

# AKARSULARDA ÇÖZÜNMÜŞ OKSİJEN KONSANTRASYONUNUN REGRESYON TABANLI YÖNTEMLERLE MODELLENMESİ: HARŞİT ÇAYI ÖRNEĞİ

*Sinan NACAR* \*<sup>id</sup>  
*Osman Tuğrul BAKİ* \*\*<sup>id</sup>  
*Adem BAYRAM* \*\*\*<sup>id</sup>

Alınma: 10.06.2021; düzeltme: 29.07.2021; kabul: 23.02.2022

**Öz:** Bu çalışmada çok değişkenli uyarlanabilir regresyon eğrileri (MARS) ve TreeNet gradyan artırma makinesi (TreeNet) isimli regresyon tabanlı yöntemler kullanılarak çözünmüş oksijen (ÇO) konsantrasyonu modellemesi amaçlanmıştır. Modelleme çalışmasında kentsel atıkları bünyesine alarak yer yer kirlenen Harşit Çayı (Gümüşhane) üzerinde belirlenmiş altı su kalitesi gözlem istasyonunda, 15 gün aralıklarla ve 24 kez yerinde gerçekleştirilen ÇO konsantrasyonu (mg/L), sıcaklık (°C), pH ve elektriksel iletkenlik (mS/cm) ölçümleri yanı sıra akarsudan alınan su örneklerinde laboratuvarında gerçekleştirilen sertlik (°dH) tayinleri neticesinde elde edilen veriler kullanılmıştır. Elde edilen veri setinin % 80'i kurulan modellerin eğitilmesinde geriye kalan % 20'si ise söz konusu modellerin test edilmesinde kullanılmıştır. Kurulan modellerin eğitim ve test veri seti performanslarını değerlendirmek amacıyla ortalama karesel hatanın karekökü (OKHK), ortalama mutlak hata (OMH), ortalama rölatif hata (ORH) ve determinasyon katsayısı ( $R^2$ ) performans istatistikleri kullanılmıştır. En düşük OKHK, OMH ve ORH ile en yüksek  $R^2$  değerleri eğitim veri seti için sırasıyla 0,2247 mg/L, 0,0666 mg/L, % 0,66 ve 0,9995 olarak TreeNet yönteminden, test veri seti için ise 0,2911 mg/L, 0,2336 mg/L, % 2,27 ve 0,9992 olarak MARS yönteminden elde edilmiştir. Her iki veri seti için ortalamalar dikkate alındığında ise, MARS yönteminden elde edilen performans değerlerinin TreeNet yönteminden elde edilenlere kıyasla daha iyi olduğu sonucuna ulaşılmıştır.

**Anahtar Kelimeler:** Çözünmüş oksijen konsantrasyonu, MARS yöntemi, Modelleme, Regresyon analizi, TreeNet yöntemi

## Modeling Dissolved Oxygen Concentration using Regression-based Methods: A Case Study from the Harşit Stream, Turkey

**Abstract:** This study aimed to model the stream dissolved oxygen (DO) concentration using the multivariate adaptive regression splines (MARS) and TreeNet gradient boosting machine (TreeNet) methods. The water quality indicators employed for the modeling studies were the stream DO concentration (mg/L), temperature (°C), pH, and electrical conductivity (mS/cm), as well as hardness (°dH). These indicators were measured semimonthly during a year for six monitoring sites selected in untreated wastewater impacted urban stream, namely Harşit, Gümüşhane Province. The stream water quality data for each indicator were 144, 80 % for the training, and the rest for the testing. To evaluate the performance for the training and testing data sets of the models, four performance statistics, i.e., root mean square error (RMSE), mean absolute error (MAE), mean relative error (MRE), and coefficient of determination ( $R^2$ ), were computed. On the one hand, the TreeNet method provided better results for the training data set. On the other hand, the MARS method provided better results for the testing data set. The lowest RMSE, MAE, and MRE and highest  $R^2$  values were calculated as 0.2247 mg/L, 0.0666 mg/L, 0.66

\* Tokat Gaziosmanpaşa Üniversitesi, Mühendislik ve Mimarlık Fakültesi, İnşaat Mühendisliği Bölümü, 60150 Taşçıiftlik/TOKAT

\*\* Karadeniz Teknik Üniversitesi, Of Teknoloji Fakültesi, İnşaat Mühendisliği Bölümü, 61830 Of/TRABZON

\*\*\* Karadeniz Teknik Üniversitesi, Mühendislik Fakültesi, İnşaat Mühendisliği Bölümü, 61080 Ortahisar/TRABZON  
İletişim Yazarı: Sinan Nacar (sinannacar@hotmail.com)

%, and 0.9995 for the training data sets, and 0.2911 mg/L, 0.2336 mg/L, 2.27 %, and 0.9992 %, for the test data sets, respectively. It was concluded that the MARS method had a better performance than to the TreeNet method considering the mean values for both data sets.

**Keywords:** Dissolved oxygen concentration, MARS method, Modeling, Regression analysis, TreeNet method

## 1. GİRİŞ

Çözünmüş oksijen (ÇO), suda bir bileşik içinde bulunmayıp, serbest halde bulunan oksijen seviyesidir. Suda yaşayan bütün canlılar ÇO'ya ihtiyaç duymaktadırlar. Sucul ortamda balıkların yanı sıra balıklar için beslenme kaynağını oluşturan diğer küçük canlılar için de ÇO önemli bir parametredir. Sucul yaşamın temelini oluşturan ÇO, suyun hava ile temas etmesi sonucunda, sıcaklığa da bağlı olarak bir miktar oksijenin suda çözünmesi ile oluşmaktadır. Su içerisinde gerçekleşen fotosentez olayları sonucunda da sucul ortam oksijen bakımından doymuş hale gelmektedir (Yalçın ve Gürü, 2010). ÇO konsantrasyonunun değişiminde etkisi olan birçok parametre bulunmaktadır. Bu parametreler önem/etki sırasına göre; ortam sıcaklığı, ortamda bulunan alg faaliyetleri, suyun tuzluluğu, atmosferik basınç, organik yüklemeler, su sirkülasyon değişiklikleri ve sucul ortamda yaşayan canlıların solunum olayları olarak sıralanabilir (Göksu, 2015). Sucul canlıların temel ihtiyaçlarından biri olan ÇO, Yerüstü Su Kalitesi Yönetmeliği (YSKY, 2012) kapsamında Kıta içi Yerüstü Su Kaynaklarının sınıflandırılmasına (I. Sınıf > 8 mg/L; II. Sınıf: 6 mg/L, III. Sınıf < 6 mg/L) dahil edilen önemli su kalitesi değişkenlerinden biridir.

