



**T.C.**

**HİTİT ÜNİVERSİTESİ**

**LİSANSÜSTÜ EĞİTİM ENSTİTÜSÜ**

**KİMYA MÜHENDİSLİĞİ ANABİLİM DALI**

**YEŞİLIRMAK NEHRİ SU KALİTE VERİLERİNDEN AYKIRI  
DEĞERLERİN TEMİZLENMESİ İÇİN KULLANICI DOSTU BİR  
ARAYÜZ GELİŞTİRİLMESİ**

**Yüksek Lisans Tezi**

**Zübeyde ZENGÜL**

**Çorum- 2023**



**YEŞİLIRMAK NEHRİ SU KALİTE VERİLERİNDEN AYKIRI DEĞERLERİN  
TEMİZLENMESİ İÇİN KULLANICI DOSTU BİR ARAYÜZ  
GELİŞTİRİLMESİ**

**Zübeyde ZENGÜL**

**Lisansüstü Eğitim Enstitüsü  
Kimya Mühendisliği Anabilim Dalı**

**Yüksek Lisans Tezi**

**TEZ DANIŞMANI**

**Dr. Öğr. Üyesi Erdal KARADURMUŞ**

**Çorum 2023**

Zübeyde ZENGÜL tarafından hazırlanan “Yeşilırmak Nehri Su Kalite Verilerinden Aykırı Değerlerin Temizlenmesi İçin Kullanıcı Dostu Bir Arayüz Geliştirilmesi” adlı tez çalışması 13/06/2023 tarihinde aşağıdaki jüri üyeleri tarafından oy birliği ile Hitit Üniversitesi Lisansüstü Eğitim Enstitüsü Kimya Mühendisliği Dalında Yüksek Lisans tezi olarak kabul edilmiştir.

Doç. Dr. Serap FINDIK

.....

Dr. Öğr. Üyesi Erdal KARADURMUŞ

.....

Dr. Öğr. Üyesi Gediz UĞUZ

.....

Hitit Üniversitesi Lisansüstü Eğitim Enstitüsü Yönetim Kurulunun .../.../..... tarih ve .....sayılı kararı ile Zübeyde ZENGÜL'ün Kimya Mühendisliği Anabilim Dalında Yüksek Lisans derecesi alması onanmıştır.

Prof. Dr. Muhammed Asif Yoldaş

Lisansüstü Eğitim Enstitüsü Müdür V.

## TEZ BİLDİRİMİ

Tez içindeki bütün bilgilerin etik davranış ve akademik kurallar çerçevesinde elde edilerek sunulduğunu, ayrıca tez yazım kurallarına uygun olarak hazırlanan bu çalışmada bana ait olmayan her türlü ifade ve bilginin kaynağına eksiksiz atıf yapıldığını beyan ederim.

Zübeyde ZENGÜL

# YEŞİLIRMAK NEHRİ SU KALİTE VERİLERİNDEN AYKIRI DEĞERLERİN TEMİZLENMESİ İÇİN KULLANICI DOSTU BİR ARAYÜZ GELİŞTİRİLMESİ

Zübeyde ZENGÜL

ORCID: 0000-0001-5851-1418

HİTİT ÜNİVERSİTESİ

LİSANSÜSTÜ EĞİTİM ENSTİTÜSÜ

Yüksek Lisans Tezi

Haziran 2023

## ÖZET

Su canlıların yaşamında önemli bir yere sahiptir. Son yıllarda hızlı nüfus artışı, endüstriyel faaliyetlerdeki artış ve bilinçsiz tarım yöntemleri nedeniyle su kaynakları ciddi tehdit altındadır. Diğer taraftan su kaynaklarının korunması ve kirlilik takibi için ciddi yaptırımlar oluşturulmuştur. Su kirliliği takibinin yapılabilmesi için son yıllarda 'on-line' istasyonların ve ölçüm yöntemlerinin kullanılması söz konusudur. Bu durum takibin daha sık aralıklarla yapılabilmesini sağladığı gibi çok sayıda verinin elde edilmesini de kolaylaştırmıştır. Ayrıca veri madenciliği ve bilgisayar teknolojilerindeki gelişmeler pek çok parametre için zahmetli ve maliyetli deneysel yöntemlerden ziyade tahmin yeteneğine sahip modellerin gelişmesini mümkün kılmıştır. Ancak gerçek veri ile geliştirilen pek çok modelde veri kalitesi önemli bir parametre olduğu için verilerin belirli ön işlemlerden geçirilmesi gereklidir.

Bu tez kapsamında, Yeşilırmak Nehri üzerinde kurulu 'on-line' izleme istasyonundan 2007-2009 yılları arasında toplanmış olan veriler kullanılarak veri seti içerisindeki aykırı değerlerin temizlenmesi çalışılmıştır. TÜBİTAK projesi kapsamında kurulan bu istasyonlardan alınan verilerden; çözülmüş oksijen (*luminescence dissolved oxygen*, LDO), sıcaklık, pH, iletkenlik, toplam organik karbon (*total organic carbon*, TOC), nitrat azotu (NO<sub>3</sub>-N) ile amonyum azotu (NH<sub>4</sub>-N) için değerlendirmeler yapılmıştır. Bu kapsamda, ortanca, ortalama, *Grubbs*, *Gesd*, çeyrekler açıklığı olmak üzere beş farklı aykırı veri temizleme yöntemi kullanılmış ve MATLAB ortamında bir kullanıcı etkileşimli grafiksel ara yüz geliştirilmiştir. Böylece kullanıcı herhangi bir veri setini yazılıma tanıtarak aykırı verileri birkaç adımda farklı yöntemlerle temizleyerek

modelleme alıřmaları iin hazır duruma getirebilmektedir. Aykırı veri temizleme yntemleri arasında ortanca algoritmasında daha fazla sayıda veri ayıklandığı grlmřtr.

**Anahtar Kavramlar:** Yeřilirmak Havzası, Nehir su kirliliđi, *on-line* izleme, modelleme, grafiksel ara yz tasarımı

**Bilim Kodu:** 91204, 91216



# **A USER-FRIENDLY INTERFACE FOR REMOVAL OF OUTLIERS FROM THE YESILIRMAK RIVER WATER QUALITY DATA**

Zübeyde ZENGÜL

ORCID: 0000-0001-5851-1418

HITIT UNIVERSITY

GRADUATE SCHOOL

Master of Science Thesis

June 2023

## **ABSTRACT**

Water plays a crucial role in the life of living organisms. Water resources have been seriously threatened in recent years due to rapid population growth, increased industrial activities, and unsustainable agricultural practices. On the other hand, significant measures have been taken to protect water resources and monitor pollution. In recent years, online stations and measurement methods have been introduced for water pollution monitoring. This allows for more frequent monitoring and facilitates the collection of a large amount of data. Furthermore, advancements in data mining and computer technologies have enabled the development of predictive models for various parameters, replacing laborious and costly experimental methods. However, data quality is important in models developed using real data, necessitating proper data preprocessing.

This thesis focuses on removing outlier values in the dataset using the data collected from an online monitoring station on the Yeşilırmak River between 2007 and 2009. Evaluations were performed for dissolved oxygen (luminescence dissolved oxygen, LDO), temperature, pH, conductivity, total organic carbon (TOC), nitrate nitrogen ( $\text{NO}_3\text{-N}$ ), and ammonium nitrogen ( $\text{NH}_4\text{-N}$ ) based on the data obtained from these stations established under a TÜBİTAK project. Five outlier data removal methods were utilised: median, mean, Grubbs, Gesd, and interquartile range. Additionally, a graphical user interface was developed in the MATLAB environment. As a result, users can introduce any dataset to the software and remove outlier values using different methods in a few steps, preparing the data for modelling studies. It was



observed that the median algorithm removed more data points among the outlier data-removing methods.

**Key Terms:** Yeşilırmak Basin, River water pollution, online monitoring, modelling, graphical interface design

**Science Code:** 91204, 91216



## TEŐEKKÖR

Bu alıőmalarda emeęi geen, baőtta lisans eęitiminden beri danıőmanım olup tecrübeleriyle yol gōsteren ok kıymetli ve saygıdeęer hocam Dr. Őęr. Üyesi Erdal KARADURMUŐ' a, ok üzücü olan depremin yıkıcı etkisini yaőamasına raęmen bilgisayar programlarının ve ara yüzün geliştirilmesinde yardımını esirgemeyen İnönü Üniversitesinden ok deęerli hocam Prof. Dr. Mehmet YÜCEER 'e, bilgi ve yardımlarıyla Ankara Üniversitesi'nden ok kıymetli hocam Arő. Gör. Eda GÖZ 'e en içten teőekkürlerimi ve saygılarımı sunarım. Eęitimim sürecimin manevi güçleri sevgili eőim Erdem ve ocuklarım Mehmet Erkam, Emirhan ve Kemalettin ZENGÖL' e teőekkür ediyorum.

Zübeyde ZENGÖL

## İÇİNDEKİLER

	Sayfa
ÖZET.....	iv
ABSTRACT.....	vi
TEŞEKKÜR.....	viii
İÇİNDEKİLER.....	ix
TABLolar DİZİNİ.....	xi
ŞEKİLLER DİZİNİ.....	xii
RESİMLER DİZİNİ.....	xiv
SİMGELER VE KISALTMALAR.....	xv

### GİRİŞ

#### 1. BÖLÜM

##### GENEL BİLGİLER

1.1. Su Dünyamız.....	3
-----------------------	---

#### 2. BÖLÜM

##### YEŞİLIRMAK HAVZASI VE ÖLÇÜM İSTASYONLARI

2.1. Yeşilirmak Havzası Genel Bilgiler.....	7
2.2. Yeşilirmak Havzası Çalışma Alanı.....	8
2.3. Yeşilirmak Havzasında Kirlilik Parametreleri.....	9
2.3.1. Çözünmüş oksijen.....	9
2.3.2. Sıcaklık.....	10
2.3.3. pH.....	10
2.3.4. İletkenlik.....	11
2.3.5. Toplam organik karbon.....	12
2.3.6. Nitrat azotu.....	12
2.3.7. Amonyum azotu.....	13
2.4. Aykırı Değer Tanımı.....	14
2.5. Aykırı Değerlerin Tespiti.....	14

2.5.1. Aykırı deęer temizlięinin önemi.....	15
2.5.2. Aykırı deęer temizleme testleri .....	20

### **3. BÖLÜM**

#### **LİTERATÜR ARAŞTIRMASI**

3.1. Nehir Kirlilięi Modelleme ve Aykırı Veri Ayıklama Çalışmaları.....	24
---	----

### **4. BÖLÜM**

#### **YÖNTEM**

4.1. Verilerin Toplanması .....	26
4.2. Grafikselsel Ara Yüz Tasarımı.....	26
4.2.1. Arayüzün oluşturulmasının görsel anlatımı ve ayrıntıları.....	27

### **5. BÖLÜM**

#### **SONUÇLAR VE TARTIŞMA**

5.1. Çözünmüş Oksijen Parametresi için Aykırı Verilerin Ayıklanması .....	32
5.2. Sıcaklık Parametresi için Aykırı Verilerin Ayıklanması .....	35
5.3. pH Parametresi için Aykırı Verilerin Ayıklanması .....	38
5.4. İletkenlik Parametresi için Aykırı Verilerin Ayıklanması.....	41
5.5. Toplam Organik Karbon Parametresi için Aykırı Verilerin Ayıklanması.....	44
5.6. Nitrat Azotu ve Amonyum Azotu Parametreleri için Aykırı Verilerin Ayıklanması.....	47
<b>SONUÇ VE ÖNERİLER.....</b>	<b>55</b>
<b>KAYNAKÇA.....</b>	<b>56</b>
<b>EKLER .....</b>	<b>62</b>
<b>EK-1. 'deneme. m' dosyasına ait MATLAB kodları.....</b>	<b>63</b>
<b>EK-2. Aykırı Veri Temizleme Yazılımı Kullanım Kılavuzu ve Yardım Dosyası .....</b>	<b>68</b>
<b>EK-3. Ölçüm İstasyonları Yerlerinin Şematik Gösterimi.....</b>	<b>73</b>
<b>EK-4. İstasyon Analiz Sonuçlarının Çevre ve Orman Bakanlığı Analiz Sonuçları ile Kıyaslanması.....</b>	<b>74</b>

## TABLULAR DİZİNİ

<b>Tablo</b>	<b>Sayfa</b>
<b>Tablo 1.1.</b> Su Kalite Parametreleri ve Sınıfları .....	4
<b>Tablo 1.2.</b> Karadeniz yerüstü suları kalite kriterleri .....	5
<b>Tablo 2.1.</b> İkinci istasyon (Durucasu İstasyonu) ham verilerinin istatistiksel değerleri .....	20
<b>Tablo 5.1.</b> LDO için orijinal ve Temizlenmiş Veri Sayıları .....	53



## ŞEKİLLER DİZİNİ

Şekil	Sayfa
Şekil 2.1. Türkiye akarsu havzaları.....	7
Şekil 2.2. Yeşilirmak Havzası ve alt havzaları haritası.....	8
Şekil 2.3. Yeşilirmak Havzası Çalışma Alanı Haritası (Yüceer, 2005).....	9
Şekil 2.4. Aykırı Veri Tespit Yöntemleri.....	15
Şekil 2.5. Yeşilirmak havzası çalışma alanı NO <sub>3</sub> -N parametresinde kayıp veri oluşumu .....	19
Şekil 2.6. Ortanca Kutu Grafiği.....	21
Şekil 4.1. GUI oluşturma penceresi .....	28
Şekil 4.2. GUI içerisinde yer alacak objeler.....	29
Şekil 4.3. Yeşil ok (Çalıştırma) butonunun bulunduğu konum .....	29
Şekil 4.4. Yeşil ok tuşu ( <i>Run Figure</i> ) .....	29
Şekil 4.5. Dosya ismi verme .....	30
Şekil 4.6. Oluşturulan arayüzün görüntüsü.....	30
Şekil 4.7. Arayüze ait genel gösterim .....	31
Şekil 4.8. Arayüze ait detaylı gösterim .....	31
Şekil 5.1. Mean yöntemi ile LDO verilerinden aykırı verilerin ayıklaması sonuçları.....	32
Şekil 5.2. Median yöntemi ile LDO verilerinden aykırı verilerin ayıklaması sonuçları.....	33
Şekil 5.3. Gesd yöntemi ile LDO verilerinden aykırı verilerin ayıklaması sonuçları .....	34
Şekil 5.4. Grubbs yöntemi ile LDO verilerinden aykırı verilerin ayıklaması sonuçları .....	34
Şekil 5.5. Quartiles yöntemi ile LDO verilerinden aykırı verilerin ayıklaması sonuçları .....	35
Şekil 5.6. Mean yöntemi ile Sıcaklık verilerinden aykırı verilerin ayıklaması sonuçları.....	36
Şekil 5.7. Median yöntemi ile Sıcaklık verilerinden aykırı verilerin ayıklaması sonuçları.....	36
Şekil 5.8. Gesd yöntemi ile Sıcaklık verilerinden aykırı verilerin ayıklaması sonuçları .....	37
Şekil 5.9. Grubbs yöntemi ile Sıcaklık verilerinden aykırı verilerin ayıklaması sonuçları .....	37
Şekil 5.10. Quartiles yöntemi ile Sıcaklık verilerinden aykırı verilerin ayıklaması sonuçları .....	38

<b>Şekil 5.11.</b> Mean yöntemi ile pH verilerinden aykırı verilerin ayıklaması sonuçları.....	39
<b>Şekil 5.12.</b> Median yöntemi ile pH verilerinden aykırı verilerin ayıklaması sonuçları.....	39
<b>Şekil 5.13.</b> Gesd yöntemi ile pH verilerinden aykırı verilerin ayıklaması sonuçları .....	40
<b>Şekil 5.14.</b> Grubbs yöntemi ile pH verilerinden aykırı verilerin ayıklaması sonuçları .....	40
<b>Şekil 5.15.</b> Quartiles yöntemi ile pH verilerinden aykırı verilerin ayıklaması sonuçları .....	41
<b>Şekil 5.16.</b> Mean yöntemi ile iletkenlik verilerinden aykırı verilerin ayıklaması sonuçları ....	42
<b>Şekil 5.17.</b> Median yöntemi ile iletkenlik verilerinden aykırı verilerin ayıklaması sonuçları	42
<b>Şekil 5.18.</b> Gesd yöntemi ile iletkenlik verilerinden aykırı verilerin ayıklaması sonuçları.....	43
<b>Şekil 5.19.</b> Grubbs yöntemi ile iletkenlik verilerinden aykırı verilerin ayıklaması sonuçları.	43
<b>Şekil 5.20.</b> Quartiles yöntemi ile iletkenlik verilerinden aykırı verilerin ayıklaması sonuçları .....	44
<b>Şekil 5.21.</b> Mean yöntemi ile TOC verilerinden aykırı verilerin ayıklaması sonuçları .....	45
<b>Şekil 5.22.</b> Median yöntemi ile TOC verilerinden aykırı verilerin ayıklaması sonuçları .....	45
<b>Şekil 5.23.</b> Gesd yöntemi ile TOC verilerinden aykırı verilerin ayıklaması sonuçları .....	46
<b>Şekil 5.24.</b> Grubbs yöntemi ile TOC verilerinden aykırı verilerin ayıklaması sonuçları .....	46
<b>Şekil 5.25.</b> Quartiles yöntemi ile TOC verilerinden aykırı verilerin ayıklaması sonuçları.....	47
<b>Şekil 5.26.</b> Mean yöntemi ile NO <sub>3</sub> -N verilerinden aykırı verilerin ayıklaması sonuçları.....	48
<b>Şekil 5.27.</b> Median yöntemi ile NO <sub>3</sub> -N verilerinden aykırı verilerin ayıklaması sonuçları.....	48
<b>Şekil 5.28.</b> Gesd yöntemi ile NO <sub>3</sub> -N verilerinden aykırı verilerin ayıklaması sonuçları .....	49
<b>Şekil 5.29.</b> Grubbs yöntemi ile NO <sub>3</sub> -N verilerinden aykırı verilerin ayıklaması sonuçları .....	49
<b>Şekil 5.30.</b> Quartiles yöntemi ile NO <sub>3</sub> -N verilerinden aykırı verilerin ayıklaması sonuçları ...	50
<b>Şekil 5.31.</b> Mean yöntemi ile NH <sub>4</sub> -N verilerinden aykırı verilerin ayıklaması sonuçları.....	50
<b>Şekil 5.32.</b> Median yöntemi ile NH <sub>4</sub> -N verilerinden aykırı verilerin ayıklaması sonuçları.....	51
<b>Şekil 5.33.</b> Gesd yöntemi ile NH <sub>4</sub> -N verilerinden aykırı verilerin ayıklaması sonuçları.....	51
<b>Şekil 5.34.</b> Grubbs yöntemi ile NH <sub>4</sub> -N verilerinden aykırı verilerin ayıklaması sonuçları .....	52
<b>Şekil 5.35.</b> Quartiles yöntemi ile NH <sub>4</sub> -N verilerinden aykırı verilerin ayıklaması sonuçları ...	52
<b>Şekil 5.36.</b> Aykırı veri ayıklama yöntemlerine göre orijinal ve temizlenmiş veri sayıları.....	53

## RESİMLER DİZİNİ

Resim	Sayfa
<b>Resim 2.1.</b> Nehir suyunda ani yükselme sonucu şamandıranın bükülmesi.....	17
<b>Resim 2.2.</b> Çalışma alanında kirliliğin temizlenmesi.....	18
<b>Resim 2.3.</b> Kirliliğe bağlı çamur birikmesi.....	18





## SİMGELER VE KISALTMALAR

### Simgeler

H	Hidrojen
N	Azot
NH <sub>3</sub> -N	Nitrat Azotu
NH <sub>4</sub>	Amonyum
NH <sub>4</sub> -N	Amonyum Azotu
NO <sub>3</sub>	Nitrat

### Kısaltmalar

CWA	Temiz Su Yasası
EPA	Çevre Koruma Ajansı
GUI	Kullanıcı Etkileşimli Arayüz
GESD	<i>Generalized Extreme Studentized Detection</i>
ESD	<i>Extreme Studentized Detection</i>
EDS	<i>Event Detection Systems</i>
NPDES	Ulusal Kirletici Deşarj Eliminasyon Sistemi
MEB	Millî Eğitim Bakanlığı
SKKY	Su Kirliliği Kontrolü Yönetmeliği
TKN	Toplam Kjeldahl Azotu
TOC	Toplam Organik Karbon
TÜBİTAK	Türkiye Bilimsel ve Teknik Araştırma Kurumu
WFD	Avrupa Su Çerçeve Direktifi
WHO	Dünya Sağlık Örgütü
WWF	Doğal Hayatı Koruma Vakfı

## GİRİŞ

Su, canlıların yaşamını sürdürebilmesi için ana unsurlardan biridir. Dünya nüfusunun artması ve insanoğlunun konfor arayışı sonucunda temiz su tüketimi artarak evsel, tarımsal ve endüstriyel atıklar ile birlikte sular kirlenmiştir. Dünyanın su varlığı ölçüsü toplamda 1,4 milyar km<sup>3</sup> olarak bilinmektedir. Bu ölçünün %97,5'i tuzlu su olarak denizlerde ve okyanuslarda bulunmaktadır. Geriye kalan %2,5'lik dilim ise göllerde, nehirlerde ve yeraltı kaynaklarında bulunmaktadır (TÜBİTAK Bilim Genç, 2016).

Son yıllarda, su kaynaklarının azalmasına bağlı olarak su kirliliğinin kontrolü için ciddi yaptırımlar söz konusu olmuştur. Yüzey suyu kalitesi, çeşitli kentsel, endüstriyel ve tarımsal, dağınık ve noktasal kirlilik kaynaklarından etkilenir (Casillas-García vd., 2021, s. 1). 1970'lere kadar Amerika halkının doğal kaynakların sınırsız ve tükenmeyeceği inancı sonucunda oluşan ağır çevre tahripleri birçok insanın tepkisini çekmiştir. Bu durumun sonucunda çevreci anlayış ilk defa ortaya çıkmıştır. Ekosistemin korunması için bilimsel çalışmalar, sulak alanların artırılıp korunması için çevre koruma politikası ve mevzuatın geliştirilmesine sebep olmuştur (Güneş, 2004, s. 89). Amerika Birleşik Devletleri 1948 yılında Federal Su Kirliliği Kontrol Yasasını çıkartmış ve bu yasaı revize ederek önce 1972 yılında, yüzey sularındaki kirleticiler için ulusal su kalitesi kriterleri önerilerini geliştirip Temiz Su Yasası (CWA)' nı çıkarmıştır. Bu yasa ek kanunlarla genişletilerek (1973, 1977, 2008, 2018 vb.) Avrupa Birliği Çevre Yasaları gibi küresel olarak örnek alınmaktadır. ABD Çevre Koruma Ajansı (*U.S. Environmental Protection Agency*, EPA), insan sağlığını ve çevreyi korumak için 1970 yılında oluşturulmuştur. EPA, endüstri için atık su standartlarını belirleyip kirlilik kontrol programları uygulamaktadır. 1972'de Temiz Su Yasasıyla oluşturulan Ulusal Kirletici Deşarj Eliminasyon Sistemi (NPDES) izin programı, eyalet hükümetlerine izin verme, idari ve uygulama yönünden yetki vermiştir (EPA, 2022). Bu durum su kaynaklarının takibi ve kontrolünü zorunlu kılmıştır. Bu amaçla özellikle yüzey sularında gerçek zamanlı kirlilik takibi yapılmaya başlanmıştır. Gerçek zamanlı kirlilik takibi projeleri ile bilgisayar teknolojisi ve veri transferindeki gelişmeler sonucu çok sayıda veriye kısa sürede ulaşma olanağı sağlamaktadır. Toplanan bu verilerin işlenerek eldeki mevcut parametreler kullanılarak diğer parametrelerin öngörülmesini yani modelleme çalışmalarını ortaya koymuştur. Son yıllarda modelleme çalışmaları oldukça popüler hale gelmiştir.

Literatürde farklı türde modelleme çalışmaları mevcuttur. Oğuz (1997), yaptığı doktora çalışmasında Erzurum Karasu Nehrinde, tek boyutlu olarak matematiksel modellemeyle, deşarj ve nehir parametrelerini segmentler arasındaki mesafeyle ilintili olarak simüle edip derişimin mesafeyle deęişimini göstermiştir. Özkan vd. (2006), Kızılırmak nehrindeki çözünmüş oksijen derişiminin modellenmesinde yapay sinir ağlarını kullanmıştır. Köklü (2010), doktora tezinde Melen Nehri su kalitesine ait birçok veriyi istatistiksel analiz yöntemleri ve yapay zekâ yöntemiyle modellemiştir. Ancak mekanistik yani cebirsel ve diferansiyel denklemleri içeren modellerde sistem tanımlanırken fazla bilgi gerektiği için çok

büyük veri girişi olmaktadır. Bu yoğun veri girişinin anlaşılması amacıyla yapay zekâ modellerine ihtiyaç söz konusu olmuştur. Model, su kalitesi reaksiyon mekanizmalarını basitleştirir, kavramsallaştırır ve sentezleyerek bunları sayısal yöntemler ve bilgisayar programları kullanarak çözülen matematiksel yönetici ve yardımcı denklemler olarak ifade eder (Lammers ve Bledsoe, 2019, s. 105; Sharma ve Kansal, 2013, s. 287; Başer vd. 2014, s. 1767).

Akarsu havzalarının su kalitesini belirleyen kirlilik parametreleri modellenirken tanımlanan çok fazla veri, çeşitli sebeplerle normalden sapabilmektedir. Su kalitesi izleme verilerinin normalden sapması bazı aykırı veriler olduğunun göstergesi olabilir (Than, vd., 2021, s. 4). Bu sapmalar, geliştirilmiş veya geliştirilme aşamasındaki modellerin performansını düşürmektedir. Bu sebeple bazı aykırı veri ayıklama yöntemlerinin uygulanması genellikle modelin öngörü başarısını artırmaktadır. Bunlar; verilerin hacmine, ölçüm dağılımına ve değişken sayısına göre değişkenlik gösteren yöntemlerdir. En çok uygulanan aykırı değer tespit yöntemleri istatistiksel yöntemler olup, Dixon, Gesd, Grubbs, en küçük kareler, standart sapma, mod, ortalama, Rosner, Walsh vb. testler aykırı değer tespit çalışmalarında uygulananlardan sadece bir kısmıdır.

Bu çalışmada, Yeşilırmak nehri üzerinde kurulu istasyonlardan ikincisi olan Duracasu istasyonundan, 2007-2009 yılları arasında beşer dakika aralıklarla elde edilen veriler kullanılmıştır (TÜBİTAK, KAMAG-105G002). Bu veriler aykırı değerler taşımaktadır. Aykırı veriler için beş farklı aykırı değer tespit yöntemi; ortalama (*mean*), ortanca (*median*), Grubbs, Gesd ve çeyrekler açıklığı (*interquartile*) yöntemleri kullanılarak çözülmüş oksijen (LDO), sıcaklık(T), pH, iletkenlik (EC), toplam organik karbon (TOC), nitrat azotu (NO<sub>3</sub>-N) ve amonyum azotu (NH<sub>4</sub>-N) olmak üzere yedi adet parametre içeren veri seti ayıklanmıştır. Veri ayıklama işlemi çalışma kapsamında geliştirilen kullanıcı dostu bir arayüz aracılığı ile sağlanmıştır. Sonuçlar sayısal ve görsel olarak arayüz üzerinden takip edilebilmektedir.

# 1. BÖLÜM

## GENEL BİLGİLER

### 1.1. Su Dünyamız

Su, genel olarak doğal kaynakların en temeli olarak kabul edilir (Ghotbizadeh vd., 2022, s. 1). Tatlı su kaynakları canlı yaşamının sürekliliği için vazgeçilmezdir. Tarihteki, Mısır, Mezopotamya, Sümer, İnkâ ve Aztekler gibi büyük medeniyetler tatlı su kaynaklarının yakınında kurulmuştur (Başer, 2006, s. 3). Tatlı su kaynaklarının hidrolojik anlamda azalmasındaki etkenler; iklim değişikliği ve küresel ısınma sonucunda kuraklık oluşmasıdır. Tarım ve Orman Bakanlığı, 22 Mart Dünya Su Günü'nde iklim değişikliği tahminlerini yayımlamıştır. Yağışların azalması ve diğer faktörlerin etkisi su sıkıntısına yol açacağından, önlem alınmaz ise 2050 yılından itibaren ülkemizin ciddi su sıkıntısı yaşayan ülkeler arasında olacağı öngörülmektedir (T.C. Tarım Orman Bakanlığı, 2023).

Dünyamızın çoğu yani %71'i -genel tabiri ile üçte ikisi- sularla kaplıdır. Bu suların gezegen üzerindeki dağılımı ise %2,5 tatlı su, %97,5 tuzlu sudur. Bütün tatlı su kaynaklarının ise %68,9 buz ve buzullar, %30'u ise yer altı suları, donmuş toprak ve toprağın içindeki nemdir. Geriye kalan sadece %0,3'ü ise yüzey suyu olan göller ve nehirlerdir (TÜBİTAK Bilim Genç, 2016). Kaliteli ve yeteri miktardaki sular; insanlık, gıda güvencesi, sürdürülebilir kalkınma ve tatlı su ekosisteminin devamı için temel unsurdur. Dünyanın kırk üç farklı ülkesinde, tahmini 700 milyon kişi su kıtlığı çekmektedir. Ayrıca, 2,6 milyar civarı kişi ise bütün bir yıl boyunca en az bir ay su kıtlığı olan havzalarda yaşamaktadır (Pegram, vd., 2013, s.6). Havzalardaki suyun kaliteli olması önemlidir. Çünkü su kalitesi insan sağlığını etkilemekte olup kirli su tüketimi sağlık sorunlarını arttırmaktadır (Kadam, vd., 2019, s.951).

Son yıllarda, kasıtlı veya kasıtsız su ile yapılan olumsuz senaryoların insanları endişeye sevk etmesi sebebiyle, kirlilik takibinin yapılması için takip sistemlerinin kurulması ve güvenlik önlemlerinin alınması zorunluluk olmuştur (Housh ve Ohar, 2017, s.180).

Doğal Hayatı Koruma Vakfı (*World Wildlife Fund, WWF*), 2013 yılında yayımladığı, Türkiye'nin Su Ayak İzi Raporunda; Kimya dalında Nobel ödülü bulunan bilim adamı Paul Josef Crutzen'in söylemlerinde, insanların son iki yüzyılda dünyamızı olumsuz anlamda değiştirdiğini belirtmişlerdir. Paul Josef Crutzen'in, bu değişimin etkilerinin, gezegenimizin görünüşünü, devamındaki milyonlarca yıl etkileyerek yeni bir jeolojik dönemin başlangıcını yaşıyor olabileceğimizi dile getirdiğini söylemektedir. Jeoloji uzmanları, sanayi devrimi ve onu izleyen ekonomik gelişmelerin hızlandığı küresel çevre sorunları yüzünden yaşadığımız dönemin, İnsan Çağı-Antroposen diye adlandırılması gerektiğini belirtmektedirler. Doğal kaynaklarımızın çok fazla tüketilmesi ile oluşan kirlilik, gezegenimizin ekosistemlerini kapasitesinin son seviyesine getirmiştir. Dünya üzerinde coğrafi olarak dağınık halde bulunan

su kaynaklarının yönetiminde de sıkıntılar vardır. Dünya nüfusunun yaklaşık 8 milyar olduğu düşünülürse bu nüfusun ortalama 3 milyarı temiz olmayan ortamda yaşamaktadır. Yaklaşık 1 milyar insan ise sağlıklı sulara erişememektedir. Yeryüzündeki nüfusun 1/6'sı ise ihtiyaçlarını karşılayacak temiz suya erişebilecekleri sıhhi bir kaynak bulamamaktadır (Pegram, vd., 2013, s.6).

Bu bağlamda bakıldığında su kaynaklarının korunması ve gelecek nesillere sağlıklı bir biçimde aktarılması elzem ve gereklidir. Yönetimler, kanun ve yönetmelikler ile çeşitli yaptırımlar uygulayarak önemli bir yerüstü su kaynağı olan akarsu havzalarını koruma altına almışlardır. Avrupa Birliği ülkeleri, çevreyi ve akarsu havzalarını korumanın önemini kabul ederek, üye ülkelerde nehir yetkilileri tarafından takip edilecek izleme ağları kurmuş ve bu çaba Avrupa Su Çerçeve Direktifi (WFD)'nin yürürlüğe girmesinden sonra yoğunlaşmış ve desteklenmiştir (Avrupa Birliği Konseyi, 2000).

**Tablo 1.1.** Su Kalite Parametreleri ve Sınıfları

Su Kalite parametreleri	Su Kalite Sınıfları		
	I (çok iyi) Mavi	II (iyi) Yeşil	III (orta) Sarı
Renk(m <sup>-1</sup> )	RES 436 ≤ nm:1,5 RES 525 ≤ nm:1,2 RES 620 ≤ nm:0,8	RES 436 nm:3 RES 525 nm:2,4 RES 620 nm:1,7	RES 436> nm:4,3 RES 525> n>m:3,7 RES 620 nm:2,5
pH	6-9	6-9	6-9
İletkenlik (µS/cm)	<400	1000	> 1000
Yağ-gres (mg/L)	<0,2	0,3	>0,3
Çözünmüş Oksijen(mg/L)	>8	6	<6
KOİ (mg /L)	<25	50	>50
BOİ <sub>5</sub> (mg /L)	<4	8	>8
Amonyum azotu (NH <sub>4</sub> -N mg /L)	<0,2	1	>1
Nitrat azotu (mg NO <sub>3</sub> -N/L)	<3	10	>10
Toplam kjeldahl azotu (mg N/L)	<0,5	1,5	>1,5
Toplam Azot (mg N/L)	<3,5	11,5	>11,5
Orto fosfat fosforu (mg o-PO <sub>4</sub> -P/L)	<0,05	0,16	>0,16
Toplam Fosfor (mg P/L)	<0,08	0,2	>0,2
Florür (µg/L)	≤ 1000	1500	>1500
Mangan (µg/L)	≤ 100	500	>500
Selenyum (µg/L)	≤ 10	15	>15
Sülfür (µg/L)	≤ 2	5	>5

Türkiye’de su kaynakları yasal olarak ticari meta olmayıp tarihi bir miras olarak kabul edilmiştir. Devlet kontrolünde olan su kaynakları ve yerüstü suları yasalar ile teminat altında olup amme hizmeti ilkesine haizdir (Meriç, 2004, s.33). Ülkemiz suları, su kalitesi yönetimine

ilişkin düzenlemeler sonucunda, sınıflarına göre 2872 Sayılı Çevre Kanunu'na bağlı, SKKY (Su Kirliliği Kontrolü Yönetmeliği, 1988) ile ilk defa kalitelerine göre ayrılmıştır. Son şeklini 2021 yılında, Kıta içi yerüstü su kaynaklarının genel kimyasal ve fizikokimyasal parametreler açısından sınıflarına göre kalite kriterleri olarak alıp Resmî Gazetede yayımlanmıştır (Resmî Gazete, Tarih:16.06.2021 Sayı: 31513). Tablo 1.1'de Türkiye'nin Kıta içi yerüstü su kaynaklarının genel kimyasal ve fizikokimyasal parametreler açısından sınıflarına göre kalite kriterleri belirtilmiştir. En son olarak güncellenip yayınlanan bu parametreler sonucu havza çalışmalarının kirlilik kontrolleri yapılmaktadır.

