



GEN İFADE PROGRAMLAMA İLE GÖKSU NEHRİ'NİN AKIM TAHMİNİ

Özlem TERZİ*, Onur ÖZCANOĞLU

Süleyman Demirel Üniversitesi, Teknoloji Fakültesi, İnşaat Mühendisliği Bölümü, Isparta, Türkiye

Anahtar Kelimeler

*Gen İfade Programlama,
Günlük akım,
Göksu Nehri,
Doğu Akdeniz Havzası,
Akım Tahmini.*

Özet

Bu çalışmada, Gen İfade Programlama (GEP) yöntemi kullanılarak Doğu Akdeniz havzasında bulunan Göksu Nehri'nin akım tahmini yapılmıştır. Akım tahmini için Göksu Nehri'nde bulunan Bucakkışla, Karahacılı, Kırkkavak, Hamam, Yeşilköy ve Gravga akım gözlem istasyonlarından 2006–2010 yıllarına ait günlük akım değerleri kullanılarak modeller geliştirilmiştir. Karahacılı istasyonuna ait modeller geliştirilirken diğer istasyonların akım değerleri ve Karahacılı istasyonunun önceki günlerine ait akım değerleri girdi kullanılarak çeşitli kombinasyonlar denenmiştir. Modellerin performansları değerlendirildiğinde, akım tahmininde GEP yönteminin başarılı sonuçlar verdiği görülmüştür.

STREAMFLOW ESTIMATION OF GÖKSU RIVER WITH GENE EXPRESSION PROGRAMMING

Keywords

*Gene Expression
Programming,
Daily flow,
Göksu River,
Eastern Mediterranean Basin,
Streamflow estimation.*

Abstract

In this study, Gene Expression Programming (GEP) technique was used to forecast the river flow of Göksu River, located in Doğu Akdeniz Basin. Various models were developed using the daily river flow data of Bucakkışla, Karahacılı, Kırkkavak, Hamam, Yeşilköy and Gravga stations, located on Göksu River, for the period of 2006-2010. While the models of Karahacılı station were developed, several combinations were tried using the flow data of other stations and also the previous flow data of Karahacılı station as input. When the performances of the models are evaluated, it is seen that GEP is a successful method in the forecast of river flow.

Alıntı / Cite

Terzi, Ö., Özcanoglu, O., (2017), Gen İfade Programlama ile Göksu Nehri'nin Akım Tahmini, *Journal of Engineering Sciences and Design*, 5(3), 483-488.

Yazar Kimliği / Author ID (ORCID Number)

O. Terzi, 0000-0001-6429-5176

O. Ozcanoglu, 0000-0001-6221-4918

Başvuru Tarihi / Submission Date	24.07.2017
Revizyon Tarihi / Revision Date	21.08.2017
Kabul Tarihi / Accepted Date	12.09.2017
Yayın Tarihi / Published Date	18.12.2017

* İlgili yazar: ozlemterzi@sdu.edu.tr, +90-246-211-8001

1. Giriş

Günümüzde, ölçüm yapılamayan yerlerdeki nehir akımlarının tahmini araştırmacıların sıklıkla karşılaştığı problemlerden biridir. Özellikle su kaynaklarının planlanması ve projelendirilmesi aşamasında, güvenilir akım tahminlerinin yapılması büyük önem taşımaktadır. Akarsularda akım, Devlet Su İşleri (DSİ) tarafından akarsuyun belirli noktalarında kurulan ölçüm istasyonları ile belirlenmektedir. Ancak, herhangi bir nedenle veri elde edilememesi durumunda ve dolayısıyla ölçüm olmayan yerlerde günlük ve aylık akımların tahmini için çeşitli yöntemlerin kullanılmasına ihtiyaç duyulmaktadır (Yılmaz, 2014).

Doğrusal olmayan problemlerin ele alınmasında etkin olarak kullanılan ve başarılı sonuçlar veren ve 2001 yılında Ferreira tarafından geliştirilen gen ifade programlama (GEP) bu yöntemlerden biridir. Günümüzde su kaynakları problemlerinde bu yöntemi kullanan çeşitli araştırmalar mevcuttur (Nourani vd., 2012; Fernando vd., 2012; Shoaib vd., 2015; Shiri vd. 2014).

