlukvaka(.) %>%
   rename(haftalik=vakalar)

  veri2$haftalik[1]<-veri$It[gunsayisi] - veri$It[1]

  if ("I_tahmin" %in% names(veri)) {
   veri2<-veri2 %>% rename(haftalik_tahmin=vakalar_tahmin)
   veri2$haftalik_tahmin[1]<-veri$I_tahmin[gunsayisi] - veri$I_tahmin[1]
  }
  return(veri2)
}

rezidu_hesapla<-function(veri) {
 # veriden yapilan tüm tahminlerin rezidulerini hesaplar
  require(tidyverse)
  veri<-veri %>%
   mutate(S_rezidu=St-S_tahmin,
        S_goreceli_rezidu=round(S_rezidu/St*100,1),
        I_rezidu=It-I_tahmin,
        I_goreceli_rezidu=round(I_rezidu/It*100,1),
        R_rezidu=Rt-R_tahmin,
        R_goreceli_rezidu=round(R_rezidu/Rt*100,1))

  if ("vakalar" %in% names(veri)) {
   veri<-veri %>%
    mutate(vakalar_rezidu=vakalar-vakalar_tahmin,
         vakalar_goreceli_rezidu=round(vakalar_rezidu/vakalar*100,1))
  }
  if ("haftalik" %in% names(veri)) {
   veri<-veri %>%
```

```
        mutate(haftalik_rezidu=haftalik-haftalik_tahmin,
               haftalik_goreceli_rezidu=round(haftalik_rezidu/haftalik*100,1))
    }
    return(veri)
}


rezidu_istatistik<-function(veri)
{
    require(tidyverse)
    n<-nrow(veri)
    snc<-veri %>%
      select(matches("rezidu")) %>%
      select(!matches("goreceli_rezidu")) %>%.^2 %>%
      mutate_if(is.numeric, list(~na_if(., Inf))) %>%
      summarize_all(list(sum=sum),na.rm=TRUE) %>%
      format(scientific=FALSE,digits=2) %>%
      unlist() %>% as.double() %>% matrix(nrow=4,ncol=1) %>%
      as.tibble() %>% rename(SSE=V1) %>% mutate(MSE=SSE/(n-2)) %>%
      as.data.frame()
    rownames(snc)<-c("S","I","R","Vakalar")
    return(snc)
}



##REZİDU
rezidu_analizi_SIR<-function(veri) {
  require(tidyverse)
  options(scipen=999)
  n<-nrow(veri)

  gv<-gunlukvaka(veri)
  hv<-haftalikvaka(veri)
  n2<-nrow(hv)
  if ((n2 - 2) < 1) n2<-4

  gv<-gv  %>%
    mutate(vakalar_rezidu=vakalar-vakalar_tahmin,
           vakalar_goreceli_rezidu=round(vakalar_rezidu/vakalar*100,1))

  hv<-hv %>%
    mutate(haftalik_rezidu=haftalik-haftalik_tahmin,
           haftalik_goreceli_rezidu=round(haftalik_rezidu/haftalik*100,1))

  snc1<-gv %>%
    select(matches("rezidu")) %>% .^2 %>%
    mutate_if(is.numeric, list(~na_if(., Inf))) %>%
```

```
  summarize_all(list(sum=sum),na.rm=TRUE) %>%
  format(scientific=FALSE,digits=2) %>%
  unlist() %>% as.double() %>% matrix(nrow=2,ncol=1) %>%
  as_tibble() %>% rename(Gunluk_SSE=V1) %>%
  mutate(Gunluk_MSE=Gunluk_SSE/abs(n-2),
         Gunluk_RMSE=sqrt(Gunluk_MSE)) %>%
  as.data.frame()

