

Sađlamcı ARIMA Modelleri ve Yapay Sinir Ađlarının Karşılaştırmalı İncelenmesi:Turizm
Örneđi

Selim Dönmez

YÜKSEK LİSANS TEZİ

İstatistik Anabilim Dalı

Ocak 2014

Comparative Examination of Robust ARIMA Models and Artificial Neural Networks: an
example on tourism

Selim Dönmez

MASTER OF SCIENCE THESIS

Department of Statistics

January 2014

Sađlamcı ARIMA Modelleri ve Yapay Sinir Ađlarının Karşılaştırılmalı İncelenmesi:Turizm
Örneđi

Selim Dönmez

Eskişehir Osmangazi Üniversitesi

Fen Bilimleri Enstitüsü

Lisansüstü Yönetmeliđi Uyarınca

İstatistik Anabilim Dalı

İstatistiksel Bilgi Sistemleri Bilim Dalında

YÜKSEK LİSANS TEZİ

Olarak Hazırlanmıştır

Danışman: Yrd.Doç.Dr. Özer Özaydın

Ocak 2014

ONAY

İstatistik Anabilim Dalı Yüksek Lisans öğrencisi Selim Dönmez'in YÜKSEK LİSANS tezi olarak hazırladığı "Sağlamcı ARIMA Modelleri ve Yapay Sinir Ağlarının Karşılaştırılmalı İncelenmesi: Turizm Örneği" başlıklı bu çalışma, jürimizce lisansüstü yönetmeliğin ilgili maddeleri uyarınca değerlendirilerek kabul edilmiştir.

Danışman : Yrd.Doç.Dr.Özer Özaydın

İkinci Danışman : -

Yüksek Lisans Tez Savunma Jürisi:

Üye : Yrd.Doç.Dr.Özer Özaydın

Üye : Prof.Dr.Veyssel Yılmaz

Üye : Prof.Dr.Zeki Yıldız

Üye : Doç.Dr.Sevil Şentürk

Üye : Yrd.Doç.Dr.Fatih Çemrek

Fen Bilimleri Enstitüsü Yönetim Kurulu'nun tarih ve sayılı kararıyla onaylanmıştır.

Prof. Dr. Nimetullah BURNAK

Enstitü Müdürü

ÖZET

Zaman serileri analizi, George E. P. Box ve Gwilyn M. Jenkins tarafından sistematik hale getirilmesiyle önemli bir çalışma alanı haline gelmiştir. İstatistiğin çeşitli branşlarında zaman serileri analizi için yeni metotlar geliştirilmiş ve hala da zaman serileri analizi için yeni yöntemler geliştirilmeye devam edilmektedir. Bu yöntemlerden biri de yapay sinir ağları yöntemidir. Yarım asır önce ortaya atılan yapay sinir ağları yöntemi, görüntü işleme, sınıflandırma, örüntü tanımlama gibi pek çok alanda etkin bir şekilde kullanılmaktadır ve şimdilerde zaman serileri analizinde etkin bir şekilde kullanılmaya çalışılmaktadır. Bu yapılırken yapay sinir ağlarının mimarisini oluşturma yolları aranmaktadır. Bu mimariyi oluşturmak, turizm verileri gibi büyük rakamlarda gözlemler içeren verilerde ayrıca zordur. Bunun nedeni analizdeki hata ne kadar küçültülürse küçültülsün, verideki büyük rakamlar nedeniyle gene de mimarideki en ufak hatalar göze batar hale gelmektedir. Bu nedenle bu çalışmada, yapay sinir ağlarında mimari oluşturulurken istatistiksel yöntemlerden faydalanılarak genel turizm analizlerine uygun modellerin bulunması amaçlanmıştır.

Anahtar Kelimeler: Zaman Serileri Analizi, Yapay Sinir Ağları, Yapay Sinir Ağları Mimarisi, İstatistiksel Yöntemler, Genel Turizm Analizi

SUMMARY

Time Series Analysis became a prominent study field by the works of George E. P. Box and Gwilyn M. Jenkins. There has been numerous studies about time series analysis originating from different branches of the discipline of Statistics and new time series analysis methods are still developed. One of these methods is the Artificial Neural Networks method. The Artificial Neural Networks method dating back to studies half a century ago is used effectively in several fields such as Image Processing Classification, Pattern Recognition etc. and is practiced in most studies to effectively analyze any time series data. In those studies, methods for the creation of architecture of Artificial Neural Networks are searched. However, datasets like the ones that regarding general tourism in a country can pose a problem in constructing an Artificial Neural Network architecture. The reason why this particular is aroused is that no matter how much we minimize the error of the analysis, the big figures in the datasets force the flaws of the architecture to look bigger than it already is. In this study, the aim is to produce proper models by using statistical methods to create an ideal architecture for Artificial Neural Networks.

Key Words: Time Series Analysis, Artificial Neural Networks, Architecture of Artificial Neural Networks, Statistical Methods, General Tourism Analysis

Teşekkür

Bu çalışmada, bana desteğini veren danışman hocam Yrd.Doc.Dr. Özer ÖZAYDIN'a teşekkürü borç bilirim.

Uygulama konusunda bana yol gösteren hocam Yrd.Doc.Dr. Fatih ÇEMREK'e ayrıca teşekkürlerimi sunarım.

Veri konusunda bana yardım eden T.C. Eskişehir Valiliğine teşekkür ederim

Annemle babama da manevi desteklerinden ötürü teşekkür ederim.

İÇİNDEKİLER

	<u>Sayfa</u>
ÖZET	v
SUMMARY	vi
TEŞEKKÜRLER	vii
ŞEKİLLER DİZİNİ	x
ÇİZELGELER DİZİNİ	xi
SİMGELER VE KISALTMALAR DİZİNİ	xii
GİRİŞ.....	1
1. ZAMAN SERİLERİNDE KULLANILAN İSTATİSTİKSEL YÖNTEMLER VE MODELLER	4
1.1.ARIMA Modelleri.....	4
1.1.1.Otoregresif Model.....	4
1.1.2.Hareketli Ortalamalar Modeli.....	5
1.1.3.Otoregresif- Hareketli Ortalamalar Karma Modeli.....	5
1.1.4. Otoregresif- Hareketli Ortalamalar Karma Entegre Modeli.....	5
1.2.Otokorelasyon ve Kısmi Otokorelasyon Tahmin Yöntemleri.....	6
1.3.Model Tahmin Yöntemleri.....	8
1.3.1.Klasik Tahmin Yöntemleri.....	8
1.3.1.1.En Küçük Kareler Yöntemi ve Yule-Walker Eşitlikleri.....	8
1.3.1.2 En Çok Olabilirlik Yöntemi.....	14
1.3.2 Sağlamcı İstatistiksel Yöntemler ve Zaman Serilerindeki Kullanımı.....	20
1.3.2.1 Modifiye Azami Olabilirlik Yöntemi.....	20
1.3.2.2 $F\tau$ Tahmin Tekniği.....	27

İÇİNDEKİLER (devam)

1.3.2.3 M-tahmin Edicileri.....	30
2.YAPAY SİNİR AĞLARI.....	32
2.1.Yapay Sinir Ağlarının Tarihçesi.....	32
2.2.Yapay Sinir Ağlarında Öğrenme Şekilleri.....	36
2.2.1.Hata Düzeltmeli Öğrenme.....	36
2.2.2.Hebbçi Öğrenme.....	37
2.2.3.Rekabete Dayalı Öğrenme.....	38
2.2.4.Boltzmann Öğrenme.....	39
2.2.5.Geri Yayılım Algoritması.....	40
2.3.Yapay Sinir Ağları ve Zaman Serileri Analizi.....	41
3.ÇEŞİTLİ TURİZM VERİLERİNİN ÖNGÖRÜSÜNDE SAĞLAMCI ARIMA VE YAPAY SİNİR AĞLARININ KULLANIMI.....	48
3.1.Yurtdışından Gelip Sınırlardan Çıkış Yapan Turist Sayısının Öngörüsü.....	50
3.2.Türkiye’de Turistlerin Geceleme Oranları Öngörüsü.....	53
3.3.Türkiyedeki Turistlerin Kişi Başına Harcama Tutarlarının Öngörüsü.....	56
SONUÇ VE ÖNERİLER.....	60
KAYNAKLAR DİZİNİ.....	63

ŞEKİLLER DİZİNİ

ŞEKİL	Sayfa
Şekil 3.1 Hataların histogramı.....	51
Şekil 3.2 MEÇO ve Yapay Sinir Ağları yöntemleriyle elde edilen öngörüler.....	52
Şekil 3.3 Hataların Histogramı.....	54
Şekil 3.4 Yapay Sinir Ağları yöntemleriyle elde edilen öngörüler.....	55
Şekil 3.5 Hataların histogramı.....	57
Şekil 3.6 Yapay Sinir Ağları yöntemleriyle elde edilen tahminler.....	58
Şekil 3.7 MEÇO yöntemiyle elde edilen öngörüler.....	58

ÇİZELGELER DİZİNİ

<u>ÇİZELGE NUMARASI VE BAŞLIĞI</u>	<u>Sayfa</u>
Çizelge 3.1 r ve d değerleri için Logolabilirlik değerleri.....	51
Çizelge 3.2 İki yonteme göre günümüzdeki dört döneme ilişkin öngörüler.....	53
Çizelge 3.3 Yapay Sinir Ağları yontemine göre önümüzdeki dört döneme ilişkin öngörüler.....	56
Çizelge 3.4 Yapay Sinir Ağları yontemine göre önümüzdeki dört döneme ilişkin tahminler.....	56
Çizelge 3.5 Harcama verisi için üretilen A_t değerleri.....	59
Çizelge 3.6 Harcama verisi için üretilen A_t değerleri.....	59

SİMGELER VE KISALTMALAR DİZİNİ

<u>Simge</u>	<u>Açıklama</u>
Y_t	Zaman Serisi
μ	Sabit Terim
A_t	Hata Terimi
B	Geri işlem Operatörü
r_k	Ampirik Otokorelasyon
N	Zaman Serisini Büyüklüğü
ρ_v	Teorik Otokorelasyon
η	Zaman Serisinin Sonlu Beklenen değeri
ϕ_i	Zaman Serisinde Gecikmeler için Katsayılar
$\hat{\cdot}$	Tahmin edici operatörü
$\hat{\Gamma}_p$	Kovaryans Tahminleri Matrisi
\hat{v}_p	Standartlaştırıcı Terim
$\hat{\sigma}(\cdot)$	Standart Hata Fonksiyonu
L	Olabilirlik fonksiyonu
$\text{Sup}(\cdot)$	Supremum fonksiyonu
$\text{Cov}(\cdot)$	Tahmin edicilerin kovaryans matrisi
∇_{θ}	Parametre uzayı
\hat{t}	Zaman serisinde t. değer için (t-1). değerden faydalanılarak elde edilen öngörü

GİRİŞ

Zaman serisi analizi, kısaca bir zaman serisini modelleyerek temsil ettiği olguyu açıklama biçimidir. Zaman serisi analizi, bir model ortaya koyduğundan dolayı içerisinde sağlıklı öngörü yapmayı da barındırmaktadır. Oluşturulan zaman serisi modeli, her zaman için elde edilen veriye bağımlı kalarak oluşturulur ve bunun veriye uygunluğundan ziyade gerçekte olgunun kendisini açıklaması beklenir. İstatistik teorisinin ve özel olarak zaman serileri analizi teorisinin sırrı büyük bir süper kitleden seçilen göreceli olarak küçük bir parçasından oluşturulan modelin, genellenebilir ve yorumlanabilir sonuçlar çıkarmasıdır. Zaman serileri analizi kapsamında, sonuç olarak oluşturulan modelden elde edilen öngörülerin doğruluğu en önemli ölçüt olmalıdır.

Zaman serilerinin analizinde kullanılan temel yöntemler, otokorelasyon analizi, çapraz korelasyon analizi, spektral analiz ve dalgacık analizidir. Bu yöntemlerden ilk ikisi zaman alanı yöntemleri olarak geri kalanı da frekans alanı yöntemleri olarak sınıflandırılır. Bunun yanında zaman serisi yöntemleri parametrik ve parametrik olmayan olmak üzere ikiye ayrılabilir. Parametrik yaklaşımlar, zaman serisinin altındaki durağan stokastik sürecin küçük miktarda parametre tarafından açıklanabilen belli bir yapısının olduğunu varsaymaktadır. Buna karşılık parametrik olmayan yaklaşımlar, özel bir yapıya sahip olup olmadığına bakılmaksızın sürecin kovaryans matrisini elde etmekten geçmektedir. Bütün bu sayılan yöntemler haricinde istatistik literatüründe bir zaman serisini analiz etmek için aşağıdaki yöntemler de kullanılabilir:

- Otokorelasyon veya spektral yoğunluk fonksiyonu aracılığıyla analiz
- Ölçeklendirilmiş çapraz ve otokorelasyon fonksiyonuyla analiz (Nikolić et al., 2012)
- Frekans alanında serileri incelemek için Fourier Dönüşümü yapmak
- Filtreleme aracılığıyla istenmeyen gürültüleri ortadan kaldırmak
- Temel bileşenler analizi
- Tekil spektrum analizi

- Yapısal modeller:
 - Genel durum uzayı modelleri
 - Gözlenmeyen bileşenler modelleri
- Makine Öğrenme
 - Yapay sinir ağları
 - Destek Vektör Makineleri
 - Bulanık Mantık
- Gizli Markov modelleme
- Kontrol grafikleri
 - Shewhart bireysel kontrol grafikleri
 - CUSUM grafiği
 - EWMA grafiği
- Trentten arındırılmış dalgalanma analizi
- Dinamik Bayezyan ağ
- Zaman Frekans analizi yöntemleri
 - Hızlı Fourier dönüşümü
 - Sürekli dalgacık dönüşümü
 - Kısa zamanlı Fourier dönüşümü
 - Chirplet dönüşümü
 - Kesirsel Fourier dönüşümü
- Kaotik analiz
 - Korelasyon boyutu
 - Tekrarlı grafikler
 - Tekrarlı nitelik grafikleri
 - Lyapunov üstelleri
 - Entropi kodlama
- ARIMA modelleri
- Durum-Uzay modelleri

Genel olarak istatistiksel analiz yöntemleri, ‘‘Sağlamcı’’ ve ‘‘Klasik’’ olmak üzere ayrı ayrı değerlendirilebilir. Bu ayrımın yapılmasındaki temel amaç, istatistiğin gelişiminin veri analizi yöntemlerini nasıl etkilediğini açıklamaktır. Örneğin klasik istatistik ya da başka bir deyişle sıklıkçı istatistik (frequentist statistics), hipotez testlerinin gerçekleştirilmesinde ve parametreler için güven aralıklarının oluşturulmasında önemli bir rol üslenmiştir. Sağlamcı istatistik (robust statistics), varsayımlardan bağımsız olarak etkin istatistiksel tahmin etme teknikleri üzerinde durmaktadır ve diğer istatistiksel yaklaşımlara göre daha yeni bir yaklaşım türüdür.

Yapay sinir ağları, yapay zekâ ile beraber gelişen bir kavramdır. Yapay sinir ağlarıyla ilgili ilk makale, 1943 yılında nörobilimci Warren Mcculloch ve matematikçi Walter Pitts tarafından yazılmıştır (Özaydın, 2009). Bunun sonrasında Donald Hebb 1949 yılında Organization of Behaviour (Davranışın organizasyonu) adlı eseriyle ilk defa öğrenebilen bir yapay sinir ağı modeli oluşturulmuştur (Özaydın, 2009). Bu gelişmeleri 1957 yılındaki Rosenblatt ortaya attığı tekli doğrusal algılayıcı ve 1959 yılında Bernard Widrow ile Marcian Hoff’un ortaya attıkları ADALINE ve MADALINE ağ modellerinin ortaya çıkışı izlemiştir. 1970’lerde her ne kadar yapay sinir ağlarına ilişkin çalışmalar durma noktasına gelse de 1980’lerde bu durum değişmiş ve çok büyük hızda gelişmeler gerçekleşmiştir. Bugünlerde popüler konulardan biri, zaman serileri analizinin yapay sinir ağlarıyla nasıl yapılabileceğidir. Bu çalışmada, istatistiksel yaklaşımların zaman serilerinde uygulanma şekillerini incelenecek ve yapay sinir ağlarıyla karşılaştırılacaktır.

Çalışmanın 2. Bölümü’nde zaman serileri analizinde daha çok uygulama alanı bulan ARIMA modelleri klasik ve sağlamcı olarak incelenmiş; 3. Bölümü’nde yapay sinir ağlarıyla ilgili gelişmeler ve tarihçesi daha detaylı bir şekilde sunulmuş; 4. Bölümü’nde ise geceleme oranlarına, kişi başı harcamalara ve ülkemize giriş yapıp sonra çıkış yapan turistlere ilişkin turizm verileri aracılığıyla ARIMA modelleri ve yapay sinir ağları modelleri arasında kıyaslama yapılmış ve son olarak tartışma bölümü ortaya konulmuştur.

1. BÖLÜM

ZAMAN SERİLERİNDE KULLANILAN ARIMA MODELLERİ

Zaman serileri analizi, istatistik literatüründe yaklaşık olarak yarım asırlık bir alandır. Bu alana ilişkin ilk eser verenlerden ikisi George Edward Pelham Box ve Gwilym Meirion Jenkins'tir. Bu iki istatistikçinin zaman serileri analizi literatürüne katkıları Box-Jenkins yöntemiyle olmuştur. Bu yöntemde modelin tipinin belirlenmesi, parametre tahmini ve son olarak modelin uygunluk testi yapılır.

1.1 ARIMA Modelleri

Box et al. (1994), kitaplarında dört farklı modele yer vermişlerdir. Bu modeller aşağıdaki gibi ifade edilmiştir:

- Otoregresif Model (AR modeli)
- Hareketli Ortalamalar Modeli (MA modeli)
- Otoregresif- Hareketli Ortalamalar Karma Modeli (ARMA modeli)
- Otoregresif- Hareketli Ortalamalar Entegre Karma Modeli (ARIMA modeli)

1.1.1 Otoregresif Model

Otoregresif model (AR(p)) eşitlik (1.1)'deki gibi ifade edilmektedir:

$$y_t = \mu + \sum_{i=1}^p \theta_i y_{t-i} + a_t \quad (1.1)$$

Bu modelin tespiti, otokorelasyon fonksiyonu ve kısmi otokorelasyon fonksiyonuna bakılarak yapılmaktadır. Şayet otokorelasyon fonksiyonunun grafiğinde p. ertelemeden sonra otokorelasyon azalıyor ve kısmi otokorelasyon fonksiyonunda etki hep belirlenmiş sınırlar içinde kalıyorsa bu durumda otoregresif model AR(p) kullanılır.

1.1.2 Hareketli Ortalamalar Modeli

Otoregresif model (MA(q)) eşitlik (1.2)'deki gibi ifade edilmektedir:

$$y_t = \mu + \sum_{i=0}^p \varphi_i a_{t-i} + a_t \quad (1.2)$$

Bu modelin tespiti, otokorelasyon fonksiyonunun ve kısmi otokorelasyon fonksiyonuna bakılarak yapılmaktadır. Şayet kısmi otokorelasyon fonksiyonunun grafiğinde q. ertelemeden sonra otokorelasyon azalıyor ve otokorelasyon fonksiyonunda etki hep belirlenmiş sınırlar içinde kalıyorsa bu durumda hareketli ortalamalar modeli MA(q) kullanılır.

1.1.3 Otoregresif-Hareketli Ortalamalar Karma Modeli

Otoregresif-Hareketli Ortalamalar modeli (ARMA(p,q)) eşitlik (1.3)'deki gibi ifade edilmektedir:

$$y_t = \mu + \sum_{i=1}^q \theta_i y_{t-i} + \sum_{i=0}^p \varphi_i a_{t-i} \quad (1.3)$$

Bu modelin tespiti, otokorelasyon fonksiyonunun ve kısmi otokorelasyon fonksiyonuna bakılarak yapılmaktadır. Şayet otokorelasyon fonksiyonunun grafiğinde p. ertelemeden sonra otokorelasyon ve kısmi otokorelasyon fonksiyonunun grafiğinde q. ertelemeden sonra kabul edilebilir seviyelere düşüyor ise otoregresif-hareketli ortalamalar modeli ARMA(p,q) kullanılır.

1.1.4 Otoregresif-Hareketli Ortalamalar Entegre Karma Modeli

Otoregresif-Hareketli Ortalamalar Entegre Karma modeli (ARIMA(p,d,q)) eşitlik (1.4)'teki gibi ifade edilmektedir:

$$(1 - \sum_{i=1}^p \phi_i B^i)(1 - B)^d y_t = \mu + (1 - \sum_{i=1}^q \theta_i B^i) a_t \quad (1.4)$$

Bu model en genelleştirilmiş model olup, veri için ARMA(p,q)'nin bütün koşullarını sağlamalıdır ve veride kaçınıcı erteleme sonunda durağanlaşma gerçekleştiği incelenmelidir. Durağanlığın gerçekleştiği erteleme d'yi vermektedir ve şayet otokorelasyon fonksiyonunun grafiğinde p. ertelemeden sonra ve kısmi otokorelasyon fonksiyonunun grafiğinde q. ertelemeden sonra düşüş yaşıyorsa otoregresif-hareketli ortalamalar entegre karma modeli ARIMA(p,d,q) kullanılır.

Bu modeller belirlenirken, otokorelasyon ve kısmi otokorelasyon tahminlerini bulmak gereklidir. Bu nedenle, şimdi otokorelasyon fonksiyonunun ve kısmi otokorelasyon fonksiyonunun tahmin edilişi incelenecektir.

