

Eksik Hidrolojik Verilerin Simbiyotik Organizmalar Arama Algoritması ile Tahmini

Kemal Saplıoğlu^{1,*}, Tülay Suğra Küçükerdem Öztürk¹, Fatih Ahmet Şenel²

¹İnşaat Mühendisliği Bölümü, Mühendislik Fakültesi, Süleyman Demirel Üniversitesi, Isparta, Türkiye

²Bilgisayar Mühendisliği Bölümü, Mühendislik Fakültesi, Süleyman Demirel Üniversitesi, Isparta, Türkiye

Makale Tarihiçesi

Gönderim: 03.10.2019

Kabul: 05.02.2020

Yayın: 22.05.2020

Araştırma Makalesi

Öz – Su kaynakları proje ve planlamalarının en etkili şekilde yapılabilmesi için düzenli olarak verilerin toplanması ve bu verilerin analiz edilmesi gerekmektedir. Ancak gerek maddi gerekse teknik nedenlerden dolayı bazı alanlarda veriler düzenli olarak toplanamamaktadır. Bu durum ise eksik veri problemini beraberinde getirmektedir. Eksik veri problemi su kaynaklarının planlanmasında, projelendirilmesinde ve yönetiminde birtakım sorunlar meydana getirmektedir. Bu problemin çözümü için ölçüm yapılan istasyona benzer nitelikteki diğer istasyon verilerine ihtiyaç duyulmaktadır. Eksik verilerin tamamlanması için literatürde çok farklı çalışmalar yapılmıştır. Bu çalışmada ise Türkiye'nin Yeşilirmak nehri üzerinde bulunan ölçüm istasyonları kullanılmıştır. Çalışmada Symbiotic Organisms Search (SOS) algoritması yardımı ile 3 farklı fonksiyon optimize edilmiştir. Ayrıca optimize edilen fonksiyonlar yapay sinir ağları, normal oran metodu ve çoklu regresyon yöntemlerinden elde edilen sonuçlar ile karşılaştırılmıştır. Oluşturulan modellerden elde edilen sonuçlar Mallows's Cp ile test edilmiş ve sonuçların kabul edilebilir düzeyde olduğu görülmüştür. Yapay sinir ağları ile oluşturulan 6 adet modelin sonuçları ise normal oran metodu ve çoklu regresyona göre daha iyi sonuç vermesine rağmen Symbiotic Organisms Search optimizasyon yöntemi kadar başarılı olamamıştır. Yapay sinir ağları modellerinden en iyisi de bu çalışma için 8 nöronlu olarak tespit edilmiştir. Çalışmada SOS olmadan oluşturulan modellerin hata değerlerinin %3-%4 seviyelerinde olduğu, SOS ile optimize edilen fonksiyon sonuçlarının diğer yöntemlere göre daha iyi olduğu görülmüştür.

Anahtar Kelimeler – Eksik veriler, simbiyotik organizmalar arama algoritması, Yeşilirmak nehri, tahminleme

Estimation of Missing Hydrological Data by Symbiotic Organisms Search Algorithm

¹Department of Civil Engineering, Faculty of Engineering, Suleyman Demirel University, Isparta, Turkey

²Department of Computer Engineering, Faculty of Engineering, Suleyman Demirel University, Isparta, Turkey

Article History

Received: 03.10.2019


Accepted: 05.02.2020


Published: 22.05.2020


Research Article

Abstract – In order for water resources projects and planning to be made in the most effective way, data must be collected and analyzed regularly. However, due to financial and technical reasons, data cannot be collected regularly in some areas. This situation brings with it the missing data problem. The missing data problem poses some problems in planning, designing and managing water resources. For the solution of this problem, other station data similar to the station under measurement is needed. Many different studies have been conducted in the literature to complete the missing data. In this study it was used measuring station located on Turkey's Yesilirmak river. In the study, 3 different functions were optimized with the help of Symbiotic Organisms Search (SOS) algorithm. In addition, the optimized functions were compared with the results obtained from artificial neural networks, normal ratio method and multiple regression methods. The results obtained from the models created were tested with Mallows's Cp and the results were found to be acceptable. Although the results of 6 models created with artificial neural networks give better results than the normal rate method and multiple regression, Symbiotic Organisms Search was not as successful as the optimization method. The best of artificial neural network models was found to be 8 neurons for this study. In the study, it was observed that the error values of the models created without SOS were at the level of 3% 4%, and the results of the functions optimized with SOS were better than other methods.

Keywords – Missing data, symbiotic organisms search, Yesilirmak river, estimating

¹  <http://orcid.org/0000-0003-0016-8690> kemalsaplioglu@sdu.edu.tr

²  <http://orcid.org/0000-0002-1102-1718> tulaykucukerdem@sdu.edu.tr

³  <http://orcid.org/0000-0003-1918-7277> fatihsenel@sdu.edu.tr

*Sorumlu Yazar / Corresponding Author

1. Giriş

İklim değişiklikleri ile birlikte artan çevresel problemlerin başında kullanılabilir su kaynaklarının azalması gelmektedir. Bu durum sınırlı bir kaynak olan suyun kullanımının planlanmasını önemli hale getirmektedir. Su kaynakları potansiyelinin belirlenmesinde doğal olaylar büyük rol oynamaktadır. Tabiat olayları çok uzun yıllardır devam etmesine rağmen bu olaylar ile ilgili kayıtlar sınırlıdır. Geçmiş tarihlere ait veriler; iklim koşullarının güçlüğü, ulaşım şartlarının zorluğu veya teknolojik yetersizlikler gibi sebeplerle bazı durumlarda kayıt altına alınamamıştır. Su kaynaklarının planlanması ve yönetiminde belirli hidrolojik verilere ihtiyaç duyulmaktadır. Planlama ve yönetim çalışmaları bu veriler üzerinden yapılacağı için; verilerin, uzun süreli ve eksiksiz olmasının önemi büyüktür. Fakat çoğu zaman çeşitli sebeplerle ölçüm yapılamayan alanlarda eksik veriler oluşmaktadır. Verilerdeki bu kesinti, su kaynaklarının etkili bir şekilde planlanması, projelendirilmesi ve yönetiminde sorun haline gelmektedir (Önöz, 2015). Eksik verilerin tamamlanması ve gelecek planlaması amacıyla ileriye dönük tahminlerin yapılması da su kaynakları ve hidrolojik çalışmalar için önemlidir (Şen, 2004).

