

Spatial modeling of airborne particles (PM_{2.5} and PM₁₀) in Tehran city using convolutional neural network

Bashardoost, A.¹  | Mesgari, M.S.²  

1. Doctoral student of civil engineering and mapping, geographic information systems, K. N. Toosi University, Tehran, Iran.
2. Associate Professor of Department of Geodesy and Geomatics, K. N. Toosi University of Technology, Tehran, Iran.

Corresponding Author E-mail: mesgari@kntu.ac.ir

(Received: 12 Des 2023, Revised: 27 Feb 2024, Accepted: 16 Mar 2024, Published online: 16 Mar 2024)

Abstract

Air pollution poses significant risks to human health and the environment, which makes it necessary to create effective strategies for air quality management. This study presents an approach for air quality management in Tehran using the Convolutional Neural Network (CNN) algorithm. The proposed method provides the possibility of spatial modeling and preparation of risk maps of two important air pollutants, namely particulate matter 2.5 (PM_{2.5}) and particulate matter 10 (PM₁₀). To develop this air pollution model, the data available in the database containing the annual average of two pollutants from 2012 to 2022 were used. In this model, various parameters affecting air pollution including altitude, humidity, distance to industrial areas, normalized difference index of plants (NDVI), population density, precipitation, distance to the street, temperature, traffic volume, wind direction, and wind speed are considered. Taken and spatial modeling of two pollutants using CNN has been done. The evaluation of the model was done using different evaluation criteria, and the findings showed that the R-squared (R²) values in this model for PM_{2.5} and PM₁₀ pollutants are 0.889 and 0.972, respectively. The accuracy of the risk map was evaluated using relative operating characteristic (ROC) for two pollutants, and the findings showed that the CNN model has an acceptable accuracy in producing the pollution risk map. In general, risk maps provide useful information about geographic areas with high pollution risks and help in decision-making and targeted pollution reduction efforts.

Keywords: Air pollutants, Spatial modeling, Risk map, Deep learning.

Cite this article: Bashardoost, A., & mesgari, M. S. (2024). Spatial modeling of airborne particles (PM_{2.5} and PM₁₀) in Tehran city using convolutional neural network. *Journal of the Nivar*, 48(124-125), 31-49. DOI: <https://doi.org/10.30467/nivar.2024.430255.1276>

E-mail: (1) abedbashardoost@gmail.com



Publisher: Iran Meteorological Organization
DOI: <https://doi.org/10.30467/nivar.2024.430255.1276>

Print ISSN: 1735-0565
Online ISSN: 2645-3347

مدل سازی مکانی ذرات معلق هوا (PM_{2.5} و PM₁₀) در شهر تهران با استفاده از شبکه عصبی کانولوشن

عابد بشردوست^۱ | محمد سعدی مسگری^۲ ✉

۱. دانشجوی دکتری مهندسی عمران نقشه برداری، گرایش سیستم های اطلاعات جغرافیایی، دانشگاه خواجه نصیرالدین طوسی، تهران، ایران.
۲. دانشیار دانشکده نقشه برداری، دانشگاه صنعتی خواجه نصیرالدین طوسی، تهران، ایران.

رایانامه نویسنده مسئول: mesgari@kntu.ac.ir

(دریافت: ۱۴۰۲/۰۹/۲۱، بازنگری: ۱۴۰۲/۱۲/۰۸، پذیرش: ۱۴۰۲/۱۲/۲۶، انتشار آنلاین: ۱۴۰۲/۱۲/۲۶)

چکیده

آلودگی هوا خطرات قابل توجهی را برای سلامت انسان و محیط زیست به همراه دارد، که این موضوع ایجاد راهبردهای موثر در مدیریت کیفیت هوا را ضروری می‌سازد. این مطالعه رویکردی برای مدیریت کیفیت هوا شهر تهران با استفاده از الگوریتم شبکه عصبی کانولوشن (CNN) ارائه می‌دهد. روش پیشنهادی امکان مدل سازی مکانی و تهیه نقشه خطر دو آلاینده هوایی مهم یعنی ذرات معلق ۲/۵ میکرومتر (PM_{2.5}) و ذرات معلق ۱۰ میکرون (PM₁₀) را فراهم می‌نماید. برای توسعه این مدل آلودگی هوا، از داده‌های موجود در پایگاه داده حاوی میانگین سالانه دو آلاینده از سال ۲۰۱۲ تا ۲۰۲۲ استفاده شده است. در این مدل، پارامترهای مختلف موثر بر آلودگی هوا شامل ارتفاع، رطوبت، فاصله تا مناطق صنعتی، شاخص تفاوت نرمال شده گیاهان (NDVI)، چگالی جمعیت، بارش، فاصله تا خیابان، دما، حجم ترافیک، جهت باد و سرعت باد در نظر گرفته شده و مدل سازی مکانی دو آلاینده با استفاده از CNN انجام گردیده است. ارزیابی مدل با استفاده از معیارهای ارزیابی مختلف انجام گردید. نتایج نشان داد مقادیر ضریب تعیین (R²) -در این مدل برای آلاینده‌های PM_{2.5} و PM₁₀ به ترتیب ۰/۸۸۹ و ۰/۹۷۲ است. دقت نقشه خطر با استفاده از مساحت زیر منحنی عملکرد مشخص کننده (ROC) برای دو آلاینده ارزیابی گردید که نتایج نشان دادند مدل CNN در تولید نقشه خطر آلودگی دقت قابل پذیرشی دارد. به طور کلی، نقشه‌های خطر اطلاعات مفیدی درباره مناطق جغرافیایی با خطرات آلودگی بالا ارائه می‌دهند و در تلاش‌ها به منظور تصمیم سازی و کاهش هدفمند آلودگی کمک می‌نمایند.

کلیدواژه‌ها: آلاینده‌های هوا، مدل سازی مکانی، نقشه خطر، یادگیری عمیق.

۱. مقدمه

حفظ سلامت عمومی را افزایش داده است. بر این مبنا نقش نقشه‌های آلودگی، راهبردها و سیاست‌های تحلیل و بهبود کیفیت هوا در مناطق شهری بسیار حائز اهمیت است (پرا، ۲۰۱۷). درک توزیع مکانی آلودگی هوا در مناطق شهری برای پرداختن به مسائل سلامت جمعیت بسیار حائز اهمیت

کاهش کیفیت هوا در مناطق شهری به دلیل فعالیت‌های صنعتی، دستگاه‌های گرمایشی و تکیه بر سوخت‌های فسیلی برای حمل و نقل، یکی از مشکلات اساسی است (گالیمووا و همکاران، ۲۰۲۲). گسترش سریع مناطق شهری و رشد صنایع در کشورهای در حال توسعه، کیفیت هوا را به شدت تحت تأثیر قرار داده و نیاز به نظارت و تنظیمات سختگیرانه برای

استناد: بشردوست، عابد و مسگری، محمد سعدی (۱۴۰۲). مدل سازی مکانی ذرات معلق هوا (PM_{2.5} و PM₁₀) در شهر تهران با استفاده از شبکه عصبی کانولوشن. مجله

نیوار، ۴۸(۱۲۴-۱۲۵)، ۳۱-۴۹. DOI: <https://doi.org/10.30467/nivar.2024.430255.1276>

رایانامه: (۱) abedbashardost@gmail.com



اثرات آنها بر سلامت انسان و اکوسیستم‌ها مورد توجه قرار گرفته است. بیماری‌هایی مانند اختلالات تنفسی، سرطانی، بیماری‌های قلبی و عروقی، تأثیر بر سازوکارهای دفاعی و مرگ زودرس از عمده‌ترین عوارض ذرات معلق بر روی انسان هستند (مامیک و همکاران، ۲۰۲۳؛ زهتاب یزدی و همکاران، ۲۰۲۱). این موضوع سبب شده است تا محققان زیادی، مطالعات مختلفی راجع به پراکنش آلاینده ذرات معلق انجام بدهند. کشاورزی شیرازی و یقار (۱۳۸۵)، در کار پژوهشی خود با عنوان «مدل سازی و تحلیل پراکنش آلاینده‌های منتشره از منابع متحرک آلودگی هوا براساس GIS در شهر تهران» انتشار آلاینده‌های منواکسیدکربن و ذرات معلق را که مربوط به منابع متحرک می‌باشند، مدل سازی نموده‌اند. بهاری و همکاران (۱۳۹۴)، مدل سازی مکانی غلظت آلاینده ذرات معلق ۲/۵ میکرومتر (PM_{2.5}) را با بهره‌گیری از مدل رگرسیون وزن‌دار جغرافیایی از خانواده مدل‌های آماری محلی انجام دادند.

پیش‌بینی ترکیب هوا به طور قابل توجهی در مدیریت کیفیت هوا کمک می‌کند. با این حال، پیش‌بینی کیفیت هوا به دلیل پیچیدگی فرایندها و جفت شدن قوی در بسیاری از پارامترها که بر عملکرد مدل‌سازی تأثیر می‌گذارد، همچنان یک چالش باقی می‌ماند. به‌طور کلی، ترکیب و غلظت ذرات معلق بسیار متفاوت است. این تنوع به احتمال زیاد ناشی از عوامل طبیعی و محیطی مانند منابع انسانی یا طبیعی، دما و تغییرات فصلی و موقعیت جغرافیایی می‌باشد (مامیک و همکاران، ۲۰۲۳). تجزیه و تحلیل ذرات معلق برای پایش کیفیت هوا به داده‌های تفکیک مکانی و زمانی بالاتر نیاز دارد. شبکه عصبی کانولوشن (CNN⁴) نوعی از معماری شبکه عصبی است که برای یادگیری عمیق استفاده می‌شود. این شبکه از لایه‌های پنهان تشکیل شده است که مهم‌ترین آنها لایه‌های CNN می‌باشد که نقاط داده را برای تشخیص و سنجش الگو در هم می‌پیچند. CNN با الهام از قشر بصری مغز انسان، برای

است. سامانه اطلاعات مکانی (GIS¹) در درک موقعیت آلودگی‌ها نقش حیاتی داشته و با ایجاد نقشه‌های مکانی از پدیده آلودگی هوا، مدل‌سازی و تجسم از آن را فراهم می‌نماید. این نقشه‌ها برای نظارت بر کیفیت هوای شهری ارزشمند بوده و می‌توانند توسط افراد غیر تخصصی به راحتی تفسیر گردند (گالیور و بریگز، ۲۰۱۱). تکنیک‌های GIS، از جمله تحلیل مکانی و تداخل، ابزارهای قدرتمندی برای نقشه‌برداری آلودگی و پیش‌بینی شرایط کیفیت هوای آینده هستند. علاوه بر این، از سنجش از راه دور (RS²) برای اندازه‌گیری سطح آلودگی هوا استفاده می‌شود که دسترسی آسان، تفکیک مکانی بالا و هزینه کمی دارد. RS اطلاعات معناداری در مورد کیفیت هوای شهری ارائه می‌دهد و به‌عنوان یک ابزار ضروری در درک و مدیریت آلودگی هوا محسوب می‌شود (لو و همکاران، ۲۰۱۹؛ رضوی ترمه و همکاران، ۲۰۲۱).

