



Spatio – Temporal Estimation of Carbon Monoxide and Nitrogen Dioxide based on Remote Sensing Data and Ancillary Data in Tehran

Ali Shamsoddini^{1*}, Wanko Ahmadi¹

¹ Department of Remote Sensing and Geographic Information System, Faculty of Humanities, Tarbiat Modares University, Tehran, Iran

ARTICLE INFO

Article Type: Research article

Article history:

Received 28 October 2020

Accepted 13 December 2020

Available online 19 December 2020

Keywords:

Machine Learning, Artificial Neural Network, Spatio-Temporal Modeling, Carbon Monoxide, Nitrogen Dioxide.

Citation: Shamsoddini, A., Ahmadi, W. (2020). Spatio – Temporal Estimation of Carbon Monoxide and Nitrogen Dioxide based on Remote Sensing Data and Ancillary Data in Tehran. *Geography and Sustainability of Environment*, 10 (3), 107-124.

doi: [10.22126/GES.2020.4227.2057](https://doi.org/10.22126/GES.2020.4227.2057)

ABSTRACT

Air pollution is one of the most important consequences of human activities, which not only threatens human health but also negatively affects all elements of the environment, including plants and animals. Tehran, the capital of Iran, and the administrative, political and economic center of the country, is no exception which is constantly struggling with these hazard. So far, many linear and nonlinear models have been applied to model air pollution. In this research, 8 pollutant measurement stations distributed over Tehran were selected according to the availability of their recorded data. In order to provide a model predicting pollutants, spatially and temporally, the combination of spatial and temporal features extracted of remote sensing data and environmental data was modeled using multilayer perceptron artificial neural network. The input data include meteorological data, topography, traffic index, population data, air pollutant concentrations for the last days, and land use map. In addition, vegetation cover, distance from heat islands, and the land surface temperature derived from remotely sensed data were used as remotely sensed attributes. In order to increase the accuracy of modeling, wavelet transform and feature selection methods were used on input attributes of the model. Random forest feature selection method was applied on the input data in order to reduce the number of input attributes,. The results of the model evaluation indicated that the model was efficient in estimating the concentrations of pollutants. Temporally, carbon monoxide and nitrogen dioxide were predicted with error estimation of 13% and 11.5%, respectively. Besides, these pollutants were spatially predicted with the estimation error less than 17%.

*. Corresponding author E-mail address:

ali.shamsoddini@modares.ac.ir



تخمین مکانی - زمانی آلاینده‌های منواکسید کربن و دی‌اکسید نیتروژن شهر تهران مبتنی بر داده‌های حاصل از سنجش از دور و داده‌های کمکی

علی شمس‌الدینی^{۱*}، وانکو احمدی^۱

^۱ گروه سنجش از دور و سیستم اطلاعات جغرافیایی، دانشکده علوم انسانی، دانشگاه تربیت مدرس، تهران، ایران

چکیده	مشخصات مقاله
<p>آلودگی هوا یکی از پیامدهای ناهنجار فعالیت‌های بشر است که نه تنها سلامت انسان را تهدید می‌کند؛ بلکه بر همه عوامل محیط‌زیست از جمله گیاهان و جانوران تأثیر نامطلوب می‌گذارد. تهران به عنوان مرکز اداری، سیاسی و اقتصادی کشور و پرجمعیت‌ترین شهر ایران، یکی از آلوده‌ترین شهرهای دنیا به‌شمار می‌رود. از مدل‌های خطی و غیر خطی متعددی تاکنون به منظور مدل‌سازی آلودگی هوا استفاده شده است. در نوشتار پیش رو از ویژگی‌های مکانی و زمانی مستخرج از تصاویر سنجش از دور و داده‌های محیطی ایستگاه‌های پایش آلودگی هوا سازمان محیط‌زیست واقع در سطح شهر تهران پس از پیش‌پردازش‌های لازم به‌مثابه ورودی مدل استفاده شد. از میان ایستگاه‌های آلاینده‌سنج موجود در سطح شهر تهران، با در نظر گرفتن پوشش سری زمانی مشترک داده‌های ثبت‌شده، تعداد هشت ایستگاه انتخاب شد. به‌منظور انجام فرایند مدل‌سازی از روش شبکه عصبی مصنوعی پرسپترون چندلایه با الگوریتم آموزش لوبزبرگ - مارکوارت و تابع فعال‌سازی سیگموئیدی استفاده شد. در پژوهش حاضر از داده‌های هواشناسی، داده‌های مربوط به غلظت آلاینده‌ها در روزهای قبل، کاربری اراضی و نیز داده‌های مستخرج از تصاویر ماهواره‌ای شامل داده‌های مربوط به پوشش گیاهی و جزایر حرارتی به‌منظور مدل‌سازی غلظت آلاینده‌ها استفاده شد. از روش تبدیل موجک بر روی مقادیر غلظت آلاینده‌ها در روزهای قبل استفاده گردید و سپس روش انتخاب ویژگی جنگل تصادفی بر روی ویژگی‌های ورودی مدل اعمال شد؛ همچنین با توجه به تغییرات مکانی آلودگی هوا سعی بر آن شد که با استفاده از اطلاعات هفت ایستگاه، مقادیر غلظت آلاینده یک ایستگاه برآورد شود. نتایج حاصل از ارزیابی مدل بیانگر کارایی مدل ارائه‌شده در تخمین مقادیر بیشینه روزانه غلظت آلاینده بود. منواکسید کربن و دی‌اکسید نیتروژن به ترتیب با خطای ۱۳٪ و ۱۱/۵٪ به‌صورت زمانی پیش‌بینی شدند؛ همچنین این دو آلاینده به‌صورت مکانی با خطای تخمین کمتر از ۱۷٪ پیش‌بینی شدند.</p>	<p>نوع مقاله: پژوهشی</p> <p>تاریخچه مقاله: دریافت ۷ آبان ۱۳۹۹ پذیرش ۲۳ آذر ۱۳۹۹ دسترسی آنلاین ۲۹ آذر ۱۳۹۹</p> <p>کلیدواژه‌ها: یادگیری ماشین، شبکه عصبی مصنوعی، مدل‌سازی مکانی - زمانی، منواکسید کربن، دی‌اکسید نیتروژن.</p> <p>استناد: شمس‌الدینی، علی؛ احمدی، وانکو (۱۳۹۹). تخمین مکانی - زمانی آلاینده‌های منواکسید کربن و دی‌اکسید نیتروژن شهر تهران مبتنی بر داده‌های حاصل از سنجش از دور و داده‌های کمکی. <i>جغرافیا و پایداری محیط</i>، ۱۰ (۳)، ۱۰۷-۱۲۴. doi: 10.22126/GES.2020.4227.2057</p>

مقدمه

هر آلاینده براساس غلظت و نوع آن، ویژگی‌های مکانی و زمانی خاص خود را دارد (هویت^۱، ۱۹۹۱) و به‌همین سبب تاکنون مطالعات گوناگونی براساس روش‌های قطعی و روش‌های تجربی - آماری به‌منظور مدل‌سازی زمانی و مکانی آلاینده‌ها صورت گرفته است (گاردنر و درلینگ^۲، ۱۹۹۹؛ هراست و همکاران^۳، ۲۰۰۹). در روش‌های قطعی از مدل‌های انتشار^۴ اتمسفری استفاده می‌شود که در آن‌ها سعی بر استفاده از روابط فیزیکی و شیمیایی حاکم بر تولید و انتشار آلاینده‌ها است تا بتوان مدلی برای شبیه‌سازی به‌دست آورد. این روش نیازمند حجم زیادی از داده‌های اندازه‌گیری شده نیست، ولی در مقابل به اطلاعات دقیق و کاملی از منابع انتشار آلودگی، چگونگی انتشار آن و فرایندهای فیزیکی و شیمیایی حاکم بر آن نیاز دارد. این درحالی است که اطلاعات مورد نیاز برای این‌نوع مدل‌سازی، به‌طور معمول به‌طور کامل وجود نداشته و به‌همین خاطر، ساده‌سازی و تخمین در مدل‌سازی لازم است که خود منجر به بروز خطا می‌شود (هراست، ۲۰۰۹). درمقابل، روش‌های تجربی - آماری به حجم زیادی از داده‌های اندازه‌گیری شده نیاز دارند تا رابطه میان متغیرهای مختلف و مقادیر اندازه‌گیری شده مدل به‌درستی تشخیص داده شود (شی و هریسون^۵، ۱۹۹۷).

روش‌های تجربی - آماری شامل مدل‌های خطی و غیر خطی هستند که هر یک مزایای خاص خود را دارند. روش‌های رگرسیون خطی یکی از پرکاربردترین روش‌های مدل‌سازی است که پایه ریاضی و محاسبات بسیار ساده‌ای دارد و برای مدل‌سازی در زمانی مناسب است که پیچیدگی رابطه بین متغیرها و مقادیر اندازه‌گیری شده کم باشد. روش‌های غیر خطی همچون شبکه عصبی مصنوعی^۶ به‌دلیل توانایی شناسایی روابط غیر خطی پیچیده، بین متغیرهای مستقل و وابسته در مدل‌سازی غیر خطی آماری استفاده می‌شود (عبودی و همکاران، ۱۳۹۴؛ ژانگ^۷ و همکاران، ۲۰۱۲ الف).