Karadeniz Bölgesi havza çalışmaları için temel alınacak olan bilgiler ise Resmî Gazetede yayımlanan, Karadeniz için genel kimyasal fizikokimyasal parametreler açısından sınıflarına göre kıyı suları alıcı ortam kalite kriterleri belgesidir (Resmî Gazete, Tarih: 16.06.2021 sayı: 31513). Karadeniz'de bu parametreler referans alınarak kirlilik kontrolleri yapılmaktadır. Tablo 1.2'de Karadeniz yerüstü sularının kalite sınıflarına göre kirlilik parametreleri verilmiştir.

**Tablo 1.2.** Karadeniz yerüstü suları kalite kriterleri

<b>Karadeniz Su Kalite Sınıfları</b>			
Parametre	I (Çok iyi) Yüksek Kaliteli	II (İyi) Az Kirlenmiş	III (Orta) Kirli
Çözünmüş Oksijen (mg O <sub>2</sub> /L)	≥6	5	<5
TP(µg/L)	<8	8-12	>12
NO <sub>2</sub> (Oksitlenmiş azotlu bileşikler) (NO <sub>2</sub> -N+NO <sub>2</sub> -N) (µg/L)	<14	14-20	>20
Yağ-gres(mg/L)	<0,2	0,3	>0,3
Yüzer Madde	Yüzer halde sıvı maddeler, çöp ve benzeri katı maddeler ile köpük bulunamaz.		

Resmî Gazetede Tarım ve Orman Bakanlığının Yönetmeliğine göre;

- Birinci sınıf yerüstü suları, yüksek kaliteli içme suyu olarak, yüzme gibi vücut teması olan aktivitelerde, alabalık üretim tesislerinde ve hayvan çiftliklerinde kullanılabilir.

- İkinci sınıf sular ise az kirli olup iyi durumda olmasından dolayı içme suyu, aktivite, alabalık harici balık üretimi, güncel mevzuata uygun sulama suyu olarak kullanılır.
- Üçüncü sınıf sular ise orta kalite olan kirlenmiş sulardır. Yiyecek, dokuma gibi kaliteli su gerektiren kuruluşlar dışında uygulanacak bir arıtım sonucu sanayi suyu ve akuakültür olarak kullanılabilir.

Su kirliliğın önlenmesi ve kirleticilerin tespiti için Avrupa Birliğı'nin öncülüğünde başlanan gerçek zamanlı izleme projeleri ve ağıları ülkemizde de uygulanmaktadır. Çevre ve Şehircilik Bakanlığının 55 tane İl Müdürlüğü'ne bağılı toplam 389 adet yerüstü suları izleme noktası bulunmakta olup, bu istasyonlardan elde edilen muhtemel kirlenmişlik nedenleri aşağıda verilmiştir:

- 297 yerüstü suyunun %76,3'ü kanalizasyon sularıyla,
- 294 yerüstü suyunun %75,6'sı tarım ilacı(pestisit) ve gübrelemeyle,
- 228 yerüstü suyunun %58,6'sı evsel olan katı atıklar ile,
- 165 yerüstü suyunun %42,4'ü hayvancılık faaliyetleri sonucunda,
- 138 yerüstü suyunun %35,5'i endüstriyel atık su sonucunda kirlenmiştir (T.C. Çevre ve Şehircilik Bakanlığı, 2019, s. 70).

2017 yılında edinilen bilgiler sonucunda Amasya ilinde, Yeşilirmak, Tersakan, Çekerek ve Deliçay'ın muhtemel kirlenme nedenleri; tarımsal ilaç ve gübreleme, evsel atık su, endüstriyel kaynaklı atık su, evsel katı atık ile hayvancılık olarak belirlenmiştir (T.C. Çevre ve Şehircilik Bakanlığı, 2019, s. 61).

Su kirliliğinin denetimi için oluşturulan pek çok önlemin ve yaptırımın varlığı özellikle yüzey sularında sıkı takibi gerekli kılmıştır. Son yıllarda veri madenciliğindeki gelişmelere bağılı olarak ve *on-line* ölçüm istasyonlarının kurulması bilgisayar destekli izleme ve kontrol sistemlerinin gelişmesine olanak sağlamıştır. Bu çalışma kapsamında Yeşilirmak Havzası üzerinde kurulu olan istasyonlardan 2007-2009 yılları arasında toplanan veriler (TÜBİTAG-KAMAG-105G002 destekli proje kapsamında) kullanılmış ve bu veri setindeki aykırı değerlerin veri setinden temizlenmesi gerçekleştirilmiştir. Bu kapsamda 5 farklı aykırı veri ayıklama yöntemi kullanıcı etkileşimli ara yüze (*user friendly graphical user interface, GUI*) entegre edilmiştir. Bölüm 2'de çalışma bölgesi, parametreler ve aykırı veri ayıklama yöntemlerine ilişkin detaylı bilgiler verilmiştir.

## 2. BÖLÜM

### YEŞILIRMAK HAVZASI VE ÖLÇÜM İSTASYONLARI

#### 2.1. Yeşilirmak Havzası Genel Bilgiler

Yerüstü sularının planlanması ve yönetimi için ülkemiz hidrolojik açıdan 26 adet akarsu havzasına ayrılmıştır (Akkaya, 2002, s. 20). Ancak, T.C. Tarım ve Orman Bakanlığı ve ona bağlı olan Devlet Su İşleri Genel Müdürlüğü akarsu havzalarının sayısını 25 olarak tekrar güncellemiştir (Devlet Su İşleri, 2019). Şekil 2.1’de Devlet Su İşleri Genel Müdürlüğü’nün 2019 yılında yayımladığı 25 adet olan Türkiye akarsu havzaları haritası gösterilmektedir.

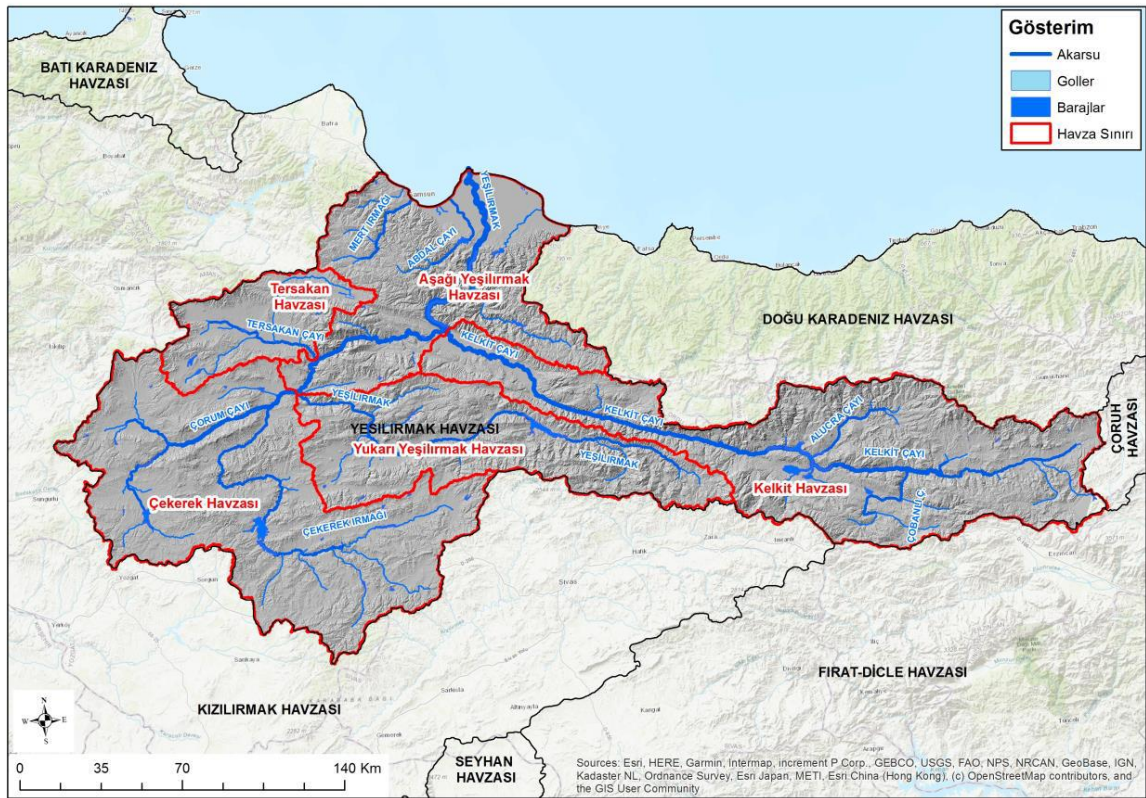


Şekil 2.1. Türkiye akarsu havzaları

Türkiye’deki suların çevrimi ve dağılımı dikkate alınarak ayırt edilmiş 25 tane akarsu havzasından biri olan Yeşilirmak Havzası çeşitli çalışmalara konu olmuştur (Yetik, 2007, s. 5; Köle ve Ataol, 2016, s. 462). Yeşilirmak Havzası, yaklaşık 39.574,38 km<sup>2</sup>’lik yağış alanı ile Sivas, Tokat, Yozgat, Çorum, Amasya, Samsun illerini tamamen veya kısmen içerisine almaktadır (Karadurmuş, 2000, s. 3; Arıkan, 2021, s. 21-22). Tarım ve Orman Bakanlığı tarafından yapılan ölçüme göre Yeşilirmak havzası 3282 m yükseltide olup Türkiye’nin toplam alanının yüzde 5’ini, yenilenebilir su potansiyelinin ise %3,12’sini içerisinde bulundurmaktadır (Kale, 2018, s. 264). Havza, Türkiye’nin kuzeyinde yer almaktadır (T.C. Tarım ve Orman Bakanlığı, 2022, s. 21-22). Havzaya ismini veren Yeşilirmak Nehri, 450 km uzunluğunda olup Sivas ili Suşehri ilçesinden doğmaktadır (Karadurmuş, 2000, s.3.) Batıya doğru Tokat ovası ile Turhal



ovalarının ardından Amasya Ovası'na akmakta sonrasında kuzeyden Çarşamba Ovasına açılarak ova içinde geniş bir delta yaparak Karadeniz'e dökülmektedir. Nehir, Tersakan Çayı, Kelkit Çayı ve Çekerek Çayı olarak üç büyük kola ayrılmaktadır. Şekil 2.2'de gösterilen Yeşilirmak Havzası ve alt havzaları haritasına göre, mevcudunda Aşağı Yeşilirmak, Yukarı Yeşilirmak, Çekerek, Tersakan ve Kelkit olmak üzere beş adet alt havza bulundurmaktadır (Karadurmuş, 2000, s. 3; Arıkan vd., 2021, s. 21-22; T.C. Tarım ve Orman Bakanlığı, 2022, s. 21-22).



Şekil 2.2. Yeşilirmak Havzası ve alt havzaları haritası

## 2.2. Yeşilirmak Havzası Çalışma Alanı

Yeşilirmak havzası üzerinde 1989 yılında balık ölümlerinin yaşanması sonucu Tarım ve Köy İşleri Bakanlığının başvurusuyla 1991-1992; 1993-1994 yıllarında TÜBİTAK destekli iki farklı proje çalışmaları yürütülmüştür (Karadurmuş, 2000, s.6; Yetik 2007, s.4). Veriler 2007-2009 yıllarında TÜBİTAK tarafından desteklenmiş olan projeden elde edilmiştir (TÜBİTAK-KAMAG-105G002). Yeşilirmak nehri üzerinde yapılan çalışma ile Merkezi Gerçek Zamanlı Nehir Su Kalitesi İzleme Sistemi ve Ankara Üniversitesi Yeşilirmak İzleme Merkezinde (YİM) iki ölçüm istasyonundan Genel Paket Radyo Servisi (GPRS) ve internet sağlayıcılar (*Asymmetric Digital*

*Subscriber*, ADSL) aracılığıyla veriler toplanmıştır (Göz vd., 2019, s. 889). Bu ölçüm istasyonları, Aynalı Mağara İstasyonu (AU1) ve Durucasu İstasyonu (AU2) olarak isimlendirilmiştir. Şekil 2.3'te verilen Yeşilirmak Havzası Çalışma Alanı Haritasında iki istasyon işaretlenerek belirtilmiştir.



Şekil 2.3. Yeşilirmak Havzası Çalışma Alanı Haritası (Yüceer, 2005)

### 2.3. Yeşilirmak Havzasında Kirlilik Parametreleri

Yeşilirmak, Resmî Gazetede, Karadeniz için yayımlanmış olan parametreler doğrultusunda sınıflarına göre üçüncü sınıf sular kategorisinde, orta kalite olan kirlenmiş sularındır.

Türkiye'nin kuzeydoğusunda Amasya'da Yeşilirmak nehri yakınlarında inşa edilen *on-line* izleme istasyonlarından iki yıl boyunca (2007-2009) beş dakikalık zaman aralığında birçok veri *on-line* olarak toplanmıştır. İzleme istasyonlarından Çözünmüş oksijen derişimi (LDO) ve sıcaklığı, pH, iletkenlik, amonyum azotu (NH<sub>4</sub>-N), nitrat azotu (NO<sub>3</sub>-N), toplam organik karbon (TOC), klorür, o- Fosfat, bağıl nem, bulanıklık, askıda katı madde, debi ve kabin sıcaklığı ile dış sıcaklık gibi parametreler ölçülmüştür (Göz vd. 2019, s. 889).

#### 2.3.1. Çözünmüş oksijen

Çözünmüş oksijen (*Luminescent Dissolved Oxygen*, LDO) derişimi, sucul yaşamın devamı için gerekli organik madde miktarıdır. Organizmaların metabolik ve fotosentez olaylarının devamı için gerekli olan ve aynı zamanda da kirliliğin ve oluşan alg seviyesinin de bir belirteçidir

(Özkan vd., 2006, s. 31). Sulardaki çözülmüş oksijen, 1 mg/L'den aşağı hipoksik, 1-3 mg/L'de suncul canlılarda ölümcül, 5 mg/L'den aşağıda ise stres etkenidir. 6 mg/L balıkların larvadan çıkması, 7 mg/L büyümeleri, popülasyondaki yaşamlarının devamı için 9 mg/L ve daha fazlası gereklidir. Yerüstü sulardaki yaşam için ise bu oran 4 ile 15 mg/L aralığında olmalıdır. Bu oranın altı veya üstü canlı yaşamı için tehlikelidir. Bu bilgiler dikkate alındığında LDO derişimi oldukça önemlidir (Nacar vd., 2020, s. 1480; *Foundriest Environmental*, 2013). Nehir, akarsu ve sığ suların kalitelerinin bir ölçüsü olan LDO derişimi, mevsimsel ve günlük bir döngüye sahiptir. Akarsulardaki biyolojik, kimyasal ve fiziksel etkilerden bu etkilerin deęişimlerinden etkilenirler (Nacar vd., 2020, s. 1480; Rajwa vd., 2014, sy.).

Yeşilirmak İzleme Merkezinde (YİM) 2007-2009 yılları arasında izlenen ilk ölçümler DSİ 'nin 1993 yılı analizleriyle bir ön inceleme olarak kıyaslandığında; su kalitesinde en kritik deęişken olan çözülmüş oksijen açısından geriye gidışin varlığı söz konusu olmuştur.

### 2.3.2. Sıcaklık

Yüzey sularındaki kalite ve kirlilik parametresi olan LDO derişimini etkileyen en önemli deęişkenlerden biri de günün belli saatleri ile mevsimlere göre deęişen su sıcaklığı (T)' dir. Sulardaki yaşamın devamı sıcaklıkla ilişkilidir. LDO ise sıcaklığın baęlı olduęu parametrelere doğrudan veya dolaylı olarak baęlıdır. LDO deęeri ile su sıcaklığı arasında ters bir baęıntı vardır (Erdoğan ve Bakioęlu, 2021, s. 96). Su sıcaklığının artması sudaki oksijenin çözümlüęünü azaltmaktadır. Yüzey sularındaki LDO derişimi, sıcaklıkların düşük seviyede olduęu kış ve ilkbahar aylarında artmakta, sıcaklıkların yüksek seviye de olduęu yaz ve sonbahar aylarında ise azalmaktadır. (Nacar vd., 2020, s. 1994). Aynı zamanda su sıcaklığı, su ortamında bulunabilecek en yüksek miktardaki LDO derişimini ifade eden doyunluk derişimini etkilemektedir. Sudaki fotosentez veya hızlı bir sıcaklık deęişimi LDO doyunluęunun %100'ün üzerine çıkmasına sebep olabilir. Su sıcaklığı ve oksijen arasındaki ilişki histeretik ilişkiyi ortaya çıkararak döngü şekli ve yayılımı, ilgili nehre, yılın mevsimine ve hava koşullarına baęlı olarak deęişim göstermektedir (Rajwa vd., 2014, s. 9).

### 2.3.3. pH

Latince "potentia hydrojenii" kelimesinin kısaltılmışı olan pH kavramı, çözeltideki hidrojen iyonu miktarının göstergesidir.  $pH = -\log [H^+]$  şeklinde tanımlanır. Nötr suyun pH'ı 7 olup, eşit miktarda hidrojen ve hidroksit iyonu içerir. Hidrojen iyonunun artışı suyu asit yapar ve suyun pH'ı düşer. Hidroksit iyon miktarındaki artış ise suyu daha bazik yani alkali yapmaktadır (Ölmez ve Saraç, 2009, s.13). Atık suların nehir havzasına etkisinin en önemli göstergelerinden biri de pH'tır. Noktasal ve noktasal olmayan kaynaklı deşarjların sonucunda yüzeysel sularda pH deęişimi gözlenir (Giwa vd., 2008, s. 1). pH'ın 5 olduęu asidik suların verimlilięi düşük, bazik

karakterdeki suların ise verimliliği yüksektir. Akarsularda pH ile çözülmüş karbondioksit ters, bikarbonat ile doğru orantılı olarak değişim gösterir (Ölmez ve Saraç, 2009, s. 12). Sucul canlıların yaşaması pH oranının 6,5-8 aralığında olmasına bağlıdır. İnsanlar için, pH 4-11 aralığı tolerans düzeyidir. Bunun dışındaki değerler, göz ve cilt tahrişi meydana getirir. Bu oran 4'ün altında borularda korozyona sebep olup mobilize toksik metal riskini artırır. pH değişikliği oligotrofik (çözülmüş oksijen bakımından zengin) sularda, ötrofik (çözülmüş oksijen içermeyen) hale geldiğinde erişilebilir besinlerle, su bitkileri ve alglerin gelişmesine ve çözülmüş oksijen talebinin artmasına sebep olur. Bu olay besinler ve bitki yaşamı açısından zengin, ancak çözülmüş oksijen derişim değişimleri açısından düşük ötrofik bir ortam oluşturur. Ötrofik bir ortamda, pH seviyeleri uygun değer aralıkta olsa bile suda yaşayan diğer organizmaların strese girmesine sebep olur (*Foundriest Environmental*, 2013).

#### **2.3.4. İletkenlik**

Elektriksel iletkenlik (*Electrical Conductivity*, EC), suyun veya çözeltilinin elektrik akımını iletebilmesinin bir ölçüsü olarak sularda bulunan iyonların elektrik akımını iletme kapasitesi veya çözeltilinin elektrik akımını geçirmeye karşı gösterdiği direncin bir göstergesidir (MEB, 2014, s. 41). Berrak sularda iletkenlik fazla iken kirliliğin yoğun olduğu bulanık sularda azdır. Elektriksel iletkenlik suyun kalitesi hakkında fikir verebilmektedir. Su içerisinde çözülmüş iyonların artması elektriksel iletkenlik değerinin artmasına sebep olur. Tatlı sularda su kalitesi bakımından düşük olan suların elektriksel iletkenlik değeri yükselir. Deniz sularının elektriksel iletkenlik değeri yüksektir. (Erdoğan ve Bakioğlu, 2021, s. 96). EC, Düşük sıcaklık(T) ve yüksek atmosfer basıncı, sudaki LDO derişimini etkileyerek artıran faktörlerdendir (Nacar vd., 2020, s.1480). Elektriksel iletkenliğin artmasıyla sudaki tuz derişimi artmakta ve sulardaki çözülmüş tuz miktarı elektriksel iletkenliği göstermektedir. Tuz artışı ile birlikte iletkenlik artarak sucul canlılarda toksik yani zehirleyici etki meydana getirmektedir. Elektriksel iletkenliğin 0,64 ile bölünmesi sonucunda toplam tuz derişimi hesaplanmaktadır (Altuntaş, 2018, s. 8; MEB, 2014, s. 41). İyonize olmuş maddenin toplam derişimi ve sıcaklığına bağlı olup çözülmüş iyonların hareketliliği, yükü ve derişimi iletkenliği değiştirmektedir. Sulardaki iletkenlik artışı, suyun kirlendiğini ya da suya deniz suyunun karıştığının göstergesidir. Suda bulunan inorganik çözülmüş maddeler tarafından elektrik iletilir. Organik maddeler elektriği yeteri kadar iletmediğinden ölçümde dikkate alınmazlar. İçerisinde inorganik maddeleri elektrolit olarak bulundurmayan sular (deiyonize, saf su gibi) elektriği çok fazla iletmemektedir. Aynı şekilde sıcaklık arttıkça elektrik geçirgenlikte azalmaktadır. Bu nedenle suların elektrik geçirgenliğinin ölçülmesi içerisindeki elektrolit miktarının da bir ölçütü olarak alınır. Doğal sular da tuz çözeltileri olduğundan elektriği iletir. Tuzlar, elektrik iletkenliği bakımından değerlendirildiğinde cinslerinden çok miktarlarının önemli olduğu söylenebilir. İletkenliği yüksek olan sular büyük ölçüde korozif (metal yüzeylerinde aşınmaya sebebiyet verici) özellik gösterir. Çözeltideki iyonların iletkenlikleri birbirinden farklıdır. Suda çözünen

iyon ne kadar fazla ise iletkenlik o kadar fazla olur. Elektriksel direncin birimi ohm olduğundan iletkenlik birimi bunun tersi olan mho ya da Siemens'dir. Su analizlerinde Siemens birimi çok büyük olduğu için sonuçlar mikroSiemens cinsinden ifade edilir. Elektrolitik iletkenlik, metalik iletkenliğin tersine sıcaklık artınca artar ve 1°C sıcaklık artışı için %1,9 artış gösterir. Bu nedenle iletkenlik ölçümlerinin 25°C sıcaklıkta verilmesi uygundur (MEB, 2014, s. 41).

### **2.3.5. Toplam organik karbon**

Bir kirlilik parametresi olan toplam organik karbon (TOC) analizi ile sulardaki organik madde miktarı belirlenmektedir (Tekin, 1997, s. 3). Yüzey suları ve atık sudaki TOC, karbon içeren tüm organik maddelerin toplam içeriğini tanımlayan önemli bir kirlilik parametresidir. TOC derişimi suda oluşan çözünmüş ve parçacıklı organik bileşiklerdeki karbon içeriğinin bir ölçüsüdür ama organik maddelerin doğası hakkında hiçbir bilgi vermemektedir. Doğal sular, hem doğal (biyosentetik) hem de antropojenik yani insan kökenli organik maddeler içerir. TOC, yüzeysel nehir sularında ve atık su ekosistemlerinde karbon varlığı hakkında bilgi sağlayarak biyolojik bozunma ve saflaştırma derecesini belirlemede önemli bir faktör olan su kirliliği derecesinin bir ölçüsüdür. (Siepak, 1999, s. 282). TOC derişiminin suyun cinsine bağlı olmakla birlikte sıcaklık, tuzluluk gibi birçok parametre, pH, mikrobiyal aktivite ve çevredeki bitki varlığına da bağlı olduğunu düşünülmektedir. TOC değeri, oldukça değişken olup göl veya nehir sularında genellikle 2–10 mg/L' lik değer aralığındadır (Bayram vd., 2011, s. 85). Nehirlerdeki TOC değeri mevsimsel sıcaklıklar ve yağışlardan dolayı değişebilmektedir. Yağışların arttığı sonbaharda yağmur sularının seyreltme etkisiyle düşerken, yazın sıcaklıkların artışı ile yükselmektedir (Bayram vd., 2011, s. 94).

### **2.3.6. Nitrat azotu**

Nehir sularına azotlu gübrelerin, organik madde atıklarının karışması, sudaki nitrat ve nitrit seviyesinin yükselmesine sebep olmaktadır. Suda nitrat ve nitritin fazla olması ötrofikasyona neden olarak su kalitesinin bozulmasını hızlandırır (Erdoğan ve Bakioğlu, 2021, s. 96). Organik ve inorganik, hayvansal ve bitkisel, artık maddelerin suların dibinde birikerek sedimentleri oluşturarak bakteri ve mikroorganizmalara yaşam ortamı sağlamaktadır. Çözünmüş oksijeni kullanarak organik maddeyi aerobik olarak parçalayan mikroorganizmalar nitratı açığa çıkarırlar (Tekin, 1997, s. 3). Başlıca nitrat kaynakları, kimyasal gübrelerin düzenli kullanımı, evsel atık sular ve evsel atık depolama alanlarıdır (Güneş, 2019, s. 762). Sularda nitrat ve nitritin kaynağı organik maddeler, doğadaki bazı mineraller ve azotlu gübrelerdir. Yerleşim yerlerindeki nitrat ve nitritin çoğunlukla artık organik maddelerden kaynaklandığı belirtilmiştir. Organik madde yüksekliğinde nitrit ve nitrat derişimi organik madde yoğunluğundaki artış ile artarken çözünmüş O<sub>2</sub> değeri de azalır (Erdoğan ve Ertan, 201, s. 65). Alıcı ortamlardaki azot döngüsü için mevsimsel değişimleri incelemek gerekir. Azotlu

bileşikler, sucul sistemlere doğal yollarla gelen azotlu bileşikler, yağışlar ve toprakla temas halindeki yüzey sularının getirdiği maddeler ile mikroorganizmalar tarafından bağlanan azottan veya insan faaliyetleri sonucu geçmektedir. Ancak, bu ortamlara insan faaliyetiyle geçen azot yükü, doğal yolla geçene oranla çok daha fazladır. İnsan faaliyetleri sonucu sularda meydana gelen azot yükü, başlıca kentsel atık su arıtma tesislerinin çıkış suyu ile koklaştırma tesisleri, sunî gübre fabrikaları ve nitroselüloz fabrikaları gibi sanayilerin drenaj sularından kaynaklanmaktadır. Özellikle kentsel atık su arıtma tesislerinin çıkış suları, yağışsız havada 40-80 g/cm<sup>3</sup> azot içermektedir. Diğer taraftan, sucul sistemlerdeki insan faaliyeti orijinli azotlu bileşiklerin önemli bir kaynağını da doğal veya sunî gübreleme işlemleri oluşturmaktadır. Çünkü toprak, nitrat iyonlarını daha güç bağlayabilmekte ve bu iyonlar sonuçta drenaj sularıyla yıkanarak sucul sistemlere karışmaktadır. Bu iş için doğal gübrelerin kullanılması halinde de toprak bakterilerinin metabolik faaliyetleri sonucunda nitrat iyonları oluşacaktır. İçme suyu kaynaklarına karışan azotlu bileşikler içerisinde insan sağlığına en olumsuz etkilere sahip olan bileşik nitrat olmaktadır (Kurucuoğlu ve Algur, 2011, s. 13; Sáez vd., 2003, s. 97). Nitrat derişimi içme sularında 500 mg/L'i geçerse sindirim sistemi ve ürogenital iltihaplanmalara, nitrite indirgenirse 6 aylıktan küçük bebeklerde ise mavi bebek hastalığına sebep olmaktadır. Bunun yanında kanserojen etkilerinin de olduğu kaydedilmektedir (Kurucuoğlu Ökmen ve Algur, 2011, s. 13; Sáez vd, 2003, s. 97; Özdestan ve Üren, 2012, s. 12; Yücel vd. 2019; s. 78). Ortamda biriken nitritin de balıklar ve sucul canlılar için son derece zehirli olduğu bilinmektedir. Organik azot biyolojik olarak mineralizasyona uğrayarak önce amonyum azotu (NH<sub>4</sub>-N)' na, amonyum azotu ise nitrifikasyona uğrayarak nitrat azotu (NO<sub>3</sub>-N)' na dönüşmektedir (Bellitürk ve Danışman, 2007, s. 64; Güneş ve Aktaş, 1992, s. 502). Çevre Koruma Ajansı (EPA) ve Dünya Sağlık Örgütü (WHO) gibi kurumlar, içme sularında sırasıyla 10 ve 11,3 NO<sub>3</sub>-N mg /L sınır değerlerini belirlemiştir. Türk içme suyu standardı TS-226'ya göre içme suyunda müsaade edilen maksimum nitrat derişimi 45 mg/L'dir (Kurucuoğlu Ökmen ve Algur, 2011, s.13,).

Yerleşim yerlerindeki nitrat ve nitritin çoğunlukla artık organik maddelerden kaynaklandığı belirtilmiştir. Organik madde yüksekliğinde nitrit ve nitrat derişimi organik madde yoğunluğundaki artık ile artarken çözünmüş O<sub>2</sub> değeri de azalır (Erdoğan ve Ertan, 2016, s. 65).

### **2.3.7. Amonyum azotu**

Organik kaynaklı üre gibi azotlu bileşikler, hidrolize uğrayarak Amonyum azotu (NH<sub>4</sub>-N)' na dönüşür. Bu bazen atık su aracılığıyla nehir suyuna karışmaktadır. Atık suların içerisindeki organik maddenin göstergesi NH<sub>4</sub>-N ve azot bileşikleri olup evsel ve endüstriyel üretim modellerindeki değişimden de etkilenmektedir (Borzooei vd., 2019. s.454). Ayrıca, azotlu gübrelerin toprağa inorganik veya organik formlarda verilmesiyle nehir suyuna karışmaktadır. Organik azot biyolojik olarak mineralizasyona uğrayarak önce amonyum azotuna, amonyum

azotu ise nitrifikasyona uğrayarak nitrat azotuna dönüşmektedir (Bellitürk ve Danışman, 2007, s. 64; Güneş ve Aktaş, 1992, s. 502). Toplam azotun bileşenleri, organik azot, amonyum azotu, nitrit azotu ve nitrat azotudur. Atık suların içinde nitrit azotu ve nitrat azotu hemen hemen bulunmamakta ve toplam azot yaklaşık olarak Kjeldahl azotuna eşit olmaktadır. TKN (Toplam Kjeldahl Azotu) olarak adlandırılan Kjeldahl azotu toplam oksitlenmemiş azot olup organik azot ve amonyum azotunun toplamıdır (Sertkaya, 2015, s. 6). Su Kirliliği Kontrol Yönetmeliğine göre nitrat azotunun 5mg/L'yi geçmediği sular I. Kalite sulardan olmaktadır (Erdoğan ve Bakioğlu, 2021, s. 96).

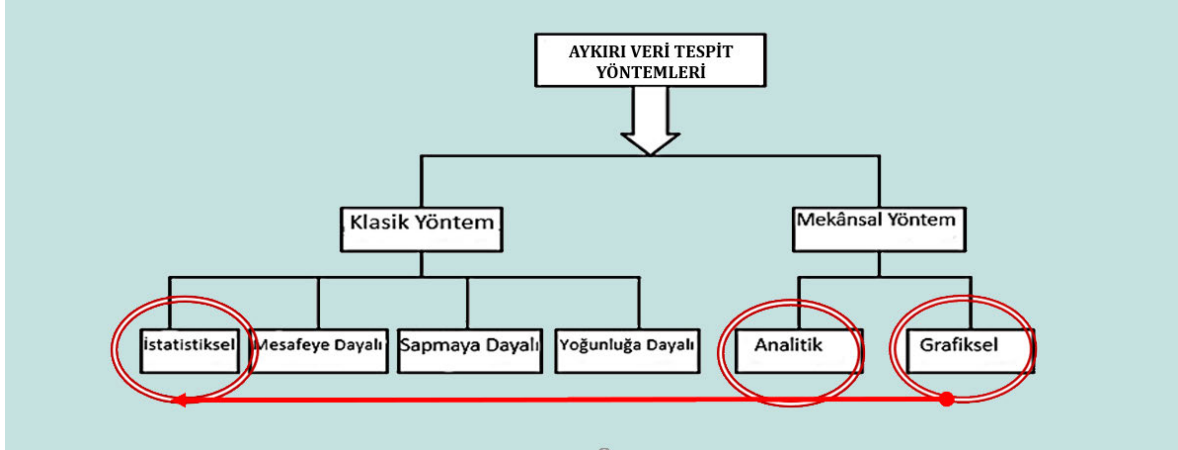
#### **2.4. Aykırı Değer Tanımı**

Aykırı değer temizleme; aykırı veri ayıklama, anormal veri (*outlier detection*), anomali gibi farklı isimlerle adlandırılmakta olup literatürde birçok tanımı mevcuttur. Yapılan araştırmaların sonucunu, oluşturulmuş veri setlerinden elde edilecek istatistiksel sonucu, değiştirebilecek olan, kayıp veya ortalamadan farklı veri alanına aykırı veri denmektedir (Atılğan ve Günay, 2012, s. 46; Hubert vd., 2008, s. 92). Aykırı değer, bir veri kümesindeki gözlemlerin çoğuyla tutarsız ve farklı görünen bir gözlemdir (Seem, 2007, s. 55). Bir aykırı gözlem veya aykırı değer, içinde meydana geldiği örneklemin diğer üyelerinden belirgin şekilde sapan gözlemdir (Chugh vd., 2014, s. 429; Grubbs, 1969). Aykırı değerler, kendi gerçek değerlerinden sapan gözlemler olarak yorumlanabilir ya da eksik gözlenen değerler veya yanlış ölçülen değerler olarak da yorumlanabilirler (Hawkins, 1980, s. 14; Wu, 2009, s.34). Aykırı değer tespiti, belirsiz bir veri kümesindeki anormal davranışları tespit eden, gürültüyü ortadan kaldırmak ya da anlamlı bilgileri keşfetmek için kullanılan önemli bir veri madenciliği tekniğidir (Aggarwal, 2013; Yan vd., 2017, s.827). Aykırı değerler, gözlemlerin ve verinin çoğunluğu ile aynı yapıyı göstermeyen ya da çözümlemelerde kullanılan genel varsayımlardan sapmalar gösterir (Atılğan vd., 2012, s.; Hubert vd. 2008, s.).