Eksik verilerin tamamlanması ve ileriye dönük tahminler yapmak için literatürde çeşitli yöntemlerin önerildiği farklı çalışmalar mevcuttur (Elshorbagy, Simonovic ve Panu, 2002; Partal, 2007; Yurdusev, Acı, Turan ve İçağa, 2008; Mwale, Adeloje ve Rustum, 2012; Bardossy ve Pegram, 2014; Tencaliec, Favre, Prieur ve Mathevet, 2015). Bu çalışmalar incelendiğinde; yapay sinir ağları, regresyon analizi ve zaman serisi analizi en çok kullanılan yöntemler olarak karşımıza çıkmaktadır.

Yapay sinir ağlarının, su kaynakları ve hidrolojik verilerin tahmininde kullanıldığı çok sayıda çalışma mevcuttur. Yıldırım ve Kandemir (2018), çalışmalarında yapay sinir ağları ile çok sayıda farklı modeller kurarak yağış miktarını tahmin etmeye çalışmışlar, elde ettikleri sonuçları çoklu doğrusal regresyon analizi sonuçları ile kıyaslamışlardır. Yapay sinir ağları ile kurdukları modellerin çoklu doğrusal regresyon analizi sonuçlarına göre daha iyi sonuçlar verdiğini belirtmişlerdir. Benzer bir çalışmada Turhan ve Çağatay (2016), Asi nehri için akım verilerini tahmin etmede ileri beslemeli geri yayımlı yapay sinir ağları modeli ile çoklu doğrusal ve doğrusal olmayan regresyon modellerini kıyaslamışlar ve yapay sinir ağları modelleri ile daha başarılı sonuçlara ulaştıklarını belirtmişlerdir. Gümüş ve Kavşut (2013), çalışmalarında ileri beslemeli geri yayımlı, radyal tabanlı ve genelleştirilmiş regresyon yapay sinir ağları yöntemlerini kullanarak modelleme yapmışlar ve yapay sinir ağı yöntemlerinin birbirine yakın sonuçlar vermekle birlikte, ileri beslemeli geri yayımlı yapay sinir ağı modelinden daha başarılı sonuçlar elde ettiklerini ifade etmişlerdir.

Regresyon modellerinden faydalanılarak yapılan bir çalışmada Tosunoğlu, İspirli, Gürbüz ve Şengül (2017), Fırat havzasında eksik akım verilerini regresyon modelleri ve debi süreklilik çizgileri kullanarak tahmin etmeye çalışmışlar, her iki yöntem ile de iyi sonuçlar elde etmelerine rağmen kurulan regresyon modellerinin daha başarılı sonuçlar verdiğini belirtmişlerdir. Bakış ve Göncü (2015), çalışmalarında Zap suyu havzasındaki istasyonlarda eksik verileri tamamlamak için bir regresyon analizi uygulamışlardır. Çalışma sonucunda istasyonlardaki eksik veri sayısının az olduğu durumlarda elde ettikleri sonuçların başarılı olduğunu, eksik veri sayısının fazla olduğu durumlarda güvenilir sonuçlar elde etmenin daha zor olduğunu ifade etmişlerdir.

Eksik verilerin tahmin edilmesinde zaman serileri analizi kullanımı söz konusu olduğunda ise, ülkemizde kısıtlı sayıda çalışma bulunmakla birlikte dünya genelinde fazla sayıda çalışma bulunmaktadır (Salas, 1993; Boakye ve Schultz, 1994). Beauchamp, Downing ve Railsback (1989), nehir akışını tahmin etmede regresyon modeli ve zaman serisi teknikleri kullanmışlar, eksik verilerin tahmininde zaman serisi modeli kullanmanın kabul edilebilir düzeyde doğru sonuçlar verdiğini belirtmişlerdir. Ayrıca eksik verilerin tahmini için Devlet Su İşleri tarafından da kullanılan normal oran metodu sıklıkla kullanılmıştır (Bayazit, 2013; İsmail ve Zin, 2017).

Bu çalışmada Türkiye'nin Yeşilirmak nehrinde yer alan 5 adet istasyon kullanılmıştır. İstasyonlara ait aylık akım verileri Elektrik İşleri Etüt İdaresi (EİE) kayıtlarından elde edilmiştir. Bu veriler yardımı 1402 nolu istasyonun verileri 1401,1412,1413 ve 1414 nolu istasyonlar yardımı ile elde edilmeye çalışılmıştır. Symbiotic Organisms Search (SOS) algoritması yardımı ile optimize edilecek 3 farklı fonksiyon

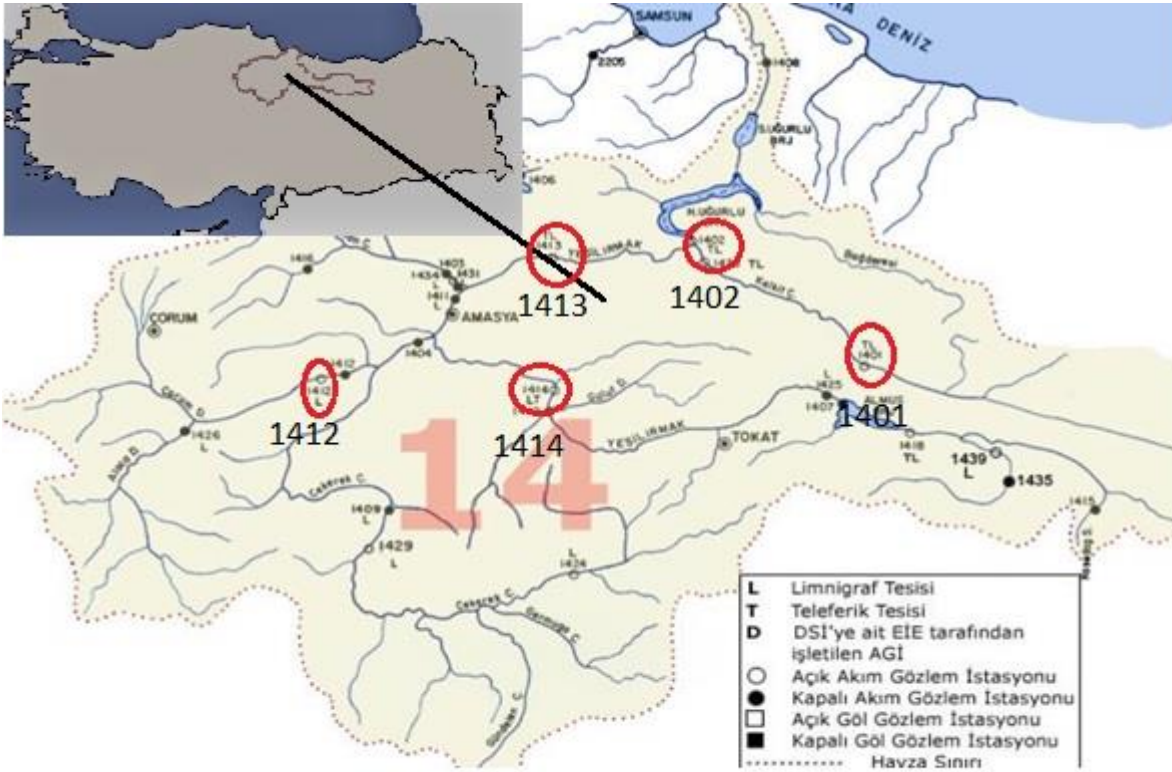
kullanılmıştır. Ayrıca bu fonksiyonların yanı sıra çoklu regresyon ve normal oran metodu ile elde edilen sonuçlar karşılaştırılmıştır.

