



## Porsuk çayı akım verilerinin değerlendirilmesi

**Ali YILDIRAN**

Bilecik Şeyh Edebali Üniversitesi, İnşaat Mühendisliği Bölümü, Bilecik  
[aliyildirann@gmail.com](mailto:aliyildirann@gmail.com) ORCID: 0000-0003-3130-6582

**Süheyla YEREL KANDEMİR\***

Bilecik Şeyh Edebali Üniversitesi, İnşaat Mühendisliği Bölümü, Bilecik  
[syerel@gmail.com](mailto:syerel@gmail.com) ORCID: 0000-0003-4056-5383, Tel: (228) 214 15 59

Geliş: 27.07.2018, Revizyon: 29.10.2018, Kabul Tarihi: 30.10.2018

### Öz

*Dünyamızda canlılığın temel ihtiyaçlarından bir tanesi sudur. Hayatımızda bu kadar önemli bir yer tutan su dünyamızda sınırsız değildir. Bu nedenden dolayı dünyamızdaki su kaynaklarının bilinmesi, kontrol altında tutulması ve doğru bir şekilde değerlendirilmesi büyük önem taşımaktadır. Su kaynaklarının büyük bir kısmı nehirlerden elde edilmektedir ve nehir akım değerlerinin bilinmesi günümüzde büyük önem taşımaktadır. Nehir akım değerleri birçok parametreye bağlıdır. Belirli sebeplerden dolayı bu parametrelerin ölçülememesi ya da eksikliği durumunda bizlere su miktarı ile ilgili bilgi verebilen ve sadece matematiksel olarak bir davranış modeli belirleyen kara kutu modelleri bulunmaktadır. Yapay sinir ağları günümüzde en çok tercih edilen kara kutu modelinden biridir. Bu çalışmada yapay sinir ağlarının, ileri beslemeli geri yayılım metodu ve çoklu doğrusal regresyon analizi ile gen ifade programlama tekniği kullanılarak 4 farklı senaryo belirlenmiş ve bu senaryoların hangilerinin doğruya en yakın tahminlerde bulunduğu incelenmiştir. Çalışma sonunda yapay sinir ağları modellerinin gerçek değerlere daha yakın sonuçlar verdiği görülmüştür.*

**Anahtar Kelimeler:** Yapay sinir ağları, regresyon analizi, akım, Porsuk Çayı, gen ifade programlama

\* Yazışmaların yapılacağı yazar

## Giriş

Dünyamızdaki su kaynakları sınırsız bir miktarda bulunmamaktadır. Toplam su miktarı yapılan araştırmalara göre 1,4 milyar  $\text{km}^3$ 'tür ve bunların sadece %2.5'lük bir kısmını nehir ve göllerde bulunan tatlı sular oluşturmaktadır. Ülkemiz yıllık ortalama 501 milyar  $\text{m}^3$  yağış almaktadır ve bu yağışın 274 milyar  $\text{m}^3$ 'lük kısmı toprak ve su yüzeyleri ile bitkilerden dolayı gerçekleşen buharlaşmadan dolayı atmosfere geri dönmektedir. 158 milyar  $\text{m}^3$ 'lük kısmı da akıma dönüşerek farklı büyüklüklerdeki akarsular ile denizlere ve kapalı havzalardaki göllere dökülmektedir (DSİ, 2018).

Tüm dünya yaşamı için önemli bir yere sahip olan ve insanlık için vazgeçilemez bir kaynak olan su; korunması ve kontrol altında tutulması gereken bir kaynaktır. Yapılan hidroloji çalışmalarında nehir akım verilerinin ölçülemediği, eksik ölçüldüğü ya da geleceğe dair bir fikir vermesi gerektiği durumlarda uygun modelleme tahmin yöntemlerine ihtiyaç duyulmaktadır. Debi ölçümlerinde kullanılan malzemelerin arazi koşullarına uygun seçilmemesi, ekipman eksikliği gibi durumlarda ölçüm değerlerinin elde edilmesinde birçok hidrolojik modeller türetilmiştir. Bunların içinde karakteristik birçok veriye ihtiyaç duyan parametrik modeller ve fazla miktarda veriye ihtiyaç duymadan matematiksel bir yaklaşımla tahmin modelleri oluşturan kara kutu modelleri bu eksik akım verilerinin hesaplanmasında ya da geleceğe dönük tahminlerin yapılmasında kullanılmaktadır (Küçük ve Ağrılioğlu, 2005-Gümüş vd., 2013).

Kara kutu modellerinden biri olan yapay sinir ağları; son yıllarda sıklıkla kullanılan bir modelleme ve analiz tahmin yöntemidir. Bazı analiz tahmin yöntemleri gibi çok sayıda veriye ihtiyaç duyulmadan doğruya yakın sonuçlar verebildiği, kullanım kolaylığı ve buna benzer sebeplerden ötürü su kaynakları tahmin ve analizlerinde son yıllarda sıklıkla kullanılan bir yöntem haline gelmiştir. Bu yüzden bu tahmin metodu ile geçmişten günümüze kadar çok sayıda çalışma yapılmıştır (Cıgızoğlu, 2003). Bu

konuda yapılan bazı çalışmalar aşağıda verilmiştir.