1.2 Otokorelasyon ve Kısmi Otokorelasyon Tahmini

Bir zaman serisinin otokorelasyonu, genel olarak zaman serisi içerisinde bir gözlemin başka bir gözlemlerle olan ilişkisini belirtmektedir. Buna karşılık kısmi otokorelasyon fonksiyonu ise zaman serisinin otokorelasyonların kaç tanesi tarafından ifade edildiğini gösterir. Bunlara örnek vermek gerekirse otokorelasyon fonksiyonu p. değerinden sonra görülen bir azalma en kayda değer ilişkilerin p. ertelemeye kadar ortaya çıktığını göstermektedir. Bunun yanında kısmi otokorelasyon q. ertelemeden sonra azalıyor ise otokorele olmuş hata serisinin sayısı q'dur. Çünkü hatalar zaman serisinden doğar. Bu nedenle kısmi otokorelasyon fonksiyonu ile otokorelasyon fonksiyonu, zaman serisini modellemede önemli yer tutar. Bu nedenden ötürü, bu fonksiyonları tahmin etmek önemlidir. Herhangi bir zaman serisi için k. dereceden otokorelasyonun tahmin edicisi eşitlik (1.5)'le hesaplanır (Box et al., 1994):

$$r_k = \frac{\sum_{i=1}^{N-k} (Y_i - \bar{Y})(Y_{i+k} - \bar{Y})}{\sum_{i=1}^N (Y_i - \bar{Y})^2} \quad (1.5)$$

Burada Y_i zaman serisindeki deęerleri, \bar{Y} zaman serisinin ortalaması ve N zaman serisinin büyüklüğünü göstermektedir. Bu tahmin edicinin Bartlett yaklaşımına dayalı varyansı da eşitlik (1.6)'da verilmiştir (Box et al.,1994):

$$\text{var}(r_k) \cong \frac{1}{N} \left(1 + 2 \sum_{v=1}^q \rho_v^2 \right) \quad (1.6)$$

Bu denklemde N zaman serisinin büyüklüğünü, ρ_v teorik otokorelasyonu temsil etmektedir. Eşitlik (1.6)'daki formülü gerçek hayatta teorik otokorelasyon yerine onun tahmin edicisi kullanılabilir. Bunun yanında otokorelasyonu bulmak için, aşağıdaki yöntemler kullanılabilir (Maronna et al., 2006):

- Sağlamcı yöntemlerle filtrelenmiş gözlem deęerlerinden klasik otokorelasyonları ve klasik kısmi korelasyonları kullanmak
- Sağlamcı yöntemlerle k otokorelasyonu hesaplamak ve bunların her birinden m tane kısmi otokorelasyon bulmak

İkinci yöntemde kullanılacak sağlamcı yöntemler için, Ma and Genton (2000)'a bakılabilir. Kısmi otokorelasyon katsayıları sağlamcı tekniklerle tahmin etmek istenirse aşağıdaki yöntemler önerilmektedir (Maronna et al., 2006):

- Sağlamcı tekniklerle filtrelenmiş gözlem deęerleri için Durbin-Levinson algoritmasını kullanmak
- $F\tau$ tahmin teknięiyle elde edilen otokorelasyon tahmin edicileriyle elde etmek

Otokorelasyon ve kısmi otokorelasyon belirlendikten sonra model tahminine geçilebilir. Model tahmin yöntemleri izleyen paragraflarda ele alınmıştır.

1.3 Model Tahmin Yöntemleri

Model tahmin yöntemleri istatistik literatüründe klasik ve sağlamcı tahmin yöntemleri olmak üzere ikiye ayrılır.

1.3.1 Klasik Tahmin Yöntemleri

Klasik tahmin yöntemleri istatistik literatüründe en eski yaklaşımları içermektedir. Bunlardan bazıları en küçük kareler yöntemi, Yule-Walker eşitlikleri ve en çok olabilirlik yöntemidir.

1.3.1.1 En Küçük Kareler Yöntemi ve Yule-Walker Eşitlikleri

En küçük kareler tahmin edicileri, istatistik literatüründeki en sık kullanılan tahmin edicilerden biridir. Bu tahmin ediciler, özellikle regresyon analizinde hatalar normal dağıldığı zaman en çok olabilirlik (EÇO) tahmin edicileriyle aynı olmakta ve bu onları daha ayrıcalıklı kılmaktadır. Ancak EÇO tahmin edicileri zaman serileri analizinde belli durumlarda elde edilemediklerinden ötürü, en küçük kareler yöntemiyle tahmin etmek bu durumlarda daha mantıklıdır. EÇO tahmin yöntemi 1.3.1.2'de ele alınacaktır.

Zaman serileri analizinde en küçük kareler tahmin edicilerinin pek çok hesaplanma yolu vardır. Açık bir şekilde ifade etmek için AR(p) modeli ele alınsın. Şayet y_t durağan kabul edilirse, $\eta = E(y_t)$ olacak şekilde bir η bulunabilmektedir ve bu durumda $\phi(z) = 1 - \sum_{i=1}^p \phi_i Y^i$ ifadesinin bütün kökleri birim çember dışında kaldığı varsayılırsa η için geçerli formül eşitlik (1.7)'deki gibidir:

$$\eta = \frac{\mu}{1 - \sum_{i=1}^p \phi_i} \quad (1.7)$$

Böyle bir durumda $\hat{\mu}$ örneklem ortalaması olarak alınırsa $\hat{\eta} = \frac{\hat{\mu}}{1 - \sum_{i=1}^p \phi_i}$ olarak hesaplanır. Hatta genel olarak herhangi bir AR(p) zaman serisi modelini eşitlik (1.8) ile ifade etmek mümkündür:

$$y = G\beta + a \quad (1.8)$$

Burada

$$\beta' = (\phi_1, \dots, \phi_p, \mu) \quad (1.9)$$

$$a' = (a_{p+1}, \dots, a_{(2p+1)}) \quad (1.10)$$

$$y' = (y_{p+1}, \dots, y_{(2p+1)}) \quad (1.11)$$

olmaktadır ve böyle durumda $G = \begin{bmatrix} y_p & y_{p-1} & \dots & y_1 & 1 \\ y_{p+1} & y_p & \dots & y_2 & 1 \\ \vdots & \vdots & \dots & \vdots & \vdots \\ y_{T-1} & y_{T-2} & \dots & y_{T-p} & 1 \end{bmatrix}$ olmak üzere en küçük

kareler tahmin edicisi $\hat{\Phi} = (G'G)^{-1}G'y$ olarak ifade edilir. Büyük örneklem için Yule Walker tahmin edicileri normal dağılıma yakınsar (Brockwell and Davis, 2002). Herhangi bir büyük ölçekli zaman serisi verisine AR(p) modeli uydurulursa bu durumda otoregresif model parametreleri için normal dağılıma yakınsaklıktan faydalanarak (1.12)'deki eşitsizlik ile ifade edilen güven aralıklarını oluşturabileceği söylenebilir (Brockwell and Davis, 2002):

$$\frac{(\phi_p - \phi)' \hat{\Gamma}_p (\phi_p - \phi)}{\hat{v}_p} \leq \frac{\chi_{(1-\alpha)}^2(p)}{n} \quad (1.12)$$

Bu formülde $\hat{\Gamma}_p [\hat{\gamma}(i-j)]_{i,j=1}^{p+1}$ kovaryans tahminleri matrisi, $\hat{\rho}_m$ otokorelasyonlar matrisi olmak üzere:

$$\hat{v}_p = \hat{\gamma}(0) [1 - \hat{\rho}_m' \hat{\Gamma}_m^{-1} \hat{\rho}_m] \quad (1.13)$$

şeklinde ifade edilmektedir. Genel olarak Yule-Walker denklemler sistemi herhangi bir ARMA(p,q) modeli için aşağıdaki biçimde ifade edilir:

$$\hat{y}(k) - \phi_1 \hat{y}(k-1) - \dots - \phi_p \hat{y}(k-p) = \sigma^2 \sum_{j=k}^q \theta_j \psi_{j-k}, \quad (1.14)$$

Burada θ_j , MA katsayılarını temsil etmek üzere $\psi_j = \theta_j/\phi_j$ olmaktadır. Bu tarz denklem sistemlerini çözmek için algoritmalar geliştirilmiştir. Bu algoritmalar Burg algoritması, Durbin-Levinson algoritması, Hannan-Rissanen algoritması ve hatalar algoritmasıdır. Durbin-Levinson algoritması ilk olarak Levinson (1947) tarafından Toeplitz sistemi çözmek için ortaya atılmıştır. Daha sonradan Durbin (1960) tarafından geliştirilerek en hızlı çalışan algoritmalarından biri olmuştur. Bu algoritma için en ilginç özellik, $n \times n$ tipinde bir Toeplitz matrisine sahip bir denklem sistemi içerisinde n yeterince küçük olursa ($n \leq 256$), algoritmanın herhangi bir süper hızlı algoritmadan daha hızlı çalışmasıdır. Durbin (1960) bu algoritmayı geliştirirken zaman serileri analizinde uygulamayı da başarmıştır. Durbin-Levinson algoritması AR(p) tipi modellerinde öngörü kolaylığı sağlayan bir algoritmadır. Algoritmanın temelinde, AR(p) modelini eşitlik (1.15) ile ifade etmek bulunmaktadır (Maronna et al., 2006):

$$\hat{y}_{t,m} = \phi_{m,1} y_{t-1} + \dots + \phi_{m,m} y_{t-m} \quad (1.15)$$

Burada $\hat{y}_{t,m}$, durağan seri y_t 'nin temsili ve $\phi_{m,i}$ kısmi otokorelasyon katsayılarını temsil etmektedir. Burada $\phi_{m,i}$ için formül eşitlik (1.16) ile verilmiştir:

$$\phi_{m,i} = \begin{cases} \frac{\rho(m) - \sum_{i=1}^{m-1} \rho(i) \phi_{m-1,i}}{1 - \sum_{i=1}^{m-1} \rho(i) \phi_{m-1,i}} & i = m \\ \phi_{m-1,i} - \left(\frac{\rho(m) - \sum_{i=1}^{m-1} \rho(i) \phi_{m-1,i}}{1 - \sum_{i=1}^{m-1} \rho(i) \phi_{m-1,i}} \right) \phi_{m-1,m-i} & 1 \leq i \leq m-1 \end{cases} \quad (1.16)$$

Durbin-Levinson algoritmasında eşitlik (1.15)'de bütün esas değerler tahminlerine dönüştürülüyor. Bu noktada:

$$\phi_{1,1} = \operatorname{argmin}_{\zeta} \sum_{t=2}^T (y_t - \zeta y_{t-1})^2 \quad (1.17)$$

olarak bulunur ve geri kalan kısmi otokorelasyonlar bir önceki eşitlikte olduğu gibi hesaplanır. Böylece:

$$\begin{aligned} \phi_{m,m} = \operatorname{argmin}_{\zeta} \sum_{t=m+1}^T (y_t - (\hat{\phi}_{m-1,1} - \zeta \hat{\phi}_{m-1,m-1})y_{t-1} - \dots - (\hat{\phi}_{m-1,m-1} - \\ \zeta \hat{\phi}_{m-1,1})y_{t-m+1} - \zeta y_{t-m})^2 = \operatorname{argmin}_{\zeta} \sum_{t=m+1}^T \hat{u}_{t,m}(\zeta)^2 \end{aligned} \quad (1.18)$$

elde edilerek tahminler oluşturulur. Durbin-Levinson algoritması iki nedenden ötürü sağlam kabul edilmemektedir:

- Durbin-Levinson algoritmasının kayıp fonksiyonu sınırlı değildir.
- $\widehat{u}_{t,m}(\zeta)$ 'ler zaman serilerindeki sapan değerlerden etkilenmektedirler.

Bu nedenlerden ötürü Maronna et al. (2006), sağlamcı bir Durbin-Levinson algoritması geliştirmiştir. Bunun için $\hat{\phi}_{m,m}$ eşitlik (1.19) ile tanımlanmıştır:

$$\hat{\phi}_{m,m} = \operatorname{argmin}_{\zeta} \hat{\sigma}(\widehat{u}_{m+1,m}(\zeta), \dots, \widehat{u}_{T,m}(\zeta)) \quad (1.19)$$

Burada $\hat{\sigma}(\cdot)$ fonksiyonu ölçek parametresinin sağlam bir tahmin edicisini temsil etmektedir. Maronna et al. (2006)'a göre $\widehat{\phi}_{m,m}$ 'i bu şekilde tanımlamak hem sapan değerlerin etkisini azaltmakta hem de kayıp fonksiyonunu kısıtlamaktadır. $\hat{\sigma}(\cdot)$ fonksiyonunun seçimi ise bayağı zorludur. Örneğin Hössjer (1992)'e göre kırılma noktası 0,5 olan sağlamcı bir şekilde filtrelenmiş otoregresif model tahmin edicisi olan S-tahmin edicisiyle bu fonksiyon ifade edilirse büyük örneklemelerde etkinliği 0.33'ten yüksek olmayan bir tahmin ediciyle çalışılmış olur. Bu da analizi olumsuz yönde etkiler.

Bu algoritma sonucunda büyük örneklem için elde edilen tahmin ediciler normal dağılıma yakınsamaktadır. Söz konusu algoritmalar dışında hatalar algoritması vardır ve aşağıdaki adımlarla hesaplanmaktadır (Brockwell and Davis, 2002):

$$1. \begin{bmatrix} \hat{\theta}_{m,q+1} \\ \vdots \\ \hat{\theta}_{m,q+p} \end{bmatrix} = \begin{bmatrix} \hat{\theta}_{mq} & \dots & \hat{\theta}_{m,q+1-p} \\ \vdots & \hat{\theta}_{mq} & \vdots \\ \hat{\theta}_{m,q+p-1} & \dots & \hat{\theta}_{mq} \end{bmatrix} \begin{bmatrix} \hat{\phi}_1 \\ \vdots \\ \hat{\phi}_p \end{bmatrix} \text{ (Burada } j < q \text{ için } \theta_j = 0 \text{ ve } j \geq q$$

için $\theta_j = 1$ olmak üzere $\hat{\theta}_{mj} = \theta_j + \sum_{i=1}^{\min(j,p)} \phi_i \hat{\theta}_{mj-i}$ şeklinde bulunur.)

$$2. \hat{\phi} \text{ hesaplandıktan sonra formülüyle } \hat{\theta}_j = \hat{\theta}_{mj} - \sum_{i=1}^{\min(j,p)} \hat{\phi}_i \hat{\theta}_{m,j-i} \text{ 'yı hesapla.}$$

$$3. \hat{\sigma}^2 = \frac{\sum_{i=1}^n (Y_t - \hat{Y}_t)}{nE[(W_t - W_{t-1})^2]} \text{ hesaplanır (} W_t \text{ için Ansley(1979)'e bakınız.).}$$

Bunlar haricinde Hannan-Rissanen algoritması ile AR(p) modeli tahmini yapılabilir. Bu algoritmanın işleyişi aşağıdaki iki maddeden oluşmaktadır (Brockwell and Davis, 2002):

$$1. \hat{\phi}_{ij} \text{ Yule-Walker tahmin edicilerinden faydalanılarak } \hat{Z}_t = Y_t - \hat{\phi}_{m1} Y_{t-1} - \dots - \hat{\phi}_p Y_{t-p} \text{ hesaplanır.}$$

$$2. S(\beta) = \sum_{t=m+1+q}^n (Y_t - \hat{\phi}_{m1} Y_{t-1} - \dots - \hat{\phi}_{m1} Y_{t-p} - \hat{\theta}_{m1} \hat{Z}_{t-1} - \dots - \hat{\theta}_{mp} \hat{Z}_{t-p})^2 \text{ 'yi minimize et}$$

Bunun sonucunda elde edilen $\beta = (\phi', \theta')'$ tahmin edicisi eşitlik (1.20) ile ifade edilir:

$$\hat{\beta} = (Z'Z)^{-1} Z' X_n \quad (1.20)$$

Burada $X_n = (X_{m+1+q}, \dots, X_n)'$ ve Z matrisi ise eşitlik (1.21)'deki gibi ifade edilen $(n-m-q) \times (p+q)$ tipinde bir matristir:

$$Z = \begin{bmatrix} X_{m+q} & \cdots & X_{m+q+1-p} \hat{Z}_{m+q} & \cdots & \hat{Z}_{m+1} \\ \vdots & \cdots & \vdots & \vdots & \vdots \\ X_{n-1} & \cdots & X_{n-p} \hat{Z}_{n-1} & \cdots & \hat{Z}_{n-q} \end{bmatrix} \quad (1.21)$$

Bu durumda varyansın Hannan-Rissanen tahmin edicisi eşitlik (1.22)'deki gibidir:

$$\hat{\sigma}_{HR}^2 = \frac{S(\hat{\beta})}{n-m-q} \quad (1.22)$$

Bu noktadan sonra oluşturulan modellerin güvenilirliğini araştırmak üzere birtakım kriterler ortaya konması gerekmektedir. Brockwell and Davis (2002), bir yaklaşım olarak eşitlik (1.23)'teki AICC (Akaike koşullu bilgi kriteri) istatistiğini kullanmak gerektiğini belirtmektedir:

$$AICC = -2\ln L(\phi_p, \theta_q, S(\phi_p, \theta_q)/n) + 2(p+q+1)n/(n-p-q-2) \quad (1.23)$$

Yenilikler algoritması, sonlu ikincil momentlere sahip olan herhangi bir zaman serisine uygulanabilmektedir (Brockwell and Davis, 2002). Algoritmanın temelinde, genelliği bozmayan eşitlik (1.24)'teki varsayım yapılmaktadır:

$$E(X_i X_j) = \kappa(i, j) \quad (1.24)$$

Buradan yola çıkılarak Akkaya and Tiku (2004), modifiye EÇÖ tahmin edicileri üzerinde birtakım oynamalar yapıldığında en küçük kareler tahmin edicilerin elde edileceğini belirtmişlerdir. Akkaya and Tiku (2004), aynı eserlerinde gerçekleştirilen simülasyonlarla en küçük kareler yönteminin kendi geliştirdikleri modifiye EÇÖ tahmin edicilerine kıyasla etkinliğinin az olduğunu göstermiştir. Bunun yanında Akkaya and Tiku (2004), en küçük kareler tahmin edicilerinin beklenen değerlerinin ve varyanslarının hesaplanmasının zor olduğunu belirtmiş ancak asimptotik olarak en küçük kareler tahmin edicileri yansızdırlar ve varyanslarını hesaplamak mümkündür.

Örneğin $y_t = \phi y_{t-1} + a_t$ AR(1) modeli için ϕ ile $\sigma = \sqrt{Var(a_t)}$ 'nın kovaryans matrisi eşitlik (1.25)'teki gibi ifade edilir (Akkaya and Tiku, 2004):

$$Cov(\hat{\phi}, \hat{\sigma}) = \begin{bmatrix} \frac{(1-\phi^2)}{n} & 0 \\ 0 & \left(1 + \frac{\lambda_4}{2}\right) \frac{\sigma^2}{2n} \end{bmatrix} \quad (1.25)$$

Burada n zaman serisindeki gözlem sayısını ve λ_4 basıklık değerini belirtmektedir. En küçük kareler yönteminin zaman serileri analizinde, kötü yanlarına karşılık iyi yanları da bulunmaktadır. Daha önceden belirtildiği gibi bu tahmin edicilerin hesaplanması oldukça kolay formülleri vardır. Bunun yanında bu tahmin edici, zaman serisindeki trendi oldukça iyi yakalamaktadır. Bütün bu sayılanların yanında nokta tahmini yapmak için kullanılabilecek en pratik tahmin ediciler, en küçük kareler tahmin edicileridir. Bunun nedeni hata kareler toplamını en küçük yapan tahmin ediciler olmasıdır.

1.3.1.2 En Çok Olabilirlik Yöntemi

En Çok Olabilirlik (EÇO) tahmin yöntemi, istatistiğin en popüler tahmin yöntemlerinden birisidir. EÇO tahmin yöntemi, 1912-1922 yılları arasında R.A. Fisher tarafından geliştirilmiştir (Pfanzagl, 1994). Bu yöntemde, olabilirlik fonksiyonu adı verilen eşitlik (1.26)'daki bir fonksiyon hesaplanır:

$$L = \prod_{i=1}^n f(x_i|\theta) \quad (1.26)$$

Burada x_i aynı dağılıma sahip birbirinden bağımsız rassal değişkenleri temsil etmektedir. Yöntemde bu fonksiyon elde edildikten sonra, genelde LogL fonksiyonu hesaplanır ve parametrelere göre türevler sıfıra eşitlenir ve eşitlikler çözülür. Ancak bu noktada zaman serisi analizi açısından, bu yöntem bazı durumlarda problem çıkartmaktadır. Örneğin zaman serisi modelinde a_t 'ler gamma dağılımına sahip olduğunda, bu denklemlerin çözülemedikleri görülmüştür (Akkaya and Tiku, 2004).

Zaman serisi analizi dışında EÇÖ yönteminin kullanılmadığı durumlar da mevcuttur ancak yine de en çok olabilirlik yönteminin faydalı yönleri fazladır. Bu faydalı özelliklerin birkaçı aşağıda ifade edilebilir:

- Teorik olarak EÇÖ tahmin edicileri olasılıksal olarak gerçek değere yakınsar.
- Örneklem büyüklüğü arttıkça, EÇÖ tahmin edicilerinin ortak dağılımı çok boyutlu normal dağılıma yakınsar ve bu dağılımda ortalama vektörü tahmin edilmek istenen parametrelerin gerçek değerlerini içermekte ve kovaryans matrisi Fisher bilgi matrisinin tersidir.
- EÇÖ tahmin edicisinin hata kareler ortalaması, tüm asimptotik olarak yansız tahmin edicilerinkinden küçüktür. Örneklem büyüklüğü arttıkça, Cramér-Rao en küçük sınırını sağlar.
- Tahmin edicilerde düzeltme yapılsa bile etkinliği yüksektir.

Bu özelliklerin sağlanması için birtakım kurallar vardır (Engle and McFadden, 1999). Bunun dışında:

1. θ 'nın değişik değerleri için tahmin edicilerin farklı dağılımlara sahip olması gerekmektedir. Şayet bu sağlanmazsa, bunlar arasındaki fark hiçbir şekilde ifade edilemez.
2. Olabilirlik fonksiyonunun tek maksimumu olmalı ve bu maksimumu elde etmek başka bir noktada mümkün olmamalı. Bu özellik, maksimum noktasının makul bir komşuluğu alınarak hafifletilebilir.
3. $\log f(\cdot|\theta)$ fonksiyonu θ 'nın uzayında sürekli olmalıdır
4. Öyle bir integrallenebilir bir $D(\cdot)$ fonksiyonu vardır ki eşitlik (1.27)'deki eşitsizliği bütün θ değerleri için sağlar:

$$|\ln f(x|\theta)| < D(x) \quad (1.27)$$

Bu koşulların üstüne aşağıdaki koşullar da sağlanırsa bu durumda asimptotik normallik sağlanır (Engle and McFadden, 1994):

1. $\theta_0 \in \text{ic}(\Theta)$ ($\text{ic}(\Theta)$ parametre uzayının iç kısmını temsil etmektedir.)
2. $f(\cdot|\theta) > 0$ ve θ_0 'ın herhangi bir N komşuluğunda iki kez türevlenebilir durumda olmalıdır.
3. $\int \sup_{\theta \in N} \|\nabla_{\theta} f(x|\theta)\| dx < \infty$ ve $\int \sup_{\theta \in N} \|\nabla_{\theta\theta} f(x|\theta)\| dx < \infty$;
4. $I = E[\nabla_{\theta} f(x|\theta_0)\nabla_{\theta} f(x|\theta_0)']$ var ve tekil değil;
5. $E[\sup_{\theta \in N} \|\nabla_{\theta\theta} f(x|\theta)\|] < \infty$.

