

TEK DEĞİŞKENLİ ZAMAN SERİLERİNDE MODEL SEÇİM ÖLÇÜTLERİ ÜZERİNE BİR İNCELEME

Hilal GÜNEY*

Reşat KASAP**

ÖZET

Bu çalışmada, Akaike Bilgi (AIC), Son Kestirim Hatası (FPE), Hannan-Quinn Bilgi (HQ), Düzeltilmiş Akaike Bilgi (AICC) ve Schwarz Bilgi (SIC) Ölçütleri dikkate alınarak, tek değişkenli zaman dizileri modellerinde uygun model derecesinin seçilmesi incelenmiştir. Ölçütlerin birbirleri ile karşılaştırılması, Monte Carlo simülasyon yöntemi kullanılarak hangi ölçütün hangi gecikme sayısını kaç kez seçtiği çalışmaları yapılmıştır.

Anahtar Kelimeler: Model seçim ölçütleri, Simülasyon, Zaman serileri.

1. GİRİŞ

Zaman serileri analizinde modelleme sürecinin en önemli aşamalarından biri uygun model seçimidir. Tek değişkenli zaman serilerinde uygun model seçimi uygun model derecesinin belirlenmesiyle aynı anlama gelir. Uygun modeli seçmek için en çok olabilirlik (Maximum likelihood) yöntemi gibi varsayımları ve ulaştıkları noktalar birbirinden farklı olan birçok yöntem kullanılabilir. En çok olabilirlik yöntemi, daima en yüksek olasılıklı seçimi verir. Bununla birlikte en çok olabilirlik yöntemi, doğru boyutu seçmek için uygun bir biçim değildir (Schwarz, 1978). Otoregresif (AR) modeller ile Hareketli Ortalamalar (MA) modellerinde uygun model derecesinin belirlenmesi için kullanılan bazı ölçütler vardır. Farklı parametre sayısına sahip farklı modeller arasından en uygun dereceli modeli seçmek için Akaike tarafından en çok olabilirlik prensibinin genişletilmiş biçimi 1974'te geliştirilmiştir. Akaike'nin önerisi, her bir j modeli için ayrı ayrı olabilirlik fonksiyonunu maksimize etmek olarak ifade edilebilir (Akaike, 1974; Schwarz, 1978).

Zaman serileri analizinde modelleme süreci için Box-Jenkins yaklaşımı çok sık kullanılır. Bu metoda göre model tespiti yapılırken öncelikle serilerin durağan olup olmadığına bakılır. Bunun için Otokorelasyon (ACF) ve Kısmi Otokorelasyon (PACF) katsayıları incelenir. Bu katsayılar istatistiksel olarak anlamsız ise seriyi durağan hale getirmek için uygun fark alma işlemleri uygulanır. İkinci olarak ACF ve PACF'nin grafiksel davranışları incelenir. ACF, otoregresif modeller için üstel veya sinüzoidal olarak azalırken, hareketli ortalamalar modelleri için q gecikmeden sonra aniden kesilir. PACF ise otoregresif modeller için p gecikmeden sonra aniden kesilirken, hareketli ortalamalar modelleri için üstel veya sinüzoidal olarak azalır. ACF ve PACF'nin ikisinde de salınımlı bir düşüş varsa bu durum Otoregresif Hareketli Ortalamalar (ARMA) sürecini belirtir. Son olarak sıfırdan farklı ACF ve PACF yapısıyla p ve q model dereceleri belirlenir. Tek değişkenli zaman serilerinde modelin derecesi kabaca

* Gazi Üniversitesi, Fen Edebiyat Fakültesi, İstatistik Bölümü, Ankara, e-posta: hguncy@gazi.edu.tr

** Gazi Üniversitesi, Fen Edebiyat Fakültesi, İstatistik Bölümü, Ankara, e-posta: rkasap@gazi.edu.tr

belirlendikten sonra parametreler tahmin edilir. Parametre tahmini genel olarak momentler yöntemi, en çok olabilirlik yöntemi ve bilinen küçük kareler yöntemiyle yapılır. Geçici modelin parametre tahminleri yapıldıktan sonra modelin uygunluğu incelenir. Modelin uygunluğunun incelenmesi, artıklara dayalı bir incelemedir. A_t sürecinin beyaz gürültü olma varsayımı, normal dağılıma sahip olması ve artıklara ilişkin otokorelasyonların sıfır olması durumları araştırılır. Uygun modelin *ACF* ve *PACF* grafiklerinin beyaz gürültü sürecinin *ACF* ve *PACF*'lerin yapısına benzemesi istenir.

Zaman serileri analizinde, bir seri için birden fazla uygun model elde edilebilir. Bunlardan en iyi modeli seçmek için ise model seçim ölçütleri kullanılır (Wei, 1990; Box ve Jenkins, 1994).