Savic vd. (1999), yağış akış modellemesi yapmak için gen ifade programlama ve yapay sinir ağları yöntemlerini kullanmışlardır. Çalışmada her iki yöntemden başarılı sonuçlar elde edilmiştir. Alp ve Cıgızoğlu (2004), ileri beslemeli geri yayılım metodu ve genelleştirilmiş regresyon sinir ağı yöntemlerini bir modele uygulamış ve tahmin metodlarının performanslarını karşılaştırmışlardır. Çalışma sonunda ileri beslemeli geri yayılım metodu genelleştirilmiş regresyon analizine göre daha iyi sonuçlar vermiştir. Demirpençe (2005), günlük akım verilerini kullanarak ileri beslemeli geri yayılım sinir ağı ile akım tahmini yapmıştır. Çalışmada 6 model oluşturularak performansları incelenmiştir. Keskin ve Taylan (2007), otoregresif model ve yapay sinir ağları modellerinin performanslarını karşılaştırdıkları bir çalışma yapmışlardır. Yapay sinir ağı modellerinin otoregresif modellere göre daha başarılı sonuçlar verdiği görülmüştür. Fay ve Ringwood (2007), dalgacık analizi ile yapay sinir ağı kullanımını birleştirerek oluşturulan dalgacık sinir ağı modelleri geliştirmişlerdir. Model başarılı tahminlerde bulunmuştur. Makkeasorn vd. (2008), iklim değişimi verilerinden yararlanarak kısa dönemli nehir akım tahmini yapmışlardır. Çalışmada gen ifade programlama ve yapay sinir ağı yöntemleri kullanılmıştır. Yurdusev vd. (2008), aylık akım ve yağış verilerini kullanarak yapay sinir ağı modelleri geliştirmiş ve bu modelleri test etmişlerdir. Çalışmalar sonucunda verilerin yeterli sayıda olduğu durumlarda yapay sinir ağlarının, yağış akış modellemesinde güvenilir sonuçlar verdiği görülmüştür. Aytek ve Kişi (2008) günlük sediment-debi ilişkisini yorumlamak amacıyla genetik programlamadan faydalanmışlardır. Çalışma sonu verileri çoklu lineer regresyon yöntemi sonuçlarıyla karşılaştırılmış ve genetik programlama ile daha verimli sonuçlar alındığı görülmüştür. Okkan ve Mollamahmutoğlu (2010), Yiğitler Çayına ait günlük akım verilerinden faydalanarak yapay sinir ağları ile modelleme yapmışlardır. Hazırlanan yapay sinir ağı modeli çoklu regresyon modeli ile

karşılaştırılmış ve yapay sinir ağı verilerinin çoklu regresyon modeline göre daha başarılı olduğu görülmüştür. Saplıoğlu ve Çimen (2010), yaptıkları çalışmada bir istasyonda mevcut ölçüm verileri yardımıyla, geçmişe dönük veriler elde edilip edilemeyeceğini irdelemiştir. Modelleme için 2000 – 2009 yıllarına ait 3170 adet günlük yağış verisi kullanılmıştır. Shiri ve Kisi (2011), genetik programlama ile sinirsel bulanık çıkarım sistemi (ANFIS) yöntemlerinden yararlanarak yeraltı su seviyesi tahminlerinde bulunmuşlardır. Her iki yönteminde yeraltı su seviyesi tahmininde başarılı sonuçlar verdiği görülmüştür. Kişi vd. (2013), Orta Sakarya havzasında yaptıkları çalışmada 1987-1991 yıllarına ait yağış ve akım verilerinden faydalanarak yağış akış modellemesi yapmışlardır. Çalışmada yapay sinir ağı, ANFIS ve gen ifade programlama yöntemleri kullanılmış, alınan sonuçlar çoklu lineer regresyon yöntemiyle elde edilmiş sonuçlar ile karşılaştırılmıştır. Kızılaslan vd. (2014), yapay sinir ağı kullanarak yaptıkları çalışmada farklı senaryolar belirleyerek en doğru sonuca ulaşmaya çalışmışlardır. Turhan ve Çağatay (2015), yaptıkları çalışmada bir istasyondan eksik olarak ölçülmüş aylık akım verilerini başka bir akım gözlem istasyonu yardımıyla tahmin etmek için yapay sinir ağı ile modelleme yapmışlardır. Ayrıca, Turhan ve Çağatay (2016)'da yaptıkları başka bir çalışmada Hatay havzasında yer alan akım gözlem istasyonu verilerinden faydalanarak eksik aylık akım verilerinin modellenmesi için ileri beslemeli geri yayımlı yapay sinir ağı yöntemini kullanmışlardır.

Bu çalışmada, Sakarya havzasında yer alan ve Porsuk Çayı üzerinde bulunan Uludere akım gözlem istasyonundan alınan akım gözlem verileri kullanılarak tahmin modelleri oluşturulmuştur. Yapay sinir ağlarının ileri beslemeli geri yayılım yöntemi ile hazırlanan bu tahmin modelleri daha sonra çoklu doğrusal regresyon analizi ve genetik programlama yöntemi sonuçlarıyla karşılaştırılmıştır.

## Materyal ve Yöntem

### Çalışma Alanı

Bu çalışmada akım verilerinin tahmin edilmesi için Sakarya havzasının alt havzası olan Aşağı Sakarya havzasında yer alan Porsuk Çayı üzerindeki istasyon verileri kullanılmıştır. 2005-2015 yıllarına ait akım verileri kullanılarak hazırlanan modeller bir yapay sinir ağı yöntemi olan ileri beslemeli geri yayılım sinir ağı kullanılarak performansları incelenmiştir. Daha sonra elde edilen sonuçlar çoklu doğrusal regresyon analizi ve gen ifade programlama sonuçlarıyla karşılaştırılmıştır.

### Akım Gözlem İstasyonu

Çalışmada kullanılan akım verileri Uludere Uluçayır akım gözlem istasyonundan temin edilmiştir. Yağış alanı 290.7 km<sup>2</sup>, rakım 826 m'dir. İstasyon 30° 24' 0" enleminde, 39° 38' 0" boylamında yer almaktadır ve Eskişehir Porsuk baraj yolunun 20. km sindeki Gökçek Işık köyünden sola ayrılan Yörük Kırkaköy yolunun 3 km ilerisinde bulunan Uluçayır köyündedir. 1964 yılında kurulan istasyon 2015 yılına kadar ölçümlerine devam etmiştir. Akım gözlem istasyonunun harita üzerinde yeri Şekil 1.'de verilmiştir.

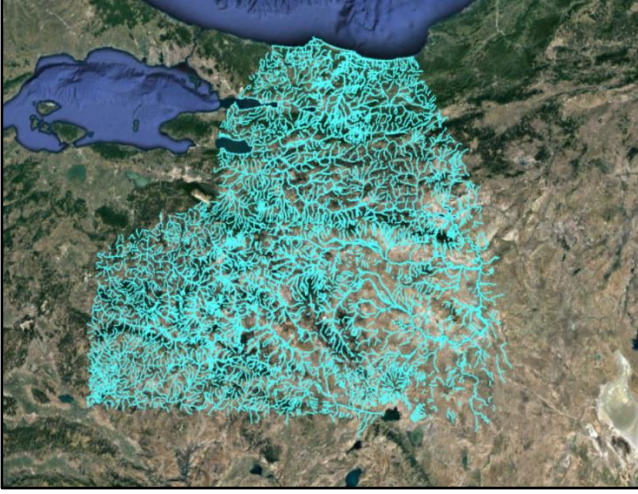


Şekil 1. Akım gözlem istasyonu (Google Maps görseli)

