

## VERİ MADENCİLİĞİ SÜRECİ KULLANILARAK ALTINAPA BARAJI'NIN REZERVUAR İŞLETME MODELİ

Özlem TERZİ \*, Onur ÖZCANOĞLU, Tahsin BAYKAL

Geliş Tarihi/ Received: 30.05.2017, Kabul tarihi/Accepted: 14.08.2017

### Özet

Rezervuarlardaki su hacminin kontrolü, suyun biriktirilmesinin ve dağıtılmasının doğru zamanda yapılmasıyla olur. Gerekli önlemlerin alınmaması durumunda can ve mal kayıpları da olabilmektedir. Rezervuar işletme çalışmasının yapılması hem su temini hem de olası zararları önlemek açısından önemlidir. Rezervuarların etkili bir şekilde işletilmesi için su hacmini doğru bir şekilde saptamak gereklidir. Çalışmada, Konya il sınırları içerisinde bulunan Meram Çayı üzerindeki Altınapa Baraj Gölü'nün rezervuar işletme çalışması için veri madenciliği süreci kullanılmıştır. Veri madenciliği süreci ile modeller geliştirilirken iki farklı veri seti kullanılmıştır. Bu veri setlerinin ilkinde, göle gelen su miktarı, toplam sarfiyat ve yağış, ikincisinde ise göle gelen su miktarı, buharlaşma ve yağış parametreleri mevcuttur. Bu parametrelerle farklı girdi kombinasyonları kullanılarak hazne hacmini tahmin etmek için çeşitli modeller geliştirilmiştir. Geliştirilen model sonuçları incelendiğinde, rezervuar işletme çalışmalarında veri madenciliği sürecinin kullanılabilir olduğu görülmüştür.

**Anahtar Kelimeler:** Altınapa Barajı, veri madenciliği, rezervuar işletme modeli, Meram Çayı, Konya

## RESERVOIR MANAGEMENT MODEL OF ALTINAPA DAM USING DATA MINING PROCESS

### Abstract

The control of the water volume in the reservoirs is done by accumulation and distribution of the water in appropriate time. If the necessary precautions cannot be taken, there can be life and property losses. Reservoir operation is important for both and prevention of possible damage. For the efficient operation of reservoirs, it is necessary to determine the water volume accurately. In this study, the data mining process has been used for the reservoir operation of the Altınapa Dam Lake located on the Meram Stream within the boundary of Konya Province. Two different data sets have been used in developing models with data mining process. The first data set consist of inflow, total discharge and precipitation, whereas the second data set consist of inflow, evaporation and precipitation parameters. Using different input combinations of these parameters, various models have been developed. When the results of the models are examined, it is seen that data mining process can be used for reservoir operation studies.

**Key Words:** Altınapa Dam, data mining, reservoir operation model, Meram Stream, Konya

\*Süleyman Demirel Üniversitesi, Teknoloji Fakültesi, İnşaat Mühendisliği, Isparta  
E-posta:ozlemterzi@sdu.edu.tr.

## 1. Giriş

Son yıllarda nüfus artışındaki hızlanmanın beraberinde getirdiği artan su ihtiyacı, günümüzde su kaynakları mühendisliğinin çözüm aradığı temel sorunlardan birini oluşturmaktadır. Sanayi ve tarımsal faaliyetlerin artması ile aşırı su kullanımından ve artan su talebini karşılayacak nitelikte su kaynaklarının yetersizliğinden dolayı su kaynakları yönetiminin önemi bir kat daha artmaktadır [1].

Dünyada gün geçtikçe su kaynakları öncelikler arasına girmektedir. Nüfus artışı, su kaynaklarının kirletilmesi ve plansız kullanım, iklim değişiklikleri etkisi ile sınırlı miktarda olan su kaynaklarının kullanılabilirliği ve sürdürülebilirliğinde azalma göstermiştir.

Su sıkıntısı, sadece tarım ve enerji üretimi açısından değil, sulama, içme suyu, hidrolojik sistemler ve etkinlikleri içeren su kaynakları yönetimi açısından da kritik bir noktaya ulaşmıştır [2]. Ayrıca su biriktirilmesinde kullanılan rezervuarların da güvenliğini ve iş görürlüğünü olumsuz yönde etkilemektedir. Uygunsuz işletme hesapları, düzensiz su salımları, kaçaklar, izinsiz çıkışlar, sel sonucu gövde önüne alüvyon birikmesi, ağır yağışlar sonucu güvenlik sıkıntıları, yağışsız dönemler gibi nedenlerden dolayı rezervuarların kullanımını kısıtlı kalmaktadır.

