



دانشگاه آزاد اسلامی واحد اهر
فصلنامه‌ی علمی-پژوهشی فضای جغرافیایی

سال شانزدهم، شماره‌ی ۵۵
پاییز ۱۳۹۵، صفحات ۱۷۶-۱۵۵

شادی اوسطی^۱

جمیل امان‌اللهی^۲

بختیار محمدی^۳

ارزیابی دقت مدل‌های رگرسیون چندگانه، ARIMA و GRNN در پیش‌بینی غلظت ذرات معلق PM_{2.5} در شهر سنندج

تاریخ پذیرش: ۱۳۹۳/۱۲/۰۹

تاریخ دریافت: ۱۳۹۳/۰۳/۱۶

چکیده

وجود مشکلات و بحران‌های زیست‌محیطی در جهان باعث شده است تا طی نیم قرن گذشته اهمیت بحث در مورد محیط‌زیست و مسایل زیست‌محیطی بیش‌تر شود. در این میان آلودگی هوا به‌عنوان یکی از معضلات زیست‌محیطی ده‌های اخیر در ردیف یکی از زیان‌بارترین بلایای طبیعی قرار گرفته است. با توجه به اثرگذاری زیان‌بار ذرات معلق بر روی سلامت انسان‌ها، پیش‌بینی مقدار غلظت این ذرات در روزهای آینده می‌تواند سبب کاهش این اثرات شود. لذا در این مطالعه به منظور افزایش صحت پیش‌بینی ذرات معلق PM_{2.5} در شهر سنندج دقت مدل‌های خطی شامل رگرسیون خطی چندگانه (MLR) و خود رگرسیون میانگین متحرک انباشته (ARIMA) و مدل غیرخطی شبکه عصبی رگرسیون عمومی (GRNN) مورد آزمون قرار گرفت. داده‌های مربوط به کیفیت هوا شامل PM₁₀، SO₂، NO₂، CO، O₃ و داده‌های مربوط به هواشناسی شامل میانگین حداقل دما (MinT)، میانگین حداکثر دما (MaxT)، میانگین فشار جوی (AP)، بارش کل روزانه (PR)، رطوبت نسبی روزانه سطح هوا (RH) و سرعت باد روزانه (WS)

۱- دانشجوی کارشناسی ارشد محیط‌زیست، دانشگاه کردستان، سنندج، ایران.

۲- استادیار گروه محیط‌زیست، دانشگاه کردستان، سنندج، ایران.

۳- استادیار گروه آب‌وهوا شناسی، دانشگاه کردستان، سنندج، ایران.

مربوط به سال ۱۳۹۳ به‌عنوان متغیرهای مستقل و میزان غلظت ذرات معلق $PM_{2.5}$ به‌عنوان متغیر وابسته در نظر گرفته شدند. نتایج نشان داد که مدل شبکه عصبی رگرسیون عمومی با مقدار $R^2=0/81$ و $RMSE=6/9468$ و $MAE=5$ در مرحله آموزش و مقدار $R^2=74$ و $RMSE=0/0725$ و $MAE=3/4874$ در مرحله آزمون بهترین عملکرد را نسبت به مدل‌های خطی در پیش‌بینی ذرات معلق $PM_{2.5}$ در شهر سنندج داشته است.

کلید واژه‌ها: ذرات معلق $PM_{2.5}$ ، آلودگی هوا، سنندج، رگرسیون خطی چندگانه، خود رگرسیون میانگین متحرک انباشته، شبکه عصبی رگرسیون عمومی.

مقدمه

افزایش مشکلات و بحران‌های زیست‌محیطی در جهان از یک طرف و درک پیامدهای بلندمدت موضوعات زیست‌محیطی در زندگی انسان‌ها از سوی دیگر، باعث شده است تا طی نیم قرن گذشته اهمیت بحث در مورد محیط‌زیست و مسایل زیست‌محیطی افزایش یابد (صالحی و قائمی‌اصل، ۱۳۹۲: ۶۸). بی‌شک زندگی انسان از روزهای آغازین تا کنون همواره با خطرات ناشی از عوامل طبیعی همراه بوده است، که یکی از این خطرات، آلودگی هوا می‌باشد که در ردیف یکی از زیان‌بارترین بلاهای طبیعی قرار گرفته است و مشکلات محیطی عدیده‌ای را به‌وجود آورده است. هوا نیز مانند سایر منابع محیط‌زیست دارای ظرفیت محدود می‌باشد و تحمل تخلیه مواد زاید و سمی مختلف را در حدی که امروزه بشر به آن تحمیل کرده است ندارد (نصراللهی و غفاری^۴، ۲۰۱۰: ۳۷۷). آلودگی هوا عبارت است از حضور یک یا بیش از یک آلاینده در هوای آزاد به مقدار کافی، با خواص مشخص و تداوم کافی که می‌تواند حیات انسان، گیاه یا جانوران و یا اموال انسانی را به مخاطره اندازد (قوامی، ۱۳۹۲: ۱۱). یکی از شاخص‌های مهم در بحث کیفیت هوا غلظت ذرات معلق (PM)^۵ می‌باشد (بالداسانو و جیمینز^۶، ۲۰۰۳: ۱۴۳). ذرات معلق موجود در هوا می‌تواند منشأ طبیعی یا انسانی داشته باشند. منابع طبیعی تشکیل دهنده ذرات معلق، شامل ذرات جامد ناشی از فعالیت آتشفشان (تیمرک و همکاران^۷، ۲۰۰۹: ۱)، ذرات نمک برخاسته از اقیانوس‌ها و گردوغبار حمل شده توسط باد از مناطق بیابانی (کافمن و همکاران^۸، ۲۰۰۲: ۲۱۶) و منابع انسانی آن شامل ذرات تولید شده توسط دودکش کارخانجات، سوزاندن زباله، وسایل گرمایشی منازل و آتش‌سوزی‌های عمدی

