



PM₁₀ Parametresinin Makine Öğrenmesi Algoritmaları ile Mekânsal Analizi, Kayseri İli Örneği

Spatial Analysis of PM₁₀ Parameter by Machine Learning Algorithms, City of Kayseri

Öznur Begüm Gökçek^{1*}, **Nuray Şaşa¹**, **Yeşim Dokuz²**, **Aslı Bozdağ³**

¹ Niğde Ömer Halisdemir Üniversitesi Mühendislik Fakültesi Çevre Mühendisliği Bölümü, Niğde, TÜRKİYE

² Niğde Ömer Halisdemir Üniversitesi Mühendislik Fakültesi Bilgisayar Mühendisliği Bölümü, Niğde, TÜRKİYE

³ Niğde Ömer Halisdemir Üniversitesi Mühendislik Fakültesi Harita Mühendisliği Bölümü, Niğde, TÜRKİYE

Sorumlu Yazar / Corresponding Author *: begumgokcek@ohu.edu.tr

Geliş Tarihi / Received: 18.04.2021

Kabul Tarihi / Accepted: 30.08.2021

Araştırma Makalesi/Research Article

DOI:10.21205/deufmd.2022247008

Atıf şekli/How to cite: GÖKÇEK, Ö.B., ŞAŞA, N., DOKUZ, Y., BOZDAĞ, A.(2022). PM10 Parametresinin Makine Öğrenmesi Algoritmaları ile Mekânsal Analizi, Kayseri İli Örneği .DEUFMD, 24(70), 65-80.

Öz

Hava kirliliğinin son yıllarda artışı ile alınacak olan erken önlemler dâhilinde hava kirliliği tahmininin yapılması insan ve çevre sağlığına verilebilecek zararın en aza indirilmesinde önemlidir. Bu çalışmada günlük ortalama hava kirliliği miktarının, önemli bir hava kirletici olan PM₁₀ konsantrasyonu üzerinden tahminlenmesi ve hava kirliliğinin çevresel ve mekânsal modellenmesi amaçlanmıştır. Tahminleme modeli, Orta Anadolu Bölgesinde yer alan Kayseri ilinde bulunan 3 istasyondan alınan 2010-2018 yılları arasında ölçülen PM₁₀ konsantrasyonu verileri kullanılarak makine öğrenmesi algoritmaları (kNN, DVR, RF, ANN, Lineer Regresyon) ile eğitilmiştir. Kayseri'deki 3 istasyonun 2010-2018 yılları arasındaki PM₁₀ konsantrasyon değerleri girdi olarak verilmiş ve 2019 yılına ait PM₁₀ konsantrasyon değerleri tahmin edilmiştir. En iyi sonuçlar 3 istasyon için de Destek Vektör Regresyonu algoritması ile elde edilmiş olup Trafik bölgesi için R²:0.85, RMSE:17.57, MAE:10.17; Hürriyet bölgesi için R²:0.73, RMSE:34.91, MAE:24.61 ve OSB bölgesi için R²:0.82, RMSE:41.71, MAE:21.62 olarak tespit edilmiştir. Ayrıca elde edilen tahmini konsantrasyon sonuçlarının mekânsal dağılımı (CBS) ve değişimi de analiz edilmiştir.

Anahtar Kelimeler: Makine öğrenmesi algoritmaları, PM₁₀, destek vektör algoritması, CBS, mekânsal analiz

Abstract

With the increase of air pollution in recent years, it is important to make an estimation of air pollution within the scope of early measures to be taken in minimizing the damage to human and environmental health. In this study, it is aimed to estimate the daily average amount of air pollution

from the PM₁₀ concentration, which is an important air pollutant, and to model the environmental and spatial air pollution. The prediction model was trained with machine learning algorithms (kNN, DVR, RF, ANN, Linear Regression) using PM₁₀ (particulate matter) concentration data measured between 2010-2018 from 3 stations in Kayseri, Central Anatolia. PM₁₀ concentration values of 3 stations in Kayseri between 2010-2018 were given as input and PM₁₀ concentration values for 2019 were estimated. The best results were obtained with Support Vector Regression algorithm for all three stations. For the Traffic region, R²: 0.85, RMSE: 17.57, MAE: 10.17; For Hurriyet region, R²: 0.73, RMSE: 34.91, MAE: 24.61 and for OSB region R²: 0.82, RMSE: 41.71, MAE: 21.62. Also, the spatial distribution and variation of the estimated concentration results were analyzed by the Geographical Information System (GIS).

Keywords: Machine learning algorithms, PM₁₀, Support vector algorithm, GIS, spatial analysis

1. Giriş

Son yıllarda gelişen sanayi, şehirleşme oranındaki artış ve ulaşımın gelişmesi ile artan hava kirliliği her geçen gün insan ve canlı yaşamını tehdit eder boyuta ulaşmaktadır.[1] Doğal olarak havada yer almayan kirleticilerin ortaya çıkması ya da zararlı boyutlarda olmayan kirletici partiküllerin (sağlığa konu olan partiküller, aerodinamik çapı 10 µm. nin altındaki partiküller) havadaki miktarındaki artışı hava kirliliğini oluşturmaktadır. En önemli hava kirletici parametreleri, Partikül Madde (PM), Kükürt dioksit (SO₂), Karbon monoksit (CO), Ozon (O₃), Azotoksitler (NO_x) ve Hidrokarbonlar (HC)'dir. Kirleticilerin kaynağı yangın, toz taşınımı gibi doğal kaynaklar olacağı gibi antropojenik faaliyetlerden olan fabrika, enerji tesisi, yakma tesisi ve inşaat kaynaklı da olabilmektedir [2].

Hava kirliliğinin özellikle insan sağlığı üzerindeki etkisinden dolayı kirliliğin belirlenmesi, kirletici düzeylerinin kontrol altında tutulması ve önlenmesi adına pek çok çalışma yapılmaktadır [3]. Hava kalitesi izleme çalışması bunlardan biridir. Bu izlemin devamlı gerçekleşmesi için belirli noktalarda izleme istasyonları kurulmuştur. Ancak bu istasyonların kurulum ve işletme maliyetleri yüksek olduğundan sınırlı sayıda bulunmaktadır. Bu durumda kişilerin hava kirliliğine gerçek maruz kalım düzeyleri saptanamamaktadır. Hava kalitesini tahmin etmek amaçlı günümüzde analitik yöntemlerin dışında yapay zekâ teknolojilerinden olan makine öğrenmesi algoritmalarına yönelik tahminleme sistemleri

giderek artmaktadır [4-6]. Bu sistem geliştirilen bilgisayar programları ile hava kirletici dağılımının matematiksel olarak tahmininin yapılması işlemine dayanmaktadır. Bu sistemler ile hava kirliliğine yönelik objektif ve daha hassas sonuçlar elde edilen çalışmalar yapılmaktadır. Bu matematiksel modelden istenen sonuçların yeni teknolojilerle belirli bir düzeyin altında tutulması ve geleceğe yönelik sağlıklı tahminlerde bulunulması gerekmektedir. Literatürde hava kalitesi parametrelerinin tahmini için çeşitli makine öğrenmesi algoritmaları kullanılarak çalışmalar yapılmıştır. Bu algoritmalara örnek olarak LASSO Regresyon [7-9], Destek Vektör Makinesi (DVM) [10-12], Rastgele Orman (RF) [13-17] ve k En Yakın Komşu algoritmaları (kNN) [18-20] sayılabilir. Hava kalitesi konsantrasyon tahmininde kullanılan yöntem ve algoritmaların başarısını etkileyen bazı etmenler bulunmaktadır. Bunlar; kullanılan verinin miktarı ve elde edildiği süre, incelenen bölgenin arazi kullanımı ve etki eden meteorolojik değişkenler olarak sıralanabilir. Literatürde farklı etmenler kullanılarak PM₁₀ konsantrasyon değişimini, yüksek doğrulukla tahmin etmeye yönelik algoritmalar kullanan çalışmalar bulunmaktadır [21-24]. Çalışmalarda tahmin değerlerinin daha iyi yorumlanabilmesi için bölgesel analizler yapılması ve mekânsal dağılımının incelenmesi önem taşımaktadır. Elde edilen verilerin haritalandırılarak görselleştirilmesi ve bunun üzerinden analizler sağlanabilmesi bir sistemi gerektirmektedir. Coğrafi Bilgi Sistemleri modelleme çalışmaları için veri sağlamada veya modelleme çalışmalarından elde edilen sonuçların

görselleştirilmesinde yaygın olarak kullanılmaktadır [25,26]. CBS, mekânsal verinin toplanması, saklanması, saklanan verilerin kullanılarak yeni verilerin elde edilmesi, sorgulanması, düzenlenmesi, analiz edilmesi ve sonuçların grafik, harita veya 3 boyutlu görüntü şeklinde görsel hâle getirilmesine dayanan bir bilgi sistemidir [27-29]. CBS ortamında sıklıkla toplanan ve depolanan mekânsal özelliği bulunan veriler (iklim değişikliği, arazi kullanım, bitki örtüsü, yer şekilleri, nüfus, yerleşme vs.) ve bu verilere ilişkin analizler makine öğrenmesi algoritmalarında kullanılmaya uygun niteliktedir [30]. CBS bu kapsamda makine öğrenimi algoritmaları için veri kaynağı sağlarken aynı zamanda sınıflandırma, kümeleme, tahminleme gibi makine öğrenme algoritmalarına yönelik sonuçlarının görselleştirilmesinde etkin olarak kullanılmaktadır. Literatürde, hava kalitesine yönelik tahmin ve modelleme çalışmalarında CBS ile oluşturulan haritalardan yararlanılarak mekânsal ve zamansal değişimin ortaya konulduğu görülmektedir [31,32].