Bu çalışmada, genetik algoritma ve genetik programlamanın geliştirilmesi ile elde edilen GEP yöntemi kullanılarak Göksu Nehri'ne ait akımların tahmini için çeşitli modeller geliştirilmiştir. Bunun için Akdeniz Bölgesi'nde yer alan Göksu Nehri üzerinde bulunan altı adet akım gözlem istasyonuna ait akım değerleri kullanılmıştır. En başarılı GEP modeline ait matematiksel formül elde edilmiştir.

2. Bilimsel Yazın Taraması

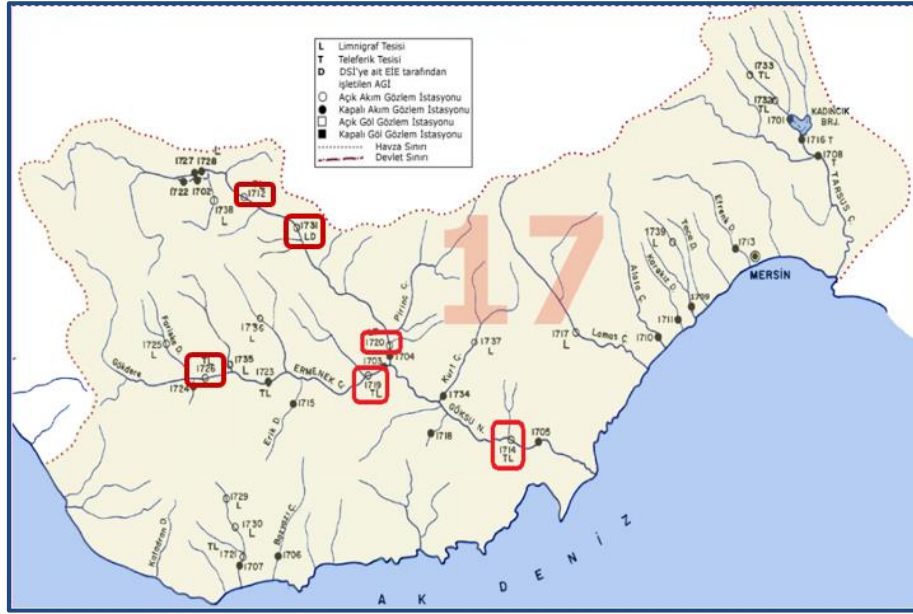
Aytek ve Kişi (2008) günlük sediment-debi ilişkisini çıkarmak için genetik programlama (GP) kullanmışlardır. Anahtar eğrisi ve çoklu lineer regresyon (ÇLR) yöntemi ile geliştirdikleri GP formülünü kıyaslamışlardır. GP formülünün iyi sonuçlar verdiğini ve kullanımının pratik olduğunu belirtmişlerdir. Shiri ve Kisi (2011) GP ile adaptif sinirsel bulanık çıkarım sistem (ANFIS) yöntemlerinin yeraltı su seviyesi tahminindeki kullanılabilirliğini incelemişlerdir. ABD'deki Bondville ve Perry istasyonlarına ait yeraltı su seviyesi verilerini kullanarak bir, iki ve üç gün sonraki yeraltı su seviyelerini tahmin edecek şekilde 5 farklı model geliştirmişlerdir. GP ve ANFIS yöntemlerinin başarılı sonuçlar verdiğini ve yeraltı su