2. Materyal ve Yöntem

Çalışmanın bu bölümünün ilk kısmında Yeşilirmak nehri ve kullanılan istasyonlar hakkında bilgi ve istatistiksel değerlendirme yapılacaktır. İkinci kısmında ise çalışmada kullanılan klasik yöntem, çoklu regresyon metodu ve Symbiotic Organisms Search ile ilgili bilgiler verilecektir. Ayrıca Symbiotic Organisms Search algoritmasında kullanılan fonksiyonlar gösterilecektir.

2.1. Çalışma Alanı ve Veriler

Yeşilirmak havzası $39^{\circ} 30'$ ile $41^{\circ} 21'$ enlemleri ve $34^{\circ} 40'$ ile $39^{\circ} 48'$ boylamları arasında bulunan Türkiye'deki 25 havzadan biridir (Şekil 1). Adını Yeşilirmak Nehri'nden alan havzadaki ana nehir kanalının uzunluğu 519 km'dir. Yeşilirmak Nehri'nin başlıca kolları Kelkit, Çekerek, Çorum Çat ve Tersakan derelelidir. Yaklaşık 3 873 280 ha olan havza, Türkiye'deki üçüncü en büyük havzadır (Kuruç, Yürekli ve Öztürk, 2005).



Şekil 1. Yeşilirmak havzası yer bulunduru haritası (E.İ.E., 2003)

Çalışmada Yeşilirmak nehrine ait 1401, 1402, 1412, 1413 ve 1414 nolu istasyonlar kullanılmıştır. İstasyonlara ait tanımlayıcı istatistik bilgileri Tablo 1'de bu istasyonların arasındaki akım verilerinin birbiri ile olan Spearman-Rank korelasyon katsayıları ise Tablo 2'de özetlenmiştir. E.İ.E (Elektrik İşleri Etüd İdaresi) tarafından açılan istasyonlardan 5 tanesi bu çalışmada kullanılmak üzere seçilmiştir. Tüm istasyonlardaki ölçümler için ortak dönem olan 1957-2012 su yılları arası aylık ortalama debi değerleri kullanılmıştır. Bu yıllardan 1957-2000 yılları arasındaki toplam 504 adet veri eğitim için geri kalan yıllara ait 132 veri ise test aşaması için ayrılmıştır. Eğitim ve test aşamasına ait istatistiksel özellikler Tablo 1, Tablo 2 ve Tablo 3'te verilmiştir.

Tablo 1
İstasyon ait eğitim verilerinin tanımlayıcı istatistiği

	1401	1412	1413	1414	1402
Ortalama	70.66	7.35	63.90	25.69	151.33
Standart Hata	3.42	0.38	2.38	0.87	5.58
Standart Sapma	76.81	8.55	53.52	19.53	125.23
Varyans	5899.65	73.09	2864.77	381.54	15682.93
Basıklık	3.46	5.31	3.13	5.28	2.54
Çarpıklık	1.90	2.12	1.64	1.88	1.54
Aralık	443.77	59.08	347.53	136.54	751.50
En Büyük	5.23	0.02	2.47	2.46	13.50
En Küçük	449.00	59.10	350.00	139.00	765.00
Toplam	35611.66	3705.22	32204.54	12946.99	76271.80
Sayı	504	504	504	504	504
Güvenirlilik Düzeyi (%95)	6.72	0.75	4.68	1.71	10.96

Tablo 2
İstasyon ait test verilerinin tanımlayıcı istatistiği

	1401	1412	1413	1414	1402
Ortalama	60.25	3.49	38.02	14.74	125.04
Standart Hata	3.59	0.36	2.98	0.76	6.64
Standart Sapma	41.26	4.09	34.29	8.68	76.26
Örnek Varyans	1702.67	16.75	1175.48	75.31	5815.13
Basıklık	2.09	5.87	2.97	3.69	0.92
Çarpıklık	1.44	2.16	1.88	1.85	1.34
Aralık	200.80	23.90	156.44	45.10	317.70
En Küçük	12.20	0.00	4.56	4.60	42.30
En Büyük	213.00	23.90	161.00	49.70	360.00
Toplam	7952.70	460.43	5018.68	1945.62	16505.00
Sayı	132	132	132	132	132
Güvenirlilik Düzeyi	7.10	0.70	5.90	1.49	13.13

Tablo 3
İstasyonlardan alınan eğitim verilerinin arasındaki debiler arası korelasyon katsayıları

	1401	1412	1413	1414	1402
1401	1.000				
1412	0.532	1.000			
1413	0.650	0.871	1.000		
1414	0.320	0.322	0.400	1.000	
1402	0.897	0.766	0.882	0.338	1.000

2.1.1. Normal Oran Metodu

Bu yöntemde her bir veri setinin yıllık ortalama değeri alınır. Verisi tamamlanacak istasyona ait veriler, bu istasyona ait ortalama debinin diğer istasyonlara oranı ile elde edilen katsayılar yardımı ile hesaplanır. Bu katsayılar ile istasyonların verileri ile çarpılır ve toplamları mevcut istasyon adedine bölünür (Ismail ve Zin, 2017).

$$Q_e = (Q_1 * \frac{Q_{e(ort)}}{Q_{1(ort)}} + Q_2 * \frac{Q_{e(ort)}}{Q_{2(ort)}} + \dots + Q_n * \frac{Q_{e(ort)}}{Q_{n(ort)}}) / n \quad (2.1)$$

Burada Q debi miktarını n ise tahmin için kullanılan istasyon adedini ifade eder.