Bu çalışmada Kayseri'deki günlük ortalama hava kirliliğinin PM₁₀ konsantrasyonu üzerinden tahminlenmesi ve hava kirliliğinin çevresel ve mekânsal modellenmesi amaçlanmıştır. PM₁₀ verisi zamansal olarak düzenli olması ile erişilebilirliği açısından bu parametrenin seçilmesinde etkilidir. Tahminleme modeli ile istasyonlar arasında birbirlerine bağımlılıklarına bakarak gelecek tahmini yapılmıştır. Bu amaçla 3 istasyondan 2010-2018 yılları arasında PM₁₀ konsantrasyonu verileri kullanılarak makine öğrenmesi algoritmaları (kNN, DVR, RF, ANN, Lineer Regresyon) ile veriler eğitilmiş her bölgenin 2019 yılına ait tahmini, gerçek verileri verilmeden diğer bölgelerin 2019 verileri kullanılarak elde edilmiştir. Bu şekilde bölgesel bağımlılık elde edilerek geleceğe yönelik tahmin yapılmıştır. Elde edilen tahmin değerlerine göre mekânsal dağılımın ve değişimin CBS yardımıyla analizi de yapılmıştır.

2. Materyal ve Yöntem

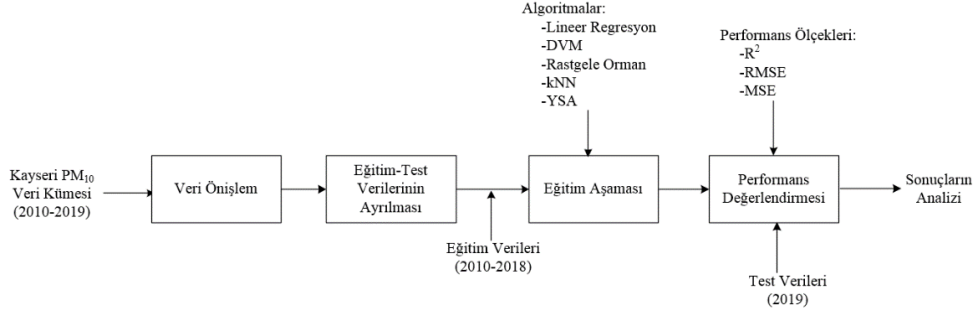
2.1. Örnekleme noktaları ve örnekleme analizleri

Türkiye'de birçok ilde olduğu gibi Kayseri il genelinde zaman zaman hava kirliliği problem oluşturmaktadır [1]. Hava kirliliğinin araştırılmasının yanında kirliliğin giderilmesi adına çözüm ortaya konulması önem arz etmektedir. Kayseri ilinde Ulusal Hava Kalitesi İzleme Ağı bünyesinde 2007 yılından itibaren anlık ve sürekli izleme ağına geçilmiştir [33]. Hava kalitesi düzeyini ölçmek amacıyla hava kalitesi Gözlem İstasyonları kurularak saatlik ortalama SO₂, PM₁₀, hava sıcaklığı, hava basıncı, nem, rüzgâr yönü ve rüzgâr hızı ölçülmektedir. Bu çalışma için PM₁₀ konsantrasyon verileri Kayseri İli hava kalitesi ölçüm istasyonları olan Trafik, Hürriyet ve OSB istasyonlarından günlük ortalama veriler olarak temin edilmiştir.

Makine öğrenmesi tekniklerinin başarılı sonuçlar verebilmesi için eksik verilerin bir önışlem aşaması olarak tamamlanması gerekmektedir. Seçilmiş olan zaman aralığında çeşitli sebeplerden dolayı istasyonlarda ölçümlerde eksiklik veya kayıp değerler bulunmaktadır. Bu eksik ve kayıp değerlerin, o niteliğe ait ortalama değerlerle doldurulması sağlanmıştır.

2.2. Makine öğrenmesi algoritmaları

Bu çalışmada makine öğrenmesi algoritmalarından Lineer Regresyon, Destek Vektör Regresyon, Rastgele Orman, k En Yakın Komşu ve Yapay Sinir Ağları algoritmaları kullanılmıştır. Kullanılan algoritmalar, Weka makine öğrenmesi yazılımı [34] kullanılarak mevcut veri üzerinde uygulanmış ve sonuçlar elde edilmiştir. Önerilen yönteme ait blok diyagram Şekil 1'de sunulmuştur. Şekil 1'de gösterildiği üzere, ilk olarak Kayseri PM₁₀ konsantrasyon verileri toplanmış ve önışlemden geçirilmiştir. Daha sonra, 2010-2018 yılları arasındaki veriler eğitim, 2019 yılındaki veriler test için ayrılmıştır. Eğitim verileri kullanılarak makine öğrenmesi algoritmaları çalıştırılmıştır. Daha sonra test verileri ve performans ölçekleri kullanılarak algoritmaların performansları değerlendirilmiştir. Son olarak algoritmaların bulmuş olduğu tahmin sonuçlarının analizi yapılmıştır.



Şekil 1. Önerilen yönteme ait blok diyagram

Çalışmada kullanılan algoritmalar farklı parametrelerle test edilmiş ve en iyi performansı gösteren parametreler kullanılmıştır.

2.2.1. Lineer Regresyon Algoritması

Lineer regresyon iki ya da daha fazla değişken arasındaki ilişkiyi ölçmek için kullanılan metottur. En basit makine öğrenme modelidir. Bu algoritma neden ve sonuç arasında lineer ilişki olup olmadığını belirler. Eğitim kümesi kurularak hipotezdeki uygun parametre seçilir ve gelecekteki farklı veri girdilerine karşılık tahmin yapılır[35]

2.2.2. Destek Vektör Makinesi Algoritması

DVM algoritması iki veya daha fazla boyuttaki noktaları birbirlerinden ayıran bir çizgi, düzlem veya hiperdüzlem üreterek sınıflandırma yapmaya çalışan ayrımcı bir sınıflandırma algoritmasıdır [36]. DVM algoritması, farklı sınıflara ait noktalar arasındaki uzaklığı maksimize edecek en uygun doğruyu bulmaya çalışarak sınıflar arasındaki ayrımı mümkün olduğu kadar iyi belirlemeye çalışmaktadır. Destek Vektör Regresyon (DVR) algoritması ise DVM algoritmasının regresyon amacıyla özelleştirilmiş türüdür.

Bu çalışmada DVR algoritması için radyal tabanlı fonksiyon kullanılmış, C ve gamma parametreleri sırasıyla 100 ve 0.01 olarak seçilmiştir.

2.2.3 Rastgele Orman Algoritması

Rastgele orman algoritmasında birden fazla karar ağacı oluşturularak bir orman meydana getirilir [37]. Rastgele orman algoritması girdi parametreler seti içerisinde rastgele parametreler seçerek bu parametreler ile çok sayıda karar ağacı oluşturmaktadır. Oluşturulan bu karar ağaçlarının sonuçları kullanılarak algoritmanın sınıflandırma çıktısı hesaplanmaktadır. Bu sayede, birbirinden bağımsız veya ilgisiz parametreler farklı karar ağaçlarıyla değerlendirildiği için rastgele orman algoritması başarılı sonuçlar üretmektedir.

Bu çalışmada rastgele orman algoritması için 200 ağaç kullanılmıştır.

2.2.4 k En Yakın Komşu Algoritması

k-En Yakın Komşu (kNN) algoritması etiketleme bilgisi bulunmayan gözlemlerin birbirlerine yakınlık derecesi göz önünde bulundurularak yapılan sınıflandırma işlemidir. Bu algoritmanın uygulama aşamasında öncelikle eğitim verisi kullanılarak sistem eğitilir. Eğitim setinde sınıflandırma bilgisi bulunan veriler mevcuttur. Eğitim seti verildikten sonra kullanıcı tarafından bir k değeri tespit edilir. Sistem öğrendiği eğitim setindeki sınıfları göz önünde bulundurarak yeni verilen gözlem verilerinden birbirine en yakın bulunan k tane komşu veriyi sınıflandırır. Verilerin birbirine olan yakınlığını ölçmek için Öklid uzaklığı, Jaccard uzaklığı, Manhattan uzaklığı gibi yöntemler kullanılır [38].

Bu çalışmada kNN algoritması için komşu sayısı olarak 20 kullanılmıştır.

2.2.5 Yapay Sinir Ağları

Yapay sinir ağları (Artificial Neural Networks – ANN), insanların beyindeki nöronlar ve bu nöronların kurduğu bağlantılardan ilham alınarak geliştirilmiş hesaplama sistemleridir [39]. Yapay sinir ağları kendilerine gösterilen örneklerden öğrenmekte ve yeni gelen örnekleri de öğrenme sistemine dahil etmektedir.

