

Türkiye’de 1977-2006 Yılları Arasında Meydana Gelen Aylık Trafik Kazalarının Zamansal Analizi

Temporal Analysis of Monthly Road Traffic Accidents Occurred in Turkey Between 1977-2006

Ahmet ATALAY^{a*}, Ahmet TORTUM^b, Mahir GÖKDAĞ^b

^aAtatürk Üniversitesi, Narman Meslek Yüksekokulu, İnşaat Bölümü, 25530, Narman, Erzurum

^bAtatürk Üniversitesi, Mühendislik Fakültesi, İnşaat Mühendisliği Bölümü, 25240, Erzurum

Geliş Tarihi/Received : 04.12.2011, Kabul Tarihi/Accepted : 04.05.2012

ÖZET

Bu çalışmada 1977-2006 yılları arasında meydana gelen aylık trafik kaza verileri (şehir içi ve şehir dışı toplamı) kullanılarak zaman serisi analiz yöntemi ile modelleme yapılmıştır. Yapılan analizler sonucunda çalışma döneminde kullanılan verilere göre en uygun modelin ARIMA(4,1,4) olduğu belirlenmiştir. Çalışmada en uygun model kullanılarak 2006:01-2007:12 dönemi için aylık kaza tahmini yapılmıştır. Tahmin edilen ve gerçek kaza değerleri kullanılarak regresyon eğrisi çizilmiş ve korelasyon katsayısı ($r=0,9163$) belirlenmiştir. Tahmin değerleri ile gerçek değerler arasında güçlü pozitif ilişki olduğu belirlenmiştir. Ayrıca ARIMA(4,1,4) modelinin başarı ölçütü ortalama karesel hataların karekökü (OKHK) değeri hesaplanarak belirlenmiştir. Çalışma dönemi boyunca en fazla trafik kazası Aralık, Ekim ve Kasım aylarında, en az trafik kazası Şubat, Mart ve Nisan aylarında meydana geldiği belirlenmiştir.

Anahtar Kelimeler: *Trafik kazaları, Zamansal analiz, Zaman serisi analizi.*

ABSTRACT

In this study, modeling with time series analysis was done method by using date of monthly road traffic accidents (RTAs) occurred between 1977-2006. According to results of analysis and date in period of this study, it was determined that ARIMA(4,1,4) was the best model. In the study monthly accident number are forecasted by appropriate model for 2006:1-2007:12 periods. Regression curve is drawn by using the estimated and actual number of RTAs and the correlation coefficient ($r = 0.9163$) were determined. A strong positive relationship was determined between estimated values and actual values. In addition, the success measure of ARIMA(4,1,4) model was identified by calculating value of root mean square errors (RMSE). It was determined that occur up to a traffic accident in December, October and November, at least a traffic accident in February, March and April

Keywords: *Road traffic accidents, Temporal analysis, Time series analysis.*

* Yazışılan yazar/Corresponding author. E-posta adresi/E-mail address : ahatalay@atauni.edu.tr (A. Atalay)

1. GİRİŞ

Tüm gelişmekte olan ülkelerde olduğu gibi Türkiye’de de trafik kazaları önde gelen ölüm nedenlerinden biridir. Çeşitli kanser türleri ve kalp hastalıklarından sonra en fazla ölümler trafik kazalarından kaynaklanmaktadır (Karagöz, 2008). 1977-2006 yılları arasında Türkiye’de toplam 7.051.746 trafik kazası meydana gelmiştir. Sadece 2006 yılında 664 539 kaza meydana gelmiş, 3.365 insan hayatını kaybetmiş ve 135.754 insan yaralanmıştır (Anon., 2006).

Yol güvenliği ile ilgili planlama ve politikaların belirlenmesinde, ileriye yönelik kaza tahminlerin bilinmesi ve modellenmesi gerekmektedir (Akgüngör ve Doğan, 2008). Zheng ve Liu, (2009) yaptıkları çalışmada kaza tahmin modellerini iki grup; zaman serileri tahmin metotları (zaman serileri metodu, Markov zincir metodu, gri model ve sinirsel ağlar) ve nedensellik metotları (senaryo analizleri, regresyon metot ve Bayesiyen ağları) olarak sınıflandırmışlardır. Trafik kaza tahmininde yapılan birçok çalışmada da zaman serileri yöntemi kullanılmıştır (Gandhi ve Hu, 1995; Van den Bossche v.d., 2004; McLeod ve Vingilis, 2008; Quddus, 2008).

Literatürde yapılan trafik kaza tahmin modelleri; toplulaştırılmış (aggregate) model ve ayrıştırılmış (disaggregate) model olarak iki farklı biçimde oluşturulmuştur. Toplulaştırılmış modellerde yıllık kaza sayıları, ayrıştırılmış modellerde aylık kaza sayıları kullanılır (Fridström ve Ingebrigtsen, 1991; Oppe, 1991; Chang ve Graham, 1993; Keeler, 1994; Van den Bossche v.d., 2004; Garcia-Ferrer, de Juan, Poncela, 2006; Quddus, 2008). Literatürde yapılan kaza tahmin modellerinde ayrıştırılmış modellerin toplulaştırılmış modellere göre daha iyi performans gösterdiklerini belirlemişlerdir (Garcia-Ferrer, de Juan, Poncela, 2006; Quddus, 2008).

