

**T.C.**

**MUĞLA SITKI KOÇMAN ÜNİVERSİTESİ  
FEN BİLİMLERİ ENSTİTÜSÜ**

**İSTATİSTİK ANABİLİM DALI**

**HALK SAĞLIĞI İZLEMESİ İÇİN BAZI  
İSTATİSTİKSEL SALGIN BELİRLEME YÖNTEMLERİ**

**YÜKSEK LİSANS**

**ÜNAL CENGİZ**

**NİSAN 2018**

**MUĞLA**

**Ü. CENGİZ**

**İSTATİSTİK ANABİLİM DALI**

**YÜKSEK LİSANS TEZİ**

**MUĞLA 2018**

**T.C.**  
**MUĞLA SITKI KOÇMAN ÜNİVERSİTESİ**  
**FEN BİLİMLERİ ENSTİTÜSÜ**

**İSTATİSTİK ANABİLİM DALI**

**HALK SAĞLIĞI İZLEMESİ İÇİN BAZI**  
**İSTATİSTİKSEL SALGIN BELİRLEME YÖNTEMLERİ**

**YÜKSEK LİSANS**

**ÜNAL CENGİZ**

**NİSAN 2018**

**MUĞLA**

**MUGLA SITKI KOÇMAN ÜNİVERSİTESİ**  
**Fen Bilimleri Enstitüsü**

**TEZ ONAYI**

ÜNAL CENGİZ tarafından hazırlanan **HALK SAĞLIĞI İZLEMESİ İÇİN BAZI İSTATİSTİKSEL SALGIN BELİRLEME YÖNTEMLERİ** başlıklı tezinin, 06/04/2018 tarihinde aşağıdaki jüri tarafından İstatistik Anabilim Dalı'nda yüksek lisans derecesi için gerekli şartları sağladığı oybirliği/oyçokluğu ile kabul edilmiştir.

---

**TEZ SINAV JURİSİ**

Doç. Dr. Dursun AYDIN (**Jüri Başkanı**)

İstatistik Anabilim Dalı,  
Muğla Sıtkı Koçman Üniversitesi, Muğla

İmza:



Dr. Öğr. Üyesi Mehmet KARAHASAN(**Danışman**)

İstatistik Anabilim Dalı,  
Muğla Sıtkı Koçman Üniversitesi, Muğla

İmza:



Dr. Öğr. Üyesi Ali Zafer DALAR (**Üye**)

Giresun Üniversitesi, Giresun

İmza:



---

**ANA BİLİM DALI BAŞKANLIĞI ONAYI**

Doç. Dr. Dursun AYDIN

İstatistik Anabilim Dalı,  
Muğla Sıtkı Koçman Üniversitesi, Muğla

İmza:



Dr. Öğr. Üyesi Mehmet KARAHASAN

Danışman, İstatistik Anabilim Dalı,  
Muğla Sıtkı Koçman Üniversitesi, Muğla

İmza:



Savunma Tarihi: 06/04/2018

---

Tez çalışmalarım sırasında elde ettiğim ve sunduğum tüm sonuç, doküman, bilgi ve belgelerin tarafımdan bizzat ve bu tez çalışması kapsamında elde edildiğini; akademik ve bilimsel etik kurallarına uygun olduğunu beyan ederim. Ayrıca, akademik ve bilimsel etik kuralları gereği bu tez çalışması sırasında elde edilmemiş başkalarına ait tüm orijinal bilgi ve sonuçlara atıf yapıldığını da beyan ederim.

Ünal Cengiz

06/04/2018

**ÖZET**  
**HALK SAĞLIĞI İZLEMESİ İÇİN BAZI İSTATİSTİKSEL SALGIN**  
**BELİRLEME YÖNTEMLERİ**

Ünal CENGİZ

Yüksek Lisans Tezi

Fen Bilimleri Enstitüsü

İstatistik Anabilim Dalı

Danışman: Dr. Öğr. Üyesi Mehmet KARAHASAN

Nisan 2018, 189 sayfa

Birey ve toplumun sağlığını olumsuz etkileyebilecek olayların örneğin hastalık salgınlarının zamanında tespiti ve önlemlerin alınması halk sağlığında önemli bir amaçtır. Bu amaçla halk sağlığı izleme sistemleri oluşturulmuştur ve istatistiksel salgın belirleme algoritmaları bu sistemler içinde kullanılmaktadır. Halk sağlığı izleme sistemlerinde istatistiksel algoritmalarının uygulamasını güçleştiren bir husus hastalık salgınlarının izlenmesi için toplanan verilerin özellikleridir. Veriler genellikle günlük ya da haftalık olay sayılarını belirttikleri için kesikli, öz ilişkili olup, mevsimsellik etkisi gösterebilmektedir. Başka bir ifadeyle, veriler standart istatistiksel yöntemlerin gerektirdiği gibi genellikle normallik, bağımsızlık ve durağanlık özelliklerine sahip değildir.

Bu tez çalışması tek değişkenli zamana bağlı ileriye dönük izleme yapabilen bazı istatistiksel salgın belirleme yöntemlerini tanıtmayı, gerçek verilere uygulamayı ve simülasyon çalışmasıyla bu yöntemlerin performanslarını incelemeyi hedeflemiştir. Bu amaca yönelik olarak bu çalışma kapsamında regresyon modellemesine, istatistiksel süreç kontrolüne, hipotez testine, zaman serisi modellemesine dayanan bazı yöntemler dikkate alınmıştır. Ayrıca bu çalışma kapsamında zaman içerisinde ileriye dönük izleme yapabilen ve UCengiz yöntemi olarak adlandırılan basit bir istatistiksel yöntem önerilmiştir.

Gerçek veri analizi için, bu tez kapsamında önerilen UCengiz yöntemiyle birlikte Serfling regresyon yöntemi, Farrington ve FarringtonFlexible yöntemleri, OutbreakP, Bayes, Cusum, Ears C1, C2, C3, Rki, Haar-Wavelets, Glnb, Shewhart I-chart, EWMA gibi süreç kontrol grafikleri, Holt-winters, Hidden Markov modeli ve ARIMA modelleme yöntemleri dikkate alınmıştır. Bu yöntemlerin her biri CDC (Centers for Disease Control and Prevention) kurumundan elde edilen 2015-2016 haftalık grip mevsimi verilerini izlemek amacıyla uygulanmıştır. Analizler R programının

surveillance paketi, çeşitli sitelerden indirilen excell, MATLAB tabanlı programlarının düzenlenmesi ve R kodlarıyla yazılan programlar aracılığıyla gerçekleştirilmiştir. Bu analizler sonucunda Hidden Markov modeli, Serfling regresyon yöntemi, Cusum, UCengiz, Ewma artıklarını kullanan I-Chart, Bayes, Glrb ve Rki yöntemlerinin salgınları daha erken ve doğru olarak belirledikleri tespit edilmiştir. Ayrıca sadece salgının başlagıcını belirlenmesi amacıyla tasarlanmış OutbreakP yöntemi, salgını erken belirleyen yöntemler arasında yer almıştır.

Ayrıca CDC kurumunda elde edilen 2013-2014, 2014-2015, 2015-2016 ve 2016-2017 grip mevsimlerine ilişkin haftalık verilerin özelliklerini yansıtacak şekilde grip verilerinin 10000 tekrarlı simülasyonları yapılmıştır. Bu simülasyonlarda; aynı ATFS (Average Time Between False Signals) değerinde, ATFOS (Average Time to First Outbreak Signals) ve FMO (Fraction of Missed Outbreaks) ölçütlerine göre 11 algoritmanın performansları çeşitli salgın büyüklükleri ve süreleri dikkate alınarak karşılaştırılmıştır. Simülasyonlarda Bayes, Cusum, Farrington, FarringtonFlexible, Ears C1, Ears C2, Ears C3, OutbreakP, Glrb, Serfling regresyon modeli ve bu çalışmada önerilen UCengiz yöntemi dikkate alınmıştır. Tüm bu işlemler tez çalışması için özel olarak oluşturulan R kodlarıyla yazılan programlar ile gerçekleştirilmiştir. Simülasyonlar sonucunda; bu tez çalışmasında önerilen UCengiz yönteminin ATFOS açısından, Serfling regresyon yönteminin FMO açısından genel olarak en iyi performansa sahip oldukları söylenebilmektedir. Her iki ölçüt birlikte dikkate alındığında ise UCengiz, Serfling Regresyon, Cusum, Farrington ve FarringtonFlexible yöntemlerinin iyi performanslara sahip oldukları gözlemlenmiştir.

**Anahtar Kelimeler:** Biyoizleme, Halk Sağlığı İzleme Sistemleri, İleriye Dönük İstatistiksel Salgın Belirleme Algoritmaları, Erken Uyarı Algoritmaları, Grip Salgınlarının Belirlenmesi

**ABSTRACT**  
**SOME STATISTICAL OUTBREAK DETECTION METHODS FOR PUBLIC  
HEALTH SURVEILLANCE**

Ünal CENGİZ

Master of Science (M.Sc.)

Graduate School of Natural and Applied Sciences

Department of Statistics

Supervisor: Asst. Prof. Dr. Mehmet KARAHASAN

April 2018, 189 pages

Identification of events that can negatively affect the health of the individual and the community, such as the detection of disease epidemics early in time and taking measures against, is an important goal in public health. For this purpose public health monitoring systems have been established and statistical epidemic detection algorithms are used in these systems. One of the matters that makes the application of statistical algorithms difficult for public health monitoring systems is inherent features of the data collected for disease outbreak surveillance. Since the data are generally daily or weekly numbers of event of interest, they are often discrete, autocorrelated and have seasonal effects. In other words, the data usually are not distributed normal, independent and stationary as required by standard statistical methods.

The aim of this thesis is to introduce some univariate temporal statistical epidemic detection methods that make real time surveillance, to apply them to real data and to study their performance with simulation studies. For this purpose, some methods based on regression modeling, statistical process control, hypothesis testing, time series modeling are taken into consideration. In addition, a simple statistical method, called UCengiz method, is proposed, which can monitor prospectively the number of events of interest over time.

To make real data analysis, Serfling regression method, Farrington and FarringtonFlexible methods, OutbreakP, Bayes, Cusum, Ears C1, C2, C3, Rki, Haar Wavelets, Glrnb, Hidden Markov model, ARIMA modeling and process control graphics such as Shewhart I-Chart, EWMA are taken into consideration together with the newly proposed UCengiz method in this study. Each of these methods has been applied to monitor 2015-2016 weekly flu seasonal data obtained from Centers for Disease Control and Prevention (CDC). Analyzes were made by using surveillance package of R program, excell macros downloaded from various sites, modification of some MATLAB based programs and the programs written in R codes. As a result of

these analyzes, Hidden Markov model, Serfling regression method, Cusum, UCengiz, I-Chart based on Ewma residuals, Bayes Glnb and Rki methods were found to identify outbreaks earlier and more accurately. Furthermore, the OutbreakP method, which was designed to determine the onset of the disease epidemic, was also one of the early methods to identify the outbreak.

In addition, 10000 repetitive simulations of influenza data were conducted to reflect the characteristics of the weekly influenza data on the influenza seasons 2013-2014, 2014-2015, 2015-2016 and 2016-2017 obtained at CDC. In these simulations; performance of 11 algorithms based on the same ATFS (Average Time Between False Signals), were compared for various outbreak sizes and durations in terms of ATFOS (Average Time to First Outbreak Signals) and FMO (Fraction of Missed Outbreaks) criteria. In the simulations Bayes, Cusum, Farrington, Farrington Flexible, Ears C1, Ears C2, Ears C3, OutbreakP, Glnb, Serfling regression model and UCengiz method proposed in this study are considered. All these simulations were performed with programs written in R codes which are specially developed for the thesis. As a result of the simulations; it can be said that the performance of the UCengiz method is generally the best in terms of ATFOS while the performance of the Serfling regression method is generally the best in terms of FMO. Furthermore, it is observed that UCengiz Method, Serfling regression, Cusum, Farrington and FarringtonFlexible methods have good performances according to both ATFOS and FMO criteria.

**Keywords:** Biosurveillance, Public Health Surveillance Systems, Prospective Statistical Outbreak Detection Algorithms, Early Event Detection, Influenza Outbreaks Detection



## ÖNSÖZ

Bu tez çalışmasının her aşamasında, öneri ve tavsiyeleriyle yol gösteren tez danışmanım Dr. Öğr. Üyesi Mehmet KARAHASAN'a ve manevi destekleriyle beni hiçbir zaman yalnız bırakmayan aileme teşekkür ederim.



## İÇİNDEKİLER

ÖNSÖZ.....	viii
İÇİNDEKİLER .....	ix
ÇİZELGELER DİZİNİ .....	xiii
ŞEKİLLER DİZİNİ .....	xvii
SEMBOLLER VE KISALTMALAR DİZİNİ .....	xx
<b>1. GİRİŞ .....</b>	<b>1</b>
1.1. Halk Sağlığı İzlemesi .....	1
1.2. Halk Sağlığı İzleme Sistemleri .....	3
1.3. Halk Sağlığı İzlemesi Verilerin Özellikleri .....	4
1.4. Amaç ve Kapsam .....	4
<b>2. BİYOİZLEME ALGORİTMALARININ GENEL YAPISI .....</b>	<b>5</b>
2.1. Beklenen Veri Davranışının Belirlenmesi Amacıyla Oranlama İşlemi .....	5
2.2. Alarm Algoritmalarının Bileşenleri ve Algoritmik İşleyiş Adımları .....	6
2.3. İzleme Verilerindeki Değişimin Kaynakları .....	8
2.4. Salgın Sinyalinin Yapısı .....	9
2.5. Salgınların Erken Belirlenmesinde Karşılaşılan İstatistiksel Zorluklar .....	9
2.6. Algoritmaların veya Verilerin Biyoizlemeye Uyarlanma Stratejileri .....	12
2.7. Salgın Belirleme Yöntemlerinin Sınıflandırılması .....	13
2.7.1 Bazı biyoizleme sistemlerinde kullanılan mevcut algoritmalar.....	14
2.7.2 Bioizleme verilerindeki sistematik yapıyı dikkate almak amacıyla kullanılan ana yaklaşımlar.....	15
2.8. Algoritmaların Performanslarını Değerlendirme Ölçütleri .....	16
2.9. Algoritmaların Değerlendirilmesinde Kullanılan Veriler .....	20
2.9.1 Yapay salgın üretme yaklaşımları .....	21
2.9.2 Yapay bioizleme verilerinin üretilmesi .....	23
<b>3. BAZI BİYOİZLEME YÖNTEMLERİ.....</b>	<b>24</b>
3.1. Regresyon Modellemesine Dayanan Yaklaşımlar.....	24
3.1.1 Serfling parametrik regresyon modeli .....	25

3.1.2 Farrington ve FarringtonFlexible Yöntemi.....	26
3.1.3 Bayesci Salgın Belirleme Yöntemi (Boda).....	29
3.1.4 Bayeşçi Bir Öngörü Sonsal Yaklaşımı .....	32
3.1.5 Yarı parametrik regresyon modelleri.....	33
3.1.6 Haar-wavelet Dönüşümler Yöntemi .....	34
3.2. Zaman Serisi Modellemesine Dayanan Yaklaşımlar .....	37
3.2.1 Holt-Winters düzleştirme yöntemi .....	38
3.2.2 Hidden Markov Modeli Yöntemi .....	39
3.3. İstatistiksel Süreç Kontrol Grafiklerine Dayanan Yaklaşımlar .....	41
3.3.1 Cusum kontrol grafiğı.....	42
3.3.2 Bireysel ölçümler için Shewhart kontrol grafiğı .....	43
3.3.3 Halk sağığı izlemesinde kullanılan süreç kontrol grafikleri.....	44
3.3.4 Ears C1, C2 ve C3 Yöntemi .....	45
3.3.5 Olaylar arasında geçen süreyi temel alan kontrol grafikleri.....	47
3.3.6 Tarama İstatistiğıne Dayanan Yöntemler .....	48
3.4 Hipotez Testlerine Dayanan Biyoizleme Yaklaşımları .....	48
3.4.1 Tam olabilirlik oranı yaklaşımı .....	50
3.4.2 Shiryayev-Roberts olabilirlik oran yaklaşımı.....	51
3.4.3 Shewhart yaklaşımı.....	52
3.4.4 Parametrik olabilirlik oran yaklaşımı .....	52
3.4.5 Parametrik olmayan ya da yarı parametrik olabilirlik oran yaklaşımı .....	53
3.4.6 GLR istatistiklerine dayalı regresyon kontrol grafikleri.....	55
<b>4. MATERYAL VE YÖNTEM.....</b>	<b>58</b>
4.1 Verilerin Özellikleri.....	58
4.2 Verilerde İzleme Amaçlı Kullanılacak İleriye Dönük İzleme Yöntemleri .....	60
4.3 Simülasyonlarda Performansları Karşılaştırılacak İleriye Dönük İzleme Yöntemleri .....	61
4.4 Karşılaştırma Ölçütleri .....	62
4.5 Tez Çalışması Kapsamında Önerilen UCengiz Yöntemi .....	63
<b>5. UYGULAMA.....</b>	<b>66</b>
5.1 Regresyon Türü Yöntemlerin CDC Grip Verilerine Uygulanması .....	66

5.1.1 Serfling regresyon yönteminin CDC grip verilerine uygulanması .....	66
5.1.2 Farrington ve FarringtonFlexible yönteminin CDC grip verilerine uygulanması .....	68
5.1.3 Boda yönteminin CDC grip verilerine uygulanması .....	71
5.2 Zaman Serisi Modellemesi Yaklaşımlarının CDC Grip Verilerine Uygulanması .....	73
5.2.1 Holt-Winters yaklaşımının CDC grip verilerine uygulanması .....	74
5.2.2 Arima modelleme yaklaşımının CDC grip verilerine uygulanması .....	78
5.2.3 Hidden Markov modeli yaklaşımının CDC grip verilerine uygulanması..	83
5.3 Kalite Kontrol Grafiklerinin CDC Grip Verilerine Uygulanması .....	85
5.3.1 Holt-Winters artıklarına Shewhart I-Chart, Ewma, Cusum grafiklerinin uygulanması .....	85
5.3.2 Mevsimselliği giderilmiş ham verilere Shewhart I-Chart grafiğinin uygulanması .....	90
5.3.3 CDC grip verilerinin Ewma yöntemiyle elde edilmiş artıklara Shewhart I-Chart grafiğinin uygulanması.....	92
5.3.4 Rossi vd. (1999) Cusum yönteminin CDC grip verilerine uygulanması...	94
5.3.5 Ears C1, C2 ve C3 Yöntemlerinin CDC Grip Verilerine Uygulanması ....	97
5.4 Hipotez Testlerine Dayanan Yaklaşımlarının CDC Grip Verilerine Uygulanması .....	99
5.4.1 OutbreakP yönteminin CDC grip verilerine uygulanması.....	99
5.4.2 Genelleştirilmiş olabilirlik oranı regresyon grafiği .....	106
5.5 Diğer Yaklaşımların CDC Grip Verilerine Uygulanması .....	109
5.5.1 Bayesçi bir sonsal öngörü yaklaşımının CDC grip verilerine uygulanması .....	109
5.5.2 Rki yaklaşımının CDC grip verilerine uygulanması .....	111
5.5.3 Wavelet yaklaşımının CDC grip verilerine uygulanması .....	113
5.6 Önerilen UCengiz Yönteminin CDC Grip Verilerine Uygulanması.....	117
5.6.1 ATFS 13 için elde edilen izleme sonuçları.....	117
5.6.2 ATFS 26 için elde edilen izleme sonuçları.....	120
<b>6. SİMÜLASYON ÇALIŞMASI .....</b>	<b>123</b>
6.1 Grip Verilerin Üretimi.....	123
6.2 Karşılaştırma Ölçütleri .....	125

6.3 İzleme Dönemine Salgınların Eklenmesi .....	128
6.4 Simülasyon Planı .....	129
6.5 Simülasyon Sonuçları .....	130
6.5.1 Dört hafta süreli salgınlar için simülasyon sonuçları .....	130
6.5.2 Sekiz hafta süreli salgınlar için simülasyon sonuçları .....	134
6.5.3 On iki hafta süreli salgınlar için simülasyon sonuçları .....	138
<b>7. SONUÇLAR VE TARTIŞMA .....</b>	<b>143</b>
<b>KAYNAKLAR .....</b>	<b>148</b>
<b>EKLER .....</b>	<b>154</b>
Ek A. 2012-2016 Yıllarına İlişkin CDC Grip Verileri .....	154
Ek B. CDC Grip Verileri İçin Serfling Regresyon Tahminleri ve Eşik Değerleri .....	155
Ek C. CDC Grip Verileri İçin Farrington ve FarringtonFlexible Yöntemlerine İlişkin Sonuçlar .....	156
Ek D. CDC Grip Verileri İçin Boda Yöntemlerine İlişkin Sonuçlar .....	157
Ek E. CDC Grip Verileri İçin Holt-Winters Yöntemine İlişkin Sonuçlar .....	158
Ek F. CDC Grip Verileri İçin Arima Modellemesine İlişkin Sonuçlar .....	159
Ek G. 2012-2016 Yıllarına İlişkin Serfling Regresyon Artıkları .....	160
Ek H. CDC Grip Verileri İçin Holt-Winters Düzleştirme Yöntemine İlişkin Tahmin Hataları .....	161
Ek I. 2012-2013, 2013-2014 ve 2014-2015 Yıllarına İlişkin Mevsimsellik Düzeltmesi İndeksleri ve Mevsimsellik Düzeltmesi Yapılmış CDC grip verileri .....	162
Ek İ. Ewma Tahmin ve Artıklarına İlişkin Sonuçlar .....	164
Ek J. CDC Grip Verileri İçin Ears C1, C2 Ve C3 Yöntemlerine İlişkin Sonuçlar .....	166
Ek K. Serfling Modeli Tahmin ve Artıkları .....	167
Ek L. Salgın Belirleme Algoritmalarına İlişkin R Kodları .....	168
Ek M. 2015-2016 Grip Mevsimi Verileri İçin UCengiz Yöntemine İlişkin Sonuçlar .....	184
<b>ÖZGEÇMİŞ .....</b>	<b>188</b>

## ÇİZELGELER DİZİNİ

Çizelge 3.1. Farrington ve FarringtonFlexible yöntemlerine ilişkin genelleştirilmiş doğrusal modeller .....	27
Çizelge 3.2. Boda yöntemindeki model kısımlarına ilişkin önsel dağılımlar .....	30
Çizelge 4.1. 2012-2016 grip mevsimlerine ilişkin ADF ve PP durağanlık testleri sonuçları.....	60
Çizelge 4.2. CDC kurumunun 2015-2016 grip mevsimi için haftalık grip faaliyet durumu.....	62
Çizelge 4.3. Yöntemlerin verdiği kararlar için sınıflandırma tablosu.....	63
Çizelge 5.1. Serfling regresyona ilişkin 2015-2016 grip mevsimi için haftalık grip faaliyet durumu.....	68
Çizelge 5.2. Serfling yönteminin verdiği kararlar için sınıflandırma tablosu.....	68
Çizelge 5.3. Farrington ve FarringtonFlexible yöntemlerine ilişkin 2015-2016 grip mevsimi için haftalık grip faaliyet durumu .....	69
Çizelge 5.4. Farrington yöntemlerinin verdiği kararlar için sınıflandırma tablosu....	70
Çizelge 5.5. Boda yöntemlerine ilişkin 2015-2016 grip mevsimi için haftalık grip faaliyet durumu.....	72
Çizelge 5.6. Boda yöntemlerinin verdiği kararlar için sınıflandırma tablosu .....	73
Çizelge 5.7. Aykırı değerler yerine tsoutliers fonksiyonu tarafından belirlenen değerler .....	73
Çizelge 5.8. CDC verileri için Holt-Winters düzeltirme yöntemine ilişkin ölçütler	75
Çizelge 5.9. Holt-Winters yöntemine ilişkin 2015-2016 grip dönemi için haftalık grip faaliyet durumu.....	77
Çizelge 5.10. Holt Winters yönteminin verdiği kararlar için sınıflandırma tablosu ..	77
Çizelge 5.11. Birinci farkı ve mevsimsel farkı alınmış veriler için aday modellerin özellikleri .....	80
Çizelge 5.12. Seçilen modele ilişkin parametre kestirimleri.....	82
Çizelge 5.13. Arima modelleme yöntemine ilişkin 2015-2016 grip mevsimi için haftalık grip faaliyet durumu .....	83
Çizelge 5.14. Arima yönteminin verdiği kararlar için sınıflandırma tablosu .....	83

Çizelge 5.15. Hidden Markov modeli yöntemlerine ilişkin 2015-2016 grip mevsimi için haftalık grip faaliyet durumu .....	85
Çizelge 5.16. Hidden Markov modeli yöntemlerinin verdiği kararlar için sınıflandırma tablosu .....	85
Çizelge 5.17. CDC verileri için Holt-Winters düzeltirme yöntemine ilişkin ölçütler .....	86
Çizelge 5.18. I-Chart, Ewma ve Cusum grafiklerine ilişkin 2015-2016 grip mevsimi için haftalık grip faaliyet durumu .....	88
Çizelge 5.19. Kalite kontrol grafikleri yöntemlerinin verdiği kararlar için sınıflandırma tablosu .....	88
Çizelge 5.20. I-Chart grafiğine ilişkin 2015-2016 grip mevsimi için haftalık grip faaliyet durumu.....	92
Çizelge 5.21. Kalite kontrol grafikleri yöntemlerinin verdiği kararlar için sınıflandırma tablosu .....	92
Çizelge 5.22. I-Chart grafiğine ilişkin 2015-2016 grip mevsimi için haftalık grip faaliyet durumu.....	94
Çizelge 5.23. Kalite kontrol grafikleri yöntemlerinin verdiği kararlar için sınıflandırma tablosu .....	94
Çizelge 5.24. Rossi vd. (1999) Cusum yönteminin verdiği kararlar için sınıflandırma tablosu.....	96
Çizelge 5.25. Rossi vd. (1999) Cusum yöntemine ilişkin 2015-2016 grip mevsimi için haftalık grip faaliyet durumu .....	96
Çizelge 5.26. Ears C1, C2 ve C3 yöntemlerine ilişkin 2015-2016 grip dönemi için haftalık grip faaliyet durumu .....	97
Çizelge 5.27. Ears yöntemlerinin verdiği kararlar için sınıflandırma tablosu .....	99
Çizelge 5.28. İzleme zamanlarının temsil ettiği haftalar.....	103
Çizelge 5.29. Referans dönemi verileri için Poisson regresyon model kestirimleri	107
Çizelge 5.30. Poisson regresyon kontrol grafiği yöntemine ilişkin 2015-2016 grip mevsimi için haftalık grip faaliyet durumu .....	108
Çizelge 5.31. Poisson regresyon kontrol grafiği yönteminin verdiği kararlar için sınıflandırma tablosu .....	109
Çizelge 5.32. Bayes yöntemine ilişkin 2015-2016 grip mevsimi için haftalık grip faaliyet durumu.....	110
Çizelge 5.33. Bayes yönteminin verdiği kararlar için sınıflandırma tablosu.....	110
Çizelge 5.34. Rki yöntemine ilişkin 2015-2016 grip mevsimi için haftalık grip faaliyet durumu.....	112

Çizelge 5.35. Rki yönteminin verdiği kararlar için sınıflandırma tablosu.....	112
Çizelge 5.36. Wavelet yöntemine ilişkin 2015-2016 grip mevsimi için haftalık grip faaliyet durumu.....	115
Çizelge 5.37. Haar-wavelets yönteminin verdiği kararlar için sınıflandırma tablosu .....	115
Çizelge 5.38. Wavelet yöntemine ilişkin 2015-2016 grip mevsimi için haftalık grip faaliyet durumu.....	116
Çizelge 5.39. Haar-wavelets yönteminin verdiği kararlar için sınıflandırma tablosu .....	116
Çizelge 5. 40. 2015-2016 grip mevsimi verileri için UCengiz yöntemine ilişkin istenen ATFS değerlerini veren parametreler.....	117
Çizelge 5. 41. UCengiz yöntemine göre ATFS 13 değeri için 2015-2016 grip mevsimine ilişkin haftalık grip faaliyet durumu.....	119
Çizelge 5. 42. UCengiz yönteminin ATFS 13 değeri için 2015-2016 grip mevsimine ilişkin verdiği kararların sınıflandırma tablosu.....	119
Çizelge 5. 43. UCengiz yöntemine göre ATFS 26 değeri için 2015-2016 grip mevsimine ilişkin haftalık grip faaliyet durumu.....	120
Çizelge 5. 44. UCengiz yönteminin ATFS 26 değeri için 2015-2016 grip mevsimine ilişkin verdiği kararların sınıflandırma tablosu.....	120
Çizelge 6.1. Simülasyon çalışmasında kullanılan veri üretim senaryoları.....	124
Çizelge 6.2. Dikkate alınan algoritmalar için yaklaşık ATFS=13 veren parametre değerleri (10000 tekrar).....	126
Çizelge 6.3. 2010-2017 grip mevsimlerine ilişkin salgın bilgileri.....	129
Çizelge 6.4. Simülasyon parametreleri .....	130
Çizelge 6.5. 4 hafta süreli ve 0.10 $\mu$ Baseline büyüklüğündeki salgınlarda algoritmaların performansları (10000 tekrar).....	131
Çizelge 6.6. 4 hafta süreli ve 0.25 $\mu$ Baseline büyüklüğündeki salgınlarda algoritmaların performansları (10000 tekrar).....	131
Çizelge 6.7. 4 hafta süreli ve 0.5 $\mu$ Baseline büyüklüğündeki salgınlarda algoritmaların performansları (10000 tekrar).....	132
Çizelge 6.8. 4 hafta süreli ve 1 $\mu$ Baseline büyüklüğündeki salgınlarda algoritmaların performansları (10000 tekrar).....	133
Çizelge 6.9. 4 hafta süreli ve 1.5 $\mu$ Baseline büyüklüğündeki salgınlarda algoritmaların performansları (10000 tekrar).....	133
Çizelge 6. 10. 4 hafta süreli salgınlar için en iyi performansa sahip algoritmalar ...	134



Çizelge 6.11. 8 hafta süreli ve 0.10 $\mu$ Baseline büyüklüğündeki salgınlarda algoritmaların performansları (10000 tekrar) .....	135
Çizelge 6.12. 8 hafta süreli ve 0.25 $\mu$ Baseline büyüklüğündeki salgınlarda algoritmaların performansları (10000 tekrar) .....	135
Çizelge 6.13. 8 hafta süreli ve 0.5 $\mu$ Baseline büyüklüğündeki salgınlarda algoritmaların performansları (10000 tekrar) .....	136
Çizelge 6.14. 8 hafta süreli ve 1 $\mu$ Baseline büyüklüğündeki salgınlarda algoritmaların performansları (10000 tekrar).....	137
Çizelge 6.15. 8 hafta süreli ve 1.5 $\mu$ Baseline büyüklüğündeki salgınlarda algoritmaların performansları (10000 tekrar) .....	137
Çizelge 6. 16. 8 hafta süreli salgınlr için en iyi performansa sahip algoritmalar ...	138
Çizelge 6.17. 12 hafta süreli ve 0.10 $\mu$ Baseline büyüklüğündeki salgınlarda algoritmaların performansları (10000 tekrar) .....	139
Çizelge 6.18. 12 hafta süreli ve 0.25 $\mu$ Baseline büyüklüğündeki salgınlarda algoritmaların performansları (10000 tekrar) .....	140
Çizelge 6.19. 12 hafta süreli ve 0.5 $\mu$ Baseline büyüklüğündeki salgınlarda algoritmaların performansları (10000 tekrar) .....	140
Çizelge 6.20. 12 hafta süreli ve 1 $\mu$ Baseline büyüklüğündeki salgınlarda algoritmaların performansları (10000 tekrar) .....	141
Çizelge 6.21. 12 hafta süreli ve 1.5 $\mu$ Baseline büyüklüğündeki salgınlarda algoritmaların performansları (10000 tekrar) .....	142
Çizelge 6. 22. 12 hafta süreli salgınlr için en iyi performansa sahip algoritmalar .	142

## ŞEKİLLER DİZİNİ

Şekil 1.1. Bioizlemenin alt dalları.....	1
Şekil 1.2. Halk sağlığı izlemesinin sınıflandırılması .....	2
Şekil 2.1. Uyarlamalı bir algoritma için referans, koruma bandı ve izleme periyodu .	7
Şekil 3.1. Ears C1, C2 ve C3 yöntemlerine ilişkin referans dönemleri .....	46
Şekil 4.1. 2012-2016 yıllarına ilişkin haftalık hasta sayısının zaman serisi, özilişki ve kısmi özilişki grafikleri.....	59
Şekil 4.2. 2012-2016 yıllarına ilişkin haftalık hasta sayısının 1. farklarının özilişki ve kısmi özilişki grafikleri.....	60
Şekil 4.3. Bazı grip sezonlarına ilişkin haftalık ILI kaynaklı ayakta tedavi edilen hasta oranlarının zaman serisi grafiği.....	63
Şekil 5.1. AIC ölçütüne göre en iyi Serfling model ve parametrelerinin kestirimleri	67
Şekil 5.2. Referans ve izleme dönemleri için tahmin ve eşik değerleri .....	67
Şekil 5.3. Farrington ve FarringtonFlexible yönteminden elde edilen izleme grafikleri .....	70
Şekil 5.4. Boda yönteminden elde edilen izleme grafikleri .....	72
Şekil 5.5. Aykırı değer analizi öncesi ve sonrası referans dönemi grafikleri .....	74
Şekil 5.6. Holt-Winters yöntemi artıklarına ilişkin grafikler .....	76
Şekil 5.7. 2015-2016 grip mevsimine ilişkin Holt-Winters yöntemiyle elde edilen 52 haftalık öngörülerin grafiği.....	77
Şekil 5.8. Aykırı değerlerden arındırılmış ve logaritması alınmış referans dönemi verilerine ilişkin özilişki ve kısmi özilişki grafikleri.....	79
Şekil 5.9. Birinci farkı alınmış verilere ilişkin özilişki ve kısmi özilişki grafikleri...	79
Şekil 5.10. Birinci farkı ve mevsimsel farkı alınmış verilere ilişkin özilişki ve kısmi özilişki grafikleri.....	80
Şekil 5.11. Aday modellerin artıklarına ilişkin Ljung-Box grafikleri.....	81
Şekil 5.12. Arima(0,1,1)(0,1,1) modeli artıklarına ilişkin özilişki, kısmi özilişki grafikleri ve 2015-2016 izleme dönemine ilişkin 52 haftalık öngörü grafiği .....	82
Şekil 5.13. CDC grip mevsimi verileri için Hidden-Markov modellerine ilişkin grafikler .....	84
Şekil 5.14. Holt-Winters yöntemi ile izleme dönemine ilişkin tahminler.....	86

Şekil 5.15. Holt-Winters yönteminden elde edilen artıkların özilişki grafiği .....	87
Şekil 5.16. Holt-Winters yönteminden elde edilen standartlaştırılmış hataların kontrol grafikleri .....	89
Şekil 5.17. Mevsimselliği giderilmiş 2015-2016 grip verileri için deneysel kontrol sınırları kullanan I-Chart( $\alpha \cong 0.01$ ) grafiği.....	91
Şekil 5.18. 2015-2016 izleme dönemi Ewma artıkları için teorik kontrol sınırları kullanan I-Chart( $\alpha = 0.01$ ) grafiği .....	93
Şekil 5.19. Mevsimsellik düzeltilmesi yapılmış artıkların 1. farklarının zaman serisi grafiği.....	95
Şekil 5.20. Rossi vd. (1999) Cusum yönteminin Serfling artıklarının 1. farklarına uygulanması.....	96
Şekil 5.21. Holt-Winters yönteminden elde edilen standartlaştırılmış hatalara ilişkin Ears C1, C2 ve C3 grafikleri .....	98
Şekil 5.22. Yeniden ölçeklendirilmiş CDC grip verilerinin Serfling regresyon eğrisi .....	100
Şekil 5.23. Yeniden ölçeklendirilmiş CDC grip verilerinin Serfling regresyon model tahmini .....	100
Şekil 5.24. PV ölçütünün 100 ve 1500 alarm sınırları için simülasyondan elde edilmiş değerleri .....	102
Şekil 5.25. 2015-2016 grip mevsimi için OutbreakP yöntemi sonuçları .....	103
Şekil 5.26. PV ölçütünün 75 ve 4000 alarm sınırları için simülasyondan elde edilmiş değerleri .....	104
Şekil 5.27. 2015-2016 grip mevsimi için OutbreakP yöntemi sonuçları .....	105
Şekil 5.28. CDC grip verilerinin Poisson regresyon ile kontrol grafiği.....	108
Şekil 5.29. Bayes yönteminden elde edilen izleme grafiği .....	110
Şekil 5.30. Rki yönteminden elde edilen izleme grafiği .....	112
Şekil 5.31. Mevsimsel olarak düzeltilmiş CDC grip verilerin zaman serisi grafiği	114
Şekil 5.32. CDC grip verilerinin 2 düzeyli Haar-wavelet dönüşümü grafikleri .....	114
Şekil 5.33. Mevsimsel olarak düzeltilmiş CDC grip verilerine ilişkin 1.farkların zaman serisi grafiği.....	115
Şekil 5.34. CDC grip verilerinin 2 düzeyli Haar-wavelet dönüşümü grafikleri .....	116
Şekil 5. 35. 2015-2016 grip mevsimi verileri için UCengiz yöntemine ilişkin haftalık sinyal grafikleri.....	118
Şekil 5. 36. 2015-2016 grip mevsimi verileri için UCengiz yöntemine ilişkin haftalık sinyal grafikleri.....	121

Şekil 6.1. Senaryolara göre gerçek ve yapay olarak üretilmiş grip verilerinin zaman serisi grafikleri .....	125
Şekil 6.2. Salgın yapısı .....	128
Şekil 6.3. 4 hafta süreli ve 0.10 $\mu$ Baseline büyüklüğündeki salgınlarda algoritmaların performans grafikleri .....	131
Şekil 6. 4. 4 hafta süreli ve 0.25 $\mu$ Baseline büyüklüğündeki salgınlarda algoritmaların performans grafikleri .....	132
Şekil 6. 5. 4 hafta süreli ve 0.5 $\mu$ Baseline büyüklüğündeki salgınlarda algoritmaların performans grafikleri .....	132
Şekil 6. 6. 4 hafta süreli ve 1 $\mu$ Baseline büyüklüğündeki salgınlarda algoritmaların performans grafikleri .....	133
Şekil 6. 7. 4 hafta süreli ve 1.5 $\mu$ Baseline büyüklüğündeki salgınlarda algoritmaların performans grafikleri .....	134
Şekil 6. 8. 8 hafta süreli ve 0.1 $\mu$ Baseline büyüklüğündeki salgınlarda algoritmaların performans grafikleri .....	135
Şekil 6. 9. 8 hafta süreli ve 0.25 $\mu$ Baseline büyüklüğündeki salgınlarda algoritmaların performans grafikleri .....	136
Şekil 6. 10. 8 hafta süreli ve 0.5 $\mu$ Baseline büyüklüğündeki salgınlarda algoritmaların performans grafikleri .....	136
Şekil 6. 11. 8 hafta süreli ve 1 $\mu$ Baseline büyüklüğündeki salgınlarda algoritmaların performans grafikleri .....	137
Şekil 6. 12. 8 hafta süreli ve 1.5 $\mu$ Baseline büyüklüğündeki salgınlarda algoritmaların performans grafikleri .....	138
Şekil 6. 13. 12 hafta süreli ve 0.1 $\mu$ Baseline büyüklüğündeki salgınlarda algoritmaların performans grafikleri .....	139
Şekil 6. 14. 12 hafta süreli ve 0.25 $\mu$ Baseline büyüklüğündeki salgınlarda algoritmaların performans grafikleri .....	140
Şekil 6. 15. 12 hafta süreli ve 0.5 $\mu$ Baseline büyüklüğündeki salgınlarda algoritmaların performans grafikleri .....	141
Şekil 6. 16. 12 hafta süreli ve 1 $\mu$ Baseline büyüklüğündeki salgınlarda algoritmaların performans grafikleri .....	141
Şekil 6. 17. 12 hafta süreli ve 1.5 $\mu$ Baseline büyüklüğündeki salgınlarda algoritmaların performans grafikleri .....	142

## SEMBOLLER VE KISALTMALAR DİZİNİ

$ARL^0$	Kontrol Altında Ortalama Çalışma Uzunluğu
$ARL^1$	Kontrol Dışında Ortalama Çalışma Uzunluğu
$ED_\tau$	Beklenen Gecikme
$t_A$	Algoritmanın Alarm Zamanı
$A(h)_J$	Alarm Durumu
$PSD(d, t)$	Başarılı Biçimde Belirleme Olasılığı
$Se(h)$	h Eşik Değerine Bağlı Duyarlılık
$Sp(h)$	h Eşik Değerindeki Belirlilik
$T(h, i)$	Gecikme Miktarı
$\tau$	Salgının Başlangıç Noktası
ACF1	1. Gecikmedeki Otokorelasyon
AIC	Akaike Information Criterion
AMOC	Activity Monitoring Operating Characteristics
ARIMA	Otoregresif Bütünleşmiş Hareketli Ortalamalar
ARL	Average Run Length
ARMA	Otoregresif Hareketli Ortalamalar
ATFOS	Average Time First Outbreak Signals
ATFS	Average Time to False Signals
BIOSENSE	ABD'deki ulusal düzeyde bir izleme sistemi
BODA	Bayesian Outbreak Detection Algorithm
CDC	Centers for Disease Control and Prevention
CUSUM	Birikimli Toplam Grafiği
EARS	Early Aberration Reporting System
EDD	Early Event Detection
EM	Expectation Maximization
ESSENCE	Electronic Surveillance System for the Early Notification of Community-based Epidemics
EWMA	Üstel Olarak Ağırlıklandırılmış Hareketli Ortalamalar

FMO	Fraction of Missed Outbreaks
FN	Yanlış Negatif
FP	Yanlış Pozitif
GLM	Genelleştirilmiş Doğrusal Model
GTM	Genelleştirilmiş Toplamsal Model
ILI	Grip Benzeri Hastalık
INAR	Tam sayılı Otoregresif Model
INLA	İç İçe Bütünleşik Laplace Yaklaşımı
LR	Likelihood Ratio
MA	Hareketli Ortalama
MAE	Mutlak Hataların Ortalaması
MAPE	Ortalama Mutlak Yüzde Hata
MASE	Ortalama Mutlak Ölçeklenmiş Hatalar
MCMC	Markov Chain Monte Carlo
ME	Hataların Ortalaması
ML	En Çok Olabilirlik
MPE	Yüzdelik Hata Ortalaması
NNDS	National Notifiable Diseases Surveillance System
PV	Predictive Value
RKI	Robert Koch Institute
RMSE	Ortalama Kareli Hatanın Karakökü
ROC	Receiver Operating Characteristics
RODS	Real-Time Public Health Surveillance
SA	Situational Awareness
SWT	Durağan Dalgacık Ayırımı
TN	Gerçek Negatif
TP	Gerçek Pozitif
TROC	Timeliness Receiver Operating Characteristics
WAVELET	Dalgacık

# 1. GİRİŞ

Halk sađlıđı konusunda önemli bir husus, birey ve toplum sađlıđını etkileyen olumsuz olayların zamanında tespiti ve önlemlerin alınabilmesidir. Örneđin; Genetik bozukluđa sahip bebeklerin doğum oranındaki artışın, çeşitli kanser türlerinin ve piyasaya yeni çıkan bir ilacın çeşitli yan etkilerinin belirlenmesi, çeşitli bulaşıcı hastalıkların vaka sayılarındaki artışın belirlenmesi önem taşımaktadır. Hızlı tespit ve önleme işlemleri hem bireyler hem de toplum için önem taşımaktadır. Bu nedenle halk sađlıđı izlemesi yapılmaktadır.

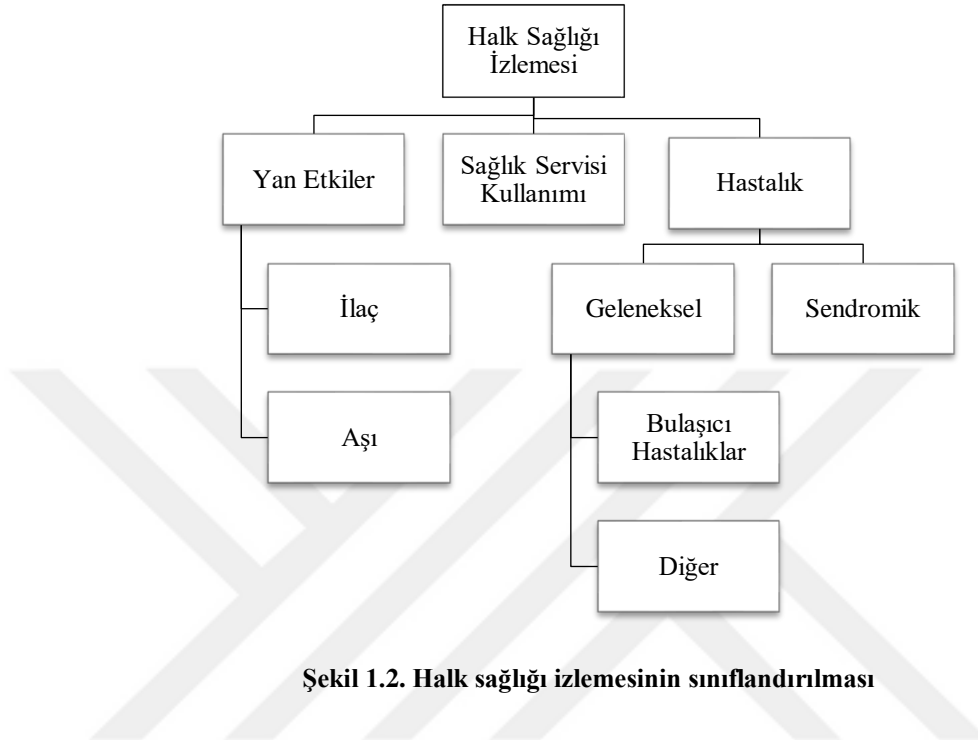
## 1.1. Halk Sađlıđı İzlemesi

Halk sađlıđından sorumlu kuruluş ve kişilerle paylaşım ve halk sađlıđı programlarının planlanması, yürütülmesi ve deđerlendirilmesi amacıyla; devamlı olarak sistematik bir biçimde veri toplanması, analizi ve yorumlanması faaliyetleri Halk Sađlıđı İzlemesi olarak adlandırılmaktadır. Thacker (2000) halk sađlıđı izlemesini; hastalıkları, yaralanmaları ve öteki sađlık problemlerini önlemek ve kontrol etmek için sađlıkla ilgili verilerin sürekli olarak sistematik biçimde toplanması, analizi, yorumlanması ve dağıtılması olarak tanımlamaktadır (Fricker, 2013). Halk sađlıđı izlemesi aslında biyoizlemenin bir alt dalı olarak görülebilir ve bioizlemede Şekil 1.1.'deki popülasyonların sađlıđının izlenmesi olarak tanımlanabilir (Fricker, 2013).



Şekil 1.1. Bioizlemenin alt dalları

Halk sađlıđı izlemesi; zellikle ilalar ve ařılar olmak zere tıbbi mdahalelere olan yan etkilerinin, sađlık hizmetlerinin nasıl kullanıldıđının ve salgınların izlenmesini kapsamaktadır (Fricker, 2013). Őekil 1.2. bu durumu tanımlamaktadır.



**Őekil 1.2. Halk sađlıđı izlemesinin sınıflandırılması**

Epidemolojik izleme yani salgınların izlenmesi Őekil 1.2.'de grldđ gibi klasik ya da sendromik izleme Őeklinde olmaktadır. Klasik izleme genellikle geriye dnk (gemiŐe dnk) izlemedir ve salgınlar oluŐtuktan sonra nedenlerini araŐtırmaya yneliktir. Klasik izleme ileriye dnk olabilmektedir. Ancak ileriye dnk klasik izleme, belirlenmesi hedeflenen ilgili hastalık salgının davranıŐının tamamen bilinmesini gerektirmektedir (Fricker, 2013). Sendromik izleme ise her zaman ileriye dnk olarak salgın oluŐmadan nce ilgili salgına iliŐkin belirtileri izlemeyi hedefleyen faaliyetleri iermektedir. Ancak klasik izlemenin tersine, bu tr izlemede belirlenmesi hedeflenen salgın tr genellikle tam olarak tanımlanamamaktadır (Fricker, 2013).



## 1.2. Halk Saęlıęı İzleme Sistemleri

Fricker (2013) biyoizleme sistemlerinin iki ana amacı olduęunu belirtmektedir. Bu amalar halk saęlıęı durumsal farkındalıęını (situational awareness SA) arttırmak ve salgınlara iliřkin erken olay belirlemeyi (early event detection EED) iyileřtirmektir. CDC (2008), durumsal farkındalıęı; bir salgına karřı etkili bir önlemin desteklenmesi ya da reddedilmesini, salgının büyüklüęünü, coęrafyasını, deęiřim hızını ve yařam döngüsünü izlemek amacıyla, ayrıntılı ve gerek zamanlı saęlık verilerini kullanma kabiliyeti olarak tanımlamaktadır. Bařka bir ifadeyle durumsal farkındalık salgının doęasını anlamak ve karřı önlemler alabilmek için gerekli olan bilgisel alt yapıyı temsil etmektedir. Erken olay belirleme ise CDC (2008) tarafından halk saęlıęına tehdit olabilecek olayları mümkün olan en kısa zamanda belirleme yeteneęi olarak tanımlanmaktadır (Fricker, 2013).

řu an kullanımda olan izleme sistemlerinden bazıları BIOSENSE, ESSENCE, EARS'dır. EARS, biyoterorizm kaynaklı saęlık problemleri için erkenden uyarı sistemi olarak tasarlanmıřtır. Dięer sistemler ise bařlangıta askerler arasındaki bulařıcı hastalık salgınlarını önceden belirlenmesi ve önlemlerin alınmasında gerekli koordinasyonun saęlanması amacıyla tasarlanmıřtır. Daha sonraları bu sistemler sivil halk arasında benzer saęlık problemlerinin izlenmesini kapsayacak řekilde genişletilmiřlerdir.

eřitli ülkelerde halk saęlıęı programları çerevesinde ok büyük miktarlardaki veri toplayan bazı kuruluşlar vardır. Örneęin ardıřık zaman noktalarındaki vaka sayıları, reetesiz ila satıřları, okullardaki devamsızlık durumları gibi ilgili veriler bu kurumlar tarafından izleme sistemlerine aktarılmaktadır. ABD'deki Centers for Disease Control and Prevention kurumu, National Notifiable Diseases Surveillance sistemine hafta hafta hem eyalet hem de ulusal düzeyde veri aktarmaktadır. İngiltere ve Galler'de ise Communicable Disease Surveillance Centre ve Public Health Laboratory Service ise benzer görevler yapmaktadır (Fricker, 2013).

### 1.3. Halk Saęlıęı İzlemesi Verilerin Özellikleri

Veriler sıklıkla kesikli bir rasgele deęişkene ilişkin zaman içinde elde edilen gözlemlerdir. Bu gözlemler genellikle rassal bir süreç ya da zaman serileri olarak görülebilir. Örneęin; bir bulaşıcı hastalığa ilişkin belirli zaman dilimlerindeki vaka sayıları, ilgili hastalığa ilişkin belirtilerle acil servise başvuran günlük hasta sayısı, ilgili hastalığa ilişkin belirtileri tedavi etmekte kullanılan reçetesiz ilaçların günlük satış sayısı, okullara devamsızlık durumları, ilgili hastalığa ilişkin web sitelerini ziyaret sayısı vb. olabilir. Gözlemlerin ardışık zaman noktalarındaki deęerlerinin yanı sıra bunların elde edildięi coęrafik bilgide (konumsal bilgide) bu tür verilerde yer alabilmektedir. Gözlemler vaka hızları biçiminde ifade edilebiliyorsa veriler sürekli de olabilir.

Salgınları izlemek için kullanılan veriler genellikle klasik istatistiksel yöntemlerin dayandıęı varsayımları ihlal etmektedir: duraęan olma, baęımsız olma, sabit varyanslılık, normallik vb. gibi. Genellikle bu tür veriler özilişkili olup, mevsimsellik ve trend etkileri göstermektedir.

### 1.4. Amaç ve Kapsam

Bu tez çalışmasında halk saęlığında hastalık salgınlarının önceden belirlenebilmesi amacıyla; zaman içinde ileriye dönük olaęan dışı durum izlemesi yapan bazı tek deęişkenli istatistiksel yöntemler tanıtılacak ve uygulamaları gerçek haftalık grip verileri üzerinde gerçekleştirilecektir. Ayrıca bu yöntemlerin bir kısmının performansları benzetim çalışması aracılıęıyla incelenecektir.

## 2. BİYOİZLEME ALGORİTMALARININ GENEL YAPISI

Bir alarm algoritması bir sađlık göstergesinin gözlemlenen frekanslarının normal seviyesinde bir deđişiklik olup olmadığının erken uyarısını vermek amacıyla tasarlanmış analitik süreçlerdir. Bu sađlık göstergeleri belirli bir bulguyla gelen hasta sayısı, bir hastalığın belirtilerini hafifletmek amacıyla kullanılan ilaç satışları ve öğrencilerin okula devamsızlık durumu gibi çeşitli veri türlerini içerebilmektedir. Etkin bir izleme için yani en az yanlış alarm ile olađandışı durumların erkenden belirlenmesi amacıyla; veri türleri, verilerin özellikleri ve bu verileri izlemekte kullanılan algoritmaların seçimi önemlidir. Algoritmalar veri türlerine uygun olarak seçilmelidir (Burkom, 2007).

### 2.1. Beklenen Veri Davranışının Belirlenmesi Amacıyla Oranlama İşlemi

Bir istatistiksel alarm algoritmasının kullanılmasıyla hedeflenen amaç, verilerdeki olađandışı durumu belirlemektir. Bu olađandışı durum bir tek zaman noktasındaki anormal bir artış olabileceđi gibi zaman içinde yavaş yavaş gelişen artışlar da olabilir. Bu amaçla ilgili veriler için olađan davranışın ne olduğunun bilinmesi önemlidir. Başka bir ifadeyle olađandışı durumun olmadığı durumu temsil eden ve referans olarak adlandırılan verilerin beklenen davranışının ortaya çıkarılması gereklidir (Burkom, 2007).

Verilerin beklenen davranışını ortaya çıkarabilmek amacıyla, frekans halinde olan veriler (time series of count data) mümkünse hızlar ya da oranlar şeklinde ifade edilmelidir. Örneđin, nüfus bilgisine ulaşılabiliyorsa veriler nüfusun büyüklüğüne bađlı hesaplanan hızlar şeklinde dikkate alınmalıdır. Çođu zaman nüfus gibi hız hesaplanmasını sađlayan bir bilgiye sahip olunmadığında, veriler anlamlı olan başka

bilgilere oranlanabilir. Örneğin, belirli bir zaman noktasında grip ve grip benzeri belirtilerle gelen hasta sayıları o zaman noktasındaki gelen toplam hasta sayısına oranlanabilir. Bu tür bir oranlama ise payda büyük olduğunda ve elde edilen oran 0.5'in altında olduğunda yararlıdır. Frekans verilerini standartlaştırmanın bir başka yolu da bunların bölüneceği paydayı zamanın fonksiyonu olarak ifade etmektir. Bunu gerçekleştirmek amacıyla şu yaklaşımlar kullanılabilir. Yaklaşımlardan biri daha önceki yaklaşımda olduğu gibi izlenen veriler içindeki ek bilgi kullanılmasıdır. Bir başka yaklaşım olarak yıl içinde bilinen tatil, okul kapanışı gibi olayların etkisi dikkate alınmasıdır. Diğer bir yaklaşım geçmişteki aynı dönemlere ilişkin ilgili frekansların dikkate alınmasıdır. Ayrıca hava sıcaklığı gibi dışsal veri kaynaklarının dikkate alınması kullanılacak yaklaşımlardan biridir. Her veri türü için bu yaklaşımların iyi sonuç vermesi beklenmemektedir (Burkom, 2007). Ayrıca en az yanlış alarmla en kısa zamanda sinyal vermesi beklenen bir algoritmanın, erişilebilecek açıklayıcı değişken bilgisini, örneğin; coğrafi bölge, yaş grubu, cinsiyet gibi bilgilerini de kullanması istenir.

## **2.2. Alarm Algoritmalarının Bileşenleri ve Algoritmik İşleyiş Adımları**

Algoritmalar uyarlamalı (adaptive) ve uyarlamasız (non-adaptive) olmak üzere 2 sınıfta gruplanabilir. Uyarlamalı bir algoritma, verilerin yakın geçmişteki davranışına uygun olarak parametrelerini güncellemektedir. Buna karşın uyarlamasız bir algoritma ise güncellenmeyen parametreler kullanmaktadır. Bir algoritma genel olarak aşağıdaki bileşenlerden oluşmaktadır (Burkom, 2007).

- Referans dönemi (eğitim dönemi)
- Test periyodu (izleme dönemi)
- Tampon dönemi (koruma bandı)
- Test değerleri
- Test istatistikleri (izleme istatistiği)
- Sıfırlama ölçütü
- Eşik değeri

- Isınma süresi

Referans dönemi, verilerin salgın olmadığı durumdaki beklenen davranışını belirlemek amacıyla dikkate alınan dönemdir. Beklenen davranış basit istatistiklerle ifade edilebileceği gibi karmaşık modellemelerle de tanımlanabilir. Test periyodu ya da izleme dönemi ise, ilgili verilerin belirli zaman aralıklarıyla beklenen davranışa sahip olup olmadığını belirlemek amacıyla izlendiği dönemdir. Tampon dönemi ya da koruma bandı olarak adlandırılan zaman dilimi, referans dönemi ile izleme dönemi arasındaki zaman dilimidir. Böyle bir zaman diliminin kullanılmasının amacı referans döneminin sonlarına doğru başlayan yavaş gelişen salgınları belirleyebilmektir. Test değerleri ise, izleme döneminde olağandışı durumların karanımı vermek için kullanılan test istatistiğini yani izleme istatistiğini oluşturmak amacıyla gözlemlenen ve izlenen niceliklerdir. Sıfırlama ölçütü ise, bir alarmdan sonra tekrarlı alarmları önlemek için izleme istatistiğinin değerinin başlangıç değerine ya da uygun bir değere eşitlenmesidir. Eşik değeri ise, olağandışı durum sinyalinin verilebilmesi için izleme istatistiğinin ulaşması gereken en düşük değerdir. Isınma süresi ise, algoritmanın ileriye dönük izleme amaçlı kullanılması için gerekli en kısa referans zaman aralığıdır. Şekil 2.1. uyarlamalı bir algoritma için referans, koruma bandını ve izleme periyodunu göstermektedir (Burkom, 2007).

Referans Dönemi	Koruma Bandı	İzleme Dönemi
-----------------	--------------	---------------

**Şekil 2.1. Uyarlamalı bir algoritma için referans, koruma bandı ve izleme periyodu**

Algoritmik bir süreçte genellikle aşağıdaki adımlar takip edilmektedir.

- İzleme dönemi için beklenen değerler güncelleştirilir. Başka bir ifadeyle, algoritma uyarlamalı bir algoritma doğasına sahipse, modelin parametreleri yakın zamanda elde edilen veriler kullanılarak yeniden tahmin edilir.
- İzleme dönemi verileriyle modelden elde edilen beklenen değerler arasındaki farklar hesaplanır.

- Bir önceki adımda elde edilen farklar standartlaştırılır. Örneğin model bir regresyona dayanan bir model ise, farklar regresyon standart hatasına bölünür.
- Bu standartlaştırılmış farklara dayanarak izleme istatistiği hesaplanır.
- Son olarak alarm kararını vermek üzere deneysel veya kuramsal eşik değeri bulunarak izleme istatistiğinin değeri ile karşılaştırılır.

### 2.3. İzleme Verilerindeki Değişimin Kaynakları

Çeşitli nedenlerden dolayı izleme verileri değişim göstermektedir. Bu değişimin bir kısmı modellenebilmektedir çünkü değişime neden olan durumlar beklenmekte ve öngörü zamanında bunlara ilişkin bilgiye ulaşılabilir. Bazıları da modellenememektedir. Bunun nedeni ise değişim hakkında gerekli bilgi yoktur veya değişim beklenmemektedir. Modellenebilen durumların bazıları aşağıda belirtilmektedir (Burkom, 2007).

- Geçmiş verilerden ölçülebilen doğal rasgele değişkenlik
- Haftanın günü etkisi (özellikle hafta içi günler ile hafta sonu günler arasındaki sistematik farklılık)
- Mevsimsel trendler
- Tatiller veya okul kapanışları gibi olaylar
- İklimsel değişiklikler (örneğin; günlük en yüksek sıcaklık gibi)

Kolaylıkla modellenemeyen değişim kaynaklarının bir kısmı ise aşağıda listelenmektedir.

- İzlenen nüfusun büyüklüğünde veya davranışında olan değişiklikler
- Planlanmamış tatiller veya kötü hava koşulları
- Sağlık hizmeti talebindeki değişiklikler
- Veri tedarikçilerinin sisteme iştirak etmelerindeki değişiklikler
- Sağlık hizmeti alma hakkındaki olan değişiklikler

## 2.4. Salgın Sinyalinin Yapısı

İzleme için uygun algoritmanın seçimi belirlenmesi istenen sinyale bağlıdır. Bir sinyal ise ilgilenilen sağlık olayına ilişkin izdir. Ani artışlar belirlemek ya da aramak dışında, sinyalin yapısı biliniyorsa, izleme verilerinden oluşan zaman serilerinin bilinen o yapıya sahip kısımları bulunmaya çalışılır (Burkom, 2007). Salgın sinyallerine ilişkin muhtemel yapılar aşağıdaki gibi maddelendirilebilir.

- Basamak biçimindeki sinyal
- Mevsimsel etki temelli salgın sinyali
- Belirli hastalıklar için kuluçka dönemlerine bağlı olarak uzun ve kısa süreli sinyaller

İzlenen verilerde basamak şeklinde bir artışı gösteren sinyal söz konusu olabilir. Ancak bu tür salgınlar muhtemelen geçici ve mevsimsellik özelliğine sahip olduğundan, bunları belirleyecek algoritmalar bu özellikleri dikkate almalıdır. Başka bir sinyal yapısı da yılın belirli bir zaman diliminde artışa geçip en yüksek bir değere veya değerlere ulaştıktan sonra azalma eğilimine girmesidir. Örneğin; grip ve benzeri belirtilerle gelen hasta sayısı sonbahar ve kış aylarında bu yapıya sahip bir davranış göstermektedir. Diğer bir sinyal yapısı da; salgın yokken verilerin fazla değişim göstermediği bir durumda kuluçka sürelerine bağlı olarak uzun veya kısa bir zaman diliminde bulunduğu seviyeden artış eğilimine geçerek zirveye ulaşması ve daha sonra azalma eğilimine geçerek normal seviyesine ulaşmasıdır (Burkom, 2007).

## 2.5. Salgınların Erken Belirlenmesinde Karşılaşılan İstatistiksel Zorluklar

Biyozlemedeki istatistiksel zorluklar; toplanan verilerin doğasından, salgın olmayan döneme ilişkin beklenen normal davranışın belirlenmesinden, salgının doğal yapısından, algoritmaların performanslarının nasıl değerlendirileceği gibi hususlardan kaynaklanmaktadır (Shmueli ve Burkom, 2010).

Modern biyoizleme ya da diđer adıyla sendromik izlemedeki amaç salgınların erkenden belirlenmesi olduđundan, ilgilenilen salgınla ilgili önceden bilgi verebilecek çeşitli göstergeler yani tanı öncesi veriler izlenir. Dolayısıyla izlenen veriler, ilgilenilen hastalığa ilişkin klinik veriler olmayıp, bu hastalığa ilişkin dolaylı olarak bilgi vermesi beklenen göstergelerdir. Örneđin ilgilenilen hastalığa yönelik günlük ya da haftalık reçetesiz ilaç satışları, internet aramaları, okul ya da işe devamsızlık durumları, doktor ya da acil servis ziyaretleri sendromik izleme verileri olarak kullanılabilir.

Zaman serileri biçiminde elde edilen bu veriler genellikle durađan deđillerdir. Başka bir ifadeyle bu serilerin ortalamaları, varyansları ve otokorelasyon yapıları zamana bađlı olarak deđişmektedir. Verilerin durađan olmama özelliđi standart yöntemlerin uygulanmasında güçlükler yol açmaktadır. Ancak bu seriler beklenen bazı sistematik yapılaraya sahip olabilmektedir. Örneđin haftanın günü etkisi, mevsimsellik etkisi, tatil öncesi ve sonrası etkisi gibi etkiler yanında, hava koşullarındaki ve nüfusun davranış yapısındaki mevsimsel yıllık deđişim nedeniyle döngüsel etkilerde verilerde görülebilmektedir. Ayrıca verilerin günlük, saatlik gibi sık aralıklarla toplanması da verilerin dikkate deđer şekilde özilişkili olmasını sağlamaktadır. Bütün bu özellikler, verilerde olabilecek tutarsızlık, eksiklik gibi hatalarla birleşince, sendromik verilerin analizi ve yorumlanması güçleşmektedir (Shmueli ve Burkom, 2010).

Sendromik verilerin izlenmesindeki amaç bir anormalliđin varlıđının belirlenmesidir. Ancak anormalliđi belirlemek normal olanın bilinmesini gerektirir. Amaç Bioterör kaynaklı bir salgını belirlemekse; referans ya da normal davranış (baseline), dođal nedenli salgınlarda oluşacak davranışları da kapsar. Başka bir ifadeyle referans verilerinin salgın içermediđi kabul edilir. Eđer amaç dođal kaynaklı salgınları belirlemekse; referans ya da normal davranış, salgın olmayan dönemlerdeki veri yapısının bilinmesini gerektirir. Salgın olmayan dönemlerin belirlenmesi zordur. Enfeksiyonun başladıđı tarih ile tedavi gereksinimi duyulan tarih arasındaki gecikmelerin sayısal olarak ifade edilmesi kolay deđildir. Bunun sonucu olarak algoritmaların istatistiksel olarak deđerlendirilmesi ve karşılaştırılması işlemini güçleşir. Referans dönemindeki salgınların belirlenmesindeki güçlüđe ek olarak, izlenen



kitlelerdeki veri bildirimindeki deęişiklikler, saęlık mevzuatındaki deęişiklikler vb. faktörler duraęan olmayan ve sürekli deęişen referans davranışına neden olurlar. Bu nedenle, bioizleme iyi tanımlanmamış bir referans (normal davranış) problemine sahiptir (Shmueli ve Burkom, 2010).

Sendromik izleme genellikle klinik olmayan verilere dayalı olduğundan, salgın dönemlerinde verilerin nasıl bir sistematik yapı göstereceğine dair bilgi yoktur. Böyle bir sistematik yapı bilgisinin yokluęunda, salgınların belirlenmesi anormallik belirlenmesi durumuna dönüşür. Dolayısıyla sendromik izleme yani bioizleme belirli bir hastalığın salgınına ilişkin belirlenmiş bir davranışı bulma işlemi deęil, her türlü olaęandışı davranışı belirleme işidir. Bu durum, kalite kontrol şemaları gibi belirlenmiş bir kontrol dışı davranışı belirlemek amacıyla kullanılan standart yöntemlerin bioizlemede kullanılmasını zorlaştırmaktadır (Shmueli ve Burkom, 2010).

Bioizlemede sıklıkla kullanılan algoritma deęerlendirme ölçütleri olarak duyarlılık (sensitivity), belirlilik (specificity) ve dakiklik (timeliness) farklı algoritmaları karşılaştırmak amacıyla kullanılmaktadır. Ancak bu ölçülerin saęlıklı şekilde hesaplanabilmesi için verilerin bazı özelliklere sahip olması gerekir. Örneğin duyarlılık ölçütünün ifade ettięi salgınları doęru belirleme olasılığını iyi bir şekilde tahmin edebilmek amacıyla verilerde yeterince salgın olayının gerçekleşmesi gerekir. Bir algoritmaya ilişkin belirlilik ölçütünün hesaplanabilmesi için salgın olmadığında salgın sinyali vermemesi olasılığını tahmin edebilmesi gerekir. Bu ise salgın olmayan dönemlerin doęru biçimde daha önceden bilinmesini gerektirir. Salgının başladığı zaman noktası ile salgın sinyali arasındaki süreyi ifade eden dakiklik ölçütü ise, salgının başladığı noktanın doęru olarak tespit edilmesi gerektirir. Ancak salgın olmayan dönemlerin ve salgının başladığı zaman noktasının belirlenebilmesi kolay deęildir. Bu problemler algoritmaların performanslarının istatistiksel olarak deęerlendirilmesi ve karşılaştırılmasında zorluklar oluşturur (Shmueli ve Burkom, 2010).

Bioizlemede genellikle mevcut veriler (zaman serileri) tek tek izlenmektedir. Ayrı ayrı izlemekten ve ayrı ayrı yöntemler kullanmaktan dolayı elde edilen sonuçların nasıl birleştirileceęi önemli bir husustur. Bir dięer önemli husus ise verilerin çok deęişkenli

izlenmesi halinde ortaya çıkabilecek problemlerdir. Örneğin; birlikte izlenen serilerin bildiriminde farklı gecikmeler yaşanması, bu seriler arasındaki ilişkilere dair bilgilerin olmayışı ve bu verilerin birlikte durağan olmama problemleri gibi sorunlar istatistiksel analizi ve yorumlamayı güçleştirmektedir (Shmueli ve Burkom, 2010).

## **2.6. Algoritmaların veya Verilerin Biyoizlemeye Uyarlanma Stratejileri**

İki genel yaklaşım söz konusudur. Bunlardan ilki klasik istatistiksel yöntemleri kullanmak amacıyla, doğrudan bioizleme verilerini kullanmak yerine bu verilerin uygun şekilde modellenmesiyle elde edilebilecek artıkların kullanılmasıdır. Artıkların durağanlık, bağımsızlık ve normallik gibi klasik yöntemlerin gerektirdiği özelliklere sahip olması beklendiğinden, istatistiksel süreç kontrol grafikleri gibi standart yöntemler bioizlemede anormallikleri belirlemek amacıyla kullanılabilir. Ancak kontrol grafiklerinin bioizlemede kullanılması tek yönlü olmakta yani artış yönündeki olağandışı durumların belirlenmesi şeklinde olmaktadır ve alt grupların ortalamaları yerine bireysel gözlemlere dayanmaktadır. Grafiklerin tek yönlü olarak kullanılması yanlış alarm hızını azaltmasına rağmen, bireysel gözlemlerin kullanılması varyansı arttırmakta ve özilişki sorununu ortaya çıkarmaktadır. Bundan dolayı yanlış alarm hızı artmaktadır. Ayrıca gözlemlerin normal dağılmaması bunun yerine sıklıkla uzun kuyruklu olan negatif binom veya aşırı yayımlı Poisson dağılımına sahip olması nedeniyle de yanlış alarm hızının artması beklenmektedir (Burkom, 2007).

Artıkları izleme yaklaşımında performans iyileştirmesini sağlamak için kullanılan yaklaşımlardan biri de, verilerin birden fazla modelleme yöntemiyle modellenmesi ve birden fazla artık serisi elde edilmesidir. Daha sonra bu artık serileri, performansı en iyi yapacak şekilde doğrusal bir fonksiyonla birleştirilerek tek artık serisi elde edilir ve kullanılacak kontrol grafiği bu doğrusal fonksiyona uygulanır. Benzer şekilde artıkların izlenmesinde bir diğer yaklaşım ise, birden fazla kontrol grafiğinin en iyi performansı verecek şekilde doğrusal bir fonksiyonla birleştirilmesidir (Yahav vd., 2011).