Bu çalışmada yapay sinir ağı modeli olarak her birinde 3 nöron bulunan 3 gizli katman kullanılmış, iterasyon sayısı 500 ve ağırlık optimizasyonu için Kısıtlı Hafıza Broyden-Fletcher-Goldfarb-Shanno (Limited Memory Broyden-Fletcher-Goldfarb-Shanno – LBFGS) algoritması seçilmiştir.

2.3. Performans metrikleri

Makine öğrenmesi algoritmalarının performanslarının belirlenebilmesi için çeşitli metrikler kullanılmıştır. Bu metrikler arasında en yaygın kullanılanları R^2 , RMSE ve MAE'dir. R^2 gerçek test değerleri ile algoritmaların çıkardığı sonuçların birbirine ne kadar uyumlu olduğunu gösteren bir metriktir. RMSE ve MAE metrikleri ise gerçek test değerleri ile algoritmaların çıktılarları arasındaki hatayı ifade etmektedir. Kullanılan metriklere ait formüller sırasıyla Eşitlik 1-3'te verilmiştir.

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

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

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \quad (3)$$

2.4. Mekansal analiz

2010-2019 yılları arasında istasyon verilerinin yıllık ortalamaları ve mevcut analizlere göre yapılan tahminleri incelemek amacıyla Coğrafi Bilgi Sistemlerinden (CBS) yararlanılmıştır. CBS uygulamalar için veri toplama, saklama, sorgulama, ilişkilendirme ve modelleme çalışmalarından elde edilen sonuçların görselleştirilmesinde yaygın olarak kullanılmaktadır [25,26]. Hava kirliliği değişiminin ortaya konulmasında CBS uygulamalarından mekânsal dağılımın analiz yöntemleri (Kriging ve IDW yöntemleri) sıklıkla kullanılmaktadır [40-44].

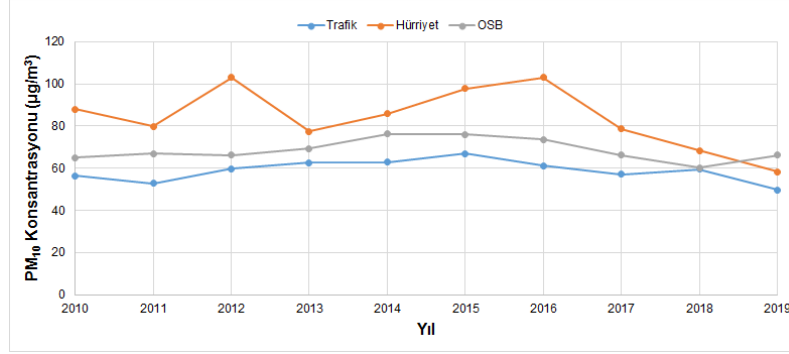
Bu çalışmada CBS yazılımlarından biri olan ArcGIS 10.6.1 yazılımı yardımıyla istasyonların yıllara göre değişen verilerinin kent üzerindeki mekânsal dağılımları IDW yöntemiyle analiz edilmiştir. IDW, birbirine yakın olan şeylerin birbirine uzak olandan daha benzer olduğunu varsayarak ölçülmemiş herhangi bir konum için tahmin konumunu çevreleyen ölçülmüş değerleri kullanarak bir değer tahmin etmektedir. IDW enterpolasyon yöntemiyle 2010, 2015 ve 2019 yıllarına ilişkin değişen veriler ortak lejand üzerinde görselleştirilerek yıllar arasındaki fark ortaya konulmuştur.

3. Bulgular ve Tartışma

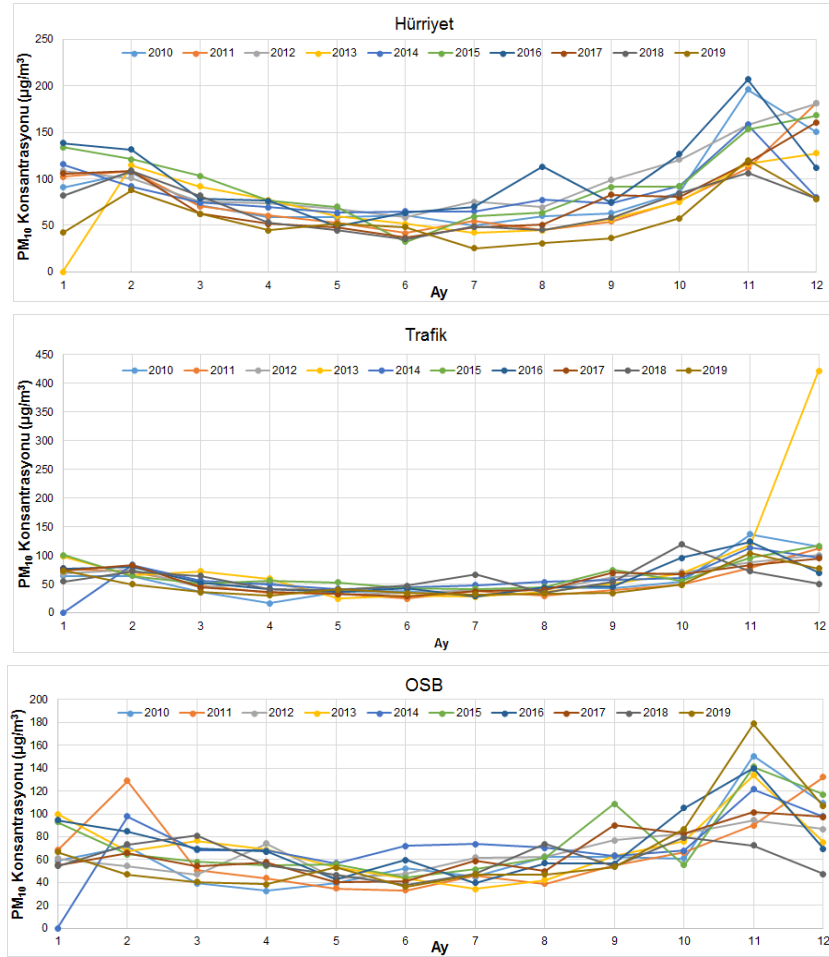
3.1. Çalışma bölgesi ve hava kirliliği verileri

Ulusal Hava Kalitesi İzleme Ağı tarafından elde edilen veriler ışığında Şekil 2'de seçilmiş olan üç bölge için (OSB, Trafik, Hurriyet) 2010-2019 yılları arası yıllık bazda PM_{10} konsantrasyonu değişimi, Şekil 3'te ise 2010-2019 yılları arasında üç bölge için PM_{10} konsantrasyonunun aylık değişimi verilmiştir.

DEÜ FMD 24(70), 65-80, 2022



Şekil 2. Hürriyet-Trafik ve OSB istasyonlarına ait 2010-2019 yılları arası PM₁₀ konsantrasyonu değişimi



Şekil 3. Hürriyet-Trafik ve OSB istasyonlarına ait 2010-2019 yılları arası PM₁₀ konsantrasyonu aylık değişimi

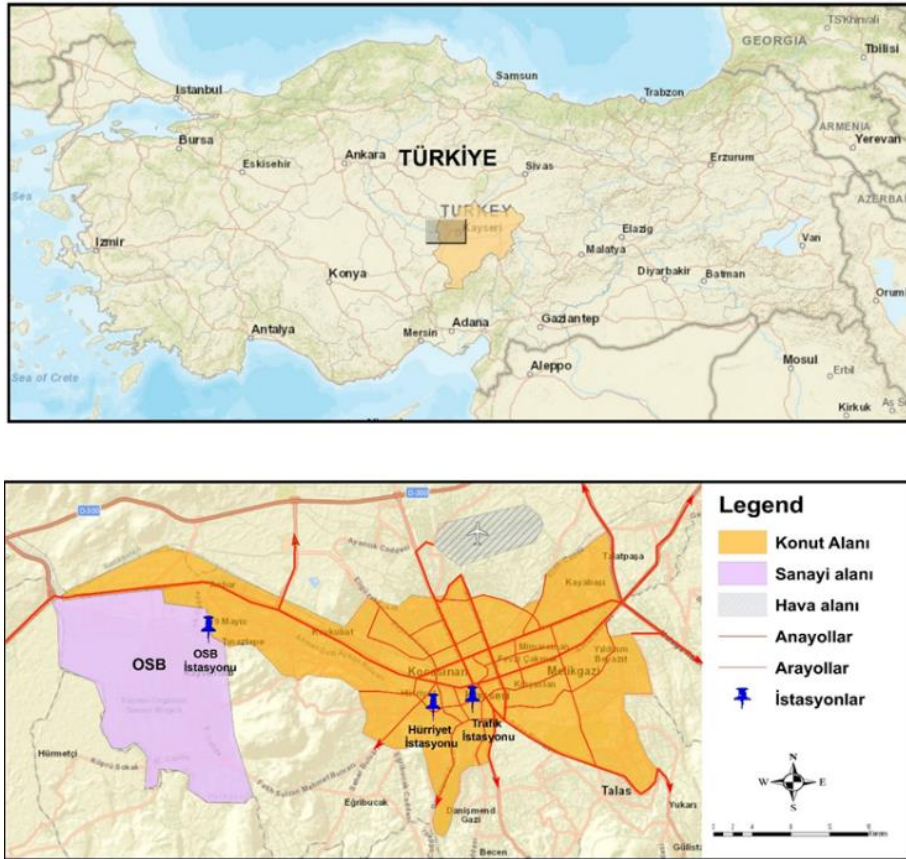
Çalışma kapsamında değerlendirilen istasyonların Şekil 2'de verilen yıllık ortalama PM₁₀ ölçüm sonuçları incelendiğinde "Hava Kalitesi Değerlendirme ve Yönetimi Yönetmeliği"nde PM₁₀ için belirtilen yıllık sınır değer olan 40 µg/m³'ü tüm istasyonlarda aştığı görülmektedir. Çalışmada değerlendirilen 3 istasyonun Şekil 3'te verilen 2019 yılına ait aylık değerlendirme yapıldığında; her 3 istasyon içinde özellikle kış ayları olan Kasım, Aralık, Ocak ve Şubat aylarında PM₁₀ değerlerinin ulusal sınır değer olan 60 µg/m³'ü aşmıştır

3.1.1 Coğrafi konumu

Kayseri, İç Anadolu Bölgesinin güneyinde Orta Kızılırmak bölümünde yer almaktadır. 37°45' ile 38°18' kuzey enlemleri ve 34°56' ile 36°58' doğu boylamları arasında yer almaktadır (Şekil 4).

