

**TRAFİK KAZALARININ ÇOK DEĞİŞKENLİ
İSTATİSTİKSEL YÖNTEMLERLE MODELLENMESİ**

Halim Ferit BAYATA

Doktora Tezi

İnşaat Mühendisliği Anabilim Dalı

Doç.Dr. Sinan HINISLIOĞLU

2010

Her Hakkı Saklıdır

ATATÜRK ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

DOKTORA TEZİ

**TRAFİK KAZALARININ ÇOK DEĞİŞKENLİ İSTATİSTİKSEL
YÖNTEMLERLE MODELLENMESİ**

Halim Ferit BAYATA

İNŞAAT MÜHENDİSLİĞİ ANABİLİM DALI

ERZURUM
2010

Her Hakkı Saklıdır

Doç. Dr. Sinan HINISLIOĞLU danışmanlığında, **Halim Ferit BAYATA** tarafından hazırlanan bu çalışma **01/10/2010** tarihinde aşağıdaki jüri tarafından, **İnşaat Mühendisliği Anabilim Dalı'nda Doktora** Tezi olarak kabul edilmiştir.

Başkan: Prof. Dr. Ömer AKBULUT

İmza : 

Üye: Doç. Dr. Sinan HINISLIOĞLU

İmza : 

Üye: Doç. Dr. Yetiş Şazi MURAT

İmza : 

Üye: Yrd. Doç. Dr. İbrahim CAN

İmza : 

Üye: Yrd. Doç. Dr. Osman Ünsal BAYRAK

İmza : 

Yukarıdaki sonucu onaylarım

(imza)

.....

Prof. Dr. Ömer AKBULUT

Enstitü Müdürü

ÖZET

Doktora Tezi

TRAFİK KAZALARININ ÇOK DEĞİŞKENLİ İSTATİSTİKSEL YÖNTEMLERLE MODELLENMESİ

Halim Ferit BAYATA

Atatürk Üniversitesi
Fen Bilimleri Enstitüsü

İnşaat Mühendisliği Anabilim Dalı

Danışman: Doç. Dr. Sinan HINISLIOĞLU

Türkiye’de meydana gelen trafik kazaları her geçen gün önemini korumaktadır. Kazaların oluş nedenlerini belirlemek için bazı çalışmalar yapılmakla beraber kazaların birden fazla değişkenden etkilenmesi, çok değişkenli istatistiksel analizi gerekli kılmaktadır.

Bu çalışmada, Türkiye’de meydana gelen trafik kazaları, trafik kaza tutanağı verileri kullanılarak, çok değişkenli zaman serileri, temel bileşenler analizi, yapay sinir ağları, çok değişkenli regresyon analizi, poisson regresyonu yöntemleri yardımıyla analiz edilmiştir. Zaman serisi analizlerinde Box-Jenkins (ARIMA(p,d,q)) ve VAR(p), yöntemleri kullanılmıştır. Mevcut verilerden, kazaları en çok etkilediği düşünülen değişkenler ile 7 adet model oluşturulmuştur. Modellerde ARIMA(p,d,q) ve YSA yöntemleri istatistiksel olarak en anlamlı yöntemler olmuştur. Nüfus, araç sayısı, sürücü sayısı ve Ağırlıklı Ortalama Günlük Trafik (AOGT) bağımsız değişkenleri ile kaza sayısı bağımlı değişkeni arasında tek yönlü bir Granger Nedenselliği bulunmuştur. Çok değişkenli regresyonda, nüfus, araç sayısı, ve AOGT arasında yüksek korelasyon gözlemlenmiştir. Yol-km uzunluğu ve Gayri Safi Milli Hasıla (GSMH), kaza sayısını en az açıklayan bağımsız değişkenler olmuştur. Kaza sayısı tahmini modellemelerinde YSA yöntemi istatistiksel olarak en anlamlı çıkmıştır.

2010, 165 sayfa

Anahtar Kelimeler: Trafik Kazaları, Zaman Serisi Analizi, Çok Değişkenli İstatistik, Yapay Sinir Ağları, Poisson Regresyonu, Temel Bileşenler Analizi

ABSTRACT

Ph. D. Thesis

MODELING OF TRAFFIC ACCIDENTS BY MULTIVARIATE STATISTICAL METHODS

Halim Ferit BAYATA

Atatürk University
Graduate School of Natural and Applied Sciences
Department of Civil Engineer

Supervisor: Assoc. Prof. Dr. Sinan HINISLIOĞLU

Traffic accidents in Turkey are maintaining their importance each passing day. Although some studies related to the reasons of accident occurrence are performed, the multivariate statistical modeling is required since more than one variable have influence on the traffic accidents.

In this study, traffic accidents in Turkey are analyzed by multivariate time series, principal component analysis, multivariate regression, Poisson regression and Artificial Neural Network (ANN) using the data obtained from traffic accident reports. Box-Jenkins and VAR(p) methods are used in time series analysis. 7 models are constructed with the variables thought to have most influence on the accidents. ARIMA(p,d,q) and ANN models are, in general, more statistically significant than the other models. It is found that there is one-way granger causality among the number of accidents, population, number of vehicle and Annually Average Daily Traffic (AADT). In the multivariate regression model, it is seen that there is high correlation among the population, number of vehicle and AADT. Highway length and Gross National Product (GNP) have least influence on the number of traffic accidents. ANN method has the most significance among multiregression model, poisson regression model and ANN model.

2010, 165 Pages

Keywords: Traffic Accidents, Time Series Analysis, Multivariate Statistics, Artificial Neural Network, Poisson Regression, Principal Component Analysis

TEŞEKKÜR

Çalışmalarımın her aşamasında desteğini esirgemeyen, bilgi ve birikimlerinden faydalandığım çok değerli hocam Sayın Doç. Dr. Sinan HINISLIOĞLU'na derin minnet ve şükranlarımı sunarım.

Tez çalışmaları süresinde değerleri fikirleri ve tecrübelerinden faydalandığım Sayın Prof. Dr. Ömer AKBULUT ve Sayın Yrd. Doç. Dr. İbrahim CAN'a teşekkürü bir borç bilirim.

Çalışmalarım esnasında yardımlarını esirgemeyen, bilgi ve tecrübelerinden faydalandığım çok değerli dostum Sayın Yrd. Doç. Dr. Yunus AKALTUN'a ve Sayın Yrd. Doç. Dr. Ahmet BARAN'a teşekkür ederim. Tez çalışmam süresinde ilgili yardımlarından dolayı Sayın Yrd. Doç. Dr. Yavuz KILAĞIZ'a, ayrıca Erzincan Üniversitesi Meslek Yüksekokulu Eski Müdürü Sayın Prof. Dr. Şakir BAYINDIR'a, Dr. Fatih HATTATOĞLU'na teşekkür ederim.

Erzincan Üniversite Vakfı'na verilerin Emniyet Genel Müdürlüğü'nden temininde sağlamış oldukları destekten dolayı teşekkür ederim.

Çalışmalarım boyunca bana göstermiş oldukları destekten dolayı çok değerli ailem ve eşime teşekkür ederim.

Halim Ferit BAYATA

Eylül, 2010

İÇİNDEKİLER

ÖZET.....	i
ABSTRACT.....	ii
TEŞEKKÜR.....	iii
SİMGELER VE KISALTMALAR DİZİNİ.....	vi
ŞEKİLLER DİZİNİ.....	vii
ÇİZELGELER DİZİNİ.....	xi
1.GİRİŞ	1
1.1. Amaç ve Kapsam	1
1.2. Kaynak Özetleri	3
1.3. Trafik Kazaları ve Dünya Analizi	6
1.4. Trafik Kazaları ve Türkiye Analizi	6
1.5. Trafik Kaza Analizleri ve Tahmin Modelleri.....	8
1.5.1. Çok değişkenli zaman serisi uygulamaları.....	8
1.5.2. Yapay sinir ağları uygulamaları	10
1.5.3. Çok değişkenli regresyon uygulamaları.....	11
1.5.4. Poisson ve negatif binomial regresyon uygulamaları	12
2. KURAMSAL TEMELLER	13
2.1. Zaman Serileri	13
2.1.1. Birim Kök Testleri	15
2.1.2. Model seçim kriterleri	18
2.1.3. Box-Jenkins yöntemi.....	20
2.1.3.a. Otoregresif süreç AR(p)	21
2.1.3.b. Hareketli ortalamalar MA(q)	22
2.1.3.c. Otoregresif bütünleşik hareketli ortalama ARIMA(p,d,q)	22
2.1.4. Çok değişkenli zaman serileri (VAR(p))	23
2.1.4.a. Vektör otoregresyon modellerinin belirlenmesi.....	25
2.1.4.b. Granger nedensellik analizi.....	27
2.1.4.c. Etki tepki fonksiyonu	27
2.1.4.d. Varyans analizi.....	28

2.2. Yapay Sinir Ağları	28
2.2.1. Aktivasyon ve transfer fonksiyonları	33
2.2.2. Yapay sinir ağları öğrenme algoritmaları	34
2.3. Çok Değişkenli Regresyon Analizi	34
2.3.1. Çok değişkenli regresyon analizi varsayımları	35
2.3.2. Çok değişkenli regresyon analizi model belirleme kriterleri	38
2.3.3. Temel bileşenler analizi	39
2.4. Poisson Regresyon Analizi	42
3. MATERYAL ve YÖNTEM.....	44
3.1. Çok Değişkenli (ÇDR), ZSA (VAR(p)) ve YSA modellemeleri	45
3.1.1. Model –I	45
3.1.2. Model –II.....	62
3.1.3. Model –III	75
3.1.4. Model –II-III YSA uygulaması	84
3.1.5. Model –IV	86
3.1.6. Model –V	96
3.1.7. Model –IV-V YSA uygulaması	104
3.2. Poisson Regresyonu ÇDR ve YSA Modellemeleri	105
3.2.1. Model –VI	106
3.2.2. Model –VII.....	110
4. ARAŞTIRMA BULGULARI ve TARTIŞMA.....	115
4.1. Çok Değişkenli ZSA (VAR(p)) YSA ve ÇDR Modellemeleri	115
4.1.1. Model –I	115
4.1.2. Model –II.....	122
4.1.3. Model –III	126
4.1.4. Model –II-III YSA uygulaması	130
4.1.5. Model –IV	133
4.1.6. Model –V	137
4.1.7. Model –IV-V YSA uygulaması	141
4.2. Poisson Regresyon ÇDR ve YSA Modellemeleri	144
4.2.1. Model –VI.....	144
4.2.2. Model –VII.....	147

5. SONUÇ ve ÖNERİLER	152
KAYNAKLAR	160
ÖZGEÇMİŞ	166

SİMGELER VE KISALTMALAR DİZİNİ

Simgeler

σ_e	Hata terimleri varyansı
Q_t	Q istatistik test değeri
Q^*	Q istatistiğinin geliştirilmiş hali
R^2	Belirtme katsayısı
$\bar{R}^2$	Düzeltilmiş belirtme katsayısı
ε_i	Hata terimi (artıklar)
T	Tolerans değeri
F	F testi tablo istatistik değeri
t	t testi tablo istatistik değeri
E(A)	Beklenen kaza sayısı
Pseudo R^2	Pseudo R^2 değeri
λ_i	Temel bileşenler analizinde özdeğerler
α_i	NBREG te dağılım parametresi
y_i	Bağımlı değişken
ρ_i	Standart sapma
μ_i	Ortalama
σ_i	Popülasyon varyansı
S_i	Örnek varyansı
β_i	Çok değişkenli regresyonda model katsayısı (tahmin edilen parametre)
X_i	Bağımsız değişkenler
$\Rightarrow$	Tek yönlü Granger Nedenselliği
$\Leftrightarrow$	Çift Yönlü Granger Nedenselliği
$\Updownarrow$	Granger Nedenselliği Yok

Kısaltmalar

ACF	Otokorelasyon fonksiyonu
ADF	Geniştirilmiş Dickey Fuller Birim kök testi
AIC	Akaike bilgi kriteri
AOGT	Ağırlıklı ortalama günlük trafik
AR(p)	Otoregresif süreç
ARIMA(p,d,q)	Bütünleşik otoregresif hareketli ortalama süreci
ÇDR	Çok değişkenli regresyon
ÇKA	Çok katmanlı ağ
DW	DurbinWatson test istatistiği
FPE	Nihai tahmin hatası (Final Prediction Error)
GSMH	Gayri safi milli hasıla
HKO	Karesel ortalama hata (mean square error) MSE
HKOK	Karesel ortalama hata karekökü (root mean square) RMS
HKT	Toplam karesel hata (sum square error) SSE
HQ	Hannan-Quinn bilgi kriteri (tutarlı gecikme seviyesi seçim kriteri)
IR	Etki-tepki analizi (Impulse-Response Analyses)
Km-taşıt	Türkiye’de km başına düşen taşıt sayısı
Km-yol	Türkiye’de km yol uzunlukları
MA(q)	Hareketli ortalamalar süreci
PACF	Kısmi Otokorelasyon fonksiyonu
SBC	Schwartz Bayes bilgi kriteri
TBA	Temel bileşenler analizi
VAR (p)	Vektörel otoregresyon modeli
VDC	Varyans ayrıştırması (Variance Decomposition)
YOGT	Yıllık ortalama günlük trafik
Yolcu-km	Türkiye’de km başına düşen yolcu sayısı
Yük-ton	Türkiye’de km başına düşen yük değeri
YSA	Yapay sinir ağı
ZSA	Zaman serisi analizi

ŞEKİLLER DİZİNİ

Şekil 2.1. YSA'da bir ağ mimarisi	30
Şekil 3.1. Model-I TBA temel bileşen sayısı için yamaç eğimi grafiği	49
Şekil 3.2. Model-I VAR(p) modeli hata terimleri ACF ve PACF grafikleri	56
Şekil 3.3. Model-I VAR(p) modeli hata terimleri Q-Q normallik analizi	56
Şekil 3.4. Model-I VAR(p) karakteristik kökler için birim çember grafiği	58
Şekil 3.5. Model-I VAR(p) modeli etki-tepki fonksiyonu grafiği	58
Şekil 3.6. Model-I ARIMA(1,1,0) hata terimleri ACF-PACF grafikleri	59
Şekil 3.7. Model-I ARIMA(1,1,0) karakteristik kök için birim çember grafiği	60
Şekil 3.8. Model-I YSA ağ mimarisi	62
Şekil 3.9. Model-II VAR(2) modeli hata terimleri Q-Q normallik analizi	72
Şekil 3.10. Model-II VAR(2) karakteristik kökler için birim çember grafiği	72
Şekil 3.11. Model-II VAR(2) modeli etki-tepki fonksiyonu grafiği	73
Şekil 3.12. Model-II VAR(2) modeli hata terimleri ACF-PACF grafikleri	73
Şekil 3.13. Model-II ARIMA(2,0,0) hata terimleri ACF-PACF grafikleri	74
Şekil 3.14. Model-II ARIMA(2,0,0) karakteristik kökler için birim çember grafiği	75
Şekil 3.15. Model-III modeli hata terimleri ACF-PACF grafikleri	80
Şekil 3.16. Model-III modeli hata terimleri Q-Q normallik analizi	81
Şekil 3.17. Model-III modeli VAR(1) karakteristik kökler için birim çember grafiği ..	82
Şekil 3.18. Model-III modeli VAR(1) modeli etki-tepki fonksiyonu grafiği	82
Şekil 3.19. Model-III modeli ARIMA(0,0,1) hata terimleri ACF-PACF grafiği	83
Şekil 3.20. Model-III ARIMA(0,0,1) karakteristik kökler için birim çember grafiği ...	84
Şekil 3.21. Model-II-III modeli YSA ağ mimarisi	85
Şekil 3.22. Model-IV modeli TBA temel bileşen sayısı için yamaç eğimi grafiği	89
Şekil 3.23. Model-IV modeli VAR(1) karakteristik kökler için birim çember grafiği ..	93
Şekil 3.24. Model-IV modeli VAR(1) hata terimleri ACF-PACF grafiği	93
Şekil 3.25. Model-IV modeli VAR(1) etki-tepki fonksiyonu grafiği	94
Şekil 3.26. Model-IV modeli ARIMA(1,0,0) hata terimleri ACF-PACF	95
Şekil 3.27. Model-IV ARIMA(1,0,0) karakteristik kökler için birim çember grafiği	96
Şekil 3.28. Model-V VAR(1) karakteristik kökler için birim çember grafiği	101
Şekil 3.29. Model-V modeli VAR(1) etki-tepki fonksiyonu grafiği	102

Şekil 3.30. Model-V modeli ARIMA(1,1,0) hata terimleri ACF-PACF grafiği	103
Şekil 3.31. Model-V ARIMA(1,1,0) karakteristik kökler için birim çember grafiği ...	104
Şekil 3.32. Model-IV-V YSA Ağ mimarisi	105
Şekil 3.33. Model-VI-VII Uygulama alanı	106
Şekil 3.34. Model-VI YSA Ağ mimarisi	110
Şekil 3.35. Model-VII YSA Ağ mimarisi	114
Şekil 4.1. Model-I TBA ile tarihi seri serpilme ve uyum grafikleri.....	116
Şekil 4.2. Model-I VAR(1) tarihi seri serpilme ve uyum grafikleri	117
Şekil 4.3. Model-I ARIMA(1,1,0) tarihi seri serpilme ve uyum grafikleri.....	118
Şekil 4.4. Model-I Örnek YSA ağ mimarisi	119
Şekil 4.5. Model-I YSA eğitim setinin tarihi seri serpilme ve uyum grafikleri.....	120
Şekil 4.6. Model-I YSA test setinin tarihi seri serpilme ve uyum grafikleri	121
Şekil 4.7. Model-II TBA ile tarihi seri serpilme ve uyum grafikleri	123
Şekil 4.8. Model-II VAR(2) tarihi seri serpilme ve uyum grafikleri	124
Şekil 4.9. Model-II ARIMA(1,1,0) tarihi seri serpilme ve uyum grafikleri	125
Şekil 4.10. Model-III TBA ile tarihi seri serpilme ve uyum grafikleri	127
Şekil 4.11. Model-III VAR(2) tarihi seri serpilme ve uyum grafikleri.....	128
Şekil 4.12. Model-III ARIMA(0,1,1) tarihi seri serpilme ve uyum grafikleri	129
Şekil 4.13. Model-II-III ölü sayıları için YSA eğitim setinin tarihi seri serpilme ve uyum grafikleri	131
Şekil 4.14. Model-II-III yaralı sayıları için YSA eğitim setinin tarihi seri serpilme ve uyum grafikleri	131
Şekil 4.15. Model-II-III ölü sayıları için YSA test setinin tarihi seri serpilme ve uyum grafikleri	134
Şekil 4.16. Model-II-III yaralı sayıları için YSA test setinin tarihi seri serpilme ve uyum grafikleri	132
Şekil 4.17. Model-II-III Örnek YSA ağ mimarisi.....	133
Şekil 4.18. Model-IV TBA ile tarihi seri serpilme ve uyum grafikleri.....	134
Şekil 4.19. Model-IV VAR(1) tarihi seri serpilme ve uyum grafikleri.....	135
Şekil 4.20. Model-IV ARIMA(1,1,0) tarihi seri serpilme ve uyum grafikleri.....	136
Şekil 4.21. Model-V TBA ile tarihi seri serpilme ve uyum grafikleri	138
Şekil 4.22. Model-VI VAR(1) tarihi seri serpilme ve uyum grafikleri.....	139

Şekil 4.23. Model-IV ARIMA(1,1,0) tarihi seri serpilme ve uyum grafikleri.....	140
Şekil 4.24. Model-IV-V ölü sayıları için YSA eğitim setinin tarihi seri serpilme ve uyum grafikleri.....	141
Şekil 4.25. Model-IV-V yaralı sayıları için YSA eğitim setinin tarihi seri serpilme ve uyum grafikleri	142
Şekil 4.26. Model-IV-V ölü sayıları için YSA test setinin tarihi seri serpilme ve uyum grafikleri	142
Şekil 4.27. Model-IV-V yaralı sayıları için YSA test setinin tarihi seri serpilme ve uyum grafikleri	143
Şekil 4.28. Model-IV-V Örnek YSA ağ mimarisi	144
Şekil 4.29. Model-VI Poisson uyum ve serpilme grafiği.....	145
Şekil 4.30. Model-VI ÇDR uyum ve serpilme grafiği	146
Şekil 4.31. Model-VI YSA uyum ve serpilme grafiği	147
Şekil 4.32. Model-VII Poisson uyum ve serpilme grafiği	149
Şekil 4.33. Model-VII ÇDR uyum ve serpilme grafiği.....	149
Şekil 4.34. Model-VII YSA test seti uyum ve serpilme grafiği.....	150
Şekil 4.35. Model-VII YSA eğitim seti uyum ve serpilme grafiği	150

ÇİZELGELER DİZİNİ

Çizelge 1.1. Çok değişkenli zaman serileri model dereceleri	8
Çizelge 1.2. YSA'da temel bileşen sayısı	10
Çizelge 3.1. Model-I değişkenler arası normallik test sonuçları.....	46
Çizelge 3.2. Model-I değişkenler arası korelasyon matrisi.....	46
Çizelge 3.3. Model-I TBA küresellik testi sonuçları	47
Çizelge 3.4. Model-I TBA özvektörler	48
Çizelge 3.5. Model-I TBA toplam varyansın açıklama oranları	48
Çizelge 3.6. Model-I TBA değişkenlerin temel bileşenler tarafından açıklanma oranı ...	50
Çizelge 3.7. Model-I TBA temel bileşenler matrisi	50
Çizelge 3.8. Model-I TBA bağımsız değişken ile TBA skorları korelasyon matrisi	51
Çizelge 3.9. Model-I TBA regresyon modeli çıktıları	51
Çizelge 3.10. Model-I TBA regresyon modeli parametreleri	51
Çizelge 3.11. Model-I İkinci TBA için katsayılar ve EKK parametreleri	53
Çizelge 3.12. Model-I VAR(p) modeli birim kök testi sonuçları	53
Çizelge 3.13. Model-I VAR(p) model derecesi belirleme kriterleri	54
Çizelge 3.14. Model-I VAR(p) Granger Nedensellik analiz sonuçları.....	55
Çizelge 3.15. Model-I VAR(p) Hata terimleri normallik analiz sonuçları	56
Çizelge 3.16. Model-I VAR(p) otokorelasyon LM testi sonuçları	57
Çizelge 3.17. Model-I VAR(p) modeli varyans ayrıştırması sonuçları	59
Çizelge 3.18. Model-I ARIMA(1,1,0) hata terimleri normallik analizi sonuçları	60
Çizelge 3.19. Model-I YSA modeli optimum model çıktıları	61
Çizelge 3.20. Model-II değişkenler arası normallik test sonuçları	63
Çizelge 3.21. Model-II değişkenler arası korelasyon matrisi	64
Çizelge 3.22. Model-II TBA küresellik testi sonuçları	64
Çizelge 3.23. Model-II TBA özvektörler	65
Çizelge 3.24. Model-II TBA toplam varyansın açıklama oranları.....	65
Çizelge 3.25. Model-II TBA değişkenlerin temel bileşenler tarafından açıklanma oranı	66
Çizelge 3.26. Model-II TBA temel bileşenler matrisi.....	67
Çizelge 3.27. Model-II TBA bağımsız değişken ile TBA skorları korelasyon matrisi ..	67
Çizelge 3.28. Model-II TBA regresyon modeli çıktıları.....	68

Çizelge 3.29. Model-II TBA regresyon modeli parametreleri.....	68
Çizelge 3.30. Model-II İkinci TBA için katsayılar ve EKK parametreleri.....	69
Çizelge 3.31. Model-II VAR(p) modeli birim kök testi sonuçları.....	69
Çizelge 3.32. Model-II VAR(p) model derecesi belirleme kriterleri.....	70
Çizelge 3.33. Model-II VAR(2) Granger Nedensellik analiz sonuçları.....	70
Çizelge 3.34. Model-II VAR(2) Hata terimleri normallik analiz sonuçları.....	71
Çizelge 3.35. Model-II VAR(2) modeli varyans ayrıştırması sonuçları.....	74
Çizelge 3.36. Model-II ARIMA(2,0,0) hata terimleri normallik analizi sonuçları.....	74
Çizelge 3.37. Model-III bağımsız değişken için normallik test sonucu.....	76
Çizelge 3.38. Model-III değişkenler arası korelasyon matrisi	76
Çizelge 3.39. Model-III TBA bağımsız değişken ile TBA skorları korelasyon matrisi .	77
Çizelge 3.40. Model-III TBA regresyon modeli çıktıları	77
Çizelge 3.41. Model-III TBA regresyon modeli parametreleri.....	77
Çizelge 3.42. Model-III İkinci TBA için katsayılar ve EKK parametreleri.....	78
Çizelge 3.43. Model-III VAR(p) modeli birim kök testi sonuçları.....	78
Çizelge 3.44. Model-III VAR(p) model derecesi belirleme kriterleri	79
Çizelge 3.45. Model-III VAR(1) Granger Nedensellik analiz sonuçları	79
Çizelge 3.46. Model-III VAR(1) Hata terimleri normallik analiz sonuçları.....	81
Çizelge 3.47. Model-III VAR(1) modeli varyans ayrıştırması sonuçları	83
Çizelge 3.48. Model-III ARIMA(0,0,1) hata terimleri normallik analizi sonuçları	83
Çizelge 3.49. Model-IV değişkenler için normallik test sonucu.....	86
Çizelge 3.50. Model-IV değişkenler arası korelasyon matrisi	87
Çizelge 3.51. Model-IV TBA küresellik testi sonuçları	87
Çizelge 3.52. Model-IV TBA özvektörler	87
Çizelge 3.53. Model-IV TBA toplam varyansın açıklama oranları	88
Çizelge 3.54. Model-IV TBA değişkenlerin temel bileşenler tarafından açıklanma oranı	89
Çizelge 3.55. Model-IV TBA temel bileşenler matrisi	89
Çizelge 3.56. Model-IV TBA bağımsız değişken ile TBA skorları korelasyon matrisi .	90
Çizelge 3.57. Model-IV TBA regresyon modeli çıktıları	90
Çizelge 3.58. Model-IV TBA regresyon modeli parametreleri	90
Çizelge 3.59. Model-IV İkinci TBA için katsayılar ve EKK parametreleri	91
Çizelge 3.60. Model-IV VAR(p) modeli birim kök testi sonuçları	91

Çizelge 3.61. Model-IV VAR(p) model derecesi belirleme kriterleri	92
Çizelge 3.62. Model-IV VAR(1) Granger Nedensellik analiz sonuçları	92
Çizelge 3.63. Model-IV VAR(1) Hata terimleri normallik analiz sonuçları	92
Çizelge 3.64. Model-IV VAR(1) modeli varyans ayrıştırması sonuçları	94
Çizelge 3.65. Model-IV ARIMA(1,0,0) hata terimleri normallik analizi sonuçları	95
Çizelge 3.66. Model-V bağımsız değişken için normallik test sonucu.....	97
Çizelge 3.67. Model-V değişkenler arası korelasyon matrisi	97
Çizelge 3.68. Model-V TBA bağımsız değişken ile TBA skorları korelasyon matrisi ..	98
Çizelge 3.69. Model-V TBA regresyon modeli çıktıları.....	98
Çizelge 3.70. Model-V TBA regresyon modeli parametreleri.....	98
Çizelge 3.71. Model-V TBA için katsayılar ve EKK parametreleri	99
Çizelge 3.72. Model-V VAR(p) modeli birim kök testi sonuçları.....	99
Çizelge 3.73. Model-V VAR(1) model derecesi belirleme kriterleri.....	99
Çizelge 3.74. Model-V VAR(1) Granger Nedensellik analiz sonuçları	96
Çizelge 3.75. Model-V VAR(1) Hata terimleri normallik analiz sonuçları.....	96
Çizelge 3.76. Model-V ARIMA(1,1,0) hata terimleri normallik analizi sonuçları.....	102
Çizelge 3.77. Model-VI Poisson regresyonu model parametreleri	107
Çizelge 3.78. Model-VI Korelasyon matrisi	108
Çizelge 3.79. Model-VI Regresyon model çıktıları	108
Çizelge 3.80. Model-VI Regresyon model parametreleri	109
Çizelge 3.81. Model-VI Optimum YSA ağ mimarisi	109
Çizelge 3.82. Model-VII Poisson regresyonu model parametreleri	111
Çizelge 3.83. Model-VII Korelasyon matrisi.....	112
Çizelge 3.84. Model-VII Regresyon model çıktıları.....	112
Çizelge 3.85. Model-VI Regresyon model parametreleri	113
Çizelge 3.86. Model-VI Optimum YSA ağ mimarisi	113
Çizelge 4.1. Model-I ÇDR model belirleme kriterleri	116
Çizelge 4.2. Model-I VAR(p) model belirleme kriterleri	118
Çizelge 4.3. Model-I ARIMA(1,1,0) model belirleme kriterleri	119
Çizelge 4.4. Model-I YSA model belirleme kriterleri	121
Çizelge 4.5. Model-I Yöntemlerin karşılaştırılması.....	121
Çizelge 4.6. Model-II ÇDR model belirleme kriterleri.....	123

Çizelge 4.7. Model-II VAR(p) model belirleme kriterleri	124
Çizelge 4.8. Model-II ARIMA(1,1,0) model belirleme kriterleri	125
Çizelge 4.9. Model-II YSA model belirleme kriterleri	126
Çizelge 4.10. Model-III ÇDR model belirleme kriterleri	127
Çizelge 4.11. Model-III VAR(p) model belirleme kriterleri.....	129
Çizelge 4.12. Model-III ARIMA(1,1,0) model belirleme kriterleri.....	130
Çizelge 4.13. Model-III YSA model belirleme kriterleri.....	130
Çizelge 4.14. Ölü sayıları için YSA modeli model belirleme kriterleri.....	132
Çizelge 4.15. Yaralı sayıları için YSA modeli model belirleme kriterleri.....	133
Çizelge 4.16. Model-IV ÇDR model belirleme kriterleri	134
Çizelge 4.17. Model-IV VAR(p) model belirleme kriterleri	135
Çizelge 4.18. Model-IV ARIMA(1,1,0) model belirleme kriterleri	136
Çizelge 4.19. Model-IV YSA model belirleme kriterleri	137
Çizelge 4.20. Model-V ÇDR model belirleme kriterleri.....	138
Çizelge 4.21. Model-V VAR(p) model belirleme kriterleri.....	139
Çizelge 4.22. Model-V ARIMA(1,1,0) model belirleme kriterleri.....	140
Çizelge 4.23. Model-V YSA model belirleme kriterleri.....	140
Çizelge 4.24. Ölü sayıları için YSA modeli model belirleme kriterleri.....	143
Çizelge 4.25. Yaralı sayıları için YSA modeli model belirleme kriterleri.....	143
Çizelge 4.26. Model-VI Yöntemlerin karşılaştırılması.....	147
Çizelge 4.27. Model-VII Yöntemlerin karşılaştırılması	151

1. GİRİŞ

1.1. Amaç ve Kapsam

Dünya genelinde yıllık ortalama 1.3 milyonun üzerinde insan ölmekte 50 milyonun üzerinde insan yaralanmakta veya sakat kalmaktadır. Ülkemizde de kaza istatistikleri pek iç açıcı değildir. Trafığe kayıtlı motorlu araç sayısı 13 milyon, ehliyetli sürücü sayısı ise 18 milyonun üzerindedir. Yıllık ortalama 1.3 milyon ölümlü, yaralanmalı ve maddi hasarlı trafik kazası meydana gelmektedir. Her yıl ülkemizde 5000 kişi trafik kazaları nedeniyle ölmektedir. Yani günde ortalama 13 kişi ölmekte ve her 1,2 dakikada bir trafik kazası meydana gelmekte, yüzbinin üzerinde insanımız da yaralanmakta veya sakat kalmaktadır. 1980–2010 yılları arasında trafik kazalarında ülkemizde 40000 den fazla insan ölmüştür. Aynı yıllar arasında ülkemizde meydana gelen terör olaylarında, trafik kazalarından daha az insanımızı kaybetmiş bulunmaktayız.

Bu kadar büyük bir problem karşısında trafik kazalarına ilişkin istatistiksel bir takım analizler yapmak ve bu analizlerle kaza sayılarını etkileyen değişkenleri tespit etmek ve bu değişkenlerin kendi aralarındaki ilişkileri ortaya incelemek, gelecek yıllardaki kaza sayılarını, bu kazalardaki ölüm ve yaralı sayılarını tespit etmek bu çalışmanın amaçları arasında yer almaktadır.

Bu çalışmada yer alan veriler Karayolları Genel Müdürlüğü, Karayolları 12. Bölge Müdürlüğü, Emniyet Genel Müdürlüğü Trafik Eğitim ve Araştırma Daire Başkanlığı, Türkiye İstatistik Kurumu (TÜİK)’ten alınmıştır. Tercan-Erzurum-Ağrı-Gürbulak sınır kapısına kadar olan 422 km uzunluğundaki Kuzey Transit Yoluna ilişkin (E–80 veya D–100), yol genişliği, YOGT, şerit sayısı, kaza sayısı, yaralı sayısı, ortalama hız, uzunluk km, bağlantı sayısı ve kavşak sayısı, değişkenleri, Karayolları 12. Bölge Müdürlüğü’nden elde edilmiştir.

Detay nitelikli deęiřkenlerimiz ile negatif binomial regresyon, çok deęiřkenli regresyon, yapay sinir aęları yöntemleri kullanılmıřtır. Türkiye genelini kapsayan deęiřkenlerden baęımlı deęiřken olarak kaza sayısı, ölü sayısı ve yaralı sayısı, baęımsız deęiřken olarak ise nüfus, araç sayısı, otobüs, kamyon, tır, ağır tonajlı araç sayısı, aęırlıklı ortalama günlük trafik (AOGT), karayollarında enerji tüketimi, yol-km, yolcu-km, yük-ton kullanılmıřtır. Modelleme aşamasına girmeden önce aralarında yüksek korelasyon bulunan baęımsız deęiřkenlere Temel Bileřenler Analizi (TBA) uygulanmıřtır. Birbirleri ile korelasyonlu p kadar deęiřken yerine bir birleri ile korelasyonsuz p kadar temel bileřen ile modeller oluřturulmuřtur. Modeller oluřturulurken, çok deęiřkenli regresyon, çok deęiřkenli Zaman Serileri Analizi (VAR(p)), ARIMA(p,d,q), ve Yapay Sinir Aęları (YSA) yöntemleri kullanılmıřtır.

Kullanılan modellerde bu yöntemler, belirleme katsayısı (R^2), ortalama karesel hata (OKH), ve Akaike Bilgi Kriteri (AIC) kriterlerine göre sınıflandırılmıřtır. Minimum OKH ve minimum AIC deęerini veren iki yöntem arasında en yüksek R^2 deęerini veren yöntem en iyi yöntem olarak belirlenmiřtir.

Çok deęiřkenli zaman serisi analizlerinde, Granger Nedensellik analizi belirleyici bir rol almıřtır. Deęiřkenler arasındaki nedensellik, modeli oluřtururken parametre sayısını etkilemiřtir. Aynı zamanda bu yöntemde etki-tepki analizi yapılmıř baęımlı deęiřkene verilecek bir birimlik bir řokun dięer deęiřkenlerde nasıl bir tepki vereceęi grafiksel olarak yorumlanmıřtır. Varyans ayrıřtırması yapılarak serilerimizin öngörü hata varyansının dięer deęiřkenler tarafından açıklanma oranları belirlenmiřtir. Baęımlı deęiřkenin baęımsız deęiřkenler tarafından etkilenme oranları da varyans ayrıřtırmasında tespit edilmiřtir.

2008–2009 yıllarında meydana gelen kaza sayıları, bu kazalardaki ölü sayıları ve yaralı sayıları kullanılan modellerde ki tahmini deęerlerle yakın çıkmıřtır. Bu yıllarda gerçeķleşen kaza sayıları ve modellerden tahmin edilen kaza sayıları %95 güven aralığında bulunmuřtur. Kazaları etkileyen deęiřkenlerin başında nüfus, araç sayısı, km-

taşıt, AOGT tespit edilmiştir. Aynı zamanda bu değişkenlere ilişkin tespitler Camkesen (1998) ve Cribbins *et al.* (1967) tarafından da çalışılmıştır.

1.2. Kaynak Özetleri

Trafik kaza modellemesi ile ilgili literatürde bir çok çalışma mevcut olup bunlardan bazıları aşağıda sunulmuştur.

Kihlberg *et al.* (1968) çalışmasında şerit sayısı, orta refüj ve bağlantı noktalarının kontrol durumuna göre sınıflandırmış homojen yol kesimlerindeki kaza şiddetini modellemede lineer regresyon analizini kullanmışlardır.

William *et al.* (1973) yaptıkları çalışmada, tek başına kaza sayılarını modellemiştir. Çalışmada kaza analizden ziyade Bayesci bir yaklaşım ile bir kesitteki kaza potansiyelinin analiz edilmesi gerektiği önerilmiştir.

Jovanis and Chang (1986) trafik kazaları gibi rastgele kesikli değişkenler için, geleneksel lineer regresyon modeli negatif olmayan olayları açıklayamayacağından genelleştirilmiş lineer regresyon modelini kullanmışlardır.

Nakatsyii (1989), Altun vd (2005) yapay sinir ağları yardımıyla trafik akım kontrolü isimli bir çalışma yapmış, bu çalışmayı Türkçeye çevirmişlerdir. Bu çalışmada sürücülerin karakteristik özellikleri (yaş, cinsiyet, alkol durumu, emniyet kemeri), araç tipi, hız, bölge özellikleri, aydınlık durumu gibi değişkenler girdi katmanında yer almıştır. Çıkış katmanında ise kazanın hasarsız, hafif hasarlı, orta şiddetli, ağır ve ölümcül olup olmadıkları yapay sinir ağları, fuzzy (bulanık mantık), probit model ile modellenmeye çalışılmış ve yöntemler kıyaslanmıştır. Aynı çalışmada trafik akımı ve hızı tahmin edilmeye çalışılmıştır. Yoğunluk, hava durumu, görüş, açıklık ve ağır taşıt yüzdesi değişkenleri alınmış, 360.000 trafik kaydının içinden 2688 adeti seçilmiş, verilerin yarısı ağırlık eğitimi yarısında ağırlık testi için kullanılmıştır.

Persaud (1991) YOGT ve kesim uzunluğu değişkenlerini kullanarak, kaza sayısını NB Regresyon yöntemini kullanarak modellemiştir.

İyınam vd (1998), tek değişkenli zaman serisi analizi ile aylık trafik kaza sayılarını modellemeye çalışmış kaza sayılarının 20 yıllık bir süreçte ne şekilde arttığını gözlemlemeye çalışmıştır.

Camkesen (1998), trafik kazalarının oluşumunda birden fazla faktörün rol aldığını ifade etmektedir. Bu faktörlerin tamamının birbirleri ile ilişkileri, yolların geometrik özellikleri, çevre koşulları ve sürücü karakteristikleri, kazaların meydana gelmesinde büyük rol oynamaktadırlar.

YOGT kullanılan birçok modelde değişken olarak kabul edilmiştir. Cribbins *et al.* (1967) yapmış olduğu çalışmada YOGT, refüj genişliği, hız limiti, gibi 8 adet değişkeni kullanarak kaza tahmin modeli geliştirmiştir (Camkesen 1998).

Delen *et al.* (2006), YSA kullanarak hangi durumlarda sürücülerin ve yolcuların öldüğü veya yaralandığını anlamak için sürüş güvenliğini artırmaya yardımcı olabilecek parametreler geliştirmek için çalışmıştır. Trafik kazalarında yaralanmaları artıran riskler birçok unsur içerir. Bunlardan sürücülerin karakteristik özellikleri (davranış ve kişiliği), çevresel faktörler, yol şartları (kaza anında), aracın teknik özelliklerini, Yapay sinir ağları kullanılarak trafik kazalarında yaralanma şiddetini tahmin etmeye çalışmışlardır.

Ferrer-Garcia *et al.* (2006), zaman serileri ile ayrıştırılmış veriler kullanarak trafik kaza tahmini yapmaya çalışmışlardır. Kullanılan modellerde aylık olarak ölçülen farklı karakteristik özellik gösteren trafik kaza verileri kullanılmıştır. Özellikle trafik güvenlik kampanyalarının değerlendirilmesinde zorunlu değişkenlerin etkileri göz önüne alınmıştır.

Mohamed *et al.* (2000), trafik kazalarının meydana geliş şekli ve ilişkilerin modellenmesi çalışmasında NB regresyon yöntemini kullanmışlardır. Yolların geometrik özellikleri ve sürücü karakteristiklerini değişken olarak almışlardır. Hız, orta refüj genişliği, şerit genişliği, şerit sayısı ve uluslar arası yollarda yol bağlantı sayılarının, kazalarla ilişkili olduğunu göstermiştir. Genç erkek sürücülerin bayan sürücülere göre ağır trafik hacmi olan yerlerde daha dikkatsiz oldukları ve daha hızlı gittiklerini tespit etmişlerdir. Genç sürücülerin yaşlı ve orta yaşlı sürücülere göre kurplarda daha hızlı araç sürdüklerini ve bunun da kazalara neden olduğunu belirlemişlerdir.

Miao *et al.* (2003), trafik kaza analizlerinde karar ağacı ve yapay sinir ağları kullanarak kaza modellemesi yapmışlardır. Sürücülerin emniyet kemeri kullanımı, yol hava şartları ve sürücülerin alkol durumları değişken olarak alınmıştır. Değişkenler ölümlü ve yaralanmalı kazalar olarak sınıflandırmış, kazalar yapay sinir ağları ve bulanık mantık ile modellenmiştir.

Greibe (2003), Danimarka'daki şehirlerarası yollar için kaza tahmin modellemesi ile çok değişkenli regresyon modeli geliştirmiştir. Çalışmasında 142 km 'lik bir yol kesiminde 1036 tali yol bağlantısı üzerinde yaptığı çalışmada, kaza sıklığı, yıllık ortalama günlük trafik, hız limiti, yol bağlantı uzunlukları, şerit sayısı, çift tek yön, yol genişliği, hız azaltan ölçümler, çıkışlar, yan yol sayıları, bisiklet imkanı, yaya yolu, park imkanı ve refüj gibi değişkenleri ele alarak bir tahmin modeli geliştirmiştir.

Lee *et al.* (2008), çalışmasında yapısal denklem modeli kullanarak Kore karayollarındaki trafik kaza büyüklüğü analizini yapmışlardır. Kaza şiddeti, toplam araç sayısı, toplam hasarlı araç sayısı, toplam ölü ve yaralı sayısı ile ifade edilir denilmiştir. Güvenli ulaşım özelliklerinin en önemli ölçüm indeksinin kaza büyüklüğü olduğu yine bu çalışmada belirtilmiştir.

Quddus (2008), çalışmasında kaza analizi uygulamalarında yeni kullanılmaya başlanılan Box-Jenkins (ARIMA) ve mevsimsel (SARIMA) zaman serisi metodu ile INAR Poisson ve NB negatif binom yöntemlerini kullanmış, 1985-2007 yılları arasında kaza sayıları km-yol uzunluklarını modelleyerek yöntemleri karşılaştırılmıştır.

1.3. Trafik Kazaları ve Dünya Analizi

Dünyada her yıl trafik kazalarından 1.3 milyon insan ölmekte 50 milyon insan da yaralanmaktadır. Dünyada yaralanmalı kazalarda trafik kazaları birinci sırada yer almaktadır. Analizciler nüfus arttıkça ve trafiğe kayıtlı karadaki motorlu araç sayısı arttıkça kazalarda ki artışın devam edeceğini bildirmişlerdir. Karayolu ulaşımı, ulaşım sistemleri içinde en karmaşık olanı olmamasına rağmen en fazla ölümlü yaralanmalı kazanın olduğu sistemdir. Bugüne kadar bu sistem üzerine yapılan çalışmalarda kazaların azaltılmasına ilişkin somut bir gelişme olmamıştır. Havayolu, demiryolu ve denizyolu kazalarında kazaya sebep olan bir veya iki değişken varken, karayollarında trafik kazalarına sebep olan birbirleri ile ilişkili birden fazla karmaşık ilişki söz konusudur.

Dünya Sağlık Örgütü 1990 yılında evrensel felaket sıralamasında trafik kazalarını 9. sırada göstermiştir. 2020 yılında bu sıralamanın üçüncü en büyük felaket olacağı sonucu öngörülmektedir (Miranda-Moreno 2006; Kibar 2008).

2007 yılı verilerine göre Avrupa ülkelerinden İngiltere’de kazalardaki ölü sayılarının kazaya oranı %1.6, İtalya’da %2.8, Fransa’da ise %6.4’tür. Gelişmiş ülkelerde de trafik kazaları önemini korumaktadır.

1.4. Trafik Kazaları ve Türkiye Analizi

Ülkemizde trafiğe kayıtlı araç sayısı 14 milyonun üzerindedir. 2010 yılında trafiğe kayıtlı araç sayısı aylık olarak bir önceki aya göre ortalama %40 artış göstererek devam

etmiştir. Ehliyetli sürücü sayısı da 18 milyonun üzerinde bulunmaktadır. Yıllık ortalama, ölümlü, yaralanmalı ve maddi hasarlı 1.3 milyon trafik kazası meydana gelmektedir. Yine ülkemizde yıllık ortalama 5000 kişi ölmektedir. Ortalama bir ilçemizin nüfus büyüklüğü 5000 kabul edilirse yılda bir ilçe nüfusu ölmektedir. 100 bin kişi bir yılda trafik kazalarında yaralanmaktadır. Son 10 yıllık maddi kayıp 100 milyar doların üzerindedir.

Rakamların bu kadar büyük olmasındaki bazı sebepler; ülkemizdeki yolların son döneme kadar standartlara uymaması, ülkemizde karayolu ulaşımının diğer ulaşım türleri içindeki kullanım oranının %90'nın üzerinde yer alması, trafik eğitiminin yeterince verilememesi, yolun geometrik özelliklerinden kaynaklanan kazalarda teknik bilirkişi raporlarının hazırlanamaması ve bu raporlar doğrultusunda iyileştirilmelerin yapılamaması sayılabilir. Ağır tonajlı araçların denetlenememesi sonucu yol aplatmasında meydana getirdikleri deformasyonların önlenememesi, kurumsal altyapıların olmaması, sorumluluk aşamasında hiçbir kurumun sorumluluğu üzerine almaması bu acı tabloyu tetiklemektedir.

İnsan ölümlerinin sebepleri olarak sıralandığında ülkemizde trafik kazaları 3. sırada yer almaktadır. Bu kadar acı sonuçlar karşısında maalesef ülkemizde trafik kazaları ile ilgili bir üst kurul veya iletişim sağlayacak bir yapı bulunmamaktadır.

Kaza analizleri yapılırken en fazla son 10 yıllık verilere ulaşılabilmekte, mevzuat gereği ve/veya kurumlar kendilerinden emin olmadıklarından bazı verilere de hiç ulaşamamaktadır. Trafik kazaları ile ilgili en sağlıklı çalışmalar Emniyet Genel Müdürlüğü, Trafik Eğitim ve Araştırma Daire Başkanlığı bünyesinde yürütülmektedir. Veri tabanları, trafik kaza tutanaklarının sayısallaştırılması ile oluşmaktadır. İsteğe bağlı en fazla üç değişkenli veriler elde edilebilmektedir. Yoldan kaynaklanan sorunlarla ilgili hiçbir bilgiye erişilememektedir.

1.5. Trafik Kaza Analizleri ve Tahmin Modelleri

1.5.1. Çok değişkenli zaman serisi uygulamaları

Çok değişkenli zaman serisi analizinin kullanıldığı yöntem VAR(p) vektörel otoregresif zaman serileri yöntemidir. Altı farklı model üzerinde bu yöntem uygulanmıştır. Veriler TÜİK'ten elde edilmiştir. Çok değişkenli uygulamaların ilk modeli, kaza sayısı bağımlı değişken olarak alınmış, bağımsız değişkenler ise yük-ton, km-taşıt, kamyon sayısı, tır sayısı, otobüs sayısı, karayollarında enerji tüketimi, yol-km uzunluğu, ağırlıklı ortalama günlük trafik (AOGT) değerleri alınmıştır.

Bütün modellerimiz için, zaman serilerinin varsayımlarından olan normal dağılıma uygunluk için değişkenler logaritmik dönüşüme tabii tutulmuşlardır. Birim kök analizlerinden ADF (Geliştirilmiş Dickey Fuller) istatistiği sonucu fark serileri alınarak durağan hale getirilmişlerdir.

Durağan seriler üzerinden Granger Nedensellik analizleri yapılmıştır. Değişkenler arasında Granger nedeni olan seriler modellere katılmışlardır. Akaike (AIC), Schwarz (SCW), nihai sonuç hatası (FPE), Hannan-Quin (HQ) istatistiklerine göre minimum olan model belirleme derecesi alınmış ve modellere VAR(p) olarak uygulanmıştır. Modellerin hepsi için model parametreleri ve sonuçlar, tartışma kısmında verilmiştir.

Çizelge 1.1. Çok değişkenli zaman serileri model dereceleri

Modeller	VAR(p)
Model-I	VAR(1)
Model-II	VAR(2)
Model-III	VAR(2)
Model-IV	VAR(1)
Model-V	VAR(1)

Etki-tepki analizi yapılarak bağımlı değişkene verilecek bir birimlik şokun diğer değişkenler üzerindeki tepkileri analiz edilmiştir. Kaza sayılarına 1. dönemde verilen şokların bağımsız değişkenlerde genellikle 4., 5. ve 6. dönemde etkilerini kaybettikleri görülmüştür. Bu fonksiyonun verilerimize daha sağlıklı bir şekilde uygulanabilmesi için verilerin yıllık değil de aylık veya haftalık olarak verilmesi, sonuçların analizini daha anlamlı hale getirecektir. Kaza sayıları yılın belli aylarında değişiklik gösterdiğinden mevsimsel bir dalgalanmaya sahiptirler. Yaz aylarında ve bayram tatillerinde kaza sayılarında bir artış gözlemlenmekte ve bu artışın etki-tepki fonksiyonuna yansımaları daha anlamlı olmaktadır. Bu analiz yöntemine ilişkin sonuçlar materyal ve yöntem kısmında grafiksel olarak verilmiştir.

Varyans ayrıştırması yapılarak bir değişkene ait hata varyansının diğer değişkenler tarafından açıklanma oranları tespit edilmiştir. Modellerin hepsinde bulunan bağımlı değişken olan kaza sayısını açıklayan bağımsız değişken birinci modelde otobüs sayısı, ikinci ve üçüncü modelde nüfus, olmuştur. Dördüncü ve beşinci modellerde de bağımlı değişken yaralı ve ölü sayısı olduğundan en fazla etkileyen değişken beklenildiği gibi kaza sayısı sonra da AOGT olmuştur.

