

Atık Su Miktarının ARIMA ve Yapay Sinir Ağları ile Tahmini

Estimation of Wastewater Amount with ARIMA and Artificial Neural Networks

Ayşegül YILDIZ , Sermin ELEVLİ , Mehmet Serhat ODABAŞ* 

***Makale Bilgisi / Article Info**

Alındı/Received: 27.08.2024

Kabul/Accepted: 16.10.2024

Yayımlandı/Published: 11.04.2025

Ondokuz Mayıs Üniversitesi, Lisansüstü Enstitüsü, Akıllı sistemler Mühendisliği Anabilim Dalı, Samsun, Türkiye

© Afyon Kocatepe Üniversitesi

© 2025 The Authors | Creative Commons Attribution-Noncommercial 4.0 (CC BY-NC) International License



Öz

Atık su akış tahmini, atık su arıtma tesislerinin doğru ve etkin bir şekilde yönetimi için anahtar rol oynamaktadır. Kontrolsüz şehirleşme, nüfus artışı, iklim değişikliğinden kaynaklı aşırı yağışlar ve altyapı yetersizlikleri gibi nedenlerden kaynaklanan tutarsız veri ve belirsizlikler atık su akış tahminini güçlendirmektedir. Bu kapsamda uzun vadeli eğilimleri kapsayacak etkili tahmin modellerinin kullanılması ihtiyacı belirgin hale gelmiştir. Bu çalışmada Samsun'un Doğu İleri Biyolojik Atık Su Arıtma Tesisi için atık su akış miktarının bir zaman serisi analiz modeli olan Otoregresif Bütünleşik Hareketli Ortalama (ARIMA) ve yapay sinir ağları ile tahmin edilmesi amaçlanmıştır. Bir yıllık süreye karşılık gelen günlük akış miktarı verileri kullanılan çalışmada modellerin performansları Kök Ortalama Kare Hatası (RMSE), Ortalama mutlak hata (MAE) ve Ortalama Mutlak yüzde Hatası (MAPE) değerleri açısından karşılaştırılmıştır. ARIMA (2, 1, 2) modeli daha yüksek doğrulukta performans göstermiştir.

Anahtar Kelimeler: ARIMA; Atıksu, Öngörü; Yapay Sinir Ağları.

Abstract

Wastewater flow estimation plays a key role for the accurate and efficient management of wastewater treatment plants. Inconsistent data and uncertainties arising from uncontrolled urbanization, population increases, excessive rainfall due to climate change and infrastructure deficiencies make wastewater flow forecasting difficult. In this context, the need to use effective forecasting models that will cover long-term trends has become evident. In this study, it is aimed to estimate the amount of wastewater flow for Samsun's East Advanced Biological Wastewater Treatment Plant with ARIMA, a time series analysis model, and artificial neural networks. Daily flow rate data corresponding to a period of one year were used and the performances of the models were compared in terms of Root Mean Square Error (RMSE), Mean Absolute Error (MAE), and Mean Absolute Percentage Error (MAPE) values. ARIMA (2, 1, 2) model showed higher accuracy.

Keywords: ARIMA; Wastewater; Forecasting; Artificial Neural Networks.

1. Giriş

Yaşamın kaynağı olan su insanların temel ihtiyaçları arasındadır. Nüfus artışı ve sanayileşmenin sebep olduğu ve doğanın absorbe edebileceğinin üzerinde gerçekleşen atık su üretimi, insan hayatını olumsuz yönde etkileyen problemlerden biridir. Dolayısıyla suyun kullanımı ve atık su miktarı yönetimi çevre bilinci ve enerji tasarrufu açısından oldukça önemlidir (Lai vd. 2023). Atık suyun etkin politikalarla verimli bir şekilde yeniden kullanılması, sürdürülebilir su yönetimi açısından büyük bir öneme sahiptir. Atık su miktarının tahmini, bu politikaların geliştirilmesi ve uygulanmasında kritik bir rol oynar (Çelik vd. 2022). Atık su, temiz suyun evsel, endüstriyel veya tarımsal kullanımdan sonra çeşitli biyolojik, fiziksel veya kimyasal maddelerle kirlenmesinin sonucudur. Son yıllarda dünya nüfusunun ve beraberinde endüstriyel faaliyetlerin önemli ölçüde artışı, doğal su döngüsüne geri döndürülmeden önce arıtılması gereken atık su miktarında da artışına neden olmuştur. Atık su arıtma

tesislerinde arıtma süreçleri esnasında kirleticileri gidermek, arıtılmış su yasal limitlerine ulaşmak ve enerji tüketimini azaltmak için atık su akış miktarı ve kalite parametrelerinin sürekli olarak izlenme ihtiyacı bulunmaktadır. Nitekim yeterli düzeyde arıtılmamış atık suyun deşarjı, nehirlerin ve akarsuların kendi kendini temizleme sürecini azaltmak gibi olumsuzlukları da beraberinde getirmektedir. Bununla birlikte atık su arıtma sürecinin dinamik ve doğrusal olmayan yapısı, tesislerin işletilmesi ve yönetilmesi konusunda belirsizliklere yol açabilmektedir (S. Al-Dahidi vd. 2024).

Günümüz dünyasında geri dönüşümün hızla artan önemi ve su kıtlığı problemi, atık su arıtma faaliyetlerinin iyileştirilmesine olan ihtiyacı artırmıştır. Atık su arıtma tesislerinde göl, deniz, toprak ve diğer kaynaklara deşarj edilen arıtılmış suların etkin bir şekilde yönetilebilmesi için atık su miktarının doğru tahmin edilmesi gerekmektedir. Bu tesislerde istikrarlı ve sürdürülebilir su kalitesine ulaşılması için atık su miktarı yol gösterici

olmakla kalmaz, aynı zamanda kapasite risklerinin saptanmasına ve mevcut sistemin iyileştirilmesine de yardımcı olabilir. Bu iyileştirmeler tesis işletiminin daha pratik, güvenilir ve ekonomik hale getirilmesini sağlar (Baki ve Aras, 2018). Tasarım parametrelerine uygun çalışılması, kaynakların etkin kullanılması ve uygun çevresel standartlara ulaşılması açısından atık su akış miktarlarının tahmin edilmesi, atık su arıtma tesislerinde önemli bir görev haline gelmiştir. Gelecek dönem atık su tahmini, arıtma süreçlerini optimize etmeye, böylece olası çevresel riskleri azaltmaya ve arıtma verimliliğini arttırmaya katkı sağlayacaktır (Zhang vd. 2019).

Otoregresif Bütünleşik Hareketli Ortalama (ARIMA) modelleri, atık su ile ilgili miktar ve parametre tahminleri için yaygın olarak kullanılan bir teknik olmuştur (Do vd. 2022). Regresyon modellerinin aksine ARIMA modeli, geçmiş veya gecikmeli değerlerini, hata terimleri ile birlikte kullanarak sonuç üretir (Rahman ve Hasan 2017). ARIMA modeli, otoregresif (AR), hareketli ortalama (MA) ve durağanlığa ulaşmak için fark (I) alma işleminin entegrasyonu sonucu oluşan bir zaman serisi formudur. ARIMA doğrusal olarak yapılandırıldığı için, verilerin doğrusal bir model izlediği varsayımına dayanır. Ancak, gerçek dünya verileri sadece doğrusal bir model oluşturmaz. Bu nedenle bazı durumlarda, doğrusal yaklaşımın tahmin performansı zayıf kalabilir (Suhermi vd. 2018).