İkinci yaklaşım ise mevcut yöntemlerin bioizleme verilerine ilişkin genellikle durağan olmama ve özilişki özelliklerini dikkate alacak şekilde uyarlanmalarıdır. Bioizleme verilerinin sürekli değişen doğasını dikkate almak için standart yöntemler, referans dönemini sürekli yenileyecek ve izleme dönemi ile referans dönemi arasına koruma bandı ekleyecek biçimde düzenlenmektedir. Örneğin durağan olmayan ve özilişkili veri durumunda, ortalamalar ve standart sapmalar kayan bir referans kullanılarak ayarlanabilmektedir. Böyle bir referans ile izleme yapılacak zaman noktası arasına koruma bandı da konulabilmektedir. Bu tür bir uyarlamalı algoritmalara örnek olarak EARS C1, EARS C2 ve EARS C3 algoritmaları örnek verilebilir. Başka bir örnek ise ESSENCE; 28 günlük referans ve 2 günlük koruma bandını kullanan uyarlamalı EWMA algoritmasıdır (Burkom, 2007).

Bioizlemede salgın sinyalleri geçici olmakta bazen sinyaller birkaç gün bazen de aylarca sürebilmektedir. Grip gibi bir hastalık için bile salgın sinyali aniden ve büyük olabileceği gibi daha yavaş ve azar azar artan şekilde olabilmektedir. Bu tür sinyalleri daha etkin bir şekilde belirlemek amacıyla, çeşitli kontrol grafikleri bir arada kullanılabilir. Örneğin Cusum-Shewart ve Ewma-Shewart grafikleri bu grafiklerdendir (Burkom, 2007).

## **2.7. Salgın Belirleme Yöntemlerinin Sınıflandırılması**

Salgın belirleme yöntemleri çeşitli şekilde sınıflandırılabilir. En basit sınıflandırma işleminde ileriye dönük ve geriye dönük izleme yöntemleri olarak ayrım yapılabilir. Bir diğer sınıflandırma ise, yöntemler sadece zaman içinde izleme yapabilen ve hem zaman hem de mekan içinde izleme yapabilen yöntemler olarak sınıflandırılır. Ayrıca aynı anda izlenen değişken sayısına bağlı olarak tek değişkenli ve çok değişkenli izleme olarak da sınıflandırma yapılabilmektedir. Başka bir sınıflandırmada ise yöntemler uyarlamalı ve uyarlamasız algoritmalar biçiminde sınıflandırılabilir. Farklı bir sınıflandırma ise yöntemleri, bioizleme verilerindeki haftanın günü etkisi, mevsimsellik ve özilişki gibi sistematik yapıları ele

alış biçimlerine bağlı olarak model tabanlı ve veri tabanlı olarak da ayırt etmek mümkündür.

### **2.7.1 Bazı biyoizleme sistemlerinde kullanılan mevcut algoritmalar**

Şu an ki izleme sistemleri çoğunlukla istatistiksel süreç kontrol yöntemleri ve regresyona dayanan yöntemler gibi klasik izleme araçlarını kullanmaktadır (Shumeli ve Burkom, 2010). Örneğin EARS sistemi C1, C2 ve C3 olarak adlandırılan uyarlanmış kontrol grafikleri kullanmaktadır; C1 ve C2 grafikleri Shewhart I-Chart'ın uyarlanmış biçimleri iken, C3 Cusum grafiğinin uyarlanmış bir biçimidir. Bu grafikler verilerin durağan olmayan yapısını dikkate almak amacıyla, kayan bir referans kullanmakta ve yavaş gelişen salgınların referans dönemi verilerini etkilemesini önlemek amacıyla C2 ve C3 bir koruma bandı kullanmaktadır. Bir diğer örnek olarak ESSENCE sistemi, doğrudan vaka sayılarına uygulanan Ewma grafiği ile doğrusal regresyon artıklarına dayanan Shewhart I-Chart arasında bir uyum iyiliği istatistiğinin değerine bakarak izleme yapmaktadır. Ayrıca bu grafikler 4 haftalık kayan bir referans ile bir haftalık kayan bir koruma bandı kullanmaktadır (Shumeli ve Burkom, 2010).

RODS sistemi de izleme için 4 algoritma kullanmaktadır. Bunlardan ilki 120 günlük kayan referans kullanan MA grafiği iken, ikincisi ise Cusum ve Ewma grafiklerinin birlikte kullanımına dayanan bir algoritmadır. Ewma ile bir sonraki gündeki ilgili olay sayısı tahmin edilmekte, Cusum ile  $h=4.08$  eşik değeri kullanılarak bu tahmin sonucu elde edilen artık izlenmektedir. Bu grafiklerde teorik eşik değerleri yerine istenilen performansı verecek deneysel eşik değerleri kullanılmaktadır. RODS izleme sisteminin kullandığı üçüncü yöntem ise ardışık en küçük kareler yöntemiyle parametrelerinin sürekli olarak güncellendiği otoresif bir modelin kullanılması ve bu modelden elde edilen artıkların 4 standart sapma eşik değerine sahip Shewhart I-Chart kullanılarak izlenmesidir. Bu algoritmaya benzer şekilde BioSense günlük izleme verilerini Poisson regresyon kullanarak modellemektedir. Bu modelleme sonucu değişik coğrafik bölgelere ilişkin  $p$  değerleri birleştirilerek elde edilen SMART skorlar izlenmektedir.

Regresyona ya da kontrol grafiğine dayanmayan ve RODS sistemi tarafından kullanılan bir algoritma da WAVELET algoritmasıdır. Bu algoritma Haar wavelet'lar ile zaman serisini ayrıştırıp, serinin uzun dönemli trendini gidermektedir. Trendi giderilmiş seriyi yani artıkları 4 standart sapmadan oluşan eşik değeri kullanarak Shewhart I-Chart ile izlemektedir (Shumeli ve Burkom, 2010).

### **2.7.2 Bioizleme verilerindeki sistematik yapıyı dikkate almak amacıyla kullanılan ana yaklaşımlar**

Bioizleme verilerindeki haftanın günü etkisi, mevsimsellik ve özilişki gibi etkilerin giderilerek bağımsız normal dağılımlı artıklar elde etmek amacıyla 2 ana yaklaşım vardır. Bunlardan ilki modellemeye dayanan (model-based) yaklaşımlar iken diğeri ise gözlemlenen izleme verisine dayanan (data-driven) yaklaşımlardır.

Modellemeye dayanan yaklaşımlar genellikle regresyon modellemesine ve Arima modellemesine dayanmaktadır. Regresyon modellemesinde sıklıkla doğrusal ve Poisson regresyon modelleri kullanılmaktadır. Regresyon modellerinin bioizleme verilerindeki sistematik yapıyı yeterince açıklayabilmelerine rağmen, verilerin durağan olduğunu varsayması ve uzun bir referans gerektirmesi dezavantajlarıdır. Bioizleme verileri genellikle durağan yapıda olmayıp, geçmişe ait uzun dönemli verilere ulaşılması kolay değildir. Geçmişe ait veriler elde edilse de şu an ki durumun temsilcisi olduğunu varsaymak makul değildir (Shumeli ve Burkom, 2010).

Modellemeye dayalı ikinci yaklaşım Arima modelleme, çok sayıda serinin aynı anda izlendiği otomatik izleme sistemine uygun değildir çünkü her bir seri ayrı bir model gerektirebilmektedir. Hatta bazı serilerin durağan hale getirilememesinden dolayı uygun bir Arima model bulunamaması ihtimali de söz konusu olmaktadır. Arima modellemesine ilişkin başka bir sakınca ise yavaş gelişen salgınların modele dahil edilme ihtimali olmasıdır. Bu durum yavaş yavaş oluşan salgınların belirlenememesi veya belirlenmesinde geç kalınması ile sonuçlanabilmektedir. Dolayısıyla Arima modellemesi ileriye dönük izleme için elverişli olmamaktadır (Shumeli ve Burkom, 2010).

Bioizleme, verilerde gözlemlenen özilişki, haftanın günü etkisi gibi sistematik bileşenlerin etkisini gidermek amacıyla verilerin dönüşümlerine dayanan yöntemlerde önerilmiştir. Bu yaklaşımlar veriye dayalı yaklaşımlar olarak bilinmektedir. Örneğin Muscatello (2004) haftanın günü etkisini gidermek amacıyla 7 günlük fark alma işlemini önermiştir (Shumeli ve Burkom, 2010).

Haftanın günü etkisini gidermek için veri düzleştirme yöntemleri de önerilmiştir. Ancak düzleştirme ile elde edilen kazanç, düzleştirme işlemi neticesinde algoritmanın salgını belirlenmesinde oluşan gecikme nedeniyle önemini yitirmiştir. Ancak Burkom vd. (2007) zaman içinde değişen trend, haftanın günü etkisi ve mevsimsellik gibi etkiler için uygun olan Holt-Winters üstel düzleştirmesinin diğer düzleştirilmelere kıyasla daha iyi bir performansa sahip olduğunu göstermiştir. Verilere dayanan bir diğer yaklaşım ise wavelet türü yöntemlerdir. Bu yaklaşım genel amaçlı bir anormallik belirleme yöntemi olup ileriye dönük tahmin ve izleme için sıklıkla kullanılmamıştır. Ayrıca ileriye dönük izleme amacıyla düzenlenmiş olan wavelet yaklaşımları da bulunmaktadır (Shumeli ve Burkom, 2010).

## **2.8. Algoritmaların Performanslarını Değerlendirme Ölçütleri**

Fricker (2013) ilgili literatürde izleme sistemlerinin performansını değerlendirmekte genelde 3 ölçütün kullanıldığını belirtmektedir. Bunlardan ilki hızlı ve doğru karar verme hızıdır (timeliness) yani ilgili olay gerçekleşmeden önce onu belirleme yeteneğini ölçen ölçüttür. İkinci ölçüt ise yanlış karar verme hızıdır (specifity). Başka bir ifadeyle ilgili olay gerçekte yok iken yanlışlıkla ilgili olayın var olduğu ikazlarının gerçekleştirme hızıdır. Son ölçüt ise ilgilenilen olayın örneğin salgının doğru olarak belirlenmesi olasılığıdır (sensitivity). Bu olasılık ise, sistemin verdiği çok sayıda ilgili olayın varlığını belirten kararlarının doğruluk oranını yansıtmaktadır. Bir izleme sistemi için bu ölçütlerden üçü aynı anda iyileştirilememektedir. Bir ölçüt iyileştirilmek istendiğinde diğer ölçüt veya ölçütlerden ödün verilme zorunluluğunda kalınmaktadır (Fricker, 2013).

Bu ölçütleri matematiksel olarak ifade etmekte kullanılan ve alarm durumu olarak adlandırılan iki durumlu bir değişken aşağıdaki gibi tanımlanır (Buckeridge vd., 2007).

$$A(h)_j = \begin{cases} 1, & S(h, j) > h \\ 0, & \text{diğer durum} \end{cases}$$

Burada  $h$  eşik değerini,  $j$  kararın verildiği zaman noktasını ve  $S(h, j)$  ise algoritmanın ilgili  $(h, j)$  ikilisi için verdiği değeri göstermektedir. Tanımlanan bu alarm durumu değişkenine göre belirlilik, duyarlılık ve dakiklik aşağıdaki gibi tanımlanmaktadır. Belirlilik salgın yokken algoritmanın alarm vermeme olasılığı olarak tanımlanır.

$$\text{Belirlilik} = P(\bar{A}|\bar{O}) = \frac{n(\bar{A}, \bar{O})}{n(\bar{O})}$$

Burada  $O$  salgın olayını,  $A$  alarm verilmesi olayını ve  $n()$  ise ilgili olayın gerçekleşme sayısını belirtmektedir.  $n(\bar{O})$  salgın olmayan zaman noktalarının sayısı yani referans dönemindeki gözlem sayısıdır. Bu durumda referans dönemi için belirlilik aşağıdaki gibi hesaplanır (Buckeridge vd., 2007).

$$Sp(h) = 1 - \frac{1}{m} \sum_{j=1}^m A(h)_j$$

Burada  $Sp(h)$ ,  $h$  eşik değerindeki belirliliği ve  $m$  referans dönemindeki gözlem sayısını göstermektedir. Belirlilik olasılığını yanlış alarm hızı olarak ifade etmek için, bu olasılığın 1'den farkı belirli bir zaman aralığında verilen kararların sayısı ile çarpılır. Örneğin, haftalık verilerde aylık yanlış alarm hızını bulmak için haftalık verilerden hesaplanmış belirlilik değerinin 1'den farkı 4 ile çarpılarak 1 ayda beklenen yanlış alarmların sayısı yani hızı bulunur. Bir diğer ölçüt duyarlılık ise gerçekte salgın varken algoritmanın alarm verme olasılığıdır. Bu olasılık aşağıdaki gibi tanımlanır (Buckeridge vd., 2007).

$$\text{Duyarlılık} = P(A|O) = \frac{n(A, O)}{n(O)}$$

Burada  $n(O)$  salgın olan zaman noktalarının sayısını ve  $n(A, O)$  salgın olduğu durumda algoritmanın da alarm verdiği zaman noktalarının sayısını göstermektedir. Her birinde bir salgının var olduğu ve bu salgının  $m_i$  adet karar noktasında devam ettiğinin kabul edildiği  $n$  adet veri kümesi için duyarlılık aşağıdaki gibi hesaplanmaktadır (Buckeridge vd., 2007).

$$Se(h) = \frac{1}{n} \sum_{i=1}^n \min \left( 1, \sum_{j=1}^{m_i} A(h)_{ij} \right)$$

Burada  $Se(h)$ ,  $h$  eşik değerindeki duyarlılığı göstermektedir. Buradaki sakınca, duyarlılık ve belirlilik ölçütlerinin aynı ölçekte hesaplanmamasından kaynaklanır: bu durum iki ölçütün doğrudan karşılaştırılmasını karmaşık hale getirmektedir. Son ölçüt olan salgın belirlemenin dakikliği, bir salgının başladığı zaman noktasına göre veya bir başka yöntemin belirlediği salgının başlangıcına göre ölçülebilir. Yapay olarak oluşturulmuş ve  $m_i$  adet karar noktası boyunca süren bir tek salgının olduğu durumda, bu salgının başlangıcına göre dakiklik aşağıdaki gibi tanımlanmaktadır (Buckeridge vd., 2007).

$$T(h, i) = \min_j (j: A(h)_{ij} = 1)$$

Burada alarm durumu değişkeni  $A(h)_{ij}$ , salgının devam ettiği  $m_i$  karar noktasının her birinde incelenmektedir ve  $\sum_{j=1}^{m_i} A(h)_{ij} = 0$  ise dakiklik tanımlanamamaktadır. Bu üç ölçütün ikiyeşerli veya tamamının birlikte yorumlanabilmesi için çeşitli eğriler veya yüzeyler çizilebilmektedir. Örneğin çeşitli eşik değerlerindeki duyarlılıklar, belirliliklerin 1'den farklarına karşı çizilerek ROC (receiver operating characteristics) eğrisi elde edilebilir ya da belirlilikler dakiklik değerlerine karşı çizilerek AMOC (activity monitoring operating characteristics) eğrisi elde edilebilir. Ayrıca dakiklik hem duyarlılıklara hem de belirliliklere karşı çizilerek TROC (timeliness receiver operating characteristics) yüzeyi elde edilebilmektedir (Buckeridge vd., 2007).



Yukarıda bahsedilen temel ölçütlere ek olarak kullanılan bazı ölçütlere ilişkin tanımlar burada verilecektir. Burada  $\tau$  salgının başlangıç zamanını,  $t_A$  salgının başlangıcına ilişkin izleme yönteminin verdiği alarmın zamanını,  $D$ , salgının henüz başlamadığı zaman periyodunu,  $C$  ise salgının başladığı ve devam ettiği zaman periyodunu göstermektedir. Alarm sıklığı ARL (Average Run Length) ile ölçülür. Sistemde değişim olmamışken yani salgın yokken ARL ölçütü yanlış alarm hızını  $ARL^0$ 'ı verir.

$$ARL^0 = E[t_A | D]$$

Yaygın olarak kullanılan bir diğer ölçüt ise yanlış alarm olasılığıdır. Bu ölçüt aşağıdaki gibi tanımlanmakta ve  $\tau$ 'nun bir dağılıma genellikle de Geometrik bir dağılıma sahip olduğunu varsaymaktadır. Bu varsayım, her zaman noktasında sistemdeki değişimin sıklığı sabit olduğunda uygundur (Sonesson ve Bock, 2003).

$$P(t_A < \tau) = \sum_{t=1}^{\infty} P(\tau = t)P(t_A < \tau | \tau = t)$$

Sistemdeki değişim  $\tau = 1$  anında yani izlemenin başlangıcı ile birlikte sistemde değişim gerçekleşiyorsa, alarmın gecikmesi  $ARL^1$  ile ölçülür.

$$ARL^1 = E[t_A | \tau = 1]$$

Sistemdeki değişimin izlemeye başladıktan sonra herhangi bir zamanda gerçekleşeceği düşünüldüğünde, daha genel bir alarm gecikme ölçüsü koşullu beklenen gecikmedir (CED).

$$CED(t) = E[t_A - \tau | t_A \geq \tau, \tau = t]$$

$\tau$ 'nun bir dağılıma sahip olduğu varsayılırsa, beklenen gecikme aşağıdaki gibi hesaplanabilir (Sonesson ve Bock, 2003).

$$ED_{\tau} = \sum_{t=1}^{\infty} P(\tau = t)P(t_A \geq t)CED(t)$$

Birçok izleme yöntemi için  $\tau$  birden çok değer olma eğiliminde olduğunda, CED(t) sabit bir değere yakınsayacaktır. Bu değer kararlı durum ortalama gecikme zamanı (Steady state Average Delay Time) olarak adlandırılır (Frisen ve Andersson, 2009). Bazen salgına karşı önlem alabilmek için zaman kısıtlıdır. Böyle durumlarda Frisen (1992) tarafından önerilmiş bir ölçütte, salgını belirli bir  $d$  zamanından önce başarılı biçimde belirleme olasılığıdır.

$$PSD(d, t) = P[t_A - \tau \leq d | t_A \geq \tau, \tau = t]$$

İzleme sistemlerini değerlendirmek amacıyla kullanılan bir diğer ölçüt ise öngörü değeri (predictive value) ölçütüdür. Bu ölçü değişik zamanlarda verilen alarmların gerçeği yansıtmaya değerini ölçmektedir.  $\tau$  bir rastgele değişken olarak dikkate alındığında, Frisen (1992) öngörü değeri ölçütünü aşağıdaki gibi tanımlamıştır (Frisen ve Andersson, 2009).

$$PV(t) = P(C | t_A = t) = \frac{\sum_{i=1}^t P(t_A = t | \tau = i) \cdot P(\tau = i)}{\sum_{i=1}^t P(t_A = t | \tau = i) \cdot P(\tau = i) + P(t_A = t | \tau > t) P(\tau > t)}$$

Bu ölçü alarm zamanının bir fonksiyonudur ve bir alarm verildiğinde sistemin kontrol dışında olma olasılığını vermektedir. Başka bir ifadeyle belirli bir zaman noktasındaki alarmın ne kadar güven duyulması gerektiğini belirtmektedir (Sonesson and Bock, 2003).

## 2.9. Algoritmaların Değerlendirilmesinde Kullanılan Veriler

Algoritmaların değerlendirilmesinde kullanılan verilerin bilinen özelliklere sahip salgın veya salgınlar içermeleri gerektiği bildirilmektedir. Bu tür verilere ulaşmak için tercih edilebilecek yaklaşımlar genel olarak 3 sınıfta toplanmaktadır. Bunlar:

- Gerçek salgınlara sahip gerçek verilerin kullanılması
- Üzerinde yapay salgınlar oluşturulmuş gerçek verilerin kullanılması
- Yapay salgınlara sahip yapay verilerin kullanılması

Gerçek salgınlara sahip gerçek verilerin kullanılması en iyi seçenek gibi gözükmemektedir. Ancak bu yaklaşım bu tür veri setlerinin az olması, bu veri setlerindeki salgınların kesinlikle belirlenemiyor olması, salgınların başlangıcının net bir şekilde belirlenememesi ve gözlemlenen vaka sayılarının ne kadarının salgınlardan ne kadarının doğal olarak gerçekleştiğinin bilinmemesi gibi dezavantajlara sahiptir (Buckeridge vd., 2007).

Bu güçlüklerden dolayı, algoritmaların performanslarının yapay salgınların enjekte edildiği gerçek verilerle ölçüldüğü belirtilmektedir. Bu yaklaşımla salgın olmadığı durumda ve salgın olduğu durumda özellikle salgının başlangıcı, büyüklüğü gibi salgını belirleyen parametreler kolaylıkla değiştirilerek ilgili algoritmanın çeşitli durumlarda duyarlılık, belirlilik ve dakiklik özelliklerinin nasıl değiştiği belirlenebilmektedir. Bu yaklaşımın bahsedilen esnekliğine rağmen kullanılan gerçek verilerde belirlenememiş doğal salgınlar mevcut olabilir ve enjekte edilen yapay salgın gerçek bir salgını yeterli bir şekilde temsil etmeyebilir (Buckeridge vd., 2007).

Yapay verilere yapay salgınlar enjekte edilmesi ise bir başka yaklaşımdır. Bu yaklaşımda hem veriler hem de salgınların özellikleri belirlenebildiğinden, bu özellikler değiştirilerek çok çeşitli durumlarda algoritmaların genel performansları, sağlamlıkları ve zaman içindeki performansları (örneğin; duyarlılık, belirlilik gibi) ölçülebilmektedir. Bu avantajlara rağmen yapay verilerin modellenmesindeki güçlük ve yapay olarak oluşturulan verilerin gerçekte karşılaşılan verileri yeterli bir şekilde temsil etmeyeceği ihtimali bu yaklaşımın dezavantajlarıdır (Buckeridge vd., 2007).

### **2.9.1 Yapay salgın üretme yaklaşımları**

Yapay salgınları üretmek için çeşitli yaklaşımlar kullanılmaktadır. Bu yaklaşımlar daha çok zaman serisi verileri içeren performans değerlendirmelerinde görülmektedir. Bunlar matematiksel fonksiyonlar, deneysel dağılım fonksiyonları ve mekanistik modellerdir.

Matematiksel fonksiyonlar, örneğin; basamak biçimli, doğrusal fonksiyonlar ya da üstel ve log-normal gibi olasılık fonksiyonları, zaman içinde ilgilenilen hastalığa

ilişkin vakalar üretmek için kullanılmaktadır. Bu yaklaşım, ilgilenilen birçok sorun ve salgın durumu için kullanılabilir basit bir yöntemdir. Özellikle log-normal fonksiyonu birçok bulaşıcı hastalığın belirtilerinin ortaya çıkması için gerekli kuluçka dönemi için iyi bir modeldir. Ancak matematiksel fonksiyonlarla üretilebilecek farklı durumlar sınırlıdır ve bu nedenle gerçek bir salgın durumu ile ilişkilendirilmesi de kolay değildir. Dolayısıyla matematiksel fonksiyonları kullanarak yapılan bir algoritma değerlendirme çalışmasının gerçek yaşama uygulanması uygun olmayabilir (Buckeridge vd., 2007).

Bir başka yaklaşımda, gerçek salgın vakalarından yararlanılarak deneysel yoğunluk fonksiyonu kestirilmekte ve bu fonksiyon yardımıyla da yapay salgınlar oluşturulmaktadır. Deneysel yoğunluk fonksiyonu, gözlemlenen salgından elde edilen orijinal haliyle doğrudan kullanılabilir gibi düzleştirme (smoothing) ya da ölçekleme işlemine tabi tutulduktan sonra da kullanılabilir. Bu fonksiyon geçmişteki veriler kullanılarak bir paket programı aracılığıyla elde edilebilir. Ancak bu yaklaşımla elde edilen salgınların daha gerçekçi olması beklenmesine rağmen, oluşturulan salgınlar performans değerlendirmesi için amaçlanmış salgın durumlarını yansıtmayabilir (Buckeridge vd., 2007).

Salgınlar üretmek için kullanılan bir başka yaklaşım ise mekanistik modellerdir. Bu modellerin mekanistik olarak adlandırılmasının nedeni salgının temelinde olan mekanizmayı tanımlamalarından dolayıdır. Dolayısıyla bu modeller hastalık, enfeksiyon ve sağlık hizmetleri kullanımı gibi unsurları içermektedir. İlgili modellerde kişiler ya ağ tabanlı bir çerçevede ya da bağımsız olasılıksal süreçler olarak modellenebilmektedir. Ağ tabanlı modelleme çok sayıda parametre içermekte ve bulaşıcı hastalıklar için kullanılabilir. Bağımsız olasılıksal süreçler şeklindeki modelleme ise daha az sayıda parametre içermekte ve bulaşıcı olmayan hastalıklar için uygun olmaktadır. Ancak bulaşıcı hastalıklar durumuna da genelleştirilebilmektedir. Hangi model kullanılırsa kullanılsın, parametre değerleri sonuçları değiştirebilmektedir. Bu nedenle sonuçların parametrelerin değişimine olan duyarlılığının da analiz edilmesi gerekmektedir (Buckeridge vd., 2007).

## 2.9.2 Yapay bioizleme verilerinin üretilmesi

Yapay bioizleme verileri üretilebilir. Ancak üretilen veriler gerçek verilerin mevsimsellik ya da haftanın günü etkisi gibi sahip olduğu sistematik yapıları da içermelidir. Bu tür bir veri üretebilmek amacıyla, ilk olarak verilerin zamana bağlı olarak değişen beklenen değerinin yapısı belirlenmelidir. Örneğin mevsimsel değişimleri içeren bir beklenen değer aşağıdaki gibi tanımlanabilir.

$$\mu(t) = \mu_0 \left( 1 + A \cos \left[ \frac{2\pi(t - t_0)}{T} \right] \right)$$

Burada,  $\mu_0$  zaman serisinin uzun dönemli ortalamasını,  $A$  ( $\leq 1$ ) mevsimsel değişimin genliğini,  $t_0$  mevsimsel değişimin tepe noktalarından birisini ve  $T$  ise mevsimsel değişimin tekrarlandığı zaman periyodunu göstermektedir. Bu beklenen değere sahip bir zaman serisi üretebilmek amacıyla, serinin her elemanı için uygun bir istatistiksel dağılımdan rastgele bir gözlem seçilir. Kullanılacak dağılım ise  $\mu_t$  beklenen değerine ve  $k = \sigma_t^2/\mu_t$  oranına bağlıdır. Eğer  $k = 1$  ise  $\mu_t$  ortalamalı Poisson dağılımı kullanılmaktadır. Varyansın ortalamaya oranı  $k < 1$  ise,  $n = \mu_t/(1 - k)$  ve  $p = 1 - k$  ile binom dağılımı kullanılmaktadır. İlgili oran  $k > 1$  ise, gözlemlenen başarı sayısını üretmek için  $r = \mu_t/(k - 1)$  ve  $p = 1/k$  ile negatif binom dağılımı kullanılmaktadır. Haftanın günü ve tatil günü etkisinden kaynaklanan değişim ise zaman serisi elemanları arasında vakaların kaydırılmasıyla sağlanır (Buckeridge vd., 2007).

### 3. BAZI BİYOİZLEME YÖNTEMLERİ

Salgınları belirlemek amaçlı yöntemler genel olarak ileriye ve geriye dönük izleme yapan yöntemler olarak iki sınıfa ayrılabilir. Bu tez çalışması kapsamında ileriye dönük izleme yapan yöntemler dikkate alındığından, bu bölümde bu yöntemlerden bazıları tanıtılacaktır. Tanıtılacak yöntemler arasında regresyon modellemesine, zaman serileri modellemesine, istatistiksel süreç kontrol grafiklerine dayanan yöntemler vardır.

#### 3.1. Regresyon Modellemesine Dayanan Yaklaşımlar

Bu yaklaşımlar laboratuvar ve doğrulanmış vakalara dayalı izleme sistemlerinde ya da sendromik izleme sistemlerinde yaygın biçimde kullanılmaktadır. Regresyon yöntemleri, Shewhart kontrol grafiklerinin genelleşmiş hali olarak görülebilir. Kontrol altındaki süreç ortalamasının ve muhtemelen süreç standart sapmasının zaman içinde değişebilmesine imkan sağlamaktadır. Ayrıca, sürece ilişkin gözlemlerin dağılımı normal dağılım olmak zorunda değildir. Regresyonu temel alan salgın belirleme yöntemlerine ilişkin performansın, Shewhart kontrol grafiklerinin performansını yansıtması beklenir (Unkel vd., 2012).

Salgın belirleme amacıyla regresyon tekniklerinin kullanımı basittir ve otomasyonda kolaydır. Ancak bu yöntemler;  $t$  zaman noktasından itibaren kullanılmaya başlanması ve büyük bir salgının bu  $t$  zaman noktasında başlaması şartıyla, salgının belirlenmesinde etkindirler. Buna karşın  $t$  zaman noktasından bir müddet sonra başlayan ve daha yavaş gelişen salgınları belirlemede daha az etkindirler (Unkel vd., 2012). Ayrıca regresyon yöntemlerinin bir eksikliğide verilerin durağan olduğunu varsayımları ve özilişkiyi dikkate almamalarıdır.

Regresyon temelli yöntemlerin çoğu  $t$  zaman noktasındaki  $\mu_t$  ortalaması için bir model belirlemektedir. Bir zaman noktasındaki gözlemlenen değer, ortalama ve varyans kullanılarak belirlenen eşik değerinden büyük olduğunda, salgın sinyali verilmektedir. Bu amaçla yaygın bir yaklaşım hataların normal dağıldığını kabul ederek,  $\mu_t$ 'yi ve  $\sigma_t^2$ 'yi tahmin etmek ve eşik değerini  $u_t = \hat{\mu}_t + k\hat{\sigma}_t$  olarak belirlemektir. Böyle bir yaklaşımdan daha doğru bir yaklaşım, eşik değerini  $t$  zamanındaki  $Y_t$  gözleminin öngörü dağılımının üst yüzdeliklerinden olarak belirlemektir. Böylelikle hem süreç varyansını hem de süreç ortalamasını (referans düzeyini) kestirmekten kaynaklanacak değişim dikkate alınır (Unkel vd., 2012).

### 3.1.1 Serfling parametrik regresyon modeli

Yaygın olarak kullanılan salgın belirleme amaçlı parametrik bir regresyon modeli Serfling (1963) tarafından kullanılmıştır. Bu model Eş. 3.1' de verilmektedir.

$$y_t = \mu_t + \varepsilon_t, \quad \varepsilon_t \sim N(0, \sigma^2) \quad (3.1)$$

Burada  $\mu_t$  aşağıdaki gibi tanımlanmaktadır.

$$\mu_t = E(y_t) = \mu + at + \sum_{i=1}^r \{\beta_i \sin(\omega_i t) + \gamma_i \cos(\omega_i t)\}$$

Serfling (1963) regresyon modeli grip salgınının başlangıcını belirlemek amacıyla kullanılmıştır. Bu yöntemi uygulayabilmek adına, ilk önce “eğitim dönemi” olarak adlandırılan geçmişe ait bir dönem dikkate alınır. Daha sonra geçmişteki salgınların etkisini gidermek amacıyla, bu döneme ilişkin en yüksek değerli verilerden başlamak kaydıyla, verilerin %20'si dışlanır. Geri kalan verilere Serfling (1963) modeli uydurulur. Bu modele ilişkin öngörü dağılımın (prediction distribution) üst yüzdelikleri örneğin 90. yüzdeliği eşik değeri olarak alınır, bu değeri aşan gelecekteki gözlemler için sistem salgın sinyali verir (Unkel vd., 2012).

Veriler seyrek olduğunda yani nadir gerçekleşen olayların sayısını gösterdiğinde, hataların normal dağılımlı olduğunu varsayan regresyon modellerinin kullanması uygun değildir. Bu durumlarda, Poisson log doğrusal model ya da iki durumlu lojistik regresyon model gibi Genelleştirilmiş Doğrusal Modeller (GLM) kullanılmakta ve bu modellerin sağladığı esneklikten yararlanılmaktadır. Örneğin Farrington vd. (1996) salgının izlenmesi amacıyla genelleştirilmiş doğrusal bir model kullanımını önermişlerdir (Unkel vd., 2012).

Bioizleme verilerinin regresyonuna ilişkin başka bir yaklaşımda, her öngörü değeri için katsayıların sürekli güncellendiği uyarlamalı regresyon yaklaşımını kullanmaktır. Bu yaklaşımda regresyon modeli, parametrelerin güncellenmesi gereken kayan bir referans dönemi kullanmaktadır. Bu referans döneminin kısa olması halinde Serfling regresyonunda olduğu gibi mevsimsellik terimleri yer alamamaktadır. Bu tür regresyon modelleri geçmişe dönük büyük miktarda referans verisi olmadığında ve/veya geçmişteki referans verisi günümüzdeki bioizleme verilerini yeterince temsil etmediği durumlarda uygun bir yaklaşımdır (Burkom, 2007).

### **3.1.2 Farrington ve FarringtonFlexible Yöntemi**

Farrington yöntemi (Farrington vd., 1996) ve FarringtonFlexible yöntemi (Noufaily vd., 2013) tek nokta bazlı olağandışı durum belirleme yöntemleridir. Başka bir ifadeyle, sadece bulunan zaman noktasındaki bilgiye dayanılarak kararlar alınmaktadır. Farrington yöntemi Avrupa'daki halk sağlığı enstitülerinde kullanılan bir yöntemdir (Salmon vd., 2016).

Bu yöntemler herhangi bir  $t_0$  anında karar verirken, ilgili noktaya özel referans değerleri oluşturur. Farrington yönteminde  $t_0$  noktasından  $b$  yıl kadar geriye gidilir ve geçmişte  $t_0$ 'ın bulunduğu noktanın civarındaki  $w$  adet zaman noktası referans değerlerinin kümesine katılır. Bu küme toplamda  $b(2w + 1)$  kadar veriyi referans olarak içermektedir ve matematiksel olarak aşağıdaki gibi temsil edilir (Manitz ve Höhle, 2013; Salmon vd., 2016)



$$R(w, b, s) = \left( \bigcup_{i=1}^b \bigcup_{j=-w}^w y_{t_0-is+j} \right)$$

Burada  $s$  bir yılın kaç zaman noktasında oluştuğunu göstermektedir. Örneğin veriler haftalık ise  $s = 52$ , aylık ise  $s = 12$ 'dir. Böylelikle verilerdeki mevsimsellik etkisi dikkate alınmaktadır. Bu model mevsimsel etkiyi tasarımı gereği dikkate almakta ve doğrusal trend açık bir şekilde modelde yer almaktadır (Unkel vd.,2012). FarringtonFlexible yöntemi de aynı referans kümesini kullanmaktadır. Ancak yavaş gelişen salgınlara önlem almak amacıyla Farrington yöntemi genellikle  $t_0$ 'dan önce 3 gözlemi, FarringtonFlexible yönteminde de 26 gözlemin dikkate alınmaması tavsiye edilmektedir. Bu referans değerlerinden oluşan küme kullanılarak Çizelge 3.1.'deki geliştirilmiş doğrusal modeller kestirilir (Salmon vd., 2016).

**Çizelge 3.1. Farrington ve FarringtonFlexible yöntemlerine ilişkin geliştirilmiş doğrusal modeller**

<b>Farrington Yöntemi için GLM</b>	<b>FarringtonFlexible Yöntemi için GLM</b>
$y_t   \mu_t \sim \text{Poisson}(\mu_t) \quad y_t \in R(b, w, s)$	$y_t   \mu_t \sim \text{NegatifBinom}(\mu_t, v) \quad y_t \in R(b, w, s)$
$E(y_t) = \mu_t$	$E(y_t) = \mu_t$
$\log(\mu_t) = \eta_t = \alpha + \beta t$	$\log(\mu_t) = \eta_t = \alpha + \beta t + \gamma_{c(t)}$
$\text{Var}(y_t) = \phi \mu_t$	$\text{Var}(y_t) = \phi \mu_t, \quad v = \frac{\mu_t}{\phi - 1}$

Çizelge 3.1.'deki  $\phi$  yayılım parametresidir ve  $\gamma_{c(t)}$  10 düzeyli bir faktör değişkenine ifade edilecek sıfırıncı dereceden 11 düğümlü eğriyi (spline) göstermektedir. Başka bir ifadeyle FarringtonFlexible yöntemi mevsimselliği açıklamak üzere, modele spline dahil etmektedir yani bir yıl 10 parçaya bölünmektedir. Ayrıca modellerdeki trend etkisi belirli koşullar altında dikkate alınmaktadır. Örneğin istenilen anlamlılık düzeyinde  $\beta$ 'nin anlamlı olması, referans kümesinin en az 3 yıllık verilerden oluşması gibi.

Modeller kestirildikten sonra; geçmişteki salgınlara etkisini gidermek amacıyla, büyük Anscombe artıklarına ( $s_i$ ) sahip referans değerleri yeniden ağırlıklandırılarak (düşük ağırlıklar verilerek) model tekrar kestirilir. Farrington yönteminde artıklar

1'den büyükse, FarringtonFlexible yönteminde artıklar 2.58'den büyükse ağırlıklandırma ( $\omega_i$ ) yapılır (Salmon vd., 2016).

$$s_i = \frac{3}{2\hat{\phi}^{1/2}} \frac{y_i^{2/3} - \hat{\mu}_i^{2/3}}{\hat{\mu}_i^6 (1 - h_{ii})^{1/2}} \quad \omega_i = \begin{cases} \gamma s_i^{-2} & \text{ise } s_i > 1 \\ \gamma & \text{d. durumlar} \end{cases}$$

Burada  $\phi$  yayılım parametresi aşağıdaki gibi tahmin edilmektedir.

$$\hat{\phi} = \max \left\{ \frac{1}{n-p} \sum_{i=1}^n \omega_i \frac{(y_i - \hat{\mu}_i)^2}{\hat{\mu}_i}, 1 \right\}$$

Bu işlemlerden sonra  $t_0$  noktasındaki beklenen değer  $\mu_{t_0}$  kestirilir. Bunu takiben Farrington yöntemi için tahmin edilen bu değer ve onun varyansına dayanarak ilgili nokta için eşik değeri  $U_{t_0} \%100(1 - \alpha)$  düzeyindeki iki yanlı tahmin aralığının üst sınırı olarak belirlenir.

$$U_{t_0} = \hat{\mu}_0 + z_{(1-\frac{\alpha}{2})} \sqrt{\widehat{Var}(y_{t_0} - \hat{\mu}_{t_0})}$$

Eşik değerinin bu şekilde hesaplanması tahmin hatasının ( $y_{t_0} - \hat{\mu}_0$ ) normal dağılması varsayımına dayanmaktadır. Normallik varsayımının sağlanabilmesi için bazı durumlarda verilere güç dönüşümü (örneğin Farrington yöntemi için sıklıkla verilerin 2/3 kuvveti alınır) uygulanır. Bazen de varyansın sabit (stabilized) hale getirilmesi için güç dönüşümü yapılır. FarringtonFlexible yöntemi için eşik değeri  $U_{t_0}$  ise NegatifBinom ( $\hat{\mu}_{t_0}, \hat{v}$ ) dağılımının  $(1 - \alpha)$ . kuantili olarak alınır. Bu yaklaşımın bir dezavantajı ilgili kuantili bulurken gerçek parametre değerleri yerine onların kestirimlerinin kullanılmasıdır (Salmon vd., 2016; Noufaily vd., 2013).

Son olarak karar verirken  $y_{t_0} > U_{t_0}$  ise, olağandışı durumu sinyali verilir. Bu modelin bir dezavantajı ise zaman serisinin gözlemlerinin bağımsız olduğunu varsaymasıdır (Amoros vd., 2015).

### 3.1.3 Bayesci Salgın Belirleme Yöntemi (Boda)

Bu yöntem bayesci düşünceye sahip bir izleme algoritması olup Farrington vd. (1996) ve Heisterkamp vd. (2006) yöntemlerinin genişletilmiş durumudur. Bu yaklaşım hem öngörüdeki hem de parametre tahminindeki belirsizliğin tüm kaynaklarını dikkate alarak eşik değerlerinin doğrudan hesaplanmasını kolaylaştırmaktadır. Daha da önemlisi salgın belirleme işlemine açıklayıcı değişkenlerin etkisini dahil etmek mümkün olmaktadır (Manitz ve Höhle, 2013).

Bu algoritmanın temeli bir bayesci hiyerarşik modeldir ve bu modelde bildirilen vakaların sayılarının  $(y_t, t = 1, 2, \dots, T)$  aşağıdaki ortalama ve varyansla negatif binom dağılımlı olduğu varsayılır.

$$E(y_t|\mu_t) = \mu_t \quad Var(y_t|\mu_t, v) = \mu_t + \mu_t^2/v$$

Burada  $v > 0$  parametresi aşırı yayılım parametresidir ve  $v \rightarrow \infty$  için negatif binom dağılımı, ortalaması ve varyansı eşit olan Poisson dağılımına yakınsamaktadır. Vaka sayılarının ortalamasını gösteren  $\mu_t$  bir log-bağ(log-link) kullanılarak modellenmektedir. Bu model Eş.3.2’de belirtilmektedir.

$$y_t|\mu_t, v \sim NB(\mu_t, v) \quad (3.2)$$

Bu modele ilişkin  $\mu_t$ 'nin Eş.3.3’ü sağladığı kabul edilmektedir.

$$\log(\mu_t) = \eta_t = a_{0t} + \beta t + \gamma_t + x_t' \delta_x + \zeta z_t, \quad t = 1, \dots, T \quad (3.3)$$

Eş.3.2 ve Eş.3.3 ile ifade edilen modelin parametreleri bir vektör ile temsil edilebilir:  $\theta_t = (a_{01}, \dots, a_{0T}, \beta, \gamma_1, \dots, \gamma_T, \delta_x', \zeta, v)'$ . Bu modelde  $a_{0t}$  zamana bağlı kesim noktasını,  $\beta$  trend etkisini,  $\gamma_t$  periyodikliği  $s$  olan mevsimsel etkiyi,  $\delta_x'$  açıklayıcı değişken etkisini ve  $\zeta$  geçmişteki salgınların etkisini düzeltmek amaçlı değeri göstermektedir. Ayrıca  $z_t$ , geçmişteki olağandışı durum olup olmadığını 0 ya da 1 değerleri alarak gösteren bir değişkendir. Bu model  $T$  zamanındaki vaka sayısının olağandışı olup olmadığına karar vermek için şu şekilde kullanılır. İlk olarak  $T - 1$  anına kadar olan ( $T - 1$  dahil) tüm veriler model parametrelerinin sonsal dağılımını

elde etmek için kullanılır. Bunu takiben  $y_T$ 'nin öngörü sonsal dağılımı, parametrelerin sonsal dağılımı kullanılarak hesaplanır. Daha sonra gözlemlenen gerçek  $y_T$  değeri,  $y_T$ 'nin öngörü sonsal dağılımının önceden belirlenmiş bir kuantili yani  $\%100(1 - \alpha)$  yüzdeliği ile karşılaştırılır. Bu karşılaştırma sonucu  $y_T$ , bu kuantili yani eşik değerini geçerse olağandışı durum sinyali verilir (Manitz ve Höhle, 2013).

Bu yönteme ilişkin önsel dağılımların, sonsal dağılımların belirlenmesi ve eşik değerlerinin elde edilmesi ayrıntılı olarak Manitz ve Höhle (2013) verilmektedir. Burada kısaca bu konulara değinilecektir. Bunun için ilk olarak negatif binom dağılımındaki büyüklük parametresi  $v$ 'nin bilgilendirici olmayan log-gamma dağılımına sahip kabul edilir. Genellikle  $v \sim \log \text{gamma}(1,100)$  alınır. İkinci olarak 5 kısım için önsel dağılımların belirlenmesi yapılır. Bu kısımlar kesim noktası, trend, mevsimsellik parametreleri, açıklayıcı değişkenlere ilişkin parametreler ve geçmişteki salgın etkilerinden oluşan kısımlardır. Böyle bir ayrışım zaman serilerinin yaygın olan ayrışım yöntemine karşılık gelmektedir (Manitz ve Höhle, 2013).

Çizelge 3.2.'de ilgili zaman serisinin 5 kısmı için önsel dağılımlar belirtilmektedir. Zamana bağlı olarak değişen kesim noktası  $\alpha_{0t}$ ,  $\lambda_\alpha$  hassaslığına sahip bir rasgele yürüyüş olarak belirlenmektedir. Buradaki  $m$ , Çizelge 3.2.'de belirtilen üç farklı modeli göstermektedir. Hangi model seçilirse seçilsin, normal dağılımlı bir hata modele dahil edilir.

**Çizelge 3.2. Boda yöntemindeki model kısımlarına ilişkin önsel dağılımlar**

Modelin Kısımları	Önsel Dağılım	
Kesim noktası	$\alpha_{0t} \sim N(g_m(t), \lambda_\alpha^{-1})$	Durağan model: $g_{iid}(t) = \alpha_0$ Komşuluk modeli: $g_{rwl}(t) = \alpha_{0,t-1}$ Doğrusal model: $g_{rw2}(t) = 2\alpha_{0,t-1} - \alpha_{0,t-2}$
Trend		$\beta \sim N(0, \lambda_\beta^{-1})$
Mevsimsellik		$\sum_{i=1}^{s-1} \gamma_{t+i} \sim N(0, \lambda_\gamma^{-1})$
Açıklayıcı değişken		$\delta_x \sim N(0, \Lambda_\delta^{-1})$
Geçmişteki salgınlar		$\zeta \sim N(0, \lambda_\zeta^{-1})$

Çizelge 3.2.'deki trend etkisinin bilgilendirici olmayan önsel dağılımla logaritma ölçeğinde doğrusal olduğu; mevsimsellik etkisinin ise  $s$  periyoduna sahip olduğu;

açıklayıcı değişken etkisinde düz bir önsel dağılımla belirlendiği ve geçmiş salgınlarmın etkisinin bilgilendirici olmayan bir önsel dağılımla belirlendiği görülmektedir (Manitz ve Höhle, 2013).

Bu etkilerin bağımsız olduğu varsayımı altında birleştirilmesi sonucu, önsel dağılım modeli aşağıdaki gibi orantılı olarak ifade edilir.

$$p(\theta_T) \propto p(\alpha_{01}, \dots, \alpha_{0T} | \lambda_\alpha, R_m) \cdot p(\beta | \lambda_\beta) \cdot p(\gamma_1, \dots, \gamma_T | \lambda_\gamma, R_s) \cdot p(\delta_x | \Lambda_\delta) \cdot p(\zeta | \lambda_\zeta) \cdot p(v)$$

Son olarak Çizelge 3.2.'deki önsel dağılımlara ilişkin varyansların tersleri yani duyarlılık için önsel dağılımlar genellikle log-gamma dağılımları olarak seçilir. Önsel dağılımlar belirlendikten sonra, bağımsız ve negatif binom dağılımlı kabul edilen gözlemlerin olabilirliği elde edilir.

$$p(y_1, \dots, y_{T-1} | \theta_T) = \prod_{t=1}^{T-1} f(y_t | \theta_T)$$

Parametrelere ilişkin önsel dağılımlar ve gözlemlenen verilerin olabilirliği kullanılarak  $y_T$  gözleminin ve  $\theta_T$  parametre vektörünün birleşik sonsal dağılımı hesaplanır. Bu birleşik dağılım daha sonra  $\theta_T$  vektörüne göre integrallendiğinde,  $y_T$  gözlemi için  $y_1, \dots, y_{T-1}$  verildiğinde sonsal dağılım yani öngörü sonsal dağılımı elde edilir.

$$\begin{aligned} p(y_T | y_1, \dots, y_{T-1}) &= \int f(y_T, \theta_T | y_1, \dots, y_{T-1}) d\theta_T \\ &= \int f(y_T | \theta_T) p(\theta_T | y_1, \dots, y_{T-1}) d\theta_T \end{aligned}$$

Bu öngörü sonsal dağılımının  $\%100(1 - \alpha)$ . kuantili eşik değeri olarak alınır. Eşik değerinin hesaplanmasında INLA olarak bilinen bütünleştirilmiş iç içe Laplace yaklaşımı kullanılır. Bu yönteminin seçilmesinin nedeni Markov Chain Monte Carlo (MCMC) yöntemlerinden daha hızlı olmasıdır. Son olarak  $y_T$  gözlemi bu eşik değerini geçtiğinde, olağandışı durum sinyali verilir (Manitz ve Höhle, 2013).

Bu algoritma ayrıntılı bir modellemeyi gerektirmesinin yanı sıra algoritmanın hızında düşük olması bunun rutin uygulamalarındaki kullanımını sınırlamaktadır. Salmon, Schumacher, Stark ve Höhle bu algoritmanın daha hızlı bir çeşidini önermişlerdir. Ancak önerilen yeni algoritma açıklayıcı değişkenlere ve kesim noktasının zamana bağlı olmasına izin vermemektedir (Salmon vd., 2016).

### 3.1.4 Bayeşçi Bir Öngörü Sonsal Yaklaşımı

Riebler (2004) bioizlemede kullanılacak bir bayeşçi yaklaşım sunmuştur. Bu yaklaşımda verilerin  $\lambda$  parametresiyle bağımsız ve aynı Poisson dağılımlı varsayılmaktadır. Ayrıca  $\lambda$ 'nın da gamma dağılımlı bir önsel dağılıma sahip olduğu kabul edilir. Başka bir ifadeyle,

$$\lambda \sim \text{Gamma}(\alpha, \beta) \quad \text{ve} \quad y_i | \lambda \sim \text{Poisson}(\lambda), i = 1, \dots, n.$$

Bu yöntemde herhangi bir  $t$  anında kararın verilebilmesi için kullanılan referans verileri kümesi  $R_{BAYES} = R(w, w_0, b) = \{y_1, \dots, y_n\}$  olarak tanımlanır. Ayrıca  $y_{0:t}$  ise  $t$  zaman noktasındaki gözlemlenen ve hakkında olağan dışı olup olmadığına karar verilecek değerdir. Burada  $b$ , geçmişe ait kaç yılın dikkate alınacağını;  $w$ , geçmişe ait ilgili zaman noktasını içeren pencere genişliğini;  $w_0$ , içinde bulunan yılda hakkında karar verilecek  $y_t$  noktasından önce dikkate alınacak gözlem sayısını göstermektedir. Matematiksel işlemler sonucu  $\lambda$  parametresine ilişkin sonsal dağılım yine gamma dağılımı olarak elde edilir (Höhle, 2007).

$$\lambda | y_1, \dots, y_n \sim \text{Gamma} \left( \alpha + \sum_{i=1}^n y_i, \beta + n \right)$$

Daha sonra bir sonraki gözlem  $y_{n+1}$  için öngörü sonsal dağılımı elde edilir.

$$f(y_{n+1} | y_1, \dots, y_n) = \int_0^{\infty} f(y_{n+1} | \lambda) f(\lambda | y_1, \dots, y_n) d\lambda$$

Bu işlem sonucu Poisson-gamma dağılımı yani negatif binom dağılımının genelleştirilmiş bir biçimine ulaşılır.

$$y_{n+1} | y_1, \dots, y_n \sim \text{NegatifBinom} \left( \alpha + \sum_{i=1}^n y_i, \frac{\beta + n}{\beta + n + 1} \right)$$

$\lambda$  parametresi için Gamma (0.5, 0) dağılımı önsel dağılım olarak kullanıldığında, negatif binom dağılım parametreleri aşağıdaki gibi elde edilir (Höhle, 2007).

$$y_{n+1} | y_1, \dots, y_n \sim \text{NegatifBinom} \left( 0.5 + \sum_{y_{i,j} \in R_{BAYES}} y_{i,j}, \frac{|R_{BAYES}|}{|R_{BAYES}| + 1} \right)$$

Bu dağılımın üst  $\alpha$  yüzdeliği yani  $P(y_{n+1} \leq y_\alpha | y_1, \dots, y_n) \geq 1 - \alpha$  eşitliğini sağlayan  $y_\alpha$  bulunur. Daha sonra  $y_{0:t} \geq y_\alpha$  ise olağandışı bir durum sinyali verilir (Höhle, 2007).

### 3.1.5 Yarı parametrik regresyon modelleri

Bu yöntemlerde bir düzleştirme (smoothing) işlemi kullanılarak bir baz (referans düzeyi) ve standart sapma belirlenir. Hataların normal dağıldığı kabul edilerek, referans düzeyinden 2 standart sapma yukarısındaki gözlemler için sistem salgın sinyali verir. Standart sapma, hatalardan yani gözlemler ile düzleştirilmiş referans değerleri arasındaki farklardan hesaplanır (Unkel vd.,2012).

Bir başka yaklaşımda,  $\mu_t$  genelleştirilmiş bir toplamsal modelle (GTM) tahmin edilir. Bu modelden elde edilen artıklara başka bir GTM uygulanarak  $\sigma_t^2$  tahmin edilir. Daha sonra tercih edilen bir  $k$  değeri kullanılarak, eşik değeri  $\hat{\mu}_t + k\hat{\sigma}_t$  olarak belirlenir. Kullanılan GTM'lerdeki düzleştirme terimleri Gaussian çekirdek düzleştiricilerini temel almakta ve bant genişliği de öngörü hata kareler ortalamasını minimum yapacak biçimde seçilmektedir (Unkel vd.,2012).

Başka bir alternatif olarak (Zhang vd., 2003) referans değerlerinin ‘wavelet dönüşümünü’ kullanmayı önermişlerdir. Böylelikle trend ve mevsimsellik gibi etkiler dikkate alınırken, geçmişteki salgınlardan veya yapay olaylardan, örneğin tatillerden gelen yüksek frekansa karşı sağlamlık kazanılmış olmaktadır. Bu yaklaşımda, wavelet dönüşümü aracılığıyla elde edilen referans değerleri verilerden çıkarılarak artıklar elde edilir ve artıkların dağılımı dikkate alınarak eşik değerleri oluşturulur (Unkel vd.,2012).

### **3.1.6 Haar-wavelet Dönüşümler Yöntemi**

Percival ve Walden (2006) wavelets ayrışımını, bir zaman serisini belirli bir zaman ve frekans ile ilişkili katsayılar cinsinden temsil etme yöntemi olarak tanımlamıştır (Shumeli, 2013). Wavelets ayrışımı yaygın bir şekilde sinyal işleme işlemlerinde, sinyallerdeki değişimi (gürültüyü) gidererek sinyaldeki asıl bilgiyi ortaya çıkarmakta kullanılmaktadır. Bu yöntem zaman serilerindeki durağan olmama özelliğini ve otokorelasyon yapısını dikkate almaktadır. Bu nedenle, uzun dönemli ilişki yapısına sahip zaman serileri için örneğin nehir debileri, atmosferik olaylar, iletişim, astronomi ve finansal piyasalar kaynaklı zaman serilerinin ayrışımında wavelets yönteminin kullanışlı olduğu belirtilmiştir (Shumeli, 2013).

Ayrıca bu yöntem zaman serilerindeki herhangi bir yapıdaki anormalliği belirleyebilmektedir. Wavelets ayrışımı yöntemi, kalite kontrol grafikleri gibi (süreçteki ani değişimleri, üstel değişimleri ya da basamak şeklindeki değişimleri) belirli bir anormallik yapısını belirlemek amacıyla tasarlanmamışlardır. Wavelets ayrıştırma yönteminin bu özellikleri dikkate alındığında; genellikle durağan olmayan, özilişkili ve değişen anormallik yapılarına sahip zaman serilerinin gözlemlendiği biyoizlemede bu yöntemin uygun bir araç olabileceği belirtilmektedir (Shumeli, 2013).

Biyozlemede zaman serileri verilerini izlemek için kesikli wavelet dönüşümleri kullanılmaktadır. Kesikli bir wavelet dönüşümü,  $n$  uzunluğundaki bir  $X$  zaman serisinin ortogonal bir dönüşümüdür. Bu dönüşümle ilgili zaman serisi  $t$  anında aşağıdaki gibi ifade edilebilir (Shumeli, 2013).



$$X_t = \sum_{j=1}^J D_{j,t} + A_{J,t}$$

Burada  $D_j = W_j' w_j$   $j$ . ölçekte  $X$ 'teki deęişimle iliřkili ayrıntıyı göstermekteyken,  $W$   $n \times n$  boyutlu dönüşümü tanımlayan ortogonal matris ve  $w = WX$  biçiminde tanımlı  $n \times 1$  boyutlu wavelet dönüşümü katsayılar vektörüdür. Ayrıca  $A_{J,t}$  ise örneklem ortalamasını göstermektedir.

Wavelets ayrışım yönteminin en basit hali olan Haar-wavelets işleyiři şu şekilde tanımlanabilir. İlk adımda zaman serisi iki ayrı süzgece(filters) tabi tutulmaktadır. Bu süzgeçlerden biri  $(g = \sqrt{2} \begin{bmatrix} 1 & 1 \\ 2 & 2 \end{bmatrix})$  serideki ardışık gözlemlerin ortalamalarını kullanarak  $cA_1$  olarak gösterilen 1. ölçek yaklaşım katsayılarını; dięer süzgeç  $(h = \sqrt{2} \begin{bmatrix} 1 & -1 \\ 2 & -2 \end{bmatrix})$  ardışık gözlemlerin farklarını kullanarak  $cD_1$  olarak gösterilen 1. ölçek ayrıntı katsayılarını elde etmektedir. Daha sonra bu iki katsayı vektörünün ardışık elemanlarının her ikisinden birisi silinir(downsampling), böylelikle bu vektörlerin gözlem sayısı, serinin gözlem sayısının yarısına düşer ( $n/2$ ). Bir sonraki adımda  $cA_1$  yine bu iki süzgece tabi tutularak  $n/4$  uzunluęunda 2. ölçeęe iliřkin yaklaşım katsayıları vektörü  $cA_2$  ve ayrıntı katsayıları vektörü  $cD_2$  elde edilmektedir. Bu adımlar son ölçek olan  $J$ . ölçek için  $cA_J$  ve  $cD_J$  vektörünü elde edene kadar devam ettirilir. Aslında  $t$  zamanında  $j$ . ölçekteki bir katsayı ( $cD_{j,t}$ ),  $t$  noktasından önce ve sonraki seride bulunan  $2j - 1$  deęerin ortalamalar arasındaki farkı göstermektedir. Bu ayrışımında maksimum ölçek sayısı  $J = \log_2(n)$  olabilmektedir. Ancak uygulamada daha az sayıda ölçek kullanılmaktadır (Shumeli, 2013).

Yukarıda tanımlanan ayrışım sonlandıktan sonra elde edilen katsayı vektörleri kullanılarak seri yeniden oluşturulabilir. Bu amaçla yapılan işlemlerin tersi yapılır. Son ölçeęe iliřkin  $cA_J$  ve  $cD_J$  katsayı vektörlerindeki ardışık elemanları arasına 0'lar yerleřtirilir (upsampling). Daha sonra bu katsayı vektörlerine yukarıda bahsedilen süzgeçlerin tersi uygulanarak sondan bir önceki katsayı vektörleri  $cA_{J-1}$  ve  $cD_{J-1}$  elde edilir. Bu işlemler orijinal serinin yeniden oluşturulmasında kullanılan  $cA_1$  ve

$cD_1$  katsayı vektörleri elde edilinceye kadar tekrarlanır. Seri yeniden oluşturulurken herhangi bir  $j$ . ölçek için elde edilen katsayı vektörleri  $A_j$  ve  $D_j$  olarak gösterilmektedir (Shumeli, 2013).

Wavelets dönüşümler yöntemi geriye ve ileriye dönük izleme amaçlı kullanılabilir. İzleme ilgili serinin ayrışımı yapıldıktan sonra elde edilen ölçeklerin  $(cD_1, cD_2, \dots, cD_j, cA_j)$  ya da seri yeniden oluşturulurken elde edilen ölçeklerin süreç kontrol yöntemleri gibi izleme yöntemleri kullanılarak gerçekleştirilebilmektedir. Bu yöntemin ileriye dönük olarak uygulaması geriye dönük uygulamasına göre daha zordur. Geriye doğru izlemede her ölçekteki katsayılar ilişkisiz (decorrelated) haline gelmektedir.

İleriye dönük izleme uygulaması için bir yaklaşım durağan wavelet ayrışımı (SWT) olarak adlandırılan downsampling (altörnekleme) işlemini uygulanmamasıdır. Bu nedenle bu yaklaşımda herhangi bir ölçek içerisindeki gözlem sayısı orijinal serinin gözlem sayısı ile aynı olmakta ancak ölçek içindeki gözlemler (katsayılar) özilişkili olarak kalmaktadır. Bununla birlikte katsayılar arasındaki bu özilişki yapısının orijinal zaman serisindeki özilişki yapısına kıyasla modellenmesinin daha kolay olduğu gözlemlenmiştir (Shumeli, 2013).

Ayrıca SWT ile elde edilen sonuçlar, kesikli wavelet ayrışımı yöntemi aksine serinin başlangıç noktasına bağlı olarak değişmemektedir. Kesikli bir wavelets dönüşümündeki katsayılar orijinal serinin başlangıcı ve sonunda dış değer bulma işlemlerinden (extrapolations) etkilenmektedir. SWT'nin bir diğer avantajı ise Haar-wavelets ayrışımının kullanılması durumunda katsayıların bu dış değer bulma işlemlerinden minimum şekilde etkilenmesidir (Shumeli, 2013). Bütün bu özellikler düşünüldüğünde, Haar-wavelets tabanlı durağan bir wavelet ayrışımının (sadece ilgili  $t$  zaman noktasının öncesindeki bilgileri kullanacak şekilde) ileriye doğru izleme için uygun bir yöntem olduğu söylenebilir.

Bu yöntemin birinci aşama (salgın olmayan dönem) ve ikinci aşama (izleme dönemi) şeklindeki verilere uygulanması makuldür. Hem ölçek içindeki hem de ölçekler arasındaki ilişkiden dolayı yanlış alarm hızı artmaktadır. Bu problemin etkisini

azaltmak için çeşitli yaklaşımlar yapılmıştır. Örneğin; ölçeklere ilişkin kontrol sınırlarının teorik olmak yerine deneysel olarak belirlenmesi ya da kontrol sınırlarının Arima modellemesi ile belirlenmesidir. Ayrıca yanlış keşif hızı düzeltmesi (FDR correction) yapılması da çoklu test yapmanın etkisini azaltmak amacıyla önerilmiştir (Shumeli, 2013).

### 3.2. Zaman Serisi Modellemesine Dayanan Yaklaşımlar

Salgın izleme verileri genellikle öz ilişkili olmakta, trend ve mevsimsellik etkileri gösterebilmektedir. Bu nedenle zaman serisi yöntemleri izleme sistemlerinde kullanılmaktadır. Trend ve mevsimsellik etkilerini kestirmek için genellikle Serfling (1963) modeli ya da Holt-Winters düzleştirme yöntemleri kullanılmaktadır (Unkel vd., 2012).

Arima ( $p, d, q$ ) ve Inar ( $p$ ) (Integer Valued Auto Regressive) modelleri bulaşıcı hastalıkların salgınlarını belirlemek için kullanılmıştır. Örneğin Cardinal vd. (1999), Kanada'nın Montreal bölgesindeki menenjit vakaları için Inar (5) ve Ar (5) modelleri kullanmışlardır. Wagner (2010) Scotland' taki kusmayla ilgili olarak sağlık merkezine olan çağrılarını, mevsimsellik etkisi giderildikten sonra Ar (6) ile modellemiştir. Reis ve Mandl (2003), 1992'den itibaren 10 yıl boyunca Boston bölgesindeki bir hastanenin acil servisine olan toplam ziyaret sayısını Arma (2,1), solunum şikayeti nedeniyle olan ziyaret sayısını Arma (1,1) ile modellemiştir. Miller vd. (2003) hafta sonları ve tatillerin gösterge değişkenle ifade edildiği Serfling (1963) modelini, Ar (6) ile birlikte kullanarak grip benzeri belirtilerle ayakta tedavi merkezine olan ziyaretleri modellemiştir (Unkel vd., 2012).

Zaman serisi modellemesi yaklaşımının bazı dezavantajları söz konusudur. Öncelikle çok sayıda seri izlenmek zorunda olduğunda, zaman serisi yöntemleri otomasyona elverişli değildir çünkü her bir seri için ayrı ayrı model bulunması gereklidir. Bazen de durağan hale getirilememesi ya da veri seyrekliğinden model kurulamamaktadır. Ayrıca, modelleme ve öngörü için uzun seriler gerekmektedir. Bunlara ek olarak her

yeni gözlem elde edildikçe, zaman serisi modellerinin güncellenmesi gerekmektedir. Güncelleme işlemi yavaş gelişen salgınları belirlemeyi engelleyebilmektedir (Unkel vd., 2012).

Salgınları belirlemek amacıyla Arima modellerinden daha karmaşık modellerde kullanılmıştır. Bu modeller ise Gizli Markov Modelleri (Hidden Markov Models), Markov Switching Models (Bayesçi markov anahtarlama modelleri olarak) ve Held vd. (2006) iki bileşenli bayes modelidir. Bu yöntemlerde ortak düşünce zaman serisini salgın yok ve salgın var şeklinde iki parçaya bölmek ve bu parçalar arasındaki geçiş noktasını tahmin etmeye çalışmaktır. Bu yöntemler arasındaki temel farklılık ise trend, mevsimsellik ve öz ilişkiyi modellemedeki farklılıklarından kaynaklanmaktadır (Unkel vd., 2012).

### 3.2.1 Holt-Winters düzeltme yöntemi

Bu yöntem zaman içinde değişen trend, değişen döngüsel etki ve değişen ortalama düzeylerine sahip zaman serilerini tahmin etmek amacıyla kullanılmaktadır. Finansal ve sanayi üretimi ile ilgili zaman serileri için yaygınlıkla kullanılan bir yöntemdir. Bu yöntem  $L_t$ ,  $T_t$ ,  $S_t$  ile ifade edilen ve ardışık olarak güncellenen ortalama düzeyi, trend bileşeni ve mevsimsel bileşen içermektedir. Bu yöntemle  $k$  adım ötesi için öngörü ya da tahmin aşağıdaki gibi elde edilir (Burkom, 2007).

$$\hat{y}_{t+k} = (L_t + kT_t)S_{t+k-M}$$

Burada  $M$  döngüyü oluşturan zaman birimi sayısıdır. Örneğin haftalık döngü için  $M = 52$ 'dir.  $L_t$ ,  $T_t$  ve  $S_t$  aşağıdaki eşitlikler kullanılarak güncellenmektedir (Burkom, 2007).

$$L_t = \alpha \frac{Y_t}{S_{t-M}} + (1 - \alpha)(L_{t-1} + T_{t-1})$$

$$T_t = \beta(L_t - L_{t-1}) + (1 - \beta)T_{t-1}$$

$$S_t = \gamma \frac{Y_t}{L_t} + (1 - \gamma)S_{t-M}$$

Sendromik zaman serilerinde genellikle mevsimsel etkiler  $S_t$ , serinin ortalama düzeyi  $L_t$ 'ye oransal olduğundan bu denklemlerdeki  $L_t$  ve  $S_t$  arasındaki ilişki çarpımsaldır. Bu ilişkinin toplamsal olduğu Holt-Winters yaklaşımı da mevcuttur. Bu yöntem  $\alpha, \beta$  ve  $\gamma$  gibi üç düzeltme parametresine ihtiyaç duymaktadır. Bu parametreler güncelleştirme miktarlarını belirlemektedirler. Seçilen güncelleme eşitlikleri verilerin beklenen davranışlarını yakalayabilecek en basit biçimde olmalıdırlar. Holt-Winters yaklaşımı ortalama düzeyi, trend ve mevsimsel bileşenler için başlangıç değerleri gerektirmektedir. Başlangıç değerlerinin seçimi öngörü için önem taşımaktadır. Bu yaklaşımla negatif değerlerin tahmin edilmesinden kaçınmak için, ardışık olarak hesaplanan ortalama düzeyi  $L_t$ 'nin alt sınırının 0 olması koşulu getirilir. Dolayısıyla sıfıra bölüm işleminden kaçınmak amacıyla da  $L_t$  ve  $S_t$  aşağıdaki gibi tanımlanır (Burkom, 2007).

$$S_{t-M} = 0 \quad \text{ise} \quad L_t = L_{t-1} + T_{t-1} \quad \text{ve} \quad L_t = 0 \quad \text{ise} \quad S_t = S_{t-M}$$

Bu yaklaşımda yapılabilecek başka bir ayarlama ise, öngörü hatası çok büyük olduğunda serinin ortalama ve trendinin güncellenmesinin yapılmamasıdır. Örneğin mutlak değerce öngörü hatası en fazla öngörü değerinin yarısı ise güncelleme yapılabilir. Böylelikle aykırı değerlerden kaynaklanan olumsuz etki azaltılabilmektedir (Burkom, 2007).

### 3.2.2 Hidden Markov Modeli Yöntemi

Le Strat ve Carrat (1999) Hidden Markov modellerinin halk sağlığında rutin olarak toplanan verilerin yani zaman serilerinin analizinde kullanılmasını önermişlerdir. Bu model, bir Markov zincirinin durumlarına bağlı olarak bir dizi ölçümün olasılık yoğunluk fonksiyonunun farklılaştığını varsaymaktadır ve bu varsayıma dayanarak ölçümlerin hangi durum altında elde edildiğini analiz etmektedir. Bu modele ilişkin parametre vektörü, Markov zincirinin durumlarına bağlı olarak değişen olasılık yoğunluk fonksiyonlarına ilişkin parametreleri ve durumlar arası geçiş olasılıklarını içermektedir. Parametre tahminleri düzenlenmiş bir EM (expectation maximization)

algoritması ile gerçekleştirilmektedir. Ayrıca verilerdeki trend ve mevsimselliği dikkate alacak şekilde algoritma düzenlenebilmektedir (Le Strat ve Carrat, 1999). Hidden Markov modelleri, örneğin; otomatik konuşma tanıma, elektrokardiyografik sinyal analizi, epileptik nöbet frekansı analizi, DNA dizi analizi, nöron halkası ve meteorolojinin modellenmesi gibi birçok alanda kullanılmaktadır (Le Strat ve Carrat, 1999).

Hidden Markov modelinde gözlemler;  $y_t, t = 1, \dots, n$  biçimindedir ve bu gözlemler stokastik bir süreç olan  $\mathbf{Y} = (Y_t; t = 1, \dots, n)$  tarafından üretilmektedir. Hidden Markov modelinin temel düşüncesi, her  $Y_t$  rasgele değişkenini gözlemlenemeyen ve durumu gösteren bir rastgele değişken  $S_t$  ile ilişkilendirmek ve bu değişkene bağlı olarak  $Y_t$ 'nin koşullu olasılık dağılımını belirlemektir.

$$(Y_t | S_t = j) \sim f_{j_t}(y_t; \theta_j) \quad j = 1, 2, \dots, m$$

Burada  $f_{j_t}$ , belirli bir parametrik dağılım ailesine ilişkin olasılık yoğunluk fonksiyonunu ve  $\theta_j$  ise tahmin edilecek parametreleri göstermektedir. Örneğin logaritmik bağ fonksiyonuyla doğrusal olabilen bir Poisson dağılımı için koşullu dağılımlar aşağıdaki gibi tanımlanır.

$$(Y_t | S_t = j) \sim Po(\mu_t^j)$$

$$\log(\mu_t^j) = \varphi_j + \beta_j t + \sum_{i=1}^{nH} \gamma_j^i \cos \left[ \frac{2i\pi}{freq} (t-1) \right] + \delta_j^i \sin \left[ \frac{2i\pi}{freq} (t-1) \right]$$

Gözlemlenemeyen  $\{S_t\}$  dizisinin, durağan geçiş olasılıkları  $\alpha_{ij}$  ile  $m$  durumlu 1. dereceden bir Markov zinciri olduğu varsayılmaktadır.

$$\alpha_{ij} = P(S_t = j | S_{t-1} = i), \quad i, j = 1, \dots, m$$

Ayrıca  $S_t$ 'ye ait bilginin verilmesi koşulu altında,  $Y_t$  rastgele değişkenlerinin bağımsız olduğu kabul edilmektedir. Markov zincirinin durumlarının dizilimi

gözlemlenemediği için, bu Markov zincirine bağlı karma model gizli bir Markov modeli olarak adlandırılır (Le Strat ve Carrat, 1999).