seviyesi tahmininde kullanılabileceğini, ancak GP yönteminin formülleri açık bir şekilde verdiği için daha üstün olduğunu vurgulamışlardır. Azamathulla vd. (2011) Malezya'da bulunan Pahang Nehri'nin seviye-debi ilişkisini tahmin etmek için GP'nin devamı olan GEP yöntemini alternatif bir yöntem olarak sunmuşlardır. Elde ettikleri sonuçları daha geleneksel yöntemler olan anahtar eğrisi ve regresyon teknikleri ile de kıyaslamışlardır. GEP modelinin performansının GP ve diğer geleneksel yöntemlere göre çok daha iyi olduğunu belirlemişlerdir. Traore ve Guven (2013) Kuzey Afrika'da yer alan Burkino Faso'daki tropik ve kurak bölgelerdeki evapotranspirasyon miktarını modellemek için GEP yönteminden yararlanmışlardır. Evapotranspirasyon tahmininde birçok meteorolojik veri kombinasyonu kullanmışlar ve performans değerlendirmesi için determinasyon katsayısı (R^2) ve karekök ortalama hata (KOH) değerlerini kıyaslamışlardır. Çalışmada elde ettikleri sonuçlarda, GEP modelinin 0.979 R^2 ve 0.108 KOH ile başarılı bir yaklaşım elde ettiğini belirtmişlerdir. Kisi vd. (2013) Orta Sakarya havzasının bir alt havzası olan Kurukavak havzasındaki yağış-akış sürecini modellemek için 1987-1991 yıllarına ait ölçülmüş yağış ve akış verilerini kullanmışlardır. Yapay Sinir Ağları (YSA), ANFIS ve GEP yöntemlerini kullanarak geliştirdikleri modellerde değişik girdi kombinasyonlarını denemişler ve geleneksel ÇLR modeli ile de kıyaslama yapmışlardır. Çalışma sonuçlarına göre GEP yönteminin yağış-akış sürecini modellemede kullanılabileceği ve diğer yapay zekâ yöntemleri ile geleneksel yöntemlere göre geçerli bir alternatif olduğu belirtilmiştir.

3. Materyal ve Yöntem

3.1. Çalışma Bölgesi ve Veriler

Konya, Karaman, Antalya ve Mersin il sınırları içerisinde bulunan 260 km uzunluğa sahip Göksu Nehri, Akdeniz'e dökülmektedir. Kaynağı Toros Dağları'nda bulunan Geyik Dağları olan nehrin Gökçay ve Gökdere olmak üzere kuzey ve güney kolları mevcuttur. Ermenek'ten sonra Mut'un güneyinde birleşen bu iki kol Göksu adını alarak Taşucu ile Silifke arasında Akdeniz'e dökülmektedir. ([http://tr.wikipedia.org/wiki/Göksu_\(Kilikya\)](http://tr.wikipedia.org/wiki/Göksu_(Kilikya))). Çalışma bölgesi Şekil 1'de verilmiş olup, çalışmada kullanılan akım gözlem istasyonları şekil üzerinde kırmızı çerçeveler ile işaretlenmiştir.



Şekil 1. Göksu Nehri havzası ve çalışmada kullanılan akım gözlem istasyonları

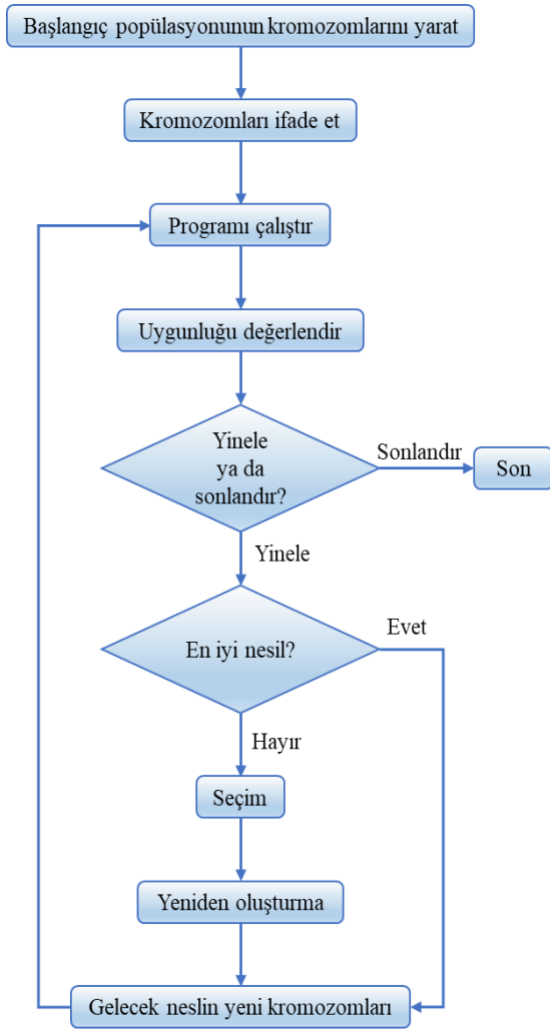
Modeller geliştirilirken, Elektrik İşleri Etüt İdaresi (EİE)'nden Göksu Nehri üzerinde bulunan Bucakkışla (1712), Karahacılı (1714), Kırkkavak (1719), Hamam (1720), Yeşilköy (1726) ve Gravga (1731) akım gözlem istasyonları için 2006–2010 yıllarına ait toplam 1826 adet günlük akım değerleri alınmıştır. İlk dört yıllık veri model geliştirmek için eğitim setine (% 80), kalan bir yıl ise modellerin uygunluğunu belirlemek için test setine (% 20) ayrılmıştır.

