



**T.C.**  
**KONYA TEKNİK ÜNİVERSİTESİ**  
**LİSANSÜSTÜ EĞİTİM ENSTİTÜSÜ**



**GNSS GÖZLEMLERİNDEKİ TROPOSFERİK  
GECİKMENİN DESTEK VEKTÖR MAKİNELERİ  
ALGORİTMASI İLE KESTİRİMİ**

**Ali Utku AKAR**

**YÜKSEK LİSANS**

**Harita Mühendisliği Anabilim Dalı**

**Temmuz-2021**  
**KONYA**  
**Her Hakkı Saklıdır**

## TEZ KABUL VE ONAYI

Ali Utku AKAR tarafından hazırlanan “GNSS GÖZLEMLERİNDEKİ TROPOSFERİK GECİKMENİN DESTEK VEKTÖR MAKİNELERİ ALGORİTMASI İLE KESTİRİMİ” adlı tez çalışması 07/07/2021 tarihinde aşağıdaki jüri tarafından oy birliği ile Konya Teknik Üniversitesi Lisansüstü Eğitim Enstitüsü Harita Mühendisliği Anabilim Dalı’nda YÜKSEK LİSANS TEZİ olarak kabul edilmiştir.

### Jüri Üyeleri

### İmza

#### Başkan-Danışman

Prof. Dr. Cevat İNAL

.....

#### Üye

Prof. Dr. Cemal Özer YİĞİT

.....

#### Üye

Dr. Öğr. Üyesi Sercan BÜLBÜL

.....

Yukarıdaki sonucu onaylarım.

Prof. Dr. Saadettin Erhan KESEN  
Enstitü Müdürü

## **TEZ BİLDİRİMİ**

Bu tezdeki bütün bilgilerin etik davranış ve akademik kurallar çerçevesinde elde edildiğini ve tez yazım kurallarına uygun olarak hazırlanan bu çalışmada bana ait olmayan her türlü ifade ve bilginin kaynağına eksiksiz atıf yapıldığını bildiririm.

## **DECLARATION PAGE**

I hereby declare that all information in this document has been obtained and presented in accordance with academic rules and ethical conduct. I also declare that, as required by these rules and conduct, I have fully cited and referenced all material and results that are not original to this work.

İmza

Ali Utku AKAR

Tarih: 07.07.2021

## ÖZET

### YÜKSEK LİSANS

## GNSS GÖZLEMLERİNDEKİ TROPOSFERİK GECİKMENİN DESTEK VEKTÖR MAKİNELERİ ALGORİTMASI İLE KESTİRİMİ

Ali Utku AKAR

Konya Teknik Üniversitesi  
Lisansüstü Eğitim Enstitüsü  
Harita Mühendisliği Anabilim Dalı

Danışman: Prof. Dr. Cevat İNAL

2021, 131 Sayfa

Jüri

Prof. Dr. Cevat İNAL  
Prof. Dr. Cemal Özer YİĞİT  
Dr. Öğr. Üyesi Sercan BÜLBÜL

Troposferik gecikme, GNSS sinyallerinin troposferdeki gaz kütleleriyle etkileşimi sonucu doğrusal hareket etmemesi ve buna bağlı olarak sinyallerin beklenen sürede alıcıya ulaşmamasından oluşmaktadır. Bu yönüyle troposferik gecikmenin erişilebilirliği ve hassas modellemesi, GNSS konumlandırma uygulamalarının yanı sıra meteorolojik çalışmalar ve hava durumu tahminlerinde önemli bir rol oynamaktadır. Günümüzde GNSS tekniği ile zenit troposferik gecikme etkisi belirlenirken, ani hava değişimine duyarlı, standartlaştırılmış ve deneysel yöntemlere dayalı parametreler içeren troposfer modelleri kullanılmaktadır. Söz konusu modeller gecikme etkisini, meteorolojik sensörü olan GNSS istasyonlarını referans alarak, diğer noktaları enterpole yöntemlerinin ya da deneysel denklemlerin kullanımı sonucu belirlemektedir. Dolayısıyla bu modeller ile gün içerisinde her zamana ait gecikme etkisinin hassas bir şekilde belirlenmesi veya erişilmesi mümkün değildir. Ayrıca teknik veya donanımsal sorunlar nedeniyle GNSS istasyonuna erişilemezse, veri arşivinde meydana gelen boşluklardan kaynaklı gecikme tahmininin kalitesi bozulmaktadır. Bu amaç doğrultusunda, ani hava değişimlerinden kaynaklanan troposferik gecikme etkisinin anlık, sürekli ve doğru bir şekilde belirlenmesi, kestirimde referans istasyonlarının birbirinden bağımsız hale getirilmesi veya teknik/donanım sorunlarının önlenmesi için yeni alternatif yaklaşımlar gerekmektedir. Son zamanlarda bilim ve mühendislik alanında yaygın kullanılan makine öğrenme algoritmalarının GNSS teknolojisine dâhil olmasıyla, güncel sorunlara yönelik yeni yaklaşımlar, ilişkilerin açıklanıp yorumlanması ve yeni karşılaştığı olaylar için çıkarımda bulunması olanakları ortaya çıkmıştır.

Bu çalışmada, zenit troposferik gecikme etkisinin, gerçek meteorolojik ve GNSS gözlem verileri ile oluşturulmuş olan makine öğrenimi modellerinden kestirilmesi amaçlanmıştır. Yöntem olarak öğrenme modellerindeki Destek Vektör Makineleri (DVM)'nin alt formu olan Destek Vektör Regresyonu (DVR) kullanılmıştır. Uygulamada, DVR modelleri oluşturulurken, gerçek ZTD değerleri ile meteorolojik parametrelerin kullanıldığı veri setine ihtiyaç vardır. Meteorolojik parametreler, EUREF (European Terrestrial Reference System) GNSS ağında çalışma bölgesi olarak seçilen "GOPE" istasyonundaki 2019-2020 yılları için elde edilmiştir. Gerçek ZTD verileri de aynı dönem için, gözlemlerle günde 6'şar saat aralıklı arazi koşullarında belirlenen VMF1 modelinden alınmıştır. Sonrasında toplanan tüm veriler düzenlenmiş ve DVR'in yapısındaki farklı matematiğe sahip; doğrusal (Doğrusal-DVR), polinomal (Polinomal-DVR) ve radyal-tabanlı (RTF-DVR) kernel fonksiyonlarının kullanımı ile gecikme modellerinden kestirimler gerçekleştirilmiştir. Kestirim sonuçlarına göre en yüksek başarı,  $R^2$  değeri 0,84 olan RTF-DVR modelinden hesaplanmıştır. RTF-DVR modeli, diğer DVR modellerine kıyasla daha iyi performans gösterdiği için çalışmada en uygun kestirim modeli olarak seçilmiştir.

Çalışmanın son aşamasında, RTF-DVR modelinden kestirilen ZTD değerleri, IGS/EPN ağından yayınlanan troposferik ürünler ve online-PPP servisleri içerisindeki CSRS-PPP’de kestirilen ZTD değerleriyle karşılaştırılmıştır. Ayrıca tüm veriler kullanılarak oluşturulan RTF-DVR modelinin iyileştirilmesi ve örneklem grupları arasında uyumsuz verilerin çıkartılmasının, model sonucunu nasıl etkilediğini görmek amacıyla çok değişkenli aykırı gözlem analizlerinden LOF (Local Outlier Factor) tekniği kullanılmıştır. LOF’a göre uyumsuz olarak tespit edilen veriler, veri setinden çıkartılıp RTF-DVR (LOF) modeli kurulmuştur. Çalışma kapsamında oluşturulan yeni modellerin performans analizleri incelenmiştir. Analiz sonuçlarına göre, RTF-DVR (LOF) ve RTF-DVR modellerinden elde edilen ZTD değerleri ile gerçek ZTD değerleri (VMF1) arasındaki karesel ortalama hata değerlerinin sırasıyla  $\pm 1,69$  cm ve  $\pm 2,08$  cm olduğu saptanmıştır. Diğer GNSS değerlendirme servislerinden olan IGS/EPN ve CSRS-PPP ile karşılaştırıldığında ise, kestirim sonuçlarının birbirine oldukça yakın olduğu ( $\sim 0,6$  cm) belirlenmiştir. Gerçekleştirilen uygulama ile GNSS ve troposfer konularına alternatif yaklaşımların sunulabileceği görülmüştür. Makine öğreniminin, GNSS uygulamalarındaki problemlere yönelik yeni çözümler veya mevcut çözümlerin verimliliğini arttırmada önemli yer tutabileceğini göstermektedir.

**Anahtar Kelimeler:** GNSS, Makine Öğrenimi, Destek Vektör Makineleri, Zenit Troposferik Gecikme, Destek Vektör Regresyonu, Troposfer Modelleri, Model Optimizasyonu



## **ABSTRACT**

## **MS THESIS**

# **ESTIMATION OF TROPOSPHERIC DELAY IN GNSS OBSERVATIONS USING SUPPORT VECTOR MACHINES ALGORITHM**

**Ali Utku AKAR**

**Konya Technical University  
Institute of Graduate Studies  
Department of Geomatics Engineering**

**Advisor: Prof. Dr. Cevat İNAL**

**2021, 131 Pages**

**Jury**

**Prof. Dr. Cevat İNAL**

**Prof. Dr. Cemal Özer YİĞİT**

**Assist. Prof. Dr. Sercan BÜLBÜL**

The tropospheric delay occurs not moving linear of GNSS signals as a result of interaction with the gas masses in the troposphere and so, the signals do not reach the receiver in the expected time. In this respect, the accessibility and precise modeling of tropospheric delay play an important role in meteorological studies and weather forecasts, as well as in GNSS positioning applications. Today, while determining the zenith tropospheric delay effect with the GNSS technique, troposphere models are used, which are insensitive to sudden air changes, include parameters based on standardized and experimental methods. These models determine the delay effect at other points (station-wise) with interpolating methods or experimental equations by referencing GNSS stations containing the meteorological sensors. Therefore, with these models, it is not possible to precisely determine or reach the delay effect of every time during the day. Besides, if a GNSS station cannot be accessed due to technical or hardware problems, the quality of the delay forecast deteriorates due to gaps in the data archive. In line with this purpose, new alternative approaches are required to determining the tropospheric delay effect due to sudden weather changes instantly, continuously and accurately, making the reference stations independent from each other in predictions or preventing technical/hardware problems. With the inclusion of machine learning algorithms, which are widely used in science and engineering, into GNSS technology, new approaches to current problems, explanation and interpretation of relationships, and inferences for new cases have been provided.

In this study, it is aimed to estimate the zenith tropospheric delay effect from machine learning models created with actual meteorological and GNSS observation data. Support Vector Regression (SVR), which is a sub-form of Support Vector Machines (SVM) in learning models, was used as a method. In application, when creating SVR models, a data set using actual ZTD values and meteorological parameters is needed. Meteorological parameters were obtained for the years 2019-2020 at the “GOPE” station selected as the study area in the EUREF (European Terrestrial Reference System) network. Actual ZTD data were also taken from the VMF1 model, which was determined by observations in field conditions with 6-hour intervals (for the years 2019-2020). Afterward, all the collected data were arranged and the model forecasts were realized by using Linear (Linear-SVR), Polynomial (Polynomial-SVR) and Radial-Based (RBF-SVR) kernel functions belong to different mathematics in the structure of SVR. According to the estimation results, the highest success was calculated from the RBF-SVR model with an  $R^2$  value of 0,84. The RBF-SVR model was chosen as the most appropriate prediction model in the study because it performs better than other SVR models.

At the last stage of the study, the ZTD values forecasted from the RBF-SVR model have compared with the tropospheric products published from the IGS/EPN network and the ZTD values forecasted from the CSRS-PPP within the online-PPP services. In addition, the LOF (Local Outlier Factor) technique, one of the multivariate outlier observation analyzes, was used to improve the RBF-SVR model created using all data and to see how the removal of outlier data between samples groups affected the model result. According to the LOF, the data detected as outlier were removed from the data set and the RBF-SVR (LOF) model was established. Performance analyzes of the newly created models were examined. According to the analysis results, root mean square error values between the ZTD values obtained from the RBF-SVR (LOF) and RBF-SVR models with the actual ZTD values (VMF1) were found to be  $\pm 1,69$  cm and  $\pm 2,08$  cm, respectively. When compared with other GNSS processing services, IGS/EPN and CSRS-PPP, the forecast results were determined to be very close to each other ( $\sim 0,6$  cm). It has been seen that alternative approaches to GNSS and troposphere topics can be presented with the realized application. It shows that machine learning can play an important role in increasing the efficiency of new solutions or existing solutions for problems in GNSS applications.

**Keywords:** GNSS, Machine Learning, Support Vector Machines, Zenith Tropospheric Delay, Support Vector Regression, Troposphere Models, Model Optimization



## ÖNSÖZ

Tezin hazırlanması süresince, tecrübelerini aktaran, fikir ve önerilerini daima benimle paylaşan danışman hocam Prof. Dr. Cevat İNAL' a, lisans eğitimimden beri hep yanımda olan, desteğini hiçbir zaman esirgemeyen, kendisinden çok şey öğrendiğim ve öğrenmeye devam ettiğim değerli hocam Doç. Dr. Şükran YALPIR'a ve ayrıca birlikte çalışma fırsatı bulduğum, her konuda bilgi ve deneyimlerini aktararak bizlere yol gösteren Prof. Dr. Esra YEL hocama teşekkürü bir borç bilirim.

Gerek çalışmalarda sürekli kendisine danıştığım gerekse fikirlerimi paylaşmam açısından daima yanımda olan, her türlü derdimi sabırla dinleyerek önerilerini ve katkılarını eksik etmeyen arkadaşım Arş. Gör. Süleyman ŞİŞMAN'a çok teşekkür ederim.

Bu zor ve sıkıntılı süreçte desteklerini hiçbir zaman esirgmeden arkamda duran, beni her konuda cesaretlendiren ve hayatımın her anında benden sevgilerini hiç eksik etmeyen çok değerli aileme sonsuz teşekkür ederim.

Ali Utku AKAR  
KONYA-2021

## İÇİNDEKİLER

|                                                                         |             |
|-------------------------------------------------------------------------|-------------|
| <b>ÖZET .....</b>                                                       | <b>iv</b>   |
| <b>ABSTRACT.....</b>                                                    | <b>vi</b>   |
| <b>ÖNSÖZ .....</b>                                                      | <b>viii</b> |
| <b>İÇİNDEKİLER .....</b>                                                | <b>ix</b>   |
| <b>ŞEKİL LİSTESİ.....</b>                                               | <b>xi</b>   |
| <b>ÇİZELGE LİSTESİ.....</b>                                             | <b>xiii</b> |
| <b>SİMGELER VE KISALTMALAR .....</b>                                    | <b>xiv</b>  |
| <b>1. GİRİŞ .....</b>                                                   | <b>1</b>    |
| 1.1. Problemin Tanımlanması .....                                       | 1           |
| 1.2. Tezin Amacı.....                                                   | 3           |
| 1.3. Tezin Önemi .....                                                  | 4           |
| 1.4. Kaynak Araştırması .....                                           | 4           |
| <b>2. GNSS VE TROPOSFERİK GECİKME ETKİSİ.....</b>                       | <b>12</b>   |
| 2.1. Uydularla Konum Belirleme Sistemleri (GNSS).....                   | 12          |
| 2.2. GNSS ile Konum Belirleme Yöntemleri .....                          | 12          |
| 2.3. Troposferik Gecikme Etkisi.....                                    | 14          |
| 2.4. Troposferik Gecikmenin GNSS Gözlemlerinde Modellenmesi .....       | 15          |
| 2.4.1. Zenit troposferik gecikme modelleri.....                         | 17          |
| 2.4.1.1. Zenit hidrostatik gecikme .....                                | 17          |
| 2.4.1.2. Zenit ıslak gecikme .....                                      | 20          |
| 2.4.1.3. Troposferik izdüşüm fonksiyonları.....                         | 20          |
| 2.4.1.3.1. VMF1 troposferik izdüşüm fonksiyonu.....                     | 22          |
| <b>3. MATERYAL VE YÖNTEM.....</b>                                       | <b>25</b>   |
| 3.1. Materyal .....                                                     | 28          |
| 3.1.1. Çalışma bölgesinin belirlenmesi .....                            | 28          |
| 3.1.2. Verilerin elde edilmesi ve düzenlenmesi .....                    | 31          |
| 3.1.3. Düzenleme sonucu elde edilen veri setleri .....                  | 34          |
| 3.2. Yöntem.....                                                        | 36          |
| 3.2.1. Makine öğrenmesi.....                                            | 36          |
| 3.2.2. Makine öğrenme algoritmaları ve öğrenme çeşitleri .....          | 38          |
| 3.2.2.1. Gözetimli öğrenme (Supervised learning) .....                  | 39          |
| 3.2.2.2. Gözetimsiz öğrenme (Unsupervised learning) .....               | 40          |
| 3.2.2.3. Pekiştirmeye dayalı öğrenme (Reinforcement learning) .....     | 41          |
| 3.2.3. Destek vektör makineleri .....                                   | 43          |
| 3.2.4. Destek vektör makinelerinde kullanılan kernel fonksiyonlar ..... | 51          |
| 3.2.5. Destek vektör regresyonu .....                                   | 54          |

|                                                                               |            |
|-------------------------------------------------------------------------------|------------|
| 3.2.5.1. Doğrusal destek vektör regresyonu.....                               | 54         |
| 3.2.5.2. Doğrusal olmayan destek vektör regresyonu .....                      | 58         |
| 3.2.6. Performans analiz yöntemleri .....                                     | 61         |
| 3.2.7. Çok değişkenli aykırı gözlem analizi.....                              | 61         |
| <b>4. ARAŞTIRMA SONUÇLARI VE TARTIŞMA.....</b>                                | <b>64</b>  |
| 4.1. Veri Setinin İstatiksel Analizleri.....                                  | 65         |
| 4.2. Meteorolojik Veriler-ZTD İlişkisinin Değerlendirilmesi .....             | 69         |
| 4.3. ZTD'deki Mevsimsel Etkilerin İstatistiki Değerlendirilmesi .....         | 70         |
| 4.4. Zenit Troposferik Gecikmenin DVM ile Kestirimi .....                     | 73         |
| 4.4.1. Doğrusal (Linear) kernel fonksiyonu ile modelleme .....                | 73         |
| 4.4.2. Polinomal (Polynomial) kernel fonksiyonu ile modelleme .....           | 77         |
| 4.4.3. Radyal tabanlı (RTF) kernel fonksiyonu ile modelleme .....             | 80         |
| 4.4.4. Model kestirimlerine ait analiz sonuçları.....                         | 84         |
| 4.4.5. Uyuşumsuz verilerin model sonuçlarına etkisi .....                     | 93         |
| 4.5. Farklı Troposferik Gecikme Kestirim Modellerinin Karşılaştırılması ..... | 94         |
| <b>5. SONUÇLAR VE ÖNERİLER.....</b>                                           | <b>96</b>  |
| <b>KAYNAKLAR .....</b>                                                        | <b>99</b>  |
| <b>EKLER .....</b>                                                            | <b>109</b> |

## ŞEKİL LİSTESİ

|                                                                                                  |    |
|--------------------------------------------------------------------------------------------------|----|
| Şekil 2.1. GNSS'deki taşıyıcı faz ve kod gözlemleri için düzeltme denklemleri.....               | 13 |
| Şekil 2.2. Troposfer katmanındaki troposferik gecikme etkisi .....                               | 14 |
| Şekil 2.3. GNSS gözlemlerinden atmosferik ürünlerin elde edilmesi .....                          | 15 |
| Şekil 2.4. Ayrık troposferik gecikme modelleri .....                                             | 24 |
| Şekil 2.5. Ampirik troposferik gecikme modelleri .....                                           | 24 |
| Şekil 3.1. Çalışmadaki iş akışı .....                                                            | 27 |
| Şekil 3.2. EUREF ağında GNSS-Meteorolojik sensörlerin entegre çalıştığı istasyonlar              | 28 |
| Şekil 3.3. EPN/IGS referans istasyonlarındaki meteorolojik gözlem (gün) sayıları .....           | 30 |
| Şekil 3.4. Çalışmada kullanılan GOPE referans istasyonu .....                                    | 31 |
| Şekil 3.5. Çalışmadaki troposferik gecikmeye etki eden kriterler .....                           | 32 |
| Şekil 3.6. Uygulama için elde edilen verilere ait örnek dosyalar .....                           | 33 |
| Şekil 3.7. Makine öğrenmesindeki temel işlem süreci .....                                        | 37 |
| Şekil 3.8. Makine öğrenme tekniğindeki öğrenme çeşitleri .....                                   | 38 |
| Şekil 3.9. Gözetimli öğrenme tekniği için sürecin işleyişi .....                                 | 39 |
| Şekil 3.10. Gözetimsiz öğrenme tekniği için sürecin işleyişi .....                               | 40 |
| Şekil 3.11. Pekiştirmeli öğrenme tekniği için sürecin işleyişi .....                             | 42 |
| Şekil 3.12. DVM ile iki boyutta doğrusal ayrılabilen verilerin sınıflandırması .....             | 44 |
| Şekil 3.13. DVM ile oluşturulan küçük (a) ve büyük (b) boşluklu hiper düzlemler .....            | 45 |
| Şekil 3.14. Destek vektörü ( $X_p$ ) ile optimum hiper düzlem arasındaki geometrik ilişki        | 46 |
| Şekil 3.15. İki hiper düzlem arasındaki dik uzaklık .....                                        | 47 |
| Şekil 3.16. Verilerin doğrusal olarak ayrılamadığı durum .....                                   | 48 |
| Şekil 3.17. Doğrusal ayrılamayan veriler için tanımlanmış hiper düzlemler .....                  | 49 |
| Şekil 3.18. Doğrusal olmayan verinin kernel fonksiyonu yardımıyla doğrusal olarak ayrılması..... | 50 |
| Şekil 3.19. Birinci, ikinci ve n. dereceden polinoma ait çekirdek fonksiyonu .....               | 52 |
| Şekil 3.20. Karmaşık yapıdaki verilerin RTF ile gösterimi .....                                  | 53 |
| Şekil 3.21. DVR'deki kayıp fonksiyonlar .....                                                    | 54 |
| Şekil 3.22. Doğrusal destek vektör regresyonunda hata toleransı ( $\epsilon$ ) .....             | 55 |
| Şekil 3.23. Doğrusal destek vektör regresyonunda hata toleransı ve esneklik parametresi .....    | 56 |
| Şekil 3.24. Doğrusal olmayan bir destek vektör regresyon fonksiyonu .....                        | 58 |
| Şekil 3.25. Doğrusal ve doğrusal olmayan DVR'de sınır genişliği .....                            | 59 |
| Şekil 3.26. LOF tekniği ile belirlenmiş olan aykırı gözlemler .....                              | 62 |
| Şekil 4.1. Model kurulum ve değerlendirmede kullanılan Python v.3.8 yazılımı .....               | 65 |
| Şekil 4.2. Saat 00:00'a ait veri setindeki değişkenlerin istatistiki bilgileri.....              | 67 |
| Şekil 4.3. Saat 06:00'a ait veri setindeki değişkenlerin istatistiki bilgileri.....              | 67 |
| Şekil 4.4. Saat 12:00'a ait veri setindeki değişkenlerin istatistiki bilgileri.....              | 68 |
| Şekil 4.5. Saat 18:00'a ait veri setindeki değişkenlerin istatistiki bilgileri.....              | 68 |
| Şekil 4.6. Basınç(a), Sıcaklık(b) ve Bağıl nem(c) değişkenlerinin ZTD ile ilişkisi .....         | 69 |
| Şekil 4.7. Değişkenler arasındaki korelasyon çaprazlamaları .....                                | 71 |
| Şekil 4.8. Mevsimsel etkilere göre ZTD'nin değişimi.....                                         | 72 |
| Şekil 4.9. Farklı (C) ceza/düzeltilme parametrelerine ait model doğrulukları.....                | 73 |
| Şekil 4.10. Farklı (C) ceza/düzeltilme parametrelerine ait model doğrulukları.....               | 74 |
| Şekil 4.11. Model ve gerçek gecikme değerleri arasındaki ilişki .....                            | 76 |
| Şekil 4.12. Farklı (C) ceza/düzeltilme parametrelerine ait model doğrulukları.....               | 77 |
| Şekil 4.13. Farklı C ve d ikili parametre değerlerine ilişkin model doğrulukları .....           | 78 |
| Şekil 4.14. Model ve gerçek gecikme değerleri arasındaki ilişki .....                            | 79 |

|                                                                                                               |    |
|---------------------------------------------------------------------------------------------------------------|----|
| Şekil 4.15. Farklı (C) ceza/düzeltilme parametrelerine ait model doğrulukları.....                            | 80 |
| Şekil 4.16. Farklı C ve $\dot{Y}$ ikili parametre değerlerine ilişkin model doğrulukları .....                | 81 |
| Şekil 4.17. Model ve gerçek gecikme değerleri arasındaki ilişki .....                                         | 83 |
| Şekil 4.18. RTF-DVR (00:00) modelinin IGS/EPN gecikme değerleriyle uyumu .....                                | 85 |
| Şekil 4.19. RTF-DVR (06:00) modelinin IGS/EPN gecikme değerleriyle uyumu .....                                | 86 |
| Şekil 4.20. RTF-DVR (12:00) modelinin IGS/EPN gecikme değerleriyle uyumu .....                                | 87 |
| Şekil 4.21. RTF-DVR (18:00) modelinin IGS/EPN gecikme değerleriyle uyumu .....                                | 88 |
| Şekil 4.22. RTF-DVR (00:00) modelinin CSRS-PPP gecikme değerleriyle uyumu.....                                | 91 |
| Şekil 4.23. RTF-DVR (06:00) modelinin CSRS-PPP gecikme değerleriyle uyumu.....                                | 91 |
| Şekil 4.24. RTF-DVR (12:00) modelinin CSRS-PPP gecikme değerleriyle uyumu.....                                | 92 |
| Şekil 4.25. RTF-DVR (18:00) modelinin CSRS-PPP gecikme değerleriyle uyumu.....                                | 92 |
| Şekil 4.26. Farklı (1-Tüm, 2-LOF) verilerden elde edilen değerler ile gerçek değerler arasındaki ilişki ..... | 94 |
| Şekil 4.27. Farklı ZTD kestirim modellerinin karesel ortalama hata değerleri .....                            | 95 |
| Şekil 5.1. Çalışmadaki farklı gecikme model sonuçlarının birbirine yakınlık derecesi .                        | 97 |



## ÇİZELGE LİSTESİ

|                                                                                                            |    |
|------------------------------------------------------------------------------------------------------------|----|
| Çizelge 1.1. Konuya ilişkin yapılan diğer çalışmalar .....                                                 | 10 |
| Çizelge 2.1. Hidrostatik ve toplam izdüşüm fonksiyonları için kullanılan parametreler .....                | 23 |
| Çizelge 2.2. Farklı izdüşüm fonksiyonlarına ait hata değerleri ( $e_0 = 5^\circ$ ).....                    | 23 |
| Çizelge 3.1. Farklı bölgedeki referans istasyonlarına ait konum ve eksik meteorolojik veri bilgisi .....   | 29 |
| Çizelge 3.2. Verilerin elde ediliş biçimi ve kullanılma şekli hakkında genel bilgi .....                   | 34 |
| Çizelge 3.3. Düzenleme sonucu elde edilen veri seti (GOPE, 00:00 için) .....                               | 35 |
| Çizelge 4.1. Veri setine ait bilgiler .....                                                                | 66 |
| Çizelge 4.2. Doğrusal (Linear) DVR modeli performans sonuçları .....                                       | 75 |
| Çizelge 4.3. Doğrusal-DVR modeline ait genel hata değerleri .....                                          | 76 |
| Çizelge 4.4. Polinomal (Polynomial) DVR modeli performans sonuçları .....                                  | 78 |
| Çizelge 4.5. Polinomal-DVR modeline ait genel hata değerleri.....                                          | 80 |
| Çizelge 4.6. Radyal tabanlı (RTF) DVR modeli performans sonuçları.....                                     | 82 |
| Çizelge 4.7. Radyal Tabanlı-DVR modeline ait genel hata değerleri .....                                    | 83 |
| Çizelge 4.8. RTF-DVR modelleri ve CSRS-PPP'den elde edilen gecikme değerleri arasındaki farklar (mm) ..... | 89 |
| Çizelge 4.9. LOF tekniğine göre aykırı olarak tespit edilen günler.....                                    | 93 |
| Çizelge 5.1. Farklı DVR modellerindeki performans sonuçları .....                                          | 96 |

## SİMGELER VE KISALTMALAR

### SİMGELER

|                        |                                                 |
|------------------------|-------------------------------------------------|
| $\Delta_{trop}$        | : Troposferik gecikme fonksiyonu                |
| $g_m^0$                | : Ortalama gravite değeri                       |
| $D_h^z$                | : Zenit hidrostatik gecikme                     |
| $D_w^z$                | : Zenit ıslak gecikme                           |
| $\rho_m$               | : Kuru hava ve su buharının toplam yoğunlukları |
| $R_d$                  | : Evrensel gaz sabiti                           |
| $\Delta L_{gec}$       | : Geometrik gecikme etkisi                      |
| $\gamma$               | : Gamma parametresi                             |
| $w$                    | : Ağırlık vektörü                               |
| $\epsilon$             | : Epsilon parametresi (Hata toleransı)          |
| $C$                    | : Ceza/Düzeltilme Parametresi                   |
| $\zeta^+$ ve $\zeta^-$ | : Esneklik Parametreleri                        |
| $K(x_i, x_j)$          | : Kernel fonksiyonları                          |
| $\lambda_i$            | : Lagrange çarpanı                              |

### KISALTMALAR

|                 |                                                                                         |
|-----------------|-----------------------------------------------------------------------------------------|
| <b>ANFIS</b>    | : Adaptive-Network Based Fuzzy Inference Systems                                        |
| <b>BDS</b>      | : Beidou Satellite Navigation System                                                    |
| <b>CSRS-PPP</b> | : Canadian Spatial Reference System-Precise Point Positioning Service                   |
| <b>DVM</b>      | : Destek Vektör Makinesi (Support Vector Machine)                                       |
| <b>DVR</b>      | : Destek Vektör Regresyonu (Support Vector Regression)                                  |
| <b>ECMWF</b>    | : European Centre for Medium-Range Weather Forecasts                                    |
| <b>EGNOS</b>    | : European Geostationary Navigation Overlay System                                      |
| <b>EPN</b>      | : European Permanent Network                                                            |
| <b>EPOSA</b>    | : Echtzeit Positionierung Austria GNSS Network                                          |
| <b>ERA5</b>     | : European Reanalysis-5                                                                 |
| <b>EUREF</b>    | : European Terrestrial Reference System                                                 |
| <b>GDOP</b>     | : Geometric Dilution of Precision                                                       |
| <b>GLONASS</b>  | : Globalnaya Navigatsionnaya Sputnikovaya Sistema (GLObal Navigation Satellite Systems) |
| <b>GMF</b>      | : Global Mapping Function                                                               |
| <b>GNSS</b>     | : Global Navigation Satellite Systems                                                   |
| <b>GPS</b>      | : Global Positioning System                                                             |
| <b>GPT</b>      | : Global Pressure and Temperature Model                                                 |
| <b>GPT3</b>     | : Generative Pre-trained Transformer 3                                                  |
| <b>GRAD</b>     | : A priori Horizontal Gradients Model                                                   |
| <b>IDW</b>      | : Inverse Distance Weighting Method                                                     |
| <b>IGS</b>      | : International GNSS Service                                                            |
| <b>IMF</b>      | : Isobaric Mapping Function                                                             |
| <b>IPWV</b>     | : Integrated Precipitable Water Vapor                                                   |
| <b>KOH</b>      | : Karesel Ortalama Hata                                                                 |
| <b>LIN</b>      | : Linear Interpolation Method                                                           |
| <b>LOF</b>      | : Local Outlier Factor                                                                  |
| <b>MF</b>       | : Mapping Function                                                                      |

|                     |                                                          |
|---------------------|----------------------------------------------------------|
| <b>NARX</b>         | : Nonlinear Autoregressive Network with Exogenous Inputs |
| <b>NMF</b>          | : Niell Mapping Function                                 |
| <b>NWP</b>          | : Numerical Weather Prediction                           |
| <b>OMH</b>          | : Ortalama Mutlak Hata                                   |
| <b>OMYH</b>         | : Ortalama Mutlak Yüzde Hata                             |
| <b>PPP</b>          | : Precise Point Positioning                              |
| <b>PWV</b>          | : Precipitable Water Vapor                               |
| <b>RINEX</b>        | : Receiver Independent Exchange Format                   |
| <b>RTF</b>          | : Radyal Tabanlı Fonksiyon (Radial-Bases Function)       |
| <b>RTK</b>          | : Real Time Kinematic                                    |
| <b>TAWES</b>        | : The New Austrian Meteorological Measurement Network    |
| <b>TEC</b>          | : Total Electron Content                                 |
| <b>TUSAGA-Aktif</b> | : Türkiye Ulusal Sabit GNSS Ağı-Aktif                    |
| <b>TUTGA</b>        | : Türkiye Ulusal Temel GPS Ağı                           |
| <b>UNB</b>          | : University of New Brunswick                            |
| <b>UTC</b>          | : Coordinated Universal Time                             |
| <b>VLBI</b>         | : Very Long Baseline Interferometry                      |
| <b>VMF</b>          | : Vienna Mapping Function                                |
| <b>WVR</b>          | : Water Vapor Radiometers                                |
| <b>YSA</b>          | : Yapay Sinir Ağları                                     |
| <b>ZHD</b>          | : Zenith Hydrostatic Delay                               |
| <b>ZTD</b>          | : Zenith Tropospheric Delay                              |
| <b>ZWD</b>          | : Zenith Wet Delay                                       |

## 1. GİRİŞ

### 1.1. Problemin Tanımlanması

Günümüzde çok fazla ihtiyaç duyulan Uydularla Konum Belirleme Sistemleri (GNSS: Global Navigation Satellite System), nokta konumlarının belirlenmesi, atmosferik çalışmalar, navigasyon hizmetleri, yerel ve global deformasyon ölçmeleri, şehir planlaması ve arazi yönetimi gibi farklı amaçlara uzanan geniş yelpazede kendisine uygulama alanı bulmaktadır. Harita Mühendisliği açısından kapsamlı bir yere sahip olan GNSS tekniği ile hassas konum bilgisinin ya da çalışma kapsamındaki atmosfer modellemelerinin elde edilebilmesi için rölatif (bağıl) ve mutlak konum belirleme başta olmak üzere farklı yöntemler kullanılmaktadır. İlgili çalışmalarda etkin olarak kullanılan bu yöntemlerin zaman içerisinde sürekli geliştirilebilmesi, kullanıcılar için emek, zaman, maliyet ve doğruluk gibi faktörler göz önüne getirildiğinde önemli faydalar sağlamıştır (Öcalan, 2015).

Özellikle uygulama alanı oldukça artan GNSS teknolojisinde kullanılan GNSS sinyalleri, atmosferde yayılırken iyonosferdeki elektron içeriği ve nötr troposferdeki hava yoğunluğundan etkilenmektedir. Bu sinyallerin atmosferdeki çeşitli kütlelerle etkileşimini en aza indirebilmek için araştırmalarda çift frekanslı GNSS alıcıları olmasına özen gösterilmektedir (Angrisano ve ark., 2013). Ancak bu durum sadece iyonosfer tabakasına dayalı hataları gidermekte olup, troposfer tabakasında meydana gelen hataları giderememektedir. Troposfer tabakası, atmosferin yeryüzüne en yakın bölgesinde olup, kalınlığı ekvatorдан kutuplara doğru gidildikçe azalmaktadır. Bu farklılığı göstermesinin temel nedeni, ekvatorial bölgelerde hem sıcaklığa bağlı olarak ısınan havanın yükselmesi hem de dış kuvvet etkilerinin kutuplara nazaran daha büyük olmasıdır (Thayer, 1974). Ayrıca kalınlık farkları mevsimlere göre (meteorolojik olaylar) çok fazla değişiklik gösterebilmektedir. Dolayısıyla bu yönüyle farklı tabaka kalınlığına sahip olan troposfer, sinyallere doğrudan etki ettiğinden, oldukça önemli bir konudur ve birtakım gecikmelere sebep olur (Pace ve ark., 2010).

Literatürde troposferik gecikme olarak adlandırılan konu, GNSS sinyallerinin troposferdeki kütleler ile etkileşimi sonucu doğrusal hareket etmemesi ve buna bağlı olarak sinyallerin ya olması gerektiğinden daha fazla yol alması ya da daha kısa zamanda alıcıya ulaşmasıdır. Zenit yönündeki indirgenmiş durumu ise zenit troposferik gecikmesi şeklinde ifade edilmektedir. Gecikme etkisi, hidrostatik (kuru) ve ıslak

bileşenden meydana gelmektedir (Boehm ve ark., 2007). Genellikle hidrostatik bileşen deneysel (ampirik) modellerle doğrudan hesaplanabilmektedir. Ancak ıslak bileşen için kesin deneysel model yoktur. Çünkü ıslak bileşendeki su buharının dağılımı düzensiz olduğu için modellenememektedir (Geiger, 1988; Tregoning ve Herring, 2006). GNSS tabanlı uygulamalarda kısıtlayıcı kriterler olduğu için konum belirlemede, meteoroloji alanında daha yüksek hassasiyetin sağlanması adına gözlemlere bileşen etkilerinin dâhil edilmesi şarttır (Leick, 2004). Ayrıca ıslak gecikmenin belirlenmesinde, öncül olarak kullanılan hidrostatik bileşen değerlerinin doğru modellenmesi, ıslak bileşenin gerçeğe yakın kestirilebilmesinde önemli bir unsurdur böylelikle bileşenlerin her ikisinin de iyi kestirilmesinde fayda vardır (Dach ve ark., 2015).

Troposferik gecikmenin; bölgeden bölgeye, yüksekliğe, ani hava değişimine ve meteorolojik olaylar vb. gibi kriterlere bağlı sürekli değişim gösterme problemini aşmak amacıyla “Boehm” önderliğinde birçok bilim insanı tarafından troposferik gecikme modelleri geliştirilmiştir (Boehm ve ark., 2006a; Boehm ve ark., 2006b; Boehm ve ark., 2015). Ancak geliştirilen modeller, arazideki gerçeğe yakın sonuçların elde edilmesi için doğru tercih olmamaktadır. Çünkü gecikme modellerinden üretilen ürünler, deneysel yöntemlere dayalı olarak belirlenmekte ve farklı bölgeler için grid katsayılarına göre elde edilen ortalama değerleri kullanmaktadır (Emardson ve Johansson, 1998). Diğer yandan küresel ısınma sonucu dünya üzerindeki troposfer tabakasının kalınlığı sürekli değişim içerisinde olması, atmosferik ürünlerin doğruluğunu olumsuz yönde etkilemektedir. Bu nedenle modellerde, yüksek hassasiyetli kestirimler zorlaşmaktadır (Stende, 2006).

Gecikmenin yüksek hassasiyette kestirimi için çözüm olarak, ya uygulamadaki her istasyona ait meteorolojik veriler yardımıyla ayrı ayrı düzeltme değerinin hesaplanması gerekir ya da arazi koşullarındaki meteorolojik değerler ile hesap yapılan bir troposfer modeli tercih edilmelidir. Ancak dünya üzerinde GNSS ölçme çalışmalarında, entegreli kullanılabilen çok az sayıda meteorolojik sensörler bulunmaktadır. Günlük hayatta her referans istasyonunda GNSS-Meteoroloji parametre hesabı yapabilmek de mümkün değildir (Osah ve ark., 2021). Dolayısıyla bu durumlar dikkate alındığında, gerçek verilerle oluşturulmuş bir troposfer modeli/metodolojisi kullanılması daha doğru olacaktır. Gün içerisinde sürekli değişen troposfer için, ani hava değişimine duyarlı ve anlık olarak arazi koşullarına göre kestirim yapan troposfer modeli kullanmak yerine, birtakım varsayımlardan oluşan (anlık değişimlere duyarlı) deneysel modellerin kullanılması problemin tanımı şeklinde ifade edilmiştir.

## 1.2. Tezin Amacı

GNSS ölçümlerinin doğruluğu, büyük ölçüde sinyalin uydudan alıcıya giden yolda yaşadığı gecikmeye bağlıdır. Gecikmeye çeşitli etkiler neden olur. Bu etkilerden birisi de troposferin sinyali kırma gücüdür. Etkinin türü çeşitli deneysel araştırmalarla modellenabilir ancak bazı durumlarda gecikmeyi doğru bir şekilde belirlemede problemler yaşanmaktadır. Troposferik gecikmeyi kestirmenin alternatif bir yaklaşımı son yirmi yılda bilim ve mühendisliğin birçok alanında yaygın olarak kullanılan makine öğrenimi algoritmalarıdır. Bu tür algoritmalar, söz konusu ilişkileri önceden özel olarak modellemeye gerek kalmadan, büyük veri kümelerindeki (Big-Data) ilişkilerin çıkarılmasına izin verir. Bu bağlamda;

Tez çalışmasının amacı; zenit troposferik gecikme (ZTD: Zenith Tropospheric Delay) değerlerinin, gerçek meteorolojik ve GNSS gözlem verileri ile kurulacak olan destek vektör makinesi (DVM) modelleri yardımıyla kestirilmesidir. Daha sonra yeni modelden kestirilen gecikme değerlerinin, diğer uluslararası GNSS değerlendirme servislerden (IGS/EPN ve CSRS-PPP) elde edilen gecikme değerleri ile karşılaştırılmasının yapılmasıdır. Çalışmanın ilk bölümünde kaynak araştırmasına dayanarak konu kapsamındaki genel bilgiler ve ulusal/uluslararası yapılan diğer çalışmalar irdelenmiştir. Bunun dışında problemin tanımlanması, tezin amacı ve önemi açıklanmıştır. İkinci bölümde GNSS ve troposferik gecikmenin modellenmesi konuları üzerinde geçmişten günümüze bilim insanları tarafından ortaya atılan yöntemler incelenmiştir. Üçüncü bölüm materyal ve yöntem diye ikiye ayrılmıştır. Materyal kısmında dünya üzerinde farklı bölgelerde dağılım sergileyen, GNSS/meteorolojik sensörlerin entegreli veri topladığı istasyonlardan, uygulamada yer alacak çalışma bölgesinden ve kullanılacak verilerden bahsedilmiştir. Yöntem kısmında ise troposferik gecikme kestirimi için kullanılacak olan makine öğrenme yaklaşımı, DVM, DVM'nin formu olan destek vektör regresyonu (DVR) ve çok değişkenli aykırı gözlem analizi (LOF: Local Outlier Factor) yöntemleri açıklanmıştır. Dördüncü bölümde model oluşturmak için gerekli olan veri setinin istatistiksel analizleri, toplanan verilerin ilişkisinin değerlendirilmesi, ZTD'nin mevsime bağlı değişimlerinin etkisinin incelenmesi, ZTD'nin yeni oluşturulan DVM/DVR modelleriyle kestirimi ve farklı troposferik gecikme kestirim modellerinin karşılaştırılması gerçekleştirilmiştir. Son olarak, beşinci bölümde araştırma neticesinde analizlerden çıkan sonuçlar değerlendirilip, tartışılarak konu ile ilgili sonuç ve önerilerde bulunulmuştur.

### 1.3. Tezin Önemi

Yapılan tez çalışması, GNSS uygulamalarında makine öğrenimi potansiyelinin araştırılması için önem arz etmektedir. Yeni makine öğrenme modelinin yüksek başarı performansı ile kurulması halinde meteorolojik sensörler bulunmayan ve ortalama değerlere göre kestirim yapan farklı istasyon noktalarında da gerçeğe daha yakın ZTD kestirimleri yapabilecektir. Ayrıca uygulanacak yeni modelde, troposfer gecikme hatası, tamamen istasyonlar arasındaki baz uzunluklarından bağımsız hale getirilecektir. İstenilen zamanda modele o anki arazi hava durumu parametreleri tanıtılarak anlık gecikme kestirimi yapılacaktır.

Diğer taraftan bazı durumlarda da referans istasyonlarında teknik ya da donanımsal problemlerden kaynaklı ölçüm hataları yapılmaktadır. Kaba ve sistematik olarak sayılabilecek bu ölçüm hataları yüzünden kestirimler hatalı elde edilmektedir. Daha sonrasında kestirimlere dayalı ölçülere getirilen düzeltmeler, analiz sonuçlarını doğrudan olumsuz etkilemektedir. Sağlıklı ölçüm yapan uluslararası GNSS referans istasyonu (EUREF, IGS vb.) verileriyle oluşturulacak yeni modelin, konu kapsamındaki bahsedilen sorunların (kaba ve sistematik ölçüm hatası, baz uzunluğu, ani hava değişimi, yetersiz istasyon sayısı vb.) çözümüne de yarar sağlayacağı tartışılmazdır.

### 1.4. Kaynak Araştırması

GNSS sinyalleri, troposferdeki kütlelerden kaynaklı etkileşim içerisinde olmasıyla beraber, saçılım ya da dağılıma uğradığından doğrusal hareket etmezler. Bu bakımdan sinyaller ya olması gerektiğinden daha fazla yol alabilir ya da daha kısa zamanda alıcıya ulaşırlar. Bu şekilde oluşan hataya troposferik gecikme adı verilir. GNSS gözlemleri için troposferik gecikme hatası çok önemlidir ve ihmal edilmemesi gerekmektedir (Atalay, 2018).

GNSS gözlemlerindeki troposfere dayalı gecikme etkileri, ilk başlarda yalnızca konum doğruluğunu ilgilendiren bir kriter olarak görülmüştür. Ancak, özellikle son zamanlarda GNSS gözlemleri; atmosfer fiziği değişimi, hava ve iklim çalışmalarında ZTD, yoğunlaşabilir su buharı (PWV: Precipitable Water Vapor) ve toplam elektron içeriği (TEC: Total Electron Content) gibi birtakım atmosferik ürünlerin elde edilmesi için bir araç olarak kullanılmaya başlamıştır (Douša ve ark., 2016). Atmosferik ürünlerden ZTD modelleri, hem GNSS ile hassas konum belirlemeye hem de su buharı değişimi ya da

meteorolojik tahmin vb. konulara yardımcı olmuştur. Konu ile ilgili birçok ulusal ve uluslararası çalışmalar gerçekleştirilmiş olup yükseklik açısıyla ilişkili izdüşüm fonksiyonları (MF: Mapping Function) kullanılarak elde edilen ZTD aracılığıyla, atmosferdeki su buharı miktarı mm hassasiyetinde belirlenebilmiştir (Tunalı, 2018). Bu nedenle, yüksek mekânsal-zamansal çözünürlüğe sahip ZTD modellerinin mevcudiyeti, GNSS uygulamalarında doğru ve yüksek duyarlılığa sahip sonuçlar elde etme açısından büyük öneme sahiptir. GNSS gözlemleriyle konum belirleme çalışmalarında, troposferik gecikme hatasının giderilebilmesi için düzeltme değeri hesaplanması gerekir. Bunun için de meteorolojik parametrelere dayalı bir modele ihtiyaç vardır. Literatüre ve yapılan araştırmalara bakıldığında, birçok bilim insanı tarafından standartlaştırılmış troposferik modeller geliştirildiği görülmektedir.

Hopfield (1969) ve Saastamoinen (1972) gibi standart troposfer modelleri, ölçülen meteorolojik parametrelere dayanmaktadır. Fakat bu modeller, meteorolojik verilerin bazı yerlerde veya gerçek zamanlı olarak elde edilme zorluğundan dolayı post-process olarak kullanılmaktadır. Yani hesaplamalar ampirik yöntemlere göre sonradan yapılmaktadır. Diğer yandan, bazı araştırmacı kuruluşlar, University of New Brunswick (UNB), European Geostationary Navigation Overlay System (EGNOS) ve Global Pressure/Temperature (GPT), meteorolojik modeller önermek için meteorolojik parametrelerin zamansal ve mekânsal özelliklerini inceledi (Penna ve ark., 2001; Boehm ve ark., 2007; Leandro ve ark., 2006). Ancak bu modellerde kullanıcı, enterpolasyon ve fonksiyona uydurma yoluyla araştırma alanında karşılık gelen meteorolojik parametreleri elde edebilir. Diğer noktalar için bu meteorolojik parametreler, klasik ZTD modelleriyle birleştirilerek ZTD değeri tahmin edilebilir. Bu yönüyle değerlendirildiğinde, modellerin ayrık kullanılması, hem hesap yükünün artması hem de modeldeki hesap hatalarını artırıcı unsur olarak karşımıza çıkacaktır. Hatta her seferinde ayrı bir modelleme yapılması veri standardını ve sonuç değerleri olumsuz etkileyecektir. Standart troposfer modellerini böyle kullanmak yerine, gerçek istasyon değerleri kullanarak tek modele dayalı şekilde karmaşık problemlerin çözülmesi güçlü bir yaklaşım olarak görülmektedir.

Son yirmi yılda, bilim ve mühendislik alanındaki doğrusal olmayan varyasyonların modellemesi ve tahmin problemlerini çözmek için sinir ağları ya da makine öğrenimi algoritmaları yaygın şekilde uygulanmaya başlamıştır. Çeşitli uygulama alanları arasından, gün geçtikçe GNSS teknolojisine olan talebin artması sonucu, atmosferik çalışmalarda ve konuya ilişkin parametre kestiriminde bu tekniklerin

kullanımının önemi ortaya çıkmıştır (Hadas ve ark., 2017). Ölçme tekniği ve jeodezi uygulamalarında, modellemesi zor üstelik karmaşık yapıdaki atmosfer için sözü edilen öğrenim tekniklerin kullanımı günümüzde daha da hız kazanmaktadır.

Çalışmalarda ilk olarak, atmosferdeki sıcaklık, basınç, kütsel gazlar ve su buharı basıncı gibi karmaşık parametrelerin ya da hava şartlarının tahmini için sinir ağlarına dayalı modellerin oluşturulması üzerine ağırlık verilmiştir (Şenkal ve ark., 2012; Putra ve Lursinsap, 2014; Polyakov ve ark., 2014). Gerçekleştirilen atmosferik çalışmalarda, genellikle geleneksel olarak kullanılan modellerin iyileştirilmesi, global bir model elde etme çabası, anlık veri gereksinimi ortadan kaldırma gibi etkenlere çözüm bulmak amacıyla yöntem arayışı olduğu görülmektedir. Ding (2018), uluslararası makalesinde, GNSS ile ölçülebilen su buharı bilgisi yerine gerçek verilerle daha doğru sonuçların elde edilebildiği çok katmanlı, ileri beslemeli sinir ağı modeli önermiştir. Bunun için, zenit ıslak gecikmenin (ZWD: Zenit Wet Delay), yoğunlaşabilir su buharına dönüştürülmesinde kullanılan ağırlıklı ortalama sıcaklık değişkenini ( $T_m$ ), GNSS gözlemleri ile değil, yeni sinir ağı modeline göre tespit etmiştir. Modeldeki iki farklı  $T_m$  özelliğinin, (mevsimsel değişim/yüzey meteorolojik unsuru) kombinasyonu sağlanarak kullanıcılar için daha doğru bir  $T_m$  modeli elde edilmiştir. Ağ modelini GPT2<sub>w</sub>, GT<sub>m</sub>, GT<sub>m</sub>-I ve PT<sub>m</sub>-I global modellerle karşılaştırmış, sonuç modelin iyi performansından dolayı dönüşümde kullanılabileceğini ifade etmiştir. Los ve ark. (2020), yaptıkları uluslararası çalışmada, Polonya için rastgele orman sınıflama yaklaşımını kullanarak 0-2 saatlik fırtınaları tahmin etmede entegreli su buharı ve tomografi tabanlı düşey profillerin sinerjisini kullanmak için bir metodoloji sunmuştur. Model doğruluğunun %87'nin üzerine çıktığını ve tahmin kesinliğinin yaklaşık %30'a ulaştığını tespit etmişlerdir. Sonuçlar, ıslak kırılma ve entegreli su buharının, fırtına yerini tahmin etme potansiyeli olan önemli parametreler arasında olduğunu göstermiştir. Analiz sonuçlarından ikisi arasında bir korelasyon olduğu sonucuna varılmıştır.

Su buharı basıncının düzensizliği üzerinde de çalışmalar gerçekleştirilmiş olup GPS'den yoğunlaşabilir su buharı ve meteorolojik sensörler kullanılarak saatlik yoğun yağış tahmini çalışmaları gerçekleştirilmiştir (Manandhar ve ark., 2018). Doğrudan yağış tahmini genellikle çok sayıda meteorolojik veri gerektirdiğinden, meteorolojik gözlemlerin eksik olduğu Antarktika, Afrika gibi ıssız bölgelerde, PWV kullanılarak dolaylı şekilde tahmin edilmektedir. Bu konuda Yue ve Ye (2019), çalışmalarında, Antarktika'daki Zhongshan istasyonundan belirlenen GPS analiz sonuçlarından yola

çıkarak dört farklı giriş şeması ile istasyondaki PWV'yi 6 ve 12 saat aralıklar ile tahmin etmek için genetik algoritmali yapay sinir ağı (YSA) kullanmışlardır.