16.917 km²'lik yüzölçümüne sahip Kayseri ilinin denizden yüksekliği 1054 m'dir [45].

Kent, eski kent merkezi etrafında örümcek ağı morfolojisine uygun nitelikte yayılmıştır (Şekil 4). OSB bölgesi geliştikçe kent bu bölgeye doğru lineer bir yayılma göstermektedir. Kentin hava kalitesi düzeyinde yoğun konsantrasyon miktarlarının gözlemlendiği kent merkezi ve OSB civarında üç istasyon (OSB, Trafik ve Hürriyet İstasyonu) yer almaktadır. Bu istasyonlardan Hürriyet İstasyonu konut sahası olarak kentin yoğun yerleşim alanlarının bulunduğu bölgede yer almaktadır. Trafik istasyonu ise eski kent merkezinin ve geleneksel ticaret merkezinin odağında ve günlük ihtiyaçlarına yönelik yaşayanların ve turistlerin uğrak noktası olduğu trafiğin en yoğun yaşandığı bölgede yer almaktadır.



Şekil 4. Kayseri kenti genel görünümü

3.2. Makine öğrenmesi algoritmaları

Makine öğrenmesi algoritmaları kullanılarak PM₁₀ konsantrasyonu tahmini metoduna ait sonuçlar Tablo 1'de verilmiştir. Tabloda, her bölgeye ait sonuçlar için diğer bölgelerdeki PM₁₀ konsantrasyon değerleri kullanılarak tahmin modelleri kurulmuş ve 2019 yılı PM₁₀ konsantrasyonu ile ilgili bölgedeki tahminine ait sonuçlar gösterilmiştir. Bu şekilde bir şehirdeki farklı bölgeler arasında mekânsal bağ kurularak istenen bir bölgedeki değerin tahmini yapılabilmektedir. Uygulanan algoritmalar ile veri setinde büyük veri ile çok sayıda özellik bulunuyorsa ve bunları yorumlayacak denklem ya da fonksiyon yok ise anlam çıkarmak veya çeşitli sebeplerden dolayı istasyonlarda ölçümlerin eksik olması, kayıp değerlerin olması durumunda veri setindeki kayıp değerler için o

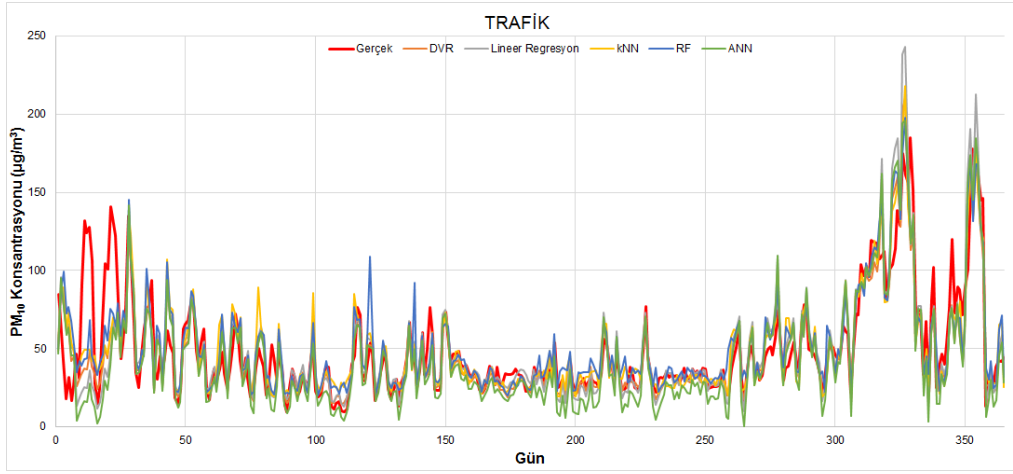
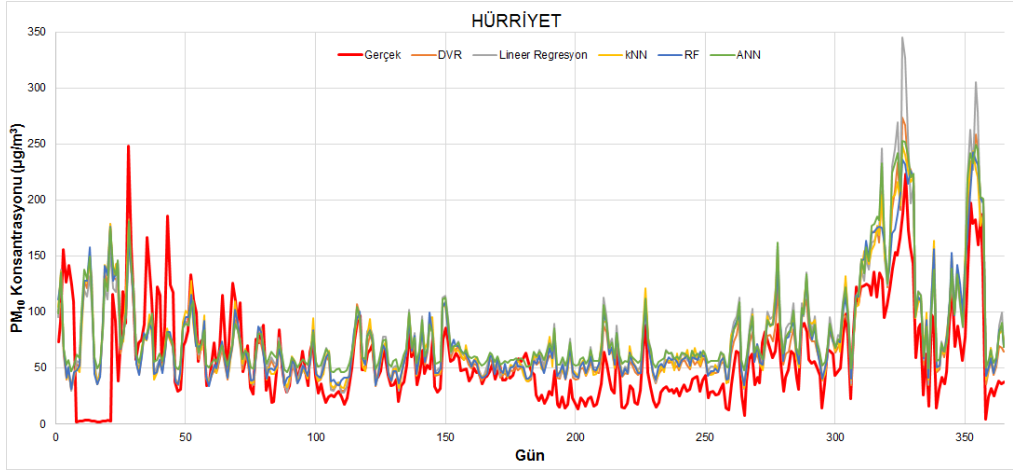
niteliğe ait ortalama değerlerle doldurulması sağlanmaktadır. Bunun için çözülmesi gereken problemi çözerken doğru veri ile algoritmayı beslemek çok önemlidir.

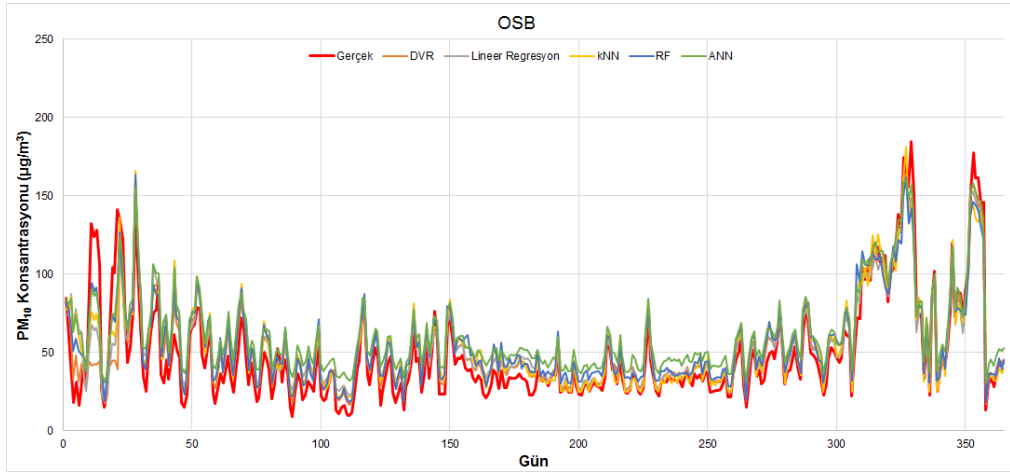
Sonuçlar incelendiğinde Trafik, Hürriyet ve OSB bölgesindeki istasyonlarda en iyi sonucun DVR algoritması tarafından üretildiği görülmektedir. En iyi sonuçlar Trafik bölgesi için DVR algoritması ile $R^2=0,85$, RMSE=17,57, MAE=10,17; Hürriyet bölgesi için DVR algoritması ile $R^2=0,73$, RMSE=34,91, MAE=24,61; OSB bölgesi için DVR algoritması ile $R^2=0,82$, MAE=21,62 ve ANN algoritması ile RMSE=41,59 olarak gözlemlenmiştir. Bununla birlikte diğer makine öğrenmesi algoritmaları da DVR algoritmasına oldukça yakın performans göstermişlerdir.