EÇÖ yönteminde tahmin edici, gözlenen veride en büyük olasılığı verecek şekilde formülleştirilir. EÇÖ tahmin edicisi bu nedenle verideki birtakım dönüşümlere karşı duyarsızdır. Bunun yanında farklı dağılımlar arasında ilişkiler varsa bu durumda EÇÖ tahmin edicileri bu ilişkilerden faydalanabilir. EÇÖ tahmin edicileri istatistik literatüründe önemli bir yere sahiptir. Bunun yanında bir bilgisayar aracılığıyla hesap edilmesi kolaylaştırılmıştır.

EÇÖ tahmin edicilerinin iteratif olarak hesaplanmasının kolaylaştırmasındaki en büyük etken EM (Expectation Maximization) algoritmasının ortaya çıkmasıdır. İlk defa Dempster, Laird and Rubin(1977) tarafından geliştirilen bu algoritma istatistik literatüründe ayrıcalıklı yere sahiptir. (EM algoritmasının yakınsaklığına ilişkin inceleme için Wu(1983)'ya bakılabilir.) Brockwell and Davis (2002) algoritmayı aşağıdaki şekilde tanımlamışlardır:

1. E-adımı. $Q(\theta|\theta_{(i)}) = E_{\theta_{(i)}} [l(\theta; X, Y)|Y]$ hesapla
2. M-adımı $Q(\theta|\theta_{(i)})$ ifadesini büyütecek θ 'yı bul

M-adımında bulduğumuz değer $\theta_{(i+1)}$ olarak atanır ve belli bir iterasyon sayısını bulunca da algoritma sonlanır. EM-algoritması pek çok konuda kullanışlıdır. Buna verilebilecek örnekler arasında zaman serileri öngörüsü haricinde kayıp veri analizi,

sınıflandırma ve kümeleme analizleri bulunmaktadır. Bunun yanında EM algoritması üzerinde oynamalar yapılmıştır. Örneğin EM algoritması, bir makalede ECM (Expectation Conditional Maximization) algoritmasına dönüşmüş başka bir makalede ECME (Expectation Conditional Maximization Either) algoritmasına çevrilmiştir. Bu iki algoritmadan ECM algoritmasında M-adımında koşullu olabilirlik fonksiyonları devreye sokulmuş böylece her ne kadar daha çok adımda işlem yapılsa da geçirilen hesap vakti açısından bir tasarrufa gidilmiştir (Everitt, 2006). Daha sonra ECM algoritmasındaki tam veriye ait kısıtlı beklenen olabilirlik fonksiyonu üzerinde optimizasyon gerçekleştirilen adımlar yerine ona karşılık gelen gerçek olabilirlik fonksiyonun üzerinde optimizasyon yapan adımlara sahip ECME algoritması geliştirilmiştir. Hem EM algoritmasından ve ECM algoritmasından daha hızlı hem de ECM gibi yaklaşımla kısıtlandırılmadan gerçek olabilirlik fonksiyonu üzerinden işlem yapan bir algoritma ortaya çıkmıştır (Everitt, 2006).

Brockwell and Davis (2002), kitabında eşitlik (1.28)'deki AR(2) modelini incelemiştir:

$$W_t - 1.0415W_{t-1} + 0.2494W_{t-2} = Z_t \quad (1.28)$$

Bu denklemde Z_t beyaz gürültülü 0 ortalamalı ve 0,4790 varyanslı bir seriyi ve W_t ise ortalama ile düzeltilmiş bir göl seviyesi verisini temsil etmektedir. Brockwell and Davis (2002), veriden 10 tane değeri atarak gerçekleştirdikleri analizde kullandıkları EM algoritması sadece 4 iterasyonda gerçek değerlere yaklaşmayı başarmıştır. Bunu gerçekleştirirken veri tamamlama işlemini eşitlik (1.29)'daki amaç fonksiyonuna göre yapmışlardır:

$$\sum_j (W_{t+j} - \hat{\phi}_1^{(1)}W_{t+j-1} - \hat{\phi}_2^{(1)}W_{t+j-2}) \quad (1.29)$$

Buna göre çözüm eşitlik (1.30)'da verilmiştir:

$$\hat{W}_t = \frac{\hat{\phi}_2^{(1)}(W_{t-2}+W_{t+2})+(\hat{\phi}_1^{(1)}-\hat{\phi}_1^{(1)}\hat{\phi}_2^{(1)})(W_{t-1}+W_{t+1})}{1+(\hat{\phi}_1^{(1)})^2+(\hat{\phi}_2^{(1)})^2} \quad (1.30)$$

Bu çözümler sayesinde algoritmada veriye bağlı olabilirlik fonksiyonu sürekli düşmüştür ve algoritma öngöründe başarılı olmuştur. EM algoritması olmadan EÇÖ tahmin edicileri elde etmek için hataların dağılımı için varsayımda bulunmak gerekmektedir. Örneğin hataların dağılımının normal olduğu varsayılırsa, olabilirlik fonksiyonu eşitlik (1.31) ile ifade edilir (Brockwell and Davis, 2002):

$$L = \frac{1}{\sqrt{(2\pi)^n \det(\Gamma_n)}} \exp\left\{-\frac{1}{2}Y_n' \Gamma_n Y_n\right\} \quad (1.31)$$

Burada Γ_n kovaryans matrisini ifade etmektedir. Bu noktada C_n 'yi eşitlik (1.32) ile ifade edilirse:

$$C_n = \begin{bmatrix} 1 & 0 & 0 & \dots & 0 \\ \theta_{11} & 1 & 0 & \dots & 0 \\ \theta_{22} & \theta_{21} & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & 0 \\ \theta_{n-1,n-1} & \theta_{n-1,n-2} & \dots & \dots & 1 \end{bmatrix} \quad (1.32)$$

Böyle bir durumda kovaryans matrisinin $D_n = \text{diag}(a_1, \dots, a_{n-1})$ olmak üzere eşitlik (1.33) ile ifade edilir (Brockwell and Davis, 2002):

$$\Gamma_n = C_n D_n C_n' \quad (1.33)$$

Buradan:

$$Y_n' \Gamma_n Y_n = \sum_{j=1}^n \frac{(Y_j - \hat{Y}_j)^2}{a_{j-1}} \quad (1.34)$$

$$\text{Det}(\Gamma_n) = (\text{Det}(C_n))^2 (\text{det}(D_n)) = a_1 \dots a_{n-1} \quad (1.35)$$

elde edilir. Böylece olabilirlik fonksiyonu eşitlik (1.36) ile ifade edilir:

$$L = \frac{\exp\left(\frac{-\sum_{j=1}^n \frac{(Y_j - \hat{Y}_j)^2}{a_{j-1}}}{2}\right)}{\sqrt{(2\pi)^n a_1 \dots a_{n-1}}} \quad (1.36)$$

Bu olabilirlik fonksiyonunda

$$\hat{X}_{j+1} = \begin{cases} \sum_{j=1}^n \theta_{nj} (Y_{n+1-j} - \hat{Y}_{n+1-j}) & 1 \leq n < m \\ \sum_{i=1}^p \phi_i Y_{n+1-i} + \sum_{j=1}^q \theta_{nj} (Y_{n+1-j} - \hat{Y}_{n+1-j}) & n \geq m \end{cases} \quad (1.37)$$

olmaktadır. Yenilikler algoritması aracılığıyla θ_{nj} ve r_j elde edilirse aşağıdaki yedek olabilirlik fonksiyonu eşitlik (1.38)'deki gibidir:

$$L = \frac{\exp\left(\frac{-\sum_{j=1}^n \frac{(Y_j - \hat{Y}_j)^2}{r_{j-1}}}{2}\right)}{\sqrt{(2\pi\sigma^2)^n r_1 \dots r_{n-1}}} \quad (1.38)$$

Buradan hataların varyansının EÇÖ tahmin edicisi eşitlik (1.39)'dan elde edilir.

$$\hat{\sigma}^2 = \sum_{j=1}^n \frac{(Y_j - \hat{Y}_j)^2}{nr_{j-1}} \quad (1.39)$$

ARMA modelindeki katsayıların tahmin edicileri $\hat{\phi}$ ve $\hat{\theta}$ eşitlik (1.40)'taki amaç fonksiyonu minimize edilerek bulunur (Brockwell and Davis, 2002):

$$F = \log\left(\sum_{j=1}^n \frac{(Y_j - \hat{Y}_j)^2}{nr_{j-1}}\right) + \frac{\log\left(\prod_{j=1}^n r_{j-1}\right)}{n} \quad (1.40)$$

1.3.2 Sağlamcı İstatistiksel Yöntemler ve Zaman Serilerindeki Kullanımı

Sağlamcı istatistiksel yöntemlerin geniş bir tarihi bulunmaktadır. Sağlamcı kelimesinin çevrildiği orjinal kelime olan robust, İngilizce sert, katı, koşullar karşısında dimdik durabilen gibi anlamlara gelmiştir ve bugün sağlamcı istatistiksel teknikler denince akla varsayımlar olmadan da etkili olabilen istatistiksel teknikler gelmektedir. Her ne kadar böyle istatistiksel teknikler bazı durumlarda mümkün olmasa da, bu konuda önemli aşamalar kaydedilmiştir. Stigler (1973), makalesinde sağlamcı istatistiksel tekniklerin ilk olarak 1886'da esas mesleği astronom olan Simon Newcomb tarafından yazılan makaleyle gündeme geldiğini belirtmiştir. Simon Newcomb 1886'daki makalesinde uzun kuyruklu dağılıma sahip olan merkür geçiş verisinin konum parametresini tahmin etmek için ilk kez normal dağılımların karma modelini kullanmış ve örneklem ortalamasından daha etkili bir istatistik ortaya koymuştur. Bunun yanında Stigler (1973), teorik çalışma olarak medyanın dağılımını çıkartan Laplace'ın çalışmasını referans olarak göstermiştir. 1960'larda ve 1970'lerin başlarında meydana gelen gelişmelerle sağlamcı istatistik büyük bir ilerleme katetmiştir (Maronna et al., 2006). Bu gelişimin başını çeken makaleler, Tukey (1960, 1962), Huber (1964, 1967), Hampel (1971, 1974), Tiku (1967a, 1967b, 1968a, 1968b, 1968c, 1970, 1973) ve Tiku and Suresh (1992) olmuştur.

1.3.2.1 Modifiye En Çok Olabilirlik Yöntemi

Temeli Tiku (1967a, 1967b, 1968a, 1968b, 1968c, 1970, 1973) ve Tiku and Suresh (1992)'in makalelerine dayanan bu yöntem, EÇO yönteminin uygulanmasına ilişkin zorluklarını telafi etmek için ortaya atılmıştır. Bu yöntemin en ilginç ve en önemli yanı, belli koşullar altında elde edilen modifiye en çok olabilirlik tahmin edicilerinin (MEÇO) EÇO tahmin edicileriyle asimptotik olarak eşdeğer olmasıdır (Akkaya and Tiku, 2004). Bu özelliğinden ötürü, MEÇO tahmin edicileri EÇO tahmin edicilerinin bütün özelliklerini elde etmektedir (Akkaya and Tiku, 2004).

Zaman serileri analizinde basit bir otoregresif model eşitlik (1.41) ve (1.42) ile ifade edilebilmektedir:

$$y_t = \mu + \delta x_t + e_t \quad (1.41)$$

$$e_t = \phi e_{t-1} + a_t \quad (1.42)$$

Burada y_t gözlenen rassal değişken y 'nin t . zamanda aldığı değer, x_t stokastik olmayan değişkenin önceden belirlenmiş değeri, a_t hata olarak adlandırılmaktadır. Burada a_t 'ler birbirinden bağımsız olarak aynı dağılıma sahip olmaktadır (Akkaya and Tiku, 2004). Eşitlik (1.41) ve (1.42)'deki formülün bir alternatifi aşağıdaki şekildedir:

$$y_t - \phi y_{t-1} = \mu + \delta(x_t - x_{t-1}) + a_t \quad (1.43)$$

Bu formüle göre, şayet $\phi = 0$ olursa elde edilecek denklem standart bir doğrusal regresyon denklemi olmaktadır ve $\phi \neq 0$ ve $\delta = 0$ olursa elde AR(1) modeli kalmaktadır. Burada y_0 için iki farklı model geliştirilebilmektedir (Vinod and Shenton, 1996). Bu modeller aşağıdaki şekilde ifade edilir:

- y_0 sabit(Model A)
- y_0 stokastik bir değişken(Model B)

Esnekliğinden ötürü y_0 için model B benimsenebilir. a_t 'ler gamma dağılımına sahip olduğunda parametrelerin EÇÖ tahmin edicileri elde edilmektedir (Akkaya and Tiku, 2004). Bu durumda MEÇÖ tahmin edicileri Pearson and Hartley (1972) tarafından elde edilen $t_{(i)}$ değerlerinden yola çıkılarak (1.44) - (1.62) eşitlikleriyle ifade edilebilmektedir:

$$\hat{\mu} = \bar{v} - \delta \bar{u} - (\Delta/m)\hat{\sigma} \quad (1.44)$$

$$\hat{\delta} = G - H\hat{\sigma} \quad (1.45)$$

$$\hat{\phi} = \frac{[A - \hat{\sigma} [\sum_{i=1}^n (\Delta_i - (\Delta/m)\beta_i)(y_{[i]} - \delta x_{[i]})]]}{[\sum_{i=1}^n \beta_i (y_{[i]} - \delta x_{[i]})^2 - (\frac{1}{m})[\sum_{i=1}^n \beta_i (y_{[i]} - \delta x_{[i]})]^2]} \quad (1.46)$$

$$\alpha_i = 2/t_{(i)} \quad (1.47)$$

$$\beta_i = 1/t_{(i)}^2 \quad (1.48)$$

$$A = \sum_{i=1}^n \beta_i (y_{[i-1]} - \hat{\delta}x_{[i-1]})(y_{[i]} - \hat{\delta}x_{[i]}) - T \quad (1.49)$$

$$T = \left(\frac{1}{m}\right) \left[\sum_{i=1}^n \beta_i (y_{[i]} - \hat{\delta}x_{[i]})\right] \left[\sum_{i=1}^n \beta_i (y_{[i-1]} - \hat{\delta}x_{[i-1]})\right] \quad (1.50)$$

$$\hat{\sigma} = \left(-B + \sqrt{B^2 + 4nC}\right) / 2\sqrt{n(n-3)} \quad (1.51)$$

$$m = \sum_{i=1}^n \beta_i \quad (1.52)$$

$$\Delta_i = \alpha_i - \frac{1}{k-1} \quad (1.53)$$

$$\Delta = \sum_{i=1}^n \Delta_i \quad (1.54)$$

$$\bar{v} = \frac{\sum_{i=1}^n \beta_i v_{[i]}}{m} \quad (1.55)$$

$$v_{[i]} = y_{[i]} - \hat{\phi}y_{[i-1]} \quad (1.56)$$

$$\bar{u} = \frac{\sum_{i=1}^n \beta_i u_{[i]}}{m} \quad (1.57)$$

$$u_{[i]} = x_{[i]} - \hat{\phi}x_{[i-1]} \quad (1.58)$$

$$G = \frac{\sum_{i=1}^n \beta_i (u_{[i]} - \bar{u})v_{[i]}}{\sum_{i=1}^n \beta_i (u_{[i]} - \bar{u})^2} \quad (1.59)$$

$$H = \frac{(\sum_{i=1}^n \Delta_i (u_{[i]} - \bar{u}))}{\sum_{i=1}^n \beta_i (u_{[i]} - \bar{u})^2} \quad (1.60)$$

$$B = (k-1) \sum_{i=1}^n \Delta_i \left[v_{[i]} - \bar{v} - G(u_{[i]} - \bar{u})^2 \right] \quad (1.61)$$

$$C = (k-1) \sum_{i=1}^n \beta_i \left[v_{[i]} - \bar{v} - G(u_{[i]} - \bar{u})^2 \right] \quad (1.62)$$

Bu eşitliklerde k gamma dağılımının konum parametresi, n zaman serisinde birim sayısıdır. Bu tahmin ediciler karmaşık formüllere sahip olsalar da varyanslarının hesaplanması kolaydır. MEÇÖ tahmin edicileri EÇÖ tahmin edicileriyle asimptotik olarak eşdeğer olduklarından ötürü ikisinin de varyansı Fisher bilgi matrisinin tersiyle bulunur (Akkaya and Tiku, 2004). Burada bir başka ilginç nokta $t_{(i)}$ 'ler hep pozitif olduklarından ötürü β_i 'ler de pozitif olmakta ve standart sapmanın MEÇÖ tahmin edicisi hep pozitif değerler almaktadır ancak aynı durum EÇÖ tahmin edicisi için geçerli değildir. Pratikte standart sapmanın EÇÖ tahmin edicisi negatif değerler

alabilmektedir (Akkaya and Tiku, 2004). MEÇO tahmin edicileri dağılımlara göre değişik formüllerle ifade edilebilmektedir. Bu sebepten Akkaya and Tiku (2004) eserlerinde üç farklı dağılımlar ailesi tanımlamış ve bunun üzerinden MEÇO tahmin edicilerinin formüllerini çıkarmışlardır. Tanımlanan dağılımlar aşağıdaki gibidir:

- Uzun kuyruklu simetrik dağılımlar ailesi
- Kısa kuyruklu simetrik dağılımlar ailesi
- Çarpık dağılımlar ailesi

Bu dağılım ailelerine göre, değişik tahmin ediciler çıkarabilmektedir. Bunlara göre çıkartılan tahmin ediciler, çok kullanışlı özelliklere sahiptir. Bu durumda tahmin edicilerin kullanılmalrı, analizde çok önemli sonuçlar elde etmekte önemli rol üstlenebilir. Örneğin a_t terimi, aşağıdaki olasılık yoğunluk fonksiyonuna sahip kısa kuyruklu simetrik bir dağılıma sahip olduğu varsayılınsın:

$$f(a) = \frac{C_1}{\sigma} \left\{ 1 + \frac{\lambda a^2}{2r\sigma^2} \right\}^r \frac{e^{-a^2/2\sigma^2}}{\sqrt{2\pi}} \quad (1.63)$$

(1.63) eşitliğinde $d < r$ olmak üzere $C_1 = \frac{1}{\sum_{j=0}^r \binom{r}{j} \left(\frac{\lambda}{2r}\right)^{2j} j!}$, $\lambda = \frac{r}{r-d}$ ve $\sigma > 0$ 'dır.

Burada, olabilirlik eşitliklerinin çözülmesi neredeyse imkansızdır (Akkaya and Tiku, 2005). $z_{(i)} = \frac{(w_{(i)} - \mu)}{\sigma}$ ve $w_{(i)} = y_i - \phi y_{i-1} - \delta(x_i - \phi x_{i-1})$ olmak üzere MEÇO tahmin edicileri (1.64) - (1.78) eşitlikleri aracılığıyla ifade edilir:

$$\hat{\mu} = \bar{v} - \delta \bar{u} \quad (1.64)$$

$$\hat{\delta} = G + H \hat{\sigma} \quad (1.65)$$

$$\hat{\phi} = K - \lambda D \hat{\sigma} \quad (1.66)$$

$$\hat{\sigma} = \frac{\{-\lambda B + \sqrt{(\lambda B)^2 + 4nC}\}}{2\sqrt{n(n-3)}} \quad (1.67)$$

$$v_{[i]} = y_{[i]} - \hat{\phi} y_{[i]-1} \quad (1.68)$$

$$u_{[i]} = x_{[i]} - \hat{\phi} x_{[i]-1} \quad (1.69)$$

$$\bar{v} = \frac{\sum_{i=1}^n \beta_i v_{[i]}}{m} \quad (1.70)$$

$$\bar{u} = \frac{\sum_{i=1}^n \beta_i u_{[i]}}{m} \quad (1.71)$$

$$m = \sum_{i=1}^n \beta_i \quad (1.72)$$

$$G = \frac{\sum_{i=1}^n \beta_i (u_{[i]} - \bar{u}) v_{[i]}}{\sum_{i=1}^n \beta_i (u_{[i]} - \bar{u})^2} \quad (1.73)$$

$$H = \frac{(\sum_{i=1}^n \alpha_i (u_{[i]} - \bar{u}))}{\sum_{i=1}^n \beta_i (u_{[i]} - \bar{u})^2} \quad (1.74)$$

$$K = \frac{1}{\Delta} [\sum_{i=1}^n \beta_i (y_{[i]} - \hat{\delta} x_{[i]}) (y_{[i]-1} - \hat{\delta} x_{[i]-1})] - Z \quad (1.75)$$

$$Z = (\sum_{i=1}^n \beta_i (y_{[i]} - \hat{\delta} x_{[i]})) (\sum_{i=1}^n (y_{[i]-1} - \hat{\delta} x_{[i]-1})) \quad (1.76)$$

$$B = \sum_{i=1}^n \alpha_i [v_{[i]} - \bar{v} - G(u_{[i]} - \bar{u})^2] \quad (1.77)$$

$$C = \sum_{i=1}^n \beta_i [v_{[i]} - \bar{v} - G(u_{[i]} - \bar{u})^2] \quad (1.78)$$

Akkaya and Tiku (2004) uzun kuyruklu dağılıma sahip a_t 'ler için eşitlik (1.79)'daki olasılık yoğunluk fonksiyonunu ele almışlardır:

$$f(a) \cong \frac{1}{\sigma} \left\{ 1 + \frac{a^2}{k\sigma^2} \right\}^{-p} \quad (1.79)$$