4- Nasrollhi and Ghaffari

5- Particulate matter

6- Baldasano and Jiménez

7- Timmreck et al

8- Kaufman et al

و غیر عمدی هستند (پستوسکا و همکاران^۹، ۲۰۱۰: ۶۱۴). ذرات معلق که در مجاورت جاده‌ها یا صنایع مولد غبار یافت می‌شوند و قطرشان بزرگتر از ۲/۵ میکرون و کوچک‌تر از ۱۰ میکرون است را PM₁₀ می‌نامند و ذرات معلق که در دود و مه‌دود یافت می‌شود و قطرشان ۲/۵ میکرون و یا کم‌تر است را PM_{2.5} می‌نامند. مضر بودن این ذرات برای سلامتی انسان‌ها توسط محققان زیادی ثابت شده است (آدامسون و همکاران^{۱۰}، ۱۹۹۹: ۴۳) که باعث بیماری‌های آسم، ذات‌الریه و عفونت‌های مجاری تنفسی می‌شود (سندستروم^{۱۱}، ۲۰۰۸: ۸۰۸). ذرات معلق علاوه بر ایجاد مشکلات سلامتی برای انسان‌ها باعث تأثیر بر روی مقدار تابش خورشید به زمین و تغییر سیستم زمین و جو، تأثیر روی الگوهای جریان جوی، تغییر دمای سطحی زمین و بارش و کاهش دید نیز می‌شوند (گوپتا و ساندر^{۱۲}، ۲۰۰۸: ۶۴۶۶). تحقیقی که توسط WHO^{۱۳} در برلین، کپنهاک و رم صورت گرفت، نشان داد PM در غلظت‌های طولانی مدت، باعث افزایش ۶٪ مرگ‌ومیر و افزایش هر $10 \mu\text{g}/\text{m}^3$ در غلظت PM_{2.5}، افزایش ۱۲٪ بیماری‌های قلبی عروقی و ۱۴٪ سرطان ریه را در پی دارد (عزیزی‌فر و همکاران، ۱۳۹۰: ۶۰). با توجه به اثرات زیان‌بار این ذرات بر سلامتی انسان‌ها، توانایی پیش‌بینی میزان غلظت آن‌ها و صدور اخطار از رسانه‌های عمومی می‌تواند سبب بالا رفتن سلامت عمومی جامعه گردد. اخیراً روش‌های متعددی برای پیش‌بینی غلظت PM_{2.5} ارائه شده است. مدل رگرسیون خطی چندگانه (MLR) روشی است که برای ارتباط خطی بین یک متغیر وابسته و یک یا چند متغیر مستقل مورد استفاده قرار می‌گیرد (کازمانوسکی و آلكسوسکا^{۱۴}، ۲۰۰۳: ۱۶۹). مدل‌های رگرسیون چندگانه به دلیل صحت تفسیرشان (ژو و همکاران^{۱۵}، ۲۰۱۴: ۲۶۵)، به طور گسترده‌ای در زمینه‌ی پیش‌بینی PM_{2.5} مورد استفاده قرار می‌گیرند (کوبورن^{۱۶}، ۲۰۱۰: ۳۰۱۷). یکی دیگر از روش‌های پیش‌بینی که در علوم جوی کاربرد زیادی دارد روش شبکه‌های عصبی مصنوعی^{۱۷} (ANN) است که به صورت پیشرونده و موفق برای سیستم‌های غیرخطی در زمینه‌های علمی و مهندسی به‌کار می‌رود. پیش‌بینی سری‌های زمانی با استفاده از مدل‌های تجزیه و تحلیل از دیرباز در منابع علمی جهان در حال انجام است (حسن^{۱۸}، ۲۰۱۴: ۴۲۲) که در این رابطه می‌توان به مطالعات زیر اشاره کرد. پرزو همکاران^{۱۹} (۲۰۰۰) مدل شبکه عصبی مصنوعی را برای پیش‌بینی PM_{2.5} در سانتیاگو به‌کار بردند و از غلظت PM_{2.5}