روش‌های مختلفی مانند رگرسیون کاربری اراضی (LUR³)، یادگیری ماشین و یادگیری عمیق برای نظارت و مدل‌سازی آلودگی هوا استفاده می‌شود. درحالی‌که رگرسیون کاربری اراضی در تعمیم به مکان‌های جدید به داده‌های تجربی نیاز دارد، یادگیری ماشین سرعت پردازش سریع و توانایی برقراری ارتباطات پیچیده بدون درک عمیق از خواص آلاینده‌ها را دارد. مطالعاتی که از الگوریتم‌های یادگیری ماشین استفاده می‌کنند، نقشه‌های قابل اعتمادی برای نظارت بر کیفیت هوا ارائه داده‌اند (بککار و همکاران، ۲۰۲۱؛ رن و همکاران، ۲۰۲۰؛ اشتاینینگر و همکاران، ۲۰۲۰). الگوریتم‌های یادگیری عمیق به دلیل انعطاف‌پذیری و دقت پیش‌بینی، به‌خصوص برای داده‌های بزرگ محبوب هستند، با این حال دارای مشکلاتی مانند بیش‌برازنده شدن و پیچیدگی مدل هستند (هو و همکاران، ۲۰۲۱).

ذرات معلق، به‌ویژه با قطر کمتر از ۲/۵ و ۱۰ میکرومتر در سال‌های اخیر به موضوعات تحقیقاتی مهمی تبدیل شده‌اند و

³ Land Use Regression

⁴ Convolutional Neural Network

¹ Geospatial Information System

² Remote Sensing

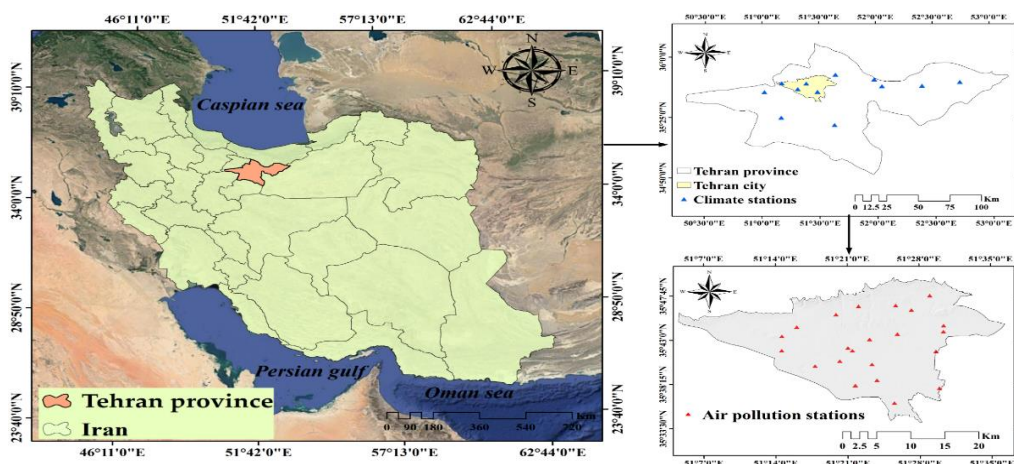
۲-۱. منطقه مطالعه

تهران پایتخت ایران در عرض جغرافیایی ۳۵ درجه و ۳۶ دقیقه تا ۳۵ درجه و ۴۴ دقیقه عرض شمالی و ۵۱ درجه و ۱۷ دقیقه تا ۵۱ درجه و ۳۳ دقیقه طول شرقی با ارتفاعی بین ۱۰۳۲ تا ۱۸۳۲ متر از سطح دریا واقع شده است. این شهر به عنوان پرجمعیت ترین شهر ایران در دو دهه گذشته با آلودگی شدید هوای محیطی مواجه شده است. این آلودگی به دلایلی نظیر توسعه ناسازگار صنعتی و شهرنشینی، رشد روزافزون ناوگان حمل و نقل و انتشار آلاینده های آن ها، استانداردهای ملی نامناسب کیفیت هوای محیطی و طوفان های گرد و خاک رخ می دهد. تهران ۲۰ درصد انرژی ایران را مصرف نموده و جغرافیای آن که توسط رشته کوه احاطه شده است، از پراکندگی هوای آلوده به ویژه در فصل زمستان جلوگیری می نماید. ساختار فضای شهری ارتباط تنگاتنگی با سیستم حمل و نقل داشته و آمارها، مصرف بالای زمین را نشان می دهد که این موضوع منجر به افزایش مسافت ها و نیازهای حمل و نقل برای فعالیت های اداری، آموزشی و اوقات فراغت می شود (یوسفیان و همکاران، ۲۰۲۰). منطقه مورد مطالعه، در شکل یک نشان داده شده است که ایستگاه های پایش کنترل کیفیت هوا و ایستگاه های هواشناسی را نشان می دهد.

پردازش داده های یکپارچه همانند تصاویر، صدا و ویدئو مناسب می باشد. در این شبکه. به جای استفاده از توابع فعال سازی، از توابع CNN و ادغام استفاده می شود (چاوهان و همکاران، ۲۰۲۱ و لی و همکاران، ۲۰۲۱).

هدف این تحقیق ارائه یک مدل مکانی آلودگی هوای قابل اطمینان با استفاده از یک CNN است. هدف این است که با تولید نقشه های خطر با تفکیک بالا، بینش های نوآورانه ای برای مدیریت کیفیت هوا ارائه نماییم. با توجه به نقاط قوت تجزیه و تحلیل مکانی CNN، این تحقیق به دنبال ارائه پیش بینی های دقیق تر مکانی از غلظت آلاینده های ذرات معلق هوا است. برای توسعه این مدل آلودگی هوا، از داده های موجود در پایگاه داده حاوی میانگین سالانه دو آلاینده از سال ۲۰۱۲ تا ۲۰۲۲ استفاده می شود. در این مدل، پارامترهای مختلف مؤثر بر آلودگی هوا شامل ارتفاع، رطوبت، فاصله تا مناطق صنعتی، شاخص تفاوت نرمال شده گیاهان، چگالی جمعیت، بارش، فاصله تا خیابان، دما، حجم ترافیک، جهت باد و سرعت باد در نظر گرفته شده و مدل سازی مکانی دو آلاینده با استفاده از CNN انجام می گردد. ارزیابی مدل با استفاده از معیارهای ارزیابی مختلف انجام می گردد. دقت نقشه خطر با استفاده از مساحت زیر منحنی عملکرد مشخص کننده ROC¹ برای دو آلاینده ارزیابی می گردد.

۲. مواد و روش ها



شکل ۱. منطقه مورد مطالعه با ایستگاه های آلودگی هوا و ایستگاه های همدیدی

۲-۵. اهمیت ویژگی عوامل با استفاده از روش

GeoDetector

GeoDetector روشی است که برای شناسایی و بهره‌برداری از تفاوت‌های جغرافیایی استفاده می‌شود و تعداد نیروهای محرک، عوامل تأثیرگذار و تعاملات چندعاملی را تعیین می‌کند (وانگ و خو، ۲۰۱۷). این روش شامل رویه‌های پیچیده تنظیم پارامتر نمی‌شود و همچنین به فرض‌های کلاسیک تکنیک‌های آماری خطی محدود نیست. در این روش، اگر یک متغیر مستقل تأثیر قابل توجهی روی یک متغیر مستقل دیگر داشته باشد، توزیع مکانی آن نشان داده خواهد شد (ژانگ و همکاران، ۲۰۲۲). GeoDetector دارای چهار عملکرد متمایز است: تشخیص عامل، تشخیص تعامل، تشخیص خطر و تشخیص محیط زیست (وانگ و خو، ۲۰۱۷). آشکارساز عامل برای تشخیص ناهمگونی مکانی متغیر وابسته Y و ارزیابی توانایی توضیحی متغیر مستقل X بر Y استفاده می‌شود. آشکارساز عامل اثربخشی مقدار q مشتق شده را با در نظر گرفتن رابطه بین متغیرها ارزیابی می‌نماید. مقادیر q به دست آمده از GeoDetector امکان اندازه‌گیری تغییرات مکانی و تحلیل عاملی را فراهم می‌نماید (جیا و همکاران، ۲۰۲۱) مقدار q_x از معادله ۲ به دست می‌آید.

$$q_x = 1 - \frac{SSW}{SST} \quad (2)$$

$$SSW = \sum_{h=1}^l N_h \sigma_h^2 \quad (3)$$

$$SST = N\sigma^2 \quad (4)$$

که در آن SSW^3 مجموع مربعات درون گروهی (محلی) است، درحالی که SST^4 نشان‌دهنده واریانس سراسری است. متغیر h مخفف تعداد دسته‌های متغیر مستقل است، N_h و N به ترتیب تعداد واحدهای منطقه h و تعداد کل منطقه را نشان می‌دهند. متغیر σ_h^2 نشان‌دهنده واریانس Y در منطقه h است و σ^2 نشان‌دهنده واریانس جهانی Y در کل منطقه است.

۲-۲. داده‌های مورد استفاده (ذرات معلق PM_{2.5}

و PM₁₀)

از اول ژانویه ۲۰۱۲ تا یک ژانویه ۲۰۲۲، میانگین غلظت سالانه ذرات معلق PM_{2.5} و PM₁₀ در تهران در ۲۳ ایستگاه پایش کیفیت هوا در سطح شهر بررسی شد.