Bir diğer çalışmada, Benevides ve ark. (2019) saatlik yağışları tahmin etmek için GNSS ve meteorolojik verileri işlemek için doğrusal olmayan bir otoregresif eksojen sinir ağı modelini (NARX) uygulamıştır. Bu modelin zamanla değişen PWV'de ve şiddetli yağmuru tespit etmek için kullanılabileceğini bulmuşlardır. Bununla birlikte, diğer sinir ağı modellerine göre daha yeterli bilgi içerdiği, tahmin sonuçlarının %63'ten %72'ye kadar iyileşme gösterdiği saptanmıştır. Bu ve buna benzer çalışmalar, GNSS teknolojisi ile yapay zekâ tekniklerinin birlikte kullanımının, yağış/hava tahminini iyi bir şekilde yorumlayabildiğini göstermektedir (Luo ve ark, 2013; Sapucci, 2014; Barindelli ve ark, 2018). Özellikle Harita Mühendisliği'nde önemli bir yere sahip olan GNSS teknolojisinde gözlemlere dayalı oluşturulabilen yapay zekâ modelleri, atmosferik parametrelerin doğru kestirilmesi için kritik rol oynamaktadır.

Aynı zamanda, sinir ağlarının bu kadar atmosfer/troposfer modellemesinde kullanılması, GNSS ile konum belirleme çalışmaları için önemli hata kaynağı olan troposferik gecikme kestirimi için de yeni modellerin oluşturulmasını tetiklemiştir (Miotti, 2019). Konu üzerinde birçok araştırma yapılarak, oluşturulan modellerin gecikme etkisini açıklama derecesi ve ilişkileri irdelenmiştir.

Katsougiannopoulos ve Pikridas (2009), zenit troposferik gecikmeyi bir, üç ve altı saatlik çeşitli zaman aralıklarında tahmin etmek için sinir ağı algoritmalarını kullanmıştır. Model tahminlerinin neticesinde, gerçek zamanlı ulusal sabit GNSS ağından türetilen ZTD ile sinir ağı modelinden elde edilen ZTD değeri arasında 2-3 cm sapmalar gözlemlenmiştir.

Troposferik ıslak gecikmenin iyileştirilmesi için sinir ağı modeline dayalı enterpolasyon uygulamaları da gerçekleştirilmiş olup, Selbesoğlu (2017) tez çalışmasında, GNSS istasyonları ve Yeni Avusturya Meteorolojik Ölçüm Ağı'ndaki meteorolojik verilerin entegre edilmesi ile YSA'ya dayalı troposferik ıslak gecikmenin enterpolasyonu için doğruluk araştırması gerçekleştirmiştir. Echtzeit Positionierung Austria GNSS Network of Austria (EPOSA) ve The New Austrian Meteorological Measurement Network (TAWES)'ten ilgili parametre verilerini elde etmiş (enlem, boylam, yükseklik ve meteorolojik parametreler) 6 farklı yapay sinir ağı oluşturmuştur. Çalışmada bir başka, lineer enterpolasyon yöntemi (LIN) ile mesafenin tersi ağırlıklı enterpolasyon (IDW) yöntemi de kullanılmıştır. Mesafelerin enterpolasyon yöntemleri üzerindeki etkisini incelemek için, farklı sayıda istasyonlara sahip iki ağ oluşturmuştur.

Yapılan analizler sonucunda, IDW, LIN ve YSA yöntemlerinin ortalama standart sapmaları sırasıyla 12,6 mm, 13,4 mm ve 5,9 mm olarak belirlenmiştir. Buradan şu çıkarım yapılmaktadır. Yeni oluşturulan sinir ağı modeliyle ıslak gecikmenin kestiriminin diğer enterpole yöntemlerine göre daha iyi sonuçlar verdiği görülmüştür. Ayrıca, troposferik gecikme etkisi istasyonlar arasındaki mesafeye bağlı olarak tamamen değişmektedir. Mesafenin artması durumunda troposfer hatasında da artışlar görülmektedir.

Farklı bir çalışmada, uyarlamalı ağ tabanlı bulanık çıkarım sistemi (ANFIS) algoritması, tek bir GNSS istasyonuna dayalı ZTD modelini kurmak için önerilmiştir, burada modelin girdi ve çıktı parametreleri sırasıyla yüzey meteorolojik verileri ve GNSS istasyonundaki tahmini ZTD değerleri olmuştur (Suparta ve Alhasa, 2015).

Buraya kadar bahsedilen mevcut ZTD modellerinin çoğu, sinir ağları algoritmasına dayanmaktadır, ancak sinir ağlarındaki algoritmalar, ileriki zamana ait tahmin kararlılığı açısından zayıf bir performansa ve hatalardan çabuk etkilenen yapıya sahiptir. Bir probleme çözüm ürettiği zaman, bunun neden ve nasıl olduğuna ilişkin bir ipucu vermez deneme-yanılma yoluyla çözer (İnal, 2010). Dolayısıyla bu yöntemler GNSS uygulamaları için mevcut durumda yeterli başarı gösterse de sonraki bir zaman dilimi için kararlı duruş gösteremezler. DVM gibi genelleştirilmiş makine öğrenme algoritmaları ise, daha kararlı ZTD modellemesini sağlamak için güçlü bir matematiksel teoriye sahiptir. Üstelik gecikme tahmininde, robust bir yöntem olarak kullanılabilir.

Xu ve ark. (2020) uluslararası bildirilerinde, geri yayımlı sinir ağına ve en küçük kareler destek vektör makinesine dayalı ZTD tahmin algoritmalarını zaman ve uzay alanında önermişlerdir. ZTD zaman serisindeki ana eğilim terimleri, polinom uydurma ile çıkartarak kalan artıkları, sırasıyla sinir ağı ve destek vektör algoritması aracılığıyla yeniden modellemişlerdir. Test sonuçlarına göre, zaman alanında 20 istasyondaki Uluslararası GNSS Hizmetinin (IGS) ZTD ürünlerini kullanması sonucu destek vektör algoritması performansının tahmin kararlılığı ve doğruluğu açısından sinir ağından daha iyi olduğunu saptamışlardır.

Deng ve ark. (2018) uluslararası makalede, uzun mesafeli az sayıdaki referans istasyonu altında troposferik gecikme enterpolasyonunun düşük kesinliği problemini çözmek için DVM teorisine dayalı bir troposferik gecikme enterpolasyon tahmini yöntemi önermiştir. Çalışmada ilk olarak, troposferin sonsuz ince tek katmanlı bir tabaka olduğu varsayıldı ve referans istasyonlarının bilinen bilgileriyle, tek katmanlı

tabakadaki her bir görünür uydu üzerindeki her bir referans istasyonunun koordinatları ve ZTD elde edildi. Daha sonra, DVM'nin uygun çekirdek parametrelerini optimize etmek ve troposferik gecikmenin DVM modelini oluşturmak için eğitim örnekleri olarak istasyon koordinatları ve zenit troposferik gecikme alındı. Son olarak, enterpolasyonlu troposferik gecikmenin etkisini karşılaştırmak ve analiz etmek için farklı uzunluklara sahip iki ağ gerçek zamanlı kinematik (RTK) verisi seti seçildi. Sonuçlar, uzun mesafe referans istasyonu altında, oluşturulan DVM troposferik gecikme modeliyle, enterpolasyon tahmininin troposferik gecikme doğruluğunun 2 cm'den daha iyi olduğunu, tahmin hatasının kararlı olduğunu ve uydunun enterpolasyon etkisinin olduğunu göstermiştir.

Radhika ve Shashi (2009) uluslararası makalede, hava tahmini için DVM'nin bir uygulamasını sunmuştur. Bir konumdaki günlük maksimum sıcaklığın zaman serisi verileri, girdi sırası olarak adlandırılan, önceki  $n$  günün bir aralığı için günlük maksimum sıcaklıklara dayalı olarak o konumda bir sonraki günün maksimum sıcaklığını tahmin etmek için analiz etmiştir. Sistemin performansı, kernel fonksiyonunun optimum değerleri kullanılarak 2 ila 10 günlük çeşitli aralıklarda gözlemlenmiştir. Doğrusal olmayan regresyon yöntemi, bu uygulama için DVM'yi eğitmek için uygun bulunmuştur. Sonuçlar, geri yayılma algoritmasında eğitilmiş çok katmanlı algılayıcı ile karşılaştırılmış ve DVM'nin performansının sürekli olarak daha iyi olduğu bulunmuştur.

Yine gecikme modelinde olduğu gibi yüksek hassasiyet gerektiren deformasyon çalışmalarında, mühendislik yapılarındaki hızlı hasar tespitini değerlendirmek, zamanında karar vermek, sağlıklarını izlemek ve nihayetinde kamu güvenliği düzeyini artırmak için optimizasyon tabanlı algoritmalar kullanılmış ve istenilen sonuçlar elde edilmiştir (Worden ve Lane, 2001; Kim ve ark., 2012; Hegde ve Rokseth, 2020). Örnek olarak, mühendislerin büyük verileri işlemesine ve hasarı/kusurları etkin bir şekilde tespit etmesine yardımcı olması ve böylece sivil altyapı yönetimi uygulamalarını desteklemek için Gui ve ark. (2017) tarafından üç optimizasyon algoritması tabanlı DVM yaklaşımı sunulmuştur. Grid tarama, kısmi sürü optimizasyonu ve genetik algoritma dâhil olmak üzere algoritmalar, ceza parametrelerini ve Gauss çekirdeği işlevi parametrelerini optimize etmek için kullanıldı. Yöntemleri doğrulamak için literatürdeki 17 farklı senaryo ile bir kıyaslama verisi kullanmışlardır. Sayısal sonuçlar, optimize edilmiş üç makine öğrenimi yönteminin de geleneksel yöntemlere göre duyarlılık, doğruluk ve etkinlik açısından önemli ölçüde gelişme gösterdiğini ortaya koydu.

Genetik algoritma tabanlı DVM, diğer yöntemlerden daha iyi bir öngöründe bulundu. Çalışmanın genel kapsamı, hasar tespiti ve yapısal sağlık koşullarını değerlendirmede duyarlılığı artırmak için optimizasyon yöntemlerinin çok önemli olduğunu göstermiştir. Uluslararası literatürde, GNSS ve troposferik gecikme konusunda makine öğrenme algoritmalarının kullanıldığı diğer farklı çalışmalar Çizelge 1.1’de belirtilmiştir.

**Çizelge 1.1.** Konuya ilişkin yapılan diğer çalışmalar

| ARAŞTIRMA KONUSU                                                                                       | KULLANILAN YÖNTEMLER                                          | KAYNAK               |
|--------------------------------------------------------------------------------------------------------|---------------------------------------------------------------|----------------------|
| Konum belirleme ve tahmin etme                                                                         | Destek Vektör Regresyonu (DVR)                                | Wu ve ark., (2007)   |
| GPS GDOP’un farklı yöntemlerle kestirilmesi                                                            | Doğrusal, hız ve izotonik regresyon; DVM; Genetik Programlama | Wu ve ark., (2010)   |
| GNSS ve ERA5 verisi kullanarak en küçük kareler destek vektör makinesine dayalı bölgesel ZTD kestirimi | En küçük kareler destek vektör makinesi                       | Li ve ark., (2021)   |
| Günlük IGS ZTD’nin tahmini için derin öğrenme modeli                                                   | TensorFlow ve Keras modeli                                    | Osah ve ark., (2021) |
| GPT3 ve YSA’ya Dayalı Bölgesel ZTD kestirimi                                                           | GPT3 ve YSA modeli                                            | Yang ve ark., (2021) |

Yapılan kaynak araştırmalarına göre, bu zamana kadar incelenen birçok çalışmada makine öğrenme yaklaşımlarının ilgili oldukları konu üzerinde gayet etkili ve yüksek başarıya sahip sonuçlar elde ettiği ortaya konulmuştur. Ayrıca bu tür makine öğrenme algoritmalarının daha da iyileştirilmesi için, korelasyon analizine dayalı yeni veri kümesi stratejisi uygulanmıştır. Stratejinin sonucunda veri modeli üzerinde yaklaşık olarak %10-15 oranında iyileşmeler görülebilmektedir (Brownlee, 2016).

Bu bağlamda, tüm çalışmaların genel kapsam çerçevesine bakılacak olursa, troposferik gecikmenin modellenmesinde birçok farklı tekniğin alternatif olarak kullanılması; uydu bazlı sistemlerin konum belirleme doğruluğu (Erkan, 2008), teknik ve teknik olmayan hataların GNSS analizleri üzerindeki etkisinin minimize edilmesi (Penna ve ark., 2001) ve GNSS meteorolojisi alanı (Gökdaş, 2014) için çok büyük avantaj sağlamaktadır. Özellikle farklı yükseklik seviyelerinde, ani hava değişimlerinde veya çevresel olguların göz önüne alınması gerektiği durumlarda troposferik gecikme etkileri daha da belirgin hale gelmektedir. Bu da doğrudan gerçek zamanlı konum

belirleme uygulamalarında, yükseklik bileşeni için “cm” ya da “dm” mertebesinde hatalara neden olabilmektedir (Guo, 2015; Juni ve Rózsa, 2019).

Günümüzde IGS ağlarıyla, her IGS istasyonu tarafından oluşturulan RINEX formatındaki veriler kullanılarak tüm IGS zemin sahalarına özgü, yaklaşık 4 mm’lik bir doğrulukta, yüksek kaliteli, sürekli, günlük zaman serisi ZTD ürünleri sağlanmaktadır (Byun ve ark., 2005). IGS gecikme ürünleri, yüksek kalitesi gerekçesiyle diğer ZTD modellerinin veya ürünlerinin performanslarının doğrulanmasında, hatta araştırma uygulamaları için birçok yerde referans olarak kullanılmıştır (Labib ve ark., 2019).

Ancak bu ağlardan elde edilen analiz sonuçlarında/ürünlerinde bazı dezavantajlar vardır. IGS-ZTD kestirimlerinin kalitesi, çeşitli IGS yer istasyonlarındaki kaliteye ve sürekli günlük veri arşivine bağlıdır. Bununla birlikte, zayıf internet bağlantısı, donanım arızası ve elektrik güç kesikliği gibi bazı sorunlardan dolayı, IGS yer alanlarına erişilemez. Veri akışındaki tutarsızlık nedeniyle veri arşivi ve troposferik gecikme tahmininin kalitesi bozulur (Combrinck ve Chin, 2001). IGS istasyonlarının bazı yerlerde (Afrika ve kutup bölgelerinde) sayı olarak sınırlı olması ya da yetersiz olmasından kaynaklı büyüyen hatalar da çokça görülmektedir (Isioye ve ark., 2015; Erol ve ark., 2020). Referans istasyonlarının birbirleriyle ilişkili olduğu ölçme tekniklerinde, istasyonlar arası mesafeler arttıkça troposferik gecikmenin kestirimindeki hata miktarı artmakta ve buna bağlı olarak konum doğruluğu düşmektedir. Bir başka husus, istasyonlardaki uydu sinyalinin geliş açısına göre farklı yükseklik açılarında da troposferik gecikmenin kuru ve ıslak bileşen üzerindeki etkilerinde ciddi farklılıklar gözlemlenmektedir (Kahveci ve Yıldız, 2017; Nistor ve ark., 2020).

Bu tür sorunlara çözüm getirmesi, istasyonların birbirine olan bağımlılıklarının ortadan kaldırılması ve tek bir öğrenme modeli oluşturulup gerekli analizlerin, oluşturulan modelden hesaplanması gibi etkenler göz önünde bulundurularak, günümüzde mühendislik problemlerinin çözümünde de sıklıkla kullanılan makine öğrenmesi teorisinin yapılan çalışmalarda güçlü bir yaklaşım olduğuna karar verilmiştir. Bundan sonraki yapılacak olan GNSS uygulamalarına da altlık olabilecek nitelik taşıması açısından çalışma büyük önem arz etmektedir.

## 2. GNSS VE TROPOSFERİK GECİKME ETKİSİ

### 2.1. Uydularla Konum Belirleme Sistemleri

GNSS, bir referans alıcının konumunu, dünyanın herhangi bir yerinde tespit etmek için kullanılan, farklı uydu sistemlerinin bir arada bulunduğunu ifade eden genel terimdir. 1980'li yıllarda uydularla konum belirleme için ABD Savunma Dairesi tarafından geliştirilmiş GPS (Global Positioning System) ile başlayan bu serüven, Rusya'nın geliştirdiği Globalnaya Navigatsionnaya Sputnikovaya Sistema (GLONASS), Avrupa'nın Galileo sistemi ve Çin'in Beidou Satellite Navigation System (BDS) ile birlikte hizmet vermeye başlaması sonucu GNSS adını almıştır. GNSS; nokta konumu belirleme, askeri uygulamalar, atmosfer fiziği gibi çalışma konularında hızlı, güvenilir ve ekonomik çözümler bulmasıyla günümüzde en çok tercih edilen ve dünyanın her yerinde yaygın olarak kullanılan bir teknoloji olmaktadır (Li, 2016).

Genel kapsamda, GNSS sistemlerinin her biri, bir yer istasyonları ağı ile birlikte gökyüzünde kepler yasalarına bağımlı bir şekilde çalışan uydu takımlarını kullanır (Bezcioglu, 2018). Böylelikle zaman içerisinde gökyüzündeki uydu sistemleri ile yeryüzündeki istasyonlar arasındaki bağlantının kurulması adına birçok global/yerel ağlar oluşturulmuştur. Bu ağlar çalışma konularına hem anlık hem de post-process olarak araştırma imkânı vermenin yanı sıra; iyonosfer, troposfer düzeltmeleri, saat düzeltmeleri, uydu yörünge efemerisi vb. bilgileri kullanıcılara aktarılmaktadır. Ağların farklı kurum/kuruluşlar tarafından iyileştirilmesi, korunması ve devamlılığı sağlanmaktadır. Bu ağlara global kullanılan IGS ve Türkiye de bölgesel kullanılan post-process ile nokta konumlaya imkan veren Türkiye Ulusal Temel GPS Ağı (TUTGA) ile hem post-process hem de anlık olarak nokta konumlaya imkan veren Türkiye Ulusal Sabit GNSS Ağı-Aktif (TUSAGA-Aktif) ağları örnek olarak verilebilir (Bülbül, 2018).

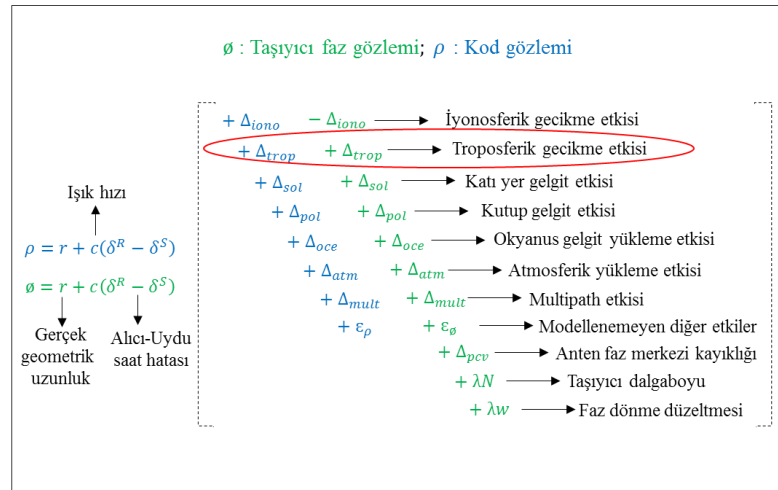
### 2.2. GNSS ile Konum Belirleme Yöntemleri

GNSS uygulamalarında, kod ve faz gözlemleri olmak üzere iki farklı ölçüm yöntemi kullanılmaktadır. Kod ölçülerine dayalı gözlemlerden düşük doğruluklu navigasyon bilgisine sahip olunurken, taşıyıcı faz ölçülerine dayalı gözlemler ile yüksek hassasiyetli bir konum bilgisi elde edilmektedir. GNSS tekniği, kod gözlemlerini sırasıyla, sivil kod ve hassas kod olarak adlandırılan C/A ve P-Code olmak üzere iki

farklı kod frekansı üzerinden, taşıyıcı faz gözlemlerini ise L1, L2 ve L5 frekanslarından gerçekleştirmeye olanak sağlar (Mirsa ve Enge, 2012).

Mutlak konum belirleme kapsamındaki kod gözlemlerine dayalı ölçüm yönteminde, uyduların bilinen koordinatları ve kod gözlemlerinden hesaplanan alıcı-uydu uzaklığı ile alıcının konumu hesaplanır. Konum doğruluğu; uydu geometrisi, uydu-alıcı saati ile GPS zamanı senkronizasyonuna ve kullanılan kod bilgisine bağlı olarak değişmektedir. Rölatif konum belirleme kapsamındaki kod ve faz gözlemlerine dayalı ölçüm yönteminde ise, alıcının uydulardan aldığı veriler ve koordinatları bilinen referans istasyonlarından aldığı düzeltmeler kullanılarak alıcı için final koordinatları belirlenmektedir (Kahveci ve Yıldız, 2017). Ayrıca bu yöntemde birden fazla alıcı kullanılarak eş zamanlı gözlem yapılmaktadır. Bu yönüyle, anında konum belirleme amacı için yeterli doğruluğa sahip ve pratiklik açısından büyük önem taşımaktadır. Ancak, kod ölçülerinden elde edilen hassasiyet, klasik uygulamalar için yeterli olup daha yüksek hassasiyet gerektiren mühendislik uygulamaları için yeterli olmamaktadır.

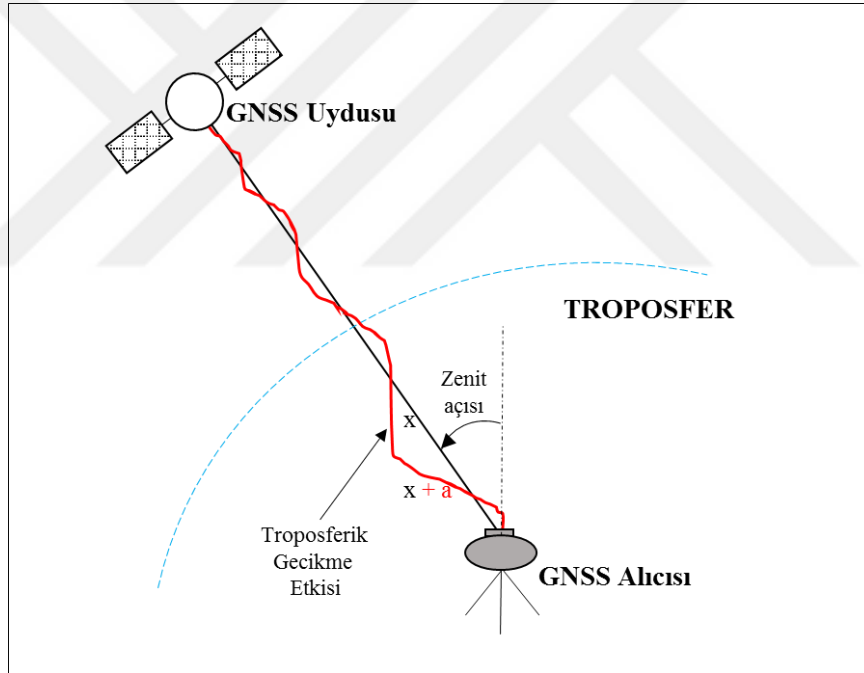
GNSS gözlemleri, doğası gereği alıcıdan, uydudan ve çeşitli ortamlardan kaynaklı hatalar ile yüklüdür. Bu hataların ortadan kaldırılması veya en aza indirilmesi için hatanın yönünün ve büyüklüğünün belirlenmesi önemlidir. Şekil 2.1’de görüldüğü gibi GNSS’deki kod ve taşıyıcı faz gözlemlerinde kullanılan düzeltme denklemi gösterilmiştir (Kouba ve Héroux, 2001). Taşıyıcı dalga fazı iyonosferden geçerken hızlandığından faz gözlemlerine getirilecek düzeltme (-) işaretlidir. Kod ölçüleri ise yavaşladığından iyonosfer düzeltmesi kod gözlemleri için (+) işaretlidir. Burada GNSS gözlemleri için önemli bir hata kaynağı olan troposferik gecikme etkisi, tezin kapsamını ve içeriğini oluşturmaktadır.



Şekil 2.1. GNSS’deki taşıyıcı faz ve kod gözlemleri için düzeltme denklemleri

### 2.3. Troposferik Gecikme Etkisi

Çok katmanlı atmosferde gazlar ve su buharıyla en yoğun tabaka olan troposfer, çoğu navigasyon, meteorolojik çalışmalar, radyo sinyalleri, uydu ve haberleşme sistemleri gibi uygulamalarda etkin rol oynamaktadır. Harita mühendisliği çalışma alanlarının ana noktasında bulunan uydu navigasyon sistemleri ile konum belirlemede, GNSS uyduları ve yayınladıkları sinyaller oldukça önemli bir unsurdur. GNSS tarafından yayınlanan sinyaller uygulamada direkt ışın içerisinde olduğundan sinyalin yayılma şekli de önem kazanmaktadır (Şekil 2.2). Sinyaller, troposferdeki kütlelerle etkileşim içerisinde olmasıyla beraber saçılım ya da dağılıma uğradığından doğrusal hareket etmezler. Sonucunda sinyaller ya olması gerektiğinden daha fazla yol alabilir ya da daha kısa zamanda alıcıya ulaşırlar (Niell, 1996).

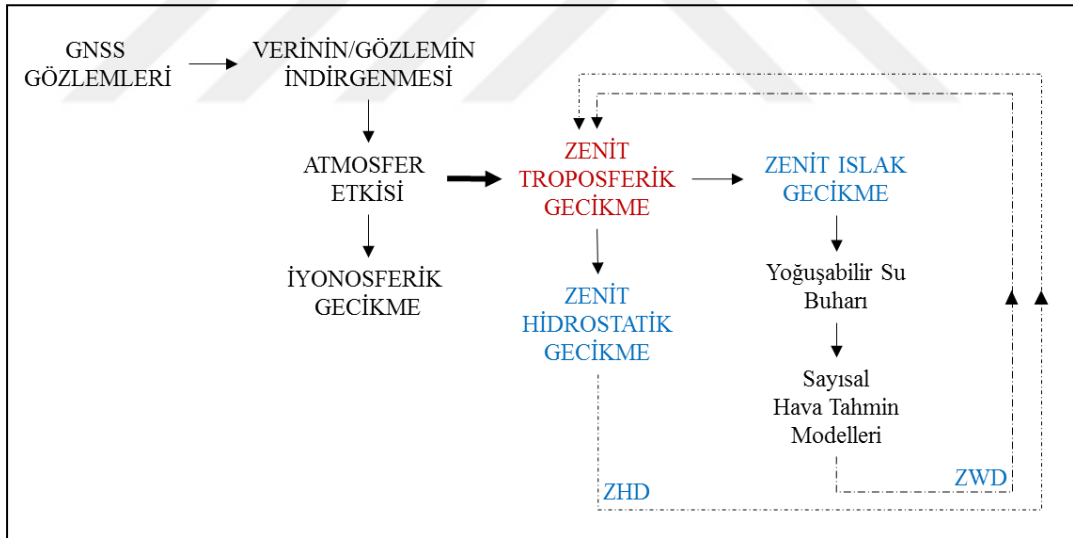


Şekil 2.2. Troposfer katmanındaki troposferik gecikme etkisi

GNSS gözlemleri için önemli hata kaynağı olan troposfer gecikmesinin modellenmesi, gerçek zamanlı ve post-process uygulamalarda oldukça karmaşık hal almaktadır. Karmaşıklığın nedeni, troposfer tabakasındaki ani su buharı değişimleri, istasyonların konumu ve gözlem yapılan zamanda havadaki gaz kütlelerinin değişiklik göstermesidir. Birçok çalışma troposferik gecikme etkisinin kolay incelenmesi ve modellenmesi amacıyla gecikmeyi hidrostatik ve ıslak bileşenden oluşacak şekilde iki kısma ayırmıştır.

Burada hidrostatik bileşen kısmı basınç ölçümleriyle hassas olarak belirlenmektedir. Ancak ıslak bileşen kısmı ıslak kırılma kriterinin zamansal ve mekânsal değişkenliğinden tam olarak modellenememektedir (Hoffmann-Wellenhof ve ark., 2000). Yüksek doğrulukla post-process GNSS verisi değerlendirmede, gerçek zamanlı kinematik (RTK) konum belirleme uygulamalarında, gerçek meteorolojik gözlemler kullanarak kuru ve ıslak bileşenin doğru modellenmesi kritik olacaktır. Bu nedenle, gecikme hatasını minimize etmek için öncül kuru bileşen değerinin doğru modellenmesi ve ıslak bileşenin gerçeğe yakın değerleri kestirebilmesi önemli bir konudur.

GNSS gözlem tekniği, sadece konu kapsamında ölçülere getirilen troposferik gecikme hata düzeltmesi olarak değil özellikle son zamanlarda atmosfer fiziği değişimi, hava ve iklim çalışmalarında entegreli yoğunlaşabilir su buharı (IPWV: Integrated Precipitable Water Vapor), sayısal hava tahmin modelleri, sıcaklık ve basınç gibi birtakım atmosferik ürünlerin elde edilmesi içinde önemli bir araç olarak kullanılmaya başlamıştır (Şekil 2.3).



Şekil 2.3. GNSS gözlemlerinden atmosferik ürünlerin elde edilmesi

## 2.4. Troposferik Gecikmenin GNSS Gözlemlerinde Modellenmesi

GNSS ve VLBI (Very Long Baseline Interferometry) gibi mikrodalga sinyalleri, troposfer olarak da adlandırılan elektriksel anlamda tarafsız atmosferdeki hava yoğunluğu nedeniyle gecikmeye (kırılmaya) uğrar. Böylelikle troposfer etkisiyle sinyaller üzerindeki elektromanyetik dalgaların hızı ve yönünde değişimler meydana

gelir. Bu olumsuz etki, iyonosferdeki gibi lineer kombinasyonlar ile çözülememektedir. Çünkü troposfer, GNSS sinyalleri için frekanstan bağımsız (L1 ve L2 frekans GNSS taşıyıcıları için aynı) ve yaklaşık 30-35 GHz'e kadar elektromanyetik dalgalarda dağıtıcı özellikte olmayan bir ortamdır. Bu nedenle alıcı ve frekans sayısından bağımsız, ayrı şekilde troposferin modellenmesi gerekmektedir. GNSS çözümlemelerinde genel kullanıma tabii olan troposferik model eşitliği aşağıdaki gibi ifade edilir.

$$\Delta_{trop} = ZHD * mf_h(ele) + ZWD * mf_w(ele) + [G_N * \cos(\alpha) + G_E * \sin(\alpha)] \quad (2.1)$$

Burada eşitlik içerisinde geçen ifadelerden;

$ZHD$  = Zenit Hidrostatik Gecikmeyi (Zenith Hydrostatic Delay),

$ZWD$  = Zenit Islak Gecikmeyi (Zenith Wet Delay),

$mf_h(ele)$  ve  $mf_w(ele)$  = Hidrostatik ve ıslak gecikme için izdüşüm fonksiyonları (MF),

$G_N$  ve  $G_E$  = Yatay yöndeki gradyent gecikmeleri (Horizontal Gradient Delay),

$ele$  ve  $\alpha$  ise sırasıyla, yükseklik ve azimut değerlerini göstermektedir.

GNSS gözlemlerinde troposferden kaynaklı gecikmeleri yüksek doğrulukta belirlemek için iki farklı yöntem uygulanabilir. Birinci yöntemde; yeterli sayıda yatay ve düşey çözünürlüğe sahip troposferik kırılma profillerine ihtiyaç vardır. Bu profillerin elde edilebilmesi için troposfer tahmin modelleri (GPT/GPT2: Global Pressure ve Temperature, VMF1/VMF3: Vienna Mapping Function, NMF: Niell Mapping Function), su buharı radyometreleri (WVR: Water Vapor Radiometers) sayısal hava tahminleri (NWP: Numerical Weather Prediction) gibi modeller kullanılarak ışıın boyunca entegrasyon ya da radyosonde gözlemleri tercih edilmektedir.

Eğer bu modeller kullanılamıyor ya da verileri toplamıyorsa ikinci bir yöntem olarak, PPP (Precise Point Positioning) ve GNSS ağı çözümlerinde  $ZHD$  parametresi, oturumun yapıldığı istasyonlarda hassas basınç ölçümleri ( $p$ ), enlem ( $\phi$ ) ve yüksekliğin ( $h$ ) bir fonksiyonu olarak Saastamoinen (1972) gibi analitik modellere dayalı olarak yüksek hassasiyetle belirlenir. Belirlenen parametre, çözüm denkleminde öncül (a priori) troposfer gecikme değeri olarak girmektedir. Modellenmesi güç olan  $ZWD$  parametresi ise, zamansal ve mekânsal değişkenliği yüksek olan su buharı ile çok yakın ilişkisinden kaynaklı en küçük kareler parametre kestirimi sonucunda bulunur (Tunalı, 2018). Sonrasında hesaplanan öncül değere, kestirilen  $ZWD$  parametresi düzeltme olarak

getirilir. Sinyallerin gözlem doğrultusu boyunca toplam gecikme değerleri, hidrostatik ve ıslak gecikmenin zenit doğrultusunda kestirilmesi ve ilgili izdüşüm fonksiyonları ile çarpılması sonucu tespit edilir (Selbesoğlu, 2017). Bunun yanında, troposfer tabaka kalınlığının her noktada farklı olmasından dolayı gradyent gecikmeleri, (2.1) eşitliğine hesaba dâhil edilir.

#### 2.4.1. Zenit troposferik gecikme modelleri

Geçmişten bugüne kadar, troposferde seyir halinde olan radyo sinyalleri üzerindeki gecikme etkisinin tespiti için çok sayıda atmosferik model ortaya atılmıştır. Bu modellerden en yaygın olarak kullanılanı, Hopfield (1969), Saastamoinen (1972), Davis ve ark. (1985), Askne ve Nordius (1987), Berrada-Baby ve ark. (1988) tarafından öne sürülen troposfer modelleridir. Bu modeller troposferik gecikmeyi hesaplarken yalnızca istasyon konumu ve yüzeyde ölçülen atmosfer değerlerini kullanmışlardır. Hatta hidrostatik ve ıslak olarak ayrı ayrı incelemeyi başarmışlardır (Kleijer, 2004). Günümüzde ulusal ve uluslararası yapılan çalışmalarda, ticari veya akademik yazılımlarda sıkça karşılaşılan sayısal troposfer modellerinin yanı sıra Saastamoinen, Hopfield gibi deneysel modellerde pratik bir kullanımı olması sebebiyle çoğu zaman tercih edilmektedir.

##### 2.4.1.1. Zenit hidrostatik gecikme

Zenit Hidrostatik Gecikme (ZHD), yansız hidrostatik atmosferin neden olduğu zirveden (zenit) gelen bir radyo sinyalinin gecikmesidir. Bu bileşende, yüzey basıncı, enlem ve ortalama deniz seviyesinin üzerindeki yüksekliğe dayalı olarak gecikme hesaplanır. ZHD, eşitlik (2.2) integraliyle tanımlanmaktadır:

$$D_h^z = 10^{-6} \int_{h_0}^{\infty} N_h(h) dh \quad (2.2)$$

Burada  $D_h^z$ : zenit hidrostatik gecikmeyi (m),  $N_h$ : hidrostatik kırılma katsayısını,  $h$ : ortalama deniz seviyesindeki yüksekliği (m) ve  $h_0$ : ortalama deniz seviyesindeki

yüzey ölçüm yüksekliğini (m) ifade eder. (2.2) eşitliği atmosferdeki kütlelere bağlı daha da genişletilecek olursa hidrostatik gecikmenin yeni formu aşağıdaki gibi ifade edilir.

$$D_h^z = 10^{-6} k_1 R_d \int_{h_0}^{\infty} \rho_m(h) dh \quad (2.3)$$

$k_1$ : ampirik kırılma sabiti,  $\rho_m$ : kuru hava ve su buharının toplam yoğunlukları ve  $R_d$ : evrensel gaz sabitini belirtmektedir. Hidrostatik denge koşulu altında hidrostatik bileşen denkliği (2.4) ile kurulur.

$$\frac{dP}{dh} = -\rho_m(h)g(h) \quad (2.4)$$

Burada formüle yeni dâhil olan  $g(h)$ , yüksekliğin bir fonksiyonu olarak yerçekimine bağlı gravite etkisidir. Denge koşulu altında kurulan (2.4) eşitliğinin integrali alınmasıyla, (2.5) eşitliğinden hava basıncının ( $P_0$ ) yüzey değeri elde edilir.

$$\int_{P_0}^0 dP = - \int_{h_0}^{\infty} \rho_m(h)g(h)dh = -P_0 \quad (2.5)$$

Yüksekliğe bağlı yerçekimindeki gravite değeri “ $g(h)$ ” yerine, ortalama gravite değeri kullanılabilir. Bu nedenle ortalama gravite değeri aşağıdaki gibi tanımlanmaktadır.

$$g_m^0 = \frac{\int_{h_0}^{\infty} \rho_m(h) g(h) dh}{\int_{h_0}^{\infty} \rho_m(h) dh} \quad (2.6)$$

(2.3) ve (2.5) eşitliğinin, (2.6) eşitliği ile kullanılması sonucu  $P_0$ 'ın yüzey değerine göre zenit hidrostatik gecikmesi (2.7) ve (2.8) eşitliğinde verilmiştir.

$$D_h^z = 10^{-6} k_1 \frac{R_d}{g_m} P_0 \quad (2.7)$$

$$f(\varphi, h) = 1 - 0,00266 * \cos(2\varphi) - 0,28 * 10^{-6} h \approx 1 \quad (2.8)$$

Saastamoinen ve Davis'e göre;  $g_m = g_m^0 * f(\varphi, h)$  formülü ile tespit edilir. Burada  $\varphi$  ve  $h$ , istasyonun jeodezik enlemi ve ortometrik (veya elipsoidal) yüksekliği olup,  $g_m^0$  (9,784) ortalama gravite değeridir. (2.7) eşitliğindeki  $g_m$ 'nin, (2.8) eşitliğine göre düzenlenmesiyle zenit hidrostatik gecikme eşitliği aşağıdaki hale gelir.

$$ZHD = 10^{-6} k_1 \frac{R_d}{g_m^0} \frac{P_0}{f(\varphi, h)} \quad (2.9)$$

$$k_1 = 77,604 \text{ K mbar}^{-1}$$

$$R_d = 8,31434 \text{ kg m}^2 \text{ s}^{-2} \text{ mol}^{-1} \text{ K}^{-1}$$

$$\rho_m = 28,9644 * 10^{-3} \text{ kg mol}^{-1}$$

$$g_m^0 = 9,784 \text{ m s}^{-2}$$

Tüm  $k_1$ ,  $R_d$ ,  $\rho_m$  ve  $g_m^0$  değerlerinin, (2.9) eşitliğinde yerine konulması sonucu, metre cinsindeki zenit hidrostatik gecikme, aşağıdaki son halini alır.

$$ZHD = (0,0022768 \text{ (m/mbar)}) \frac{P_0}{f(\varphi, h)} \quad (2.10)$$

Burada  $P_0$ , hPa (hectopascal) biriminde olup tüm troposfer gecikme modelleri için belirli bir yüksekliğe (90~100 km) kadar aynıdır. Zenit hidrostatik gecikmesini belirlerken hatalar genelde,  $k_1$  ampirik kırılma sabiti ve yüzey basınç ölçümlerinden kaynaklanmaktadır. Ölçme tekniğinde basınç değerinin elde edilmesi için birkaç yöntem mevcuttur. Bu yöntemlerden ilki, istasyonun olduğu konumda meteorolojik istasyondan basınç değerlerinin kaydedilmesidir. Eğer meteorolojik istasyon mevcut değilse, numerik hava modelleri veya ampirik modeller kullanılarak basınç değerleri elde edilir. Kullanılan bu atmosfer modellerinde genellikle, deniz düzeyindeki yüzey değerleri tanımlanmaktadır. Bu değerlerin yükseklikle doğrusal değiştiği kabul edilerek, GNSS gözlemi yapılan nokta yüksekliğindeki atmosferik parametre değerleri hesaplanmakta olup bu değerler zamandan ve gerçek hava koşullarından bağımsızdır. Bu yönüyle modeller gerçek değere ulaşma açısından dezavantaj sağlamaktadır.

#### 2.4.1.2. Zenit ıslak gecikme

Zenit Islak Gecikme (ZWD), ilgili istasyondaki düşey profil boyunca ıslak kırılmanın integrali şekilde (2.11) eşitliğindeki gibi tanımlanır.

$$D_w^z = ZWD = 10^{-6} \int_{h_0}^{\infty} N_w(h) dh = 10^{-6} \int_{h_0}^{\infty} (k_2' \frac{e}{T} + k_3 \frac{e}{T^2}) dh \quad (2.11)$$

Bu eşitlikteki,  $N_w$ : ıslak kırılma katsayısı,  $e$ : su buharı basıncı (hPa),  $T$ : sıcaklık (K),  $k_2'$ :  $22,1 \text{ K hPa}^{-1}$  ve  $k_3$ :  $373 900 \text{ K}^2 \text{ hPa}^{-1}$  olarak tanımlanmıştır (Bevis ve ark., 1994). Buradan ZWD'nin buhar basıncı ve sıcaklıkla değiştiği kolaylıkla görülebilir.

Su buharı basıncı ve sıcaklık profilleri kullanarak ışın boyunca entegrasyon ya da radyosonde gözlemlerinden doğru ZWD değerleri elde edilmektedir. Ancak pratik uygulamalarda (örneğin, uçak navigasyon ve konumlandırma, geniş alan güçlendirme sistemi: wide area augmentation), meteorolojik veri profillerinin mevcut olmaması nedeniyle ıslak gecikmenin elde edilebilmesi için günümüze kadar birçok ampirik ve analitik model ortaya konulmuştur (Mendes, 1999). Çoğu model, yüzey ZWD, sıcaklık ve su buharı basıncı arasındaki ilişkiyi dikkate alan ampirik bir yaklaşım kullanır (Ifadis, 1992). Ancak ZWD, ZHD'den farklı olarak, atmosferdeki su buharı içeriğinin yüksek mekânsal ve zamansal değişkenliği nedeniyle bu tür modellerin doğruluğunu sınırlamaktadır. Şimdiye kadar, en doğru ve yaygın olarak kullanılan ZWD modeli, Askne ve Nordius (1987) tarafından ortaya atılan modeldir. Model atmosferdeki düşey nem dağılımını hesaba katar ve yükseklik arttıkça su buharı basıncının katlanarak azaldığını varsayar (Dousa ve Vaclavovic, 2014). Yine de, ortaya koyulan modeller yeterli doğrulukta başarı sağlayamamaktadırlar. Bu sebeple ıslak gecikme, GNSS ölçülerinin değerlendirilmesinde bilinmeyen olarak ele alınıp, en küçük kareler yöntemiyle kestirilmektedir.

#### 2.4.1.3. Troposferik izdüşüm fonksiyonları

Troposferik izdüşüm fonksiyonları (MF), farklı yükseklik açılarındaki zenit gecikme değerlerinin elde edilmesinde kullanılır. Herhangi bir uydudan gelen sinyale etki eden troposferik gecikme miktarının zenit doğrultusundaki gecikme miktarına

indirgenmesini sağlarlar. Bu zamana kadar farklı araştırmacılar tarafından çok sayıda izdüşüm fonksiyonu yayımlanmıştır. Ancak burada tezin kapsamını oluşturan VMF izdüşüm fonksiyon modeli üzerinde daha detaylı durulacaktır.

Niell (1996) izdüşüm fonksiyonunu (NMF) yüzey meteorolojik parametreleri yerine istasyonların koordinatlarını ve zamansal değişiklikleri kullanarak geliştirmiştir. Kullanılan fonksiyonlar hidrostatik ve ıslak bileşenlerdir. Bu fonksiyonun katsayıları olan  $a$ ,  $b$  ve  $c$ ; ABD Standart atmosferinin sıcaklık ve bağıl nem profillerine (26 farklı radyosonde istasyonu verisi kullanılmıştır) dayalı olarak  $3^\circ$  sinyal yüksekliğine kadar türetilmiştir. Bu katsayıların hesapları, kuzey yarımkürede  $15^\circ$ ,  $30^\circ$ ,  $45^\circ$ ,  $60^\circ$ ,  $75^\circ$  enlemlerinde Ocak ve Haziran aylarında gerçekleştirilmiştir. Yüksekliğe bağlı etkinin düzenli olmamasından dolayı ıslak kısım için hesaplara yükseklik düzeltmesi getirilmemektedir.

Son otuz yıldır numerik hava modelleri (NWM) verilerine dayalı farklı izdüşüm fonksiyonları geliştirilmiştir. Bu izdüşüm fonksiyonlarından olan izobarik izdüşüm fonksiyonu (IMF) ile Niell tarafından devam eden kesir formunun yeni katsayılarını elde etmek için 6 saatlik çözünürlüğe sahip bir numerik hava modeli ortaya konmuştur. Hidrostatik izdüşüm fonksiyonunun atmosferik kalınlığına dayalı olarak, sabit basınç seviyelerinin jeopotansiyel yüksekliği ile olası IMF katsayılarını deneysel yolla incelemişlerdir. IMF katsayılarının hesabında, istasyon enlem ve yüksekliğinin yanı sıra, 200 hPa basınç seviyesinin jeopotansiyel yüksekliğini ve ıslak troposferik gecikmenin  $3.3^\circ$  yükseklik açısındaki değeri ile zenit doğrultusundaki gecikme değerinin oranlarını kullanmaktadır. Islak IMF izdüşüm fonksiyonunun hesaplanmasında bazı pratik ve kavramsal kısıtlamalar (düşük düşey çözünürlük) Vienna izdüşüm fonksiyonu (VMF) ile giderilmiştir (Boehm ve Schuh, 2004).

VMF, azimuta değil, sadece yükseklik açısına bağlıdır. Yani, troposferin istasyonlar etrafında simetrik olduğunu varsayarlar. VMF fonksiyonunun  $a$  katsayısı, ışıın boyunca entegrasyon verilerinden belirlenmiş,  $b$  ve  $c$  katsayıları ise o sırada mevcut olan en iyi değerler hidrostatik kısım için izobarik izdüşüm fonksiyonlarından (IMF) ve ıslak kısım için  $45^\circ$  enlemde Niell izdüşüm fonksiyonlarından (NMF) alınmıştır (Boehm ve ark., 2006b). Simetrik olarak hesaplanan  $b$  ve  $c$  katsayıları, Antarktika ve ekvatorial bölgelerde istasyon yükseklik bilgisinde bozulmalara neden olmuştur. Bu nedenle hidrostatik izdüşüm fonksiyonları için yeni  $b$  ve  $c$  katsayılarına dayanan ve bundan sonra VMF1 olarak adlandırılacak olan güncellenmiş bir versiyon geliştirilmiştir.

VMF1, küresel olarak mevcut olan troposfer gecikmeleri ve uzay jeodezik gözlemlerinin tüm tarihi için en doğru izdüşüm fonksiyonudur. Ancak dezavantajı diğer izdüşüm fonksiyonlarına nazaran hesaplamalardaki yük fazlalığıdır. Bu nedenle, VMF1 ile tutarlı, ancak uygulanması basitleştirilmiş, derecesi ve genliği 9'a kadar olan küresel harmonik açılımına dayanan bir Global izdüşüm fonksiyonu (GMF) geliştirilmiştir.

#### 2.4.1.3.1. VMF1 troposferik izdüşüm fonksiyonu

Avrupa Orta Menzilli Hava Tahminleri Merkezi (ECMWF: European Centre for Medium-Range Weather Forecasts) analizleri sonucunda hesaplanan ve 2.5° (boylam) x 2.0° (enlem) grid formatında günde 4 kez yayınlanan (00:00, 06:00, 12:00, 18:00 UTC) VMF1; içeriğinde kuru ve ıslak izdüşüm fonksiyonlarının parametreleri ile verilen grid noktaları için hesaplanmış ZHD ve ZWD değerlerini bulundurmaktadır.

Bu modelde b ve c katsayıları hesaplanırken, ilk olarak 10 farklı yükseklik açısında ( $e_0 = 3,2^\circ - 5^\circ - 7^\circ - 10^\circ - 15^\circ - 20^\circ - 30^\circ - 50^\circ - 70^\circ - 90^\circ$ ) 7488 profil (12 ay \* günde 4 kere \* 156 grid noktası) için toplam ve hidrostatik izdüşüm fonksiyonu belirlenmiştir (Boehm ve ark., 2006a). Geometrik gecikme etkisi ( $\Delta L_{gec}$ ), (2.12) ve (2.13) eşitliklerindeki hidrostatik ve toplam izdüşüm fonksiyonlarına eklenmiştir.

$$mf_h(e) = \frac{\Delta L_h(e) + \Delta L_{gec}(e)}{\Delta L_h^z} \quad (2.12)$$

$$mf_t(e) = \frac{\Delta L_h(e) + \Delta L_w(e) + \Delta L_{gec}(e)}{\Delta L_h^z + \Delta L_w^z} \quad (2.13)$$

Daha sonra, toplam ve hidrostatik haritalama fonksiyonlarının üç katsayısı; a, b ve c, en küçük kareler kestirimi ile her bir profilin 10 ayrı izdüşüm fonksiyonu değerine uydurulur. Uyumlarının artıkları genellikle 0,5 mm'den daha küçüktür. Hem toplam hem de hidrostatik izdüşüm fonksiyonu için bütün b katsayılarının ortalama değerleri,  $b_t = b_h = 0,0029$  bulunur ve tüm analizlerde sabit tutulur. Bir sonraki adımda en küçük kareler kestirimi tekrarlanır, ancak tüm profiller için sadece a ve c katsayıları tahmin edilirken hidrostatik ve toplam izdüşüm fonksiyonları için b katsayısı  $b_t = b_h$  ortalama değere sabit tutulur. Daha sonra c katsayıları, mevsime ve enleme bağlı olarak açık bir değişkenlik gösterir, ancak önceki izdüşüm fonksiyonlarının (eski VMF) aksine, c

katsayıları ekvatora göre simetrik değildir. Bu nedenle  $c$  katsayısını modellemek için (2.14) eşitliği kullanılır.

$$c = c_0 + \left[ \left( \cos \left( \frac{\text{doy} - 28}{365} * 2\pi + \Psi \right) + 1 \right) * \frac{c_{11}}{2} + c_{10} \right] * (1 - \cos \varphi) \quad (2.14)$$

Çizelge 2.1, hidrostatik ve toplam izdüşüm fonksiyonları için elde edilen eşitlik (2.14)'deki parametreleri içerir. Islak izdüşüm fonksiyonlarının  $b$  ve  $c$  katsayılarının değiştirilmesi gerekmez çünkü ıslak zenit gecikmeleri hidrostatik zenit gecikmelerinden ~10 kat daha küçüktür ve  $b$  ve  $c$ 'deki varyasyonların etkisi önemli değildir. Islak izdüşüm fonksiyonunun  $b$  ve  $c$  katsayıları  $45^\circ$  enlemdeki NMF katsayılarına hala sabittir ve her profil için  $a$  katsayısı tahmin edilir; yani, önerilen ıslak izdüşüm fonksiyonu hala VMF'dir.

**Çizelge 2.1.** Hidrostatik ve toplam izdüşüm fonksiyonları için kullanılan parametreler

| <i>Yarım Küre</i>   | $c_0$ | $c_{10}$ | $c_{11}$ | $\Psi$ |
|---------------------|-------|----------|----------|--------|
| Kuzey (toplam)      | 0,062 | 0,000    | 0,006    | 0      |
| Güney (toplam)      | 0,062 | 0,001    | 0,006    | $\pi$  |
| Kuzey (hidrostatik) | 0,063 | 0,000    | 0,004    | 0      |
| Güney (hidrostatik) | 0,063 | 0,001    | 0,006    | $\pi$  |

Aşağıda Çizelge 2.2'de,  $5^\circ$  yükseklik açısında farklı izdüşüm fonksiyonlarına ait ortalama bias ve standart sapma değerleri verilmiştir. Parantez içindeki değerler ise farklı modellerde eşdeğer istasyon yüksekliğindeki sapmaları ifade etmektedir.