Model seçim ölçütleriyle ilgili literatürde birçok çalışma vardır. Bunlardan ilki olan Akaike, Son Kestirim Hatası (Final Prediction Error) *FPE* ölçütünü otoregresif modeller için denemiştir. Kestirim hata karesi ortalaması olarak tanımlanan *FPE* ölçütünün, tahminin yanlılığı ve varyansı arasında karşılıklı iyi bir denge oluşturduğunu göstererek bu yöntemin iyi sonuçlar verdiğini belirtmiştir (Akaike, 1969). Akaike, bu ölçütün durağan ve bağımsız beyaz gürültü sürecinden üretilen otoregresif modeller için tanımlamış ve en küçük *FPE* değerine sahip modeli en iyi model olarak önermiştir. Bu çalışmada Akaike, *FPE* ölçütünün tutarlı olduğunu da göstermiştir (Akaike, 1970). Akaike, *ARMA* modeller için en çok olabilirlik yöntemini denemiş ve bu yöntemi geliştirerek uygunluğunu göstermiştir (Akaike, 1973). Ayrıca klasik en çok olabilirlik yöntemini yeniden gözden geçirmiş ve Akaike Bilgi Ölçütü (Akaike Information Criterion) *AIC*'yi tanımlamıştır (Akaike, 1974). Akaike bundan başka, *AIC*, *FPE* ve Bayes yöntemini otoregresif modellerde deneyerek, *AIC* model seçim ölçütünün kullanılabilirliğini örneklerle göstermiştir (Akaike, 1979). Hannan ve Quinn, güçlü tutarlı bir tahmin etme yönteminin tekrarlı logaritma yöntemi kullanılarak geliştirilebileceğini göstermiştir. Ayrıca bu çalışmada önerilen Hannan-Quinn (*HQ*) Ölçütünün, diğer ölçütlere oranla daha az eksik tahmin ettiği gösterilmiştir. Sonuç olarak kesin bir yargıya ulaşamamış olsa da bu ölçütün, örnek hacminin büyük olduğu otoregresyon modellerinde daha iyi bir yaklaşımda bulunduğunu ortaya koymuştur (Hannan ve Quinn, 1979). Quinn, Hannan ve Quinn'in 1979'da yaptıkları çalışmada geliştirdikleri ölçütü, çok değişkenli *AR* süreçlerinde denemiştir. Tekrarlı logaritma kullanılan bu ölçüt, çok değişkenli *AR* süreçleri için tavsiye edilmiştir (Quinn, 1980). Hurvich ve Tsai, *AIC*'ye yanlılık için düzeltme uygulanarak oluşturulan düzeltilmiş *AIC*'yi regresyon modelleri ve otoregresif zaman serileri için denemişlerdir. Bu düzeltmenin yalnızca küçük örnek hacmine sahip model için geçerli olduğu belirtilmiştir. Ayrıca gerçek model sınırlı boyutlu olduğunda, düzeltilmiş *AIC*'nin diğer ölçütlerden daha iyi tahmin sonuçları verdiği gösterilmiştir (Hurvich ve Tsai, 1989). Bedrick ve Tsai, küçük örnek hacimli, çok değişkenli regresyon modelleri için düzeltilmiş *AIC* ölçütünü geliştirmişlerdir. Ayrıca bu çalışma, düzeltilmiş *AIC*'nin Kullback-Leibler bilgisinin yansız tahmin edicisi olduğunu açıklamaktadır. Küçük örnek hacimli tarımsal gözlemler üzerinde yapılan karşılaştırmada, düzeltilmiş *AIC* diğer ölçütlerden daha iyi sonuç vermiştir (Bedrick ve Tsai, 1994). Schwarz, farklı boyutlu modellerden birini seçme problemi için Bayes çözümünü önermiştir (Schwarz, 1978). Stone, çalışmasında *AIC* ve *SIC* ölçütünü karşılaştırmıştır. Bu çalışmada parametre sayısı sonsuz kabul edilmiştir. Gerçek modellerde parametre sayısı sonsuz olamayacağından bu karşılaştırmanın bir gerçekliği bulunmadığını belirtmiştir (Stone, 1979).

Bu çalışmanın konusu, ele alınan beş farklı model seçim ölçütünün tek değişkenli durağan otoregresif ve hareketli ortalamalar serileri açısından karşılaştırılmasıdır. Bu karşılaştırma simülasyon programları vasıtası ile yapılmıştır.

2.YÖNTEM

2.1 Otoregresif (AR) Modeller ve Hareketli Ortalamalar (MA) Modelleri

Gecikmeleri cinsinden ifade edilebilen zaman dizisi, *AR* model olarak adlandırılır (Box ve Jenkins, 1994). Eğer geçmiş dönemden sadece bir tane gözleme bağlı bir *AR* modeli söz konusu ise “birinci dereceden *AR* modeli”, *p* tane geçmiş gözlem değerine bağlı *AR* modeli söz konusu ise “*p*. dereceden *AR* modeli” olarak adlandırılır ve *AR(p)* şeklinde gösterilir. *p*. dereceden otoregresif süreç genel olarak aşağıdaki gibi ifade edilir.

$$Z_t = \phi_1 Z_{t-1} + \dots + \phi_p Z_{t-p} + A_t \quad (1)$$

Burada $\phi_1, \phi_2, \dots, \phi_p$ modelin bilinmeyen fakat tahmin edilecek parametreleridir. *p*, modelin derecesi ve A_t ise sıfır ortalamalı, σ^2 sabit varyanslı beyaz gürültü sürecidir (Enders, 2004).