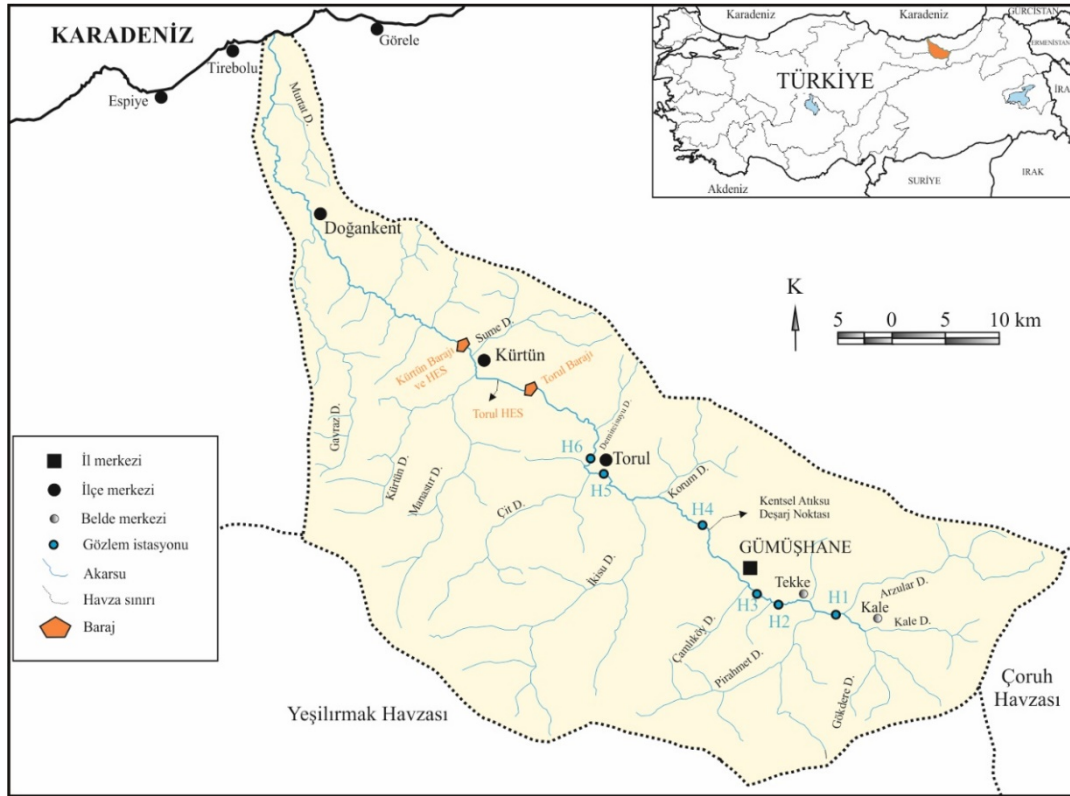
Canlılar üzerinde etkili bir parametre oluşu (Nacar, ve diğ., 2020b) sebebiyle son yıllarda ülkemiz yerüstü su kaynaklarında ÇO takibi araştırmacıların ilgisini çekmiştir (Dede ve diğ., 2013; Kose, ve diğ., 2014; Yuceer ve Coskun, 2016; Sener, ve diğ., 2017; Könez, 2019; Kukrer ve Mutlu, 2019; Küçükler, 2020; Karadeniz, 2020; Ustaoglu, ve diğ., 2020). Literatürde bu anlamda gerçekleştirilen ÇO tahmin çalışmaları irdelendiğinde; geleneksel regresyon analizi (CRA) (Olyaie, ve diğ., 2017), yapay sinir ağları (YSA) (Soyupak ve diğ., 2003; Sengorur ve diğ., 2005; Kisi ve diğ., 2013) ve uyarlamalı bulanık sinir ağı çıkarım sistemi (Altunkaynak ve diğ., 2005; Altuntaş, 2018; Elkiran, ve diğ., 2019) gibi yöntemlerle yapılan çalışmaların yer aldığı görülmektedir. Bu yöntemlerden başka yapay arı kolonileri (ABC) ve öğretim-öğrenme esaslı optimizasyon (TLBO) gibi farklı algoritmalar; çok değişkenli uyarlanabilir regresyon eğrileri (MARS) ve TreeNet gradyan artırma makinesi (TreeNet) gibi farklı regresyon tabanlı yöntemler kullanılarak da ÇO tahmin çalışmaları yapılmıştır (Bayram ve diğ., 2015; Heddami ve Kisi, 2018; Nacar ve diğ., 2020a; Nacar ve diğ., 2020b).

Bu çalışmanın amacı kentsel atıksuların etkisi altında kalan Harşit Çayı'nda (Gümüşhane) ÇO konsantrasyonu değerlerinin MARS ve TreeNet yöntemleri ile modellenmesidir. Yapılmış olan ÇO modelleme çalışmaları incelendiğinde birçok su kalitesi değişkeninin bağımsız değişken olarak kullanıldığı görülmektedir (Nacar ve diğ., 2020a; Nacar ve diğ., 2020b; Nacar ve diğ., 2020c). Bu çalışmada ÇO modellemesinde yaygın olarak kullanılan; su sıcaklığı (T), pH, elektriksel iletkenlik (Eİ) ve su sertliği (SS) değişkenleri bağımsız değişkenler olarak seçilmiştir. Kullanılan ÇO, T, pH, Eİ ve SS verileri, içinden geçtiği yerleşim birimlerinin atıksularını alarak yer yer kirlenen Harşit Çayı'nın Bayram (2011) tarafından su kalitesinin takibine yönelik yürütülen arazi ve laboratuvar çalışmaları kapsamında elde edilmiş olan su kalitesi verilerindedir. Bayram ve Kankal (2015) söz konusu su kalitesi parametrelerinden ÇO, T ve pH, tahmin yöntemlerinden ise CRA ve YSA, Bayram ve diğ., (2015) ise parametrelerin tamamını ve yöntemlerden CRA yanı sıra TLBO ve ABC algoritmaları kullanımı ile ÇO modelleme çalışmaları yapmıştır. Bu çalışma kapsamında ise MARS ve TreeNet yöntemleri kullanılarak kurulan modellerin performansları Bayram ve diğ., (2015) tarafından yapılan çalışmada CRA, TLBO ve ABC yöntemleri kullanılarak kurulan modellerin performansları ile karşılaştırılmıştır.

## 2. MATERYAL VE METOT

### 2.1. Çalışma Alanı

Türkiye hidrolojik olarak 25 havzaya ayrılmaktadır. Hidrolojik havzaların 22. si olan Doğu Karadeniz Havzası, Çoruh (23) ve Aras (24) havzaları ile birlikte, Kafkasya Ekolojik Bölgesi'nin Türkiye kısmını oluşturmaktadır. Doğu Karadeniz Havzası batıdan doğuya Terme Çayı (Samsun), Bolaman ve Melet çayları (Ordu), Aksu ve Harşit çayları (Giresun), Değirmendere ve Karadere (Trabzon) ile İyidere ve Fırtına Deresi (Rize) gibi birbirine paralel olarak uzanan akarsuların oluşturduğu alt havzalardan meydana gelmektedir. Bu alt havzaların en büyüğü, Vauk Dağı'nın kuzey eteklerinden ve Sifon Deresi ismiyle kaynağını alan, Gümüşhane'nin içinden geçen ve Giresun İli, Tirebolu İlçesi'nin doğusundan Karadeniz'e dökülen, 143 km ana kol uzunluğu ve 3.280 km<sup>2</sup> drenaj alanına sahip olan Harşit Çayı Havzasıdır (Bayram ve diğ., 2010a; Bayram ve diğ. 2010b). İçinden geçtiği ve atıksularını bünyesine aldığı yerleşim birimlerinden Tekke Beldesi, Gümüşhane il merkezi ve Torul İlçesi dikkate alınarak ve bu yerleşim birimlerine göre memba ve mansap olarak nitelendirilen kısımlarında H1 (Tekke, memba), H2 (Tekke, mansap), H3 (Gümüşhane, memba), H4 (Gümüşhane, mansap), H5 (Torul, memba) ve H6 (Torul, mansap) olarak isimlendirilen toplamda altı su kalitesi gözlem istasyonu (Şekil 1) belirlenmiştir (Bayram 2010a; Bayram, 2011). Çalışma alanı Gümüşhane İl sınırları içerisinde yer almaktadır.



**Şekil 1:**  
Harşit Çayı Havzası ve su kalitesi gözlem istasyonları

### 2.2. Su Kalitesi Değişkenlerinin Ölçümü

Çalışma kapsamında su kalitesi değişkenlerinden ÇÖ (mg/L), T (°C), pH ve Eİ (mS/cm) ölçümleri Japon menşeli Horiba U-10 (water quality checker) isimli taşınabilir cihaz yardımıyla

akarsuda gerçekleştirilmiştir. Ölçümlerinin bir dakika aralıklarla ve 15 dakika boyunca gerçekleştirildiği söz konusu su kalitesi değişkenleri bir başka taşınabilir cihaz yardımıyla da (Amerikan menşeli HQ 40d) eş zamanlı olarak takip edilmiştir. HQ 40d ölçümlerine ait veriler bu çalışmada kullanılmamış olup Horiba U-10 ölçümlerinin doğruluklarının kontrolü amaçlanmıştır.

Akarsudan alınan ham su örneklerine SS (°dH) tayini öncesi filtrasyon işlemi uygulanmıştır. Bu işlemde gözenek boyutu 0,45 µm olan 50 mm çaplı Sartorius marka selüloz nitrat filtreler ile 47 mm çaplı süzme haznesine sahip Sartorius marka vakumlu filtrasyon seti kullanılmıştır (Bayram, 2011).

UV-VIS spektrofotometre (Dr. Lange Cadas 200) ile 0,01 °dH hassasiyetle ve 1-20 °dH aralığında SS tayinleri gerçekleştirilebilmektedir. Tayininin temel prensibi, numunedeki kalsiyum ve magnezyum iyonlarının menekşe renkli ftaleyn kompleksini oluşturmasıdır. Oluşan bu renkli kompleks 10 mm'lik bir küvette ve 572 nm dalga boyunda ölçülmüştür. Tayinin bu prensiple (metalftaleyn) yapılabilmesi için su örneklerinin sıcaklığı 15-25 °C ve pH'sı 4-9 aralıklarında tutulmuştur (Bayram, 2011).

