

ARMAX Modelleri ve Porsuk Barajı Su Seviyesinin Öngörüsü

Hülya Şen^a ve Özer Özaydın^a

^aEskişehir Osmangazi Üniversitesi, Fen-Edebiyat Fakültesi, İstatistik Böl., 26480, Eskişehir
e-posta: hsen@ogu.edu.tr , oozaydin@ogu.edu.tr

Geliş Tarihi: 24.05.2010; Kabul Tarihi: 17.08.2010

Özet

Bu çalışmada, istatistikte öngöründe kullanılan dışsal değişken kaynaklı otoregresif hareketli ortalamalar (autoregressive moving average with exogenous input, ARMAX) modelleri incelenmiş ve söz konusu modeller Porsuk Barajı su seviyesinin öngörülmesinde kullanılmıştır. Çalışmada öncelikle öngörü modellemesi olan ARMAX modelleri açıklanmıştır. Çalışmanın sonraki bölümünde ise Porsuk Barajı su seviyesinin ARMAX modelleri ile öngörülmesi elde edilmiştir. Çalışmadan elde edilen Porsuk barajının su seviyesinin öngörüsü, ileride su sıkıntısı yaşanıp yaşanmayacağı veya su taşkınlarının olup olmayacağına bir göstergesi olabilecektir. Yapılan öngörüler, Porsuk barajı su seviyesinin bir önceki gerçekleşen değerlerine göre düşüşlerin olduğunu göstermiştir.

Anahtar Kelimeler: ARMAX, Öngörü, Porsuk Barajı, Su Seviyesi

ARMAX Models and Forecasting Water Level of Porsuk Dam

Abstract

In this study, ARMAX (autoregressive moving average with exogenous input) models which are used for forecasting in statistics have been investigated and this models have been applied to forecast the water level of Porsuk dam. First of all, ARMAX models among the forecasting models have been explained in this study. After that, forecasts of water level of Porsuk dam have been obtained with ARMAX models. The results of forecasting of water level in Porsuk dam obtained in this study could give significant information whether there would be any water shortage or water flood in the near future. The forecasts of the application demonstrated that generally there was a decrease in water levels of Porsuk dam in contrast to their previous values.

Key Words: ARMAX, Forecasting, Porsuk Dam, Water Level

1. Giriş

Zamana bağlı bir değişkenin gelecekte hangi değeri alacağı çeşitli istatistiksel yöntemlerle belirlenebilmektedir. Bunlardan bir tanesi Box ve Jenkins tarafından geliştirilen öngörü modellemesi olan otoregresif birleştirilmiş hareketli ortalamalar (autoregressive integrated moving average, ARIMA) modelleridir. ARIMA modelleri ile doğru modelleme yapılırsa söz konusu seri için oldukça iyi öngöründe

bulunulabilmektedir. Ancak ARIMA modellerinin kurulumunda dikkate alınmayan ve modelde yer alması daha iyi öngörülere sebep olacak seriyi etkileyen değişkenler söz konusu olabilir. Böyle değişkenlerin modelde yer alması modeli karmaşıklaştırmasına rağmen eğer uygun modelleme yapılırsa ARIMA modellerine göre daha iyi öngörüler yapılabilmektedir. Eğer zamana bağlı bir değişkeni etkileyen bir veya birden fazla böyle değişken varsa bu defa dışsal değişken kaynaklı otoregresif hareketli

ortalamalar (autoregressive moving average with exogenous input, ARMAX) söz konusu olmaktadır.

Literatürde karmaşık tekniklerin, basit öngörü tekniklerine kıyasla, karmaşık tekniklere başvurmayı gerektirecek kadar bir doğruluk kazancı sağlayamadıklarını ve hatta basit tekniklerin karmaşık tekniklerden daha uygun öngöründe bulunduğunu Makridakis ve Hibon (1997) ileri sürmüştür. Böyle bir sonuca ancak karmaşık tekniklerin yeterince denenmemesi ve uygun bir metodoloji takip edilmemesi sonucu ulaşılabilir. Ayrıca Makridakis ve Hibon (1997) AR(1) ve ARMA(1,1) modellerinin ARIMA modelleri kadar ve hatta daha doğru öngöründe bulduklarını, bu sebepten dolayı en uygun ARMA modelini belirlemede oto korelasyon, kısmi otokorelasyon grafiklerini incelemeye ve bilmeye gerek olmadığını ileri sürmektedir. Çünkü Box-Jenkins metodolojisine göre tahmin edilen herhangi bir modelin hata terimleri rassal bir dağılıma sahipse o model uygun model olarak kabul edilmektedir. Diğer taraftan Mahmoud (1984), Box-Jenkins tekniğinin regresyona üstün geldiğini işaret etmiştir (Akai, 2003). Ayrıca Akai (2002) çalışmasında ARMAX tekniğinin diğerlerine üstün geldiğini belirlemiştir.

Öngörü modelleri özellikle ekonomi ve mühendislik alanları olmak üzere birçok alanda kullanılmaktadır. Tek değişkenli zaman serileri ile öngörü yapılabileceği gibi çok değişkenli zaman serileri ile de öngörü yapılabilmektedir. Çok değişkenli zaman serisinin öngörüsünde ise tek bağımlı değişkenli zaman serisinin öngörüsü yapılabildiği gibi birden fazla bağımlı değişkenli zaman serisinin de öngörüsü yapılabilmektedir. Tek değişkenli öngörü modellemesinde en yaygın kullanım alanına sahip ARIMA modelleridir. Bu modellere Box-Jenkins modelleri de denmektedir.

2. Sistem tanımlama ve ARMAX modelleri

Özellikle mühendislik alanlarında öngörü modelleri sistem tanımlama başlığı altında incelenmektedir. Sistem tanımlama çeşitli alanlarda birçok farklı modelle gösterilmektedir (Ljung, 1999).

Sistem tanımlama, veriden elde edilen farklı yapılarla modellerin tanımlanması ve model performansının karşılaştırılmasının yapıldığı iteratif bir süreçtir (Ljung, 2008a,b). Bir sistemin tanımlanmasından genellikle mühendislik alanında bahsedildiği için, istatistiksel olarak bağımlı ve bağımsız olarak bilinen değişkenler birçok kaynakta girdi ve çıktı olarak isimlendirilmektedir. Basit model yapılarının parametrelerinin tahminlenmesi ile sürece başlanır. Model performansı düşük olursa, model yapısının karmaşıklığı yavaş yavaş artırılır. Eninde sonunda sistem dinamiklerini en iyi tanımlayan basit model seçilir. Basit model yapıyla başlanmasının nedeni, yüksek dereceden modeller her zaman tam doğru değildir. Model karmaşıklığının yükseltilmesi model parametrelerindeki belirsizliği artırır ve genellikle daha fazla veri gerektirir (Ljung, 2008a,b).