Bu çalışmada 1977-2006 yılları arasında meydana gelen aylık trafik kaza verileri kullanılarak Box-Jenkins metodu ile modelleme yapıp 2007 yılı için aylık trafik kaza sayıları tahmin edilmesi amaçlanmıştır. Böylece yapılacak trafik güvenliği ile ilgili planlamalarında trafik kaza sayılarının önceden

tahmin edilmesi önemli katkı sağlayacağı umulmaktadır.

2. MATERYAL METOT

Çalışmada kullanılan veriler 1977-2006 yıllarında meydana gelen aylık trafik kaza verileridir. Bu çalışmadaki verilerin 1977-1996 dönemi Öğüt ve İyınam’ın (1998) yaptıkları çalışmadan, 1997-2006 dönemi ise Türkiye İstatistik Kurumu (TUIK)’undan elde edilmiştir. Çalışmada kullanılan veriler Atalay, (2010) tarafından yapılan çalışmada detaylı olarak verilmiştir. Bu çalışmada, çalışma dönemi için kullanılan veriler kentiçi ve kentdışı olarak detaylı istatistik veri elde edilemediğinden dolayı aylık toplam trafik kaza verileri kullanılmıştır.

Çalışmada kullanılan yöntem ise zaman serisi analiz yöntemi olan Box-Jenkins yöntemi kullanılmıştır. Zaman serisi yöntemleri, bir olaya ait geçmişteki verilerin incelenmesi ve belirli eğilimlerin ortaya çıkarılarak ileriye yönelik tahminlerin yapılması temeline dayanmaktadır. Bu yöntemlerin amacı geçmiş gözlem değerlerindeki veri kalıplarını kullanarak istatistiksel modeller oluşturmak ve bu modellerle geleceği tahmin etmektir (De Lurgio, 1998).

Zaman serisi yöntemlerinde geleceğin tahmini yanında geçmiş dönemlerin incelenmiş olması geçmişteki olumlu ve olumsuz gelişmelerin tespit edilmesine nedenlerinin araştırılmasına ve yapılan yanlışların tekrarlanmaması için gerekli tedbirlerin alınmasına da imkan sağlamaktadır. Sayılan bu özellikleri nedeniyle bilhassa orta ve kısa dönem tahminlerine ihtiyaç duyulan her alanda yaygın olarak kullanılmaktadır (Akgül, 2003).

Zaman serileri genel olarak eğer gözlemler zaman içerisinde sürekli ise “sürekli” eğer belli bir zaman aralığındaki gözlemler söz konusu ise “kesikli” zaman serisi adını alırlar (Akgül, 1994). Kesikli zaman serileri birkaç şekilde ortaya çıkabilir. Sürekli bir zaman serisi veri iken zamanın eşit aralıklarında kesikli bir seri olarak tanımlanan dizi, örneklem serisi olarak adlandırılır. Kesikli serilerin bir diğer türü ise anlık değerlere sahip olmayan fakat eşit zaman aralıkları boyunca biriken bütüncül verilerdir.

Örneğin aylık ihracat ve ithalat rakamları ya da yağış miktarları bu tür verilerdir (Sevüktekin ve Nargeleçekenler, 2005).

2.1. Zaman Serisinin Bileşenleri

Bir zaman serisinin gerçek değerlerini trend, mevsimsel değişiklikler, konjoktürel değişiklikler ve rassal değişiklikler olarak dört faktör etkilemektedir. Bu faktörler serinin gerçek değerleri üzerindeki etkileri çarpımsal veya toplamsal olarak ifade edilmektedirler. Toplamsal modelde faktörlerin birbirlerinden bağımsız oldukları, çarpımsal modelde ise faktörlerin birbirleri ile ilişkili oldukları varsayılmaktadır (Bülbül, 2001).

- a) Trend: Zaman serisi gözlem değerinin uzun dönemde artma ya da azalma şeklinde gösterdiği genel eğilime Trend adı verilir. Trend bileşeni zamana bağlı değişken üzerindeki genel eğilime neden olan uzun dönemli etkileri açıklar.
- b) Mevsimsel değişim: Mevsimsel değişimler birbirini izleyen yılların çeyrek yılların mevsimlerin ve ayların aynı dönemlerine ait gözlem değerlerindeki artış ya da azalış şeklinde ortaya çıkan düzenli değişimleri ifade eder. Bu değişimleri doğuran nedenler arasında iklim, alışkanlıklar, sosyal olaylar vb sayılabilir.
- c) Konjoktürel değişim: Konjoktür dalgalanmaları mevsim dalgalanmalarına göre daha uzun bir zaman periyodunda ortaya çıkan dalgalanmalardır. Bu dalgalanmalar ekonomide görülen genel bir büyümenin ardından bir gerileme ve dalgalanma dönemi arkasından tekrar büyümenin görülmesi şeklinde karakterize edilebilir (Çakıcı v.d., 1999).
- d) Rassal değişim: Rassal (düzensiz) değişimler beklenmedik olayların zaman serileri üzerindeki etkisiyle meydana gelen değişimlerdir. Deprem siyasal karışıklıklar savaşlar grev, hava şartlarında görülebilen mevsim normalleri dışındaki değişiklikler buna örnek olarak verilebilir (Özmen, 2003).