آلودگی هوای شهری از شرایط اقلیمی مانند میزان دما، رطوبت هوا، خصوصیات باد، میزان بارندگی و همچنین موقعیت جغرافیایی، افزایش جمعیت، منابع آلوده‌ساز، استفاده بیش از حد از سوخت‌های فسیلی و نیز ترافیک و الگوی نامناسب سیستم حمل‌ونقل شهری، تأثیر می‌پذیرد (اجتهادی، ۱۳۸۶)؛ همچنین عوامل هواشناسی از جمله وارونگی دمایی و استقرار مداوم سامانه‌های پرفشار همراه با هوای پایدار، به‌ویژه در دوره سرد سال، مانع جابه‌جایی هوا و باعث افزایش آلاینده‌های هوا می‌شوند (شرعی‌پور، ۱۳۸۸). توپوگرافی از قبیل کوه‌ها و ساختمان‌ها موجب انسداد بادهای و محبوس کردن آلاینده‌ها می‌شوند؛ همچنین این موانع ارتفاعی منجر به جلوگیری از رسیدن نور خورشید به سطح زمین شده و در منطقه سایه هوای سرد و خنک مستقر شده و منجر به راکدشدن هوا به‌صورت منطقه‌ای و محلی می‌شوند که در نتیجه این عوامل بر میزان تجمع آلاینده‌ها در آن نواحی افزوده می‌شود (خبری و همکاران، ۱۳۹۲). یکی دیگر از عوامل مؤثر بر کیفیت هوا، پدیده جزیره حرارتی شهری^۸ است که نه‌تنها منجر به کاهش کیفیت هوا شده، بلکه بر میزان بارندگی و افزایش آلودگی نیز اثر می‌گذارد (ژانگ و همکاران، ۲۰۱۲ ب). با توجه به عوامل متعدد تأثیرگذار بر میزان غلظت آلاینده‌ها، می‌توان دریافت که مدل‌سازی آن‌ها با استفاده از

1- Hewitt

2- Gardner & Dorling

3- Hrust

4- Diffusion

5- Shi & Harrison

6- Artificial Neural Network

7- Zhang

8- Urban Heat Island

روش‌های غیر خطی همچون شبکه عصبی مصنوعی می‌تواند نتایج بهتری را به دنبال داشته باشد. چنانچه در پژوهش‌های بسیاری برتری مدل شبکه عصبی پرسپترون چندلایه^۱ نسبت به مدل‌های رگرسیون خطی به منظور مدل‌سازی غلظت آلاینده‌هایی همچون اکسیدهای نیتروژن (گاردنر و درلینگ، ۱۹۹۹؛ الانگاسینقه^۲ و همکاران، ۲۰۱۴)، منواکسید نیتروژن (پرز و تریر^۳، ۲۰۰۱؛ مونترو-لورنزو^۴ و همکاران، ۲۰۱۳)، دی‌اکسید گوگرد (چلانی^۵ و همکاران، ۲۰۰۲)، ازن و ذرات معلق در هوا (مکندری^۶، ۲۰۰۲) نشان داده شده است؛ البته باید به این نکته اشاره شود که با توجه به نوع آلاینده، رابطه بین غلظت آلاینده‌ها و متغیرهای مورد استفاده می‌تواند تا حدودی متفاوت باشد (آگیره-باسارکو^۷ و همکاران، ۲۰۰۶).

با استفاده از قابلیت‌های داده‌های سنجش‌ازدور می‌توان به بررسی آلاینده‌های موجود در هوا، بررسی فرایندهای فیزیکی، تخمین غلظت آن و بررسی انتشار گازها و ذرات معلق موجود در اتمسفر پرداخت (مارتین^۸، ۲۰۰۸). در یکی از این پژوهش‌ها، تغییرات مکانی و زمانی آلاینده PM_{2.5} با استفاده از مدل‌های رگرسیون کاربری اراضی و رگرسیون خطی چندمتغیره و از داده‌های کاربری زمین، جمعیت، ترافیک، ویژگی‌های توپوگرافی و متغیرهای AOD داده‌های مادیس مربوط به سال ۲۰۰۵ به عنوان داده‌های مستقل در شهر فلوریدا به منظور مدل‌سازی داده‌های میانگین ماهانه PM_{2.5} استفاده شد (مائو^۹ و همکاران، ۲۰۱۲).

در پژوهشی دیگر، با استفاده از داده‌های هواشناسی لیدار و طیف‌سنج خورشیدی سنجنده مادیس مربوط به سال ۲۰۱۱ میزان ذرات معلق و عمق نوری آئروسول برای چهار شهر قبرس بررسی شدند (نیسانتزی^{۱۰} و همکاران، ۲۰۱۳)؛ همچنین غلظت متوسط سالانه آلاینده دی‌اکسید نیتروژن در منطقه‌ای در جنوب شرق اوکراین با استفاده از داده‌های سنجش‌ازدور اتمسفری مربوط به سال‌های ۲۰۰۱، ۲۰۰۴، ۲۰۰۵، ۲۰۰۶ و ۲۰۱۰ از سنسورهای ENVISAT/SCIAMACHY تعیین شد (خاریتانوف^{۱۱} و همکاران، ۲۰۱۳). در همین راستا، رابطه بین داده‌های عمق نوری آئروسول و غلظت زمینی PM₁₀ از سال ۲۰۰۸ تا ۲۰۱۲ در کشور کرواسی با استفاده از داده‌های سنجنده مادیس و داده‌های هواشناسی انجام گرفت (گرگوریچ^{۱۲} و همکاران، ۲۰۱۴)؛ همچنین در پژوهشی دیگر، به منظور بررسی و پیش‌بینی تغییرات مکانی و زمانی غلظت آلاینده PM₁₀ در چین طی سال‌های ۲۰۰۴ تا ۲۰۰۵ از داده‌های سنجنده مادیس، داده‌های هواشناسی و کاربری زمین استفاده شد (چن^{۱۳} و همکاران، ۲۰۱۸). براساس مطالعات پیش‌گفته، داده‌های سنجش‌ازدور می‌توانند اطلاعات مفیدی را به منظور مدل‌سازی آلاینده‌های هوا فراهم آورند.

به منظور مدل‌سازی غلظت آلاینده‌های هوا به عنوان متغیرهای وابسته، تاکنون متغیرهای مستقل متنوعی در مطالعات گوناگون استفاده شده‌اند که از آن جمله می‌توان به متغیرهایی همچون غلظت آلاینده‌ها در ساعت‌ها و روزهای گذشته (آل‌شیخ و همکاران، ۱۳۹۱؛ الانگاسینقه و همکاران، ۲۰۱۴؛ عبودی و همکاران، ۱۳۹۴)،

- 1- Multi Layer Perceptron (MLP)
- 2- Elangasinghe
- 3- Perez & Trier
- 4- Montero-Lorenzo
- 5- Chelani
- 6- McKendry
- 7- Agirre-Basurko
- 8- Martin
- 9- Mao
- 10- Nisantzi
- 11- Kharytonov
- 12- Grgurić
- 13- Chen

پارامترهای هواشناسی از جمله دما، سرعت و جهت باد (آل‌شیخ و همکاران، ۱۳۹۱؛ شمس‌الدینی^۱ و همکاران، ۲۰۱۷)، داده‌های مربوط به موقعیت ایستگاه نسبت به خیابان‌ها و بزرگراه‌های اصلی (رفیع‌پور و همکاران، ۱۳۹۲؛ شمس‌الدینی و همکاران، ۲۰۱۷)، ویژگی‌های محیطی از قبیل نزدیکی به فضاهای سبز (شمس‌الدینی و همکاران، ۲۰۱۷)، اشاره کرد. به‌طور مشخص این متغیرها در طی زمان و مکان تغییر می‌کنند که این مسئله خود می‌تواند بر پیچیدگی روابط بین متغیرهای وابسته و متغیر مستقل (غلظت آلاینده‌های هوا) بیفزاید. این مسئله در برآورد غلظت آلاینده‌ها برای مکان و زمانی خاص با استفاده از مدل‌هایی که برای زمان‌ها و مکان‌های دیگر توسعه یافته‌اند، اهمیت می‌یابد؛ برای مثال زمانی که دستگاه آلاینده‌سنج برای منطقه خاصی وجود ندارد، ناگزیر از مدل تخمین غلظت آلاینده‌ها که برای مناطق دیگر توسعه یافته‌اند برای تخمین میزان غلظت آلاینده‌ها در آن فصل استفاده شود. با توجه به بررسی‌های انجام‌شده، در زمینه امکان‌سنجی بسط مدل‌های زمانی و مکانی تخمین غلظت آلاینده‌ها به سایر زمان‌ها و مکان‌ها، مطالعات چندانی صورت نگرفته است و به همین سبب هدف نوشتار پیش رو بررسی این مسئله است. به دلیل عملکرد مناسب روش شبکه عصبی پرسپترون چندلایه در مدل‌سازی‌های غلظت آلاینده‌های هوا، در پژوهش حاضر از این مدل به منظور مدل‌سازی‌های مکانی و زمانی استفاده می‌شود.

مواد و روش‌ها

منطقه مورد مطالعه در جستار پیش رو شهر تهران با مساحتی حدود ۸۰۰ کیلومترمربع است که در محدوده جغرافیایی ۳۵° ۳۴' تا ۳۵° ۴۹' شمالی و ۵۱° ۰۴' تا ۵۱° ۳۶' شرقی در دامنه جنوبی رشته کوه البرز و حاشیه شمالی کویر مرکزی قرار گرفته است. ارتفاع این شهر در جنوب ۱۰۳۱ متر و در شمال ۱۸۳۸ متر از سطح دریا است که از جنوب به شمال افزایش یافته است. متوسط دمای تهران در ماه مرداد ۲۹/۸ و در دی به ۲/۸ درجه سلسیوس می‌رسد. دمای هوای تهران در طول سال در مرکز بیشتر از حومه بوده و سبب ایجاد جزیره حرارتی شده است. موقعیت جغرافیایی شهر تهران به همراه موقعیت ایستگاه‌های پایش آلودگی هوای شهر تهران متعلق به سازمان محیط‌زیست و ایستگاه‌های هواشناسی در شکل ۱ نمایش داده شده است.

داده‌های مورد استفاده

داده‌های مورد استفاده در پژوهش حاضر شامل دو دسته داده‌های محیطی و داده‌های مستخرج از تصاویر ماهواره‌ای هستند.

داده‌های محیطی

داده‌های محیطی مورد استفاده در نوشتار پیش رو شامل داده‌های غلظت آلاینده‌ها، داده‌های هواشناسی، داده‌های ترافیکی، داده‌های کاربری اراضی و جمعیت هستند.