### **3.3. İstatistiksel Süreç Kontrol Grafiklerine Dayanan Yaklaşımlar**

İleriye dönük olarak hastalıklara ilişkin vaka sayılarındaki olağan dışılığı belirlemek, sanayideki üretim süreçlerinde olağan dışı durumları belirlemeye benzerdir (Unkel vd., 2012). Bu nedenle, bir üretim sürecindeki ürünlerin kalitesini denetim altına almak ve izlemek amacıyla geliştirilen yöntemler örneğin Shewhart, Ewma, Cusum gibi grafikler ya da bu yöntemlerden esinlenmiş çeşitli yöntemler halk sağlığı izlemelerinde kullanılmaktadır.

İstatistiksel süreç kontrol yaklaşımlarının kullanımının halk sağlığı izlemesindeki oluşturduğu problem, halk sağlığı verilerinin genellikle bu yöntemlerin dayandığı varsayımları sağlamamasıdır. Örneğin durağan olma, bağımsız olma, sabit varyanslılık ve normallik gibi varsayımlar halk sağlığı izlemesi verileri için gerçekçi olmamaktadır (Unkel vd., 2012).

İstatistik süreç kontrol grafiklerini kullanımından doğan varsayım ihlali problemini hafifletmek amacıyla, verilere çeşitli modeller uydurulmakta (örneğin doğrusal ya da Poisson regresyon modeli, Kalman Filtrelemesi, zaman serisi modeli, wavelet analizi gibi) ve artıklar elde edilmektedir. Elde edilen artıklara klasik istatistiksel süreç kontrol yöntemleri uygulanmaktadır. Bunlara ek olarak bazen bu süreç kontrol yöntemlerinin eşik değerleri geçmişe ait veriler kullanılarak deneysel olarak elde edilmekte ve salgının olmadığı zaman dilimiyle, salgının olup olmadığı kontrol edileceği zaman dilimi arasına koruma boşluğu eklenmektedir.

Halk sağlığı izlemesi için sıklıkla kullanılan yöntemler yüzeysel olarak şu şekilde sınıflandırılabilir: Klasik süreç kontrol grafikleri (ya da bu grafiklerin halk sağlığı verilerine uyarlanmış biçimleri), zamana bağlı tarama istatistiği ve başarısızlık zamanını yani belirli bir olay gerçekleşene kadar geçen süreyi temel alan yöntemlerdir (Unkel vd., 2012).

### 3.3.1 Cusum kontrol grafiđi

Rossi vd. (1999) zamanla deđiřebilecek ortalamalara sahip Poisson dađılımlı veriler için kullanılacak yaklařık bir Cusum yöntemi önermiřlerdir. Bu yöntem için izlenen olay sayılarının aynı dađılımlı Poisson olduđu kabul edilir:  $y_1, \dots, y_n \sim \text{Poisson}(m)$ . Poisson dađılımlı verilerden dolayı, önerilmiř Cusum yönteminin performansının düşmemesi için veriler normal dađılımlı olacak řekilde dönüřtürölmektedir. Rossi vd. (1999) verilerde normalliđi sađlamak amacıyla ařađıdaki dönüřümü önermiřlerdir (Höhle, 2007).

$$x_t = \frac{y_t - 3m + 2\sqrt{my_t}}{2\sqrt{m}}$$

Pierce ve Schafer (1986) deviance, Pearson veya Anscombe artıkları gibi normalliđi sađlayan bařka dönüřümlerden de söz etmektedir. Dönüřümden sonra kontrol altındaki verilerin  $N(0,1)$  ve kontrol dıřındaki verilerin  $N(\mu, 1)$  dađıldıđı kabul edilmektedir (Höhle, 2007).

Verilerde mevsimsellik gibi zamana bađlı etkilerin var olması durumunda; verilerin aynı dađılımlı ve bađımsız olma varsayımı, Poisson dađılımının ortalaması  $m$ 'nin zamana bađlı olarak ifade edilmesiyle esnetilebilir. Bu amaçla mevsimselliđi dikkate almak amacıyla bir Poisson regresyon modeli kullanılır:

$$\log(m_t) = \alpha + \sum_{s=1}^S [\gamma_s \cos(\omega_s t) + \delta_s \sin(\omega_s t)]$$

Burada  $\omega_s, \frac{2\pi}{52}s$  biřimindeki Fourier frekanslarıdır. Bu yöntem için Cusum istatistiđi,  $S_0 = 0$  ve  $k$  kullanıcı tarafından belirlenmek üzere ařađıdaki gibi hesaplanır:

$$S_t = \max(0, S_{t-1} + x_t - k), \quad t = 1, \dots, n$$

Daha sonra arzu edilen ARL performanslarını verecek řekilde eřik deđer  $h$  belirlenir ve  $S_t > h$  olduđunda bir sinyal verilir. Verilen ARL'ler için  $k$  ve  $h$  deđerleri örneđin



R programındaki spc paketindeki bir fonksiyonla hesaplanabilmektedir. Kenett ve Pollak (1983) bu Cusum yönteminde sinyal verildikten sonra Cusum istatistiğini sıfırlanmamasını yani kendiliğinden eşik değerinden düşük oluncaya kadar sinyal verilmesini önermişlerdir (Höhle, 2007).

### 3.3.2 Bireysel ölçümler için Shewhart kontrol grafiği

Süreçlerin izlenmesi için birçok durumda süreçlerin değişik zaman noktalarında alınacak örneklemin büyüklüğü 1 'dir (Montgomery, 2009). Böyle durumlarda süreçteki değişkenliği kontrol etmek için; süreç ortalamasındaki orta ve büyük seviyedeki kaymaları tespit etmek amacıyla, Shewhart I-Chart yani Shewhart bireysel kontrol grafiği kullanılabilir. Bireysel kontrol grafiği aşağıdaki gibi ifade edilmektedir.

$$AKS = \mu - k\sigma, \quad \text{Süreç ortalaması} = \mu, \quad ÜKS = \mu + k\sigma$$

Burada  $\mu$  bireysel gözlemlere ilişkin süreç ortalamasını ve  $\sigma$  bu gözlemlere ilişkin süreç standart sapmasını göstermektedir. Bu grafiği oluşturmak için  $\mu$  bireysel gözlemlerin ortalaması  $\bar{X}$  ile tahmin edilirken,  $\sigma$  parametresi ise  $\frac{\overline{MR}}{d_2}$  ile tahmin edilmektedir. Burada  $d_2$  değerleri çeşitli  $n$ 'ler için tablolatırılmış bir sabittir.  $\bar{R}$  ise  $MR_i = |X_i - X_{i-1}|$ ,  $i = 1, 2, \dots, n - 1$  biçiminde tanımlanan ardışık gözlemler arasındaki genişliğin ortalamasını göstermektedir. Ayrıca  $k$  ise genellikle 3 olarak alınmaktadır. Bu durumda bireysel kontrol grafiği parametreleri aşağıdaki gibi belirlenmektedir (Montgomery, 2009).

$$AKS = \bar{X} - 3 \frac{\overline{MR}}{d_2}, \quad \text{Süreç ortalama çizgisi} = \bar{X}, \quad ÜKS = \bar{X} + 3 \frac{\overline{MR}}{d_2}$$

Bu grafik bireysel gözlemlerin normal dağıldığını varsaymaktadır. Gözlemlerin normal dağılmaması durumunda kontrol altındaki  $ARL_0$ 'ın dikkate değer ölçüde etkilendiği belirlenmiştir. Dolayısıyla bu sürece ait izlemenin yapıldığı ikinci aşama için sorun oluşturmaktadır. Böyle bir sorunu gidermek amacıyla değişkene normal

dağılmayı sağlayacak şekilde dönüşüm yapılması ya da verilerin gerçek dağılımdan elde edilen yüzdelerinin üst ve alt kontrol sınırlarında kullanılması önerilmektedir. Özellikle en az 100 tercihen de 200 gözlem mevcutsa, üst ve alt kontrol sınırlarında kullanılacak yüzdelerinin bu verilerden tahmin edilebileceği belirtilmektedir (Montgomery, 2009).

### 3.3.3 Halk sağlığı izlemesinde kullanılan süreç kontrol grafikleri

Shewhart kontrol grafikleri, genellikle normal dağılımlı bir sürecin mevcut en son gözlemine ( $y_t$ ) ilişkin bilgiyi kullanarak sürecin kontrol dışı olup olmadığına karar vermektedir. Süreçteki en son gözlemin yanı sıra geçmişteki gözlemleri de kullanan süreç kontrol grafikleri mevcuttur. Bunlardan en yaygın olanlardan ikisi Cusum ve Ewma süreç kontrol grafikleridir. Tek yönlü standartlaştırılmış Cusum istatistiği aşağıdaki gibi tanımlanmaktadır.

$$C_t = \max \left\{ 0, C_{t-1} + \left( \frac{y_t - \mu_t}{\sigma_t} - k \right) \right\}$$

Burada  $C_0 = 0$  ve genellikle  $k = 1/2$  alınır.  $C_t > h$  ise sürecin kontrol dışında olduğu söylenir. Eşik değeri  $h$ , salgın yokken yanlış sinyaller arasındaki ortalama çalışma uzunluğu  $ARL_0$ 'ın istenen değere ulaşmasını sağlayacak biçimde belirlenir. Nadir olayların sayısının izlenmesinde, normal dağılımlı Cusum yerine Poisson Cusum daha uygundur. Eğer veriler oranlar biçiminde ifade edilebiliyorsa, binom Cusum grafiği kullanılabilir (Unkel vd., 2012).

Başka bir grafik Ewma ise geçmişteki gözlemlere, üstel olarak azalan bir hızda daha az ağırlık vermektedir.

$$Z_t = \lambda \left( \frac{y_t - \mu_t}{\sigma_t} \right) + (1 - \lambda)Z_{t-1}$$

Burada  $Z_0 = 0$  ve  $\lambda \in (0,1]$ .  $\lambda = 1$  ise Ewma, Shewhart grafiğine dönüşür.  $Z_t > L\sigma_{Z_t}$  ise sürecin kontrol dışında olduğuna karar verilir.  $L$  ise istenen  $ARL_0$

değerini verecek biçimde seçilir. Süreç gözlemlerindeki trend ve mevsimselliği açıklamak amacıyla, Ewma grafiklerinde Holt-Winters yöntemi önerilmiştir. Poisson ve binom dağılımlı veriler durumuna uyarlanmış Ewma grafikleri de mevcuttur. Ayrıca frekanslardan oluşan verileri kullanan regresyon süreç kontrol grafikleri de önerilmiştir. Bu yöntemler bulaşıcı hastalıkların ortalama vaka sayısına olan mevsimsel değişimleri dikkate almaktadır. Vaka sayıları örneğin negatif binom dağılımlı kabul edilebilmektedir. Frekans verilerine dayalı regresyon grafikleri gözlemler arasında olabilecek öz ilişkiyi de dikkate alacak biçimde uyarlanmışlardır (Unkel vd., 2012).

### 3.3.4 Ears C1, C2 ve C3 Yöntemi

Ears C1, C2 ve C3 yöntemleri BioSense izleme sisteminde kullanılan ve istatistiksel süreç kontrol grafiklerinden esinlenmiş erken anormallik belirleme yöntemleridir. Cusum benzeri yöntemler olarak geliştirilmelerine rağmen, aslında C1 ve C2 yöntemleri hareketli ortalama ve standart sapmayı kullanan Shewhart türü yöntemlerdir (Fricker vd., 2008).

Bu yöntemlerden C1, her karar verme noktasından önceki 7 gözlemi referans olarak almakta ve kullandığı izleme istatistiğini aşağıdaki gibi oluşturmaktadır.

$$C_1(t) = \frac{Y(t) - \bar{Y}_1(t)}{S_1(t)} \quad \bar{Y}_1(t) = \frac{1}{7} \sum_{i=t-7}^{t-1} Y(i) \quad S_1^2(t) = \frac{1}{6} \sum_{i=t-7}^{t-1} [Y(i) - \bar{Y}_1(i)]^2$$

Başka bir ifadeyle izleme istatistiği oluşturulurken gözlemler, son 7 gözleme ilişkin ortalama ve standart sapma kullanılarak standartlaştırılmaktadır. Olağandışı bir durum olmadığı hipotezi altında  $C1 \sim N(0,1)$  olduğu varsayılmakta ve bu dağılımın  $\%100(1 - \alpha)$ . kuantili kullanılarak eşik değeri  $U_t = \bar{y}_t + z_{1-\alpha} s_t$  olarak hesaplanmaktadır. Karar verilecek  $t$  anındaki gözlem için  $y_t > U_t$  ise, olağandışı durum sinyali verilir (Salmon vd., 2016). Genellikle bu yöntemde olağandışılık sinyali  $C_1(t) > 3$  ise verilmektedir (Fricker vd., 2008).

C2 yöntemi ise, karar verme noktasından önceki son 9 gözlemin ilk 7 gözlemini referans ve son 2 gözlemini koruma bandı olarak almaktadır. İzleme istatistiği yine referans alınan gözlemlerin ortalaması ve standart sapması kullanılarak aşağıdaki gibi oluşturulmaktadır.

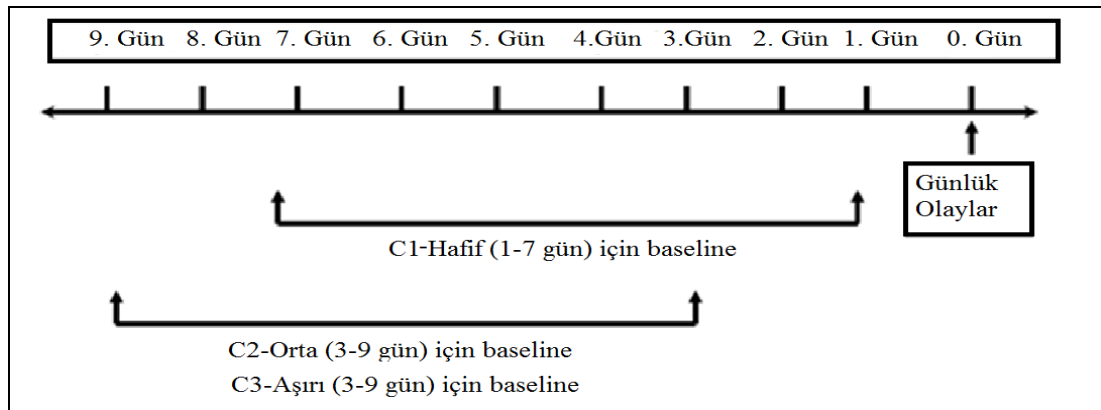
$$C_2(t) = \frac{Y(t) - \bar{Y}_3(t)}{S_3(t)} \quad \bar{Y}_3(t) = \frac{1}{7} \sum_{i=t-9}^{t-3} Y(i) \quad S_3^2(t) = \frac{1}{6} \sum_{i=t-9}^{t-3} [Y(i) - \bar{Y}_3(i)]^2$$

Başka bir ifadeyle, karar verilecek t noktasındaki gözlem, koruma bandı olan son 2 gözlemden önceki 7 gözlemin ortalaması ve standart sapması kullanılarak standartlaştırılmaktadır. Eşik değeri C1 yönteminin eşik değerine benzer şekilde oluşturulmaktadır;  $C_2(t) > 3$  ise olağandışı durum sinyali verilmektedir (Fricker vd., 2008).

C3 yöntemindeki izleme istatistiği ise; t anındaki karar verme noktasındaki gözlem ve önceki 2 gözlem için, C2 yöntemine ilişkin izleme istatistiklerini kullanarak aşağıdaki gibi oluşturulmaktadır.

$$C_3(t) = \sum_{i=t}^{t-2} \max[0, C_2(i) - 1]$$

Bu yöntem için  $C_3(t) > 2$  olduğunda olağandışı durum sinyali verilir (Fricker vd., 2008). Bu üç izleme yöntemine ilişkin referans dönemi Şekil 3.1.'de verilmektedir.



Şekil 3.1. Ears C1, C2 ve C3 yöntemlerine ilişkin referans dönemleri

Bu yöntemler yukarıda görüldüğü gibi kısa dönem referans verisi kullanmaktadır. Örneğin; olimpiyatlar, ulusal veya uluslararası toplantılar sırasındaki olağandışı durumları belirlemek amacıyla kullanılan standart izleme yöntemleri arasındadır. Ayrıca bu yöntemler verilerin bağımsız olduğunu varsaymakta, trend, mevsimsellik ve haftanın günü etkisi gibi etkileri dikkate almamaktadır. Bu nedenle bu yöntemler kullanılırken bu tür etkilerin giderilmesi gerekmektedir (Fricker vd., 2008; Salmon vd., 2016). Fricker vd. (2008), bu yöntemleri Cusum yöntemiyle çeşitli koşullar arasında simülasyon çalışmaları aracılığıyla karşılaştırmış ve Cusum yönteminin daha iyi performans gösterdiğini tespit etmiştir.

### **3.3.5 Olaylar arasında geçen süreyi temel alan kontrol grafikleri**

İlgilenilen olaylar nadir gerçekleşiyorsa, Chen'in (1978) önerdiği ve daha sonra çeşitli durumlara genelleştirilmiş olan kümeler yöntemi kullanılmaktadır. Kümeler yönteminde, nadir olaylar arasındaki olayların tamamına kümeler denir. Bu kümelerin büyüklükleri geometrik dağılımlı rasgele değişkenlerdir. Ardışık küme büyüklüklerinin küçük olması nadir olayların daha sık gerçekleştiği yani süreçte bir olağan dışılık gerçekleştiği anlamına gelir. Ayrıca nadir sağlık olaylarının izlenmesi için Bernoulli Cusum kontrol grafiğinin kullanımında önerilmiştir.

Kümeler yöntemi, homojen Poisson süreçlerine yani olaylar arasında geçen sürenin üstel dağılımlı olduğu duruma da uyarlanabilmektedir. Ardışık  $n + 1$  olay arasındaki süre yani  $n$  adet zaman dilimi gözlemlenir. Bu zaman dilimlerinin genişliği önceden belirlenen bir eşik değerinden ( $\tau$ ) küçükse, ilgilenilen olayların meydana gelişlerinde olağan dışı şekilde artış olduğuna karar verilir. Burada  $(n, \tau)$  çifti, yanlış karar verme olasılığının ve doğru karar verme olasılığının (duyarlılığın) hedeflenen değerlerde olmasını sağlayacak şekilde belirlenir. Ayrıca olaylar arasındaki süre üstel dağılımlı olduğunda, üstel Cusum ya da üstel Ewma grafikleri kullanılabilir (Unkel vd., 2012).

### 3.3.6 Tarama İstatistiğine Dayanan Yöntemler

Tarama istatistikleri, ya tamamen zaman içinde ya tamamen mekansal ya da tamamen hem zamansal hem de mekansal olarak oluşan hastalık vakalarının yığılmalarını belirlemek için kullanılmaktadır. Zamana bağlı tarama ise belirli genişlikte pencereyi zaman ekseninde hareket ettirerek, bu pencere içindeki gözlemlenen ve olması beklenen gözlemlerin incelenmesiyle yapılmaktadır. Bu yöntem geriye dönük salgın belirleme yöntemi olmasına rağmen, Kulldorff (2001), İsmail vd. (2003), Naus ve Wallenstein (2006) bu yöntemi ileriye dönük salgın belirleme durumuna uyarlamışlardır. İleriye dönük olarak zamana bağlı taramayı temel alan iki farklı yöntem mevcuttur. Bunlardan ilki belirli bir bölgede en son zaman dilimindeki ya da sabit bir genişlikteki pencere içindeki vakaların sayısını bulmayı içermektedir. En son zaman noktasındaki vaka sayısını gösteren gözlem  $y_n$  ve  $L$  pencerenin genişliği olmak üzere, tarama istatistiği;

$$S_n = \sum_{t=n-L+1}^n y_t$$

ağırlıklandırılmamış hareket eden toplam olarak tanımlanmaktadır. İstenen  $ARL_0$ 'ı veren eşik değeri  $h$  belirlenir ve  $S_n > h$  ise salgın sinyali verilir. İkinci yöntemde ise yani Kulldorff (2001) ileriye dönük tarama yönteminde, pencere genişliği sabit olmayıp, zaman ekseninde bulunulan noktaya bağlı olarak değişmektedir. Halk sağlığı izleme verileri sıklıkla mevsimsellik ve çeşitli etkilerle durağan değildir. Bu nedenle Wallenstein ve Naus (2004) mevsimsellik etkilerini dikkate alan zamana bağlı bir tarama yöntemi önermişlerdir (Unkel vd., 2012).

### 3.4 Hipotez Testlerine Dayanan Biyoizleme Yaklaşımları

Bir  $y_t$  gözleminin salgının habercisi olup olmadığını belirlemenin bir yolu, referans değerlerle aynı dağılıma sahip olup olmadığını testin yapılmasıdır. Frisen ve

Anderson (2009), Frisen vd. (2010) bir salgının başlangıcını belirlemek amacıyla, parametrik ve yarı parametrik yöntemler önermişlerdir. Yarı parametrik bir model, hastalık vakalarının sayısının düzenli üstel dağılım ailesinden bir dağılıma uyduğunu varsaymaktadır. Zaman noktası  $t$ 'ye kadar ki vaka sayıları  $y_s$  ( $s < t$ ), ya aynı ortalamaya sahiplerdir (yokluk hipotezi) ya da  $d$  bilinmek üzere  $t = d$  noktasından itibaren monoton olarak artmaktadır (alternatif hipotez). İzlemenin başlangıcından ( $d = 1'den$ ) itibaren artan Poisson dağılımlı vaka sayıları için, olabilirlik oranı istatistiği aşağıdaki gibi tanımlanır:

$$LR = \prod_{s=1}^t \left( \frac{\hat{\mu}_s^1}{\hat{\mu}_s^0} \right)^{y_s}$$

Burada  $\hat{\mu}^0$ , yokluk hipotezi doğruyken süreç ortalamasının ML kestiricisi;  $\hat{\mu}^1$  ise alternatif hipotez doğruyken süreç ortalamasının ML kestiricisidir. Önceden belirlenmiş bir  $k$  değeri için,  $LR > k$  ise salgın alarmı verilir (Unkel vd., 2012).

Bu alt bölümde bir salgının başlangıcını belirlemek amacıyla önerilmiş olabilirlik oran testini temel alan bazı yöntemler tanıtılmaya çalışılacaktır. Özellikle ilgilenilen hastalığın salgın durumunda olup olmadığının göstergesi olan gözlemlerin ortalaması sabit bir düzeydeyken, bu düzeyin artış eğilimine geçmesi durumu ile ilgilenilmektedir. Başka bir ifadeyle başlangıçta sabit olup daha sonra monoton artan bir eğriye sahip regresyon modelleri ile salgının başlangıcı modellenmeye çalışılmaktadır.

İlgilenilen olabilirlik oranı yaklaşımları; tam olabilirlik oranı yaklaşımı, Shiryayev Roberts olabilirlik oran yaklaşımı ve Shewhart yaklaşımı şeklinde üç sınıfa ayrılabilir. Ayrıca yaklaşımlar parametrik ve parametrik olmayan olabilirlik oran yaklaşımları biçiminde de sınıflandırılabilir (Frisen ve Anderson, 2009; Bock vd. 2007).

İleriye dönük olarak izleme amaçlı olan bazı yöntemler, bir tür olabilirlik oranı biçiminde ifade edilebilmektedir. Bundan dolayı bazı iyilik kriterlerini (ölçütlerini)

sağlamaktadır. Bu yöntemler arasında Shiryaev-Roberts, Ewma, Cusum ve olabilirlik oranı yöntemine doğrusal yaklaşımları gibi yöntemler bulunmaktadır.

Tanıtilacak yaklaşımları açıklamak amacıyla Frisen ve Anderson (2009)'ın gösterimi kullanılacaktır.  $X$  gözlemlenen değişkene ait süreci,  $x(t)$  ise bu sürecin  $t$  zamanındaki değerini gösterebilir. Karar verme zamanı  $s$  olmak üzere,  $s$  noktasına kadar ki mevcut bilgi  $x_s = \{x(1), \dots, x(s)\}$  ile gösterilsin. Ayrıca  $D$ , salgın öncesinde süreç ortalamasının sabit olduğu dönemi,  $C$  ise süreç ortalamasının monoton olarak artmaya başladığı salgının başlangıç dönemini temsil etsin. Bu gösterimlere ek olarak  $\tau$  salgının bilinmeyen başlangıç noktasını gösterebilir.  $\tau$  kesikli bir rasgele değişken olarak görülebilir. Bu durumda  $C = \{C1 \cup C2 \cup \dots, Cs\}$  ya da eşdeğer olarak  $C = \{\tau \leq s\}$  biçiminde ifade edilebilirken,  $D$  ise  $D = \{\tau > s\}$  biçiminde tanımlıdır.

Değişim noktası  $\tau'$  ya bağlı olarak süreç ortalaması  $\mu(t)$  genel şekliyle,  $C$  ve  $D$  durumları için aşağıdaki gibi ifade edilebilmektedir.

$$\mu(t) = \begin{cases} \mu^D(t), & t < \tau \\ \mu^{C_j}(t), & t \geq \tau, \tau = j \end{cases}$$

$$D \text{ durumu} : \mu(1) = \mu(2) = \dots = \mu(s)$$

$$C1 \text{ durumu} : \mu(1) \leq \mu(2) \leq \dots \leq \mu(s)$$

$$C2 \text{ durumu} : \mu(1) < \mu(2) \leq \dots \leq \mu(s)$$

$$Cj \text{ durumu} : \mu(1) = \dots = \mu(j-1) < \mu(j) \leq \dots \leq \mu(s), \quad j > 2$$

İzleme problemi ise karar verilecek  $s$  noktasında; sürecin başlangıcından itibaren ulaşılabilen tüm bilgiye dayanarak yani  $x_s$  gözlem vektörünü temel alarak, sürecin  $C$  veya  $D$  döneminde olup olmadığına karar vermektir.

### 3.4.1 Tam olabilirlik oranı yaklaşımı

$C$  ve  $D$  durumları için  $s$  karar verme noktasındaki tam olabilirlik oranı Eş. 3.4'deki gibi tanımlıdır (Frisen vd., 2009; Bock vd. 2007).



$$LR(s) = \frac{f(x_s/C)}{f(x_s/D)} = \sum_{j=1}^s w_j \cdot \frac{f(x_s|\tau = j)}{f(x_s|\tau > s)} = \sum_{j=1}^s w_j \cdot L(s, j) > k_s \quad (3.4)$$

Burada ağırlıklar  $w_j = P(\tau = j)/P(\tau \leq s)$ , kısmi olabilirlik oranları  $L(s, j) = \frac{f(x_s|\tau=j)}{f(x_s|\tau>s)}$ , yoğunluk ya da olabilirlik fonksiyonu  $f$  ile gösterilmektedir. Ayrıca  $x_s = \{x(1), \dots, x(s)\}$ ,  $s$  zamanına kadar olan gözlem vektörüdür ve  $k = \left(\frac{P(D(s))}{P(C(s))}\right)$  oranını göstermek üzere,  $k_s = k/(1 - k)$  biçimindedir.  $C$  olayı  $C = \{\{\tau = 1\}, \{\tau = 2\} \dots, \{\tau = s\}\}$  biçiminde bileşik olduğundan, Eş. 3.4'deki olabilirlik oranı kısmi olabilirlik oranları  $LR(s, j)$  'lerin ağırlıklı toplamı şeklinde yazılabilmektedir. Eş. 3.4'deki durum gerçekleştiğinde salgın başlangıcı alarmı verilir.

Shiryaev (1963) tam olabilirlik oranı yönteminin, verilen bir yanlış alarm olasılığı için en kısa gecikme zamanını veren yaklaşım olduğunu göstermiştir (Frisen ve Anderson, 2009). Tam olabilirlik yaklaşımı hem  $\tau$ 'nun hem de  $x_s$  gözlem vektörünün olasılık dağılımının bilinmesini gerektirmektedir.

### 3.4.2 Shiryaev-Roberts olabilirlik oran yaklaşımı

Shiryaev (1963) ve Roberts (1966) olabilirlik oran yönteminin özel bir durumunu önermişlerdir. Bu özel durumda  $\tau$ 'nun geometrik olasılık dağılıma sahip olduğu ve  $\lim_{j \rightarrow \infty} P(\tau = j | \tau \geq j) = 0$  eşitliğinin sağlandığı kabul edilir. Böylelikle Eş. 3.4'deki  $\omega_j$  ağırlıklarının eşit olduğu söylenebilmektedir. Dolayısıyla Eş. 3.4'deki olabilirlik oranı (SR yöntemine ilişkin istatistik) Eş. 3.5'teki gibi ifade edilir.

$$LR_{SR}(s) = \sum_{j=1}^s \frac{f(x_s|\tau = j)}{f(x_s|\tau > s)} = \sum_{j=1}^s L(s, j) > k \quad (3.5)$$

Eş. 3.5'te görüldüğü gibi SR yaklaşımına ilişkin olabilirlik oran istatistiği sabit bir limit değeri  $k$ 'yı aştığında, salgın sinyali verilir (Bock vd., 2007).

### 3.4.3 Shewhart yaklaşımı

Frisen (2003) ve Frisen (2007), Shewhart yönteminin  $C = \{\tau = s\}$  durumunda olabilirlik oran yönteminin özel bir durumu olduğunu belirtmiştir. Frisen ve Mare (1991) bağımsız normal dağılımlı süreç gözlemleri ve süreç ortalamasında kayma söz konusu olduğunda, Eş. 3.4'deki olabilirlik oranına ilişkin eşitsizliğin Eş. 3.6'daki eşitsizliğe dönüştüğünü belirtmiştir.

$$(x(s) - \mu^D) > k \quad (3.6)$$

Salgın sinyali verme sınırı  $k$  ise genellikle  $3\sigma$  olarak seçilir (Frisen ve Andersson, 2009). Genelleşmiş bir Shewhart yöntemi,  $G$  sabit olmak üzere aşağıdaki alarm kriteri ile ifade edilebilir.

$$L(s, s) > G$$

Frisen (2007) kararın verildiği anda büyük bir değişim gerçekleşiyorsa, Shewhart yönteminin kullanılmasının uygun olacağını çünkü bu tür durumlarda Shewhart yönteminin en iyi belirleme kabiliyetine sahip olacağını ifade etmiştir. Başka bir ifadeyle karar verilen anda süreçte büyük bir değişim bekleniyorsa, Shewhart yöntemi alarm vermek için en kısa gecikme zamanına sahip olacaktır (Frisen ve Andersson, 2009).

### 3.4.4 Parametrik olabilirlik oran yaklaşımı

Parametrik yaklaşım süreç ortalama vektörünün parametrik bir fonksiyonla ve süreci tanımlayan dağılımın parametrik olarak ifade edilmesini gerektirmektedir. Grip salgınlarının belirlenmesi amacıyla süreç ortalaması zamanın fonksiyonu olarak çeşitli şekillerde parametrik olarak modellenmişlerdir. Bu modeller arasında parçalı doğrusal model, Serfling (1963) modeli ve parçalı üstel model vardır. Örneğin parçalı bir üstel model aşağıdaki gibi ifade edilebilmektedir.

$$\mu(t|\tau) = \begin{cases} \mu^D(t), & t < \tau \\ \mu^{C_j}(t), & t \geq \tau, \tau = j \end{cases}$$

$$\mu(t|\tau) = \begin{cases} \mu_0, & t < \tau \\ \exp(\beta_0 + \beta_1 \cdot (t - j + 1)), & t \geq \tau, \tau = j \end{cases}$$

Burada  $\mu_0, \beta_0$  ve  $\beta_1$  sabit değerlerdir. Örneğin süreç gözlemleri  $X$ , Poission dağılımlı,  $\tau$ 'da  $v$  parametresi ile geometrik dağılımlı ise, olabilirlik oran istatistiği aşağıdaki gibi ifade edilebilir.

$$LR(s) = \sum_{j=1}^s \omega_j \frac{\prod_{t=1}^s \exp(-\mu^{C_j}(t)) \cdot (\mu^{C_j}(t))^{x(t)}}{\prod_{t=1}^s \exp(-\mu^D(t)) \cdot (\mu^D(t))^{x(t)}}$$

Burada  $\omega_j = \frac{(v(1-v)^{i-1})}{1-(1-v)^s}$  şeklinde tanımlıdır. Parametrik olabilirlik oran yaklaşımı,  $\omega_j$  ağırlıklarının ve  $\mu(t)$  fonksiyonun parametrik biçiminin bilinmesi şartıyla, tam olabilirlik oranına ilişkin iyilik özelliklerini taşımaktadır. Çoğu zaman bunların bilgisine ulaşılabilir değildir. Bu nedenle bu bilgiler olmaksızın olabilirlik oran yaklaşımları önerilmiştir (Frisen ve Andersson, 2009).

Bu yaklaşımlardan birisi de Shiryaev-Roberts olabilirlik oran yöntemidir. Bu yöntem  $\omega_j$  ağırlıklarının bilgisini gerektirmeyen bir yaklaşımdır ve olabilirlik oran testi yaklaşımının özel bir durumudur. Alt Bölüm 3.4.2'de bahsedildiği gibi değişimin yoğunluğunu gösteren  $P((\tau = j | \tau \geq j))$  olasılığı sıfıra yaklaştıkça, Eş. 3.4'teki olabilirlik oran yaklaşımı Shiryaev Roberts'ın Eş. 3.5'deki olabilirlik oran yaklaşımına dönüşmektedir (Frisen and Anderson, 2009).

### 3.4.5 Parametrik olmayan ya da yarı parametrik olabilirlik oran yaklaşımı

Shiryaev-Roberts yönteminin kullanımı ile parametrik yaklaşımının gerekliliklerinden biri olan ağırlıkların bilinmesine ihtiyaç kalmamaktadır. Ancak halen süreç ortalamasının  $t$  zaman noktasının parametrik fonksiyonu olarak belirlenmesi koşulu

devam etmektedir. Böyle bir parametrik modelleme ise, grip vakaları gibi yıldan yıla değişken yapı gösteren hastalıkların vaka sayılarının modellenmesinde uygun değildir.

Bu amaçla Frisen ve Anderson (2009) süreç ortalamasının parametrik bir fonksiyonla modellenmesini gerektirmeyen yarı parametrik bir olabilirlik oran yaklaşımı önermişlerdir. Bu yaklaşım, süreç dağılımının düzenli üstel dağılım ailesinin üyesi olmasını ancak salgın başlangıcını belirleyen regresyon modelinin parametrik olmamasını gerektirmektedir. Başka bir ifadeyle, süreç ortalamalarının  $C$  ve  $D$  durumlarındaki sıralarını dikkate alan monotonik regresyon ile ifade edilmesini gerektirmektedir. Kullanılan parametrik olmayan regresyon modeli ise süreç ortalamasının; salgın öncesi sabit, salgın başlangıcı ve sonrasında ise monoton artan olduğunu kabul etmektedir.

$$D \text{ durumu} : \mu(1) = \mu(2) = \dots = \mu(s)$$

$$C1 \text{ durumu} : \mu(1) \leq \mu(2) \leq \dots \leq \mu(s)$$

$$C2 \text{ durumu} : \mu(1) < \mu(2) \leq \dots \leq \mu(s)$$

$$C_j \text{ durumu} : \mu(1) = \dots = \mu(j-1) < \mu(j) \leq \dots \leq \mu(s), \quad j > 2$$

Bu yaklaşım için olabilirlik oranı,

$$\frac{\max f(x_s|C)}{\max f(x_s|D)}$$

biçimindedir. Düzenli bir üstel aileye ait olasılık dağılımı için, olabilirlik fonksiyonlarına maksimum değerleri veren en çok olabilirlik kestiricileri  $\hat{\mu}^D$  ve  $\hat{\mu}^{C_j}$ ,  $D$  ve  $C_j$  durumları için süreç ortalamasına ilişkin sıra kısıtları altında bulunabilmektedir (Frisen and Anderson, 2009).  $\tau = j$  için en çok olabilirlik oranı,

$$\frac{\max f(x_s|C_j)}{\max f(x_s|D)} = \frac{f(x_s; \mu = \hat{\mu}^{C_j})}{f(x_s; \mu = \hat{\mu}^D)}$$

biçimindedir.  $C1$  kümesi diğer tüm  $C_j$  kümelerini kapsadığından,

$$\max f(x_s|C) = \max f(x_s|C1) = f(x_s; \mu = \hat{\mu}^{C1})$$

Frisen ve Anderson (2009) aşağıdaki test istatistiğini önermişlerdir.

$$\frac{\max f(x_s|C)}{\max f(x_s|D)} = \frac{\max f(x_s|C1)}{\max f(x_s|D)}$$

Önerilen bu istatistiğin genelleştirilmiş Cusum istatistiği ile benzerliklere sahip olduğu bu yazarlar tarafından belirtilmektedir. Örneğin Poisson dağılımı için önerilen istatistik aşağıdaki gibidir.

$$\frac{f(x_s; \mu = \hat{\mu}^{C1})}{f(x_s; \mu = \hat{\mu}^D)} = \exp \left\{ \sum_{t=1}^s (\hat{\mu}^D(t) - \hat{\mu}^{C1}(t)) \right\} \cdot \prod_{t=1}^s \left( \frac{\hat{\mu}^{C1}(t)}{\hat{\mu}^D(t)} \right)^{x(t)}$$

### 3.4.6 GLR istatistiklerine dayalı regresyon kontrol grafikleri

Höhle ve Paul (2008), biyoizleme için genelleştirilmiş olabilirlik oran istatistiklerini kullanan regresyon kontrol grafiklerini önermişlerdir. Bu yöntemle ilişkin regresyon modeli genelleştirilmiş doğrusal modellere dayanmakta ve dağılım olarak verilerdeki aşırı yayılım durumuna bağlı olarak Poisson ya da negatif binom dağılımını kullanmaktadır.

Önerilen yöntemi tanımlamak için  $y_1, y_2, \dots$  gözlemlerinin  $f_\theta$  yoğunluk fonksiyonuna sahip bir parametrik dağılımdan geldikleri ve  $\tau$ 'nin sürecin değişim noktası olduğu varsayalım. Bu durumda  $x_t$  bilinen açıklayıcı değişkenler olmak üzere,  $y_t$  için koşullu dağılım aşağıdaki gibi ifade edilir.

$$(y_t|x_t, \tau) \sim \begin{cases} f_{\theta_0}(\cdot|x_t), & t = 1, \dots, \tau - 1 \text{ (kontrol altı durum)} \\ f_{\theta_1}(\cdot|x_t), & t = \tau, \tau + 1, \dots \text{ (kontrol dışı durum)} \end{cases}$$

Biyozlemenin yapıldığı birçok durumda  $f$ 'nin zamana bağlı olarak değişmeyen ve bilinen yayılım parametresi  $\alpha$  ile negatif binom dağılımlı olduğu kabul edilmektedir. Ancak  $\alpha \rightarrow 0$  iken bu dağılım Poisson dağılımına dönüşür. Ayrıca kontrol altı

ortalama  $\mu_{0,t}$  ve kontrol dışı ortalama  $\mu_{1,t}$  ile gösterilsin. Bu ortalamalar sırasıyla  $\theta_0$  ve  $\theta_1$ 'in fonksiyonları olsun. Regresyon kontrol grafikleri uygulamalarında genellikle kontrol dışı ortalamanın  $\mu_{1,t} = \mu_{0,t}e^{\lambda t}$  ya da  $\lambda > 0$  olduğu durumda  $\mu_{1,t} = \mu_{0,t} + \lambda y_{t-1}$  biçiminde olduğu kabul edilir. Kontrol altındaki ortalama  $\mu_{0,t}$  için regresyon modeli genel olarak aşağıdaki gibi ifade edilir.

$$\log \mu_{0,t} = \beta_0 + \beta_1 t + \sum_{s=1}^S [\beta_{2s} \cos(\omega st) + \beta_{2s+1} \sin(\omega st)]$$

Burada  $S$  harmonik dalgaların sayısını,  $\omega = 2\pi/T$  ve  $T$  ise verilerin döngü periyodunu göstermektedir. Ayrıca bu modele başka açıklayıcı değişkenlerde dahil edilebilir ve aday modeller arasında en iyi model seçilir.

Oluşturulacak bir regresyon grafiğinde amaç değişim noktası  $\tau$ 'yu online olarak belirlemektir. Başka bir ifadeyle  $H_0: \mu_t = \mu_{0,t}, t = 1, 2, \dots$  hipotezinin ret edildiği ana kadar süreci izlemektir (Höhle ve Paul, 2008).  $y_s = \{y_t: t \leq s\}$ ,  $s$  karar anına kadar olan gözlemleri göstermek üzere;  $y_t$  gözlemlerinin bağımsız olması şartıyla,  $H_0$  hipotezinin testi için olabilirlik oran istatistiği aşağıdaki gibidir.

$$L(s, t) = \frac{f(y_s | \tau = t)}{f(y_s | \tau > s)} = \frac{\prod_{i=1}^{t-1} f(y_i; \theta_0) \prod_{i=t}^s f(y_i; \theta_1)}{\prod_{i=1}^s f(y_i; \theta_0)} = \prod_{i=t}^s \frac{f(y_i; \theta_1)}{f(y_i; \theta_0)}$$

Shewhart türü grafikler,  $L(s, s)$  istatistiğini değişim noktasını bulmak amacıyla kullanılmaktadırlar. Bu grafikler büyük süreç kaymalarını kısa sürede belirlemede iyidirler. Ancak süreçteki kayma küçük ve kalıcı ise, zaman içindeki süreç kaymalarının birikimli etkisi değişim noktasının belirlenmesi için önem arz etmektedir. Bu amaçla önerilen yöntemlerden biri de Page (1954) olabilirlik oranına dayalı Cusum grafiğidir. Bu grafik istatistik olarak  $\max\{1 \leq t \leq s: L(s, t)\}$  ifadesini kullanmaktadır (Höhle ve Mazick, 2010).

Ancak biyoizleme verileri, aynı dağılımlı olma varsayımını ihlal eden mevsimsel değişimler ve trend içerebilmektedir. Bundan dolayı Höhle ve Paul (2008) Cusum

grafiğinde,  $L(s, t)$  olabilirlik oranının doğal logaritmasının maksimumunun istatistik olarak kullanılmasını önermişlerdir. Başka bir ifadeyle bu olabilirlik oran istatistiğini kullanan grafiğin sinyal verme kuralını aşağıdaki gibi tanımlamışlardır.

$$N = \inf \left\{ s \geq 1: \max_{1 \leq k \leq s} \left[ \sum_{t=k}^s \log \left\{ \frac{f_{\theta_1}(y_t|x_t)}{f_{\theta_0}(y_t|x_t)} \right\} \right] \geq c_\gamma \right\}$$

Kontrol altı parametresi  $\theta_0$  ve kontrol dışı parametresi  $\theta_1$  (yani  $\mu_{1,t}$ 'nin  $\mu_{0,t}$  cinsinden ifadesi edilişindeki  $\kappa$  ya da  $\lambda$ ) bilinmek şartıyla, bu istatistik ve bu istatistiğe bağlı karar verme kuralı aşağıdaki gibi ifade edilebilmektedir.

$$l_0 = 0, \quad l_s = \max \left( 0, l_{s-1} + \log \left\{ \frac{f_{\theta_1}(x_s)}{f_{\theta_0}(x_s)} \right\} \right), \quad s \geq 1 \quad \text{ve} \quad N = \inf \{s: l_s \geq c_\gamma\}$$

Bu tür regresyon kontrol grafiklerinin uygulamasında, veriler genellikle birinci aşama ve ikinci aşama olmak üzere ikiye bölünmektedir. Birinci aşamada genelleştirilmiş doğrusal regresyon modeline ilişkin parametreler tahmin edilmektedir. İkinci aşamada ise, bu tahminlere dayanarak kontrol dışı durumlarının varlığını tespit etmek amacıyla izleme yapılmaktadır. Kontrol dışı parametresi  $\theta_1$  (yani  $\mu_{1,t}$ 'nin  $\mu_{0,t}$  cinsinden ifadesi edilişindeki  $\kappa$  ya da  $\lambda$ ) bilinmiyorsa, izleme aşamasında tahmin edilmesi gerekmektedir. Böyle bir durumda kullanılan istatistik genelleştirilmiş olabilirlik oranı haline dönüşmektedir ve bu istatistiğe ilişkin sinyal verme kuralı aşağıdaki gibi ifade edilmektedir (Höhle ve Paul, 2008).

$$N_G = \inf \left\{ s \geq 1: \max_{1 \leq k \leq s} \sup_{\theta \in \Theta} \left[ \sum_{t=k}^s \log \left\{ \frac{f_\theta(y_t|x_t)}{f_{\theta_0}(y_t|x_t)} \right\} \right] \geq c_\gamma \right\}$$

Bu grafiklerde kullanılacak eşik değeri ise Monte Carlo simülasyonları ya da Markov zinciri yardımıyla, kontrol altında istenilen bir ARL değerini sağlayacak şekilde hesaplanabilir.

## 4. MATERYAL VE YÖNTEM

Bu tez kapsamında kullanılacak veriler ABD'nin CDC (Center for Disease Control and Prevention) kurumunun internet sitesinden alınmıştır. Veriler 2012-2016 yılları arasındaki haftalık grip ya da grip benzeri hastalıklarla ayakta tedavi olunan hasta sayılarından oluşmaktadır. Ek A.'da ilgili veriler verilmektedir. ABD'de grip sezonu bir yılın 40. haftasında başlayıp diğer yılın 39. haftasında sonlanmaktadır. Bu çalışma kapsamında 2012-2013, 2013-2014 ve 2014-2015 grip mevsimlerine ilişkin haftalık veriler referans (baseline) olarak alınarak, 2015-2016 grip mevsimi için haftalık ileriye dönük izleme yapılacaktır. İzleme amaçlı kullanılacak yöntemler bu bölümde belirtilecektir. Ayrıca bir simülasyon çalışmasıyla performansı karşılaştırılacak yöntemler de bu bölümde ifade edilecektir.

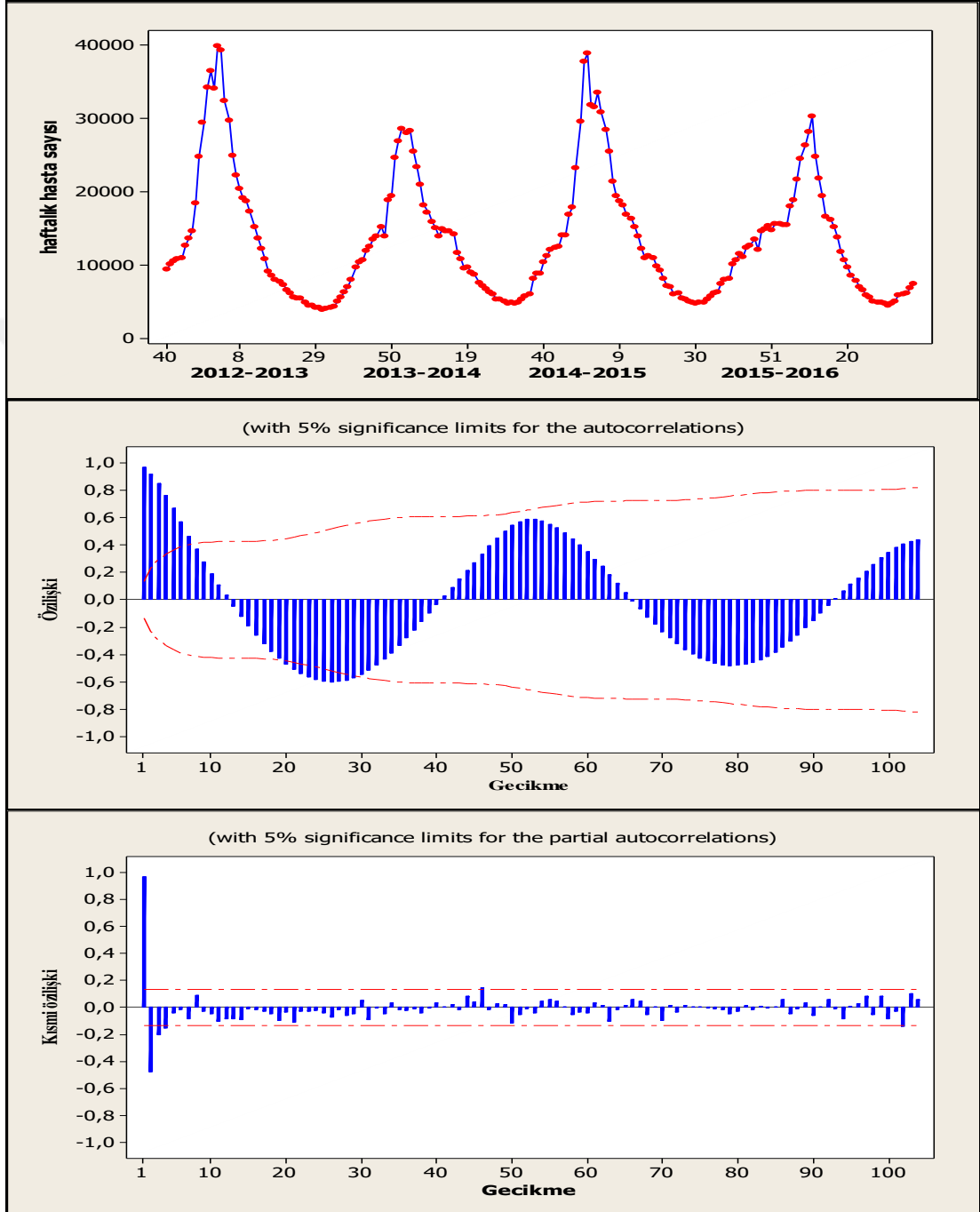
### 4.1 Verilerin Özellikleri

Analiz edilecek veriler kesikli olup, haftalık grip ve grip benzeri belirtilere sahip kişi sayısını göstermektedir. Ayrıca veriler mevsimsellik özelliğine sahiptir. Bu durum Şekil 4.1.'deki zaman serisi grafiğinden görülmektedir. Şekil 4.1. incelendiğinde gözle görülebilir trend eğilimi yoktur. Bir başka özellik ise Şekil 4.1.'deki özilişki ve kısmi özilişki grafiklerinden anlaşılacağı gibi verilerin özilişkili olmasıdır. Ayrıca gecikmeler arttıkça özilişki değerlerinin hızlıca erimemesi verilerin durağan olmadığı izlemine uyandırmaktadır. Çizelge 4.1.'deki durağanlık testlerinden ADF (Augmented Dickey-Fuller) testi verilerin durağan olduğunu belirtirken, PP (Phillips Perron) testi verilerin durağan olmadığını belirtmektedir.

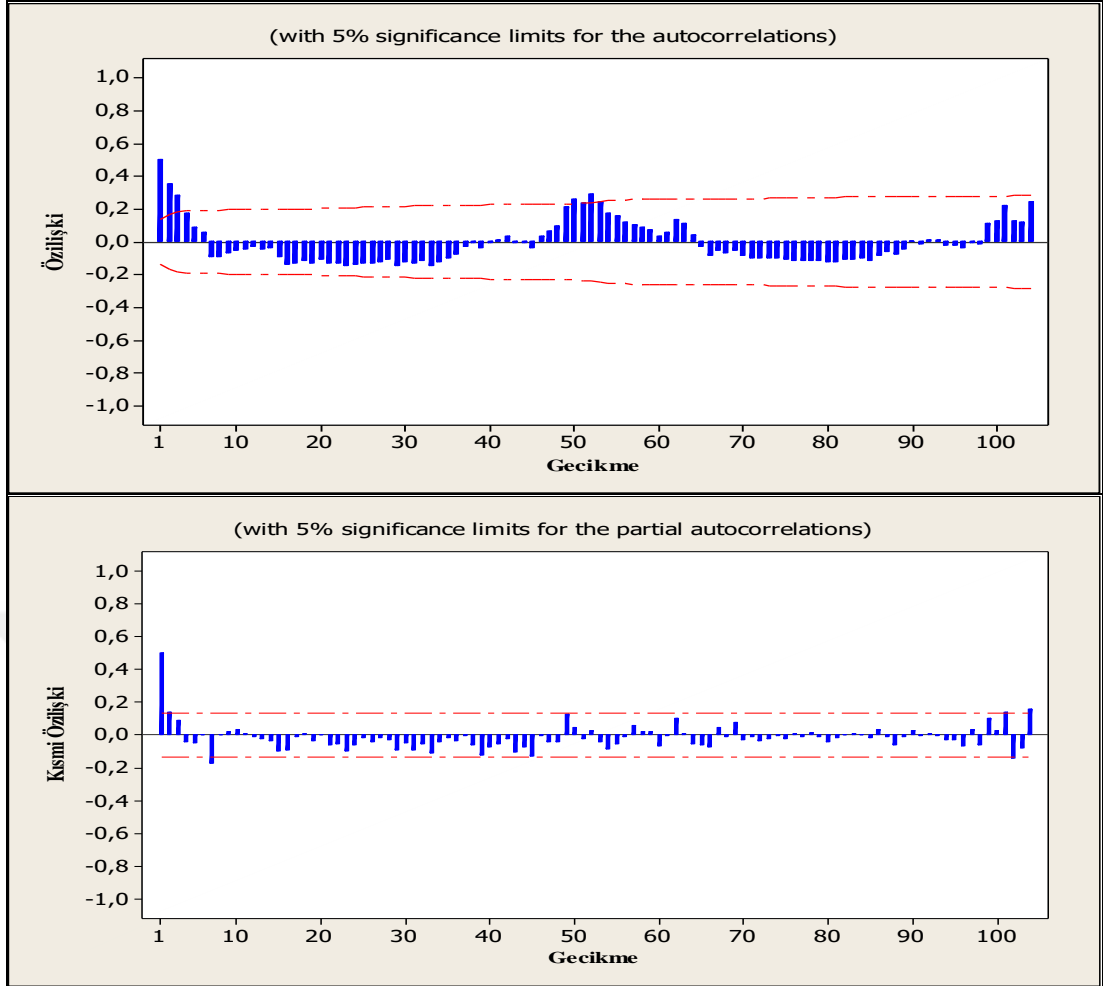
İlgili verilerin 1. farklarına ilişkin özilişki ve kısmi özilişki grafikleri Şekil 4.2.'de verilmektedir. Şekil 4.2. 1.fark işleminin verileri durağan olma eğilimine yaklaştırdığı



izlemi vermektedir. Çizelge 4.1.'de sonuçları verilen ADF ve PP testleri de 1. farkların durağan olduğunu desteklemektedir. Ayrıca Şekil 4.2.'deki özilişki ve kısmi özilişki grafiklerinden mevsimsel bir etkinin olduğuda anlaşılmaktadır.



Şekil 4.1. 2012-2016 yıllarına ilişkin haftalık hasta sayısının zaman serisi, özilişki ve kısmi özilişki grafikleri



Şekil 4.2. 2012-2016 yıllarına ilişkin haftalık hasta sayısının 1. farklarının özilişki ve kısmi özilişki grafikleri

Çizelge 4.1. 2012-2016 grip mevsimlerine ilişkin ADF ve PP durağanlık testleri sonuçları

Test edilen veri	ADF Test			PP test		
	test istatistiği	gecikme derecesi	p değeri	test istatistiği	gecikme derecesi	p değeri
orjinal veriler	-3.7334	5	0.02342	-13.983	4	0.3197
1. fark alınmış veriler	-4.6121	5	0.01	-108.9	4	0.01

#### 4.2 Verilerde İzleme Amaçlı Kullanılacak İleriye Dönük İzleme Yöntemleri

Bu tez çalışması kapsamında 2015-2016 grip mevsimi verilerini haftalık olarak ileriye dönük izlemek amacıyla 5. Bölüm’de aşağıdaki yöntemler dikkate alınacaktır.

- Serfling Regresyonu
- Süreç Kontrol Grafikleri
- Ears C1, C2 ve C3
- Farrington ve FarringtonFlexible
- Bayeşçi Salgın Belirleme Algoritması (Boda)
- Bayeşçi bir sonsal öngörü yaklaşımı (Bayes)
- Rki
- Holt-Winters
- Arima modelleme yaklaşımı
- Wavelet yaklaşımı
- OutbreakP
- GLR regresyon grafiđi yaklaşımı
- Gizli Markov Modeli yaklaşımı
- UCengiz yöntemi (Bu çalıřma kapsamında önerilen izleme yöntemi)

#### **4.3 Simülasyonlarda Performansları Karşılaştırılacak İleriye Dönük İzleme Yöntemleri**

Bu tez çalıřması kapsamında yapılacak simülasyonlarda ařađıda belirtilen yöntemlerin performansları ATFOS ve FMO ölçütlerine göre 6. Bölüm’de karşılaştırılacaktır.

- Serfling Regresyonu
- Cusum
- Ears C1, C2 ve C3
- Farrington ve FarringtonFlexible
- Bayeşçi bir sonsal öngörü yaklaşımı (Bayes)
- OutbreakP
- GLR regresyon grafiđi yaklaşımı
- UCengiz yöntemi (Bu çalıřma kapsamında önerilen izleme yöntemi)

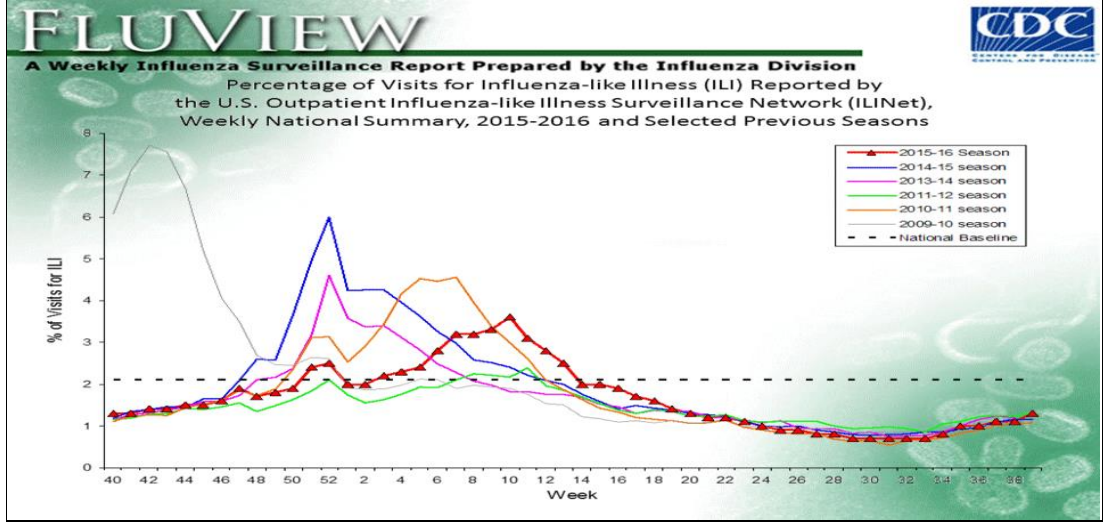
#### 4.4 Karşılaştırma Ölçütleri

2015-2016 grip mevsimi verilerine ilişkin ileriye dönük izleme yöntemlerinden elde sonuçlar, CDC'nin 2015-2016 yılına ilişkin haftalık salgın kararları ile (altın standart, benchmarking) karşılaştırılacaktır. Başka bir ifadeyle ilgili izleme döneminde CDC tarafından alınan kararların gerçek salgın durumunu belirttiği kabul edilecektir. CDC kurumu, herhangi bir grip izleme dönemine ilişkin haftalık kararlarını son 3 grip mevsimine dayandırmaktadır. Bu 3 grip mevsiminde grip olmayan haftalardaki grip ve benzeri bulgular nedeniyle ayakta tedavi gören hasta sayısı oranının ortalamasını dikkate almaktadır ([www.cdc.gov/flu/weekly/overview.htm](http://www.cdc.gov/flu/weekly/overview.htm)). Bu ortalama 2 standart sapma eklenerek ulusal referans bulunmakta ve bu referansı geçen haftalar için olağan dışı grip faaliyeti olduğu kararı verilmektedir. Çizelge 4.2.'de 2015-2016 grip izleme dönemi için haftalık olarak grip faaliyeti durumu verilmektedir. Bu bilgiler görsel olarak da Şekil 4.3.'de ifade edilmektedir.

**Çizelge 4.2. CDC kurumunun 2015-2016 grip mevsimi için haftalık grip faaliyet durumu**

Hafta	Karar	Hafta	Karar	Hafta	Karar	Hafta	Karar
40	Normal	1	Normal	14	Normal	27	Normal
41	Normal	2	Normal	15	Normal	28	Normal
42	Normal	3	Olağandışı	16	Normal	29	Normal
43	Normal	4	Olağandışı	17	Normal	30	Normal
44	Normal	5	Olağandışı	18	Normal	31	Normal
45	Normal	6	Olağandışı	19	Normal	32	Normal
46	Normal	7	Olağandışı	20	Normal	33	Normal
47	Normal	8	Olağandışı	21	Normal	34	Normal
48	Normal	9	Olağandışı	22	Normal	35	Normal
49	Normal	10	Olağandışı	23	Normal	36	Normal
50	Normal	11	Olağandışı	24	Normal	37	Normal
51	Olağandışı	12	Olağandışı	25	Normal	38	Normal
52	Olağandışı	13	Olağandışı	26	Normal	39	Normal

2015-2016 grip mevsimi verilerine ilişkin ileriye dönük izleme yöntemlerinden elde sonuçlar, CDC'nin 2015-2016 yılına ilişkin haftalık salgın kararları ile karşılaştırılırken Çizelge 4.3.'deki ölçütler dikkate alınacaktır.



Şekil 4.3. Bazı grip sezonlarına ilişkin haftalık ILI kaynaklı ayakta tedavi edilen hasta oranlarının zaman serisi grafiği

Çizelge 4.3. Yöntemlerin verdiği kararlar için sınıflandırma tablosu

Verilen Karar	Gerçek Durum		Toplam
	Normal	Olağandışı	
Normal	TN=D	FN=C	C+D
Olağandışı	FP=B	TP=A	A+B
Toplam	B+D	A+C	N
Seçicilik=D/(B+D)		Duyarlılık=A/(A+C)	

Burada TP, gerçek pozitif; FP, yanlış pozitif; TN, gerçek negatif; FN, yanlış negatiftir. Ayrıca duyarlılık ve seçicilik de Çizelge 4.3.'de belirtildiği gibi hesaplanacaktır

#### 4.5 Tez Çalışması Kapsamında Önerilen UCengiz Yöntemi

Bu çalışma kapsamında basit bir istatistiksel izleme yöntemi önerilmektedir. Bu yöntem zaman içinde ileriye doğru izleme yapabilmekte ve basit bir yapısı nedeniyle salgın belirtilerini ifade eden çok sayıda zaman serisinin birlikte izlenmesine olanak sağlamaktadır. Ayrıca, UCengiz olarak adlandırılan bu yöntem, az sayıda gözlemden oluşan ve kayan kısa bir baseline (referans dönemi) kullanmaktadır.

UCengiz yöntemi iki parametreye bağlı olarak uygulanmaktadır. Bu parametrelerden ilki  $k$ , referans döneminin kaç gözlemden oluştuğunu belirlerken, ikinci parametre

duruma bağılı olarak Poisson ya da negatif binom dağılımının üst kaçıncı yüzdeliğinin kullanılacağını belirleyen  $\alpha$  değeridir. Bu yöntemin işleyişi adım adım aşağıda belirtilmektedir.

1. İzleme başlamadan hemen önceki  $k$  adet gözlemin ya da ilk izleme öncesine en yakın  $k$  adet gözlemin, salgın olmayan döneme ilişkin gözlemler olduğu kabul edilir. Başka bir ifadeyle, ilk izlemenin yapılacağı  $Y_1$  gözlemi için bu gözlemler referans oluşturur.
2. İlk izleme gözlemi için ortalama ve varyans bilgisinin, kendisinden önce gelen referans dönemi gözlemlerinden elde edileceği varsayılır. İlk referans dönemine ilişkin  $k$  adet gözlem, ilk gözlem olan  $Y_1$  gözlemine yakınlıklarına göre ağırlıklandırılarak ortalama ve varyans aşağıdaki gibi hesaplanır:

$$\bar{Y}_1 = \frac{\sum_{i=-(k-1)}^0 (k+i)Y_i}{\sum_{i=1}^k i} = \frac{Y_{1-k} + 2Y_{2-k} + 3Y_{3-k} + \dots + kY_0}{k(k+1)/2}$$

$$S_1^2 = \frac{\sum_{i=-(k-1)}^0 (Y_i - \bar{Y}_1)^2 (k+i)}{\sum_{i=1}^k i} = \frac{(Y_{1-k} - \bar{Y}_1)^2 + 2(Y_{2-k} - \bar{Y}_1)^2 + 3(Y_{3-k} - \bar{Y}_1)^2 + \dots + k(Y_0 - \bar{Y}_1)^2}{k(k+1)/2}$$

3.  $S_1^2 \leq \bar{Y}_1$  ise; ilk izleme gözlemi  $Y_1$  için kritik değer,  $\lambda_1 = \bar{Y}_1$  parametrelili Poisson dağılımının üst  $\alpha$  yüzdeliği olarak belirlenir:  $kd_1 = Poisson_{\lambda_1, \alpha}$ . Ancak  $S_1^2 > \bar{Y}_1$  ise; ilk izleme gözlemi  $Y_1$  için kritik değer,  $p_1 = \bar{Y}_1/S_1^2$  ve  $r_1 = \bar{Y}_1^2/(S_1^2 - \bar{Y}_1)$  parametrelili negatif binom dağılımının üst  $\alpha$  yüzdeliği olarak belirlenir:  $kd_1 = NegatifBinom_{r_1, p_1, \alpha}$ .
4. İlk izleme gözlemi  $Y_1$  bir önceki adımda belirlenen kritik değerden büyükse yani  $Y_1 > kd_1$  ise olağan dışı durum sinyali verilir.
5. Eğer olağan dışılık sinyali verilirse,  $k$  adet gözlemden oluşan referans dönemi aynı kalır. Olağan dışılık sinyali verilmezse, ilk izleme gözlemi  $Y_1$  referans dönemine dahil edilirken, referans döneminin ilk gözlemi  $Y_{1-k}$  referans döneminden çıkartılır. Böylelikle sinyal verilmediği sürece zaman ekseninde kayan ve  $k$  gözlemden oluşan bir referans kullanılır.

6. Herhangi bir izleme gözlemi  $Y_t$  için olağan olup olmadığı kararı verilirken, benzer adımlar izlenir:  $i_1 > i_2 > \dots > i_k$  pozitif tam sayılar olsun.  $Y_t$  için  $k$  gözlemlik referans dönemi; bu gözlemden önce gelen ve olağan dışı durum sinyali verilmemiş  $Y_t$  gözlemine en yakın  $k$  adet  $Y_{t-i_1}, Y_{t-i_2}, \dots, Y_{t-i_k}$  gözlemlerinden oluşsun.
7. İlk olarak ortalama ve varyans hesaplanır:

$$\bar{Y}_t = \frac{Y_{t-i_1} + 2Y_{t-i_2} + 3Y_{t-i_3} + \dots + kY_{t-i_k}}{k(k+1)/2}$$

$$S_t^2 = \frac{(Y_{t-i_1} - \bar{Y}_t)^2 + 2(Y_{t-i_2} - \bar{Y}_t)^2 + 3(Y_{t-i_3} - \bar{Y}_t)^2 + \dots + k(Y_{t-i_k} - \bar{Y}_t)^2}{k(k+1)/2}$$

8.  $S_t^2 \leq \bar{Y}_t$  ise;  $Y_t$  gözlemi için kritik değer,  $\lambda_t = \bar{Y}_t$  parametrelili Poisson dağılımını üst  $\alpha$  yüzdeliği olarak belirlenir:  $kd_t = Poisson_{\lambda_t, \alpha}$ . Ancak  $S_t^2 > \bar{Y}_t$  ise;  $Y_t$  gözlemi için kritik değer,  $p_t = \bar{Y}_t/S_t^2$  ve  $r_t = \bar{Y}_t^2/(S_t^2 - \bar{Y}_t)$  parametrelili negatif binom dağılımının üst  $\alpha$  yüzdeliği olarak belirlenir:  $kd_t = NegatifBinom_{r_t, p_t, \alpha}$ .
9.  $Y_t$  kendisi için belirlenen kritik değerden büyükse yani  $Y_t > kd_t$  ise, olağan dışı durum sinyali verilir.
10. Eğer olağan dışılık sinyali verilirse,  $k$  adet gözlemden oluşan referans dönemi aynı kalır. Olağan dışılık sinyali verilmezse;  $Y_t$  gözlemi referans dönemine dahil edilir ve referans döneminin ilk gözlemi  $Y_{t-i_1}$  referans döneminden çıkartılır. İzleme işlemleri benzer biçimde devam eder.

UCengiz yöntemi için referans dönemlerinin, izleme yapılacak herhangi bir  $t$  anından hemen önceki gözlemlerden oluşması ya da oluşma eğiliminden dolayı, bu yöntemin verilerdeki olası mevsimsellik etkisini, özilişkiyi ve aşırı yayılım özelliği gibi özellikleri kısmen dikkate alma eğilimi vardır. Ayrıca, UCengiz yöntemi verilerin Poisson ya da negatif binom dağılımlı olduğunu varsayarak, izleme verilerin kesikli olan doğasını dikkate almaktadır.

## 5. UYGULAMA

Bu bölümde 2012-2013, 2013-2014, 2014-2015 grip mevsimleri referans alınarak 2015-2016 grip mevsimi verileri haftalık olarak çeşitli yöntemler kullanılarak izlenecektir. Böylelikle bioizlemede kullanılan bazı ileriye dönük salgın belirleme yöntemleriyle, bu tez çalışması kapsamında önerilen UCengiz yönteminin uygulamaları yapılacaktır. Bu uygulamaları gerçekleştirmek amacıyla kullanılan R ve Matlab program komutları Ek L.'deki çizelgelerde verilmektedir.

### 5.1 Regresyon Türü Yöntemlerin CDC Grip Verilerine Uygulanması

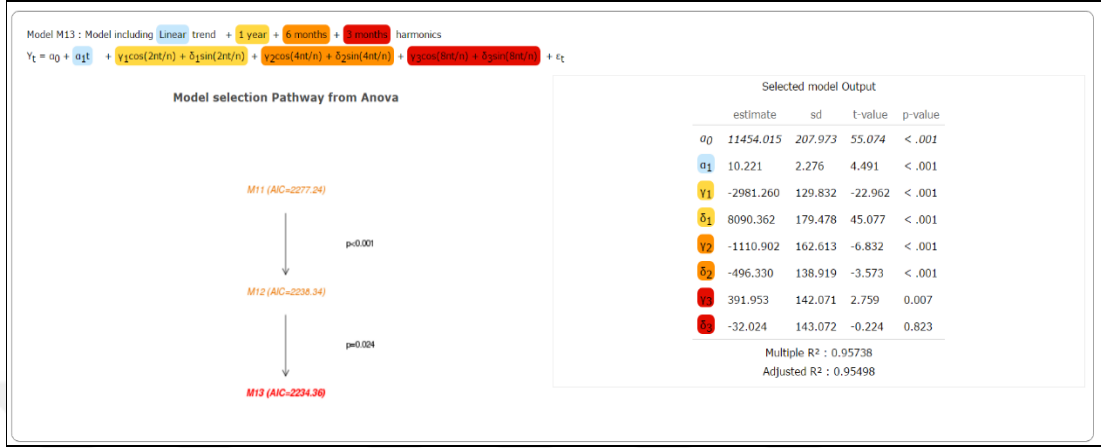
Bu alt bölümde Serfling regresyon, Farrington, FarringtonFlexible ve Boda yöntemleri, CDC kurumunun 2015-2016 grip mevsimi haftalık verilerinin izlenmesi amacıyla uygulanmaktadır.