Modellemede kullanılan seriler zaman temelli olabileceği gibi frekans temelli de olabilmektedir. Frekans temelli modellerde Fourier dönüşümü kullanılmaktadır. Zaman temelli modelleme ise sürekli ve kesikli olarak ikiye ayrılmaktadır. Ayrıca modeller, probleme bağlı olarak, tek bağımlı değişkene (tek çıktılı) veya çok bağımlı değişkene (çok çıktılı) göre oluşturulabilmektedir (Ljung, 2008a,b).

Bazı durumlarda sistemin davranışı fiziksel yasalar ile ideal bileşenler dikkate alınarak matematiksel model ile tanımlanabilmektedir. Bu tanımlama beyaz kutu modelleme adını almaktadır. Bu tip modelde sürecin bütün bilgisi ele alınmaktadır. Ancak bu her zaman mümkün olmamaktadır. Beyaz kutu modeller ile sistem

tanımlama çok faydalı bir yaklaşım olmasına rağmen, sistem çok karmaşık tanımlanabilmekte ya da yeterince öncü bilgiye ihtiyaç duyulmaktadır. Bu yüzden girdi ve çıktı değişkenlerinden faydalanılarak tanımlanmaktadır. Bu tür modeller de kara-kutu modelleri olarak adlandırılmaktadır. Kara-kutu modelleri parametre sayısına bağlı olarak parametrik ve parametrik olmayan modeller olmak üzere ikiye ayrılmaktadır. Eğer modeldeki parametre sayısı sonlu ise parametrik modeller tercih edilmektedir. Çünkü böyle modellerle çalışmak çok daha kolay olmaktadır (Erdoğan v.d., 2005).

Zaman temelli parametrik modeller yazılmak istenirse, aşağıdaki eşitlik gibi genel bir doğrusal model yapısı oluşturulur.

$$A(q)y(t) = \sum_{i=1}^n \frac{B(q)}{F(q)} x_i(t - nk_i) + \frac{C(q)}{D(q)} e(t) \quad (1)$$

Burada, A, B, C, D ve F sırasıyla n_a , n_b , n_c , n_d ve n_f boyutlarıyla modelin polinomlarını; nk_i , i . bağımsız değişkenin gecikmesini; q , zamanda geriye öteleme operatörünü ($q^{-n}y(t) = y(t-n)$) (Z dönüşüm formuyla tamamen aynıdır.); t zaman indeksini; y , bağımlı değişkeni (çıkıtı); x_i , bağımsız değişkenleri, e ise ortalaması 0 varyansı σ^2 olan hata terimini ifade etmektedir.

Aşağıda 2, 3, 4 ve 5 eşitlikleri ile ifade edilen parametrik modeller 1 eşitliğinde tanımlanan genel modelin bazı özel durumlarıdır. Model yapıları arasındaki temel fark, bozucu etkilerin modellenme şekline kaynaklanmaktadır (Erdoğan ve ark., 2005; Ljung, 2008a,b).

Eşitlik 1’de C, D ve F polinomlarının 1’e eşit olduğu durumlarda dışsal kaynaklı otoregresif model (ARX(n_a , n_b , nk), Autoregressive with Exogenous Input) söz konusudur. Bir bağımlı ve bir bağımsız değişkenli ARX modeli aşağıdaki eşitlik 2’deki gibi ifade edilmektedir.

$$A(q)y(t) = B(q)x(t - nk) + e(t) \quad (2)$$

Bozucu etkiler otoregresif ile filtrelenmiş hata terimi olarak modellenmektedir. n_a ve n_b modelin derecelerini (n_a , y ile ifade edilen bağımlı değişkenin parametre sayısı, n_b ise x ile ifade edilen bağımsız değişkenin parametre sayısıdır), nk ise gecikmeyi ifade etmektedir.

Yine eşitlik 1’de A, C ve D polinomlarının 1’e eşit olduğu durumda çıktı hatalı model (OE, Output Error) elde edilir (Eşitlik 3).

$$y(t) = \frac{B(q)}{F(q)} x(t - nk) + e(t) \quad (3)$$

Aynı eşitlikte A polinomunun 1’e eşit olduğu durumda Box-Jenkins (ARIMA) modeli söz konusu olmaktadır (Eşitlik 4).

$$y(t) = \frac{B(q)}{F(q)} x(t - nk) + \frac{C(q)}{D(q)} e(t) \quad (4)$$

Aynı genel model yapısında F ve D polinomlarının 1 olduğu durumda aşağıdaki eşitlikte gösterilen ARMAX model elde edilir.

$$A(q)y(t) = B(q)x(t - nk) + C(q)e(t) \quad (5)$$

Bu tip modellere transfer fonksiyonu modelleri de denilmektedir. Transfer fonksiyonu, bir dinamik regresyon modeli formu olarak açıklayıcı zaman serisi filtresi anlamına gelmektedir. Eşitlik 5’te eşitliğin her iki tarafı da $A(q)$ ’ya bölüldüğünde t anındaki bağımlı değişkenin değeri elde edilmektedir. Bağımsız değişkenin polinomu olan $\frac{B(q)}{A(q)}$, ya bağımsız değişkenin bağımlı değişken üzerindeki etkisinden dolayı transfer fonksiyonu veya doğrusal filtre adını almaktadır. Transfer fonksiyonu olarak isimlendirilmesinin sebebi ise bağımsız değişken üzerindeki değişkenliğin bağımlı değişken

üzerindeki değişkenliğe transfer edilmesidir (Beltran,1993).

Eşitlik 5'teki A, B ve C polinomları aşağıda verilmiştir:

$$A(q) = 1 + a_1q^{-1} + a_2q^{-2} + \dots + a_{na}q^{-na} \quad (6)$$

$$B(q) = b_0 + b_1q^{-1} + b_2q^{-2} + \dots + b_{nb}q^{-nb} \quad (7)$$

$$y(t) + a_1y(t-1) + \dots + a_{na}y(t-na) = b_0x(t) + b_1x(t-1) + \dots + b_{nb}x(t-nb) + e(t) + c_1e(t-1) + \dots + c_{nc}e(t-nc) \quad (9)$$

Yukarıdaki eşitlik matris şeklinde yazılacak olursa sırasıyla aşağıdaki parametre vektörü ve veri vektörü elde edilir (Eşitlik 10, 11).

$$\theta^T = [a_1 \dots a_{na} \ b_1 \dots b_{nb} \ c_1 \dots c_{nc}] \quad (10)$$

$$\varphi^T = [-y(t-1) \dots -y(t-na) \ -x(t-1) \dots -x(t-nb) \ -e(t-1) \dots -e(t-nc)] \quad (11)$$

Böylece $\hat{y}(t)$ değişkenini veri ve parametre vektörüne bağlayan eşitlik 12 elde edilir.

$$\hat{y}(t) = \varphi^T(t)\hat{\theta}^T + e(t) \quad (12)$$

Burada $\varphi^T(t)$ ile θ^T arasında doğrusal bir ilişki olduğundan bu eşitlik parametreler bakımından doğrusal bir eşitliktir ve parametre hesaplama algoritmalarının başlangıç noktasıdır (Ljung, 1999).