۲-۳. عوامل مؤثر

عوامل مؤثر بر غلظت ذرات معلق PM_{2.5} و PM₁₀ در هوا در این تحقیق شامل داده‌های هواشناسی (بارندگی، دما، رطوبت، جهت باد و سرعت باد)، ارتفاع، شاخص تفاوت نرمال شده گیاهان (NDVI¹)، فاصله از خیابان، فاصله از مناطق صنعتی، حجم ترافیک و تراکم جمعیت است.

۲-۴. تجزیه و تحلیل هم خطی چندگانه

(Multicollinearity)

یکی از مفاهیم آماری در تحلیل رگرسیون، هم خطی چندگانه است. هم خطی چندگانه یعنی بین چند متغیر مستقل رابطه خطی وجود دارد. مشکل چندخطی به دلیل همبستگی (رابطه قوی) بین پیش‌بینی‌کننده‌ها و عدم استقلال آنها در یک مجموعه داده به وجود می‌آید. در مدل‌های به دست آمده از این داده‌ها، اگر چندخطی بودن بررسی نشود، ممکن است منجر به تحلیل‌های اشتباه شود (گارگ و تای، ۲۰۱۳). ضریب تورم واریانس (VIF²) روشی است که برای شناسایی چندخطی در مدل رگرسیون استفاده می‌شود (کیم، ۲۰۱۹)، و VIF بیشتر از ۱۰ نشان‌دهنده وجود چندخطی بودن است (چن و همکاران، ۲۰۱۸) (معادله ۱).

$$VIF = \frac{1}{\text{Tolerance}} = \frac{1}{1-R^2} \quad (1)$$

در معادله فوق، تلورانس نماد نشان‌دهنده تحمل است و R² مقدار R-squared رگرسیون است.

³ Sum of squares within

⁴ Sum of squares total

¹ Normalized Difference Vegetation Index

² Variance Inflation Factor

۲-۶. الگوریتم شبکه عصبی کانولوشن (CNN)

(۱) شبکه عصبی مصنوعی سنتی در تجزیه و تحلیل شبکه‌های پیچیده با چالش کاهش سرعت فرایند یادگیری روبرو بود که با پیشنهاد بنزیو همکاران در سال ۱۹۹۵ مبنی بر استفاده از CNN، یک شبکه عصبی که ارتباطات محلی بین لایه‌ها را پیدا می‌کند، این مشکل برطرف شد (لو و همکاران، ۲۰۱۷). CNN در زمینه‌های مختلف تشخیص الگو به نتایج قابل توجهی دست یافته است و به‌ویژه در کاهش تعداد پارامترها در ANN مفید است (آلبوی و همکاران، ۲۰۱۷). CNN یکی از پرکاربردترین الگوریتم‌های یادگیری عمیق مناسب برای تحلیل داده‌های مکانی است (خسروی و همکاران، ۲۰۲۰). ساختار (معماری) CNN عموماً از لایه‌های کانولوشن، ادغام و کاملاً متصل تشکیل شده است (مانند لایه‌های استاندارد در ANN) (فام و همکاران، ۲۰۲۰). در ادامه ساختار هر لایه توضیح داده شده است:

(۲) لایه کانولوشن اساسی‌ترین و ضروری‌ترین لایه در معماری CNN است. این لایه عملیات کانولوشن یا ضرب را بر روی ماتریس پیکسل تولید شده برای تصویر هدف انجام می‌دهد و در نتیجه یک نقشه فعال‌سازی برای آن تصویر ایجاد می‌شود. نقشه فعال‌سازی تمام ویژگی‌های منحصربه‌فرد تصویر را ذخیره نموده و به کاهش حجم داده‌های پردازش شده کمک می‌نماید که این موضوع یکی از مزایای اصلی آن است.

(۳) لایه ادغام^۱ یک لایه حیاتی است که به کاهش ابعاد نقشه فعال‌سازی کمک نموده و درعین حال ویژگی‌های اساسی را حفظ می‌نماید و عدم تغییر

فضایی را کاهش می‌دهد. با کاهش تعداد ویژگی‌های قابل یادگیری، این لایه به موضوع بیش از حد برازش می‌پردازد. لایه ادغام همچنین CNN را قادر می‌سازد تا تمام ابعاد یک تصویر را ترکیب نموده و به آن اجازه می‌دهد تا شی موردنظر را به‌درستی شناسایی کند، حتی اگر شکل آن در موقعیت صحیح نباشد.

(۴) لایه نهایی در شبکه عصبی، لایه کاملاً متصل است که ورودی لایه‌های قبلی را دریافت می‌کند. تمام محاسبات و استدلال در این لایه از داده‌ها انجام می‌شود (رای و ریواس، ۲۰۲۰).

۲-۷. روش‌های درست‌سنجی

برای گسترش کاربرد مدل به خروجی‌های ناآشنا، لازم است عملکرد آن با مقایسه نتایج پیش‌بینی شده از هر مدل با نتایج واقعی (مشاهداتی) ارزیابی شود (ممبینی و یزدانی چمرینی، ۱۳۹۴). در این مطالعه از شاخص‌های مختلفی مانند جذر میانگین مربعات خطا ($RMSE^2$)، میانگین خطای مطلق^۳ (MAE)، R^2 و ناحیه زیر منحنی (AUC^4) - منحنی مشخصه عملکرد سیستم (ROC^5) برای ارزیابی اثربخشی ساخت مدل استفاده شده است.

۲-۷-۱. MAE و RMSE

MAE و RMSE شاخص‌هایی هستند که خطای بین مقادیر واقعی و پیش‌بینی شده را محاسبه می‌نمایند (فراهانی و همکاران، ۲۰۲۲). تفاوت اصلی بین شاخص‌های MAE و RMSE این است که MAE وزن یکسانی را به همه خطاها اختصاص می‌دهد. برعکس، RMSE با دادن وزن بیشتر به خطاهای با مقادیر مطلق بزرگتر از خطاهای با مقادیر کوچکتر، واریانس را جریمه می‌کند (چای و دراکسلر،

⁴ Area Under the Curve

⁵ Receiver Operating Characteristic

¹ Pooling

² Root Mean Square Error

³ Mean Absolute Error

در این معادله، چهار دسته داده در ماتریس سردرگمی عبارت‌اند از TN (منفی واقعی)، TP (مثبت واقعی)، FN (منفی کاذب)، و FP (مثبت کاذب). AUC بین ۰ و ۱ است. شاخص J^۳ یک آمار واحد است که عملکرد یک تست تشخیصی دوگانه را نشان می‌دهد و اغلب همراه با (ROC) استفاده می‌شود. این شاخص برای تمام نقاط یک منحنی ROC تعریف می‌شود و حداکثر مقدار شاخص به‌عنوان معیاری برای انتخاب نقطه برش بهینه استفاده شود. این شاخص به‌صورت گرافیکی معادل حداکثر فاصله عمودی بین منحنی ROC و خط مورب می‌باشد.

۳. بحث و نتایج

۳-۱. مشخصات ذرات معلق PM₁₀ و PM_{2.5}

مشخصات ذرات معلق PM₁₀ و PM_{2.5} در جدول ۱ به تفصیل ارائه شده و روند داده‌ها از سال ۲۰۱۲ تا ۲۰۲۲ در شکل ۲ نشان داده شده است. علاوه بر این، نقشه‌هایی که غلظت آلاینده را نشان می‌دهند با استفاده از درون‌یابی کریجینگ در نرم افزار ArcGIS 10.8 با اندازه پیکسل ۳۰*۳۰ متر تولید شده است (شکل ۳). مناطق پرخطر برای هر آلاینده به نقاط وقوع (با مقدار هدف ۱) تبدیل شدند، در حالی که مناطق کم خطر به عنوان نقاط غیر وقوع (با مقدار هدف ۰) برای اهداف مدل‌سازی تعیین شدند.

RMSE و MAE با توجه به معادلات ۵ و ۶ محاسبه شدند.

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (A_i - P_i)^2}{N}} \quad 5$$

$$MAE = \frac{\sum_{i=1}^n |A_i - P_i|}{N} \quad 6$$

در معادلات فوق، A_i نشان دهنده مقدار مشاهده شده، P_i نشان دهنده مقدار پیش‌بینی شده و N تعداد نمونه‌ها است.

۲-۲-۲. R²

R² نسبت واریانس در متغیر وابسته است که متغیرهای مستقل می‌توانند توضیح دهند.

R² طبق معادله ۷ محاسبه می‌شود (چیکو و همکاران، ۲۰۲۱).

$$R^2 = 1 - \frac{\sum_{i=1}^n (A_i - P_i)^2}{\sum_{i=1}^n (A_i - \bar{A}_i)^2} \quad 7$$

در این معادله، A_i مقدار مشاهده شده، P_i مقدار پیش‌بینی شده و \bar{A}_i میانگین مجموعه مشاهده شده است.