#### **2.5. Aykırı Değerlerin Tespiti**

Aykırı değer tespiti (aykırı veri ayıklama), verilerin hacmine, verilerin dağılımına ve verilerin değişken sayısına bağlıdır (Esen ve Timor, 2019, s. 268). Tüm veri kümelerine uyan evrensel bir yaklaşım olmadığından, uygulanabilir aykırı değer tespit yönteminin seçilmesi insan müdahalesi ve analizi gerektirir (Koren, 2023, s. 2). Aykırı değerler temelde, noktasal aykırı değer, içeriğe dayalı aykırı değer ve toplu aykırı değer olmak üzere üç kategoriye ayrılmaktadır. (Ceyhan, 2019, s. 8; Bansal vd., 2016, s. 374). Noktasal aykırı değerde belirli bir küme içerisindeki herhangi bir veri noktası kalan verilerle tutarsız olursa bu veri noktasal aykırı değer olarak tanımlanmaktadır. Veri setinin içeriğinin, diğerlerine göre tutarsız olduğu durum içeriğe dayalı aykırı değerdir. Toplu aykırı değer veri setinin içeriğinin kendi içinde tutarlı olup, tüm veri setine göre tutarsız olmasıdır.

Aykırı veri türleri literatürde çok farklı olarak ele alınmış ve işlenmiştir. Bunun genel sebebi her sektörün verilerinin kendine has değerler barındırması ve çözüm olarak çoklu yöntemlerin bulunmasıdır. Aykırı verilerin ayıklanmasındaki yöntemler klasik ve mekânsal olarak iki sınıfa ayrılmaktadır. Yapılan bu ayırım Şekil 2.4'de Aykırı Veri Tespit Yöntemlerinin sınıflandırılması ile gösterilmiştir (Xi, 2008, s. 95; Bansal vd., 2016, s. 374).



Şekil 2.4. Aykırı Veri Tespit Yöntemleri

Mekânsal değerlendirmelerde grafik yöntemler ve analitik yöntemler olmak üzere iki durum söz konusudur (Chelishchev, 2018, s.1-2). Analitik yöntemler, yollar, binalar ve şehirler gibi analitik olarak tanımlanabilen nesnelerin oluşturduğu veri setinin analiz edilebilmesi için gerekli olan üç boyutlu yaklaşımlardır. Analitik veriler konum, şekil, geometrik veya topolojik unsurlar barındırmaktadır. Analitik verilerdeki aykırı verilerin tespit edilmesi birçok uygulamada önemli bilgiler ihtiva eder. Orman yangınları, su taşkını, fırtına gibi şiddetli meteorolojik olaylar, karayolu tıkanıklıkları, uydu görüntüleri ile askeri hedeflerin saptanması, petrol rezervlerinin potansiyel konumlarının bulunması ve su kirliliğinin tespit edilmesi bu uygulamaların çalışma alanlarındandır (Ceyhan, 2019, s. 8; Xi, 2008, s. 96). Grafiksel yöntemler ise kutu grafikleri ve histogramlar yardımı ile aykırı verilerin ayıklanmasında bir diğer yöntemdir.

### 2.5.1. Aykırı değer temizliğinin önemi

Günümüz dünyasındaki teknoloji, sanayi devriminin sonucudur. Buharlı makinelerden, elektrikli makinelere, silikon vadisinden, yapay zekâ programlarına kadar her geçen gün teknoloji ilerlemektedir. Bu gelişmeler her alanda olduğu gibi veri analiz cihazlarının ölçümlerini daha hassas yapmasına ve büyük verilerin toplanıp depolanmasına imkân veren çalışmaların yapılmasını sağlamıştır. Çok büyük kapasiteli akan ve depolanan veriler, doğru sonuçlardan saparak ölçümlerinin yanlış, kayıp veya eksik olması gibi çözülmesi gereken



problemleri de beraberinde getirmiştir. Yıllar önceki veri analiz teknolojisi ile yapılan analizler, içinde birçok veri problemini barındırmaktadır. Depolanan verilerin sağlıklı ve anlamlı olması gerektiği aşikârdır. Çünkü analiz sonuçlarının doğruluk oranı ne kadar yüksek ise bilimsel analiz sonuçlarındaki başarı da o kadar yüksek olacaktır.

Aykırı değerler farklı sebeplerle oluşabilmektedir, bunlar; ölçüm hatası, örnekleme hatası, örnekleme sonuçlarının kasıtlı veya kasıtsız olarak eksik veya fazla raporlanması, mekânsal ve zamansal ölçüm değişkenlikleri, yanlış kayıt, veri kümesinin yanlış dağılımı veya model varsayımları, nadir gözlemler, cihaz ve sensör hataları gibi sebeplerdir.

Aykırı değerlerin belirlenmesi ve temizlenmesi ile verilerin yorumlanması kolaylaşacaktır. Aykırı değerler, veri kaynağında veya oluşturma mekanizmasında bulunan bilgileri bozmakta ve azaltmaktadır. İmalat endüstrisinde süreç ve ürün tasarımı ile kalite kontrol prosedürlerinin etkinliğini zayıflatacaktır. Olası aykırı değerler mutlaka kötü veya aykırı değildir. Aykırı değer tespiti, kredi kartı dolandırıcılığının algılanmasında, ağa izinsiz girişin tespit edilmesinde, sermaye piyasasında, hava tahminlerinde gürültülü görüntülerin algılanmasında çok önemlidir (Yan vd. 2017, s. 827). Örneğin adli bilimler, güvenlik, bankacılık gibi sektörlerde aykırı veri dolandırıcılık teşebbüslerini ortaya çıkarmaktadır. Güvenlik açığının önlenmesi aykırı veri varlığının belirlenmesiyle olacaktır. Aykırı değerler bazen temel bilgileri taşıyabilmekte olup doğru bir şekilde tanımlanmalıdır. Ölçüm süreçlerinde aykırı değerlerin incelenmesi, takip edilmesi ve saptanması, süreçlerin daha iyi anlaşılabilir ve iyileştirilmiş çıkarımlarla sonuçlanacak uygun veri analizine yol açacaktır (ISO16269-4, 2010). Aykırı değer tespiti, veri setinde beklenen modele uymayan veya beklenen davranıştan büyük ölçüde sapma gösteren kalıpları bulmak için son yıllarda fazlaca kullanılan bir prosedürdür (Liu vd., 2020, s.1). Bu veriler kayıt hatasından, üretim aşamasında yaşanan bir bozulmadan veya insanlarla ilişkili sebeplerden meydana gelebilirler. Aykırı veriler model kurma hatasına, yanlış parametre tahminlerine veya yanlış analiz sonuçlarına sebep olabilirler (Liu vd. 2004, s.1635). Aykırı değerlerin, diğer normal verilerden sapma gösteren gerçek veri içerisinde rastgele oluşmuş olan gürültüden ayırt edilmesi gerekmektedir. Rastgele oluşan gürültü araştırmacı için önemli bir konu olmayabilirken gürültü ve aykırı değer tespiti aynı derece önemli bir veri madenciliği görevidir. Bu nedenle etkili veri analizi için hem gürültü hem de aykırı değer tespiti aynı derecede öneme sahiptir. Bu bağlamda aykırı noktaları tespit ederek bu noktaları gürültüden ayırmak için daha öncesinde geri bildirim alınmış bilinen aykırı değerleri kullanmak en iyi çözümü sunar (Ceyhan E, 2019. s.2). Aykırı veri ayıklama, geniş uygulama alanına sahiptir. Çeşitli bilim dallarında aykırı veri analizi ile ilgili çalışmalar mevcuttur (Bansal vd. 2016; Moustafa vd., 2019; Zhao vd., 2022; Dilmaç ve Şişman, 2023; Ding vd., 2021; Aslam, 2020).

Son yıllarda nehirlerdeki kirlilik parametrelerinin değişimini takip ve tespit etmek, bunun sonucunda da gerekli önlemlerin alınması ve yaptırımların uygulanması için merkezi gerçek zamanlı nehir izleme ve kirlilik kontrol sistemleri ve nehir kirlilik verisi elde eden benzetim

modelleri geliştirilmiştir. Çalışan sistemdeki potansiyel kirlenmeyi tespit etmek için verileri aykırı ve normal değer olarak belirlemek gerekmektedir (Housh ve Ostfeld, 2015, s.183). Yeşilırmak Nehri üzerinde kurulu olan iki ölçüm istasyonundan analitik yöntemlerle toplanan veriler, Merkezi Gerçek Zamanlı Nehir İzleme ve Kirlilik Kontrol Sistemi projesi kapsamında Ankara Üniversitesi merkezli bir ofiste bulunan bilgisayarlarda kayıt altına alınmıştır. Bu çalışma alanında çeşitli sebeplerle oluşan olumsuz durumlar, istasyon verilerinde anlamsız, kayıp, eksik veya fazla birçok veriye sebep olmuştur. Aykırı verilerin oluşma sebepleri aşağıda sıralanmıştır.

- Veri iletişimindeki aksaklıklar, GPRS ve ADSL kesintileri, kayıp verilere sebep olmuştur.
- İklim koşullarında oluşan anormal değişimler. Örneğin, Ocak 2008 yılında ilk defa oluşan buzlanma numune alınmasının önüne geçmiştir.
- Nehir seviyesinde oluşan ani yükselmeler, mevsimsel yağışların artması, baraj kapaklarının açılması sonucu debinin artması taşkın sonucu istasyonda mekanik sistemin zarar görmesi sonucu numune alınmasında sorunlar oluşmuştur. 22.04.2008 tarihindeki nehir seviyesinde yaşanan ani yükselmenin Resim 2.1.'de şamandıra sistemini bükteğü görülmektedir. Bunun sonucunda su çekme sistemi zarar görmüş ve bir müddet numune alınamamıştır.



**Resim 2.1.** Nehir suyunda ani yükselme sonucu şamandıranın bükülmesi

- Nehir suyunda biriken kirlilik (doğal, evsel, endüstriyel, tarımsal) ve bırakılan molozlar sonucu oluşan tıkanıklık Resim 2.2.'de görüldüğü gibidir. Temiz bırakılan çalışma alanı kirlenince istasyonda aykırı veriler oluşturmaktadır.



**Resim 2.2.** Çalışma alanında kirliliğin temizlenmesi

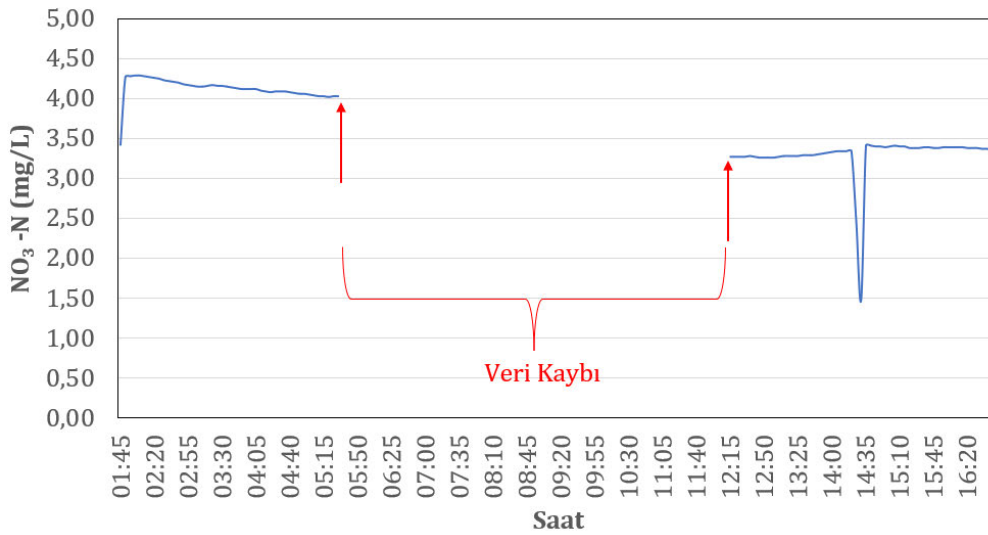
- Kirliliğe bağlı olarak aşırı çamur birikimi oluşmaktadır. Resim 2.3.'te görülen bu durumun sonucunda nehirlerde meydana gelen aşırı çamurlanma analitik ölçümler ile cihaz değerleri arasında büyük farka sebep vermektedir. Örneğin çözülmüş oksijen derişimi gerçek değerinin oldukça altında ölçülmüştür.



**Resim 2.3.** Kirliliğe bağlı çamur birikmesi

- Mevsimsel şartlardan etkilenerek, örneğin yazın çok yükselen sıcaklık sonucu oluşan buhar cihazların ve sensörlerin arızalanmasına ve kayıp verilere sebebiyet vermektedir.
- İnsanların ve alıcı ortama kirli su deşarj eden fabrikaların izlenmemek için istasyona kasıtlı müdahalesi.

Sistematik olarak oluşan cihaz ve sensör hatalarının sebeplerinden biri de kalibrasyon zamanlarının gelmesidir. Çalışmaların yapıldığı Yeşilırmak Havzasında Şekil 2.5'ten de anlaşılacağı üzere  $\text{NO}_3$  azotu deęerinde istasyon verilerinin kesintisi bulunmaktadır. 14.07.2008 tarihinde kaydedilen bu verilerde saat 05.30 ile 12:15 arasında veri kesintisi meydana gelmiştir. Bu kesintinin oluşturduğu boşluklar aykırı veri olarak deęerlendirilmekte ve veri setinden atılmadığı takdirde sonuçları olumsuz etkilemektedir.



**Şekil 2.5.** Yeşilırmak havzası çalışma alanı  $\text{NO}_3\text{-N}$  parametresinde kayıp veri oluşumu

- Rastlantısal hatalar
- Eysel kullanımın yoğun olduğu zaman olan sabah ve akşam saatlerinde amonyak deęerinin normalden saparak aykırı veri oluşturması yanlış ölçümler gibi sebeplerle aykırı veriler oluşmuştur.

Tablo 2.1' de 2. istasyon olan Durucasu' dan alınan kirlilik parametrelerine ait ham verilerde istatistiksel deęerler incelendiğinde pH deęerinin alt deęerlerde negatif olması aykırı deęeri göstermektedir. İzleme istasyonunda toplanan kirlilik parametrelerine ait ham verilerin gerçek ve ortalama deęerden sapması sonucu oluşan aykırı deęerler; yanlış analiz sonuçlarına, yanlış parametre tahminlerine ve modelleme hatasına sebep olabileceği için uygun yöntemlerle ayıklanmalıdır. Modelleme çalışmaları öncesinde grafiksel yöntemler veya klasik yöntemlerden olan istatistiksel yöntemlerle elde edilen parametreler için aykırı verilerin ayıklanması çalışmaları yapılmıştır.

**Tablo 2.1.** İkinci istasyon (Durucasu İstasyonu) ham verilerinin istatistiksel değerleri

2. İSTASYON DURUCASU HAM VERİLERİNİN İSTATİSTİKSEL DEĞERLERİ							
Parametre	LDO (mg/L)	T (°C)	pH	EC (µS/cm)	TOC (mg/L)	NO <sub>3</sub> -N (mg/L)	NH <sub>4</sub> -N (mg/L)
Üst Değer	10	31,4	8,41	797	22,58	6,12	4,24
Alt Değer	0	0	-3,46	0	0	0	0
Ortalama Değer	6,933	21,435	7,261	624,729	20,277	2,031	0,0313
Standart Sapma	2,772	3,186	1,320	185,644	3,669	1,076	0,1038

### 2.5.2. Aykırı değer temizleme testleri

İstatistiksel testler, parametrik ve parametrik olmayan testler olarak temelde ikiye ayrılmaktadır. Veri seti normal dağılım özelliği gösterirse parametrik, normal olmayan bir dağılım gösterir ise, parametrik olmayan bir test olmaktadır (Gençoğlu vd., 2017, s. 627). Çalışmada kullanılan Grubbs, GESD (*Generalized Extreme Studentized Detection*) testler normal dağılım gösteren parametrik testlerdir (Üçkardeş vd., 2010, s. 43.) Medyan, ortalama, Grubbs, GESD yöntemleri ve çeyrekler açıklığının ötesinde gözlemler gibi aykırı değerlerin bulunması için çeşitli yöntemler geliştirilmiştir ve her yöntemin kendine has özellikleri vardır (Ding vd., 2021, s. 4). Aykırı değerleri otomatik olarak temizlemek için yaygın olarak kabul edilen bir sistematik yöntem yoktur. Her çalışmada farklı metodolojiye sahip yöntemler ile testler tercih edilmekte ve geliştirilmektedir.

Aykırı verilerin ayıklanmasında kullanılan yöntemlere ait detaylı bilgiler aşağıda sunulmuştur.

#### 2.5.2.1. Ortalama

Ortalama, (Aritmetik ortalama, *mean*), bir grup verinin dağılımındaki değerlerin toplamının, dağılımdaki veri sayısına bölümüdür. Aykırı değerler, ortalamadan üçten fazla standart sapma olan öğeler olarak tanımlanır. Bu yöntem daha hızlıdır ancak ortanca yöntemine göre sağlamlığı daha azdır.



Önce bir hipotez kurulur;

$H_0$ : Veri seti aykırı değere sahip değildir.

$H_1$ : Veri seti tek aykırı değere sahiptir.

$$Y_{\min} = \min(x_i)$$

$$Y_{\max} = \max(x_i)$$

Tek taraflı sol kuyruklu test minimum değer in aykırı değer olma ihtimalini,

$$G_{\min} = \frac{\bar{x} - x_{\min}}{\sigma}$$

Tek taraflı sağ kuyruklu test maksimum değer in aykırı değer olma ihtimalini,

$$G_{\max} = \frac{x_{\max} - \bar{x}}{\sigma}$$

Çift taraflı test ise minimum ve maksimum iki değer in aykırı değer olma ihtimalini araştırır.

$$G = \max(G_{\max}, G_{\min})$$

$$G_{\text{crit}} \approx \frac{(N-1) t_{\alpha/(2N), N(N-2)}}{\sqrt{N(N-2) + t_{\alpha/(2N), N(N-2)}^2}}$$

$x_i$	:	Aykırı değer olma ihtimali olan veri
$\bar{x}$	:	Aritmetik Ortalama
$\sigma$	:	Standart sapma
$N$	:	Toplam veri sayısı
$A$	:	Önem düzeyi
$\frac{\alpha}{2N}$	:	Anlamlılık düzeyi
$t_{(\alpha/(2N), N)}$	:	t dağılımın kritik tablo değeri
$N - 2$	:	Serbestlik derecesi

Grubbs yöntemin olumsuz yanı maskeleye etkisidir (Chelishchev vd. 2018, s.3). Birden fazla aykırı değer içeren veri kümeleri, maskeleye ve bataklık etkilerine tabidir (Ben-Gal, 2005, s. 8). Maskeleye etkisi, tanımlanan bir aykırı değer başka bir aykırı değeri kapsadığında gerçekleşir. İlk aykırı değer kaldırıldığında, ikinci şüpheli değer bir aykırı değer olarak ortaya çıkar. Bataklık etkisi h, ikinci aykırı değer yalnızca birinci aykırı değer in varlığında bir aykırı değer olduğunda ortaya çıkar (Vera vd, 2013, s.1320). Maskeleye; birden fazla aykırı değer varlığının, her bir aykırı değer in tespit edilmesini zorlaştırmasıdır (ISO16269-4, 2010).

#### 2.5.2.4. Gesd Testi

GESD (*Generalized Extreme Studentized, Genelleştirilmiş Uç Student Sapma*) testi, Grubbs (1969) istatistiğinin Rosner (1975, 1983) tarafından genelleştirilmesine dayanan sıralı ve iki taraflı aykırı değer testidir (Cohn vd., s. 5048). Rosner testi olarakta bilinen GESD testi, aykırı veri sayısının tam olarak bilinmediği, neredeyse normal dağılım gösterdiği durumda kullanılır. Tekli ve çoklu iki uçlu aykırı verileri hesaplayan parametrik bir aykırı değer testidir.

Grubbs testinde olduğu gibi bir hipotez kurulur;

$H_0$ : Veri seti aykırı değere sahip değildir.

$H_1$ : Veri seti tek aykırı değere sahiptir.

$$R_i = \frac{x_i - \bar{x}}{\sigma}$$

$x_i$	:	Aykırı değer olma ihtimali olan veri
$\bar{x}$	:	Aritmetik Ortalama
$\sigma$	:	Standart sapma

GESD testi, uzaktaki gözlemleri teker teker tanımlar. Bu test, Grubbs testinin yapmadığı, dış gözlemlerin sayısına dayalı olarak kritik değerler için uygun ayarlamalar yapar. GESD testinin bir başka avantajı, şüpheli aykırı değer sayısında yalnızca bir üst sınır gerektirmesidir (Grubbs, 1969).

#### 2.5.2.5. Çeyrekler açıklığı

Çeyrekler açıklığı (*Interquartile, IQ*), veri setinin %75'inci çeyrek (Q3) ile %25'inci çeyrek (Q1) arasındaki farkı olan %50'lik kısımdır. Bu kısım veri setinde normal, temiz veri olarak kabul edilir. Çeyrekler açıklığı verilerin ortancadan (Q2) ne kadar saptığını da göstermektedir. Veri setinin, çeyrekler açıklığı değerinin 1.5 kat altı ve çeyrekler açıklığı değerinin 1,5 kat üstü aykırı değer olarak kabul edilir. Matematiksel hesaplamalarda olduğu gibi MATLAB ortamında da aykırı değerler için eşik, üst çeyreğin üzerinde veya alt çeyreğin altında 1,5 çeyrekler arası aralıkla ayarlanmalıdır. Ayarlanan eşik haricindeki bütün değerler aykırı veri olarak kabul edilmelidir (Bonakdari and Zeynoddin, 2022, s.27). Bu yöntem verilerin dağılımını normal değilmiş gibi algılar (Mathworks, R2023a, ty).



## 3. BÖLÜM

### LİTERATÜR ARAŞTIRMASI

#### 3.1. Nehir Kirliliği Modelleme ve Aykırı Veri Ayıklama Çalışmaları

Liu vd. (2020), Amerika Birleşik Devletleri'nin bir eyaleti olan Batı Virginia'daki Potomac Nehri'nin yüzey suyu kalitesindeki değişimler ile anormalliklerin analizi ve tespiti için çalışmışlardır. Nehir çevre kirliliğiyle önceden başa çıkmak ve erken uyarı sağlamak üzere yüzey suyu kalitesi parametreleri, bulanıklık (TURB), özgül iletkenlik (SC), çözünmüş oksijen (LDO) için aykırı veri ayıklamada etkili bir veri odaklı çerçeve geliştirilmişlerdir. Geliştirilen çerçeve ile yüzey suyu kalitesi değişimi tahmini için BAR (*Bayesian autoregressive*) modelini, su kalitesinde aykırı verileri ayıklamak için ise IF (*Isolation Forest*) algoritmasını oluşturmuşlardır. Aykırı veri temizliği için BAR modelinden elde edilen tahmin kalıntılarına IF algoritmasını uygulamışlardır. IF algoritmasına ait aykırı değer verileri, ROC eğrisinde; yanlış alarm olasılığı FPR (*False Positive Rate*, Yanlış Pozitif Oranı) değerlerine karşı apsis(x), tespit olasılığı olan TPR (*True Positive Rate*, Gerçek Pozitif Oranı) değerleri ordinat(y), olarak çizdirilmiştir. ROC eğrilerinin altında kalan alan AUC değerleri, doğruluğu değerlendirmek için kullanılmışlardır. AUC aykırı değer tespit doğruluğunu, TURB ve SC özniteliklerinin kullanımıyla maksimum (0,919), TURB-DO öznitelikleri için (0,797) ve SC -DO öznitelikleri için (0,805) bulmuşlardır. En yüksek doğruluk oranının TURB-SC özniteliği olması dolayısıyla aykırı verilerin temizliğinde etkin olacağını tespit etmişlerdir. Geliştirilen uyum çerçevesiyle, nehir çevre kirliliğinin yaptığı olumsuz etkiyi önlemesi ve acil müdahale için erken uyarı vermesi için Potomac Nehri'nin yüzey suyu kalitesindeki aykırı veri ölçüm sonuçlarına göre izlenebileceği sonucunu elde etmişlerdir (s.1-17).

Casillas-García vd. (2021), çalışmalarında, Meksika'da bulunan Santiago- Guadalajara Nehri (SGR)'nin su kaynağının yönetiminde daha iddialı kararlar alınmasının sağlanması amacıyla 2009-2019 yıllarının izleme verileri toplanmıştır. Nehirden elde edilen su kalite verilerinden BOD<sub>5</sub>, NH<sub>3</sub>-N, NO<sub>3</sub>, Pb, Zn, Cr, Cd, F, Hg 44 adet parametre kullanılarak öncelikle aykırı veri temizleme çalışması yapılmıştır. Kutu grafiği ile 1. Çeyrek(Q<sub>1</sub>), 3. çeyrek(Q<sub>3</sub>), çeyrekler açıklığı (IQ) ve ortanca hesaplanıp Grupps Testi uygulayarak aykırı verileri temizlemişlerdir. Santiago Guadalajara Nehri (SGR-WQI) için su yaşamının korunmasına yönelik özel yönergeleri etkili bir şekilde içeren ekosisteme özgü bir WQI (Su kalitesi indeksi) geliştirmişlerdir. Geliştirilen endeks, büyük bir kapasiteye sahip olarak örnekleme noktaları ve mevsimler arasında ayırım yapmak için tasarlanmıştır. Bu uzamsal ve zamansal değişimleri tam olarak yansıttığı için, Santiago-Guadalajara Nehri'nin özel bağlamındaki su kalitesi eğilimlerini değerlendirmek için kullanılabilir ve kolay tasarlanacak bir arayüze uygun bir yöntemdir (s.1-14).

Than vd (2021), çalışmalarında Vietnam'da bulunan Dong Nai Nehri üzerinde kurulu izleme istasyonlarından 2007'den 2019'a kadar toplanan izleme verilerini zaman serilerine göre su

kalitesini sınıflandırmak amacıyla kullanarak Hibrit LSTM-MA modelini geliştirmişler ve çeşitli modelleme çalışmaları ile karşılaştırmışlardır. Modelleme yapmadan önce aykırı verileri temizlemek için Grubbs testini kullanmışlardır. Sonuç olarak, tahmin etkinliğini kontrol etmek için kullanmış oldukları modeller ARIMA, NAR, NAR-MA ve LSTM, LSTM-MA ile çalıştıkları Hibrit LSTM-MA modeli ile karşılaştırıp elde ettikleri modelin daha hızlı eğitim süresine ve daha istikrarlı ve güvenilir tahmin edilebilirliğe sahip olduğunu ortaya koymuşlardır (s.1-17).

Zanoni vd. (2022), İtalya'da bulunan Adige Nehrinde, geliştirmiş oldukları iki makine öğrenimi algoritması olan *Deep feed-forward Neural Network* (DNN) ve *Random Forest* (RF) kullanarak su sıcaklığı, çözünmüş oksijen, arsenik, sülfat ve klorür derişimleri ile elektriksel iletkenliğin altı adet bölgesel modelini geliştirmişlerdir. Modelleme öncesinde, ampirik olasılık dağılımının 99. yüzdelik dilimini aşan verileri aykırı değerler olarak kabul edip sistemden kaldırmışlardır. Bu modelleri su kalitesi parametrelerini tahmin etmek için test ederek çalışmalarını standart Doğrusal Regresyon (LR) modeli ile karşılaştırmışlardır. RF ve DNN, tüm su kalitesi parametreleri için LR modelinden daha iyi performans gösterdiğini saptamışlardır. DNN ve RF arasındaki karşılaştırmada ise, birincisinin doğrusal olmayan ilişkileri tespit etmede daha esnek ve etkili olduğunu ve bu nedenle araştırılan su kalitesi değişkenlerinin modellenmesinde tercih edildiği sonucuna varmışlardır (s. 1-15).

## 4. BÖLÜM

### YÖNTEM

#### 4.1. Verilerin Toplanması

Türkiye Bilimsel ve Teknik Araştırma Kurumu (TÜBİTAK), Kamu Araştırmaları Destek Grubu (KAMAG) 105G002 kodlu, Merkezi Gerçek Zamanlı Nehir İzleme ve Kirlilik Kontrol Sistemi projesi kapsamında Yeşilirmak Nehri üzerinde kurulu olan Aynalı Mağara (AU1) ve Duracasu (AU2) *on-line* ölçüm istasyonlarından 5'er dakikalık sürelerle 12 tane parametre için 2007-2009 yılları arasında veriler toplanmıştır. Bu veriler GPRS ağı yardımıyla Ankara Üniversitesine aktarılmıştır. Bu parametreler; çözülmüş oksijen, pH, iletkenlik, nitrat azotu, amonyum azotu, toplam organik karbon, klor, orto fosfat, sıcaklık, iletkenlik, askıda katı madde ve akış hızı olarak sıralanabilir (Göz vd., 2019, s. 889). Tez kapsamında; çözülmüş oksijen, sıcaklık, pH, iletkenlik, toplam organik karbon, nitrat azotu ve amonyum azotu için aykırı verilerin ayıklaması gerçekleştirilmiştir.

#### 4.2. Grafikselsel Ara Yüz Tasarımı

Uygulamada grafikselsel arayüz geliştirmek için çeşitli programlama dilleri (Fortran, C++, C#, Java, Phyton, MATLAB gibi) mevcuttur.

**MATLAB** (*Matrix Laboratory*): Her mühendislik alanında kullanılabilen, ileri analiz, veri görüntüleme ve algoritma geliştirme sürelerini C, C++ veya Fortran gibi programlama dillerine göre oldukça kısaltan bir programlama dilidir (Baykal, 2014, s.46). 1970'li yıllardan günümüze gelişmekte olan MATLAB programı ve eklentileri, şu anda C, C++, Phyton, Java vb. programa dillerinin farklı sürümleriyle uyumlu olup MATLAB'ı başka bir program dilinden veya diğer program dilleriyle onlara ait sistem bileşenlerini MATLAB'dan çağırarak mümkün hale getirilmiştir (Mathworks, 1994-2023, sy.). Dezavantajı ise ücretli olmasıdır.

**Fortran**: İlk olarak 1957 yılında IBM (*International Business Machines*, Uluslararası İş Makineleri) tarafından geliştirilerek çeşitli sürümleri mevcut olan dil C yazılım dilinin gelişmesine öncü olan ilk yüksek seviyeli derlenebilir yazılım dilidir. C, C++ gibi dillerden üstün olsa da günümüzde çok tercih edilmemektedir (Ersoy ve Ersoy, 2018, s.1-2).

**C++**: 1969 yılında Unix sistemlerine programlama yapmak amacıyla çalışılan C programlama dilinin kapsamını temel alan C++ programı, 1979 yılında C'ye nesneye dayalı programlamayı ilave ederek günümüze geliştirilerek gelmiştir. Aynı zamanda C'nin standart kütüphanesindeki fonksiyonları destekleyen nesneye dayalı kendi kütüphanesini oluşturmuştur (Çakır, 2015. s.4).

C#: 1990'ların sonunda Microsoft tarafından C, C++ programlarını ve bu dilleri temel alan Java dilinin eksik tarafı olan diller arası uyumluluk probleminin giderilmesiyle geliştirilmiş, Windows platformuna uyumlu bir programlama dilidir (Schildt, 2011, s.17).

Java: 1991 yılından itibaren, C++ programının söz dizimi ve kullanımını alarak (Akçelik, 1995, s.16; Schildt, 2011, s.16) yılında tasarlanmış olan ücretsiz, *inovatif* ve dijital hizmetler (çevrimiçi bir şekilde aynı anda birçok kullanıcının etkileşimiyle işlem yapması) için geliştirilmiş olan bilgi işlem platformu, yazılım ve programlama dilidir (Java, ty, sy).

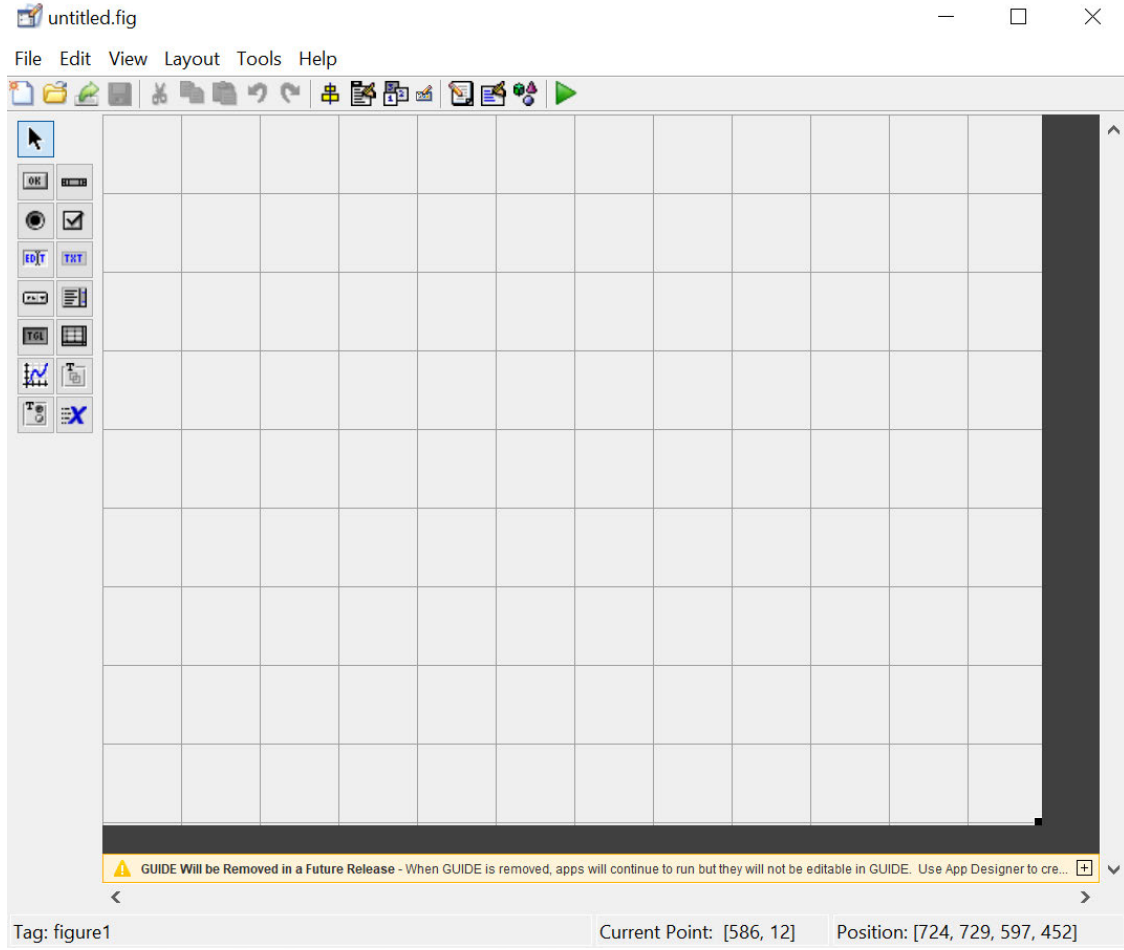
Phyton: 1990'larda geliştirilmeye başlanmış olan bir yazılımdır (Phyton, 2023, sy.) Veri bilimleri, kurumsal uygulama ve web tasarımı gibi kapsamlı kullanım alanları için geniş bir kitaplık koleksiyonuna sahip programlama dilidir. Phyton ve MATLAB programları arasındaki fark; Phyton'ın genel amaçlı bir bilgi işlem platformu olup MATLAB'ın ise mühendislik ve bilimsel uygulamalar için kullanılmasıdır (Mathworks, 1994-2023, sy.). Ücretsiz kullanıma açıktır.