Bütün VAR(p) modellerinin hata terimlerinde otokorelasyon olmadığı LM testi ile test edilmiştir. Aynı şekilde hata terimlerinin normal dağılıma sahip olup olmadıkları da Jarque Bera testi ile test edilerek normal dağılıma uygun oldukları anlaşılmıştır. Hata terimlerinin otokorelasyon fonksiyonu (ACF) ve kısmi otokorelasyon fonksiyonu (PACF) grafikleri çizilerek otokorelasyonun olup olmadığı gözlemlenmiştir.

Seçilen modellerin parametrelerinin tutarlılıkları analiz edilip en uygun model belirlendikten sonra uyum grafikleri ve serpilme grafikleri çizilerek yorumlandı. Modellere ilişkin VAR(p) yönteminin R^2 'si, OKH'sı ve AIC kriteri belirlenerek tartışma ve sonuçlar kısmında verildi.

1.5.2. Yapay sinir ağı uygulamaları

Yedi farklı model üzerinde YSA uygulaması yapılmıştır. Modelleme aşamasına geçmeden önce bağımsız değişkenler için korelasyon matrisleri hazırlanarak incelenmiş ve bağımsız değişkenler arasında ilk beş modelde yüksek korelasyon gözlemlenmiştir. Bu modeller için YSA-TBA uygulaması için Matlab’te “prepca” kodu yardımıyla eğitim seti içerisinde %2’den daha az katkısı olan elemanların elendiği geri kalan elemanlar kadar temel bileşen ile giriş tabakasındaki nöronlar oluşturulmuştur.

Çizelge 1.2. Çok değişkenli zaman serileri model dereceleri

Modeller	Temel Bileşen Sayısı	İnput Nöron Sayısı
Model-I	2	2
Model-II	2	2
Model-III	2	2
Model-IV	1	1
Model-V	1	1

Yine Matlab’te yazılan bir yazılım ile ağı ezberlemesini engellemek için minimum parametre ile 3 tabakalı ağlar oluşturuldu. Aktivasyon fonksiyonu “trainbr” olarak belirlendi. Gizli katmandaki nöron sayısı 1-2-3 olarak sınırlandırıldı. Maksimum iterasyon sayısı 1000 olarak belirlendi. Dördüncü ve beşinci modellerde çıkış tabakası iki nöronlu olarak (ölü sayısı-yaralı sayısı) alındı. Transfer fonksiyonu logsig-purelin-tansig olarak alınmıştır. Verilerimiz, transfer fonksiyonunu özelliklerine göre [0-1] arası veya [+1,-1] arası bir ölçeklemeye tabi tutulmuştur. Eğitim seti, verilerin %70’ini, test seti ise verilerin %30’unu oluşturulmuştur. Program çalıştırıldığında birden fazla ağ yapısı içerisinde minimum AIC kriterinin veren ağ ile minimum OKH veren ağ mimarisi arasında maksimum R^2 ’yi veren ağ mimarisi optimum olarak kabul edilmiştir. Bütün modellere ilişkin YSA uygulamalarının test değerlerinin R^2 ’si, OKH’sı ve AIC kriterleri belirlenerek tartışma ve sonuçlar kısmında verilmiştir.

1.5.3. Çok deęişkenli regresyon (ÇDR)

Modellerde kullanılan deęişkenlerin hepsi, logaritmik dönüşüm uygulanarak normal dağılıma uygun hale getirilmiştir. Bu modeller için bütün deęişkenleri kapsayan korelasyon matrisleri oluşturulmuştur. Tercan-Erzurum-Ağrı-Gürbulak E-80 karayolunun 422 km uzunluęundaki verileri kapsayan modeller hariç dięer 5 modeldeki bağımsız deęişkenler için yüksek korelasyon, yüksek R^2 , çok düşük tolerans çok yüksek VIF oranları çıkmıştır. Bütün bu göstergeler ilk beş modelde Çok deęişkenli regresyon uygulamasında çoklu bağlantı probleminin olduğunu göstermiştir.

Çok deęişkenli regresyonun varsayımlarından biri olan çoklu bağlantı probleminin kaldırılması için TBA uygulaması yapılmıştır. P kadar deęişken yerine p kadar temel bileşen bulunmuştur. Bu temel bileşenlerin toplam varyansı açıklama oranı, %95 ve üzeri olarak kabul edilmiş ve hemen hemen bütün modellerde iki temel bileşen sayısı bulunmuştur. Bu temel bileşen sayısının tespiti için yamaç eğimi grafikleri çizilmiş ve yorumlanmıştır.

Yeni bulunan temel bileşen skorları ile bağımlı deęişken, lineer regresyona tabii tutulmuş model parametreleri istatistiksel olarak anlamlı bulunmuştur. Fakat regresyona giren bağımsız deęişkenlerin skorları ile model oluşturmak yerine, orjinal verilerin kullanılması daha anlamlı olacağından, en küçük kareler kestiricileri yardımıyla özvektörler kullanılarak ters dönüşüm uygulanmış ve çoklu bağlantı probleminin olmadığı regresyon modelleri oluşturulmuştur.

Modellerin hepsinde tahmin ve orijinal bağımlı deęişkenin uyum grafikleri ve serpilme grafikleri çizilmiştir. Bütün modellere ilişkin ÇDR uygulamalarının R^2 'si, OKH'sı ve AIC kriterleri belirlenerek tartışma ve sonuçlar kısmında verilmiştir.

1.5.3. Poisson regresyonu ve negatif binomial regresyon uygulamaları

E-80 Karayolunun, Tercan-Erzurum-Ağrı-Gürbulak sınır kapısına kadar olan 422 km'lik yol uzunluğundaki kaza sayılarının Poisson veya Negatif Binom Regresyon ile modellemeleri yapılmıştır. Bu yöntem ile daha çok sayma verileri ve kesikli rastsal değişkenler kullanılmaktadır. Kaza sayıları da kesikli rastsal değişkenler olduğundan trafik kazaları için çok uygun bir yöntemdir. Literatürde de bu yöntemin kullanılması sonucu birçok çalışma yapıldığı görülmüştür. Veriler 2005–2008 yıllarını kapsamaktadır.

Bu güzergâh ilk önce Karayolları 12. Bölge Müdürlüğü tarafından kısımlara ayrılmış ve kesimlerin başlangıç km'leri tespit edilmiştir. Uzunluk diye tabir edilen bu değişkenler kesimlerin km'leridir. Diğer değişkenler YOGT, kesimlere ilişkin şerit sayısı, şerit genişliği, kavşak sayısı, ortalama hız ve bağlantı sayısı, olarak alınmıştır. Aynı yöntemde bağımlı değişken olarak bir modelde kaza sayısı, diğer modelde ise yaralı sayısı alınmıştır.

Modellerde bazı değişkenler istatistiksel olarak anlamlı çıkmamış ve modelden çıkarılmışlardır. Poisson regresyonu uygulanmış, aşırı yayılım göstermesi durumunda negatif binom regresyon uygulaması yapılmıştır. Uyum iyiliği testi yapılmış, araştırma bulguları ve tartışma kısmında belirtilmiştir. Yöntem, YSA ve çok değişkenli regresyon yöntemi ile karşılaştırılmıştır.

2. KURAMSAL TEMELLER

2.1. Zaman Serileri

Bir zaman serisi, o serinin kapsadığı zaman içindeki hareketini gözlemler. (Bozkurt 2007). Tek bir serinin zaman içindeki hareketleri inceleniyor ve gelecek tahmini yapılıyorsa, bu tip serilere ve analize tek değişkenli zaman serisi analizi denir. Birden fazla serinin zaman içinde hareketleri inceleniyorsa ve bu serilerin birbirleri ile olan ilişkileri de bu analiz içine giriyorsa, bu tip serilere de çok değişkenli zaman serileri denir. Zaman serileri, zaman içerisinde ara vermeden kaydedilebilen verilere sahip seriler (elektrik sinyalleri voltaj vb.) olan sürekli zaman serileri ve zaman içerisinde belli aralıklarla verileri kaydedilebilen seriler (kaza verileri, faiz oranları, döviz oranları vb.) olan kesikli zaman serileri olmak üzere iki türdür.

Zaman serileri analizine geçmeden önce bazı varsayımların kabul edilmesi gerekmektedir. Bir zaman serisinin normal dağılıma uygun ve durağan olması temel varsayımlardan biridir. Durağan olan seri ile gelecek tahmini yapmak sahte regresyonu önlemekte oluşan modellerin istatistiksel olarak anlamlı olmasını sağlamaktadır. Zaman serilerini oluşturan bazı bileşenler mevcuttur. Ayrıca zaman serileri; trend, mevsimsel bileşen, çevrimsel bileşen ve düzensiz bileşen olmak üzere dört ana bileşenden oluşur (Kadırlar 2005).

Zaman serilerinin trafik kaza analizinde kullanılması düşünülen modeli Box-Jenkins ve VAR modelleridir. Box-Jenkins yöntemi sırasıyla bir seriye uygun modeli belirleme, modelin tahmini, tanısal denetimi ve tahmin işlemlerini içermektedir. Serinin tamamıyla kendi geçmiş bilgisine dayanarak tahmin yapmaya yönelik bir metottur. Zaman serilerinin analizi için serinin durağan bir seri olması yani eğilimden arındırılması gerekmektedir.

Zaman serileri analizlerinde iki önemli fonksiyon vardır. Bunlar otokorelasyon fonksiyonu ve grafiği ACF(k) ile kısmi otokorelasyon fonksiyonu ve grafiği PACF(k) dir. Araştırmacılar, bu iki fonksiyonun grafiklerinden modelin derecesini ve durağanlığını test etmektedirler. Bunun yerine istatistiksel testlerin yapılarak yorumların yapılması daha uygun olur.

Otokorelasyon, bir değişkenin birden fazla gecikmeli dönemi arasındaki ilişki olarak tanımlanmaktadır. Tek bir değişkenin kendi gecikmeli değerleri arasındaki otokorelasyonu denklem 2.1. ile hesaplanabilir (Sevüktekin ve Nargileçekenler 2007).

$$r_k = r_{Y_t Y_{t-k}} = \frac{Cov(Y_t, Y_{t-k})}{S_{Y_t} S_{Y_{t-k}}} = \frac{\sum_{t=1+k}^T (Y_t - \bar{Y})(Y_{t-k} - \bar{Y})}{\sqrt{\sum_{t=1}^T (Y_t - \bar{Y})^2} \sqrt{\sum_{t=1}^T (Y_{t-k} - \bar{Y})^2}} \quad (2.1)$$

Otokorelasyon katsayısı, kovaryansların iki standart sapmasının çarpımları oranına eşittir.

$$ACF(k) = \frac{\sum_{t=1+k}^T (Y_t - \bar{Y})(Y_{t-k} - \bar{Y})}{\sum_{t=1}^T (Y_t - \bar{Y})^2} \quad (2.2)$$

Kısmi otokorelasyonlar ise diğer zaman gecikmelerinin etkileri ($t=1,2,\dots,k-1$) yok edildiğinde Y_t ve Y_{t-k} arasındaki birlikteliğin derecesini ölçmede kullanılır. Aynı serilerin 1 (Y_{t-1}) gecikmeli ve 2 (Y_{t-2}) gecikmeli serileri arasında da bir otokorelasyon olduğunu varsayalım. Bu durumda serinin kendisi (Y_t), 1 (Y_{t-1}) gecikmeli değeri ve iki (Y_{t-2}) gecikmeli değeri arasında bir otokorelasyon ilişkisi vardır. Serinin kendisi ile iki (Y_{t-2}) gecikmeli değeri arasındaki bir korelasyonu ölçmek için aradaki 1 (Y_{t-1}) gecikmeli değer etkisini bulma gereği ortaya çıkar. Bu işlem kısmi otokorelasyon olarak ortaya çıkar (Sevüktekin ve Nargileçekenler 2007).

k'nıncı dereceden kısmi otokorelasyon katsayısı ϕ_{kk} olarak ifade edilirse, $Y_{t-1}, \dots, Y_{t-k}$ 'ya karşı Y_t 'nin regresyon modeli oluşturulursa otokorelasyon katsayısı hesaplanabilir. (Sevüktekin ve Nargileçekenler 2007)

$$Y_t^* = \phi_{21}Y_{t-1}^* + \phi_{22}Y_{t-2}^* + e_t \quad (2.3)$$

Burada Y_t ile Y_{t-2} arasındaki kısmi otokorelasyon ϕ_{22} 'dir. Yukarıda da bahsedildiği gibi Y_{t-1} 'in etkisinin kaldırılması sonucunda Y_t ile Y_{t-2} arasındaki korelasyondur. Örnek büyüklüğü T ise gecikme sayısı belirlenirken $T/4$ sayıda gecikme PACF 'nin bulunmasında kullanılabilir. Her bir ACF(k) değeri %95 güvenirlilik için $\pm 1.96(1/\sqrt{n})$ değeri bulunarak yapılır. Eğer ACF(k) değeri güven aralığı sınırları dışında kalıyorsa otokorelasyon vardır. Kısmi korelasyon fonksiyonu gecikmeli değişkenler arasındaki ilişkiyi ifade eder.

$$\phi_{kk} = \frac{\rho_k - \sum_{j=1}^{k-1} \phi_{k-1,j} \rho_{k-j}}{1 - \sum_{j=1}^{k-1} \phi_{k-1,j} \rho_j} \quad k=3,4,5,\dots \quad (2.4)$$

$$\text{Burada; } \phi_{kj} = \phi_{k-1,j} - \phi_{kk} \phi_{k-1,k-j}, j=1,2,3,\dots,k-1 \text{ dir.} \quad (2.5)$$

2.1.1. Birim kök testleri

Bir serinin durağanlığını kontrol etmede kullanılan testlere birim kök testleri denir. Çalışmalarımızda birim kök testlerinden Genişletilmiş Dickey Fuller (ADF) Testi kullanılmıştır. Bu test Monte Carlo yöntemi kullanılarak 1981 yılında Dickey ve Fuller tarafından bulunmuştur. Dickey Fuller (1981), durağan olmama $H_0 : \rho = 1$ yokluk

hipotezinin bir deterministik trend etrafında durağan olma $H_1 : \rho < 1$ seçenек hipotezine karşı testi ele almıştır (Kadılar 2000).

$$y_t = \beta_0 + \beta_1 t + \rho y_{t-1} + \varepsilon_t \quad (2.6)$$

$$\Delta y_t = \beta_0 + \beta_1 t + \gamma y_{t-1} + \varepsilon_t \quad (2.7)$$

Burada $H_0 : \rho = 1$ yokluk hipotezi $H_0 : \gamma = 0$ hipotezine dönüşmektedir. Çünkü $\gamma = \rho - 1$ olmaktadır. γ 'nin t istatistiği, T_t şeklinde gösterilmektedir. Fark serisi alınmış serinin yine trende sahip olduğu görülüyor ise en az kısıtlı model olan (2.7)'nolu $(\beta_0 \neq 0, \beta_1 \neq 0)$ denklemden işe başlanmalıdır. Yokluk hipotezi red ediliyorsa işlemlere devam etmeye gerek yoktur. Eğer red edilmiyorsa trendin önemli olup olmadığının kontrolü gereklidir. Dickey-Fuller testi için geliştirilen denklem (2.8)

$$\Delta y_t = \beta_0 + \beta_1 t + \gamma_1 y_{t-1} + \sum_{i=1}^p \gamma_{2i} \Delta y_{t-i} + \varepsilon_t \quad (2.8)$$

Burada p, ε_t hata terimlerinin akgürültü olabilmesi için gerekli büyüklükteki bir değerdir (Kadılar 2000).

Zaman serileri analizlerinde yorumcular serinin durağan olması gerektiği varsayımı üzerinde durmuşlardır. Durağanlık kavramı serinin değerlerinin belli bir değere yaklaşmasını ya da beklenen değeri etrafında dalgalanmasını ifade eder. Durağanlık koşulları aşağıda verilmiştir (Bozkurt 2007).

$$- E(y_t) = \mu \rightarrow \text{ortalama}$$

$$- \text{Var}(y_t) = E(y_t - \mu)^2 = \sigma^2 \rightarrow \text{varyans}$$

$$- \gamma_k = E[(y_t - \mu)(y_{t-k} - \mu)] \rightarrow \text{kovaryans}$$

Tüm gecikmeler için ortalama, varyans ve kovaryans değerleri sabit olmalıdır. Bir başka ifade ile serinin durağanlığı, zaman içinde varyansın ve ortalamanın sabit olması ve gecikmeli iki zaman periyodundaki değişkenlerin ko-varyansının değişkenler arasındaki gecikmeye bağlı olmasıdır. Bu tip durağanlığa zayıf durağanlık denir. Güçlü durağanlık ise değişkenlerin koşullu ihtimal dağılımlarının zaman içinde değişmediği bir süreci gösterir (Enders 1995; Bozkurt 2007).

Durağanlığı tespit yöntemlerinden biri de korelogramlardan faydalanmaktır. ACF eğer çok yüksek bir değerden başlayıp çok yavaş küçülüyorsa, bu serinin durağan olmadığına bir göstergesidir.

Diğer bir ifade ile hesaplanan “t” değerinin mutlak değeri Dickey-Fuller veya McKinnon Dickey-Fuller kritik değerlerinin mutlak değerini aşıyorsa, zaman serisinin durağan olduğu hipotezini reddedemeyiz. ‘Ho: p=1’ reddedilirse zaman serisi durağandır (Şimşek ve Halaç 2003).

Dickey-Fuller’in ortaya koyduğu üç denklem türü bulunmaktadır;

Sabitsiz trendsiz Dickey-Fuller denklemi: $\Delta Y_t = \gamma Y_{(t-1)} + u_t$

Sabitli trendsiz Dickey-Fuller denklemi: $\Delta Y_t = a + \gamma Y_{(t-1)} + u_t$

Sabitli trendli Dickey-Fuller denklemi: $\Delta Y_t = a + b_t + \gamma Y_{(t-1)} + u_t$

Birim kökün varlığının denmesi için kullanılan iki hipotez kullanılmaktadır. Bunlar;

$H_1 : \gamma < 0$ ($p < 1$) (seride birim kök yoktur, seri durağandır.)

$H_0 : \gamma = 0$ ($p = 1$) (seride birim kök vardır, seri durağan değildir.)

2.1.2. Model seçim kriterleri

Yapılan analizlerde bulunan birçok model içinde en uygununun bulunabilmesi bazı kıstasların yerine gelmesine bağlıdır. Uygulamalarda birden çok uygun model tespit edilebilir. Bu uygun modeller içinde istatistiksel olarak en uygun olanını tespit etmek gerekir. Zaman serileri analizlerinde, model seçim kriteri en fazla kullanılan hata terimleri varyansına bağlı olan Akaike (AIC) bilgi kriteridir.

Model seçilirken gecikme sayısı, model derecesi olan p ve q ne kadar büyük olursa hataların karelerinin toplamı o kadar küçük olacaktır (Kutlar 2000).

$$AIC = N \ln(\sigma_e^2) + 2(p + q) \quad (2.9)$$

Burada N , ölçü sayısı, σ_e^2 hata terimlerinin varyansı, p ve q da model derecelerini temsil etmektedir (Salas *et. al.*). Aynı kıstasın bir başka gösterim şekli;

$$AIC = T \ln(\text{hata kareleri toplamı}) + 2n \quad (2.10)$$

$$SBC = T \ln(\text{hata kareleri toplamı}) + n \ln(T) \quad (2.11)$$

şeklindedir. İkinci kıstas ise Schwartz Bayesian kriteridir (SBC). Burada n , tahmin edilen parametre sayısıdır (Kutlar 2000).

AIC (2.9) Formülüyle hesaplanmaktadır. Modelde sabit terim olduğunda parametre sayısına 1 eklemek gerekmektedir. Akaike kriterine alternatif olarak geliştirilen Schwartz Bayes kriteri (SBC) değeri

$$SBC = T \ln \sigma_{\varepsilon}^2 + M \ln T \quad (2.12)$$

formülüyle hesaplanmaktadır. Bu iki formülün aralarındaki tek fark Akaike bilgi kriteri formülündeki 2 yerine $\ln T$ teriminin gelmesidir ki bu terim küçük örneklemeler için ($50 < t < 150$) 4 ile 5 arasında değer almaktadır. Bu fark Schwartz kriterinin Akaike bilgi kriterine göre daha az parametrelili modelleri seçme eğilimli olmasını sağlamaktadır. Bu açıdan bakıldığında Schwartz kriteri Akaike kriterine göre tercih edilebilmektedir (Kadılar 2005).

Uygun olan modeller içinde bu kriter değerlerinin en küçüğüne sahip olan model, seriye en uygun model olmaktadır. Sonuç olarak ACF ve PACF grafikleriyle belirlenen uygun modellerden incelenen seri için en doğru modeli bulabilmek amacıyla Akaike ya da Schwarz bilgi kriterleri kullanılmaktadır.

Model seçim kriterleri içinde üçüncü bir kriter de Nihai Öngörülen Hata Kriteri (Final Prediction Error-FPE) dir.

$$FPE = \frac{T + n}{T - n} \hat{\sigma}^2 \quad (2.13)$$

dir. Burada asıl olan, hata kareleri toplamının minimum olmasıdır.

Belirleme kriterlerinin minimum olduğu değeri veren model en uygun model olarak kabul görür. Fakat bu da yeterli değildir. Modelin tahmininden sonra aşağıda belirtilen maddelerin de yerine getirilmesi gerekmektedir (Hınıslioğlu ve Bayata 2009).

- Modelin tahminlerinin orjinal seriye uyumu görsel olarak görülmelidir.
- Modelin tahmin edilen değerlerinin alt ve üst güven aralıklarının içinde olduğu görülmelidir.
- Hata terimlerinin otokorelasyon ve kısmi otokorelasyon grafikleri çizilmeli serinin akgürültü bir seri olduğu görülmelidir.
- Hata terimlerinin normal dağılıma uygunluğu test edilmelidir.

- Orijinal seri ile tahmin serisi arasındaki serpilme grafiği çizilmeli ve uyum gözlemlenmelidir.
- Modelin karakteristik köklerinin birim çember içinde kalıp kalmadığı gözlemlenmelidir.

Başka bir deyişle, model seçim kriterleri model belirleme aşamasının son adımı olmakta, dolayısıyla tanısıl denetim aşaması model seçim kriterleri uygulandıktan sonra muhakkak yapılmalıdır (Kadılar 2005).

2.1.3. Box-Jenkins yöntemi

Box-Jenkins Modelleri mevsimsel ve mevsimsel olmayan, durağan ve durağan olmayan modeller olmak üzere dört farklı duruma ayrılmaktadır. Mevsimsel olmayan modeller $ARIMA(p,d,q)$ şeklinde gösterilmektedir. Burada p otoregresyon (AR) modelinin derecesi, d fark alma işlemi sayısı, q hareketli ortalama modelinin (MA) derecesi olmaktadır. Mevsimsel Box-Jenkins modelleri ise $ARIMA(p,d,q)(P,D,Q)_s$ olmaktadır. Burada ise P mevsimsel otoregresyon (SAR) modelinin derecesi, D mevsimsel fark alma sayısı, Q mevsimsel ortalama (SMA) modelinin derecesi ve S ise periyot olmaktadır. Bu modeller daha çok tek değişkenli zaman serileri analizinde kullanılmaktadır (Kadılar 2005).

Box-Jenkins 1970 yılında geliştirmiş olduğu bu yönteme kendi adını vermiştir. Bu yöntemde üç tip modelleme söz konusudur. Otoregresif süreç $[AR(p)]$, hareketli ortalamalar süreç $[MA(q)]$ süreci ve ikisinin birleşiminden oluşan Otoregresif Hareketli Ortalama (ARMA) sürecidir. Durağan olmayan bir seride fark alınması gerekiyor ise fark alma derecesi olan $d(I)$ sürece eklenir ve yöntem $ARIMA(p,d,q)$ haline döner. Yöntemin en önemli özelliği en uygun modelin belirlenmesini sağlamaktır. Yöntemin bazı aşamaları vardır. Bozkurt (2000)'a göre belirleme, tahmin ve modelin testi, aşamaları mevcuttur. Serilerin ilk önce durağanlıkları test edilir. Durağan olmayan serilerin fark serileri alınarak analiz tekrar yapılır. Birim Kök Testi dediğimiz bu analiz

ve fark serilerin ACF-PACF grafiklerinden yorumlamalar yapılır. AR(p), MA(q) veya fark alma sayısı kadar ARIMA(p,d,q) modelleri belirlenir ve modelin parametrelerinin tutarlılıkları analiz edilir.

Süreç belirlendikten sonra katsayıların tahmini yapılır. En küçük kareler yöntemi (EKK) veya en çok olabirlik (Maksimum Likelihood) yöntemi kullanılır. Modelin seriyi iyi açıklayıp açıklamadığını anlamak için Q istatistikleri otokorelasyon katsayılarının anlamlı olup olmadıkları ve hata terimlerinin normallik varsayımına uyup uymadıkları analiz edilmelidir.

Modelin uygun bir model olduğuna karar verilirse, ileriye dönük olarak tahmin amaçlı kullanılabilir. Chatfield and Prothero (1973), Bozkurt (2007)'ye göre uygun modele karar verdikten sonra eğer tahmin sonuçları tatmin edici değil ise farklı ARIMA(p,d,q) süreçleri denenmelidir. Bu alternatifler arasında minimum Q istatistiğine sahip yöntem seçilebilir.

2.1.3.a. Otoregresif (AR(p)) süreci

Otoregresif süreçte, seri kendi geçmiş değerlerinin ağırlık toplamı ile hata terimine bağlıdır (Tarı 2008). Bu tip sürece AR(p) süreci denilmektedir. Burada p 'nin ne olacağına karar vermek için değişkenin korelogram görünümüne bakmak veya minimum AIC ve SC değerlerine bakmak gerekmektedir.

$$y_t = \phi_1 y_{t-1} + \phi_2 y_{t-2} + \phi_3 y_{t-3} + \dots + \phi_p y_{t-p} + \varepsilon_t \quad (2.16)$$

$$(y_t - \delta) = \phi_1 (y_{t-1} - \delta) + \phi_2 (y_{t-2} - \delta) + \phi_3 (y_{t-3} - \delta) + \dots + \phi_p (y_{t-p} - \delta) + \varepsilon_t \quad (2.17)$$

AR(p) sürecinin açık hali (2.17) nolu denklemdir. Burada y_t ortalaması δ etrafında ve p gecikme ile ifade edilen bir sürece sahip olacaktır. PACF de anlamlı bir p, ACF de ise

ağır ağır azalan bir görüntü AR(1) konusunda bir fikir verebilir (Bozkurt 2007). Basit bir AR(p) süreci grafiksel olarak analiz edilmek istenirse ACF grafiğinin ağır ağır azalması, PACF grafiğinde ise ilk gecikmede ani bir azalma söz konusu ise AR(p) sürecinden bahsedilebilir.

2.1.3.b. Hareketli ortalama (MA(q)) süreci

Bir hareketli ortalama süreci, bağımlı değişkenin beyaz gürültü sürecine sahip geçmiş hata terimleri ile ifade edilmesidir. MA(q) modelinin açık biçimi aşağıda sunulmuştur.

$$z_t = \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \theta_3 \varepsilon_{t-3} - \dots - \theta_q \varepsilon_{t-q} \quad (2.18)$$

Bu eşitlikte sağ taraf anlamlı q sayısı kadar hata terimini ifade etmektedir. Hata terimi sıfır ortalamaya ve sabit varyansa sahiptir. MA sürecinde ACF ve PACF grafikleri analiz edilirken PACF’de geometrik bir azalma ACF’de ise anlamlı çıkışlar gözlemlenerek gecikme sayısı olan (q) belirlenir (Gujarati 2004; Bozkurt 2007).

2.1.3.c. Otoregresif bütünleşik hareketli ortalama (ARIMA(p,d,q)) süreci

ARIMA modelleri homojen ve durağan olmayan süreçlerdir. Durağan olmayan seriler ile çalışıldığından I(d) fark alma derecesi olarak kullanılır ve birinci fark ikinci fark d dereceden durağan oluyorsa da I(d) olur (Tarı 2008). d kadar fark alındıktan sonra durağan olan seriye d. dereceden fark durağan seri adı verilir. Daha önceki iki süreçte ACF ve PACF grafiklerine bakılarak model derecesi hakkında bilgi sahibi olunabiliyorken ARIMA modellerinde bu iki grafiğin geometrik olarak beraberce azalıyor olması gözlemlenebilir. Eğer seri fark almadan durağan halde ise süreç ARMA(p,q) yapısına dönüşür. ARIMA(p,d,q) modelleri zaman serileri ilgili üç unsuru da kapsamaktadır. Tarı (2008)’e göre buradaki asıl mesele p,d,q değerlerinin

bulunmasıdır. Bu durumda cimrilik prensibinin benimsendiği Box-Jenkins metodu uygulanmaktadır.

$$y_t = \alpha_0 + \sum_{i=1}^p \alpha_i y_{t-i} + \sum_{i=0}^q \beta_i e_{t-i} \quad (2.19)$$

$$z_t = \phi_1 z_{t-1} + \phi_2 z_{t-2} + \phi_3 z_{t-3} + \dots + \phi_p z_{t-p} + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \theta_3 \varepsilon_{t-3} - \dots - \theta_q \varepsilon_{t-q} \quad (2.20)$$

Modelde tüm karakteristik köklerin birim çember içinde kalması beklenir. ARIMA süreci dört kısımdan meydana gelmektedir.

- Belirleme: Durağanlık için fark serisi alınarak I(d) derecesinin tespitinin Birim Kök testi yapılarak analiz edilmesi, AR(p) ve MA(q) gecikmelerinin belirlenmesidir.
- Tahmin: Birinci aşamada belirlenen parametrelerin istatistiksel tutarlılığı ve ARIMA(p,d,q) modellerinin tespiti yapılır. Birden fazla uygun model belirlenir.
- Uygunluk Testi (Ayırt Edici Kontrol) : Hata terimlerinin normal dağılıma uygun olup olmadığını test etmek için Jarque-Bera testi ve Q testleri yapılır. Hata terimlerinin korelogramı incelenir.
- Öngörü: Tahmin edilen ARIMA(p,d,q) modeli ile gelecek tahmini yapılır.

Sürecin bu dört aşamasında fark serisi olarak durağanlaşan serimizin korelogramı yardımıyla AR(p) ve MA(q) ya da ARMA(p,q) süreçlerinin parametreleri p,q belirlenir. Köklerin mutlak değer olarak birim çember içinde kalması serinin durağan olduğunu göstermektedir. Eğer serinin birim kökü mutlak değer olarak 1'e eşit ise o zamanda bu tip serilere birim köklü seriler denir (Akdi 2003).

2.1.4. Çok değişkenli zaman serileri

Her bir değişkenin kendi gecikmeli değerleri ve sistemdeki diğer tüm değişkenlerin gecikmeli değerleri ile açıklanan çok değişkenli bir model olmaktadır. Kısa adı VAR

olarak ifade edilen vektör otregresyon modeli, Sims (1980) tarafından geliştirilmiştir. Sistemde VAR(1) yerine VAR(p) kadar bir gecikmeye sahip bir model geliştirilirse modeldeki parametre sayısı $(d^2p)+d$ tane bilinmeyen parametre olur (Kadılar 2000). X ve Y gibi basit iki değişkenin VAR modeli;

$$Y_t = \alpha_{10} + \sum_{i=1}^p \alpha_{11i} Y_{t-i} + \sum_{i=1}^p \alpha_{12i} X_{t-i} + u_{1t} \quad (2.21)$$

$$X_t = \alpha_{20} + \sum_{i=1}^p \alpha_{21i} Y_{t-i} + \sum_{i=1}^p \alpha_{22i} X_{t-i} + u_{2t} \quad (2.22)$$

Burada α_{i0} sabit terimi, α_{ijk} i'nci denklemdeki j'nci değişkenin k gecikmesine ait parametre, u_{it} hata terimi, p gecikme sayısıdır. Sabit terim, modele değişkenlerin sıfırdan farklı ortalamalara sahip olması durumunda dahil edilir (Tarı 2008). Modelin matrislerle ifadesi ise;

$$\begin{bmatrix} Y_t \\ X_t \end{bmatrix} = \begin{bmatrix} \alpha_{10} \\ \alpha_{20} \end{bmatrix} + \sum_{i=1}^p \begin{bmatrix} \alpha_{11i} & \alpha_{12i} \\ \alpha_{21i} & \alpha_{22i} \end{bmatrix} \begin{bmatrix} Y_{t-i} \\ X_{t-i} \end{bmatrix} + \begin{bmatrix} u_{1t} \\ u_{2t} \end{bmatrix} \quad (2.23)$$

Biçiminde veya kısaca,

$$y_t = c + \sum_{i=1}^p A_i y_{t-i} + u_t \quad (2.24)$$

şeklinde yazılabilir.

Bu model iki değişkenli olduğundan 2 boyutlu bir VAR modelidir. Daha genel bir ifade ile k sayıda değişken için,

$$y_t = c + A_1 y_{t-1} + A_2 y_{t-2} + \dots + A_p y_{t-p} + u_t \quad (2.25)$$

VAR(p) olarak ifade edilen modelde p gecikme değeri modelin derecesini de belirler, p'inci vektörel dereceden otoregresif model denir.

VAR modelleri, özellikle az sayıda veri ile çalışılacaksa ve nedensellik ilişkisi neticesinde karşılıklı ilişkiler saptanmış ise, uygundur. Ayrıca değişkenlerin sadece gecikmeli değerlerinin yer alması geleceğe yönelik tahminlerin başarısını artıracaktır (Kumar vd 1995; Bozkurt 2007).

VAR(p) modellerinin analizi yapılırken sisteme giren bütün değişkenlerin durağan olması bir zorunluluktur. Bu durumda durağan değişkenin şoku geçici bir etkiye sahip olacaktır. Başlangıçtaki bir şokun etkisi bir süre sonra kalkacaktır (Tarı 2008).

VAR modellerinin dezavantajları ise bütün değişkenlerin durağan olmasının beklenmesidir. VAR modellerinin oluşmasında değişkenlerin seçimi önemlidir. Çok fazla parametre içerdiklerinden serbestlik derecesi çok büyük olmaktadır (Tarı 2008).

2.1.4.a. Vektör otoregresif modellerinin belirlenmesi

Değişkenlerin seçimi, özelliklerinin belirlenmesi ve sıralaması önemlidir. Durağanlık koşulunun da sağlanması gerekmektedir. Gecikme uzunluğunun belirlenmesi ve öngörünün yapılması VAR modeli aşamalarını oluşturmaktadır.

VAR(p) modelleri oluşturulurken en önemli kıstaslardan biri modelin derecesinin belirlenmesidir. Eğer gereğinden küçük bir derece ile çalışılırsa model parametrelerinin tahmini tutarlı olmamakta ve tahminin varyansı gereğinden büyük çıkmaktadır (Shibata 1976, Kadılar 2000). Model derecesinin tespiti için en çok olabilirlik yönteminin geliştirilmiş bir şeklini Akaike (1974)'te önermiştir. Farklı parametrelere sahip farklı modeller içinde seçim yapmak için bütün modellerin en çok olabilirlik fonksiyonunu bulup bu fonksiyon değerinin en büyük olan modelin en uygun model olarak seçmek olduğu vurgulanmıştır (Schwarz 1978; Kadılar 2000).

Değişkenlerin sıralaması ve normal dağılıma uygunluğu test edilir. Birim kök testlerinden Genişletilmiş Dickey Fuller testi uygulanması tercih edilerek serilerin durağanlıkları analiz edilir ve durağan olmayan seriler durağanlaştırılır. Normal dağılıma uygun olan verilerimiz durağanlaştıktan sonra ortalama hatanın minimizasyonu ile nihai tahmin hatası FPE, veya AIC Akaike bilgi kriterinin minimum olduğu p değerini veren gecikme seviyesi belirlenir. Bir başka gecikme seviyesini belirleme kriteri de Hannan-Quinn (HQ) kriteri ve Schwarz (SC) bilgi kriteridir. Bu son iki kriter tutarlı gecikme seviyesinin belirlenmesinde daha etkindir.

Durağan olan değişkenler üzerinde Granger nedensellik analizi yapılarak değişkenler üzerinde birbirlerinin etkileşimleri analiz edilir. Gecikme sayısı belirlenen VAR(p) modelinin parametrelerinin uygunluğu için hata terimlerine, LM testi yapılır ve hata terimlerinde otokorelasyon olup olmadığı test edilir. LM olabilirlik değerinin %5'den büyük olması beklenir. Bu durumda H_0 hipotezi kabul edilir.

$$H_0 : P_1 = P_2 = P_3 = \dots = P_{12} = 0 \Rightarrow \text{otokorelasyon yok}$$

$$H_1 : P_1 \neq P_2 \neq P_3 \neq \dots \neq P_{12} \neq 0 \Rightarrow \text{otokorelasyon var}$$

Bu durumda modelin hata terimlerinde otokorelasyon yoktur denir. White testi uygulaması yapılarak modelde değişen varyanslık probleminin olup olmadığı kontrol edilir. Joint olabilirlik değeri %5'ten büyük ise modelde değişen varyans yoktur denir. Hata terimleri sabit varyanslıdır denir. H_0 hipotezi reddedilmez.

$$H_0 \Rightarrow \text{değişen varyans yok}$$

$$H_1 \Rightarrow \text{değişen varyans var}$$

Hata terimlerinin normal dağılıma uygun olup olmadığı ise uygulamalarımızda kullanılan Jarque Bera testi ile analiz edilir. Bu testte olabilirliklerin %5'ten büyük

olması beklenir. Aynı zamanda bu test χ^2 tablo değerin test istatistik değerlerinden küçük olması beklenir bu durumda H_0 hipotezi reddedilmez.

$H_0 : \Rightarrow$ hata terimleri normal değişen varyans yok

$H_1 : \Rightarrow$ değişen varyans var

Hataların otokorelasyon grafikleri çizilerek akgürültü seri halini aldıkları gözlemlenmelidir. Bütün bu aşamalardan sonra modelin uygun bir model olduğu kabul edilir. Değişkenler arasındaki Granger nedenselliği analizi, etki-tepki analizi ve varyans ayrıştırması, VAR(p) modellerinin önemli amaçları arasındadır.

2.1.4.b. Granger nedensellik analizi

Granger (1969) ve Sims (1972) tarafından geliştirilen bu analizde değişkenler arasında karşılıklı olarak bir nedenselliğin olabileceği düşünülmüştür. Bu nedenselliğin tek yönlü veya çift yönlü olabileceği daha doğrusu nedenselliğin yönünün gecikme sayısına bağlı olarak tespit edilebileceğini belirtmişlerdir. Bu nedensellik testine de “Granger Nedensellik Testi” adını vermişlerdir.

Granger nedensellik testi için serilerin durağan olması gerekmektedir. Nedensellik analizi aslında ekonomik veriler için daha çok önem arz etmektedir. Fakat bu çalışma da bir yöntem olarak kullanılmıştır.

2.1.4.c. Etki-Tepki fonksiyonu

VAR(p) modellerinde katsayıların yorumlanması oldukça güçtür. Bunun yerine bağımsız değişkenlere verilecek bir birimlik şokların bağımlı değişken ve diğer değişkenler üzerindeki etkileri analiz edilir. Sistemdeki değişkenlerin durağan olması nedeniyle verilen bir birimlik şokların (hata payı) diğer değişkenler üzerindeki etkileri

bir süre sonra sona erecektir. Eğer seriler durağan bir yapıda değil iseler şoklar devam edecek ve verilen tepki sağlıklı ölçülmeyecektir (Bozkurt 2007).

Etki tepki analizi, şokların etkilerinin ve bunlara verilen tepkilerin ölçüldüğü bir analiz olduğundan değişkenler arasında bir Granger Nedenselliğinin olması gerekir. Eğer bir değişken diğerinin nedeni değil ise birinin üzerine verilecek bir birimlik şok diğeri üzerinde bir etki yapmayacaktır.

2.1.4.d. Varyans ayrıştırması

Değişkenlerdeki değişimin, kendisi ve diğer değişkenler tarafından kaynaklanma oranlarını araştırır. Başka bir deyişle varyans ayrıştırması bir değişkene ilişkin öngörü hata varyansının diğer değişkenler tarafından açıklanma oranıdır. Eğer bir şok, bir değişkene ilişkin ileriye yönelik tahmin hata varyansını açıklayabiliyor ise, o değişken içsel olarak kabul edilir (Lütkepohl 1993; Bozkurt 2007).

Varyans ayrıştırması VAR modellerinde beklenen ikinci fonksiyondur. Varyans ayrıştırması ile serinin varyansındaki değişmeye şok süresince her bir değişkenin katkısı ölçülür.

2.2. Yapay Sinir Ağları (YSA)

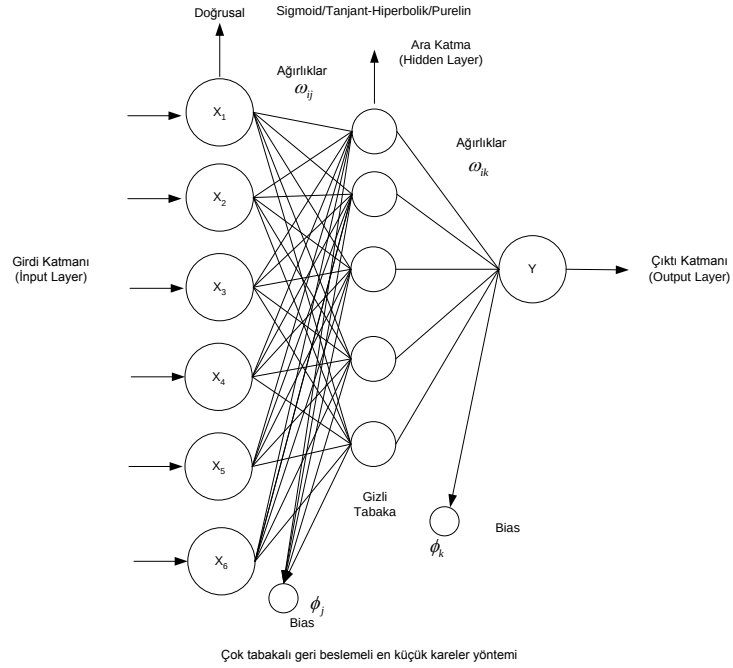
YSA, beynin bir işlevini yerine getirme yöntemini modellemek için tasarlanan bir sistem olarak tanımlanabilir. Bir YSA, yapay sinir hücrelerinin birbirleri ile çeşitli şekillerde bağlanmasından oluşur. YSA, öğrenme algoritmaları ile öğrenme sürecinden geçtikten sonra, bilgiyi toplama, hücreler arasındaki bağlantı ağırlıkları ile bu bilgiyi saklama ve genelleme yeteneğine sahip olurlar. YSA, yapılarına göre farklı öğrenme yaklaşımları kullanırlar. YSA, insan beyninin çalışma ve düşünebilme yeteneğinden yola çıkılarak oluşturulmuş bir bilgi işlem teknolojisidir. YSA işleyiş özelliklerine

dayanan ikinci tür tanımı ise ilk ticari yapay sinir ağının geliştiricisi olan Saraç (2004) tarafından NIELSEN'e ait bir atfında “Yapay sinir ağı, dışarıdan gelen girdilere dinamik olarak yanıt oluşturma yoluyla bilgi işleyen, birbiriyle bağlantılı basit elemanlardan oluşan bilgi işlem sistemidir. YSA paralel olarak bağlantılı ve çok sayıdaki basit elemanın, gerçek dünyanın nesneleriyle biyolojik sinir sisteminin benzeri yolla etkileşim kuran, hiyerarşik bir organizasyonudur” (Saraç 2004).

Biyolojik sinir ağlarında olduğu gibi yapay sinir ağlarında da temel unsur, yapay sinir hücresidir. Yapay sinir hücresi, YSA’ nın çalışmasına esas teşkil eden en küçük ve temel bilgi işleme birimidir. Ağ içinde yer alan tüm nöronlar bir veya birden fazla girdi alırlar ve tek bir çıktı verirler. Bu çıktı yapay sinir ağının dışına verilen çıktılar olabileceği gibi başka nöronlara girdi olarak da kullanılabilirler. Genel özellikleri ile bir yapay hücre modeli 5 bileşenden oluşmaktadır (Saraç 2004).

Bunlar;

- Girdiler
- Ağırlıklar
- Birleştirme Fonksiyonu
- Aktivasyon Fonksiyonu
- Çıktı



Şekil 2.1. YSA’da bir ağ mimarisi

Yapay sinir ağlarının özellikleri ve üstünlükleri aşağıda belirtildiği şekildedir.

- Doğrusal olmama
- Paralellik
- Gerçeklenme kolaylığı
- Yerel bilgi işleme
- Hata tolerans
- Öğrenebilirlik
- Genelleme
- Uyarlanabilirlik
- Donanım hızı
- Analiz ve tasarım kolaylığı

$$v = \sum x_i w_i + \theta, \quad y = F(v) \quad (2.26)$$

Burada;

w: Hücrenin ağırlıklar matrisi

x: Hücre giriş vektörü

v: Hücre net girişi

y: Hücre çıkışı'dır.

YSA'ların verilerin transfer fonksiyonunun özelliğine göre belli bir ölçeklemeye tabii tutulması gerekmektedir. Bunun için veri setinde bulunan minimum ve maksimum değerlerin bulunması ve aşağıda verilen ölçeklemelere tabii tutulması gerekmektedir. Bir Y veri setinin [0,1] aralığında ölçekleme için (2.27) No'lu eşitlikten [-1,1] aralığında ölçekleme için ise (2.28) No'lu eşitlikten yararlanılmalıdır. [a,b] gibi bir aralıkta ölçeklemek için de ölçekleme için (2.29) No'lu eşitlikten yararlanılmalıdır.

$$Y_{yeni} = \frac{Y - y_{\min}}{y_{maks} - y_{\min}} \quad (2.27)$$

$$Y_{yeni} = 2 \frac{Y - y_{\min}}{y_{maks} - y_{\min}} - 1 \quad (2.28)$$

$$Y_{yeni} = \frac{Y - y_{\min}}{y_{maks} - y_{\min}} (b - a) + a \quad (2.29)$$

Eğitim esnasında normalize edilmiş giriş ve çıkış değerleri kullanılır. İşlem sonunda ters dönüşüm yapılarak gerçek değerlere dönüşüm sağlanmalıdır. Öğrenme algoritmasını etkileyen en önemli unsurlardan biri de performans fonksiyonudur. En çok kullanılan fonksiyonlar aşağıda belirtilmişlerdir.

İleri beslemeli ağlarda ki performans fonksiyonu karesel ortalama hatadır. (MSE mean square error);

$$MSE = \frac{1}{N} \sum_{i=1}^N (t_i - td_i)^2 \quad (2.30)$$

Toplam karesel hata SSE (sum square error)

$$SSE = \sum_{i=1}^N (t_i - td_i)^2 \quad (2.31)$$

Karesel ortalama hata karekökü RMSE (root mean square error)

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (t_i - td_i)^2} \text{ ile ifade edilir.} \quad (2.32)$$

Performans fonksiyonuna ağırlık biaslarının ve ağırlıklarının karelerinin toplamının ortalama bir değerini ifade eden bir terim eklenerek genelleştirme iyileştirilebilir. Bu ağı daha hassas davranmaya zorlar ve ağırlık öğrenme kümesi haricindeki veriler içinde ezberleme yapmasını engeller (Sağıroğlu vd 2003).

Bazen ağa giren değişkenler arasındaki ilişki çok yüksek olabilir. Aynı zamanda giriş vektörünün boyutları da çok yüksek olabilir. Bu durumda giriş vektörünün boyutlarını azaltmak ve giriş vektöründeki değişkenler arasındaki korelasyonu kaldırmak için YSA’da da kullanılan temel bileşenler analizini uygulamak gerekir. Bu yöntem, giriş vektörlerini ilişkisiz hale getirir, ortogonal bir sıralama meydana getirir ve veri kümesi içinde en son türe uyan bileşenleri alır. Temel bileşenler analizi Matlab’de “Principal Component Analysis” kısaltılmış şekli ile ifade edilen “prepca” komutu ile oluşturulmaktadır (Sağıroğlu vd 2003).

2.2.1. Aktivasyon ve transfer fonksiyonu

Aktivasyon veya transfer fonksiyonları bir YSA'da nöronun çıkış genliğini, istenilen değerler arasında sınırlar. Bu değerler çoğunlukla $[0,1]$ veya $[-1,+1]$ aralığındadır. YSA'larda kullanılan fonksiyonların türevi alınabilir olmalı ve süreklilik arz etmeli. Uygulamalarda daha çok tanjant hiperbolik veya sigmoid fonksiyonu kullanılmaktadır. Lineer ve doğrusal olmayan transfer fonksiyonlarının kullanımı YSA'ların çok karmaşık problemlere uygulanmasını sağlamıştır (Sağıroğlu vd 2003).

Doğrusal fonksiyon, YSA'ların çıkış katmanında kullanılır. Basamak fonksiyonu ise perseptron olarak bilinen işlemci eleman bu fonksiyon ile işlem görür. Kutupsamalı basamak fonksiyonu, parçalı doğrusal fonksiyon aktivasyon fonksiyonlarından bazılarıdır (Sağıroğlu vd 2003).

Sigmoid fonksiyonu YSA uygulamalarında en çok kullanılan aktivasyon fonksiyonudur (2.33). Bu fonksiyonun en aktif bölgesinin 0.2 ile 0.8 arasında olduğu bilinmektedir.

$$y = \frac{1}{1 + e^{-v}} = \frac{1}{2}(\tanh(v/2) + 1) \quad (2.33)$$

Tanjant Hiperbolik fonksiyonu ise YSA'larda en çok kullanılan ikinci fonksiyondur. Giriş uzayının genişletilmesinde etkili bir fonksiyondur (Sağıroğlu vd 2003).

$$y = \frac{1 - e^{-2v}}{1 + e^{-2v}} = \tanh(\beta v) \quad (2.34)$$

2.2.2. Yapay sinir ağlarında öğrenme algoritmaları

Öğrenme algoritması bir YSA ana çatısını oluşturur. Birçok öğrenme algoritması vardır ama en çok kullanılan trainlm ve trainbr dir. Bu iki algoritma Levenberg-Marguardt algoritmasıdır.