Rezervuarlar, hem mekân hem de zaman içinde suyun kararsız dağılımının değiştirilmesinde en önemli ve etkili su depolama tesisleridir. Sadece içme suyu, enerji ve sulama suyu sağlamak için değil, aynı zamanda sel veya kuraklığı azaltmak için aşırı su girişlerini düzenlemek için de kullanılırlar. Mevcut suyun en iyi şekilde kullanımı için, rezervuarların optimal işletimi çok önemlidir. Rezervuar işletimi zaman içinde suyun biriktirilmesi ve bırakılmasını belirleyen bir takım kararlar gerektirir. Hazneye gelen suyun gelecek için biriktirilmesi ve rezervuara gelen fazla sel sularının rezervuar güvenliğini olumsuz yönde etkilememesi için işletim çok önemlidir. Ayrıca sınırlı su kaynakları giderek artan bir nüfus ile birlikte su ihtiyacı problemini arttırdığından su kaynaklarının yönetimini dikkatli bir şekilde incelemeyi gerektirir. Su kaynaklarının optimal olarak planlanması ve işletilmesi, ülke ekonomisi, hayat standardı, ulaşım ve çevre kirliliği gibi etkenleri içerisinde barındıran karmaşık bir yapıya ve sürece dayanmaktadır. Buradaki karmaşık yapının içerisine, amaçlar ve öngörülen kriterler de yerleştirildikten sonra, su kaynakları sisteminden maksimum faydayı elde etmek ilk hedef olarak karşımıza çıkmaktadır. Bu kapsamda gözetilecek amaçlar enerji üretimi (güvenilir enerji ve toplam enerji), sulama, içme suyu ve üretim gibi alanlarda en yüksek faydayı elde edebilmekle birlikte, taşkın zararları, kuraklık ve kirlilik gibi ekolojik dengenin bozulmasının önüne geçerek en düşük zararlarla sistemden en yüksek faydayı elde edebilmek olmalıdır. Böyle bir durumda, mevcut rezervuarların uygun kullanımı, yeni rezervuarların inşasından daha ekonomik olabilir. Rezervuara giren akımların rastgele karakterli olmasından dolayı gelecekteki miktarını öngörmek ve haznenin işletme çalışmasını doğru bir şekilde yapmak genel problemlerden biridir [3]. Bu sebeple günümüzde pek çok çalışma yapılmıştır. Russell ve Campbell (1996), rezervuar işletme kurallarını belirlemek için bulanık mantık yöntemini kullanmışlardır. Bulanık mantık yönteminin tek başına daha geleneksel optimizasyon tekniklerine alternatif olamayacağını fakat bunların tamamlayıcı olabileceğini belirtmişlerdir [4]. Shrestha vd. (1996), çok amaçlı bir rezervuarın işletme kurallarını elde etmek için bulanık kural tabanlı modelleme çalışması yapmışlardır. Model geliştirirken, rezervuar depolama düzeyi, giren akım ve su talebi parametrelerini girdi olarak, rezervuardan bırakılan suyu çıktı parametresi olarak kullanmışlardır. Geliştirilen modelin karmaşık rezervuar sistemlerine kolaylıkla uygulanabileceğini göstermişlerdir [5]. Mujumdar ve Ramesh (1997), sulama amaçlı gerçek zamanlı rezervuar işletme modeli geliştirmişlerdir.