## **2.3. Su Kalitesi Değişkenlerinin Modellenmesinde Kullanılan Yöntemler**

### **2.3.1. Çok Değişkenli Uyarlanabilir Regresyon Eğrileri (MARS)**

Freidman (1991) tarafından geliştirilen MARS yöntemi, parametrik olmayan, bağımlı ve bağımsız değişkenler arasındaki fonksiyonel ilişkiye dair varsayımlarda bulunmayan bir regresyon yöntemidir. Bu varsayım yerine bağımlı ve bağımsız değişkenler arasında kendisinin çıkardığı temel fonksiyonlara dayanarak yeni bir ilişki oluşturmaktadır. MARS yöntemi bağımsız değişkenlerin farklı aralıklarına karşılık gelen temel fonksiyonları kullanarak esnek bir regresyon modeli kurmaktadır (Özbalcı, 2008; Toprak, 2011; Oğuz, 2014; Nacar ve diğ., 2018b). MARS yönteminde en uygun çıktıları veren modelin elde edilebilmesi için iki aşamalı bir süreç vardır. İlk aşamada mümkün olan tüm temel fonksiyonlar oluşturulur. Çok büyük bir model bulununcaya kadar yani modelin karmaşıklığı maksimum seviyeye ulaşıncaya kadar eklenen temel fonksiyonlarla model geliştirilir. İkinci aşamada oluşturulan maksimum model budanarak, yani önemli bağımsız değişkenler ve bu değişkenlerin karşılıklı etkileşimleri belirlenerek hata miktarı en küçük olan en uygun model oluşturulur (Ünal, 2009). MARS yöntemi kullanılarak kurulan modellerin anlaşılması ve yorumlanması oldukça kolaydır. Bu modellerde bağımlı ve bağımsız değişkenlerin dağılımları kategorilere ayrılmış veya sürekli olabilir. Modeller taban fonksiyonları ile parçalı doğrusal bir yapıya sahip olduğu için doğrusal modellere kıyasla daha esnek bir yapıya sahiptir. Bunun yanında kayıp ve uç değerlerden az etkilenir ve model oluşturulurken verilerde bir ön hazırlık gerektirmez. MARS yöntemi büyük veri setlerine uygulanabilir, bağımsız değişkenlerin birbirleri ile olan ilişkilerini tanımlar ve bu ilişkilerin anlaşılmasına yönelik grafikler üretir (Friedman, 1991; <http://www.utstat.utoronto.ca/reid/sta450/Mar16.pdf>; Yılmaz ve diğ., 2018; <https://www.minitab.com/content/dam/www/en/uploadedfiles/content/products/spm/IntroMARS.pdf>). MARS yöntemi bu üstünlüklerinin yanı sıra bazı dezavantajlara da sahiptir. Model kurulurken çok sayıda veri seti gerektirmesi ve her verinin sadece bir kez işlenebilmesi (<http://www.utstat.utoronto.ca/reid/sta450/Mar16.pdf>) bu dezavantajlar arasında gösterilebilir.

### 2.3.2. TreeNet Gradyan Arttırma Makinesi (TreeNet)

Friedman (2001) tarafından geliştirilen ve gradyan artırma olarak adlandırılan stokastik gradyan artırma, makine öğrenimi teknolojisinde devrim niteliğinde bir gelişme olarak kabul edilmektedir. TreeNet yöntemi sınıflandırma ve regresyon problemlerinin çözümünde gradyan artırma algoritmasını uygulayan çok etkili bir tekniktir. Bu yöntem modellemede hızlı sonuç üretme, eksik veri kümelerinde yüksek derecede hata toleransı sağlama ve yüksek doğrulukta veriler üretmesi gibi avantajlara sahiptir. Bir TreeNet modeli, birkaç 100 ila 1000 küçük ağaçtan oluşan ve bu ağaçların her biri için altı terminal düğümü içeren bir modeldir. Bu ağaçlardan her biri son modelin oluşmasına katkıda bulunur ve nihai model bu küçük ağaçların katkılarının toplamı olarak meydana gelir (Nacar ve diğ., 2020b). Bu karmaşık yapı ile TreeNet modeli, hassas sonuçlar veren bir kara kutu olarak düşünülebilir. TreeNet modeli birçok yönden karmaşık olsa da, modelin özü, genişleme devam ettikçe aşamalı olarak daha doğru hale gelen faktörlerin bir toplamı olan Fourier veya Taylor serileri gibi uzun seri genişlemesine benzer. YSA'nın aksine, TreeNet metodolojisi zaman alan veri hazırlığı, ön işleme veya eksik değerlerin ispatlanmasını da gerektirmez (Friedman, 2001). Bu avantajların yanı sıra elde ettiği sonuçların doğru olmaması gibi bir dezavantaja sahiptir. Ağaç tabanlı modeller bazen rekabetçi sonuçlar elde etmesine karşın bazı uygulamalarda gerçeğe en yakın sonuçları elde edemezler (Friedman, 2003).

### 2.4. Çözünmüş Oksijen Modelleme Çalışmaları

ÇO (mg/L) modelleme çalışmalarında kullanılacak bağımsız değişkenlerin belirlenmesi model sonuçlarını da doğrudan etkileyen önemli bir konudur. Bayram ve diğ. (2015) tarafından yapılan çalışmada kullanılan T (°C), pH, Eİ (µS/cm) ve SS (°dH) su kalitesi değişkenleri bu çalışmada bağımsız ve ÇO (mg/L) ise bağımlı değişken olarak kullanılmıştır. Mart 2009 ile Şubat 2010 döneminde Harşit Çayı (Gümüşhane) ana kolu üzerinde belirlenen altı istasyonda, 15 gün aralıklarla yürütülen 24 çalışma neticesinde elde edilen 144 adet veri setinin 114'ü modellerin eğitim aşamasında geriye kalan 30'u ise test aşamasında kullanılmıştır. Model performanslarının karşılaştırılabilmesi için eğitim ve test aşamalarında kullanılan veriler Bayram ve diğ. (2015) çalışmasında olduğu gibi seçilmiştir (Tablo 1). Veri seçimi yapılırken her iki veri seti için hem her mevsimden hem de her istasyondan eşit sayıda veri alınmasına dikkat edilmiştir. Böylece tüm mevsimlerin yanı sıra istasyonların da hem eğitim hem de test veri setinde temsil edilebilmesi sağlanmıştır.

**Tablo 1. Modelleme çalışmalarında kullanılan veri setlerinin (eğitim ve test) seçimi**

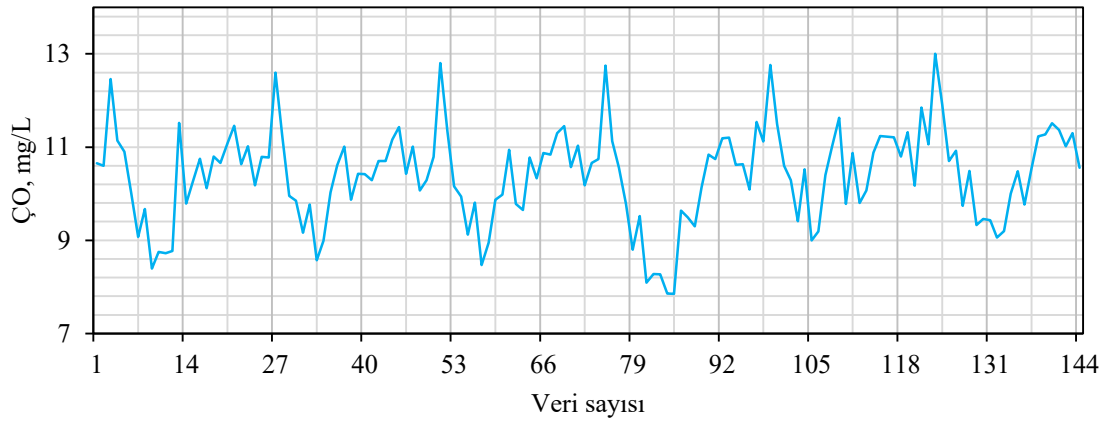
İstasyon	Verilerin ait olduğu dönem (Mart 2009 - Şubat 2010)																								
	İlkbahar						Yaz						Sonbahar						Kış						
H1	▲	■	■	■	▲	■	■	■	■	▲	■	■	■	■	▲	■	■	■	■	▲	■	■	■	■	▲
H2	■	▲	■	■	■	■	■	■	▲	■	■	■	■	■	■	▲	▲	■	■	▲	■	■	■	■	■
H3	■	▲	■	■	▲	■	■	■	■	▲	■	■	■	■	■	■	■	■	▲	■	■	■	■	■	▲
H4	■	▲	■	■	■	■	■	▲	■	■	■	■	■	■	■	■	■	■	▲	▲	■	▲	■	■	■
H5	■	■	■	▲	■	■	■	■	■	▲	▲	■	■	■	■	■	■	■	▲	■	■	■	■	■	▲
H6	■	■	■	▲	■	■	■	■	■	■	▲	■	▲	▲	■	■	■	■	■	■	■	■	■	■	▲

■ Eğitim veri seti ▲ Test veri seti

Çalışmada kullanılan verilerin tümü ve ayrıca eğitim ve test veri setleri için en küçük ( $x_{min}$ ), ortalama ( $\bar{X}$ ), en büyük ( $x_{mak}$ ), standart sapma ( $\sigma$ ), varyasyon katsayısı (VK) ve korelasyon katsayısı ( $r$ ) temel istatistikler Tablo 2'de verilmektedir. Çalışma kapsamında kullanılan ÇO değerlerinin zamanla değişiminin daha net bir şekilde görülebilmesi için de tüm veri setine ait zaman serisi Şekil 2'de verilmektedir.