### Havza Bilgileri

Sakarya havzası Devlet Su İşleri Genel Müdürlüğü tarafından belirlenmiş 26 akarsu havzasından biridir. Havza Aşağı Sakarya havzası ve Yukarı Sakarya havzası olmak üzere 2 bölüme ayrılmıştır (DSİ, 2018). Çalışma Aşağı Sakarya Havzasında yer alan Porsuk Çayı akım

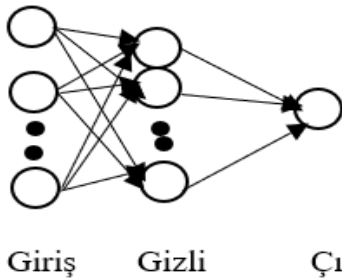
verileri kullanılarak yapılmıştır. Havzanın harita görseli Şekil 2.'de verilmiştir



Şekil 2. Sakarya Havzası (Google Maps görseli)

### Yapay Sinir Ağları

Yapay Sinir Ağları; insan beynindeki sinir hücrelerinden ilham alınarak geliştirilmiş yapay sinir hücrelerinin farklı geometrileri ile birbirlerine bağlanarak meydana getirilmiş karmaşık sistemler olarak tanımlanmaktadır. Giriş bilgilerini alır, kendi içinde ilişkilendirir ve belirli bir işlem sıralamasından geçirerek sonuca ulaşır. Genel olarak yapay sinir ağları (YSA), verilen girdilere karşı sonuçlar üreten bir kara kutu modelidir (Kohonen, 1988; Terzi ve Köse, 2012). Yapay sinir ağının yapısı Şekil 3.'de verilmiştir.



Şekil 3. Yapay sinir ağının yapısı

Yapay sinir ağı şeklinde tanımlanan yapı genel olarak sinir hücresi adı verilen katmanlar, bu katmanları birbirleri ile ilişkilendiren ağırlık değerleri, verilerle işlem yapan bir matematiksel fonksiyon, aktivasyon fonksiyonu ve çıkış değerleri şeklinde adlandırılan beş bölümden

oluşur. Her bir ağırlık değeri kendinden önceki ve sonraki katmanların arasındaki ilişkinin derecesini gösterir. Girdiler sonuç değerleriyle bu ağırlık değerlerine göre ilişkilendirilir ve toplama fonksiyonlarında gerekli işlemleri yapılır. Toplam fonksiyonunu gösteren matematiksel işlem denklem 1'de verilmiştir (Terzi ve Barak, 2015).

$$F_{\text{net}} = \sum_{i=1}^N w_{ij} x_i + c \quad (1)$$

Buradaki N değerleri giriş katmanındaki hücre sayısı,  $w_{ij}$  ağırlık değerleri,  $x_i$  giriş katmanında bulunan değerler ve c eşik düzey değeridir.

Girdilerin çıkış katman değerlerine dönüştürme işlemi için aktivasyon fonksiyonları kullanılmaktadır. Bu fonksiyonlar toplama fonksiyonlarından gelen değerleri alır ve çıkış katmanına bu değerleri aktarır. Genel olarak 0-1 değerleri arasında tanımlanan aktivasyon fonksiyonlarının lineer, step, sigmoid ve hiperbolik tanjant gibi birçok farklı çeşidi bulunmaktadır. Uygulamalarda sıklıkla sigmoid aktivasyon fonksiyonu seçilmektedir (Öztemel, 2003). Sigmoid fonksiyonu denklem 2' de verilmiştir.

$$F(y) = 1/[1+(e)^{-y}] \quad (2)$$

Yapay sinir ağları hesaplamaları arasında sürekli ileri doğru girdileri çıktılar haline dönüştüren ve hataların azaltılması için ağırlıkları geriye doğru yenilemek olmak üzere iki aşama vardır. Bir yapay sinir ağları modelinin geleceğe dair güvenilir tahminler verebilmesi için verilerin ağı çok iyi öğretilmesi gerekmektedir (Şen, 2004).

Ağın katmanlar arasındaki ağırlık değerlerini güncellediği formül denklem 3'de verilmiştir.

$$A_{ij}^y = A_{ij}^e - \eta \frac{\partial T_a}{\partial A_{ij}} \quad (3)$$

Burada  $A_{ij}^y$  ağın yeniden güncellediği yeni ağırlık değerlerini,  $A_{ij}^e$  ağın önceki ağırlık değerlerini,  $\eta$  ağın öğrenme oranını,  $\partial T_a \setminus \partial A_{ij}$  ağ hatasının eğimini göstermektedir. Yapay sinir ağları modellerinde ağın genelleme yeteneğinin

değerlendirilebilmesi için ağın giriş ve çıkış değerlerinin hepsi kullanılmamaktadır. Belirli bir kısım eğitim amaçlı diğer kısım ise test amacıyla kullanılmaktadır (Kişi, 2004; Gümüş vd., 2010).

### Çoklu Doğrusal Regresyon Analizi

Çoklu doğrusal regresyon analizi; bir değer değişiminin başka bağımsız değerlere bağlı olduğu durumlarda o değer tahmin edilmesi için kullanılan bir yöntemdir. Burada bağımlı değişken olan  $y$  değeri; bağımsız değişken  $x$  değerinin değişimine göre değişkenlik gösteriyorsa bu en basit tanımıyla  $y = ax + b$  şeklinde ifade edilir ve bu lineer regresyon denklemi olarak tanımlanır (Gümüş ve Kavşut, 2013). Değişken sayısı arttıkça farklı hesaplamalar için değişken seçim yöntemlerine ihtiyaç duyulur. Eğer bağımsız değişken sayısı birden fazla olursa ( $x_1, x_2, x_3, \dots, x_n$ ) yapılması gereken işlem denklem 4’de verilmiştir.

$$y = a + b_1x_1 + b_2x_2 + b_3x_3 + \dots + b_nx_n \pm c \quad (4)$$

Bu denklem sayesinde birden fazla değişken içeren veri dizilerinin anlamlı bir şekilde tahmin modelleri hazırlanabilmektedir (Alpar, 2003).

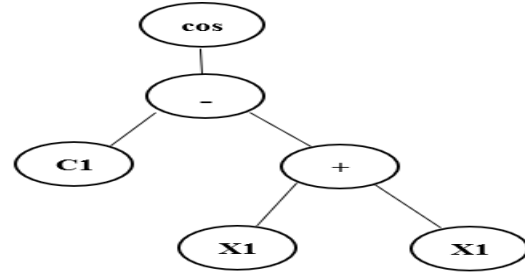
### Gen İfade Programlama (GEP)

Genetik programlama; Koza (1992) tarafından tasarlanmış ve bir sistemde genetik algoritmayı kullanarak çeşitli görevleri yerine getiren sistemlerdir. (Demir vd., 2014).

