

**T.C.
AKDENİZ ÜNİVERSİTESİ**



**KONTROL GÜNÜ SÜT VERİMLERİNİN BAYESIAN ZAMAN SERİSİ
YÖNTEMİ İLE MODELLENMESİ**

Özge KOZAKLI

FEN BİLİMLERİ ENSTİTÜSÜ

ZOOTEKNİ

ANABİLİM DALI

YÜKSEK LİSANS

TEMMUZ 2020

ANTALYA

T.C.
AKDENİZ ÜNİVERSİTESİ



**KONTROL GÜNÜ SÜT VERİMLERİNİN BAYESIAN ZAMAN SERİSİ
YÖNTEMİ İLE MODELLENMESİ**

Özge KOZAKLI

FEN BİLİMLERİ ENSTİTÜSÜ

ZOOTEKNİ

ANABİLİM DALI

YÜKSEK LİSANS

TEMMUZ 2020

ANTALYA

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

**KONTROL GÜNÜ SÜT VERİMLERİNİN BAYESIAN ZAMAN SERİSİ
YÖNTEMİ İLE MODELLENMESİ**

**Özge KOZAKLI
FEN BİLİMLERİ ENSTİTÜSÜ
ZOOTEKNİ
ANABİLİM DALI
YÜKSEK LİSANS**

**Bu tez Akdeniz Üniversitesi Bilimsel Araştırma Projeleri Koordinasyon Birimi
tarafından FYL-2020-5127 proje numarası ile desteklenmiştir.**

TEMMUZ 2020

T.C.
AKDENİZ ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

KONTROL GÜNÜ SÜT VERİMLERİNİN BAYESIAN ZAMAN SERİSİ
YÖNTEMİ İLE MODELLENMESİ

Özge KOZAKLI

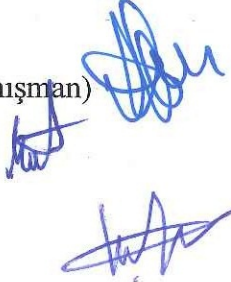
ZOOTEKNİ
ANABİLİM DALI
YÜKSEK LİSANS TEZİ

Bu tez 01/07/2020 tarihinde jüri tarafından Oybirliği ile kabul edilmiştir.

Prof. Dr. Mehmet Ziya FIRAT(Danışman)

Doç. Dr. Burak KARACAÖREN

Doç. Dr. Özgür KOŞKAN



ÖZET

KONTROL GÜNÜ SÜT VERİMLERİNİN BAYESİAN ZAMAN SERİSİ YÖNTEMİ İLE MODELLENMESİ

Özge KOZAKLI

Yüksek Lisans, Zootekni Anabilim Dalı

Danışman: Prof. Dr. Mehmet Ziya FIRAT

TEMMUZ 2020; 76 Sayfa

Bu çalışmanın amacı zaman serisi modelleri ile kontrol günü süt verimleri kullanılarak laktasyon eğrisi modelleme çalışmalarına Bayesian yaklaşım ile yeni bir alternatif geliştirmek ve Frekansçı zaman serisi modeli ile karşılaştırmaktır. Çalışmada 1.070 hayvana ait 10'ar laktasyon süt verimi kaydı kullanılmış ve hayvanlar rastgele sıralanarak veri seti oluşturulmuştur. Oluşan veri seti eşit sayıda hayvan içerecek şekilde ikiye ayrılarak ilk grup veri analiz için, ikinci grup veri, öngörü başarısının incelenmesi için 535'er gözleme sahip olacak şekilde oluşturulmuştur. Veri oluşturma yönteminden kaynaklı mevsimsel bileşenin deterministik yapısı gereği çarpımsal olarak klasik ayrıştırma ile modelden ayrıştırılmıştır. Yapılan analiz sonucunda *ARIMA*(1,0,1) modeline karar verilerek, modelim parametreleri her iki yaklaşım içinde tahmin edilerek karşılaştırılmış ve Bayesian yaklaşım ile tahmin edilen modelin küçük farklarla veri seti için daha uygun olduğuna karar verilmiştir. Bayesian zaman serisi modeli ile yapılan tahmin değerlerinin gerçek değerler ile korelasyon katsayıları literatürden daha yüksek bulunmuştur. Laktasyon eğrisinin tahmin edilmesi için ideal kontrol günü sayısına 5 olarak karar verilmiştir. Böylece Bayesian zaman serisi modelinin laktasyon eğrisi tahmininde alternatif bir yaklaşım olarak kullanılabilceği sunulmuştur.

ANAHTAR KELİMELER: ARIMA, Bayesian Zaman Serisi, Box-Jenkins Yöntemi, Kontrol Günü, Laktasyon Eğrisi, Süt Verimi, Öngörü

JÜRİ: Prof. Dr. Mehmet Ziya FIRAT

Doç. Dr. Burak KARACAÖREN

Doç. Dr. Özgür KOŞKAN

ABSTRACT

MODELING MILK YIELDS OF CONTROL DAY BY BAYESIAN TIME SERIES METHOD

Özge KOZAKLI

MSc Thesis in Animal Science

Supervisor: Prof. Dr. Mehmet Ziya FIRAT

July 2020; 76 pages

The aim of this study is to develop a new Bayesian approach to lactation curve modeling studies using test-day milk yields with time series and to compare it with the Frequentist time series model. 10 test day records of 1,070 daily cattle were used in this study and the data set was created by randomly ordering the animals. Then, the data set was split into two, and the first dataset was used to analyze, second was used to examine forecast accuracy. The seasonal component which occurs due to the form of the dataset was separated from data with classical decomposition. As a result of the analysis, ARIMA(1,0,1) was selected and the parameters of this model were estimated with Frequentist and Bayesian approaches and although there were small differences, it was decided that the model which estimated the Bayesian approach was more suitable for the dataset. The correlation coefficients of the predictions of the Bayesian time series model were calculated and better correlation coefficients found than literature and the optimal test day count was found as 5 to estimate lactation curve. Thus, it is presented that the Bayesian time series model can be used as an alternative approach in estimating the lactation curve.

KEYWORDS: ARIMA, Bayesian Time Series, Box-Jenkins Method, Predict, Test day, Milk yield, Lactation curve

COMMITTEE: Prof. Dr. Mehmet Ziya FIRAT

Assoc. Prof. Dr. Burak KARACAÖREN

Assoc. Prof. Dr. Özgür KOŞKAN

ÖNSÖZ

Her zaman desteđini hissettiđim, bilgi ve tecrübeleri ile bana yol gösteren deđerli danışman hocam Prof. Dr. Mehmet Ziya FIRAT'a, çalışmamı inceleyerek önemli katkı ve düzeltmelerde bulunan Prof. Dr. Mehmet MERT'e, her ihtiyaç duyduğumda bilgisini esirgemeyen ve bilgi birikimi ile bana destek olan aynı alanda çalıştığım deđerli arkadaşım Murat KOPTUR'a, tecrübelerini benimle paylaşan ve deđerli zamanını ayıran Aarhus Üniversitesi post doktora öğrencisi deđerli hocam Emre KARAMAN'a, çalışma için maddi destek sağlayan Akdeniz Üniversitesi Bilimsel Araştırma Projeleri Koordinasyon Birimi'ne, çalışma arkadaşlarıma, Akdeniz Üniversitesi Ziraat Fakültesi Zootekni bölümü akademisyen kadrosuna ve hayatımın her döneminde desteklerini hiçbir zaman esirgemeyerek kişisel gelişimimden sorumlu deđerli aileme teşekkür ederim.

İÇİNDEKİLER

ÖZET	i
ABSTRACT	ii
İÇİNDEKİLER.....	iv
AKADEMİK BEYAN	vii
SİMGELER VE KISALTMALAR	viii
ŞEKİLLER DİZİNİ	xi
ÇİZELGELER DİZİNİ	xiii
1. GİRİŞ	1
2. KAYNAK TARAMASI.....	4
2.1. Deterministik ve Stokastik Modeller.....	4
2.2. Tek Değişkenli Zaman Serileri	4
2.2.1. Zaman Serilerine Etki Eden Faktörler	5
2.2.1.1. Trend Bileşeni (T)	5
2.2.1.2. Mevsimsel dalgalanmalar (S).....	6
2.2.1.3. Konjonktürel (döngüsel-çevrimsel) dalgalanmalar (C)	7
2.2.1.4. Düzensiz hareketler (Hata Terimi)(E)	7
2.2.1.5. Ayırıştırma yöntemleri.....	8
2.2.2. Zaman Serilerinin Sınıflandırılması	9
2.2.2.1. Kesikli ve Sürekli Zaman Serileri	10
2.2.2.2. Tek Değişkenli Zaman Serilerinde Durağanlık.....	10
2.2.2.3. Mevsimsel ve Mevsimsel Olmayan Zaman Serileri.....	10
2.2.3. Zaman Serilerinde Durağanlık Sınamaları.....	11
2.2.3.1. Arttırılmış Dickey-Fuller Testi (ADF)	12
2.2.3.2. Phillips-Perron Testi (PP) (1988)	13
2.2.3.3. Kwiatkowski, Phillips, Schmidt, Shin Testi (KPSS)(1992).....	14
2.2.4. Zaman Serilerinde Otokorelasyon	14
2.2.4.1. Otokovaryans Fonksiyonu	15
2.2.4.2. Otokorelasyon Fonksiyonu (ACF)	15
2.2.4.3. Kısmi Otokorelasyon Fonksiyonu	16
2.2.5. Zaman Serisi Modelleri	17
2.2.5.1. Otoregresif Model AR(p)	18
2.2.5.2. Hareketli Ortalama Modeli MA(q).....	20

2.2.5.3. Karma Otoregresif Hareketli Ortalama Modeli ARMA(p, q)	21
2.2.5.4. Bütünlenen Otoregresif Hareketli Ortalama Modeli ARIMA(p, d, q) .	22
2.2.5.5. Mevsimsel Otoregresif Modeller SAR(P).....	22
2.2.5.6. Mevsimsel Hareketli Ortalama Modeli SMA(Q).....	23
2.2.5.7. Mevsimsel Karma Otoregresif Hareketli Ortalama Model SARMA(P, Q)s	23
2.2.5.8. Mevsimsel Bütünlenen Otoregresif Hareketli Ortalama Modeli ARIMA(P, D, Q)s	24
2.2.6. Zaman Serisi Modeli Seçimi	24
2.2.6.1. Belirlenme	25
2.2.6.2. Tahmin	26
2.2.6.3. Ayırt Edici Kontrol (Artık Analizi)	27
2.2.6.4. Büyük Ayrım	27
2.2.6.5. Öngörü.....	28
2.3. Bayesian Analizi	28
2.3.1. Bayesian Yaklaşımı ile Zaman Serileri Analizi	29
2.3.1.1. Önsel dağılım.....	30
2.3.1.2. Sonsal dağılımın tahmini.....	32
2.3.1.3. Tanı testleri.....	37
2.3.2. Bayesian Zaman Serisi Modelleri	38
2.3.2.1. Bayesian Otoregresif Model AR(1).....	39
2.3.2.2. Hareketli ortalama Modeli MA(1).....	40
2.3.2.3. Karma Otoregresif Hareketli Ortalama Modeli ARMA(1,1)	41
2.4. Zaman Serilerinin Parametre Tahmin Yöntemleri.....	42
2.4.1. Frekansçı Yaklaşım ile Parametre Tahmini	43
2.4.2. Bayesian Yaklaşım ile Parametre Tahmini	46
3. METERYAL VE METOT	49
3.1. Materyal.....	49
3.2. Metot	49
4. BULGULAR	52
4.1. Belirleme	52
4.2. Tahmin: Frekansçı Yaklaşım ile Parametre Tahmini.....	57
4.3. Ayırt Edici Kontrol: Frekansçı Yaklaşım Modeli Kalıntıların Kontrolü	58
4.4. Tahmin: Bayesian Yaklaşım ile Parametre Tahmini	61

4.5. Ayırt Edici Kontrol: Bayesian Yaklaşım Modeli Kalıntılarının Kontrolü	66
4.6. Öngörü: Frekanscı Model ile Öngörü	69
4.7. Öngörü: Bayesian Model ile Öngörü	70
5. TARTIŞMA.....	72
6. SONUÇLAR.....	75
7. KAYNAKLAR.....	76
ÖZGEÇMİŞ	

AKADEMİK BEYAN

Yüksek Lisans Tezi olarak sunduğum “Kontrol Günü Süt Verimlerinin Bayesian Zaman Serisi Yöntemi ile Modellenmesi ” adlı bu çalışmanın, akademik kurallar ve etik değerlere uygun olarak yazıldığını belirtir, bu tez çalışmasında bana ait olmayan tüm bilgilerin kaynağını gösterdiğimi beyan ederim.

01/07/2020

Özge KOZAKLI

SİMGELER VE KISALTMALAR

Simgeler

- a : model artıkları
- B : geri kaydırma işlemcisi
- β : regresyon katsayılarına ait parametre vektörü
- d : regular fark işlemcisi
- D : mevsimsel fark işlemcisi
- e : doğal logaritma tabanı
- E : beklenen değer
- k : gecikme sayısı
- K : kinetik enerji gösterimi
- L : örnekleme için adım sayısı
- \mathcal{L} : olabilirlik fonksiyonu
- μ : popülasyon ortalaması
- m : Markov Zinciri sayısı
- \mathcal{M} : simetrik kareler toplamı
- n : Markov Zinciri uzunluğu
- \hat{n}_{eff} : efektif örneklem büyüklüğü istatistiği
- p : mevsimsel olmayan otoregresif parametre sayısı
- P : olasılık fonksiyonu
- P_j : Jeffreys önseli ile elde edilen olasılık fonksiyonu
- P_I : Improper önseli ile elde edilen olasılık fonksiyonu
- P_U : uniform önseli ile elde edilen olasılık fonksiyonu
- \hat{R} : Gelman Rubin istatistiği
- s : mevsimsel fark periyodu
- S : kalıntı kareler toplamı

- t : gözlem sırası
T : gözlem sayısı
U : potansiyel enerji gösterimi
I : Fisher bilgi matrisi
 δ : yığılım parametresi
 σ^2 : varyans
 γ : kovaryans
 ρ : popülasyon korelasyon katsayısı
 $\hat{\rho}$: örnek korelasyon katsayısı
 φ : otoregresif parametre
 θ : hareketli ortalama parametresi
 Δ : fark alma işlemcisi
 ∂ : türev alma işlemcisi
 ε : örnekleme için adım büyüklüğü
 Σ : varyans-kovaryans matrisi

Tezde ondalık ayıraç olarak “,” virgül kullanılmıştır.

Kısaltmalar

- ADF : Artırılmış Dickey, Fuller Testi
PP : Phillips, Perron Testi
KPSS : Kwiatkowski, Phillips, Schmidt, Shin Testi
AR : Otoregresif (Autoregressive) Model
ARMA : Karma Otoregresif Hareketli Ortalama Modeli
ARIMA : Bütünlenen Otoregresif Hareketli Ortalama Modeli
B : Geri kaydırma işlemcisi (Backward shift operator)

E-BFMI	: Energy-Bayesian Fraction of Missing Information
HMC	: Hamiltonian Monte Carlo
KOKF	: Kısmi otokorelasyon fonksiyonu
MA	: Hareketli Ortalama Modeli
MAPE	: Hataların mutlak yüzde ortalaması
MAE	: Hataların mutlak ortalaması
MCMC	: Markov Zinciri Monte Carlo
NUTS	: No-U-Turn Örnekleycisi
OK	: Otokorelasyon
OKF	: Otokorelasyon fonksiyonu
OKVF	: Otokorelasyon fonksiyonu
ÖOKF	: Örnek otokorelasyon fonksiyonu
ÖKOKF	: Örnek kısmi otokorelasyon fonksiyonu
RMSE	: Hata kareler ortalamasının karekökü
SAR	: Mevsimsel Otoregresif Model
SARMA	: Mevsimsel Karma Otoregresif Hareketli Ortalama Modeli
SARIMA	: Mevsimsel Bütünlenen Otoregresif Hareketli Ortalama Modeli
SMA	: Mevsimsel Hareketli Ortalama Modeli
TÜİK	: Türkiye İstatistik Kurumu

ŞEKİLLER DİZİNİ

Şekil 1.1. İneklerde tipik laktasyon eğrisi.....	2
Şekil 2.1. Trend yapıları; a) Doğrusal artan trend yapısı; b) Doğrusal azalan trend yapısı; c) Doğrusal olmayan artan trend yapısı; d) Doğrusal olmayan azalan trend yapısı	6
Şekil 2.2. Saf mevsim bileşenin zaman yolu grafiği	7
Şekil 2.3. Hata teriminin zaman yolu grafiği	8
Şekil 2.4. Bileşenlerinin toplamından ve çarpımından oluşan seri örnekleri; a) Toplamsal model grafiği; b) Çarpımsal model grafiği	9
Şekil 2.5. Model belirleme algoritması.....	17
Şekil 2.6. Koşullu olasılık şeması.....	29
Şekil 2.7. Önsel seçimi kriterleri	31
Şekil 2.8. Başlıca sonsal dağılım hesaplama yöntemleri	33
Şekil 2.9. Üç farklı örnekleme yöntemine ait örnekleme uzayları grafikleri; a) Metropolis-Hastings Yöntemi; b) Gibbs Örneklemesi; c) NUTS yöntemi d) Gerçek örnekleme uzayı	34
Şekil 2.10. İki boyutlu bir olasılık yoğunluk fonksiyonu	35
Şekil 2.11. Örnekleme uzayı gösterimleri; a) HMC yöntemi örneklemleri; b) NUTS düzeltmesinden sonra örnekleme uzayı	36
Şekil 2.12. Birden fazla merkezli sonsal dağılım için NUTS yönteminin örnekleme uzayı	36
Şekil 2.13. Bağımsız 4 Markov Zincirine ait tek Trace grafiği.....	37
Şekil 2.14. Yakınsama problemi taşıyan Markov Zinciri	37
Şekil 4.1. Veri setinin orijinal zaman serisi grafiği	53
Şekil 4.2. İlk 50 gözlem için orijinal seri için zaman serisi grafiği.....	53
Şekil 4.3. Mevsimsellikten arındırılmış zaman serisi grafiği.....	54
Şekil 4.4. İlk 50 gözlem için arındırılmış seri için zaman serisi grafiği	55
Şekil 4.5. Seriyeye ait otokorelasyon grafiği	56
Şekil 4.6. Seriyeye ait kısmi otokorelasyon grafiği	57
Şekil 4.7. R paket programı en iyi model seçimi kodu ve sonucu	57
Şekil 4.8. Frekanscı modele ait tanımlayıcı istatistikler	58
Şekil 4.9. Gerçek (kırmızı) ve öngörü (mavi) değerlerinin zaman yolu grafikleri	58
Şekil 4.10. Frekanscı zaman serisi modeline ait kalıntıların zaman serisi grafiği	59
Şekil 4.11. Frekanscı modelin kalıntılarına ait histogram grafiği ve normal dağılım eğrisi	59
Şekil 4.12. Frekanscı modelin kalıntılarına ait serisinin otokorelasyon katsayıları grafiği.....	60
Şekil 4.13. Frekanscı modelin kalıntıları serisine ait kısmi otokorelasyon katsayıları grafiği	60
Şekil 4.14. Frekanscı modelin kalıntılarına ait Box-Pierce Test sonuçları.....	61
Şekil 4.15. Sabit terime ait sonsal dağılım grafiği.....	62

Şekil 4.16. Tahminlere ait otoregresif parametre için sonsal dağılım grafiği	62
Şekil 4.17. Kalıntılara ait otoregresif parametre için sonsal dağılım grafiği	62
Şekil 4.18. Kalıntılar için sonsal dağılım grafiği.....	63
Şekil 4.19. Markov Zincirleri grafikleri a) Sabit terim; b) Tahminlere ait otoregresif parametreler; c) Kalıntılara ait otoregresif parametreler; d) Kalıntılar	63
Şekil 4.20. Iraksama, Ağaç Derinliği ve E-BFMI sına R raporu.....	64
Şekil 4.21. Bayesian yaklaşım ile model parametrelerinin tahminin R programı sonuçları	65
Şekil 4.22. Gerçek ve öngörü değerlerinin zaman yolu grafikleri	66
Şekil 4.23. Bayesian zaman serisi modeline ait kalıntıların zaman serisi grafiği	66
Şekil 4.24. Bayesian modelin kalıntılarına ait histogram grafiği ve normal dağılım eğrisi	67
Şekil 4.25. Bayesian modelin kalıntılarına ait serisine ait otokolerasyon katsayıları grafiği	67
Şekil 4.26. Bayesian modelin kalıntılarına serisine ait kısmi otokolerasyon katsayıları grafiği	68
Şekil 4.27. Bayesian modelin kalıntılarına ait Box-Pierce Test sonuçları.....	68
Şekil 5.1. Macciotta vd. (2002) çalışmasında kontrol günleri süt miktarlarının öngörüsü için kullanılan zaman serisi modeli tahminlerinin gerçek değerleri ile ilişki matrisi	72
Şekil 5.2. Karaman (2010) çalışmasında kontrol günleri süt miktarlarının öngörüsü için kullanılan zaman serisi modeli tahminlerinin gerçek değerleri ile ilişkisi matrisi	73

ÇİZELGELER DİZİNİ

Çizelge 1.1. Sığırların adlandırılma kriterleri	1
Çizelge 2.1. ARIMA modellerinde gecikme uzunluklarının belirlenmesi	26
Çizelge 2.2. ARIMA modellerinde gecikme uzunluğu belirlenmesinde örnek durumlar	26
Çizelge 2.3. Parametre tahmini yaklaşımlarının karşılaştırılması	42
Çizelge 2.4. Zamanın fonksiyonu olan veri setleri ile laktasyon dönemi modelleri	45
Çizelge 2.5. Laktasyon dönemi mevsimsel modelleri	46
Çizelge 3.1. Veri seti	50
Çizelge 3.2. Öngörü algoritması	51
Çizelge 4.1. Veri setinin tanımlayıcı istatistikleri	52
Çizelge 4.2. Toplamsal ve çarpımsal ayrıştırma yöntemlerine ait istatistikler	54
Çizelge 4.3. Durağanlık sınamalarının özet sonuçları	55
Çizelge 4.4. Etkili örneklem büyüklüklerinin zincir uzunlukları toplamına oranı	64
Çizelge 4.5. Öngörü algoritmasında kullanılan gözlem numaraları	69
Çizelge 4.6. Mevcut kontrol günleri ile Frekanscı öngörü değerleri arasındaki korelasyonlar	70
Çizelge 4.7. Mevcut kontrol günleri ile Bayesian öngörü değerleri arasındaki korelasyonlar	71

1. GİRİŞ

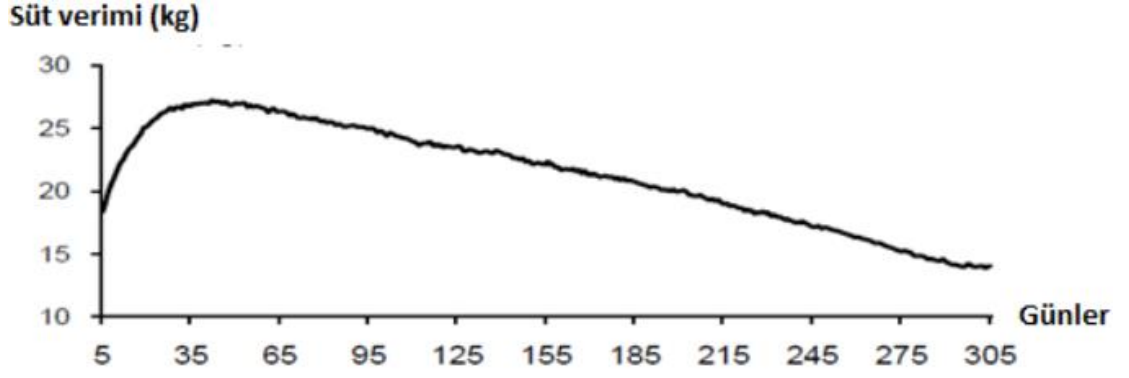
Tarım sektörü için oldukça büyük bir potansiyele sahip olan hayvancılığın, ülke ekonomisi içerisinde yeri oldukça önemlidir. Hayvancılık, yan ürünlerin değerlendirilmesinde, işgücü verimliliğinin artırılmasında ve çok yönlü işletmelerin çeşitli sebeplerden kaynaklanan risk faktörünün azalmasında olumlu katkılar sağlamayan bir faktördür. Hayvancılık, artan nüfusun protein içerikli beslenmesi yönünden ayrıca önem taşımaktadır. Gelişmiş ülkelerde endüstri kolu olan hayvancılık, ekonominin önemli bir parçası olmuştur. Bu demek oluyor ki hayvancılık insan hayatında önemini artırarak sürdürecektir. Türkiye'nin ekonomisinde ve insan beslenmesinde hayvancılığın yeri yadsınmaz. Türkiye'de hayvancılığın iyileştirilmesi için mevcut durumun en iyi şekilde incelenmesi ve hayvancılık faaliyetlerinin sürdürülebilirliğinin sağlanması gerekmektedir (Aydemir ve Pıçak 2007; Vural ve Fidan 2007).

Hayvansal üretimde üretimi etkileyen birçok faktör ve ürün olarak birçok çıktı bulunmaktadır. Bu çalışmada hayvansal ürünlerden biri olan inek sütünü ele alarak kontrol günü süt verimlerinin tahmini ile ilgilenildi. 2019 yılı TÜİK verilerine göre; 2018 yılında süt üretimi bir önceki yıla göre yüzde 6,86 artarak 22 milyon 120 bin 716 ton olarak gerçekleşmiştir. Bu miktarın yüzde 90,58'ini sığır sütü, yüzde 6,54'unu koyun sütü, yüzde 2,54'ünü keçi sütü ve yüzde 0,33'ünü manda sütü oluşturmuştur. Türkiye'de süt üretiminin çok büyük bir kısmı sığırlardan sağlanmaktadır (Anonim1). Kavram kargaşasına neden olmamak için yaş ve çiftleştirilmiş olma durumlarına göre sığırların nasıl adlandırıldıkları Çizelge 1.1'de bahsedilmiştir.

Çizelge 1.1. Sığırların adlandırılma kriterleri (Anonim2)

Buzağı	0 – 6 aylık inek yavrusuna denir. Erkek veya dişi buzağı olabilir.
Dana	6 – 15 aylık inek yavrusuna denir. Erkek dana veya dişi dana olabilir.
Düve	Henüz doğum yapmamış dişi inek yavrusuna denir.
Tosun	15 aylık – 2 yaş arası erkek inek yavrusuna denir.
İnek	2 yaş ve üzeri dişi sığırlara denir.
Boğa	2 yaş ve üzeri erkek sığırlara denir.
Öküz	Kısırlaştırılmış boğalara denir.

Sığırlarda süt verimi, hayvanın doğum yapmasıyla başlar ve 305 gün devam eder. Ortalama 305 gün olarak kabul edilen bu süreye *laktasyon dönemi* denir. Laktasyon döneminde süt verimi çevresel ve genetik faktörlerden etkilense de genel olarak benzer bir seyir izler. 7. ve 8. haftada maksimum süt verimine ulaşılmış olur ve daha sonraki haftalarda belirli bir oranda azalarak devam eder. Süt miktarının zamana karşı oluşturduğu bu seyre *laktasyon eğrisi* adı verilir. Şekil 1.1'de tipik laktasyon eğrisi grafiği örneği yer almaktadır. Laktasyon eğrisi sivri yapılı olan inekler, daha düz laktasyon eğrisine sahip ineklere göre daha az kesif yeme ihtiyaç gösterirler. Bu nedenle besleme daha ekonomik ve az (üreme ve metabolik) hastalık riski taşıdıkları için daha yüksek döl verimine sahiptirler. Bu gibi nedenler ile laktasyon eğrisi hayvancılık çalışmalarında önemli bir yol göstericidir (Sönmez Oskay, 2016).



Şekil 1.1. İneklerde tipik laktasyon eğrisi (Sönmez Oskay 2016)

Laktasyon eğrisinin tipik şekline etki eden faktörlerden bazıları; servis periyodu, buzağılama mevsimi ve laktasyon sırasındır. Bu faktörlerle birlikte laktasyon eğrisinin genel seyri kalıtsal bir özellik olduğu ve eğrinin eğiminin büyük ölçüde kalıttan etkilendiği bilinmektedir. Süt verimi yüksek nesillerin, elde edilebilmesi için yapılan seleksiyon çalışmalarında, laktasyon eğrisinin güvenilirliği önemli bir gerekliliktir. Ayrıca sürü ve işletmedeki verimi düşük hayvanlar saptanarak ayıklanabilir ve sürüde hayvan sağlığı ile ilgili erken bilgi sağlanabilir. Laktasyon eğrisinin oluşturulabilmesi için inekten elde edilen günlük süt miktarının, günlük olarak kayıt edilmesi gerekir. Bu işlem maliyetli olduğu gibi fazla emek gerektiren bir süreçtir, işletmelerin bu kayıtları her inek için tutmaları mümkün değildir. Ancak laktasyon eğrisi çeşitli yöntemlerle tahmin edilebilir. Genellikle ortalama 10 ay süren 305 günlük dönem içerisinde ayda bir yapılan kontrollerle laktasyon eğrisi elde edilir. Ay içerisinde seçilen bu günlere *kontrol günleri* adı verilmektedir. Kontrol günleri bilgilerinin tamamı ve kısmi bilgisi kullanılarak laktasyon eğrisi ile süt verimi arasındaki ilişkiyi açıklayabilmek için farklı modeller geliştirilmiştir (Karaman, 2010).

Hayatta sürekli olarak kullandığımız; sezgi, sağduyu gibi hissi yeteneklerimizden bilimsel amaçlar için yararlanılabilir. Bu tür hislerin bilimsel temeller çerçevesinde kullanımı Bayesian çıkarımını ortaya çıkarır. 1763 yılında R.Thomas Bayesian tarafından ortaya çıkartılmış olan bu teorem Bayesian Zaman Serisi analizinin temellerini oluşturmaktadır. (Eğrioğlu, 2002).

Zaman serisi analizlerinde Bayesian yaklaşımı üzerine yapılan önemli çalışmadan Zellner (1971) $AR(1)$ ve $AR(2)$ modelli parametrelerini Uniform önsel seçerek Bayesian analizini gerçekleştirmişlerdir. Diğer bir önemli çalışma Box-Jenkins (1970) çalışmasında Jeffreys önselini kullanarak tüm $ARMA$ modellerinin Bayesian analizini gerçekleştirmiştir. Sonrasında Monahan (1983) çalışmasında bilgi içermeyen önsel seçimini eleştirmiş ve modelde bulunan parametre sayısına bağlı bilgi içeren önsel seçimi ile $ARMA$ modellerinin Bayesian analizini gerçekleştirmiştir. Mevcut çalışmalar içerisinde önemli bir yeri olan çalışmalardan iki tanesi ise McCullach ve Tsoy (1994) ve Chip ve Greenberg (1994) çalışmalarıdır. Bu iki çalışmada nümerik integrasyon yöntemlerinden Gibbs Örnekleme yöntemi kullanılarak $ARMA$ modellerinin Bayesian analizi gerçekleştirilmiştir. Sonrasında çeşitli veri setleri üzerinde Bayesian yaklaşım ile zaman serisi model tahminleri çalışmaları gerçekleştirilmiştir.

Klasik Zaman Serisi yöntemi örnek verilerinden yararlanılarak bilgi sahibi olmak istediğimiz popülasyona ait sonuçlar ortaya çıkartılmasıdır. Ancak Bayesian yöntemler; nedenlerden sonuçlara ulaşmayı hedefleyen klasik yöntemlere kıyasla sonuçlardan nedenlere doğru bir mantık zinciri izlemektedir (Eğrioğlu, 2002).

Kontrol günleri bilgilerinin tamamı ve kısmi bilgisi kullanılarak laktasyon eğrisi ile süt verimi arasındaki ilişki modellenmesinde kullanılmış deterministik modeller Gamma modelleri; Wood Modeli, Ali-Schaeffer Modeli, Glasbey Modeli, Logaritmik Model, Wilmink Modeli, Goodall Modelidir. Sıklıkla kullanılan Gamma modellerindeki deterministik yapılara stokastik bir yapı kazandırılması ile geliştirilmiş doğrusal tek değişkenli zaman serisi modeli olan ARIMA modellerinden de yararlanılmıştır (Karaman, 2010; Orhan ve Kaygısız, 2002; Wood, 1967).

Genel olarak laktasyon eğrileri modellemesinde; tek değişkenli zaman serisi yöntemlerinden olan ARIMA modellerinin Gamma modellerinden daha az sapma ile süt verim tahmini yaptığı tespit edilmiştir. Yani laktasyon eğrilerinde kontrol günlerine ait süt verimlerinin modellenmesinde tek değişkenli zaman serileri yöntemlerinin, Gamma modellerinden daha uygun olduğu söylenebilir. Teke (2017) çalışmasında elde edilen sonuçların, Deluyker vd. (1990), Lark vd. (1999), Macciotta vd. (2000), Macciotta vd. (2002), Cappio-Borlino vd. (2004), Burki vd. (2005), Bek (2008); Karaman (2010); Berberoğlu (2010) yaptıkları çalışmaların sonuçları ile paralellik gösterdiği belirtilmiştir (Teke, 2017).

