KONYA İLİ SICAKLIK VERİLERİNİN ÇİFTDOĞRUSAL ZAMAN SERİSİ MODELİ İLE MODELLENMESİ

İsmail KINACI[1], Aşır GENÇ[1], Galip OTURANÇ[2], Aydın KURNAZ[2], Şefik BİLİR[3]
[1]Selçuk Üniversitesi, Fen-Edebiyat Fakültesi İstatistik Bölümü Kampüs Konya
[2]Selçuk Üniversitesi, Fen-Edebiyat Fakültesi Matematik Bölümü Kampüs Konya
[3]Selçuk Üniversitesi, Mühendislik-Mimarlık Fak. Makine Müh. Bölümü Kampüs Konya

Özet: *Gerçek hayatta karşılaşılan bir zaman serisi verisi için genelde ilk önce lineer bir zaman serisi modeli düşünülmektedir. Çünkü lineer modeller ile istatistiksel sonuç çıkarımlarda ve öngörülerde bulunmak oldukça kolaydır. Ancak bazı durumlarda lineer zaman serisi modelleri seriyi açıklamakta yetersiz kalmakta ve lineer olmayan zaman serisi modellerinin kullanılması kaçınılmaz olmaktadır. Çalışmada Konya ili için Devlet Meteoroloji Genel Müdürlüğü (DMGM)'nden alınan 1997-2004 yılları arasındaki günlük ortalama sıcaklık değerlerinden yararlanılarak elde edilen bir yıllık günlük ortalama sıcaklık verisi çiftdoğrusal zaman serisi modeli ile modellenmeye çalışılmış ve öngörü performansı açısından lineer zaman serisi modelleri ile karşılaştırılmıştır.*
Anahtar Kelimeler: *Lineer zaman serisi, Lineer olmayan zaman serisi, çiftdoğrusal moel, parametre tahmini*

1. GİRİŞ

Zamana bağlı verilerin analizinde son zamanlarda zaman serisi tekniklerinden önemli ölçüde faydalanılmaktadır. Zaman serisi analizlerinde amaç öncelikle seri için bir model belirleyip ileriye yönelik güvenilir öngörüler elde etmektir. Bu öngörülerin güvenilir olabilmesi, seri için önerilen modelin bazı varsayımları sağlamasına bağlıdır. Temel olarak bir zaman serisi verisi için öncelikle lineer ve lineer olmayan zaman serisi modelleri olmak üzere iki tip model düşünülebilir. Uygulamada genellikle bir zaman serisi verisi için lineer bir model düşünülmekte ve bu model yardımıyla öngörü değerleri bulunmaktadır. Bunun sebebi lineer modeller ile tahmin ve öngörülerde bulunmanın lineer olmayan zaman serisi modellerine göre daha kolay olmasıdır. Ancak bazı durumlarda seri için lineer modeller yeterince açıklayıcı bulunmamakta ve bu durumda lineer olmayan modellerin kullanılması kaçınılmaz olmaktadır. Bununla birlikte bir zaman serisi verisi için lineer olmayan bir zaman serisi modelini kullanarak tahmin ve öngörülerde bulunmak zor ve zahmetli uğraşlar gerektirebilmektedir. Bu sebeple bir zaman serisi verisi için lineer olmayan bir model kullanmadan önce serideki lineer olmayan yapının kanıtlarının mutlaka iyi bir şekilde ortaya konması gerekmektedir. Serideki lineer olmayan yapının belirlenebilmesi için literatürde grafiksel ve testler olarak birçok yöntem bulunmaktadır [4, 5, 6].
Bu çalışmada amaç Konya ili günlük ortalama sıcaklık verileri için lineer bir zaman serisi modeli ile lineer olmayan modeller sınıfından olan çiftdoğrusal (*BL*) zaman serisi modeli karşılaştırılacaktır. Ancak seriyi çiftdoğrusal model ile modellemeden önce, seriye lineerlik testinin yapılması ve lineerliğin reddedilmesi gerekmektedir. Bundan sonraki kesimde, serideki lineer olmayan yapının belirlenmesinde kullanılabilen Lagrange çarpanları yöntemi tanıtılacaktır.