Gen ifade programı, genetik programlama ve genetik algoritma ile arasında çok küçük farklılıklar bulunan bir sistemdir. Gen ifade programı genetik programlamadan farklı olarak karmaşık ve doğrusal olmayan veriler arasındaki ilişkiyi kurmada başarılıdır. Genetik algoritmadan farkı ise değişik boyut ve şekle sahip doğrusal olmayan veriler arasındaki ilişkiyi tanımlamak için oluşturulan “işlem ağaçları” ile uygun çözümler verir (Lopes ve Weinert, 2004; Sarıdemir, 2016).

Genetik programlamada bireyler ya da veriler farklı boyut ve şekildeki lineer olmayan birimler (ağaçlar) olarak ifade edilen sabit uzunluktaki lineer kromozomlar şeklinde kodlanır. Bundan

dolayı gen ifade programlamadaki ana parametreler kromozom ve ağaçlardır (Kök vd., 2014). Bu ağaç yapılarına örnek bir çalışma Şekil 4’de verilmiştir.



Şekil 4. Gen ifade programlama işlem ağacı

Şekil 4’de yer alan C1 ifadesi sabit bir değer, X1 ise değişken değerimiz olup diğer matematiksel semboller ise yapılacak olan işlem sıralamasını belirtmektedir. Şekilde verilen işlem ağacının matematiksel ifadesi denklem 5’de verilmiştir.

$$\text{Cos} (C1 - (X1 + X1)) \quad (5)$$

### Uygulama

Çalışmada Uludere Uluçayır AGİ’ye ait 2005-2015 yılları arasında ölçülmüş 3620 adet günlük akım ölçüm verileri kullanılmıştır. Akım verileri ile ilgili ortalama, standart sapma, minimum ve maksimum değerleri hesaplanmıştır. En yüksek akım verisi  $33.3 \text{ m}^3/\text{s}$ , en düşük akım verisi ise akımın gözlenmediği  $0 \text{ m}^3/\text{s}$  yani yatağın kuru olduğu andır. 2005-2015 tarihleri arasında ölçülmüş günlük akım değerlerinin aritmetik ortalaması  $0.5839 \text{ m}^3/\text{s}$  ve standart sapma değeri 0.882 olarak hesaplanmıştır.

### Yapay Sinir Ağları ile Modellerinin Hazırlanması

Uludere Uluçayır AGİ’den temin edilen 3620 adet akım gözlem verisinin %80’i ağın eğitimi için %20’lik kısmı ise test aşaması için hazırlanmıştır. Verilerin analizi için bir yapay sinir ağı yöntemi olan ileri beslemeli geri yayımlı sinir ağı kullanılmıştır. Aktivasyon fonksiyonu olarak sigmoid fonksiyonu kullanılmıştır.

Veriler sigmoid fonksiyonunun kullanılması için  $(x - x_{\min}) / (x_{\max} - x_{\min})$  bağıntısının yardımıyla

0-1 aralığına normalize edilmiştir (Cobaner vd., 2008). Normalize edilen veriler sonuçlar alındıktan sonra yapılan normalize işleminin tersi uygulanmış ve gerçek veriler elde edilmiştir. Bu sayede veriler arasındaki uzaklık azaltılmış ve ağırlık eğitimi performansı artırılmıştır. Çalışmada 4 farklı senaryo hazırlanmıştır ve bu senaryoların performansları incelenmiştir. Birinci senaryoda bir gün önceki öteleme değerleriyle gerçek akım değerlerine ulaşmaya çalışılmıştır. İkinci senaryoda değerler bir gün ve 2 gün ötelenmiş şekilde kullanılmıştır. Üçüncü senaryoda giriş katman değerlerinin sayısı 3'e çıkarılmış ve akım değerleri sırasıyla 3 gün, 2 gün ve 1 gün

ötelenecek kullanılmıştır. Son olarak da 4. senaryoda veriler sırasıyla 4 gün öteleme, 3 gün öteleme, 2 gün öteleme ve son olarak da 1 gün öteleme yapılarak giriş katman değerleri olarak kullanılmıştır. Hazırlanan senaryolar Tablo 1'de verilmiştir.

Verilerin birbiriyle ilişkisini incelemek için ötelenmiş 4 verinin birbirleri arasındaki korelasyon ilişkisi hesaplanmış ve değerleri Tablo 2'de verilmiştir.

**Tablo 1.** Yapay sinir ağı senaryoları

Durum Çıktılar	Girdiler	
1	Q-1	Q
2	Q-2, Q-1	Q
3	Q-3, Q-2, Q-1	Q
4	Q-4, Q-3, Q-2, Q-1	Q

**Tablo 2.** Korelasyon analizi sonuçları

Veriler	Q	Q-1	Q-2	Q-3	Q-4
Q	1	0.9261	0.9256	0.8334	0.7285
Q-1		1	0.9169	0.8223	0.7592
Q-2			1	0.9185	0.8134
Q-3				1	0.9070
Q-4					1

### Çoklu Doğrusal Regresyon Analizi ile Modellerinin Hazırlanması

Çoklu doğrusal regresyon analizi için mevcut 3620 veri kullanılmıştır. 4 farklı model için 4 farklı regresyon denklemi kurulmuştur ve giriş

değerlerine göre de bağımsız değişken (x) sayıları belirlenmiştir.

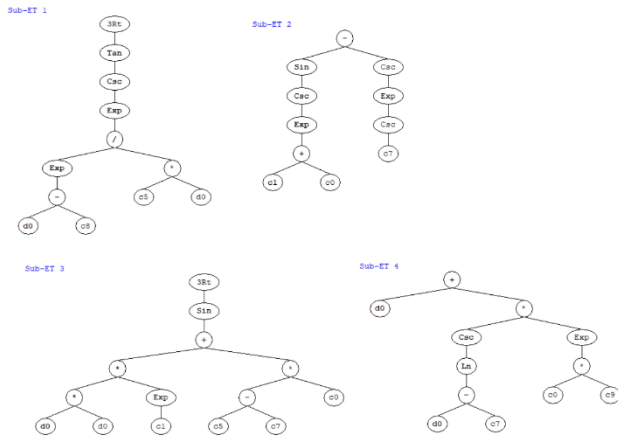
Analiz sonucu verileri tahmin edecek regresyon denklemleri oluşturulmuştur. Denklemler Tablo 3'de verilmiştir.