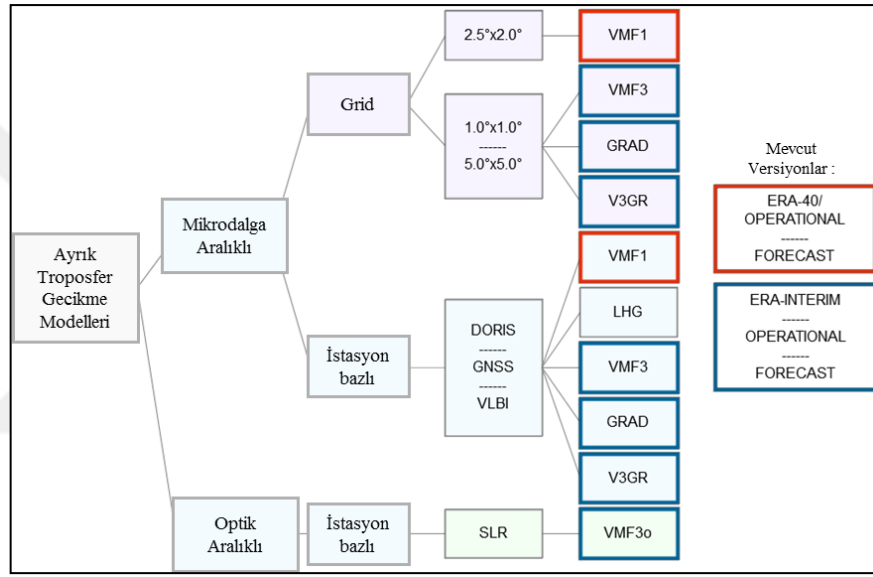
**Çizelge 2.2.** Farklı izdüşüm fonksiyonlarına ait hata değerleri ( $e_0 = 5^\circ$ )

|             | <b>Ortalama Bias (mm)</b> | <b>Standart Sapma (mm)</b> |
|-------------|---------------------------|----------------------------|
| <b>NMF</b>  | 21,8 (4,4)                | 35,0 (7,0)                 |
| <b>VMF</b>  | 3,3 (0,7)                 | 11,2 (2,2)                 |
| <b>VMF1</b> | 2,3 (0,5)                 | 2,0 (0,4)                  |

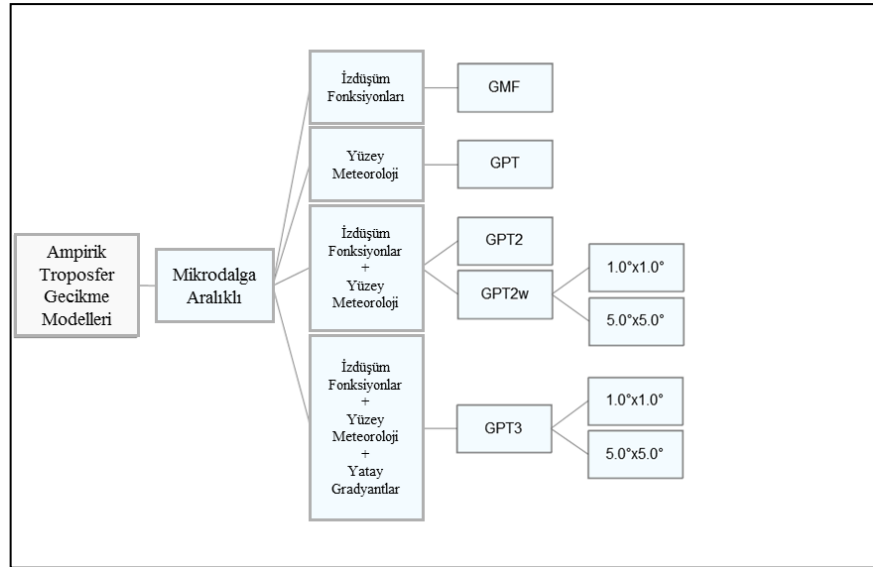
Buradan da anlaşılacağı üzere VMF1 modelinin hem gecikme değerini belirlemede en yüksek doğruluğa sahip olması hem de atmosferik parametreleri yerinde

gözlemler veya gridler ile iyi derecede tespit etmesi çalışma kapsamında uygulanacak olan yeni model için nokta atışı bir tercih olarak görülebilir.

Günümüzde (2021 itibariyle) ECMWF tarafından oluşturulan VMF veri sunucusunda sağlanan ürünler “troposferik gecikme modelleri” ayırık ve ampirik olarak 2’ye ayrılmıştır (Şekil 2.4 ve 2.5). Ayırık modeller için, parametreler her dönem için belirlenir (00, 06, 12, 18, günlük) ve ardından, gözlem zamanı ve aletin konumu için enterpolasyon yapılması gerekir. Ampirik modeller, uzun vadeli ortalamalara dayanmaktadır. Tüm parametreler, herhangi bir rastgele zaman ve konum için model katsayılarından hesaplanabilir (VMF Data Server, 2021).



Şekil 2.4. Ayırık troposferik gecikme modelleri



Şekil 2.5. Ampirik troposferik gecikme modelleri

### 3. MATERYAL VE YÖNTEM

Tez kapsamındaki temel aşamalar, problemin tanımını ortaya koyma, uygulama konusu ve çalışmanın amacı, kaynak araştırması, çalışma bölgesinin belirlenmesi, verinin düzenlenmesi ve değerlendirilmesi, uygulama için veri setlerinin oluşturulması, öğrenme modellerinin oluşturulması, troposferik gecikme modellerine ait sonuç çıktıları ve model sonuçlarının değerlendirilmesidir (Şekil 3.1). Bu çalışmada, GNSS uygulamalarındaki önemli hata kaynağı olan troposferik gecikme hatası için son zamanlarda sıkça karşımıza çıkan ve gelecekte de önemli bir yere sahip olacak olan makine öğrenimi algoritmaları ile troposferik gecikme tahmin modeli oluşturulup hata kestirimi yapılmaktadır. Problemin tanımı, tezin konusu ve amacı önceki bölümlerde belirtilmiş olup detaylı olarak açıklanmaktadır.

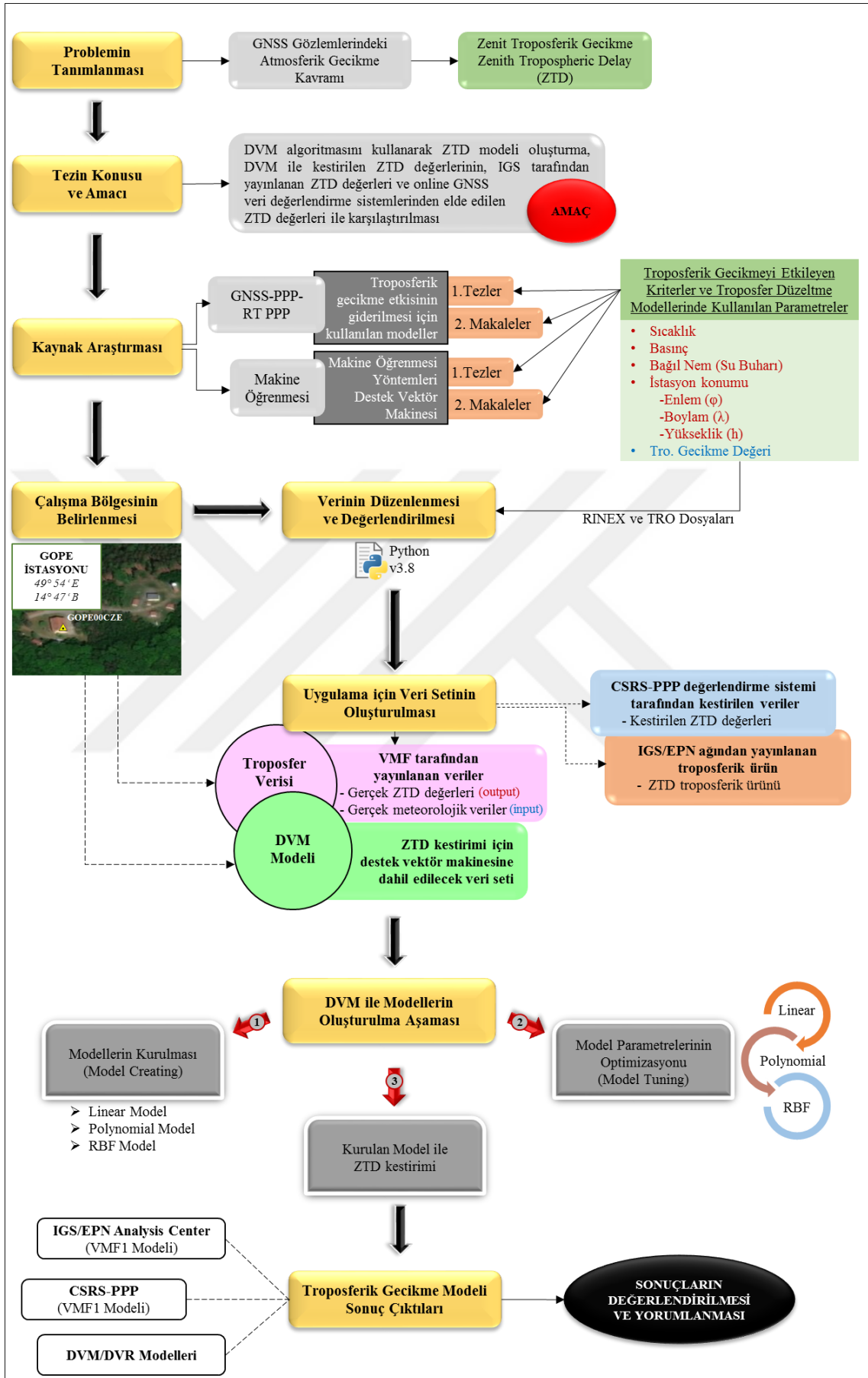
Geçmişten günümüze troposferik gecikme hatasının giderilmesi için çok sayıda bilim insanı çeşitli atmosferik modeller üretmişler ve ürettikleri modelleri geliştirmeye çalışmışlardır. Bu modeller çoğunlukla sıcaklık, basınç, bağıl nem, istasyon konumu gibi kriterleri kullanarak hata hesaplaması yapmaktadır. Ancak Hopfield, Saastamoinen gibi bazı deneysel troposferik gecikme modelleri, mevsimsel kriterleri ortalama değer şeklinde kullanmaktadır. Ya da VMF, GMF, NMF gibi ayrık modeller, hesap yükünün fazla olması nedeniyle mevsimsel ortalama değerlere göre belirlenmiş gridler kullandığı için istasyona (konuma) bağlı ani hava değişikliklerini hesaba dâhil edememektedir. Anlık değişen atmosferde hata giderilmesi amacıyla modellerde gerçek meteorolojik verilerin kullanılması gerekliliği vardır. Bu gereklilik için çalışmada, zenit troposferik gecikmeyi etkileyen kriterler ve troposferik düzeltme modellerinde kullanılan parametreler irdelenerek kullanılacak olan kriterler belirlenmiştir. Yöntemin uygulanmasında makine öğrenimi yaklaşımlarından olan destek vektör makinesi/regresyonu (DVM/DVR) tercih edilmiştir.

Destek vektör regresyonuyla zenit troposferik gecikmenin kestirilebilmesi için istasyona ait verilere ihtiyaç duyulmaktadır. Veriler, bağımsız değişken için çalışma bölgesi olarak belirlenmiş olan GOPE istasyonundaki meteorolojik veriler ve bağımlı değişken için VMF Veri Merkezi'nde istasyona ait 6 saat aralıklarla belirlenen gerçek gecikme analizi sonuçlarıdır. Oluşturulacak olan modeller için veriler, açık kaynak kodlu Python v3.8 yazılım programlamada düzenlenip istatistiki anlamda değerlendirilerek veri setleri oluşturulmuştur. Veri setlerinde troposferik gecikmeyi etkileyen tüm kriterler modele dâhil edilmiştir. Ek olarak veri setindeki örneklem

grupları arasında aykırı verilerin sonuç üzerindeki etkisini görmek amacıyla Çok Değişkenli Aykırı Gözlem Analizi (LOF: Local Outlier Factor) tekniği kullanılmıştır. Değişkenler arasındaki aykırılığı açıklamada kullanılan analiz türünde, bir noktanın yerel yoğunluğu ile komşu noktaların yoğunluğu karşılaştırılır. Seyrek veya sık görülme durumuna göre verinin uyumlu/uyumsuz olması yorumlanır. Böylelikle belirli bir gün içerisinde kullanılan meteorolojik verilerde olmaması gereken bir durum söz konusu olduğunda, bu veri aykırı olarak seçilip veri setinden çıkartılır. Tüm veriler kullanılarak ve uyumsuz verilerden arındırılmış olarak iki farklı veri seti modelleme için hazır hale getirilmiştir.

Çalışmanın model oluşturma aşamasında ise öncelikle modelleri farklı fonksiyonlarla kurma işlemi tamamlanmış ardından kurulan modelin optimizasyonu ve optimize edilmiş sonuç modelden troposferik gecikme kestirimi gerçekleştirilmiştir. Modellerden elde edilen sonuçların kıyaslanması ve troposferik gecikme etkisinin farklı kestirim algoritmalarıyla değerlendirilmesi açısından IGS/EPN ağı ürünü ve bir diğer web tabanlı yazılım olan CSRS-PPP değerlendirme servisi kullanılmıştır. CSRS-PPP’de veriler değerlendirildikten sonra analiz raporunda çalışmanın konusunu oluşturan zenit troposferik gecikme etkileri incelenmiştir.

DVM/DVR modellerinden elde edilen gecikme değerleri, IGS/EPN istasyonları tarafından çözümleme sonucunda yayınlanan gecikme ürünleri ve online PPP servisi kapsamındaki CSRS-PPP’de kestirilen gecikme değerleriyle karşılaştırılmıştır. Üç farklı kestirim yöntemiyle elde edilen değerlere ait istatistiksel sonuçlar değerlendirilip yorumlanmıştır. DVM/DVR modellerinin kendi içerisindeki model başarısı analiz edilirken Ortalama Mutlak Hata (OMH), Ortalama Mutlak Yüzde Hata (OMYH), Karesel Ortalama Hata (KOH) ve  $R^2$  değeri hesaplamalarına dayalı performans ölçütleri kullanılmıştır.

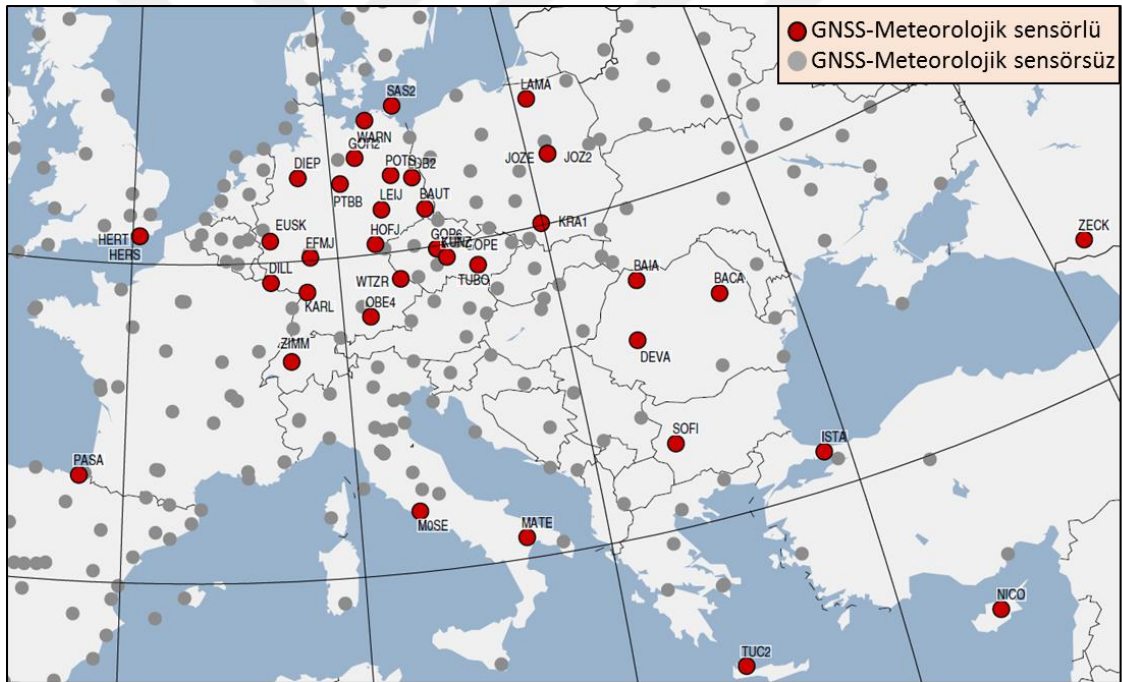


Şekil 3.1. Çalışmadaki iş akışı

### 3.1. Materyal

#### 3.1.1. Çalışma bölgesinin belirlenmesi

Çalışma kapsamında gerçek meteorolojik veriler ile zenit troposferik gecikme kestirimi için, 355'i aktif 7'si aktif olmayan toplam 362 adet GNSS referans istasyonuna sahip ve Avrupa'yı kaplayan EUREF (European Terrestrial Reference System) GNSS ağındaki, EPN (EUREF Permanent Network) istasyon verileri incelenmiştir. Sabit GNSS ağındaki yaklaşık 46 adet EPN istasyonunda GNSS ve meteorolojik sensörlü donanımlar entegreli olarak çalışıp gerçek meteorolojik parametreler üretmektedir. Ağdaki GNSS-Meteorolojik sensörlerin birlikte bulunduğu istasyonlara ait konumsal dağılım aşağıdaki şekilde görülmektedir (Şekil 3.2).



Şekil 3.2. EUREF ağında GNSS-Meteorolojik sensörlerin entegre çalıştığı istasyonlar

EPN istasyonlarının orta kuşak enlemlerinde diğer bölgelere kıyasla daha çok dağılım sergilediği açık bir şekilde görülmektedir. Diğer taraftan duruma Türkiye sınırları içerisinde bakıldığında bazı kurum/kuruluşlar tarafından tesis edilen ve sürekli ölçü toplayan birçok GNSS ağı mevcut olmasına karşın GNSS ile meteorolojik sensörün entegre çalıştığı tek bir uluslararası referans istasyonu bulunmaktadır (ISTA00TUR: Referans istasyonu). Hava ile ilgili araştırma konularının önem arz ettiği

günümüzde Türkiye için bu istasyon sayısının yetersiz kaldığı ortaya çıkmaktadır. Dolayısıyla atmosfer araştırmaları ve projelerinin çoğunlukla meteorolojik istasyonların sık bulunduğu bölgelerde yapıldığı gözlemlenmektedir.

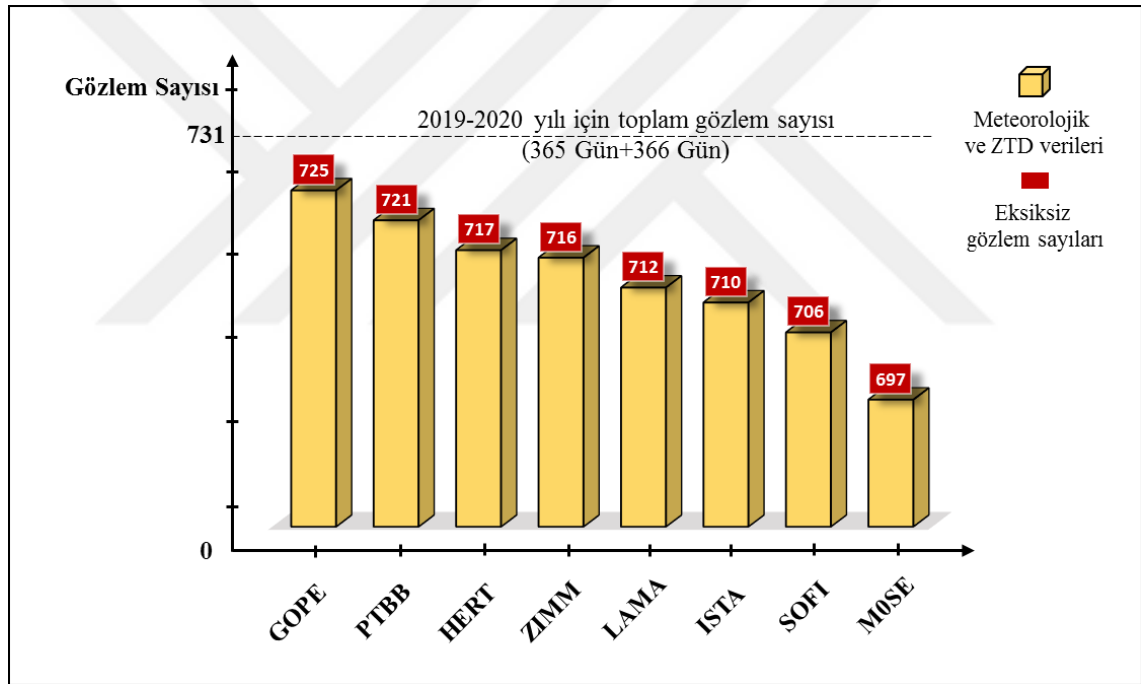
Uygulamada 2019-2020 yılları arasındaki günleri kapsayan gerçek meteorolojik veriler ile modeller oluşturulacak olup zenit troposferik gecikme kestirimi yapılacaktır. Regresyon modellerinin yüksek tahmin başarısı için yeterli veri sayısının sağlanması, hatasız ve doğru parametrelerin matematiksel modelde kullanılması gibi etkenler dikkate alınarak, çok genel anlamda modelde kullanılacak istasyonun diğer istasyonları açıklamada da yeterli olacağı düşünülen IGS/EPN istasyon seçimi planlanmıştır. Referans istasyonlarının IGS/EPN şeklinde ifade edilmesinin nedeni, çalışma bölgesi belirlemede kullanılan istasyonların her iki GNSS ağına da aktif olarak kullanılıyor olmasıdır.

Kestirim modeli için ağ üzerinde farklı bölgelerde yer alan referans istasyonları (GOPE00CZE, PTBB00DEU, HERT00GBR, ZIMM00CHE, LAMA00POL, ISTA00TUR, SOFI00BGR, M0SE00ITA) incelenmiş ve referans istasyonlarına ait araştırma bulguları Çizelge 3.1’de verilmiştir. 2019-2020 yılı olmak üzere iki yılı kapsayan veri kümesinde istasyonlardaki eksik meteorolojik gözlem sayıları azdan çoğa doğru sıralanmıştır.

**Çizelge 3.1.** Farklı bölgedeki referans istasyonlarına ait konum ve eksik meteorolojik veri bilgisi

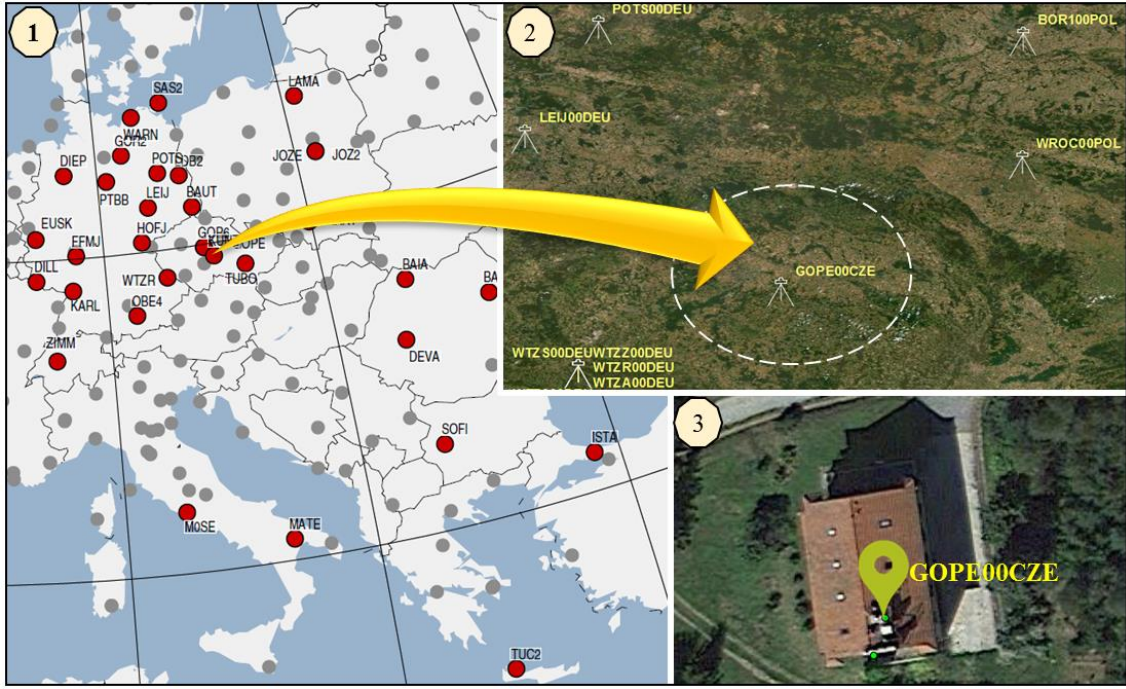
| NO | ÜLKE<br>KOD/ŞEHİR                                                                 | İSTASYON<br>ADI | ENLEM<br>(°) | BOYLAM<br>(°) | YÜKSEKLİK<br>(m) | Eksik<br>Meteorolojik<br>Gözlem (Gün)<br>Sayısı |
|----|-----------------------------------------------------------------------------------|-----------------|--------------|---------------|------------------|-------------------------------------------------|
| 1  | <b>Çek Cumhuriyeti</b><br>CZE, Ondrejov                                           | GOPE00CZE       | 49.9137      | 14.7856       | 592.6            | 6                                               |
| 2  | <b>Almanya</b><br>DEU, Braunschweig                                               | PTBB00DEU       | 52.2962      | 10.4597       | 130.2            | 10                                              |
| 3  | <b>Büyük Britanya<br/>ve Kuzey İrlanda<br/>Birleşik Krallığı</b><br>GBR, Hailsham | HERT00GBR       | 50.8675      | 0.3344        | 83.3             | 14                                              |
| 4  | <b>İsviçre</b><br>CHE, Zimmerwald                                                 | ZIMM00CHE       | 46.8771      | 7.4653        | 956.4            | 15                                              |
| 5  | <b>Polonya</b><br>POL, Olsztyn                                                    | LAMA00POL       | 53.8924      | 20.6699       | 186.7            | 19                                              |
| 6  | <b>Türkiye</b><br>TUR, İstanbul                                                   | ISTA00TUR       | 41.1044      | 29.0193       | 147.2            | 21                                              |
| 7  | <b>Bulgaristan</b><br>BGR, Sofia                                                  | SOFI00BGR       | 42.5561      | 23.3947       | 1119.6           | 25                                              |
| 8  | <b>İtalya</b><br>ITA, Roma                                                        | M0SE00ITA       | 41.8931      | 12.4933       | 120.6            | 30+                                             |

Modele dâhil edilecek veri setinin modeli açıklama başarısının yüksek veya düşük olması veri hacminin yeterli sayıda olmasına bağlıdır. Ayrıca model içerisinde yeterli hacimde veri kümesi olmasına rağmen veri çeşitliliği bulunmuyorsa yine yöntem başarısız kalacaktır. Çünkü makine öğrenmesi algoritmaları kendilerine gösterilen örnek veri grubunu inceleyerek bu veri grubunun meydana geliş biçimlerini öğrenir ve algoritmada gösterilmeyen veri grubunun tahmininde genelleme yapma yeteneği kazanır. Bu nedenle veri uzunluğunun (gün sayısı) yeterli olması, ölçüm hatası olmayan istasyonun tercih edilmesi, web tabanlı GNSS veri değerlendirme servisinde ilgili istasyonun çözülebilmesi konuları dikkate alınmıştır. Gözlem sayılarına bakıldığında en fazla meteorolojik veriye sahip olanı GOPE00CZE istasyonu, en az meteorolojik veriye sahip olanı ise M0SE00ITA referans istasyonu olarak tespit edilmiştir (Şekil 3.3).



Şekil 3.3. EPN/IGS referans istasyonlarındaki meteorolojik gözlem (gün) sayıları

Nihai referans istasyonu seçimi için yukarıda bahsedilen gerekçeler göz önünde bulundurularak araştırma ve incelemeler sonucunda “GOPE00CZE” referans istasyon olarak tercih edilmiştir (Şekil 3.4). Ayrıca EUREF ağında bulunan sağlam bir nokta olması, üzerinde anlık meteorolojik veri toplayacak sistemin bulunması ve toplanan verilerin GMF, GPT gibi atmosferik veri kaynaklarında kabul görmesi, istasyon seçiminde etken olmaktadır. Noktaya etki eden atmosferik koşulların ne olduğu konusunda fikir elde etmek açısından uygulama önem arz etmektedir.

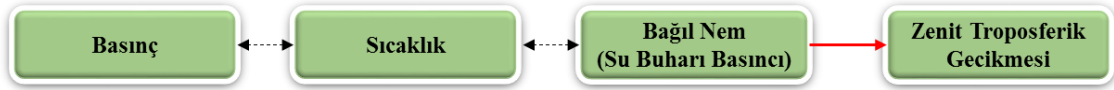


Şekil 3.4. Çalışmada kullanılan GOPE referans istasyonu ( $\varphi = 49^{\circ}54'$ ,  $\lambda = 14^{\circ}47'$ )

### 3.1.2. Verilerin elde edilmesi ve düzenlenmesi

Zenit troposferik gecikmesi ile ilgili çalışmalar ve araştırmalarda farklı atmosferik modeller için ele alınan farklı kriterler ve düzeltme parametre katsayıları mevcuttur. Bu kısım tamamen çalışmalarda kullanılan yöntemdeki matematiksel denkleme bağlı kalmaktadır. Gecikme etkisinin tespitinde farklı modele göre değişkenlik gösteren düzeltme katsayıları, kimi zaman ampirik denklemlere göre belirlenirken, kimi zamanda farklı metotlarla (ışın izleme, radyosonde vb.) dönemsel olarak belirli zaman aralıklarına göre belirlenmektedir.

Çalışmada ilgili modellemeler için kullanılacak olan ana materyaller/veriler istasyona ait gerçek meteorolojik veriler ve gözlemlerle elde edilen 6 saat aralıklı (00:00, 06:00, 12:00, 18:00, günlük) gerçek zenit troposferik gecikme değerleridir. Hem IGS'in gecikme değerini elde etmede kullandığı troposfer modelinin "VMF1 modeli" olması hem de online GNSS veri değerlendirme servisi CSRS-PPP yazılımında veri çözümleme aşamasında, gecikme değeri VMF modeline göre kestirilmesinden dolayı modellemede kullanılacak olan kriterler ve parametre katsayıları; VMF1'de olduğu gibi basınç, sıcaklık, bağıl nem (su buharı basıncı) şeklinde sıralanmıştır (Şekil 3.5).



**Şekil 3.5.** Çalışmadaki troposferik gecikmeye etki eden kriterler

2019-2020 yılları arasında, ilk günü 01 Ocak 2019 (01.01.2019) olmak üzere 31 Aralık 2020 (31.12.2020) tarihine kadar tüm ardışık günlere ait veriler ilgili istasyon için temin edilmiştir. Burada verilerin iki yıl içerisinde ardışık günleri içermesi, troposferik gecikmenin hem mevsimsel etkilere göre modellenmesini hem de son zamanlarda küresel ısınma etkisiyle ortaya çıkan ani hava değişimlerinin makine öğrenme yöntemlerine öğretilerek anlık olayların analiz edilmesi açısından önemlidir.

Tez kapsamında makine öğrenme tekniği için kullanılacak olan meteorolojik veriler, EUREF GNSS ağındaki EPN istasyonlarına ait ham veri ve analiz sonuçlarının paylaşıldığı internet sitesinden alınmıştır. Gerçek zenit troposferik gecikme değerleri ise, IGS ilişkili VMF analiz merkezlerinden (Vienna Mapping Functions Open Access Data) elde edilen analiz sonuçlarının sonradan paylaşıldığı (yaklaşık 2-3 hafta sonrası) internet sitesinden elde edilmiştir. Ayrıca hassas nokta konumlandırma servisi CSRS-PPP'ye gönderilen verilerin çözülmesi adına IGS tarafından yayınlanan “.o uzantılı” gözlem dosyaları indirilip troposfer rapor sonuçlarının elde edilmesi için hazırlanmıştır. Uygulamada kullanılmak üzere işlenen verilere ilişkin örnek dosyalar aşağıda gösterilmiştir (Şekil 3.6).

(a) IGS Data Analysis' deki  
ZTD troposferik ürünü

DESCRIPTION CODE, Astronomical Institute, University of Bern  
OUTPUT CODE IGS 3-day solution  
CONTACT code@aiub.unibe.ch  
SOFTWARE Bernese GNSS Software Version 5.3  
HARDWARE UBELIX: Linux, x86\_64  
INPUT CODE IGS 1-day solutions  
FILE/REFERENCE

\*TROP/DESCRIPTION  
\*KEYWORD VALUE(S)  
ELEVATION CUTOFF ANGLE 3  
SAMPLING INTERVAL 180  
SAMPLING TROP 7200  
TROP MAPPING FUNCTION WET VMF  
SOLUTION\_FIELDS\_1 TROTOT STDDEV TGNTOT STDDEV TGETOT  
\*TROP/DESCRIPTION

GOPE 19:001:03600 2269.2 0.4 -0.223 0.042 -0  
GOPE 19:001:10800 2261.2 0.3 -0.267 0.039 -0  
GOPE 19:001:18000 2247.6 0.4 -0.310 0.035 -0  
GOPE 19:001:25200 2235.7 0.4 -0.354 0.032 -0  
GOPE 19:001:32400 2235.9 0.3 -0.397 0.030 -0

GOPE 2.11 METEOROLOGICAL DATA  
GOPE METRNX RIGTC, GO PECNY 20190102 001514 UTC  
GOPE METEOROLOGICAL DATA FROM STATION GOPE (RIGTC - GO PECNY)  
GOPE EPOCHS AFTER RECOMPUTING FROM UTC TO GPS TIME SCALE  
GOPE GOPE00CZE  
GOPE 11502M002

GOPE 4 PR TD HR  
GRAS COMET SYSTEM, CZ T7410  
GRAS COMET SYSTEM, CZ NS520  
GRAS COMET SYSTEM, CZ HG421  
GRAS METEOSERVIS, CZ MR3H  
GRAS 3979316.8350 1050313.7  
GRAS 3979316.9620 1050313.4  
GRAS 3979316.8970 1050313.5  
GRAS 3979310.8410 1050332.4  
GRAS  
GRAS

19 1 1 0 0 18 962.4  
19 1 1 0 1 18 962.5  
19 1 1 0 2 18 962.4  
19 1 1 0 3 18 962.4  
19 1 1 0 4 18 962.4  
19 1 1 0 5 18 962.4  
19 1 1 0 6 18 962.4  
19 1 1 0 7 18 962.4  
19 1 1 0 8 18 962.4  
19 1 1 0 9 18 962.4  
19 1 1 0 10 18 962.4  
19 1 1 0 11 18 962.4  
19 1 1 0 12 18 962.3  
19 1 1 0 13 18 962.3  
19 1 1 0 14 18 962.4  
19 1 1 0 15 18 962.4  
19 1 1 0 16 18 962.4

(c) Modelde kestirimin için gerekli  
meteorolojik veriler ve VMF1 gerekç  
troposferik gecikme deęerleri

\*TROP/DESCRIPTION  
\*KEYWORD  
ELEVATION CUTOFF ANGLE 7.5  
SAMPLING INTERVAL 30  
TROP MAPPING FUNCTION VMF1  
SOLUTION\_FIELDS\_1 TROTOT STDDEV TGNTOT STDDEV TGETOT  
\*TROP/DESCRIPTION  
\*TROP/STA\_COORDINATES  
\*SITE PT SOLN T STA X STA Y STA Z SYSTEM REMRK  
GOPE A 1 P 3979315.901 1050312.708 4857067.240 IGS14 CSRS-PPP  
\*TROP/STATION  
\*SITE EPOCH TROTOT STDDEV TGNTOT STDDEV TGETOT STDDEV  
GOPE 19:001:00000 2195.5 77.0 11.0 -0.3 1.2 -0.2 1.0  
GOPE 19:001:00030 2195.5 77.0 11.0 -0.3 1.2 -0.2 1.0  
GOPE 19:001:00060 2196.0 77.1 10.9 -0.3 1.2 -0.2 1.0  
GOPE 19:001:00090 2195.9 77.0 10.7 -0.3 1.2 -0.2 1.0  
GOPE 19:001:00120 2195.9 77.0 10.6 -0.3 1.2 -0.2 1.0  
GOPE 19:001:00150 2196.0 77.0 10.4 -0.3 1.2 -0.2 1.0  
GOPE 19:001:00180 2196.0 77.0 10.3 -0.3 1.2 -0.2 1.0  
GOPE 19:001:00210 2195.9 76.9 10.2 -0.3 1.2 -0.2 1.0  
GOPE 19:001:00240 2195.9 76.9 10.1 -0.3 1.2 -0.2 1.0  
GOPE 19:001:00270 2195.9 76.9 10.0 -0.3 1.2 -0.2 1.0  
GOPE 19:001:00300 2195.9 76.9 9.9 -0.3 1.2 -0.2 1.0

(b) CSRS-PPP' de kestirilen  
troposferik gecikme deęeri

RINEX VERSION / TYPE 3 1.2 -0.2 1.0  
PGM / RUN BY / DATE 3 1.2 -0.2 1.0  
COMMENT 3 1.2 -0.2 1.0  
COMMENT 3 1.2 -0.2 1.0  
MARKER NAME 3 1.2 -0.2 1.0  
MARKER NUMBER 3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2 1.0  
3 1.2 -0.2

Şekil 3.6. Uygulama için elde edilen verilere ait örnek dosyalar

Tüm veriler ilgili yerlerden temin edildikten sonra verinin düzenlenmesi işlemine geçilmiştir. Burada verinin bütünleştirilmesi, standartlaştırılması ve en önemlisi gerçek gözlemlerin 6'şar saat aralıklı belirlendiği diğer zaman aralıklarının interpolasyon ile bulunduğu ayrık troposfer gecikme modellerinde (VMF1, VMF3, GRAD) olduğu gibi tercih edilen modelin de dönemsel olarak oluşturulmasına karar verilmiştir. Yani gün içerisinde "Gece 00:00, Sabah 06:00, Öğle 12:00 ve Akşam 18:00" şeklinde sıcaklık değişiminin yaşandığı 4 farklı zaman dilimine ait kestirim yapılması düşünülmüştür.

Ancak ilgili kurum/kuruluşların bünyesinden toplanan istasyonlara ait bilgilerin günlük (daily) ya da saatlik (hourly) şeklinde ve farklı epok aralığında yayınlanması söz konusudur. Örneğin EPN istasyonundan alınan meteorolojik veriler gün içerisinde 1'er

dakika aralıklarla elde edilmiş gözlemleri içermektedir. VMF veri merkezinden elde edilen gerçek zenit troposferik gecikme değerleri gün içerisinde 6 saat aralıklı gözlemler olarak yayınlanmaktadır. Benzer şekilde IGS tarafından troposferik gecikme değerleri 2'şer saat aralıklarla ve CSRS-PPP'de ise gecikme değerleri 30 sn gözlem aralığıyla elde edilmektedir. Bu nedenle verilerin çalışmada belirlenen zaman aralıklarına göre belirli bir formatta dosyalardan çekilmesini gerekli kılmaktadır. Çözüm için açık kaynak kodlu yazılımlardan birisi olan PYTHON v3.8 programlama dili kullanılmıştır. İlgili sitelerden indirilen tüm veriler farklı klasörler altında gruplandırılmıştır. Örneğin meteorolojik veriler için “2019\_2020\_STATION\_METEO” klasörü, gerçek troposferik gecikme değerleri için “2019\_2020\_STATION\_ZTD” klasörü şeklindedir. Gruplanan klasörler yazılıma tanıtılarak klasör içindeki her bir veri dosyası gün sırasına göre okutulmuştur. PYTHON programı yardımıyla okutulan her bir veri dosyasında 00:00, 06:00, 12:00 ve 18:00 saatlerindeki değerleri oluşturan bölümler, doğrudan çıktı dosyası olarak dışarıya çekilmiştir. Çekilen çıktı dosyaları liste haline dönüştürülmüştür. Çizelge 3.2’de istasyon adı, troposfer modeli, indirilen verilerin elde ediliş biçimi ve düzenleme sonucu verinin kullanılacağı epok aralığı (veri kayıt aralığı) ifade edilmiştir.

**Çizelge 3.2.** Verilerin elde ediliş biçimi ve kullanılma şekli hakkında genel bilgi

| <b>Veriler</b><br>2019/2020 yılı<br>Toplam: 725 gün |                                       | <b>İstasyon<br/>Adı</b> | <b>Troposfer<br/>Modeli</b> | <b>Elde edildiği<br/>yer</b> | <b>İndirilen<br/>dosyada<br/>epok aralığı</b> | <b>Düzenlenen<br/>dosyada<br/>epok aralığı</b> |
|-----------------------------------------------------|---------------------------------------|-------------------------|-----------------------------|------------------------------|-----------------------------------------------|------------------------------------------------|
| Model için                                          | Meteorolojik<br>Veriler               | GOPE                    |                             | EUREF<br>GNSS Ağı            | 1 dk                                          | 6 Saat                                         |
|                                                     | Zenit Troposferik<br>Gecikme Verileri | GOPE                    | VMF1                        | VMF Veri<br>Merkezi          | 6 saat                                        | 6 Saat                                         |
| IGS’den elde edilen<br>Gecikme Verileri             |                                       | GOPE                    | VMF1                        | IGS Ağı                      | 2 saat                                        | 6 Saat                                         |
| CSRS-PPP’den elde<br>edilen Gecikme Verileri        |                                       | GOPE                    | VMF1                        | CSRS-PPP<br>servisi          | 30 sn                                         | 6 Saat                                         |

### 3.1.3. Düzenleme sonucu elde edilen veri setleri

Bu çalışmada oluşturulacak olan troposfer modeli için veriler; meteorolojik veriler ve gerçek zenit troposferik gecikme değerleri olmak üzere iki şekilde ele alınmış ve önceki bölümde anlatıldığı düzende hazırlanmıştır. Meteorolojik veriler, “bir

istasyon veya gözlem noktasında belirli bir zaman diliminde ölçülebilen dönemsel özellikler” olarak ifade edilirken, bu verilerin sürekli olarak bir periyot halinde takip edilerek kayıt altına alınması sonucu elde edilen veri topluluğuna meteorolojik veri seti denilmektedir. Bir diğer veri olan gerçek zenit troposferik gecikme değerlerinde de söz konusu durum benzerlik göstermektedir.

Meteorolojik verilerde; basınç değişkeni ölçüleri “**milibar (mbar)**” biriminde, sıcaklık değişkeni ölçüleri “**Kelvin (K)**” biriminde ve su buharı basıncı değişkeni ölçüleri “**g/m<sup>3</sup>’ün yüzdesel (%)**” biriminde hesaba katılırken, troposfere ait gecikme değerleri ise “**milimetre (mm)**” biriminde kullanılmıştır. Çizelge 3.3’de verilerin düzenlenmesi sonucu elde edilen veri setine ait bir kısım tablo halinde verilmiştir.

**Çizelge 3.3.** Düzenleme sonucu elde edilen veri seti (GOPE, 00:00 için)

| YIL | İNDEKS | GÜN | SAAT  | ZTD    | $a_h$<br>Katsayı | $a_w$<br>Katsayı | BASINÇ | SICAKLIK | BAĞIL<br>NEM |
|-----|--------|-----|-------|--------|------------------|------------------|--------|----------|--------------|
| 19  | 1      | 1   | 00:00 | 2269,2 | 0,00121          | 0,00053          | 962,4  | 276,05   | 97,3         |
| 19  | 2      | 2   | 00:00 | 2209,4 | 0,00119          | 0,00042          | 949,7  | 273,95   | 76,5         |
| 19  | 3      | 3   | 00:00 | 2212,3 | 0,00118          | 0,00041          | 960,1  | 269,35   | 67,7         |
| 19  | 4      | 4   | 00:00 | 2222,4 | 0,00119          | 0,00047          | 963,0  | 268,15   | 90,6         |
| 19  | 5      | 5   | 00:00 | 2253,4 | 0,00120          | 0,00055          | 955,3  | 272,85   | 96,7         |
| 19  | 6      | 6   | 00:00 | 2219,5 | 0,00119          | 0,00047          | 954,0  | 274,15   | 97,4         |
| 19  | 7      | 7   | 00:00 | 2243,5 | 0,00119          | 0,00057          | 960,8  | 271,35   | 91,0         |
| 19  | 8      | 8   | 00:00 | 2246,8 | 0,00120          | 0,00052          | 953,1  | 274,05   | 97,2         |
| 19  | 9      | 9   | 00:00 | 2187,6 | 0,00119          | 0,00050          | 935,4  | 273,85   | 95,2         |
| ... | ...    | ... | ...   | ...    | ...              | ...              | ...    | ...      | ...          |
| ... | ...    | ... | ...   | ...    | ...              | ...              | ...    | ...      | ...          |
| ... | ...    | ... | ...   | ...    | ...              | ...              | ...    | ...      | ...          |
| ... | ...    | ... | ...   | ...    | ...              | ...              | ...    | ...      | ...          |
| ... | ...    | ... | ...   | ...    | ...              | ...              | ...    | ...      | ...          |
| 20  | 717    | 23  | 00:00 | 2298,3 | 0,00121          | 0,00055          | 951,2  | 282,25   | 98,4         |
| 20  | 718    | 24  | 00:00 | 2258,7 | 0,00121          | 0,00054          | 945,4  | 282,15   | 97,6         |
| 20  | 719    | 25  | 00:00 | 2218,5 | 0,00119          | 0,00053          | 945,2  | 274,65   | 90,7         |
| 20  | 720    | 26  | 00:00 | 2204,4 | 0,00118          | 0,00046          | 954,3  | 271,75   | 84,5         |
| 20  | 721    | 27  | 00:00 | 2194,8 | 0,00119          | 0,00056          | 951,9  | 270,15   | 74,9         |
| 20  | 722    | 28  | 00:00 | 2147,9 | 0,00120          | 0,00049          | 924,4  | 272,95   | 84,5         |
| 20  | 723    | 29  | 00:00 | 2153,7 | 0,00119          | 0,00050          | 919,9  | 273,05   | 90,0         |
| 20  | 724    | 30  | 00:00 | 2182,1 | 0,00119          | 0,00052          | 935,7  | 275,95   | 59,7         |
| 20  | 725    | 31  | 00:00 | 2189,5 | 0,00119          | 0,00049          | 942,0  | 273,05   | 87,4         |

## 3.2. Yöntem

### 3.2.1. Makine öğrenmesi

Makine öğrenmesi, veriler ve bilgiler arasındaki temel ilişkileri sentezlemek için sistematik olarak algoritmalar uygulayan bir yapay zekâ dalıdır. Ortaya çıkmasındaki amaç en başta, bilgisayarların bilmediği gelecekteki olayları veya senaryoları tahmin etmek şeklindeydi. Ancak sonraki dönemlerde farklı alanlardaki yaygın kullanımı ve kullanım sonucunda yüksek başarı performansı sergilemesi birçok yeni formunun da ortaya çıkmasına neden olmuştur.

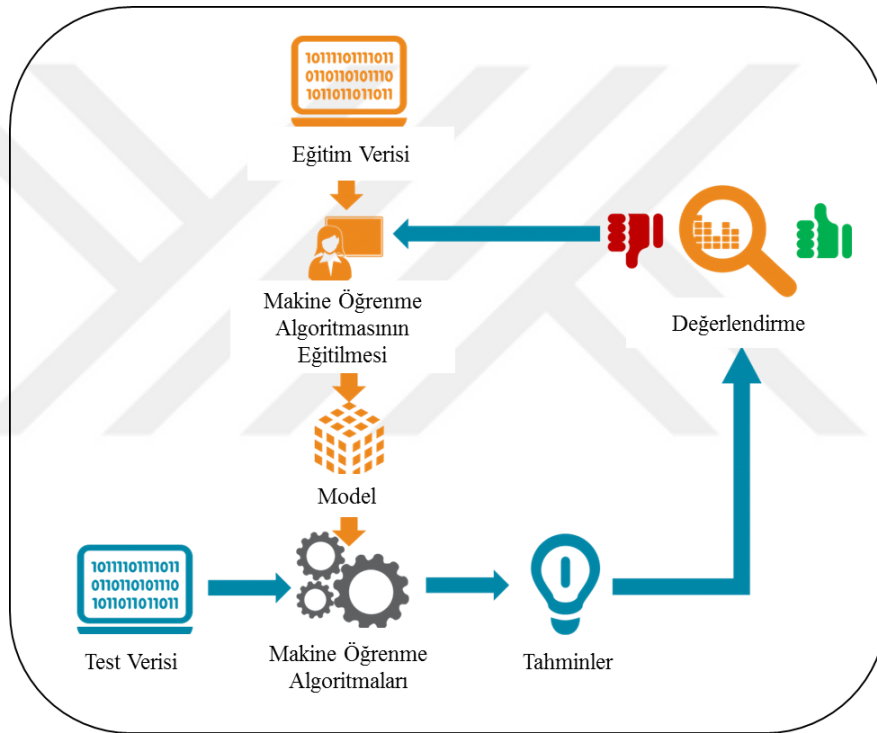
İlk olarak 1960'lı yıllarda Arthur Samuel tarafından makine öğrenimi “bilgisayarlara açıkça programlanmadan öğrenme yeteneği veren çalışma alanı” olarak tanımlanmıştır (Samuel, 1959). Samuel'in tanımına göre; bilgisayarları deneyimlerden öğrenmek için programlamanın, ayrıntılı programlama çabasının çoğuna olan ihtiyacı ortadan kaldırdığı sonucuna varılmaktadır. Diğer taraftan bilgisayarın temelini oluşturan Turing, makine zekâsını göstermek için bir kıyaslama standardı getirmiştir, “öyle ki bir makinenin bir insanınkinden ayırt edilemeyecek şekilde akıllı ve duyarlı olması gerekir” ifadesi ile makine öğrenimi kavramının temellerini atmıştır (Turing, 1950). Literatürdeki en genel makine öğrenimi kavramı ise; “Uygun, anlaşılır ve düzenli yapıdaki verinin işlenerek farklı bilgilere/olaylara dönüştürülmesini sağlayan, bilgisayar algoritmalarının geliştirilmesine dayalı uygulama alanıdır” şeklindedir (Shalev-Shwartz ve Ben-David, 2014).

Günümüzde farklı çalışma alanlarında yapılacak olan uygulamalar için olmazsa olmaz verinin hem miktarının sürekli artması hem de ele alınması gereken değişkenlerin farklılık göstermesi mevcut durumu karmaşıklştırmaktadır. Veri ve bilgi arasındaki ilişkilerin karmaşaya uğramadan, sahip oldukları yapının korunması için zaman zaman düzenlerini anlamaya ihtiyaç duyulabilmektedir. İşte tam olarak bu hedefin odak noktasında yer alan makine öğrenmesi, bilgisayarların belirli bir görevi yerine getirmek için veri odaklı kararlar almalarını sağlar. Çeşitli yöntemler veya algoritmalar kullanarak, zaman içinde sisteme yeni verilerin dâhil edilebildiği öğrenici ve geliştirilebilir sistem tasarımı oluşturur (Mitchell, 1997).

Makine öğrenimi sistemleri, yenilikçi ve sürekli gelişen yapısını, geniş veri dizileri içindeki temel ilişkileri, büyük veri analitiği, davranışsal kalıpları tanıma ve bilgi evrimindeki sorunları çözecek şekilde karakterize etme becerisine borçludur.

Ayrıca, çalışma davranışındaki farklılıkları modellemek için bir sürecin değişen koşullarını kategorize etme olarak eğitilebilir. Bilgi yapıları, yeni fikirlerin ve teknolojilerin etkisi altında geliştikçe, makine öğrenimi sistemleri mevcut modellerdeki aksaklıkları belirleyebilir ve yeni bilgiye uyum sağlamak için kendilerini yeniden tasarlayabilir (Awad ve Khanna, 2015).

Makine öğrenmesindeki temel işlem süreci Şekil 3.7'deki sıraya göre oluşmaktadır. İlk aşamada öğrenme algoritmasıyla eğitim veri seti kullanılarak model eğitilir. İkinci aşamada makine öğrenmesi algoritması, eğitim veri setinde bulunmayan yeni bir test verisi ile karşılaştığında oluşturduğu model temelinde bir tahmin yapar.



Şekil 3.7. Makine öğrenmesindeki temel işlem süreci (URL, 2021)

Yaptığı tahminin doğruluk oranı değerlendirilir. Eğer doğruluk oranı kabul edilebilir düzeyde ise model başarılı sayılır ve uygulamada yer alır. Ancak doğruluk oranı kabul edilebilir düzeyde değilse, makine öğrenme algoritması yeniden oluşturulan bir veri seti ile modeli baştan eğitir.

Bilgisayarlara insana benzer öğrenim stratejisi kazandırmak için çeşitli algoritma ve tekniklerle donatılan makine öğrenmesi, son yıllarda web araması, otomasyon, telekomünikasyon, davranış analizi, hava durumu tahmini, büyük veri analitiği ve daha birçok uygulamada yaygın kullanım alanı bulmaktadır. Gelecekte ise, bir dizi kullanıcı merkezli yeniliğin geliştirilmesinde belirleyici bir rol oynaması muhtemeldir.