Aynı dönemin hata terimi ile belirli sayıda geçmiş dönemin hata terimlerinin doğrusal bir bileşimi olarak ifade edilebilen bir zaman dizisi söz konusu ise bu modeller, *MA* modelleri olarak adlandırılır. *MA* modelleri de *AR* modellerine benzer şekilde içerdikleri geçmiş hata terimi sayısına göre belirlenir. Eğer geçmiş dönemden sadece bir tane hata terimine bağlı bir *MA* modeli söz konusu ise “birinci dereceden *MA* modeli”, *q* tane geçmiş hata terimine bağlı bir *MA* modeli söz konusu ise “*q*. dereceden *MA* modeli” olarak adlandırılır.

MA(q) modeli genel olarak aşağıdaki gibi ifade edilir.

$$Z_t = A_t - \theta_1 A_{t-1} - \theta_2 A_{t-2} - \dots - \theta_q A_{t-q} \quad (2)$$

Burada, $\theta_1, \theta_2, \dots, \theta_q$ modelin tahmin edilecek parametrelerini, *q* ise *MA* modelinin derecesini göstermektedir (Enders, 2004).

2.2 Model Seçim Ölçütleri

Zaman serilerinde model seçim ölçütleri üzerine yapılan pek çok çalışmada genellikle belli başlı seçim ölçütleri dikkate alınmıştır. Bu seçim ölçütlerinin çok sayıda türevi üzerinde yeni bakış açıları verilmiştir. Bu makalede bilinen beş model seçim ölçütü üzerinde durulacağından, bunlara ilişkin kuramsal yapı aşağıda özetlenmiştir.

2.2.1 Akaike Bilgi Ölçütü

Akaike, Kullback-Leibler bilgisini log-olabilirlik fonksiyonunun maksimum noktasına dayandıran bir çözüm yolu bulmuştur (Burnham ve Anderson, 2004). Çünkü en çok olabilirlik tahminleri, uygun koşullar altında asimptotik olarak etkindir ve gerçek değer çevresinde parametrelerin küçük değişimlerine karşı duyarlıdır. Ayrıca en çok olabilirlik fonksiyonu nicelik olarak da ifade edilmeye uygundur (Akaike, 1974).

AIC 'nin uygulamadaki yaygın kullanım şekli,

$$AIC = n \ln \hat{\sigma}^2 + 2k \quad (3)$$

biçimindedir. Bu eşitlikte, modeldeki toplam parametre sayısı k ile örnek hacmi ise n ile gösterilmiştir. $\hat{\sigma}^2$, beyaz gürültü teriminin varyansı olan σ^2 'nin en çok olabilirlik tahmin edicisidir. Bu varyansın tahmini,

$$\hat{\sigma}^2 = \frac{RSS}{n} = \frac{\sum (\hat{e}_i^2)}{n} \quad (4)$$

eşitliği ile elde edilir. Burada RSS , modelden elde edilen artıkların kareleri toplamını, n , örnek hacmini ifade etmektedir.

Farklı modeller için hesaplanan AIC değerlerinden en küçük olanına sahip model en iyi model olarak seçilir (Box ve Jenkins, 1994).

2.2.2 Düzeltilmiş Akaike Bilgi Ölçütü

Akaike'nin yaklaşımı, temel bir teoriye dayanan model seçimine ve başka teorik çalışmalara da izin vermiştir. AIC ölçütü, parametre sayısı örnek hacmine göre fazla olduğunda, başka bir ifadeyle küçük örnek hacmiyle çalışıldığında iyi sonuçlar vermemektedir. Genelde yaklaşık olarak

$$\frac{n}{k} < 40 \quad (5)$$

olduğunda AIC ölçütü sapma için ek bir düzeltmeye ihtiyaç duyar (Burnham ve Anderson, 2004).

$$AIC_c(k) = \log \hat{\sigma}_k^2 + \frac{n+k}{n-k-2} \quad (6)$$

n örnek hacmi ve p uydurulan modelin gecikme sayısıdır. En küçük AIC_c değerine sahip model en iyi model olarak alınır (Luna, 1995).

n , k 'ya göre büyüdüğünde diğer bir ifadeyle $\frac{n}{k}$ oranı yaklaşık olarak 40'tan büyük olduğunda, AIC_c , AIC 'ye yaklaşır ve AIC iyi sonuçlar vermeye başlar (Burnham ve Anderson, 2004). AIC_c kullanılması gereken durumlarda AIC kullanmak literatürde çok yapılan bir hatadır (Burnham ve Anderson, 2004).