2.1. ARMAX model parametrelerinin elde edilmesi

Model parametrelerinin belirlenmesi için farklı teknikler mevcuttur. Bunlar; yinelemeli en küçük kareler tekniği, filtrelenmiş yinelemeli en küçük kareler tekniği, yinelemeli geliştirilmiş en küçük kareler tekniğidir.

$$C(q) = 1 + c_1q^{-1} + c_2q^{-2} + \dots + a_{nc}q^{-nc} \quad (8)$$

Eşitlik 7'deki b_0 model sabitidir ve genelde model gösterimlerinde kullanılmaz. Yukarıdaki eşitliklerden yararlanılarak eşitlik 5'teki ARMAX model açık şekilde yazılacak olursa, eşitlik 9 elde edilir.

2.1.1. Yinelemeli en küçük kareler tekniği

Sistem tanımlama için kullanılan en uygun tekniklerden biridir. Kendinden ayarlamalı kontrol yöntemlerinde, hesaplanan sistem model parametreleri her örnekleme zamanında gelen yeni veriler kullanılarak tekrar hesaplanmaktadır. Yinelemeli en küçük kareler algoritması aşağıda verilmiştir.

t+1 anında;

a) Yeni bağımlı ve bağımsız değişken verileri kullanılarak $\varphi(t+1)$ vektörü oluşturulur.

b) $\varepsilon(t+1)$ tahmin hatası hesaplanır.

$$\varepsilon(t+1) = y(t+1) - \varphi^T(t+1)\hat{\theta}(t) \quad (13)$$

c) Kovaryans matrisi $\rho(t+1)$ hesaplanır.

$$\rho(t+1) = \rho(t) - \frac{\rho(t)\varphi(t+1)\varphi^T(t+1)\rho(t)}{1 + \varphi^T(t+1)\rho(t)\varphi(t+1)} \quad (14)$$

d) Parametre vektörü güncelleştirilir.

$$\hat{\theta}(t+1) = \hat{\theta}(t) + \rho(t+1)\varphi(t+1)\varepsilon(t+1) \quad (15)$$

Bir sonraki zamana kadar beklenir ve ilk adıma geri dönülür (Akay, 2004).

2.1.2. Filtrelenmiş yinelemeli en küçük kareler tekniği

Bu yöntemde bağımlı ve bağımsız değişkenlerden oluşan veri vektörü (φ^T) uygun bir polinom ile filtrelenir. y_f ve x_f sırasıyla filtrelenmiş bağımlı ve bağımsız değişkenleri gösterirse eşitlik 16 elde edilir.

$$Ay_f(t) = Bx_f(t-1) + e(t) \quad (16)$$

Bu durumda veri ve parametre vektörü sırasıyla aşağıdaki 17 ve 18 eşitlikleri gibi;

$$\varphi^T = [-y_f(t-1), \dots, -y_f(t-n) \quad x_f(t-1), \dots, x_f(t-m-1)] \quad (17)$$

$$\theta^T = [a_1, \dots, a_n \quad b_0, \dots, b_m] \quad (18)$$

yazılabilir (Akay, 2004).

2.1.3. Yinelemeli geliştirilmiş en küçük kareler tekniği

Yinelemeli Geliştirilmiş En Küçük Kareler algoritmasında hata teriminin parametreleri de hesaplanmaktadır. ARMAX modelinde yer alan C polinomunun katsayılarının tahmini için $e(t-1)$, $e(t-2)$,, $e(t-nc)$ değerlerinin bilinmesi gerekmektedir. Pratikte $e(t)$ değerleri ölçülemediğinden tahmin hatası $\varepsilon(t)$, eşitlik 19'dan hesaplanarak $e(t)$ değerlerinin yerine kullanılır.

$$\varepsilon(t) = y(t) - \varphi^T(t)\hat{\theta}(t-1) \quad (19)$$

Yinelemeli Geliştirilmiş En Küçük Kareler Yöntemi için algoritma aşağıda verilmiştir.

$t+1$ anında;

- a) Yeni bağımlı ve bağımsız değişkenlerin verileri kullanılarak $\varphi(t+1)$ vektörü oluşturulur.
- b) Eşitlik 20'den tahmin hatası hesaplanır.

$$\varepsilon(t+1) = y(t+1) - \varphi^T(t+1)\hat{\theta}(t) \quad (20)$$

- c) Eşitlik 21'den kovaryans matrisi yenilenir.

$$\rho(t+1) = \rho(t) - \frac{\rho(t)\varphi(t+1)\varphi^T(t+1)\rho(t)}{1 + \varphi^T(t+1)\rho(t)\varphi(t+1)} \quad (21)$$

- d) 22 eşitliği ile parametre vektörü yenilenir.

$$\hat{\theta}(t+1) = \hat{\theta}(t) + \rho(t+1)\varphi(t+1)\varepsilon(t+1) \quad (22)$$

Bir sonraki zamana kadar beklenir ve ilk adıma geri dönülür (Akay, 2004, Köksal ve Öztürk, 2004).

2.2. ARMAX Tahmin Algoritması

ARMAX parametre tahmini gerçekleştirilirken dikkat edilmesi gereken bir dizi ek noktalar vardır. Bunlar, izleyen paragraflarda ele alınmıştır.

2.2.1. Örnekleme oranı veya veri sayısı

Model parametrelerinin kestiriminde en küçük kareler tekniğinin kullanıldığı durumda bağımlı ve bağımsız değişkenlerin veri vektörlerindeki örneklerin sayısı önemlidir. Minimum veri sayısı modelin büyüklüğüne göre belirlenir (Moore et al., 2007).

Örnekleme oranı ARMAX modelinin kestiriminde ve daha sonradan ortaya çıkan alt parametrelerin hesaplanmasında da kritik bir öneme sahiptir. Ljung 1999, tahmin edilen modeldeki yanlışlık ve varyans güven aralıklarının örnekleme etkisini tartışmış ve çok yüksek örnekleme oranlarının kesikli zaman modellerinde

sayısal problemler çıkaracağını belirtmiştir. Diğer yandan da çok düşük örneklem oranı sistem dinamiklerinin temsil edilmesi açısından çok düşük yoğunluklu sonuçlar ortaya çıkarmaktadır. Sonuç olarak örneklem oranının en uygun seçimi sistem zamanı kısıtları çerçevesinde belirlenmelidir.

2.2.2. Model parametrelerinin sayısının ve gecikmenin belirlenmesi (na, nb, nc, nk)

Fassois 2001, model parametrelerinin sayısını ve gecikmeyi belirlemede bir yaklaşımda bulunmuştur. $na = na_{min} \dots , na_{max}$ kriterinden hareketle en iyi modeli belirlemek için $nc=na$ yapmıştır. Daha sonra da nc 'deki değişimin etkinliğini test etmiştir (Moore et al., 2007). Model parametrelerini belirlemek için birkaç adım kullanılmaktadır. Birinci adımda nk 'yi belirlemek için na , nb ve nc 'ye aynı değerleri vererek nk 'nın hangi değerinde model seçim yaklaşımlarından biri (Örneğin AIC, BIC, FPE, HKO vb.) en düşük değeri veriyorsa gecikme değeri olarak o değer alınır. Genellikle söz konusu değer 2'yi geçmemektedir. İkinci adımda ise yine model seçimi yaklaşımlarından biri ile olabilecek tüm kombinasyonlar denenerek uygun na , nb , nc ve nk belirlenmeye çalışılır (Koulocheris and Dertimanis, 2008). Alternatif olarak ise istatistiksel model derecesi belirleme kriterleri kullanılarak uygun model dereceleri belirlenebilir.

