

Gamma Otoregresif Modeller ve Kızılırmak Havzasına Uygulama

Nermin ŞARLAK*
A. Ünal ŞORMAN**

ÖZ

Su kaynakları projelerinde uzun akım verilerinin sentetik olarak türetilmesi için stokastik modellere gereksinim duyulur. Bu tip modeller akım kayıtlarının kısa olduğu akarsular için ayrıca önemlidir. Otoregresif zaman serileri modelleri (AR) bu alanda en çok kullanılan modellerdir. Son yıllarda gamma dağılımı için geçerli olan AR modelleri geliştirilmektedir. Bu çalışmalarda, model parametrelerinin tahmininde momentler yöntemi ile uyarlanmış maksimum olabilirlik yöntemi kullanılmaktadır. Bu çalışmanın amacı, gamma dağılımı varsayımı altında iki AR modelinin karşılaştırılması ve Kızılırmak havzasında yer alan EIE 1501 ve EIE 1517 aylık akım gözlem istasyonu verilerine uygulanmasıdır. Her iki modelden elde edilen sonuçlar ileride bu modelleri kullanacak araştırmacılar için yorumlanmıştır.

ABSTRACT

Gamma Autoregressive Models and Application on Kızılırmak Basin

Stochastic models are required to generate synthetic values of flows statistically similar to observed ones to use in the simulation studies of water related structures. Such models are extremely important for generation of flows of streams with short record length. Autoregressive time series (AR) models are the most used models in this area. In recent years, AR models valid for gamma distribution are developed. In these studies, method of moments and modified maximum likelihood methods are used to estimate the model parameters. The aim of this study are comparing two AR models under gamma distribution assumption and applying these models for monthly data set of EIE 1501 and EIE 1517 streamflow gauging stations located in Kızılırmak Basin. The results of each model are interpreted for researchers using these models in their further studies.

Not: Bu yazı

- Yayın Kurulu'na 26.06.2006 günü ulaşmıştır.
- 30 Eylül 2007 gününe kadar tartışmaya açıktır.

* İnönü Üniversitesi, İnşaat Mühendisliği Bölümü, Malatya – nsarlak@inonu.edu.tr

** ODTÜ, İnşaat Mühendisliği Bölümü, Ankara – sorman@metu.edu.tr

1. GİRİŞ

Dünyada kurak ve yarı-kurak bölgelerdeki su ihtiyacı sulak bölgelerdekinden daha fazladır. Nüfus ve endüstrinin artması ile gerekli su miktarı ihtiyacı artmaktadır. Bu nedenle, gerekli su ihtiyacını karşılamak için mevcut akarsu, göl ve yeraltı sularından faydalanılması gerekmektedir. Doğal su kaynaklarındaki suların ihtiyaç duyulan bölgelerde kullanılabilmesi için akarsuların üzerine bazı yapıların yapılması gereklilik arz etmektedir.

Bu amaçla, akarsuların geçmiş hidrolojik bilgilerinin bilinmesi zorunludur ve akarsularda gözlemlenmiş verileri kullanarak iyi bir şekilde tanımlanması gerekmektedir. Bu sebepten gözlemlerin matematiksel yapısının ve uygun matematiksel modellerin araştırılarak sağlıklı sentetik serilerin türetilmesi için geliştirilmesi hidroloji alanında uzun yıllardır araştırılan bir konudur. Bu amaçla kullanılan matematiksel modellerden hidroloji alanında en yaygın olarak kullanılan otoregresif zaman serisi (AR) modelleridir. Bu modeller ise çeşitli varsayımlar üzerine dayalıdır. En önemli varsayımda serilerin veya hata terimlerinin normal dağıldığı varsayımdır. Söz konusu varsayımın yapılmasındaki başlıca neden, normal dağılım dışındaki dağılımlar için model parametrelerini tahmin etmede karşılaşılan zorluklardır.