Bu çalışmada geliştirilen kullanıcı etkileşimli grafiksel arayüz aşağıda belirtilen nedenlerden dolayı MATLAB (*Matrix Laboratory*) ortamında oluşturulmuştur.

- Kullanım kolaylığı sağlaması,
- Bilimsel çalışmalarda referans kabul edilmesi,
- Yaygın bir ağa sahip olması,
- Kimya mühendisliğinden, bilgisayar mühendisliğine, istatistiksel çalışmalardan tıp alanına kadar olan tüm meslek dallarının kullanması,
- Programın kendini güncel tutması,
- Üniversitelerde programlama eğitiminin verilmesi,
- Üniversitelerin dersliklerde öğrenciler için ücretsiz kullanıma açması,
- Birçok veri seti yükünü kaldırabilecek olması,
- İçerisinde birçok hata ayıklama yöntemlerinin tanımlı olması.

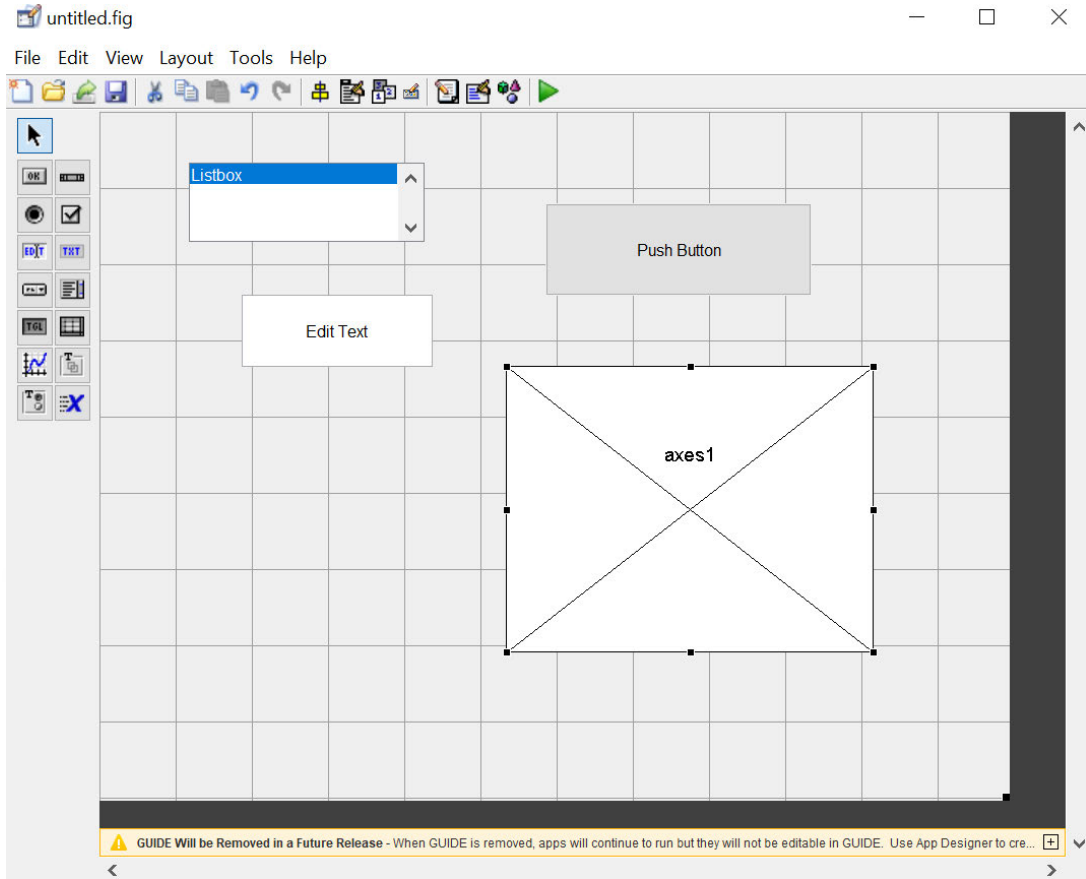
#### **4.2.1. Arayüzün oluşturulmasının görsel anlatımı ve ayrıntıları**

MATLAB Komut ortamına >>guide yazılarak enter tuşuna basılır ve GUI (*Graphical User Interface*) oluşturma penceresi açılır. Açılan GUI penceresi aşağıda Şekil 4.1 ile gösterilmiştir.



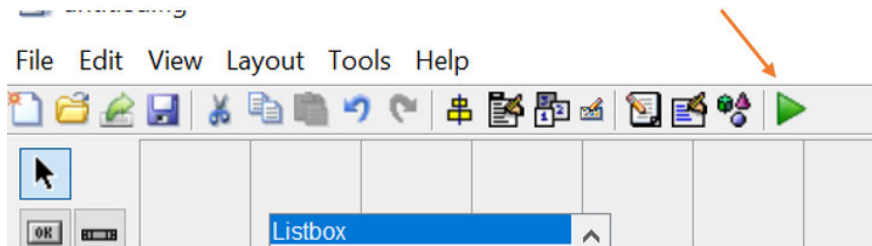
**Şekil 4.1.** GUI oluşturma penceresi

Burada GUI içinde yer alınması istenen objeler mouse yardımıyla sürükleyip bırakılarak açılan GUI penceresi içerisine Şekil 4.2 de gösterildiği gibi yerleştirilir.



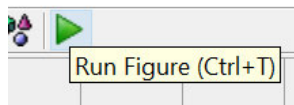
**Şekil 4.2.** GUI içerisinde yer alacak objeler

İstenilen tüm objeler pencere içerisine yerleştirildikten sonra aşağıda Şekil 4.3'te gösterilen şekilde olduğu çubuk üzerindeki gibi yeşil ok tuşuna (*Run Figure*) basılır.



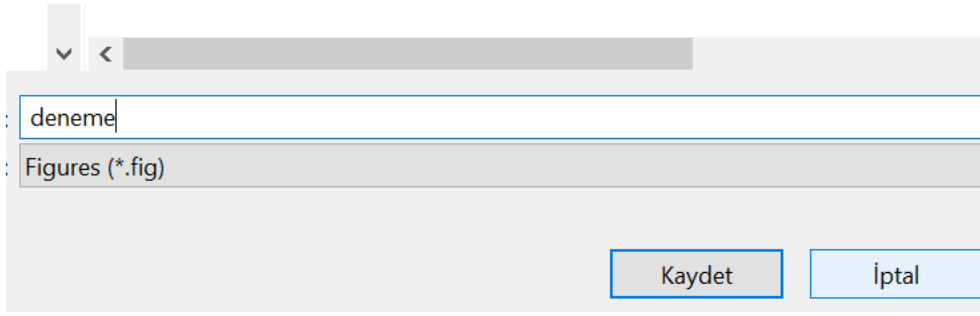
**Şekil 4.3.** Yeşil ok (Çalıştırma) butonunun bulunduğu konum

Şekil 4.4 ile yeşil ok tuşu (*Run Figure*) gösterilmiştir.



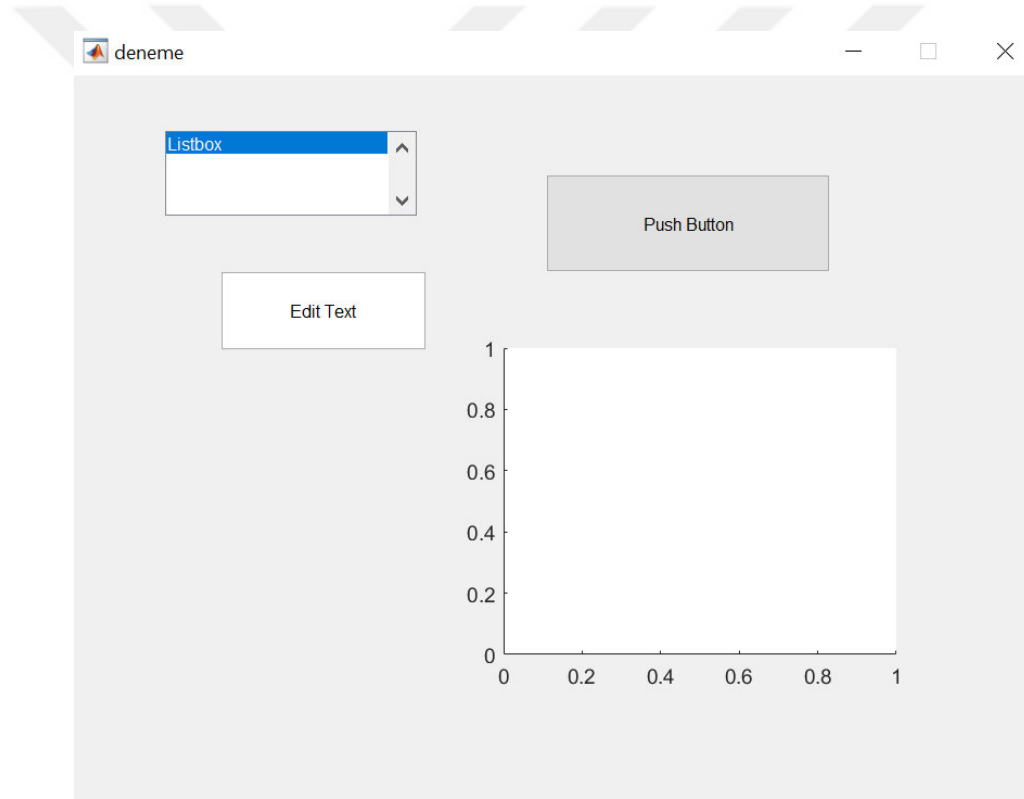
**Şekil 4.4.** Yeşil ok tuşu (*Run Figure*)

Hazırlanan Figure dosyasına bir isim verilerek Şekil 4.5'te olduğu gibi kaydedilir.



Şekil 4.5. Dosya ismi verme

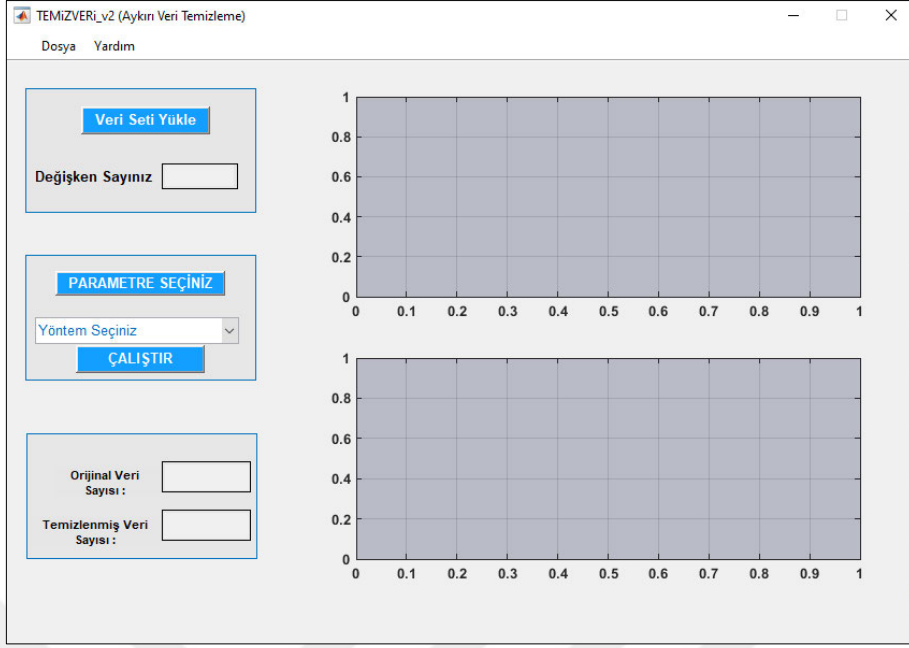
Sonrasında aşağıda bulunan Şekil 4.6'daki gibi bir görüntü elde edilir.



Şekil 4.6. Oluşturulan arayüzün görüntüsü

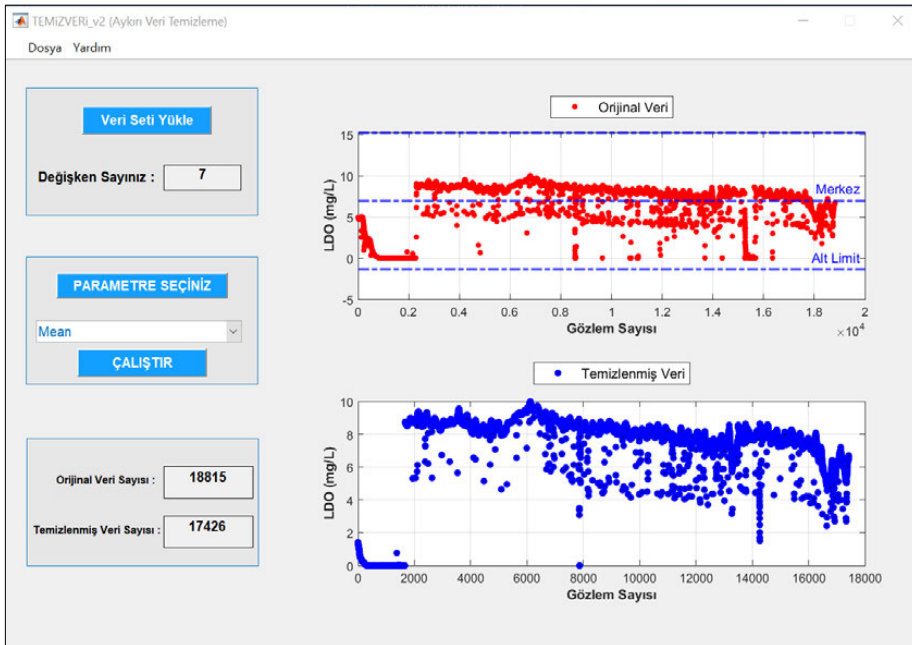
Böylece verilen isimle (burada deneme.m) figür dosyasına ilave olarak bir deneme.m dosyası daha oluşur. 'deneme. m' dosyasına ait örnek MATLAB kodları EK-1'de verilmiştir.

Bu çalışmada oluşturulan arayüzün genel görünümü Şekil 4.7'de verilmiştir.



**Şekil 4.7.** Arayüze ait genel gösterim

Şekil 4.8’de gösterildiği gibi Excel de kayıtlı olan veri seti yüklendikten sonra değişken sayısı ara yüzde görülmektedir. Temizlenecek değişkene ait sütun numarası girilmesinin ardından istenilen yöntem seçilip veriler temizlenebilmektedir.



**Şekil 4.8.** Arayüze ait detaylı gösterim

Hazırlanmış olan arayüze elimizdeki parametreler için farklı yöntemlerle hata ayıklama yapılmıştır.



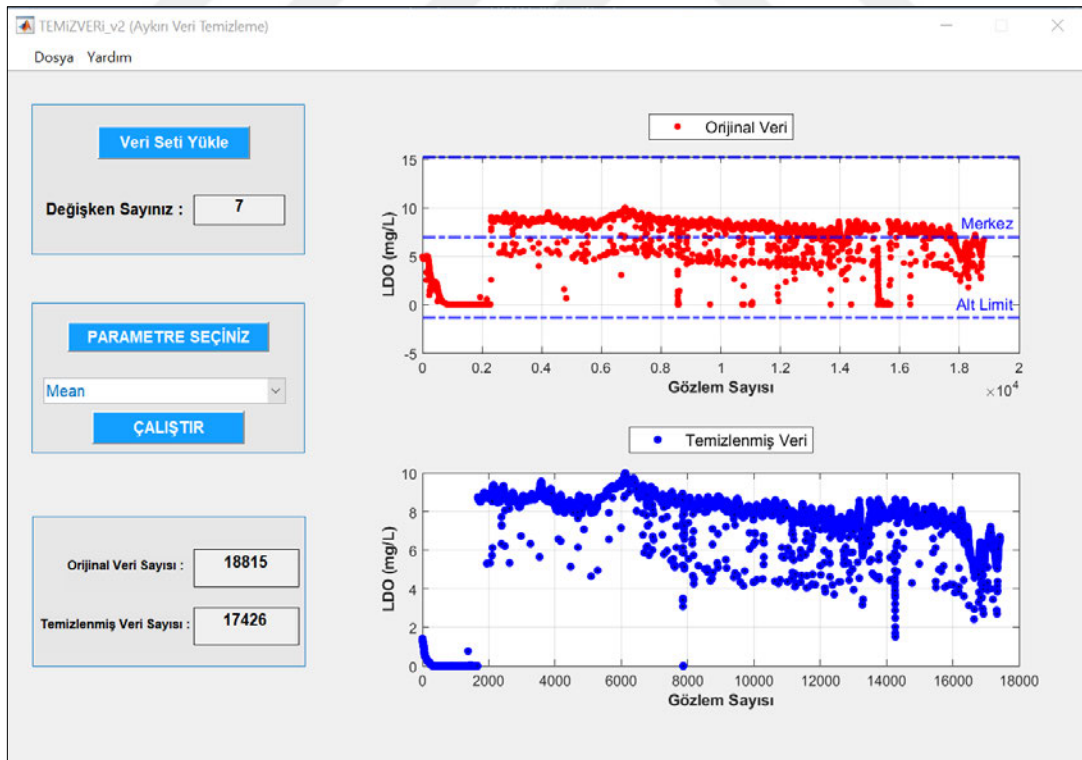
## 5. BÖLÜM

### SONUÇLAR VE TARTIŞMA

Yöntem bölümünde oluşturulan grafiksel arayüz ile bu çalışma kapsamında değerlendirilen parametreler için seçimli yöntemler kullanılarak elde edilen sonuçlar aşağıda verilmiştir. Ayrıca, tez çalışması kapsamında geliştirilen yazılımın kullanım kılavuzu EKLER bölümünde detaylı olarak açıklanmıştır.

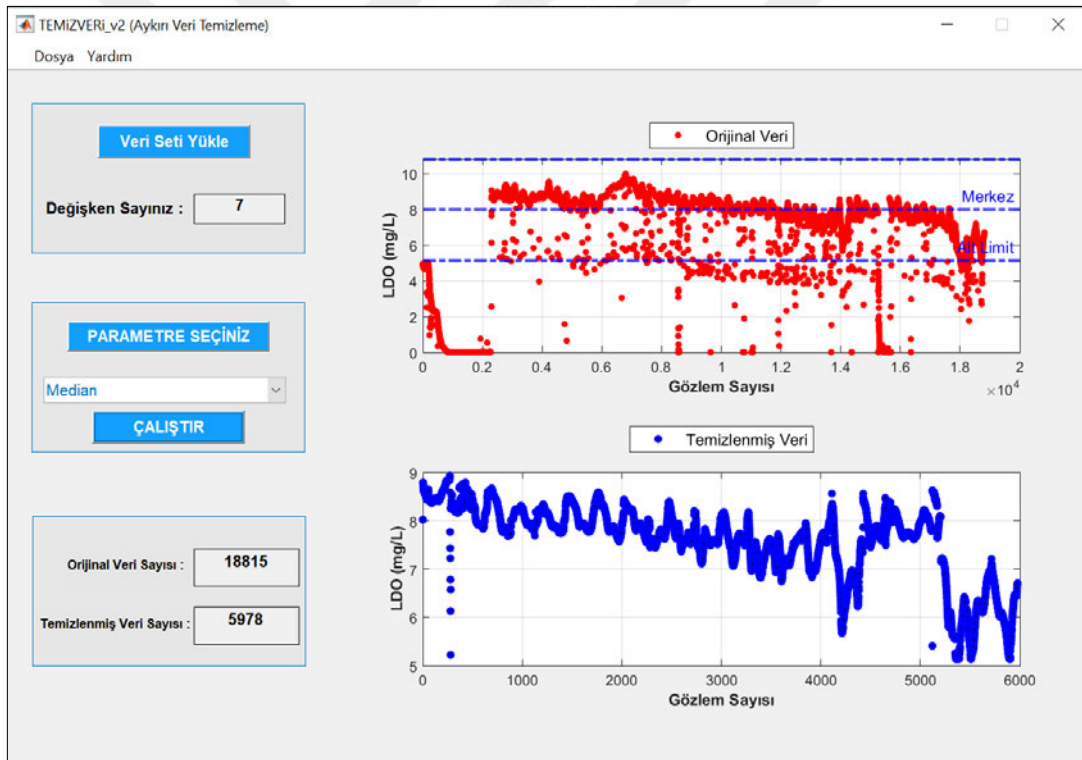
#### 5.1. Çözünmüş Oksijen Parametresi için Aykırı Verilerin Ayıklanması

Farklı yöntemler seçilerek elde edilen LDO sonuçları aşağıda Şekil 5.1, Şekil 5.2, Şekil 5.3, Şekil 5.4 ve Şekil 5.5 'te sunulmuştur. Ortalama yöntemi kullanılarak aykırı verilerin ayıklanması sonucunda, Şekil 5.1'den de görüleceği üzere orijinal veri sayısı 18815 iken ayıklanmış veri sayısı 17426'ya düşmüştür. Orijinal veri grafiği üzerinde çizilen alt sınır, üst sınır ve merkez çizgilerine göre alt ve üst sınırlar dışında kalan veriler aykırı veriler olarak değerlendirilmiş ve temizlenmiş veri grafiğinden de anlaşılacağı üzere söz konusu aykırı değerler veri setinden atılmıştır.

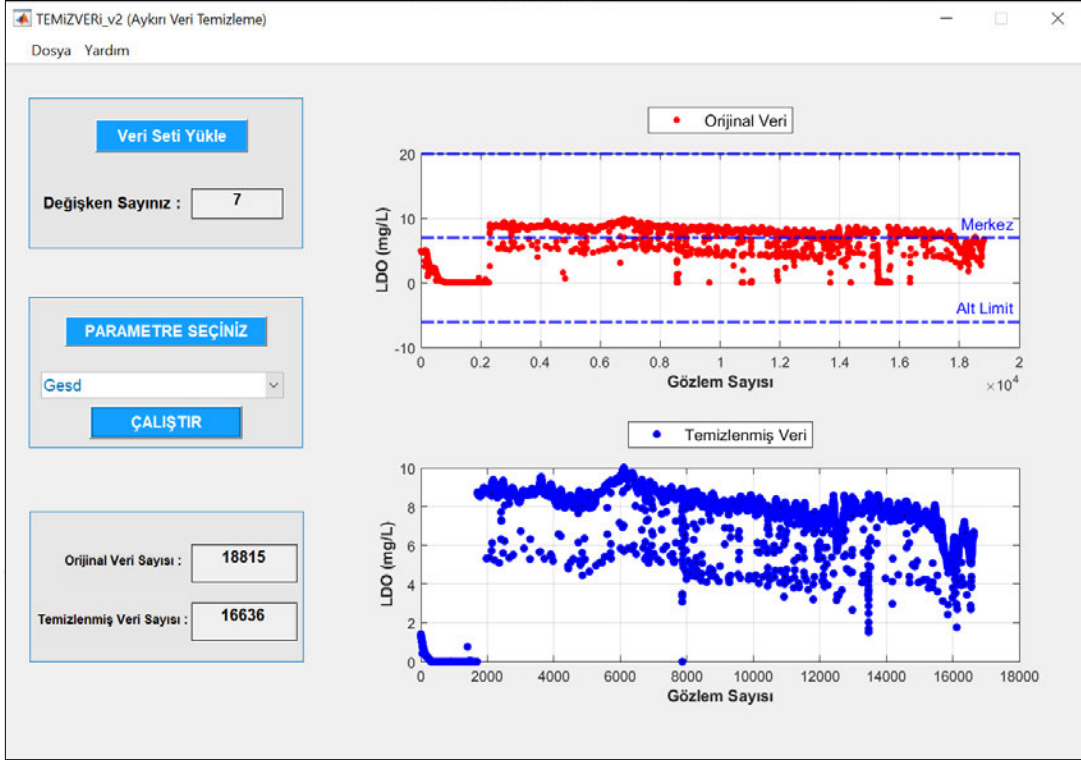


Şekil 5.1. Mean yöntemi ile LDO verilerinden aykırı verilerin ayıklanması sonuçları

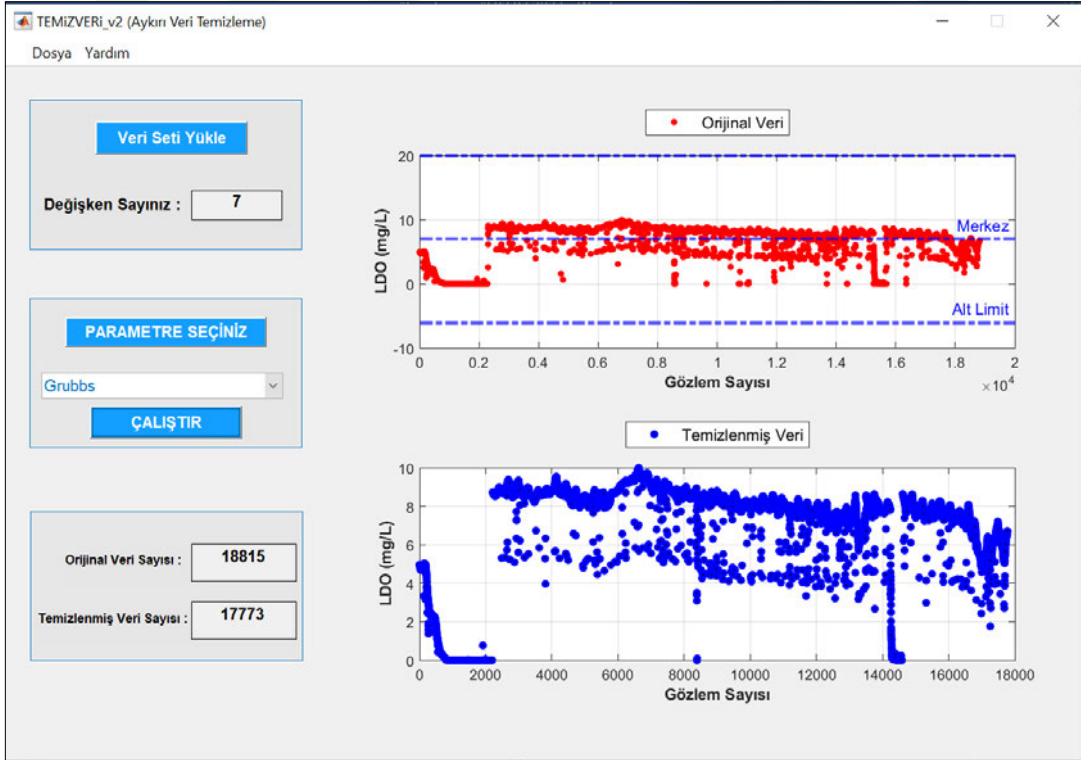
Bu aşamada dikkat edilmelidir ki bu çalışmada kullanılan yedi değişken için de aynı yöntem kullanıldığında aynı sayıda temizlenmiş veri kalacaktır. Çünkü, veri seti matris formatında alınmakta ve parametrelerin birindeki aykırı değere karşılık gelen diğer parametrelerin değerleri de atılmak zorundadır. Daha sonra yapılacak olan modelleme çalışmalarında aynı zamana karşılık gelen verilerin bütünlüğü oldukça önemlidir. Şöyle ki 20. dakikada LDO verisinde aykırı değer var ise yine 20. dakikadaki diğer parametrelerin verileri de aykırı olarak değerlendirilmelidir. Aksi taktirde modelleme yapılırken veri kayması ile karşı karşıya kalınır ve doğru bir model geliştirilmesi mümkün kılınmaz. Kullanıcı parametreleri ayrı ayrı analiz etmek isterse programa tanıttığı Excel dosyasında sadece ilgili parametrenin yer almasını sağlamalıdır. Bu durumda sadece ilgili parametre değerlendirilmiş olunur. Bunun yanında, eğer birden fazla parametre değerlendirilmek isteniyorsa yine Excel dosyasında değerlendirilecek parametreler ve onlara ait veriler yer almalıdır. Bu durumda, doğal olarak temizlenmiş veri sayısı farklı olabilecektir.



