



Geomatik

<https://dergipark.org.tr/tr/pub/geomatik>

e-ISSN 2564-6761



PM10, SO₂ hava kirleticilerinin çoklu doğrusal regresyon ve yapay sinir ağları ile sezonsal tahmini

Burak Kotan*¹, Arzu Erener²

¹Kocaeli Üniversitesi, Fen Bilimleri Enstitüsü, Jeodezi ve Jeoinformasyon Mühendisliği, Kocaeli, Türkiye

²Kocaeli Üniversitesi, Mühendislik Fakültesi, Harita Mühendisliği, Kocaeli, Türkiye

Anahtar Kelimeler

Coğrafi Bilgi Sistemleri
Çoklu Doğrusal Regresyon
Hava Kirliliği
Tahmin Modeli
Yapay Sinir Ağı

Araştırma Makalesi

Geliş: 06.08.2022
Revize: 13.12.2022
Kabul: 19.12.2022
Yayınlanma: 15.02.2023



Öz

Nüfus artışıyla birlikte kentleşme, sanayileşme ve taşıt sayısındaki artışlar hava kirliliğinin artmasına sebep olmaktadır. Hava kirliliği insan ve çevre sağlığına zarar vermektedir. Bu nedenle bu çalışmada hava kirliliğinin önlenmesi, tedbirlerin alınması ve planlamaların yapılabilmesi için 1 yıl önceden hava kirliliğinin tahmin edilmesi amaçlanmıştır. Hava kirliliği parametrelerinden olan PM10 ve SO₂ parametrelerinin mevsimsel ortalamalarının tahmin edilmesi için çoklu doğrusal regresyon analizi ve yapay sinir ağı yöntemleri kullanılmıştır. Kocaeli, Türkiye'nin sayılı sanayi bölgelerinden olması dolayısı ile çalışma alanı olarak seçilmiştir. Tahmin modellerinde meteorolojik veriler, kirletici konsantrasyonları, kentleşme, sanayileşme, topografik ve demografik veriler kullanılmıştır. Çalışmada enterpolasyon yöntemleri ile verilerin sürekliliğinin sağlanmasında, veri setlerinin oluşturulmasında ve haritaların yapılmasında coğrafi bilgi sistemleri kullanılmıştır. Çalışmada 2008 ile 2017 yılları arasındaki veriler ile tahmin modelleri oluşturulmuştur. Oluşturulan modellere 2018 verileri sunularak 2019 yılına ait tahmin değerleri elde edilmiştir. Elde edilen tahmin değerleri 2019 yılına ait gerçek değerler ile karşılaştırılmıştır. Model performansları ortalama mutlak yüzde hata (OMYH) değerine göre değerlendirilmiştir. Çalışma sonucunda yapay sinir ağının çoklu regresyon analizine göre performansının daha yüksek olduğu tespit edilmiştir. Ayrıca, mekânsal verilerin hava kirliliği tahmin modellerinde coğrafi bilgi sistemleri ile kullanılabilirliği gösterilmiştir.

Seasonal forecasting of PM10, SO₂ air pollutants with multiple linear regression and artificial neural networks

Keywords

Geographic Information Systems
Multiple Linear Regression
Air Pollution
Prediction Model
Artificial Neural Network

Research Article

Received: 06.08.2022
Revised: 13.12.2022
Accepted: 19.12.2022
Published: 15.02.2023

Abstract

Increases in air pollution are caused by population growth, urbanization, industrialization, and a rise in the number of cars. Human and environmental health are harmed by air pollution. According to this scope, the goal is to forecast air pollution one year ahead of time in order to prevent pollution, take safeguards, and plan ahead. The seasonal averages of PM10 and SO₂ values, which are air pollution metrics, were calculated using multiple linear regression analysis and artificial neural network methods. Because Kocaeli is one of Turkey's few industrial regions, it was chosen as the study area. In the forecast models, meteorological data, pollution concentrations, urbanization, industrialization, topography, and demographic data were all employed. Geographic information systems were employed in the study to assure data continuity using interpolation methods, construct data sets, and create maps. Prediction models were built using data from 2008 to 2017 in the study. The forecast values for 2019 were produced by feeding the 2018 data to the constructed models. The anticipated values were compared to the actual values for the year 2019. The mean absolute percent error (MAPE) was used to assess model performance. As a result of the research, it was discovered that the artificial neural network outperformed the multiple regression analysis. Furthermore, the use of spatial data in air pollution forecasting models using geographic information systems has been established.

*Sorumlu Yazar

*bkotan19@gmail.com) ORCID ID 0000-0003-2105-7498
(arzu.erenner@kocaeli.edu.tr) ORCID ID 0000-0002-9168-4254

Kaynak Göster (APA)

Kotan, B., & Erener, A. (2023). PM10, SO₂ hava kirleticilerinin çoklu doğrusal regresyon ve yapay sinir ağları ile sezonsal tahmini. *Geomatik, 8(2)*, 163-179

1. Giriş

Hava, insanların ve birçok canlıının solunumunda en temel unsurlardan biri olduğu için hava kirliliği de tüm insanları ilgilendiren bir konu haline gelmektedir. Hava kirliliğine sebep olan kirletici konsantrasyonlarında meydana gelen artışlar da insan, canlı ve çevre sağlığını büyük oranda etkilemektedir. Dünya Sağlık Örgütü'ne (DSÖ) göre her yıl dünyada yaklaşık 7 milyon insanın ölümüne hava kirliliğinin sebep olduğu söylenmekte ve hava kirliliğinin akciğer kanserine, kalp hastalıklarına ve akut solunum yolu enfeksiyonlarına bağlı olarak ölüm sayılarının artmasına sebebiyet verdiği söylenmektedir (Du ve ark., 2021).

Dünya nüfusunun çoğunluğu şehirlerde yaşamını sürdürdüğü için kentlerde hava kirliliği insan sağlığı açısından büyük sorun teşkil etmektedir (Bannister ve ark., 2021). Özellikle yaşlılar ve çocuklar üzerindeki etkisinin fazla olmasıyla birlikte sağlık hizmetlerindeki ekstra maliyetlere dolayısıyla ekonomiye de olumsuz etkileri olmaktadır (Gonzalez-Enrique ve ark. 2021). İnsanların hava kirliliğine maruz kalmama ihtimalinin az olmasından dolayı ulusal veya bölgesel düzeyde yerel yönetimler tarafından tedbirlerin alınması gereklidir. Hava kirliliğinin sebep olduğu insan sağlığı üzerindeki etkisinin azaltılması konusunda uzun vadeli politikaların geliştirilip uygulanmaya başlanması, yerleşim alanları, ulaşım, enerji üretimi ve sanayi alanları gibi birçok sektörde gerekli olduğu belirtilmektedir (Kolasa-Wiecek ve Suszanowicz, 2019).

Hava kalitesinin belirlenmesinde birçok kirletici olmasına rağmen DSÖ partikül madde 10 (PM₁₀) ve kükürtdioksit (SO₂) ölçümlerinin kullanılmasını yeterli görmektedir (Yılmaz, 2017). PM₁₀, havada asılı duran 10 mikrometre (µm) ve daha küçük boyuttaki katı ve sıvı partikülleri ifade etmektedir. SO₂ ise renksiz, kokusuz ve doğrudan zehirleyici bir gaz olma özelliği taşımaktadır. Hava kirliliği doğal nedenler ve insan faaliyetleri sonucunda olmak üzere 2 şekilde gerçekleşmektedir. Volkanik aktiviteler, orman yangınları ve bataklıklar doğal nedenlere, ısınma, ulaşım, enerji üretimi, sanayileşme ise insana bağlı olarak hava kirliliğine sebep olan faktörler arasında yer almaktadır. Isınma ihtiyacı doğrultusunda oluşan hava kirliliği, düşük kalitedeki yakıtların kullanılması, konutlarda fosil yakıtların kullanılması, yüksek oranlarda kül ve kükürt içermeleri, yanlış yakma sistemleri ve bu sistemlerin düzenli bakımlarının yapılmamasından kaynaklanmaktadır. Motorlu taşıt sayısındaki artışla birlikte trafik yoğunluğuna bununla beraber kentlerde hava kirliliğine neden olmaktadır. Nüfus artışının doğal bir sonucu olarak kentleşme ve sanayileşme bunların yanında topografik özellikler, iklimsel ve meteorolojik olaylar da hava kirliliğini büyük oranda etkilemektedir. Sanayi alanlarının yanlış konumlandırılması veya zaman içerisinde kent alanlarının içinde kalması, arıtma sistemlerinin olmaması veya yeterli olmaması da hava kalitesini son derece etkilemektedir.

Türkiye'de hava kirliliğini doğru tespit edebilmek, temiz hava politikaları oluşturabilmek ve hava kalitesini iyileştirebilmek amacıyla illerde hava kalitesi izleme istasyonları kurularak Ulusal Hava Kalitesi İzleme Ağı oluşturulmuştur. Bu ağ sayesinde saatlik ortalamalar

şeklinde veriler elde edilmektedir. Belirli ve sınırlı sayıdaki istasyonlardan elde edilen veriler sebebiyle jeostatistiksel yöntemler kullanılarak gözlem veya ölçüm yapılmayan konumlarda enterpolasyon ile değerler tahmin edilerek sürekli veri oluşturulabilmektedir (Akyürek ve ark., 2013).

Hava kirliliğine etki eden parametreleri ile kirlilik ilişkilerini inceleyen çalışmalara bakıldığında; (Erener ve ark., 2019), Kocaeli'nde hava kirliliğinin mevsimsel olarak meteorolojik verilerle ilişkisini incelemişlerdir. 2015 yılına ait günlük SO₂ ve PM₁₀ parametreleri ile sıcaklık, rüzgâr yönü, rüzgâr hızı, bağıl nem ve hava basıncı gibi meteorolojik verileri ile regresyon analizi yapılarak değerlendirmeler yapmışlardır. Bai ve ark., 2019, Yangtze Nehri Ekonomik Bölgesinde 2015 yılına ait kentsel hava kalitesi izleme verileri ile birlikte hava kirliliğine etki eden doğal ve sosyoekonomik faktörler üzerine çalışma yapmışlardır. Çalışma sonucunda topografik, meteorolojik, ekonomik kalkınma ve kentleşme faktörleri hava kirliliği ile ilişkili olduğu sonucuna varmışlardır. Birtakım bilim insanları yaptıkları araştırmalarda hava kirliliğine meteorolojik faktörlerin (sıcaklık, yağış, nem, atmosferik basınç, rüzgâr hızı vb.) önemli etkisinin olduğunu göstermişlerdir (Liu ve ark., 2017; Li ve ark., 2014; Lu ve ark., 2017). Lin ve Wang (2016) Çin'de yaptıkları çalışmada ise hava kirliliğinin enerji tüketimine, sanayileşmeye ve teknolojik gelişmelere önemli derecede bağlı olduğunu söylemişlerdir. Bazı bilim insanlarının yaptığı çalışmalarda ise topografik yapının kirleticilerin kümelenmesini ve dağılmasını engellediği gibi sonuçlara ulaşırken nüfus ve gelir seviyesinin artmasıyla birlikte fosil yakıt kullanımı, araç sayısındaki artışlar, kentleşme ve sanayileşme gibi faktörlerin artarak kirletici emisyonların artmasına sebebiyet verdiği dolayısıyla bu faktörlerin hava kirliliğini etkilediği ortaya koyulmuştur (Alvarez ve ark., 2013; Jiang ve ark., 2018).

Kirletici maddelerin insan sağlığına, hayvanlara, bitkilere, doğal çevreye, ticari ve özel mülklere zarar verebilecek seviyelerde ve sürede atmosferde olması olarak tanımlanan hava kirliliğinin izlenmesi, değerlendirilmesi, dağılımının ve değişiminin haritalanması önemlidir (Çetin ve ark., 2019; Kaplan ve Avdan 2020; Aydınoglu ve ark. 2022). Hava kirliliğinin tahmin edilmesi ve bu tahmin doğrultusunda tehlikeleri öngörebilmek ve tedbirler almak insanların, hayvanların, bitki ve tüm canlıların sağlığı olmak üzere doğal çevrenin sürdürülebilirliği açısından yerel yönetimlerce önem arz etmektedir (Yüksek ve ark., 2007).

Hava kirliliğinin tahmini için deterministik yöntemler ve istatistiksel yöntemler olmak üzere 2 farklı yöntemden söz edilmektedir. Deterministik yöntemlerde meteorolojik, emisyon ve kimya modelleri kullanılarak kirleticilere ait deşarj, biriktirme veya transfer süreçleri gibi fiziksel ve kimyasal süreçleri modelleyen bir yöntem olmakla birlikte bu yöntemlerin doğruluğu, kullanılan verinin kalitesi ve ölçüğüne bağlı olarak değişen aynı zamanda bu yöntemlerin büyük veriler ile uygulanmasının oldukça zaman alıcı olduğu söylenmektedir. Deterministik yöntemlerinin eksik ve zayıf noktalarını kapatan istatistiksel yöntemlerin tahmin çalışmalarında öneminin arttığı söylenmektedir.

