



Sentinel-5P Uydu Verileri ve Makine Öğrenmesi Yöntemiyle İzmir Atmosferinde Saatlik NO₂ Konsantrasyonlarının Tahmini

Estimation of Hourly NO₂ Concentrations in İzmir's Atmosphere Using Sentinel-5P Satellite Data and a Machine Learning Approach

Efem Bilgiç^{1*} , Tolga Elbir^{1,2} 

¹Dokuz Eylül Üniversitesi, Mühendislik Fakültesi, Çevre Mühendisliği Bölümü, İzmir, Türkiye

²Dokuz Eylül Üniversitesi Çevre Araştırma ve Uygulama Merkezi (ÇEVMER), İzmir, Türkiye

Öz

Bu çalışmada, hava kalitesi tahmininde uzaktan algılama verilerinin ve makine öğrenmesi yöntemlerinin sunduğu potansiyelden yararlanılarak 2022-2023 yılları arasında İzmir'deki 10 hava kalitesi izleme istasyonundan elde edilen saatlik NO₂ konsantrasyonlarının tahmini amaçlanmıştır. Bu kapsamda, Sentinel-5P uydu verileri, yeniden analiz ürünleri ve sayısal yükselti modeli parametreleri kullanılarak rastgele orman algoritması uygulanmıştır. Model performansı, 10 katmanlı çapraz doğrulama yöntemiyle değerlendirilmiş; R², uyuma indeksi (Uİ), kök ortalama kare hatası (KOKH) ve ortalama mutlak hata (OMH) metrikleri sırasıyla 0.71, 0.91, 14.52 ve 8.88 olarak hesaplanmıştır. Bu doğruluk değerleri, ulusal ölçekte hava kalitesi modelleme çalışmalarında güçlü bir referans sunmaktadır. Analiz sonuçları, model performansının mevsimsel ve mekânsal farklılıklardan etkilendiğini ortaya koymuştur. Özellikle uç değerlerin varlığı ve mevsimsel koşulların Sentinel-5P verilerinin doğruluğu üzerindeki etkisi, performans değişikliklerinde önemli bir rol oynamaktadır. Özellik önemi analizleri, model tahminlerinde en etkili parametrelerin vejetasyon, Sentinel-5P verileri, albedo ve yükselti olduğunu göstermiştir. Bu çalışma, mevcut ölçüm ağlarının kısıtlı olduğu bölgelerde NO₂ konsantrasyonlarının güvenilir şekilde tahmin edilmesine yönelik yenilikçi bir yaklaşım sunarak, hava kirliliği ile mücadelede karar alıcılar için önemli bir katkı sağlamaktadır. Gelecekte Türkiye genelinde gerçekleştirilecek geniş kapsamlı modelleme çalışmaları ile diğer hava kirlleticilerinin tahminine yönelik uygulamaların, hava kirliliği analitiği ve çevresel politika geliştirme süreçlerine önemli katkılar sağlaması öngörülmektedir.

Anahtar Kelimeler: Hava kalitesi tahmini, İzmir, Makine öğrenmesi, NO₂, Sentinel-5P.

Abstract

This study, leveraging the potential of remote sensing data and machine learning methods in air quality prediction, aimed to estimate hourly NO₂ concentrations obtained from 10 air quality monitoring stations in İzmir between 2022 and 2023. In this context, the random forest algorithm was applied using Sentinel-5P satellite data, reanalysis products, and digital elevation model parameters. The model's performance was evaluated using 10-fold cross-validation, with the metrics R², index of agreement (IA), root mean square error (RMSE), and mean absolute error (MAE) calculated as 0.71, 0.91, 14.52, and 8.88, respectively. These accuracy values provide a firm reference for national air quality modeling studies. The analysis revealed seasonal and spatial variations influenced the model's performance. Specifically, the presence of outliers and the impact of seasonal conditions on the accuracy of Sentinel-5P data played a significant role in the performance variations. Feature importance analyses indicated that the most influential parameters in the model's predictions were vegetation, Sentinel-5P data, albedo, and elevation. This study offers a comprehensive approach for reliably estimating NO₂ concentrations in areas where measurement networks are limited and serves as a valuable resource for decision-makers in combating air pollution. It is foreseen that future large-scale modeling studies across Türkiye for other air pollutants will contribute significantly to air pollution analytics and environmental policy development processes.

Keywords: Air quality prediction, İzmir, Machine learning, NO₂, Sentinel-5P.

*Sorumlu yazarın e-posta adresi: efem.bilgic@deu.edu.tr

Efem Bilgiç  orcid.org/0000-0002-6855-4750

Tolga Elbir  orcid.org/0000-0001-6760-3955



1. Giriş

Yanma reaksiyonları sonucu oluşan ve antropojenik faaliyetlerin bir göstergesi olarak kabul edilen azot dioksit (NO₂) atmosferdeki nitrata, 2.5 mikrometreden küçük çaplı partikül maddelere (PM_{2.5}) ve ozona katkı sağlayan önemli bir hava kirleticidir (Fu vd. 2023, Hanif vd. 2021, Kang vd. 2021, Kharol vd. 2017, Qin vd. 2020, Zhang vd. 2021). Yüksek dozda veya sürekli maruz kalınması halinde NO₂ kardiyovasküler ve solunum sistemlerine önemli zararlar verebilmektedir (Castro vd. 2017, Fu vd. 2023, Ghahremanloo vd. 2021, Liu vd. 2015, Weinmayr vd. 2010, Zhu vd. 2019). Bu nedenle, NO₂'nin insan yaşamının yoğun olduğu yerleşim alanlarında sürekli izlenmesi büyük önem taşımaktadır. Günümüzde bu izleme işlemleri genellikle ülkelerin hava kalitesi izleme ağlarına bağlı sabit istasyonlarda yapılan ölçümlerle gerçekleştirilmektedir. Bu istasyonların konumları büyük oranda kent merkezlerinde yoğunlaşmakta olup, yüksek kurulum ve işletme maliyetleri nedeniyle kırsal bölgelerde istasyonlar sayıca azalmakta ya da hiç bulunmamaktadır. Bu nedenle, kırsal alanlarda NO₂ ölçümleri kentsel alanlara göre daha sınırlı kalmaktadır.

Son yıllarda teknolojinin gelişmesiyle birlikte sabit istasyonlarda yer seviyesinde yapılan ölçümler, uydular üzerine yerleştirilen çeşitli enstrümanlar tarafından yapılan atmosferik gözlemlerle desteklenmeye başlamıştır. Örneğin; ozon (O₃), NO₂, formaldehit (HCHO), kükürt dioksit (SO₂), metan (CH₄) ve karbon monoksit (CO) gibi bazı gazların uydulardan izlenmesi amacıyla Amerika Birleşik Devletleri Ulusal Havacılık ve Uzay Dairesi (NASA)'nın Aura uydusuna yerleştirilmiş ve atmosferdeki ozon tabakasını izlemek için geliştirilen OMI (Ozone Monitoring Instrument) enstrümanı (Levelt vd. 2006) ile Avrupa Uzay Ajansı (ESA)'nın Sentinel 5 Precursor (Sentinel-5P) uydusuna yerleştirilen bir troposferik izleme aracı olan TROPOMI (TROPOspheric Monitoring Instrument) enstrümanı (Veefkind vd. 2012) bunların başında gelmektedir. Yüksek mekânsal ve zamansal çözünürlüğü, geniş kapsama alanı ve çok sayıda gazı izleyebilme kapasitesi ile TROPOMI, hizmet vermeye başladığı 2017 yılından itibaren atmosferik izleme ve hava kirliliği çalışmaları için diğer enstrümanlara kıyasla daha yaygın kullanılır hale gelmiştir (Cooper vd. 2022, Goldberg vd. 2020, Griffin vd. 2019, Hashim vd. 2021, Ialongo vd. 2020, Zhang vd. 2020). Pasif bir görüntüleme spektrometresi olan TROPOMI, ultraviyole (UV), görünür (VIS), yakın kızılötesi (NIR) ve kısa dalga kızılötesi (SWIR) bantlarında ölçümler yapmakta olup 3.5 km × 5.5 km'lik yüksek mekânsal çözünürlüğe ve 2600 km genişliğinde tarama alanına sahiptir (Veefkind vd. 2012).

Troposfer boyunca gazların dikey dağılımı heterojen olduğu için bazı gazlar yer seviyesinde yoğunlaşırken bazıları atmosferin üst katmanlarında yoğun olarak bulunur. TROPOMI gibi enstrümanlar ise gazların troposferik kolon yoğunluğunu ölçer, diğer bir ifade ile, bir gazın yer yüzeyinden uydunun görüş hattı boyunca tüm troposferdeki toplam miktarını belirler. Bu nedenle, uydu tabanlı ölçümler doğrudan yer seviyesindeki konsantrasyonları yansıtmazlar (Shetty vd. 2024). Ancak, son yıllarda uydu verileri yardımıyla gazların yer seviyesi konsantrasyonlarını tahmin etmek mümkün olabilmektedir. Özellikle makine öğrenmesi algoritmalarının çeşitlenmesi ve yeteneklerinin artmasıyla, bu tahmin başarısı oldukça artmıştır (Dokuz vd. 2020, Grzybowski vd. 2023, Gündoğdu ve Elbir 2024a, Gündoğdu ve Elbir 2024b, Li vd. 2022, Ünalı ve Yalçın 2022).

Literatürdeki tahmin çalışmalarının bir kısmı uydu verilerini kullanarak çoğunlukla partikül madde tahmini üzerine (Gui vd. 2020, Gündoğdu vd. 2022, Wei vd. 2022, Yao vd. 2018) odaklanmış olup gaz kirleticiler için bu tür çalışmalar sınırlı sayıda kalmaktadır (Chen vd. 2019, Ghahremanloo vd. 2021, Hu vd. 2024, Li vd. 2022, Long vd. 2022, Zhao vd. 2023). Chen vd. (2019) uydu tabanlı olarak geliştirdikleri Aşırı Gradyan Takviyeli Model (XGBoost) modeliyle NO₂ tahmini yapmış ve kriging kalibrasyonu kullanarak R² değerini 0.81'den 0.85'e yükseltmiştir. Ghahremanloo vd. (2021) uydu gözlemleri ve diğer yardımcı verilerden faydalanarak geliştirdikleri Derin Konvolüsyonel Sinir Ağı (Derin-KSA) modeliyle R²'yi 0.83 olarak hesaplamıştır. Li vd. (2022) uydu gözlemleri ve ek yardımcı verilerle geliştirdikleri rastgele orman (RO) modeliyle yer seviyesindeki NO₂ konsantrasyonunu R²=0.78 korelasyonuyla tahmin etmiştir. Long vd. (2022) dört farklı algoritma ile NO₂ tahmin modelleri geliştirmiş ve en yüksek korelasyonu Ekstra Ağaçlar (EA) ve RO modellerinde elde etmiştir. Zhao vd. (2023) XGBoost kullanarak geliştirdikleri NO₂ konsantrasyonu tahmin modelinde iki farklı uydu veri setini karşılaştırmış ve TROPOMI'nin OMI'ye kıyasla daha yüksek korelasyon sağladığını göstermiştir. Hu vd. (2024) ise uydulardan elde edilen farklı seviyedeki verilerin NO₂ tahmini üzerindeki etkilerini incelemek için bir RO modeli geliştirmiştir. Ayrıca, modele alan bilgisine dayalı ek değişkenler (örneğin, güneş zenit açısı, aerosol optik derinliği, yüzey albedosu, yüzey basıncı) dahil edildiğinde korelasyonun arttığı tespit edilmiştir.

Türkiye'de uydu verilerine dayalı çalışmalar genellikle hava kalitesindeki değişimlerin incelenmesiyle sınırlı kalmıştır (Sümsüli ve Kalkan 2022, Ünal Çilek 2022, Yavaşlı ve Ölgen 2022). Yer seviyesinde ise yalnızca toz kirleticisi konsantras-

yonlarının tahminine yönelik çalışmalar mevcuttur (Gündoğdu vd. 2022, Tuna Tuygun vd. 2021, Tuna Tuygun ve Elbir 2023). Türkiye’de uydu verilerinden gaz kirleticilerin tahminine yönelik bir çalışma bulunmamakla birlikte, bazı uluslararası araştırmalarda Türkiye’deki NO₂ ölçümlerinin girdi olarak kullanıldığı görülmektedir (Shetty vd. 2024).

Meteorolojik parametreler, kirleticilerin atmosferdeki dağılımını, taşınımını ve yer seviyesindeki konsantrasyonlarını doğrudan etkilediği için hava kirliliği tahmininde önemli rol oynar. Özellikle rüzgâr, sıcaklık, basınç, nem ve yağış gibi parametreler, hava kalitesinin zamansal ve mekânsal değişiminde belirleyici olmaktadır. Bu nedenle, hava kirliliği tahmin çalışmalarında meteorolojik verilerin kullanılması zorunludur. Ancak, meteoroloji istasyonlarından elde edilen veriler, istasyonların yakın çevresini temsil etmekte olup istasyonların bulunmadığı dağlık veya kırsal bölgelerde yetersiz kalmaktadır. Bu nedenle, son yıllarda küresel ölçekte daha homojen ve sürekli bir veri akışı sağlayan yeniden analiz (reanalysis) verileri sıklıkla tercih edilmektedir (Chi vd. 2022, Deng vd. 2024, Gündoğdu vd. 2022, Tuna Tuygun ve Elbir 2023, Zhao vd. 2023). Kullanılan yeniden analiz veri setine göre mekânsal çözünürlüğü değişen (örneğin ERA5 için 31 km, JRA-55 için 55 km; ERA-Interim için 80 km; NCEP-NCAR Reanalysis-1 için 250 km) bu veriler, hava kalitesi modelleme ve tahmin çalışmaları için önemli bir veri kaynağını oluşturmaktadır. Avrupa Orta Vadeli Hava Tahmin Merkezi (ECMWF) tarafından üretilen bir yeniden analiz veri seti olan ERA5, yaklaşık 31 km (0.25°) çözünürlükle atmosfer, deniz ve kara yüzeyini kapsayan küresel bir veri seti olup atmosferin dikey yapısını 137 seviyede temsil eder (Hersbach vd. 2023). Bu veri setinden türetilen ERA5-Land veri seti ise, ERA5’in kara yüzeyine odaklanan daha yüksek çözünürlüklü (0.1° × 0.1°, yaklaşık 11 km) bir alt veri seti olarak sunulmuş olup özellikle kara yüzeyi süreçlerine (toprak nemi, yüzey sıcaklıkları, kar örtüsü, vb.) detaylı veri sağlamak amacıyla geliştirilmiştir (Muñoz Sabater 2019). Makine öğrenmesi yöntemlerini kullanarak hava kalitesi tahmin etme çalışmalarında yüksek çözünürlüklü ERA5-Land verisinin kullanımına dair sınırlı sayıda uluslararası çalışma (de la Cruz Libardi vd. 2024, Grzybowski vd. 2023, Ibrahim vd. 2022, Schneider vd. 2020) bulunmakta olup ülkemizde henüz böyle bir çalışma gerçekleştirilmemiştir.