Bu denklemde $k=2p-3$, $p \geq 2$ ve $\sigma > 0$ 'dır. Burada, olabilirlik eşitlikleri (1.80-1.83) çıkmaktadır (Akkaya and Tiku, 2004):

$$\frac{\partial \ln L}{\partial \mu} = \frac{2p}{k\sigma} \sum_i g \left(\frac{y_i - \mu - \phi y_{i-1} - \delta (x_i - \phi x_{i-1})}{\sigma} \right) \quad (1.80)$$

$$\frac{\partial \ln L}{\partial \delta} = \frac{2p}{k\sigma} \sum_i (x_i - \phi x_{i-1}) g \left(\frac{y_i - \mu - \phi y_{i-1} - \delta (x_i - \phi x_{i-1})}{\sigma} \right) \quad (1.81)$$

$$\frac{\partial \ln L}{\partial \phi} = \frac{2p}{k\sigma} \sum_i (y_{i-1} - \delta x_{i-1}) g \left(\frac{y_i - \mu - \phi y_{i-1} - \delta (x_i - \phi x_{i-1})}{\sigma} \right) \quad (1.82)$$

$$\frac{\partial \ln L}{\partial \sigma} = \frac{-n}{\sigma} + Z \quad (1.83)$$

Burada $Z = \frac{2p}{k\sigma} \sum_i \left(\frac{y_i - \mu - \phi y_{i-1} - \delta(x_i - \phi x_{i-1})}{\sigma} \right) g \left(\frac{y_i - \mu - \phi y_{i-1} - \delta(x_i - \phi x_{i-1})}{\sigma} \right)$ ve $g(x) = \frac{x}{\left(1 + \frac{x^2}{k}\right)}$ 'dir. Akkaya and Tiku (2004), bu eşitliklerin çözülmesinin neredeyse

imkansız olduklarını iddia etmektedir. Bu nedenle $z_{(i)} = \frac{(w_{(i)} - \mu)}{\sigma}$ ve $w_{(i)} = y_i - \phi y_{i-1} - \delta(x_i - \phi x_{i-1})$ olmak üzere MEÇÖ tahmin edicilerine ilişkin çözümler (1.84) - (1.87) eşitliklerinde verilmektedir:

$$\hat{\mu} = \bar{v} - \delta \bar{u} \quad (1.84)$$

$$\hat{\delta} = G + H \hat{\sigma} \quad (1.85)$$

$$\hat{\phi} = K + D \hat{\sigma} \quad (1.86)$$

$$\hat{\sigma} = \frac{\{B + \sqrt{B^2 + 4nC}\}}{2\sqrt{n(n-3)}} \quad (1.87)$$

Bu denklemlerdeki $v_{[i]}$, $u_{[i]}$, \bar{v} , \bar{u} , G, H, D ve K daha önceden uzun kuyruklu dağılımlar için elde edilenlerle aynıdır. B ve C ifadeleri, eşitlik (1.88) ve (1.89) ile verilmiştir:

$$B = \frac{2p}{k} \sum_{i=1}^n \alpha_i \left[v_{[i]} - \bar{v} - G(u_{[i]} - \bar{u})^2 \right] \quad (1.88)$$

$$C = \frac{2p}{k} \sum_{i=1}^n \beta_i \left[v_{[i]} - \bar{v} - G(u_{[i]} - \bar{u})^2 \right] \quad (1.89)$$

İçinde bulunan formüller açısından çok karmaşık görünen bu tekniğin en güzel yanı küçük verilerde bile iyi sonuçlar vermesidir (Akkaya and Tiku, 2004). MEÇÖ yöntemiyle zaman serileri analizi yapılırken dikkat edilmesi gereken husus, hataların dağılımının iyi belirlenmesidir. Örneğin, Hata dağılımını dikkate almadan $\alpha_{(i)}$ ve $\beta_{(i)}$ 'ler sıfır ve bir alındığında otomatik olarak en küçük kareler tahmini elde edilmektedir (Akkaya and Tiku, 2004). Böyle bir durumda, şayet hata dağılımı kısa kuyruklu bir dağılım ise örneklem büyüklüğü arttıkça eşitlik (1.41) ve (1.42)'deki model katsayılarının tahmin edicilerinin etkinliği düşmektedir (Akkaya and Tiku, 2004). Bunun haricinde şayet hata dağılımı uzun kuyruklu dağılım ailesinden ise, bu durumda

etkinlik uzun kuyruklu parametresi p 'ye bağlıdır ve etkinlik %50 ile %100 (p büyüdükçe) arasında değişim gösterir. Bu bilgiler ışığında hata dağılımının belirlenmesi çok önemli yer tutmaktadır. Hata dağılımının belirlenmesinden sonra $t_{(i)} = E(x_{(i)})$ değerlerinin bulunması gerekir. Kısa kuyruklu dağılımlar için $t_{(i)}$ eşitlik (1.90)'dan hesaplanmaktadır (Akkaya and Tiku, 2004; Tiku et al., 2001):

$$\int_{-\infty}^{t_{(i)}} \frac{C_1}{\sigma} \left\{ 1 + \frac{\lambda a^2}{2r\sigma^2} \right\}^r \frac{e^{-a^2/2\sigma^2}}{\sqrt{2\pi}} da = \frac{i}{n+1} \quad (1.90)$$

Bu integralin hesaplanması için Tiku et al.(2001)'e bakılabilir. Uzun kuyruklu dağılımlar için $t_{(i)} = E(x_{(i)})$ 'nin bulunması daha zordur. Uzun kuyruklu dağılımlar için olasılık yoğunluk fonksiyonu eşitlik (1.91) ile ifade edilmektedir:

$$f(x) = \frac{1}{\sigma \sqrt{2p-3} \beta(0.5, p-0.5)} \left(1 + \frac{(x-\mu)^2}{(2p-3)\sigma^2} \right)^{-p} \quad (1.91)$$

Buna göre $t_{(i)}$ eşitlik (1.92) aracılığıyla hesaplanır:

$$t_{(i)} = \frac{Cn!(k\pi)}{4(r-1)!(n-r)!} \int_{-1}^1 \sin(\pi u) \cos^q(\pi u/2) (F(u))^{i-1} (1-F(u))^{n-i} du \quad (1.92)$$

Burada $q=2p-4$, $C = \frac{\Gamma(p)}{\sqrt{(2p-3)\pi} \Gamma(p-0.5)}$ ve birikimli dağılım fonksiyonunun değeri $F(\cdot)$, Student t dağılımı aracılığıyla bulunmaktadır. Çünkü uzun kuyruklu dağılıma sahip herhangi bir x değişkeni için $\sqrt{\frac{v}{2p-3}} \left(\frac{x-\mu}{\sigma} \right)$ değişkeni v serbestlik dereceli t dağılımına sahiptir (Trawinski et al., 1985) ancak integralin çözümü oldukça zordur ve Simpson yönteminin alternatif bir kullanımı gerekmektedir (Rowland and Varol, 1972). Hesaplama zorluklarından doğan tekniklerden ötürü student t dağılımından sayı üretilip $t_{(i)}$ hesaplanabilir.

1.3.2.2 F_τ Tahmin Tekniđi

F_τ tahmin tekniđi, Maronna et al. (2006) tarafından bulunan gözlem deđerlerini filtreleyerek onlardan tahmin edici çıkararak yöntemdir. (Bu tahmin edicilerle ilgili daha fazla bilgi edinmek için Bianco, Garcia Ben, Martinez and Yohai(1996)'nin eserine bakılabilir.) Bu bakımdan F_τ tahmin tekniđi filtrelenmiş τ tahmin tekniđi olarak da ifade edilebilir. Standart τ tahmin yöntemine göre ARMA modelinde hatalar şayet normal dağılırsa bu durumda τ tahmin edicisi eşitlik (1.93) ile bulunabilir (Maronna et al., 2006):

$$\hat{\tau} = \hat{y}_{t|t-1}(\lambda) = E(y_t | y_1, \dots, y_{t-1}) \quad (1.93)$$

Bu durumda tahmin hataları eşitlik (1.94) ile ifade edilir:

$$\hat{e}_t(\lambda) = y_t - \hat{y}_{t|t-1}(\lambda) \quad (1.94)$$

$\hat{e}_t(\lambda)$ 'nin varyansı eşitlik (1.95) ile ifade edilir:

$$var(\hat{e}_t(\lambda)) = k_t^2(\lambda)\sigma_u^2 \quad (1.95)$$

Burada $k_t(\lambda)$ varyansı düzeltmek için uygulanan bir katsayıdır ve t büyüdükçe 1'e yaklaşmaktadır (Brockwell and Davis, 1991). Maronna et al.(2006) yukarıdaki formülü verdikten sonra, eşitlik (1.96)'daki olabilirlik formülünü tanımlamıştır:

$$Q(\lambda) = -2max_{\sigma_u} \log L(y_1, \dots, y_T, \lambda, \sigma_u) \quad (1.96)$$

Bu eşitlikteki L(.) fonksiyonu olabilirlik fonksiyonunu temsil etmektedir. Brockwell and Davis (1991)'ten faydalanılarak bu eşitlik (1.97) ile ifade edilmektedir:

$$Q(\lambda) = \sum_{i=1}^T \log(k_t^2(\lambda)) + T \log\left(\frac{\sum_{i=1}^T \hat{u}_t^2(\lambda)}{T}\right) \quad (1.97)$$

Bu eşitlikten sonra τ ölçek tahmin edicisi (1.98) ile ifade edilir:

$$\tau^2(u) = s^2(u) \sum_{t=1}^T \rho_2\left(\frac{u_t}{s(u)}\right) \quad (1.98)$$

Burada $u = (u_1, \dots, u_T)$ olmak üzere:

$$\rho_k(x) = \begin{cases} x^2 & |x| \leq k \\ 2k|x| - k^2 & |x| > k \end{cases} \quad (1.99)$$

olarak ifade edilir ve $s(u)$, $\rho_2(x)$ sınırlı fonksiyonuna dayalı M-ölçek tahminidir. M-tahmin edicilerine ilişkin daha geniş bilgi 1.3.2.3 kısmında verilecektir. Sağlamcı filtreleme yöntemleri Maronna et al. (2006)'ya göre toplamcı sapan değere sahip olan ARMA modellerinde yüksek bir kırılma noktası sağlamanın bir yoludur. Maronna et al. (2006), herhangi bir ARMA(p,q) modelini $k = \max(p, q + 1)$ olmak üzere eşitlik (1.100) ile ifade etmenin bir yolu olduğunu belirtmektedir:

$$\alpha_t = \phi \alpha_{t-1} + d u_t \quad (1.100)$$

$p > q$ olduğu durumda $i > q$ için $\theta_i = 0$ olmak üzere d değişkeni eşitlik (1.101) ile ifade edilmektedir:

$$d = (1, -\theta_1, \dots, -\theta_{k-1})' \quad (1.101)$$

Bu durumda durum geçiş matrisi eşitlik (1.102) ile ifade edilir:

$$\phi = \begin{bmatrix} \phi_{k-1} & I_{k-1} \\ \phi_k & 0_{k-1} \end{bmatrix} \quad (1.102)$$

Burada $\phi_{k-1} = (\varphi_1, \dots, \varphi_{k-1})$ ve $i > p$ için $\varphi_i = 0$ 'dır. Maronna et al. (2006)'ya göre şayet model $y_t = x_t + v_t$ şeklinde bir toplamcı sapan değere sahip ise bu durumda $E[\alpha_t | y_1, \dots, y_{t-1}]$ ve $E[\alpha_t | y_1, \dots, y_t]$ 'ye ilişkin tahminler sırayla eşitlik (1.103) ve (1.104) ile ifade edilir:

$$\hat{\alpha}_{t|t} = (\hat{\alpha}_{t,1|t}, \dots, \hat{\alpha}_{t,k|t}) \quad (1.103)$$

$$\hat{\alpha}_{t|t-1} = (\hat{\alpha}_{t,1|t-1}, \dots, \hat{\alpha}_{t,k|t-1}) \quad (1.104)$$

Buna göre;

$$\hat{x}_{t|t} = \hat{x}_{t|t}(\lambda) = \hat{\alpha}_{t,1|t}(\lambda) + \mu \quad (1.105)$$

olmaktadır ve bu tahmin ediciler $E[\alpha_t | y_1, \dots, y_{t-1}]$ ve $E[\alpha_t | y_1, \dots, y_t]$ için en basit tahmin ediciler olarak ortaya çıkmaktadır. Bunun yanında, $\hat{\alpha}_{t|t}$ ile $\hat{\alpha}_{t|t-1}$ tekrar tekrar hesaplanması için eşitlik (1.106) - (1.108)'den faydalanılır (Maronna et al., 2006):

$$\hat{\alpha}_{t|t-1} = \phi \hat{\alpha}_{t-1|t-1} \quad (1.106)$$

$$\tilde{u}_t(\lambda) = y_t - \hat{x}_{t|t-1} = y_t - \hat{\alpha}_{t,1|t}(\lambda) - \mu \quad (1.107)$$

$$\hat{\alpha}_{t|t} = \hat{\alpha}_{t|t-1} + \frac{m_t \psi\left(\frac{\tilde{u}_t(\lambda)}{s_t}\right)}{s_t} \quad (1.108)$$

Burada $\psi(\cdot)$ fonksiyonu tekil ve $x \geq 0$ için $\psi(x) \geq 0$ olan amaç fonksiyonu ρ fonksiyonunu temsil etmektedir. Bunun yanında m_t eşitlik (1.109) ile ifade edilen kovaryans matrisi M_t 'nin birinci sütununu ifade etmektedir:

$$M_t = E[(\hat{x}_{t|t-1} - x_t)(\hat{x}_{t|t-1} - x_t)'] \quad (1.109)$$

s_t ise eşitlik (1.110) ile ifade edilen M_t 'nin birinci sütun ve birinci satırda ($M_{t,11}$) bulunan değerinin karekökü olmaktadır:

$$s_t = \sqrt{M_{t,11}} \quad (1.110)$$

Filtrelenmiş τ tekniğinde ise eşitlik (1.96)'daki filtrelenmemiş artıkları filtrelenmiş artıklarla değiştirilerek tahmin edici elde edilir. Böyle bir işlem gerçekleştiğinde $Q(\lambda)$ eşitlik (1.111) ile ifade edilir:

$$Q^*(\lambda) = \sum_{i=1}^T \log(a_i^2(\lambda)) + T \log\left(\tau \left(\frac{\hat{u}_1^2(\lambda)}{\hat{a}_1^2(\lambda)}, \dots, \frac{\hat{u}_T^2(\lambda)}{\hat{a}_T^2(\lambda)}\right)\right) \quad (1.111)$$

Bu durumda filtrelenmiş τ tahmin edicisi

$$\hat{\lambda} = \operatorname{argmin}_{\lambda} Q^*(\lambda) \quad (1.113)$$

formunda ifade edilir. (Pratik bir algoritma için Maronna et al.(2006)'ya bakınız.)

1.3.2.3 M-tahmin Edicileri

M-tahmin edicileri genel bir tahmin ediciler grubu olarak tanımlanmıştır. Herhangi bir g amaç fonksiyonunu minimize eden tahmin ediciye M- tahmin edicisi denir. Örnek olarak, bir regresyon modelindeki katsayılar için en küçük kareler tahmin edicisi hata kareler toplamı fonksiyonu açısından bir M-tahmin edicisidir. Keza EÇO fonksiyonunu minimize ederken de M-tahmin edicisi hesaplanmaktadır. O nedenle bu yöntemi uygulamakta önemli olan tek şey, kullanılacak amaç fonksiyonudur. Örneğin Maronna et al. (2006), ARMA(p,q) modelleri için amaç fonksiyonunu eşitlik (1.114) ile tanımlamıştır:

$$\sum_{t=p+1}^T D \left(\frac{\hat{u}_t(\lambda)}{\hat{\sigma}} \right) \quad (1.114)$$

Buradaki D fonksiyonu sınırlı herhangi bir fonksiyon olup λ ARMA(p,q) modelindeki tahmin edilecek olan bütün parametreleri ifade etmekte ve λ 'nın tahmin sonucu elde edilen artıkları $\hat{u}_t(\lambda)$ ile temsil edilmiştir. Bunun haricinde $\hat{\sigma}$ Maronna et al. (2006) tarafından tahmin edilen bir ölçek parametresidir. Maronna et al.(2006), elde edilen tahmin edicinin koşullu bir EÇÖ tahmin edicisi olduğunu ve asimptotik olarak etkin olduğunu belirtmiştir. Bunun yanında AR(3) modeli için sapan değerlerin sayısının bulunan M-tahmin edicisi üzerindeki etkisi ile sıradan en küçük kareler tahmin edicisi üzerindeki etkisi arasında kıyaslama yapmıştır. Bu kıyaslama sonucunda en küçük kareler tahmin edicisinin sapan değerlerden daha çok etkilendiğini gözlemlemiştir. Maronna et al. (2006), sonlu olmayan varyansa sahip Cauchy dağılımlı a_t 'lere sahip modelleri simüle edip en küçük kareler tahmin edicisiyle kendi M-tahmin edicisinin davranışlarını incelemiştir. Bunun sonucunda elde ettiği sonuçlarla kendi tahmin edicisinin çok daha etkin olduğunu gözlemlemiştir. Ancak buna rağmen zaman serileri analizinde analizi etkileyecek pek çok etken bulunmaktadır ve bunlardan bir tanesi kaldıraç noktalarıdır (leverage points). Kaldıraç noktaları, regresyon modellerinde parametreleri etkileyen verideki noktalara denir (Everitt, 2006). Maronna et al. (2006), kaldıraç noktalarının etkisini kısıtlamak için eşitlik (1.115)'in çözülebilineceğini belirtmiştir:

$$\sum_{t=p+1}^T z_{t-1} \psi \left(\frac{\hat{u}_t(\lambda)}{\hat{\sigma}_u} \right) = 0 \quad (1.115)$$

Burada $\hat{u}_t(\lambda)$, ile $\hat{\sigma}_u$ daha önceden tanımlandığı gibidir. ψ sınırlı bir fonksiyon ve $z_t = (1, y_t, y_{t-1}, \dots, y_{t-p+1})'$ 'dür. Bu denklemde z_t ile $\hat{u}_t(\lambda)$ kısıtlanırsa daha sağlam tahmin ediciler elde edilir. Bu tahmin ediciler sınıfına genelleştirilmiş M-tahmin ediciler sınıfı denir (M-tahmin edicilerinin literatürdeki kullanımları için Denby and Martin (1979)'in, Martin(1980, 1981)'in makalelerine bakılabilir). Bustos(1982), AR(p) modelleri için genelleştirilmiş M-tahmin edicilerinin asimptotik olarak normal dağılıma sahip olduklarını ispatlamıştır.

2.BÖLÜM

YAPAY SİNİR AĞLARI

2.1 Yapay Sinir Ağlarının Tarihçesi

Sinir ağlarına ilişkin ilk çalışma Mcculloch ve Pitts (1943)'e aittir ve makalelerinde, nörofizyoloji ve matematiksel mantık çalışmalarını birleştiren sinir ağlarının mantıksal analizi tanımlanmıştır. Onların resmi nöron modeli ya hep ya hiç prensibine göre çalışıyordu. Bu makaleden esinlenilerek Von Neumann ENIAC (Elektronik nümerik birleştirici ve bilgisayar) ve EDVAC (Elektronik kesikli zaman otomatik bilgisayar)'ın yapımında kullanılan açma-gecikme elemanlarını tasarladı (Haykin, 1998). Bundan sonra 1948 yılında, denetim, iletişim ve istatistiksel sinyal işleme gibi konuları işleyen Wiener (1948)'in meşhur kitabı “Sibernetik: Makine ya da Hayvanda İletişim ve Kontrol” basıldı. Bu kitabın öğrenme ve kendi kendini organize etme konularını da içeren ikinci baskısı daha sonra 1961 yılında basıldı. Wiener istatistiksel mekaniğin yapay sinir ağları için önemini vurgulamış ancak istatistiksel mekaniğin öğrenme sistemlerini geliştirmesi Hopfield tarafından sağlanmıştır (Haykin, 1998). Sinir ağlarında Wiener'in atılımından sonraki gelişme, içinde sinir ağlarında sinaptik ağırlıklarda değişiminin izah edileceği Hebb (1949)'da basılan “Davranışın Organizasyonu: Nöropisokolojik Teori” adlı kitabıyla gelmiştir. Bu kitapta Hebb (1949), beyinde değişik görevler üstlenen yapıların nasıl oluştuğunu izah etmekteydi. Hebb'in kitabındaki meşhur öğrenme teoremi iki nöron arasındaki bağlantının iki nöronun da aynı anda ateşlenmesiyle güçlendiğini belirtmektedir. Esas mesleği psikologluk olan Hebb'in kitabı psikologlar arasında oldukça popüler olmasına karşılık yapay sinir ağları alanında yok denecek kadar az etkisi oldu (Haykin, 1998).

Minsky (1954), “sinirsel benzerlik takviye sistemleri teorisi ve beyin-model problemine uygulanması” adında doktora tezini yazdı. 1967'de Minsky'nin kitabı, “Hesaplama: Sonlu ve Sonsuz Makineler” basıldı ve kitap, Haykin(1998)'in belirttiğine göre Mcculloch ve Pitts'in 1943'te bulduğu sonuçları genişleterek hesaplama teorisine

yerleřtirdi. Bu grőe gre yapay sinir ađlarının hesaplamaya iliőkin uygulamaları, 1967 senesine dayanmakta olduđu sylenbilir.

Dijital bilgisayarın yapı temelini oluőturmuő olan Von Neumann yapısı, onun onuruna isimlendirilmiőtir (Haykin, 1998). 1955'te Yale niversitesi tarafından Siliman seminerlerini vermek zere davet edilmiőtir (Haykin, 1998). 1957'de vefat ettikten sonra verdiđi Siliman seminerleri, "Bilgisayar ve Beyin" adında bir kitap haline getirildi ve bir yıl sonra basıldı. Bugn John von Neumann, 20. yzyılın en nemli isimlerden biri olarak grlr.

Sinir ađlarındaki ele alınan bir baőka konu, nronlardan gvenilir bir ađ oluőturulmasıdır. Bu problem, Von Neumann tarafından Shannon ve Mccarthy (1956)'nin Automata Studies adlı kitabında yazılan bir blmnde gereksizlik kavramı baőka yerlere çekilmiőtir. Rosenblatt (1958)'in yayınlanmış makalesi ile algılayıcı kavramı ortaya çıkmıőtir. Rosenblatt'in makalesinde ortaya koyduđu teorem ile algılayıcının performansını teorik olarak stnlđn gstermiőtir. Bir algılayıcı ile onun bir benzeri model olan ADALINE(ADaptive LINear Element) arasındaki fark, eđitim sreçlerinden kaynaklanmaktadır (Haykin, 1998).

Bir dnem algılayıcıların her Őeyi yapabileceđine inanılıyordu (Haykin, 1998). Ancak daha sonra Minsky ve Papert tarafından bu konuda yazılan bir kitapta tek tabakalı algılayıcıların hesaplayabileceklerinin bir sınırı olduđunu gsterdiler. Çok tabakalı algılayıcılar iēin kitapta herhangi bir sıkıntı olmadıđı belirtilmiőtir (Haykin, 1998).

1980'lerde teoriye ve sinir ađlarının tasarımına bazı cephelerde ilgi yeniden canlandı (Haykin, 1998). Grossberg 1970'lerde rekabete dayalı đrenimi tanıtırken (Grossberg,1972, 1976a, b), ART adındaki yeni bir prensibini ortaya koydu.