Daha önce kontrol günlerinin modellenmesinde klasik zaman serisi yöntemleri kullanılmış olup Bayesian zaman serisi yöntemleri kullanılmamıştır. Bu çalışmanın amacı farklı alanlarda kullanılmış olan ve diğer yöntemlere göre daha isabetli öngörülerde bulunabilmiş olan Bayesian zaman serisi yöntemi ile kontrol günlerinin modellenmesi ve uygun kontrol günü sayısının belirlenmesidir. Bu çalışmanın diğer bir amacı Frekansçı yaklaşım ile parametreleri tahmin edilen model ile Bayesian yaklaşımla parametreleri tahmin edilen iki modeli karşılaştırmaktır.

2. KAYNAK TARAMASI

2.1. Deterministik ve Stokastik Modeller

Modelleme çalışmalarının temel amacı incelenen değişkene dair en isabetli öngörülerini elde edebilmektir. Bu amaç doğrultusunda araştırmada incelenen değişken en iyi şekilde tanımlanmalıdır. Modelleme çalışmalarında kullanılan modeller iki sınıfta ele alınırlar: *deterministik ve stokastik modeller* (Karaman, 2010).

Deterministik modeller; değişkene ait yapılan öngörülerde hiçbir sapmanın olmadığı, kesin ilişkilerin açıklandığı modellerdir. Deterministik modeller genellikle fizik veya geometrik hesaplamalar için kullanılır. Bir karenin çevresi ($\Ç$) ile kenar uzunları (a) arasında ki ilişki,

$$\Ç = 4 \times a$$

şeklinde değişkenler arasındaki bu net ilişki *deterministik modeller* yardımı ile tam bir şekilde formüle edilebilir (Karaman, 2010).

Ancak doğada ve canlılarda gözlenen olaylara ait öngörülerin net ve kesin olarak hesaplanması neredeyse imkânsız kabul edilir. Biyolojik ve sosyal olaylar aynı koşullar altında tekrarlırsa, sebep ve sonuçlar matematiksel olarak tanımlansalar bile öngörülenden farklı sonuçlar gözlenebilmektedir. Böylece kesin olarak ilişkilerin modellenmesi mümkün olmamaktadır.

Stokastik veya olasılıksal modeller; incelenen değişken ile ilgili rastgeleliğin modele dâhil edildiği sistemlerdir. Bu rastgelelik faktörünün etkisi ile olay tekrarlırsa bile aynı değerler elde edilmeyecektir ancak sonuçların hangi aralıklarda ve olasılıklarda ortaya çıkabileceği hesaplanabilecektir.

Zaman serileri kapsamında yapılan modelleme çalışmalarında deterministik öğelere stokastik öğelerin eklenmesi ile hem kesin ilişkilerin modelde bulunduğu hem de değişkene ait rastgeleliğin yani stokastik etkilerin modele dâhil edildiği karma modeller ile çalışılmaktadır (Akgül, 2003). Örneğin,

$$Y_i = \beta_1 X_i + \varepsilon_i$$

olan modelde bulunan $\beta_1 X_i$ modelin deterministik kısmını ε_i stokastik kısmını temsil etmektedir (Şengül, 1986). Tek değişkenli zaman serisi modelleri hem deterministik bileşenlere hem de stokastik bileşenler içerebilmektedir.

2.2. Tek Değişkenli Zaman Serileri

Zaman serileri, sayısal olarak ifade edilebilen herhangi bir değişkenin nümerik değerlerinin belirli bir düzende ardı ardına sıralanması ile elde edilmiş dizilere verilen isimdir (Kaya 2019). Zaman serilerine örnek olarak bir çiftliğin aylık-haftalık-günlük kazançları ile elde edilen bir seri veya bir bölgenin yıllık-aylık aldığı yağış miktarı gibi seriler örnek gösterilebilirler (Karaman, 2010).

Ancak zamanın bir fonksiyonu olmamasına karşın belirli bir sistem çerçevesinde ardı ardına sıralanmış veriler de zaman serisi analiz yöntemleri ile analiz edilebilirler. Örneğin homojen veri setlerinde birden çok birime ait eşit zaman aralıklarında elde edilen ölçüm verileri ardı ardına sıralanarak tıpkı bir zaman serisi gibi bazı koşullar altında analiz edilebilirler. Bu neden ile *zaman serisi* olarak adlandırılabilirler (Karaman 2010; Macciotta vd. 2002).

Tahmin yöntemi fark etmeksizin zaman serilerinin analizinde kullanılan belirli kavramlar ve istatistiksel bir çerçeve söz konusudur. Bu kavramlar ve istatistiklerin iyi şekilde anlaşılması modelleme çalışmaları için oldukça önemlidir (Akgül, 2003). Bir zaman serisini diğer serilerden ayıran öncelik, seriye etki eden faktörlerin tespit edilmesidir. Böylece seri doğru şekilde sınıflandırılır. Zaman serilerinin analizinde tek değişken ile ilgilenildiği için otokolerasyon kavramı oldukça önemlidir. Analizlerin temelinde incelenen serinin durağanlık koşulu sınanmaktadır. Yapılan bu ayırt edici durumların tespiti ve koşulları sağlandıktan sonra uygun model seçimi yapılmaktadır. Bu genel çerçeve altında bu kavramlar ve istatistikler özetlendi.

2.2.1. Zaman Serilerine Etki Eden Faktörler

Serinin analiz edilebilmesinin temeli seriyi meydana getiren faktörlerin en iyi şekilde tespit edilmesine dayanır. Aynı zamanda bu faktörler serilerin karakteristik özelliklerini oluşturur. Bir seriyi diğer serilerden ayıran temel bileşenleri; *Trend*, *Mevsimsel Dalgalanmalar*, *Konjonktürel Dalgalanmalar* ve *Düzensiz Hareketler*dir (Kaya, 2019).

Zaman serileri analiz yöntemleri gereği seriyi oluşturan deterministik yapıdaki bileşenler önce ayrıştırılıp sonrasında öngörü yapılabilmesi için modelden ayrıştırılma yöntemine göre modele eklenebilmektedirler (Mert ve Çağlar 2019). Bu başlık kapsamında zaman serilerini meydana getiren temel bileşenlere ve deterministik unsurlara ait (toplamsal ve çarpımsal) ayrıştırma prosedürlerine yer verilmiştir.

2.2.1.1. Trend Bileşeni (T)

Trend bileşeni serinin ortalamasında meydana gelen sürekli artma veya azalma eğilimidir. Trend bileşeni meydana geliş tipine göre iki yapıda karşımıza çıkabilir. İlki stokastik yapıda (yığılımsal), ikincisi deterministik (dönemler arası bağımsız) yapıda trend yapısıdır. Stokastik yapıdaki trend bileşeni yığılımsal olarak artış veya azalış gösterirken, deterministik yapıdaki trend bileşeni serinin ortalamasında meydana gelen doğal değişimden kaynaklanmaktadır. Seriyeye ait trend yapısının en doğru şekilde belirlenmesi ve analizi oldukça önemlidir (Mert ve Çağlar 2019).

Stokastik yapıdaki trend bileşeni fark alınarak seri analizinden önce trend bileşeninden arındırılmalıdır. y_t zaman serisinin stokastik yapıda trendinin tespit edilmesi durumunda seriye,

$$\Delta y_t = y_t - y_{t-1}$$

işlemi uygulanarak fark alma işlemi gerçekleştirilir. Stokastik yapıdaki trend bileşeni ortadan kalkmış ise analize farkı alınmış durağan seri ile devam edilir ve fark

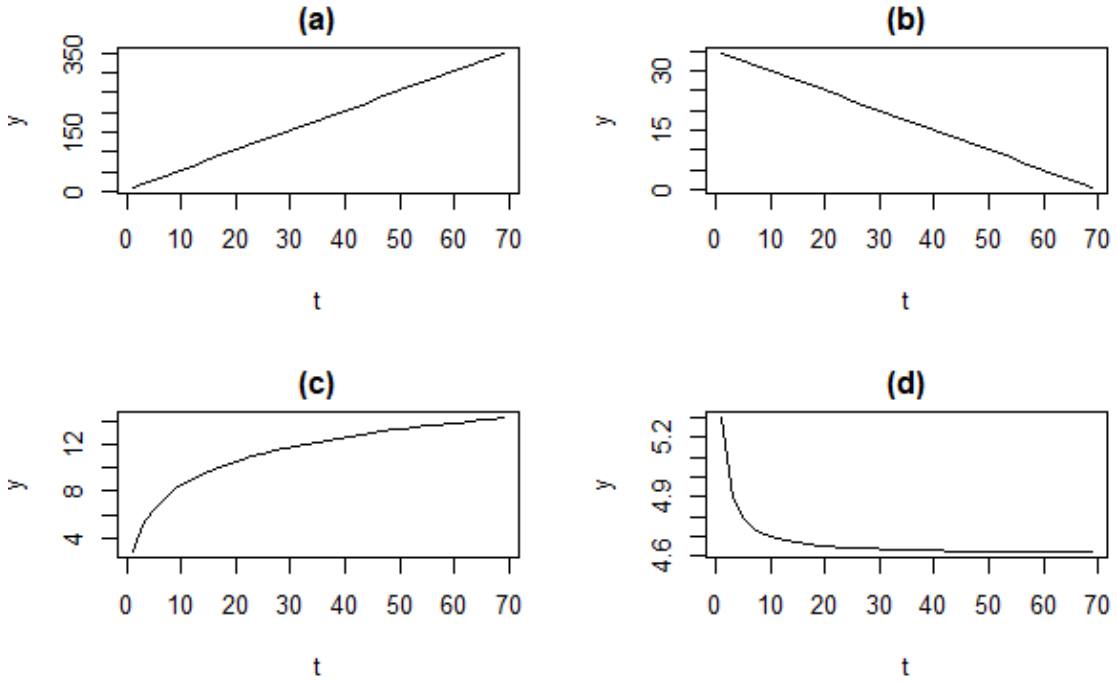
parametresi 1 olarak belirlenir. Ancak stokastik trend bileşeni hala var ise bu trend bileşenin doğrusal olmadığını gösterir ve ikinci fark alma işlemi uygulanarak,

$$\Delta^2 y_t = y_t - 2y_{t-1} - y_{t-2}$$

ikinci farkı alınarak durağanlaştırılmış seri ile analize devam edilir ve fark parametresi 2 olarak belirlenir (Akgül, 2000).

Seride bulunan trend yapısının deterministik özellikte olması durumunda ayrıştırma yöntemleri kullanılarak *saf trend bileşeni* elde edilir ve seriden ayrıştırma yöntemleri ile saf trend bileşeni ayrıştırılır. Ayrıştırma işlemi y_t serisindeki deterministik trend bileşenin toplamsal veya çarpımsal olma durumuna göre belirlenmektedir. (Mert and Çağlar, 2019).

Grafik yardımı ile görsel sınamada karşılaşılabilecek trend yapıları Şekil 2.1’de sunulmuştur. Şekil 2.1a’da ortalamada sürekli artış olması durumunda doğrusal artan trend yapısı, Şekil 2.1b’de sürekli azalış olması durumunda doğrusal azalan trend grafiği görseli yer almaktadır. Ayrıca trend bileşeni ortalamaya göre eğrisel değişim yapıları sergiliyor olabilir. Şekil 2.1c’de saf trend bileşenin eğrisel artan ve Şekil 2.1d’de eğrisel azalan yapısına ait görünüm yer almaktadır.



Şekil 2.1. Trend yapıları;**a)** Doğrusal artan trend yapısı;**b)** Doğrusal azalan trend yapısı;**c)** Doğrusal olmayan artan trend yapısı;**d)** Doğrusal olmayan azalan trend yapısı

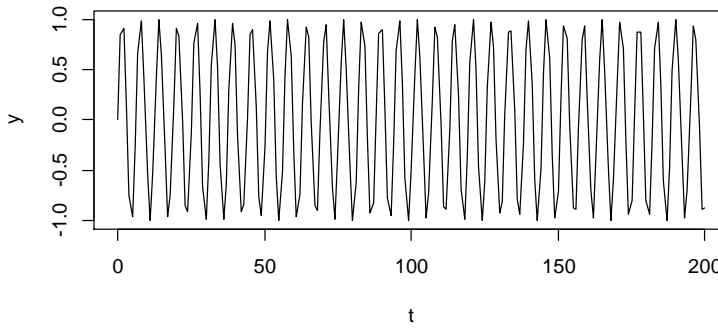
2.2.1.2. Mevsimsel dalgalanmalar (S)

Mevsim bileşeni serinin seyrinde sabit aralıklarla meydana gelen tekrarlı döngülerdir. Bir serinin değerleri trendin üstünde veya altında tekrarlı biçimde değerler almasıyla mevsimsel etkiler meydana gelir. Mevsimsel dalgalanmalar 3'er aylık

olabileceği gibi 5 günlük hatta 2’şer saatlik bile gerçekleşebilir. Bu dalgalanmaların süresi ne kadar olursa olsun *periyod* olarak ifade edilmektedir (Kadılar, 2005).

Mevsim bileşeni stokastik veya deterministik yapıda olabilir. Stokastik yapıdaki mevsimsel bileşen geçmiş dönemlere bağımlı bir şekilde yığılımlı özellik sergilemektedir. Ancak deterministik mevsim bileşeninde stokastik yapıya göre her periyotta benzer dalgalar bulunur ve yığılımlı etki nispeten bulunmamaktadır.

Deterministik mevsimsel bileşen seriden önce uygun yöntem ile ayrıştırılıp sonrasında yöntem gereğince modele eklenebilir. Seriden ayrıştırılmış mevsimselliğe *saf mevsim bileşeni* adı verilir (Mert and Çağlar, 2019) Şekil 2.2’de saf mevsim bileşeni zaman grafiği örneği bulunmaktadır.



Şekil 2.2. Saf mevsim bileşenin zaman yolu grafiği

Seride bulunan mevsimsel bileşenin stokastik yapıda olması durumunda seriye mevsimsel fark alma işlemi uygulanmalıdır. Örneğin her 10 gözlemlerde stokastik mevsimsellik barındıran bir zaman serisi y_t serisi için,

$$\Delta_{10}y_t = y_t - y_{t-10}$$

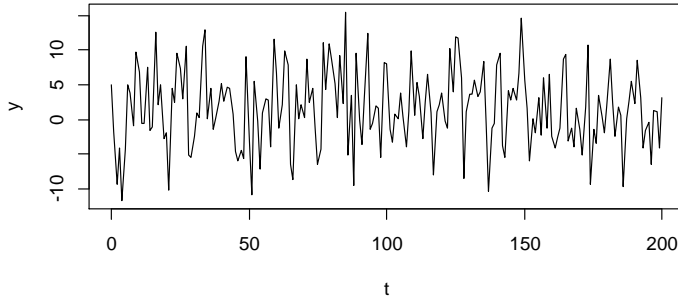
işlemi mevsimsel fark alma işlemi gerçekleştirilerek seri mevsimsellikten arındırılarak durağanlığı sağlanır ve mevsimsel fark mertebesi 1, mevsimsel fark periyodu 10 olarak belirlenir (Akgül, 2003).

2.2.1.3. Konjonktürel (döngüsel-çevrimsel) dalgalanmalar (C)

Değişkenin içinde bulunduğu sektörün veya ekonominin depresyon ve refah dönemlerindeki sistematik olmayan düzensiz durumlarda ki değişimleri kapsar. Bu tür salınımların devirleri mevsimsel periyotlardan çok daha uzun sürer (Berberoğlu, 2010).

2.2.1.4. Düzensiz hareketler (Hata Terimi)(E)

Önceden tahmin edilemeyen olayların etkilerini yansıtır. Bu faktör zaman serisi modelinin sistematik kısmını oluşturur. Diğer bir deyişle modelin deterministik faktörleri ile açıklanamayan (stokastik) rassal kısmını ifade eder (Akgül, 2003).



Şekil 2.3. Hata teriminin zaman yolu grafiği

Şekil 2.3’de sıfır ortalamalı, sabit varyanslı normal dağılıma sahip rassal hata bileşeni zaman yolu grafiği bulunmaktadır. Şekil 2.3’de görüldüğü gibi hata terimi, bir düzensiz bir seyir izlemelidir. Bir düzen içeriyor olması hata bileşenin doğru bir şekilde ayrıştırılmadığının göstergesi olacaktır (Kaya, 2019).

2.2.1.5. Ayrıştırma yöntemleri

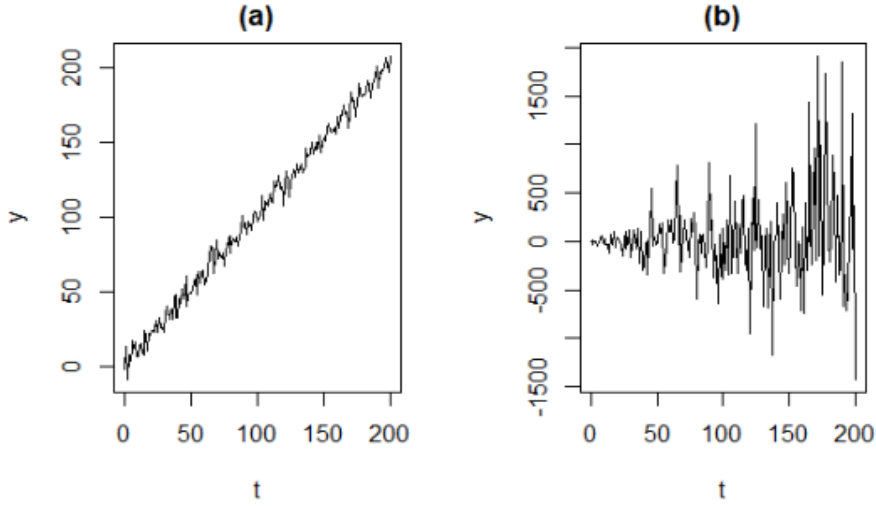
Ayrıştırma işlemleri ancak deterministik bileşenleri seriden ayırtmak için uygulanabilmektedir. Seride öncelikle deterministik bileşenlerin ayrıştırılması gerekmektedir. Bu neden ile seriye (Y) ait tüm bileşenler tespit edilerek zaman serisi modelinin yapısı incelenmelidir. Zaman serisine ait bileşenlerin iki olası yapısını inceleyelim. Bu yapılardan ilki toplamsal model yapısıdır ve gösterimi

$$Y = T + S + C + E$$

şeklinde ifade edilmektedir. Bir diğer olası model yapısı çarpımsal model ise

$$Y = T \times S \times C \times E$$

şeklinde ifade edilir. Görsel karşılaştırma için toplamsal ve çarpımsal bileşenli modellerin zaman yolu grafikleri Şekil 2.4’de sunulmuştur. Şekil 2.4a’da toplamsal özelliklerde bileşenlere sahip, Şekil 2.4b’de çarpımsal özelliklere sahip bileşenlere sahip zaman serisi zaman yolu grafikleri bulunmaktadır.



Şekil 2.4. Bileşenlerinin toplamından ve çarpımından oluşan seri örnekleri;**a)** Toplamsal model grafiği;**b)** Çarpımsal model grafiği

Genel olarak,

$$Y_t = f(T_t, S_t, C_t, E_t)$$

şeklinde gösterilebilen serinin bileşenlerine ayrılmış halidir. Burada T, S, C ile simgelenen unsurlar seriye ait deterministik bileşenlerin varlığını ortaya koyar bu 3 bileşene ait ayrıştırımlar, E bileşeni ile gösterilen stokastik bileşen pür rassal süreç (sabit ortalama, sabit varyans ve otokorelasyona sahip olmama) varsayımlarını sağlayana kadar devam ettirilir (Mert ve Çağlar 2019)

Seriye oluşturan bileşenlerin ayrıştırılması esnasında bileşenlerin toplamsal veya çarpımsal olduklarına karar vermek için hatalardan elde edilen bazı istatistiklerden yararlanır. Genellikle Hata Kareler Ortalamasının Karekökü (RMSE), Hataların Mutlak Ortalaması (MAE), Hataların Mutlak Yüzde Ortalaması (MAPE) değerlerinden faydalanılır. Ayrıştırma işleminin ardından modellere ait hata terimleri ile hesaplanan istatistiklerden en küçük olan ayrıştırma yöntemi tercih sebebidir (Mert ve Çağlar 2019). RMSE, MAE ve MAPE istatistikleri aşağıdaki gibi ifade edilirler:

$$RMSE = \sqrt{\frac{\sum_{t=1}^T E_t^2}{T}}, MAE = \frac{\sum_{t=1}^T |E_t|}{T}, MAPE = \frac{\sum_{t=1}^T \frac{|E_t|}{y_t}}{T}$$

2.2.2. Zaman Serilerinin Sınıflandırılması

Zaman serisi gözlem değerlerinin elde edilme biçimine göre kesikli ve sürekli seriler; seriyi oluşturan değerlerin ortalama değerinden farklarına göre durağan ve durağan olmayan seri; barındırdıkları devirli hareketlere göre ise mevsimsel ve mevsimsel olmayan zaman serisi olarak sınıflandırılır (Kaya 2019).

2.2.2.1. Kesikli ve Sürekli Zaman Serileri

Seriye oluşturan gözlemler sürekli bir şekilde elde edilebiliyorsa sürekli olarak adlandırılırlar. Yem fiyatları sürekli zaman serilerine bir örnek olarak verilebilir. Eğer seriyi oluşturan gözlemler sadece zamanın belli aralıklarında gözleniyorsa *kesikli zaman serisi* olarak adlandırılırlar. Tavuklara ait günlük yumurta miktarı kesikli zaman serisi örnekleridir (Karaman 2010).

Kesikli zaman serileri sürekli zaman serilerine dönüştürülemezler ancak sürekli zaman serilerine ait gözlem değerlerine belirli zaman aralıklarda temsil edecek değerler atanarak kesikli zaman serilerine dönüştürülebilmektedir (Mert ve Çağlar 2019). Bu çalışmadaki kullanılan süt miktarları sürekli bir değişken olduğu halde kontrol günlerindeki süt miktarları ile seri oluşturulduğu için kesikli zaman serilerine dönüştürülmüştür.

2.2.2.2. Tek Değişkenli Zaman Serilerinde Durağanlık

Bir zaman serisi için seriyi meydana getiren stokastik sürecin zamana göre değişmediği varsayımına *durağanlık* adı verilmektedir. Başka bir ifade ile seride herhangi bir yığılma bulunmaması sonucunda seride zamana bağlı değişen ortalama ve değişen varyansın bulunmaması durumudur (Akgül 2003).

Seride zamana bağlı sabit ortama ve sabit varyansın olmaması durumunu uygulamada çok sık karşılaşılan bir durumdur. Seriyeye ait varyansın durağanlaştırılması için seriyeye logaritmik dönüşüm yapılabilen, ortalamayı durağan hale getirmek için seriyeye fark alma işlemi uygulanmaktadır (Karaman 2010).

Fark alma işlemi seriyi oluşturan her gözlemden fark derecesi kadar önce gözlemin çıkartılması işlemidir. Durağan olmayan Y serisinin 1. Derece fark alma işlemi,

$$\Delta Y_t = Y_t - Y_{t-1}$$

şeklinde gerçekleştirilir. Aynı işlem geri kaydırma işlemcisi olan B yardımıyla,

$$\Delta Y_t = Y_t - Y_{t-1} = Y_t - BY_t = (1 - B)Y_t$$

şeklinde ifade edilir (Box-Jenkins, 1976).

2.2.2.3. Mevsimsel ve Mevsimsel Olmayan Zaman Serileri

Zamanın bir fonksiyonu olarak elde edilmiş zaman serileri bir yılı oluşturan mevsimlerin etkisinden veya veriyi oluşturan dönemin her periyodundan kaynaklı devri hareketler içeriyor olabilir. Örnek olarak Türkiye’de mera faktörü ile hayvancılık maliyetlerinin her bahar ve her yaz aylarında diğer aylara göre benzer şekilde azalma davranışı sergiliyor olması verilebilir. Bu dalgalanmalar *mevsimsel etki* olarak adlandırılmaktadır.

Ancak serinin oluşturulma metoduna bağlı periyodik dalgalanmalar içeriyor olabilir. Örneğin bu çalışmada zaman serisi zamanın bir fonksiyonu olarak değil de

tekrarlı ölçüm yöntemi ile elde edilmiş gözlemlerin ardı ardına sıralanması sonucunda meydana gelmiştir. Bu dalgalanmalar tıpkı mevsim etkisi gibi analiz edilir ancak farklı yorumlanırlar (Karaman 2010; Macciotta vd. 2000)

2.2.3. Zaman Serilerinde Durağanlık Sınamaları

Zaman serilerinde durağanlık analizi veya birim kök sınamalarında bahsedilen; zayıf ya da kovaryans durağanlığın sınamasıdır. Zayıf durağanlığın üç koşulu vardır. Bu koşullar sağlandığında serinin durağan olduğu söylenebilir. Bu koşullar sırayla,

$$E(y_t) = \mu,$$

$$Var(y_t) = \sigma^2,$$

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

şeklinde sıralanabilir. Durağanlık veya birim kök sınamaları temelde rassal yürüyüş sürecine dayanmaktadır. Bir seri için rassal yürüyüş süreci,

$$y_t = \varphi y_{t-1} + \varepsilon_t, \quad \varepsilon_t \sim iid(0, \sigma^2)$$

eşitliğindeki $\varphi = 1$ için sabit terimsiz ve trendsiz süreç olarak tanımlanmaktadır. Bu süreç t zaman için,

$$y_1 = \varepsilon_1$$

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

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

⋮

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

olarak elde edilecektir. Dolayısıyla $E(y_t) = E(\varepsilon_1 + \varepsilon_2 + \varepsilon_3 + \dots + \varepsilon_t) = 0$ elde edilir ve ilk koşul sağlanmış olur. Ancak ikinci koşul olan varyans $\delta = 1$ olduğu durumda $Var(y_t) = Var(\varepsilon_1 + \varepsilon_2 + \varepsilon_3 + \dots + \varepsilon_t) = t\sigma^2$ olarak hesaplanır. Durağanlık koşulunun sağlanabilmesi için varyansın zamandan bağımsız olması gerekir. Bu neden ile $|\delta| < 1$ olduğu durumda durağanlıktan bahsedilebilir. Sonuç olarak durağanlık testlerinde temel mantık $|\delta| < 1$ durumunun sınanmasıdır (Mert ve Çağlar 2019).

Bu başlık kapsamında birim kök (durağanlık) testlerinden bazılarına ait test denklemleri, hipotezleri, test istatistikleri formu ve karar aşaması hususlarına yer verilmiştir. Sıklıkla kullanılan birim kök testlerinden ADF (Arttırılmış Dickey-Fuller), PP (The Phillips-Perron) ve KPSS (Kwiatkowski, Phillips, Schmidt, Shin) testlerine ait bilgiler yer almaktadır.

2.2.3.1. Arttırılmış Dickey-Fuller Testi (ADF)

Dickey-Fuller (1979) test denklemi rassal yürüyüş sürecinin birinci farkının alınması ile elde edilir. Yani eşitliğin her iki tarafından y_{t-1} çıkartılarak elde edilir. Sabit terimsiz ve trendsiz model için,

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

$$\Delta y_t = \delta y_t + \varepsilon_t, \quad \varepsilon_t \sim iid(0, \sigma^2)$$

Sabit terimli (μ) ve trendsiz model için,

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

$$\Delta y_t = \mu + \delta y_t + \varepsilon_t, \quad \varepsilon_t \sim iid(0, \sigma^2)$$

Sabit terimli (μ) ve trendli (T) model için;

$$y_t - y_{t-1} = \mu + \beta T + \varphi y_{t-1} - y_{t-1} + \varepsilon_t$$

$$\Delta y_t = \mu + \beta T + \delta y_t + \varepsilon_t, \quad \varepsilon_t \sim iid(0, \sigma^2)$$

Böylece birim kökün varlığı durumunda $\delta = 0$ yani $\varphi = 1$ olurken, durağanlık durumunda $\delta = \varphi - 1 < 0$ olacaktır (Mert ve Çağlar 2019). Arttırılmış (genişletilmiş) Dickey-Fuller test denklemleri birden fazla gecikmeli fark denklemleri ile durağanlık sınaması yapılabilmesine olanak sağlamıştır. Böylece 1 den fazla merteye için birim kök sınaması testi yapılabilmektedir (Kaya, 2019).

Test denklemleri:

$$\text{(sabitli, trendsiz model)} \Delta y_t = \delta y_t + \sum_{i=1}^p \beta_i \Delta y_{t-i} + \varepsilon_t, \quad \varepsilon_t \sim iid(0, \sigma^2)$$

$$\text{(sabitli, trendsiz model)} \Delta y_t = \mu + \delta y_t + \sum_{i=1}^p \beta_i \Delta y_{t-i} + \varepsilon_t, \quad \varepsilon_t \sim iid(0, \sigma^2)$$

$$\text{(sabitli, trendli model)} \Delta y_t = \delta y_t \mu + \beta T + \sum_{i=1}^p \beta_i \Delta y_{t-i} + \varepsilon_t, \quad \varepsilon_t \sim iid(0, \sigma^2)$$

Test hipotezleri:

$$H_0: \delta = 0 \quad (\text{Birim kök vardır, seri durağan değildir})$$

$$H_1: \delta < 0 \quad (\text{Birim kök yoktur, seri durağandır})$$

Test istatistiği:

$$t_\delta = \frac{\hat{\delta}}{S_\delta} \sim \tau au$$

test istatistiği τau dağılır ve $\hat{\delta}$ değeri tahmin ediciyi, S_δ değeri tahmin edicinin tahmin edicisini simgeler.

Karar:

Seçilen test denkleminde göre, tercih edilecek güven düzeyine ve serbestlik derecesine göre τ dağılımından kritik değer hesaplanmaktadır. Eğer test istatistik değeri ilgili kritik değerden küçük ise H_0 hipotezi red edilecektir ve serinin zayıf durağan olduğuna karar verilecektir. Test denklemlerinde bulunan p gecikme uzunluğu hata terimi için gerekli varsayımların sağlandığı durumda tutumluluk (parsimony) prensibi gereği en küçük değer olarak ya da AIC ve BIC gibi bilgi kriterlerine göre belirlenir (Akgül 2000; Mert ve Çağlar 2019).

2.2.3.2. Phillips-Perron Testi (PP) (1988)

Dickey-Fuller test denklemlerinde bulunan $\sum_{i=1}^p \beta_i \Delta y_{t-i}$ teriminin varlığı ile hata terimleri arasındaki korelasyonun ortadan kaldırıldığı varsayılır. Phillips ve Perron hata terimine ait otokorelasyon problemi için alternatif olarak parametrik olmayan yöntemler önermişlerdir. Bu alternatif yaklaşıma ait hipotez, test denklemleri, test istatistiği ve kritik değerleri aşağıdaki gibidir (Mert ve Çağlar 2019).

Test denklemleri:

$$\text{(sabitli, trendsiz model)} \quad \Delta y_t = \delta y_t + \varepsilon_t, \quad \varepsilon_t \sim iid(0, \sigma^2)$$

$$\text{(sabitli, trendsiz model)} \quad \Delta y_t = \mu + \delta y_t + \varepsilon_t, \quad \varepsilon_t \sim iid(0, \sigma^2)$$

$$\text{(sabitli, trendli model)} \quad \Delta y_t = \delta y_t \mu + \beta T + \varepsilon_t, \quad \varepsilon_t \sim iid(0, \sigma^2)$$

Test hipotezleri:

$$H_0: \delta = 0 \quad (\text{Birim kök vardır, seri durağan değildir})$$

$$H_1: \delta < 0 \quad (\text{Birim kök yoktur, seri durağandır})$$

Test istatistiği:

$$\bar{t}_\delta = t_\delta \left(\frac{\gamma_0}{f_0} \right)^{1/2} - \frac{T(f_0 - \gamma_0) S_\delta}{2 f_0^{1/2} s} \sim \tau$$

Test istatistiğinde; t_δ değeri Dickey-Fuller test istatistiği, S_δ tahmin ediciye ait standart hata, s değeri test regresyonunun standart hatası, γ_0 değeri hata varyansının tahmin edicisi $[(T-k)s^2/T]$ ve f_0 değeri sıfır frekansta kalıntı spektrumu tahmincisidir.

Karar:

Eğer test istatistik değeri ilgili MacKinnon kritik değerinden küçük ise veya hesaplanan p istatistik değeri kritik değerden küçük ise H_0 hipotezi reddedilecektir ve serinin zayıf durağan olduğuna karar verilecektir.

2.2.3.3. Kwiatkowski, Phillips, Schmidt, Shin Testi (KPSS)(1992)

Diğer test hipotezlerinden farklı hipotezlere sahiptir. Diğer durağanlık sınamalarından farklı bir yaklaşım izleyerek benzer sonuçlara ulaşır. Diğer testlerde H_0 hipotezi reddedilerek seri durağan sonucuna varılırken, bu test de durum tam tersi şekilde gerçekleşerek, serinin durağanlığı sınanmaktadır.

Test denklemi:

$$y_t = x_t' \alpha + \varepsilon_t \xrightarrow{EKK} \hat{\varepsilon}_t, \quad \varepsilon_t \sim iid(0, \sigma^2)$$

Test denkleminde bulunan x_t terimi sabit terimi (μ) veya hem sabit hem trend terimini (T) içermektedir.