2.2. Box-Jenkins (ARIMA) Yöntemi

George Box Jenkins tarafından 1970 yılında geliştirilmiştir. Box-Jenkins yönteminde

temel olarak iki ayrı yöntemin (Oto regresyon ve Hareketli Ortalama) bir kombinasyonu oluşturulmaya çalışılmaktadır. Bu kombinasyonu ifade etmek için kısaca ARMA (Auto Regressive Moving Averages) ifadesi kullanılmaktadır. Ancak sözkonusu modeller sadece durağan serilerde kullanılabildiği için seriye fark alma (differencing) işlemi uygulanması gerekmektedir. Fark alma işlemlerinin sayısını belirleyen “entegresyon indeksi” nin de ifadeye katılması ile birlikte ARIMA (Auto Regressive Integrated Moving Average) modelleri ortaya çıkmaktadır (Fretchling, 2001). ARIMA modellerinde temel yaklaşım incelenen değişkenin bugünkü değerinin geçmiş değerlerinin ağırlıklı toplamı ve rassal şokların birleşimine dayandığı şekilde ifade edilmektedir (Akgül, 2003).

Box-Jenkins modellerinde dikkate alınan durağanlık kovaryans durağanlığıdır. Kovaryans zamana bağlı bir değişkenin farklı dönemlere ilişkin değerleri arasındaki karşılıklı ilişkinin bir ölçüsüdür (Makridakis v.d., 1998). Genel olarak ortalamasıyla varyansı zaman içinde değişmeyen ve iki dönem arasındaki kovaryansı bu kovaryansın hesaplandığı döneme değil de yalnızca iki dönem arasındaki uzaklığa bağlı olan olasılıklı (stokastik) bir süreç durağandır denilir. Belirli bir dönem için gözlenen bir seriyi (Yt) ortaya çıkararak stokastik sürecin durağan olması şartları aşağıdaki gibi ifade edilebilir;

$$\text{Ortalama } E(Y_t) = \mu \quad (1)$$

$$\text{Varyans } \text{var}(Y_t) = E(Y_t - \mu)^2 = \sigma^2 \quad (2)$$

$$\text{Kovaryans } \gamma_k = E[(Y_t - \mu)(Y_{t+k} - \mu)] \quad (3)$$

Burada γ_k , k gecikme değerindeki kovaryans, Yt ile Yt+k arasındaki yani aralarında k dönem farkı olan iki Y arasındaki kovaryanstır. Eğer k=0 ise γ_0 bulunur ki bu Y nin varyansıdır $Y(=\sigma^2)$ eğer k=1 ise γ_1 Ynin ardışık iki değeri arasındaki kovaryanstır. Eğer bir zaman serisi yukarıdaki anlamda durağan değilse durağan olmayan zaman serisi adını alır (Gujarati, 1995).

Durağan olmayıp farkı alınarak durağan hale getirilmiş serilere uygulanan modellere “durağan olmayan doğrusal stokastik modeller” veya kısaca entegre modeller denir. Bu entegre modeller belirli sayıda farkı alınmış serilere

uygulanan AR ve MA modellerinin birleşimidir. Eğer AR modelinin derecesi p, MA modelinin derecesi q ve serinin d kez farkı alınmış bu modele (p, d, q) dereceden “otoregresif entegre hareketli ortalama modeli” denir ve ARIMA (p, d, q) şeklinde gösterilir.

Durağan olmayan Y_t serisinin d inci mertebeden farkı alınarak durağanlaştırıldığında yeni seri W_t olarak tanımlanırsa dönüşüm;

$$W_t = \Delta^d Y_t = (1 - B)^d Y_t \quad (4)$$

şeklinde gösterilmektedir.

Burada;

Δ = Fark alma işlemcisi,
d= Fark derecesi,

$W_t, W_{t-1}, \dots, W_{t-p}$ = Farkı alınmış seriyi göstermektedir.

Bu durumda ARIMA (p, d, q) modelinin genel ifadesi aşağıdaki gibi olmaktadır (Bircan ve Karagöz, 2003).

$$W_t = \phi_1 W_{t-1} + \phi_2 W_{t-2} + \dots + \phi_p W_{t-p} + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \dots - \theta_q \varepsilon_{t-q} \quad (5)$$

ARIMA(p,d,q) modelinin orijinal veri cinsinden genel gösterimi $\delta \neq 0$ varsayımı ile;

$$\Delta^d Y_t = \delta + \gamma_1 \Delta^d Y_{t-1} + \gamma_2 \Delta^d Y_{t-2} + \dots + \gamma_p \Delta^d Y_{t-p} + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \dots - \theta_q \varepsilon_{t-q} \quad (6)$$

şeklinde yapılmaktadır. İncelenen Y_t serisinin durağan olmaması nedeniyle yapılan (4)'nolu dönüşüm ile serinin durağanlığı sağlanmakta ve (5)'deki model ile gösterilmektedir.