Bu eksikliği gidermek amacıyla, Türkiye’de ilk kez uydu verisi destekli ve makine öğrenmesi tabanlı bir NO₂ tahmin modeli geliştirilmiştir. Geliştirilen bu model ile, mevcut ölçüm ağlarının kısıtlı olduğu bölgelerde NO₂ konsantrasyonlarının güvenilir şekilde tahmin edilmesi sağlanmıştır.

Bu çalışma kapsamında, İzmir ilinde hava kalitesi izleme ağına ait 10 istasyonda saatlik NO₂ konsantrasyonları tahmin edilmiştir. Bu amaçla, yüksek çözünürlüklü ERA5-Land meteoroloji verileri ve Sentinel-5P TROPOMI’den elde edilen NO₂ kolon verileri kullanılarak RO algoritması ile bir tahmin modeli geliştirilmiştir. Bu model ile yapılan tahminler, istasyonlarda yapılan eş zamanlı ölçümlerle karşılaştırılarak modelin performansı istatistiksel olarak ortaya konmuştur. Bu çalışma, bu kapsamı ile ulusal ölçekte ilk örnek olmasının yanı sıra, yüksek çözünürlüklü ERA5-Land verilerinin kullanımıyla uluslararası literatüre de önemli katkılar sağlamaktadır.

2. Gereç ve Yöntem

Bu çalışmada, yer seviyesindeki atmosferik NO₂ konsantrasyonlarını etkileyebilecek farklı veri setleri kullanılarak bir makine öğrenmesi modeli geliştirilmiştir. Çevresel, meteorolojik ve sosyo-ekonomik özellikleri bu tür bir model geliştirmek için uygun bir altyapı sunduğu için çalışma alanı olarak İzmir seçilmiştir. Akdeniz iklimine sahip olan İzmir’de yazlar oldukça sıcak, kurak ve güneşli; kışlar ise ılık ve orta derecede yağışlı geçmektedir. Ayrıca, Ege Bölgesi’nin denize dik uzanan dağlarının etkisiyle oluşan hava akımları, şehirde farklı meteorolojik dinamikler yaratmaktadır. İzmir, Türkiye’nin en kalabalık üçüncü şehri olup, yoğun nüfus, sanayi ve trafik kaynaklı emisyonlar nedeniyle NO₂ kirliliği açısından önemli bir örnek teşkil etmektedir (Elbir vd. 2011, Elbir ve Muezzinoglu 2004). Bu çeşitlilik, NO₂ kaynaklarının yayılımını incelemek ve halk sağlığı üzerindeki etkileri geniş bir nüfus üzerinde değerlendirmek için uygun bir çalışma sahası sunmaktadır. Bu çalışmada kullanılan veri setlerinin özellikleri ve değerlendirme kriterleri ayrıntılı şekilde ele alınmıştır.

2.1. Veri Temini

Makine öğrenmesi modellerinde, bağımlı ve bağımsız değişkenlerin seçimi büyük önem taşımaktadır. Bu çalışmada, bağımsız değişkenlerin küresel düzeyde zamansal ve mekânsal erişilebilirliği olan veri setlerinden seçilmesine öncelik verilmiştir. Bu doğrultuda, uluslararası literatürde yaygın olarak kullanılan veri setleri arasından, tahmin edilmek istenen bağımlı değişkenle en güçlü ilişkiye sahip parametreler belirlenmiş olup bunlar aşağıda detaylı olarak özetlenmiştir.

2.1.1. Hava Kalitesi Verisi

Bu çalışmada, İzmir il sınırları içinde bulunan hava kalitesi izleme istasyonlarından temin edilen NO₂ konsantrasyonları, tahmin modelinin geliştirilmesinde referans veri kay-

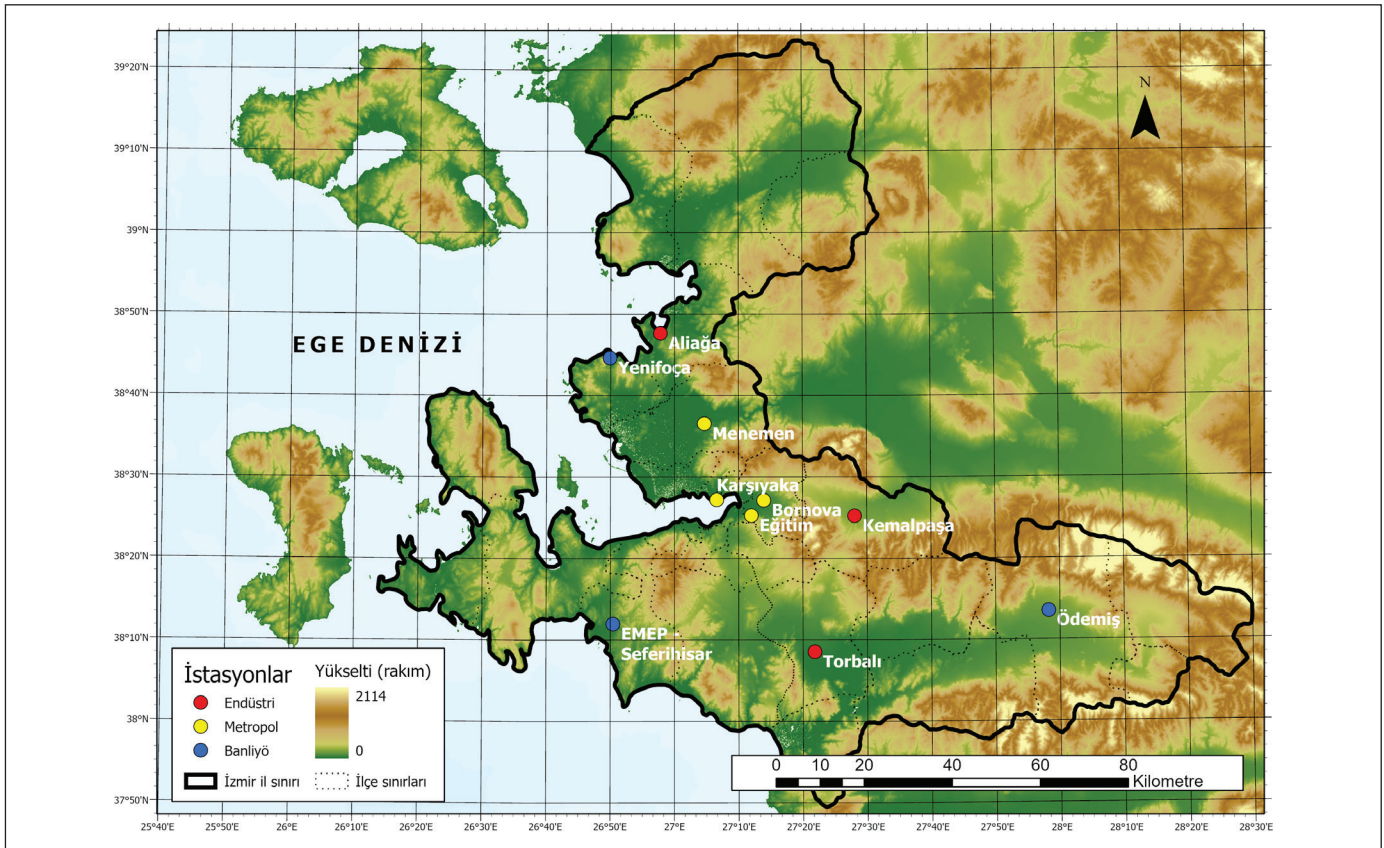
nağı olarak kullanılmıştır. Veriler T.C. Çevre, Şehircilik ve İklim Değişikliği Bakanlığı'nın ulusal hava kalitesi izleme ağından temin edilmiştir (<https://sim.csb.gov.tr/Services/AirQuality>). Çalışmada İzmir il sınırları içindeki tüm istasyonlar taranmış ve Sentinel-5P TROPOMI'nin NO₂ verisi üretmeye başladığı 28 Haziran 2018 tarihinden itibaren NO₂ verisi olan istasyonlara odaklanılmıştır. Tespit edilen 18 istasyondan elde edilen saatlik NO₂ konsantrasyonlarının tanımlayıcı istatistikleri belirlenmiş olup yapılan değerlendirmeler sonucunda bazı istasyonlar (Alsancak İBB, Çiğli İBB, Çeşme, Güzelyalı İBB, Karabağlar, Karaburun ve Konak) yeterli sayıda ve kalitede veriye sahip olmaması nedeniyle çalışmaya dahil edilememiştir. Sonuç olarak, çalışma kapsamında toplam 10 istasyondan kullanılabilir veri elde edilmiştir. Bu istasyonların 3'ü nüfus yoğunluğunun fazla olduğu kent merkezinde (Bornova, Eğitim ve Karşıyaka) iken geriye kalan 7 istasyon ise Aliğa, Foça, Kemalpaşa, Menemen, Ödemiş, Seferihisar ve Torbalı gibi kent merkezine uzak ilçelerde bulunmaktadır. Bu istasyonların konumları Şekil 1'de görülmektedir. Bu 10 istasyon için veri sürekliliği ve miktarı dikkate alındığında tüm istasyonlarda eş zamanlı ölçümlerin ağırlıklı olarak 2022 ve 2023 yıllarında yoğun-

laştığı görülmüştür. Bu nedenle bu çalışma kapsamında çalışma dönemi olarak 1 Ocak 2022 ile 31 Aralık 2023 arası seçilmiştir.

2.1.2. Troposferik Dikey NO₂ Kolon Verisi

Çalışma kapsamında, Sentinel-5P TROPOMI enstrümanı ile elde edilen troposferik dikey NO₂ kolon verisi tahmin modeline girdi olarak kullanılmıştır. Sentinel-5P, ESA tarafından 13 Ekim 2017'de hava kirliliğini izlemek amacıyla fırlatılmıştır. Sentinel-5P uydusu üzerindeki TROPOMI enstrümanı, UV ile SWIR arasındaki bantlarda, atmosferin üst kısmında (TOA) yeryüzünden yansıyan ve yayılan güneş radyasyonunu pasif uzak algılama teknikleriyle ölçen nadir bakışlı bir görüntüleme spektrometresidir. Görevine başladığı dönemde 3.5 × 7.0 km olan yatay çözünürlüğü, 6 Ağustos 2019'dan itibaren 3.5 × 5.5 km'ye geliştirilmiştir. Bu çözünürlük, Sentinel-5P TROPOMI verisinin atmosferik bileşenlerinin izlenmesinde diğer mevcut uydulara göre daha yüksek bir hassasiyet sağlamasını mümkün kılmaktadır (Copernicus 2024).

Bu çalışmada, Sentinel-5P TROPOMI NO₂ verisi Google Earth Engine (GEE) platformu üzerinden temin edilmiştir.



Şekil 1. İzmir'de NO₂ konsantrasyonu ölçen hava kalitesi istasyonlarının konumları.

GEE, Sentinel-5P uydusundan elde edilen troposferik dikey NO₂ kolon verilerini, orijinal Seviye-2 (L2) düzeyinden Seviye-3 (L3) düzeyine dönüştürmektedir. L2 verisi zaman dilimlerine göre organize edilmişken, GEE'de işlenebilmesi için harpconvert aracıyla mekânsal bir gride oturtulmuş ve yaklaşık 1113.2 metre çözünürlükle L3 düzeyine dönüştürülmüştür. Dönüşüm sürecinde, güvenilirliği artırmak amacıyla, troposferik dikey NO₂ kolon bandında kalite güvence (QA) değeri %75'in altında olan pikseller filtrelenmiştir. Bu süreç, L2 verisini görselleştirme ve analiz açısından daha erişilebilir hale getirerek uydu verisinin etkili kullanımını sağlamıştır (Google Earth Engine 2024).

2.1.3. Yeniden Analiz Verisi

Çalışmada ihtiyaç duyulan meteorolojik parametreler, ECMWF Copernicus programı kapsamında sağlanan 5. nesil yeniden analiz ürünlerinin yer aldığı ERA5-Land veri setinden elde edilmiştir (<https://cds.climate.copernicus.eu/>). Bu veri setinden; sıcaklık (t2m), işba/çiğ noktası sıcaklığı (d2m), rüzgârın u (u10) ve v (v10) bileşenleri, yüzey basıncı (sp), toplam yağış (tp), toplam buharlaşma (te), tahmini albedo (fal), yaprak alanı indeksi-yüksek vejetasyon (lai_hv) ve yaprak alanı indeksi-düşük vejetasyon (lai_lv) parametreleri kullanılmıştır. Rüzgâr hızı ve rüzgâr yönü (ws10 ve wd10) u10 ve v10 bileşenlerinden hesaplanarak modele girilmiştir. Diğer yandan, ERA5-Land kara dışında (su kütlelerinde) ve atmosferin üst tabakalarında veri sağlamadığı için bulutluluk (tcc) ve sınır tabakası yüksekliği (blh) parametreleri ERA5-Single Level veri setinden elde edilmiştir.