2.2.3. Model seçimi

Literatürde öngörude kullanılacak modelin seçimi ile ilgili birçok yaklaşım vardır. Akaike kriteri (AIC) ve Schwarz Bayesyen kriteri (SBC veya BIC) kullanımı kolay ve etkin olduğu için literatürde en çok kullanılan kriterlerdir (Eşitlik 23, 25).

$$n \ln(HKO) + 2d \quad (\text{AIC- Akaike Bilgi Kriteri}) \quad (23)$$

Burada HKO hata karaler ortalamasını (Eşitlik 24), n örneklemdeki birim sayısını, d ise modeldeki tahmin edilen parametre sayısını göstermektedir.

$$HKO = \frac{1}{n} \sum_{t=1}^n (y_t - \hat{y}_t)^2 \quad (24)$$

Eşitlik 24'te n örneklemdeki veri sayısını, \hat{y}_t ise ilk adımda elde edilen öngörüü göstermektedir.

$$n \ln(HKO) + k \ln(n) \quad (\text{SBC veya BIC, Schwarz bayesyen bilgi kriteri}) \quad (25)$$

BIC yeni geliştirilen metotların büyüklüğünü ve elemanlar arasındaki korelasyonu inceler. Ayrıca büyük modeller için cezalandırmalar da içerir (Moore et al., 2007; <http://www.technion.ac.il/docs/sas/ets/chap30/sect19.htm>, 2008).

AIC ve BIC kriterleri dışında model seçiminde aşağıdaki kriterler kullanılmaktadır (<http://www.technion.ac.il/docs/sas/ets/chap30/sect19.htm>, 2008).

Hata Kareler Toplamı (HKT, SSE)

$$HKT = \sum_{t=1}^n (y_t - \hat{y}_t)^2 \quad (26)$$

Hata Kareler Ortalaması (HKO, MSE) Eşitlik 24

Hata Kareler Ortalamasının Karekökü

$$HKOK = \sqrt{HKO} \quad (27)$$

Ortalama Mutlak Yüzde Hata (OMYH, MAPE)

$$OMYH = \frac{100}{n} \sum_{t=1}^n |(y_t - \hat{y}_t) / y_t| \quad (28)$$

Ortalama Mutlak Hata (OMH, MAE)

$$OMH = \frac{1}{n} \sum_{t=1}^n |(y_t - \hat{y}_t)| \quad (29)$$

Akaike'nin Son Öngörü Hatası (ASÖH, FPE)

$$ASÖH = FPE = V \left(\frac{1+d/n}{1-d/n} \right) \quad (30)$$

Eşitlik 30'da V kayıp fonksiyonu (Eşitlik 31), d tahmin edilen parametre sayısını, n ise tüm eşitliklerde örnekleme birim sayısını göstermektedir.

$$V = \det \left(\frac{1}{n} \sum_{t=1}^n \varepsilon(t, \theta_n) (\varepsilon(t, \theta_n))' \right) \quad (31)$$

Burada θ_n tahmin edilen parametreleri göstermektedir.

Çeşitli model dereceleri ile (na ve nc) çok sayıdaki modelin tahmini, yapının davranışlarını doğru olarak betimleyen modelin tanımlanma şansını arttırmakta olup bu konuda bir uzlaşma sağlanmıştır (Moore et al., 2007).

Model seçim kriterlerinden sadece birine bağlı kalarak model belirlenmemelidir. Çünkü kriterler arasında tutarsızlıklar gözlenebilir. Örneğin AIC ve BIC'ye bakıldığında bir tutarsızlıkla karşılaşılabılır, bu durumda bunların yanında üçüncü veya daha fazla seçim kriterine bakarak model seçilmelidir. Örneğin Ortalama Mutlak Yüzde Hata (OMYH, MAPE) en iyi ölçüm aracı olarak önerilmektedir (Akal, 2003).

3. Porsuk barajı su seviyesinin ARMAX modelleri ile öngörüsü

Su kaynakları sistemleri lineer olmayan ve pek çok parametreye sahip karmaşık ilişkilerden oluşur. Söz konusu ilişkileri modellemek ve bu modeller aracılığı ile ileriye yönelik öngörüler yapmak oldukça zordur. Bilgisayar teknolojisinin gelişmesiyle birlikte karmaşık problemlere daha kısa zaman içerisinde çözüm bulunabilmektedir. Söz konusu karmaşık ilişkilere uygun modelleme ARMAX modelleri ile yapılarak öngörülerde bulunulabilmektedir.

Ülkemiz, doğal su kaynakları bakımından oldukça zengin olmasına rağmen pek çok bölgede düzensiz yağış rejimi ve coğrafik koşulların etkisiyle su problemleri yaşanmaktadır. Su probleminin çözümünde barajlar oldukça önemli bir yere sahiptir. Baraj gölleri genellikle enerji üretimi, sulama, içme suyu ve taşkından korunma gibi amaçlarla kurulmaktadır. Türkiye'deki büyük şehirlerden biri olan Eskişehir'de yer alan ve baraj kapasitesinin büyüklüğü nedeniyle önemli bir yere sahip olan Porsuk barajı da tüm bu amaçlara hizmet etmek üzere Sakarya nehrinin bir kolu olan Porsuk çayı üzerinde kurulmuştur (Altın v.d., 2008).

3.1. Porsuk Barajı

Bu çalışmada Porsuk barajının incelenmesinin sebebi, Türkiye'nin büyük şehirlerinden biri olan Eskişehir'e su sağlayan dört barajın kapasitesi içerisinde %85'lik bir orana sahip olmasıdır. Diğer barajlarla birlikte bu durum Çizelge 1'de özetlenmektedir (Altın v.d., 2008).

Porsuk Barajı Eskişehir'in 40km güneybatısında Porsuk çayı üzerinde kuruludur.

Porsuk barajının bulunduğu bölge Şekil 1'de gösterilmektedir (Altın v.d., 2008).

1973 yılında barajın gövde hacminin yükseltilmesi için yapılan inşaatı kadar birçok kez taşmış ve Eskişehir'de zarara yol açmıştır. Şu

anda Porsuk Barajı'nın maksimum depolayabileceği su miktarı 525.000.000 m³ olmasına rağmen, genelde baraj kapakları açılarak su seviyesinin bu miktara ulaşmasına izin verilmemekte baraj seviyesi genellikle 400.000.000 m³ – 450.000.000 m³ civarında tutulmaktadır. Bunun nedeni ise daha sonra ansızın gelebilecek olan su için yer ayrılması ve böylelikle taşkınlara ve kontrolsüz su salınmasına engel olunmaya çalışılmasıdır. Taşkın olmasa bile kontrolsüz su salınması durumunda tarımsal alanlar ve buna bağlı olarak ürünler büyük zarara uğramaktadır.