Son yıllarda normal dağılım dışındaki dağılımlar için geçerli olan AR modelleri üzerine çalışmalara rastlanmaktadır. Obeysekera ve Yevjevich (1985), eldeki serinin ortalama, varyans ve çarpıklık değerlerini kullanarak gamma dağılımına uyan serilerin AR modeli ile türetilmesine yönelik yöntem önermişlerdir. Fernandez ve Salas (1986), normal dağılıma uymayan periyodik seriler için AR(1) modeli geliştirmişlerdir ve önerdikleri modeli haftalık akım serilerine uygulayıp, test etmişlerdir. Sim (1987), ARMA (1,1) modelini oldukça çarpık akım verilerinin türetilmesinde kullanmışlardır. Gamma dağılımı varsayımı altında model parametreleri, momentler yöntemi (MOM) ile tahmin edilmiştir. 100 aylık akım verilerine söz konusu model uygulanarak seriler türetilmiştir. Türetilen serilerin moment değerlerinin ölçülmüş verilerin moment değerlerine oldukça yakın olduğu saptanmıştır. Fernandez ve Salas (1990), akım verilerinin modellenmesinde gamma otoregresif (GAR(1)) modeli üzerinde çalışmışlardır. Momentler yöntemi ile tahmin edilen parametrelerin bağımlı ve normal olmayan değişkenler için yansız hale getirmek için bazı düzeltmeler önermişlerdir. Önerilen yöntem çeşitli nehirlerdeki yıllık akım verilerine uygulanmıştır ve GAR(1) modelinin sentetik akım türetmede oldukça cazip bir model olduğu iddia edilmiştir. Cığızoğlu ve Bayazıt (1998), Türkiye'deki 10 yıllık akım verisini kullanarak istatistiksel geçiş ve oran parametrelerini belirlemek için GAR(1) modelini kullanmışlardır. Söz konusu çalışmada, geçiş ve oran parametrelerini bulmada yansız model parametrelerinin bulunmasının ne kadar önemli olduğu saptanmıştır. Tiku ve ark. (1999), gamma hata terimli AR(1) modeli üzerinde çalışmışlardır. Model parametrelerini bulmak için uyarlanmış maksimum olabilirlik (MML) yöntemini önermişlerdir. Aynı yöntem simetrik olmayan hata terimli diğer dağılımlara da uygulanmıştır [1, 3, 14, 15].

Bu çalışmada, gamma dağılımı varsayımı altında GAR(1) ve gamma hata terimli AR(1) zaman serisi modellerinin Kızılırmak havzasında yer alan EIE 1501 ve EIE 1517 istasyonları 40*12 aylık (1956-1995) akım veri setine uygulanarak elde edilen sonuçların kıyaslanması amaçlanmaktadır. Farklı parametre tahmin yöntemleri (MOM ve MML) kullanılarak tahmin edilen model parametreleri ile kurulan her iki AR(1) modelinden sentetik seriler türetilmiştir. Türetilen 1000 adet sentetik serilerin ortalama, standart sapma gibi ilk iki moment değerleri hesaplanmış ve gözlemlenmiş akım serisinin moment

değerleri ile kıyaslanmıştır. Ayrıca, model kurmada kullanılmayan akım verilerini tahmin etmede modellerin performanslarını göstermek amacıyla sentetik seri de türetilmiştir ve model kurmada kullanılmayan gözlemlenmiş akım değerleri (1995-2002) ile grafiksel ve istatistiksel olarak karşılaştırılmıştır. Gamma hata terimli AR(1) modelinden elde edilen sonuçların GAR(1) modelinden elde edilen sonuçlara göre daha iyi sonuç verdiği saptanmıştır.

2. OTOREGRESİF ZAMAN SERİSİ MODELLERİ

Otoregresif zaman serisi modelleri hidroloji alanında en yaygın olarak kullanılan stokastik modellerden biridir. Bunun başlıca nedeni, otoregresif formun zaman bağımlılığına sahip olması (bugünkü zamandaki değişkenin bir önceki zamana bağımlı olması) ve kullanımlarının kolay olmasıdır [9]. AR(1) modeli aşağıdaki gibi formüle edilmektedir [2]:

$$x_i = \phi x_{i-1} + \varepsilon_i, \quad i=1,2,\dots,n \quad (1)$$

eşitlikteki, ϕ otoregresyon katsayısını ve ε_i ise bağımsız rasgele değişkeni veya hata terimini ifade etmektedir. Genelde literatürde eldeki mevcut serinin veya hata teriminin normal dağıldığı varsayılır $N(0, \sigma^2)$. Söz konusu varsayımın yapılmasının en önemli sebebi normal dağılım dışındaki dağılımlar için model parametrelerini tahmin etmede karşılaşılan zorluklardır. Bu sebepten normal dağılıma uymayan zaman serilerine çeşitli dönüşüm fonksiyonlarının uygulanarak normal hale getirilmesi önerilmektedir. Serinin logaritmasının alınarak normal hale getirilmesi bu dönüşüm fonksiyonlarından biridir. Serinin dönüşüm fonksiyonları kullanılarak orjinal yapısının bozulması istenmeyen bir durum olduğu için son yıllarda çarpık dağılımlar için geçerli olacak AR modelleri üzerinde çalışmalara rastlanmaktadır.