داده‌های غلظت آلاینده‌ها

غلظت آلاینده‌های منواکسید کربن و دی‌اکسید نیتروژن ایستگاه‌های پایش آلودگی هوا مربوط به سازمان حفاظت محیط‌زیست واقع در سطح شهر تهران برای تمامی روزهای سال ۱۳۹۴ و نیمه اول سال ۱۳۹۵ به صورت ساعتی از سازمان محیط‌زیست استان تهران تهیه شد که با توجه به نقص داده‌ها در برخی از ایستگاه‌ها، از میان دوازده

ایستگاه موجود، هشت ایستگاه دانشگاه شهید بهشتی، چشمه، دانشگاه علم و صنعت، شهرداری منطقه ۱۵، پاسداران، پارک رازی، پارک سلامت و پارک شکوفه انتخاب شد.

داده‌های هواشناسی

داده‌های هواشناسی مورد استفاده در جستار پیش رو شامل سرعت و جهت باد، دمای هوا، فشار هوا، رطوبت نسبی و پوشش کلی ابر است؛ از این میان، اطلاعات دما و رطوبت نسبی مربوط به هشت ایستگاه ثابت پایش آلودگی هوا متعلق به سازمان محیط زیست، سرعت و جهت باد از ایستگاه‌های سازمان هواشناسی واقع در سطح شهر تهران (مهرآباد، ژئوفیزیک، چیتگر و شمیران) و داده‌های فشار و پوشش کلی ابر از سایت مرکز پیش‌بینی‌های متوسط هواشناسی اروپا^۱ استخراج و میانگین آن‌ها محاسبه شد. با توجه به وجود چندین ایستگاه در منطقه مورد مطالعه، برای اختصاص اطلاعات ایستگاه‌های هواشناسی به ایستگاه‌های پایش آلودگی هوا از روش پلیگون‌بندی تیسن استفاده شد.

داده‌های ترافیکی

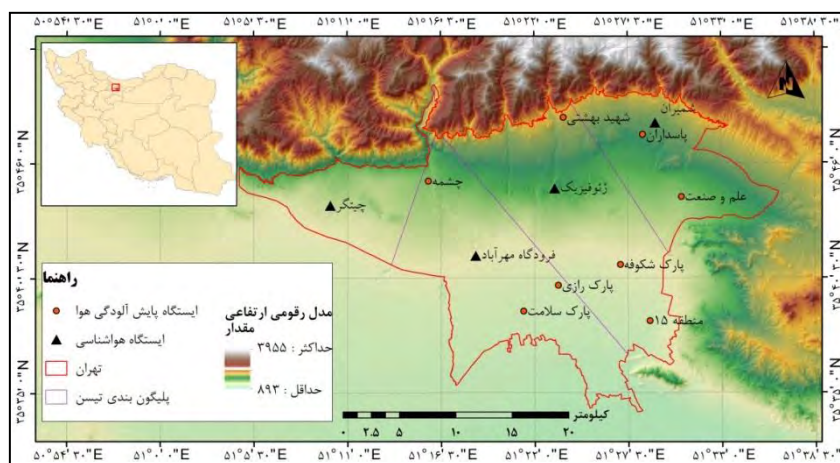
به دلیل قابل تهیه نبودن اطلاعات ترافیکی شهر تهران به جای اطلاعات واقعی ترافیک، از شاخص ترافیک استفاده شد. برای محاسبه شاخص ترافیک، بافرهایی به شعاع صد متر تا شعاع یک کیلومتری اطراف هر ایستگاه ایجاد شد. با توجه به اینکه غلظت آلاینده‌ها با فاصله از خیابان‌ها تغییر می‌کند، فاصله خیابان‌ها از هر ایستگاه محاسبه و در عرض خیابان ضرب و در نهایت میانگین وزنی آن‌ها محاسبه شد (شمس‌الدینی و همکاران، ۲۰۱۷).

مساحت کاربری زمین

با توجه به تأثیر نوع کاربری سطح زمین همچون فضای سبز، صنعتی و غیره بر کیفیت هوا، به منظور اخذ اطلاعات کاربری زمین، فایل رقومی شهر تهران از شهرداری تهیه شد و مساحت هر کاربری در بافرهای با شعاع صد متر تا یک کیلومتر اطراف هر ایستگاه محاسبه شد.

جمعیت

لایه رقومی سرشماری جمعیت سال ۱۳۹۰ تهران در مقیاس محله‌ای از مرکز آمار ایران تهیه شد و جمعیت اطراف هر ایستگاه پایش آلودگی تا شعاع یک کیلومتری استخراج شد.



شکل ۱. موقعیت جغرافیایی شهر تهران و ایستگاه‌های پایش آلودگی هوا و هواشناسی

داده‌های ماهواره‌ای

در نوشتار پیش رو به منظور استخراج میانگین پوشش گیاهی و جزایر حرارتی شهری از تصاویر سنجنده‌های OLI/TIRS ماهواره لندست-۸ به شماره مسیر ۱۶۴ و ردیف ۳۵ که از سایت سازمان زمین‌شناسی آمریکا^۱ و محصول دمای سطح زمین سنجنده مادیس که از سایت سیستم داده مشاهدات زمینی اداره کل ملی هوانوردی و فضایی آمریکا^۲ به شماره افقی ۲۲ و عمودی ۵ دریافت شدند، استفاده شد. تعداد شش تصویر بدون ابر ماهواره لندست-۸ به‌مثابه نماینده هر فصل (جدول ۱) و تعداد ۵۴۲ محصول دمای سطح زمین سنجنده مادیس به‌صورت روزانه برای بازه زمانی مورد مطالعه دریافت شد. پس از دانلود این داده‌ها، ضمن بررسی تصاویر مادیس، تعداد ۳۸۱ تصویر به‌عنوان تصاویر بدون ابر برای محدوده شهر تهران انتخاب شده و در ادامه پژوهش، از آن‌ها استفاده شد.

پیش‌پردازش تصاویر ماهواره‌ای لندست-۸

تصاویر لندست نسبت به هم رجیستر بودند بنابراین، نیازی به تصحیح هندسی وجود نداشت^۳. به‌منظور تصحیح اتمسفری از روش تصحیح اتمسفری و توپوگرافی^۴ استفاده شد. این روش برای انجام تصحیح اتمسفری ابتدا با استفاده از مدل رقومی ارتفاعی، زمان تصویربرداری، زاویه زینت و آزیموت خورشید تصحیح توپوگرافی را انجام می‌دهد. از مدل رقومی ارتفاعی حاصل از تصاویر سنجنده پالسار^۵ ماهواره آلوس^۶ با قدرت تفکیک مکانی ۱۲/۵ متر برای انجام تصحیح اتمسفری استفاده شد. این داده از سایت پلت‌فرم داده‌های ماهواره‌ای آلاسکا^۷ دریافت شد.

شاخص پوشش گیاهی تفاضلی نرمال شده

به‌منظور محاسبه میانگین پوشش گیاهی از شاخص پوشش گیاهی تفاضلی نرمال شده^۸ استفاده و در نهایت میانگین آن در بافرهایی با شعاع صد متر تا یک کیلومتر اطراف هر ایستگاه استخراج شد.

فاصله از جزیره حرارتی

روش‌های مختلفی برای استخراج دما از باند حرارتی ارائه شده است؛ که در پژوهش حاضر از باند ۱۰ در محدوده طیفی ۱۰/۶۰ تا ۱۱/۱۹ میکرومتر سنجنده مادون قرمز حرارتی^۹ ماهواره لندست-۸ و به‌روش تک‌باندی (آرتیس و کارناهان^{۱۰}، ۱۹۸۲)، دمای سطح زمین استخراج شد.

جدول ۱. تاریخ تصاویر اخذ شده ماهواره لندست-۸

سنجنده	تاریخ تصویر
OLI/TIRS	۲۰۱۵/۰۵/۱۷
OLI/TIRS	۲۰۱۵/۰۸/۰۵
OLI/TIRS	۲۰۱۵/۱۱/۲۵
OLI/TIRS	۲۰۱۶/۰۲/۱۳
OLI/TIRS	۲۰۱۶/۰۵/۱۹
OLI/TIRS	۲۰۱۶/۰۸/۰۷

1- <https://earthexplorer.usgs.gov/>2- <https://ladsweb.modaps.eosdis.nasa.gov/search/order/>3- <https://www.usgs.gov/land-resources/nli/landsat/landsat-level-1-processing-details>

4- Atmospheric and Topographic Correction (ATCOR)

5- PALSAR

6- ALOS

7- <https://vertex.daac.asf.alaska.edu/>

8- Normalized Difference Vegetation Index (NDVI)

9- TIRS

10- Artis & Carnahan

پس از انجام پیش‌پردازش‌های لازم برای تبدیل دمای درخشندگی به دمای جنبشی سطح لازم است که گسیل‌مندی سطح زمین^۱ در تصویر محاسبه شود. روش‌های مختلفی برای محاسبه گسیل‌مندی وجود دارد، با توجه به اینکه در فضای شهر سطوح مختلف دارای گسیل‌مندی متفاوت هستند؛ بنابراین با استفاده از روش گسیل‌مندی برپایه طبقه‌بندی^۲ ارائه‌شده اسنایدر^۳ و همکاران (۱۹۹۸) گسیل‌مندی سطح زمین در تصاویر ماهواره‌ای محاسبه شد (اسنایدر و همکاران، ۱۹۹۸). در این روش، کاربری اراضی/ پوشش زمین از تصاویر ماهواره‌ای و با روش حداکثر شباهت^۴ که روش رایجی در سنجش‌ازدور است، استخراج شد. صحت نقشه‌های تولیدشده با استفاده از برداشت داده‌های آزمایشی در گوگل‌ارث^۵، تعیین شد و ضرایب کاپا بین ۰/۸۶ تا ۰/۹ برای تصاویر مختلف به دست آمد که صحت قابل قبولی است؛ سپس با استفاده از کتابخانه گسیل‌مندی پدیده‌ها، مقدار آن به هر طبقه کاربری اختصاص داده و با استفاده از روش تک‌کاناله (رابطه ۱) دمای سطح زمین^۶ بازیابی شد.

$$LST = \frac{T_b}{1 + \left(\lambda \times \frac{T_b}{\rho}\right) \ln \varepsilon} \quad \text{رابطه ۱}$$

$$\rho = h \times \frac{c}{\sigma}$$

که در این معادله، T_b عبارت از تصویر دمای درخشندگی؛ λ : نشان‌دهنده طول موج رادیانس تابشی؛ h : ثابت پلانک^۷ (6.626×10^{-34} ژول ثانیه)، σ : ثابت استفان بولتزمن^۸ (5.67×10^{-8} وات بر متر مربع بر درجه کلوین به توان چهارم)، c : سرعت نور^۹ (2.998×10^8 متر بر ثانیه) و ε : تصویر گسیل‌مندی سطح زمین است. در نهایت برای استخراج جزایر حرارتی، تصاویر دمای سطح زمین بازیابی و با استفاده از رابطه ۲ نرمال شده و سپس با استفاده از جدول ۲ طبقه‌بندی شدند:

$$N_i = \frac{T_i - T_{\min}}{T_{\max} - T_{\min}} \quad \text{رابطه ۲}$$

که در این رابطه N_i : نشان‌دهنده مقدار نرمال‌شده پیکسل i ، T_i : بیان‌کننده مقدار دمای مطلق پیکسل i ، T_{\min} : حداقل دمای هر تصویر و T_{\max} : حداکثر دمای هر تصویر است.