Çizelge 1. Eskişehir'e su sağlayan barajların kapasiteleri ve toplam içindeki oranları

Baraj Adı	Kapasite (1000 m ³)	Toplam İçindeki Oran (%)
Porsuk	525000	85,50
Dodurga	19210	3,13
Kunduzlar	22800	3,71
Çatören	47000	7,66
TOPLAM	614010	100

En önemlisi Eskişehir'in şehir merkezinden geçen su kanallarındaki suyun aniden artması durumunda, Eskişehir halkı da su baskınları gibi büyük sorunlarla karşı karşıya kalmaktadır (Altın v.d., 2008).

Baraja giren suyun 19.000.000 m³'ü sulama kapaklarının altında olduğu için kullanılamamakta ve ölü hacim olarak adlandırılmaktadır. Bu

Çizelge 2. Porsuk barajının faaliyet amaçları

AMAÇ	Tesisten Beklenen Fayda (%)
Sulama	36
İçme suyu	28
Taşkın	21
Enerji	15

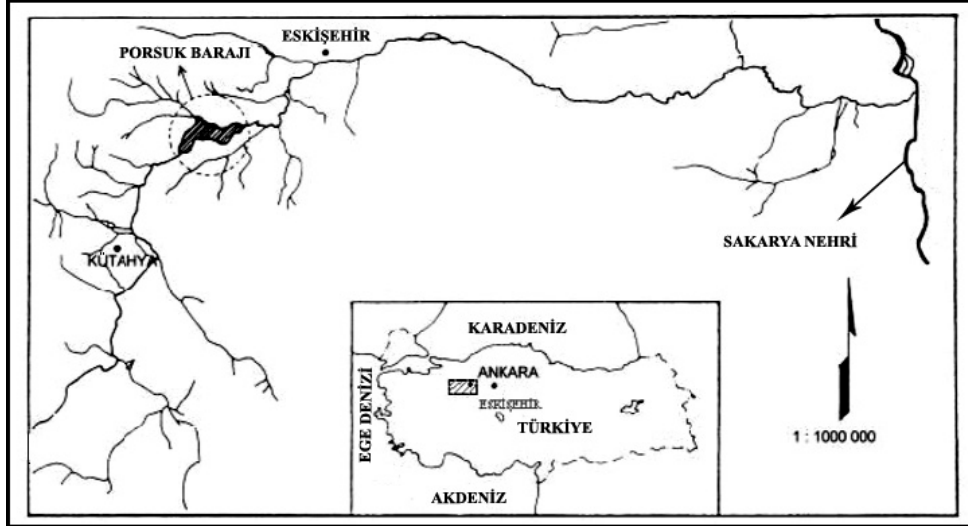
Porsuk Barajı'ndan sulama, içme suyu ve diğer (taşkın ve enerji) amacıyla kullanılması beklenen su miktarları, daha önceki yıllardaki deneyimler göz önünde bulundurularak, Devlet Su İşleri yetkilileri tarafından tahmin edilmektedir. Bu miktarlar ise Çizelge 3'te belirtilmiştir (Altın v.d., 2008).

Çizelge 3. Porsuk barajından kullanım amacına uygun karşılanması beklenen yıllık su miktarı

Kullanım Amacı	Miktar (1000 m ³ /yıl)
Sulama	125.850
İçme Suyu	100.000
Diğer	49.500

kullanılamayan su miktarı dışında kalan ve aktif hacim olarak adlandırılan su miktarı ise 446.250.000 m³'tür.

Porsuk barajından çeşitli amaçlarla yararlanılmakta ve Devlet Su İşleri yetkilileri tarafından verilen bilgilere göre bu amaçlara yönelik olarak kullanım oranları Çizelge 2'de yer almaktadır (Altın v.d., 2008).



Şekil 1. Porsuk barajı konum haritası

Kullanılabilir su kaynaklarının giderek azalması ve mevsimsel değişimler nedeniyle var olan su kaynaklarının yukarıda belirtilen amaçlar doğrultusunda planlı bir şekilde kullanılması gerekir. Bu amaç doğrultusunda gelecekte baraj su seviyesinin ne olacağı önemli bir konudur. Bu sebeple ileriye yönelik öngörü yapabilmek için ARMAX modelleri ve yapay sinir ağları kullanılmıştır. İstatistiksel teknikler içinden ARMAX modellerinin seçilmesinin sebebi, baraj su seviyesinin öngörüsünde baraja gelen su miktarı ile sarfiyat miktarı arasındaki fark değişkeninin ve m^2 'ye düşen yağmur miktarı değişkeninin dışsal değişken olarak modelde yer almasının gerekliliğinin söz konusu olmasıdır.

Belirtilen amaçlar doğrultusunda, Ocak 1973 - Kasım 2009 dönemi arasında Porsuk Barajı'nın su seviyesi (m), baraj gölüne gelen su miktarı (hm^3) ve baraj gölündeki su sarfiyatı (taşkın, sulama, buharlaşma ve diğer sarfiyatların toplamı) (hm^3) DSİ 3. Bölge Müdürlüğü'nden ve m^2 'ye düşen yağış miktarı (mm) Eskişehir Meteoroloji Bölge Müdürlüğü'nden aylık ortalama veriler halinde alınmıştır. Söz konusu değişkenlerden Porsuk

Barajı'nın su seviyesi bağımlı değişken olarak kabul edilmiş, bunun yanında baraj gölüne gelen su miktarı ile sarfiyat arasındaki fark ile m^2 'ye düşen yağış miktarı bağımsız değişken olarak ele alınmıştır.

Çalışmada, Porsuk barajında su tutulmasından itibaren günümüze kadar gelen süreç içerisinde ve gelecekte söz konusu değişkenler için elde edilen değerlerin tamamı evreni oluşturmaktadır. Bunun yanında ele alınan Ocak 1973 - Kasım 2009 dönemi içinde söz konusu değişkenler itibarıyla elde edilen değerler örneklemini oluşturmaktadır.

Belirlenen değişkenlerle çeşitli paket programlarda (SAS, SPSS, Statistica ve MATLAB) analizler yapılarak Porsuk Barajı'nın yüzey su seviyesi en uygun şekilde öngörülmeye çalışılmıştır.