2.1 GAR(1) Modeli

Gamma otoregresif modeller, modellenecek serinin gamma dağıldığı varsayımı üzerine kuruludur. Tarihi serinin moment değerleri (ortalama, standart sapma, çarpıklık ve içsel bağımlılık katsayısı) toplumun moment değerlerini tahmin etmek için kullanılarak, model parametreleri MOM ile tahmin edilmektedir. Ancak, normal dağılım dışındaki dağılımlar için momentler yönteminden tahmin edilen parametreler yanlı (bias) olduğundan yansız model parametrelerinin tahmin etmek için bazı düzeltmelerin yapılması gerekmektedir. Söz konusu düzeltme formülasyonları literatürde mevcuttur [5].

GAR(1) model parametreleri toplum moment değerlerinden tahmin edildiği için türetilen serinin gamma dağılımına sahip olması gerekliliği kaçınılmazdır. Normal dağılımın aksine gamma dağılımına sahip değişkenlerin toplamı yine gamma dağılımı özellikleri göstermemektedir. Bu bakımdan momentler yöntemi ile tahmin edilen GAR(1) modeli parametreleri ile kurulan zaman serisinden, sentetik seri türetilmesi için Lawrance ve Lewis (1981) tarafından geliştirilen yöntem hata terimlerini hesaplamak için kullanılmıştır.

2.2 Gamma Hata Terimli AR(1) Modeli

Konum parametresinin, λ eşitlik 1'deki otoregresif zaman serisi modeline eklenmesi ile aşağıdaki eşitlik elde edilmektedir.

$$x_i = \lambda + \phi x_{i-1} + \varepsilon_i \quad (2)$$

GAR(1) modelinin aksine bu modelde eşitlik 2'deki hata teriminin gamma dağıldığı varsayılmaktadır. Maksimum olabilirlik yöntemi ile gamma dağılımı için geçerli olabilirlik denkleminin iteratif yöntem gerektirmeyen (explicit) çözümü olmadığından, iteratif çözüm yöntemlerinin kullanılmasını gerektirmektedir. Bu da çözümde çeşitli problemlere yol açmaktadır. Bu problemler arasında i) tek bir çözüm değerinin elde edilememesi veya çözüm değeri elde edilememesi ii) çözüm değerine çok yavaş yakınsaması veya yanlış değere yakınsaması ve iii) hiç bir değere yakınsamaması sayılabilir. Bundan dolayı, gamma dağılımı için geçerli olabilirlik denklemlerinin iteratif yöntem gerektirmeyen çözümlerini elde edebilen MML eşitlikleri literatürde mevcuttur [13]. Model parametrelerini tahmin etmede momentler yöntemi kullanılmadığı için sentetik serilerin üretiminde Lawrance ve Lewis (1981) tarafından geliştirilen yöntem kullanılmaksızın gamma dağılmış rasgele değişkenlerin modele eklenmesi yeterli olmaktadır [11].