#### 5.1.1 Serfling regresyon yönteminin CDC grip verilerine uygulanması

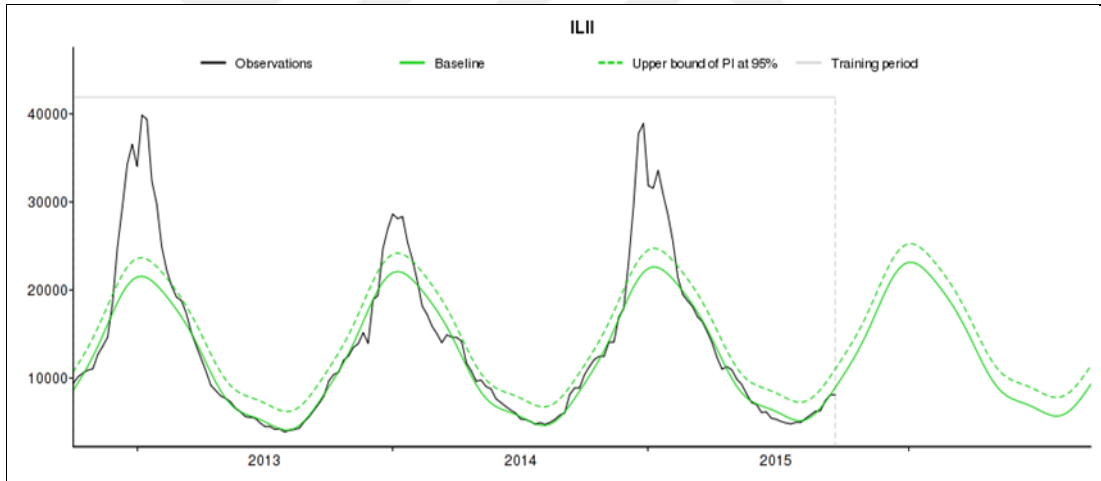
Ek A.'daki 2012-2013, 2013-2014 ve 2014-2015 grip mevsimleri verileri referans yılları alınarak, AIC kriterine göre en iyi Serfling regresyon modeli ve bu modele ilişkin parametre kestirimleri Şekil 5.1.'de verilmektedir. Bu model kullanılarak izleme dönemi olan 2015-2016 grip mevsimi için 52 haftalık öngörü (prediction) ve eşik değerleri elde edilmiştir (Ek B.). Eşik değerleri %95 düzeyli tahmin aralığının üst sınırını göstermektedir. Referans ve izleme dönemlerine ilişkin tahmin ve eşik değerleri Şekil 5.2.'de gösterilmektedir. Bütün bu işlemler Pierre Louis Araştırma Enstitüsü'ndeki bir ekip tarafından geliştirilen <http://marne.u707.jussieu.fr/periodic> adresindeki web uygulaması üzerinden gerçekleştirilmiştir. Ayrıca, referans



yıllarındaki olası salgınlardan kaynaklanabilecek yüksek vaka sayılarının etkilerini azaltmak amacıyla, model parametreleri kestirilmeden önce verilerin üst %15'lik kısmı analizden dışlanmıştır.



Şekil 5.1. AIC ölçütüne göre en iyi Serfling model ve parametrelerinin kestirimleri



Şekil 5.2. Referans ve izleme dönemleri için tahmin ve eşik değerleri

İzleme dönemi verileri ile eşik değerleri farkları alınarak (Ek B.), pozitif farklara karşılık gelen haftalar olağandışı grip faaliyeti olarak belirlenmiştir (Çizelge 5.1.). Elde edilen sonuçlar CDC kurumunun 2015-2016 grip mevsimi için belirlediği haftalık salgın durumlarıyla karşılaştırılmıştır (Çizelge 5.2.).

**Çizelge 5.1. Serfling regresyona ilişkin 2015-2016 grip mevsimi için haftalık grip faaliyet durumu**

Hafta	Karar	Hafta	Karar	Hafta	Karar	Hafta	Karar
40	Normal	1	Normal	14	Normal	27	Normal
41	Normal	2	Normal	15	Olağandışı	28	Normal
42	Normal	3	Normal	16	Olağandışı	29	Normal
43	Normal	4	Normal	17	Normal	30	Normal
44	Normal	5	Normal	18	Normal	31	Normal
45	Normal	6	Normal	19	Normal	32	Normal
46	Normal	7	Olağandışı	20	Normal	33	Normal
47	Normal	8	Olağandışı	21	Normal	34	Normal
48	Normal	9	Olağandışı	22	Normal	35	Normal
49	Normal	10	Olağandışı	23	Normal	36	Normal
50	Normal	11	Olağandışı	24	Normal	37	Normal
51	Normal	12	Olağandışı	25	Normal	38	Normal
52	Normal	13	Olağandışı	26	Normal	39	Normal

**Çizelge 5.2. Serfling yönteminin verdiği kararlar için sınıflandırma tablosu**

Yöntem	Karar	Gerçek Durum(CDC)		Toplam	Duyarlılık	Seçicilik
		Normal	Olağandışı			
Serfling	Normal	TN=36	FN=6	42	0.5385	0.9231
	Olağandışı	FP=3	TP=7	10		
	Toplam	39	13	52		

### 5.1.2 Farrington ve FarringtonFlexible yönteminin CDC grip verilerine uygulanması

Bu alt bölümde Farrington vd. (1996)'nin bir nevi bağımsızlık gerektiren log doğrusal regresyon modeli ve bu modelin iyileştirilmiş hali Noufaily vd. (2012) tarafından önerilen FarringtonFlexible, CDC verilerine uygulanacaktır. Bu yöntemler mevsimselliği, trendi, verilerdeki aşırı yayılımı (overdispersion) dikkate almaktadır. Ayrıca, referans değerlerinde bulunabilecek geçmişteki salgınların etkisini azaltmak amacıyla, genelleştirilmiş doğrusal modelden elde edilen büyük artıklara sahip olan referans değerleri düşük değerlerle ağırlıklandırılmaktadır.

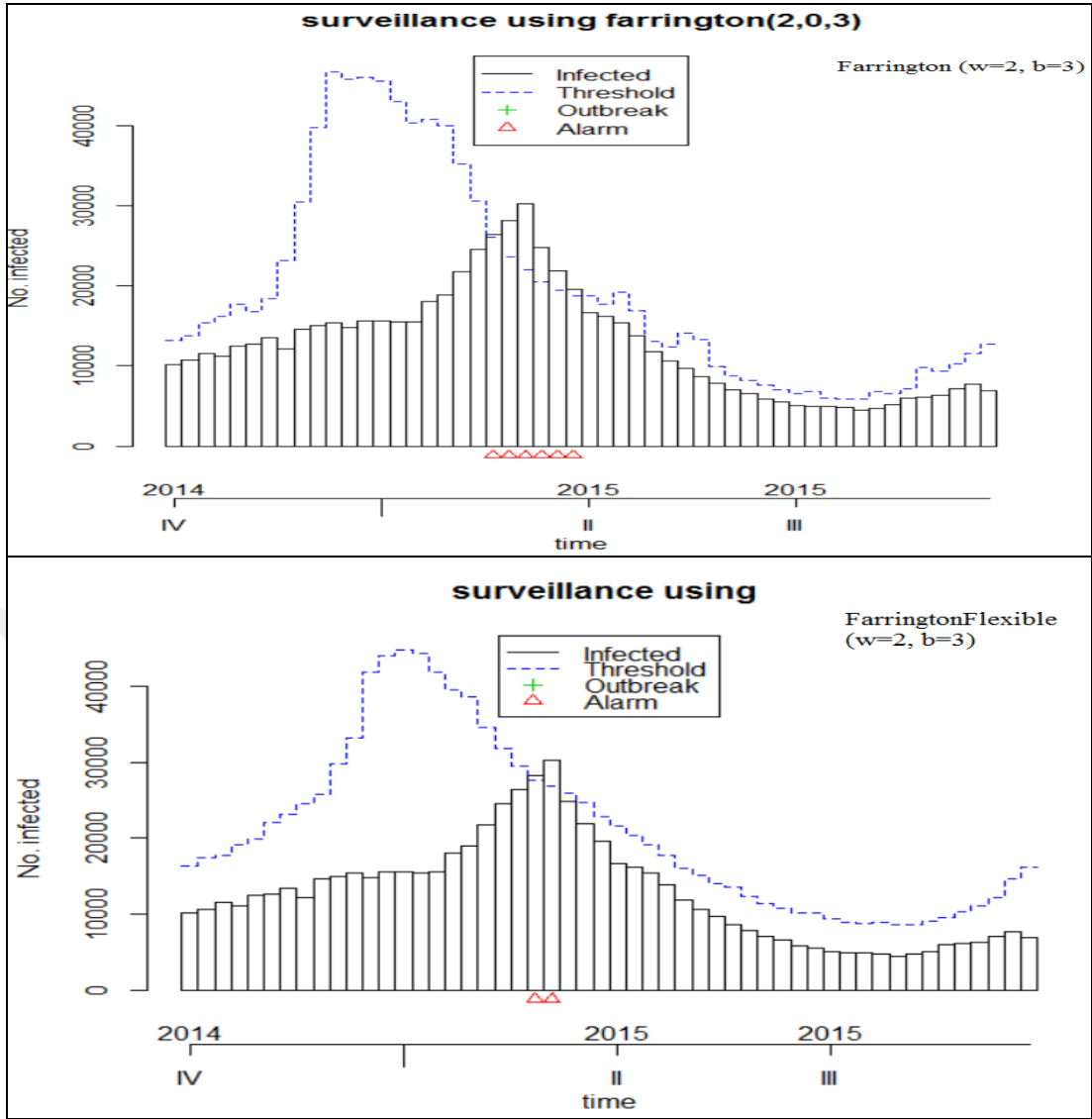
İlk olarak orijinal Farrington yöntemi daha sonra da bu yöntemin geliştirilmiş hali FarringtonFlexible yöntemiyle, 2012-2013, 2013-2014 ve 2014-2015 grip mevsimleri referans dönemi alınarak, 2015-2016 yılı grip mevsiminin ileriye dönük olarak izlenmesi gerçekleştirilmiştir. Pencere genişliği parametresi  $w = 2$ , zamanda geriye

gidilen yıl sayısı  $b = 3$ , öngörü aralığı için  $\alpha = 0.01$  alınmıştır. Bu işlem için R yazılımındaki “surveillance” paketi kullanılmıştır. Şekil 5.3. bu yöntemlere ilişkin izlemeleri görsel olarak göstermekte ve Çizelge 5.3. ise bu izleme sonucunda ilgili yöntemler için 2015-2016 grip mevsimine ilişkin haftalık faaliyet durumunu vermektedir. Ayrıca Ek C.’de her iki yönteme ilişkin eşik değerleri verilmektedir.

**Çizelge 5.3. Farrington ve FarringtonFlexible yöntemlerine ilişkin 2015-2016 grip mevsimi için haftalık grip faaliyet durumu**

Hafta	Karar	Hafta	Karar	Hafta	Karar	Hafta	Karar
40	Normal	1	Normal	14	Normal	27	Normal
41	Normal	2	Normal	15	Normal	28	Normal
42	Normal	3	Normal	16	Normal	29	Normal
43	Normal	4	Normal	17	Normal	30	Normal
44	Normal	5	Normal	18	Normal	31	Normal
45	Normal	6	Normal	19	Normal	32	Normal
46	Normal	7	Olağandışı(F)	20	Normal	33	Normal
47	Normal	8	Olağandışı(F, FF)	21	Normal	34	Normal
48	Normal	9	Olağandışı(F, FF)	22	Normal	35	Normal
49	Normal	10	Olağandışı(F)	23	Normal	36	Normal
50	Normal	11	Olağandışı(F)	24	Normal	37	Normal
51	Normal	12	Olağandışı(F)	25	Normal	38	Normal
52	Normal	13	Normal	26	Normal	39	Normal

Çizelge 5.3. incelendiğinde, FarringtonFlexible yönteminin beklenildiği gibi Farrington yöntemine göre daha az sinyal verdiği dikkat çekmektedir. Şekil 5.3.’den grip ve grip benzeri hasta sayısının 3. haftadan itibaren dikkate değer bir artışa geçtiği görülmesine rağmen, yöntemlerin bu artışa olan reaksiyonun yavaş olduğu ve neredeyse en yüksek seviyeye ulaştığı haftalarda olağandışı sinyallerin verildiği gözlemlenmektedir. Ayrıca elde edilen sonuçlar Çizelge 5.4.’te CDC kurumunun 2015-2016 grip mevsimi için belirlediği haftalık salgın durumlarıyla karşılaştırılmıştır.



Şekil 5.3. Farrington ve FarringtonFlexible yönteminden elde edilen izleme grafikleri

Çizelge 5.4. Farrington yöntemlerinin verdiği kararlar için sınıflandırma tablosu

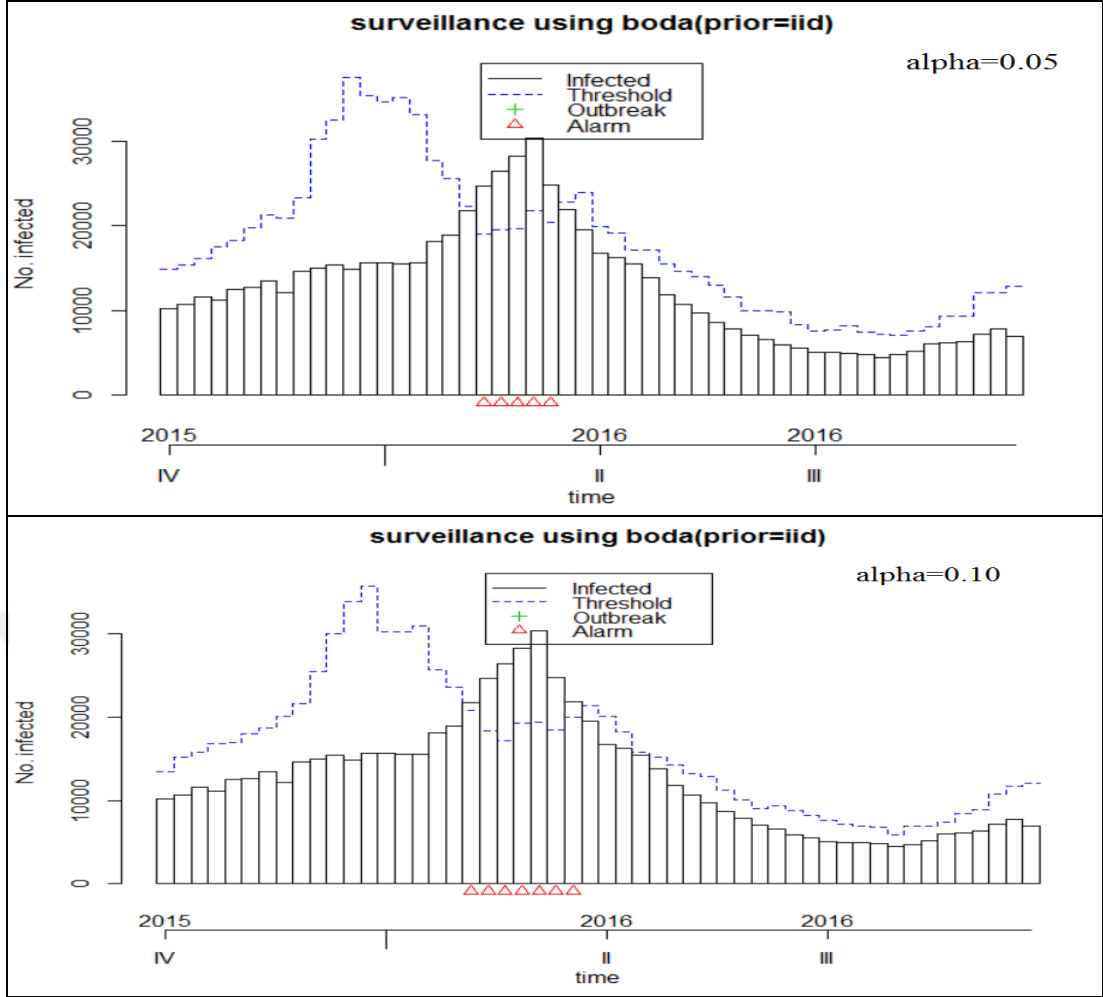
Yöntem	Karar	Gerçek Durum(CDC)		Toplam	Duyarlılık	Seçicilik
		Normal	Olağandışı			
Farrington	Normal	TN=39	FN=7	46	0.4615	1.0000
	Olağandışı	FP=0	TP=6	6		
	Toplam	39	13	52		
F. Flexible	Normal	TN=39	FN=11	50	0.1538	1.0000
	Olağandışı	FP=0	TP=2	2		
	Toplam	39	13	52		

### 5.1.3 Boda yönteminin CDC grip verilerine uygulanması

Bu bölümde bayesci bir yaklaşım olan boda yöntemi CDC grip verilerine uygulanmaktadır. Bu yöntem Farrington ve FarringtonFlexible yöntemlerinin Bayesci yaklaşıma sahip olan karşılığıdır ve açıklayıcı değişkenler istenirse modele kolaylıkla dahil edilebilmektedir (Salmon vd., 2016). Ayrıca verilerdeki trend, mevsimsellik etkisi, geçmiş salgınların etkisi, zamana bağlı süreç seviyesi ayarlaması ve istenirse açıklayıcı değişkenlerin etkisi dikkate alınmaktadır.

2015-2016 grip mevsimi izleme dönemi olarak alınmıştır. Geriye dönük olarak gidilebilecek en son zaman noktası 2012-2013 grip mevsiminin başlangıcı olan 40. hafta olmak üzere, R programındaki surveillance paketindeki boda algoritması kullanılmıştır. Bu yöntemde 2012-2013, 2013-2014, 2014-2015 geçmiş grip mevsimlerine ilişkin haftalık grip faaliyet durumu CDC sitesinden elde edilerek boda yönteminin algoritmasına girdi olarak verilirken, 2015-2016 izleme dönemi için ise haftalık grip faaliyet durumları olağan olarak bu algoritmaya verilmiştir.

Öngörü sonsal dağılımının  $(1 - \alpha)$ 'ncü kuantilini belirlemek amacıyla algoritmaya  $\alpha = 0.05$  ve  $\alpha = 0.10$  değerleri verilerek iki defa çalıştırılmıştır. Elde edilen izleme sonuçları Şekil 5.4. ve Çizelge 5.5.'te gösterilmektedir. İki ayrı durum için elde edilen eşik değerleri ise Ek D.'de verilmektedir. Şekil 5.4. incelendiğinde  $\alpha = 0.10$  için boda algoritması sonuçlarının daha erken uyarı verdiği, dolayısıyla daha iyi performansa sahip olduğu görülmektedir. Çizelge 5.6.'da ise elde edilen izleme sonuçları CDC kurumunun sonuçlarıyla karşılaştırılmaktadır.



Şekil 5.4. Boda yönteminden elde edilen izleme grafikleri

Çizelge 5.5. Boda yöntemlerine ilişkin 2015-2016 grip mevsimi için haftalık grip faaliyet durumu

Hafta	Karar	Hafta	Karar	Hafta	Karar	Hafta	Karar
40	Normal	1	Normal	14	Normal	27	Normal
41	Normal	2	Normal	15	Normal	28	Normal
42	Normal	3	Normal	16	Normal	29	Normal
43	Normal	4	Normal	17	Normal	30	Normal
44	Normal	5	Olağandışı(B0.10)	18	Normal	31	Normal
45	Normal	6	Olağandışı(B.05,B0.10)	19	Normal	32	Normal
46	Normal	7	Olağandışı(B.05,B0.10)	20	Normal	33	Normal
47	Normal	8	Olağandışı(B.05,B0.10)	21	Normal	34	Normal
48	Normal	9	Olağandışı(B.05,B0.10)	22	Normal	35	Normal
49	Normal	10	Olağandışı(B.05,B0.10)	23	Normal	36	Normal
50	Normal	11	Olağandışı(B0.10)	24	Normal	37	Normal
51	Normal	12	Normal	25	Normal	38	Normal
52	Normal	13	Normal	26	Normal	39	Normal

**Çizelge 5.6. Boda yöntemlerinin verdiği kararlar için sınıflandırma tablosu**

Yöntem	Karar	Gerçek Durum(CDC)		Toplam	Duyarlılık	Seçicilik
		Normal	Olağandışı			
Boda(0.05)	Normal	TN=39	FN=8	47	0.3846	1.000
	Olağandışı	FP=0	TP=5	5		
	Toplam	39	13	52		
Boda(0.10)	Normal	TN=39	FN=6	45	0.5385	1.000
	Olağandışı	FP=0	TP=7	7		
	Toplam	39	13	52		

## 5.2 Zaman Serisi Modellemesi Yaklaşımlarının CDC Grip Verilerine Uygulanması

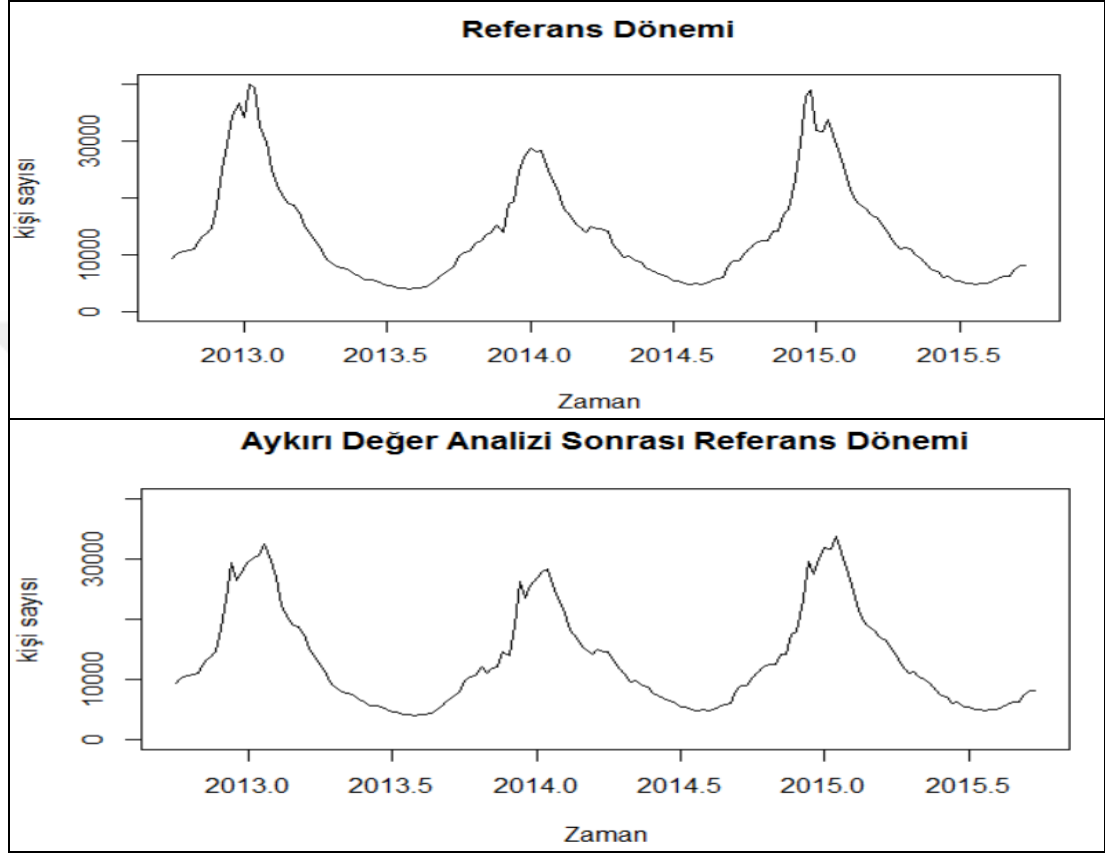
Bu bölümde Holt-Winters yöntemi ve Arima modelleme yöntemleri CDC grip verilerine uygulanmaktadır. Ayrıca Gizli Markov modelleri yaklaşımı da bu bölümde uygulanacaktır. 2012 yılının 40. haftasından başlayıp 2015 yılının 39. haftasına kadar olan kişi sayılarının oluşturduğu zaman serisindeki yani referans dönemindeki salgınların etkisini azaltmak amacıyla, aykırı değer analizi R programındaki “tsoutliers” fonksiyonu aracılığıyla yapılmıştır. Bu analiz sonucu aykırı değerlerin yerini “tsoutliers” fonksiyonunun belirlediği değerler (Çizelge 5.7.) almıştır. Aykırı değerler analizi öncesi ve sonrasına ilişkin referans döneminin zaman serisi grafikleri Şekil 5.5.’te gösterilmektedir.

**Çizelge 5.7. Aykırı değerler yerine tsoutliers fonksiyonu tarafından belirlenen değerler**

Gözlem No	Belirlenen Değerler	Gözlem No	Belirlenen Değerler
10	23294.95	63	26178.97
12	26529.57	64	23511.73
13	28344.92	65	25549.48
14	29491.84	66	26907.42
15	30232.54	67	27870.00
16	30507.26	76	14207.30
19	26788.37	80	12931.88
57	10934.16	112	17511.69
58	11826.51	116	27476.95
59	12058.68	117	30019.06
60	14500.11	135	10379.34

Aykırı değerlerden arındırılmış referans dönemi verilerine dönüşüm gerekli olup olmadığını tespit etmek amacıyla, R programı aracılığıyla Box-Cox analizi uygulandı.

Bu analiz sonucu veriler için logaritmik dönüşümün uygun olduğu görüldü. Bu nedenle aykırı değerlerden arınmış referans dönemleri verilerinin doğal logaritması alındı. Doğal logaritması alınmış bu veriler için uygun modeller bulmak amacıyla, ilk önce Holt-Winters yöntemi daha sonrada Arima modelleme yaklaşımı uygulandı.



Şekil 5.5. Aykırı değer analizi öncesi ve sonrası referans dönemi grafikleri

### 5.2.1 Holt-Winters yaklaşımının CDC grip verilerine uygulanması

Verilerin logaritması alınmış olduğundan toplamsal Holt-Winters modeli R programı aracılığıyla verilere uygulanmıştır (Çizelge 5.8.). Elde edilen artıklara ilişkin özilişki, kısmi özilişki ve Ljung-Box testi  $p$  değerleri grafikleri Şekil 5.6.'da verilmektedir.

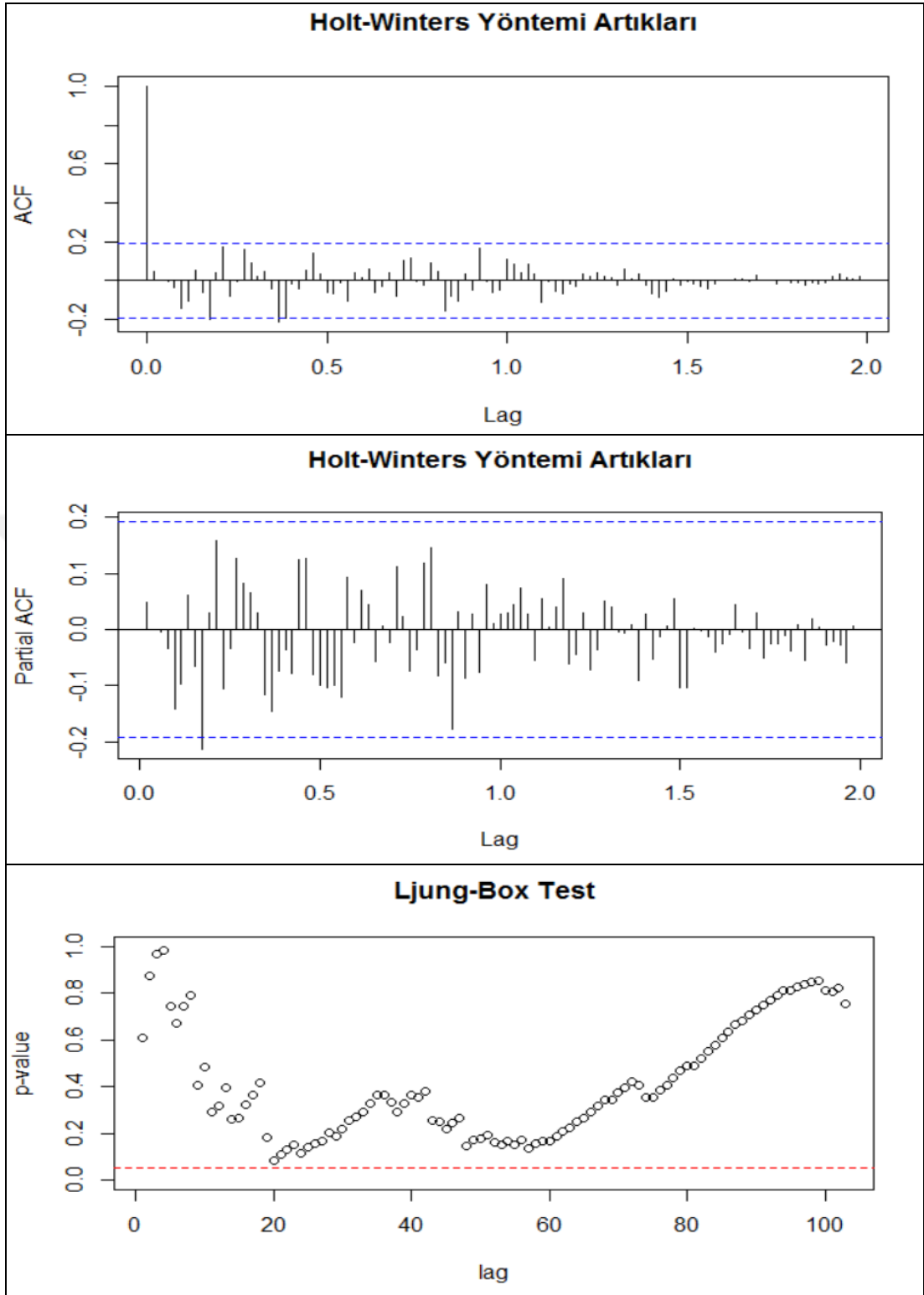


**Çizelge 5.8. CDC verileri için Holt-Winters düzeltirme yöntemine ilişkin ölçütler**

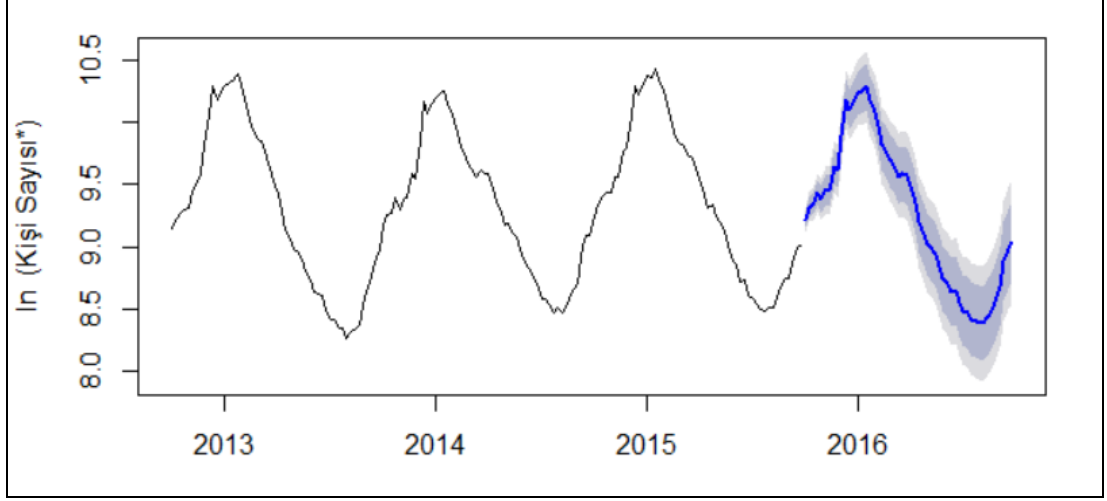
Ölçütler	Toplamsal Holt-Winters		
	$\alpha = 0.702$	$\beta = 0$	$\gamma = 1$
ME			-0.0008
RMSE			0.0492
MAE			0.0359
MPE			-0.0134
MAPE			0.3927
MASE			0.2766
ACF1			0.0494

Şekil 5.6.'daki acf, pacf ve Ljung-Box testi  $p$  değerleri incelendiğinde, artıkların ilişkisiz oldukları kabul edilebilir. Başka bir ifadeyle artıkların beyaz gürültü (White-noise) olduğu söylenebilir. Ayrıca Anderson-Darling ( $AD = 1.6166, p = 0.0004$ ) ve Shapiro-Wilk ( $W = 0.9644, p = 0.0068$ ) normallik testleri artıklara uygulandığında, artıkların normal dağılmadığı görülür. Artıkların rastgele dağılıp dağılmadıklarının kontrolü için Koşu (runs) testi ( $Z_\alpha = 0.8106, p = 0.2088$ ) uygulandığında, artıkların rastgele dağıldığı kabul edilebilir.

2015-2016 grip mevsimi olarak belirlenen izleme dönemi için Holt-Winters yöntemi kullanılarak 52 haftalık öngörü değerleri elde edilmiştir (Şekil 5.7.). Öngörülerin değerleri ve eşik değeri olarak kullanılacak %95 öngörü aralığının üst sınırı Ek E.'de verilmektedir. Ayrıca Ek E.'de izleme dönemi verilerinin doğal logaritması ve bu doğal logaritmaların eşik değerleriyle farkları yer almaktadır. Farkların pozitif olduğu haftalar olağandışı durumu belirtmektedir (Çizelge 5.9.). Son olarak CDC kurumunun sonuçları ile Holt-Winters yönteminin verdiği sonuçlar karşılaştırılmıştır (Çizelge 5.10.).



Şekil 5.6. Holt-Winters yöntemi artıklarına ilişkin grafikler



Şekil 5.7. 2015-2016 grip mevsimine ilişkin Holt-Winters yöntemiyle elde edilen 52 haftalık öngörülerin grafiği

Çizelge 5.9. Holt-Winters yöntemine ilişkin 2015-2016 grip dönemi için haftalık grip faaliyet durumu

Hafta	Karar	Hafta	Karar	Hafta	Karar	Hafta	Karar
40	Normal	1	Normal	14	Normal	27	Normal
41	Normal	2	Normal	15	Normal	28	Normal
42	Normal	3	Normal	16	Normal	29	Olağandışı
43	Normal	4	Normal	17	Normal	30	Olağandışı
44	Normal	5	Normal	18	Normal	31	Normal
45	Normal	6	Normal	19	Normal	32	Normal
46	Normal	7	Normal	20	Normal	33	Normal
47	Normal	8	Olağandışı	21	Normal	34	Normal
48	Normal	9	Olağandışı	22	Normal	35	Normal
49	Normal	10	Olağandışı	23	Normal	36	Normal
50	Normal	11	Olağandışı	24	Normal	37	Normal
51	Normal	12	Olağandışı	25	Normal	38	Normal
52	Normal	13	Normal	26	Normal	39	Normal

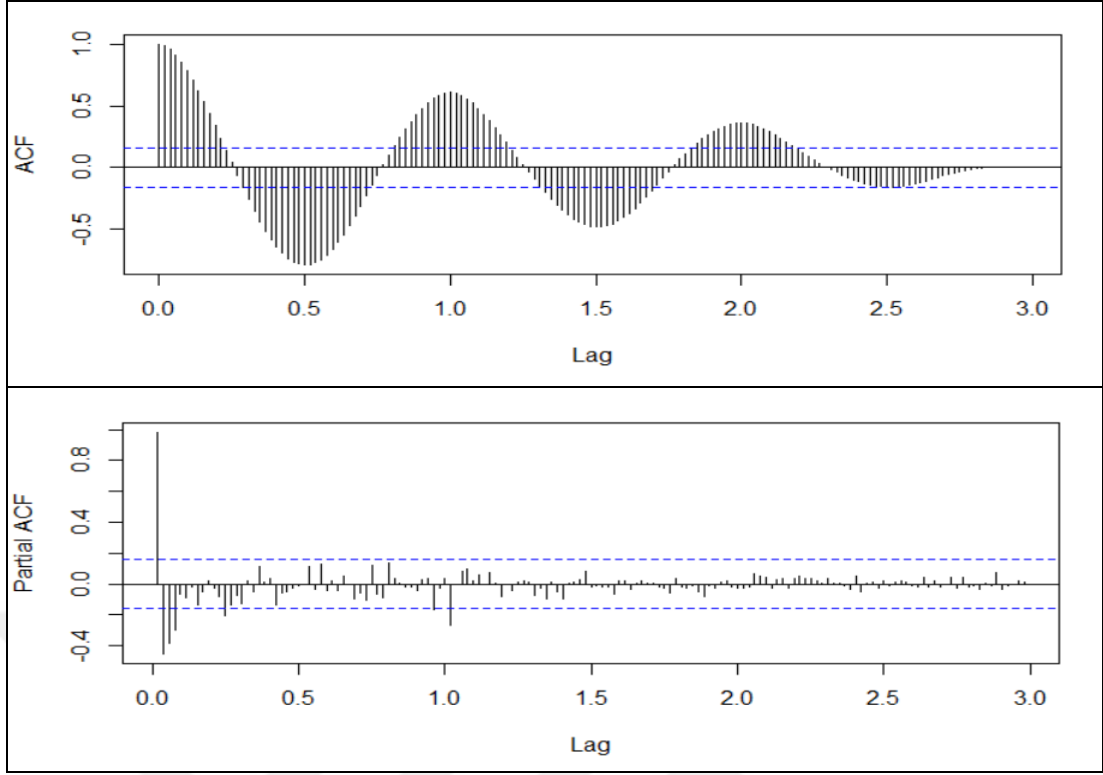
Çizelge 5.10. Holt Winters yönteminin verdiği kararlar için sınıflandırma tablosu

Yöntem	Karar	Gerçek Durum(CDC)		Toplam	Duyarlılık	Seçicilik
		Normal	Olağandışı			
Holt-Winters	Normal	TN=37	FN=8	45	0.3846	0.9487
	Olağandışı	FP=2	TP=5	7		
	Toplam	39	13	52		

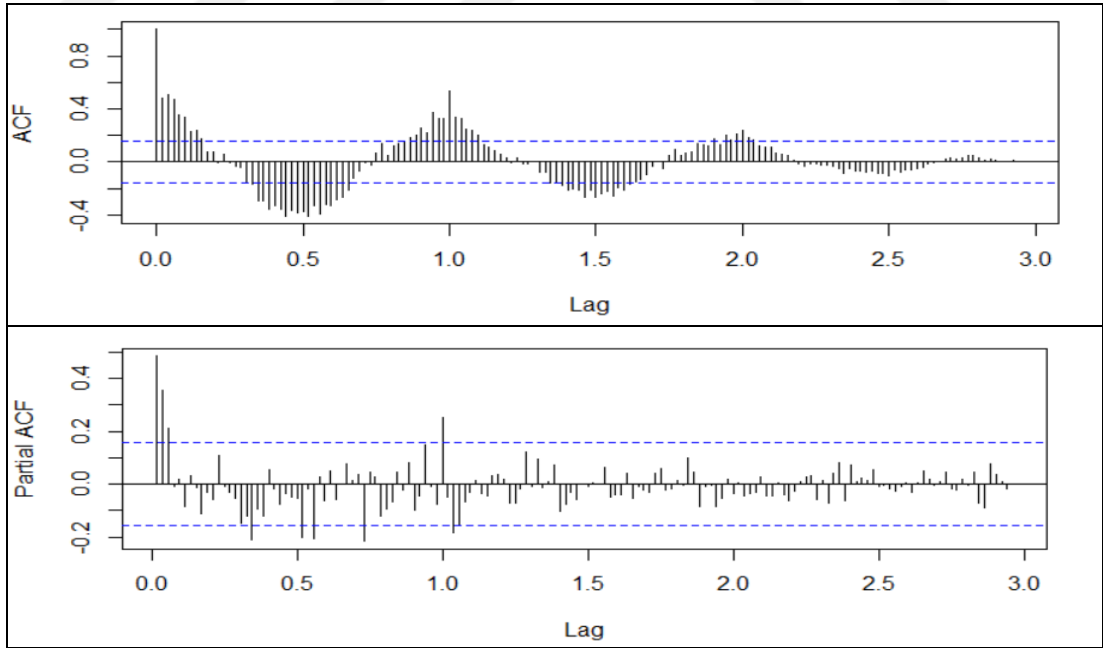
### 5.2.2 Arima modelleme yaklaşımının CDC grip verilerine uygulanması

Bu alt bölümdeki tüm analizler ve grafikler R programı kullanılarak gerçekleştirilmiştir. Aykırı değerlerden arındırılmış ve doğal logaritması alınmış referans dönemine ilişkin verilerin özilişki ve kısmi özilişki grafikleri Şekil 5.8.'de verilmektedir. Bu grafiklerden serinin durağan olmadığı izlenimi elde edilmektedir. Bu izlenimi test etmek amacıyla Augmented Dickey-Fuller testi ( $adf = -3.7623$ ,  $p = 0.0227$ ) ve Phillips-Perron testi ( $Z_\alpha = -6.723$ ,  $p = 0.729$ ) uygulanmıştır. Phillips-Perrons testi serinin durağan olmadığını ve Augmented Dickey-Fuller testi ise  $\alpha = 0.0227$  düzeyi ve daha düşük düzeylerde serinin durağan olmadığını belirtmektedir.

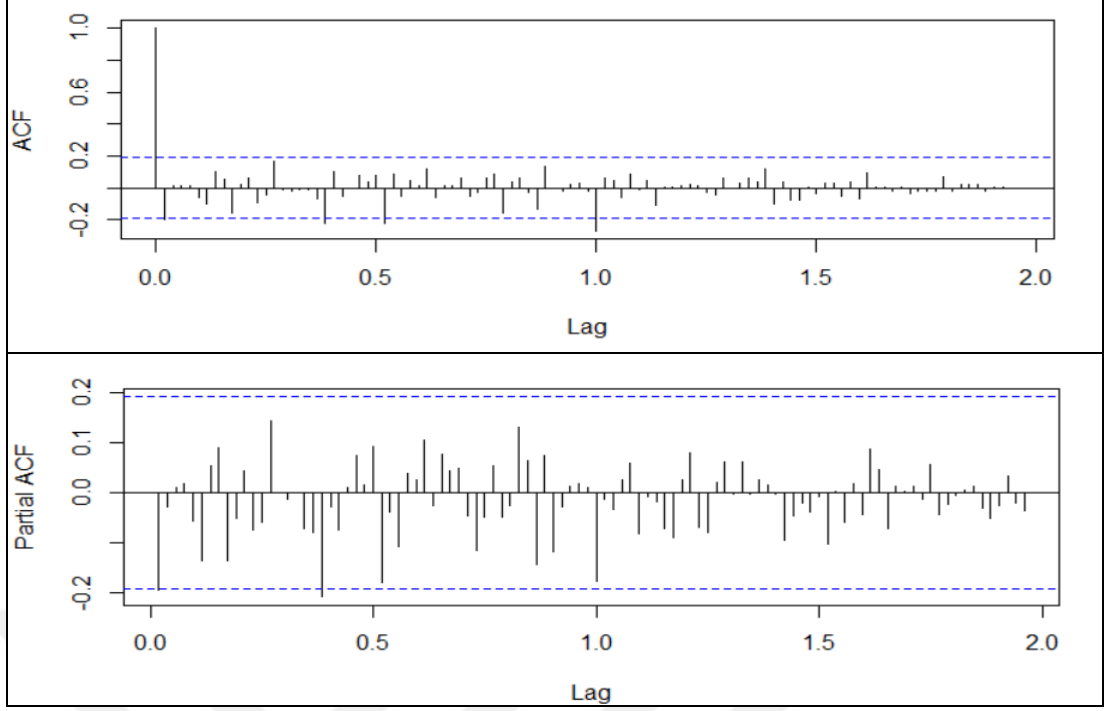
Serinin durağan hale getirilebilmesi amacıyla fark alma işlemi uygulanmıştır. Birinci fark işlemi sonucu, özilişki ve kısmi özilişki grafikleri serinin mevsimsel kaynaklı durağan olmadığını belirtmektedir (Şekil 5.9.). Bu durumu test etmek amacıyla Augmented Dickey-Fuller testi; ( $adf = -3.0642$ ,  $p = 0.1328$ ) ve Phillips-Perron testi; ( $Z_\alpha = -89.527$ ,  $p < 0.01$ ) uygulanmıştır. ADF testi birinci farkın durağan olmadığını desteklerken, PP testi durağan olduğunu desteklemektedir. Bu nedenle 1. farkı alınmış seriye mevsimsel fark işlemi uygulanmıştır. Bu işlem sonucu durağan bir serinin elde edildiği özilişki ve kısmi özilişki grafiklerinden (Şekil 5.10.) görülmektedir. Durağanlık, Augmented Dickey-Fuller testi ( $adf = -4.6267$ ,  $p < 0.01$ ) ve Phillips-Perron testi ( $Z_\alpha = -120.87$ ,  $p < 0.01$ ) tarafından desteklenmektedir.



**Şekil 5.8. Aykırı değerlerden arındırılmış ve logaritması alınmış referans dönemi verilerine ilişkin özilişki ve kısmi özilişki grafikleri**



**Şekil 5.9. Birinci farkı alınmış verilere ilişkin özilişki ve kısmi özilişki grafikleri**



**Şekil 5.10. Birinci farkı ve mevsimsel farkı alınmış verilere ilişkin özilişki ve kısmi özilişki grafikleri**

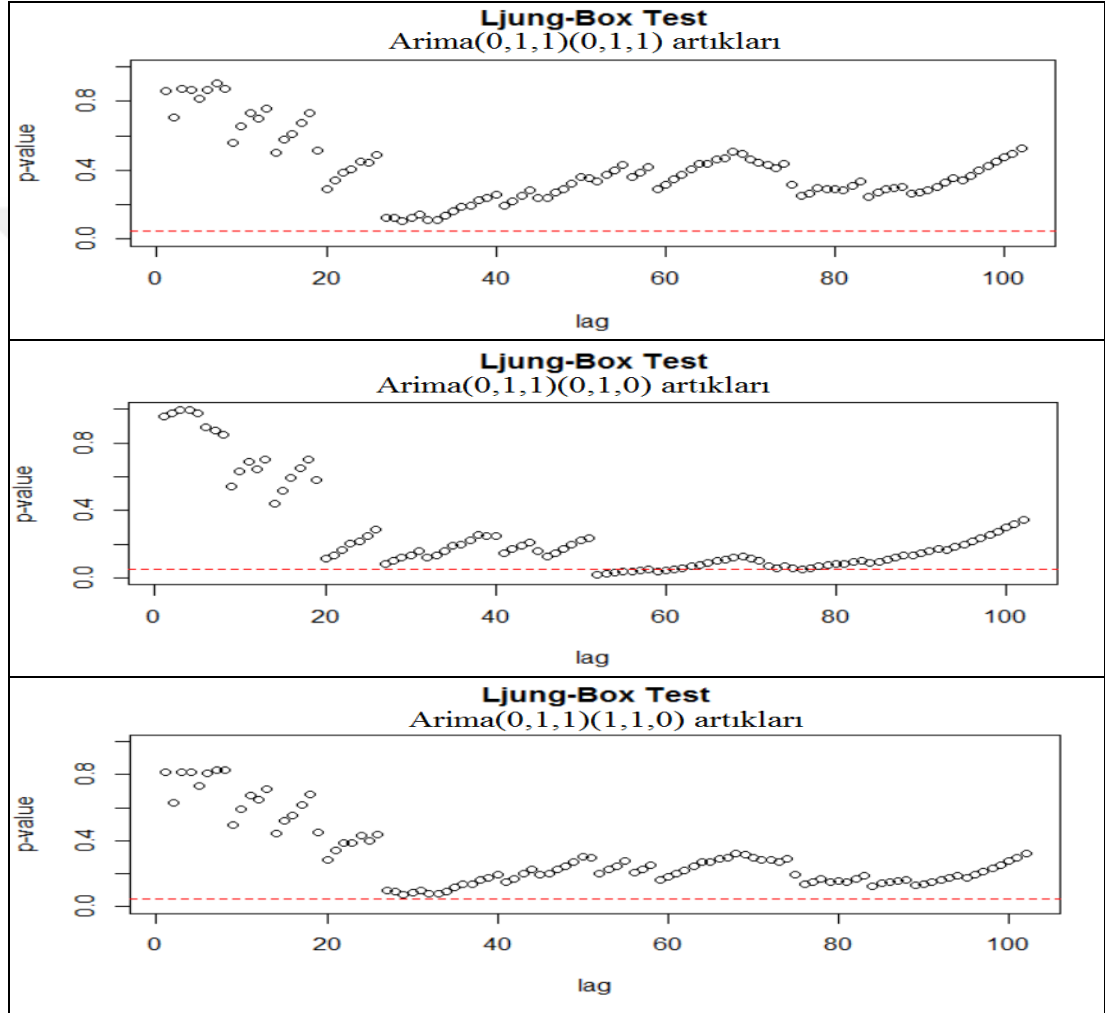
Şekil 5.10.'daki acf ve pacf grafiklerinden veriler için uygun olabilecek modeller ve bazı özellikleri Çizelge 5.11.'de verilmektedir. Parametre kestirimi yapılabilen modellere ilişkin artıkların normal dağılmadığı (Anderson Darling test) ancak rastgele oldukları (koşu testi) kabul edilebilmektedir. Ayrıca bu artıkların beyaz gürültü özelliğini sağlayıp sağlamadığı Ljung-Box  $p$  değerleri grafiklerinden (Şekil 5.11.) denetlenebilir.

**Çizelge 5.11. Birinci farkı ve mevsimsel farkı alınmış veriler için aday modellerin özellikleri**

Modeller	Me	Rmse	Mape	A-D test	A-D test p değeri	Koşu testi	Koşu testi p değeri
Arima(1,1,0)(1,1,0)				Parametre kestirimi yapılamadı.			
Arima(2,1,0)(1,1,0)				Parametre kestirimi yapılamadı.			
Arima(0,1,1)(0,1,1)	-0.0010	0.0375	0.2546	5.3021	<0.0001	-0.1179	0.5469
Arima(0,1,1)(0,1,0)	-0.0014	0.0540	0.3602	5.9674	<0.0001	0.0308	0.4877
Arima(0,1,1)(1,1,0)	-0.0010	0.0444	0.3012	5.4756	<0.0001	-0.1697	0.5674

Şekil 5.11.'den Arima (0,1,1)(0,1,0) modeline ilişkin artıkların beyaz gürültü özelliğini sağlamadığı görülmektedir. Bu nedenle bu model dikkate alınmaz. Geriye kalan iki model arasında seçim yapmak için Çizelge 5.11.'deki RMSE ve MAPE

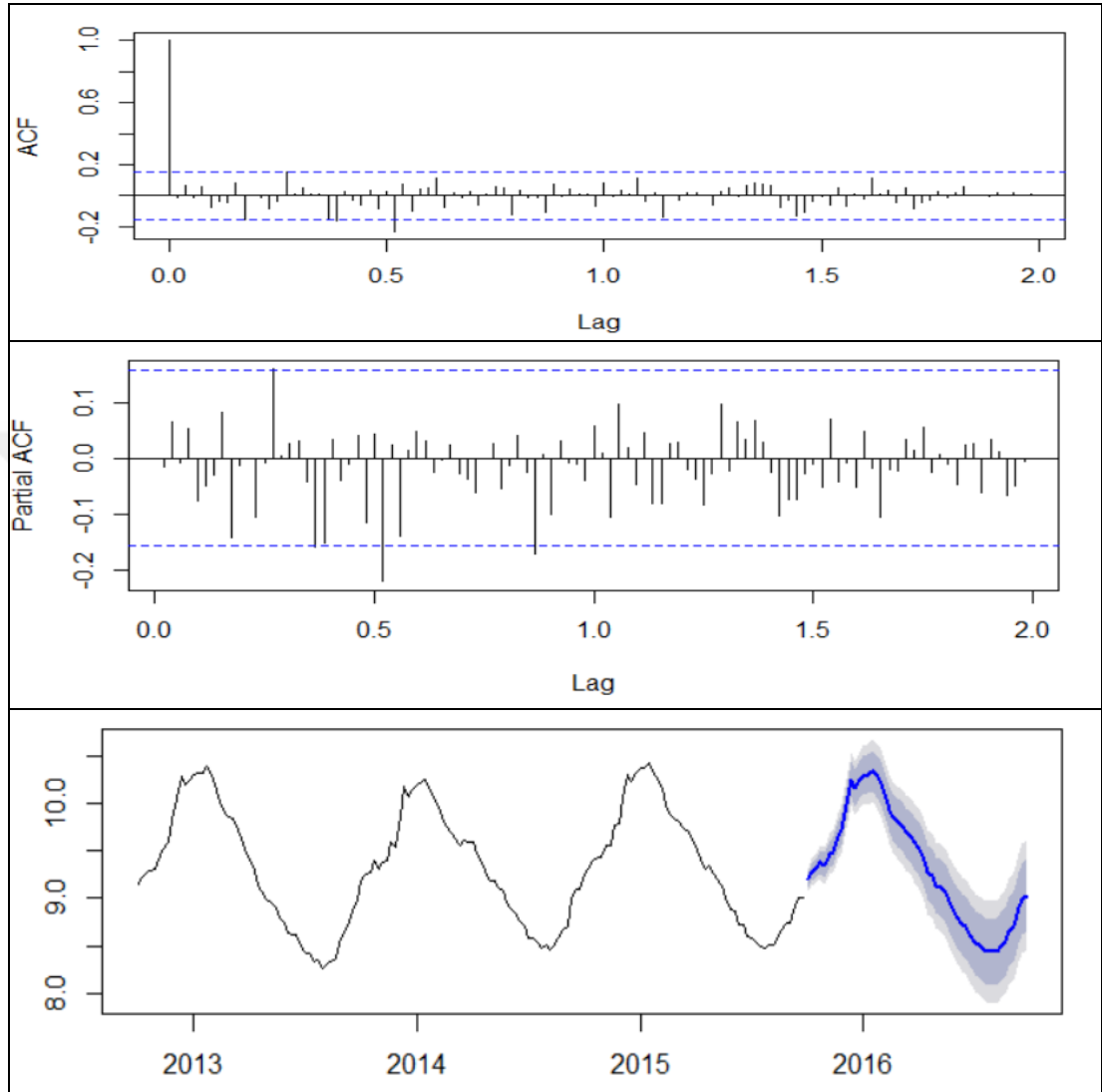
ölçütleri incelenirse, Arima (0,1,1)(0,1,1) modelinin daha uygun olduğu karara ulaşılır. Bu modellerin AIC kriterleri de incelenirse (Arima (0,1,1)(0,1,1): AIC = -280.82; Arima (0,1,1)(1,1,0): AIC = -281.34); bu ölçütlerin değerlerinin birbirine oldukça yakın olduğu ve modellerin eşit sayıda parametre sahip olduğu dikkate alınır, Arima (0,1,1)(0,1,1) modelinin veriler için daha uygun olduğu kararı desteklenir.



Şekil 5.11. Aday modellerin artıklarına ilişkin Ljung-Box grafikleri

Bu nedenle izleme dönemine ilişkin öngörüler bu model kullanılarak yapılacaktır. Uygun modelin artıklarına ilişkin özilişki ve kısmi özilişki grafikleri Şekil 5.12.'de gösterilmekte ve model parametrelerinin kestirimleri ise Çizelge 5.12.'de

verilmektedir. Bu modelle elde edilen 52 haftalık öngörüler Şekil 5.12.'de gösterilmekte ve Ek F.'de verilmektedir.



Şekil 5.12. Arima (0,1,1)(0,1,1) modeli artıklarına ilişkin özilişki, kısmi özilişki grafikleri ve 2015-2016 izleme dönemine ilişkin 52 haftalık öngörü grafiği

Çizelge 5. 12. Seçilen modele ilişkin parametre kestirimleri

Model	Parametre kestirimleri ve standart hatalar	
	ma1	sma1
	-0.2273	-0.9987
Arima(0,1,1)(0,1,1)	s.e. 0.0905	0.5450
sigma <sup>2</sup> estimated as 0.002086: log likelihood = 143.41, aic = -280.82		



Ayrıca Ek F.'de %90 güven düzeyli öngörü aralığının üst sınırı (eşik değerleri) ve doğal logaritması alınmış izleme dönemi verilerinin eşik değerlerinden farklarında verilmektedir. Pozitif farklar ilgili haftalarda olağandışı durumu ifade etmektedir. Modele ilişkin artıklar normal olmadığından, öngörü aralığı güven düzeyi %90 olarak seçilmiştir. Çizelge 5.13. Arima modellemesine ilişkin 52 haftalık olağandışı durum karar tablosunu vermektedir. Bu çizelgedeki kararların Holt-Winters yöntemine ait kararlarla paralellik gösterdiği görülmektedir. Çizelge 5.14. elde edilen sonuçlar ile CDC kurumunun sonuçlarını karşılaştırmaktadır.

**Çizelge 5.13. Arima modelleme yöntemine ilişkin 2015-2016 grip mevsimi için haftalık grip faaliyet durumu**

Hafta	Karar	Hafta	Karar	Hafta	Karar	Hafta	Karar
40	Normal	1	Normal	14	Normal	27	Normal
41	Normal	2	Normal	15	Normal	28	Normal
42	Normal	3	Normal	16	Normal	29	Olağandışı
43	Normal	4	Normal	17	Normal	30	Normal
44	Normal	5	Normal	18	Normal	31	Normal
45	Normal	6	Normal	19	Normal	32	Normal
46	Normal	7	Normal	20	Normal	33	Normal
47	Normal	8	Olağandışı	21	Normal	34	Normal
48	Normal	9	Olağandışı	22	Normal	35	Normal
49	Normal	10	Olağandışı	23	Normal	36	Normal
50	Normal	11	Olağandışı	24	Normal	37	Normal
51	Normal	12	Olağandışı	25	Normal	38	Normal
52	Normal	13	Normal	26	Normal	39	Normal

**Çizelge 5.14. Arima yönteminin verdiği kararlar için sınıflandırma tablosu**

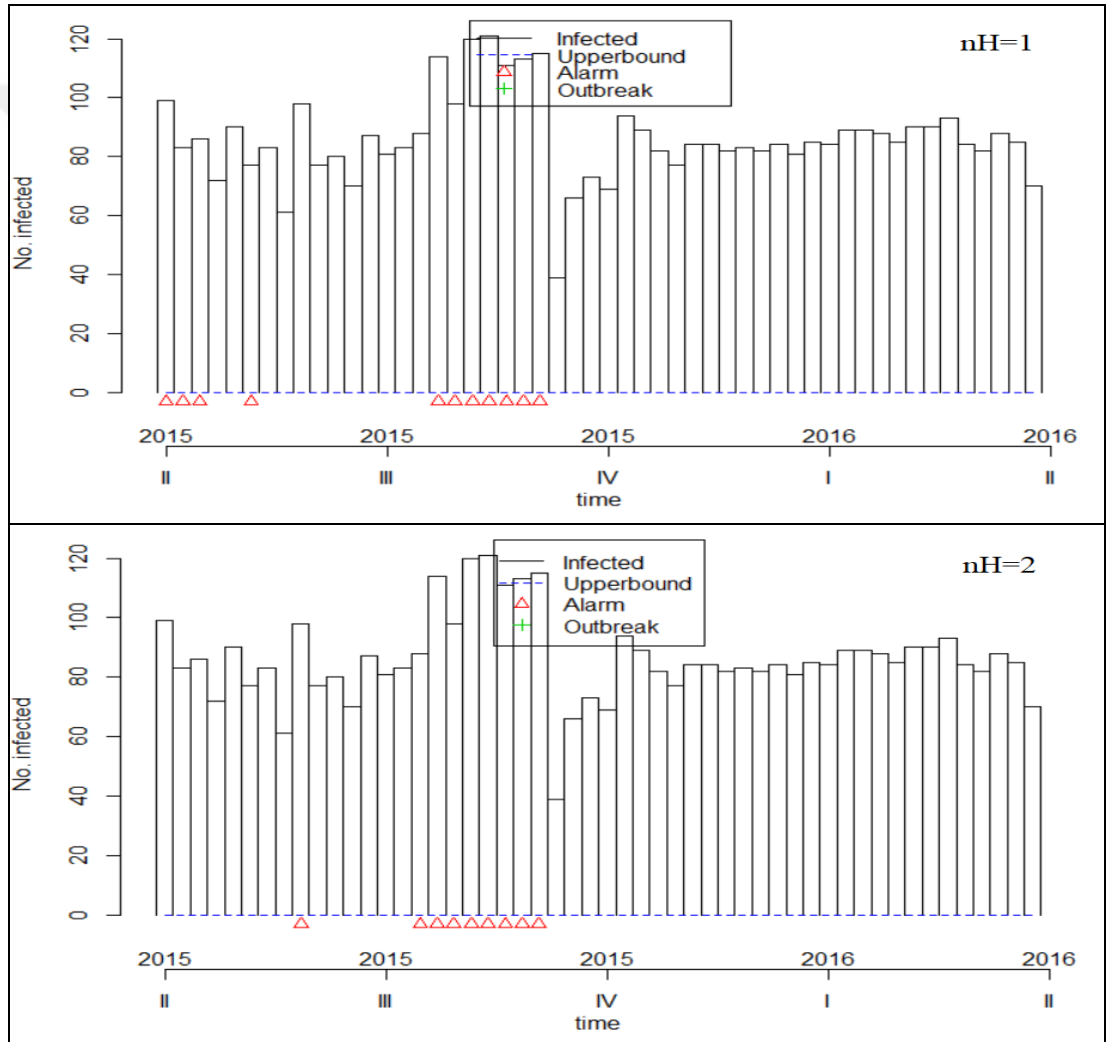
Yöntem	Karar	Gerçek Durum(CDC)		Toplam	Duyarlılık	Seçicilik
		Normal	Olağandışı			
Arima	Normal	TN=38	FN=8	46	0.3846	0.9744
	Olağandışı	FP=1	TP=5	6		
	Toplam	39	13	52		

### 5.2.3 Hidden Markov modeli yaklaşımının CDC grip verilerine uygulanması

Bu alt bölümde; Le Strat ve Carrat (1999) tarafından önerilen Hidden Markov (Gizli Markov) modeli kullanılacaktır. Bu yöntem, Alt Bölüm 5.3.4'de elde edilen

mevsimsellik düzeltmesi yapılmış artıkların pozitif hale getirilmiş 1. farklarına (Ek G.) uygulanmaktadır.

Her ne kadar verilerde mevsimsellik düzeltmesi yapılsa da verilerde zamana bağlı etkiler halen mevcuttur. Bu nedenle Hidden Markov modeli uygulanırken kalan mevsimsellik etkilerini dikkate alabilmek için, kullanılan dalga sayısı (harmoniklik)  $nH = 1$  ve  $nH = 2$  olacak biçimde 2 ayrı şekilde alınmıştır. Bu model, CDC grip verilerine R programının surveillance paketi aracılığıyla uygulanmıştır. Elde edilen sonuçlar Şekil 5.13., Çizelge 5.15. ve Çizelge 5.16.'da verilmektedir.



Şekil 5.13. CDC grip mevsimi verileri için Hidden-Markov modellerine ilişkin grafikler

**Çizelge 5.15. Hidden Markov modeli yöntemlerine ilişkin 2015-2016 grip mevsimi için haftalık grip faaliyet durumu**

Hafta	Karar	Hafta	Karar	Hafta	Karar	Hafta	Karar
40	Olağandışı(HMM1)	1	Normal	14	Normal	27	Normal
41	Olağandışı(HMM1)	2	Normal	15	Normal	28	Normal
42	Olağandışı(HMM1)	3	Olağandışı(HMM2)	16	Normal	29	Normal
43	Normal	4	Olağandışı(HMM1, HMM2)	17	Normal	30	Normal
44	Normal	5	Olağandışı(HMM1, HMM2)	18	Normal	31	Normal
45	Olağandışı(HMM1)	6	Olağandışı(HMM1, HMM2)	19	Normal	32	Normal
46	Normal	7	Olağandışı(HMM1, HMM2)	20	Normal	33	Normal
47	Normal	8	Olağandışı(HMM1, HMM2)	21	Normal	34	Normal
48	Olağandışı(HMM2)	9	Olağandışı(HMM1, HMM2)	22	Normal	35	Normal
49	Normal	10	Olağandışı(HMM1, HMM2)	23	Normal	36	Normal
50	Normal	11	Normal	24	Normal	37	Normal
51	Normal	12	Normal	25	Normal	38	Normal
52	Normal	13	Normal	26	Normal	39	Normal

HMM1, nH=1 iken; HMM2 ise nH=2 iken alınan kararları göstermektedir.

**Çizelge 5.16. Hidden Markov modeli yöntemlerinin verdiği kararlar için sınıflandırma tablosu**

Yöntem	Karar	Gerçek Durum(CDC)		Toplam	Duyarlılık	Seçicilik
		Normal	Olağandışı			
Hidden Markov Modeli-1	Normal	TN=35	FN=6	41	0.5385	0.8974
	Olağandışı	FP=4	TP=7	11		
	Toplam	39	13	52		
Hidden Markov Modeli-2	Normal	TN=38	FN=5	43	0.6154	0.9744
	Olağandışı	FP=1	TP=8	9		
	Toplam	39	13	52		

### 5.3 Kalite Kontrol Grafiklerinin CDC Grip Verilerine Uygulanması

Bu alt bölümde bioizlemede kullanılan bazı kontrol grafiklerinin uygulaması CDC grip verileri üzerinde gösterilmektedir. Burada dikkate alınan grafikler Shewhart I-Chart, Ewma, Cusum, Ears C1, Ears C2 ve Ears C3 yöntemleridir.

#### 5.3.1 Holt-Winters artıklarına Shewhart I-Chart, Ewma, Cusum grafiklerinin uygulanması

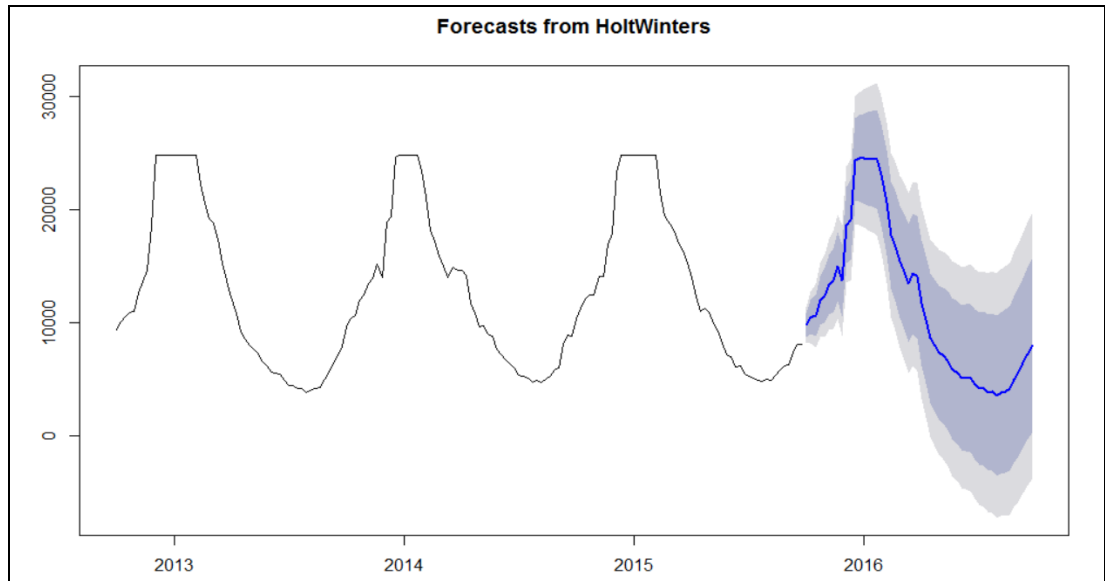
Referans yıllarını oluşturan 2012-2013, 2013-2014 ve 2014-2015 grip mevsimleri verilerinde, salgınlardan kaynaklanabilecek vaka sayıları yüksekliğinin etkisinin giderilmesi amacıyla, verilerin dağılımının üst %15'lik kısmı 85'inci yüzdeliğine eşitlenmiştir. Standart kalite kontrol grafiklerinin uygulanabilmesi için verilerdeki

olası mevsimsellik, trend ve özilişki yapısı gibi özelliklerin dikkate alınması gerekir. Bu amaçla verilere Holt-Winters düzeltirme yöntemi R programı aracılığıyla uygulanmıştır. Elde edilen sonuçlara göre toplamsal mevsimsel etkiye sahip olan Holt-Winters düzeltilmesi seçilmiş (Çizelge 5.17.) ve 2015-2016 yılına ait ileriye dönük tahminler toplamsal Holt-Winters düzeltilmesi kullanılarak elde edilmiştir (Ek H.). Bu tahminler Şekil 5.14.'de görsel olarak ifade edilmektedir.

**Çizelge 5.17. CDC verileri için Holt-Winters düzeltirme yöntemine ilişkin ölçütler**

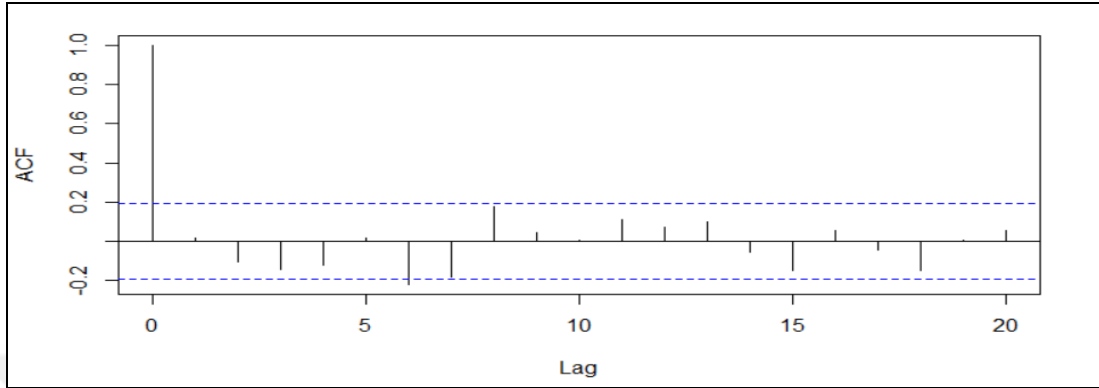
Ölçütler	Çarpımsal Holt-Winters			Toplamsal Holt-Winters		
	$\alpha = 0,90$	$\beta = 0$	$\gamma = 1$	$\alpha = 1$	$\beta = 0$	$\gamma = 0$
ME	7.5295			-6.0598		
RMSE	936.4991			825.2849		
MAE	487.9188			461.4289		
MPE	0.2339			-0.4503		
MAPE	4.2069			4.1404		
MASE	0.3719			0.3517		
ACF1	0.0764			0.0134		

Çizelge 5.17.'de Holt-Winters yöntemine ilişkin  $\alpha, \beta, \gamma$  düzeltirme katsayıları R programının uyguladığı bir optimizasyon algoritması sonucu elde edilmiştir. Burada  $\alpha$  serinin ortalama düzeyini,  $\beta$  serinin trendini ve  $\gamma$  serinin mevsimselliğini güncellemek amacıyla kullanılan katsayılardır.