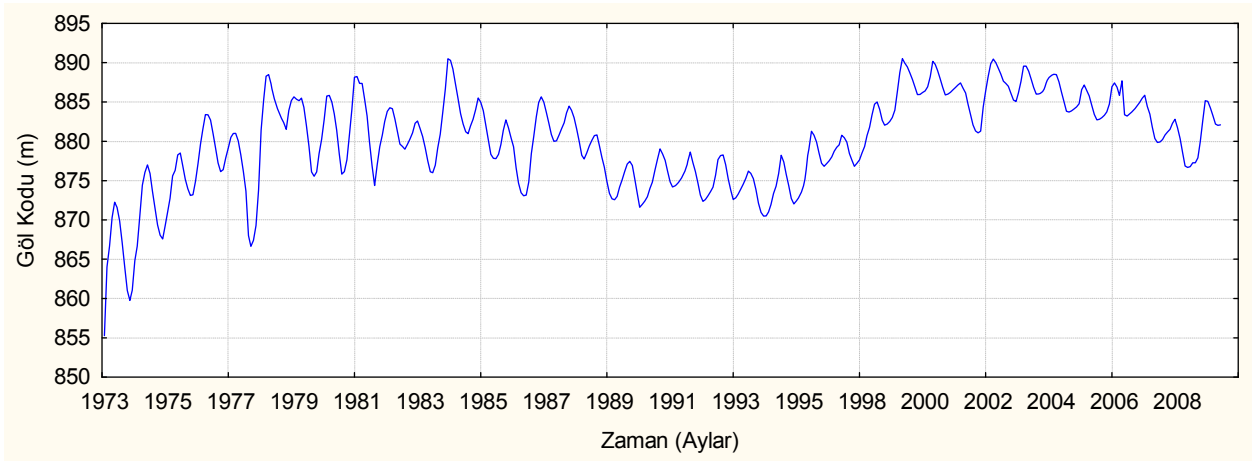
3.2. Porsuk barajı su seviyesinin ARMAX ve ARIMA modelleri ile öngörüsü

Ocak 1973-Kasım 2009 dönemleri arasındaki Porsuk Barajı'nın yüzey suyu seviyesinin değerlerinin grafiği Şekil 2'de verilmiştir.

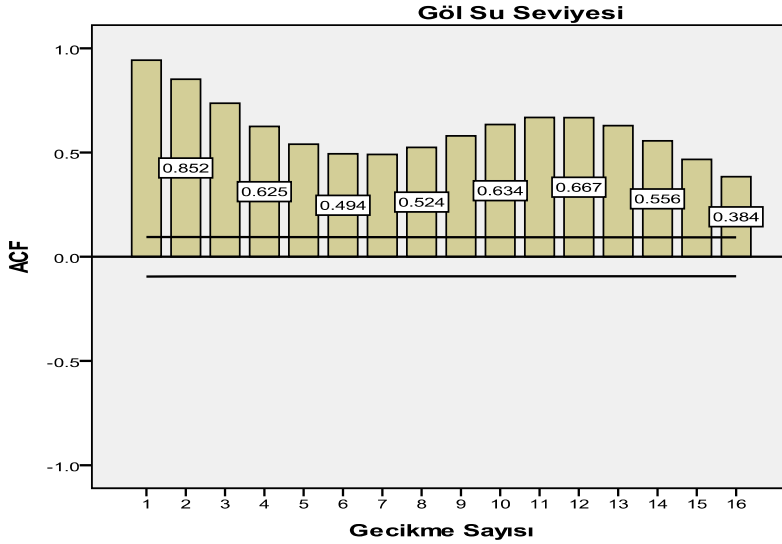
Şekil 2'den de görüleceği gibi Porsuk barajı su seviyesi serisinin mevsimsellik içerdiği ve

durağan olmadığı söylenebilir. Durağanlık, zaman içinde varyansın ve ortalamasının sabit olması ve gecikmeli iki zaman aralığında değişkenlerin kovaryansının değişkenler arasındaki gecikmeye bağlı olup zamana bağlı olmamasıdır (Gujarati, 1999). Durağanlığın olmaması zaman serisinin trend içermesi demektir. Bu da zaman serisini etkileyen faktörlerden bir tanesidir ve

arındırılması gerekir. Durağanlığı belirlemek için otokorelasyon fonksiyonundan yararlanılacağı gibi birim kök testleri de yapılabilir (Dikey-Fuler testi, ADF testi, Philips-Perron testi v.b.). Bu çalışmada otokorelasyon fonksiyonuna bakıldığında durağanlığın olup olmadığının belirlenebileceği düşünülmüştür. Otokorelasyon fonksiyonu Şekil 3'te gösterilmiştir.



Şekil 2. Porsuk barajı aylık yüzey su seviyesi (Ocak 1973-Kasım 2009)



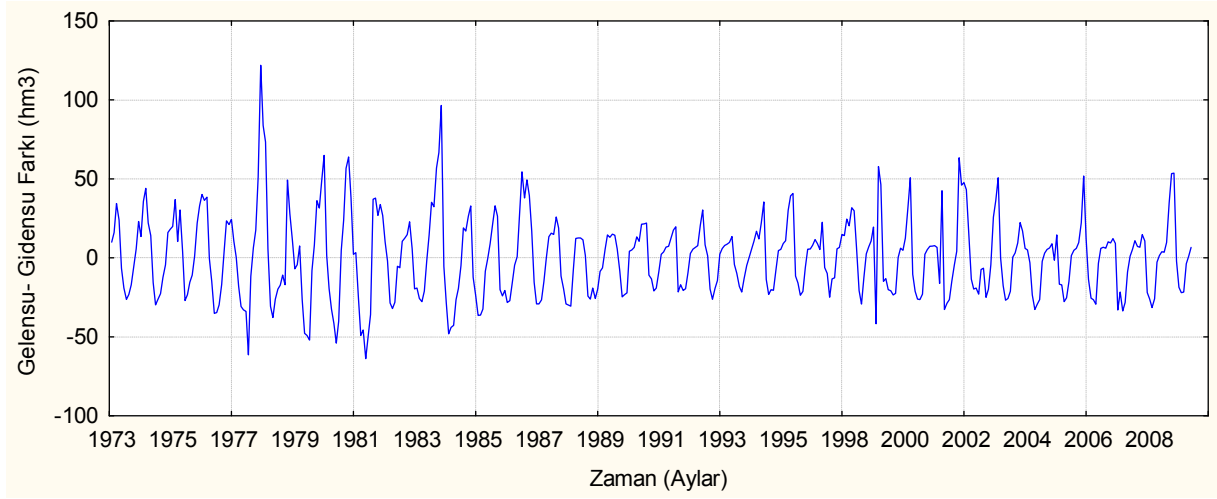
Şekil 3. Porsuk barajı su seviyesi serisinin otokorelasyon fonksiyonu

Şekil 3 incelendiğinde ilk gecikmelerde fonksiyon değerleri yüksek başlayıp yüksek gecikmelere doğru azaldığından dolayı trendin olduğunu ve durağanlığın olmadığı söylenebilir. Bunun için seride genellikle 1 derece fark alınması yeterlidir.