2.1.2. Çoklu Regresyon Analizi

Birbirleriyle sebep-sonuç ilişkisi içinde bulunan iki veya daha fazla değişken arasındaki ilişkiyi çözümlenmek ve bu ilişkiyi matematiksel anlamda modellemek için yapılan istatistiksel analize regresyon analizi denir. Bu analiz sonucunda elde edilen modelde bağımlı değişken bağımsız değişkenin bir fonksiyonu olarak ifade edilerek regresyon denklemi oluşturulur. Regresyon denklemi değişkenler arasındaki bağımsız değişkenlerde meydana gelecek bir birimlik artışın bağımlı değişkende getireceği değişimi ve kurulan fonksiyon biçimi hakkında bilgi verir. Regresyon analizi için iki değişken inceleniyorsa basit regresyon analizi, daha fazla değişken inceleniyorsa çoklu regresyon analizi olarak adlandırılır (Sun ve Trevor, 2018).

$$Y = a_1X_1 + a_2X_2 + \dots + a_kX_k + U \quad (2.2)$$

Burada Y bağımlı (tahmin edilen) değişkeni, X bağımsız (açıklayıcı) değişkeni, a kısmi regresyon katsayılarını k girdi parametresi sayısını ve U hata terimini ifade etmektedir.

Çoklu doğrusal regresyon analizinin uygulanabilmesi için verilerin normal dağılım sergilemesi, bağımsız değişkenlerle bağımlı değişken arasındaki ilişkinin doğrusal olması ve her bir bağımsız değişken için hataların varyansının sabit olması gereklidir (Hair, Black, Babin ve Anderson, 2009).

2.1.3. Simbiyotik Organizmalar Arama Algoritması

Bir probleme en uygun çözümü bulmak için son yıllarda sezgisel yöntemler kullanılmaktadır. Sezgisel yöntemler iyiye çok yakın çözümler sunan, çözüm süresini oldukça kısaltan yöntemlerdir. Gerçek hayatta her gün kullandığımız bu yaklaşımlar genel olarak doğadaki olaylardan esinlenir (Özdemir, Aydemir, Olgun ve Mulbay, 2016; Çatal ve Saplıoğlu, 2018; Uzundurukan ve Saplıoğlu, 2019).

Her ne kadar yalnız yaşayan organizma türleri olsa da güvenlik ve yardım gibi konular nedeniyle organizmalar çoğunlukla ortak bir yaşamı paylaşırlar. Bazı durumlarda zorunlu bir ortak yaşam olabildiği gibi, organizmaların çıkarları için kendileri de ortak bir yaşamı tercih etmektedirler. Organizmalar arasındaki bu birlikte yaşam üç farklı şekilde gerçekleşebilmektedir. Bunlar; karşılıklı fayda birlikteliği (mutualizm), tek taraflı birliktelik (kommensalizm) ve asalaklık (parazitizm)'dir.

Simbiyotik Organizmalar Arama algoritması (SOS), Cheng ve Prayogo tarafından 2014 yılında geliştirilmiş güçlü bir sezgi ötesi optimizasyon algoritmasıdır (Cheng ve Prayogo, 2014). SOS, doğada yaşayan organizmaların hayatta kalmak için birbirleri ile olan etkileşimlerinden esinlenilerek geliştirilmiş ve literatürde farklı problemlerin çözümünde kullanılmıştır (Baysal ve Altas, 2017; Çelik ve Öztürk, 2017). SOS başlangıçta bir grup organizmanın rasgele oluşturulduğu bir ekosistem ile başlatılmaktadır. Her bir organizma aday çözümleri temsil etmekte ve uygunluk değerleri adaptasyon derecelerini ifade etmektedir. Başka diğer optimizasyon algoritmalarında da olduğu gibi, SOS algoritmasında da organizmaların yaşamlarını sürdürebilmesi için biyolojik etkileşim adımları bulunmaktadır. SOS algoritması mutualizm, kommensalizm ve parazitizm adımları olmak üzere üç adımdan oluşmaktadır.

Mutualizm adımımda, birliktelik kuran organizmaların her ikisinin de karşılıklı olarak fayda elde ettikleri adımdır. Ekosistemdeki i . organizma (X_i) ile ekosistemden rasgele seçilen j . organizma (X_j)'nın fayda sağlayacakları bir ilişkinin kurulduğu aşamadır. Bu aşamada X_i ve X_j organizmaları için yeni aday çözümler üretilmektedir. Denklem 2.3, 2.4 ve 2.5'te gösterildiği şekilde aday çözümler oluşturulmaktadır (Çelik ve Öztürk, 2017).

$$X'_i = X_i + rand(0,1) * (X_{eniyi} - MV * BF_1) \quad (2.3)$$

$$X'_j = X_j + rand(0,1) * (X_{eniyi} - MV * BF_2) \quad (2.4)$$

$$MV = \frac{X_i + X_j}{2} \quad (2.5)$$

Burada, BF_1 ve BF_2 katsayıları sırasıyla i . ve j . bireylerin ilişkiden elde edeceği fayda faktörünü ifade etmektedir. BF_1 ve BF_2 katsayıları 1 veya 2 olarak rasgele seçilmektedir. MV değeri mutual_ vektör olarak adlandırılmakta ve organizmalar arasındaki ilişki karakteristiğini ifade etmektedir. Mutualizm adımımda üretilen yeni aday bireylerinin (X'_i ve X'_j) uygunluk değerleri eğer eski bireylerin (X_i ve X_j) uygunluk değerlerinden daha iyi ise, yeni aday bireyler ile eski bireyler yer değiştirmektedir. X_{eniyi} değeri ise en iyi adaptasyon sağlamış olan bireyi temsil etmektedir.

SOS algoritmasının kommensalizm adımımda, ekosistemdeki organizmalardan biri fayda sağlarken diğeri herhangi bir fayda veya zarar sağlamamaktadır. Ekosistemdeki bir X_i organizmasına fayda sağlayacak X_j organizması rasgele olarak seçilmektedir. Denklem 2.6'da gösterildiği gibi X_i organizmasının yeni aday çözümü hesaplanmaktadır.

$$X'_i = X_i + rand(-1,1) * (X_{eniyi} - X_j) \quad (2.6)$$

Denklem 2.4'ten görüldüğü üzere X_j organizması hiçbir şekilde etkilenmemektedir. X'_i yeni aday organizmanın uygunluk değeri daha iyi olursa, mevcut X_i organizması yeni aday organizma ile değiştirilerek bu adım gerçekleştirilir.