2.2.3 Son Kestirim Hatası

Modelin optimum gecikme sayısı p 'yi tahmin etmenin bir yolu da p 'nin değerini seçmektir. Bu seçim, tahmin edilmiş ortalama hata karelerini minimize eden bazı üst sınırların altındaki değerlerden birini seçerek yapılır. Akaike bu sonucu kullanarak, FPE ölçütünü önermiştir (Luna, 1995). FPE ölçütü, AR modellerin derecesinin sınanması amacıyla önerilmiştir (Akaike, 1969; Baran ve Bacanlı, 2006).

AR modellerinin gecikme sayısını belirlemek için ilk tanımlama Akaike tarafından 1969'da yapılmıştır. Bu yöntem, gelecek gözlemlerin kestirimini yapmak için geçmiş gözlemlerin doğrusal kombinasyonunu kullanır. t zamanı için $Z_t, Z_{t-1}, Z_{t-2}, \dots$ gözlemleri verilmiş ve gelecek Z_{t+1} gözlemi kestirilmek isteniyorsa, en iyi bir adım öte kestirim ya da son kestirim $\hat{Z}_t(1)$, geçmiş gözlemlerin lineer kombinasyonları olarak

$$Z_t(1) = -\phi_1 Z_t - \dots - \phi_p Z_{t-p} \quad (7)$$

şeklinde verilir. Bu eşitlik son kestirim hatası $\varepsilon_t(1)$ 'in kare ortalamasını minimize ettiği için en iyi lineer kombinasyondur. Ortalama hata karesi ise aşağıdaki gibi verilebilir,

$$E[\varepsilon_t(1)]^2 = E[Z_{t+1} - \hat{Z}_t(1)]^2 = \sigma_\varepsilon^2 \quad (8)$$

Bu yüzden son kestirim ortalama hata karesi, artık varyansı olarak da değerlendirilir. Bir adım ileri kestirim hatası sadece katsayılar kesin olarak bilindiğinde kullanılabilir. Katsayıların yerine en küçük kareler tahminleri koyulduğunda, tahmin edilmiş tek adım kestirimin ortalama hata karesi σ_ε^2 olarak ifade edilir (Parkhurst, 1992). Akaike, bu hata karesi kestirimini, artıkların en çok olabilirlik tahminini kullanarak aşağıdaki gibi göstermiştir,

$$FPE(j) = \frac{n+k}{n-k} \hat{\sigma}_\varepsilon^2 \quad (9)$$

(Akaike, 1969). Burada n , uydurulan model için gözlem sayısıdır. Artıkların varyansının en çok olabilirlik tahmin edicisi

$$\hat{\sigma}_\varepsilon^2 = \hat{\rho}_0 + \hat{\phi}_1 \hat{\rho}_1 + \dots + \hat{\phi}_p \hat{\rho}_p \quad (10)$$

şeklinde. Eğer $FPE(p) - p$ grafiği çizilirse, grafiğin minimum noktasındaki p değeri, AR modeli için uygun dereceyi gösterir (Parkhurst, 1992). Diğer bir deyişle en küçük FPE değerine sahip model en uygun model olarak seçilir.

2.2.4 Schwarz Bilgi Ölçütü

Schwarz, değişik boyutlara sahip modellerden bir tanesini seçme problemini, Bayes çözümüyle denemiştir. Model derecesini belirlemek için Bayes çözümü altında bir seçim ölçütü sunmuştur. AIC gibi Schwarz Bilgi Ölçütü (SIC) de cimrilik prensibinin (principle of parsimony) nitel olarak matematiksel formülasyonunu verir.

Schwarz bilgi ölçütü aşağıdaki gibi verilir,

$$SIC = \ln(\text{ençokolabilirlik}) - \frac{1}{2} \ln(n)k \quad (11)$$

Burada, k tahmin edilen parametre sayısını, n ise gözlem sayısını ifade eder (Schwarz, 1978).

SIC , Bayes bilgi ölçütüne asimptotik bir yaklaşımdır. Bu yaklaşım, örnek hacmi arttıkça en yüksek sonsal (posterior) olasılığın etkilerini gösterir. İki varsayım göz önüne alınır. Bunlardan ilki, aynı gecikme sayısına sahip fakat farklı parametrelili modellerin oluşturduğu sınıfın, sıfır olmayan önsel (priori) olasılıklı olması, ikincisi ise ceza terimi sabit tutularak uygun olmayan modelin saptanmasıdır. Schwarz bu varsayımlara dayanarak SIC ölçütünü önermiştir (Neftçi, 1982).

Diğer model seçim ölçütlerinde olduğu gibi, SIC her model için hesaplanır ve en küçük SIC değerine sahip model en iyi model olarak seçilir (Koehler ve Murphree, 1988; Burnham ve Anderson, 2004).

AIC ve SIC modelin uygunluğu için iki objektif ölçüttür. Fakat bu iki ölçüt model derecesini belirlemek için eklenen ceza (penalty) terimi açısından farklıdır. Bu ceza terimi SIC ölçütünde $\frac{1}{2} \log(n)$ 'dir ve AIC 'deki $2k$ ceza teriminden farklı olarak daha düşük boyutlu modeller için daha uygun bir belirleme ölçütüdür. Bu düşük boyutlu model, en az 8 ya da daha fazla gözlem içermelidir (Schwarz, 1978).