3.2. Gen İfade Programlama (GEP)

Matematiksel ifade, karar ağacı, polinom yapı ve mantıksal ifade gibi bilgisayar programlarını içeren bir arama tekniği olan Genetik Programlama (GP) ilk olarak Koza tarafından 1992 yılında önerilmiştir (Goldberg, 1989). Genetik algoritmalar (GA) genelleştirilerek elde edilmiş olan GP, problemin yapısına uygun fonksiyon ve terminalleri (girdi) içeren bilgisayar programlarının popülasyonu ile işleme başlar. Gen ifade programlama da (GEP), GA ve GP gibi bireylerin popülasyonlarını kullanarak uygunluklarına göre seçim yapar. Genetik varyasyonu bir ya da birden fazla genetik operatör kullanarak tanıttığı için genetik algoritmaya benzerdir (Aytekin ve Kişi, 2008). Bu üç algoritma arasındaki asıl fark bireylerin doğasından kaynaklanmaktadır. GP'de bireyler farklı boyutlarda ve şekillerde (potansiyel çözümleri temsil eden kromozomlar) doğrusal olmayan varlıklar iken, GA'da, sabit uzunlukta (kromozomlar) olan doğrusal dizelerdir. Amaç fonksiyonuna göre, her bir varlık farklı uygunlukları sergiler. Aynı zamanda, doğrusal olmayan varlıklar, diyagramlar veya ağaçlar şeklinde temsil edilebilirler. GEP'de ise bireyler, farklı şekillerde (basit diyagram gösterimleri veya ifade ağaçları) ve boyutlarda doğrusal olmayan varlıklar olarak ifade edilen sabit uzunlukta (kromozomlar) doğrusal dizeler olarak kodlanmıştır (Ferreira, 2001).

Genetik programlama kullanılarak bir problemin çözümünde beş temel adım mevcuttur. İlki, bilgisayar programlarında kullanılmak üzere, popülasyonda terminal (girdi) setini tespit etmektir. İkinci adım ise, fonksiyonları belirlemektir. Fonksiyon setinde, aritmetik operatörler (*, /, -, +), matematiksel fonksiyonlar (sin, cos, log), mantıksal ifadeler (eğer-daha sonra-başka) ve boolean operatörleri (ve, ya da, yoksa) veya kullanıcının tanımlandığı başka fonksiyonlar bulunur. Popülasyondaki bilgisayar programlarının oluşturulmasını sağlayan unsurlar, terminaller ve fonksiyonlardır. Üçüncü adım, mevcut problemin çözümünde bilgisayar programının başarı seviyesini değerlendirme şeklini belirlemektir. Sayısal parametrelerin ve nitel değişkenlerin değerlerinin verildiği kontrol parametrelerinin olduğu dördüncü adımda ise, çalışma aşaması kontrol edilir. Son adımda, çalışmanın sonlandırıldığı ve sonuç tayininin yapıldığı bir kriter belirlenir (Koza, 1992).

