



Available online: <http://ijhe.tums.ac.ir>

مقاله پژوهشی

## تخمین غلظت ازن در سطح زمین با استفاده از روش رگرسیون خطی چند متغیره و تعیین پارامترهای موثر

کیوان عزى مند، عطاءاله عبداله‌ی کاکرودی\*

گروه سنجش از دور و سیستم اطلاعات جغرافیایی، دانشکده جغرافیا، دانشگاه تهران، تهران، ایران

### چکیده

### اطلاعات مقاله:

زمینه و هدف: ازن سطح زمین ( $O_3$ ) به‌عنوان یکی آلاینده‌های بسیار خطرناک دارای اثرات قابل توجهی برای سلامت ساکنان مناطق شهری است. هدف از این مطالعه شناسایی عوامل مؤثر بر تشکیل ازن و مدلسازی تغییرات مکانی و زمانی غلظت ازن در کلان شهر تهران است. روش بررسی: داده‌های مورد استفاده در این تحقیق شامل داده‌های هواشناسی و داده‌های مربوط به غلظت آلاینده‌ها برای سال ۱۳۹۳ است. در این پژوهش ابتدا تاثیر و همبستگی پارامترها با غلظت ازن با استفاده از ضریب همبستگی پیرسون (coefficient of Pearson) و سپس مدلسازی غلظت ازن، با روش رگرسیون خطی چند متغیره (multivariate linear regression) انجام شد. یافته‌ها: مدل توسعه داده شده توانایی توصیف ۷۹ درصد تغییرات داده برای سال ۱۳۹۳ را داشته است. آنالیز زمانی غلظت ازن نشان می‌دهد که بهترین عملکرد مدل برای فصل تابستان ( $R^2=0.771$ ) و تیر ماه ( $R^2=0.778$ ) بوده است. همچنین نتایج مدلسازی مکانی برای ایستگاه‌های سنجش آلودگی نشان می‌دهد که ایستگاه شهرداری منطقه ۴ دارای کمترین ضریب تشخیص مکانی ( $R^2=0.16$ ) و ایستگاه اقدسیه دارای بالاترین ضریب تشخیص ( $R^2=0.79$ ) هستند. در نهایت توزیع مکانی مقدار غلظت ازن تخمین زده شده با مقدار غلظت ازن اندازه گیری در سطح ایستگاه همخوانی کاملی داشت. نتیجه‌گیری: براساس نتایج این مطالعه، تمامی پارامترهای مربوط به غلظت آلاینده‌ها همراه با پارامترهای هواشناسی از عوامل مؤثر در مدلسازی غلظت ازن سطح زمین بوده‌اند. همچنین پراکنش مکانی غلظت ازن در شهر تهران نشان از غلظت بیشتر ازن در جنوب و شرق تهران نسبت به شمال و غرب تهران بوده است.

۹۶/۰۸/۲۰

تاریخ دریافت:

۹۶/۱۱/۰۸

تاریخ ویرایش:

۹۶/۱۱/۱۵

تاریخ پذیرش:

۹۶/۱۲/۱۶

تاریخ انتشار:

واژگان کلیدی: غلظت آلاینده‌ها، پارامترهای هواشناسی، تخمین غلظت ازن، رگرسیون خطی چند متغیره

پست الکترونیکی نویسنده مسئول:

a.a.kakroodi@ut.ac.ir

## مقدمه

افزایش غلظت آلاینده‌های جوی و ارتباط آن با سلامتی انسان‌ها در کلان‌شهرها و کشورهای در حال توسعه در سال‌های اخیر به مسئله‌ای مهم تبدیل شده است (۱). در میان تمامی آلاینده‌های جوی، غلظت آلاینده ازن سطح زمین که در نتیجه فرایند فتوشیمیایی تولید می‌شود، در طول قرن اخیر دو برابر شده است (۲). قرار گرفتن در معرض آن برای سلامتی انسان بسیار مضر بوده و به ریه‌ها بشدت آسیب می‌رساند و برای افرادی که در فضای باز ورزش می‌کنند خطرناک است (۳). همچنین این گاز باعث کاهش تولید محصولات (۴)، اثرات گسترده بر اکوسیستم‌های زمینی (۵-۷) و در نهایت سبب شدت یافتن اثر گلخانه‌ای و افزایش دمای کره زمین می‌شود. مقدار غلظت ازن دارای تغییرات زمانی و مکانی است (۸) و در نتیجه تغییر عرض‌های جغرافیایی، ارتفاع، و میزان فعالیت‌های انسانی متغیر است (۹). آلاینده فتوشیمی ازن سطح زمین به دلیل انتشار گاز نیتریک اکساید (NO)، تغییرات پارامترهای اقلیمی (۱۰) و در شهرهای بزرگ که در مسیر باد هستند، تغییرات چرخه‌ای نامنظم دارند (۱۱)؛ در نتیجه تعیین غلظت آلاینده ازن و مدلسازی تغییرات آن در نقاط و فصول مختلف جهت ارائه راهکارهای کنترلی و مدیریتی امری ضروری است. محققان همواره سعی کرده‌اند که برای تخمین مقدار غلظت ازن از مدل‌های ساده خطی تا مدل‌های غیرخطی پیچیده استفاده نمایند. در برخی از این مطالعات ابتدایی، از مدل رگرسیون خطی (۱۲) استفاده کرده‌اند. جهت بهبود فرایند مدلسازی و تخمین غلظت ازن از مدل‌های پیچیده همچون شبکه عصبی (۱۳) و رگرسیون خطی چند متغیره با بهره‌گیری از پارامترهای NO، NO<sub>x</sub> (۱۴)، پارامترهای اقلیمی (۱۰) و تاثیر دمای سطح (۱۵) استفاده شده است. Abdul-Wahab و همکاران در سال ۲۰۰۵ از مدل رگرسیون خطی چند متغیره و تحلیل مؤلفه‌های اصلی (Principal component analysis) و به کارگیری پارامترهای مربوط به آلاینده‌های جوی و پارامترهای هواشناسی به پیش‌بینی مقدار غلظت ازن در منطقه مسکونی Khaldiya پرداخته‌اند و به این نتیجه رسیده‌اند

که پارامترهای NO، SO<sub>2</sub> و دما بیشترین تاثیر بر افزایش غلظت ازن داشته‌اند (۱۶). Sousa و همکاران در سال ۲۰۰۷ با استفاده از شبکه عصبی و رگرسیون خطی چند متغیره به پیش‌بینی مقدار غلظت ازن برای شهر Oporto با استفاده از پارامترهای رطوبت نسبی، سرعت باد، دما، NO و NO<sub>2</sub> پرداخته‌اند و به این نتیجه رسیده‌اند که هر دو مدل در نتیجه ورود تمامی این پارامترها به خوبی توانسته‌اند غلظت ازن را پیش‌بینی کنند (۱۷). همچنین در رابطه با پیش‌بینی مقدار غلظت سایر آلاینده‌های جوی با استفاده از رگرسیون خطی چند متغیره (۱۸، ۱۹) در شهر تهران نیز مطالعاتی انجام گرفته است.