1982'de Hopfield enerji fonksiyonu fikrini kullanarak, simetrik sinapslara sahip tekrarlı sinir ađlarında gerēekleően sinaptik deđiőimi aēıklamaya ēalıőmıőtir. Buna ek

olarak, istatistiksel fizik alanında kullanılan Ising modeli ile tekrarlı sinir ağıları arasında bir benzerlik oluşturdu (Haykin, 1998). Bununla, fiziğin sinir ağlarında kullanımına yol açtı. Sinir ağlarının bu türü, 1980'lerde büyük dikkat çekti ve zaman içerisinde bu ağlara Hopfield ağları denmeye başlandı. Bu ağların ortaya çıkışı aşağıdaki öncü çalışmalara dayanmaktadır (Haykin, 1998):

- Cragg ve Tamperley (1954, 1955), nöronlar nasıl ateşlenebiliyor ya da sakin kalabiliyorsa atomların kafes dönüşlerinde ya yukarı doğru ya da aşağıya doğru dönüşleri vardır.
- Grossberg (1967, 1968), nöronun doğrusal olmayan fark ve diferansiyel eşitlikler içeren toplamsal modelini tanıttı ve kısa dönem hafıza için modelin kullanımını araştırdı.
- Amari (1972), nöron için toplamsal modeli bağımsız olarak tanımlamıştır ve rassal olarak bağlantılı nöron benzeri elemanların dinamik davranışını çalışmıştır.
- Wilson and Cowan (1972), ateşleyici veya önleyici modeldeki nöronlar için yerleşmiş kitlelerin dinamiklerini inceleyen doğrusal olmayan diferansiyel eşitlikleri çıkarmıştır.
- Little and Shaw (1975), hem ateşleyen hem de ateşlemeyen nöronların olasılıksal modelini tanımlamış ve modeli kullanarak kısa dönem hafızayı açıklamaya çalışmıştır.
- Anderson, Silverstein, Ritz ve Jones (1977) makalelerinde doğrusal olmayan dinamiklerle ilişkili ilişkilendirici hafızadan oluşan bir kutu içerisinde beyin modelini sundu.

Hopfield ağlarının yanında Kohonen'in 1982 yılında tanıttığı kendini düzenleyen ağ da 1980'lerin önemli keşiflerinden sayılmaktadır (Haykin, 1998).

Kirkpatrick, Gelatt ve Vecchi (1983) kombinatorik optimizasyon problemlerini çözmek için yeni bir yöntem geliştirmiştir. Yöntemde, Ackley, Hinton ve

Sejnowski(1985) tarafından Boltzmann makinesi adı verilen stokastik bir makinenin tasarlanmasında kullanılmıştır (Haykin, 1998).

1983'te Barto, Sutton ve Anderson tarafından takviyeli öğrenme konusunda makale basıldı. Haykin (1998), kitabında böyle bir öğrenme türüyle çalışan yapay sinir ağı örneğini, bir el arabası üzerinde süpürgeyi dengede tutabilen bir makineyle vermiştir. Ağın tek gereksinim duyduğu şey, kritik bir açıdan düşmeye başladığında veya araba hareketi kestiğinde ona bunu bildirecek bir sinyaldi.

Rumelhart, Hinton ve Williams(1986) tarafından yazılan makaleye kadar pek kullanılmadı. Bunun yanısıra 1986'da, Rumelhart ve McClelland(1986)'ın Paralel Dağılmış İşlemci: Bilişin mikroyapılarında araştırmalar kitabı basılmıştır. Bu kitapla geri yayılım algoritması daha sonra çok tabakalı algılayıcılar için en popüler öğrenme algoritmalarından birisi olmuştur (Haykin, 1998). Gerçekte, geri yayılım algoritması aynı zamanda iki farklı yerde daha ortaya çıkmıştır. Geri yayılımın arkasındaki temel fikir için, Bryson ve Ho (1969)'nun uygulamalı optimal kontrol adındaki kitabına bakılabilir. Haykin (1998)'in kitabında 1.2 kısmı da geri yayılım öğrenme algoritması hakkında bilgi alınabilir.

Linsker, algılayıcı ağda kendini düzenleme için yeni bir prensip tanımlamıştır (Linsker, 1988). Bu prensip sinaptik bağlantılar ve sinaps dinamik genişliğine göre, azami bilgiyi korumak üzere tasarlanmıştır (Haykin, 1998). Linsker, (Shannon(1948) tarafından atılan)bunu yaparken Shannon istatistiksel bilgi teorisini kullanmıştır (Haykin, 1998).

1990'ların başında Vapnik ve arkadaşları, Destek Vektör Makinelerini keşfetti (Boser, Guyon ve Vapnik, 1992; Cortes and Vapnik, 1995; Vapnik, 1998, 2000). Bu yeni yöntem sonlu örneklem büyüklükleriyle öğrenme teorisindeki sonuçlar üzerine kuruludur. Destek Vektör Makinelerindeki en önemli özellik Vapnik-Chervonenkis(VC) boyutunun tasarımda doğal olarak yer almasıdır (Haykin, 1998).

VC boyutu, sinir ağının örnekler kümesinden öğrenilebilecek kapasitesidir (Vapnik ve Chervonankis, 1971; Vapnik, 1982).

1982’de çıkan Hopfield’in makalesi 1980’lerde sinir ağlarına ilişkin ilgiyi canlandırmıştır. Haykin(1998)’e göre bu dal, nörobilimlerde, psikolojide, matematikte, fiziksel bilimlerde ve mühendislikte derin kökleri olan disiplinler arası konu olmuştur.

2.2 Yapay Sinir Ağlarında Öğrenme Şekilleri

Yapay sinir ağlarında öğrenme, sinir ağının kendi hatasını yapısındaki bozukluğu düzeltmek için kullanması anlamına gelir. Hata düzeltmeli öğrenme, Hebb’çi öğrenme, rekabete dayalı öğrenme, Boltzmann öğrenmesi ve geri yayılım algoritmasıyla öğrenme olmak üzere beşe ayrılır.

2.2.1 Hata Düzeltilmeli Öğrenme

Hata düzeltilmeli öğrenmede, ağda sürekli olarak hata hesaplanmakta ve bunlara göre sinaptik ağırlıklar güncellenmektedir. Buna göre adıma göre hata eşitlik (2.1) ile ifade edilir (Haykin, 1998):

$$e_k(n) = d_k(n) - y_k(n) \quad (2.1)$$

Burada $d_k(n)$ ve $y_k(n)$ sırasıyla istenen çıktı ile ağın çıktısını temsil etmektedir. Bu noktada sinaptik ağırlıklar (2.2)’ye göre güncellenir:

$$w_{kj}(n+1) = w_{kj}(n) + \eta e_k(n)x_j(n) \quad (2.2)$$

Burada $x_j(n)$, sinaptik ağırlığın girdi tabakasında karşılık geldiği yerdir. η ise 0 ile 1 arasında değer alan bir parametre olup öğrenme parametresi olarak geçmektedir. Öğrenme parametresinin büyük olması, katsayılarda büyük ölçekli dalgalanmalara

sebeup olmakla birlikte öğrenme hızının nisbi olarak yüksek olmasını sağlamaktadır. Şayet öğrenme parametresi küçük alınırsa dalgalanma çok fazla olmamakla birlikte öğrenme hızı düşük kalacaktır:

2.2.2 Hebb'çi Öğrenme

Hebb'çi öğrenme, öğrenme şekillerinin en eskisidir (Haykin, 1998). Hebb(2002), Organization of Behaviour adlı kitabında eğer birbirine yakın iki nöron arasında sürekli bir şekilde senkronize bir ateşlenme oluyorsa bu durumda bu iki nöron arasındaki sinaptik bağlantı kendisini güçlendirecek şekilde metabolik değişimlere gider. Bu kurala ek olarak yapay sinir ağları kuramcıları, birbirine yakın iki nöron arasında senkronize olmayan bir ateşlenme gerçekleştiğinde sinaptik bağlantının zayıflaması kuralını koymuş ve sinir ağlarında Hebbçi öğrenme adında yeni bir öğrenme teorisi ortaya çıkmıştır. Buna göre sinaptik bağlantılardaki öğrenme faaliyeti en basit formu eşitlik (2.3) ile ifade edilir (Haykin, 1998):

$$w_{kj}(n+1) = w_{kj}(n) + \eta y_k(n)x_j(n) \quad (2.3)$$

Haykin (1998), burada formülün zaman içinde sinaptik ağırlıkları aşırı derecede büyüteceği için Sejnowski (1977 a, b) tarafından ifade edilen eşitlik (2.4)'ün kullanılabileceğini belirtmiştir:

$$w_{kj}(n+1) = w_{kj}(n) + \eta(y_k(n) - \bar{y})(x_j(n) - \bar{x}) \quad (2.4)$$

Burada \bar{y} ve \bar{x} , x ve y'nin zamana göre ortalamalarını temsil etmektedir. Bu formül sayesinde sinaptik ağırlıkların çok büyük olması engellenmektedir. Hebbçi bir sinapsın belirgin dört tane karakteristiği vardır (Brown et al.,1990):

- Zamana bağlı bir yapısı vardır. Sinaps öncesi sinyal ile sinaps sonrası sinyal aynı anda olup olmaması sinapsın yapısını etkilemektedir.

- Girdi sinyaline bağılı olarak kendini düzenleyen bir mekanizması vardır.
- Sinaps öncesi sinyal ile sinaps sonrası sinyal arasındaki zaman ilişkisi nedeniyle nöronların etkileşimi sinapsta önemli bir rol oynar. Bu yönüyle tamamen etkileşime bağılı bir yapısı olduğu söylenebilir.
- Aynı anda iki nöronda ateşleme olması sinaptik bağlantının kendini geliştirmesi için yeterlidir.

Bu özellikler Hebbçi öğrenme anlayışını önemli kılmaktadır.

2.2.3 Rekabete Dayalı Öğrenme

Rekabete dayalı öğrenmede birden fazla çıktı nöronu bulunmakta ve bu nöronların herbiri ateşlenmek için birbirleriyle yarışmaktadır. Rumelhart and Zipser (1985), makalelerinde bu tarz öğrenme için aşağıdaki üç temel elemanı ortaya koymuşlardır:

- Rassal dağıtılmış sinaptik ağırlıklar dışında belli miktarda girdi nöronları
- Her nöronun gücüne konmuş bir limit
- Tek bir çıktı nöronu aktive edecek bir mekanizma

Etkinleştirilen çıktı nöronuna, özel olarak winner-takes-all (kazanan hepsini alır) nöronu denir. Eğer bir çıktı nöronu etkinleştirilirse 1 değerini, aksi takdirde 0 değerini alır. Rekabete dayalı öğrenmede, sinaptik ağırlıklardaki değişim olan Δw_{kj} eşitlik (2.5) ile ifade edilir:

$$\Delta w_{kj} = \begin{cases} \eta(x_{kj} - w_{kj}) & k.\text{nöron etkinleşirse} \\ 0 & k.\text{nöron etkinleşmezse} \end{cases} \quad (2.5)$$

Bu formülden anlaşılacağı gibi, ağırlıklar zaman içerisinde girdi değerlerine kaymaktadır. Bu tarz sinaptik değişimlerin etkili olduğu alanlardan biri kümeleme

analizidir ama Haykin(1998), rekabetçi öğrenmenin başarıya ulaşması için her girdi yapısının farklı farklı kümelere girmesini şart koşturmuştur. Yazara göre bunun aksinin olması durumunda, birbirleriyle ilgisi olmayan girdi yapılarının aynı kümelere girmesi görülebilir.

2.2.4 Boltzmann Öğrenme

Boltzmann öğrenme, Ludwig Boltzmann onuruna isimlendirilmiş bir öğrenme türüdür. Boltzmann öğrenmenin gerçekleştiği sinir ağına, Boltzmann makinesi denmektedir. Boltzmann makinesi, eşitlik (2.6) ile ifade edilir (Haykin, 1998):

$$E = -\frac{1}{2} \sum_j \sum_{k \neq j} w_{kj} x_k x_j \quad (2.6)$$

Burada x_k yalnızca +1 ile -1 değerlerini almaktadır ve sinyalini değiştirme olasılığı eşitlik (2.7) ile sağlanmaktadır (Haykin, 1998):

$$P(x_k \rightarrow -x_k) = \frac{1}{1 + \exp\left(\frac{\Delta E_k}{T}\right)} \quad (2.7)$$

Burada ΔE_k ve T, enerji değişimini ve sıcaklığı simgelemektedir. Boltzmann makinesinin nöronları görünür ve gizil olmak üzere ikiye ayrılır (Haykin, 1998). Görünür nöronlar ağ ile çalıştığı çevre arasında bir arayüz sağlarken, gizil nöronlar özgürce çalışırlar (Haykin, 1998). Boltzmann makinesinde işlemler yürütülürken iki durumunun düşünülmesi gerekir:

- Çevre tarafından belirlenmiş belli hallerde kenetlenmiş durum
- Görünür ya da gizil bütün nöronların özgürce çalıştığı durum

ρ_{kj}^+ , kenetlenmiş durumlardaki j. nöron ve k. nöron arasındaki korelasyonu ifade etsin. ρ_{kj}^- , özgür durumlarındaki j. nöron ve k. nöron arasındaki korelasyonu ifade etsin. Bu iki korelasyon sinir ağında termal denge hakimken bütün olası durumlar için

ortalamları alınır ve böylece Boltzmann öğrenme kuralına göre, w_{kj} 'ye uygulanan değişim Δw_{kj} eşitlik (2.8) ile tanımlanır (Hinton ve Sejnowski, 1986);

$$\Delta w_{kj} = \eta(\rho_{kj}^+ - \rho_{kj}^-), j \neq k \quad (2.8)$$

Boltzmann makinesi hakkında daha fazla bilgi edinmek için Haykin (1998)'in *Neural Networks: A Comprehensive Foundation Second Edition* kitabının 11. bölümü incelenebilir.

2.2.5 Geri Yayılım Algoritması

Geri yayılım algoritması, sinir ağı üzerinde genel bir öğrenmeyi gerçekleştiren bir algoritmadır. Bu tarz öğrenmede sinaptik ağırlıkları güncellemek için (2.9) küçültülmektedir (Haykin, 1998):

$$E_{av} = \frac{1}{2N} \sum_{n=1}^N \sum_{j \in C} e_j^2(n) \quad (2.9)$$

Burada C kümesi çıktı nöronlarının indislerinin kümesidir ve bu fonksiyonun küçültülmesi için eşitlik (2.10) küçültülmeye çalışılır:

$$E(n) = \frac{1}{2} \sum_{j \in C} e_j^2(n) \quad (2.10)$$

Bu fonksiyon küçültülürken matematikteki zincir kuralından yola çıkılır ve eşitlik (2.11) elde edilir:

$$\frac{\partial E(n)}{\partial w_{ji}(n)} = \frac{\partial E(n)}{\partial e_j(n)} \frac{\partial e_j(n)}{\partial y_j(n)} \frac{\partial y_j(n)}{\partial v_j(n)} \frac{\partial v_j(n)}{\partial w_{ji}(n)} \quad (2.11)$$

Birtakım işlemler neticesinde eşitlik (2.11) (2.12)'ye dönüşür (Haykin, 1998):

$$\frac{\partial E(n)}{\partial w_{ji}(n)} = -e_j(n) \phi_j'(v_j(n)) y_i(n) \quad (2.12)$$

Böylece:

$$\Delta W_{ji}(n) = -\eta \frac{\partial E(n)}{\partial w_{ji}(n)} \quad (2.13)$$

elde edilir. Burada η öğrenme hızını ifade etmektedir. Genel olarak geri yayılım algoritmasında ağırlık değişimlerinin formülü $\delta_j(n)$ yerel gradyan fonksiyonunu ifade etmek üzere eşitlik (2.14) ile ifade edilir (Haykin, 1998):

$$\Delta W_{ji}(n) = \eta \delta_j(n) y_i(n) \quad (2.14)$$

2.3 Yapay Sinir Ağları Ve Zaman Serileri Analizi

Yapay sinir ağları, yarım asırlık bir geçmişi olan veriden öğrenme üzerine kurulu bir analiz yöntemidir. Veriyle ilgili hiçbir varsayımda bulunmaması onu bir parametrik olmayan yöntem kılmaktadır ve bunun yanında tasarımındaki yapılar sayesinde hesap hatasını düşürmektedir. Bu bahsedilen yapılar, sinir ağları içerisindeki gizil tabakalardır. Bu tabakalar, verilerin ifade edildiği girdi tabakası ile tahminlerin elde edildiği çıktı tabakası arasında yer alan tabakalardır. Bütün bunların yanında sinir ağlarının istatistiksel metodolojisinden daha etkili sonuçlar verdiği Werbos (1989, 1990) tarafından ispatlanmıştır. Bunun yanısıra Lapedes and Farmer (1988), doğrusal olmayan zaman serilerini yapay sinir ağlarıyla analiz etmeyi başarmıştır. Bunlar haricinde günümüzde zaman serileri analizi ile ilgili pek çok yapay sinir ağları çalışması olmuştur.

Coulibaly et al. (2000), makalesinde erken durdurulan eğitim yaklaşımını çok tabakalı ileri beslemeli sinir ağlarında eğitmek için kullanmıştır. Önerilen yöntemle,

ileri beslemeli ađın uyum g¼c¼n¼ en uygun hale getirmek iin hem Levenberg-Marquardt geri beslemesinden hem de apraz onaylamadan faydalanmıř ve ađın genelleřtirme becerisini geliřtirmiřtir. Modelin performansı, Kanada'da kuzey Quebec'teki su řebekesi verisi ¼zerinde istatistiksel model ile kıyaslanarak test edilmiřtir. Bu y¼ntem ile sadece ¼ng¼r¼ kesinliđi artmakla kalmamıř aynı zamanda dinamik uyarlamalı ¼ng¼r¼ye bir alternatif sađlanmıřtır.

Yamin et al. (2004), makalesinde yeniden yapılandırılmıř g¼ piyasalarında yapay sinir ađlarını kullanarak kısa vadeli elektrik fiyatı ¼ng¼r¼s¼nde bulunmak iin bir model ¼nermektedir. Modelde zaman, y¼k, rezerv ve tarihsel fiyat fakt¼rleri dikkate alınmıř ve MAPE iin yeni bir tanım getirilerek y¼ntemin performansı iki verisetinde incelenmiřtir.

Huang et al. (2005), destek vekt¼r makinelerini kullanarak NIKKEI225 endeksi adı verilen ¼zel bir endekste ¼ng¼r¼ gerekleřtirmiřtir. Makalede, destek vekt¼r makinesinin etkinliđini deđerlendirmek iin dođrusal ayırma analizi, kuadratik ayırma analizi ve Elman geri yayılım sinir ađlarıyla kıyaslama yapılmıřtır. Destek vekt¼r makineleri genel olarak sınıflandırmada kullanılmasına karřın sınıflandırma yapılarak bir ¼ng¼r¼ yapma ¼zerinde durulmuřtur ve makalede destek vekt¼r makinelerinin diđer sınıflandırma y¼ntemlerine g¼re daha iyi sonular verdiđi deneysel olarak kanıtlanmıřtır. Bunlara ek olarak b¼t¼n sınıflandırma modelleri biraraya getirilerek daha ¼st¼n bir model ortaya atılmıřtır. Bahsieen model t¼m ¼ng¼r¼ y¼ntemlerinden daha iyi bir iř ıkardıđı g¼r¼lm¼řt¼r.

Vahidinasab et al. (2008), kısa vadede elektrik fiyatı ¼ng¼r¼s¼ iin deđiřik y¼ntemlerin etkinliđini arařtırmıřtır. Makalede en uygun y¼ntem olarak yapay sinir ađları seilmiř ve duyarlılık analizi gerekleřtirilerek yapay sinir ađlarının en uygun girdileri seilmiř ve bulanık c-ortalamlar algoritması ile g¼nl¼k yapı k¼melemesi gerekleřtirilmiřtir. Son olarak modifiye edilmiř Levenberg-Marquardt ¼đrenme algoritması ile Pensilvanya-New Jersey-Maryland elektrik fiyatları tahmin edilmeye

çalışılmıştır. Daha önceden yapılan çalışmalarla kıyaslanıldığında sonuçlar gayet iyi çıkmıştır.

Aladağ et al. (2010), çalışmalarında ileri beslemeli sinir ağlarının yardımıyla bulanık ilişkiler tanımlayarak uyarlamalı beklenti modeliyle öngörülerde düzenleme getirmiştir. Bu makaledeki yenilik, öngörülerin düzenlenmesinin ağırlıklı parametrenin kısıtlı optimizasyonu ile yapılmasıdır. Önerilen yöntem Alabama üniversitesinden alınan yıllara göre kayıt verisi üzerinde uygulanmış ve de sonuçlara göre diğer yöntemlere göre bu öngörü yönteminin daha üstün olduğu ortaya çıkmıştır.

Behrang et al. (2010), çalışmasında güneşten kaynaklanan küresel radyasyonu tahmin etmeye çalışmıştır. Bunu gerçekleştirirken değişik yapay sinir ağları tasarımlarından faydalanılmış değişkenler olarak hava sıcaklığı nisbi nem, güneşin vurduğu saatler, buharlaşma ve rüzgâr hızı değerleri kullanılmıştır. Önerilen yöntem konvansiyonel yöntemlerle kıyaslanmış ve MAPE açısından %5,21 ile en iyi yöntem olarak öne çıkmıştır.

Hadavandi et al. (2010), genetik bulanık sistemler ile yapay sinir ağları metodolojisinin piyasa fiyatları öngörüsünde kullanımını araştırmıştır. Önerdiği yöntemde adımsal regresyon analizi kullanarak piyasa fiyatlarında etkili olan faktörleri ele almıştır. Sonra ham veriyi kendi kendini organize eden ağla k tane kümeye ayırmıştır. En sonunda genetik bulanık sistemler modeli kural bazlı çıkarım ile veri bazlı ayarlamayla kümeler çözümlenmiştir. IT ve havaalanı sektörlerinden elde edilen piyasa verilerindeki kıyaslamayla önerilen yöntemin daha önceki yöntemlerden daha etkili olduğu belirtilmiştir.

Yang et al. (2011), makalesinde iç mekanizma ile dış faktörleri öğrenerek öngörü yapan bir model geliştirmiş ve Baltık Panama Endeksinde (BPE) uygulamıştır. Önerilen yöntemde göre öncelikle veri dalgacık dönüşümüne tabi tutularak gürültüden arındırılıyor sonra destek vektör makinesiyle birleştirilmiş model ile öngörü yapılıyor. Modeldeki girdi değişkenleri önceki beş ay BPE değeri ve çıktısı ise beşinciden sonraki

ayın BPE değeridir. Model destek vektör makinesi eğitiminden de geçirilmiş ve performansının hem dalgacık dönüşümünden hem destek vektör makinesinkinden daha iyi olduğu görülmüştür.

Oğcu et al. (2012), Türkiye'deki elektrik tüketim verisini destek vektör regresyonu ile yapay sinir ağlarını kıyaslayarak analiz etmiştir. Kıyaslama kriteri olarak MAPE seçilen çalışmada destek vektör regresyonu yapay sinir ağına göre daha iyi sonuçlar vermiştir. Ancak makalede kullanılan sinir ağının yapısına ilişkin bir şey anlatılmamıştır.

Zhang et al. (2012), partikül yığını optimizasyonu ile çalışan en küçük kareler destek vektör makinesi, dalgacık dönüşümü ve ARIMA tekniklerini birleştirerek yeni bir teknik sunmuştur. Önerilen yöntemle Avustralya elektrik piyasasında Yeni Güney Wales'ten elde edilen veri analiz edilmiş ve bahsi geçen tekniklerden daha etkili sonuçlar elde edilmiştir.