İstatistiksel yöntemler arasında kriging, regresyon ve yapay sinir ağları (ANN) gibi yöntemler sayılmaktadır (Wang ve ark., 2015; Ghaemi ve ark., 2018).

Hava kirliliğinin tahmini üzerine yapılmış çalışmalara bakıldığında: (Caselli ve ark., 2009), meteorolojik veriler kullanarak hava kirliliğinin tahmini için hem çok değişkenli regresyon analizi hem de yapay sinir ağı yöntemlerini kullanarak 1 ve 2 günlük tahmin modelleri oluşturmuşlar ve çalışma sonucunda yapay sinir ağı modelinin daha doğru sonuçlar verdiğini tespit etmişlerdir. (Campos ve ark., 2021), yaptıkları çalışmada meteorolojik veriler kullanarak yankı durumu ağları (ESN-Echo State Networks) ve aşırı öğrenme makineleri (ELM-Extreme Learning Machines) yöntemleri kullanarak tahmin sonuçlarını çok katmanlı algılayıcı sinir ağı ve çok değişkenli regresyon modelleri ile karşılaştırmaları yapılmıştır. Karşılaştırma sonucunda aşırı öğrenme makineleri ve çok değişkenli regresyon modellerinin en iyi sonuçları verdiği kanısına ulaşmışlardır. (Moustris ve ark., 2012), çalışmalarında hava kirliliği ve meteorolojik verileri kullanarak ozon konsantrasyonunu tahmin etmek için yapay sinir ağı ve çoklu regresyon yöntemi kullanmışlar ve yapay sinir ağının daha üstün olduğunu saptamışlardır. (Mansor ve ark., 2021), saatlik kirlilik verileri ve meteorolojik verileri kullanarak 1, 2 ve 3 saat sonrası için çoklu doğrusal regresyon ile tahmin modelleri geliştirmişlerdir. (Maleki ve ark., 2019), çalışmasında meteorolojik parametreler, kirlenici konsantrasyonları, zaman ve tarih verilerini kullanarak yapay sinir ağları ile hava kalitesi indeks değerinin tahmini için model oluşturmuşlardır. Çalışma sonucunda yapay sinir ağlarının, yetkililerin ve karar vericilerin hava kalitesi indeksinin tahmini için kullanılabilmesi sonucuna varmışlardır. (Dahari ve Wahid, 2017), ileri beslemeli sinir ağı kullanarak O₃, NO₂ ve SO₂ konsantrasyonlarını tahmin modeli için girdi verileri CO, SO₂, NO₂, O₃, partikül madde olan kirlenici parametreler ve rüzgâr hızı, sıcaklık gibi meteorolojik veriler kullanmışlardır. Girdi verilerinin farklı kombinasyonları ile modelleri denemişlerdir. Zaman ve meteorolojik verilerin etkisinin fazla olduğunu gözlemlemişlerdir. (Yadav ve Nath, 2020a), hava kirlenicilerinden olan PM₁₀, SO₂ ve NO₂ parametrelerinin 1 ay öncesinden konsantrasyon değerlerinin tahmin edilmesi için yapay sinir ağı ile model oluşturmuşlardır. (Özcan ve ark., 2006), İstanbul ili için 6 kirlenici parametre kullanmışlar ve 8 meteorolojik parametre ile troposferik O₃ konsantrasyonunu yöntem olarak seçtikleri hücreli yapay sinir ağı ile modellemişlerdir. (Dutta ve Jinsart, 2021), PM₁₀ konsantrasyonunun 1 yıl önceden tahmini için çoklu doğrusal regresyon (ÇDR, MLR-multiple linear regression), yapay sinir ağı doğrusal olmayan çok katmanlı algılayıcısı (ÇKA, MLP ANN) ve doğrusal olmayan sınıflandırma yöntemi ile regresyon ağacı (CART) metotları olmak üzere 3 yöntem kullanmışlar. Çalışma sonucunda en iyi tahmin değerlerini yapay sinir ağı doğrusal olmayan çok katmanlı algılayıcısı yönteminin verdiğini tespit etmişlerdir.

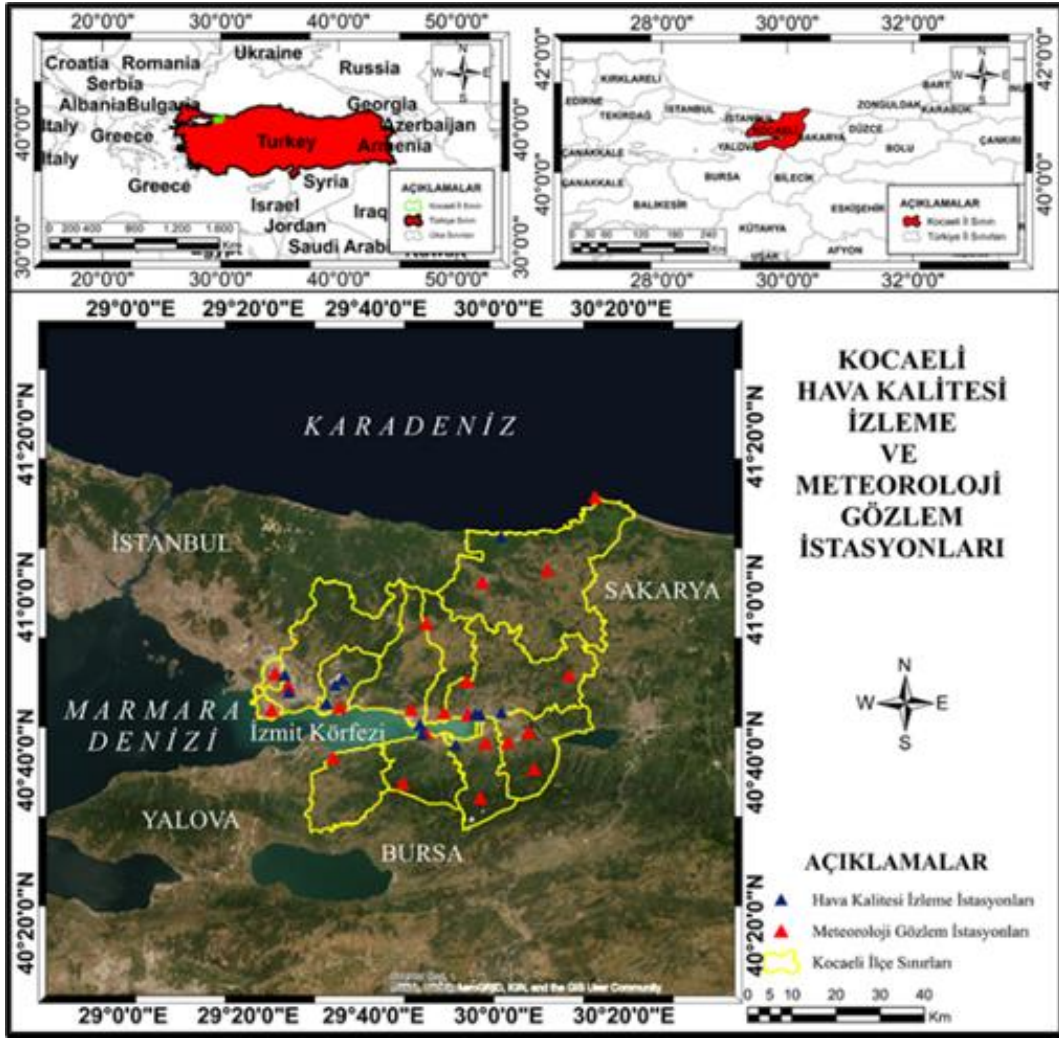
2. Amaç ve Yöntem

Literatürde yapılan çalışmalara bakıldığında hava kirliliğini etkileyen parametrelerde meteorolojik

verilerin yanında topografik ve demografik yapının, trafik emisyonlarının, yerleşim ve sanayi alanlarının da hava kirliliğindeki etkisi görülmektedir. Yapay sinir ağları ve regresyon ile hava kirliliği tahmin modelleme çalışmalarına bakıldığında ise girdi parametrelerinde ağırlıklı olarak meteorolojik veriler kullanılarak modellerin oluşturulduğu görülmektedir. Bu çalışmada Türkiye'nin sanayi açısından önde gelen kentleri arasında yer alan ve nüfus yoğunluğu bakımından İstanbul'dan sonra Türkiye'de 2. kent olan Kocaeli ili için kirlenici maddelerden PM₁₀ ve SO₂ parametrelerine ait mevsimsel ortalama konsantrasyon değerlerinin literatürdeki çalışmaların aksine saatlik veya günlük olarak değil konsantrasyon değerlerinin 1 yıl önceden tahmin edilmesi amaçlanmıştır. Bu amaç doğrultusunda çalışmada; meteorolojik veriler ve hava kirliliği parametre konsantrasyon verilerine literatürdeki çalışmalara ek olarak coğrafi bilgi sistemleri (CBS) kullanılarak ve uzaktan algılama (UA) teknolojilerinden faydalanılarak nüfus, sayısal yükseklik modeli (SYM) verileri, yol ağı, sanayi alanları ve bina izi verileri kullanılarak hem ÇRA stepwise yöntem hem de ileri beslemeli geri yayılım algoritması ile YSA yöntemi kullanılarak PM₁₀ ve SO₂ parametrelerinin mevsimlik ortalama konsantrasyon değerleri için tahmin modellerinin oluşturulması ve model performanslarının karşılaştırılması amaçlanmıştır.

2.1. Çalışma Alanı

Kocaeli, Türkiye'nin Marmara Bölgesi'nde bulunan 29°22'- 30°21' doğu boylamları ve 40°31'- 41°13' kuzey enlemleri arasında yer alan bir büyükşehirdir. Kocaeli Şekil 1'de görüldüğü gibi doğusunda Sakarya; batısında Yalova, İstanbul ve Marmara Denizi; güneyinde Bursa, kuzeyinde Karadeniz bulunmaktadır. Yüz ölçümü 3397 km² olan kent 12 ilçeye sahiptir (URL-1). Kocaeli nüfusu 2021 verilerine göre 2033441 kişidir (URL-2). Türkiye'nin nüfus bakımından 10. kenti; nüfus yoğunluğu bakımından Türkiye'nin en yoğun 2. kent özelliğini taşımaktadır. Kocaeli sanayi alanında Türkiye için önde gelen illeri arasında yer almaktadır. Kocaeli sanayisi il ekonomisine katkısı %51 iken Türkiye sanayisine %13'lük katkı sağlamaktadır. Türkiye'deki araç üretiminin %36,4'ü, kimya sanayisinin %27'si, metal sanayinin %19'u Kocaeli tarafından karşılanmaktadır. Ayrıca bölgede kurulan Bilişim Vadisi ile ar-ge ve inovasyonun merkezi olma konusunda da ilerlemektedir (URL-3). Asya ve Avrupa kıtasını birbirine bağlayan yol üzerindedir ayrıca doğal liman özelliği taşıyan İzmit Körfezi 35 limanı ile işlek bir deniz yoluna sahiptir. Asya ve Avrupa'yı birbirine bağlayan D-100 ve Trans Avrupa Otoyolu (TEM) Kocaeli içerisinde yer almaktadır. Ayrıca TEM yük ve yolcu taşımacılığı açısından Türkiye'nin trafik talebinin en yoğun olduğu trafik koridorları arasındadır. Kocaeli'nin Dilovası ilçesinde bulunan Dil Burnu ile Yalova'nın Altınova ilçesinde bulunan Hersek Burnu arasında inşa edilen dünyanın en uzun açıklıklı asma köprüleri arasında gösterilen özelliği taşıyan otoyol kapsamında inşa edilen İzmit Körfezi'nin geçme süresi 1 saatten 6 dakikaya indiren Osmangazi Köprüsü de mevcuttur (URL-4).



Şekil 1. Kocaeli lokasyon haritası ve meteorolojik gözlem ile hava kalitesi izleme istasyonları

2.2. Veri Seti

Çalışmada kullanılan PM_{10} ve SO_2 parametrelerine ait konsantrasyon değerleri, 2008 yılından 2019 yılına kadar günlük olarak Çevre, Şehircilik ve İklim Değişikliği Bakanlığı tarafından oluşturulan Hava Kalitesi İzleme Ağından (Şekil 1), 12 istasyondan temin edilmiştir. Çalışmada kullanılan bir diğer veri seti meteorolojik (rüzgâr hızı, hava sıcaklığı, buhar basıncı, rüzgâr yönü, bağıl nem) verilerdir. Bu veriler Kocaeli sınırları içinde bulunan (Şekil 1) 21 istasyondan 2008 yılından 2019 yılına kadar günlük olarak temin edilmiştir. Günlük olarak 2008 ile 2019 yılları arası için temin edilen hava kirlenici parametrelerinden olan PM_{10} ve SO_2 konsantrasyon değerleri ile meteorolojik veriler için aylık ortalama değerleri hesaplanarak, aylık ortalama değerlerinden de mevsimlik ortalama değerleri hesaplanmıştır. İstasyonlardan temin edilen kirlenici parametrelerine ait ortalama mevsimlik değerleri PM_{10} için Şekil 2'de, SO_2 için Şekil 3'te gösterilmektedir. İstasyonlardan temin edilen meteorolojik parametrelere ait ortalama mevsimlik değerleri kış, ilkbahar, yaz ve sonbahar mevsimleri sırasıyla Şekil 4, Şekil 5, Şekil 6 ve Şekil 7'de gösterilmektedir.