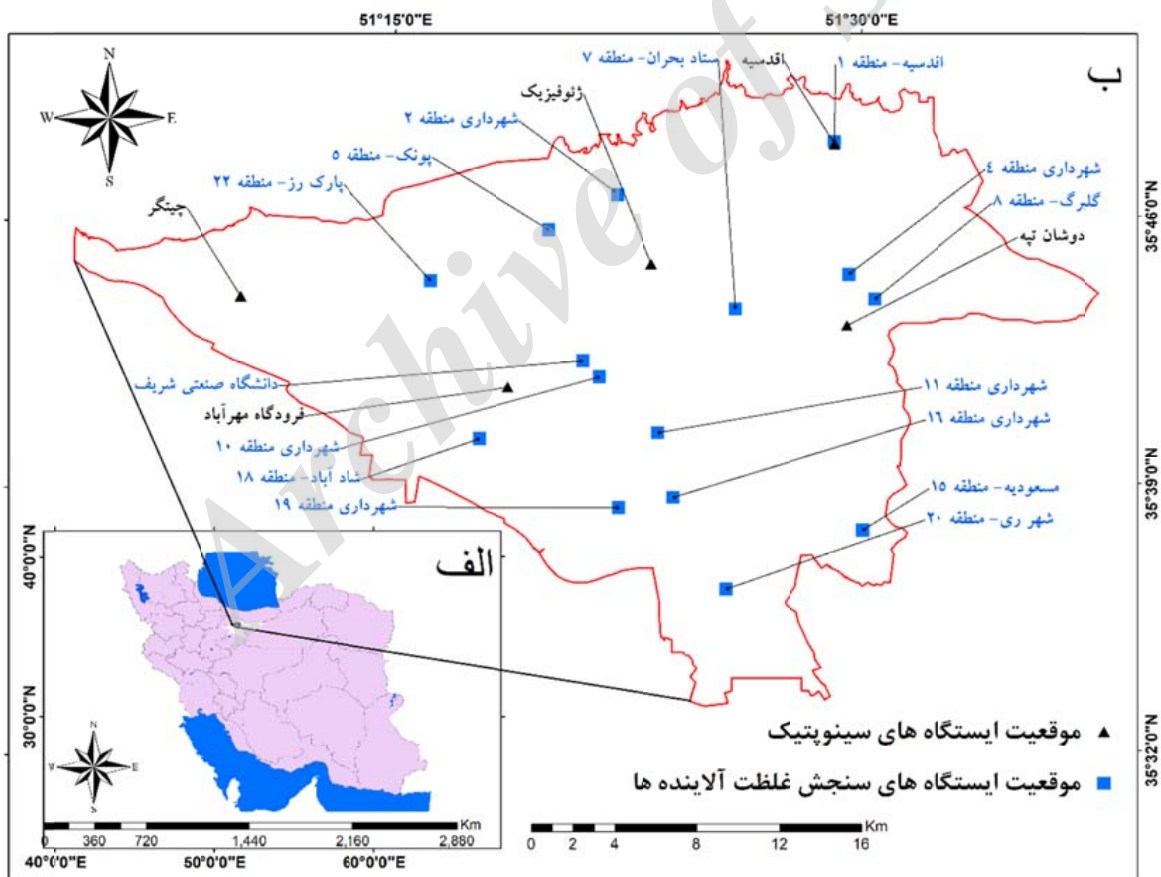
تهران یکی از آلوده‌ترین شهرهای جهان است، جمعیت زیاد، شبکه متراکم شهری، طوفان‌های گردوغبار از عوامل این آلودگی هستند (۱۸). همچنین توپوگرافی خاص شهر تهران که با کوه‌های اطراف احاطه شده است مشکل آلودگی هوای شهر تهران را تشدید می‌کند. بنابراین پیش‌بینی و تخمین غلظت ازن با توجه به افزایش غلظت آن در دهه‌های اخیر (۲۰) شهر تهران می‌تواند نقش عمده‌ای در مدیریت بهینه آن داشته باشد. مروری بر ادبیات تحقیق نشان داد که مطالعات متفاوتی در رابطه با مدلسازی مقدار غلظت آلاینده‌های جوی از جمله ازن در شهر تهران صورت گرفته است ولی مطالعه‌ای جامع به منظور پیش‌بینی غلظت ازن براساس کل ایستگاه‌های سنجش غلظت آلاینده شهر تهران و همچنین در نظر گرفتن پارامترهای هواشناسی در کنار پارامترهای غلظت آلاینده‌ها به صورت هم‌زمان انجام نگرفته است. بر این اساس و با توجه به اثرات سوء غلظت ازن بر سلامتی انسان و محیط زیست هدف این مطالعه تخمین غلظت ازن با استفاده از پارامترهای هواشناسی و غلظت آلاینده‌های جوی در شهر تهران و نشان دادن تغییرات مکانی و زمانی آن با استفاده از مدل رگرسیون خطی چند متغیره است. به این منظور، برای رسیدن به این هدف از داده‌های ایستگاه‌های سینوپتک و داده‌های ایستگاه‌های سنجش آلاینده‌های جوی سال ۱۳۹۳ مربوط به شهر تهران استفاده شده است.

## مواد و روش‌ها

– منطقه مورد مطالعه و داده‌های مورد استفاده

مجموعه داده‌هایی که برای این مطالعه مورد استفاده قرار می‌گیرند مربوط به سال ۱۳۹۳ برای کلان شهر تهران است که شامل پارامترهای رطوبت نسبی بر حسب درصد، دما بر حسب °C و سرعت باد بر حسب m/s از ایستگاه‌های سینوپتیک مربوط به سازمان آب و هواشناسی کل کشور (شکل ۱) و غلظت آلاینده‌های دی اکسید نیتروژن (NO<sub>2</sub>) بر حسب ppb، نیتریک اکساید (NO) بر حسب ppb، مونو اکسید کربن (CO) بر حسب ppm، دی اکسید گوگرد (SO<sub>2</sub>) بر حسب ppb و ذرات معلق μ ۲/۵ (PM<sub>2.5</sub>) بر حسب μg/m<sup>3</sup> مربوط به شرکت کنترل کیفیت هوای شهر تهران (شکل ۱) است.