 snc2<-hv %>%
  select(matches("rezidu")) %>% .^2 %>%
  mutate_if(is.numeric, list(~na_if(., Inf))) %>%
  summarize_all(list(sum=sum),na.rm=TRUE) %>%
  format(scientific=FALSE,digits=2) %>%
  unlist() %>% as.double() %>% matrix(nrow=2,ncol=1) %>%
  as_tibble() %>% rename(Haftalik_SSE=V1) %>%
  mutate(Haftalik_MSE=Haftalik_SSE/abs(n2-2),
         Haftalik_RMSE=sqrt(Haftalik_MSE)) %>%
  as.data.frame()
 snc<-cbind(snc1,snc2)
 # snc<-snc %>% select(!matches("SSE"))
 rownames(snc)<-c("Rezidu","Goreceli Rezidu")
 return(snc)
}


rezidu_analizi_SIR_tr<-function(gun,bas1,kac1,bas2,kac2,bas3,kac3) {

 tr1<-tumtahmin(turkiye,bas1,gun,kacgunluk = kac1)
 tr2<-tumtahmin(turkiye,bas2,gun,kacgunluk = kac2)
 tr3<-tumtahmin(turkiye,bas3,gun,kacgunluk = kac3)
 n<-nrow(tr1)+nrow(tr2)+nrow(tr3)
 n2<-nrow(rbind(haftalikvaka(tr1),haftalikvaka(tr2),haftalikvaka(tr3)))
 snc_gun<-rezidu_analizi_SIR(tr1) %>% select(matches("Gunluk_SSE")) +
  rezidu_analizi_SIR(tr2) %>% select(matches("Gunluk_SSE")) +
  rezidu_analizi_SIR(tr3) %>% select(matches("Gunluk_SSE"))
 snc_haf<-rezidu_analizi_SIR(tr1) %>% select(matches("Haftalik_SSE")) +
  rezidu_analizi_SIR(tr2) %>% select(matches("Haftalik_SSE")) +
  rezidu_analizi_SIR(tr3) %>% select(matches("Haftalik_SSE"))
 snc_gun<-snc_gun %>% mutate(
  Gunluk_MSE=Gunluk_SSE/(n-2),
  Gunluk_RMSE=sqrt(Gunluk_MSE)
 )
 snc_haf<-snc_haf %>% mutate(
  Haftalik_MSE=Haftalik_SSE/(n2-2),
  Haftalik_RMSE=sqrt(Haftalik_MSE)
 )
 snc<-cbind(snc_gun,snc_haf) %>% select(matches("RMSE"))
 return(snc)
```

```
}


##ARIMA
arima_fonk<-function(ot,start,frekans,bas,bit,kacgun) {
 require(forecast)
 require(tidyverse)
 veri<-ts(ot,start=start,frequency = frekans)
 model<-auto.arima(veri[bas:bit])
 ongoruler<-forecast(model,h=kacgun)

 vakalar<-ot[bas:(bit+kacgun)]
 tah_ongoru<-c(model$fitted,ongoruler$mean)
 gtah_ongoru<-rbind(matrix(c(NA),ncol=4,nrow=(bit-
bas+1)),cbind(ongoruler$lower,ongoruler$upper))
 veritum<-cbind(vakalar,tah_ongoru,gtah_ongoru)
 veritum<-data.frame(veritum)
 colnames(veritum)<-c("Vakalar","Tahminler","lo80","lo95","hi80","hi95")

 veritum<-veritum %>%
  mutate(rezidu=Vakalar - Tahminler,
       goreceli_rezidu = round(rezidu / Vakalar * 100 , 1) )

 bit<-bit-bas
 bas<-1
 snc1<-rezidu_hesapla_ARIMA(veritum,bas,bit) %>% select(RMSE)
 snc2<-rezidu_hesapla_ARIMA(veritum,bit+1,bit+kacgun) %>% select(RMSE)
 snc3<-rezidu_hesapla_ARIMA(veritum,bas,bit+kacgun) %>% select(RMSE)
 snc<-cbind(snc1,snc2,snc3)
 colnames(snc)<-c("RMSE_tahmin","RMSE_Ongoru","RMSE_Toplam")
 snclar<-ongoruler
 snclar$reziduler<-snc
 snclar$veri<-veritum
 snclar$model<-model