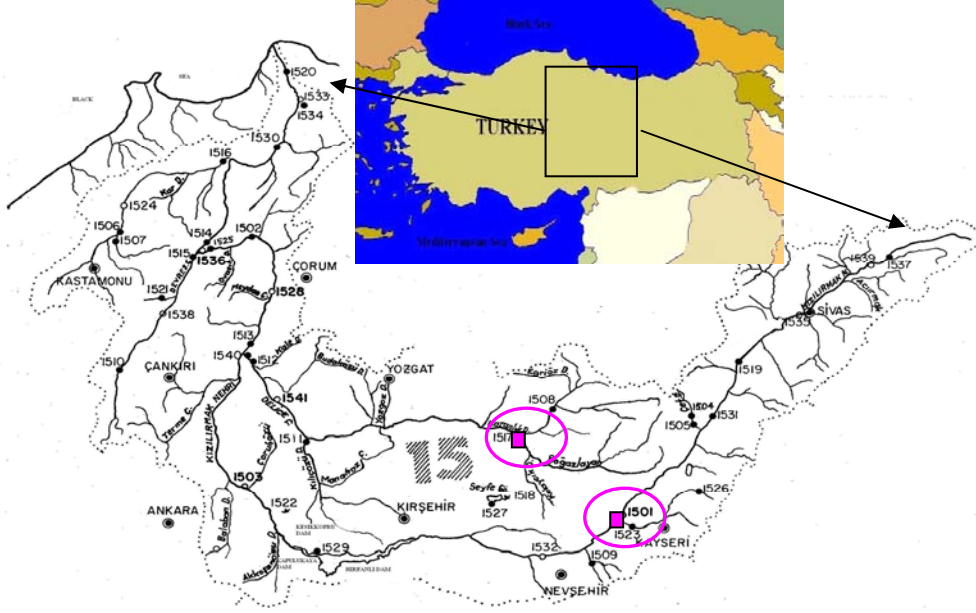
3. KIZILIRMAK HAVZASINA UYGULAMALAR

Modellerin uygulamaları için Kızılırmak havzasında Kızılırmak nehri üzerindeki EIE 1501 ve EIE 1517 akım gözlem istasyonlarının aylık akım verileri seçilmiştir. Şekil 1'de gösterilen EIE 1501 ve EIE 1517 aylık akım gözlem istasyonu verileri, gözlem sürelerinin uzun olması, ayrı ana kol üzerinde olmaları ve hiç bir su yapısından etkilenmiyor olmaları nedeniyle tercih edilmiştir. Belirtilen istasyonlar Elektrik İşleri Etüd İdaresi (EIE) tarafından işletilmektedir.

Gözlemlenmiş 1955-2002 yılına kadar olan 48*12 aylık akım veri seti mevcuttur. Çalışmanın amacı, farklı model parametre tahmin yöntemleri ile kurulan AR(1) model yapısının yapay seri türetme üzerindeki etkilerini göstermektir. Bu amaçla 1955 yılının ocak ayında ölçülen akım değeri başlangıç değeri olarak alınması durumunda 492 (1955-1995) adet aylık gözlem veri seti model yapılarını oluşturmak için, 84 aylık (1996-2002) veri seti ise kurulan AR(1) modellerinin geçerliliğini tartışmak için kullanılmıştır.

492 aylık veri seti kullanılarak GAR(1) modeli ve gamma hata terimli AR(1) modeli parametreleri MOM ve MML yöntemleri kullanılarak tahmin edilmiştir. Tahmin edilen parametreler Çizelge 1'de gösterilmiştir. Bilindiği gibi ML yöntemi tüm model parametrelerinin olabilirlik fonksiyonunu maksimum yapacak model parametrelerinin elde edilebilmesi üzerine kuruludur.

Bu sebepten, MML yöntemi ile tahmin edilen parametrelerle olabilirlik fonksiyonunun maksimum hale gelip gelmediğinin kontrol edilemesi gerekir. Bu çalışmada, olabilirlik fonksiyonunu maximum yapan herhangi bir şekil parametre değeri elde edilememiştir.



Şekil 1. Kızılırmak Havzası ve Akım Gözlem İstasyonunun Yeri

Bu da eldeki veriler için gamma dağılımının uygun olmadığını bir işaretidir. Bu çalışma da iki farklı parametre tahmin yöntemini karşılaştırmak amaçlandığından Gamma hata terimli AR(1) modelinde uygun şekil parametresi elde edilememesine rağmen GAR(1) modelinden elde edilen şekil parametresi değerinin doğru olduğu kabul edilmiştir. Model parametreleri her iki yöntem ile tahmin edilmiş ve sentetik seriler türetilmiştir. Gerçekte eldeki verilere en uygun dağılımlar belirlenip MML yöntemi ile parametrelerin tahmin edilmesi ve uygun modellerin kurulması durumunda daha sağlıklı sonuçların elde edilmesi gerekmektedir [12].