Tablo 1. Makine öğrenmesi algoritmalarıyla PM₁₀ konsantrasyon tahmini deney sonuçları

		DVR	Lineer Regresyon	kNN	RF	ANN
TRAFİK	R ²	0,8534	0,8218	0,84	0,8377	0,8281
	RMSE	17,5746	20,4584	18,1333	18,3478	21,1978
	MAE	10,1744	11,5951	11,6919	11,9915	13,5875
HÜRRİYET	R ²	0,7326	0,7212	0,7185	0,7148	0,7178
	RMSE	34,9185	39,2278	36,5521	36,0552	39,0553
	MAE	24,6184	28,9708	27,0007	26,5406	30,5481
OSB	R ²	0,8261	0,7945	0,788	0,7911	0,8068
	RMSE	41,7104	43,1906	42,6345	42,7542	41,5974
	MAE	21,6296	23,3571	22,7533	22,187	23,1481

Üç bölge için PM₁₀ konsantrasyonlarının 2019 yılı için çalıştırılan algoritmalar tarafından yapılan tahminlerin gerçek PM₁₀ değerleriyle karşılaştırmalı olarak incelenmesi için algoritmalar tarafından tahmin edilen PM₁₀ konsantrasyon değerleri ile gerçek değerlerin 2019 yılı için gün bazında karşılaştırması Şekil 5'te sunulmuştur.





Şekil 5. Algoritmalar sonucu ile gerçek değerlerin yıllık olarak karşılaştırılması

Her üç bölge için sonuçlar incelendiği zaman, algoritmaların genel olarak doğru değerler üretebildikleri görülmektedir. Lineer Regresyon, RF ve ANN algoritmalarının gerçek sonuçlara göre yüksek ve düşük değerler ürettikleri ve genel olarak gerçek değerlere en uygun sonuçların DVR algoritması tarafından üretildiği görülmektedir. OSB bölgesi için yılın ilk ayları için algoritmaların tahminlerinin genel olarak düşük kaldığı ve ilerleyen aylarda sonuçların gerçek değerlere yaklaştığı gözlemlenmektedir. Algoritmaların hepsinin birbirine benzer sonuçlar ürettiği görülmektedir. Hürriyet bölgesi için genel olarak algoritmaların ya düşük veya da yüksek sonuçlar ürettiği görülmektedir. Bu algoritmalar içerisinde en iyi sonucun DVR algoritması tarafından elde edilebildiği de görülmektedir. DVR algoritması çok çeşitli verilerde daha verimli çalışmakta ve daha iyi sonuç vermektedir [10–12]. En kötü sonucun da kNN algoritması tarafından elde edildiği görülmektedir. Bunun sebebi olarak, kNN algoritmasının eğitim ve test verileri arasındaki farklılıklara daha duyarlı olduğu ve bu nedenle elde ettiği sonuçların daha kötü olduğu söylenebilir.

Literatüre bakıldığında son yıllarda hava kirliliği üzerine çeşitli makine öğrenmesi algoritmaları kullanılarak farklı çalışmalar yapılmıştır. Fakat yapılan çalışmaların çoğunda daha az veri

kullanılarak çalışmalar yapılmış ayrıca bölgesel bir tahminden ziyade gelecek tahmini yapılmıştır. Yapılan bu çalışmaya benzer algoritmalar ve hava kirletici parametreler kullanılarak yapılan çalışmalardan biri olan Ashayeri vd., PM_{2.5} konsantrasyonlarını tahmin etmek için Chicago'daki farklı kentsel mahallelerdeki hava kalitesi verilerini kullanarak en etkili algoritmayı bulmak için dokuz farklı makine öğrenmesi algoritmasını test etmişlerdir. Sonuç olarak, yapılan bu çalışmaya benzer şekilde DVR algoritmasının test edilen en iyi algoritma olduğu sonucuna varmışlardır. DVR algoritması model performansını yıl boyunca %18,4 ve aylık veri kümeleri için %98,7'ye kadar arttırdığı belirtilmiştir [46]. Bu çalışmada tahmin edilen hava kalitesi parametresi ile aynı parametrenin kullanıldığı Choubin vd., tarafından yapılan çalışmada Barselona Eyaletindeki 75 istasyondaki yıllık PM₁₀ için tehlikeli alanları tahmin etmek adına RF, Torbalı Sınıflandırma ve Regresyon Ağaçları (Torbalı CART) ve Karışım Ayırım Analizi (MDA) gibi makine öğrenmesi algoritmaları kullanmışlardır. Sonuçlarda üç makine öğrenmesi modelinin de mükemmel sonuç (doğruluk: >%87 ve hassasiyet: >%86) verdiği belirtilmiştir. Torbalı CART ve RF modellerinin aynı ve MDA modeline göre daha yüksek performansa sahip olduğu belirtilmiştir [22]. Yapılan başka bir çalışmada, Avusturya

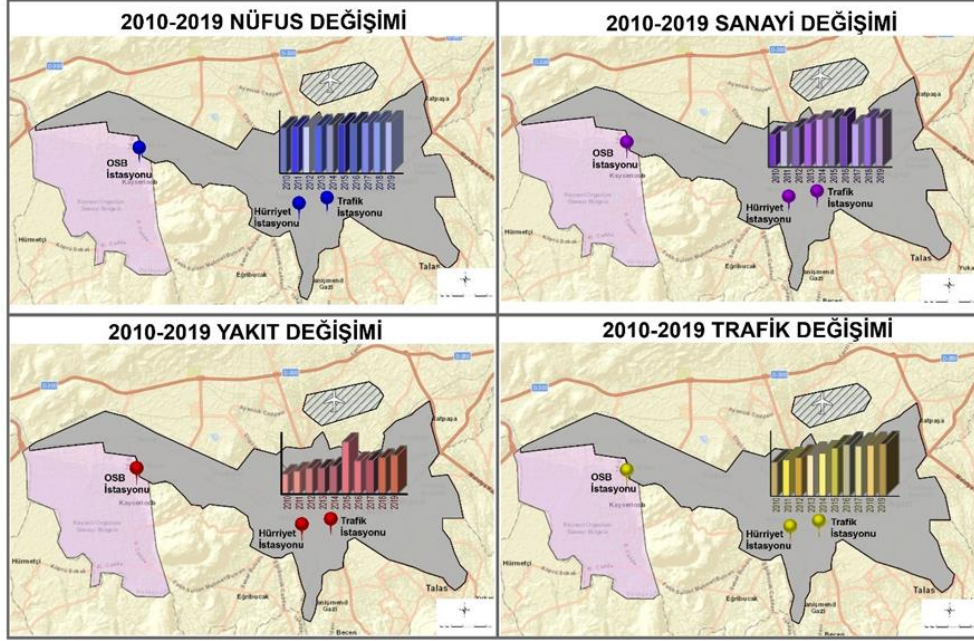
Graz'daki COVID-19 kısıtlaması sırasında hava kalitesi iyileştirmelerini analiz etmek için bir makine öğrenmesi modeli geliştirilmiştir. Graz'daki beş ölçüm istasyonundan NO₂, PM₁₀ ve O₃ konsantrasyonları seçilmiş ve Rastgele Orman (RF) modeli için hedef değişken olarak kullanılmıştır. Makine öğrenmesi modellerinin konsantrasyonları tahmin etmede yüksek performans gösterdiği sonucuna varılmıştır. Kentin kısıtlama dönemlerinde ortalama konsantrasyon düşüşleri NO₂ ve PM₁₀ için sırasıyla %36,9-41,6 ve %6,6-14,2 olduğu ayrıca O₃ için %11,6-33,8'lik bir artış olduğu tahmini yapılmıştır [47]. PM₁₀ parametresi üzerine Avustralya'nın Moranbah maden kasabasında yapılan başka bir çalışmada, 2011 yılından beri PM₁₀ hava kalitesinin standartların üzerinde seyretmesinden dolayı ölçülen PM₁₀ konsantrasyonlarını modellemek ve meteorolojik olarak normalleştirmek için makine öğrenmesi algoritmalarından Gradyan Artırılmış Regresyon ve Rastgele Orman modelleri kullanılmıştır. Sonuçta en iyi performansı Rastgele Orman algoritmasının sağladığı belirtilmiştir [16]. Masood ve Ahmed (2020) yılında Delhi'de yaptıkları çalışmada bu çalışmada da kullanılan algoritmalarından DVR ve YSA algoritmalarını kullanarak günlük PM_{2.5} konsantrasyonlarını tahmin etmişlerdir. Sonuç olarak YSA'nın (eğitim ve test için korelasyon değeri sırasıyla 0,856 ve 0,730) PM_{2.5} tahmini için DVR'den daha iyi tahmin doğruluğu gösterdiği belirtilmiştir [48].