جدول ۲. چگونگی طبقه‌بندی دمای سطح زمین بازیابی‌شده

شماره طبقه	نام طبقه	دامنه مربوط به هر طبقه
۱	نواحی بسیارخنک	$T \leq T_{mean} - 1.5 \text{ std}$
۲	نواحی خنک	$T_{mean} - 1.5 \text{ std} < T \leq T_{mean} - \text{std}$
۳	نواحی با دمای متوسط	$T_{mean} - \text{std} < T \leq T_{mean} + \text{std}$
۴	نواحی گرم	$T_{mean} + \text{std} < T \leq T_{mean} + 1.5 \text{ std}$
۵	نواحی بسیارگرم	$T > T_{mean} + 1.5 \text{ std}$

- 1- Land Surface Emissivity
- 2- Classification-Based Emissivity Method (CBEM)
- 3- Snyder
- 4- Maximum Likelihood Classification (MLC)
- 5- Google Earth
- 6- Land Surface Temperature
- 7- Planck's constant
- 8- Boltzmann constant
- 9- Velocity of light

دمای سطح زمین

برای دریافت اطلاعات دمای سطح زمین در هر روز در بازه زمانی مربوط به داده‌های غلظت آلاینده‌ها، از محصول MOD11A1 سنجنده مادیس که روی سکوی ترا قرار دارد، استفاده شد که از سایت مربوطه به صورت روزانه برای بازه زمانی مورد مطالعه دریافت شد. محصول MOD11A1 دربرگیرنده دمای سطح و ضریب گسیل‌مندی با قدرت تفکیک مکانی یک کیلومتر است که با شرایط آسمان صاف تولید می‌شود؛ سپس میانگین پیکسل‌های آن در بافری با شعاع یک کیلومتر اطراف هر ایستگاه پایش آلودگی هوا محاسبه شد.

تبدیل موجک

هدف از تبدیل موجک گسسته، تجزیه سیگنال اصلی به موجک‌هایی در سطوح مختلف است؛ به طوری که سیگنال اصلی براساس رابطه ۳، به دو دسته سیگنال کلی و سیگنال جزئی تجزیه می‌شود:

$$s(n) = D_1(n) + D_2(n) + \dots + D_j(n) + A_j(n) \quad \text{رابطه ۳}$$

در این رابطه، $s(n)$: عبارت از سیگنال اصلی؛ $A_j(n)$: نشان‌دهنده کلیات سیگنال اصلی در سطح j ام و $D_j(n)$: بیان‌کننده جزئیات سیگنال اصلی در سطح j ام است.

در پژوهش حاضر به منظور افزایش صحت تخمین حاصل از مقادیر بیشینه غلظت آلاینده، از تبدیل موجک بر روی مقادیر بیشینه غلظت آلاینده‌ها تا چهار روز قبل استفاده شد. از میان توابع تبدیل موجک، تابع DB به دلیل ارائه کوچک‌ترین تغییرات سیگنال در سطوح مختلف انتخاب شد (سیوک و اسوسک^۱، ۲۰۱۲). در این تابع موج، موجک مادر به صورت متعامل و دو متعاملی است و می‌تواند در تحلیل موجک گسسته و پیوسته استفاده شود (حسینی شفیق و همکاران، ۱۳۹۵). به منظور تعیین سطح بهینه j ، از رابطه ۴ استفاده شد:

$$\frac{\text{std}(A_j)}{\text{std}(x)} < 0.1 \quad \text{رابطه ۴}$$

در این رابطه، $\text{std}(A_j)$: انحراف معیار کلیات سیگنال و $\text{std}(x)$: انحراف معیار سیگنال اصلی است. با توجه به این رابطه، مقدار $j=5$ به دست آمد و روی داده‌های بیشینه غلظت آلاینده‌های روزهای قبل تبدیل موجک DB در پنج سطح اعمال شد. در این راستا با توجه به رابطه تبدیل موجک، مقادیر سیگنال کلیات در سطح پنجم و سیگنال جزئیات در پنج سطح برای ورود به مدل‌سازی آماده شدند.

انتخاب ویژگی

انتخاب ویژگی یکی از مراحل مهم و اساسی در زمینه شناسایی الگو، یادگیری ماشین و داده‌کاوی است. هدف از انتخاب ویژگی، انتخاب بهینه‌ترین زیرمجموعه ویژگی از کل فضای ویژگی‌های اصلی مسئله مورد نظر است، به طوری که ضمن کاهش ابعاد بتوان به صحت مدل‌سازی مطلوبی دست یافت (کولر و سهامی^۲، ۱۹۹۶). برای انتخاب ویژگی، الگوریتم‌های فراوانی ارائه شده است. در پژوهش پیش رو، به منظور انتخاب ویژگی‌های بهینه از روش انتخاب ویژگی با استفاده از جنگل تصادفی^۳ استفاده شد. در این روش، در گام اول، الگوریتم جنگل تصادفی بر روی تمامی متغیرهای مستقل اعمال می‌شود و با انجام مدل‌سازی با تمامی متغیرها، میانگین مربعات خطای مدل محاسبه می‌شود. در گام بعد با محاسبه اهمیت متغیرها براساس میزان باقی‌مانده‌های خطا با روش جنگل

1- Siwek & Osowsk

2- Koller & Sahami

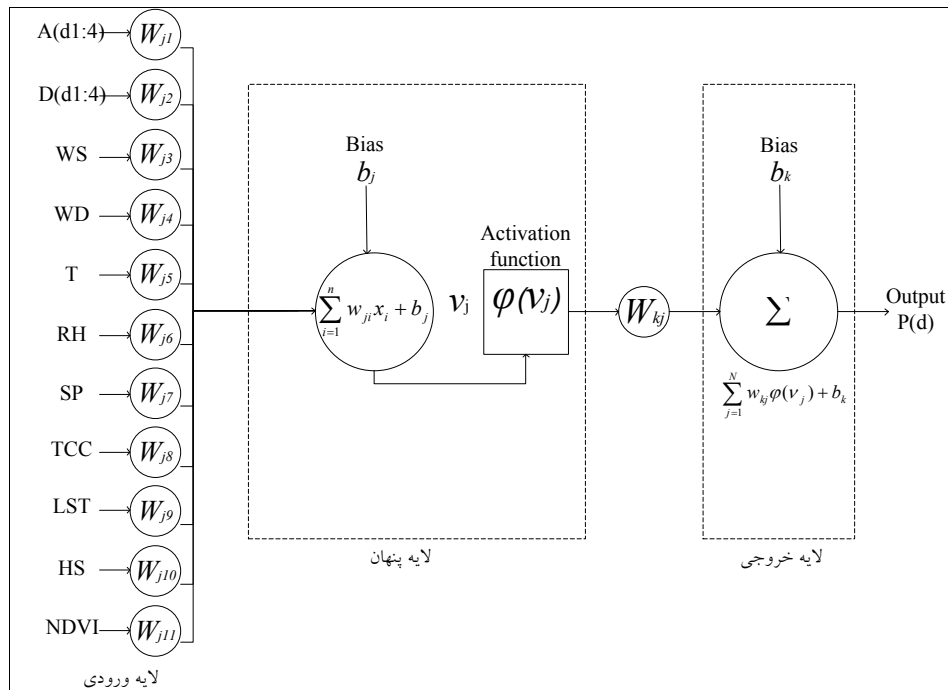
3- Random Forest

تصادفی، متغیرهای کم‌اهمیت حذف می‌شوند. حذف متغیرهای کم‌اهمیت به صورت مرحله‌ای است. با حذف متغیرهای کم‌اهمیت در هر مرحله، روش جنگل تصادفی با سایر متغیرها اجرا می‌شود و میزان خطای مدل محاسبه می‌شود.

برای دستیابی به تعداد نهایی متغیرها، در هر مرحله، از آزمون تی زوجی روی مقادیر قدر مطلق باقی‌مانده‌ها در آن مرحله و مرحله قبل (در صورت افزایش میزان خطا) استفاده می‌شود. این آزمون به منظور بررسی معنی‌دار بودن میزان افزایش خطای دو مدل استفاده می‌شود. در صورت معنی‌دار بودن این اختلاف، تعداد نهایی متغیرها همان تعداد متغیرهای مرحله قبل می‌شود، اما اگر که این اختلاف معنی‌دار نبود، تعداد متغیرهای همان مرحله به‌عنوان تعداد نهایی انتخاب می‌شوند (شمس‌الدینی و همکاران، ۲۰۱۴). پس از اعمال این روش، از ۱۵۶ خصیصه ورودی، ۴۲ متغیر مستقل شامل پنج سطح کلیات و پنج سطح جزئیات حاصل از تبدیل موجک مقادیر آلاینده‌ها در چهار روز قبل، سرعت باد، جهت باد، دمای هوا، رطوبت نسبی، درصد پوشش کالی ابر، فشار هوا، دمای سطح زمین، فاصله از جزیره حرارتی و شاخص پوشش گیاهی باقی ماند.