Gerçek zamanlı adaptif modelin kritik düşük akış yıllarında daha iyi performans sağladığını belirtmişlerdir [6]. Panigrahi ve Mujumdar (2000), Hindistan Karnataka'daki Malaprabha sulama rezervuarı için rezervuar işletmesini bulanık mantık yöntemiyle modellemişlerdir. Modellemede, rezervuar depolaması, giren akım ve su talebi girdi olarak, salınan su çıktı olarak kullanılmıştır. Rezervuar işletme problemlerinde bulanık mantık yönteminin tek amaçlı yapılan rezervuar sistemleri için sınırlı kalabileceğini belirtmişlerdir [7]. Neelakantan ve Pundarikanthan (2000), rezervuar işletme problemini çözmek için sinir ağı tabanlı simülasyon-optimizasyon modeli kullanmışlardır. Sinir ağı tabanlı simülasyon-optimizasyon modelinin performansını, geleneksel modelin performansı ile kıyasladıklarında tatmin edici sonuçlar verdiğini bulmuşlardır [8]. Kumar vd. (2001), Hindistan'ın Orissa eyaletindeki Mahanadi nehri üzerindeki Hirakud rezervuarına ait optimum işletme modeli geliştirmek için bulanık mantık yaklaşımını kullanmışlardır. Optimum işletme politikasını çok amaçlı bulanık lineer programlama kullanarak maksimum doygunluk düzeyi için belirlemişlerdir [9]. Loaiciga (2002), göl hidrolojisi değişkeni ile rezervuar tasarımı ve işletimi çalışması yapmışlardır. Çalışmada optimum rezervuar kapasitesi ve su verimi üzerinde göl buharlaşması ve yağış etkilerini incelemişlerdir. Göl hidrolojisinin optimum rezervuar kapasitesi ve ortalama yıllık bırakılan su tahmini üzerinde önemli bir etkisinin olduğunu göstermişlerdir. Ayrıca göl hidrolojisi hakkında bilgi eksikliğinin olması durumunda daha büyük hata yapılacağını belirtmişlerdir [10]. Tilmant vd. (2002), bulanık ve bulanık olmayan optimum rezervuar işletme politikalarını kıyaslamışlardır. Her ikisinin de sistem performansının ölçümlerine benzer sonuçlar verdiğini belirtmişlerdir [11]. Chang ve Chang (2004) Tayvan'daki Taban deresi üzerine kurulu Shihmen rezervuarının sel dönemlerinde su seviyesini tahmin etmek için yapay sinir ağı tabanlı bulanık çıkarım esaslı melez sistem oluşturmuşlardır. Sonuç olarak oluşturulan modelin su seviyesi tahmininde başarıyla kullanılabilirliğini göstermişlerdir [12]. Keskin vd. (2009) veri madenciliği sürecini kullanarak Eğirdir, Kovada ve Karacaören Baraj Gölleri için ayrı ayrı ve bu üç gölü kapsayan entegre tava buharlaşma modelleri geliştirmişlerdir. Ölçülmüş tava buharlaşma değerleri modellerin sonuçları ile kıyasladıklarında, REP tree modelinin sonuçlarının diğer modellerden daha uygun sonuçlar verdiğini belirtmişlerdir [13]. Terzi vd. (2011) veri madenciliği süreci ile güneş ışınımı modelleri geliştirmişlerdir. Model sonuçlarını ölçüm değerleri ile kıyasladıklarında çok tabakalı algılayıcı algoritması ile kurulan üç girdili (P-Ta-Rh) -modelin en uygun sonucu verdiğini görmüşlerdir [14]. Terzi (2012) Isparta'nın yağış tahmini için veri madenciliği süreci ile modeller geliştirmiştir. Modellerde girdi olarak Isparta'nın Senirkent, Uluborlu, Eğirdir ve Yalvaç ilçelerindeki istasyonların yağış değerlerini kullanmıştır. Geliştirdiği modeller içinde en uygun modeli 0,81 determinasyon katsayısı değeri ile çoklu lineer regresyon algoritmasında elde edildiğini ifade etmiştir [15].

Bu çalışmada, veri madenciliği süreci kullanılarak Konya ilinin içme suyunun bir kısmını karşılayan ve Meram Çayı üzerinde kurulan Altınapa Barajı'na ait rezervuar işletme modelleri geliştirilmiştir. Veri madenciliği sürecinin en önemli avantajı, veriler arasında tanımlanamayan ilişkileri tespit etme, eksik ya da hatalı verileri tamamlama ve bunlara ait tahmin modelleri geliştirme yeteneğine sahip olmasıdır. Altınapa Barajı için veri madenciliği sürecinde kullanılan algoritmalar denenerek çeşitli rezervuar işletme modelleri geliştirilmiş ve aralarında kıyaslama yapılarak olayı temsil eden en uygun rezervuar işletme modeli belirlenmiştir.

## 2. Materyal ve Yöntem

### 2.1. Çalışma Bölgesi ve Veriler

Altınapa Barajı, Konya ilinde, Meram Çayı üzerinde, sulama ve içme suyu temini amacı ile 1963 - 1967 yılları arasında inşa edilmiş bir barajdır (Şekil 1). Altınapa Barajı, Devlet Su İşleri IV. Bölge Müdürlüğü'ne bağlıdır. Baraj hizmete 1968 yılında açılmıştır. Konya Kapalı Havzası içinde yer alıp, Meram Çayının beslediği bir barajdır. Kaya gövde dolgu tipi olan barajın gövde hacmi 325.000 m<sup>3</sup>, akarsu yatağından yüksekliği 32 m, normal su kotunda göl hacmi 15 hm<sup>3</sup>, normal su kotunda gölalanı 2,20 km<sup>2</sup>'dir. 1.400 hektarlık bir alana sulama hizmeti vermekte, yılda 38 hm<sup>3</sup> içme-kullanma suyu temini sağlamaktadır. Sahip olduğu 598 km<sup>2</sup> yağış alanına ortalama 364 mm/yıl yağış düşmektedir.

Bu çalışmada, 1975-2008 yılları arasında ölçülen göle gelen su miktarı, toplam sarfiyat (enerji, sulama, taşkın ve diğer kullanımlar), yağış ve buharlaşma oluşturulan modelde girdi parametresi olarak hazne hacmi ise çıktı parametresi olarak kullanılmıştır. Verilerin % 80'i modeli eğitmek için, kalan % 20'si ise modelin doğruluğunu test etmek için kullanılmıştır.