Çevre ve Şehircilik Bakanlığı'nın 81 ildeki hava kalitesi izleme istasyonlarından aldığı verilerle oranını tespit ettiği havadaki partikül maddeler (PM₁₀) "akciğerlere ulaşarak iltihaplanmaya ya da insanları çok olumsuz etkileyecek kalp ve

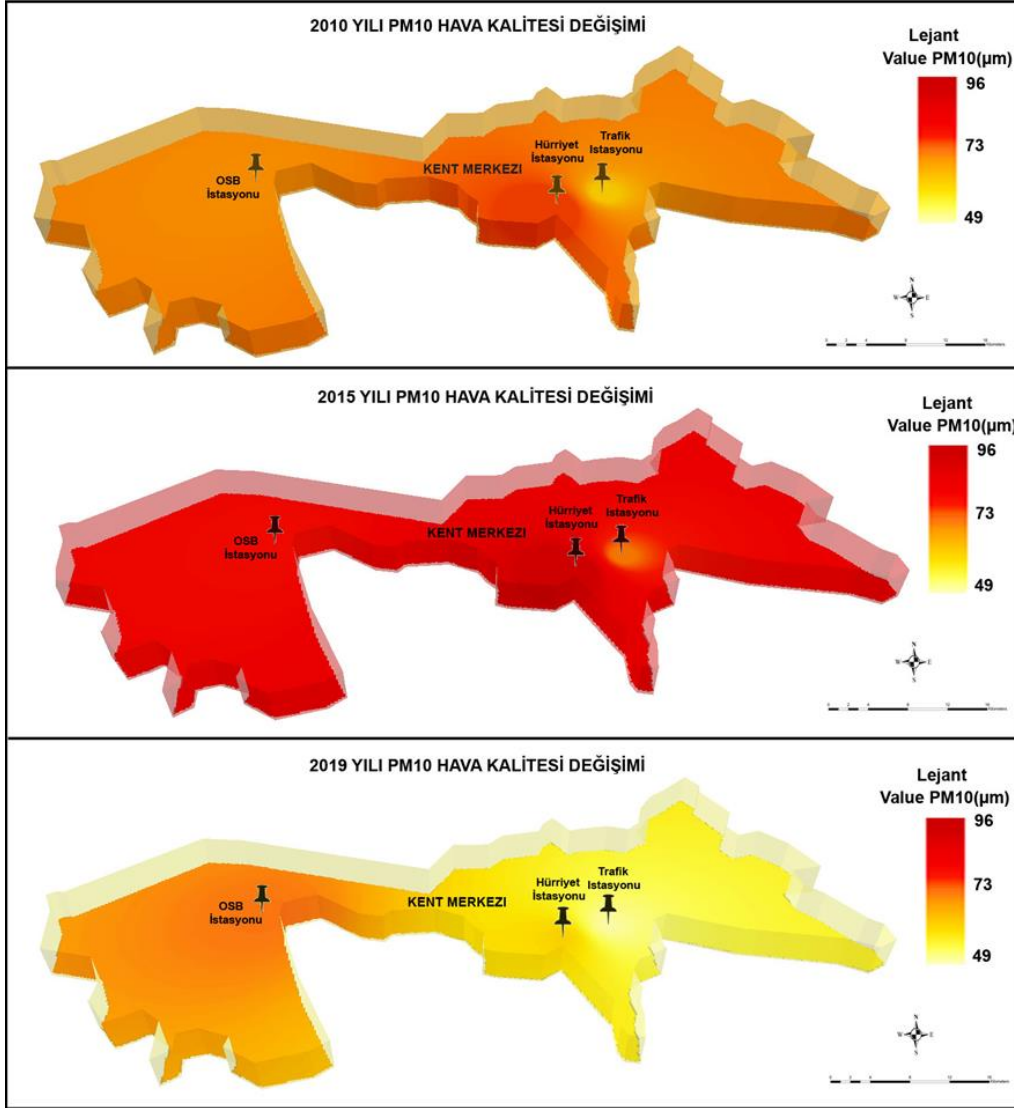
akciğer hastalıklarına neden" olabilmektedir. Dolayısıyla bu partiküllerin havadaki oranı, kirliliğinin en önemli göstergelerinden biri olarak kabul edilmektedir. Ancak Türkiye'nin mevcut yönetmelik ve uygulamalarda hava kalitesi ölçümünde sağlıklı seviye olarak kabul ettiği değerler(SO₂:250 µg/m³, PM₁₀:100 µg/m³, Avrupa Birliği (AB) için (SO₂:125 µg/m³, PM₁₀:50 µg/m³) ve DSÖ'nün kabul ettiği sınır değerlerden (SO₂: 500 µg/m³, PM₁₀: 50 µg/m³) çok daha yüksek olduğu belirtilmektedir [49].

3.3. Mekansal Analiz Sonuçları

Çalışmada PM₁₀ değerine etkileyen unsurlar grafik ve haritalar ile analiz edilmiştir (Şekil 6). Bu analizler 2010-2019 yılı kent merkezinin nüfus yoğunluğu, sanayi sektörü, yakıt kullanımı ve trafik yoğunluğunun değişimi olarak ele alınmıştır. Kent merkezinin nüfus yoğunluğunun 2010 yılından 2019 yılına kadar giderek arttığı tespit edilmiştir. Büyükşehir statüsündeki bir kent için beklenen nüfus artışına yönelik bu değişimin hava kalitesinin değişimine etkisi istasyon verileri ile değerlendirilmiştir. Sanayi sektöründe 2016 yılına kadar sanayi tesis sayısında artış gözlenirken daha sonraki yıllarda yeni açılan sanayi tesisi sayısında bir azalma tespit edilmiştir. Yakıt kullanımı değişiminde ise 2015 yılı yakıt kullanımının en fazla arttığı yıl olarak belirlenmiştir. Taşıt kullanımında ise 2015 ve 2016 yılında diğer yıllara oranla en yüksek artış gerçekleşmiştir.



Şekil 6. Kayseri kent merkezinin yıllara göre mevcut durum analizi



Şekil 7. 2010, 2015 ve 2019 yıllarına göre PM₁₀ hava kalitesi değişiminin mekânsal dağılımı

Şekil 7'ye göre mekânsal değişimi izlenen 2010, 2015 ve 2019 yılları arasında PM₁₀ konsantrasyon miktarının en yüksek olan yılın 2015 yılı olduğu tespit edilmiştir. Analiz sonuçları konumsal değişim ve mevcut durum analizlerine göre incelenmiştir. Buna göre 2010 ve 2015 yılında kentin en eski yerleşim noktası olan Hürriyet ve Trafik İstasyonları yoğun bir hava kirliliği içinde bulunmaktadır. Hürriyet istasyonu bu yıllarda diğer tüm istasyonlardan daha fazla hava kirliliği etkisi altındadır. 2019

yılında ise bu yoğunluğun kent merkezinden OSB'ne doğru konum değiştirdiği görülmektedir. Bunun nedeni 2016 yılından sonra azalan yakıt ve taşıt kullanımının bu bölgede yer alan istasyonlardaki (Hürriyet ve Trafik) konsantrasyon miktarlarında azalmaya neden olmasıdır. Ayrıca 2016 yılından sonra kentte sanayi sektöründe yıllara göre yeni bir artışa geçilmiş olması OSB bölgesinde hava kirliliğini artırmış ve bu nedenle hava kirliliğinin odak

noktası kent merkezinden OSB bölgesine konum değiştirmiştir.

Diğer bir etken olan Nüfus Artışı ise kent merkezinin doygunluğa ulaşmış olması nedeniyle kent merkezi dışında özellikle hava alanı ve kentin doğu bölgesine yönelmiştir. Bu nedenle nüfus artışının doğrudan mevcut istasyonlardaki etkisi diğer unsurlara göre daha azdır.

Yapılan bu çalışmaya benzer olarak Tayvan'da Chang vd., (2020) tarafından yapılan çalışmada görünür iki boyutlu topolojik haritada uzun süreli bölgesel PM_{2.5} konsantrasyon tahmini için bir makine öğrenmesi modeli oluşturulmuştur. 25 istasyondan çeşitli zaman ölçeklerinde (yıllık, mevsimsel, saatlik) PM_{2.5} konsantrasyon verilerini kullanmışlardır. Sonuç olarak nüfus ve trafik yoğunluğu ve mevsimsel değişikliklerin PM_{2.5} konsantrasyon değişiminde etkisi olduğunu ortaya koymuşlardır. Makine öğrenmesi tekniklerinin PM_{2.5} özelliklerini mükemmel özetleyebildiği ve hava kalitesi tahmin doğruluğunu arttırabildiğini belirtmişlerdir [50].

4.Sonuç

Yapılan çalışma kapsamında;

- Hava kirliliği değişimini makine öğrenim algoritmalarına (kNN, DVR, RF, ANN, Lineer Regresyon) göre yüksek doğrulukla tahmin etme,
- Elde edilen tahmin değerlerine göre mekânsal dağılımın CBS yardımıyla analizi ve mekânda alınacak hava kirliliğinin iyileştirilmesine ilişkin çevresel önlemlerin ve arazi kullanım stratejilerinin oluşturulması gerçekleştirilmiştir.