مدل‌سازی

از ویژگی‌های اصلی شبکه‌های عصبی مصنوعی ایجاد مدل رگرسیونی بین داده‌های ورودی و خروجی بدون لحاظ کردن روابط فیزیکی بین آن‌ها است (کولیبالی^۱ و همکاران، ۲۰۰۵). مدل شبکه عصبی پرسپترون چندلایه متشکل از شبکه‌ای از عناصر ساده و ارتباطها است. تعداد نورون ورودی و خروجی شبکه‌های عصبی مصنوعی با ماهیت مسئله تعیین می‌شود. هر نورون مجموع وزن‌داری از ورودی‌های نورون مجاور را محاسبه می‌کند. این پردازش‌ها با استفاده از یک تابع فعال‌سازی انجام و در لایه بعدی توزیع می‌شود (کاکونن^۲ و همکاران، ۲۰۰۳). شکل ۲ ساختار شبکه عصبی را نشان می‌دهد.



شکل ۲. ساختار شبکه عصبی به‌همراه تبدیل موجک

1- Coulibaly
2- Kukkonen

به‌منظور انجام مدل‌سازی، متغیرهای انتخاب‌شده با روش جنگل تصادفی به‌عنوان ورودی مدل و بیشینه غلظت آلاینده‌ها در هرروز به‌منزله خروجی مدل معرفی شد. مجموعه داده‌ها برای مدل‌سازی زمانی آلاینده‌های منواکسید کربن و دی‌اکسید نیتروژن شامل ۴۴۱۵ نمونه از بیشینه غلظت روزانه آلاینده و ۴۲ خصیصه از هشت ایستگاه پایش کیفیت هوای ذکرشده در بازه زمانی هجده ماه شد. در این میان، روزهایی که تصاویر ماهواره‌ای دمای سطح زمین سنجنده مادیس به‌دلیل وجود ابر فاقد داده بودند؛ از مجموعه داده‌ها حذف شدند و درنهایت تعداد ۳۰۴۸ نمونه به‌عنوان ورودی مدل باقی ماندند؛ همچنین برای آموزش، اعتبارسنجی و ارزیابی مدل‌ها، داده‌ها به سه مجموعه شامل ۷۰٪ برای آموزش، ۱۵٪ برای اعتبارسنجی و ۱۵٪ برای آزمایش مدل به‌صورت تصادفی تقسیم شدند.

با توجه به مقیاس مختلف داده‌ها برای جلوگیری از ناهمگونی شبکه و از بین بردن ابعاد مختلف داده‌ها، ورودی‌های شبکه عصبی باید نرمال شوند؛ بدین منظور خصیصه‌های مورد استفاده با رابطه ۵ نرمال شده و در بازه [۰ و ۱] قرار گرفتند.

$$X_{norm} = \frac{(x_i - x_{min})}{(x_{max} - x_{min})} \quad \text{رابطه ۵}$$

در این رابطه، X_{norm} : مقدار نرمال‌شده یک خصیصه؛ x_i : مقادیر اصلی؛ x_{min} : مقادیر کمینه و x_{max} : مقادیر بیشینه هر خصیصه است.

به‌منظور اجرای شبکه عصبی پرسپترون چندلایه از تابع آموزشی لوبز-مارکوارت (LM)، با حد بالای نرخ آموزش ۰/۱۰ و ثابت مومنتوم ۰/۰۱ استفاده شد؛ همچنین برای یافتن تعداد نورون‌های مناسب، شبکه چندین بار اجرا شد و تعداد نورون مناسب در لایه پنهان برای هر آلاینده تعیین شد که با توجه به تنوع آلاینده‌ها بین ۳ تا ۵ نورون متغیر بود.

برای قابلیت‌سنجی مدل‌های مکانی به‌منظور پیش‌بینی آلاینده‌ها، داده‌های مربوط به یک ایستگاه به‌عنوان داده آزمایشی و داده‌های هفت ایستگاه دیگر به‌مثابه داده‌های آموزشی شامل ۹۰٪ و اعتبارسنجی شامل ۱۰٪ قرار داده شد و این مدل‌سازی برای هر هشت ایستگاه انجام شد.

ارزیابی عملکرد مدل

به‌منظور ارزیابی عملکرد مدل‌ها، از رابطه‌های ۶ تا ۹ استفاده شد که به ترتیب شاخص‌های ریشه میانگین مربعات خطا، خطای استاندارد تخمین^۲، ضریب تعیین^۳ و درصد خطا هستند.

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (P_i - O_i)^2} \quad \text{رابطه ۶}$$

$$SEE = \sqrt{\frac{1}{N-m-1} \sum_{i=1}^N (P_i - O_i)^2} \quad \text{رابطه ۷}$$

$$R^2 = \frac{\sum_{i=1}^N (P_i - \bar{O})^2}{\sum_{i=1}^N (O_i - \bar{O})^2} \quad \text{رابطه ۸}$$

1- Root Mean Square Error (RMSE)

2- Standard Error of Estimation (SEE)

3- R2

$$\%Error = \left(\frac{RMSE}{\bar{O}} \right) \times 100$$

رابطه ۹

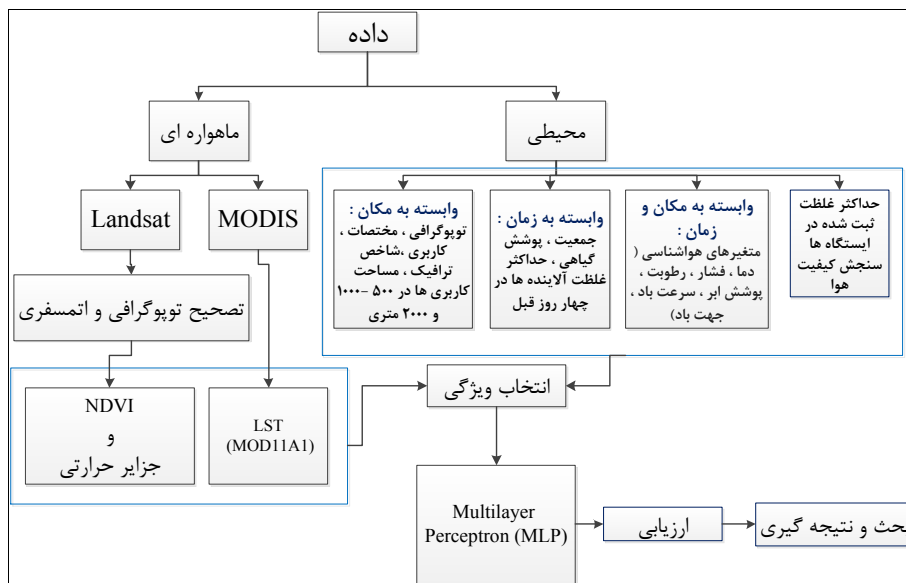
در معادلات بالا، P_i : عبارت از مقادیر پیش‌بینی شده؛ O_i : بیان‌کننده مقادیر مشاهده‌شده؛ \bar{O} : نشان‌دهنده میانگین مقادیر مشاهده‌شده؛ N : تعداد نمونه‌ها و m : تعداد پیش‌بینی‌کننده‌ها است. برای ارزیابی آماری مدل به‌دست‌آمده از آزمون تی نمونه جفت‌شده استفاده شد. مراحل انجام پژوهش در ادامه آورده شده است (شکل ۳).

نتایج

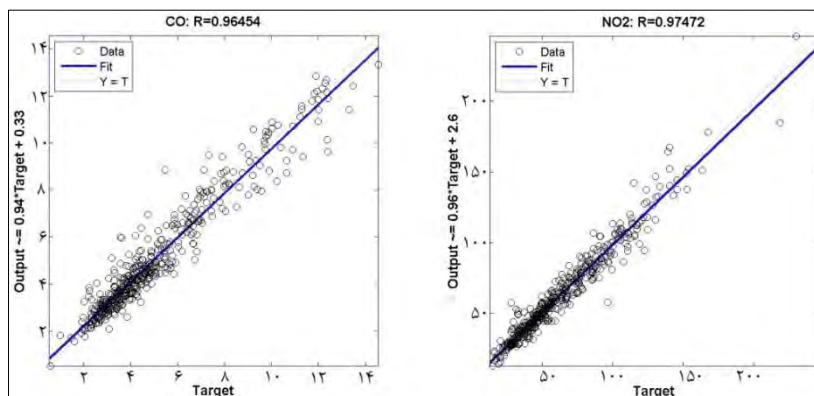
همان‌گونه که گفته شد، هدف اصلی جستار پیش‌رو، بررسی امکان بسط زمانی - مکانی مدل‌های حاصل از روش شبکه عصبی پرسپترون چندلایه به‌منظور پیش‌بینی میزان آلاینده‌های موجود در هوای شهر تهران است. پس از اجرای روش شبکه عصبی پرسپترون چندلایه، نتایج به‌دست‌آمده برای آلاینده‌ها به‌صورت زیر است:

نتایج حاصل از بررسی قابلیت‌سنجی مدل به‌منظور پیش‌بینی زمانی آلاینده‌ها

پراکندگی مقادیر اندازه‌گیری‌شده و مقادیر داده‌های پیش‌بینی‌شده (شکل ۴) و نتایج عملکرد مدل در مدل‌سازی زمانی آلاینده‌های منواکسید کربن و دی‌اکسید نیتروژن (جدول ۲) در ادامه ارائه شده است.



شکل ۳. نمودار جریان‌ی مراحل انجام پژوهش



شکل ۴. پراکندگی مقادیر اندازه‌گیری‌شده و مقادیر داده‌های پیش‌بینی‌شده آلاینده‌های منواکسید کربن (چپ) و دی‌اکسید نیتروژن (راست)

با توجه به نتایج، مدل‌های پیش‌بینی‌کننده بیشینه غلظت آلاینده‌های منواکسید کربن و دی‌اکسید نیتروژن توانسته‌اند تا به خوبی تغییرات این آلاینده‌ها را مدل‌سازی کنند (شکل ۴)؛ همچنین ضریب تعیین این مدل‌ها بالاتر از ۰/۹۳ است (جدول ۲). براساس نتایج به دست آمده، مدل پیش‌بینی‌کننده بیشینه غلظت آلاینده دی‌اکسید نیتروژن با ۱۱/۴۶٪ خطای پیش‌بینی بهتر از مدل پیش‌بینی‌کننده بیشینه غلظت منواکسید کربن عمل کرده است (جدول ۳).