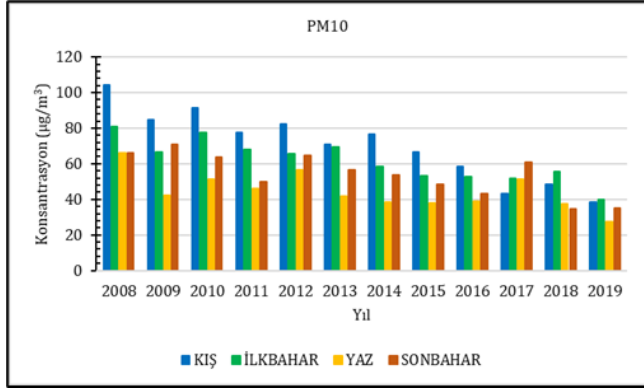
Kocaeli iline ait nüfus verileri TÜİK web sayfasından, kullanılan yol ağı verisi Karayolları Genel Müdürlüğü'nden temin edilmiştir. Meteorolojik veriler ise Meteoroloji Genel Müdürlüğü'nden talep edilmiştir.

Yol ağı şehirlerarası yolları içeren otoyol, şehir içi yolları ve köy yolları olmak üzere 3 ayrı kategoride sınıflandırılmıştır. Çalışma alanına ait topografik bilgi sağlayacak sayısal yükseklik modeli (SYM) verisi <https://search.earthdata.nasa.gov/search> sayfasından temin edilmiştir. Nüfus verileri Şekil 8'de, binalar, sanayi alanları, yol ağı ve sayısal yükseklik modeli ise Şekil 9'da gösterilmektedir.

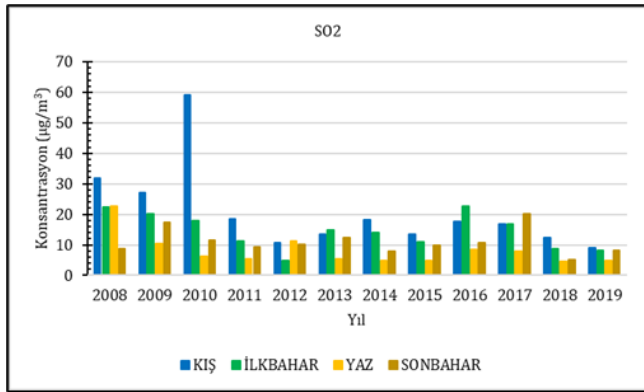
Sınırlı sayıda istasyondan veri temin edildiği için sürekli verinin oluşturulması amacıyla çalışma alanını kapsayacak şekilde ölçüm ve gözlem yapılmayan alanlarda ters mesafe ağırlıklandırma (TMA) metodu ile bilinen komşu değerler dikkate alınarak tüm çalışma alanı için hem PM_{10} ve SO_2 değerleri hem de meteorolojik değerler üretilmiştir. TMA tekniğinin hava kirliliği parametrelerinin enterpolasyonu için uygun olduğu söylenmekte ve buna örnek olarak 2019 yılında hava kalite endeksinin tahmin edilmesi üzerine yapılan çalışma kapsamında parametrelerin enterpolasyonunda TMA tekniği kullanılmıştır (Jumaah ve ark., 2019). Ayrıca başka bir çalışmada ise PM_{10} değerlerini 3 farklı enterpolasyon yöntemleri ile üreterek en iyi performansla sahip yöntemin TMA olduğu tespit edilmiştir (Vorapracha ve ark., 2015). Çalışmada kullanılan kaynak veri ve üretilen veriler Tablo 1'de sunulmaktadır.

Bina, sanayi ve yol ağı verileri için (Şekil 10 ve Şekil 11) kernel yoğunluk (density) analizi ve öklid mesafesi

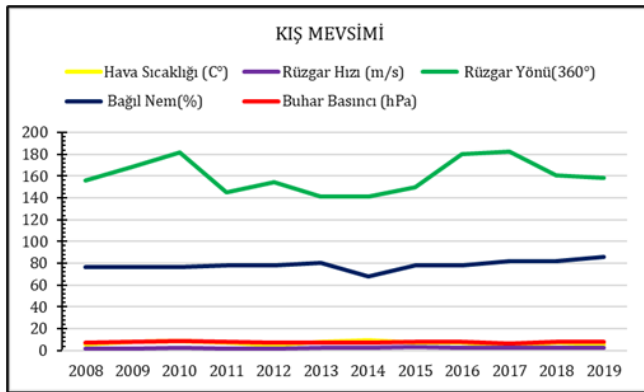
(euclidean distance) analizleri yapılarak, veriler çoklu regresyon analizi ve yapay sinir ağı modelinde girdi olarak kullanılacak hale getirilmiştir.



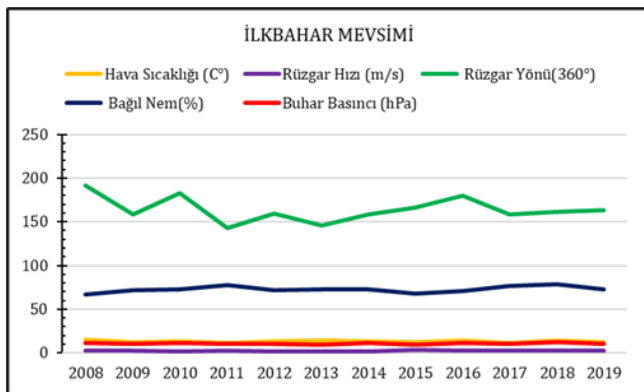
Şekil 2. PM10 mevsimsel ortalama değerleri



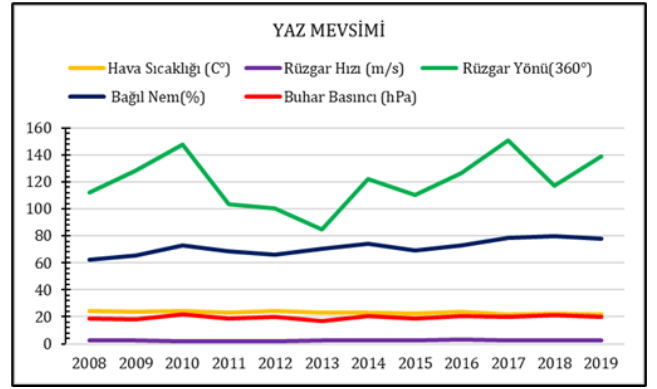
Şekil 3. SO2 mevsimsel ortalama değerleri



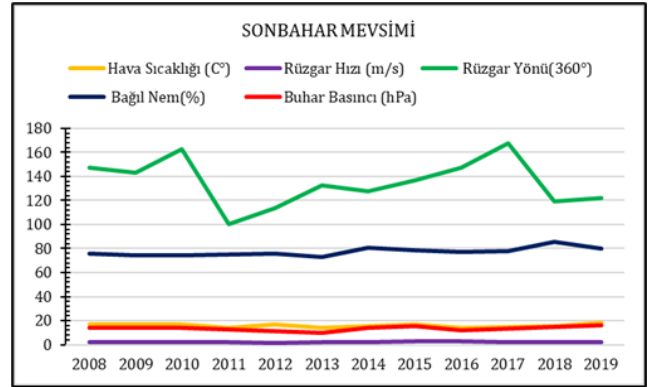
Şekil 4. Kış mevsimi meteorolojik parametrelerin ortalama değerleri



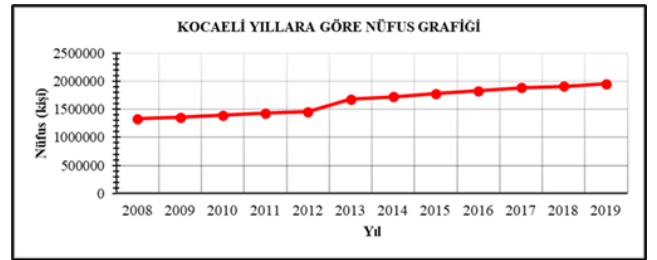
Şekil 5. İlkbahar mevsimi meteorolojik parametrelerin ortalama değerleri



Şekil 6. Yaz mevsimi meteorolojik parametrelerin ortalama değerleri



Şekil 7. Sonbahar mevsimi meteorolojik parametrelerin ortalama değerleri



Şekil 8. Kocaeli ilinin yıllara göre nüfus bilgisi

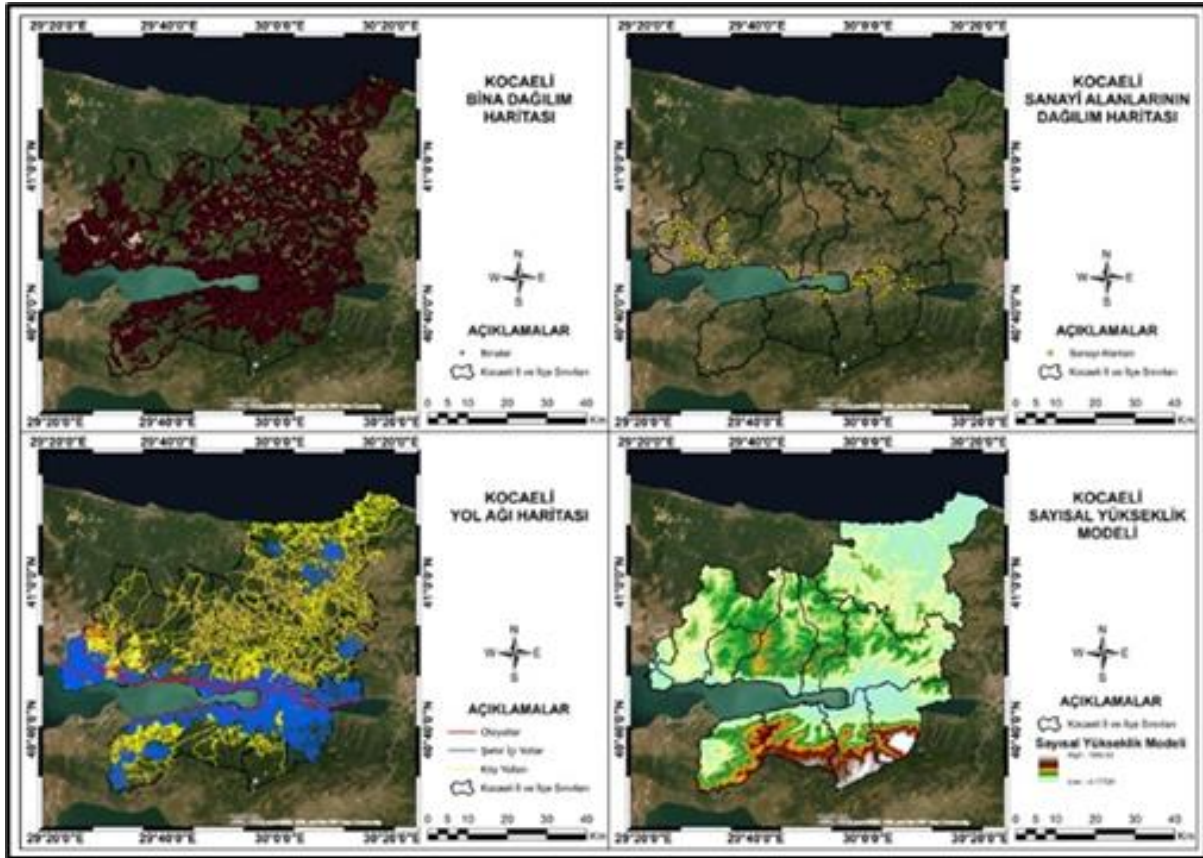
Tablo 1. CBS yazılımı kullanılarak kaynak verilerden mekansal değişkenlerin üretilmesi

Veri Tipi	Kaynak Veri	Kaynak Veri Sınıflandırması	Üretilen Veri	
Hava Kirlenmeleri	Kirlenme Parametreleri	PM10	TMA	
		SO2		
Meteorolojik	Meteorolojik Parametreler	Hava Sıcaklığı	TMA	
		Bağıl Nem		
		Rüzgâr Hızı		
		Rüzgâr Yönü		
Demografik	Nüfus	Nüfus	TMA	
		Otoyol	Mesafe	
		Yoğunluk		
Arazi Kullanımı	Yol Ağı	Şehir İçi Yollar	Mesafe	
		Yoğunluk		
	Endüstriyel Alan	Köy Yolları	Mesafe	
			Yoğunluk	
Topografik	SYM	Sanayi Alanları	Mesafe	
		Yerleşim Alanı	Yoğunluk	
		Bina İzi	Yükseklik Değerleri	

Modelde kullanılacak nüfus verisi Kocaeli’ndeki tüm mahalleler için 2008 yılından 2019 yılına kadar temin edilmiştir. Çalışma alanının tümünde veri sürekliliğini sağlamak amacıyla her mevsime ait nüfus, meteorolojik ve kirletici parametre verileri için TMA yöntemi ile tüm alan üzerinde veriler üretilmiştir. Çalışma alanı üzerinde bina verisi, yol ağı ve sanayi alanları için Kernel yoğunluk ve Öklid mesafeleri kullanılarak sürekli veriler üretilmiştir. Daha sonra yaklaşık mahalle merkezleri kullanılarak 485 örneklem noktasına ait veri tabanı oluşturulmuştur (Şekil 12).