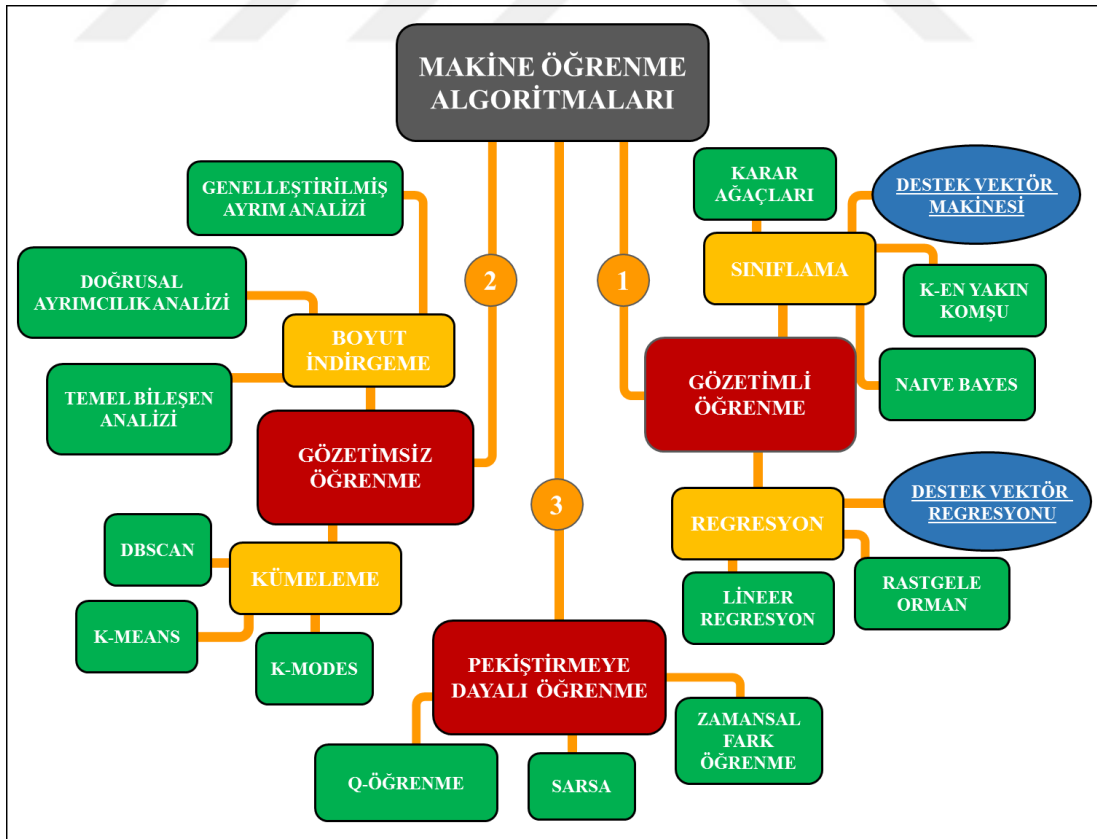
### 3.2.2. Makine öğrenme algoritmaları ve öğrenme çeşitleri

Makine öğrenmesinde kullanılacak algoritma makinenin kararına bırakılmamakta, bunun yerine öğrenilmesi hedeflenen bilgi ve analize konu olan verinin yapısı bu hususu belirlemektedir (Barnes, 2015). Çünkü çok sayıda çeşitli yaklaşımlar ve algoritmaların bulunması neticesinde, öğrenme anında esas olacak ana unsurların temelinde, veri setinin özelliklerine ve öğrenilmesi amaçlanan konuya göre değişmektedir. Dolayısıyla uygulama sonucu elde edilecek başarı, seçilecek algoritmanın türüne bağlıdır.

Makine öğreniminin olguyu öğrenme aşamasında sunulan giriş verileri ile beklenen çıktı arasındaki temel eşleştirmelere dayalı olarak, makine öğrenimi algoritmaları üç ana kategoriye ayrışabilir (Mitchell, 1997). Bunlar;

- Gözetimli Öğrenme (Supervised Learning)
- Gözetimsiz Öğrenme (Unsupervised Learning)
- Pekiştirmeye Dayalı Öğrenme (Reinforcement Learning)

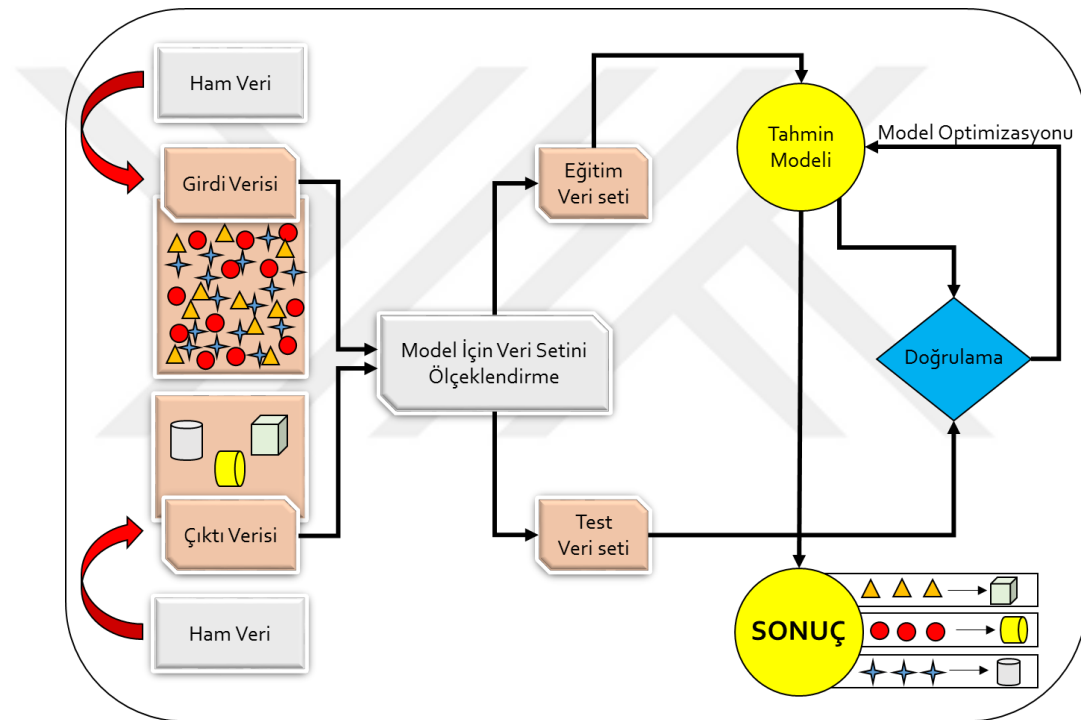
olmak üzere kategorize edilmiş ana sınıflardır (Şekil 3.8).



Şekil 3.8. Makine öğrenme tekniğindeki öğrenme çeşitleri (Galimberti, 2017)

### 3.2.2.1. Gözetimli öğrenme (Supervised learning)

Gözetimli öğrenme, girdi verileri ile tahmine tabi tutulan bir hedef değişken (çıktı verisi) arasındaki temel ilişkiyi belirleyen öğrenme mekanizmasıdır. Mekanizmada yeni olayın tahmini için Şekil 3.9'dan yola çıkılarak, eğitim veri setindeki hem girdi hem de istenen çıktı veriler sağlanarak model oluşturulur. Eğitim setindeki girdi ve çıktı verileri, oluşturulan modelin öğrenme temelini sağlamak amacıyla önceden etiketlidir. Diğer bir ifadeyle hangi girdi verisinin hangi çıktı verisine karşılık geldiğinin bilinmesi veri setinin etiketli olduğu anlamına gelmektedir.



Şekil 3.9. Gözetimli öğrenme tekniği için sürecin işleyişi

Bu öğrenme tekniğindeki algoritmalar, girdi ve çıktı değerleri arasındaki ilişkiyi ifade etmesi beklenen fonksiyonları üretmeye çalışır. Öğrenme fonksiyonu, girdi değerleri ve çıktı değerleri arasındaki temel ilişkiyi genelleştirmeye çalışan modeli sentezlemek için eğitim verilerini kullanır. Eğitim verileri, gözlemlenen girdi değerlerini ve istenen çıktı değerini içerir. Denetimli öğrenme algoritmasına dayalı iyi eğitilmiş bir işlev modeli, yeni örnek ve örneği oluşturan girdiler verildiğinde hangi çıktı değerinin oluşacağını tahminini yapar (Bonaccorso, 2017).

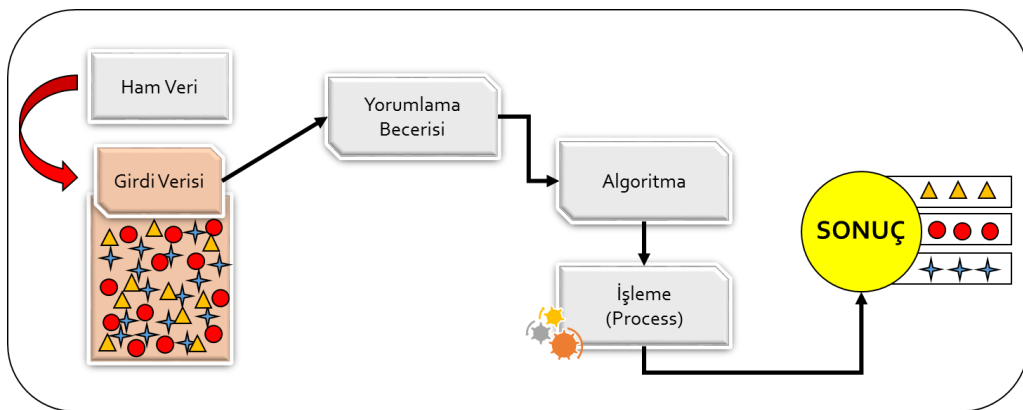
Algoritmaları öğrenmenin amacı, belirli bir girdi kümesi (eğitim seti) için hatayı en aza indirmektir. Bununla birlikte, etiketli örneklerin doğruluğu ve çok

yönlülüğünden etkilenen kalitesiz bir eğitim seti için, model tipik olarak zayıf genelleme ve hatalı sınıflandırmayı temsil eden aşırı uyum sorunuyla karşılaşabilir.

Gözetimli öğrenme algoritmaları en genel anlamda, regresyon ve sınıflandırma olmak üzere kendi içinde ikiye ayrılır. Regresyon iki veya daha fazla değişkenin arasındaki ilişkiyi bulmayı amaçlarken, sınıflandırma ilk defa karşılaşılan bir verinin eğitim seti kullanılarak sınıflandırılmasını yapar. Regresyon formu için lineer regresyon modeli, rastgele orman, destek vektör regresyonu örnek algoritmalar olarak verilebilirken; sınıflandırma formu için karar ağaçları, DVM, K-en yakın komşu, YSA örnek sayılabilir (Niculescu-Mizil ve Caruana, 2005).

### 3.2.2.2. Gözetimsiz öğrenme (Unsupervised learning)

Gözetimsiz öğrenme algoritmaları, gözetimli öğrenmenin aksine eğitim verisinde çıktı verileri olmadan sadece girdi verilerini kullanarak etiketsiz veri kümelerindeki gizli yapıları keşfetmek için tasarlanmıştır (Awad ve Khanna, 2015). Etiketsiz veri kümesi denilerek anlatılmak istenen durum, bir girdi verisine karşılık gelen çıktı verisinin olmayışıdır. Bu mekanizma, veri sıkıştırma, aykırı değer tespiti, sınıflandırma, insan öğrenimi vb. alanlarda birçok yaygın kullanım bulmuştur. Öğrenmeye genel yaklaşım, olasılıksal veri modelleri aracılığıyla eğitimi içerir. Genel olarak, gözetimsiz bir öğrenme veri seti  $x_1, x_2, x_3, x_4, \dots, x_n$  girdilerinden oluşur, ancak ne hedef çıktıları (gözetimli öğrenmede olduğu gibi) ne de çevresinden gelen pekiştirmeye dayalı bilgileri içermezler. Bu durumda öğrenimin amacı, verimli karar verme, tahmin, bilgi filtreleme ve kümeleme için girdi verilerinin temsillerini yorumlayarak varsaymaktır (Şekil 3.10).



Şekil 3.10. Gözetimsiz öğrenme tekniği için sürecin işleyişi

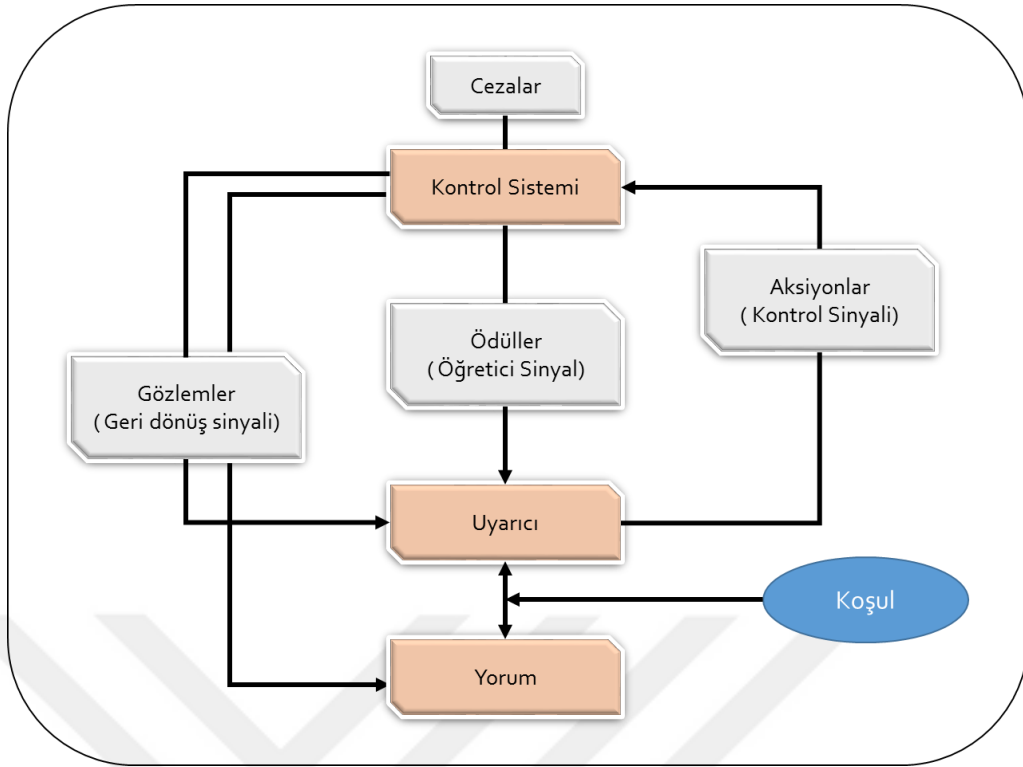
Gözetimsiz öğrenmenin iki popüler kategorisi, kümeleme ve boyut indirgeme formudur. Özellikle verilerdeki gizli kalıpları veya gruplamaları bulmak için kullanılan kümeleme (Clustering) yönteminde modele bir veri seti verildiğinde, içinde kümeler oluşturarak veri setindeki kalıpları ve veriler arasındaki ilişkileri bulur (Yüksek, 2020). Bu tarz gruplar küme olarak adlandırılmakla birlikte, bu noktada amaç spesifik ve tek bir doğruya yakınsama elde edilmesi değildir. Sonuçların doğruluğu, yapılan kümeleme işleminde, her bir birimin dâhil edildiği gruba, ne ölçüde benzediği ve diğer gruplardan ne ölçüde farklı olduğu ile ölçülmektedir (Berry ve Linoff, 2004).

### 3.2.2.3. Pekiştirmeye dayalı öğrenme (Reinforcement learning)

Pekiştirmeli öğrenme tekniği, kümelenmiş ödülü en üst düzeye çıkarmak için belirli bir ortamda, uyarıcı tarafından takip edilecek şekilde eylem veya davranış dizisinin araştırılmasını içerir (Şekil 3.11). Uyarıcının eylemi, ortamın durumunda gözlemlenebilir bir değişikliği tetikler. Öğrenme tekniği, belirli bir dizi deneysel eylem ve çevrenin durumuna yönelik gözlemlenen tepkiler için kendisini eğiterek bir adaptasyon modelini sentezler.

Genel olarak, bir dizi eylemle ilişkili ödüller ve cezalar içeren bir kontrol-teorik, deneme-yanılma öğrenme paradigması olarak görülebilir (Mohri ve ark., 2018). Uyarıcı, deneyimleri ve deneme-yanılma sonucunda ortaya çıkan ödüllere göre politikasını değiştirir. Araştırdığı ve ödüllere sonuçlanan geçmiş eylemleri arar. Olası tüm eylem ödül tahminlerinin kapsamlı bir veri tabanını veya modelini oluşturmak için birçok testi olmayan eylemi kullanır. Sonucunda, yeni olası eylemlerin keşfi ile bu eylemlerden kaynaklanan başarısızlık olasılığı arasında bir denge kurarak olayları öğrenir (Sayıt ve Sönmez, 2012).

Şekil 3.11’de izlenen politikaya göre, ortamda gerçekleşen aksiyonlar kontrol sistemi tarafından değerlendirilip ödüller ve cezalar şeklinde karara bağlanmaktadır. Uyarıcı, mevcut yapıyı yorumlayan tahmini bir değer işlevini temsil eder. Alternatif olarak, bir olaydaki koşulları da dikkate alarak mevcut durumun performansını değerlendirir. Burada uyarıcı, sürekli denemeler ve düzeltmeler yaparak politikayı şekillendirir. Karardaki ödüllendirme eylemi, girişilen bir olayda istenen/beklenen durum iken cezalandırma eylemi istenmeyen/beklenmeyen durumu ifade eder.



Şekil 3.11. Pekiştirmeli öğrenme tekniği için sürecin işleyişi (Awad ve Khanna, 2015)

Literatürde çalışma amacına göre kullanılan birçok makine öğrenme algoritması ve farklı matematiğe dayalı öğrenme yöntemleri bulunmaktadır. Yukarıda genel başlıklar altında bahsedilen yöntem ve algoritma çeşitleri hakkında bilgiler verilmiştir.

Amaç doğrultusunda bu çalışmada uygulanabilecek yöntemler araştırılmıştır. Araştırma sonucuna göre, farklı çalışma alanlarında da yaygın kullanım alanına sahip gözetimli öğrenme sınıfından DVM'deki DVR formu yöntem olarak seçilmiştir. DVR yönteminin seçilmesinin en temel sebepleri;

- Gerçek ve tahmin edilen değer arasındaki hatayı en aza indirmek için DVR, bir eşik değeri içindeki en iyi çizgiyi uydurmaya çalışır. Yani eşik değeri (sınır) içerisinde kalan veri kümesinin maksimum olmasını hedefler.
- Aykırı verilere karşı sağlamdır (robusttur). Dolayısıyla bu şu anlama gelmektedir: GNSS gözlemleri için önemli hata kaynağı olan troposferik gecikmenin belirlenmesinde kullanılacak olan meteorolojik veriler hatalı ölçülmüş olsa da DVR'de belirli bir eşik değeri olduğu için sonuç üzerinde çok büyük etkisi olmayacaktır.
- DVR modeli, her veri noktası için değişken sayısının az olduğu durumda yüksek performans, değişken sayısının fazla olduğu durumda ise düşük performans

göstermektedir. Çalışmada kullanılan değişken sayısının az olması model performansını yüksek tutmaktadır.

- Model, veri kümesinin daha az gürültülü olduğu durumda iyi performans göstermektedir. Çalışmadaki sıcaklık, basınç, bağıl nem gibi değişkenlerin belirli bir aralıkta bulunmasından (max. ve min. arasında çok fazla fark bulunmaması) kaynaklı daha az gürültülü olduğu için veri kümesi modele uygunluk göstermektedir.
- Karar verici, modeli kolaylıkla güncelleyebilir. Ayrıca genelleme kabiliyeti yüksektir.

şeklinde sıralanabilir.

### 3.2.3. Destek vektör makineleri

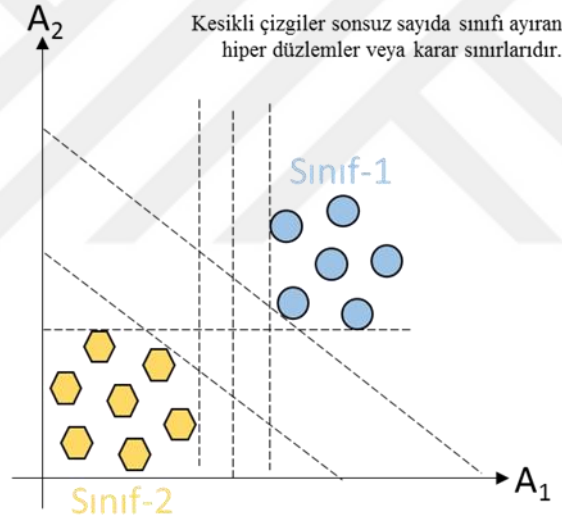
Regresyon ve sınıflama problemlerinin çözümü için günümüzde makine öğrenme algoritmaları önemli yer tutmaktadır. Özellikle mühendislik, finansal yönetim, kimya, tıp, endüstri gibi farklı alanlarında kullanılan bu algoritmalar birçok sınıflama ve tahmin problemlerinin çözümüne katkı sağlamaktadır (Trafalis ve Ince, 2000; Chen, 2004; Rostami ve ark., 2015). Son zamanlarda makine öğrenmede genelleme başarısı yüksek ve birçok probleme çözüm geliştiren en başarılı algoritmalarından birisi de Destek Vektör Makineleri (DVM) olmuştur. Regresyon ve sınıflandırma formlarını içeren DVM, Vapnik ve arkadaşları tarafından istatistiksel öğrenme teorisine dayalı olarak geliştirilen sağlam ve önemli bir öğrenme yöntemi olarak tanıtılmıştır (Vapnik, 1995).

DVM, öğrenme, sınıflandırma, kümeleme, çekirdek yoğunluğu tahmini ve son olarak da verilerden regresyon kuralları üretmek için kullanılan bir eğitme algoritması şeklindedir. Teorik temelleri ilk olarak doğrusal veriyi sınıflandırma konusunda atılmış ve sonrasında bu algoritmalar hem doğrusal olmayan hem de çok sınıflı verileri sınıflandırma amacıyla geliştirilmiştir (Gunn, 1998). DVM algoritmalarının temel amacı, iki veya daha fazla sayıdaki veri kümesine sahip bir veri setinden sınıflama algoritmasını öğrenerek, modele dışarıdan dâhil edilecek yeni bir verinin hangi sınıflamada yer alması gerektiğini tahmin etmesidir.

Son zamanlarda özellikle uzaktan algılama konusunda da güncel kullanılmakta olup görüntülerin sınıflandırılması şeklinde uygulanan bir sınıflandırma algoritmasıdır

(Li ve ark., 2006; Çölkesen, 2015). Örüntü-nesne tanıma, ses tanıma ve yüz tanıma gibi birçok uygulamalarda da yerini almıştır (Burges, 1998; Eray, 2008; Deng ve ark., 2015). Yöntemdeki temel süreç iki sınıf arasındaki ayrımın optimum olmasını sağlayacak hiper düzlemin (ayırt etme yüzeyinin) bulunması veya veriyi birbirinden ayırmak için en uygun karar fonksiyonunun belirlenip modellenmesidir (Vapnik ve Chappelle, 2000; Cervantes, 2020).

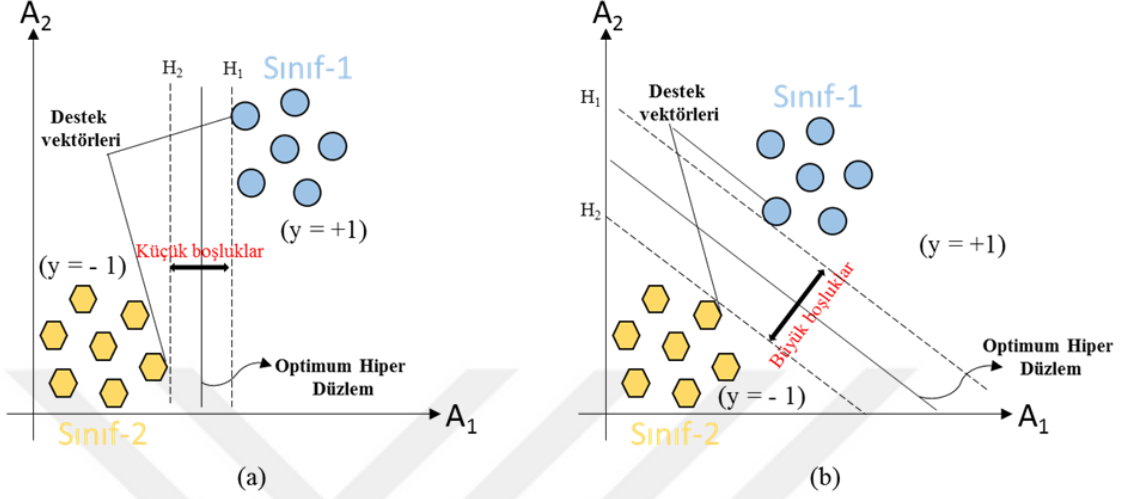
DVM’de, verilerin doğrusal olarak ayrılabilir olduğu veya doğrusal olarak ayrılamaz olduğu iki durum söz konusudur. Doğrusal olarak ayrılabilen veri yapılarında, iki sınıf birbirinden bir doğru ile ayrılabilirken, doğrusal olarak ayrılamayan verileri sınıflandırmak için, veri çeşitli yollarla farklı bir boyuta taşınır ve burada en iyi ayırıcı hiper düzlem bulunarak veri sınıflandırılır. DVM’nin doğrusal, ayrılabilir ve iki sınıflı olduğu problemdeki çözüm arayışı Şekil 3.12’de gösterildiği gibidir.



Şekil 3.12. DVM ile iki boyutta doğrusal ayrılabilen verilerin sınıflandırılması

Şekil 3.12’de sınıflama işlemi için  $A_1$  ve  $A_2$ , iki giriş özelliğine dayalı bir grafiği ifade etmektedir. Kümelenmiş halde bulunan mavi renkteki semboller Sınıf-1’i temsil ederken, sarı renkteki semboller ise Sınıf-2’yi temsil etmektedir. Grafiğe bakıldığında, iki boyutlu verilerin doğrusal olarak ayrılabilir olduğu görülmektedir. Üstelik bu verileri (Sınıf-1’i) diğerinden (Sınıf-2) direkt olarak ayırabilen birçok hiper düzlem (doğru) çizilebilir. Ancak DVM’nin asıl amacı ve buna bağlı olarak sağladığı avantaj, karşısına bilinmeyen bir veri seti geldiğinde sınıflama hatasını en küçük yapacak optimum hiper düzlemi seçmesidir.

Bir DVM, maksimum marjinal (büyük boşluklu) hiper düzlemi arayarak bu amaca yaklaşır. İki farklı sınıfı ayıran hiper düzlemler ve bunlarla ilişkili kenar boşlukları Şekil 3.13’de gösterilmiştir.



Şekil 3.13. DVM ile oluşturulan küçük (a) ve büyük (b) boşluklu hiper düzlemler

Burada, büyük boşluklu hiper düzlemler, sınıfı bilinmeyen diğer veri noktalarını küçük boşluklu hiper düzlemlerden daha doğru sınıflamaktadır. Kısaca özetlenecek olursa DVM birbirinden en uzak iki hiper düzlem belirlemektedir. En uzak iki hiper düzlemin ortasında yer alan hiper düzleme ise optimum hiper düzlem adı verilmektedir. DVM’nin hedefi hiper düzlem ile en yakın örnek nokta (destek vektörü) arasındaki mesafeyi maksimum yapmaktır.

Destek vektördeki temel eşitlikler iki sınıfa ait doğrusal olarak ayrılan veri setleri göz önüne alınarak geliştirilmiştir. DVM’deki problem çözümünde etken olan ayırıcı bir alt düzlem (3.1) eşitliğindeki gibi yazılmaktadır:

$$W \cdot x_i + b = 0 \quad (3.1)$$

Söz konusu eşitlikte, iki sınıflı  $n$  sayıda doğrusal yapıdaki veriden oluşan bir eğitim verisi seti  $\{x_i, y_i\}$ ,  $i=1, \dots, k$  şeklinde tanımlanmaktadır. Burada  $x_i = (x_1, x_2)$ , Şekil 3.12’de bahsedilen  $A_1$  ve  $A_2$ ’ye ait giriş bilgileri,  $y_i \in \{-1, +1\}$  ise sınıfı belirtmektedir. Hiper düzleme ait ağırlık vektörü  $W = (w_1, w_2, \dots, w_n)$  ve hiper düzlemin uzayda yönelimini tanımlayan  $b$  sayısal büyüklüğünün hesaplanmasıyla iki sınıfı birbirinden ayıran hiper düzlem belirlenmektedir (Çölkesen, 2015). Şekil 3.13’de görüldüğü gibi

( $H_1$  ve  $H_2$ ) sözü geçen hiper düzlemler (3.2) ve (3.3) eşitlikleri yardımıyla hesaplanmaktadır.

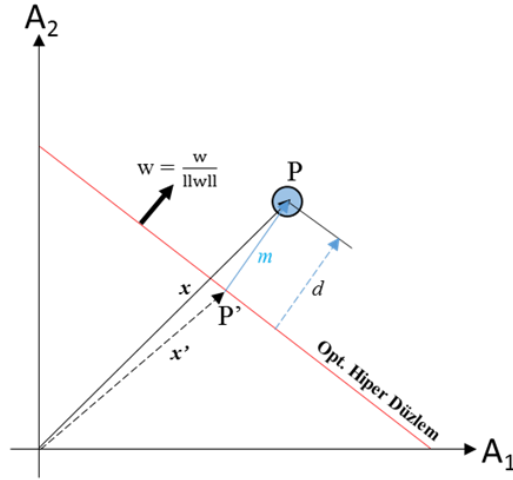
$$H_1: w_0 + w_1x_1 + w_2x_2 \geq 1 \text{ her } y_i = +1 \text{ için} \quad (3.2)$$

$$H_2: w_0 + w_1x_1 + w_2x_2 \leq -1 \text{ her } y_i = -1 \text{ için} \quad (3.3)$$

Eşitlik sonucuna bağlı olarak,  $H_1$  hipotezinin gerçekleşmesi durumunda modele dâhil edilen (sınıfı bilinmeyen) veri seti +1 sınıfına dâhil olur. Ancak  $H_2$  hipotezi gerçekleşirse bu durum tam tersi olarak gerçekleşir ve veri seti -1 sınıfına yönelir. Bu iki eşitlik düzenlenerek tek eşitlik ile gösterilecek olursa genel eşitlik,

$$y_i (w \cdot x_i + b) - 1 \geq 0 \text{ ya da } y_i (w_0 + w_1x_1 + w_2x_2) \geq 1 \quad (3.4)$$

şeklinde elde edilir. Eşitliklerden elde edilecek sonuçlar,  $w$  ve  $b$  parametrelerine bağlı değişim göstermesinden dolayı (değiştirme/ölçekleme) birçok hiper düzlem ortaya çıkabilmektedir (Osuna ve ark., 1997). Bir  $X_p$  destek vektörü ile seçilmiş olan optimum hiper düzlem arasındaki dik uzaklık aşağıdaki gibi belirlenir (Şekil 3.14).



**Şekil 3.14.** Destek vektörü ( $X_p$ ) ile optimum hiper düzlem arasındaki geometrik ilişki

Burada  $d$  (marjin); veri örneği veya destek vektörü ile hiper düzlem arasındaki öklid uzaklığıdır.  $x$  ve  $x'$ ; vektörler,  $w$ ; bir birim vektördür.  $P'$  noktası;  $P$  noktasının hiper düzlem üzerine dik izdüşümüdür.  $m$  ise  $w$  birim vektörü ile aynı yöndedir ve normu  $d$  mesafesi kadardır.

$$m = d \cdot \frac{w}{\|w\|} \quad (3.5)$$

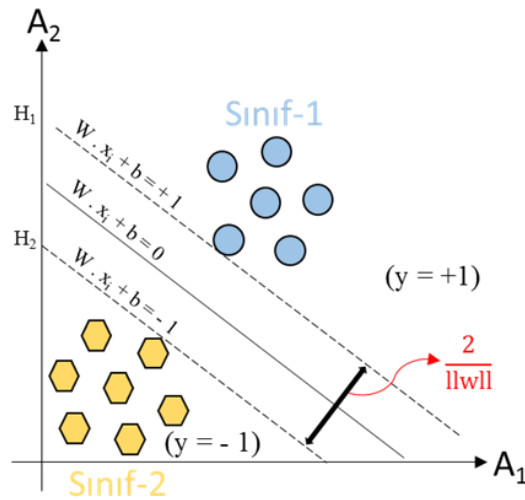
Vektörlerin çıkarma özelliğinden yararlanıp;

$$x' = x - m, \text{ işleminden yola çıkarak } x' = x - d \cdot \frac{w}{\|w\|} \quad (3.6)$$

eşitliği elde edilir.  $P'$  noktası hiper düzlem üzerinde olduğu için  $x'$  vektörü hiper düzlemin denklemini sağlayacaktır. (3.5) ve (3.6) eşitlikleri için kullanılan formüller hiper düzlemin bir denklemi olan (3.1) eşitliğinde yerine yazılırsa;

$$\begin{aligned} \langle W, x' \rangle + b &= 0 \\ \langle W, x - d \cdot \frac{w}{\|w\|} \rangle + b &= 0 \\ \langle W, x \rangle - d \cdot \frac{\langle w, w \rangle}{\|w\|} + b &= 0 \\ \langle W, x \rangle - d \cdot \frac{\|w\|^2}{\|w\|} + b &= 0 \\ d &= \frac{|\langle W, x \rangle + b|}{\|w\|} = \frac{1}{\|w\|} \end{aligned} \quad (3.7)$$

eşitliğine dönüşür. Buradan  $d$  uzaklığı, yani  $X_p$  destek vektörü ile hiper düzlem arasındaki dik uzaklık elde edilmiş olur.  $d = \frac{1}{\|w\|}$  ile hiper düzlemin maksimum olması istendiği için  $2d$  mesafesi gerekli olduğundan değerin maksimizasyonu gerekir ve böylelikle  $2d = \frac{2}{\|w\|}$  elde edilir (Şekil 3.15).



Şekil 3.15. İki hiper düzlem arasındaki dik uzaklık

Destek vektörleri arasındaki sınırın maksimum hale getirilmesi ağırlık vektörünün ( $w$ ) minimum hale getirilmesi ile elde edilmektedir. Bu durumda, optimizasyon problemi aşağıdaki şekilde ifade edilebilir.

$$\text{Min}_{w,b} = \frac{1}{2} \|w\|^2 \quad (3.8)$$

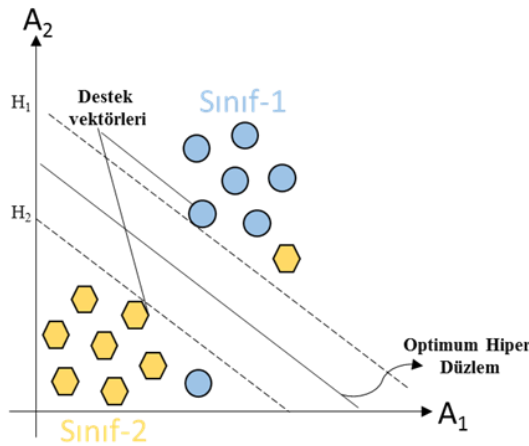
$$y_i (w \cdot x_i + b) - 1 \geq 0 \quad i = 1, \dots, k \quad (3.9)$$

Söz konusu koşul ve eşitlikler (3.8 ve 3.9) ile ifade edilen doğrusal olmayan optimizasyon problemi, Lagrange teorisi kullanılarak çözülebilmektedir. Lagrange teorisinin temel amacı, başlangıçta hiçbir eşitsizlik sabiti olmayan bir optimizasyon probleminin çözümünü karakterize etmektir (Yıldırım, 2006).

$F(x)$  amaç fonksiyonlu ve  $h_i(x)=0$ ,  $i = 1, \dots, m$  eşitlik sabiti verilen bir optimizasyon problemi için,  $L(x, \lambda) = f(x) + \sum_{i=1}^m \lambda_i h_i(x)$  lagrange denklemidir ve  $\lambda_i$  lagrange çarpanını ifade eder (Savaş, 2019). Sonuç olarak DVM ile iki sınıflı doğrusal veriyi birbirinden ayırmak için en uygun karar fonksiyonu,

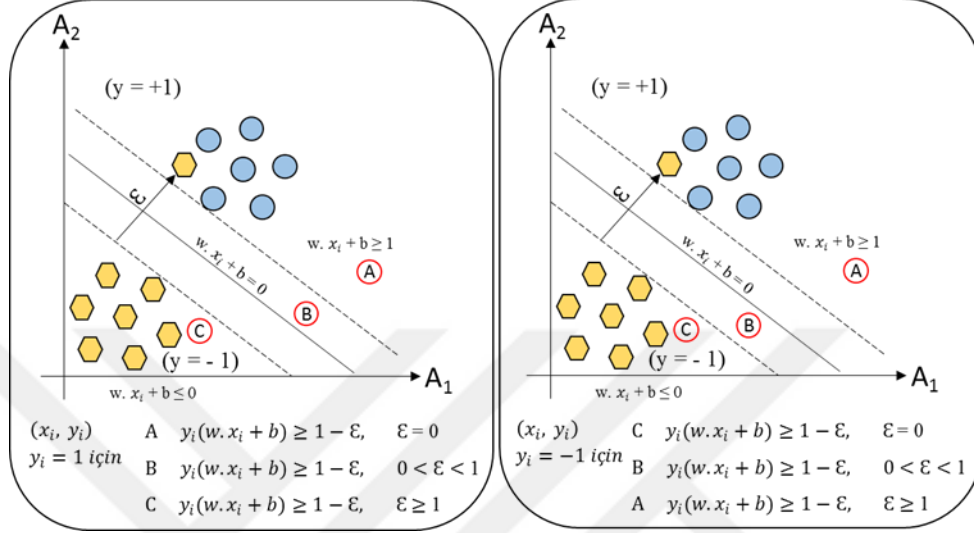
$$f(x) = \text{sign} \left( \sum_{i=1}^m \lambda_i y_i (x \cdot x_i) + b \right) \quad (3.10)$$

eşitliği ile belirlenmektedir (Tso ve Mather, 2009). Genel olarak karşılaşılan çoğu problemde verilerin doğrusallığından ziyade doğrusal olmadığı durumlar söz konusu olmaktadır. Diğer bir ifade ile Şekil 3.16'da belirtildiği gibi doğrusal bir düzlem yardımıyla verilerin birbirinden ayrılabilmesi durumu vardır.



Şekil 3.16. Verilerin doğrusal olarak ayrılabilmediği durum

Cortes ve Vapnik (1995), doğrusal olmayan verilerin sınıflandırılması için Şekil 3.17'deki bölgelerde sınıflandırma hatalarını ifade eden değişkenin ( $\epsilon$ -epsilon) tanımlanması ve optimizasyon modeline eklenmesi ile çözümün bulunabileceğini ifade etmişlerdir.



Şekil 3.17. Doğrusal ayrılmayan veriler için tanımlanmış hiper düzlemler

Hatalı sınıflandırılan veri noktaları için tanımlanan değişkenler ( $\epsilon$ ) göz önüne alındığında, optimum hiper düzlem, destek vektörler arasındaki sınırı maksimum yaparken hatalı sınıflandırılmış veri noktasını, minimum yapacak şekilde tespit etmelidir (Huang ve ark, 2002). Sınırın maksimum hale getirilmesi ve hatalı sınıflandırmaların minimum hale getirilmesi arasındaki denge, kullanıcı tarafından belirlenen düzeltme parametresi ( $C$ ) yardımıyla ayarlanmaktadır. Dolayısıyla koşula,  $C \sum_{i=1}^m \epsilon$  ( $i = 1, \dots, k$ ) gibi terimin eklenmesine sebep olmaktadır. Burada, eğer  $C$  küçük ise doğru olarak sınıflandırılmayan birçok veri kabul edilir (Özkan, 1998; Çölkesen, 2009).  $C$  parametresi,  $\epsilon$  değişkeninin büyük değer alması sonucu çözüme ceza vererek minimum koşulunun  $\frac{1}{2} \|w\|^2$  sağlanmasında kullanılır. Söz konusu optimizasyon problemi, doğrusal olmayan verilerde (3.11) ve (3.12) eşitliğine dönüşür.

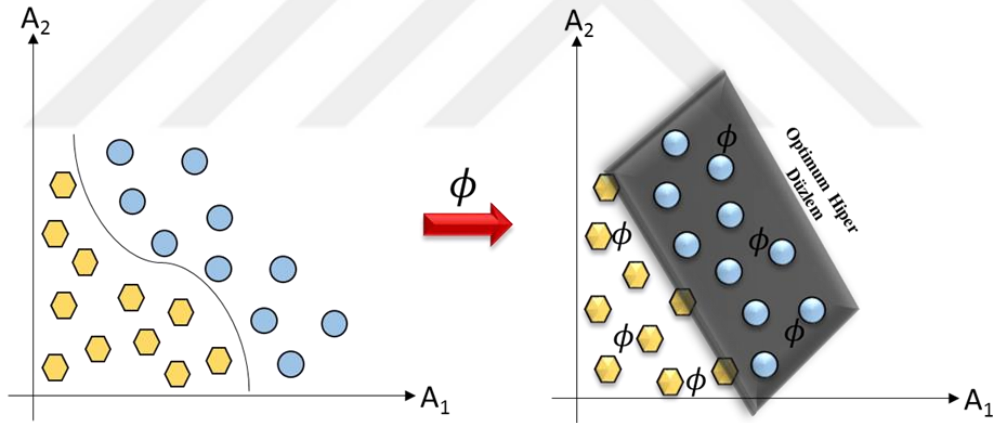
$$\min_{w,b,\epsilon} \left[ \frac{1}{2} \|w\|^2 + C \sum_{i=1}^m \epsilon \right] \quad (3.11)$$

$$y_i(w \cdot x_i + b) - 1 + \epsilon \geq 0 \quad ; \quad \epsilon \geq 0, i = 1, \dots, k \quad (3.12)$$

Sonuç olarak Lagrange fonksiyonu ile doğrusal olmayan veriyi birbirinden ayırmak için en uygun karar fonksiyonu, (3.13) eşitliği yardımıyla verilmiştir.

$$f(x) = \text{sign} \left( \sum_{i=1}^m \lambda_i y_i \phi(x) \cdot \phi(x_i) + b \right) \quad (3.13)$$

Karar fonksiyonundan anlaşılaçağı üzere doğrusal olmayan problemlerde yani doğrusal olarak ayırt edilemeyen verilerin olduğı durumlarda, verileri daha yüksek boyutlu bir özellik uzayına taşıyarak çözümlmek için doğrusal olmayan fonksiyonlar kullanılmaktadır. Özellik uzayındaki optimum hiper düzlemin belirlenmesi için eşitlik 3.13'teki  $\phi(x) \cdot \phi(x_i)$  çarpımlarının yerine kernel (çekirdek) fonksiyonları kullanılarak dönüşümler gerçekleştirilir (Mathur ve Foody, 2008). Böylelikle kernel fonksiyonlarındaki dönüşümler sonucu optimum hiper düzlem belirlenerek veriler doğrusal olarak ayrılabilir (Şekil 3.18).



**Şekil 3.18.** Doğrusal olmayan verinin kernel fonksiyonu yardımıyla doğrusal olarak ayrılması

DVM, doğrusal olmayan haritalama fonksiyonu yardımıyla, giriş uzayından doğrusal olarak kolayca sınıflandırma yapabileceğı yüksek boyutlu nitelik uzayına dönüşümü, kernel fonksiyonları kullanılarak bütün değerlerin tekrar tekrar çarpım değerlerinin hesaplanarak bulunması yerine doğrudan kernel fonksiyonunda değerin yerine koyularak nitelik uzayındaki değerin bulunması sağlanır. Bu sayede, son derece yüksek boyutlu bir nitelik uzayı ile uğraşma olasılığı kalmaz. Kernel fonksiyonlarının diğ er bir avantajı da, eğitime aşamasında bir eğitim örneğı için fonksiyon kurulup değerler bulunduktan sonra diğ er örnekler için artık kalıp değerleri

eğitim örneği dışında tamamen hazır olduğu için çok daha kolay hesaplanmasıdır (Kecman, 2001).

### 3.2.4. Destek vektör makinelerinde kullanılan kernel fonksiyonlar

DVM'nin içyapısında yaygın olarak kullanılan 4 farklı kernel fonksiyonu bulunmaktadır. Bu fonksiyonlar doğrusal, polinom, radyal tabanlı ve sigmoid kernel fonksiyonları şeklinde ayrılmaktadır. Her bir kernel fonksiyonu kendine göre vektörel noktalarını üreterek yüksek boyuta dönüşüm yapmakta ve sonuçları da bu dönüşümlerden etkilenmektedir.

#### ➤ Doğrusal kernel fonksiyonu

Doğrusal fonksiyon, en basit çekirdek türüdür. Bu fonksiyondaki temel yapının matematiksel gösteriminde,

$$K(x, y) = (x \cdot y) \quad (3.14)$$

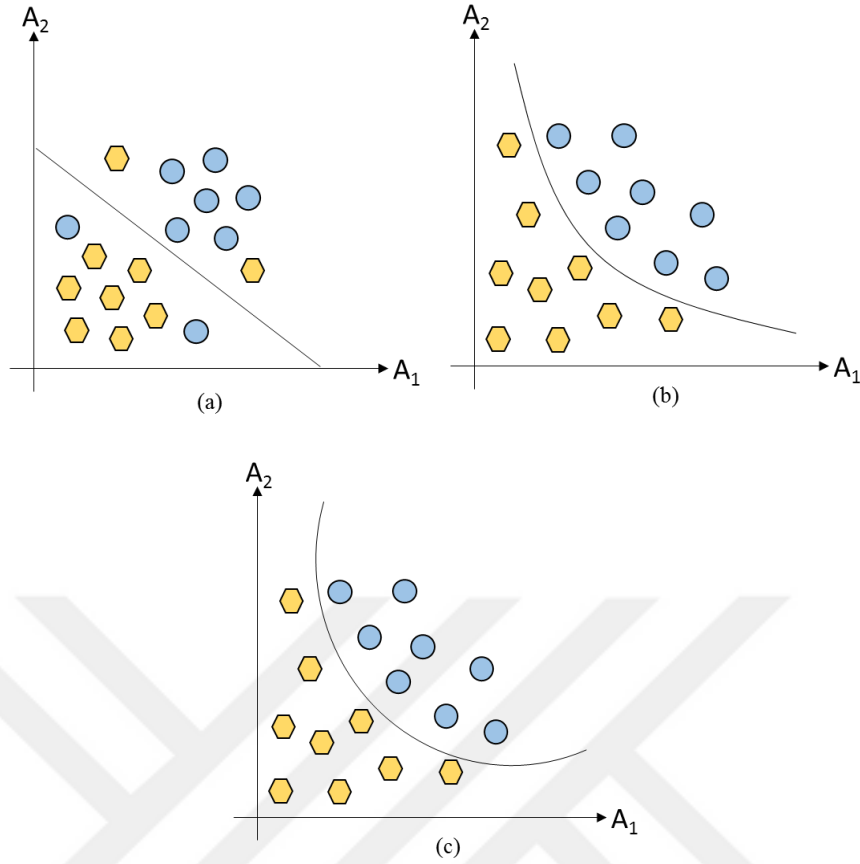
eşitliği kullanılmaktadır.

#### ➤ Polinomal kernel fonksiyonu

Bu çekirdek fonksiyonunun işlevi (3.15) eşitliği ile ifade edilmektedir.

$$((x, y) + 1)^d \quad (3.15)$$

Bu işlevde iki parametre vardır; biri sabit olan C değeri ve diğeri çekirdeğin derecesini belirten d polinom derecesidir. Bir polinom çekirdeğinin derecesi 1 ise ve sabit bir çekirdek yoksa çekirdek doğrusal çekirdek ile aynı olacaktır, ancak derece arttığında hiper düzlem daha karmaşık hale gelir ve polinomal ilişkiye geçilir. İki boyutlu örnek uzayında farklı dereceden polinomal fonksiyonların sınıflandırılması aşağıda gösterilmiştir (Şekil 3.19).



Şekil 3.19. Birinci (a), ikinci (b) ve n. dereceden (c) polinoma ait çekirdek fonksiyonu

#### ➤ Sigmoid fonksiyonu

Sigmoid fonksiyon ya da iki katmanlı yapay sinir ağı ile yapılan sınıflandırmada,

$$\tanh(b(x, y) - c) \quad (3.16)$$

eşitliği kullanılmaktadır. Burada  $b$  ve  $c$ , kullanıcı tarafından belirlenen parametrelerdir.

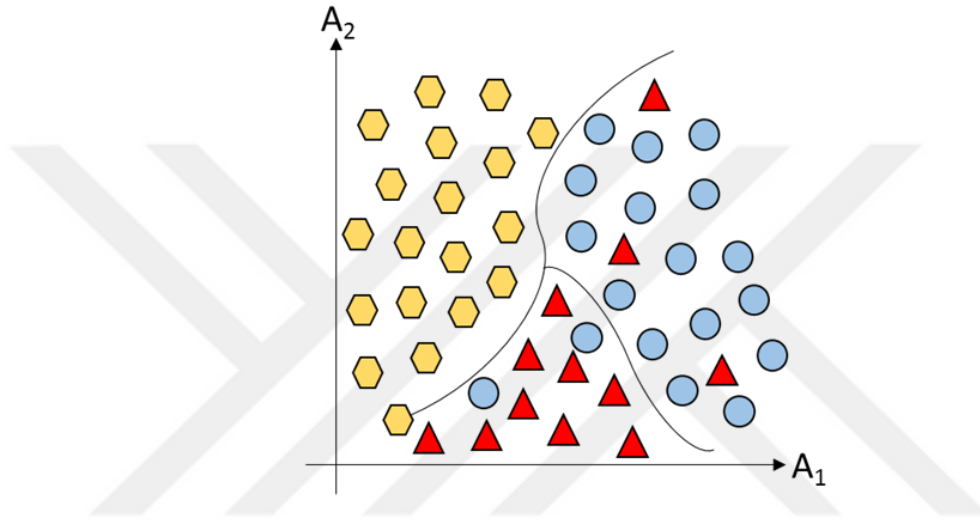
#### ➤ Radyal tabanlı fonksiyon

Polinom çekirdeğin bile çözemediği bazı problemler vardır. Bunun sebebi veri kümesi, Şekil 3.20'deki problemler gibi çok karmaşıktır. Bu gibi durumlarda radyal tabanlı fonksiyondaki (RTF) sınıflandırma işlemi, doğrusal olmadığından örneği daha yüksek boyutlu bir uzaya taşıyarak gerçekleştirir. Bu fonksiyonlardaki temel yapının matematiksel gösterimi aşağıda ifade edildiği gibidir.

$$e^{-\gamma \|x-y\|^2}$$

(3.17)

Buradaki  $\gamma$  parametresi, kullanıcı tarafından belirlenmektedir. Buraya kadar olan kısımda doğrusal, polinomal, sigmoid ve radyal tabanlı olmak üzere farklı yapıdaki kernel fonksiyonlarından bahsedilmiştir. Ancak araştırmalarda bu bahsedilenler dışında birçok çekirdek türü olmasına rağmen yeterli başarı seviyesine erişemedikleri için kullanılmamaktadırlar (Souza, 2010; Kowalczyk, 2017).



Şekil 3.20. Karmaşık yapıdaki verilerin RTF ile gösterimi

Genel kapsamda değerlendirilecek olursa, fonksiyonların her biri, girdi uzayında farklı bir doğrusal olmayan sınıflandırıcıyla sonuçlanır. Sinir ağlarındaki gibi, doğrusal olmayan DVM'ler için bulunan optimum hiper düzlemlerinin, diğer iyi bilinen sinir ağı sınıflandırıcıları tarafından bulunanlarla aynı tipte olduğu tespit edilmiştir (Xiaoyan, 2003). Örneğin, radyal tabanlı işleve (RTF) sahip bir DVM, radyal temel işlev ağı olarak bilinen bir tür sinir ağı ile aynı optimum hiper düzlemini verir. Sigmoid çekirdeğe sahip bir DVM, çok katmanlı algılayıcı olarak bilinen (gizli katmanları olmayan) basit iki katmanlı bir sinir ağına eşdeğerdir. Hangi kabul edilebilir çekirdeğin en doğru DVM ile sonuçlanacağını belirlemenin, altın kuralları yoktur. Uygulamada, seçilen çekirdek genellikle, sonuçta doğrulukta büyük bir fark yaratmaz. DVM eğitimi, çoğu yerel minimumun genellikle mevcut olduğu geri yayılım gibi sinir ağlarının aksine her zaman küresel bir çözüm bulur (Han ve ark., 2011).

### 3.2.5. Destek vektör regresyonu

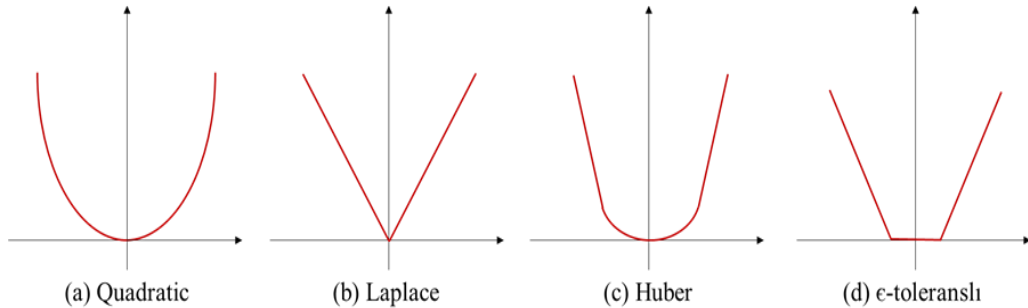
DVM'nin bir formu olan regresyon problemlerinin çözümünde kullanmaya ilişkin en kapsamlı araştırmalar Smola ve Schölkopf (2004) tarafından gerçekleştirilmiştir. DVM ile sınıflandırma işleminde temel husus, özellik uzayında eğitim verileri göz önünde bulundurarak optimal hiper düzlemi belirlemektir. DVR'de ise, gözlenen eğitim hatasını en aza indirmek yerine, genelleştirilmiş hata sınırını en aza indirmeye çalışmasıdır. Yani, regresyon için optimal hiper düzlem ve destek vektörleri arasındaki maksimum marj yerine bir “yaklaşım hatası” kullanılmaktadır. Sınıflandırmada olduğu gibi regresyon için de sınır üzerindeki en uygun genelleme araştırılır (Palaniswami ve Shilton, 2002). Sınıflandırma çözümlerindeki koşulların regresyon çözümlerinde de sağlanması beklenmektedir. Aşağıda hem sınıflama hem de regresyondaki destek vektörlerine ait genel tanımlama yapılmıştır.

$$(x_1, y_1), (x_2, y_2), \dots, (x_i, y_i); \quad x \in R^d; \quad y \in R \quad f(x) = w \cdot x_i + b \quad (3.18)$$

Buradaki  $x_i$ ;  $d$  boyutlu uzayda girdi (öğrenilen değer) vektörünü,  $y_i$ ; girdi vektörlerine karşılık gelen çıktı (tahmini değer) vektörlerini,  $w$ ; ağırlık vektörünü ve  $b$  sapmayı temsil etmektedir.