Test hipotezleri;

H_0 : Seri Durağandır.

H_1 : Seri Durağan Değildir.

Test istatistiği:

$$LM = \sum_t S(t)^2 / (T^2 f_0)$$

Test istatistiğinde; f_0 değeri sıfır frekansta kalıntı spektrumu tahmincisi, $S(t) = \sum_{r=1}^t \hat{\varepsilon}_t$ ifadesi birikimli kalıntı fonksiyonudur ve kalıntılar $\hat{\varepsilon}_t = y_t - x_t' \alpha(0)$ formülü ile elde edilir. Burada α orjinal fonksiyonun kalıntılara dayalı elde edilir.

Karar:

Kritik değerler KPSS (1992) tarafından belirlenmişlerdir. Hesaplanan test istatistiği belirlenen kritik değerden küçük ise H_0 hipotezi red edilemeyecektir ve serinin durağan olduğuna karar verilecektir (Mert ve Çağlar 2019).

2.2.4. Zaman Serilerinde Otokorelasyon

Otokorelasyon kelimesi oto ve korelasyon kelimelerinin birleşmesi ile oluşmuş bir terimdir. Kendi anlamına gelen 'oto' kelimesi ve ilişki anlamına gelen 'korelasyon' kelimesi anlamına uygun olarak serinin kendi geçmiş değerlerinin ilişkisi ile ilgilenir.

Tek değişkene ait modelleme çalışmalarında Otokorelasyon ve Kısmi Otokorelasyon fonksiyonları; serinin durağanlaştırılması işlemlerinde, mevsimsellik ilişkisi incelenirken, en iyi modelin seçimi sırasında, hata terimlerine ait varsayımlar incelenirken ve tahmin işlemi sırasında kullanılmaktadır (Akgül, 2003; Karaman, 2010).

Bir zaman serisi kendi gecikmeli dönemlerinden etkileniyor olması sıklıkla karşılaşılabilen bir durumdur. Bu ilişkinin tespit edilmesi ve modele uyarlanması seriye ait stokastik sürecin tanımlanabilmesine yardımcı olmaktadır. Seriyeye ait stokastik

sürecin kısmen açıklanabilmesi için; *otokovaryans fonksiyonu*, *otokorelasyon fonksiyonu* ve *kısmi otokorelasyon fonksiyonu* kullanılan araçlardır (Akgül, 2003).

2.2.4.1. Otokovaryans Fonksiyonu

Otokovaryans; seriden rasgele seçilmiş olan herhangi y_{x_1} değişkeni ile y_{x_2} değişkeni arasındaki ilişkinin ölçüsüdür. Kovaryans formülü,

$$Cov[y_{x_1}, y_{x_2}] = E \left[(y_{x_1} - E(y_{x_1})) (y_{x_2} - E(y_{x_2})) \right]$$

şeklinde ifade edilir. Otokovaryans fonksiyonu teorik olarak,

$$\gamma_t = Cov[y_t, y_{t+k}] = E[(y_t - E(y_t))(y_{t+k} - E(y_{t+k}))] = E[(y_t - \mu)(y_{t+k} - \mu)]$$

gösterilmektedir. Uygulamada \bar{y} seri değerlerinin ortalaması, T gözlem sayısı ve $K = 0, 1, 2, \dots, k$ için,

$$\hat{\gamma}_t = \frac{1}{T} \sum_{t=1}^{T-k} (y_t - \bar{y})(y_{t+k} - \bar{y})$$

formülü kullanılmaktadır (Akgül, 2003).

2.2.4.2. Otokorelasyon Fonksiyonu (ACF)

Otokovaryans değerlerindeki farklı ölçme birimlerinin etkisi ortadan kaldırmak için normalleştirme işlemi gerçekleştirilir. Elde edilen bu yeni fonksiyona *otokorelasyon fonksiyonu* adı verilir (Akgül, 2003).

Farklı zamanlardaki seri değerlerinden hesaplanan ilişki katsayılarına OKK denir. k =Seçilen gecikme sayısı olmak üzere OKK teorik olarak,

$$\rho_t = \frac{Cov[y_t, y_{t-k}]}{\sigma_{y_t} \sigma_{y_{t-k}}} = \frac{E[(y_t - \mu)(y_{t+k} - \mu)]}{\sqrt{E[(y_t - \mu)^2 E(y_{t+k} - \mu)^2]}}$$

formüle edilmiştir. Örnek otokorelasyon katsayıları;

$$\hat{\rho}_t = \frac{\sum_{t=k+1}^n (y_t - \bar{y})(y_{t-k} - \bar{y})}{\sum_{t=1}^n (y_t - \bar{y})^2}$$

formülü ile hesaplanmaktadır. Ancak hesaplanan örnek Otokorelasyon katsayılarının istatistiksel olarak anlamlılıkları sınanmalıdır.

Eşitlikte $k = 0, 1, 2, \dots, T$ için hesaplanan örnek otokorelasyon katsayılarının ortalaması sıfırdır ve standart sapması yaklaşık olarak $1/\sqrt{T}$ olan bir dağılıma sahip olması ortalaması 0 olan bir popülasyondan geldiğini gösterir. Bu varsayım Bartlett Testi, Box-Pierce testi gibi otokorelasyon testlerinin temelini oluşturmaktadır. Örnek otokorelasyon katsayıları model seçiminde ve serinin sabit varyansa, sabit ortalamaya

sahip olması için gerekli dönüşümün yapılması aşamalarında kullanılmaktadır (Akgül, 2003).

Seriye ait durağanlık analizi için (ÖOKF) örnek otokorelasyon katsayıları fonksiyonundan yararlanılır. k =Gecikme katsayıları arttıkça:

- Eğer örnek Otokorelasyon katsayıları fonksiyonu hızlı bir şekilde azalarak sıfır eksenini küçük farklarla kesiyorsa seriyi oluşturan sürecin durağan olmadığına karar verilir.

- Eğer örnek Otokorelasyon katsayıları fonksiyonu hızla azalmıyorsa ve sıfır eksenini uzun gecikmelerle kesiyorsa serinin durağan olmadığına karar verilir. (Akgül, 2003).

Zaman serisi modellerinde (100 gözlemden fazla değere sahip) kalıntı serilerinde otokorelasyon (serisel korelasyon) sınaması için Box-Pierce test istatistiğinden yararlanılır. Test için n : Gözlem sayısı, m : Test edilecek otokorelasyon katsayısı, $\hat{\varepsilon}_t$: kalıntı serisi q : MA modelinin derecesi, p : AR modelinin derecesidir.

Test hipotezleri;

$$H_0: \rho_1 = \rho_2 = \dots = \rho_k \text{ (Otokorelasyon yok)}$$

$$H_1: \rho_1 \neq \rho_2 \neq \dots \neq \rho_k \text{ (Otokorelasyon var)}$$

Test istatistiği;

$$Q = n \sum_{k=1}^m \rho_k^2 \sim \chi_{m-p-q}^2$$

Örnek otokorelasyon katsayıları;

$$\rho_k = \frac{\sum \hat{\varepsilon}_t \cdot \hat{\varepsilon}_{t-k}}{\sum \hat{\varepsilon}_t^2} \sim \chi_{m-p-q}^2$$

Karar;

Hesaplanan test istatistiği χ_{m-p-q}^2 değerinden büyük ise H_0 red edilir ve otokorelasyon olduğuna karar verilir. Hesaplanan test istatistiği χ_{m-p-q}^2 değerinden küçük ise H_0 red edilemez ve serisel korelasyon problemi olmadığına karar verilir.

2.2.4.3. Kısmi Otokorelasyon Fonksiyonu

Seriye ait herhangi iki gecikmenin diğer gecikmelerin etkileri sabit tutularak elde edilen ilişkilerine ait fonksiyona *kısmi otokorelasyon fonksiyonu* denir. Yani diğer bir ifade ile seriye ait diğer gecikmelerin etkisi sabit tutulduğunda iki gecikme arasındaki ilişki katsayısıdır (Kadılar, 2005). Serinin k . derece kısmi otokorelasyon katsayısı φ_{kk} ile gösterilir. Örnek kısmi otokorelasyon katsayısı,

$$\hat{\varphi}_{kk} = \frac{\hat{\rho}_k - \sum_{i=1}^{k-1} (\hat{\varphi}_{k-1,i})(\hat{\rho}_{k-i})}{1 - \sum_{i=1}^{k-1} (\hat{\varphi}_{k-1,i})(\hat{\rho}_i)}$$

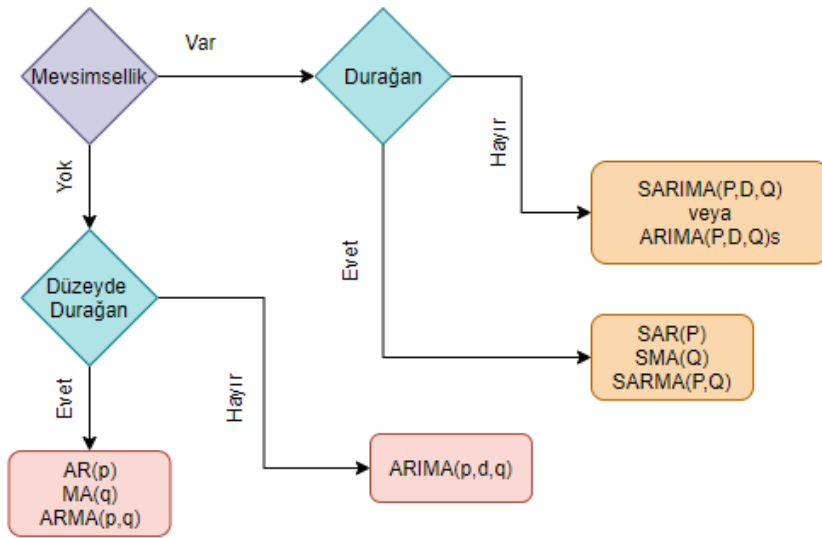
formülü ile hesaplanır. Eşitlikte $\hat{\rho}_k = k$. gecikme sonrası otokorelasyon katsayılarıdır. k . gecikmenin kısmi otokorelasyon katsayısı için i 'inci gecikmenin etkisi ortadan kaldırıldığında kısmi otokorelasyon katsayısı;

$$\hat{\varphi}_{ki} = \hat{\varphi}_{k-1,i} - (\hat{\varphi}_{kk})(\varphi_{k-1,k-i})$$

formülü ile hesaplanmaktadır. Örnek Otokorelasyon katsayıları fonksiyonu gibi örnek kısmi otokorelasyon katsayıları fonksiyonu da aynı varsayımlar ile serinin durağan olup olmadığına karar verilmesi aşamasında ve model seçiminde kullanılır (Kadılar, 2005).

2.2.5. Zaman Serisi Modelleri

Hem stokastik hem de deterministik unsur barındıran modellerden biri olan 'Tek değişkenli zaman serisi modelleri' farklı değişkenlerin tek modelde incelenmesi sonucunda sağlanması gereken ön koşulları gerektirmiyor olmasıyla ve kısa dönemli iyi öngörüler elde edilebiliyor olması bu modelin diğer öngörü amaçlı geliştirilen yöntemlerden üstünlüğüdür (Akgül, 2003).



Şekil 2.5. Model belirleme algoritması

Şekil 2.5'de ki akış şemasında görüldüğü gibi öncelikle seride mevsimsellik araştırması yapılır. Daha sonra durağanlık sınaması yapılarak uygun model yapısına karar verilebilir. Mevsimsellik içermeyen ve durağan veri yaratma süreçleri; Otokoregresif $AR(p)$, Hareketli ortalama $MA(q)$ ve Karma Otokoregresif hareketli ortalama $ARMA(p,q)$ süreçleridir. Veriyi yaratan sürecin durağan olmadığı ancak d ile temsil edilen fark alma işlemcisi ile durağanlaştırılarak tanımlanan; Bütünlenen Otokoregresif-Hareketli Ortalama $ARIMA(p, d, q)$ sürecidir (Akgül, 2003).

Mevsimsellik içeren ve durağan zaman serisi süreçleri; Mevsimsel Otoregresif $SAR(P)$ veya $AR(P)_s$, Mevsimsel Hareketli ortalama $SMA(Q)$ veya $MA(Q)_s$ ve Mevsimsel Karma Otoregresif hareketli ortalama $SARMA(P, Q)$ veya $ARMA(P, Q)_s$ süreçleridir. Veriyi yaratan sürecin mevsimsellikten kaynaklı durağan olmadığı ancak D ile temsil edilen mevsimsel fark alma işlemcisi ile durağanlaştırılarak tanımlanan; Mevsimsel Bütünlenen Otoregresif-Hareketli Ortalama $SARIMA(P, D, Q)$ veya $ARIMA(P, D, Q)_s$ sürecidir (Akgül, 2003).

2.2.5.1. Otoregresif Model AR(p)

Temelleri 1921 yıllarında Yule tarafından ortaya atılmış bir modeldir (Kaya, 2019). İncelenen değişkene ait zaman serisinin bir noktada aldığı değer y_t , $t = 1, 2, \dots, n$ için geçmiş p dönem değerleri ile tanımlanan bir fonksiyon ise veri yaratma süreci *Otoregresif süreç* olarak tanımlanmaktadır (Karaman, 2010). Otoregresif sürece ait model,

$$y_t = \varphi_1 y_{t-1} + \varphi_2 y_{t-2} + \dots + \varphi_p y_{t-p} + \varepsilon_t, \quad \varepsilon_t \sim ND(0, \sigma^2), \quad \sum_{i=1}^p \varphi_i < 1$$

ile ifade edilir. Geri kaydırma işlemcisi olan $B^p = y_{t-p}$ ile model gösterimi;

$$y_t = \varphi_1 B + \varphi_2 B^2 + \dots + \varphi_p B^p + \varepsilon_t, \quad \varepsilon_t \sim ND(0, \sigma^2), \quad \sum_{i=1}^p \varphi_i < 1$$

veya

$$y_t = \frac{1}{1 - \varphi_1 B - \varphi_2 B^2 - \dots - \varphi_p B^p} a_t = \varphi^{-1}(B) \varepsilon_t, \quad \varepsilon_t \sim ND(0, \sigma^2), \quad \sum_{i=1}^p \varphi_i < 1$$

ile yapılmaktadır. ε_t = modele ait hata terimidir ve kendinden önceki ($\varepsilon_{t-1}, \varepsilon_{t-2}, \dots$) hata terimlerinden bağımsız olması gibi varsayımları karşılıyor olması gerekmektedir. Aynı zamanda durağanlık varsayımı gereği model bulunan $AR(p)$ katsayıları (φ_i) toplamı 1'den küçük olmalıdır (Akgül, 2003; Karaman, 2010).

Model parametrelerini tahmin aşamasında sıklıkla olabilirlik fonksiyonundan yararlanılır. Otoregresif model için olabilirlik fonksiyonunun nasıl elde edildiğini inceleyelim. Böylece diğer modeller içinde olabilirlik fonksiyonunu elde edilme mantığı otoregresif model üzerinden açıklanmış olacaktır. Otoregresif modelin olabilirlik fonksiyonu hesabı için kareler toplamı fonksiyonu,

$$S(\varphi) = \sum_{t=1}^n \varepsilon_t^2$$

olarak yazılabilir. Bu toplamın hesaplanabilmesi için $\varepsilon_1, \varepsilon_2, \dots, \varepsilon_p$ değerleri hesaplanmalıdır. Ancak

$$\varepsilon_1 = y_1 - \varphi_1 y_0 - \varphi_2 y_{-1} - \cdots - \varphi_p y_{1-p},$$

$$\varepsilon_2 = y_2 - \varphi_1 y_1 - \varphi_2 y_0 - \cdots - \varphi_p y_{2-p},$$

$$\vdots$$

$$\varepsilon_p = y_p - \varphi_1 y_{p-1} - \varphi_2 y_{p-2} - \cdots - \varphi_p y_0$$

eşitliklerinin hesaplanabilmesi için ilk p tane artık sıfır kabul edilerek ve $\varepsilon_t \sim ND(0, \sigma^2)$ varsayımı altında ortak olasılık fonksiyonu;

$$f(\varepsilon_{p+1}, \varepsilon_{p+2}, \dots, \varepsilon_n) = (2\pi\sigma_\varepsilon^2)^{-(n-p)/2} \exp \left[\frac{-1}{2\sigma_\varepsilon^2} \left(\sum_{t=p+1}^n \sigma_\varepsilon^2 \right) \right]$$

elde edilir (Eğrioğlu, 2002). $\varepsilon_2 = y_t - \varphi_1 y_{t-1} - \cdots - \varphi_p y_{t-p}$ dönüşümü ile $\varepsilon_{p+1}, \varepsilon_{p+2}, \dots, \varepsilon_p$ değerlerini $y_{p+1}, y_{p+2}, \dots, y_n$ 'lere dönüştürürsek,

$$\varepsilon_{p+1} = y_{p+1} - \varphi_1 y_p - \varphi_2 y_{p-1} - \cdots - \varphi_p,$$

$$\varepsilon_{p+2} = y_{p+2} - \varphi_1 y_{p+1} - \varphi_2 y_p - \cdots - \varphi_p,$$

$$\vdots$$

$$\varepsilon_n = y_n - \varphi_1 y_{n-1} - \varphi_2 y_{n-2} - \cdots - \varphi_p y_{n-p}$$

eşitlikleri elde edilir. Burada y_1, y_2, \dots, y_p değerleri göz ardı edilmiştir. Böylece y_{p+1}, \dots, y_n 'ler için ortak olasılık fonksiyonu,

$$f(y_{p+1}, \dots, y_n) = (2\pi\sigma_\varepsilon^2)^{-(n-p)/2} \exp \left[\frac{-1}{2\sigma_\varepsilon^2} \left(\sum_{t=p+1}^n (y_t - \varphi_1 y_{t-1} - \cdots - \varphi_p y_{t-p})^2 \right) \right]$$

olarak yazılır. Göz ardı edilen değerlerinde $y_t \sim ND(0, \sigma^2)$ varsayımı altında ihmal edilen y_1, y_2, \dots, y_p değerlerinin ortak olasılık yoğunluk fonksiyonu ile çarparsak tam olabilirlik fonksiyonu elde edilir. Aşağıdaki eşitlikte $\sigma_\varepsilon^2 \Sigma_p^{-1} = y' = (y_1, y_2, \dots, y_p)$ varyans-kovaryans matrisi olduğu durum için ortak olasılık fonksiyonu,

$$f(y_1, \dots, y_p) = (2\pi\sigma_\varepsilon^2)^{-p/2} |\Sigma_p|^{1/2} \exp \left[\frac{-1}{2\sigma_\varepsilon^2} \left(\sum_{t=p+1}^n (y' \Sigma_p y)^2 \right) \right]$$

olarak hesaplanır. Şartsız kareler toplamı $\tilde{Q}(\varphi_1, \varphi_2, \dots, \varphi_p) = \sum_{i=1}^p \sum_{j=1}^p \Sigma_{ij} y_i y_j$ olduğu durum için y_1, y_2, \dots, y_p 'lerin ortak olasılık fonksiyonu ile y_{p+1}, \dots, y_n 'lerin ortak olasılık fonksiyonunu ile çarpılarak tam olabilirlik fonksiyonu,

$$f(y_1, \dots, y_n) = (2\pi\sigma_\varepsilon^2)^{-n/2} |\Sigma_p|^{1/2} \exp \left[\frac{-1}{2\sigma_\varepsilon^2} \tilde{Q}(\varphi_1, \varphi_2, \dots, \varphi_p) \right]$$

olarak elde edilir. Şartlı kareler toplamı $Q(\varphi_1, \varphi_2, \dots, \varphi_p)$ olduğu durumda tam logaritmik fonksiyonu;

$$\mathcal{L}(\varphi_1, \dots, \varphi_n) = -\frac{n}{2} \log 2\pi\sigma_\varepsilon^2 + \frac{1}{2} \log |\Sigma_p| - \frac{\tilde{Q}(\varphi_1, \varphi_2, \dots, \varphi_p)}{2\sigma_\varepsilon^2}$$

olarak elde edilen olabilirlik fonksiyonunda bulunan $\frac{1}{2} \log |\Sigma_p|$ ifadesinin türevini sıfıra eşitlediğimizde denklem çözülemez hale geldiği için ihmal edilebilir ve Frekansçı yaklaşım ile parametre tahmini gerçekleştirilebilir (Eğrioğlu, 2002).

Örnek olarak $AR(1)$ modeli için olabilirlik fonksiyonunu hesaplayalım. Kovaryans matrisi,

$$\sigma_a^2 |\Sigma_1^{-1}| = \frac{\sigma_a^2}{1-\varphi_1^2} \quad \Sigma_1^{-1} = \frac{1}{1-\varphi_1^2} \quad \Sigma_1 = (1-\varphi_1^2)$$

olur ve şartlı kareler toplamı; $Q(\varphi_1) = \sum_{t=2}^n (y_t - \varphi_1 y_{t-1})^2$ eşitliğine bağlı olarak şartsız kareler toplamı; $\tilde{Q}(\varphi_1) = (1-\varphi_1^2)y_1^2 + \sum_{t=2}^n (y_t - \varphi_1 y_{t-1})^2$ yardımı ile $AR(1)$ modelinin tam olabilirlik fonksiyonu,

$$\mathcal{L}(\varphi_1) = -\frac{n}{2} \log 2\pi\sigma_a^2 + \frac{1}{2} \log(1-\varphi_1^2) - \frac{\tilde{Q}(\varphi_1)}{2\sigma_a^2}$$

olarak elde edilir.

2.2.5.2. Hareketli Ortalama Modeli MA(q)

1927 yılında Shutsky tarafından ortaya çıkartılmış modeldir (Kaya, 2019). İncelenen değişkene ait zaman serisinin bir noktada aldığı değer y_t , $t = 1, 2, \dots, T$ adet mevcut ve q adet geçmiş dönem rassal şokun ağırlıklı toplamı ise süreç q . mertebeden hareketli ortalama sürecidir (Karaman, 2010). Hareketli ortalama modeli,

$$y_t = \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \dots - \theta_q \varepsilon_{t-q} \quad \varepsilon_t \sim ND(0, \sigma^2)$$

olarak ifade edilmektedir. θ_i ifadesi MA parametrelerini, $\varepsilon_{t-1}, \varepsilon_{t-2}, \dots, \varepsilon_{t-q}$ 'ler geçmiş dönem öngörü hatalarını simgelemektedir. Aşağıdaki eşitlikte $k = 1, 2, \dots$ geri kaydırma işlemcisi ($B^k = y_{t-k}$) ile model gösterimi,

$$y_t = (1 - \theta_1 B - \theta_1 B^2 - \dots - \theta_q \varepsilon_{t-q}) \varepsilon_t = \theta(B) \varepsilon_t, \quad \varepsilon_t \sim ND(0, \sigma^2)$$

şeklinindedir ve θ_i parametreleri ile ilgili istatistiksel olarak anlamlı olmaları dışında bir kısıt söz konusu değildir ancak durağanlık varsayımı gereği gecikme sayısı arttıkça θ_i değerlerinin küçülmesi gerekmektedir. Bu gereklilik gecikme sayısı olan k arttıkça korelasyon katsayılarının giderek küçülecek olması ve otokorelasyon fonksiyonun sıfıra yaklaşıyor olması ile tutarlıdır (Akgül, 2003; Karaman, 2010).

Hareketli ortalamalar modeli olabilirlik fonksiyonu için $a_t \sim ND(0, \sigma^2)$ varsayımı altında $\sigma_a^2 \Sigma_q^{-1} = (y_1, y_2, \dots, y_q)$ varyans-kovaryans matrisi olduğu durum için tam olabilirlik fonksiyonu,

$$\mathcal{L}(\theta \cdot \sigma_a / y_t) = (2\pi\sigma_a^2)^{-n/2} + |\Sigma_q|^{1/2} \exp \left\{ \frac{1}{2\sigma_a^2} \sum_{t=1-q}^n [a_t]^2 \right\}$$

olarak ifade edilir. $MA(1)$ modeli için olabilirlik fonksiyonu,

$$\mathcal{L}(\theta \cdot \sigma_a / y_t) = (2\pi\sigma_a^2)^{-T/2} + (1 - \theta^2)^{1/2} \exp \left\{ \frac{1}{2\sigma_a^2} \sum_{t=0}^T [a_t]^2 \right\}$$

ifadesine eşittir.

2.2.5.3. Karma Otoregresif Hareketli Ortalama Modeli ARMA(p, q)

1954 yıllarında Wold tarafından AR ve MA modellerinin birleştirilmesi ile oluşturulmuş bir modeldir (Kaya, 2019). İncelenen zaman serisi verisi yaratan süreç her zaman sadece AR veya MA olmayıp her iki veri yaratma süreci özelliklerini taşıyor olabilir. Böyle süreçlere Karma Otoregresif Hareketli Ortalama Süreci adı verilmektedir. $ARMA(p, q)$ modeli,

$$y_t = \varphi_1 y_{t-1} + \varphi_2 y_{t-2} + \dots + \varphi_p y_{t-p} + a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} - \dots - \theta_q a_{t-q},$$

$$a_t \sim ND(0, \sigma^2)$$

ifade edilmektedir. Geri kaydırma işlemcisi ($B^k = y_{t-k}$) ile gösterimi,

$$(1 - \varphi_1 B - \varphi_2 B^2 - \dots - \varphi_p B^p) y_t = (1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q) a_t,$$

$$a_t \sim ND(0, \sigma^2)$$

veya kapalı formda

$$\varphi(B) y_t = \theta(B) a_t,$$

$$a_t \sim ND(0, \sigma^2)$$

şeklinde yapılır. Modelin,

$$y_t = \varphi^{-1}(B) \theta(B) a_t$$

veya

$$y_t = \frac{\theta(B)}{\varphi(B)} a_t = \frac{1 - \varphi_1 B - \varphi_2 B^2 - \dots - \varphi_p B^p}{1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q} a_t, \quad a_t \sim ND(0, \sigma^2), \quad \sum_{i=1}^p \varphi_i < 1$$

olarak gösterimi mevcuttur (Akgül, 2003; Karaman, 2010). Karma Otoregresif Hareketli Ortalama Modeli olabilirlik fonksiyonu için $a_t \sim N(0, \sigma^2)$ varsayımı altında $\sigma_a^2 \Sigma_{p,q}^{-1}$ varyans-kovaryans matrisi olduğu durum için tam olabilirlik fonksiyonu,

$$\mathcal{L}(\varphi \cdot \theta \cdot \sigma_a / y) = (2\pi\sigma_a^2)^{-n/2} + |\Sigma_{p,q}|^{1/2} \exp \left\{ -\frac{\tilde{Q}(\varphi, \theta)}{2\sigma_a^2} \right\}$$

olarak hesaplanır. Eşitlikte yer alan geriye dönük kalıntı kareler toplamı,

$$\tilde{Q}(\varphi, \theta) = \sum_{t=1-p-q}^T a_t^2$$

eşitliğini simgelemektedir. $ARMA(1,1)$ modeli için olabilirlik fonksiyonu kovaryans matrisinin $\Sigma_{(1,1)} = \frac{1-\varphi^2}{\theta^2+2\theta\varphi+1}$ denkleminde eşit olduğu durum için,

$$\mathcal{L}(\varphi \cdot \theta \cdot \sigma_a / y) = (2\pi\sigma_a^2)^{-T/2} \left(\frac{1-\varphi^2}{\theta^2+2\theta\varphi+1} \right)^{1/2} \exp \left\{ \frac{1}{2\sigma_a^2} \sum_{t=-1}^T [a_t]^2 \right\}$$

olarak ifade edilir (Eğrioğlu, 2002).

2.2.5.4. Bütünlenen Otoregresif Hareketli Ortalama Modeli ARIMA(p, d, q)

Durağan olmayan ancak d. mertebe farkı alınarak durağanlaştırılmış ARMA modeline aynı zamanda Otoregresif Entegre Hareketli Ortalama Modelleri adı da verilmektedir. Model gösterimi,

$$\begin{aligned} \Delta^d y_t &= \varphi_1 \Delta^d y_t + \varphi_2 \Delta^d y_t + \dots + \varphi_p \Delta^d y_t + a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} - \dots \\ &\quad - \theta_q a_{t-q}, \quad a_t \sim ND(0, \sigma^2) \end{aligned}$$

İfade edilmektedir. Geri kaydırma işlemcisi ($B^k = y_{t-k}$) ile gösterimi,

$$\varphi(B)(1-B)^d y_t = \theta(B)a_t, \quad a_t \sim ND(0, \sigma^2)$$

şeklinde ifade edilir (Akgül, 2003; Karaman, 2010).

2.2.5.5. Mevsimsel Otoregresif Modeller SAR(P)

Mevsimsel modellerin analizleri mevsimsel olmayan modellerin istatistiksel analizleri ile aynı şekilde gerçekleştirilmektedir (Kadılar, 2005). Durağan zaman serisinin t dönemindeki gözlemlerinin, önceki yılına veya periyoduna karşılık gelen dönemi gözlemlerinin artı rassal şokun doğrusal bir fonksiyonu olarak ifade edildiği durumlarda sürecin mevsimsel otoregresif süreç olduğu ifade edilmektedir (Akgül, 2003; Karaman, 2010). Mevsimsel otoregresif model,

$$y_t = \Phi_1 y_{t-s} + \Phi_2 y_{t-2s} + \dots + \Phi_P y_{t-Ps} + a_t \quad a_t \sim ND(0, \sigma^2)$$

yada kaydırma işlemcisi ($B^k = y_{t-k}$) kullanarak,

$$(1 - \Phi_1 B^s + \Phi_2 B^{2s} + \dots + \Phi_P B^{Ps})y_t = a_t \quad a_t \sim ND(0, \sigma^2)$$

şeklinde ifade edilmektedir. Kapalı form gösterimi,

$$\Phi_1(B^s)y_t = a_t \quad a_t \sim ND(0, \sigma^2)$$

şeklinde ifade edilmektedir. Modellerde s =Mevsimsel dönem sayısını, P =Mevsimsel AR derecesini, Φ_i =Mevsimsel AR parametrelerini simgelemektedir. Burada dikkat edilmesi gereken nokta modelin kısmi otokorelasyon ve otokorelasyon fonksiyonu katsayıları ardışık gözlemleri ile değil belirli periyodlarına ilişkin gecikmelere ilişkisini ortaya koyacaktır (Akgül, 2003; Karaman, 2010).

2.2.5.6. Mevsimsel Hareketli Ortalama Modeli SMA(Q)

Düzeyle durağan zaman serisinin gözlemlerinde meydana gelen rassal şokun artı s -dönem önceki meydana gelmiş rassal şokun doğrusal bir fonksiyonu şeklinde ifade edildiği sürece mevsimsel hareketli ortalama süreci adı verilmektedir. Model $SMA(Q)$ veya $MA(Q)_s$ olarak ifade edilebilmektedir. (Akgül, 2003; Karaman, 2010). $MA(Q)_s$ modeli genel olarak,

$$y_t = a_t - \Theta_1 a_{t-s} - \Theta_2 a_{t-2s} - \dots - \Theta_Q a_{t-Qs} \quad a_t \sim ND(0, \sigma^2)$$

tanımlanmaktadır. Geri kaydırma işlemcisi ($B^k = y_{t-k}$) kullanılarak,

$$y_t = (1 - \Theta_1 B^s - \Theta_2 B^{2s} - \dots - \Theta_Q B^{Qs}) a_t \quad a_t \sim ND(0, \sigma^2)$$

veya kapalı formda

$$y_t = \Theta(B^s)a_t \quad a_t \sim ND(0, \sigma^2)$$

olarak yazılabilir. Modellerde SAR modelinde olduğu gibi s =Mevsimsel dönem sayısını, Q =Mevsimsel MA derecesini, Θ_i =Mevsimsel MA parametrelerini simgelemektedir. Burada dikkat edilmesi gereken nokta modelin kısmi otokorelasyon ve otokorelasyon fonksiyonu katsayıları ardışık gözlemleri ile değil belirli periyodlarına ilişkin gecikmelere ilişkisini ortaya koyacaktır (Akgül, 2003; Karaman, 2010).

2.2.5.7. Mevsimsel Karma Otoregresif Hareketli Ortalama Model SARMA(P, Q)_s

İncelenen mevsimsel zaman serisi verisi yaratan süreci hem AR hem MA olmayıp her iki veri yaratma süreci özelliklerini taşıyor olabilir. Böyle süreçlere Karma Otoregresif Hareketli Ortalama Süreci adı verilmekte ve $SARMA(P, Q)$ veya $ARMA(P, Q)_s$ olarak ifade edilebilmektedir (Akgül, 2003; Karaman, 2010). $ARMA(P, Q)_s$ modeli,

$$y_t = \Phi_1 y_{t-1} + \Phi_2 y_{t-2} + \dots + \Phi_P y_{t-P} + a_t - \Theta_1 a_{t-1} - \Theta_2 a_{t-2} - \dots - \Theta_Q a_{t-Q}$$

$$a_t \sim ND(0, \sigma^2)$$

ifade edilmektedir. Geri kaydırma işlemcisi ($B^k = y_{t-k}$) ile gösterimi,

$$(1 - \Phi_1 B^s - \Phi_2 B^{2s} - \dots - \Phi_P B^P) y_t = (1 - \Theta_1 B^s - \Theta_2 B^{2s} - \dots - \Theta_Q B^{Qs}) a_t$$

$$a_t \sim ND(0, \sigma^2)$$

veya kapalı formda,

$$\Phi_P(B^s) y_t = \Theta_Q(B^s) a_t$$

$$a_t \sim ND(0, \sigma^2)$$

şeklinde ifade edilir.