Trainlm, öğrenme kümesi çok büyük ise hafıza ihtiyacını alzan bir parametreye sahip Levenberg-Marguardt algoritmasıdır. Modern öğrenme algoritmalarının en hızlısıdır. Trainbr ise bayesian düzenleyici öğrenme metodur. Levenberg-Marguardt algoritmasının, genelleme yeteneğini (ezberleme) iyileştirmek için geliştirilmiş bir öğrenme algoritmasıdır. Optimum ağ yapısının ne kadar olması gerektiği problemini kısmen azaltmaktadır (Sağıroğlu vd 2003).

2.3. Çok Değişkenli Regresyon Analizi

Bağımlı değişken ile bağımsız değişkenler arasındaki ilişkiyi matematiksel bir modelle istatistik temellere oturtturularak açıklamak, bağımsız değişkenler yardımıyla bağımlı değişken için kestirimde bulunabilmek çok değişkenli istatistik analizin amaçlarından biridir (Alpar 2003).

Bayazıt (2006)'ya göre bir rastgele değişken ile birden fazla rastgele değişkenler arasında istatistikî ilişkinin belirlenmesi amaçlanmaktadır. Söz konusu rastgele değişkenin varyansının olabildiğince büyük kısmını açıklayacak şekilde belirlenecek bu ilişki (regresyon denklemi) ile söz konusu değişkenin bir olayda alacağı değer, bilinen diğer rastgele değişkenlere bağlı olarak tahmin edilecek ve tahminlerin güven aralıkları belirlenecektir.

Çok değişkenli regresyon analizinde değişkenler arasındaki ilişkinin doğrusal olduğu kabul edilmektedir. Tahmin edilen değerlerle bilinmeyen gerçek değerler arasındaki

farklar olan kalıntı terimlerinin bağımsız, varyansı sabit ve normal dağılmış oldukları da yapılan kabuller arasındadır. Son kabuller istatistik hipotezlerin kontrolünde ve güven aralıklarının belirlenmesinde gereklidir (Bayazıt 2006).

Bağımlı değişkeni bağımsız değişkenlerden hangisinin daha çok etkilediğini tespit etmek, çok değişkenli istatistiğin amaçlarındandır. Birden fazla bağımsız değişkenle açıklanabilen model içinde, bağımsız değişkenlerin sayılarını azaltarak az sayıda parametre ile bağımsız değişkeni açıklamakda başka bir amacdır.

Y bağımlı değişkeni ile X_j ($j=1,2,\dots,k$) bağımsız değişkenleri arasındaki doğrusal istatistiki ilişki

$$y = b_0 + b_1x_1 + b_2x_2 + \dots + b_kx_k + \varepsilon \quad (2.35)$$

denklemini ifade edilir. Bu denklemde b_j ($j=0,1,2,\dots,k$) regresyon katsayıları, ε kalıntı (hata) terimidir. Denklemde parametre sayısı (regresyon katsayılarının sayısı) $p=k+1$ dir. Gözlemlerin sayısı n ise serbestlik derecesi $s.d. = n-p$ olur.

$$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} 1 & x_{11} & x_{12} & \dots & x_{1k} \\ 1 & x_{21} & x_{22} & \dots & x_{2k} \\ \dots & \dots & \dots & \dots & \dots \\ 1 & x_{n1} & x_{n2} & \dots & x_{nk} \end{bmatrix} \begin{bmatrix} b_0 \\ b_1 \\ \vdots \\ b_k \end{bmatrix} + \begin{bmatrix} e_1 \\ e_2 \\ \vdots \\ e_n \end{bmatrix} \quad (2.36)$$

Y değişkeninin alacağı değer, X_j değişkenlerinin bilinen x_{ij} değerlerine bağlı olarak tahmin edilebilir (Bayazıt 2006).

2.3.1. Çok değişkenli regresyon analizi varsayımları

Çoklu doğrusal regresyon modellerinde varsayımlar;

- Normal dağılım

- Doğrusallık
- Hata terimlerinde otokorelasyonun olmaması
- Sabit varyans
- Bağımsız değişkenler arasında çoklu bağlantı (kolinearite) olmaması

olarak sınıflandırılabilir.

Normal dağılıma uygunluk için istatistiksel testler ve grafiksel görüntüler incelenebilir. İstatistiksel testlerde Ki-kare uygunluk testi, Kolmogorov-Smirnov testi (K-S), Z testi ve Shapiro-Wilks (W istatistiği) sayılabilir. Uygulamalarımızda genellikle Kolmogorov-Smirnov testi uygulanmış ve bu testte olabirlik değeri %5 anlamlılık düzeyinden büyük ise H_0 hipotezi reddedilmiştir. Grafik testler ise histogram, gövde–yaprak, kutu, Q-Q ve P-P grafikleridir.

Doğrusal olmayan etkiler için hesaplanacak doğrusal korelasyonlar gerçek ilişkiyi her zaman düşük gösterebilir. İki değişken arasında doğrusallığı sağlamada birçok dönüşüm uygulanır. Bazı dönüşümler logaritmik dönüşüm ($\log(x)$), karekök dönüşümü ($\sqrt{x}$), hiperbolik dönüşüm ($\frac{1}{x}$), kare dönüşümü (x^2), arcsin dönüşümü ($\arcsin \sqrt{x} = \sin^{-1} \sqrt{x}$), logit dönüşümü $\left[\text{Logit}(p) = 0,5 \log \left(\frac{p}{1-p} \right) \right]$, fisher z dönüşümü $\left[\text{FisherZ} = 0,5 \log \left(\frac{1+r}{1-r} \right) \right]$ olarak sayılmaktadır; (Albayrak 2006)

Hata terimlerinin bağımsızlığı, hata terimlerinin normal dağılıma uygun, geçmiş değerleri ile ilişkisi olmayan ak gürültü bir seri olduğudur. Bunun için de Durbin-Watson testine başvurulur. Aynı zamanda hata terimlerinin otokorelasyon grafikleri çizilerek gecikmelerin güven aralıklarını geçip geçmediği görülebilir.

$$DW = \frac{\sum_{t=2}^n (e_t - e_{t-1})^2}{\sum_{t=1}^n (e_t)^2} \cong 2(1 - \rho) \quad (2.39)$$

Bu amaçla DW hesaplanır ve genellikle 1.5–2.5 civarında olan DW test sonucu otokorelasyon olmadığını gösterir.

Tek değişkenli analiz durumunda kovaryans matrisi sabit bir sayı ve bağımlı değişkenin varyansı bütün hücreler için eşit olduğu zaman sabit varyanstan söz edilebilir (Albayrak 2006).

Çoklu bağlantı problemi (kolinearite) bağımsız değişkenler arasındaki yüksek korelasyondur. Bir modelde çoklu bağlantı olması durumunda modelin R^2 'si çok yüksek çıkmasına rağmen modele giren bağımsız değişkenlerin katsayılarının istatistiksel anlamlılıkları uygun olmamaktadır. Çoklu bağlantı olması durumunda bağımsız değişkenlerin işaretleri ters işaretli olabilmektedir. Çoklu bağlantı halinde modele katılan değişkenlerin varyans ve kovaryansları beklenmedik şekilde artabilir (Gujarati 1995; Kalaycı 2006).

Çoklu bağlantı problemini belirlemenin birkaç yolu vardır. Bağımsız değişkenlerin korelasyon matrisleri incelendiğinde yüksek bir ilişki söz konusu ise çoklu bağlantının varlığı söz konusu olabilir. Bu yüksek ilişkiye rağmen değişkenlerin katsayılarının olabilirliklerini test eden “t” testi sonuçları, istatistiksel olarak anlamlı çıkmayabilir. Bir diğer parametre olan varyans artış faktörü denilen VIF (Variance Inflation Factor) değerinin çok yüksek toleranslarında çok düşük ise modelde çoklu bağlantı probleminin olduğu anlaşılır (Kalaycı 2006). Çoklu bağlantı probleminin diğer bir tespit yöntemi de tolerans değeridir. Tolerans değeri $T = 1 - R_i^2$ şeklinde hesaplanır. Daha küçük tolerans daha büyük VIF demektir (Gujarati 1995).

Bu durumda diğer bağımsız değişkenlerle korelasyonu yüksek olan değişkenler modelden çıkarılmalıdır. Bunu yaparken fazla bilgi kaybedilmemesine dikkat edilmelidir. Aralarındaki korelasyon katsayısı r olan iki bağımsız değişkenden biri modelden çıkarılırsa bilgi kaybı oranı $1-r^2$ olur (Bayazıt 2006).

Çoklu bağlantı problemini ortadan kaldırmak için temel bileşenler analizi uygulaması yapmak gerekir.

2.3.2. Çok değişkenli regresyon analizinde model belirleme kriterleri

Belirlilik katsayısı (R^2) hemen hemen bütün istatistik modellerinde bir belirleyici kriterdir. Çoklu belirlilik katsayısı olarak da isimlendirilen R^2 , (multiple coefficient of determination) birden çok bağımsız değişkenin bağımlı değişkeni açıklama oranını verir. 0-1 arasında değişen R^2 , 0'a yaklaşırsa verinin model verilere uygun değil modelin verileri açıklamadığı anlaşılır. R^2 değerini analiz ederken dikkatli olunmalıdır. Çünkü bağımsız bir değişkenin eklenmesi bile R^2 'de artış meydana getirebilir. Düzeltilmiş R^2 ise bir bağımsız değişkenin modele eklenmesiyle, artarak değişen R^2 belli bir miktar azaltılarak düzeltilir ve düzeltilmiş $\bar{R}^2$ elde edilir. Bir modeli analiz ederken R ile $\bar{R}^2$ arasındaki fark belirli bir dereceden fazla ise $\bar{R}^2$ değerine bakılması gerekir (Alpar 2003).

$$R^2 = \frac{\sum_{i=1}^n (\hat{y}_i - \bar{y})^2}{\sum_{i=1}^n (y_i - \bar{y})^2} = \frac{RKT}{KT_Y} = \frac{\hat{\beta}' X' Y - n \bar{y}^2}{Y' Y - n \bar{y}^2} \quad (2.38)$$

$$\bar{R}^2 = \left(R^2 - \frac{p}{n-1} \right) \left(\frac{n-1}{n-p-1} \right) \quad (2.39)$$

Model belirleme kriterlerinden biri de (2.32) No'lu denklemde bahsedilen ortalama karesel hatanın (OKH “MSE mean squer error”) minimum olduğu durumdur. AIC Akaike bilgi kriteri de yine belirleyici bir kriterdir. Modelin tamamının istatistiksel olarak uygunluğunu test etmek için F testi uygulanır. $F_{hesap} > F_{tablo}$ ve olabilirlikleri % 5’ ten küçük ise modelin tamamının uygun olduğu kabul edilir. Aynı şekilde model içinde bulunan parametrelerin yani bağımsız değişkenlerin ve sabit terimin katsayılarının istatistiksel olarak anlamlılıkları, t testi ile analiz edilir ve yine olabilirliklerin % 5’ten küçük olması beklenir.

Ayrıca; tahmin serisi ile tarihi serinin %95 güven aralıkları ile grafikleri çizilerek uyum gözlenir. Aynı şekilde bağımsız değişkenin tarihi değerleri ile tahmin değerleri arasında serpilme (saçılma) grafikleri çizilir. İki seri arasındaki uyum R^2 ile gözlemlenir.

2.3.3. Temel bileşen analizi (TBA)

Çoklu bağlantı probleminin olması sahte regresyonu ortaya çıkarmakta modelden elde edilen tahminlerin yanıltıcı olmasını sağlamaktadır. Çoklu bağlantı problemini çözmenin en iyi yolu temel bileşenler analizi uygulamaktır. Temel bileşen analizi kısaca, bağımlı değişkeni, birbirleriyle çok ilişkili bütün bağımsız değişkenler yerine birbirleriyle ilişkisiz daha az sayıda temel bileşen ile açıklamaktır. Bu analize geçmeden önce temel bileşenler analizinin uygun bir yöntem olup olmadığını analiz etmek için küresellik testi denilen bir test uygulanır. Bu teste Barlett testi de denir.

$H_0 : R = I$ Değişkenler arasındaki ilişki önemsizdir.

$H_1 : R \neq I$ Değişkenler arasındaki ilişki önemlidir.

Olabilirlik $P < 0.05$ ise H_0 reddedilir yani küresellik varsayımı kabul edilmez. Değişkenler arasındaki ilişkiler önemlidir. Bu yüzden TBA (temel bileşenler analizi) uygulanabilir.

Temel bileşen (principal component) aralarında istatistiksel olarak ilişki bulunan P adet X_j ($j = 1, 2, 3, 4, \dots, p$) değişkenin aralarında istatistiksel olarak ilişki bulunmayan Z_j kadar temel bileşenle açıklanması olayıdır (Bayazıt 2006).

$$Z_j = \sum_{k=1}^p a_{kj} X_k \quad j = 1, 2, 3, \dots, q \quad q \leq p \quad (2.42)$$

X_j değişkenlerinin S kovaryans matrisinin yapısını açıklayarak anlamlı temel bileşen sayısı $q < p$ ise veri sayısını azaltmak temel bileşenlerin amaçlarındandır.

P boyutlu uzaydaki toplam varyans (özdeğerler toplamı), her biri özvektörlerle tanımlanan öyle yeni değişkenlerle ifade edilir ki, en küçük varyans sonuncu özvektöre ait olur ($\lambda_1 > \lambda_2 > \lambda_3 > \dots > \lambda_p$). Bağımsız değişkenler kümesindeki toplam değişimin büyük bir bölümü birinci özvektör, ondan daha az bir bölümü ikinci öz vektör tarafından açıklanır (Alpar 2003).

Temel bileşen sayısını belirlemede bazı kıstaslar mevcuttur. Standart değişkenlerle belirlenen λ_j 'lerden 1'den (ya da 0.7'den) küçük olanlara karşılık gelen temel bileşenler gözönüne alınmayabilir. Bu kritere Kaiser kriteri denir. Cettell döküntü (scree) testi ise λ_j değerlerinin j'ye göre değişimini gösteren çizginin yatıklaşmaya başladığı j değeri q olarak alınır. Bu grafik aynı zamanda scree plot denilen bir grafik yardımıyla analiz edilir. $S_j = \lambda_j + \lambda_{j+1} + \dots + \lambda_p$ hesaplandıktan sonra λ_j / S_j daha önce temel bileşenlerin açıklamadığı varyansın Z_j tarafından açıklanan oranını gösterir. Bu oranın 0.1'den küçük olduğu Z_j ve bundan sonraki temel bileşenler göz önüne alınmayabilir (Bayazıt 2006).

Temel bileşenler analizinde her yeni değişken, orjinal değişkenlerin doğrusal bir bileşimidir. Birinci yeni değişken (temel bileşen) verilerdeki maksimum varyansı

açıklayıcı bir şekildedir. İkinci temel bileşen geriye kalan toplam varyansa maksimum katkıda bulunmaktadır. Maksimum sayıda üretilebilecek temel bileşen sayısı orjinal değişken sayısı kadardır. Yeni bileşenler bağımsız üretilirler (Albayrak 2006).

Veri matrisinde yer alan p değişkeninin doğrusal bileşenlerini bulmak için kovaryans matrisinin ya da korelasyon matrisinin özdeğerleri ve özvektörleri kullanılır. Eğer değişkenler aynı birim veya karşılaştırılabilir birimlerdeyse ve değişken varyansları aynı boyutta ise varyans-kovaryans matrisi kullanılır. Bu durumlar sağlanmadığında varyans-kovaryans matrisi yerine korelasyon matrisi kullanılır (Kocabaş 2008).

TBA'nın 3 temel amacı vardır:

1. Verilerin boyutunu azaltmak,
2. Tahminleme yapmak,
3. Veri setini, bazı analizler için görüntülemek.

TBA uyguladığımızda sürecin sonunda bu p boyutlu uzayın gerçek boyutu belirlenir. Bu gerçek boyuta temel bileşenler adı verilir.

TBA, sonuç olarak p sayıda bağımsız değişken yerine p sayıda temel bileşenle bağımlı değişkeni açıklamaktadır. Bu temel bileşenler sırasıyla toplam varyansa maksimum katkıda bulunurlar. Yani, birinci temel bileşen en çok, diğer bileşenler ise gittikçe azalan miktarda toplam varyansa katkıda bulunurlar. Bu nedenle az sayıdaki temel bileşen ile toplam varyansın büyük bir çoğunluğu açıklanmış olur. Temel bileşen analizi aslında bir boyut indirgeme tekniğidir. Temel bileşenler analizinde kullanılan veriler ortalamaya göre düzeltilmiş veya standartlaştırılmış verilerdir. Ortalamaya göre düzeltilmiş veriler kullanılırsa temel bileşen ağırlıkları değişkenlerin varyanslarından etkilenmektedirler. Varyansı büyük olan değişkenlere daha büyük ağırlıklar verilebilmektedir. Standart verilerle çalışılması durumunda değişkenlerin ağırlıkları varyanslarından etkilenmemektedirler (Albayrak 2006).

2.4. Poisson Regresyonu (PR)

Bağımlı değişkenin sayma verisi aldığı (0,1,2,3,4.....gibi) kesikli değer aldığı kategorik olmadığı birkaç durum vardır. Bu tür değişkenlere kaza sayıları, yangın sayısı patent sayısı gibi örnekler verilebilir. Kesikli ve kategorik olmayan olaylarla ilişkili bağımlı değişkenli model bazı varsayımlar altında Poisson regresyon modeli olarak adlandırılır. Poisson regresyon modeli daha çok sayma verilerini analiz etmek için kullanılır (Maddala 1983; Akın 2002).

$$P(Y_i = y_i | X_{i1}, X_{i2}, X_{i3}, \dots, X_{iq}) = \frac{e^{-\lambda_i} \lambda_i^{y_i}}{y_i!}, y_i = 0, 1, 2, \dots \quad (2.41)$$

$$\lambda_i = e^{\beta' x_i}$$

Greene (1993)'e göre verilen Poisson regresyon modelini kaza sayılarına göre yorumlamak gerekirse;

$P(Y_i = y_i | X_{i1}, X_{i2}, X_{i3}, \dots, X_{iq})$: Kazaların olma olasılığı

y_i : Mevcut t zaman dilimi içinde olan kaza sayısı

λ_i : t zaman dilimi içindeki beklenen kaza sayısı

λ_i , logaritmik doğrusal olarak bağımsız değişkenlere bağlıdır (Akın 2002).

$$\ln \lambda_i = \beta' x_i = \beta_0 + \sum_{j=1}^p \beta_j x_{ij} \quad (2.42)$$

$$\lambda_i = \exp(\beta' x_i = \beta_0 + \sum_{j=1}^p \beta_j x_{ij}) \quad (2.43)$$

Poisson regresyonu doğrusal olmayan bir regresyondur ve parametrelerin katsayıları maksimum olabilirlik metodu ile kolayca bulunabilir. Olabilirlik fonksiyonu 2.44 No'lu denklemde ki gibi ifade edilir.

$$L(\lambda_i) = \prod_{i=1}^n \frac{\lambda_i^{y_i} e^{-\lambda_i}}{y_i!} \quad (2.44)$$

Parametreler, olabilirliğin maksimizasyonu veya daha genel olarak, olabilirliğin logaritmasının (log-olabilirlik diye belirtilir) maksimizasyonu ile tahmin edilir. Buna eşdeğer olarak, tahmin, log-olabilirliğin negatifinin minimize edilmesiyle de bulunabilir (Kibar 2008).

Kazalar kesikli rastgele olaylar olduğundan bunların normal dağılım kullanılarak modellenmesi yapılamaz (Al-Ghirbal and Al-Ghamdi 2006). Poisson dağılımı, aralıklı olan ve negatif olmayan rasgele, kesikli olaylar olan kazaların ifade edilmesi için uygun bir metottur (Hedayeghi 2002). İlk defa Fransız matematikçi Simeon Denis Poisson (1781–1840) tarafından incelenmeye başlanan bu dağılım, daha sonra Bartkrewicz tarafından ender görülen olaylara uygulanmış ve çok uygun sonuçlar elde edilmiştir (Çömlekçi 1989; Filiz 1995; Kibar 2008).

Poisson regresyonunda karşılaşılan sıkıntılardan biri de aşırı yayılımdır. Aşırı yayılım sonucunda tahmin edilen parametrelere ilişkin katsayıların varyansı çok küçük değerlere eğilim gösterir ve tahmin edilen katsayılar yanlı olurlar. Bu problem kaza verilerinin analizinde Negatif Binom dağılımlı olduğu kabul edilerek çözülebilir (Hedayeghi 2002; Kibar 2008)

Aşırı yayılımın iki sebebi olabilir bunlardan biri kazaların kesikli rassal değişken olmaları münasebetiyle veri setindeki çok fazla sıfır olmasıdır. Diğer ise trafik hacminin, doğru olmaması, bilgi eksikliği, hava koşullarında ölçsüz varyans, görsellik ve sürücü davranışları gibi etkilerin unutulmuş olmasıdır (Hauer 2001; Shankar 1997; Qin vd. 2004; Lord vd. 2005; Kibar 2008).

3. MATERYAL ve YÖNTEM

Kaza analizleri için çeşitli yöntemler vardır. Giriş bölümünde bu yöntemlerle yapılmış araştırmalara yer verilmiştir. Kullanılan bazı yöntemler aşağıdaki gibidir.

- 1) Çok değişkenli regresyon yöntemi
- 2) Zaman serisi yöntemi
- 3) Yapay sinir ağları yöntemi
- 4) Poisson regresyonu

Çok değişkenli regresyon, çok değişkenli zaman serileri analizi, yapay sinir ağları yöntemi, poisson regresyonu yöntemleri aynı verilere uygulanmış sonuçlar ortalama karesel hata (OKH), belirlilik katsayısı (R^2) ve Akaike (AIC) bilgi kriterine göre değerlendirilmiştir.

Uygulamalarda kullanılan veriler Emniyet Genel Müdürlüğü, Trafik Eğitim ve Araştırma Daire Başkanlığından, yol ve sürücü karakteristiklerine ilişkin detay nitelik veriler ücreti karşılığında, diğer veriler ise yıllık yayınlanan Trafik İstatistik Yıllıklarından (1996-2002) elde edilmiştir. 2003-2010 yılları arasını kapsayan veriler ise Türkiye İstatistik Kurumu'nun yayınladığı Trafik Kaza İstatistikleri Karayolu, Motorlu Kara Taşıtları İstatistikleri, Ulaştırma İstatistikleri Özeti, kitaplarında elde edilmiştir.

Çalışmalarımızda kullanılan bağımsız değişkenler arasında lineer bir ilişki olmamasından dolayı, lineer ilişki olmayan değişkenlerin modellenmesinde kullanılan ve en iyi sonucu veren YSA uygulamalarımızda tercih edilmiştir.

YSA Uygulamasında Matlab-2008'de bir program yazıldı. Program transfer fonksiyonları (tansig-purelin-logsig), eğitim fonksiyonları (trainbr-trainlm), gizli tabakadaki nöron sayısı (1,2,3,4.....), arasında bir döngü yapabilen bir yazılım olarak

geliştirildi. Yazılımın diğer özelliği ise ağın performans belirleme kistası olarak ortalama karesel hatayı alması, iterasyon sayısını isteğe bağlı olarak değiştirebilmesi ve ağın eğitiminin istenilen hassasiyette sonlandırabilmesidir. Alternatif ağ yapılarını, girdi tabakasındaki nöron sayısı, gizli tabakadaki nöron sayısı, çıktı tabakasındaki nöron sayısı, eğitim fonksiyonu, tabakalardaki transfer fonksiyonları, R^2 , OKH, AIC, değerlerinin excell ortamında okunabilmesi sağlanmıştır.

3.1. Çok Değişkenli Regresyon Analizi, Çok Değişkenli Zaman Serisi Analizi, ARIMA ve Yapay Sinir Ağı Modellemeleri

3.1.1. Model-I

Yıllık kaza sayılarının; yük-ton değeri, yolcu-km, km-taşıt, kamyon sayısı, tır sayısı, otobüs sayısı, karayollarındaki enerji tüketimi, yol-km uzunluğu, AOGT değerlerine göre modellenmesi;

Y: Yıllık Kaza Sayıları

X_1 : Yük-Ton Değeri

X_2 : Yolcu-Km

X_3 : Km-Taşıt

X_4 : Kamyon sayısı

X_5 : Tır Sayısı

X_6 : Otobüs Sayısı

X_7 : Karayollarında enerji tüketimi (bin/ton)

X_8 : Yol-Km Uzunluğu

X_9 : AOGT

Çok değişkenli regresyon modellerinde normallik bir varsayım olduğu için değişkenleri modellemeyen önce normal dağılıma uygunluk analizi yapıldı. (Çizelge 3.1)

$$k_i = \log(X_i)$$

$$t_i = \log(Y_i)$$

Çizelge 3.1. Değişkenler için normallik test sonuçları

Değişken	Kolmogorov-Smirnov			Kolmogorov-Smirnov (Logaritmik)		
	İstatistik	Sd	Olabilirlik	İstatistik	Sd	Olabilirlik
Y	0.221	34	0.000	0.148	34	0.056
X ₁	0.215	34	0.000	0.181	34	0.060
X ₂	0.143	34	0.078	0.143	34	0.074
X ₃	0.161	34	0.026	0.143	34	0.077
X ₄	0.229	34	0.000	0.121	34	0.200
X ₅	0.159	34	0.029	0.094	34	0.200
X ₆	0.129	34	0.161	0.101	34	0.200
X ₇	0.093	34	0.200	0.133	34	0.137
X ₈	0.132	34	0.142	0.129	34	0.129
X ₉	0.092	34	0.168	0.08	34	0.200

H₀: Değişkenler normal dağılım gösterir.

H₁: Değişkenler normal dağılım göstermez.

Kolmogorov-Smirnov testinde olabilirlikler %5 ten küçük ise seçenek hipotezi H₀ ret edilmez.

Çizelge 3.2. Korelasyon matrisi

Değiş.	t	k ₁	k ₂	k ₃	k ₄	k ₅	k ₆	k ₇	k ₈	k ₉
t	1									
k ₁	0.995	1								
k ₂	0.976	0.973	1							
k ₃	0.988	0.987	0.962	1						
k ₄	0.963	0.958	0.916	0.976	1					

Çizelge 3.2. (devam)

k_5	0.953	0.947	0.936	0.959	0.982	1				
k_6	0.985	0.980	0.977	0.978	0.973	0.985	1			
k_7	0.952	0.956	0.967	0.944	0.936	0.967	0.983	1		
k_8	0.702	0.706	0.569	0.714	0.782	0.702	0.664	0.591	1	
k_9	0.959	0.958	0.970	0.951	0.937	0.960	0.981	0.987	0.955	1

Bağımlı ve bağımsız değişkenlerimiz arasındaki yüksek korelasyon göze çarpmaktadır (Çizelge 3.2).

TBA uygulaması yapmadan önce verilerin bu yöntemle uygunluğu analiz edildi. Küresellik Testi yapılarak olabilirlik $P < 0.05$ olduğundan ve Kaiser-Meyer-Olkin test istatistik değeri 0.70 ve üzeri olduğundan, H_0 seçeneğin hipotezi red edilerek, değişkenler arası ilişkinin önemli olduğu ve TBA'nın uygun bir yöntem olduğu anlaşıldı (Çizelge 3.3).

Çizelge 3.3. TBA Küresellik testi sonuçları

Kaiser-Meyer-Olkin	0.834	Mükemmel
Barlet Test İstatistiği	869.684	
Olabilirlik	0.000	

Çoklu bağlantı problemini çözmek için TBA uygulaması yapılarak p adet değişken yerine p adet temel bileşenle modelimiz kuruldu. Özdeğerler ve öz vektörler Çizelge 3.4'de sunulmuştur.

Çizelge 3.4. Temel bileşenler analizi, Özvektörler

Değişkenler	Vektör-1	Vektör-2	Vektör-3	Vektör-4	Vektör-5	Vektör-6	Vektör-7	Vektör-8	Vektör-9
k ₁	-0.3439	0.0181	0.4865	-0.0694	0.0442	-0.4639	0.4429	0.4293	-0.2055
k ₂	-0.3373	0.2567	0.4295	-0.1121	-0.4881	0.4588	-0.1331	-0.2498	-0.3074
k ₃	-0.3441	-0.0112	0.3482	0.4762	0.3280	-0.0058	-0.5630	0.0987	0.3150
k ₄	-0.3429	-0.1498	-0.2494	0.4701	0.2029	-0.1180	0.2333	-0.5143	-0.4501
k ₅	-0.3428	0.0066	-0.5342	0.2481	-0.3451	0.1685	-0.0731	0.6074	-0.1175
k ₆	-0.3467	0.1016	-0.0811	0.0661	-0.2354	0.0243	0.4489	-0.2389	0.7375
k ₇	-0.3395	0.2243	-0.2748	-0.4381	-0.1030	-0.5735	-0.4291	-0.2082	-0.0018
k ₈	-0.2518	-0.8959	0.0494	-0.3204	-0.0736	0.1089	-0.0871	-0.0300	0.0554
k ₉	-0.3399	0.2188	-0.1598	-0.4176	0.6485	0.4365	0.1274	0.0937	-0.0247

Aynı şekilde temel bileşenler analizi sonucunda bulunan özdeğerler ve bunların toplam varyansı açıklama oranları Çizelge 3.5’te sunulmuştur.

Çizelge 3.5. Temel bileşenlerin bulunması ve toplam varyansı açıklama oranları

Temel Bileşenler	Başlangıç Özdeğerler					
	Toplam	%Varyans	%Kümülatif	Toplam	%Varyans	%Küm.
1	8.237	91.520	91.520	8.237	91.520	91.520
2	0.589	6.542	98.062	0.589	6.542	98.062
3	0.087	0.972	99.034			
4	0.046	0.506	99.540			
5	0.018	0.203	99.743			
6	0.012	0.135	99.878			
7	0.006	0.066	99.943			
8	0.004	0.044	99.988			
9	0.001	0.012	100.000			

Öz değerler;

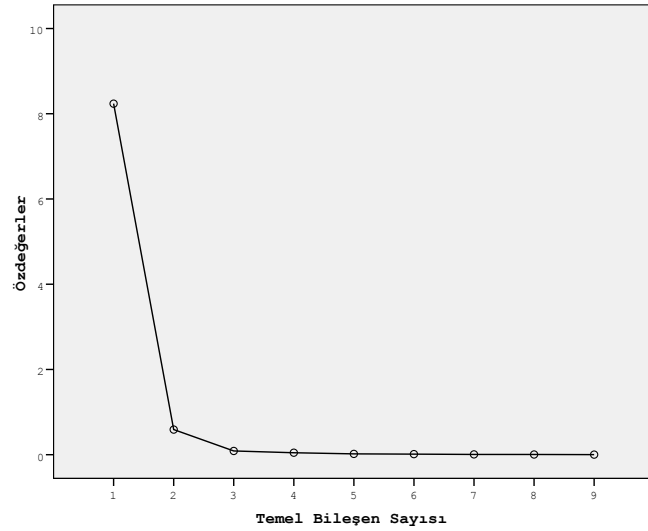
$$\lambda_1 = 8.237, \lambda_2 = 0.589, \lambda_3 = 0.087, \lambda_4 = 0.046, \lambda_5 = 0.018, \lambda_6 = 0.012, \lambda_7 = 0.006$$

$$\lambda_8 = 0.004, \lambda_9 = 0.001$$

olarak bulunmuştur.

$$\sum_{j=1}^m \frac{\lambda_j}{p} = \frac{8.237 + 0.589}{9} = 0.98 > \frac{2}{3}$$

olduğundan 2 adet temel bileşen vardır. Bu oranın haricinde özdeğerlerden (λ_j) 1'in altında olanlar elenebilir. 9 Bağımsız değişken yerine 2 temel bileşen ile devam edilecektir. Çizelge 3.5 incelendiğinde birinci temel bileşen toplam varyansın % 91.52'sini açıklarken, ikinci temel bileşen % 6.54'ünü açıklamaktadır. Bu iki temel bileşen toplam varyansın % 98.06'lık oranını açıklamaktadır. Aynı şekilde yamaç eğimi grafiği çizilerek de temel bileşen sayısı bulunabilir. Eğimin bittiği noktada sol tarafta kalan sayı temel bileşen sayısı olarak kabul edilebilir. Şekil 3.1'de görüldüğü gibi ikinci temel bileşenden itibaren eğim azalmaktadır. Bu durumda temel bileşen sayımız 1 veya 2 olarak alınabilir.



Şekil 3.1. Temel bileşen sayısı için yamaç eğimi grafiği (scree grafiği)

Çizelge 3.6. Değişkenlerin temel bileşenler tarafından açıklanma oranı

Değişkenler	Oran
k ₁ (yük-ton)	0.974
k ₂ (yolcu-km)	0.976
k ₃ (km-taşıt)	0.975
k ₄ (kamyon sayısı)	0.982
k ₅ (tır sayısı)	0.968
k ₆ (otobüs sayısı)	0.996
k ₇ (enerji tüketimi)	0.979
k ₈ (yol-km uz)	0.995
k ₉ (AOGT)	0.980

Değişkenlerin temel bileşenler tarafından açıklanma oranı Çizelge 3.6’da verilmiştir. Temel bileşenler tarafından en çok açıklanan bileşenler %99.5 ile yol-km uzunluğu ve %99.6 ile otobüs sayısı, %97.9 ile enerji tüketimi, %97.6 ile yolcu-km ve km-taşıt sayıları, %96.8 ile tır sayısı, %97.4 ile de yük-ton, %98 ile AOGT’dir.

Çizelge 3.7. Temel bileşenler matrisi

Değişkenler	Temel Bileşenler	
	1	2
k ₆ (otobüs sayısı)	0.995	-0.078
k ₃ (km-taşıt)	0.987	0.009
k ₁ (yük-km)	0.987	-0.014
k ₄ (kamyon sayısı)	0.984	0.115
k ₅ (tır sayısı)	0.984	-0.005
k ₇ (enerji tüketimi)	0.975	-0.168
k ₂ (yolcu-km)	0.968	-0.172
k ₈ (yol-km uz)	0.723	-0.197
k ₉ (AOGT)	0.975	0.687

Birinci temel bileşenler matrisimizde, % 99.5 ile otobüs sayısı %98.7 ile km-taşıt, %98.7 ile yük-km ve kamyon sayısı, %98.4 ile tır sayısı, %97.5 ile enerji tüketimi, %96.8 ile yolcu-km, %72.3 ile yol-km uzunluğu, %97.5 ile AOGT etkili olmuştur (Çizelge 3.7).

TBA sonucunda birbirleri ile ilişkili 9 bağımsız değişken yerine birbirleriyle ilişkisiz 2 adet temel bileşen ile model oluşturulmuştur. Bu model ile ilgili istatistiksel sonuçlar Çizelge 3.9, ve Çizelge 3.10'da sunulmuştur.

Çizelge 3.8. Bağımsız değişken ve temel bileşen skorları ile korelasyon matrisi

Değişkenler	Y	Z ₁	Z ₂
Y	1		
Z ₁	0.890	1	0.000
Z ₂	0.432	0.000	1

Görüldüğü üzere temel bileşen skorları arasındaki korelasyon sıfır olmuştur (Çizelge 3.8). Aynı regresyonun diğer parametreleri Çizelge 3.9'da gösterilmiştir.

Çizelge 3.9. Regresyon modeli çıktıları

Model	R	R ²	Düzeltilmiş R ²
1	0.988	0.977	0.975

Çizelge 3.10. Bağımsız değişken ve temel bileşen skorları ile regresyon modeli parametreleri

Değişkenler	B (Katsayı)	t	Olab. P	Tolerans	VIF	F	Olab. P
Sabit	5.180	438.498	0.000	1.00	1.00	655.609	0.000
SZ ₁	0.388	32.320	0.000	1.00	1.00		
SZ ₂	0.196	16.330	0.000	1.00	1.00		

Çoklu regresyon modelleri, birinci temel bileşen için $y = \beta_0 + \beta_1(sz_1) + \varepsilon$, ikinci temel bileşen için

$$y = \beta_0 + \beta_1(sz_1) + \beta_2(sz_2) + \varepsilon \text{ dir.} \quad (3.1)$$

Y orjinal bağımlı değişkenimizin standartlaştırılmış hali ile (SY), ikinci temel bileşenin standartlaştırılmış Z_1 , Z_2 arasında bir regresyon denklemi elde edildi (Çizelge 3.10).

$$SY = \beta_1 * Z_1 + \beta_2 * Z_2 \quad (3.2)$$

Denklemimizde bulunan katsayılar (-0.344, 0.024) sırasıyla birinci ve ikinci temel bileşenle çarpılarak ilk iki temel bileşen için regresyon katsayıları $\tilde{b}_1, \tilde{b}_2, \tilde{b}_3, \dots$ elde edildi.

$$\tilde{b}_i = V_i * \beta_i \quad (3.3)$$

Burada V_i i'ninci temel bileşen vektörünü, β_i 'de regresyon denklemindeki katsayıyı ifade etmektedir.

$$b_i = \tilde{b}_i(S_y / S_{x_i}) \quad (3.4)$$

Burada, S_y bağımlı değişkenin standart sapmasını ifade etmekte S_{x_i} bağımsız değişkenlerin standart sapmasını ifade etmektedir.

$$b_0 = \bar{t} - [b_1 \bar{x}_1 + b_2 \bar{x}_2 + b_3 \bar{x}_3 + b_4 \bar{x}_4 + b_5 \bar{x}_5 + b_6 \bar{x}_6] \quad (3.5)$$

Ters dönüşüm sonucunda model parametreleri aşağıdaki gibi olmuştur (Çizelge 3.11).

Çizelge 3.11. İkinci temel bileşen için bulunan katsayılar ve en küçük kareler kestiricileri

Değişken	2. Temel Bileşen İçin Denklem	
	$\bar{b}_i$	b_i
	0.118746	0.185252
	0.122194	0.322975
	0.118086	0.221942
	0.114377	0.135824
	0.118097	0.226115
	0.121705	0.192535
	0.122183	0.289786
	0.065115	3.896268
	0.122171	0.242882
b_0	-21.8063	
R^2	0.975	

Çok değişkenli zaman serisi analizinde logaritmik dönüşüm ile normal dağılıma uygun olan değişkenlerimize birim kök testlerinden, Genişletilmiş Dickey Fuller testi uygulanmış ve Çizelge 3.12’deki sonuçlar elde edilmiştir.

Çizelge 3.12. Çok değişkenli zaman serisi analizi için birim kök testi sonuçları

Değişkenler	Fark Serisi Alınmadan Önce			1. Fark Serisi Alındıktan Sonra		
	ADF t istatistik	%5 Tablo Değeri	Olabilirlik	ADF t istatistik	%5 olabilirlik değeri	Olabilirlik
Değişkenler t(kaza sayısı)	-2,73	-3,56	0,23	-3,57	-3,55	0,05
k_1 (yük-ton)	-2,46	-3,57	0,34	-5,72	-3,55	0,003
k_2 (yolcu-km)	-1,32	-3,55	0,86	-4,43	-3,56	0,0069
k_3 (km-taşıt)	-2,85	-3,55	0,19	-4,24	-3,55	0,0107
k_4 (kamyon sayısı)	-2,12	-3,59	0,51	-4,18	-3,60	0,004
k_5 (tır sayısı)	-2,11	-3,57	0,52	-5,56	-3,55	0,0028
k_6 (otobüs sayısı)	-2,59	-3,55	0,29	-4,79	-3,56	0,0004
k_7 (enerji tüketimi)	-2,76	-3,55	0,22	-5,55	-3,60	0,005
k_8 (yol-km uz)	-8,07	-3,60	0,00	-	-	-
k_9 (AOGT)	-1.19	-3.60	0.88	-4.26	-3.60	0.01

Mutlak deęer olarak $t_{ADF} > t_{0.05}$ olabilirlik deęeri ise seri duraęandır. Deęişkenlerimiz, birinci fark serisi alınarak duraęan hale getirilmiştir.

En uygun gecikme sayısının tespiti için gerekli olan kriterler ve sonuçları Çizelge 3.13’de verilmiştir. Bu durumda nihayi tahmin hatası, Akaike bilgi kriteri, Schwarz bilgi kriteri, Hannan-Quin bilgi kriterlerine bakılarak minimum olan deęerler karşılaştırılmış olup bütün bu kriterlerden AIC deęeri 1. gecikmede minimum olduğundan VAR(p) modelinin derecesi 1 olarak alınmıştır.

Çizelge 3.13. Model derecesi belirleme kriterleri

Gecikme Sayısı	FPE (NTH)	AIC	SC	HQ
0	2.21e-32	-47.34830	-46.93198*	-47.21259
1	4.30e-32	-51.18189*	-42.73413	-45.54022
2	2.93e-33*	-46.89732	-43.27183	-48.60341*

Model derecesi belirlendikten sonra deęişkenler arasındaki Granger Nedensellik analiz sonuçları Çizelge 3.14’de sunulmuştur.

$F_{hesaplanan} > F_{tablo}$ ise bir nedensellik var demektir.

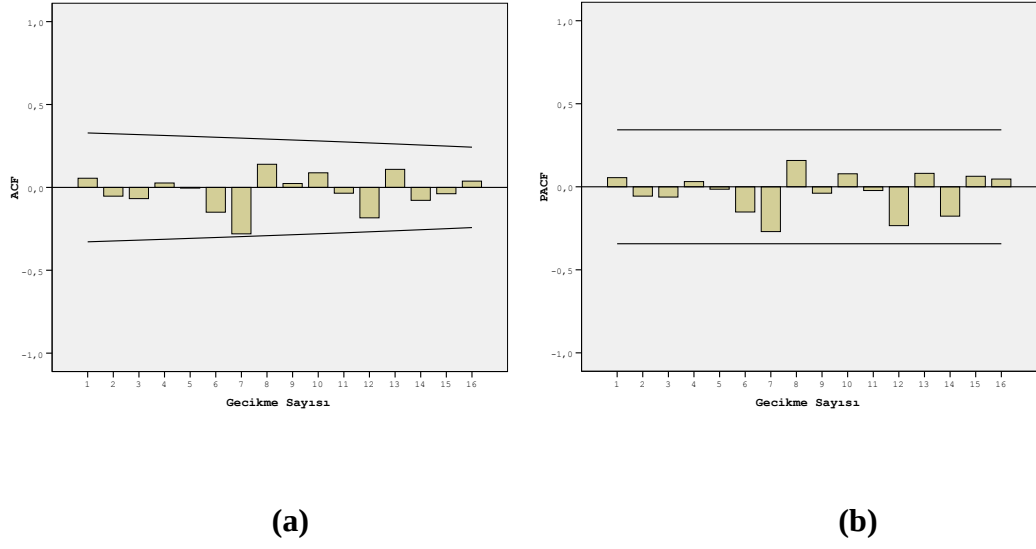
$H_0 : a_i = 0 (Y \rightarrow X)$ Y’den X’e nedensellik ilişkisi yoktur

$H_1 : a_i \neq 0 (Y \rightarrow X)$ Y’den X’e nedensellik ilişkisi vardır

Çizelge 3.14. Granger Nedensellik analiz sonuçları

Değişkenler	F-İstatistik	Olabilirlik
Ton-Km $\Updownarrow$ Kaza Sayısı Kaza Sayısı $\Rightarrow$ Ton-Km	0.25635 7.79298	0.7757 0.0021
Yolcu-km $\Updownarrow$ Kaza Sayısı Kaza Sayısı $\Updownarrow$ Yolcu-km	1.65671 1.14847	0.2096 0.3321
Kaza Sayısı $\Updownarrow$ Km-Taşıt Km-Taşıt $\Rightarrow$ Kaza Sayısı	0.14107 10.7680	0.8691 0.0004
Kaza Sayısı $\Updownarrow$ Kamyon Sayısı Kamyon Sayısı $\Updownarrow$ Kaza Sayısı	0.37672 0.97406	0.6897 0.3904
Kaza Sayısı $\Rightarrow$ Tır Sayısı Tır Sayısı $\Updownarrow$ Kaza Sayısı	3.85538 0.78179	0.0337 0.4677
Kaza Sayısı $\Rightarrow$ Otobüs Sayısı Otobüs Sayısı $\Updownarrow$ Kaza Sayısı	7.10407 0.25459	0.0033 0.7771
Karayollarında En. Tük. $\Rightarrow$ Kaza Sayısı Kaza Sayısı $\Updownarrow$ Karayollarında En. Tük.	4.22412 0.48498	0.0253 0.6210
Yol-km $\Updownarrow$ Kaza Sayısı Kaza Sayısı $\Updownarrow$ Yol-km	1.33644 1.50170	0.2796 0.2408
AOGT $\Rightarrow$ Kaza Sayısı Kaza Sayısı $\Updownarrow$ AOGT	5.51035 2.14319	0.02042 0.13681

Hata terimlerinin otokorelasyon ve kısmi otokorelasyon grafikleri çizilerek akgürültü bir seri oldukları görüldü (Şekil 3.2).

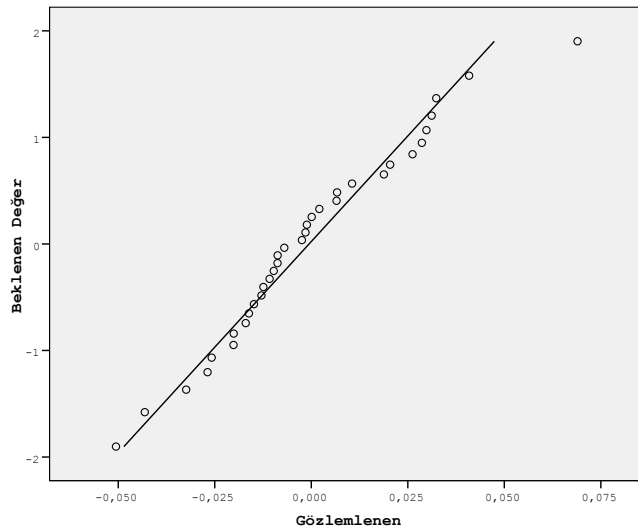


Şekil 3.2. Hata terimleri otokorelasyon (a) ve kısmi otokorelasyon (b) grafikleri

Çizelge 3.15. Hata terimleri normallik analizi

	Kolmogorov-Smirnov		
	İstatistik	sd	Olabilirlik
ε_i Hata Terimi	0.121	34	0.704

Olabilirlik $P > 0.05$ olduğundan seri normal dağılımlıdır (Çizelge 3.15).



Şekil 3.3. Hata terimleri normallik analizi Q-Q grafiği

Çizelge 3.16. Hata terimlerinde otokorelasyon testi LM

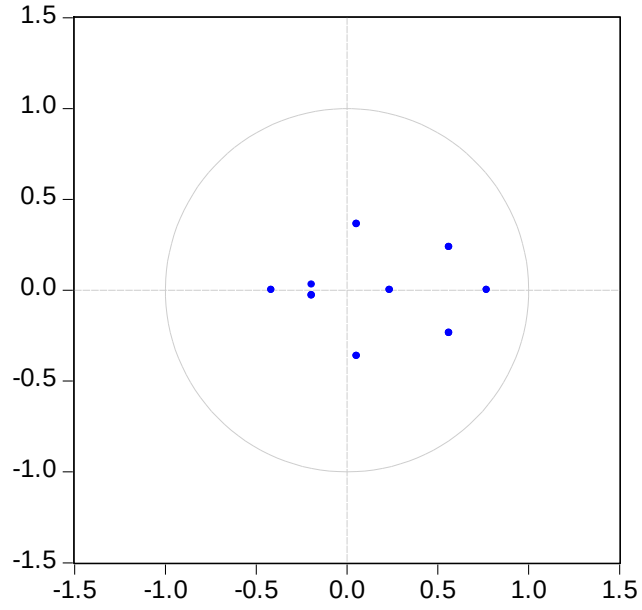
Gecikme Sayısı	LM-İstatistik	Olabilirlik
1	111.2789	0.0544
2	88.49305	0.2665
3	92.95525	0.1714
4	90.14755	0.2281
5	91.32902	0.2029
6	93.24808	0.1662
7	129.6872	0.5005
8	152.9377	0.6570
9	89.91583	0.2332
10	84.22255	0.3813
11	94.14238	0.1508
12	66.95759	0.8689

$H_0 : P_1 = P_2 = P_3 = \dots = P_{12} = 0 \Rightarrow$ Otokorelasyon yok

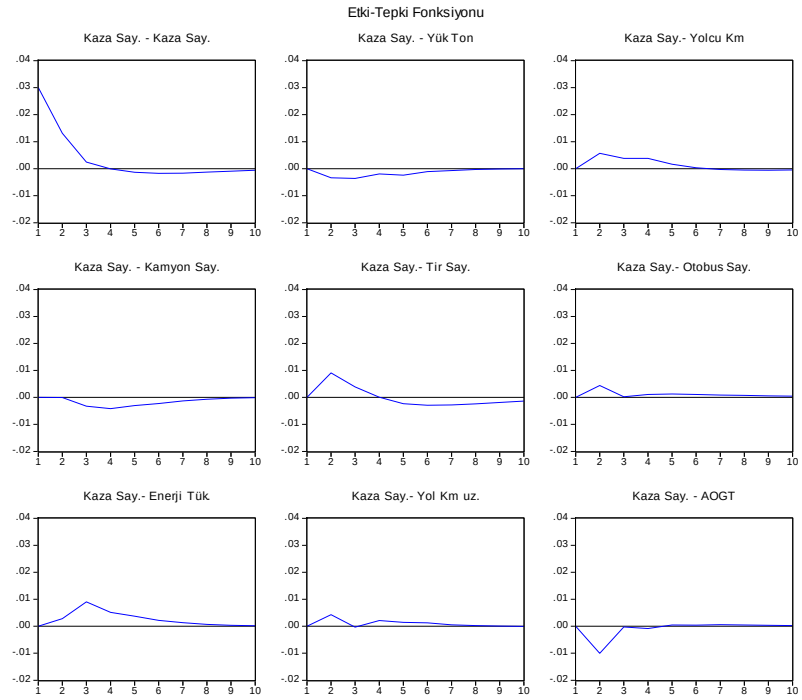
$H_1 : P_1 \neq P_2 \neq P_3 \neq \dots \neq P_{12} \neq 0 \Rightarrow$ Otokorelasyon var

$LM_{Olabilirlik} > 0.05$ bütün gecikmeler için olduğundan H_0 reddedilmez.