2. LAGRANGE ÇARPANLARI LİNEERLİK TESTİ

Burada bir zaman serisinde lineer olmayan yapının olup olmadığını belirleyebilmek için kullanılabilen Lagrange çarpanları testi diğer bazı testlerden farklı olarak lineer bir zaman serisi modelini lineer olmayan bir alternatif modeline karşı test etmektedir. Bu sebeple bu testin yapılabilmesi için öncelikle zaman serisi verisi için uygun bir ARMA modelinin belirlenmesi gerekmektedir. ARMA modelleri Box-Jenkins modelleri olarak da bilinir ve bu modeller ayrıntılı bir şekilde Box ve ark. [1] tarafından incelenmiştir. LM istatistiği hemen hemen her önemli lineer olmayan modele karşı lineer modelleri test etmek için oluşturulabilir. LM testi için hipotez,

$$\begin{aligned} H_0 &: \text{"}\mathit{Model\ lineerdir}\text{"} \\ H_1 &: \text{"}\mathit{Model\ lineer\ de}\text{ğ}\mathit{ildir}\text{"} \end{aligned} \tag{1}$$

şeklinde oluşturulur. Sadece, yokluk hipotezinde verilen lineer modelin tahmin edilmesi ile LM testinin işlem adımları gerçekleştirilebilmektedir. Bu, tek bir alternatif için oluşturulan bir LM testinin çeşitli alternatiflere karşı kullanılabileceği anlamına gelmektedir.
Yukarıdaki hipotezin test edilebilmesi için hesaplanması gereken Lagrange çarpanları test istatistiği

$$LM = \hat{\sigma}^2 \left(\sum_{t=1}^{T} \hat{z}_{1t} e_t \right)' \left(\hat{M}_{11} - \hat{M}_{12} \hat{M}_{22}^{-1} \hat{M}_{21} \right)^{-1} \left(\sum_{t=1}^{T} \hat{z}_{1t} e_t \right) \tag{2}$$

şeklinde tanımlanmaktadır. Yokluk hipotezi altında LM istatistiği asimptotik olarak p serbestlik dereceli χ^2 dağılımına sahiptir (burada p yokluk hipotezi altında sınırlanan parametre sayısıdır) [4, 6].

3. ÇİFTDOĞRUSAL (BL) MODEL

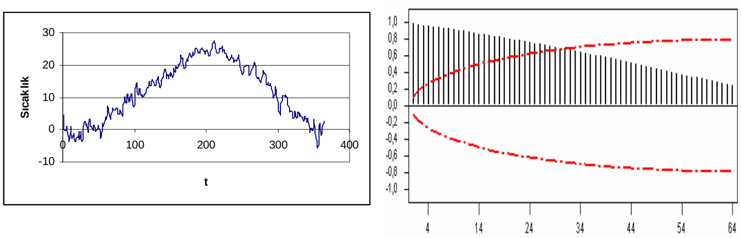
Çift doğrusal modellerin genel yapısı,

$$X_t = \alpha + \sum_{i=1}^{p} a_i X_{t-i} + \sum_{i=1}^{q} b_i \varepsilon_{t-i} + \sum_{i=1}^{P} \sum_{j=1}^{Q} c_{ij} X_{t-i} \varepsilon_{t-j} + \varepsilon_t \tag{3}$$

şeklindedir. Bu model kısaca $BL(p,q,P,Q)$ şeklinde gösterilmektedir. Çift doğrusal modeller, hata terimleri verildiğinde X'lere göre, X'ler verildiğinde hata terimlerine göre lineer bir modeldir. Fakat aynı anda her ikisine göre lineer olmayan bir modeldir. $ARMA(p,q)$ modeli, (3) ile verilen çift doğrusal modelde $P=0$ ve $Q=0$ alınarak elde edilebilir. Yani, ARMA modelleri çift doğrusal modellerin özel bir halidir [2, 3, 5].