2.2.5.8. Mevsimsel Bütünlenen Otoregresif Hareketli Ortalama Modeli $ARIMA(P, D, Q)_s$

Durağan olmayan ancak D. mertebe mevsimsel farkı alınarak durağanlaştırılmış ARMA modeline mevsimsel bütünlenen otoregresif hareketli ortalama modelleri adı da verilmektedir. Model $ARIMA(P, D, Q)_s$ veya $SARIMA(P, D, Q)$ olarak gösterimleri mevcuttur. $ARIMA(P, D, Q)_s$ modeli geri kaydırma işlemcisi ($B^k = y_{t-k}$) ile;

$$\Phi_P(B^s) \Delta_s^D y_t = \Theta_Q(B^s) a_t, \quad a_t \sim ND(0, \sigma^2)$$

ifade edilmektedir (Karaman, 2010).

2.2.6. Zaman Serisi Modeli Seçimi

İncelenen değişkene ait zaman serisinin en iyi tanımlanması model başarısının temel gereksinimidir. En iyi tanımlanmasından kasıt değişkeni oluşturan deterministik ve stokastik unsurların, artan veya azalan eğimin, mevsimsel etki altında olup olmadığının en doğru şekilde tespit edilmesi ve sonrasında kendi geçmiş değerlerinden, cari ve geçmiş dönem rassal artıklarından ne derece etkileniyor olduğunun tespit edilmesidir (Akgül, 2003).

Tek değişkenli zaman serisi modeli inşası için geliştirilen yöntemler; Hareketli Ortalamalar Yöntemi, Trend Analizi Yöntemi, Uyarlayıcı Arındırma Tahmin Yöntemi, Üssel Düzeltme Yöntemi ve Box-Jenkins Yöntemleri olarak sıralayabilir (Kaya, 2019).

Bu yöntemlerden Box-Jenkins yöntemleri en son geliştirilmiş ve diğer yöntemlerin eksikleri üzerine çalışmış bir yöntemdir. Newbold ve Granger (1974) çalışmasında 50 veri seti üzerinde yaptıkları çalışma sonucunda diğer yöntemlerden daha başarılı bulunmuştur (Çağıl, 2017). Bu çalışmada da Box-Jenkins model seçme stratejisi adımları ile en uygun modele karar verilmiştir.

Box-Jenkins modelleme yönteminin bazı varsayımları vardır. Bunlar verinin kesikli, durağan ve stokastik bir yapıda olması gibi varsayımlardır.

- İncelenen değişken kesikli bir veri seti olmalıdır yani gözlemler eşit zaman aralıkları ile elde edilmiş olmalıdır.
- Seri durağan olmasa bile durağan hale getirilmiş olmalıdır.
- Stokastik bir yapıda olması gerekliliğinde ki felsefe verinin açıklanmasında kendi cari ve geçmiş değerlerinin kullanılmasına dayanmaktadır.

Bu varsayımlar nedeni ile Box-Jenkins modellerine *doğrusal stokastik modeller* adı da verilmektedir. Box-Jenkins karar verme stratejisi 5 basamak olarak özetlenmektedir. Bunlar sırası ile Belirlenme, Tahmin, Ayırt Edici Kontrol (Artık Analizi), Büyük Ayrım, Öngörüdür (Kaya, 2019).

2.2.6.1. Belirlenme

Belirlenme basamağını adımlar halinde gerçekleştirmek belirleme doğruluğu açısından önem arz etmektedir. Mert ve Çağlar (2019) çalışmasında bu 5 basamağın ilki olan Belirlenme basamağını üç adıma ayırmışlardır:

Adım 1: Seride deterministik yapıda mevsimsellik veya trend var ise önce bu bileşenden seri ayrıştırılmalı, daha sonra ayrıştırma işleminin toplamsal veya çarpımsal olmasına göre öngörü aşamasında mevsimsel bileşen modele eklenmeli veya çarpılmalıdır.

Adım 2: Birim kök (Durağanlık) testleri ile serinin durağanlığı araştırılmalıdır. Durağanlık sınamaları serinin birim kök sürecinde durağan olmadığı sonucunu verdiği durumda seriye fark alma işlemi uygulanarak farkı alınmış seri için birim kök testi yapılır. d değeri durağanlık mertebesi değeridir. Seriyeye kaç kez fark alındığını gösterir. Seriyeye uygulanan durağanlık sınamaları sonucunda seri düzey halinde durağan bulunmuş ise d değerinin sıfır olduğuna karar verilir.

Eğer seride stokastik mevsimsellikten kaynaklı durağan olmama problemi var ise mevsimsel fark alınır. Mevsimsel fark alma işlemi sayısı D değerini, periyod değeri s ile simgelenerek SARIMA modelleri ile çalışılır.

Adım 3: $ARIMA(p, d, q)$ modelindeki p ve q gecikme uzunlukları belirlenmesi adıımıdır. Bu değerlerin belirlenmesi adıımında otokorelasyon ve kısmi otokorelasyon grafiklerinden yararlanılmaktadır (Akgül, 2003). Otokorelasyon fonksiyonunda ki anlamlı ilişki katsayıları gecikme arttıkça yavaşça azalıyor ise q değerinin sıfırdan farklı olduğuna, aniden azalmış ise p değerinin sıfıra eşit olduğuna karar verilir. Aynı şekilde kısmi otokorelasyon grafiğindeki anlamlı ilişki katsayıları gecikme arttıkça yavaşça azalıyor ise q değeri sıfırdan farklıdır, aniden azalmış ise sıfıra eşittir. 1. ve 2. adımlar uygulanarak durağanlaştırılmış seriye ait p ve q değerinin belirlenme prosedürü Çizelge 2.1'de ki gibidir.

Otokorelasyon katsayıların anlamlılığının kontrolü için grafikler incelenirken katsayıların %95 güven sınırlarının grafik üzerinde yer alması katsayılara ilişkin istatistiksel anlamlılıklarının incelenmesi için gereklidir (Akgül, 2003).

Çizelge 2.1. ARIMA modellerinde gecikme uzunluklarının belirlenmesi (Sevüktekin ve Nargeleçekenler 2007; Akgül 2003;Karaman 2010).

Model	OKF	KOKF
$AR(p)$	Üstel olarak veya sinüs dalgaları şeklinde azalır	p gecikme sonrası keser
$MA(q)$	q gecikme sonrası keser	Üstel olarak veya sinüs dalgaları şeklinde azalır
$ARMA(p, q)$	Üstel olarak veya sinüs dalgaları şeklinde azalır, p gecikme sonrası keser	Üstel olarak veya sinüs dalgaları şeklinde azalır, q gecikme sonrası keser

Çizelge 2.1’de p ve q değerlerinin genel belirlenme prensibi ifade edilmiştir. Örnek gecikme uzunluklarının belirlenme prensipleri Çizelge 2.2’deki gibidir.

Çizelge 2.2. ARIMA modellerinde gecikme uzunluğu belirlenmesinde örnek durumlar (Eğrioğlu, 2002)

Model	OKF	KOKF
$AR(1)$	Üstel ve sinüsel azalma	Yalnızca 1. Gecikme farklı
$MA(1)$	Yalnızca 1. Gecikme farklı	Üstel ve sinüsel azalma
$ARMA(1, 1)$	Üstel azalma	Üstel azalma
$AR(2)$	Üstel ve sinüsel azalma	Yalnızca 1. Ve 2. Gecikme farklı
$MA(2)$	Yalnızca 1. Ve 2. Gecikme farklı	Üstel ve sinüsel azalma

Çizelge 2.2’de ki çerçeve neticesinde karar verilebildiği gibi olası durumların sınanması yöntemi de kullanılmaktadır. Mert ve Çağlar (2019) çalışmasında maksimum (genellikle maksimum 2) p ve q değeri ile başlanıp tüm kombinasyonların tahmin aşamasına taşınmasını önermişlerdir.

2.2.6.2. Tahmin

Deneme model parametrelerinin en iyi şekilde (sapmasız, tutarlı ve etkin) tahminlerin elde edildiği basamaktır. Akgül (2010) çalışmasında önerildiği gibi otokorelasyon grafikleri yardımı ile tek bir modele ait parametre tahmini gerçekleştirilebilir veya Mert ve Çağlar (2019) çalışmasında önerildiği gibi tüm olası gecikmeler için parametre tahmini gerçekleştirilebilir.

2.4. Zaman Serilerinin Parametre Tahmin Yöntemleri başlığında bu çalışmada kullanılan parametre tahmin yöntemi detaylı bir şekilde anlatılmıştır. Frekanscı ve Bayesian yaklaşımları çerçevesinde parametre tahmini gerçekleştirilebilir.

Bu basamakta dikkat edilmesi gereken husus tüm parametre değerlerinin istatistiksel olarak anlamlılıklarının incelenmesidir. Anlamlı olmayan parametreler modelden çıkartılır (Akgül, 2000).

2.2.6.3. Ayırt Edici Kontrol (Artık Analizi)

Parametre tahminlerinin ardından model artıkları elde edilir ve kalıntı serisi için bazı testler gerçekleştirilerek temiz dizi (White Noise) özellikte olup olmadığı sınanır. Kalıntı serisinin sıfır ortalamalı sabit varyansa sahip olması gerekir. Ayrıca kalıntılarda otokorelasyon kontrolü gerçekleştirilir.

Modelden elde edilen kalıntılarda herhangi bir serisel korelasyon belirtisinin olmaması varsayımı kalıntılara ait otokorelasyon ve kısmi otokorelasyon grafikleri incelenerek de gerçekleştirilebilir. Gecikme uzunluğu (k) artıkça kalıntılara ait otokorelasyon değerlerinin sıfıra yaklaşarak sönüp gitmesi beklenir.

Akgül (2010) çalışması için kalıntı serisi varsayımları sağlamıyorsa belirleme basamağına geri dönülerek başka olası gecikme uzunlukları ile basamaklar tekrar edilir. Mert ve Çağlar (2019) çalışması için gerekli varsayımları sağlamayan olası modeller elenerek sağlayan modeller ile Büyük Ayrım basamağına geçilir.

Bayesian yaklaşım ile parametre tahmini gerçekleştirildiği durumlarda hata terimi varsayımlarının sınanmasından önce sonsal dağılımının hesaplanmasında kullanılan simülasyon çalışmasının başarısı incelenmelidir. 2.3.1.3. Tanı Testleri başlığında simülasyona ait kontrollerde kullanılan tanı testleri sıralanmış ve açıklanmıştır.

2.2.6.4. Büyük Ayrım

Mert ve Çağlar (2019) çalışmasına göre Ayırt Edici Kontrol basamağından geçen modeller arasında en iyi modele karar verildiği aşamadır. En iyi modele karar verilmesi için birçok kriterlerden yararlanılmaktadır. Ancak başlıca iki bilgi kriteri olan Akaike bilgi kriteri ve Schwarz Bilgi Kriteri hakkında bilgi verilebilir.

Sıklıkla kullanılan ilk kriter Akaike bilgi kriteri (AIC) için; k modeldeki parametre değeri, n örnek genişliği ve AKT artık kareler toplamı olsun. Bu durumda,

$$AIC(k) = n \cdot \ln\left(\frac{AKT}{n}\right) + 2k$$

olarak hesaplanır (Karaman, 2010). Bir diğer kriter Schwarz (1978)'in geliştirdiği kriter dayanan Schwarz Bilgi Kriteri (BIC veya SBC veya $SBIC$)'dir. n gözlem sayısı olduğu durum için,

$$BIC(k) = n \cdot \ln\left(\frac{AKT}{n}\right) + k \cdot \ln(n)$$

işlemine eşittir. AIC ve BIC değerlerinden en küçüğüne sahip olan model için veri seti ile daha uyumlu olduğu söylenebilir ve o model seçilir.

2.2.6.5. Öngörü

Model ile öngörülerin yapıldığı aşamadır. 2.2.6.4. Büyük Ayrım başlığında seçilen model kullanılarak kısa dönemli tahmin işlemi gerçekleştirilir. Zamanın bir fonksiyonu olan zaman serilerinin kullanıldığı çalışmalarda gelecek dönemlerin tahmin edilmesi anlamlı olacaktır ancak bu çalışmada olduğu gibi zamanın bir fonksiyonu olmayan veri seti için gelecek tahmini yapmak faydalı olmayacaktır.

Ancak zamanın bir fonksiyonu olmayan veri setleri için Macciotta vd. (2002) ve Karaman (2010) çalışmalarında tercih edilen, Çizelge 3.2’de özetlenen öngörü algoritması ile öngörü değerleri elde edilerek kontrol grubuna ait değerler ile kıyaslanmaları faydalı olacaktır.

2.3. Bayesian Analizi

Bayesian analizi Thomas Bayes’in koşullu olasılık teorisine dayanmaktadır. Koşullu olasılık teorisini basit bir örnek ile anlatacak olursak; Kanseri teşhiste kullanılan bir metot kanserli hastaların %98’inde pozitif sonuç vermekte; kansersiz hastaların ise % 4’ünde pozitif sonuç vermekte olsun. Hastanenin kayıtlarına göre gelen kanserli hasta oranı %3 olduğuna göre hastaneye gelen bir bireye kanser testi uygulandığı durumda test sonucu pozitif çıkmış ise bireyin kanserli olma olasılığını araştıralım.

$$P(B) = \{\text{Şahsın kanserli olması}\} \rightarrow 0.03$$

$$P(A) = \{\text{Şahsın teste pozitif reaksiyon göstermesi}\}$$

$$P(A|B) = \{\text{Kanserli bireyin teste + reaksiyon göstermesi}\} \rightarrow 0.98$$

$$P(A|\bar{B}) = \{\text{Kansersiz bireyin teste + reaksiyon göstermesi}\} \rightarrow 0.04$$

$$P(B|A) = \{\text{Kanserli bireyin testinin pozitif çıkması}\}$$

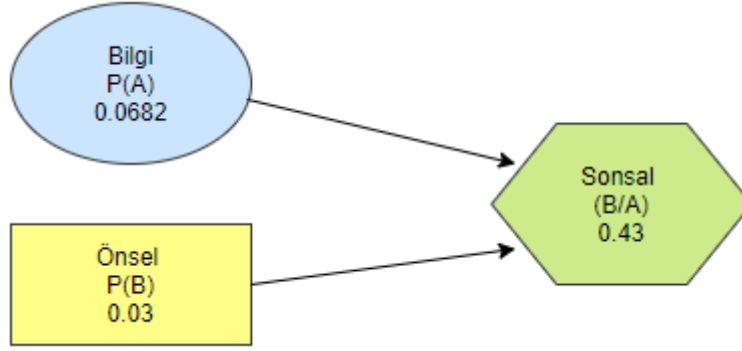
$$P(A) = P(A|B) \times P(B) + P(A|\bar{B}) \times P(\bar{B}) = 0.0682$$

$$P(B) + P(\bar{B}) = 1$$

$$P(\bar{B}) = 1 - 0.03 = 0.97$$

$$P(B|A) = \frac{P(A|B)P(B)}{P(A)} = \frac{0.98 \times 0.03}{0.0682} = 0.43$$

Bu durumda hastaneye gelen bireyin testinin pozitif çıkması durumunda gerçekten kanserli olma olasılığı %43’tür. Koşullu olasılık teorisi Şekil 2.6’da ki gibi görselleştirilebilir.



Şekil 2.6. Koşullu olasılık şeması

Bir olaya ait koşullu olasılığın tekrarlanması ile sonsal bilgi kendinden sonraki tekrar için önsel bilgiye dönüşecektir. Bu örnek için bireyin testi tekrar yaptırdığını düşünelim. Bu durumda önsel bilgi $P(B|A) = 0.43$ olacaktır.

$$P(B|A, C) = \{2.kez testinin pozitif çıkmış bireyin kanserli olması\}$$

$$P(B|A, C) = \frac{0.98 \cdot 0.43}{0.43 \cdot 0.98 + 0.57 \cdot 0.04} = 0.9487$$

Bireyin 2. testinin pozitif çıkması durumunda gerçekten hasta olma olasılığı %94.87 olarak hesaplanmıştır. Test her tekrarlandığında sonsal olasılık güncellenecek ve yeteri kadar tekrarlandıktan sonra %98 değerine yakınsayacaktır.

Bayesian teorisi çerçevesinde olasılıksal durumlardan dağılımların elde edilmesi Bayesian analizini meydana getirmiştir. Zaman serisi modellerinde Bayesian Yaklaşımdan faydalanılabilmektedir (Eğrioğlu, 2002).

2.3.1. Bayesian Yaklaşımı ile Zaman Serileri Analizi

Bir olaya ait örnek için anlatılan durumu biraz daha genelleyerek y_t zaman serisini oluşturan $t = 1, 2, \dots, T$ adet gözlemi ele alalım. Veri setindeki her gözlemin anakütlenin olabilirlik fonksiyonundan elde edilmiş birer olay olduğunu düşünürsek elimizdeki veri setinin olabilirlik fonksiyonunu tespit etmek demek popülasyonun da olabilirlik fonksiyonunu elde etmek demek olacaktır.

- Modele ait kalıntı serisi; $a_t \sim N(0, \sigma^2)$,
- Otoregresif fark değeri; p ,
- Kalıntı otoregresif fark değeri ; q ,
- Sürecin derecesine bağlı değişen bir fonksiyon; $f(\varphi, \theta)$,
- Kareler toplamı; $S(\varphi, \theta) = \sum_{t=1-(p+q)}^T [a_t / \varphi, \theta, y]^2$

olduğu durum için olabilirlik fonksiyonu,

$$\mathcal{L}(\varphi, \theta, \sigma_a) = f(\varphi, \theta) - \frac{T}{2} \ln(\sigma_a^2 2\pi) - \frac{S(\varphi, \theta)}{\sigma_a^2 2\pi}$$

şeklinde ifade edilir (Eğrioğlu, 2002). Bayesian yaklaşım olabilirlik fonksiyonun maksimum noktasını değil tamamını kullanarak araştırdığımız parametrelerin olabilirlik fonksiyonlarını elde eder. Ancak Bayesian yaklaşım popülasyonun tüm bilgisini kullanma prensibi gereği sadece örnek (dolayısıyla anakütlenin) olabilirlik fonksiyonunu kullanamaz ve popülasyon üzerindeki önsel olabilirlik fonksiyonuna da ihtiyaç duyar. Yani Bayesian yaklaşım olasılık teorisi gereği elindeki mevcut tüm olasılıkları çarparak popülasyonun sonsal olasılığını (mümkün en genel olabilirlik fonksiyonunu) elde etmeyi hedefler. Bu felsefeyi zaman serisi modeli değerleri ile ifade edecek olursak;

- Gözlem vektörü: y ,
 - Parametre vektörü: β , Parametre vektörü dağılımı: $P(\beta)$
 - Modelin olabilirlik fonksiyonu: $\mathcal{L}(\varphi, \theta, \sigma_a) = \mathcal{L}(\beta|y)$
 - Olasılık teorisi: $P(y, \beta) = P(y|\beta)P(\beta) = P(\beta|y)P(y)$
- eşitliklerinden yola çıkılarak sonsal dağılım,

$$P(\beta|y) = \frac{P(\beta)P(y|\beta)}{P(y)}$$

eşitliği ile ifade edilir. Gözlem vektörü bilindiği için $P(y)$ bir sabit olup *normalleştirme sabiti* olarak adlandırılır. $P(y) \neq 0$ olma şartı ile,

$$P(\beta|y) \propto P(\beta)P(y|\beta)$$

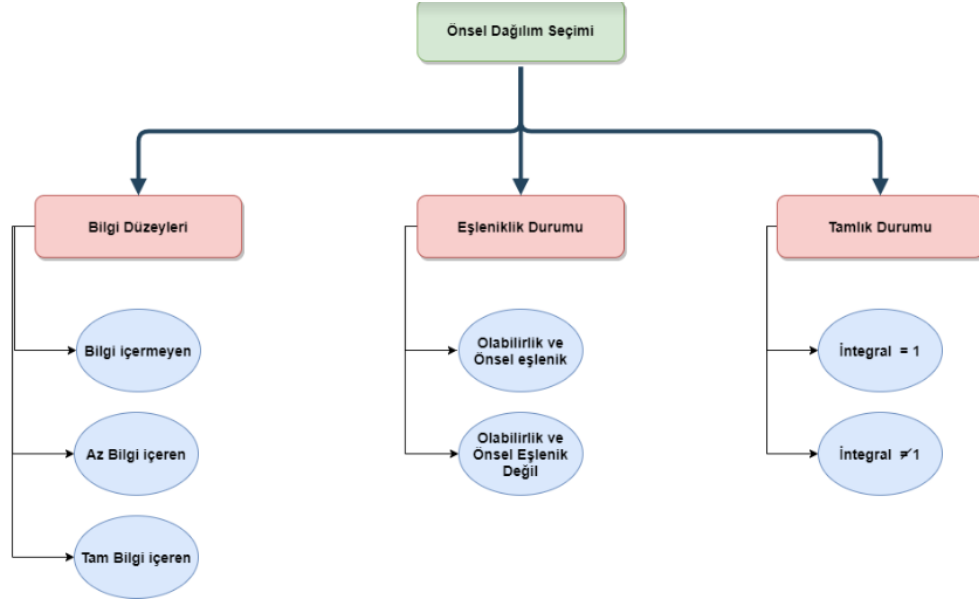
sonsalsal dağılım gösterilir veya $P(y|\beta)$ olabilirlik fonksiyonu gibi tavır sergilediği için yerine olabilirlik fonksiyonu $\mathcal{L}(\beta|y)$ kullanılabilir böylece

$$\text{Sonsal Dağılım} \propto \text{Önsel Dağılım} \times \text{Olabilirlik fonksiyonu}$$

olarak Bayesian analizinin 3 temel bileşeni gösterilebilir (Eğrioğlu, 2002). Frekansçı yöntemde de kullanılan olabilirlik fonksiyonundan bahsetmiştik. Bayesian yaklaşımı diğer yöntemlerden ayıran temel özellik önsel dağılımı analize dâhil etmesidir. Öncelikle önsel dağılım adımını, sonrasında sonsalsal dağılıma ait marjinal dağılımların tahmin edilmesi ve doğruluğunun kontrol edildiği tanı testlerinin uygulanma adımlarını inceleyelim.

2.3.1.1. Önsel dağılım

Analize yön veren önsel bilgi, çalışılan konu üzerinde çalışma yapan uzmanların bilgilerini ve çalışmalarının bir sonucudur. Önsel seçimindeki önsel kararın subjektif olabilen özelliği Bayesian Yaklaşımın eleştirilmesinin başlıca sebebidir (Koptur, 2020b). Ancak parametre dağılımındaki ön bilgimiz fazla ve veri seti yetersiz olabilir. Özellikle bu durumlarda oldukça faydalı bir yaklaşım olacaktır (Saçaklı Saçıldı, 2011). Şekil 2.7’de belirtilen önsel seçim kriterlerinden biri içerdikleri bilgi düzeylerine göre seçim yapılmasıdır.



Şekil 2.7. Önsel seçimi kriterleri

Önsel bilginin fazla olduğu durumlarda rahatlıkla spesifik aralıklarda bir önsel dağılım seçimi yapılabilir ve bu dağılıma *tam bilgi içeren önsel dağılım* adı verilir. Örneğin normal dağıldığı düşünülen tam bilgi içeren önsel için standart sapma değeri küçük $\beta \sim N(0, 0,0001)$ dağılımı örnek olarak gösterilebilir. İkinci grup az bilgi içeren önsel dağılımlardır; parametrenin aldığı değerlerden daha geniş aralıkta değerler almasına izin veren dağılımlardır. Yine normal dağılım örneği için $\beta \sim N(0, 10)$ dağılımı örnek gösterilebilir (Box ve Tiao 1973; Koptur 2020b). Parametreler hakkında çok daha az bilgiye sahip olunan durumlarda bilgi içermeyen önsel dağılımlar kullanılır. Bilgi içermeyen önsel dağılımlar üzerine Jeffreys (1962) önemli çalışmalar yapmıştır. *Jeffreys önseli* olarak bilinen Fisher bilgi matrisinin karekökü olarak alınabileceğini belirtmiştir. Başka bir bilgi içermeyen önsel ise Jeffreys önselinin karesi olan *Improper önsel dağılımıdır*. Bilgi içermeyen önsel dağılımlardan yaygın kullanılan önsellerden bir tanesi de *uniform önsel dağılımıdır*. Genel olarak sadece değişkenin pozitif-negatif olduğu veya belirli bir sayıdan büyük-küçük olduğu bilgisini taşıyan önsellerdir (Eğrioğlu, 2002). Herhangi bir model için olabilirlik fonksiyonu \mathcal{L} , β tahmin edilecek parametre $i \times j$ matrisi ve σ modelin standart hatası olmak üzere Fisher bilgi matrisi

$$I(\beta) = -E \left[\frac{\partial^2 \mathcal{L}}{\partial \beta_i \partial \beta_j} \right]$$

olarak ifade edilir ve bilgi içermeyen önsel dağılımlar şöyledir.

- Jeffreys Önseli: $P(\beta) \propto |I(\beta)|^{1/2}$
- Improper önsel: $P(\beta) \propto |I(\beta)|$
- Uniform önsel: \mathcal{L}/σ

ifadelerine eşit olacaktırlar.

Eşlenik olma durumuna göre önsel seçimi istatistiksel bilgisayar programları kullanılmadığı durumlar için gereklidir. Ana dağılım aileleri vardır ve diğer dağılımlar ana dağılımların özel durumlarıdır. Aynı ana dağılımın özel durumları olan iki dağılıma *eşlenik dağılımlar* denir. Önsel dağılım olabilirlik dağılımı ile aynı aileden seçilen önsel dağılımın matematiksel olarak çarpılması durumunda hesaplanan yaklaşık sonsal dağılım aynı aileye ait bir dağılıma sahip olacaktır. Bu bilgi yardımı ile olabilirlik fonksiyonu ile aynı aileden bir önsel dağılım bilgisi kullanılarak sonsal dağılımın tespit edilmesi işlem kolaylığı sağlamaktadır. Bu neden ile benzetim (simülasyon) çalışması yapılamadığı durumlar için eşlenik dağılım seçimi tercih sebebi olabilmektedir. Bu neden ile bu çalışmada önsel seçimi için bir kıstas olmamıştır.

Tamlık durumuna göre önseller ikiye ayrılır. Tanım aralığında integrali veya toplamı 1 olan dağılımlara *tam dağılım* denir. Tam olan önsel seçimi (sonsuz çözümü olan eşitsizlikler dışında) sonsal dağılımın tam olmasını sağlar. Tam olmayan önsel dağılımlar tanım aralığında integrali veya toplamı negatif sonsuz veya negatif sonsuz değerler alabilmektedir. Yine tamlık durumu eşleniklik durumu gibi sonsal dağılımın analitik hesaplanma probleminin üstesinden gelmek için dikkat edilen bir özelliktir. Bu neden ile tamlık durumu da bu çalışmada önsel seçimi için bir kıstas olmamıştır.

2.3.1.2. Sonsal dağılımın tahmini

Sonsal dağılım olabilirlik fonksiyonu ile önsel dağılımın çarpımından yaklaşık olarak elde edilir. Frekanscı yaklaşım ile yapılan En Çok Olabilirlik yönteminde kullanılan olabilirlik fonksiyonu Bayesian çıkarım için de kullanılır. Ancak Bayesian çıkarımda β değerleri birer rassal değişkendir. Gözlem vektörünün y , parametre vektörünün β , parametre vektörü dağılımının $P(\beta)$, modelin olabilirlik fonksiyonunun $\mathcal{L}(\beta|y)$ eşitlikleri ile ifade edildiği durum için sonsal dağılımı,

$$P(\beta|y) \propto P(\beta)\mathcal{L}(\beta|y)$$

olarak gösterilebilir Bu durumda β parametresini tahmin etmek için sonsal dağılımın β için beklenen değeri olarak,

$$\hat{\beta} = \int \beta \cdot P(\beta|y) \cdot d\beta$$

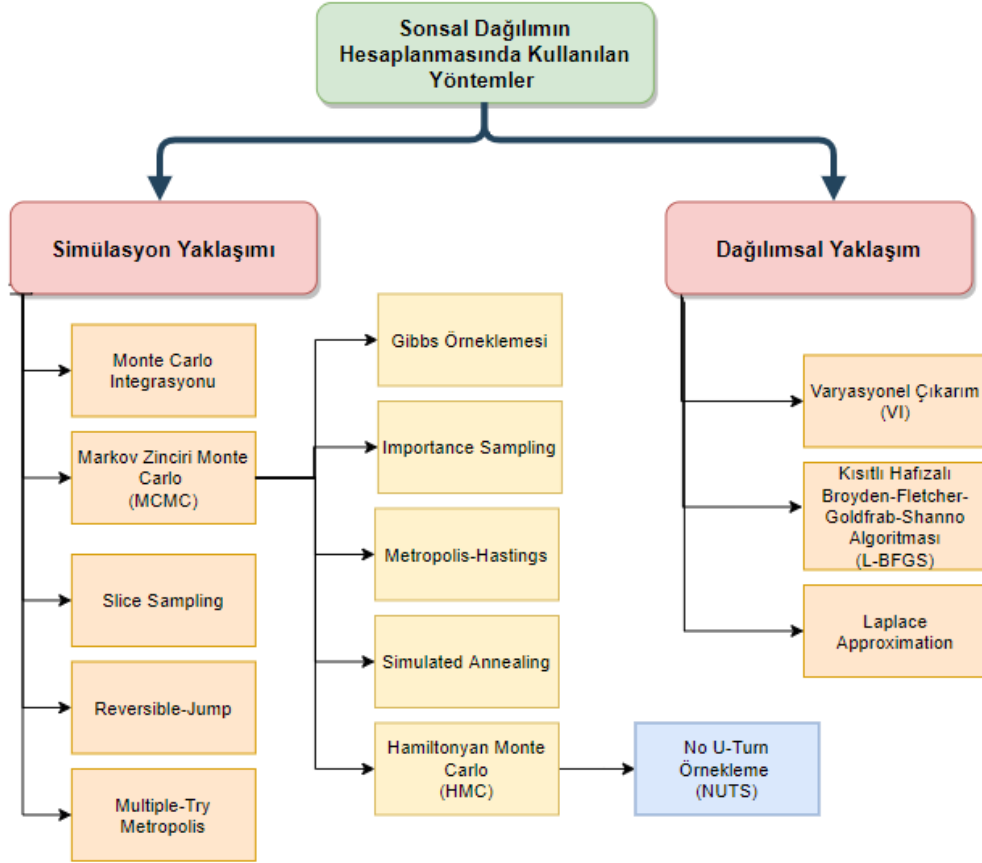
eşitliği hesaplanır. $\beta = (\beta_1)$ için Bayesian çıkarım yapılırken bu integral işlemi hesaplanabilir. Örneğin Eğrioğlu (2002) yaptığı çalışmada $ARIMA(1,1,0)_{12}$ modelinin parametre tahminleri Bayesian Yaklaşım ile hesaplayarak, matematiksel olarak açıklamıştır.

Ancak tahmin edilmek istenen birden fazla parametrenin $\beta = (\beta_1, \beta_2, \dots, \beta_p)$ olduğu durum için sonsal dağılım $P(\beta|y)$ çok değişkenli bir olasılık dağılımı olacaktır. Bu durumda β_1 parametresini tahmin etmek istediğimizde marjinal olasılık fonksiyonunu elde etmemiz gerekir ve bu

$$P(\beta_1|X) = \iint \dots \int P(\beta|y) d\beta_2 \dots d\beta_p$$

çözümünü yapmamız gerekecektir. Bu durumda ise bu integralin analitik olarak bir çözümü bulunmadığı durumlarda nümerik yöntemlere başvurmamız gerekmektedir (Koptur, 2020b).

Kullanılan istatistik paket programına bağlı olarak parametre tahmin yöntemi farklılık göstermektedir. Sonsal dağılımdan marjinal dağılımlar elde edilmesinde kullanılan iki yaklaşım bulunmaktadır: Simülasyon yaklaşımı ve dağılımsal yaklaşımdır. Kullanılan yöntemlerden başlıcaları Şekil 2.8’de yer almaktadır.