Simbiyotik Organizmalar Arama Algoritması

```

1: Başlangıç ekosistemini rasgele olarak belirle
2: while Durdurma kriteri? do
3:   for  $i \leftarrow 1$  to  $eko_{boyut}$  do
4:     for  $i \leftarrow 1$  to  $eko_{boyut}$  do
5:        $f(X_i)$   $\triangleright$  Organizmaların uygunluk değerlerini hesapla
6:     end for
7:      $X_{eniye}$ 'yi belirle
8:     //Mutualizm aşaması
9:      $X_i$  organizmasından farklı  $X_j$  organizmasını rasgele seç
10:     $[X'_i, X'_j] = mutualizm(X_i, X_j, X_{eniye})$   $\triangleright$  Adayları oluştur
11:    if  $f(X'_i) < f(X_i)$  then
12:       $X_i = X'_i$ 
13:    end if
14:    if  $f(X'_j) < f(X_j)$  then
15:       $X_j = X'_j$ 
16:    end if
17:    //Kommensalizm aşaması
18:     $X_i$  organizmasından farklı  $X_j$  organizmasını rasgele seç
19:     $X'_i = kommensalizm(X_i, X_j, X_{eniye})$   $\triangleright$  Adayı oluştur
20:    if  $f(X'_i) < f(X_i)$  then
21:       $X_i = X'_i$ 
22:    end if
23:    //Parazitizm aşaması
24:     $X_i$  organizmasından farklı  $X_j$  organizmasını rasgele seç
25:     $X_i$  organizmasının parazitini ( $P_{X_i}$ ) oluştur
26:    if  $f(P_{X_i}) < f(X_j)$  then
27:       $X_j = P_{X_i}$ 
28:    end if
29:  end for
30: end while

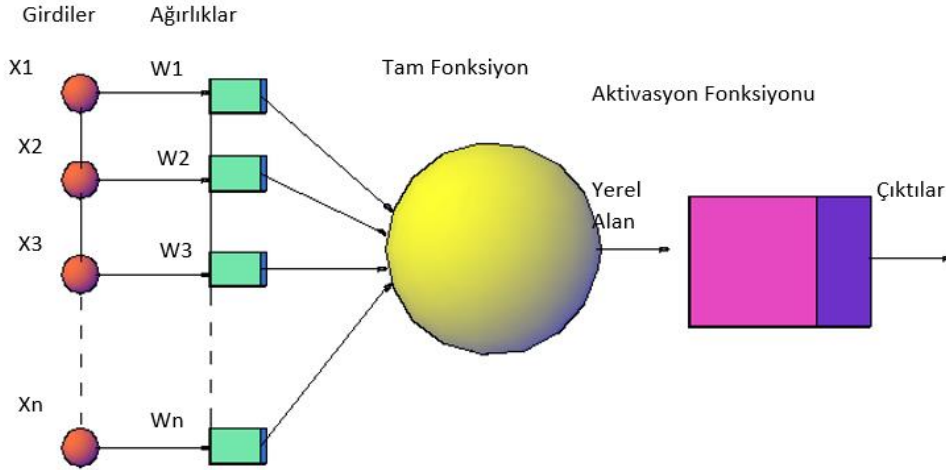
```

Şekil 2. Simbiyotik organizmalar arama algoritması

Parazitizm aşamasında bir organizma diğer organizma tarafından zarara uğratılmaktadır. X_i organizmasının bazı parametrelerinde rasgele değişiklikler yapılarak X_i organizmasının paraziti (P_{X_i}) oluşturulmaktadır. Daha sonra ekosistemden zarara uğratılacak olan organizma (X_j) rasgele olarak seçilmektedir. P_{X_i} ve X_j organizmalarının uygunluk değerleri hesaplanarak karşılaştırılır. Eğer P_{X_i} organizmasının uygunluk değeri daha iyi ise, X_j organizması ekosistemden çıkarılarak yerine P_{X_i} organizması yerleştirilir. Böylece X_j organizması bu birliktelikten ekosistemden ayrılarak zarar etmiş olmaktadır (Şekil 2).

2.1.4. Yapay Sinir Ağları

Yapay Sinir Ağı (YSA) canlı yapılarından ilham alınan biyolojik sinir ağlarından etkilenmiş programdır (Avcar ve Saplıoğlu, 2015). YSA kendisine verilen veri setindeki girdi ve çıktılar arasında bir ilişki kurmaya çalışan bu işlemi de bir takım öğrenme algoritmaları ile yürütmeye çalışan bir modeldir. En büyük dezavantajı verilen verileri ezberleyerek verilere göre ağırlık atmasıdır. Böyle bir durumda gerçek öğrenme olup olmadığını anlayabilmek için verilerin en az % 20 lik kısmı test için ayrılır ve hiçbir şekilde eğitime katılmaz. Test verilerinin de iyi sonuç vermesi modelin doğruluğu hakkında bir fikir verebilir.



Şekil 3. YSA model örneği (Avcar ve Saplıoğlu, 2015)

Birçok sinir ağı modeli vardır. Bazıları diğerlerinden daha fazla tercih edilmektedir ve en çok tercih edileni ise Geri Yayılım Modelidir. Bu modelin tahmin ve sınıflandırma süreçlerinde iyi sonuçlar verdiği tecrübe edilmiştir (Elmas, 2003).

Geril Yayılımlı Sinir Ağı lineer olmayan ve çok karmaşık yapıdaki problemleri çözümedeki üstünlüğünden ve kabiliyetinden dolayı çokça tercih edilen modeldir. Geril Yayılımlı Sinir Ağı, Geril Besleme Algoritmaları tarafından eğitilmiş çok katmanlı ve ileri beslemeli sinir ağıdır (Chandwani, Agrawal ve Nagar, 2015). Bu model girdileri ve sonuçları tekrar tekrar işler ve bu operasyonu kullanarak, model en küçük kareler hatalarını en aza indirmeye çalışır.

3. Bulgular ve Tartışma

Çalışmada Yeşilirmak havzasında yer alan 5 adet istasyondan yararlanılmıştır. 1957-2012 su yılları arasındaki aylık akım verilerinden 504 tanesi eğitim 132 tanesi ise test aşamasında kullanılmıştır. 5 istasyon içerisinde en mansapta bulunan 1402 nolu istasyonun verileri, diğer 4 istasyonun verileri kullanılarak tahmin edilmeye çalışılmıştır. Çalışmada 3 farklı fonksiyonun SOS ile optimizasyonu sağlanarak bu fonksiyonlara ait katsayılar belirlenmiştir. Ayrıca bu 3 fonksiyonun yanı sıra modelleri, çoklu regresyon ve normal oran metodu da kullanılarak tahmin modelleri oluşturulmuş ve SOS ile optimize edilen fonksiyonlarla kıyaslanmıştır.