نتایج حاصل از بررسی قابلیت‌سنجی مدل به‌منظور پیش‌بینی مکانی آلاینده‌ها

نتایج عملکرد مدل در پیش‌بینی مکانی مقادیر آلاینده منواکسید کربن (جدول ۴) و نتایج عملکرد مدل در پیش‌بینی مکانی مقادیر دی‌اکسید نیتروژن (جدول ۵) در ادامه ارائه شده‌اند.

با توجه به نتایج به دست آمده، مدل‌های مکانی حاصل، بیشینه غلظت آلاینده منواکسید کربن را در ایستگاه پارک سلامت و ایستگاه چشمه به ترتیب با بیشترین و کمترین میزان درصد خطا برآورد کرده‌اند (جدول ۴). به‌منظور بررسی بیشتر وضعیت این ایستگاه‌ها، بین متغیرهای مختلف ورودی به صورت مجزا و مقادیر بیشینه غلظت آلاینده‌ها از همبستگی اسپیرمن استفاده شد. نتایج حاصل از همبستگی اسپیرمن نشان داد در ایستگاه پارک سلامت رابطه معنی‌داری بین متغیرهای دمای هوا، دمای سطح زمین، درصد پوشش ابر، فاصله از جزیره حرارتی، میانگین پوشش گیاهی تا شعاع ۴۰۰ متری وجود دارد؛ همچنین در ایستگاه چشمه، رابطه معنی‌داری بین متغیرهای سرعت باد، جهت باد، فشار هوا، رطوبت نسبی، درصد پوشش ابر و میانگین پوشش گیاهی در شعاع ۵۰۰ تا ۱۰۰۰ متری با غلظت آلاینده منواکسید کربن وجود دارد؛ همچنین نتایج نشان داد که ایستگاه دانشگاه شهید بهشتی بیشترین و ایستگاه پارک شکوفه کمترین درصد خطای برآورد را در پیش‌بینی مقدار غلظت آلاینده دی‌اکسید نیتروژن داشته است (جدول ۵).

نتایج حاصل از همبستگی اسپیرمن نشان داد در ایستگاه شهید بهشتی رابطه معنی‌داری بین متغیرهای سرعت باد، جهت باد، دمای هوا، فشار هوا، درصد پوشش ابر، دمای سطح زمین و میانگین پوشش گیاهی و در ایستگاه پارک شکوفه رابطه معنی‌داری بین متغیرهای سرعت باد، دمای هوا، فشار هوا، رطوبت نسبی، درصد پوشش ابر، فاصله از جزیره حرارتی و دمای سطح زمین و میانگین پوشش گیاهی تا شعاع ۳۰۰ متر با غلظت آلاینده دی‌اکسید نیتروژن وجود دارد.

جدول ۳. نتایج عملکرد مدل در مدل‌سازی زمانی آلاینده‌های منواکسید کربن و دی‌اکسید نیتروژن

آلاینده	خطای استاندارد تخمین	ریشه میانگین مربعات خطا	ضریب تعیین	درصد خطای برآورد
منواکسید کربن	۰/۷۰	۰/۶۶	۰/۹۳	۱۳/۱۳
دی‌اکسید نیتروژن	۷/۶۳	۷/۲۶	۰/۹۵	۱۱/۴۶

جدول ۴. نتایج عملکرد مدل در مدل‌سازی به‌منظور پیش‌بینی مکانی مقادیر آلاینده منواکسید کربن

ایستگاه	خطای استاندارد تخمین	ریشه میانگین مربعات خطا	ضریب تعیین	درصد خطای برآورد
دانشگاه شهید بهشتی	۰/۴۶	۰/۴۳	۰/۸۶	۱۳/۹۲
چشمه	۰/۳۲	۰/۳۰	۰/۸۳	۸/۳۴
دانشگاه علم و صنعت	۰/۴۸	۰/۴۶	۰/۸۳	۱۲/۳۷
شهرداری منطقه ۱۵	۱/۱۰	۱/۰۴	۰/۸۴	۱۶/۷۰
پاسداران	۱/۰۲	۰/۹۶	۰/۸۵	۱۱/۸۲
پارک رازی	۰/۸۴	۰/۷۹	۰/۸۷	۱۵/۶۵
پارک سلامت	۰/۹۲	۰/۸۷	۰/۸۶	۱۷/۰۳
شکوفه	۰/۸۲	۰/۷۷	۰/۸۲	۱۳/۸۸

جدول ۵. نتایج عملکرد مدل در مدل‌سازی به‌منظور پیش‌بینی مکانی مقادیر آلاینده دی‌اکسید نیتروژن

ایستگاه	خطای استاندارد تخمین	ریشه میانگین مربعات خطا	ضریب تعیین	درصد خطای برآورد
دانشگاه شهید بهشتی	۶/۲۹	۵/۹۳	۰/۹۱	۱۴/۳۹
چشمه	۴/۷۹	۴/۵۱	۰/۹۳	۱۰/۵۶
دانشگاه علم و صنعت	۷/۲۵	۶/۸۳	۰/۹۰	۱۲/۳۸
شهرداری منطقه ۱۵	۶/۳۶	۵/۹۹	۰/۹۱	۱۱/۱۷
پاسداران	۹/۸۲	۹/۲۵	۰/۹۱	۱۰/۶۹
پارک رازی	۱۰/۱۳	۹/۵۴	۰/۹۲	۱۲/۰۵
پارک سلامت	۱۱/۳۴	۱۰/۶۸	۰/۹۱	۱۱/۹۰
شکوفه	۵/۴۷	۵/۱۵	۰/۸۹	۱۰/۴۰

همچنین به‌منظور بررسی آماری عملکرد مدل‌های ایجادشده، آزمون تی جفت‌شده برای مقایسه آماری مقادیر پیش‌بینی شده و اندازه‌گیری شده بیشینه غلظت آلاینده‌ها استفاده شد. نتایج حاصل از این آزمون نشان داد که اختلاف معنی‌داری بین مقادیر پیش‌بینی شده حاصل از مدل‌ها و مقادیر غلظت ثبت‌شده آلاینده‌های منواکسید کربن و دی‌اکسید کربن در سطح معنی‌داری ۵٪ وجود ندارد که خود نشان‌دهنده عملکرد مطلوب مدل‌های توسعه‌داده‌شده در نوشتار پیش رو است.

بحث

پس از انجام مدل‌سازی زمانی و تعیین صحت عملکرد مدل‌های به‌دست‌آمده برای پیش‌بینی بیشینه غلظت منواکسید کربن و دی‌اکسید نیتروژن، مشخص شد که دی‌اکسید نیتروژن با صحت بیشتری نسبت به منواکسید کربن پیش‌بینی می‌شود (جدول ۳)؛ همچنین در مدل‌سازی مکانی این آلاینده‌ها نیز، دی‌اکسید نیتروژن نسبت به آلاینده دیگر، در بیشتر ایستگاه‌ها خطای کمتری داشته و در تمامی موارد، مدل پیش‌بینی‌کننده این آلاینده توانسته است با ضریب تعیین بالاتری، غلظت این آلاینده را پیش‌بینی کند (جدول ۴ و ۵). با توجه به اینکه واکنش‌های شیمیایی و فتوشیمیایی که منجر به تولید و انتشار این آلاینده‌ها می‌شود، متفاوت است (کاراتزاس و کالتساتوس^۱، ۲۰۰۷).

مهم‌ترین عامل مؤثر بر غلظت آلاینده منواکسید کربن حجم ترافیک است و براساس گزارش شرکت کنترل کیفیت هوای سال ۱۳۹۴ تهران سهم منابع متحرک در تولید آلاینده‌های منواکسید کربن ۹۸٪ است و این درحالی است که این عامل در تولید آلاینده اکسیدهای نیتروژن ۴۷٪ بوده و بقیه با منابع ساکن تولید می‌شود (متکان و همکاران، ۱۳۸۸). منابعی همچون حجم ترافیک، منبع متغییری است که می‌تواند باعث تغییرپذیری بالای منواکسید کربن شود. تفاوت در منابع ایجاد این آلاینده‌ها می‌تواند در نوع الگوی روزانه تغییرات این آلاینده‌ها مؤثر باشد. در مناطق شهری، به‌طور معمول تغییرات غلظت اکسیدهای نیتروژن از یک الگوی روزانه منظم پیروی می‌کند (نلسون و لی^۲، ۱۹۷۷). این درحالی است که با توجه به بررسی‌های انجام‌شده روی تغییرات غلظت آلاینده‌ها در این مطالعه، تغییرات غلظت منواکسید کربن بیشتر از دی‌اکسید نیتروژن بوده و پایداری کمتری دارد؛ بنابراین می‌توان گفت که ناپایداری بیشتر غلظت منواکسید کربن و تغییرپذیری بالاتر این آلاینده می‌تواند یکی از دلایل مهم صحت پایین‌تر مدل‌های پیش‌بینی‌کننده این آلاینده نسبت به دی‌اکسید نیتروژن باشد.

استفاده از شبکه عصبی مصنوعی در کنار استفاده از روش تبدیل موجک بر روی داده‌های غلظت آلاینده‌ها در

1- Karatzas & Kaltsatos

2- Knelson & Lee

روزهای قبل، توانست نتایج مطلوبی (خطای تخمین کمتر از ۱۴٪ برای مدل‌های زمانی و کمتر از ۱۷٪ برای مدل‌های مکانی) به‌دنبال داشته باشد؛ همچنین تست آماری مورد استفاده در نوشتار پیش رو بیان‌کننده عملکرد مطلوب روش مورد استفاده به‌منظور برآورد مقادیر منواکسید کربن و دی‌اکسید نیتروژن است. عملکرد شبکه عصبی مصنوعی تا حدود زیادی به معماری این شبکه شامل تعداد لایه‌ها، تعداد نورون‌ها و توابع فعال‌سازی مورد استفاده وابسته است (فیفر^۱ و همکاران، ۲۰۰۹) که در پژوهش حاضر با استفاده از روش آزمون و خطا، تلاش شد تا مقادیر بهینه برای آن‌ها تعیین شود. توانایی پیش‌بینی مدل حاصل از شبکه عصبی مصنوعی افزون بر معماری مطلوب این شبکه، به عوامل دیگری همچون تراکم ایستگاه‌های نظارت آلودگی هوا، به‌ویژه در زمان مدل‌سازی مکانی، وابسته است (آلیمیسیس^۲ و همکاران، ۲۰۱۸).