**Şekil 5.14. Holt-Winters yöntemi ile izleme dönemine ilişkin tahminler**

Holt-Winters uygulaması sonucu elde edilen artıklar özilişkisiz durumuna gelmiş olup mevsimsellik, trend gibi etkilerden arınmışlardır (Şekil 5.15.). Şekil 5.15.'de 6. gecikmedeki ACF değerinin anlamlı gözükmemesi tesadüf eseri olarak kabul edilebilir çünkü  $20 * 0.05 = 1$  adet anlamlılık özilişki yokken beklenmektedir.



Şekil 5.15. Holt-Winters yönteminden elde edilen artıkların özilişki grafiği

Kalite kontrol grafiklerinde kullanılacak verileri oluşturmak amacıyla; 2015-2016 grip mevsimi (izleme dönemi) verileri ile Holt-Winters tahminleri arasında fark alınmıştır. Daha sonra tahmin hataları, referans yıllarına ilişkin artıkların standart sapması kullanılarak standartlaştırılmıştır (Ek H.).

$$z_t = (x_t - tahmin_t) / \hat{\sigma}_{artık}$$

Standartlaştırılmış artıklara ilişkin I-Chart, Cusum ve Ewma kalite kontrol grafikleri Minitab 16 programı ile elde edilmiştir (Şekil 5.16.). I-Chart grafiği için  $3\sigma$  kontrol sınırları, Cusum grafiği parametreleri  $h = 4, k = 1$  ve Ewma grafiği için parametreler ise  $L = 3, \lambda = 0.60$  alınmıştır. Cusum ve Ewma grafiklerindeki  $k$  ve  $\lambda$  değerlerinin seçimi ortalamadan olan büyük sapmaları belirlemeye yöneliktir ve pozitif özilişkiden dolayı  $k$  değeri sanayideki uygulamaların aksine 0.5 değil 1 alınmıştır. Ayrıca Şekil 5.16.'daki Cusum grafiği her olağandışı sinyalden sonra birikimli toplamı gösteren istatistiği sıfırlamaktadır. Grafiklerin oluşturulması için kullanılan grafik parametrelerinin değerleri kısmen bu grafiklerin sanayi ortamındaki uygulamalarıyla uyumludur. Ancak parametre değerlerinin bu şekilde seçimiyle, sanayi uygulamalarında ulaşılabilecek ARL performansları buradaki gibi biyoizleme amaçlı

uygulamalarda beklenmemektedir. Bunun en önemli nedeni izleme dönemi için standartlaştırılmış tahmin hataları kullanılmasına rağmen, bu hataların halen özilişkili olmasıdır. Başka bir ifadeyle izleme dönemi için bazı grafik varsayımları ihlal edilmektedir. Dolayısıyla parametrelerin bu şekilde seçiminden kaynaklanacak yüksek eşik değerleri özilişkiden kaynaklı olabilecek yanlış alarmların sayısını azaltma etkisine sahip olacaktır.

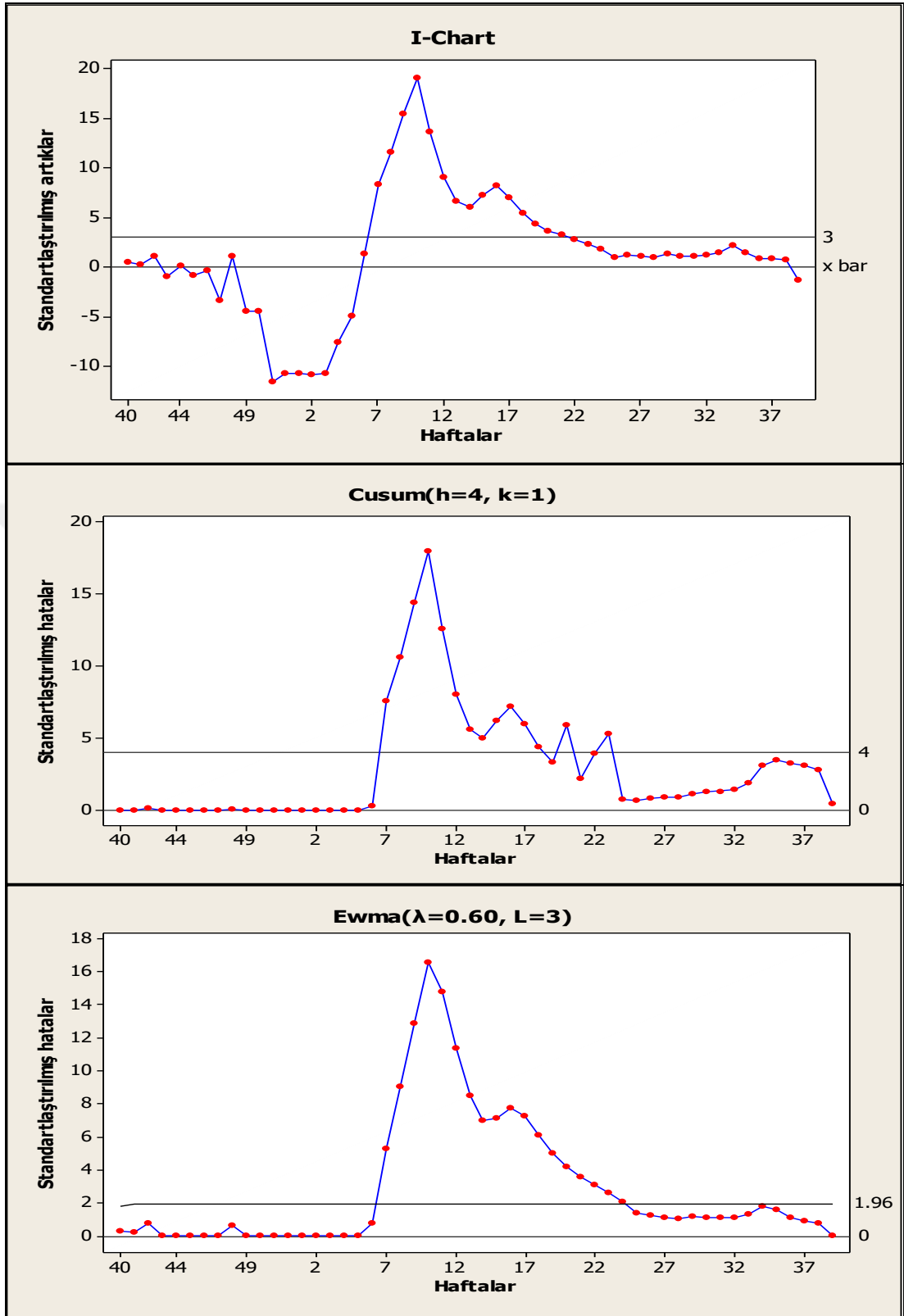
Şekil 5.16. incelendiğinde I-Chart, Cusum ve Ewma grafiklerinin benzer sonuçlar verdiği görülmektedir. Olağandışı sinyallerin başlangıcı (7. Hafta) her üç grafikte de aynıdır ve bu olağandışı sinyallerin bitiş zaman noktaları da birbirlerine yakındır. Bu grafiklerden izleme dönemine ilişkin 52 haftalık kararlar Çizelge 5.18.'de verilmektedir. CDC kurumunun sonuçları ile elde edilen sonuçlar Çizelge 5.19.'da karşılaştırılmaktadır.

**Çizelge 5.18. I-Chart, Ewma ve Cusum grafiklerine ilişkin 2015-2016 grip mevsimi için haftalık grip faaliyet durumu**

Hafta	Karar	Hafta	Karar	Hafta	Karar	Hafta	Karar
40	Normal	1	Normal	14	Olağandışı(I,E,C)	27	Normal
41	Normal	2	Normal	15	Olağandışı(I,E,C)	28	Normal
42	Normal	3	Normal	16	Olağandışı(I,E,C)	29	Normal
43	Normal	4	Normal	17	Olağandışı(I,E,C)	30	Normal
44	Normal	5	Normal	18	Olağandışı(I,E,C)	31	Normal
45	Normal	6	Normal	19	Olağandışı(I,E)	32	Normal
46	Normal	7	Olağandışı(I,E,C)	20	Olağandışı(I,E,C)	33	Normal
47	Normal	8	Olağandışı(I,E,C)	21	Olağandışı(I,E)	34	Normal
48	Normal	9	Olağandışı(I,E,C)	22	Olağandışı(E)	35	Normal
49	Normal	10	Olağandışı(I,E,C)	23	Olağandışı (E,C)	36	Normal
50	Normal	11	Olağandışı(I,E,C)	24	Olağandışı (E)	37	Normal
51	Normal	12	Olağandışı(I,E,C)	25	Normal	38	Normal
52	Normal	13	Olağandışı(I,E,C)	26	Normal	39	Normal

**Çizelge 5.19. Kalite kontrol grafikleri yöntemlerinin verdiği kararlar için sınıflandırma tablosu**

Yöntem	Karar	Gerçek Durum(CDC)		Toplam	Duyarlılık	Seçicilik
		Normal	Olağandışı			
I-Chart	Normal	TN=31	FN=6	37	0.5385	0.7949
	Olağandışı	FP=8	TP=7	15		
	Toplam	39	13	52		
Ewma	Normal	TN=28	FN=6	34	0.5385	0.7179
	Olağandışı	FP=11	TP=7	18		
	Toplam	39	13	52		
Cusum	Normal	TN=32	FN=6	38	0.5385	0.8205
	Olağandışı	FP=7	TP=7	14		
	Toplam	39	13	52		



Şekil 5.16. Holt-Winters yönteminden elde edilen standartlaştırılmış hataların kontrol grafikleri

### 5.3.2 Mevsimselliği giderilmiş ham verilere Shewhart I-Chart grafiğinin uygulanması

Ham biyoizleme verileri üzerinde kalite kontrol grafiklerinin doğrudan kullanıldığı uygulamalar mevcuttur. Örneğin Ivanov vd. (2003) solunum ve sindirim şikayetlerine sahip 5 yaş altı çocukların günlük hastahane ziyaretleri sayısını Ewma yöntemini kullanarak izlemiştir (Shumeli, 2013). Bu uygulamalarda yanlış alarm hızı sanayideki süreç kontrol uygulamalarına kıyasla daha yüksek tutulmaktadır. Ayrıca verilerdeki olası özilişkinin etkisini azaltmak amacıyla, ya teorik kontrol sınırları yüksek tutulmakta ya da birinci aşama verilerinden (salgın olmayan dönem) deneysel eşik değerleri hesaplanmaktadır. Bunlara ilaveten yanlış alarm hızını yükseltmemek amacıyla, verilerdeki mevsimsellik giderildikten sonra süreç kontrol yöntemleri uygulanabilmektedir.

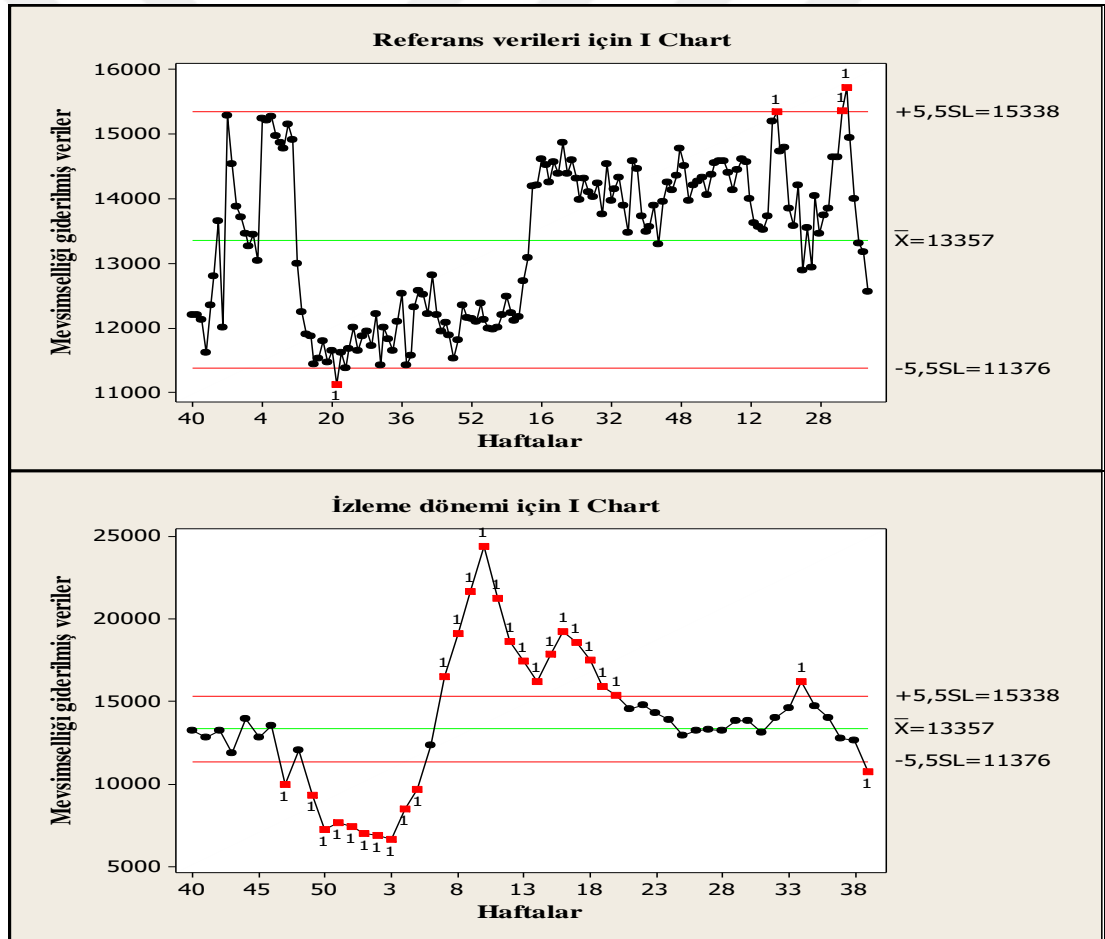
İlk olarak referans dönemi verilerine Alt Bölüm 5.2’de yapıldığı gibi aykırı değer analizi gerçekleştirilerek, geçmişteki olası salgınların veriler üzerindeki etkisi azaltılmaya çalışılmıştır. Daha sonra aykırı değerlerden arındırılmış haftalık CDC referans dönemi verilerindeki çarpımsal mevsimsel etkiyi gidermek amacıyla, hareketli ortalamalara oranlama yöntemi kullanılmıştır. Bu işlem için ilk önce ilgili verilerin 52 genişliğindeki hareketli ortalamaları bulunmuştur, daha sonra ise bu hareketli ortalamaların da 52 genişliğindeki merkezi hareketli ortalamaları alınmıştır. Aykırı değerlerden arındırılmış veriler, bunlara karşılık gelen hareketli ortalamalara bölünüp 100 ile çarpılmıştır. Elde edilen katsayılar haftalık ve yıllık olacak şekilde  $52 \times 3$  boyutta bir matrise dönüştürülmüş ve haftalık ortalamalar bulunmuştur. Bu ortalamaların toplamları 5200 olacak şekilde düzenlenerek 52 haftalık mevsimsel indeksler elde edilmiştir (Çizelge Ek I.1.). Ayrıca her hafta için elde edilen mevsimsellik katsayıları kullanılarak, 2015-2016 grip mevsimindeki haftalık izleme verilerine de mevsimsellik düzeltilmesi yapılmıştır (Çizelge Ek I.2.).

Referans dönemini oluşturan aykırı değerlerden arındırılmış ve mevsimselliği giderilmiş veriler kullanılarak, iki yanlı yanlış alarm hızı yaklaşık 0.02, tek yanlı alarm hızı yaklaşık 0.01 olacak biçimde I-Chart için deneysel kontrol sınırları bulunmuştur.



Bu kontrol sınırları 5.5 standart sapma, ortalaması 13356.8 ve standart sapması 360.149 parametre değerleri ile ifade edilmektedir (Şekil 5.17.). Dolayısıyla, referans verilerini oluşturan mekanizma aynı kalmak şartıyla; bu sınırlar ile deneysel olarak her 100 haftada 1, artış yönünde gerçekleşen yanlış bir alarm beklenmektedir. Şekil 5.17’de referans verileri için 4 yanlış alarm görülmekte, yani iki yanlış yanlış alarm hızı  $4/156 \cong 0.0256$ , tek yanlış yanlış alarm hızı 0.0128’dir.

Deneysel olarak belirlenen bu kontrol sınırları kullanılarak, 2015-2016 grip mevsimi haftalık verileri I-Chart ile izlenmiştir (Şekil 5.17.). İzleme sonucu elde edilen kararlar Çizelge 5.20.’de verilmekteyken, bu kararların CDC kurumunun kararlarıyla karşılaştırılması Çizelge 5.21.’de verilmektedir.



Şekil 5.17. Mevsimselliği giderilmiş 2015-2016 grip verileri için deneysel kontrol sınırları kullanan I-Chart( $\alpha \cong 0.01$ ) grafiği

**Çizelge 5.20. I-Chart grafiğine ilişkin 2015-2016 grip mevsimi için haftalık grip faaliyet durumu**

Hafta	Karar	Hafta	Karar	Hafta	Karar	Hafta	Karar
40	Normal	1	Normal	14	Olağandışı	27	Normal
41	Normal	2	Normal	15	Olağandışı	28	Normal
42	Normal	3	Normal	16	Olağandışı	29	Normal
43	Normal	4	Normal	17	Olağandışı	30	Normal
44	Normal	5	Normal	18	Olağandışı	31	Normal
45	Normal	6	Normal	19	Olağandışı	32	Normal
46	Normal	7	Olağandışı	20	Olağandışı	33	Normal
47	Normal	8	Olağandışı	21	Normal	34	Olağandışı
48	Normal	9	Olağandışı	22	Normal	35	Normal
49	Normal	10	Olağandışı	23	Normal	36	Normal
50	Normal	11	Olağandışı	24	Normal	37	Normal
51	Normal	12	Olağandışı	25	Normal	38	Normal
52	Normal	13	Olağandışı	26	Normal	39	Normal

**Çizelge 5.21. Kalite kontrol grafikleri yöntemlerinin verdiği kararlar için sınıflandırma tablosu**

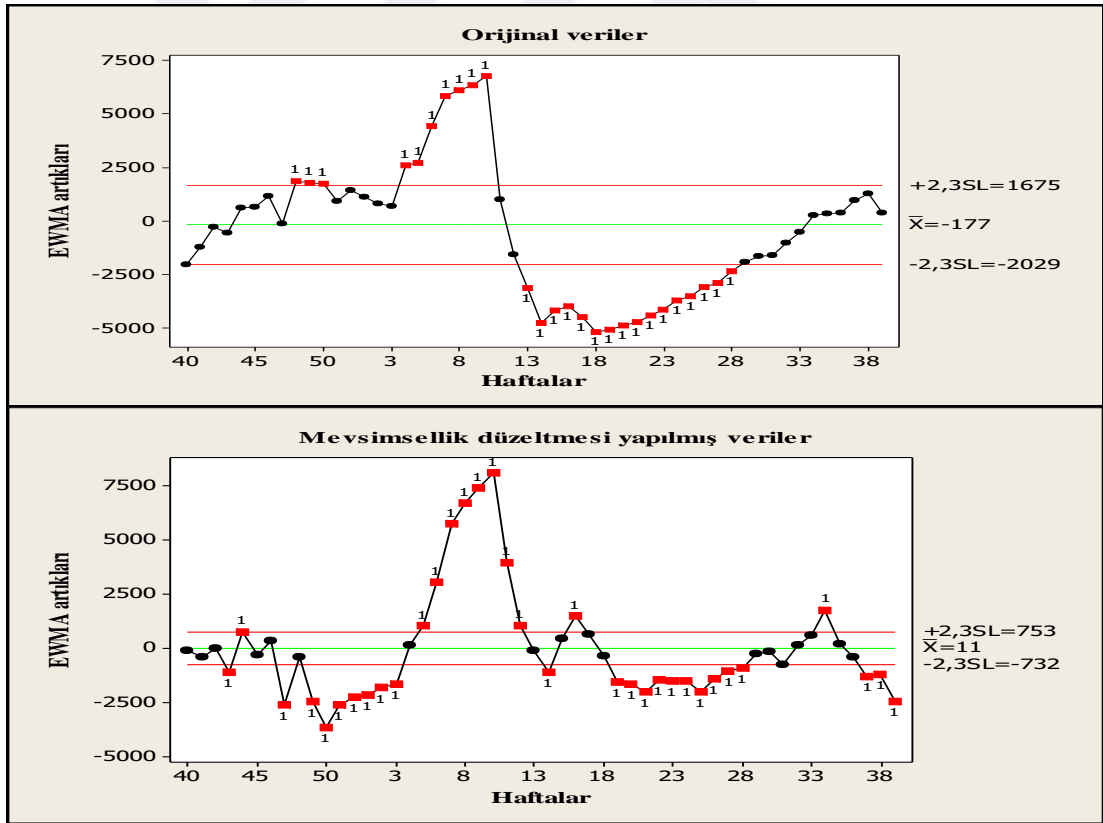
Yöntem	Karar	Gerçek Durum(CDC)		Toplam	Duyarlılık	Seçicilik
		Normal	Olağandışı			
I-Chart	Normal	TN=31	FN=6	37	0.5385	0.7949
	Olağandışı	FP=8	TP=7	15		
	Toplam	39	13	52		

### 5.3.3 CDC grip verilerinin Ewma yöntemiyle elde edilmiş artıklara Shewhart I-Chart grafiğinin uygulanması

Süreç kontrol yöntemlerinin biyoizleme verilerine uygulanmasındaki bir yaklaşımda, Ewma yöntemiyle elde edilen artıkların I-Chart ya da Cusum grafik yöntemleriyle izlenmesidir (Shumeli ve Burkom, 2010). Bu bağlamda hem orijinal CDC grip verilerine hem de bu verilerin mevsimsellik düzeltilmesi yapılmış haline ilk olarak Ewma yöntemi uygulanarak tahminler elde edilmiştir. Daha sonra bu tahminler kullanılarak elde edilen tahmin hataları, teorik yanlış alarm hızı 0.01 yani kontrol sınırları  $2.33\sigma$  olacak şekilde I-Chart ile izlenmiştir.

Bu işlemleri özetlemek gerekirse; ilk olarak orijinal referans dönemi verilerine ( $y_i$   $i = 1,2,3, \dots, 156$ ) Ewma kontrol grafiği ( $\lambda = 0.20$  ve  $3\sigma$  kontrol sınırları) uygulanmış ve Ewma istatistiğinin ( $\hat{y}_i$   $i = 1,2,3, \dots, 156$ ) değerleri, ortalama, standart sapmanın tahmini elde edilmiştir. Daha sonra ilgili ortalama ve standart sapma

ile Ewma kontrol grafiği 2015-2016 grip mevsimine ilişkin izleme dönemi verilerine uygulanarak, Ewma istatistiğinin değerleri elde edilmiştir. Hem referans hem de izleme dönemi verileri, hesaplanan Ewma istatistiğinin değerlerinden çıkarılarak Ewma yöntemine ilişkin tahmin hataları elde edilmiştir. Referans dönemine ilişkin tahmin hataları I-Chart ile izlenerek ortalama ve standart sapma tahminleri elde edilmiştir. Son olarak bir önceki adımda bulunan referans dönemi artıklarına ilişkin ortalama ve standart sapma tahminleri kullanılarak, teorik yanılı düzeyi  $\alpha = 0.01$  ( $2.33\sigma$  kontrol sınırları) olacak şekilde izleme dönemi artıkları I-Chart ile izlenmiştir. Bütün bu işlemler mevsimselliği giderilmiş referans dönemi verileri için de uygulanmıştır (Şekil 5.18.). Elde edilen sonuçlara ilişkin karar tablosu Çizelge 5.22.'de verilmektedir. Çizelge 5.23. elde edilen sonuçlar CDC kurumunun sonuçları ile karşılaştırılmaktadır. Bu işlemlerde kullanılan Ewma tahminleri ve artıkları EK İ.'de verilmektedir.



Şekil 5.18. 2015-2016 izleme dönemi Ewma artıkları için teorik kontrol sınırları kullanan I-Chart( $\alpha = 0.01$ ) grafiği

**Çizelge 5.22. I-Chart grafiğine ilişkin 2015-2016 grip mevsimi için haftalık grip faaliyet durumu**

Hafta	Karar	Hafta	Karar	Hafta	Karar	Hafta	Karar
40	Normal	1	Normal	14	Normal	27	Normal
41	Normal	2	Normal	15	Normal	28	Normal
42	Normal	3	Normal	16	Olağandışı(B)	29	Normal
43	Normal	4	Olağandışı(A)	17	Normal	30	Normal
44	Olağandışı(B)	5	Olağandışı(A,B)	18	Normal	31	Normal
45	Normal	6	Olağandışı(A,B)	19	Normal	32	Normal
46	Normal	7	Olağandışı(A,B)	20	Normal	33	Normal
47	Normal	8	Olağandışı(A,B)	21	Normal	34	Olağandışı(B)
48	Olağandışı(A)	9	Olağandışı(A,B)	22	Normal	35	Normal
49	Olağandışı(A)	10	Olağandışı(A,B)	23	Normal	36	Normal
50	Olağandışı(A)	11	Olağandışı(B)	24	Normal	37	Normal
51	Normal	12	Olağandışı(B)	25	Normal	38	Normal
52	Normal	13	Normal	26	Normal	39	Normal

A: Orijinal veriler B: Mevsimsellik düzeltmesi yapılmış veriler

**Çizelge 5.23. Kalite kontrol grafikleri yöntemlerinin verdiği kararlar için sınıflandırma tablosu**

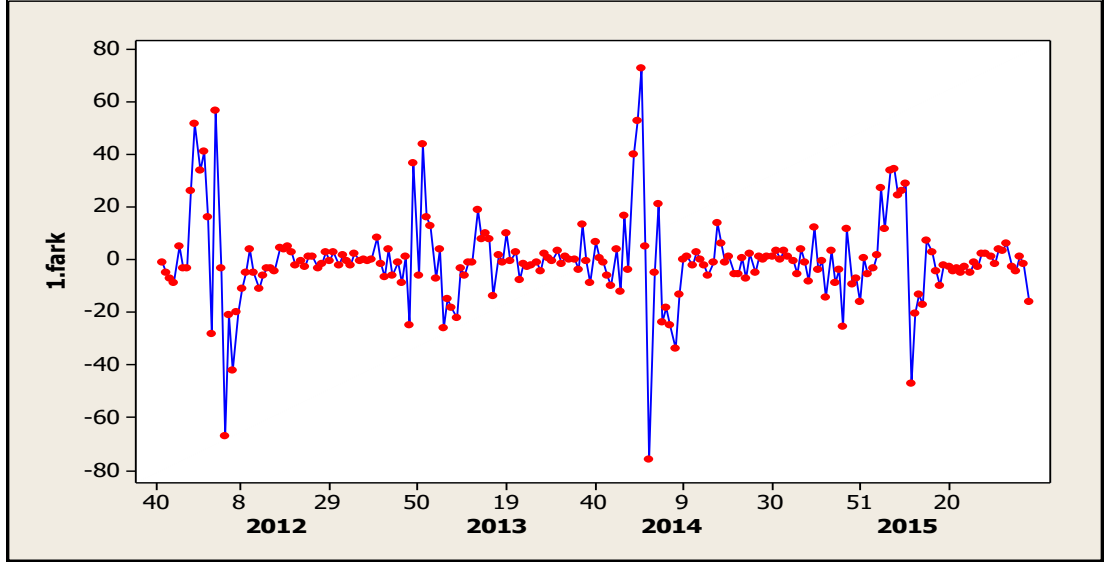
Yöntem	Karar	Gerçek Durum(CDC)		Toplam	Duyarlılık	Seçicilik
		Normal	Olağandışı			
Ewma artıklı I-Chart(A)	Normal	TN=36	FN=6	42	0.5385	0.9231
	Olağandışı	FP=3	TP=7	10		
	Toplam	39	13	52		
Ewma artıklı I-Chart(B)	Normal	TN=36	FN=5	41	0.6154	0.9231
	Olağandışı	FP=3	TP=8	11		
	Toplam	39	13	52		

A: Orijinal veriler B: Mevsimsellik düzeltmesi yapılmış veriler

### 5.3.4 Rossi vd. (1999) Cusum yönteminin CDC grip verilerine uygulanması

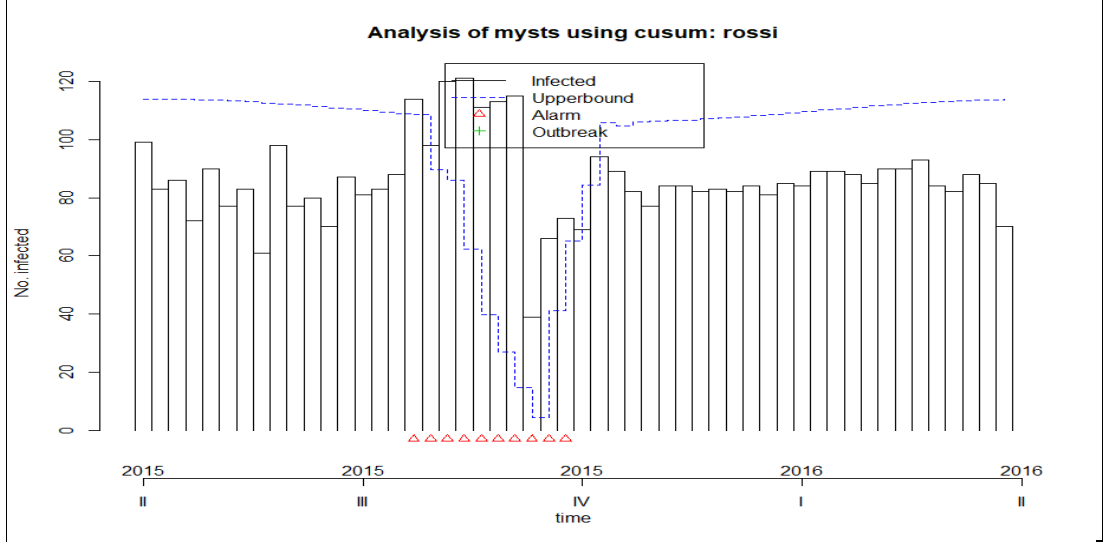
Bu alt bölümde Rossi vd. (1999) zamana bağlı olarak ortalaması değişen Poisson dağılımlı veriler için önerilen Cusum yöntemi, CDC grip verilerine R programının surveillance paketi aracılığıyla uygulanmaktadır. Süreç kontrolünde kullanılan Cusum grafiği verilerin bağımsız ve normal dağılımı olmasını gerektirmektedir. Bu nedenle Rossi vd. (1999) aynı dağılımlı bağımsız Poisson verilerini, normal dağılımlı hale getirmek için dönüşüm önermişlerdir. Başka dönüşümlerde söz konusudur (Höhle, 2007).

Cusum yönteminin sağlıklı bir biçimde uygulanabilmesi için CDC grip verilerinin mevsimsellik gibi zamana bağlı etkilerinden arındırılması gerekmektedir. Bu amaçla uygulanan Serfling regresyon modeli ile artıklar elde edilmiştir. Elde edilen artıklar özilişkili olduğundan Cusum algoritmasının işleyişine olabilecek olumsuz etkisini azaltmak için bu artıkları 1. farkları alınmıştır (Şekil 5.19.).



Şekil 5.19. Mevsimsellik düzeltilmesi yapılmış artıkların 1. farklarının zaman serisi grafiği

Şekil 5.19.'da zamana bağlı etkilerin halen var olduğu görülmektedir. Böyle bir durumu dikkate alabilmek için ortalamaların zamana bağlı olarak değiştiği Poisson regresyon modeli Cusum algoritmasında kullanılmaktadır. Böylelikle verilerin bağımsız aynı dağılımlı Poisson olma varsayımı esnetilebilmektedir. Algoritmanın uygulanabilmesi için verilerin tamsayı ve pozitif olması gerektiğinden, referans dönemine ilişkin artıkların standart sapmasının 5 katı ( $5 \times 17.2850$ ) artıkların tamamına eklenmiş ve en yakın tamsayıya yuvarlanmıştır (Ek G.). Daha sonra  $ARL_0 = 100$  ve  $ARL_1 = 2$  olacak biçimde Cusum grafiğinin  $h$  ve  $k$  parametreleri R programı aracılığıyla 1.39 ve 1.09 olarak bulunmuştur. Bu parametreler ile R programında surveillance paketinde bulunan “algo.cusum” algoritması çalıştırılmıştır. Elde edilen sonuçlar Şekil 5.20. ve Çizelge 5.24.'de verilmektedir. CDC kurumunun sonuçları ile Cusum yönteminden elde edilen sonuçlar Çizelge 5.25.'de karşılaştırılmalı olarak verilmektedir.



Şekil 5.20. Rossi vd. (1999) Cusum yönteminin Serfling artıklarının 1. farklarına uygulanması

Çizelge 5.24. Rossi vd. (1999) Cusum yönteminin verdiği kararlar için sınıflandırma tablosu

Yöntem	Karar	Gerçek Durum(CDC)		Toplam	Duyarlılık	Seçicilik
		Normal	Olağandışı			
Rossi vd. (1999)	Normal	TN=39	FN=3	42	0.7692	1.0000
	Olağandışı	FP=0	TP=10	10		
Cusum	Toplam	39	13	52		

Çizelge 5.25. Rossi vd. (1999) Cusum yöntemine ilişkin 2015-2016 grip mevsimi için haftalık grip faaliyet durumu

Hafta	Karar	Hafta	Karar	Hafta	Karar	Hafta	Karar
40	Normal	1	Normal	14	Normal	27	Normal
41	Normal	2	Normal	15	Normal	28	Normal
42	Normal	3	Normal	16	Normal	29	Normal
43	Normal	4	Olağandışı	17	Normal	30	Normal
44	Normal	5	Olağandışı	18	Normal	31	Normal
45	Normal	6	Olağandışı	19	Normal	32	Normal
46	Normal	7	Olağandışı	20	Normal	33	Normal
47	Normal	8	Olağandışı	21	Normal	34	Normal
48	Normal	9	Olağandışı	22	Normal	35	Normal
49	Normal	10	Olağandışı	23	Normal	36	Normal
50	Normal	11	Olağandışı	24	Normal	37	Normal
51	Normal	12	Olağandışı	25	Normal	38	Normal
52	Normal	13	Olağandışı	26	Normal	39	Normal

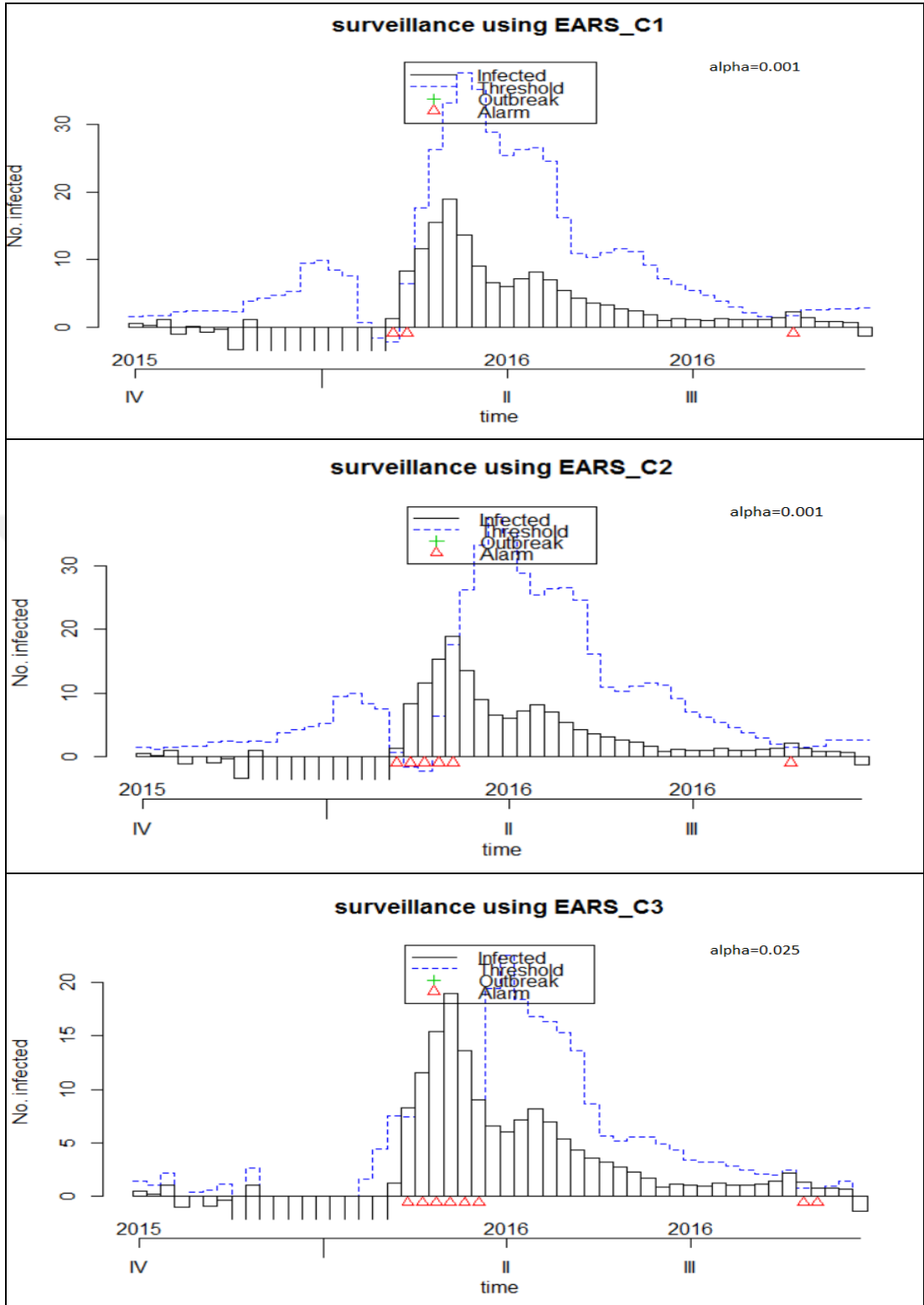
### 5.3.5 Ears C1, C2 ve C3 Yöntemlerinin CDC Grip Verilerine Uygulanması

Bir tür Shewhart grafikleri olarak kabul edilebilen ve kısa referans dönemi verileriyle izleme yapabilen Ears C1, C2 ve C3 yöntemleri, daha çok günlük izleme verilerine uygulanmasına rağmen burada 2015-2016 grip mevsimine ilişkin haftalık grip ve grip benzeri belirtilere sahip kişi sayısını izlemek amacıyla kullanılmıştır. Ears yöntemleri verilerdeki mevsimsellik, trend vb. gibi sistematik yapılardan etkilendiği için, bu sistematik yapılar dikkate alınarak yapılan tahminlerin standartlaştırılmış hatalarına uygulanmıştır. Standartlaştırılmış tahmin hataları Alt Bölüm 5.3.1’de uygulanan Holt-Winters yöntemi ile elde edilmiştir. Ears C1, C2 ve C3 yöntemlerinin uygulanması R programı aracılığıyla gerçekleştirilmiştir. Bu yöntemlerde  $\alpha$ ; C1 ve C2 yöntemleri için 0.001 olarak seçilirken, C3 yönteminde 0.025 olarak seçilmiştir. Bu değerler tavsiye edilen eşik değerlerine yaklaşık olarak karşılık gelmektedir.

Elde edilen sonuçlar grafiksel olarak Şekil 5.21.’de verilmekteyken, 52 haftalık karar tablosu ise Çizelge 5.26.’da verilmektedir. Ayrıca bu yöntemlere ilişkin hesaplanan eşik değerleri ve alarm durumları EK J.’de verilmektedir. Çizelge 5.27. Ears yöntemlerinin sonuçlarının CDC kurumunun sonuçlarıyla karşılaştırılmasını göstermektedir.

**Çizelge 5.26. Ears C1, C2 ve C3 yöntemlerine ilişkin 2015-2016 grip dönemi için haftalık grip faaliyet durumu**

Hafta	Karar	Hafta	Karar	Hafta	Karar	Hafta	Karar
40	Normal	1	Normal	14	Normal	27	Normal
41	Normal	2	Normal	15	Normal	28	Normal
42	Normal	3	Normal	16	Normal	29	Normal
43	Normal	4	Normal	17	Normal	30	Normal
44	Normal	5	Normal	18	Normal	31	Normal
45	Normal	6	Olağandışı(C1,C2)	19	Normal	32	Normal
46	Normal	7	Olağandışı(C1,C2,C3)	20	Normal	33	Normal
47	Normal	8	Olağandışı(C2,C3)	21	Normal	34	Olağandışı(C1,C2)
48	Normal	9	Olağandışı(C2,C3)	22	Normal	35	Olağandışı(C3)
49	Normal	10	Olağandışı(C2,C3)	23	Normal	36	Olağandışı(C3)
50	Normal	11	Olağandışı(C3)	24	Normal	37	Normal
51	Normal	12	Olağandışı(C3)	25	Normal	38	Normal
52	Normal	13	Normal	26	Normal	39	Normal



Şekil 5.21. Holt-Winters yönteminden elde edilen standartlaştırılmış hatalara ilişkin Ears C1, C2 ve C3 grafikleri



**Çizelge 5.27. Ears yöntemlerinin verdiği kararlar için sınıflandırma tablosu**

Yöntem	Karar	Gerçek Durum(CDC)		Toplam	Duyarlılık	Seçicilik
		Normal	Olağandışı			
Ears C1	Normal	TN=38	FN=11	49	0.1538	0.9744
	Olağandışı	FP=1	TP=2	3		
	Toplam	39	13	52		
Ears C2	Normal	TN=38	FN=8	46	0.3846	0.9744
	Olağandışı	FP=1	TP=5	6		
	Toplam	39	13	52		
Ears C3	Normal	TN=37	FN=7	44	0.4615	0.9487
	Olağandışı	FP=2	TP=6	8		
	Toplam	39	13	52		

#### **5.4 Hipotez Testlerine Dayanan Yaklaşımların CDC Grip Verilerine Uygulanması**

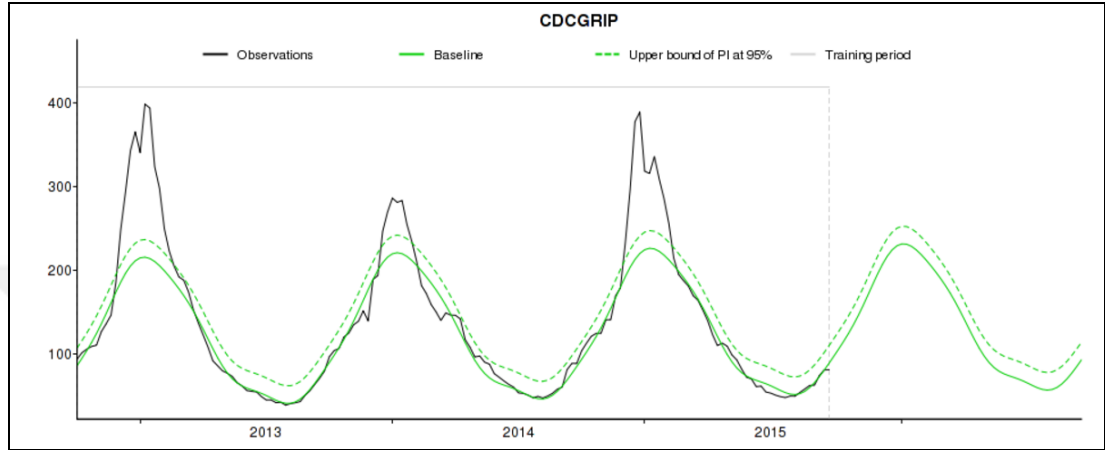
Bu alt bölümde OutbreakP ve Glnb yöntemleri, CDC kurumunun 2015-2016 grip mevsimi haftalık verilerinin izlenmesi amacıyla uygulanmaktadır. Bu yöntemler olabilirlik oranı ya da genelleştirilmiş olabilirlik oranına dayanan testler uygulamaktadır.

##### **5.4.1 OutbreakP yönteminin CDC grip verilerine uygulanması**

Salgın başlangıcını belirlemek amacıyla önerilmiş yöntemlerden biri de Frisen ve Anderson (2009) tarafından OutbreakP olarak adlandırılan ve referans dönemi gerektirmeyen yöntemdir. Alt Bölüm 3.4'de belirtildiği gibi bu yöntem salgının başlangıcı için monoton artan regresyon modelini kullanırken, verilerin Poisson dağılımlı olduğunu kabul etmektedir. OutbreakP en çok olabilirlik temelli yarı parametrik bir izleme yöntemi olarak değerlendirilmektedir. Bu yöntem mevsimsellik özelliği göstermeyen ve bağımsız kabul edilebilen izleme verilerine doğrudan uygulanabileceği gibi, mevsimsellik özelliği giderilmiş izleme verilerinin artıklarına da uygulanabilmektedir (Frisen vd, 2009).

Tez kapsamında dikkate alınan CDC grip verileri mevsimsellik etkisi gösterdiğinden OutbreakP yönteminin bu verilere doğrudan uygulanması uygun değildir. Ayrıca

Poisson dağılımı ile ilgili hesaplamaları kolaylaştırabilmek amacıyla haftalık veriler 100'e bölünerek yeniden ölçeklendirilmiştir. Daha sonra 2012-2013, 2013-2014 ve 2014-2015 referans dönemi verileri için Serfling regresyon modeli ile hem referans dönemi verileri hem de 2015-2016 izleme dönemi verileri için artıklar elde edilmiştir (Ek K.).



Şekil 5.22. Yeniden ölçeklendirilmiş CDC grip verilerinin Serfling regresyon eğrisi



Şekil 5.23. Yeniden ölçeklendirilmiş CDC grip verilerinin Serfling regresyon model tahmini

Elde edilen model tahminleri ve regresyon eğrisi Şekil 5.22. ve Şekil 5.23.'de verilmektedir. Bütün bu işlemler Pierre Louis Araştırma Enstitüsü'ndeki bir ekip

tarafından geliştirilen <http://marne.u707.jussieu.fr/periodic> adresindeki web uygulaması üzerinden gerçekleştirilmiştir. Ayrıca, referans yıllarındaki olası salgınlardan kaynaklanabilecek yüksek vaka sayılarının etkilerini azaltmak amacıyla, model parametreleri kestirilmeden önce verilerin üst %15'lik kısmı analizden dışlanmıştır.

2015-2016 izleme dönemi artıkları negatif değerler içermektedir. Poisson dağılımı pozitif ve tam sayı değerler gerektirdiğinden, referans verilerine ilişkin artıkların standart sapmasının 3 katı ( $3 \times 41.2569 \cong 124$ ) izleme dönemi artıklarına eklenmiş ve artıklar tamsayı haline getirilmiştir. OutbreakP yöntemini uygulayabilmek için alarm sınırlarının belirlenmesi gerekir. Bu amaçla aşağıda tanımlanan  $PV(t)$  ya da  $CED(t)$  gibi,

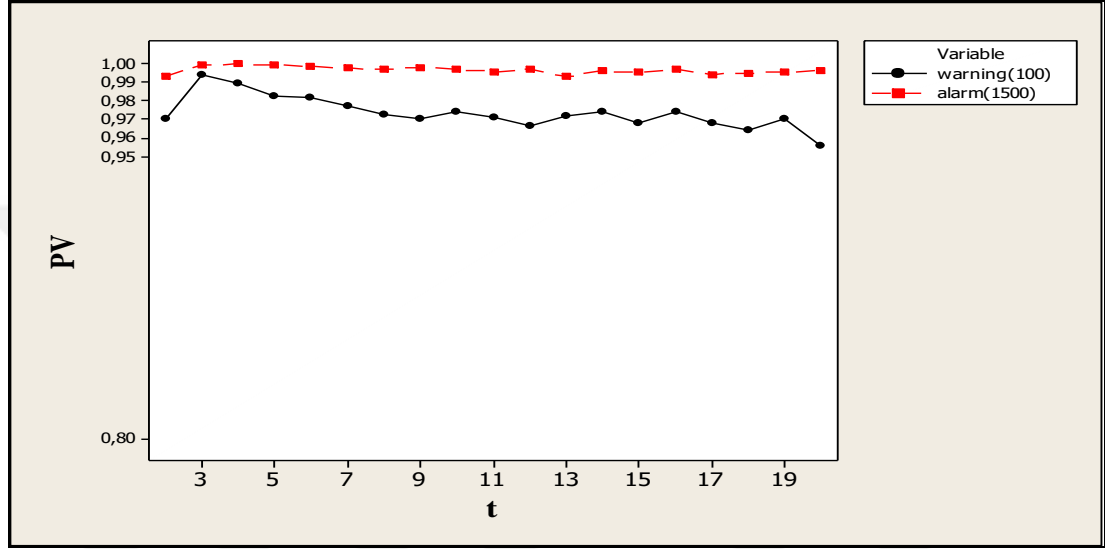
$$PV(t) = P(C|t_A = t) \quad CED(t) = E[t_A - \tau | t_A \geq \tau, \tau = t]$$

belirli bir ölçüte kabul edilebilen bir değer verecek alarm sınırları seçilir. Frisen vd. (2009) İsveç'teki grip verilerine OutbreakP yöntemini uygularken, bu verileri modelleyen bir simülasyon çalışmasıyla  $PV(t)$  ölçütüne sırasıyla 0.95 ve yaklaşık 1 değerlerini veren 150 uyarı sınırını ve 9000 alarm sınırını belirlemişlerdir. Bu tez çalışmasında da  $PV(t)$  ölçütü sınırları belirlemek amacıyla kullanılmıştır. CDC grip verilerine ilişkin salgın başlangıcı Frisen vd. (2009) gibi modellenmiştir.

$$X(t) \sim \begin{cases} Poisson(\mu_0), & t < \tau \\ Poisson(\mu(t)), & t \geq \tau \end{cases}$$

Burada;  $\mu(t) = e^{(\beta_0 + \beta_1(t - \tau + 1))}$  ve  $\tau$  ise salgının başlangıcını gösteren değişkendir. Frisen vd. (2009),  $\tau$  değişkenini geometrik dağılımlı kabul etmiş ve bu dağılımın parametresi, referans verilerindeki salgın anına kadar geçen sürenin ortalamasını kullanarak CDC grip verileri için  $v = 0.1428$  olarak tahmin edilmiştir. İzleme dönemine en yakın olan 2014-2015 grip mevsimi verileri  $\beta_0$  ve  $\beta_1$  parametrelerinin tahmini için dikkate alınmıştır. Bu amaçla ilgili döneme ilişkin salgının başladığı dönemde; verilerin pozitif hale getirilmiş artıklarının logaritmaları ile  $(t - \tau + 1)$  arasındaki doğrusal regresyon modeli sonucu  $\hat{\beta}_0 = 4.53$  ve  $\hat{\beta}_1 = 0.21$  parametre

tahminleri elde edilmiştir. Ayrıca  $\mu_0$ , referans dönemlerine ilişkin 40. haftaların ortalamalarına 124 sabit sayısının eklenmesi ile  $\hat{\mu}_0 = 131$  bulunmuştur. R kodlarıyla yazılan ve salgının başlangıcını modelleyen bir program aracılığıyla, 100000 tekrarlı simülasyon çalışmaları yapılmıştır. Bu çalışmalar sonucunda alarm sınırını 100 olarak alınmasının  $PV(t)$  değerini 0.95 ve üzerine çıkardığı, 1500 olarak alınması ise  $PV(t)$  değerini 0.99 ve üzerine çıkardığı tespit edilmiştir (Şekil 5.24.).



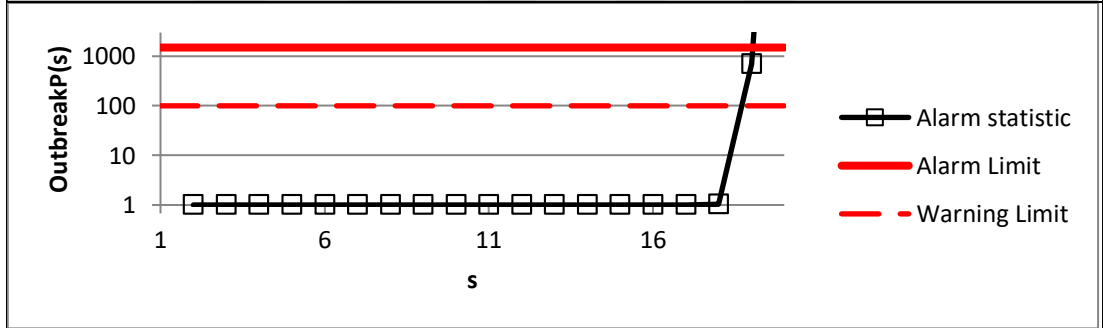
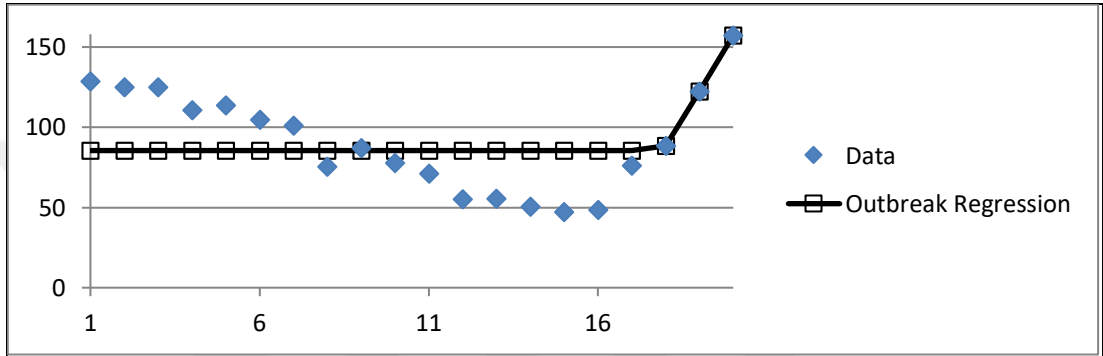
Şekil 5.24. PV ölçütünün 100 ve 1500 alarm sınırları için simülasyondan elde edilmiş değerleri

İzleme dönemi olan 2015-2016 grip mevsimi haftalık verilerine ilişkin pozitif hale getirilmiş artıklara, bu sınırlar kullanılarak Frisen vd. (2009) tanımlanan OutbreakP programı ([www.statistics.gu.se/surveillance](http://www.statistics.gu.se/surveillance)) uygulanmıştır (Şekil 5.25.). Amaç salgının başlangıcını belirlemek olduğundan, sinyal verilmeye başlandığında izleme durdurulmuştur yani ilk 20 hafta izleme yapılmıştır.

Şekil 5.25. incelendiğinde; OutbreakP algoritması salgının başlangıcını alarm sınırı 100 iken 6. hafta, alarm sınırı 1500 iken 7. hafta olarak belirlemiştir. CDC kurumuna göre salgın 3. haftada başladığından, bu algoritma 3-4 haftalık bir gecikme ile salgının başlangıcını tespit etmektedir. Şekil 5.25. ve Şekil 5.27.'de OutbreakP algoritmasının çıktısında görülen izleme zamanlarının temsil ettiği haftalar Çizelge 5.28.'de verilmektedir.

Çizelge 5.28. İzleme zamanlarının temsil ettiği haftalar

Sıra	Hafta	Sıra	Hafta
1	40	11	50
2	41	12	51
3	42	13	52
4	43	14	1
5	44	15	2
6	45	16	3
7	46	17	4
8	47	18	5
9	48	19	6
10	49	20	7



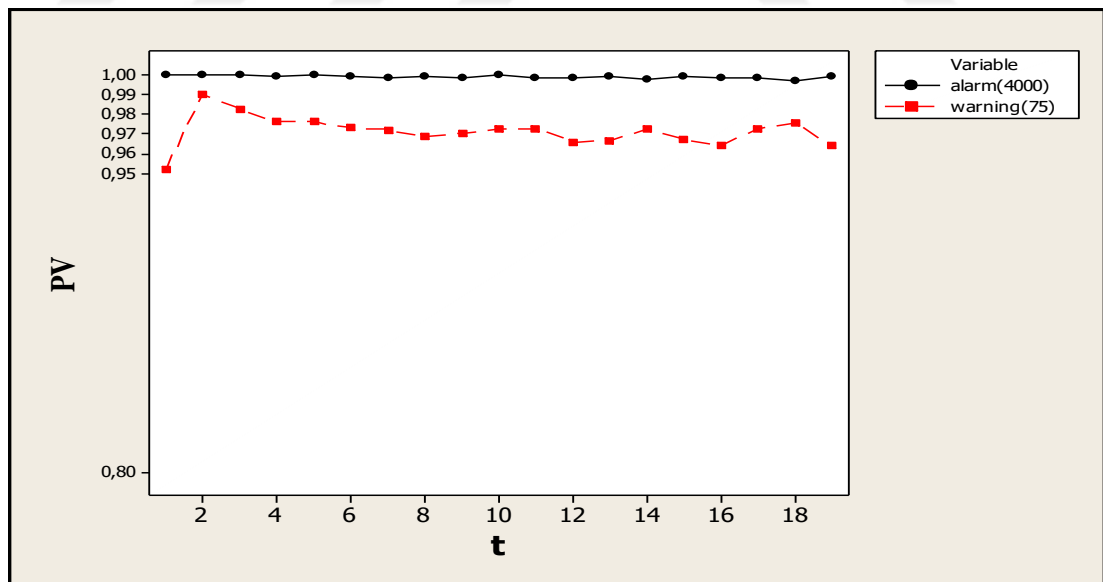
Time	Alarmstat	Warning	Alarm
1-1	1	0	0
1-2	1	0	0
1-3	1	0	0
1-4	1	0	0
1-5	1	0	0
1-6	1	0	0
1-7	1	0	0
1-8	1	0	0
1-9	1	0	0
1-10	1	0	0
1-11	1	0	0
1-12	1	0	0
1-13	1	0	0
1-14	1	0	0
1-15	1	0	0
1-16	1	0	0
1-17	1	0	0
1-18	1,0434897	0	0
1-19	706,6359	1	0
1-20	9,061E+11	1	1

Şekil 5.25. 2015-2016 grip mevsimi için OutbreakP yöntemi sonuçları

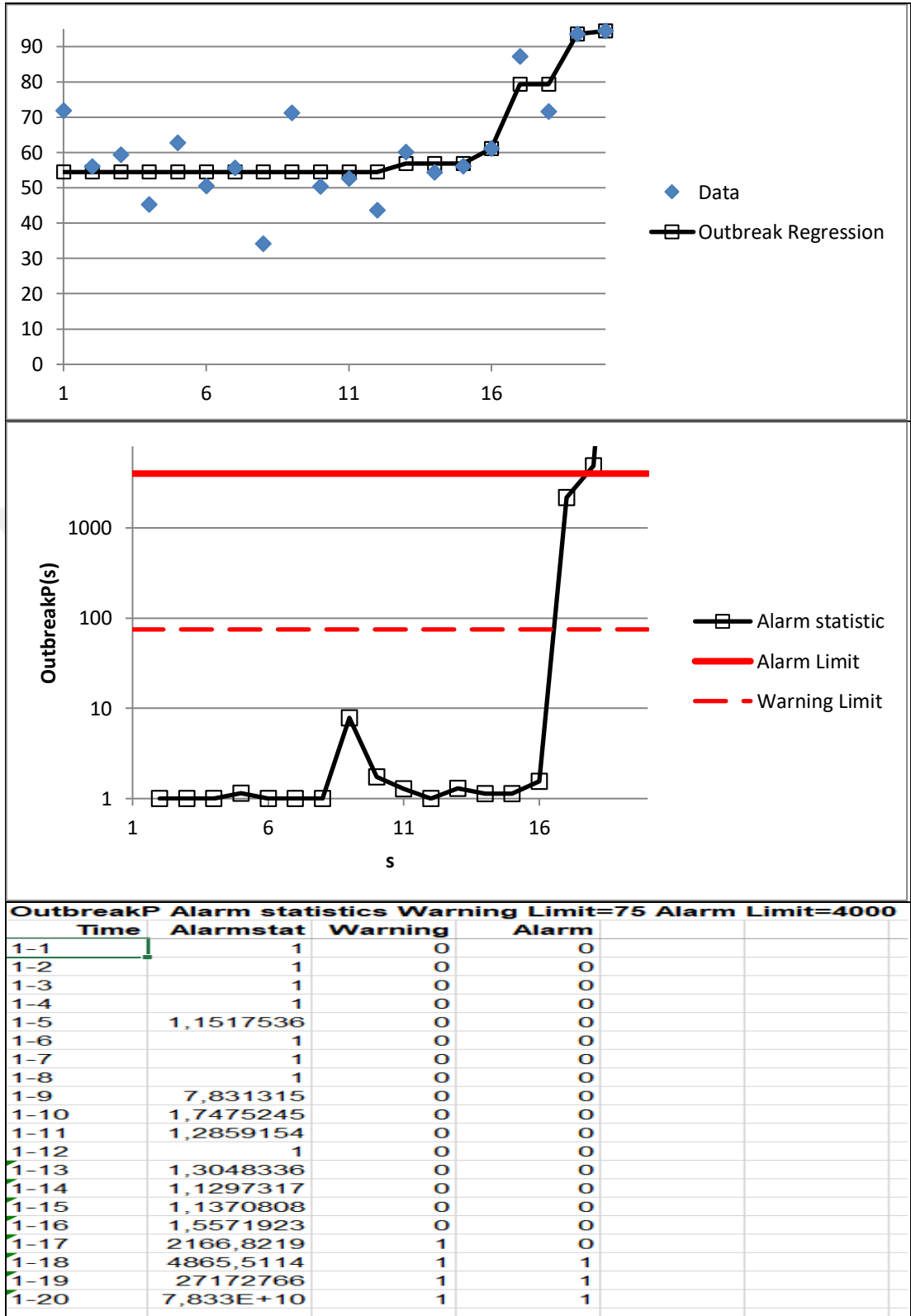
İzlenen pozitif hale getirilmiş artıklar özilişkilidir. Bu özilişki yapısının OutbreakP yöntemine olabilecek olumsuz etkisini azaltmak amacıyla, artıkların birinci farklarının

izlenmesi daha uygun olduğu düşünülebilir. Ayrıca fark alma işlemi salgın öncesi sabit bir ortalama ve salgın başlangıcında monoton artan bir ortalama gerektiren monoton regresyon modelini de bozmamaktadır.

2014-2015 grip mevsimine ilişkin artıkların 1. farkları dikkate alınarak elde edilen değerler; bu farklara ilişkin standart sapmanın 3 katı ( $3 \times 19.85 \cong 59.55$ ) eklenerek pozitif hale getirilmiştir. Bu verilerin doğal logaritması  $y = \ln(y)$  alınmış ve her bir veri için  $x = (t - \tau + 1)$  hesaplanmıştır. Daha sonra ilgili  $\ln(y)$  verilerine ilişkin monoton artışın olduğu zaman diliminde; verilerin pozitif hale getirilmiş artıklarının logaritmaları ile  $(t - \tau + 1)$  arasındaki doğrusal regresyon modeli sonucu  $\hat{\beta}_0 = 3.60$  ve  $\hat{\beta}_1 = 0.273$  parametre tahminleri elde edilmiştir. Ayrıca  $\mu_0$ , referans dönemlerine ait 40. haftalara ilişkin fark ortalamalarına 59.55 sabit sayısının eklenmesi ile  $\hat{\mu}_0 = 68.83$  bulunmuştur. Bu parametreleri kullanan ve salgının başlangıcını modelleyen R kodlarıyla yazılmış bir program aracılığıyla, 100000 tekrarlı simülasyon çalışmaları yapılmıştır. Bu çalışmalar sonucunda alarm sınırını 75 olarak alınmasının  $PV(t)$  değerini 0.95 ve üzerine taşıdığı, 4000 olarak alınması ise  $PV(t)$  değerini 0.99 ve üzerine taşıdığı görülmüştür (Şekil 5.26.).



Şekil 5.26. PV ölçütünün 75 ve 4000 alarm sınırları için simülasyondan elde edilmiş değerleri



Şekil 5.27. 2015-2016 grip mevsimi için OutbreakP yöntemi sonuçları

Bu sınırlar kullanılarak izleme dönemi olan 2015-2016 grip mevsimi haftalık verilerine ilişkin pozitif hale getirilmiş farklara, Frisen vd. (2009) tarafından tanımlanan OutbreakP programı ([www.statistics.gu.se/surveillance](http://www.statistics.gu.se/surveillance)) uygulanmıştır (Şekil 5.27.). İlk 20 hafta için izleme yapılmıştır. Şekil 5.27. incelendiğinde; salgının başlangıcını alarm sınırı 75 iken 4. hafta, alarm sınırı 4000 iken 5. hafta olarak belirlemiştir. CDC kurumuna göre salgın 3. haftada başladığından, OutbreakP algoritması 1-2 haftalık bir gecikme ile salgının başlangıcını tespit etmektedir.

#### 5.4.2 Genelleştirilmiş olabilirlik oranı regresyon grafiği

Bu alt bölümde zaman içinde değişen ortalamalara sahip Poisson ya da negatif binom dağılımlı zaman serileri için; Höhle ve Paul (2008) tarafından önerilmiş olan genelleştirilmiş olabilirlik oranı istatistiklerinin kullanıldığı Cusum grafiği, CDC grip verilerine uygulanacaktır. Bu tür yöntemler regresyon kontrol grafikleri olarak da adlandırılmaktadır. Veri olarak Alt Bölüm 5.3.4’de kullanılan Serfling artışlarının 1. farklarından oluşan ve pozitif hale getirilmiş olan veriler kullanılmıştır (Ek G.).

Bu yöntem uygulanırken verilerin Poisson dağılımlı olduğu kabul edildi ve izleme dönemi verilerinde kontrol dışılığı gösterecek artış yönündeki değişim miktarı önceden belirlenmedi. Ayrıca verilerde trend etkisinin olmadığı ancak olası mevsimsellik etkisini dikkate alacak Poisson regresyonu için çeşitli genelleştirilmiş doğrusal modeller ( $S = 1, S = 2$  ve  $S = 3$ ) denendi (Çizelge 5.29.). Veriler haftalık olduğundan  $period = 52$  olarak alındı.

$$Y_t \sim \text{Poisson}(\mu_{0,t})$$

$$\log \mu_{0,t} = \beta_0 + \sum_{s=1}^S \left[ \beta_{2s} \cos\left(\frac{2\pi st}{\text{Period}}\right) + \beta_{2s+1} \sin\left(\frac{2\pi st}{\text{Period}}\right) \right]$$

Çizelge 5.29. incelendiğinde, bu 3 model arasında  $AIC$  değeri en küçük olan model ( $S = 3$ ) seçildi.

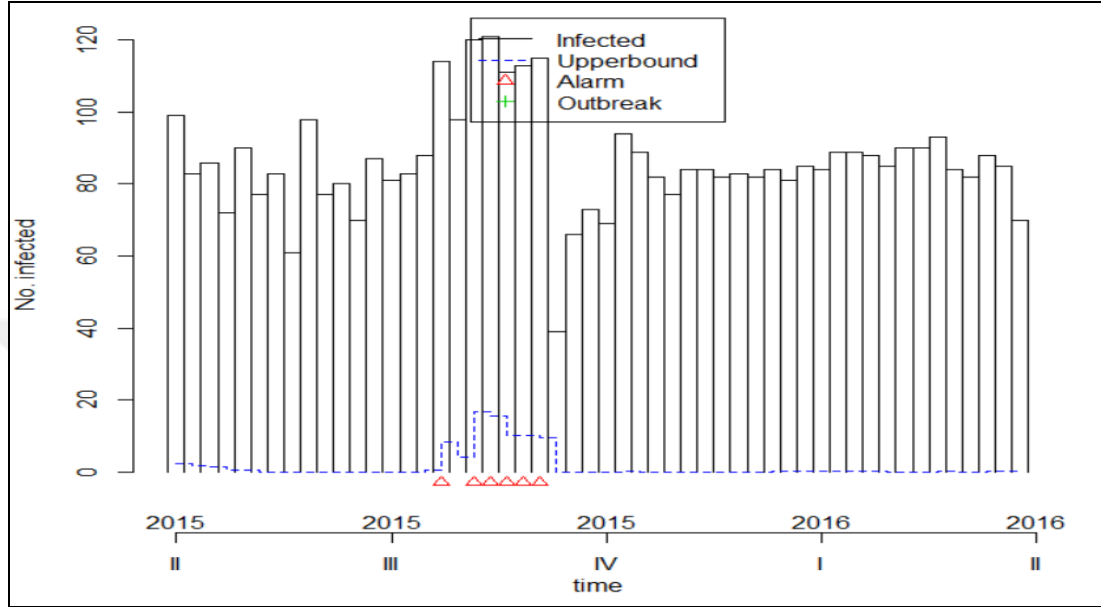


Çizelge 5.29. Referans dönemi verileri için Poisson regresyon model kestirimleri

<pre>Call: glm(formula = y ~ 1 + x1 + x2, family = poisson())  Deviance Residuals:     Min       1Q   Median       3Q      Max -10.457   -0.538   -0.073    0.328    6.902  Coefficients:             Estimate Std. Error z value Pr(&gt; z ) (Intercept)  4.455470   0.008657  514.64 &lt; 2e-16 *** x1           0.040900   0.012284    3.33 0.00087 *** x2           0.001467   0.012202    0.12 0.90432 --- Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1  (Dispersion parameter for poisson family taken to be 1)      Null deviance: 574.7  on 154  degrees of freedom Residual deviance: 563.6  on 152  degrees of freedom AIC: 1541.2</pre>					
<pre>Call: glm(formula = y ~ 1 + x1 + x2 + x3 + x4, family = poisson())  Deviance Residuals:     Min       1Q   Median       3Q      Max -10.5749  -0.7057    0.0409    0.6126    6.4089  Coefficients:             Estimate Std. Error z value Pr(&gt; z ) (Intercept)  4.454e+00  8.671e-03  513.677 &lt; 2e-16 *** x1           4.102e-02  1.235e-02    3.322 0.000895 *** x2           8.507e-06  1.216e-02    0.001 0.999442 x3          -1.900e-02  1.229e-02   -1.546 0.122126 x4           7.114e-02  1.222e-02    5.824 5.74e-09 *** --- Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1  (Dispersion parameter for poisson family taken to be 1)      Null deviance: 574.70  on 154  degrees of freedom Residual deviance: 527.26  on 150  degrees of freedom AIC: 1508.9</pre>					
<pre>Call: glm(formula = y ~ 1 + x1 + x2 + x3 + x4 + x5 + x6, family = poisson())  Deviance Residuals:     Min       1Q   Median       3Q      Max -10.8470  -0.5444   -0.1140    0.6112    5.6623  Coefficients:             Estimate Std. Error z value Pr(&gt; z ) (Intercept)  4.450352   0.008703  511.377 &lt; 2e-16 *** x1           0.040550   0.012376    3.277 0.001051 ** x2          -0.003957   0.012194   -0.325 0.745553 x3          -0.018343   0.012305   -1.491 0.136027 x4           0.072297   0.012247    5.903 3.56e-09 *** x5          -0.099582   0.012322   -8.082 6.38e-16 *** x6          -0.045027   0.012228   -3.682 0.000231 *** --- Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1  (Dispersion parameter for poisson family taken to be 1)      Null deviance: 574.70  on 154  degrees of freedom Residual deviance: 448.25  on 148  degrees of freedom AIC: 1433.9</pre>					

Bu modele ilişkin anlamlı olmayan terimler modelden dışlandı. Modele ilişkin parametre kestirimleri kullanılarak 52 haftalık 10000 seri zamanla değişen ortalamalara sahip Poisson olasılık dağılımlı olarak simülasyon ile üretildi. Bu simülasyon programı EK L.'de verilmektedir. Kullanılacak Poisson regresyon grafiğinin süreç kontrol altındayken, 52 haftalık izlemede yanlış sinyal verme

olasılığını 0.05'ten küçük yapan eşik değeri bu simülasyon ile 6 olarak belirlenmiştir. CDC grip verileri için; ilgili parametrelerle Poisson regresyon grafiğini elde etmek amacıyla, R programındaki ilgili algoritma çalıştırılmış ve sonuçlar elde edilmiştir (Çizelge 5.30., Şekil 5.28., Çizelge 5.31.).