**Tablo 2. Modelleme çalışmalarında kullanılan Harşit Çayı su kalitesi veri setlerine ait temel istatistikler (Bayram, 2011; Bayram ve diğ., 2015)**

Veri Seti	Tüm veriler					Eğitim veri seti					Test veri seti				
	Değişken	ÇO	T	pH	Eİ	SS	ÇO	T	pH	Eİ	SS	ÇO	T	pH	Eİ
Birim	mg/L	°C		mS/cm	°dH	mg/L	°C		mS/cm	°dH	mg/L	°C		mS/cm	°dH
X <sub>min</sub>	7,85	2,00	6,93	0,14	3,75	7,85	2,20	6,93	0,14	4,15	8,57	2,00	7,21	0,15	3,75
$\bar{X}$	10,39	9,94	7,78	0,32	10,79	10,42	10,06	7,77	0,32	10,74	10,30	9,50	7,81	0,33	10,97
X <sub>mak</sub>	13,00	21,40	8,40	0,66	18,80	13,00	21,40	8,40	0,66	18,80	11,91	20,60	8,35	0,53	16,70
$\sigma$	1,01	5,44	0,36	0,11	3,52	1,04	5,35	0,37	0,11	3,69	0,84	5,65	0,32	0,09	2,76
VK	9,67	54,52	4,60	33,85	32,59	10,00	53,20	4,72	35,35	34,32	8,20	59,52	4,06	26,64	25,20
r	1,00	-0,61	0,47	-0,40	-0,31	1,00	-0,58	0,43	-0,35	-0,27	1,00	-0,79	0,71	-0,66	-0,56

**Şekil 2:**

*Harşit Çayı çözünmüş oksijen (mg/L) verilerine ait zaman serisi*

Modelleme çalışmalarında veri seti içerisinde bulunan uç değerlerin modellenmesi zor olabilmektedir. Uç değerlerin tahminini kolaylaştırmak, farklı boyutların etkisini en aza indirmek ve daha etkili sonuçlar elde etmek için bağımlı ve bağımsız değişkenlere ait verilerin tamamı Eşitlik (1) kullanılarak normalize edilmiştir (Bayram ve diğ., 2015; Fetene ve diğ., 2018; Nacar ve diğ., 2018a). Su kalitesi modelleme çalışmalarında normalizasyon için farklı bağıntılar kullanılmaktadır. Yapılan çalışmada verilerin normalizasyonu için Eşitlik (1) kullanılmıştır.

$$X_n = \frac{X_i - X_{\min}}{X_{\max} - X_{\min}} \times a + b \quad (1)$$

Burada,  $X_n$  normalize değeri,  $X_{\min}$  en küçük değeri ve  $X_{\max}$  en büyük değeri ifade etmektedir. Bayram ve diğ., (2015) çalışmasında olduğu gibi bu çalışmada da a ve b katsayıları sırasıyla 0,8 ve 0,1 olarak seçilmiştir. Böylece modellemede kullanılan tüm veriler 0,1-0,9 arasında normalize edilmiştir. Veriler normalize edildikten sonra eğitim veri seti kullanılarak MARS ve TreeNet modelleri kurulmuş ve en düşük hata değerlerini veren modele ait temel fonksiyonlar ve denklem belirlenmiştir. Daha sonra test veri setinde yer alan bağımsız değişkenler belirlenen temel fonksiyonlar ve denklemlerdeki yerlerine yazılarak model çıktıları olan ÇO (mg/L) değerleri elde edilmiştir. Modelleme çalışmaları tamamlandığında elde edilen çıktılar ters normalize işleminin ardından eski boyutlarına dönüştürülmüştür. Çalışmada kullanılan yöntemlerin ve kurulan modellerin performanslarının değerlendirilmesi için Bayram ve diğ. (2015) tarafından yapılan çalışmada kullanılan performans istatistikleri hesaplanmıştır.

Böylelikle bu çalışmadan elde edilen performans istatistikleri ile Bayram ve diğ. (2015) tarafından elde edilen performans istatistikleri karşılaştırılabilmiştir. Çalışmada kullanılan performans istatistikleri ortalama karesel hatanın karekökü (OKHK), ortalama mutlak hata (OMH), ortalama rölatif hata (ORH) ve determinasyon katsayısı ( $R^2$ ) olup hesaplanmalarında aşağıda verilen Eşitlikler (2-5) kullanılmıştır. Bu eşitliklerde, ölçüm değerlerini  $t_i$ , tahmin değerlerini  $td_i$  ve veri adedini  $N$  temsil etmektedir.

$$OKHK = \sqrt{\frac{1}{N} \sum_{i=1}^N (td_i - t_i)^2} \quad (2)$$

$$OMH = \frac{1}{N} \sum_{i=1}^N |(td_i - t_i)| \quad (3)$$

$$ORH = \frac{1}{N} \sum_{i=1}^N \left| \frac{(td_i - t_i)}{td_i} \right| \times 100 \quad (4)$$

$$R^2 = 1 - \frac{\sum_{i=1}^N (td_i - t_i)^2}{\sum_{i=1}^N (td_i)^2} \quad (5)$$

### 3. BULGULAR VE İRDELEME

Çalışma kapsamında MARS ve TreeNet yöntemleri ile ÇO (mg/L) modellemesinde kullanılabilecek en uygun modelin belirlenmesinde yöntem temelli bazı parametrelerin model performanslarını etkilediği görülmüş ve bu parametreler deneme yanılma yolu ile değiştirilerek en uygun sonuçları veren modele ait parametre değerleri belirlenmeye çalışılmıştır. MARS ve TreeNet yöntemleri için model performansında etkili olduğu belirlenen parametreler ve denenen değer aralıkları Tablo 3'te verilmektedir.

**Tablo 3. MARS ve TreeNet yöntemlerinde kullanılan parametreler ve değer aralıkları**

MARS yöntemi		TreeNet yöntemi	
Parametreler	Değer aralıkları	Parametreler	Değer aralıkları
Hız faktörü	1-5	Aşırı öğrenme koruma parametresi	0,5-1
Düğüm optimizasyonu için serbestlik derecesi	1-5	Ağaç sayısı	10000
Maksimum temel fonksiyon sayısı	5-50	Ağaç başına maksimum düğüm sayısı	1-10
Maksimum etkileşim	1-10	Maksimum ağaç derinliği	10000-50000
Düğümmler arasındaki minimum gözlem sayısı	1-10	Minimum terminal düğüm sayısı	1-10
Ridge	-9-0		

Yapılan modelleme çalışmalarında yöntemlerin model performanslarını etkileyen parametreler Tablo 3'te verilen değer aralıklarında kullanılarak en yüksek performansı veren model belirlenmiştir. Bu modeller arasından diğerlerine kıyasla daha yüksek performans değerlerine sahip olan altı model için kullanılan parametreler ve bu parametrelere ait değerler Tablo 4'te verilmektedir.

**Tablo 4. MARS ve TreeNet modellerine ait parametreler ve en iyi model için kullanılan değerler**

Modeller	MARS yöntemi						TreeNet yöntemi					
	Hız faktörü	Düğüm optimizasyonu için serbestlik derecesi	Maksimum temel fonksiyon sayısı	Maksimum etkileşim	Düğümler arasındaki minimum gözlem sayısı	Ridge	Aşırı öğrenme koruma parametresi	Ağaç sayısı	Ağaç başına maksimum düğüm sayısı	Maksimum ağaç derinliği	Minimum terminal düğümü	
M1	4	4	20	3	5	-5	0,5	10000	6	50000	3	
M2	4	4	20	3	7	-5	0,5	10000	10	50000	1	
M3	4	4	20	3	5	-2	1,0	10000	10	100000	4	
M4	4	4	26	3	5	-2	1,0	10000	10	100000	10	
M5	4	4	30	4	4	-2	1,0	10000	7	10000	7	
M6	4	4	13	4	5	-2	1,0	10000	3	10000	5	

Tablo 4'te verilen bazı parametrelerin artırılıp azaltılmasının model performanslarına etkilerinin olduğu belirlenmiş ancak bu modeller yüksek performans değerleri vermediği için sunulmamıştır. Sunulan modeller için hesaplanan performans istatistikleri her iki yöntem için Tablo 5'te verilmektedir.