Normal oran metodunda kullanılan (Denklem 1) katsayılar istasyonlara ait uzun dönem ortalamaları dikkate alınarak oluşturulmaktadır. Bu çalışmada bu katsayıların uzun dönem ortalamaları yerine sezgisel optimizasyon tekniklerinden olan SOS ile yapılması düşünülmüştür. Optimize edilecek fonksiyonlar oluşturulurken normal oran metodunun oluşturulma mantığı temel alınmıştır (Denklem 3.1). Temel alınan bu birinci fonksiyonu iyileştirebilmek içinse taban akışını temsilen bir sabit sayı (Denklem 3.2) ve fonksiyonun sonucunu (Denklem 3.3) oluşturulmuştur. Oluşturulan bu denklemler eğitim verileri yardımı ile SOS kullanılarak optimize edilmiştir.

$$Q_{1402} = X_1 * Q_{1401} + X_2 * Q_{1412} + X_3 * Q_{1413} + X_4 * Q_{1414} \quad (3.1)$$

$$Q_{1402} = X_1 * Q_{1401} + X_2 * Q_{1412} + X_3 * Q_{1413} + X_4 * Q_{1414} + C \quad (3.2)$$

$$Q_{1402} = (X_1 * Q_{1401} + X_2 * Q_{1412} + X_3 * Q_{1413} + X_4 * Q_{1414} + C)^m \quad (3.3)$$

Çoklu regresyon (Denklem 2) ve Normal oran metodu (Denklem 1) yöntemleri için elde edilen denklemler de Denklem 3.4 ve Denklem 3.5 de gösterilmiştir.

$$Q_{1402} = 0.93 * Q_{1401} + 0.941 * Q_{1412} + 1.119 * Q_{1413} + 0.363 * Q_{1414} + 16.521 \quad (3.4)$$

$$Q_{1402} = 2.14 * Q_{1401} + 20.59 * Q_{1412} + 2.37 * Q_{1413} + 25.68 * Q_{1414} \quad (3.5)$$

Denklem 3.1, 3.2 ve 3.3 'un SOS yöntemiyle optimize edilmesi ile elde edilen katsayılar Tablo 4' de gösterilmiştir. Burada dikkat edilmesi gereken bir nokta ise Tablo 3'de verilen korelasyonu en düşük olan 1414 nolu istasyonu SOS ile yöntem dikkate almamış ve katsayısını 0 olarak tespit etmiştir.

Tablo 4
SOS ile optimize edilmiş denklem katsayıları

	X_1	X_2	X_3	X_4	C	m
Denklem 3.1	0.9855	0.1914	1.2024	0		
Denklem 3.2	0.9825	0.2632	1.7712	0	1.302	
Denklem 3.3	1.2435	0.5067	1.4027	0	0	0.9632

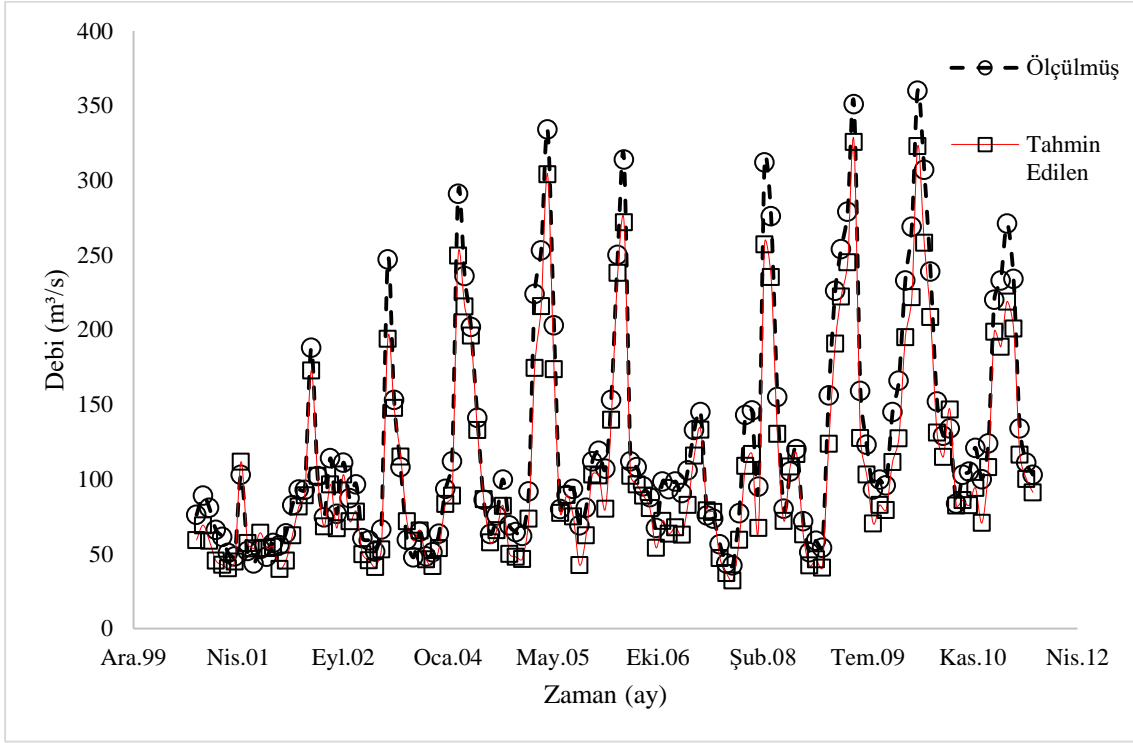
Elde edilen sonuçlar ve bu sonuçlara göre elde edilmiş değerlerin hem eğitim hem de test verileri için R^2 değerleri ile ortalama karesel hataları Tablo 5' de verilmiştir. Bu sonuçlara göre SOS ile oluşturulmuş tüm fonksiyonlar ile yapılan tahminlerin diğer iki yöntemle göre çok daha iyi olduğu görülmüştür. Optimize edilen bu üç denklemden ise Denklem 3.3 ile yapılan tahminlerin diğer optimize edilen denklemlere göre çok farklı olmadığı ancak diğerlerine oranla %0.5 ila %0.7 arasında bir doğruluk sağlamıştır.