Oluşturulan veri tabanı 12 yılı (2008-2019) kapsayan her yıl için 485 örneklem noktasından oluşan toplamda 5820 satırlık veridir. Her satır ise 42 parametreyi (x

koordinatı, y koordinatı, nüfus, bina mesafe, bina yoğunluk, sanayi mesafe, sanayi yoğunluk, şehir içi yolları mesafe, şehir içi yolları yoğunluk, otoyol mesafe, otoyol yoğunluk, köy yolları mesafe, köy yolları yoğunluk, SYM, kış sıcaklık, kış nem, kış basınç, kış rüzgâr hızı, kış rüzgâr yönü, kış PM₁₀, kış SO₂, ilkbahar sıcaklık, ilkbahar nem, ilkbahar basınç, ilkbahar rüzgâr hızı, ilkbahar rüzgâr yönü, ilkbahar PM₁₀, ilkbahar SO₂, yaz sıcaklık, yaz nem, yaz basınç, yaz rüzgâr hızı, yaz rüzgâr yönü, yaz PM₁₀, yaz SO₂, sonbahar sıcaklık, sonbahar nem, sonbahar basınç, sonbahar rüzgâr hızı, sonbahar rüzgâr yönü, sonbahar PM₁₀, sonbahar SO₂) içeren değerlerden oluşmaktadır.

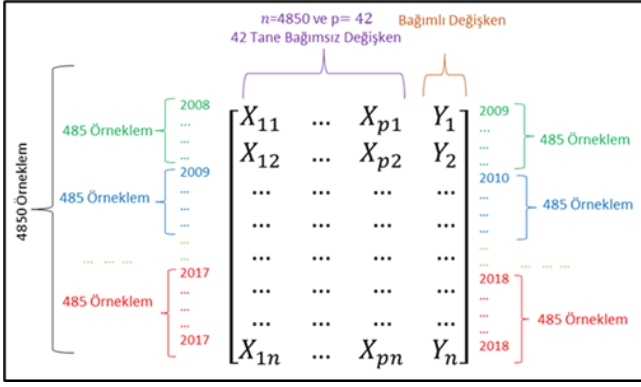


Şekil 9. Kocaeli bina ve sanayi alanlarının dağılımı, yol ağı ve sayısal yükseklik modeli (SYM)

2.3. Çalışmada İş Akışı ve Uygulama

Literatürde hava kirliliğinin tahmin edilmesi amacıyla çoklu regresyon analizi (ÇRA) ve YSA yöntemleri ile tahmin modellerinin oluşturulduğu görülmektedir. Oluşturulan tahmin modellerinde genel olarak meteorolojik parametrelerin ve kirletici parametrelerin girdi olarak kullanıldığı, ayrıca saatlik ve günlük olarak tahmin modellerin oluşturulduğu görüldü (Caselli ve ark., 2009; Campos ve ark., 2021; Moustris ve ark., 2012; Mansor ve ark., 2021; Maleki ve ark., 2019; Dahari ve Wahid, 2017; Yadav ve Nath, 2020b; Özcan ve ark., 2006; Dutta ve Jinsart, 2021). Bunun yanında hava kirliliğine etki eden parametreler ve kirlilik ilişkisini inceleyen çalışmalarda ise meteorolojik verilerin yanında kentleşme, sanayileşme, topografyanın ve sosyoekonomik faktörlerinde hava kirliliğine etki etti söylenmektedir (Erener ve ark., 2019; Bai ve ark., 2019;

Liu ve ark., 2017; Li ve ark., 2014; Lu ve ark., 2017; Lin ve Wang, 2016; Alvarez ve ark., 2013; Jiang ve ark., 2018). Bu çalışma kapsamında hava kirliliğinin tahmini için meteorolojik ve kirletici parametrelerin yanında, topografik, demografik, sanayileşme ve kentleşmeyi de dahil edecek verilerle birlikte, ayrıca 1 yıl önceden mevsimsel ortalama tahmin değerlerini elde etme amacıyla tahmin modelleri oluşturuldu. Çalışma amacında belirtildiği üzere 1 yıl sonrası için tahmin modeli üretilmek istendiği için her bağımsız değişken grubu 1 yıl sonraki bağımlı değişken ile ilişkilendirildi. Modelin oluşturulması ve analizlerin yapılması için Şekil 13'te görüldüğü gibi 2008-2017 yıllarına ait bağımsız değişkenler 2009-2018 yıllarına ait bağımlı değişken ile ilişkilendirilip veri setleri her mevsim için hem PM₁₀ hem de SO₂ için oluşturuldu. Modeller oluşturulduktan sonra 2018 yılına ait bağımsız değişkenler modele sunularak 2019 yılına ait tahmin değerleri elde edildi.



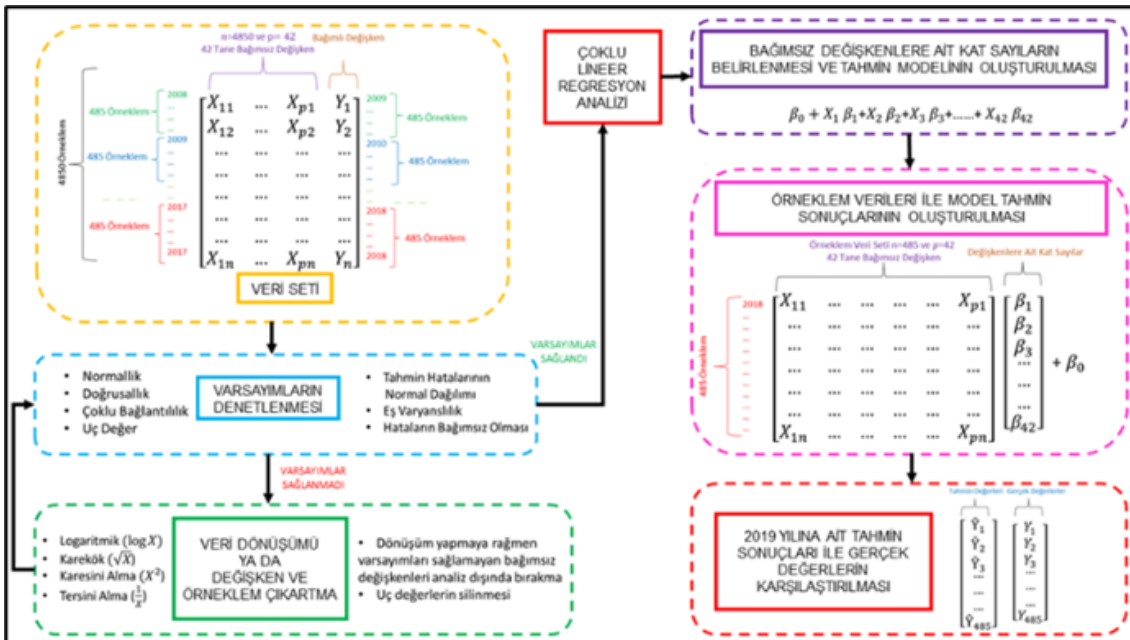
Şekil 13. Veri seti yapısı

2.3.1. Çoklu doğrusal regresyon analizi modelleri

Sorunlar ve sorunlara sebebiyet veren faktörler incelendiğinde bunların birbirleri ile ilişkili oldukları görülmektedir. Bu doğrultuda sorunlara sebebiyet veren faktörlerin sorunla ilişkisini açıklayan ya da modelleyen matematiksel ifadeye regresyon denklemi denilirken regresyon denkleminde yer alan değişken veya değişkenlerin diğer bir değişkeni en doğru şekilde modellemesini sağlayarak regresyon denkleminin elde edilmesine ise regresyon analizi denmektedir. Eğer bir bağımsız değişken kullanılarak bağımlı değişken tahmin edilmeye çalışılırsa basit doğrusal regresyon adını alırken, birden fazla bağımsız değişken ile bağımlı değişken tahmin edilmeye çalışılırsa çoklu doğrusal regresyon adını almaktadır (Ghatak, 2017; Adachi, 2020). Regresyon denklemlerindeki model parametrelerinden olan katsayıların ve sabit sayının tahmini tespit edilmesi için en küçük kareler yöntemi en fazla kullanılan yöntemdir (Lee ve ark., 2019; Forsyth, 2019).

Regresyon analizlerinde parametrik testlerin yapılabilmesi, yapılan analizin doğruluğu ve güvenilirliği açısından birtakım varsayımlar mevcuttur. Değişkenlerin normallik varsayımı için basıklık ve çarpıklık değerlerinin ± 2 içinde olması gerektiği

söylenmektedir (George ve Mallery, 2010). Ayrıca normallik varsayımı için farklı bir görüş olarak da tahminlere ait hataların normal dağılması gerektiği de söylenmektedir (Field, 2009). Bu çalışmada her mevsim için yapılan varsayım analizlerinde normallik için 2 varsayımında sağlanmasına dikkat edildi. Diğer bir varsayım ise bağımlı ve bağımsız değişkenlerin doğrusallığı varsayımıdır. Bağımlı ve bağımsız değişkenler arasında doğrusal bir ilişkinin olmaması beklenir. Bu kapsam doğrultusunda tüm mevsimler için hazırlanan veri setleri için bağımlı ve bağımsız değişkenler arasında korelasyon analizi yapılmıştır. Çok değişkenli regresyon analizlerinde çoklu bağlantılılık (multicollinearity) problemi ile karşılaşmaktadır. Bu problemin denetlenmesi bağımsız değişkenler arasında yapılacak korelasyon analizi sonucunda analizde kullanılan değişkenler arasında güçlü bir ilişkinin olmaması gerekmektedir (Field, 2009). Bu problemi tespit etmenin başka bir yolu ise varyans tümsek faktörü (VIF, Variance Inflation Factor) değerinin 10'dan büyük olmasıdır. Bu değer 10'dan büyük olması çoklu bağlantılılık probleminin olduğunu göstermektedir (Field, 2009; Tsagris ve Pandis, 2021; Voss, 2005). Yapılacak testlerin gücüne ve tahmin modellerinin sonucuna etki eden uç değerlerin tespiti için Cook distance değerinin 1'den büyük olması ve standartlaştırılmış artık değerinin $\pm 3,29$ içinde olmaması uç değerlerin göstergesidir (Field, 2009). Eşvaryanslılık (homoscedasticity) varsayımında ise bağımsız değişkenlere karşılık gelen hataların eşit varyansa sahip olması beklenmektedir. Bu kapsam dahilinde bu varsayım için grafik yöntem kullanılmaktadır. Son olarak otokorelasyon varsayımı ise hataların birbirinden bağımsız olması gerekliliğidir. Bu varsayım için Durbin-Watson değerinden faydalanılır. Bu değer 1'den küçük ve 3'den büyük olması otokorelasyon sorununu göstermektedir (Field, 2009). Çoklu lineer regresyon yöntemi ile oluşturulacak tahmin modeli iş akışı Şekil 14'te gösterilmektedir. Analizler SPSS Statistics 26 yazılımında gerçekleştirilmiştir.