9- Pastuszka et al

10- Adamson et al

11- Sandstrom

12- Gupta and Sundar

13- World health organization

14- Kuzmanovski and Aleksavska

15- Zhou et al

16- Cobourn

17- Artificial Neural Network

18- Hassan

19- Pérez et al

سرعت باد و جهت باد به‌عنوان متغیرهای پیش‌بینی‌کننده استفاده کردند. آن‌ها نتایج را با مدل رگرسیون خطی مقایسه کردند و به این نتیجه رسیدند که شبکه عصبی نتایجی دقیق‌تری را ارائه می‌دهد. الکامل و همکاران^{۲۰} (۲۰۰۱) سطوح ازن را با استفاده از شبکه عصبی مصنوعی پیش‌بینی کردند. به همین منظور اطلاعات مربوط به هواشناسی (سرعت و جهت باد، رطوبت نسبی و درجه حرارت خورشید) و غلظت آلاینده‌های متان، منوکسیدکربن، دی‌اکسیدکربن و گردوغبار در یک منطقه صنعتی در کویت جمع‌آوری شد و عملکرد شبکه عصبی با مدل‌های رگرسیون خطی و غیرخطی مقایسه گردید. نتایج نشان داد که شبکه‌های عصبی پیش‌بینی بهتری را ارائه می‌دهد. چالولاکو و همکاران^{۲۱} (۲۰۰۳) مدل شبکه عصبی و MLR را برای پیش‌بینی غلظت PM₁₀ در آتن پیشنهاد دادند. به همین منظور متغیرهای هواشناسی طی یک دوره‌ی دو ساله به‌عنوان ورودی‌های شبکه عصبی و MLR مورد بررسی قرار گرفتند. نتایج نشان داد اگر شبکه عصبی مصنوعی به درستی اجرا شود می‌تواند گزینه مناسبی برای پیش‌بینی این آلاینده باشد. در مطالعه‌ای صدرموسوی و رحیمی (۱۳۸۸) از مدل MLR و شبکه عصبی مصنوعی به منظور پیش‌بینی کوتاه‌مدت غلظت ازن برای شهر تبریز استفاده کردند. داده‌های مورد نیاز شامل داده‌های هواشناسی (سرعت باد، رطوبت نسبی، جهت باد، درجه حرارت، بارندگی، فشار هوا، مقدار تابش و مقدار تبخیر) و داده‌های آلودگی هوا (غلظت ازن) است. نتایج به‌دست آمده نشان می‌دهد که مدل شبکه‌های عصبی توانایی‌های بیش‌تری نسبت به روش‌های رگرسیونی خطی داشته است. به‌طوری که ضریب همبستگی در مدل رگرسیون چندگانه ۰/۴۵ در حالی که ضریب همبستگی در شبکه‌های عصبی ۰/۹۱ بوده است. کای و همکاران^{۲۲} (۲۰۰۹) غلظت آلاینده‌های نزدیک جاده‌های شهری را با استفاده از شبکه عصبی مصنوعی محاسبه کردند. به‌این منظور غلظت منوکسیدکربن، نیتروژن‌دی‌اکسید و ازن را اندازه‌گیری کردند و نتایج نشان داد که این مدل قادر به پیش‌بینی دقیق آلاینده‌ها تا ۱۰ ساعت قبل از آن می‌باشد. بوداقپور و چرخستانی (۱۳۹۰) برای پیش‌بینی میزان غلظت آلاینده‌های شهر تهران از شبکه عصبی مصنوعی استفاده کردند. برای این منظور از آمار غلظت آلاینده هوای ثبت شده توسط حسگرهای نصب شده در سال‌های ۱۳۸۱ تا ۱۳۸۶ استفاده کردند (گاز NO_x). همچنین این پیش‌بینی با استفاده از رگرسیون غیرخطی انجام گرفته و خطای جذر میانگین مربعات مدل شبکه عصبی با معادلات رگرسیون مقایسه شده و مشاهده می‌شود که خطای مدل شبکه عصبی کم‌تر از روش رگرسیون است. منصور و همکاران (۱۳۹۲) به پیش‌بینی و بررسی وضع آلاینده‌های هوای شهر کرمان با استفاده از مدل خود رگرسیون میانگین متحرک انباشته (ARIMA) پرداختند. نتایج نشان می‌دهد که آلاینده‌ها در هوای کرمان روند تقریباً ثابتی داشتند به جز منواکسیدکربن که به طور