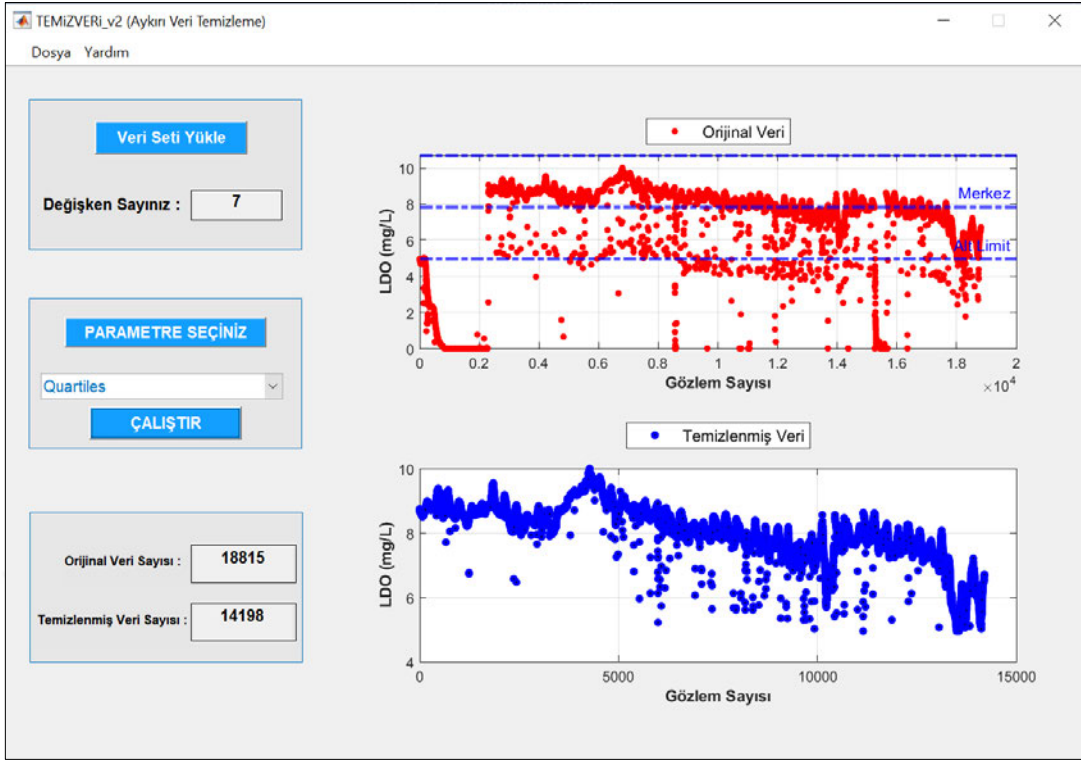
Şekil 5.2. Median yöntemi ile LDO verilerinden aykırı verilerin ayıklaması sonuçları



Şekil 5.3. Gesd yöntemi ile LDO verilerinden aykırı verilerin ayıklaması sonuçları



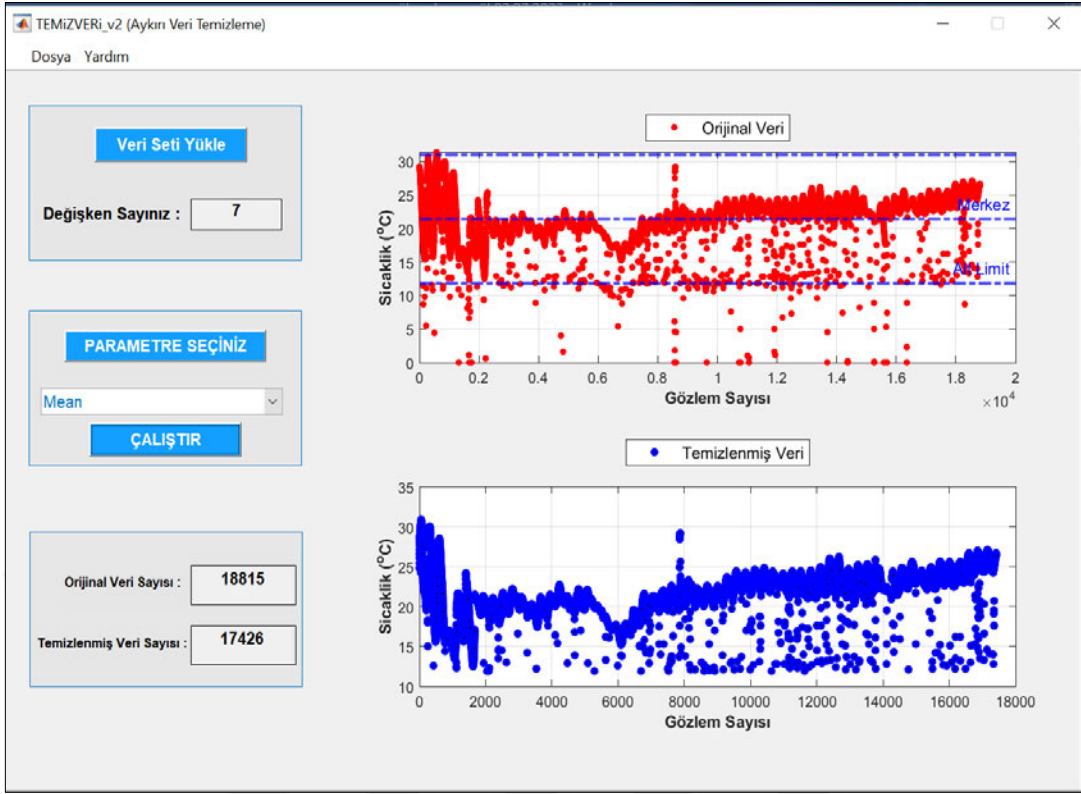
Şekil 5.4. Grubbs yöntemi ile LDO verilerinden aykırı verilerin ayıklaması sonuçları



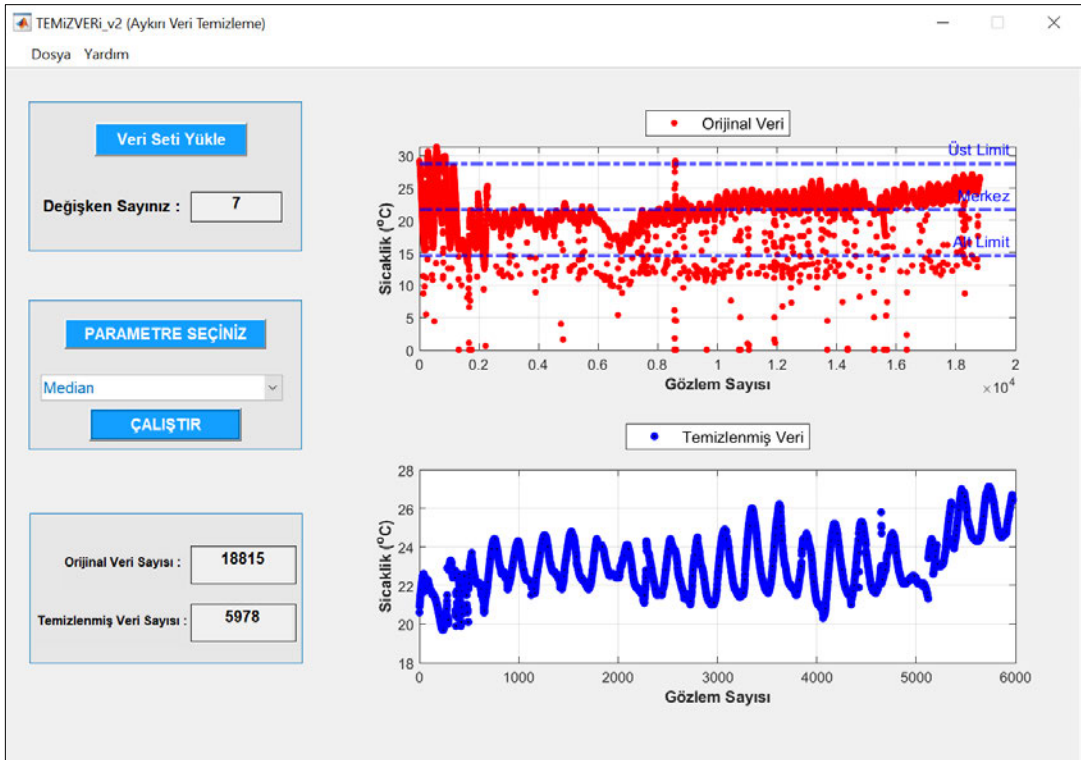
**Şekil 5.5.** Quartiles yöntemi ile LDO verilerinden aykırı verilerin ayıklaması sonuçları

## 5.2. Sıcaklık Parametresi için Aykırı Verilerin Ayıklanması

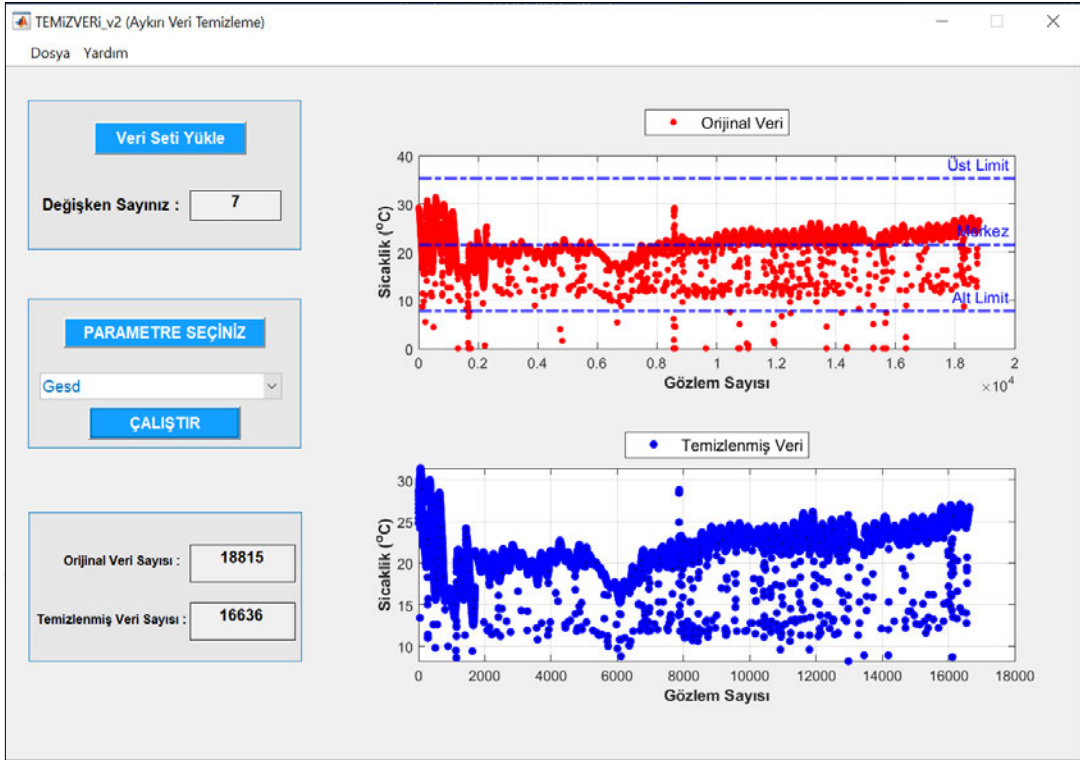
Sıcaklık verilerinin mean, median, Gesd, Grubbs ve Quartiles yöntemlerine göre elde edilen aykırı veri ayıklama sonuçları sırasıyla Şekil 5.6, Şekil 5.7, Şekil 5.8, Şekil 5.9 ve Şekil 5.10'da verilmiştir. Ortanca yöntemi kullanılarak aykırı verilerin ayıklanması sonucunda, Şekil 5.7'den de anlaşılacağı üzere orijinal veri sayısı 18815 iken ayıklanmış veri sayısı 5978 olmuştur. Orijinal veri grafiği üzerinde çizilen alt ve üst sınırlar dışında kalan veriler aykırı veriler olarak değerlendirilmiş ve temizlenmiş veri grafiğinden de görüldüğü gibi aykırı değerler veri setinden atılmıştır.



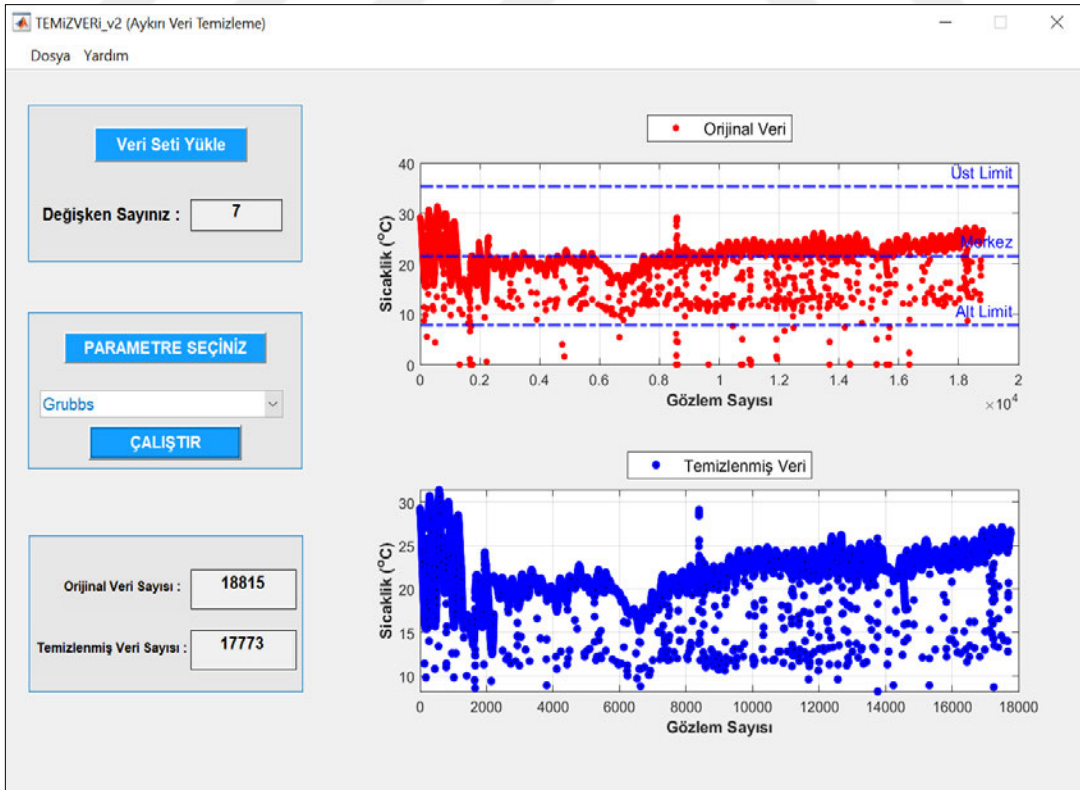
Şekil 5.6. Mean yöntemi ile Sıcaklık verilerinden aykırı verilerin ayıklaması sonuçları



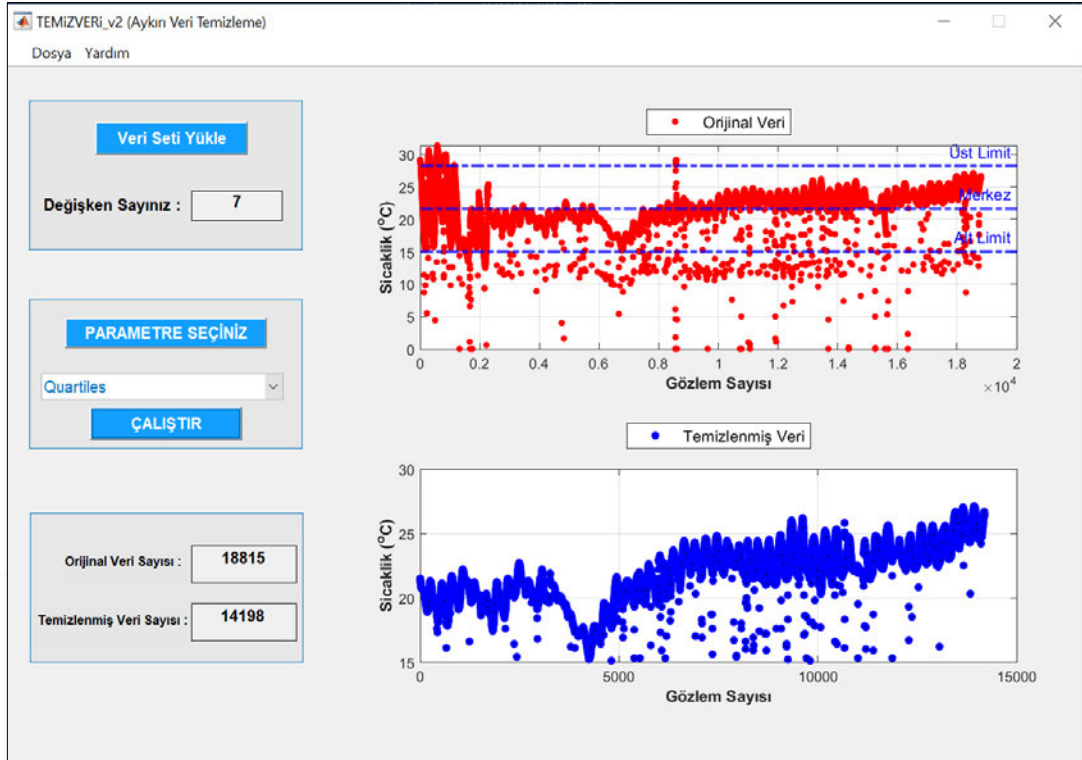
Şekil 5.7. Median yöntemi ile Sıcaklık verilerinden aykırı verilerin ayıklaması sonuçları



Şekil 5.8. Gesd yöntemi ile Sıcaklık verilerinden aykırı verilerin ayıklaması sonuçları



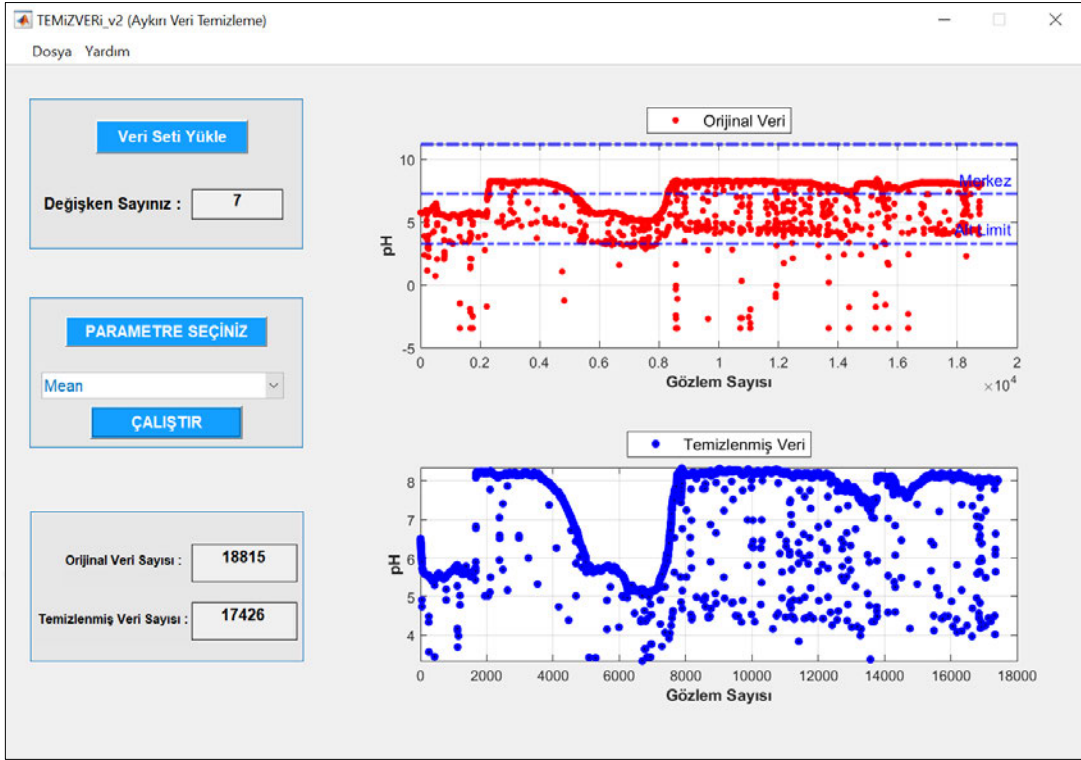
Şekil 5.9. Grubbs yöntemi ile Sıcaklık verilerinden aykırı verilerin ayıklaması sonuçları



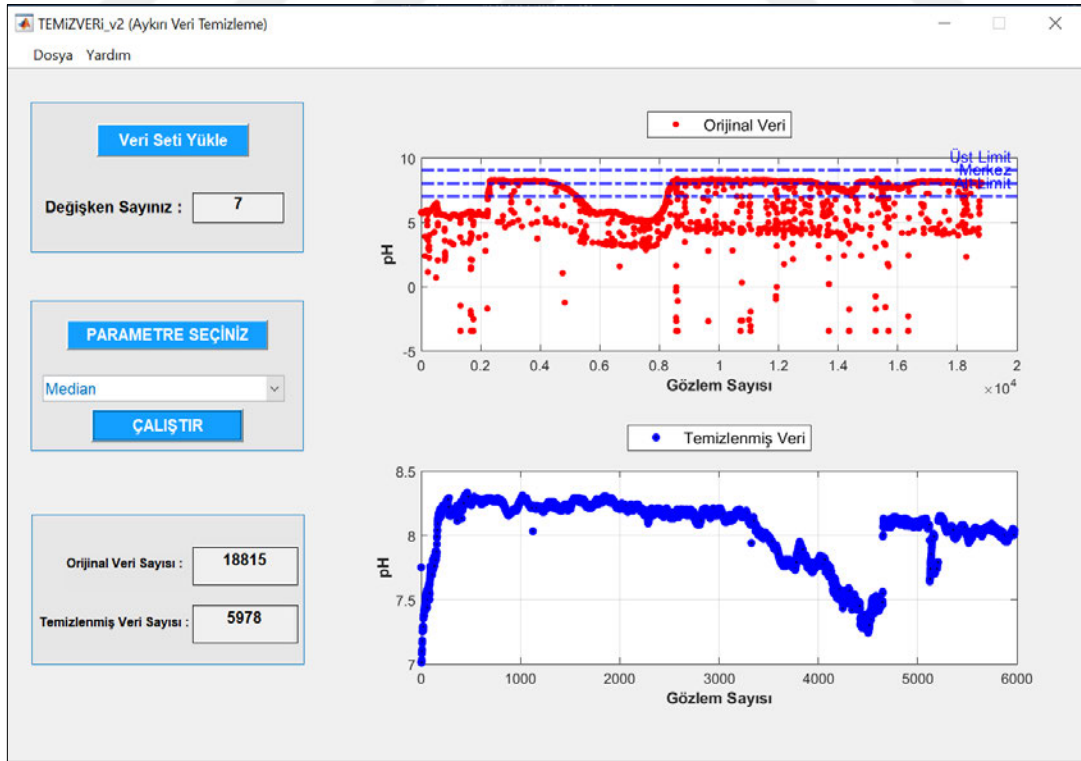
**Şekil 5.10.** Quartiles yöntemi ile Sıcaklık verilerinden aykırı verilerin ayıklanması sonuçları

### 5.3. pH Parametresi için Aykırı Verilerin Ayıklanması

MATLAB ortamından alınan pH verileri için aykırı veri ayıklama programı sonuçları Şekil 5.11, Şekil 5.12, Şekil 5.13, Şekil 5.14 ve Şekil 5.15,'te verilmiştir. *Gesd* yöntemi kullanılarak aykırı verilerin ayıklanması neticesinde, Şekil 5.13'te görüldüğü gibi orijinal veri sayısı 18815 iken aykırı veri ayıklama işleminden sonra geriye kalan veri sayısı 16636 olmuştur. Orijinal veri grafiği üzerinde görülen alt ve üst sınır çizgileri dışında kalan veriler aykırı veriler olarak değerlendirilerek bu aykırı veriler atılmış ve temizlenmiş veri grafiği oluşturulmuştur.

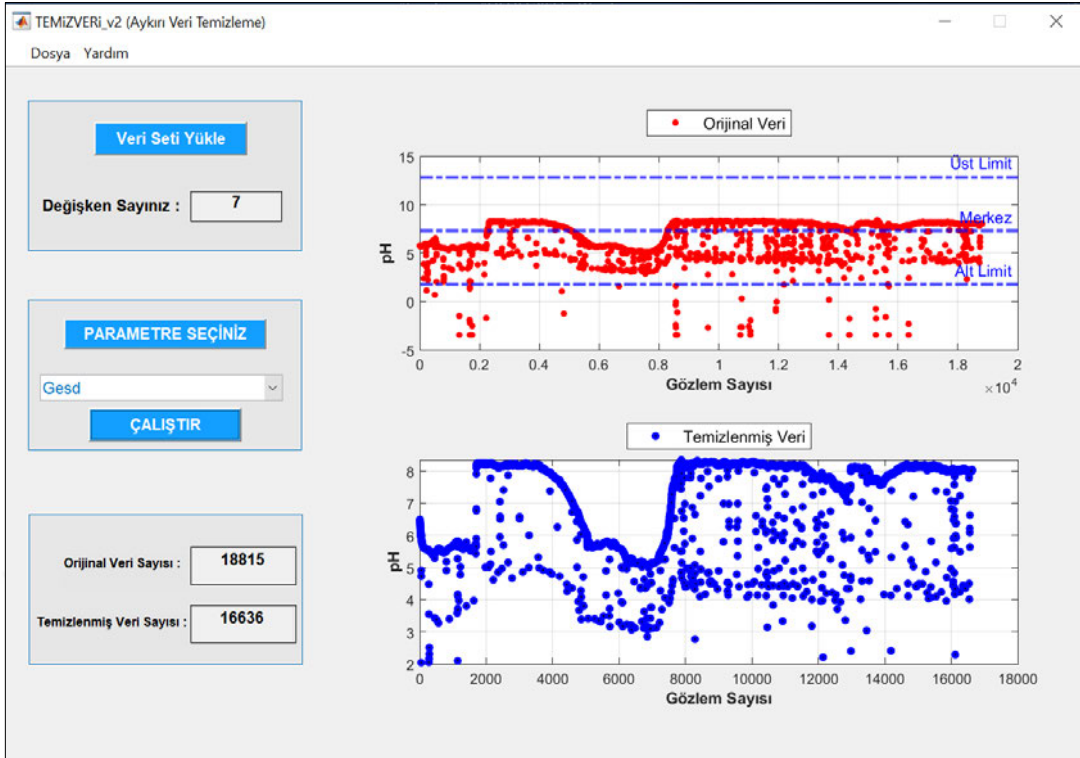


Şekil 5.11. Mean yöntemi ile pH verilerinden aykırı verilerin ayıklaması sonuçları

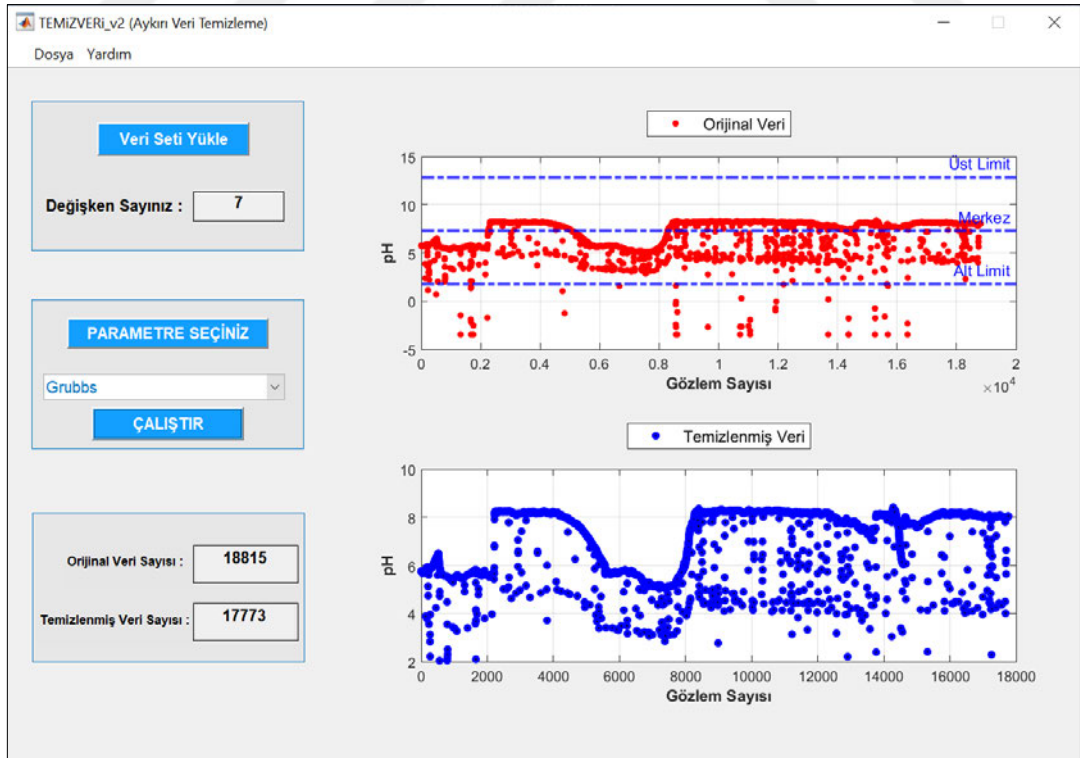


Şekil 5.12. Median yöntemi ile pH verilerinden aykırı verilerin ayıklaması sonuçları

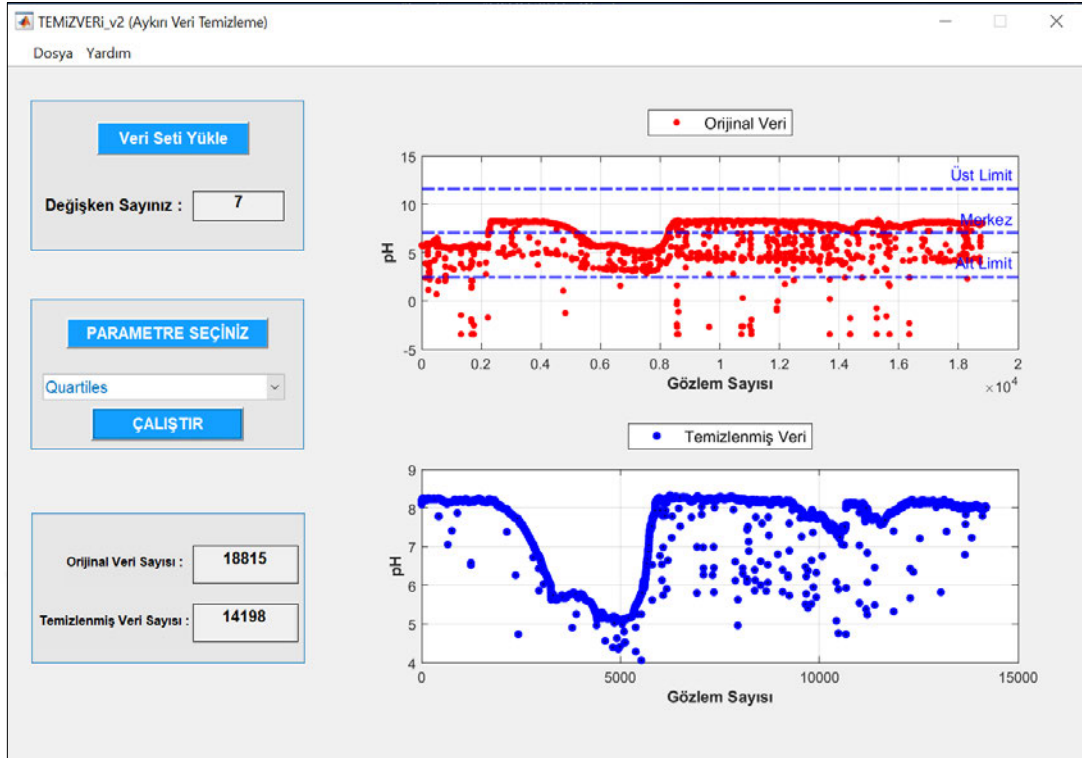




Şekil 5.13. Gesd yöntemi ile pH verilerinden aykırı verilerin ayıklaması sonuçları



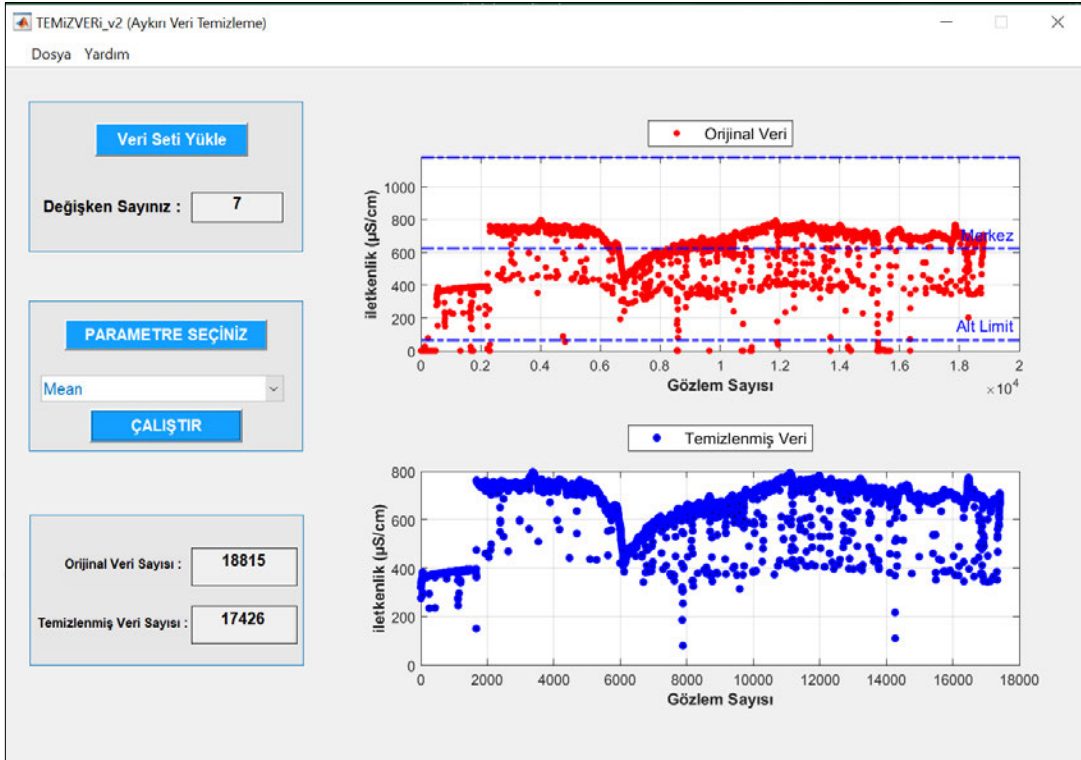
Şekil 5.14. Grubbs yöntemi ile pH verilerinden aykırı verilerin ayıklaması sonuçları



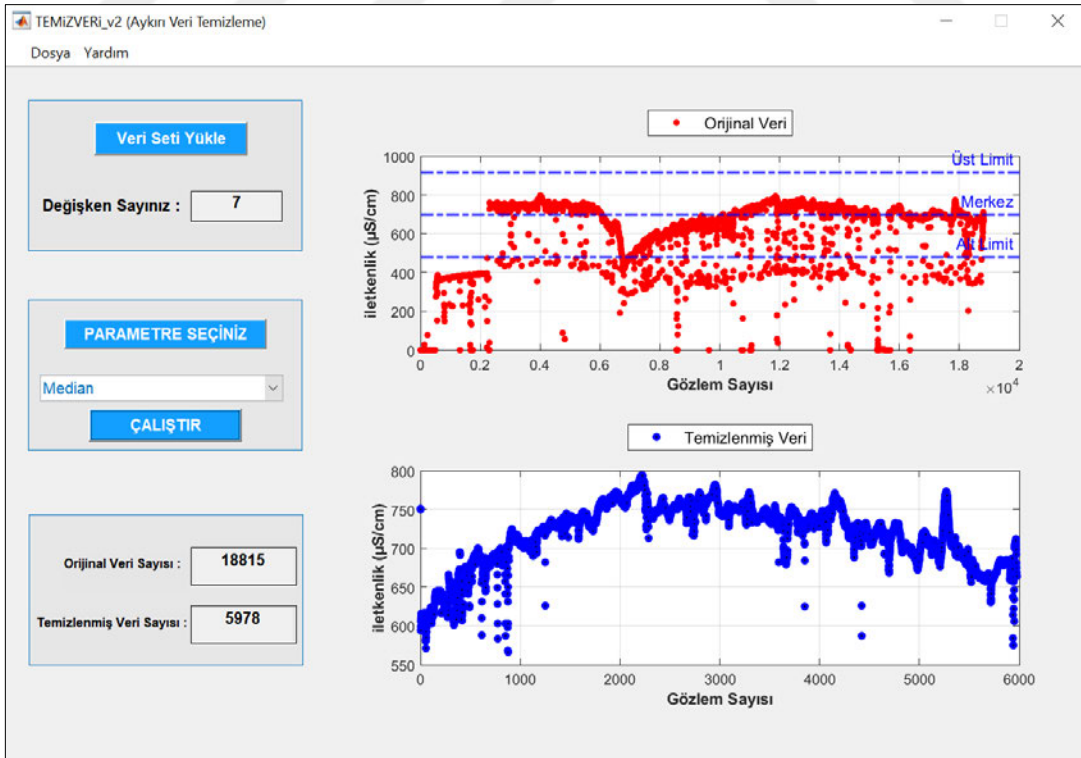
**Şekil 5.15.** Quartiles yöntemi ile pH verilerinden aykırı verilerin ayıklaması sonuçları

#### 5.4. İletkenlik Parametresi için Aykırı Verilerin Ayıklanması

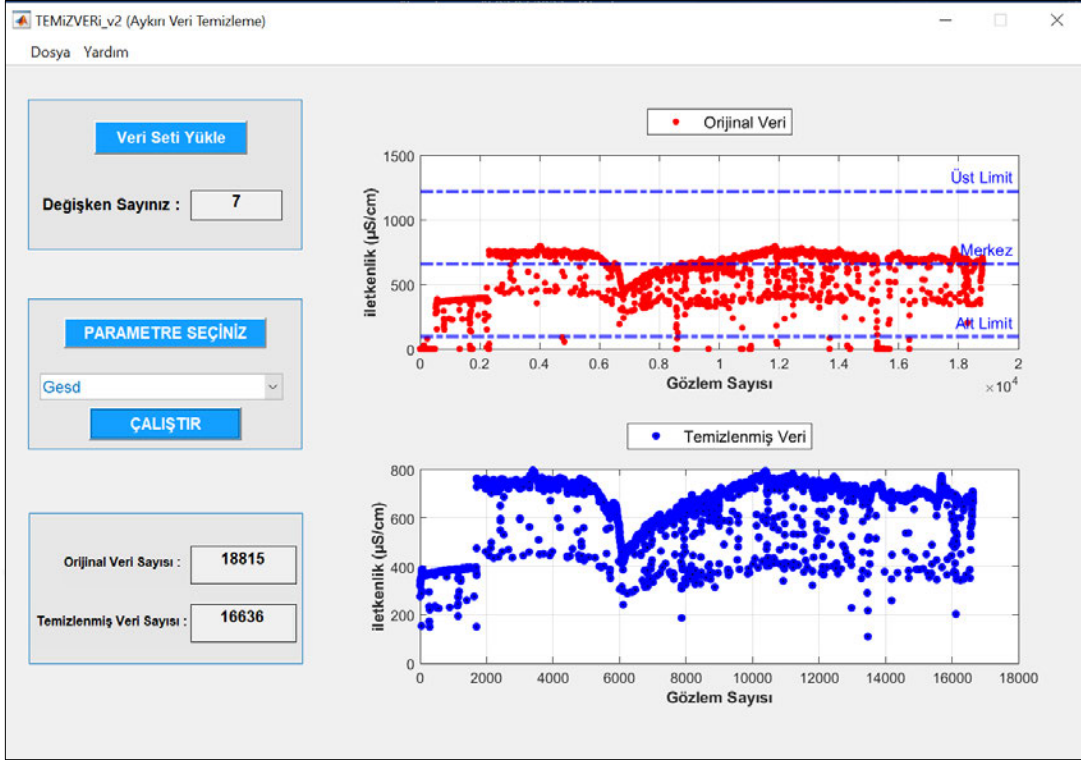
İletkenlik verilerinin mean, median, Gesd, Grubbs ve Quartiles yöntemlerine göre elde edilen aykırı veri ayıklama sonuçları sırasıyla Şekil 5.16, Şekil 5.17, Şekil 5.18, Şekil 5.19 ve Şekil 5.20’de görüldüğü gibidir. Grubbs yöntemi kullanılarak aykırı verilerin ayıklanması ile elde edilen sonuçlar Şekil 5.19’da yer aldığı gibidir. 18815 olan orijinal veri sayısı, aykırı veriler ayıklandıktan sonra 17773’e düşmüştür. Yine diğer yöntemlerde olduğu gibi, orijinal veri grafiğinde yer alan alt ve üst sınır çizgileri dışındaki veriler aykırı veriler olarak değerlendirilmiştir.