Yapay Sinir Ağı (YSA), biyolojik sinir sistemlerinin işleyişini taklit eden ve öğrenme, desen tanıma, sınıflandırma gibi görevlerde kullanılan hesaplama modelleridir (Kaya vd. 2023; Palabıyık ve Akkan, 2024). Yapay sinir ağları çok katmanlı algılayıcı bir yapıya sahiptir. Bu çok katmanlı algılayıcı modeli, bir giriş katmanı, bir veya daha fazla gizli katmanı ve bir çıkış katmanı içerir. Her bir katmanda en az bir işlem elemanı bulunur. Bir üst katmandaki her işlem elemanı bir alt katmandaki her işlem elemanı ile bağlantılıdır. Bilgi sadece ileri doğru akar ve giriş katmanında herhangi bir veri işlenmez. Giriş ve çıkış katmanlarında bulunan işlemci elemanı sayısı, uygulanan probleme bağlıdır. Ara katmanlardaki işlemci sayısı ise deneme yanılma yoluyla hesaplanır. Ağ, girdilere karşılık doğru çıktıları öğrenerek verilerle eğitim yapmış olabilir. Böylece ağırlıklarını ayarlar ve belirli görevleri yerine getirebilir (Ataseven 2013; Işık vd. 2024). YSA çeşitli doğrusal olmayan problemleri modellemek için esnek hesaplama yapılarıdır (Odabas vd. 2017). YSA modellerinin diğer doğrusal olmayan model sınıflarına göre önemli bir üstünlüğü, geniş bir fonksiyon sınıfına yüksek doğrulukla yaklaşabilen evrensel yaklaşımıcı olmalarıdır (Zhang 2003; Palabıyık ve Akkan, 2024). Yapay Sinir Ağları, görüntü tanıma, doğal dil işleme ve genel

regresyon görevleri gibi çeşitli uygulamalarda yaygın olarak kullanılan çok yönlü modellere sahiptir. Farklı veri türlerini işleme esnekliği sayesinde çeşitli girişleri yönetebilirler. Buna karşın, Doğrusal Olmayan Otonom Sinir Ağları (NAR) özellikle zaman serisi tahmini için tasarlanmıştır ve geçmiş gözlemleri kullanarak gelecekteki değerleri tahmin eder. Bu otoregresif yaklaşım, zamansal bağımlılıkları doğal olarak yakaladığı için NAR ağları, verinin sıralamasının kritik olduğu görevlerde son derece etkili hale gelir. Her iki ağ türü de benzer eğitim yöntemlerine dayanmasına rağmen, NAR ağları geçmiş sıralar biçiminde giriş verilerinin dikkatlice hazırlanmasını gerektirir, bu da onların zamana bağlı kalıplara odaklanan özel bir işlevselliğe sahip olduğunu vurgular (Namasudra vd. 2023).

Atık su ile ilgili geleceğe yönelik tahmin çalışmalarında atık su parametreleri ve akış miktarı ele alınmıştır. Örneğin, Baki ve Aras (2018) tarafından yapılan çalışmada, BOİ parametresi değerleri, diğer kirletici parametrelere bağlı olarak tahmin edilmiştir. En iyi tahmin modeli olarak seçilen Çok Değişkenli Uyarlanabilir Regresyon Eğrileri modelinde, en düşük ortalama hata karesi karekökü (RMSE) ve ortalama mutlak hata verileri (MAE) sırasıyla, 47,7378 ve 37,7566 olarak bulunmuştur. Çelik vd. (2022)'ye göre farklı kaynaklara deşarj edilen atık suların tahmini için ARIMA zaman serisi analizini kullanmıştır. Do vd. (2022) tarafından, mevsimsellik içeren atık su akışı zaman serisi için mevsimsel (S) ARIMA modeli kullanılmıştır. SARIMA (1,0,3) (2,1,2)₂₄, ortalama hata karesi karekökü (RMSE = 5,508) ve ortalama mutlak yüzde hata (MAPE = 20,78) performans göstergeleri açısından atık su tahmini için en iyi model olarak bulunmuştur. Zhang vd. (2019) tarafından yapılan çalışmada, atık su akışını tahmin etmek için ARIMA ve çok katmanlı yapay sinir ağı (MLPNN) modeli kullanılmıştır. MLPNN ve ARIMA modeline ait RMSE değerleri sırasıyla 115,85 ve 109,03 olarak bulunmuştur. Lai vd. (2023)'göre, COVID-19 vakalarını atık su ve çevresel değişkenlerden tahmin etmek için bir zaman serisi makine öğrenimi (TSML) yöntemi kullanmıştır. TSML modeli uzun dönem tahminlerinde RMSE (20,95), MAE (15,61) ve MAPE (112,45) değerleri ile ARIMAX (dışsal değişkenli ARIMA)'dan daha iyi performans göstermiştir. Al-Dahibi vd. (2024) tarafından yapılan çalışmada, atık su arıtma tesisinin performansı için tahmin yapılmış ve destek vektör makinesi (SVM), MAE'ye göre 8,8 değeri ile en iyi tahmin performansını vermiştir.

Bu çalışmada Samsun ili nüfusunun yarıya yakınına hizmet veren ve kapladığı yüz ölçümü sebebiyle Karadeniz Bölgesi'nin en büyük arıtma tesisi olma özelliği gösteren Doğu İleri Biyolojik Atık Su Arıtma Tesisi için atık su akış

miktarı tahmini yapılması amaçlanmıştır. Bu tahmin için karşılaştırma yapabilmek ve en uygun modeli seçebilmek için bir zaman serisi analiz modeli olan ARIMA modeli ve YSA ile kullanılmıştır. Model performansları MAPE (Ortalama Mutlak Yüzde Hata), RMSE (Ortalama Hata Karesi Karekökü), MAE (Ortalama Mutlak Hata) üzerinden değerlendirilmiştir.

2. Materyal ve Metot

2.1 Çalışma Alanı

Samsun Doğu İleri Biyolojik Atık Su Arıtma Tesisi (SDİBAT), 1 Eylül 2014'te faaliyete geçmiş ve Samsun ili nüfusunun % 44'üne hizmet vermektedir. Tesis, 220 dönümlük bir alanda günlük ortalama 105000 m³ atık suyu arıtarak yaklaşık 2,5 kilometre açık denize deşarj etmektedir (Şekil 1). Başlangıçta 4 ilçe belediyesine hizmet veren tesis, 31 Mart 2014'ten itibaren 13 ilçe belediyesinin eklenmesiyle tüm il sınırına hizmet sunmaktadır. Bu özellikleriyle Karadeniz Bölgesi'nin en büyük atık su arıtma tesisidir. Çevre ve Şehircilik Bakanlığı tarafından hassas bölge

olarak belirlenen Bafra-Ünye sahil bandında üretilen atık sular, yönetmelik gereğince ileri biyolojik yöntemlerle arıtılmaktadır. Otomatik olarak işletilen tesis, arıtılmış suları 23,5 m derinlikte ve 2,45 km mesafede denize deşarj etmektedir. Akış miktarı öngörüsü için SDİBAT'tan elde edilen 2023 yılı günlük atık su giriş debisi (m³) verilerinin özet istatistikleri Çizelge 1'de sunulmuştur.