Çalışmada ERA5 veri setinden temin edilen tüm bu parametrelerin hava kirlleticilerin atmosferdeki yayılımı ve taşınımı mekanizmalarına doğrudan ya da dolaylı katkı sağlamasına dikkat edilmiştir. Öyle ki; rüzgârın düşük hızlarında kirleticiler atmosferde birikirken, yüksek rüzgâr hızlarında kirleticiler daha geniş alanlara yayılırlar. Rüzgâr yönü ise doğrudan kirleticilerin taşındığı bölgeleri belirler. Bununla birlikte, yüksek basınç sistemleri durgun hava koşulları yaratarak kirleticilerin birikmesine yol açarken, düşük basınç sistemleri atmosferde daha fazla karışım ve seyrelme etkisi sağlar. Diğer bir parametre olan sıcaklığın yüksek olması durumunda O₃ gibi kirleticilerin fotokimyasal reaksiyonları hızlanarak bölgesel hava kirliliği sorunları görülebilir. Ayrıca atmosferde düşey sıcaklığın farklılaşması (sınır tabaka yüksekliğinin azalması) inversiyon oluşumuna ve kirleticilerin yüzeye yakın tabakalarda birikmesine neden olur. Bulutluluk da inversiyon süreçlerini etkileyerek atmosferdeki fotokimyasal süreçleri doğrudan etkiler ve bu nedenle hava kirliliği üzerinde önemli bir rol oynar. Yağış, kirleticilerin yıkanarak

yüzeye indirilmesi yoluyla atmosferden uzaklaştırılmasında büyük rol oynar. Benzer şekilde, çiğ noktası sıcaklığı atmosferdeki nem oranını gösteren önemli bir parametre olup aerosollerin ve gazların atmosferdeki davranışını etkiler. Yüksek buharlaşma oranlarının ise kirleticilerin yayılımını ve atmosferdeki kimyasal reaksiyonların hızını arttırdığı bilinmektedir. Albedo, bir yüzeyin güneş ışığını yansıtma kapasitesini ifade etmekte olup yüzey sıcaklığı ile hava kirliliği arasındaki ilişkiyi açıklayan önemli bir faktördür. Yaprak alanı indeksi, yüksek (ağaçlar ve çalılar gibi uzun boylu ve odunsu bitki örtüsü) ve düşük (otlar, çalılar ve çimenler gibi kısa ve otsu bitkiler) vejetasyon olarak iki farklı grupta vejetasyon yoğunluğunu yansıtan bir parametredir. Bitki örtüsü, atmosferik kirleticileri tutarak ve kentsel ısı adası etkisini azaltarak hava kalitesini iyileştirme özelliğine sahiptir. Ancak, yoğun vejetasyon düşük rüzgâr hızlarına neden olarak kirleticilerin birikimine de katkıda bulunabilir.

2.1.4. Sayısal Yükselti Modeli (dem)

Meteorolojik verilerin yanı sıra mekânsal farklılığı temsil eden ve hava kalitesi üzerinde önemli etkiye sahip diğer bir etken ise yükseltidir. Çalışmada yükselti verisi, NASA'nın 30 m mekânsal çözünürlüğe sahip Mekik Radar Topografya Misyonu (SRTM- Shuttle Radar Topography Mission) veri setinden kullanılmıştır (<https://www.earthdata.nasa.gov/sensors/srtm>). Bu veri seti ile elde edilen bölgenin topoğrafik haritası Şekil 1'de verilmiştir.

Tahmin modeli için toplam 14 parametre (bağımsız değişken) girdi olarak kullanılmıştır. Çalışmada kullanılan tüm parametreler topluca Çizelge 1'de özetlenmiştir.

2.2. Veri Hazırlama

Çalışmada tüm veriler 1 Ocak 2022 – 31 Aralık 2023 tarihleri arasındaki dönemi kapsayacak şekilde temin edilmiştir. Tüm veriler istasyonların bulunduğu koordinatlar temel alınarak elde edilmiş ve her istasyonun koordinatlarının içine girdiği, diğer bir ifade ile istasyonlara konum olarak en yakın gridteki değerler kullanılmıştır. Bu süreçte herhangi bir mekânsal ölçek indirgeme veya interpolasyon yöntemiyle ara değer hesaplama gibi işlemler uygulanmamıştır. Bu yöntem, farklı çözünürlüklere sahip parametrelerin uyumlu bir şekilde analiz edilmesini sağlamıştır. Zamansal eşleştirmede ise, Sentinel-5P TROPOMI NO₂ verisi temel alınmıştır. Kullanılabilir NO₂ verisinin temin edilebildiği saatler için ERA5 veri setinden de eş zamanlı ilgili parametreler temin edilmiştir. Arasında zamansal dilim farklılığı olan veri setlerinde gerekli zamansal düzenlemeler yapılmıştır. Eşleştirme sonunda, İzmir'deki tüm istasyonlarda saatlik bazda toplam 5320 eşleştirilmiş veri elde edilmiştir.

Çizelge 1. Çalışmada kullanılan veriler

Parametre	Kısaltma	Birim	Çözünürlük		Kaynak
			Mekânsal*	Zamansal	
Yer seviyesi NO ₂ ölçümleri	no2	µg.m ⁻³		Saatlik	Ulusal Hava Kalitesi İzleme Ağı
Troposferik dikey NO ₂ kolonu	no2_uydu	mol.m ⁻²	1113.2 m	Günlük	ESA Sentinel 5P Uydusu
Sıcaklık (2m)	t2m	°C	11132 m	Saatlik	ECMWF ERA5-Land
İşba/çiğ noktası sıcaklığı (2m)	d2m	°C	11132 m	Saatlik	ECMWF ERA5-Land
Rüzgâr hızı (u10 ve v10'dan üretildi)	ws10	m.s ⁻¹	11132 m	Saatlik	ECMWF ERA5-Land
Rüzgâr yönü (u10 ve v10'dan üretildi)	wd10	derece	11132 m	Saatlik	ECMWF ERA5-Land
Yüzey basıncı	sp	hPa	11132 m	Saatlik	ECMWF ERA5-Land
Toplam yağış	tp	mm	11132 m	Saatlik	ECMWF ERA5-Land
Toplam buharlaşma	te	mm su eşdeğeri	11132 m	Saatlik	ECMWF ERA5-Land
Tahmini albedo	fal	birimsiz	11132 m	Saatlik	ECMWF ERA5-Land
Yaprak alanı indeksi - Yüksek vejetasyon	lai_hv	m ² .m ⁻²	11132 m	Saatlik	ECMWF ERA5-Land
Yaprak alanı indeksi - Düşük vejetasyon	lai_lv	m ² .m ⁻²	11132 m	Saatlik	ECMWF ERA5-Land
Sınır tabakası yüksekliği	blh	m	0.25° (~31 km)	Saatlik	ECMWF ERA5-Single Level
Toplam bulutluluk	tcc	birimsiz	0.25° (~31 km)	Saatlik	ECMWF ERA5-Single Level
Yükselti (rakım)	dem	m	30 m		NASA SRTM

*GEE platformundan temin edilen parametreler için, GEE'nin belirttiği çözünürlük değerleridir.

2.3. Makine Öğrenmesi

Uydu verileri kullanılarak makine öğrenmesi ile yer seviyesi NO₂ konsantrasyonu tahmini üzerine yapılan çalışmalarda, öğrenme modeli olarak sıklıkla RO (Ghahremanloo vd. 2021, Grzybowski vd. 2023, Hu vd. 2024, Kang vd. 2021, Li vd. 2022), XGBoost (Chi vd. 2022, Deng vd. 2024, Fu vd. 2023, Kang vd. 2021, Kim vd. 2021, Shetty vd. 2024, Zhao vd. 2023), EA (Long vd. 2022) ve Hafif Gradyan Takviyeli Model (LightGBM) (Kang vd. 2021) gibi ağaç tabanlı yöntemler kullanılmaktadır. Bu çalışmada, saatlik NO₂ konsantrasyonlarını tahmin etmek amacıyla RO algoritması kullanılmıştır. RO, birden fazla karar ağacının oluşturduğu bir topluluk öğrenme yöntemidir. Her bir ağaç, eğitim verisinin farklı alt kümeleri üzerinde eğitilir ve bağımsız olarak üretilen tahminlerin ortalaması alınarak nihai sonuçlar elde edilir. Bu yöntem, değişkenler arasındaki karmaşık ilişkileri öğrenme ve aşırı öğrenmeyi önleme konusunda etkinliği ile bilinmektedir (Breiman 2001).

Bu çalışmada, RO modeli hiper parametrelerinin belirlenmesi için python scikit-learn kütüphanesinin GridSearchCV sınıfından faydalanılmıştır. Karar ağacı sayısı (n_estimators) için 100-2000 aralığı, her bir düğümde kullanılacak özellik sayısı (max_features) için 6-8 aralığı ve bir yaprakta bulunması gereken minimum örnek sayısı (min_samples_leaf) ise 2-3 aralığında tanımlanmıştır. Tespit edilen optimum parametrelerin listesi Çizelge Ek 1'de listelenmiştir. Ayrıca, veri setinde yer alan bağımsız değişkenler farklı birim ve ölçeklere sahip olduğundan, model performansını iyileştirmek ve değişkenlerin eşit katkıda bulunmasını sağlamak amacıyla standartlaştırılmıştır.

Model çalışmasında sınırlayıcı parametre olarak TROPOMI verilerinin kullanılması nedeniyle, tüm analizler Sentinel-5P uydusunun İzmir üzerinden geçtiği 09.00–12.00 (UTC) saat aralığındaki verilere dayanmaktadır.

2.4. Model Parametrelerinin Değerlendirilmesi

Modelin tahmin sonuçlarını yorumlamak ve hangi bağımsız değişkenlerin NO₂ tahmininde daha etkili olduğunu belirlemek amacıyla özellik önemi analizi gerçekleştirilmiştir. Bu analizde, RO modelindeki her bir karar ağacının düğüm ayrımlarında kullanılan özelliklerin, model başarısına katkısı değerlendirilmiştir. Özellik önemi, bir düğümde belirli bir özelliğin kullanımıyla ortaya çıkan Ortalama Mutlak Hata (OMH) azalmasının hesaplanmasıyla belirlenir. Her özellik için bu azalmalar, tüm karar ağaçları boyunca toplanır ve bir özellik için toplam OMH azalması, o özelliğin göreceli önemini ifade eder. Bu değerler, tüm önem skorlarının toplamı 1 olacak şekilde normleştirilir. Daha yüksek önem değerine sahip özellikler, modelin tahminlerinde daha etkili olurken, düşük değerlere sahip özelliklerin etkisi daha sınırlıdır.

Bu çalışmadaki analiz, python scikit-learn kütüphanesinin RandomForestRegressor modelinin sağladığı feature_importances_ özelliği kullanılarak gerçekleştirilmiştir. Analiz sonucunda, NO₂ tahmini için kritik öneme sahip değişkenler belirlenmiş ve bu değişkenlerin hava kirliliği bağlamındaki etkileri değerlendirilmiştir. Özellik önemi analizi, değişkenlerin modele olan etkisini anlamak için önemli bir araç olarak kullanılmıştır.

2.5. Model Performansının Belirlenmesi

Modelin genelleme kapasitesini değerlendirmek ve aşırı öğrenmeyi önlemek amacıyla 10 katmanlı çapraz doğrulama (10-fold cross-validation) yöntemi kullanılmıştır. Bu yöntemde veri seti, her iterasyonda farklı bir bölümü test verisi olarak kullanmak üzere 10 eşit alt parçaya bölünmüştür. Her iterasyonda model, geri kalan 9 parça ile eğitilmekte ve ayrılan 1 parça ile test edilmektedir. Bu süreç, tüm veri setinin hem eğitim hem de test amacıyla kullanılması sayesinde modelin performansının daha güvenilir bir şekilde değerlendirilmesine olanak tanır. Çapraz doğrulama ile elde edilen sonuçlar, modelin genelleme yeteneği üzerinde daha doğru bir fikir sunmaktadır.

Bu çalışmada, modelin performansını değerlendirmek için çeşitli istatistiksel parametreler kullanılmıştır. Tahmin edilen ve gerçek değerler arasındaki ilişkinin boyutunu gösteren Pearson Korelasyon Katsayısı (R) hesabı Eşitlik 1'de gösterilmiştir.

$$R = \frac{\sum_{i=1}^n (y_{ölçüm,i} - \bar{y}_{ölçüm})(y_{tahmin,i} - \bar{y}_{tahmin})}{\sqrt{\sum_{i=1}^n (y_{ölçüm,i} - \bar{y}_{ölçüm})^2 \sum_{i=1}^n (y_{tahmin,i} - \bar{y}_{tahmin})^2}} \quad (1)$$

Burada $y_{ölçüm,i}$ ölçülen i nolu NO₂ konsantrasyonunu, $\bar{y}_{ölçüm}$ ölçülen NO₂ konsantrasyonlarının ortalamasını, $y_{tahmin,i}$ mo-

del tarafından tahmin edilen i nolu NO₂ konsantrasyonunu, $\bar{y}_{tahmin}$ model tarafından tahmin edilen NO₂ konsantrasyonlarının ortalamasını temsil etmektedir. Bu çalışmada modelin tahminlerinin gözlemlerle ne derece uyumlu olduğunu ölçen bir başka metrik olan Uyuşma İndeksi (UI) de hesaplanmıştır (Eşitlik 2).

$$U = \frac{\sum_{i=1}^n (y_{ölçüm,i} - y_{tahmin,i})^2}{\sum_{i=1}^n (|y_{tahmin,i} - \bar{y}_{ölçüm,i}| + |y_{ölçüm,i} - \bar{y}_{tahmin,i}|)^2} \quad (2)$$

Model performansı değerlendirmesinde, modelin ne kadar yanılma payına sahip olduğunu ölçen OMH metriği de kullanılmıştır. OMH, modelin tahmin ettiği değerler ile gerçek değerler arasındaki farkların mutlak değerlerinin ortalamasını alarak hesaplanmaktadır (Eşitlik 3).

$$OMH = \frac{1}{n} \sum_{i=1}^n |y_{ölçüm,i} - y_{tahmin,i}| \quad (3)$$

Model tahminlerinin doğruluğunu ölçen bir diğer performans ölçütü olan Kök Ortalama Kare Hatası (KOKH) ile tahmin hatalarının karesinin ortalamasının karekökü hesaplanır ve bu gösterge büyük hatalara karşı daha duyarlıdır. KOKH hesabı Eşitlik 4'te gösterilmiştir.

$$KOKH = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_{ölçüm,i} - y_{tahmin,i})^2} \quad (4)$$

Bu performans metrikleri, modelin doğruluk ve hata oranlarını değerlendirme olanağı sunarak, modelin NO₂ tahminindeki başarısının kapsamlı bir şekilde analiz edilmesini sağlamaktadır. 10 katmanlı çapraz doğrulama yöntemi kapsamında her bir parça için hesaplanan metriklerin ortalaması alınarak modelin genel performansı daha güvenilir bir şekilde değerlendirilmiştir.