#### 3.2.5.1. Doğrusal destek vektör regresyonu

Sınıflandırma çözümünde kolaylıkla uygulanan DVM, alternatif bir kayıp fonksiyonunun verilmesiyle regresyon problemlerine uygulanabilmektedir (Gunn, 1998). Şekil 3.21'de yaygın olarak kullanılan dört kayıp fonksiyonu (a) Quadratic, (b) Laplace, (c) Huber, (d)  $\epsilon$ -toleranslı verilmiştir.

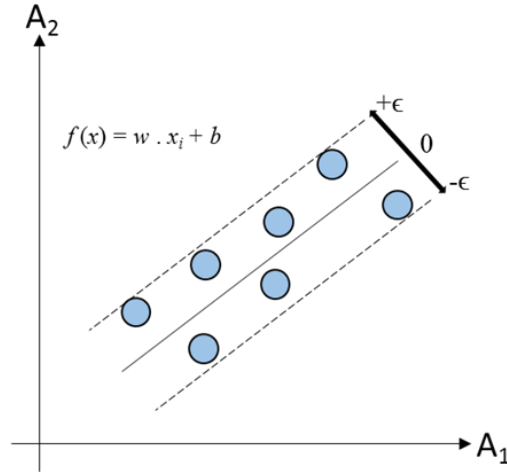


Şekil 3.21. DVR'deki kayıp fonksiyonlar

(Şekil 3.21a), Quadratic kayıp fonksiyonu klasik en küçük kareler hata kriterine karşılık gelmektedir. (Şekil 3.21b), Laplace ise hatalı verilere uç noktalarda Quadratic kayıp fonksiyona göre daha az duyarlıdır. (Şekil 3.21c), Huber, robust (sağlam) ve verilerin temel dağılımı bilinmediğinde optimal özelliklere sahip olan kayıp fonksiyonuna benzer bir fonksiyon önermiştir. Ancak bu üç kayıp fonksiyonu destek vektörünün az olduğu durumlarda veya bütün verilerini kullanma gerekçesiyle pek tercih edilmez. Bu nedenle çalışmalarda daha çok (Şekil 3.21d),  $\epsilon$ -toleranslı kayıp fonksiyonu kullanılır (Awad ve Khanna, 2015). Örneğin  $x$  ve  $y$  vektörleri arasında doğrusal ilişkinin olduğu bir doğrusal destek vektör regresyonunu ele alalım. Regresyondaki esas konu her bir eğitim verisi olan  $x_i$  değerinden, daha önce kestirilen bir hata toleransı ( $\epsilon$ )’ndan büyük olmayan bir uzaklıkta  $y_i$  değerini tespit edebilecek  $f(x)$  fonksiyonunu bulmaktır.

$$f(x) = |y_i - (w \cdot x_i + b)| \leq \epsilon \quad (3.19)$$

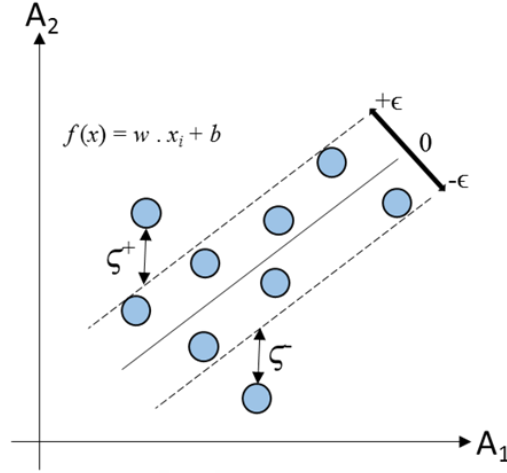
Regresyon algoritması,  $\epsilon$ ’den küçük olduğu yani veri noktasının sınır genişliğinin içinde kaldığı durumda hata olmadığını ama  $\epsilon$ ’den büyük yani sınır genişliğinin içerisinde olmayan durumlarda hataya sahip olduğunu tespit etmektedir (Şekil 3.22).



Şekil 3.22. Doğrusal destek vektör regresyonunda hata toleransı ( $\epsilon$ )

Günlük hayatta yukarıda bahsedilen durumun gerçekleşmesi söz konusu değildir. Çünkü karşılaşılan problemlerin çoğunda doğrusallık gözlenmez ve çok fazla sayıda karmaşık veri setinin bulunmasından dolayı modelin yapılandırılması gerekir. Tüm veri setine dayalı bir  $f(x)$  fonksiyonu bulunamayacağı için regresyona esneklik

parametresi  $\zeta^+$  ve  $\zeta^-$  eklenir ( $\zeta^+ \geq 0$  ve  $\zeta^- \geq 0$ ). Regresyona ait hata toleransı ve esneklik parametresi veri noktaları üzerinde gösterilmiştir (Şekil 3.23).



Şekil 3.23. Doğrusal destek vektör regresyonunda hata toleransı ve esneklik parametresi

Yukarıda geçen (3.19) eşitliğindeki fonksiyona esneklik parametresi eklenirse (3.20) eşitliği elde edilir.

$$\begin{aligned} y_i - (w \cdot x_i + b) &\leq \epsilon + \zeta^+ \\ (w \cdot x_i + b) - y_i &\leq \epsilon + \zeta^- \end{aligned} \quad (3.20)$$

Doğrusal olmayan verileri sınıflandırmadaki optimizasyon probleminin aynısı regresyon için de geçerlidir. Ancak buradaki en temel fark esneme payının da algoritma içerisine dahil edilmesidir. Hatalar ve regresyon fonksiyonunun düzlüğü arasındaki denge, kullanıcı tarafından belirlenen düzeltme parametresi (C) ile ayarlanmaktadır.

$$f(x) = C \sum_{i=1}^m (\zeta^+ + \zeta^-) + \frac{1}{2} \|w\|^2 \quad (3.21)$$

Parametre ayarlamasının iç yapısında, aşağıdaki Lagrange fonksiyonu kullanılır.

$$\begin{aligned} Lp = & C \sum_{i=1}^L (\zeta^+ + \zeta^-) + \frac{1}{2} \|w\|^2 - \sum_{i=1}^L (\mu_i^+ \zeta^+ + \mu_i^- \zeta^-) - \sum_{i=1}^L a_i^+ (\epsilon + \zeta^+ + y_i - f(x_i)) \\ & - \sum_{i=1}^L a_i^- (\epsilon + \zeta^- - y_i - f(x_i)) \end{aligned} \quad (3.22)$$

$Lp$ 'deki  $w$ ,  $b$ ,  $\zeta^+$  ve  $\zeta^-$  ifadelerinin kısmi türevi alınarak;

$$\frac{\partial Lp}{\partial w} = 0 = w = \sum_{i=1}^N (a_i^+ - a_i^-)x_i \quad (3.23)$$

$$\frac{\partial Lp}{\partial b} = 0 = \sum_{i=1}^N (a_i^+ - a_i^-) = 0 \quad (3.24)$$

$$\frac{\partial Lp}{\partial \zeta^+} = 0 = C = a_i^+ + \mu_i^+ \quad (3.25)$$

$$\frac{\partial Lp}{\partial \zeta^-} = 0 = C = a_i^- + \mu_i^- \quad (3.26)$$

lagrange fonksiyonunda kullanılacak eşitlikler oluşturulur. Veri setinin koşulları sağlaması kabulüyle eşitlik (3.23) ve (3.24), eşitlik (3.22)'de yerine koyulur ve  $Lp$ ,  $a_i^+$  ve  $a_i^-$  ye göre (3.27) eşitliğinde gösterildiği şekilde maksimize edilir.

$$\begin{aligned} 0 \leq a_i^+, a_i^- \leq C \quad & \sum_{i=1}^L (a_i^+ - a_i^-) = 0 \\ Lp = \sum_{i=1}^L (a_i^+ - a_i^-)f(x_i) - \epsilon \sum_{i=1}^L (a_i^+ - a_i^-) - \frac{1}{2} \sum_{i,j}^L (a_i^+ - a_i^-)(a_j^+ - a_j^-)x_i \cdot x_j \end{aligned} \quad (3.27)$$

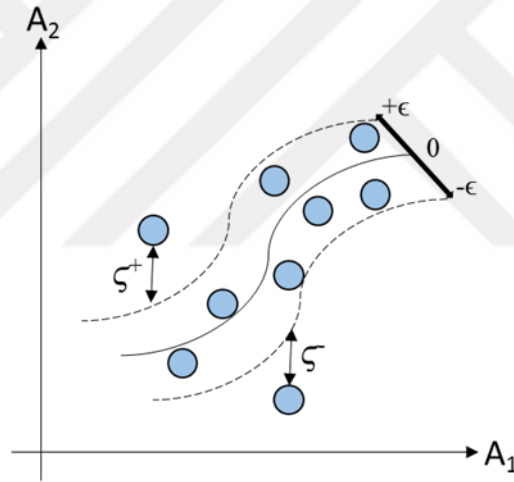
(3.23) eşitliği  $f(x) = w \cdot x_i + b$  formülünde yerine koyulursa tahmin fonksiyonu aşağıdaki gibi olur.

$$f(x) = \sum_{i=1}^L (a_i^+ - a_i^-)x_i \cdot x_j + b \quad (3.28)$$

Bu ifade destek vektör açılımı olarak isimlendirilmektedir. Yukarıdaki eşitliklerden görüldüğü üzere  $w$  ağırlıkları  $x_i$  örneklerinin doğrusal kombinasyonu şeklinde yazılabilmektedir. Böylece destek vektörler ile temsil edilen fonksiyonun karmaşıklığı girdi örneklerinin boyutundan bağımsız olacak ve sadece destek vektörlerin sayısına bağlı olacaktır (Yalçın, 2016).

### 3.2.5.2. Doğrusal olmayan destek vektör regresyonu

DVM ile sınıflandırma problemlerinin çözümündeki gibi doğrusal olmayan regresyon bölümünde de, girdi uzaydaki doğrusal olmayan (non-linear) noktalara ilişkin veri setleri daha üst boyuta taşınarak doğrusal verilere haritalanması söz konusudur (Şekil 3.24). Daha üstteki boyutlu uzaylar “özellik uzayı” olarak tanımlanmaktadır. Fakat boyut büyütme durumu bazı dezavantajlara neden olacaktır. Boyut sayısının büyümesi özellik uzayındaki hesaplama işlemlerini daha zor ve karmaşık yapacaktır. Ayrıca tahmin performansı, giriş uzayındaki boyutun artması nedeniyle düşecektir. Böyle durumları çözebilmek amacıyla alternatif kernel yöntemleri kullanılmaktadır (Kecman, 2001). Bu yöntemler hem doğrusal hem de doğrusal olmayan sınıflama ya da regresyon formu için genelleştirilebilir niteliktedir.



Şekil 3.24. Doğrusal olmayan bir destek vektör regresyon fonksiyonu

Kernel fonksiyonu  $K(x_i, x_j) = \phi(x_i) \cdot \phi(x_j)$  kullanılarak formüller doğrusal olmayan regresyon modeline uygulanabilir. Doğrusal olmayan DVR çözümünde hata toleransı ( $\epsilon$ ) dikkate alınarak (3.29) eşitliği elde edilir.

$$\begin{aligned} \max \frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N (a_i^+ - a_i^-)(a_j^+ - a_j^-) K(x_i, x_j) - \epsilon \sum_{i=1}^N (a_i^+ - a_i^-) - \sum_{i=1}^N y_i (a_i^+ - a_i^-) \\ \sum_{i=1}^N (a_i^+ - a_i^-) = 0 ; \quad 0 \leq a_i^+, a_i^- \leq C \end{aligned} \quad (3.29)$$

Optimizasyon problemlerinde Lagrange fonksiyonu kullanılarak yapılan çözümlemede doğrusal olmayan DVR modelleri için (3.30) eşitliği elde edilmiş olur.

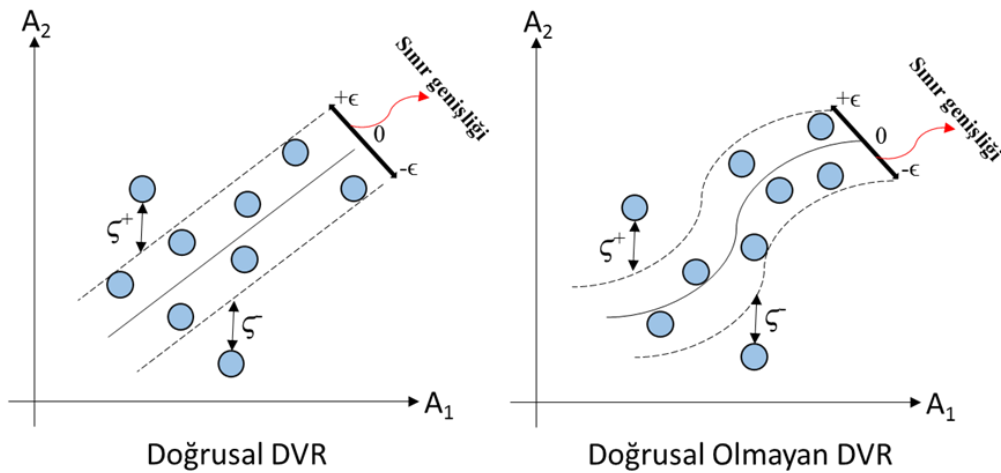
$$(x) = \sum_{i=1}^l (a_i^+ - a_i^-) K(x_i, x_j) + b \quad (3.30)$$

(3.28) ve (3.30) eşitlikleri birbirine benzerdir. Ancak burada ortaya çıkan en önemli fark; kernel fonksiyonunun,  $K(x_i, x_j)$ , doğrusal olmayan noktalara ilişkin tespitlerde kullanılmasıdır. DVR’de modelin tahmin performansı üzerinde etkili olan üç parametre bulunmaktadır. Bu parametreler sırasıyla;

- Hata toleransı ( $\epsilon$ )
- Kernel türü ve fonksiyonu
- Ceza/düzeltilme parametresi (C)

şeklinde. Dolayısıyla destek vektör regresyon modeli oluşturulurken model parametrelerinin ve kullanılan kernel denklemlerinin seçimi kritik bir konudur. Aşağıda tahmin performansını etkileyen üç parametrenin model üzerinde nasıl etkiye sahip olduğu belirtilmiştir.

Hata toleransı ( $\epsilon$ ), destek vektörlerin düzlem içerisinde oluşturduğu sınır genişliğini kontrol etme görevine sahiptir (Şekil 3.25). Eğer model tahmini sonucu değer, sınırın içinde kalıyorsa hata tolere edilmiş demektir ve kayıp fonksiyonundaki karşılığı “0” dır. Tam tersi sonuç çıkarsa hata tolere edilemez ve kayıp fonksiyonu tahmini değerle sınırın yarıçapı arasındaki uzunluk olur.



Şekil 3.25. Doğrusal ve doğrusal olmayan DVR’de sınır genişliği

Diğer taraftan hata toleransı değeri eğer büyük seçilirse, DVR modelini oluşturmak için daha az sayıda destek vektörü kullanılacaktır. Ancak tahmin sonuçlarının performansı düşecektir. Hata toleransı değeri küçük seçilirse, DVR modeli daha çok destek vektörü ile kurulacaktır. Ancak bu seferde tahmin sonuçları aşırı uyum gösterecektir. Sonuç olarak hata toleransı değerinin modele uygun olarak seçilmesi DVR performansı açısından önemlidir.

Kernel türü ve fonksiyonu seçiminde ilk olarak RTF ve polinom kernelin düşünülmesi oldukça uygundur. Sebebi, RTF ve polinom kernelin doğrusal ayrılamayan veri noktalarını daha yüksek boyutlu bir uzayda eşleştirmesidir. Bu nedenle doğrusal kernel fonksiyonunun tersine giriş ve çıkış özellikli verilerin doğrusal olmadığı durumlarda daha belirgin yargılara varabilmektedir (Debnath ve Takahashi, 2004; James ve ark., 2013). Polinom kernelde polinoma ait dereceler sürekli değişime uğrayacağından matematiksel anlamda problem çözümünü zorlaştırmaktadır. Buna karşın RTF kernelinde gamma ( $\gamma$ ) parametresindeki değişimin tahmin performansına etkisinin daha az olduğu tespit edilmiştir (Han ve ark., 2012).

RTF kernelinde iki parametre olan ( $C$  ve  $\gamma$ ) ile polinom kernelinde iki parametre olan ( $C$  ve  $d$ )'nin regresyon tahmin sonucuna etkisi kesinlikle söz konusudur. Ancak kullanılan hangi parametre çiftinin modelde en iyi sonucu verdiği bilinmemektedir. Çözüm için ilgili parametreler fonksiyon içerisinde işleme tabii tutularak en uygun olanının belirlenmesi gerekmektedir. Bu amaçla kullanılan birkaç yöntem mevcuttur. Yaygın olarak kullanılan yöntemlerden birisi eğitim verisini iki kısma ayırmaktır (Chang ve Lin, 2002). Birinci kısımda eğitim verisi doğrudan eğitime amaçlı kullanılırken ikinci kısımda modelin görmediği yani bilinmeyen veriler test amaçlı kullanılır. Bu sayede modelin tahmin doğruluğuna ait performansı daha hassas bir şekilde belirlenmiş olacaktır. Literatürde sözü geçen yöntem çapraz geçerlilik (cross-validation) olarak isimlendirilmektedir. ( $C$ ,  $\gamma$ ) veya ( $C$ ,  $d$ )'yı çapraz geçerlilikle belirlerken grid tarama yöntemi kullanılır. Grid yöntemindeki çalışma prensibi, parametre çiftlerinin çapraz geçerlilik yoluyla denenmesi ve en iyi çapraz geçerlilik doğruluğuna sahip olanın belirlenmesidir. İşlem için katlanarak büyüyen diziler ( $C=2^{-5}, \dots, 2^{15}$ ;  $\gamma=2^{-15}, \dots, 2^3$ ) kullanılmaktadır (Ito ve Nakano, 2003).

Kısaca özetlenecek olursa destek vektör regresyonu son yıllarda kullanımının gittikçe artmasının yanı sıra, örneklerden öğrenme yapabilmesi, genelleme yapabilmesi ve istatistiksel yöntemlere kıyasla daha az varsayımları içermesinden dolayı analizlerde çalışma yöntemleri olarak seçilmiştir.

### 3.2.6. Performans analiz yöntemleri

Tez kapsamında kullanılan yöntem ile elde edilen model tahmin değerleri ve uluslararası GNSS servislerinin yayınlamış olduğu gerçek gecikme değerlerini karşılaştırmak üzere Ortalama Mutlak Hata (OMH), Ortalama Mutlak Yüzde Hata (OMYH), Karesel Ortalama Hata (KOH) ve  $R^2$  değeri sonuçları performans analizlerinde kullanılmıştır. Özellikle yöntemlerin karşılaştırıldığı çalışmalarda bu hata oranları ile modelin performansı incelenmektedir (Lin, 2010; İlhan ve Şener, 2013). OMH değeri, gerçek ile tahmin edilen veri arasındaki farkın mutlak değerlerinin ortalaması (3.31) eşitliği kullanılarak hesaplanır. OMYH, tahmin hatalarının mutlak değerlerinin gerçek değerlerine oranının ortalaması (3.32) eşitliği kullanılması sonucu hesaplanan hatadır. KOH, tahmin hatalarının kareleri alınıp ortalamalarının karekökten çıkarılmak suretiyle (3.33) eşitliği ile hesaplanan hata türüdür.  $R^2$  ise, regresyon modeli için bağımsız değişkenler tarafından açıklanan bağımlı değişkendeki değişim oranını (3.34) eşitliği ile ölçen değerdir.

$$OMH = \frac{1}{n} \sum_{i=1}^n |y_i - \bar{y}_i| \quad (3.31)$$

$$OMYH = \frac{\sum_{i=1}^n \frac{|y_i - \bar{y}_i|}{y_i}}{n} * 100 \quad (3.32)$$

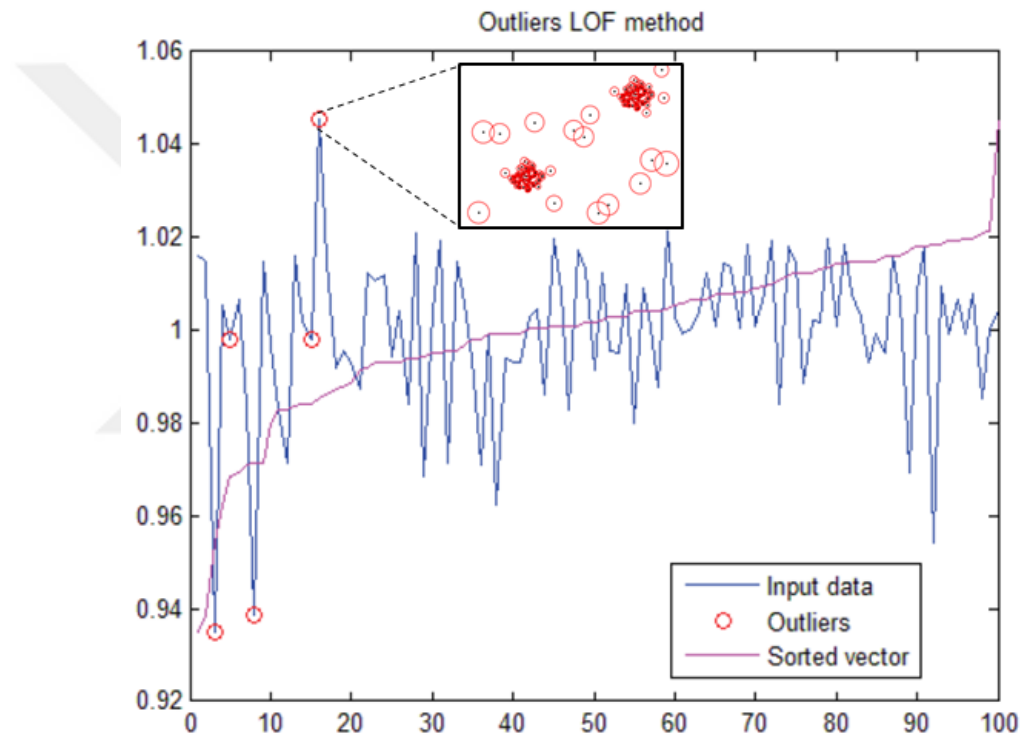
$$KOH = \pm \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \bar{y}_i)^2} \quad (3.33)$$

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \bar{y}_i)^2}{\sum_{i=1}^n (\bar{y}_i - \hat{y})^2} \quad (3.34)$$

$y_i$  = Gerçek Gecikme Değeri,  
 $\bar{y}_i$  = Modeldeki Gecikme Değeri,  
 $\hat{y}$  = Ortalama Gecikme Değeri,  
 $i = 1, 2, \dots, n, i = \text{Günler}$

### 3.2.7. Çok deęişkenli aykırı gözlem analizi

Çok deęişkenli aykırı gözlem analizlerinden birisi olan ve çalışmada kullanılan yerel aykırı deęer faktörü (LOF: Local Outlier Factor), Breunig ve ark. (2000) tarafından önerilmiştir. Dağılımdaki veri noktalarının yoğunluğunu, aykırı deęerleri tespit etmek için anahtar faktör olarak kullanan, denetimsiz bir makine öğrenme algoritmasıdır. Adından da anlaşılacağı gibi, yerel yoğunluk yaklaşımı, k-en yakın komşular aracılığıyla hesaplanır ve nispeten düşük yoğunluklu alanlarda bulunan veri noktaları, aykırı deęerler olarak sınıflandırılır (Şekil 3.26).



Şekil 3.26. LOF teknięi ile belirlenmiş olan aykırı gözlemler (Cateni ve ark., 2008)

Dięer algoritmalar aykırı deęer için ikili bir sınıflandırma üretirken, LOF ayrıca bir noktanın aykırı olduęu dereceyi de üretebilir. Bu yönüyle, LOF algoritması dört kısma ayrılmaktadır. Sırasıyla,

- K-mesafe ve k-komşu ( $k_d$ ),
- Erişilebilirlik mesafesi ( $rd_k$ ),
- Yerel erişilebilirlik yoğunluęu ( $lrd_k$ ) ve
- LOF hesaplanması ( $LOF_k$ )

olarak birbirini takip etmektedir.

İlk olarak k-mesafe ve k-komşu bölümü ile gözlemler arasındaki mesafeyi bulmak için k hiper parametresi belirlenmektedir. Buradaki k değeri kaç tane komşuya bakılacağını ifade eder. Erişilebilirlik mesafesi, iki nokta arasındaki mesafe ve ikinci noktanın k-mesafesinden maksimum olanını ifade eder (3.35). Burada iki nokta arasındaki mesafe öklid, minkowski, manhattan ya da diğer uzaklık ölçülerinden herhangi biri olabilir.

$$rd_k(a, b) = \max\{k_d(b), \text{normal-mesafe}(a, b)\} \quad (3.35)$$

Yerel erişilebilirlik yoğunluğu, bir sonraki noktaya veya noktalar kümesine ulaşmak için bulunan noktadan ne kadar uzağa gidilmesi gerektiğini ifade eder. Hesaplanan erişilebilirlik mesafesi yerel erişilebilirlik yoğunluğunun hesaplanmasında kullanılır. (3.36) eşitliğinde, tüm k en yakın komşuları için erişilebilirlik mesafesi hesaplanır ve bulunan değerler toplanır ve k değerine bölünür. Bu değer tersi alındığında da aranan yoğunluk hesaplanmış olur.

$$lrd_k(A) = 1/(\text{toplaml}(rd_k(a, n))/k) \quad (3.36)$$

Buradaki n, k'ya kadar olan komşuları ifade etmektedir. Bulunan yerel erişilebilirlik yoğunlukları a'nın en yakın k komşusunun yerel erişilebilirlik yoğunluklarıyla kıyaslanır. (3.37) eşitliği ile her bir komşunun yoğunluğu toplanıp a'nın yoğunluğuna bölünür. Bulunan değer komşu sayısı yani k'ya bölünür.

$$LOF_k(A) = [(lrd_k(1. \text{komşu}) + lrd_k(2. \text{komşu}) + \dots + lrd_k(k. \text{komşu})) / lrd_k(a)] / k \quad (3.37)$$

Burada,

$LOF_k(A) \sim 1$ , komşularla benzer yoğunluk anlamına gelir.

$LOF_k(A) < 1$ , komşulardan daha yüksek yoğunluk anlamına gelir (Aykırı olmayan).

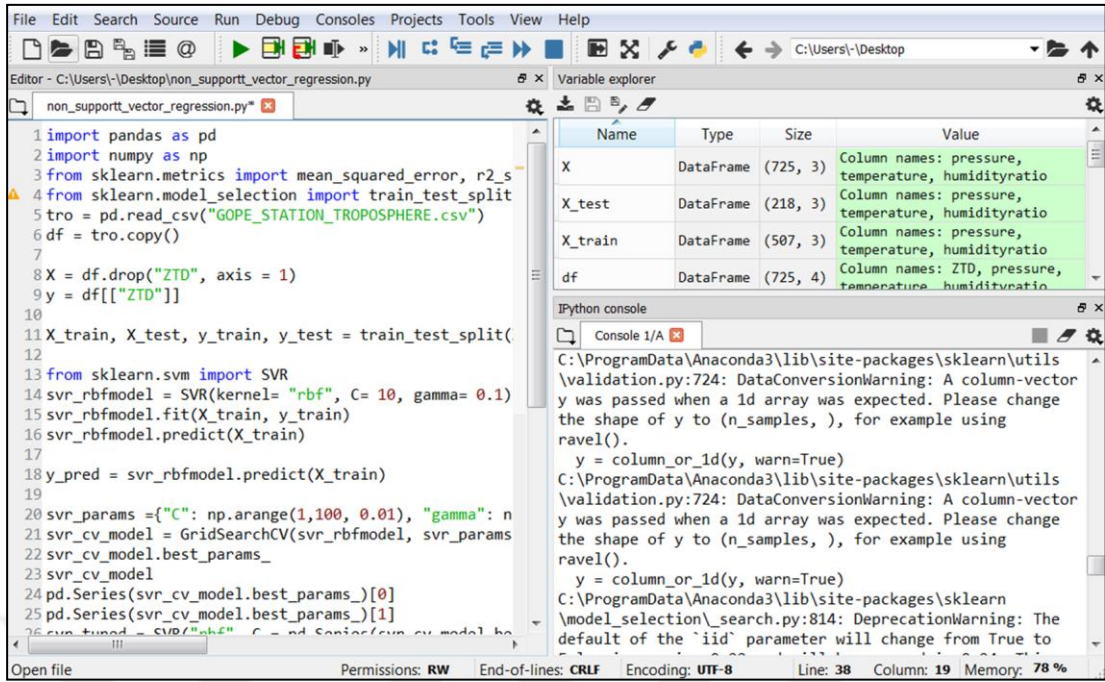
$LOF_k(A) > 1$ , komşulardan daha düşük yoğunluk anlamına gelir (Aykırı).

#### 4. ARAŞTIRMA SONUÇLARI VE TARTIŞMA

DVR modelleri sonucunda elde edilen ZTD değerlerinin, hem IGS tarafından yayınlanan gerçek ZTD ürünleri hem de online PPP servisi kapsamındaki CSRS-PPP’de kestirilen ZTD değerleri ile karşılaştırılabilmesi için üç farklı veri setinin oluşturulması gerekmiştir. Birinci veri seti, DVR modellerini kurmak için gerekli olan girdi ve çıktı değişkenlerinden oluşmaktadır. İkinci veri seti, DVR model kestirim sonuçlarının ve VMF veri merkezindeki (VMF1) gerçek gecikme değerlerinin karşılaştırılması için IGS’in yayınladığı analiz sonuçlarından oluşmaktadır. Üçüncü veri seti, yine karşılaştırmak için CSRS-PPP’ye gönderilen verilerin değerlendirilmesi sonucu elde edilmiş olan troposfer rapor sonuçlarından oluşmaktadır. Yeni oluşturulacak olan DVR modellerinde kullanmak üzere; hazırlanan veri setindeki bütün bilgiler, sayısal formda olup veride herhangi bir işlem (normalizasyon, dönüşüm vb.) yapılmamıştır. Bu konuda, ham verilerle çözüm üretmek istenmesi ve verilerin çalışma için uygun nitelikte olması etken olmuştur.

DVR ile yapılan modelleme işleminde; doğrusal, polinomal, radyal tabanlı kernel fonksiyonları kullanılmıştır. Sigmoid kernel fonksiyonu yapısı gereği (0 ve 1 arasında döngü) çalışmaya uygun olmamasından dolayı tercih edilmemiştir. Fonksiyonlar için gerekli parametreler çapraz geçerlilik yöntemi ile eğitim sırasında tespit edilmiş ve parametre değerlerindeki değişimlerin tahmin doğruluğuna etkileri araştırılmıştır. Modellerin en iyi parametrelerle kurulması amacıyla parametreler arasında, optimizasyon işlemi gerçekleştirilmiştir. Optimizasyon sonucu tespit edilen en iyi parametrelerle troposferik gecikme değeri tahmin edilmiştir. Sonrasında elde edilen tahmin sonuçlarına ait performans analizleri detaylıca incelenmiştir.

Modellerin kurulum aşamasının gerçekleştirilmesi ve optimizasyon çalışması için Python v.3.8 programlama dili kullanılmıştır. Bu dildeki *numPy*, *scikit-learn*, *matplotlib*, *pandas*, *seaborn*, *statsmodels*, *sciPy* kütüphanelerinden yararlanılmaktadır (Şekil 4.1). Ayrıca modelin kurulmasında büyük rol oynayan veri setlerinin, model yapısını oluşturmadaki izleyeceği yolun ortaya çıkarılması ve sürecin hangi aşamasında, nelere etki ettiğinin bilinmesinde de Python v.3.8 yazılımı değerlendirmeye dâhil edilmiştir. Uygulamadaki temel süreç; veri setinin değerlendirilmesi, modelleme ve optimizasyon, sonuçların yorumlanması şeklinde birbirini takip etmektedir.



Şekil 4.1. Model kurulum ve değerlendirmede kullanılan Python v.3.8 yazılımı

Model oluşturmada kullanılacak veri setlerinden anlamlı sonuçların elde edilmesi, faydalı bilgilerin bulunması, model sonuçlarının doğru yorumlanması ve karar vermeyi desteklemesi üzerine verilerin incelemesi çok önemlidir. Bu nedenle çalışmada veri setlerinin barındırdığı her bir değişkenin betimsel istatistikleri en başta değerlendirilmiş ve istatistiksel veri analizi gerçekleştirilmiştir.

#### 4.1. Veri Setinin İstatistiksel Analizleri

Verinin incelenmesinde kullanılan istatistiki analizler, farklı kullanım alanlarına sahip olurken, çeşitli başlıklar altında farklı teknikleri kapsayan çok yönlü yaklaşımları bünyesinde bulundurmaktadır. Günümüzde veri analizi, kararları daha akademik veya bilimsel forma dönüştürmede, sistemlerin en etkin bir şekilde çalışmasına yardımcı olmada ve veri yapısını kolay tanımda kritik rol oynamaktadır.

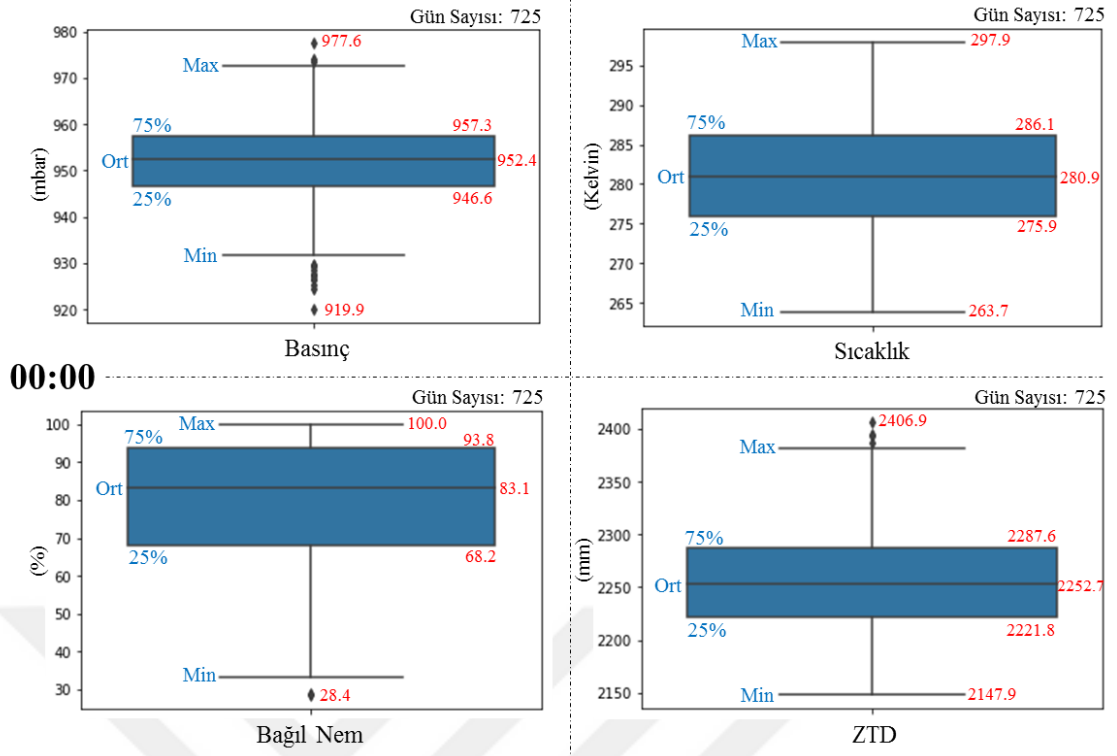
Bu bölümde modellemeye geçmeden önce kullanılacak olan veri seti ve bu veri setindeki değişkenler ile ilgili yapılması gereken analizler gerçekleştirilmektedir. Genellikle veri setini modele entegre ederken ne yapılabileceği ile ilgili bir kararsızlık mevcuttur. Bu tür belirsizliklerde nereye bakmalı, hangi noktaya odaklanmalı ve yapısının nasıl olduğu konularında çözüm üretmek amacıyla yapılmıştır. İlk olarak çalışmada kullanılacak olan veri setindeki değişkenlerin veri yapısı tanımlanmıştır (Çizelge 4.1).

**Çizelge 4.1.** Veri setine ait bilgiler

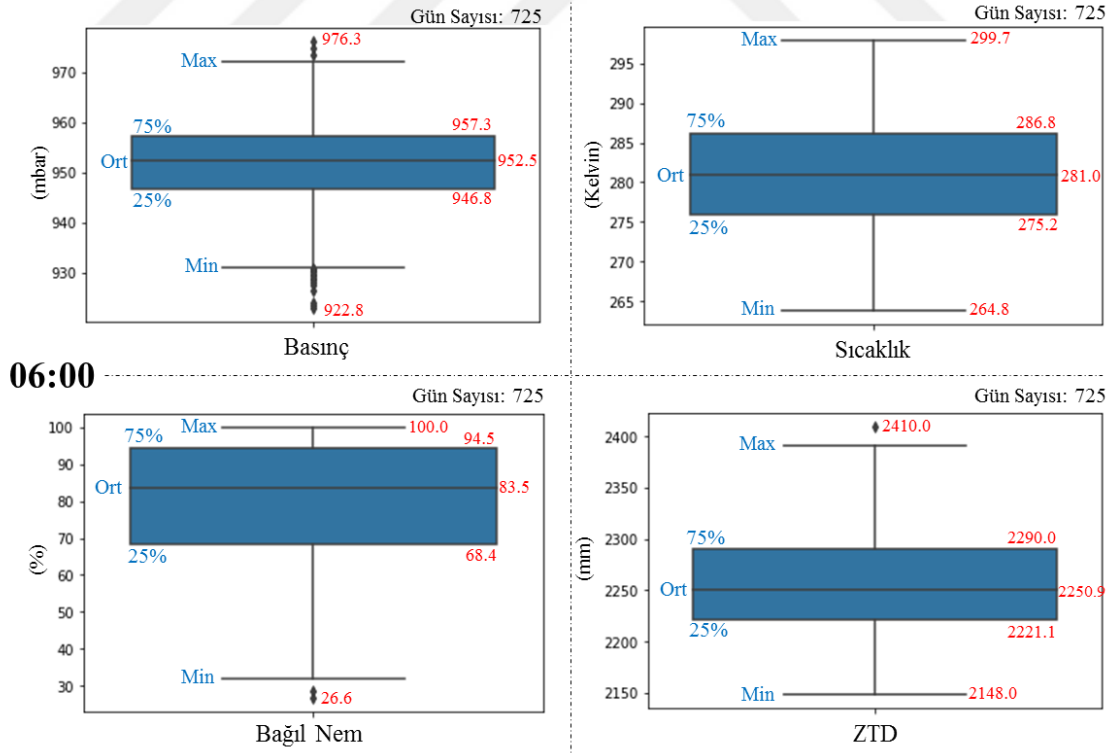
| <b>00:00 (UTC)</b>                                                                                            |                      | <b>06:00 (UTC)</b>                                                                                            |                      |
|---------------------------------------------------------------------------------------------------------------|----------------------|---------------------------------------------------------------------------------------------------------------|----------------------|
| <class 'pandas.core.frame.DataFrame'><br>RangeIndex: 725 entries, 0 to 724<br>Data columns (total 4 columns): |                      | <class 'pandas.core.frame.DataFrame'><br>RangeIndex: 725 entries, 0 to 724<br>Data columns (total 4 columns): |                      |
| <b>Basınç (mbar)</b>                                                                                          | 725 non-null float64 | <b>Basınç (mbar)</b>                                                                                          | 725 non-null float64 |
| <b>Sıcaklık (K)</b>                                                                                           | 725 non-null float64 | <b>Sıcaklık (K)</b>                                                                                           | 725 non-null float64 |
| <b>Bağımlı Nem (%)</b><br><b>Su Buharı Basıncı</b>                                                            | 725 non-null float64 | <b>Bağımlı Nem (%)</b><br><b>Su Buharı Basıncı</b>                                                            | 725 non-null float64 |
| <b>ZTD</b>                                                                                                    | 725 non-null float64 | <b>ZTD</b>                                                                                                    | 725 non-null float64 |
| dtypes: float64(4)<br>memory usage: 22.8 KB                                                                   |                      | dtypes: float64(4)<br>memory usage: 22.8 KB                                                                   |                      |
| <b>12:00 (UTC)</b>                                                                                            |                      | <b>18:00 (UTC)</b>                                                                                            |                      |
| <class 'pandas.core.frame.DataFrame'><br>RangeIndex: 723 entries, 0 to 722<br>Data columns (total 4 columns): |                      | <class 'pandas.core.frame.DataFrame'><br>RangeIndex: 723 entries, 0 to 722<br>Data columns (total 4 columns): |                      |
| <b>Basınç (mbar)</b>                                                                                          | 723 non-null float64 | <b>Basınç (mbar)</b>                                                                                          | 723 non-null float64 |
| <b>Sıcaklık (K)</b>                                                                                           | 723 non-null float64 | <b>Sıcaklık (K)</b>                                                                                           | 723 non-null float64 |
| <b>Bağımlı Nem (%)</b><br><b>Su Buharı Basıncı</b>                                                            | 723 non-null float64 | <b>Bağımlı Nem (%)</b><br><b>Su Buharı Basıncı</b>                                                            | 723 non-null float64 |
| <b>ZTD</b>                                                                                                    | 723 non-null float64 | <b>ZTD</b>                                                                                                    | 723 non-null float64 |
| dtypes: float64(4)<br>memory usage: 22.7 KB                                                                   |                      | dtypes: float64(4)<br>memory usage: 22.7 KB                                                                   |                      |

Belirli saat aralıkları için oluşturulmuş olan veri setleri, 00:00 ve 06:00 saatlerinde 725 güne ait değişkenleri tanımlıyorken, 12:00 ve 18:00 saatlerinde 723 güne ait değişkenler tanımlanmıştır. Bunun sebebi saat 12:00-18:00 için ölçülen meteorolojik verinin veya zenit troposferik gecikme değerinin iki gün için eksik olması anlamına gelmektedir.

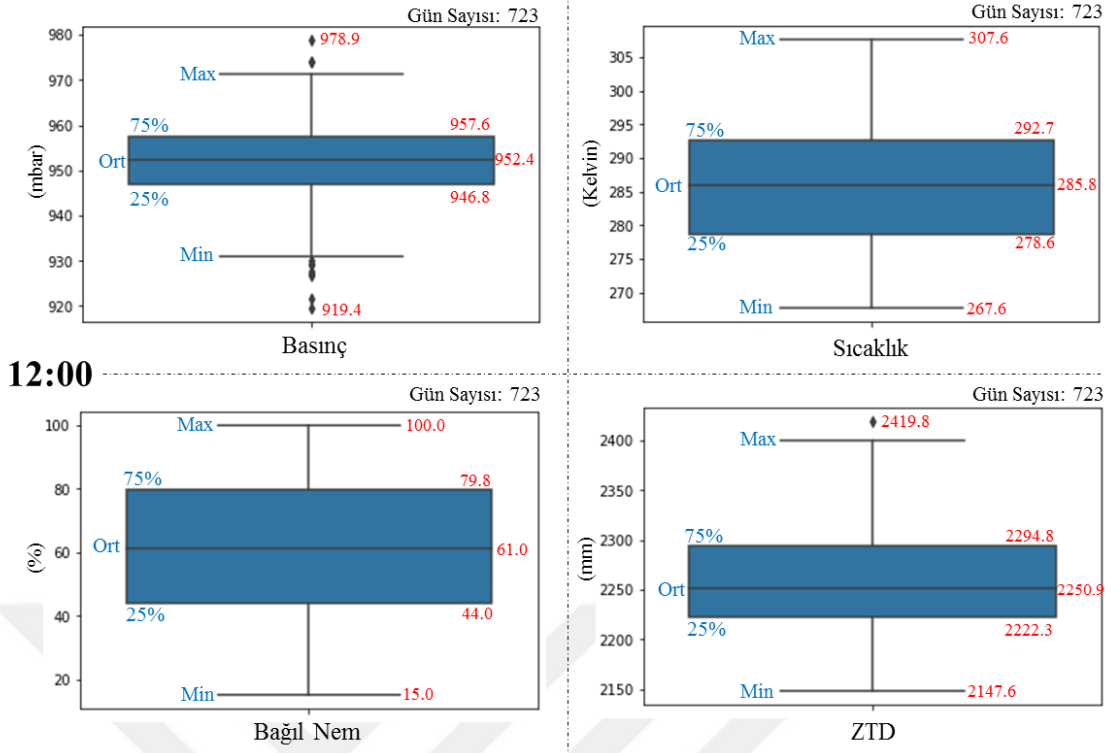
Aşağıda değişkenlere ait istatistiksel analiz sonuçları kutu grafiği oluşturularak gösterilmiştir (Şekil 4.2, Şekil 4.3, Şekil 4.4, Şekil 4.5). Kutu grafiklerine bakıldığında, 725 gün boyunca 00:00-06:00 saatlerindeki basınç, sıcaklık, bağımlı nem ve ZTD değerlerinden oluşan veri seti için maksimum değer, minimum değer, çeyreklikler (25%-75%), ortanca değer bilgilerinin ayrı bir şekilde hesaplandığı görülmektedir (Şekil 4.2, Şekil 4.3). Aynı durum 723 gün boyunca 12:00-18:00 saatlerinde elde edilmiş veri seti içinde geçerlidir (Şekil 4.4, Şekil 4.5).



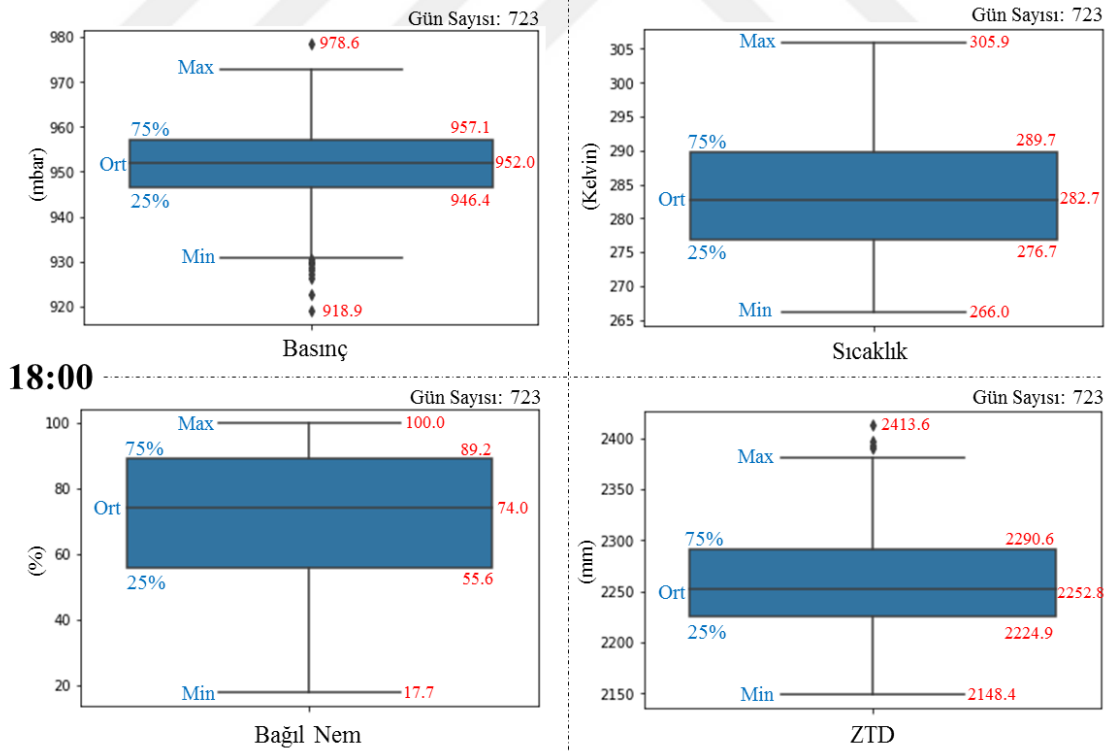
Şekil 4.2. Saat 00:00'a ait veri setindeki değişkenlerin istatistiki bilgileri



Şekil 4.3. Saat 06:00'a ait veri setindeki değişkenlerin istatistiki bilgileri



Şekil 4.4. Saat 12:00'a ait veri setindeki değişkenlerin istatistiki bilgileri

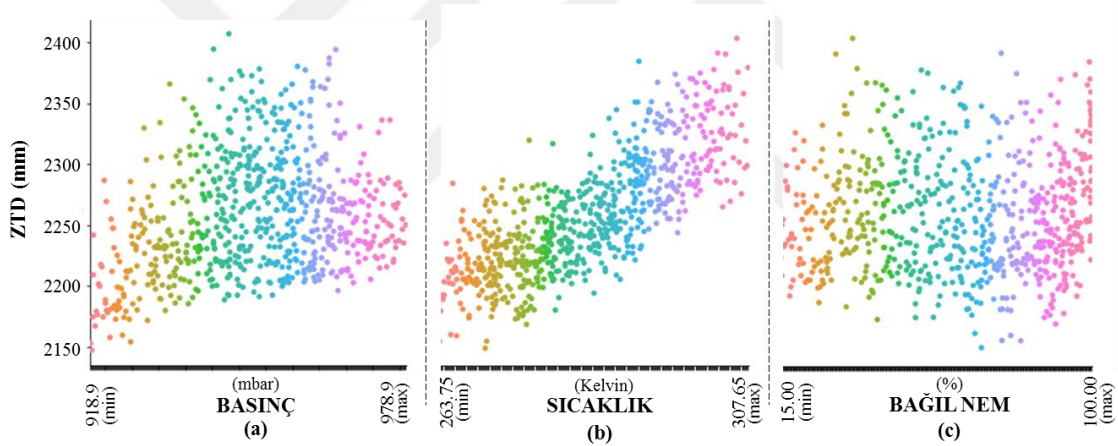


Şekil 4.5. Saat 18:00'a ait veri setindeki değişkenlerin istatistiki bilgileri

Oluşturulan kutu grafiklerindeki bazı değişkenlerde uç sınırların dışında kalan sayılı siyah noktalar bulunmaktadır. Bu noktalar gözlem yapılan gündeki verinin genel dağılıma uymadığını yani aykırı olduğunu ifade etmektedir. Normalde aykırı olan verilerin veri setinden çıkartılması gerekir, ancak çalışma kapsamında ani hava değişiminden kaynaklı etkilerin olması, küresel ısınma sonucu hava şartlarında beklenilmeyen koşulların gerçekleşmesi gibi etmenlerden dolayı bireysel bazda aykırı veri çıkartılmamıştır. Bunun yerine çok değişkenli aykırı gözlem analizi kullanılmıştır.

#### 4.2. Meteorolojik Veriler-ZTD İlişkisinin Değerlendirilmesi

Meteorolojik verilerdeki her bir bağımsız değişkenin zenit troposferik gecikme üzerindeki etkisini görmek amacıyla, her bir ölçü günündeki değişkene ait veriler kullanılmıştır (Şekil 4.6).