در این پژوهش ابتدا داده‌ها برای سال ۱۳۹۳ آماده‌سازی گردید و در مرحله بعد به منظور تعیین میزان و نوع ارتباط هر یک از پارامترها با غلظت ازن از ضریب همبستگی پیرسون (coefficient of Pearson) استفاده شده است. مدل رگرسیون خطی چند متغیره (multivariate linear regression) با استفاده از ۷۰ درصد داده‌ها برای سال ۱۳۹۳ توسعه داده شده و با استفاده از ۳۰ درصد داده‌ها میزان ضریب تشخیص (Coefficient of determination) و جذر میانگین مربعات خطا (Root Mean Square Error) برای فصول و ماه‌های مختلف مورد بررسی قرار گرفته است و در نهایت تغییرات مکانی غلظت ازن در شهر تهران نمایش داده شده است.



شکل ۱- محدوده مورد مطالعه (الف) موقعیت شهر تهران نسبت به کشور ایران؛ (ب) محدوده شهر به همراه پراکنش ایستگاه‌های سنجش غلظت آلاینده‌ها و ایستگاه‌های سینوپتیک در شهر تهران

در معادله فوق  $h\nu$  اشعه خورشید است. رابطه بین ازن و دما نیز از طریق اشعه فرابنفش خورشید و ارتباط آن نیز با رطوبت نسبی بدین صورت است که در شرایط رطوبت نسبی بالا و یا هوای بارانی مقدار ازن کاهش پیدا می‌کند. علاوه بر این باد نیز به‌عنوان یک عامل مهم در پراکندگی سایر آلودگی‌های جو است. می‌توان گفت که رابطه عکس بین سرعت باد و رطوبت نسبی با غلظت آلاینده ازن وجود دارد (۲۳).  $PM_{2.5}$  و  $SO_2$  که از دیگر آلاینده‌های هوا هستند، همانند گازهای اولیه تشکیل ازن از پیش‌نیازهای تولید این گاز هستند (۲۶)، و در مطالعات بیان شده است که این پارامترها با میزان غلظت ازن رابطه معکوس دارند (۲۳).

جهت تعیین همبستگی میان متغیرهای مستقل و وابسته مورد استفاده از ضریب پیرسون استفاده شده است. ضریب پیرسون با معادله ۲ محاسبه می‌شود (۲۷).

$$R = \frac{\sum(X-\bar{X})(Y-\bar{Y})}{\sqrt{\sum(X-\bar{X})^2 \sum(Y-\bar{Y})^2}} \quad (2)$$

$X$  و  $Y$  مقادیر دو متغیر مورد نظر،  $\bar{X}$  و  $\bar{Y}$  میانگین مقادیر دو متغیر است.

- فرایند مدل‌سازی

رگرسیون خطی چند متغیره از تکنیک‌های رایج در مطالعات آلودگی هوا است. این تکنیک در این مطالعه به دلیل قابلیت‌های آن برای پیش‌بینی مشارکت پارامترهای انتخاب شده در تغییرات غلظت ازن مورد استفاده قرار گرفت (۱۶). رگرسیون خطی چند متغیره با استفاده از روش کمترین مربعات مدل‌سازی انجام می‌دهد، که اختلاف مربع بین غلظت ازن مشاهده شده و غلظت ازن پیش‌بینی شده به حداقل برسد. معادله کلی رگرسیون خطی چند متغیره به‌صورت معادله ۳ بیان می‌شود (۲۸).

$$y = b_0 + \sum_{i=1}^p b_i x_i + \varepsilon \quad (3)$$

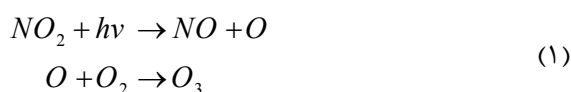
در معادله فوق  $b_0$  عرض از مبدا،  $b_i$  ضریب متغیرهای مستقل،  $x_i$

داده‌های مربوط به ایستگاه‌های سینوپتیک هر سه ساعت یکبار برداشت می‌شوند به همین منظور داده‌های مربوط به غلظت آلاینده‌ها نیز با میانگین سه ساعته تهیه شده و در مدل‌سازی مورد استفاده قرار گرفته‌اند. پس از آماده‌سازی داده‌های غلظت آلاینده‌ها براساس داده‌های هواشناسی، درنهایت داده‌های ایستگاه سینوپتیک با ایستگاه‌های سنجش آلودگی ادغام شده‌اند. به‌منظور کنترل کیفیت داده‌ها و حذف داده‌های مزاحم مربوط به برداشت‌های نادرست، پس از محاسبه انحراف معیار داده‌ها، تمامی داده‌هایی که انحراف معیار بین اعداد  $-35$  تا  $+35$  را داشته‌اند در مدل‌سازی مدنظر قرار گرفته‌اند و سایر داده‌های مزاحم پس از بررسی دوباره که کمتر از ۱ درصد کل داده‌ها بودند در فرایند مدل‌سازی مورد استفاده قرار نگرفتند. تعداد کل داده‌های مورد استفاده در این تحقیق برای ایستگاه‌های هواشناسی ۱۷۰۵ داده و برای ایستگاه‌های سنجش غلظت آلاینده‌ها ۳۳۲۸ داده بوده است.

- توسعه مدل

ارتباط بین غلظت ازن با سایر آلاینده‌ها و داده‌های هواشناسی

همان‌طور که در قبل هم اشاره شد مطالعات زیادی به بررسی ارتباط بین پارامترهای هواشناسی و پارامترهای آلودگی با غلظت ازن انجام گرفته است (۲۱). گازهای  $NO$ ،  $NO_2$  و  $CO$  از گازهای اولیه در فرایند تولید ازن هستند (۲۲). Sullivan و همکاران با بررسی سری زمانی سه آلاینده فوق نشان دادند تغییرات غلظت این آلاینده‌ها با آلاینده ازن دارای ارتباط غیرمستقیم است (۲۳). اگرچه اشعه فرابنفش B (Ultraviolet B) خورشید نقش مهمی در فرایند تشکیل ازن دارد اما بدون گازهای فوق این فرایند تکمیل نمی‌شود (۲۴). در نتیجه تابش اشعه خورشید با گازهای  $NO_2$ ، این گازها شکسته شده و سبب آزاد شدن اتم‌های  $O$  می‌شود که با مولکول‌های  $O_2$  واکنش نشان می‌دهد و گاز ازن تشکیل می‌گردد. فرایند تشکیل این گاز در معادله ۱ بیان شده است (۲۵).



در معادله فوق  $n$  حجم نمونه،  $p_i$  غلظت ازن پیش‌بینی شده،  $O_i$  غلظت ازن مشاهده شده است.  $\bar{p}$  میانگین غلظت ازن پیش‌بینی،  $\bar{O}$  میانگین غلظت ازن مشاهده شده،  $S_{pred}$  انحراف معیار غلظت ازن پیش‌بینی و  $S_{obs}$  انحراف استاندارد غلظت ازن بین مقادیر ورودی و خروجی است. معادله جذر میانگین مربعات خطا نیز به صورت معادله ۶ است (۳۰).