Şematik olarak GEP'in bu önemli adımları Şekil 2'de verilmiştir. Belli bir sayıda bireyin kromozomlarının rastgele üretilmesi ile işlem başlamaktadır. Sonra bu kromozomlar ifade edilir ve her bir bireyin uygunluğu, fitness kümesi ile değerlendirilmektedir. Yeni özellikleri ile üremek ve modifikasyon ile yeniden oluşmak için bireyler uygunluklarına göre seçilir. Aynı gelişme sürecine, yeni bireyler de tabi tutulur. İfade genomları, seçim ortamının kıyaslanması, modifikasyon ve seçim ile üretilir. İyi bir çözüm bulunana kadar veya belli bir sayıda işlem yinelenir (Ferreira, 2006).



Şekil 2. Bir gen ifade algoritmasının akış şeması (Ferreira, 2006)

İfade ağaçlarına dönüştürülen doğrusal kromozomlarda, GEP'in bireyleri kodlanmıştır. Böylece, GEP'de doğrusal kromozomlar ve ifade ağaçları hem yapısal hem de işlevsel olarak birlikte çalışan farklı varlıklar olmuştur. GEP, kromozomlar ve kromozomlarda kodlanan genetik bilgiyi gösteren ifade ağaçları olmak üzere iki temel kısımdan oluşmaktadır (Grosan ve Abraham, 2006).

4. Araştırma Bulguları

Göksu Nehri'ne ait akım tahmin modelleri geliştirmek için havzada bulunan altı adet akım gözlem istasyonunun akım değerleri ile GEP yöntemi kullanılmıştır. Akım tahmin modellerinde, havzanın çıkış noktasına en yakın istasyon olan Karahacılı istasyonunun günlük akım değerleri çıktı olarak alınmıştır. GEP modelleri geliştirilirken determinasyon katsayısı (R^2) ve karekök ortalama hata (KOH) değerleri uygunluk fonksiyonu olarak seçilmiştir. Terminal (T) setinde bağımsız değişkenlerden oluşan akım miktarları kullanılırken, fonksiyon setinde ise, aritmetik operatörler (*, /, -, +) ve bazı matematiksel fonksiyonlar ($\sqrt{\quad}$, \wedge , \ln , $10x$, \log) kullanılmıştır. Kromozom mimarisi, gen sayısı 3 ve

başlık uzunluğu 8 olacak şekilde belirlenmiştir. Bağlantı fonksiyon türü olarak toplam (+) seçilmiştir. Ayrıca, modellerde kullanılan genetik operatörler seti Tablo 1'de verilmiştir.

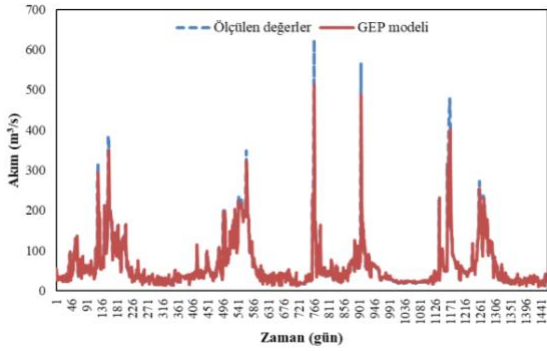
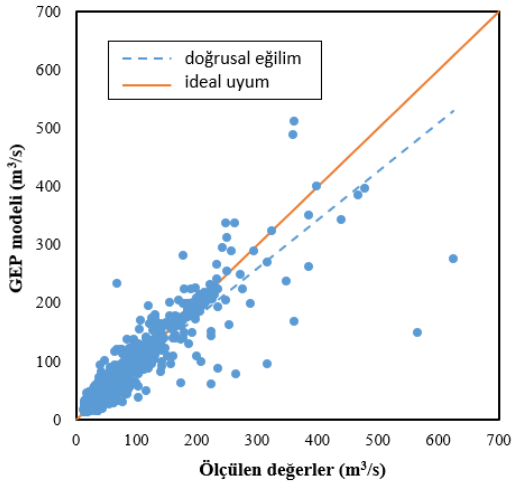
Tablo 1. GEP modeli parametreleri