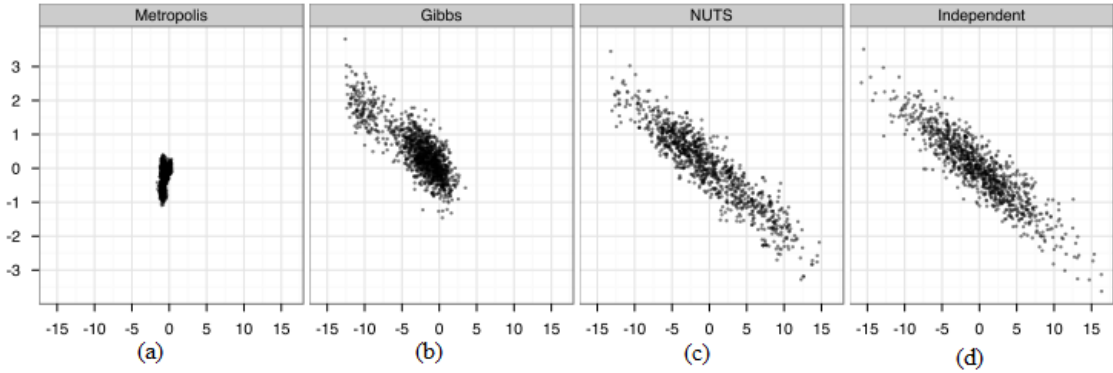
Şekil 2.8. Başlıca sonsal dağılım hesaplama yöntemleri

Şekil 2.8’de Markov Zinciri Monte Carlo (MCMC) yönteminin farklı varyasyonları bulunmaktadır. MCMC yöntemi temel alınarak geliştirilen başlıca yöntemler Önem Örnekleme (importance sampling), Gibbs örnekleme, Simulated Annealing yöntemleri sayılabilir. Ayrıca John Von Neumann ve Nicholas Metropolis tarafından MCMC yöntemi geliştirilerek Metropolis-Hastings yöntemi geliştirilmiştir. Yine MCMC yöntemi geliştirilerek HMC yöntemi ve düzeltilmesi NUTS yöntemleri geliştirilmiştir (Hoffman ve Gelman, 2014).

Zaman serisi model parametrelerini Bayesian Yaklaşım ile tahmin eden çalışmalarda genellikle fazla boyut ve birden çok merkezli sonsal dağılımdan örneklem çekilmesi yapılmamıştır. Bu çalışmalarda sonsal dağılımın hesaplanması aşamasında genellikle Metropolis-Hastings ve Gibbs örnekleme gibi yöntemleri ile sonsal dağılım hesaplanmıştır.

MCMC yöntemleri varyasyonlarında önsel dağılımı atanmış olasılıksal durum için Markov Zinciri bir başlangıç değeri ile başlatılır ve bir zincir oluşması için çalıştırılır. Zincir bir kez durağan duruma geldiğinde artık durağan durumda kalır. Başlangıçta durağan hale gelinceye dek oluşan kısım (warm-up periyod) ihmal edilir (Koptur, 2020b).

Hoffman ve Gelman (2014) çalışmalarında 250 boyutlu bir sonsal dağılım için MCMC yöntemlerinden; Metropolis-Hastings Yöntemini, Gibbs Örnekleme yöntemi ve No U-turn Örnekleme Yöntemini karşılaştırmışlardır ve yöntemlerin örneklem uzayları grafiksel olarak Şekil 2.9'da ki gibi elde edilmiştir.



Şekil 2.9. Üç farklı örnekleme yöntemine ait örneklem uzayları grafikleri; **a)** Metropolis-Hastings Yöntemi; **b)** Gibbs Örnekleme; **c)** NUTS yöntemi **d)** Gerçek örnekleme uzayı (Hoffman and Gelman, 2014)

Şekil 2.9'da Metropolis-Hastings Yöntemi aralarında yüksek korelasyon bulunan çok boyutlu hesaplamalarda diğer iki yöntemle göre daha zayıf kalmıştır. En iyi performansı güncel yöntemlerden biri olan ve bu çalışmada kullanılan (NUTS) No U-Turn Örnekleme yöntemi göstermiştir. NUTS yöntemi HMC yönteminde karşılaşılan örnekleme uzayını dolaşma problemine çözüm olarak sunulmuş bir düzeltme yöntemidir.

Hamiltonyan Monte Carlo yöntemi ilgilendiğimiz sonsal dağılımdan parametre örnekleme çekebilme için olasılık yoğunluk fonksiyona bir fiziksel sistemin potansiyel fonksiyonu olarak yaklaşmaktadır. HMC yöntemi dinamikleri; olasılık yoğunluk fonksiyonunu hiper sürtünmesiz bir ortam kabul eder ve bu ortamda bir topun var olduğunu varsayar. Toplam enerjiyi tanımlamak için pozisyon ve momentumdan yararlanır. Toplam enerjiyi bu iki bileşim ile açıklamak için

$$H(q, p) = U(q) + K(p)$$

eşitliğinden yararlanır. Hamiltonyan denkleminde $U(q) = -\log(\text{Posterior})$ potansiyel enerjiyi tanımlar. $K(p) = -\log(\text{olabilirlik})$ kinetik enerjiyi tanımlar (Robert, 2013). Hamiltonyan denkleminin hesaplanabilmesi için bazı nümerik hesaplamalar yapmak gerekmektedir. $j = 1, 2, \dots, L$ için

$$1. \quad p_{t+\varepsilon/2} = p_t - (\varepsilon/2) \frac{\partial U}{\partial q}(qt)$$

$$2. \quad q_{t+\varepsilon} = q_t + \varepsilon \frac{\partial K}{\partial p} (p_{t+\varepsilon/2})$$

$$3. \quad q_{t+\varepsilon} = p_{t+\varepsilon/2} - (\varepsilon/2) \frac{\partial U}{\partial q} (q_{t+\varepsilon}), \quad t = t + \varepsilon$$

HMC yönteminin performansını denklemden bulunan $\varepsilon =$ adım büyüklüğü ve $L =$ adım sayısı değerleri belirler. Eğer,

ε değeri çok büyük ise \rightarrow Yanlış simülasyon ve çok reddetme gerçekleşecektir,

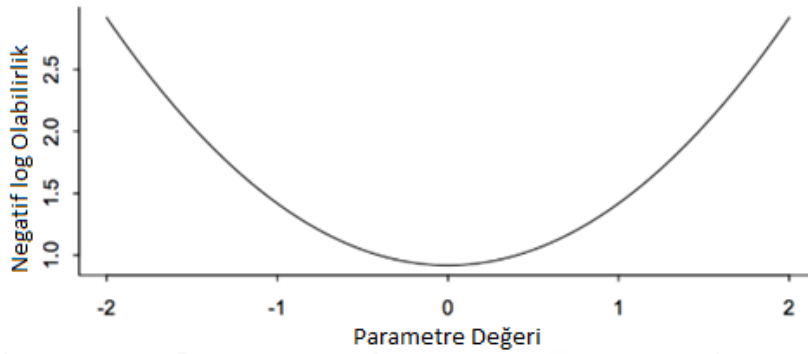
ε değeri çok küçük ise \rightarrow Örneklem uzayını az dolaşacaktır,

L değeri çok küçük ise \rightarrow Rassal davranıyor ve yavaş yakınsama gerçekleşecektir,

L değeri çok büyük ise \rightarrow Yörüngeler adımlarını takip ediyor (Yönteminde olduğu gibi)

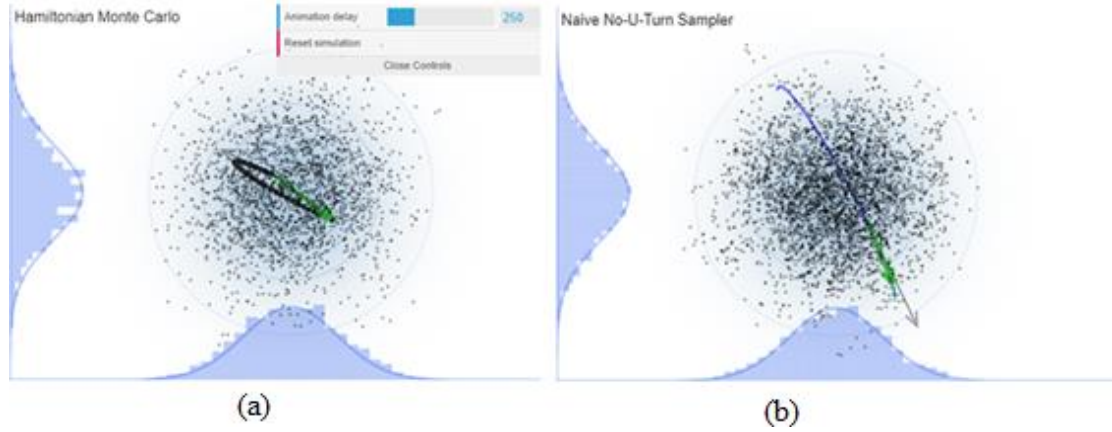
ε ve L değerleri ile ilgili problemlerin yanı sıra periyodik olarak tekrarlanma problemi gerçekleşmektedir (Robert, 2013).

Şekilsel olarak iki boyutlu bir olasılık yoğunluk fonksiyonumuz (fiziksel sistem) olsun ve olası durumlarımızın bu fiziksel sistemde sürtünmesiz olarak bir topun hareketlerine benzer şekilde sistemde hareket ettiğini varsayalım.



Şekil 2.10. iki boyutlu bir olasılık yoğunluk fonksiyonu (McElreath, 2017)

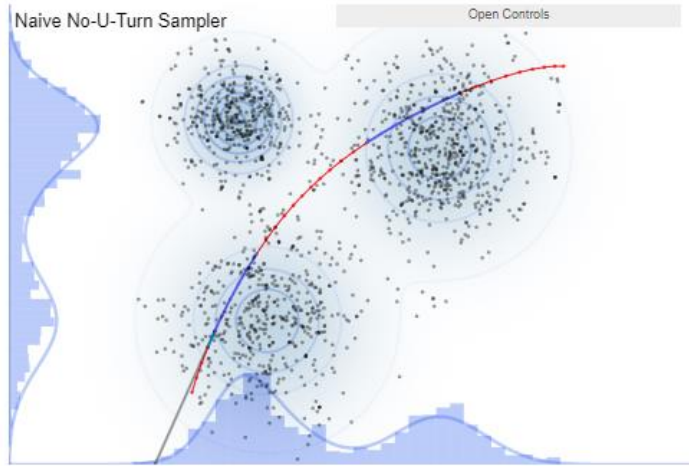
Şekil 2.10'da ki alanda hayali top farklı noktalardan bırakılmış olduğu durumda orta (dip) alanında topun dolaşma sıklığı fazla olacaktır (Robert, 2013). Böylece dağılımın merkezi diğer bölgelere daha fazla dolaşılacak ve yanlış örneklem elde edilmiş olacaktır. Bu durumun uygulamada gösterimi için Şekil 2.11'de ki 250 boyutlu sonsal dağılım için HMC yöntemi ve NUTS yöntemi simülasyonları karşılaştırılmıştır.



Şekil 2.11. Örneklem uzayı gösterimleri;**a)** HMC yöntemi örneklemeleri;**b)** NUTS düzeltmesinden sonra örneklem uzayı (McElreath, 2017)

Şekil 2.11’de iki yöntemde sonsal dağılıma ait örneklem uzayını dolaşma simülasyonlarının başarılı olduğu görülmektedir. Ancak HMC yöntemi simülasyonun merkez dışı alanları NUTS yöntemi simülasyonuna göre daha az dolaştığı ve NUTS yöntemine göre daha az başarılı olduğu görülmektedir. HMC yönteminin bu gibi eksikliklerini düzeltmek için NUTS yöntemi geliştirilmiştir. Yolun ne zaman U dönüşü yapacağını bularak U dönüşü yapmasına mani olur. Böylece yolun her iki tarafını da dolaşmasına izin vererek dağılımda daha çok alanının dolaşılmasına imkan sağlar (Robert, 2013).

NUTS yöntemi tek merkezli durumlarda kullanışlı bir örnekleyicidir. Ancak daha fazla merkezli sonsal dağılımlarda merkezler arasında geçişinin yavaş olması bu yöntemin eksikliği olarak sayılabilir. Şekil 2.12 bu duruma örnek olarak sunulabilir.



Şekil 2.12. Birden fazla merkezli sonsal dağılım için NUTS yönteminin örneklem uzayı (McElreath, 2017)

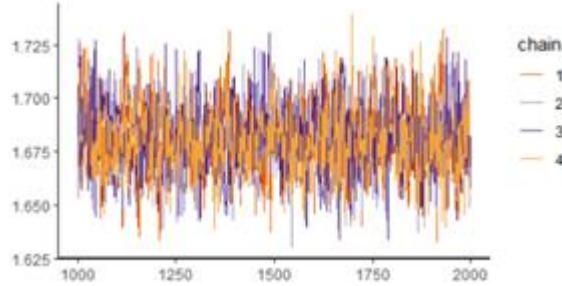
Şekil 2.12’de 3 tepeli bir sonsal dağılımından NUTS yöntemi ile örneklem elde edilmesi simule edilmiştir. Ancak alttaki merkezin diğer merkezlere göre daha az

dolaşıldığı görülmektedir (McElreath, 2017). Bu çalışmada tek merkezli sonsal dağılıma ait örneklem çekilmiştir ve böylece bu problemten kaçınılmıştır.

2.3.1.3. Tanı testleri

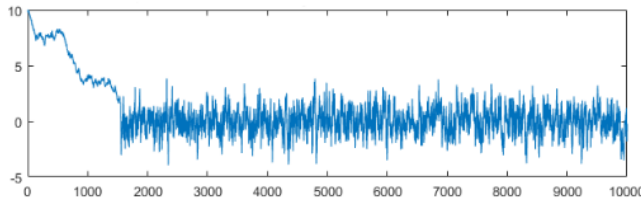
Sonsal dağılımdan örneklem çekilmesi için kullanılan yöntemlerden bahsedilmiştir. Tanı testleri kullanılan yöntemlerin kontrol edildiği aşamadır. Simülasyon yaklaşımı yöntemlerinden MCMC yöntemi algoritmasından türetilen Markov Zincirinin sonsal dağılıma yakınsayıp yakınsanmadığının kontrolü oldukça önemlidir (Saçaklı Saçıldı, 2011). Bu nedenle öncelikle Markov Zinciri (veya Zincirleri) kontrolünü bu çalışmanın kapsamında ilk 4 kriterde inceleyeceğiz. Sonrasında örneklem uzayının yeteri kadar dolaşılıp dolaşılmadığını V. ve VI. Tanı testleri yardımı ile inceleyeceğiz. V. ve VI. kriterler HMC gibi yöntemlerin kullanıldığı durumda gerekli kriterlerdir (Koptur, 2020a).

- I. Traceplots:** MCMC yöntemlerinin varyasyonları için bir birinden bağımsız Markov zincirleri oluşturulur ve Trace grafikleri yardımı ile simülasyon çalışmasının geçerliliği sınanır. Seçilen algoritma birden fazla zincir için çalıştırılır ve her zincirin birbirine benzer sonuçları vermesi beklenir. Zincirlerin grafikleri ayrı ayrı veya üst üste Şekil 2.13’de ki gibi çizdirilebilir. Her zincirin birine benzer salınımda ve rassal bir süreç sergiliyor olması beklenir. Böylece Şekil 2.13’de ki gibi serisel korelasyon içermediklerine ve rassal yürüyüş sergilediklerine karar verilebilir (Koptur, 2020b).



Şekil 2.13. Bağımsız 4 Markov Zincirine ait tek Trace grafiği

- II. Iraksama:** Markov Zincirlerinin yakınsaması incelenir. Warm-up periyodunun kısa olması ve zincirin hızlı bir şekilde durağanlığa ulaşması istenen bir durumdur. Uzun Warm-up periyoduna sahip örnek bir Markov Zinciri örneği Şekil 2.14’de sunulmuştur.



Şekil 2.14. Yakınsama problemi taşıyan Markov Zinciri (Koptur, 2020a)

- III. Efektif Örneklem Büyüklüğü İstatistiği (\hat{n}_{eff}):** Efektif örneklem büyüklüğü değerinin zincirlerin toplam uzunluğuna ($m \cdot n$) oranı 0.001 değeri ile kıyaslanır. 0.001'den büyük olması tercih sebebidir. m : Zincir sayısı, n : Her bir zincirin uzunluğu, $\hat{\rho}_t$: 1. Gecikmedeki otokorelasyon, T : $\hat{\rho}_{T+1} + \hat{\rho}_{T+2}$ toplamını negatif yapan ilk pozitif tam sayı olduğu durum için Efektif Örneklem Büyüklüğü,

$$\hat{n}_{eff} = \frac{m \cdot n}{1 + 2 \sum_{t=1}^T \hat{\rho}_t}$$

formülasyonu ile hesaplanır (Koptur, 2020b).

- IV. Gelman-Rubin İstatistiği (\hat{R}):** Gelman ve Rubin (1992) çalışmasında öne sürdüğü zincirler arası varyansın, zincirler içi varyans ile karşılaştırıldığı bir tanı testidir. Zincir içi varyansın, zincirler arası varyansa çok yakın olması beklenir. Bu neden ile bir birlerine oranlarının 1'e çok yakın olması istenir ve bu oran 1'e uzaklığı ile değerlendirilir. Böylece 1'e yakın orana sahip olan bağımsız Markov Zincirlerinin hedef dağılıma yakınsadıkları ve kararlı olduklarına karar verilir (Saçaklı Saçıldı, 2011).
- V. Maksimum Ağaç Derinliği:** HMC algoritması özyinelemeli hesaplamalar yaptığı için bu hesaplamaların derinliğinin 15'in üzerinde olması istenir.
- VI. Energy-Bayesian Fraction of Missing Information (E-BFMI):** HMC yöntemi gibi Momentum hesaplamaları ile pozisyonlar belirleyen yöntemlerde pozisyonlar güncellenirken sıçrama büyüklüğüne göre örnekleyicinin uzayı ne kadar dolaştığını ölçen bir tanı testine ihtiyaç duyulmaktadır. Bu istatistiğin 0.2'den az olmaması beklenir (Koptur, 2020b).

2.3.2. Bayesian Zaman Serisi Modelleri

Bayesian analizi için öncelikle olabilirlik fonksiyonuna ihtiyaç vardır. 2.2.5. Zaman Serisi Modelleri başlığında tek değişkenli zaman serileri modellerine ait olabilirlik fonksiyonunun doğal logaritma tabanında genel olarak ifade edecek olursak ARMA(p,q) için;

- Kareler toplamı; $S(\varphi, \theta) = \sum_{t=1-(p+q)}^T [a_t / \varphi, \theta, y]^2$
- Sürecin derecesine bağlı değişen bir fonksiyon; $f(\varphi, \theta)$, Örneğin ARMA(1,1) için $f(\varphi, \theta) = \sqrt{\frac{1-\varphi^2}{\theta^2+2\theta\varphi+1}}$

olduğu durum için olabilirlik fonksiyonu,

$$\mathcal{L}(\varphi, \theta | y) = (2\pi\sigma_a^2)^{-T/2} f(\varphi, \theta) \cdot \exp \left\{ -\frac{S(\varphi, \theta)}{2\sigma_a^2} \right\}$$

olarak ifade edilebilir. Olabilirlik fonksiyonunun önsel dağılım ile çarpılması ile sonsal dağılım elde edilebilir (Eğrioğlu, 2002). Bu durumu bilgi içermeyen üç farklı önsel için ifade edebiliriz. ARIMA modellerinin bilgi matrisi için, $\Sigma_{p,q}$ varyans kovaryans matrisi olmak üzere,

$$I(\varphi, \theta) = n\Sigma_{p,q}^{-1}$$

önerilmiştir (Box ve Jenkins 1976; Eğrioğlu 2002). İlk olarak önsel dağılımın Jeffreys önseli olduğu durumu inceleyelim. Önsel dağılım,

$$P(\varphi, \theta, \sigma_a) \propto |I(\varphi, \theta)|^{1/2} \cdot \sigma_a^2$$

olarak ifade edilir. Sonsal dağılım Jeffreys önseli için,

$$P_j(\varphi, \theta|y) \propto |I(\varphi, \theta)|^{1/2} \sigma_a^{-(n+1)} f(\varphi, \theta) \left\{ \frac{S(\varphi, \theta)}{2\sigma_a^2} \right\}$$

ifadesine eşit olacaktır. φ ve θ parametrelerinin marjinal sonsal dağılımı için sonsal dağılımın σ_a 'ye göre integralenmesi ile,

$$P_j(\varphi, \theta|y) \propto |I(\varphi, \theta)|^{1/2} f(\varphi, \theta) \{S(\varphi, \theta)\}^{-n/2}$$

eşitliği elde edilir. İkinci olarak Improper önseli kullanıldığı durumu inceleyelim. Önsel dağılım,

$$P(\varphi, \theta, \sigma_a) \propto |I(\varphi, \theta)| \cdot \sigma_a^{-1}$$

olarak ifade edilir. φ ve θ parametrelerinin marjinal sonsal dağılımı,

$$P_l(\varphi, \theta|y) \propto |I(\varphi, \theta)| f(\varphi, \theta) \{S(\varphi, \theta)\}^{-n/2}$$

ifadesine eşit olacaktır. Son olarak Uniform önsel seçimi yapıldığı durumu inceleyelim. Önsel dağılım,

$$P(\varphi, \theta, \sigma_a) \propto \sigma_a^{-1}$$

olarak ifade edilir. φ ve θ parametrelerinin marjinal sonsal dağılımı,

$$P_u(\varphi, \theta|y) \propto f(\varphi, \theta) \{S(\varphi, \theta)\}^{-n/2}$$

ifadesine eşit olacaktır. Aynı durumu farklı zaman serisi modelleri için inceleyebiliriz.

2.3.2.1. Bayesian Otoregresif Model AR(1)

Otoregresif modele ait Bayesian parametre tahmini prensibi gereği olabirlik fonksiyonu ve önsel dağılımın yaklaşık matematiksel çarpımlarından yararlanır. Otoregresif model için az bilgi içeren üç örnek dağılımın önsel olarak seçildiği durumu inceleyelim.

Otoregresif süreç için Jeffreys önseli ile olabirlik fonksiyonundan elde edilen sonsal dağılımı,

$$P_j(\varphi|y) \propto |I(\varphi)|^{1/2} f(\varphi) \{S(\varphi)\}^{-n/2}$$

şeklinde yazılır. Bu eşitlikte etkisi az olduğu için $|I(\varphi)|^{1/2}f(\varphi)$ ifadesi ihmal edildiğinde Sonsal dağılım,

$$P_j(\varphi|y) \propto \{S(\varphi)\}^{-n/2}$$

olarak yazılabilir. Parametre değerlerinin $\varphi_i = (1, \varphi_1, \varphi_2, \dots, \varphi_p)$ olduğu, simetrik kareler toplamlarının $\mathcal{M}_{ij} = \mathcal{M}_{ij} = y_i y_j + y_{i+1} y_{j+1} + \dots + y_{T+1-j} y_{T+1-i}$ olduğu ve $\hat{\varphi} = \frac{D_{i,i+1}}{D_{i+1,i+1}}$ olarak ifade edildiği durum için Jeffreys önseli kullanılan birinci derece sonsal dağılım,

$$P_j(\varphi|y) \propto \left(1 + \frac{(\varphi - \hat{\varphi})\mathcal{M}_p(\varphi - \hat{\varphi})}{v\sigma_a^2}\right)^{-n/2}$$

ifadesine eşittir (Eğrioğlu, 2002). Birinci derece Improper önseli için sonsal dağılım,

$$P_I(\varphi|y) \propto (1 - \varphi)^{-1/2}\{S(\varphi)\}^{-n/2}$$

olarak ve birinci derece Uniform önseli için sonsal dağılım,

$$P_U(\varphi|y) \propto (1 - \varphi)^{1/2}\{S(\varphi)\}^{-n/2}$$

olarak yazılır (Eğrioğlu, 2002).

2.3.2.2. Hareketli ortalama Modeli MA(1)

Hareketli ortalama modeli için az bilgi içeren üç örnek dağılımın önsel olarak seçildiği durumu inceleyelim. Jeffreys önseli ile olabilirlik fonksiyonundan elde edilen sonsal dağılımı;

$$P_j(\theta|y) \propto |I(\theta)|^{1/2}f(\theta)\{S(\theta)\}^{-n/2}$$

şeklinde yazılır. Aynı şekilde bu eşitlikte de etkisi az olduğu için $|I(\theta)|^{1/2}f(\theta)$ ifadesi göz ardı edilebilmektedir (Eğrioğlu, 2002). Bu durumda sonsal dağılım,

$$P_j(\theta|y) \propto \{S(\theta)\}^{-n/2}$$

olarak yazılabilir. Parametre değerlerinin $\theta_i = (1, \theta_1, \theta_2, \dots, \theta_p)$ olduğu, simetrik kareler toplamlarının $\mathcal{M}_{ij} = \mathcal{M}_{ij} = y_i y_j + y_{i+1} y_{j+1} + \dots + y_{T+1-j} y_{T+1-i}$ olduğu ve $\hat{\theta} = \frac{D_{i,i+1}}{D_{i+1,i+1}}$ olarak ifade edildiği durum için Jeffreys önseli kullanıldığı durumda hareketli ortalamalar sürecinin sonsal dağılımında kullanılan kareler toplamı,

$$S(\theta) = v\sigma_a^2 \propto \left(1 + \frac{(\varphi - \hat{\varphi})\mathcal{M}_p(\varphi - \hat{\varphi})}{v\sigma_a^2}\right)^{-n/2}$$

ifadesine eşit olacaktır (Eğrioğlu, 2002). Birinci derece Improper önseli için sonsal dağılım,

$$P_I(\theta|y) \propto (1 - \theta^2)^{-1/2} \{S(\theta)\}^{-n/2}$$

olarak ve birinci derece Uniform önseli için sonsal dağılım,

$$P_U(\theta|y) \propto (1 - \theta^2)^{1/2} \{S(\theta)\}^{-n/2}$$

olarak yazılır (Eğrioğlu, 2002).

2.3.2.3. Karma Otoregresif Hareketli Ortalama Modeli ARMA(1, 1)

Karma otoregresif hareketli ortalama modeli için Jeffreys önseli ve olabilirlik fonksiyonu kullanılarak elde edilen sonsal dağılım,

$$P_j(\varphi, \theta|y) \propto |I(\varphi, \theta)|^{1/2} \sigma_a^{-(n+1)} f(\varphi, \theta) \left\{ \frac{S(\varphi, \theta)}{2\sigma_a^2} \right\}$$

şeklinde yazılır. Eşitlikteki $|I(\varphi, \theta)|^{1/2} f(\varphi, \theta)$ ifadesi tamamen göz ardı edilemez ancak daha sade bir ifade ile yer değiştirilebilir (Eğrioğlu, 2002).

$$\begin{aligned} & (y_t - \varphi_1^* y_{t-1} - \varphi_2^* y_{t-2} - \dots - \varphi_{p+q}^* y_{t-(p+q)}) \\ & = (y_t - \varphi_1 y_{t-1} - \varphi_2 y_{t-2} - \dots - \varphi_p y_{t-p}) \\ & \times (y_t - \varphi_1 y_{t-1} - \varphi_2 y_{t-2} - \dots - \varphi_q y_{t-q}) \end{aligned}$$

eşitliğinden φ^* elde edilir ve $|I(\varphi, \theta)|^{1/2} f(\varphi, \theta) = J(\varphi^*|\varphi, \theta)$ eşitliği kullanılarak yeni sonsal dağılım,

$$P_j(\varphi, \theta|y) \propto J(\varphi^*|\varphi, \theta) \{S(\varphi, \theta)\}^{-n/2}$$

eşitliği ile ifade edilebilir. ARMA(1,1) modeli için $\varphi^* = \varphi + \theta$, $\varphi_2^* = -\varphi\theta$ ve $J = |\varphi - \theta|$ eşitlikleri yardımı ile

$$P_j(\varphi, \theta|y) \propto |\varphi - \theta| \{S(\varphi, \theta)\}^{-n/2}$$

orantısı elde edilir. ARMA(1,1) modeli için Improper önseli için sonsal dağılım,

$$P_I(\varphi, \theta|y) \propto |\varphi - \theta| \left(\frac{1 - \varphi^2}{\theta^2 + 2\varphi\theta + 1} \right)^{-1/2} \{S(\varphi, \theta)\}^{-n/2}$$

olarak ve birinci derece Uniform önseli için sonsal dağılım,

$$P_U(\varphi, \theta|y) \propto \left(\frac{1 - \varphi^2}{\theta^2 + 2\varphi\theta + 1} \right)^{1/2} \{S(\varphi, \theta)\}^{-n/2}$$

olarak elde edilir (Eğrioğlu, 2002).

2.4. Zaman Serilerinin Parametre Tahmin Yöntemleri

Gecikme uzunlukları belirlenmiş deneme niteliğindeki model parametrelerinin belirlenmesi aşamasında çeşitli yöntemler kullanılmaktadır. Tahmin aşamasında kullanılan yöntemler temel olarak iki kategoriye ayrılırlar. Bu yaklaşımlar Frekansçı Yaklaşım ve Bayesian Yaklaşım'dır. İki kategoriye ayrılmış bu yaklaşımların başlıca farkları Çizelge 2.3'de ki gibi özetlenebilir.

Çizelge 2.3. Parametre tahmini yaklaşımlarının karşılaştırılması

Frekansçı Yaklaşım	Bayesian Yaklaşım
Hipotezler ve sonuçlar için eşitlikleri kullanır. Yani parametreler bilinmeyen bir sabittir, veriler rastgeledir.	Hipotezler ve sonuçlar için olasılıkları kullanır. Yani parametreleri tanımlamak için olasılık dağılımlarını kullanır bu nedenle parametreler rastgeledir, veriler sabittir.
Hem gözlemler hem de öngörüler sadece mevcut veri setine bağlıdır.	Gözlemlerin önsel bilgisine ve mevcut bilgisine bağlı çıkarımlar yapar.
Tümdengelim özelliği sergiler.	Tümevarım özelliği sergiler.
Herhangi bir ön bilgi ihtiyacı gerektirmez.	Objektif veya sübjektif önsel bir bilgi gerektirir.
Hesaplama yoğunluğu azdır. Daha az işlem kapasitesi gerektirir.	Hesaplama yoğunluğu yüksektir. Daha fazla işlem kapasitesi gerektirir.

(Koptur, 2020a; Orloff and Bloom, 2014)

Çizelge 2.3' de özetlendiği gibi Frekansçı yaklaşımda parametre değerinin bilinmeyen bir değer olduğu düşünülür ve bu katsayıya en yakın nokta tahminleri yapılması için tahminleme çalışmaları yapılır veya katsayının bir noktaya eşitliği test edilir. Ancak Bayesian yaklaşım parametre değerlerinin tamamının olasılık dağılımını tespit ederek olabilecek tüm mümkün durumların olasılık yoğunluk fonksiyonu hesaplanır. Bir nokta tahminin sonsal dağılım içinde olup olmadığını tespit eder (Wakefield, 2005). Önsel olarak kullanılan bilginin güvenilirliğinin yüksek olduğu durumlar için Bayesian yaklaşım ile yapılan parametre tahminleri Frekansçı yöntem sonuçlarına göre daha iyi sonuçlar verecektir.

Frekansçı yaklaşım kullanılan veri setini inceleyerek tahminleme çalışmaları yapar. Bayesian yaklaşım önsel bir dağılım bilgisini kullanarak mevcut veriye ait bilgileri kullanarak analiz yapar. Bayesian tahmin yönteminde kullanılan bilgiler dağılımlara dönüştürülür: önsel bilgiden önsel dağılıma, mevcut bilgiden olabilirlik fonksiyonuna bir geçiş vardır. Sonrasında elde edilmiş sonsal dağılımlardan bilgi edinilmesi söz konusudur (Ünal, 2003).

Frekanscı yaklaşım veri setinin tüm bilgi olduğu varsayımı ile spesifik bilgilere ulaşmayı hedefler. Buna karşın Bayesian yaklaşım mevcut bilginin tümün bir parçası olduğunu ve kısıtlı bilgi ile tüm bilgiye ulaşılabileceğini varsayar (Wakefield, 2005).

ARIMA model parametreleri Frekanscı yaklaşım ile belirlenirken genellikle En Küçük Kareler (EKK) tahmin yöntemi ile hata terimlerinde görülen otokorelasyon etkisi nedeni ile etkin olmayan tahminler elde edileceğinden Doğrusal Olmayan EKK yöntemi veya hata terimlerinin normal dağıldığı varsayımı ile En Çok Olabilirlik Yöntemi kullanılmaktadır (Akgül, 2003; Karaman, 2010). Bilgisayar işlem gücündeki güncel yenilik ve gelişmeler sonucunda yoğun işlem kapasitesi gerektiren Bayesian yaklaşım ile çoklu parametre tahmini yaygın hale gelmiştir. (Eğrioğlu, 2002; Fan and Yao, 2008; Ghosh and Basu ve O'Mahony vd, 2007; Tripathi vd., 2017).

Laktasyon dönemi süt verimine ait modelleme çalışmalarında şimdiye dek Frekanscı yaklaşım yöntemleri Doğrusal Olmayan EKK yöntemi veya En Çok (Maksimum) Olabilirlik yöntemleri kullanılmıştır. Bu çalışmada hem Bayesian yaklaşım ile parametre tahmini hem de Frekanscı En Çok (Maksimum) Olabilirlik yöntemi ile parametre tahmini gerçekleştirilmiştir.

2.4.1. Frekanscı Yaklaşım ile Parametre Tahmini

1990 yılına kadar birçok alanda tek değişkenli zaman serilerinin modellenmesinde Frekanscı yaklaşım kullanılarak *ARIMA* model parametreleri belirlenmiştir. Ancak süt veriminin modellenmesinde ilk bilenen çalışma 1990 yılında Deluyker vd. (1990) yapmış olduğu laktasyon dönemindeki 513 Holstein ırkı ineğe ait günde üç kez toplanan veriler ile gerçekleştirilmiştir. Bu çalışmada laktasyon eğrisi modellenmesi için kısmi ve tam kayıtlar kullanılarak *ARIMA* model parametreleri belirlenmiştir. Zaman serisi yöntemlerinin laktasyon eğrilerinin modellenmesinde deterministik modellere iyi bir alternatif olarak sunulabileceğini belirtmişlerdir.