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (o_i - p_i)^2} \quad (۶)$$

در معادله فوق  $n$  تعداد دفعات مشاهده،  $O_i$  مقدار غلظت ازن اندازه‌گیری شده و  $p_i$  مقدار غلظت ازن مدلسازی شده است.

### یافته‌ها

#### - توصیف آماری داده‌ها

توصیف آماری داده‌های مورد استفاده در این تحقیق در جدول ۱ نشان داده شده است. مقادیر غلظت ازن در تمامی ایستگاه‌ها با میانگین ۳۳/۳۱ برحسب ppb از ۱ تا ۱۲۵ متغیر بوده است که بیانگر مقدار بالای این آلاینده است. همچنین بررسی تغییرات سایر آلاینده‌های جوی و همچنین میانگین مربوط به این آلاینده‌ها حاکی از غلظت بالای متغیرهای فوق است. متوسط رطوبت نسبی با میانگین ۳۴/۰۴ درصد نشان‌دهنده یک شرایط جوی خشک و مجاورت آن با مناطق کویری است (۱۹). مقادیر دما با بازه تغییرات  $7^{\circ}\text{C}$  تا  $36/3^{\circ}\text{C}$  و مقادیر سرعت باد با بازه تغییرات  $0\text{ m/s}$  تا  $12/25\text{ m/s}$  نشان‌دهنده تغییرات این دو پارامتر در طی فصول مختلف است.

نتایج همبستگی میان پارامتر وابسته و متغیرهای مستقل مورد استفاده نشان می‌دهد (جدول ۲) که همانند مطالعات پیشین (۳۱، ۳۲) بین میزان غلظت ازن و پارامترهای  $\text{CO}$ ،  $\text{SO}_2$ ،  $\text{NO}$ ،  $\text{NO}_2$ ،  $\text{PM}_{2.5}$ ، با اینکه از گازهای اولیه تولید ازن هستند همبستگی منفی وجود داشته است که دلیل آن ناشی از تمایل بیشتر اتم اکسیژن به ترکیب با گازهای فوق نسبت به عنصر اکسیژن است. همچنین بین میزان غلظت ازن با رطوبت نسبی و سرعت باد همبستگی منفی وجود دارد که می‌تواند ناشی از پخش نمودن آلاینده‌ها به وسیله باد و کاهش

متغیر مستقل و  $\varepsilon$  خطای تصادفی مرتبط با رگرسیون است. مدل ارائه شده برای مدلسازی مقدار غلظت ازن با استفاده از سایر آلاینده‌ها و پارامترهای هواشناسی به صورت زیر است (معادله ۴).

(۴)

$$o_{ij} = b_0 + (b_1 \times \text{NO}_{2ij}) + (b_2 \times \text{CO}_{ij}) + (b_3 \times \text{RH}_{ij}) + (b_4 \times \text{NO}_{ij}) + (b_5 \times \text{PM}_{2.5ij}) + (b_6 \times \text{SO}_{2ij}) + (b_7 \times \text{WS}_{ij}) + (b_8 \times \text{Temp}_{ij}) + \varepsilon_{ij}$$

در مدل فوق  $O_{ij}$  مقدار غلظت  $\text{O}_3$  در ایستگاه پایش  $i$  ام و روز  $j$  ام است.  $\text{PM}_{2.5ij}$ ،  $\text{NO}_{ij}$ ،  $\text{RH}_{ij}$ ،  $\text{CO}_{ij}$ ،  $\text{NO}_{2ij}$ ،  $\text{SO}_{2ij}$ ،  $\text{WS}_{ij}$  و  $\text{Temp}_{ij}$  به ترتیب غلظت آلاینده‌های دی‌اکسید نیتروژن، مونواکسید کربن، رطوبت نسبی، نیتریک اکساید، ذرات معلق  $2.5 \mu\text{m}$ ، دی‌اکسید گوگرد، سرعت باد و دما مربوط به ایستگاه پایش  $i$  ام و روز  $j$  ام است. مدل فوق با استفاده از نرم‌افزار آماری SPSS و مدل رگرسیون خطی چند متغیره بین متغیر وابسته  $\text{O}_3$  و متغیرهای مستقل هواشناسی و سایر آلاینده‌های جوی با استفاده از الگوریتم انتخابی گام‌به‌گام (Stepwise) توسعه داده شده است.

#### - اعتبارسنجی مدل

پس از آماده‌سازی داده‌ها با استفاده از دستور Create Random Points در محیط نرم‌افزار ArcGIS از ۷۰ درصد داده‌ها به صورت تصادفی برای مدلسازی و از ۳۰ درصد داده‌ها برای ارزیابی مدل استفاده شده است. میزان دقت مدلسازی برای تمامی ماه‌ها و همچنین سه ماهه مورد بررسی قرار گرفته است. در نهایت میزان دقت مکانی مدلسازی نیز در سطح ایستگاه بررسی شده است. روش‌های مختلفی برای ارزیابی دقت مدل‌ها وجود دارد (۲۹، ۳۰). در این مطالعه با استفاده ضریب تشخیص و جذر میانگین مربعات خطا میزان دقت مدلسازی مورد بررسی قرار گرفته است. ضریب تشخیص برای تعیین همبستگی میان مقادیر اندازه‌گیری شده و پیش‌بینی شده است (۲۹). معادله کلی ضریب تشخیص به صورت معادله ۵ است (۳۰).

$$R^2 = \frac{\sum_{i=1}^n (p_i - \bar{p})(o_i - \bar{o})}{n \cdot S_{pred} \cdot S_{obs}} \quad (۵)$$

جدول ۱- توصیف آماری پارامترهای مستقل و وابسته مورد استفاده در مدل سازی

| متغیر             | آماره | Min   | Max   | Mean  | SD |
|-------------------|-------|-------|-------|-------|----|
| O <sub>3</sub>    | ۱     | ۱۲۵   | ۳۳/۳۱ | ۱۶/۲۶ |    |
| SO <sub>2</sub>   | ۲     | ۱۲۱   | ۲۰/۰۹ | ۱۲/۲۳ |    |
| CO                | ۰     | ۹/۱   | ۱/۴۲  | ۱/۲۹  |    |
| NO <sub>2</sub>   | ۴     | ۱۵۱   | ۵۲    | ۲۹/۳۷ |    |
| PM <sub>2.5</sub> | ۲     | ۱۳۱   | ۲۷/۷  | ۱۷/۶۴ |    |
| NO                | ۱۰/۷۲ | ۲۲۳   | ۵۷    | ۳۴    |    |
| دما               | -۷    | ۳۶/۳  | ۱۷/۷  | ۱۰/۲۹ |    |
| رطوبت             | ۷/۸۷  | ۷۰/۵  | ۳۴/۰۴ | ۱۵/۱۹ |    |
| سرعت باد          | ۰     | ۱۲/۲۵ | ۲/۵۵  | ۱/۶۸  |    |