Gen sayısı	3
Başlık uzunluğu	8
Kromozom sayısı	50
Bağlantı fonksiyonu	+
Mutasyon oranı	0,044
Tek noktada rekombinasyon oranı	0,3
İki noktada rekombinasyon oranı	0,3
Gen rekombinasyon oranı	0,1
Gen aktarım hızı	0,1

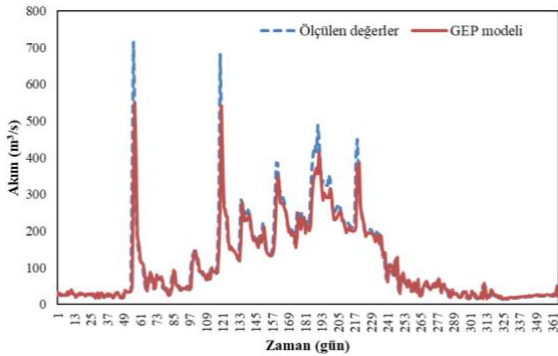
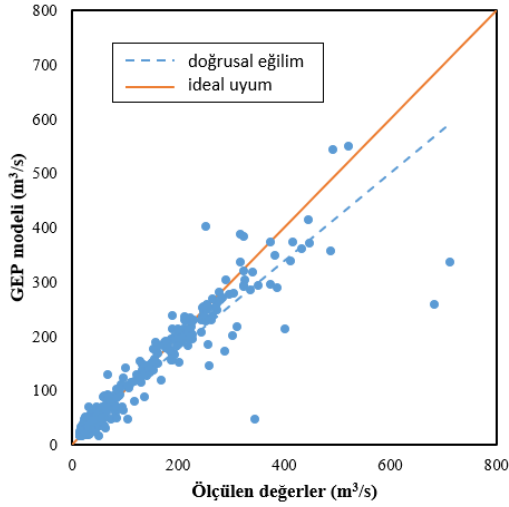
Çalışmada geliştirilen GEP modellerinin eğitim ve test setlerine ait KOH ve R^2 değerleri Tablo 2'de verilmiştir. Tablo 2 incelendiğinde, aynı istasyonun önceki günlere ait akım değerleri kullanılarak geliştirilen modellerin R^2 değerleri, diğer istasyonların verileri kullanılarak geliştirilen modellerden daha yüksek bulunmuştur. Buna göre, geliştirilen modeller arasında Q_{t-1} ve Q_{t-2} verilerinin girdi olarak kullanıldığı model 7'nin diğerlerine göre daha iyi sonuç verdiği görülmüştür. Bu modelin eğitim setine ait KOH 25,04 m^3/s ve R^2 değeri 0,839 olarak hesaplanırken, test setine ait KOH 43,00 m^3/s ve R^2 değeri ise 0,868 olarak bulunmuştur. Ayrıca, model 7'nin eğitim ve test setlerine ait saçılma diyagramları ile zaman serileri Şekil 3 ve Şekil 4'de gösterilmiştir. Şekil 3 ve 4'den görüldüğü gibi ölçülen değerler ile model sonuçları arasında iyi bir uyum olduğu ve noktaların 45 derecelik ideal uyum çizgisi etrafında toplandığı görülmüştür. Doğrusal eğilim çizgisi ile ideal uyum çizgileri arasındaki açının küçük olması da modelin başarısının yüksek olduğunu göstermektedir.