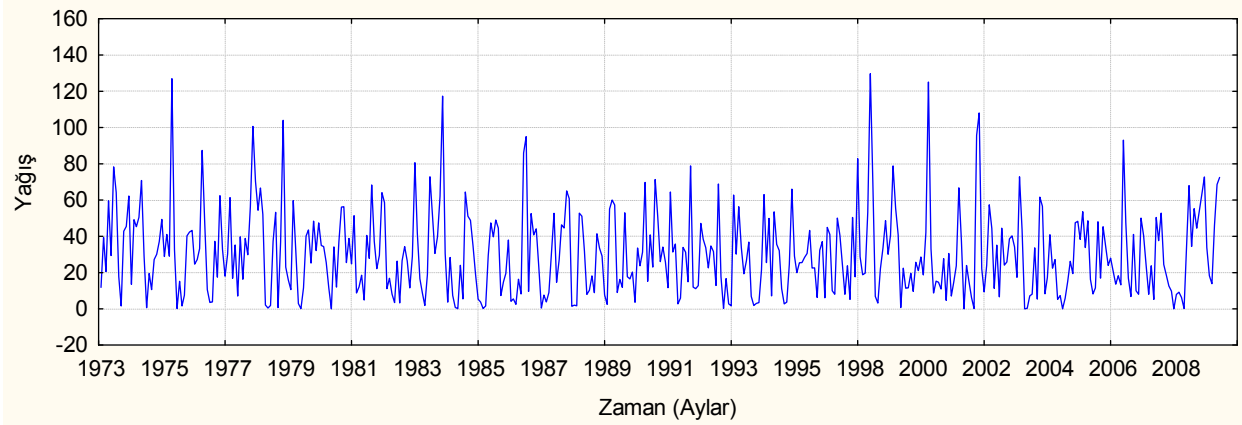
Bir zaman serisinde her yılın (üçer ayın, ayın, günün, saatin vb.) belirli dönemlerinde aynı salınım söz konusu oluyorsa mevsimsellikten söz edilebilir. Nitekim Şekil 2.'ye bakıldığında mevsimselliğin söz konusu olduğu görülebilir. 1

derece de mevsimsel fark alınması mevsimsel durağan olmama durumunu düzeltecektir.

Bağımsız değişkenler olan baraja gelen su miktarı ile barajdaki su sarfiyatı arasındaki fark değişkeni ve m^2 'ye düşen yağmur miktarı değişkeninin zamana bağlı aldıkları değerleri gösteren grafikleri çizildiğinde (Şekil 4, 5) bu seriler hakkında da bir ön bilgiye sahip olunabilmektedir. Bu ön bilgiler daha sonra modelin oluşturulmasında ön bilgi olarak kullanılmaktadır.



Şekil 4. Porsuk barajı aylık gelen su ile sarfiyat arasındaki fark (hm^3) (Ocak 1973-Kasım 2009)



Şekil 5. Porsuk barajı aylık m^2 'ye düşen yağış miktarı (mm) (Ocak 1973- Kasım 2009)

Şekil 4 ve 5 incelendiğinde bağımsız değişkenler olan gelen su ile sarfiyat arasındaki fark ile m^2 'ye düşen yağış miktarı değişkenlerinin zamana bağlı serilerinin durağan olmadığı ve mevsimsellik içerdiği görülebilir. Buradan da her iki serinin de modelde durağanlığı sağlamak için bir derece farkı aynı zamanda mevsimsel olarak bir derece farkı alınacağı söylenebilir.

Daha sonra belirtilen ön bilgiler göz önüne alınarak öngörüler elde edebileceğimiz ARMAX modelinin model dereceleri belirlenmeye çalışılmıştır. Sonuç olarak birkaç model arasından ARMAX(1 1 0 [1 1 0 1 1 0]) (0 1 1 [0 1 0 0 1 0])₁₂ model, uygun model olarak bulunmuştur. Sözü edilen modelde ilk parantez içinde gösterilen (1 1 0) değerleri, istatistiksel olarak bilinen bağımlı değişken Porsuk barajı göl su seviyesi için sırasıyla p, d, q değerlerini, sonraki köşeli parantez içinde verilen [1 1 0 1 1 0] değerleri ise iki bağımsız değişken olan baraj gölüne gelen su

ile sarfiyat arasındaki fark ve m^2 'ye düşen yağış miktarı değişkenlerinin sırasıyla p, d, q değerlerini göstermektedir. Burada d fark alma operatörü olarak kullanılmaktadır. Bunun yanında ikinci parantez içinde verilen (0 1 1) değerleri bağımlı değişkenin mevsimsel parametreleri P, d, Q değerlerini göstermektedir. İkinci köşeli parantez içinde verilen [0 1 0 0 1 0] değerleri sırasıyla iki bağımsız değişkenin mevsimsel parametrelerini göstermektedir. Belirtilen modelin uygun model olduğuna karar verirken birkaç model seçim kriterine bakılmıştır. Ayrıca karşılaştırmak amacıyla ARIMA modelleri de göl kodu değişkeni için göz önüne alınmıştır. En uygun ARIMA modeli de aynı belirleme aşamaları kullanılarak ARIMA (1 1 0)(0 1 1)₁₂ modeli uygun model olarak seçilmiştir. Söz konusu model seçim kriterlerinin, karar verilen ARMAX ve ARIMA modelleri için değerleri Çizelge 4'te verilmiştir.

Çizelge 4. Seçilen ARMAX ve ARIMA modellerinin model seçim kriterleri değerleri

	ARMAX	ARIMA
Durağan R^2	0,680	.454
R^2	0,976	.968
HKOK	0,885	.926
OMYH	0,061	.068
MMYH	1,035	.795
OMH	0,534	.599
MMH	8,943	6.890
BIC	-0,161	-.125

Çizelge 4'te HKOK, hata kareler ortalamasının karekökünü, OMYH, ortalama mutlak yüzde hatayı, MMYH, maksimum mutlak yüzde hatayı, OMH, ortalama mutlak hatayı, MMH, maksimum mutlak hatayı, BIC ise Bayesyan bilgi kriterini

göstermektedir. Modelin uygunluğu belirlendikten sonra aynı modelin parametre tahminleri yapılabilir. Çizelge 5 ve Çizelge 6'da sırasıyla seçilen ARIMA ve ARMAX modellerinin parametre tahmin değerleri bulunmaktadır.

Çizelge 5. Seçilen ARIMA modelinin parametre tahmin değerleri

					Tahmin	Standart Hata	t	p
Gol_Kodu-Model_1	Gol_Kodu	Dönüştürme yok	AR	Lag 1	.457	.041	11.161	.000
			Difference		1			
			Seasonal Difference		1			
			MA, Seasonal	Lag 1	.839	.032	26.339	.000

Çizelge 6. Seçilen ARMAX modelinin parametre tahmin değerleri

					Tahmin	Standart Hata	t	p
Gol_Kodu-Model_1	Gol_Kodu	Dönüştürme yok	AR	Lag 1	.595	.039	15,397	.000
			Difference		1			
			Seasonal Difference		1			
			MA, Seasonal	Lag 1	.804	.039	20,685	.000
Gel_Gid	Dönüştürme yok	Dönüştürme yok	Numerator	Lag 0	-.019	.002	-9,933	.000
			Difference		1			
			Seasonal Difference		1			
			Numerator	Lag 0	-.005	.001	-4,784	.000
Yagis	Dönüştürme yok	Dönüştürme yok	Difference		1			
			Seasonal Difference		1			

ARMAX ve ARIMA modellerinin katsayıları belirlendikten sonra ileriye dönük öngörüler yapılabilir. ARMAX ve ARIMA modelinden elde edilen öngörü değerleri Çizelge 7’de verilmiştir.