20- Elkamel et al

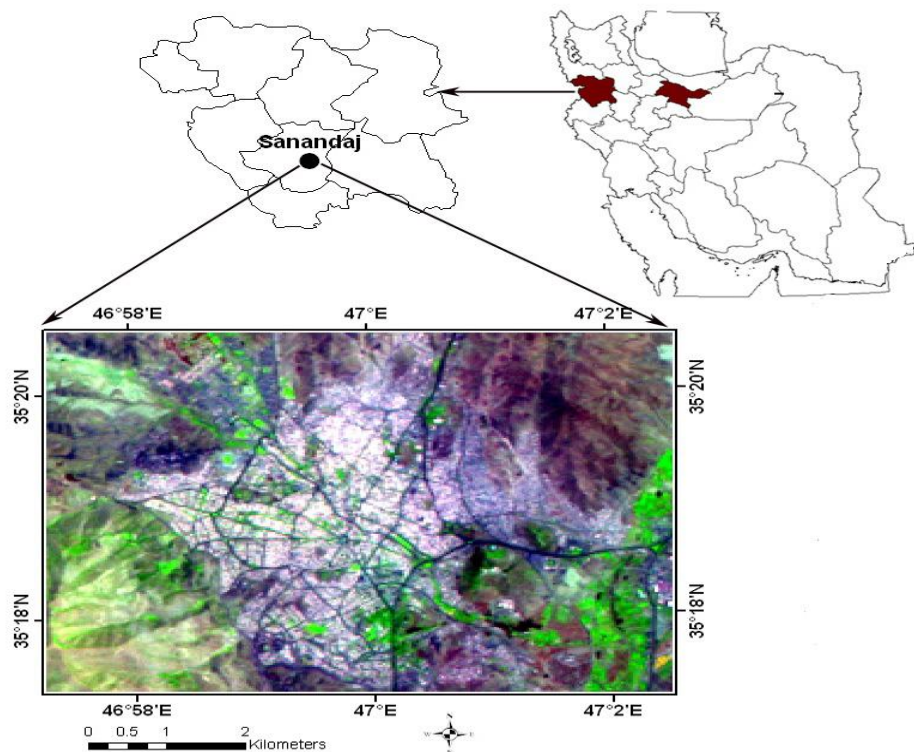
21- Chaloulakou et al

22- Cai et al

معنی داری در حال کاهش و گردوغبار که روند افزایشی داشت همه آلاینده‌ها الگوی فصلی داشتند. از آنجایی که سنندج به‌عنوان یکی از شهرهای آلوده‌ی کشور به حساب می‌آید، پیش‌بینی انواع آلودگی‌های آن به‌ویژه پیش‌بینی ذرات معلق $PM_{2.5}$ که عامل بسیاری از بیماری‌ها می‌باشد، می‌تواند در اعلام به موقع و قبل از افزایش میزان $PM_{2.5}$ کمک شایانی به سلامت جامعه کند. از این رو برای افزایش دقت پیش‌بینی مدل‌های خطی شامل ARIMA، MLR و مدل غیر خطی شبکه عصبی رگرسیون عمومی (GRNN) استفاده شدند و در نهایت دقیق‌ترین مدل از میان این مدل‌ها برای پیش‌بینی ذرات معلق $PM_{2.5}$ معرفی گردید.

منطقه مورد مطالعه

شهرستان سنندج با مساحت ۲۹۰۶ کیلومتر مربع مرکز استان کردستان با ۳۵ درجه و ۲۰ دقیقه عرض شمالی و ۴۷ درجه طولی شرقی از نصف‌النهار گرینویچ و در ارتفاع ۱۵۳۸ متری از سطح دریا قرار دارد (شکل ۱).



شکل ۱: موقعیت جغرافیایی شهر سنندج در کشور

مواد و روش‌ها

در این مطالعه از داده‌های مربوط به کیفیت هوای شهر سنندج شامل $PM_{2.5}$ ، PM_{10} ، SO_2 ، NO_2 ، CO و O_3 و داده‌های مربوط به هواشناسی شامل میانگین حداقل دما (MinT)، میانگین حداکثر دما (MaxT)، میانگین فشار جوی

(AP)، بارش کل روزانه (PR) و سرعت باد روزانه (WS) مربوط به سال ۱۳۹۳ برای پیش‌بینی غلظت ذرات معلق $PM_{2.5}$ شهر سنندج استفاده شد (جدول ۱).

جدول ۱- میانگین و انحراف معیار پارامترهای مورد استفاده شده

	میانگین	انحراف معیار
SO ₂	۲۱/۸۷	۷/۷۸
NO ₂	۱۷/۴	۶/۴۸
PM ₁₀	۶۸/۵۵	۳۰/۰۹
PM _{2.5}	۳۱/۰۲	۱۵/۳۱
CO	۱/۲۴	۱/۳۷
O ₃	۱۲/۲۶	۸/۳۷
MinT	۹/۲۶	۷/۲۰
MaxT	۲۷/۲۵	۱۰/۱۱
AP	۱۰۱۰/۰۸	۹/۰۷
PR	۰/۷	۲/۸۲
WS	۵/۹۲	۲/۳۶

(جدول ۱) میانگین و انحراف معیار پارامترهای استفاده شده در شهر سنندج در سال ۱۳۹۳ را نشان می‌دهد. PM_{10} و $PM_{2.5}$ دارای میانگین غلظت بالایی نسبت به دیگر آلاینده‌ها می‌باشند. در این سال میانگین بارش باران خیلی کم و در حدود ۰/۷ میلی‌متر برای هر روز است.