ARIMA modelleri ile durağan olmayan zaman serilerinin (p, d, q) mertebesi ile modellenmesi mümkün olmaktadır. Bu aşamada yapılan fark alma işlemi ise durağan olmayan serilerin durağan hale getirilmesinde dönüşüm aracı olarak kullanılmaktadır. Uygulamada yaygın

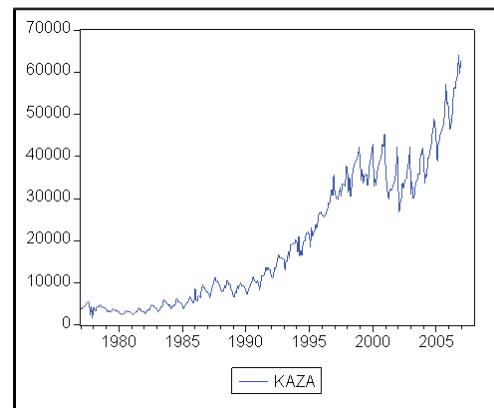
kullanılan fark alma mertebelerinin d=1 ve d=2 olduğu görülmektedir. Serinin mertebesi belirlendikten sonra modeldeki otoregresif terim sayısı p ve gecikmeli hata terim sayısı q belirlenmektedir.

Sonuçta ARIMA modelleri durağan olmayan serilerin durağan olana kadar kaç kere farklarının alındığını gösteren d mertebesine AR terim sayısı p ve MA terim sayısı q'nun ilave edilmesi ile belirlenmekte ve her üç değerın seçilmesinin ARIMA modellerinde en önemli basamak olduğu ifade edilmektedir (Akgül, 2003).

3. BULGULAR

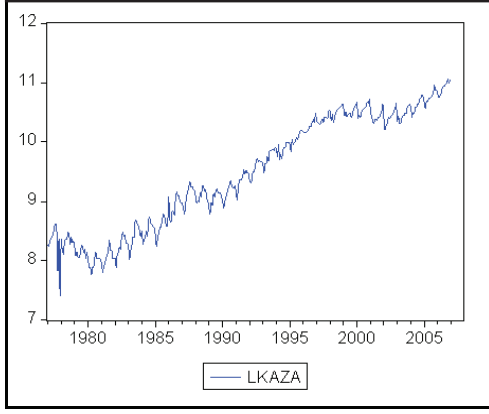
Çalışmanın bu bölümünde 1977-2006 yılları arasında Türkiye’de meydana gelen trafik kaza sayıları (şehir içi ve şehir dışı toplam kaza sayıları) kullanılarak zaman serileri analizi yöntemi ile zamansal analiz yapılmıştır. Çalışmanın bu bölümünde 1977-2006 trafik kaza sayıları (kentiçi + kentdışı toplam kaza sayıları) ve Box-Jenkins yöntemi kullanılarak modelleme yapılmıştır. Zaman serileri analizi için uygulamada EViews paket programı kullanılmıştır.

Öncelikli olarak trafik kaza verilerinin zaman yolu grafiği aşağıdaki gibidir (Şekil 1). Bu grafikten anlaşılacağı üzere seri ortalama ve varyansda durağan olmayan bir seridir. Bunun için ilk olarak trafik kaza verilerinin doğal logaritması alınarak kaza serisinin dönüşümü yapılmış ve LKAZA serisi elde edilmiştir.



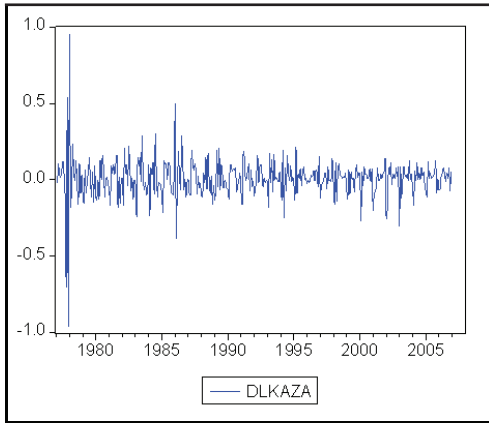
Şekil 1. Kaza zaman serisinin zaman yolu grafiği.

1977: 01-2006: 12 dönemine LKAZA serisinin zaman yolu grafiği Şekil 2’deki gibi elde edilmiştir.



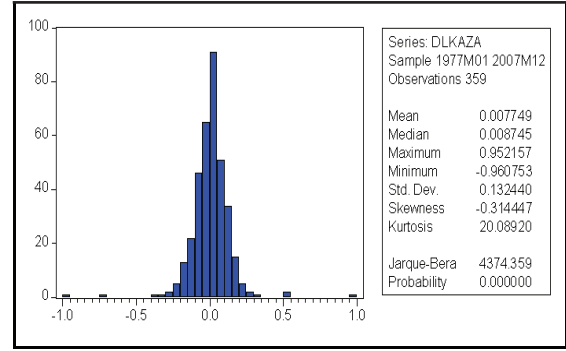
Şekil 2. Doğal logaritması alınmış kaza serisinin zaman yolu grafiği.