**Tablo 3.** Bulunan regresyon denklemleri

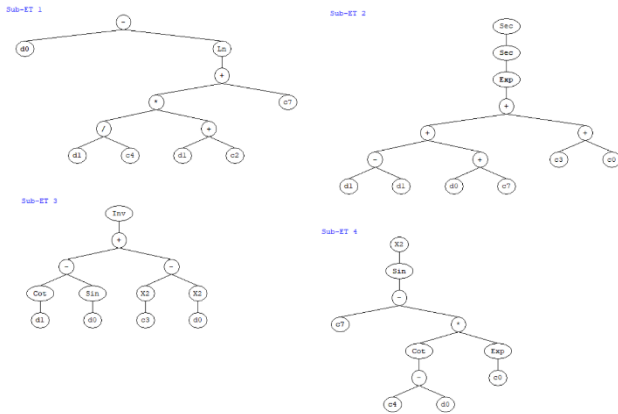
Model girdileri	Denklemler
Q-1	$y = 0.075161 + 0.8857x$
Q-1, Q-2	$y = 0.046 + 0.9337x_1 - 0.0853x_2$
Q-1, Q-2, Q-3	$y = 0.6858 + 0.9238x_1 - 0.185x_2 + 0.1482x_3$
Q-1, Q-2, Q-3, Q-4	$y = 0.0486 + 0.9005x_1 - 0.2204x_2 + 0.1471x_3 + 0.0328x_4$

## Gen İfade Programlama (GEP) ile Modellerinin Hazırlanması

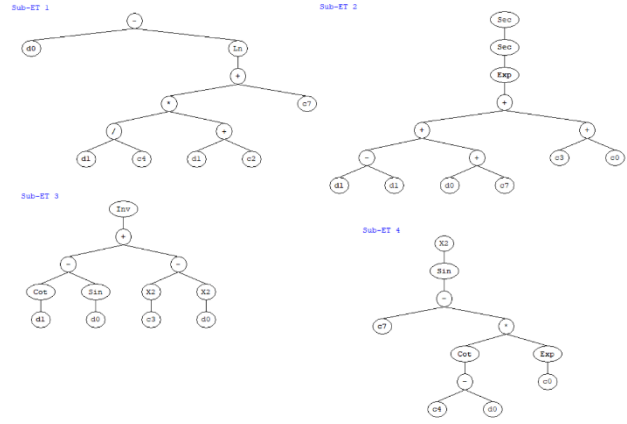
Genetik programlama için kullanılacak 3620 adet verinin %80'lik kısmı eğitim safhasında, %20'lik kısmı ise test safhasında kullanılmak üzere ayrılmıştır. Model girdileri karşılaştırma yapılabilmesi amacıyla yapay sinir ağları ve çoklu doğrusal regresyon analizindeki girdilerle aynı seçilmiştir ve 4 model girdisi için işlem ağaçları bulunmuştur. Gen ifade programlamada kromozom sayısı 50, gen sayısı 4, başlık sayısı 8 ve mutasyon oranı 0.0044 olarak seçilmiştir. Bulunan işlem ağaçları Şekil 5, Şekil 6, Şekil 7 ve Şekil 8'de gösterilmiştir.



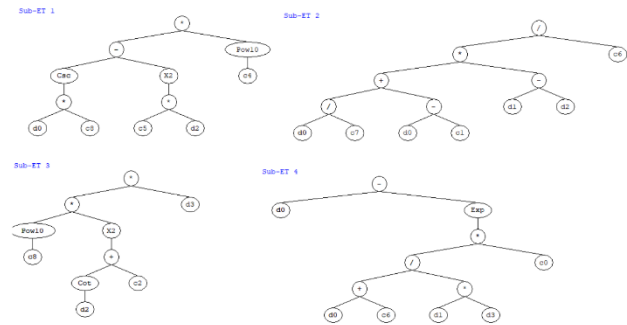
Şekil 5. Model-II GEP işlem ağacı



Şekil 6. Model-I GEP işlem ağacı



Şekil 7. Model-III GEP işlem ağacı



Şekil 8. Model-IV GEP işlem ağacı

## Bulgular ve Tartışma

Oluşturulan modellerin yapay sinir ağları (YSA) çoklu regresyon analizi ve gen ifade programlama (GEP) ile tahminleri yapılmıştır. Bulunan veriler gerçek verilerle karşılaştırılmış ortalama karesel hata ve  $R^2$  değerleri hesaplanmıştır. Ortalama karesel hata (OKH), veri serisinde gözlenen ve tahmin edilen veri değerlerinin farkının toplanıp, toplam veri sayısına bölünerek elde edilen bir değerdir. İstenilen değer ile bulunan değerler arasındaki hatayı karesel olarak gösterir. Bu değer sıfıra yakın olursa, tahmin edilen değer doğruya yakın olduğunu gösterir.  $R^2$  ise tahmin edilen verilerin gerçek verilerle olan uyumunu gösteren bir değerdir. 0 ile 1 arasında değer alan  $R^2$  değerinin 1'e yaklaşması bağımlı değişkendeki varyansın büyük bir kısmının modeldeki bağımsız olan değişkeni açıklayabildiği sonucunu verir (Ünver, 1999). Yapay sinir ağları ve çoklu regresyon analizi sonucu elde edilen verilerin OKH ve  $R^2$  değerleri Tablo 4.' de verilmiştir.

Modellerin OKH ve  $R^2$  değerlerinin yer aldığı Tablo 5. incelendiğinde, veri çokluğu bakımından en fazla veriye sahip Model-IV' ün en az dataya sahip Model-I' e göre daha iyi sonuçlar verdiği görülmüştür. Yapay sinir ağlarında model-IV çıkış katmanı verilerinden 0.9245  $R^2$  ve 25.98 OKH değerleri elde edilmişken; Model-I çıkış katmanı verilerinden 0.9185  $R^2$  ve 27.35 OKH değerleri elde edilmiştir. Gen ifade programlamada (GEP) model-IV çıkış katmanı verilerinden 0.8116  $R^2$  ve 25.98 OKH değerleri elde edilmişken; Model-I çıkış katmanı verilerinden 0.77  $R^2$  ve 34.04 OKH değerleri elde edilmiştir. Çoklu regresyon analizinde Model-IV çıkış verilerinden 0.7736  $R^2$  ve 34.82 OKH değerleri elde edilmişken; Model-I çıkış katmanı verilerinden 0.7540  $R^2$  değerine ve 36.57 OKH değerleri elde edilmiştir.