مدل‌سازی مکانی در نوشتار پیش رو نشان داد که صحت مدل‌سازی می‌تواند برای ایستگاه‌های مختلف تا حدودی متفاوت باشد (جداول ۴ و ۵). تفاوت در عملکرد مدل‌های مکانی برای ایستگاه‌های مختلف در پژوهش‌های دیگر نیز به‌تأیید رسیده است (آلیمیسیس و همکاران، ۲۰۱۸). براساس گزارش شرکت کنترل کیفیت هوای تهران در سال ۱۳۹۴، میانگین غلظت سالانه آلاینده‌ها در ایستگاه‌های واقع در مرکز و جنوب شهر تهران بیشتر از سایر مناطق بوده است (شرکت کنترل کیفیت هوای تهران، ۱۳۹۵). این تفاوت در غلظت آلاینده‌ها می‌تواند بیانگر این مسئله باشد که عملکرد عوامل انتشار آلاینده‌ها و نیز نوع این عوامل، می‌تواند برای ایستگاه‌های مختلف، متفاوت باشد. در واقع عوامل زیادی ویژگی‌های اقلیمی ایستگاه، ارتفاع ساختمان‌های اطراف ایستگاه، وضعیت ترافیک اطراف ایستگاه، وضعیت پوشش گیاهی، وضعیت استقرار واحدهای صنعتی بر میزان آلاینده‌های یک ایستگاه تأثیرگذار هستند. به‌طور مشخص، این عوامل برای یک ایستگاه نسبت به ایستگاهی دیگر می‌تواند متفاوت باشد و به‌دنبال آن، سازوکار انتشار و توسعه آلاینده‌ها نیز در یک ایستگاه با ایستگاه‌های دیگر تا حدودی متفاوت باشد. خطای زیاد تخمین آلاینده‌ها در تعدادی از ایستگاه‌ها، می‌تواند ناشی از تفاوت در شرایط انتشار آلاینده‌ها، تحولات شیمیایی آلاینده‌ها، منابع تولیدکننده و همچنین متغیرهای مورد استفاده برای تخمین آلاینده‌ها در آن منطقه نسبت به سایر مناطق باشد (آگیره-باسارکو و همکاران، ۲۰۰۶).

همان‌گونه که در نتایج نوشتار پیش رو نیز بیان شد، نتایج حاصل از همبستگی اسپیرمن نشان داد که رابطه متغیرهای مختلف مورد استفاده در پژوهش حاضر با مقادیر بیشینه غلظت آلاینده‌ها در ایستگاه‌های مختلف سطح شهر تهران متفاوت است؛ بنابراین انتظار می‌رود تا مدل مکانی که با استفاده از چند ایستگاه توسعه یافته است، عملکرد مناسبی برای تخمین بیشینه غلظت آلاینده در یک ایستگاه داشته باشد؛ به شرط آنکه شرایط مختلف تولید و انتشار آلاینده برای آن ایستگاه با ایستگاه‌های مورد استفاده برای توسعه مدل مشابه باشد؛ همچنین صحت روش‌های مدل‌سازی به صحت، تعداد و تراکم ایستگاه‌های پایش آلودگی هوا بستگی دارد (تامزاک^۳، ۱۹۹۸) که خود می‌تواند دلیل دیگری بر عملکرد متفاوت مدل‌های مکانی در ایستگاه‌های مختلف باشد؛ افزون بر این، با توجه به اثرات پیچیده توپوگرافی، سازوکارهای حمل‌ونقل، واکنش‌های شیمیایی، منابع آلودگی هوا و انتشار گازها عملکرد مدل می‌تواند تحت تأثیر شرایط و ویژگی‌های ایستگاه مورد نظر برای انجام تخمین و نیز نوع آلاینده باشد (آلیمیسیس و همکاران، ۲۰۱۸). تفاوت در عملکرد شبکه عصبی مصنوعی در رابطه با آلاینده‌های مختلف و نیز ایستگاه‌های مختلف در پژوهش حاضر نیز رخ داد.

1- Pfeiffer
2- Alimissis
3- Tomczak

نتیجه‌گیری

در نوشتار پیش رو با به کارگیری یک مدل شبکه عصبی پرسپترون چندلایه و استفاده از خصیصه‌های محیطی و سنجش‌ازدوری به مدل‌سازی مقادیر بیشینه روزانه آلاینده‌های منواکسید کربن و دی‌اکسید نیتروژن به صورت زمانی و مکانی پرداخته شد. مهم‌ترین نتایج نوشتار پیش رو را می‌توان به صورت زیر بیان کرد:

- نتایج حاصل از مدل‌سازی زمانی و مکانی آلاینده دی‌اکسید نیتروژن به دلیل پایداری بیشتر این آلاینده در جو، از صحت کمابیش بهتری نسبت به آلاینده منواکسید کربن برخوردار است؛ بنابراین مدل حاصل از شبکه عصبی مصنوعی پرسپترون چندلایه برای آلاینده‌های پایدارتر در هوا صحت پیش‌بینی بهتری را نسبت به آلاینده‌های ناپایدار ارائه می‌دهد.

- شبکه عصبی مصنوعی در مدل‌سازی زمانی نسبت به مدل‌سازی مکانی، صحت کمابیش بهتری ارائه می‌کند.

- معماری شبکه عصبی مصنوعی در مدل‌سازی مکانی نسبت به مدل‌سازی زمانی می‌تواند متفاوت باشد.

- نتایج حاصل از مدل‌سازی مکانی آلاینده‌ها نشان داد که عملکرد مدل‌های مکانی حاصل از شبکه عصبی مصنوعی می‌تواند برای ایستگاه‌های مختلف متفاوت باشد.

نتایج پژوهش و ارزیابی آماری نشان داد که استفاده از مدل‌سازی شبکه عصبی مصنوعی در کنار استفاده از تبدیل موجک به منظور پیش‌پردازش داده‌های غلظت آلاینده‌ها و نیز روش انتخاب ویژگی جنگل تصادفی می‌تواند در پیش‌بینی مقادیر آلاینده‌ها به صورت مکانی و زمانی نتایج خوبی را به همراه داشته باشد.

در پایان با توجه به وجود برخی مسائل همچون نقص فنی دستگاه‌های ثبت‌کننده آلاینده هوا در یک ایستگاه که موجب عدم ثبت غلظت آلاینده می‌شود، پیشنهاد می‌شود که از روش ارائه‌شده در پژوهش پیش رو برای رفع نقص داده‌های موجود استفاده شود.

منابع

آل‌شیخ، علی‌اصغر؛ قراگوزلو، علیرضا؛ سجادیان، مهیار (۱۳۹۱). بهره‌گیری از شبکه عصبی به منظور استفاده در فرایند مدیریت ریسک زیست‌محیطی ناشی از آلودگی هوای منتج از ترافیک در کلان‌شهر تهران. *جغرافیایی چشم‌انداز زاگرس*، ۴ (۱۴)، ۲۵-۳۸.

اجتهادی، مرجان (۱۳۸۶). بررسی آلودگی هوای شهری ناشی از سامانه حمل‌ونقل با تأکید بر ذرات معلق و ارائه راهکارهای مدیریتی (مطالعه موردی، تهران). *دهمین همایش ملی بهداشت محیط، همدان: دانشگاه علوم پزشکی همدان*.
حسینی شفیق، روجا؛ علیمحمدی، عباس؛ قاسمیان یزدی، محمدحسن (۱۳۹۵). ارزیابی آثار موجک پایه و تعداد سطوح تجزیه جهت تخمین نقشه تغییرات با استفاده از الگوریتم موجک. *سنجش‌ازدور و GIS/یران*، ۱ (۲)، ۱۷-۳۴.
خبری، زهرا؛ موسویان ندوشن، نرجس السادات؛ نژاد کورکی، فرهاد؛ منصوری، نبی‌ا... (۱۳۹۲). تأثیر مدل رقومی ارتفاعی در مدل‌سازی آلودگی هوا با استفاده از اثرمود (AERMOD). *سنجش‌ازدور و سامانه اطلاعات جغرافیایی در منابع طبیعی*، ۴ (۴)، ۲۵-۳۳.

رفیع‌پور، مهرداد؛ آل‌شیخ، علی‌اصغر؛ علیمحمدی سراب، عباس؛ صادقی نیارکی، ابوالقاسم (۱۳۹۲). مدل‌سازی مکانی غلظت منواکسید کربن در تهران با استفاده از رگرسیون چندمتغیره و شبکه‌های عصبی. *همایش ملی ژئوماتیک*، ۲۰. تهران: دانشگاه آزاد اسلامی.

شرعی‌پور، زهرا (۱۳۸۸). بررسی تغییرات فصلی و روزانه آلاینده‌های هوا و ارتباط آن با پارامترهای هواشناسی. *مجله فیزیک زمین و فضا*، ۲ (۲)، ۱۱۹-۱۳۷.

شرکت کنترل کیفیت هوای تهران (۱۳۹۵). گزارش سالانه کیفیت هوای تهران در سال ۱۳۹۴. تهران: شهرداری تهران.

عبودی، محمدرضا؛ کرمی، جلال؛ شمس‌الدینی، علی (۱۳۹۴). مدل‌سازی خطی و غیر خطی آلاینده‌های هوای شهر تهران با استفاده از خصیصه‌های محیطی و ترافیک. *اولین کنفرانس ملی مهندسی فناوری اطلاعات مکانی*. تهران: دانشگاه صنعتی خواجه نصیر طوسی.

متکان، علی‌اکبر؛ شکیبا، علیرضا؛ پورعلی، سید حسین؛ بهارلو، ایمان (۱۳۸۸). تعیین تغییرات مکانی و زمانی آلودگی‌های منواکسید کربن و ذرات معلق با استفاده از تکنیک‌های GIS در شهر تهران. *سنجش‌ازدور و GIS/ایران*، ۱(۱)، ۵۷-۷۲.