LKAZA serisi duran olmadığı zaman yolu grafiğinden de anlaşılmaktadır. LKAZA serisini durağanlaştırmak için 1. dereceden farkı alınmıştır. Farkı alınan seri DLKAZA olarak adlandırılmıştır. DLKAZA serisinin zaman yolu grafiği Şekil 3’deki gibi elde edilmiştir.



Şekil 3. DLKAZA serisinin zaman yolu grafiği.

Elde edilen DLKAZA serisinin tanımlayıcı istatistik değerleri ve histogram Şekil 4’de gösterilmiştir. Bu istatistik değerlerinden ortalama değerinin 0,007749 olması ve bu ortalama değer etrafında serinin salınım yapması serinin durağan olmasına işaret etmektedir.



Şekil 4. DLKAZA serisinin tanımlayıcı istatistikleri.

Zaman serisinin durağanlaştığını istatistiksel olarak belirlemek için birim kök testi yapılmıştır. DLKAZA serisinin birim kök testi sonuçları aşağıda Tablo 1’de verilmektedir.

Tablo 1. DLKAZA serisinin birim kök testi sonuçları.

	t- İstatistiği	Önem
Artırılmış Dickey-Fuller(ADF) test İstatistiği	-5,048653	0.0000
Test Kritik Değerleri	%1	-3.449053
	%5	-2.869677
	%10	-2.571174

Tablo 1’de ADF test istatistiği -5,048653 kritik değerlerden küçük olduğu için seri birim kök içermemektedir. Başka bir ifade ile DLKAZA serisi durağandır denilir. Serinin durağanlığı belirlendikten sonra en uygun modelin belirlenmesi aşamasına geçilir. Deneme yapılan değişik modellerin katsayılarının anlamlılığı dikkate alınarak aşağıda Tablo 2’de verilen DLKAZA serisi için anlamlı dört model belirlenmiştir.

Tablo 2. DLKAZA serisi için alternatif model tahmin sonuçları.

	ARMA(2,1)	ARMA(3,3)	ARMA(4,4)	MA(4)
R2 (Belirleme Katsayısı)	0.197623	0.230763	0.429499	0.225260
\bar{R}^2 (Düzeltilmiş Belirleme Katsayısı)	0.190804	0.217538	0.416308	0.216506
S.E. of regression	0.119375	0.117546	0.101667	0.117230
Hata Kalıntı Kareler Toplamı (Sum squared resid)	5.030.423	4.822.169	3.576.302	4.864.937
Olabilirlik Oranı (Log likelihood)	2.542.473	2.605.617	3.123.843	2.626.788
AIC(Akaike Bilgi kriteri)	-1.401.946	-1.424.504	-1.709.207	-1.435.536
SIC (Schwarz Bilgi Kriteri)	-1.358.498	-1.348.311	-1.611.041	-1.381.451
F-İstatistiği	2.898.087	1.744.941	3.256.050	2.573.183
Prob.(F-İstatistiği)	0.000000	0.000000	0.000000	0.000000

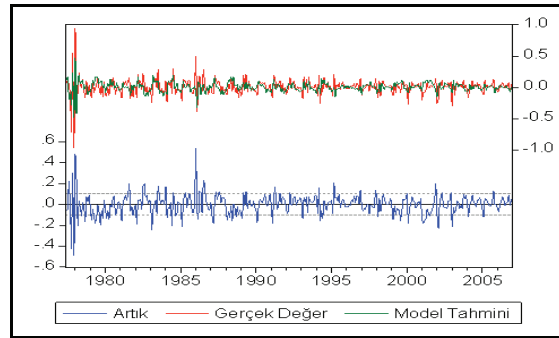
En uygun modeli belirlemek için aşağıdaki kriterler dikkate alınmalıdır;

1. Öncelikle tahmin edilen parametrelerin anlamlı olmasıdır. Yukarıdaki alternatif modellerin parametreleri anlamlıdır,
2. Daha sonra yüksek bir belirleme katsayısına (veya düzeltilmiş belirleme katsayısına) sahip olmalıdır,
3. Modelin F istatistiğinin anlamlı olmasıdır,
4. AIC ve SIC bilgi kriterleri küçük olmalıdır,
5. Hata kalıntı kareler toplamı küçük olmalıdır,
6. Olabilirlik oranı mümkün olduğunca yüksek olmalıdır.
7. Hataların temiz-dizi olduğu belirlenmelidir. Bu kriterler bir arada göz önüne alındığında DLKAZA serisi için en uygun modelin ARMA(4,4) yapısı olduğu ortaya çıkmaktadır. Ayrıca hataların temiz dizi olduğunu belirlemek için Breusch-Godfrey serisel korelasyon LM testi uygulanmıştır.

Tablo 3. LM testi sonuçları.

Breusch-Godfrey Serisel Korelasyon LM Test:			
F-istatistiği	0.694479	Önem	0.500033
Obs*R-kare	1.427458	Önem	0.489814

Tablo 3'den anlaşılacağı üzere LM testi önem değeri (Prb.= 0,500033) olması modelin hatalarının serisel olarak korelasyon olmadığına işaret etmektedir. Başka bir ifade ile hata teriminin temiz dizi olduğunu belirtmektedir. Bu değerlerin zaman yolu grafiği Şekil 5'deki gibi belirlenmiştir.