4. UYGULAMA

Bu çalışmada Konya ili için Devlet Meteoroloji Genel Müdürlüğü (DMGM)'nden alınan 1997-2004 yılları arasındaki günlük ortalama sıcaklık değerlerinden yararlanılarak elde edilen 1 yıllık günlük ortalama sıcaklık verileri kullanılmıştır. günlük ortalama sıcaklık verilerin zamana göre grafiği ve otokorelasyon değerleri Şekil 1'de verildiği gibi elde edilmiştir.

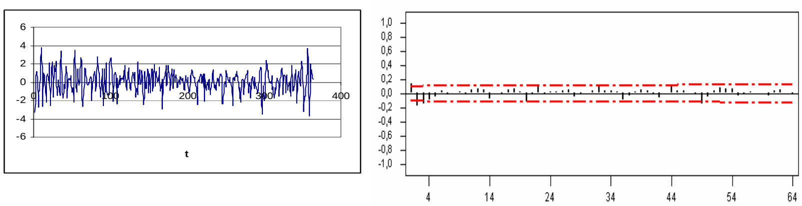


(a) **(b)**

Şekil 1. Günlük ortalama sıcaklık serisi (a) ve otokorelasyon değerleri (b)

Şekil 1.a'dan görüldüğü gibi serinin ortalaması önce artarak daha sonra da azalarak gitmektedir. Yani serimizin ortalaması zamana bağlı olarak değişmektedir. Bu durum bize serinin durağan olmadığını göstermektedir. Çünkü durağan bir zaman serisinin ortalaması ve varyansı zaman içinde değişmemektedir. Serinin durağan olmadığı Şekil 1.b' de verilen otokorelasyon fonksiyonunun grafiğinden de kolaylıkla anlaşılabilmektedir. Şekil 1.b'den de görüldüğü gibi otokorelasyon değerleri, gecikme değeri arttıkça oldukça yavaş bir şekilde azalmaktadır. Bu da serinin durağan olmadığını göstermektedir.

Seriye bir derece fark uygulayarak elde edilen yeni serinin zamana karşı nokta diyagramı ve otokorelasyon değerleri Şekil 2'de verildiği gibi elde edilmiştir.



(a) **(b)**

Şekil 2. Fark operatörü uygulanmış seri (a) ve otokorelasyon değerleri (b)

Şekil 2.a'dan yeni serinin ortalamasının ve değişim aralığının zaman içinde fazla değişmediği görülmektedir. Yani serinin ortalaması ve varyansı zaman içinde hemen hemen değişmemektedir. Bu durum bize serinin durağan olduğunu işaret etmektedir. Ayrıca bu serinin Şekil 2.b'de verilen otokorelasyon katsayılarına bakıldığında, otokorelasyon katsayılarının tamamına yakınının güven bandının içinde olduğu yani istatistiksel olarak anlamsız olduğu anlaşılmaktadır. Bu da bize yeni serinin durağan olduğunu göstermektedir.
Uygulamanın bundan sonraki kısımlarında durağan hale getirilen yeni seri üzerinde çalışılacaktır ve ilk baştaki günlük ortalama sıcaklık serisinden bahsedilirken orijinal seri tabiri kullanılacaktır.

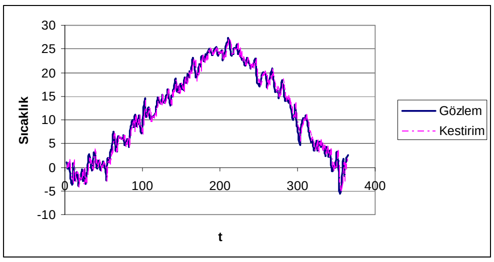
4.1. Lineer Modelin Belirlenmesi

Artık durağan hale getirilen günlük ortalama sıcaklık serisi (Y_t) için zaman serisi analizlerine geçilebilir. İlk önce seri için en uygun lineer zaman serisi modeli araştırılacaktır. Daha önceden de bahsedildiği gibi tahmin edilen modeller arasından en uygun modelin seçiminde AIC kriteri kullanılabilmektedir. Burada en uygun lineer zaman serisi modeli $ARMA(p,q)$. $p = 0,1,\ldots,5$, $q = 0,1,\ldots,5$ modelleri arasından AIC kriterine göre seçilmiş ve en uygun modelin $ARMA(4,3)$ modeli olduğu anlaşılmıştır. Bu modelin parametre tahminine ilişkin ayrıntılı sonuçlar aşağıda Çizelge 1'de verilmiştir.