2000 yılında Macciotta vd. çalışmalarında laktasyon döneminin modellenmesi için kontrol günlerini zamanın bir fonksiyonu olarak değil de farklı bir metot ile analize dahil etmişlerdir. 1200 Sarda koyuna ait aylık ölçülmüş 7'şer süt verim kayıtlarını ardı ardına dizerek zaman serisi analizi için kullanılmıştır. 1200 Sarda koyunu 6 gruba ayrılmışlardır ve grupların ilk 3 tanesi doğum sayılarına göre, diğer 3 grup ise çiftliklerin buldukları yerlere (dağ, ova ve tepede olmasına) göre oluşturulmuştur. Her koyuna ait 7'şer laktasyon kaydı kullanılmasına ek olarak 2'şer ve 4'er laktasyon kaydını kullanıldığı durumlar için farklı Box-Jenkins modelleri denenmiştir. Çalışmanın sonucunda mevsimsel *ARIMA* modellerinin süt verim modellerinin deterministik ve stokastik özelliğine uyumlu olduğunu, kayıp gözlemleri tahmin etmede ve kısa dönemli öngörü yapmada diğer yöntemlere göre daha kullanışlı olduğunu ve geliştirilmesi gerektiğini belirtmişlerdir.

Yine Macciotta vd. (2002) çalışmalarında 2000 İtalyan Simmental ırkı ineğine ait süt verimlerini, süt yağ ve protein değerlerini kullanmışlardır. Bu çalışmada süt verim kayıtları 2000 yılındaki çalışmalarındaki metot ile düzenlenmiştir ve 2000 adet ineğin 8'er kaydı birleştirilerek veri seti oluşturulmuştur. İneklerin buzağılama (1. 2. ve 3. laktasyon dönemi) sayılarına göre 3 farklı zaman serisini modellemişlerdir. Çalışmada toplam süt verimi için 3, yağ ve protein değerleri için 6 model tahmin

edilmiştir. ARMA modellerinin süt ve protein değerlerini modellenmesinin yağ değerlerinin modellenmesinden daha başarılı olduğunu belirtmişlerdir. Sonuç olarak mevsimsel ARMA modelinin süt verimlerinin öngörüsünde karmaşık modellerle kıyaslandığında uygulamasının daha kolay olduğunu ve dikkate değer sonuçlar verdiğini ifade etmişlerdir.

Cappio-Borlino vd. (2004) çalışmalarında Macciotta vd. (2002) çalışmalarındaki koyunların laktasyon dönemine ait süt verim kayıtları yardımıyla tahmin edilmiş model ile 2. ve 4. kontrol günü bilgilerini kullanarak tahmin edilen laktasyon sürecini incelemişlerdir. Kısmi kontrol günü kayıtlarıyla kurulan uygun zaman serisi modelinin tüm laktasyona ait süt verimini tahmin etmede yeterli ve başarılı olduğunu belirtmişlerdir.

Berberoğlu (2010) çalışmasında 41.058 adet siyah alaca ırkı ineğin laktasyon dönemi süt verim kayıtlarını buzağılama sayısına göre 7 gruba ayırmıştır. Süt verimi verileri günlük toplandığı için veri zamanın bir fonksiyonu olarak oluşturulmuştur. Her bir laktasyon sayısı için bir model tahmin edilerek toplamda 7 *ARIMA* modeli tahmin edilmiştir. Örneğin 1. laktasyonda olan ineklerin kontrol günleri verimleri analiz edilerek *ARIMA*(10,1,7) modeline karar verilmiş ve aynı kontrol günleri kayıtları kullanılarak Wood'un geliştirmiş olduğu Gamma modeli tahmin edilerek karşılaştırılmıştır. Karşılaştırma sonuçlarına göre tüm laktasyon dönemlerinde *ARIMA* modellerinin ortalama hata kareler toplamı değerleri daha küçük bulunmuştur. Çalışma sonucunda stokastik unsur barındıran Zaman serisi modellerinin Laktasyon dönemi için kısa dönemli öngörülerde, deterministik model olan Gamma modeline göre daha başarılı olabileceğini belirtmiştir.

Karaman (2010) çalışmasında kaçınıcı laktasyon döneminde oldukları bilgisi bulunmayan 23.873 ineğe ait 10'ar kontrol günü verim kayıtlarını kullanmıştır. Farklı laktasyon dönemlerinde olabilecek hayvanlara ait modelleme yapabilmek için benzer laktasyon özelliğindeki 1.070 ineğin 10'ar kontrol günü verim değerlerini Macciotta vd. (2000) ve Macciotta vd. (2002) çalışmalarında kullanılan yöntem ile bir zaman serisine dönüştürmüştür.

En uygun model olarak *ARIMA*(2,0,0)(1,1,1)₁₀ modeli seçilmiştir. Macciotta vd. (2000) ve Macciotta vd. (2002) çalışmalarında kullanılan öngörü algoritması ile kısmi kontrol günü kayıtları ile eksik kayıtları tahmin etmiştir. Gerçek değerler ile tahmin edilen değerler arasında yüksek ilişki katsayıları hesaplanmıştır. Sonuçlar zaman serisi yöntemlerinin laktasyon eğrilerini modellenmede kullanışlı olabileceğini ve farklı veri oluşturma yöntemleri ile zaman serisi analiz yöntemlerinin kullanılabileceğine örnek bir çalışmadır.

Teke (2017) çalışmasının veri seti 80 siyah alaca ırkı ineğinin günlük verim kayıtlarından oluşmaktadır. Böylece Berberoğlu (2010) çalışmasında ki gibi veri seti zamanın bir fonksiyonu olarak oluşturulmuştur. Teke çalışmasında ineklerin 1 den 5'e laktasyon (buzağılama) sayısı için 5 farklı modelleme çalışması yapmıştır. Çalışmasında her laktasyon dönemini hem tek değişkenli zaman serisi modeli olan *ARIMA* modelleri ile hem de doğrusal olmayan modellerden olan Gamma modeli ve *MARS* (Çok değişkenli uygulanabilir regresyon modeli) ile tahmin etmiştir.

Zaman serisi modelleri tahmini her bir laktasyon için ayrı ayrı yapılmıştır ve en uygun modele mümkün durumlar denenerek karar verilmiştir. Her bir laktasyon için tahmin edilen tek değişkenli zaman serisi modeli, tahmin edilen Gamma fonksiyonları ile karşılaştırılmış ve 5 laktasyonun her birinde uygun tek değişkenli zaman serisi modellerinin Gamma fonksiyonlarına göre daha fazla açıklama oranına sahip olduğu tespit edilmiştir. *MARS* Modelleri *ARIMA* modellerine göre laktasyon dönemlerinin modellenmesinde bu çalışma için daha başarılı bulunmuştur.

Laktasyon döneminin tek değişkenli zaman serisi ile modelleme çalışmaları veri oluşturma metotlarına göre 2'ye ayrılarak özetlenmiştir. Çizelge 2.4'de laktasyon dönemi süt verim kayıtları günlük olan ve veri setinin zamanın bir fonksiyonu olan çalışmalarda öngörü için kullanılan *ARIMA* modellerinin gecikme değerleri özetlenmiştir.

Çizelge 2.4. Zamanın fonksiyonu olan veri setleri ile laktasyon dönemi modelleri

Yazar(lar)	Yıl	İnek Sayısı	Buzağılama (Laktasyon) Sayısı	Model
Deluyker vd.	1990	513 (Holstein)		<i>ARIMA</i> (0,1,1)
Berberoğlu	2010	41 058 (Siyah Alaca)	1	<i>ARIMA</i> (10, 1, 7)
			2	<i>ARIMA</i> (4,1,9)
			3	<i>ARIMA</i> (10, 1, 10)
			4	<i>ARIMA</i> (6,1,5)
			5	<i>ARIMA</i> (8, 1, 9)
			6	<i>ARIMA</i> (9,1,9)
			+7	<i>ARIMA</i> (9, 1, 10)
Teke	2017	80 (Siyah Alaca)	1	<i>ARIMA</i> (1, 1, 2)
			2	<i>ARIMA</i> (2,1,0)
			3	<i>ARIMA</i> (1, 1, 2)
			4	<i>ARIMA</i> (1,1,2)
			5	<i>ARIMA</i> (2, 1, 1)

Veri oluşturulma metodu zamanın bir fonksiyonu olmayan ve süt verimleri ayda bir kontrol edilerek oluşturulmuş veri setleri ile yapılan modelleme çalışmalarında karar verilen modeller Çizelge 2.5'de özetlenmiştir.

Çizelge 2.5. Laktasyon dönemi mevsimsel modelleri

Yazar(lar)	Yıl	Canlı Türü-Sayısı	Box-Jenkins Modeli
Macciotta vd.	2000	1200 (Sarda Koyunu)	ARIMA(2,0,0)(1,0,1) ₇
Macciotta vd.	2002	2000 (Simmental Sığırı)	ARIMA(1,0,1)(1,0,1) ₈
Karaman	2010	10 070 (Süt Sığırı)	ARIMA(2,0,0)(1,1,1) ₁₀

Çizelge 2.5’de özetlenen çalışmalarda kontrol günü sayısı adedi mevsimsel periyod sayısını (s) belirlemiştir. Yalnızca Karaman (2010) çalışmasında mevsimsel fark alınmıştır. Üç çalışmada da mevsimsel parametreler kullanılmıştır.

Laktasyon döneminin zaman serisi modelleri ile modellenmesi çalışmalarının genelinde En Çok Olabilirlik (Maximum likelihood) yöntemi kullanılmıştır. Bu yöntemin mantığı olabilirlik fonksiyonuna dayanmaktadır. $y_t, t = 1, 2, \dots, T$ zaman serisi için olabilirlik fonksiyonu varsayımsal bir durum için bazı açıklamalar şöyledir.

- Sürecin derecesine bağlı değişen bir fonksiyon; $f = (\varphi, \theta)$,
- Kareler toplamı; $S(\varphi, \theta) = \sum_{t=1-(p+q)}^T [a_t / \varphi, \theta, Y]^2$

olduğu durum için olabilirlik fonksiyonu,

$$\mathcal{L}(\varphi, \theta, \sigma_a) = f = (\varphi, \theta) - \frac{T}{2} \ln(\sigma_a^2 2\pi) - \frac{S(\varphi, \theta)}{\sigma_a^2 2\pi}$$

şeklinde ifade edilir (Eğrioğlu, 2002). Olabilirlik fonksiyonunda yer alan $\theta = \theta_1, \theta_2, \dots, \theta_q$ ve $\varphi = \varphi_1, \varphi_2, \dots, \varphi_p$ parametre değerlerinin elde edilebilmesi için olabilirlik dağılımını maksimum yapan parametre değerleri araştırılır. (Olabilirlik fonksiyonunun parametre değerleri için türevleri alınarak sıfıra eşitlenir.)

Genel olarak En Çok Olabilirlik yöntemi Frekansçı Yaklaşım nedeni ile yığının olasılık dağılımını oluşturan parametre değerlerini bir nokta olarak araştırdığı için olabilirlik fonksiyonunu maksimum yapan parametre değerlerini tahminci olarak elde eder.

2.4.2. Bayesian Yaklaşım ile Parametre Tahmini

Elde edilen sonsal dağılımlar parametrelere ait tüm bilgileri içermektedir. Pratikte kullanılabilmesi için genellikle özetleyici göstergeler kullanılır. Parametre sayısı ikiden fazla ise grafiksel olarak gösterilemez ancak marjinal dağılımlarına ait grafikler çizilebilir. Ayrıca aralık ve nokta tahminleri kullanılabilir (Koptur, 2020a).

Sonsal dağılımdan çıkarım yapmak için dağılımı en iyi temsil eden değer hesaplanabilir. Dağılımın yapısına göre dağılıma ait ortalama, medyan veya mod değerleri tercih edilebilir (Koptur 2020b; Mathy ve Chekaf 2018).

En iyi temsil eden parametre değerini tespit etmek için bir değer seçilebileceği gibi bir değer geçireliliği incelenebilir. Bunun için sınanan değer dağılıma ait %95 yüksek yoğunluk bölgesinde olup olmadığı kontrol edilebilir. Bu incelemede eger değer sonsal dağılımın güvenilir aragında değilse geçireliliği red edilir. Ancak geçireliliğinin kabul edilmesinin incelenmesi (ROPE) pratik eşitlik bölgesi kavramını ortaya çıkarmıştır. İncelenen değer yerine değere ait herhangi bir aralığın en yüksek aralığın içersinde olup olmadığı kontrol edilir (Gelman vd. 2013; Koptur 2020b).

Zaman serisi modellerinin parametrelerinin ve parametresinin tahmin edildiği bazı çalışmalar yapılmıştır. Örneğin $ARIMA(1,1,0)_{12}$ modeline ait parametreleri belirleyen Eğrioğlu (2002) çalışmasında Bayesian Zaman Serisi analizi üzerine teorik ve matematiksel olarak detaylı bir anlatıma yer vermiştir.

Uygulama olarak da Samsun kükürdioksit verilerini kullanmıştır. Aylık frekanstaki veriyi öncelikle stokastik özellikteki mevsimsel bileşeninden arındırmak için 12. periyotta 1. mevsimsel fark almıştır. ACF ve PACF grafikleri yardımı ile gecikme uzunluklarına karar verilmiştir. Çalışma sonucunda $ARIMA(1,1,0)_{12}$ modelinin parametre tahminleri önce Frekanscı yaklaşım ile tahmin edilmiştir.

Daha sonra Bayesian yaklaşımda bilgi içermeyen önsel dağılımlardan; Jeffreys önsel, Improper ve Uniform önselleri için parametrelere ilişkin sonsal dağılımlar hesaplanmıştır. İşlem kolaylığı ön planda tutulan bu çalışmada karmaşık hesaplamalardan kaçınılmıştır.

Detaylı matematiksel açıklamalarla Bayesian Zaman Serisi analizlerinin yapıldığı çalışmaya bir diğer örnek Fan ve Yao (2008) çalışmasıdır. Zaman serisi modelleri olan $AR(p)$, $MA(q)$ ve $ARMA(p, q)$ modellerinin Bayesian yaklaşım ile parametrelerinin tahmininin matematiksel formlarını incelemişlerdir ve bir uygulama örneği vermişlerdir.

Uygulama olarak, Şanghai emlak fiyat endeksi için öngörü modeli oluşturulmuştur. Zaman seri modellerinden $ARIMA(0,0,3)$ modeli uygun model olarak seçilmiştir. Bayesian zaman serisi modeli ile yapılan öngörülerin dijital sonuçlara göre iyi bir performans gösterdiği raporlanmıştır.

İstatistik paket programları kullanılarak Bayesian zaman serisi analizlerinin yapıldığı çalışmalara örnek Ghosh ve ark (2007) yaptıkları çalışmada gösterilebilir. Çalışmada trafik ışıklarını kısa süre için tahmin etmede Bayesian zaman serisi analizi gerçekleştirmişlerdir.

Veri seti 15'er dakikada bir toplanmış verilerden oluşmaktadır. Bu neden ile $\left(\frac{24\text{Saat}}{15\text{Dakika}}\right)$ 96. mevsimsel fark alınarak seri mevsimsellikten arındırılmıştır. Uygun model olarak $ARIMA(1,0,0)(0,1,1)_{96}$ belirlenmiştir.

Önsel olarak bilgi içermeyen Uniform Önsel tercih edilmiştir. Analizde sonsal dağılımın hesaplanması için MCMC yöntemlerinin bir türevidir olan Metropolis-Hastings yöntemi kullanılmıştır. Dublin şehir merkezinden elde edilen veriler ile modele ait dalgalanmaları iyi temsil edecek, başarılı sonuçlar elde edilmiştir.

Litaratüre bakıldığında genellikle bilgi içermeyen önsel seçilerek zaman serisi parametreleri belirlenmiştir. Bilgi içermeyen önsel seçimi yapan çalışmalardan biride Tripathi ve ark (2017) çalışmalarıdır. Çalışmalarında Hindistan'a ait yıllık frekansta toplanmış gayrisafi yurtiçi hasıla (GDP) verilerini kullanarak ARMA zaman serisi sürecini Bayesian yaklaşım ile analiz etmişlerdir.

Gecikme uzunluklarına maksimum 2 gecikmeye kadar olası kombinasyonlar denenerek en iyi modele bilgi kriterleri yardımı ile karar verilmiştir. En iyi model olarak *ARIMA(0,0,1)* modeli belirlenmiş ve parametre tahminleri için önsel seçiminde bilgi içermeyen önsel seçimi yapılmıştır. Sonsal dağılımdan örneklem çekmek için Gibbs örneklemesinden faydalanmışlardır. Verilerin kısa vadeli tahminleri gerçekleştirilmiştir ve sonuçlar tatmin edici bulunmuştur (Ghosh et al., 2007).

Bayesian yaklaşım ile parametre tahmini gerçekleştirilen tek değişkenli zaman serisi modelleme çalışmalarında genellikle bilgi içermeyen önsel seçimi gerçekleştirilmiştir. Az bilgi içeren önsel seçimine rağmen sonuçlar veri setleri ile uyumlu elde edilmiş ve kısa dönemli öngörü için yeterli bulunmuştur.

3. METERYAL VE METOT

3.1. Materyal

Çalışmanın materyalini 7.973 İngiliz Siyah Alaca Sığırı sürüsünden elde edilen 10'arlı kontrol günlerine ait litrelik verim kayıtları oluşturmaktadır. Kayıtlar İngiltere Ulusal Süt Kayıtları Birliği tarafından Kasım 1988 tarihi ile Ekim 1989 tarihleri arasında elde edilmiştir. Her bir kontrol günü ayın bir gününde toplanan süt miktarından oluşmaktadır. Veri seti 40 denenmiş, 649 denenmemiş boğanın toplam 23.873 dişi yavrusundan oluşmaktadır.

Analize dahil edilecek hayvanlara karar verilirken laktasyon kayıtları incelenmiş ve kaydında problem olduğu düşünülen veya tipik laktasyon sürecine uygun olmayan üretim değerlerine sahip ineklerden oluşan 3.031 adet inek veri setinden çıkartılmıştır.

Son olarak veri setindeki homojenliği artırmak için ortalama süt veriminden ± 1 standart sapmadan fazla sapsmaya sahip verim kaydı olan 3.872 hayvan veri setinden çıkartılarak 1.070 ineğe ait verim kayıtları ile çalışma materyali elde edilmiştir. Çalışma için veri seti oluşturulurken hayvanların kaçınıcı laktasyonda oldukları bilgisi var olmadığı için veri 23.873 ineğe ait kayıtlardan benzer verime sahip 1.070 sığıra ait kayıtlar kullanılmıştır. Veri seti ikiye bölünerek 535 inek analiz için, 535 inek analizin kontrolü için kullanılmıştır.

3.2. Metot

Bu çalışmada veriler hem Frekanslı yaklaşım hem Bayesian yaklaşım ile zaman serileri analiz yöntemleri çerçevesinde analiz edilmiştir. Model parametrelerinin tahmin edilmesine kadar izlenen prosedürler klasik zaman serisi analiz yöntemlerine aittir. Ancak hayvan materyali analize hazırlanırken zamanın bir fonksiyonu olarak oluşturulan zaman serilerinden farklı olarak Macciotta vd. (2000), Macciotta vd. (2002) ve Karaman (2010) çalışmalarında uygulanan bir yaklaşımla hazırlanmıştır. Macciotta vd (2000) çalışmalarında index değerinin sadece zaman ifadesi olmak zorunda olmadığını ve ancak veri setini sıralamak için bir araç olarak kullanıldığını ifade etmişlerdir.

Zamanın bir fonksiyonu olarak oluşturulan zaman serilerinden farklı olan bu veri seti ile çalışılması bazı terminolojik farklılıkları beraberinde getirmiştir. Bu çalışmada ve Karaman (2010) çalışmasında kullanılan bazı terminolojik eşitlikler şöyledir.

Zaman → İndeks

Ay → Kontrol Günü

Yıl → Laktasyon Dönemi

Mevsim → Tipik Laktasyon Davranışı (pik ve azalış dönemleri dalgaları)

Bu terminolojik kıyaslama ile birlikte analiz basamakları ve sonuçlar daha anlaşılır olacaktır. Bu çalışmanın materyalini oluşturan veri setin zamanın bir

fonksiyonu olmadan Çizelge 3.1'deki gibi oluşturulmuştur ve veri setini kullanarak seri klasik zaman serisi yöntemleri ile analiz edilmiştir.

Çizelge 3.1. Veri seti (Karaman, 2010; N. P. P. Macciotta and Pulina, 2000; N. P. P. Macciotta et al., 2002)

İnek	Kontrol Günü	Verim	İndeks
1	1	*	1
1	2	*	2
1	3	*	3
1	4	*	4
1	5	*	5
1	6	*	6
1	7	*	7
1	8	*	8
1	9	*	9
1	10	*	10
2	1	*	11
2	2	*	12
2	3	*	13
2	4	*	14
2	5	*	15
2	6	*	16
2	7	*	17
2	8	*	18
2	9	*	19
2	10	*	20
.	.	.	.
.	.	.	.
.	.	.	.
535	1	*	5341
535	2	*	5342
535	3	*	5343
535	4	*	5344
535	5	*	5345
535	6	*	5346
535	7	*	5347
535	8	*	5348
535	9	*	5349
535	10	*	5350

Analiz ve kontrol grubu olarak veri seti ikiye ayrılmadan önce veri setinde bulunan 1.070 inek rastgele olarak sıralanmıştır. Daha sonra ilk 535 sığır analiz için, son 535 sığır kontrol için ayrılmıştır. Her iki gruptaki sığırlara ait 10'ar laktasyon kaydı alt alta yazılmıştır. Analiz veri seti (535 × 10) 5350 gözlemden oluşmaktadır. Aynı şekilde kontrol veri seti de (535 × 10) 5350 gözlemden oluşmaktadır. Böylece ilk on

gözlem bir hayvana ait sıralı on laktasyon süt veriminden oluşmaktadır. Örneğin 12. sıradaki gözlem 2. ineğe ait ikinci laktasyon kayıdır.

Veri seti R programının 3.6.3 versiyonu x86_64-w64-mingw32 platform üzerinde çalıştırılmıştır. R programında kullanılan kütüphanelerin isimleri ve versiyonları şunlardır: urca_1.3-0, tidyr_1.0.2, shinystan_2.5.0, shiny_1.4.0.2, ruler_0.2.2, rstantools_2.0.0, rstan_2.19.3, ggplot2_3.3.0, StanHeaders_2.21.0-1, Metrics_0.1.4, magrittr_1.5, readr_1.3.1, forecast_8.11, dplyr_0.8.5, bayesplot_1.7.1, brms_2.12.0, Rcpp_1.0.3, aTSA_3.1.2, assertr_2.7.

Model inşası tek değişkenli zaman serisi Box-Jenkins model kurma stratejisine uygun gerçekleştirilmiştir. Veri setinin oluşturulmasında kullanılan metot gereği her on gözlemlerde tipik laktasyon eğrisi meydana geldiği için modelde toplamsal özellikte olduğu tespit edilen deterministik mevsimsel bileşen çarpımsal olarak ayrıştırılmıştır. Böylece veri setinin oluşturulma metodundan kaynaklı meydana gelen dalgalanmalar seriden önce ayrıştırılmış daha sonra (model inşasının ardından) modele dahil edilmiştir. Öngörü algoritması oluşturulurken öncelikle Çizelge 3.2 de ki adımlar takip edilmiştir.

Çizelge 3.2. Öngörü algoritması

Kullanılan Gerçek Kontrol günü Sayısı	Tahmin Edilen Kontrol Günü Sayısı	Kalıntı Seri Numarası
İlk 2	Son 3-10→8 Adet	1
İlk 3	Son 4-10→7 Adet	2
İlk 4	Son 5-10→6 Adet	3
İlk 5	Son 6-10→5 Adet	4
İlk 6	Son 7-10→4 Adet	5
İlk 7	Son 8-10→3 Adet	6
İlk 8	Son 9-10→2 Adet	7

Yedi adet kalıntı serisi oluşturulmuş ve kalıntılar için tanımlayıcı istatistik değerleri hesaplanmıştır. Böylece geriye kalan kontrol günleri tahmini değerleri ile gerçek değerleri arasındaki korelasyonlar incelenerek optimum kontrol gününe karar verilmiştir.

4. BULGULAR

Veri setine bulunan 1.070 sığıra ait 10'ar adet laktasyon kaydı bulunmaktadır. Kontrol günlerine ait ortalama değerler incelendiğinde en yüksek süt verimi ortalaması ikinci aya ait kontrol gününde gözlenmiştir. Sığırların süt verimindeki en fazla farklılık 1. kontrol gününde gerçekleşmiştir. Veri setinin tanımlayıcı istatistikleri Çizelge 4.1'de sunulmuştur.

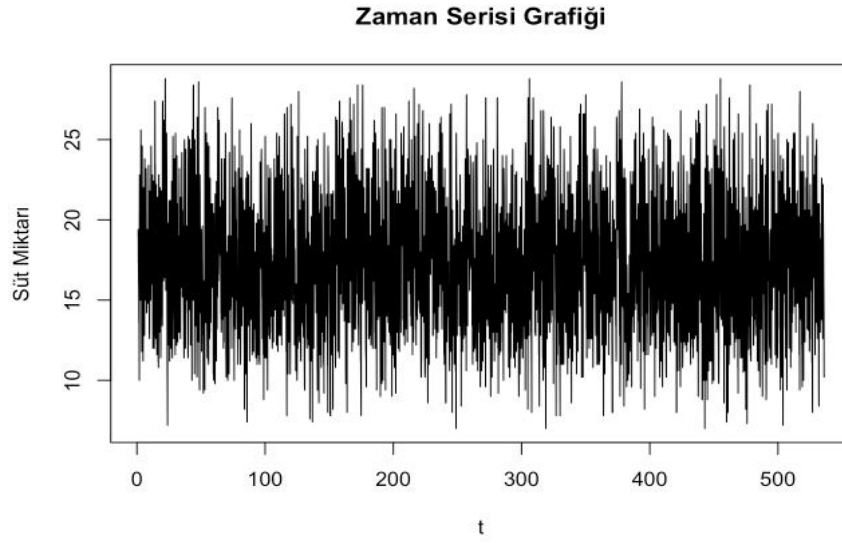
Çizelge 4.1. Veri setinin tanımlayıcı istatistikleri

Kontrol Günü	N	Ortalama	Std. Sapma	Min. Değer	Max. Değer
1	1070	19,42	3,1321	5,80	28,20
2	1070	22,59	2,7548	15,20	29,40
3	1070	20,46	2,4378	14,20	25,80
4	1070	18,72	2,3739	12,60	25,00
5	1070	17,07	2,2787	12,00	23,40
6	1070	15,43	2,2693	11,00	22,60
7	1070	15,35	2,5203	10,40	21,80
8	1070	14,91	2,4650	9,80	21,20
9	1070	14,27	2,6218	8,80	20,80
10	1070	13,03	2,8995	6,80	19,40

Çizelge 4.1'de özetlenen tanımlayıcı istatistikler tipik laktasyon dönemi özellikleri ile örtüşmektedir. Veride rastgele sıralanan gözlemlerin ilk 5.350 adedi analiz için 5.350 adedi kontrol için ayrılmıştır.

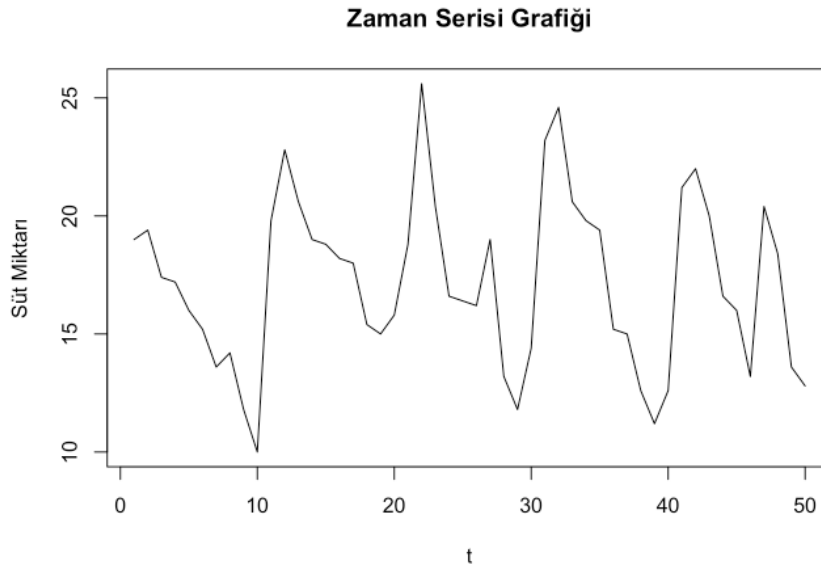
4.1. Belirleme

Belirlenme basamağı için öncelikle analiz grubunun zaman yolu grafiği incelenmiştir. Analiz grubu serisinin index değerlerine karşılık gelen görseli (zaman serisi grafiği) Şekil 4.1'de sunulmuştur.



Şekil 4.1. Veri setinin orijinal zaman serisi grafiđi

Şekil 4.1’de veri seti oldukça uzun olduđu için serideki mevsimsel dalgalanmalar tam olarak görülememektedir. Bu neden ile Şekil 4.2’de veri setinin ilk 50 gözlemi için oluşturulmuş zaman serisi grafiđi yer almaktadır.



Şekil 4.2. İlk 50 gözlem için orijinal seri için zaman serisi grafiđi

Şekil 4.2’de her on gözlemden aynı sistematik artış azalış davranışı görülmektedir. Serinin grafiđinde trend bileşeni bulunmamaktadır. Ancak her on gözlemden benzer deterministik dalgalar görülmektedir. Bu veri oluşturulma metodunun kaçınılmaz bir sonucudur.

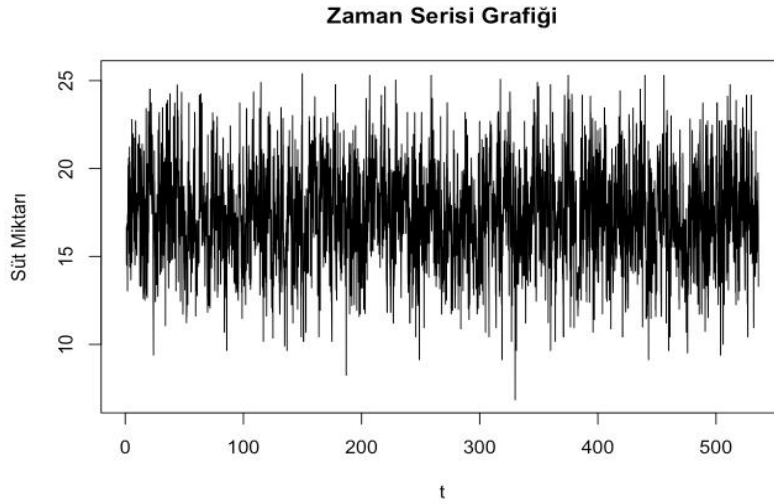
Seriye, seride bulunan deterministik yapıdaki mevsimsel bileşenden arındırma işlemi uygulanır. Arındırma işlemi mevsimsel bileşenin seriden ayrıştırılması ile gerçekleşir. İki tür ayrıştırmadan bahsedilmiştir; toplamsal ve

çarpımsal ayrıştırma. Seriyeye ait toplamsal ve çarpımsal ayrıştırmaya ait hata kareler toplamları Çizelge 4.2’de özetlenmiştir.

Çizelge 4.2. Toplamsal ve çarpımsal ayrıştırma yöntemlerine ait istatistikler

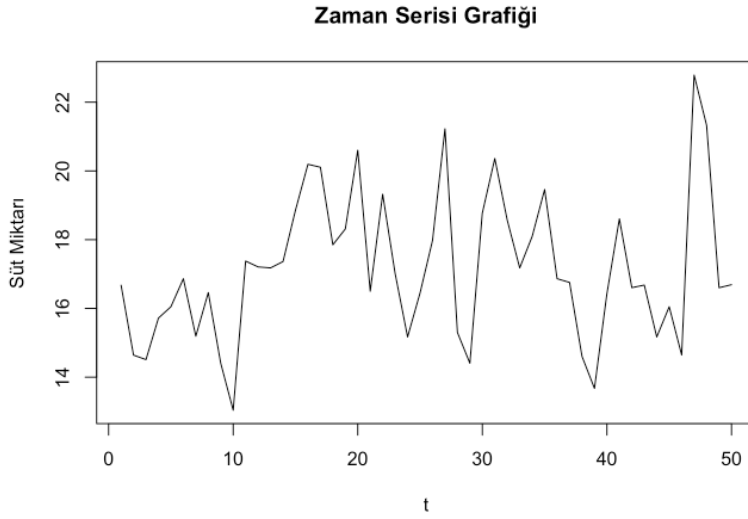
Model: MY_t	RMSE	MAE
Toplamsal Ayrıştırma	1,773063	1,367484
Çarpımsal Ayrıştırma	1,003961	0,9968151

Çizelge 4.2’de görüldüğü gibi toplamsal ayrıştırmaya için hesaplanan istatistikler daha büyüktür ve bu çarpımsal ayrıştırma yapılmaya dayanağıdır. Ayrıştırma işleminden sonra elde edilen seriyeye ait zaman grafiği Şekil 4.3’de sunulmuştur.