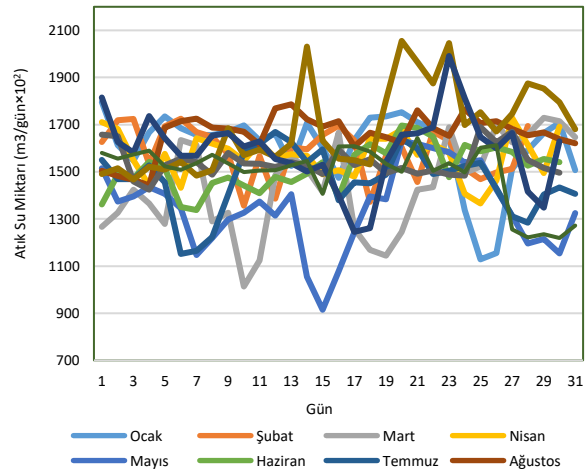


Şekil 1. Samsun doğu ileri biyolojik atık su arıtma tesisi

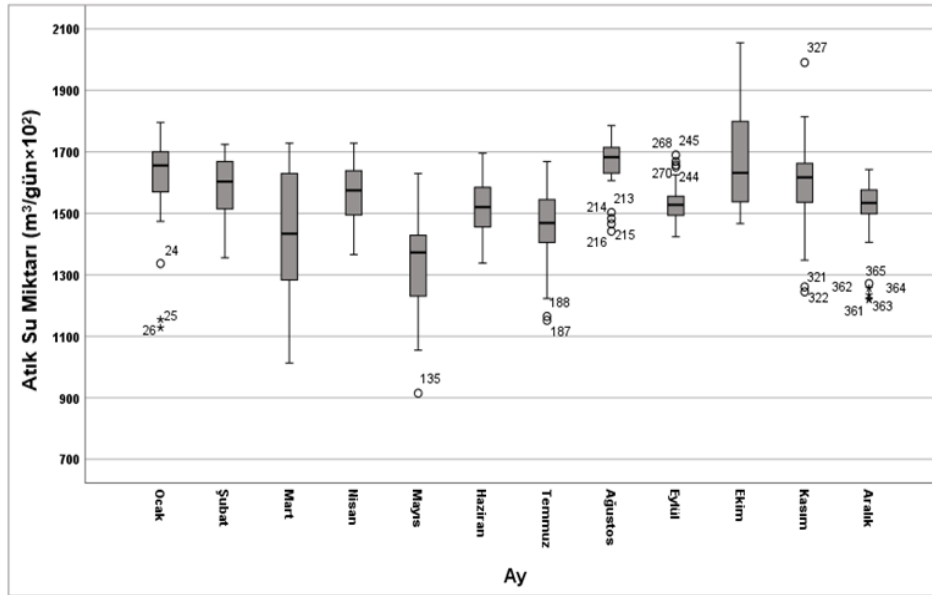
Çizelge 1. 2023 yılı günlük atık su verileri özet istatistikleri

Aylar	Gözlem sayısı	Ortalama	Standart Sapma	Q1	Q2 (medyan)	Q3	Min.	Max.	R
1	31	160703,00	15598,50	154759,00	165648,00	170605,00	112842,00	179580,00	66738,00
2	28	158562,00	10734,32	151334,00	160353,50	166932,50	135613,00	172509,00	36896,00
3	31	144375,00	19897,95	127837,00	143438,00	163380,00	101319,00	172897,00	71578,00
4	30	156975,00	9289,03	149337,75	157511,00	163969,75	136563,00	172862,00	36299,00
5	31	134736,00	17349,75	121873,00	137318,00	143481,00	91461,00	163018,00	71557,00
6	30	151632,00	9432,99	145507,75	152047,50	158525,00	133755,00	169630,00	35875,00
7	31	146395,00	13192,10	140377,00	146904,00	155038,00	115158,00	166894,00	51736,00
8	31	165986,00	8604,18	162409,00	168265,00	171520,00	144211,00	178624,00	34413,00
9	30	154108,00	6389,29	149371,00	152798,00	156172,75	142404,00	168986,00	26582,00
10	31	168337,00	17959,60	153158,00	163229,00	180393,00	146726,00	205526,00	58800,00
11	30	159279,00	15523,15	153499,75	161683,50	166305,50	124528,00	199130,00	74602,00
12	31	149674,00	12344,38	149907,00	153399,00	158049,00	121894,00	164326,00	42432,00

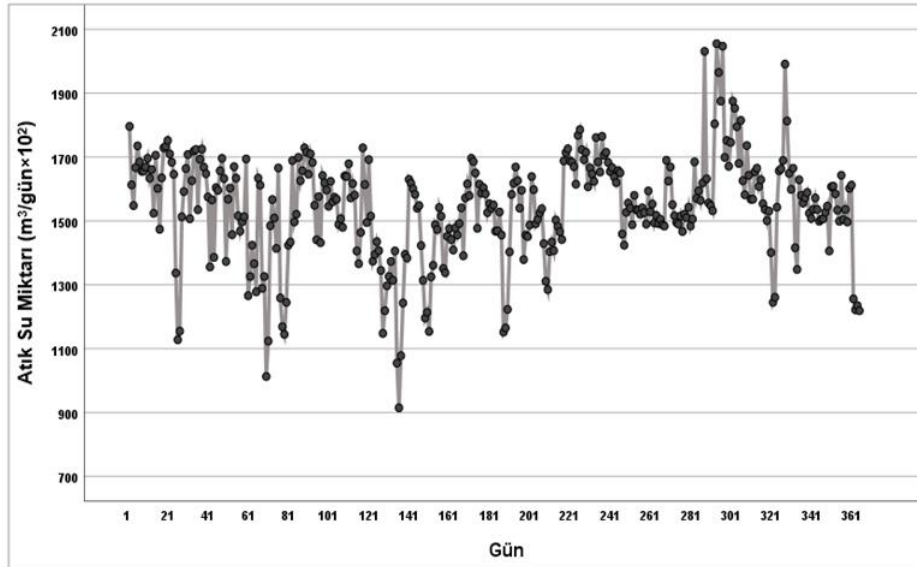
SDİBAT'ın 2023 yılı için günlük atık su giriş debisi (m³) verileri kullanılarak oluşturulan ve aylık bazda değişimlerin seyrini veren grafik Şekil 2'de görülmektedir. Özellikle meteorolojik gelişmelerin etkisiyle akış miktarında bazı aylarda ciddi bir azalma bazı aylarda ise artış göze batmaktadır. Ani değişimler özellikle Ocak, Mart, Mayıs, Temmuz, Ekim ve Kasım aylarında daha belirgin bir şekilde görülmektedir. Atık suyun aylar bazında günlük debilerinin merkezsel konumu ve yayılmasını gözlemlemek için oluşturulan kutu grafiği Şekil 3'de görülmektedir. Bu grafikte verilerin %50'sinin yer aldığı kutu yükseklikleri ve medyan (ortanca) değerleri, aylar bazında debilerin önemli ölçüde değişkenlik içerdiğine işaret etmektedir. Ayrıca Ocak, Mayıs, Temmuz, Ağustos, Eylül, Kasım ve Aralık aylarında aykırı değerler göze çarpmaktadır.



Şekil 2. 2023 yılı Samsun atık su miktarının aylara göre dağılımını gösteren çizgi grafiği



Şekil 2. 2023 yılı Samsun atık su miktarının aylara göre dağılımını gösteren kutu grafiği



Şekil 4. 2023 Yılı Atık Su Miktarı Zaman Serisi

Samsun Doğu İleri Biyolojik Atık Su Arıtma Tesisi 'ne ait 2023 yılı atık su zaman serisi Şekil 4'de görüldüğü gibidir. Ani iniş ve çıkışların olduğu, belirli bir mevsimsellik ve trend içermeyen bu seri durağan değildir.