Şekil 14. Çoklu lineer regresyon tahmin modeli iş akışı

Tablo 2. Tahminlerde kullanılan bağımsız değişkenler

Bağımlı Değişkenler	Kış		İlkbahar		Yaz		Sonbahar	
	PM ₁₀	SO ₂	PM ₁₀	SO ₂	PM ₁₀	SO ₂	PM ₁₀	SO ₂
Bağımsız Değişkenler	Kış PM ₁₀	Yaz SO ₂	Kış PM ₁₀	Kış SO ₂	Yaz SO ₂	Yaz SO ₂	İlkbahar PM ₁₀	Yaz SO ₂
	Yaz Sıcaklık	İlkbahar PM ₁₀	Yaz PM ₁₀	Yaz SO ₂	Sonbahar Sıcaklık	Kış Rüzgâr Hızı	Sonbahar Sıcaklık	Kış PM ₁₀
	Kış Rüzgâr Hızı	Kış Sıcaklık	Kış Sıcaklık	Sonbahar Sıcaklık	Yaz PM ₁₀	Kış Sıcaklık	İlkbahar SO ₂	İlkbahar SO ₂
	İlkbahar PM ₁₀	İlkbahar Sıcaklık	Sonbahar Sıcaklık	Sonbahar SO ₂	Sonbahar SO ₂	Sonbahar PM ₁₀	Yaz Sıcaklık	Kış Rüzgâr Hızı
	Yaz SO ₂	Yaz Rüzgâr Hızı	Sonbahar SO ₂	İlkbahar Sıcaklık	İlkbahar PM ₁₀	Sonbahar SO ₂	Yaz SO ₂	Yaz Sıcaklık
	Sonbahar Sıcaklık	Kış PM ₁₀	Yaz SO ₂	Yaz Sıcaklık	Kış Sıcaklık	Kış PM ₁₀	Otoyollara Mesafe	Sonbahar PM ₁₀
	Sonbahar PM ₁₀	Sonbahar PM ₁₀	Kış SO ₂	Yaz Rüzgâr Hızı	İlkbahar SO ₂	Kış SO ₂	Sonbahar Nem	Sonbahar Rüzgâr Hızı
	Kış SO ₂	İlkbahar SO ₂	Sonbahar PM ₁₀	İlkbahar Rüzgâr Hızı	Yaz Sıcaklık	Sonbahar Sıcaklık	Kış PM ₁₀	İlkbahar Rüzgâr Hızı
	İlkbahar Sıcaklık	Kış Rüzgâr Hızı	Yaz Sıcaklık	İlkbahar SO ₂	İlkbahar Rüzgâr Hızı	İlkbahar Rüzgâr Hızı	Nüfus	İlkbahar Sıcaklık
	Yaz PM ₁₀	Sonbahar SO ₂	Otoyollara Olan Mesafe		Sonbahar PM ₁₀	Şehir İçi Yol Yoğunluğu	İlkbahar Rüzgâr Hızı	Kış SO ₂
	Sonbahar Nem	İlkbahar Rüzgâr Hızı	Nüfus			X Koordinatı		Sonbahar SO ₂
	Şehir İçi Yolu Yoğunluğu	Yaz Sıcaklık	İlkbahar PM ₁₀			SYM		Otoyol Yoğunluğu
	Nüfus	Sonbahar Sıcaklık	İlkbahar Rüzgâr Hızı			İlkbahar PM ₁₀		İlkbahar PM ₁₀
		Sanayi Alanlarının Yoğunluğu	Yaz Rüzgâr Hızı			Sonbahar Rüzgâr Hızı		Sanayi Alanlarının Yoğunluğu
		Kış SO ₂						
	SYM							
	Otoyollara Olan Mesafe							

Tablo 3. Çoklu regresyon analizi model performansları

Mevsimler	Kış		İlkbahar		Yaz		Sonbahar	
	PM ₁₀	SO ₂	PM ₁₀	SO ₂	PM ₁₀	SO ₂	PM ₁₀	SO ₂
Kirletici Model	15	19	14	9	12	14	10	14
R	0.9	0.9	1.0	0.9	0.7	0.7	0.8	0.8
R ²	0.8	0.9	0.9	0.8	0.5	0.5	0.7	0.7
Düzeltilmiş R ²	0.8	0.9	0.9	0.8	0.5	0.5	0.6	0.7
Tahminin Standart Hatası	8.4	0.1	2.9	2.6	0.5	0.1	6.3	0.1
Ortalama Gerçek Değer	64.4	8.4	42.1	7.9	101.6	29.5	60.6	21.6
Ortalama Tahmin Değer	36.2	0.2	34.7	16.5	30.0	5.6	60.1	6.3
Korelasyon	-0.7	0.0	0.9	0.7	-0.7	0.9	-0.5	0.7
OKH	0.3	0.1	0.0	0.0	0.2	0.1	0.2	0.1
KOH	0.5	0.3	0.1	0.1	0.5	0.2	0.5	0.3
OH	-0.2	0.2	0.0	-0.1	-0.2	0.1	-0.1	0.1
OYH	42.4	98.2	19.9	-109.5	70.2	81.5	-2.9	71
OMYH	44.0	100.1	20.2	109.5	70.2	81.5	19.4	71
OMH	0.5	0.2	0.1	0.1	0.4	0.2	0.4	0.2

Varsayımların sağlanması açısından bağımlı değişkenlerin tahminlerinde kullanılacak olan bağımsız değişkenler **Tablo 2**'de gösterilmektedir. Çoklu regresyon analizinde bağımsız değişkenler arasında eş doğrusallık problemine karşılık bir çözüm olduğu

söylenen stepwise metodu kullanılmıştır ([Chen ve ark., 2013](#)).

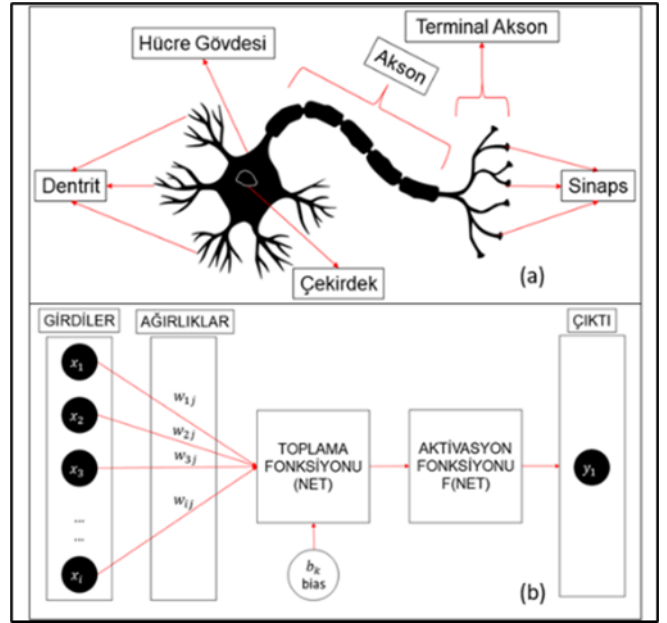
Her mevsim için hem PM₁₀ hem de SO₂ için toplamda 8 veri seti ile 8 tahmin modeli oluşturuldu. Yapılan regresyon analizi sonuçlarına göre p<0,01 ile anlamlı

modeller oluşturulduğu görülmüştür. Tablo 3'te her mevsim ve her parametreye ait oluşturulan 8 tahmin modelinin stepwise metodu ile yapılan regresyon analizi sonuçları ve elde edilen tahmin değerleri ile gerçek değerlerin karşılaştırılarak model performansları görülmektedir. 2019 yılı için tahmin değerleri ve gerçek değerlerin karşılaştırılması için gerçek değerler ile hesaplanan tahmin değerlerinin ortalama değeri, korelasyon katsayısı, ortalama yüzde hata (OYH), ortalama mutlak yüzde hata (OMYH) ve oluşturulan modelin performansının daha kolay anlaşılması açısından min-max normalizasyonu uygulanan gerçek ve tahmin değerlerin ortalama karesel hata (OKH), karekök ortalama hata (KOH), ortalama hata (OH) ve ortalama mutlak hata (OMH) hesaplanarak değerler Tablo 3'te gösterilmektedir.

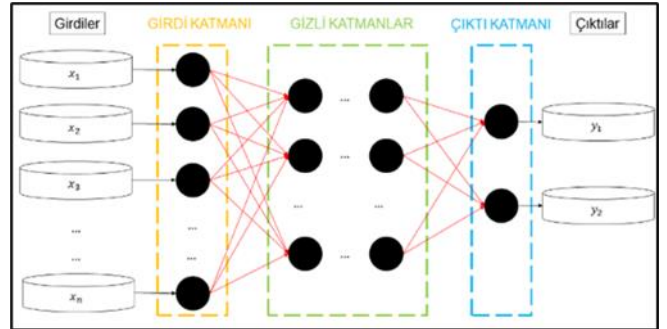
2.3.2. Yapay sinir ağı modelleri

Yapay zekanın bir dalı olan YSA, biyolojik beyin çalışma sisteminden faydalanılarak oluşturulmuş veri işleme ve bilgisayar sistemlerindedir. YSA geçmiş verileri kullanarak karmaşık problemlerin çözümünde kullanılmaktadır. Örüntü tanımlama, tahminde bulunma ve sınıflandırma yapma yeteneğine sahip matematiksel yapı olarak tanımlanmaktadır (Moustris ve ark., 2010; Haglin ve ark., 2019; Yadav ve Nath, 2020b). Biyolojik sinir ağından esinlenerek geliştirilen YSA Şekil 15'te gösterilen yapay sinir hücrelerinden meydana gelmektedir.

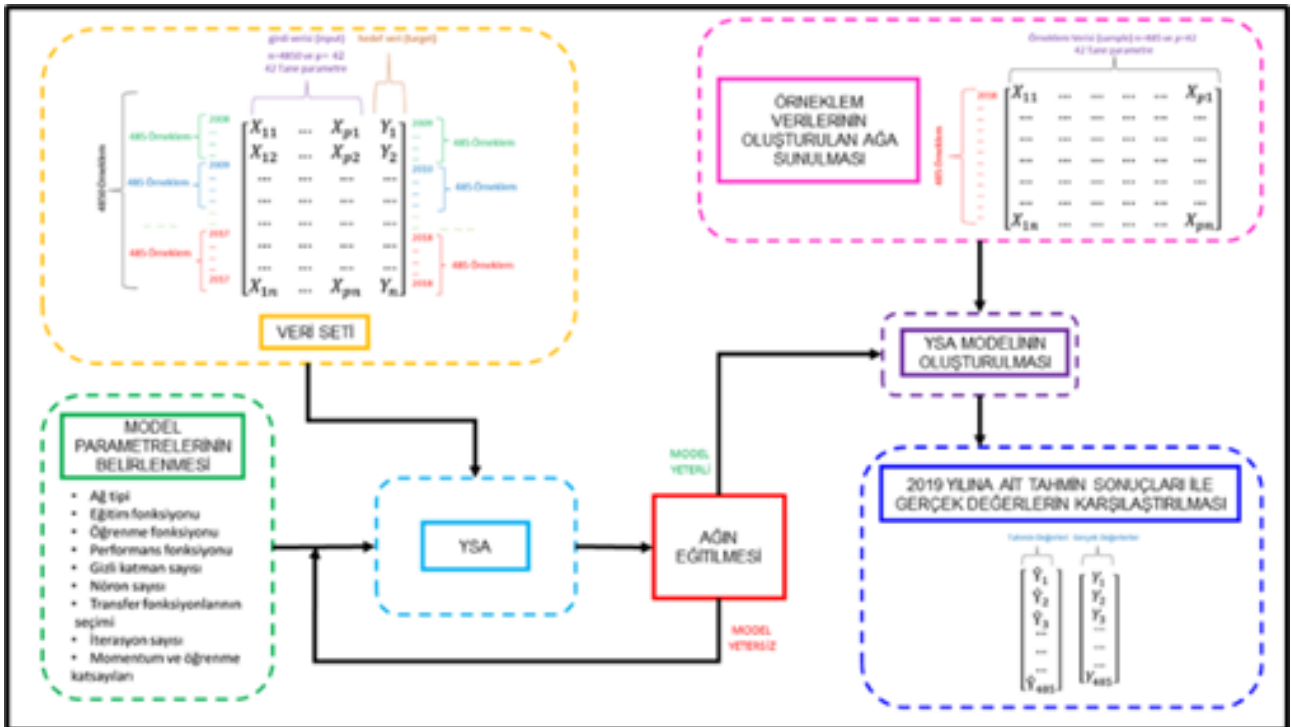
YSA da Şekil 16'da gösterildiği gibi birbirine bağlı şekilde bulunan katmanlar, yapay sinir ağlarının içerisinde barındırdığı yapay sinir hücrelerinin bir araya gelmesiyle oluşturulur.