Çizelge 1. $ARMA(4,3)$ modeline ilişkin tahmin sonuçları

Parametre	Parametre Tahminleri	Standart Hatalar	t-Değeri	p-Değeri
AR(1)	-0.9088	0.0520	-17.49	0.000
AR(2)	0.2035	0.0704	2.89	0.004
AR(3)	0.3073	0.0694	4.43	0.000
AR(4)	-0.1930	0.0524	-3.68	0.000
MA(1)	-1.0746	0.0065	-165.27	0.000
MA(2)	0.2796	0.0090	30.99	0.000
MA(3)	0.7103	0.0005	1569.00	0.000
$\boldsymbol{SSE} = 463.665$. $\boldsymbol{S} = 1.299$				

Çizelge 1'den görüldüğü gibi $ARMA(4,3)$ modelindeki tüm parametreler için elde edilen p-değerleri $\alpha = 0.05$ anlam seviyesi değerinden daha küçüktür. Bu da $ARMA(4,3)$ modelindeki tüm parametrelerin $\alpha = 0.05$ anlam seviyesinde istatistiksel olarak anlamlı olduğunu göstermektedir. Ayrıca $ARMA(4,3)$ modelinin hata kareler toplamı $SSE = 463.665$ ve hata kareler ortalaması $MSE = 1.299$ olarak elde edilmiştir.
Günlük ortalama sıcaklık serisi için uygun bir model olan $ARMA(4,3)$ modeli kullanılarak elde edilen kestirim değerleri ve gözlem değerleri Şekil 3'de verildiği gibidir.



Şekil 3. Gözlemler ve $ARMA(4,3)$ ile elde edilen kestirim değerleri

Sonuç olarak serimiz için $ARMA(4,3)$ modeli, gerek çeşitli lineer modeller arasında en küçük AIC değerine sahip olması. gerek parametrelerinin istatistiksel olarak anlamlı olması gerekse bu modelden elde edilen artık sürecinin beyaz gürültü sürecine sahip olması bakımından uygun ve anlamlı bir modeldir. Ancak bu model sadece lineer modeller arasında uygun görülen bir modeldir. Acaba lineer olmayan zaman serisi modelleri içinde, performansı $ARMA(4,3)$ modeline göre daha iyi olan bir model belirleyebilir miyiz? Uygulamanın bundan sonraki kısmında bu sorunun cevabı bulunmaya çalışılacaktır.

4.2. Lineer Olmayan Modelin Belirlenmesi

Seri için lineer olmayan bir model belirleyip bu modeli tahmin etmeden önce lineer modeller arasından seri için uygun gördüğümüz $ARMA(4,3)$ modeli farklı derecelerden ($P = 1,2,3,4$ ve $Q = 1,2,3,4$) $BL(4,3,P,Q)$ lineer olmayan alternatif modellerine karşı test edilmiş ve elde edilen LM istatistiğinin değerleri ve karar için test istatistiğinin karşılaştırılacağı PQ serbestlik dereceli χ^2 dağılım tablosundan $\alpha = 0.05$ anlam seviyesi için belirlenen tablo değerleri Çizelge 2'de verilmiştir.

Çizelge 2. $ARMA(4,3)$ modelinin $BL(m,k)$ alternatifine karşı LM testi sonuçları

P	Q	**LM**	$\chi^2_{0.95;PQ}$	**Karar**	P	Q	**LM**	$\chi^2_{0.95;PQ}$	**Karar**
1	**1**	0.578	3.841	$ARMA(4,3)$	**3**	**1**	10.049	7.815	$BL(m,k)$
1	**2**	11.850	5.991	$BL(m,k)$	**3**	**2**	16.320	12.592	$BL(m,k)$
1	**3**	17.754	7.815	$BL(m,k)$	**3**	**3**	34.715	16.919	$BL(m,k)$
1	**4**	19.734	9.488	$BL(m,k)$	**3**	**4**	57.467	21.026	$BL(m,k)$
2	**1**	2.985	5.991	$ARMA(4,3)$	**4**	**1**	17.875	9.488	$BL(m,k)$
2	**2**	12.843	9.488	$BL(m,k)$	**4**	**2**	49.701	15.507	$BL(m,k)$
2	**3**	32.414	12.592	$BL(m,k)$	**4**	**3**	57.447	21.026	$BL(m,k)$
2	**4**	35.415	15.507	$BL(m,k)$	**4**	**4**	59.955	26.296	$BL(m,k)$