Modelin hata terimleri analiz edilirken otokorelasyonun olup olmadığı LM testi ile test edilmiş ve olabirliklerin 0.05 üzerinde olması münasebetiyle otokorelasyon olmadığı anlaşılmıştır (Çizelge 3.16). Hata terimlerinin normal dağılıma uygun olduğu Q-Q grafiği analiz edildi (Şekil 3.3). Karakteristik köklerin birim çember içinde kalıp kalmadığı da analiz edilmiştir. Şekil 3.4’de karakteristik köklerin birim çember içinde kalması modelde birim kökün olmadığını göstermiştir.



Şekil 3.4. VAR(1) Modeli için karakteristik köklerin gösterimi

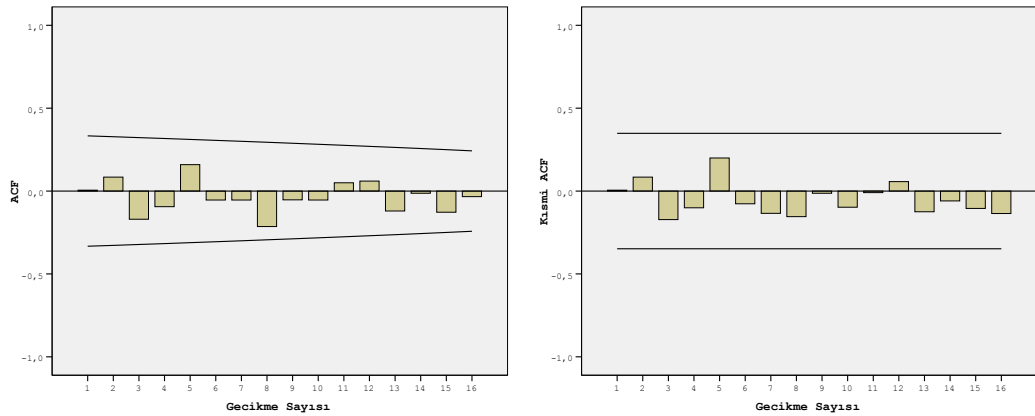


Şekil 3.5. VAR(1) Modeli için Etki-Tepki Fonksiyonu

Çizelge 3.17. VAR(1) Modeli varyans ayrıştırması

Peryot	Kaza Sayısı	Ton-Km	Yolcu-Km	Km-Taşıt	Kamyon Sayısı	Tır Sayısı	Otobüs Sayısı	Enerj.Tük	Yol-Km	AOGT
1	100.0000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
2	46.45062	2.296032	14.25545	0.370144	11.22103	1.523233	8.650276	12.75385	1.737083	0.742277
3	49.36089	2.830059	10.83301	0.616573	9.565389	1.700782	7.121807	15.1793	1.854336	0.937855
4	47.25326	2.956046	9.873576	1.164632	8.296411	3.743723	11.22064	13.07651	1.607587	0.807618
5	43.51109	2.908381	9.038587	3.154866	7.829557	7.030609	12.18661	12.01197	1.571993	0.756345
6	43.71823	2.858635	9.598183	3.109218	7.841626	6.88884	11.94406	11.748	1.547015	0.746191
7	41.23003	2.821452	10.66706	6.118685	7.629576	6.792952	11.46559	11.08041	1.490951	0.703286
8	38.81777	2.895803	10.46561	10.35776	7.468792	6.387718	11.11248	10.43328	1.400451	0.660335
9	35.62096	3.339571	12.75971	13.46533	7.043486	5.865724	10.17877	9.825281	1.294758	0.606415
10	32.06227	3.517501	17.364	13.5455	7.172323	5.304191	9.175033	10.0439	1.239487	0.575787

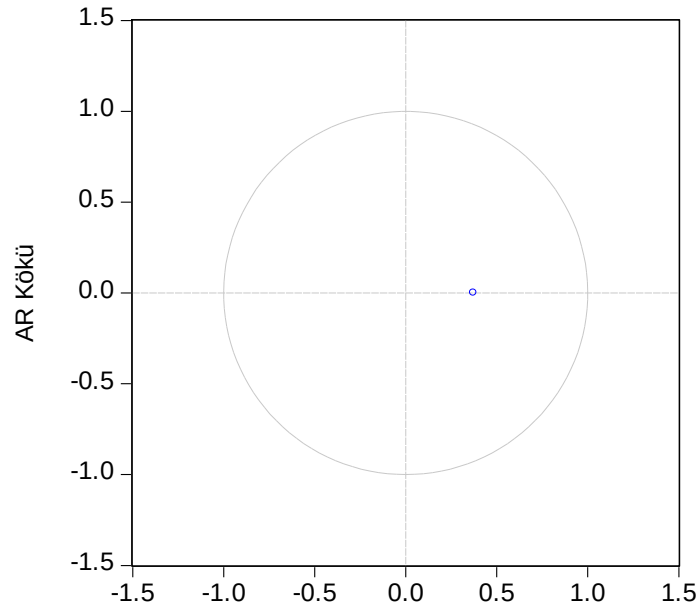
ARIMA(p,d,q) modelleri ile de bir analiz yapılmıştır. Analiz yapılırken SPSS 15.0 kullanılmıştır. Zaman serisi analizi menüsünde bağımlı değişken kısmına kaza sayısı (Y), bağımsız değişkenler kısmına da sırayla, yük-ton (k_1), yolcu-km (k_2), km-taşıt (k_3), kamyon sayısı (k_4), tır sayısı (k_5), otobüs sayısı (k_6), enerji tüketimi (k_7), yol-km (k_8), AOGT (k_9), değerleri girilerek birçok alternatif denenmiştir. Modelde I(d) fark alma derecesi 1 olarak belirlenmiştir. Serinin durağan olmadığı Dickey-Fuller birim kök testi ile analiz edilmiş ve birinci dereceden farkı alınıp durağanlaştırıldığından d(1) olmuştur. Model parametreleri en tutarlı olan model ARIMA(1,1,0) olarak bulunmuştur. Model hata terimleri analizi yapılmış Şekil 3.6'da sunulmuştur.

**Şekil 3.6.** ARIMA(1,1,0) Modeli için hata terimleri ACF PACF Grafikleri

Hata terimlerinin ACF ve PACF grafikleri çizildi. Grafiklerde gecikmeler güven aralıklarının içinde kaldığı görüldü (Şekil 3.6). Hata terimlerin akgürültü bir seri olduğu anlaşıldı. Hata terimlerinin normallik testi olarak Kolmogorov-Smirnov testi ile yapılmış, serinin olabilirlik P değeri $0.093 > 0.05$, %95 güven seviyesinden büyük olduğundan normal dağılıma uygundur denilmiştir (Çizelge 3.18). Son olarak modelin karakteristik köklerinin birim çember içinde kaldığının grafiği çizilmiş ve analiz edilmiştir (Şekil 3.7).

Çizelge 3.18. ARIMA(1,1,0) Modeli hata terimleri normallik testi

Değişken	Kolmogorov-Smirnov		
	İstatistik	sd	Olabilirlik
Model ARIMA(1,1,0)	0.149	33	0.061



Şekil 3.7. ARIMA (1,1,0) Model için karakteristik köklerin gösterimi

Son yöntem olarak YSA kullanılmıştır. YSA'da 9 adet girdi kullanılmış bu giriş vektörlerinin daha önceki yöntemlerde birbirleri ile korelasyonlu olması nedeniyle

YSA'da temel bileşenler analizi uygulanmıştır. Birbirleri ile ilişkisiz 2 temel bileşen girdi olarak kabul edilmiştir.

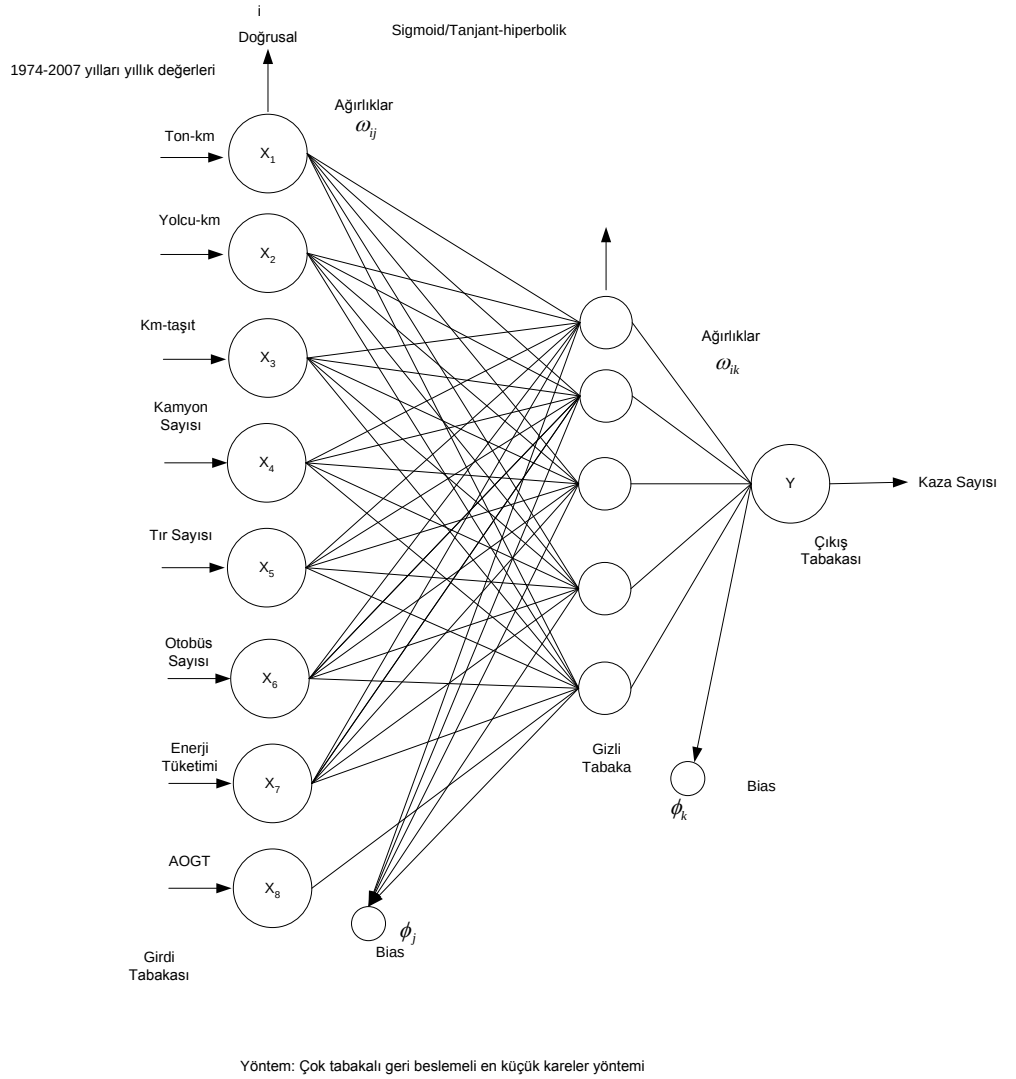
Aktivasyon fonksiyonu olarak trainbr kullanılmıştır. Bu fonksiyonun özelliği eğitim verisinden başka test verisine ihtiyaç duymamasıdır. Eldeki veri tamamen YSA'nın eğitimine kullanılmaktadır. YSA'nın aşırı eğitiminden dolayı ezberleme olayına izin vermemesidir (Can 2001).

Matlab'te yazılan bir yazılım, ağın transfer fonksiyonu olarak logsig, purelin, tanh hiperbolik fonksiyonlarını alarak gizli tabakadaki nöron sayısını 1-10 arasında iterasyon yapma özelliğine sahiptir. R^2 , MSE, ve AIC kriterlerine göre ağ mimarilerini sıralayan 1000 iterasyon yapma özelliğine sahip program en uygun iki ağ yapısı arasında bu belirtilen kriterlere göre tercih yapmayı sağlamıştır (Çizelge 3.19).

Çizelge 3.19. Optimum YSA ağ mimarisi

Ağ Yapısı	AIC (Akaike)	Ortalama Karese Hata (MSE)	R^2	Aktivasyon Fonksiyonu	Transfer Fonksiyonu	İterasyon Sayısı	Tabaka Sayısı	Gizli Tabakadaki nöron sayısı
Yöntem1	-113.439	0.009	0.98	Trainbr	Purelin-purelin-logsig	500	3	2
Yöntem2	-105.437	0.009	0.97	Trainbr	Purelin-purelin-logsig	500	2	3

Şekil 3.8'de örnek ağ mimarisi verilmiştir.



Şekil 3.8. YSA örnek ağ mimarisi

3.1.2. Model-II

II Numaralı yaklaşımda yıllık trafik kazalarındaki yaralı sayıları; nüfus, yol-km değeri, km-taşıt, ağır tonajlı araç sayısı, otomobil sayısı, otobüs sayısı, AOGT, GSMH değerlerine göre modellenmiştir.

Y: Yıllık trafik kazalarındaki yaralı sayıları

X_1 : Nüfus

X_2 : Yol-Km

X_3 : Km-Taşıt

X_4 : Ağır tonajlı araç sayısı

X_5 : Otomobil sayısı

X_6 : Otobüs sayısı

X_7 : AOGT

X_8 : GSMH

Model-II için normal dağılıma uygunluk analiz edilmiş, normal dağılıma uygun olmayan verilere doğal logaritmik dönüşüm uygulanmış ve test sonuçları Çizelge 3.20’de gösterilmiştir.

Çizelge 3.20. Değişkenler için normallik test sonuçları

Değişken	Kolmogorov-Smirnov			Kolmogorov-Smirnov (Logaritmik)		
	İstatistik	Sd	Olabilirlik	İstatistik	Sd	Olabilirlik
Y	0.219	34	0.087	1.275	34	0.078
X_1	0.082	34	0.142	0.522	34	0.948
X_2	0.132	34	0.026	0.755	34	0.619
X_3	0.161	34	0.002	0.832	34	0.493
X_4	0.197	34	0.007	0.561	34	0.912
X_5	0.179	34	0.161	0.698	34	0.714
X_6	0.129	34	0.200	0.587	34	0.881
X_7	0.092	34	0.081	0.632	34	0.819
X_8	0.142	34	0.002	0.661	34	0.661

$$k_i = \log(X_i)$$

$$t_i = \log(Y_i)$$

H_0 : Değişkenler normal dağılım gösterir.

H_1 : Değişkenler normal dağılım göstermez.

Çizelge 3.21. Korelasyon matrisi

Değişkenler	t	k ₁	k ₂	k ₃	k ₄	k ₅	k ₆	k ₇	k ₈
t	1								
k ₁	0.944	1							
k ₂	0.397	0.645	1						
k ₃	0.899	0.974	0.714	1					
k ₄	0.867	0.967	0.763	0.973	1				
k ₅	0.927	0.994	0.687	0.981	0.971	1			
k ₆	0.941	0.996	0.664	0.978	0.979	0.994	1		
k ₇	0.950	0.987	0.595	0.951	0.945	0.981	0.981	1	
k ₈	0.831	0.981	0.695	0.912	0.922	0.911	0.914	0.968	1

Bağımlı ve bağımsız değişkenlerin korelasyon matrisinde bağımsız değişkenler arasında yüksek korelasyon göze çarpmaktadır (Çizelge 3.21).

Çizelge 3.22. Temel bileşenler küresellik testi sonuçları

Keiser-Meyer-Olkin	0.804	Çok İyi
Barlett's	750.150	
Sd	21	
Olabilirlik	0.000	

Küresellik analizi yapılmış verilerin TBA uygulamasına uygun olduğu anlaşılmıştır (Çizelge 3.22).

Çizelge 3.23. TBA, Özvektörler

Değişkenler	Vektör-1	Vektör-2	Vektör-3	Vektör-4	Vektör-5	Vektör-6	Vektör-7	Vektör-8
k_1	-0.365607	0.198069	0.201418	0.087095	0.044218	0.437642	0.585893	0.492127
k_2	-0.276566	-0.913289	0.196804	0.153364	0.130800	-0.003366	0.084114	-0.054477
k_3	-0.366216	0.040235	0.055641	-0.822499	0.279122	-0.299094	0.115778	-0.062406
k_4	-0.367405	-0.059815	0.023557	-0.119202	-0.840925	-0.152669	-0.231841	0.249918
k_5	-0.368365	0.119199	0.121290	0.074806	0.365603	0.288275	-0.751228	0.220634
k_6	-0.367995	0.157765	0.070924	0.033703	-0.179696	0.412552	0.058472	-0.792222
k_7	-0.358933	0.283202	0.249872	0.499481	0.137186	-0.664056	0.092609	-0.103346
k_8	-0.347556	-0.046626	-0.913713	0.148211	0.105337	-0.037710	0.078487	0.038925

TBA sonucunda bulunan özvektörler, özdeğerler ve bunların toplam varyansı açıklama oranı Çizelge 3.23 ile Çizelge 3.24'te sunulmuştur.

Çizelge 3.24. TBA ve varyansı açıklama oranları

Temel Bileşenler	Başlangıç Özdeğerler					
	Toplam	%Varyans	%Kümülatif	Toplam	%Varyans	%Küm.
1	7.255	90.689	90.689	7.255	90.689	90.689
2	0.525	6.567	97.256	0.525	6.567	97.256
3	0.145	1.818	99.074			
4	0.033	0.418	99.492			
5	0.025	0.313	99.806			
6	0.012	0.148	99.954			
7	0.002	0.030	99.894			
8	0.001	0.016	100.000			

Öz değerler;

$$\lambda_1 = 7.255, \lambda_2 = 0.525, \lambda_3 = 0.145, \lambda_4 = 0.033, \lambda_5 = 0.025, \lambda_6 = 0.012, \lambda_7 = 0.002$$

$$\lambda_8 = 0.001$$

olarak bulunmuştur.

$$\sum_{j=1}^m \frac{\lambda_j}{p} = \frac{7.255 + 0.525}{8} = 0.973 > \frac{2}{3}$$

olduğundan 2 adet temel bileşen vardır. 8 Bağımsız değişken yerine 2 temel bileşen ile devam edilecektir. Çizelge 3.24 incelendiğinde birinci temel bileşen toplam varyansın % 90.689'unu açıklarken, ikinci temel bileşen % 6.567'sini açıklamaktadır. Bu iki temel bileşen toplam varyansın % 97.256'lık bir oranını açıklamaktadır.

Şekil 3.9'da temel bileşen sayısını tespit etmekte kullanılan yamaç eğimi grafiği çizilmiştir. Grafikte eğimin 2. temel bileşenden sonra yaklaşık değişmediği görüldüğünden temel bileşen sayısı 2 olarak kabul edilmiştir.

Çizelge 3.25. Değişkenlerin temel bileşenler tarafından açıklanma oranı

Değişkenler	Oran
k ₁ (nüfus)	0.990
k ₂ (yol-km)	0.993
k ₃ (km-taşıt)	0.974
k ₄ (ağır tonajlı araç)	0.981
k ₅ (otomobil sayısı)	0.992
k ₆ (otobüs sayısı)	0.996
k ₇ (AOGT)	0.977
k ₈ (GSMH)	0.878

Temel bileşenler tarafından en çok açıklanan bileşen, %99.6 ile otobüs sayısı, %99.3 ile yol-km uzunluğudur, %99.2 ile otomobil sayısı, %99 ile nüfus, %98.1 ile ağır tonajlı araç, %97.7 ile AOGT, %97.4 ile km-taşıt, %87.8 ile GSMH temel bileşenler tarafından açıklanmaktadır (Çizelge 3.25).

Çizelge 3.26. Temel bileşenler matrisi

Değişkenler	Temel Bileşenler	
	1	2
k ₁ (nüfus)	0.985	-0.144
k ₂ (yol-km)	0.745	0.662
k ₃ (km-taşıt)	0.986	-0.029
k ₄ (ağır tonajlı araç)	0.990	0.043
k ₅ (otomobil sayısı)	0.992	-0.086
k ₆ (otobüs sayısı)	0.991	-0.114
k ₇ (AOGT)	0.967	-0.205
k ₈ (GSMH)	0.936	0.034

Çizelge 3.26’da birinci temel bileşenler matrisimize göre; kaza sayısı üzerinde, % 99.2 ile otomobil sayısı %99.0 ile ağır tonajlı araç sayısı, %99.1 ile otobüs sayısı, %98.5 ile nüfus, %98.6 ile km-taşıt, %96.7 ile AOGT, %93.6 ile GSMH, %74.5 ile yol-km uzunluğu etkili olmuştur.

Temel bileşenler analizi sonucunda birbirleri ilişkili 8 bağımsız değişken yerine birbirleriyle ilişkisiz 2 adet temel bileşen ile modelimiz oluşturuldu. Bu model ile ilgili istatistiksel sonuçlar Çizelge 3.28 ve Çizelge 3.29’da sunulmuştur.

Çizelge 3.27. Bağımsız değişken ve temel bileşen skorları ile korelasyon matrisi

Değişkenler	Y	Z ₁	Z ₂
Y	1		
Z ₁	0.986	1	
Z ₂	0.062	0.000	1

Görüldüğü üzere temel bileşen skorları arasındaki korelasyon sıfır olmuştur (Çizelge 3.27). Regresyonun diğer parametreleri, Çizelge 3.28’de verilmiştir.

Çizelge 3.28. Regresyon modeli çıktıları

Model	R	R ²	Düzeltilmiş R ²
1	0.988	0.976	0.975

Çizelge 3.29. Bağımsız değişken ve temel bileşen skorları ile regresyon modeli parametreleri

Değişkenler	B (Katsayı)	t	Olab. P	Tolerans	VIF	F	Olab. P
Sabit	4.834	701.982	0.000	1.00	1.00	636.849	0.000
SZ ₁	0.249	35.620	0.000	1.00	1.00		
SZ ₂	0.016	2.223	0.034	1.00	1.00		

$$y = 4.834 + 0.249(sz_1) + 0.016(sz_2)$$

Orjinal değişkenler üzerinden kestirim yapabilecek denklemleri bulmak istatistiksel açıdan daha kullanışlıdır. Bu durumda en küçük kareler kestiricileri yardımıyla yapılan bir ters dönüşüm ile orijinal verilerden elde edilecek model bulunmuştur.

Y orjinal bağımlı değişkenimizin standartlaştırılmış hali ile (SY), ikinci temel bileşenin standartlaştırılmış Z₁-Z₂ arasında bir regresyon denklemi bulunmuştur. Denklem (3.2), (3.3) ve (3.4) yardımıyla

$$SY = -0.334 * Z_1 + 0.567 * Z_2$$

denklemleri elde edilmiştir. Orijinal verilerin modelde kullanılabilmesi için ters dönüşüm sonucunda bulunan model parametreleri Çizelge 3.30’da sunulmuştur.

Çizelge 3.30. İkinci temel bileşen için bulunan katsayılar ve en küçük kareler kestiricileri

Değişken	2. Temel Bileşen İçin Denklem	
	$\bar{b}_i$	b_i
	0.23442	0.68888
	-0.42546	-14.6312
	0.14513	0.15677
	0.08880	0.07021
	0.19206	0.12397
	0.21236	0.19308
	0.28046	0.32045
	0.08965	0.11101
b_0	67.23424	
R^2	0.975	

Çizelge 3.31. Çok değişkenli zaman serisi analizi için birim kök testi sonuçları.

Değişkenler	ADF t istatistiği	%5 Tablo Değeri	P	ADF t istatistiği	%5 olabilirlik değeri	Olabilirlik	ADF t istatistiği	%5 olabilirlik değeri	P
	Fark Serisi Alınmadan Önce			1. Fark Serisi Alındıktan Sonra			2. Fark Serisi Alındıktan Sonra		
t (yaralı say.)	-1.40	-3.55	0.84	-3.97	-3.55	0.02			
k ₁ (nüfus)	1.68	-3.57	1.00	-3.42	-3.57	0.06	-6.58	-3.58	0.00
k ₂ (yol-km)	-8.07	-3.60	0.00						
k ₃ (km-taşıt)	-2.85	-3.56	0.19	-4.24	-3.56	0.01			
k ₄ (ağır ton.arac)	0.07	-3.59	0.99	-4.40	-3.56	0.009			
k ₅ (oto.say.)	-1.55	-3.60	0.78	-3.42	-3.61	0.07	-6.09	-3.56	0.0001
k ₆ (otob.say.)	-2.59	-3.56	0.28	-4.79	-3.56	0.003			
k ₇ (AOGT)	-1.19	-3.60	0.88	-4.79	-3.56	0.003			
k ₈ (GSMH)	-3.65	-3.60	0.05						

Mutlak değer olarak $t_{ADF} > t_{\%5}$ ise seri durağandır. Değişkenlerimiz birinci fark serisi ve ikinci fark serisi alınarak durağan hale dönüşmüşlerdir (Çizelge 3.31).

En uygun gecikme sayısının tespiti için gerekli olan kriterler ve sonuçları Çizelge 3.32’de verilmiştir. Bu durumda nihai tahmin hatası, Akaike bilgi kriteri, Schwarz bilgi

kriteri, Hannan-Quin bilgi kriterlerine bakılarak minimum olan değerler karşılaştırılmış olup bütün bu kriterlerden 2. gecikmeye ait değerler minimum olduğundan VAR(p) modelinin derecesi 2 olarak alınmıştır (Çizelge 3.32).

Çizelge 3.32. Model derecesi belirleme kriterleri

Gecikme Sayısı	FPE (NTH)	AIC	SC	HQ
0	7.04e-33	-48.49315	-48.07279*	-48.35868
1	2.18e-34	-52.20417	-48.00058	-50.85941
2	7.14e-35*	-55.12923*	-47.14240	-52.57418*

Model derecesi belirlendikten sonra değişkenler arasındaki Granger Nedensellik analiz sonuçları Çizelge 3.33'te sunulmuştur.

$F_{\text{hesaplanan}} > F_{\text{tablo}}$ ise bir nedensellik var demektir.

$H_0 : a_i = 0 (Y \rightarrow X)$ Y'den X'e nedensellik ilişkisi yoktur

$H_1 : a_i \neq 0 (Y \rightarrow X)$ Y'den X'e nedensellik ilişkisi vardır

Çizelge 3.33. Model-II Granger Nedensellik analiz sonuçları

Değişkenler	F-İstatistiği	P
Nüfus $\Rightarrow$ Yaralı Sayısı	4.89959	0.0153
Yaralı Sayısı $\Updownarrow$ Nüfus	1.21476	0.3125
Yol-km $\Updownarrow$ Yaralı Sayısı	0.56606	0.5744
Yaralı Sayısı $\Updownarrow$ Yol-km	4.17839	0.0262
Yaralı Sayısı $\Updownarrow$ Km-Taşıt	0.76293	0.4761
Km-Taşıt $\Updownarrow$ Yaralı Sayısı	1.28574	0.2928
Yaralı Sayısı $\Updownarrow$ Ağır Tonajlı Araç Sayısı	1.04984	0.3638
Ağır Tonajlı Araç Sayısı $\Updownarrow$ Yaralı Sayısı	0.24562	0.7840

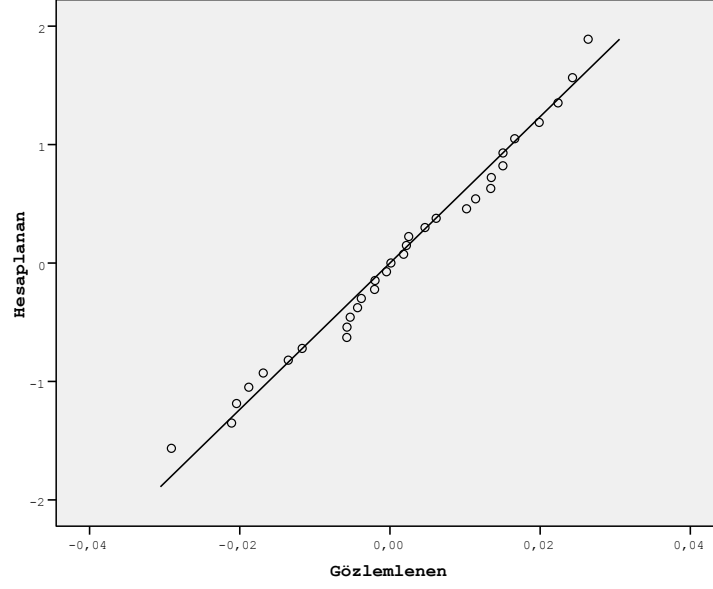
Çizelge 3.33. (devam)

Yaralı Sayısı ⇕ Otomobil Sayısı	0.96458	0.3939
Otomobil Sayısı ⇕ Yaralı Sayısı	2.18559	0.1319
Yaralı Sayısı ⇕ Otobüs Sayısı	1.32705	0.2820
Otobüs Sayısı ⇕ Yaralı Sayısı	0.28672	0.7530
AOGT ⇒ Yaralı Sayısı	3.60660	0.0409
Yaralı Sayısı ⇕ AOGT	0.75773	0.4784
GSMH ⇕ Yaralı Sayısı	0.83183	0.4461
Yaralı Sayısı ⇕ GSMH	1.47328	0.2470

Çizelge 3.34. Hata terimleri normallik analizi

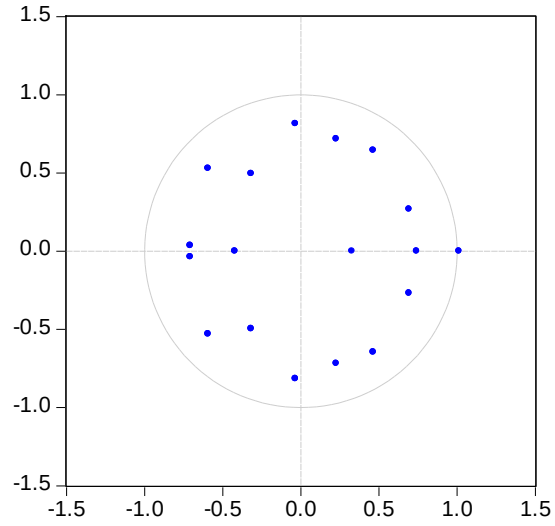
	Kolmogorov-Smirnov		
	İstatistik	sd	Olabilirlik
ε_i Hata Terimi	0.104	34	0.200

Serinin olabilirlik P değeri $0.200 > 0.05$, %95 güven seviyesinden büyük olduğundan normal dağılıma uygundur (Çizelge 3.34). Serinin hata terimlerinin normal dağılıma uygunluğu Q-Q grafiği ilede tespit edilmiştir (Şekil 3.9).

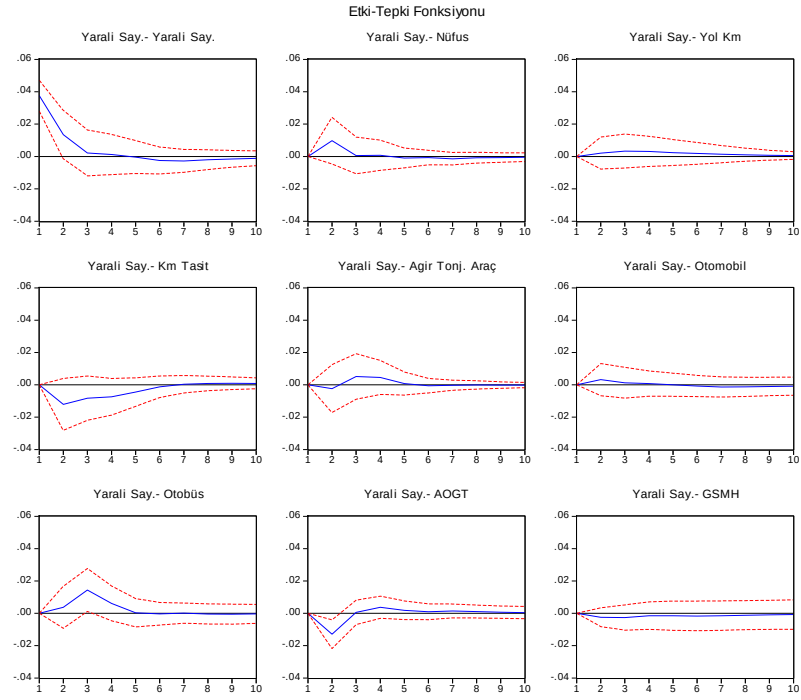


Şekil 3.9. Var(2) modeli hata terimleri normallik analizi Q-Q grafiği

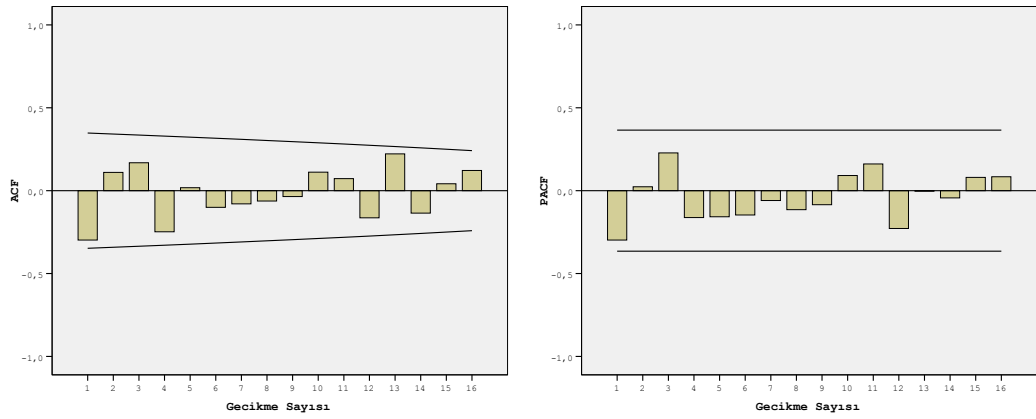
Karakteristik köklerin birim çember içinde kaldığı analiz edilmiştir (Şekil 3.10).



Şekil 3.10. VAR(2) Modeli için karakteristik köklerin gösterimi



Şekil 3.11. VAR(2) Modeli için Etki-Tepki Fonksiyonu

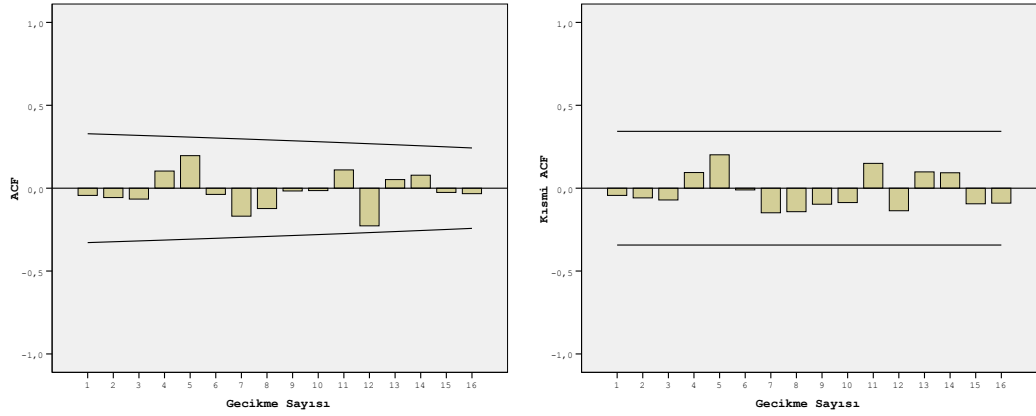


Şekil 3.12. VAR(2) Modeli için hata terimleri ACF-PACF grafikleri

Çok değişkenli zaman serisi uygulamasında ve ARIMA(2,1,0) uygulamasında hata terimlerinin ACF-PACF grafikleri çizildi. Bütün gecikmelerin güven aralığı içinde kaldığı ve serinin akgürültü bir seri olduğu anlaşıldı (Şekil 3.12-3.13).

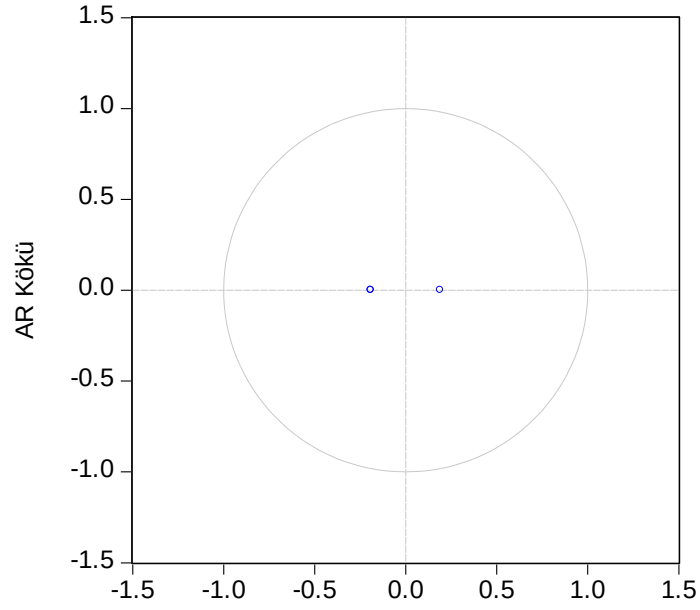
Çizelge 3.35. VAR(2) Modeli varyans ayrıştırması

Peryot	Yaralı Sayısı	Nüfus	Yol-Km	Km-Taşıt	Ağır Tonajlı Araç Say.	Otomobil Sayısı	Otobüs Sayısı	AOGT	GSMH
1	100.00	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
2	86.222	3.6409	7.2299	0.3005	0.0107	0.6157	0.2194	0.0480	1.7121
3	69.664	8.3026	5.7872	3.7854	0.4857	1.5475	7.8305	0.6884	1.9076
4	66.717	8.6406	6.0758	4.9964	0.5019	2.2861	8.2043	0.7027	1.8739
5	62.231	8.0594	6.6041	7.8626	0.9058	3.2055	7.6736	1.5408	1.9159
6	58.518	7.7772	8.4587	7.2867	2.3285	3.9890	7.3276	2.3790	1.9342
7	56.092	7.4748	10.333	8.8976	2.3876	3.7201	6.6867	2.5388	1.8680
8	57.776	7.2503	9.7871	8.5453	2.2967	3.7102	6.3462	2.4655	1.8222
9	58.107	7.2123	9.7063	8.4245	2.2735	3.7841	6.2652	2.4206	1.8059
10	57.584	7.1703	9.6809	8.3150	2.2428	3.9887	6.2529	2.7812	1.9839

**Şekil 3.13** ARIMA (2,1,0) Hata terimleri ACF ve PACF grafikleri**Çizelge 3.36** ARIMA(2,1,0) Modeli normallik testi

Değişken	Kolmogorov-Smirnov		
	İstatistik	Sd	Olabilirlik
Model ARIMA(2,1,0)	1.177	34	0.125

ARIMA(2,1,0) modelinin hata terimlerine normallik analizi uygulandı. Kolmogorov Smirnov test sonucunda serinin normal dağılıma uygun olduğu anlaşıldı (Çizelge 3.36). Modelin karakteristik kökleri birim çember içinde olduğu görüldü (Şekil 3.14).



Şekil 3.14. ARIMA (2,1,0) Model için karakteristik köklerin gösterimi

3.1.3. Model-III

Model III’de yıllık Trafik kazalarında ki ölü sayıları; nüfus, yol-km değeri, km-taşıt, ağır tonajlı araç sayısı, otomobil sayısı, otobüs sayısı, AOGT, GSMH, değerlerine göre modellenmiştir.

Y:Yıllık Trafik Kazalarındaki Ölü Sayıları

X_1 : Nüfus

X_2 : Yol-Km

X_3 : Km-Taşıt

X_4 : Ağır Tonajlı Araç Sayısı

X_5 : Otomobil Sayısı

X_6 : Otobüs Sayısı

X_7 : AOGT

X_8 : GSMH

$$k_i = \log(X_i)$$

$$t_i = \log(Y_i)$$

Model-III için normallik analizi yapıldı. Normal dağılıma uygun olmayan verilere doğal logaritmik dönüşüm uygulanmış ve test sonuçları Çizelge 3.37’de gösterilmiştir.

Çizelge 3.37. Bağımsız değişken için normallik test sonucu

Değişken	Kolmogorov-Smirnov			Kolmogorov-Smirnov (Logaritmik)		
	İstatistik	sd	Olabilirlik	İstatistik	sd	Olabilirlik
Y	0.095	34	0.200	0.123	34	0.200

H_0 : Değişkenler normal dağılım gösterir.

H_1 : Değişkenler normal dağılım göstermez.

Çizelge 3.38. Korelasyon matrisi

Değişkenler	t	k ₁	k ₂	k ₃	k ₄	k ₅	k ₆	k ₇	k ₈
t	1								
k ₁	-0.479	1							
k ₂	-0.864	0.645	1						
k ₃	-0.545	0.974	0.714	1					
k ₄	-0.607	0.967	0.763	0.973	1				
k ₅	-0.497	0.994	0.687	0.981	0.971	1			
k ₆	-0.487	0.996	0.664	0.978	0.979	0.994	1		
k ₇	-0.407	0.987	0.595	0.951	0.945	0.981	0.981	1	
k ₈	-0.472	0.891	0.695	0.912	0.922	0.911	0.914	0.868	1

Bağımlı ve bağımsız değişkenlerimiz arasındaki korelasyon matrisinde, bağımsız değişkenler arasındaki yüksek korelasyon göze çarpmaktadır (Çizelge 3.38). Model-

II’de kullanmış olduğumuz bağımsız değişkenler Model III’ten farklılık göstermediğinden TBA uygulamalarındaki bazı tablolarda değişiklik yoktur. Çizelge 3.22, 3.23, 3.24, 3.25 ve 3.26 aynıdır. Aynı zamanda Şekil 3.9’daki yamaç eğimi grafiği de aynıdır. Model-III uygulamasında da 8 bağımsız değişken yerine iki temel bileşenle model açıklanmaya çalışılacaktır.

Temel bileşenler analizi sonucunda birbirleri ile ilişkili 8 bağımsız değişken yerine, birbirleriyle ilişkisiz 2 adet temel bileşen kullanarak model oluşturuldu. Bu model ile ilgili istatistiksel sonuçlar Çizelge 3.40 ve Çizelge 3.41’de sunulmuştur.

Çizelge 3.39. Bağımsız değişken ve temel bileşen skorları ile korelasyon matrisi

Değişkenler	Y	Z ₁	Z ₂
Y	1		
Z ₁	-0.182	1	
Z ₂	-0.847	0.000	1

Görüldüğü üzere temel bileşen skorları arasındaki korelasyon sıfır olmuştur (Çizelge 3.39).

Çizelge 3.40. Regresyon modeli çıktıları

Model	R	R ²	Düzeltilmiş R ²
1	0.866	0.751	0.735

Çizelge 3.41. Bağımsız değişken ve temel bileşen skorları ile regresyon modeli parametreleri

Değişkenler	B (Katsayı)	t	P	Tolerans	VIF	F	P
Sabit	3.690	340.743	0.000	1.00	1.00	46.684	0.000
SZ ₁	-0.022	-2.035	0.051	1.00	1.00		
SZ ₂	-0.104	-9.446	0.000	1.00	1.00		

Orijinal verilerin modelde kullanılabilmesi için ters dönüşüm sonucunda bulunan model parametreleri Çizelge 3.42’de sunulmuştur.

Çizelge 3.42. İkinci temel bileşen için bulunan katsayılar ve en küçük kareler kestiricileri

Değişken	2. Temel Bileşen İçin Denklem	
	$\bar{b}_i$	b_i
	0.10446	0.149225
	-0.89105	-14.8667
	-0.03948	-0.0207
	-0.13097	-0.05159
	0.03297	0.010132
	0.06734	0.029724
	0.18362	0.101859
	-0.11481	-0.06903
b_0	74.11142	
R^2	0.735	

Çok değişkenli zaman serisi analizinde logaritmik dönüşüm ile normal dağılıma uygun olan bağımlı değişkenimize birim kök testlerinden Genişletilmiş Dickey Fuller testi uygulanmış ve Çizelge 3.43’deki sonuçlar elde edilmiştir. Bağımsız değişkenler Model-II uygulamasındakiyle aynı olması münasebetiyle Çizelge 3.31’de verilmişti.

Çizelge 3.43. Çok değişkenli zaman serisi analizi için birim kök testi sonuçları

Değişkenler	ADF t istatistik	%5 Tablo Değeri	Olabilirlik	ADF t istatistik	%5 olabilirlik değeri	Olabilirlik
	Fark Serisi Alınmadan Önce			1. Fark Serisi Alındıktan Sonra		
t	-1.88	-3.56	0.63	-3.83	-3.56	0.03

En uygun gecikme sayısının tespiti için gerekli olan kriterler ve sonuçları Çizelge 3.44’de verilmiştir. Bu durumda nihai tahmin hatası, Akaike bilgi kriteri, Schwarz bilgi kriteri, Hannan Quin bilgi kriterlerine bakılarak minimum olan değerler karşılaştırılmış

olup bütün bu kriterlerden 2. gecikmeye ait değerler minimum olduğundan VAR(p) modelinin derecesi 2 olarak alınmıştır.

Çizelge 3.44. Model derecesi belirleme kriterleri

Gecikme Sayısı	FPE (NTH)	AIC	SC	HQ
0	5.99e-33	-48.65383	-48.23348	-48.51936
1	1.77e-34*	-52.41472	-48.21113	-51.06996
2	8.47e-36	-57.22298*	-49.24305*	-54.67482*

Model derecesi belirlendikten sonra değişkenler arasında ki Granger Nedensellik analiz sonuçları Çizelge 3.45'te sunulmuştur.

$F_{\text{hesaplanan}} > F_{\text{tablo}}$ ise bir nedensellik var demektir.

$H_0 : a_i = 0 (Y \rightarrow X)$ Y'den X'e nedensellik ilişkisi yoktur

$H_1 : a_i \neq 0 (Y \rightarrow X)$ Y'den X'e nedensellik ilişkisi vardır

Çizelge 3.45. Model Granger Nedensellik analiz sonuçları

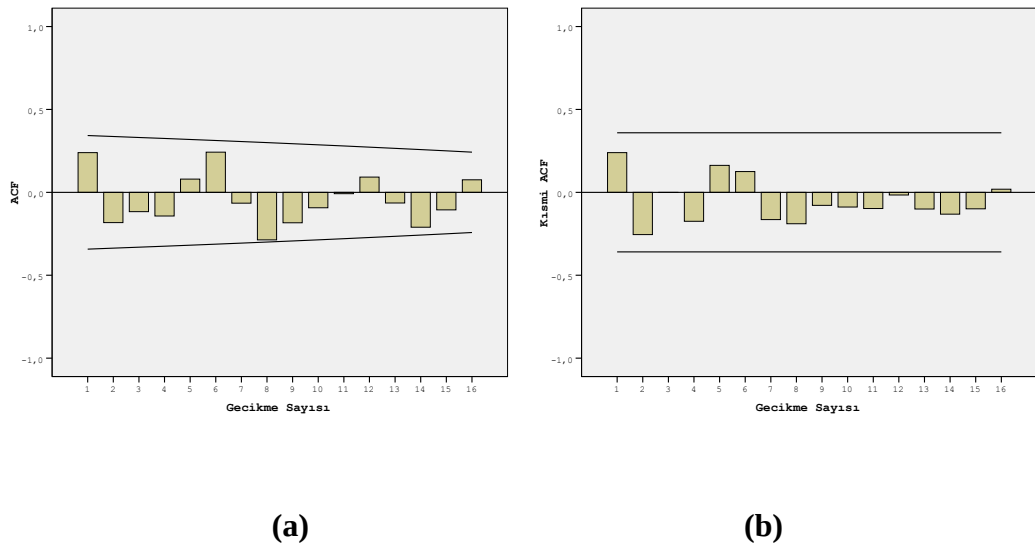
Değişkenler	F-İstatistik	Olabilirlik
Nüfus $\Rightarrow$ Ölü Sayısı	4.56670	0.0409
Ölü Sayısı $\Updownarrow$ Nüfus	2.08029	0.1596
Yol-km $\Rightarrow$ Ölü Sayısı	4.49415	0.0424
Ölü Sayısı $\Updownarrow$ Yol-km	1.29947	0.2633
Ölü Sayısı $\Rightarrow$ Km-Taşıt	9.94288	0.0037
Km-Taşıt $\Updownarrow$ Ölü Sayısı	0.18462	0.6705
Ölü Sayısı $\Rightarrow$ Ağır Tonajlı Araç Sayısı	5.10487	0.0313
Ağır Tonajlı Araç Sayısı $\Updownarrow$ Ölü Sayısı	5.63889	0.2142
Ölü Sayısı $\Rightarrow$ Otomobil Sayısı	6.42133	0.0167
Otomobil Sayısı $\Updownarrow$ Ölü Sayısı	0.56235	0.4592

Çizelge 3.45. (devam)

Ölü Sayısı $\Rightarrow$ Otobüs Sayısı	4.63269	0.0395
Otobüs Sayısı $\Updownarrow$ Ölü Sayısı	0.72426	0.4015
AOGT $\Rightarrow$ Ölü Sayısı	6.00818	0.0203
Ölü Sayısı $\Updownarrow$ AOGT	0.64804	0.4271
GSMH $\Rightarrow$ Ölü Sayısı	2.77692	0.3022
Ölü Sayısı $\Updownarrow$ GSMH	1.06469	0.3104

Granger nedensellik analizi yapıldıktan sonra model parameterleri belirlenmiştir. Modelin uygun bir model olup olmadığını anlayabilmek için hata terimleri üzerindeki varsayımların analizinin yapılması gerekmektedir.

Hata terimlerinin otokorelasyon ve kısmi otokorelasyon grafikleri çizilerek akgürültü bir seri oldukları görüldü (Şekil 3.15).

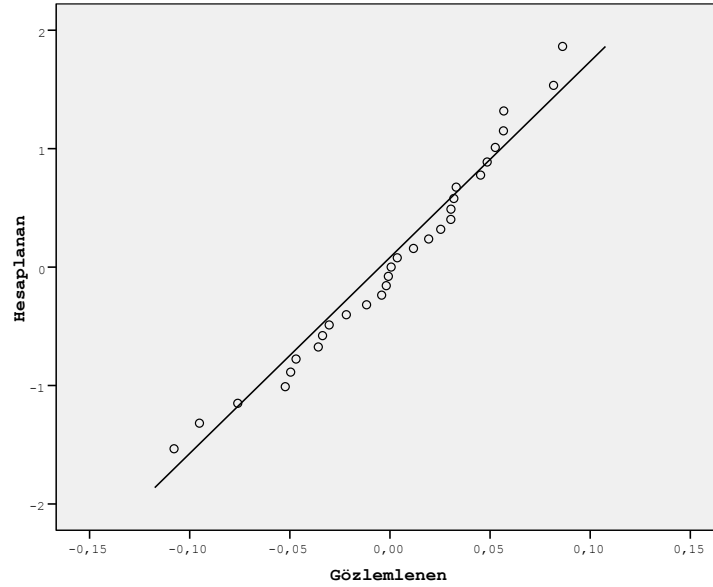


Şekil 3.15. Hata terimleri otokorelasyon (a) ve kısmi otokorelasyon (b) grafikleri. Hata terimleri normal dağılıma uygunlukları test edildi.