Tablo 5
Elde edilen sonuçların karşılaştırılması

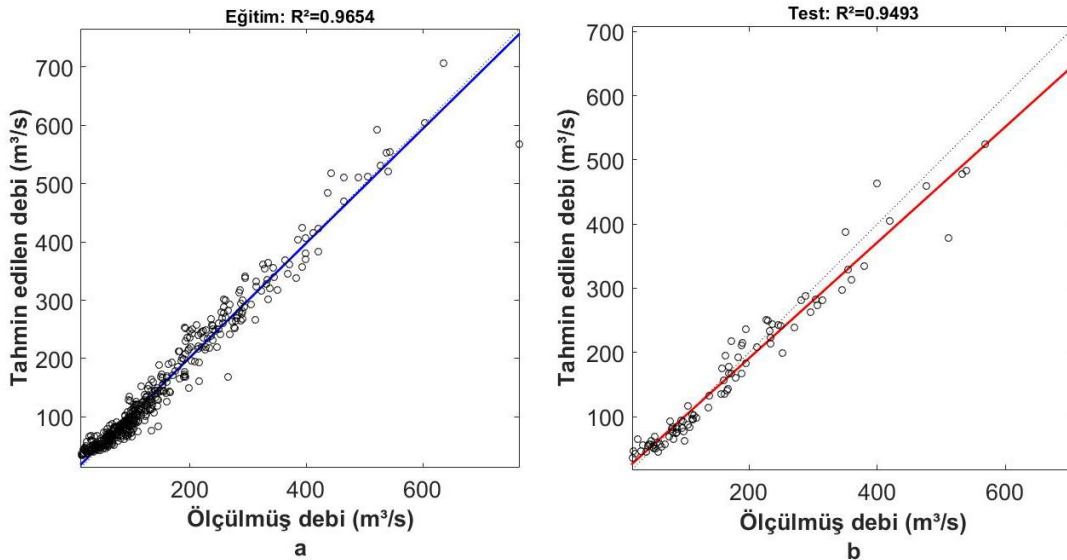
	Eğitim			Test		
	R^2	Adj R^2	OKH	R^2	Adj R^2	OKH
Denklem 3.1	0.9601	0.9598	0.043	0.9773	0.9766	0.0388
Denklem 3.2	1.0000	1.000	0.0461	1.0000	1.000	0.0364
Denklem 3.3	0.9997	0.9997	0.0360	0.9996	0.9996	0.0311
5 N. YSA	0.9435	0.9433	0.0481	0.9332	0.9328	0.0492
6 N. YSA	0.9502	0.9498	0.04752	0.9435	0.9432	0.05221
7 N. YSA	0.9632	0.9625	0.04685	0.9598	0.9593	0.05132
8 N. YSA	0.9654	0.9651	0.04621	0.9641	0.9637	0.04728
9 N. YSA	0.9621	0.9617	0.04681	0.9603	0.9599	0.04792
10 N. YSA	0.9618	0.9614	0.04793	0.9594	0.9591	0.05345
Çoklu Reg.	0.8809	0.8923	0.0566	0.9212	0.9058	0.0566
N.Oran Met.	0.8932	0.8800	0.01035	0.9086	0.9187	0.0750

Denklem 3.3 ile elde edilen debi gidiş çizgisi ile ölçülen debi gidiş çizgisi kıyaslandığında özellikle diğer yöntemlerde hataların yoğunlaştığı maksimum debilerde hata oranının oldukça az olduğu görülmektedir (Şekil 4). Yapılan çalışmada oluşturulan 6 farklı nöron sayısına sahip YSA modelleri (5,6,7,8,9,10 nöronlu) ile elde edilen sonuçlar incelendiğinde tüm YSA modellerinin çoklu regresyon modelinden ve normal oran

metodundan daha iyi sonuçlar verdiği görülmüştür (Tablo 5). Ancak tüm YSA modellerinin de Denklem 3.2 ve Denklem 3.3 den elde edilen sonuçlar kadar başarılı olmadığı da tespit edilmiştir. YSA modelleri içerisinde en verimli modelin ise 8 nöronlu model olduğu söylenebilir (Şekil 5). Ayrıca hem eğitim hem test verilerinin P değerleri 0 olarak tespit edilmiştir. Ayrıca Mallows Cp değeri tüm modellerde 2 çıkmıştır. Karşılaştırmalar 2 parametre ile yapıldığından sonuçların kabul edilebilir olduğu düşünülmüştür.



Şekil 4. Test verileri debi gidiş çizgisi ile Denklem 9'dan elde edilen debi gidiş çizgisinin karşılaştırılması



Şekil 5. a) 8 nöronlu YSA modeli eğitim verileri saçılım diyagramı b) 8 nöronlu YSA modeli test verileri saçılım diyagramı

4. Sonuçlar

Küresel ısınma ve nüfus artışı sebebiyle suya olan talebin artmasından dolayı proje planlama çalışmalarının optimum şekilde yapılması gerekmektedir. Optimum bir analiz yapabilmek ise doğru ve güvenilir veri ile mümkün olabilmektedir. Ancak bazen fiziki ve maddi şartlar verilerin eksik alınmasına

veya hiç alınamamasına sebep olmaktadır. Bu çalışma, eksik olan akım verilerinin ölçüm yapılan diğer istasyonlar yardımı ile tamamlanması üzerine yapılmıştır. Bu amaçla Yeşilirmak nehri üzerinde mevcut olan 5 adet istasyondan faydalanılmıştır. Bu istasyonlardan bir tanesinin verileri diğer 4 istasyondan alınan veriler yardımı ile tahmin edilmeye çalışılmıştır. Bunun için literatürde mevcut çoklu regresyon ve normal oran metodunun yanı sıra çalışmada 3 farklı fonksiyon kullanılmış ve bu fonksiyonların katsayılarının tespitinde ise sezgisel bir optimizasyon tekniği olan SOS algoritmasından faydalanılmıştır. Sonuçlar incelendiğinde, SOS algoritması yardımı ile optimize edilen 3 fonksiyondan elde edilen sonuçların tamamının 6 farklı YSA modelinin yanı sıra hem çoklu regresyon hem de normal oran metoduna göre çok daha iyi olduğu görülmüştür. Ayrıca bu 3 fonksiyon irdelendiğinde düzeltme katsayılı fonksiyonun diğerlerine oranla %0.1-%0.7 arası daha iyi sonuçlar verdiği tespit edilmiştir. Ancak çalışmada kullanılan 3 fonksiyonun tamamının hata değerleri %3-%4 aralığındadır. Bu nedenle bu fonksiyonlar ve bu fonksiyonları optimize etmek için kullanılan SOS algoritmasının eksik akım verilerinin tahmininde kullanılabileceği, hatta yeni açılmış istasyonların geçmişe yönelik verilerinin tamamlanmasında da kullanılabileceği düşünülmüştür. Çalışmada elde edilen sonuçlar göstermektedir ki; sezgisel optimizasyon tekniklerinden olan SOS algoritması bu fonksiyonların optimize edilmesi için iyi sonuçlar vermesine karşın en iyi sonucu verecek başka optimizasyon tekniklerinin olabileceği de unutulmamalıdır. Bu nedenle çalışmanın diğer optimizasyon teknikleri ile de denenmesinin yararlı olabileceği düşünülmektedir.

Yazar Katkıları

Kemal SAPLIOĞLU: Analiz planlamış, matlab yazılımı ve istatistiksel karşılaştırma yapmıştır.

Tülay Suğra KÜÇÜKERDEM ÖZTÜRK: Literatür taramış ve YSA modelini kurmuştur.

Fatih Ahmet ŞENEL: Optimizasyon modelini kurmuştur.

Çıkar Çatışması

Yazarlar arasında herhangi bir çıkar çatışması yoktur.