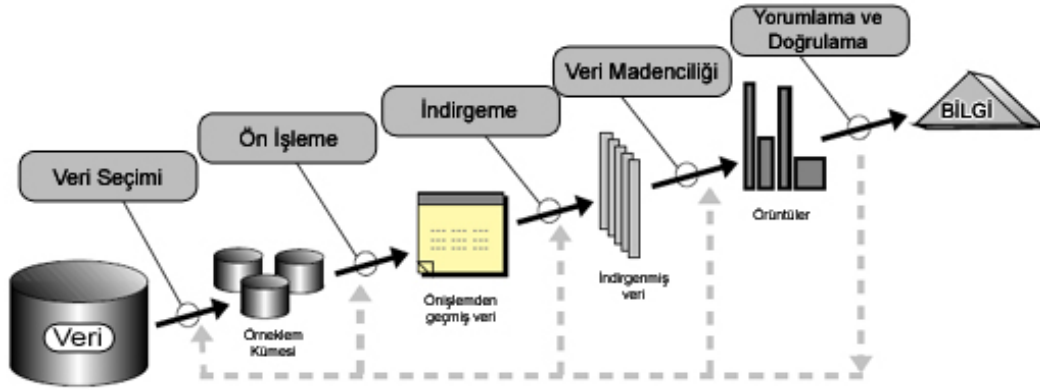


Şekil 1. Çalışma Alanı [16]

### 2.2. Veri Madenciliği

Büyük veri tabanlarından gizli kalmış örüntüleri çıkarma sürecine veri madenciliği adı verilmektedir. Geleneksel yöntemler kullanılarak çözülmesi çok zaman olan problemlere veri madenciliği süreci kullanılarak daha hızlı bir şekilde çözüm bulunabilir [17]. Veri madenciliğinin ana amacı elimizde bulunan veriden gizli kalmış örüntüleri çıkarmak, elimizdeki verinin değerini arttırmak ve veriyi bilgiye dönüştürmektir [18]. Veri madenciliği süreci, geniş veri yığınları içerisinde, yararlı olma potansiyeline sahip, aralarında beklenmedik/bilinmedik ilişkilerin olduğu verilerin keşfedilerek, veri sahibi için hem anlaşılır hem de kullanılabilir bir biçime getirilmesine yönelik geliştirilmiş yöntemler topluluğu olarak açıklanmıştır [19].

Veri Madenciliği bir yöntem değildir bir süreçtir. Bu süreçte ana unsur süreci gerçekleştiren uygulamadır. Süreçte bulunan adımlar doğru olarak yerine getirilmediği sürece istenilen sonuca ulaşılması mümkün değildir. Şekil 2’de veri madenciliği süreci gösterilmiş ve aşağıda adımları açıklanmıştır.



Şekil 2. Bilgi keşfi sürecinde veri madenciliğinin yeri [20]

### 2.1. Veri toplama

Veri madenciliğinin birinci adımı veri toplamadır. Birçok kaynaktan gelen veriler, tek bir alanda toplanmaktadır. Toplanan verilerin analizi yapılarak uygun veriler seçilir. Veri toplama işlemi bittikten sonra, veriler test ve analiz veri seti olarak iki gruba ayrılır. Genellikle yapılan uygulamalarda verilerin %80'i analiz %20'si ise test verisi olarak ayrılır.

### 2.2. Veri temizleme ve dönüştürme

Veri temizleme işleminin amacı, veriler içindeki uygun olmayan veya hatalı girilmiş verileri ayıklamaktır. Bu işlemde eksik veriler geçmiş yılların ortalaması alınarak tamamlanır. Eğer eksik veri çok ise bu kaydın silinmesi gerekir. Veri dönüşümü ise, elimizdeki kaynak veriyi farklı formatlara veya değerlere dönüştürmektir. Örneğin; veri tabanındaki boolean bir alan integer bir tipe dönüştürülebilir. Bunun sebebi ise kullanılan bazı veri madenciliği algoritmalarının integer veri tipiyle boolean veri tipine göre daha başarılı sonuçlar üretmesidir [20].

### 2.3. Model kurma

Model kurma veri madenciliğinin çekirdeğidir. Modeli doğru bir şekilde kurabilmek için yapılacak projenin amacı çok iyi bir şekilde kavranmış olmalıdır. Her amaçla ilgili birden fazla algoritma mevcuttur. Bu durumda eldeki veriler üzerinde uygun algoritmaların hepsi çalıştırılır ve en doğru sonucu veren algoritma kullanılır.

### 2.4. Model değerlendirme

Veriden elde edilen analiz sonuçlarının incelenmesi işlemidir. Eldeki veriler üzerinde uygun algoritmalar çalıştırıldıktan sonra en doğru sonucu hangisinin verdiğini bulmak için çeşitli

yöntemler mevcuttur. Örneğin, tahmine yönelik sayısal veriler varsa ve kullanılan modelin doğruluğu test edilmek isteniyorsa ortalama mutlak hata, ortalama karesel hata gibi yöntemler kullanılabilir.

## 2.5. Raporlama

Raporlama veri madenciliği bulgularını göstermek için önemli bir dağıtım kanalıdır. Birçok veri madenciliği aracı elde edilen modelden kullanıcıların daha önceden tanımladığı raporları göstermek için gerekli araçlara sahiptir.