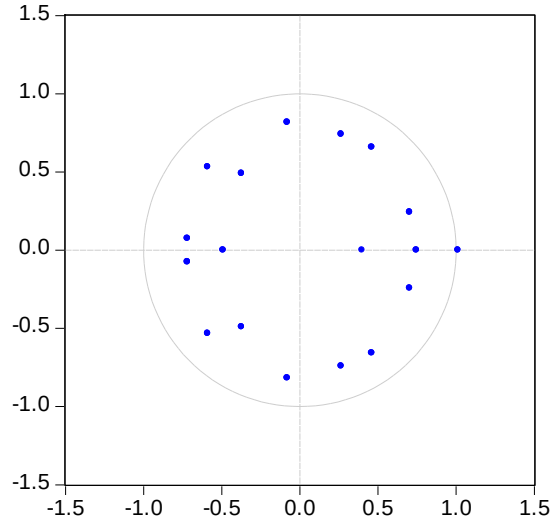
Çizelge 3.46. Hata terimleri normallik analizi

	Kolmogorov-Smirnov		
	İstatistik	Sd	Olabilirlik
ε_i Hata Terimi	0.118	33	0.200

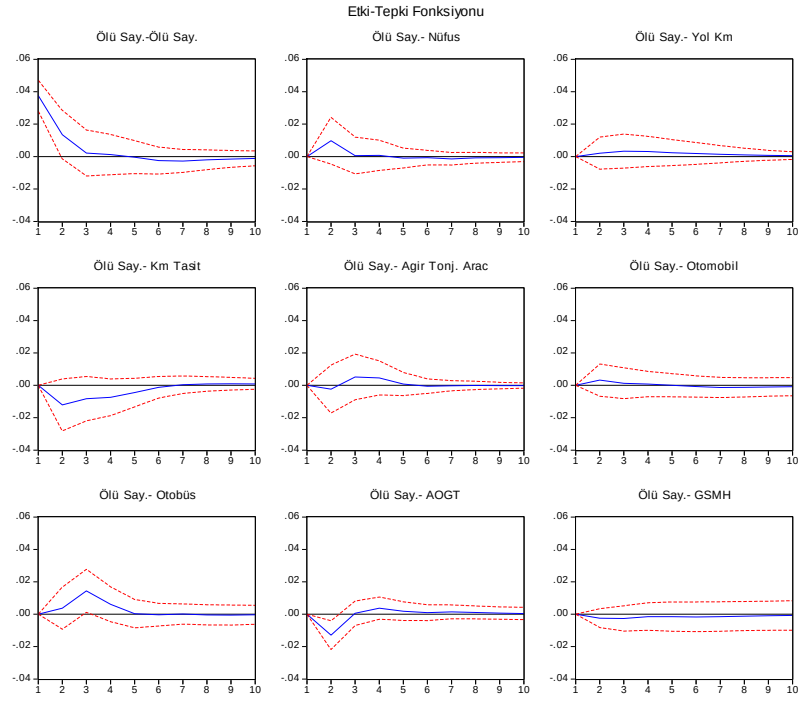
Olabilirlik $P > 0.05$ olduğundan hata terimleri normal dağılımlıdır (Çizelge 3.46-Şekil 3.16).

**Şekil 3.16.** Hata terimleri normallik analizi Q-Q grafiği

Karakteristik köklerin birim çember içinde kalıp kalmadığı analiz edildi. Bütün karakteristik kökler birim çember içinde olduğundan modelde birim kök olmadığı anlaşıldı (Şekil 3.17).



Şekil 3.17. VAR(2) Modeli için karakteristik köklerin gösterimi

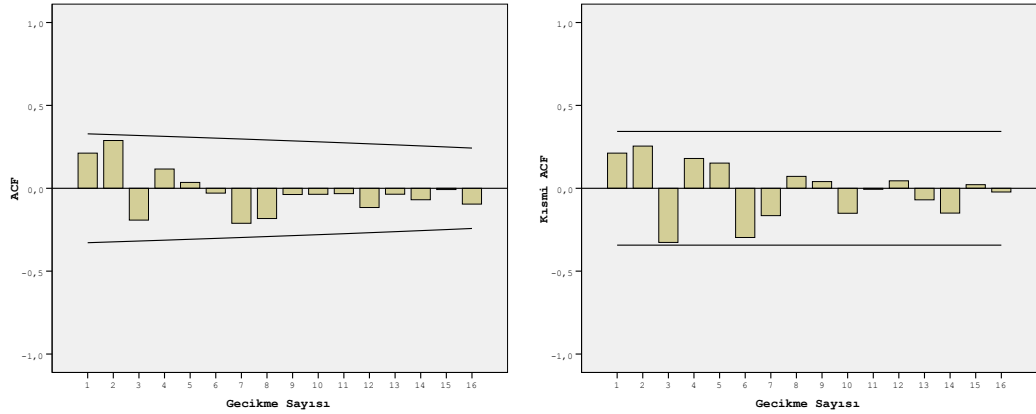


Şekil 3.18. VAR(2) Modeli için Etki-Tepki Fonksiyonu

Çizelge 3.47. VAR(2) Modeli varyans ayrıştırması

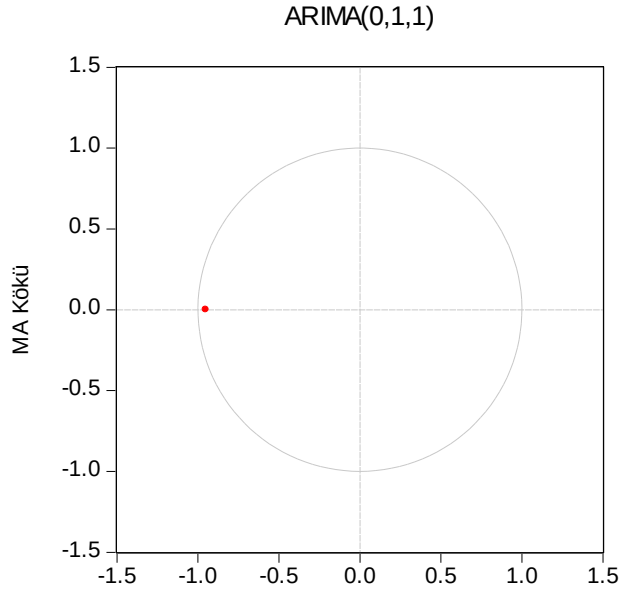
Peryot	Ölü Sayısı	Nüfus	Yol-Km	Km-Taşıt	Ağır Tonajlı Araç Say.	Otomobil Sayısı	Otobüs Sayısı	AOGT	GSMH
1	100.0000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.0000
2	78.10667	4.622056	0.213871	7.153873	0.265287	0.504103	0.660243	8.187221	0.2866
3	67.51323	3.997529	0.651505	9.044366	1.337053	0.501688	9.337100	7.068831	0.5487
4	63.78269	3.794757	0.994411	10.72341	2.079621	0.497605	10.29408	7.228576	0.6048
5	62.95964	3.780598	1.205117	11.37190	2.076007	0.491180	10.16355	7.267797	0.6842
6	62.80851	3.773967	1.330116	11.35701	2.075642	0.509933	10.10272	7.254957	0.7871
7	62.67116	3.825031	1.392099	11.28427	2.064248	0.576314	10.03351	7.287450	0.8659
8	62.56947	3.833847	1.428448	11.26297	2.055719	0.633170	9.998842	7.300900	0.9166
9	62.48681	3.843595	1.446473	11.26629	2.051246	0.673503	9.983709	7.298012	0.9503
10	62.43555	3.844987	1.455140	11.27777	2.048992	0.699563	9.973139	7.292261	0.9725

Model parametreleri en tutarlı olan model ARIMA(0,1,1) olarak bulunmuştur. Model hata terimleri akgürültü bir seri olduğu gecikmelerin güven aralıklarının içinde kaldığı Şekil 3.19’da görülmüştür.

**Şekil 3.19.** ARIMA(0,1,1) Modeli için hata terimleri ACF PACF Grafikleri**Çizelge 3.48.** ARIMA(0,1,1) Modeli hata terimleri normallik testi

Değişken	Kolmogorov-Smirnov		
	İstatistik	Sd	Olabilirlik
Model ARIMA(0,1,1)	0.132	34	0.139

Olabilirlik $P > 0.05$ olduğundan seri normal dağılımlıdır (Şekil 3.48).



Şekil 3.20. ARIMA (0,1,1) Model için karakteristik köklerin gösterimi

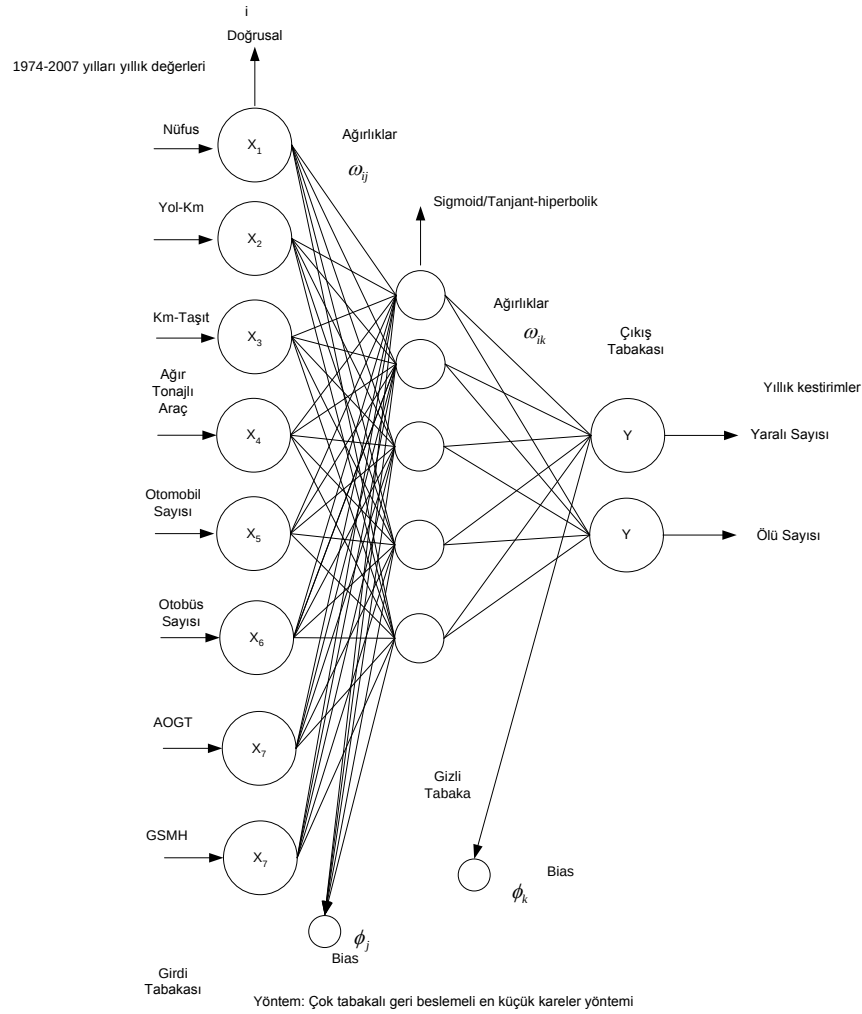
Hareketli ortalama modeli olan ARIMA(0,1,1) karakterisitk kökü birim çember içinde kalmıştır (Şekil 3.20).

3.1.4. Model II-III için yapay sinir ağı uygulaması

Modellerden Model-II ve Model-III'ün yapay sinir ağı uygulaması beraber yapılmıştır. İlk önce 8 bağımsız değişken Şekil 3.22'deki gibi giriş vektörünü oluşturmuştur. Giriş-gizli ve çıkış tabakası olmak üzere toplam üç tabakadan oluşan bir ağ mimarisi üzerinde çıkış tabakasında 2 nöron bulunmaktadır. Ağın %60'ı eğitim seti, %40'ıda test seti olmak üzere sınıflandırılmıştır. Çıkış nöronları ölü sayısı ve yaralı sayısıdır. Ağın eğitim setinde $R^2=0.95$ olarak tespit edilmiştir. Aktivasyon fonksiyonu trainbr olarak alınmıştır. Transfer fonksiyonları, logsig, purelin, tansig olarak tespit edilmiştir.

Giriş nöronları arasında yüksek korelasyon olduğundan 8 bağımsız değişken üzerinde TBA uygulaması yapılmış ve eğitim setine %2'den daha az katkısı olan elemanların elendiği yeni giriş vektörleri oluşturulmuştur. Birbirleri ile ilişkili 8 nöron yerine

ilişkisz 2 nöron giriş nöronlarını meydana getirmiştir. Gizli tabakada da 2 nöron kabul edilmiş çıktı tabakasında ise ölü ve yaralı sayılarının olduğu 2 nöron alınmış Şekil 4.17’de ağ yapısı sunulmuştur. Ağın eğitim setinde ölü sayıları için $R^2= 0.95$, yaralı sayıları içinde $R^2= 0.98$ olarak bulunmuştur. Eğitim seti de ölü sayısı için $OKH= 0.001$, yaralı sayısı için ise $OKH= 0.002$ olarak bulunmuştur.



Şekil 3.21. Model II-III genel YSA ağ mimarisi

3.1.5. Model-IV

Bu modelde; yıllık trafik kazalarındaki yaralı sayılarının, kaza sayısı, AOGT, GSMH değerlerine göre modellenmesi yapılmıştır.

Y: Yıllık Trafik Kazalarındaki Yaralı Sayıları

X_1 : Kaza Sayısı

X_2 : AOGT

X_3 : GSMH

$$k_i = \log(X_i)$$

$$t_i = \log(Y_i)$$

Çizelge 3.49. Bağımsız değişken için normallik test sonucu

Değişken	Kolmogorov-Smirnov			Kolmogorov-Smirnov (Logaritmik)		
	İstatistik	sd	Olabilirlik	İstatistik	sd	Olabilirlik
Y	0.819	34	0.001	1.275	34	0.078
X_1	1.289	34	0.007	0.864	34	0.444
X_2	0.534	34	0.013	0.632	34	0.819
X_3	0.827	34	0.045	0.730	34	0.661

Logaritmik dönüşümden sonra yapılan Kolmogorov-Smirnov test sonuçları Çizelge 3.49'da sunulmuştur.

H_0 : Değişkenler normal dağılım gösterir.

H_1 : Değişkenler normal dağılım göstermez.

Çizelge 3.50. Korelasyon matrisi

Değişkenler	t	k ₁	k ₂	k ₃
t	1			
k ₁	0.915	1		
k ₂	0.950	0.959	1	
k ₃	0.831	0.903	0.868	1

Bağımlı ve bağımsız değişkenlerimiz arasındaki korelasyon matrisinde bağımsız değişkenler arasında yüksek korelasyon göze çarpmaktadır (Çizelge 3.50).

Çizelge 3.51. Temel bileşenler küresellik testi sonuçları

Keiser-Meyer-Olkin	0.729	İyi
Barlett's	131.270	
Sd	3	
Olabilirlik	0.000	

Çizelge 3.52. Temel bileşenler analizi Özvektörler

Değişkenler	Vektör-1	Vektör-2	Vektör-3
k ₁	-0.586107	-0.255629	-0.768851
k ₂	-0.579017	-0.531629	0.618151
k ₃	-0.566761	0.807481	0.163578

Küresellik analizi sonucunda veriler TBA uygulamsına uygun olduğu sonucuna varılmıştır (Çizelge 3.51). Temel bileşenler analizi sonucunda bulunan özvektörlert Çizelge 3.52'de, özdeğerler ve bunların toplam varyansı açıklama oranı Çizelge 3.53'te sunulmuştur.

Çizelge 3.53. Temel bileşenlerin bulunması ve toplam varyansı açıklama oranları

Temel Bileşenler	Başlangıç Özdeğerler					
	Toplam	%Varyans	%Kümülatif	Toplam	%Varyans	%Küm.
1	2.821	94.017	94.017	2.821	94.017	94.017
2	0.143	4.755	98.772			
3	0.037	1.228	100.000			

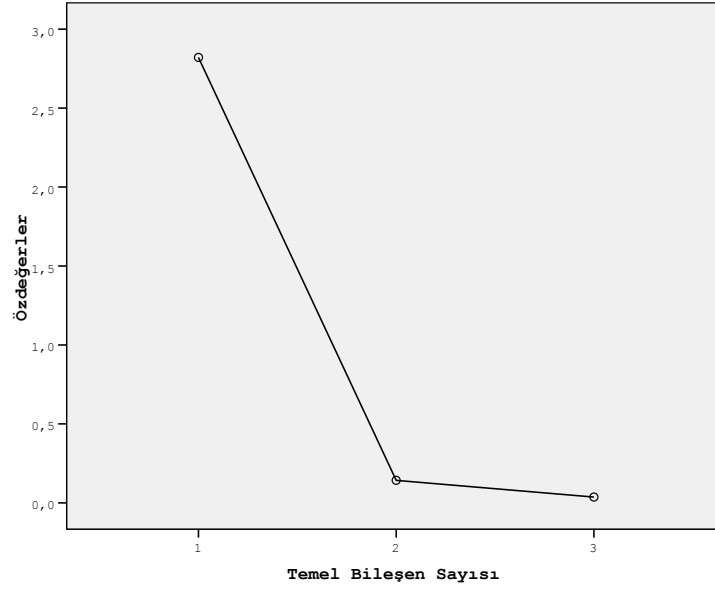
Özdeğerler;

$$\lambda_1 = 2.821, \lambda_2 = 0.143, \lambda_3 = 0.037$$

olarak bulunmuştur.

$$\sum_{j=1}^m \frac{\lambda_j}{p} = \frac{2.821 + 0.143}{3} = 0.988 > \frac{2}{3}$$

olduğundan 1 adet temel bileşenimiz vardır. 3 Bağımsız değişken yerine 1 temel bileşen ile devam edilecektir. Çizelge 3.53 incelendiğinde birinci temel bileşen toplam varyansın % 94.017'sini açıklarken, ikinci temel bileşen % 4.755'ini açıklamaktadır. Yamaç eğimi grafiği incelendiğinde 1 temel bileşenle modelin açıklanacağı görülebilmektedir (Şekil 3.22).



Şekil 3.22. Temel bileşen sayısı için yamaç eğimi (çizgi) grafiği

Çizelge 3.54. Değişkenlerin temel bileşenler tarafından açıklanma oranı

Değişkenler	Oran
k_1 (kaza sayısı)	0.969
k_2 (AOGT)	0.946
k_3 (GSMH)	0.906

Değişkenlerin temel bileşenler tarafından açıklanma oranı Çizelge 3.54'te verilmiştir. Temel bileşenler tarafından en çok açıklanan bileşenimiz %96.9 ile kaza sayısıdır.

Çizelge 3.55. Temel bileşenler matrisi

Değişkenler	Temel Bileşenler
	1
k_1 (kaza sayısı)	0.984
k_2 (AOGT)	0.972
k_3 (GSMH)	0.952

Temel bileşenler matrisine, % 98.4 ile kaza sayısı, %97.2 ile AOGT, %95.2 ile GSMH etkili olmuştur (Çizelge 3.55).

Temel bileşenler analizi sonucunda birbirleri ilişkili 3 bağımsız değişken yerine 1 adet temel bileşen ile modelimiz oluşturuldu. Bağımlı değişken ile temel bileşen faktörü arasındaki korelasyon matrisi Çizelge 3.56’da verilmiştir. Bu model ile ilgili istatistiksel sonuçlar Çizelge 3.57 ve Çizelge 3.58’de sunulmuştur.

Çizelge 3.56. Bağımsız değişken ve temel bileşen skorları ile korelasyon matrisi

Değişkenler	Y	Z ₁
Y	1	
Z ₁	0.927	1

Çizelge 3.57. Regresyon modeli çıktıları

Model	R	R ²	Düzeltilmiş R ²
1	0.927	0.860	0.856

Çizelge 3.58. Bağımsız değişken ve temel bileşen skorları ile regresyon modeli parametreleri

Değişkenler	B (Katsayı)	t	Olab. P	Tolerans	VIF	F	Olab. P
Sabit	4.834	293.902	0.000	1.00	1.00	196.699	0.000
SZ ₁	0.234	14.025	0.000	1.00	1.00		

Çizelge 3.59. İkinci temel bileşen için bulunan katsayılar ve en küçük kareler kestiricileri

Değişken	1. Temel Bileşen İçin Denklem	
	$\bar{b}_i$	b_i
	0.323531	0.185939
	0.319617	0.365186
	0.312852	0.387412
b_0	1.356429	
R^2	0.856	

TBA sonucunda ters dönüşüm yapılmış ve modelin katsayıları Çizelge 3.59’da sunulmuştur.

Çizelge 3.60. Çok değişkenli zaman serisi analizi için birim kök testi sonuçları.

Değişkenler	ADF t istatistik	%5 Tablo Değeri	P	ADF t istatistik	%5 olabilirlik değeri	P
	Fark Serisi Alınmadan Önce			1. Fark Serisi Alındıktan Sonra		
t (yaralı say.)	-1.40	-3.55	0.84	-3.97	-3.55	0.02
k_1 (kaza sayısı)	-2.73	-3.56	0.23	-3.57	-3.55	0.05
k_2 (AOGT)	-1.19	-3.60	0.88	-4.79	-3.56	0.003
k_3 (GSMH)	-3.65	-3.60	0.05			

Mutlak değer olarak $t_{ADF} > t_{\%5}$ ise seri durağandır. Değişkenlerimiz birinci fark serisi ve ikinci fark serisi alınarak durağan hale getirilmiştir (Çizelge 3.60).

En uygun gecikme sayısının tespiti için gerekli olan kriterler ve sonuçları Çizelge 3.61’de verilmiştir. Bu durumda nihai tahmin hatası, Akaike bilgi kriteri, Schwarz bilgi kriteri, Hannan Quin bilgi kriterlerine bakılarak minimum olan değerler karşılaştırılmış olup bütün bu kriterlerden 1. gecikmeye ait değerler minimum olduğundan VAR(p) modelinin derecesi 1 olarak alınmıştır (Çizelge 3.61).

Çizelge 3.61. Model derecesi belirleme kriterleri

Gecikme Sayısı	FPE (NTH)	AIC	SC	HQ
0	8.89e-11	-11.79	-11.61	-11.73
1	1.26e-11*	-13.76*	-12.83*	-13.46*
2	1.91e-11	-13.40	-11.74	-12.86

Model derecesi belirlendikten sonra değişkenler arasındaki Granger Nedensellik analiz sonuçları Çizelge 3.62’de sunulmuştur.

$F_{\text{hesaplanan}} > F_{\text{tablo}}$ ise bir nedensellik var demektir.

$H_0 : a_i = 0 (Y \rightarrow X)$ Y’den X’e nedensellik ilişkisi yoktur.

$H_1 : a_i \neq 0 (Y \rightarrow X)$ Y’den X’e nedensellik ilişkisi vardır.

Çizelge 3.62. Model Granger Nedensellik analiz sonuçları

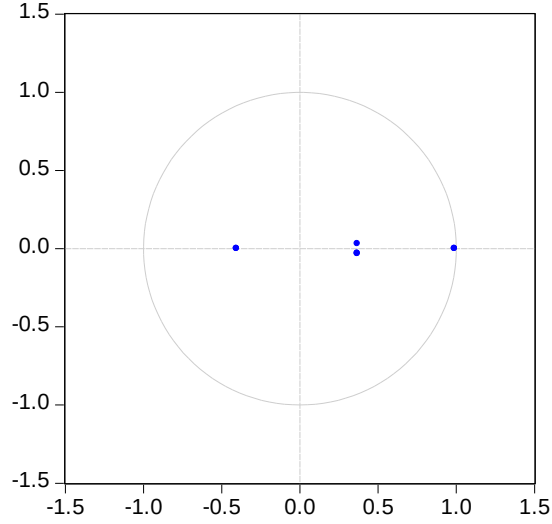
Değişkenler	F-İstatistik	Olabilirlik
Kaza Sayısı $\Downarrow$ Yaralı Sayısı	0.00018	0.9893
Yaralı Sayısı $\Rightarrow$ Kaza Sayısı	7.60745	0.0098
AOGT $\Downarrow$ Yaralı Sayısı	0.23826	0.6290
Yaralı Sayısı $\Rightarrow$ AOGT	2.08297	0.0159
Yaralı Sayısı $\Downarrow$ GSMH	0.08772	0.7691
GSMH $\Downarrow$ Yaralı Sayısı	1.25893	0.2708

Hata terimleri normal dağılıma uygunlukları test edildi (Çizelge 3.63).

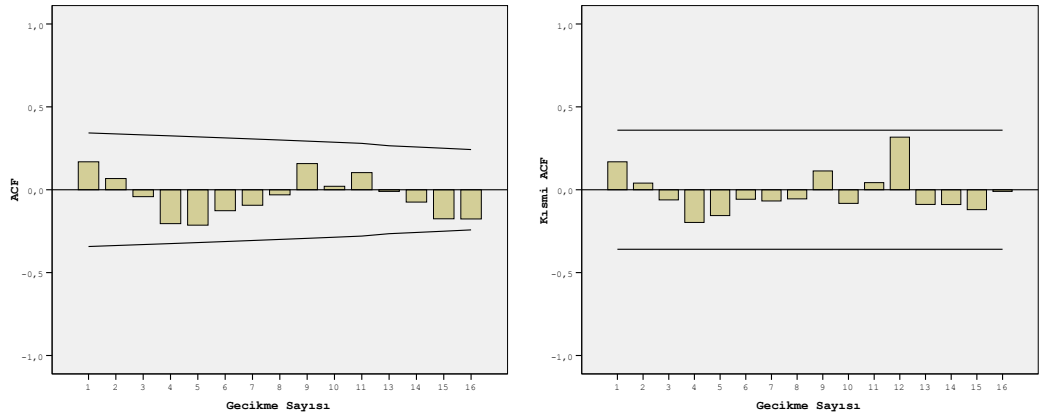
Çizelge 3.63. Hata terimleri normallik analizi

	Kolmogorov-Smirnov		
	İstatistik	Sd	Olabilirlik
ε_i Hata Terimi	0.512	31	0.956

Olabilirlik $P > 0.05$ olduğundan seri normal dağılımlıdır (Çizelge 3.63). Karakteristik köklerin birim çember içinde kaldığından seride birim kök olmadığı anlaşıldı (Şekil 3.23).

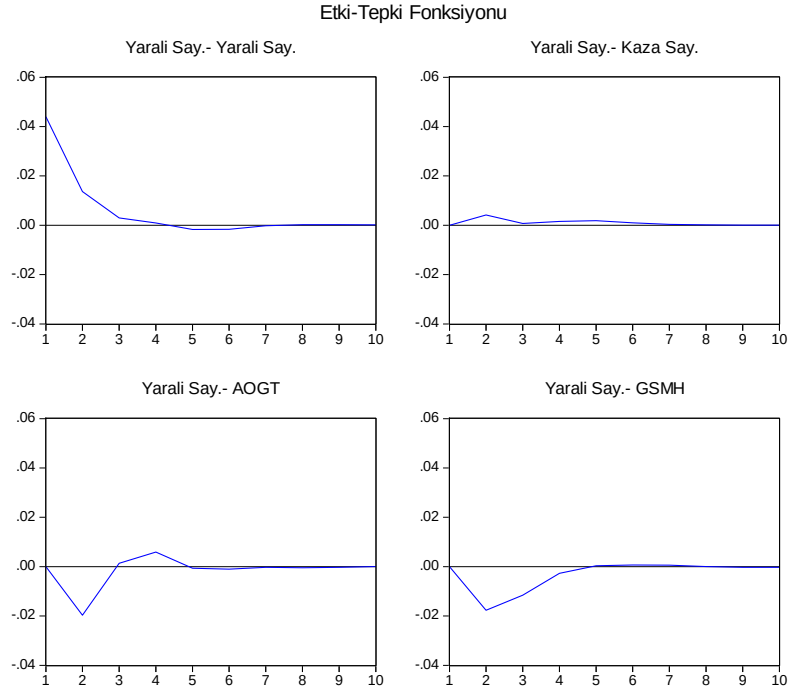


Şekil 3.23. VAR(1) Modeli için karakteristik köklerin gösterimi



Şekil 3.24. VAR(1) Modeli için hata terimleri ACF-PACF grafikleri

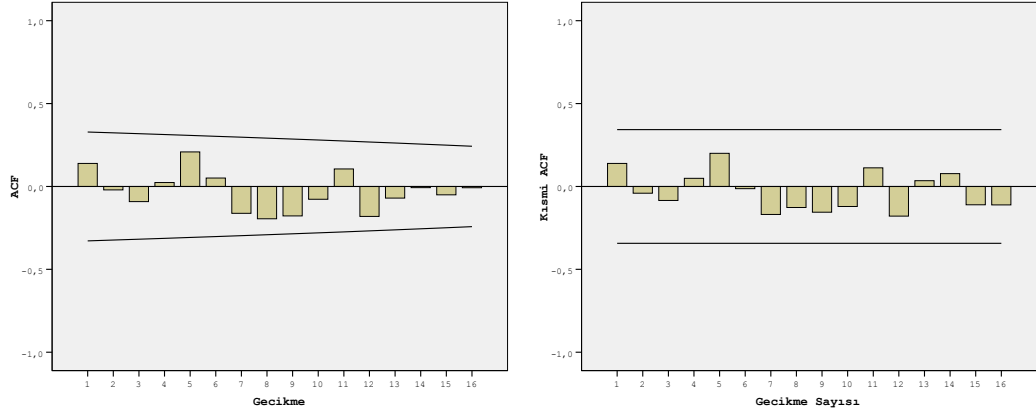
VAR(1) modeline ilişkin hata terimleri ACF ve PACF grafikleri çizilerek modelin hata terimlerinin güven aralıkları içinde kaldığı görüldü. Hata terimleri otokorelasyon olmadığı ve akgürültü bir seri oldukları anlaşıldı (Şekil 3.24).



Şekil 3.25. VAR(1) Modeli için Etki-Tepki Fonksiyonu

Çizelge 3.64. VAR(1) Modeli varyans ayrıştırması

Peryot	Ölü Sayısı	Kaza Sayısı	AOGT	GSMH
1	100.0000	0.000000	0.000000	0.000000
2	74.74605	0.611930	13.61899	11.02303
3	71.41362	0.600674	13.01725	14.96845
4	70.38023	0.667310	13.96429	14.98817
5	70.31579	0.781184	13.94552	14.95751
6	70.28618	0.810303	13.95477	14.94875
7	70.27411	0.814320	13.95464	14.95693
8	70.26859	0.815297	13.96063	14.95548
9	70.26597	0.815400	13.96209	14.95654
10	70.26468	0.815414	13.96170	14.95821



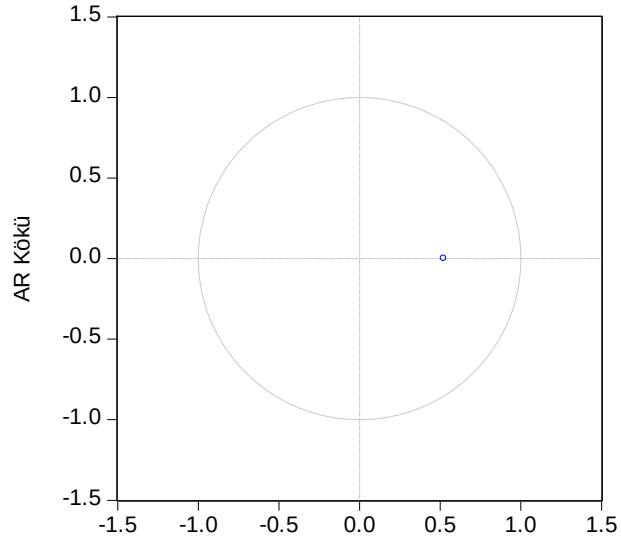
Şekil 3.26. ARIMA (1,1,0) Hata terimleri ACF ve PACF grafikleri

ARIMA(1,1,0) modeline ilişkin hata terimleri ACF ve PACF grafikleri çizilerek modelin hata terimlerinin güven aralıkları içinde kaldığı görüldü. Hata terimleri otokorelasyon olmadığı ve akgürültü bir seri oldukları anlaşıldı (Şekil 3.26).

Çizelge 3.65. ARIMA(1,1,0) Modeli normallik testi

Değişken	Kolmogorov-Smirnov		
	İstatistik	Sd	Olabilirlik
Model ARIMA(1,1,0)	0.992	34	0.278

Olabilirlik $P > 0.05$ olduğundan seri normal dağılımlıdır (Çizelge 3.65).



Şekil 3.27. ARIMA (1,1,0) Model için karakteristik köklerin gösterimi

Karakteristik köklerin birim çember içinde kaldığı görüldü. Modelin birim kök taşımadığı anlaşıldı (Şekil 3.27).

3.1.6. Model-V

Model-V’de yıllık trafik kazalarındaki ölü sayıları, kaza sayısı, AOGT ve GSMH değerlerine göre modellenmiştir.

Y: Yıllık Trafik Kazalarındaki Ölü Sayıları

X_1 : Kaza Sayısı

X_2 : AOGT

X_3 : GSMH

$$k_i = \log(X_i)$$

$$t_i = \log(Y_i)$$

Çizelge 3.66. Bağımsız değişken için normallik test sonucu

Değişken	Kolmogorov-Smirnov			Kolmogorov-Smirnov (Logaritmik)		
	İstatistik	sd	Olabilirlik	İstatistik	sd	Olabilirlik
Y	0.095	34	0.200	0.123	34	0.200

Olabilirlik $P > 0.05$ olduğundan seri normal dağılımlıdır (Çizelge 3.66).

H_0 : Değişkenler normal dağılım gösterir.

H_1 : Değişkenler normal dağılım göstermez.

Çizelge 3.67. Korelasyon matrisi.

Değişkenler	t	k ₁	k ₂	k ₃
t	1			
k ₁	-0.528	1		
k ₂	-0.407	0.959	1	
k ₃	-0.472	0.903	0.868	1

Bağımlı ve bağımsız değişkenlerimiz arasındaki korelasyon matrisinde Çizelge 3.67’de bağımsız değişkenler arasında yüksek korelasyon göze çarpmaktadır. Model IV’de kullanmış olduğumuz bağımsız değişkenler, Model V’de farklılık göstermediğinden, TBA uygulamalarındaki bazı tablolarda değişiklik yoktur. Çizelge 3.51, 3.52, 3.53, 3.54 ve 3.55 ile aynıdır. Aynı zamanda Şekil 3.22’deki yamaç eğimi grafiğide aynıdır. Model IV uygulamasında 3 bağımsız değişken yerine bir temel bileşenle model açıklanmaya çalışılacaktır.

Temel bileşenler analizi sonucunda birbirleri ilişkili 3 bağımsız değişken yerine 1 adet temel bileşen ile model oluşturuldu. Bu model ile ilgili istatistiksel sonuçlar Çizelge 3.69 ve Çizelge 3.70’te sunulmuştur.

Çizelge 3.68. Bağımsız değişken ve temel bileşen skorları ile korelasyon matrisi

Değişkenler	Y	Z ₁
Y	1	
Z ₁	-0.484	1

Model-V için bağımlı değişken ve TBA skoru arasındaki korelasyon negatif bulunmuştur (Çizelge 3.68).

Çizelge 3.69. Regresyon modeli çıktıları

Model	R	R ²	Düzeltilmiş R ²
1	0.484	0.432	0.421

Çizelge 3.70. Bağımsız değişken ve temel bileşen skorları ile regresyon modeli parametreleri

Değişkenler	B (Katsayı)	t	Olab. P	Tolerans	VIF	F	Olab. P
Sabit	3.690	197.505	0.000	1.00	1.00	9.784	0.000
SZ ₁	-0.059	-3.128	0.051	1.00	1.00		

Model-V'te belirlilik katsayısı düzeltilmiş R² oldukça küçük bulunmuştur (Çizelge 3.69). Modelin diğer parametreleri ise istatistiksel olarak anlamlı bulunmuştur (Çizelge 3.70).

Çizelge 3.71. Temel bileşen için bulunan katsayılar ve en küçük kareler kestiricileri

Değişken	Temel Bileşen İçin Denklem	
	$\bar{b}_i$	b_i
	-0.16880	-0.0471
	-0.16676	-0.0925
	-0.16323	-0.0981
b_0	4.570631	
R^2	0.48	

Çok değişkenli zaman serisi analizinde logaritmik dönüşüm ile normal dağılıma uygun olan bağımlı değişkenimize birim kök testlerinden Genişletilmiş Dickey Fuller testi uygulanmış ve Çizelge 3.71'deki sonuçlar elde edilmiştir.

Çizelge 3.72. Çok değişkenli zaman serisi analizi için birim kök testi sonuçları

Değişkenler	ADF t istatistik	%5 Tablo Değeri	Olabilirlik	ADF t istatistik	%5 olabilirlik değeri	Olabilirlik
	Fark Serisi Alınmadan Önce			1. Fark Serisi Alındıktan Sonra		
t	-1.88	-3.56	0.63	-3.83	-3.56	0.03

Fark serisi alınarak duğanlaşan serilen ADF test sonuçları Çizelge 3.72'de sunulmuştur. En uygun gecikme sayısının tespiti için gerekli olan kriterler ve sonuçları Çizelge 3.73'de verilmiştir. Bu durumda nihayi tahmin hatası, Akaike bilgi kriteri, Schwarz bilgi kriteri, Hannan Quin bilgi kriterlerine bakılarak minimum olan değerler karşılaştırılmış olup bütün bu kriterlerden 1. gecikmeye ait değerler minimum olduğundan VAR(p) modelinin derecesi 1 olarak alınmıştır.

Çizelge 3.73. Model derecesi belirleme kriterleri

Gecikme Sayısı	FPE (NTH)	AIC	SC	HQ
0	8.35e-11	-11.85435	-11.66932	-11.79404
1	1.00e-11*	-13.98772*	-13.06257*	-13.68614*
2	1.82e-11	-13.44685	-11.78158	-12.90402

Model derecesi belirlendikten sonra değişkenler arasında ki Granger Nedensellik analiz sonuçları Çizelge 3.74'te sunulmuştur.

$F_{\text{hesaplanan}} > F_{\text{tablo}}$ ise bir nedensellik var demektir.

$H_0 : a_i = 0 (Y \rightarrow X)$ Y'den X'e nedensellik ilişkisi yoktur

$H_1 : a_i \neq 0 (Y \rightarrow X)$ Y'den X'e nedensellik ilişkisi vardır

VAR(1) modeline ilişkin uygun gecikme sayısı tespit edilmiştir. Gecikme sayısına bağlı olarak Granger Nedensellikleri bulunmuştur (Çizelge 3.74).

Çizelge 3.74. Model Granger Nedensellik analiz sonuçları

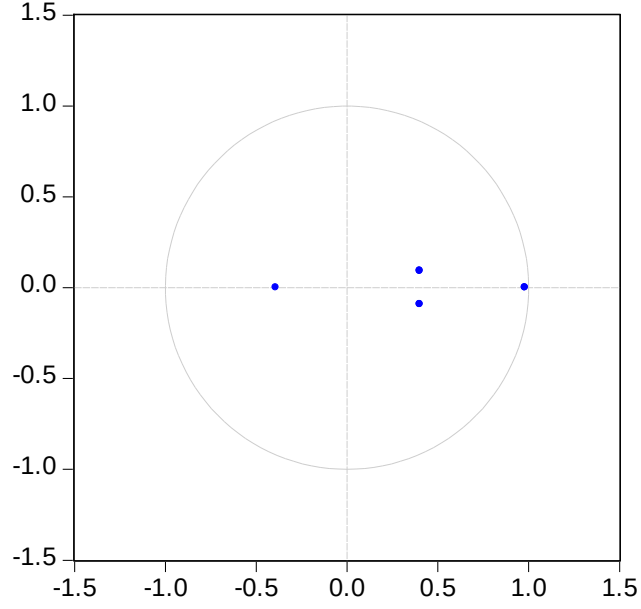
Değişkenler	F-İstatistik	Olabilirlik
Kaza Sayısı $\Rightarrow$ Ölü Sayısı	6.84949	0.0138
Ölü Sayısı $\Rightarrow$ Kaza Sayısı	3.69149	0.0542
AOGT $\Rightarrow$ Ölü Sayısı	6.00818	0.0203
Ölü Sayısı $\Updownarrow$ AOGT	0.64804	0.4271
GSMH $\Updownarrow$ Ölü Sayısı	2.77692	0.0226
Ölü Sayısı $\Updownarrow$ GSMH	1.06469	0.3104

Çizelge 3.75. Hata terimleri normallik analizi

	Kolmogorov-Smirnov		
	İstatistik	Sd	Olabilirlik
ε_i Hata Terimi	0.113	32	0.200

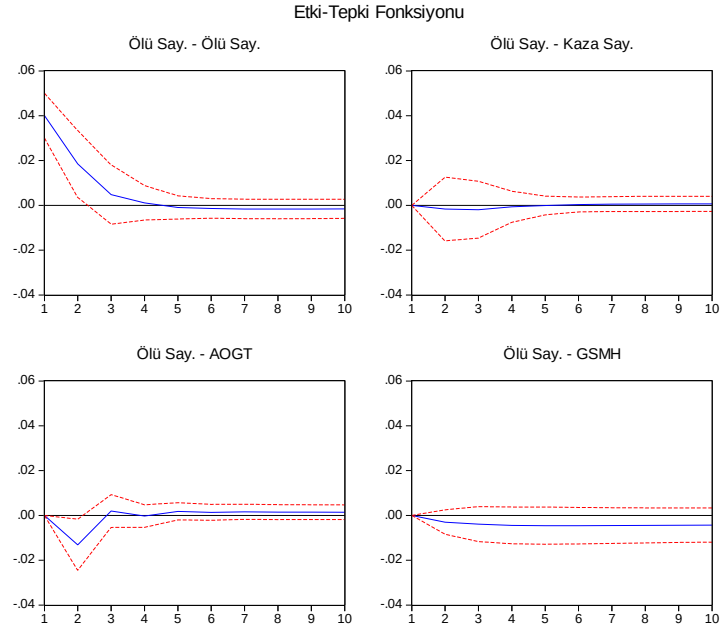
Olabilirlik $P > 0.05$ olduğundan seri normaldir (Çizelge 3.75)

Karakteristik köklerin birim çember içinde kalıp kalmadığı analiz edilmiştir (Şekil 3.28).



Şekil 3.28. VAR(1) Modeli için karakteristik köklerin gösterimi

Etki tepki analizi yapılmıştır. Ölü sayılarının hata terimleri varyansına bir birimlik bir şok kaza sayısı, AOGT ve GSMH değişkenleri 4. dönemde etkisini kaybetmektedir. O halde değişkenlerimiz için 4. dönemde şokların etkisi azalmaktadır (Şekil 3.29).

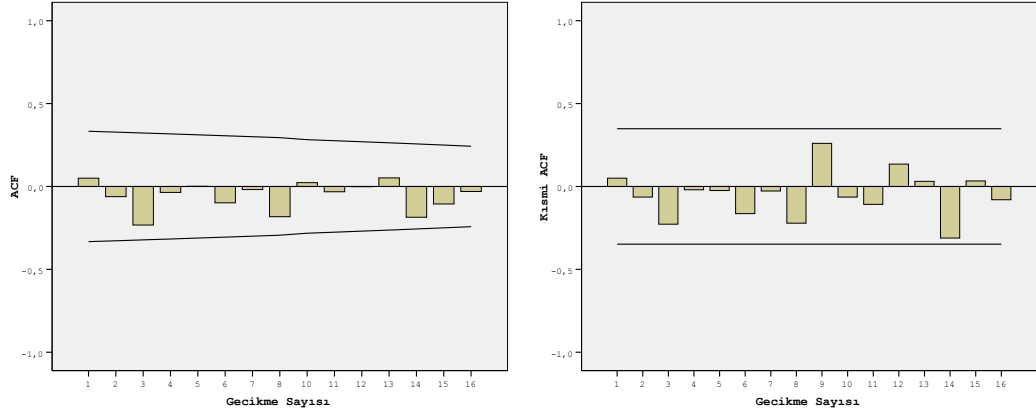


Şekil 3.29. VAR(1) Modeli için Etki-Tepki Fonksiyonu

Çizelge 3.76. VAR(1) Modeli varyans ayrıştırması

Peryot	Ölü Sayısı	Kaza Sayısı	AOGT	GSMH
1	100.0000	0.000000	0.000000	0.000000
2	91.39659	0.127956	8.055317	0.420137
3	90.52961	0.298917	8.060754	1.110720
4	89.69245	0.316083	7.985256	2.006214
5	88.71725	0.312777	8.043403	2.926572
6	87.81612	0.315077	8.033811	3.834989
7	86.92586	0.324003	8.052471	4.697663
8	86.09126	0.337332	8.057558	5.513854
9	85.30076	0.351883	8.065371	6.281986
10	84.55749	0.366567	8.070524	7.005419

Model hata terimlerinin ACF ve PACf grafikleri çizilmiş gecikmelerin güven aralıklarının içinde kaldığı görülmüştür (Şekil 3.30).

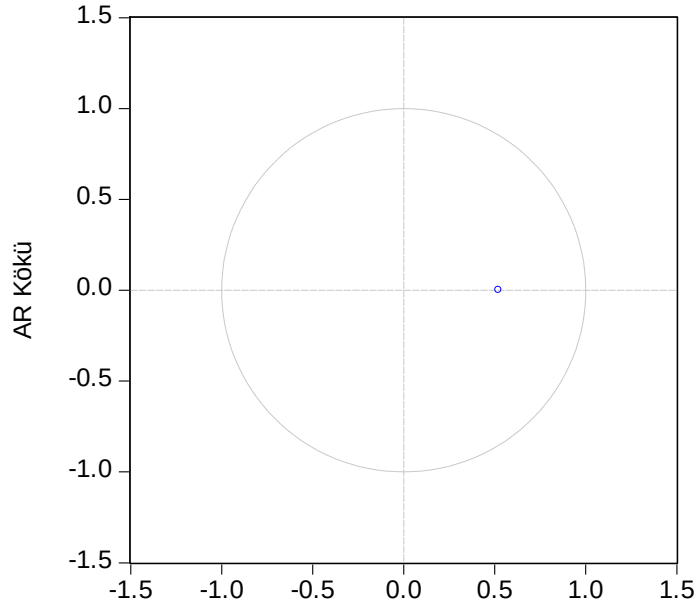


Şekil 3.30. ARIMA(1,1,0) modeli için hata terimleri ACF PACF Grafikleri

Çizelge 3.77. ARIMA(1,1,0) modeli hata terimleri normallik testi

Değişken	Kolmogorov-Smirnov		
	İstatistik	Sd	Olabilirlik
Model ARIMA(1,1,0)	0.118	33	0.155

Modelin hata terimlerinin normallik analizi yapılmıştır. Hata terimlerinin normal dağılıma uygun olduğu görülmüştür (Çizelge 3.76). ARIMA(1,1,0) modelinin karakteristik kökleri birim çember içinde kalmıştır ve model, birim kök içermemektedir (Şekil 3.31).

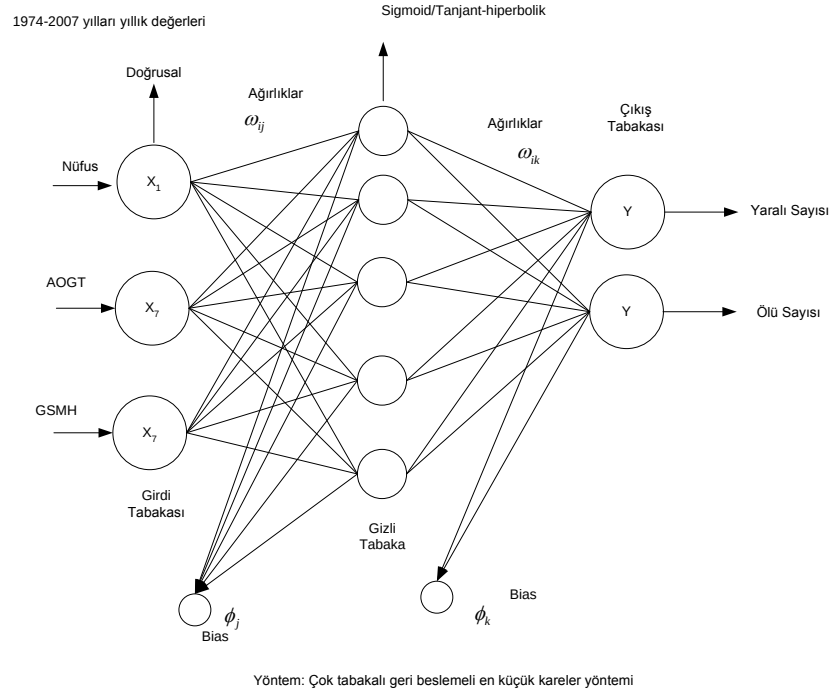


Şekil 3.31. ARIMA (1,1,0) Model için karakteristik köklerin gösterimi

3.1.7. Model IV-V için yapay sinir ağı uygulaması

Modellerden Model-IV ve Model-V'in yapay sinir ağı uygulaması beraber yapılmıştır. İlk önce 3 bağımsız değişken Şekil 3.33'deki gibi giriş vektörünü oluşturmuştur. Giriş-gizli ve çıkış tabakası olmak üzere toplam üç tabakadan oluşan bir ağ mimarisi üzerinde çıkış tabakasında 2 nöron bulunmaktadır. Ağın %60'ı eğitim seti, %40'ında test seti olmak üzere sınıflandırılmıştır. Çıkış nöronları ölü sayısı ve yaralı sayısıdır. Ağın eğitim setinde $R^2 = 0.98$ olarak tespit edilmiştir. Aktivasyon fonksiyonu trainbr olarak alınmıştır. Transfer fonksiyonları, tanjant-sigmoid olarak tespit edilmiştir.

Gizli tabakada da 2 nöron kabul edilmiş çıktı tabakasında ise ölü ve yaralı sayılarının olduğu 2 nöron alınmış ve şekil 4.28'de sunulmuştur. Ağın eğitim setinde ölü sayıları için $R^2 = 0.95$, yaralı sayıları içinde $R^2 = 0.98$ olarak bulunmuştur. Eğitim setinde ölü sayısı için OKH= 0.001, yaralı sayısı içinde OKH= 0.002 olarak bulunmuştur.