پیش‌بینی ذرات معلق $PM_{2.5}$ شهر سنندج

مدل رگرسیون خطی چندگانه (MLR)^{۲۳}

در مدل MLR مقادیر یک متغیر وابسته از روی مقادیر دو یا چند متغیر مستقل دیگر برآورد می‌شود. مدل‌های MLR در زمینه کیفیت هوا توسط محققین زیادی مورد استفاده قرار گرفته شده است (بارسناس و همکاران^{۲۴}، ۲۰۰۵: ۱۵۱؛ باسورکو و همکاران^{۲۵}، ۲۰۰۶: ۴۳۲). در این پژوهش آنالیز خطی دادها با استفاده از نرم‌افزار SPSS ورژن ۲۰ انجام شد. در مدل خطی، متغیر وابسته، میزان غلظت ذرات معلق $PM_{2.5}$ و متغیرهای مستقل، $PM_{2.5}$ روز قبل، SO_2 ، PM_{10} ، NO_2 ، CO ، O_3 ، $MinT$ ، $MaxT$ ، AP ، PR و WS در نظر گرفته شد (ژو و همکاران، ۲۰۱۴: ۲۶۶).

23- Multiple linear regression

24- Barcenás et al

25- Basurko et al

فرمول کلی مدل MLR به شکل زیر است (تیریکی و همکاران^{۲۶}، ۲۰۱۴: ۳۱):

$$Y = \beta_0 + \beta_1 X_1 + \dots + \beta_n X_n + \varepsilon$$

که در آن Y متغیر وابسته، X_1, \dots متغیرهای مستقل، β_0, \dots, β_n ثابت های معلوم و ε مقدار باقی مانده می باشد. در مدل MLR از روش گام به گام^{۲۷} استفاده شده است که براساس آن متغیرهای مستقل تک تک وارد معادله می شوند و اگر نقش معنی داری در رگرسیون نداشته باشند از آن خارج می شوند.

مدل خود رگرسیون میانگین متحرک انباشته (ARIMA)^{۲۸}

یک سری زمانی مجموعه مشاهداتی است که بر حسب زمان مرتب شده اند. پیش بینی سری های زمانی یکی از مهم ترین زمینه های پیش بینی است که در آن مشاهدات گذشته یک متغیر جمع آوری و به منظور به دست آوردن روابط اساسی بین مشاهدات و تعیین یک مدل توصیفی، مورد تجزیه و تحلیل قرار می گیرد (خاشعی و بیجاری، ۱۳۸۷: ۸۴). یکی از مهم ترین و پرکاربردترین مدل های سری های زمانی، مدل ARIMA می باشد. این مدل به الگوی باکس-جنکینز نیز معروف است و در نرم افزار MINITAB قابل بررسی است. مدل های پیش بینی باکس-جنکینز با سری های زمانی ایستا سروکار دارند. سری های زمانی به دو دسته کلی ایستا و غیرایستا تقسیم بندی می شوند. در سری های ایستا میانگین و پراکندگی در طول زمان ثابت می باشد، در غیر این صورت سری را غیرایستا می نامند. نایستایی در واریانس را با استفاده از تبدیل توانی باکس و کاکس^{۲۹} و نایستایی در میانگین را با استفاده از تفاضلی کردن می توان برطرف کرد (احمدی و همکاران، ۱۳۸۶: ۲۲). فرمول کلی آریما به شرح زیر است:

$$y_t = \sum_{i=1}^p f_i y_{t-1} + \sum_{i=1}^q \theta_i e_{t-i} + \varepsilon_t$$

در این رابطه y_t مقادیر مشاهده شده در زمان t، f_i پارامتر اتورگرسیو، θ_i پارامتر میانگین متحرک و ε_t خطا در زمان t است (کهنسال و پرمه، ۱۳۹۱: ۵۱). برای ساختن یک مدل آریما از یک روش سه مرحله ای تکراری استفاده می شود. به این ترتیب که ابتدا یک مدل آزمایشی از طریق تجزیه و تحلیل داده های تاریخی مشخص می شود. سپس پارامترهای نامعلوم مدل تخمین زده می شود. در نهایت آزمون خطا اجرا می شود تا شایستگی مدل را تعیین کند. چنانچه مدل مورد تایید قرار گرفت می توان آن را مبنای پیش بینی رفتار آینده سری قرار داد (خرمی و بزرگنیا، ۱۳۸۶: ۲۲). مدل آریما دارای سه پارامتر به صورت ARIMA (p, d, q) می باشد. پارامتر d مربوط به مرتبه

26- Tiryaki et al

27- Stepwise

28- Autoregressive Integrated moving average

29- Box- Cox

تفاضل‌گیری برای تبدیل سری نایستا به ایستا، پارامتر p مربوط به مرتبه جمله خودرگرسیون و q مرتبه جمله میانگین متحرک می‌باشد (زرانژاد و همکاران، ۱۳۹۲: ۳۸).