۲-۲-۳. منحنی ROC

ROC یک روش برجسته برای ارزیابی مدل‌های مکانی و یک ابزار استاندارد برای تعیین دقت نقشه‌های خروجی است (شگرخدایی و همکاران، ۲۰۲۱). منحنی ROC نرخ مثبت کاذب (FPR^۱) را روی محور x (معادله ۸) در برابر نرخ مثبت واقعی (TPR^۲) در محور y (معادله ۹) ترسیم می‌کند تا سطح زیر منحنی (AUC) را به‌عنوان مقدار واقعی اندازه‌گیری کند به صورتی که آستانه‌های کاذب تغییر می‌کند (فام و همکاران، ۲۰۲۰).

$$x = 1 - \frac{TN}{FP+TN} \quad (8)$$

$$y = \frac{TP}{FN+TP} \quad (9)$$

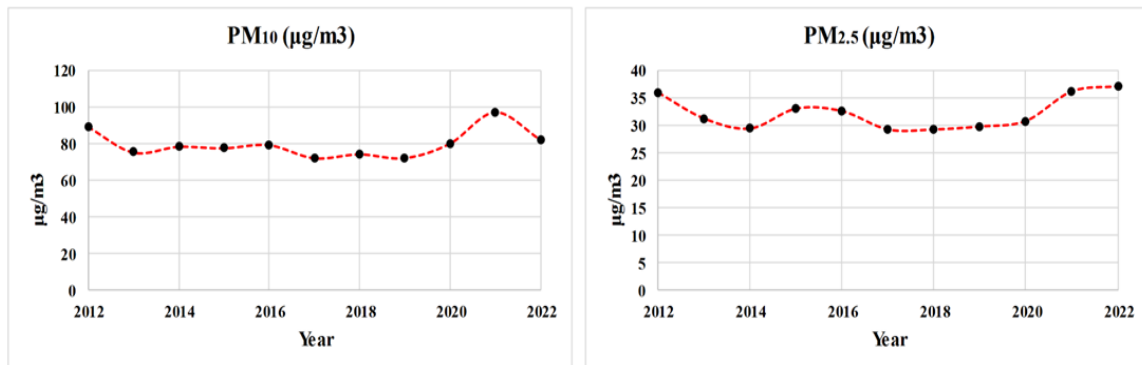
³ Youden index J

¹ False Positive Rate

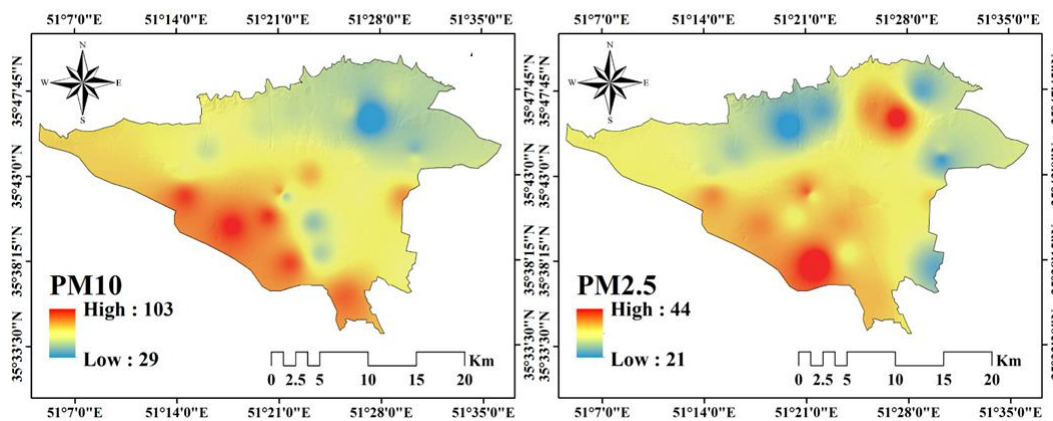
² True Positive Rate

جدول ۱- مشخصات ذرات معلق PM₁₀ و PM_{2.5}

پارامتر	کمترین مقدار سالانه (میکروگرم بر مترمکعب)	بیشترین مقدار سالانه (میکروگرم بر مترمکعب)	میانگین سالانه (میکروگرم بر مترمکعب)	انحراف معیار
PM _{2.5}	۲۱	۴۵	۳۱/۶	۶/۴
PM ₁₀	۲۹	۱۰۴	۸۴/۰۸	۱۸/۳۹



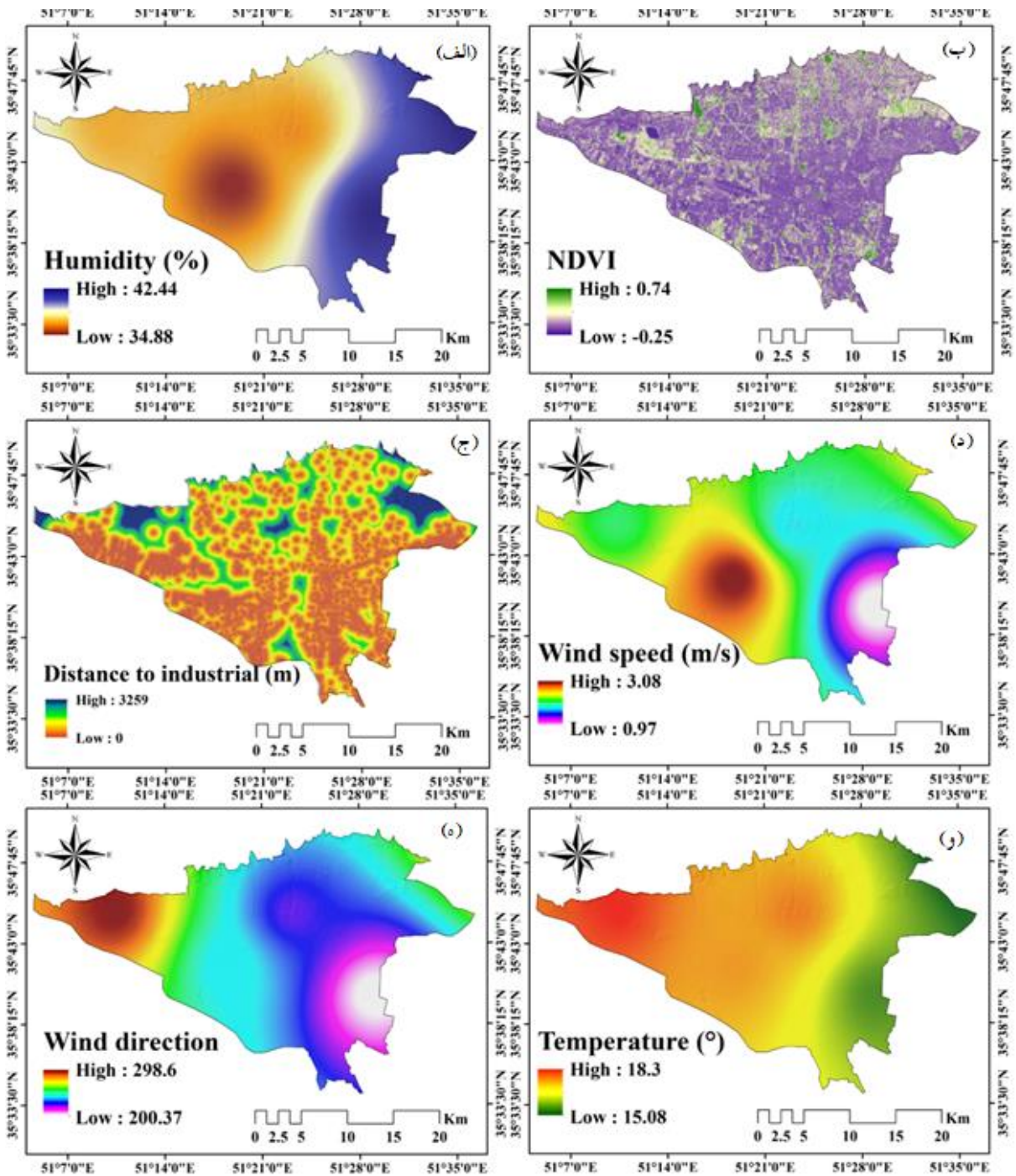
شکل ۲- روند غلظت سالانه PM₁₀ و PM_{2.5} از سال ۲۰۱۲ تا ۲۰۲۲.

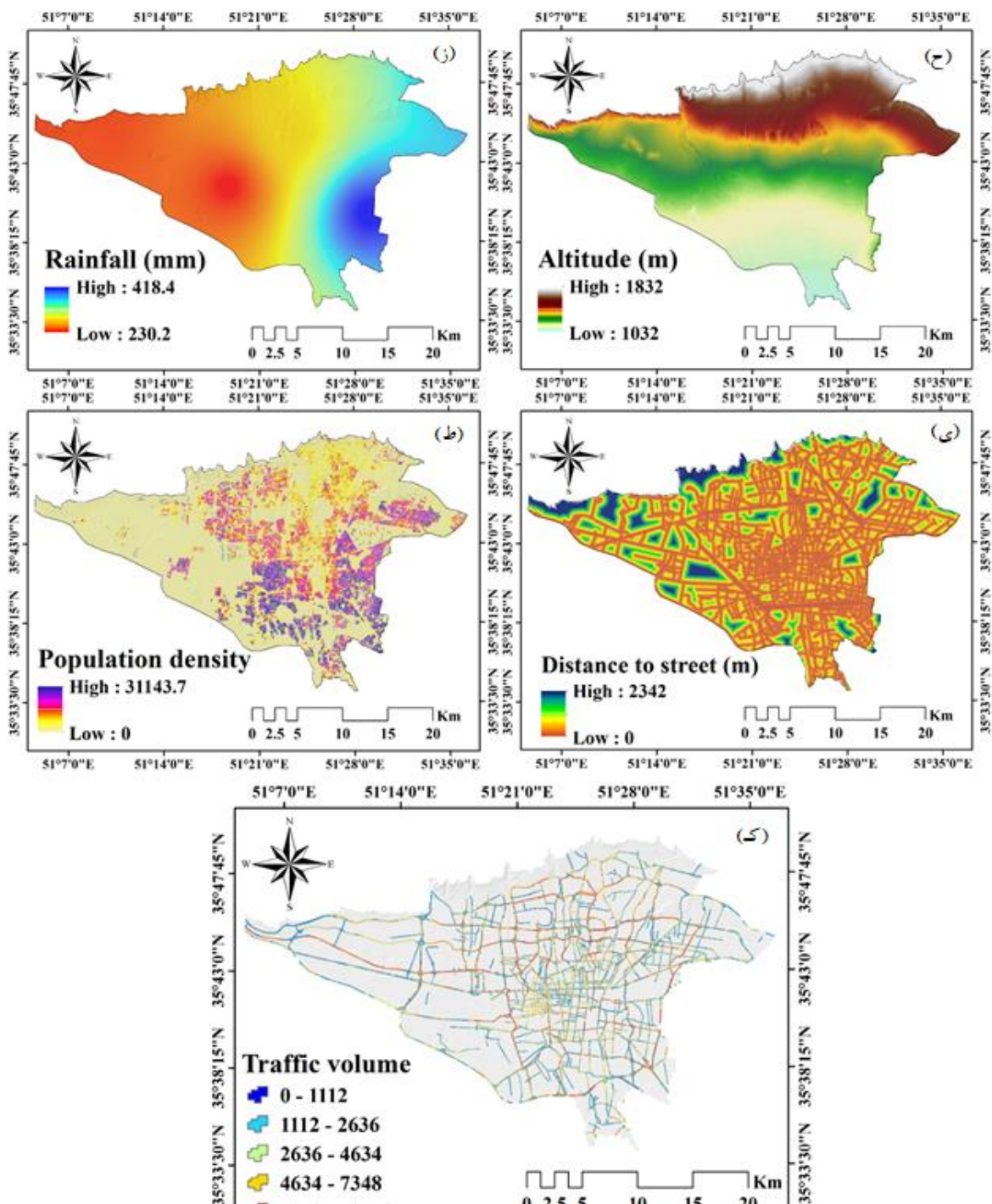


شکل ۳- نقشه میانگین غلظت سالانه ذرات معلق PM₁₀ و PM_{2.5} در تهران از ۲۰۱۲ تا ۲۰۲۲.