Tablo 5 incelendiğinde MARS yönteminde eğitim veri seti için M6 dışındaki modellerin tamamı M6'ya kıyasla daha iyi performans göstermiştir. Ancak test veri seti için hesaplanan performans değerleri incelendiğinde M6 modelinin diğerlerine kıyasla en az %40 en fazla ise %200 daha iyi performans gösterdiği görülmüştür. Benzer bir değerlendirme TreeNet yöntemi kullanılarak kurulan modeller için yapıldığında, eğitim ve test veri setleri için en iyi performans değerlerini veren modelin M3 olduğu görülmektedir. Bu değerlendirmeler ışığında her iki yöntem için bir değerlendirme yapıldığında en yüksek performansı gösteren modellerin MARS yöntemi için M6, TreeNet yöntemi için ise sırasıyla M2 ve M3 olduğu sonucuna varılmıştır. Model performans değerleri incelendiğinde yöntemlerde yer alan parametrelerin farklı değerlerine karşılık elde edilen çıktıların farklı performans değerlerine sahip olduğu görülmektedir. Bu durum yöntemlerin kurulmasında kullanılan parametrelerin model performansları üzerindeki etkilerini açıkça ortaya koymaktadır.

**Tablo 5. Eğitim ve test veri setlerine ait performans istatistikleri**

	Modeller (Eğitim)						Modeller (Test)					
	M1	M2	M3	M4	M5	M6	M1	M2	M3	M4	M5	M6
MARS												
OKHK	0,6022	0,5056	0,6026	0,6110	0,6096	<b>0,6264</b>	0,4588	0,8273	0,4613	0,3985	0,4746	<b>0,2911</b>
OMH	0,4897	0,3959	0,4908	0,4884	0,4828	<b>0,4971</b>	0,3454	0,5300	0,3464	0,3130	0,3907	<b>0,2336</b>
ORH	4,7500	3,8400	4,7700	4,7400	4,6600	<b>4,8100</b>	3,3900	4,9800	3,4000	3,0700	3,7100	<b>2,2700</b>
R <sup>2</sup>	0,9967	0,9977	0,9967	0,9966	0,9966	<b>0,9964</b>	0,9981	0,9938	0,9980	0,9985	0,9979	<b>0,9992</b>
TreeNet												
OKHK	0,2556	0,2394	<b>0,2247</b>	0,2960	0,3767	0,3437	0,5263	0,5150	<b>0,5023</b>	0,5710	0,5632	0,5236
OMH	0,0949	0,0889	<b>0,0666</b>	0,1178	0,1912	0,1656	0,3773	0,3754	<b>0,3758</b>	0,4250	0,3884	0,3674
ORH	0,9300	0,8700	<b>0,6600</b>	1,1700	1,9000	1,6300	3,7600	3,7400	<b>3,7300</b>	4,2200	3,8600	3,6300
R <sup>2</sup>	0,9994	0,9995	<b>0,9995</b>	0,9992	0,9987	0,9989	0,9974	0,9975	<b>0,9976</b>	0,9969	0,9970	0,9974

MARS yöntemi için en yüksek performansı gösteren M6 modeline ait temel fonksiyonlar ve bu temel fonksiyonlar kullanılarak oluşturulan genel denklem Tablo 6'da verilmektedir. Bu temel fonksiyonlar ve genel eşitlik kullanılarak ÇO değerleri test veri seti için üretilmiş ve performansları bu üretilen değerler üzerinden hesaplanmıştır. TreeNet yönteminde ise üretilen genel denklem ağaç dal yapılanmasına benzediği için MARS yönteminden elde edilen genel denkleme kıyasla çok daha uzun ve karmaşık bir yapıda olduğundan verilememiştir.



**Tablo 6. MARS yöntemi M6 modeline ait temel fonksiyonlar ve denklemleri**

Temel fonksiyonlar	
TF02	$\text{mak} (0, 0,789 - T)$
TF03	$\text{mak} (0, SS - 0,402) \times \text{TF02}$
TF04	$\text{mak} (0, 0,402 - SS) \times \text{TF02}$
TF05	$\text{mak} (0, \text{pH} - 0,573) \times \text{TF04}$
TF07	$\text{mak} (0, E\dot{I} - 0,579) \times \text{TF03}$
TF10	$\text{mak} (0, 0,557 - \text{pH})$
TF12	$\text{mak} (0, 0,739 - T) \times \text{TF10}$
ÇO	$= 0,406923 + 0,253035 \times \text{TF02} + 13,2586 \times \text{TF05} + 82,8305 \times \text{TF07} - 0,91322 \times \text{TF10} + 2,05572 \times \text{TF12}$

MARS ve TreeNet yöntemleri kullanılarak en iyi performans değerlerini veren M6 ve M3 modellerine ait performans değerleri Bayram ve diğ. (2015) tarafından hesaplanan performans değerleri ile sırasıyla eğitim ve test veri setleri için karşılaştırılmıştır (Tablo 7).

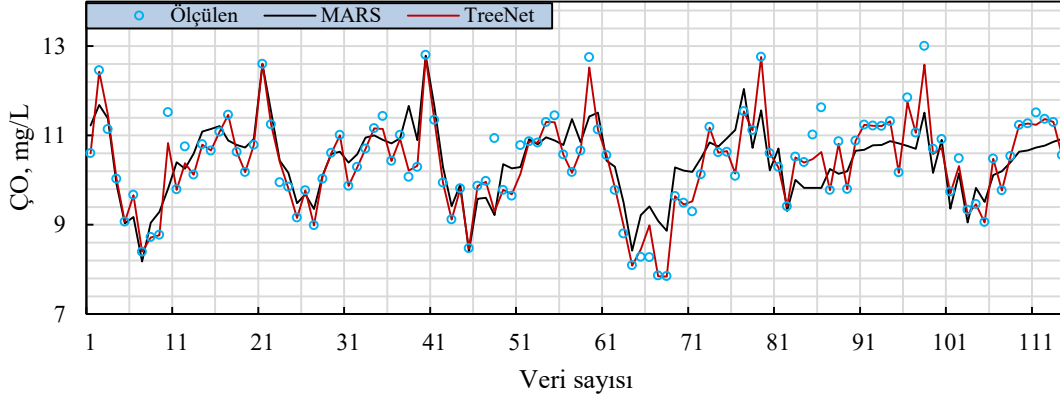
**Tablo 7. Eğitim veri seti için hesaplanan performans istatistiklerinin karşılaştırılması**

Bu çalışma									
		Eğitim				Test			
Yöntem	Model	OKHK (mg/L)	OMH (mg/L)	ORH (%)	R <sup>2</sup>	OKHK (mg/L)	OMH (mg/L)	ORH (%)	R <sup>2</sup>
MARS	M6	<b>0,6264</b>	<b>0,4971</b>	<b>4,8100</b>	<b>0,9964</b>	<b>0,2911</b>	<b>0,2336</b>	<b>2,2700</b>	<b>0,9992</b>
TreeNet	M3	<b>0,2247</b>	<b>0,0666</b>	<b>0,6600</b>	<b>0,9995</b>	<b>0,5023</b>	<b>0,3758</b>	<b>3,7300</b>	<b>0,9976</b>
Bayram ve diğ. (2015)									
		Eğitim				Test			
Yöntem	Fonksiyon	OKHK (mg/L)	OMH (mg/L)	ORH (%)	R <sup>2</sup>	OKHK (mg/L)	OMH (mg/L)	ORH (%)	R <sup>2</sup>
TLBO	Kuadratik	<b>0,6411</b>	<b>0,5052</b>	<b>4,8400</b>	<b>0,9962</b>	<b>0,3294</b>	<b>0,2616</b>	<b>2,6100</b>	<b>0,9990</b>
	Üstel	0,7973	0,6489	6,2600	0,9942	0,6930	0,5280	4,9600	0,9955
	Doğrusal	0,7963	0,6484	6,2500	0,9942	0,3989	0,3291	3,2400	0,9985
	Üs	0,8127	0,6744	6,5300	0,994	0,6807	0,5205	4,8900	0,9957
ABC	Kuadratik	0,8389	0,7002	6,6800	0,9936	0,4003	0,3182	3,1300	0,9985
	Üstel	0,8126	0,6576	6,3900	0,9940	0,6436	0,5078	4,7700	0,9961
	Doğrusal	0,8157	0,6765	6,5000	0,9939	0,3689	0,3115	3,0800	0,9987
	Üs	0,8198	0,6895	6,7100	0,9939	0,6119	0,4844	4,5500	0,9965
CRA	Kuadratik	0,7569	0,6202	6,0100	0,9948	0,6086	0,4856	4,7700	0,9965
	Üstel	0,7973	0,6504	6,3000	0,9942	0,4163	0,3538	3,5000	0,9984
	Doğrusal	0,7964	0,6481	6,2500	0,9942	0,5527	0,3260	3,2100	0,9985
	Üs	0,8013	0,6629	6,4000	0,9941	0,5691	0,3188	3,1500	0,9984