References

- Aboodi, M., Karami, J. & Shamsoddini, A. (2015). Linear and nonlinear modeling of air pollutants in Tehran using environmental and traffic characteristics. In: *First National Conference on Spatial Information Technology Engineering*. Tehran: K. N. Toosi University of Technology. (In Persian)
- Agirre-Basurko, E., Ibarra-Berastegi, G. & Madariaga, I. (2006). Regression and multilayer perceptron-based models to forecast hourly O3 and NO2 levels in the Bilbao area. *Environmental Modelling & Software*, 21 (4), 430-446.
- Alesheikh, A., Qara Gozlu, A. & Sajjadian, M. (2012). Use of neural network for use in environmental risk management process due to air pollution caused by traffic in Tehran metropolis. *Zagros Landscape Geography and Planning*, 4 (14), 25-38. (In Persian)
- Alimissis, A., Philippopoulos, K., Tzani, C. G. & Deligiorgi, D. (2018). Spatial estimation of urban air pollution with the use of artificial neural network models. *Atmospheric environment*, 191, 205-213.
- Artis, D. A. & Carnahan, W. H. (1982). Survey of emissivity variability in thermography of urban areas. *Remote Sensing of Environment*, 12 (4), 313-329.
- Chen, G., Knibbs, L. D., Zhang, W., Li, S., Cao, W., Guo, J., ... & Hamm, N. A. S. (2018). Estimating spatiotemporal distribution of PM1 concentrations in China with satellite remote sensing, meteorology, and land use information. *Environmental pollution*, 233, 1086-1094.
- Chelani, A. B., Rao, C. C., Phadke, K. M. & Hasan, M. Z. (2002). Prediction of sulphur dioxide concentration using artificial neural networks. *Environmental Modelling & Software*, 17 (2), 159-166.
- Coulibaly, P., Dibike, Y. B. & Anctil, F. (2005). Downscaling precipitation and temperature with temporal neural networks. *Journal of Hydrometeorology*, 6 (4), 483-496.
- Ejtehadi, M. (2007). Investigation of urban air pollution caused by transportation system with emphasis on suspended particles and presentation of management solutions (Case study, Tehran). In: *10th National Conference on Environmental Health*. Hamadan: Hamadan University of Medical Sciences. (In Persian)
- Elangasinghe, M. A., Singhal, N., Dirks, K. N. & Salmond, J. A. (2014). Development of an ANN-based air pollution forecasting system with explicit knowledge through sensitivity analysis. *Atmospheric Pollution Research*, 5 (4), 696-708.
- Gardner, M. W. & Dorling, S. R. (1999). Neural network modelling and prediction of hourly NOx and NO2 concentrations in urban air in London. *Atmospheric Environment*, 33 (5), 709-719.
- Grgurić, S., Krizan, J., Gasparac, G., Antonić, O., Spiric, Z.; Mamouri, R. E., Christodoulou, A.; Nisantzi, A., Agapiou, A., Themistocleous, K., Fedra, K., Panayiotou, C. & Hadjimitsis, D. (2014). Relationship between MODIS based Aerosol Optical Depth and PM10 over Croatia. *Central European Journal of Geosciences*. 6 (1), 2-16.
- Hewitt, C. N. (1991). Spatial variations in nitrogen dioxide concentrations in an urban area. *Atmospheric Environment. Part B. Urban Atmosphere*, 25 (3), 429-434.
- Hosseini Shafi, R., Ali Mohammadi, A. & Qasemian Yazdi, M. (2016). Evaluation of base wavelet effects and number of decomposition levels to estimate change map using wavelet algorithm. *Iranian Remote sensing & GIS*, 8 (2), 17-34. (In Persian)
- Hrustić, L., Klaić, Z., B. Križan, J., Antonić, O. & Hercog, P. (2009). Neural network forecasting of air pollutants hourly concentrations using optimised temporal averages of meteorological

- variables and pollutant concentrations. *Atmospheric Environment*, 43 (35), 5588-5596.
- Karatzas, K. D. & Kaltsatos, S. (2007). Air pollution modelling with the aid of computational intelligence methods in Thessaloniki, Greece. *Simulation Modelling Practice and Theory*, 15 (10), 1310-1319.
- Khabari, Z., Mousavian Nodoshan, N., Nejad Korki, F. & Mansoori, N. (2013). The effect of digital elevation model on air pollution modeling using AERMOD. *Remote sensing and GIS in natural resources*, 4 (4), 25-33. (In Persian)
- Kharytonov, M. M., Khlopova, V. M., Stankevich, S. A. & Titarenko, O. V. (2013). Remote and ground-based sensing of air polluted by nitrogen dioxide in the Dnepropetrovsk region (Ukraine). In *Disposal of Dangerous Chemicals in Urban Areas and Mega Cities* (pp. 291-298). Springer, Dordrecht.
- Knelson, J. H. & Lee, R. E. (1977). Oxides of nitrogen in the atmosphere: origin, fate and public health implications. *Ambio*, 6 (2/3), 126-130.
- Koller, D. & Sahami, M. (1996). Toward optimal feature selection. *Stanford InfoLab*.
- Kukkonen, J., Partanen, L., Karppinen, A., Ruuskanen, J., Junninen, H., Kolehmainen, M., ... & Cawley, G. (2003). Extensive evaluation of neural network models for the prediction of NO₂ and PM₁₀ concentrations, compared with a deterministic modelling system and measurements in central Helsinki. *Atmospheric Environment*, 37 (32), 4539-4550.
- Mao, L., Qiu, Y., Kusano, C. & Xu, X. (2012). Predicting regional space-time variation of PM 2.5 with land-use regression model and MODIS data. *Environmental Science and Pollution Research*, 19 (1), 128-138.
- Martin, R. V. (2008). Satellite remote sensing of surface air quality. *Atmospheric environment*, 42 (34), 7823-7843.
- Matkan, A., Shakiba, A., Pour Ali, S. & Baharlou, I. (2009). Determination of spatial and temporal variations of carbon monoxide and particulate pollutants using GIS techniques in Tehran. *Iranian Remote sensing & GIS*, 1 (1), 57-72. (In Persian)
- McKendry, I. G. (2002). Evaluation of artificial neural networks for fine particulate pollution (PM₁₀ and PM_{2.5}) forecasting. *Journal of the Air & Waste Management Association*, 52 (9), 1096-1101.
- Montero-Lorenzo, J. M., Fernández-Avilés, G., Mondéjar-Jiménez, J. & Vargas-Vargas, M. (2013). A spatio-temporal geostatistical approach to predicting pollution levels: The case of mono-nitrogen oxides in Madrid. *Computers, Environment and Urban Systems*, 37, 95-106.
- Nisantzi, A., Hadjimitsis, D. G., Akylas, E., Agapiou, A., Panayiotou, M., Michaelides, S., & Paronis, D. (2013). Study of air pollution with the use of modis data, lidar and sun photometers in Cyprus. In *Advances in Meteorology, Climatology and Atmospheric Physics* (pp. 1133-1139). Springer, Berlin, Heidelberg.
- Perez, P. & Trier, A. (2001). Prediction of NO and NO₂ concentrations near a street with heavy traffic in Santiago, Chile. *Atmospheric Environment*, 35 (10), 1783-1789.
- Pfeiffer, H., Baumbach, G., Sarachaga-Ruiz, L., Kleanthous, S., Poulida, O. & Beyaz, E. (2009). Neural modelling of the spatial distribution of air pollutants. *Atmospheric Environment*, 43 (20), 3289-3297.
- Rafiepour, M., Alesheikh, A., Ali Mohammadi Sarab, A. & Sadeghi Niaraki, A. (2013). Spatial modeling of carbon monoxide concentration in Tehran using multivariate regression and neural networks. In: *National Geomatics Conference*. Tehran: Islamic Azad University. (In Persian)
- Shamsoddini, A., Aboodi, M. R. & Karami, J. (2017). Tehran air pollutants prediction based on random forest feature selection method. *International Archives of the Photogrammetry, Remote Sensing & Spatial Information Sciences*, 483-488.
- Shamsoddini, A., Raval, S. & Taplin, R. (2014). Spectroscopic analysis of soil metal contamination around a derelict mine site in the Blue Mountains, Australia. *ISPRS Annals of Photogrammetry, Remote Sensing & Spatial Information Sciences*, 75-79.
- Sharie pour, Z. (2009). Investigation of seasonal and daily changes of air pollutants and its relationship with meteorological parameters. *Journal of Earth and Space Physics*, 35 (2), 119-

137. (In Persian)
- Shi, J. P. & Harrison, R. M. (1997). Regression modelling of hourly NO_x and NO₂ concentrations in urban air in London. *Atmospheric Environment*, 31 (24), 4081-4094.
- Snyder, W. C., Wan, Z., Zhang, Y. & Feng, Y. Z. (1998). Classification-based emissivity for land surface temperature measurement from space. *International Journal of Remote Sensing*, 19 (14), 2753-2774.
- Siwek, K. & Osowski, S. (2012). Improving the accuracy of prediction of PM₁₀ pollution by the wavelet transformation and an ensemble of neural predictors. *Engineering Applications of Artificial Intelligence*, 25 (6), 1246-1258.
- Tehran Air Quality Control Company (2016). *Annual Report of Tehran Air Quality In 1394*. Tehran: Municipality of Tehran. (In Persian)
- Tomczak, M. (1998). Spatial interpolation and its uncertainty using automated anisotropic inverse distance weighting (IDW)-cross-validation/jackknife approach. *Journal of Geographic Information and Decision Analysis*, 2 (2), 18-30.
- Zhang, H., Wu, W. & Yao, M. (2012 a). Boundedness and convergence of batch back-propagation algorithm with penalty for feedforward neural networks. *Neurocomputing*, 89, 141-146.
- Zhang, Y., Yiyun, C., Qing, D. & Jiang, P. (2012 b). Study on urban heat island effect based on Normalized Difference Vegetated Index: A case study of Wuhan City. *Procedia environmental sciences*, 13, 574-581.