گرفته‌اند دارای سطح معنی داری مقبولی هستند ( $P < 0.001$ ). بدین معنا که تمامی پارامترهای مورد استفاده در این مدل سازی به بهبود عملکرد مدل برای تخمین غلظت ازن در سطح زمین کمک خواهند کرد. مدل توسعه داده شده فوق توانایی توصیف ۷۹ درصد داده‌ها برای سال ۱۳۹۳ را داشته است و مقدار RMSE به دست آمده طی دوره مدل سازی ۱۱/۹۳ بوده است (نمودار ۱). پس می‌توان گفت که مدل توسعه داده شده توانایی توصیف در مدل سازی غلظت ازن داشته است. همچنین این مدل حداقل ۷۴ درصد داده برای بهار، و حداکثر ۷۷ درصد برای فصل تابستان را داشته است. لازم به ذکر است که در فرایند مدل سازی برای فصل‌های مختلف از داده‌های ۶۰ تا ۷۵ روز استفاده شده است. همچنین این مدل کمترین و بیشترین

دادن اثر خورشید توسط رطوبت نسبی باشد. غلظت ازن تنها با پارامتر دما همبستگی مثبت داشته است و دلیل آن این است که فرایند تشکیل ازن در نتیجه افزایش دما بیشتر می‌شود. بیشترین همبستگی منفی و مثبت بین ازن و سایر پارامترها به ترتیب مربوط به NO<sub>2</sub> و دما بوده است. بنابراین در مناطقی که بیشترین مقدار دما و ساعات آفتابی (۳۳، ۳۴) و یا فصولی که بیشترین میزان انرژی دریافتی از خورشید داشته باشند بیشترین میزان غلظت ازن نیز وجود خواهد داشت.

#### - عملکرد مدل توسعه داده شده

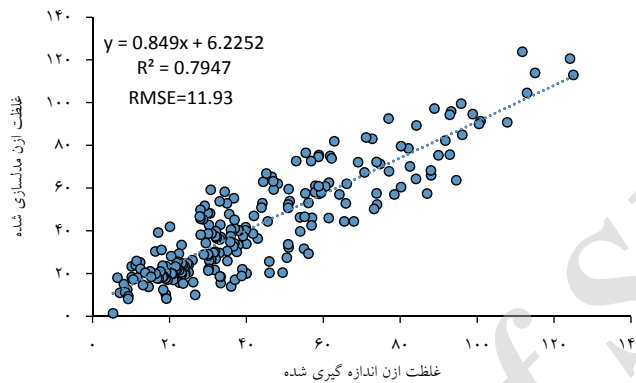
نتایج مربوط به عملکرد فرایند مدل سازی به صورت زمانی و مکانی در نمودار ۱، ۲ و ۳ نشان داده شده است. تمامی متغیرهای مستقل که در این مدل سازی مورد استفاده قرار

جدول ۲- میزان ضریب همبستگی (R) بین آلاینده‌ها و هواشناسی با غلظت ازن

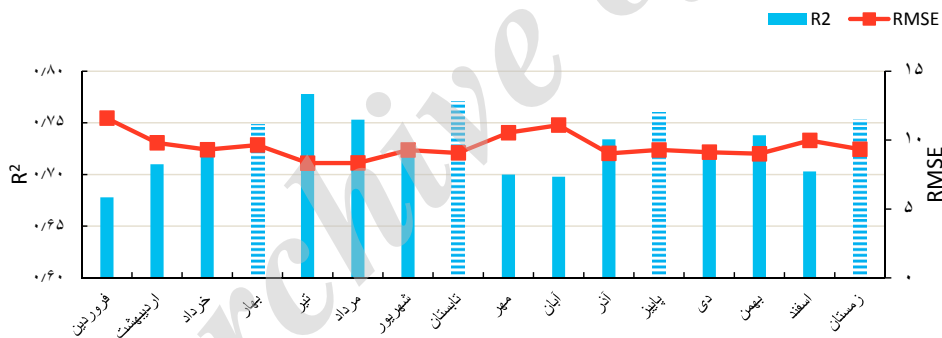
| متغیر          | CO     | SO <sub>2</sub> | NO     | NO <sub>2</sub> | PM <sub>2.5</sub> | دما    | رطوبت نسبی | سرعت باد | آلاینده |
|----------------|--------|-----------------|--------|-----------------|-------------------|--------|------------|----------|---------|
| O <sub>3</sub> | -۰/۲۷۶ | -۰/۰۱۴          | -۰/۲۱۴ | -۰/۴۰۳          | -۰/۱۲۲            | +۰/۲۸۹ | -۰/۲۸۳     | -۰/۱۴۸   |         |

است (نمودار ۲). بیشترین میزان ضریب تشخیص (۰/۷۷) و کمترین RMSE (۸/۳۳) برحسب ppb مربوط به تیر ماه و کمترین میزان ضریب تشخیص (۰/۶۷) و بالاترین RMSE (۱۱/۵۹) برحسب ppb مربوط به فروردین ماه است. نتایج ارزیابی دقت مکانی در نمودار ۳ نشان می‌دهد که

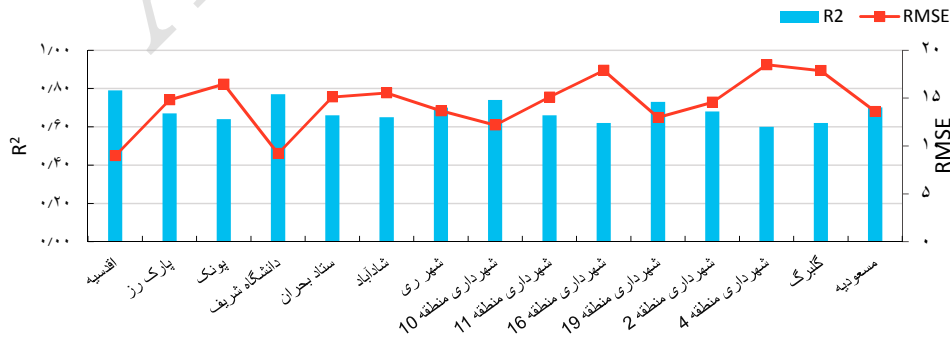
میزان RMSE را برحسب ppb به ترتیب برای فصل تابستان (۹/۰۷) و بهار (۹/۶۴) داشته است. مدل توسعه داده شده برای ماه‌های مختلف نیز مورد ارزیابی قرار گرفته است و برای هر ماه از ۱۷ تا ۲۵ روز برای مدلسازی استفاده شده است. نتایج مدلسازی برای ماه‌های مختلف نشان می‌دهد که میزان  $R^2$  اختلاف بین ۰/۶۷ تا ۰/۷۷ در برآورد مقدار غلظت ازن داشته



نمودار ۱- مقایسه مقادیر غلظت ازن اندازه‌گیری شده و مدلسازی شده



نمودار ۲- نتایج مدلسازی غلظت ازن برحسب ppb



نمودار ۳- نتایج دقت مکانی غلظت ازن

پیش‌بینی و همچنین RMSE کم بیانگر سطح قابل قبول این مدل‌سازی است. اگرچه مطالعات گوناگونی در رابطه با پیش‌بینی غلظت ازن انجام گرفته است اما برای شهر تهران مطالعاتی جامع و به‌گونه‌ای که تمامی پارامترهای هواشناسی و غلظت آلاینده‌های موثر را مدنظر قرار دهد انجام نگرفته است. بررسی ضریب همبستگی پیروسون بین غلظت ازن و سایر آلاینده‌ها و داده‌های هواشناسی نشان داد که میزان این همبستگی در بازه  $-0.403$  تا  $+0.289$  متغیر بوده است که عامل دما بیشترین همبستگی مثبت و  $\text{NO}_2$  بیشترین همبستگی منفی با غلظت ازن را داشته است.