Tabloda sunulan değerler bu çalışmadan elde edilen ve Bayram ve diğ. (2015) çalışmasından alınan performans değerlerini göstermektedir. Bayram ve diğ. (2015) tarafından yapılan çalışmada, en yüksek performansı gösteren modelin belirlenmesinde dört farklı regresyon eşitliği ve üç farklı yöntem olmak üzere toplamda 12 farklı model uygulanmıştır. Bayram ve diğ. (2015) tarafından kuadratik, üstel, doğrusal ve üs fonksiyonlarına ait katsayılar TLBO ve ABC yöntemleri ile optimize edilmiş ve hesaplanan performans istatistikleri CRA'dan elde edilen performans istatistikleri ile karşılaştırılmıştır. Tablo 7 incelendiğinde, Bayram ve diğ. (2015) tarafından yapılan çalışmada en iyi performans değerlerinin kuadratik fonksiyonun ağırlıklarının TLBO yöntemi ile optimizasyonundan elde edilen katsayıların kullanıldığı modele ait olduğu anlaşılmaktadır. Söz konusu çalışmada elde edilen en iyi modele ait performans değerleri kalın olarak sunulmuştur. ABC yöntemi kullanılarak elde edilen performans değerlerinin CRA yönteminden daha kötü performans değerlerine sahip olduğu da görülmektedir.

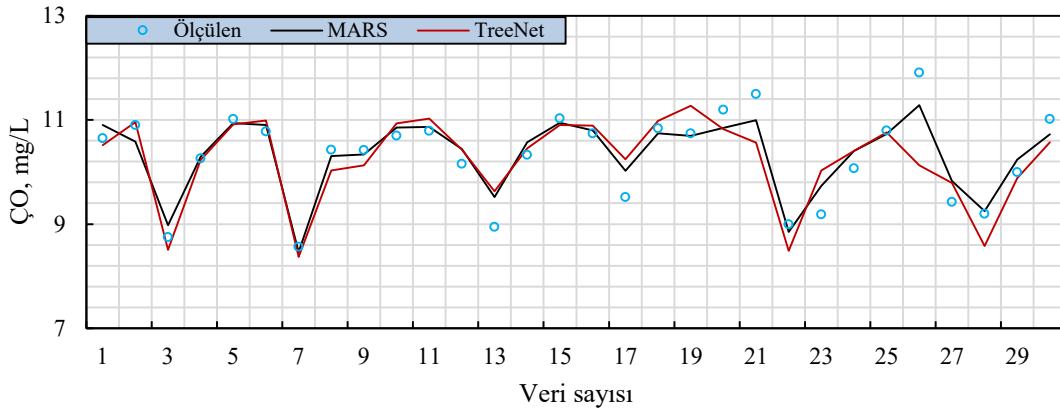
MARS ve TreeNet yöntemleri kullanılarak kurulan M6 ve M3 modellerinden elde edilen performans değerleri TLBO, ABC ve CRA yöntemlerine ait modellerden elde edilen değerlerle

(Bayram ve diğ., 2015) karşılaştırıldığında eğitim veri seti için en yüksek performans değerlerinin sırasıyla TreeNet ve MARS yöntemlerinden elde edildiği görülmektedir (Tablo 7). Ancak test seti için en yüksek performans değerleri MARS yönteminden elde edilmiştir. Bayram ve diğ. (2015) tarafından kullanılan TLBO yöntemine kıyasla bu çalışmada kullanılan TreeNet yönteminde eğitim veri seti için OKHK performans istatistiğinde yaklaşık % 65'lik bir artışı ve test veri seti için yaklaşık % 52'lik bir düşüş söz konusudur. MARS yönteminde ise yine aynı performans istatistiğinde eğitim veri setinde yaklaşık % 2'lik ve test veri setinde ise % 11'lik bir performans artışı sağlanmıştır. Eğitim ve test veri setleri için hesaplanan performans istatistikleri birlikte değerlendirildiğinde, en yüksek performansa sahip yöntemin MARS yöntemi olduğu anlaşılmaktadır. MARS yönteminde model performans değerlerindeki iyileşme miktarı her ne kadar düşük olsa da TLBO ve ABC yöntemlerine kıyasla uygulanması kolay ve çok kısa sürede daha yüksek performans değerleri vermektedir. MARS ve TreeNet yöntemlerinde en yüksek performans değerlerine sahip ÇO modellerinin eğitim ve test veri setleri çıktıları ile ölçüm değerlerinin görsel olarak karşılaştırılabilmesi amacıyla eğitim veri seti için Şekil 3 ve test veri seti için Şekil 4'te zaman serileri verilmektedir.



**Şekil 3:**

*Çözünmüş oksijen eğitim veri seti model çıktıları ile ölçüm değerlerinin karşılaştırılması*

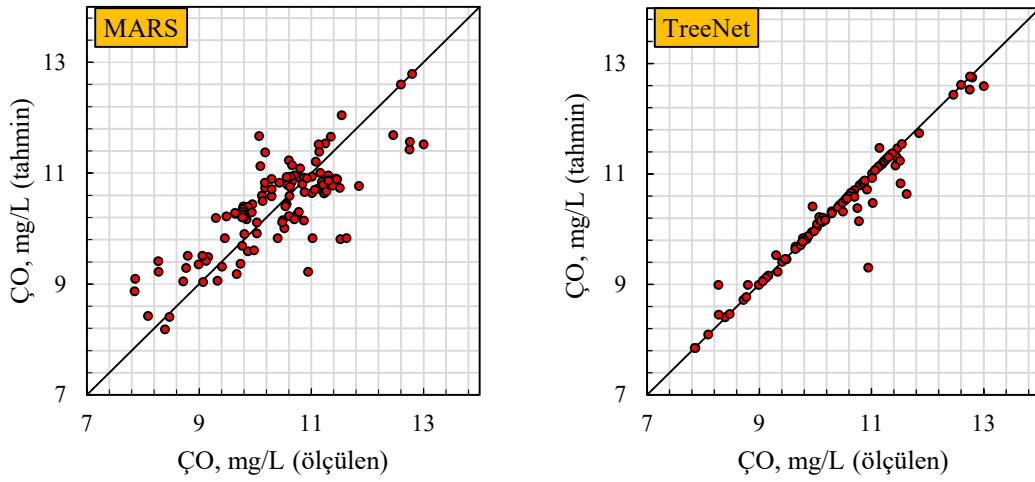


**Şekil 4:**

*Çözünmüş oksijen test veri seti model çıktıları ile ölçüm değerlerinin karşılaştırılması*

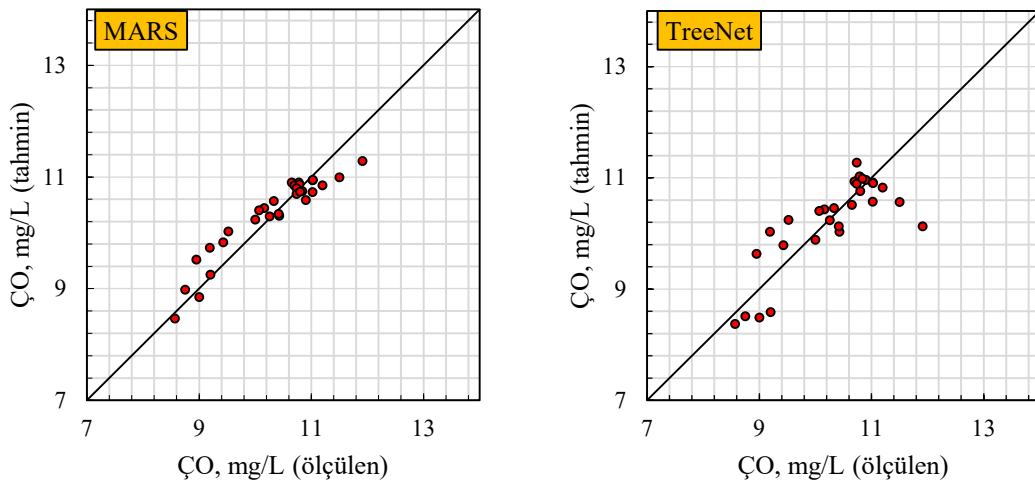
Eğitim ve test veri setleri için hazırlanmış olan zaman serileri (Şekil 3 ve 4) incelendiğinde TreeNet yönteminin MARS yöntemine kıyasla en büyük ve en küçük değerleri daha iyi modellediği görülmektedir. Ancak test veri seti için MARS yöntemi TreeNet yöntemine kıyasla ölçüm değerlerine daha yakın değerler üretmiştir. Model performanslarının daha net görülebilmesi amacıyla eğitim veri seti için Şekil 5 ve test veri seti için Şekil 6'da saçılım grafikleri de verilmektedir.