Şekil 4.6. Basınç(a), Sıcaklık(b) ve Bağıl nem(c) değişkenlerinin ZTD ile ilişkisi

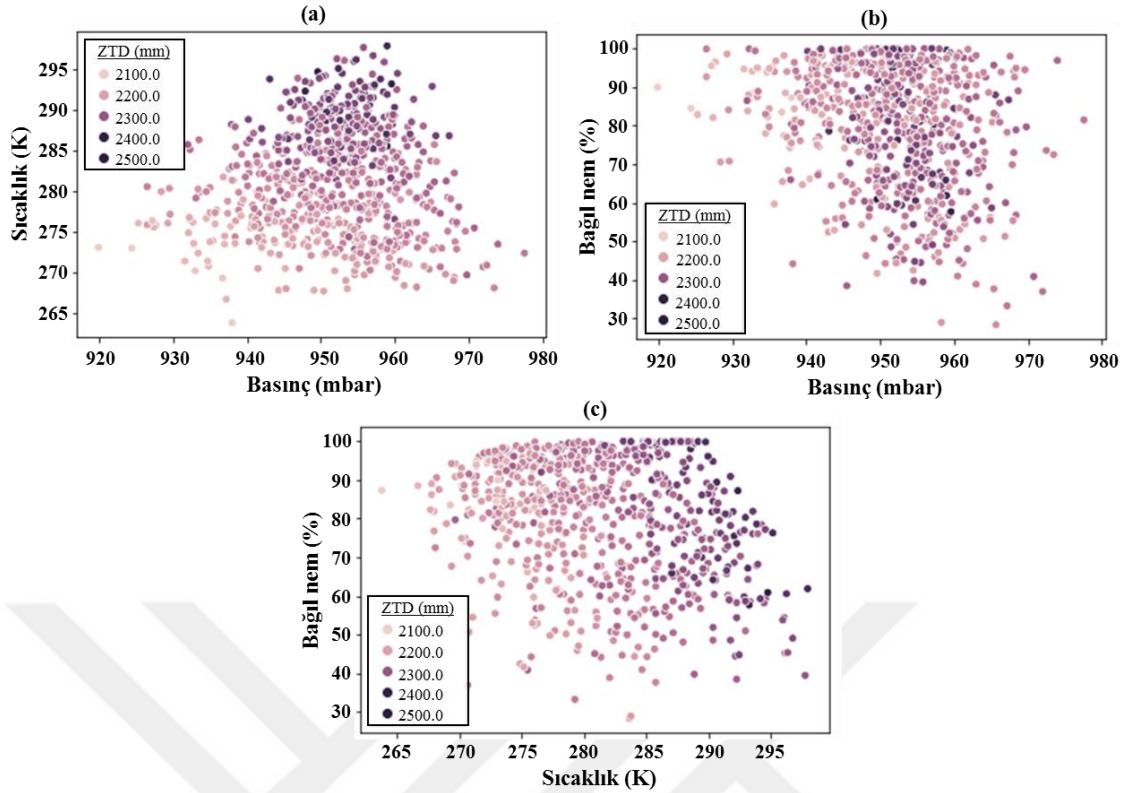
Grafiklerde sıcaklık etkeni baz alınırsa değerler maksimum seviyelere doğru yaklaştığı sırada troposferde gerçekleşen sinyal kırılmaları ve buna bağlı olarak gecikme etkisi de artmaktadır (Şekil 4.6b). Dolayısıyla yüksek sıcaklığa sahip zaman diliminde elde edilecek zenit troposferik gecikme değerinin düşük sıcaklıktaki zaman diliminde elde edilene göre daha fazla olduğu çıkarımı yapılabilmektedir. En genel ifadesi ise sıcaklık ile ZTD arasında doğrusal bir ilişki olduğu şeklindedir. Durum bağıl nem ve basınç etkeninde farklı olma eğilimindedir. Bağıl nem, havadaki su buharı miktarı olarak ifade edilirken, zamansal ve mekânsal dağılımı normal değil yani heterojen, ani değişkenlik gösteren ve önceden tahmin edilmesi mümkün olmayan bir yapıdadır (Şekil 4.6c). Basınç değişkeni de bağıl nem gibi normal dağılım

sergilememekte ancak kıyaslandığında birazda olsa sıcaklığa benzer şekilde saçılıma karşılık gelmektedir. Basınç değerinin artması durumunda gecikme etkisinde de artış belirginleşmektedir. Fakat yine de en yüksek gecikme etkileri ortalama basınç değeri aralığında gözlemlenmektedir (Şekil 4.6a). Değişkenler arasındaki ilişkiden yola çıkılarak tez kapsamında şu çıkarım yapılabilir: Çeşitli deneysel troposfer modelleri ile elde edilen troposfer gecikmelerin hesabında kullanılan parametreler mevsime ait ortalama değerler olduğu için sonuçlarda sapmaların yaşanması muhtemeldir. Çözüm olarak anlık ölçüm yapılan gerçek değerlerle sonuç elde etmek daha makul yaklaşımdır.

#### 4.3. ZTD'deki Mevsimsel Etkilerin İstatistiki Değerlendirilmesi

İstatiksel veri değerlendirmesi veya keşifçi veri analizi uygulamaları; model performanslarını optimize etme, doğru model seçimi, modeldeki parametrelerin doğru belirlenmesi, yüksek başarıda model sonuçları, modellerin daha açık yorumlanması, hatayı minimize etme vb. konulara yardımcı olmaktadır. Bunun için model oluşturmadan önce verilerin/değişkenlerin, iyi bir şekilde istatistiki anlamda tanınması ve aralarındaki ilişkilerin detaylıca irdelenmesi gerekmektedir. Çalışmada, daha yüksek başarı performansı ile model oluşturulması, verilerin model üzerinde nasıl kullanılırsa en etkin ve verimli şekilde fayda sağlayacağı hakkında bilgi sahibi olunması, birbirleri arasındaki ilişkinin sonuca etkisinin ne olduğunun bilinmesi amacıyla veriler/değişkenler arasında ilişkisel korelasyon çaprazlamaları gerçekleştirilmiştir.

Korelasyon, değişkenler arasındaki bağlantıyı açıklayan istatiksel bir terim olup söz konusu değişken örneklemeleri arasındaki ilişkiyi inceler. Ardından ilişkileri değerlendirerek veri seti içerisinde ortaya çıkan yapı hakkında bilgi sağlar. Önceki bölümde yapılan değerlendirmeler sonucu, meteorolojik ve ZTD değişkenleri arasındaki ilişki tek tek açıklanmıştı. Ancak şimdi saçılım grafiği yaklaşımıyla üçüncü boyuta göre birlikte çaprazlama yapılarak çoklu değişkenler arası korelasyon ilişkisi incelenmektedir (Şekil 4.7).

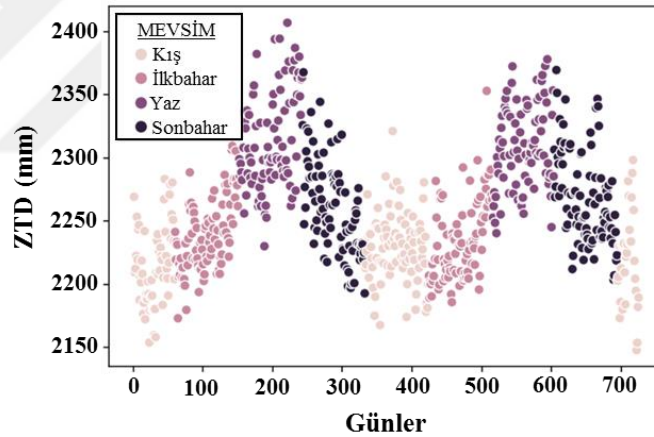


Şekil 4.7. Değişkenler arasındaki korelasyon çaprazlamaları

Şekil 4.7a için sıcaklık değerinin yüksek, basınç değerinin ortalama seviyelerde (950-960 mbar) olduğu durum troposferik gecikmenin en yüksek olduğu kesimi işaret etmektedir. Sıcaklık ve basınçın düşük olduğu aralıkta, troposferik gecikmenin az olduğu görülmektedir. Basınç değerinin genellikle ortalama değerlerde yoğunlaşmasının sebebi çalışmada seçilmiş istasyona ait tek bir yükseklikte ölçümlerin gerçekleşmesidir. Dolayısıyla basınç-yükseklik ilişkisi düşünüldüğünde buradaki basınç değeri değişiminin yükseklik farklarından değil sadece hava koşullarından kaynaklı olduğu göz ardı edilmemelidir. Şekil 4.7b’de basınç ve bağıl neme bağlı troposferik gecikme etkisi incelendiğinde, yine basınç değerinin ortalama seviyelerde olduğu durumda meydana gelen gecikmenin en yüksek olduğu kesimleri işaret etmektedir. Ancak daha önceden belirtildiği gibi bağıl nemin hava olaylarındaki düzensiz dağılımdan kaynaklı net bir aralığı söz konusu değildir. Bu yönüyle söz konusu bağıl nemin makine öğrenme modeli oluşturmada kullanılacak olması, hem avantaj sağlarken hem de dezavantaj oluşturmaktadır. Şöyle ki standart troposfer modellerinde bağıl nem için deneye dayalı ortalama değerler kullanılırken çoğu zaman düzensiz oluşundan sebep, hesaba katılamamakta ve tam olarak doğru modellenememektedir. Yeni oluşturulacak makine öğrenmesi modelinde ise durum farklıdır. Modelde bir zaman dilimi için belirlenmiş

olan bağıl nem değeri troposfer gecikme üzerinde girdi verisi olarak kullanılacağı için makinenin öğrenme kısmında farklı kombinasyonlar oluşacaktır. Yani herhangi bir bağıl nem değerinde oluşabilecek troposfer gecikme değeri, o andaki ölçülmüş olan bağıl nem değerine göre makineye öğretilmiş olacaktır. Bu yaklaşım avantaj olarak sayılabilir ancak bağıl nem ve troposfer gecikme arasındaki korelasyonun düşük olması model başarısına olumsuz etki ettiğinden dezavantajlı yönü de bulunmaktadır. Son olarak ise Şekil 4.7c’de sıcaklık ile bağıl nem arasında güçlü korelasyon ilişkisi olduğu açıkça görülmektedir. Sıcaklığın artış gösterdiği aralıkta bağıl nemin de artmasıyla gecikme etkisinin en fazla olduğu gözlemlenmektedir.

GNSS gözlemlerinin analizi sonucu elde edilen ZTD değerlerinde hava şartları önemli faktör olması sebebiyle mevsimsel etkiler önemli bir yere sahiptir. Bunun için seçilen istasyonda iki yıl boyunca elde edilen ZTD değerleri ile zaman serisi grafiği oluşturulmuştur. Oluşturulan seride uzun dönem için meydana gelen ZTD değişimlerinin hangi mevsimde ne kadar etkilendiği incelenebilmektedir (Şekil 4.8).



Şekil 4.8. Mevsimsel etkilere göre ZTD’nin değişimi

Aralık, Ocak ve Şubat ayından oluşan kış mevsiminde gecikme etkisinin en düşük olduğu görülmektedir. En yüksek gecikme etkisi Haziran, Temmuz ve Ağustos ayından oluşan yaz mevsimindedir. Diğer mevsimler ise geçiş evresi (gecikme etkisinin artıp azalması) olarak düşünülebilir.

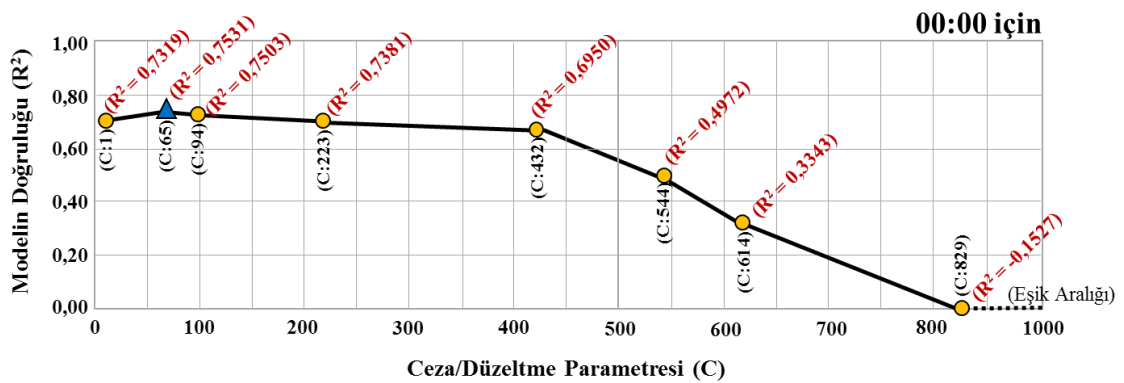
Bu değerlendirmeler, model kurmadan önce; doğru, güncel, hassas ve kaliteli veri toplayarak troposfer gecikme kestirimi konusunda doğru karar vermeye katkı sağlamaktadır. Ayrıca verilerin/değişkenlerin aralarındaki ilişkilerin detaylıca irdelenmesi çalışmadaki ana hedefe daha çok yaklaştıracaktır.

#### 4.4. Zenit Troposferik Gecikmenin DVM ile Kestirimi

Tez çalışması kapsamında, GOPE referans istasyonuna ait iki yıllık verileri kullanarak DVM modelleri ile zenit troposferik gecikmenin kestirimi gerçekleştirilmiştir. DVM modeli ile ilgili zamanlar (epok) için yapılan kestirim işleminde; doğrusal (linear), polinomal (polynomial), radyal tabanlı (radial-basis) kernel fonksiyonları kullanılmıştır. Fonksiyonlar yardımıyla kurulan modellerin optimizasyonu ise k-katlı ( $k=10$ ) cross-validation ve grid tarama yöntemleri kullanılarak yapılmıştır. Bu model optimizasyon yöntemleri ile parametre değerlerindeki değişimlerin model doğruluğuna etkileri araştırılmıştır. Ayrıca optimize edilmiş parametrelerle troposferik gecikme hatası kestirilmiş ve elde edilen kestirim sonuçları analiz edilerek performansları değerlendirilmiştir.

##### 4.4.1. Doğrusal (Linear) kernel fonksiyonu ile modelleme

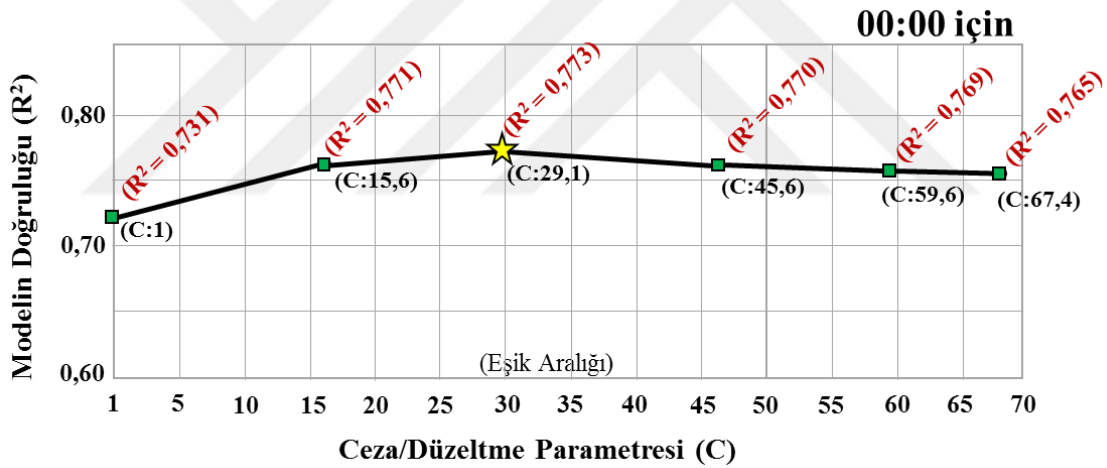
Doğrusal (linear) kernel kullanımı için kullanıcı tarafından belirlenmesi gereken ilk parametre ceza/düzeltilme ( $C$ ) parametresidir. Optimizasyon yöntemleri ile bu parametre, oluşturulacak olan gecikme tahmin modeli için eğitim ve test verileri kullanılarak tespit edilmiştir. İlk aşamada  $C$  parametresi, optimize edilmeden önce varsayılan ayarı “ $C=1$ ” olacak şekilde model kurularak eğitim için  $R^2$  değeri (0,7319) bulunmuştur. Optimize işlemi esnasında farklı eşik aralıklarında  $C$  değeri 1 arttırılarak denenmiş ve sonucunda uygun parametrenin “ $C=65$ ” eşik aralığında olduğu tespit edilmiştir. Seçilmiş aralıklarda farklı  $C$  düzeltme parametresinin model doğruluğunda meydana getirdiği etkiler,  $R^2$  değeri üzerinden aşağıda gösterilmiştir (Şekil 4.9).



Şekil 4.9. Farklı ( $C$ ) ceza/düzeltilme parametrelerine ait model doğrulukları

C'nin 65 olduğu durumda model en yüksek  $R^2$  değerine (0,7531) ulaşmıştır. Modelin en düşük  $R^2$  değerine (-0,1527) sahip olduğu durumda ise C parametresi 829 olarak belirlenmiştir. Böylelikle C parametresinin küçük eşik aralığında daha iyi sonuçlar verdiği söylenebilir ancak en iyi C parametresinin bulunan 65 değeri olduğu hala belirsizdir.

Bunun için ikinci aşamada grid tarama yöntemi ile ilk aşamada belirlenen parametrelerin yakınında daha küçük bir aralıkta ikinci bir arama yapılarak daha iyi parametre değerlerine ulaşılmaya çalışılmıştır. C parametresi, eşik değeri 1-70 arasında 0,1 artışla denenmiştir. 1-70 eşik aralığında denenme sebebi, hem bu aşamaya kadar uygun bulunan C=65 değerinin bu ölçek aralığında bulunması hem de optimize süresinin çok fazla uzun sürmesinin önüne geçmektir. Aşağıda daha dar ölçekte C parametresinin model doğruluğuna olan etkileri,  $R^2$  değeri üzerinden gösterilmiştir (Şekil 4.10).



Şekil 4.10. Farklı (C) ceza/düzeltilme parametrelerine ait model doğrulukları

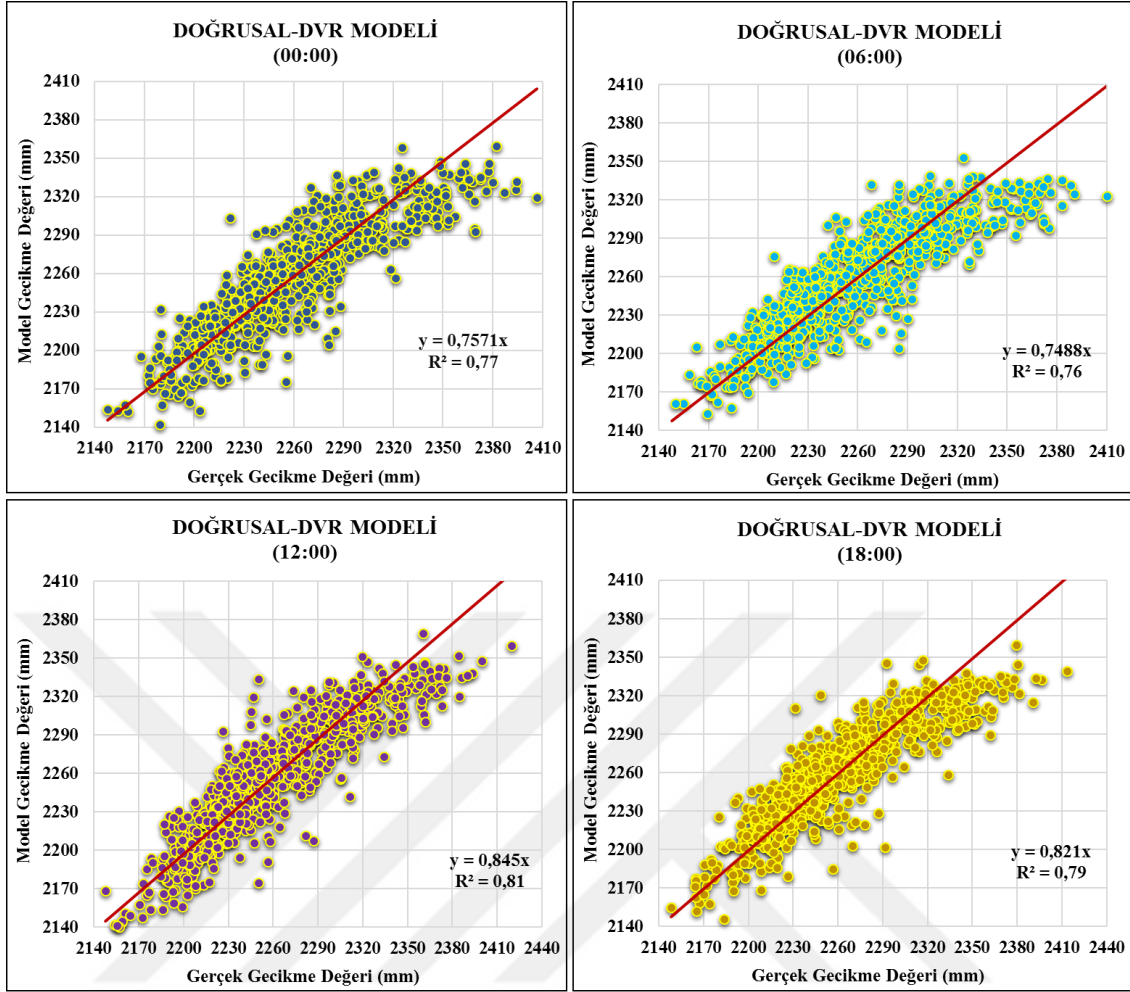
Grafiğe göre, belirlenmiş aralıkta C'nin 29,1 olduğu durumda model en yüksek  $R^2$  değerine (0,773) ulaşmıştır. Modelin en düşük  $R^2$  değerine (0,731) sahip olduğu durumda ise C parametresi 1'dir. Sonuç olarak, troposferik gecikme kestirimi için oluşturulacak olan Doğrusal-DVR modelinde kullanılan parametre, C için "29,1" olarak belirlenmiştir.

Diğer saatler (06:00, 12:00 ve 18:00) için modellemede aynı fonksiyon ve veri türü kullanıldığından belirlenmiş olan parametre bu saat dilimleri içinde geçerlidir. Çizelge 4.2'de doğrusal kernel fonksiyonu ile kurulan DVR modellerinin 4 farklı zaman için elde edilen model sonuçları özet halinde verilmiştir.

Çizelge 4.2. Doğrusal (Linear) DVR modeli performans sonuçları

| Model<br>C:29,1  |                | Linear DVR<br>00:00 için | Linear<br>DVR<br>06:00 için | Linear<br>DVR<br>12:00 için | Linear<br>DVR<br>18:00 için |
|------------------|----------------|--------------------------|-----------------------------|-----------------------------|-----------------------------|
| Eğitim<br>Hatası | R <sup>2</sup> | 0,773                    | 0,764                       | 0,816                       | 0,796                       |
|                  | OMH (mm)       | 17,5942                  | 18,1068                     | 16,4827                     | 16,5253                     |
|                  | OMYH           | 0,7746                   | 0,7979                      | 0,7258                      | 0,7279                      |
|                  | KOH (mm)       | ± 22,5356                | ± 22,9880                   | ± 21,0371                   | ± 21,5184                   |
| Test<br>Hatası   | R <sup>2</sup> | 0,751                    | 0,734                       | 0,761                       | 0,750                       |
|                  | OMH (mm)       | 18,4783                  | 19,9955                     | 18,8091                     | 18,1045                     |
|                  | OMYH           | 0,8145                   | 0,8776                      | 0,8277                      | 0,7954                      |
|                  | KOH (mm)       | ± 24,3798                | ± 25,4834                   | ± 24,0915                   | ± 23,7141                   |

R<sup>2</sup> değerinin hem eğitim hem de test verisinde en yüksek olduğu model, saat 12:00 için kurulan model olduğu görülmüştür. Aynı zamanda bu modeldeki hata oranları, diğer saatlerde kurulan modellere göre azalmıştır. Model-12:00'daki verilerin diğer saatlerde oluşturulan modellerdeki verilere göre daha az dağınık (yayılmış) halde bulunduğu yani, veri dağılımlarının, eğilim çizgisinin yakın çevresinde toplandığı saptanmıştır (Şekil 4.11). Hatalara ait mutlak ortalamaların, kurulan Model-00:00 için 17,59 mm ve sırasıyla diğer modeller için 18,11 mm, 16,48 mm ve 16,52 mm olduğu, karesel ortalama hataların ise, eğitim hatası için ±21-22 mm ve test hatası için ±23-25 mm olarak tespit edilmiştir. Model-18:00'ın Model-12:00 sonuçlarına göre R<sup>2</sup> değerinin 0,02 oranında düştüğü, OMYH oranının 0,002 ve KOH değerinin 0,481 arttığı gözlemlenmiştir. Model-00:00 ve Model-18:00 arasındaki R<sup>2</sup> değeri ve hata oranları farkının az olduğu belirlenmiştir ve değerler birbirine çok yakındır. Model-06:00 ise eğitim/test verisi için en düşük R<sup>2</sup> değerine sahiptir ve hata oranları diğer saatlerdeki modellere nazaran yüksektir. Şekil 4.11'de görüldüğü üzere, Model-06:00'daki verilerin diğer saatlerde oluşturulan modellerdeki verilere göre daha dağınık (yayılmış) halde bulunduğu saptanmıştır.



Şekil 4.11. Model ve gerçek gecikme değerleri arasındaki ilişki

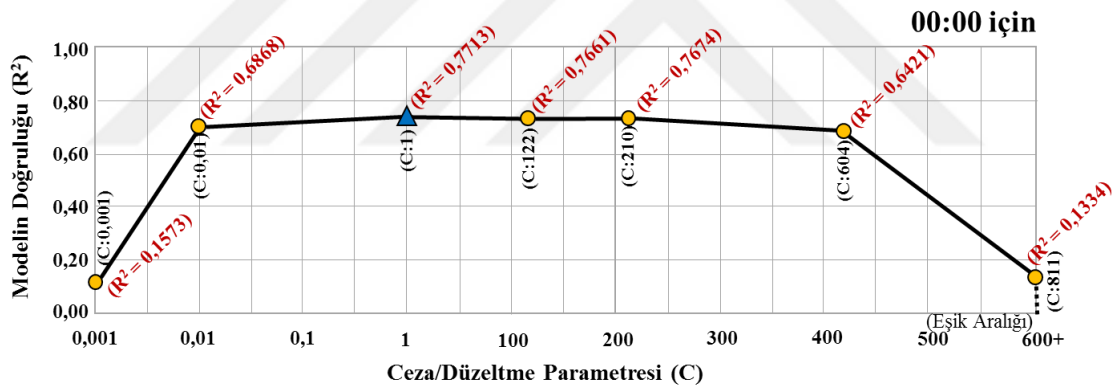
Bu kısımda 4 farklı zaman dilimi için oluşturulmuş olan modeller karşılaştırılıp yorumlanmıştır (Çizelge 4.2 ve Şekil 4.11). Ancak tez kapsamındaki amaç, en iyi troposferik saat modelini seçmek değil, belirlenmiş tüm ilgili zamanlar (epoklar) için en uygun DVR modeliyle troposferik gecikme kestirimi yapmaktır. Bu nedenle Doğrusal-DVR modelinin diğer kernel fonksiyonlarındaki modellerle kıyaslanabilmesi adına 4 farklı zamana ait OMH, OMYH ve KOH hata türleri incelenerek, hata yayılma kuralına göre genel model değerlendirmesi yapılmıştır (Çizelge 4.3).

Çizelge 4.3. Doğrusal-DVR modeline ait genel hata değerleri

| MODEL        |           | OMH DEĞERİ<br>(mm) | OMYH ORANI<br>(%) | KOH DEĞERİ<br>(mm) |
|--------------|-----------|--------------------|-------------------|--------------------|
| Doğrusal-DVR | Eğitim H. | 17,1772            | 0,7566            | ± 22,0198          |
|              | Test H.   | 18,8468            | 0,8288            | ± 24,4172          |

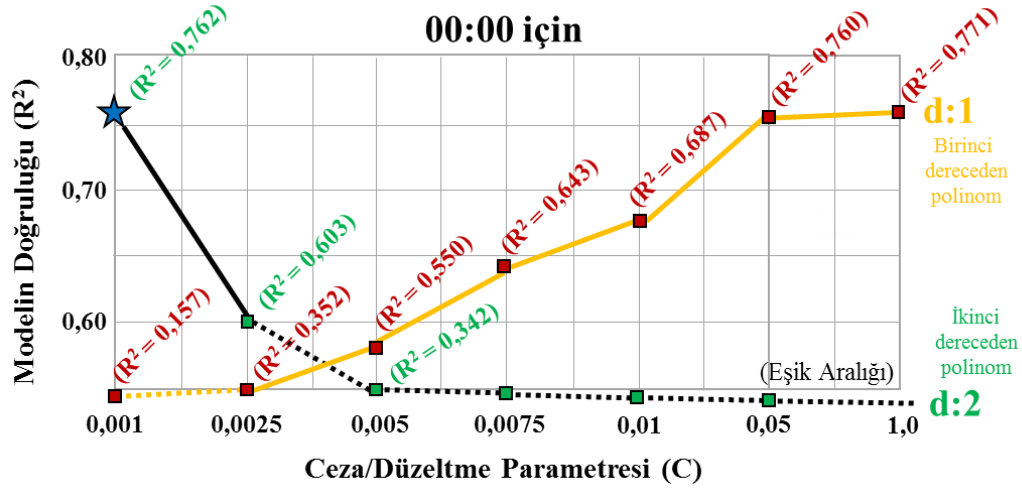
#### 4.4.2. Polinomal (Polynomial) kernel fonksiyonu ile modelleme

Polinomal (polynomial) kernel kullanımı için kullanıcı tarafından belirlenmesi gereken parametreler, ceza/düzeltilme (C) ve polinom derecesi (d) parametreleridir. Polinom kernelinde iki parametre olan C ve d'nin model tahmin sonucuna etkisi kesinlikle söz konusudur. Ancak kullanılan hangi parametre çiftinin modelde en iyi sonucu verdiği bilinmemektedir. Dolayısıyla ilgili fonksiyonda en uygun parametrelerin belirlenmesi gerekmektedir. Optimizasyon yöntemleri ile bu parametreler, oluşturulacak olan gecikme tahmin modeli için eğitim ve test verileri kullanılarak tespit edilmiştir. Çözüm olarak, polinomal fonksiyon için en uygun C ve d parametrelerinin seçimi, grid tarama yöntemine göre araştırılmıştır. İlk aşamada C parametresi, belirlenmiş eşik aralıklarında değeri 1 artırılarak denenmiş ve en uygun eşik değeri “C=1” olarak tespit edilmiştir. Aşağıda farklı C parametrelerinin model doğruluğuna olan etkileri,  $R^2$  değeri üzerinden gösterilmiştir (Şekil 4.12).



Şekil 4.12. Farklı (C) ceza/düzeltilme parametrelerine ait model doğrulukları

Grafiğe göre, C'nin 1 olduğu durumda model en yüksek  $R^2$  değerine (0,7713) ulaşmıştır. Modelin en düşük  $R^2$  değerine (0,1334) sahip olduğu durumda ise C parametresi 811 olarak belirlenmiştir. İkinci aşamada ise ilk aşamada belirlenen parametrelerin yakınında daha küçük bir aralıkta ikinci bir arama yapılarak daha iyi parametre değerlerine ulaşılmaya çalışılmıştır. Ancak bu kısımda sadece C parametresi değil d polinom derecesi de hesaba katılmıştır. C, eşik değeri  $10^{-3}$  ve  $10^0$  arasında 0,1 artışla denenmiştir. Ayrıca polinom derecesi sırayla arttırılmıştır (Şekil 4.13).



Şekil 4.13. Farklı C ve d ikili parametre değerlerine ilişkin model doğrulukları

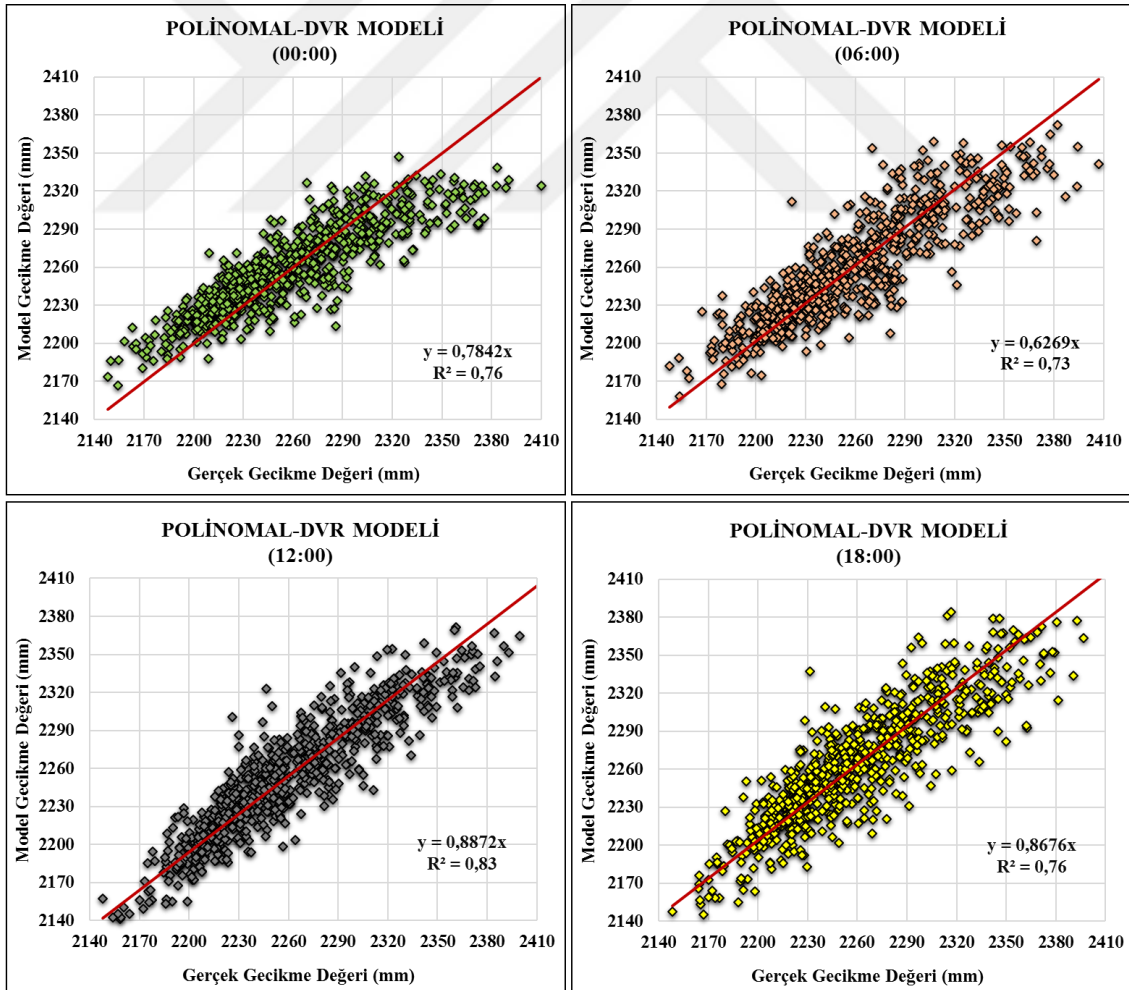
Şekil 4.13'deki sonuçlara göre, C'nin  $10^0$  (1,0) ve d polinom derecesinin 1 olduğu durumda model en yüksek  $R^2$  değerine (0,771) ulaşmıştır. Ancak burada farklı bir durumla karşılaşmıştır. Modeli kurmak için gerekli olan parametrelerin en uygunu C=1 ve d=1 olsa da modeli test ederken başarılar oldukça düşük elde edilmiştir. Bu nedenle eğitim sırasında aşırı öğrenme olayı meydana gelmiştir. Bunun için 2. dereceden elde edilen sonuçlarda dikkate alınarak, en uygun modelin “C= $10^{-3}$ ” ve “d=2” parametre değerleriyle kurulması sonucuna varılmıştır.

Diğer saatler (06:00, 12:00 ve 18:00) için modellemede aynı fonksiyon ve veri türü kullanıldığından tespit edilen parametreler bu saat dilimleri içinde geçerlidir. Çizelge 4.4'de polinomal kernel fonksiyonu ile kurulan DVR modellerinin 4 farklı zaman için elde edilen model sonuçları özet halinde verilmiştir.

Çizelge 4.4. Polinomal (Polynomial) DVR modeli performans sonuçları

| Model<br>C= $10^{-3}$ , d=2 |                | Polynomial<br>DVR<br>00:00 için | Polynomial<br>DVR<br>06:00 için | Polynomial<br>DVR<br>12:00 için | Polynomial<br>DVR<br>18:00 için |
|-----------------------------|----------------|---------------------------------|---------------------------------|---------------------------------|---------------------------------|
| Eğitim<br>Hatası            | R <sup>2</sup> | 0,762                           | 0,737                           | 0,830                           | 0,759                           |
|                             | OMH (mm)       | 17,7550                         | 19,1511                         | 15,7972                         | 18,2563                         |
|                             | OMYH           | 0,7823                          | 0,8457                          | 0,6994                          | 0,8043                          |
|                             | KOH (mm)       | ± 23,0850                       | ± 24,2940                       | ± 20,2156                       | ± 23,3853                       |
| Test<br>Hatası              | R <sup>2</sup> | 0,775                           | 0,743                           | 0,793                           | 0,728                           |
|                             | OMH (mm)       | 17,7341                         | 20,1986                         | 17,5052                         | 19,2482                         |
|                             | OMYH           | 0,7828                          | 0,8939                          | 0,7754                          | 0,8460                          |
|                             | KOH (mm)       | ± 23,1808                       | ± 25,0423                       | ± 22,4615                       | ± 24,7388                       |

$R^2$  değerinin hem eğitim hem de test verisinde en yüksek olduğu model, saat 12:00 için kurulan model olmuştur. Aynı zamanda bu modeldeki hata değerleri, diğer saatlerdeki modellere göre azalmıştır. Model-12:00'daki verilerin diğer saatlerde oluşturulan modellerdeki verilere göre daha az dağınık halde bulunduğu yani, veri dağılımlarının, eğilim çizgisinin yakın çevresinde toplandığı saptanmıştır (Şekil 4.14). Hatalara ait mutlak ortalamaların, kurulan Model-00:00 için 17,76 mm ve sırasıyla diğer modeller için 19,15 mm, 15,80 mm ve 18,26 mm olduğu, karesel ortalama hatanın ise, eğitim hatası için  $\pm 22$ -24 mm ve test hatası için  $\pm 24$ -26 mm olarak tespit edilmiştir. Model-18:00'in Model-12:00 sonuçlarına göre  $R^2$  değerinin 0,07 düştüğü, OMYH oranının 0,010 ve KOH değerinin 3,170 arttığı gözlemlenmiştir. Model-00:00 ve Model-18:00 arasındaki  $R^2$  değeri ve hata oranları birbirine çok yakındır. Şekil 4.14'de görüldüğü üzere, Model-06:00'daki verilerin diğer saatlerdekilere göre daha dağınık halde bulunduğu saptanmıştır.



Şekil 4.14. Model ve gerçek gecikme değerleri arasındaki ilişki

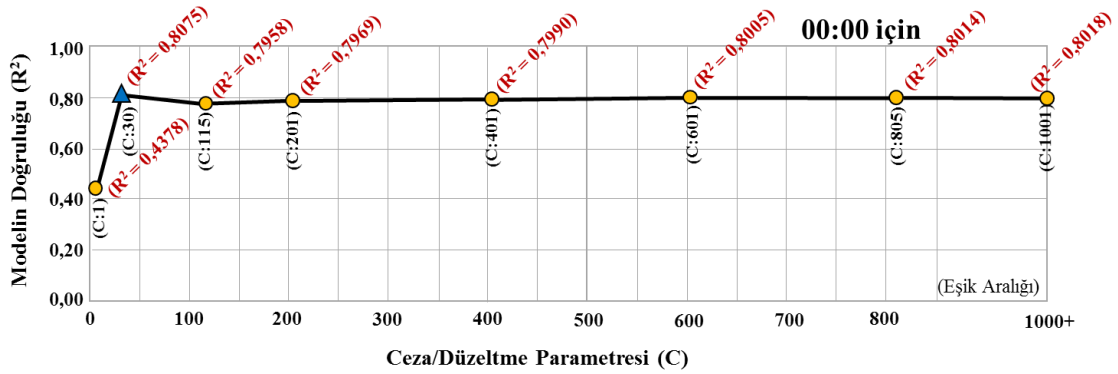
Bu kısımda 4 farklı zaman dilimi için oluşturulmuş olan modeller karşılaştırılıp yorumlanmıştır (Çizelge 4.4 ve Şekil 4.14). Daha önce de ifade edildiği şekilde, Polinomal-DVR modelinin diğer fonksiyonlardaki modellerle kıyaslanabilmesi için 4 farklı zamana ait OMH, OMYH ve KOH hata türleri irdelenerek, genel model değerlendirmesi yapılmıştır (Çizelge 4.5).

**Çizelge 4.5.** Polinomal-DVR modeline ait genel hata değerleri

| MODEL         |           | OMH DEĞERİ<br>(mm) | OMYH ORANI<br>(%) | KOH DEĞERİ<br>(mm) |
|---------------|-----------|--------------------|-------------------|--------------------|
| Polinomal-DVR | Eğitim H. | 17,7399            | 0,7829            | ± 22,7445          |
|               | Test H.   | 18,6715            | 0,8245            | ± 23,8558          |

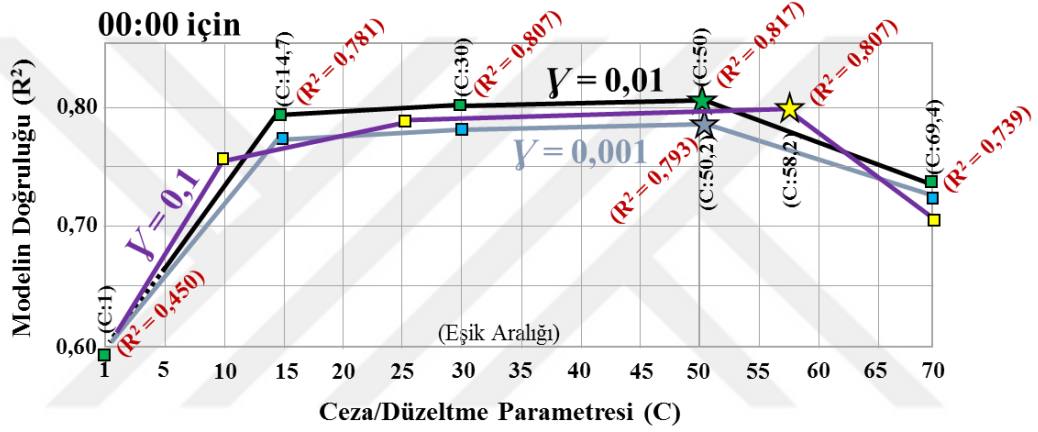
#### 4.4.3. Radyal tabanlı (RTF) kernel fonksiyonu ile modelleme

Radyal Tabanlı (RTF) kernel kullanımı için kullanıcı tarafından belirlenmesi gereken parametreler, ceza/düzeltilme (C) ve gamma ( $\gamma$ ) parametreleridir. RTF kernelinde iki parametre olan C ve  $\gamma$ 'nın model tahmin sonucuna etkisi kesinlikle söz konusudur. Ancak kullanılan hangi parametre çiftinin modelde en iyi sonucu verdiği bilinmemektedir. Optimizasyon yöntemleri ile bu parametreler, oluşturulacak olan gecikme tahmin modeli için eğitim ve test verileri kullanılarak tespit edilmiştir. Çözüm olarak, radyal tabanlı fonksiyon için en uygun C ve  $\gamma$  parametrelerinin seçimi, grid tarama yöntemine göre araştırılmıştır. İlk aşamada C parametresi, belirlenmiş eşik aralıklarında değeri 1 arttırılarak denenmiş ve en uygun eşik değeri “C=30” olarak tespit edilmiştir. Aşağıda farklı C parametrelerinin model doğruluğuna olan etkileri,  $R^2$  değeri üzerinden gösterilmiştir (Şekil 4.15).



**Şekil 4.15.** Farklı (C) ceza/düzeltilme parametrelerine ait model doğrulukları

Şekil 4.15'e göre, C'nin 30 olduğu durumda model en yüksek  $R^2$  değerine (0,8075) ulaşmıştır. Modelin en düşük  $R^2$  değerine (0,4378) sahip olduğu durumda ise C parametresi 1 olarak belirlenmiştir. İkinci aşamada ise ilk aşamada belirlenen parametrelerin yakınında daha küçük bir aralıkta ikinci bir arama yapılarak daha iyi parametre değerlerine ulaşılmaya çalışılmıştır. Ancak bu kısımda sadece C parametresi değil  $\gamma$  gamma parametresi de hesaba katılmıştır. C, eşik değeri 1 ve 70 arasında 0,1 artışla denenmiştir. Belirlenen (1-70) aralığın kullanılma sebebi, doğrusal modelde açıklanan sebepler ile aynıdır. Diğer taraftan  $\gamma$  parametresinin 0,001, 0,01 ve 0,1 aralıklarındaki farklılıkları izlenmiştir (Şekil 4.16).



Şekil 4.16. Farklı C ve  $\gamma$  ikili parametre değerlerine ilişkin model doğrulukları

C ve  $\gamma$  ikili parametresi tüm değerler için irdelenmiş olup değerlendirmeye alınmıştır. Değerlendirme sonucu tüm farklı  $\gamma$  parametresinde, C değerindeki artışa karşılık elde edilen model doğruluklarının belirli bir parametreye kadar arttığı, sonrasında düşmeye başladığı görülmüştür. Sonuç olarak, Şekil 4.16'dan da görüleceği üzere, C'nin 50 ve  $\gamma$  parametresinin 0,01 olduğu durumda model en yüksek  $R^2$  değerine (0,817) ulaşmıştır. Yani, en uygun RTF modelin "C=50" ve " $\gamma=0,01$ " parametre değerleriyle kurulması sonucuna varılmıştır.

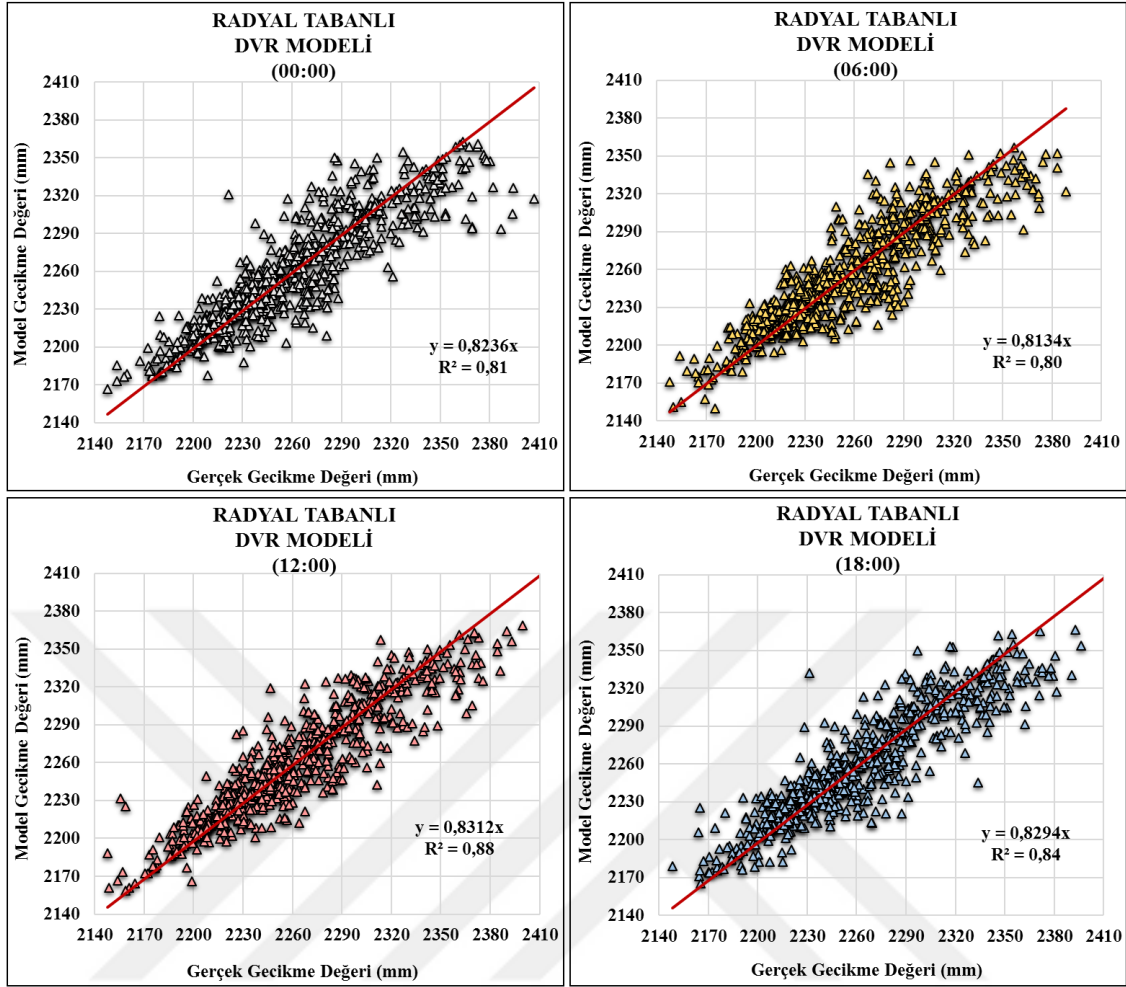
Diğer saatler (06:00, 12:00 ve 18:00) için modellemede aynı fonksiyon ve veri türü kullanıldığından tespit edilen parametreler bu saat dilimleri içinde geçerlidir. Çizelge 4.6'da radyal tabanlı kernel fonksiyonu ile kurulan DVR modellerinin 4 farklı zaman için elde edilen model sonuçları özet halinde verilmiştir.

**Çizelge 4.6.** Radyal tabanlı (RTF) DVR modeli performans sonuçları

| <b>Model</b><br>C=50, $\gamma=0,01$ |                      | <b>RTF-DVR</b><br><b>00:00 için</b> | <b>RTF-DVR</b><br><b>06:00 için</b> | <b>RTF-DVR</b><br><b>12:00 için</b> | <b>RTF-DVR</b><br><b>18:00 için</b> |
|-------------------------------------|----------------------|-------------------------------------|-------------------------------------|-------------------------------------|-------------------------------------|
| <b>Eğitim</b><br><b>Hatası</b>      | <b>R<sup>2</sup></b> | 0,817                               | 0,796                               | 0,884                               | 0,847                               |
|                                     | <b>OMH (mm)</b>      | 13,9862                             | 14,7389                             | 11,3743                             | 12,3084                             |
|                                     | <b>OMYH</b>          | 0,6161                              | 0,6503                              | 0,5008                              | 0,5421                              |
|                                     | <b>KOH (mm)</b>      | ± 20,2677                           | ± 21,3860                           | ± 16,7358                           | ± 18,6638                           |
| <b>Test</b><br><b>Hatası</b>        | <b>R<sup>2</sup></b> | 0,778                               | 0,752                               | 0,789                               | 0,770                               |
|                                     | <b>OMH (mm)</b>      | 16,6662                             | 18,6398                             | 16,2627                             | 17,3759                             |
|                                     | <b>OMYH</b>          | 0,7357                              | 0,8263                              | 0,7168                              | 0,7679                              |
|                                     | <b>KOH (mm)</b>      | ± 22,0603                           | ± 23,6196                           | ± 21,6873                           | ± 22,7697                           |

R<sup>2</sup> değerinin hem eğitim hem de test verisinde en yüksek olduğu model, saat 12:00 için kurulan model olmuştur. Aynı zamanda bu modeldeki hata değerleri, diğer saatlerdeki modellere göre azalmıştır. Model-12:00'daki verilerin diğer saatlerde oluşturulan modellerdeki verilere göre daha az dağınık halde bulunduğu yani, veri dağılımlarının, eğilim çizgisinin yakın çevresinde toplandığı saptanmıştır (Şekil 4.17). Hatalara ait mutlak ortalamaların, kurulan Model-00:00 için 13,99 mm ve sırasıyla diğer modeller için 14,74 mm, 11,37 mm ve 12,31 mm olduğu tespit edilmiştir. Karesel ortalama hataların ise, eğitim hatası için ±16-21 mm ve test hatası için ±21-24 mm olarak tespit edilmiştir. Model-18:00'in Model-12:00 sonuçlarına göre R<sup>2</sup> değerinin 0,037 düştüğü, OMYH oranının 0,041 ve KOH değerinin 1,938 arttığı gözlemlenmiştir.

Şekil 4.17'de görüldüğü üzere, Model-06:00'daki verilerin diğer saat dilimindekilere göre daha dağınık halde bulunduğu saptanmıştır. Model-12:00 ise verilerde dağınıklığın en az gözlemlendiği yani gerçek değer ile model değerinin birbirine en yakın olduğu saat modelidir.



Şekil 4.17. Model ve gerçek gecikme değerleri arasındaki ilişki

Bu kısımda 4 farklı zaman dilimi için oluşturulmuş olan modeller karşılaştırılıp yorumlanmıştır (Çizelge 4.6 ve Şekil 4.17). Daha önce de ifade edildiği şekilde, Radyal Tabanlı-DVR modelinin diğer fonksiyonlardaki modellerle kıyaslanabilmesi için 4 farklı zamana ait OMH, OMYH ve KOH hata türleri irdelenerek, hata yayılım kuralına göre genel model değerlendirmesi yapılmıştır (Çizelge 4.7).