Şekil 3.32. Model-IV ve Model-V’te kullanılan bağımlı ve bağımsız değişkenlere ilişkin örnek YSA ağ mimarisi

3.2. Poisson Regresyonu Çok Değişkenli Regresyon ve Yapay Sinir Ağı Modelleri

Uygulamada, (Mercan – Tercan – Aşkale – Erzurum – Pasinler – Horasan – Eleşkirt – Ağrı - Doğubayazıt) güzergahında bulunan 422 km uzunluğundaki E-80 (D – 100) Karayolu (100-27), (100-28), (100-29), (100-30), (100-31), (100-32), (100-33) kesimlerine ait kaza istatistikleri ve yolun geometrik özelliklerine ait istatistikler kullanılmıştır. Veriler Erzurum Karayolları 12. bölge Müdürlüğü’nden elde edilmiştir. Poisson Regresyonu, Çok Değişkenli İstatistik ve Yapay Sinir Ağları yöntem olarak kullanılacaktır.



Şekil 3.33. Model VI-VII uygulama alanı

3.2.1. Model – VI

Model-VI’da kullanılan parametreler aşağıda sunulmuştur.

Y_1 :Toplam Kaza Sayıları

X_1 :Yıllık Ortalama Günlük Trafik YOGT)

X_2 :Kavşak Sayısı

X_3 :Şerit Sayısı

X_4 :Şerit Genişliği

X_5 :Uzunluk km

X_6 :Ortalama Hız

X_7 :Bağlantı Sayısı

Poisson veya Negatif Binom hata yapılarından hangisinin kullanılacağına karar şu yöntemle belirlenmiştir. İlk olarak model parametreleri Poisson regresyonu ile tahmin edilmiş, sonra, aşırı yayılım göstergesi hesaplanmıştır.

Aşırı Yayılım Testi:

$$Pearson\chi^2 = \sum_{i=1}^n \frac{(y_i - \mu_i)^2}{Var(Y_i)}$$

Dağılım parametresi α önemli derecede 0'dan büyük ise dağılım aşırı yayılım göstermiş ve Negatif Binomial Regresyon uygun yöntemdir denir. Dağılım parametresi uygulamamızda $\alpha = 0,5340817$ bulunmuştur.

Dağılım parametresinin testi içinde;

$LR = -2(\log olabilirlik(poisson) - \log olabilirlik(negatifbinomial)) > \chi^2_{(1-2d,1sd)}$ ise $\alpha = 0$ hipotezi ret edilir ve negatif binomial regresyonunu uygun yöntem olduğu anlaşılır.

Uyum iyiliği değerlendirmesi için

$$0,8 < \frac{Pearson\chi^2}{n-p} < 1,2$$

Değerine bakılması uygundur ve bu aralıkta 0,8-1,2 aralığında olması beklenmektedir. Uygulamada uyum iyiliği değeri 0.66 bulunmuştur. Modele giren bağımsız değişkenlerin katsayılarının %95 güven aralığı içinde kaldıkları gözlemlenmiştir. Değişkenlere ait olabilirlik değeri “P” değeri 0,05’den küçüktür. Yapılan çalışmada denenen birçok modelden en iyi sonucu veren model parametreleri Çizelge 3.77’de sunulmuştur.

Çizelge 3.77. Poisson regresyonu model parametreleri

n=120 sd=n-p=120-6=114 Pearson $\chi^2 = 115,7460982$ $\chi^2_{0.05,113} = 124,342$				
Parametre	Katsayı	Standart Hata	z oranı	P
loguz	1.2396	0.136	9.09	0,001
Yogt	0.0002328	0.00002	9.29	0,001
sg	0.1746494	0.0808	2.16	0,031
Hız	-0.011652	0.006	9.88	0,000
bs	0.0604	0.004	-2.86	0.004

Veriler için ikinci olarak çok değişkenli regresyon yöntemi kullanılmıştır. Bu aşamada modele giren bütün bağımsız değişkenler için korelasyon matrisi incelendi. Bağımsız değişkenler arasında yüksek korelasyon gözlemlenmiştir (Çizelge 3.78).

Çizelge 3.78. Korelasyon matrisi

Değişkenler	ks	uz	yogt	ss	kavs	sg	bs	hız
ks	1							
uz	0.253	1						
yogt	0.252	-0.025	1					
ss	0.072	-0.115	0.225	1				
kavs	0.259	0.146	-0.044	0.177	1			
sg	0.111	0.038	-0.030	0.269	0.236	1		
bs	0.389	-0.004	0.065	0.097	0.695	0.290	1	
hız	-0.045	-0.107	0.116	-0.035	-0.105	0.190	0.038	1

Oluşturulan model incelendiğinde şerit sayısı, kavşak sayısı, şerit genişliği ve hız değişkenlerinin katsayılarının istatistiksel olarak anlamlı olmadıkları görülmektedir. Modelin tamamını istatistiksel olarak test eden “F” testi anlamlı çıkmıştır (Çizelge 3.80). Buna paralel düşük tolerans, düşük varyans artış faktörü bulunmuştur. Anlamsız değişkenlerin modelden çıkarılarak yeniden modellenmesi yapılmıştır.

Yeni modeldeki değişkenlerimiz hız, bağlantı sayısı, uzunluk km, YOGT olmuştur. Bu değişkenlere ilişkin parametreler Çizelge 3.79’da sunulmuştur.

Çizelge 3.79. Regresyon modeli çıktıları

Model	R	R ²	Düzeltilmiş R ²
1	0.788	0.620	0.607

Model parametrelerinin tutarlığının analizi için, varyans tablosu ve katsayılar tablosu ile “F” istatistik değeri, “t” değeri, tolerans ve VIF değeri Çizelge 3.80’de sunulmuştur.

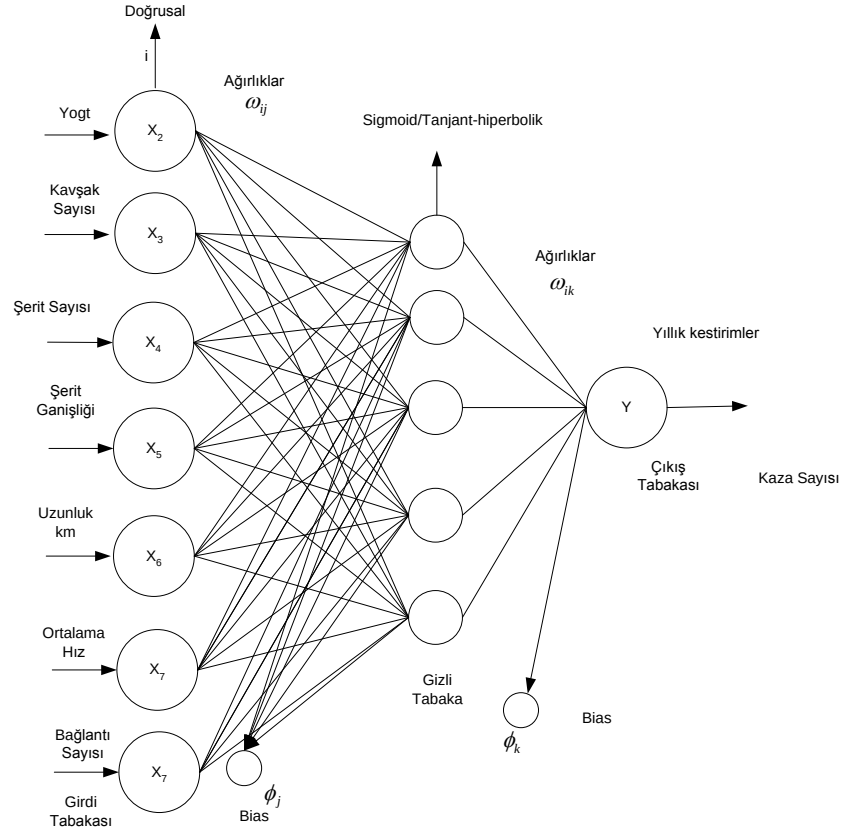
Çizelge 3.80. Regresyon modeli parametre çıktıları

Değişkenler	B (Katsayı)	T	Olab. P	Tolerans	VIF	F	Olab. P
uz	8.654	3.177	0.002	0.819	1.078	44.124	0.000
yogt	13.157	3.012	0.003	0.892	1.116		
bs	0.707	4.753	0.000	0.878	1.134		
hız	-24.866	-4.374	0.002	0.565	1.196		

En uygun YSA mimarilerine ilişkin kıstaslar Çizelge 3.81’de sunulmuştur.

Çizelge 3.81. Optimum YSA mimarisi

Ağ Yapısı	AIC (Akaike)	Ortalama Karesel Hata (MSE)	R ²	Aktivasyon Fonksiyonu	Transfer Fonksiyonu	İterasyon Sayısı	Tabaka Sayısı	Gizli Tabakadaki nöron sayısı
Yöntem 1	-239.894	0.008	0.96	Trainbr	Tansig tansig purelin	500	3	3
Yöntem 2	-222.869	0.008	0.96	Trainbr	Tansig tansig purelin	500	3	2



Şekil 3.34. Model-VI ve Model-VII için Optimum YSA mimarisi

3.2.2. Model – VII

Model-VII’de kullanılan parametreler aşağıda sunulmuştur.

Y_1 :Yaralı Sayısı

X_1 :Kaza Sayısı (ks)

X_2 :Yıllık Ortalama Günlük Trafik (YOGT)

X_3 :Kavşak Sayısı (kavs)

X_4 :Şerit Sayısı (şs)

X_5 :Şerit Genişliği (şg)

X_6 :Uzunluk km (uz)

X_7 :Ortalama Hız (hz)

X_8 :Bağlantı Sayısı (bs)

Bağımlı değişken olarak yaralı sayısı alınmıştır. Bütün bağımsız değişkenler modele katılmıştır. Bazı değişkenler istatistiksel olarak anlamlı çıkmadığından modelden çıkarılmıştır. Aşırı yayılım göstergesi hesaplanmıştır.

Aşırı Yayılım Testi:

Dağılım parametresi $\alpha = 0,2293242 < 1$ olduğundan dağılım aşırı yayılım göstermemiş olduğundan Negatif Binomial Regresyon uygun bir yöntem olmadığından Poisson regresyonu ile modellenmenin yapılacağına karar verilmiştir.

Uyum iyiliği değerlendirmesi için

Uygulamada uyum iyiliği değeri 5.85 çıkmıştır. Bu sonuç bir kez daha yöntem olarak Negatif Binomial Regresyon yönteminin uygun bir yöntem olmadığını ve yerine Poisson regresyonunun kullanılması gerektiğini göstermiştir.

Yapmış olduğumuz çalışmada denenen birçok modelden en iyi sonucu veren model parametreleri Çizelge 3.82’de verilmiştir.

Çizelge 3.82. Poisson regresyonu model parametreleri

n=120 sd=n-p=120-6=114 Pearson $\chi^2 = 759.445$ $\chi^2_{0.05,113} = 124,342$				
Parametre	Katsayı	Standart Hata	z oranı	P
Ks	0.0196034	0.002	8.25	0.000
Ss	0.04661516	0.017	2.60	0.009
Kavs	-0.0848338	0.022	-3.79	0.000
Bs	0.0512742	0.007	7.16	0.000
Hiz	-0.008345	0.004	-1.91	0.055
Sabit	3.207331	0.374	8.57	0.000

Km-uzunluk, şerit genişliği, değişkenleri istatistiksel olarak anlamlı çıkmadıklarından modelden çıkarılmışlardır.

Veriler için ikinci olarak çok değişkenli regresyon yöntemi kullanılmıştır. Bu aşamada modele giren bütün bağımsız değişkenler için korelasyon matrisi incelenmiştir. Bağımsız değişkenler arasında yüksek korelasyon gözlemlenmiştir (Çizelge 3.83).

Çizelge 3.83. Korelasyon matrisi

Değişkenler	ys	ks	uz	yogt	ss	kavs	sg	bs	hız
ys	1								
ks	0.430	1							
uz	-0.830	0.030	1						
yogt	0.011	0.222	-0.201	1					
ss	0.099	-0.111	-0.258	0.174	1				
kavs	0.113	0.146	0.038	0.251	0.160	1			
sg	0.166	0.032	-0.050	0.211	0.324	0.341	1		
bs	0.357	-0.004	-0.042	0.289	0.054	0.695	0.503	1	
hız	0.066	-0.107	-0.121	-0.101	-0.024	-0.105	0.208	0.038	1

Model parametreleri incelendiğinde şerit sayısı, kavşak sayısı, şerit genişliği, uzunluk-km, sabit terim ve hız değişkenlerinin katsayılarının istatistiksel olarak anlamlı olmadıkları görülmektedir. Modelin tamamını istatistiksel olarak test eden “F” testi anlamlı çıkmıştır. Buna paralel düşük tolerans, düşük varyans artış faktörü bulunmuştur. Anlamsız değişkenlerin modelden çıkarılarak yeniden modellenmesi yapılacaktır.

Yeni modeldeki değişkenler hız, bağlantı sayısı, uzunluk km ve YOGT olmuştur. Bu değişkenlerimize ilişkin belirlilik katsayısı Çizelge 3.84’te sunulmuştur.

Çizelge 3.84. Regresyon modeli çıktıları

Model	R	R ²	Düzeltilmiş R ²
1	0.86	0.77	0.76

Model parametrelerinin tutarlılığının analizi için, varyans tablosu ve katsayılar tablosu ile “F” istatistik değerleri, “t” değerleri, tolerans ve VIF değerleri Çizelge 3.84’de sunulmuştur.

Çizelge 3.85’te regresyon parametreleri, Çizelge 3.86’da ise optimum YSA ağ yapılarına ilişkin özellikler sunulmuştur.

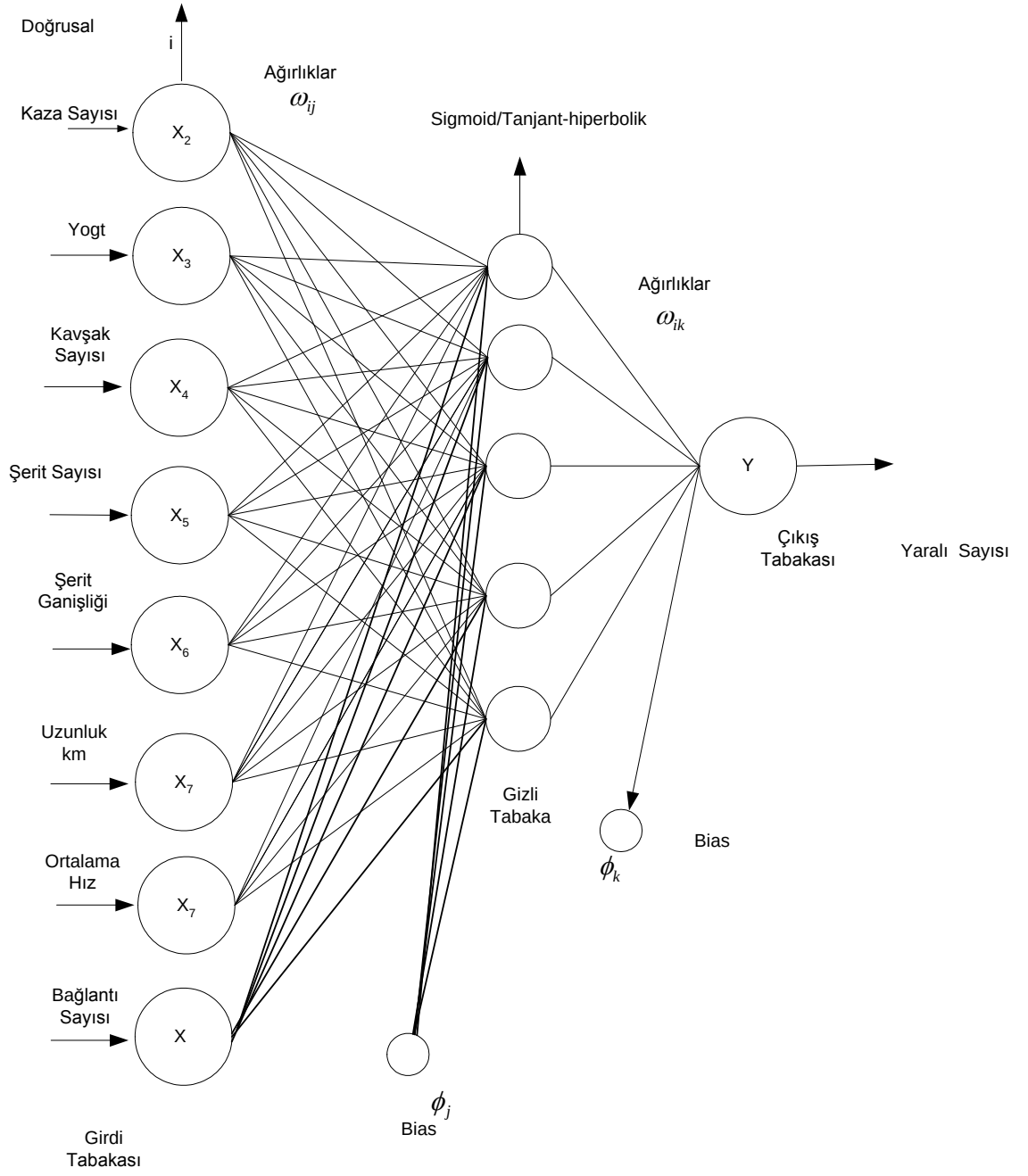
Çizelge 3.85. Regresyon modeli parametre çıktıları

Değişkenler	B (Katsayı)	T	Olab. P	F	Olab. P
Ks	0.271	4.620	0.000	694.949	0.000
Kavs	-0.062	-2.637	0.010		
Bs	0.031	4.048	0.000		
Hız	-0.504	20.700	0.000		

Çizelge 3.86. Optimum YSA ağ mimarisi

Ağ Yapısı	AIC (Akaike)	Ortalama Karesel Hata (MSE)	R ²	Aktivasyon Fonksiyonu	Transfer Fonksiyonu	İterasyon Sayısı	Tabaka Sayısı	Gizli Tabakadaki nöron saayısı
Yöntem 1	-327.563	0.18	0.95	Trainbr	Logsig purelin logsig	390	3	2
Yöntem 2	-310.862	0.28	0.92	Trainbr	Logsig purelin logsig	285	3	1

Şekil 3.36’da Model-VI ve Model-VII için optimum YSA ağ mimarisi verilmiştir.



Şekil 3.35 Optimum YSA mimarisi

4. ARAŞTIRMA BULGULARI ve TARTIŞ MA

4.1. Çok Değişkenli Regresyon, Çok Değişkenli Zaman Serisi Analizi, ARIMA (p,d,q) ve Yapay Sinir Ağı Model Bulguları

4.1.1. Model-I bulguları

Yıllık kaza sayılarının, yük-ton değeri, yolcu-km değeri, km-taşıt değeri, kamyon sayısı, tır sayısı, otobüs sayısı, karayollarındaki enerji tüketimi, yol-km uzunluğu ve AOGT değerlerine göre modellendi.

Yapılan bu çalışmada çok değişkenli regresyon denklemi uygulamasında, temel bileşenler analizi birbirleri ile ilişkili 9 bağımsız değişken yerine birbirleri ile ilişkisiz 2 temel bileşenle regresyon modeli oluşturulmuştur. Standartlaştırılmış temel bileşen skorları ile yapılan regresyon modeli aşağıda verilmiştir.

$$t = 5.180 + (0.388 * Z_1) - (0.196 * Z_2)$$

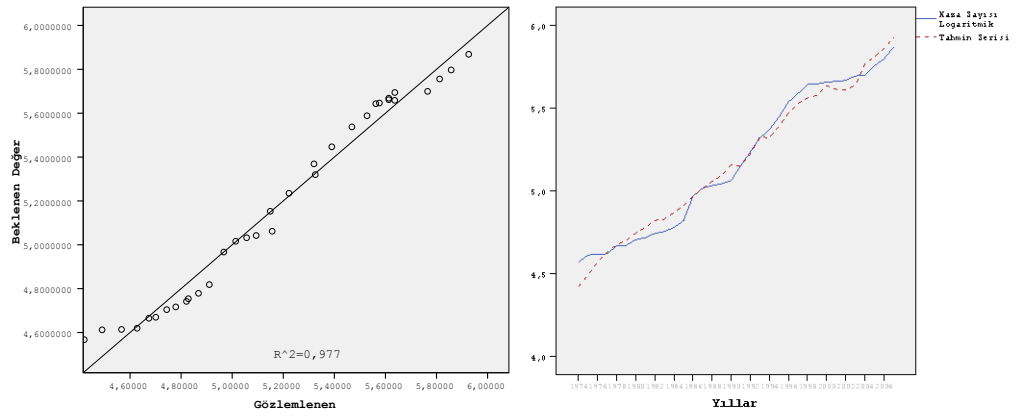
2008 yılı için kaza sayısı tahmini aşağıdaki gibi elde edilmiştir.

$$t = 5.180 + (0.388 * 1.22294) + (0.196 * 1.41213) = 5.92$$

En küçük kareler kestiricileri yardımıyla yapılan ters dönüşüm sonucu, orijinal verilerin kullanıldığı regresyon modeli aşağıdaki biçimde elde edilmiştir.

$$t = -21.8063 + (0.185252 * \text{ton} - \text{km}) + (0.322975 * \text{yolcu} - \text{km}) + (0.221942 * \text{km} - \text{arac}) \\ + (0.135824 * \text{kamyon}) + (0.226115 * \text{tir}) + (0.192535 * \text{otobüs}) + (0.289786 * \text{enerji}) \\ + (3.896268 * \text{yol} - \text{km}) + (0.242882 * \text{AOGT}) = 5.91$$

ÇDR uygulaması sonucunda elde edilen uyum grafiği ve serpilme grafiği Şekil 4.1’de model belirleme kriterlerinden R^2 değeri Çizelge 4.1’de sunulmuştur.



(a)

(b)

Şekil 4.1. Temel bileşen skorları ile tarihi seri arasında yapılan regresyon modeli serpilme grafiği (a) ve uyum grafiği (b)

Çizelge 4.1. Çok değişkenli regresyon için model belirleme kriterleri

Yöntem	R^2	OKH	AIC
Çok Değişkenli Regresyon	0.97	0.004	-73.198

2008–2009 yıllarında meydana gelen ölümlü yaralanmalı ve maddi hasarlı trafik kaza sayısı 2008 yılı için 950120, 2009 yılı için 1034435 olarak gerçekleşmiştir. Yapılan logaritmik ters dönüşüm sonucunda modelden bulunan değerler ise güven aralıkları içine girerek 2008 yılında tahmini kaza sayısının 950.000 üzerinde, 2009 yılında ise 1040000’nin üzerinde olabileceği anlaşılmıştır.

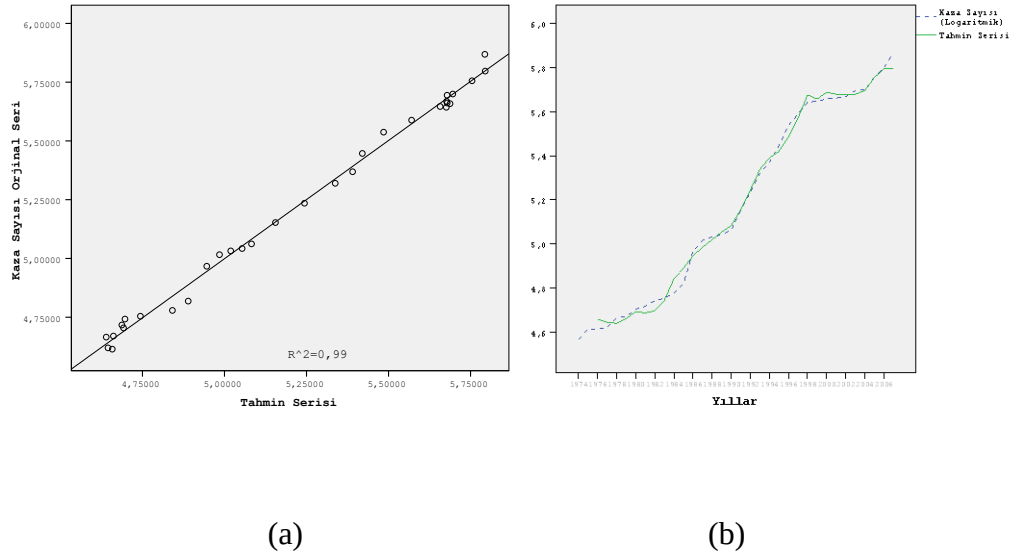
Çok değişkenli zaman serisi analizi uygulamasında VAR modelleri için varsayım olan durağanlık sonucu, durağan hale gelen seri ile tahminler yapılmış ve tahmin serisi ile tarihi seri arasındaki serpilme grafikleri ve uyum grafiği Şekil 4.2’de VAR(1) yöntemine ilişkin kıstaslar ise Çizelge 4.2’de sunulmuştur. VAR(1) modelinin genel ifadesi aşağıda verilmiştir.

$$t_t = C_1 \Delta t_{t-1} + C_2 \Delta k1_{t-1} + C_3 \Delta k2_{t-1} + C_4 \Delta k3_{t-1} + C_5 \Delta k4_{t-1} + C_6 \Delta k5_{t-1} + C_7 \Delta k6_{t-1} + C_8 \Delta k7_{t-1} + C_9 \Delta k8_{t-1} + C_{10} \Delta k9_{t-1} + C_{11}$$

Çalışmada elde edilen VAR(1) modeli;

$$t_t = 0.3298 \Delta t_{t-1} - 0.0414 * \Delta k1_{t-1} + 0.2141 * \Delta k2_{t-1} - 0.1420 * \Delta k3_{t-1} + 0.2028 * \Delta k4_{t-1} - 0.0641 \Delta k5_{t-1} + 0.7906 \Delta k6_{t-1} - 0.1403 \Delta k7_{t-1} + 0.0717 \Delta k9_{t-1} + 0.3611$$

olarak bulunmuştur.



Şekil 4.2. Orijinal seri ile tahmin serisi için yapılan VAR(1) modeli serpilme grafiği (a) ve uyum grafiği (b)

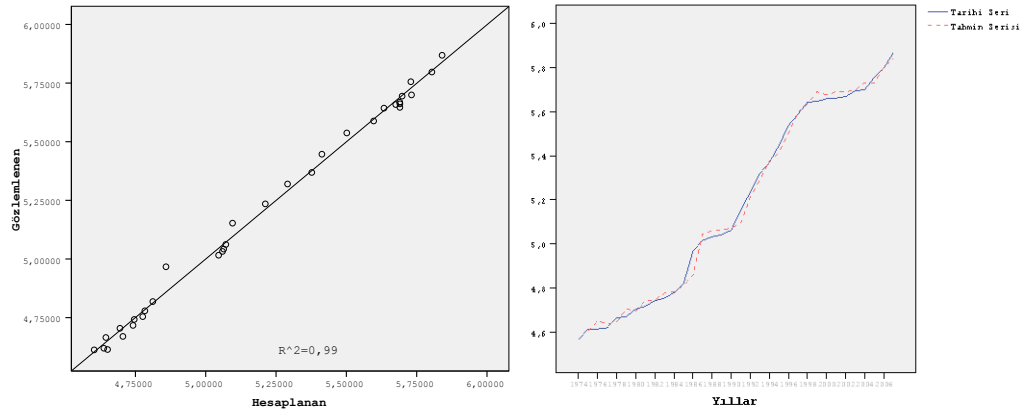
Çizelge 4.2. Çok değişkenli zaman serisi modeli model belirleme kriterleri

Yöntem	R ²	OKH	AIC
Vektörel Otoregresyon (VAR)	0.99	0.006	-41.88

Üçüncü yöntem olarak ARIMA(p,d,q) modelleri uygulanmıştır. Yapılan çalışmada bağımlı değişken kaza sayısı (Y), bağımsız değişken olarak AOGT (X₉) alınmıştır. Diğer bağımsız değişkenlerin katsayıları istatistiksel olarak anlamlı bulunmamıştır. Uyum grafiği ve serpilme grafikleri Şekil 4.3'te, model kıstasları ise Çizelge 4.3'te sunulmuştur.

$$y_t = \beta_0 + \beta_1 y_{t-1} + \beta_2 y_{t-2} + \beta_3 y_{t-3} + e_t$$

$$y_t = 0.360y_{t-1} + 0.012X_9$$



(a)

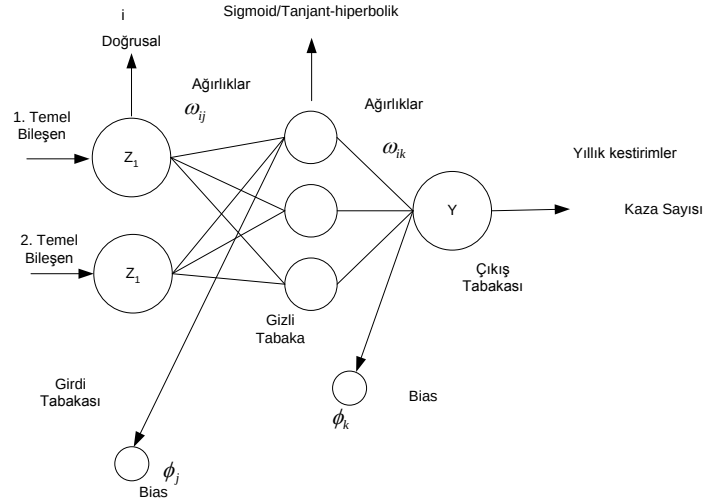
(b)

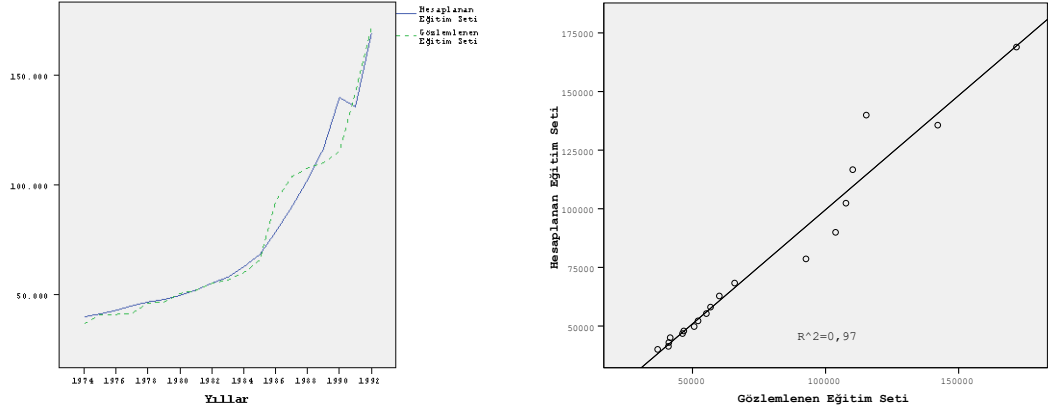
Şekil 4.3. Orijinal seri ile tahmin serisi arasında yapılan ARIMA(1,1,0) modeli uyum grafiği ve uyum grafiği

Çizelge 4.3. ARIMA (1,1,0) zaman serisi modeli model belirleme kriterleri

Yöntem	R^2	OKH	AIC
ARIMA(1,1,0)	0.99	0.02	-142.82

Dördüncü yöntem olarak yapay sinir ağları kullanılmıştır. Bu yöntemde birbirleri ilişkili 9 bağımsız değişken yerine prepca kodu kullanılarak Matlab temel bileşenler analizi yapılmış ve eğitim seti içinde toplam değişime %2'den daha az katkısı olan değişkenler alınmamıştır. Bu durumda birbirleri ilişkili 9 bağımsız değişken yerine birbirleri ile ilişkisiz 2 temel bileşen input olarak kabul edilmiştir (Şekil 4.4). Veriler [0,1] arasında bir normalize edilmiştir. Ağın eğitimi sonucu bulunan uyum ve serpilme grafikleri Şekil 4.5'te sunulmuştur. Ağın testi sonucu bulunan uyum ve serpilme grafikleri Şekil 4.6'da YSA yöntemi kıstasları ise Çizelge 4.4'te sunulmuştur.

**Şekil 4.4.** Örnek ağ mimarisi



(a)

(b)

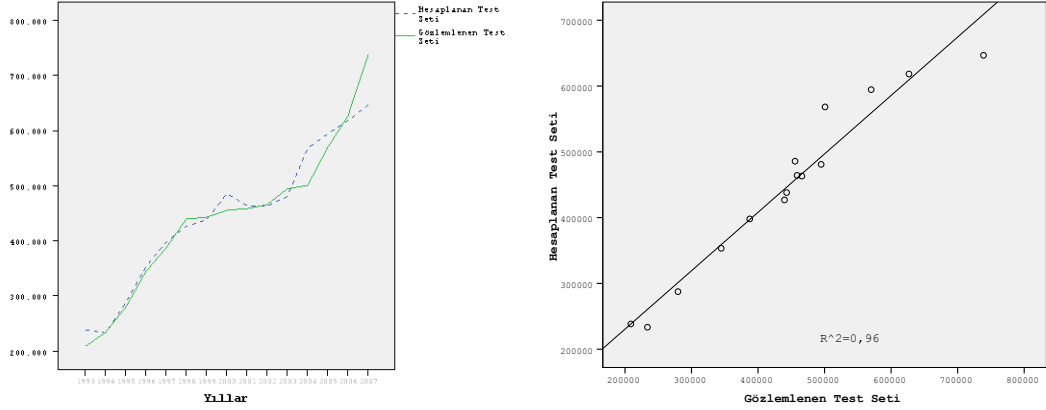
Şekil 4.5. Yapay sinir ağının eğitimi sırasında gözlemlenen ve hesaplanan kaza sayıları (a), ağı eğitimi aşamasındaki serpilme grafiği (b)

$$Y(t) = \phi_0 + \sum_{j=1}^m v_j \left(\int \sum_{i=1}^k w_{ij} \begin{bmatrix} p_i \\ r_j \end{bmatrix} + \phi_i \right) + e_i \quad (4.1)$$

$$Y(t) = \{v\}_{1 \times 3} \left(\tan \operatorname{sig} \left([w]_{3 \times 2} \begin{bmatrix} p_i \\ r_j \end{bmatrix}_{2 \times 1} + \{\phi\}_{3 \times 1} \right) \right) + \{\phi_0\}_{1 \times 1} \quad (4.2)$$

YSA ile elde edilen modelin matris şeklinde gösterimi aşağıda verilmiştir.

$$Y(t) = \begin{bmatrix} 0.5745 & 0.2600 & 0.6449 \end{bmatrix} \tan \operatorname{sig} \left(\begin{bmatrix} -0.5703 & 0.0519 \\ -0.2578 & 0.0235 \\ -0.6401 & 0.0589 \end{bmatrix}_{3 \times 2} \begin{bmatrix} p_i \\ r_j \end{bmatrix}_{2 \times 1} + \begin{bmatrix} -0.0443 \\ -0.0200 \\ -0.0497 \end{bmatrix}_{3 \times 1} \right) - [0.0768]$$



(a)

(b)

Şekil 4.6. Yapay sinir ağının testi sırasındaki gözlemlenen ve hesaplanan kaza sayıları (a), ağın testi aşamasındaki serpilme grafiği (b)

Çizelge 4.4. YSA modeli için belirleme kriterleri

Yöntem	Eğitim Seti			Test Seti		
	R^2	OKH	AIC	R^2	OKH	AIC
YSA	0.97	0.009	-113.439	0.96	0.002	-47.198

Çizelge 4.5. ÇDR, VAR(1), ARIMA (1,1,0) YSA yöntemlerinin karşılaştırılması

Yöntem	R^2	OKH	AIC
ÇDR	0.97	0.004	-73.93
VAR(1)	0.67	0.086	-62.30
ARIMA(1,1,0)	0.99	0.008	-131.26
YSA	0.95	0.002	-47.198

R^2 belirlilik katsayısı en yüksek olan ve Akaike AIC bilgi kriteri en küçük olan yöntem ARIMA(1,1,0) olmuştur (Çizelge 4.5).

4.1.2. Model-II bulguları

Yıllık trafik kazalarındaki yaralı sayılarının; nüfus, yol-km değeri, km-taşıt, ağır tonajlı araç sayısı, otomobil sayısı, otobüs sayısı, AOGT, GSMH değerlerine göre modellendi.

Yapılan bu çalışmada çok değişkenli regresyon denklemi uygulamasında temel bileşenler analizi uygulanmış olup birbirleri ile ilişkili 8 bağımsız değişken yerine birbirleri ile ilişkisiz 2 temel bileşenle regresyon modeli oluşturulmuştur.

$$t_i = 4.834 + (0.249 * SZ_1) + (0.016 * SZ_2)$$

Temel bileşenler skorları ile yapılan bu modelin 2007 yılı kestirim değerini bulmak için;

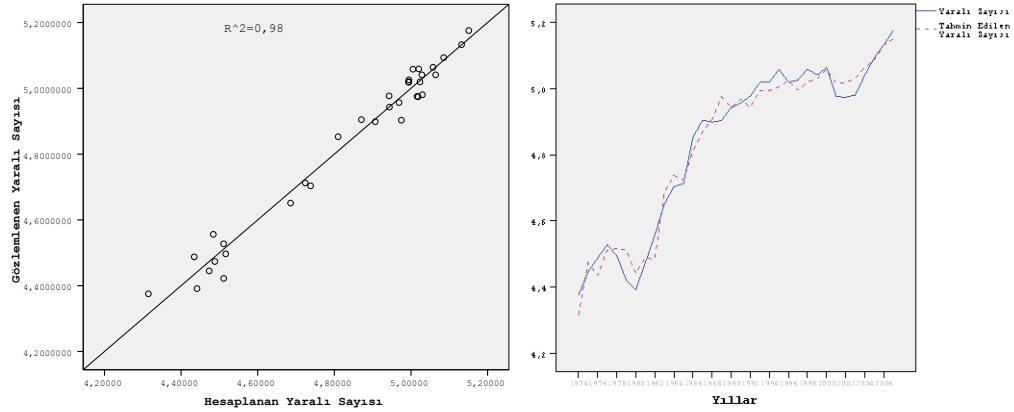
$$t = 4.834 + (0.249 * 1.7758) + (0.016 * 1.531159) = 5.15098 \text{ 2007 yılı kestirimi,}$$

2007 yılı gerçek değer ise $t = 5.175552$ dir.

Aynı tahmin aşağıda ters dönüşüm uygulanmış denklemlerde yapılmış ve yaklaşık aynı sonuç elde edilmiştir.

$$t = 67.23424 + (0.6889 * k_1) - (14.6312 * k_2) + (0.1568 * k_3) + (0.0721 * k_4) \\ + (0.1240 * k_5) + (0.1931 * k_6) + (0.3205 * k_7) + (0.1110 * k_8) = 5.15$$

ÇDR uygulamasında modelin uyum grafiği ve serpilme grafiği Şekil 4.7’de, yöntemin kıstasları ise Çizelge 4.6’da sunulmuştur.



(a)

(b)

Şekil 4.7. Temel bileşen skorları ile tarihi seri arasında yapılan regresyon modeli serpilme grafiği (a) ve uyum grafiği (b)

Çizelge 4.6. Çok değişkenli regresyon modeli model belirleme kriterleri

Yöntem	R^2	OKH	AIC
Çok Değişkenli Regresyon	0.98	0.004	-73.198

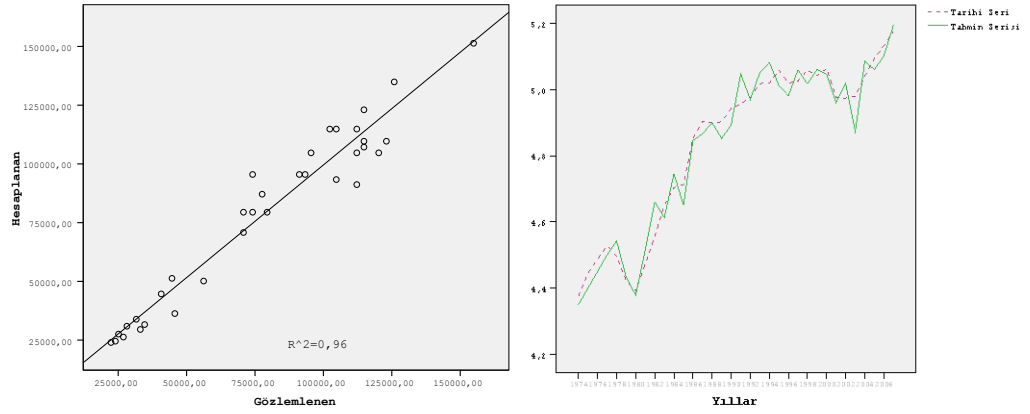
2008–2009 yıllarında meydana gelen trafik kazalarındaki yaralı sayısı 2008 yılı için 184468, 2009 yılı için 200405 olarak gerçekleşmiştir. Yapılan logaritmik ters dönüşüm sonucunda modelimizden bulunan değerler güven aralıkları içine girerek 2008 yılı için 170 binin üzerinde 2009 yılı için ise 185 binin üzerinde olduğu görülmüştür.

Çok değişkenli zaman serisi analizi uygulamasında, VAR modellerinin bir varsayımı olan durağanlık için, durağan hale gelen seriler ile tahminler yapılmış ve tahmin serisi ile tarihi seri arasındaki serpilme grafikleri ve uyum grafiği Şekil 4.8’de, yöntemin kıstasları Çizelge 4.7’de sunulmuştur.

VAR(2) modelinin genel ifadesi ve çalışmada elde edilen VAR(2) modeli aşağıda verilmiştir.

$$t_t = C_1 \Delta t_{t-1} + C_2 \Delta t_{t-2} + C_3 \Delta k1_{t-1} + C_4 \Delta k1_{t-2} + C_4 \Delta k2_{t-1} + C_5 \Delta k2_{t-2} + C_6 \Delta k3_{t-1} + C_7 \Delta k3_{t-2} + C_8 \Delta k4_{t-1} + C_8 \Delta k4_{t-2} + C_9 \Delta k5_{t-1} + C_{10} \Delta k5_{t-2} + C_{11} \Delta k6_{t-1} + C_{12} \Delta k6_{t-2} + C_{13} \Delta k7_{t-1} + C_{14} \Delta k7_{t-2} + C_{15} \Delta k8_{t-1} + C_{15} \Delta k8_{t-2} + C_{16}$$

$$t_t = 0.3374 \Delta t_{t-1} + 0.1883 \Delta t_{t-2} + 5.2225 \Delta k1_{t-1} + 2.8583 \Delta k1_{t-2} + 3.6373 \Delta k2_{t-1} - 0.4762 \Delta k2_{t-2} - 0.2304 \Delta k3_{t-1} - 0.2264 \Delta k3_{t-2} + 1.0966 \Delta k4_{t-1} - 1.9278 \Delta k4_{t-2} + 0.7558 \Delta k5_{t-1} + 0.2453 \Delta k5_{t-2} - 1.7218 \Delta k6_{t-1} + 2.9249 \Delta k6_{t-2} - 0.3020 \Delta k7_{t-1} - 0.3240 \Delta k7_{t-2} - 0.1914 \Delta k8_{t-1} + 0.1222 \Delta k8_{t-2} - 14.8425$$



(a)

(b)

Şekil 4.8. Orijinal seri ile tahmin serisi arasında yapılan VAR(2) modeli serpilme grafiği ve uyum grafiği

Çizelge 4.7. Çok değişkenli zaman serisi modeli için belirleme kriterleri

Yöntem	R^2	OKH	AIC
Vektörel Otoregresyon (VAR)	0.96	0.0006	-55.13

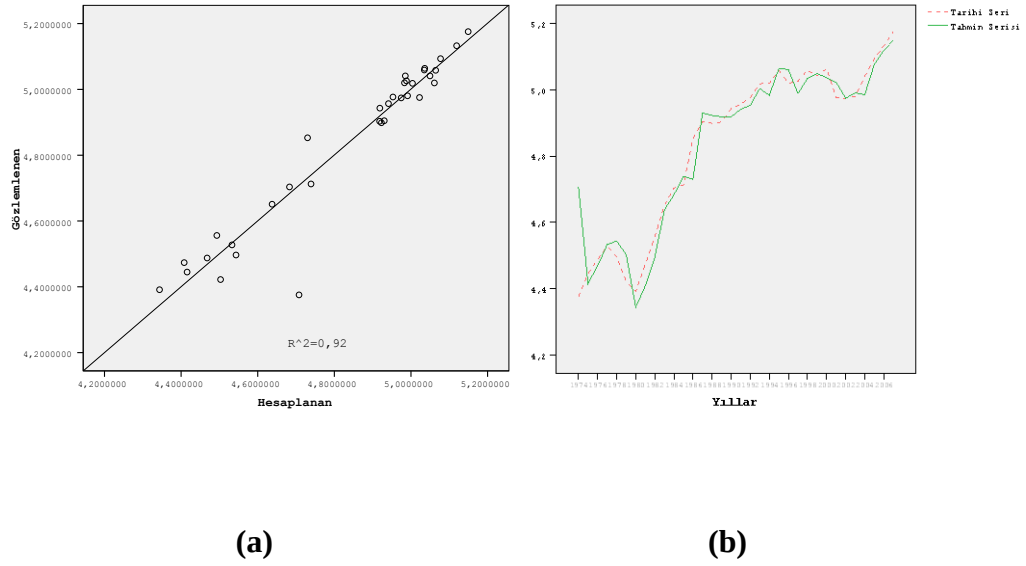
Üçüncü yöntem olarak ARIMA(p,d,q) modelleri uygulanmıştır. Yapılan çalışmada bağımlı değişken kaza sayısı (Y), bağımsız değişkenler ise; yol-km uzunluğu (X_2) ve (X_8) AOGT alınmıştır. Diğer bağımsız değişkenlerin katsayıları istatistiksel olarak

anlamlı bulunmamıştır. Yöntemin uyum grafiği ve serpilme grafiği Şekil 4.9.'da, kıstasları ise Çizelge 4.8'de sunulmuştur.

Elde edilen model aşağıda verilmiştir.

$$y_t = \beta_0 + \beta_1 y_{t-1} + \beta_2 y_{t-2} + \beta_3 y_{t-3} + e_t$$

$$y_t = 1.483y_{t-1} - 0.520y_{t-2} + 0.280X_2 + 0.067X_8$$



Şekil 4.9. Orijinal seri ile tahmin serisi arasında yapılan ARIMA(2,1,0) modeli serpilme grafiği (a) uyum grafiği (b)

Çizelge 4.8. ARIMA (2,1,0) zaman serisi modeli için belirleme kriterleri

Yöntem	R^2	OKH	AIC
ARIMA(2,1,0)	0.92	0.0021	-70.3525

Çizelge 4.9. ÇDR, VAR(2), ARIMA (2,1,0) yöntemlerinin karşılaştırılması

Yöntem	R ²	OKH	AIC
ÇDR	0.98	0.004	-73.198
VAR(1)	0.96	0.0006	-55.13
ARIMA(2,1,0)	0.92	0.0021	-70.35

Belirlilik katsayısı (R²) en yüksek olan ve AIC bilgi kriteri en küçük olan yöntem çok değişkenli regresyon yöntemi olmuştur (Çizelge 4.9).

4.1.3. Model-III bulguları

Yıllık trafik kazalarında ki ölü sayılarının, nüfus, yol-km değeri, km-taşıt, ağır tonajlı araç sayısı, otomobil sayısı, otobüs sayısı, AOGT, GSMH değerlerine göre modellendi.

Yapılan çalışmada çok değişkenli regresyon denklemi uygulamasında temel bileşenler analizi uygulanmış olup, birbirleri ile ilişkili 8 bağımsız değişken yerine, birbirleri ile ilişkisiz 2 temel bileşenle regresyon modeli oluşturulmuştur.

$$t_i = 3.690 - (0.022 * SZ_1) - (0.104 * SZ_2)$$

Temel bileşenler skorları ile yapılan bu modelin 2007 yılı kestirim değeri;

$$t = 3.690 - (0.022 * 1.17758) - (0.005 * 1.53159) = 3.5044 \text{ olarak hesaplanmıştır.}$$

2007 yılı gerçek değer ise $t = 3.539327$ dir.

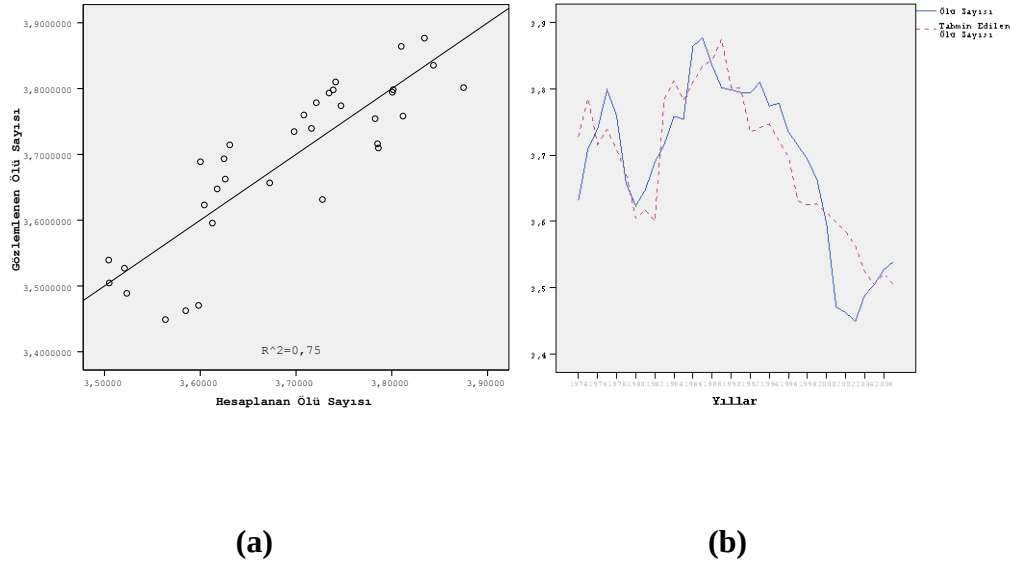
$$SY = \beta_1 * Z_1 + \beta_2 * Z_2$$

$$SY = 0.208 * Z_1 + 0.912 * Z_2$$

$$t_i = 74.11142 + (0.149225 * k_1) - (14.8667 * k_2) - (0.0207 * k_3) - (0.05159 * k_4) + (0.010132 * k_5) + (0.029724 * k_6) + (0.101859 * k_7) - (0.06903 * k_8)$$

Ters dönüşüm yapılarak elde edilen yukarıdaki denklemden elde edilen 2007 yılı için kesitirim değeri 3.50 dir.