Bu kapsamda Kayseri İli Trafik, OSB, Hürriyet bölgeleri için 2010-2018 yılları arasındaki yıllık PM₁₀ konsantrasyon verileri kullanılarak 2019 yılına ait hava kirliliği tahmini yapılmıştır. 3 istasyondan 2010-2018 yılları arasında PM₁₀ konsantrasyonu verileri kullanılarak makine öğrenmesi algoritmaları (kNN, DVR, RF, ANN, Lineer Regresyon) ile veriler eğitilmiş her bölgenin 2019 yılına ait tahmini, gerçek verileri verilmeden diğer bölgelerin 2019 verileri kullanılarak elde edilmiştir. Bu şekilde bölgesel bağımlılık elde edilerek geleceğe yönelik tahmin yapılmıştır. Bu amaçla makine öğrenmesi algoritmaları olan DVR, Rastgele Orman, kNN, ANN algoritmaları tarafından üretilmiş olan

tahmin modelleri kullanılmış ve model sonuçları karşılaştırılmıştır. Elde edilen sonuçlara göre hava kirletici tahmini için en yüksek performansı üç istasyon için de DVR algoritması vermiştir. Modeller karşılaştırıldığında ise en iyi performans Trafik bölgesi için DVR algoritması ($R^2=0.85$, $RMSE=17.57$, $MAE=10.17$) ile sağlanmıştır. Ayrıca algoritmaların tahmin sonuçları ile gerçek değerler karşılaştırıldığı zaman, tüm algoritmaların genel olarak iyi sonuçlar üretebildikleri ve gerçek değerlere yakın tahminlerde buldukları görülmüştür. Uygulanan algoritmalar ile veri setinde büyük veri ile birlikte çok sayıda özellik bulunuyorsa ve bunları yorumlayacak denklem ya da fonksiyon yok ise anlam çıkarmak veya çeşitli sebeplerden dolayı istasyonlarda ölçümlerin eksik olması, kayıp değerlerin olması durumunda veri setindeki kayıp değerler için o niteliğe ait ortalama değerlerle doldurulması sağlanmaktadır. Bunun için çözülmesi gereken problemi çözerken doğru veri ile algoritmayı beslemek çok önemlidir.

Algoritmalar ile elde edilen tahmini konsantrasyon sonuçlarının CBS ile mekânsal analizi yapılarak konumsal değişim ve mevcut durum analizlerine göre inceleme yapılmıştır. Buna göre; 2010 ve 2015 yılında kentin en eski yerleşim noktası, nüfus ve taşıt yoğunluğu olan Hürriyet ve Trafik İstasyonları yoğun bir hava kirliliği içinde bulunmaktadır. 2016 yılından sonra ve 2019 yılında kentte sanayi sektöründe yıllara göre artışa geçilmiş olması OSB bölgesinde hava kirliliğini arttırmış ve bu nedenle hava kirliliğinin odak noktasının kent merkezinden OSB bölgesine konum değiştirmiştir.

Kayseri İl'inde hava kirliliği kaynakları özellikle kış aylarında kırsal kesimin konutlarda ısınma amaçlı olarak yakılan kalitesiz kömür, motorlu taşıtlardan kontrolsüz olarak yayılan salımlar ve OSB bölgesinde faaliyet gösteren sanayi tesislerinin çıkardığı emisyonlardır. Özellikle kırsal kesimlerde yoğun olarak kullanılan fosil yakıtların yerine doğalgaz ya da yenilenebilir enerji kaynaklarının kullanımının teşvik edilmesi, kentleşme ve düzenli yapılaşmanın bir plan dahilinde yapılması, yeşil alanların arttırılması, sanayi tesislerinde emisyon değerlerine sıkı denetimler getirilmesi bu düzenlemeler arasında sıralanabilir. Yapılan bu çalışma, yüksek yoğunluklu kentlerin hava kalitesinin makine öğrenmesi algoritmaları ile

modellenmesi ve hava kalitesini etkileyen arazi kullanımının sürdürülebilir planlanmasına ilişkin öneriler sunması ile literatüre katkı sağlamaktadır.

Kaynaklar

- [1] A. Alkan, Hava Kirliliğinin Ciddi Boyutlara Ulaştığı Kentlere Bir Örnek : An Example of Cities Where Air Pollution Has Reached Serious Dimensions: Siirt, (2018) 641–666.
- [2] Y. DOKUZ, A. BOZDAĞ, B. GÖKÇEK, HavKalitesiParametreleriniTahmini Ve Mekansal DağılımıçinMakineÖğrenmesiYöntemlerinin Kullanılması, Ömer Halisdemir Üniversitesi Mühendislik Bilim. Derg. 9 (2020) 37–47. <https://doi.org/10.28948/ngumuh.654092>.
- [3] G.Ç. SÜMER, Hava Kirliliği Kontrolü: Türkiye’de Hava Kirliliğini Önlemeye Yönelik Yasal Düzenlemele-rin ve Örgütlenmelerin İncelenmesi, Uluslararası İktisadi ve İdari İncelemeler Derg. 13 (2014) 37. <https://doi.org/10.18092/ijeas.51643>.
- [4] A. Rahimpour, J. Amanollahi, C.G. Tzanis, Air quality data series estimation based on machine learning approaches for urban environments, Air Qual. Atmos. Heal. 14 (2021) 191–201. <https://doi.org/10.1007/s11869-020-00925-4>.
- [5] C. Amuthadevi, D.S. Vijayan, V. Ramachandran, Development of air quality monitoring (AQM) models using different machine learning approaches, J. Ambient Intell. Humaniz. Comput. (2021). <https://doi.org/10.1007/s12652-020-02724-2>.
- [6] H. Tian, Y. Zhao, M. Luo, Q. He, Y. Han, Z. Zeng, Estimating PM2.5 from multisource data: A comparison of different machine learning models in the Pearl River Delta of China, Urban Clim. 35 (2021) 100740. <https://doi.org/10.1016/j.uclim.2020.100740>.
- [7] Y. Son, Á.R. Osornio-vargas, M.S.O. Neill, P. Hystad, J.L. Texcalac-sangrador, P. Ohman-strickland, Q. Meng, S. Schwander, Land use regression models to assess air pollution exposure in Mexico City using finer spatial and temporal input parameters, Sci. Total Environ. 639 (2018) 40–48. <https://doi.org/10.1016/j.scitotenv.2018.05.144>.
- [8] Y. Xu, H. Liu, Z. Duan, A novel hybrid model for multi-step daily AQI forecasting driven by air pollution big data, Air Qual. Atmos. Heal. 13 (2020) 197–207. <https://doi.org/10.1007/s11869-020-00795-w>.
- [9] S. Van Roode, J.J. Ruiz-Aguilar, J. González-Enrique, I.J. Turias, An artificial neural network ensemble approach to generate air pollution maps, Environ. Monit. Assess. 191 (2019). <https://doi.org/10.1007/s10661-019-7901-6>.
- [10] J.Y. Yang, W.F. Ip, C.M. Vong, P.K. Wong, Effect of choice of kernel in support vector machines on ambient air pollution forecasting, Proc. 2011 Int. Conf. Syst. Sci. Eng. ICSSE 2011. (2011) 552–557. <https://doi.org/10.1109/ICSSE.2011.5961964>.
- [11] W. Lu, W. Wang, A.Y.T. Leung, S.M. Lo, R.K.K. Yuen, Z. Xu, H. Fan, Air pollutant parameter forecasting using support vector machines, Proc. Int. Jt. Conf. Neural Networks. 1 (2002) 630–635. <https://doi.org/10.1109/ijcnn.2002.1005545>.
- [12] A. Sotomayor-Olmedo, M.A. Aceves-Fernández, E. Gorrostieta-Hurtado, C. Pedraza-Ortega, J.M. Ramos-Arreguín, J.E. Vargas-Soto, Forecast Urban Air Pollution in Mexico City by Using Support Vector Machines: A Kernel Performance Approach, Int. J. Intell. Sci. 03 (2013) 126–135. <https://doi.org/10.4236/ijis.2013.33014>.
- [13] H. Sun, D. Gui, B. Yan, Y. Liu, W. Liao, Y. Zhu, C. Lu, Assessing the potential of random forest method for estimating solar radiation using air pollution index, Energy Convers. Manag. 119 (2016) 121–129. <https://doi.org/10.1016/j.enconman.2016.04.051>.
- [14] D. Kumar, ScienceDirect ScienceDirect Evolving Differential evolution method with random forest for Evolving Differential evolution method with random forest for prediction of Air Pollution prediction of Air Pollution, Procedia Comput. Sci. 132 (2018) 824–833. <https://doi.org/10.1016/j.procs.2018.05.094>.
- [15] C. Gariazzo, G. Carlino, C. Silibello, M. Renzi, S. Finardi, N. Pepe, P. Radice, F. Forastiere, P. Michelozzi, G. Viegi, M. Stafoggia, A multi-city air pollution population exposure study: Combined use of chemical-transport and random-Forest models with dynamic population data, Sci. Total Environ. 724 (2020) 138102. <https://doi.org/10.1016/j.scitotenv.2020.138102>.
- [16] M.D. Mallet, Meteorological normalisation of PM10 using machine learning reveals distinct increases of nearby source emissions in the Australian mining town of Moranbah, Atmos. Pollut. Res. (2020) 0–1. <https://doi.org/10.1016/j.apr.2020.08.001>.
- [17] V. Chaudhary, A. Deshbhratar, V. Kumar, D. Paul, Time Series Based LSTM Model to Predict Air Pollutant ’ s Concentration for Prominent Cities in India, Udm’18. (2018).
- [18] C. Wen, S. Liu, X. Yao, L. Peng, X. Li, Y. Hu, T. Chi, A novel spatiotemporal convolutional long short-term neural network for air pollution prediction, Sci. Total Environ. 654 (2019) 1091–1099. <https://doi.org/10.1016/j.scitotenv.2018.11.086>.
- [19] Z. Qin, C. Cen, X. Guo, Prediction of Air Quality Based on KNN-LSTM, J. Phys. Conf. Ser. 1237 (2019). <https://doi.org/10.1088/1742-6596/1237/4/042030>.
- [20] A. Marjovi, A. Arfire, A. Martinoli, High resolution air pollution maps in urban environments using mobile sensor networks, Proc. - IEEE Int. Conf. Distrib. Comput. Sens. Syst. DCOSS 2015. (2015) 11–20. <https://doi.org/10.1109/DCOSS.2015.32>.
- [21] F. Taşpınar, Improving artificial neural network model predictions of daily average PM10 concentrations by applying principle component analysis and implementing seasonal models, J. Air Waste Manag. Assoc. 65 (2015) 800–809. <https://doi.org/10.1080/10962247.2015.1019652>.
- [22] B. Choubin, M. Abdolshahnejad, E. Moradi, X. Querol, A. Mosavi, S. Shamsirband, P. Ghamisi, Spatial hazard assessment of the PM10 using machine learning models in Barcelona, Spain, Sci. Total Environ. 701 (2020) 134474. <https://doi.org/10.1016/j.scitotenv.2019.134474>.
- [23] E. Sharma, R.C. Deo, R. Prasad, A. V. Parisi, A hybrid air quality early-warning framework: An hourly forecasting model with online sequential extreme learning machines and empirical mode decomposition algorithms, Sci. Total Environ. 709 (2020) 135934. <https://doi.org/10.1016/j.scitotenv.2019.135934>.