2.2 Metot

2.2.1. Otoregresif Bütünleşik Hareketli Ortalama (ARIMA)

Belirli bir zaman diliminde toplanan verilerden oluşan zaman serisinin analizinde yaygın olarak kullanılan ARMA (p, q) modeli iki bileşene sahiptir. Otoregresif (AR) bileşen, mevcut zaman serisinin önceki değerlerinin mevcut değeri etkilediği kısımdır. AR(p) modeli, p parametresi kadar önceki değerlere dayanan bir modeldir. Hareketli Ortalama (MA) bileşeni ise mevcut zaman serisindeki hataların geçmiş değerlerinin mevcut değeri etkilediği

kısımdır. MA(q) modeli, q kadar önceki hata terimlerine dayanan bir modeldir. Bütünleşik bir model olan ARIMA modeli ise durağan olmayan zaman serilerini modellemek için zaman serisi analizinde kullanılan genişletilmiş bir modeldir. ARIMA (p, d, q) modeli, ARMA modeline seriyi d kez fark alarak durağan hale getirmeyi ifade eden I(d) teriminin eklenmesiyle elde edilir (Elevli 2020). Bir ARIMA modelinde, bir değişkenin gelecekteki değerinin, birkaç geçmiş gözlemin ve rastgele hataların doğrusal bir fonksiyonu olduğu varsayılır (Zhang 2003). ARIMA modellerinin Eşitlik (1)'de verilen genel matematiksel gösteriminde α sabit değeri, β gözlem değerleri için katsayıları, Y_t t anındaki zaman serisinin değerini, ϕ hata terimlerinin katsayılarını, ϵ hata terimlerini, p otoregresif katsayısını ve q değeri ise hareketli ortalama parametresini göstermektedir (Erden 2023).

$$Y_t = \alpha + \beta_1 Y_{t-1} + \beta_2 Y_{t-2} + \dots + \beta_p Y_{t-p} + \varphi_1 \varepsilon_{t-1} + \varphi_2 \varepsilon_{t-2} + \dots + \varphi_q \varepsilon_{t-q} \quad (1)$$

ARIMA modellerinin tasarlanmasına başlanmadan önce durağanlığın kontrol edilmesi gerekmektedir. ARIMA modelleri bünyesinde bulundurduğu fark alma işlemi sayesinde durağanlaştırma yeteneğine sahip olduğu için, durağan olmayan serilerde de kullanılır.

Box-Jenkins modeli olarak bilinen ARIMA modeli, üç aşamada tamamlanır. Birinci aşamada, mevcut zaman serisinin durağan bir yapıda olup olmadığını anlamak için otokorelasyon (ACF) ve kısmi otokorelasyon (PACF) fonksiyonlarının grafikleri kullanılır. İkinci aşamada, ACF ve PACF grafiklerinden yola çıkarak p, d ve q değerleri belirlenir ve model tahmini yapılır. Seçilen ARIMA modeline en iyi uyumu sağlayan parametreleri belirlemek için çeşitli hesaplama algoritmaları kullanılır. Parametre tahmini yaparken en yaygın kullanılan yöntemler arasında Maksimum Olasılık Tahmini (MLE) ve Doğrusal Olmayan En Küçük Kareler (Non-Linear Least Squares) yer alır. Bu yöntemler, modelin parametrelerinin en uygun değerlerini hesaplamada etkilidir. Üçüncü aşamada ise tahmin edilen modelin durağan tek değişkenli bir sürecin özelliklerine uyup uymadığının test edilmesi, istatistiksel modelin kontrol edilmesi ve tahmin uygulamasının yapılması yer almaktadır. ARIMA'nın model seçiminde farklı (p, d, q) kombinasyonları sonucu elde edilen performans kriterlerinin karşılaştırılması gerekmektedir. Ortalama hata karesi karekökü (RMSE), ortalama yüzde hata (MPE), ortalama mutlak yüzde hata (MAPE), ortalama mutlak hata (MAE), ortalama mutlak ölçekli hata (MASE) ve Bayesian bilgi kriterleri (BIC) olan seçim kriterlerinin minimum değerine göre seçim yapılmalıdır (Rahman ve Hasan 2017).

2.2.2. Yapay Sinir Ağları (YSA)

İnsanın bedeni ve özellikle beyin bölümünün fizyolojik yapısı incelendiğinde sinir hücrelerinin nasıl hareket ettiği ve verileri birbirine nasıl aktardıkları konusu bilimsel araştırmaların konusu olmuştur. İnsandaki bu sinir hücrelerinin yapısı taklit edilmeye çalışılarak makine öğrenmesi adı altında bir alan geliştirilmiştir. Makine öğrenmesi altında ele alınan yapay sinir ağları, beyin bir işlevini yerine getirmek için izlediği yolu modellemek ve onu taklit etmek için tasarlanmıştır (Er Aydın vd. 2022).

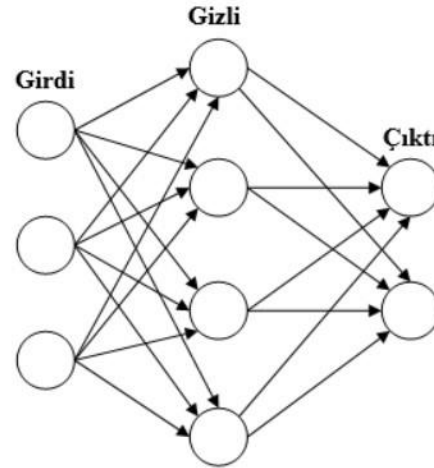
Yapay sinir ağları, geçmiş deneyimlerden bilgi elde ederek tahmin modelleri ve bir çıktı oluşturmak için kullanılan bir makine öğrenmesi tekniğidir (Do vd. 2022). Makine öğrenmesi ve derin öğrenme gibi alanlarda veri örüntülerini öğrenme ve genelleme yetenekleri yaygın olarak kullanılmaktadır. YSA'lar, çok sayıda düğümden (neuron) oluşan katmanlar aracılığıyla veriyi işlemek ve

belirli bir çıktı üretmek için çalışır. Her nöron, bir ağırlık (weight) ve bir aktivasyon fonksiyonu kullanarak girdilerini işler. Aktivasyon fonksiyonları, nöronun aldığı toplam girdiği işleyip çıktı haline getiren matematiksel fonksiyonlardır. Bahsi geçen ağırlık ise mevcut sinir ağının öğrenme kapasitesini temsil eden ve girdi verilerinin nöronlar arasındaki iletimini etkileyen parametrelerdir. Ağırlıklar, öğrenme süreci boyunca ayarlanarak ağın performansı optimize edilir. Şekil 5'de gösterilen ağın içerisindeki katmanların açıklaması aşağıda verilmiştir:

Giriş Katmanı (Input Layer): Ham verinin ağa ilk kez sunulduğu katmandır.

Gizli Katmanlar (Hidden Layers): Girdi verilerini işleyen ve öğrenme sürecinin çoğunun gerçekleştiği katmanlardır. Çeşitli sayıda gizli katman sayısı olabilmektedir.

Çıkış Katmanı (Output Layer): Çıktıların üretildiği katmandır.



Şekil 5. Temel bir yapay sinir ağı modeli ve katmanları

Atık su miktarlarının tahmini için 365 günlük giriş verisi olarak YSA sistemine tanımlanmıştır. Toplamda 365 adet veri alınmıştır. Bu verilerin %70'i 255 veri olarak öğrenme, %30'u ise (108) test ve doğrulamada kullanılmıştır. Analiz için gerekli kod yazımı ve analiz aracı olarak MATLAB 2023a programının Neural Network Time Series aracı ile gerçekleştirilmiş olup gerekli kodlar yazılmıştır. Yapay sinir ağı modelini geliştirmek için, orijinal veri kümesi %70 eğitim, %15 test ve %15 doğrulama olacak şekilde ağın eğitiminde kullanılmıştır. Gizli katman sayısı ise 10 olarak belirlenmiştir. Bu çalışma kapsamında Levenberg-Marquardt (LM) geri yayılım algoritması kullanılmıştır. Levenberg-Marquardt algoritması, genellikle yapay sinir ağlarının eğitimi ve doğrusal olmayan en küçük kareler problemlerinin çözümü için kullanılan bir optimizasyon

algoritması olup hızlı bir şekilde sonuç vermektedir. LM algoritması maksimum komşuluk düşüncesi üzerine temellenmiştir. Bu algoritma, diğer öğrenme algoritmalarının en iyi özelliklerini barındırır (Ataseven 2013). ARIMA ve YSA gibi farklı öngörü modellerinin karşılaştırılmasında tahmin hatasını içeren doğruluk ölçülerinin kullanılması gerekmektedir. En yaygın kullanılan ölçüler arasında yer alan MAPE (Ortalama Mutlak Yüzde Hata), RMSE (Ortalama Hata Karesi Karekökü) ve MAE (Ortalama Mutlak Hata) Eşitlik (2)-(4)'te verilmiştir. Tüm doğruluk ölçüleri geçmiş verilere dayanmakta olup, farklı modeller için bu ölçülerin en düşük olduğu öngörü modelinin seçilmesi gerekmektedir.