2.2.5 Hannan-Quinn Bilgi Ölçütü

Hannan ve Quinn tek değişkenli otoregresif modeller için kullanılacak tutarlı tahmin edicinin tekrarlı logaritma yöntemine dayandırıldığını göstermişlerdir. Diğer tutarlı model seçim yöntemleriyle kıyaslayarak Hannan-Quinn (HQ) Bilgi Ölçütünün derece seçiminde eksik tahmine daha az meyilli olduğunu açıklamışlardır (Hannan ve Quinn, 1979). Quinn bu yöntemi çok değişkenli otoregresif modeller için tavsiye etmiştir (Quinn, 1980). Hannan $ARMA$ modelleri için bu ölçütün güçlü tutarlılığını göstermiştir (Hannan, 1980).

HQ model seçim ölçütü aşağıda verilmiştir,

$$HQ = n \log \hat{\sigma}_e^2 + 2k \log \log(n) \quad (12)$$

Burada, n örnek hacmini k ise modeldeki parametre sayısını göstermektedir (Hannan ve Quinn, 1979).

3. BULGULAR

Bu bölümde, AR ve MA modellerinde uygun gecikme sayısının belirlenebilmesi amacıyla kullanılan bazı bilgi ölçütlerinin karşılaştırması için Monte Carlo yöntemiyle bir simülasyon programına yer verilmiştir. Bu simülasyon çalışmasında AIC , FPE , HQ , $AICC$, SIC ölçütleri ele alınmıştır.

Çalışmanın her üç durumunda da dört ayrı örnek hacmi ($n=60$, $n=120$, $n=240$ ve $n=600$) ele alınmıştır. Bu örnek hacimleri, gerçek bir uygulamaya benzetebilmek açısından 12'nin katları olarak tercih edilmiştir. Bu 4 ayrı örnek hacmi ve yukarıda ayrıntılı olarak verilen modeller için 1000 tekrarlı simülasyon 12 gecikmeye kadar gerçekleştirilmiş ve her model için ele alınan model seçim ölçütlerinin hangi gecikme sayısını kaç kez seçtiği hesaplanmıştır. Bu deneysel çalışmada Matlab 2006 bilgisayar programı kullanılmıştır.

Simülasyon çalışması aşağıda verilen modeller üzerinde yapılmıştır. Önce ele alınan AR modelinden $n=60$ hacminde seri elde edilmiş ve bu seri için incelenen tüm model seçim ölçütlerinin hangi gecikme sayısını seçtiği belirlenmiştir. Bu işlem 1000 kez tekrarlanarak hangi ölçütün hangi model derecesini $n=60$ örnek hacmi için kaç kez seçtiği hesaplanmıştır. Ardından aynı döngü $n=120$, $n=240$ ve $n=600$ hacmindeki örnekler için tekrarlanmıştır. Her bir AR modeli için yapılan bu simülasyon MA modeller için de uygulanmıştır. AR ve MA modeller için ortalama 0, varyans ise 1 olarak alınmıştır. Hata terimleri ise 0 ortalamalı ve σ^2 sabit varyanslı normal dağılımdan üretilmiştir. Simülasyon sonuçları Tablo 1, Tablo 2 ve Tablo 3'te verilmiştir.

$$AR(3) : Z_t = 0.3Z_{t-1} - 0.7Z_{t-2} + 0.2Z_{t-3} + A_t \quad (13)$$

$$AR(4) : Z_t = 0.2Z_{t-1} + 0.2Z_{t-2} - 0.1Z_{t-3} - 0.1Z_{t-4} + A_t \quad (14)$$

$$AR(5) : Z_t = 0.2Z_{t-1} + 0.1Z_{t-2} - 0.1Z_{t-3} - 0.3Z_{t-4} + 0.6Z_{t-5} + A_t \quad (15)$$

$$MA(3) : Z_t = A_t - 0.3A_{t-1} + 0.7A_{t-2} - 0.2A_{t-3} \quad (16)$$

$$MA(4) : Z_t = A_t - 0.2A_{t-1} - 0.2A_{t-2} - 0.1A_{t-3} + 0.1A_{t-4} \quad (17)$$

$$MA(5) : Z_t = A_t - 0.2A_{t-1} - 0.1A_{t-2} + 0.1A_{t-3} + 0.3A_{t-4} - 0.6A_{t-5} \quad (18)$$