OKH değeri sıfıra en yakın olan model YSA ve çoklu regresyon analizinde 4 ötelemeli modeldir ve yine 1 değerine en yakın  $R^2$  modeli 4 ötelemeli model olmuştur. Gen ifade programlamada ise Model-III, Model-IV' e göre daha başarılı sonuçlar vermiştir.

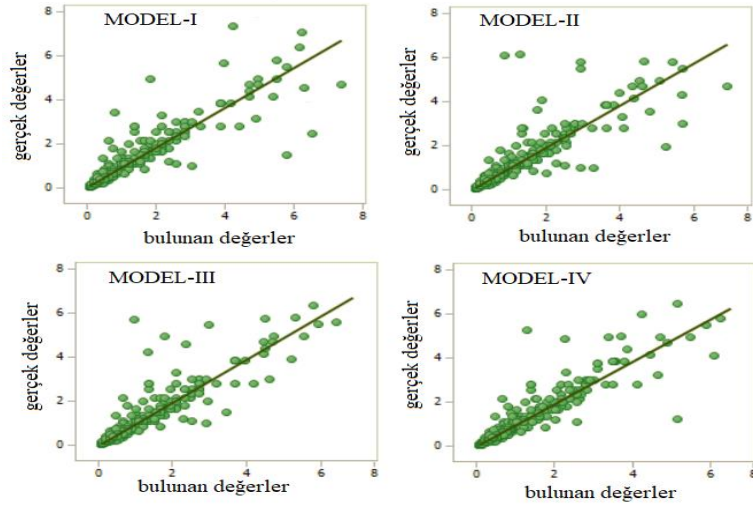
Gen ifade programlamada en iyi sonucu veren Model-III' e ait matematiksel gösterim aşağıda verilmiştir.

$$y=1/\ln[4.446-(\cot x_2/5.56)+1/8.04]+\ln(\sec(-3.28)-2.111)+(\cos(0.5)*x_2-2.11))+\tan(\exp(\sin(-0.436+7.525+\sec(3.14))^{1/3}))+\ln(\exp(\exp(\cos(x_3-x_1)*\ln 1.11*x_2))) \quad (6)$$

Modellere ilişkin OKH ve  $R^2$  değerleri Tablo 4' de verilmiştir.

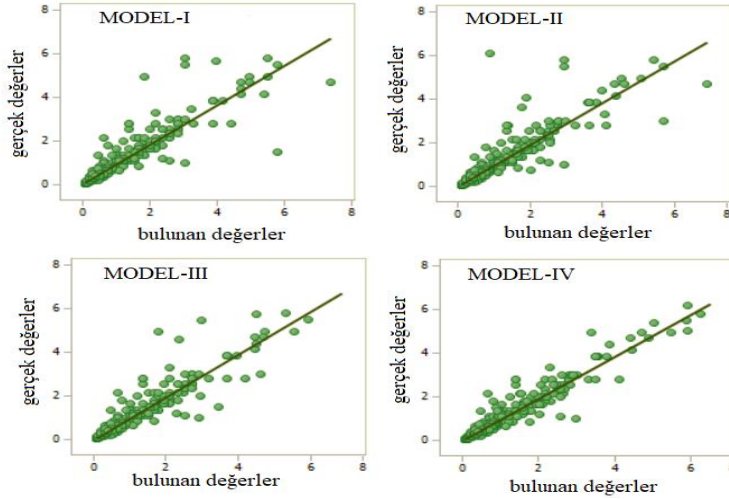
**Tablo 4.** Modellerin OKH ve  $R^2$  değerleri

Modeller	YSA		ÇDR		GEP	
	$R^2$	OKH	$R^2$	OKH	$R^2$	OKH
I- Q-1	0.9185	27.35	0.7540	36.57	0.7700	34.04
II- Q-1, Q-2	0.9168	26.84	0.7660	36.11	0.7920	33.67
III- Q-1, Q-2, Q-3	0.9221	26.77	0.7730	35.99	<b>0.8310</b>	<b>31.88</b>
IV- Q-1, Q-2, Q-3, Q-4	<b>0.9245</b>	<b>25.98</b>	<b>0.7736</b>	<b>34.82</b>	0.8116	32.07

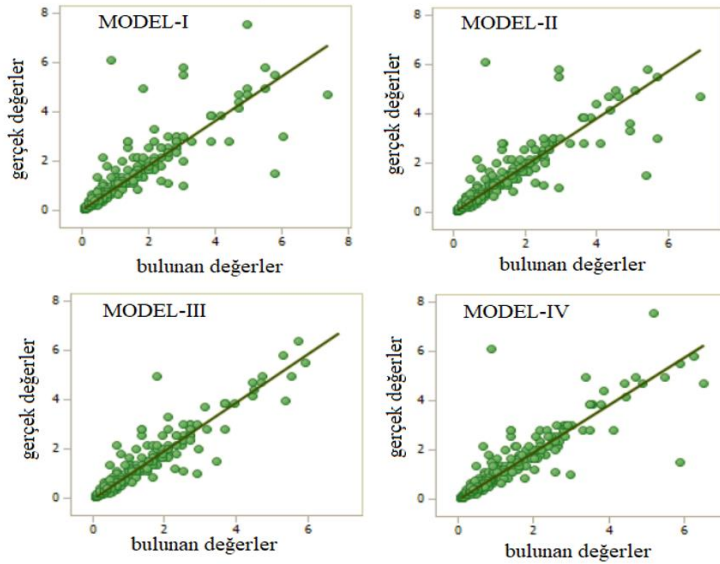


**Şekil 9.** Regresyon sonuçları saçınım diyagramları





Şekil 10. YSA sonuçları saçılım diyagramları



Şekil 11. GEP sonuçları saçılım diyagramları

Grafikler incelendiğinde YSA modellerinin regresyon modellerine ve gen iletim programlama modellerine göre verdiği değerlerin gerçek değerlere daha yakın olduğu görülmektedir.