نرمال شده گیاهان (NDVI)، فاصله از خیابان، فاصله از مناطق صنعتی، حجم ترافیک و تراکم جمعیت است. هر یک از فاکتورهای ذکر شده با اندازه پیکسل ۳۰*۳۰ در نرم افزار ArcGIS 10.8 تهیه شده است.

۲-۳ عوامل مؤثر بر غلظت ذرات معلق PM_{2.5} و PM₁₀ شکل ۴ نشان دهنده ی عوامل مؤثر بر غلظت ذرات معلق PM₁₀ و PM_{2.5} در هوا شامل داده‌های هواشناسی (بارندگی، دما، رطوبت، جهت باد و سرعت باد)، ارتفاع، شاخص تفاوت





شکل ۴- عوامل مؤثر بر سطوح آلودگی هوا: الف) رطوبت، ب) NDVI، ج) فاصله تا مکان‌های صنعتی، د) سرعت باد، ه) جهت باد، و) دما، ز) بارندگی، ح) ارتفاع، ط) تراکم جمعیت، ی) فاصله تا خیابان، و ک) حجم ترافیک.

۳-۳. نتیجه آزمون چندخطی

یک آزمون چندخطی برای ارزیابی وجود چندخطی در میان متغیرهای مستقل مورد استفاده در مدل‌سازی مکانی و نقشه خطر ذرات معلق $PM_{2.5}$ و PM_{10} در هوا انجام شد. نتایج

آزمون ارائه شده در جدول ۲، سطوح VIF را برای هر متغیر مستقل نشان می‌دهد. از نتایج، می‌توان مشاهده کرد که هیچ یک از متغیرهای مستقل دارای مقادیر VIF بیش از حد آستانه ۱۰ نیستند. این نشان دهنده عدم وجود چند خطی

مختلف بر آلاینده‌های هوا به کار گرفته شد (شکل ۵) و تجزیه و تحلیل مربوطه یافته‌های متمایز را برای هر آلاینده نشان داد. برای PM₁₀، ارتفاع، جهت باد و سرعت باد تأثیرگذارترین پارامترها در نظر گرفته شدند. این عوامل بر پراکندگی و انتقال ذرات PM₁₀ تأثیر گذاشتند. برای PM_{2.5}، ارتفاع، بارندگی و دما به عنوان پارامترهای کلیدی شناسایی شدند. ارتفاع بر توزیع عمودی ذرات PM_{2.5} تأثیر می‌گذارد، در حالی که بارش و دما در شکل‌گیری و پراکندگی آنها بسیار مهم بودند.

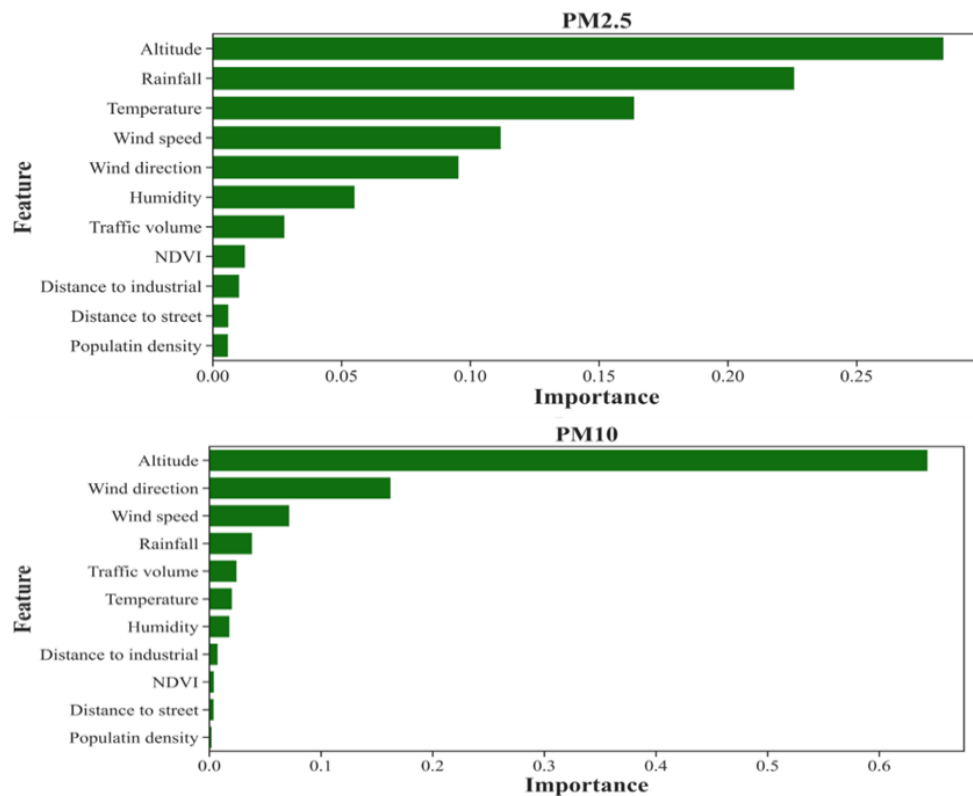
شدید در بین متغیرهای مستقل است، که نشان می‌دهد می‌توان آنها را به صورت جداگانه و جمعی در مدل سازی مکانی و نقشه خطر در نظر گرفت. تحلیل و بررسی. مقادیر VIF از ۱/۰۹ تا ۶/۸۱ نشان می‌دهد که متغیرهای مستقل دارای سطوح نسبتاً کم تا متوسطی از همبستگی با یکدیگر هستند. این نشان می‌دهد که متغیرها اطلاعات منحصر به فردی را ارائه می‌دهند و قدرت پیش‌بینی یکدیگر را بیش از حد تکرار نمی‌کنند.

۳-۴. نتایج اهمیت ویژگی عوامل

روش GeoDetector برای تعیین اهمیت پارامترهای

جدول ۲- نتایج تست چند خطی بر روی عوامل موثر بر میزان PM_{2.5} و PM₁₀ در هوا.

متغیرهای مستقل	PM _{2.5}	PM ₁₀
ارتفاع	۱/۸	۱/۷۷
رطوبت	۴/۴	۴/۳
فاصله تا صنعت	۱/۲	۱/۳
NDVI	۱/۱	۱/۱۲
تراکم جمعیت	۱/۰۹	۱/۱۶
بارش باران	۵/۳	۵/۲
فاصله تا خیابان	۱/۱۳	۱/۱۳
درجه حرارت	۵/۰۹	۳/۶
حجم ترافیک	۱/۲	۱/۰۹
جهت باد	۳/۹	۳/۴
سرعت باد	۶/۷	۶/۸



شکل ۵- نتایج اهمیت ویژگی با استفاده از GeoDetector

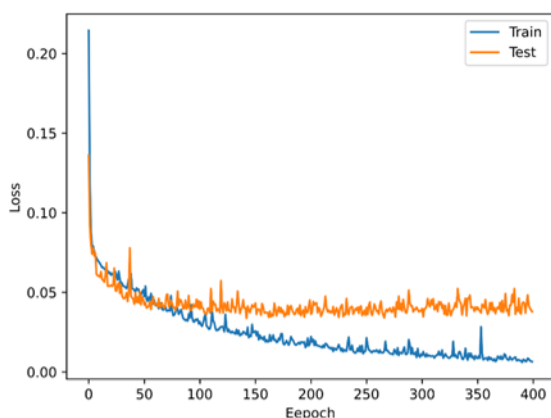
۳-۵. مدل سازی آلودگی هوا

مدل یکپارچه با استفاده از پایتون در Google Colab، یک محیط توسعه پایتون مبتنی بر ابر پیاده‌سازی شد. داده‌های ورودی بین صفر و یک نرمال‌سازی شدند تا از مقیاس‌پذیری در ویژگی‌های فضایی مختلف اطمینان حاصل شود. این مرحله نرمال‌سازی به بهبود کارایی آموزشی و همگرایی مدل کمک می‌کند. آزمایش‌ها و تجزیه و تحلیل‌ها بر روی یک رایانه رومیزی ویندوز ۱۰ با پردازنده Intel i7 و ۱۶ گیگابایت رم انجام شد. داده‌های ورودی با استفاده از تقسیم ۳۰-۷۰ تصادفی به مجموعه‌های آموزشی و آزمایشی تقسیم شدند. مجموعه آموزشی که ۷۰ درصد داده‌ها را شامل می‌شود، برای آموزش مدل و بهینه‌سازی پارامترها استفاده شد. ۳۰ درصد باقی‌مانده از داده‌ها برای آزمایش عملکرد مدل آموزش‌دیده و ارزیابی دقت پیش‌بینی آن ذخیره شده است. CNN برای ثبت الگوهای فضایی و وابستگی‌ها در