مدل‌های شبکه عصبی

روش شبکه عصبی برای اولین بار توسط مک کلاچ و پیترز در سال ۱۹۴۳ معرفی گردید. شبکه عصبی مصنوعی یک روش شبیه‌سازی بوده که از سیستم مغز و شبکه عصبی موجودات زنده الهام گرفته و از متداول‌ترین روش‌های پیش‌بینی در علوم مختلف است (سرمدیان و همکاران، ۱۳۸۹: ۲۱۴). شبکه عصبی متشکل از سه لایه است، لایه ورودی که داده‌ها را در شبکه توزیع می‌کند، لایه پنهان مسئول پردازش داده‌ها و لایه خروجی که نتایج را به ازای ورودی‌های مشخص استخراج می‌کند. انواع مختلفی از شبکه‌های عصبی مصنوعی وجود دارند که به منظور پیش‌بینی مورد استفاده قرار می‌گیرند که از آن جمله می‌توان به شبکه‌های عصبی چند لایه پیشخور (FFNN)، شبکه‌های عصبی شعاعی محور (RBF)، شبکه‌های عصبی احتمالی (PNN) و شبکه‌های عصبی رگرسیون عمومی (GRNN) اشاره کرد (خاشعی و همکاران، ۱۳۹۲: ۴). برای پیش‌بینی غلظت ذرات معلق $PM_{2.5}$ شهر سنج، از مدل GRNN استفاده گردید. مدل GRNN در این مطالعه برای اولین بار است که برای پیش‌بینی آلودگی اتمسفر در ایران استفاده می‌شود.

مدل شبکه عصبی رگرسیون عمومی (GRNN)^{۳۰}

به منظور تخمین روابط رگرسیونی سری‌های زمانی، نوع خاصی از شبکه‌های عصبی به نام GRNN مطرح شدند که قادرند با دقت بالایی پارامترهای روابط رگرسیونی حاکم بر سری‌های زمانی را تخمین زنند. GRNN توسط اسپچت^{۳۱} برای شناسایی و مدل‌سازی ارائه شد (اسپچت، ۱۹۹۱: ۵۶۹). شبکه‌های عصبی رگرسیون عمومی می‌توانند هر تابع را با در اختیار داشتن داده‌های گذشته ورودی و خروجی تقریب بزنند (له آنگ و همکاران^{۳۲}، ۲۰۰۰: ۱۷۶). از ویژگی‌های بارز آن زمان آموزش سریع و مدل‌سازی توابع غیرخطی است. این شبکه دارای فاکتور هموارساز^{۳۳} است که مقدار بهینه آن در طی چندین اجرا با توجه به میانگین مربعات خطا به دست می‌آید (مرادزاده و همکاران، ۱۳۹۰: ۵۷) (شکل ۲).

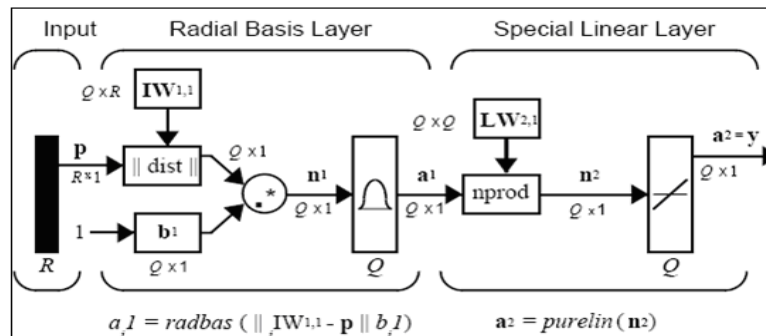
توپولوژی GRNN به شرح زیر است:

30- General Regression Neural Network

31- Specht

32- L- eung et al

33- Smooth Factor



شکل ۲: ساختار GRNN (دمیوت و بیل، ۲۰۰۲)

مدل GRNN دارای چهار لایه است:

لایه ورودی: این لایه وظیفه ورود اطلاعات به شبکه را بر عهده دارد. تعداد نرون های ورودی در یک GRNN برابر با تعداد متغیرهای موجود در بردار ورودی است.

لایه الگو: نرون های موجود در این لایه نرون های الگو نام دارند. این نرون ها با ترکیب و پردازش داده های ورودی آن ها را تنظیم می کنند. تعداد نرون های الگو در این لایه برابر با تعداد دسته های آموزشی خواهد بود. خروجی نرون الگوی نام از تابع زیر محاسبه می شود:

$$P_i = \exp\left(\frac{(X-X_i)^T(X-X_i)}{2\sigma^2}\right)$$

در این رابطه σ پارامتر هموارسازی، X متغیر ورودی و X_i نمونه آموزشی برای نرون نام در لایه الگو است.