Şekil 5. Modelin artık, gerçek değer ve modelin tahminlerinin grafiği.

En iyi model belirlenip ve hata terimlerinin de temiz dizi olduğu belirlendikten sonra öngörü yapma aşamasına geçilir. En uygun model ARIMA(4, 1, 4) modelinin katsayıları Tablo 4'deki gibi elde edilmiştir. Bu katsayılar kullanılarak DLKAZA serisi için oluşturulan modelin bağıntısı aşağıdaki gibidir.

$$DLKAZA_t = 0,007903 + 0,384325DLKAZA_{t-1} + 0,931679DLKAZA_{t-2} - 0,595341DLKAZA_{t-3} - 0,344234DLKAZA_{t-4} + \varepsilon_t - 1,021131\varepsilon_{t-1} - 0,615780\varepsilon_{t-2} + 1,020090\varepsilon_{t-3} - 0,157005\varepsilon_{t-4}$$

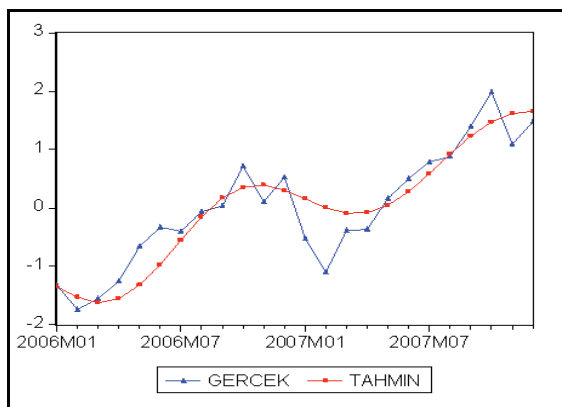
Tablo 4. ARIMA(4,1,4) modelinin katsayıları.

Değişken	Katsayı	Std. Hata	t-İstatistiği	Önem
C	0.007903	0.001961	4.030384	0.0001
AR(1)	0.384325	0.077347	4.968863	0.0000
AR(2)	0.931679	0.069596	13.38700	0.0000
AR(3)	-0.595341	0.060275	-9.877041	0.0000
AR(4)	-0.344234	0.067503	-5.099561	0.0000
MA(1)	-1.021131	0.070116	-14.56344	0.0000
MA(2)	-0.615780	0.085248	-7.223434	0.0000
MA(3)	1.020090	0.068895	14.80642	0.0000
MA(4)	-0.157005	0.069814	-2.248906	0.0251

Bu model 1977: 01-2006: 12 verileri kullanılarak oluşturulmuştur. Oluşturulan model ile 2006: 01- 2007: 12 dönemi için öngörü yapıp ve öngörü değerleri Tablo 5’de gösterilmiştir. Yapılan öngörü değerleri öncelikle DLKAZA değerleridir. Orijinal kaza değerlerini elde etmek için önce öngörü DLKAZA kaza değerlerini bir önceki değer ile toplayarak LKAZA serisinin öngörü değerleri hesaplanmıştır. Daha sonra LKAZA serilerinin antilogaritması alınarak orijinal kaza sayıları hesaplanmıştır (Tablo 5). 2006: 01- 2007: 12 dönemi için Tablo 5’de gösterilen gerçek ve tahmin kaza değerlerinin zamansal akış diyagramı Şekil 6’da verilmiştir.

Tablo 5. LKAZA Serisi için ARIMA (4, 1, 4) modelinin öngörü sonuçları.

AYLAR	GERÇEK KAZA DEĞERİ	DLKAZA	LKAZA	TAHMİN KAZA DEĞERİ
Oca.06	49434	-0,031540	10,845165	51286
Şub.06	46476	-0,014545	10,830620	50545
Mar.06	47804	-0,008128	10,822492	50136
Nis.06	49936	0,006005	10,828497	50438
May.06	54227	0,019179	10,847676	51415
Haz.06	56572	0,027740	10,875416	52861
Tem.06	56055	0,032681	10,908097	54617
Ağu.06	58468	0,029847	10,937944	56272
Eyl.06	59164	0,023730	10,961674	57623
Eki.06	64055	0,012851	10,974525	58368
Kas.06	59652	0,002956	10,977481	58541
Ara.06	62696	-0,006365	10,971116	58169
Oca.07	55173	-0,010583	10,960533	57557
Şub.07	51054	-0,011253	10,949280	56913
Mar.07	56157	-0,006485	10,942795	56545
Nis.07	56285	0,000443	10,943238	56570
May.07	60113	0,009398	10,952636	57104
Haz.07	62511	0,016687	10,969323	58065
Tem.07	64556	0,022066	10,991389	59361
Ağu.07	65244	0,023208	11,014597	60755
Eyl.07	68937	0,021236	11,035833	62059
Eki.07	73129	0,015831	11,051664	63049
Kas.07	66759	0,009385	11,061049	63643
Ara.07	69516	0,002652	11,063701	63812

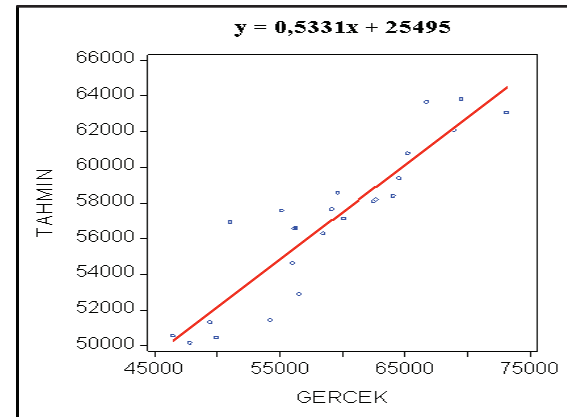


Şekil 6. Gerçek ve tahmin kaza değerlerinin zamansal akış diyagramı.