Kaynaklar

- Avcar, M. ve Saphioğlu, K. (2015). An artificial neural network application for estimation of naturel frequencies of beams. *International Journal of Advanced Computer Science and Applications*, 6(6), 94-102.
- Bakış, R. ve Göncü, S. (2015). Akarsu debi ölçümlerinde eksik verilerin tamamlanması: Zap suyu havzası örneği. *Anadolu Üniversitesi Bilim ve Teknolojisi Dergisi A- Uygulamalı Bilimler ve Mühendislik*, 16(1):63-79.
- Bardossy, A. ve Pegram, G. (2014). Infilling missing precipitation records- A comparison of a new copula-based method with other techniques. *Journal of Hydrology*, (519), 1162-1170.
- Bayazıt, M. (2013). *Hidroloji*. İstanbul: Birsen Yayınevi.
- Baysal, YA. ve Altas, IH. (2017). Power quality improvement via optimal capacitor placement in electrical distribution systems using symbiotic organisms search algorithm. *Mugla Journal of Science and Technology*, 3(1) ,64-68.
- Beauchamp, JJ. Downing, DJ. Ve Railsback, SF. (1989). Comparison of regression and time-series methods for synthesizing missing streamflow records. *Journal of the American Water Resources Association*, 25(5), 961-975.
- Chandwani, V. Agrawal, V. ve Nagar, R. (2015). Modeling slump of ready mix concrete using genetic algorithms assisted training of artificial neural networks. *Expert Systems with Applications*, 42(2), 885-893.
- Cheng, M. ve Prayogo, D. (2014). Symbiotic Organisms Search: A new metaheuristic optimization Algorithm. *Computers and Structures*, (139), 98-112.
- Çatal, Y. ve Saphioğlu, K. (2018). Comparison of adaptive neuro-fuzzy inference system, artificial neural networks and non-linear regression for bark volume estimation in Brutian Pine (*Pinus brutia ten.*) *Applied Ecology and Environmental Research*, 16(2), 2015-2027.
- Çelik, E. ve Öztürk, N. (2017). Doğru akım motor sürücülerini için PI parametrelerinin simbiyotik

- organizmalar arama algoritması ile optimal ayarı. *Bilişim Teknolojileri Dergisi*, 10(3), 311-318.
- E.İ.E. (2003). Elektrik İşleri Etüd İdaresi
- Elmas, Ç. (2003). *Yapay sinir ağları (Kuram, mimari, eğitim, uygulama)*. Ankara: Seçkin Yayınevi.
- Elshorbagy, A. Simonovic, SP. ve Panu, US. (2002). Estimation of missing streamflow data using principles of chaos theory. *Journal of Hydrology*, (255),123-133.
- Gümüş, V. ve Kavşut, ME, (2013). Zamanti Nehri-Ergenüşağı istasyonu eksik aylık akım verilerinin tahmini. *Gazi Üniversitesi Fen Bilimleri Dergisi Part:C*, 1(2), 81-91.
- Hair, JF. Black, WC. Babin, BJ. ve Anderson, RE. (2009). *Multivariate data analysis*. Pearson Education.
- Ismail, WNW. ve Zin, WZW. (2017). Estimation of rainfall and stream flow missing data for Terengganu, Malaysia by using interpolation technique methods. *Malaysian Journal of Fundamental and Applied Sciences*, 13(3), 213-217.
- Kuruç, A. Yürekli, K. ve Öztürk, F. (2005). Effect of discharge fluctuation on water quality variables from the Yeşilirmak river. *Tarım Bilimleri Dergisi*, 11(2), 189-195.
- Mwale, FD. Adeloye, AJ. ve Rustum, R. (2012). Infilling of missing rainfall and streamflow data in the Shire river basin, Malawi- A self organizing map approach. *Physics and Chemistry of the Earth, Parts A/B/C*, 50(52), 34-43.
- Özdemir, G. Aydemir, E. Olgun, MO. ve Mulbay, Z. (2016). Forecasting of Turkey natural gas demand using a hybrid algorithm. *Energy Sources Part B- Economics Planning and Policy*, 11(4), 295-302.
- Önöz, B. (2015). Ölçüm olmayan havzalarda akım tahmin yöntemleri, <http://www.dsi.gov.tr/docs/tuhk/ölçüm-olmayan-havzalarda-akım-tahmin-yöntemleri.pdf?sfvrsn=2> (Erişim Tarihi:23.05.2019).
- Partal, T. (2007). *Türkiye yağış miktarlarının yapay sinir ağları ve dalgacık dönüşümü yöntemleri ile tahmini* (Yayımlanmış doktora tezi). İstanbul Teknik Üniversitesi Fen Bilimleri Enstitüsü, İstanbul, Türkiye.
- Salas, JD. (1993). *Analysis and modeling of hydrologic time series*. New York: McGraw-Hill Book Company.
- Sun, W. ve Trevor, B. (2017). Multiple model combination methods for annual maximum water level prediction during River Ice breakup. *Hydrological Processes*, 32, 421-435.
- Şen, Z. (2004). *Yapay sinir ağı ilkeleri*. İstanbul:Su Vakfı Yayınları .
- Tencaliec, P. Favre, A. Prieur, C. ve Mathevet, T. (2015). Reconstruction of missing daily streamflow data using dynamic regression models. *Water Resources Research*, 51, 9447-9463.
- Tosunoğlu, F. İspirli, MN. Gürbüz, F. ve Şengül, S. (2017). Fırat Havzası'ndaki eksik akım verilerinin debi süreklilik çizgileri ve regresyon modelleri ile tahmin edilmesi. *Iğdır Üniversitesi Fen Bilimleri Enstitüsü Dergisi*, 7(4), 85-94.
- Turhan, E. ve Çağatay, HÖ. (2016). Eksik akım verilerinin tahmin modelinin oluşturulmasında yapay sinir ağlarının kullanımı: Asi Nehri-Demirköprü akım gözlem istasyonu örneği. *Çukurova Üniversitesi Mühendislik Mimarlık Fakültesi Dergisi*, 31(1), 93-106.
- Uzundurukan, S. Saphoğlu, K. (2019). Bilimsel çalışmalarda kullanılan yapay zekâ uygulamalarının ve trendlerinin incelenmesi. *Dicle Üniversitesi Mühendislik Fakültesi Mühendislik Dergisi*, 10(1), 249-262.
- Yıldıran, A. Yerel Kandemir, S. (2018). Yağış miktarının yapay sinir ağları ile tahmini. *Bilecik Şeyh Edebali Üniversitesi Fen Bilimleri Dergisi*, 5(2), 97-104.
- Yurdusev, MA. Acı, M. Turan, ME. ve İçağa, Y. (2008). Akarçay nehri aylık akımlarının yapay sinir ağları ile tahmini. *Celal Bayar Üniversitesi Fen Bilimleri Dergisi*, 4(1), 73-88.