لایه جمع: خروجی نرون های الگو به لایه سوم یعنی نرون های جمع کننده وارد می شوند. دو نوع جمع S_S و S_W در این لایه محاسبه می شود:

$$S_S = \sum_{i=1}^n P_i$$

$$S_W = \sum_{i=1}^n W_i P_i$$

در این رابطه W_i وزن نرون نام در لایه الگو است که به لایه جمع وصل شده است.

لایه خروجی: حاصل جمع های محاسبه شده در لایه جمع، به لایه خروجی فرستاده می شوند و نرون های این لایه خروجی را مطابق رابطه زیر محاسبه می نمایند:

$$y = S_W / S_S$$

ارزیابی مدل‌ها

یکی از معیارهای مورد بررسی اعتبار نتایج به‌دست آمده از مدل‌ها، آماره R^2 می‌باشد. آماره R^2 به‌عنوان یک شاخص برای بیان دقت خط رگرسیون به‌کار رفته و مقدار آن بین صفر و یک تغییر می‌کند و مقادیر نزدیک به یک تطابق بهتر داده‌های مشاهده‌ای و برآورده شده را نشان می‌دهد. ولی با توجه به این مسئله که مقدار R^2 تحت تأثیر داده‌های پرت است باید علاوه بر آن از پارامترهای دیگر استفاده کرد (لگاتس و مک کیب^{۳۵}، ۱۹۹۹: ۲۳۵). به همین دلیل از دیگر پارامترها نظیر ریشه متوسط مربعات خطا ($RMSE^{36}$) و میانگین مطلق خطا (MAE^{37}) نیز برای بررسی اعتبار نتایج استفاده می‌شود (نوری و همکاران، ۱۳۸۷: ۱۴۰) (جدول ۲). ریشه متوسط مربعات خطا و میانگین مطلق خطا طبق زیر محاسبه می‌شود:

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (Y_{da} - Y_{dp})^2}$$

$$MAE = \frac{1}{n} \sum_{i=1}^n |Y_{da} - Y_{dp}|$$

در این رابطه‌ها Y_{da} مقدار واقعی، Y_{dp} مقدار پیش‌بینی شده و n تعداد داده‌های موجود است.

نتایج پیش‌بینی ذرات معلق $PM_{2.5}$ شهر سنندج

نتایج مرحله آموزش مدل MLR در پیش‌بینی غلظت ذرات معلق $PM_{2.5}$ برای شهر سنندج

نتایج آموزش کاربرد مدل MLR در پیش‌بینی غلظت ذرات معلق $PM_{2.5}$ در شهر سنندج در (شکل ۳) نشان داده شده است و معادله به‌دست آمده به شکل زیر است:

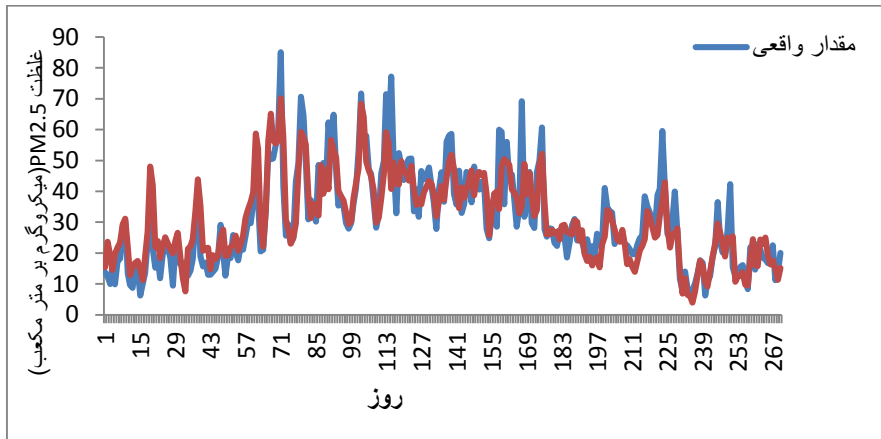
$$PM_{2.5F} = 3.403 - 0.353SO_2 + 0.25PM_{10} + 0.452PM_{2.5P} + 0.122O_3 + 0.244MAX - 0.67WS$$

در این بخش مقدار $R^2=0.76$ ، $RMSE=7/3539$ و $MAE=2555/5$ بود.