Çalışmada öngörü başarısı en yüksek olan 2006:1-2007:12 dönemi için yapılan öngörü modelidir. 2006:1-2007:12 dönemi için yapılan öngörünün başarı ölçütü olarak ortalama karesel hataların karekökü (OKHK) değeri 0,046237’dir.

En küçük OKHK değeri modelin tüm olarak uyum iyiliğinin bir ölçüsü olarak kabul edilmektedir. Öngörü başarısını karşılaştırmak amacıyla kullanılan tüm kriterler de hesaplanan değerlerin küçük olması modelin öngörüsünün yüksek olduğunu göstermektedir.

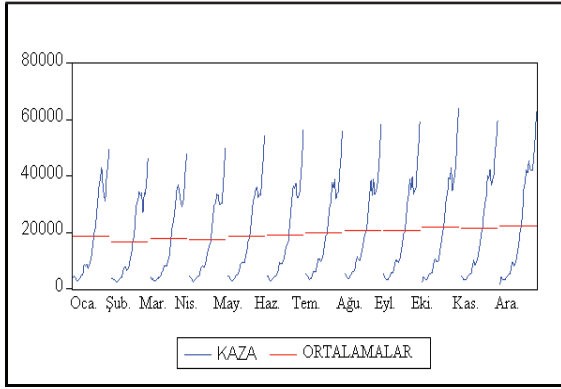
Gerçek ve tahmin kaza sayısı değerleri arasındaki ilişkiyi belirlemek için 2006: 1-2007: 12 dönemi için serpilme diyagramı grafiği çizilmiş ve regresyon eğrisi oluşturulmuştur (Şekil 7).



Şekil 7. Gerçek ve tahmin kaza değerlerinin serpilme diyagramı ve regresyon eğrisi.

Gerçek ve tahmin kaza sayısı değerleri arasındaki korelasyon katsayısı $r = 0,9163$ olarak belirlenmiştir. Buna göre gerçek ve tahmin kaza sayısı değerleri arasında pozitif ve güçlü bir ilişki olduğu belirlenmiştir. Şekil 7’de regresyon çizgisinin denklemi $y = 0,5331x + 25495$ ve belirleme katsayısı $R^2 = 0,8396$ olarak belirlenmiştir.

Trafik kazalarının aylara göre değişimlerini gösteren grafik Şekil 8’deki gibi oluşturulmuştur. Bu grafikten Ekim, Kasım ve Aralık aylarında en fazla kaza meydana gelmiştir. Şubat, Mart ve Nisan aylarında en az kaza meydana gelmiştir.



Şekil 8. 1977-2006 yıllarındaki trafik kazaların aylara göre değişimi.

4. SONUÇLAR

Yol güvenliği ile ilgili planlama ve politikaların belirlenmesinde, ileriye yönelik kaza tahminlerinin bilinmesi ve modellenmesi gerekmektedir. Bu amaç doğrultusunda bu çalışma yapılmıştır.

Bu çalışmada 1977-2006 yıllarında meydana gelen aylık trafik kaza verileri kullanılarak Box-Jenkins metodu ile modelleme yapılmıştır. Bu çalışmada aylık trafik kaza verilerine göre en uygun modelin ARIMA(4,1,4) olduğu belirlenmiştir. Oluşturulan model ile 2006:1-2007:12 dönemi için oluşabilecek kaza sayıları tahmin edilmiştir. Kurulan modelin tahmin değerleri gerçek değerlere oldukça yakın olduğu belirlenmiştir. Gerçek ve tahmini kaza değerleri arasında korelasyon katsayısı 0,9163 olarak oldukça yüksek korelasyon olduğu belirlenmiştir.

Ayrıca çalışma döneminde trafik kazalarının aylık değişimleri incelenmiştir. Aylık ortalama trafik kazası sayılarına göre en fazla trafik kazası Ekim, Kasım ve Aralık aylarında meydana geldiği belirlenmiştir. En az trafik kazası ise Şubat, Mart ve Nisan aylarında meydana geldiği belirlenmiştir. Bunun nedeni öncelikli olarak yaz aylarından kışa geçiş ayları olduğundan sürücüler kışa hazırlıksız olmalarından dolayı, ayrıca son on yılda dini bayramların bu aylara tekabül etmesidir. Ayrıca Şubat, Mart ve Nisan aylarında tatil ve bayram denk gelmemesi sonucu şehirlerarası trafikte yoğunluk yaşanmaması kazaların sayısında azalma meydana getirmektedir.