Şekil 4.3. Mevsimsellikten arındırılmış zaman serisi grafiği

Arındırılmış serideki düzenli ritmik dalgalanmalar azalmıştır ancak gözlem değeri ($t = 5350$) yüksek olduğu için grafikte tam olarak ayırt edilememektedir. Arındırma işlemi verinin elde edilme yönteminin seride oluşturduğu yapay dalgalanma halinden arındırılması olarak ifade edilebilir. Arındırılmış serinin ilk 50 gözlemi için zaman serisi grafiği Şekil 4.4’de sunulmuştur.



Şekil 4.4. İlk 50 gözlem için arındırılmış seri için zaman serisi grafiği

Şekil 4.4’de görüldüğü gibi her on gözlemde görülen sistemik dalgalanmadan seri arındırılmıştır. Böylece arındırılmış seriye durağanlık sınaması yapılarak doğru sonuçlar elde edilebilir. Durağanlık sınamasına ait sonuçlar Çizelge 4.3’de sunulmuştur. ADF ve PP testlerinin yokluk hipotezi seri durağan değil iken KPSS test hipotezi seri durağandır şeklindedir.

Çizelge 4.3. Durağanlık sınamalarının özet sonuçları

y_t = Mevsimsellikten arındırıldı, Sabitli, Trendsiz Model

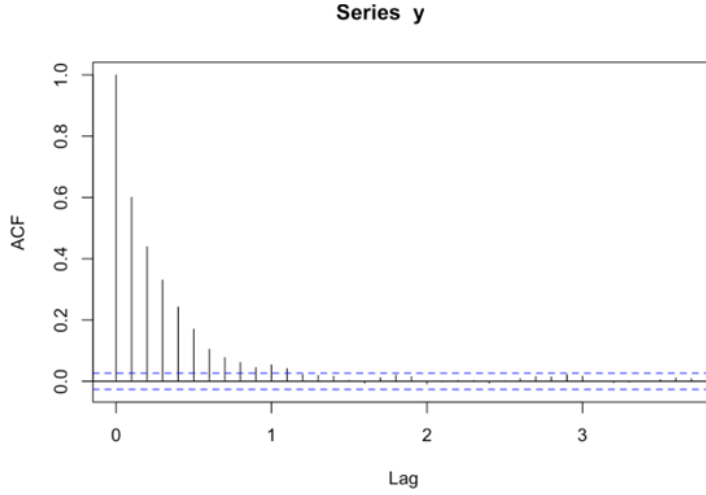
Test	Seri	Test İstatistiği	P-Değeri(Prob)	Karar(%95)
ADF	y_t	-36,5	0.01	Seri durağandır.
PP	y_t	-2411	0.01	Seri durağandır.
KPSS	y_t	0.0475	0,10	Seri durağandır.

Çizelge 4.3’de ilk iki test Artırılmış Dickey-Fuller ve Phillips-Perron Testidir. Test denkleminde sabitli, trendsiz test denklemini kullanılmış ve trendsiz, sabitsiz denklemler için kritik değerler hesaplatılmıştır. Otokolerasyon problemi için eklenen (10’a kadar) tüm gecikme uzunlukları için P-değerleri 0,05’in altında hesaplanmıştır. Böylece ADF testi için zaman serisinin 0,95 güvenirlilikte durağan olduğuna karar verilmiştir.

Çizelge 4.3’de 3. sıradaki test Kwiatkowski-Phillips-Schmidt-Shin Testidir. Sınama için kullanılan diğer iki testten farklı olarak serinin durağanlığını sınanan bir testtir. Bu nedenle sabitli, trendsiz model için P-değeri 0,10, kritik değer olan 0,05’ten daha büyük olduğu için serinin 0,95 güvenirlilikte KPSS testi içinde

durağan olduğu sonucuna varılmıştır. Sonuç olarak üç durağanlık sınaması test sonucunda da serinin %95 güven düzeyinde durağan olduğuna karar verilmiştir.

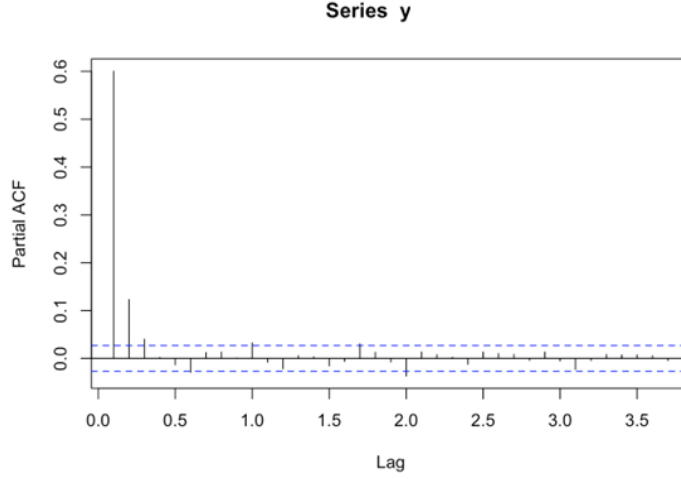
Sonraki aşama otokorelasyon grafikleri incelenerek gecikme uzunluklarına karar verilmesi aşamasıdır. Seriyeye ait otokorelasyon grafiği Şekil 4.5’de sunulmuştur. Şekilde bulunan yatay eksen gecikme uzunluklarını diğer eksen ise otokorelasyon katsayılarına karşılık gelmektedir. Ayrıca yatay eksen etrafındaki kesikli yatay eksen %95 güvenilirlik düzeyini göstermektedir. Bu güven aralığının üstünde kalan katsayılar istatistiksel olarak anlamlı kabul edilmiştir.



Şekil 4.5. Seriyeye ait otokorelasyon grafiği

Şekil 4.5’de görülebileceği gibi 1 gecikme sonrasındaki gecikmelerde anlamlı bir otokorelasyon katsayısı bulunmamaktadır. Bu durumda q değerinin 1 olduğuna karar verilebilir.

Seriyeye ait kısmi otokorelasyon fonksiyonuna ait grafik Şekil 4.6’da verilmiştir. Yatay eksen gecikme uzunluklarına, dikey eksen kısmi otokorelasyon katsayılarına aittir. Yatay eksen etrafındaki kesikli çizgiler ile belirlenmiş alan %95 güvenilirlikte anlamlılık düzeyidir. Anlamlılık seviyesinin dışında kalan katsayıların istatistiksel olarak anlamlı olduklarını simgelemektedir.



Şekil 4.6. Seriyeye ait kısmi otokorelasyon grafiği

Şekil 4.6’da ki kısmi otokorelasyon fonksiyonu grafiğinde görüldüğü gibi 1. gecikmeden sonra anlamlı bir katsayı bulunmamaktadır. Bu görünüm p değerinin 1’e eşit olduğunu göstermektedir. Otokorelasyon ve kısmi otokorelasyon fonksiyonu grafiklerinin görsel inceleme sonucuna göre $ARIMA(1,0,1)$ modeli deneme modeli olarak tespit edilmiştir. Ancak gecikme uzunluklarını doğrulayabilmek için tüm olası gecikmelerin denendiği Şekil 4.7’de ki R paket programı komutu ile olası modellerden en iyi model seçimi hesaplatılmıştır. Böylece bilgi kriterlerinden AIC ve BIC yardımı ile $ARIMA(1,0,1)$ modeli doğrulanmıştır.

```
fit.auto.nonseasonal <- auto.arima(y , seasonal = FALSE)
fit.auto.nonseasonal

## Series: y
## ARIMA(1,0,1) with non-zero mean
```

Şekil 4.7. R paket programı en iyi model seçimi kodu ve sonucu

Deneme modeli gecikme uzunluklarına karar verildikten sonra model parametrelerinin tahmin edilerek geçerliliklerinin sınanması gerekmektedir. Model öncelikle Frekansçı yaklaşım yöntemi ile daha sonra Bayesian yaklaşım yöntemi ile tahmin edilmiştir.

4.2. Tahmin: Frekansçı Yaklaşım ile Parametre Tahmini

$ARIMA(1,0,1)$ modeli parametreleri Frekansçı yaklaşım ile analiz edilerek model parametreleri En Çok Olabilirlik (Maksimum Likelihood) yöntemi ile tahmin edilmiştir. Modele ait tanımlayıcı istatistikler Şekil 4.8’de sunulmaktadır.

```

Series: ymult
ARIMA(1,0,1) with non-zero mean

Coefficients:
      ar1      ma1  intercept
      0.7355  -0.2153   17.2568
      ---
      z value  Pr(>|z|)
ar1      48.9477 < 2.2e-16 ***
ma1      -9.9037 < 2.2e-16 ***
intercept 200.6774 < 2.2e-16 ***
      ---
signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```

Şekil 4.8. Frekanscı modele ait tanımlayıcı istatistikler

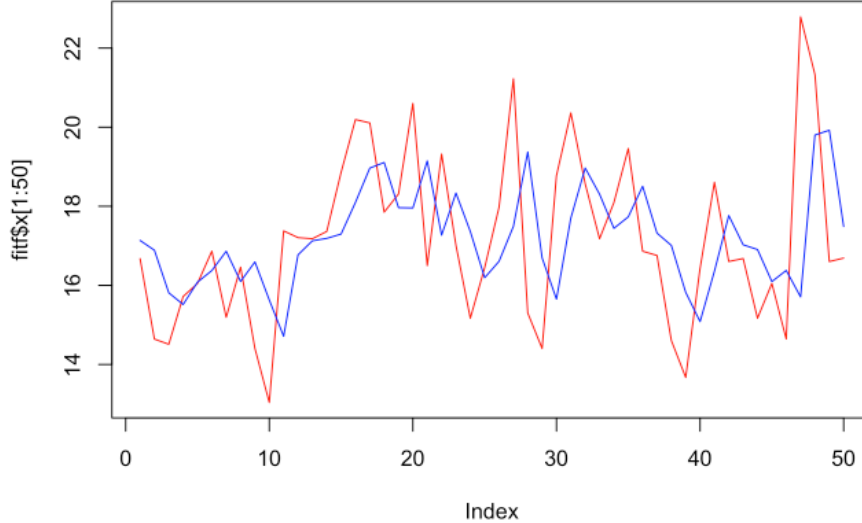
Frekanscı yaklaşım modelinin açık hali;

$$(1 - 0,7355B)(y_t - 17,2568) = (1 + 0,2153B)\varepsilon_t$$

veya

$$\hat{y}_t = 17,2568 - 0,7355 \times 17,2568 + 0,7355 \cdot y_{t-1} - 0,2153 \cdot \varepsilon_{t-1}$$

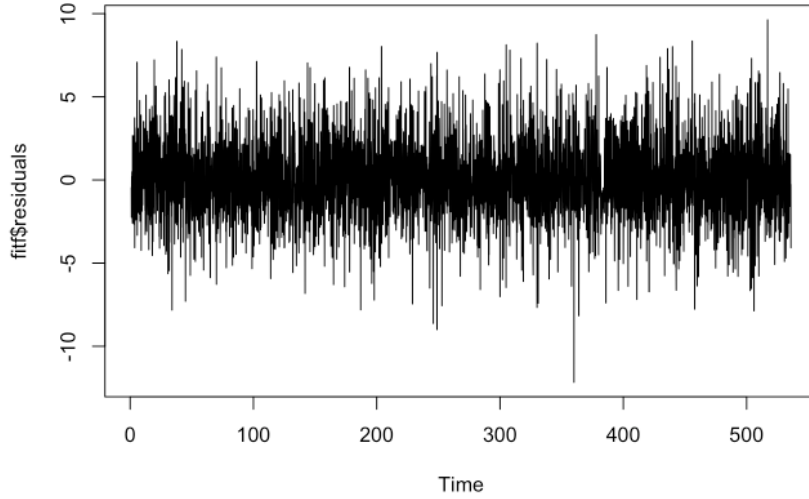
olarak ifade edilebilir. İlk 50 gözlem için gerçek değerlere ait (kırmızı) zaman yolu grafiği ve tahmin edilen değerlere ait (mavi) zaman yolu grafiği birlikte çizdirilmiştir ve sonuçlar Şekil 4.9'da sunulmuştur.



Şekil 4. 9. Gerçek (kırmızı) ve öngörü (mavi) değerlerinin zaman yolu grafikleri

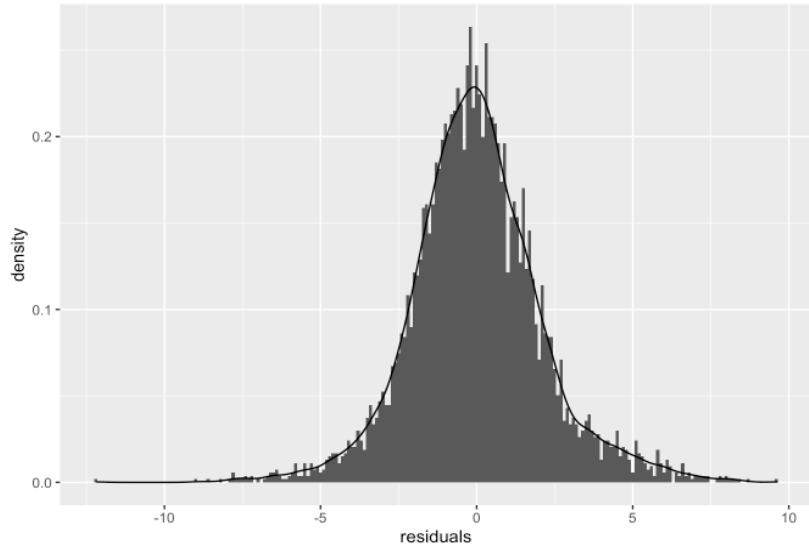
4.3. Ayırt Edici Kontrol: Frekanscı Yaklaşım Modeli Kalıntılarının Kontrolü

Sonraki adım model geçerliliği için $y_t - \hat{y}_t = \varepsilon_t$ kalıntı serinin sağlaması gereken varsayımların sınanmasıdır. Kalıntı serisinin otokolerasyonsuz ve temiz dizi özelliğinde olması özelliği sınanmıştır. Şekil 4.10'de Frekanscı zaman serisi modeline ait hataların zaman yolu grafiği sunulmuştur.



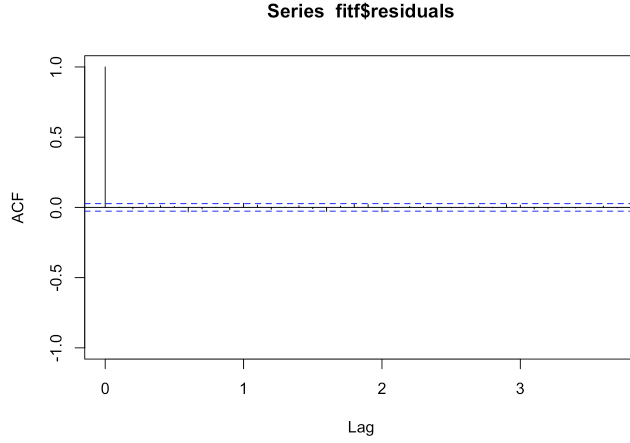
Şekil 4.10. Frekanscı zaman serisi modeline ait kalıntıların zaman serisi grafiği

Şekil 4.10'da kalıntılara ait zaman yolu grafiğine göre hataların 0 ortalama etrafında rassal olarak salındığı gözlemlenmiştir. Şekil 4.11'de hatalara ait histogram grafiği verilmiştir.



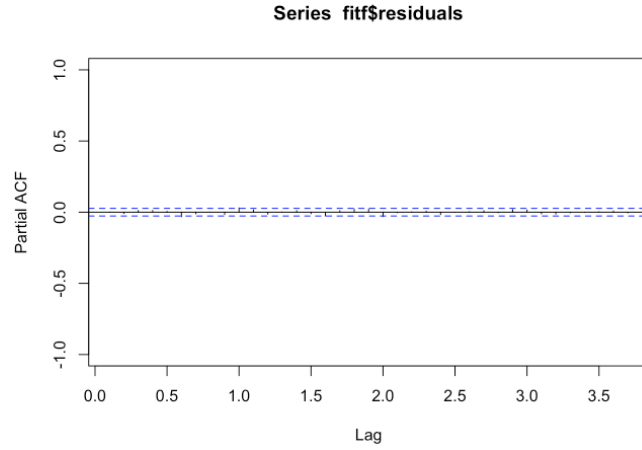
Şekil 4.11. Frekanscı modelin kalıntılarına ait histogram grafiği ve normal dağılım eğrisi

Şekil 4.11'e göre hata terimlerinin 0 ortalamaya sahip olduğu ve normal çan eğrisine uyumlu bir dağılım gösterdiği tespit edilmiştir. Kalıntı serisinde serisel korelasyon incelemesi için Şekil 4.12'de otokolerasyon katsayıları grafiği ve Şekil 4.13'de kısmi otokolerasyon katsayıları grafiği incelenmiştir.



Şekil 4. 12. Frekanscı modelin kalıntılarına ait serisinin otokolerasyon katsayıları grafiği

Şekil 4.12’de kalıntı serisinin geçmiş dönemleri ile anlamlı bir ilişkisi bulunmadığı görülmektedir. Kalıntı serisi için kısmi otokolerasyon katsayıları grafiği Şekil 4.13’de sunulmuştur.



Şekil 4.13. Frekanscı modelin kalıntılar serisine ait kısmi otokolerasyon katsayıları grafiği

Şekil 4.13 incelendiğinde kalıntı serisinde anlamlı kısmi otokolerasyon katsayısı gözlenmemiştir. Kalıntılarda serisel otokolerasyonsuz olması varsayımı istatistiksel olarak sınanması için kalıntı serisi için Box-Pierce test istatistikleri hesaplanmıştır ve sonuçlar Şekil 4.14’de verilmiştir.

Autocorrelation Check of Residuals									
To Lag	Chi-Square	DF	Pr > ChiSq	Autocorrelations					
6	855.92	4	<.0001	-0.392	0.071	-0.005	0.006	0.013	-0.029
12	868.32	10	<.0001	-0.000	0.011	-0.029	0.031	0.007	-0.018
18	873.89	16	<.0001	0.004	0.011	-0.002	-0.026	0.013	0.006
24	886.79	22	<.0001	0.024	-0.034	0.010	-0.002	0.013	-0.020
30	890.83	28	<.0001	0.005	-0.002	0.012	-0.014	0.013	0.014
36	892.67	34	<.0001	-0.013	-0.005	0.001	-0.003	-0.002	0.011
42	901.64	40	<.0001	-0.011	0.014	-0.005	0.015	-0.026	0.021
48	906.81	46	<.0001	-0.018	0.011	-0.004	-0.003	0.011	-0.019

Şekil 4.14. Frekanscı modelin kalıntılarına ait Box-Pierce Test sonuçları

Şekil 4.14’de özetlenen sonuçlarda tüm gecikmeler için Pr değerlerinin tamamı 0,05 den küçük olduğu için %95 güvenirlilikte kalıntı serisinde serisel korelasyon olmadığına karar verilmiştir.

Bu durumda $ARIMA(1,0,1)$ modeli kalıntıları temiz dizi özelliğindedir ve bu modelin uygun bir model olduğuna karar verilmiştir. Modelin açık hali;

$$(1 - 0,7355B)(y_t - 17,2568) = (1 + 0,2153B)\varepsilon_t$$

veya

$$y_t = 4,5644 + 0,7355 \cdot y_{t-1} - 0,2153 \cdot \varepsilon_{t-1} + \varepsilon_t, \quad \varepsilon_t \sim N(0, \sigma^2)$$

şeklinde elde edilir.

4.4. Tahmin: Bayesian Yaklaşım ile Parametre Tahmini

$ARIMA(1,0,1)$ modelinin açık hali

$$(1 - \Phi B)(y_t - \mu) = (1 + \Theta B)\varepsilon_t$$

veya

$$y_t = \mu - \Phi\mu + \Phi y_{t-1} + \Theta \varepsilon_{t-1} + \varepsilon_t, \quad \varepsilon_t \sim N(0, \sigma^2)$$

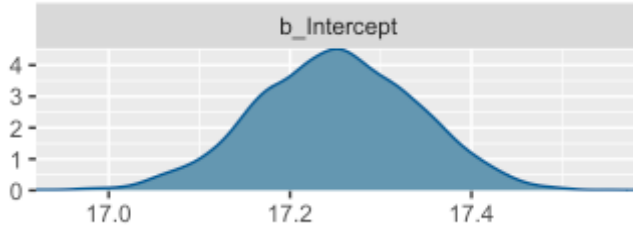
formüldeki gibidir. Parametreleri Bayesian yaklaşım ile tahmin edilebilmesi için ilk adım önsel belirlenmesi ve olabilirlik fonksiyonu ile çarpılarak sonsal dağılımına ait olasılık yoğunluk fonksiyonunun elde edilmesi aşamasıdır.

Literatürde parametre tahmini için önsel belirlemesi hakkında yapılmış bir çalışma bulunmamaktadır. Ayrıca literatürde kontrol günlerinin modellenmesinde $ARIMA(1,0,1)$ modeli daha önce kullanılmamıştır. Bu durumda parametrelerin tüm değerleri alabilmesine imkân verecek az bilgi içeren önseller tercih edilmiştir.

Önsel Dağılımlar için;

■ Sabit terim için: $\mu \sim Normal(0,10)$

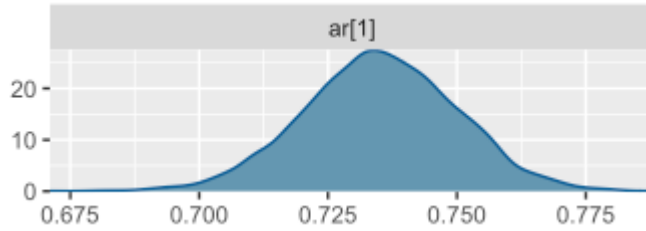
Popülasyon ortalaması μ değeri için 0 ortalamalı ve 10 standart sapmalı normal dağılım kullanıldı ve sonsal dağılımı Şekil 4.14'de sunulduğu gibi tahmin edilmiştir.



Şekil 4.15. Sabit terime ait sonsal dağılım grafiği

■ Tahminlere ait otoregresif parametre için: $\Phi \sim Normal(0,2)$

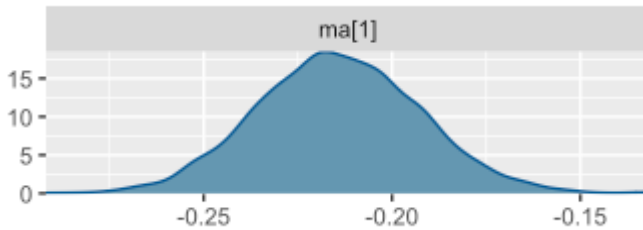
Φ değeri için 0 ortalamalı 2 standart sapmalı normal dağılım kullanılmıştır ve Φ değerinin ± 1 aralığında olması için kısıtlandı. Sonsal dağılımı Şekil 4.15'de sunulduğu gibi tahmin edilmiştir.



Şekil 4.16. Tahminlere ait otoregresif parametre için sonsal dağılım grafiği

■ Kalıntılara ait otoregresif parametre için: $\Theta \sim Normal(0,2)$

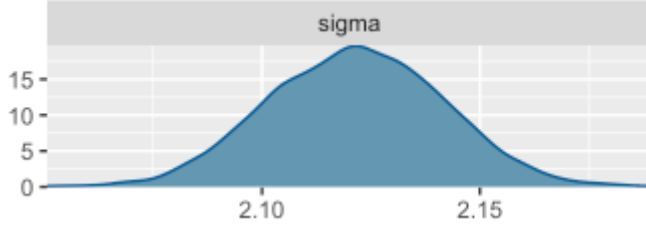
Θ değeri için 0 ortalamalı 2 standart sapmalı normal dağılım kullanılmıştır ve Θ değerinin ± 1 aralığında olması için kısıtlandı. Sonsal dağılımı Şekil 4.16'de sunulduğu gibi tahmin edilmiştir.



Şekil 4.17. Kalıntılara ait otoregresif parametre için sonsal dağılım grafiği

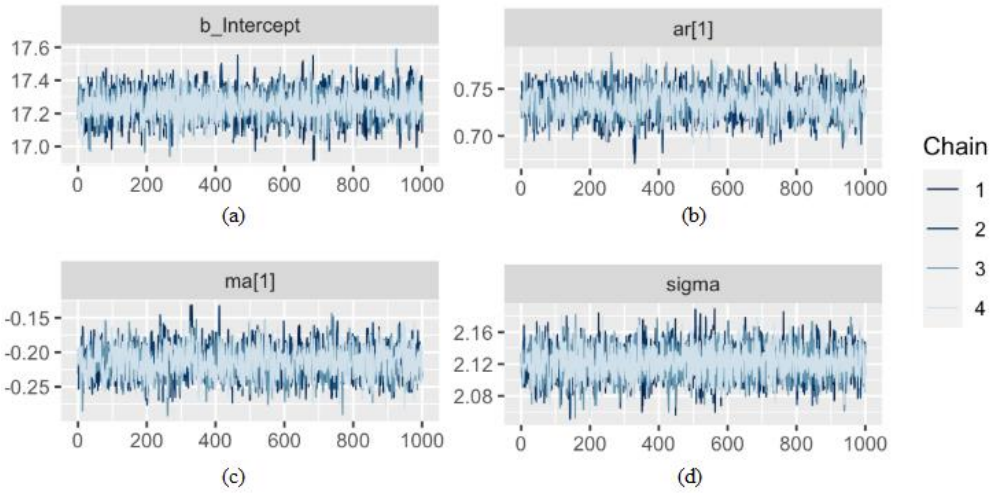
■ Kalıntılar için: $\varepsilon_t \sim Student_t(3,0,10)$

ε_t için 3 serbestlik dereceli, sıfır ortalamalı ve 10 standart sapmalı Student-t dağılımı önsel olarak tercih edilmiştir. Standart sapma için sıfırdan küçük olmama şartı ile sınırlandırılmıştır. Sonsal dağılımı Şekil 4.17’de sunulduğu gibi tahmin edilmiştir.



Şekil 4.18. Kalıntılar için sonsal dağılım grafiği

Tahmin edilmiş olan sonsal dağılımlardan çıkarım yapılabilmesi için gerekli varsayımlar vardır. Sonsal dağılımdan MCMC yöntemi varyasyonlarından biri olan NUTS yöntemi ile örneklem çekilerek sonsal dağılımlar tahmin edildiği için örneklem seçimi için kullanılan Markov zincirleri incelenir. İlk olarak Şekil 4.19’deki Trace grafikleri incelenerek yakınsama problemi veya serisel otokolerasyon davranışı bulunup bulunmadığı kontrol edilmiştir.



Şekil 4.19. Markov Zincirleri grafikleri **a)**Sabit terim; **b)** Tahminlere ait otoregresif parametreler; **c)** Kalıntılara ait otoregresif parametreler; **d)** Kalıntılar

Şekil 4.19’deki Trece grafiklerinde her bir değişken için 4’er (m) Markov Zinciri oluşturulmuştur. Her bir zincir uzunluğu 2000 olarak seçilmiş ancak ilk 1000 iterasyon Warm-up değeri olarak zincirden çıkartılmış ve kullanılan zincir uzunluğu (n) 1000 olarak belirlenmiştir. Her parametre için 4 zincire ait grafikler üst üste çizdirilmiştir. 4 parametre içinde yakınsama problemi yaşanmadığı, rassal seyir izledikleri ve birbirlerine benzer olduğu görülebilmektedir.

Sonrasında Çizelge 4.4’de tüm parametreler için zincirler arası varyansın zincirler içi varyansa oranları 1’den büyük olma şartı incelenmiştir. Etkili örneklem büyüklükleri ($m \times n$) 4000 değerine bölünerek 0,01’den büyüklükleri incelenmiştir.

Çizelge 4.4. Etkili örneklem büyüklüklerinin zincir uzunlukları toplamına oranı

Parametre	$n_{eff}/m \cdot n$
μ Sabit Terim	3971/4000 = 0,9927
Φ ar[1]	2554/4000 = 0,6385
Θ ma[1]	2562/4000 = 0,6405
Sigma	2610/4000 = 0,6525

Çizelge 4.4’de görüldüğü gibi tüm etkili örneklem büyüklükleri yeterli (0,01’den büyük) bulunmuştur. Zincirler için iraksama, ağaç derinliği ve E-BFMI kontrol edilmiştir. Sonuçlar Şekil 4.20’de sunulmuştur.

```

Divergences:
0 of 4000 iterations ended with a divergence.

Tree depth:
0 of 4000 iterations saturated the maximum tree depth of 10.

Energy:
E-BFMI indicated no pathological behavior.

```

Şekil 4.20. Iraksama, Ağaç Derinliği ve E-BFMI sınama R raporu

Şekil 4.20’de ki rapora göre zincirde iraksak değer bulunmamaktadır, zincirde ağaç derinliği 10’dan az olan değer bulunmamaktadır ve E-BFMI göstergesine göre örneklem uzayı yeterli bir şekilde dolaşmıştır. Dört zincir için E-BFMI değerleri; [1.1035 1.1787 1.2692 0.9649] olarak hesaplanmıştır. Çalışmalarda 0,2 den büyük olmaları beklenir. Görüldüğü gibi zincirlere ait E-BFMI değerleri oldukça yeterlidir.

Parametre tahmini için kullanılan simülasyon çalışmaları ve NUTS yöntemi varsayımları sağlanmıştır. Model parametreleri Bayesian yaklaşım ile zaman serisi modeli parametreleri elde edilmiştir ve modele ait tanımlayıcı istatistikler Şekil 4.21’de verilmiştir.