Azadeh et al. (2013), makalesinde yapay sinir ağlarını, bulanık doğrusal regresyonu ve konvansiyonel doğrusal regresyonunu birleştirerek yeni bir yöntem oluşturmuştur. Önerilen yöntemin değişik verilere uygulanabilecek kadar esnek olduğu belirtilmektedir. Önerilen yöntem Amerika'daki ve İran'daki demir fiyatları verilerine uygulanmış ve olumlu sonuçlar elde edilmiştir.

Wu and Chau (2013), makalelerinde yağmur yağışına ilişkin öngöründe bulunmak için birtakım esnek hesaplama (soft computing) yaklaşımlarını incelemiştir. Bu yaklaşımları daha etkili kılmak için yazarlar veriyi ön işlemeden geçirmek ve yöntemin kendini veriye uyarlaması gibi seçenekleri gözden geçirmişlerdir. Bu bağlamda, hareketli ortalamalar ve tekil spektrum analizi ile ön işlem yapıp yapay sinir ağlarının destek vektör regresyon modeli birleşimi ve sadece yapay sinir ağları ile esas analiz gerçekleştirilmiştir. Bu şekilde dört adet veri seti incelenmiş ve yapay sinir ağları-destek vektör regresyon kombine modeli en iyi sonuçları vermiştir.

Ongena et al. (2013), makalesinde zaman serisinde özellik kestirimi, destek vektör makineleri, yankı-durum ağı (Echo-State Network), saf Bayes, özellik seçimi gibi sınıflandırma tekniklerini bir tane gerçek yaşam verisi üzerinde kıyaslamıştır. Makalede kullanılan gerçek yaşam verisinde yapılan analiz, bir hastanedeki kaydolan hastalardan beşinci veya onuncu günde diyaliz makinesine ihtiyaç duyanların sayısını öngörmeye yöneliktir. Elde edilen sonuçlara göre sınıflandırmada en başarılı sonuç %87,4 ile saf Bayes-özellik seçimi- özellik kestirimi birleşimiyle elde edilmiştir.

Kheirkah et al. (2013), elektrik tüketiminde sezonluk ve aylık değişimlerle ilgili yeni bir algoritma önermiştir. Önerilen algoritma yapay sinir ağlarını, temel bileşenler analizini, veri zarflama tekniğini ve ANOVA yöntemlerini kullanarak elektrik tüketimindeki değişimleri modellemektedir. Yöntemin etkinliğini göstermek için değişik yapılardaki yapay sinir ağlarını, İran'ın Nisan 1992'den Şubat 2004'e kadar aylık elektrik tüketim verisini analiz için kullanmışlardır. Önerilen yöntem çok büyük doğrulukla öngörüle bulunmayı başarmıştır.

Jhajharia et al. (2013), makalede sıtma vakaları üzerine yapay sinir ağları metodolojisini uygulamıştır. Bundaki temel amaç sıtma yayılımının hava koşullarından etkilenip etkilenmediğini test etmektir. Hindistan'ın kuzeyindeki çölde, değişik virüsler tarafından yayılan sıtma salgınındaki trendler Mann-Kendall testi ile test edilmiş sonrasında sıtma verisi ak gürültülü hale getirilmiştir. Bundan sonra iç bağıntılılık problemi incelenmiş ve temel bileşenler analiziyle iç bağıntı problemi tamamen ortadan kaldırılmıştır. Analiz sonunda öngörü için yapay sinir ağları kullanılmış ve virüslerden biri olan P. Falciparum'un çöl şartlarından etkilendiği görülmüştür.

Khatibi et al. (2013), çalışmasında Tebrizdeki hava kirliliği verisini kullanarak çoklu doğrusal regresyonu, yapay sinir ağlarını, genetik programlama, doğrusal olmayan yerel tahmin modeli ile ARIMA modelini kıyaslamıştır. İlk üç modelleme tekniğinde açıklayıcı değişkenler, sıcaklık, güneşin radyasyonu, çiğ oluşma derecesi ve rüzgâr hızı olarak alınmıştır. Ozon tabakasına ilişkin veri, 2010 Ağustosundan 2011 Martına kadarki saatlik verilerden oluşan bir zaman serisidir. Makalede çoklu doğrusal

regresyon, yapay sinir ağıları, genetik programlama modelleri etkili sonuçlar veremese de doğrusal olmayan yerel programlama ile ARIMA modeli etkili olmuştur.

Ramedani et al. (2013), Tahrandaki İran meteoroloji ofisinden elde ettiği küresel güneş radyasyonu (KGR) verisini yapay sinir ağılarıyla analiz etmiştir. Analizi gerçekleştirirken en yüksek sıcaklık, en düşük sıcaklık, ortalama sıcaklık, günışığı alım süresi ve yağmuru girdi değişkenler olarak, KGR değişkenini ise çıktı değişkeni olarak almıştır. Bu tarz çalışmalarda kullanılan bir başka model olan ampirik Hargreaves and Samani eşitliği (HS eşitliği) kıyaslama için ele alınmıştır. İncelemeler sonucunda yapay sinir ağının HS eşitliği modelinden daha etkili olduğu tespit edilmiştir.

Zhu and Wei (2013), makalelerinde en küçük kareler yöntemine dayalı çalışan destek vektör makinesi ile ARIMA modelini birleştirerek karbon fiyatları verisi üzerinde analiz gerçekleştirmiştir. Bunun yanında en küçük kareler destek vektör makinesinin en uygun parametrelerini bulmak için partikül yığını optimizasyonu denilen (particle swarm optimization) yöntemle tahmin doğruluğu geliştirilmiştir. Yeni yöntemin test edilmesi için avrupa birliği karbon fiyatları kuruluşundan iki karbon fiyat serisi kullanılmış ve olumlu sonuçlar alınmıştır.

Stepnicka et al. (2013), makalesinde çok adımlı sezonluk zaman serileri öngörüsü için temel bir yöntem geliştirmiştir. Önerilen yöntem, yapay sinir ağıları, destek vektör makineleri ve bulanık kurallar aracılığıyla ortaya çıkmıştır. Geliştirilen yöntemde sekiz veri üzerinde test edilmiş ve kombine yöntemin diğer yöntemlerden daha iyi sonuç verdiği belirtilmiştir.

Yeh (2013), makalesinde literatürde basitleştirilmiş yığın optimizasyonu adı verilen bir yöntemi geliştirerek yapay sinir ağıları metodolojisine uygulamıştır. Bu yöntem bir yapay sinir ağında ağırlık ayarlama yöntemi olup bu yöntem zaman serisi analizinde kullanılan en popüler beş yöntem (geri yayılım algoritması, partikül yığın algoritması, genetik algoritma, kooperatif öğrenmeli partikül yığın algoritması ve

basitleştirilmiş yığın optimizasyonu) ile kıyaslanmaktadır. Deneysel sonuçlar, önerilen yöntemin sağlam ve diğer algoritmalara göre daha etkili olduğunu göstermiştir.

3.BÖLÜM

ÇEŞİTLİ TURİZM VERİLERİNİN ÖNGÖRÜSÜNDE SAĞLAMCI ARIMA VE YAPAY SİNİR AĞLARININ KULLANIMI

Turizm, dünya ekonomisi üzerinde önemli bir etkisi olan bir gelir kaynağıdır. 1997 yılında dünya turizm ve seyahat sektörünün dünyadaki gayrisafi hasılanın %10'una karşılık geldiği ve bu sektörde 262 milyon insanın çalıştığı belgelenmiştir (United Nations World Tourism Organization, 1998). 1997 yılında doğu Asya-Pasifik bölgesindeki ekonomik krizden ötürü, turizmde gelişme düşmüştür. Aynı zamanda amerikan dolarına karşılık gelen majör para birimlerinin amortismanı uluslararası turizm sektörünü vurmuş ve bu da Avustralya turizmini etkilemiştir. Genel turist akımları 1998 yılının erken dönemlerinde bir düşüş gerçekleşmiştir. Japonya, Kuzeydoğu Asya ve Güneydoğu Asya'da 1998'in ilk yarısında genel turist akımları sırayla %6, %30 ve %21 oranında düşüş sergilenmiştir (Australian Bureau of Statistics, 1998). Buna rağmen Tourism Forecasting Council (1998) kuruluşu Avrupalı ve Kuzey Amerikalı turist akımlarının 1998 yılının ikinci yarısında büyüyeceğini tahmin etmiştir. Bütün bunların yanında öyle ülkeler vardır ki ekonomileri yerel turizmden önemli ölçüde etkilenmektedir. Örneğin yerel turizm Avustralya'da önemli bir sektördür. 1996 yılında Avustralya içerisinde yerel turizm akımları %1,5 oranında büyüme kaydederken, aynı dönemdeki yerel turizm içerisindeki geceleme oranı %2 oranında artmıştır.

İş için seyahat edenler, tatil amaçlı seyahat edenler, turla gelen turistler, konferans katılımcıları, hükümet yetkilileri ve diğer türden ziyaretçiler Avustralya'nın turizm sektörü için müşteri portföyünü temsil etmektedir (Ahmed and Cross, 1999). Öyle ki bu ihtiyaçların karşılanması amacıyla otel sayısı ve otel odalarının sayısı 1988-1998 yılları arasında arttırılmıştır (Australian Bureau of Statistics, 1998 b). Otel odalarının sayısında artırma, daha fazla turisti ağırlamak için gerçekleşmiştir. İş hacmindeki gelişmeler sayesinde Avustralya'da turizm büyük bir gelişim kaydederek otel odasının %60'ın üstünde bir oranda dolu olması sağlanmıştır (Law, 1998). Bunun yanında turizm ulaşım, barınma alanlarındaki yatırımlara etki ettiğinden turizm için öngörülerde bulunmak önemli olmuştur (Ahmed and Cross,1999).

Bu çalışmada, yurt dışından tatil için gelip sınırlardan çıkan yabancı sayısı, yurtdışından gelen yabancıların Türkiye genelinde geceleme oranları ve son olarak yabancı turistlerin kişi başına yaptığı harcamaları öngörülme çalışılmıştır. Veriler TÜİK'ten alınmıştır. Söz konusu çalışmada doğal logaritması alınmış veriler üzerinden analizler gerçekleştirilmiştir.

Bu çalışmada turizm verisinin kullanılmasındaki temel sebep, turizmin ülkemiz açısından önemli gelir kaynakları arasında yer alması ve ülkemizde 90'lı yıllardan itibaren önemli bir değişim geçirmesidir. Örneğin, Eskişehir'in geçtiğimiz yıllarda belediyeçilik faaliyetleriyle, kültürel altyapısı, sanatsal faaliyetleri avrupa kentleri seviyesine gelmesi iç turizm açısından önemli bir kazanç olmuştur (Eskişehir Valiliği, 2011). Deniz-Güneş turizmi dışında başka alternatif turizm çeşitlerinin çıkması bu alanda elde edilen kazançları etkilemiştir. Bugün bir şehir açısından pek çok faktör şehir turizmine etki etmektedir ve yavaş yavaş bu konuda bir inceleme yapılması gerekliliği ortaya çıkmaktadır. Henüz çıkarsamalar ve standart öngörüler için turizm ile ilgili yeterli miktarda veri olmamasına rağmen sağlamcı istatistiksel teknik olan MEÇO tekniği veya yapay sinir ağları için yeterli derecede veri bulunmaktadır.

Yapay sinir ağı yöntemi uygulanırken yurtdışından gelip çıkış yapan turist sayısı hariç bütün verilerde MINITAB 16 programıyla model elde edilmiş ve buna göre girdi tabakasındaki nöron sayısı belirlenmiştir. Örneğin, zaman serisinde AR(2) tipi bir model varsa, yapay sinir ağı tasarımında 3 girdi nöronu ve 1 çıktı nöronu (Bu durumda tasarım 3-0-1 şeklinde ifade edilecektir) bulunmaktadır. Bunun yanında tasarım içerisinde hata düzeltmeli öğrenme gerçekleştirilmiş ve öğrenme katsayısı düşük tutulmuştur. Öğrenme katsayısının düşük tutulması tamamen bir metodolojik yenilik getirme çabası olup, yapay sinir ağı modellemesini bir çeşit yöneylem araştırması problemi olarak ele almayı ifade etmektedir. Daha açık olmak gerekirse, elimizde yapay sinir ağıyla analiz edebileceğimiz bir veri varsa ona ilişkin birtakım kısıtlamalarla ideal modele varmamız daha kolay olacaktır. Bu çalışmada kısıt koymadan kastedilen olay, model katsayı tahminleri ile oluşturulan güven aralıklarıdır ve turizm verilerinde bu

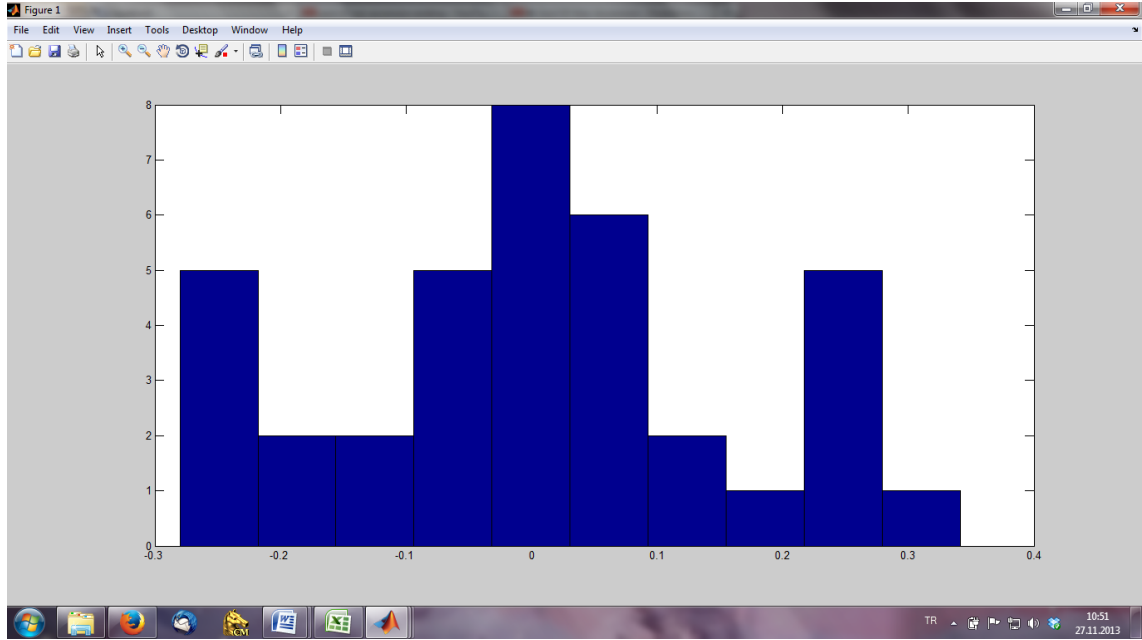
sayede daha iyi sonuçlar elde edilmiştir. Yöneylem araştırmasında bu kısıtları koyma tekniğine TABU araması adı verilmiştir. İlk olarak Glover and McMillan (1986) tarafından geliştirilen ve Glover (1989, 1990) tarafından formüle edilen bu yöntem herhangi bir çözüm arama yaklaşımının performansını geliştirerek en ideal sonuca varmasını sağlamaktır. Çalışmada sinaptik ağırlıklar rasgele dağıtılırken daha önceden geliştirilmiş olan modellerde katsayıların güven aralıkları kullanılarak en uygun çözüm bulmaya çalışılmıştır.

3.1 Yurtdışından Gelip Sınırlardan Çıkış Yapan Turist Sayısının Öngörüsü

Bu veri yılı dört döneme bölerek (2003-I döneminden 2013-III dönemine kadar) toplanmış bir veridir. Verinin büyüklüğü 43 birim olup uygulamada verinin önce doğal logaritması ve sonradan seri farkları alınarak analiz edilmiştir. Yapay sinir ağları tekniği uygulanırken MEÇO tekniğiyle katsayılar hesaplanmış ve ona göre aralıklar belirlenmiştir. Bu aralıklardan rasgele seçim yapılarak sinaptik katsayılar belirlenmiştir.

MEÇO tekniği uygulanırken en başta en küçük kareler ile hataları bulunmuş ve Şekil 3.1'deki histogram çizdirilmiştir:

Şekil 3.1: Hataların histogramı



Görüldüğü gibi hata dağılımında birden fazla mod göze çarpmaktadır. Bu durum hataların dağılımının $r > d > 0$ olan bir kısa kuyruklu dağılıma sahip olduğunu göstermektedir (Akkaya and Tiku, 2004). Bunun yanında hatalarının dağılımının çarpıklığı ve basıklığı 0,1339 ile 1,4534 bulunmuştur ve $d > 0$ için basıklık nispeten daha düşük çıkmaktadır. Bu nedenle Çizelge 3.1'deki Logolabilirlik fonksiyon değerlerine bakılması gerekir:

Çizelge 3.1 r ve d değerleri için Logolabilirlik değerleri

LogL	d=-0,5	d=0	d=0,5	d=1	d=1,5
r=2	77,1571	69,4743	59,3802	43,7135	18,1407
r=4	72,4904	68,3410	61,1838	53,5916	43,8441

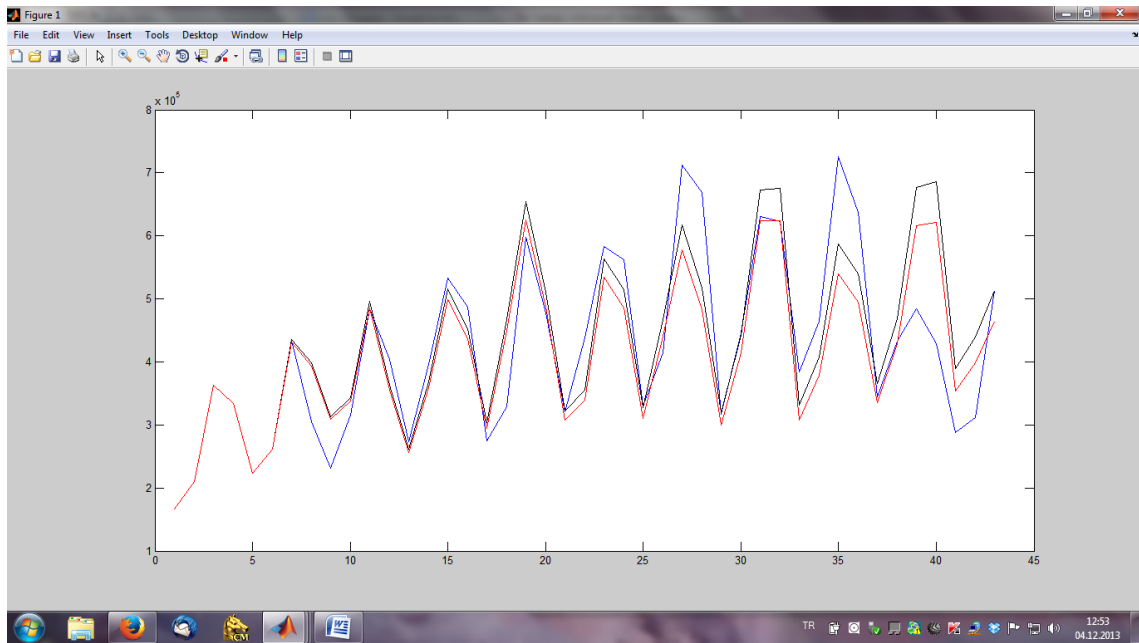
Çizelge 3.1'e göre hatalar için en uygun dağılım $r=2$ ve $d=-0,5$ için kısa kuyruklu dağılımdır. Buna rağmen uygun kabul edilen hata dağılımıyla elde edilen modelin MAPE değeri 93 civarı çıktığından dışlanmış ve onun yerine $r=4$ ve $d=-100$ parametrelili kısa kuyruklu dağılım seçilmiştir. Bu seçilen dağılım sonucunda elde edilen RMSE'ler her iki teknik için de 0,1578 (Yapay sinir ağlarında test verisi için RMSE 0,091'dir. Son üç hata sırayla 0,1645, -0,0428 ve 0,3476'dır ve MEÇÖ modeli için son üç hata sırayla 0,169, -0,0426 ve 0,3419'dur) bulunmuştur. MAPE değerleri ise yapay sinir ağları için 1,7407 ve MEÇÖ tekniği için 2,3876'dır. Buna karşılık gelen gerçek RMSE değerleri 81760 ile 83974 olup MAPE değerleri 12,7433 ile 13,2712 bulunmuştur. Modeller sırasıyla yapay sinir ağları ve MEÇÖ tekniği için aşağıdaki şekilde oluşturulmuştur:

$$Y_t = -0,0115 - 0,4731Y_{t-1} + 0,8999Y_{t-4} + 0,425Y_{t-5} + \varepsilon_t \quad (3.1)$$

$$Y_t + 0,4815Y_{t-1} = -0,0092 + 0,9099(Y_{t-4} + 0,4815Y_{t-5}) + \varepsilon_t \quad (3.2)$$

Bunun sonucunda gerçek verilerle beraber Şekil 3.2'deki grafik elde edilmiştir:

Şekil 3.2: MEÇÖ ve Yapay Sinir Ağları yöntemleriyle elde edilen öngörüler



Mavi renk gerçek veriyi, kırmızı renk yapay sinir ağları yöntemiyle elde edilen öngörülerini ve siyah renk MEÇO tekniğiyle elde edilen öngörülerini temsil etmektedir. Bütün bulunan sonuçlardan hareketle yapay sinir ağları tekniğinin MEÇO tekniğine göre daha iyi olduğu söylenebilir. Elde edilen modellerle gelecek 4 dönem için öngörüler Çizelge 3.2'deki gibidir:

Çizelge 3.2: İki yöntemle göre günümüzdeki dört döneme ilişkin öngörüler

Dönemler	MEÇO tekniği için	Yapay sinir ağları tekniği için
2013-IV	561848	505858
2014-I	320519	290158
2014-II	361853	326491
2014-III	411147	369928

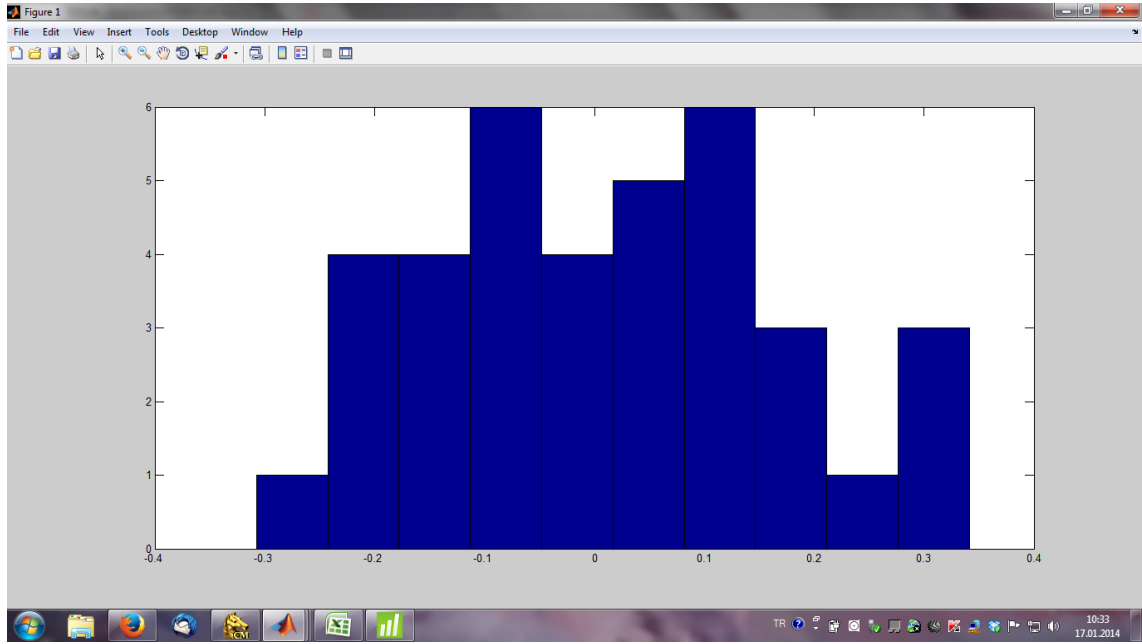
Anlaşılabacağı gibi kurulacak iyi bir yapay sinir ağı modeli sağlamcı öngörü tekniğine göre daha iyi öngörüler yapabilmektedir. İlerleyen bölümlerde Türkiye'de turistlerin geceleme oranı ve kişi başına harcama tutarları yapay sinir ağları modelleri ile öngörülmüştür.