$$MAPE = \frac{\sum |(A_t - F_t)/A_t|}{n} \quad (1)$$

$$RMSE = \sqrt{\frac{\sum (A_t - F_t)^2}{n}} \quad (2)$$

$$MAE = \frac{\sum |A_t - F_t|}{n} \quad (3)$$

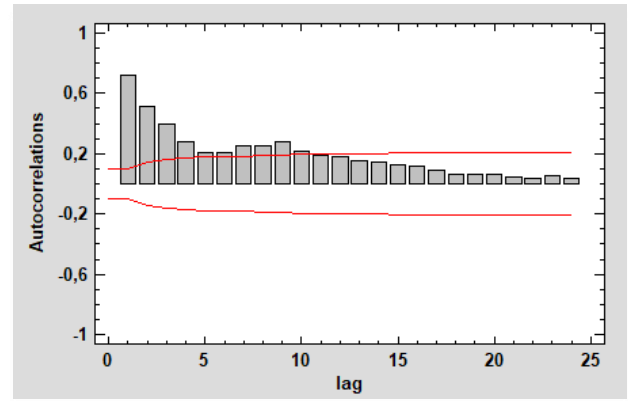
Formüllerde gösterilen, A_t gerçek değeri, F_t tahmin edilen değeri, n ise toplam gözlem sayısını ifade etmektedir.

3. Bulgular

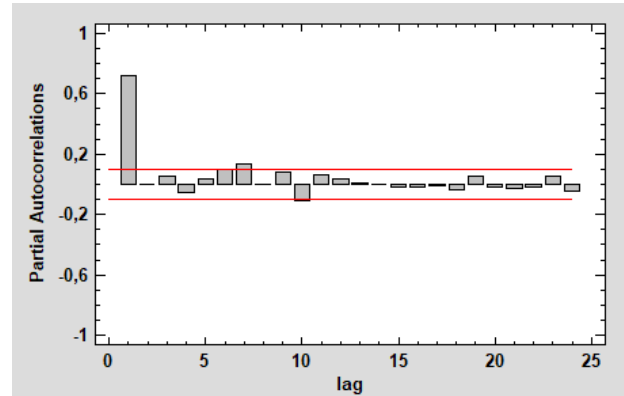
Bu çalışma temel olarak iki aşamadan oluşmaktadır. Birinci aşamada zaman serisi analizi yapılarak Statgraphics 19 yazılımı deneme sürümü ile ARIMA modeli kurulmuştur. İkinci aşamada ise Matlab 2023a yazılımı ile yapay sinir ağı modeli oluşturulmuştur. Böylece doğrusal ve doğrusal olmayan ilişkileri ele almakta kullanılabilen iki farklı yöntemden yararlanılmıştır.

3.1. ARIMA Modelinin Sonuçları

ARIMA modeli oluşturulurken zamana göre art arda elde edilmiş gözlem kümesinin (serinin) durağan olup olmadığının ve model parametrelerinin belirlenmesinde, otokorelasyon ve kısmi otokorelasyon fonksiyonlarının grafiklerinden yararlanılmıştır. Otokorelasyon katsayısı, zaman serisinde farklı zaman aralıklarına (gecikme) sahip değerler arasındaki ilişkinin, kısmi otokorelasyon katsayısı ise zaman serisinin diğer gecikmeli serilerin etkileri ihmal edilerek ilgili gecikmeye ait zaman serisi ile arasındaki ilişkinin katsayısını göstermektedir. Atık su akış verileri için elde edilen otokorelasyon ve kısmi otokorelasyon fonksiyonu grafikleri sırasıyla Şekil 6 ve Şekil 7'de görülmektedir. Kural olarak, 1. gecikmede yüksek bir otokorelasyon ve takip eden birkaç gecikmede yavaş yavaş azalan bir eğilim olması, verilerin durağan olmadığına işaret etmektedir.

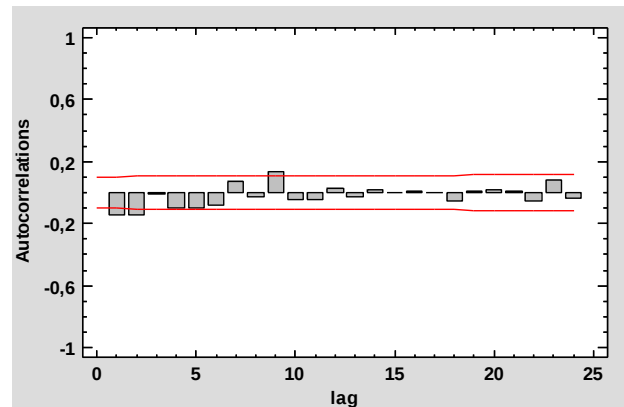


Şekil 6. Otokorelasyon değerleri

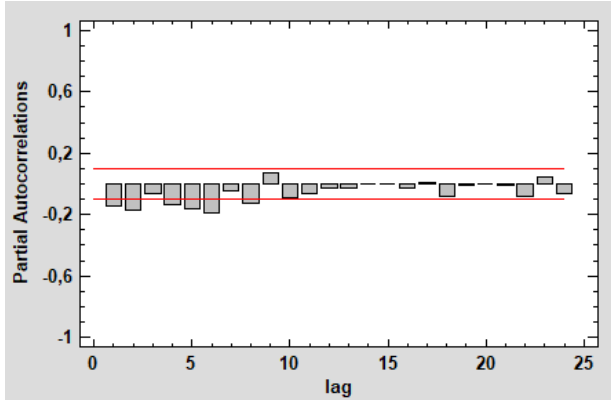


Şekil 3. Kısmi otokorelasyon değerleri

Şekil 6'daki otokorelasyon grafiği serinin durağan olmadığı yönünde yorumlandığından, serinin farkları alınıp ortalama, varyans ve otokorelasyon gibi istatistiksel özelliklerin zaman içerisinde değişmediği durağan bir zaman serisine dönüştürülmüştür. Serinin farkları alındıktan sonra elde edilen otokorelasyon ve kısmi otokorelasyon fonksiyonu grafikleri ise sırasıyla Şekil 8 ve Şekil 9'da verilmektedir. Kısmi otokorelasyon fonksiyonunda ilk 1 ve 2. gecikmelerin sınır dışında kalması, serinin 1. dereceden farkının alınması ile durağan hale gelmesi ve otokorelasyon fonksiyonunda ilk iki gecikmenin sınırın dışında kalması nedeni ile ARIMA (2,1,2) modeli uygun bir model olarak öngörülmüştür.