Tablo 2. Geliştirilen modellerin performansları

Model no	Model girdileri	Eğitim seti		Test seti	
		KOH (m^3/s)	R^2	KOH (m^3/s)	R^2
1	$Q_{(1720)t}$	42.02	0.546	84.45	0.492
2	$Q_{(1720)t}, Q_{(1719)t}$	42.16	0.543	86.58	0.466
3	$Q_{(1720)t}, Q_{(1719)t}, Q_{(1712)t}$	42.98	0.525	85.69	0.477
4	$Q_{(1720)t}, Q_{(1719)t}, Q_{(1712)t}, Q_{(1731)t}$	42.84	0.529	86.28	0.469
5	$Q_{(1720)t}, Q_{(1719)t}, Q_{(1712)t}, Q_{(1731)t}, Q_{(1726)t}$	39.62	0.597	89.46	0.429
6	Q_{t-1}	25.27	0.836	44.01	0.862
7	Q_{t-1}, Q_{t-2}	25.04	0.839	43.00	0.868
8	$Q_{t-1}, Q_{t-2}, Q_{t-3}$	25.94	0.827	44.25	0.860
9	$Q_{t-1}, Q_{t-2}, Q_{t-3}, Q_{t-4}$	25.76	0.830	43.24	0.867
10	$Q_{t-1}, Q_{t-2}, Q_{t-3}, Q_{t-4}, Q_{t-5}$	22.09	0.875	47.99	0.836



Şekil 3. Model 7'nin eğitim setine ait saçılma diyagramı ve zaman serisi

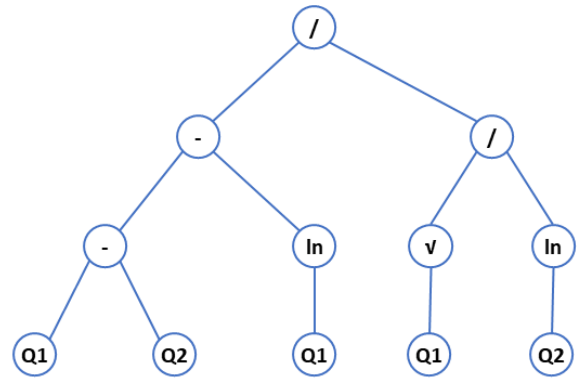


Şekil 4. Model 7'nin test setine ait saçılma diyagramı ve zaman serisi

Çalışmada elde edilen modellerden en uygun sonuçları veren model 7'ye ait matematiksel formül Eşitlik(1)'de verilmiştir. Bu modele ait olarak genetik kod kurallarına göre çıkarılan ifade ağacı ise Şekil 5'de gösterilmiştir. Formülde ve ifade ağacında yer alan Q sembolü Karahacılı istasyonunun t günündeki akım değerini, Q1 sembolü (t-1) günündeki akım değerini, Q2 sembolü ise (t-2) günündeki akım değerini ifade etmektedir.

$$Q = \left(\frac{(Q_1 - Q_2) - \ln Q_1}{\frac{\sqrt{Q_1}}{\ln Q_2}} \right) + Q_1 + \left(\sqrt{Q_1} - (\ln Q_1) \sqrt{\sqrt{Q_1}} \right) \quad (1)$$