Bu çalışmanın sonucunda Box-Jenkins metodu ile trafik kaza tahmin modelleri oluşturulabilirliği belirlenmiştir. Bundan sonra yapılacak çalışmada kent içi ve kent dışı trafik kazaları verileri elde edildiği takdirde hem kent içi hemde kent dışı yerler için ayrı ayrı zaman serisi yöntemi kullanılarak modelleme yapılacaktır.

5. KAYNAKLAR

Akgül, I. 1994. Zaman Serisi Analizi ve Öngörü Modelleri, Öneri, Marmara Üniversitesi Sosyal Bilimler Enstitüsü Dergisi, Cilt 1, Sayı 1, s. 52-69.

Akgül, I. 2003. Zaman Serilerinin Analizi ve ARIMA Modelleri, Der Yayınları, İstanbul.

Akgüngör, A.P., Doğan, E. 2008. Smeed ve Andreassen kaza modellerinin Türkiye uygulaması: Farklı senaryo analizleri, Gazi Üniv., Müh. Mim. Fak. Der. 23/4, 821-827, Ankara.

Anonim, 2006. Trafik Kaza İstatistikleri, Türkiye İstatistik Kurumu, Ankara, Türkiye. (www.tuik.gov.tr).

Atalay, A. 2010. Türkiye'deki trafik kazalarının mekansal ve zamansal analizi, Doktora Tezi, Fen Bilimleri Enstitüsü, Atatürk Üniversitesi, Erzurum.

Bircan, H. ve Karagöz, Y. 2003. Box-Jenkins Modelleri ile aylık döviz kuru tahmini üzerine bir uygulama, Kocaeli Üniversitesi Sosyal Bilimler Enstitüsü Dergisi, Cilt 6, Sayı 2, s. 49-62.

Bülbül, S.E. 2001. Çözümsel İstatistik. Alfa Basın Yayın Dağıtım Ltd. Şti. İstanbul.

Chang, B.-H., Graham, J.D. 1993. A new method for making interstate comparisons of highway fatality rates. Accident Analysis and Prevention 25 (1), 85-90.

Çakıcı, M., Oğuzhan, A. ve Özdil, T. 1999. Temel İstatistik, İzmir.

De Lurgio, S.A. 1998. Forecasting Principles and Applications, Irwin/Mc Graw-Hill, Boston.

Fretchling, D.C. 2001. Forecasting Tourism Demand: Methods and Strategies, Butterworth-Heinemann.

- Fridström, L., Ingebrigtsen, S. 1991. An aggregate accident model based on pooled, regional time-series data. *Accident Analysis and Prevention* 23 (5), 363–378.
- Gandhi, U. N. & Hu, S. J. 1995. Data-based approach in modeling automobile crash, *International Journal of Impact Engineering*, 16 (1), 95–118.
- Garcia-Ferrer, A., A. de Juan and Poncela, P. 2006. Forecasting traffic accidents using disaggregated data. *International Journal of Forecasting* 22 (2), 203-222.
- Gujarati, D.N. 1995. *Basic Econometrics*, International Edition, McGraw-Hill Inc., Literatür Yayıncılık, İstanbul.
- Karagöz K. 2008. Bölünmüş yol uygulamasının trafik kazaları üzerindeki etkisi: istatistiksel bir yaklaşım, 1.Karayolu Ulusal Kongresi, Ankara.
- Keeler, T.E. 1994. Highway safety, economic behavior, and driving enforcement. *The American Economic Review*, 84 (3), 684–693.
- Makridakis, S., Wheelwright, S.C., ve Hyndman, R.H., 1998. *Forecasting: Methods and Applications*, Third Edit., John Wiley and Sons, Inc., New York.
- McLeod, A. I. & Vingilis, E. R. 2008. Power computations in time series analyses for traffic safety interventions. *Accident Analysis and Prevention*, 40 (3), 1244–1248.
- Oppe, S. 1991. Development of traffic and traffic safety: global trends and incidental fluctuations. *Accident Analysis and Prevention* 23 (5), 413–422.
- Öğüt, K.S. ve İyınam, A.F. 1998. Türkiye’de Trafik Kazalarının Modellenmesi, 2. Uluslararası Ulaşım Sempozyumu, İstanbul, ss. 441-449.
- Özmen, A. 2003. İstatistik içinde zaman serisi çözümlemesi, Editör: Ali Fuat Yüzer, Anadolu Üniversitesi Yay. No:1448, Eskişehir.
- Quddus, M. A. 2008. Time series count data models: An empirical application to traffic accidents, *Accident Analysis and Prevention*, (40), 1732–1741.
- Sevüktekin, M., ve Nargeleçekenler, M. 2005. *Zaman Serileri Analizi*, Nobel Yayın Dağıtım Ltd., Ankara.
- Van den Bossche, F., Wets, J. & Brijs, T. 2004. A regression model with ARIMA errors to investigate the frequency and severity of road traffic accidents. In *Proceedings of the 83rd annual Meeting of the Transportation research Board* (pp. 11–15), Washington, DC, USA, January 2004.
- Zheng X., Liu M. 2009. An overview of accident forecasting methodologies, *Journal of Loss Prevention in the Process Industries*, (22), 484–491.