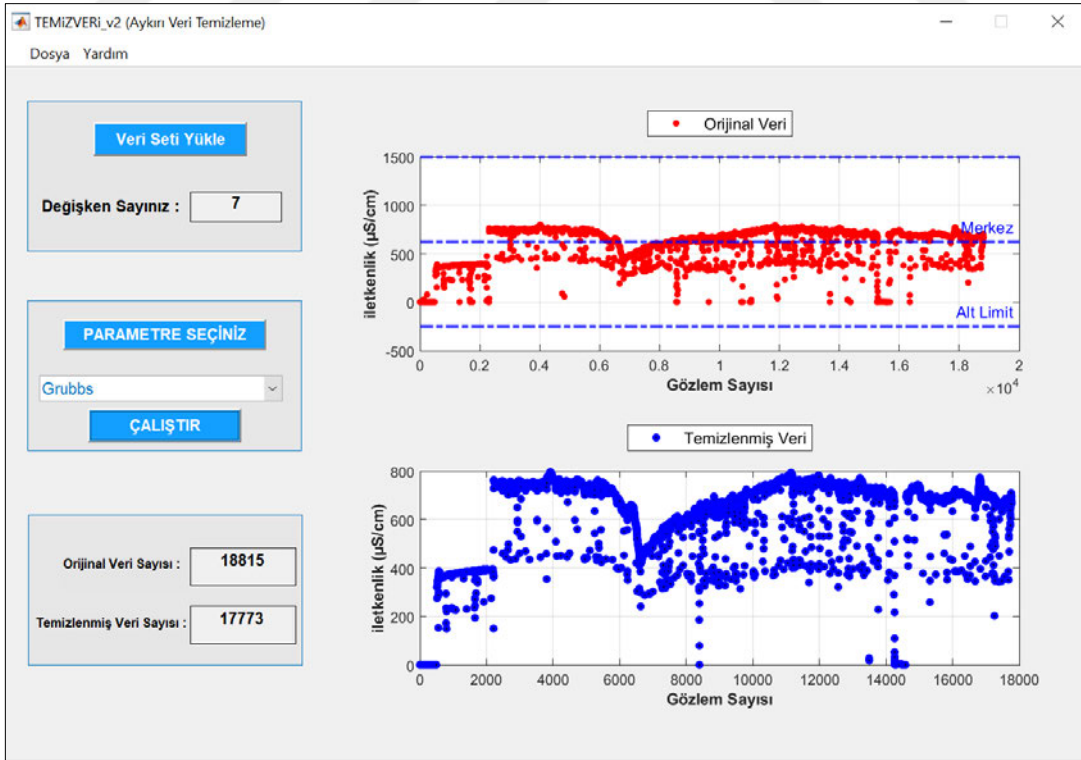
Şekil 5.16. Mean yöntemi ile iletkenlik verilerinden aykırı verilerin ayıklaması sonuçları



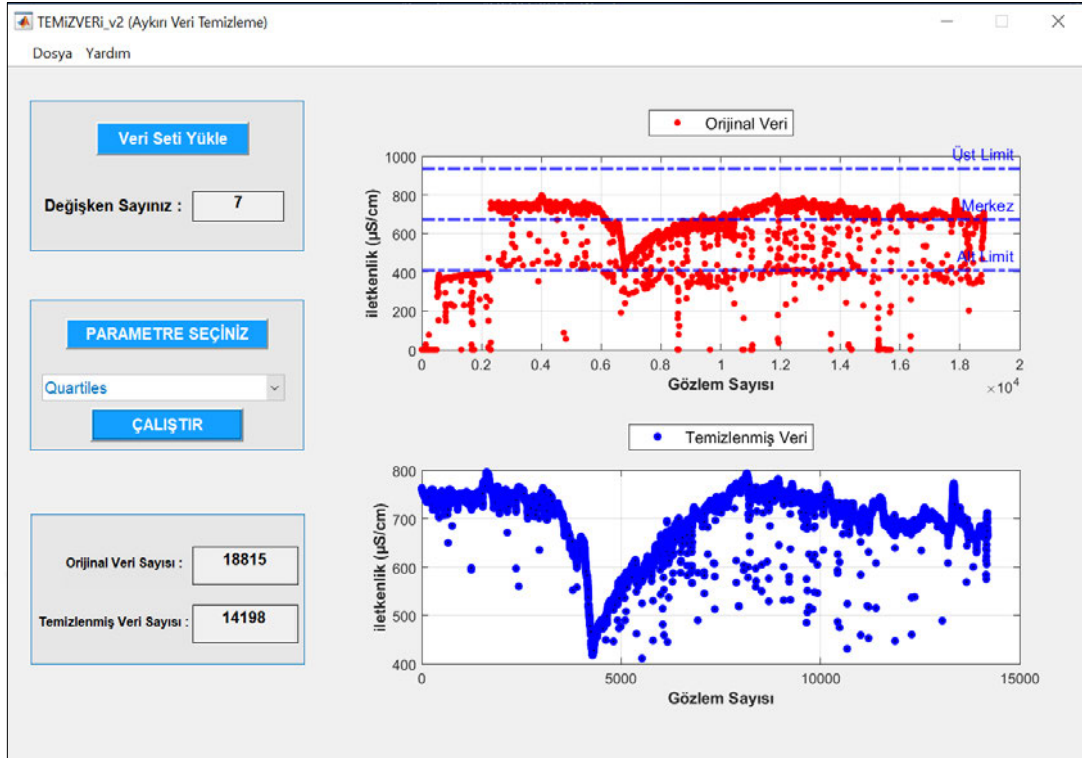
Şekil 5.17. Median yöntemi ile iletkenlik verilerinden aykırı verilerin ayıklaması sonuçları



Şekil 5.18. Gesd yöntemi ile iletkenlik verilerinden aykırı verilerin ayıklaması sonuçları



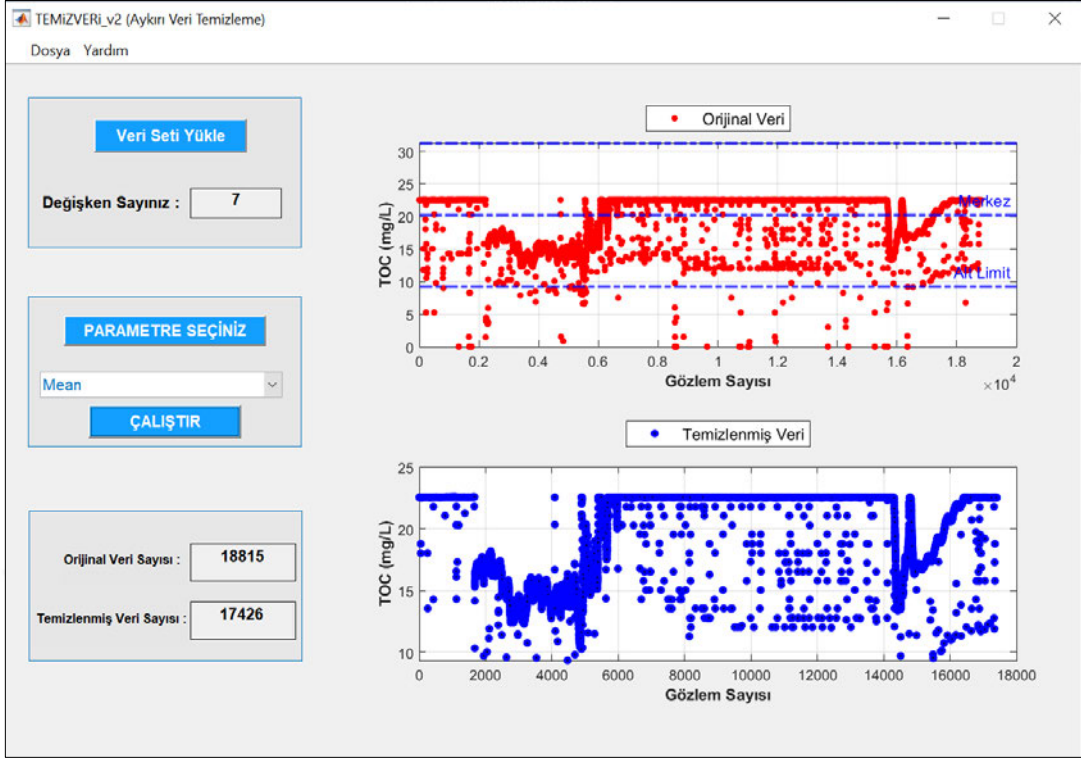
Şekil 5.19. Grubbs yöntemi ile iletkenlik verilerinden aykırı verilerin ayıklaması sonuçları



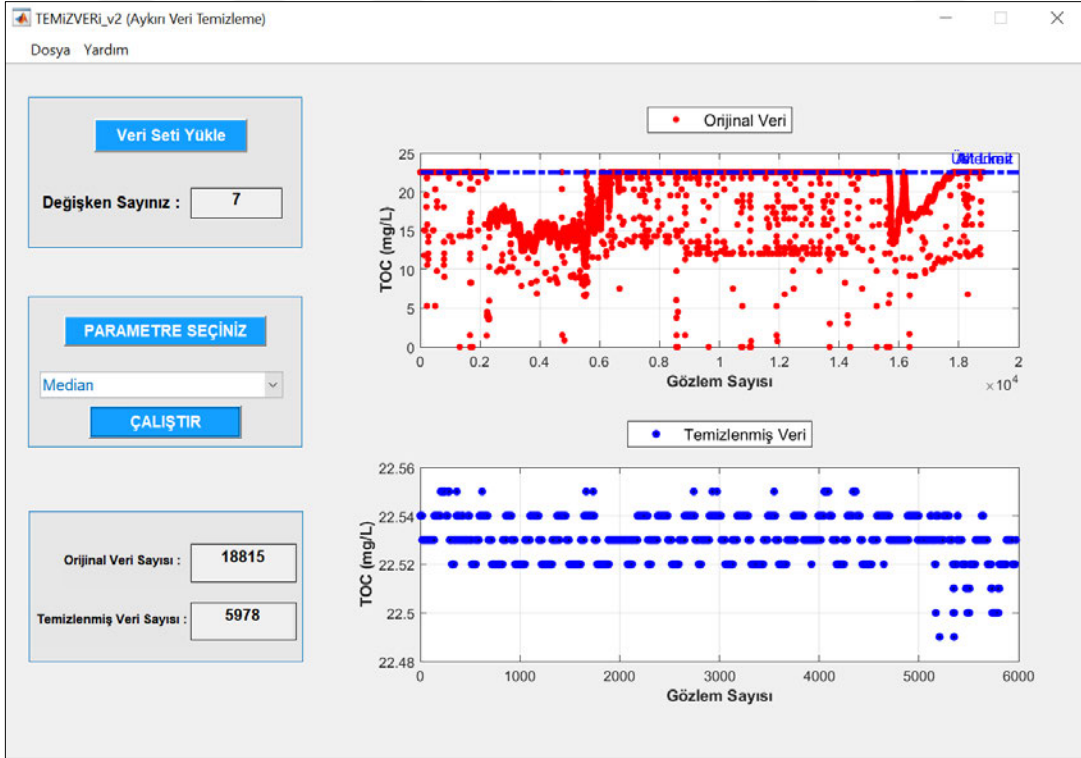
**Şekil 5.20.** Quartiles yöntemi ile iletkenlik verilerinden aykırı verilerin ayıklanması sonuçları

### 5.5. Toplam Organik Karbon Parametresi için Aykırı Verilerin Ayıklanması

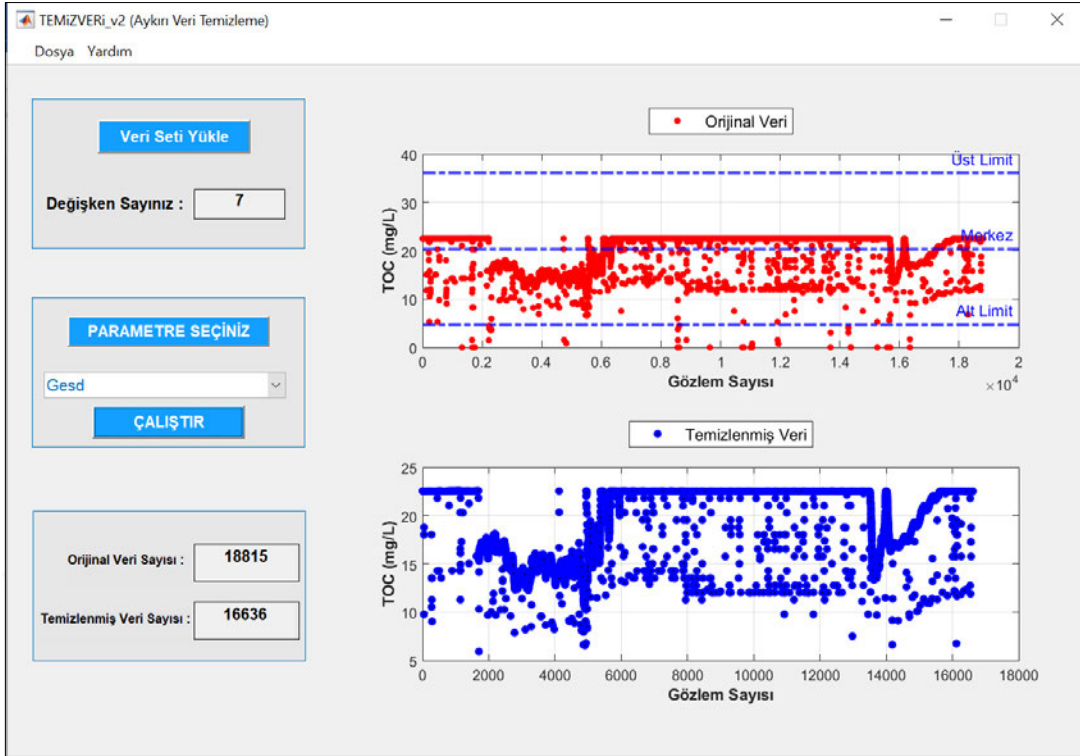
Toplam Organik Karbon (TOC) verileri için aykırı veri ayıklama programı sonuçları Şekil 5.21-Şekil 5.25 arasında gösterilmiştir. Çeyrekler Açıklığı (*Quartiles*) yöntemi kullanılarak aykırı verilerin ayıklanması ile elde edilen sonuçlar Şekil 5.25'te gösterildiği gibidir. Aykırı veriler ayıklandıktan sonra 18815 olan orijinal veri sayısı, 14198 olmuştur. Burada da orijinal veri grafiğinde çizilmiş olan alt ve üst sınırlar dışında kalan veriler aykırı olarak tanımlanmış ve temizlenmiştir.



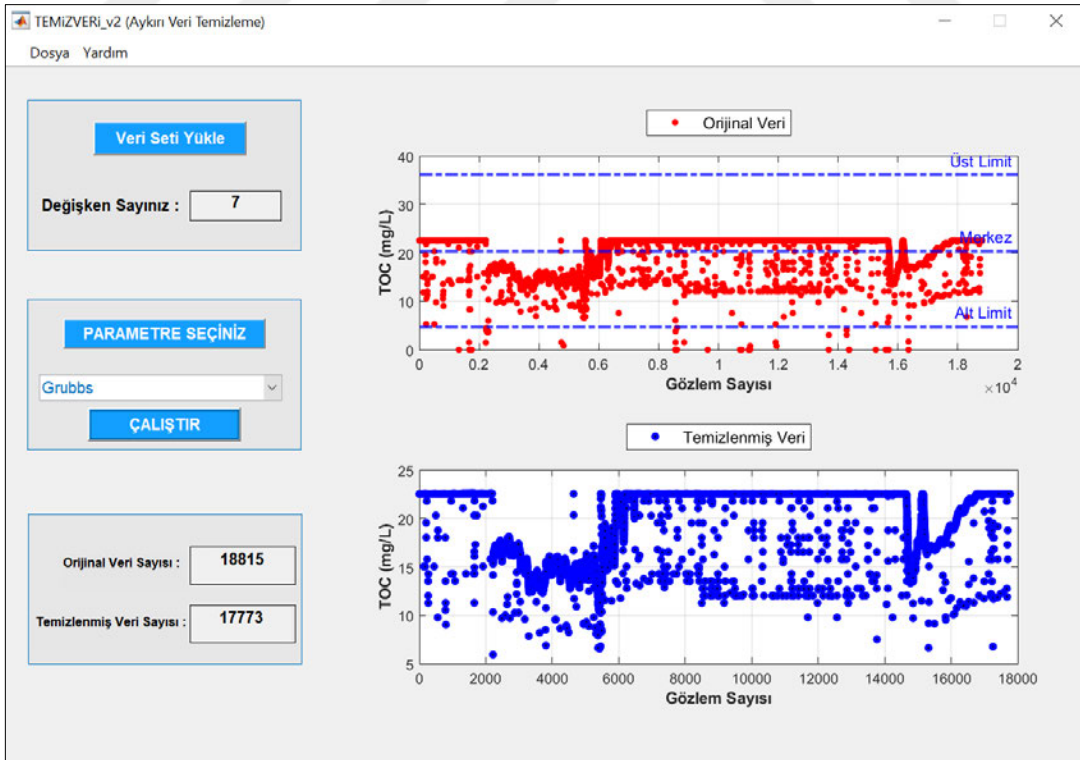
Şekil 5.21. Mean yöntemi ile TOC verilerinden aykırı verilerin ayıklaması sonuçları



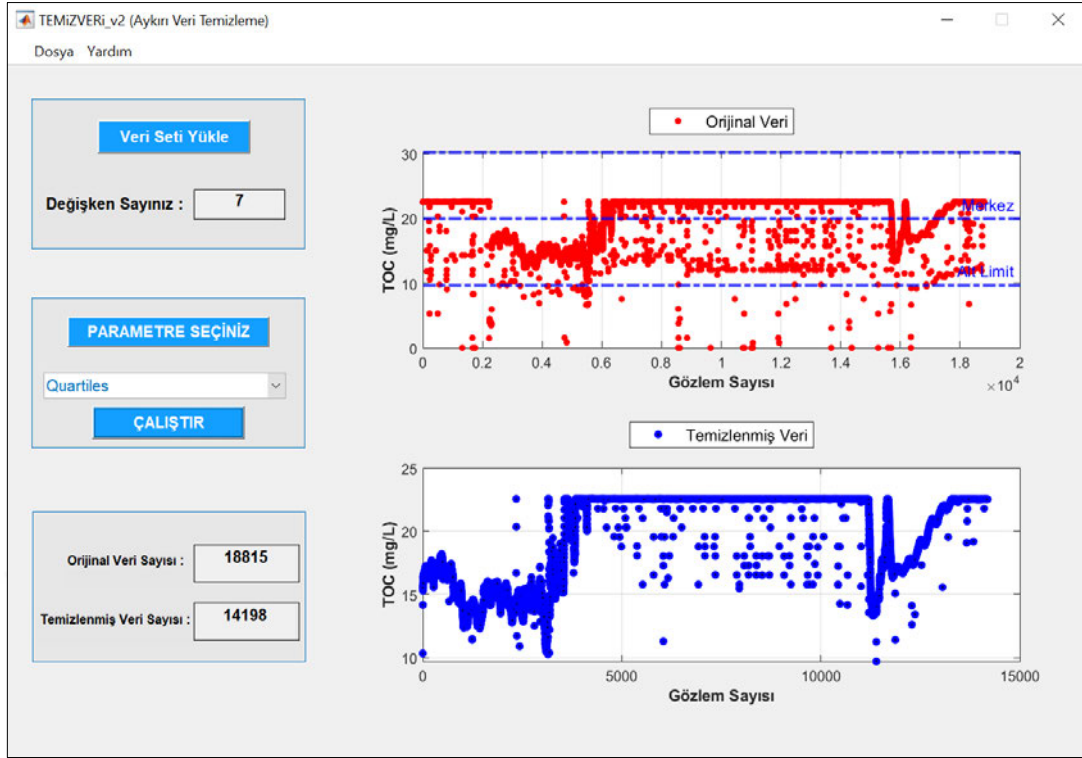
Şekil 5.22. Median yöntemi ile TOC verilerinden aykırı verilerin ayıklaması sonuçları



Şekil 5.23. Gest yöntemi ile TOC verilerinden aykırı verilerin ayıklaması sonuçları



Şekil 5.24. Grubbs yöntemi ile TOC verilerinden aykırı verilerin ayıklaması sonuçları

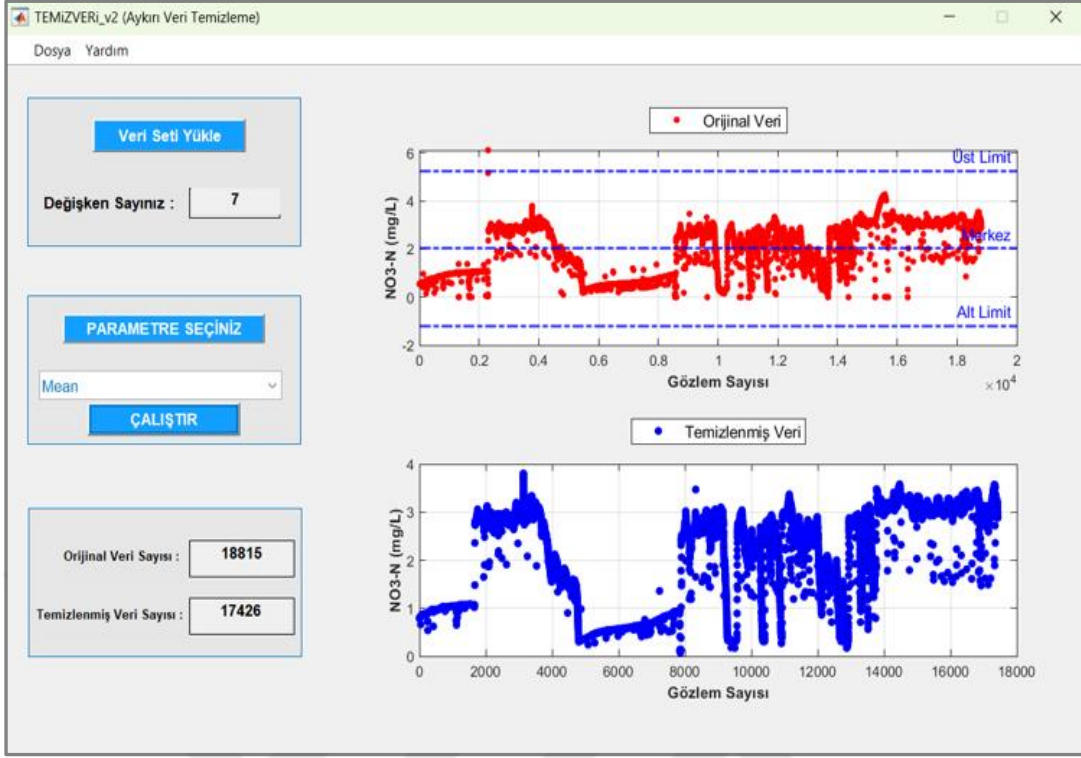


**Şekil 5.25.** Quartiles yöntemi ile TOC verilerinden aykırı verilerin ayıklanması sonuçları

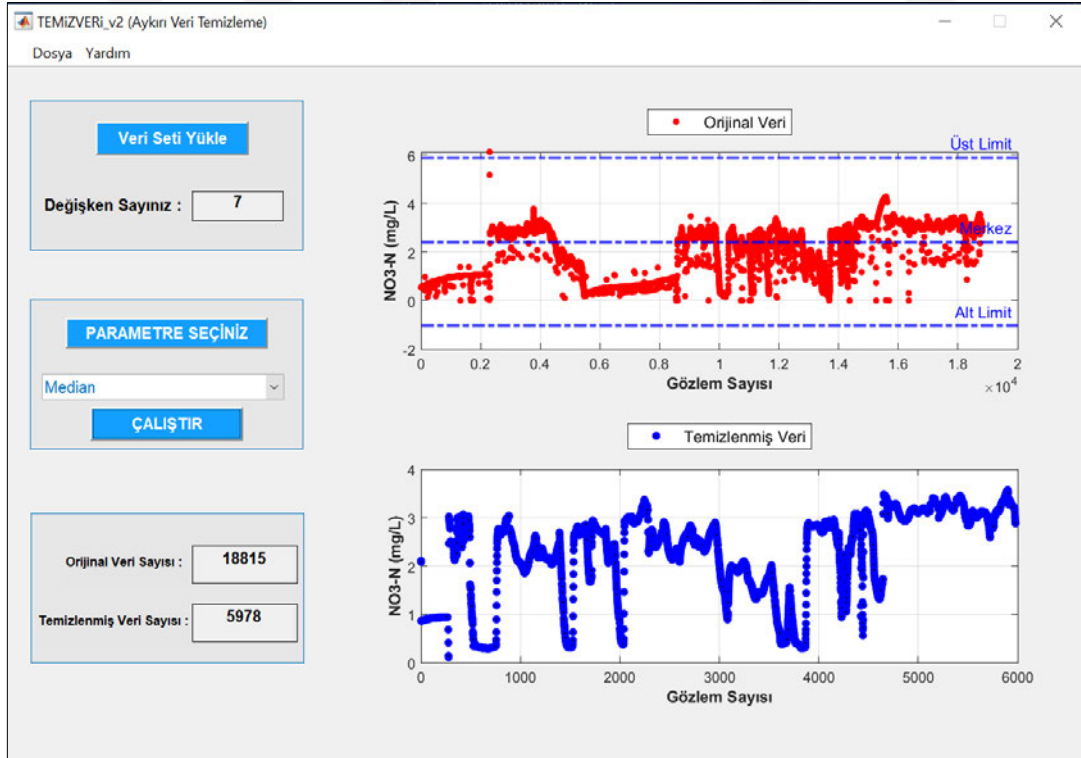
## 5.6. Nitrat Azotu ve Amonyum Azotu Parametreleri için Aykırı Verilerin Ayıklanması

Her iki azot bileşeni için orijinal verilere beş farklı yöntem (ortalama-*mean*, ortanca-*median*, Gesd, Grubbs ve çeyrekler açıklığı-*quartiles*) uygulandıktan sonra elde edilen sonuçlar Şekil 5.26-Şekil 5.30 arasında sunulmuştur. Azot verilerine ait sonuçlar; nitrat azotu (NO<sub>3</sub>-N) için Şekil 5.26, Şekil 5.27, Şekil 5.28, Şekil 5.29 ve Şekil 5.30'da, amonyum azotu (NH<sub>4</sub>-N) için Şekil 5.31, Şekil 5.32, Şekil 5.33, Şekil 5.34 ve Şekil 5.35'te gösterilmiştir.

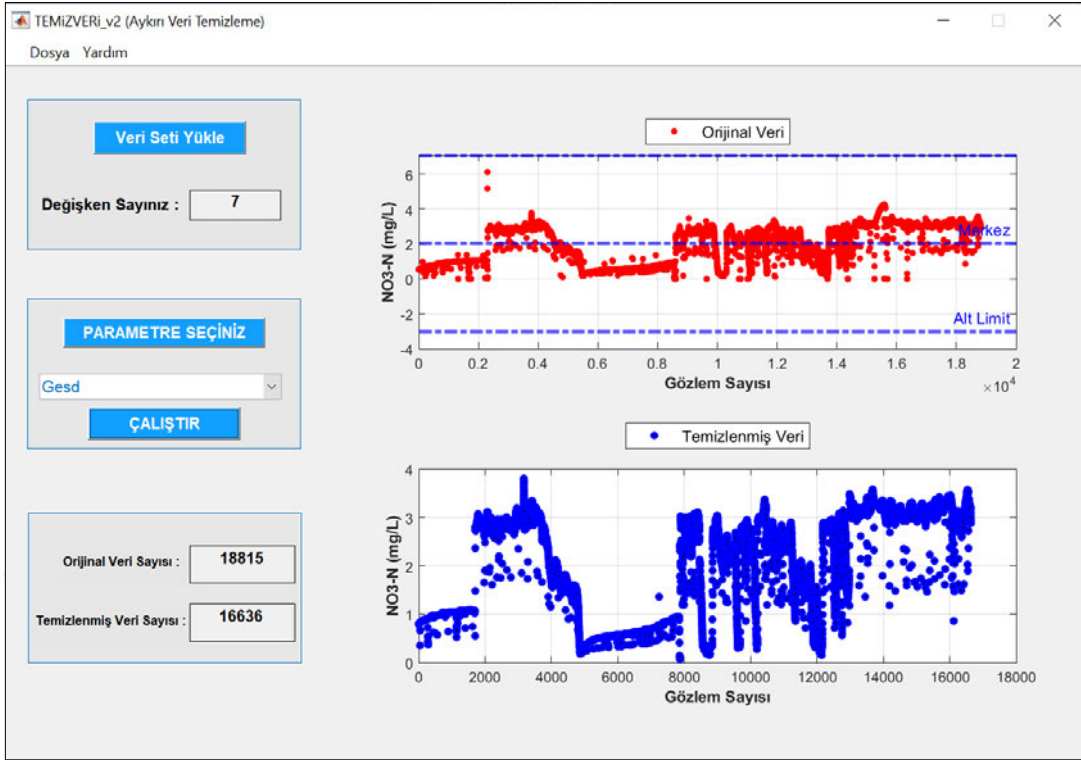




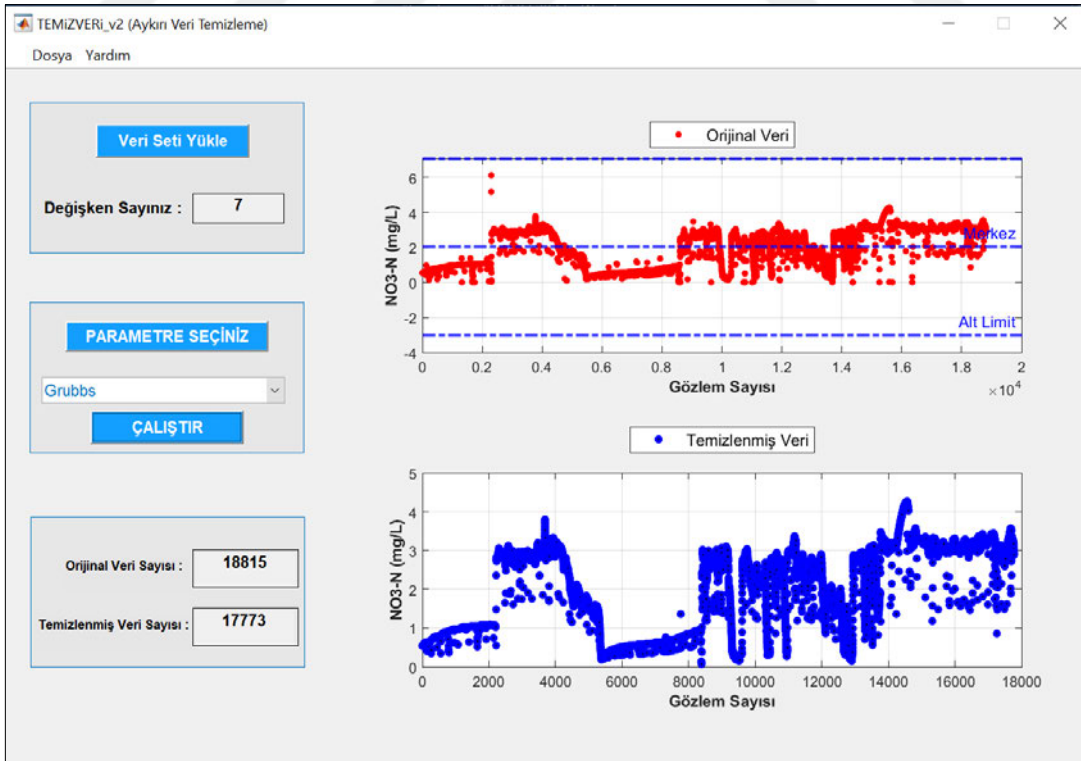
Şekil 5.26. Mean yöntemi ile NO<sub>3</sub>-N verilerinden aykırı verilerin ayıklaması sonuçları