Çizelge 1. EIE 1501 ve EIE 1517 istasyonu verilerinin model parametre tahmin değerleri

İstasyon	Modeller / Parametreler	k şekil	$\hat{\lambda}$ konum	$\hat{\phi}$ otoregresyon	$\hat{\sigma}$ ölçek
EIE 1501	GAR(1)	2.05	-1.421	0.6634	0.6933
	Gamma hata terimli AR(1) modeli	2.05	-1.195	0.4232	0.6099
EIE 1517	GAR(1)	3.05	-1.738	0.7495	0.5694
	Gamma hata terimli AR(1) modeli	3.05	-1.272	0.4742	0.4241

3.1 Modellerin Değerlendirilmesi

Tahmin edilen parametreler ile kurulan modeller sentetik serilerin türetilmesinde kullanılmıştır. Yazılan fortran programı ile 1000 adet sentetik seri türetilmiş ve ortalama, standart sapma değerleri elde edilerek tarihi serinin ortalama, standart sapma değerleri ile karşılaştırılması amacı ile Çizelge 2’de gösterilmiştir.

Çizelge 2. Her iki istasyon için modellerden ve tarihi seriden elde edilen moment değerleri

İstasyon No.	Moment	Modeller		
		GAR(1)	Gamma hata terimli AR(1)	Tarihi seri
EIE 1501	Ortalama (m ³ /s)	68.27	71.58	68.48
	Standart sapma (m ³ /s)	78.68	80.45	78.87
	Çarpıklık katsayısı	2.40	2.23	2.31
EIE 1517	Ortalama (m ³ /s)	11.65	11.96	11.68
	Standart sapma (m ³ /s)	12.04	11.30	12.07
	Çarpıklık katsayısı	1.74	1.45	1.71

Modellerin performanslarını değerlendirmek için tarihi seri moment değerleri ile göreceli hatalar her bir istasyon için elde edilmiş ve Çizelge 3’de gösterilmiştir. Göreceli hata değerleri modelden elde edilen moment değerlerinden tarihi serinin moment değerlerinin çıkartılıp modelden elde edilen moment değerlerine bölünerek 100 ile çarpılmasıyla hesaplanmıştır.

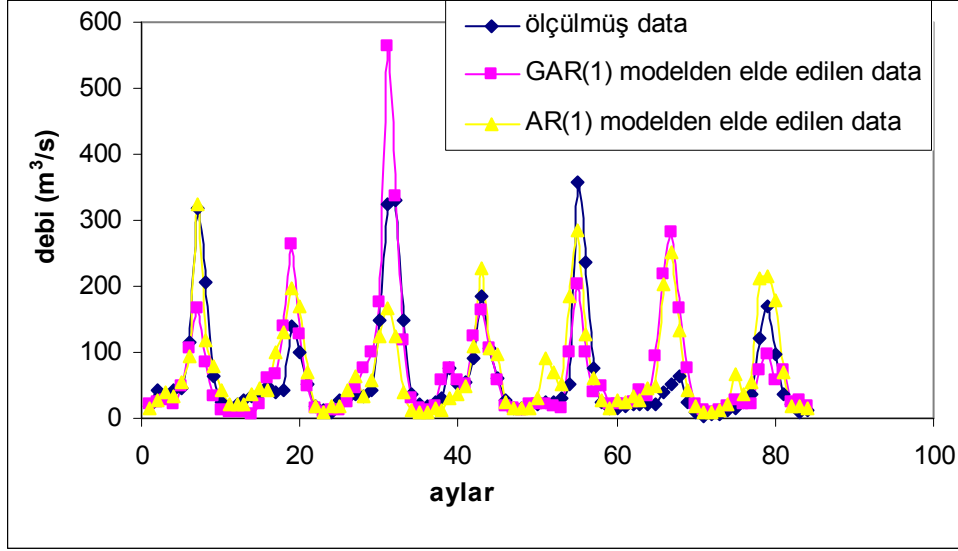
Çizelge 3. Her iki istasyon için modellerden elde edilen göreceli hata değerleri

İstasyon No.	Moment	Modeller	
		GAR(1) %	Gamma hata terimli AR(1) %
EIE 1501	Ortalama (m ³ /s)	-0.31	4.33
	Standart sapma(m ³ /s)	-0.24	1.96
	Çarpıklık katsayısı	3.75	-3.58
EIE 1517	Ortalama (m ³ /s)	-0.26	2.34
	Standart sapma(m ³ /s)	-0.25	-6.81
	Çarpıklık katsayısı	1.72	-17.93