3. Sonuçlar ve Tartışma

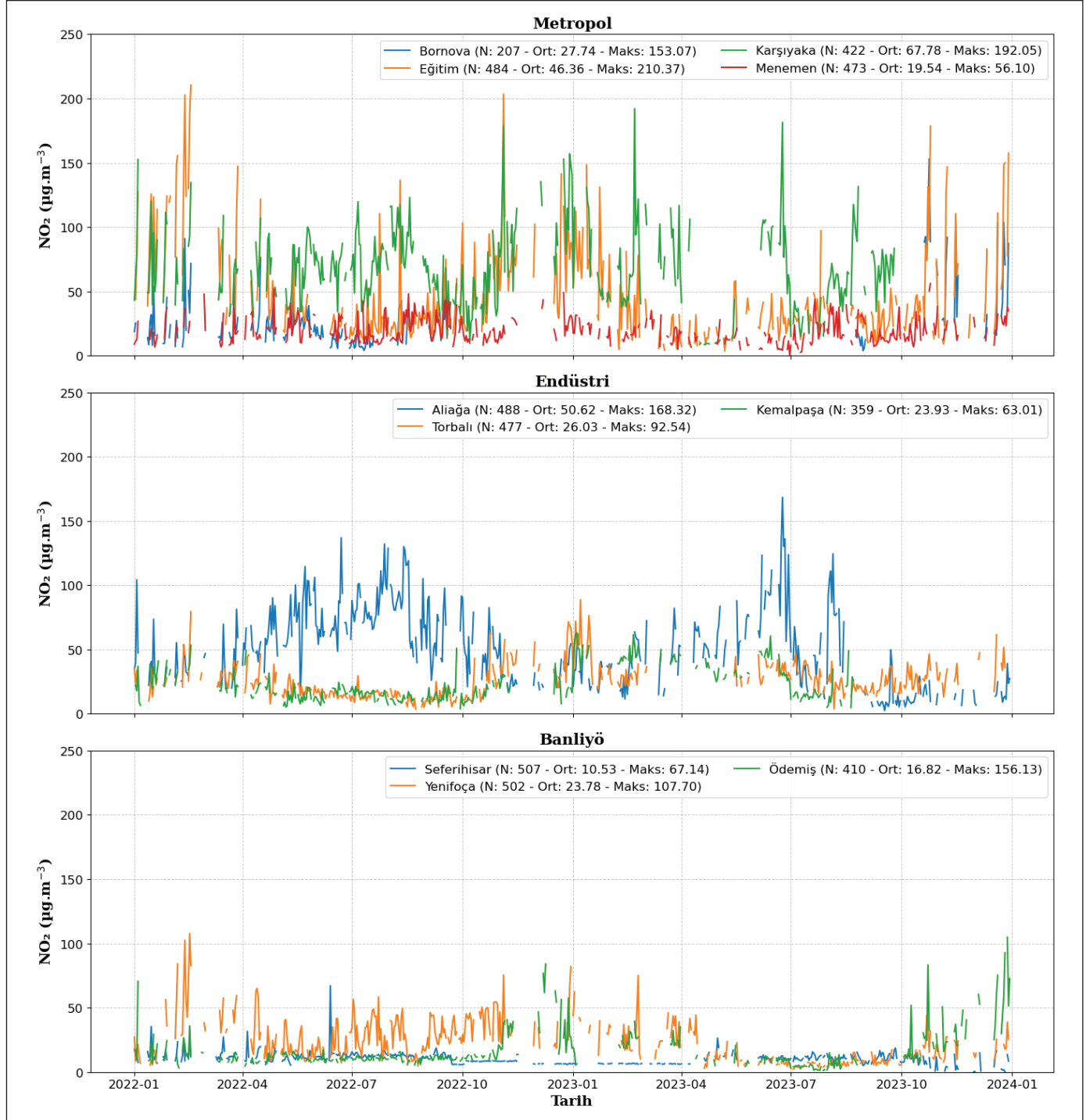
3.1. Tanımlayıcı İstatistikler

NO₂ konsantrasyonları, istasyonların mekânsal özellikleri dikkate alınarak üç farklı grupta incelenmiştir: Metropol, Endüstri ve Banliyö. Metropol grubu, İzmir'in iç körfez çevresinde kentleşmenin ve nüfusun yoğun olduğu merkez ilçeleri temsil etmektedir. Bu gruba ait istasyonlar Bornova, Karşıyaka, Menemen ve Eğitim istasyonlarıdır. Endüstri grubu, sanayinin yoğun olduğu Aliaga, Kemalpaşa ve Torbalı istasyonlarını kapsamaktadır. Banliyö grubu ise kentleşmenin ve sanayinin daha az olduğu kırsal bölgelerde yer alan Yenifoça, Ödemiş ve Seferihisar istasyonlarından oluşmaktadır.

Şekil 2'de verilen zaman serisi grafikleri, Metropol ve Endüstri gruplarındaki istasyonlarda, Banliyö istasyonlarına

kıyasla daha yüksek NO₂ konsantrasyonlarının ölçüldüğünü göstermektedir. Özellikle kentleşmenin ve sanayinin yoğun olduğu Karşıyaka, Eğitim ve Aliğa istasyonlarında yüksek konsantrasyonlar tespit edilmiştir. Bu üç istasyonda ölçülen yıllık ortalama NO₂ konsantrasyonları, Hava Kalitesi De-

ğerlendirme ve Yönetimi Yönetmeliği'nde yer alan yıllık ortalama limit değerin (40 µg/m³) üzerinde çıkmıştır. Eğitim istasyonunda ise saatlik limit değeri (200 µg/m³) sadece 3 kere aşılmış, diğer istasyonlarda kullanılan NO₂ konsantrasyonları bu değerin altında kalmıştır.



Şekil 2. Modelde öğretilen NO₂ değerlerinin istasyon bazlı olarak zamansal değişimleri (Ort: Ortalama, Maks: Maksimum, N: Veri Sayısı).

Mevsimsel değişim incelendiğinde, Aliağa dışındaki tüm istasyonlarda kış aylarında daha yüksek, yaz aylarında ise daha düşük NO₂ konsantrasyonları gözlemlenmiştir. Bu durum, kış mevsiminde doğalgaz ve kömür tüketimindeki artış gibi ısınma kaynaklı emisyonlarla ilişkilendirilmektedir. Ayrıca, modelde kullanılan blh verilerinin mevsimsel değişimi de bu sonuçları desteklemektedir. Kış aylarında düşen blh değerleri, kirleticilerin atmosferde dikey doğrultuda dağılmasını engelleyerek NO₂ konsantrasyonlarının artmasına neden olmaktadır. Yaz aylarında ise blh değerlerindeki artış, daha düşük NO₂ seviyelerine katkı sağlamaktadır.

Modelde kullanılan ve zamansal olarak değişen parametrelerin histogramları Şekil 3'te sunulmuştur. no2 ve no2_uydu verileri, birbirleriyle uyumlu bir şekilde düşük değerlerde yoğunlaşmıştır. Ayrıca, çok düşük konsantrasyon aralıklarında veri sayısının nispeten az olduğu gözlemlenmiştir. Sentinel-5P'nin İzmir üzerinden öğle saatlerinde geçmesi nedeniyle t2m değerleri, günün en yüksek sıcaklıklarını temsil etmekte ve veri setindeki değerler normal dağılıma yakınlık göstermektedir. Daha düşük sıcaklıklara sahip olan d2m, t2m ile benzer bir dağılım ve standart sapma göstermektedir. ws10, ortalama 3.31 m/s olarak hesaplanmış ve düşük değerlerde yoğunlaşmıştır. wd10 ise en düşük ve en yüksek değerlere doğru yoğunlaşarak rüzgarın genellikle kuzeyden estiğini göstermektedir. sp genellikle 1000 hPa civarında yoğunlaşmış, ancak zaman zaman alçak basınç koşulları da gözlemlenmiştir. te, negatif değerlere sahip olduğundan, sağa çarpık bir grafik sunmakla birlikte düşük değerlerde yoğunlaşma göstermektedir. blh, sola çarpık bir dağılım sergilese de oldukça küçük değerlere sahip değildir. Veri setindeki blh değerleri özellikle 400-1200 m aralığında benzer sıklıklarla dağılmıştır. fal, çok dar bir değişim aralığına sahip olup ortalama değer etrafında yoğunlaşmıştır. lai_hv, bulunduğu değer aralığında dengeli olarak dağılırken, lai_lv mekânsal farklılıklar nedeniyle dengesiz bir dağılım sergilemiştir. Çalışmada kullanılan verilerin çoğunlukla bulutsuz ve yağışsız günlere ait olması nedeniyle tcc ve tp değerlerinin büyük bölümü 0 (sıfır) veya yakın değerlerden oluşmuştur. Bu durum, histogramların logaritmik eksende çizilmesini gerektirmiş ve her iki parametrenin de düşük değerlerde yoğunlaştığını ortaya koymuştur.

3.2. Özellik Önemi

RO modeline ait girdilerin model performansına katkıları değerlendirildiğinde, en yüksek önem düzeyine sahip parametreler sırasıyla lai_lv, no2_uydu, dem, lai_hv ve fal olarak tespit edilmiştir. Şekil 4'te tüm parametrelerin önem düzeyleri verilmiştir.

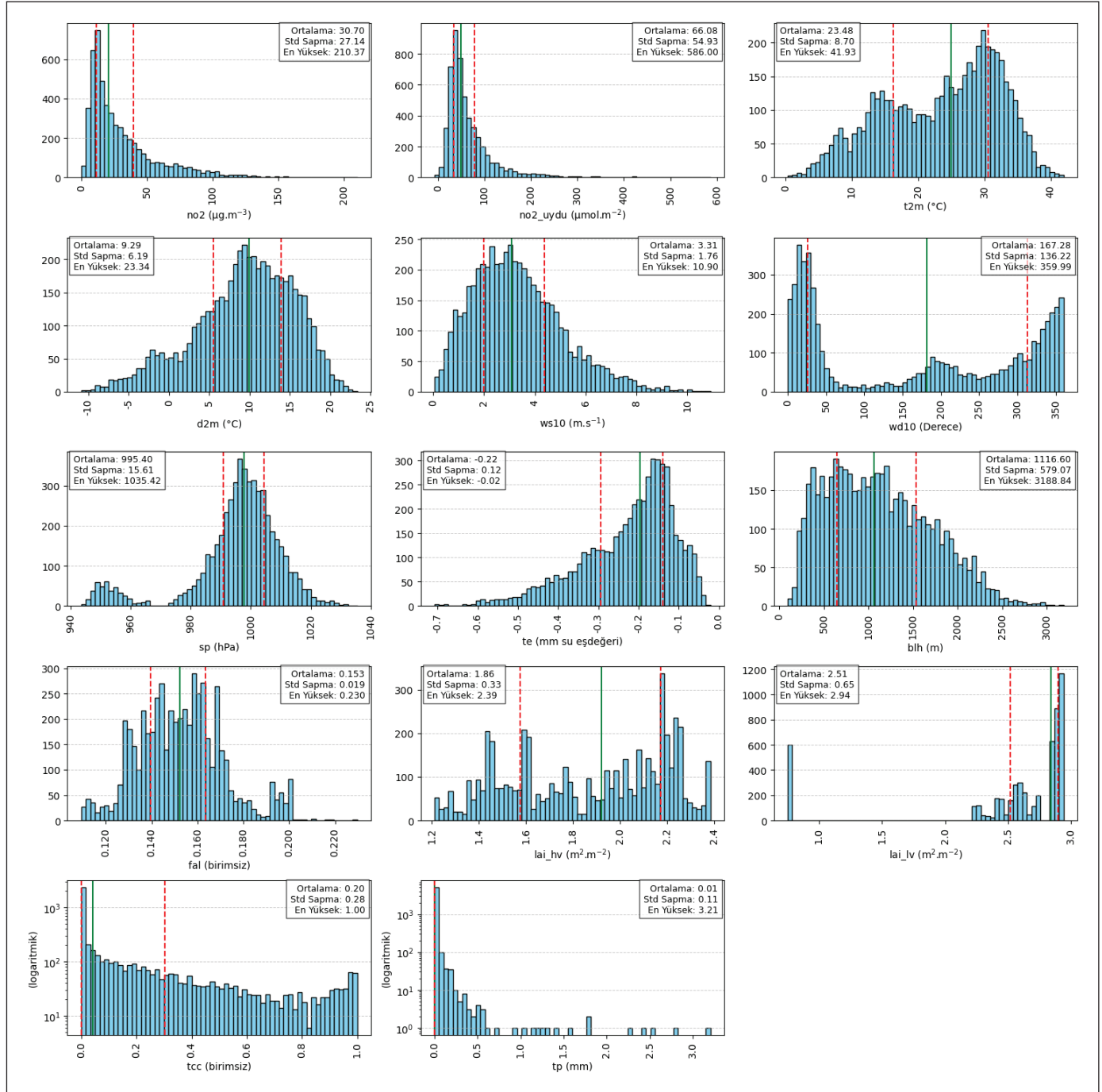
Literatürde, lai_lv ve lai_hv gibi parametreler yaygın olarak kullanılmamakla birlikte, bunlara benzer şekilde vejetasyonu temsil eden NDVI (Normalized Difference Vegetation Index) veya EVI (Enhanced Vegetation Index) gibi çeşitli parametrelerin model sonuçları üzerindeki yüksek etkisine işaret eden çok sayıda çalışma bulunmaktadır (Chen vd. 2019, Shetty vd. 2024). Vejetasyonun yoğunluğu veya eksikliği, modellemede kritik bir rol oynamaktadır. Yüksek veya düşük vejetasyon tiplerinin varlığı, kirleticilerin giderilmesi ve kentsel ısı adası etkisinin azaltılması yoluyla NO₂ konsantrasyonlarının azalmasına katkı sağlamaktadır (Cichowicz ve Bochenek 2024). Buna karşılık, vejetasyonun bulunmadığı alanlar, genellikle yüksek NO₂ emisyonlarının yoğun olduğu kentsel ve endüstriyel bölgeleri işaret etmektedir.

En yüksek etkiye sahip iki parametreden biri olan no2_uydu, troposferdeki NO₂ yoğunluğunu doğrudan ölçmesi nedeniyle emisyon kaynaklarının yoğun olduğu bölgelerde daha yüksek değerlere sahiptir. Literatürde, no2_uydu benzeri çalışmaların tamamında bu parametrenin en etkili girdiler arasında yer aldığı rapor edilmektedir (Ahmad vd. 2024, Ghahremanloo vd. 2021, Kim vd. 2021). Bu durum, no2_uydu parametresinin, antropojenik emisyon yoğunluğunu yansıtan güvenilir bir gösterge olarak önemini vurgulamaktadır.

Topoğrafik bir gösterge olan dem, model performansına sağladığı önemli katkılar nedeniyle NO₂ tahmini yapan çalışmalarda yaygın olarak kullanılan diğer bir parametre olarak öne çıkmaktadır (Becerra-Rondón vd. 2023, Shetty vd. 2024). dem, hem kirleticilerin taşınımı hem de kentleşme dinamikleri (örneğin, engebeli arazilerde kırsallaşma eğilimleri) üzerindeki etkileri nedeniyle model öğrenmesinde öne çıkan değişkenlerden biri olarak belirlenmiştir.