Şekil 15. Biyolojik hücre yapısı (a), yapay sinir hücre yapısı (b)



Şekil 16. Yapay sinir ağı yapısı, çok katmanlı algılayıcı



Şekil 17. Yapay sinir ağı tahmin modeli iş akışı

Tablo 4. Çoklu regresyon analizi model performansları

Deneme	Katman	Gizli Katman	1.Katman Transfer Fonk	Nöron Sayısı	2.Katman Transfer Fonk	Nöron Sayısı	Çıktı Katman Transfer Fonk	Nöron Sayısı
1	2	1	tansig	10			tansig	1
2	2	1	tansig	15			tansig	1
3	2	1	tansig	30			tansig	1
4	2	1	logsig	10			tansig	1
5	2	1	logsig	15			tansig	1
6	2	1	logsig	30			tansig	1
7	2	1	logsig	10			logsig	1
8	2	1	logsig	15			logsig	1
9	2	1	logsig	30			logsig	1
10	2	1	tansig	10			logsig	1
11	2	1	tansig	15			logsig	1
12	2	1	tansig	30			logsig	1
13	3	2	tansig	10	tansig	10	tansig	1
14	3	2	tansig	15	tansig	10	tansig	1
15	3	2	tansig	30	tansig	10	tansig	1
16	3	2	tansig	10	tansig	15	tansig	1
17	3	2	tansig	15	tansig	15	tansig	1
18	3	2	tansig	30	tansig	15	tansig	1
19	3	2	tansig	10	tansig	30	tansig	1
20	3	2	tansig	15	tansig	30	tansig	1
21	3	2	tansig	30	tansig	30	tansig	1
22	3	2	logsig	10	tansig	10	tansig	1
23	3	2	logsig	15	tansig	10	tansig	1
24	3	2	logsig	30	tansig	10	tansig	1
25	3	2	logsig	10	tansig	15	tansig	1
26	3	2	logsig	15	tansig	15	tansig	1
27	3	2	logsig	30	tansig	15	tansig	1
28	3	2	logsig	10	tansig	30	tansig	1
29	3	2	logsig	15	tansig	30	tansig	1
30	3	2	logsig	30	tansig	30	tansig	1
31	3	2	logsig	10	logsig	10	tansig	1
32	3	2	logsig	15	logsig	10	tansig	1
33	3	2	logsig	30	logsig	10	tansig	1
34	3	2	logsig	10	logsig	15	tansig	1
35	3	2	logsig	15	logsig	15	tansig	1
36	3	2	logsig	30	logsig	15	tansig	1
37	3	2	logsig	10	logsig	30	tansig	1
38	3	2	logsig	15	logsig	30	tansig	1
39	3	2	logsig	30	logsig	30	tansig	1

Yapay sinir ağı tahmin modelleri Matlab R2021a yazılımı nntool ile oluşturulmuştur. Yapay sinir ağı ile oluşturulacak hava kirliliği tahmin modeli iş akışı Şekil 17’de gösterilmektedir. Yapay sinir ağı için performansına normalizasyon tekniklerinin etki ettiği söylenmektedir (Jayalakshmi ve Santhakumaran, 2011). Verilerin çok farklı ölçeklerde olmasından dolayı ve literatürde çokça kullanılan yöntem olmasından dolayı min-max normalizasyonu tercih edilerek verilere uygulandı. Yapay sinir ağı modellerinde ileri beslemeli geri yayılım algoritması ile YSA kullanılarak 42 parametrenin hepsi girdi olacak şekilde tahmin modelleri oluşturulmuştur.

Yapay sinir ağı modellerinde en çok tercih edilen ağ tipi olan feed-forward backprop kullanılırken, eğitim fonksiyonunda trainlm, öğrenme fonksiyonu olarak learnnrm ve performans fonksiyonu olarak OKH (mean squared error) tercih edildi. Oluşturulan 8 model için yapılan denemelerde kullanılan katman sayıları, nöron sayıları ve transfer fonksiyonları Tablo 4’te gösterilmektedir.

Her mevsim için hem PM₁₀ hem de SO₂ parametreleri için oluşturulan 8 veri seti ile yapay sinir ağı modelleri oluşturuldu. Bu modeller oluşturulurken her bir veri seti için YSA model parametrelerinde yapılan değişiklikler ile farklı denemeler yapıldı. Yapılan denemelerde en iyi performansa sahip model tespit edildi. Tespit sırasında hem performans ölçütlerine dikkat edildi hem de ağın eğitilmesi sırasında eğitim, doğrulama ve test verilerinin grafiklerine göre yorumlanarak seçildi. Denemeler sonucunda en iyi performansa sahip modelin 2019 yılı için tahmin değerleri ve gerçek değerlerin karşılaştırılması için gerçek değerler ile hesaplanan tahmin değerlerinin ortalama değeri, korelasyon katsayısı, OYH, OMYH ve oluşturulan modelin performansının daha kolay anlaşılması açısından min-max normalizasyonu uygulanan gerçek ve tahmin değerlerin OKH, KOH, OH ve OMH hesaplanarak Tablo 5’te gösterilmektedir.

Tablo 5. Yapay sinir ağı tahmin model performansları

Mevsim	Kış		İlkbahar		Yaz		Sonbahar	
	PM ₁₀	SO ₂	PM ₁₀	SO ₂	PM ₁₀	SO ₂	PM ₁₀	SO ₂
Deneme	28	15	29	19	23	31	13	22
Ort. Gerçek	64.4	8.4	42.1	7.9	101.6	29.5	60.6	21.6
Ort. Tahmin	60.6	7.8	42.3	9.3	99.3	5.2	50.8	20.0
Corr.	-0.2	0.9	1.0	0.8	-0.1	0.8	0.1	-0.3
OKH	0.2	0.0	0.0	0.1	0.5	0.1	0.2	0.2
KOH	0.4	0.1	0.1	0.2	0.7	0.2	0.4	0.5
OH	-0.2	0.0	0.1	0.2	-0.5	-0.1	-0.2	-0.1
OYH	5.3	10.3	-1.2	-15	1.9	83.0	14.0	5.2
OMHY	15.6	21.8	4.9	20.9	24.0	83.0	19.5	44.9
OMH	0.4	0.1	0.1	0.2	0.6	0.2	0.3	0.4

3. Bulgular

Çalışmada denemeler ve modeller arasındaki karşılaştırmalar OMYH değerine göre yapılmıştır. OMYH değerinin %10’un altında olan modellerin yüksek doğrulukta olduğu söylendiği gibi başka bir görüş ise

%10’un altındaki modellerin çok iyi dereceli olduğu, %10-%20 arasında olan modellerin iyi dereceli modeller olduğu, %20-%50 arasında olan modellerin kabul edilebilir dereceli modeller olduğu ve OMYH değeri %50’den fazla olan modellerin ise hatalı ve yanlış modeller olduğu söylenmektedir (Çuhadar ve Kayacan,

2005; With ve With, 1992; Lewis, 1982). Bu kapsamda CRA ve YSA modellerinin OMYH değerlerine göre performans dereceleri PM₁₀ için **Tablo 6**'da SO₂ için **Tablo 7**'de gösterilmektedir. PM₁₀ konsantrasyonunun 1 yıl önceden mevsimsel ortalama değerinin tahmin edilmesi için çoklu regresyon analizi ile oluşturulan tahmin modellerinde sonbahar mevsimi için %19.41 OMYH değeri ile iyi derecede model; kış ve ilkbahar mevsimleri için sırasıyla %44 ve %20.15 ile kabul edilebilir derecede bir model ve yaz için ise %70.20 ile yanlış ve hatalı bir model oluşturulduğu görüldü. PM₁₀ için yapay sinir ağı ile oluşturulan modellerde ise ilkbahar mevsimi için %4.91 OMYH değeri ile çok iyi derecede bir model; kış ve sonbahar mevsimleri için sırasıyla %15.61 ve %19.46 ile iyi derecede model ve yaz mevsimi için %24.02 ile kabul edilebilir derecede bir model oluşturulmuştur. SO₂ konsantrasyonunun 1 yıl önceden mevsimsel ortalama değerinin tahmin edilmesi için çoklu regresyon analizi ile oluşturulan tahmin modelleri kış mevsimi %100.14, ilkbahar mevsimi %109.51, yaz mevsimi %81.46 ve sonbahar mevsimi %70.99 OMYH değerleri ile yanlış ve hatalı modellerin oluşturulduğu görülmüştür. SO₂ için yapay sinir ağı ile oluşturulan modeller ise kış mevsimi için %21.76, ilkbahar mevsimi için %20.88 ve sonbahar mevsimi için %44.93 OMYH değerleri ile kabul edilebilir modeller oluşturulduğu ancak yaz mevsimi için %82.99 OMYH değeri ile yanlış ve hatalı bir model oluşturulduğu tespit edilmiştir.

Tablo 6. PM₁₀ parametresi model performans dereceleri

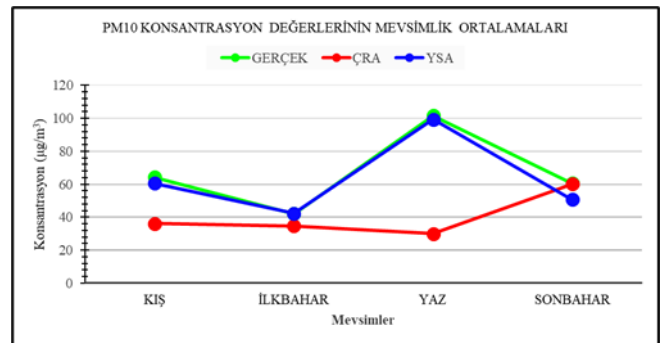
Mevsim	Gerçek Ort.	PM ₁₀					
		CRA			YSA		
		Tahmin Ort.	OMYH	Derecesi	Tahmin Ort.	OMYH	Derecesi
Kış	64.4	36.2	44.0	Kabul Edilebilir	60.6	15.6	İyi
İlkbahar	42.1	34.7	20.2	Kabul Edilebilir	42.3	4.9	Çok İyi
Yaz	101.6	30.0	70.2	Yanlış ve Hatalı	99.3	24.0	Kabul Edilebilir
Sonbahar	60.6	60.1	19.4	İyi	50.8	19.5	İyi

Tablo 7. SO₂ parametresi model performans dereceleri

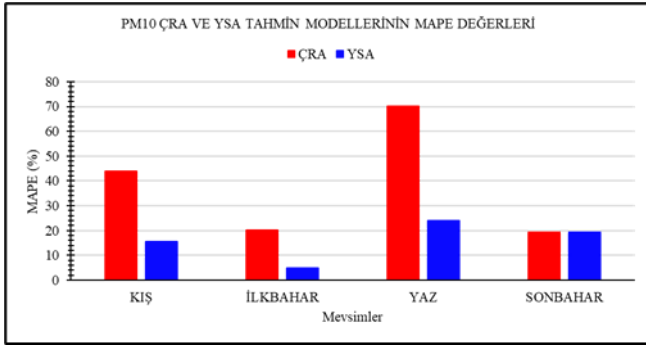
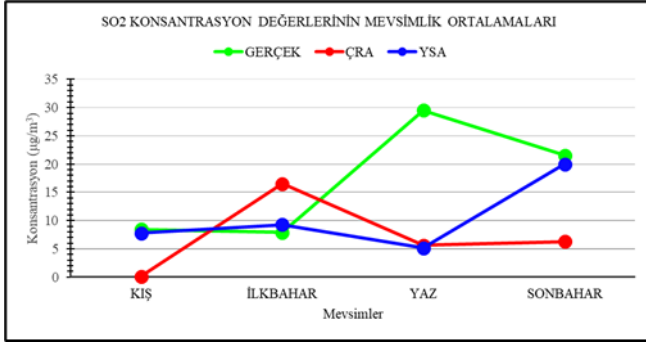
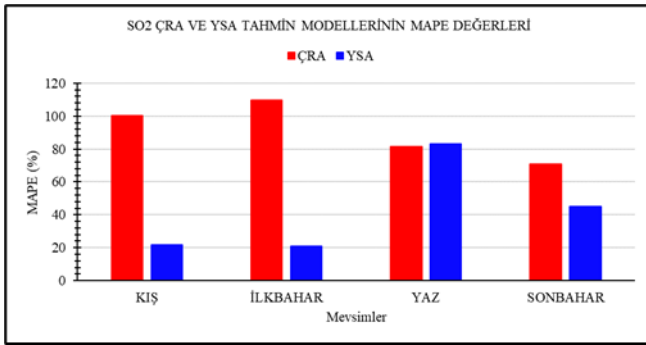
Mevsim	Gerçek Ort.	SO ₂					
		CRA			YSA		
		Tahmin Ort.	OMYH	Derecesi	Tahmin Ort.	OMYH	Derecesi
Kış	8.4	0.2	100.1	Yanlış ve Hatalı	7.8	21.8	Kabul Edilebilir
İlkbahar	7.9	16.5	109.5	Yanlış ve Hatalı	9.3	20.9	Kabul Edilebilir
Yaz	29.5	5.6	81.5	Yanlış ve Hatalı	5.2	83.0	Yanlış ve Hatalı
Sonbahar	21.6	6.3	71.0	Yanlış ve Hatalı	20.0	44.9	Kabul Edilebilir

ÇRA ve YSA yöntemleri ile oluşturulan modellere göre 2019 yılı için PM₁₀ ve SO₂ parametrelerinin tahmin ve gerçek değerlerinin mevsimsel ortalama değer ve OMYH grafikleri çizildi. PM₁₀ için mevsimsel ortalama değerleri **Şekil 18**'de, OMYH değerleri **Şekil 19**'da, SO₂ için mevsimsel ortalama değerleri **Şekil 20**'de ve OMYH değerleri **Şekil 21**'de gösterilmektedir.

Parametrelerin tahminine yönelik bu çalışmanın daha kolay yorumlanabilmesi ve gerçek değerler ile karşılaştırılmasını kolaylaştırması açısından PM₁₀ konsantrasyonuna ait tahmin ve gerçek değerlerin yoğunluk haritaları **Şekil 22**'de gösterilmektedir.



Şekil 18. PM₁₀ konsantrasyonu için mevsimsel ortalama değerleri

Şekil 19. PM₁₀ konsantrasyonu OMYH değerleriŞekil 20. SO₂ konsantrasyonu için mevsimsel ortalama değerleriŞekil 21. SO₂ konsantrasyonu OMYH değerleri

SO₂ konsantrasyonuna ait tahmin ve gerçek değerlerin yoğunluk haritaları Şekil 23'de gösterilmektedir.