## 2.6. Tahmin (scoring)

Birçok veri madenciliği projesinde, örüntüleri bulmak çalışmanın yarısıdır. Esas amaç, tahmin için modeli kullanmaktır. Tahmin veri madenciliği terminolojisinde scoring olarak da adlandırılır. Tahmin yapabilmek için eğitilen model ve yeni durumları içeren veri setinin olması gerekir [20]. Böylece, eğitilen model kullanılarak yeni durumlar için tahminde bulunulabilir.

## 2.7. Uygulama entegrasyonu

Bu aşamada kurulan veri madenciliği modeli gerçek zamanlı olarak çalıştırmak üzere geliştirilen uygulama içerisine gömülür. Kurulan ve geçerliliği kabul edilen model doğrudan veya alt parça olarak kullanılabilir.

## 2.8. Model yönetimi

Her bir veri madenciliği modeli bir yaşam döngüsüne sahiptir. Bazı uygulamalarda işler, özellikler durağandır ve modelin yeniden eğitilmesine gerek yoktur. Fakat birçok iş özellikleri sık sık değişir. Yeni veriler geldikçe modelin yeniden eğitilmesine gerek vardır. Yani bir model kurulduktan sonra eğer çok sık olarak veri setinde değişiklik yapılıyorsa model sık sık güncellenmelidir [20].

## 3. Araştırma ve Bulgular

Rezervuar işletme modelleri geliştirilirken 1994 yılından sonra buharlaşma verileri eksik olduğu için iki farklı veri seti oluşturulmuştur. Birinci veri setinde, 1975-2008 yılları arasında göle gelen su miktarı, yağış ve toplam sarfiyat (enerji, sulama, taşkın ve diğer kullanımlar) girdi parametreleri olarak göl hacmi ise çıktı parametresi olarak kullanılmıştır. İkinci veri setinde ise girdi parametrelerini 1975-1994 yılları arasında ölçülen göle gelen su miktarı, yağış ve buharlaşma oluşturmaktadır.

Modellerin eğitim ve test setlerine ait determinasyon katsayısı ( $R^2$ ), karekök ortalama hata (KOH) ve ortalama mutlak hata (OMH) değerleri denklem 1, 2 ve 3'deki bağıntılarla hesaplanmıştır.

$$R^2 = 1 - \frac{\sum_{i=1}^N (V_{i(\text{gerçek})} - V_{i(\text{model})})^2}{\sum_{i=1}^N (V_{i(\text{gerçek})} - V_{i(\text{ort})})^2} \quad (1)$$

$$KOH = \sqrt{\frac{1}{N} \sum_{i=1}^N (V_{i(\text{gerçek})} - V_{i(\text{model})})^2} \quad (2)$$

$$OMH = \frac{1}{N} \sum_{i=1}^N |V_{i(\text{gerçek})} - V_{i(\text{model})}| \quad (3)$$

Burada, N toplam veri sayısı,  $V_{i(\text{gerçek})}$  ve  $V_{i(\text{model})}$  sırasıyla ölçülen göl hacim değerleri ve geliştirilen modelin sonuçlarıdır.  $V_{\text{ort}}$  ise ortalama göl hacim değerini belirtmektedir.

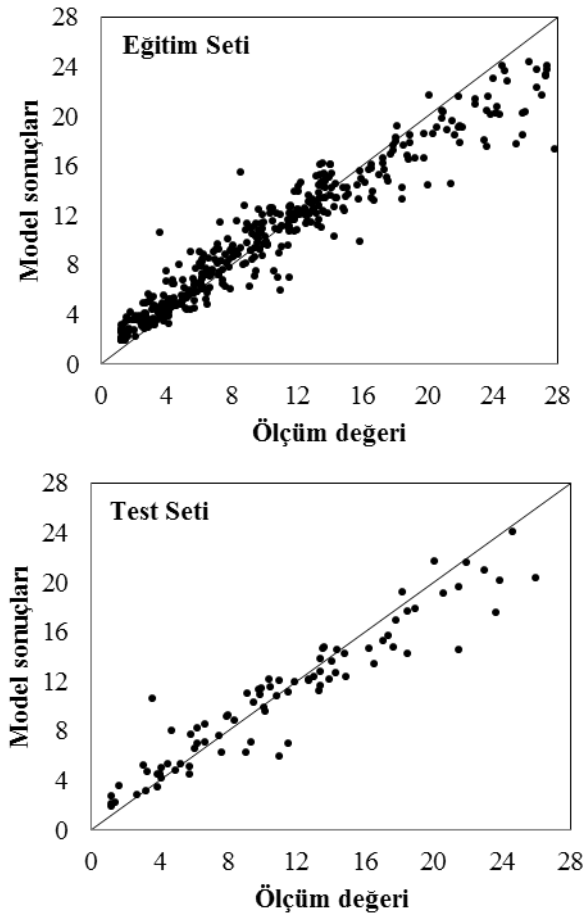
Veri madenciliği süreci ile model kurma aşamasında KStar, REPTree, M5P, DecisionTable, LinearRegression ve MultilayerPerceptron algoritmaları kullanılmıştır. Birinci veri setinde bulunan girdi parametrelerinin farklı kombinasyonları denenerek geliştirilen modeller için hesaplanan  $R^2$ , KOH ve OMH değerleri Tablo 1'de verilmiştir.