En etkili beş değişkenden dört tanesi (lai_lv, no2_uydu, fal ve lai_hv), hem doğrudan NO₂ emisyon kaynaklarıyla ilişkilidir hem de mevsimsel koşullara bağlı olarak düzenli değişkenlik göstermektedir. Bu özellikleri nedeniyle, bu parametreler modelin öğrenme kapasitesine en yüksek katkıyı sağlamıştır.

Bu temel parametrelerin yanı sıra, t2m, te, ws10, sp, d2m, wd10, ve blh gibi meteorolojik parametrelerin de modele anlamlı katkılar sağladığı tespit edilmiştir. Bu meteorolojik değişkenler, kirleticilerin düşey ve yatay taşınım süreçlerini ve çökelme mekanizmalarını doğrudan etkilemektedir. Dolayısıyla, modelin öğrenmesi sürecinde önemli katkıları olmuştur. Özellikle, t2m, te ve d2m gibi parametrelerin NO₂ ile benzer şekilde mevsimsel koşullardan doğrudan

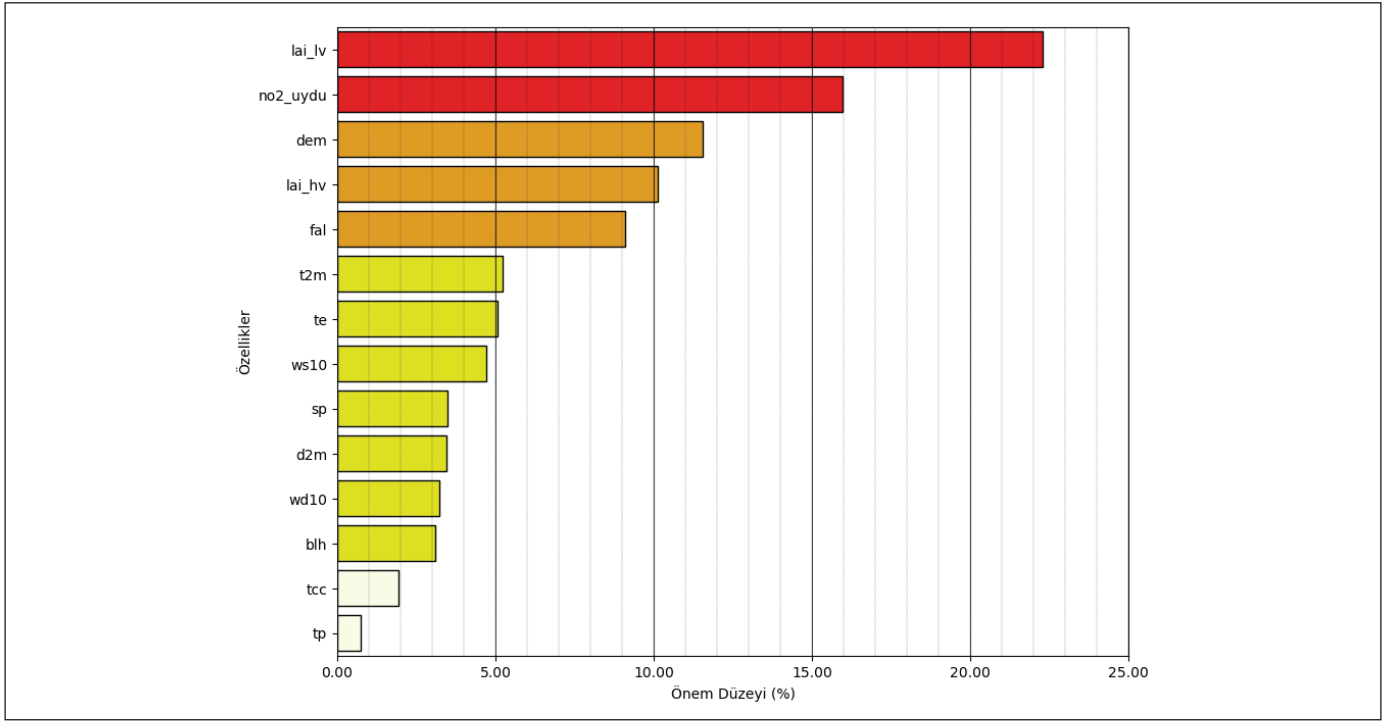


Şekil 3. Modelde kullanılan parametrelerin histogramları. Yeşil çizgi medyanı, kırmızı kesikli çizgiler çeyreklikleri temsil etmektedir.

etkilenmeleri, bu değişkenlerin model üzerindeki önemini artırmaktadır. ws10 ve wd10, NO₂'nin yatay yayılımını belirlerken, sp ve blh ise düşey taşınım süreçlerinde belirleyici bir rol oynamaktadır.

Bu çalışmada, çökeltme yoluyla kirletici giderimi üzerinde önemli bir etkisi olan tp ve inversiyon süreçlerini etkileyen tcc parametrelerinin katkılarının düşük olduğu belirlenmiştir.

tir. Bunun temel nedeni, uyduların bulutlu ve yağışlı günlerde ölçüm yapamaması ya da düşük kaliteli veri sağlamasıdır. Bu durum, verilerin büyük çoğunluğunun yağışsız ve bulutsuz günlerden elde edilmesine yol açmış ve bu iki parametrenin model üzerindeki istatistiksel önemini sınırlamıştır (Deng vd. 2024).



Şekil 4. Tahmin modeli için girdi olarak kullanılan tüm bağımsız değişkenlerin özellik önem düzeyleri.

3.3. Model Performansı

Bu çalışmada, İzmir ili için NO₂ konsantrasyonlarını tahmin etmek amacıyla geliştirilen RO modeli, 10 katlı çapraz doğrulama yöntemiyle test edilmiştir. 2 yıllık periyotta, 10 farklı istasyon için temin edilen verilerle geliştirilen model ile R² değeri 0.72 olarak hesaplanmıştır. Uİ'nin 0.91 ve OMH ile KOKH değerlerinin de sırasıyla 8.87 ve 14.42 olarak hesaplanması model tahminlerinin başarısını desteklemektedir. Bu sonuçlar modelin il genelinde NO₂ tahminlerinde başarılı olduğunu göstermektedir (Şekil 5). İzmir'de elde edilen bu tahmin başarısı literatürdeki birçok benzer çalışmada elde edilenlere yakın olup, İzmir'de tahmin edilen NO₂ değerlerinin ölçülen konsantrasyonlarla büyük bir uyuma sahip olduğunu göstermiştir.

Çizelge 2'de literatürde çeşitli makine öğrenmesi yöntemleriyle NO₂ tahmin eden çalışmalar ve bu çalışmalarda geliştirilen modellerin başarıları listelenmiştir. Benzer çalışmalarda R² değerlerinin 0.30 ile 0.91 arasında geniş bir aralıkta değiştiği görülmektedir. Bu çalışmadaki sonuçlar ile literatürdeki bazı çalışmalardan (Becerra-Rondón vd. 2023, Grzybowski vd. 2023, Ngo vd. 2023) daha başarılı model performansı elde edilmiş olup bazı çalışmalarla (Shao vd. 2023, Zhao vd. 2023) ise çok yakın sonuçlar elde edilmiştir. Bununla beraber Long vd. (2022)'nin geliştirdiği bazı modeller bu çalışmanın gerisinde kalırken, literatürdeki birkaç

çalışmada (Chen vd. 2019, Deng vd. 2024, Fu vd. 2023, Tao vd. 2024) daha iyi sonuçlar elde edilmiştir.

Geliştirilen modelin performansı ayrıca istasyon grupları üzerinden de analiz edilmiştir. Şekil 5'te görüldüğü üzere, Endüstri grubu istasyonları için geliştirilen modelde hesaplanan R² ve Uİ değerleri, tüm istasyonları kapsayan modeldeki değerlerle aynıdır. Ayrıca OMH ve KOKH değerleri sırasıyla 8.55 µg/m³ ve 12.91 µg/m³ olarak hesaplanarak hata metriklerinde diğer gruplara göre küçük azalmalar gerçekleştiği tespit edilmiştir. Metropol grubu istasyonları için geliştirilen modelde sırasıyla R² ve Uİ değerleri; 0.66 ve 0.88 olarak hesaplanmıştır. Ayrıca OMH (12.32 µg/m³) ve KOKH (18.93 µg/m³) değerlerinde de artış olmuştur. Model performansındaki bu düşüş büyük oranda kentleşme kaynaklıdır. Özellikle Karşıyaka ve Eğitim istasyonlarındaki bazı uç değerler model performansının zayıflamasına neden olmuştur. En düşük performans ise Banliyö grubu istasyonlarında gerçekleşmiştir. Sırasıyla R² ve Uİ; 0.60 ve 0.85 olarak hesaplanırken, OMH (5.29 µg/m³) ve KOKH (9.09 µg/m³) ise diğer 3 senaryoya göre daha düşük tespit edilmiştir.

Model performansının mevsimsel olarak değerlendirilmesi, her mevsimin kendine özgü atmosferik koşullarına bağlı olarak modelin başarısında farklılıklar ortaya koymuştur. Veri sayıları incelendiğinde en çok veri yaz mevsimi (N=1741) için temin edilebilirken, en az veri kış mevsiminde (N=1044)

Çizelge 2. Literatürdeki NO₂ tahmini yapan bazı çalışmaların performans ölçütleri.

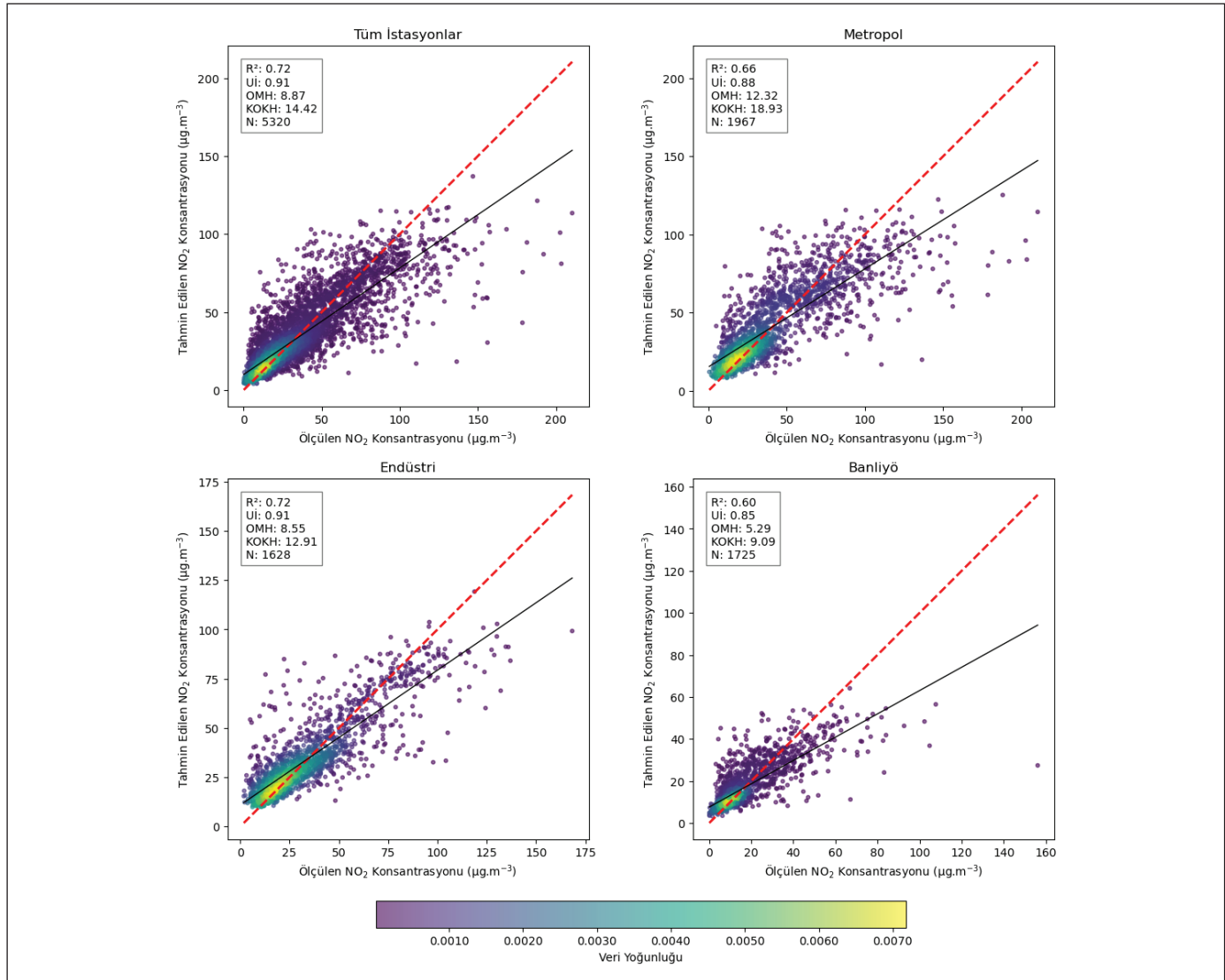
Kullanılan Algoritma	Modele Girdi Olarak Kullanılan Parametreler		İstatistiksel Göstergeler			Kaynak
	Uydu verisi (dikey NO ₂ kolonu)	Diğer parametreler	R ²	KOKH	OMH	
RO	OMI	MERRA 2'den 4 meteoroloji parametresi, dem	0.68	6.43	4.37	Becerra-Rondón vd. 2023
XGBoost	OMI	Meteoroloji istasyonlarından 8 parametre, dem, vejetasyon	0.81	12.53	-	Chen vd. 2019
XGBoost	Sentinel-5P (TROPOMI)	ERA5'ten 6 parametre, dem, nüfus, vejetasyon, arazi kullanımı	0.87	5.36	3.96	Deng vd. 2024
XGBoost	Sentinel-5P (TROPOMI)	Meteoroloji istasyonlarından 7 parametre, dem, nüfus, vejetasyon, arazi kullanımı yol uzunluğu	0.89	9.30	-	Fu vd. 2023
Derin-KSA	Sentinel-5P (TROPOMI)	NLDAS'dan 8 ve MERRA-2'den 2 meteoroloji parametresi, dem, nüfus, vejetasyon, arazi kullanımı	0.83	2.40	1.75	Ghahremanloo vd. 2021
DR*	Sentinel-5P (TROPOMI)	ERA5'ten 7 meteoroloji parametresi, dem, nüfus, yol yoğunluğu	0.32	7.30	4.90	Grzybowski vd. 2023
ÇDR*			0.45	6.60	4.20	
RO			0.53	5.90	3.70	
DVM*			0.54	6.10	3.70	
RO	Sentinel-5P (TROPOMI)	ERA5'ten 7 meteoroloji parametresi, vejetasyon)	0.80	6.28	4.38	Hu vd. 2024
RO	Sentinel-5P (TROPOMI)	GFS'den 5 meteoroloji parametresi, dem, nüfus, arazi kullanımı	0.78	7.04	5.00	Li vd. 2022
KA*	Sentinel-5P (TROPOMI)	ERA5'ten 9 meteoroloji parametresi, MEIC'den 10 farklı emisyon girdisi, dem, arazi kullanımı	0.61	4.91	3.22	Long vd. 2022
GAKA*			0.65	4.67	3.24	
RO			0.80	3.57	2.41	
EA			0.80	3.52	2.38	
ST-KA*			0.62	4.90	3.21	
ST-GAKA*			0.66	4.63	3.20	
ST-RO*			0.81	3.45	2.33	
ST-EA*			0.81	3.45	2.31	
SA*	Sentinel-5P (TROPOMI), OMI	ERA5'ten 4 meteoroloji parametresi, vejetasyon, yol yoğunluğu	0.30	10.43	-	Ngo vd. 2023
LightGBM			0.59	7.93	-	
XGBoost	OMI, GOME-2	ERA5'ten 12 meteoroloji parametresi, arazi kullanımı	0.75	9.11	6.72	Shao vd. 2023
XGBoost	Sentinel-5P (TROPOMI)	ERA5'ten 5 meteoroloji parametresi, dem, nüfus, vejetasyon, arazi kullanımı, yol verisi	0.72	5.40	-	Zhao vd. 2023
	OMI		0.73	5.40	-	
RO	Sentinel-5P (TROPOMI)	ERA5'ten 10 meteoroloji parametresi, dem, vejetasyon	0.72	14.42	8.87	Bu çalışma