Şekil 4. Durağan hale getirilmiş zaman serisinin otokorelasyonu



Şekil 5. Durağan hale getirilmiş zaman serisinin kısmi otokorelasyonu

Durağanlığın ve ARIMA derecelerinin grafiksel yorumu dayalı tespiti büyük ölçüde sübjektif kararlara dayandığı için farklı modeller de karşılaştırma amacıyla kullanılmıştır. Karşılaştırılacak model alternatiflerinin belirlenmesinde durağanlık öncesi ve sonrası otokorelasyon ve kısmi otokorelasyon grafiklerine bakılarak sınır dışı kalan gecikme sayıları dikkate alınmıştır. Karşılaştırılan ARIMA modellerine ait hata ölçütleri Çizelge 2'de yer almaktadır.

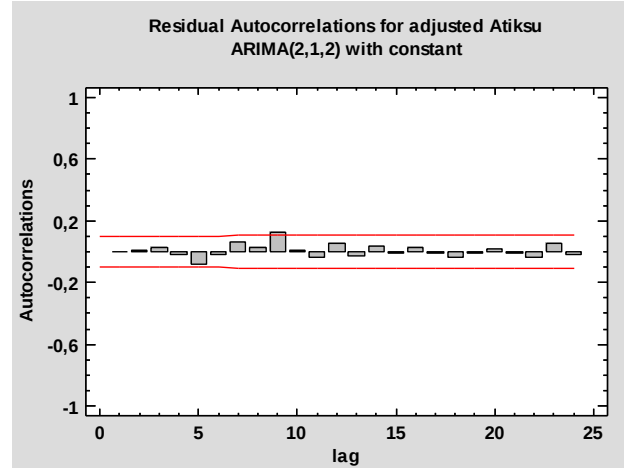
Çizelge 2. Karşılaştırılan ARIMA modellerinin hata ölçütleri

Model	RMSE	MAE	MAPE
ARIMA (1,0,4)	11354,10	8048,50	5,41
ARIMA (1,0,5)	11355,50	8022,87	5,39
ARIMA (6,1,2)	11321,20	8122,96	5,44
ARIMA (6,1,0)	11452,80	8242,94	5,52
ARIMA (2,1,2)	11244,60	7996,35	5,37

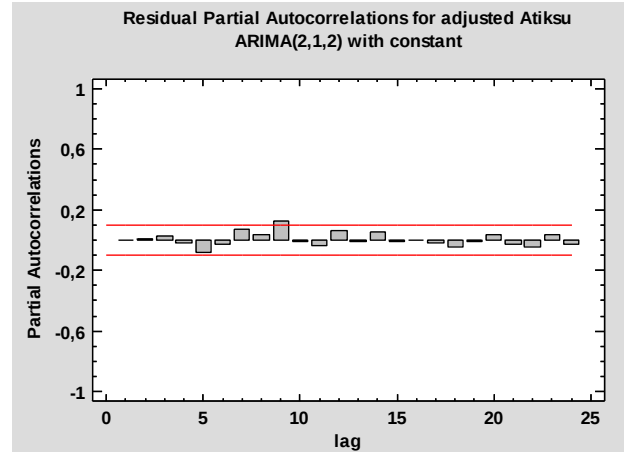
ARIMA modelinin hata ölçütlerini içeren Çizelge 2, beş farklı tahmin modelinin sonuçlarını karşılaştırmaktadır. Hata ölçütlerine bakıldığında, tahmin dönemi boyunca RMSE, MAE ve MAPE değerlerinin en küçük olduğu model ARIMA (2,1,2) modeli en uygun model olarak tespit edilmiştir. Bu modelin hata terimlerine ait otokorelasyon grafiği (ACF) Şekil 10'da ve kısmi otokorelasyon grafiği (PACF) Şekil 11'de gösterilmektedir. Bu grafikler hata terimlerinin rastgele dağılmış olduğunu ve belirgin bir otokorelasyon yapısı göstermediğini ortaya koymaktadır. Bir başka deyişle, ARIMA modelinin veriyi iyi yakaladığını ve atık su miktarının gelecekteki tahminlerinin güvenilir olduğunu göstermektedir.

ARIMA (2,1,2) modelinin tahmin edilen parametreleri ve parametrelerin anlamlılığı Çizelge 3'de verilmiştir. AR (2) ve MA (2) terimleri için P-değerinin 0,05'ten küçük olması istatistiksel açıdan öneme işaret etmektedir.

Seçilen ARIMA modeli ile elde edilen 12 günlük gelecek dönem tahminleri, alt ve üst güven sınırları ile birlikte Çizelge 4'de yer almaktadır.



Şekil 6. ARIMA (2,1,2) modeli otokorelasyon grafiği



Şekil 11. ARIMA (2,1,2) modeli kısmi otokorelasyon grafiği

Çizelge 3. ARIMA modeli özeti

Parametre	Tahmin	Standart Hata	T	P-Değeri
AR(1)	-0,277	0,0495	-5,598	0,0000
AR(2)	0,625	0,0417	15,017	0,0000
MA(1)	-0,032	0,0183	-1,791	0,0741
MA(2)	0,964	0,0039	246,137	0,0000
Mean	-33,407	67,8343	-0,4925	0,6227
Constant	-21,768			

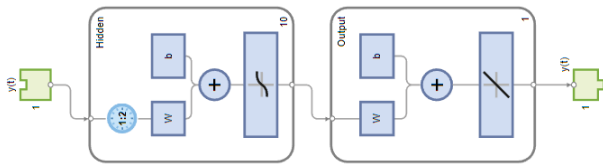
Çizelge 4. ARIMA (2,1,2) modelinin gelecek dönem tahmin değerleri

Dönem	Tahmin	Alt Sınır 95%	Üst Sınır 95%
366	135397	113261	157534
367	139432	111686	167178
368	143417	113670	173163
369	144814	113735	175893
370	146898	115296	178500
371	147173	115086	179260
372	148379	116096	180661
373	148195	115665	180725
374	148978	116342	181614
375	148624	115823	181425
376	149190	116313	182068
377	148790	115781	181800

3.2. Yapay Sinir Ağı Sonuçları

Yapay sinir ağları (YSA), yapay zekâ alanının bir alt dalını oluşturur. İnsan beyninin çalışma prensiplerini örnek alan ve tıpkı gerçek bir sinir hücresi gibi aldığı sinyalleri bir başka sinir hücresine ileten bu yapay sistemler, öğrenme ve hızlı işlem yapabilme yetenekleri sayesinde bilim dünyasında popülerliğini kazanmıştır. Yapay sinir ağları öğrenme yoluyla bilgi ve deneyimlerini arttırmakta ve öğrenilen bilgilerden yararlanarak bir çıktı üretme prensibi ile çalışmaktadır (Öztemel ve Dügenci 2016).

Bu çalışma kapsamında kullanılan veri setine uygun ağ seçeneği NAR Network olduğu için analize, bu seçenek seçilerek devam edilmiştir ve genel gösterimi Tablo 5'te verilmiştir. Model eğitildikten sonra çıkan analiz sonucu özeti Çizelge 5'de, zaman serisi ve tahmin sonuçlarının birlikte olduğu grafik Şekil 12'de yer almaktadır.