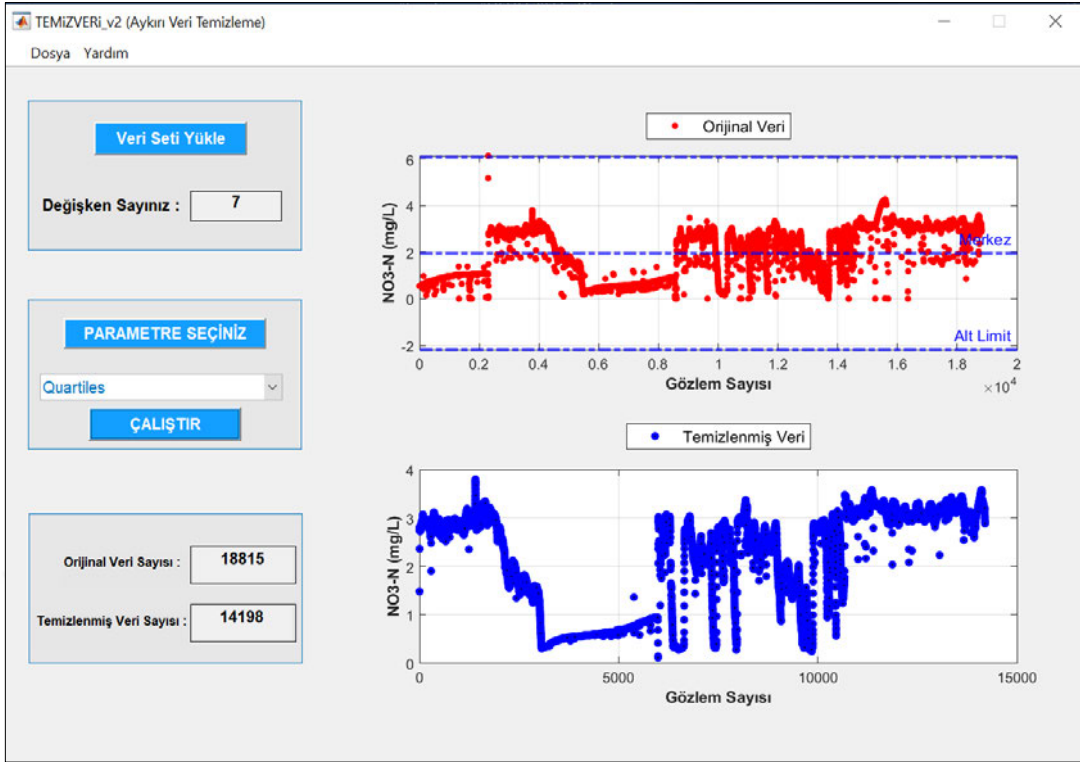
Şekil 5.27. Median yöntemi ile NO<sub>3</sub>-N verilerinden aykırı verilerin ayıklaması sonuçları



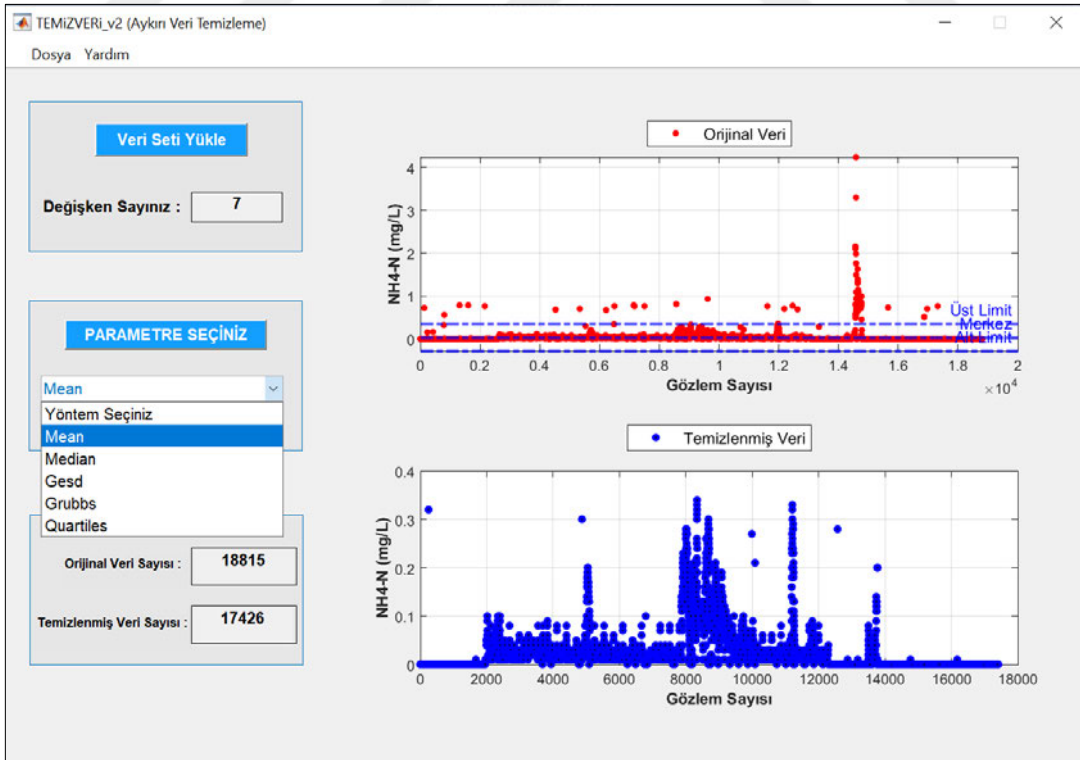
Şekil 5.28. Gest yöntemi ile NO<sub>3</sub>-N verilerinden aykırı verilerin ayıklaması sonuçları



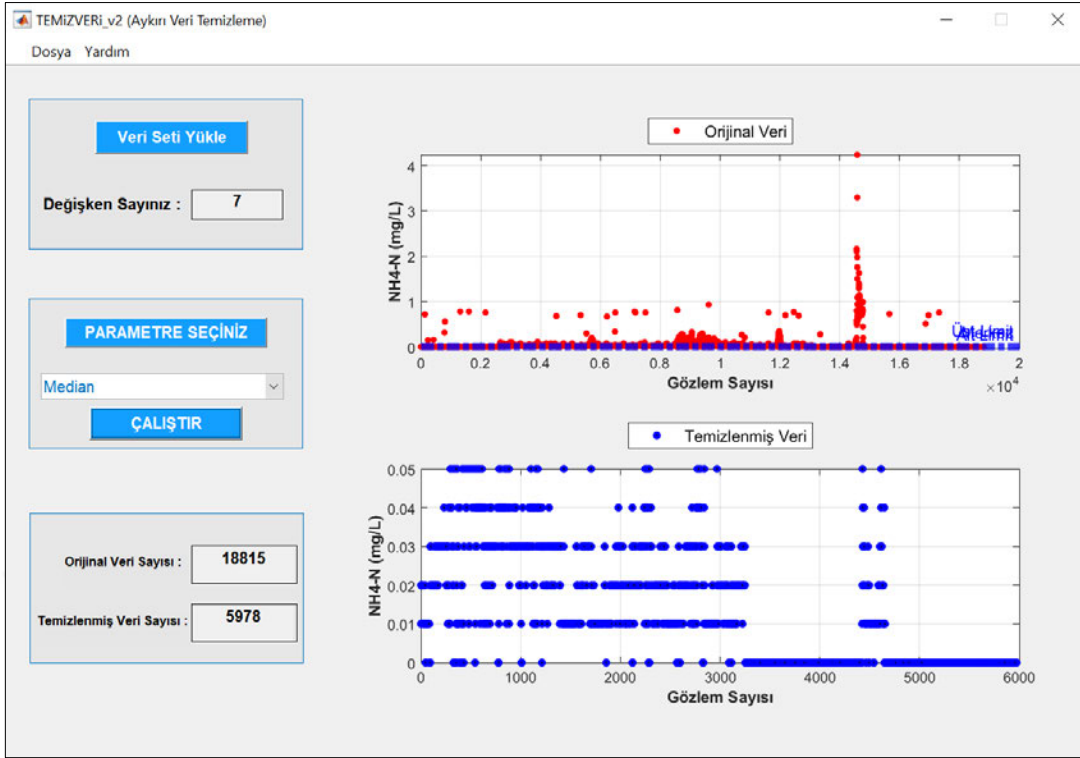
Şekil 5.29. Grubbs yöntemi ile NO<sub>3</sub>-N verilerinden aykırı verilerin ayıklaması sonuçları



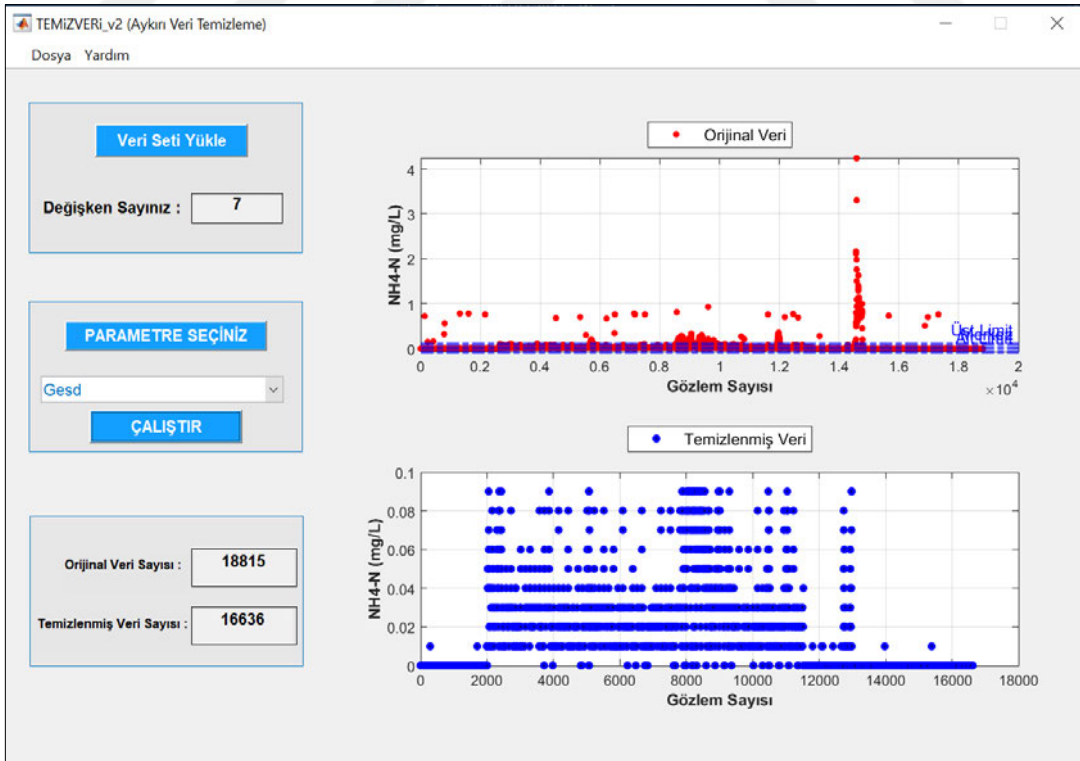
Şekil 5.30. Quartiles yöntemi ile  $\text{NO}_3\text{-N}$  verilerinden aykırı verilerin ayıklaması sonuçları



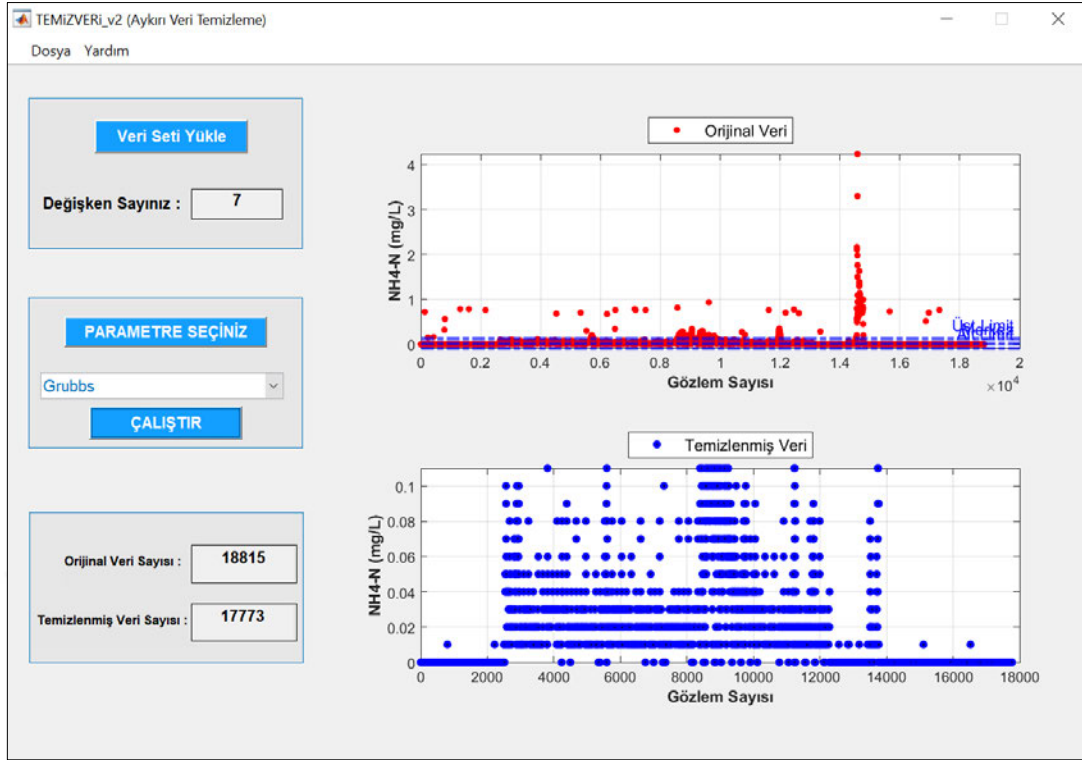
Şekil 5.31. Mean yöntemi ile  $\text{NH}_4\text{-N}$  verilerinden aykırı verilerin ayıklaması sonuçları



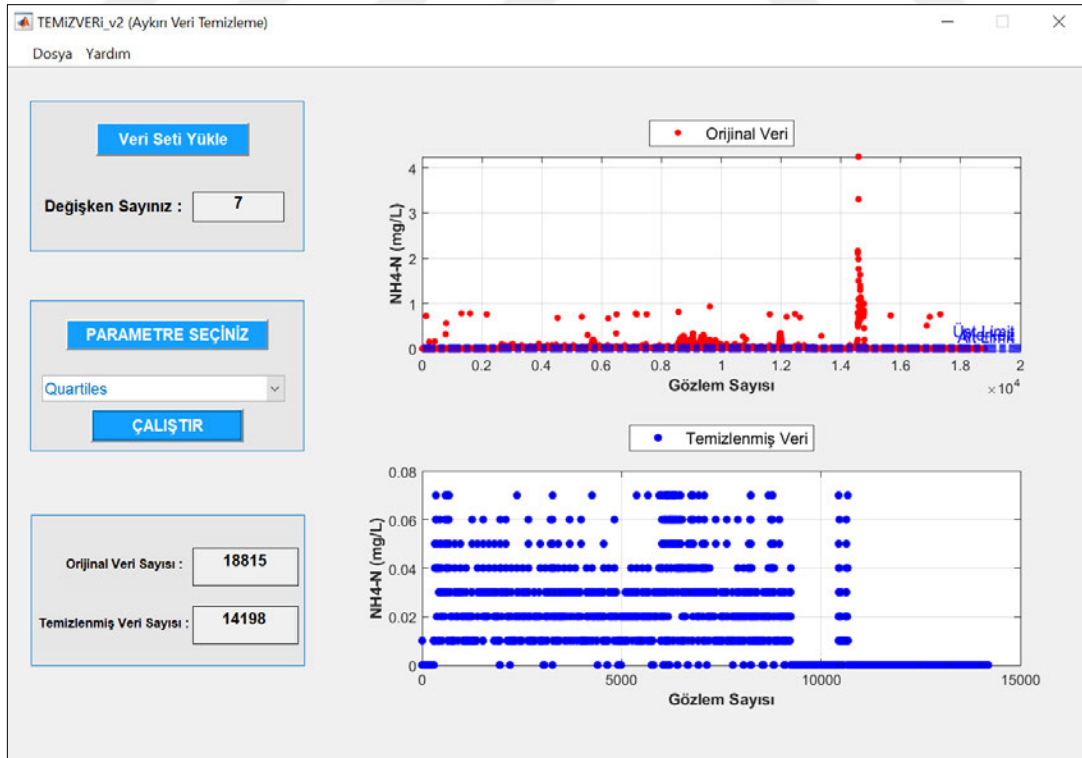
Şekil 5.32. Median yöntemi ile  $\text{NH}_4\text{-N}$  verilerinden aykırı verilerin ayıklaması sonuçları



Şekil 5.33. Gesd yöntemi ile  $\text{NH}_4\text{-N}$  verilerinden aykırı verilerin ayıklaması sonuçları



Şekil 5.34. Grubbs yöntemi ile  $\text{NH}_4\text{-N}$  verilerinden aykırı verilerin ayıklaması sonuçları

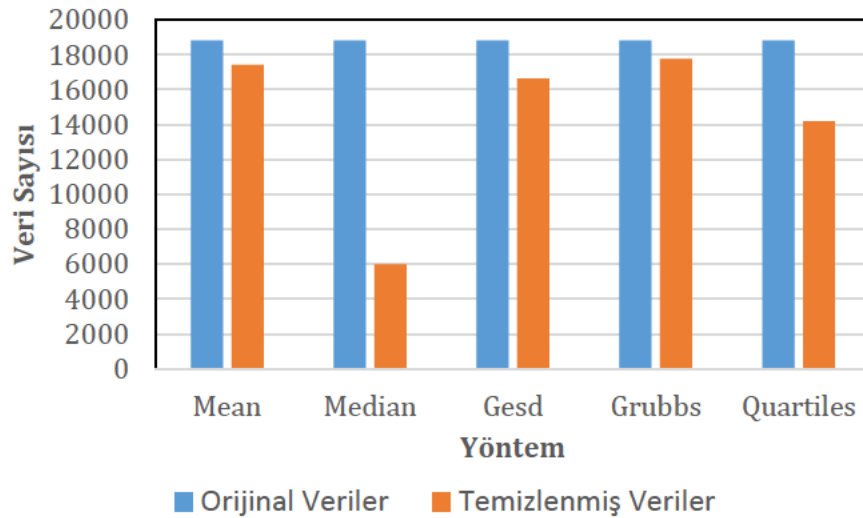


Şekil 5.35. Quartiles yöntemi ile  $\text{NH}_4\text{-N}$  verilerinden aykırı verilerin ayıklaması sonuçları

Yukarıda verilen tüm şekiller değerlendirildiğinde orijinal veri sayısı ve aykırı veriler temizlendikten sonra geriye kalan veri sayısı her bir aykırı veri ayıklama yöntemine göre Tablo 5.1’de özetlenmiştir. Tablodan da anlaşılacağı üzere en az aykırı veri *Grubbs* ile belirlenirken en fazla sayıda aykırı veri ise ortanca (*median*) ile belirlenmiştir. Aykırı veri ayıklama yöntemlerine göre orijinal ve temizlenmiş veri sayıları çubuk grafiği olarak Şekil 5.36’da verilmiştir. Burada kullanılan aykırı veri temizleme yöntemlerinin başarısı, bu tez çalışmasının bir sonraki aşaması olan modelleme çalışmaları neticesinde belirlenebilir. Her bir yöntem sonucu elde edilen temiz veri seti ile modelleme çalışmaları yapıldıktan sonra en başarılı sonuçlar üreten model için hangi aykırı veri temizleme yöntemi sonucu elde edilen veriler kullanıldıysa o yöntemin uygunluğundan bahsedilebilir.

**Tablo 5.1.** LDO için orijinal ve Temizlenmiş Veri Sayıları

Yöntem	Orijinal Veri Sayısı	Temizlenmiş Veri Sayısı
Mean	18815	17426
Median	18815	5978
Gesd	18815	16636
Grubbs	18815	17773
Quartiles	18815	14198



**Şekil 5.36.** Aykırı veri ayıklama yöntemlerine göre orijinal ve temizlenmiş veri sayıları

Şekil 5.36'da aykırı veri ayıklama yöntemlerine göre orijinal ve aykırı veri temizliği sonrası geriye kalan veri sayıları verilmiş olup; Mean yöntemi 1389 veriyi, Median yöntemi 12837 veriyi, Gesd yöntemi 2179 veriyi, Grubbs yöntemi 1042 veriyi ve Quartiles yöntemi ise 4617 veriyi ayıklayıp atmıştır. Bundan dolayı en fazla veri Median yöntemi ile temizlenmiş ve geriye 5978 veri kalmıştır. Grubbs yöntemi ile en az sayıda veri ayıklanmış olup, ayıklama sonrası kalan veri sayısı 17773'tür. Çalışmada kullanılan veri seti için, en etkili yöntemin hangisi olduğu modelleme çalışmaları yapıldıktan sonra değerlendirilebilir.



## SONUÇ VE ÖNERİLER

Tez çalışması kapsamında Yeşilirmak nehri üzerinde kurulu olan istasyonlardan toplanan verilerden farklı yöntemlerle aykırı verilerin ayıklanması çalışması yapılmıştır. Aykırı veri ayıklama yöntemleri kullanıcı etkileşimli arayüze entegre edilmiştir. Geliştirilen ara yüz sayesinde farklı yöntemler kullanılarak kolay bir şekilde aykırı verilerin ayıklanması gerçekleştirilmiştir. Bu çalışma kapsamında;

- Literatürde bu tarz bir yazılıma rastlanmamıştır.
- 18815 adet çözülmüş oksijen (LDO), sıcaklık(T), pH, iletkenlik (EC), toplam organik karbon (TOC), nitrat azotu (NO<sub>3</sub>-N) ve amonyum azotu (NH<sub>4</sub>-N) parametrelerine ait verilerden normalden sapan değerler ayıklanmıştır.
- Bu kapsamda ortanca (*median*), ortalama(*mean*), *Grubbs*, *Gesd* ve çeyrekler açıklığı (*quartiles*) yöntemleri veri setine uygulanmıştır.
- Geliştirilen model kullanıcı etkileşimli ara yüze entegre edilmiştir.
- Algoritmalar kıyaslandığında ortanca (*median*) algoritması daha çok sayıda veri ayıklarken *Grubbs* algoritması daha az sayıda veri ayıklamıştır.
- Geliştirilen kullanıcı etkileşimli ara yüz ile veri setleri kolaylıkla sisteme yüklenip istenilen yöntemin seçilmesinin ardından tek tıkla sonuca ulaşılmasını sağlayabilmektedir.
- Veri kalitesi modelleme başarısını etkileyen önemli bir etkidir. Bundan dolayı kolay bir arayüzle yöntem seçilerek hata ayıklama programlarının kullanılması ile elde edilen verilerin modelleme başarısını arttıracığı tahmin edilmektedir.
- Tez kapsamında geliştirilmiş ara yüzün farklı disiplinlerden kullanıcılar için kolaylık sağlayacağı öngörülmektedir.

Bu yöntemlerle aykırı veriler ayıklandıktan sonra geriye kalan veri seti çeşitli modelleme yöntemleri kullanılarak modelleme çalışmaları daha başarılı bir şekilde yapılabilir. Modellerin geliştirilmesi aşamasında, kullanıcılar aykırı değer temizleme yöntemlerinden hangisi ile ayıklanan verilerden daha başarılı model elde ettiklerini değerlendirerek kullanılan aykırı veri temizleme yönteminin performanslarını da kıyaslayabilir. Bu çalışma kapsamında geliştirilen kullanıcı etkileşimli grafiksel arayüz, MATLAB ortamında geliştirilmiş olup, çeşitli programlama dilleri (Fortran, C++, C#, Java, Python vb.) ile uygulama yapılmasına da açıktır.



## KAYNAKÇA

- AB (ty). Avrupa Birliđi- EU Water Initiative, Su Çerçeve Direktifi. Eriřim Tarihi: 14.05.2022 [https://environment.ec.europa.eu/topics/water/water-framework-directive\\_en](https://environment.ec.europa.eu/topics/water/water-framework-directive_en)
- Aggarwal, C. C. (2013). On the Analytical Properties of High-Dimensional Randomization, IEEE Transactions on Knowledge and Data Engineering, 7(25), 1628-1642.
- Akkaya, C. (2002). Su Ve Toprak Kaynakları Yönetiminde Havza Yönetim Modelinin Önemi. Türkiye Mühendislik Haberleri, 421-422-423(4-5-6),20-23.
- Altuntař, M. (2018) Yeřilirmak Havzası su kalite parametrelerinin yapay zekâ teknikleriyle modellenmesi, (Yüksek Lisans Tezi), Kayseri: Erciyes Üniversitesi, Fen Bilimleri Enstitüsü.
- Arıkan, O. (2021), Yeřilirmak Nehir Havzası Yönetim Planı, Stratejik Çevresel Deđerlendirme Taslak Raporu,1-11.
- Aslam, M. (2020). Introducing Grubbs's test for detecting outliers under neutrosophic statistics – An application to medical data, Journal of King Saud University, 32(6), 2696-2700.
- Atılğan Günay, Y. Günay, S. 2012. Regresyon Analizinde Gözlemlerin Aykırı Deđer Haritası İle Sınıflandırılması. TÜİK, İstatistik Arařtırma Dergisi, 9(1), 45-51.
- Bansal, R. Gaur, N. Singh, S.N. (2016). Outlier detection: applications and techniques in data mining, 2016 6th International Conference on In Cloud System and Big Data Engineering (Confluence). India, 373-377.
- Bařer K. (2006). Sazlı Dere'nin azot ve fosfor kirliliđinin izlenmesi ve etkisinin İrdelenmesi. (Yüksek Lisans Tezi), İstanbul: Marmara Üniversitesi Fen Bilimleri Enstitüsü.
- Baykal, R.(2014).*Örnek Sorular ve Çözümleriyle Temel Matlab*, Bursa: Ekin Yayınevi.
- Bayram, A. Önsoy, H. Akıncı, G. ve Bulut V. N. (2011). Akarsu boyunca toplam organik karbon içeriđinin deđiřimi Harsit, Dođu Karadeniz Havzası, Türkiye. Çevre İzleme Deđerlendirmesi, 182,85–95.
- Bellitürk, K. ve Danıřman, F. (2007). Üre Uygulamasının Topraklarda Amonyum Ve Nitrat Oluřumuna Etkisi. OMÜ Zir. Fak. Dergisi, ,22(1), 64-72.
- Ben-Gal, I. (2005). Outlier detection, in Data Mining and Knowledge Discovery Handbook: A Complete Guide for Practitioners and Researchers, O. Maimon and L. Rockach, eds., Kluwer Academic Publishers, Israel, 131–146.
- Bonakdari, H. and Zeynoddin, M. (2022). Stochastic Modeling: A Thorough Guide to Evaluate, Pre-Process, Model and Compare Time Series with MATLAB Software. Elsevier Science.
- Borzooei, S. Miranda, G. H. B. Teegavarapu, R. Scibilia, G. Meucci, L. ve Zanetti, M.C. (2019). Assessment of weather-based influent scenarios for a WWTP: Application of a pattern recognition technique. Journal of Environmental Management.242,450-456.
- Büyükoztürk, ř. Çokluk, Ö ve Köklü, N. (2010). Sosyal Bilimler İin İstatistik, Pegem Akademi, (40), Ankara.
- Casillas-García. L.F. Anda, J. Yebra-Montes, Carlos. Shear, H. Díaz-V'azquez, D. ve Gradilla-Hern'andez, M.S. (2021). Development of a specific water quality index for the protection of aquatic life of a highly polluted urban river. Ecological Indicators, 129, 1-14.
- Ceyhan, N. (2019). Akan Verilerde Aykırı Deđer Tespiti Yaklařımları. (Yüksek Lisans Tezi), İzmir: Ege Üniversitesi, Fen Bilimleri Enstitüsü.

- Chelishchev, P. Popov, A. Sorby, K. (2018). An investigation of Outlier Detection Procedures for CMM Measurement Data. MATEC Web of Conferences, 220, 1-6.
- Chugh, N. Chugh, M. ve Agarwal, A. (2014), Outlier detection in streaming data a research perspective. 2014 International Conference on In Parallel, Distributed and Grid Computing (PDGC), India, 429-432.
- Cohn T. A. England J. F. Berenbrock C. E. Mason R. R. Stedinger J. R. ve Lamontagne, J. R. (2013). A generalized Grubbs-Beck test statistic for detecting multiple potentially influential low outliers in flood series. Water Resources Research, 49, 5047-5058.
- Coşkun, Ç. Pulatsü, S. ve Coşkun, T. (2018). *Evsel Atıksulardan Azot ve Fosforun Biyolojik Giderilme Yöntemleri*. Kilis 7 Aralık Üniversitesi Fen ve Mühendislik Dergisi, 2(2), 53-63.
- Çakır, A.(2015). Bülent Ecevit Üniversitesi, Mühendislik Fakültesi, Maden Mühendisliği Bölümü, C++ Programlama Dili Ders Notları. Erişim Tarihi: 04.07.2023. <http://www.kocaelimakine.com/wp-content/uploads/2018/01/c-ders-notlari-alaaddin-cakir.pdf>
- Devlet Su İşleri, (2019). DSİ 2019 Yılı Resmi Su Kaynakları İstatistikleri. *Türkiye Havza Numaraları ve Havzaları, 2019*. Erişim Tarihi: 25 Mart 2023. <https://www.dsi.gov.tr/Sayfa/Detay/1344>
- Dilmaç, H. ve Şişman, Y. (2023). New Approaches For Outlier Detection: The Least Trimmed Squares Adjustment. International Journal of Engineering and Geosciences, 8(1), 26-31.
- Ding, Y. Brattebø, H. ve Nord, N. (2021). A Systematic Approach For Data Analysis And Prediction Methods For Annual Energy Profiles: An Example For School Buildings İn Norway. Energy and Buildings, 247, 1319-1332.
- Erdoğan, Ö. ve Bakioğlu, B. (2021). Boğa Çayı (Antalya) Su Kalitesinin Değerlendirilmesi. Uluborlu Mesleki Bilimler Dergisi, 4(2), 91-98.
- Erdoğan, Ö. ve Ertan Ö. O. (2016). Köprüçay Nehri Nehirağzı Bölgesinin Bazı Su Kalitesi Parametrelerinin İncelenmesi. Süleyman Demirel Üniversitesi Yalvaç Akademi Dergisi, 1 (1), 58-66.
- Ersoy, E. ve Ersoy, E. (2018). Fortran Programlama Diline Giriş. Erişim Tarihi: 14 Mayıs 2023. [http://dosyalar.ersoykardesler.net/yayinlar/Fortran\\_Programlama\\_Diline\\_Giris.pdf](http://dosyalar.ersoykardesler.net/yayinlar/Fortran_Programlama_Diline_Giris.pdf).
- Esen, M. F. ve Timor, M. (2019). Çok Değişkenli Aykırı Değer Tespiti İçin Klasik Ve Dayanıkl Mahalanobis Uzaklık Ölçütleri: Finansal Veri İle Bir Uygulama. International Journal of Economic and Administrative Studies, 25, 267-282.
- Fondriest Environmental, Inc. "Suyun pH'ı." Çevresel Ölçümlerin Temelleri. 19 Kasım 2013. Erişim Tarihi: 20 Mayıs 2023. <https://www.fondriest.com/environmental-measurements/parameters/water-quality/ph/>>.
- Fondriest Environmental, Inc. "Dissolved Oxygen." Fundamentals of Environmental Measurements. Erişim Tarihi: 14 Mayıs 2023. <https://www.fondriest.com/environmental-measurements/parameters/water-quality/dissolved-oxygen/>.
- Gençoğlu, P. Bağlıtaş, H. H. ve Kuşkaya, S. (2017) Sosyal Pazarlama Aracı Olarak Kamu Spotlarının Birey Davranışları Üzerindeki Etkileri: Parametrik Olmayan İstatistiksel Bir Analiz. The Journal of International Social Research, 10(48), 622-630.
- Ghotbizadeh, M., Cuss, C.W., Grant-Weaver, I., Markov, A., Noernberg, T., Ulrich, A. ve Shoty, W. (2022). Spatiotemporal variations of total and dissolved trace elements and their distributions amongst major colloidal forms along and across the lower Athabasca River. Journal of Hydrology: Regional Studies, 40, 1-23.
- Giwa, A. Yüceer, M. Karadurmuş, E. ve Berber, R. (2008) Pkt-07 Yeşilirmak Havzasına Deşarj Edilen Endüstriyel Bir Atıksu İçin pH Modeli Ve Kalibrasyonu. Sekizinci Ulusal Kimya Mühendisliği Kongresi, İnönü Üniversitesi, Malatya.

- Göz, E. Yüceer, M. ve Karadurmuş, E. (2019). Total Organic Carbon Prediction with Artificial Intelligence Techniques. 29th European Symposium on Computer Aided Process Engineering, 46, 889-894.
- Grubbs, F.E. (1969). Procedures for detecting outlying observations in samples, *Technometrics*, 11, 1-21.
- Grubbs, F. E. ve G. Beck (1972), Extension of sample sizes and percentage points for significance tests of outlying observations, *Technometrics*, 14(4), 847-854.
- Güneş, A. ve Aktaş, M. (1992). *Kireçli bir toprakta N-servin nitrifikasyon oranı ve azot kaybı üzerine etkisi*. Doğa-Tr.J. Agri. For. 16, 501-506.
- Güneş, G. (2019). *Bartın Nehri'nin Fizikokimyasal Özelliklerinin Yağışlı Ve Kurak Dönemlerdeki Değişimi* Dokuz Eylül Üniversitesi Fen ve Mühendislik Fakültesi Fen ve Mühendislik Dergisi, 21(63), 761-774.
- Güneş, Y. (2004) Amerika Birleşik Devletleri Çevre Hukukunun Gelişimi Ve Türk Çevre Hukuku İle Karşılaştırılması. İstanbul Üniversitesi Hukuk Fakültesi Mecmuası, 62 (1-2), 81 - 122.
- Hawkins, D. M. (1980), Identification of Outliers, Chapman and Hall, (11), London.
- Housh, M. Ostfeld, A. (2015). An integrated logit model for contamination event detection in water distribution systems. *Water Res.* 75, 210-223.
- Housh, M. ve Ohar, Z. (2017). Integrating physically based simulators with Event Detection Systems: multi-site detection approach. *Water Res.* 110, 180-191.
- Hubert, M. Rousseeuw, P. J. ve Aelst, S. V. (2008). High-Breakdown Robust Multivariate Methods, *Statistical Science*, 23(1), 92-119.
- ISO16269-4, (2010). Aykırı değerlerin tespiti ve işlenmesi, Bölüm 4: Verilerin istatistiksel olarak yorumlanması.
- Java, Java teknolojisi nedir ve neden ihtiyacım var? Erişim tarihi: 04.07.2023. [https://www.java.com/tr/download/help/whatis\\_java.html](https://www.java.com/tr/download/help/whatis_java.html)
- Kadam, A.K. Wagh, V.M. Muley, A.A. Umrikar, B.N. ve Sankhua, R.N. (2019). Prediction of water quality index using artificial neural network and multiple linear regression modelling approach in Shivganga River basin, India. *Modeling Earth Syst. Environ.* 5 (3), 951-962.
- Kale, M.M.(2018) *Yeşilirmak Havzası Mekânsal Yağış Dağılımına ait Değişiminin Deterministik ve Stokastik Yöntemlerle Belirlenmesi*. *Yerbilimler Dergisi*, 39(3),263-276.
- Karadurmuş, E. (2000). Endüstriyel Atıklarla Kirlenmiş Akarsu Kullanarak Azaltma Ve Parametre Belirleme: Yeşilirmak Örnek Uygulaması, (Doktora Tezi), Ankara: Ankara Üniversitesi Fen Bilimleri Enstitüsü.
- Koren, O., Koren, M. ve Peretz, O. (2023). A procedure for anomaly detection and analysis. *Engineering Applications of Artificial Intelligence*, 117, 1-8.
- Köklü, R. (2010), *Melen Nehri Su Kalitesinin İstatistiksel Analiz Yöntemleri Ve Yapay Zekâ Teknikleri Kullanılarak Değerlendirilmesi*, (Doktora Tezi), Sakarya: Sakarya Üniversitesi, Fen Bilimleri Enstitüsü.
- Köle, M. M. ve Ataoğlu, M. (2016). Coastal Erosion of Yeşilirmak Delta Between 2000 - 2016 /Yeşilirmak Deltası'nda 2000- 2016 Yılları Arasında Gözlenen Kıyı Erozyonu. 4th International Geography Symposium- GEOMED 2016 (s. 460-468). Antalya: 4th International Geography Symposium- GEOMED2016.
- Kurucuoğlu Ökmen, G. ve Algur, Ö. F. (2011). Denitrifikasyon ve Mikroorganizmalar. *Türk Bilimsel Derlemeler Dergisi*, 4 (2), 11-15.
- Lammers, R.W. ve Bledsoe, B.P.(2019). *Quantifying pollutant loading from channel sources: watershed-scale application of the River erosion model*. *Journal of Environmental Management*. 234, 104-114.
- Liu, H. Shah S. ve Jiang, W. (2004). On-line outlier detection and data cleaning. *Computers & Chemical Engineering*, 28 (9), 1635-1647.