*DR: Doğrusal regresyon, ÇDR: Çoklu doğrusal regresyon, DVM: Destek vektör makinesi, KA: Karar Ağacı, GAKA: Gradyan Artırmalı Karar Ağacı, ST-: Mekansal ve zamansal destekli, -SA: *Sinir ağı*

kullanılabilmiştir. Yazın havanın açık (bulutsuz) olması nedeniyle veri kalitesi ve miktarında artış olurken, kış aylarında tam tersi durum gözlemlenmektedir. Bu durum modelin başarısını da etkilemektedir. Modelin en yüksek başarısı yaz mevsiminde gerçekleşirken, sonbahar ve kış mevsimlerinde performansın göreceli olarak düştüğü tespit edilmiştir.

Yaz mevsimi için R^2 ve $U\bar{I}$ değerleri sırasıyla 0.81 ve 0.94 olarak hesaplanmış olup, tahmin edilen NO₂ konsantrasyonlarının oldukça yüksek uyumda olduğu görülmüştür. OMH (7.00 $\mu\text{g}/\text{m}^3$) ve KOKH (12.17 $\mu\text{g}/\text{m}^3$) değerleri de tüm verilerle geliştirilen modele göre daha düşük olup, yaz mevsiminde model performansının arttığını ortaya koymaktadır. İlkbaharda model performansı yaz mevsimine göre düşük

olmasına rağmen, tüm verilerle geliştirilen model ile benzer bir performans sergilemiştir. Korelasyon parametreleri ($R^2=0.72$ ve $U\bar{I}=0.91$), tüm verilerle geliştirilen model için hesaplananlar ile aynı elde edilirken, hata göstergelerinin de daha küçük değerler olduğu tespit edilmiştir (OMH = 8.22 $\mu\text{g}/\text{m}^3$ ve KOKH=12.54 $\mu\text{g}/\text{m}^3$). Kış ve sonbahar mevsimlerinde model performansında azalma görülmüştür. Kış mevsiminde R^2 ve $U\bar{I}$ sırasıyla 0.62 ve 0.87 olarak hesaplanırken, sonbaharda da benzer bir eğilim tespit edilmiş olup bu değerler sırasıyla 0.65 ve 0.87 olarak tespit edilmiştir. Korelasyondaki bu azalma hem sonbahar ve kış aylarında antropojenik kirletici kaynaklar nedeniyle atmosferde görülen yüksek kirletici konsantrasyonlarından hem de bu

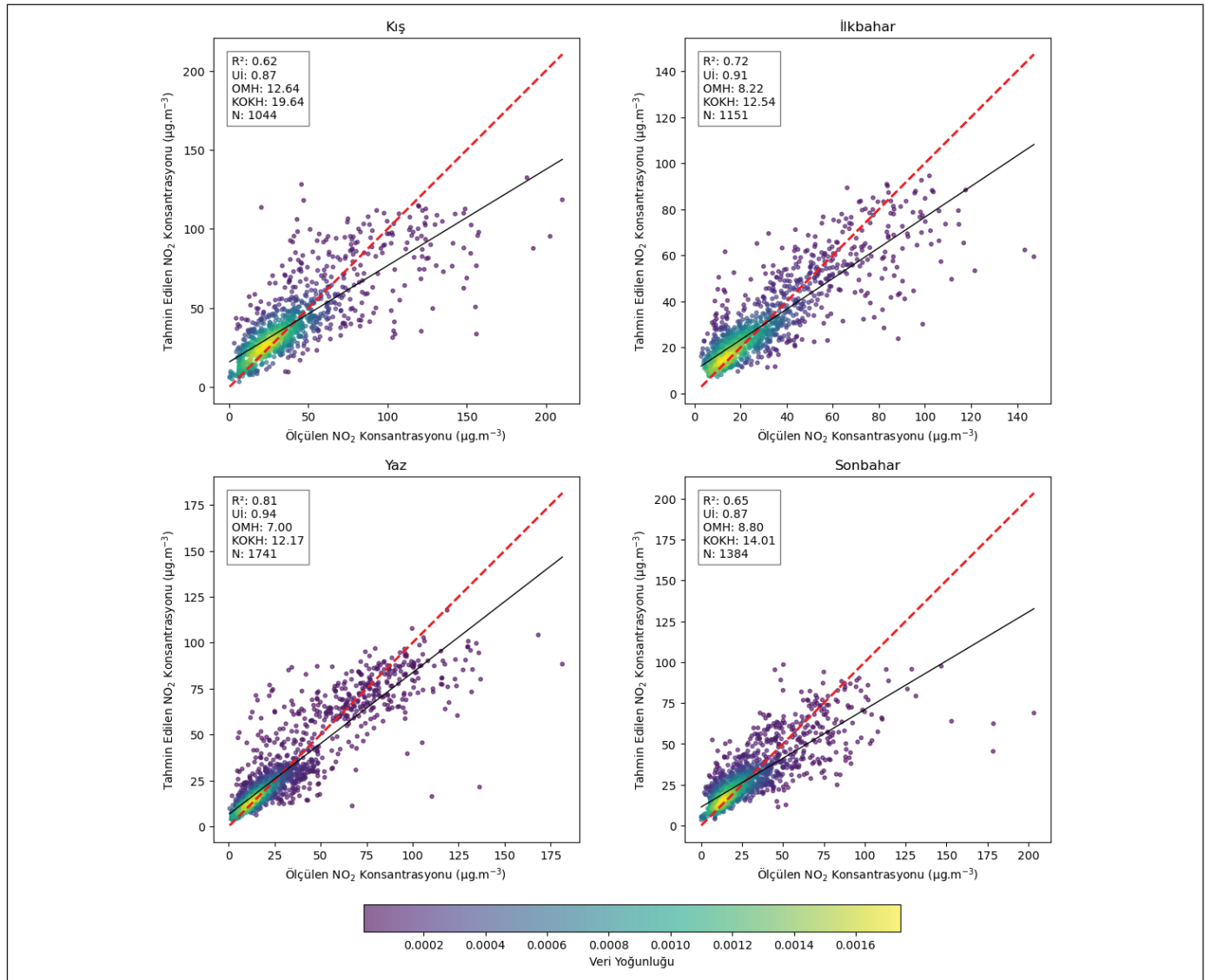


Şekil 5. İzmir'de farklı istasyon grupları için ölçülen ve tahmin edilen saatlik NO₂ konsantrasyonlarına ait dağılım grafikleri (N: Veri sayısı).

mevsimde bulutluluk vb. nedenlerle veri sayısı ve kalitesinin düşmesinden kaynaklanmaktadır. Hava kalitesi seviyelerinin özellikle kış aylarında daha fazla olması nedeniyle OMH (Kış: 12.64 $\mu\text{g}/\text{m}^3$ ve Sonbahar: 8.80 $\mu\text{g}/\text{m}^3$) ve KOKH (Kış: 19.64 $\mu\text{g}/\text{m}^3$ ve Sonbahar: 14.01 $\mu\text{g}/\text{m}^3$) değerleri de diğer mevsimlere göre yüksek hesaplanmıştır (Şekil 6).

Modelin uç değerlerdeki başarısının değerlendirilmesi için, farklı yüzdelik dilimler için performans analizi yapılmıştır. Saatlik NO₂ konsantrasyonları %0-10, %10-90 ve %90-100 yüzdelik dilimleri için üç gruba ayrılmış olup bu gruplar için modelin tahmin başarısını gösteren dağılım grafikleri Şekil 7'de verilmiştir. Buna göre; ikinci grup (%10-%90) için geliştirilen model en iyi performansı verirken ($R^2 = 0.65$ ve

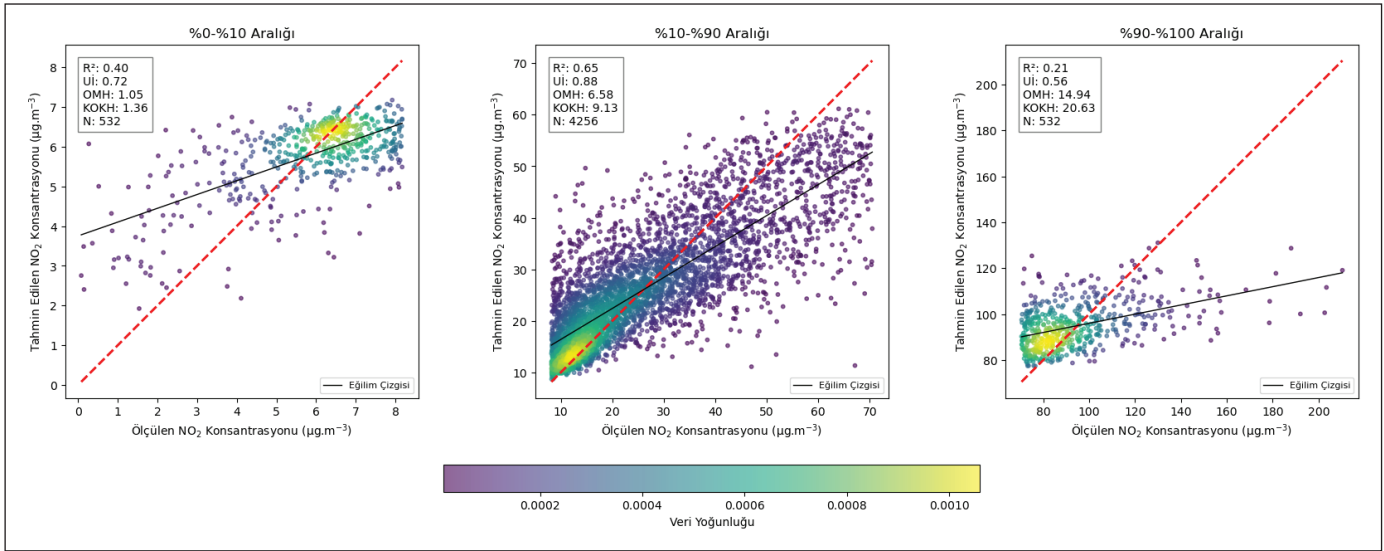
$UI = 0.88$) üçüncü grup (>%90) için geliştirilen model en başarısız performansı sunmuştur ($R^2 = 0.21$ ve $UI = 0.56$). Diğer yandan birinci grup (<%10) için geliştirilen modelde diğer iki durum arasında bir korelasyon elde edilmiştir ($R^2 = 0.40$ ve $UI = 0.72$). Bu durum, modellerin çok büyük ve çok küçük NO₂ konsantrasyon değerlerini tahmin etmekte başarısız olduğunu göstermektedir. En iyi performans, veri setinin daha dengeli dağılım gösteren aralıklarında elde edilmektedir. OMH ve KOKH hesapları, tıpkı diğer analizlerde olduğu gibi, küçük değerlerle geliştirilen modellerde daha düşük, yüksek konsantrasyonlarla ile geliştirilenlerde daha yüksek hesaplanmıştır.



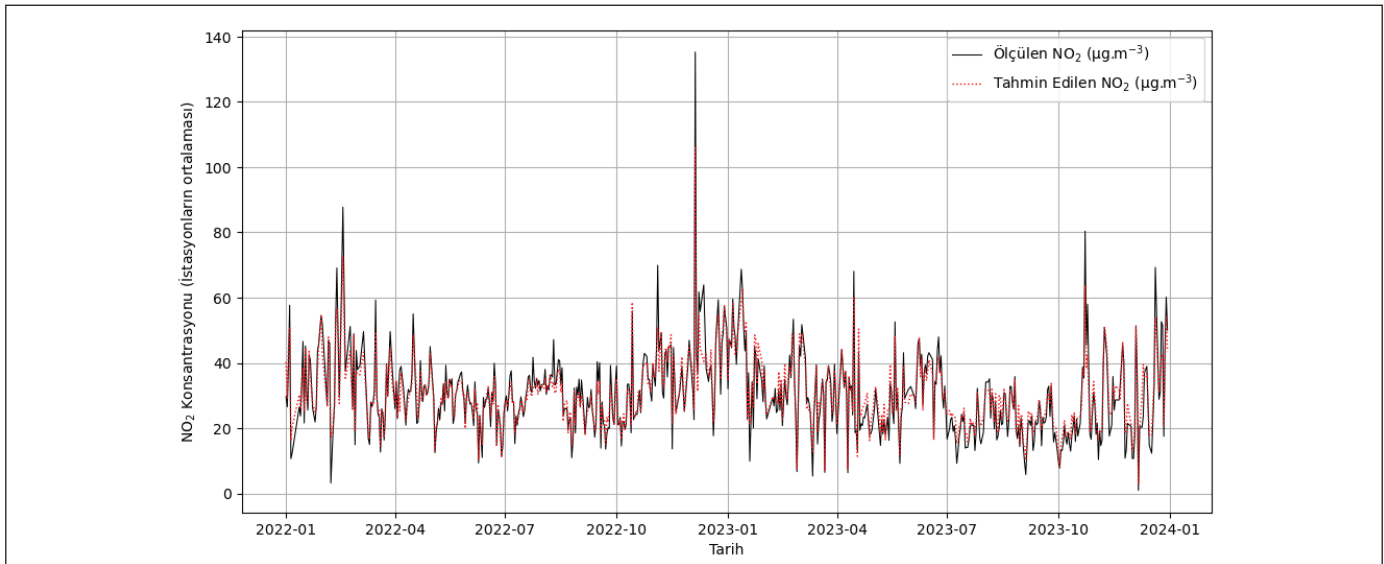
Şekil 6. Ölçülen ve tahmin edilen NO₂ konsantrasyonlarının mevsimsel dağılım grafikleri (N: Veri sayısı).