Şekil 7. Doğrusal olmayan otoregresif sinir ağı (NAR)

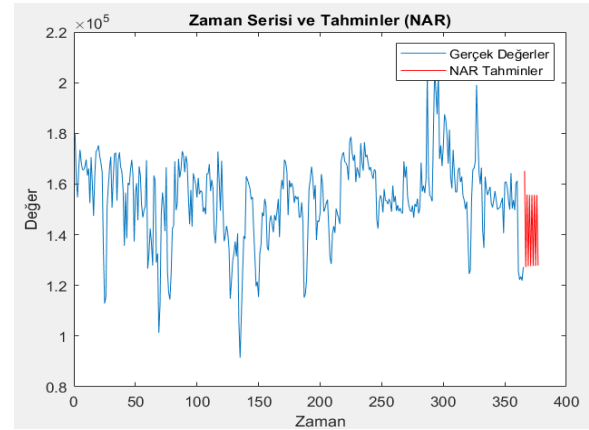
Doğrusal olmayan otoregresif sinir ağı girdi verileri ve gizli katman aracılığıyla doğrusal olmayan bir işlemden geçerek bir çıktı değeri üretir ve sonuçlanır. Bu genel yapı, zaman serisinin gelecek değerlerini tahmin eder. Çalışmamızda NAR (Nonlinear Autoregressive) Network'ün tercih edilmesinin temel nedeni, atık su miktarının tahmin edilmesinde zaman serisinin kendi geçmiş değerlerinin belirleyici olmasıdır. NAR Network, yalnızca geçmiş veri değerleri ile çalışarak dışsal faktörler olmadan tahmin yapabilme yeteneğine sahiptir. Bu özellik, atık su miktarı gibi dışsal etmenlerden sınırlı şekilde etkilenen tek değişkenli zaman serileri için ideal bir seçim olmasını sağlamaktadır.

Ayrıca, NAR Network'ün Levenberg-Marquardt algoritması ile hızlı yakınsama ve yüksek performans gösterme yeteneği, modelin etkinliğini artırmaktadır. MATLAB yazılımının sağladığı entegre araçlar ve kapsamlı dokümantasyon desteği, modelin uygulanabilirliğini ve geliştirme sürecini kolaylaştırmaktadır. NAR Network'ün benzer problemler için sağladığı kanıtlanmış başarılarında, bu modelin çalışmamız için en uygun ağ yapısı olarak seçilmesinde etkili olmuştur. Analizin özetinde hangi tür zaman serisinin eğitildiği, verilerin özellikleri, hangi eğitim algoritmasının kullanıldığı, model performansının hangi ölçütlerle ölçüldüğü, katman sayısı, test ve eğitim verilerinin nasıl paylaştırıldığı bilgileri yer almaktadır (Çizelge 5).

Çizelge 5. Yapay sinir ağı model özeti

	Veri sayısı	MSE	R ²
Eğitim	255	1.0800 e-8	0.62
Doğrulama	54	1.3795 e-8	0.51
Test	54	1.5734 e-8	0.52

2023 yılı atık su verilerine ait zaman serisi ve yapay sinir ağı modelinin gelecek dönemler için bulduğu tahmin değerlerinin birleştirilmesi ile elde edilen grafik Şekil 13'deki gibidir. Eğitim durumları Levenberg-Marquardt (LM) geri yayılım algoritması kullanılarak oluşturulmuştur (AL-Zubaidi vd. 2019). YSA sonucuna göre tahmin modelinin %55 oranında doğru olduğunu göstermektedir.



Şekil 13. Zaman serisi ve tahmin değerleri

Yapay Sinir Ağı (YSA) modelimizin sonuçlarına göre, tahmin modeli %55 oranında doğru tahminler yapabilmektedir. Bu sonuç, modelin eğitim ve test süreçleri sonrasında elde edilen doğruluk oranını göstermektedir. %55'lik bir doğruluk oranı, modelin zaman serisi verilerindeki eğilimleri ve kalıpları bir miktar yakalayabildiğini ancak daha fazla iyileştirmeye ihtiyaç duyduğunu göstermektedir.

4. Sonuçlar ve Tartışma

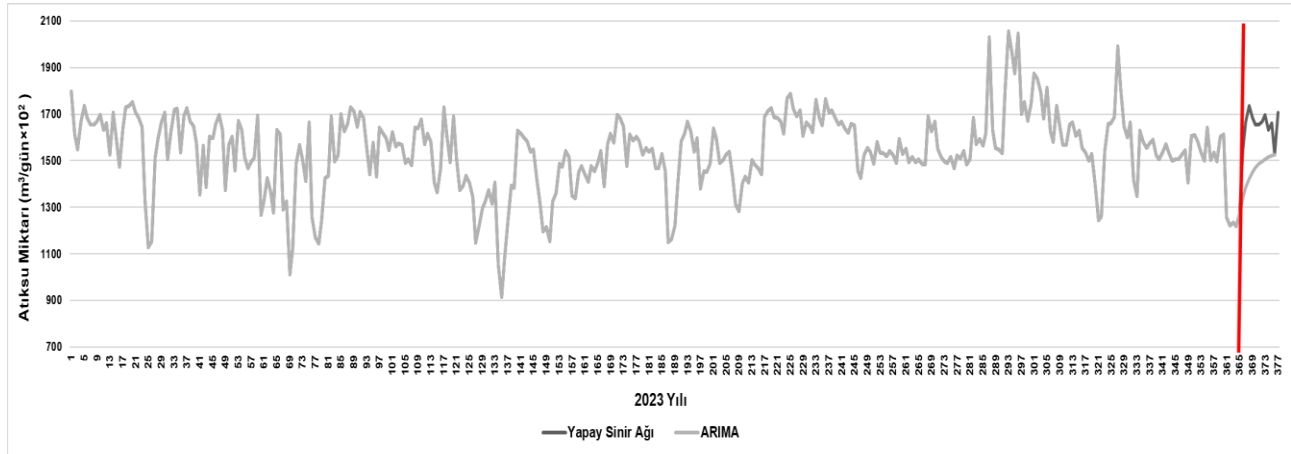
Atık su arıtma tesisleri için arıtılacak atık suyun miktarının tahmin edilmesi kaynak yönetimi ve işletme verimliliği açısından önemlidir. Tesisin kapasitesinin doğru hesaplanması ve gelecek dönemlerdeki yeterliliği için doğru planlanması gerekmektedir. Aksi takdirde tesis kapasitesi yetersiz kalıp arıtılması gereken atık sular arıtılmazsa çevresel kirliliğe sebep olabilir. Bununla birlikte nüfus artışı, sanayileşme gibi faktörler göz önüne alındığında tesis planlanması, kapasite hesabı ve yönetiminde kritik rol oynayan atık su miktarı doğru tahmin edilmelidir. Çalışmada ele alınan konunun atık su tahmini olması, bu elzem hususlar içindir. Bu çalışmada yapay sinir ağı ve ARIMA modelinin kıyaslanması, bir atık su arıtma tesisine ait zaman serisi tahmin çalışması için gerçekleştirilmiştir. Analizler sonucunda ARIMA ve Yapay

sinir ağı modeli RMSE, MAE ve MAPE değerleri açısından Çizelge 6'da olduğu gibi karşılaştırılmıştır. ARIMA modeli ile elde edilen sonuçlar yapay sinir ağlarına kıyasla daha yüksek performans göstermiştir. ARIMA modeli analizi sonucunda, 12 gün için tahmin değerleri sunmuştur.

Yapay sinir ağı ve ARIMA modellerini karşılaştırmak adına yapay sinir ağı ile elde edilen tahmin sonuçlarından ilk 12 gün baz alınmıştır. Şekil 14'te 2023 yılı verisine ilave olarak 365. ve 377. Günler arasında ARIMA ve yapay sinir ağı modeli tahmin değerleri birlikte verilmiştir.

Çizelge 6. ARIMA ve Yapay Sinir Ağları Sonuçları Karşılaştırılması

Model	RMSE	MAE	MAPE
ARIMA(2,1,2)	11244,60	7996,35	5,37
Yapay Sinir Ağı	21856,48	16844,66	11,81



Şekil 14. 365 günlük zaman serisi ve devamında yer alan 12 günlük yapay sinir ağı ve arima modeli tahmin değerleri kıyaslanması

Sonuç olarak, atık su arıtma tesisleri için kritik öneme sahip olan atık su miktarının tahmin edilmesi, yapay sinir ağı (YSA) ve ARIMA modelleri kullanılarak analiz edilmiştir. Çalışmanın amacı, her iki modelin performansını karşılaştırarak, atık su tahmini için en uygun yöntemi belirlemektir. YSA ve ARIMA modelleri, farklı özelliklere ve metodolojik yaklaşımlara sahip olmaları nedeniyle seçilmiştir.