Saçılım grafiklerinde verilerin köşegene yakın dağılması model performansının arttığını göstermektedir. Ayrıca köşegen üzerinde rölatif hata sifira yaklaşmaktadır (Bayram ve diğ., 2015). Şekil 5 incelendiğinde, eğitim veri seti için TreeNet yöntemi kullanılarak modellenen ÇO değerlerinin ölçüm değerlerine daha yakın olduğu, Şekil 6'da ise test veri seti için MARS yöntemi kullanılarak elde edilen ÇO değerlerinin ölçüm değerlerine daha yakın olduğu görülmektedir.



**Şekil 5:**

*Çözünmüş oksijen eğitim veri seti model çıktıları ile ölçüm değerleri saçılım grafikleri*



**Şekil 6:**

*Çözünmüş oksijen test veri seti model çıktıları ile ölçüm değerleri saçılım grafikleri*

#### 4. SONUÇLAR

Bu çalışmada çok değişkenli uyarlanabilir regresyon eğrileri (MARS) ve TreeNet gradyan arttırma makinesi (TreeNet) yöntemlerinin çözünmüş oksijen (ÇO) konsantrasyonu modellenmesindeki performansları araştırılmıştır. Bu amaçla modelleme çalışmalarında kentsel atıksuları bünyesine alan Harşit Çayı (Gümüşhane) ana kol boyunca seçilen altı istasyonda yerinde ölçülen ÇO (mg/L), sıcaklık (T, °C), pH ve elektriksel iletkenlik (Eİ, µS/cm) ile su sertliği (SS, °dH) su kalitesi değişkenleri kullanılmıştır. Çalışmadan elde edilen sonuçlar aşağıda verilmektedir.

Model performanslarının karşılaştırılması sonucunda çalışma kapsamında kullanılan TreeNet yönteminin eğitim veri seti için MARS yöntemine kıyasla daha doğru sonuçlar verdiği görülmüştür. Ancak test veri seti için MARS yönteminin TreeNet'e kıyasla daha iyi performans değerleri verdiği belirlenmiştir. MARS yöntemi temel olarak veri setine özel taban fonksiyonları üretir. Bağımsız değişkenlerin tüm serisine bir katsayı üretmek yerine daha esnek küçük doğrusal fonksiyonlarla tüm seri için bir denklem elde edilir. Bu nedenle test veri setinde diğer regresyon yöntemlerinden daha başarılı sonuçlar üretebilmektedir. Eğitim ve test veri setleri için hesaplanan performans değerlerine göre MARS yönteminden, daha önce aynı veri setleriyle yapılmış olan çalışmada kullanılan öğretme-öğrenme tabanlı optimizasyon (TLBO), yapay arı kolonisi (ABC) ve geleneksel regresyon analizi (CRA) yöntemlerine kıyasla daha iyi performans değerleri elde edilmiştir. MARS yönteminin kullanım kolaylığı, TLBO ve ABC yöntemlerine kıyasla daha az bilgisayar gücü ve zaman gerektirmesi ile CRA yöntemine göre % 100'e varan oranlarda performans artışı sağlaması üstünlükleri nedeniyle ÇO modellemesi çalışmalarında tercih edilebilir bir yöntem olduğu sonucuna ulaşılmıştır. Bu çalışmada ÇO modellenmesinde dört farklı su kalitesi değişkeni kullanılmış ancak daha az sayıda su kalitesi değişkeni ve farklı bağımsız değişken kombinasyonları kullanılarak benzer sonuçların elde edilebilirliği araştırılmamıştır.

#### ÇIKAR ÇATIŞMASI

Yazarlar, bilinen herhangi bir çıkar çatışması veya herhangi bir kurum/kuruluş ya da kişi ile ortak çıkar bulunmadığını onaylamaktadırlar.

#### YAZAR KATKISI

Dr. Adem Bayram arazi ölçümlerinin gerçekleştirilmesinde ve verilerin toplanmasında, Dr. Adem Bayram ve Dr. Sinan Nacar çalışmanın kavramsal ve tasarım süreçlerinin belirlenmesinde, Dr. Sinan Nacar ve İnş. Yük. Müh. Osman Tuğrul Baki çalışmanın kavramsal ve tasarım süreçlerinin yönetiminde, Dr. Adem Bayram, Dr. Sinan Nacar ve İnş. Yük. Müh. Osman Tuğrul Baki verilerin analizi, modelleme çalışmalarının yapılması, yorumlanması ve makalenin yazılmasında, Dr. Adem Bayram ve Dr. Sinan Nacar gerekli düzeltmelerin yapılmasında yer almışlardır.

#### TEŞEKKÜR

Yazarlar, modelleme çalışmalarında kullanılan Salford Predictive Modeler 8 yazılımının sağlayıcılarına desteklerinden ötürü teşekkür eder. Harşit Çayında yürütülen su kalitesi takibi çalışmalarındaki özveri desteklerinden ötürü, Dr. Volkan Numan Bulut, Dr. Tayfun Dede, Dr. Mustafa Durmaz ve Hidrolik Laboratuvarı Teknisyeni Yüksel Hardal'ın yanı sıra 19.02.2013 tarihinde aramızdan ayrılan kıymetli büyüğümüz ve değerli hocamız merhum Dr. Murat İhsan Kömürçü'ye teşekkür ederler. Mekanı cennet olsun.

## KAYNAKLAR

1. Altunkaynak, A., Ozger, M. ve Cakmakci, M. (2005) Fuzzy logic modeling of the dissolved oxygen fluctuations in Golden Horn, *Ecological Modelling*, 189(3-4), 436-446. <https://doi.org/10.1016/j.ecolmodel.2005.03.007>
2. Altuntaş, M. (2018) Yeşilirmak Havzası su kalitesi parametrelerinin yapay zeka teknikleriyle modellenmesi, *Yüksek Lisans Tezi*, Erciyes Üniversitesi, Fen Bilimleri Enstitüsü, Kayseri.
3. Bayram, A. (2011) Harşit Çayı su kalitesinin mevsimsel değişiminin incelenmesi ve askı madde konsantrasyonunun yapay sinir ağları yöntemi ile tahmin edilmesi, *Doktora Tezi*, Karadeniz Teknik Üniversitesi, Fen Bilimleri Enstitüsü, Trabzon.
4. Bayram, A. ve Kankal, M. (2015) Artificial neural network modeling of dissolved oxygen concentration in a Turkish watershed, *Polish Journal of Environmental Studies*, 24(4), 1507-1515.
5. Bayram, A., Kankal, M., Önsoy, H. ve Bulut, V.N. (2010a) Harşit Çayı hidrolik yapılarının askı madde hareketine etkileri, *VI. Ulusal Hidroloji Kongresi*, Denizli, 873-882.
6. Bayram, A., Onsoy, H., Bulut, V.N. ve Tufekci, M. (2010b) Dissolved oxygen levels in the stream Harşit (Turkey), *2010 In: 9th International Congress on Advances in Civil Engineering*, Trabzon, (Full text in CD: ACE 2020HYD-041).
7. Bayram, A., Uzlu, E., Kankal, M. ve Dede, T. (2015) Modeling stream dissolved oxygen concentration using teaching-learning based optimization algorithm, *Environmental Earth Sciences*, 73(10), 6565-6576. <https://doi.org/10.1007/s12665-014-3876-3>
8. Dede, O., Ilker, T. ve Aral, M. (2013) The use of water quality index models for the evaluation of surface water quality: A case study for Kirmir Basin, Ankara, Turkey, *Water Quality Exposure and Health*, 5(1), 41-56. <https://doi.org/10.1007/s12403-013-0085-3>
9. Elkiran, G., Nourani, V. ve Abba, S.I. (2019) Multi-step ahead modelling of river water quality parameters using ensemble artificial intelligence-based approach, *Journal of Hydrology*, 577, 123962. <https://doi.org/10.1016/j.jhydrol.2019.123962>
10. Fetene, B.N., Shufen, R. ve Dixit, U.S. (2018) FEM-based neural network modeling of laser-assisted bending, *Neural Computing and Applications*, 29(6), 69-82. <https://doi.org/10.1007/s00521-016-2544-9>
11. Friedman, J.H. (1991) Multivariate adaptive regression splines, *The Annals of Statistics*, 19(1), 1-67.
12. Friedman, J.H. (2001) Greedy function approximation: a gradient boosting machine. *The Annals of Statistics*, 29, 1189-1232. <https://doi.org/10.1214/aos/1013203451>
13. Friedman, J.H. ve Meulman, J.J. (2003) Multiple additive regression trees with application in epidemiology, *Statistics in Medicine*, 22(9), 1365-1381. <https://doi.org/10.1002/sim.1501>
14. Göksu, M.Z.L. (2015) *Su Kirliliği*, 2. Basın, Akademisyen Kitapevi, Ankara.
15. Heddam, S. ve Kisi, O. (2018) Modelling daily dissolved oxygen concentration using least square support vector machine, multivariate adaptive regression splines and M5 model tree, *Journal of Hydrology*, 559, 499-509. <https://doi.org/10.1016/j.jhydrol.2018.02.061>
16. <https://www.minitab.com/content/dam/www/en/uploadedfiles/content/products/spm/IntroMARS.pdf>, Erişim Tarihi: 24.02.2022, Konu: *Introducing MARS*.