ÇDR uygulamasında tahmin serisi ile tarihi seri arasındaki serpilme grafikleri ve uyum grafiği Şekil 4.10'da, yöntemin kıstasları ise Çizelge 4.10'da gösterilmiştir.



Şekil 4.10. Temel bileşen skorları ile tarihi seri arasında yapılan regresyon modeli serpilme grafiği (a) ve uyum grafiği (b)

Çizelge 4.10. Çok değişkenli regresyon modeli için belirleme kriterleri

Yöntem	R^2	OKH	AIC
Çok Değişkenli Regresyon	0.75	0.004	-77.50

2008–2009 yıllarında meydana gelen trafik kazalarında ölü sayısı 2008 yılı için 4236, 2009 yılı için 4300 olarak gerçekleşmiştir. Yapılan logaritmik ters dönüşüm sonucunda modelimizden bulunan değerlerin ise güven aralıkları içine girerek 2008 yılı için 4150'nin üzerinde 2009 yılı için ise 4200'ün üzerinde olabileceği anlaşılmıştır.

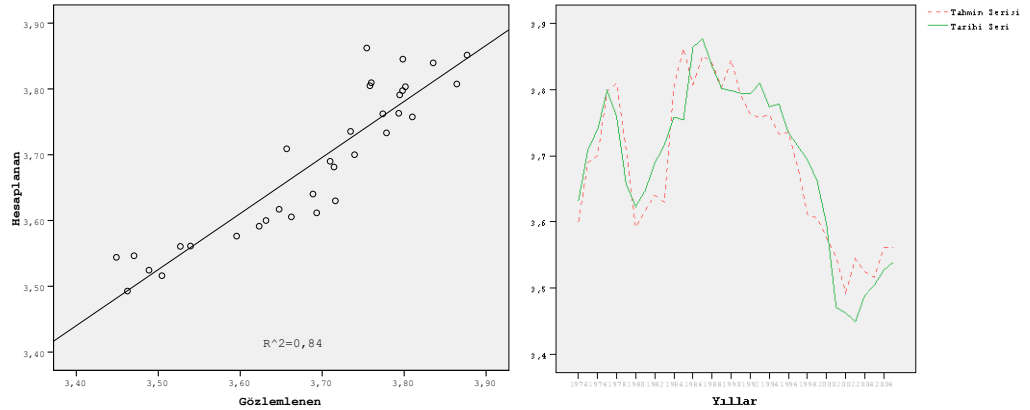
Çok değişkenli zaman serisi analizi uygulamasında VAR modelleri için varsayım olan durağanlık sonucu durağan hale gelen seri ile tahminler yapılmış ve tahmin serisi ile tarihi seri arasındaki serpilme grafikleri ve uyum grafiği Şekil 4.11’de, yöntemin kıstasları ise Çizelge 4.11’de gösterilmiştir.

VAR(2) modelinin genel ifadesi aşağıdaki gibi olmuştur.

$$t_t = C_1 \Delta t_{t-1} + C_2 \Delta t_{t-2} + C_3 \Delta k1_{t-1} + C_4 \Delta k1_{t-2} + C_5 \Delta k2_{t-1} + C_6 \Delta k2_{t-2} + C_7 \Delta k3_{t-1} + C_8 \Delta k3_{t-2} + C_9 \Delta k4_{t-1} + C_{10} \Delta k4_{t-2} + C_{11} \Delta k5_{t-1} + C_{12} \Delta k5_{t-2} + C_{13} \Delta k6_{t-1} + C_{14} \Delta k6_{t-2} + C_{15} \Delta k7_{t-1} + C_{16} \Delta k7_{t-2} + C_{17} \Delta k8_{t-1} + C_{18} \Delta k8_{t-2} + C_{19}$$

Çalışmada elde edilen model aşağıda verilmiştir.

$$t_t = 0.4158 \Delta t_{t-1} + 0.1790 \Delta t_{t-2} - 4.3984 \Delta k1_{t-1} + 1.8080 \Delta k1_{t-2} + 2.1302 \Delta k2_{t-1} + 0.8316 \Delta k2_{t-2} - 0.1334 \Delta k3_{t-1} + 0.2914 \Delta k3_{t-2} + 0.5473 \Delta k4_{t-1} - 1.5641 \Delta k4_{t-2} - 1.2213 \Delta k5_{t-1} + 0.7618 \Delta k5_{t-2} + 0.1783 \Delta k6_{t-1} + 1.0547 \Delta k6_{t-2} - 0.6987 \Delta k7_{t-1} - 0.0810 \Delta k7_{t-2} + 0.0054 \Delta k8_{t-1} - 0.0716 \Delta k8_{t-2} - 13.9133$$



(a)

(b)

Şekil 4.11. Orijinal seri ile tahmin serisi arasında yapılan VAR(2) modeli serpilme grafiği (a) ve uyum grafiği (b)

Çizelge 4.11. Çok değişkenli zaman serisi modeli için belirleme kriterleri

Yöntem	R ²	OKH	AIC
Vektörel Otoregresyon (VAR)	0.84	0.0008	-52.4288

Üçüncü yöntem olarak ARIMA(p,d,q) modeli uygulanmıştır. Yapılan çalışmada bağımlı değişken kaza sayısı (Y), bağımsız değişken ise; yol-km uzunluğu olarak (X₂) alınmıştır. Diğer bağımsız değişkenlerin katsayıları istatistiksel olarak anlamlı bulunmamıştır.

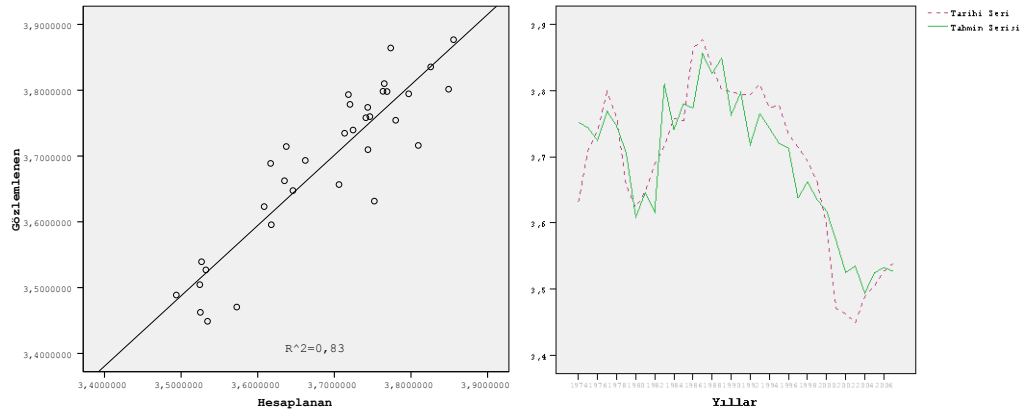
Modelin genel ifadesi aşağıda verilmiştir.

$$y_t = \beta_0 + \beta_1 e_{t-1} + \beta_2 e_{t-2} + \dots + \varepsilon_t$$

Çalışmada elde edilen model aşağıdaki gibi olmuştur.

$$y_t = 18.086 - 0.517e_{t-1} - 3.382X_2$$

ARIMA(0,1,1) modeline ilişkin uyum ve serpilme grafikleri Şekil 4.12’de, modelin kıstasları ise Çizelge 4.12’de sunulmuştur.



(a)

(b)

Şekil 4.12. Orijinal seri ile tahmin serisi arasında yapılan ARIMA(0,1,1) modeli serpilme grafiği (a) ve uyum grafiği (b)

Çizelge 4.12. ARIMA (0,1,1) zaman serisi modeli için belirleme kriterleri

Yöntem	R^2	OKH	AIC
ARIMA(0,1,1)	0.83	0.002	-120.924

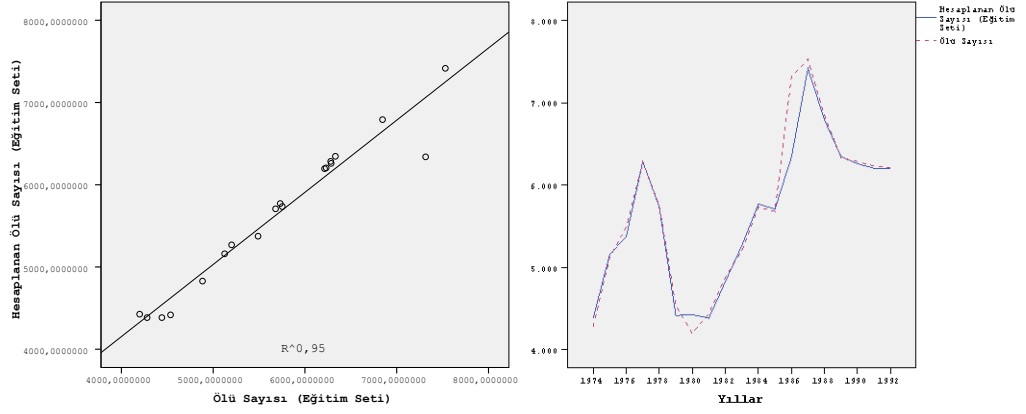
Çizelge 4.13. ÇDR, VAR(1), ARIMA (0,1,1) YSA yöntemlerinin karşılaştırılması

Yöntem	R^2	OKH	AIC
ÇDR	0.75	0.004	-77.50
VAR(2)	0.84	0.0008	-52.4288
ARIMA(0,1,1)	0.83	0.002	-120.92

Belirlilik katsayısı (R^2) en yüksek olan yöntem VAR(2) ve Akaile AIC bilgi kriteri en küçük olan yöntem ARIMA(0,1,1) olmuştur (Çizelge 4.13).

4.1.4. Model II-III için yapay sinir ağı uygulaması

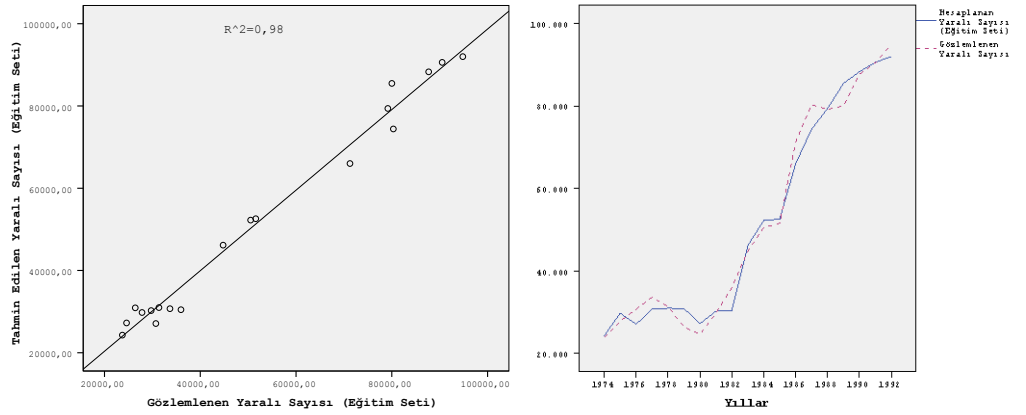
YSA uygulamasında verilerin %60'ı eğitim, %40'da test amaçlı kullanıldığından ölü ve yaralı sayılarına ilişkin eğitim ve test setindeki uyum ve serpilme grafikleri Şekil 4.13 - 4.14 – 4.15 - 4.16'da sunulmuştur. Yaralı ve ölü sayılarına ilişkin model kıstasları ise Çizelge 4.14–4.15'te sunulmuştur.



(a)

(b)

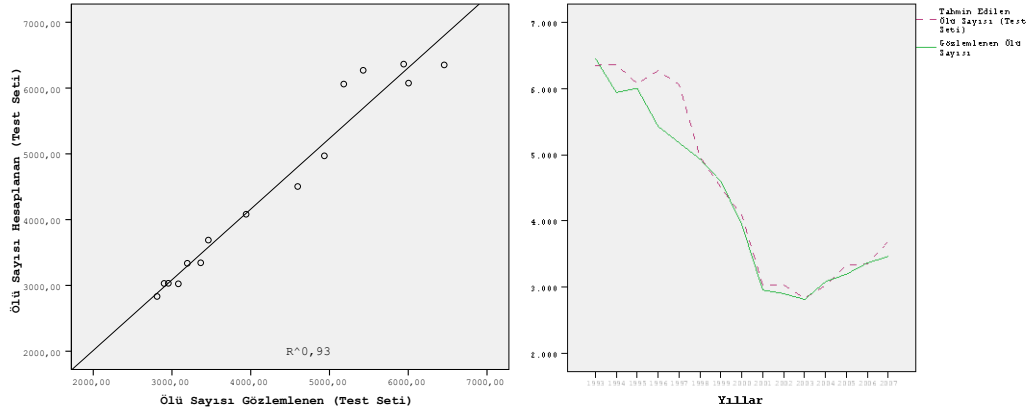
Şekil 4.13. Ölü sayıları için yapay sinir ağının eğitimi sırasında gözlemlenen ve hesaplanan kaza sayıları serpilme grafiği (a), ağın eğitimi aşamasındaki uyum grafiği (b)



(a)

(b)

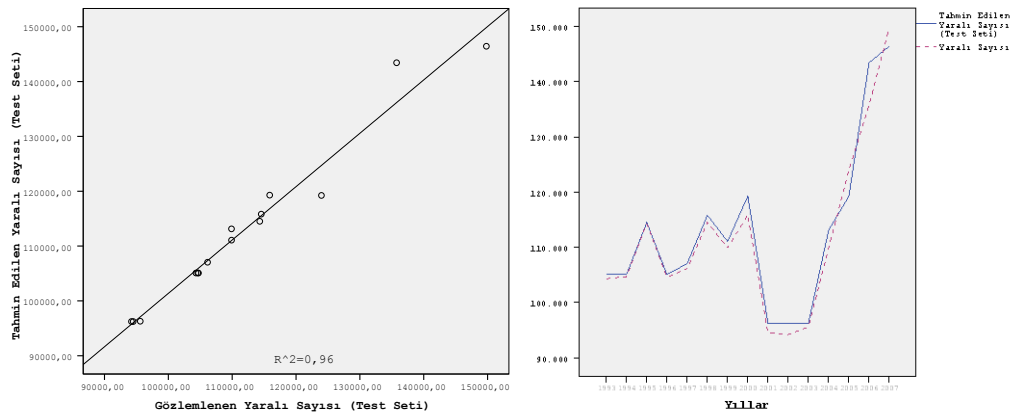
Şekil 4.14. Yaralı sayıları için yapay sinir ağının eğitimi sırasında gözlemlenen ve hesaplanan kaza sayıları serpilme (a), ağın eğitimi aşamasındaki uyum grafiği (b)



(a)

(b)

Şekil 4.15. Ölü sayılarının yapay sinir ağının testi sırasındaki gözlemlenen ve hesaplanan kaza sayıları serpilme (a), ağın testi aşamasındaki uyum grafiği (b)



(a)

(b)

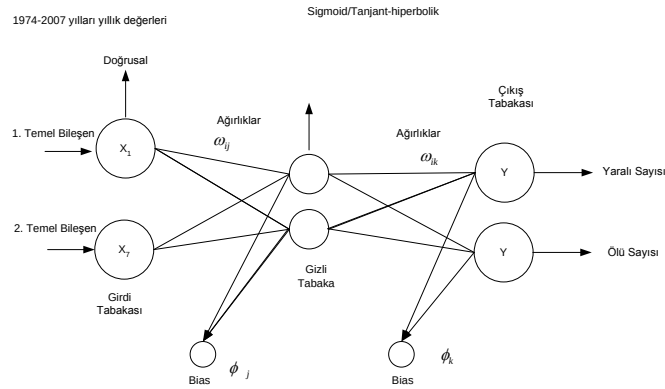
Şekil 4.16. Yaralı sayılarının yapay sinir ağının testi sırasındaki gözlemlenen ve hesaplanan kaza sayıları serpilme (a), ağın testi aşamasındaki uyum grafiği (b)

Çizelge 4.14. Ölü sayıları için YSA modeli için belirleme kriterleri

Yöntem	Eğitim Seti			Test Seti		
	R^2	OKH	AIC	R^2	OKH	AIC
YSA	0.95	0.0014	-30.088	0.93	0.004	-23.83

Çizelge 4.15. Yaralı sayıları için YSA modeli için belirleme kriterleri

Yöntem	Eğitim Seti			Test Seti		
	R^2	OKH	AIC	R^2	OKH	AIC
YSA	0.98	0.0017	-28.02	0.96	0.0002	-49.31

**Şekil 4.17.** Model II-III için YSA ağ mimarisi

4.1.5. Model-IV bulguları

Yıllık trafik kazalarındaki yaralı sayıları; kaza sayısı, AOGT ve GSMH değerlerine göre modellendi.

Çalışmada çok değişkenli regresyon uygulamasında temel bileşenler analizi ile olup birbirleri ile ilişkili 3 bağımsız değişken yerine 1 temel bileşenle model oluşturulmuştur.

$$t_i = 4.834 + (0.234 * SZ_1)$$

Temel bileşenler skorları ile yapılan bu modelin 2007 yılı kestirim değeri;

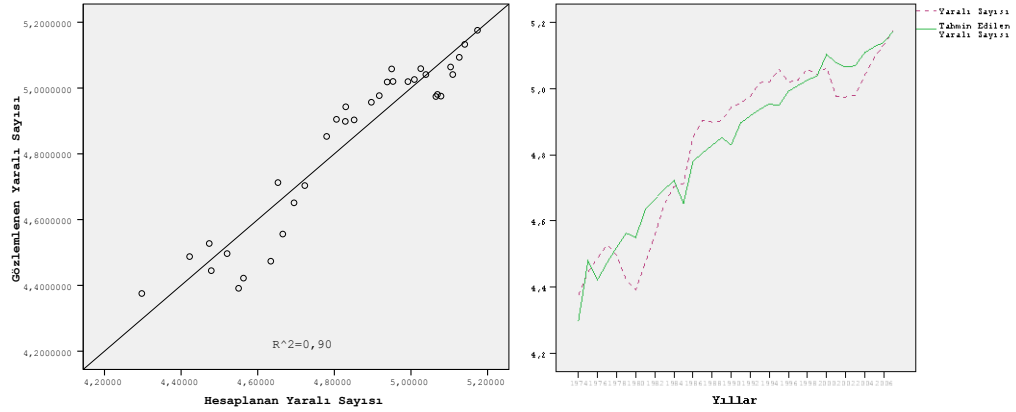
$$t = 4.834 + (0.234 * 1.75965) = 5.24851 \text{ olarak hesaplanmıştır.}$$

2007 yılı gerçek değer ise $t = 5.175552$ dir.

$$SY = \beta_1 * Z_1 + \beta_2 * Z_2$$

$$SY = -0.552 * Z_1$$

$$t = 1.356429 + (0.185939 * k_1) + (0.365186 * k_2) + (0.387412 * k_3)$$



(a)

(b)

Şekil 4.18. Temel bileşen skorları ile tarihi seri arasında yapılan regresyon modeli serpilme grafiği (a) ve uyum grafiği (b)

Çizelge 4.16. Çok değişkenli regresyon modeli için belirleme kriterleri

Yöntem	R^2	OKH	AIC
Çok Değişkenli Regresyon	0.86	0.00018	-67.003

ÇDR uygulamasında tahmin serisi ile tarihi seri arasındaki serpilme grafikleri ve uyum grafiği Şekil 4.18’de, yöntemin kıstasları ise Çizelge 4.16’da gösterilmiştir.

2008–2009 yıllarında meydana gelen trafik kazalarındaki yaralı sayısı 2008 yılı için 184468, 2009 yılı için 200405 olarak gerçekleşmiştir. Yapılan logaritmik ters dönüşüm sonucunda modelimizden bulunan değerlerin ise güven aralıkları içine girerek 2008 yılı

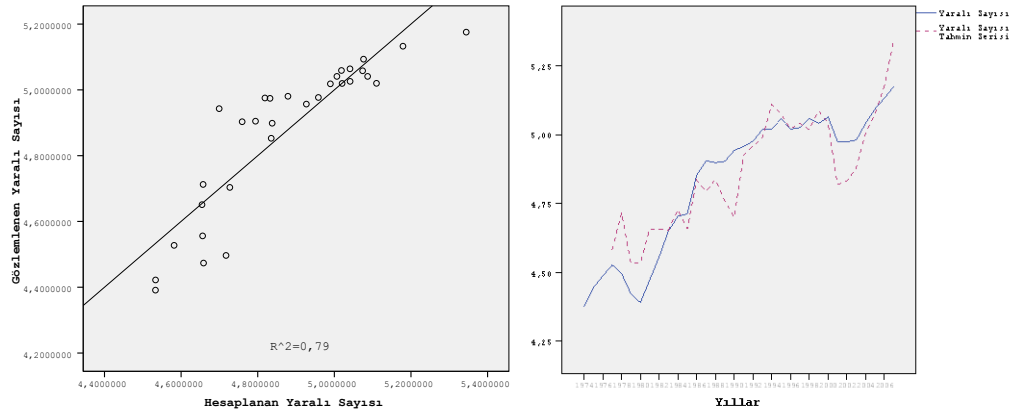
için 171 binin üzerinde 2009 yılı içinde 183 binin üzerinde olabileceği tahmin edilmiştir.

Çok değişkenli zaman serisi analizi uygulamasında durağan hale gelen seriler ile tahminler yapılmış ve tahmin serisi ile tarihi seri arasındaki serpilme grafikleri ve uyum grafiği Şekil 4.19'da yöntemin kıstasları ise Çizelge 4.17'de gösterilmiştir. Modelin genel ifadesi aşağıdaki gibi olmuştur.

$$t_t = C_1 \Delta t_{t-1} + C_2 \Delta k1_{t-1} + C_3 \Delta k2_{t-1} + C_4 \Delta k3_{t-1} + C_{16}$$

Çalışmada elde edilen model aşağıda verilmiştir.

$$t_t = 0.3986 \Delta t_{t-1} + 0.0641 \Delta k1_{t-1} - 0.3665 \Delta k2_{t-1} - 0.0447 \Delta k3_{t-1} + 0.1692$$



(a)

(b)

Şekil 4.19. Orijinal seri ile tahmin serisi arasında yapılan VAR(1) modeli serpilme grafiği (a) ve uyum grafiği (b)

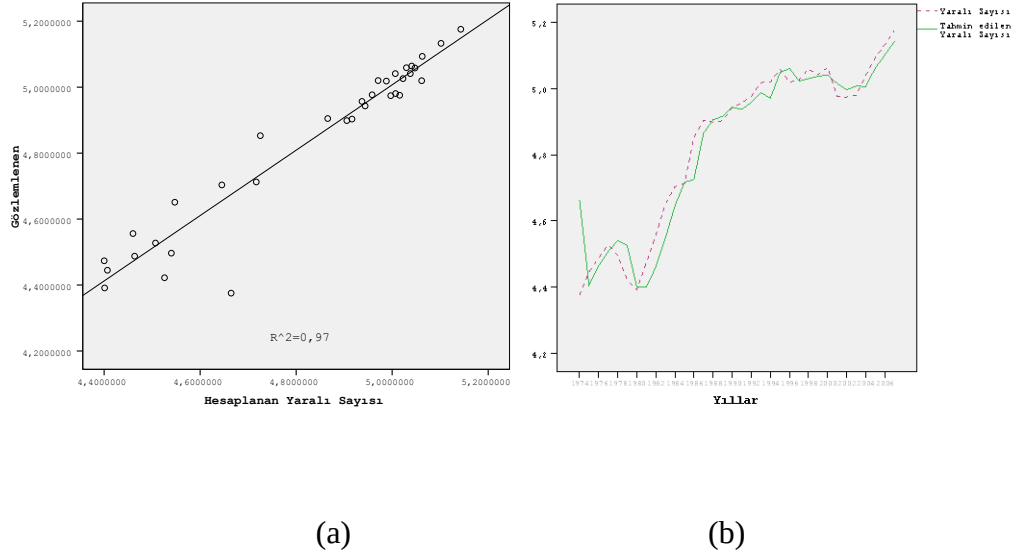
Çizelge 4.17. Çok değişkenli zaman serisi modeli için belirleme kriterleri

Yöntem	R^2	OKH	AIC
Vektörel Otoregresyon (VAR)	0.79	0.003	-13.82

Üçüncü yöntem olarak ARIMA(p,d,q) modelleri verilere uygulanmıştır. Yapılan çalışmada bağımlı değişken kaza sayısı (Y), bağımsız değişkenlerden AOGT (X_1) alınmıştır. Diğer bağımsız değişkenlerin parametreleri istatistiksel olarak anlamlı bulunmamıştır. Yöntemin uygulamsı sonucunda tahmin edilen değerler ile gerçek değerler arasındaki uyum ve serpilme grafikleri Şekil 4.20’de yöntemin kıstasları ise Çizelge 4.18’de sunulmuştur.

$$y_t = \beta_0 + \beta_1 y_{t-1} + \beta_2 y_{t-2} + \beta_3 y_{t-3} + e_t$$

$$y_t = 3.717 + 0.981y_{t-1} + 0.317X_2$$



Şekil 4.20. Orijinal seri ile tahmin serisi arasında yapılan ARIMA(1,1,0) modeli serpilme (a) grafiği ve uyum grafiği (b)

Çizelge 4.18. ARIMA (1,1,0) zaman serisi modeli için belirleme kriterleri

Yöntem	R^2	OKH	AIC
ARIMA(1,1,0)	0.92	0.0052	-75.93

Çizelge 4.19. ÇDR, VAR(1), ARIMA (1,1,0) yöntemlerinin karşılaştırılması

Yöntem	R ²	OKH	AIC
ÇDR	0.86	0.00018	-67.003
VAR(1)	0.79	0.003	-13.82
ARIMA(1,1,0)	0.92	0.0052	-75.93

Belirlilik katsayısı (R²) en yüksek olan ve Akaile AIC bilgi kriteri en küçük olan yöntem çok değişkenli regresyon yöntemi olmuştur (Çizelge 4.19).

4.2.1. Model-V bulguları

Yıllık trafik kazalarındaki ölü sayılarının, kaza sayısı, AOGT ve GSMH değerlerine göre modellendi.

Yapılan çalışmada çok değişkenli regresyon denklemi uygulamasında temel bileşenler analizi uygulanmış olup birbirleri ile ilişkili 3 bağımsız değişken yerine 1 temel bileşenle regresyon modeli oluşturulmuştur.

$$t_i = 3.690 - (0.059 * SZ_1)$$

Temel bileşenler skorları ile yapılan bu modelin 2007 yılı kestirim değeri,

$$t = 3.690 - (0.059 * 1.75965) = 3.5862 \text{ olarak hesaplanmıştır.}$$

2007 yılı gerçek değer ise $t = 3.539327$ dir.

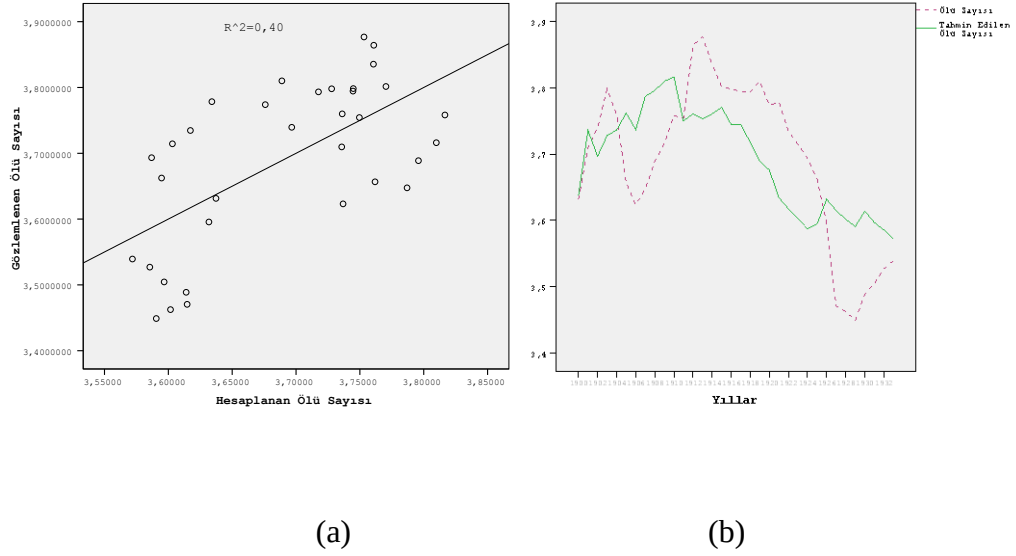
$$SY = \beta_1 * Z_1$$

$$SY = 0.288 * Z_1$$

$$Y = 4.570631 - (0.0471 * k_1) - (0.0925 * k_2) - (0.0981 * k_3)$$

Ters dönüşüm yapılarak eld edilen yukarıdaki denklemden 2007 yılı için kesitirim değeri 3.58'dir.

ÇDR uygulamasında tahmin serisi ile tarihi seri arasındaki serpilme grafikleri ve uyum grafiği Şekil 4.21’de, yöntemin kıstasları ise Çizelge 4.20’de gösterilmiştir.



Şekil 4.21. Temel bileşen skorları ile tarihi seri arasında yapılan regresyon modeli serpilme grafiği (a) ve uyum grafiği (b)

Çizelge 4.20. Çok değişkenli regresyon modeli için belirleme kriterleri

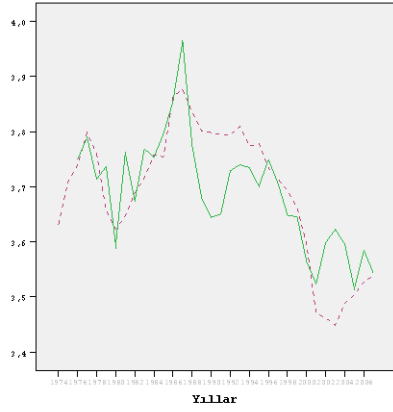
Yöntem	R^2	OKH	AIC
Çok Değişkenli Regresyon	0.40	0.010	-65.59

2008–2009 yıllarında meydana gelen trafik kazalarındaki ölü sayısı 2008 yılı için 4236, 2009 yılı için 4300 olarak gerçekleşmiştir. Yapılan logaritmik ters dönüşüm sonucunda modelimizden bulunan değerler ise güven aralıkları içine girerek 2008 yılı için tahmini kaza sayısı 4100 ün üzerinde bir ölü sayısı tahmin edilmiş 2009 yılı içinde 4320’nin üzerinde bir kazanın olabileceği anlaşılmıştır.

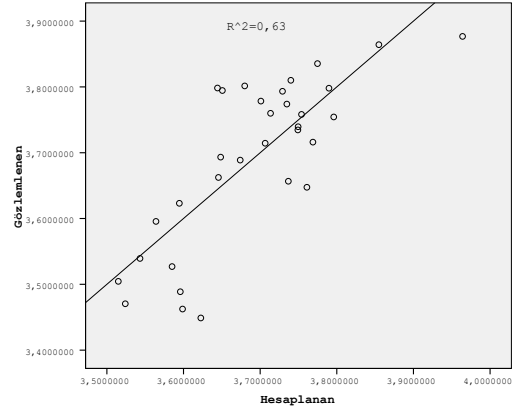
Çok değişkenli zaman serisi analizi uygulamasında durağan hale gelen seri ile tahminler yapılmış ve tahmin serisi ile tarihi seri arasındaki serpilme grafikleri ve uyum grafiği Şekil 4.22’de yöntemin kıstasları ise Çizelge 4.21’de gösterilmiştir.

$$t_t = C_1 \Delta t_{t-1} + C_2 \Delta k1_{t-1} + C_3 k2_{t-1} + C_4 \Delta k3_{t-1} + C_5$$

$$t_t = 0.4916 \Delta t_{t-1} + 0.00389 * \Delta k1_{t-1} - 0.4453 * \Delta k2_{t-1} - 0.0463 * \Delta k3_{t-1} + 0.1614$$



(a)



(b)

Şekil 4.22. Orijinal seri ile tahmin serisi arasında yapılan VAR(1) modeli serpilme grafiği (a) ve uyum grafiği (b)

Çizelge 4.21. Çok değişkenli zaman serisi modeli için belirleme kriterleri

Yöntem	R ²	OKH	AIC
Vektörel Otoregresyon (VAR)	0.63	0.0019	-14.02116

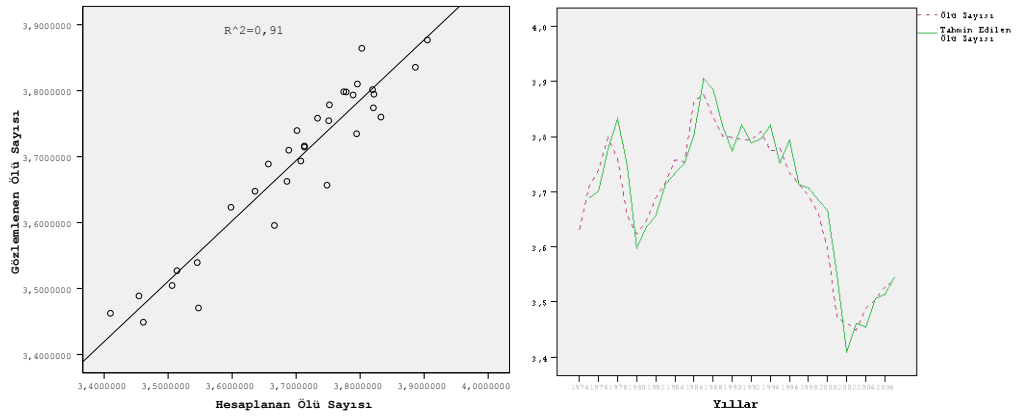
Üçüncü yöntem olarak ARIMA(p,d,q) modeli uygulanmıştır. Yapılan çalışmada bağımlı değişken kaza sayısı (Y), bağımsız değişken ise; AOGT (X₂) alınmıştır. Diğer bağımsız değişkenlerin katsayıları istatistiksel olarak anlamlı bulunmamıştır. Yöntem sonucu tahmin değerleri ile gerçek değerler arasındaki uyum ve serpilme grafikleri Şekil 4.23'te yöntemin kıstasları ise Çizelge 4.22'de sunulmuştur.

Modelin genel ifadesi aşağıda verilmiştir.

$$y_t = \beta_0 + \beta_1 e_{t-1} + \beta_2 e_{t-2} + \dots + \varepsilon_t$$

Çalışmada elde edilen model aşağıdaki gibidir.

$$y_t = 0.471y_{t-1} + 0.328X_2$$



(a)

(b)

Şekil 4.23. Orijinal seri ile tahmin serisi arasında yapılan ARIMA(1,1,0) modeli uyum grafiği (a) ve uyum grafiği (b)

Çizelge 4.22. ARIMA (1,1,0) zaman serisi modeli için belirleme kriterleri

Yöntem	R^2	OKH	AIC
ARIMA(1,1,0)	0.91	0.0016	-92.78

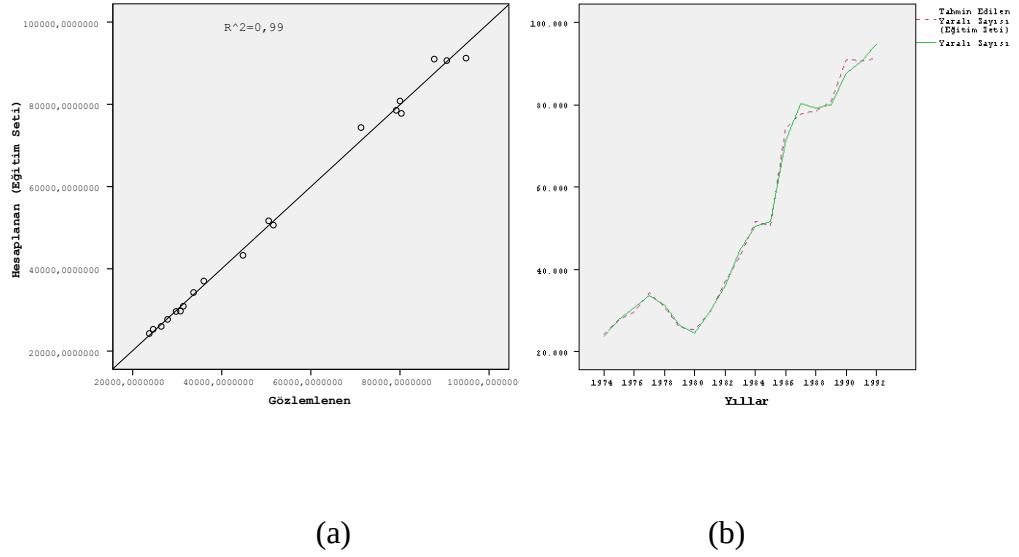
Çizelge 4.23. ÇDR, VAR(1), ARIMA (1,1,0) YSA yöntemlerinin karşılaştırılması

Yöntem	R^2	OKH	AIC
ÇDR	0.40	0.010	-65.59
VAR(1)	0.63	0.0019	-14.02116
ARIMA(1,1,0)	0.91	0.0016	-92.78

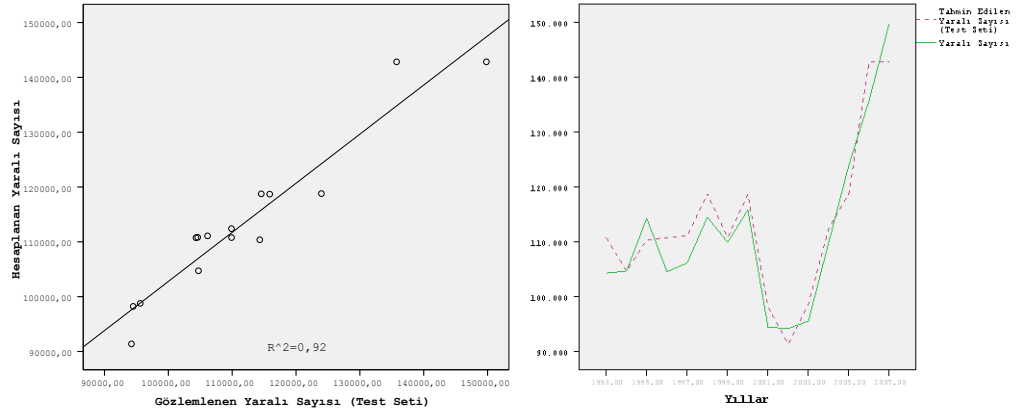
Belirlilik katsayısı (R^2) en yüksek olan yöntem ARIMA(1,1,0), ve AIC en küçük olan yöntem yine ARIMA(1,1,0) olmuştur (Çizelge 4.23).

4.1.7. Model IV-V için yapay sinir ağı uygulaması

YSA uygulamasında verilerin %60'ı eğitim, %40'da test amaçlı kullanıldığından ölü ve yaralı sayılarına ilişkin eğitim ve test setindeki uyum ve serpilme grafikleri Şekil 4.24 – 4.25 – 4.26 – 4.27'de sunulmuştur. Yaralı ve ölü sayılarına ilişkin model kıstasları ise Çizelge 4.24 – 4.25'de sunulmuştur.



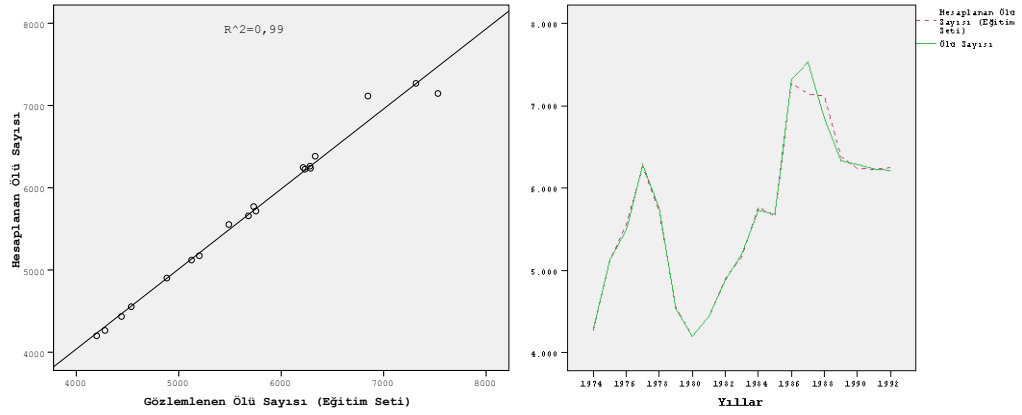
Şekil 4.24. Yaralı sayıları için yapay sinir ağına eğitimi sırasında gözlemlenen ve hesaplanan kaza sayılarının ağı eğitimi aşamasındaki serpilme grafiği (a), uyum grafiği (b)



(a)

(b)

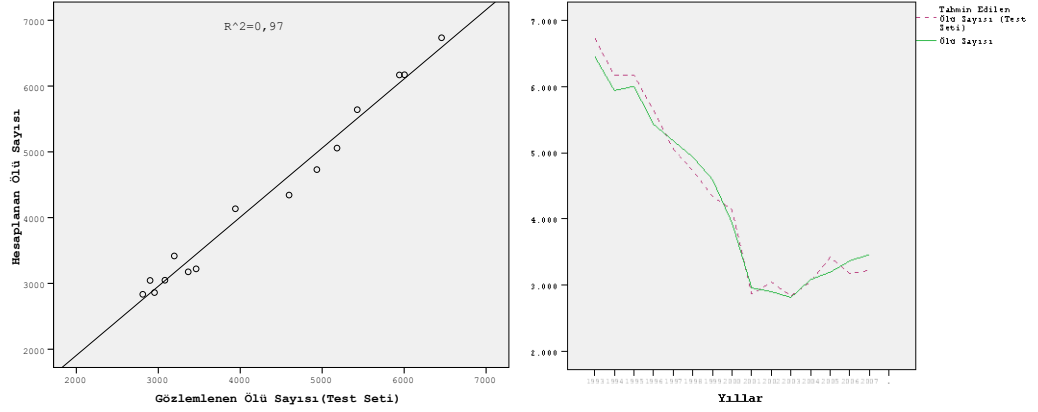
Şekil 4.25. Yaralı sayıları için yapay sinir ağının eğitimi sırasında gözlemlenen ve hesaplanan kaza sayılarının, ağın testi aşamasındaki serpilme grafiği (a), uyum grafiği (b)



(a)

(b)

Şekil 4.26. Ölü sayılarının yapay sinir ağının eğitimi sırasında gözlemlenen ve hesaplanan kaza sayılarının, ağın testi aşamasındaki serpilme grafiği (a), uyum grafiği (b)



(a)

(b)

Şekil 4.27. Ölü sayılarının yapay sinir ağının testi sırasındaki gözlemlenen ve hesaplanan kaza sayılarının, ağın testi aşamasındaki serpilme grafiği (a), uyum grafiği (b)

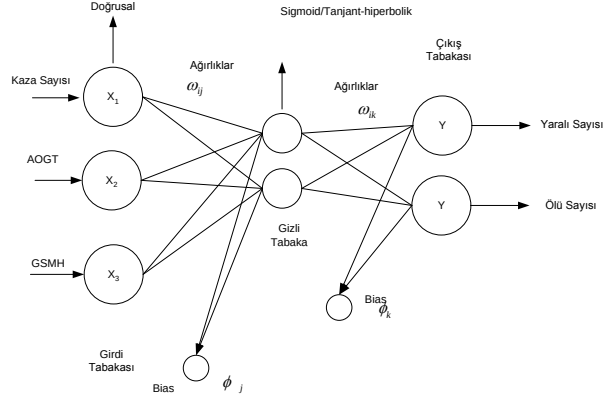
Çizelge 4.24. Ölü sayıları için YSA modeli için belirleme kriterleri

Yöntem	Eğitim Seti			Test Seti		
	R^2	OKH	AIC	R^2	OKH	AIC
YSA	0.99	0.00026	-39.78	0.97	0.0021	-22.49

Çizelge 4.25. Yaralı sayıları için YSA modeli için belirleme kriterleri

Yöntem	Eğitim Seti			Test Seti		
	R^2	OKH	AIC	R^2	OKH	AIC
YSA	0.99	0.00018	-42.51	0.92	0.00045	-36.30

1974-2007 yılları yıllık değerleri



Şekil 4.28. Model IV-V için YSA mimarisi

4.2. Poisson Regresyonu, Çok Değişkenli İstatistik, Yapay Sinir Ağı Modellemeleri

4.2.1. Model-VI bulguları

Denenen birçok model içinde parametreleri en anlamlı model aşağıdaki gibi bulunmuştur.

$$E(A) = \exp(b_0 + b_1 X_1 + b_2 X_2 + \dots + b_q X_q)$$

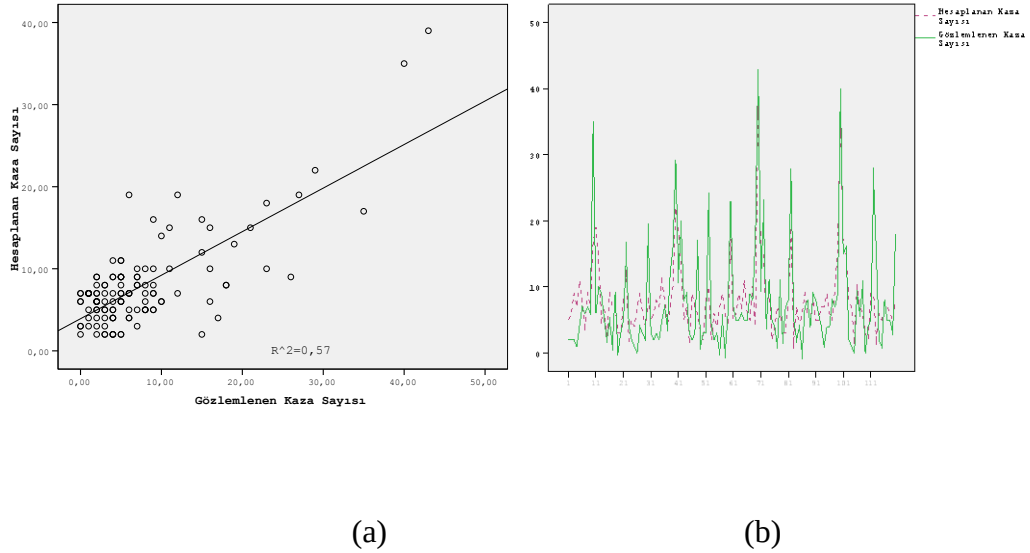
$$E(A) = \exp(1.239572 * \log uz + 0,0002328 * yogt + 0.1746494 * sg + 0.0604205 * bs - 0.011652 * hiz)$$

Burada;

- n: Gözlem sayısı
- p: Parametre sayısı
- sd: Serbestlik derecesi
- sabit: Sabit terim

ks:	Kaza sayısı
yogt:	Yıllık ortalama günlük trafik
log(uz):	İlgili yolun km yol uzunluğunun doğal logaritması
ss:	Şerit sayısı
kavs:	Kavşak sayısı
sg:	Şerit genişliği
bs:	bağlantı sayısı
hız :	Ortalama hız

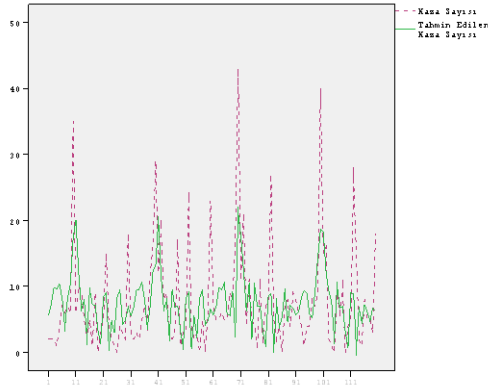
Poisson regresyonu bütün bağımsız değişkenler modele katılmıştır. Yapılan analizlerde bütün bağımsız değişkenler modele katıldığında P değerleri bazı değişkenler için 0.05 değerinden büyük olduklarından anlamlı bulunmamış ve modelden çıkarılmışlardır. Modelleme aşamasında, kavşak sayısı (kavs) ve şerit sayısı (ss), bağımsız değişken olarak kullanılmış fakat bu değerler %95 güven aralığında çıkmadığından modelden çıkarılmışlardır. Yöntemin uyum ve serpilme grafikleri Şekil 4.29’da sunulmuştur.



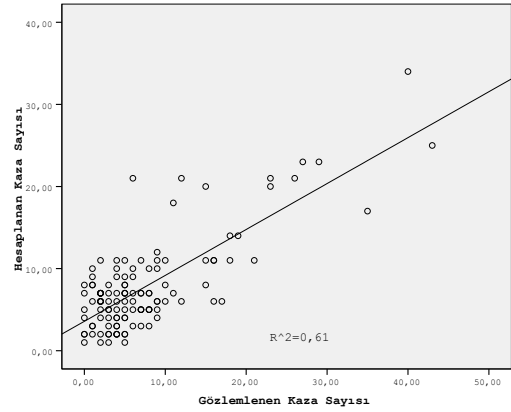
Şekil 4.29. Model VI-Poisson regresyon modeli için serpilme grafiği (a) ve uyum grafiği (b)

Çok değişkenli regresyon uygulamasında ise $R^2 = 0.61$ olarak bulunmuş ve modelin matematiksel ifadesi aşağıdaki gibi olmuştur (Şekil 4.30).

$$ks = 8.654 * uz + 13.157 * yogt + 0.707 * bs - 24.866hiz$$



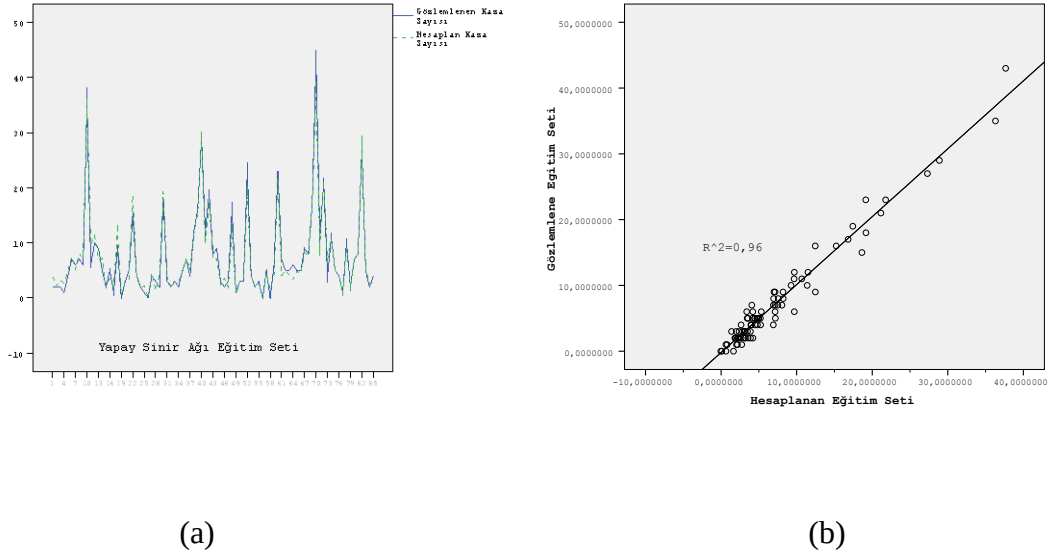
(a)



(b)

Şekil 4.30. Model VI-Çok değişkenli regresyon modeli için serpilme grafiği (a) ve uyum grafiği (b)

YSA uygulamasında eğitim seti için uyum ve serpilme grafikleri Şekil 4.31’de sunulmuştur.