Şekil 5.28. CDC grip verilerinin Poisson regresyon ile kontrol grafiği

Çizelge 5.30. Poisson regresyon kontrol grafiği yöntemine ilişkin 2015-2016 grip mevsimi için haftalık grip faaliyet durumu

Hafta	Karar	Hafta	Karar	Hafta	Karar	Hafta	Karar
40	Normal	1	Normal	14	Normal	27	Normal
41	Normal	2	Normal	15	Normal	28	Normal
42	Normal	3	Normal	16	Normal	29	Normal
43	Normal	4	Olağandışı <sub>1</sub>	17	Normal	30	Normal
44	Normal	5	Normal	18	Normal	31	Normal
45	Normal	6	Olağandışı <sub>1</sub>	19	Normal	32	Normal
46	Normal	7	Olağandışı <sub>1</sub>	20	Normal	33	Normal
47	Normal	8	Olağandışı <sub>1</sub>	21	Normal	34	Normal
48	Normal	9	Olağandışı <sub>1</sub>	22	Normal	35	Normal
49	Normal	10	Olağandışı <sub>1</sub>	23	Normal	36	Normal
50	Normal	11	Normal	24	Normal	37	Normal
51	Normal	12	Normal	25	Normal	38	Normal
52	Normal	13	Normal	26	Normal	39	Normal

**Çizelge 5.31. Poisson regresyon kontrol grafiği yönteminin verdiği kararlar için sınıflandırma tablosu**

Yöntem	Karar	Gerçek Durum(CDC)		Toplam	Duyarlılık	Seçicilik
		Normal	Olağandışı			
Poisson regresyon kontrol grafiği	Normal	TN=39	FN=7	46	0.4615	1.0000
	Olağandışı	FP=0	TP=6	6		
	Toplam	39	13	52		

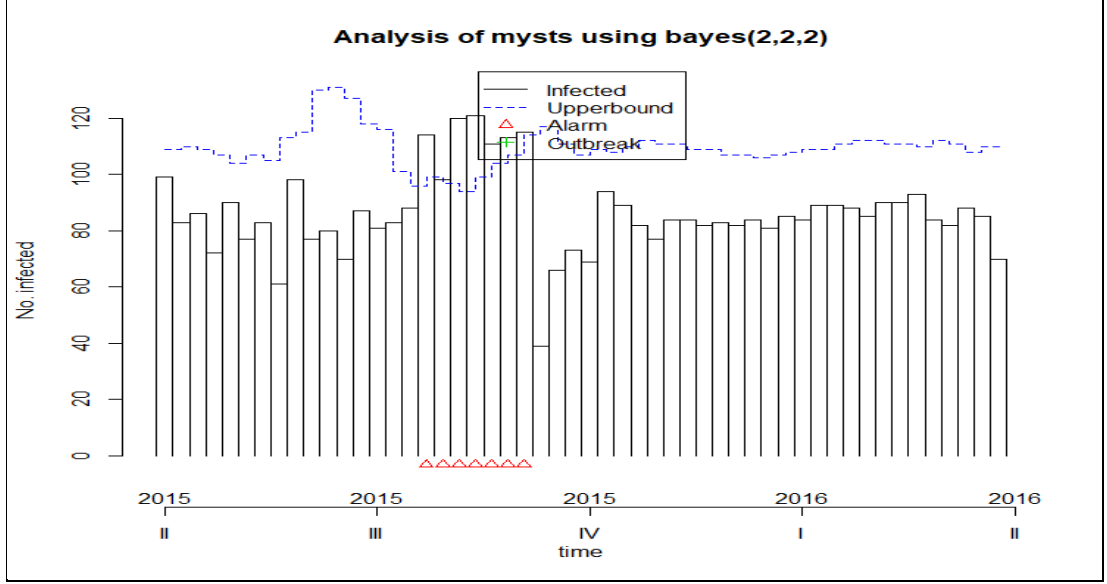
## 5.5 Diğer Yaklaşımların CDC Grip Verilerine Uygulanması

Bu alt bölümde, bayesçi bir yaklaşım, Rki yöntemi ve wavelet yöntemi, CDC kurumunun 2015-2016 grip mevsimi haftalık verilerinin izlenmesi amacıyla uygulanmaktadır.

### 5.5.1 Bayesçi bir sonsal öngörü yaklaşımının CDC grip verilerine uygulanması

Ortalaması aynı Poisson dağılımlı veriler için Riebler (2004) tarafından önerilen bir bayes yöntemi, CDC grip verilerine R programının surveillance paketi aracılığıyla uygulanmaktadır. Höhle (2007) bu yöntemin ayrıntılarını açıklamaktadır. Bayes yönteminin CDC grip verilerine uygulanabilmesi için zamana bağlı olarak değişen etkilerin giderilmesi gerekmektedir. Bu nedenle Alt Bölüm 5.3.4'deki elde edilen mevsimsellik düzeltmesi yapılmış artıkların 1. farklarının pozitif hali (Ek G.) veri olarak kullanılmaktadır.

Öngörü sonsal dağılımının 0.99'uncü kuantili ( $\alpha = 0.01$ ) eşik değeri olmak kaydıyla, geçmişe dönük 2 yıl ( $b = 2$ ) ve izlenecek her nokta için referans dönemindeki yarı pencere genişliği 2 ( $w = 2$ ) alınarak bayes algoritması çalıştırılmıştır. Burada  $b = 2$  seçilmesinin nedeni referans döneminin geçmişe dönük 3 grip mevsimi verilerden oluşmasından dolayıdır. Herhangi bir haftadaki verinin; bu haftanın içinde bulunduğu bir aylık dönemden etkileneceği düşüncesiyle, veriye ilişkin ilgili haftanın öncesindeki 2 haftalık ve sonrasındaki 2 haftalık zaman periyodu dikkate alınmıştır ( $w = 2$ ). Bu parametrelerle elde edilen sonuçlar Şekil 5.29., Çizelge 5.32. ve Çizelge 5.33.'de verilmektedir.



Şekil 5.29. Bayes yönteminden elde edilen izleme grafiği

Çizelge 5.32. Bayes yöntemine ilişkin 2015-2016 grip mevsimi için haftalık grip faaliyet durumu

Hafta	Karar	Hafta	Karar	Hafta	Karar	Hafta	Karar
40	Normal	1	Normal	14	Normal	27	Normal
41	Normal	2	Normal	15	Normal	28	Normal
42	Normal	3	Normal	16	Normal	29	Normal
43	Normal	4	Olağandışı	17	Normal	30	Normal
44	Normal	5	Olağandışı	18	Normal	31	Normal
45	Normal	6	Olağandışı	19	Normal	32	Normal
46	Normal	7	Olağandışı	20	Normal	33	Normal
47	Normal	8	Olağandışı	21	Normal	34	Normal
48	Normal	9	Olağandışı	22	Normal	35	Normal
49	Normal	10	Olağandışı	23	Normal	36	Normal
50	Normal	11	Normal	24	Normal	37	Normal
51	Normal	12	Normal	25	Normal	38	Normal
52	Normal	13	Normal	26	Normal	39	Normal

Çizelge 5.33. Bayes yönteminin verdiği kararlar için sınıflandırma tablosu

Yöntem	Karar	Gerçek Durum(CDC)		Toplam	Duyarlılık	Seçicilik
		Normal	Olağandışı			
Bayes	Normal	TN=39	FN=6	45	0.5385	1.0000
	Olağandışı	FP=0	TP=7	7		
	Toplam	39	13	52		

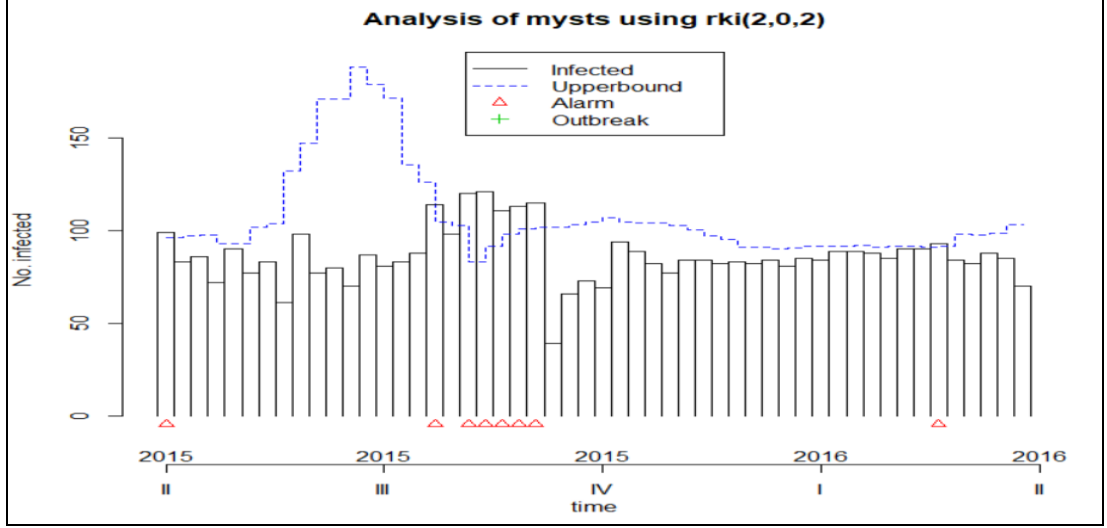
### 5.5.2 Rki yaklaşımının CDC grip verilerine uygulanması

Robert Koch Institute (RKI) tarafından kullanılan Rki yöntemine ilişkin bilgi veren bir yayın bulunmamaktadır (Salmon vd., 2016). Bu yöntem Farrington, Bayes ve CDC yöntemleri gibi, referans kümesi  $R(w, w_0, b)$  olarak geçmişe ait verilerin belirli bir genişlikteki alt kümelerini kullanmaktadır:

$$R(w, w_0, b) = \left( \bigcup_{i=1}^b \bigcup_{j=-w}^w y_{-i:t+j} \right) \cup \left( \bigcup_{k=-w_0}^{-1} y_{0:t+k} \right)$$

Burada  $b$ , geçmişe ilişkin kaç yıl verinin kullanılacağını;  $w$ , her yıldaki ilgili zaman noktasını içeren zaman diliminin yarı genişliğini;  $w_0$  ise içinde bulunan yılda ilgili zaman noktasından hemen önce kaç noktanın referans kümesine dahil edileceğini göstermektedir. Bu yöntem referans kümesi kullanarak mevsimselliği dikkate alabilmektedir. Ancak doğrusal trendi dikkate almamakta ve geçmişteki salgınlardan etkisini azaltmak amacıyla, verilerde ağırlıklandırma yapmamaktadır. Ayrıca verilerde herhangi bir dönüşüm uygulamamaktadır (Salmon vd., 2016).

CDC grip verilerini Rki yöntemiyle analiz edebilmek için, Alt Bölüm 5.3.4'deki gibi Serfling regresyonu artıklarının 1.farklarının pozitif ve tamsayı hale getirilmiş hali kullanılmaktadır (Ek G.). Bu yöntem R programının surveillance paketi aracılığıyla uygulanmıştır. Rki yöntemine ilişkin referans kümesi oluşturmak amacıyla geçmişe dönük 2 yıl ( $b = 2$ ) ve izlenecek her nokta için referans dönemindeki yarı pencere genişliği 2 ( $w = 2$ ) alınmıştır. Burada  $b = 2$  seçilmesinin nedeni referans döneminin geçmişe dönük 3 grip mevsimi verilerden oluşmasından dolayıdır. Mevsimsellik düzeltilmesi yapıldığı için; herhangi bir haftadaki verinin içinde bulunacağı bir aylık dönemden etkileneceği düşüncesiyle, veriye ilişkin ilgili haftanın öncesindeki 2 haftalık ve sonrasındaki 2 haftalık zaman periyodu dikkate alınmıştır ( $w = 2$ ). Bu parametrelerle elde edilen sonuçlar Şekil 5.30., Çizelge 5.34. ve Çizelge 5.35.'te verilmektedir.



Şekil 5.30. Rki yönteminden elde edilen izleme grafiği

Çizelge 5.34. Rki yöntemine ilişkin 2015-2016 grip mevsimi için haftalık grip faaliyet durumu

Hafta	Karar	Hafta	Karar	Hafta	Karar	Hafta	Karar
40	Olağandışı	1	Normal	14	Normal	27	Normal
41	Normal	2	Normal	15	Normal	28	Normal
42	Normal	3	Normal	16	Normal	29	Normal
43	Normal	4	Olağandışı	17	Normal	30	Normal
44	Normal	5	Normal	18	Normal	31	Normal
45	Normal	6	Olağandışı	19	Normal	32	Normal
46	Normal	7	Olağandışı	20	Normal	33	Normal
47	Normal	8	Olağandışı	21	Normal	34	Olağandışı
48	Normal	9	Olağandışı	22	Normal	35	Normal
49	Normal	10	Olağandışı	23	Normal	36	Normal
50	Normal	11	Normal	24	Normal	37	Normal
51	Normal	12	Normal	25	Normal	38	Normal
52	Normal	13	Normal	26	Normal	39	Normal

Çizelge 5.35. Rki yönteminin verdiği kararlar için sınıflandırma tablosu

Yöntem	Karar	Gerçek Durum(CDC)		Toplam	Duyarlılık	Seçicilik
		Normal	Olağandışı			
Rki	Normal	TN=37	FN=7	44	0.4615	0.9487
	Olağandışı	FP=2	TP=6	8		
	Toplam	39	13	52		

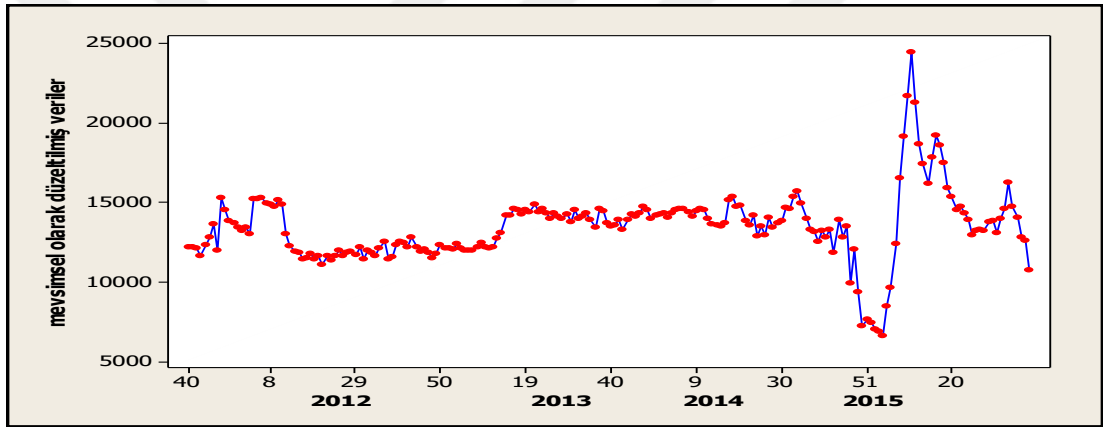
### 5.5.3 Wavelet yaklaşımının CDC grip verilerine uygulanması

Bu alt bölümde CDC grip verilerine wavelet yaklaşımı ileriye dönük olarak uygulanmaktadır. İleriye dönük wavelet yaklaşımından kastedilen Shumeli (2013) tanımladığı geriye doğru uygulamalı durağan kesikli wavelet dönüşümüdür (backward-implemented stationary discrete wavelet transform).

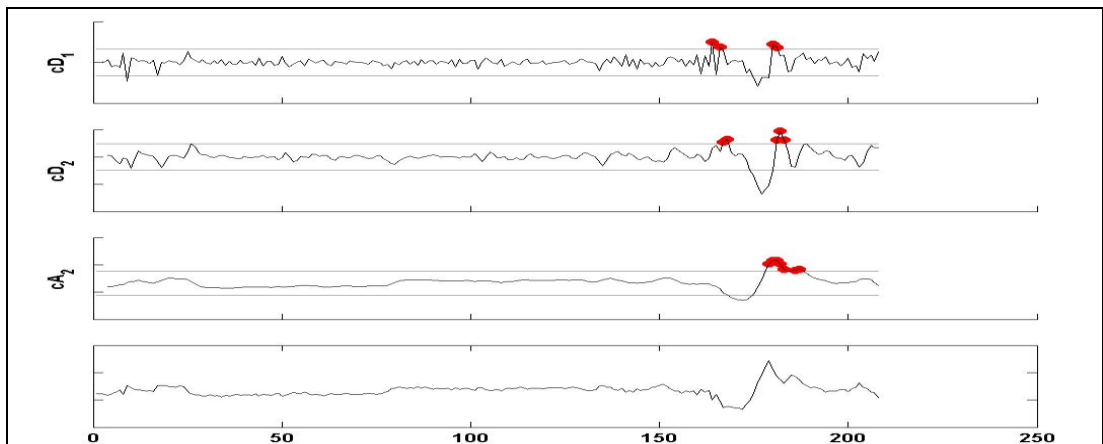
Shumeli (2013); herhangi bir gündeki verinin, ilgili günün içinde bulunduğu 30 günlük bir zaman dilimindeki verilerden etkileneceği düşüncesiyle ( $2^5 = 32$ ) Haar dönüşümünü 5 düzeyli olarak uygulamıştır. Buna paralel olarak herhangi bir haftanın verisinin, ilgili haftanın içinde bulunduğu 4 haftalık zaman dilimindeki verilerden etkileneceği düşüncesiyle bu tez çalışmasında ( $2^2 = 4$ ) 2 düzeyli Haar ayrışımının uygun olacağı düşünülmüştür.

İlk olarak referans dönemi verilerine Alt Bölüm 5.2’de yapıldığı gibi aykırı değer analizi (Çizelge 5.7.) gerçekleştirilerek, geçmiş olası salgınların veriler üzerindeki etkisi azaltılmaya çalışılmıştır. Shumeli (2013), günlük verilere wavelet yöntemini uygulamadan önce haftanın günü etkisi ile oluşan mevsimsellik yapısını hareketli ortalamalara oranlama yöntemi (ratio to moving average methods) ile gidermeye çalışmıştır. Bu nedenle aykırı değerlerden arındırılmış haftalık CDC referans dönemi verilerindeki çarpımsal mevsimsel etkiyi gidermek amacıyla, hareketli ortalamalara oranlama yöntemi kullanılmıştır. Bu işlem için ilk önce ilgili verilerin 52 genişliğindeki hareketli ortalamaları bulunmuştur, daha sonra ise bu hareketli ortalamaların da 52 genişliğindeki merkezi hareketli ortalamaları alınmıştır. Aykırı değerlerden arındırılmış veriler, bunlara karşılık gelen hareketli ortalamalara bölünüp 100 ile çarpılmıştır. Elde edilen katsayılar haftalık ve yıllık olacak şekilde  $52 \times 3$  boyutta bir matrise dönüştürülmüş ve haftalık ortalamalar bulunmuştur. Bu ortalamaların toplamları 5200 olacak şekilde düzenlenerek 52 haftalık mevsimsel indeksleri elde edilmiştir (Çizelge Ek I.1.). Ayrıca her hafta için elde edilen mevsimsellik katsayıları kullanılarak, 2015-2016 grip sezonundaki haftalık izleme verilerine de mevsimsellik düzeltmesi yapılmıştır (Çizelge Ek I.2.). Mevsimsellik düzeltmesi yapılan verilerin zaman serisi grafiği Şekil 5.31.’de verilmektedir.

Mevsimsellik etkisi giderilmiş olan veriler öz ilişkilidir. Ancak wavelet yöntemi verilerin durağan ve bağımsız olmasını gerektirmemektedir (Shumeli, 2013). Özilişki nedeniyle kaynaklanabilecek yanlış alarmları azaltmak amacıyla, oluşturulacak Shewhart grafiklerinde  $4\sigma$  sınır değerleri kullanılmıştır. Bu nedenle mevsimselliği giderilmiş CDC verilerine, wavelet analizi Shumeli (2013) tarafından sağlanan program aracılığıyla uygulanmıştır (<http://www.galitshmueli.com/content/matlab-code-biosurveillance>). Bu amaçla, günlük verilere yönelik MATLAB kodlarıyla yazılan ve 5 düzeyli Haar-wavelet ayrışımı yapan bu program düzenlenerek, CDC'nin haftalık grip verilerine 2 düzeyli Haar-wavelet dönüşümü gerçekleştirilecek şekilde uyarlanmıştır. Elde edilen sonuçlar Şekil 5.32., Çizelge 5.36. ve Çizelge 5.37.'de verilmektedir.



Şekil 5.31. Mevsimsel olarak düzeltilmiş CDC grip verilerin zaman serisi grafiği



Şekil 5.32. CDC grip verilerinin 2 düzeyli Haar-wavelet dönüşümü grafikleri



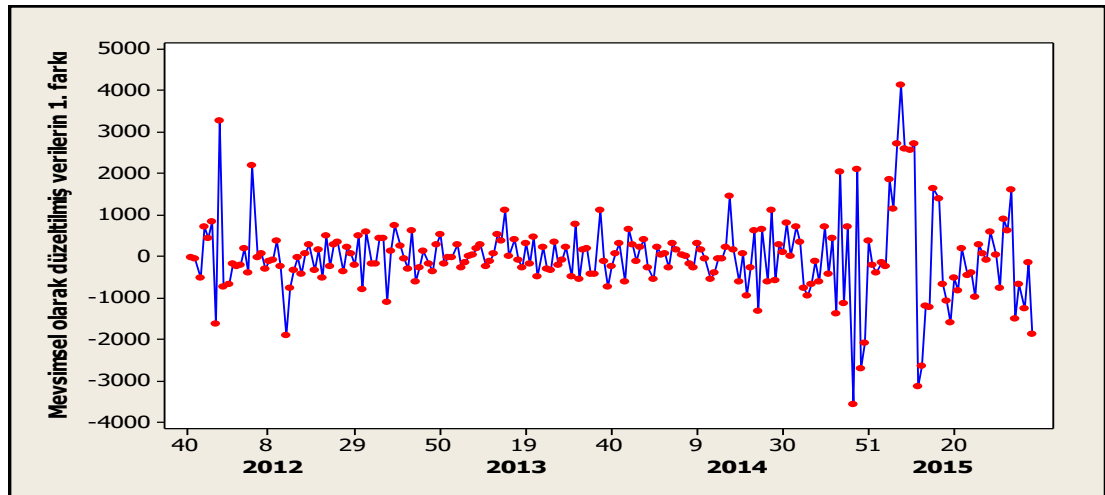
**Çizelge 5.36. Wavelet yöntemine ilişkin 2015-2016 grip mevsimi için haftalık grip faaliyet durumu**

Hafta	Karar	Hafta	Karar	Hafta	Karar	Hafta	Karar
40	Normal	1	Normal	14	Olağandışı	27	Normal
41	Normal	2	Normal	15	Normal	28	Normal
42	Normal	3	Normal	16	Normal	29	Normal
43	Normal	4	Normal	17	Olağandışı	30	Normal
44	Normal	5	Normal	18	Olağandışı	31	Normal
45	Normal	6	Normal	19	Normal	32	Normal
46	Normal	7	Normal	20	Normal	33	Normal
47	Olağandışı	8	Normal	21	Normal	34	Normal
48	Normal	9	Normal	22	Normal	35	Normal
49	Olağandışı	10	Olağandışı	23	Normal	36	Normal
50	Olağandışı	11	Olağandışı	24	Normal	37	Normal
51	Olağandışı	12	Olağandışı	25	Normal	38	Normal
52	Normal	13	Olağandışı	26	Normal	39	Normal

**Çizelge 5.37. Haar-wavelets yönteminin verdiği kararlar için sınıflandırma tablosu**

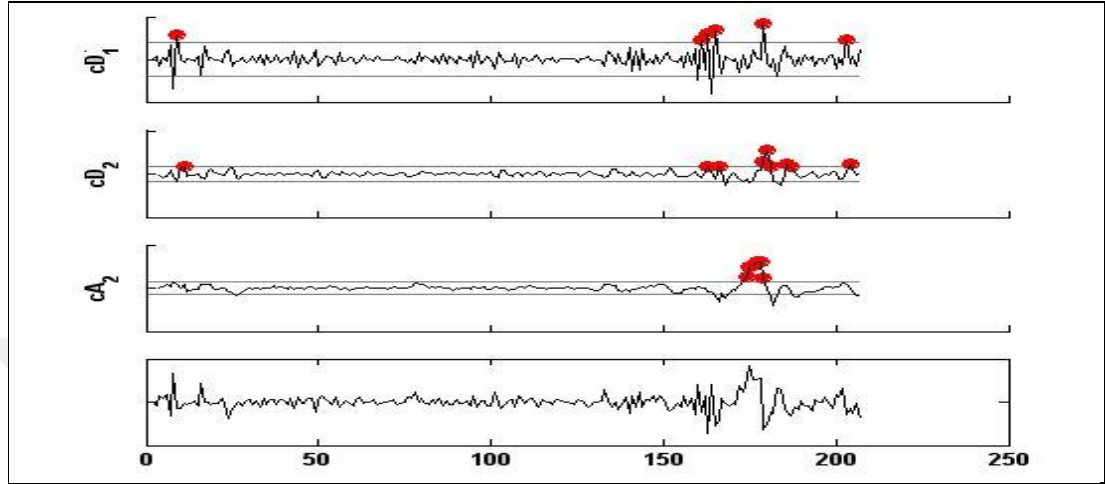
Yöntem	Karar	Gerçek Durum(CDC)		Toplam	Duyarlılık	Seçicilik
		Normal	Olağandışı			
Haar-wavelets	Normal	TN=33	FN=8	41	0.3846	0.8462
	Olağandışı	FP=6	TP=5	11		
	Toplam	39	13	52		

Mevsimsellik etkisi azaltılmasına rağmen veriler (Şekil 5.31.) öz ilişkilidir. Bu öz ilişkinin algoritmanın işleyişine olabilecek olumsuz etkisini azaltmak amacıyla, mevsimsellik düzeltmesi yapılan verilerin 1. farkları alınmıştır (Şekil 5.33.).



**Şekil 5.33. Mevsimsel olarak düzeltilmiş CDC grip verilerine ilişkin 1.farkların zaman serisi grafiği**

Shumeli (2013) tarafından verilen program aracılığıyla, farkı alınmış bu verilere 2 düzeyli Haar-wavelet dönüşümünü kullanan wavelet analizi uygulanmıştır. Kontrol sınırları Shewhart kalite kontrol grafiklerinde  $3\sigma$  olarak verilerek elde edilen sonuçlar Şekil 5.34., Çizelge 5.38. Çizelge 5.39.'da gösterilmektedir.



Şekil 5.34. CDC grip verilerinin 2 düzeyli Haar-wavelet dönüşümü grafikleri

Çizelge 5.38. Wavelet yöntemine ilişkin 2015-2016 grip mevsimi için haftalık grip faaliyet durumu

Hafta	Karar	Hafta	Karar	Hafta	Karar	Hafta	Karar
40	Normal	1	Normal	14	Normal	27	Normal
41	Normal	2	Normal	15	Normal	28	Normal
42	Normal	3	Normal	16	Normal	29	Normal
43	Normal	4	Normal	17	Olağandışı	30	Normal
44	Olağandışı	5	Olağandışı	18	Olağandışı	31	Normal
45	Normal	6	Olağandışı	19	Normal	32	Normal
46	Olağandışı	7	Olağandışı	20	Normal	33	Normal
47	Normal	8	Olağandışı	21	Normal	34	Normal
48	Olağandışı	9	Olağandışı	22	Normal	35	Olağandışı
49	Olağandışı	10	Olağandışı	23	Normal	36	Olağandışı
50	Normal	11	Olağandışı	24	Normal	37	Normal
51	Normal	12	Olağandışı	25	Normal	38	Normal
52	Normal	13	Normal	26	Normal	39	Normal

Çizelge 5.39. Haar-wavelets yönteminin verdiği kararlar için sınıflandırma tablosu

Yöntem	Karar	Gerçek Durum(CDC)		Toplam	Duyarlılık	Seçicilik
		Normal	Olağandışı			
Haar-wavelets	Normal	TN=31	FN=5	36	0.6154	0.7949
	Olağandışı	FP=8	TP=8	16		
	Toplam	39	13	52		

## 5.6 Önerilen UCengiz Yönteminin CDC Grip Verilerine Uygulanması

Bu alt bölümde uzun referans dönemi gerektirmeyen ve tez kapsamında önerilen UCengiz yöntemi, CDC kurumundan elde edilen 2015-2016 grip mevsimi 52 haftalık verilerini izlemek amacıyla uygulanmaktadır. Bu yöntem 4. Bölüm’de tanımlandığı gibi 2 parametreye bağlıdır. Bunlardan ilki dağılımın üst yüzdeliğini belirleyen  $\alpha$  değeri iken, ikincisi referans döneminin genişliğini gösteren gözlem sayısıdır ( $k$ ).

Uygulamada hem orijinal veriler hemde işlenmiş veriler için; UCengiz yöntemine ilişkin 4 ve 8 gözlemlik referans dönemleri dikkate alındığında, ATFS ölçütünü 13 ve 26 değerlerine ulaştıran  $\alpha$  değerleri 10000 tekrarlı simülasyon çalışması sonucu belirlenmiştir (Çizelge 5.40.). Verilerin yeniden ölçeklenerek Serfling regresyon yöntemi sonucu elde edilen artıkların 1. farklarının pozitif hali, işlenmiş verileri ifade etmektedir.

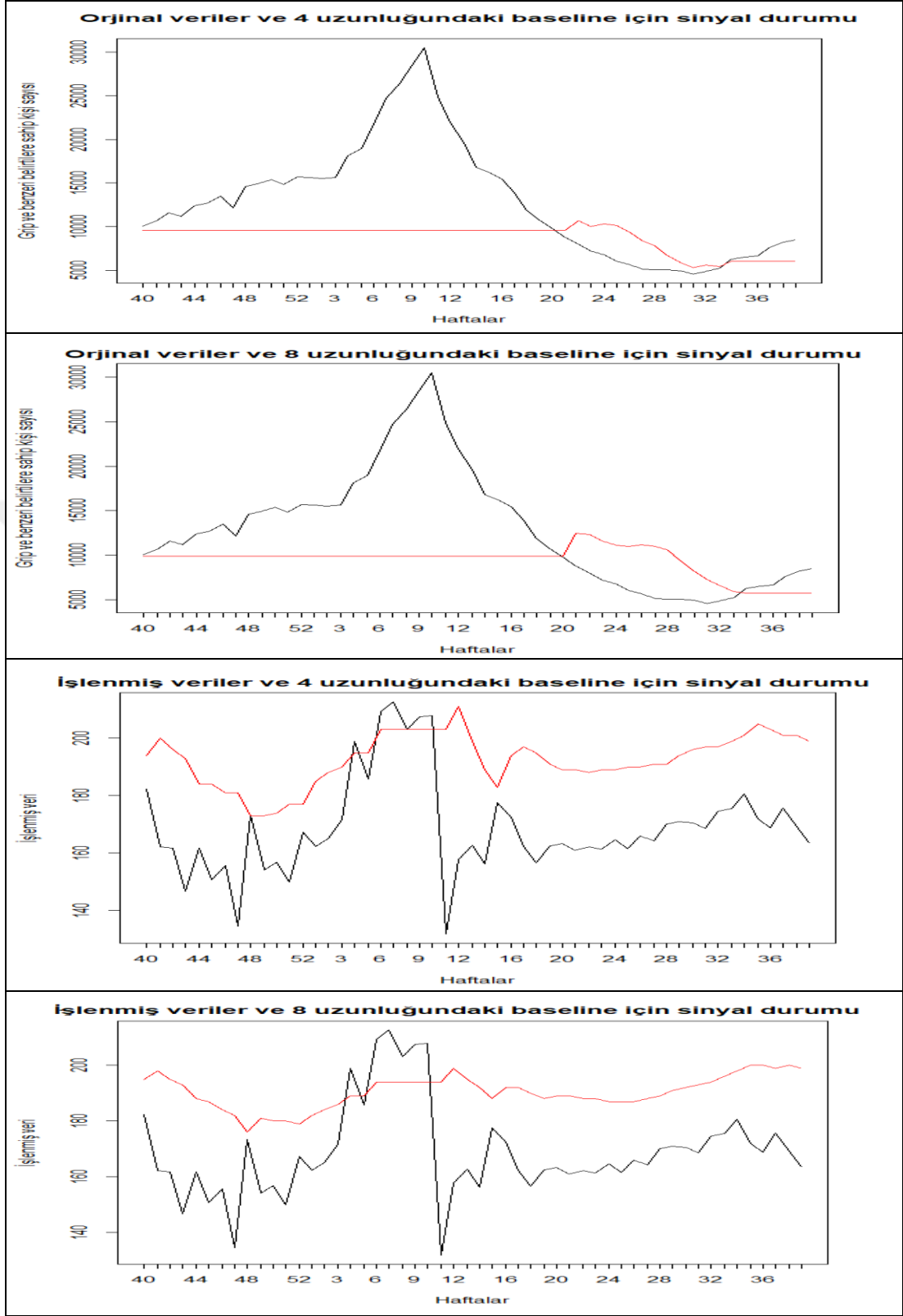
**Çizelge 5. 40. 2015-2016 grip mevsimi verileri için UCengiz yöntemine ilişkin istenen ATFS değerlerini veren parametreler**

	Referans verilerinin sayısı (baseline)	ATFS (Yanlış sinyaller arasındaki ortalama süre)	
		13	26
Orijinal veri	$k = 4$	0.00006	0.0000002
	$k = 8$	0.0014	0.00012
İşlenmiş veri	$k = 4$	0.0190	0.0034
	$k = 8$	0.0260	0.0050

Bu parametrelerle elde edilen izleme sonuçları ve CDC kurumunun sonuçlarıyla karşılaştırılmaları aşağıdaki alt bölümlerde verilmektedir.

### 5.6.1 ATFS 13 için elde edilen izleme sonuçları

R kodlarıyla yazılan bir program aracılığıyla UCengiz yöntemi, ATFS 13 değerini veren parametreler ile çalıştırılmıştır ve elde edilen sonuçlar Şekil 5.35. ile Çizelge 5.41.’de verilmektedir. Elde edilen sonuçların CDC kurumu ile karşılaştırılması ise Çizelge 5.42.’de yer almaktadır. Ayrıca yapılan uygulamalarda UCengiz yöntemiyle elde edilen eşik değerleri Ek M.’deki çizelgelerde verilmektedir.



Şekil 5. 35. 2015-2016 grip mevsimi verileri için UCengiz yöntemine ilişkin haftalık sinyal grafikleri

**Çizelge 5. 41. UCengiz yöntemine göre ATFS 13 değeri için 2015-2016 grip mevsimine ilişkin haftalık grip faaliyet durumu**

Hafta	Karar	Hafta	Karar	Hafta	Karar	Hafta	Karar
40	Odiş(i04,o8)	1	Odiş(i04,o8)	14	Odiş(i04,o8)	27	Normal
41	Odiş(i04,o8)	2	Odiş(i04,o8)	15	Odiş(i04,o8)	28	Normal
42	Odiş(i04,o8)	3	Odiş(i04,o8)	16	Odiş(i04,o8)	29	Normal
43	Odiş(i04,o8)	4	Odiş(i04,o8,i4,i8)	17	Odiş(i04,o8)	30	Normal
44	Odiş(i04,o8)	5	Odiş(i04,o8)	18	Odiş(i04,o8)	31	Normal
45	Odiş(i04,o8)	6	Odiş(i04,o8,i4,i8)	19	Odiş(i04,o8)	32	Normal
46	Odiş(i04,o8)	7	Odiş(i04,o8,i4,i8)	20	Odiş(i04)	33	Normal
47	Odiş(i04,o8)	8	Odiş(i04,o8,i4,i8)	21	Normal	34	Odiş(i04,o8)
48	Odiş(i04,o8,i4)	9	Odiş(i04,o8,i4,i8)	22	Normal	35	Odiş(i04,o8)
49	Odiş(i04,o8)	10	Odiş(i04,o8,i4,i8)	23	Normal	36	Odiş(i04,o8)
50	Odiş(i04,o8)	11	Odiş(i04,o8)	24	Normal	37	Odiş(i04,o8)
51	Odiş(i04,o8)	12	Odiş(i04,o8)	25	Normal	38	Odiş(i04,o8)
52	Odiş(i04,o8)	13	Odiş(i04,o8)	26	Normal	39	Odiş(i04,o8)

o4:orijinal veri k=4; o8:orijinal veri k=8; i4:işlenmiş veri k=4; i8:işlenmiş veri k=8

**Çizelge 5. 42. UCengiz yönteminin ATFS 13 değeri için 2015-2016 grip mevsimine ilişkin verdiği kararların sınıflandırma tablosu**

Yöntem	Karar	Gerçek Durum(CDC)		Toplam	Duyarlılık	Seçicilik
		Normal	Olağandışı			
o4	Normal	TN=13	FN=0	13	1.0000	0.3333
	Olağandışı	FP=26	TP=13	39		
	Toplam	39	13	52		
o8	Normal	TN=14	FN=0	14	1.0000	0.3590
	Olağandışı	FP=25	TP=13	38		
	Toplam	39	13	52		
i4	Normal	TN=38	FN=7	45	0.4615	0.9744
	Olağandışı	FP=1	TP=6	7		
	Toplam	39	13	52		
i8	Normal	TN=39	FN=7	46	0.4615	1.0000
	Olağandışı	FP=0	TP=6	6		
	Toplam	39	13	52		

Şekil 5.35. ile Çizelge 5.41. ve Çizelge 5.42. incelendiğinde; UCengiz yönteminin orijinal verilere uygulamalarında yüksek duyarlılığın ön plana çıktığı görülürken, işlenmiş verilere uygulamalarında ise yüksek seçicilik dikkat çekmektedir. Bu durum 13 ATFS değeriyle 52 haftalık karar verme işlemlerinde; ortalama 4 yanlış sinyal beklentisinin orijinal veriler için çok aşıldığını, işlenmiş veriler için ise bu beklentinin çok gerisinde kaldığını göstermektedir.

## 5.6.2 ATFS 26 için elde edilen izleme sonuçları

R kodlarıyla yazılan bir program aracılığıyla UCengiz yöntemi, ATFS 26 değerini veren parametreler ile çalıştırılmıştır ve elde edilen sonuçlar Şekil 5.36. ile Çizelge 5.43.'de verilmektedir. Elde edilen sonuçların CDC kurumu ile karşılaştırılması ise Çizelge 5.44.'de yer almaktadır. Ayrıca yapılan uygulamalarda UCengiz yöntemiyle elde edilen eşik değerleri Ek M.'deki çizelgelerde verilmektedir.

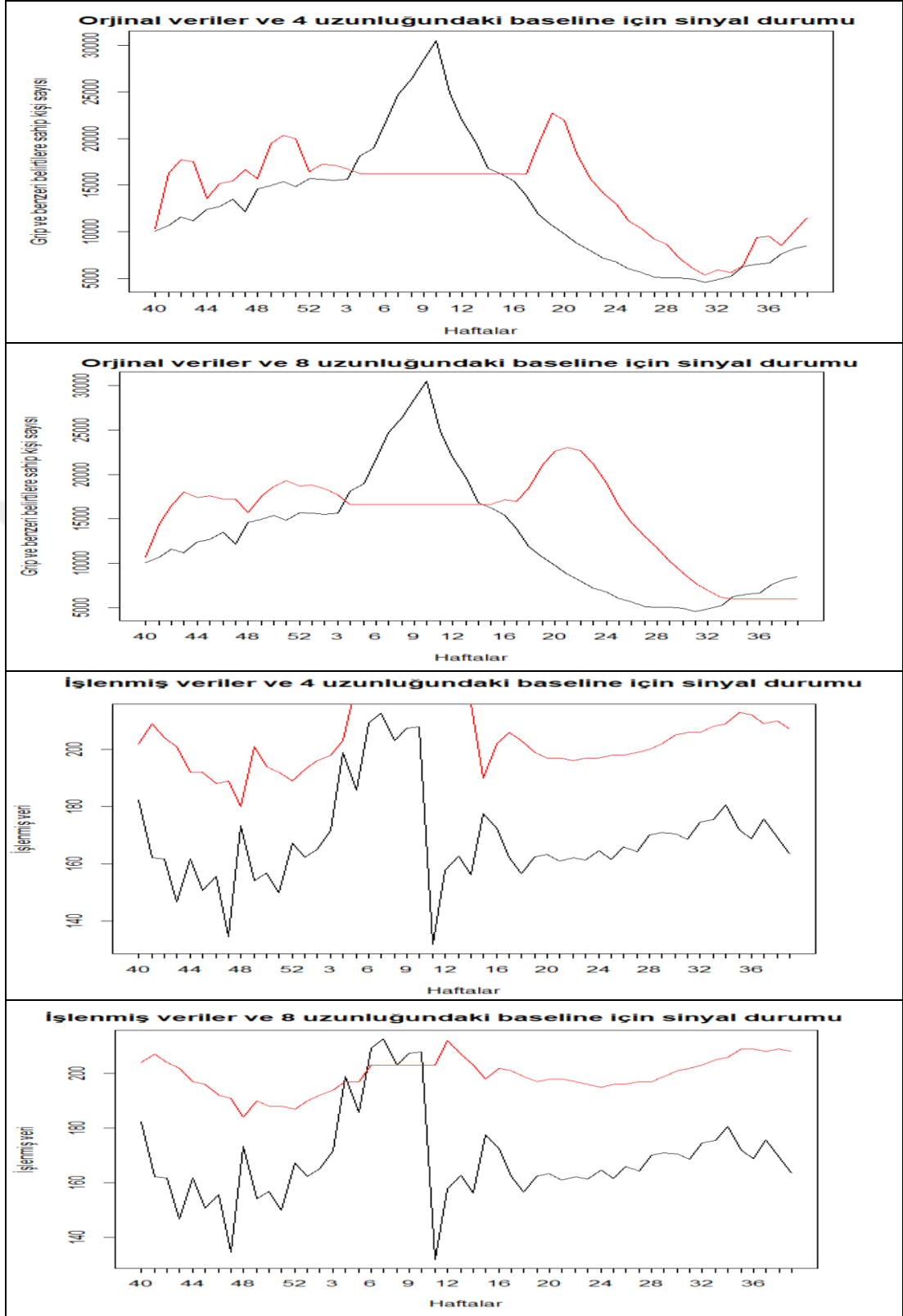
**Çizelge 5. 43. UCengiz yöntemine göre ATFS 26 değeri için 2015-2016 grip mevsimine ilişkin haftalık grip faaliyet durumu**

Hafta	Karar	Hafta	Karar	Hafta	Karar	Hafta	Karar
40	Normal	1	Normal	14	Odişli(o4,o8)	27	Normal
41	Normal	2	Normal	15	Odişli(o4)	28	Normal
42	Normal	3	Normal	16	Normal	29	Normal
43	Normal	4	Odişli(o4,o8,i8)	17	Normal	30	Normal
44	Normal	5	Odişli(o4,o8)	18	Normal	31	Normal
45	Normal	6	Odişli(o4,o8,i8)	19	Normal	32	Normal
46	Normal	7	Odişli(o4,o8,i8)	20	Normal	33	Normal
47	Normal	8	Odişli(o4,o8,i8)	21	Normal	34	Odişli(o8)
48	Normal	9	Odişli(o4,o8,i8)	22	Normal	35	Odişli(o8)
49	Normal	10	Odişli(o4,o8,i8)	23	Normal	36	Odişli(o8)
50	Normal	11	Odişli(o4,o8)	24	Normal	37	Odişli(o8)
51	Normal	12	Odişli(o4,o8)	25	Normal	38	Odişli(o8)
52	Normal	13	Odişli(o4,o8)	26	Normal	39	Odişli(o8)

o4:orijinal veri k=4; o8:orijinal veri k=8; i4:işlenmiş veri k=4; i8:işlenmiş veri k=8

**Çizelge 5. 44. UCengiz yönteminin ATFS 26 değeri için 2015-2016 grip mevsimine ilişkin verdiği kararların sınıflandırma tablosu**

Yöntem	Karar	Gerçek Durum(CDC)		Toplam	Duyarlılık	Seçicilik
		Normal	Olağandışı			
o4	Normal	TN=37	FN=3	40	0.7692	0.9487
	Olağandışı	FP=2	TP=10	12		
	Toplam	39	13	52		
o8	Normal	TN=32	FN=3	35	0.7692	0.8205
	Olağandışı	FP=7	TP=10	17		
	Toplam	39	13	52		
i4	Normal	TN=39	FN=13	52	0	1.00
	Olağandışı	FP=0	TP=0	0		
	Toplam	39	13	52		
i8	Normal	TN=39	FN=7	46	0.4615	1.00
	Olağandışı	FP=0	TP=6	6		
	Toplam	39	13	52		



Şekil 5. 36. 2015-2016 grip mevsimi verileri için UCengiz yöntemine ilişkin haftalık sinyal grafikleri

Şekil 5.36. ile Çizelge 5.43. ve Çizelge 5.44. incelendiğinde; UCengiz yönteminin orijinal verilere uygulamalarında hem seçiciliğin hemde duyarlılığın birlikte iyi olması dikkat çekmektedir. Ancak bu yöntemin işlenmiş verilere uygulamalarında ise seçicilik çok yüksek iken duyarlılığın düşük oluşu görülmektedir.





## 6. SİMÜLASYON ÇALIŞMASI

Bu bölümde bu tez çalışmasında önerilen salgın belirleme algoritması da dahil olmak üzere bazı salgın belirleme algoritmalarının performansları, simülasyon ile üretilecek haftalık grip verileri kullanılarak karşılaştırılacaktır. Simülasyonla üretilecek verilerin mevsimsellik, özilişkili vb. gibi gerçek grip verilerinin özelliklerini taşıyabilmesi amacıyla, CDC kurumunun internet sayfasındaki 2010-2017 grip mevsimleri ilişkin veriler dikkate alınarak simülasyon gerçekleştirilmiştir. Böylelikle üretilen grip verileri, gerçek grip verilerinin temel özelliklerini yaklaşık temsil edebilecek biçimde üretilmişlerdir.

### 6.1 Grip Verilerin Üretimi

Haftalık grip verileri Çizelge 6.1.'de ifade edilen 4 farklı senaryoya göre, grip verilerinin özelliklerini yansıtacak şekilde Poisson dağılımı veya negatif binom dağılımı kullanılarak salgın etkisi olmaksızın üretilmektedir. Veri üretiminde; zamana bağlı olarak değişen olasılık dağılımlarının dizisine ilişkin parametre kestirimlerinden yararlanan Fricker (2012) yarı sentetik yöntemi kullanılmaktadır. Bütün bu işlemler R programı kodlarıyla gerçekleştirilmektedir.

Çizelge 6.1.'de belirtilen her bir senaryo için, ilk olarak R programı kullanılarak aykırı değer analizi yapılmaktadır. Böylelikle verilerdeki olası salgınlardan kaynaklı etkiler azaltılmaya çalışılmaktadır. Daha sonra verilerin aşırı yayımlı olup olmamasına bağlı olarak Poisson dağılımı ( $\lambda$ ) ya da negatif binom ( $r, p$ ) dağılımıyla veriler üretilmektedir.

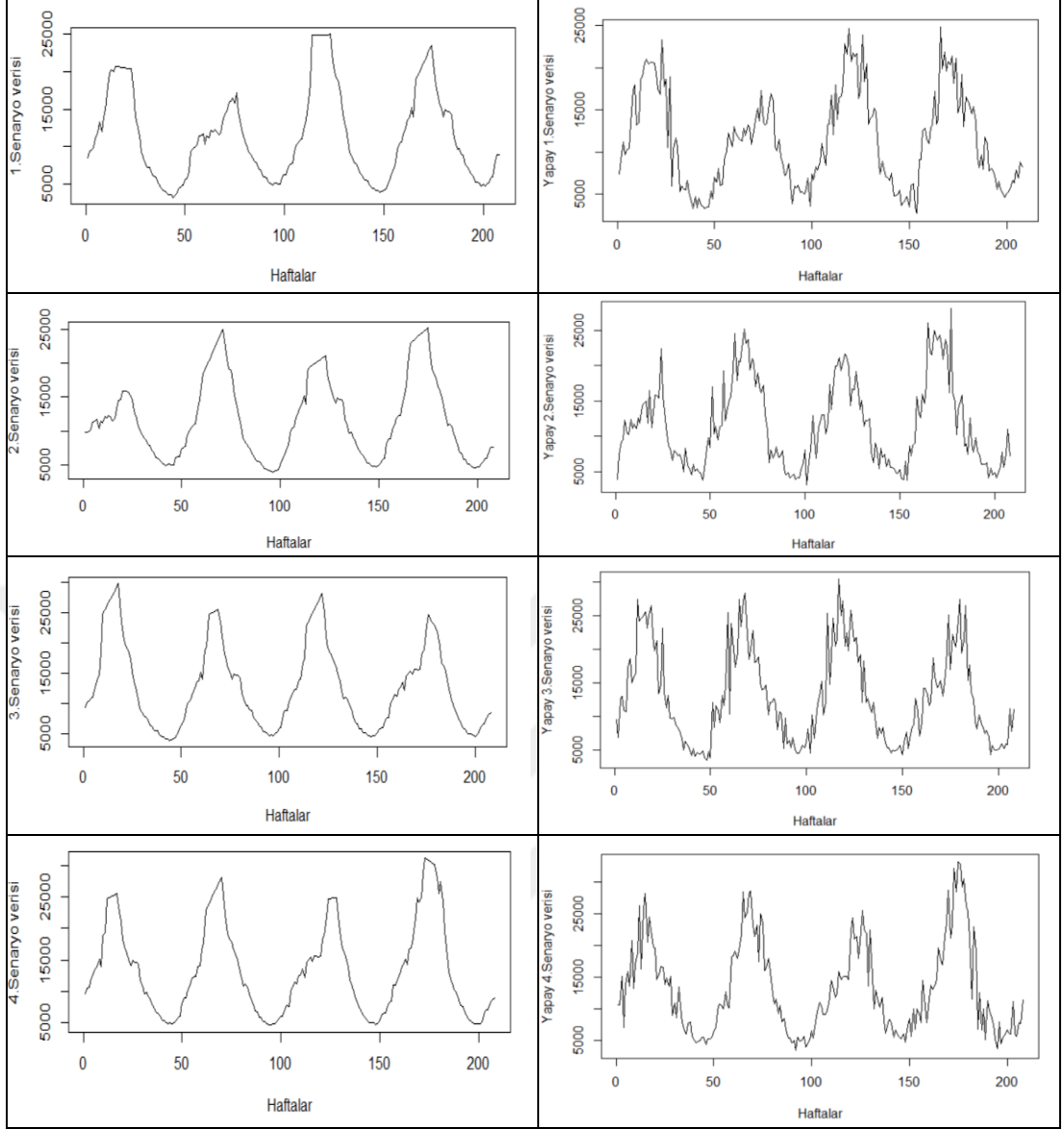
**Çizelge 6.1. Simülasyon çalışmasında kullanılan veri üretim senaryoları**

Senaryo	Referans dönemleri	İzleme dönemi
1	2010-2011 2011-2012 2012-2013	2013-2014
2	2011-2012 2012-2013 2013-2014	2014-2015
3	2012-2013 2013-2014 2014-2015	2015-2016
4	2013-2014 2014-2015 2015-2016	2016-2017

Ancak grip verilerinin durağan olmayan yapısını dikkate almak amacıyla, bu dağılımların haftalık parametre kestirimleri için 9 genişliğinde hareketli ortalamalar ( $\bar{x}_t$ ) ve hareketli varyanslar ( $s_t^2$ ) oluşturulmaktadır. Oluşturulan her hafta için hareketli varyans değeri ile hareketli ortalama değeri karşılaştırılmaktadır:

- $s_t^2 > \bar{x}_t$  ise negatif binom( $\hat{r}_t, \hat{p}_t$ ),  $\hat{r}_t = \bar{x}_t^2 / (s_t^2 - \bar{x}_t)$  ve  $\hat{p}_t = \bar{x}_t / s_t^2$
- $s_t^2 \leq \bar{x}_t$  ise Poisson( $\hat{\lambda}_t$ ),  $\hat{\lambda}_t = \bar{x}_t$

Yukarıdaki parametre kestirimi yöntemi kullanılarak ilgili hafta için veriler üretilmektedir. Bu yöntemle 4 senaryoya göre üretilmiş grip verilerine ilişkin zaman serilerinin birer örneği, Şekil 6.1.'de ilgili senaryolara ilişkin gerçek verilerin zaman serileri grafikleriyle birlikte gösterilmektedir. Şekil 6.1.'deki zaman serileri grafiklerinden simülasyon ile üretilen verilerin gerçek grip verilerine benzedikleri söylenebilir. Bu şekilde üretilen verilerin izleme dönemine daha sonra algoritmaların performans karşılaştırılması için yapay salgınlar eklenecektir.



**Şekil 6.1. Senaryolara göre gerçek ve yapay olarak üretilmiş grip verilerinin zaman serisi grafikleri**

## 6.2 Karşılaştırma Ölçütleri

Salgın belirleme algoritmalarına ilişkin performans değerlendirmeleri için Fricker vd. (2008) çalışmasındaki ölçütler dikkate alınmıştır. Algoritmalar ilk olarak referans dönemlerinde aynı ATFS (average time between false signals) performansına sahip

olacak biçimde parametreleri ayarlanmıştır. ATFS ölçütünün hesaplanması aşağıda verilmektedir.

$$ATFS = \frac{\sum_{i=1}^r t_i}{r}$$

Burada  $r$ , simülasyondaki tekrar sayısını ve  $t_i$ ,  $i$ . tekrarda ilk sinyale kadar olan süreyi göstermektedir. Fricker (2012) ATFS değerinin ilgili kurumun kaynakları dahilinde olabildiğince en küçük değer olarak seçilmesi gerektiğini belirtmektedir. Böylelikle ilgili algoritmanın salgınları belirlemeye karşı daha duyarlı olacağı ifade edilmektedir. Bu simülasyon çalışmasında ATFS değeri 13 olarak belirlenmiştir. Böyle bir belirleme, izleme yapan kurumun 52 haftalık zaman diliminde ortalama 4 yanlış sinyali araştırmak için yeterli kaynaklara sahip olduğunun varsayılması anlamına gelir. Daha sonra  $ATFS = 13$  değerini elde edebilmek amacıyla, çeşitli algoritmalar için gerekli parametre değerleri 10000 tekrarlı simülasyon çalışması aracılığıyla belirlenmiştir (Çizelge 6.2.).

**Çizelge 6.2. Dikkate alınan algoritmalar için yaklaşık ATFS=13 veren parametre değerleri (10000 tekrar)**

Algoritmalar	Senaryo1(2014)	Senaryo2(2015)	Senaryo3(2016)	Senaryo4(2017)
Bayes	$wz = 2$	$wz = 2$	$wz = 2$	$wz = 2$
	$b1 = 2$	$b1 = 2$	$b1 = 2$	$b1 = 2$
	$\alpha = 0.0025$	$\alpha = 0.00025$	$\alpha = 0.015$	$\alpha = 0.0005$
Cusum	$aa = 33$	$aa = 69$	$aa = 39$	$aa = 59$
	$ar = 13$	$ar = 17$	$ar = 3$	$ar = 15$
Farrington	$wz = 2$	$wz = 2$	$wz = 2$	$wz = 2$
	$b1 = 2$	$b1 = 2$	$b1 = 2$	$b1 = 2$
	$\alpha = 0.05$	$\alpha = 0.015$	$\alpha = 0.275$	$\alpha = 0.02$
F. Flexible	$wz = 2$	$wz = 2$	$wz = 2$	$wz = 2$
	$b1 = 2$	$b1 = 2$	$b1 = 2$	$b1 = 2$
	$\alpha = 0.12$	$\alpha = 0.045$	$\alpha = 0.35$	$\alpha = 0.05$
C1	$\alpha = 0.02$	$\alpha = 0.015$	$\alpha = 0.015$	$\alpha = 0.01$
C2	$\alpha = 0.018$	$\alpha = 0.01$	$\alpha = 0.01$	$\alpha = 0.008$
C3	$\alpha = 0.045$	$\alpha = 0.035$	$\alpha = 0.02$	$\alpha = 0.035$
OutbreakP	$alimit = 55$	$alimit = 600$	$alimit = 20$	$alimit = 700$
Glrb	$kd = 2$	$kd = 5.5$	$kd = 0.5$	$kd = 2.75$
Serfling	$\alpha = 0.08$	$\alpha = 0.005$	$\alpha = 0.35$	$\alpha = 0.02$
UCengiz	$\alpha = 0.00021$	$\alpha = 0.00009$	$\alpha = 0.00006$	$\alpha = 0.000039$
	$k = 4$	$k = 4$	$k = 4$	$k = 4$

Çizelge 6.2.'de belirtilen bazı yöntemlerin performansları birden fazla parametreye bağlıdır. Bu parametrelerden bazıları bu simülasyon çalışması için sabit bir değer almak zorunda kalmıştır. Örneğin referans dönemi 3 yıl olarak dikkate alındığından bayes, Farrington, FarringtonFlexible algoritmaları için  $b$  (geçmişe dönük dikkate alınacak dönem sayısı) parametresi 2 olarak alınmıştır.

Salgın olmayan dönemi belirten referans dönemleri için algoritmaların ATFS performansları eşitlendikten sonra, izleme dönemine gerçekçi salgınlar eklenip ATFOS (the average time to first outbreak signal) ve FMO (fraction missed outbreak) karşılaştırma ölçütlerine göre algoritmaların performansları incelenmiştir. ATFOS ölçütü aşağıdaki gibi tanımlanmaktadır (Fricker vd., 2008).

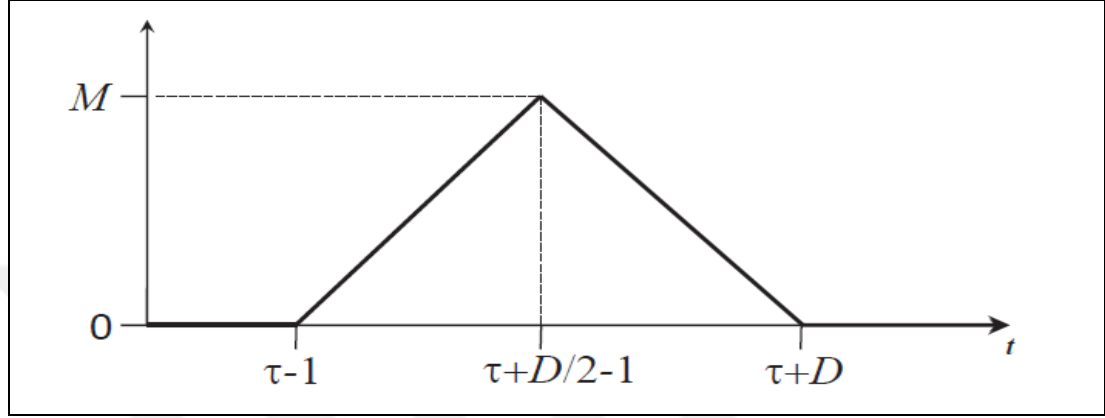
$$ATFOS = \frac{\sum_{i=1}^s (t_i - o)}{s}$$

Burada  $s$ ,  $r$  tekrarda salgın sinyali verilen tekrar sayısı;  $t_i$ ,  $i$ . tekrarda salgın olan dönemdeki ilk sinyale kadar olan süreyi ve  $o$ , salgının başlangıcına kadar geçen süreyi göstermektedir. Salgını yakalayamama oranı ise  $FMO = \frac{r-s}{r} \times 100$  olarak ifade edilir (Fricker vd., 2008).

ATFS eşitlemelerinde ve performans simülasyon çalışmalarında Serfling, Glrb, Farrington, FarringtonFlexible yöntemleri ve bu tez çalışmasında önerilen UCengiz yöntemi ham verileri kullanırken; Bayes, Cusum, Ears C1, Ears C2, Ears C3 ve OutbreakP yöntemleri ise hamverilerin işlenmiş haline kullanmaktadır. Grip verilerinin mevsimsellik etkisi gibi özelliklerini dikkate alan yöntemlerin ham verilere doğrudan; grip verilerindeki mevsimsellik etkisi ve özilişki gibi özelliklerden olumsuz etkilenebilecek yöntemlerin ise, bu etkilerin kısmen giderildiği işlenmiş verilere uygulanmasının daha uygun olduğu düşünülmüştür. Verilerin ön işlenmesi ise; verilerin yeniden ölçeklenmesi, mevsimsellik etkisini azaltmak amacıyla Serfling regresyon artıklarının elde edilmesi, ilgili artıkların 1. farklarının alınması ve pozitif hale getirilmesi işlemlerini içermektedir.

### 6.3 İzleme Dönemine Salgınların Eklenmesi

Bu simülasyon çalışmasındaki 4 senaryo ve CDC kurumunun ilgili senaryolar için belirlediği salgın haftaları incelendiğinde, salgın yapısının Şekil 6.2.'deki yapıyla yaklaşık olarak uyumlu olabileceği söylenebilir.



Şekil 6.2. Salgın yapısı

Şekil 6.2.'de  $\tau$  salgının başlangıç noktasını,  $M$  salgının eriştiği en büyük değeri ve  $D$  ise salgının süresini göstermektedir. Bu salgın yapısının matematiksel ifadesi aşağıda verilmektedir (Fricker, 2013).

$$o_t = \begin{cases} \frac{2M(t - \tau + 1)}{D + 1}, & \tau - 1 < t \leq \tau + \frac{D}{2} - 1 \\ \frac{2M(D + \tau - t)}{D + 1}, & \tau + \frac{D}{2} - 1 < t < \tau + D \\ 0, & \text{diğer durumlar} \end{cases}$$

Burada  $o_t$ ,  $t$  zamanında Şekil 6.2.'deki ideal salgın yapısına ilişkin ortalama büyüklüktür. İlgili zaman noktalarındaki ortalamaya sahip Poisson dağılımından veriler üretilmekte ve aynı zaman noktalarındaki referans dönemi verilerine eklenerek salgınlar oluşturulmaktadır. Bu simülasyon çalışması için kullanılacak yukarıdaki salgın yapısının parametreleri olan  $\tau$ ,  $M$  ve  $D$  hakkında fikir edinmek amacıyla CDC kurumuna ait 2010-2017 yılları arasındaki grip mevsimlerine ilişkin bilgiler Çizelge

6.3.'te verilmektedir. Çizelge 6.3. dikkate alındığında, yapılacak simülasyon çalışmasında salgın parametrelerinin aşağıdaki gibi belirlenmesi uygun görülmüştür.

- Salgının başlangıç haftası ( $\tau$ ), sıklıkla görülen 47, 47, 48, 49, 50, 51, 51, 52, 1, 2, 3, 4 haftalarından rastgele olarak seçilmektedir.
- Salgının süresi ( $D$ ), 4, 8 ve 12 hafta olarak alınmaktadır.
- Salgının alabileceği en büyük değer ( $M$ ), referans dönemi ortalamasının 0.10, 0.25, 0.50, 1 ve 1.5 katı olarak alınmaktadır.

**Çizelge 6.3. 2010-2017 grip mevsimlerine ilişkin salgın bilgileri**

Grip mevsimleri	CDC kurumunun belirlediği salgın haftaları	Salgının başlangıcı ( $\tau$ )	Salgının süresi ( $D$ )	Salgının büyüklüğü ( $M$ )
2010-2011	51,52,1,2,3,4,5,6,7,8,9,10,11	$\tau_1 = 51$ $\tau_2 = 1$	$D_1 = 3$ $D_2 = 10$	$M_1 = 0.20$ $M_2 = 0.71$
2011-2012	Salgın yok	-	-	-
2012-2013	47,49,50,51,52,1,2,3,4,5,6,7,8,9,10,11	$\tau = 47$	$D = 15$	$M = 0.79$
2013-2014	48,49,50,51,52,1,2,3,4,5,6,7,8,9	$\tau = 48$	$D = 13$	$M = 0.65$
2014-2015	47,48,49,50,51,52,1,2,3,4,5,6,7,8,9,10,11,12	$\tau = 47$	$D = 18$	$M = 0.96$
2015-2016	51,52,3,4,5,6,7,8,9,10,11,12,13	$\tau = 3$	$D = 10$	$M = 0.42$
2016-2017	50,51,52,1,2,3,4,5,6,7,8,9,10,11,12,13,14	$\tau = 50$	$D = 16$	$M = 1.15$

#### 6.4 Simülasyon Planı

Önceki alt bölümlerde belirlenen parametreler ve performans ölçütlerine göre, 10000 tekrarlı simülasyonlar Çizelge 6.4.'teki koşullarda R kodlarıyla yazılan programlar ile gerçekleştirilmektedir. Çizelge 6.4.'teki plana göre toplam 15 performans çizelgesi oluşturulmaktadır.

Çizelge 6.4. Simülasyon parametreleri

Senaryolar	Simülasyonda kullanılan yöntemler	Referans dönemi ATFS	Salgın süresi (D)	Salgın büyüklüğü (M)	Karşılaştırma ölçütü
Senaryo1	Bayes Cusum Farrington F. Flexible C1	13	4 hafta	$0.10\mu_{Baseline}$	ATFOS FMO
Senaryo2	C2		8 hafta	$0.25\mu_{Baseline}$	
Senaryo3	C3		12 hafta	$0.50\mu_{Baseline}$	
Senaryo4	OutbreakP Glrnb Serfling UCengiz			$1\mu_{Baseline}$ $1.50\mu_{Baseline}$	

## 6.5 Simülasyon Sonuçları

Bu alt bölümde elde edilen performans sonuçları; salgın süresi ve büyüklüğüne bağlı olarak ayrı ayrı verilmektedir.

### 6.5.1 Dört hafta süreli salgınlar için simülasyon sonuçları

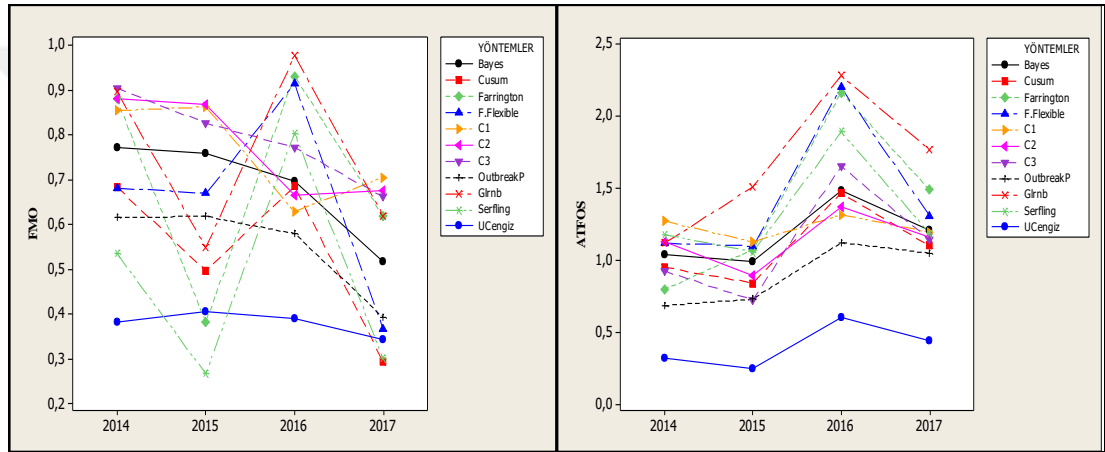
Salgın büyüklüğü, referans dönemi ortalamasının ( $\mu_{baseline}$ ) 0.10, 0.25, 0.50, 1.00 ve 1.50 katı olacak biçimde dört haftalık süreli salgınlar için 10000 tekrarlı simülasyonlarda elde edilen FMO ve ATFOS değerleri, Çizelge 6.5.-Çizelge 6.9. ve Şekil 6.3.-Şekil 6.7.'de verilmektedir. İlgili çizelgeler ve şekiller incelendiğinde değişik senaryolarda en iyi performansa sahip algoritmalar Çizelge 6.10.'da verilmektedir.

Çizelge 6.10. incelendiğinde; dört hafta süreli salgınlarda ve tüm salgın büyüklüklerinde ATFOS açısından bu çalışmada önerilen UCengiz yönteminin performansının üstünlüğü dikkat çekerken, FMO açısından ise sırasıyla en iyi performansı Serfling, Cusum, UCengiz, F. Flexible, Glrnb ve OutbreakP yöntemleri sergilemektedir. Ayrıca bu çalışmada önerilen UCengiz yönteminin küçük salgın büyüklüklerinde FMO açısından üstün olma eğilimi anlaşılmaktadır.