Şekil 8’de 1 Ocak 2022 ile 31 Aralık 2023 tarihleri arasında model tarafından tahmin edilen saatlik NO₂ konsantrasyonları ile ölçülen konsantrasyonlar bir zaman serisi üzerinde karşılaştırılmıştır. Bu karşılaştırmada aynı tarihte farklı istasyonlardan gelen verilerin ortalaması alınmıştır. Elde edilen zaman serisinde, tahmin edilen konsantrasyonlar ile gerçek konsantrasyonların değişimleri arasında genel olarak çok güçlü bir uyum görülmektedir. Diğer yandan siyah çizgi ile gösterilen model tahminlerinin, kırmızı çizgi ile gösterilen ölçüm değerlerine, pik konsantrasyonların ölçüldüğü noktalarda eşlik edemediği görülmektedir. Bu da modelin hem çok küçük hem de çok büyük konsantrasyonlar için tahmin başarısının yeterli olmadığına işaret etmektedir.

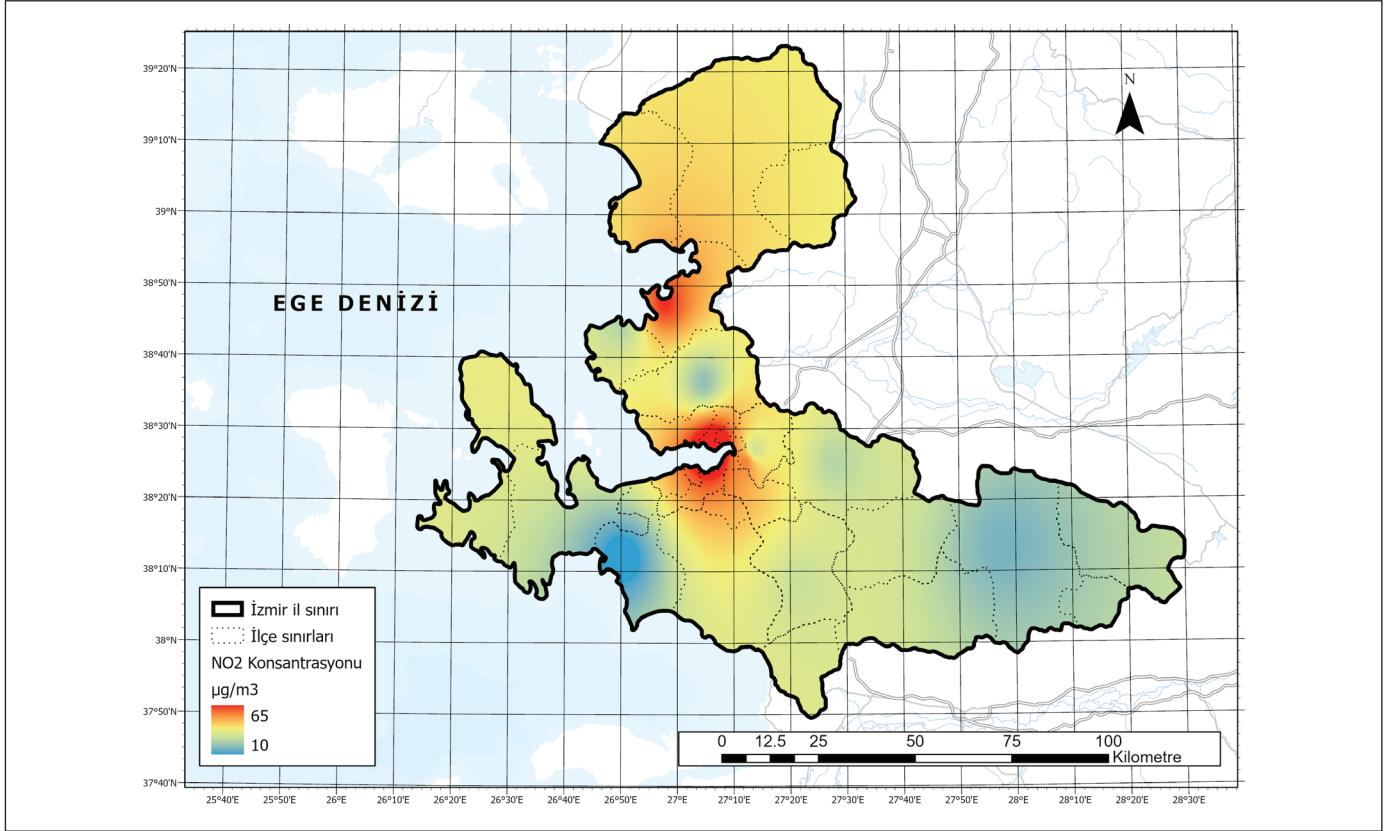
Geliştirilen modelle her bir istasyon için 2022-2023 yılları arasında günlük olarak NO₂ tahminleri yapılmıştır. Tahminlerin çalışma dönemi için zamansal ortalamaları alınarak görselleştirilmiştir. Bu ortalamalar ArcGIS Pro’nun IDW aracıyla interpolate edilerek Şekil 9’daki harita elde edilmiştir. Haritada da görüldüğü gibi en yüksek NO₂ konsantrasyonları yerleşim ve trafiğin yoğun olduğu körfez etrafındaki metropol bölgesinde ve Aliğa civarındaki endüstriyel bölgede tespit edilirken banliyö bölgelerinde daha düşük konsantrasyonlar tespit edilmiştir.



Şekil 7. Farklı yüzdelik dilimleri için ölçülen ve tahmin edilen NO₂ konsantrasyonlarının dağılım grafikleri (N: Veri sayısı).



Şekil 8. Ölçülen ve tahmin edilen saatlik NO₂ konsantrasyonları.



Şekil 9. Çalışma alanında 2022-2023 yılları arasında model ile tahmin edilen ortalama NO₂ konsantrasyonlarının dağılımı

4. Değerlendirme ve Öneriler

Bu çalışma ile elde edilen sonuçlar, makine öğrenmesi yöntemlerinin İzmir bölgesinde yüksek tahmin başarıları sağlayabileceğini ortaya koymaktadır. Hesaplanan korelasyon ve hata metrikleri literatürle genel olarak uyumlu olup, kent ölçeğinde bir çalışma olmasına rağmen yurtdışındaki birçok bölgeden daha iyi performans göstermiştir (Becerra-Rondón vd. 2023, Grzybowski vd. 2023, Ngo vd. 2023). Bu bulgular, ülkemizde ölçüm yapılmayan bölgelerde NO₂ konsantrasyonlarının başarılı bir şekilde tahmini için umut vaat etmektedir.

Makine öğrenmesi modelinde bölgeye özgü çeşitli meteorolojik ve çevresel parametrelerin girdi olarak kullanılması, tahmin başarısını önemli düzeyde arttırmıştır. İlk kez bu çalışmada girdi olarak kullanılan ERA5-Land ürünlerinden lai_{lv} ve lai_{hv} değişkenlerinin, NO₂ tahminlerinde belirgin bir etkisi olduğu tespit edilmiştir. Ayrıca, uydu gözlemlerinin kalitesini yansıtan ve albedoyu temsil eden fal değişkeninin yüksek etkisi, ERA5-Land ürünlerinin başarısını daha da pekiştirmiştir. Bunun yanı sıra, Türkiye’de bir NO₂ tahmini çalışmasında ilk kez kullanılan Sentinel-5P TRO-

POMI verisi, tahmin doğruluğunu önemli ölçüde artırmış ve en etkili iki girdiden biri olarak öne çıkmıştır.

Model performansının mekânsal ve mevsimsel değişimlerden etkilendiği gözlemlenmiştir. Çok küçük ya da çok büyük uç değerlerin görüldüğü lokasyon ve zamanlarda model performansı düşerken, dağılımın normal aralığında iyileştiği belirlenmiştir. Özellikle, Sentinel-5P TROPOMI veri kalitesinin ve miktarının azaldığı kış aylarında model performansı düşmüş, yaz mevsiminde ise veri bolluğu sayesinde performans artmıştır.

Bununla birlikte, ERA5-Land ürünlerinin sunduğu ayrıntılı mekânsal veri model performansına önemli katkı sağlasa da kara verisi olması nedeniyle birçok kıyı istasyonunda veri alınamamasına neden olmuştur. Bu durum, İzmir gibi kıyı şehirleri için önemli bir dezavantaj oluşturmaktadır. Bu nedenle, daha düşük çözünürlüklü olsa da ERA5 Single Level, GFS veya MERRA-2 gibi alternatif yeniden analiz ürünleri kullanılarak modellerin geliştirilmesi gerekmektedir. Ayrıca, bu çalışmada yalnızca RO algoritması kullanılmıştır. Model kaynaklı hataları minimize etmek ve aşırı uyumun önüne geçmek amacıyla farklı model varyasyonlarının denenmesi,

daha güçlü modellerin geliştirilmesi için kritik bir ileri çalışma alanı olacaktır.

Ulusal ölçekte gerçekleştirilecek geniş çaplı çalışmalar, modelin genel başarısını test etmek ve NO₂ konsantrasyonlarının bölgesel dağılımını anlamak açısından faydalı olacaktır.

Yazar katkısı: Efem BİLGİÇ: Çalışma hakkında verileri toplamış, modeli geliştirmiş, analiz etmiş ve makaleyi yazmıştır, Tolga ELBİR: çalışmayı planlamış, tasarlamış ve yazımına katkı sağlamıştır.

5. Kaynaklar

- Ahmad, N., Lin, C., Lau, AKH., Kim, J., Zhang, T., Yu, F., ... Lao, XQ. 2024. Estimation of ground-level NO₂ and its spatiotemporal variations in China using GEMS measurements and a nested machine learning model. *Atmospheric Chemistry and Physics*, 24(16), 9645-9665. <https://doi.org/10.5194/ACP-24-9645-2024>
- Becerra-Rondón, A., Ducati, J., Haag, R. 2023. Satellite-based estimation of NO₂ concentrations using a machine-learning model: A case study on Rio Grande do Sul, Brazil. *Atmosfera*, 37, 175-190. <https://doi.org/10.20937/ATM.53116>
- Breiman, L. 2001. Random forests. *Machine Learning*, 45(1), 5-32. <https://doi.org/10.1023/A:1010933404324>
- Castro, A., Künzli, N., Götschi, T. 2017. Health benefits of a reduction of PM₁₀ and NO₂ exposure after implementing a clean air plan in the Agglomeration Lausanne-Morges. *International Journal of Hygiene and Environmental Health*, 220(5), 829-839. <https://doi.org/10.1016/J.IJHEH.2017.03.012>
- Chen, ZY., Zhang, R., Zhang, TH., Ou, CQ., Guo, Y. 2019. A kriging-calibrated machine learning method for estimating daily ground-level NO₂ in mainland China. *Science of The Total Environment*, 690, 556-564. <https://doi.org/10.1016/J.SCITOTENV.2019.06.349>
- Chi, Y., Fan, M., Zhao, C., Yang, Y., Fan, H., Yang, X., ... Tao, J. 2022. Machine learning-based estimation of ground-level NO₂ concentrations over China. *Science of The Total Environment*, 807, 150721. <https://doi.org/10.1016/J.SCITOTENV.2021.150721>
- Cichowicz, R., Bochenek, AD. 2024. Assessing the effects of urban heat islands and air pollution on human quality of life. *Anthropocene*, 46, 100433. <https://doi.org/10.1016/J.ANCENE.2024.100433>
- Cooper, MJ., Martin, RV., Hammer, MS., Levelt, PF., Veeffkind, P., Lamsal, LN., ... McLinden, CA. 2022. Global fine-scale changes in ambient NO₂ during COVID-19 lockdowns. *Nature*, 601, 380-387. <https://doi.org/10.1038/s41586-021-04229-0>
- Copernicus. 2024. Sentinel-5P NO₂ Data Product. https://sentinels.copernicus.eu/web/sentinel/data-products/-/asset_publisher/fp37fc19FN8F/content/sentinel-5-precursor-level-2-nitrogen-dioxide
- de la Cruz Libardi, A., Masselot, P., Schneider, R., Nightingale, E., Milojevic, A., Vanoli, J., ... Gasparini, A. 2024. High resolution mapping of nitrogen dioxide and particulate matter in Great Britain (2003–2021) with multi-stage data reconstruction and ensemble machine learning methods. *Atmospheric Pollution Research*, 15(11), 102284. <https://doi.org/10.1016/J.APR.2024.102284>
- Deng, F., Chen, Y., Liu, W., Li, L., Chen, X., Tiwari, P., Qin, K. 2024. Satellite-Based Estimation of Near-Surface NO₂ Concentration in Cloudy and Rainy Areas. *Remote Sensing*, 16(10), 1785. <https://doi.org/10.3390/RS16101785>
- Dokuz, Y., Bozdağ, A., Gökçek, ÖB. 2020. Hava kalitesi parametrelerinin tahmini ve mekansal dağılımı için makine öğrenmesi yöntemlerinin kullanılması. *Niğde Ömer Halisdemir Üniversitesi Mühendislik Bilimleri Dergisi*, 9(1), 37-47. <https://doi.org/10.28948/NGUMUH.654092>
- Elbir, T., Kara, M., Bayram, A., Altıok, H., Dumanoglu, Y. 2011. Comparison of predicted and observed PM₁₀ concentrations in several urban street canyons. *Air Quality, Atmosphere and Health*, 4, 121-131. <https://doi.org/10.1007/S11869-010-0080-9>
- Elbir, T., Muezzinoglu, A. 2004. Estimation of emission strengths of primary air pollutants in the city of Izmir, Turkey. *Atmospheric Environment*, 38(13), 1851-1857. <https://doi.org/10.1016/J.ATMOENV.2004.01.015>
- Fu, J., Tang, D., Grieneisen, ML., Yang, F., Yang, J., Wu, G., ... Zhan, Y. 2023. A machine learning-based approach for fusing measurements from standard sites, low-cost sensors, and satellite retrievals: Application to NO₂ pollution hotspot identification. *Atmospheric Environment*, 302, 119756. <https://doi.org/10.1016/J.ATMOENV.2023.119756>
- Ghahremanloo, M., Lops, Y., Choi, Y., Yeganeh, B. 2021. Deep Learning Estimation of Daily Ground-Level NO₂ Concentrations From Remote Sensing Data. *Journal of Geophysical Research: Atmospheres*, 126(21), e2021JD034925. <https://doi.org/10.1029/2021JD034925>
- Goldberg, DL., Anenberg, SC., Griffin, D., McLinden, CA., Lu, Z., Streets, DG. 2020. Disentangling the Impact of the COVID-19 Lockdowns on Urban NO₂ From Natural Variability. *Geophysical Research Letters*, 47(17), e2020GL089269. <https://doi.org/10.1029/2020GL089269>
- Google Earth Engine. 2024. Earth Engine Data Catalog - Sentinel-5P OFFL NO₂: Offline Nitrogen Dioxide. https://developers.google.com/earth-engine/datasets/catalog/COPERNICUS_S5P_OFFL_L3_NO2#description