**Tablo 1.** Birinci veri seti için geliştirilen modellere ait  $R^2$ , KOH ve OMH değerleri

Modeller		KStar		RepTree		M5P		Decision Table		Linear Regression		Multilayer Perceptron	
		Eğitim	Test	Eğitim	Test	Eğitim	Test	Eğitim	Test	Eğitim	Test	Eğitim	Test
3 girdili (yağış – gelen akım – buharlaşma)	$R^2$	0,925	0,900	0,549	0,483	0,483	0,326	0,478	0,303	0,303	0,239	0,460	0,302
	KOH	2,006	2,151	4,451	4,651	4,768	5,338	4,786	5,406	5,532	5,680	5,249	5,859
	OMH	1,421	1,544	3,364	3,585	3,714	4,027	3,676	4,058	4,362	4,509	3,691	4,251
2 girdili (gelen akım - yağış)	$R^2$	0,501	0,419	0,278	0,183	0,222	0,109	0,283	0,123	0,162	0,091	0,210	0,098
	KOH	4,968	5,012	5,630	5,781	5,845	6,119	5,608	6,021	6,065	6,131	6,475	6,886
	OMH	3,838	3,940	4,281	4,585	4,534	4,722	4,261	4,657	4,771	4,774	4,682	5,111
2 girdili (gelen akım - sarfiyat)	$R^2$	0,649	0,540	0,528	0,419	0,471	0,324	0,478	0,303	0,299	0,243	0,399	0,257
	KOH	4,059	4,460	4,553	4,935	4,823	5,358	4,786	5,406	5,547	5,697	5,422	6,064
	OMH	3,114	3,347	3,475	3,679	3,746	4,024	3,676	4,058	4,397	4,536	3,948	4,448
2 girdili (sarfiyat - yağış)	$R^2$	0,590	0,514	0,424	0,328	0,401	0,292	0,366	0,190	0,303	0,239	0,380	0,220
	KOH	4,397	4,644	5,027	5,323	5,128	5,463	5,276	5,891	5,532	5,680	5,590	6,374
	OMH	3,413	3,593	3,993	4,250	4,121	4,306	4,255	4,690	4,362	4,509	4,075	4,744

Tablo 1 incelendiğinde, kullanılan yöntemler içerisinde en yüksek  $R^2$  değerleri genel olarak K-Star algoritması ile geliştirilen modellerde elde edilmiştir. Diğer algoritmalar ile geliştirilen modellerin genelde düşük determinasyon katsayıları verdiği görülmüştür. Ayrıca, girdi kombinasyonları değerlendirildiğinde ise, üç girdili modellerin diğer modellerden daha uygun sonuçlar verdiği görülmüştür. KStar algoritması ile geliştirilen üç girdili modele ait test seti için  $R^2$ , OMH ve KOH değerleri sırasıyla 0.900, 1.544 ve 2.151 olarak belirlenmiştir. Bu modele ait eğitim ve test setlerinin saçılma diyagramları Şekil 3'de verilmiştir. Şekil 3'e bakıldığında noktaların 45°lik çizgi etrafında toplandığı ve model sonuçları ile ölçüm değerlerinin uyduğu görülmüştür.





Şekil 3. Birinci veri seti için KStar modeline ait saçılma diyagramları

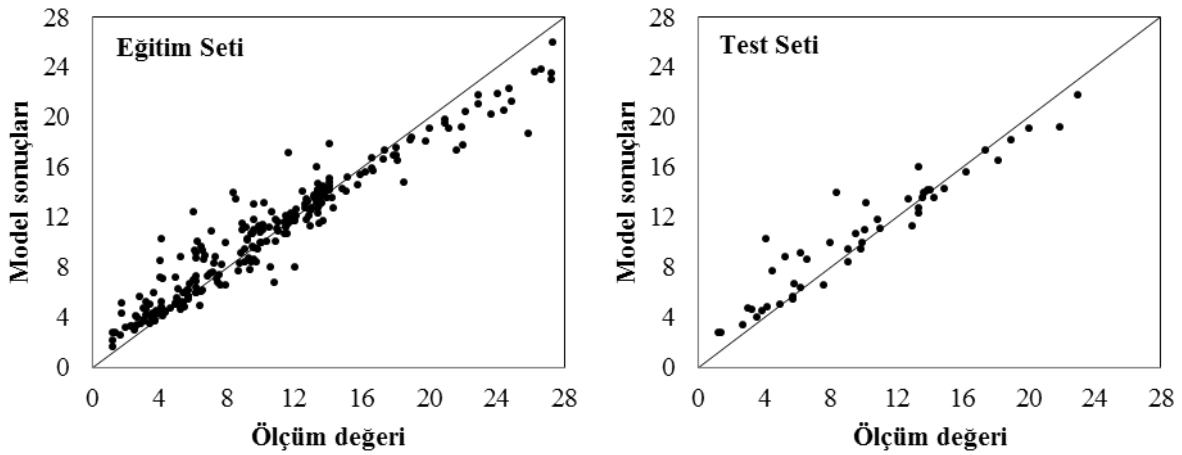
İkinci veri setinde bulunan 1975-1994 yılları arasında ölçülen göle gelen su miktarı, yağış, buharlaşma ve göl hacim verileri kullanılarak geliştirilen modeller için  $R^2$ , KOH ve OMH değerleri Tablo 2’de verilmiştir.