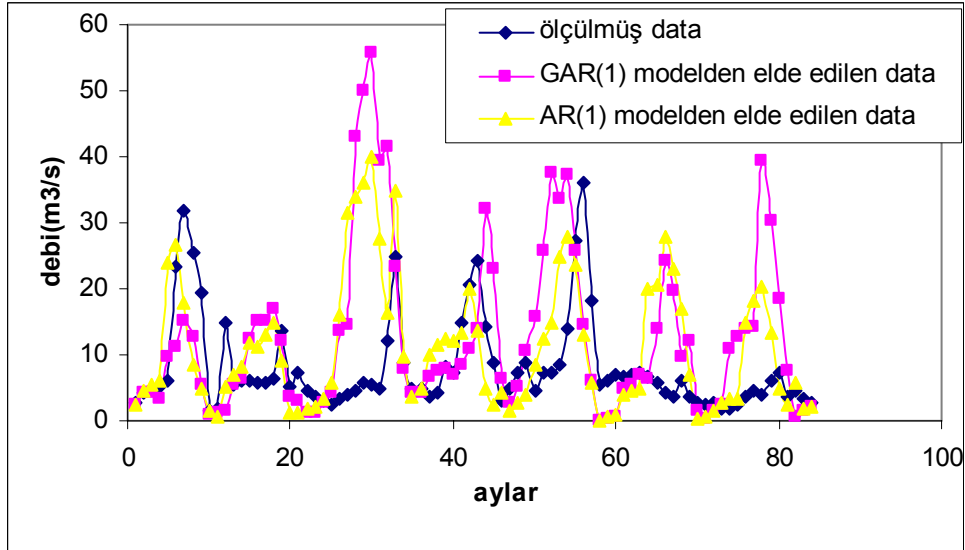
Çizelge 3’den görüleceği gibi GAR(1) modelinden elde edilen moment değerleri tarihi serinin 1. ve 2. moment değerlerine diğer modele göre daha yakındır. Bunun nedeni açıktır. GAR(1) modelinde, model parametreleri tarihi serinin moment değerleri kullanılarak elde edilmiştir. Tarihi serinin moment değerlerini muhafaza eden modelden tarihi serinin moment değerlerine yakın sonuçlar elde edilmesi beklenen bir sonuçtur.

Modellerin parametre tahmininde kullanılmayan verileri tahmin etmedeki başarısını göstermek amacı ile 1995 yılı aralık ayı gözlenmiş değer başlangıç değeri olarak kullanılarak 1996’dan 2002 yılına kadar gözlenmiş ve model kurmada kullanılmayan

veriler modellerden tahmin edilerek gözlenmiş verilerle grafiksel olarak kıyaslama amacı ile Şekil 2 ve Şekil 3’de gösterilmiştir.



Şekil 2. EIE 1501 istasyonu gözlemlenmiş ve her iki modelden elde edilen sentetik seri değerleri



Şekil 3. EIE 1517 istasyonu gözlemlenmiş ve her iki modelden elde edilen sentetik seri değerleri

İstatistiksel olarak kıyaslama ise ortalama hata karelerinin kare kökü (RMSE) ile yapılmıştır. RMSE, hedeften sapma olarak tanımlanan ve modellerin değerlendirilmesinde kullanılan bir metottur. Her iki modelden elde edilen değerler Çizelge 4’de verilmiştir.

Çizelge 4. Modellerin ortalama hata karelerinin kare kökü

İstasyon	Modeller	RMSE
EIE 1501	GAR(1)	59.36
	Gamma hata terimli AR(1)	55.01
EIE 1517	GAR(1)	14.20
	Gamma hata terimli AR(1)	10.87

Çizelge 4’den görüleceği gibi her iki modelden elde edilen RMSE değerleri birbirine yakın sonuçlar vermekle birlikte Şekil 2 ve Şekil 3’den görüleceği üzere gamma hata terimli AR(1) modelden elde edilen sentetik serilerin gözlenmiş veri dizinini daha iyi izlediği görülmektedir.

4. SONUÇLAR

Bu çalışmada hidroloji literatüründe çalışılmış olan gamma dağılımı varsayımı altında AR(1) model parametrelerinin diğer bir yöntem olan MML ile tahmin edilmesi amaçlandığından gözlenmiş serinin hangi dağılıma daha iyi uyduğu araştırılmamıştır. Sunulan çalışmada, gamma hata terimli AR(1) modelinden elde edilen sonuçların GAR(1) modelinden elde edilen sonuçlar kadar iyi olduğu gösterilmeye çalışılmıştır.