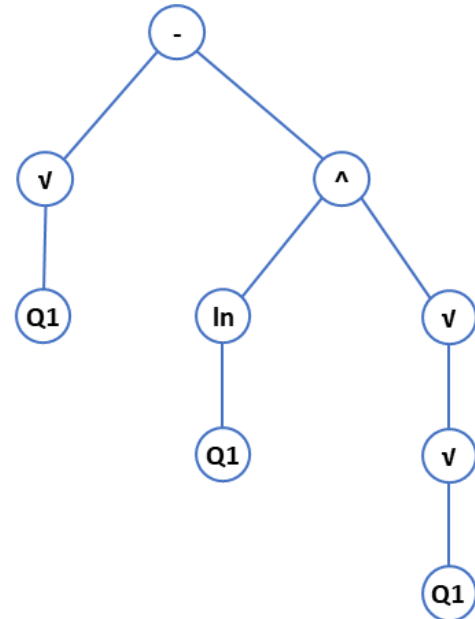
Alt ifade ağacı 1



Alt ifade ağacı 2



Alt ifade ağacı 3



Şekil 5. Model 7'ye ait ifade ağacı

5. Sonuç ve Tartışma

Bu çalışmada, Göksu Nehri'ne ait akım tahmini için gen ifade programlama (GEP) yöntemi ile çeşitli modeller geliştirilmiştir. Bu nehir üzerinde bulunan Karahacılı istasyonuna ait günlük akım tahmin modelleri geliştirmek için bu istasyonun önceki günlerine ait akım değerleri ve nehir üzerinde bulunan diğer beş adet akım gözlem istasyonunun akım değerleri kullanılmıştır. Önceki günlere ait değerler kullanılarak geliştirilen modellerin uygunluğu, diğer istasyonların değerleri kullanılarak geliştirilen modellere göre daha yüksek bulunmuştur. Bu modeller içerisinde, Q_{t-1} ve Q_{t-2} girdilerinin kullanıldığı modelde oldukça yüksek determinasyon katsayısı elde edilmiştir. Sonuç olarak, GEP yönteminin nehir akım tahmininde kullanılabilir olduğu görülmüştür. Ayrıca, GEP yöntemi ile geliştirilen modellerin matematiksel formüllerinin elde edilebiliyor olması, yeni veriler eklenmesi durumunda daha sonraki çalışmalarda akım tahmininde büyük kolaylık sağlayacaktır.

Teşekkür

3026-YL-11 no.lu proje ile çalışmamı maddi olarak destekleyen SDÜ Araştırma Projeleri Yönetim Birimi'ne teşekkür ederiz.

Conflict of Interest / Çıkar Çatışması

Yazarlar tarafından herhangi bir çıkar çatışması beyan edilmemiştir.
No conflict of interest was declared by the authors.

Kaynaklar

- Aytek, A., Kisi, O. 2008. A genetic programming approach to suspended sediment modelling. *Journal of Hydrology*, 351(3), 288-298.
- Azamathulla, H. M., Ghani, A. A., Leow, C. S., Chang, C. K., Zakaria, N. A. 2011. Gene-expression programming for the development of a stage-discharge curve of the Pahang River. *Water resources management*, 25(11), 2901-2916.
- Fernando, A. K., Shamseldin, A. Y., Abraham, R. J. 2011. Use of gene expression programming for multimodel combination of rainfall-runoff models. *Journal of Hydrologic Engineering*, 17(9), 975-985.
- Ferreira, C. 2001. Gene expression programming: A new adaptive algorithm for solving problems. *Complex Systems* 13(2), 87-129.
- Ferreira, C. 2006. Gene-expression programming: Mathematical modeling by an artificial intelligence. Springer-Verlag Berlin Heidelberg.
- Goldberg, D.E. 1989. Genetic algorithms in search, optimization and machine learning. Addison-Wesley.

- Göksu (Kilikya). [http://tr.wikipedia.org/wiki/Göksu_\(Kilikya\)](http://tr.wikipedia.org/wiki/Göksu_(Kilikya)) (Erişim Tarihi: 02.04.2017)
- Grosan, C., Abraham, A. 2006. Evolving computer programs for knowledge discovery. *International Journal of System Management* 4(2), 7-24.
- Kisi, O., Shiri, J., Tombul, M. 2013. Modeling rainfall-runoff process using soft computing techniques. *Computers & Geosciences*, 51, 108-117.
- 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.
- Nourani, V., Komasi, M., Alami, M. T. 2012. Hybrid wavelet-genetic programming approach to optimize ANN modeling of rainfall-runoff process. *Journal of Hydrologic Engineering*, 17(6), 724-741.
- 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.
- Shiri, J., Sadraddini, A. A., Nazemi, A. H., Kisi, O., Landaras, G., Fard, A. F., Marti, P. 2014. Generalizability of Gene Expression Programming-based approaches for estimating daily reference evapotranspiration in coastal stations of Iran. *Journal of hydrology*, 508, 1-11.
- Shoaib, M., Shamseldin, A. Y., Melville, B. W., Khan, M. M. 2015. Runoff forecasting using hybrid wavelet gene expression programming (WGEP) approach. *Journal of Hydrology*, 527, 326-344.
- Traore, S., Guven, A. 2013. New algebraic formulations of evapotranspiration extracted from gene-expression programming in the tropical seasonally dry regions of West Africa. *Irrigation Science*, 31(1), 1-10.
- Yılmaz, M.U. 2014. Performans Ağırlıklı Yöntemlerle Aylık Akımların Tahmini: Orta Fırat Havzası Uygulaması. İTÜ Fen Bilimleri Enstitüsü YL Tezi, 89 s.