Tablo 2. İkinci veri seti için geliştirilen modellere ait  $R^2$ , KOH ve OMH değerleri

Modeller		KStar		RepTree		MSP		Decision Table		Linear Regression		Multilayer Perceptron	
		Eğitim	Test	Eğitim	Test	Eğitim	Test	Eğitim	Test	Eğitim	Test	Eğitim	Test
3 girdili (yağış – gelen akım – buharlaşma)	$R^2$	0,922	0,910	0,500	0,604	0,582	0,707	0,630	0,570	0,335	0,374	0,408	0,416
	KOH	1,806	1,800	4,268	3,488	3,903	3,060	3,651	3,589	4,895	4,396	4,881	4,610
	OMH	1,245	1,241	3,198	2,719	2,871	2,520	2,708	2,672	3,528	3,384	3,981	3,931
2 girdili (gelen akım - yağış)	$R^2$	0,625	0,653	0,339	0,397	0,354	0,374	0,248	0,222	0,180	0,307	0,335	0,339
	KOH	3,880	3,367	4,878	4,249	4,828	4,331	5,203	4,841	5,436	4,802	4,928	4,521
	OMH	2,839	2,739	3,689	3,349	3,649	3,552	4,062	3,953	4,181	3,911	3,899	3,787
2 girdili (gelen akım - buharlaşma)	$R^2$	0,657	0,676	0,505	0,560	0,544	0,660	0,630	0,570	0,335	0,374	0,287	0,260
	KOH	3,757	3,361	4,275	3,675	4,072	3,262	3,651	3,589	4,895	4,396	5,109	4,774
	OMH	2,815	2,721	3,222	2,958	3,063	2,681	2,708	2,672	3,528	3,384	3,995	3,894
2 girdili (buharlaşma - yağış)	$R^2$	0,639	0,623	0,499	0,554	0,425	0,479	0,360	0,365	0,153	0,161	0,019	0,012
	KOH	4,219	3,889	4,434	3,876	4,586	4,034	4,802	4,392	5,524	5,042	6,066	5,485
	OMH	3,205	3,053	3,289	3,047	3,533	3,168	3,787	3,658	4,215	4,086	4,735	4,518



Tablo 2'ye bakıldığında, kullanılan algoritmaların birinci veri setine göre daha iyi sonuçlar verdiği gözlenmiştir. Birinci veri seti ile karşılaştırıldığında toplam sarfiyat yerine buharlaşma parametresi girdi olarak modellere dahil edilmiştir. Sonuçlar incelendiğinde, birinci veri seti ile yapılan üç girdili modelde de olduğu gibi en iyi sonucu KStar algoritması vermiştir. Bu yöntemin test seti için  $R^2$  değeri 0.91, OMH değeri 1.245 ve KOH değeri ise 1.806 olarak bulunmuştur. Bu veri seti için en iyi modele ait eğitim ve test setlerinin saçılma diyagramları Şekil 4'de verilmiştir. Saçılma diyagramları, model ile ölçüm değerlerinin uyumlu olduğunu göstermektedir. Ayrıca, en düşük  $R^2$  değerleri, göle gelen su miktarı ve yağış parametrelerinin girdi olarak kullanıldığı modellerde elde edilmiştir.



Şekil 4. İkinci veri seti için KStar modeline ait saçılma diyagramları

Geliştirilen tüm modeller ele alındığında, birinci veri seti için KStar haricindeki yöntemlerin düşük sonuçlar verdiği görülmüştür. İkinci veri seti ile oluşturulan modeller ise birinci veri setine nispeten daha iyi sonuçlar vermiştir. KStar yönteminde birinci ve ikinci veri setlerine ait  $R^2$  değerleri birbirlerine çok yakın çıkmasına rağmen ikinci veri seti ile geliştirilen modelin OMH ve KOH değerleri birinci veri setine göre daha düşük elde edilmiştir.