Liu, J. Wang, P. Jiang, D. Nan, J. ve Zhu, W. (2020). An integrated data-driven framework for surface water quality anomaly detection and early warning. *Journal of Cleaner Production*, 251, 1-12.

Mathworks, ,(1994-2023). MATLAB ve Python: Sizin İçin Hangisi Doğru? ErişimTarihi:04.07.2023. <https://ch.mathworks.com/products/matlab/matlab-vs-python.html>

Mathworks, ,(1994-2023). Rmoutliers Verilerdeki Aykırı Değerleri Tespit Edin Ve Kaldırın2018b'den Beri R2023a.ErişimTarihi:04.06.2023. <https://ch.mathworks.com/help/matlab/ref/rmoutliers.html>

Mathworks,(1994-2023). MATLAB ve Simulink Gereksinimleri, Diğer Dillere MATLAB Arabirimleri. ErişimTarihi:04.07.2023. <https://ch.mathworks.com/support/requirements/language-interfaces.html>

Meriç, B.T. (2004). Su Kaynakları Yönetimi ve Türkiye. *Jeoloji Mühendisliği Dergisi*, 28 (1), 27-38.

Moustafa, N. Choo, K.K.R. Radwan, İ. ve Seyit Camtepe, S. (2019). Outlier Dirichlet Mixture Mechanism: Adversarial Statistical Learning for Anomaly Detection in the Fog. *IEEE Transactions on Information Forensics and Security*, 14, 8.

Nacar, S. Mete. B. ve Bayram A. (2020). Günlük Çözünmüş Oksijen Konsantrasyonunun Çok Değişkenli Uyarlanabilir Regresyon Eğrileri İle Tahmin Edilmesi. *Uludağ Üniversitesi Mühendislik Fakültesi Dergisi*. 25(3), 1479-1498.

Oğuz, E. (1997). Erzurum Karasu Nehri Kirlilik Parametrelerinin Matematik Modellemesi.(Doktora Tezi) Erzurum: Atatürk Üniversitesi, Fen Bilimleri Enstitüsü.

Ovla, H. D. ve Taşdelen, B. (2012). Aykırı Değer Yönetimi. *Mersin Üniversitesi Sağlık Bilimleri Dergisi*, 5(3), 1-8.

Ölmez, M. ve Saraç, D. (2009). Su Ürünleri İçin pH'nın Önemi. *Ziraat Mühendisliği Dergisi*, 353,12-17. *Akademik Gıda* 10(4) (2012) 11-18.

Özdestan, Ö. ve Üren, A. (2012), Nitrate and Nitrite Contents of Baby Foods. *Akademik Gıda*, 10(4), 11-18.

Özkan, O. Kınacı. C. ve Sağıroğlu. Ş. (2006) Çözünmüş oksijen değişiminin Yapay Sinir Ağları ile belirlenmesi: Kızılırmak Nehri örneği. *İTÜ Dergisi*, Cilt:5(3), Kısım:1, 30-38.

Pegram, G. Conyngham, S. Aksoy. Dıvrak, B, B. ve Öztok D, (2013). Türkiye'nin Su Ayak İzi Raporu Su, Üretim ve Uluslararası Ticaret İlişkisi. Ofset Yapımevi, (68) Türkiye.

Python,(2023)3.11.4 Dokümantasyon, Tarih ve Lisans. ErişimTarihi:04.07.2023. <https://docs.python.org/3/license.html>

Rajwa, A. Rowinski, P.M. Bialik, R.J. ve Karpinski, M. (2014). Stream diurnal profiles of dissolved oxygen-case studies. In *Proceedings of the 3rd IAHR Europe Congress, Porto*.

Sáez F, Pozo C, Gómez MA, Rodelas B. ve González-López J. (2003). Growth and nitrite and nitrous oxide accumulation of *Paracoccus denitrificans* ATCC 19367 in the presence of selected pesticides. *Environ Toxicol Chem*. 22,1993-1997.

Schildt, H. (2011). *C# 4.0 - Herkes İçin*, çev: C, Akçelik. İstanbul: Alfa Yayıncılık. Erişim Tarihi: 04.07.2022. [https://www.academia.edu/29007687/C\\_kitab%C4%B1](https://www.academia.edu/29007687/C_kitab%C4%B1)

Seem, J.E. (2007). Using intelligent data analysis to detect abnormal energy consumption in buildings, *Energy and Buildings*, 39 (1), 52-58.

Sertkaya, S.(2015). Balıkesir'de katı atık depolama sahası örnekleri sularından jet loop reaktör kullanarak amonyak giderimi / Jet loop reaktörü ile Balıkesir'den katı atık depolama sahası sızıntı suyundan amonyum giderimi (Yüksek Lisans Tezi), Balıkesir: Balıkesir Üniversitesi, Fen Bilimleri Enstitüsü.

Sharma, D. ve Kansal, A. (2013). *Assessment of river quality models: a review*. *Reviews in Environmental Science and Biotechnology*. 12, 285-311.

Siepak, J. (1999). Total Organic Carbon (TOC) as a Sum Parameter of Water Pollution in Selected Polish Rivers (Vistula, Odra, and Warta). *Clean Soil Air Water*. 27(5), 251-396.

Su Kirliliği Kontrolü Yönetmeliği (1988). 4.10.1988 tarih ve 19919 sayılı Resmî Gazete, T.C. Çevre Bakanlığı. Erişim Tarihi: 14 Mayıs 2023. <https://www.resmigazete.gov.tr/arsiv/19919.pdf>

T.C. MEB. (2011). Çevre Sağlığı, Suların Analiz Parametreleri, 850CK0011, Ankara, 64.

T.C. Tarım ve Orman Bakanlığı, Su Yönetimi Genel Müdürlüğü, Taşkın Ve Kuraklık Yönetimi Dairesi Başkanlığı. (2022). Yeşilirmak Ve Batı Karadeniz Havzaları Kuraklık Yönetim Planının Hazırlanması Projesi, *Yeşilirmak Havzası Kuraklık Yönetim Planı Stratejik Çevresel Değerlendirme Kapsam Belirleme Raporu*, Ankara: y.e.y.

Tarım ve Orman Bakanlığı, (2023). *22 Mart Dünya Su Günü Kutlanıyor*. Erişim Tarihi: 25 Mart 2023. [www.tarimorman.gov.tr/Haber/5792/22-Mart-Dunya-Su-Gunu-Kutlaniyor](http://www.tarimorman.gov.tr/Haber/5792/22-Mart-Dunya-Su-Gunu-Kutlaniyor)

Tekin, S. (1997). Yeşilirmak nehri Amasya Şehir Merkezi bölümünde ve Tersakan çayında kirlilik araştırılması. (Yüksek Lisans Tezi), Samsun: Ondokuz Mayıs Üniversitesi, Fen Bilimleri Enstitüsü.

Temiz Su Yasasının Özeti, (2022). EPA, Temiz Yasalar ve Yönetmelikler, 33 USC §1251 ve devamı (1972). Erişim Tarihi: 21 Mayıs 2023. <https://www.epa.gov/laws-regulations/summary-clean-water-act>

Than, N. H., Ly, C. D. ve Tat, P.V.(2021). The performance of classification and forecasting Dong Nai River water quality for sustainable water resources management using neural network techniques. *Journal of Hydrology*, 596, 1-19.

Tietjen, G. ve Moore, R. (1972). Some Grubbs-type statistics for the detection of several outliers, *Technometrics*, 14(3), 583-597.

Tubitak Bilim Genç (2016). Dünya'daki Su. Erişim Tarihi:26.05.2023 <https://bilimgenc.tubitak.gov.tr/dergimiz/dunyadaki-su>

Türkiye Cumhuriyeti Çevre Ve Şehircilik Bakanlığı.(2019). *İl Çevre Sorunları ve Öncelikleri Araştırma Formları Türkiye Çevre Sorunları Ve Öncelikleri Değerlendirme Raporu (2017 Yılı Verileriyle)*. Ankara: Çevresel Etki Değerlendirmesi, İzin Ve Denetim Genel Müdürlüğü Çevre Envanteri ve Bilgi Yönetimi Dairesi Başkanlığı Veri Değerlendirme Şube Müdürlüğü, 43, 1-141.

Üçkardeş, F., Şahinler, S. ve Efe., E. (2010). *Aykırı Gözlemlerin Belirlenmesinde Kullanılan Bazı İstatistikler*. KSÜ Doğa Bilimleri Dergisi, 13(1), 42-45.

Vera, M. J., Dubravka, B., Nikola, J., Vojin, I. ve Bojana, P. B. (2013). *Detecting and removing outlier (s) in electromyographic gait-related patterns*. *Journal of Applied Statistics*, 40(6), 1319-1332.

Wu, L. (2009). *Mixed Effects Models for Complex Data*, Chapman and Hall/CRC, 440.

Xi, J. (2008). Outlier detection algorithms in data mining, 2008 Second International Symposium on Intelligent Information Technology Application, (1), China, 94-97.

Yan, Y., Cao, L. ve Rundensteiner, E.A. (2017). Distributed Top-N local outlier detection in big data, 2017 IEEE International Conference on Big Data (Big Data), USA, 827-836.

Yapar, A. ve Karadurmuş, E. (2017). İçme Suyunda İdari ve Uygulanabilir Çözüm Önerileri. *Türk Meslek Dergisi/Çevre Güvenliği ve Tıbbı*, 1(3), 158-164.

Yerüstü Su Kalitesi Yönetmeliğinde Değişiklik (2021). 16.06.2021 tarih ve 31513 sayılı Resmî Gazete, T.C. Tarım ve Orman Bakanlığı Erişim Tarihi: 14.05.2022 <https://www.resmigazete.gov.tr/eskiler/2021/06/20210616-1.htm>

Yetik, M. K. (2007). Akarsuların Dinamik Benzetiminde CBS Tekniklerinden Yararlanmak, (Doktora Tezi), Ankara: Ankara Üniversitesi Fen Bilimleri Enstitüsü.

Yetik, M.K., Yüceer, M., Karadurmuş, E., Semizer, E., Çalimli, A. ve Berber, R. (2014). *An interactive GIS-based software for dynamic monitoring of rivers*. *J. Environ. Prot.Ecol*. 15, 1767-1778.

Yüceer, M. (2005). Akarsularda Su Kalitesinin İzlenmesine Yönelik Yeni Bir Dinamik Benzetim Yazılımı, (Doktora Tezi), Ankara: Ankara Üniversitesi Fen Bilimleri Enstitüsü.

Yücel, A. Demir, Ö. ve Uçar, D. (2019). Yeraltı sularından biyolojik nitrat giderimde kullanılan reaktör tipleri. Dicle Üniversitesi Fen Bilimleri Enstitüsü Dergisi, 8(3), 77-83.

Zanoni, M. G. Majone, B. ve Bellin A. (2022). A catchment-scale model of river water quality by Machine Learning. Science of the Total Environment. 838, 1-17.

Zhao, Y. Li, H. Yu, X. N, Ma. Yang T. ve Zhou, J. (2022). An independent central point OPTICS clustering algorithm for semi-supervised outlier detection of continuous glucose measurements. Biomedical Signal Processing and Control, 71(B).



## EKLER



**EK-1.** 'deneme. m' dosyasına ait MATLAB kodları

```
function varargout = deneme(varargin)
% DENEME MATLAB code for deneme.fig
% DENEME, by itself, creates a new DENEME or raises the existing
% singleton*.
%
% H = DENEME returns the handle to a new DENEME or the handle to
% the existing singleton*.
%
% DENEME('CALLBACK',hObject,eventData,handles,...) calls the local
% function named CALLBACK in DENEME.M with the given input arguments.
%
% DENEME('Property','Value',...) creates a new DENEME or raises the
% existing singleton*. Starting from the left, property value pairs are
% applied to the GUI before deneme_OpeningFcn gets called. An
% unrecognized property name or invalid value makes property application
% stop. All inputs are passed to deneme_OpeningFcn via varargin.
%
% *See GUI Options on GUIDE's Tools menu. Choose "GUI allows only one
% instance to run (singleton)".
%
% See also: GUIDE, GUIDATA, GUIHANDLES
%
% Edit the above text to modify the response to help deneme
```



```
% Last Modified by GUIDE v2.5 02-Jun-2023 20:07:56
```

```
% Begin initialization code - DO NOT EDIT
```

```
gui_Singleton = 1;
```

```
gui_State = struct('gui_Name',    mfilename, ...  
                  'gui_Singleton', gui_Singleton, ...  
                  'gui_OpeningFcn', @deneme_OpeningFcn, ...  
                  'gui_OutputFcn', @deneme_OutputFcn, ...  
                  'gui_LayoutFcn', [], ...  
                  'gui_Callback', []);
```

```
if nargin && ischar(varargin{1})
```

```
    gui_State.gui_Callback = str2func(varargin{1});
```

```
end
```

```
if narginout
```

```
    [varargout{1:narginout}] = gui_mainfcn(gui_State, varargin{:});
```

```
else
```

```
    gui_mainfcn(gui_State, varargin{:});
```

```
end
```

```
% End initialization code - DO NOT EDIT
```

```
% --- Executes just before deneme is made visible.
```

```
function deneme_OpeningFcn(hObject, eventdata, handles, varargin)
```

```
% This function has no output args, see OutputFcn.
```

```
% hObject    handle to figure
```

```
% eventdata reserved - to be defined in a future version of MATLAB
```

```

% handles  structure with handles and user data (see GUIDATA)
% varargin  command line arguments to deneme (see VARARGIN)

% Choose default command line output for deneme
handles.output = hObject;

% Update handles structure
guidata(hObject, handles);

% UIWAIT makes deneme wait for user response (see UIRESUME)
% uiwait(handles.figure1);

% --- Outputs from this function are returned to the command line.
function varargout = deneme_OutputFcn(hObject, eventdata, handles)
% varargout  cell array for returning output args (see VARARGOUT);
% hObject    handle to figure
% eventdata  reserved - to be defined in a future version of MATLAB
% handles    structure with handles and user data (see GUIDATA)

% Get default command line output from handles structure
varargout{1} = handles.output;

% --- Executes on button press in pushbutton1.
function pushbutton1_Callback(hObject, eventdata, handles)
% hObject    handle to pushbutton1 (see GCBO)

```

```

% eventdata reserved - to be defined in a future version of MATLAB
% handles structure with handles and user data (see GUIDATA)

function edit1_Callback(hObject, eventdata, handles)
% hObject handle to edit1 (see GCBO)
% eventdata reserved - to be defined in a future version of MATLAB
% handles structure with handles and user data (see GUIDATA)

% Hints: get(hObject,'String') returns contents of edit1 as text
% str2double(get(hObject,'String')) returns contents of edit1 as a double

% --- Executes during object creation, after setting all properties.
function edit1_CreateFcn(hObject, eventdata, handles)
% hObject handle to edit1 (see GCBO)
% eventdata reserved - to be defined in a future version of MATLAB
% handles empty - handles not created until after all CreateFcns called

% Hint: edit controls usually have a white background on Windows.
% See ISPC and COMPUTER.
if ispc && isequal(get(hObject,'BackgroundColor'), get(0,'defaultUicontrolBackgroundColor'))
    set(hObject,'BackgroundColor','white');
end

% --- Executes on selection change in listbox1.
function listbox1_Callback(hObject, eventdata, handles)
% hObject handle to listbox1 (see GCBO)

```

```

% eventdata reserved - to be defined in a future version of MATLAB
% handles structure with handles and user data (see GUIDATA)

% Hints: contents = cellstr(get(hObject,'String')) returns listbox1 contents as cell array
% contents{get(hObject,'Value')} returns selected item from listbox1
% --- Executes during object creation, after setting all properties.
function listbox1_CreateFcn(hObject, eventdata, handles)
% hObject handle to listbox1 (see GCBO)
% eventdata reserved - to be defined in a future version of MATLAB
% handles empty - handles not created until after all CreateFcns called
% Hint: listbox controls usually have a white background on Windows.
% See ISPC and COMPUTER.
if ispc && isequal(get(hObject,'BackgroundColor'), get(0,'defaultUicontrolBackgroundColor'))
    set(hObject,'BackgroundColor','white');
end

```

## EK-2. Aykırı Veri Temizleme Yazılımı Kullanım Kılavuzu ve Yardım Dosyası

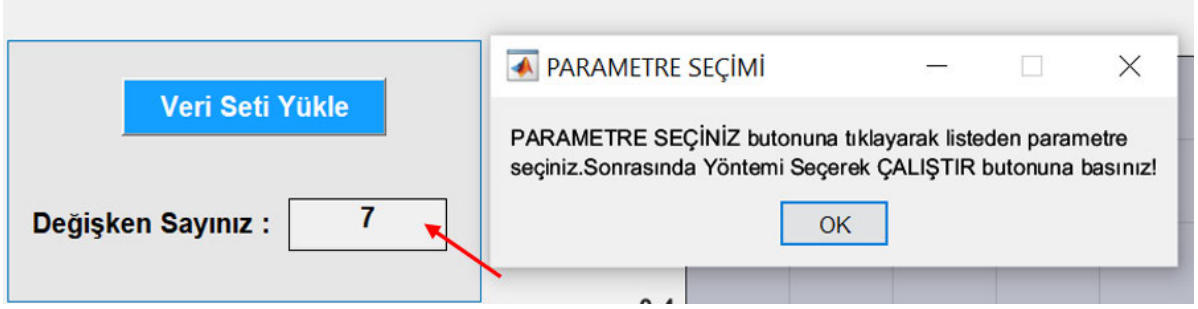
**Adım 1:** “Veri Seti Yükle” butonu ile temizlenecek verilerin yer aldığı Excel dosyasından veriler MATLAB ortamına aktarılır.



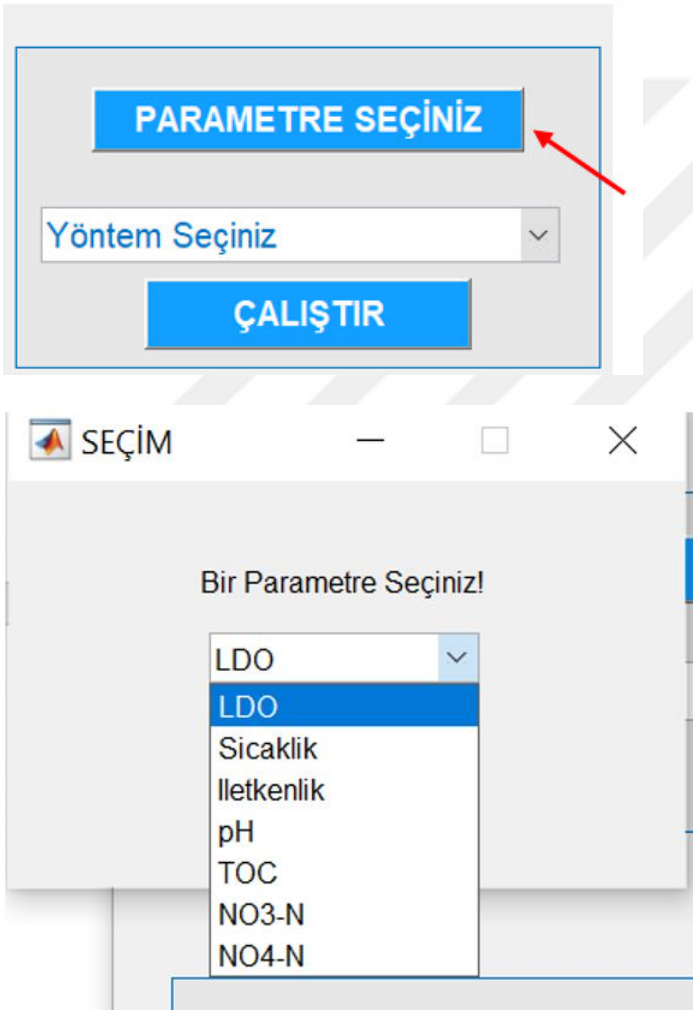
Burada dikkat edilmesi gereken önemli bir nokta excel veri setinin başlık satırının olmasıdır. Yani her bir veri kolonunun hangi parametreye/değişkene karşılık geldiği excel dosyasının ilk satırında yazılı olmalıdır. Çünkü veri setine ait her bir kolonun hangi parametreye ait olduğu buradan okunacak ve grafik çizimi için parametrelerin ismi buradan belirlenecektir. Aşağıdaki şekilde örnek dosya gösterilmiştir.

	A	B	C	D	E	F	G
1	LDO	Sicaklik	Iletkenlik	pH	TOC	NO3-N	NO4-N
2	17,86	8,15	661,194	8,38	15,56	2,65	0,36
3	18,84	6,19	692,611	7,59	6,01	1,34	0,16
4	18,27	6,19	684,533	7,39	6,03	1,41	0,18

Veri seti yüklendikten sonra MATLAB ortamına alınan bilgiler arka tarafta okunmakta ve değişken sayısı belirlenerek, kullanıcıya bilgi verilmektedir. Ayrıca bir mesaj kutusu açılarak bir sonraki aşamada ne yapılacağı konusunda bilgilendirme yapılmaktadır.



Adım 2: Gelen uyarı mesajına göre bir alt bölümde yer alan “PARAMETRE SEÇİNİZ” butonuna tıklanarak hangi parametreye ait ayıklanmış ve orijinal veri grafiği ve veri sayısı gözlenmek isteniyorsa o parametre seçilir.



Bu seçim yapıldıktan sonra (yukarıdaki örnek şekilde LDO seçilmiştir) seçim penceresi kapatılır.

**Adım 3:** Bu adımda aykırı veri ayıklama yöntemlerinden biri seçilir.

Yöntem Seçiniz

Yöntem Seçiniz

Mean

Median

Gesd

Grubbs

Quartiles

**PARAMETRE SEÇİNİZ**

Grubbs

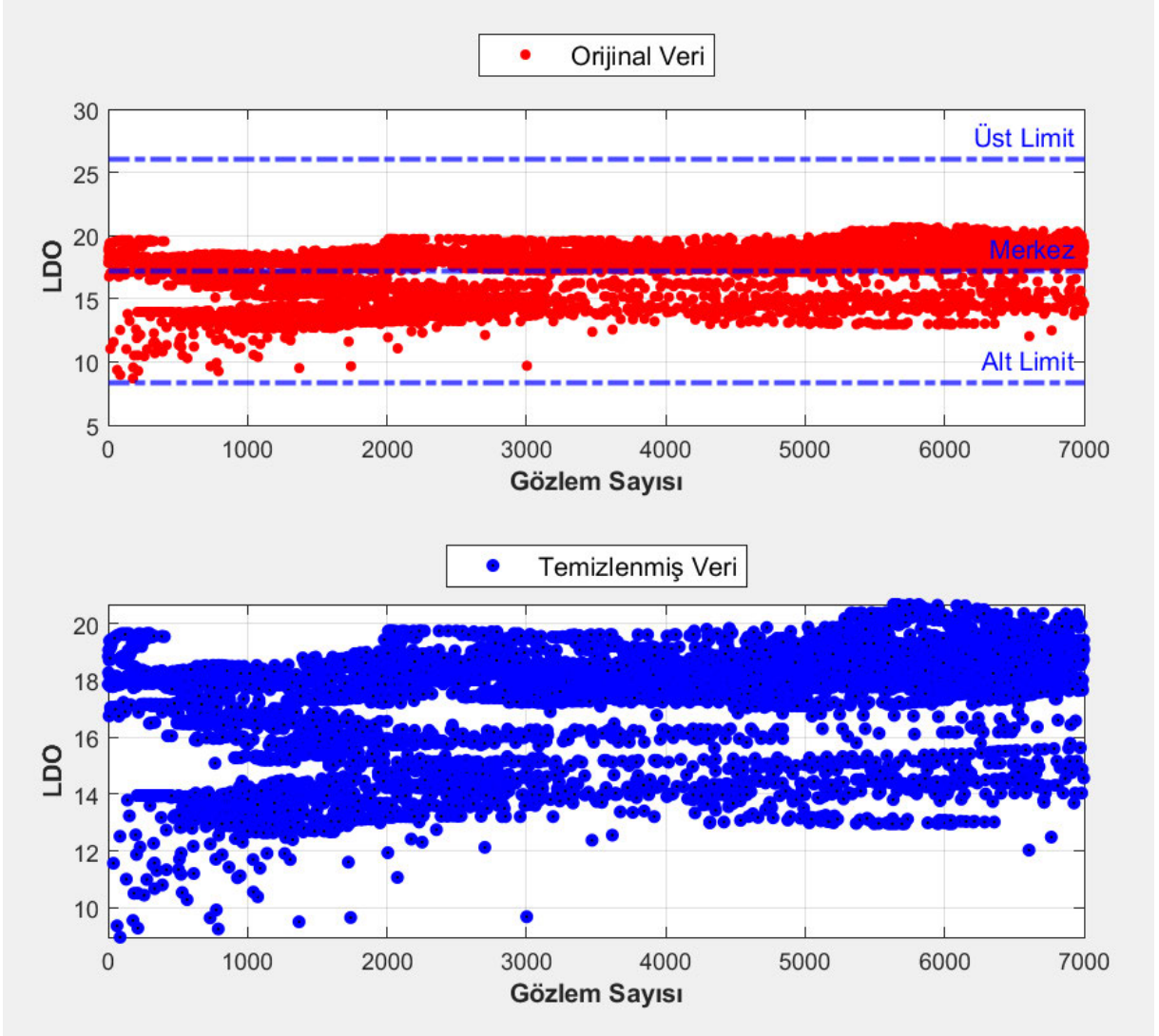
**ÇALIŞTIR**

Daha sonra “ÇALIŞTIR” butonuna basılarak orijinal verilere seçilen yöntem uygulanır. Orijinal veri sayısı ve ayıklanan veri sayısı aşağıda görüldüğü gibi yazılım üzerinden okunabilir.

**Orijinal Veri Sayısı :** **7000**

**Temizlenmiş Veri Sayısı :** **6998**

Ayrıca yazılımın sağ tarafında yer alan grafiklerde de görsel olarak izlenebilir.



Orijinal veri grafiği üzerinde çizilmiş olan Alt Limit ve Üst Limit çizgilerinin dışında kalan veriler ayıklanmış olur.

Temizlenmiş veri, "Temizlenmiş\_Data" isimli excel dosyası olarak otomatik bir şekilde kaydedilir.

Temizlenmiş veriler herhangi bir modelleme tekniği için kullanıma hazır hale getirilmiş olur. Yöntem değiştirilerek elde edilecek veri setleri ayrı ayrı modellenebilir ve model başarısına göre en uygun aykırı veri temizleme yöntemi belirlenebilir.

Burada önemli olan noktalardan biri aşağıda açıklanmıştır.

Modelleme için çok parametrelili veriler kullanıldığında tüm veri matrisindeki veriler birlikte değerlendirilerek aykırı veriler temizlendiğinden her bir değişken için kalan temizlenmiş veri sayısı aynı olmalıdır. Aksi halde bir parametrede 300 veri diğerinde 305 veri olursa veri sayıları eşit olmadığında modelleme aşamasında sorun yaşanacaktır. Örneğin bir satırda LDO

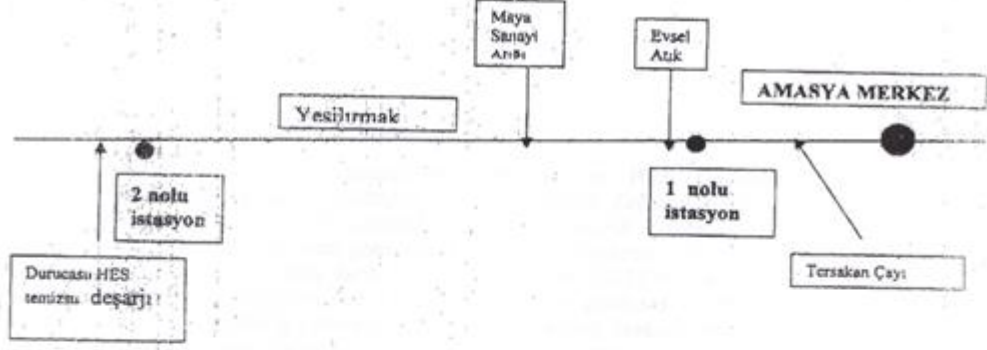


da aykırı veri varsa o satırda yer alan tüm değişkenlere ait veriler (aykırı olmasa bile) atılmış olur.

Ancak, kaç değişken için uygulama isteniyorsa orijinal veri seti ona göre ayarlanmalıdır. Örneğin 2 değişken için uygulama yapılmak isteniyorsa yüklenecek excel dosyasında sadece ilgili iki sütun yer almalıdır.



### EK-3. Ölçüm İstasyonları Yerlerinin Şematik Gösterimi



İstasyon yerleri şematik gösterimi

EK-4. İstasyon Analiz Sonuçlarının Çevre ve Orman Bakanlığı Analiz Sonuçları ile Kıyaslanması

T.C. ÇEVRE VE ORMAN BAKANLIĞI Çevre Yönetimi Genel Müdürlüğü													
ÇEVRE REFERANS LABORATUVARI KİMYA LABORATUVARI ANALİZ RAPORU													
<b>Rapor No</b>	: 08/76/15												
<b>Rapor Tarihi</b>	: 14.01.2008												
<b>Numuneyi Alan Kurum/Kişi</b>	: ÇRL 06 THL76 Mobil Ekip												
<b>Numune Alınan Yerin Adı ve Adresi</b>	: 2. İstasyon (Şehir Merkezinin 30km. Dışı) Durucusu												
<b>Numunenin Cinsi</b>	: Su												
<b>Numunenin Alındığı Nokta</b>	: İstasyonun Numune Girişi												
<b>Numune Alma Yöntemi</b>	: Anlık												
<b>Numune Alınış Tarihi</b>	: 05.01.2008												
<b>Numunenin Laboratuvara Ulaşma Tarihi</b>	: 09.01.2008												
<b>Numune Miktarı</b>	: 500 mL.												
<b>Numunenin Kabul Anındaki Durumu</b>	: Numune miktarının yetersiz olmasından dolayı şartlı kabul												
PARAMETRE	BİRİM	TABLO Özel SINIR DEĞER		ÖLÇÜLEN DEĞER	KULLANILAN METOT	İSTASYON TARAFINDAN ÖLÇÜLEN DEĞER							
		2 SAATLİK KOMPOZİT	24 SAATLİK KOMPOZİT										
pH		-	-	8,39	SM(2005)4500-H'B	8,36							
Elektriksel İletkenlik	(µS/cm)	-	-	834	-	854							
Çözülmüş oksijen	(mg/L)	-	-	11,69	SM(2005)4500-OG	12,60							
Sıcaklık	(°C)	-	-	4,8	-	4,8							
Amonyum Azotu (NH <sub>4</sub> -N)	(mg/L)	-	-	0,18	Merck Kit Metodu	0,20							
Nitrat Azotu (NO <sub>3</sub> -N)	(mg/L)	-	-	2,8	Merck Kit Metodu	2,6							
Serbest Klor	(mg/L)	-	-	0,06	Merck Kit Metodu	0,10							
Kimyasal Oksijen İhtiyacı (KOİ)	(mg/L)	-	-	23	Merck Kit Metodu	24,1							
<b>AÇIKLAMALAR</b> : pH, Elektriksel İletkenlik, Çözülmüş oksijen yerinde ölçülmüştür. KOİ, Amonyum Azotu, Serbest klor, Nitrat Azotu parametrelerinin analizleri Mobil Su Kirliliği Ölçüm Aracı'nda yapılmıştır.													
KONTROL EDEN			ONAYLAYAN										
[Redacted Signature Area]													
<ul style="list-style-type: none"> <li>- Bu analiz raporu yalnızca laboratuvara getirilen numuneyi temsil eder, örnekleme anından laboratuvarımıza teslimine kadar olan prosedürlerin ve bakılması istenilen grup ve parametrelerin belirlenmesinde teknik ve hukuki sorumluluk numuneyi alana aittir.</li> <li>- Bu rapor ve sonuçları Çevre Yönetimi Genel Müdürlüğünün izni olmadan ticari ve reklam amaçlı tamamen veya kısmen çoğaltılamaz</li> <li>- Raporlarda(*) işaretli analizler akredite parametrelerdir.</li> <li>- Rapor Numarasının başında yer alan (G) harfi raporun güncellendiğini belirtir.</li> <li>- İmzasız ve mühürlü analiz raporları geçersizdir.</li> </ul>													
<table border="0"> <tr> <td style="vertical-align: top;">H T e-1</td> <td colspan="6" style="background-color: black; color: white; text-align: center;">[Redacted Signature Area]</td> </tr> </table>							H T e-1	[Redacted Signature Area]					
H T e-1	[Redacted Signature Area]												