Çizelge 4.7. Radyal Tabanlı-DVR modeline ait genel hata değerleri

| MODEL   |           | OMH DEĞERİ<br>(mm) | OMYH ORANI<br>(%) | KOH DEĞERİ<br>(mm) |
|---------|-----------|--------------------|-------------------|--------------------|
| RTF-DVR | Eğitim H. | 13,1020            | 0,5773            | 19,2633            |
|         | Test H.   | 17,2362            | 0,7617            | 22,5342            |

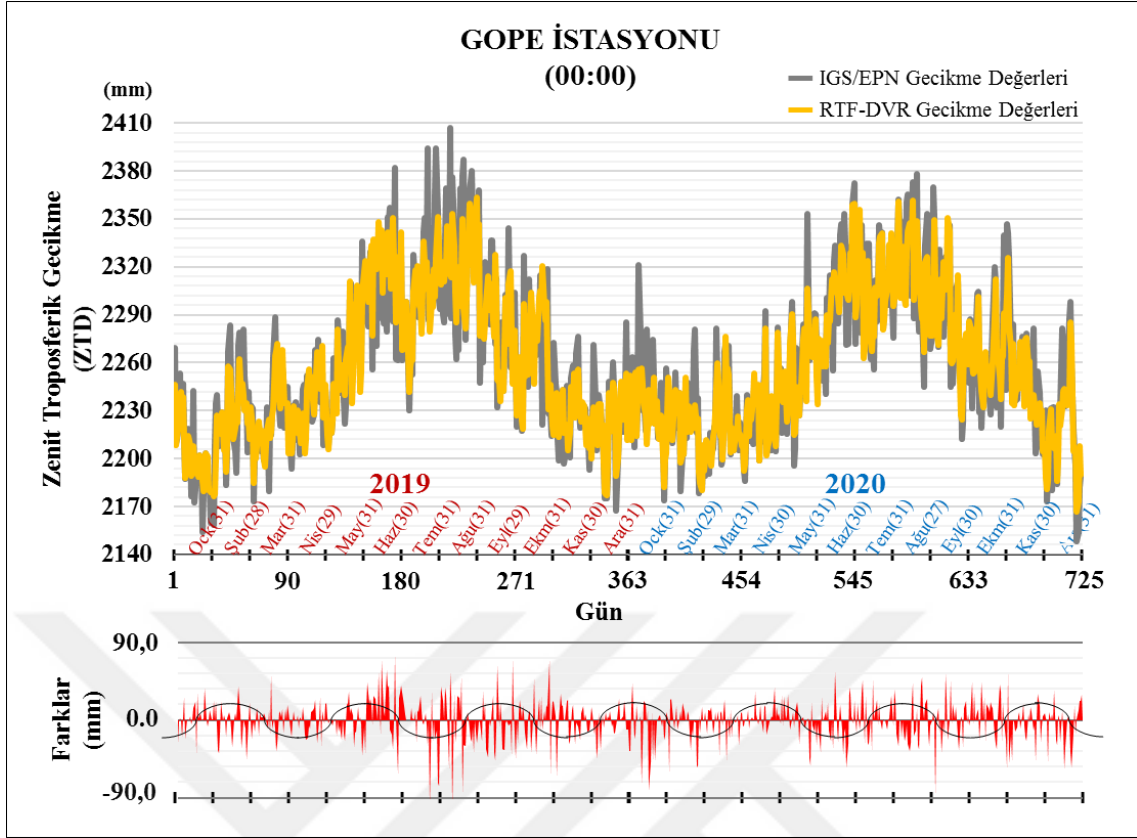
Çalışmadaki genel kapsam çerçevesinde, en uygun troposfer modeli seçimi için farklı fonksiyondaki destek vektör regresyonları ile modeller kurulmuştur.

#### 4.4.4. Model kestirimlerine ait analiz sonuçları

DVR modellerinin kullanımı sonucunda elde edilen analiz sonuçlarına göre model gecikme değerleri hesaplanmıştır. Farklı fonksiyonlardaki tüm modeller için gerçek gecikme değerleri ile model gecikme değerlerinin veri dağılımı incelenerek eğilim çizgisinin oluşturduğu doğru ile veriler arasındaki ilişkiye bakılmıştır (Şekil 4.11, Şekil 4.14 ve Şekil 4.17). Ayrıca modellerin oluşturulma ve test edilme aşamasında elde edilen hata değerleri karşılaştırılarak model doğruluğu üzerindeki etkileri incelenmiştir. İnceleme sonuçlarından yola çıkarak en iyi ve  $R^2$ 'nin 1'e en yakın sonuçları, radyal tabanlı kernel fonksiyonu ile oluşturulan DVR modelinden elde edildiği ortaya çıkmıştır (Çizelge 4.3, Çizelge 4.5 ve Çizelge 4.7). Dolayısıyla çalışmada troposferik gecikme modeli kestirimi için en iyi ve en uygun model "RTF-DVR" modeli olarak seçilmiştir.

Seçilen "RTF-DVR" modelinde, belirli zaman dilimleri için kestirilen gecikme değerleri ile diğer modellerden (IGS/EPN ve CSRS-PPP) elde edilen gecikme değerleri karşılaştırılmıştır. RTF-DVR modelinin 00:00'daki kestirim sonuçlarına ilişkin verilere (Şekil 4.18) bakıldığında, model değerlerinin özellikle 2019'daki haziran, temmuz ve ağustos aylarında, IGS/EPN istasyonlarındaki gecikme değerlerinden farklılıklar gösterdiği saptanmıştır. Bu farklılıkların nedeni, Bölüm 4.3'te de ifade edildiği gibi sıcaklık ve nemin yükselmesiyle, su buharı basıncı ve gecikme etkisinin büyümesidir. Buna bağlı olarak gecikme değerinin maksimum seviyelere ulaşma eğilimine girmesi sonucu makine öğrenme modellerinin istemediği durum gerçekleşmiştir. Çalışmadaki yaz aylarında elde edilen gecikme değerlerinin tüm yıl içerisinde elde edilen değerlere nazaran normal dağılım sergilemediği tespit edilmiştir. Dolayısıyla kullanılan makine öğrenme modeli, arazi/hava koşullarına göre belirlenen gerçek troposferik gecikme değerlerinde olduğu gibi yaz ayları için daha çok sapmalara sebebiyet vermiştir.

2020 yılı için, ocak ayındaki günlere ait değerlerde de biraz farklar bulunmakta olup diğer zamanlarda çok fazla fark gözlenmemiştir. Sıcaklık, basınç ve bağıl nem parametrelerine dayalı oluşturulan destek vektör regresyonu modeli ile IGS/EPN modelindeki değerler arasında en büyük farklar, 25 Haziran 2019 günü için 64,67 mm ve 10 Ağustos 2019 günü için -89,39 mm olarak belirlenmiştir. En küçük farklar ise, 96 farklı gün için 0,01 mm şeklinde bulunmuştur. Genel olarak, Model-00:00 için tüm değerlere ait ortalama farklar 14,78 mm olarak hesaplanmıştır (Şekil 4.18).



Şekil 4.18. RTF-DVR (00:00) modelinin IGS/EPN gecikme değerleriyle uyumu

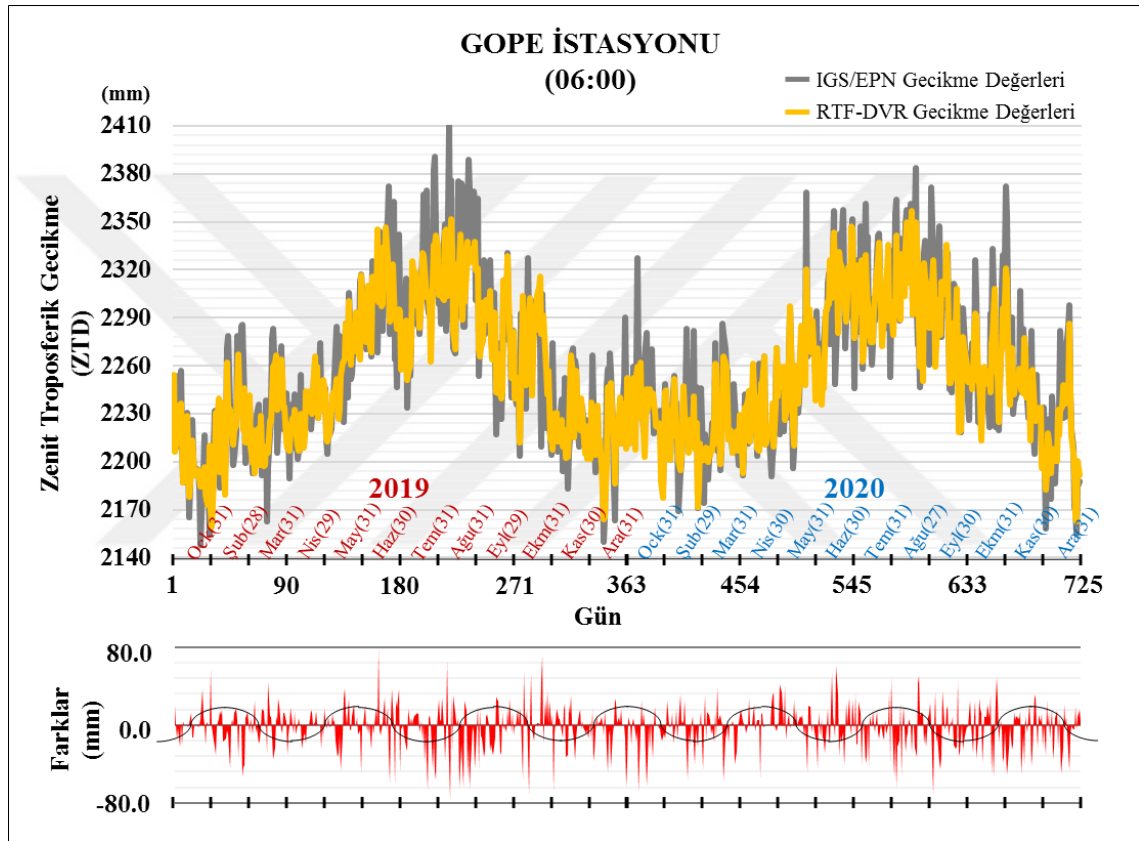
RTF-DVR modeli ile IGS/EPN'den elde edilen sonuçların birbirine ne kadar yakın olduğunu araştırmak için modellerin kendi içerisinde ortalama mutlak sapmaları hesaplanmıştır. Ortalama mutlak sapma değerleri (4.1) eşitliğiyle elde edilmiştir.

$$\frac{1}{n} \sum_{i=1}^n |x_{gün} - \bar{x}| = \frac{|x_{1.gün} - \bar{x}| + |x_{2.gün} - \bar{x}| + \dots + |x_{n.gün} - \bar{x}|}{n} \quad (4.1)$$

Modellerdeki tüm günler için hesaplanan ortalama mutlak sapmaların, IGS/EPN için 47,87 mm ve RTF-DVR yöntemi için 43,92 mm olduğu görülmüştür. Burada ortalama mutlak sapmaların düşük ya da yüksek olması önemli değil birbirine yakın değerler olması çalışmada kritik rol oynamaktadır.

Şekil 4.19 göz önüne alınarak, RTF-DVR modelinin 06:00'daki kestirim sonuçları incelendiğinde, model değerlerinin yine 2019'daki haziran, temmuz ve ağustos aylarında büyük farklılıklar gösterdiği saptanmıştır. Gözlem sonuçlarına göre 2019 ve 2020 ağustos aylarına ait günlerde, RTF-DVR modeli ile IGS-EPN değerleri arasındaki farklar (ort: 26,84 mm) yüksek çıkmıştır. Ancak Model-00:00'dan farklı

olarak 2020 yılı eylül ve ekim aylarında da modelde yüksek boyutta farklılıklar meydana gelmiştir. İlkbahar mevsimlerini kapsayan günlerde ise değerlerde anlamlı değişim bulunmayıp farklar çok az değişim göstermiştir. Bu sonuçlar neticesinde; 06:00 saati için oluşturulan destek vektör regresyonu modeli ile IGS/EPN modelindeki değerler arasında en büyük farklar, 14 Haziran 2019 günü için 77,03 mm ve 10 Ağustos 2019 günü için -76,55 mm olarak bulunmuştur. Model-06:00 için tüm değerlere ait ortalama farklar 15,91 mm olarak hesaplanmıştır (Şekil 4.19).

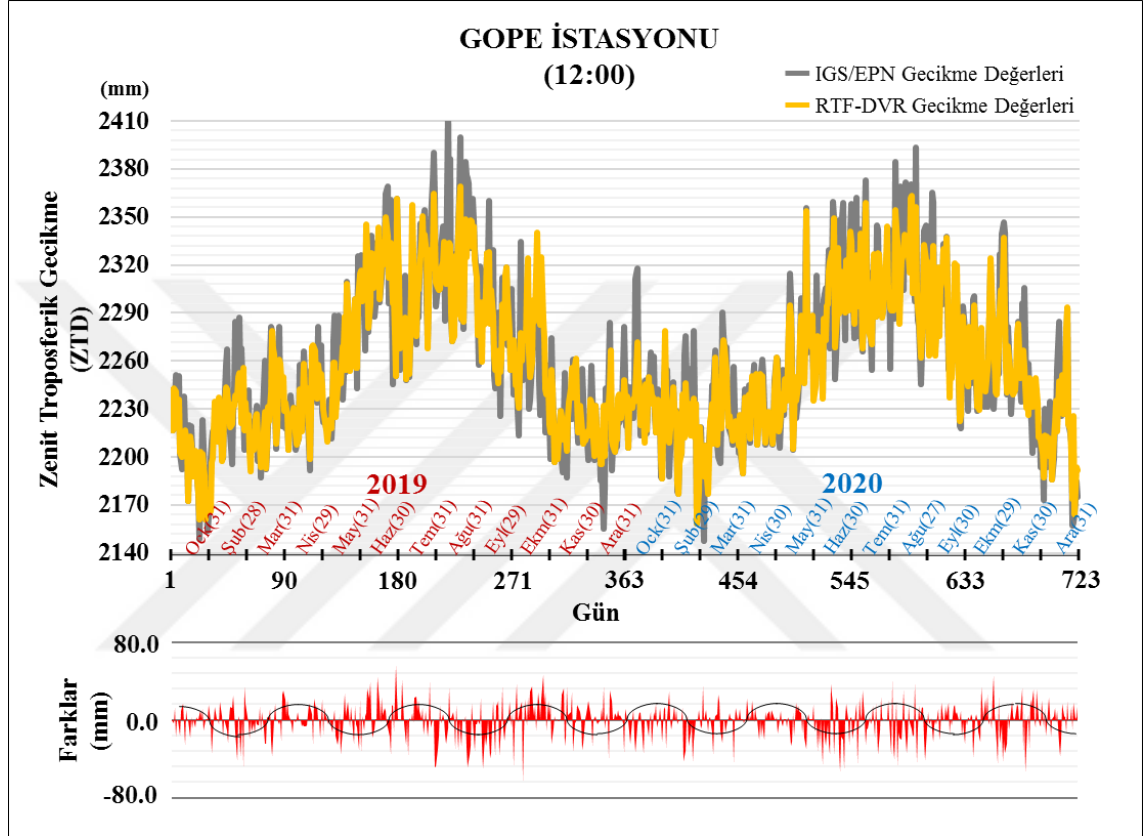


Şekil 4.19. RTF-DVR (06:00) modelinin IGS/EPN gecikme değerleriyle uyumu

Model-00:00'da olduğu gibi RTF-DVR modeli ile IGS/EPN sonuçlarının benzerliğini incelemek için ortalama mutlak sapmalar hesaplanmıştır. Modellerdeki tüm günler için hesaplanmış olan ortalama mutlak sapmaların, IGS/EPN için 48,05 mm ve RTF-DVR yöntemi için 42,59 mm olduğu görülmüştür.

Şekil 4.20 RTF-DVR modelinin 12:00'daki kestirim sonuçlarına göre, model değerlerinin IGS/EPN değerleri ile diğer saatlere kıyasla daha az farklılıklar gösterdiği açıkça belli olmaktadır. Analiz sonuçları incelendiğinde; 12:00 saati için oluşturulan destek vektör regresyonu modeli ile IGS/EPN modelindeki değerler arasında en büyük farklar, 28 Aralık 2020 günü için 66,90 mm ve 8 Ocak günü için -68,55 mm olarak

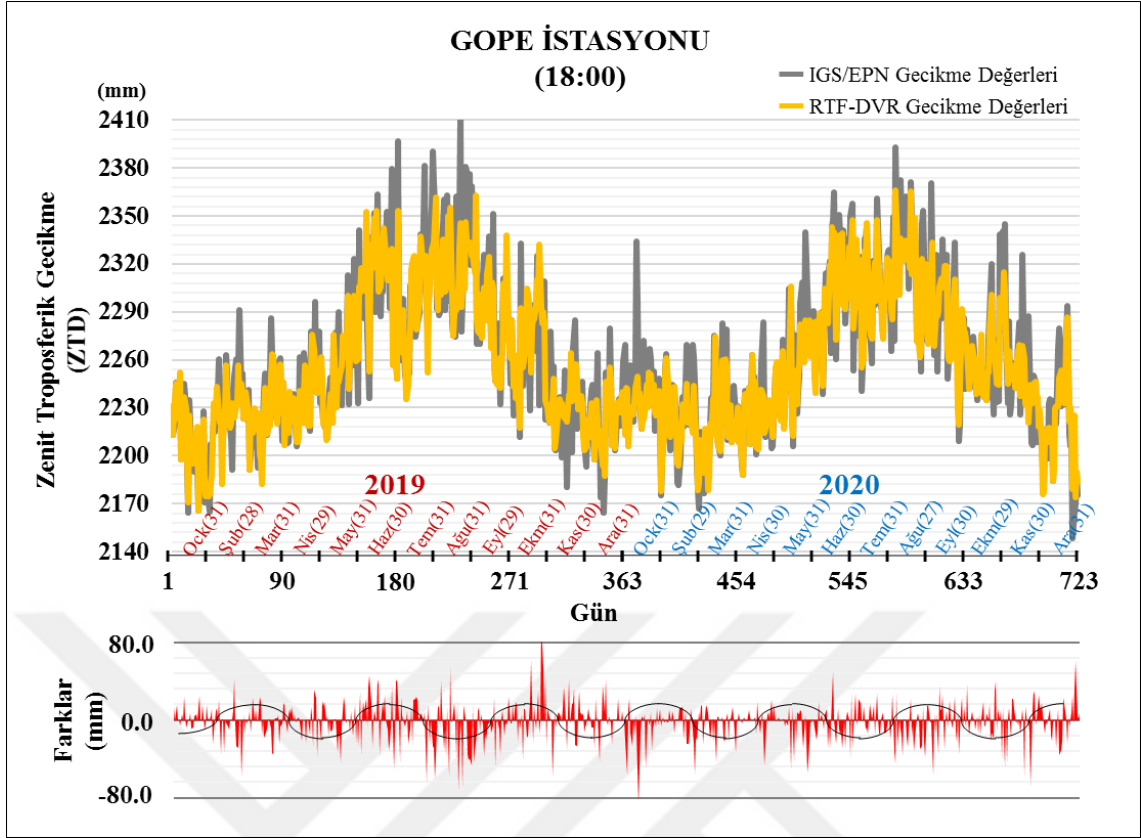
bulunmuştur. Model-12:00 için tüm değerlere ait ortalama farklar 13,10 mm olarak hesaplanmıştır (Şekil 4.20). RTF-DVR modeli ile IGS/EPN sonuçlarının benzerliğini incelemek için diğer saatlerde olduğu gibi ortalama mutlak sapmalar tespit edilmiştir. Modellerdeki tüm günler için hesaplanan ortalama mutlak sapmaların, IGS/EPN için 49,20 mm ve RTF-DVR yöntemi için 44,50 mm olduğu görülmüştür.



Şekil 4.20. RTF-DVR (12:00) modelinin IGS/EPN gecikme değerleriyle uyumu

Son olarak Şekil 4.21’de görüldüğü üzere, RTF-DVR modelinin 18:00’daki kestirim sonuçları incelendiğinde, model ile IGS/EPN değeri arasında en büyük farklar, 23 Eylül 2019 günü (100,43 mm) ve 8 Ocak 2020 günü (-88,71 mm) meydana gelmiştir. Model-18:00 için tüm değerlere ait ortalama farklar 13,57 mm olarak hesaplanmıştır. Ayrıca modellerdeki tüm günler için hesaplanan ortalama mutlak sapmaların, IGS/EPN için 47,65 mm ve RTF-DVR yöntemi için 42,41 mm olduğu görülmüştür (Şekil 4.21).

Burada farklı olarak şu göze çarpmaktadır. Model-12:00 ile Model-18:00 için değerdeki sapmanın en fazla yaşandığı gün 8 Ocak olarak tespit edilmiştir. Bu durum, model için uyumsuz verinin olabilme ihtimali anlamını taşımaktadır. Çalışmanın en son aşamasında (Bölüm 4.4.5) uyumsuz olarak düşünülen veriler ayıklanarak tekrardan modeller kurulup sonuç üzerindeki etkisi incelenmiştir.



Şekil 4.21. RTF-DVR (18:00) modelinin IGS/EPN gecikme değerleriyle uyumu

Yukarıda iki yıl boyunca IGS/EPN istasyonundan elde edilen gecikme değerleri ile kurulan modellerle kestirilen model gecikme değerlerinin karşılaştırma grafikleri verilmiştir. Oluşturulan modellerde sonuç üzerinde başarıyı düşüren günler farklı olsa bile modellerin iyi sonuçlar ürettiği görülmektedir. Bununla birlikte modellerde uyuşumsuz veriler ve verinin ham halleri kullanılmasına rağmen gecikme kestirimi için yeterli uyum sağladığı ve gerçeğe yakın sonuçlar elde ettiği görülmüştür.

RTF-DVR modelinden elde edilen gecikme değerleri, bir diğer değerlendirme modeli CSRS-PPP'den elde edilen analiz sonuçları ile karşılaştırılmıştır. Bu karşılaştırma sonuçlarından elde edilen farklar ile IGS/EPN ve RTF-DVR'nin karşılaştırılması sonucu elde edilen farkların birbirine çok benzer olduğu gözlemlenmiştir. Bunun nedeni çevrimiçi hizmet veren CSRS-PPP'nin ZTD değerini, IGS troposferik ürününe yakın bir değerde tahmin etmesi, çoğu durumda da iki değer arasındaki hata farklarının 1 cm'ye yakın veya daha az olmasıdır. Bu çalışmada, IGS/EPN istasyonlarından elde edilen ZTD değerleri ile CSRS-PPP'den elde edilen değerler arasındaki ortalama hata değeri 1,30 cm olarak hesaplanmıştır. RTF-DVR modelleri ile CSRS-PPP'nin karşılaştırılması sonucu elde edilen istatistiki bilgiler Çizelge 4.8'de verilmiştir.

Çizelge 4.8. RTF-DVR modelleri ve CSRS-PPP'den elde edilen gecikme değerleri arasındaki farklar (mm)

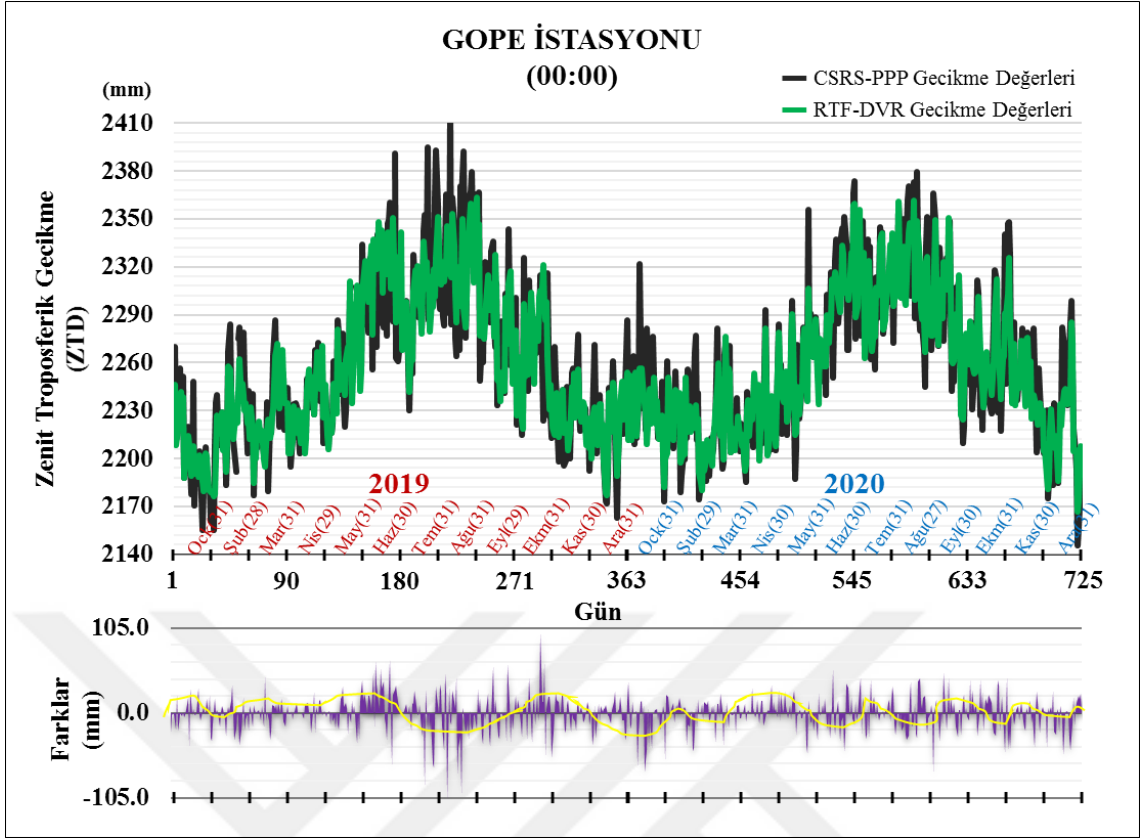
|       |         | 00:00                                                   |                             |                      | 06:00                                                   |                             |                      | 12:00                                                   |                             |                      | 18:00                                                   |                             |                      |
|-------|---------|---------------------------------------------------------|-----------------------------|----------------------|---------------------------------------------------------|-----------------------------|----------------------|---------------------------------------------------------|-----------------------------|----------------------|---------------------------------------------------------|-----------------------------|----------------------|
|       |         | RTF-DVR İLE<br>CSRS-PPP'DEKİ GECİKME<br>DEĞERLERİ ARASI |                             |                      | RTF-DVR İLE<br>CSRS-PPP'DEKİ GECİKME<br>DEĞERLERİ ARASI |                             |                      | RTF-DVR İLE<br>CSRS-PPP'DEKİ GECİKME<br>DEĞERLERİ ARASI |                             |                      | RTF-DVR İLE<br>CSRS-PPP'DEKİ GECİKME<br>DEĞERLERİ ARASI |                             |                      |
| AYLAR |         | EN<br>KÜÇÜK<br>FARK<br>(mm)                             | EN<br>BÜYÜK<br>FARK<br>(mm) | ORT.<br>FARK<br>(mm) | EN<br>KÜÇÜK<br>FARK<br>(mm)                             | EN<br>BÜYÜK<br>FARK<br>(mm) | ORT.<br>FARK<br>(mm) | EN<br>KÜÇÜK<br>FARK<br>(mm)                             | EN<br>BÜYÜK<br>FARK<br>(mm) | ORT.<br>FARK<br>(mm) | EN<br>KÜÇÜK<br>FARK<br>(mm)                             | EN<br>BÜYÜK<br>FARK<br>(mm) | ORT.<br>FARK<br>(mm) |
| 2019  | Ocak    | -0,10                                                   | -40,79                      | <b>0,31</b>          | 0,20                                                    | 54,72                       | <b>0,67</b>          | -0,10                                                   | -40,49                      | <b>-2,04</b>         | -0,17                                                   | 34,00                       | <b>2,85</b>          |
|       | Şubat   | 0,03                                                    | -39,46                      | <b>-3,82</b>         | 0,50                                                    | -59,50                      | <b>-10,62</b>        | 0,30                                                    | -47,81                      | <b>-4,10</b>         | 0,01                                                    | -55,09                      | <b>-3,03</b>         |
|       | Mart    | 0,01                                                    | 44,40                       | <b>-0,51</b>         | 2,01                                                    | 46,77                       | <b>1,88</b>          | 0,46                                                    | 23,26                       | <b>-2,69</b>         | -0,76                                                   | 33,60                       | <b>-0,36</b>         |
|       | Nisan   | -0,10                                                   | -24,67                      | <b>2,24</b>          | 0,30                                                    | -29,18                      | <b>-0,45</b>         | 0,00                                                    | -24,33                      | <b>1,86</b>          | 0,34                                                    | 25,30                       | <b>1,76</b>          |
|       | Mayıs   | 0,20                                                    | -36,97                      | <b>-2,29</b>         | 0,39                                                    | -49,29                      | <b>-0,96</b>         | 0,30                                                    | -42,65                      | <b>-3,11</b>         | -0,28                                                   | -43,28                      | <b>-3,49</b>         |
|       | Haziran | -1,48                                                   | 64,67                       | <b>16,77</b>         | -0,95                                                   | 76,43                       | <b>3,86</b>          | -0,25                                                   | 50,43                       | <b>6,32</b>          | -3,00                                                   | -61,36                      | <b>5,72</b>          |
|       | Temmuz  | 0,10                                                    | -88,98                      | <b>-5,50</b>         | -0,10                                                   | -63,86                      | <b>-11,42</b>        | 0,46                                                    | -50,88                      | <b>-3,38</b>         | -1,05                                                   | -61,04                      | <b>-10,57</b>        |
|       | Ağustos | -1,41                                                   | -105,29                     | <b>-10,36</b>        | -0,73                                                   | -76,35                      | <b>-11,19</b>        | 1,10                                                    | -98,56                      | <b>-13,40</b>        | -4,14                                                   | -71,85                      | <b>-17,67</b>        |
|       | Eylül   | -0,79                                                   | 59,20                       | <b>-3,21</b>         | -0,10                                                   | -45,27                      | <b>-5,20</b>         | -1,10                                                   | -52,44                      | <b>-7,26</b>         | -0,62                                                   | -44,62                      | <b>-7,13</b>         |
|       | Ekim    | 0,98                                                    | 97,18                       | <b>5,42</b>          | 0,69                                                    | -73,56                      | <b>6,58</b>          | 0,40                                                    | 73,04                       | <b>12,32</b>         | -0,79                                                   | 85,93                       | <b>8,93</b>          |
|       | Kasım   | 0,01                                                    | 25,35                       | <b>0,20</b>          | -0,50                                                   | -69,07                      | <b>-2,84</b>         | 0,00                                                    | -40,50                      | <b>-1,09</b>         | 1,31                                                    | -41,40                      | <b>4,95</b>          |
|       | Aralık  | -0,20                                                   | -39,15                      | <b>-3,21</b>         | -0,02                                                   | -64,35                      | <b>-5,53</b>         | 0,00                                                    | 79,49                       | <b>2,52</b>          | -0,29                                                   | -69,57                      | <b>2,81</b>          |
| 2020  | Ocak    | 0,20                                                    | -72,24                      | <b>-9,93</b>         | -0,50                                                   | -69,63                      | <b>-10,61</b>        | 0,51                                                    | -67,45                      | <b>-5,43</b>         | 0,10                                                    | -88,11                      | <b>-8,72</b>         |
|       | Şubat   | 0,00                                                    | -47,20                      | <b>-1,33</b>         | -0,10                                                   | -60,39                      | <b>-5,33</b>         | 0,09                                                    | -51,10                      | <b>-5,13</b>         | 0,20                                                    | -56,18                      | <b>-3,56</b>         |
|       | Mart    | 0,10                                                    | -39,52                      | <b>-1,19</b>         | 0,01                                                    | -57,38                      | <b>-2,04</b>         | 0,03                                                    | 44,32                       | <b>-0,42</b>         | 1,00                                                    | -42,12                      | <b>-3,38</b>         |
|       | Nisan   | 0,00                                                    | 33,41                       | <b>4,08</b>          | 0,00                                                    | -53,01                      | <b>4,27</b>          | 0,00                                                    | -21,42                      | <b>0,20</b>          | 0,20                                                    | -45,39                      | <b>0,19</b>          |
|       | Mayıs   | 0,10                                                    | -49,50                      | <b>-3,46</b>         | -0,30                                                   | -48,82                      | <b>2,61</b>          | -0,10                                                   | -39,00                      | <b>-3,23</b>         | -0,30                                                   | -56,29                      | <b>-8,45</b>         |
|       | Haziran | 0,01                                                    | 53,22                       | <b>-1,24</b>         | 0,70                                                    | 55,43                       | <b>1,53</b>          | -1,05                                                   | -59,25                      | <b>-6,42</b>         | 0,10                                                    | 53,78                       | <b>-4,91</b>         |
|       | Temmuz  | 0,40                                                    | 42,53                       | <b>3,24</b>          | 0,50                                                    | -53,98                      | <b>-5,30</b>         | 0,48                                                    | -43,63                      | <b>-1,73</b>         | 0,00                                                    | -45,23                      | <b>-3,39</b>         |
|       | Ağustos | -0,20                                                   | 42,33                       | <b>-3,68</b>         | -0,14                                                   | 65,65                       | <b>-4,25</b>         | -2,50                                                   | -62,89                      | <b>-4,25</b>         | 0,80                                                    | -41,87                      | <b>-7,53</b>         |
|       | Eylül   | 0,00                                                    | -72,00                      | <b>-0,32</b>         | 0,00                                                    | 72,55                       | <b>-5,26</b>         | 0,00                                                    | -68,24                      | <b>-5,26</b>         | -1,30                                                   | -45,79                      | <b>-1,28</b>         |
|       | Ekim    | 0,70                                                    | -49,09                      | <b>3,58</b>          | 1,19                                                    | 50,51                       | <b>0,28</b>          | -0,40                                                   | 62,16                       | <b>0,28</b>          | 0,30                                                    | -52,58                      | <b>-1,65</b>         |
|       | Kasım   | 1,36                                                    | -44,26                      | <b>-4,00</b>         | 0,80                                                    | -56,65                      | <b>-1,01</b>         | 0,21                                                    | -50,16                      | <b>-1,02</b>         | -0,80                                                   | -66,27                      | <b>-0,52</b>         |
|       | Aralık  | 0,10                                                    | -50,88                      | <b>-4,89</b>         | 0,28                                                    | 49,63                       | <b>2,79</b>          | 0,40                                                    | 69,1                        | <b>2,79</b>          | -1,10                                                   | 61,89                       | <b>0,00</b>          |

RTF-DVR modeli ile CSRS-PPP'nin karşılaştırma sonuçlarına ilişkin istatistiksel verilere (Çizelge 4.8) bakıldığında, model değerleri ile CSRS-PPP'den elde edilen değerlerin, mart aylarında ve 2019 yılının ilk ayında en iyi uyum gösterdiği gözlemlenmiştir. Bu aylarda RTF-DVR gecikme değerleri ile CSRS-PPP gecikme değerleri arasındaki farkların  $\pm 0,3-0,5$  mm arasında değiştiği tespit edilmiştir. Değerlerin, yaz mevsimini kapsayan haziran ve temmuz aylarında ise en az uyum sergilediği tablodan görülmektedir. Genelde, gecikme etkisinin en fazla yaşandığı yaz döneminde elde edilen farkların  $\pm 1,03-1,77$  cm arasında değişim gösterdiği saptanmıştır. Geri kalan günlerde ise ortalama gecikme farkları  $\pm 0,5-0,8$  cm arasında değişmektedir. Sıcaklık, basınç ve bağıl nem parametrelerine dayalı oluşturulan destek vektör regresyonu modeli ile CSRS-PPP modelindeki değerler arasındaki en büyük farklar, 00:00 için 10 Ağustos 2019 gününde -105,29 mm ve 24 Ekim 2019 gününde 97,18 mm olarak belirlenmiştir (Şekil 4.22).

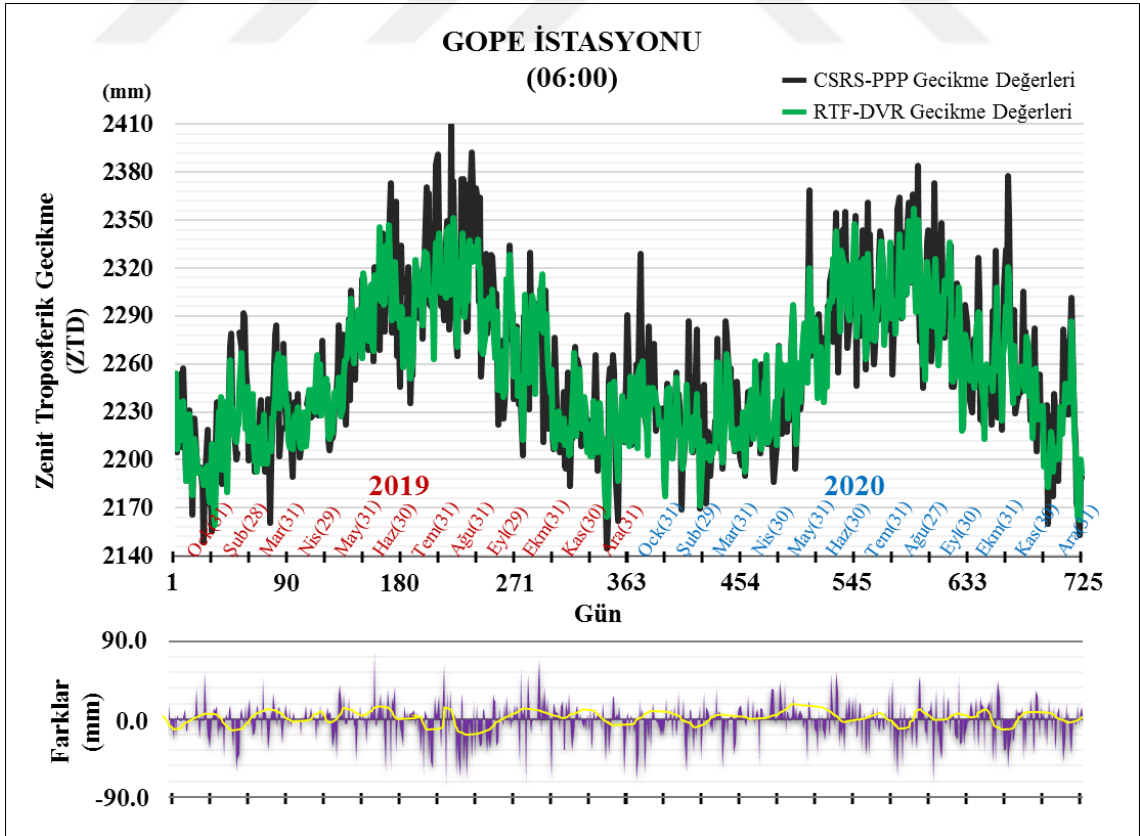
06:00'daki istatistiki sonuçlarına göre, 2019 ağustos ayına ait günlerde, RTF-DVR modeli ile CSRS-PPP değerleri arasındaki ortalama farklar (-11,19 mm) yüksek çıkmıştır. Ancak 00:00'dan farklı olarak iki yıldaki eylül ve ekim aylarında da gecikme değerleri arasında yüksek miktarda farklılıklar meydana gelmiştir. İlkbahar mevsimlerini kapsayan günlerde ise değerlerde anlamlı değişim bulunmayıp farklar çok az değişim göstermiştir. İki model arasındaki en büyük fark, 14 Haziran 2019 günü 76,43 mm civarında görülmüştür (Şekil 4.23).

12:00 için RTF-DVR modeli CSRS-PPP kestirim değerlerine çok yakın sonuçlar vermiştir. İki yıl boyunca elde edilen sonuçlara göre yeni kurulan model ile CSRS-PPP arasında  $\pm 1,0-1,2$  cm ortalama gecikme farkları gözlemlenmiştir. İki model arasındaki en büyük fark, 2020 için 13 Aralık günü 79,49 mm olarak belirlenmiştir. Ocak-Nisan arası ardışık aylarda ise gecikme değerlerinin iki farklı yöntemde de en uyumlu olduğu görülmektedir (Şekil 4.24). Son olarak 18:00'daki sonuçlara bakıldığında, model ile CSRS-PPP değeri arasında en büyük farklar, 23 Ekim 2019 günü 85,93 mm ve 8 Ocak 2020 günü -88,11 mm şeklinde meydana gelmiştir. Mart ve Nisan aylarında ise gecikme değerleri birbirleriyle en uyumlu sonuçları yakalamıştır (Şekil 4.25).

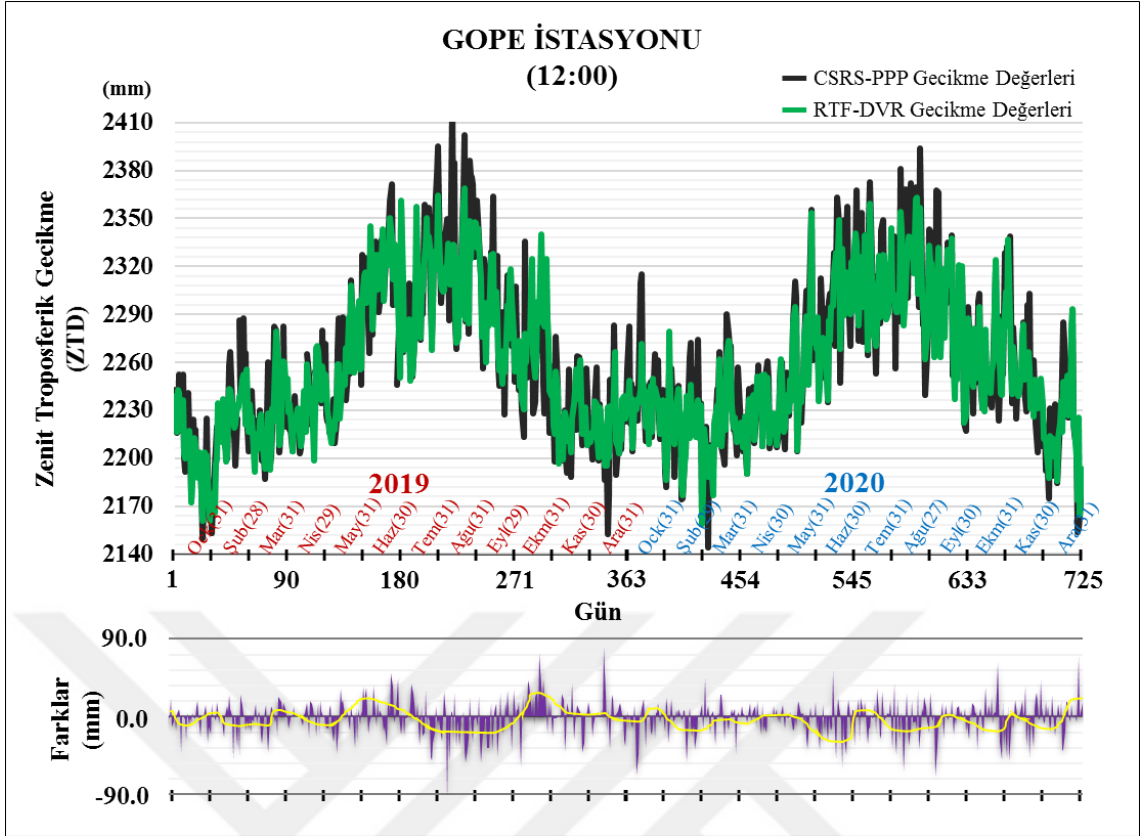
İki yıl boyunca CSRS-PPP'den elde edilen gecikme değerleri ile kurulan modellerle kestirilen model gecikme değerlerinin karşılaştırma grafikleri ve istatistikleri yorumlanmıştır. Tüm sonuçlar neticesinde, genelleme yapılacak olursa yeni oluşturulan DVR modellerinin IGS/EPN ve CSRS-PPP'deki kestirilen değerlere yakın, uyumlu ve iyi sonuçlar ürettiği görülmüştür.



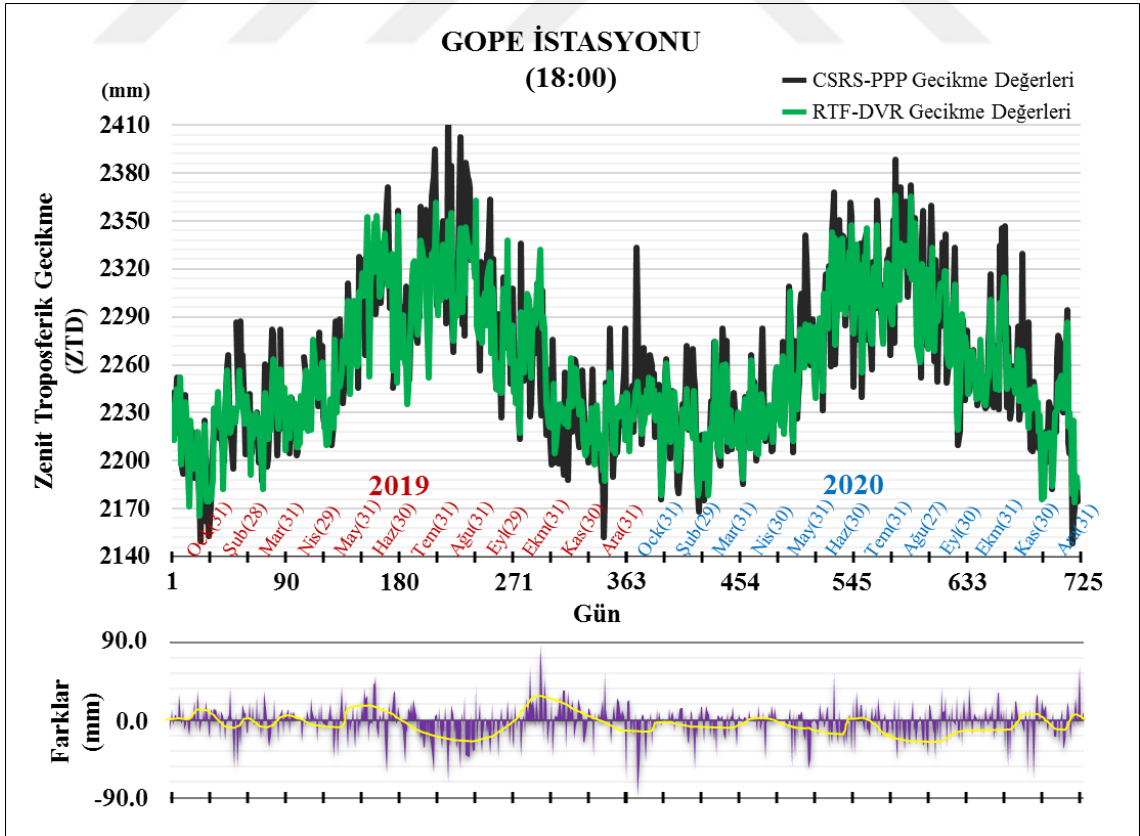
Şekil 4.22. RTF-DVR (00:00) modelinin CSRS-PPP gecikme değerleriyle uyumu



Şekil 4.23. RTF-DVR (06:00) modelinin CSRS-PPP gecikme değerleriyle uyumu



Şekil 4.24. RTF-DVR (12:00) modelinin CSRS-PPP gecikme değerleriyle uyumu



Şekil 4.25. RTF-DVR (18:00) modelinin CSRS-PPP gecikme değerleriyle uyumu

#### 4.4.5. Uyuşumsuz verilerin model sonuçlarına etkisi

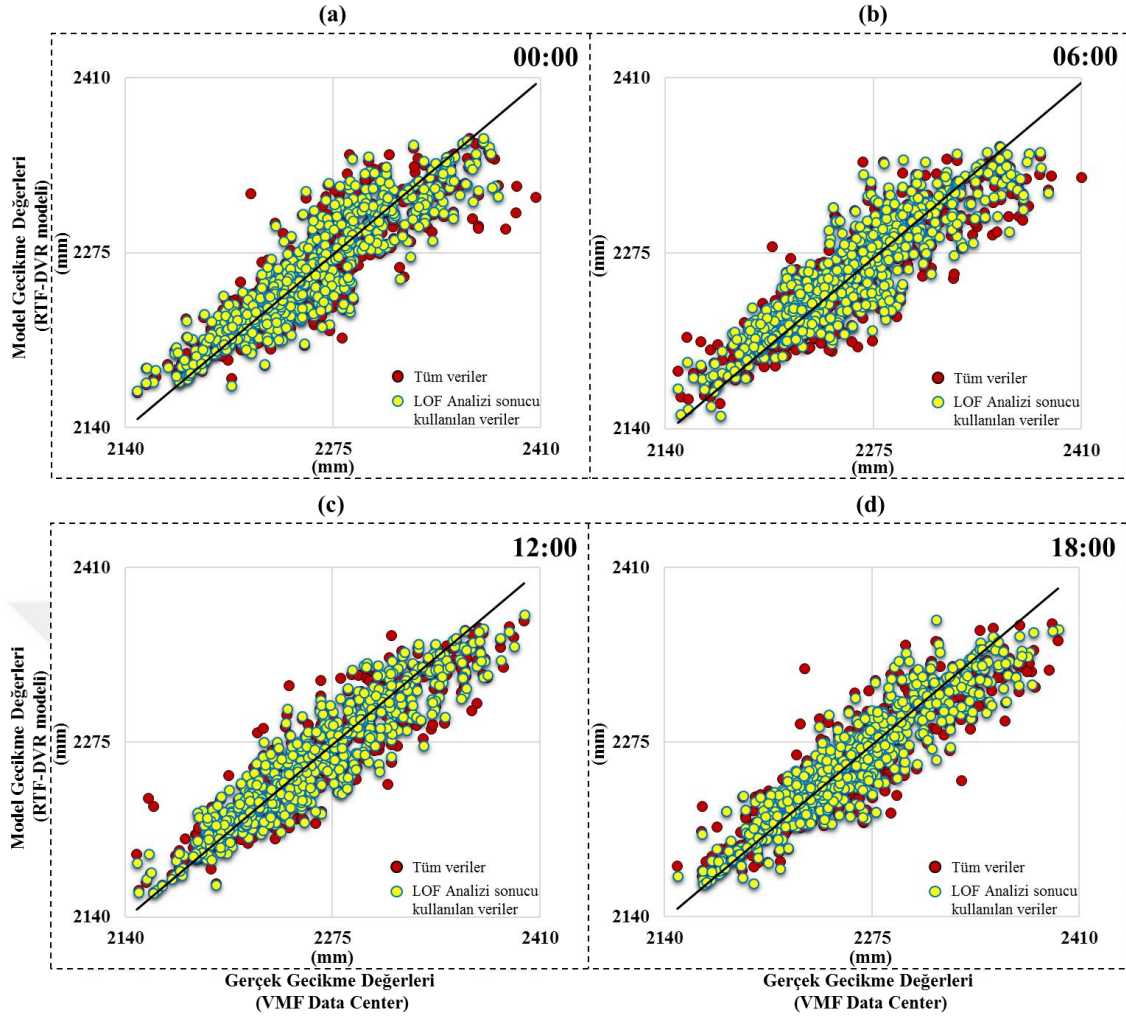
Veri setindeki örneklem grupları arasında uyumsuz verilerin model sonuçları üzerindeki etkisini görmek amacıyla çok değişkenli aykırı gözlem analizlerinden Local Outlier Factor (LOF) tekniği kullanılmıştır. Daha öncesinde sıcaklık, basınç ve bağıl nem değişkenleri için istatistiki değerlendirme yapılmış olup uyumsuz verilerin olduğu saptanmıştır. Normalde uyumsuz olan verilerin veri setinden çıkartılması gerekir. Ancak çalışma kapsamında ani hava değişiminden kaynaklı etkilerin olması, küresel ısınma sonucu hava şartlarında beklenilmeyen koşulların gerçekleşmesi gibi etmenlerden dolayı bireysel olarak uyumsuz veri çıkartılmamıştır. Bunun yerine tüm değişkenler birlikte düşünülerek gruplamanın dışında kalan uyumsuz veriler/günler uygulamadan çıkartılmıştır.

Çalışmadaki en iyi model olarak seçilen RTF-DVR’de, uyumsuz olduğu tespit edilen veriler ayıklanmış ve analizler tekrarlanarak sonuçları irdelenmiştir. LOF analiz sonuçlarına göre toplam 9 güne ait verinin uyumsuz veri olduğu tespit edilmiştir. Çizelge 4.9’a bakıldığında, özellikle 2019’daki temmuz ve ağustos aylarında veriler uyumsuz çıkmıştır. Bunun nedeni, sıcaklık ve nemin yükselmesiyle, su buharı basıncı ve gecikme etkisinin büyümesi şeklinde belirlenmiştir. Bu günlerde bu durum normal dağılım sergilememiştir. Aşağıda uyumsuz olarak tespit edilen günler sıralanmıştır.

**Çizelge 4.9.** LOF tekniğine göre aykırı olarak tespit edilen günler

| Günler |           |            |         |      |         |         |
|--------|-----------|------------|---------|------|---------|---------|
| 2019   | 23 Temmuz | 10 Ağustos | 23 Ekim | 2020 | 4 Eylül | 8 Ocak  |
|        | 30 Temmuz | 18 Ağustos |         |      |         | 15 Ocak |
|        |           | 21 Ağustos |         |      |         |         |
|        |           |            |         |      |         |         |

Uyuşumsuz olan verilerin elimine edilmesi sonucunda modellerden elde edilen analiz sonuçları EK-A’da sunulmuştur. Sonuçlara göre, uyumsuz verilerin veri setinden çıkması durumunda model sonuçlarının %5-6 civarında iyileştiği görülmüştür. Aşağıda tüm verilerin kullanılması sonucu elde edilen değerler ile gerçek değerler arasındaki ilişki ve LOF analizi sonucu kullanılan verilerden elde edilen değerler ile gerçek değerler arasındaki ilişki gösterilmiştir. Burada LOF analizi uygulamasından sonra elde edilen değerlerin gerçek gecikme değerleriyle daha uyumlu olduğu belirlenmiştir (Şekil 4.26).