#### 4. Sonuçlar

Meram Çayı üzerinde kurulu olan ve Konya İli'nin içme suyunun bir kısmını karşılayan Altınapa Baraj Gölü'nün rezervuar işletme çalışması için veri madenciliği süreci kullanılarak çeşitli modeller geliştirilmiştir. Modeller geliştirilirken KStar, REPTree, M5P, DecisionTable, LinearRegression ve MultilayerPerceptron algoritmaları kullanılmıştır. Bu modeller iki farklı veri seti ile iki ve üç girdili olarak geliştirilmiştir. Bu veri setlerinden ilkinde göle gelen su miktarı, toplam sarfiyat ve yağış verileri, ikinci veri setinde ise göle gelen su miktarı, buharlaşma ve yağış verileri girdi parametreleri olarak kullanılmıştır. Değerlendirme kriterlerine göre her iki veri setinde de en uygun modeller KStar algoritması ile elde edilmiştir. Dolayısıyla, rezervuar işletme çalışmalarında eksik verilerin tamamlanmasında veri madenciliği sürecinin kullanılabilir olduğu ve yeni veriler ilavesi ile modellerin geliştirilebilir olduğu görülmüştür.

## Teşekkür

Bu çalışma, 2189-YL-10 no.lu proje ile Süleyman Demirel Üniversitesi Bilimsel Araştırma Projeleri Koordinasyon Birimi tarafından desteklenmiştir.

## Kaynaklar

- [1] Meriç, B.T., 2004. Su Kaynakları Yönetimi ve Türkiye. Jeoloji Mühendisliği Dergisi, 28 (1), 27-38.
- [2] Kapluhan, E., 2013. Türkiye’de Kuraklık ve Kuraklığın Tarıma Etkisi. Marmara Coğrafya Dergisi, 27, 487-510.
- [3] Sattari, M. T., Fakher-Fard, A., Docherkhesaz, M., Öztürk, F., 2007. Yapay Sinir Ağları Yöntemi ile Savalan Sulama Rezervuarının Simülasyonu. Ankara Üniversitesi Ziraat Fakültesi Tarım Bilimleri Dergisi, 13(4), 337-345.
- [4] Russell S.O., Campbell P.F., 1996. Reservoir Design and Operation with Variable Lake Hydrology. Journal of Water Resources Planning and Management, 122, 165-170.
- [5] Sherestha P., Duckstein L., Stakhiv Z., 1996. Fuzzy Rule Based Modeling of Reservoir Operation. Journal of Water Resources Planning and Management, 122, 262-269.
- [6] Mujumdar P.P., Ramesh T.S.V., 1997. Real-Time Reservoir Operation for Irrigation. Water Resources Research, 33, 1157-1164.
- [7] Panigrahi D.P, Mujumdar P.P., 2000. Reservoir Operation Modelling with Fuzzy Logic. Water Resources Management, 33, 89-109.
- [8] Neelakantan T.R., Pundarikanthan N.V., 2000. Neural Network Based Simulation Optimization Model for Reservoir Operation. Journal of Water Resources Planning and Management, 126, 57-64.
- [9] Kumar, D.N., Prasad, D.S.V. Raju, K.S., 2001. Optimal Reservoir Operation Using Fuzzy Approach. In: Proceedings of the International Conference on Civil Engineering, ICCE.
- [10] Loaiciga H.A., 2002. Reservoir Design and Operation with Variable Lake Hydrology. Journal of Water Resources Planning and Management, 128, 399-405.
- [11] Tilmant A., Vanclooster M., Duckstein L., Persoons E. 2002. Comparison of Fuzzy and Nonfuzzy Optimal Reservoir Operating Policies. Journal of Water Resources Planning And Management, 128, 390-398.
- [12] Chang F.J., Chang Y.T., 2004. Adaptive Neuro-Fuzzy Inference System for Prediction of Water Level in Reservoir. Advances in Water Resources, 29, 1-10.
- [13] Keskin, M.E., Terzi, Ö., Küçüksille, E.U., 2009. Data Mining Process for Integrated Evaporation Model. Journal of Irrigation and Drainage Engineering, 135(1), 39-43.

[14] Terzi Ö., Küçüksille, E.U., Ergin, G., İlker, A., 2011. Veri Madenciliği Süreci Kullanılarak Güneş Işınımı Tahmini. SDU International Technologic Science, 3(2), 29-37.

[15] Terzi Ö., 2012. Veri Madenciliği Süreci Kullanılarak Yağış Tahmini. Akıllı Sistemlerde Yenilikler ve Uygulamaları Sempozyumu, Trabzon, 126-129.

[16] Google Earth, 2010.

[17] Hung, S., Yen, D.C., Wang, H., 2006. Applying Data Mining to Telecom Churn Management. Expert Systems with Applications, 31(3), 515-524.

[18] <http://www.sqlnedir.com/Members/ArticleDetail.aspx?Id=81>

[19] Ögüt, S., 2009. Veri Madenciliği Kavramı ve Gelişim Süreci. Yeditepe Üniversitesi Fen Bilimleri Enstitüsü, 12s., İstanbul.

[20] Tang, Z., MacLennan, J., 2005. Data Mining with Sql Server. Wiley, 480 pages, ISBN: 978-0-471-46261-3.