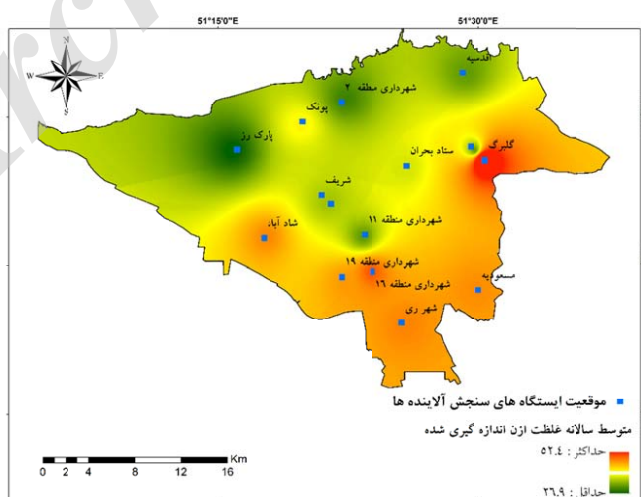
مدل توسعه داده شده در این تحقیق با ضریب تشخیص بالا ( $R^2=0.79$ ) و همچنین مقدار RMSE پایین ( $\text{RMSE}=11/93$ ) از توانایی قابل قبولی در مدل‌سازی غلظت ازن برخوردار است. بررسی تغییرات زمانی و مکانی غلظت ازن پیش‌بینی شده نشان می‌دهد که این مدل از توانایی بالایی جهت مدل‌سازی و تخمین غلظت ازن سطح زمین برخوردار بوده است. به‌گونه‌ای که این مدل توانایی توصیف حداقل ۶۷ درصد برای فروردین ماه و حداکثر ۷۷ درصد برای تیر ماه را داشته باشد. همچنین این مدل توانسته است تغییرات مکانی غلظت ازن را در بازه بین ۶۰ درصد برای ایستگاه شهرداری منطقه ۴ و ۷۹ درصد برای ایستگاه اقدسیه

بیشترین میزان  $R^2$  برای ایستگاه اقدسیه (۰/۷۹) و کمترین مقدار آن برای ایستگاه شهرداری منطقه ۴ (۰/۶) بوده است. همچنین نتایج این مدل حداقل RMSE را برای ایستگاه اقدسیه (۸/۹۸) و حداکثر آن را نیز برای ایستگاه شهرداری منطقه ۴ (۱۸/۴۸) برحسب ppb داشته است. لازم به ذکر است که تعداد کل داده‌های مورد استفاده برای ارزیابی دقت مدل‌سازی مکانی ۳۳۲۸ داده بوده است.

توزیع مکانی غلظت ازن اندازه‌گیری شده و مدل‌سازی شده برای دوره یکساله در شکل ۲ و ۳ نشان داده شده است. میانگین سالانه غلظت ازن اندازه‌گیری شده برای شهر تهران بین ۲۶/۹ تا ۵۲/۴ و مقدار مدل‌سازی شده در بازه ۲۸/۹ تا ۵۲/۹ برحسب ppb است. بیشترین مقادیر غلظت ازن مدل‌سازی شده مربوط به مناطق جنوبی و شرق شهر تهران است که با مقدار غلظت ازن اندازه‌گیری شده در سطح ایستگاه‌ها تطابق دارد.

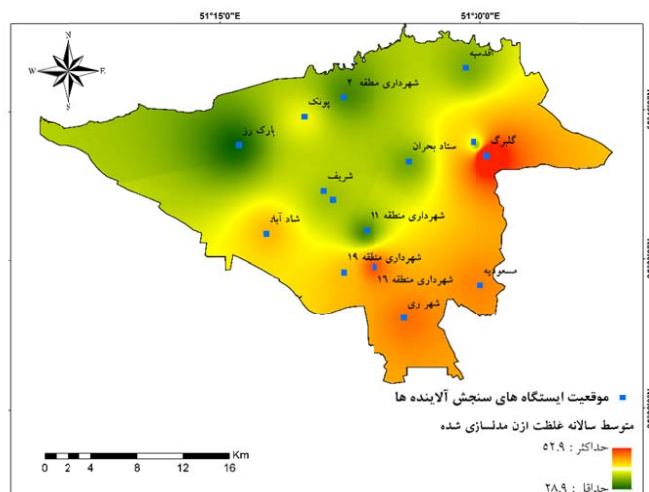
## بحث

مقایسه پژوهش انجام گرفته با نتایج سایر مطالعات پیشین بیانگر این موضوع است که نتایج این تحقیق و ضرایب تشخیص میان مقادیر اندازه‌گیری شده و مدل‌سازی شده در سطح قابل قبولی نسبت به سایر مطالعات بوده است (۱۲، ۱۷، ۳۵، ۳۶). یافته‌های این تحقیق و ضریب تشخیص بالای