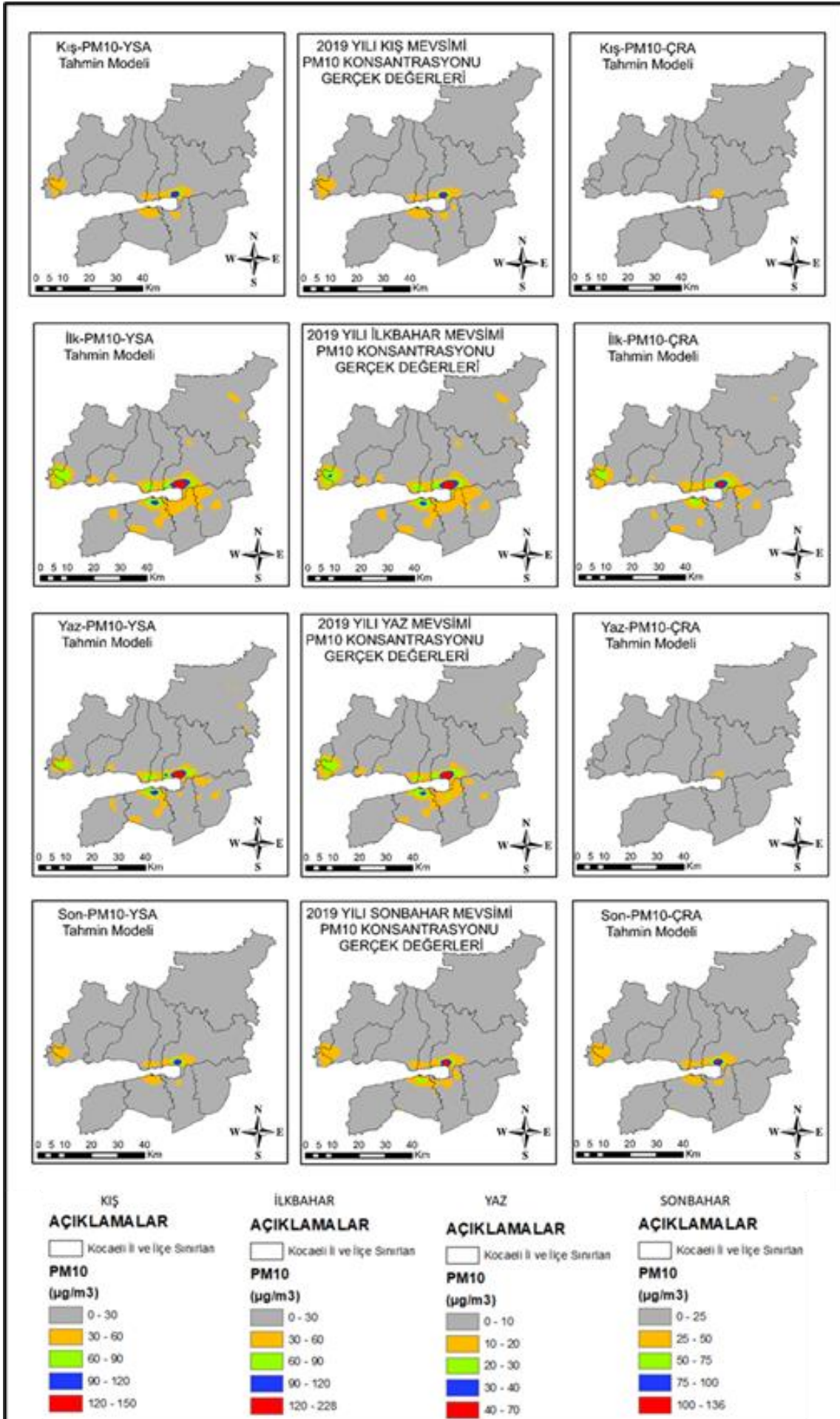
Yapılan çalışmada elde edilen bulgular doğrultusunda, grafiklerde ve haritalarda gözüktüğü gibi YSA modellerinin ÇRA modellerine göre performansının daha yüksek olduğu tespit edilmiştir.

4. Tartışma

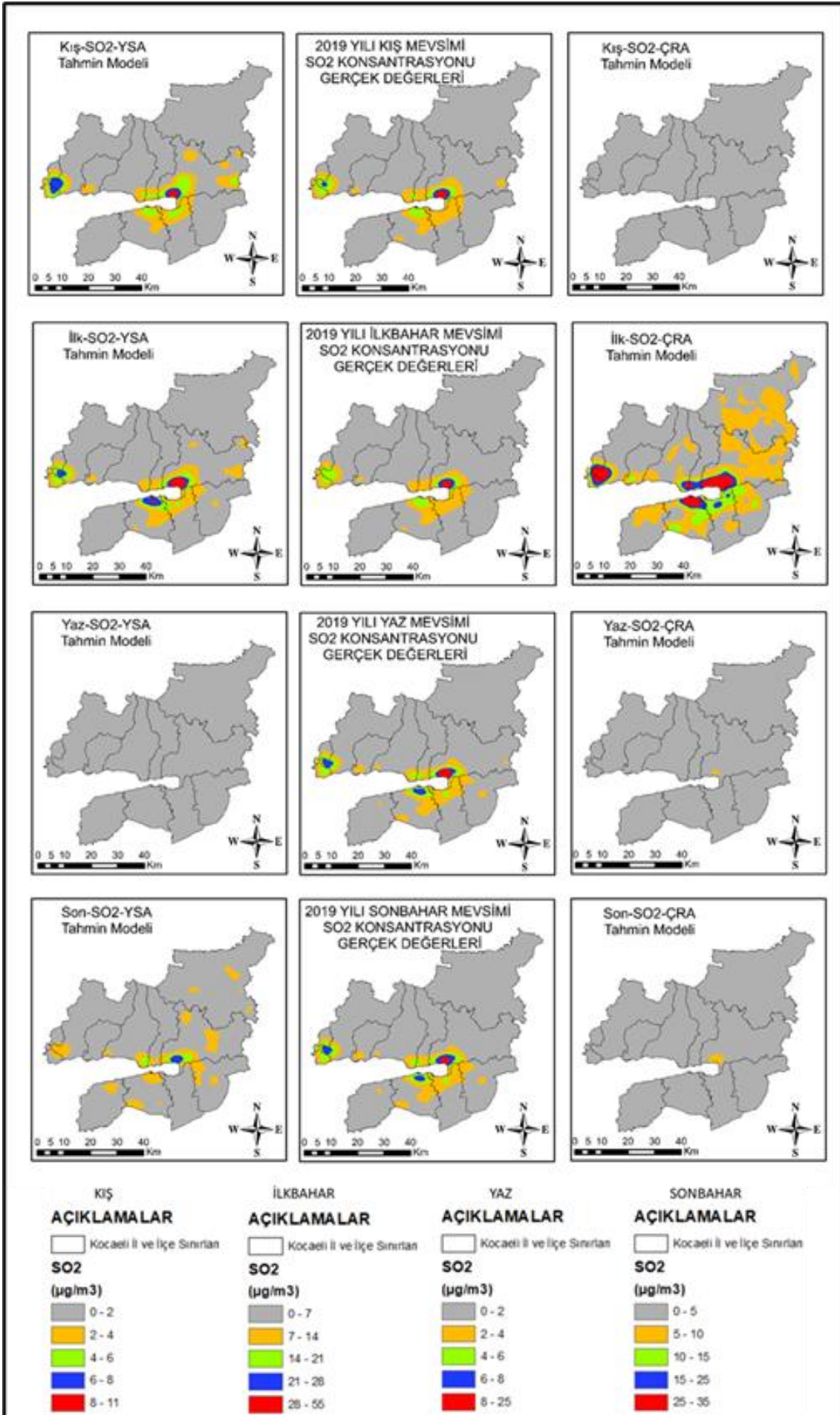
Bu çalışmanın amacı hava kirliliğinin izlenmesi, kirlilik verilerinin toplanması, haritalarının oluşturulması ve bunlar doğrultusunda hava kirliliği tahmin modellerinin oluşturularak insan ve çevre sağlığının korunması açısından hava kirliliği hakkında öngörü için karar vericilere fikir vererek tedbirlerin

alınmasını ve planlamaların yapılmasını kolaylaştırmaktır. Bu amaç doğrultusunda literatürde hava kirliliğine etki eden çalışmalardan esinlenerek çalışmada kullanılacak olan veriler elde edildiği gibi yine literatürde hava kirliliğine yönelik tahmin çalışmalarında kullanılan yöntemlerden olan çoklu doğrusal regresyon analizi ve yapay sinir ağı yöntemleri tercih edildi. Hava kirliliğine etki eden çalışmalarda meteorolojik veriler, kirlenici konsantrasyonlarının yanında trafik emisyonları, kentleşme, sanayileşme, topografik ve demografik verilerinde etkilediği söylenmekte bunun aksine hava kirliliğine yönelik yapılan tahmin çalışmalarında genellikle meteorolojik ve kirlenici konsantrasyonlarının kullanıldığı görülmektedir. Bu çalışmada ise meteorolojik ve kirlenici parametrelerle birlikte CBS kullanılarak ve UA teknolojileri yardımıyla mekânsal veriler (bina izi, sanayi alanları, yol ağı), topografik veri (SYM), demografik veri (nüfus) ile birlikte kullanılarak literatürdeki bu boşluğun doldurulması düşünüldü. Başka bir açıdan bakıldığında ise literatürde hava kirliliği tahmin modellerinin genellikle saatlik veya günlük olarak tahmin edilirken bu çalışmada 1 yıl önceden mevsimsel ortalamaların tahmin edilmesi üzerine çalışıldı. Bu amaç doğrultusunda ise tahminler sonucunda hava kirliliği konusunda kırmızı alarm verilecek kadar tehlikeli bir durumun 1 yıl önceden öngörülmesi tedbirlerin alınması, planlamaların yapılarak hayata geçirilmesi açısından hayati önem taşımaktadır.

Çalışmada belirli sayıdaki istasyonlardan veriler temin edildiği için çalışmaların güvenilirliği ve doğruluğu açısından bu istasyonların sayısının artırılması model performanslarını artıracak düşünülmemektedir. Son dönemlerde hava kirliliği verileri sürekli olarak uzaktan algılama teknolojisiyle birlikte uydu verileri yardımıyla elde edilebilmektedir. Noktasal olarak istasyonlardan alınarak yapılan bu çalışmalar uydu verilerinden yararlanılarak alınan hava kirliliği verileri ile çalışma sonuçları karşılaştırılabilir. Ayrıca çalışma alanında sürekli verinin sağlanması açısından birçok veri setine TMA tekniği ile enterpolasyona başvuruldu. Diğer enterpolasyon yöntemleri denenerken de bunların performanslarına bakılıp daha iyi tahmin modelleri elde edilebileceği düşünülmektedir. Ayrıca bu çalışma kapsamında regresyon analizinde çoklu doğrusal regresyon yönteminde stepwise metodu kullanıldığı gibi diğer regresyon yöntemleri (lojistik, ridge, lasso vb.) de kullanılarak oluşturulan modellerin performansları incelenebilir. Bunların yanında tahmin modellerinin oluşturulmasında kullanılan diğer bir yöntem olan yapay sinir ağlarında hiper parametrelerin seçimi konusunda manuel denemeler yapmak yerine en iyi modeli yakalamak adına hem zamandan tasarruf sağlayacak hem de daha farklı kombinasyonların denemesi açısından grid search gibi yöntemlerin kullanılarak ileri çalışmalarda kolaylık sağlanabilir.



Şekil 22. PM₁₀ konsantrasyonuna ait tahmin ve gerçek değerlerin yoğunluk haritaları



Şekil 23. SO₂ konsantrasyonuna ait tahmin ve gerçek değerlerin yoğunluk haritaları

5. Sonuç

Bu çalışma sonucunda hava kirliliği tahmin modellerinin oluşturulmasında YSA, ÇRA modellerine göre performansının daha iyi olduğu, modellerde mekânsal verilerin (bina izi, sanayi alanları, yol ağı), topografik (SYM) ve demografik (nüfus) verilerin kullanılabilirliği gösterilmiştir. Bu verilerin modellerde kullanılabilir veri formatına getirilmesinde CBS teknolojilerinin aktif bir şekilde kullanılabilirliği ve hava kirliliğinin haritalanmasında ve yorumlanmasında CBS'nin kullanılabilirliği gösterilmiştir. Ayrıca 1 yıl önceden yapılan tahminler ile yerel yönetimler ve karar vericiler açısından insan ve çevre sağlığının korunması amacıyla hava kirliliğinin önlenmesi için tedbirlerin alınması ve planlamaların yapılabilmesi; ulaşım, yerleşim, enerji üretimi, sanayileşme gibi birçok alanı kapsayan uzun vadeli politikaların geliştirilmesine ve bunların sürdürülmesinde altlık oluşturulabileceği; sürdürülebilir kentler oluşturma amacına büyük katkı sağlayabileceği sonuçlarına ulaşılmıştır.