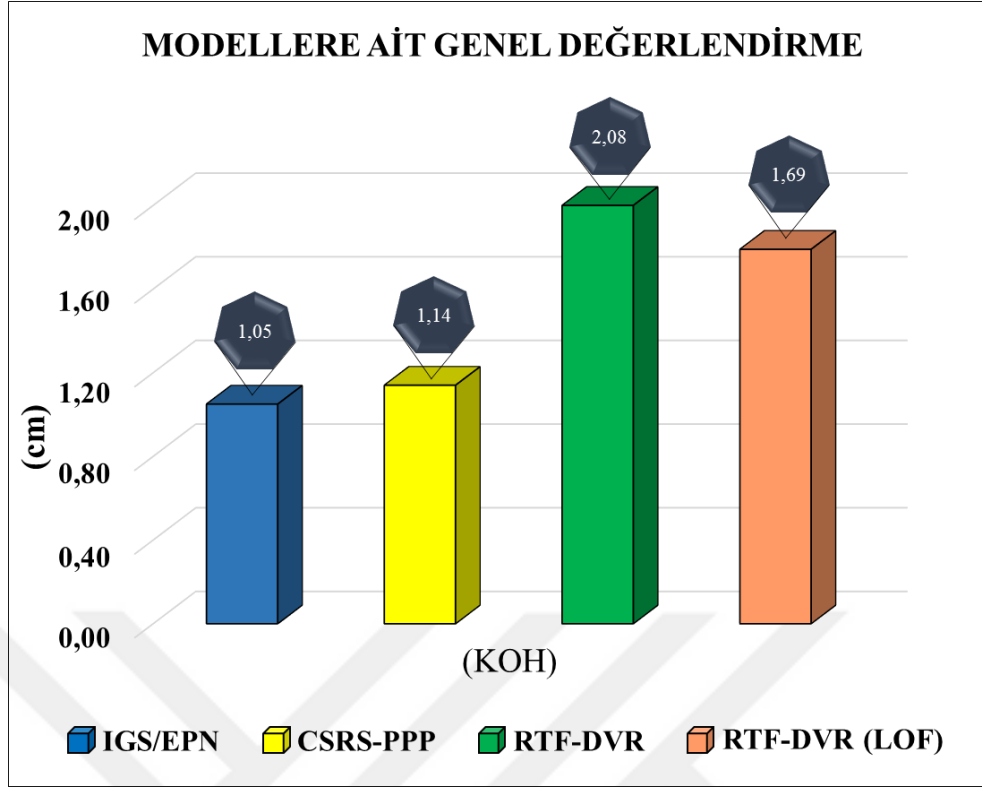


Şekil 4.26. Farklı (1-Tüm, 2-LOF) verilerden elde edilen değerler ile gerçek değerler arasındaki ilişki

#### 4.5. Farklı Troposferik Gecikme Kestirim Modellerinin Karşılaştırılması

Bu bölümde, troposferik gecikme değerinin kestirilmesinde RTF kernel fonksiyonları ile kurulan destek vektör regresyon yönteminin başarısı araştırılmıştır. Çalışma için gerçek ZTD değerleri (VMF1), VMF veri merkezinden elde edilmiştir. Bu gerçek ZTD değerleri, diğer kestirim modellerinden elde edilen ZTD değerleriyle karşılaştırılmıştır.

Karşılaştırma amacıyla kullanılan diğer modeller ise RTF-DVR ile RTF-DVR (LOF) kestirim modeli, IGS/EPN kestirim modeli ve son olarak CSRS-PPP kestirim modeli şeklinde sıralanmaktadır. Farklı kestirim modelleriyle elde edilen değerlere ait istatistiksel sonuçlar değerlendirilip yorumlanmıştır. Modellerin genel değerlendirmesi yapılırken gerçek ZTD değerleri ve modeller arasındaki karesel ortalama hata farkları incelenmiştir (Şekil 4.27).



Şekil 4.27. Farklı ZTD kestirim modellerinin karesel ortalama hata değerleri

Şekil 4.27’de verilen sonuçlara göre, IGS/EPN modelinden elde edilen ZTD değerleri ile gerçek ZTD değerleri arasındaki KOH farkının 1,05 cm olduğu belirlenmiştir. CSRS-PPP modelinden elde edilen ZTD değerleri ile gerçek ZTD değerleri arasındaki KOH farkı ise 1,14 cm olarak bulunmuştur. İki farklı modelin de, gerçek ZTD değerlerini  $\pm 1$  cm’e kadar yakın sonuçlarla elde ettiği görülmüştür. Ayrıca, IGS/EPN ve CSRS-PPP ZTD kestirim modellerinin kendi aralarındaki karesel ortalama hataları incelendiğinde, birbirine çok yakın sonuçlar ( $\sim 0,09$  cm) ürettiği görülmektedir. RTF-DVR modelleri ele alındığında, uyuşumsuz verilerden arındırılmış modelin, tüm verilerle kurulan modele göre 0,39 cm KOH farkıyla daha iyi sonuçlar verdiği saptanmıştır. RTF-DVR ve RTF-DVR (LOF) modelinden elde edilen ZTD değerleri ile gerçek ZTD değerleri arasındaki KOH farkları sırasıyla 2,08 cm ve 1,69 cm olarak elde edilmiştir. RTF-DVR (LOF) modeli baz alınacak olursa, yeni modelin diğer değerlendirme servisleri olan IGS/EPN ve CSRS-PPP’deki değerlere de oldukça yakın sonuçlar ( $\pm 0,6$  cm) elde ettiği sonucuna varılmıştır. Sonuç olarak, makine öğrenmesine dayalı yeni RTF modellerinin uluslararası kullanımı olan VMF1 modeline uyumlu, yakın ve gerçekçi sonuçlar elde ettiği görülmüştür.

## 5. SONUÇLAR VE ÖNERİLER

Troposferik gecikme kestirim modelleri, GNSS tekniği ile konum belirleme çalışmaları, GNSS meteorolojisi ve atmosfer fiziği alanında yapılan uygulamalar açısından çok önemli katkılar sunmaktadır. Özellikle uydu bazlı jeodezik ölçmelerdeki gecikme etkileri, ölçüm yapılan noktanın yeryüzü üzerindeki konumu ile yüksek oranda ilişkili olduğundan kullanılacak olan kestirim modelinin de önemi büyüktür.

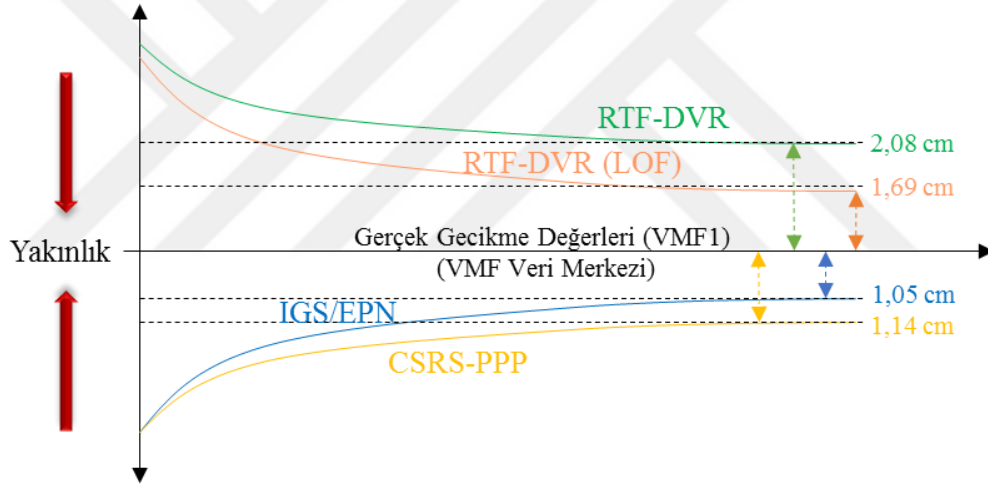
Çalışmada, zenit troposferik gecikmenin (ZTD) kestirimi için ilk defa uygulanmış olan, makine öğrenimi yaklaşımlarından DVR yöntemleri ile analizler gerçekleştirilmiştir. DVR yöntemi için 3 farklı kernel fonksiyonu (doğrusal, polinomal ve radyal tabanlı) kullanılmış olup 4 farklı zaman dilimi için modeller oluşturulmuştur. Bu modellerin oluşturulmasında, istasyona ait gerçek meteorolojik veriler ve gerçek gecikme değerleri kullanılmıştır. Uygulamada kullanmak üzere hazırlanan veri setindeki bütün bilgiler, sayısal formda olup veride herhangi bir işlem (normalizasyon, dönüşüm vb.) yapılmamıştır. Bu konuda, gerçek değerlerin ham verilerle kestirilmek istenmesi ve verilerin çalışma için uygun nitelikte olması etken olmuştur. Farklı zaman dilimleri ve kernel fonksiyonları ile elde edilen kombinasyonlar için toplam 12 model meydana gelmiştir. Bu modellerin kıyaslanması sonucu, troposferik gecikme modeli kestirimi için en iyi ve en uygun model “RTF-DVR” modeli olarak tercih edilmiştir (Çizelge 5.1).

**Çizelge 5.1.** Farklı DVR modellerindeki performans sonuçları

| MODEL                |                  | MODEL<br>BAŞARISI<br>(R <sup>2</sup> ) | OMH DEĞERİ<br>(mm) | OMYH<br>ORANI<br>(%) | KOH DEĞERİ<br>(mm) |
|----------------------|------------------|----------------------------------------|--------------------|----------------------|--------------------|
| <b>Doğrusal-DVR</b>  | <i>Eğitim H.</i> | 0,78                                   | 17,1772            | 0,7566               | 22,0198            |
|                      | <i>Test H.</i>   | 0,74                                   | 18,8468            | 0,8288               | 24,4172            |
| <b>Polinomal-DVR</b> | <i>Eğitim H.</i> | 0,77                                   | 17,7399            | 0,7829               | 22,7445            |
|                      | <i>Test H.</i>   | 0,76                                   | 18,6715            | 0,8245               | 23,8558            |
| <b>RTF-DVR</b>       | <i>Eğitim H.</i> | <b>0,84</b>                            | <b>13,1020</b>     | <b>0,5773</b>        | <b>19,2633</b>     |
|                      | <i>Test H.</i>   | <b>0,77</b>                            | <b>17,2362</b>     | <b>0,7617</b>        | <b>22,5342</b>     |

Eğitim ve test verileri birlikte düşünülerek, RTF-DVR modeli, gecikme değerlerini VMF veri merkezinin yayınladığı gerçek gecikme değerlerine göre  $\pm 20$  mm'ye kadar yakın sonuçlar ile kestirdiği gözlenmiştir (Çizelge 5.1). Ham verilerin işlenmeden ve uyuşumsuz verilerin ayıklanmadan doğrudan kurulan modellerle bu sonuçların elde edilmesi gayet başarılı olarak görülebilir. Ancak yine de uyuşumsuz

verilerin sonuç üzerindeki etkisini görmek ve daha iyi kestirim modeli elde etmek amacıyla uyuşumsuz verilerden arındırılmış veri setleriyle de modeller kurulmuştur. LOF analizi ile uyuşumsuz verilerin ayıklanması sonucu oluşturulan RTF-DVR (LOF) modelinin, tüm verilerin modelde kullanılmasıyla oluşturulan RTF-DVR modeline göre yaklaşık % 6 (KOH oranlarına göre) civarında iyi sonuçlar verdiği tespit edilmiştir. Bu nedenle modelin, uyuşumsuz veriler/günler olmadan kurulmasının gerçeğe yakın sonuçlar elde etmede daha iyi başarı yakalayacağı görülmüştür. Yeni oluşturulan RTF-DVR (LOF) modelinden elde edilen gecikme değerleri, gerçek ZTD değerlerine  $\pm 1,69$  cm yaklaşıkyla belirlenebilmiştir. Diğer değerlendirme servislerinden olan IGS/EPN ve CSRS-PPP'deki değerlere de oldukça yakın sonuçlar ( $\sim 0,6$  cm) elde edilmiştir (Şekil 5.1). Farklı kestirim modellerinden elde edilen gecikme değerlerinin grafikleri Ek-B'de detaylı şekilde ayrıca verilmiştir.



Şekil 5.1. Çalışmadaki farklı gecikme model sonuçlarının birbirine yakınlık derecesi

Günümüzde standart troposferik modeller deneysel yöntemlere dayalı olarak elde edilmektedir. Modellerde, belirli bir kesim için arazi koşullarına göre değerler hesaplanmakta olup geri kalan noktalar (meteoroloji sensör olmayan istasyonlar) için enterpolasyon ile elde edilen ortalama değerler kullanılmaktadır. Ancak gün içerisinde sürekli değişen troposfer göz önüne alındığında ani hava değişimleri meydana gelmekte ve atmosfer sürekli farklı etkilere sahip olmaktadır. Ayrıca istasyonlar arasındaki baz uzunlukları arttıkça troposferik gecikme hatası da bu doğrultuda artmaktadır. Bu da troposferik modellerin dezavantajı olarak karşımıza çıkmaktadır.

Diğer troposferik gecikme kestirimi uygulamaları incelendiğinde ise literatürde lineer enterpolasyon ve sinir ağlarına dayalı çalışmalar göze çarpmaktadır. Az sayıda uygulama alanı bulan lineer enterpolasyon yöntemleri ve sinir ağlarında, istasyon sayısı çok sayıda tutularak belirli günler için değerlendirme yapılırken, bu çalışmadaki amaç istasyon sayısından bağımsız, farklı yıllardaki tüm gün verilerini kullanarak genel çözüm üreten bir gecikme modeli kurmak olmuştur. Burada yeni oluşturulan modelin diğer modellere göre avantajları söz konusudur.

- İlk olarak çalışma için uygulanan yeni modelde troposfer gecikme hatası, tamamen istasyonlar arasındaki baz uzunluklarından bağımsız hale gelmiştir. Çünkü bu metodolojide modeller oluşturulurken gerçek veriler kullanıldığı için o bölgedeki iki yıl boyunca tüm arazi koşulu senaryoları (o anki sıcaklık, basınç, bağıl nem değerleri) modele entegre edilmiştir.
- Seçilen GOPE istasyonundan uzaklaşıp başka bir noktada gecikme kestirimi yapılmak istendiğinde herhangi bir entepole işlemi yapılmadan veya entegreli istasyonlara gerek kalmadan, sadece sıcaklık, basınç ve nem değerleri taşınabilir cihazlarla ölçülerek yeni modelden gecikme değeri kestirilebilmektedir.
- DVR modellerinin farklı bölgede bulunan istasyonlar için de kullanılabilir olması ve enterpolasyondan bağımsız olması modelin stabil olduğu anlamına gelmektedir.
- Bu çalışmada elde edilen makine öğrenmesi modelleri ve model sonuçları akademik yazılımlara da dâhil edilebilir niteliktedir.
- Ayrıca şuana kadar uygulama yapılan GOPE istasyonunda tümüyle gerçek değerler kullanılarak gecikme kestirimi yapılmıştır. Eğer bu istasyon GNSS-meteorolojik entegreli istasyon olmayıp entepole ile değerlerin taşındığı istasyon olsaydı yeni kurulmuş olan RTF-DVR modelin diğer modellere (CSRS-PPP ve IGS/EPN) göre daha iyi sonuçlar vermesi beklenmektedir.
- Makine öğrenimi ve modelleme için önemli bir konu olan veri sayısı (hacmi) dikkate alındığında, eğer modeller, uzun vadeli ve farklı arazi koşullarına ait veriler ile kurulursa model başarılarında ciddi iyileşmeler görülebilir.
- Yeni DVR modelleri, toplam gecikme değerlerinin belirlenmesi için tek bir modelle bileşenlere (hidrostatik+ıslak) ayrılmaksızın kullanılmaktadır. Ancak bu bileşenleri etkileyen ek parametreler de (sera gazı, çeşitli atm. kütleler) ayrı ayrı tek model içerisinde değerlendirilse daha başarılı sonuçlar çıkacağı muhtemeldir.

## KAYNAKLAR

- Angrisano, A., Gaglione, S., Gioia, C., Massaro, M. ve Robustelli, U., 2013, Assessment of NeQuick ionospheric model for Galileo single-frequency users, *Acta Geophysica*, 61(6), 1457-1476.
- Askne, J. ve Nordius, H., 1987, Estimation of tropospheric delay for microwaves from surface weather data, *Radio science*, 22: 379-386.
- Atalay, M., 2018, Anlık Şiddetli Yağışların GPS Konum Doğruluğuna Etkisi Üzerine Bir Çalışma, *Necmettin Erbakan Üniversitesi, Fen Bilimleri Enstitüsü, Yüksek Lisans Tezi*, Konya.
- Awad, M. ve Khanna, R., 2015, Efficient-learning machines: theories, concepts, and applications for engineers and system designers, *Springer nature*, p. 268.
- Barindelli, S., Realini, E., Venuti, G., Fermi, A. ve Gatti, A., 2018, Detection of water vapor time variations associated with heavy rain in northern Italy by geodetic and low-cost GNSS receivers, *Earth, Planets and Space*, 70(1), 1-18.
- Barnes, J., 2015, Azure Machine Learning: Microsoft Azure Essentials, *Microsoft Press, 1st ed*, Washington.
- Benevides, P., Catalao, J. ve Nico, G., 2019, Neural Network Approach to Forecast Hourly Intense Rainfall Using GNSS Precipitable Water Vapor and Meteorological Sensors, *Remote Sensing*, 11(8), 966.
- Berrada-Baby, H., Gole, P. ve Lavernat, J., 1988, A model for the tropospheric excess path length of radio waves from surface meteorological measurements, *Radio science*, 23(6), 1023-1038.
- Berry, M. J. ve Linoff, G. S., 2004, Data mining techniques: for marketing, sales, and customer relationship management, *John Wiley & Sons*.
- Bevis, M., Businger, S., Chiswell, S., Herring, T. A., Anthes, R. A., Rocken, C. ve Ware, R. H., 1994, GPS meteorology: Mapping zenith wet delays onto precipitable water, *Journal of applied meteorology*, 33(3), 379-386.
- Bezicioğlu, M., 2018, Sınırlı uydu görüşünde yüksek frekanslı kinematik GNSS-PPP yönteminin performansı: Sarsma tablası deneyleri, *Gebze Teknik Üniversitesi, Fen Bilimleri Enstitüsü, Yüksek Lisans Tezi*, Kocaeli.
- Boehm, J. ve Schuh, H., 2004, Vienna mapping functions in VLBI analyses, *Geophysical Research Letters*, 31(1).
- Boehm, J., Heinkelmann, R. ve Schuh, H., 2007, Short note: a global model of pressure and temperature for geodetic applications, *Journal of Geodesy*, 81(10), 679-683.

- Boehm, J., Möller, G., Schindelegger, M., Pain, G. ve Weber, R., 2015, Development of an improved empirical model for slant delays in the troposphere (GPT2w), *GPS solutions*, 19(3), 433-441.
- Boehm, J., Werl, B. ve Schuh, H., 2006a, Troposphere mapping functions for GPS and very long baseline interferometry from European Centre for Medium-Range Weather Forecasts operational analysis data, *Journal of geophysical research: solid earth*, 111(B2).
- Boehm, J., Niell, A., Tregoning, P. ve Schuh, H., 2006b, Global Mapping Function (GMF): A new empirical mapping function based on numerical weather model data, *Geophysical Research Letters*, 33(7).
- Bonaccorso, G., 2017, Machine learning algorithms, *Packt Publishing Ltd*.
- Breunig, M. M., Kriegel, H. P., Ng, R. T. ve Sander, J., 2000, LOF: identifying density-based local outliers, *In Proceedings of the 2000 ACM SIGMOD international conference on Management of data*, pp. 93-104.
- Brownlee, J., 2016, Machine learning mastery with python, *Machine Learning Mastery Pty Ltd*, 527, 100-120.
- Burges, C. J., 1998, A tutorial on support vector machines for pattern recognition, *Data mining and knowledge discovery*, 2(2), 121-167.
- Bülbül, S., 2018, TUSAGA-aktif noktalarında renkli gürültülerden arındırılmış hız bileşenlerinin belirlenmesi, *Selçuk Üniversitesi, Fen Bilimleri Enstitüsü, Doktora Tezi*, Konya.
- Byun, S. H., Bar-Sever, Y. E. ve Gendt, G., 2005, The new tropospheric product of the International GNSS Service, *In Proceedings of the 18th International Technical Meeting of the Satellite Division of The Institute of Navigation*, pp. 241-249.
- Cateni, S., Colla, V., Vannucci, M., Aramburo, J. ve Trevino, A. R., 2008, Outlier detection methods for industrial applications, *Advances in Robotics, Automation and Control*, 265-282.
- Cervantes, J., Garcia-Lamont, F., Rodriguez-Mazahua, L. ve Lopez, A., 2020, A comprehensive survey on support vector machine classification: Applications, challenges and trends, *Neurocomputing*, 408, 189-215.
- Chang, C. C. ve Lin, C. J., 2002, Training v-support vector regression: theory and algorithms, *Neural computation*, 14(8), 1959-1977.
- Chen, N., 2004, Support vector machine in chemistry, *World Scientific*.
- Combrinck, W. L. ve Chin, M., 2001, IGS stations: station and regional issues, *Physics and Chemistry of the Earth, Part A: Solid Earth and Geodesy*, 26(6-8), 539-544.

- Cortes, C. ve Vapnik, V., 1995, Support-vector network, *Machine Learning*, 20, 273–297.
- Çölkesen, İ., 2009, Uzaktan algılamada ileri sınıflandırma tekniklerinin karşılaştırılması ve analizi, *Gebze Yüksek Teknoloji Enstitüsü, Mühendislik ve Fen Bilimleri Enstitüsü, Yüksek Lisans Tezi*, Kocaeli.
- Çölkesen, İ., 2015, Yüksek Çözünürlüklü Uydu Görüntüleri Kullanarak Benzer Spektral Özelliklere Sahip Doğal Nesnelerin Ayırt Edilmesine Yönelik Bir Metodoloji Geliştirme, *İstanbul Teknik Üniversitesi, Fen Bilimleri Enstitüsü, Doktora Tezi*, İstanbul.
- Dach, R., Lutz, S., Walser, P. ve Fridez, P., 2015, Bernese GNSS Software Version 5.2, User manual, *Astronomical Institute, University of Bern, Bern Open Publishing*, DOI: 10.7892/boris.72297, ISBN: 978-3-906813-05-9.
- Davis, J. Herring, T. Shapiro, I. Rogers, A. ve Elgered, G., 1985, Geodesy by radio interferometry: Effects of atmospheric modeling errors on estimates of baseline length, *Radio science*, 20: 1593-1607.
- Debnath, R. ve Takahashi, H., 2004, Kernel selection for the support vector machine, *IEICE transactions on information and systems*, 87(12), 2903-2904.
- Deng, F., Guo, S., Zhou, R. ve Chen, J., 2015, Sensor multifault diagnosis with improved support vector machines, *IEEE Transactions on Automation Science and Engineering*, 14(2), 1053-1063.
- Deng, J., Xu, M., Yu, X. ve Zhang, A., 2018, Interpolation Estimation Method of Tropospheric Delay for Long Baseline Network RTK Based on Support Vector Machine, *In IOP Conference Series: Earth and Environmental Science*, 192(1), 012069.
- Ding, M., 2018, A neural network model for predicting weighted mean temperature, *Journal of Geodesy*, 92(10), 1187-1198.
- Douša, J. ve Vlacavovic, P., 2014, Real-time zenith tropospheric delays in support of numerical weather prediction applications, *Advances in Space Research*, 53: 1347-1358.
- Douša, J., Dick, G., Kačmařík, M., Brožková, R., Zus, F., Brenot, H., ... ve Kaplon, J., 2016, Benchmark campaign and case study episode in central Europe for development and assessment of advanced GNSS tropospheric models and products, *Atmospheric Measurement Techniques*, 9(7), 2989-3008.
- Emardson, T.R. ve Johansson, J.M., 1998, Spatial interpolation of the atmospheric water vapor content between sites in a ground-based GPS Network, *Geophysical Research Letters*, 25: 3347-3350.
- Eray, O., 2008, Destek Vektör Makineleri ile Ses Tanıma Uygulaması, *Pamukkale Üniversitesi, Fen Bilimleri Enstitüsü, Yüksek Lisans Tezi*, Denizli.

- Erkan, Y., 2008, Troposferik Gecikme Modellerinin GPS Nokta Konumlarına Etkisi, *Zonguldak Karaelmas Üniversitesi, Fen Bilimleri Enstitüsü, Yüksek Lisans Tezi*, Zonguldak.
- Erol, S., Mutlu, B., Erol, B., Katıgöz, S. ve Alkan, R. M., 2020, Antarktika Kıtasında Hassas Nokta Konumlama (Precise Point Positioning-PPP) Tekniğinin Performansının İncelenmesi, *Afyon Kocatepe Üniversitesi Fen ve Mühendislik Bilimleri Dergisi*, 20(5), 844-856.
- Galimberti, E., 2017, The Tree of Machine Learning Algorithms, *Teradata Blog*.
- Geiger, A., 1988, Atmospheric refraction and other important biases in GPS carrier phase observations, *Atmospheric effects on geodetic space measurements, Monograph*, 12, 1543.
- Gökdaş, Ö., 2014, GPS Meteorolojisi: İstanbul İçin Bir Uygulama, *İstanbul Teknik Üniversitesi, Fen Bilimleri Enstitüsü, Yüksek Lisans Tezi*, İstanbul.
- Gui, G., Pan, H., Lin, Z., Li, Y. ve Yuan, Z., 2017, Data-driven support vector machine with optimization techniques for structural health monitoring and damage detection, *KSCE Journal of Civil Engineering*, 21(2), 523-534.
- Gunn, S. R., 1998, Support Vector Machines for Classification and Regression, *ISIS Technical Report*, 14 May 1998.
- Guo, Q., 2015, Precision comparison and analysis of four online free PPP services in static positioning and tropospheric delay estimation, *GPS solutions*, 19(4), 537-544.
- Hadas, T., Teferle, F. N., Kazmierski, K., Hordyniec, P. ve Bosy, J., 2017, Optimum stochastic modeling for GNSS tropospheric delay estimation in real-time, *GPS Solutions, Springer*, 21(3), pp. 1069–1081.
- Han, J., Kamber, M. ve Pei, J., 2011, Data mining concepts and techniques third edition, *The Morgan Kaufmann Series in Data Management Systems*, 5(4), 83-124.
- Han, S., Qubo, C. ve Meng, H., 2012, Parameter selection in SVM with RBF kernel function, *In World Automation Congress*, pp. 1-4, IEEE.
- Hegde, J. ve Rokseth, B., 2020, Applications of machine learning methods for engineering risk assessment–A review, *Safety science*, 122, 104492.
- Hoffmann-Wellenhof B., Lichtenegger H. ve Collins J., 2000, GPS: Theory and Practice, *4th Edition, Springer, Ver-lag/Wien New York*.
- Hopfield, H., 1969, Two-quartic tropospheric refractivity profile for correcting satellite data, *Journal of Geophysical Research*, 74: 4487-4499.

- Huang, C., Davis, L. S. ve Townshend, J. R. G., 2002, An assessment of support vector machines for land cover classification, *International Journal of remote sensing*, 23(4), 725-749.
- Ifadis, I. M., 1992, The excess propagation path of radio waves: Study of the influence of the atmospheric parameters on its elevation dependence, *Survey Review*, 31(243), 289-298.
- Isioye, O. A., Combrinck, L., Botai, J. O. ve Munghemezulu, C., 2015, The potential for observing African weather with GNSS remote sensing, *Advances in Meteorology*, 2015, pp. 1–16.
- Ito, K. ve Nakano, R., 2003, Optimizing support vector regression hyperparameters based on cross-validation, *In Proceedings of the International Joint Conference on Neural Networks*, Vol. 3, pp. 2077-2082, IEEE.
- İlhan, E. G. E. ve Şener, Z., 2013, Performans ölçümünde kullanılan yöntemler: performans karnesi ve kumanda paneli karşılaştırması, *Muhasebe ve Finansman Dergisi*, (57), 107-120.
- İnal, F., 2010, Artificial neural network prediction of tropospheric ozone concentrations in Istanbul, Turkey, *CLEAN–Soil, Air, Water*, 38(10), 897-908.
- James, G., Witten, D., Hastie, T. ve Tibshirani, R., 2013, An introduction to statistical learning, *Springer*, Vol. 112, p. 18, New York.
- Juni, I. ve Rózsa, S., 2019, Validation of a new model for the estimation of residual tropospheric delay error under extreme weather conditions, *Periodica Polytechnica Civil Engineering*, 63(1), 121-129.
- Kahveci, M. ve Yıldız, F., 2017, Uydularla Konum Belirleme Sistemleri (GPS/GNSS): Teori-Uygulama, 8. Basım, Nobel Yayıncılık, Ankara.
- Katsougiannopoulos, S. ve Pikridas, C., 2009, Prediction of zenith tropospheric delay by multi-layer perceptron. *Journal of Applied Geodesy*, 3, 223–229.
- Kecman, V., 2001, Learning and Soft Computing: Support Vector Machines, *Neural Networks and Fuzzy Logic Models*, MIT Press, Cambridge-Massachusetts.
- Kim, Y., Chong, J. W., Chon, K. H. ve Kim, J., 2012, Wavelet-based AR–SVM for health monitoring of smart structures, *Smart Materials and Structures*, 22(1), 015003.
- Kleijer, F. 2004, Troposphere Modeling and Filtering for Precise GPS Leveling, *Delft University of Technology, Department of Mathematical Geodesy and Positioning, Ph.D. Dissertation*, Kluyverweg-1, Netherlands.
- Kouba, J. ve Héroux, P., 2001, Precise point positioning using IGS orbit and clock products, *GPS solutions*, 5(2), 12-28.

- Kowalczyk, A., 2017, Support Vector Machines Succinctly, *Morrisville: Syncfusion*.
- Labib, B., Yan, J., Barriot, J. P., Zhang, F. ve Feng, P., 2019, Monitoring Zenithal Total Delays over the three different climatic zones from IGS GPS final products: A comparison between the use of the VMF1 and GMF mapping functions, *Geodesy and Geodynamics*, 10(2), 93-99.
- Leandro, R., Santos, M. ve Langley, R., 2006, UNB neutral atmosphere models: development and performance, *In Proceedings of the 2006 National Technical Meeting of The Institute of Navigation*, pp. 564-573.
- Leick, A., 2004, GPS Satellite Surveying, *3rd Edition*, Wiley, New Jersey, USA.
- Li, J., Allinson, N., Tao, D. ve Li, X., 2006, Multitraining support vector machine for image retrieval, *IEEE Transactions on Image Processing*, 15(11), 3597-3601.
- Li, S., Xu, T., Jiang, N., Yang, H., Wang, S. ve Zhang, Z., 2021, Regional Zenith Tropospheric Delay Modeling Based on Least Squares Support Vector Machine Using GNSS and ERA5 Data, *Remote Sensing*, 13(5), 1004.
- Li, Y., 2016, Real-Time Ambiguity-fixed Precise Point Positioning Using Global and Regional Reference Networks, *University of Calgary, Graduate Program In Geomatics Engineering, Doctoral Thesis*, Canada.
- Lin, C. C., 2010, Critical analysis and effectiveness of key parameters in residential property valuations, *State University of New York, The Faculty of The Graduate School of The University at Buffalo, Ph.D. Dissertation*, New York.
- Los, M., Smolak, K., Guerova, G. ve Rohm, W., 2020, GNSS-Based Machine Learning Storm Nowcasting, *Remote Sensing*, 12(16), 2536.
- Luo, M., Zeng, M., Jing, Y. ve Xiong, S., 2013, Change characteristics of GPS PWV and its relationship with precipitation, *Journal of the Meteorological Sciences*, 33, 418-423.
- Manandhar, S., Lee, Y. H., Meng, Y. S., Yuan, F. ve Ong, J. T., 2018, GPS-derived PWV for rainfall nowcasting in tropical region, *IEEE transactions on geoscience and remote sensing*, 56(8), 4835-4844.
- Mathur, A. ve Foody, G. M., 2008, Multiclass and binary SVM classification: Implications for training and classification users, *IEEE Geoscience and remote sensing letters*, 5(2), 241-245.
- Mendes, V., 1999, Modeling the neutral-atmospheric propagation delay in radiometric space techniques, *University of New Brunswick, Department of Geodesy and Geomatics Engineering, Ph.D. Dissertation*, Canada.
- Miotti, L., 2019, Machine-learning approaches for estimating tropospheric zenith path delay, *Zurich University, Institute for Geodesy and Photogrammetry, Master Thesis*, Zurich.

- Mirsa, P. ve Enge, P., 2012, Global Positioning System Signals, *Measurements, and Performance, Revised 2nd edition, Ganga-Jamuna Press, USA.*
- Mitchell, T. M., 1997, Machine learning, *WCB/McGraw-Hill, ABD.*
- Mohri, M., Rostamizadeh, A. ve Talwalkar, A., 2018, Foundations of machine learning, *MIT press.*
- Niculescu-Mizil, A. ve Caruana, R., 2005, Predicting good probabilities with supervised learning, *In Proceedings of the 22nd international conference on Machine learning*, pp. 625-632.
- Niell, A. E., 1996, Global mapping functions for the atmosphere delay at radio wavelengths, *Journal of geophysical research: solid earth*, 101(B2), 3227-3246.
- Nistor, S., Norbert-Szabolcs, S. U. B. A. ve Aurelian-Stelian, B. U. D. A., 2020, The Impact of Tropospheric Mapping Function on PPP Determination for One-Month Period, *Acta Geodynamica et Geomaterialia*, 17(2), 237-253.
- Osah, S., Acheampong, A. A., Fosu, C. ve Dadzie, I., 2021, Deep Learning model for predicting daily IGS Zenith Tropospheric Delays in West Africa using TensorFlow and Keras, *Advances in Space Research*, 10-1016.
- Osuna, E. E., Freund, R. ve Girosi, F., 1997, Support vector machines: Training and applications, *Artificial Intelligence Laboratory Technical Report, (Report No: A.I. Memo: 1602-144), Massachusetts Institute of Technology, USA.*
- Öcalan, T., 2015, GNSS ağlarında GPS hassas nokta konumlama (GPS-PPP) tekniği yaklaşımli çözümler, *Yıldız Teknik Üniversitesi, Fen Bilimleri Enstitüsü, Doktora Tezi, İstanbul.*
- Özkan, C., 1998, Uzaktan algılama verileriyle orman yangını analizi, *İstanbul Teknik Üniversitesi, Fen Bilimleri Enstitüsü, Yüksek Lisans Tezi, İstanbul.*
- Pace, B., Pacione, R., Sciarretta, C. ve Vespe, F., 2010, Estimating Zenith Total Delay Residual Fields by using Ground-Based GPS network, *In Proceedings of XX EUREF Symposium, Gävle, Sweden.*
- Palaniswami, M. ve Shilton, A., 2002, Adaptive support vector machines for regression, *In Proceedings of the 9th International Conference on Neural Information Processing (ICONIP'02), Vol. 2, pp. 1043-1049.*
- Penna, N., Dodson, A. ve Chen, W., 2001, Assessment of EGNOS tropospheric correction model, *The Journal of Navigation*, 54(1), 37-55.
- Polyakov, A., Timofeyev, Y. M. ve Virolainen, Y., 2014, Comparison of different techniques in atmospheric temperature-humidity sensing from space, *International Journal of Remote Sensing*, 35(15), 5899-5912.

- Putra, A. W. ve Lursinsap, C., 2014, Cumulonimbus prediction using artificial neural network back propagation with radiosonde indeces, *In Seminar Nasional Penginderaan Jauh*, p: 153.
- Radhika, Y. ve Shashi, M., 2009, Atmospheric temperature prediction using support vector machines, *International Journal of Computer Theory and Engineering*, 1(1), 55.
- Rostami, H., Dantan, J. Y. ve Homri, L., 2015, Review of data mining applications for quality assessment in manufacturing industry: support vector machines, *International Journal of Metrology and Quality Engineering*, 6(4), 401.
- Saastamoinen, I. I., 1972, Contribution to the theory of atmospheric refraction, *Bulletin Géodésique*, 105(1): 279–298.
- Samuel, A. L., 1959, Some Studies in Machine Learning Using the Game of Checkers, *IBM Journal of Research and Development*, 44:1.2 (1959): 210–229.
- Sapucci, L. F., 2014, Evaluation of modeling water-vapor-weighted mean tropospheric temperature for GNSS-integrated water vapor estimates in Brazil, *Journal of Applied Meteorology and Climatology*, 53(3), 715-730.
- Savaş, B., 2019, Makine öğrenme algoritmalarının konut değer tahmininde kullanımı: Ankara Gölbaşı uygulaması, *Konya Teknik Üniversitesi, Lisansüstü Eğitim Enstitüsü, Yüksek Lisans Tezi*, Konya.
- Sayıt, M. ve Sönmez, O., 2012, Reinforcement learning for peer to peer video streaming applications, *In 2012 20th Signal Processing and Communications Applications Conference (SIU)*, pp. 1-4, IEEE.
- Selbesoğlu, M. O., 2017, Global Navigasyon Uydu Sistemleri (GNSS) Gözlemlerinden Elde Edilen Islak Troposfer Gecikmesinin Yapay Sinir Ağları ile Modellenmesi, *Yıldız Teknik Üniversitesi, Fen Bilimleri Enstitüsü, Doktora Tezi*, İstanbul.
- Shalev-Shwartz, S. ve Ben-David, S., 2014, Understanding machine learning: From theory to algorithms, *Cambridge University press*.
- Smola, A. J. ve Schölkopf, B., 2004, A tutorial on support vector regression, *Statistics and computing*, 14(3), 199-222.
- Souza, C. R., 2010, Kernel functions for machine learning applications, *Creative Commons Attribution-Noncommercial-Share Alike*, 3, 29.
- Stende, M., 2006, Monitoring climate variability and change by means of GNSS data, *In Atmosphere and climate*, Springer, Berlin, Heidelberg, 275-285.
- Suparta, W. ve Alhasa, K. M., 2015, Modeling of zenith path delay over Antarctica using an adaptive neuro fuzzy inference system technique, *Expert systems with applications*, 42(3), 1050-1064.

- Şenkal, O., Yıldız, B. Y., Şahin, M. ve Pestemalcı, V., 2012, Precipitable water modelling using artificial neural network in Cukurova region, *Environmental monitoring and assessment*, 184(1), 141-147.
- Thayer, G. D., 1974, An improved equation for the radio refractive index of air, *Radio Science*, 9(10), 803-807.
- Trafalis, T. B. ve Ince, H., 2000, Support vector machine for regression and applications to financial forecasting, *In Proceedings of the IEEE-INNS-ENNS International Joint Conference on Neural Networks (IJCNN), Neural Computing: New Challenges and Perspectives for the New Millennium*, Vol. 6, pp. 348-353, IEEE.
- Tregoning, P. ve Herring, T. A., 2006, Impact of a priori zenith hydrostatic delay errors on GPS estimates of station heights and zenith total delays, *Geophysical Research Letters*, 33(23).
- Tso, B. ve Mather, P.M., 2009, Classification Methods for Remotely Sensed Data, *Second edition*, Boca Raton, FL: CRC Press.
- Tunalı, E., 2018, PPP ve IGS Gerçek Zaman Servisi ile Su Buharı Hareketlerinin İzlenmesi, *Afyon Kocatepe Üniversitesi, Fen ve Mühendislik Bilimleri Dergisi*, 18(1), 343-356.
- Turing, A. M., 1950, Computing machinery and intelligence, *Mind*, 433-460.
- URL, 2021, <https://berkarat.com/makine-ogrenmesi-model-olusturma/>, Erişim Tarihi: 27 Ocak 2021.
- Vapnik, V. N., 1995, The nature of statistical learning theory, *Springer-Verlag*, New York, USA.
- Vapnik, V. ve Chapelle, O., 2000, Bounds on error expectation for support vector machines, *Neural computation*, 12(9), 2013-2036.
- VMF Data Server, 2021, re3data.org: VMF Data Server; editing status 2020-12-14, <http://doi.org/10.17616/R3RD2H>, Erişim Tarihi: 28 Mart 2021.
- Worden, K. ve Lane, A. J., 2001, Damage identification using support vector machines, *Smart materials and structures*, 10(3), 540.
- Wu, C. H., Su, W. H. ve Ho, Y. W., 2010, A study on GPS GDOP approximation using support-vector machines, *IEEE Transactions on Instrumentation and Measurement*, 60(1), 137-145.
- Wu, Z. L., Li, C. H., Ng, J. K. Y. ve Leung, K. R., 2007, Location estimation via support vector regression, *IEEE Transactions on mobile computing*, 6(3), 311-321.
- Xiaoyan, W. P. Z., 2003, Model Selection of SVM with RBF Kernel and its Application, *Computer Engineering and Applications*, 24, 021.

- Xu, T., Li, S. ve Jiang, N., 2020, Zenith Troposphere Delay Prediction based on BPNN and LSSVM, *In EGU General Assembly Conference Abstracts*, p. 5471.
- Yalçın, C., 2016, Destek vektör regresyon ile uçaklarda anlık yakıt tüketiminin analizi, *Mimar Sinan Güzel Sanatlar Üniversitesi, Fen Bilimleri Enstitüsü, Doktora Tezi*, İstanbul.
- Yang, F., Guo, J., Zhang, C., Li, Y. ve Li, J., 2021, A Regional Zenith Tropospheric Delay (ZTD) Model Based on GPT3 and ANN, *Remote Sensing*, 13(5), 838.
- Yıldırım, S., 2006, Arıza Teşhisinde Destek Vektör Makinelerinin Kullanımı, *Fırat Üniversitesi, Fen Bilimleri Enstitüsü, Yüksek Lisans Tezi*, Elazığ.
- Yue, Y. ve Ye, T., 2019, Predicting precipitable water vapor by using ANN from GPS ZTD data at Antarctic Zhongshan Station, *Journal of Atmospheric and Solar-Terrestrial Physics*, 191, 105059.
- Yüksek, E., 2020, Makine öğrenmesi yaklaşımı ile iç mekân konum tespit sisteminin tasarımı, *Sivas Cumhuriyet Üniversitesi, Sosyal Bilimler Enstitüsü, Yüksek Lisans Tezi*, Sivas.

**EKLER**

**EK-A.** RTF-DVR (RBF-SVR) modellerine ait performans analiz sonuçları

**EK-B.** Farklı kestirim modellerinden elde edilen gecikme değerlerinin grafikleri



## EK A. RTF-DVR (RBF-SVR) modellerine ait performans analiz sonuçları

| SVR MODEL                         | RBF-SVR MODEL (UTC 00:00)                                                     |                               |                                                                               |                               | RBF-SVR MODEL (UTC 06:00)                                                     |                               |                                                                               |                               |
|-----------------------------------|-------------------------------------------------------------------------------|-------------------------------|-------------------------------------------------------------------------------|-------------------------------|-------------------------------------------------------------------------------|-------------------------------|-------------------------------------------------------------------------------|-------------------------------|
| VERİ SETİ                         | Eğitim-Test (%70-%30) Veri Seti ile Oluşturulan Model (cv:10)                 |                               | Eğitim-Test (%70-%30) Veri Seti ile Oluşturulan Model (cv:10)                 |                               | Eğitim-Test (%70-%30) Veri Seti ile Oluşturulan Model (cv:10)                 |                               | Eğitim-Test (%70-%30) Veri Seti ile Oluşturulan Model (cv:10)                 |                               |
| AYKIRI VERİ ANALİZİ               | -                                                                             |                               | <u>Local Outlier Factor (LOF) Analysis</u>                                    |                               | -                                                                             |                               | <u>Local Outlier Factor (LOF) Analysis</u>                                    |                               |
| AYKIRI VERİ SAYISI (OUTLIER DATA) | Toplam Veri Sayısı/Gün Sayısı: 725<br>Toplam Aykırı Veri Sayısı/Gün Sayısı: 0 |                               | Toplam Veri Sayısı/Gün Sayısı: 716<br>Toplam Aykırı Veri Sayısı/Gün Sayısı: 9 |                               | Toplam Veri Sayısı/Gün Sayısı: 725<br>Toplam Aykırı Veri Sayısı/Gün Sayısı: 0 |                               | Toplam Veri Sayısı/Gün Sayısı: 716<br>Toplam Aykırı Veri Sayısı/Gün Sayısı: 9 |                               |
| ÖRNEKLEM SAYISI                   | 507 Eğitim-218 Test Verisi                                                    |                               | 501 Eğitim-215 Test Verisi                                                    |                               | 507 Eğitim-218 Test Verisi                                                    |                               | 501 Eğitim-215 Test Verisi                                                    |                               |
| MODEL OLUŞTURMA AŞAMASI           | Optimizasyon Öncesi                                                           | Optimizasyon Sonrası          | Optimizasyon Öncesi                                                           | Optimizasyon Sonrası          | Optimizasyon Öncesi                                                           | Optimizasyon Sonrası          | Optimizasyon Öncesi                                                           | Optimizasyon Sonrası          |
| EN UYGUN PARAMETRE                | C parametresi: 1<br>γ : 0.1                                                   | C parametresi: 50<br>γ : 0.01 | C parametresi: 1<br>γ : 0.1                                                   | C parametresi: 50<br>γ : 0.01 | C parametresi: 1<br>γ : 0.1                                                   | C parametresi: 50<br>γ : 0.01 | C parametresi: 1<br>γ : 0.1                                                   | C parametresi: 50<br>γ : 0.01 |
| EĞİTİM HATALARI                   | R <sup>2</sup> = 0.767                                                        | R <sup>2</sup> = 0.817        | R <sup>2</sup> = 0.784                                                        | R <sup>2</sup> = 0.841        | R <sup>2</sup> = 0.757                                                        | R <sup>2</sup> = 0.796        | R <sup>2</sup> = 0.773                                                        | R <sup>2</sup> = 0.824        |
|                                   | MAE= 16.4031                                                                  | MAE= 13.9862                  | MAE= 15.4278                                                                  | MAE= 13.0275                  | MAE= 17.0802                                                                  | MAE= 14.7389                  | MAE= 16.7758                                                                  | MAE= 13.9924                  |
|                                   | MSE= 521.2752                                                                 | MSE= 410.7790                 | MSE= 455.5563                                                                 | MSE= 346.4190                 | MSE= 544.8317                                                                 | MSE= 457.3629                 | MSE= 501.4007                                                                 | MSE= 389.3132                 |
|                                   | RMSE= 22.8314                                                                 | RMSE= 20.2677                 | RMSE= 21.3438                                                                 | RMSE= 18.6123                 | RMSE= 23.3416                                                                 | RMSE= 21.3860                 | RMSE= 22.3920                                                                 | RMSE= 19.7310                 |
| TEST HATALARI                     | R <sup>2</sup> = 0.722                                                        | R <sup>2</sup> = 0.778        | R <sup>2</sup> = 0.758                                                        | R <sup>2</sup> = 0.801        | R <sup>2</sup> = 0.712                                                        | R <sup>2</sup> = 0.752        | R <sup>2</sup> = 0.750                                                        | R <sup>2</sup> = 0.764        |
|                                   | MAE= 18.4937                                                                  | MAE= 16.6662                  | MAE= 17.5141                                                                  | MAE= 15.9307                  | MAE= 19.8554                                                                  | MAE= 18.6398                  | MAE= 18.1618                                                                  | MAE= 17.6711                  |
|                                   | MSE= 664.1455                                                                 | MSE= 531.7793                 | MSE= 559.1935                                                                 | MSE= 432.3683                 | MSE= 704.2096                                                                 | MSE= 606.1229                 | MSE= 542.5317                                                                 | MSE= 510.5899                 |
|                                   | RMSE= 25.7710                                                                 | RMSE= 22.0603                 | RMSE= 23.6473                                                                 | RMSE= 20.7935                 | RMSE= 26.5369                                                                 | RMSE= 23.6196                 | RMSE= 23.2923                                                                 | RMSE= 22.5962                 |

| SVR MODEL                         | RBF-SVR MODEL (UTC 12:00)                                                     |                               |                                                                               |                               | RBF-SVR MODEL (UTC 18:00)                                                     |                               |                                                                               |                               |
|-----------------------------------|-------------------------------------------------------------------------------|-------------------------------|-------------------------------------------------------------------------------|-------------------------------|-------------------------------------------------------------------------------|-------------------------------|-------------------------------------------------------------------------------|-------------------------------|
| VERİ SETİ                         | Eğitim-Test (%70-%30) Veri Seti ile Oluşturulan Model (cv:10)                 |                               | Eğitim-Test (%70-%30) Veri Seti ile Oluşturulan Model (cv:10)                 |                               | Eğitim-Test (%70-%30) Veri Seti ile Oluşturulan Model (cv:10)                 |                               | Eğitim-Test (%70-%30) Veri Seti ile Oluşturulan Model (cv:10)                 |                               |
| AYKIRI VERİ ANALİZİ               | -                                                                             |                               | <u>Local Outlier Factor (LOF) Analysis</u>                                    |                               | -                                                                             |                               | <u>Local Outlier Factor (LOF) Analysis</u>                                    |                               |
| AYKIRI VERİ SAYISI (OUTLIER DATA) | Toplam Veri Sayısı/Gün Sayısı: 723<br>Toplam Aykırı Veri Sayısı/Gün Sayısı: 0 |                               | Toplam Veri Sayısı/Gün Sayısı: 714<br>Toplam Aykırı Veri Sayısı/Gün Sayısı: 9 |                               | Toplam Veri Sayısı/Gün Sayısı: 723<br>Toplam Aykırı Veri Sayısı/Gün Sayısı: 0 |                               | Toplam Veri Sayısı/Gün Sayısı: 714<br>Toplam Aykırı Veri Sayısı/Gün Sayısı: 9 |                               |
| ÖRNEKLEM SAYISI                   | 506 Eğitim-217 Test Verisi                                                    |                               | 499 Eğitim-215 Test Verisi                                                    |                               | 506 Eğitim-217 Test Verisi                                                    |                               | 499 Eğitim-215 Test Verisi                                                    |                               |
| MODEL OLUŞTURMA AŞAMASI           | Optimizasyon Öncesi                                                           | Optimizasyon Sonrası          | Optimizasyon Öncesi                                                           | Optimizasyon Sonrası          | Optimizasyon Öncesi                                                           | Optimizasyon Sonrası          | Optimizasyon Öncesi                                                           | Optimizasyon Sonrası          |
| EN UYGUN PARAMETRE                | C parametresi: 1<br>γ : 0.1                                                   | C parametresi: 50<br>γ : 0.01 | C parametresi: 1<br>γ : 0.1                                                   | C parametresi: 50<br>γ : 0.01 | C parametresi: 1<br>γ : 0.1                                                   | C parametresi: 50<br>γ : 0.01 | C parametresi: 1<br>γ : 0.1                                                   | C parametresi: 50<br>γ : 0.01 |
| EĞİTİM HATALARI                   | R <sup>2</sup> = 0.819                                                        | R <sup>2</sup> = 0.884        | R <sup>2</sup> = 0.796                                                        | R <sup>2</sup> = 0.881        | R <sup>2</sup> = 0.791                                                        | R <sup>2</sup> = 0.847        | R <sup>2</sup> = 0.799                                                        | R <sup>2</sup> = 0.866        |
|                                   | MAE= 15.2076                                                                  | MAE= 11.3743                  | MAE= 15.1121                                                                  | MAE= 11.3080                  | MAE= 15.3131                                                                  | MAE= 12.3084                  | MAE= 15.2631                                                                  | MAE= 11.8592                  |
|                                   | MSE= 436.5684                                                                 | MSE= 280.0868                 | MSE= 461.4934                                                                 | MSE= 270.8016                 | MSE= 475.1043                                                                 | MSE= 348.3370                 | MSE= 473.1043                                                                 | MSE= 293.9615                 |
|                                   | RMSE= 20.8942                                                                 | RMSE= 16.7358                 | RMSE= 21.4824                                                                 | RMSE= 16.4560                 | RMSE= 21.7969                                                                 | RMSE= 18.6638                 | RMSE= 21.3969                                                                 | RMSE= 17.1453                 |
| TEST HATALARI                     | R <sup>2</sup> = 0.710                                                        | R <sup>2</sup> = 0.789        | R <sup>2</sup> = 0.759                                                        | R <sup>2</sup> = 0.830        | R <sup>2</sup> = 0.725                                                        | R <sup>2</sup> = 0.770        | R <sup>2</sup> = 0.745                                                        | R <sup>2</sup> = 0.801        |
|                                   | MAE= 19.2346                                                                  | MAE= 16.2627                  | MAE= 18.3339                                                                  | MAE= 16.0699                  | MAE= 18.4811                                                                  | MAE= 17.3759                  | MAE= 17.8861                                                                  | MAE= 15.4049                  |
|                                   | MSE= 704.8620                                                                 | MSE= 509.4654                 | MSE= 592.2009                                                                 | MSE= 418.2789                 | MSE= 619.7039                                                                 | MSE= 518.4585                 | MSE= 605.3839                                                                 | MSE= 405.4086                 |
|                                   | RMSE= 26.5492                                                                 | RMSE= 21.6873                 | RMSE= 24.3352                                                                 | RMSE= 20.4517                 | RMSE= 24.8938                                                                 | RMSE= 22.7697                 | RMSE= 23.7831                                                                 | RMSE= 20.1348                 |

**EK B.** Farklı kestirim modellerinden elde edilen gecikme değerlerinin grafikleri