- Griffin, D., Zhao, X., McLinden, CA., Boersma, F., Bourassa, A., Dammers, E., ... Wolde, M. 2019.** High-Resolution Mapping of Nitrogen Dioxide With TROPOMI: First Results and Validation Over the Canadian Oil Sands. *Geophysical Research Letters*, 46(2), 1049-1060. <https://doi.org/10.1029/2018GL081095>
- Grzybowski, PT., Markowicz, KM., Musiał, JP. 2023.** Estimations of the Ground-Level NO₂ Concentrations Based on the Sentinel-5P NO₂ Tropospheric Column Number Density Product. *Remote Sensing* 2023, Vol. 15, Page 378, 15(2), 378. <https://doi.org/10.3390/RS15020378>
- Gui, K., Che, H., Zeng, Z., Wang, Y., Zhai, S., Wang, Z., ... Zhang, X. 2020.** Construction of a virtual PM_{2.5} observation network in China based on high-density surface meteorological observations using the Extreme Gradient Boosting model. *Environment International*, 141, 105801. <https://doi.org/10.1016/J.ENVINT.2020.105801>
- Gündoğdu, S., Elbir, T. 2024a.** A data-driven approach for PM_{2.5} estimation in a metropolis: random forest modeling based on ERA5 reanalysis data. *Environmental Research Communications*, 6(3), 035029. <https://doi.org/10.1088/2515-7620/AD352D>
- Gündoğdu, S., Elbir, T. 2024b.** Elevating hourly PM_{2.5} forecasting in Istanbul, Türkiye: Leveraging ERA5 reanalysis and genetic algorithms in a comparative machine learning model analysis. *Chemosphere*, 364, 143096. <https://doi.org/10.1016/J.CHEMOSPHERE.2024.143096>
- Gündoğdu, S., Tuna Tuygun, G., Li, Z., Wei, J., Elbir, T. 2022.** Estimating daily PM_{2.5} concentrations using an extreme gradient boosting model based on VIIRS aerosol products over southeastern Europe. *Air Quality, Atmosphere and Health*, 15(12), 2185-2198. <https://doi.org/10.1007/S11869-022-01245-5>
- Hanif, A., Sun, M., Wang, T., Shang, S., Tsang, DCW., Shang, J. 2021.** Ambient NO₂ adsorption removal by Mg–Al layered double hydroxides and derived mixed metal oxides. *Journal of Cleaner Production*, 313, 127956. <https://doi.org/10.1016/J.JCLEPRO.2021.127956>
- Hashim, BM., Al-Naseri, SK., Al-Maliki, A., Al-Ansari, N. 2021.** Impact of COVID-19 lockdown on NO₂, O₃, PM_{2.5} and PM₁₀ concentrations and assessing air quality changes in Baghdad, Iraq. *Science of The Total Environment*, 754, 141978. <https://doi.org/10.1016/J.SCITOTENV.2020.141978>
- Hersbach, H., Bell, B., Berrisford, P., Biavati, G., Horányi, A., Muñoz Sabater, J., ... Thépaut, JN. 2023.** ERA5 hourly data on single levels from 1940 to present. Copernicus Climate Change Service (C3S) Climate Data Store (CDS).
- Hu, M., Bai, K., Li, K., Zheng, Z., Sun, Y., Shao, L., ... Liu, C. 2024.** Improving machine-learned surface NO₂ concentration mapping models with domain knowledge from data science perspective. *Atmospheric Environment*, 322, 120372. <https://doi.org/10.1016/J.ATMOSENV.2024.120372>
- Ibrahim, S., Landa, M., Pešek, O., Brodský, L., Halounová, L. 2022.** Machine Learning-Based Approach Using Open Data to Estimate PM_{2.5} over Europe. *Remote Sensing* 2022, Vol. 14, Page 3392, 14(14), 3392. <https://doi.org/10.3390/RS14143392>
- Ialongo, I., Virta, H., Eskes, H., Hovila, J., Douros, J. 2020.** Comparison of TROPOMI/Sentinel-5 Precursor NO₂ observations with ground-based measurements in Helsinki. *Atmospheric Measurement Techniques*, 13(1), 205-218. <https://doi.org/10.5194/AMT-13-205-2020>
- Kang, Y., Choi, H., Im, J., Park, S., Shin, M., Song, CK., Kim, S. 2021.** Estimation of surface-level NO₂ and O₃ concentrations using TROPOMI data and machine learning over East Asia. *Environmental Pollution*, 288, 117711. <https://doi.org/10.1016/J.ENVPOL.2021.117711>
- Kharol, SK., McLinden, CA., Sioris, CE., Shephard, MM., Fioletov, V., Van Donkelaar, A., ... Martin, RV. 2017.** OMI satellite observations of decadal changes in ground-level sulfur dioxide over North America. *Atmospheric Chemistry and Physics*, 17(9), 5921-5929. <https://doi.org/10.5194/ACP-17-5921-2017>
- Kim, M., Brunner, D., Kuhlmann, G. 2021.** Importance of satellite observations for high-resolution mapping of near-surface NO₂ by machine learning. *Remote Sensing of Environment*, 264, 112573. <https://doi.org/10.1016/J.RSE.2021.112573>
- Levelt, PF., Van Den Oord, GHJ., Dobber, MR., Mälikki, A., Visser, H., De Vries, J., ... Saari, H. 2006.** The ozone monitoring instrument. *IEEE Transactions on Geoscience and Remote Sensing*, 44(5), 1093-1100. <https://doi.org/10.1109/TGRS.2006.872333>
- Li, M., Wu, Y., Bao, Y., Liu, B., Petropoulos, GP. 2022.** Near-Surface NO₂ Concentration Estimation by Random Forest Modeling and Sentinel-5P and Ancillary Data. *Remote Sensing* 2022, Vol. 14, Page 3612, 14(15), 3612. <https://doi.org/10.3390/RS14153612>
- Liu, Y., Chen, X., Huang, S., Tian, L., Lu, Y., Mei, Y., ... Xiang, H. 2015.** Association between Air Pollutants and Cardiovascular Disease Mortality in Wuhan, China. *International Journal of Environmental Research and Public Health* 2015, Vol. 12, Pages 3506-3516, 12(4), 3506-3516. <https://doi.org/10.3390/IJERPH120403506>

- Long, S., Wei, X., Zhang, F., Zhang, R., Xu, J., Wu, K., ... Li, W. 2022.** Estimating daily ground-level NO₂ concentrations over China based on TROPOMI observations and machine learning approach. *Atmospheric Environment*, 289, 119310. <https://doi.org/10.1016/J.ATMOSENV.2022.119310>
- Muñoz Sabater, J. 2019.** ERA5-Land hourly data from 1950 to present. Copernicus Climate Change Service (C3S) Climate Data Store (CDS).
- Ngo, TX., Phan, HDT., Nguyen, TTN. 2023.** Development of ground-level NO₂ models in Vietnam using machine learning and satellite observations with ancillary data. *Frontiers in Environmental Science*, 11, 1187592. <https://doi.org/10.3389/FENVS.2023.1187592>
- Qin, K., Han, X., Li, D., Xu, J., Li, D., Loyola, D., ... Yuan, L. 2020.** Satellite-based estimation of surface NO₂ concentrations over east-central China: A comparison of POMINO and OMNO2d data. *Atmospheric Environment*, 224, 117322. <https://doi.org/10.1016/J.ATMOSENV.2020.117322>
- Schneider, R., Vicedo-Cabrera, AM., Sera, F., Masselot, P., Stafoggia, M., de Hoogh, K., ... Gasparrini, A. 2020.** A Satellite-Based Spatio-Temporal Machine Learning Model to Reconstruct Daily PM_{2.5} Concentrations across Great Britain. *Remote Sensing* 2020, Vol. 12, Page 3803, 12(22), 3803. <https://doi.org/10.3390/RS12223803>
- Shao, Y., Zhao, W., Liu, R., Yang, J., Liu, M., Fang, W., ... Ma, Z. 2023.** Estimation of daily NO₂ with explainable machine learning model in China, 2007–2020. *Atmospheric Environment*, 314, 120111. <https://doi.org/10.1016/J.ATMOSENV.2023.120111>
- Shetty, S., Schneider, P., Stebel, K., David Hamer, P., Kylling, A., Koren Berntsen, T. 2024.** Estimating surface NO₂ concentrations over Europe using Sentinel-5P TROPOMI observations and Machine Learning. *Remote Sensing of Environment*, 312, 114321. <https://doi.org/10.1016/J.RSE.2024.114321>
- Sünsüli, M., Kalkan, K. 2022.** Sentinel-5p Uydu Görüntüleri İle Azot Dioksit (NO₂) Kirliliğinin İzlenmesi. *Türkiye Uzaktan Algılama Dergisi*, 4(1), 1-6. <https://doi.org/10.51489/TUZAL.1056261>
- Tao, C., Jia, M., Wang, G., Zhang, Y., Zhang, Q., Wang, X., ... Wang, W. 2024.** Time-sensitive prediction of NO₂ concentration in China using an ensemble machine learning model from multi-source data. *Journal of Environmental Sciences*, 137, 30-40. <https://doi.org/10.1016/J.JES.2023.02.026>
- Tuna Tuygun, G., Elbir, T. 2023.** Estimation of particulate matter concentrations in Türkiye using a random forest model based on satellite AOD retrievals. *Stochastic Environmental Research and Risk Assessment*, 37(9), 3469-3491. <https://doi.org/10.1007/s00477-023-02459-4>
- Tuna Tuygun, G., Gündoğdu, S., Elbir, T. 2021.** Estimation of ground-level particulate matter concentrations based on synergistic use of MODIS, MERRA-2 and AERONET AODs over a coastal site in the Eastern Mediterranean. *Atmospheric Environment*, 261, 118562. <https://doi.org/10.1016/J.ATMOSENV.2021.118562>
- Ünal Çilek, M. 2022.** Troposferik Nitrojen Dioksitin (NO₂) COVID-19 Pandemisinde Mekânsal ve Zamansal Analizi: Adana-Mersin Bölgesi. *Yüzüncü Yıl Üniversitesi Fen Bilimleri Enstitüsü Dergisi*, 27(3), 581s-594. <https://doi.org/10.53433/YYUFBED.1119418>
- Ünal, S., Yalçın, N. 2022.** Hava Kirliliğinin Makine Öğrenmesi Tabanlı Tahmini: Başakşehir Örneği. *Mühendislik Bilimleri ve Araştırmaları Dergisi*, 4(1), 35-44. <https://doi.org/10.46387/BJESR.1055946>
- Veefkind, JP., Aben, I., McMullan, K., Förster, H., de Vries, J., Otter, G., ... Levelt, PF. 2012.** TROPOMI on the ESA Sentinel-5 Precursor: A GMES mission for global observations of the atmospheric composition for climate, air quality and ozone layer applications. *Remote Sensing of Environment*, 120, 70-83. <https://doi.org/10.1016/J.RSE.2011.09.027>
- Wei, J., Li, Z., Sun, L., Xue, W., Ma, Z., Liu, L., ... Cribb, M. 2022.** Extending the EOS Long-Term PM_{2.5}Data Records since 2013 in China: Application to the VIIRS Deep Blue Aerosol Products. *IEEE Transactions on Geoscience and Remote Sensing*, 60. <https://doi.org/10.1109/TGRS.2021.3050999>
- Weinmayr, G., Romeo, E., de Sario, M., Weiland, SK., Forastiere, F. 2010.** Short-Term effects of PM₁₀ and NO₂ on respiratory health among children with asthma or asthma-like symptoms: A systematic review and Meta-Analysis. *Environmental Health Perspectives*, 118(4), 449-457. <https://doi.org/10.1289/ehp.0900844>
- Yao, F., Si, M., Li, W., Wu, J. 2018.** A multidimensional comparison between MODIS and VIIRS AOD in estimating ground-level PM_{2.5} concentrations over a heavily polluted region in China. *Science of The Total Environment*, 618, 819-828. <https://doi.org/10.1016/J.SCITOTENV.2017.08.209>
- Yavaşlı, DD., Ölgün, MK. 2022.** Impacts of COVID-19 Pandemic on Tropospheric NO₂ over Turkey. *Aegean Geographical Journal*, 31(2), 255-264. <https://doi.org/10.51800/ECD.1109104>
- Zhang, Q., Pan, Y., He, Y., Walters, WW., Ni, Q., Liu, X., ... Jiang, C. 2021.** Substantial nitrogen oxides emission reduction from China due to COVID-19 and its impact on surface ozone and aerosol pollution. *Science of The Total Environment*, 753, 142238. <https://doi.org/10.1016/J.SCITOTENV.2020.142238>

- Zhang, R., Zhang, Y., Lin, H., Feng, X., Fu, T.M., Wang, Y. 2020.** NO_x Emission Reduction and Recovery during COVID-19 in East China. *Atmosphere* 2020, Vol. 11, Page 433, 11(4), 433. <https://doi.org/10.3390/ATMOS11040433>
- Zhao, Z., Lu, Y., Zhan, Y., Cheng, Y., Yang, F., Brook, J.R., He, K. 2023.** Long-term spatiotemporal variations in surface NO₂ for Beijing reconstructed from surface data and satellite retrievals. *Science of The Total Environment*, 904, 166693. <https://doi.org/10.1016/J.SCITOTENV.2023.166693>
- Zhu, F., Ding, R., Lei, R., Cheng, H., Liu, J., Shen, C., ... Cao, J. 2019.** The short-term effects of air pollution on respiratory diseases and lung cancer mortality in Hefei: A time-series analysis. *Respiratory Medicine*, 146, 57-65. <https://doi.org/10.1016/J.RMED.2018.11.019>