Çizelge 7. Seçilen ARMAX modelinden elde edilen öngörü değerleri

Aylar(2010)	ARMAX Öngörü Değerleri	ARIMA Öngörü Değerleri
Aralık	883,426	882.209
Ocak	884,234	882.186
Şubat	885,456	882.348
Mart	886,487	882.896
Nisan	887,325	883.591
Mayıs	886,698	884.226
Haziran	886,180	884.268
Temmuz	885,391	884.148
Ağustos	883,656	883.982
Eylül	883,725	883.830
Ekim	883,799	883.801
Kasım	883,912	883.822

4. Sonuç ve öneriler

İnsan, gelecekle ilgili kaygılarını göz önünde bulundurarak yaşamını sürdürmeye çalışmaktadır. Bu kaygılarını giderebilmek için, günlük yaşamında birçok kez ileriye dönük öngörüler yapmaktadır. İleriye dönük öngörülerin daha tutarlı dayanaklarının olması için istatistik kullanılmaktadır. İstatistiksel analizler yardımıyla önceki verilerden hareketle öngörü yapılarak geleceğe ışık tutulmaya çalışılmaktadır. Öngörülerin niteliğinin artması yaşamı da kolaylaştırmaktadır.

Çalışmanın başında öncelikle diğer tek değişkenli öngörü modelleri (ARIMA modelleri, üssel düzeltme yöntemi) ve regresyon yöntemi ile öngörüler yapılmıştır. Ancak daha sonra yapılan ARMAX modelleri ile öngörü performansının arttığı görülmüştür. Çalışmanın uygulama kısmında ARIMA ve ARMAX modellerinden uygun olanlarının sonuçları elde edilmiştir. Daha önce yapılan çalışmalarda ve bu çalışmada görülmüştür ki ARMAX modeli diğer öngörü modellerine göre daha iyi sonuçlar vermiştir.

Tüm istatistiksel tekniklerde olduğu gibi ARMAX modellerinde de veri sayısının yeterliliği

önemlidir. Veri sayısı ne kadar çok olursa modelin doğruluğu o kadar iyi olduğu görülmektedir.

Çalışmadan elde edilen Porsuk barajının su seviyesinin öngörüsü, ileride su sıkıntısı yaşanıp yaşanmayacağına veya su taşkınlarının olup olmayacağına bir göstergesi olabilecektir. Yapılan öngörüler, Porsuk barajı su seviyesinin bir önceki gerçekleşen değerlerine göre düşüşlerin olduğunu göstermiştir. Bir yıllık öngörü ile su sıkıntılarının yaşanacağı söylenemeyebilir ancak bir ön bilgi olabilecektir. Bunun yanında eğer yeterli derecede veriye sahip olunursa, her yılın aynı ayı bir zaman serisi olarak ayrı ayrı incelenebilir. Bu da, ileriye dönük birkaç yılın aynı aylar için öngörüsünün elde edilmesini sağlayacaktır. Söz konusu öngörülerle su sıkıntısının yaşanıp yaşanmayacağına veya su taşmalarının gerçekleşip gerçekleşmeyeceğinin cevabı verilebilir. Böylelikle konunun ilgilileri önceden uyarılabilecektir.

Porsuk barajının su seviyesinin öngörülerinin elde edilmesinde kullanılan yöntemlerden hangisinin daha iyi olduğunu belirlemek için model seçim kriteri olarak Ortalama Mutlak Yüzde Hata (OMYH) kullanılmıştır. Bunun yanında birçok seçim kriteri mevcuttur. Diğer seçim kriterleri de model seçiminde etkin bir şekilde kullanılabilir.

Not:

Bu çalışma Özer Özaydın'ın "Porsuk Barajı Su Seviyesinin Öngörülerinin Elde Edilmesinde Yapay Sinir Ağları ve ARMAX Modellerinin Karşılaştırmalı İncelenmesi" başlıklı doktora tezinden üretilmiştir.

Kaynaklar

- Akal, M., 2002, Accuracy Comparison of Forecasting Techniques with Variables on Exchange Rate Series: Turkish Liras Versus United States Dolar, Sakarya University Press House, Adapazarı.
- Akal, M., 2003, Öngörü Tekniklerinin Doğruluk Kıyaslaması: Basit Ekonometrik, ARMA ve ARMAX Teknikleri, Uludağ Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dergisi, Cilt XXII, Sayı 1, 233-269.
- Akay, B., 2004, Sistem Tanımlama Yöntemlerinin Karşılaştırmalı Kesikli Reaktöre Uygulanması,

- F.Ü. Fen ve Mühendislik Bilimleri Dergisi, 16 (2), 273-280.
- Altın, A., Çemrek, F. ve Özaydın, Ö., 2008, Statistical Modelling of Water Quantity in the Porsuk Dam, Ekoloji, 17, 67, 52-59
- Beltran, N. D. R., 1993, Prediction of Water Use in Puerto Rico, Phaze I: Mayagüez, Project No: 14-08-0001-G-2043, University of Puerto Rico.
- Erdoğan, H., Güral, E., Ata, E. ve Akpınar, B., 2005, Dinamik Sistemlerin Tanımlanması, Harita ve Kadastro Mühendisleri Odası, 10. Türkiye Harita Bilimsel ve Teknik Kurultayı, 1-9.
- Gujarati, D.N., 1999, Temel Ekonometri, (Çev. Ü. Şenesen, G.G. Şenesen), Literatür Yayıncılık, İstanbul, 850 s.
- Koulocheris, D. and Dertimanis, V., 2008, Evolutionary Parametric Identification of Dynamic Systems, In Iba H. (ed.), Frontiers in Evolutionary Robotics, I-Tech Education and Publishing, Vienna, Austria.
- Köksal, E. ve Öztürk, F., 2004, ARMAX Modellerinde Parametre Tahmini ve Kontrol, 2003 İstatistik Araştırma Sempozyumu, Özel Sayı, 73-87.
- Ljung, L., 1999, System Identification Theory for the User Second Edition, Prentice Hall PTR, New Jersey, USA, 672 p.
- Ljung, L., 2008a, System Identification Toolbox™ 7 Getting Started Guide, The MathWorks Inc., 3 Apple Hill Drive Natick, MA, 221 p.
- Ljung, L., 2008b, System Identification Toolbox™ 7 User's Guide, The MathWorks Inc., 3 Apple Hill Drive Natick, MA, 531 p.
- Makridakis, S. and Hibon, M., 1997, ARMA models and the Box-Jenkins Methodolgy, Journal of Forecasting, Volume 16, Issue 3, 147-163.
- Moore, S. M., Lai, J. C. S. and Shankar, K., 2007, ARMAX Modal Parameter Identification in the Presence of Unmeasured Excitation—I: Theoretical Background, Mechanical Systems and Signal Processing, Volume 21, Issue 4, 1601-1615.
- SAS web sitesi, 2008, <http://www.technion.ac.il/docs/sas/ets/chap30/sect19.htm>.