Tablo 1. Seçilmiş model dereceleri

AR(3) modeli							AR(4) modeli						
n	p	AIC	FPE	HQ	AICC	SIC	n	p	AIC	FPE	HQ	AICC	SIC
60	1	1	1	3	2	1	60	1	483	450	639	544	446
	2	459	418	593	518	417		2	59	59	42	53	59
	3	361	370	304	349	361		3	40	39	39	42	38
	4	91	87	60	79	85		4	281	284	213	272	276
	5	38	48	19	27	48		5	53	65	30	44	62
	6	15	23	11	13	22		6	32	36	15	21	37
	7	10	14	5	7	15		7	17	18	8	9	16
	8	8	12	3	3	13		8	9	14	6	6	18
	9	5	6	1	1	9		9	12	15	5	6	17
	10	5	8	1	1	10		10	5	4	1	1	5
	11	5	7	0	0	10		11	4	7	1	1	7
	12	2	6	0	0	9		12	5	9	1	1	19
120	1	0	0	0	0	0	120	1	131	121	265	148	215
	2	228	219	362	244	321		2	27	25	29	30	30
	3	572	564	541	576	546		3	26	23	28	26	22
	4	95	94	60	93	73		4	650	642	604	648	616
	5	48	52	21	45	35		5	86	87	53	83	68
	6	23	26	10	19	13		6	42	44	18	41	32
	7	12	13	3	11	5		7	15	21	2	11	9
	8	4	5	1	1	1		8	8	14	1	5	3
	9	6	6	1	3	2		9	9	11	0	5	3
	10	5	11	1	4	1		10	2	3	0	0	0
	11	5	7	0	3	2		11	3	3	0	3	2
	12	2	3	0	1	1		12	1	6	0	0	0
240	1	0	0	0	0	0	240	1	2	2	13	2	15
	2	60	58	118	64	119		2	1	1	5	1	5
	3	720	715	792	730	791		3	1	1	5	1	5
	4	114	114	68	111	68		4	780	768	894	801	892
	5	44	47	15	43	15		5	101	106	49	93	49
	6	27	27	3	24	3		6	49	49	25	49	25
	7	14	15	2	13	2		7	31	32	7	26	7
	8	12	12	1	9	1		8	17	18	2	15	2
	9	2	3	0	0	0		9	6	7	0	3	0
	10	4	5	1	3	1		10	2	3	0	2	0
	11	0	1	0	0	0		11	3	5	0	2	0
	12	3	3	0	3	0		12	7	8	0	5	0
600	1	0	0	0	0	0	600	1	0	0	0	0	0
	2	0	0	2	0	4		2	0	0	0	0	0
	3	773	770	908	782	936		3	0	0	0	0	0
	4	115	116	61	113	44		4	802	797	936	806	953
	5	58	58	20	54	14		5	108	110	48	108	39
	6	26	26	5	25	1		6	39	40	11	36	7
	7	12	13	3	11	1		7	21	21	3	21	1
	8	3	3	0	3	0		8	14	14	1	13	0
	9	6	6	1	5	0		9	5	6	1	5	0
	10	3	4	0	3	0		10	6	6	0	6	0
	11	3	3	0	3	0		11	2	2	0	2	0
	12	1	1	0	1	0		12	3	4	0	3	0

Tablo 2. Seçilmiş model dereceleri

AR(5) modeli

MA(3) modeli

AR(5) modeli							MA(3) modeli						
n	p	AIC	FPE	HQ	AICC	SIC	n	q	AIC	FPE	HQ	AICC	SIC
60	1	62	53	118	86	51	60	1	1	0	2	1	0
	2	20	15	26	25	15		2	399	350	526	469	324
	3	1	0	2	1	0		3	252	230	270	273	214
	4	4	4	6	4	4		4	87	90	67	88	83
	5	718	686	725	745	659		5	56	57	46	57	54
	6	83	85	61	73	83		6	46	45	28	38	37
	7	41	45	28	32	48		7	35	42	20	26	39
	8	28	35	16	16	39		8	31	39	15	19	45
	9	14	18	7	7	25		9	28	40	10	13	45
	10	15	26	6	6	30		10	30	41	5	5	51
	11	8	17	3	4	22		11	18	34	9	9	54
	12	6	16	2	1	24		12	17	32	2	2	54
120	1	0	0	4	2	3	120	1	0	0	0	0	0
	2	0	0	0	0	0		2	293	282	423	314	370
	3	0	0	0	0	0		3	473	459	478	476	480
	4	0	0	0	0	0		4	70	73	46	66	50
	5	783	754	873	810	835		5	49	50	20	46	30
	6	117	117	86	112	98		6	29	30	14	29	22
	7	51	63	28	46	40		7	16	18	8	14	11
	8	21	25	3	12	10		8	19	20	3	13	7
	9	14	17	4	11	7		9	11	14	2	8	5
	10	7	13	2	3	3		10	13	16	4	9	8
	11	3	4	0	3	3		11	12	17	1	13	6
	12	4	7	0	1	1		12	15	21	1	12	11
240	1	0	0	0	0	0	240	1	0	0	0	0	0
	2	0	0	0	0	0		2	105	104	183	111	184
	3	0	0	0	0	0		3	700	695	738	712	737
	4	0	0	0	0	0		4	96	97	57	94	57
	5	804	788	919	819	919		5	38	38	15	36	15
	6	105	109	61	103	61		6	23	25	3	18	3
	7	38	42	12	31	12		7	15	16	2	13	2
	8	17	19	5	14	5		8	9	9	1	7	1
	9	15	16	2	16	2		9	2	4	1	2	1
	10	15	16	0	12	0		10	3	3	0	2	0
	11	2	5	0	2	0		11	7	7	0	5	0
	12	4	5	1	3	1		12	2	2	0	0	0
600	1	0	0	0	0	0	600	1	0	0	0	0	0
	2	0	0	0	0	0		2	1	1	5	1	6
	3	0	0	0	0	0		3	805	801	921	812	941
	4	0	0	0	0	0		4	104	105	59	103	44
	5	788	784	926	795	954		5	45	45	9	43	5
	6	121	123	61	120	39		6	25	27	4	23	3
	7	33	33	7	31	4		7	7	7	1	6	1
	8	25	25	4	24	3		8	5	5	0	5	0
	9	15	14	1	16	0		9	4	4	0	4	0
	10	9	9	0	8	0		10	2	3	1	2	0
	11	7	8	0	4	0		11	1	1	0	1	0
	12	2	4	1	2	0		12	1	1	0	0	0