3.2 Türkiye'de Turistlerin Geceleme Oranlarının Öngörüsü

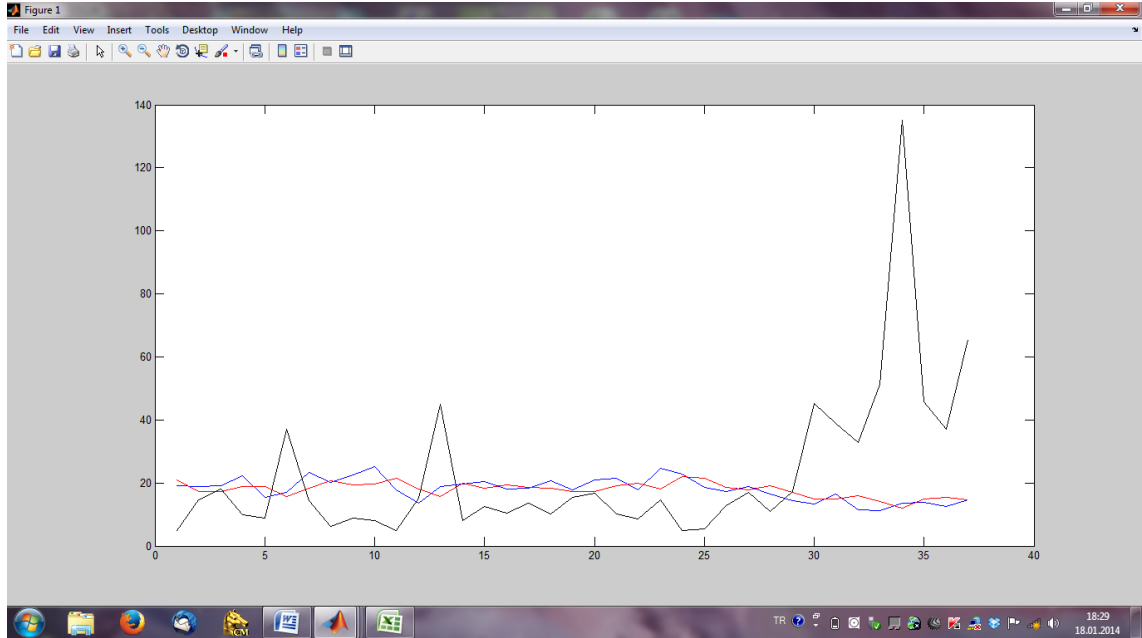
Burada kullanılan veri, 3 aylık dilimlere ayrılarak toplanmış geceleme oranları verisidir. Veri analiz edilmeden önce doğal logaritması alınmış ve sonradan yapay sinir ağlarıyla analizi gerçekleştirilmiştir. Analiz sonunda gerçek verilerle aşağıdaki grafik ortaya çıkartılmış ve doğal logaritması alınmamış haliyle RMSE 1,1716 bulunurken

MAPE=9,2202 olarak tespit edilmiştir (Son üç hata 0,9202; 0,4401; -0,410 olarak tespit edilmiştir). MEÇO yöntemine göre hatalar Şekil 3.3'teki histograma sahiptir. Buna göre A_t değerleri üretilmiş (Bu değerler için Çizelge 3.4'e bakınız) ve gerçek MAPE ve RMSE değerleri 108 ve 26,6456 elde edilmiştir (Ön işleme yapılmış verilerle MAPE 22,1934, RMSE 0,8524 elde edilmiştir ama ön işlemsiz ele alındığında durum daha kötü hale gelmiştir.). Şekil 3.4 ile yapay sinir ağları ve MEÇO yöntemlerine göre elde edilen öngörüler çizilmiştir.

Şekil 3.3 Hataların histogramı



Şekil 3.4 Yapay Sinir Ağları yöntemleriyle elde edilen öngörüler



Mavi renk gerçek veriyi, kırmızı renk yapay sinir ağları yöntemiyle elde edilen öngörülerini ve siyah renk MEÇO tekniğiyle elde edilen öngörülerini temsil etmektedir. Böylece elde edilen modelin iyi bir model olduğu söylenebilir. Elde edilen model aşağıdaki şekildedir:

$$Y_t = 1,2161 + 0,5766Y_{t-1} + 0,0109A_t + 0,0499A_{t-1} \quad (3.3)$$

$$Y_t = 12,9216 - 3,5286Y_{t-1} + 0,0704A_t - 0,2483A_{t-1} \quad (3.4)$$

Çizelge 3.3 elde edilen öngörülerini içermektedir:

Çizelge 3.3: Yapay sinir ağları yöntemine göre önümüzdeki dört döneme ilişkin öngörüler

Dönemler	Yapay sinir ağları tekniğine göre öngörüler	MEÇO tekniğine göre öngörüler
2013-IV	15,8565	47,7830
2014-I	15,4818	0,6518
2014-II	16,0861	2112960,8615
2014-III	17,7106	2,3251e-17

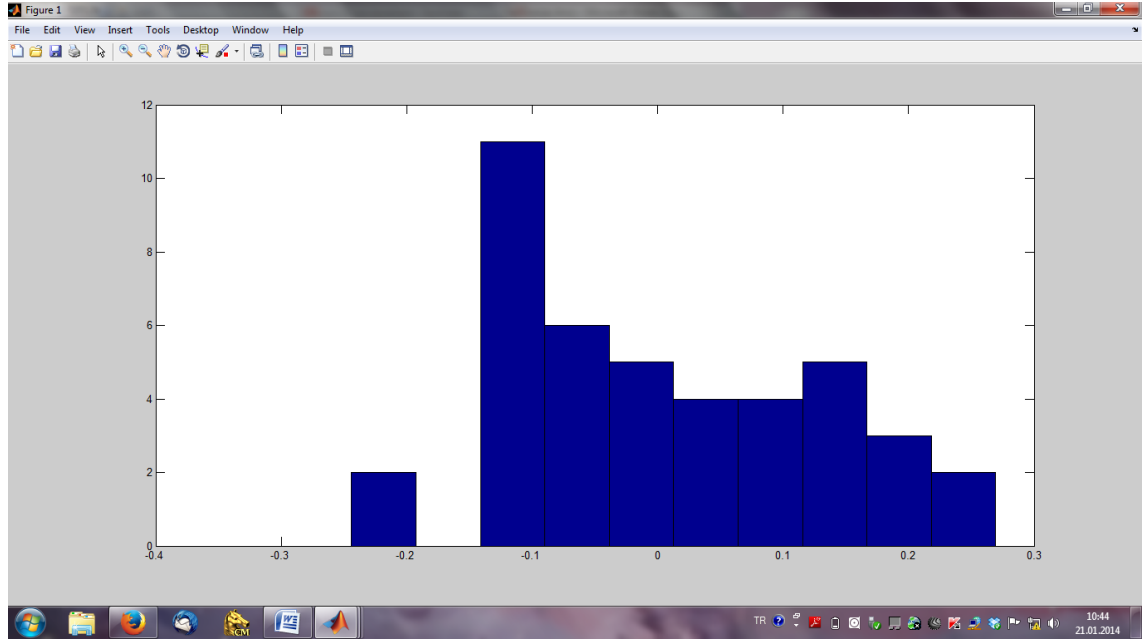
Çizelge 3.4: Üretilen A_t değerleri

t	A_t	T	A_t	t	A_t	t	A_t
1	1,0340	11	1,0307	21	1,0139	31	1,0216
2	0,9786	12	1,0160	22	1,0316	32	1,0287
3	1,0240	13	1,0194	23	1,0225	33	1,0138
4	1,0322	14	1,0339	24	1,0318	34	1,0326
5	1,0331	15	1,0096	25	1,0144	35	1,0316
6	1,0331	16	1,0149	26	1,0340	36	1,0312
7	1,0155	17	1,0324	27	1,0288	37	1,0066
8	1,0015	18	1,0309	28	1,0286	38	1,0314
9	1,0317	19	1,0300	29	1,0285	39	1,0216
10	1,0261	20	1,0057	30	1,0328	40	1,0287

3.3 Türkiye'deki Turistlerin Kişi Başına Harcama Tutarlarının Öngörüsü

Bu bölümde Türkiye'deki turistlerin kişi başına yaptığı harcama tutarları öngörülme çalışılmıştır. Veri 3 aylık dilimlerle toplanmış olup, dolar cinsinden ifade edilmiştir ve analizden önce verinin doğal logaritması alınarak ön işlemeden geçirilmiştir. Bu sayede yapılan analiz sonucunda gerçek veriler için MAPE=7,8009 ve RMSE=67,4734 çıkartılmıştır (Son üç hata 141,649, 36,849, -52,149 olarak tespit edilmiştir). Şekil 3.5'teki histograma bakıldığında hataların çarpık bir dağılıma sahip olduğu görülmektedir.

Şekil 3.5 Hataların histogramı



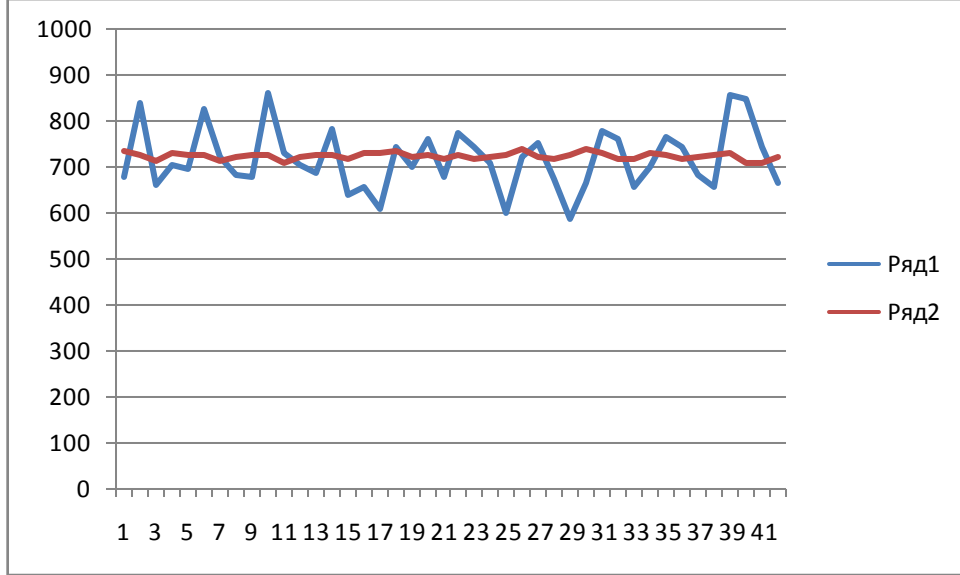
MEÇO yönteminde önce hataların dağılımı kısa kuyruklu simetrik dağılım varsayılarak, daha sonra ise geliştirilmiş lojistik dağılım varsayılarak işlem yapılmıştır. Yapay sinir ağları modeli için Çizelge 3.6'daki değerler üretilmiş ve buna göre model gerçekleştirilmiştir. İlkinde Logolabilirlik fonksiyon değeri 217 civarı ikincisinde ise bu değer 7,445 olarak elde edilmiştir. Histogramdaki dağılımdan ötürü çarpık dağılıma göre model ifade edilmiş olup bu model için MAPE RMSE 10 haneli rakamlarda elde edilmiştir (diğer model için elde edilen MAPE değeri ve RMSE daha kötü). Bu sonuçları daha açıkça yorumlamak gerekirse, ortalama hata payı yaklaşık %7'dir ve 70 dolara karşılık gelmektedir. Elde edilen modeller eşitlik (3.5) ve (3.6)'da verilmiştir:

$$Y_t = 0,5953 - 0,1072Y_{t-1} + \varepsilon_t \quad (3.5)$$

$$Y_t = 687,9069 - 103,9681Y_{t-1} - 0,016(A_t - 103,9681A_{t-1}) + \varepsilon_t \quad (3.6)$$

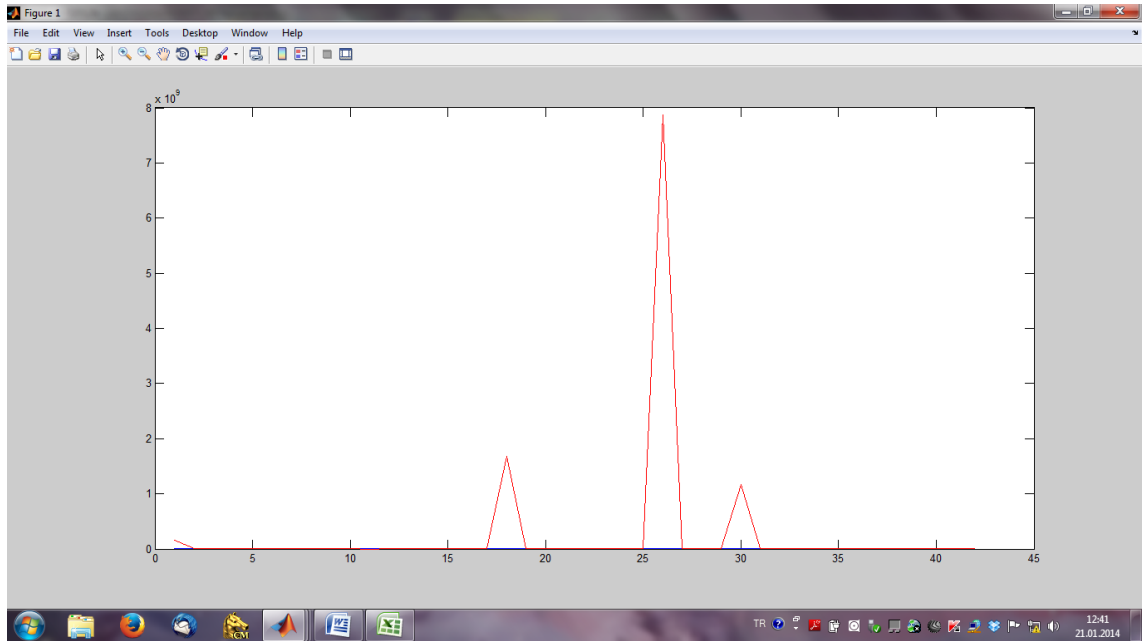
Bulunan katsayılarla elde edilen grafik Şekil 3.6'daki gibi çıkmıştır:

Şekil 3.6 Yapay Sinir Ağları yöntemleriyle elde edilen tahminler



Mavi renk gerçek veriyi, kırmızı renk yapay sinir ağları yöntemiyle elde edilen öngörülerini temsil etmektedir. Çizelge 3.5 ile yapay sinir ağlarıyla elde edilen öngörüler gösterilmiştir. Böylece elde edilen modelin iyi olduğu söylenebilir: Şekil 3.7’de MEÇO ile elde edilen öngörüler gösterilmiştir.

Şekil 3.7 MEÇO yöntemiyle elde edilen öngörüler



Çizelge 3.5: Yapay sinir ağırları yöntemine göre önümüzdeki dört döneme ilişkin tahminler (MEÇO için öngörüler modelden ötürü kötü çıkmıştır)

Dönemler	Yapay sinir ağırları tekniğine göre öngörüler
2013-IV	721,9817
2014-I	721,6773
2014-II	721,7099
2014-III	721,7064

Çizelge 3.6: Üretilen A_t değerleri

T	A_t	t	A_t	t	A_t	t	A_t	t	A_t
1	0,8351	11	-0,0825	21	0,4900	31	1,4367	41	1,0984
2	-0,2437	12	-1,9330	22	0,7393	32	-1,9609	42	-0,2779
3	0,2157	13	-0,4390	23	1,7119	33	-0,1977	43	0,7015
4	-1,1658	14	-1,7947	24	-0,19419	34	-1,2078		
5	-1,1480	15	0,8404	25	-2,1384	35	2,9080		
6	0,1049	16	-0,8880	26	-0,8396	36	0,8252		
7	0,7223	17	0,1001	27	1,3546	37	1,37897		
8	2,5855	18	-0,5445	28	-1,0722	38	-1,0581		
9	-0,6669	19	0,3035	29	0,9610	39	-0,4686		
10	0,1873	20	-0,6003	30	0,1240	40	-0,2725		

BÖLÜM 5

SONUÇ VE ÖNERİLER

Yapay sinir ağları metodu, gözlemlerden doğruluk payı yüksek çıktı elde eden modern bir analiz yöntemidir. İlk başlarda teorik altyapısı kurulduğunda birgün insan beyninin bütün işlevlerini yerine getirmesi amaçlanırken, bugün geldiği noktada değişik bir yöntem olarak ortaya çıkmıştır. Bu durum zaman serileri analizi alanında yapay sinir ağları yöntemini, geliştirilen diğer istatistiksel yöntemlerden daha avantajlı bir konuma getirmiştir. Bununla beraber istatistiğin yapay sinir ağlarını geliştirmede hiçbir faydasının olmayacağını söylemek doğru değildir.

Herhangi bir yapay sinir ağı tasarımının başarıya ulaşabilmesi için istatistiksel bilginin kullanımı kaçınılmaz olduğu çalışmada gösterilmiştir. Yapay sinir ağları literatürüne bakıldığında Vapnik (1971) tarafından bu durum daha önce ifade edilmiştir. Ancak bugün zaman serileri analizinde yapay sinir ağları kullanılırken, istatistik biliminde önceden geliştirilmiş olan model belirleme tekniklerinin faydalı olabileceği çalışmada gösterilmiştir. Özellikle model katsayı tahminleri ile oluşturulan güven aralıkları, turizm verilerinde önemli işler yapmıştır. Bu nedenle yapay sinir ağları mimarisi oluşturulurken model kısıtlamalarını referans olarak almak, çalışmayı başarıya ulaştırmaktadır.

Turizm öngörülerinde bulunulurken, doğruluk payının yükseltilmesine önem gösterilmiştir. Bunun yanında doğru modeli oluşturmak, öngörü doğruluğu kadar önemli bir faktördür. Bu çalışmada hem doğru modeli oluşturmak hem öngörü doğruluğunu sağlamak açısından önemli oranda başarı elde edilmiştir. Bu iki konuda başarı sağlamak için istatistik ve yapay sinir ağları alanında bilginin yanısıra, ideal modele ulaşmak için bir çözüm arama yöntemi gerekmektedir. Şayet yapay sinir ağları mimarisinde bir problem başgösterirse, mimari için uygun çözümü aramayı halledecek bir yöneylem araştırması önemli bir çalışma olur.

Elde edilen sonuçlara göre yapay sinir ağıları modeli teorik istatistiksel modeli iyileştirecek şekilde olumlu bir etki sunmaktadır. Bu noktada teorik modelin oluşturulmasında kullanılan MEÇO yöntemine ilişkin getirilecek en iyi öneri, veride standartlaştırma yapılmadan uygulanmasıdır. Geceleme oranları verisinde bu durum net bir şekilde görülmektedir. Yine de çıkış yapan turistler verisinde görüldüğü gibi bu durum her zaman baş göstermemektedir. Bu nedenle bu durum çelişki yaratmaktadır.

Bu çalışmada, MEÇO tekniğinin yapay sinir ağıları tekniğine yakın sayılabilecek güçlü bir teknik olduğu görülmüştür. MEÇO yönteminin zaman serileri analizinde daha başarılı olmasının önündeki en önemli engel, hata dağılımını veriden tahmin etmek olarak gözükmektedir. Bu çalışmada, hata dağılımının sadece çarpıklık ve basıklık üzerinden tahmin edilmesinin çok tehlikeli dezavantajları olduğu gösterilmiştir. Bu nedenle bu teknik için yapılması gereken en önemli iş, Logolabilirlik fonksiyonunda aramayı geliştirmektir. Böyle bir durumda yöntemin çok daha iyi işlediği, çalışmanın bu kısmında gösterilmiştir. Bunun yanında kısa kuyruklu dağılımlar ailesinin basıklık değerlerinin tekrar incelenmesi yöntem açısından faydalı olabilir. Bunun belirtilmesindeki amaç, kısa kuyruklu dağılımlarla ilgili elde edilmiş olan basıklık tablosunu ileriki çalışmalar için genişletmek ve hata dağılımlarını daha iyi tahmin etmektir. Bunun yanında bir zaman serisinde uzun kuyruklu dağılıma sahip bir hata dağılımı sezilmişse, bu durumda MEÇO tekniğini uygulamak çok zorlaşmaktadır. Bunun nedeni, istatistiksel hesaplamalarda kullanılan yazılımlar ve uzun kuyruklu dağılımdan sıralı istatistik elde etme yöntemidir. Böyle bir durumda sıralı istatistikleri elde etmek için en iyi yol simülasyon yapmaktır. Daha önceden bahsedildiği gibi, Trawinski et al. (1985) uzun kuyruklu dağılımın t-dağılımıyla ilişkili olduğunu belirtmiştir ve t-dağılımından veri üretilerek bu konuda bir çözüme ulaşmak mümkündür.