17. <http://www.utstat.utoronto.ca/reid/sta450/Mar16.pdf>, Erişim Tarihi: 24.02.2022, Konu: STA450S/4000S: Topics in Statistics. Statistical Aspects of Data Mining, Ders Notu.
18. Karadeniz, S. (2020) Giresun yaylaları kaynak suları kalitesinin belirlenmesi, *Yüksek Lisans Tezi*, Giresun Üniversitesi, Fen Bilimleri Enstitüsü, Giresun.
19. Kisi, O., Akbari, N., Sanatipour, M., Hashemi, A., Teimourzadeh, K. ve Shiri, J. (2013). Modeling of dissolved oxygen in river water using artificial intelligence techniques, *Journal of Environmental Informatics*, 22(2), 92-101. <https://doi.org/10.3808/jei.201300248>
20. Kose, E., Tokatli, C. ve Cicek, A. (2014). Monitoring stream water quality: a statistical evaluation, *Polish Journal of Environmental Studies*, 23(5), 1637-1647.
21. Könez, H. (2019) Çarşıbaşı Deresi'nin fizikokimyasal su kalitesi ve kirlilik seviyesinin araştırılması, *Yüksek Lisans Tezi*, Recep Tayyip Erdoğan Üniversitesi, Fen Bilimleri Enstitüsü, Rize.
22. Kukrer, S. ve Mutlu, E. (2019). Assessment of surface water quality using water quality index and multivariate statistical analyses in Saraydüzü Dam Lake, Turkey, *Environmental Monitoring and Assessment*, 191(2), 1-16. <https://doi.org/10.1007/s10661-019-7197-6>
23. Küçükler, E. (2020) Yanıklar Deresi'nin (Fethiye-Muğla) limnolojik ve su kalitesi yönünden incelenmesi, *Yüksek Lisans Tezi*, Muğla Sıtkı Koçman Üniversitesi, Fen Bilimleri Enstitüsü, Muğla.
24. Nacar, S., Bayram, A., Baki, O.T., Kankal, M. ve Aras, E. (2020a) Spatial forecasting of dissolved oxygen concentration in the Eastern Black Sea Basin, Turkey, *Water*, 12(4), 1041. <https://doi.org/10.3390/w12041041>
25. Nacar, S., Hınıs, M.A. ve Kankal, M. (2018a) Forecasting daily streamflow discharges using various neural network models and training algorithms, *KSCE Journal of Civil Engineering*, 22(9), 3676-3685. <https://doi.org/10.1007/s12205-017-1933-7>
26. Nacar, S., Kankal, M. ve Hınıs, M.A. (2018b) Çok değişkenli uyarlanabilir regresyon eğrileri (ÇDURE) ile günlük akarsu akımlarının tahmini - Haldizen Deresi örneği, *Gümüşhane Üniversitesi Fen Bilimleri Enstitüsü Dergisi*, 8(1), 38-47. <https://doi.org/10.17714/gumusfenbil.311188>
27. Nacar, S., Mete, B. ve Bayram, A. (2020b) Günlük çözünmüş oksijen konsantrasyonunun çok değişkenli uyarlanabilir regresyon eğrileri ile tahmin edilmesi, *Bursa Uludağ University Journal of The Faculty of Engineering*, 25(3), 1479-1498. <https://doi.org/10.17482/uumfd.750518>
28. Nacar, S., Mete, B. ve Bayram, A. (2020c) Estimation of daily dissolved oxygen concentration for river water quality using conventional regression analysis, multivariate adaptive regression splines, and TreeNet techniques, *Environmental Monitoring and Assessment*, 192(12), 1-21. <https://doi.org/10.1007/s10661-020-08649-9>
29. Oğuz, A. (2014) Çok değişkenli uyarlanabilir regresyon zincirlerinin irdelenmesi ve bir uygulama, *Yüksek Lisans Tezi*, Erzincan Üniversitesi, Fen Bilimleri Enstitüsü, Erzincan.
30. Olyaei, E., Abyaneh, H.Z. ve Mehr, A.D. (2017) A comparative analysis among computational intelligence techniques for dissolved oxygen prediction in Delaware River, *Geoscience Frontiers*, 8(3), 517-527. <https://doi.org/10.1016/j.gsf.2016.04.007>
31. Özfalcı, Y. (2008) Çokdeğişkenli uyarlanabilir regresyon kesitleri: MARS, *Yüksek Lisans Tezi*, Gazi Üniversitesi, Fen Bilimleri Enstitüsü, Ankara.

32. Sener, S., Sener, E. ve Davraz, A. (2017) Evaluation of water quality using water quality index (WQI) method and GIS in Aksu River (SW-Turkey), *Science of the Total Environment*, 584, 131-144. <https://doi.org/10.1016/j.scitotenv.2017.01.102>
33. Sengorur, B., Dogan, E., Koklu, R. ve Samandar, A. (2005) Dissolved oxygen estimation using artificial neural network for water quality control, *Electronic Letters on Science and Engineering*, 1(2), 13-16.
34. Soyupak, S., Karaer, F., Gurbuz, H., Kivrak, E., Senturk, E. ve Yazici, A. (2003) A neural network-based approach for calculating dissolved oxygen profiles in reservoirs, *Neural Computing and Applications*, 12(3), 166-172. <https://doi.org/10.1007/s00521-003-0378-8>
35. Toprak, S. (2011) Çok değişkenli uyarlamalı regresyon eğrileri ve konik programlama ile zaman serilerinin modellenmesi, *Yüksek Lisans Tezi*, Dicle Üniversitesi, Fen Bilimleri Enstitüsü, Diyarbakır.
36. Ustaoglu, F., Tepe, Y. ve Tas, B. (2020) Assessment of stream quality and health risk in a subtropical Turkey river system: A combined approach using statistical analysis and water quality index, *Ecological Indicators*, 113, 105815. <https://doi.org/10.1016/j.ecolind.2019.105815>
37. Ünal, B. (2009) Çok değişkenli uyarlamalı regresyon uzanımları, *Yüksek Lisans Tezi*, Hacettepe Üniversitesi, Fen Bilimleri Enstitüsü, İstanbul.
38. Yalçın, H. ve Gürü, M. (2010) *Su Teknolojisi*, 2. Baskı, Palme Yayıncılık, Ankara.
39. Yerüstü Su Kalitesi Yönetmeliği (YSKY), 2012, Resmi Gazete, Başbakanlık Basımevi, 28483, 9-33.
40. Yilmaz, B., Aras, E., Nacar, S. ve Kankal, M. (2018) Estimating suspended sediment load with multivariate adaptive regression spline, teaching-learning based optimization, and artificial bee colony models, *Science of the Total Environment*, 639, 826-840. <https://doi.org/10.1016/j.scitotenv.2018.05.153>
41. Yuceer, M. ve Coskun, M.A. (2016) Modeling water quality in rivers: A case study of Beylerderesi River in Turkey, *Applied Ecology and Environmental Research*, 14(1), 383-395. [https://doi.org/10.15666/aeer/1401\\_383395](https://doi.org/10.15666/aeer/1401_383395)