```

Family: gaussian
Links: mu = identity; sigma = identity
Formula: y ~ arma(p = 1, q = 1)
Data: df (Number of observations: 5350)
Samples: 4 chains, each with iter = 2000; warmup = 1000; thin = 1;
         total post-warmup samples = 4000

Correlation Structures:
      Estimate Est.Error l-95% CI u-95% CI   Rhat Bulk_ESS
ar[1]  0.7354   0.0152  0.7043  0.7642 1.0010    2554
ma[1] -0.2147   0.0216 -0.2564 -0.1715 1.0012    2562
      Tail_ESS
ar[1]    2346
ma[1]    2328

Population-Level Effects:
      Estimate Est.Error l-95% CI u-95% CI   Rhat Bulk_ESS
Intercept 17.2520   0.0855 17.0823 17.4193 0.9997    3971
      Tail_ESS
Intercept 2963

Family Specific Parameters:
      Estimate Est.Error l-95% CI u-95% CI   Rhat Bulk_ESS
sigma    2.1225   0.0201  2.0847  2.1616 1.0025    2610
      Tail_ESS
sigma    2726

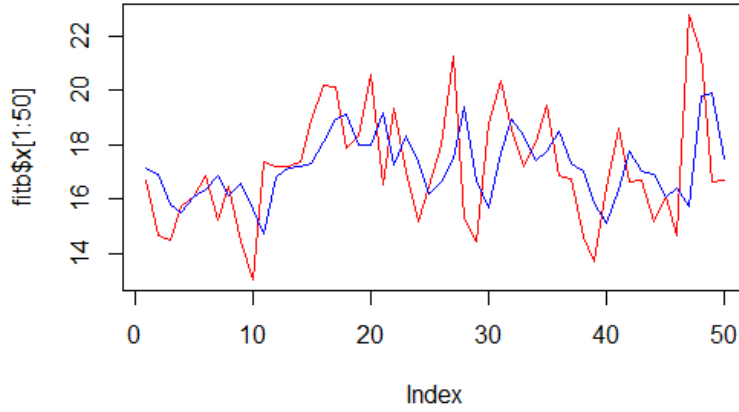
```

Şekil 4.21. Bayesian yaklaşım ile model parametrelerinin tahminin R programı sonuçları

Şekil 4.21’de R sonuçlarına göre Bayesian zaman serisi modeli açık hali;

$$\hat{y}_t = 17,252 - 0,7354 \times 17,252 + 0,7354y_{t-1} - 0,2147 \cdot \varepsilon_{t-1}$$

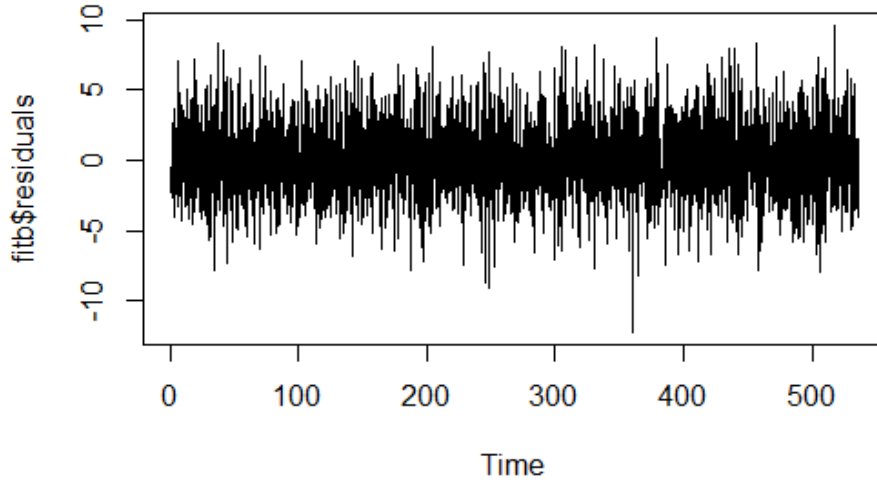
olarak tahmin edilmiştir. İlk 50 gözlem için gerçek değerlere ait (kırmızı) zaman yolu grafiği ve tahmin edilen değerlere ait (mavi) zaman yolu grafiği birlikte çizdirilmiştir ve sonuçlar Şekil 4.22’de sunulmuştur.



Şekil 4.22. Gerçek ve öngörü değerlerinin zaman yolu grafikleri

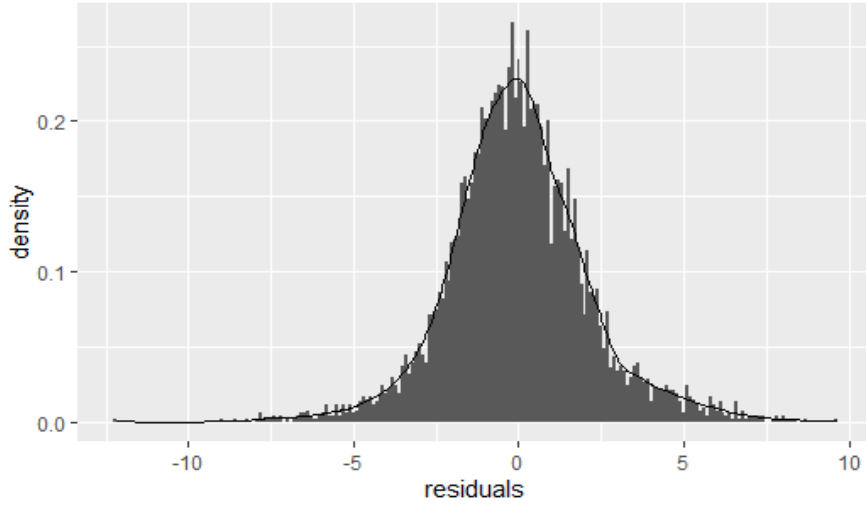
4.5. Ayırt Edici Kontrol: Bayesian Yaklaşım Modeli Kalıntılarının Kontrolü

Sonraki adım model geçerliliği için $y_t - \hat{y}_t = \varepsilon_t$ kalıntı serinin sağlaması gereken varsayımların sınanmasıdır. Kalıntı serisinin otokolerasyonsuz ve temiz dizi özelliğinde olması özelliği sınanmıştır. Şekil 4.23’de Bayesian zaman serisi modeline ait hataların zaman yolu grafiği verilmiştir.



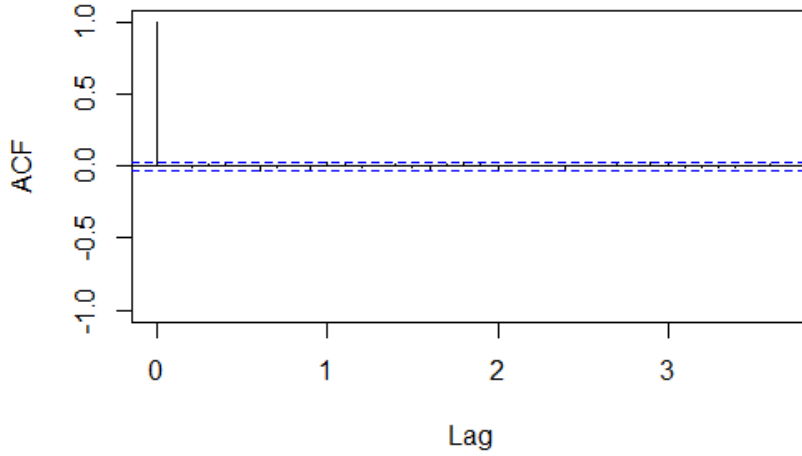
Şekil 4.23. Bayesian zaman serisi modeline ait kalıntıların zaman serisi grafiği

Şekil 4.23’de kalıntılara ait zaman yolu grafiğine göre hataların 0 ortalama etrafında rassal olarak salındığı gözlemlenmiştir. Şekil 4.24’de hatalara ait histogram grafiği verilmiştir.



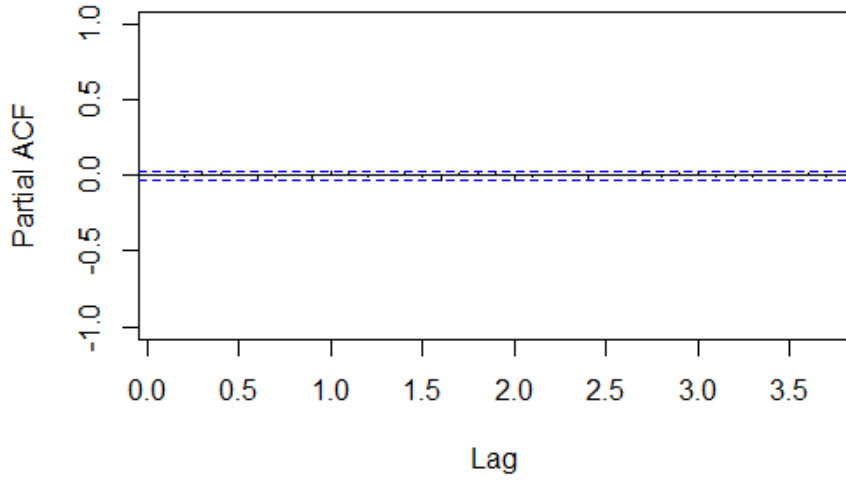
Şekil 4.24. Bayesian modelin kalıntılarına ait histogram grafiği ve normal dağılım eğrisi

Şekil 4.24'e göre hata terimlerinin 0 ortalamaya sahip olduğu ve normal eğrisine uyumlu bir dağılım gösterdiği tespit edilmiştir. Kalıntı serisinde serisel korelasyon incelemesi için Şekil 4.25'de otokolerasyon katsayıları grafiği ve Şekil 4.26'da kısmi otokolerasyon katsayıları grafiği sunulmuştur.



Şekil 4.25. Bayesian modelin kalıntılarına ait serisine ait otokolerasyon katsayıları grafiği

Şekil 4.25'de kalıntı serisinin geçmiş dönemleri ile anlamlı bir ilişkisi bulunmamaktadır. Kalıntı serisi için kısmi otokolerasyon katsayıları grafiği Şekil 4.26'da ki gibi elde edilmiştir.



Şekil 4.26. Bayesian modelin kalıntılarına serisine ait kısmi otokolerasyon katsayıları grafiği

Şekil 4.26'ı incelendiğinde kalıntı serisinin anlamlı kısmi otokolerasyon katsayısı göstermediği gözlenmiştir. Kalıntılarda serisel otokolerasyonsuz olması varsayımı sınanması için kalıntı serisinin Box-Pierce test istatistikleri hesaplanmıştır ve sonuçlar Şekil 4.27'de sunulmuştur.

Autocorrelation Check of Residuals									
To Lag	Chi-Square	DF	Pr > ChiSq	Autocorrelations					
6	854.98	4	<.0001	-0.392	0.071	-0.005	0.006	0.013	-0.029
12	867.39	10	<.0001	-0.000	0.011	-0.029	0.031	0.007	-0.018
18	872.96	16	<.0001	0.004	0.011	-0.002	-0.026	0.013	0.006
24	885.86	22	<.0001	0.024	-0.034	0.010	-0.001	0.013	-0.020
30	889.90	28	<.0001	0.005	-0.002	0.012	-0.014	0.013	0.014
36	891.74	34	<.0001	-0.013	-0.005	0.001	-0.003	-0.002	0.011
42	900.71	40	<.0001	-0.011	0.014	-0.005	0.015	-0.026	0.021
48	905.87	46	<.0001	-0.018	0.011	-0.004	-0.003	0.011	-0.019

Şekil 4.27. Bayesian modelin kalıntılarına ait Box-Pierce Test sonuçları

Şekil 4.27'de özetlenen sonuçlarda tüm gecikmeler için P değerlerinin tamamı 0,05 den küçük olduğu için %95 güvenirlilikte kalıntı serisinde serisel korelasyon olmadığına karar verilmiştir.

Bu durumda $ARIMA(1,0,1)$ modeli kalıntıları temiz dizi özelliğindedir ve bu modelin uygun bir model olduğuna karar verilmiştir. Modelin açık halı;

$$y_t = 17,252 - 0,7354 \times 17,252 + 0,7354 \cdot y_{t-1} - 0,2147 \cdot \varepsilon_{t-1} + \varepsilon_t, \quad \varepsilon_t \sim N(0, \sigma^2)$$

şeklinde elde edilir.

4.6. Öngörü: Frekanscı Model ile Öngörü

Öngörü algoritması veri setinin analiz grubu verisi kullanılarak tahmin edilen zaman serisi modelleri ile veri setinin kontrol grubunda bulunan sırasıyla 2,3,4,5,6,7,8 kontrol günleri bilgisi kullanılarak sırasıyla geriye kalan 8,7,6,5,4,3,2 kontrol günleri tahmin edilmiştir. Öngörü algoritmasında tahmin için kullanılan gözlem numaraları ve tahmini yapılan gözlem numaraları Çizelge 4.5’de gösterilmiştir.

Çizelge 4.5. Öngörü algoritmasında kullanılan gözlem numaraları

Tahmin İçin Kullanılan Gözlem Aralığı	Tahmin Edilen Gözlem Aralığı
1-12	13-20
1-22	23-30
⋮	⋮
1-5332	5333-5340
1-5342	5343-5350
⋮	⋮
1-13	14-20
1-23	24-30
⋮	⋮
1-5333	5334-5340
1-5343	5344-5350
⋮	⋮
⋮	⋮
1-18	14-20
1-28	24-30
⋮	⋮
1-5338	5339-5340
1-5348	5349-5350

Öncelikle 2 adet kontrol günü kaydı kullanılarak $ARIMA(1,0,1)$ modeli;

$$y_t = 17,2568 - 0,7355 \times 17,2568 + 0,7355 \cdot y_{t-1} - 0,2153 \cdot \varepsilon_{t-1} + \varepsilon_t, \quad \varepsilon_t \sim N(0, \sigma^2)$$

kullanılarak geriye kalan 8 kontrol günü tahmin edilmiştir. Aynı işlem kontrol günü adedi 1’er adet artırılarak gerçek kontrol günlerine ait bilgiler kullanılarak geriye kalan kontrol günleri tahmin edilmiştir. Sonrasında tahmin edilen her kontrol gününe ait veri setindeki kontrol grubunda bulunan gerçek değerler ile korelasyon katsayıları hesaplanarak korelasyon katsayıları matrisi oluşturularak Çizelge 4.6’da sunulmuştur.

Çizelge 4.6. Mevcut kontrol günleri ile Frekanscı öngörü değerleri arasındaki korelasyonlar (i*: Bayesian yaklaşım modeli korelasyon katsayısından büyük katsayılar)

		Mevcut Kontrol Günü Sayısı						
		2	3	4	5	6	7	8
Öngörüsü yapılan Kontrol günü Numarası	3	0.9817						
	4	0.8352	0.9005*					
	5	0.7214	0.9928	0.7848*				
	6	0.6502	0.8635	0.9161*	0.6955*			
	7	0.5446	0.7675	0.9945	0.7124	0.5906*		
	8	0.4064*	0.6577	0.8754	0.9204*	0.7124*	0.4551*	
	9	0.3010	0.4922	0.7496	0.9950	0.8050*	0.5659*	0.3539*
	10	0.2738	0.3691	0.5432	0.8498	0.9036*	0.6208*	0.4388*

Çizelge 4.6'da ilk 4 kontrol gününün kullanıldığı durum için kalan kontrol günü tahminlerinin gerçek değerler ile ilişkileri yeterlidir ancak 10. kontrol günü tahminlerinin gerçek değerler ile ilişkisi azalarak 0,54 olarak hesaplanmıştır. Bu neden ile tercih edilmeyebilir.

Çizelge 4.6'da ilk 5 adet kontrol gününün kullanıldığı durum da tüm hayvanların 6. kontrol günlerinin gerçek süt verim değerleri tahmin değerleri ile ilişki katsayısı 0,69 olarak hesaplanmıştır. Yine 5 adet kontrol günü üretilen tahmin değerlerinin gerçek değerler ile ilişki katsayısı 7. kontrol gününde 0,71, 8. kontrol gününde 0,92, 9. kontrol gününde 0,99, 10. kontrol gününde 0,85 ilişki katsayısı hesaplanmıştır. İlk 5 kontrol günü bilgisinin kullanıldığı durumda geriye kalan kontrol günlerinin tahmin değerleri ilişkisi genel olarak yüksek olduğu için optimal kontrol günü sayısına 5 olarak karar verilebilmektedir.

4.7. Öngörü: Bayesian Model ile Öngörü

Öngörü algoritmasında tahmin için kullanılan gözlem numaraları ve tahmini yapılan gözlem numaraları Çizelge 4.5'de gösterilmiştir. Öncelikle 2 adet kontrol günü kaydı kullanılarak $ARIMA(1,0,1)$ modeli;

$$y_t = 17,2520 - 0,7354 \times 17,2520 + 0,7354 \cdot y_{t-1} - 0,2147 \cdot \varepsilon_{t-1} + \varepsilon_t, \quad \varepsilon_t \sim N(0, \sigma^2)$$

kullanılarak geriye kalan 8 kontrol günü tahmin edilmiştir. Aynı işlem kontrol günü adedi 1'er adet artırılarak gerçek kontrol günlerine ait bilgiler kullanılarak geriye kalan kontrol günleri tahmin edilmiştir. Sonrasında tahmin edilen her kontrol gününe ait veri setindeki kontrol grubunda bulunan gerçek değerler ile korelasyon katsayıları hesaplanarak korelasyon katsayıları matrisi Çizelge 4.7'da sunulmuştur.

Çizelge 4.7. Mevcut kontrol günleri ile Bayesian öngörü değerleri arasındaki korelasyonlar (i*: Frekanscı yaklaşım modeli korelasyon katsayısından büyük katsayılar)

		Mevcut Kontrol Günü Sayısı							
		2	3	4	5	6	7	8	
Öngörüsü yapılan Kontrol günü Numarası	3	0.9818*							
	4	0.8353*	0.9005						
	5	0.7215*	0.9929*	0.7846					
	6	0.6503*	0.8636*	0.9160	0.6964				
	7	0.5447*	0.7676*	0.9946*	0.8174*	0.5904			
	8	0.4064	0.6577*	0.8755*	0.9203	0.7123	0.4549		
	9	0.3011*	0.4922*	0.7496*	0.9951*	0.8049	0.5657	0.3537	
	10	0.2738*	0.3692*	0.5432*	0.8499*	0.9035	0.6205	0.4386	

Çizelge 4.7’de ilk 4 kontrol gününün kullanıldığı durum için kalan kontrol günü tahminlerinin gerçek değerler ile ilişkileri yeterlidir ancak 10. kontrol günü tahminlerinin gerçek değerler ile ilişkisi azalarak 0,54 olarak hesaplanmıştır. Bu neden ile tercih edilmeyebilir.

Çizelge 4.7’de ilk 5 adet kontrol günün kullanıldığı durum da tüm hayvanların 6. kontrol günlerinin gerçek süt verim değerleri tahmin değerleri ile ilişki katsayısı 0,70 olarak elde edilmiştir. Yine 5 adet kontrol günü üretilen tahmin değerlerinin gerçek değerler ile ilişki katsayısı 7. kontrol gününde 0,82, 8. kontrol gününde 0,92, 9. kontrol gününde 0,99, 10. kontrol gününde 0,85 ilişki katsayısı hesaplanmıştır. İlk 5 kontrol günü bilgisinin kullanıldığı durumda geriye kalan kontrol günlerinin tahmin değerleri ilişkisi genel olarak yüksek olduğu için optimal kontrol günü sayısına 5 olarak karar verilebilmektedir.

5. TARTIŞMA

Laktasyon dönemini kontrol günleri süt verim kayıtları ile modelleyen ve bu çalışmadaki öngörü algoritması ile aynı olan çalışmalardan bir tanesi Macciotta vd (2002) çalışmasıdır. Çalışma Sarda koyunları verileri kullanılarak gerçekleştirilmiştir. Çalışma sonucunda yapılan öngörü değerleri ile gerçek değerler arasındaki korelasyon katsayıları tablosunun orijinalinden alıntısı Şekil 5.1’de sunulmuştur.

Test day	Plan				
	1	2	3	4	5
	Milk				
3	0.85				
4	0.77	0.85			
5	0.72	0.80	0.86		
6	0.63	0.72	0.78	0.85	
7	0.55	0.63	0.69	0.78	0.85
8	0.44	0.50	0.56	0.64	0.70

Şekil 5.1. Macciotta vd. (2002) çalışmasında kontrol günleri süt miktarlarının öngörüsü için kullanılan zaman serisi modeli tahminlerinin gerçek değerleri ile ilişki matrisi

Macciotta vd (2002) çalışmasında 8 kontrol günü bulunan koyunlar için mevsimsel ARIMA modeli ile kontrol günleri için süt verimi tahminleri gerçekleştirilmiştir. Beş durum için model başarısı korelasyon katsayıları ile değerlendirilmiştir. Birinci durum; ilk 2 kontrol günü bilgisi kullanılarak geriye kalan 6 kontrol günü süt verimi tahmin edilmiştir ve maksimum 0,85 ilişki, minimum 0,44 ilişki katsayısı hesaplanmıştır. İkinci durum; ilk 3 kontrol günü bilgisi kullanılarak geriye kalan 5 kontrol günü süt verimi tahmin edilmiştir ve maksimum 0,85 ilişki, minimum 0,50 ilişki katsayısı hesaplanmıştır. Üçüncü durum; ilk 4 kontrol günü bilgisi kullanılarak geriye kalan 4 kontrol günü süt verimi tahmin edilmiştir ve maksimum 0,86 ilişki, minimum 0,56 ilişki katsayısı hesaplanmıştır. Dördüncü durum; ilk 5 kontrol günü bilgisi kullanılarak geriye kalan 3 kontrol günü süt verimi tahmin edilmiştir ve maksimum 0,85 ilişki, minimum 0,64 ilişki katsayısı hesaplanmıştır. Beşinci durum; ilk 6 kontrol günü bilgisi kullanılarak geriye kalan 2 kontrol günü süt verimi tahmin edilmiştir ve maksimum 0,85 ilişki, minimum 0,70 ilişki katsayısı hesaplanmıştır.

Çalışmaya ait bulgular bilinen kontrol günü sayısı arttıkça artmış ve ilerleyen kontrol günlerinde ilişki katsayıları azalma eğilimi göstermiştir. Macciotta vd. (2002) çalışması ile bu çalışmanın farklı özelliklerinden ilki çalışmada kullanılan hayvan türlerindeki farklılıktır. Ayrıca bu çalışmada kontrol günleri aralıkları 30’ar gün iken Macciotta vd (2002) çalışmasında kullanılan kontrol günleri 42 gün aralıklarla elde edilmiştir.

Laktasyon dönemini kontrol günleri süt verim kayıtları ile modelleyen ve öngörü algoritması bu çalışma ile benzer olan çalışmalardan bir tanesi Karaman (2010) çalışmasıdır. Çalışmada bu çalışma ile benzer süt sığırlarına ait veri seti

kullanılmıştır. Çalışma sonucunda yapılan öngörü değerleri ile gerçek değerler arasındaki korelasyon katsayıları tablosunun orijinalinden alıntısı Şekil 5.2’de sunulmuştur.

	Mevcut Gözlem Sayısı							
	2	3	4	5	6	7	8	
Öngörüsü	0,74							
Yapılan	0,74	0,82						
Kontrol	0,74	0,79	0,81					
Günü	0,75	0,78	0,80	0,85				
	0,76	0,78	0,80	0,85	0,87			
	0,76	0,78	0,79	0,83	0,87	0,87		
	0,75	0,77	0,78	0,81	0,83	0,84	0,87	
	0,74	0,75	0,76	0,79	0,80	0,80	0,80	0,80

Şekil 5.2. Karaman (2010) çalışmasında kontrol günleri süt miktarlarının öngörüsü için kullanılan zaman serisi modeli tahminlerinin gerçek değerleri ile ilişkisi matrisi

Karaman (2010) çalışmasında yaklaşık 30 günde bir olmak üzere toplam 10’ar kontrol günü bulunan süt sığırları için mevsimsel ARIMA modeli ile süt verimi tahminleri gerçekleştirmiştir. Yedi farklı durum için model başarısı korelasyon katsayıları ile değerlendirilmiştir. İlk sütunda; ilk 2 kontrol günü bilgisi kullanılarak geriye kalan 8 kontrol günü süt verimi tahmin edilmiştir ve maksimum 0,76 ilişki, minimum 0,74 ilişki katsayısı hesaplanmıştır. İkinci sütunda; 3 kontrol günü bilgisi kullanılarak geriye kalan 7 kontrol günü süt verimi tahmin edilmiştir ve maksimum 0,82 ilişki, minimum 0,75 ilişki katsayısı hesaplanmıştır. Üçüncü sütunda; ilk 4 kontrol günü bilgisi kullanılarak geriye kalan 6 kontrol günü süt verimi tahmin edilmiştir ve maksimum 0,81 ilişki, minimum 0,76 ilişki katsayısı hesaplanmıştır. Dördüncü sütunda; ilk 5 kontrol günü bilgisi kullanılarak geriye kalan 5 kontrol günü süt verimi tahmin edilmiştir ve maksimum 0,85 ilişki, minimum 0,79 ilişki katsayısı hesaplanmıştır. Beşinci sütunda; ilk 6 kontrol günü bilgisi kullanılarak geriye kalan 4 kontrol günü süt verimi tahmin edilmiştir ve maksimum 0,87 ilişki, minimum 0,80 ilişki katsayısı hesaplanmıştır. Altıncı sütunda; ilk 7 kontrol günü bilgisi kullanılarak geriye kalan 3 kontrol günü süt verimi tahmin edilmiştir ve maksimum 0,87 ilişki, minimum 0,80 ilişki katsayısı hesaplanmıştır. Yedinci sütunda; ilk 8 kontrol günü bilgisi kullanılarak geriye kalan 2 kontrol günü süt verimi tahmin edilmiştir ve maksimum 0,87 ilişki, minimum 0,80 ilişki katsayısı hesaplanmıştır.

Karaman (2010) çalışması ile bu çalışmanın korelasyon katsayıları değişimlerindeki farklılığın sebebi veri seçim yöntemlerindeki farklılıklar ve Karaman (2010) çalışmasında kullanılan veri seti için logaritmik dönüşüm uygulanmış veri setinin kullanılması ve serideki mevsimselliğin stokastik yorumlanması ile seriye mevsimsel fark uygulanmış olmasından kaynaklanıyor olabilir. Ayrıca Karaman (2010) çalışmasında yine Macciotta vd (2002) çalışmasında olduğu gibi en iyi modelin bu çalışmadan farklı olarak mevsimsel ARIMA modeline karar verilmiştir.

Karaman (2010) alıřmasında ve bu alıřmadaki veri setinde kayıtları bulunan st sıęırlarının kaıncı laktasyon dneminde oldukları bilinmemektedir. Laktasyon dneminin kontrol gn kayıtları ile modellenmesi yapılacak alıřmalarda bu bilgi ile aynı laktasyon dneminde olan hayvan materyali kullanılarak veri setinde homojenlik saęlanabilir ve daha isabetli ngrlerin elde edilebileceęi bir modelleme alıřması gerekleřtirilebilir.

6. SONUÇLAR

Zaman serisi modeli parametreleri öncelikle Frekanscı yaklaşım ile tahmin edilmiştir. Model;

$$y_t = 17,2568 + 0,7355.y_{t-1} - 0,2153.\varepsilon_{t-1} + \varepsilon_t, \quad \varepsilon_t \sim N(0, \sigma^2)$$

olarak hesaplanmıştır. Sonrasında zaman serisi model parametreleri Bayesian yaklaşım ile tahmin edilmiştir. Model;

$$y_t = 17,2514 + 0,7358.y_{t-1} - 0,2149.\varepsilon_{t-1} + \varepsilon_t, \quad \varepsilon_t \sim N(0, \sigma^2)$$

olarak hesaplanmıştır. İki model için Çizelge 4.5' de öngörü algoritması çerçevesinde tahmin değerleri ile gerçek değerler karşılaştırılarak korelasyon katsayıları hesaplanmıştır. Çalışma sonucunda Frekanscı ve Bayesian yaklaşım ile elde edilen korelasyon matrislerinin karşılaştırılmasında hesaplanan korelasyon katsayılarının %57'sinde Bayesian zaman serisi modeli daha başarılı olmuştur. Az bilgi kullanıldığı durumlar için Bayesian Yaklaşım ile elde edilen öngörü değerlerinin gerçek değerler ile daha yüksek korelasyon katsayılarına sahip olduğu gözlenmiştir. Bayesian yaklaşım ile daha az bilgi kullanıldığı durumların Frekanscı yaklaşıma göre daha isabetli tahminlerde bulunması literatür ile örtüşen bir sonuçtur.

Bayesian zaman serisi modeli ile 5 kontrol günü kullanılan durumda, gerçek değerlerle ilişki katsayılarının tamamı 0,70'den fazla olduğu için optimal kontrol sayısına 5 olarak karar verilmiştir.

Kontrol günleri kullanılarak yapılan laktasyon dönemi tahminlerinde Bayesian zaman serisi modelleri kullanılması için modelleme çalışmasının sürü bazında ve aynı laktasyon döneminde olan hayvanlar üzerinde yapılması öngörü başarısını artıracaktır. Böylece benzer fenotipte sürüler için başarılı bir öngörü yöntemi olarak Bayesian zaman serisi model tahmini önerilebilir.

7. KAYNAKLAR

- Anonim1: <http://www.tuik.gov.tr/PreTabloArama.do?metod=search&araType=vt> [Son erişim tarihi: 08.12.2019]
- Anonim2: <https://www.tarimbilgisi.com/haber/hayvanlar/buzagi-nedir-dana-nedir-duve-nedir-sigir-nedir-boga-nedir//> [Son erişim tarihi: 08.05.2020]
- Akgül, I. 2000. Geleneksel Zaman Serisi Yöntemleri. Der Yayınları.
- Akgül, I. 2003. Zaman Serilerinin Analizi ve ARIMA Modelleri. İstanbul.
- Aydemir, C. ve & Pıçak, M. 2007. "GAP Bölgesi'nde Hayvancılığın Gelisimi ve Türkiye İçindeki Konumu". Elektronik Sosyal Bilimler Dergisi, 22: 13–37.
- Berberoğlu, E. 2010. Süt Sığırlarında Zaman Serisi Yöntemleriyle Laktasyon Eğrilerinin Modellenmesi. Ege Üniversitesi, Fen Bilimleri Enstitüsü, Doktora Tezi.
- Çağıl, G. 2017. "Mevsimlik Olmayan Box-Jenkins Modellerinde İki Aşamalı Yapay Sinir Ağlarının Kullanılması". Academic Platform-Journal of Engineering and Science, 3: 119–126.
- Eğrioğlu, E. 2002. ARMA Modellerinin Bayes Analizi ve Bir Uygulama. Ondokuz Mayıs Üniversitesi, Fen Bilimleri Enstitüsü, Yüksek Lisans Tezi.
- Fan, C. ve & Yao, S. 2008. "Bayesian Approach for ARMA Process and Its Application", 1(October): 49–55.
- Ghosh, B. ve Basu, B. ve & O'Mahony, M. 2007. "A Bayesian Time-Series Model for Short-Term Traffic Flow Forecasting". Journal of Transportation Engineering, 133(3): 1–37.
- Hoffman, M. D. ve & Gelman, A. 2014. "The No-U-Turn Sampler: Adaptively Setting Path Lengths in Hamiltonian Monte Carlo". Journal of Machine Learning Research, 15: 1593–1623.
- Kadılar, C. 2005. SPSS Uygulamalı Zaman Serileri Analizine Giriş. Ankara: Bizim Büro Yayınevi.
- Karaman, E. 2010. Kontrol Günü Süt Verimlerinin Zaman Serisi Yöntemi ile Modellenmesi. Akdeniz Üniversitesi, Fen Bilimleri Enstitüsü, Yüksek Lisans Tezi.
- Kaya, E. 2019. Zaman Serileri Analizinde Box-Jenkins Yöntemi ile Savunma Sanayi Verileri Üzerine Bir Uygulama. Karamanoğlu Mehmetbey Üniversitesi, Sosyal Bilimleri Enstitüsü, Yüksek Lisans Tezi.
- Koptur, M. 2020a. "Bayesci İstatistiğe Giriş Bayesci İstatistiğe Giriş", (February).
- Koptur, M. 2020b. Doğa Sporlarının Yaşam Doyumu Üzerine Etkisi: Bayesci Yaklaşım. Akdeniz Üniversitesi, Sosyal Bilimler Enstitüsü, Yüksek Lisans Dönem

Projesi.

- Macciotta, N. P. P. P. ve Cappio Borlino, A. ve & Pulina, G. 2000. "Time Series Autoregressive Integrated Moving Average Modeling of Test-Day Milk Yields of Dairy Ewes". *Journal of Dairy Science*, 83(5): 1094–1103.
- Macciotta, N. P. P. ve & Pulina, G. 2000. "Time Series Autoregressive Integrated Moving Average Modeling of Test-Day Milk Yields of Dairy Ewes". *Journal of Dairy Science*, 83(5): 1094–1103.
- Macciotta, N. P. P. ve Vicario, D. ve Pulina, G. ve & Cappio-Borlino, A. 2002. "Test Day and Lactation Yield Predictions in Italian Simmental Cows by ARMA Methods". *Journal of Dairy Science*, 85(11): 3107–3114.
- Mathy, F. ve & Chekaf, M. 2018. *Introduction to Bayesian Analysis. Experiments and Modeling in Cognitive Science*. <https://doi.org/10.1016/b978-1-78548-284-7.50010-1>
- Mert, M. ve & Çağlar, A. E. 2019. *Eviews ve Gauss Uygulamalı Zaman Serileri Analizi*.
- Orhan, H. ve & Kaygısız, A. 2002. "Siyah Alaca Sığırlarda Farklı Laktasyon Eğrisi Modellerinin Karşılaştırılması", 43(1): 94–99.
- Orloff, J. ve & Bloom, J. 2014. *Comparison of frequentist and Bayesian inference . Jeremy Orloff and Jonathan Bloom Bayes ' formula as touchstone Critiques and defenses Critique of Bayesian inference*.
- Saçaklı Saçıldı, İ. 2011. *Gelişmiş ve Gelişmekte olan Piyasalarda Hisse Senedi Getiri Volatilitelerinin Bayes Yaklaşımı ile Analizi*. Marmara Üniversitesi, Sosyal Bilimler Enstitüsü, Doktora Tezi.
- Şengül, H. 1986. *Zaman Serileri Analizinde Box-Jenkins Modelleri*. İstanbul Üniversitesi, Sosyal Bilimleri Enstitüsü, Yüksek Lisans Tezi.
- Sönmez Oskay, G. 2016. *Siyah Alaca Süt Sığırlarında Laktasyon Biyometrisi Üzerine Bir Araştırma*. IOSR Journal of Economics and Finance. Namık Kemal Üniversitesi, Fen Bilimleri Enstitüsü, Yüksek Lisans Tezi. <https://doi.org/https://doi.org/10.3929/ethz-b-000238666>
- Teke, E. 2017. *Siyah Alaca İneklerin Laktasyon Eğrilerinin Modellenmesinde Zaman Serileri ve Non-Linear Yöntemlerin Karşılaştırması*. Süleyman Demirel Üniversitesi, Fen Bilimleri Enstitüsü, Doktora Tezi.
- Tripathi, P. K. ve Ranjan, R. ve Pant, R. ve & Upadhyay, S. K. 2017. "An Approximate Bayes Analysis of ARMA Model for Indian GDP Growth Rate data". *Journal of Statistics and Management Systems*, 20(3): 399–419.
- Ünal, E. 2003. *Bayesci Vektör Otoregrasyon Modeller ve Türkiye’de Enflasyon Üzerine bir Uygulama*. Mimar Sinan Üniversitesi, Fen Bilimleri Enstitüsü, Yüksek

Lisans Tezi.

Vural, H. ve & Fidan, H. 2007. "Türkiye’de Hayvansal Üretim ve Hayvancılık İşletmelerinin Özellikleri", 13(2): 49–59.

Wakefield, J. 2005. Springer series in statistics. Modern multidimensional scaling: Theory and applications (2nd ed.). New York, NY, US: Springer Science+ Business Media.

Wood, P. D. P. 1967. "Algebraic model of the lactation curve in cattle". Nature, 216(5111): 164–165.

ÖZGEÇMİŞ

ÖZGE KOZAKLI
ozgekozakli94@hotmail.com



ÖĞRENİM BİLGİLERİ

Yüksek Lisans 2018-2020	Akdeniz Üniversitesi Ziraat Fakültesi, Zootečni Bölümü, Antalya
Lisans 2015-2018	Marmara Üniversitesi (Yatay Geçiş) İktisat Fakültesi, Ekonometri Bölümü, İstanbul
Lisans 2012-2015	Akdeniz Üniversitesi İ.İ.B.B. Fakültesi, Ekonometri Bölümü, Antalya