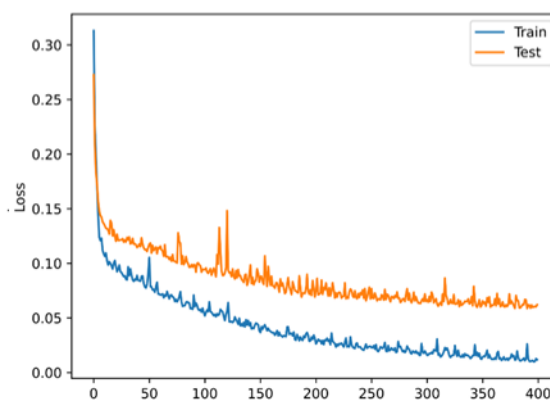
داده‌های آلاینده طراحی شده است. داده‌های ورودی شامل ویژگی‌های فضایی مختلف مانند ارتفاع، رطوبت، فاصله تا مناطق صنعتی، NDVI، تراکم جمعیت، بارندگی، فاصله تا خیابان، دما، حجم ترافیک، جهت باد و سرعت باد است. وزن‌ها، سوگیری‌ها، نرخ‌های یادگیری، تکنیک‌های منظم‌سازی و نرخ ترک تحصیل (dropout) به طور تصادفی در طول فرایند آموزش با استفاده از بهینه‌ساز Adam مقادیر اولیه و به‌روزرسانی می‌شوند. یک تابع تلفات برای اندازه‌گیری تفاوت بین غلظت آلاینده‌های پیش‌بینی شده و اندازه‌گیری‌های واقعی برای ارزیابی عملکرد مدل استفاده می‌شود. نتایج توابع تلفات برای آلاینده‌ها، همان‌طور که در شکل ۶ نشان داده شده است، نشان‌دهنده همگرایی و اثربخشی مدل ارائه شده است. مقادیر تابع تلفات برای داده‌های آموزش و آزمایش در طول آموزش کاهش می‌یابد و توانایی مدل را در یادگیری و ثبت الگوهای اساسی

تلفات روبه کاهش در داده‌های آموزشی و آزمایشی از این ایده پشتیبانی می‌کند که مدل ارائه شده با موفقیت یاد می‌گیرد و به داده‌های دیده نشده تعمیم می‌دهد، و توانایی آن در گرفتن الگوهای فضایی و وابستگی‌های آلاینده‌های هوا را برجسته می‌کند.

در داده‌های آلاینده نشان می‌دهد. روند کاهشی مقادیر تابع تلفات نشان می‌دهد که این مدل با موفقیت اختلاف بین غلظت آلاینده پیش‌بینی شده و اندازه‌گیری‌های واقعی را در طول آموزش به حداقل می‌رساند. این نشان می‌دهد که مدل در حال یادگیری پیش‌بینی‌های دقیق است و به طور مؤثر روابط پیچیده درون داده‌ها را ضبط می‌نماید. مقادیر تابع



A



B

شکل ۶- عملکرد تلفات مدل CNN (A) PM₁₀ و (B) PM_{2.5}.

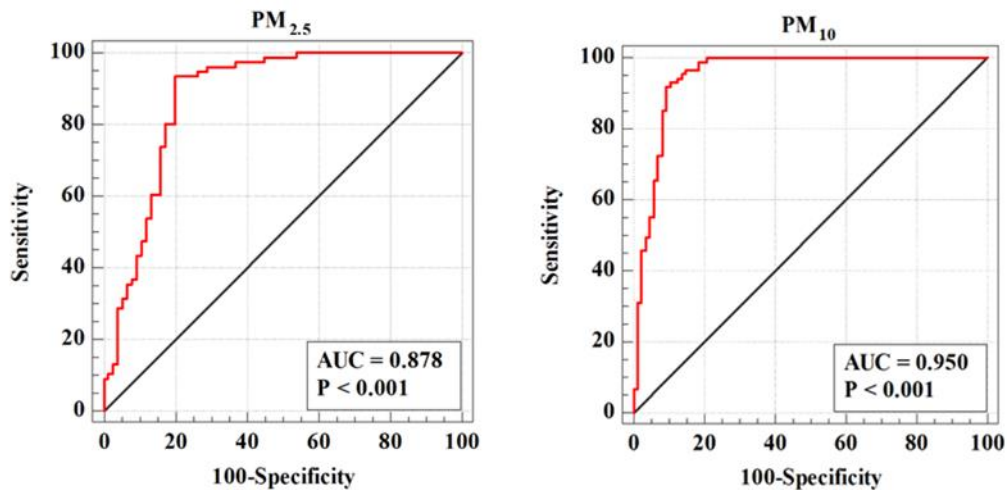
مدل CNN به R^2 چشمگیر ۰/۹۷۲ دست یافت که نشان می‌دهد این مدل می‌تواند تقریباً ۹۷.۲٪ از واریانس غلظت PM₁₀ را توضیح دهد. علاوه بر این، مدل CNN مقادیر RMSE و MAE پایین را به ترتیب ۰/۰۸۲ و ۰/۰۵۳ نشان داد.

علاوه بر این، معیارهایی مانند MAE، RMSE و R^2 برای ارزیابی دقت و قدرت پیش‌بینی مدل محاسبه می‌شوند (جدول ۳ و شکل ۷). برای آلاینده PM_{2.5}، مدل CNN عملکرد نسبتاً خوبی از خود نشان داد و به R^2 برابر با ۰/۸۸۹ دست یافت. مقادیر RMSE و MAE مربوطه به ترتیب ۰/۱۶۶ و ۰/۰۴۶ بود.

همچنین برای آلاینده PM₁₀.

جدول ۳- نتیجه مدل سازی آلودگی هوا در مرحله آموزش و آزمایش.

آلاینده	مرحله آموزش			مرحله آزمایش		
	R^2	RMSE	MAE	R^2	RMSE	MAE
PM _{2.5}	۰/۸۸۹	۰/۱۶۶	۰/۰۴۶	۰/۹۶۹	۰/۰۸۲	۰/۱۵۷
PM ₁₀	۰/۹۷۲	۰/۰۸۲	۰/۰۵۳	۰/۹۸	۰/۰۷۱	۰/۰۴۵

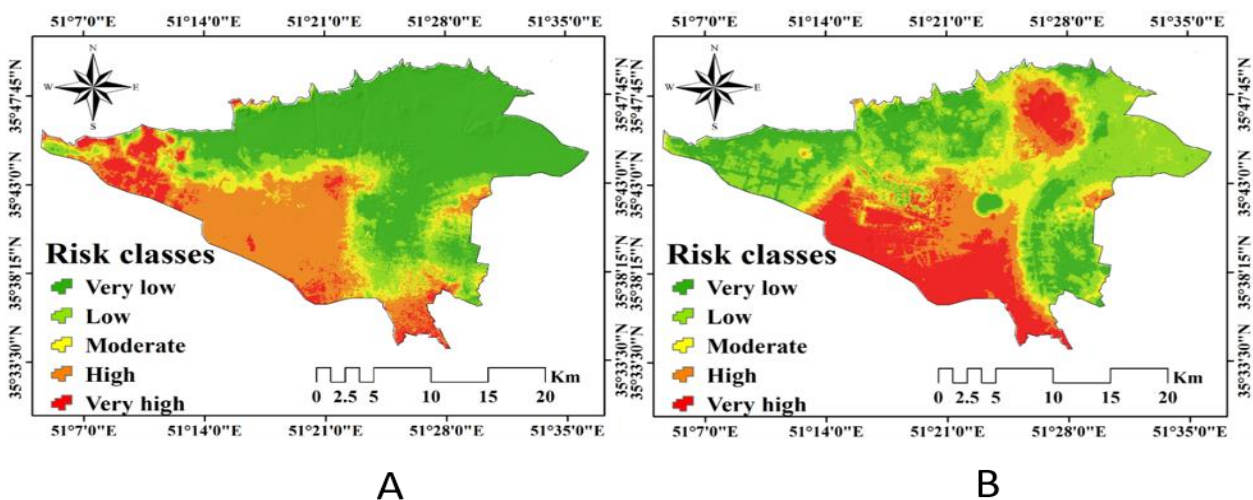


شکل ۷- درست سنجی نقشه‌های ریسک توسط منحنی ROC.

مطالعه ایجاد گردید. هر منطقه با توجه به سطح خطر اختصاص داده شده، رنگی شد، که تصویری بصری از نقاط داغ آلودگی و مناطق مورد توجه ارائه می‌دهد. با توجه به نقشه‌های خطر تولید شده از مدل CNN، آلودگی PM_{10} خطر بالاتری را در مناطق غرب و جنوب غربی ایجاد می‌نماید. در حالی که آلودگی $PM_{2.5}$ در قسمت جنوبی بیشتر می‌باشد. لازم به ذکر است نقشه داده‌های میانگین ۱۰ ساله واقعی در شکل ۳ ارائه شده است.

۳-۶. ایجاد نقشه ریسک و درست‌سنجی

با استفاده از مدل CNN آموزش دیده، غلظت آلاینده برای هر مکان در منطقه مورد مطالعه تخمین زده شد. سپس این غلظت‌های تخمین زده شده بر اساس معیارهای طبقه‌بندی، سطوح خطر را به مناطق مختلف اختصاص دادند. سطوح خطر را می‌توان به‌عنوان بسیار کم، کم، متوسط، زیاد و بسیار زیاد طبقه‌بندی نمود که نشان‌دهنده درجات مختلفی از شدت آلودگی است (شکل ۸). نقشه‌های خطر با همپوشانی غلظت آلاینده‌های برآورد شده بر روی نقشه جغرافیایی منطقه مورد



شکل ۸- نقشه خطرذرات معلق: (A) PM_{10} و (B) $PM_{2.5}$. خروجی مدل CNN میانگین سالانه ۲۰۱۲ تا ۲۰۲۲

شاخص J برابر با ۰/۷۳۶۸ است که نشان دهنده توانایی مدل در شناسایی آستانه‌های ریسک مناسب است. خطاهای استاندارد برابر ۰/۰۲۹۸ می‌باشد که دقت نقشه‌های ریسک را نشان می‌دهد. فواصل اطمینان ۹۵٪ از ۰/۸۱۵ تا ۰/۹۲۶ برای PM_{2.5} متغیر است که باعث تقویت بیشتر قابلیت اطمینان تخمین خطر می‌شود. مقادیر آماره Z به ترتیب ۱۲/۶۹ و سطح معنی‌داری برای آلاینده $P < ۰/۰۰۰۱$ است که بر اهمیت آماری نتایج مشاهده شده تأکید می‌نماید.

بنابراین الگوریتم CNN در مدل‌سازی مکانی و نقشه خطر دو آلاینده هوا مؤثر می‌باشد. مقادیر AUC بالا، مقادیر شاخص J قابل توجه، فواصل اطمینان باریک و مقادیر P پایین نشان‌دهنده توانایی مدل در تمایز بین سطوح مختلف خطر آلاینده و قابلیت اطمینان آماری آن است. این نتایج به درک ما از توزیع مکانی و پتانسیل خطر کمک می‌نماید.