شکل ۲- توزیع مکانی متوسط سالانه غلظت ازن اندازه‌گیری شده برحسب ppb





شکل ۳- توزیع مکانی متوسط سالانه غلظت ازن مدلسازی شده بر حسب ppb

داشته‌اند در همین راستا در این مطالعه از مدل رگرسیون خطی چند متغیره و همچنین با استفاده از متغیرهای مربوط به ایستگاه‌های زمینی سنجش آلاینده های جوی و داده‌های هواشناسی مقدار غلظت ازن مدلسازی گردید. برای نشان دادن اهمیت و تاثیر هر یک از این متغیرها بر مقدار غلظت ازن از همبستگی پیرسون استفاده شده است. نتایج این همبستگی نشان داد که تمامی متغیرهای مورد استفاده به غیر از دما همبستگی منفی با غلظت ازن داشته‌اند. نتایج مدل توسعه داده شده در این تحقیق حاکی از عملکرد قابل قبول آن برای تخمین غلظت ازن در سطح زمین بوده است. مدل رگرسیون خطی ضریب تشخیص بالای  $(R^2=0.79)$  و RMSE پایین  $(RMSE=11.93)$  برای کل داده‌های سال ۱۳۹۳ داشته است. نتایج مدل رگرسیون خطی چند متغیره برای فصول مختلف نشان داد که بالاترین ضریب تشخیص برای فصل تابستان  $(R^2=0.77)$  بوده است. همچنین نتایج مدلسازی مکانی نشان داد که بهترین ضریب تشخیص برای ایستگاه اقدسیه  $(R^2=0.79)$  و کمترین ضریب تشخیص برای ایستگاه شهرداری منطقه ۴  $(R^2=0.60)$  است. توزیع مکانی مقدار غلظت ازن مدلسازی در سطح شهر تهران با مقدار غلظت اندازه‌گیری شده در سطح ایستگاه‌ها همخوانی داشته و مناطق جنوبی و شرق شهر تهران به‌عنوان مناطق آلوده مشخص شده‌اند به‌عبارت دیگر با مقایسه نقشه غلظت ازن اندازه‌گیری

توصیف کند. مشکلات مربوط به عدم برداشت و یا برداشت‌های اشتباه غلظت آلاینده‌ها در سطح ایستگاه‌ها از محدودیت‌های انجام این تحقیق بوده است.

نتایج مربوط به توزیع مکانی غلظت ازن در شهر تهران نشان می‌دهد که بیشترین میزان غلظت ازن مربوط به ایستگاه‌های جنوبی و شرقی است که عمدتاً ناشی از تغییرات محلی دما، الگوی سایر آلاینده‌های جو و وزش بادهای محلی است (۱۹)، (۳۷). موقعیت جغرافیایی بخش‌های جنوبی و شرقی شهر تهران به دلیل نزدیکی آن با بخش‌های کویری ایران و در نتیجه افزایش دما و نقش مهمی که این پارامتر در فرایند تشکیل ازن دارد دلیلی بر افزایش غلظت ازن در این دو بخش است. رطوبت نسبی پایین‌تر بخش‌های جنوبی به دلیل همجواری با مناطق کویری، سرعت وزش باد کمتر در قسمت‌های جنوبی (۳۸) و جهت وزش باد نیز تاثیر مهمی در پراکنش آلاینده‌ها دارد (۳۹). جهت غالب باد در شهر تهران به‌طور معمول از غرب به شرق است در نتیجه آلاینده‌های عمده ناشی از حمل و نقل در مرکز شهر و برخی مناطق صنعتی را به‌سوی شرق انتقال می‌دهد (۱۹).

### نتیجه‌گیری

همان‌طور که اشاره شده مدل‌های مختلفی برای مدلسازی غلظت ازن ارائه شده و هر کدام از این مدل‌ها نتایج مختلفی

موثر باشند بنابراین مطالعات بیشتری برای درک بهتر اثرات پارامترهای دیگر بر غلظت ازن، مانند طیف UV-B خورشید و همچنین پدیده وارونگی حرارتی (Thermal inversion phenomena) که منجر به انباشت گازهای جوی می‌شود لازم است.

### منابع

1. Kampa M, Castanas E. Human health effects of air pollution. *Environmental Pollution*. 2008;151(2):362-67.
2. Collins WJ, Stevenson DS, Johnson CE, Derwent RG. The European regional ozone distribution and its links with the global scale for the years 1992 and 2015. *Atmospheric Environment*. 2000;34(2):255-67.
3. Devlin RB, Duncan KE, Jardim M, Schmitt MT, Rappold AG, Diaz-Sanchez D. Controlled exposure of healthy young volunteers to ozone causes cardiovascular effects. *Circulation*. 2012;126(1):104-11.
4. Holloway MJ, Arnold SR, Challinor AJ, Emberson LD. Intercontinental trans-boundary contributions to ozone-induced crop yield losses in the Northern Hemisphere. *Biogeosciences*. 2012;9(1):271-92.
5. Fuhrer J, Val Martin M, Mills G, Heald CL, Harmens H, Hayes F, et al. Current and future ozone risks to global terrestrial biodiversity and ecosystem processes. *Ecology and Evolution*. 2016;6(24):8785-99.
6. Mills G, Hayes F, Simpson D, Emberson L, Norris D, Harmens H, et al. Evidence of widespread effects of ozone on crops and (semi-)natural vegetation in Europe (1990-2006) in relation to AOT40- and flux-based risk maps. *Global Change Biology*. 2011;17(1):592-613.
7. Wilkinson S, Mills G, Illidge R, Davies WJ. How is ozone pollution reducing our food supply? *Journal of Experimental Botany*. 2012;63(2):527-36.
8. Li X, Rappenglück B. A WRF-CMAQ study on spring time vertical ozone structure in Southeast Texas. *Atmospheric Environment*. 2014;97:363-85.
9. Vingarzan R. A review of surface ozone background levels and trends. *Atmospheric Environment*. 2004;38(21):3431-42.
10. Hidy G. Ozone process insights from field experiments—part I: overview. *Atmospheric Environment*. 2000;34(12):2001-22.
11. Klein PM, Hu X-M, Xue M. Impacts of Mixing Pro-

cesses in Nocturnal Atmospheric Boundary Layer on Urban Ozone Concentrations. *Boundary-Layer Meteorology*. 2013;150(1):107-30.

12. Draxler RR. Meteorological factors of ozone predictability at Houston, Texas. *Journal of the Air & Waste Management Association*. 2000;50(2):259-71.
13. Abdul-Wahab SA, Al-Alawi SM. Assessment and prediction of tropospheric ozone concentration levels using artificial neural networks. *Environmental Modelling & Software*. 2002;17(3):219-28.
14. Blanchard CL. Ozone process insights from field experiments – Part III: extent of reaction and ozone formation. *Atmospheric Environment*. 2000;34(12):2035-43.
15. Ryu YH, Baik JJ, Kwak KH, Kim S, Moon N. Impacts of urban land-surface forcing on ozone air quality in the Seoul metropolitan area. *Atmospheric Chemistry and Physics*. 2013;13(4):2177-94.
16. Abdul-Wahab SA, Bakheit CS, Al-Alawi SM. Principal component and multiple regression analysis in modelling of ground-level ozone and factors affecting its concentrations. *Environmental Modelling & Software*. 2005;20(10):1263-71.
17. Sousa SIV, Martins FG, Alvim-Ferraz MCM, Pereira MC. Multiple linear regression and artificial neural networks based on principal components to predict ozone concentrations. *Environmental Modelling & Software*. 2007;22(1):97-103.
18. Sotoudeheian S, Arhami M. Estimating ground-level PM10 using satellite remote sensing and ground-based meteorological measurements over Tehran. *Journal of Environmental Health Science and Engineering*. 2014;12(1):122.
19. Sotoudeheian S, Arhami M. Using linear mixed effect model to estimate ground-level PM2.5: case study for Tehran. *Iranian Journal of Health and Environment*. 2017;10(2):213-24 (in Persian).