Turizm öngörülerini yapılırken verinin niteliği, öngörülerin doğruluğunu etkilemiştir. Örneğin, yurtdışından gelip çıkış yapan turist sayısı öngörülürken ortalama hata payı 80000 civarında tespit edilmiştir. Ancak genel olarak yurtdışından gelip çıkış yapan turist sayısı epey büyük olduğundan bunun makul kabul etmek mümkündür. Bu

konuda hata payına ilişkin konsensüs oluşturulması, gelecek çalışmalar için önemli olabilir. Geceleme oranlarına ilişkin öngörüler daha güvenilir durumdadır ve bu öngörülerle otel işletmelerinde isabetli düzenlemeler yapmak mümkündür. Harcama tutarlarına ilişkin öngörülerde 67 dolarlık ortalama hata payı bulunmuştur. Ülkemize gelecek turist miktarı ile birlikte turizm'den elde edilecek geliri incelemek mümkündür. Ancak böyle bir çalışmaya gidildiği zaman turist miktarını doğru öngörmek çok önemli ve zorlu olabilir. Örneğin, gelecek turist miktarını 10000'lik hata payıyla tahmin edilirse, bu durumda 670000 dolarlık bir zararı gözden çıkarmak gerekecektir. Bu zarar Türk Lirası cinsinden ifade edildiğinde bugün itibariyle 1432125 TL olmaktadır. Bu nedenle ülkemize gelecek turist sayısı öngörülürken, yapılması gereken en önemli şey pratik bir model oluşturmaktır. Pratik model geliştirmek açısından en iyi tercih, gücünü istatistiksel modellemekten alan bir yapay sinir ağı modelidir. Hata payını düşürmek için gizil tabaka bulundurmamak iyi bir tercih olabilir ancak sağlamcı istatistiksel bir modellemekten oluşturulan bir yapay sinir ağı mimarisinde gizil tabakalar hata payını yükseltebilmektedir. Bunun nedeni, gizil tabakanın sinaptik ağırlıkların dağılımını bozması ve sonunda elde edilen öngörülerin bundan etkilenmesidir.

Son olarak yapay sinir ağı yönteminin istatistiksel yöntemlerle beraber kullanıldığında, zaman serileri öngörülerinde daha başarılı olabileceği bu çalışmada gösterilmiştir.

KAYNAKLAR DİZİNİ

Ackley, D.H., Hinton, G. E. and Sejnowski, T.J., 1985, A learning algorithm for Boltzmann Machines, *Cognitive Science*, 9, 147-169.

Akkaya, A. and Tiku, M. L., 2005, Time series AR(1) model for short-tailed distributions, *Statistics: A Journal of Theoretical and Applied Statistics*, 39, 2, 117-131.

Akkaya, A. and Tiku, M.L., 2004, *Robust Estimation and Hypothesis Testing*, New Delhi: New Age International (P) Ltd., Publisher, 354 p.

Aladag, C.H., Yolcu, U. and Egrioglu, E., 2010, A high order fuzzy time series forecasting model based on adaptive expectation and artificial neural networks, *Mathematics and Computers in Simulation*, 81, 875–881.

Amari, S., 1972, Characteristics of random nets of analog neuron like elements, *IEEE Transactions on Systems, Man and Cybernetics*, SMC-2, 643-657.

Anderson, J.A., Silverstein, J.W., Ritz, S.A. and Jones, R.S., 1977, Distinctive features, categorical perception and probability learning: Some applications of a neural model, *Psychological Review*, 84, 413-451.

Ansley, C.F., 1979, An algorithm for the exact likelihood of a mixed autoregressive-moving-average process, *Biometrika*, 66, 59–65.

KAYNAKLAR DİZİNİ (devam ediyor)

Australian Bureau of Statistics, 1998a, Overseas Arrivals and Departures, 20 p.

Australian Bureau of Statistics, 1998b. Tourism Indicators, 20 p.

Azadeh, A., Neshat, N., Mardan, E. and Saberi, M., 2013, Optimization of steel demand forecasting with complete and uncertain economic inputs by an integrated neural network–fuzzy mathematical programming approach, *Int. Journal of Advanced Manufactured Technology*, 65, 833–841.

Bauwens, L., Lubrano, M. and Richard, J. F., 1999, *Bayesian Inference in Dynamic Econometric Models*, Oxford: Oxford University Press, 376 p.

Bass, F. M., and Clarke, D. G., 1972, Testing distributed lag models of advertising effect, *J. Marketing Research*, 9, 298-308.

Behrang, M.A., Assareh, E., Ghanbarzadeh, A. and Noghrehabadi, A.R., 2010, The potential of different artificial neural network (ANN) techniques in daily global solar radiation modeling based on meteorological data, *Solar Energy*, 84, 8, 1468–1480.

Bianco, A.M., Garcia Ben, M., Martinez, E.J and Yohai, V.J., 1996, Robust procedures for regression models with ARIMA errors, *Proceedings in Computational Statistics*, Albert Prat, 27–38.

Boashash, B., 2003, *Time-Frequency Signal Analysis and Processing: A Comprehensive Reference*, Elsevier Science, Oxford, 770 p.

KAYNAKLAR DİZİNİ (devam ediyor)

Box, G.; Jenkins, G. R., Gregory C., 1994, Time series analysis: forecasting and control, Prentice Hall, 612 p.

Brillinger, D. R., 1975, Time series: Data analysis and theory, New York: Holt, Rinehart. & Winston, 500 p.

Brigham, E. O., 1974, The fast Fourier transform, Prentice-Hall, 252p.

Brockwell, P.J. and Davis, R.A., 1991, Introduction to Time Series and Forecasting, New York: Springer.

Bryson, A.E. and Ho, Y.C., 1969, Applied Optimal Control: Optimization, Estimation and Control, Blaisdell, 496 p.

Bustos, O.H., 1982, General M-estimates for contaminated pth-order autoregressive processes: Consistency and asymptotic normality, Robustness in autoregressive processes. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete* , 59, 491–503.

Cheng, J. and Baldi, P., 2005, Three-stage prediction of protein β -sheets by neural networks, alignments and graph algorithms, *Bioinformatics*, 21, 75–83.

Coulibaly, P., Anctil, F. and Bobée, B., 2000, Daily reservoir inflow forecasting using artificial neural networks with stopped training approach, *Journal of Hydrology*, 230, 3–4, 244–257.

KAYNAKLAR DİZİNİ (devam ediyor)

Cortes, C. and Vapnik, V., 1995, Support vector networks, *Machine Learning*, 20, 273-297.

Cragg, B.G. and Tamperley, H.N.V., 1954, The organization of neurons: A cooperative analogy, *EEG Clinical Neurophysiology*, 6, 85-91.

Cragg, B.G. and Tamperley, H.N.V., 1955, Memory: The analogy with ferromagnetic hysteresis, *Brain*, 78, 304-316.

Dempster, A.P., Laird, N.M. and Rubin, D.B., 1977, Maximum likelihood from incomplete data via the EM algorithm, *J. R. Stat. Soc. B*, 39, 1-38.

Denby, L. and Martin, R.D., 1979, Robust estimation of the first-order autoregressive parameter, *Journal of the American Statistical Association*, 74, 140-146.

DeJong, D. and Shephard, N., 1995, The simulation smoother for time series models, *Biometrika*, 82, 339-350.

Durbin, J., 1960, The fitting of time series models, *Rev. Inst. Int. Stat.*, 28, 233-242.

Durbin J., and Koopman S.J., 2001, *Time Series Analysis by State Space Methods*, Oxford University Press, 346 p.

Elliott, D. F. and Rao, K. R., 1982, *Fast transforms: Algorithms, analyses, applications*, New York: Academic Press, 448 p.

KAYNAKLAR DİZİNİ (devam ediyor)

Everitt, B.S., 2006. The Cambridge Dictionary of Statistics, Cambridge University Press, 448 p.

Eskişehir Valiliği, 2011, Eskişehir turizm master planı, Anadolu Üniversitesi yayınları, 2426, 113 s.

Fleishman, A. I., 1978, A method for simulating non-normal distributions, Psychometrika, 43, 4, 521-531.

Filippov, A.F. and Arscott, F.M., 1988, Differential Equations with Discontinuous Right-hand Sides: Control Systems, Kluwer, 316 p.

Gelman, A., Carlin, J.B., Stern, H.S. and Rubin, D.B., 2000, Bayesian Data Analysis. Chapman & Hall, 556 p.

Gershenfeld, N., 1999, The nature of mathematical modeling, Cambridge University Press, 356 p.

Glover, F. and McMillan, C., 1986, The general employee scheduling problem: an integration of MS and AI, Computers and Operations Research, 13, 5, 563 – 573

Glover F., 1989, Tabu Search - Part 1, ORSA Journal on Computing, 1, 2, 190–206.

Glover F., 1990, Tabu Search - Part 2, ORSA Journal on Computing, 2, 1, 4–31.

KAYNAKLAR DİZİNİ (devam ediyor)

Grossberg, S., 1967, Nonlinear difference-differential equations in prediction and learning theory, Proceedings of the National Academy of Sciences, 58, 1329-1333.

Grossberg, S., 1968, A prediction theory for some nonlinear function-difference equations, Journal of Mathematical Analysis and Applications, 21, 643-693.

Grossberg, S., 1972, Neural expectation: Cerebellar and retinal analogs of cells fired by learnable or unlearned pattern classes, Cybernetik, 10, 49-57.

Grossberg, S., 1976a, Adaptive pattern classification and universal recoding: I Parallel development and coding of neural detectors, Biological Cybernetics, 23, 121-133.

Grossberg, S., 1976b, Adaptive pattern classification and universal recoding: II Feedback, expectation, olfaction, illusions, Biological Cybernetics, 23, 187-201.

Hadavandi E., Shavandi H. and Ghanbari A., 2010, Integration of genetic fuzzy systems and artificial neural networks for stock price forecasting, Knowledge-Based Systems, 23, 8, 800–808.

Hamilton, J., 1994, Time Series Analysis, Princeton Univ. Press, 820 p.

Hampel, F.R., 1971, A general definition of qualitative robustness, The Annals of Mathematical Statistics, 42, 1887–1896.

Hampel, F.R., 1974, The influence curve and its role in robust estimation, The Annals of Statistics, 69, 383–392.

KAYNAKLAR DİZİNİ (devam ediyor)

Haykin, S., 1998, Neural Networks: A Comprehensive Foundation, 2nd Edition, Prentice-Hall, 897 p.

Hebb, D.O., 2002, The Organization of Behavior: A Neuropsychological Theory, Psychology Press, 378 p.

Hopfield, J.J., 1982, Neural networks and Physical Systems with emergent collective computational abilities, Proceedings of the National Academy of Sciences, 79, 2554-2558

Hössjer, O., 1992, On the optimality of S-estimators, Statistics and Probability Letters, 14, 413–419.

Huang, W., Nakamori, Y. W. and Shou Y., 2005, Forecasting stock market movement direction with support vector machine, Computers & Operations Research, 32, 10, 2513–2521.

Huber, P.J., 1964, Robust estimation of a location parameter, The Annals of Mathematical Statistics, 35, 73–101.

Huber, P.J., 1967, The behavior of maximum likelihood estimates under nonstandard conditions, Proceedings of the Fifth Berkeley Symposium on Mathematics and Statistics Probability, University of California Press, 1, 221–232.

KAYNAKLAR DİZİNİ (devam ediyor)

Jhajharia, D., Chattopadhyay, S., Choudhary, R. R., Dev, Vas S., Vijay P. and Lal, S., 2013, Influence of climate on incidences of malaria in the Thar Desert, Northwest India. *Int. J. Climatol.*, 33, 312 – 325.

Khatibi, R., Naghipour, L., Ghorbani, M. A., Smith, M. S., Karimi, V., Farhoudi, R., Delafrouz, H. and Arvanaghi, H., 2013, Developing a predictive tropospheric ozone model for Tabriz Atmospheric Environment, 68, 286–293.

Kheirkhah, A., Azadeh, A., Saberi, M., Azaron, A. and Shakouri, H., 2013, Improved estimation of electricity demand function by using of artificial neural network, principal component analysis and data envelopment analysis, *Computers & Industrial Engineering*, 64, 1, 425–441.

Kirkpatrick, S, Gelatt, C.D. and Vecchi, M. P., 1983, Optimization by simulated annealing, *Science*, 220, 671-680.

Koop, G., 2003, *Bayesian Econometrics*. John Wiley & Sons, 376 p.

Law, R., 1988, Room Occupancy Rate Forecasting: A Neural Network Approach, *International Journal of Contemporary Hospitality Management*, 10 (6), 234-239.

Lawson, C. L., Hanson, R. J., 1987, *Solving Least Squares Problems*, Society for Industrial and Applied Mathematics, 350 p.

KAYNAKLAR DİZİNİ (devam ediyor)

Levinson, N., 1947, The Wiener RMS error criterion in filter design and prediction, J. Math. Phys., 25, 261–278.

Lin, J. and Keogh, E., Lonardi, S. and Chiu, B., 2003, A symbolic representation of time series, with implications for streaming algorithms, Proceedings of the 8th ACM SIGMOD workshop on Research issues in data mining and knowledge discovery.

Linsker, R., 1988, Self-organization in a perceptual network, Computer, 21,105-117.

Little, W.A. and Shaw, G.L., 1975, A statistical theory of short and long term memory, Behavioural Biology, 14, 115-132.

Ma, Y. and Genton, M. G., 2000, Highly robust estimation of the autocovariance function, Journal of Time Series Analysis, 21, 663–683.

Maronna, R. A., Martin, R.D. and Yohai, V.J., 2006, Robust Statistics: Theory and Methods. John Wiley & Sons, 423 p.

Martin, R.D., 1980, Robust estimation of autoregressive models, Directions in Time Series, 228–254, Haywood, CA: Institute of Mathematical Statistics.

Martin, R.D., 1981, Robust methods for time series, Applied Time Series Analysis II, 683–759.

KAYNAKLAR DİZİNİ (devam ediyor)

Minsky, M. L., 1954, Theory of neural-analog reinforcement systems and its application to the brain-model problem, Doktora tezi, Princeton Üniversitesi.

Mormann, F., Andrzejak, R. G., Elger, C. E. and Lehnertz, K., 2007, Seizure prediction: the long and winding road, *Brain*,130, 2, 314-332.

Engle, R., McFadden, D., 1999, *Handbook of Econometrics*, Vol.4, Holler, 1078 p.

Nikolić D., Muresan R.C., Feng W. and Singer W., 2012, Scaled correlation analysis: a better way to compute a cross-correlogram, *European Journal of Neuroscience*, 1–21

Nilsson, N.J., 1965, *Learning Machines:Foundations of Trainable Pattern-Classifying Systems*, New York:McGraw-Hill, 1078 p.

Ongenaë, F., Van Looy, S., Verstraeten, D., Verplancke, T., Benoit, D., De Turck, F., Dhaene, T., Schrauwen, B. and Decruyenaere, J., 2013, Time series classification for the prediction of dialysis in critically ill patients using echo-state networks, *Engineering Applications of Artificial Intelligence*, 26, 3, 984–996.

Oğcu, G., Zaim, S., Demirel, F. Ö., 2012, Forecasting Electricity Consumption with Neural Networks and Support Vector Regression, *Procedia Social and Behavioral Sciences* , 58, 1576-1585

Özaydın Ö., 2009, Porsuk barajı su seviyesinin öngörülerinin elde edilmesinde yapay sinir ağları ve ARMA modellerinin karşılaştırmalı incelenmesi, Doktora Tezi, Eskişehir Osmangazi Üniversitesi.

KAYNAKLAR DİZİNİ (devam ediyor)

Pearson, E. S. and Hartley, H. O., 1972, *Biometrika Tables for Statisticians*, Vol. 1., Cambridge University Press: Cambridge.

Pfanzagl, J., 1994, *Parametric statistical theory: with the assistance of R, Hamböcker*, Berlin DE: Walter de Gruyter., 207–208.

Qiao, C., Jing, W., Xu, Z., 2013, The UPPAM continuous-time RNN model and its critical dynamics study, *Neurocomputing*, 106, 158–166.

Ramedani, Z., Omid, M. and Keyhani, A., 2013, Modeling Solar Energy Potential in a Tehran Province Using Artificial Neural Networks, *International Journal of Green Energy*, 10, 4, 427-441.

Ropella, G.E.P., Nag, D.A. and Hunt, C.A., 2003, Similarity measures for automated comparison of in silico and in vitro experimental results, *Engineering in Medicine and Biology Society, Proceedings of the 25th Annual International Conference of the IEEE*, 3, 2933- 2936.

Rowland, J.H. and Varol, Y.L., 1972, Exit criteria for Simpson's compound rule, *Math Computation* 26, 119, 699-702.

Rosenblatt, F., 1958, *The Perceptron: A Probabilistic Model for Information Storage and Organization in the Brain*, Cornell Aeronautical Laboratory, *Psychological Review*, 65, 6, 386–408

KAYNAKLAR DİZİNİ (devam ediyor)

Rumelhart, D.E. and McClelland, J.L., 1986, *Parallel Distributed Processing: Explorations in the microstructure of cognition*, MIT press, MA.

Shannon, C.E., 1948, A mathematical theory of communication, *Bell System Technical Journal*, 27, 379-423 and 623-656 p.

Shannon, C.E. and McCarthy, 1956, *Automata Studies*, Princeton, NJ: Princeton University Press.

Stepnicka, M., Cortez, P., Donate J.P. and Stepnickova, L., 2013, Forecasting seasonal time series with computational intelligence: On recent methods and the potential of their combinations, *Expert Systems with Applications*, 40, 6, 1981–1991.

Shasha, D. , 2004, *High Performance Discovery in Time Series*, Berlin: Springer, 190 p.

Shamsuddin A., Cross, J., 1999, A Tourist Growth Model to Predict Accommodation Nights Spent in Australian Hotel Industry, SIRC 99-The 11th Annual Colloquium of the Spatial Information Research Centre University of Otago, Dunedin, New Zealand December 13-15th 1999

Shumway, R. H., 1988, *Applied statistical time series analysis*, NJ: Prentice Hall. 379 p.

Stigler, S., 1973, Simon Newcomb, Percy Daniell, and the history of robust estimation 1885–1920, *Journal of the American Statistics Association* , 68, 872–879.

KAYNAKLAR DİZİNİ (devam ediyor)

Tiku, M. L., 1967a, Estimating the mean and standard deviation from a censored normal sample, *Biometrika* 54, 155-165.

Tiku, M. L., 1967b, A note on estimating the location and scale parameters of the exponential distribution from a censored sample, *Austral. J. Statist.*, 9, 49-53.

Tiku, M. L., 1968a, Estimating the parameters of log-normal distribution from censored samples, *J. Amer. Stat. Assoc.*, 63, 134-140.

Tiku, M. L., 1968b, Estimating the parameters of normal and logistic distributions from censored samples, *Austral. J. Statist.*, 10, 64-73.

Tiku, M. L., 1968c, Estimating the mean and standard deviation from progressively censored normal samples, *J. Ind. Soc. Agric. Stat.*, 20, 20-25.

Tiku, M. L., 1970, Monte Carlo study of some simple estimators in censored normal samples, *Biometrika*, 57, 207-210.

Tiku, M. L., 1973, Testing group effects from Type II censored normal samples in experimental design, *Biometrics*, 29, 25-32.

Tiku, M. L., Islam, M. Q. and Selçuk, A. S., 2001, Nonnormal regression. II. symmetric distributions, *Communications in Statistics - Theory and Methods*, 30, 6,1021-1045

Tourism Forecasting Council, 1998, Forecast Visitor Statistics, 20 p.

KAYNAKLAR DİZİNİ (devam ediyor)

United Nations World Tourism Organization, Tourism Industry Trends, 1998, 20 p.

Trawinski, B.J., Bechhofer, R.E., Tamhane, A.C., Tiku, M.L., Kumra, S., 1985, Selected tables in mathematical statistics, American Mathematical Society, 270p.

Tukey, J.W., 1960, A survey of sampling from contaminated distributions, Contributions to Probability and Statistics, Stanford, CA: Stanford University Press.

Tukey, J.W., 1962, The future of data analysis, The Annals of Mathematical Statistics, 33, 1–67.

Vahidinasab V., Jadid S. and Kazemi A., 2008, Day-ahead price forecasting in restructured power systems using artificial neural networks, Electric Power Systems Research, 78, 8, 1332–1341.

Vapnik, V.N. and Chervonenkis, A. Y., 1971, On the uniform convergence of relative frequencies of events to their probabilities, Theoretical Probability and Its Applications, 17, 264-280.

Vapnik, V.N., 1982, Estimation of Dependences Based on Empirical Data, New York: Springer-Verlag.,505 p.

Vapnik, V.N., 1992, Principles of risk minimization for learning theory, Advances in Neural Information Processing Systems, 4, 831-838.

Vapnik, V.N., 2000, The Nature of Statistical Learning Theory, Springer, 314 p.

KAYNAKLAR DİZİNİ (devam ediyor)

Vapnik, V.N., 1998, Statistical Learning Theory, New York:Wiley, 736 p.

Vinod, H. D. and Shenton, L. R., 1996, Exact moment for autoregressive and random walk models for a zero or stationary initial value, *Econometric Theory* 12, 481-499.

Weigend, A. S. and N. A. Gershenfeld, 1994, Time Series Prediction: Forecasting the Future and Understanding the Past, Proceedings of the NATO Advanced Research Workshop on Comparative Time Series Analysis MA: Addison-Wesley, 26, 2, 302-304

Wiener, N., 1964, Extrapolation, Interpolation, and Smoothing of Stationary Time Series, The MIT Press, 163 p.

Wiener, N., 1961, Cybernetics: Or Control and Communication in the Animal and the Machine, New York: Wiley, 232 p.

Wilson, H.R. and Gowan J.D., 1972, Excitatory and inhibitory interactions in localized populations of model neurons, *Journal of Biophysics*, 12, 1-23.

Wu, C.F.J., 1983, On the convergence of the EM algorithm, *Annals of Statistics*, 11, 95–102.

Yamin, H.Y., Shahidehpour, S.M. and Li, Z., 2004, Adaptive short-term electricity price forecasting using artificial neural networks in the restructured power markets, *International Journal of Electrical Power & Energy Systems*, 26, 8, Pages 571–581.

KAYNAKLAR DİZİNİ (devam ediyor)

Yang, Z., Jin, L. and Wang, M., 2011, Forecasting Baltic Panama Inde with Support Vector Machine, *Journal of Transportation Systems Engineering and Information Technology*, 11, 3, 50–57.

Yeh, W. C., 2013, New Parameter-Free Simplified Swarm Optimization for Artificial Neural Network Training and its Application in the Prediction of Time Series, *IEEE Trans. Neural Netw. Learning Syst.*, 24, 4, 661-665.

Zhang, J., Tan, Z. and Shuxia Y., 2012, Day-ahead electricity price forecasting by a new hybrid method, *Computers & Industrial Engineering*, 63, 3, 695–701.

Zhu, B. and Wei, Y., 2013, Carbon price forecasting with a novel hybrid ARIMA and least squares support vector machines methodology, *Omega*, 41, 3, 517–523.