چندین معیار ارزیابی برای ارزیابی اثربخشی نقشه‌های ریسک تولید شده با روش CNN، از جمله منحنی ROC، AUC، و شاخص J استفاده شد. این معیارها برای تجزیه و تحلیل عملکرد نقشه‌های ریسک از نظر توانایی آنها استفاده شد. برای تمایز دقیق بین سطوح مختلف خطر. نتایج ارزیابی، همان‌طور که در جدول ۴ و شکل ۸ ارائه شده است برای PM₁₀، AUC برابر با ۰.۹۵ به دست آمد که نشان دهنده قدرت تبعیض خوب است. شاخص J برابر با ۰/۸۲۷۶ است که اثربخشی مدل را در شناسایی آستانه‌های ریسک برجسته می‌کند. خطای استاندارد ۰.۰۱۷۵ است و فاصله اطمینان ۹۵٪ از ۰/۹۰۷ تا ۰/۹۷۷ متغیر می‌باشد که نشان دهنده سطح اطمینان بالایی در نقشه خطر است. مقدار آماره Z برابر با ۲۵/۷۶ و سطح معنی‌داری $P < ۰/۰۰۰۱$ می‌باشد که اهمیت آماری یافته‌ها را تایید می‌کند. مقدار AUC برای PM_{2.5} برابر با ۰/۸۷۸ می‌باشد که قدرت تمایز متوسط تا بالا را نشان می‌دهد.

جدول ۵- نتیجه درست سنجی نقشه خطر آلاینده‌های هوا.

Parameter	AUC	شاخص یودن J	Standard Error	فاصله اطمینان ۹۵٪	آماره Z	سطح معنی‌داری
PM ₁₀	۰/۹۵	۰/۸۲۷۶	۰/۰۱۷۵	۰/۹۰۷ تا ۰/۹۷۷	۲۵/۷۶	$P < ۰/۰۰۰۱$
PM _{2.5}	۰/۸۷۸	۰/۷۳۶۸	۰/۰۲۹۸	۰/۸۱۵ تا ۰/۹۲۶	۱۲/۶۹	$P < ۰/۰۰۰۱$

است، CNN از روابط و الگوهای فضایی در داده‌ها استفاده کرد و مدل‌سازی دقیقی از غلظت آلاینده‌ها در سراسر منطقه مورد مطالعه را امکان‌پذیر کرد (جیانگ و همکاران، ۲۰۲۲). روش GeoDetector اهمیت پارامترهای مختلف را در آلاینده‌های مختلف هوا ارزیابی کرد و بینش‌های مهمی را برای سیاست‌گذاران و محققان آشکار کرد. برای PM₁₀، ارتفاع، جهت باد و سرعت باد تأثیرات قابل توجهی دارند. ارتفاع بر پراکندگی و انتقال ذرات PM₁₀ تأثیر می‌گذارد، با ارتفاعات بالاتر اغلب منجر به افزایش اختلاط جوی و رقیق شدن آلاینده‌ها می‌شود (لی و همکاران، ۲۰۱۹). جهت و سرعت باد مسیرها و مسافت‌هایی را که ذرات PM₁₀

۳-۷. بحث

نتایج این مطالعه نشان می‌دهد که الگوریتم‌های CNN یک رویکرد موفق برای مدل‌سازی فضایی و نقشه‌برداری ریسک دو آلاینده مهم هوای شهر تهران یعنی PM_{2.5} و PM₁₀ است. با استفاده از نقاط قوت این تکنیک، به پیش‌بینی‌های دقیقی از غلظت آلاینده‌های هوا دست یافتیم. این بخش یافته‌های کلیدی، مفاهیم، محدودیت‌ها و جهت‌های بالقوه آینده تحقیق را مورد بحث قرار می‌دهد. یکی از یافته‌های اصلی این مطالعه، دقت قابل توجه در مدل‌سازی است که از طریق رویکرد CNN به دست آمده است (دایری و همکاران، ۲۰۲۱). همان‌طور که با مقادیر کم MAE و RMSE مشهود

اطمینان و تعمیم‌پذیری نقشه‌های ریسک تولید شده ارائه دهد. از نظر جهت‌گیری‌های آتی، این تحقیق راه‌هایی را برای کاوش تکنیک‌ها و روش‌های اضافی برای بهبود مدل‌سازی کیفیت هوا و نقشه‌برداری ریسک باز می‌کند. به عنوان مثال، ترکیب رویکردهای مدل‌سازی مکانی-زمانی می‌تواند ماهیت پویای آلودگی هوا را به تصویر بکشد و دقت پیش‌بینی‌ها را بهبود بخشد. علاوه بر این، ادغام سایر الگوریتم‌های یادگیری ماشین یا مدل‌های ترکیبی می‌تواند نتایج بهتری را با استفاده از نقاط قوت تکنیک‌های مختلف به همراه داشته باشد.

۴. نتیجه‌گیری

این تحقیق یک رویکرد دقیقی برای مدل‌سازی فضایی و نقشه‌برداری ریسک دو آلاینده هوا با استفاده از الگوریتم CNN ارائه می‌کند. برای آلاینده $PM_{2.5}$ ، مدل CNN عملکرد نسبتاً خوبی از خود نشان داد و به R^2 برابر با 0.889 دست یافت. مقادیر $RMSE$ و MAE مربوطه به ترتیب 0.166 و 0.046 بود. همچنین برای آلاینده PM_{10} ، مدل CNN به R^2 چشمگیر 0.972 دست یافت که نشان می‌دهد این مدل می‌تواند تقریباً 97.2% از واریانس غلظت PM_{10} را توضیح دهد. علاوه بر این، مدل CNN مقادیر $RMSE$ و MAE پایین را به ترتیب 0.082 و 0.053 نشان داد. با توجه به مقادیر R^2 ، $RMSE$ و MAE نتیجه‌گیری می‌شود که مدل CNN از نظر دقت پیش‌بینی کارایی قابل قبولی ارائه می‌کند. در منطقه مورد مطالعه، آلاینده‌ها بیشتر تحت تأثیر پارامترهای خاص، یعنی ارتفاع، جهت باد، سرعت باد، بارندگی و دما قرار گرفتند که با استفاده از روش GeoDetector تعیین شد. نقشه‌های خطر تولید شده توسط مدل CNN الگوهای آلودگی متمایز را در مناطق مختلف نشان می‌دهد. آلودگی PM_{10} خطر بیشتری را در مناطق غربی و جنوب غربی ایجاد می‌کند، در حالی که آلودگی $PM_{2.5}$ در

می‌تواند طی کنند تعیین می‌کند و بر توزیع فضایی و غلظت آنها تأثیر می‌گذارد (وانگ و همکاران، ۲۰۱۰). برای $PM_{2.5}$ ، ارتفاع، بارندگی و دما اثرات قابل توجهی از خود نشان می‌دهند. ارتفاع بر توزیع عمودی ذرات $PM_{2.5}$ تأثیر می‌گذارد، با انتشار و منابع در ارتفاعات مختلف بر غلظت آنها در سطح زمین تأثیر می‌گذارد (پنگ و همکاران، ۲۰۱۵). بارندگی می‌تواند ذرات $PM_{2.5}$ را از جو حذف کند و سطح آلاینده‌ها را کاهش دهد (نواک و همکاران، ۲۰۱۳). دما می‌تواند بر واکنش‌های شیمیایی و فرآیندهای فیزیکی درگیر در تشکیل، تبدیل و پراکندگی ذرات $PM_{2.5}$ تأثیر بگذارد (سو و همکاران، ۲۰۲۰). علی‌رغم بینش‌های ارزشمند به دست آمده از این تحقیق در مورد مدیریت کیفیت هوا با استفاده از مدل‌سازی فضایی، نقشه‌برداری ریسک CNN، ادعان به محدودیت‌های خاص و ارائه پیشنهادهایی برای تحقیقات آتی مهم است. اولاً، دقت مدل‌ها به شدت به کیفیت و نمایندگی داده‌های ورودی بستگی دارد. هر گونه نادرستی یا سوگیری در داده‌های نظارت می‌تواند بر قابلیت اطمینان مدل‌ها و نقشه‌های ریسک تأثیر بگذارد. علاوه بر این، معیارهای فضایی مورد استفاده در تجزیه و تحلیل مبتنی بر دانش و فرضیات موجود در مورد عوامل مؤثر بر آلودگی هوا است. ممکن است پارامترهای فضایی نامشخص دیگری وجود داشته باشد که می‌تواند بر دقت مدل‌ها تأثیر بگذارد. مطالعات آینده می‌تواند ترکیب مجموعه داده‌های جامع‌تر و تکنیک‌های انتخاب ویژگی پیشرفته‌تر را برای افزایش دقت مدل‌سازی مورد بررسی قرار دهد. علاوه بر این، معیارهای ارزیابی مورد استفاده در این مطالعه، مانند MAE و $RMSE$ ، یک ارزیابی کلی از عملکرد مدل‌سازی ارائه می‌دهند. با این حال، در نظر گرفتن اقدامات ارزیابی اضافی، مانند تکنیک‌های درست‌سنجی فضایی، برای ارزیابی خوب بودن تناسب و توانایی مدل برای گرفتن دقیق الگوهای فضایی ضروری است. این می‌تواند بینش بیشتری در مورد قابلیت