## Sonuçlar

Bu çalışmada 3620 adet akım verisi yapay sinir ağları, çoklu regresyon analiz yöntemi ve gen ifade programlama ile modellenmiştir. Akım verileri 1 gün, 2 gün, 3 gün ve 4 gün ötelenerek modeller oluşturulmuş ve ileri beslemeli geri yayılım sinir ağı ile çoklu doğrusal regresyon

analizi ve gen ifade programlama yardımıyla 4 adet model hazırlanmış ve 3 farklı yöntem ile başarıları sınanmıştır. Veri ağının birbirleriyle olan ilişkisi ne kadar yüksekse tahminlerin doğruya daha yakın yapıldığı görülmüştür. Birbirleri ile ilişkisi yüksek olan veri sayısı artırıldığında her üç yöntem içinde bulunan tahmin verileri gerçek verilere daha fazla yaklaşmıştır. Gen ifade programlamada Model-III en iyi sonuçları vermiş 4 gün öteleme verilerinin de eklendiği Model-IV de ise gen ifade programlama tahmin değerleri gerçek değerlerden uzaklaşmıştır. Model-IV, yapay sinir ağları ve çoklu doğrusal regresyon analizinde

Model-I, Model-II ve Model-III tahmin verilerine göre en başarılı sonuçların alındığı model olmuştur. Yapay sinir ağları özellikle yüksek miktardaki verinin modellenmesi ve gerçek verilerin tahmininde çoklu regresyon analiz yöntemine ve gen ifade programlamaya göre daha yüksek başarı göstermiştir. Yapay sinir ağlarının bu başarısı mevcut verilerin birbiriyle olan ilişkilerini iyi bir şekilde analiz etmesinden kaynaklanmaktadır.

Sonuç olarak akım gözlem istasyonlarında ölçülemediği, yanlış ölçülmüş ya da ileriye dönük tahmin edilmesi gereken verilerin bulunmasında yapay sinir ağlarının, çoklu doğrusal regresyon analizi ve gen ifade programlamaya göre daha başarılı sonuçlar veren bir analiz tahmin yöntemi olduğunu görülmüştür.

## Kaynaklar

- Alp, M., Cıgızoğlu, H. K., (2004). Farklı Yapay Sinir Ağı Metotları İle Yağış-Akış İlişkisinin Modellenmesi, *İTE İnşaat Fakültesi, İnşaat Mühendisliği Bölümü Dergisi*, 3:1, 80-88.
- Alpar, R., (2003). Uygulamalı Çok Değişkenli İstatistiksel Yöntemlere Giriş 1, Detay Yayıncılık, Ankara, Türkiye.
- Aytek, A., Kisi, O., 2008. A genetic programming approach to suspended sediment modelling, *Journal of Hydrology*, 351:3, 288-298.
- Cıgızoğlu, H. K., (2003). Incorporation of ARMA models into flow forecasting by artificial neural networks, *Environmetrics*, 14:4, 417-427.
- Cobaner, M., Seckin, G., Kisi, O., (2008). Initial assessment of bridge backwater using an artificial neural network approach, *Canadian Journal of Civil Engineering*, 35, 500-510.
- Demir, M., Ceritbinmez, F., Kanca, E., (2014). Plastik Parça Üretiminde Çapaklanma Miktarının Genetik Algoritma Yöntemiyle Tahmin Edilmesi, Uluslararası Mühendislik ve Fen Bilimlerinde Yenilikçi Teknolojiler Sempozyumu, Karabük.
- Demirpençe, H., (2005). Köprüçay Akımlarının Yapay Sinir Ağları ile Tahmini, Antalya Yöresinin İnşaat Mühendisleri Sorunları Kongresi, Antalya.
- Fay, D., Ringwood, J. V., (2007). A wavelet transfer model for time-series forecasting, *International Journal Bifurcation and Chaos*, 17:10, 3691-3696.
- Gümüş, V., Kavşut M. E., Yenigün K., (2010). Yağış-Akış İlişkisinin Modellenmesinde YSA Kullanımının Değerlendirilmesi: Orta Fırat Havzası Uygulaması, Bilimde Modern Yöntemler Sempozyumu, Diyarbakır.
- Gümüş, V., Soydan, N. G., Şimşek, O., Aköz, M. S., Kırkgöz, M., S., (2013). Yağış Akış İlişkisinin Belirlenmesinde Farklı Yapay Sinir Ağı Yöntemlerinin Karşılaştırılması, *Çukurova Üniversitesi Mühendislik ve Mimarlık Fakültesi Dergisi*, 28:1, 37-49.
- Gümüş, V., Kavşut M. E., (2013). Zamanti Nehri Ergenusağı İstasyonu Eksik Aylık Akım Verilerinin Tahmini, *Gazi Üniversitesi Fen Bilimleri Dergisi*, 1:2, 81-91.
- Keskin, M. E., Taylan, E. D., (2007). Orta Akdeniz Havzasındaki Akımların Stokastik Modellemesi, *İMO Teknik Dergi*, 282, 4271-4291.
- Kızılaslan, M., Anıl, F.S., Doğan, E., Sönmez, O., (2014). Aşağı Sakarya Nehri Akımlarının Yapay Sinir Ağları ile Tahmin Edilmesi, *Sakarya Üniversitesi Fen Bilimleri Dergisi*, 18:2, 99-103.
- Kiş, Ö., (2004). Yapay Sinir Ağları ve Regresyon Teknikleri ile Filyos Çayı Akımlarının Tahmini, IV. Hidroloji Kongresi, İstanbul, Türkiye.
- Kisi, O., Shiri, J., Tombul, M. (2013). Modeling rainfallrunoff process using soft computing techniques, *Computers & Geosciences*, 51, 108-117.
- Kohonen, T., (1988). An introduction to neural computing, *Neural Networks*, 1, 3-6.
- Küçük, M., Ağrılioğlu, N., (2005). Dalgacık dönüşüm tekniği kullanılarak hidrolojik akım serilerinin modellenmesi, *İstanbul Teknik Üniversitesi Dergisi*, 5:2, 69-80.
- Koza, J. R., (1992). Genetic Programming: On the Programming of Computers by Means of Natural Selection, A Bradford Book The MIT Press Cambridge, Massachusetts London, England.
- Kök, M., Kanca, E., Doğan, M. A., Çavdar, F., (2014). AL2 O3 Takviyeli AA6061 MMK'lerde Genetik Algoritma Yöntemiyle Yüzey Pürüzlülüğünün Tahmin Edilmesi, *Mühendis ve Makina*, 55:651, 57-64.
- Lopes HS, Weinert WR., (2004). A gene expression programming system for time series modeling. In: Proceedings of XXV Iberian Latin American Congress on Computational Methods in Engineering (CILAMCE), Recife (Brazil), 10- 12/November, 1-13, 2004.
- Makkeasorn A., Chang N. B., Zhou X., (2008). Short-term streamflow forecasting with global climate change implications – A comparative study between genetic programming and neural network models, *Journaş of Hydrology*, 352, 336-354.
- Okkan, U., Mollamahmutoğlu, A., (2010). Yiğitler Çayı Günlük Akımlarının Yapay Sinir Ağları ve Regresyon Analizi İle Modellenmesi, *Dumlupınar Üniversitesi Fen Bilimleri Enstitüsü Dergisi*, 23, 33-48.
- Öztemel, E., (2003). Yapay Sinir Ağları, Papatya Yayıncılık, İstanbul, Türkiye.
- Saplıgolu, K., Çimen M., (2010). Yapay Sinir Ağlarını Kullanarak Günlük Yağış Tahmini, *Mühendislik Bilimleri Tasarım Dergisi*, 1:1, 14-21.
- Sarıdemir, M., (2017). Metakaolin ve silis dumanı içeren betonların basınç dayanımının gen ifadeli programlama