20. Ahmadi Moghadam M, Mahmoudi P. Analysis of tehran air pollution data in recent decade (2000-2009). *Iranian Journal of Health and Environment*. 2013;6(1):33-44 (in Persian).
21. Zhou Y, Mao H, Demerjian K, Hogrefe C, Liu J. Regional and hemispheric influences on temporal variability in baseline carbon monoxide and ozone over the Northeast US. *Atmospheric Environment*. 2017;164:309-24.
22. Ahamad F, Latif MT, Tang R, Juneng L, Dominick D, Juahir H. Variation of surface ozone exceedance around Klang Valley, Malaysia. *Atmospheric Research*. 2014;139:116-27.
23. Sullivan JT, Rabenhorst SD, Dreessen J, McGee TJ, Delgado R, Twigg L, et al. Lidar observations revealing transport of O<sub>3</sub> in the presence of a nocturnal low-level jet: Regional implications for "next-day" pollution. *Atmospheric Environment*. 2017;158(Supplement C):160-71.
24. Tiwary A, Colls J. *Air pollution: measurement, modelling and mitigation*. New York: Taylor & Francis; 2009.
25. Reeves CE, Penkett SA, Bauguitte S, Law KS, Evans MJ, Bandy BJ, et al. Potential for photochemical ozone formation in the troposphere over the North Atlantic as derived from aircraft observations during ACSOE. *Journal of Geophysical Research: Atmospheres*. 2002;107(D23):ACH 14-1-ACH 14-14.
26. Awang NR, Ramli NA, Yahaya AS, Elbayoumi M. Multivariate methods to predict ground level ozone during daytime, nighttime, and critical conversion time in urban areas. *Atmospheric Pollution Research*. 2015;6(5):726-34.
27. Özbay B. Modeling the Effects of Meteorological Factors on SO<sub>2</sub> and PM<sub>10</sub> Concentrations with Statistical Approaches. *CLEAN - Soil, Air, Water*. 2012;40(6):571-77.
28. Gvozdić V, Kovač-Andrić E, Brana J. Influence of meteorological factors NO<sub>2</sub>, SO<sub>2</sub>, CO and PM<sub>10</sub> on the concentration of O<sub>3</sub> in the urban atmosphere of Eastern Croatia. *Environmental Modeling & Assessment*. 2011;16(5):491-501.
29. Elbayoumi M, Ramli NA, Md Yusof NFF, Yahaya ASB, Al Madhoun W, Ul-Saufie AZ. Multivariate methods for indoor PM<sub>10</sub> and PM<sub>2.5</sub> modelling in naturally ventilated schools buildings. *Atmospheric Environment*. 2014;94:11-21.
30. Ul-Saufie AZ, Yahaya AS, Ramli NA, Rosaida N, Hamid HA. Future daily PM<sub>10</sub> concentrations prediction by combining regression models and feedforward backpropagation models with principle component analysis (PCA). *Atmospheric Environment*. 2013;77:621-30.
31. Duan J, Tan J, Yang L, Wu S, Hao J. Concentration, sources and ozone formation potential of volatile organic compounds (VOCs) during ozone episode in Beijing. *Atmospheric Research*. 2008;88(1):25-35.
32. San José R, Stohl A, Karatzas K, Bohler T, James P, Pérez JL. A modelling study of an extraordinary night time ozone episode over Madrid domain. *Environmental Modelling & Software*. 2005;20(5):587-93.
33. Chelani AB. Statistical persistence analysis of hourly ground level ozone concentrations in Delhi. *Atmospheric Research*. 2009;92(2):244-50.
34. Shan W, Yin Y, Lu H, Liang S. A meteorological analysis of ozone episodes using HYSPLIT model and surface data. *Atmospheric Research*. 2009;93(4):767-76.
35. Barrero MA, Grimalt JO, Cantón L. Prediction of daily ozone concentration maxima in the urban atmosphere. *Chemometrics and Intelligent Laboratory Systems*. 2006;80(1):67-76.
36. Moustiris KP, Nastos PT, Larissi IK, Paliatsos AG. 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;2012: Article ID: 894714. <http://dx.doi.org/10.1155/2012/894714>
37. Seaman NL. Meteorological modeling for air-quality assessments. *Atmospheric Environment*. 2000;34(12):2231-59.
38. Ghotbi S, Sotoudeheian S, Arhami M. Estimating urban ground-level PM<sub>10</sub> using MODIS 3km AOD product and meteorological parameters from WRF model. *Atmospheric Environment*. 2016;141:333-46.
39. Nayeb Yazdi M, Delavarrafiee M, Arhami M. Evaluating near highway air pollutant levels and estimating emission factors: Case study of Tehran, Iran. *Science of the Total Environment*. 2015;538:375-84.



Available online: <http://ijhe.tums.ac.ir>

Original Article



## Ozone estimation on the ground surface using multivariate linear regression and parameter determination

K Ezimand, AA Kakroodi \*

Department of Remote Sensing and GIS, Faculty of Geography, University of Tehran, Tehran, Iran

### ARTICLE INFORMATION:

**Received:** 11 November 2017  
**Revised:** 28 January 2018  
**Accepted:** 4 February 2018  
**Published:** 7 March 2018

**Key words:** Pollution concentration, Meteorological parameters, Estimation of ozone concentration, Multivariate linear regression

**\*Corresponding Author:**  
a.a.kakroodi@ut.ac.ir

### ABSTRACT

**Background and Objective:** Ground level ozone ( $O_3$ ) is one of most dangerous pollutants for human health in urban areas. The aim of this study was to identify the factors affecting the formation of ozone and modeling the spatial and temporal variations of ozone concentration in Tehran metropolitan area.

**Materials and Methods:** The data used in this research included meteorological data and pollution concentration data for 2014. First, we studied the impact and correlation of parameters to ozone concentration using the coefficient of Pearson, and then we did modeling of ozone concentration using a multivariate linear regression method.

**Results:** The developed model had the ability to describe 79% of the data changes for 2014. The temporal analysis of the ozone concentration showed that the best coefficient of determination of the model was  $R^2 = 0.771$  in the summer and  $R^2 = 0.778$  in July. These results also showed that among the air quality monitoring station of Tehran, station 4 had the lowest coefficient of determination ( $R^2 = 0.6$ ) and Aqdasieh station had the highest coefficient of determination ( $R^2 = 0.79$ ). Finally, the spatial distribution of the estimated ozone concentration was consistent with the measured ozone concentration at the station level.

**Conclusion:** According to the results, all the parameters related to air pollution concentration and meteorological parameters were effective parameters on modeling of ozone concentration on the ground level. The spatial distribution of ozone concentration in Tehran showed a greater concentration of ozone in the South and East than the North and West of the city.