- (MAE)?-Arguments against avoiding RMSE in the literature. Geoscientific model development, 7, pp. 1247-1250.
7. Chauhan, R., Kaur, H., and Alankar, B., 2021. Air quality forecast using convolutional neural network for sustainable development in urban environments. Sustainable Cities and Society, 75, pp.103239.
 8. Li, Z., Liu, F., Yang, W., Peng, S., Zhou, J., 2021. A survey of convolutional neural networks: analysis, applications, and prospects. IEEE transactions on neural networks and learning systems, 33, pp.6999-7019.
 9. Chen, W., Li, H., Hou, E., Wang, S., Wang, G., Panahi, M., Li, T., Peng, T., Guo, C., Niu, C., 2018. GIS-based groundwater potential analysis using novel ensemble weights-of-evidence with logistic regression and functional tree models. Science of The Total Environment, 634, pp.853-867.
 10. Chicco, D., Warrens, M.J., Jurman, G., 2021. The coefficient of determination R-squared is more informative than SMAPE, MAE, MAPE, MSE and RMSE in regression analysis evaluation. PeerJ Computer Science, 7, pp.e623.
 11. Dairi, A., Harrou, F., Khadraoui, S., Sun, Y., 2021. Integrated multiple directed attention-based deep learning for improved air pollution forecasting. IEEE Transactions on Instrumentation and Measurement, 70, pp.1-15.
 12. Farahani, M., Razavi-Termeh, S.V., Sadeghi-Niaraki, A., 2022. A spatially based machine learning algorithm for potential mapping of the hearing senses in an urban environment. Sustainable Cities and Society, 80, pp.103675.
 13. Garg, A., Tai, K., 2013. Comparison of statistical and machine learning methods in modelling of data with multicollinearity. International Journal of Modelling, Identification and Control, 18, pp. 295-312.
 14. Gulliver, J., Briggs, D., 2011. STEMS-Air: A simple GIS-based air pollution dispersion model for city-wide exposure assessment. Science of The Total Environment, 409, pp. 2419-2429.
- قسمت جنوبی برجسته تر است. نقشه‌های ریسک تولید شده از طریق روش شناسی یکپارچه، بینش‌های ارزشمندی را برای مدیریت کیفیت هوا ارائه می‌دهد. با تجسم توزیع فضایی غلظت آلاینده‌ها، این نقشه‌های خطر به شناسایی مناطق پرخطر و نقاط داغ آلودگی کمک می‌کند. این اطلاعات برای سیاست گذاران، آژانس‌های زیست محیطی و ذینفعان برای اولویت بندی تلاش‌های کاهش و تخصیص موثر منابع بسیار مهم است. نقشه‌های ریسک همچنین می‌توانند از فرآیندهای تصمیم‌گیری پشتیبانی کرده و توسعه مداخلات هدفمند را برای کاهش سطوح آلاینده و محافظت از سلامت عمومی تسهیل کنند.
- ### ۵. منابع
۱. بهاری، ر.، عباس‌پور، ر. و پ. پهلوانی، (۱۳۹۴)، پهنه بندی آلودگی ذرات معلق با استفاده از مدل‌های آماری محلی در GIS (مطالعه موردی، شهر تهران). علوم و فنون نقشه برداری، دوره ۵، صفحات ۱۶۵ تا ۱۷۳.
 ۲. کشاورزی شیرازی، ه. و ا.م. یدقار، (۱۳۸۵)، مدل‌سازی و تحلیل پراکنش آلاینده‌های منتشره از منابع متحرک آلودگی هوا بر پایه‌ی GIS مطالعه موردی: شهر تهران، اولین همایش تخصصی مهندسی محیط زیست، تهران، <https://civilica.com/doc/11826>
 3. Albawi, S., Mohammed, T.A., Al-Zawi, S., 2017. Understanding of a convolutional neural network, international conference on engineering and technology (ICET). Ieee , pp.1-6.
 4. Bekkar, A., Hssina, B., Douzi, S., Douzi, K., 2021. Air-pollution prediction in smart city, deep learning approach. Journal of big Data, 8, pp.1-21.
 5. Bengio, Y., Lecun, Y., 1995. Convolutional networks for images, speech, and time-Series. The Handbook of Brain Theory and Neural Networks, Cambridge: MIT Press. 3361:pp.310.
 6. Chai, T., Draxler, R.R., 2014. Root mean square error (RMSE) or mean absolute error

- prediction models on a national and regional scale using open-source remote sensing data. *Environmental Monitoring and Assessment*, 195, pp. 644.
25. Mombeini, H., Yazdani-Chamzini, A., 2015. Modeling gold price via artificial neural network. *Journal of Economics, Business and Management*, 3, pp. 699-703.
 26. Nowak, D.J., Hirabayashi, S., Bodine, A., Hoehn, R., 2013. Modeled PM_{2.5} removal by trees in ten US cities and associated health effects. *Environmental Pollution*, 178, pp.395-402.
 27. Peng, Z.-R., Wang, D., Wang, Z., Gao, Y., Lu, S., 2015. A study of vertical distribution patterns of PM_{2.5} concentrations based on ambient monitoring with unmanned aerial vehicles: A case in Hangzhou, China. *Atmospheric Environment*, 123, pp. 357-369.
 28. Perera, F., 2017. Pollution from Fossil-Fuel Combustion is the Leading Environmental Threat to Global Pediatric Health and Equity: Solutions Exist. *Int J Environ Res Public Health*, 15, pp.16.
 29. Pham, V.D., Nguyen, Q.-H., Nguyen, H.-D., Pham, V.-M., Bui, Q.-T., 2020. Convolutional neural network—optimized moth flame algorithm for shallow landslide susceptible analysis. *IEEE Access*, 8, pp. 32727-32736.
 30. Rai, M., Rivas, P., 2020. A review of convolutional neural networks and Gabor filters in object recognition, 2020 International Conference on Computational Science and Computational Intelligence (CSCI). *IEEE*, pp.1560-1567.
 31. Razavi-Termeh, S.V., Sadeghi-Niaraki, A., Choi, S.-M., 2021. Spatial modeling of asthma-prone areas using remote sensing and ensemble machine learning algorithms. *Remote Sensing*, 13, pp.3222.
 32. Ren, X., Mi, Z., Georgopoulos, P.G., 2020. Comparison of Machine Learning and Land Use Regression for fine scale spatiotemporal estimation of ambient air pollution: Modeling ozone concentrations across the contiguous United States. *Environment International*, 142, pp.105827.
 33. Shogrkhodaei, S.Z., Razavi-Termeh, S.V., Fathnia, A., 2021. Spatio-temporal
 15. Hu, X., Chu, L., Pei, J., Bian, J., Liu, W., 2021. Deep Learning Model Complexity: Concepts and Approaches.
 16. Jia, W.-J., Wang, M.-F., Zhou, C.-H., Yang, Q.-H., 2021. Analysis of the spatial association of geographical detector-based landslides and environmental factors in the southeastern Tibetan Plateau, China. *PLoS One*, 16, p. e0251776.
 17. Jiang, Z., Zheng, T., Bergin, M., Carlson, D., 2022. Improving spatial variation of ground-level PM_{2.5} prediction with contrastive learning from satellite imagery. *Science of Remote Sensing*, 5, pp.100052.
 18. Khosravi, K., Panahi, M., Golkarian, A., Keesstra, S.D., Saco, P.M., Bui, D.T., Lee, S., 2020. Convolutional neural network approach for spatial prediction of flood hazard at national scale of Iran. *Journal of Hydrology*, 591, pp.125552.
 19. Kim, J.H., 2019. Multicollinearity and misleading statistical results. *Korean journal of anesthesiology*, 72, pp.558-569.
 20. Li, X., Ma, Y., Wang, Y., Wei, W., Zhang, Y., Liu, N., Hong, Y., 2019. Vertical distribution of particulate matter and its relationship with planetary boundary layer structure in Shenyang, Northeast China. *Aerosol and Air Quality Research*, 19, pp.2464-2476.
 21. Li, Z., Liu, F., Yang, W., Peng, S., & Zhou, J. 2021. A survey of convolutional neural networks: analysis, applications, and prospects. *IEEE transactions on neural networks and learning systems*, 33, p. 6999-7019.
 22. Lü, G., Batty, M., Strobl, J., Lin, H., Zhu, A.-X., Chen, M., 2019. Reflections and speculations on the progress in Geographic Information Systems (GIS): a geographic perspective. *International journal of geographical information science*. 33, pp.346-367.
 23. Lu, H., Fu, X., Liu, C., Li, L.-g., He, Y.-x., Li, N.-w., 2017. Cultivated land information extraction in UAV imagery based on deep convolutional neural network and transfer learning. *Journal of Mountain Science*, 14, pp.731-741.
 24. Mamić, L., Gašparović, M., Kaplan, G., 2023. Developing PM_{2.5} and PM₁₀

- A., 2016. Stream-flow forecasting using extreme learning machines: a case study in a semi-arid region in Iraq. *Journal of Hydrology*, 542, pp.603-614.
39. Yousefian, F., Faridi, S., Azimi, F., Aghaei, M., Shamsipour, M., Yaghmaeian, K., Hassanvand, M.S., 2020. Temporal variations of ambient air pollutants and meteorological influences on their concentrations in Tehran during 2012–2017. *Scientific reports*, 10, pp.292.
40. Zehtab Yazdi, Y., Mansouri, N., Atabi, F., Aghamohammadi, H., 2021. Dispersion Modeling of Particulate Matters (PM_{2.5}, PM₁₀) from Asphalt Plants in the Southwest of Tehran. *Journal of Environmental Health Engineering*, 8, pp.375-390.
41. Zhang, S., Zhou, Y., Yu, Y., Li, F., Zhang, R., Li, W., 2022. Using the Geodetector Method to Characterize the Spatiotemporal Dynamics of Vegetation and Its Interaction with Environmental Factors in the Qinba Mountains, China. *Remote Sensing*, 14, pp.579.
- modeling of PM_{2.5} risk mapping using three machine learning algorithms. *Environmental Pollution*, 289, pp.117859.
34. Steininger, M., Kobs, K., Zehe, A., Lautenschlager, F., Becker, M., Hotho, A., 2020. Maplur: Exploring a new paradigm for estimating air pollution using deep learning on map images. *ACM Transactions on Spatial Algorithms and Systems (TSAS)*, 6, pp.1-24.
35. Su, H., Cheng, Y., Pöschl, U., 2020. New multiphase chemical processes influencing atmospheric aerosols, air quality, and climate in the anthropocene. *Accounts of chemical research*, 53, pp.2034-2043.
36. Wang, F., Chen, D., Cheng, S., Li, J., Li, M., Ren, Z., 2010. Identification of regional atmospheric PM₁₀ transport pathways using HYSPLIT, MM5-CMAQ and synoptic pressure pattern analysis. *Environmental Modelling & Software*, 25, pp.927-934.
37. Wang, J., Xu, C., 2017. Geodetector: Principle and prospective. *Acta Geogr. Sin.*, 72, 116-134.
38. Yaseen, Z.M., Jaafar, O., Deo, R.C., Kisi, O., Adamowski, J., Quilty, J. and El-Shafie,