 return(snclar)
}

veriArima_tr<-function(ot,veri,periyod,bas1,bit1,bas2,bit2,bas3,bit3,kacgun) {
 require(tidyverse)
 snc1<-arima_fonk(ot,2020,periyod,bas1,bit1,5)
 snc2<-arima_fonk(ot,2020,periyod,bas2,bit2,5)
 snc3<-arima_fonk(ot,2020,periyod,bas3,bit3,kacgun)
 tarih_ve_vakalar1<-veri[bas1:bit1,c(1,6)]
 tarih_ve_vakalar2<-veri[bas2:bit2,c(1,6)]
 tarih_ve_vakalar3<-veri[bas3:(bit3+kacgun),c(1,6)]
 tarih_ve_vakalar<-
rbind(tarih_ve_vakalar1,tarih_ve_vakalar2,tarih_ve_vakalar3)
```

```
  mtahminler1<-snc1$fitted
  mtahminler2<-snc2$fitted
  mtahminler3<-snc3$fitted
  mtahminler<-c(mtahminler1,mtahminler2,mtahminler3)
  mtsize<-length(mtahminler)
  mtahminler<-c(mtahminler,snc3$mean)
  gtahminler<-
rbind(matrix(c(NA),ncol=4,nrow=mtsize),cbind(snc3$lower,snc3$upper))
  veritum<-cbind(tarih_ve_vakalar,mtahminler,gtahminler)
  colnames(veritum)<-
c("Tarih","Vakalar","Tahminler","lo80","lo95","hi80","hi95")
  return(veritum)
 }


rezidu_hesapla_ARIMA<-function(veri,bas,bit) {
  n<-bit-bas+1
  snc<-veri[bas:bit,] %>%
   select(matches("rezidu")) %>% .^2 %>%
   mutate_if(is.numeric, list(~na_if(., Inf))) %>%
   summarize_all(list(sum=sum),na.rm=TRUE) %>%
format(scientific=FALSE,digits=2) %>%
   unlist() %>% as.double() %>% matrix(nrow=2,ncol=1) %>%
   as_tibble() %>% rename(SSE=V1) %>%
   mutate(MSE=SSE/abs(n-2),
        RMSE=sqrt(MSE)) %>%
   as.data.frame()
  rownames(snc)<-c("Rezidu","Goreceli Rezidu")
  return(snc)
 }

 ARIMA_rezidu<-function(veri) {
  options(scipen=999)
  require(tidyverse)
  n<-nrow(veri)
  fc_n<-sum(!is.na(veri$lo80))
  th_n<-n - fc_n
  veri<-veri %>%
   mutate(rezidu=Vakalar - Tahminler,
        goreceli_rezidu = round(rezidu / Vakalar * 100 , 1) )
  snc1<-rezidu_hesapla_ARIMA(veri,1,th_n) %>% select(RMSE)
  snc2<-rezidu_hesapla_ARIMA(veri,th_n+1,n) %>% select(RMSE)
  snc3<-rezidu_hesapla_ARIMA(veri,1,n) %>% select(RMSE)
  snc<-cbind(snc1,snc2,snc3)
  colnames(snc)<-c("RMSE_tahmin","RMSE_Ongoru","RMSE_Toplam")
  return(round(snc,digits=2))
 }
```

# ÖZGEÇMİŞ

1992 yılında Adana'da doğdum. Üniversiteye kadar olan eğitimimi Adana'da tamamladıktan sonra 2011 yılında girdiği Orta Doğu Teknik Üniversitesi İstatistik bölümünden 2017 yılında mezun oldum. 2017-2021 seneleri arasında iki farklı firmada dış ticaret uzmanı olarak görev aldım. 2020 güz döneminde, Çukurova Üniversitesi Sağlık Bilimleri Enstitüsü'ne bağlı Biyoistatistik Anabilim Dalı'nda Yüksek Lisans öğrenimine başladım. Evli ve bir çocuk sahibiyim.