- [24] M. Stafoggia, T. Bellander, S. Bucci, M. Davoli, K. De Hoogh, F. De Donato, C. Gariazzo, A. Lyapustin, P. Michelozzi, M. Renzi, M. Scortichini, A. Shtein, G. Viegi, I. Kloog, J. Schwartz, Estimation of daily PM₁₀ and PM_{2.5} concentrations in Italy, 2013 – 2015, using a spatiotemporal land-use random-forest model, *Environ. Int.* 124 (2019) 170–179. <https://doi.org/10.1016/j.envint.2019.01.016>.
- [25] A. Ertürk, A. Ekdal, M. Gurel, K. Yuceil, A. Tanik, Use of mathematical models to estimate the effect of nutrient loadings on small streams, *Fresen. Environ. Bull.* 13 (2004) 1361–1370.
- [26] M. ATAOL, Burdur Gölü'nde Seviye Değişimleri, *Co. 8* (2010) 077–092. https://doi.org/10.1501/cogbil_0000000105.
- [27] T. Yomraloğlu, *Coğrafi Bilgi Sistemleri: Temel Kavramlar ve Uygulamalar, Seçil Ofset, İstanbul, 2000.*
- [28] S. İncekara, M. Karakuyu, A. Karaburun, ORTAÖĞRETİM Coğrafya Derslerinde Yaparak Öğrenmeye Bir Örnek: Coğrafi Bilgi Sistemlerinin Proje Temelli Öğrenimde Kullanılması, *Elektron. Sos. Bilim. Derg.* 8 (2009) 305–322. <https://doi.org/10.17755/esosder.76302>.
- [29] E. Kapluhan, COĞRAFİ BİLGİ SİSTEMLERİ'NİN (CBS) COĞRAFYA ÖĞRETİMİNDE KULLANIMININ ÖNEMİ VE GEREKLİLİĞİ, *MARMARA COĞRAFYA DERGİSİ.* 29 (2014) 34–59.
- [30] C. Güney, Mekansal zekanın getirdiği paradigma değişimi, *J. Geod. Geoinf.* 6 (2019) 128–142. <https://doi.org/10.9733/jgg.2019r0006.t>.
- [31] W.J. Requia, B.A. Coull, P. Koutrakis, Evaluation of predictive capabilities of ordinary geostatistical interpolation, hybrid interpolation, and machine learning methods for estimating PM_{2.5} constituents over space, *Environ. Res.* 175 (2019) 421–433. <https://doi.org/10.1016/j.envres.2019.05.025>.
- [32] C.C. Lim, H. Kim, M.J.R. Vilcassim, G.D. Thurston, T. Gordon, L. Chen, K. Lee, M. Heimbinder, S. Kim, Mapping urban air quality using mobile sampling with low-cost sensors and machine learning in Seoul, South Korea, *Environ. Int.* 131 (2019) 105022. <https://doi.org/10.1016/j.envint.2019.105022>.
- [33] ÇED ve Çevre İzinleri Şube Müdürlüğü, *Çevre ve Şehircilik İl Müdürlüğü Kayseri İli 2018 Yılı Çevre Durum Raporu*, (2019).
- [34] G.H. Marks Hall, *WEKA: Practical Machine Learning Tools and Techniques with Java Implementations*, (1994). <https://researchcommons.waikato.ac.nz/bitstream/handle/10289/1040/uow-cs-wp-1999-11.pdf?sequence=1&isAllowed=y>.
- [35] M. Yılmaz, R. Kaniç, M. Erdal, S. Yıldız, A. Bakış, K.H. Okulu, İ.M. Bölümü, Bina Bakım Onarım Ödeneklerinin Etkin Kullanımı Maksadıyla İhale Bedelini Etkileyen Faktörlerin Yapay Sinir Ağları ve Lineer Regresyon Yöntemleri ile Belirlenmesi Determination of The Factors Effecting The Tender Price by way of Artificial Neural Networks, 19 (2016) 461–470.
- [36] H. Drucker, C.J.C. Burges, L. Kaufman, A. Smola, V. Vapnik, W. Long, B. Nj, *Support Vector Regression Machines* w) tw , 1 (n.d.).
- [37] L.E.O. Breiman, *Random Forests*, (2001) 5–32.
- [38] B. Lantz, *Machine Learning with R*, Packt Publishing Ltd, UK, 2013.
- [39] A. Jain, J. Mao, *Artificial Neural Networks: A Tutorial*, *Comput. (Long Beach, Calif.)* 29 (1996) 31–44.
- [40] J.Y. Son, M.L. Bell, J.T. Lee, Individual exposure to air pollution and lung function in Korea: Spatial analysis using multiple exposure approaches, *Environ. Res.* 110 (2010) 739–749. <https://doi.org/10.1016/j.envres.2010.08.003>.
- [41] M. Wu, J. Huang, N. Liu, R. Ma, Y. Wang, L. Zhang, A Hybrid Air Pollution Reconstruction by Adaptive Interpolation Method, in: *Proc. 16th ACM Conf. Embed. Networked Sens. Syst.*, 2018: pp. 408–409.
- [42] W.J. Requia, M.D. Adams, A. Arain, S. Papatheodorou, P. Koutrakis, M. Mahmoud, *Global Association of Air Pollution and Cardiorespiratory Diseases: A Systematic Review, Meta-Analysis, and Investigation of Modifier Variables*, *Am. J. Public Health.* 108 (2018) S123–S130. <https://doi.org/10.2105/AJPH.2017.303839>.
- [43] K. Shukla, P. Kumar, G.S. Mann, M. Khare, Mapping spatial distribution of particulate matter using Kriging and Inverse Distance Weighting at supersites of megacity Delhi, *Sustain. Cities Soc.* 54 (2020) 101997. <https://doi.org/10.1016/j.scs.2019.101997>.
- [44] A. Bozdağ, Ö.B. Gökçek, Spatial prediction of PM₁₀ concentration using machine learning algorithms in Ankara, Turkey, *Environ. Pollut.* 263 (2020). <https://doi.org/10.1016/j.envpol.2020.114635>.
- [45] Url, Kayseri İl Kültür ve Turizm Müdürlüğü, (2020). <https://kayseri.ktb.gov.tr/TR-54966/cografyayapi.html>.
- [46] M. Ashayeri, N. Abbasabadi, M. Heidarinejad, B. Stephens, Predicting intraurban PM_{2.5} concentrations using enhanced machine learning approaches and incorporating human activity patterns, *Environ. Res.* (2020) 110423. <https://doi.org/10.1016/j.envres.2020.110423>.
- [47] M. Lovrić, K. Pavlović, M. Vuković, S.K. Grange, M. Haberl, R. Kern, Understanding the true effects of the COVID-19 lockdown on air pollution by means of machine learning, *Environ. Pollut.* (2020). <https://doi.org/10.1016/j.envpol.2020.115900>.
- [48] S. Kumar, S. Mishra, S.K. Singh, A machine learning-based model to estimate PM_{2.5} concentration levels in Delhi's atmosphere, *Heliyon.* 6 (2020) e05618. <https://doi.org/10.1016/j.heliyon.2020.e05618>.
- [49] A.O. Karababa, B. ATLI, Ç. Çağlayan, G. Varol, G. Ersoy, F. Gacal, N. Etiler, P. Özfirat, S. Ayta, *Kara Rapor 2020: Hava Kirliliği ve Sağlık Etkileri, Temiz Hava Hakkı Platformu.* (2020). <https://www.temizhavahakki.com/kara-rapor/>.
- [50] F.J. Chang, L.C. Chang, C.C. Kang, Y.S. Wang, A. Huang, Explore spatio-temporal PM_{2.5} features in northern Taiwan using machine learning techniques, *Sci. Total Environ.* 736 (2020) 139656. <https://doi.org/10.1016/j.scitotenv.2020.139656>.