- ile tahmin edilmesi Pamukkale Üniversitesi Mühendislik Bilimleri Dergisi 23:3, 238-244.
- Savic D. A., Walters, G. A., Davidson J. W., (1999), A Genetic programming gapproach to rainfall-runoff modelling, *Water Resour Manag*, 13, 219-231.
- Shiri, J., Kisi, Ö., (2011). Comparison of genetic programming with neuro-fuzzy systems for predicting short-term water table depth fluctuations, *Computers & Geosciences*, 37:10, 1692-1701.
- Şen, Z., (2004). Yapay Sinir Ağları İlkeleri, Su Vakfı Yayınları, İstanbul, Türkiye.
- Terzi, Ö., Köse, M., (2012). Yapay Sinir Ağları Yöntemi İle Göksu Nehri'nin Akım Tahmini, *SDU Uluslararası Teknolojik Bilimler Dergisi*, 4:3, 1-7.
- Terzi, Ö., Barak M., (2015). Dalgacık-Sinir Ağı ile Yağış-Akış Tahmini: Kızılırmak Nehri Örneği, *Tarım Bilimleri Dergisi*, 21, 546-557.
- Turhan, E., Çağatay, H.Ö., (2015). Yapay Sinir Ağları (YSA) Yöntemi ile Göksun Nehri Poskoflu Akım Gözlem İstasyonu (AGİ) Eksik Akım Verilerinin Tahmin Edilmesi, VIII. Ulusal Hidroloji Kongresi, Şanlıurfa.
- Turhan, E., Çağatay, H.Ö., (2016). Eksik Akım Verilerinin Tahmin Modelinin Oluşturulmasında Yapay Sinir Ağlarının Kullanımı: Asi Nehri Demirköprü Akım Gözlem İstasyon Örneği, *Çukurova Üniversitesi Mühendislik Mimarlık Fakültesi Dergisi*, 31, 93-106.
- Ünver, Ö., (1999). Uygulamalı İstatistik Yöntemler, Siyasal Kitabevi, Ankara, Türkiye.
- Yurdusev, M. A., Acı, M., Turan, M. E., İçağa, Y., (2008). Akarçay Nehri Aylık Akımlarının Yapay Sinir Ağları ile Tahmini, *C.B.Ü. Fen Bilimleri Dergisi*, 4:1, 73-88.

---

Devlet Su İşleri Genel Müdürlüğü Web Sitesi.  
<http://www.dsi.gov.tr/toprak-ve-su-kaynaklari>,  
(05.04.2018).

## Assessment of streamflow data in Porsuk Creek

### Extended abstract

*The water resources in our world are not unlimited. The total amount of water is 1.4 billion km<sup>3</sup> and only 2.5% of them are fresh waters in rivers and lakes. Water, which is an indispensable resource for mankind, has an important place for all world life; protection and control is a resource that must be kept in fact.*

*In this study, prediction models were developed by using 3620 observation data obtained from Uludere flow observation station located in Porsuk Creek. Models prepared using current data for 2005-2015 are compared using prediction methods. These prediction models, which were prepared by forward feedback propagation method of artificial neural networks, were then compared with the results of multiple linear regression analysis and Gene Expression Programming results.*

*Artificial neural networks; It is defined as complex systems which are formed by connecting with different geometries of artificial nerve cells which are inspired by nerve cells in human brain. The nerve cells consist of five sections called layers, called the weight values, mathematical functions, activation function and output values that connect and these layers. The feed-forward inversion neural network used in the study is an artificial neural network method and has two stages, which transform the forward input into outputs and restore the weights backwards to reduce errors. The data are correlated to each other and the result values are compared to the actual values, then the system returns to the beginning and tries to give the closest results to the actual values by changing the weight values from the beginning.*

*Multiple linear regression analysis; is a method used to estimate that value when a value is dependent on other independent values. In its simplest form, it is mathematical expression of an unknown with a variable and a constant value. As the number of variables increases, different calculations and mathematical functions are needed. This analysis method is used in many analyzes in daily life both in terms of simplicity and processing speed.*

*Genetic programming; are the systems that are used to solve difficult problems by using genetic algorithm in a system. The most important feature of genetic programming is that its outputs data outputs as computer programs but not as mathematical values. The system scans all possible analyzes, continues to repeat the process even though it is enough and provides the most suitable solution to the problem as a computer program. The gene expression programming, which can deliver these results in different programming languages, is successful in establishing the relationship between complex and nonlinear data, unlike genetic programming. The results of the analysis of the models used in the study are given in the form of a processing tree with the help of gene expression programming. This process is transformed into mathematical functions and numerical data are obtained.*

*In this study, A total of 4 different input data were prepared by 1 day, 2 days, 3 days and 4 days. The prepared scenarios were evaluated by using artificial neural networks, multiple linear regression analysis and gene expression program. When the number of data associated with each other increases, the estimated data in each of the three methods takes a closer look at the actual data. Model-IV has been the most successful model in artificial neural network and multiple linear regression analysis. The most successful results in gene expression programming were obtained from Model III.*

*Artificial neural networks have shown a higher success rate than the multiple regression analysis method and Gene Expression Programming results, especially in modeling high-level data and predicting true data. This success of artificial neural networks is since existing data analyze well their relations with each other. As a result, it has been observed that artificial neural networks are an analytical method of predicting successful results in the presence of unmeasured data at monitoring stations.*

**Keywords:** Artificial neural networks, regression analysis, runoff, Porsuk creek, Gene Expression Programming