Kullanımının kolay olması, iteratif çözüm gerektirmemesi ve orjinal seride herhangi bir dönüşüm yapılmasına gerek duyulmaması gibi özellikler MML metodunun yapay akım serilerinin türetilmesinde oldukça cazip, yansız ve sağlam (robust) alternatif bir yöntem olduğunu göstermektedir. Ayrıca MML yönteminin gözlenmiş serilere uygun dağılımlar elde edilmesi halinde model parametrelerinin elde edilmesinde esneklik ve avantajlar sağlayacağı açıktır.

Semboller

- AR(1) : Otoregresif zaman serisi modeli
- GAR(1) : Gamma otoregresif zaman serisi modeli
- k : Şekil parametresi
- MOM : Momentler yöntemi
- ML : Maximum olabilirlik yöntemi
- MML : Uyarlanmış maximum olabilirlik yöntemi
- RMSE : Hata karelerinin kare kökü
- x_i : Akım değişkeni

- ε_i : Bağımsız rasgele değişken veya hata terimi
 ϕ : Otoregresyon katsayısı
 σ : Ölçek parametresi veya varyans

Kaynaklar

- [1] Akkaya, A.D., Tiku, M.L., Corrigendum: Time Series Models with Asymmetric Innovations. *Commun. Stat.- Theory Meth.*, 30 (10), 2227-2230, 2001.
- [2] Box, G. E. P., Jenkins, G. M., *Time Series Analysis: Forecasting and Control*, Holden-Day, San Francisco, Calif., 1976.
- [3] Akkaya, A.D., Tiku, M.L., Time Series AR(1) Model for Short Tail Distribution. *Statistics*, 39(2), 117-132, 2005.
- [4] Cıgızoğlu, K., Bayazıt, M., Application of Gamma Autoregressive Model to Analysis of Dry Periods. *Journal of Hydrologic Engineering*, 3(3), 218-221, 1998.
- [5] Fernandez, B., Salas, J.D., Periodic Gamma Autoregressive Processes for Operational Hydrology. *Water Resources Research*, 22(10), 1385-1396, 1986.
- [6] Fernandez B., Salas J.D., Gamma Autoregressive Models for Stream- Flow Simulation. *Journal of Hydraulic Engineering*, 116 (11), 1403-1414, 1990.
- [7] Lawrance, A.J., Lewis, P.A.W., A New Autoregressive Time Series Model in Exponential Variables[NEAR(1)]. *Adv. Appl. Prob.*, 13(4), 826-845, 1981.
- [8] Obeysekera, J.T.B., Yevjevich, V., A Note on Simulation of Samples of Gamma-Autoregressive Variables. *Water Resources Research*, 21(10), 1569-1572, 1985.
- [9] Salas, J.D., Delleur, J.W., Yevjevich V., Lane W.L., *Applied Modeling of Hydrologic Time Series*, Water Resources Publications, 1980.
- [10] Sim C. H., A Mixed Gamma ARMA(1,1) Model for River Flow Time Series. *Water Resources Research*, 23(1), 32-36., 1987.
- [11] Sarlak, N., Sorman A.U., Farklı Parametre Tahmin Metotları ile Gamma Otoregresif Modeller ve Uygulaması. II. Ulusal Su Mühendisliği Sempozyumu, İzmir, 89-94, 2005.
- [12] Sarlak, N., Evaluation and Modeling of Streamflowdata: Entropy Method, Autoregressive Models with Asymmetric Innovations and Artificial Neural Networks., Ph. D thesis, METU, Ankara, 2005.
- [13] Tiku, M.L., Wong, W.K, Bian, G., Time Series Models with Asymmetric Innovations. *Commun. Statist.- Theory Meth.*, 28(6), 31-1360, 1999.
- [14] Tiku, M.L., Wong, W.K. Vaughan, D.C., Bian, G., Time Series Models in Non-normal Situations: Symmetric Innovations. *Journal Time Series Analysis*, 21(5), 571-596, 2000.
- [15] Tiku, M.L., Akkaya A.D., *Robust Estimation and Hypothesis Testing*, New Age International Publishers(P): New Delhi, 2004.