Şekil 4.31. Model VI-YSA eğitim seti uyum grafiği (a) ve serpilme grafiği (b)

Çizelge 4.26. Yöntemlerin karşılaştırılması

Yöntem	R^2	OKH	Değerlendirme
Poisson Reg.	0.57	0.33	
ÇDR	0.61	0.26	
YSA	0.93	0.004	En anlamlı yöntem

Belirlilik katsayısı (R^2) en yüksek olan yöntem YSA, OKH en küçük olan yöntem yine YSA olmuştur (Çizelge 4.26).

4.2.2. Model-VII bulguları

Modelde yaralı sayıları bağımlı değişken olarak alınmıştır. Denenen birçok model içinde parametreleri en anlamlı model aşağıdaki gibi bulunmuştur.

$$E(\mu_i) = \exp(b_0 + b_1 X_1 + b_2 X_2 + \dots + b_q X_q)$$

$$\mu_i = \exp(3.207331 + 0.0196034 * (ks) + 0,0461516 * (ss) - 0.0848338(kavs) + 0.0512742(bs) - 0.008345(hiz))$$

Burada;

n: Gözlem sayısı

p: Parametre sayısı

sd: Serbestlik derecesi

Ys (μ): Yaralı sayısı

sabit: Sabit terim

ks: Kaza sayısı

yogt: Yıllık ortalama günlük trafik

ss: Şerit sayısı

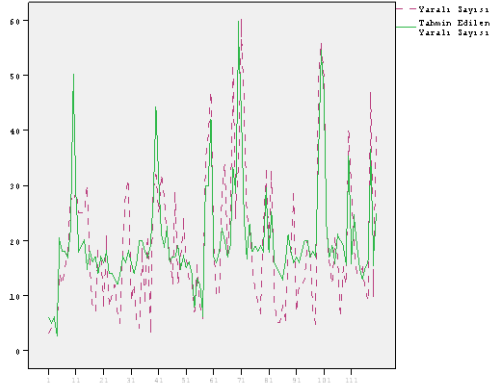
kavs: Kavşak sayısı

sg: Şerit genişliği

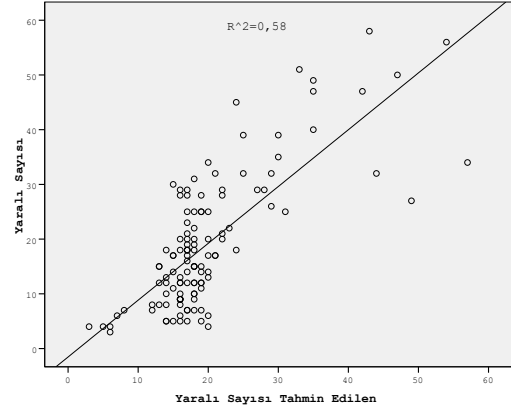
bs: Bağlantı sayısı

hız : Ortalama hız

Poisson regresyonu modelinde bütün bağımsız değişkenler modele katıldığında P değerleri bazı değişkenler için 0.05 değerinden büyük olduklarından anlamlı bulunmamış ve modelden çıkarılmışlardır. Modelleme aşamasında; kavşak sayısı (kavs), şerit sayısı (ss), bağımsız değişken olarak kullanılmış fakat değişkenlerin katsayıları anlamlı çıkmadığından modelden çıkarılmışlardır. Şekil 4.32 ve 4.33’de modelin uyum ve serpilme grafikleri verilmiştir.

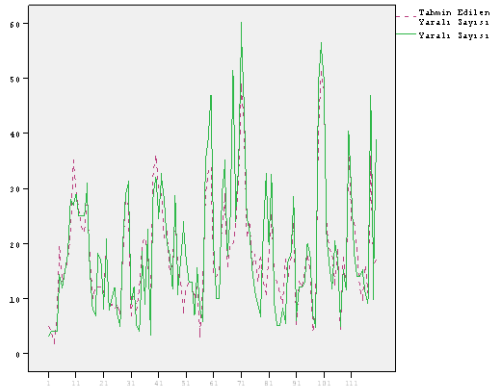


(a)

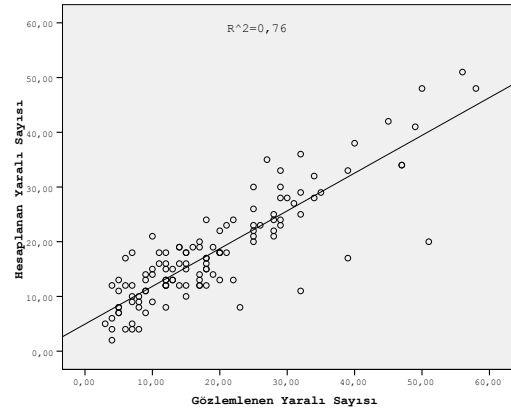


(b)

Şekil 4.32. Model VII-Poisson regresyonu uyum grafiği (a) ve serpilme grafiği (b)



(a)

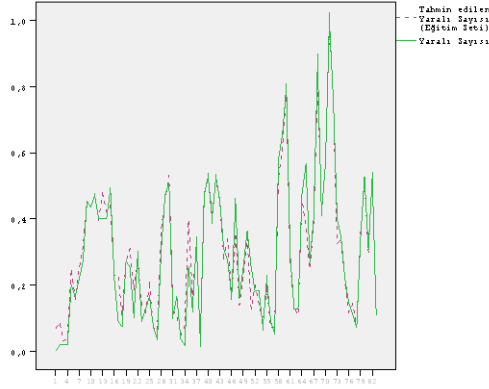


(b)

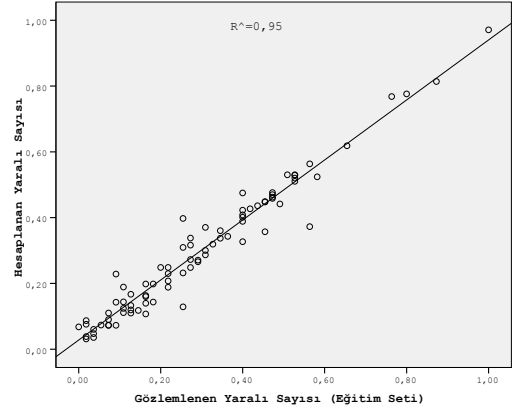
Şekil 4.33. Model VII-Çok değişkenli regresyon uyum grafiği (a) ve serpilme grafiği (b)

Çok değişkenli regresyon sonucu bulunan model;

$$\text{Yaralı Sayısı} = 0.509 \cdot (ks) + 1.031 \cdot (bs) - 0.062 \cdot (kavs) - 0.504 \cdot (\text{hız})$$

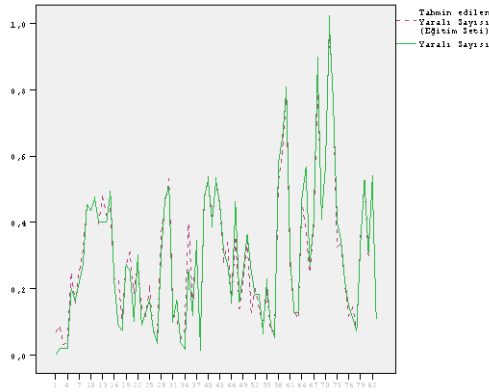


(a)

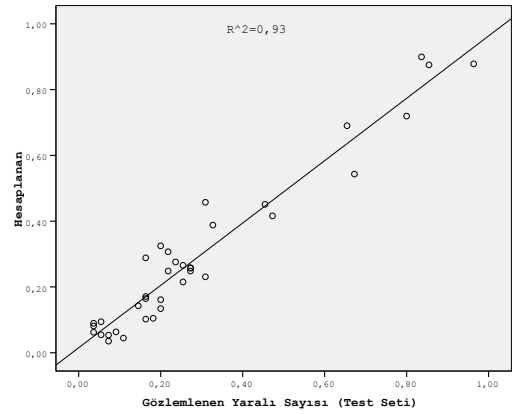


(b)

Şekil 4.34. Model VII-YSA eğitim seti uyum grafiği (a) ve serpilme grafiği (b)



(a)



(b)

Şekil 4.35. Model VII-YSA test seti uyum grafiği (a) ve serpilme grafiği (b)

YSA uygulamasında eğitim ve test setleri için uyum ve serpilme grafikleri Şekil 4.34. ve 4.35.'de sunulmuştur.

Belirlilik katsayısı (R^2) en yüksek olan yöntem YSA ve OKH en küçük olan yöntem yine YSA olmuştur (Çizelge 4.27.)

Çizelge 4.27. Yöntemlerin karşılaştırılması

Yöntem	R²	OKH	Değerlendirme
Poisson Reg.	0.58	0.086	
ÇDR	0.76	0.062	
YSA	0.93	0.018	<i>En uygun yöntem</i>

5. SONUÇ ve ÖNERİLER

Model I uygulamasında, çok değişkenli regresyon yapılmış olup değişkenler arasında çok fazla korelasyon çıkmıştır. Çok yüksek belirlilik katsayısı (R^2) çıkmış, düşük tolerans ve yüksek VIF değerleri bulunmuştur. Bu bilgiler ışığında değişkenlerimiz arasında çoklu bağlantı problemi olduğu anlaşılmıştır. TBA yapılarak iki temel bileşen ile modeller oluşturulmuş ve kaza sayısı açıklanmıştır.

Granger Nedensellik analizinde tablo değeri $F_{(m,n-k)}=3.49$ olarak bulunmuştur. F_{hesap} değeri tablo değerinden büyük olan ve olabilirlikleri 0.05'ten küçük olan değişkenler tespit edilmiştir. Bu durumda olan değişkenler arasında bir nedensellik olduğu anlaşılmıştır. Kaza sayısı ile yük-ton değişkeni arasında tek yönlü bir nedenselliğin olduğu görülmüştür. Karayollarındaki taşınan yükün artış göstermesi araç yoğunluğunu artıracığından kaza sayılarında bir değişime sebep olacağı anlaşılmıştır. Kaza sayısı ile km-taşıt değişkeni, tır sayısı, otobüs sayısı ile karayollarında tüketilen enerji miktarı, AOGT değerleri ile de tek yönlü bir Granger Nedenselliği olduğu anlaşılmıştır. Nedensellik analizleri beklendiği gibi çıkmıştır. Yol-km uzunluğu değerinde ise nedensellik bulunamamıştır. Yolcu-km değeri ile de Granger Nedenselliği bulunamamış bunun nedeninin, taşınan yolcu sayısı değerlerinin 1994 yılından öncesinde sağlıklı gözlemlenememesi olarak yorumlanmıştır.

Granger Nedensellik analizinden sonra aynı verilere eşbütünleşme analizi uygulamasında yapılması uygun olacaktır. Gelecekte bu değişkenlerin birbirlerini nasıl etkilediği bu analiz yöntemi ile belirlenebilir.

Etki tepki fonksiyonun ise kaza sayısına verilecek bir birimlik bir şok bağımsız değişkenleri 7. ve 8. dönemlerinde etkisini kaybetmektedir. Kaza sayısındaki bir birimlik şokun, km-taşıt, yolcu-km, kamyon sayısı, otobüs sayısı, enerji tüketimi, yol-km uzunlukları üzerindeki değişimler analiz edilmiş ve ortalama 7. peryotta etkilerin kaybolduğu gözlemlenmiştir (Şekil 3.11). Yıllık verilerle değil de aylık veya haftalık

verilerle çalışıldığında etki-tepki analiz sonuçları daha verimli olacaktır. Örneğin 2010 yılının nisan ayında trafiğe yeni çıkan araç sayısı, bir önceki aya göre %62 oranında artış göstermiştir. 2009 yılı aylık verilerinde küresel kriz nedeniyle trafiğe kayıtlı araç sayılarında, bir önceki yılın aynı aylarında negatif yönde bir artış olmuştur. Mevsimsel dalgalanmalar, bayram tatilleri, yaz aylarındaki YOGT artışı ve kaza sayılarındaki değişimler etki-tepki analizinde daha iyi anlaşılacaktır.

Varyans ayrıştırmasında ise yolcu-km, km-taşıt, enerji tüketimi, kaza sayısını en çok açıklayan değişkenler olmuştur. Bu modelde AOGT değerinin de kaza sayısını en çok açıklayan değişkenler arasında yer alması beklenirken bu sonuç elde edilememiştir (Çizelge 3.35).

Dört farklı yöntem içinde ARIMA(1,1,0) en başarılı yöntem olmuştur. Fakat bu yöntemde bağımsız değişken olarak sadece AOGT istatistiksel olarak anlamlı çıkmıştır. Bütün değişkenlerin modele katıldığı yöntemlerden biri de VAR(1) modeli olmuştur. Bu modelinde kıstasları diğer yöntemlere göre daha anlamlı çıkmıştır. $R^2=0.67$, OKH=0.086, AIC=-62.30 değerleri ile diğer yöntemler arasında en uygun yöntem olmuştur. YSA yönteminin test setine uygulanması sonucunda $R^2=0.95$ çıkmıştır. YSA yönteminde modele bütün değişkenler katılmıştır. Sonuç olarak YSA, OKH= 0.002 ve AIC=-47.198 ile YSA en başarılı yöntem olarak kabul edilmiştir (Çizelge 4.5).

Gelecek yıllar içinde ne yazıkki kaza sayılarındaki artışın azalacağı düşünülmemektedir. Tahmin sonuçları da bunu destekler niteliktedir. Yıllık 1.2 milyonun üzerinde ölümlü, yaralanmalı ve maddi hasarlı trafik kazası olacaktır. Bölünmüş yolların artması, ortalama hızın 110 km/h'te çıkması gibi etkenlerin kaza sayıları üzerindeki etkisi incelendiğinde 2010 yılı ilk yarısındaki kaza sayılarında bir önceki yıla göre artış olduğu görülmüştür.

Model II uygulamasında, çok değişkenli regresyon yapılmış olup değişkenler arasında yüksek korelasyon çıkmıştır. Düşük tolerans ve yüksek VIF değerleri bulunmuştur. Bu

bilgiler ışığında 8 adet birbirleri ile ilişkili bağımsız değişken yerine 2 adet ilişkisiz bağımsız temel bileşen ile modelleme yapılmıştır. Temel bileşenleri en fazla etkileyen değişken otomobil sayısı olmuştur.

Yaralı sayıları üzerinde Granger Nedenselliği analizi yapılmıştır. F tablo değeri olan $F_{(m,n-k)}=3.49$ ile hesaplanan F değeri karşılaştırılmış, olabilirlikleri $P<0.05$ olan değişkenler tespit edilmiştir. Bu analiz sonucunda trafik kazalarındaki yaralı sayısı ile nüfus değişkeni arasında tek yönlü bir Granger Nedenselliği'nin olduğu ispatlanmıştır (Çizelge 3.33). AOGT değeri ile yaralı sayısı arasında da beklenildiği gibi tek yönlü nedensellik söz konusudur. Yol-km değeri ile yaralı sayısı arasında tek yönlü bir nedensellik olduğu görülmektedir. GSMH, km-taşıt, ağır tonajlı araç sayıları ile yaralı sayıları arasında bir nedensellik bulunamamıştır.

Etki-tepki analizi yapılmış yaralı sayılarının hata terimleri varyansına bir birimlik bir şok verilerek nüfus, yol-km, km-taşıt, ağır tonajlı araç, otomobil sayısı, otobüs, AOGT ve GSMH değişkenleri üzerindeki etkileri analiz edilmiştir. Bu modelde performanslı sonuç elde edilmemiştir. Grafiksel olarak yorumlanması güç bir tablo çıkmıştır. Yaralı sayısının birinci döneminde yapılacak 0.05'lik bir etki, km-taşıt, otobüs, GSMH değişkenlerinin 2. döneminde negatif bir tepki meydana getirmiştir. Değişkenlerimiz için genel olarak 9. dönemde şokların etkisi azalmaktadır (Şekil 3.11).

Varyans ayrıştırması yapılarak yaralı sayısındaki değişimin yani modelin tahmin edilemeyen hareketi gözlemlenmiştir. 10 dönemlik periyot içerisinde yaralı sayısı değişkeni kendi hata varyansının %57.584'ünü açıklamaktadır. Yol-km uzunluğu, km-taşıt, ağır tonajlı araç sayısı, otobüs sayısı, AOGT, GSMH, otomobil ve nüfus değişkenlerinin varyansı açıklama oranları sırasıyla, %9.68, %8.31, %2.24, %6.25, %2.78, %1.98, %3.98, %7.17 olmuştur (Çizelge 3.35).

Model-II'de üç farklı yöntem kullanılmış, YSA yöntemi farklı olarak değerlendirilmiştir. Üç yöntem içinde çok değişkenli regresyon yönteminin belirlilik

katsayısı ($R^2 = 0.98$) ile en büyük, OKH'sı (0.0006) ve AIC bilgi kriteri (-73.198) en küçük olan model olmuştur.

Model-III'ün Model-II'den farkı bağımlı değişkenin ölü sayısı olarak alınmış olmasıdır. Bu uygulamada da Granger Nedenselliği analizi yapılmıştır. F tablo değeri olan $F_{(m,n-k)} = 3.49$ ile hesaplanan F değeri karşılaştırılmıştır. $F_{\text{tablo}} < F_{\text{hesap}}$ olan ve olabilirlikleri $P < 0.05$ olan değişkenler tespit edilmiştir. Bu istatistiksel değerlere sahip değişkenler arasında Granger Nedenselliği olduğu anlaşılmıştır. Bu analiz sonucunda trafik kazalarındaki ölü sayıları ile nüfus, yol-km, km-taşıt, ağır tonajlı araç sayısı, otomobil sayısı, otobüs sayısı, AOGT değişkenleri arasında tek yönlü bir Granger Nedenselliği'nin olduğu ispatlanmıştır. GSMH değeri ile ölü sayısı arasında bir nedensellik bulunamamıştır (Çizelge 3.45).

Etki-tepki analizi bu modelimizde de yapılmıştır. Ölü sayılarının hata terimleri varyansına bir birimlik bir şok verilerek; nüfus, yol-km, km-taşıt, ağır tonajlı araç, otomobil sayısı, otobüs, AOGT ve GSMH değişkenleri üzerindeki etkileri incelenmiştir. 6. dönemde şokların etkisi sona ermiştir. Ölü sayısının birinci döneminde yapılacak 0.014'lük bir şok etkisi, km-taşıt, ağır tonajlı araç, AOGT ve GSMH değişkenlerinin 2. döneminde negatif bir tepki meydana getirmiştir. Değişkenlerimiz için genel olarak 9. dönemde şokların etkisi azalmaktadır (Şekil 3.18).

Varyans ayrıştırması yapılarak ölü sayısı modelinde, tahmin edilemeyen hareketler incelenmiştir. 10 dönemlik peryot içerisinde ölü sayısı değişkeni kendi hata varyansının %62.43'ünü açıklamıştır. Yol-km uzunluğu, km-taşıt, otomobil sayısı, otobüs sayısı, AOGT, GSMH, nüfus değişkenlerinin toplam varyansı açıklama oranları sırasıyla, % 1.45, %11.27, %0.69, %9.97, %7.29, %0.97, %3.84 olmuştur. Bağımlı değişkeni en fazla açıklayan bağımsız değişkenler, km-taşıt, otobüs sayısı, AOGT ve nüfus değerleri olmuştur (Çizelge 3.47).

Model-III’de üç farklı yöntem kullanılmış, YSA yöntemi farklı olarak değerlendirilmiştir. Üç yöntem içinde çok değişkenli zaman serisi VAR(2) yönteminin belirlilik katsayısı ($R^2 = 0.84$) ile en büyük, OKH’sı (0.0008) ile en küçük olan model olmuştur. AIC bilgi kriteri en az olan model ise ARIMA(0,1,1) olmuştur. Bu bilgiler ışığında VAR(2) modeli istatistiksel olarak daha başarılı olduğu görülmüştür (Çizelge 4.13).

Ölü sayılarının tespitinde performans ölçüleri $R^2=0.93$, OKH=0.004, AIC=-23.83 olarak tespit edilmiştir. Yaralı sayıları içinde $R^2=0.96$, OKH=0.0002, AIC=-49.31 olarak tespit edilmiştir. Bu sonuçlar itibarıyla Model II-III ‘te yapılan diğer yöntemler içinde YSA en başarılı yöntem olarak tespit edilmiştir (Çizelge 4.14).

Model IV uygulamasında yaralı sayısının, kaza sayısı, AOGT, GSMH bağımsız değişkenleri ile çok değişkenli regresyon modeli oluşturulmuş olup bağımsız değişkenler arasında yüksek korelasyon çıkmıştır. Çok yüksek belirlilik katsayısı (R^2) çıkmış, düşük tolerans ve yüksek VIF değerleri bulunmuştur. Bu bilgiler ışığında 3 adet birbirleri ile ilişkili bağımsız değişken yerine 1 adet temel bileşen ile modelleme yapılmıştır. Temel bileşenimizi en fazla etkileyen değişkenimiz kaza sayısı olmuştur. Yaralı sayıları üzerinde Granger Nedenselliği analizi yapılmıştır. Bu analiz sonucunda trafik kazalarındaki yaralı sayısı ile kaza sayısı ve AOGT değişkeni arasında tek yönlü bir Granger Nedenselliği’nin olduğu ispatlanmıştır. GSMH ile yaralı sayıları arasında bir nedensellik bulunamamıştır (Çizelge 3.62).

Etki-tepki analizi ile ölü sayılarının hata terimleri varyansına bir birimlik bir şok verildiğinde, kaza sayısı, AOGT ve GSMH değişkenleri üzerinde 4. dönemde etkisini kaybetmektedir. Yaralı sayısının birinci döneminde yapılacak 0.014’lük bir şok etkisi, AOGT ve GSMH değişkenlerinin 2. döneminde negatif bir tepki meydana getirmiştir. Değişkenlerimiz için genel olarak 4. dönemde şokların etkisi azalmaktadır (Şekil 3.25).

Varyans ayrıştırması yapılarak 10 dönemlik periyot içerisinde yaralı sayısı değişkeni kendi hata varyansının %70.26’sını, kaza sayısı %0.82’sini, AOGT %13.97’sini, GSMH

%14.96'sını açıkladığı görülmüştür. Yaralı sayısını en fazla açıklayan değişkenlerimiz, AOGT ve GSMH değerleri olmuştur (Çizelge 3.65).

Model-IV'de üç farklı yöntem kullanılmış YSA yöntemi farklı olarak değerlendirilmiştir. Üç yöntem içinde ARIMA(1,1,0) yöntemi belirlilik katsayısı ($R^2=0.92$) en fazla, AIC bilgi kriteri (-75.93) ise en az olan model olmuştur. OKH'sı 0.00018 ile en az olan model ise ÇDR olmuştur. Bu bilgiler ışığında ARIMA(1,1,0) modeli istatistiksel olarak başarılı olmuştur (Çizelge 4.19).

Model-V'in Model-IV'den farkı bağımlı değişkenin ölü sayısı olarak alınmış olmasıdır. Ölü sayıları üzerinde Granger Nedenselliği analizi yapılmıştır. Bu analiz sonucunda trafik kazalarındaki ölü sayıları ile kaza sayısı, AOGT değişkenleri arasında tek yönlü bir Granger Nedenselliği'nin olduğu ispatlanmıştır. GSMH değeri ile ölü sayısı arasında bir nedensellik bulunamamıştır (Çizelge 3.74).

Etki tepki analizi ölü sayılarının hata terimleri varyansına bir birimlik bir şok; nüfus, kaza sayısı, AOGT ve GSMH değişkenlerinin 4. döneminde şokların etkisini sona erdirmiştir. Ölü sayısının birinci döneminde yapılacak 0.019'luk bir şok etkisi, kaza sayısı, AOGT ve GSMH değişkenlerinin 2. döneminde negatif bir tepki meydana getirmiştir. Değişkenlerimiz için genel olarak 4. dönemde şokların etkisi azalmaktadır (Şekil 3.29).

Varyans ayrıştırması ile 10 dönemlik periyot içerisinde ölü sayısı değişkeni kendi hata varyansının %84.56'sını, kaza sayısı %0.37'sini, AOGT %8.07'sini ve GSMH ise %7.01'ini, açıklamaktadır. Bağımlı değişkenimizi en fazla açıklayan değişkenler, AOGT ve GSMH değerleri olmuştur (Çizelge 3.76).

Model-V uygulamasında da üç farklı yöntem kullanılmış, YSA yöntemi farklı olarak değerlendirilmiştir. Üç yöntem içinde ARIMA(1,1,0) modeli, belirlilik katsayısı ($R^2=0.91$) en fazla, AIC bilgi kriteri (-92.78) ve OKH'sı (0.0016) en az olan model

olmuştur. Bu bilgiler ışığında ARIMA(1,1,0) modeli istatistiksel olarak daha başarılı olmuştur (Çizelge 4.23).

Model IV-V YSA uygulamalarında ölü sayılarının tespitinde performans ölçüleri $R^2=0.97$, OKH=0.0021, AIC=-22.49 olarak tespit edilmiştir (Çizelge 4.24). Yaralı sayıları içinde $R^2=0.92$, OKH=0.00045, AIC=-36.30 olarak tespit edilmiştir (Çizelge 4.25). Bu sonuçlar itibariyle Model IV-V'te kullanılan yöntemler içinde YSA en başarılı yöntem olarak tespit edilmiştir.

Model-VI uygulamalarında model aşırı yayılım göstermemiştir. Bu yüzden yöntem olarak Poisson Regresyonu kullanılmıştır (Çizelge 3.77). Modelde hız değişkeninin katsayısı negatif çıkmıştır. Bunun sebebi olarak güzergah boyunca ilgili tarihlerde yapılan bölünmüş yol çalışmaları gösterilebilir. Ortalama hızın 78-95 km/h arası değer aralığında olması negatif bir katsayıya sahip olabileceğini göstermektedir. Bununla birlikte çok değişkenli regresyon uygulanmış bu yöntemde ilk önce bütün değişkenler modele katılmış sabit gözardı edilmemiştir. Bunun sonucu olarak belirlilik katsayısı (R^2) çok düşük çıkmıştır. Sabit ve parametreleri tutarlı olmayan değişkenler modelden çıkarılmışlardır (Çizelge 3.80). Böylece model istatistiksel olarak daha anlamlı olmuştur. YSA uygulamasında ise bütün değişkenler girdi olarak alınmıştır (Çizelge 3.81). Üç yöntem içinde istatistiksel olarak en anlamlı olan yöntem YSA yöntemi olmuştur (Çizelge 4.26).

Model VII uygulamamızda yaralı sayıları bağımlı değişken olarak alınmış Poisson Regresyonu ile modelleme yapılmıştır. Dağılım parametresi birden küçük olduğundan Negatif Binomial Regresyon yerine Poisson Regresyonu ile modelleme yapılmıştır. YOGT değerinin doğal logaritması modele katılmıştır. Kavşak sayısının katsayısı negatif çıkmıştır. Bu durum kavşak sayısı ne kadar artarsa hızın düşeceği kaza sayısının azalacağı ve yaralı sayısına da negatif bir etki yapması olarak yorumlanmıştır. Şerit genişliği, uzunluk km, değişkenlerinin parametreleri anlamlı çıkmadığından modelden çıkarılmışlardır (Çizelge 3.82). Çok değişkenli regresyon uygulamasında $R^2=0.96$ bulunmuştur (Çizelge 3.84-3.85). YSA uygulamasında bütün değişkenler girdi olarak

alınmıştır. Eğitim setinde $R^2=0.95$ bulunmuştur. Test setinde ise $R^2=0.93$ bulunmuştur (Çizelge 3.86). YSA üç yöntem içinde istatistiksel olarak en anlamlı yöntemdir.

Belirli bölgelerdeki trafik kazaları analiz sonuçlarının Coğrafi Bilgi Sistemleri (CBS) aracılığıyla veri tabanlarına işlenerek, araçlardaki navigasyon cihazlarına uydular aracılığıyla iletilmesi ve sürücülerin özellikle kaza yoğunluğu fazla olan yol kesimlerinde uyarılabilmesi, yazar tarafından planlanan çalışmalar arasındadır.

Özellikle kaza yoğunluğu fazla olan bölgelerimiz için, yol geometrik elemanları, iklim şartları ve yol yüzey karakteristiklerinin, kazalar üzerindeki etkilerinin modellenmesi, ilgili bölgelerdeki kaza sayılarının azaltılması yönünde alınacak tedbirlerin daha gerçekçi olmasını sağlayacaktır.

Ülkemizde meydana gelen trafik kazalarının azaltılmasına yönelik şu ana kadar bir çok çalışma yapılmıştır. Fakat kazalara ait verilerin toplanması, analizi ve çözüm önerilerinin hayata geçirilmesi ile ilgili, kurumsal bir yetki karmaşası olduğu da bir gerçektir. Ülkemizin en önemli problemlerinden biri olan bu problemi daha etkin çözmek için, Ulaştırma Bakanlığı nezdinde kurulacak tam ve tek yetkili bir kuruma ihtiyaç olduğu açıkça görülmektedir. Böylece kaza sayılarının azaltılması yolunda kısa zamanda büyük mesafeler alınacağı muhakkaktır.

KAYNAKLAR

- Akaike, H., 1974, A New Look at the Statistical Model Identification, IEEE Transactions on Automatic Control, AC-19,716-723.
- Akdi, Y., 2003. Zaman Serileri Analizi (Birim Kökler ve Kointegrasyon), Bıçakçılar Kitabevi, Ankara
- Akın, F.,2002..Kalitatif Tercih Modelleri Analizi.Ekin Kitapevi, Bursa.
- Akgüngör, A.P., Doğan, E., 2010, Farklı Yöntemler Kullanarak Geliştirilen trafik Kaza Tahmin Modelleri ve Analizi, Int. J. Eng. Research & Development, Vol. 2 No.1 January 2010
- Albayrak, A.S., 2006. Uygulamalı Çok Değişkenli İstatistik Teknikleri, Asil yayın Dağıtım, Ankara.
- Al-Ghirbal, A. S. Ve Al-Ghamdi, A. S., 2006. Predicting Severe Accident Rates at Roundabouts Using Possition Distribution, TRB Annual Meeting, TRB paper 06-1684
- Alp, M., Cıgızlıoğlu, K.H.,2004. Farklı Yapay Sinir Ağları İle Yağış-Akış İlişkisinin Modellenmesi, İTÜ Dergisi, Mühendislik, Cilt:3, Sayı:1, 80-88, Şubat 2004
- Alpar, R., 2003. Uygulamalı Çok Değişkenli İstatistiksel Yöntemlere Giriş-1, Nobel Yayınları, Ankara.
- Altun İ., Dündar, S., 2005. Yapay SinirAğları ile trafik Akım Kontrolü, Deprem Sempozyumu, 23-25 Mart 2005, Kocaeli,
- Anas, J., ve Ferrara, L. 2004. Detecting Cyclical Turning Points: The ABCD Approach and Two Probabilistic Indicators, Journal of Business Cycle Measurement and Analysis, Vol. 2004, Issue.3, pp. 193-225.
- Arslan, O., 2007. Sismik Veri Analizinde Yapay Sinir Ağı Tabanlı Ana Bileşenler Analizi Modeli, İnternaitonal Earthquake Symposium, 22-26- October, Kocaeli
- Bayazıt, M., 2006. Çok Değişkenli İstatistik Analiz ve Hidrolojide Uygulamaları,Su Vakfı Yayınları, İstanbul.
- Bayata, H.F., Hınıslioğlu, S., 2009. Türkiye'deki Yıllık Trafik Kaza Sayılarının Box-Jenkins Metodu İle Modellenmesi, 10. Ekonomi ve İstatistik Sempozyumu, Erzurum
- Bayram, N., 2004. sosyal Bilimlerde SPSS ile Veri Analizi, Ezgi Kitabevi, Bursa
- Bektaş, S., 2002. Aksaray İli ve Çevre Karayollarında Meydana Gelen Trafik Kazalarının Çok Yönlü Analizi ve Kaza Tahmin Modeli, Gazi Üniversitesi, Fen Bilimleri Enstitüsü, Doktora Tezi, Ankara
- Bell, W.R. 1995. X12-Time Series Modeling Course, Washington D.C.: U.S. Dept. of Commerce, Bureau of The Census.
- Bijleved , F., Commandeur, J., Kopman, J.S., Montfort, K. V., 2005. Multivariate Nonlinear Time Series Modeling of Exposure and Risk in Road Safety Reserch,
- Biçen, C., 2006. Box-Jenkins Zaman Serisi Analiz Yöntemi ile İleri Beslemeli Yapay Sinir Ağları Tahminlerinin Karşılaştırılması, Y. Lisans Tezi, Hacettepe Ün., Sağlık Bilimleri Enstitüsü, Biyoistatistik Programı, Ankara
- Box, G.E.P. ve Jenks, G.M. 1970. Time Series Analysis: Forecasting and Control, San Francisco: Holden-Day.
- Bozkurt, H., 2007. Zaman Serileri Analizi, Ekin Kitapevi, Bursa.

- Can, İ., Şengül S., 2001. Stochastic modeling of mean montly flows of Karasu River, in Turkey, Science Citation Index,
- Camkesen, N., 1998. Trafik Kaza Analizleri ve Tahmin Modelleri, Yıldız Teknik Üniversitesi, Fen Bilimleri Enstitüsü, Doktora Tezi, İstanbul
- Cano S., Getz P., Kropf J., Scott S., Stamas G. 1996. Adjusting for a Calendar Effect in Employment Time Series, U.S. Bureau of Labour Statistics.
- Chatfield, C. and D.L. Prothero, 1973. Box-Jenkins Seasonal Forecasting: Problems in a Case-Study, Journal of the Royal Statistical Society, Series A (General), Vol:136, No:3, ss.295-336
- Cribbins, P.D., Arey J.M., Donaldson J.K., 1967 effect of Selected Roadway and Operational Characteristics an Accidents on Multilane Highways, Highway Research Record 188, Highway Research Board, Washington D.C., s.8-25
- Çömlekci, N., 1989. Temel İstatistik, Bilim ve Teknik Yayınevi, Eskişehir.
- Delen, D., et. al., 2005. Identifying significant predictors of injury severty in traffic accidents using a series of artifical neural ntworks, Accident Analysis&Prevention, 38, 434-444
- Elmas, Ç., 2007. Yapay Zeka Uygulamaları, Seçkin Yayıncılık, Ankara
- Ekonomik Göstergeler, 2007. Türkiye İstatistik Kurumu, Ankara
- Ekonomik Göstergeler, 2008. Türkiye İstatistik Kurumu, Ankara
- Enders, W., 1995. Applied Econometric Time Series, New York, John Wiley ve Sons. Inc.
- Ferrer, A., G., Juan, A., Poncela, P., 2006.Forecasting Traffic Accidents Using Diagggregated Data, İnternatioanl Journal of Forecasting, Vol. 22 203-222
- Filiz, Z., 1995. Trafik Kazalarındaki Ölümelerde Yaş ve Cinsiyet için Risk Guruplarının Poisson Regresyonu Tekniğiyle Belirlenmesi, Fen Bilimleri Enstitüsü, Osmangazi Üniversitesi, Eskişehir.
- Granger, C. W. J.,1969, Investigating Causal Relations by Econometric Models and Cross Spectral Methods, Econometrica, 37, 3, 424-438.
- Greene, W.H., 1993. Econometric Analysis, Macmilan Publishing Company, New York, USA.
- Greibe, P., 2003. Accident Prediction Models For Urban Roads, Accident Analiysis& Prevention 35 (2003) 273-285 Danish Transport Research Institute, Denmark
- Gujarati, D.N., 2004. Basic Econometrics, Fourth Edition, McGraw-Hill Companies, USA
- Hadayeghi, A., Shalaby, A.S., Persaud, B. N. Ve Cheung, C., 2006. Temporal Transferability and Updating of Zonal Level Accident Prediction Models, Accident Analysis and Prevention, 38, 3 , 579-589
- Hadayegni, A., 2002. Accident Prediction Models for Safety Evaluation of Urban Safety Transportation,
- Hadayeghi, A., Shalaby, A.S., Persaud, B. N., 2003. Macro Level Accident Prediction Models For Evaluating The Safety Of Urban Transportation System, March 31, 2003
- Hauer, E., 2001. Overdispersion in Modeling Accidents on Road Section and in Empirical Bayes Estimation, Accident Analysis and Prevention, 33, 6 ,799-808
- Hamzaçebi, C., Kutay, F., 2004. Yapay Sinir Ağları İle Türkiye Elektrik enerjisi Tüketiminin 2010 Yılına Kadar tahmini, Gazi Üniv. Müh. Mim. Fak. Der. Cilt 19. No: 3, 227-233

<http://elohab.freetzi.com/joomla/index.php/makaleler/28-the-mathworks-matlab/77-yapay-snr-alari-ve-matlab-ile-kullanimi.html>

- Jovanis, P., Chang, H., 1986. Modeling the relationship of accidents to miles traveled. Transportation Research Record 1068.
- Kadılar, C., 2000. Uygulamalı Çok Değişkenli Zaman Serileri Analizi, Bizim Büro Yayınevi, Ankara
- Kadılar, C., 2005. SPSS Uygulamalı Zaman Serileri Analizine Giriş, Bizim Büro Basımevi, Ankara.
- Kalaycı, Ş., 2006. SPSS Uygulamalı Çok Değişkenli İstatistik Teknikleri, Asil yayın Dağıtım, Ankara.
- Karayolu Ulaşım İstatistikleri, 2006, T.C. Ulaştırma Bakanlığı, Karayolları genel Müd., Strateji Geliştirme Daire Başkanlığı, Ulaşım Maliyet Etütleri Şb.Müd. Ankara
- Karayolu Ulaşım İstatistikleri, 2007, T.C. Ulaştırma Bakanlığı, Karayolları genel Müd., Strateji Geliştirme Daire Başkanlığı, Ulaşım Maliyet Etütleri Şb.Müd. Ankara
- Karayolu Ulaşım İstatistikleri, 2008, T.C. Ulaştırma Bakanlığı, Karayolları genel Müd., Strateji Geliştirme Daire Başkanlığı, Ulaşım Maliyet Etütleri Şb.Müd. Ankara
- Kibar Türe F., 2008. Trafik Kazaları ve Trabzon Bölünmüş Sahil Yolu Örneğinde Kaza Tahmin Modelinin Oluşturulması, Karadeniz Teknik Üniversitesi, Fen Bilimleri Enstitüsü, Y. Lisans Tezi, Trabzon
- Kumar, V., Robert P. Leona and John N. Gaskins., 1995. Aggregate and Disaggregate Sector Forecasting Using Consumer Confidence Measures, International Journal of Forecasting, 11, s.361-377.
- Kutlar, A., 2000. Ekonomik Zaman Serileri teori ve Uygulama, Gazi Kitabevi, Ankara
- Laitila T., 1993. A Pseudo- R^2 Measure For Limited and Qualitative Dependent Variable Models, Journal of Econometrics, Vol.56. 1993, 341-356
- Lawless, J. F., 1987. Negative Binomial and Mixed Poisson Regression, The Canadian Journal of Statistics, 15, 3, 209-225
- Lee, D.G., Kim, J., 2008. Accident analysis of a surveillance camera system through a frequency based process, WSEAS Transactions on Circuits and Systems, Vol.7 356-367
- Lord, D. ve Persaud, B. N., 2000. Accident Prediction Models With and Without Trend: Application of the Generalized Estimating Equations (GEE) Procedure, transportation Research Board 79. Years Meeting, 9-15 January 2000, Washington, D.C.
- Ljubic, P., Todorovski, L., Lavrac, N., Bullas, J.C., 2004. Time series Analysis of UK Traffic Accident Data, <http://www.fatblokeracing.org/CVREFS/IS2002-Ljubic-Traffic.pdf>
- Lütkepohl, H., 1985, Comparison of Criteria for Estimating the Order of a Vector Autoregressive Process, Journal of Time Series Analysis, 6, 1, 35-52, Düzeltme:1987-373.
- Lütkepohl, H., 1993. Introduction to Multiple Time Series Analysis, Springer-Verlag, Berlin.
- Maddala, G., 1983. Introduction to Econometrics, Macmillan Publishing Company, Second Edition, New York, USA.
- Maddala G.S., 1983. Limited Dependent and Qualitative Variables in Econometrics, Cambridge University Press, Cambridge.

- Miao M. C., Abraham, A., Paprzycki, M., 2005. Traffic Accident Analysis Using Decision Trees and Neural Networks
- Miranda-Moreno, L.F., 2006. Statistikal Models and Methods for Identifity Hazardous Location for Safety Improvements, Doktora Tezi, Waterloo Unv., Kanada
- Mohamed, A. A. A., Radwan E., A., 2000. Modeling traffic accident Occurrence and İnvolment, Accident Analysisi & Prevention, Vol. 32, 633-642
- Motorlu Kara Taşıtları İstatistikleri, 2005. Türkiye İstatistik Kurumu, Ankara
- Motorlu Kara Taşıtları İstatistikleri, 2006. Türkiye İstatistik Kurumu, Ankara
- Motorlu Kara Taşıtları İstatistikleri, 2007. Türkiye İstatistik Kurumu, Ankara
- Montgomery, D.C., Johnson A.L., Gardiner, J.S., 1990. Forecating & Time Series Analysis, McGraw-Hill İnternational Editions, İndustiral Engineering Eeries, Singapore
- Muller, A., 2004. Florida's Motorcycle Helmet Law Repeal and Fatality Rates, American Journal of Public Health, Vol. 84, No. 4
- Öğüt, S. K., İyınam, F.A., 1998, Türkiye'de Trafik Kazalarının Modellenmesi, 2. Uluslararası Ulaşım Sempaozyumu, İstanbul
- Ozgan E., 2008. Karayolu araç Tipi ve Kaza Şekli İle Kaza Sonuçları Arasındaki İlişkilerin Analizi, Gazi Ün., Müh. Mim. Fak. Dergisi, Cilt 23, No:3, 97-104
- Özmen, İ., 1998. Poisson Regresyon Çözümlemeleri, Hacettepe Üniversitesi, Fen Bilimleri Enstitüsü, Doktora Tezi, Ankara
- Öztemel, E., 2006. Yapay Sinir Ağları, Papatya Yayınları, İstanbul
- Persaud, B.N., 1992. Accident Potential Models for Ontario Road Section, Ministry of Transportation on Ontario, Canada
- Persaud, B.N., Dzbik L., 1990. Accident Predistion Models for Freeways, Transportation Research Record, Vol. 2, s. 201-207
- Qin, X., Ivan, J.N. ve Ravishenker, N., 2004. Selecting Exposure Measures in Crash Rate Prediction for Two-lane Highway Segments, Accident Analysis and Prevention, 36,2,183-191
- Quddus, A. M., 2008. Time Series Count Models: An Empirical Applcaiton To Traffic Accidents, Accident Analysis and Prevention, 40, 1732-1741
- Reuters, (2007). Argentine growth above 9 pct two months in a row, 15 Kasım, www.reuters.com/article/economicNews/idUSN1531694920071115
- Sağiroğlu, Ş., Beşdok, E., Erler, M., 2003. Mühendislikte Yapay Zeka Uygulamaları-I Yapay Sinir Ağları, Ufuk Kitap Kırtasiye Yayıncılık Tic. Ltd. Şti., Kayseri
- Salas, J.D., Delleur, J.W., Yevyevich, V., Lane, W.L., 1997. Apllied Modeling Of Hydrologic Time Series, Water Resources Publications, Michigan, USA
- Sawalha, Z. A., 2002. Traffic Accident Modeling: Statistical İssues and Safety Applications, Doctor Tesis, The University of British Colombia, Kanada
- Shankar, V.N., Milton, J. and Mannering, F., 1997. Modeling Accident Frequency as Zeroaltered Probabilty Process: An empirical Inquiry, accident Analysis and Prevention, 29, 6, 829-837
- Schwarz, G., 1978. Estimating the Dimension of a Model, Annals of Statistics, 6, 461-464
- Semiz, M., Fidan E.B., Aydın, Ö., 2007. İstatiksel Modelleme ve Eviews, Dizgi Ofset Matbaacılık, Konya
- Semiz, M., Ocak, Bengü., Aydın, Ö., 2008. Bilgisayarda İstatistik Uygulamaları SPSS, Dizgi Ofset Mat., Konya

- Sevüktekin M.,2007. Nargeleçekenler M., Ekonomik Zaman Serileri Analizi, Nobel Yayın Dağıtım, Ankara
- Shibata, R., 1976, Selection of the Order of an Autoregressive Model by Akaike's Information Criterion, Biometrika, 63,1,117-126
- Sims, C. A., 1972. Money, Income, and Causality, The American Economic Review, 62,540-552
- Sims, C. A., 1980. Macroeconomics and Reality, Econometrica, 48, 1, 1-48.
- Sipahi, B., Yurtkoru, S.E., Çinko, M., 2008, Sosyal Bilimlerde SPSS'le Veri Analizi, Beta Basım Yayın Dağıtım A.Ş., İstanbul
- Şimşek, Halaç, 2003. Birim Kök Testleri ve Trend Sunum Planı, <http://www.deu.edu.tr/userweb/onder.hanedar/dosyalar/Metin.pdf>
- Tarı R., 2008. Ekonometri, Kocaeli Ün. Yayınları Yayın No:172, Avcı Ofset, İstanbul
- Temel, F. Ve Özcebe, H., 2007. Türkiye'de Karayollarında Trafik Kazaları, <http://www.ttb.org.tr/STED/2006/kasim/turkiye.pdf> 07.10.2007
- Trafik İstatistik Yıllığı, 1996. Emniyet Genel Müdürlüğü, Trafik Hizmetleri Başkanlığı Ankara
- Trafik İstatistik Yıllığı, 1997. Emniyet Genel Müdürlüğü, Trafik Hizmetleri Başkanlığı Ankara
- Trafik İstatistik Yıllığı, 1998. Emniyet Genel Müdürlüğü, Trafik Hizmetleri Başkanlığı Ankara
- Trafik İstatistik Yıllığı, 1999. Emniyet Genel Müdürlüğü, Trafik Hizmetleri Başkanlığı Ankara
- Trafik İstatistik Yıllığı, 2000. Emniyet Genel Müdürlüğü, Trafik Hizmetleri Başkanlığı Ankara
- Trafik İstatistik Yıllığı, 2001. Emniyet Genel Müdürlüğü, Trafik Hizmetleri Başkanlığı Ankara
- Trafik İstatistik Yıllığı, 2002. Emniyet Genel Müdürlüğü, Trafik Hizmetleri Başkanlığı Ankara
- Trafik Kaza İstatistikleri Karayolu, 2003, Türkiye İstatistik Kurumu, Ankara
- Trafik Kaza İstatistikleri Karayolu, 2004, Türkiye İstatistik Kurumu, Ankara
- Trafik Kaza İstatistikleri Karayolu, 2005, Türkiye İstatistik Kurumu, Ankara
- Trafik Kaza İstatistikleri Karayolu, 2006, Türkiye İstatistik Kurumu, Ankara
- Trafik Kaza İstatistikleri Karayolu, 2007, Türkiye İstatistik Kurumu, Ankara
- Tramo (Time Series Regression with ARIMA Noise, Missing Observations and Outliers) and SEATS (Signal Extraction in ARIMA Time Series), Instruction for the user, Working Paper 9628, Bank of Spain, Madrid.
- Türkiye İstatistik Yıllığı, 2007. Türkiye İstatistik Kurumu, Ankara
- Türkiye İstatistik Yıllığı, 2008. Türkiye İstatistik Kurumu, Ankara
- www.tuik.gov.tr 2008-2009
- Tombaklar H., 2005. Çocuklar Ve Trafik Kazaları, Selçuk Üniversitesi Kazaları Araştırma, Önleme Ve Uygulama Merkezi
- Ulaştırma İstatistikleri Özeti, 2007, Türkiye İstatistik Kurumu, Ankara
- Ulaştırma İstatistikleri Özeti, 2008, Türkiye İstatistik Kurumu, Ankara

- WHO, 2004, Karayollarında Trafik Kazalarının Önlenmesi Dünya Raporu-I Özet, World Health Organization.
[http://undp.un.org.tr/who/traffic/Road%20Traffic%20-%20World%20Report%20\(Summary\).pdf](http://undp.un.org.tr/who/traffic/Road%20Traffic%20-%20World%20Report%20(Summary).pdf) 21.04.2006
- Willam, A., Preusser, D. F., Ferguson, A.S., Ulmer, Weinstein, H.B., 1973. Fatal crash risk for older drivers at intersections
- Williford W.O., Barton H.E., 1975. A Bayesian Analysis o the “Accident Rate Potential” Parameter, Accident Analysis Prevntion, Vol.7, s.55-62
- Yurtoğlu, H., 2005. Yapay Sinir Ağları Metodololijisi ile Tahmin Modellemesi Bazı Makroekonomik Değişkenler İçin Türkiye Örneği., DPT, Ankara.

ÖZGEÇMİŞ

1976 yılında Erzincan'da doğdu. İlk orta ve lise öğrenimini Erzincan'da tamamladı. 1994 yılında Yıldız Teknik Üniversitesi, İnşaat Fakültesi, Jeodezi ve Fotogrametri Mühendisliği bölümünü kazandı ve aynı bölümden 1998 yılında mezun oldu. Yüksek Lisansını Gebze Yüksek Teknoloji Enstitüsü, Fen Bilimleri Enstitüsü, Jeodezi ve Fotogrametri Mühendisliği bölümünde 2002 yılında tamamladı. 2003 yılında Atatürk Üniversitesi Fen Bilimleri Enstitüsü İnşaat Mühendisliği Bölümünde doktora öğremine başladı.

2000 yılında Gebze Yüksek Teknoloji Enstitüsü'nde Araştırma Görevlisi olarak göreve başladı. 2002 yılında Erzincan Üniversitesi Meslek Yüksekokulu'nda Okutman olarak çalışmaya başladı halen bu görevini sürdürmektedir.