Çizelge 2'ye göre $P = 1, Q = 1$ ve $P = 2, Q = 1$ dereceleri için LM istatistiğinin aldığı değerler $\chi^2_{0.95;PQ}$ tablo değerinden daha küçük olduğundan bu dereceler için $ARMA(4,3)$ modeli, $BL(4,3,1,1)$ ve $BL(4,3,2,1)$ modellerine karşı kabul edilmektedir. Benzer şekilde $P = 1, Q = 1$ ve $P = 2, Q = 1$ derecelerinin dışındaki tüm model dereceleri için $ARMA(4,3)$ modeli, $BL(4,3,P,Q)$ modellerine karşı reddedilmektedir.

Burada çeşitli derecelerden $BL(p,q,P,Q)$ modelleri tahmin edilmiş ve bu modeller arasından en uygununun, AIC değeri en küçük olan $BL(1,3,2,4)$ modeli olduğu anlaşılmıştır. Bu modelin parametre tahminine ilişkin ayrıntılı sonuçlar aşağıda Çizelge 3'de verilmiştir.

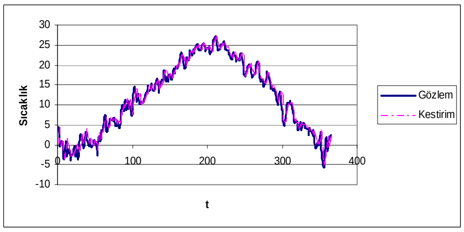
Çizelge 3. $BL(1,3,2,4)$ modeline ilişkin tahmin sonuçları

Parametre	**Tahmin**	**Standart Sapma**	**Güven Aralıkları**	
			Alt sınır	**Üst sınır**
α	0.111	0.027	0.058	0.163
a_1	-0.151	0.068	-0.284	-0.018
b_1	0.317	0.070	0.179	0.455
b_2	-0.168	0.021	-0.209	-0.127
b_3	-0.147	0.023	-0.191	-0.102
c_{11}	-0.021	0.011	-0.041	0.000
c_{12}	0.126	0.034	0.059	0.192
c_{13}	-0.002	0.014	-0.029	0.025
c_{14}	-0.201	0.017	-0.234	-0.167

c_{21}	-0.226	0.033	-0.290	-0.162
c_{22}	-0.092	0.011	-0.114	-0.070
c_{23}	0.005	0.018	-0.030	0.039
c_{24}	0.041	0.014	0.015	0.068
SSE=444.937 , ***S***=1.267				

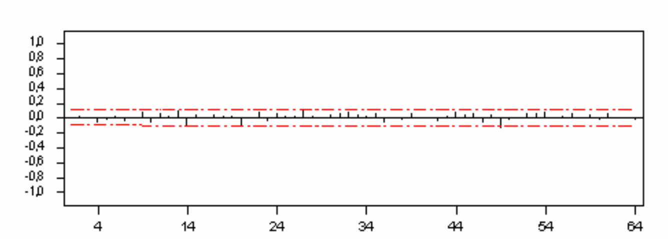
Çizelge 3'e bakıldığında c_{11} ve c_{13} parametreleri için oluşturulan $\alpha = 0.05$ anlam seviyeli güven aralıklarının 0'ı içerdiği görülür. Bu, parametreler için parametrenin sıfıra eşitliği hipotezinin $\alpha = 0.05$ anlam seviyesinde kabul edilmesi anlamına gelmektedir. Yani c_{11} ve c_{13} parametreleri istatistiksel olarak anlamsızdır. Diğer parametreler için ise oluşturulan güven aralıkları 0'ı içermediğinden bu parametreler istatistiksel olarak anlamlıdır.