**Çizelge 6.5. 4 hafta süreli ve  $0.10\mu_{Baseline}$  büyüklüğündeki salgınlarda algoritmaların performansları (10000 tekrar)**

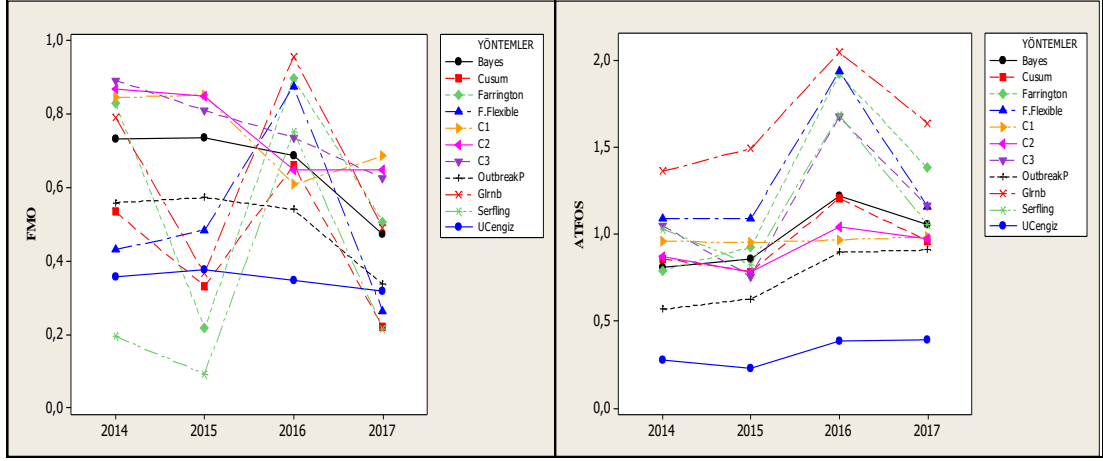
Algoritmalar	Senaryo1(2014)		Senaryo2(2015)		Senaryo3(2016)		Senaryo4(2017)	
	FMO	ATFOS	FMO	ATFOS	FMO	ATFOS	FMO	ATFOS
Bayes	0.7715	1.0411	0.7592	0.9880	0.6974	1.4805	0.5171	1.2133
Cusum	0.6838	0.9510	0.4975	0.8424	0.6857	1.4642	0.2946	1.1016
Farrington	0.8802	0.7980	0.3827	1.0753	0.9315	2.1650	0.6193	1.4886
F. Flexible	0.6818	1.1210	0.6704	1.0989	0.9143	2.1995	0.3667	1.3082
C1	0.8550	1.2703	0.8631	1.1278	0.6293	1.3154	0.7055	1.1946
C2	0.8822	1.1282	0.8689	0.8932	0.6655	1.3677	0.6771	1.1632
C3	0.9045	0.9246	0.8270	0.7266	0.7720	1.6553	0.6624	1.1431
OutbreakP	0.6161	0.6861	0.6190	0.7323	0.5800	1.1231	0.3947	1.0499
Glnb	0.8975	1.1171	0.5504	1.5098	0.9764	2.2797	0.6205	1.7679
Serfling	0.5350	1.1768	0.2683	1.0586	0.8046	1.8915	0.3017	1.1860
UCengiz	0.3821	0.3188	0.4063	0.2529	0.3897	0.6065	0.3432	0.4444



**Şekil 6.3. 4 hafta süreli ve  $0.10\mu_{Baseline}$  büyüklüğündeki salgınlarda algoritmaların performans grafikleri**

**Çizelge 6.6. 4 hafta süreli ve  $0.25\mu_{Baseline}$  büyüklüğündeki salgınlarda algoritmaların performansları (10000 tekrar)**

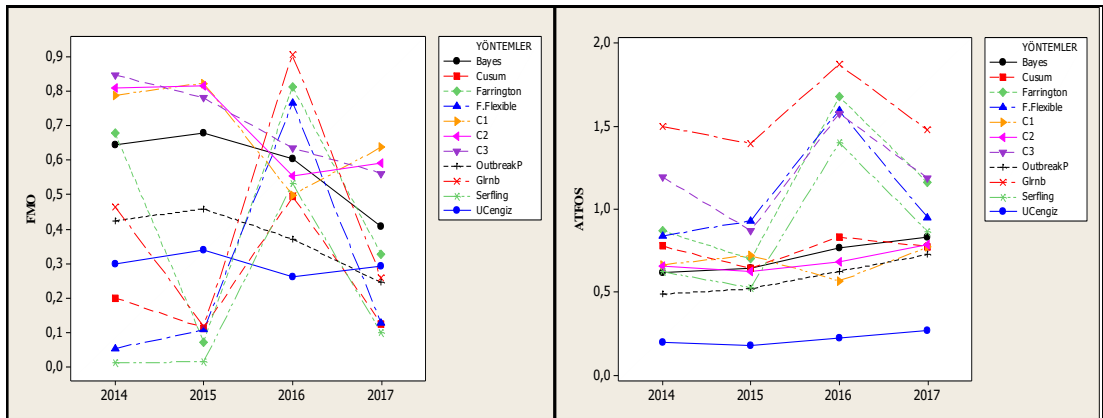
Algoritmalar	Senaryo1(2014)		Senaryo2(2015)		Senaryo3(2016)		Senaryo4(2017)	
	FMO	ATFOS	FMO	ATFOS	FMO	ATFOS	FMO	ATFOS
Bayes	0.7319	0.8113	0.7351	0.8615	0.6884	1.2211	0.4748	1.0543
Cusum	0.5342	0.8620	0.3328	0.7866	0.6620	1.2077	0.2228	0.9629
Farrington	0.8285	0.7942	0.2194	0.9270	0.8990	1.9248	0.5071	1.3824
F. Flexible	0.4309	1.0938	0.4845	1.0923	0.8757	1.9372	0.2648	1.1581
C1	0.8455	0.9612	0.8522	0.9540	0.6093	0.9693	0.6863	0.9863
C2	0.8671	0.8758	0.8494	0.7829	0.6475	1.0440	0.6497	0.9772
C3	0.8925	1.0521	0.8101	0.7583	0.7376	1.6780	0.6258	1.1638
OutbreakP	0.5595	0.5694	0.5733	0.6295	0.5405	0.8971	0.3373	0.9113
Glnb	0.7922	1.3638	0.3673	1.4957	0.9573	2.0492	0.4911	1.6412
Serfling	0.1946	1.0318	0.0941	0.8268	0.7523	1.6867	0.2167	1.0577
UCengiz	0.3564	0.2797	0.3762	0.2328	0.3494	0.3903	0.3195	0.3924



Şekil 6. 4 hafta süreli ve  $0.25\mu_{Baseline}$  büyüklüğündeki salgınlarda algoritmaların performans grafikleri

Çizelge 6.7. 4 hafta süreli ve  $0.5\mu_{Baseline}$  büyüklüğündeki salgınlarda algoritmaların performansları (10000 tekrar)

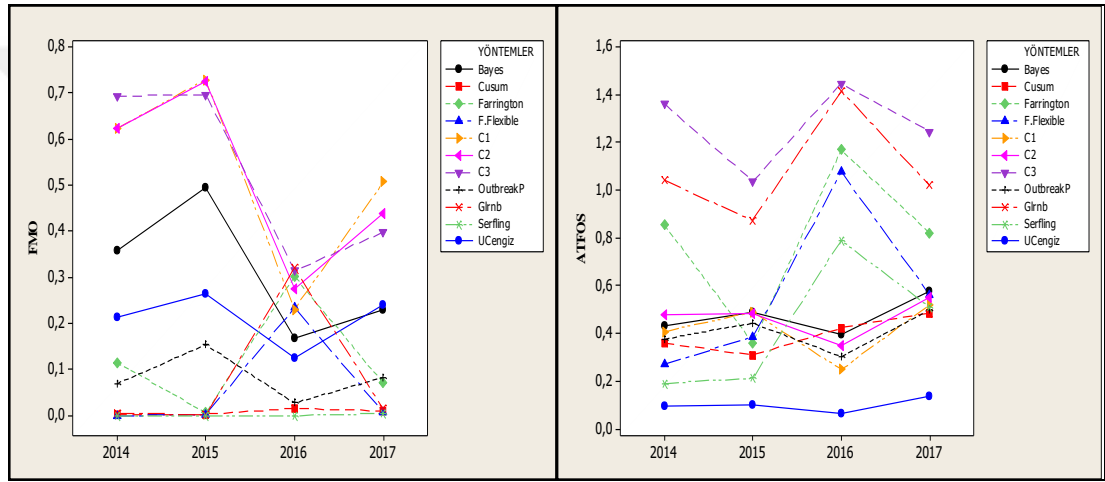
Algoritmalar	Senaryo1(2014)		Senaryo2(2015)		Senaryo3(2016)		Senaryo4(2017)	
	FMO	ATFOS	FMO	ATFOS	FMO	ATFOS	FMO	ATFOS
Bayes	0.6458	0.6208	0.6782	0.6461	0.6049	0.7677	0.4096	0.8305
Cusum	0.2017	0.7775	0.1140	0.6438	0.4954	0.8327	0.1266	0.7750
Farrington	0.6788	0.8705	0.0740	0.7004	0.8130	1.6754	0.3280	1.1612
F. Flexible	0.0546	0.8384	0.1094	0.9315	0.7672	1.5975	0.1275	0.9481
C1	0.7895	0.6670	0.8236	0.7211	0.5003	0.5675	0.6382	0.7736
C2	0.8097	0.6600	0.8173	0.6284	0.5538	0.6844	0.5907	0.7855
C3	0.8477	1.1930	0.7818	0.8712	0.6351	1.5760	0.5626	1.1882
OutbreakP	0.4250	0.4889	0.4594	0.5216	0.3718	0.6270	0.2457	0.7256
Glnb	0.4657	1.4980	0.1205	1.3964	0.9058	1.8726	0.2589	1.4755
Serfling	0.0141	0.6240	0.0156	0.5264	0.5328	1.4024	0.1019	0.8673
UCengiz	0.3004	0.1970	0.3415	0.1819	0.2623	0.2220	0.2945	0.2730



Şekil 6. 5. 4 hafta süreli ve  $0.5\mu_{Baseline}$  büyüklüğündeki salgınlarda algoritmaların performans grafikleri

**Çizelge 6.8. 4 hafta süreli ve  $1\mu_{Baseline}$  büyüklüğündeki salgınlarda algoritmaların performansları (10000 tekrar)**

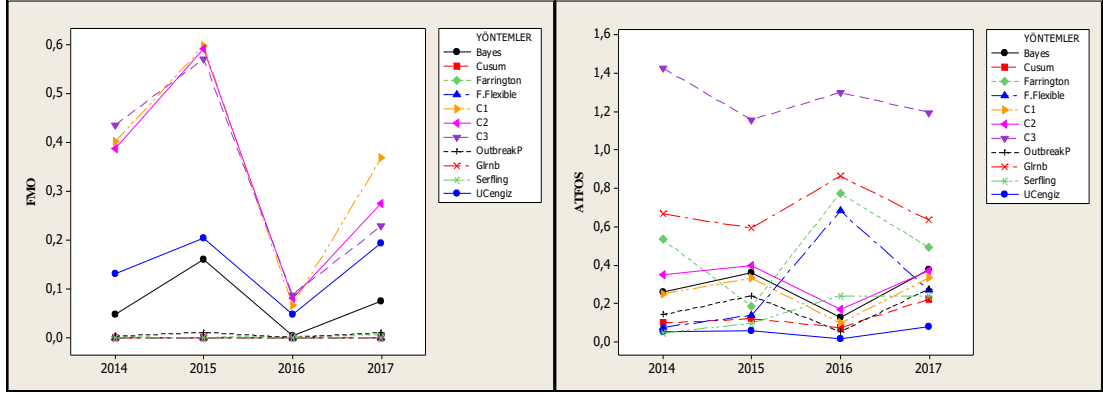
Algoritmalar	Senaryo1(2014)		Senaryo2(2015)		Senaryo3(2016)		Senaryo4(2017)	
	FMO	ATFOS	FMO	ATFOS	FMO	ATFOS	FMO	ATFOS
Bayes	0.3589	0.4313	0.4959	0.4874	0.1683	0.3943	0.2298	0.5750
Cusum	0.0010	0.3597	0.0029	0.3101	0.0154	0.4245	0.0104	0.4870
Farrington	0.1153	0.8555	0.0062	0.3616	0.3020	1.1701	0.0708	0.8218
F. Flexible	0.0000	0.2739	0.0014	0.3851	0.2340	1.0779	0.0076	0.5646
C1	0.6237	0.4090	0.7300	0.4904	0.2303	0.2501	0.5080	0.5205
C2	0.6246	0.4808	0.7268	0.4854	0.2745	0.3475	0.4387	0.5514
C3	0.6941	1.3622	0.6979	1.0394	0.3153	1.4481	0.3981	1.2464
OutbreakP	0.0688	0.3780	0.1541	0.4437	0.0273	0.3040	0.0828	0.5015
Glrnb	0.0053	1.0412	0.0015	0.8750	0.3205	1.4171	0.0148	1.0220
Serfling	0.0000	0.1917	0.0002	0.2137	0.0000	0.7910	0.0045	0.5128
UCengiz	0.2138	0.0955	0.2655	0.1037	0.1247	0.0660	0.2406	0.1387



**Şekil 6. 6. 4 hafta süreli ve  $1\mu_{Baseline}$  büyüklüğündeki salgınlarda algoritmaların performans grafikleri**

**Çizelge 6.9. 4 hafta süreli ve  $1.5\mu_{Baseline}$  büyüklüğündeki salgınlarda algoritmaların performansları (10000 tekrar)**

Algoritmalar	Senaryo1(2014)		Senaryo2(2015)		Senaryo3(2016)		Senaryo4(2017)	
	FMO	ATFOS	FMO	ATFOS	FMO	ATFOS	FMO	ATFOS
Bayes	0.0483	0.2605	0.1608	0.3608	0.0048	0.1272	0.0752	0.3774
Cusum	0.0000	0.0995	0.0000	0.1215	0.0000	0.0755	0.0000	0.2254
Farrington	0.0008	0.5362	0.0004	0.1854	0.0024	0.7767	0.0058	0.4918
F. Flexible	0.0000	0.0775	0.0000	0.1394	0.0003	0.6840	0.0000	0.2708
C1	0.4014	0.2521	0.5975	0.3327	0.0658	0.1002	0.3678	0.3328
C2	0.3865	0.3518	0.5925	0.3971	0.0819	0.1687	0.2746	0.3726
C3	0.4357	1.4260	0.5715	1.1587	0.0881	1.3007	0.2280	1.1957
OutbreakP	0.0029	0.1425	0.0108	0.2415	0.0000	0.0555	0.0101	0.2768
Glrnb	0.0000	0.6674	0.0000	0.5940	0.0006	0.8643	0.0001	0.6352
Serfling	0.0000	0.0489	0.0000	0.0969	0.0000	0.2422	0.0001	0.2424
UCengiz	0.1312	0.0546	0.2032	0.0596	0.0478	0.0169	0.1943	0.0799



Şekil 6. 7. 4 hafta süreli ve  $1.5\mu_{Baseline}$  büyüklüğündeki salgınlarda algoritmaların performans grafikleri

Çizelge 6. 10. 4 hafta süreli salgınlar için en iyi performansa sahip algoritmalar

Salgın (M)	Senaryo1(2014)		Senaryo2(2015)		Senaryo3(2016)		Senaryo4(2017)	
	FMO	ATFOS	FMO	ATFOS	FMO	ATFOS	FMO	ATFOS
0.1	UCengiz	UCengiz	Serfling	UCengiz	UCengiz	UCengiz	Cusum	UCengiz
0.25	Serfling	UCengiz	Serfling	UCengiz	UCengiz	UCengiz	Serfling	UCengiz
0.50	Serfling	UCengiz	Serfling	UCengiz	UCengiz	UCengiz	Serfling	UCengiz
1.00	Serfling F. Flex.	UCengiz	Serfling	UCengiz	Serfling	UCengiz	Serfling	UCengiz
1.50	Cusum F. Flex. Glrnb Serfling	UCengiz	Cusum F. Flex. Glrnb Serfling	UCengiz	Cusum OutbreakP Serfling	UCengiz	Cusum F. Flex.	UCengiz

### 6.5.2 Sekiz hafta süreli salgınlar için simülasyon sonuçları

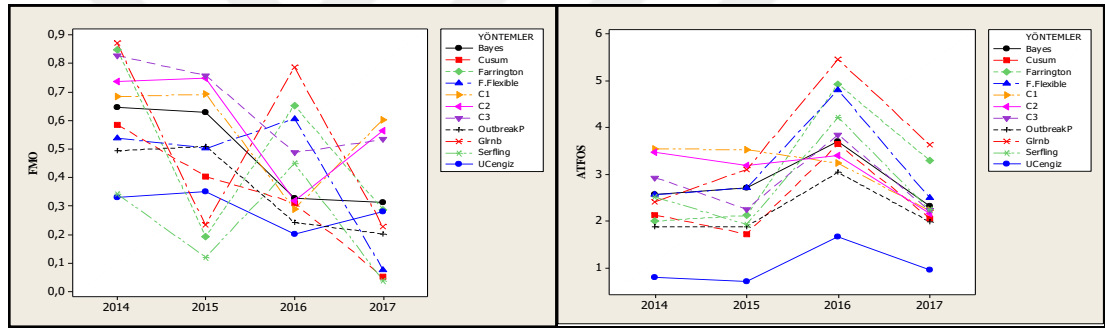
Salgın büyüklüğü, referans dönemi ortalamasının 0.10, 0.25, 0.50, 1.00 ve 1.50 katı olacak biçimde sekiz haftalık süreli salgınlar için; 10000 tekrarlı simülasyonlarda elde edilen FMO ve ATFOS değerleri, Çizelge 6.11.-Çizelge 6.15. ve Şekil 6.8.-Şekil 6.12.'de verilmektedir. İlgili çizelgeler ve şekiller incelendiğinde değişik senaryolarda en iyi performansa sahip algoritmalar Çizelge 6.16.'da verilmektedir.

Çizelge 6.16. incelendiğinde; sekiz hafta süreli salgınlarda ve tüm salgın büyüklüklerinde ATFOS açısından bu çalışmada önerilen UCengiz yönteminin performans üstünlüğü sürerken, FMO açısından ise Serfling, F. Flexible, Glrnb, Cusum, Farrington yöntemleri ve bu tez kapsamında önerilen UCengiz yöntemi sırasıyla en iyi performansı sergilemektedir. Dört hafta süreli salgınlarda olduğu gibi,

bu çalışmada önerilen UCengiz yönteminin küçük salgın büyüklüklerinde FMO açısından üstün olma eğilimi vardır.

**Çizelge 6.11. 8 hafta süreli ve  $0.10\mu_{Baseline}$  büyüklüğündeki salgınlarda algoritmaların performansları (10000 tekrar)**

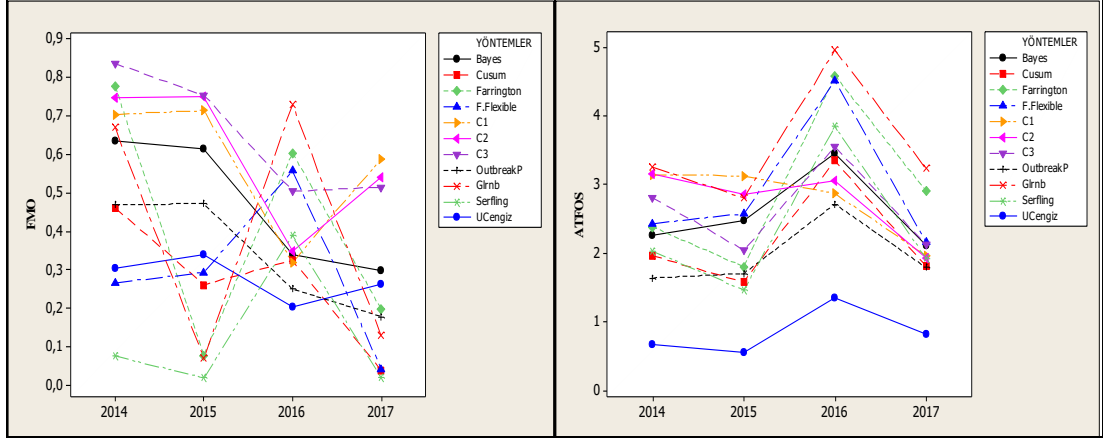
Algoritmalar	Senaryo1(2014)		Senaryo2(2015)		Senaryo3(2016)		Senaryo4(2017)	
	FMO	ATFOS	FMO	ATFOS	FMO	ATFOS	FMO	ATFOS
Bayes	0.6461	2.5685	0.6278	2.7168	0.3294	3.7049	0.3139	2.3262
Cusum	0.5863	2.1378	0.4047	1.7294	0.3118	3.6507	0.0522	2.0689
Farrington	0.8487	2.0073	0.1945	2.1286	0.6537	4.9347	0.2867	3.2999
F. Flexible	0.5396	2.5734	0.5031	2.7112	0.6071	4.8073	0.0779	2.5118
C1	0.6838	3.5478	0.6924	3.5299	0.2906	3.2447	0.6029	2.2322
C2	0.7365	3.4748	0.7485	3.1857	0.3203	3.4050	0.5650	2.1710
C3	0.8280	2.9366	0.7584	2.2649	0.4875	3.8505	0.5347	2.2304
OutbreakP	0.4942	1.8863	0.5099	1.8906	0.2438	3.0444	0.2026	1.9931
Glrnb	0.8712	2.4161	0.2333	3.1041	0.7876	5.4647	0.2279	3.6414
Serfling	0.3433	2.5273	0.1221	1.9312	0.4501	4.2215	0.0378	2.2517
UCengiz	0.3319	0.7999	0.3502	0.7196	0.2031	1.6750	0.2815	0.9592



**Şekil 6. 8 hafta süreli ve  $0.1\mu_{Baseline}$  büyüklüğündeki salgınlarda algoritmaların performans grafikleri**

**Çizelge 6.12. 8 hafta süreli ve  $0.25\mu_{Baseline}$  büyüklüğündeki salgınlarda algoritmaların performansları (10000 tekrar)**

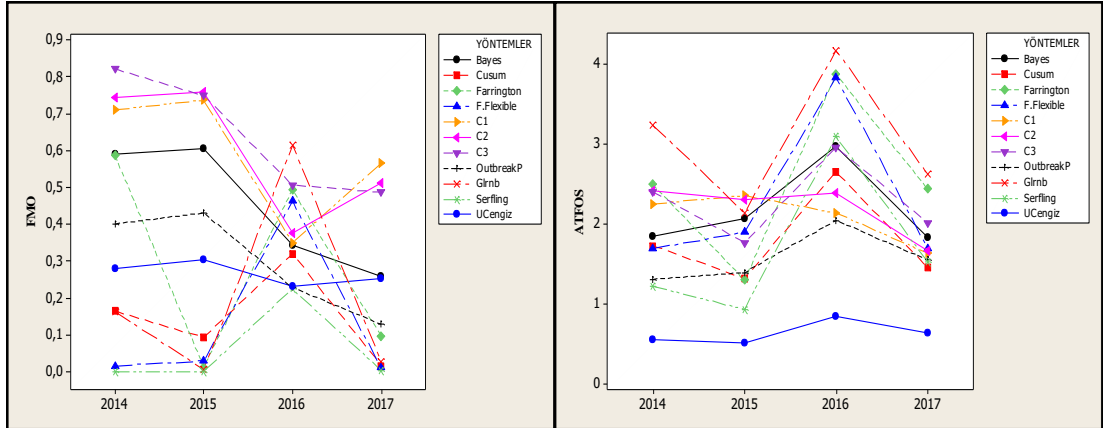
Algoritmalar	Senaryo1(2014)		Senaryo2(2015)		Senaryo3(2016)		Senaryo4(2017)	
	FMO	ATFOS	FMO	ATFOS	FMO	ATFOS	FMO	ATFOS
Bayes	0.6366	2.2570	0.6141	2.4841	0.3417	3.4636	0.2978	2.1139
Cusum	0.4622	1.9643	0.2623	1.5805	0.3253	3.3529	0.0405	1.8129
Farrington	0.7779	2.3854	0.0791	1.8052	0.6046	4.5852	0.1999	2.9086
F. Flexible	0.2680	2.4372	0.2942	2.5776	0.5577	4.5089	0.0429	2.1591
C1	0.7048	3.1463	0.7148	3.1283	0.3208	2.8732	0.5893	1.9586
C2	0.7492	3.1643	0.7511	2.8614	0.3494	3.0553	0.5421	1.9244
C3	0.8371	2.8158	0.7547	2.0465	0.5053	3.5569	0.5158	2.1314
OutbreakP	0.4694	1.6412	0.4736	1.7073	0.2521	2.7177	0.1794	1.7954
Glrnb	0.6700	3.2530	0.0728	2.8098	0.7300	4.9637	0.1300	3.2398
Serfling	0.0775	2.0349	0.0205	1.4673	0.3900	3.8456	0.0205	1.9253
UCengiz	0.3053	0.6660	0.3407	0.5579	0.2038	1.3561	0.2651	0.8203



Şekil 6. 9. 8 hafta süreli ve  $0.25\mu_{Baseline}$  büyüklüğündeki salgınlarda algoritmaların performans grafikleri

Çizelge 6.13. 8 hafta süreli ve  $0.5\mu_{Baseline}$  büyüklüğündeki salgınlarda algoritmaların performansları (10000 tekrar)

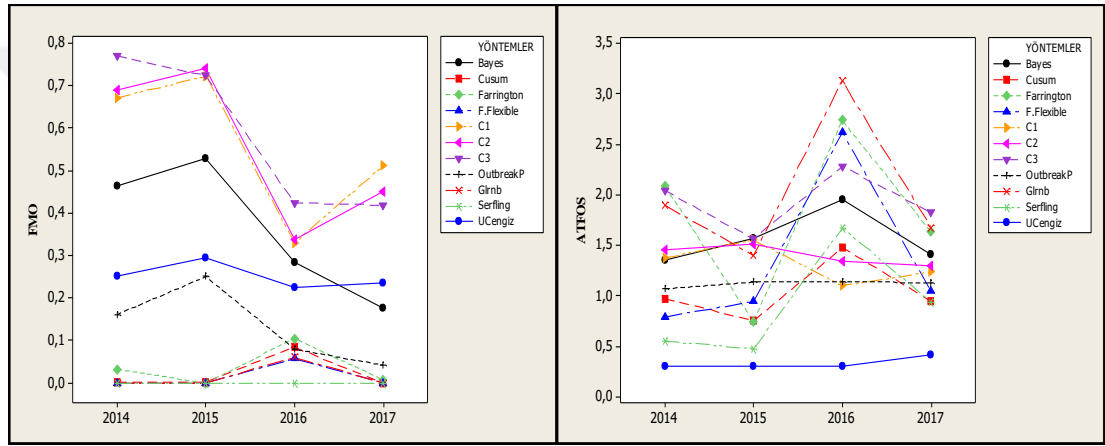
Algoritmalar	Senaryo1(2014)		Senaryo2(2015)		Senaryo3(2016)		Senaryo4(2017)	
	FMO	ATFOS	FMO	ATFOS	FMO	ATFOS	FMO	ATFOS
Bayes	0.5919	1.8405	0.6059	2.0759	0.3450	2.9722	0.2607	1.8335
Cusum	0.1653	1.7272	0.0935	1.3216	0.3211	2.6589	0.0166	1.4578
Farrington	0.5872	2.4966	0.0139	1.2976	0.4938	3.8746	0.0967	2.4386
F. Flexible	0.0168	1.6925	0.0299	1.8977	0.4632	3.8329	0.0118	1.6979
C1	0.7113	2.2546	0.7376	2.3586	0.3509	2.1382	0.5665	1.6339
C2	0.7449	2.4187	0.7594	2.3084	0.3783	2.3856	0.5141	1.6722
C3	0.8225	2.4011	0.7497	1.7659	0.5067	2.9659	0.4875	2.0162
OutbreakP	0.4020	1.3087	0.4315	1.3926	0.2291	2.0480	0.1292	1.5512
Glnb	0.1622	3.2449	0.0064	2.1434	0.6138	4.1735	0.0286	2.6315
Serfling	0.0015	1.2219	0.0014	0.9315	0.2255	3.0957	0.0043	1.5201
UCengiz	0.2798	0.5467	0.3036	0.5138	0.2334	0.8445	0.2533	0.6430



Şekil 6. 10. 8 hafta süreli ve  $0.5\mu_{Baseline}$  büyüklüğündeki salgınlarda algoritmaların performans grafikleri

**Çizelge 6.14. 8 hafta süreli ve  $1\mu_{Baseline}$  büyüklüğündeki salgınlarda algoritmaların performansları (10000 tekrar)**

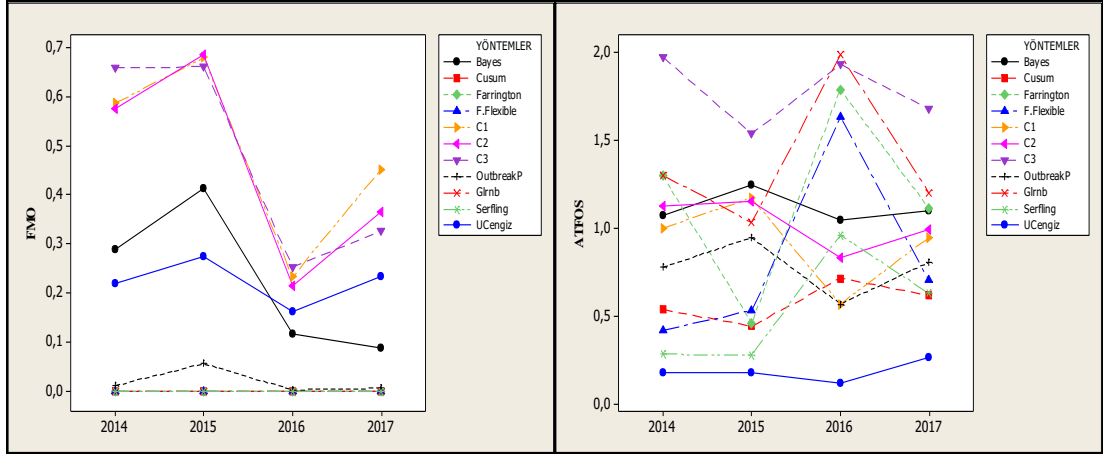
Algoritmalar	Senaryo1(2014)		Senaryo2(2015)		Senaryo3(2016)		Senaryo4(2017)	
	FMO	ATFOS	FMO	ATFOS	FMO	ATFOS	FMO	ATFOS
Bayes	0.4635	1.3530	0.5279	1.5730	0.2849	1.9494	0.1777	1.4064
Cusum	0.0013	0.9763	0.0022	0.7535	0.0844	1.4786	0.0004	0.9488
Farrington	0.0328	2.0839	0.0000	0.7431	0.1051	2.7383	0.0075	1.6320
F. Flexible	0.0000	0.7957	0.0000	0.9452	0.0577	2.6198	0.0001	1.0479
C1	0.6708	1.3764	0.7218	1.5561	0.3292	1.1011	0.5113	1.2372
C2	0.6879	1.4611	0.7395	1.5159	0.3367	1.3406	0.4505	1.2930
C3	0.7691	2.0398	0.7234	1.5651	0.4240	2.2809	0.4183	1.8257
OutbreakP	0.1615	1.0708	0.2505	1.1370	0.0802	1.1446	0.0415	1.1276
Glrnb	0.0001	1.8980	0.0000	1.3990	0.0612	3.1280	0.0001	1.6748
Serfling	0.0000	0.5487	0.0000	0.4797	0.0000	1.6731	0.0001	0.9427
UCengiz	0.2515	0.3029	0.2949	0.3027	0.2259	0.3068	0.2365	0.4195



**Şekil 6.11. 8 hafta süreli ve  $1\mu_{Baseline}$  büyüklüğündeki salgınlarda algoritmaların performans grafikleri**

**Çizelge 6.15. 8 hafta süreli ve  $1.5\mu_{Baseline}$  büyüklüğündeki salgınlarda algoritmaların performansları (10000 tekrar)**

Algoritmalar	Senaryo1(2014)		Senaryo2(2015)		Senaryo3(2016)		Senaryo4(2017)	
	FMO	ATFOS	FMO	ATFOS	FMO	ATFOS	FMO	ATFOS
Bayes	0.2895	1.0778	0.4129	1.2470	0.1173	1.0491	0.0890	1.1012
Cusum	0.0000	0.5409	0.0000	0.4455	0.0000	0.7179	0.0000	0.6196
Farrington	0.0000	1.3000	0.0000	0.4641	0.0000	1.7897	0.0001	1.1133
F. Flexible	0.0000	0.4209	0.0000	0.5351	0.0000	1.6319	0.0000	0.7055
C1	0.5867	1.0031	0.6798	1.1758	0.2327	0.5717	0.4520	0.9507
C2	0.5761	1.1248	0.6853	1.1538	0.2141	0.8339	0.3655	0.9945
C3	0.6583	1.9713	0.6614	1.5381	0.2535	1.9303	0.3263	1.6825
OutbreakP	0.0110	0.7819	0.0560	0.9474	0.0026	0.5713	0.0061	0.8091
Glrnb	0.0000	1.3023	0.0000	1.0353	0.0000	1.9878	0.0000	1.2001
Serfling	0.0000	0.2869	0.0000	0.2808	0.0000	0.9642	0.0000	0.6325
UCengiz	0.2205	0.1840	0.2744	0.1821	0.1610	0.1205	0.2342	0.2677



Şekil 6. 12. 8 hafta süreli ve  $1.5\mu_{Baseline}$  büyüklüğündeki salgınlarda algoritmaların performans grafikleri

Çizelge 6. 16. 8 hafta süreli salgınlar için en iyi performansa sahip algoritmalar

Salgın (M)	Senaryo1(2014)		Senaryo2(2015)		Senaryo3(2016)		Senaryo4(2017)	
	FMO	ATFOS	FMO	ATFOS	FMO	ATFOS	FMO	ATFOS
0.1	UCengiz	UCengiz	Serfling	UCengiz	UCengiz	UCengiz	Serfling	UCengiz
0.25	Serfling	UCengiz	Serfling	UCengiz	UCengiz	UCengiz	Serfling	UCengiz
0.50	Serfling	UCengiz	Serfling	UCengiz	Serfling	UCengiz	Serfling	UCengiz
1.00	F. Flex. Serfling	UCengiz	Farring. F. Flex. Glnb Serfling	UCengiz	Serfling	UCengiz	F. Flex. Glnb Serfling	UCengiz
1.50	Cusum Farring. F. Flex. Glnb Serfling	UCengiz	Cusum Farring. F. Flex. Glnb Serfling	UCengiz	Cusum Farring. F. Flex. Glnb Serfling	UCengiz	Cusum F. Flex. Glnb Serfling	UCengiz

### 6.5.3 On iki hafta süreli salgınlar için simülasyon sonuçları

Salgın büyüklüğü, referans dönemi ortalamasının 0.10, 0.25, 0.50, 1.00 ve 1.50 katı olacak biçimde on iki haftalık süreli salgınlar için; 10000 tekrarlı simülasyonlarda elde edilen FMO ve ATFOS değerleri, Çizelge 6.17.-Çizelge 6.21. ve Şekil 6.13.-Şekil 6.17.'de verilmektedir. İlgili çizelgeler ve şekiller incelendiğinde değişik senaryolarda en iyi performansa sahip algoritmalar Çizelge 6.22.'de verilmektedir.

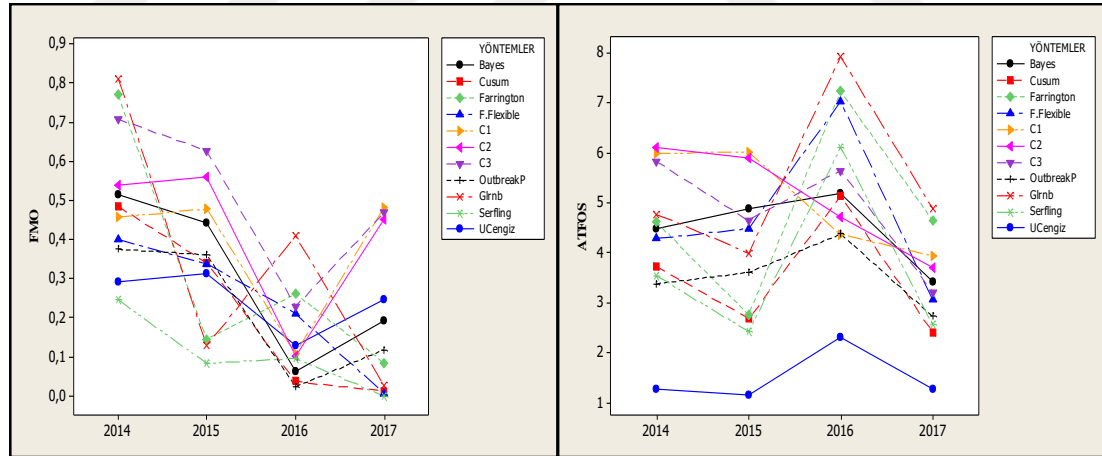
Çizelge 6.22. incelendiğinde; on iki hafta süreli salgınlarda ve tüm salgın büyüklüklerinde ATFOS açısından bu çalışmada önerilen UCengiz yönteminin



performansı genel olarak üstünlüğü ön plana çıkarken, bazı durumlarda da Serfling yöntemi en iyi performansa sahip olduğu görülmektedir. FMO açısından ise Serfling, Glrnb, F. Flexible, Farrington, Cusum, OutbreakP yöntemleri sırasıyla en iyi performansı vermektedir. Dört hafta ve sekiz hafta süreli salgınların aksine, bu tez kapsamında önerilen UCengiz yöntemi on iki haftalık süreli tüm salgın büyüklüklerinde FMO performansı yüksek olma eğiliminde değildir.

**Çizelge 6.17. 12 hafta süreli ve  $0.10\mu_{Baseline}$  büyüklüğündeki salgınlarda algoritmaların performansları (10000 tekrar)**

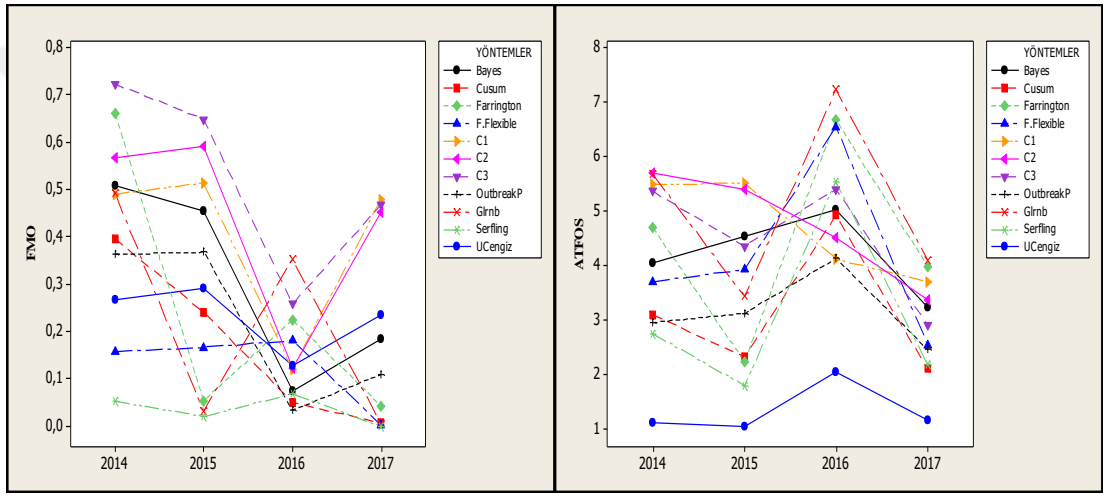
Algoritmalar	Senaryo1(2014)		Senaryo2(2015)		Senaryo3(2016)		Senaryo4(2017)	
	FMO	ATFOS	FMO	ATFOS	FMO	ATFOS	FMO	ATFOS
Bayes	0.5162	4.4833	0.4417	4.8931	0.0636	5.1979	0.1937	3.4141
Cusum	0.4853	3.7332	0.3407	2.6901	0.0378	5.1412	0.0123	2.4068
Farrington	0.7705	4.6200	0.1451	2.7581	0.2620	7.2369	0.0847	4.6527
F. Flexible	0.4004	4.2877	0.3389	4.4913	0.2114	7.0218	0.0073	3.0670
C1	0.4580	5.9928	0.4785	6.0213	0.1077	4.3675	0.4830	3.9460
C2	0.5384	6.1053	0.5616	5.9056	0.1023	4.7305	0.4515	3.7025
C3	0.7082	5.8321	0.6276	4.6541	0.2293	5.6484	0.4689	3.2203
OutbreakP	0.3755	3.3835	0.3613	3.6114	0.0241	4.3816	0.1190	2.7334
Glrnb	0.8122	4.7785	0.1295	3.9974	0.4101	7.9200	0.0276	4.8927
Serfling	0.2463	3.5460	0.0839	2.4337	0.0968	6.1138	0.0005	2.5773
UCengiz	0.2931	1.2734	0.3134	1.1675	0.1292	2.3099	0.2462	1.2798



**Şekil 6. 13. 12 hafta süreli ve  $0.1\mu_{Baseline}$  büyüklüğündeki salgınlarda algoritmaların performans grafikleri**

**Çizelge 6.18. 12 hafta süreli ve  $0.25\mu_{Baseline}$  büyüklüğündeki salgınlarda algoritmaların performansları (10000 tekrar)**

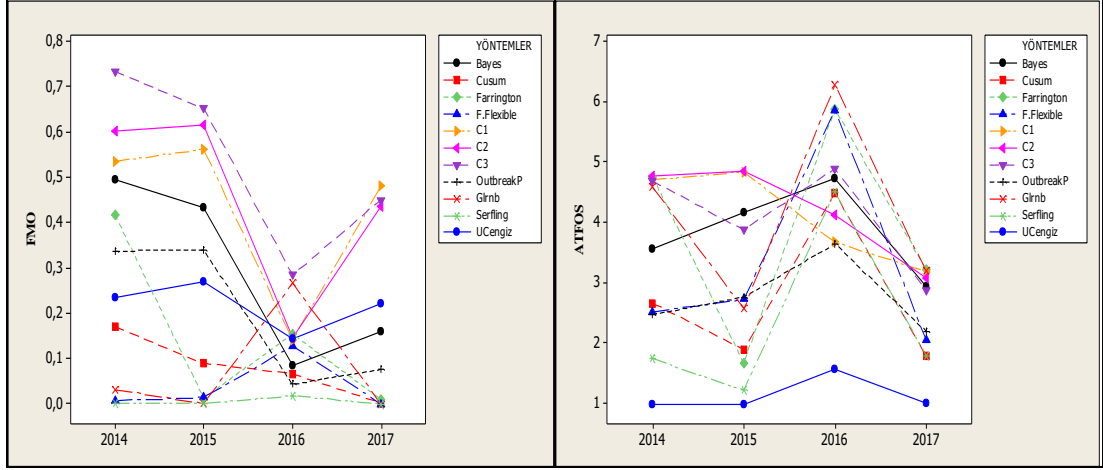
Algoritmalar	Senaryo1(2014)		Senaryo2(2015)		Senaryo3(2016)		Senaryo4(2017)	
	FMO	ATFOS	FMO	ATFOS	FMO	ATFOS	FMO	ATFOS
Bayes	0.5080	4.0435	0.4560	4.5452	0.0752	5.0322	0.1837	3.2356
Cusum	0.3965	3.1014	0.2409	2.3254	0.0508	4.9272	0.0063	2.1144
Farrington	0.6609	4.7016	0.0543	2.2476	0.2256	6.6746	0.0434	3.9934
F. Flexible	0.1585	3.7136	0.1669	3.9336	0.1817	6.5503	0.0026	2.5525
C1	0.4910	5.4923	0.5147	5.5205	0.1208	4.1237	0.4795	3.7091
C2	0.5687	5.7116	0.5925	5.4054	0.1222	4.5185	0.4519	3.3897
C3	0.7239	5.3705	0.6493	4.3581	0.2584	5.4072	0.4688	2.9143
OutbreakP	0.3634	2.9632	0.3682	3.1301	0.0334	4.1388	0.1103	2.4650
Glrnb	0.4916	5.6772	0.0323	3.4382	0.3523	7.2365	0.0054	4.1115
Serfling	0.0528	2.7541	0.0207	1.8101	0.0694	5.5509	0.0000	2.1860
UCengiz	0.2665	1.1181	0.2920	1.0534	0.1274	2.0639	0.2346	1.1723



**Şekil 6. 14. 12 hafta süreli ve  $0.25\mu_{Baseline}$  büyüklüğündeki salgınlarda algoritmaların performans grafikleri**

**Çizelge 6.19. 12 hafta süreli ve  $0.5\mu_{Baseline}$  büyüklüğündeki salgınlarda algoritmaların performansları (10000 tekrar)**

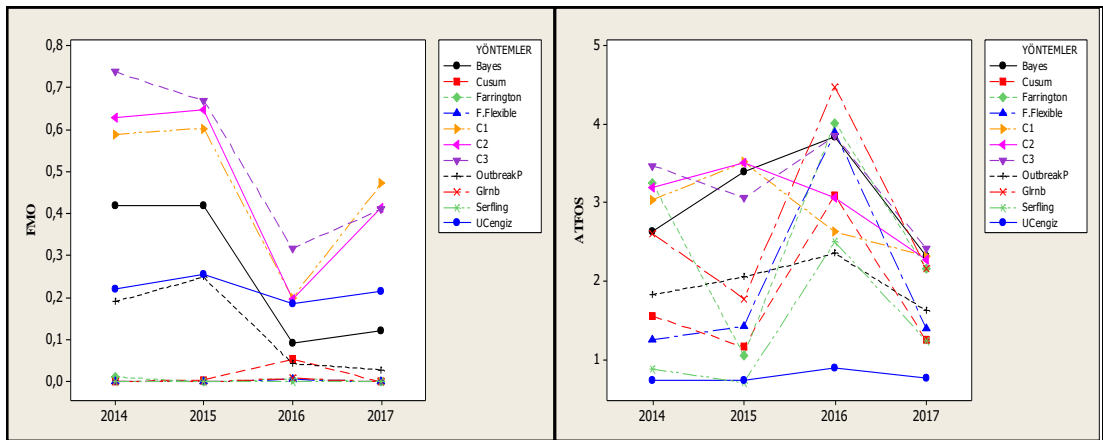
Algoritmalar	Senaryo1(2014)		Senaryo2(2015)		Senaryo3(2016)		Senaryo4(2017)	
	FMO	ATFOS	FMO	ATFOS	FMO	ATFOS	FMO	ATFOS
Bayes	0.4959	3.5620	0.4348	4.1617	0.0845	4.7381	0.1609	2.9415
Cusum	0.1702	2.6576	0.0901	1.8982	0.0661	4.4907	0.0017	1.7790
Farrington	0.4170	4.7554	0.0127	1.6712	0.1541	5.8797	0.0105	3.2162
F. Flexible	0.0069	2.5151	0.0143	2.7263	0.1280	5.8584	0.0001	2.0523
C1	0.5358	4.7042	0.5627	4.8390	0.1406	3.6738	0.4809	3.1871
C2	0.6029	4.7688	0.6161	4.8570	0.1470	4.1209	0.4371	3.0714
C3	0.7335	4.6889	0.6532	3.8740	0.2850	4.8883	0.4509	2.8796
OutbreakP	0.3382	2.4725	0.3389	2.7601	0.0433	3.6456	0.0760	2.1957
Glrnb	0.0323	4.5980	0.0006	2.5648	0.2662	6.2766	0.0002	3.1905
Serfling	0.0006	1.7449	0.0011	1.2150	0.0182	4.4991	0.0000	1.7890
UCengiz	0.2354	0.9846	0.2705	0.9782	0.1438	1.5693	0.2208	0.9997



Şekil 6. 15. 12 hafta süreli ve  $0.5\mu_{Baseline}$  büyüklüğündeki salgınlarda algoritmaların performans grafikleri

Çizelge 6.20. 12 hafta süreli ve  $1\mu_{Baseline}$  büyüklüğündeki salgınlarda algoritmaların performansları (10000 tekrar)

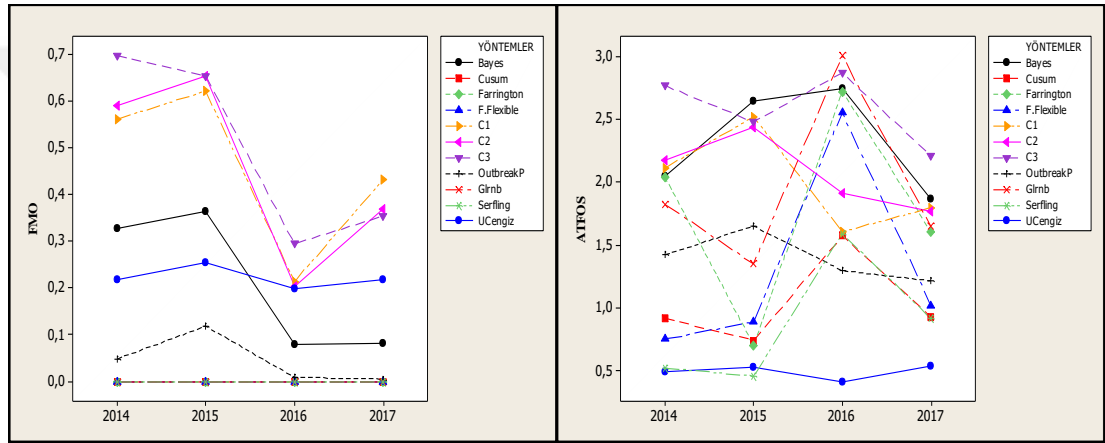
Algoritmalar	Senaryo1(2014)		Senaryo2(2015)		Senaryo3(2016)		Senaryo4(2017)	
	FMO	ATFOS	FMO	ATFOS	FMO	ATFOS	FMO	ATFOS
Bayes	0.4213	2.6420	0.4200	3.3964	0.0935	3.8459	0.1210	2.3192
Cusum	0.0011	1.5604	0.0049	1.1654	0.0548	3.0983	0.0000	1.2538
Farrington	0.0119	3.2549	0.0000	1.0576	0.0080	4.0184	0.0003	2.1633
F. Flexible	0.0000	1.2570	0.0000	1.4338	0.0070	3.9002	0.0000	1.4034
C1	0.5902	3.0361	0.6035	3.5203	0.1999	2.6297	0.4728	2.3253
C2	0.6297	3.1917	0.6476	3.5111	0.1967	3.0692	0.4144	2.2751
C3	0.7383	3.4731	0.6695	3.0611	0.3186	3.8487	0.4117	2.4187
OutbreakP	0.1911	1.8272	0.2498	2.0556	0.0438	2.3561	0.0290	1.6291
Glnb	0.0000	2.6097	0.0000	1.7811	0.0088	4.4775	0.0000	2.1649
Serfling	0.0000	0.8871	0.0000	0.7086	0.0000	2.5051	0.0000	1.2459
UCengiz	0.2215	0.7434	0.2555	0.7424	0.1871	0.9028	0.2162	0.7701



Şekil 6. 16. 12 hafta süreli ve  $1\mu_{Baseline}$  büyüklüğündeki salgınlarda algoritmaların performans grafikleri

**Çizelge 6.21. 12 hafta süreli ve  $1.5\mu_{Baseline}$  büyüklüğündeki salgınlarda algoritmaların performansları (10000 tekrar)**

Algoritmalar	Senaryo1(2014)		Senaryo2(2015)		Senaryo3(2016)		Senaryo4(2017)	
	FMO	ATFOS	FMO	ATFOS	FMO	ATFOS	FMO	ATFOS
Bayes	0.3284	2.0451	0.3639	2.6471	0.0800	2.7439	0.0817	1.8711
Cusum	0.0000	0.9195	0.0000	0.7441	0.0006	1.5759	0.0000	0.9316
Farrington	0.0000	2.0446	0.0000	0.7074	0.0000	2.7199	0.0000	1.6053
F. Flexible	0.0000	0.7538	0.0000	0.8920	0.0000	2.5544	0.0000	1.0217
C1	0.5619	2.1125	0.6209	2.5212	0.2145	1.6080	0.4316	1.7970
C2	0.5907	2.1715	0.6525	2.4386	0.2036	1.9145	0.3681	1.7693
C3	0.6965	2.7690	0.6532	2.4792	0.2953	2.8709	0.3542	2.2094
OutbreakP	0.0485	1.4247	0.1189	1.6537	0.0093	1.2999	0.0055	1.2178
Glrnb	0.0000	1.8274	0.0000	1.3525	0.0000	3.0047	0.0000	1.6485
Serfling	0.0000	0.5247	0.0000	0.4625	0.0000	1.5910	0.0000	0.9191
UCengiz	0.2192	0.4977	0.2549	0.5309	0.1990	0.4170	0.2185	0.5387



**Şekil 6.17. 12 hafta süreli ve  $1.5\mu_{Baseline}$  büyüklüğündeki salgınlarda algoritmaların performans grafikleri**

**Çizelge 6.22. 12 hafta süreli salgınlarda için en iyi performansa sahip algoritmalar**

Salgın (M)	Senaryo1(2014)		Senaryo2(2015)		Senaryo3(2016)		Senaryo4(2017)	
	FMO	ATFOS	FMO	ATFOS	FMO	ATFOS	FMO	ATFOS
0.1	Serfling	UCengiz	Serfling	UCengiz	OutbreakP	UCengiz	Serfling	UCengiz
0.25	Serfling	UCengiz	Serfling	UCengiz	OutbreakP	UCengiz	Serfling	UCengiz
0.50	Serfling	UCengiz	Glrnb	UCengiz	Serfling	UCengiz	Serfling	UCengiz
1.00	F. Flex Glrnb Serfling	UCengiz	Farring. F. Flex. Glrnb Serfling	Serfling	Serfling	UCengiz	Cusum F. Flex. Glrnb Serfling	UCengiz
1.50	Cusum Farring. F. Flex. Glrnb Serfling	UCengiz	Cusum Farring. F. Flex. Glrnb Serfling	Serfling	Farring. F. Flex. Glrnb Serfling	UCengiz	Cusum Farring. F. Flex. Glrnb Serfling	UCengiz

## 7.SONUÇLAR VE TARTIŞMA

Halk sađlıđı aısından salgınların erkenden belirlenerek önlemlerin alınması önem arz etmektedir. Bu nedenle salgınların doğrudan veya dolaylı göstergesi olan olaylara ilişkin bilgiler toplanmakta ve halk sađlıđı izleme sistemleri olarak adlandırılan süreçlerde değerlendirilmektedir. Halk sađlıđı izleme sistemlerinde verilerdeki olađan dışılıđı ya da anormalliđi belirlemek için bazı istatistiksel yöntemler kullanılmaktadır.

İdeal olarak hastalık salgınlarını izlemekte kullanılan yöntemlerin en az yanlış alarm ile kısa sürede doğru sinyal vermeleri beklenir. Hastalık salgınlarını izlemenin sađlıklı yapılabilmesi için; yöntemlerin çalışma prensiplerinin anlaşılması, doğru olarak uygulanması, yöntemlerin izleme yapılan duruma uyarlanması ya da mevcut yöntemlerin dezavantajlarını azaltan ya da bu dezavantajları taşımayan yeni yöntemlerin önerilmesi gerekmektedir.

Bu tez çalışması, zaman içinde ileriye dönük izleme yapabilen salgın belirleme algoritmalarından bazılarını tanıtmayı ve performanslarını simülasyon çalışmasıyla karşılaştırmayı hedeflemiştir. Ayrıca bu çalışma kapsamında UCengiz yöntemi olarak adlandırılan ve ileriye doğru gerçek zamanlı izleme yapabilen basit bir istatistiksel salgın belirleme yöntemi önerilmiştir.

Bu çalışmada dikkate alınan yöntemlerin uygulamaları CDC kurumundan elde edilen 2012-2016 yıllarına ilişkin haftalık grip mevsimi verileri üzerinde gerçekleştirilmiştir. Uygulanan yöntemler ise; Serfling regresyon, Farrington, FarringtonFlexible, Boda, Holt-Winters, Arima modelleme, gizli markov modeli, model artıklarını kullanan Shewhart I-Chart, Ewma, Cusum kontrol grafikleri, Rossi vd. (2009) Cusum grafiđi, Ears C1, Ears C2, Ears C3, OutbreakP, Glnb, Bayesçi sonsal bir öngörü yaklaşımı, Rki yaklaşımı, Wavelet yaklaşımı ve bu çalışmada önerilen UCengiz yöntemidir. Elde

edilen sonuçlar, CDC kurumunun ilgili haftalar için verdiği salgın durumu sonuçlarıyla karşılaştırılmıştır.

Simülasyon çalışması için ise; Bayesçi bir öngörü yaklaşımı, Rossi vd. (1999) Cusum grafiği, Farrington, FarringtonFlexible, Ears C1, Ears C2, Ears C3, OutbreakP, Glnb, Serfling regresyon modeli ve bu çalışmada önerilen UCengiz yöntemi dikkate alınmıştır. İlk olarak CDC kurumunun haftalık grip verilerinin özelliklerini yansıtacak şekilde veriler 10000 tekrarlı olarak oluşturulmuştur. Daha sonra bu verilere yapay salgınlar, çeşitli büyüklükte ve çeşitli zaman dilimlerinde geçerli olmak üzere eklenmiştir. Dikkate alınan yöntemlerin salgın olmadığı durumlardaki performansları  $ATFS = 13$  olarak eşitlenerek, salgın koşulu altındaki performansları FMO ve ATFOS ölçütleriyle değerlendirilmiştir.

Bu tez çalışmasındaki sonuçlar, CDC kurumundan elde edilen verilere yapılan uygulamalardan elde edilen sonuçlar ile simülasyon çalışmasından elde edilen sonuçlar olmak üzere iki başlık altında toplanabilir.

**Veri analizinden elde edilen sonuçlar:** 2012-2013, 2013-2014, 2014-2015 grip mevsimlerine ilişkin haftalık veriler referans kabul edilerek, 2015-2016 grip mevsimine ilişkin veriler hafta hafta yani ileriye dönük olarak salgın olup olmadığını belirlemek amacıyla ilgili yöntemler kullanılarak izlenmiştir. Ayrıca yöntemlerin dayandığı varsayımları dikkate alabilmek amacıyla; izleme için bazı yöntemlerde ham veriler, bazı yöntemlerde işlenmiş veriler kullanılmıştır.

Bu çalışmadaki veri analizindeki amaç yöntemlerin uygulamalarını tanıtmak olduğundan, elde edilen sonuçlar doğrudan karşılaştırılabilir değildir. Sadece fikir edinebilmek amacıyla; her yöntem için elde edilen sonuçlar, CDC kurumundan elde edilen sonuçlar ile karşılaştırılarak duyarlılık ve seçicilik hesaplanmıştır. Seçicilik, duyarlılık değerleri ve salgının başlangıcını belirlemekteki gecikmeler dikkate alındığında aşağıdaki yorumlar kısmen yapılabilmektedir.

- Yöntemlerin bir kısmı için; işlenmiş verilerin kullanılması, salgının belirlenmesindeki gecikmeyi azaltma eğiliminde olduğu söylenebilir.

- Gizli markov modeli yaklaşımı, Boda yöntemi ve Serfling regresyon yöntemi; salgının başlangıcını belirlemedeki gecikme süresi, seçicilik ve duyarlılık ölçütlerine göre, ham verilere uygulanan yöntemler arasında diğer yöntemlere göre ön plana çıkmaktadır.
- FarringtonFlexible, Holt-Winters ve Arima modelleme yöntemleri, salgının başlangıcını belirlemede diğer yöntemlere göre daha uzun gecikme süreleriyle ham verilere uygulanan yöntemler arasında dikkat çekmektedir.
- Salgın başlangıcını belirlemedeki dakiklik, seçicilik ve duyarlılık ölçütlerine göre; Rossi vd. (1999) Cusum grafiği, 26 ATFS değerinde UCengiz yöntemi, Ewma artıklarını kullanan I-Chart, Bayesçi bir sonsal öngörü yaklaşımı, Glnb, Wavelet (mevsimselliği giderilmiş ve farkı alınmış veriler için) ve Rki yöntemleri işlenmiş verilere uygulanan yöntemler arasında ön plana çıkmaktadır.
- İşlenmiş verilere uygulanan yöntemler arasında; Wavelet (mevsimsellik düzeltilmesi yapılmış veriler için), Ears C1, Ears C2, Ears C3, Holt-Winters artıklarını kullanan I-Chart, Ewma, Cusum grafikleri ve mevsimselliği giderilmiş verilere uygulanan I-Chart grafiği, salgının başlangıcını belirlemedeki daha uzun gecikmelerle dikkat çekmektedir.
- Ayrıca sadece salgının başlangıcını belirlemek için tasarlanmış OutbreakP yöntemi, işlenmiş veriler için 1 ya da 2 hafta gecikme ile salgının başlangıcını belirlemektedir.

**Simülasyon çalışmasından elde edilen sonuçlar:** Simülasyon çalışmalarında 4 farklı senaryo dikkate alınmıştır. Bunlar 2013-2014, 2014-2015, 2015-2016 ve 2016-2017 grip mevsimlerine ilişkin haftalık verilerdir. Belirtilen grip mevsimlerindeki veriler simüle edilerek, CDC kurumunun ilgili mevsimlerdeki gerçek verilerin özelliklerine yaklaştırılmıştır. Simülasyonlarda dikkate alınan yöntemlerin salgın olmadığı dönemlerdeki performansı ATFS=13 olacak biçimde eşitlenmiştir. Daha sonra simüle edilen verilere eklenen çeşitli yapay salgın koşulları altında, dört farklı senaryo için 10000 tekrarlı simülasyon çalışmalarında FMO ve ATFOS performansları elde edilmiştir. Ayrıca yöntemlerin dayandığı varsayımlar dikkate alabilmek amacıyla;

simülasyonlarda bazı yöntemlerde ham veriler, bazı yöntemlerde işlenmiş veriler kullanılmıştır. Seçilen simülasyon parametreleri ve bu simülasyon çalışmasında algoritmalar için belirlenen parametre değerleri çerçevesinde yöntemlere ilişkin performans sonuçlarına hakkında aşağıdaki yorumlar yapılabilmektedir.

- ATFOS ölçütü açısından tüm senaryolar ve tüm salgın durumları için genel olarak bu çalışmada önerilen UCengiz yönteminin diğer yöntemlere göre dikkat çekici üstünlüğü görülmektedir. UCengiz yöntemi diğer yöntemlere göre çok kısa sürede salgını belirlemektedir.
- ATFOS ölçütü açısından bu çalışmada önerilen UCengiz yönteminden sonra Serfling regresyon yöntemi, OutbreakP yöntemi ve Rossi vd. (1999) Cusum grafiği yöntemlerinin genel olarak diğer yöntemlere göre daha iyi performans sergilemektedir.
- Glrnb, Ears C1, Ears C2 ve Ears C3 yöntemlerinin genel olarak ATFOS performansı açısından en kötü olma eğiliminde olduğu söylenebilir. Başka bir ifadeyle, bu yöntemler salgının belirlenmesini diğer yöntemlere göre daha geç gerçekleştirmektedir.
- FMO ölçütü açısından tüm senaryolar ve tüm salgın durumları için genel olarak Serfling regresyon yönteminin en iyi performansa sahip olduğu görülmektedir. Başka bir ifadeyle, Serfling regresyon yöntemini büyük olasılıkla salgınları tespit edebilmektedir.
- FMO ölçütü açısından bu çalışmada önerilen UCengiz yöntemi kısa süreli ve küçük salgın büyüklüklerinde diğer yöntemlerden üstün performansa sahip olma eğiliminde olduğu söylenebilir.
- Ayrıca FMO ölçütü açısından Farrington, FarringtonFlexible, Glrnb, Rossi vd. (1999) Cusum grafiği, OutbreakP yöntemleri genel olarak çeşitli senaryolar ve çeşitli salgın durumları için iyi bir performans sergilediği söylenebilir.
- Ears C1, Ears C2, Ears C3 ve Glrnb yöntemleri genel olarak FMO performansı açısından en kötü olma eğiliminde olduğu söylenebilir. Başka bir ifadeyle, bu yöntemler diğer yöntemlere göre daha küçük olasılıkla salgınları belirleme eğiliminde olduğu gözükmektedir.



- Her iki ölçüt performansı birlikte dikkate alındığında; UCengiz, Serfling regresyon, Rossi vd. (1999) Cusum grafiđi, Farrington ve FarringtonFlexible yöntemleri, çeşitli senaryolar ve salgın durumları için diđer yöntemlere göre daha iyi performanslarıyla ön plana eğilimindedir.

Bu simülasyon çalışması sonucunda CDC kurumunun elde edilen haftalık grip verilerinin izlenmesi için; kısa sürede ve yüksek bir olasılıkla salgınların belirlenmesi açısından, UCengiz, Serfling regresyon, Rossi vd. (1999) Cusum grafiđi, Farrington ve FarringtonFlexible yöntemleri bu simülasyon çalışması kısıtlamaları altında tercih edilebilir yöntemler olarak gözükmektedir. Ayrıca bu tez çalışmasında önerilen UCengiz yönteminin;

- kısa süreli salgın ve küçük salgın büyüklüklerindeki dikkat çeken üstün performansı,
- çođu salgın durumu için iyi olan ATFOS performansı,
- basit yapısı nedeniyle halk sađlığı sistemlerinde çok sayıda hastalığa ilişkin zaman serilerinin otomatik izlenmesine elverişli olması,

özellikleri dikkate alındığında, halk sađlığı izlemesine olumlu katkılar yapacağı izlenimi oluşturmaktadır. Bu tez çalışmasında dikkate alınan yöntemlerin performansları hakkında daha doğru değerlendirme yapabilmek için; bu yöntemler çok sayıda gerçek izleme durumuna uygulanmalı ve deđişik simülasyon koşulları altında denenmelidirler. Ayrıca yöntemleri uygularken seçilen parametre deđerlerinin, yöntemlerin performanslarını nasıl etkilendiđinin araştırılması gerekir.

## KAYNAKLAR

- Amorós, R., Conesa, D., Martínez-Beneito, M. A., & López-Quilez, A. (2015). Statistical methods for detecting the onset of influenza outbreaks: a review. *REVSTAT–Statistical Journal*, 13(1), 41-62.
- Buckeridge, D. L., Burkom, H., Campbell, M., Hogan, W. R., & Moore, A. W. (2005). Algorithms for rapid outbreak detection: a research synthesis. *Journal of biomedical informatics*, 38(2), 99-113.
- Buckeridge, D. L. (2007). Outbreak detection through automated surveillance: a review of the determinants of detection. *Journal of biomedical informatics*, 40(4), 370-379.
- Buckeridge, D. L., Thompson, M. W., Babin, S., & Sikes, M. L. (2007). Evaluating automated surveillance systems. *Disease surveillance: A public health informatics approach*. Hoboken, NJ: Wiley, 399-424.
- Burkom, H. (2007), Alerting algorithms for biosurveillance, 143-192. Lombardo J.S. & Buckeridge D.L. (Eds). *Disease surveillance: a public health approach*. John Wiley and Sons, New Jersey.
- Burkom, H. S., Murphy, S. P., & Shmueli, G. (2007). Automated time series forecasting for biosurveillance. *Statistics in medicine*, 26(22), 4202-4218.
- Cardinal, M., Roy, R., and Lambert, J. (1999), “On the application of integer-valued time series models for the analysis of disease incidence,” *Statistics in Medicine*, 18, 2025–2039.
- CDC. 2008 Accessed at <http://www.cdc.gov/BioSense/publichealth.htm> on October 11, 2008.
- Chen, R. (1978), “A surveillance system for congenital malformations,” *Journal of the American Statistical Association*, 73, 323–327.

- Closas, P., Coma, E., & Méndez, L. (2012). Sequential detection of influenza epidemics by the Kolmogorov-Smirnov test. *BMC medical informatics and decision making*, 12(1), 112.
- Farrington, C. P., Andrews, N. J., Beale, A. D., & Catchpole, M. A. (1996). A statistical algorithm for the early detection of outbreaks of infectious disease. *Journal of the Royal Statistical Society. Series A (Statistics in Society)*, 547-563.
- Fricke Jr, R. D., Hegler, B. L., & Dunfee, D. A. (2008). Comparing syndromic surveillance detection methods: EARS' versus a CUSUM-based methodology.
- Fricke, R. D. (2012, 2013). Introduction to statistical methods for biosurveillance: with an emphasis on syndromic surveillance. Cambridge University Press.
- Frisén, M. and de Maré, J. (1991). Optimal Surveillance, *Biometrika*78: 271–280.
- Frisén, M. (1992). Evaluations of methods for statistical surveillance. *Statistics in Medicine*, 11(11), 1489-1502.
- Frisén, M. (2003). Statistical Surveillance. Optimality and Methods, *International Statistical Review*71: 403–434.
- Frisén, M. (2007). Properties and Use of the Shewhart Method and Followers, *Sequential Analysis*26: 171–193.
- Frisén, M., & Andersson, E. (2009). Semiparametric surveillance of monotonic changes. *Sequential Analysis*, 28(4), 434-454.
- Frisén, M., Andersson, E., & Schiöler, L. (2009). Robust outbreak surveillance of epidemics in Sweden. *Statistics in Medicine*, 28(3), 476-493.
- Frisén, M., Andersson, E., and Pettersson, K. (2010), “Semiparametric estimation of outbreak regression,” *Statistics*, 44, 107–117.

- Heisterkamp, S. H., Dekkers, A. L. M. and Heijne, J. C. M. (2006). Automated detection of infectious disease outbreaks: hierarchical time series models. *Statistics in Medicine* 25, 4179–4196.
- Held, L., Hofmann, M., Höhle, M., and Schmid, V. (2006), “A two-component model for counts of infectious diseases,” *Biostatistics*, 7, 422–437.
- Höhle, M. (2007).: An R package for the monitoring of infectious diseases. *Computational Statistics*, 22(4), 571-582.
- Höhle, M., & Paul, M. (2008). Count data regression charts for the monitoring of surveillance time series. *Computational Statistics & Data Analysis*, 52(9), 4357-4368.
- Höhle, M., & Mazick, A. (2010). Aberration detection in R illustrated by Danish mortality monitoring. *Biosurveillance: Methods and Case Studies*, 215-238.
- Ismail, N. A., Pettitt, A. N., and Webster, R. A. (2003), “‘Online’ monitoring and retrospective analysis of hospital outcomes based on a scan statistic,” *Statistics in Medicine*, 22, 2861–2876.
- Ivanov, O., Gesteland, P. H., Hogan, W., Mundorff, M. B., & Wagner, M. M. (2003). Detection of pediatric respiratory and gastrointestinal outbreaks from free-text chief complaints. In *AMIA Annual Symposium Proceedings* (Vol. 2003, p. 318). American Medical Informatics Association.
- Kass-Hout, T. A., Xu, Z., McMurray, P., Park, S., Buckeridge, D. L., Brownstein, J. S., ... & Groseclose, S. L. (2012). Application of change point analysis to daily influenza-like illness emergency department visits. *Journal of the American Medical Informatics Association*, 19(6), 1075-1081.
- Kenett R, Pollak M (1983) On sequential detection of a shift in the probability of a rare event. *J Am Stat Assoc* 78(382):389–395.

- Kulldorff, M. (2001). Prospective time periodic geographical disease surveillance using a scan statistic. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 164(1), 61-72.
- Le Strat, Y., & Carrat, F. (1999). Monitoring epidemiologic surveillance data using hidden Markov models. *Statistics in medicine*, 18(24), 3463-3478.
- Manitz, J., & Höhle, M. (2013). Bayesian outbreak detection algorithm for monitoring reported cases of campylobacteriosis in Germany. *Biometrical Journal*, 55(4), 509-526.
- Miller, B., Kassenborg, H., Dunsmuir, W., Griffith, J., Hadidi, M., Nordin, J., and Danila, R. (2003), "Syndromic surveillance for influenzalike illness in ambulatory care network," *Emerging Infectious Diseases*, 10, 1806–1811.
- Montgomery, D. C. (2009). *Statistical quality control: a modern introduction*. John Wiley & Sons, Aisa, 734s.
- Muscatello, D. (2004), "An Adjusted Cumulative Sum for Count Data With Day-of-Week Effects: Application to Influenza-Like Illness," presentation at Syndromic Surveillance Conference, Boston, MA.
- Naus, J. and Wallenstein, S. (2006), "Temporal surveillance using scan statistics," *Statistics in Medicine*, 25, 311–324.
- Noufaily, A., Enki, D. G., Farrington, P., Garthwaite, P., Andrews, N., & Charlett, A. (2013). An improved algorithm for outbreak detection in multiple surveillance systems. *Statistics in medicine*, 32(7), 1206-1222.
- Page, E. S. (1954). Continuous inspection schemes. *Biometrika*, 41(1/2), 100-115.
- Percival, D. B., & Walden, A. T. (2006). *Wavelet methods for time series analysis* (Vol. 4). Cambridge university press.