Bilgilendirme/Teşekkür

Bu çalışma, Burak KOTAN tarafından 2022 yılında hazırlanan, Kocaeli Üniversitesi Fen Bilimleri Enstitüsü Jeodezi ve jeoinformasyon ABD'da Prof. Dr. Arzu Erener'in danışmanlığında yürütülen yüksek lisans tezinden üretilen sonuçlara dayanarak oluşturulmuştur.

Araştırmacıların katkı oranı

Burak Kotan: Kavramsallaştırma, literatür taraması, verilerin analiz ve değerlendirilmesi, sonuçların incelenmesi, **Arzu Erener:** Kavramsallaştırma, verilerin temin edilmesi, yazım denetimi, içerik açısından makalenin kontrol edilmesi, sonuçların incelenmesi

Çatışma Beyanı

Herhangi bir çıkar çatışması bulunmamaktadır.

Kaynakça

- Adachi, K. (2020). *Matrix-Based Introduction to Multivariate Data Analysis*, Singapore, Springer.
- Akyürek, Ö., Arslan, O., & Karademir, A. (2013). SO₂ ve PM₁₀ hava kirliliği parametrelerinin CBS ile konumsal analizi: Kocaeli örneği, TMMOB Coğrafi Bilgi Sistemleri Kongresi, 11-13 Kasım 2013, Ankara, Türkiye.
- Alvarez, H., Echeverria, R., Alvarez, P., & Krupa, S. (2013). Air quality standards for particulate matter (PM) at high altitude cities, *Environmental Pollution*, 173, 255-256.
- Aydinoğlu, A. Ç., Bovkır, R., & Bulut, M. (2022). Akıllı şehirlerde büyük coğrafi veri yönetimi ve analizi: hava kalitesi örneği. *Geomatik*, 7(3), 174-186.
- Bai, L., Jiang, L., Yang, D., & Liu, Y. (2019). Quantifying the spatial heterogeneity influences of natural and socioeconomic factors and their interactions on air pollution using the geographical detector method: A

- case study of the Yangtze River Economic Belt, China". *Journal of Cleaner Production*, 232, 692-704.
- Bannister, E., Cai, X., Zhong, J., & MacKenzie, A. (2021). Neighbourhood-scale flow regimes and pollution transport in cities, *Boundary-Layer Meteorology*, 179(2), 259-289.
- Campos, D. S., de Souza Tadano, Y., Alves, T. A., Siqueira, H. V., & de Nóbrega Marinho, M. H. (2020). Unorganized machines and linear multivariate regression model applied to atmospheric pollutant forecasting. *Acta Scientiarum. Technology*, 42, e48203-e48203.
- Caselli, M., Trizio, L., de Gennaro, G., & Ielpo, P. (2009). A Simple Feedforward Neural Network for the PM₁₀ Forecasting: Comparison with a Radial Basis Function Network and a Multivariate Linear Regression Model, *Water Air and Soil Pollution*, 201(1-4), 365-377.
- Çetin, M., Onac, A., Sevik, H., & Sen, B. (2019). Temporal and regional change of some air pollution parameters in Bursa, *Air Quality, Atmosphere & Health*, 12(3), 311-316.
- Chen, Y., Shi, R., Shu, S., & Gao, W. (2013). Ensemble and enhanced PM₁₀ concentration forecast model based on stepwise regression and wavelet analysis, *Atmospheric Environment*, 74, 346-359.
- Çuhadar, M., & Kayacan, C. (2005). Yapay sinir ağları kullanılarak konaklama işletmelerinde doluluk oranı tahmini: Türkiye'deki konaklama işletmeleri üzerine bir deneme, *Anatolia: Turizm Araştırmaları Dergisi*, 16(1), 24-30.
- Dahari, N., & Wahid, H. (2017). Determination of modeling parameters for a low cost air pollution measurement system using feed forward neural networks., Editors: Mohamed Ali MS, Wahid H, Mohd Subha NA, Sahlan S, Yunus MA, Wahap, AR. *Modeling, Design and Simulation of Systems*, 685-696, Singapore, Springer Nature.
- Du, M., Liu, W., & Hao, Y. (2021). Spatial correlation of air pollution and its causes in Northeast China, *International Journal of Environmental Research and Public Health*, 18(21).
- Dutta, A., & Jinsart, W. (2021). Air pollution in Indian cities and comparison of MLR, ANN and CART models for predicting PM₁₀ concentrations in Guwahati, India, *Asian Journal of Atmospheric Environment*, 15(1).
- Erener, A., Sarp, G., & Yıldırım, Ö. (2019). Seasonal air pollution investigation and relation analysis of air pollution parameters to meteorological data (Kocaeli/Turkey). Editors: El-Askary HM, Lee S, Heggy E, Pradhan B. *Advances in Remote Sensing and Geo Informatics Applications*, 355-358, Cham, Switzerland, Springer.
- Field, A. (2009). *Discovering Statistics Using SPSS*, Dubai, SAGE Publications.
- Forsyth, D. (2019). *Applied Machine Learning*. Cham, Springer.
- George, D., & Mallery, P. (2010). *SPSS for Windows step by step: a simple guide and reference*, Boston, Allyn & Bacon.
- Ghaemi, Z., Alimohammadi, A., & Farnaghi, M. (2018). LaSVM-based big data learning system for dynamic

- prediction of air pollution in Tehran, *Environmental Monitoring and Assessment*, 190(5).
- Ghatak, A. (2017). *Machine Learning with R*, Singapore, Springer.
- Gonzalez-Enrique, J., Ruiz-Aguilar, J. J., Moscoso-Lopez, J. A., Urda, D., Deka, L., & Turias, I. J. (2021). Artificial neural networks, sequence-to-sequence LSTMs, and exogenous variables as analytical tools for NO₂ (air pollution) forecasting: A case study in the bay of algeciras (Spain), *Sensors*, 21(5), 1770.
- Haglin, J. M., Jimenez, G., & Eltorai, A. E. (2019). Artificial neural networks in medicine, *Health and Technology*, 9(1), 1-6.
- Jayalakshmi, T., & Santhakumaran, A. (2011). Statistical Normalization and Back Propagation for Classification, *International Journal of Computer Theory and Engineering*, 3(1), 89-93.
- Jiang, L., Zhou, H., Bai, L., & Zhou, P. (2018). Does foreign direct investment drive environmental degradation in China? An empirical study based on air quality index from a spatial perspective, *Journal of Cleaner Production*, 176, 864-872.
- Jumaah, H. J., Ameen, M. H., Kalantar, B., & Rizzei, H. M. (2019). Air quality index prediction using IDW geostatistical technique and OLS-based GIS technique in Kuala Lumpur, Malaysia, *Geomatics, Natural Hazards and Risk*, 10(1), 2185-2199.
- Kaplan, G., & Avdan, Z. Y. (2020). Space-borne air pollution observation from sentinel-5p tropomi: Relationship between pollutants, geographical and demographic data. *International Journal of Engineering and Geosciences*, 5(3), 130-137.
- Kolasa-Wiecek, A., & Suszanowicz, D. (2019). Air pollution in European countries and life expectancy—modelling with the use of neural network, *Air Quality Atmosphere and Health*, 12(11), 1335-1345.
- Lee, C. F., Chen, H. Y., & Lee, J. (2019). *Financial Econometrics, Mathematics and Statistics*, New York, Springer.
- Lewis, C. D. (1982). *Industrial and Business Forecasting Methods*, Londra, Butterworths Publishing.
- Li, L., Qian, J., Ou, C. Q., Zhou, Y. X., Guo, C., & Guo, Y. (2014). Spatial and temporal analysis of air pollution index and its timescale-dependent relationship with meteorological factors in Guangzhou, China, 2001–2011, *Environmental Pollution*, 190, 75-81.
- Lin, X., & Wang, D. (2016). Spatiotemporal evolution of urban air quality and socioeconomic driving forces in China, *Journal of Geographical Sciences*, 26, 1533–1549.
- Liu, H., Fang, C., Zhang, X., Wang, Z., Bao, C., & Li, F. (2017). The effect of natural and anthropogenic factors on haze pollution in Chinese cities: A spatial econometrics approach. *Journal of Cleaner Production*, 165, 323-333.
- Lu, D., Xu, J., Yang, D., & Zhao, J. (2017). Spatio-temporal variation and influence factors of PM_{2.5} concentrations in China from 1998 to 2014, *Atmospheric Pollution Research*, 8(6), 1151-1159.
- Maleki, H., Sorooshian, A., Goudarzi, G., Baboli, Z., Birgani, Y., & Rahmati, M. (2019). Air pollution prediction by using an artificial neural network model, *Clean Technologies and Environmental Policy*, 21(6), 1341-1352.
- Mansor, A., Abdullah, S., Che Dom, N., Mohd Napi, N., Ahmed, A., Ismail, M., & Zulkifli, M. (2021). Three-hour-ahead of multiple linear regression (MLR) models for particulate matter, *International Journal of Design & Nature and Ecodynamics*, 16(1), 53-59.
- Moustris, K. P., Ziomas, I. C., & Paliatsos, A. G. (2010). 3-Day-Ahead Forecasting of Regional Pollution Index for the Pollutants NO₂, CO, SO₂, and O₃ Using Artificial Neural Networks in Athens, Greece, *Water, Air, & Soil Pollution*, 209(1), 29-43.
- Moustris, K., Nastos, P., Larissi, I., & Paliatsos, A. (2012). Application of multiple linear regression models and artificial neural networks on the surface ozone forecast in the greater Athens Area, Greece, *Advances in Meteorology*, 2012.
- Özcan, H. K., Şahin, Ü., Bayat, C., & Uçan, O. N. (2006). Modeling of trophospheric ozone (O₃) concentrations using cellular neural networks, *Journal of the Faculty of Engineering and Architecture of Gazi University*, 21(2), 239-245.
- Tsagris, M., & Pandis, N. (2021). Multicollinearity, *American Journal of Orthodontics and Dentofacial Orthopedics*, 159(5), 695-696.
- URL-1:<https://www.harita.gov.tr/urun/il-ve-ilce-yuz-olcumleri/176> Erişim tarihi: 16 Mart 2022
- URL-2:<https://data.tuik.gov.tr/Bulten/Index?p=Adrese-Dayali-Nufus-Kayit-Sistemi-Sonuclari-2020-37210> Erişim tarihi: 16 Mart 2022
- URL-3:<http://www.kocaeli.gov.tr/kocaeli-ekonomisinde-sanayinin-yeri> Erişim tarihi: 16 Mart 2022
- URL-4:<http://www.kocaeli.gov.tr/lojistigin-merkezinde-bir-kent-kocaeli> Erişim tarihi: 16 Mart 2022
- Vorapracha, P., Phonprasert, P., Khanaruksombat, S., & Pijarn, N. (2015). A Comparison of Spatial Interpolation Methods for predicting concentrations of Particle Pollution (PM₁₀), *International Journal of Chemical, Environmental & Biological Sciences*, 3(4), 302-306.
- Voss, D. S. (2005) Multicollinearity. Editors: Kempf-Leonard K. *Encyclopedia of Social Measurement*, 759-770, New York, Elsevier.
- Wang, P., Liu, Y., Qin, Z., & Zhang, G. (2015). A novel hybrid forecasting model for PM₁₀ and SO₂ daily concentrations, *Science of the Total Environment*, 505, 1202-1212.
- Witt, S. F., & Witt, C. A. (1992). *Modeling and Forecasting Demand in Tourism*, Londra, Academic Press.
- Yadav, V., & Nath, S. (2020a). Artificial neural networks based condition monitoring of air pollutants for Allahabad cities in India, Editors: Malik H, Iqbal A, Yadav AK. *Soft Computing in Condition Monitoring and Diagnostics of Electrical and Mechanical Systems*, 423-437, Singapore, Springer.
- Yadav, V., & Nath, S. (2020b). Novel Application of Artificial Neural Network Techniques for Prediction of Air Pollutants Using Stochastic Variables for Health Monitoring: A Review. Editors: Malik H, Iqbal A, Yadav AK. *Soft Computing in Condition Monitoring and Diagnostics of Electrical and Mechanical Systems*:

Novel Methods for Condition Monitoring and Diagnostics, 231-245, Singapore, Springe.

Yılmaz, A. (2017). İklim parametrelerinin hava kirliliği parametreleri üzerine etkisi: Bolu ili örneği, *Journal of Current Researches on Social Sciences*, 7(2), 413-436.

Yüksek, A. G., Bircan, H., Zontul, M., & Kaynar, O. (2007). Sivas ilinde yapay sinir ağları ile hava kalitesi modelinin oluşturulması üzerine bir uygulama, *Cumhuriyet Üniversitesi İktisadi ve İdari Bilimler Dergisi*, 8(1), 97-112.



© Author(s) 2023. This work is distributed under <https://creativecommons.org/licenses/by-sa/4.0/>