Günlük ortalama sıcaklık serisi için $BL(1,3,2,4)$ modeli kullanılarak elde edilen kestirim değerleri ve gözlem değerlerinin grafiği Şekil 4'de verildiği gibidir.



Şekil 4. Gözlemler ve $BL(1,3,2,4)$ ile elde edilen kestirim değerleri

$BL(1,3,2,4)$ lineer olmayan zaman serisi modelinden elde edilen artıklara ilişkin otokorelasyon katsayıları Şekil 5'de verildiği gibi elde edilmiştir.



Şekil 5. $BL(1,3,2,4)$ model artıklarının otokorelasyon katsayıları

Şekil 5'den görüldüğü gibi $BL(1,3,2,4)$ modelinden elde edilen artıkların tüm gecikme değerleri için otokorelasyon katsayıları güven bandının içinde kalmaktadır. Yani bu modelden elde edilen artıkların beyaz gürültü sürecine sahip olduğu söylenebilir.

5. SONUÇ

Yapılan çalışmada DMGM'den alınan Konya ili için günlük ortalama sıcaklık verileri kullanılmıştır. Bu veri için, zaman serisi analizlerine başlamadan önce durağanlık sınaması yapılmış ve serinin durağan olmadığı görülmüştür. Seriyi durağan hale getirebilmek amacıyla seriye, bir derece fark operatörü uygulanmış ve görülmüştür ki bir derece farkı alınan seri durağan hale gelmiştir. Uygulamanın ikinci aşamasında durağan hale gelen günlük ortalama sıcaklık serisi için $ARMA(p,q),\ p,q = 0,1,\ldots,5$ lineer zaman serisi modelleri arasından AIC kriterine göre en uygun model olarak $ARMA(4,3)$ modeli belirlenmiş ve Çizelge 1'de bu modele ilişkin ayrıntılı sonuçlar sunulmuştur. Uygulamanın daha sonraki aşamasında, seri için uygun lineer zaman serisi modeli olarak belirlenen $ARMA(4,3)$ modeli çeşitli derecelerden BL alternatif lineer olmayan zaman serisi modeline karşı Lagrange çarpanları (*LM*) testi kullanılarak test edilmiş ve hemen hemen tüm alternatif modeller $ARMA(4,3)$ modeline karşı kabul edilmiştir. Daha sonra çeşitli derecelerden *BL* modelleri tahmin edilmiş ve *AIC* kriterine göre en uygun modelin *BL*(*1,3,2,4*) modeli olduğu görülmüştür. Günlük ortalama sıcaklık serisi için belirlenen bu iki modelin yani, $ARMA(4,3)$ ve *BL*(*1,3,2,4*) modelleri arasında, *BL*(*1,3,2,4*) lineer olmayan zaman serisi modelinin daha küçük *AIC* değerine sahip olduğu görülmüştür. Sonuçta, günlük ortalama sıcaklık serilerinin analizlerinde lineer olmayan zaman serisi modelleri kullanılabilirdir. Belki de, bu çalışmada incelenen çiftdoğrusal modellerin dışında başka bir model daha iyi performans sağlayabilir.

KAYNAKLAR

[1] Box, G.E., Jenkins, G.M., G.C. Reinsel, (1994), *Time Series Analysis: Forecasting and Control*, Holden-Day, San Fransisco, 1994.
[2] Gabr, M. M., ve Subba Rao, T., "The Estimation and Prediction of Subset Bilinear Time Series Models with Applications", Journal of Time Series Analysis, 2: 155-171, 1981.
[3] Granger, C., ve Anderson, A., *Introduction to Bilinear Time Series Models,* Vandenhoeck and Ruprect, Göttingen, 1978.
[4] Saikkonen, P., ve Luukkonen, R., "Lagrange Multiplier Test for Testing Non-linearities in Time Series Models", Scandinavian Journal of Statistics, 15: 55-68, 1988.
[5] Tong, H., *Non-linear Time Series: A Dynamical System Approach*, New York, Oxford University Press, 1990.
[6] Zhang, Y., Testing for nonlinearities in time series with an application to exchange rates, Doktora tezi, Iowa State University, Iowa, 2002.