Tablo 3. Seçilmiş model dereceleri

MA(4) modeli							MA(5) modeli						
n	q	AIC	FPE	HQ	AICC	SIC	n	q	AIC	FPE	HQ	AICC	SIC
60	1	286	242	434	356	229	60	1	357	300	531	424	282
	2	97	86	110	112	76		2	47	39	46	55	38
	3	57	53	55	61	50		3	46	40	51	52	34
	4	216	187	209	235	172		4	206	196	189	223	183
	5	88	83	68	81	71		5	80	75	55	76	69
	6	60	64	39	51	59		6	51	57	31	50	57
	7	48	51	30	39	46		7	61	62	35	47	60
	8	47	62	22	29	62		8	44	50	20	30	48
	9	29	46	15	17	48		9	34	47	13	14	53
	10	30	42	8	10	54		10	30	47	15	16	60
	11	26	44	8	7	60		11	23	46	8	8	53
	12	16	40	2	2	73		12	21	41	6	5	63
120	1	112	111	223	127	175	120	1	165	151	343	182	266
	2	54	47	91	60	78		2	18	17	23	18	21
	3	42	38	46	45	43		3	32	31	35	31	30
	4	558	535	527	571	535		4	548	540	502	560	519
	5	94	96	55	83	67		5	83	86	44	84	67
	6	30	35	17	26	22		6	39	39	23	37	29
	7	29	33	15	25	23		7	27	28	12	24	19
	8	13	18	9	16	15		8	21	24	8	16	12
	9	17	19	7	14	12		9	18	21	2	13	9
	10	13	14	2	8	7		10	17	21	6	15	11
	11	19	25	5	15	14		11	19	25	1	10	7
	12	19	29	3	10	9		12	13	17	1	10	10
240	1	3	3	10	3	10	240	1	10	10	51	11	52
	2	3	3	9	4	10		2	1	1	2	1	2
	3	7	7	8	7	8		3	1	1	4	2	4
	4	785	770	883	804	882		4	786	769	853	796	851
	5	103	104	61	96	61		5	105	111	66	102	67
	6	47	50	21	43	21		6	41	46	15	38	15
	7	20	22	3	14	3		7	21	19	3	20	3
	8	10	13	1	10	1		8	19	21	6	17	6
	9	9	8	2	8	2		9	4	5	0	3	0
	10	5	9	1	5	1		10	6	9	0	6	0
	11	5	5	0	3	0		11	3	5	0	2	0
	12	3	6	1	3	1		12	3	3	0	2	0
600	1	0	0	0	0	0	600	1	0	0	0	0	0
	2	0	0	0	0	0		2	0	0	0	0	0
	3	0	0	0	0	0		3	0	0	0	0	0
	4	815	812	939	818	955		4	794	789	927	802	954
	5	98	99	42	98	35		5	113	113	54	112	37
	6	42	41	13	41	7		6	45	46	13	44	8
	7	26	28	4	26	2		7	18	21	3	16	0
	8	3	3	1	3	1		8	12	12	1	11	1
	9	7	8	1	7	0		9	8	9	1	5	0
	10	6	6	0	5	0		10	6	6	1	6	0
	11	2	2	0	1	0		11	3	3	0	3	0
	12	1	1	0	1	0		12	1	1	0	1	0

4. TARTIŞMA VE SONUÇ

Bu çalışmada, tek değişkenli zaman otoregresif ve hareketli ortalamalar modellerinde uygun model derecesinin seçilmesinde kullanılan bazı model seçim ölçütleri incelenmiştir. Bu amaçla; *AIC*, *FPE*, *HQ*, *AICC* ve *SIC* ölçütleri dikkate alınmıştır. Seçilen farklı model yapıları ve örnek hacimleri kullanılarak ölçütlerin doğru gecikme sayısını seçme başarılarının karşılaştırılması hedeflenmiş ve bu karşılaştırma sümülasyon programları vasıtasıyla yapılmıştır.