- Pierce D, Schafer D (1986) Residuals in generalized linear models. *J Am Stat Assoc* 81(396):977–986.
- Reis, B. Y. and Mandl, K. D. (2003), “Time series modeling for syndromic surveillance,” *BMC Medical Informatics and Decision Making*, 3, No. 2.
- Riebler A (2004) Empirischer Vergleich von statistischen Methoden zur Ausbruchserkennung bei Surveillance Daten. Bachelor’s thesis, Department of Statistics, University of Munich
- Roberts, S. W. (1966). A Comparison of Some Control Chart Procedures, *Technometrics*8: 411–430.
- Rossi G, Lampugnani L, Marchi M (1999) An approximate CUSUM procedure for surveillance of health events. *Stat Med* 18:2111–2122.
- Salmon, M., Schumacher, D., & Höhle, M. (2016). Monitoring count time series in R: Aberration detection in public health surveillance.
- Serfling, R. (1963), “Methods for current statistical analysis of excess pneumonia-influenza deaths,” *Public Health Reports*, 78, 494–506.
- Shiryayev, A. N. (1963). On Optimum Methods in Quickest Detection Problems, *Theory of Probability & Its Applications*8: 22–46.
- Shmueli, G. (2013). Wavelet-Based Monitoring for Biosurveillance. *Axioms*, 2(3), 345-370.
- Shmueli, G., & Burkom, H. (2010). Statistical challenges facing early outbreak detection in biosurveillance. *Technometrics*, 52(1), 39-51.
- Sonesson, C., & Bock, D. (2003). A review and discussion of prospective statistical surveillance in public health. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 166(1), 5-21.

- Thacker, S.B (2000) Historical Development 1-16. Teutsch, S. M., & Churchill, R. E. (Eds.). Principles and practice of public health surveillance. Oxford University Press, USA.
- Unkel, S., Farrington, C., Garthwaite, P. H., Robertson, C., & Andrews, N. (2012). Statistical methods for the prospective detection of infectious disease outbreaks: a review. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 175(1), 49-82.
- Wagner, A. (2010), "Extending Scottish Exception Reporting Systems Spatially and Temporally," Ph.D. thesis, University of Strathclyde.
- Wallenstein, S. and Naus, J. (2004), "Scan statistics for temporal surveillance for biologic terrorism," *Morbidity and Mortality Weekly Report (Supplement)*, 53, 74-78.
- Yahav, I., Lotze, T., Shumeli, G. (2011) Algorithm Combination for Improved Performance in Biosurveillance 173-189. Zeng, D., Chen, H., Castillo-Chavez, C., Lober, W. B., & Thurmond, M. (Eds.) *Infectious disease informatics and biosurveillance* (Vol. 27). Springer Science & Business Media, Arizona.
- Zeng, D., Chen, H., Castillo-Chavez, C., Lober, W. B., & Thurmond, M. (Eds.). (2010). *Infectious disease informatics and biosurveillance* (Vol. 27). Springer Science & Business Media.
- Zhang, J., Tsui, F.-C., Wagner, M. M., and Hogan, W. (2003), "Detection of outbreaks from time series data using wavelet transform," *AMIA Annual Symposium*, Omnipress: Madison, Wisconsin, 748-552.

## EKLER

### Ek A. 2012-2016 Yıllarına İlişkin CDC Grip Verileri

Çizelge Ek A.

Hafta	2012-2013	2013-2014	2014-2015	2015-2016	Hafta	2012-2013	2013-2014	2014-2015	2015-2016
40	9395	9677	10374	10205	14	12273	14637	13988	16711
41	10158	10420	11297	10697	15	10818	14206	12311	16249
42	10585	10658	12126	11582	16	9188	11723	11021	15447
43	10913	12036	12474	11159	17	8589	10814	11311	13826
44	11061	12502	12490	12479	18	7986	9648	10946	11847
45	12674	13498	14102	12680	19	7704	9775	9891	10680
46	13623	13927	14109	13492	20	7317	9033	9289	9670
47	14650	15168	16967	12149	21	6589	8797	8200	8624
48	18521	13975	17911	14613	22	6190	7658	7227	7861
49	24877	18920	23247	14992	23	5638	7229	7034	7089
50	29432	19398	29627	15400	24	5539	6783	6108	6602
51	34332	24697	37800	14805	25	5474	6369	6170	5900
52	36588	26988	38951	15666	26	4915	6038	5455	5576
1	34103	28654	31870	15648	27	4506	5352	5329	5045
2	39896	28147	31613	15507	28	4510	5295	5082	4991
3	39441	28370	33630	15565	29	4186	5083	4908	4932
4	32426	25502	30943	18124	30	4233	4765	4798	4799
5	29808	23484	28563	18927	31	3886	4942	4982	4457
6	24987	21085	25577	21732	32	4075	4744	4971	4754
7	22321	18195	21469	24618	33	4168	4980	5409	5148
8	20519	17234	19512	26397	34	4317	5305	5822	6012
9	19231	15914	18788	28228	35	5053	5799	6240	6147
10	18829	15044	18143	30319	36	5626	6050	6281	6304
11	17392	14024	16988	24821	37	6374	8130	7420	7132
12	15254	14926	16425	21874	38	7110	8888	8099	7762
13	13714	14661	15253	19537	39	7963	8874	8121	6959



## Ek B. CDC Grip Verileri İçin Serfling Regresyon Tahminleri ve Eşik Değerleri

Çizelge Ek B. 2015-2016 grip mevsimine ilişkin Serfling regresyondan elde edilen sonuçlar

Hafta	Tahmin	Eşik Değeri	Fark	Hafta	Tahmin	Eşik Değeri	Fark
40	9749	11856	-1651	14	14021	17275	-564
41	10616	12723	-2026	15	12855	16128	<b>121</b>
42	11511	13618	-2036	16	11719	14962	<b>485</b>
43	12456	14564	-3405	17	10668	13827	-1
44	13474	15581	-3102	18	9746	12775	-928
45	14579	16686	-4006	19	8985	11853	-1173
46	15772	17879	-4387	20	8395	11092	-1422
47	17032	19139	-6990	21	7962	10502	-1878
48	18318	20425	-5812	22	7654	10069	-2208
49	19571	21678	-6686	23	7425	9761	-2672
50	20722	22829	-7429	24	7228	9533	-2931
51	21703	23810	-9005	25	7022	9335	-3435
52	22459	24567	-8901	26	6784	9129	-3553
1	22955	25062	-9414	27	6511	8891	-3846
2	23178	25285	-9778	28	6224	8619	-3628
3	23144	25251	-9686	29	5960	8331	-3399
4	22887	24994	-6870	30	5766	8067	-3268
5	22456	24563	-5636	31	5687	7873	-3416
6	21898	24005	-2273	32	5762	7795	-3041
7	21257	23364	<b>1254</b>	33	6009	7869	-2721
8	20562	22669	<b>3728</b>	34	6428	8116	-2104
9	19823	21930	<b>6298</b>	35	6999	8535	-2388
10	19037	21144	<b>9175</b>	36	7690	9106	-2802
11	18189	20297	<b>4524</b>	37	8463	9797	-2665
12	17266	19373	<b>2501</b>	38	9284	10570	-2808
13	16257	18364	<b>1173</b>	39	6511	11391	-4432

**Ek C. CDC Grip Verileri İçin Farrington ve FarringtonFlexible Yöntemlerine İlişkin Sonuçlar**

**Çizelge Ek C. 2015-2016 grip mevsimine ilişkin Farrington ve FarringtonFlexible yöntemlerinden elde edilen sonuçlar**

Hafta	Farrington Eşik Değerleri	FarringtonFlexible Eşik Değerleri	Hafta	Farrington Eşik Değerleri	FarringtonFlexible Eşik Değerleri
40	13182.127	16430	14	18820.125	21595
41	13823.874	17451	15	17752.457	20450
42	15367.570	17789	16	19276.703	19160
43	16178.638	19080	17	16899.155	17721
44	17717.600	19846	18	13114.086	16009
45	16794.480	22152	19	12369.792	15066
46	18394.292	23105	20	14147.923	14074
47	23185.371	24575	21	13330.626	13548
48	30490.236	25810	22	9919.538	12317
49	39802.567	29820	23	8814.423	11491
50	46745.387	33249	24	8145.114	10760
51	45796.737	41799	25	7605.159	10201
52	46108.699	44034	26	7088.917	10138
1	45561.554	44724	27	6603.888	9449
2	43048.526	44263	28	6755.156	8893
3	40368.457	41918	29	5949.239	8742
4	40883.866	39550	30	5814.633	8907
5	40008.980	38659	31	5825.628	8620
6	35237.043	34614	32	6744.243	8610
7	30612.736	31792	33	6572.022	9032
8	26117.619	29525	34	7116.818	9502
9	23616.790	27677	35	9786.128	10272
10	22032.233	26919	36	9310.049	11089
11	20557.160	25877	37	10246.725	12121
12	19414.884	24766	38	11561.493	14596
13	18728.920	22925	39	12700.426	16238

## Ek D. CDC Grip Verileri İçin Boda Yöntemlerine İlişkin Sonuçlar

Çizelge Ek D. 2015-2016 grip mevsimine ilişkin Boda yöntemlerinden elde edilen sonuçlar

Hafta	Boda( $\alpha = 0.05$ ) Eşik Değerleri	Boda( $\alpha = 0.10$ ) Eşik Değerleri	Hafta	Boda( $\alpha = 0.05$ ) Eşik Değerleri	Boda( $\alpha = 0.10$ ) Eşik Değerleri
40	14795	13487	14	19899	20160
41	15322	15230	15	19071	18210
42	16065	15848	16	17147	15838
43	17472	16805	17	17088	15170
44	18297	16912	18	15524	14228
45	19782	18041	19	14566	13190
46	21294	18678	20	13948	12834
47	20853	20078	21	12922	11276
48	23315	21615	22	11621	10042
49	30232	25515	23	9942	9070
50	32451	30073	24	9952	9380
51	37519	33920	25	9823	8746
52	35329	35730	26	8267	8206
1	34625	30268	27	7603	7649
2	35048	30273	28	7655	7140
3	33111	30963	29	8199	6986
4	27722	25759	30	7490	6785
5	25576	23556	31	7218	5921
6	22287	20857	32	7083	6915
7	19062	18320	33	7502	6969
8	19565	17226	34	8053	7405
9	19587	19245	35	9314	8387
10	21788	19459	36	9325	8881
11	20372	18530	37	12075	10736
12	22755	20013	38	12048	11669
13	23935	21389	39	12887	12081

## Ek E. CDC Grip Verileri İçin Holt-Winters Yöntemine İlişkin Sonuçlar

Çizelge Ek E. 2015-2016 grip mevsimine ilişkin Holt-Winters yönteminden elde edilen sonuçlar

Hafta	Öngörü Değerleri	Ln(İzleme)	Eşik Değerleri	Fark	Hafta	Öngörü Değerleri	Ln(İzleme)	Eşik Değerleri	Fark
40	9.2158	9.2306	9.3126	-0.0820	14	9.4892	9.7238	9.8491	-0.1253
41	9.3119	9.2777	9.4303	-0.1525	15	9.3457	9.6958	9.7120	-0.0162
42	9.3530	9.3572	9.4894	-0.1322	16	9.2009	9.6452	9.5734	<b>0.0718</b>
43	9.4362	9.3200	9.5886	-0.2686	17	9.1368	9.5343	9.5155	<b>0.0189</b>
44	9.3834	9.4318	9.5503	-0.1185	18	9.0230	9.3798	9.4077	-0.0279
45	9.4561	9.4478	9.6364	-0.1886	19	8.9916	9.2761	9.3823	-0.1062
46	9.4529	9.5099	9.6456	-0.1357	20	8.9319	9.1768	9.3285	-0.1517
47	9.6428	9.4050	9.8471	-0.4421	21	8.8422	9.0623	9.2446	-0.1823
48	9.6130	9.5897	9.8283	-0.2386	22	8.7465	8.9697	9.1546	-0.1849
49	9.8850	9.6153	10.1108	-0.4955	23	8.7007	8.8663	9.1144	-0.2481
50	10.1859	9.6421	10.4217	-0.7796	24	8.6310	8.7951	9.0502	-0.2551
51	10.1033	9.6027	10.3487	-0.7460	25	8.6445	8.6827	9.0692	-0.3865
52	10.1839	9.6592	10.4385	-0.7792	26	8.5559	8.6262	8.9860	-0.3598
1	10.2336	9.6581	10.4971	-0.8391	27	8.4878	8.5262	8.9233	-0.3971
2	10.2510	9.6490	10.5231	-0.8741	28	8.4636	8.5154	8.9043	-0.3890
3	10.2875	9.6528	10.5681	-0.9153	29	8.4191	8.5035	8.8651	-0.3616
4	10.1786	9.8050	10.4673	-0.6623	30	8.3996	8.4762	8.8507	-0.3746
5	10.0874	9.8483	10.3839	-0.5356	31	8.3904	8.4022	8.8466	-0.4444
6	9.9734	9.9865	10.2776	-0.2911	32	8.3792	8.4667	8.8404	-0.3737
7	9.8148	10.1112	10.1265	-0.0153	33	8.4350	8.5464	8.9012	-0.3548
8	9.7510	10.1810	10.0701	<b>0.1109</b>	34	8.4808	8.7015	8.9519	-0.2504
9	9.6914	10.2481	10.0176	<b>0.2304</b>	35	8.5878	8.7237	9.0638	-0.3401
10	9.6338	10.3195	9.9670	<b>0.3525</b>	36	8.6747	8.7489	9.1555	-0.4066
11	9.5657	10.1194	9.9058	<b>0.2137</b>	37	8.8870	8.8723	9.3727	-0.5003
12	9.5872	9.9931	9.9341	<b>0.0590</b>	38	8.9801	8.9570	9.4705	-0.5135
13	9.5671	9.8801	9.9205	-0.0405	39	9.0355	8.8478	9.5305	-0.6827

## Ek F. CDC Grip Verileri İçin Arima Modellemesine İlişkin Sonuçlar

Çizelge Ek F. 2015-2016 grip mevsimine ilişkin Arima modellemesinden elde edilen sonuçlar

Hafta	Öngörü Değerleri	Ln(İzleme)	Eşik Değerleri	Fark	Hafta	Öngörü Değerleri	Ln(İzleme)	Eşik Değerleri	Fark
40	9.1917	9.2306	9.2834	-0.0528	14	9.5191	9.7238	9.8728	-0.1489
41	9.2708	9.2777	9.3844	-0.1067	15	9.3932	9.6958	9.7532	-0.0574
42	9.3157	9.3572	9.4475	-0.0903	16	9.2692	9.6452	9.6354	<b>0.0098</b>
43	9.3758	9.3200	9.5237	-0.2037	17	9.2286	9.5343	9.6008	-0.0665
44	9.3488	9.4318	9.5111	-0.0793	18	9.1377	9.3798	9.5159	-0.1361
45	9.4608	9.4478	9.6364	-0.1886	19	9.1141	9.2761	9.4981	-0.2220
46	9.4915	9.5099	9.6795	-0.1696	20	9.0497	9.1768	9.4396	-0.2628
47	9.6492	9.4050	9.8488	-0.4438	21	8.9644	9.0623	9.3600	-0.2977
48	9.7226	9.5897	9.9331	-0.3434	22	8.8554	8.9697	9.2566	-0.2869
49	9.9870	9.6153	10.2079	-0.5926	23	8.7961	8.8663	9.2028	-0.3365
50	10.2540	9.6421	10.4848	-0.8427	24	8.7219	8.7951	9.1342	-0.3391
51	10.1585	9.6027	10.3989	-0.7961	25	8.7004	8.6827	9.1181	-0.4354
52	10.2378	9.6592	10.4873	-0.8281	26	8.6058	8.6262	9.0288	-0.4026
1	10.2882	9.6581	10.5466	-0.8885	27	8.5289	8.5262	8.9572	-0.4310
2	10.3056	9.6490	10.5724	-0.9234	28	8.5099	8.5154	8.9434	-0.4280
3	10.3351	9.6528	10.6103	-0.9575	29	8.4599	8.5035	8.8985	-0.3950
4	10.2922	9.8050	10.5754	-0.7704	30	8.4346	8.4762	8.8783	-0.4021
5	10.2100	9.8483	10.5011	-0.6527	31	8.4309	8.4022	8.8796	-0.4774
6	10.1017	9.9865	10.4004	-0.4138	32	8.4324	8.4667	8.8861	-0.4194
7	9.9335	10.1112	10.2395	-0.1283	33	8.4843	8.5464	8.9430	-0.3966
8	9.8555	10.1810	10.1688	<b>0.0122</b>	34	8.5417	8.7015	9.0052	-0.3037
9	9.7948	10.2481	10.1152	<b>0.1329</b>	35	8.6471	8.7237	9.1154	-0.3917
10	9.7574	10.3195	10.0847	<b>0.2348</b>	36	8.6993	8.7489	9.1724	-0.4234
11	9.6900	10.1194	10.0241	<b>0.0954</b>	37	8.8950	8.8723	9.3728	-0.5005
12	9.6515	9.9931	9.9923	<b>0.0008</b>	38	8.9905	8.9570	9.4729	-0.5159
13	9.5854	9.8801	9.9327	-0.0527	39	9.0287	8.8478	9.5158	-0.6680

## Ek G. 2012-2016 Yıllarına İlişkin Serfling Regresyon Artıkları

**Çizelge Ek G.**

Hafta	2012- 2013	2013- 2014	2014- 2015	2015- 2016	Hafta	2012- 2013	2013- 2014	2014- 2015	2015- 2016
40	*	95	93	99	14	83	96	84	69
41	86	85	87	83	15	83	94	80	94
42	81	80	85	86	16	82	72	85	89
43	79	90	80	72	17	91	88	100	82
44	77	80	76	90	18	90	85	93	77
45	91	85	90	77	19	92	96	85	84
46	83	77	74	83	20	89	86	88	84
47	83	87	103	61	21	84	89	81	82
48	112	61	82	98	22	86	79	81	83
49	138	123	126	77	23	84	85	87	82
50	120	80	139	80	24	87	84	79	84
51	127	130	159	70	25	88	84	89	81
52	102	102	91	87	26	83	85	81	85
1	58	99	10	81	27	85	82	88	84
2	143	79	81	83	28	89	89	87	89
3	83	90	107	88	29	86	87	88	89
4	19	60	62	114	30	89	86	88	88
5	65	71	68	98	31	84	90	90	85
6	44	68	61	120	32	88	85	87	90
7	66	64	52	121	33	86	88	90	90
8	75	83	73	111	34	85	87	88	93
9	81	80	86	113	35	89	87	86	84
10	90	85	87	115	36	86	83	81	82
11	81	85	84	39	37	86	100	91	88
12	75	105	89	66	38	86	86	85	85
13	80	94	86	73	39	87	78	78	70

**Ek H. CDC Grip Verileri İçin Holt-Winters Düzleştirme Yöntemine İlişkin Tahmin Hataları**

**Çizelge Ek H. 2015-2016 grip mevsimine ilişkin Holt-Winters düzleştirmeden elde edilen tahmin hataları**

Hafta	Tahmin	Tahmin Hatası	Standartlaştırılmış Hata	Hafta	Tahmin	Tahmin Hatası	Standartlaştırılmış Hata
40	9802.0	403.0	0.4859	14	11739.2	4971.80	5.9955
41	10488.6	208.4	0.2513	15	10277.8	5971.16	7.2006
42	10668.5	913.5	1.1015	16	8643.5	6803.50	8.2043
43	11999.6	-840.6	-1.0137	17	8031.9	5794.13	6.9871
44	12427.1	51.9	0.0625	18	7403.1	4443.90	5.3589
45	13386.1	-706.1	-0.8515	19	7098.2	3581.81	4.3193
46	13777.6	-285.6	-0.3444	20	6699.2	2970.77	3.5824
47	14979.7	-2830.7	-3.4136	21	5962.2	2661.80	3.2099
48	13750.3	862.7	1.0404	22	5600.8	2260.19	2.7256
49	18664.7	-3672.7	-4.4289	23	5148.7	1940.33	2.3398
50	19114.3	-3714.3	-4.4791	24	5158.5	1443.49	1.7407
51	24391.7	-9586.7	-11.5605	25	5146.8	753.20	0.9083
52	24551.1	-8885.1	-10.7145	26	4588.4	987.59	1.1909
1	24531.1	-8883.1	-10.7121	27	4178.3	866.72	1.0452
2	24514.3	-9007.3	-10.8618	28	4181.2	809.84	0.9766
3	24497.0	-8932.0	-10.7710	29	3856.0	1075.96	1.2975
4	24482.1	-6358.1	-7.6672	30	3901.9	897.09	1.0818
5	23072.7	-4145.7	-4.9993	31	3567.2	889.81	1.0730
6	20656.0	1076.0	1.2975	32	3804.9	949.08	1.1445
7	17750.6	6867.4	8.2813	33	3972.9	1175.07	1.4170
8	16771.2	9625.8	11.6077	34	4192.1	1819.93	2.1947
9	15433.4	12794.6	15.4289	35	4990.4	1156.58	1.3947
10	14551.0	15768.0	19.0145	36	5630.6	673.41	0.8121
11	13509.0	11312.0	13.6411	37	6446.2	685.76	0.8270
12	14375.8	7498.2	9.0420	38	7216.7	545.34	0.6576
13	14083.9	5453.1	6.5759	39	8062.6	1103.58	1.3308

**Ek I. 2012-2013, 2013-2014 ve 2014-2015 Yıllarına İlişkin Mevsimsellik Düzeltmesi İndeksleri ve Mevsimsellik Düzeltmesi Yapılmış CDC grip verileri**

**Çizelge Ek I.1. 2012-2013, 2013-2014 ve 2014-2015 yıllarına ilişkin mevsimsellik düzeltmesi indeksleri**

Hafta	2012-2013	2013-2014	2014-2015	M.İ	D.M.İ	Hafta	2012-2013	2013-2014	2014-2015	M.İ	D.M.İ
40	*	75.657	79.796	77.727	76.937	14	92.4931	115.865	*	104.179	103.120
41	*	81.612	86.628	84.120	83.265	15	81.5115	102.445	*	91.978	91.044
42	*	83.605	92.706	88.156	87.260	16	69.2188	92.936	*	81.077	80.254
43	*	94.541	95.088	94.815	93.851	17	64.6892	85.792	*	75.240	74.476
44	*	85.985	94.937	90.461	89.542	18	60.1285	76.595	*	68.362	67.667
45	*	93.096	106.887	99.992	98.976	19	57.9987	77.653	*	67.826	67.137
46	*	95.004	106.645	100.825	99.800	20	55.0933	71.797	*	63.445	62.801
47	*	114.322	132.007	123.164	121.913	21	49.6252	69.952	*	59.788	59.181
48	*	110.248	134.662	122.455	121.211	22	46.6459	60.911	*	53.778	53.232
49	*	149.334	174.331	161.833	160.188	23	42.5316	57.500	*	50.016	49.507
50	*	206.715	221.617	214.166	211.990	24	41.8421	53.941	*	47.891	47.405
51	*	185.720	205.028	195.374	193.389	25	41.4139	50.627	*	46.021	45.553
52	*	201.877	223.455	212.666	210.505	26	37.2447	47.966	*	42.606	42.173
1	*	212.659	236.669	224.664	222.381	27	34.2022	42.481	*	38.342	37.952
2	*	220.312	234.209	227.261	224.952	28	34.2902	41.985	*	38.138	37.750
3	*	224.302	248.576	236.439	234.037	29	31.8803	40.254	*	36.067	35.701
4	*	201.654	228.192	214.923	212.739	30	32.2985	37.678	*	34.988	34.633
5	*	185.715	210.164	197.940	195.929	31	29.7160	39.005	*	34.361	34.011
6	*	166.755	187.772	177.264	175.463	32	31.2370	37.361	*	34.299	33.950
7	*	143.904	157.264	150.584	149.054	33	32.0316	39.124	*	35.578	35.216
8	*	136.305	142.613	139.459	138.043	34	33.2623	41.569	*	37.416	37.035
9	*	125.863	137.021	131.442	130.107	35	39.0335	45.314	*	42.174	41.745
10	*	118.978	132.029	125.504	124.229	36	43.5722	47.136	*	45.354	44.893
11	*	112.353	123.360	117.857	116.659	37	49.4934	63.143	*	56.318	55.746
12	*	118.023	119.025	118.524	117.320	38	55.3465	68.807	*	62.077	61.446
13	*	115.965	110.311	113.138	111.988	39	62.1292	68.476	*	65.303	64.639



**Ek I. (devam)****Çizelge Ek I.2. 2012-2013, 2013-2014 ve 2014-2015 yıllarına ilişkin mevsimsellik düzeltmesi yapılmış veriler**

Hafta	2012-2013	2013-2014	2014-2015	2015-2016	Hafta	2012-2013	2013-2014	2014-2015	2015-2016
40	12211.3	12577.8	13483.8	13264.1	14	11901.6	14194.1	13564.7	16205.3
41	12199.6	12514.3	13567.5	12846.9	15	11882.2	14204.0	13522.1	17847.5
42	12130.4	12214.1	13896.4	13273.0	16	11448.7	14607.4	13732.7	19247.7
43	11628.0	12824.5	13291.2	11890.1	17	11532.6	14520.1	15187.4	18564.4
44	12352.8	12211.2	13948.7	13936.4	18	11801.8	14258.0	15338.8	17507.7
45	12805.2	11948.9	14247.9	12811.2	19	11475.1	14559.8	14732.6	15907.8
46	13650.3	12082.8	14137.2	13519.0	20	11651.2	14383.6	14791.3	15397.9
47	12016.8	11893.8	14364.1	9965.3	21	11133.6	14864.6	13855.8	14572.2
48	15280.0	11529.5	14776.7	12055.8	22	11628.3	14386.1	13576.4	14767.4
49	14542.2	11811.1	14512.3	9359.0	23	11388.2	14601.8	14208.0	14319.1
50	13883.6	12349.1	13975.6	7264.5	24	11684.5	14308.7	12884.8	13926.9
51	13718.2	12157.7	14208.1	7655.5	25	12016.7	13981.5	13544.6	12951.9
52	13465.2	12137.2	14260.5	7442.1	26	11654.5	14317.3	12934.9	13221.8
1	13261.8	12099.7	14331.2	7036.6	27	11872.9	14102.0	14041.4	13293.1
2	13439.6	12389.3	14053.2	6893.5	28	11947.0	14026.4	13462.2	13221.1
3	13035.2	12122.0	14369.5	6650.7	29	11725.2	14237.8	13747.6	13814.8
4	15242.1	11987.4	14545.0	8519.4	30	12222.5	13758.6	13853.8	13856.7
5	15213.7	11986.0	14578.3	9660.1	31	11425.6	14530.4	14648.0	13104.4
6	15267.3	12016.8	14576.9	12385.5	32	12002.8	13973.3	14641.9	14002.7
7	14975.1	12207.0	14403.5	16516.1	33	11835.4	14141.1	15359.3	14618.2
8	14864.3	12484.6	14134.8	19122.4	34	11656.4	14324.1	15720.1	16233.1
9	14780.9	12231.5	14440.5	21696.0	35	12104.4	13891.4	14947.8	14725.1
10	15156.7	12109.9	14604.5	24405.8	36	12532.0	13476.4	13991.0	14042.2
11	14908.4	12178.5	14562.1	21276.5	37	11434.0	14584.0	13310.4	12793.7
12	13002.0	12722.4	14000.1	18644.7	38	11571.1	14464.7	13180.6	12632.2
13	12245.9	13091.5	13620.2	17445.5	39	12319.1	13728.5	12563.6	10765.9

## Ek İ. Ewma Tahmin ve Artıklarına İlişkin Sonuçlar

**Çizelge Ek İ.1. 2015-2016 grip mevsimi orijinal verilere ilişkin Ewma tahminleri ve artıkları**

Hafta	Orijinal İzleme Verileri(y)	Ewma( $\hat{y}$ )	Ewma Artık ( $y - \hat{y}$ )	Hafta	Orijinal İzleme Verileri(y)	Ewma( $\hat{y}$ )	Ewma Artık ( $y - \hat{y}$ )
40	10205	12221.6	-2016.62	14	16711	21464.8	-4753.84
41	10697	11916.7	-1219.69	15	16249	20421.7	-4172.67
42	11582	11849.8	-267.75	16	15447	19426.7	-3979.73
43	11159	11711.6	-552.60	17	13826	18306.6	-4480.59
44	12479	11865.1	613.92	18	11847	17014.7	-5167.67
45	12680	12028.1	651.93	19	10680	15747.7	-5067.74
46	13492	12320.9	1171.15	20	9670	14532.2	-4862.19
47	12149	12286.5	-137.48	21	8624	13350.6	-4726.55
48	14613	12751.8	1861.21	22	7861	12252.6	-4391.64
49	14992	13199.8	1792.17	23	7089	11219.9	-4130.91
50	15400	13639.9	1760.14	24	6602	10296.3	-3694.33
51	14805	13872.9	932.11	25	5900	9417.1	-3517.06
52	15666	14231.5	1434.49	26	5576	8648.9	-3072.85
1	15648	14514.8	1133.19	27	5045	7928.1	-2883.08
2	15507	14713.2	793.75	28	4991	7340.7	-2349.66
3	15565	14883.6	681.40	29	4932	6858.9	-1926.93
4	18124	15531.7	2592.32	30	4799	6446.9	-1647.95
5	18927	16210.7	2716.26	31	4457	6049.0	-1591.96
6	21732	17315.0	4417.01	32	4754	5790.0	-1035.97
7	24618	18775.6	5842.40	33	5148	5661.6	-513.57
8	26397	20299.9	6097.12	34	6012	5731.7	280.34
9	28228	21885.5	6342.50	35	6147	5814.7	332.27
10	30319	23572.2	6746.80	36	6304	5912.6	391.42
11	24821	23822.0	999.04	37	7132	6156.5	975.54
12	21874	23432.4	-1558.37	38	7762	6477.6	1284.43
13	19537	22653.3	-3116.29	39	6959	6573.9	385.14

Referans Ewma artıkları için  $\bar{x} = -177.071$  Referans Ewma artıkları için  $\sigma = 795.014$

Ek İ. (devam)

Çizelge Ek İ.2. 2015-2016 grip mevsimi için mevsimsellik düzeltmesi yapılmış verilere ilişkin Ewma tahminleri ve artıkları

Hafta	Düzeltilmiş İzleme Verileri(y)	Ewma( $\hat{y}$ )	Ewma Artık ( $y - \hat{y}$ )	Hafta	Düzeltilmiş İzleme Verileri(y)	Ewma( $\hat{y}$ )	Ewma Artık ( $y - \hat{y}$ )
40	13264.1	13338.3	-74.16	14	16205.3	17284.8	-1079.49
41	12846.9	13240.0	-393.09	15	17847.5	17397.3	450.17
42	13273.0	13246.6	26.41	16	19247.7	17767.4	1480.29
43	11890.1	12975.3	-1085.19	17	18564.4	17926.8	637.60
44	13936.4	13167.5	768.89	18	17507.7	17843.0	-335.28
45	12811.2	13096.3	-285.05	19	15907.8	17455.9	-1548.15
46	13519.0	13180.8	338.20	20	15397.9	17044.3	-1646.44
47	9965.3	12537.7	-2572.40	21	14572.2	16549.9	-1977.71
48	12055.8	12441.3	-385.52	22	14767.4	16193.4	-1426.01
49	9359.0	11824.9	-2465.86	23	14319.1	15818.5	-1499.45
50	7264.5	10912.8	-3648.29	24	13926.9	15440.2	-1513.32
51	7655.5	10261.3	-2605.83	25	12951.9	14942.6	-1990.65
52	7442.1	9697.5	-2255.38	26	13221.8	14598.4	-1376.60
1	7036.6	9165.3	-2128.71	27	13293.1	14337.3	-1044.24
2	6893.5	8710.9	-1817.44	28	13221.1	14114.1	-892.99
3	6650.7	8298.9	-1648.20	29	13814.8	14054.2	-239.44
4	8519.4	8343.0	176.40	30	13856.7	14014.7	-158.03
5	9660.1	8606.4	1053.68	31	13104.4	13832.7	-728.26
6	12385.5	9362.2	3023.27	32	14002.7	13866.7	136.03
7	16516.1	10793.0	5723.09	33	14618.2	14017.0	601.22
8	19122.4	12458.9	6663.51	34	16233.1	14460.2	1772.90
9	21696.0	14306.3	7389.69	35	14725.1	14513.2	211.92
10	24405.8	16326.2	8079.59	36	14042.2	14419.0	-376.78
11	21276.5	17316.3	3960.23	37	12793.7	14093.9	-1300.23
12	18644.7	17582.0	1062.75	38	12632.2	13801.6	-1169.38
13	17445.5	17554.7	-109.16	39	10765.9	13194.4	-2428.55
Referans Ewma artıkları için $\bar{x} = 10.5339$				Referans Ewma artıkları için $\sigma = 318.596$			

**Ek J. CDC Grip Verileri İçin Ears C1, C2 Ve C3 Yöntemlerine İlişkin  
Sonuçlar**

**Çizelge Ek J. 2015-2016 grip mevsimine ilişkin Ears C1, C2 Ve C3 yöntemlerinden elde edilen sonuçlar**

Hafta	Ears C1 Eşik Değerleri	Ears C2 Eşik Değerleri	Ears C3 Eşik Değerleri	Hafta	Ears C1 Eşik Değerleri	Ears C2 Eşik Değerleri	Ears C3 Eşik Değerleri
40	1.5042	1.5787	1.4741	14	25.5166	35.2068	22.5018
41	1.6080	1.2317	1.0860	15	26.3813	28.8637	18.4116
42	1.6132	1.5042	2.1719	16	26.6265	25.5166	16.8141
43	2.2967	1.6080	0.0000	17	24.5705	26.3813	16.3525
44	2.4387	1.6132	0.4109	18	16.2414	26.6265	15.2791
45	2.3386	2.2967	0.5661	19	10.9453	24.5705	13.5901
46	2.4227	2.4387	1.1903	20	10.3359	16.2414	8.6603
47	2.2733	2.3386	0.0000	21	11.0490	10.9453	5.6792
48	3.8189	2.4227	2.6631	22	11.5670	10.3359	5.2363
49	4.2579	2.2733	0.0000	23	11.2383	11.0490	5.5740
50	4.7434	3.8189	0.0000	24	9.1228	11.5670	5.5558
51	5.2781	4.2579	0.0000	25	7.1194	11.2383	4.9204
52	9.4538	4.7434	0.0000	26	6.2400	9.1228	4.3926
1	9.9126	5.2781	0.0000	27	5.3614	7.1194	3.4516
2	8.3696	9.4538	0.0000	28	4.6528	6.2400	3.2285
3	7.5311	9.9126	0.0000	29	3.8057	5.3614	3.2756
4	0.7231	8.3696	1.6522	30	2.9465	4.6528	2.8404
5	-1.6374	7.5311	4.4631	31	2.0433	3.8057	2.4967
6	-2.2359	0.7231	7.5121	32	1.4829	2.9465	2.1527
7	6.3744	-1.6374	7.4851	33	1.4418	2.0433	1.9663
8	17.6268	-2.2359	0.0000	34	1.6278	1.4829	2.4490
9	26.2861	6.3744	0.0000	35	2.5994	1.4418	0.8020
10	33.2460	17.6268	0.0000	36	2.5756	1.6278	0.0000
11	37.6694	26.2861	0.0000	37	2.6756	2.5994	0.9585
12	35.2068	33.2460	0.0000	38	2.7334	2.5756	1.4210
13	28.8637	37.6694	19.4060	39	2.8327	2.6756	0.0000

## Ek K. Serfling Modeli Tahmin ve Artıkları

**Çizelge Ek K. 2015-2016 grip mevsimi için Serfling regresyon tahminleri ve artıkları**

Hafta	Orijinal İzleme Verileri(y)	Serfling ( $\hat{y}$ )	Serfling Artık ( $y - \hat{y}$ )	Hafta	Orijinal İzleme Verileri(y)	Serfling ( $\hat{y}$ )	Serfling Artık ( $y - \hat{y}$ )
40	102.05	97.5	4.55	14	167.11	152	15.11
41	106.97	106	0.97	15	162.49	140	22.49
42	115.82	115	0.82	16	154.47	129	25.47
43	111.59	125	-13.41	17	138.26	117	21.26
44	124.79	135	-10.21	18	118.47	107	11.47
45	126.80	146	-19.2	19	106.80	97.5	9.3
46	134.92	158	-23.08	20	96.70	89.8	6.9
47	121.49	170	-48.51	21	86.24	83.9	2.34
48	146.13	183	-36.87	22	78.61	79.6	-0.99
49	149.92	196	-46.08	23	70.89	76.5	-5.61
50	154.00	207	-53	24	66.02	74.3	-8.28
51	148.05	217	-68.95	25	59.00	72.3	-13.3
52	156.66	225	-68.34	26	55.76	70.2	-14.44
1	156.48	230	-73.52	27	50.45	67.8	-17.35
2	155.07	232	-76.93	28	49.91	65.1	-15.19
3	155.65	231	-75.35	29	49.32	62.2	-12.88
4	181.24	229	-47.76	30	47.99	59.6	-11.61
5	189.27	225	-35.73	31	44.57	57.7	-13.13
6	217.32	219	-1.68	32	47.54	56.9	-9.36
7	246.18	213	33.18	33	51.48	57.6	-6.12
8	263.97	206	57.97	34	60.12	60.1	0.02
9	282.28	198	84.28	35	61.47	64.3	-2.83
10	303.19	190	113.19	36	63.04	70	-6.96
11	248.21	182	66.21	37	71.32	76.9	-5.58
12	218.74	173	45.74	38	77.62	84.6	-6.98
13	195.37	163	32.37	39	69.59	92.8	-23.21

## Ek L. Salgın Belirleme Algoritmalarına İlişkin R Kodları

Çizelge Ek L.1. Farrington ve FarringtonFlexible yöntemlerine ilişkin kullanılan R kodları

<b>Farrington</b>
<ul style="list-style-type: none"><li>• <code>far&lt;-as.matrix(read.table("kkk.txt"))</code></li><li>• <code>far1&lt;-new("sts",start=c(2011,27),freq=52,observed=far,state=matrix(0,nrow(far),ncol(far)))</code></li><li>• <code>controlfar&lt;-list(range=170:221,b=3,w=2,alpha=0.01,reweight=T,trend=T,weightThreshold=1,pastWeeksNotIncluded=3,pThresholdTrend=0.05,thresholdMethod="delta",noPeriods=1,powertrans="2/3",glmWarnings=T)</code></li><li>• <code>izlemefar&lt;-farrington(far1,control=controlfar)</code></li><li>• <code>plot(izlemefar)</code></li><li>• <code>upperbound(izlemefar)</code></li></ul>
<b>FarringtonFlexible</b>
<ul style="list-style-type: none"><li>• <code>far&lt;-as.matrix(read.table("kkk.txt"))</code></li><li>• <code>far1&lt;-new("sts",start=c(2011,27),freq=52,observed=far,state=matrix(0,nrow(far),ncol(far)))</code></li><li>• <code>controlfar1&lt;-list(range=170:221,b=3,w=2,alpha=0.01,reweight=T,trend=T,weightThreshold=2.58,pastWeeksNotIncluded=26,pThresholdTrend=0.05,thresholdMethod="muan",noPeriods=10,powertran="2/3",glmWarnings=T)</code></li><li>• <code>izlemefar1&lt;-farringtonFlexible(far1,control=controlfar1)</code></li><li>• <code>plot(izlemefar1)</code></li><li>• <code>upperbound(izlemefar1)</code></li></ul>

## Ek L. (devam)

**Çizelge Ek L.2. Ears C1, C2 ve C3 yöntemlerine ilişkin kullanılan R kodları**

<b>Ears C1</b>
<ul style="list-style-type: none"><li>• <code>x&lt;-as.matrix(read.table("x.txt"))</code></li><li>• <code>mysts&lt;-new("sts",start=c(2014,27),observed=x)</code></li><li>• <code>control&lt;-list(range=12:63,method="C1",alpha=0.001)</code></li><li>• <code>surv&lt;-earsC(mysts,control=control)</code></li><li>• <code>plot(surv)</code></li></ul>
<b>Ears C2</b>
<ul style="list-style-type: none"><li>• <code>x&lt;-as.matrix(read.table("x.txt"))</code></li><li>• <code>mysts&lt;-new("sts",start=c(2014,27),observed=x)</code></li><li>• <code>control1&lt;-list(range=12:63,method="C2",alpha=0.001)</code></li><li>• <code>surv1&lt;-earsC(mysts,control=control1)</code></li><li>• <code>plot(surv1)</code></li></ul>
<b>Ears C3</b>
<ul style="list-style-type: none"><li>• <code>x&lt;-as.matrix(read.table("x.txt"))</code></li><li>• <code>mysts&lt;-new("sts",start=c(2014,27),observed=x)</code></li><li>• <code>control2&lt;-list(range=12:63,method="C3",alpha=0.025)</code></li><li>• <code>surv2&lt;-earsC(mysts,control=control2)</code></li><li>• <code>plot(surv2)</code></li></ul>

## Ek L. (devam)

Çizelge Ek L.3. Boda yöntemine ilişkin kullanılan R kodları

Boda
<ul style="list-style-type: none"><li>• <code>x&lt;-as.matrix(read.table("boda1.txt"))</code></li><li>• <code>xstate&lt;-as.matrix(read.table("boda.txt"))</code></li><li>• <code>bodasts&lt;-new("sts",start=c(2012,40),freq=52,observed=x,state=xstate)</code></li><li>• <code>controlboda&lt;-</code> <code>list(range=157:208,X=NULL,trend=T,season=T,prior='iid',alpha=0.05,mc.munu=100,mc.</code> <code>y=10,samplingMethod="joint")</code></li><li>• <code>**controlboda*-&lt;</code> <code>list(range=157:208,X=NULL,trend=T,season=T,prior='iid',alpha=0.10,mc.munu=100,mc.</code> <code>y=10,samplingMethod="joint")</code></li><li>• <code>izlemeboda&lt;-boda(bodasts,control=controlboda)</code></li><li>• <code>plot(izlemeboda)</code></li><li>• <code>upperbound(izlemeboda)</code></li><li>• <code>alarms(izlemeboda)</code></li></ul>



## Ek L. (devam)

### Çizelge Ek L.4. Holt-Winters yöntemine ilişkin kullanılan R kodları

Holt-Winters
<ul style="list-style-type: none"><li>• <code>grip&lt;-ts(scan("w.txt"),start=c(2012,40),frequency=52)</code></li><li>• <code>tsoutliers(grip)</code></li><li>• <code>gript&lt;-tsclean(grip)</code></li><li>• <code>plot(grip,xlab="Zaman",ylab="kişi sayısı",main="Referans Dönemi",ylim=c(0,40000))</code></li><li>• <code>plot(gript,xlab="Zaman",ylab="kişi sayısı",main="Aykırı Değer Analizi Sonrası Referans Dönemi",ylim=c(0,40000))</code></li><li>• <code>loggript&lt;-log(gript)</code></li><li>• <code>HWgrip&lt;-HoltWinters(loggript)</code></li><li>• <code>HWtahmin&lt;-forecast.HoltWinters(HWgrip,52)</code></li><li>• <code>accuracy(HWtahmin)</code></li><li>• <code>acf(resid(HWgrip),lag=103,main="Holt-Winters Yöntemi Artıkları")</code></li><li>• <code>pacf(resid(HWgrip),lag=103,main="Holt-Winters Yöntemi Artıkları")</code></li><li>• <code>LBQPlot(resid(HWgrip),lag.max = 103, StartLag =1, SquaredQ = FALSE)</code></li><li>• <code>ad.test(resid(HWgrip))</code></li><li>• <code>shapiro.test(resid(HWgrip))</code></li><li>• <code>wawotest(resid(HWgrip))</code></li></ul>

Çizelge Ek L.5. Arima yöntemine ilişkin kullanılan R kodları

Arima
<ul style="list-style-type: none"><li>• <code>grip&lt;-ts(scan("w.txt"),start=c(2012,40),frequency=52)</code></li><li>• <code>tsoutliers(grip)</code></li><li>• <code>gript&lt;-tsclean(grip)</code></li><li>• <code>lngript&lt;-log(gript)</code></li><li>• <code>acf(lngript,lag=155)</code></li><li>• <code>pacf(lngript,lag=155)</code></li><li>• <code>adf.test(lngript)</code></li><li>• <code>pp.test(lngript)</code></li><li>• <code>diff(lngript,differences=1)</code></li><li>• <code>acf(diff(lngript,differences=1),lag=154)</code></li><li>• <code>pacf(diff(lngript,differences=1),lag=154)</code></li><li>• <code>adf.test(diff(lngript,differences=1))</code></li><li>• <code>pp.test(diff(lngript,differences=1))</code></li><li>• <code>lngriptf1&lt;-diff(lngript,differences=1)</code></li><li>• <code>lngriptf1s1&lt;-diff(lngript,differences=1,lag=52)</code></li><li>• <code>acf(lngriptf1s1,lag=103)</code></li><li>• <code>pacf(lngriptf1s1,lag=103)</code></li><li>• <code>adf.test(lngriptf1s1)</code></li><li>• <code>pp.test(lngriptf1s1)</code></li><li>• <code>LBQPlot(lngriptf1s1, lag.max = 102, StartLag =1,SquaredQ = FALSE)</code></li><li>• <code>modell&lt;-arima(lngript,order=c(0,1,1),seasonal=c(0,1,1))</code></li><li>• <code>LBQPlot(resid(modell), lag.max = 102, StartLag =1,SquaredQ = FALSE)</code></li><li>• <code>accuracy(modell)</code></li><li>• <code>modelltah&lt;-forecast.Arima(modell,52,level=c(80,90))</code></li><li>• <code>modelltah</code></li><li>• <code>plot(modelltah)</code></li></ul>

Çizelge Ek L.6. OutbreakP yöntemine ilişkin kullanılan R kodları

OutbreakP
<pre>#uyarı ve alarm sınırı giriliyor alimit&lt;-700 # tao değişim noktası değerleri oluşturuyor #seriler tekrarlı olarak üretiliyor x&lt;-0 sonuc&lt;-0 ist&lt;-0 dongusayisi&lt;-100000 say&lt;-matrix(0,20,3)   for( i in 1:dongusayisi) {     print(i)     tao&lt;-rgeom(1,0.1428)     for(t in 1:20){       if( t&lt;tao) x[t]&lt;-rpois(1,131)       if( t&gt;=tao) x[t]&lt;-rpois(1,exp(4.53+0.21*(t-tao+1)))     }     y &lt;- matrix(x,ncol=1)     mysts &lt;- new("sts", observed=y, epoch=1:length(y), alarm=y*0,start=c(2015,40), freq=52)     upperbound(outbreakP(mysts, control=list(range=1:length(y),k=alimit,ret="value")))     res &lt;- outbreakP(mysts, control=list(range=1:length(y),k=alimit,ret="value"))     for( j in 1:20) {       if(alarms(res)[j,1]==1) {         say[j,1]&lt;-say[j,1]+1         if(j&gt;=tao) say[j,2]&lt;-say[j,2]+1         break       }     }   }   for(j in 1:20){     say[j,3]&lt;-say[j,2]/say[j,1]   }   print(say)   ts.plot(say[,3])</pre>

## Ek L. (devam)

Çizelge Ek L.7. Rossi vd. (1999) Cusum yöntemine ilişkin kullanılan R kodları

Rossi vd. (1999) Cusum
<ul style="list-style-type: none"><li>• <code>x&lt;-scan("cusum.txt")</code></li><li>• <code>y&lt;-matrix(x,ncol=1)</code></li><li>• <code>mysts&lt;-create.disProg(week=1:nrow(y),observed=y,start=c(2012,41))</code></li><li>• <code>kh&lt;-find.kh(ARLa=100,ARLr=2)</code></li><li>• <code>cdc.cusum&lt;- algo.cusum(mysts,control=list(k=kh\$k,h=kh\$h,m="glm",trans="rossi",range=15 6:207))</code></li><li>• <code>cdc.cusum\$alarm</code></li><li>• <code>plot(cdc.cusum,firstweek=40,startyear=2016)</code></li></ul>

## Ek L. (devam)

**Çizelge Ek L.8. Bayes yöntemine ilişkin kullanılan R kodları**

Bayes
<ul style="list-style-type: none"><li>• <code>x&lt;-scan("bayes.txt")</code></li><li>• <code>y&lt;-matrix(x,ncol=1)</code></li><li>• <code>mysts&lt;-create.disProg(week=1:nrow(y),observed=y,start=c(2012,41))</code></li><li>• <code>cntrl&lt;-list(range=156:207,w=2,b=2,alpha=0.01)</code></li><li>• <code>cdc.bayes&lt;-algo.bayes(mysts,control=cntrl)</code></li><li>• <code>plot(cdc.bayes,firstweek=40,startyear=2016)</code></li><li>• <code>print(cdc.bayes\$alarm)</code></li></ul>

## Ek L. (devam)

**Çizelge Ek L.9. Gizli Markov yöntemine ilişkin kullanılan R kodları**

Hidden Markov Model
<ul style="list-style-type: none"><li>• <code>x&lt;-scan("hmm.txt")</code></li><li>• <code>y&lt;-matrix(x,ncol=1)</code></li><li>• <code>mysts&lt;-create.disProg(week=1:nrow(y),observed=y,start=c(2012,41))</code></li><li>• <code>cntrl&lt;-list(range=156:207,Mtilde=-1, noStates=2, trend=FALSE, noHarmonics=1, covEffectEqual=FALSE,saveHMMs=FALSE,extraMSMargs=list())</code></li><li>• <code>cdc.hmm&lt;-algo.hmm(mysts,control=cntrl)</code></li><li>• <code>plot(cdc.hmm,firstweek=40,startyear=2016)</code></li><li>• <code>print(cdc.hmm\$alarm)</code></li></ul>

## Ek L. (devam)

**Çizelge Ek L.10. Rki yöntemine ilişkin kullanılan R kodları**

Rki
<ul style="list-style-type: none"><li>• <code>x&lt;-scan("rki.txt")</code></li><li>• <code>y&lt;-matrix(x,ncol=1)</code></li><li>• <code>mysts&lt;-create.disProg(week=1:nrow(y),observed=y,start=c(2012,41))</code></li><li>• <code>cntrl&lt;-list(range=156:207,w=2,b=2,alpha=0.01)</code></li><li>• <code>cdc.rki&lt;-algo.rki(mysts,control=cntrl)</code></li><li>• <code>plot(cdc.rki,firstweek=40,startyear=2016)</code></li><li>• <code>print(cdc.rki\$alarm)</code></li></ul>

**Ek L.** (devam)

**Çizelge Ek L.11. GLR yöntemine ilişkin kullanılan R kodları**

<b>Genelleştirilmiş Olabilirlik Oranı İstatistiğine Dayalı Poisson Regresyon Kontrol Grafiği</b>	
<b>GLM ile Poisson regresyon modellerinin kesitilmesi</b>	
<ul style="list-style-type: none"><li>• <code>x&lt;-scan("glrn.txt")</code></li><li>• <code>y&lt;-x[1:155]</code></li><li>• <code>t&lt;-1:155</code></li><li>• <code>omega&lt;-2*pi/52</code></li><li>• <code>omega1&lt;-4*pi/52</code></li><li>• <code>omega2&lt;-6*pi/52</code></li><li>• <code>x1&lt;-cos(omega*t)</code></li><li>• <code>x2&lt;-sin(omega*t)</code></li><li>• <code>x3&lt;-cos(omega1*t)</code></li><li>• <code>x4&lt;-sin(omega1*t)</code></li><li>• <code>x5&lt;-cos(omega2*t)</code></li><li>• <code>x6&lt;-sin(omega2*t)</code></li><li>• <code>yglm1&lt;-glm(y~1+x1+x2,family=poisson())</code></li><li>• <code>print(summary(yglm1))</code></li><li>• <code>yglm2&lt;-glm(y~1+x1+x2+x3+x4,family=poisson())</code></li><li>• <code>print(summary(yglm2))</code></li><li>• <code>yglm3&lt;-glm(y~1+x1+x2+x3+x4+x5+x6,family=poisson())</code></li><li>• <code>print(summary(yglm3))</code></li></ul>	



## Ek L. (devam)

Çizelge Ek L.11. (devam) GLR yöntemine ilişkin kullanılan R kodları

### Şeçilen uygun GLM modeli eşik değeri belirleyen simülasyon çalışması

```
#Generate data
say<-0
n<-10000
for(i in 1:n){
print(i)
S <- 2 ; t <- 1:52 ; m <- length(t)
beta <- c(4.450556,0.041178,0.072032,-0.099567,-0.045020)
omega <- 2*pi/52
omega1<-4*pi/52
omega2<-6*pi/52
#log mu_{0,t}
base <- beta[1] + beta[2] * cos(omega*t) + beta[3] * sin(omega1*t)+beta[4] * cos(omega2*t)+beta[5]
* sin(omega2*t)
mu0<-exp(base)
#alpha <- 0.06273533
x <- rpois(52,mu0)
#x <- rnbinom(52,mu=mucontrol,size=1/alpha)
y <- as.matrix(x,ncol=1)
mysts <- create.disProg(week = 1:nrow(y), observed = y,start=c(2015,40))
#Run GLR based detection
cntrl=list(range=1:52,c.ARL=5,mu0=mu0,alpha=0,change="intercept",ret="value",dir="inc")
glr.ts <- algo.glrnb(mysts,control=c(cntrl))
#plot(glr.ts)
if(sum(glr.ts$alarm)!=0) {
say<-say+1
}
}
p<-say/n
print(p)
```

## Ek L. 11<sup>(devam)</sup>

### Çizelge Ek L.11.<sup>(devam)</sup> GLR yöntemine ilişkin kullanılan R kodları

İlgili algoritma ile R programından sonuçların elde edilmesi
<ul style="list-style-type: none"><li>• <code>x&lt;-scan("glrnb.txt")</code></li><li>• <code>y&lt;-matrix(x,ncol=1)</code></li><li>• <code>mysts&lt;-create.disProg(week=1:nrow(y),observed=y,start=c(2012,41))</code></li><li>• <code>mu0&lt;-list(s=3, trend=FALSE, refit=FALSE)</code></li><li>• <code>cntrl&lt;-list(range=156:207,c.ARL=6,mu0=mu0,alpha=0,Mtilde=1,M=-1,change="intercept",theta0=NULL,dir=c("inc"),ret="value")</code></li><li>• <code>cdc.glrnb&lt;-algo.glrnb(mysts,control=cntrl)</code></li><li>• <code>plot(cdc.glrnb,firstweek=40,startyear=2016)</code></li><li>• <code>print(cdc.glrnb\$alarm)</code></li></ul>

Çizelge Ek L.12. Haar-wavelet yöntemine ilişkin kullanılan Matlab kodları

```

Haar-wavelet
function SWTchart(s,startdate,in_control_last_date,confidence,fdr,AR);
% Creates a set of control charts to monitor a series, using a 5-level
% redundant wavelet decomposition using the Haar wavelet.
% The signal should start with a no-outbreak period.
% The user inputs the series, the start date, and the end date of the no-outbreak period.
% Optional arguments are the level of confidence, whether an FDR correction should be used across scale, and whether detail control limits should be based on
AR(7).
% Call function by typing: SWTchart(s,startdate,in_control_last_date,confidence,fdr,AR)
% s = series (vector)
% startdate = start date of series. Format is '30-Jan-1999'.
% in_control_last_date = the last date where the series is known to be in-control. Default is length(s)-1.
% confidence = level of confidence. This is equal to 1-(1-alpha)/2, where alpha is the false alarm rate. Default is 0.9973 (alpha=0.00135).
% fdr = whether an fdr correction is used (1) or not (0). Default is 0.
% AR = whether to use AR(7)-based control limits for the detail coeffs (1) or ordinary Shewhart limits (0). Default is 0.
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% This program is free software; you can redistribute it and/or modify
% it under the terms of the GNU General Public License as published by
% the Free Software Foundation; either version 2 of the License, or
% (at your option) any later version.
% This program is distributed in the hope that it will be useful,
% but WITHOUT ANY WARRANTY; without even the implied warranty of
% MERCHANTABILITY or FITNESS FOR A PARTICULAR PURPOSE. See the
% GNU General Public License for more details.
% You should have received a copy of the GNU General Public License
% along with this program; if not, write to the Free Software
% Foundation, Inc., 51 Franklin St, Fifth Floor, Boston, MA 02110-1301 USA
n = length(s); %Length of the time series
s=s(1:n);
alarm_dates = [];
%T=[datenum(startdate):datenum(startdate)+n-1]; % dates of in-control period
T=[1:n];
lambda=0.4;
UCL=NaN*ones(6,n);
LCL=UCL;
Forecast = UCL;
if nargin<3,
    in_control_idx = n;
else
    % in_control_idx = datenum(in_control_last_date)-datenum(startdate)+1 ;
    in_control_idx = 156;
end
if nargin<4,
    confidence = 0.9973; % corresponds to 3-sigma limits
end

if nargin<5,
    fdr=0;
end
if nargin<6,
    AR=0;
end
mult=4;
%mult = norminv(confidence);
%Perform the decomposition of s at level 5, using haar-backward wavelet.
[cA0,cD0] = swt_haar_back(s(1:in_control_idx));
%Compute in-control std for each scale
for i = 1:2
    StdApp(i) = std(cA0(2^i:in_control_idx,i));
    StdDet(i) = std(cD0(2^i:in_control_idx,i));
    if AR == 1,
        % estimate AR parameters from first half of in-control data
        [const, ARparams0(i,:),sig_sq0(i)] = arfit(cD0(2^i:floor(in_control_idx/2)),7,7,'zero');
        % estimate standard error of forecast error from second half of in-control data
        for t=8:in_control_idx,
            Forecast(t,i) = cD0(t-1:-1:t-7,i)' * ARparams0(i,:);
        end
        sig0(i) = std( Forecast(floor(in_control_idx/2):in_control_idx,i)-cD0(floor(in_control_idx/2):in_control_idx,i) );
    end
end
in_control_mean=mean(cA0(32:in_control_idx,2));
    
```

## Ek L. (devam)

### Çizelge Ek L.12. (devam) Haar-wavelet yöntemine ilişkin kullanılan Matlab kodları

```
[cA,cD] = swt_haar_back(s);
%% Merge cA5 coefficients with detail coefficients
AD = [cD , cA(:,2)];
% Compute in-control mean for each scale
mean_AD = [0 0 in_control_mean];
std_AD = [StdDet ; StdApp(2)];
% Ordinary Shewhart control limits (w/o FDR correction), if we ignore autocorrelation
if AR == 0,
    UCL = [mean_AD + mult*std_AD] * ones(1,n);
    LCL = [mean_AD - mult*std_AD] * ones(1,n);
else
%% Forecast each detail series by an AR(7) model, estimated from the in-control period
    for t=in_control_indx+1:n,
        for i = 1:2
            [const, ARparams(i,:),sig,SBC,FPE,th] = arfit(AD(i,2^i:t-1),7,7,'zero');
            Forecast(t,i) = AD(i,t-1:-1:t-7)*ARparams(i,:);
        end
    end
    Forecast(:,6) = mean_AD(6);
    sig0(6) = std_AD(6)*sqrt(lambda/(2-lambda));
    for i=1:2,
        UCL(i,:) = Forecast(:,i) + mult* sig0(i);
        LCL(i,:) = Forecast(:,i) - mult* sig0(i);
    end
    LCL(6,:) = 0;
    UCL(6,:) = Forecast(:,6) - norminv(2*(1-confidence))*sig0(6); %% EWMA
end
Alarm_levels = NaN*ones(6,n);
j=1;
if fdr==1,
    for t = 1:n;
        if AR ==1,
            p_values = [2*(1-normcdf(abs( (AD(1:2,t)-Forecast(t,1:2))./sig0(1:2) ))), 1-normcdf((AD(6,t)-Forecast(t,6))/sig0(6))];
        else,
            p_values = 2*(1-normcdf(abs( (AD(:,t)-mean_AD)./std_AD )));
        end
        [threshold,locations] = fdr_DS(p_values, 2*(1-confidence));
        Alarm_levels(locations,t) = 1;
        if threshold<0,
            alarm_dates(j)=t;
            j=j+1;
        end
    end
else
    for i=1:size(AD,1),
        for t=1:n
            if (AD(i,t) > UCL(i,t)),
                Alarm_levels(i,t) = 1;
                alarm_dates(j)=t;
                j=j+1;
            end
        end
    end
end
end
%% PLOT WAVELET DECOMPOSITION (WITH THRESHOLDS, if no FDR correction used)
ADtitles = {'cD_1', 'cD_2', 'cA_2'};
figure;
for i = 1:size(AD,1),
    AlarmCoef = AD(i,:).*Alarm_levels(i,:);
    subplot(4,1,i);
    hold on;
    plot(T, UCL(i,:), 'color',[0.5,0.5,0.5]);
    plot(T, LCL(i,:), 'color',[0.5,0.5,0.5]);
    plot(T, AlarmCoef, 'ro', 'markerfacecolor', 'r');
    plot(T, AD(i,:), 'k');
    ylabel(ADtitles(i), 'fontweight', 'bold');
    set(gca, 'xticklabel', '', 'yticklabel', '');
    hold off;
end
subplot(4,1,4); plot(T,s,'k');
% datetick('x',12,'keeplimits','keepticks');
set(gca, 'FontSize', 10, 'FontWeight', 'bold', 'Yticklabel', '');
%% PRINT ALARM DATES
alarm_dates = sort(unique(alarm_dates));
% datestr(alarm_dates+T(1)-1)
% (alarm_dates)
```

## Ek L. (devam)

### Çizelge Ek L.12. (devam) Haar-wavelet yöntemine ilişkin kullanılan Matlab kodları

```
function[cA,cD] = swt_haar_back(s)
%%% Implements a 5-level Haar wavelet redundant decomposition (SWT), with
%%% coefficients being computed only from the past.
n = length(s);
s = s(1:n);
cA = NaN*ones(3,n);
cD = cA;
cA(1,:) = s;
for j=1:2,
    cA(j+1, 2^j:n) = ( cA(j, 2^j-2^(j-1):n-2^(j-1)) + cA(j, 2^j:n))/sqrt(2);
    cD(j+1, 2^j:n) = ( cA(j, 2^j-2^(j-1):n-2^(j-1)) - cA(j, 2^j:n))/sqrt(2);
end;
cA = cA(2:3,:);
cD = cD(2:3,:);
%%%
function [pID,alarm_scales] = fdr_DS(p,q)
% False Discovery Rate correction
% p - vector of p-values
% q - False Discovery Rate level
%
% pID - p-value threshold based on independence or positive dependence
% alarm_scales - gives the locations of the p-values that sound an alarm
%%% based on code by Tom Nichols
[p,order] = sort(p(:));
V = length(p) - sum(isnan(p));
I = (1:V);
cVID = 1;
if V>0,
    pID = max(find(p(1:V)<=I/V*q/cVID));
else,
    pID = [];
end
alarm_scales = sort(order(1:pID));
```

**Ek M. 2015-2016 Grip Mevsimi Verileri İçin UCengiz Yöntemine İlişkin  
Sonuçlar**

**Çizelge Ek M.1. 4 ve 8 gözlemlik referans uzunluklarında orijinal veriler için ATFS 13  
değerinde UCengiz yönteminden elde edilen sonuçlar**

Hafta	Orijinal veri	Eşik değeri (4gözlemlik)	Eşik değeri (8gözlemlik)	Hafta	Orijinal veri	Eşik değeri (4gözlemlik)	Eşik değeri (8gözlemlik)
40	10049	9558	9876	14	16769	9558	9876
41	10715	9558	9876	15	16261	9558	9876
42	11584	9558	9876	16	15437	9558	9876
43	11164	9558	9876	17	13836	9558	9876
44	12423	9558	9876	18	11870	9558	9876
45	12676	9558	9876	19	10707	9558	9876
46	13484	9558	9876	20	9798	9558	9876
47	12158	9558	9876	21	8787	9558	12483
48	14617	9558	9876	22	7997	10685	12322
49	14986	9558	9876	23	7201	9989	11599
50	15396	9558	9876	24	6768	10293	11121
51	14839	9558	9876	25	6024	10147	10965
52	15679	9558	9876	26	5686	9398	11191
1	15654	9558	9876	27	5118	8390	11016
2	15517	9558	9876	28	5078	7838	10577
3	15661	9558	9876	29	5044	6688	9339
4	18143	9558	9876	30	4925	5879	8198
5	18944	9558	9876	31	4596	5294	7239
6	21774	9558	9876	32	4868	5637	6592
7	24681	9558	9876	33	5265	5421	5931
8	26423	9558	9876	34	6260	6026	5765
9	28468	9558	9876	35	6504	6026	5765
10	30481	9558	9876	36	6623	6026	5765
11	24894	9558	9876	37	7639	6026	5765
12	21940	9558	9876	38	8229	6026	5765
13	19603	9558	9876	39	8505	6026	5765

**Ek M.**(devam)**Çizelge Ek M.2. 4 ve 8 gözlemlik referans uzunluklarında işlenmiş veriler için ATFS 13 değerinde UCengiz yönteminden elde edilen sonuçlar**

Hafta	İşlenmiş veri	Eşik değeri (4gözlemlik)	Eşik değeri (8gözlemlik)	Hafta	İşlenmiş veri	Eşik değeri (4gözlemlik)	Eşik değeri (8gözlemlik)
40	182.2692	194	195	14	156.2610	189	192
41	162.0998	200	198	15	177.6229	183	188
42	161.7120	196	195	16	172.3487	194	192
43	146.7196	193	193	17	162.3780	197	192
44	161.8470	184	188	18	156.5665	195	190
45	150.6716	184	187	19	162.5889	191	188
46	155.7376	181	184	20	163.3726	189	189
47	134.6010	181	182	21	160.9218	189	189
48	173.3654	173	176	22	162.0742	188	188
49	154.0807	173	181	23	161.3484	189	188
50	156.7625	174	180	24	164.6932	189	187
51	149.9446	177	180	25	161.6398	190	187
52	167.2421	177	179	26	166.0338	190	187
1	162.2666	185	182	27	164.2608	191	188
2	165.0233	188	184	28	170.1617	191	189
3	171.7576	190	186	29	170.8293	194	191
4	198.9536	195	189	30	170.4664	196	192
5	185.7021	195	189	31	168.6320	197	193
6	209.1580	203	194	32	174.5991	197	194
7	212.5871	203	194	33	175.4304	199	196
8	203.0018	203	194	34	180.5731	201	198
9	207.4453	203	194	35	171.7927	205	200
10	207.8639	203	194	36	168.8547	203	200
11	131.9375	203	194	37	175.7644	201	199
12	157.7183	211	199	38	169.1450	201	200
13	162.7875	199	195	39	163.4442	199	199

**Ek M.**(devam)**Çizelge Ek M.3. 4 ve 8 gözlemlik referans uzunluklarında orijinal veriler için ATFS 26 değerinde UCengiz yönteminden elde edilen sonuçlar**

Hafta	Orijinal veri	Eşik değeri (4gözlemlik)	Eşik değeri (8gözlemlik)	Hafta	Orijinal veri	Eşik değeri (4gözlemlik)	Eşik değeri (8gözlemlik)
40	10049	10336	10715	14	16769	16256	16628
41	10715	16319	14402	15	16261	16256	16628
42	11584	17705	16503	16	15437	16256	17161
43	11164	17543	18026	17	13836	16178	16977
44	12423	13577	17407	18	11870	19447	18489
45	12676	15170	17571	19	10707	22703	21102
46	13484	15463	17226	20	9798	21932	22603
47	12158	16652	17245	21	8787	18397	23019
48	14617	15705	15717	22	7997	15724	22639
49	14986	19476	17542	23	7201	14222	21203
50	15396	20371	18618	24	6768	13003	18995
51	14839	19968	19302	25	6024	11171	16450
52	15679	16408	18713	26	5686	10400	14581
1	15654	17217	18797	27	5118	9203	12996
2	15517	17182	18412	28	5078	8646	11749
3	15661	16757	17697	29	5044	7174	10275
4	18143	16256	16628	30	4925	6130	8894
5	18944	16256	16628	31	4596	5385	7732
6	21774	16256	16628	32	4868	5910	6975
7	24681	16256	16628	33	5265	5623	6168
8	26423	16256	16628	34	6260	6387	5953
9	28468	16256	16628	35	6504	9384	5953
10	30481	16256	16628	36	6623	9550	5953
11	24894	16256	16628	37	7639	8577	5953
12	21940	16256	16628	38	8229	10156	5953
13	19603	16256	16628	39	8505	11545	5953



**Ek M.**(devam)**Çizelge Ek M.4. 4 ve 8 gözlemlik referans uzunluklarında işlenmiş veriler için ATFS 26 değerinde UCengiz yönteminden elde edilen sonuçlar**

Hafta	İşlenmiş veri	Eşik değeri (4gözlemlik)	Eşik değeri (8gözlemlik)	Hafta	İşlenmiş veri	Eşik değeri (4gözlemlik)	Eşik değeri (8gözlemlik)
40	182.2692	203	204	14	156.2610	217	203
41	162.0998	209	207	15	177.6229	191	198
42	161.7120	205	204	16	172.3487	202	202
43	146.7196	202	202	17	162.3780	206	201
44	161.8470	193	197	18	156.5665	204	199
45	150.6716	192	196	19	162.5889	199	197
46	155.7376	189	192	20	163.3726	197	198
47	134.6010	189	191	21	160.9218	197	198
48	173.3654	181	184	22	162.0742	197	197
49	154.0807	202	190	23	161.3484	198	196
50	156.7625	194	188	24	164.6932	197	195
51	149.9446	192	188	25	161.6398	198	196
52	167.2421	190	187	26	166.0338	198	196
1	162.2666	194	190	27	164.2608	200	197
2	165.0233	196	192	28	170.1617	200	197
3	171.7576	199	194	29	170.8293	203	199
4	198.9536	203	197	30	170.4664	205	201
5	185.7021	224	197	31	168.6320	206	202
6	209.1580	223	203	32	174.5991	206	203
7	212.5871	235	203	33	175.4304	208	205
8	203.0018	245	203	34	180.5731	210	206
9	207.4453	245	203	35	171.7927	214	209
10	207.8639	247	203	36	168.8547	212	209
11	131.9375	247	203	37	175.7644	209	208
12	157.7183	291	212	38	169.1450	210	209
13	162.7875	256	207	39	163.4442	208	208

## ÖZGEÇMİŞ

### Kişisel Bilgiler

Ad Soyad : Ünal Cengiz  
Uyruk : T.C.  
Doğum Yeri ve Tarihi: Düzce 11/03/1992  
Medeni Hali : Bekar  
E-posta : unalcengiz07@hotmail.com

### Eğitim

Alınan Derece	Aldığı Kurum/Üniversite	Mezuniyet Yılı
Lise	Cumhuriyet Anadolu Lisesi	2010
Lisans	Muğla Sıtkı Koçman Üniversitesi	2014
Yüksek Lisans	Muğla Sıtkı Koçman Üniversitesi	2018

### Yabancı Dil(ler)

Dil (İngilizce, vs)	Başlangıç	Orta	İleri
Yazma		X	
Konuşma		X	
Anlama		X	
Okuma		X	

### Bilimsel Faaliyetler

1. KESKIN.,K.,CENGİZ.,U.,TASKIN.,T.,AVCI.,F., 2014. Predictions of Scores and Goal Numbers of Future Seasons of Four Grandins in Spor Toto Süper Lig by Time Series. XVth INTERNATIONAL SYMPOSIUM ON ECONOMETRICS, OPERATION RESEARCH AND STATISTICS
2. CENGİZ.,U., AVCI.,F., KARAHASAN.,M., 2015. Applied to Time Series Clustering of Gross National Income OECD Countries and Long-Term

Investment Rate Forecasts. XVth INTERNATIONAL SYMPOSIUM ON  
ECONOMETRICS, OPERATION RESEARCH AND STATISTICS

3. YALCIN.,M., CENGİZ.,U., 2016. Examination of the Relations of Norwegian Kron with Other Most Used Currencies Other than US Dollar and Euro: Cointegration Analysis. XVIIth INTERNATIONAL SYMPOSIUM ON  
ECONOMETRICS, OPERATION RESEARCH AND STATISTICS