YSA, zaman serisi verilerindeki karmaşık ilişkileri modellemek için kullanılırken, ARIMA modeli, geçmiş verilerin lineer kombinasyonlarını kullanarak gelecekteki değerleri tahmin etme yeteneğine sahip olduğu için kullanılmıştır. ARIMA modelinin atık su miktarının tahmini için daha güvenilir bir seçenek olabileceğini göstermektedir. Çalışmamız, atık su arıtma tesislerinin kapasite planlaması ve yönetimi için daha doğru tahmin modelleri geliştirilmesine yönelik önemli katkılar sağlamaktadır. Gelecekteki çalışmalar, model performansını artırmak amacıyla daha fazla veri kullanarak ve farklı model kombinasyonlarını deneyerek daha kapsamlı analizler yapılabilir. Böylece, çevresel etkilerin azaltılması ve kaynakların verimli kullanılması yönünde önemli adımlar atılabilir.

Etik Standartlar Bildirgesi

Yazarlar tüm etik standartlara uyduklarını beyan ederler.

Yazarlık Katkı Beyanı

Yazar-1: Kaynaklar, Araştırma, Deney, Yazma

Yazar-2: Araştırma, Fikir Sahibi, Biçimsel analiz, Doğrulama

Yazar-3: Deney tasarımı, Görselleştirme, Yazma –orijinal taslak

Çıkar Çatışması Beyanı

Yazarların bu makalenin içeriğiyle ilgili olarak beyan edecekleri hiçbir çıkar çatışması yoktur.

Verilerin Kullanılabilirliği

Bu çalışma sırasında oluşturulan veya analiz edilen tüm veriler, yayınlanan bu makaleye dahil edilmiştir.

5. Kaynaklar

Al-Dahidi, S., Alrbai, M., Al-Ghussain, L., Alahmer, A. and Hayajneh, H.S., 2024. Data-driven analysis and prediction of wastewater treatment plant performance: Insights and forecasting for sustainable operations. *Bioresource Technology*, **391**, 129937. <https://doi.org/10.1016/j.biortech.2023.129937>

AL-Zubaidi, E.D.A., Yas, A.H. and Abbas, H.F., 2019. Guess the time of implementation of residential construction projects using neural networks ANN. *Periodicals of Engineering and Natural Sciences*, **7(3)**: 1218-1227, 2019. <https://doi.org/10.21533/pen.v7i3.680>

Ataseven, B., 2013. Forecasting by using artificial neural networks. *Öneri Dergisi*, **10(39)**, 101-115.

- <https://doi.org/10.14783/od.v10i39.1012000311>
- Baki, O.T. and Aras, E., 2018. Estimation of Bod in wastewater treatment plant with different regression models. *Engineering Sciences*, **13(2)**, 96-105.
- Çelik, N., Coşar, D.N. and Konyalıoğlu, A.K., 2022. Wastewater Forecasting Application by an Integrated Interpolation and Box-Jenkins Modelling Approach in Turkey. *The International Symposium for Production Research*, Springer.
https://doi.org/10.1007/978-3-031-24457-5_10
- Do, P.C., Chow, W.K., Rameezdeen, R. and Gorjian, N., 2022. Wastewater inflow time series forecasting at low temporal resolution using SARIMA model: a case study in South Australia. *Environmental Science and Pollution Research*, **29(47)**, 70984-70999
<https://doi.org/10.1007/s11356-022-20777-y>
- Elevli, S., 2020. The utilisation of special cause control charts in the presence of autocorrelated data. *Sigma Journal of Engineering and Natural Sciences*, **38(2)**, 787-793
- Er Aydın, B., Odabas, M.S., Senyer, N. and Ardali, Y. 2022. Evaluation of Deep Sea discharge systems efficiency in the eastern black sea using artificial neural network: a case study for Trabzon, Turkey. *Brazilian Archives of Biology and Technology*, 65: e22210397
<https://doi.org/10.1590/1678-4324-2022210397>
- Erden, C., 2023. Performance Comparisons of Deep Learning and ARIMA: A Borsa Istanbul Stock Example. *Yönetim ve Ekonomi Dergisi*, **30(3)**, 419-438.
<https://doi.org/10.18657/yonveek.1208807>
- Işık, H.E., Bas, E., Egrioglu, T. and Akkan, A., 2024. A new single multiplicative neuron model artificial neural network based on black hole optimization algorithm: forecasting the amounts of clean water given to metropolis. *Stoch Environ Res Risk Assess.*
<https://doi.org/10.1007/s00477-024-02802-3>
- Kaya, N.S., Pacci, S., Demirağ Turan, İ., Odabas, M.S. and Dengiz, O., 2023. Comparing geographic information systems-based fuzzy-analytic hierarchical process approach and artificial neural network to characterize soil erosion risk indexes. *Rendiconti Lincei. Scienze Fisiche e Naturali*, **34(4)**, 1089-1104.
<https://doi.org/10.1007/s12210-023-01201-0>
- Lai, M.S., Wulff, S., Cao, Y., Robinson, T.J. and Rajapaksha, R., 2023. An interpretable time series machine learning method for varying forecast and nowcast lengths in wastewater-based epidemiology. *MethodsX*, **11**, 102382
<https://doi.org/10.1016/j.mex.2023.102382>
- Namasudra, S., Dhamodharavadhani, S. and Rathipriya, R., 2021. Nonlinear neural network based forecasting model for predicting COVID-19 cases. *Neural Processing Letters*, **55**, 171-191.
<https://doi.org/10.1007/s11063-021-10495-w>
- Odabas, M.S., Senyer, N., Kayhan, G. and Ergun, E., 2017. Estimation of chlorophyll concentration index at leaves using artificial neural networks. *Journal of Circuits, Systems and Computers*, **26(2)**, 1750026.
<https://doi.org/10.1142/S0218126617500268>
- Öztemel, E. and Dügenci, M., 2016. Atıksu arıtma tesis kontrolde yapay sinir ağı ile kirlilik parametre tahmini. *3rd International Symposium on Environment and Morality*, Alanya, Türkiye.
- Palabıyık, S. and Akkan, T., 2024. Evaluation of water quality based on artificial intelligence: performance of multilayer perceptron neural networks and multiple linear regression versus water quality indexes. *Environment, Development and Sustainability*
<https://doi.org/10.1007/s10668-024-05075-6>
- Rahman, A. and Hasan, M.M. 2017. Modeling and forecasting of carbon dioxide emissions in Bangladesh using Autoregressive Integrated Moving Average (ARIMA) models. *Open Journal of Statistics*, **7(4)**, 560-566
<https://doi.org/10.4236/ojs.2017.74038>
- Suhermi, N., Suhartono, D.D. and Prastyo, B.A., 2018. Roll motion prediction using a hybrid deep learning and ARIMA model. *Procedia computer science*, **144**, 251-258
<https://doi.org/10.1016/j.procs.2018.10.526>
- Zhang, G.P., 2003. Time series forecasting using a hybrid ARIMA and neural network model. *Neurocomputing*, **50**, 159-175.
[https://doi.org/10.1016/S0925-2312\(01\)00702-0](https://doi.org/10.1016/S0925-2312(01)00702-0)
- Zhang, Q., Li, Z., Snowling, S., Siam, A. and El-Dakhakhni, W. 2019. Predictive models for wastewater flow forecasting based on time series analysis and artificial neural network. *Water Science and Technology*, **80(2)**, 243-253
<https://doi.org/10.2166/wst.2019.263>