Tüm sonuçlara genel olarak bakıldığında örnek hacmi n arttıkça, model seçim ölçütlerinin çok az değişkenlik gösterdiği ve doğru gecikme sayısını seçme oranının arttığı görülmektedir. *SIC*'de değişkenlik çok daha azdır. Bu durum bu ölçütün sınırlı otoregresyon için tutarlı bir model seçim ölçütü olmasının direk bir sonucudur (Luna, 1995). $n/k < 40$ olan durumlar için *AICC* en iyi sonuçları verir ve en az değişkenlik gösterir. Bu sonuç, Hurvich ve Tsai (1989)'in sonuçları ile uyusmaktadır. Bununla birlikte örnek hacmi arttıkça *AICC*, *AIC*'ye yaklaşmaktadır ve bu iki ölçüt hemen hemen aynı sonuçları vermektedir. *FPE*'nin aşırı tanımlama eğilimi dikkat çekmektedir ve bu durum Hurvich ve Tsai (1989) ile uyusmaktadır. Yeterince büyük örnek hacmi için ise *HQ* diğer ölçütlere nazaran daha iyi sonuçlar vermektedir ve bu sonuç da Hannan ve Quinn (1979) ile uyusmaktadır. *HQ* yine *MA* modelleri için tüm örnek hacimlerinde diğer ölçütlere göre daha iyi sonuç vermektedir. Tüm bunlara ilave olarak örnek hacmi arttıkça Akaike tarafından önerilen *FPE*, *AIC* ve *AICC* ölçütlerinin sayısal değerlerinin birbirine yaklaştığı ve aşırı tanımlama eğilimlerinin fazla olduğu gözle çarpmaktadır.

5. KAYNAKLAR

Akaike, H., 1969. Fitting autoregressive model for prediction. *Annals of the Institute of Statistical Mathematics*, 21: 243-247.

Akaike, H., 1974. A new look at the statistical model identification. *I.E.E.E Transactions on Automatic Control*, AC-19: 716-722.

Baran, T., Bacanlı, Ü. G., 2006. Uygun stokastik model seçim ölçütlerinin değerlendirilmesi. *IMO Teknik Dergi*, 264: 3987-4002.

Bedrick, E. J. and Tsai, C. L. , 1994. Model selection for multivariate regression in small samples", *Biometrics*, 50: 226-231.

Burnham, K. P., Anderson, D. R., 2004. Multimodel inference understanding AIC and BIC in model selection. *Sociological Methods & Research*, 33: 261-304.

Box, G. E. P., Jenkins, G. M., Reinsel, G. C., 1994. *Time series analysis: Forecasting and a control*, Prentice Hall, New Jersey, 10-100, 200-202 .

Enders, W., 2004. *Applied econometric time series*. Wiley, United States of America, 96-97.

- Hannan E. J., Quinn B. G. , 1979. The determination of the order of an autoregression. *Journal of the Royal Statistical Society*, 41: 190-195.
- Hannan E. J., 1980. The estimation of the order of an ARMA process. *The Annals of Statistics*, 8: 1071-1081.
- Hurvich, C. M., Tsai C. T., 1989. Regression and time series model selection in small samples. *Biometrika*, 76: 297-307.
- Luna, X., 1995. An improvement of Akaike's FPE criterion to produce its variability. *Journal of Time Series Analysis*, 19: 457-471.
- Koehler, A. B., Murphree, E. S., 1998. A comparison of the Akaike and Schwarz criteria for selecting model order. *Applied Statistics*, 37: 187-195.
- Neftçi, S. N., 1982. Specification of economic time series models using Akaike's criterion. *Journal of the American Statistical Association*, 77: 537-540.
- Parkhurst, A. M., 1992. Evaluation of order determination procedures in ARMA models. Doktora Tezi, Presented to the Faculty of The Graduate Collage at The University of Nebraska, 91-93(İngilizce).
- Quinn, B. G., 1980. Order determination for a multivariate autoregression. *Journal of Time Series Analysis*, 42, 182-185.
- Schwarz, G. D., 1978. Estimating the dimension of a model. *The Annals of Statistics*, 6: 461-464.
- Stone, M., 1979. Model selection criteria of Akaike and Schwarz. *Journal of the Royal Statistical Society. Series B (Methodological)*, 41: 276-278.
- Wei, W. W. S., 1990. *Time series analysis: Univariate and multivariate analysis*. Addison-Wesley Publishing company, Inc. , England, 32-57, 67-84.

EXAMINING MODEL SELECTION CRITERIA FOR SINGLE VARIABLE TIME SERIES

ABSTRACT

In this study, by using Akaike Information Criteria (AIC), Final Prediction Error (FPE), Hannan-Quinn Information Criteria (HQ), Adjusted Akaike Information Criteria (AICC) and Schwarz Information Criteria (SIC), in selecting appropriate model degree in single variable time series are examined. In order to compare these criteria, Monte Carlo simulation method is employed to compute the lag lengths selected by each criteria.

Keywords: Model selection criteria, Simulation, Time series.