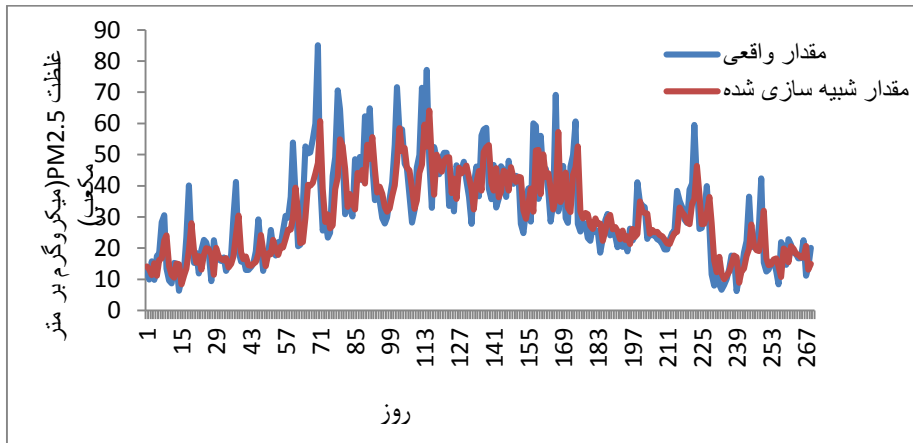
35- Legatesand McCabe

36- Root mean square error

37- Mean absolute error



شکل ۳: نتایج مدل MLR در مرحله آموزش مدل



شکل ۴: نتایج مدل ARIMA در مرحله آموزش

- نتایج مرحله آموزش مدل ARIMA انباشته در پیش بینی غلظت ذرات معلق $PM_{2.5}$ برای شهر سنندج مدل استفاده شده برای تجزیه سری زمانی $PM_{2.5}$ مدل $ARIMA(1, 1, 1)$ می باشد. معادله به دست آمده در این قسمت به صورت زیر است:

$$y_t = 0.572y_{t-1} - 0.9148e_{t-j} + \epsilon_t$$

نتایج آموزش مدل ARIMA در پیش بینی غلظت ذرات معلق $PM_{2.5}$ در شهر سنندج در (شکل ۴) نشان داده شده است.

در این قسمت مقدار $R^2 = 0.63$ ، $RMSE = 9/4137$ و $MAE = 6/767$ بود.

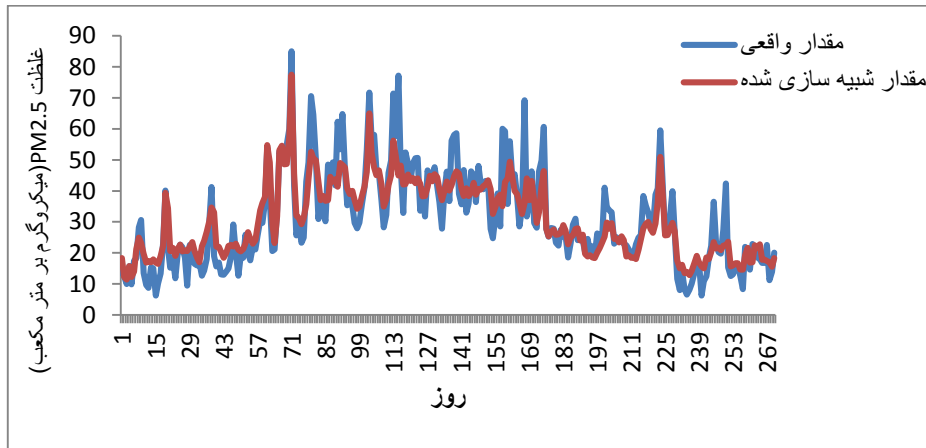
جدول ۲- مقادیر ضرایب همبستگی و جذر میانگین مربعات خطا با فاکتورهای هموارساز متفاوت

فاکتور هموارساز	آموزش R	آزمون R	آموزش RMSE	آزمون RMSE
۰/۰۵	۰/۹۹۹۹	۰/۶۸۰۳۵	۰/۲۲۰۹	۷/۳۷۱۷
۰/۰۹	۰/۹۹۵۱۲	۰/۷۲۰۹	۱/۵۲۵۲	۶/۷۵۵۸
۰/۱۲	۰/۹۸۲۵	۰/۷۷۹۳	۲/۸۶۸	۵/۹۱۴
۰/۱۶	۰/۹۶۴۴	۰/۸۱۹۹	۴/۰۸۳۶	۵/۳۶۰۹
۰/۲	۰/۹۴۷۶	۰/۸۳۸۴	۴/۹۵۸	۵/۱۲۸۸
۰/۲۴	۰/۹۳۲۲	۰/۸۴۹۲	۵/۶۴۳۴	۵/۰۲۳۴
۰/۲۸	۰/۹۱۸۹	۰/۸۵۳۴	۶/۱۸۷۸	۵/۰۴۳۲
۰/۳	۰/۹۱۳۱	۰/۸۵۴۲	۶/۴۲۲۲	۵/۰۴۳۲
۰/۳۵	۰/۹۰۰۱	۰/۸۵۹	۰/۹۴۶۵	۵/۰۷۲۵

- نتایج مرحله آموزش مدل GRNN در پیش‌بینی غلظت ذرات معلق $PM_{2.5}$ برای شهر سنندج ورودی‌های این مدل، خروجی مدل MLR می‌باشد. یعنی از متغیرهای مستقلی برای آموزش استفاده می‌کند که بر روی غلظت متغیر وابسته ($PM_{2.5}$) تأثیر داشته باشد. آموزش شبکه برای برآورد خروجی با فاکتورهای متفاوت هموارساز صورت گرفت. مقدار فاکتور هموارساز بین ۰ و ۱ می‌باشد (قیاس و اسمیت^{۳۸}، ۲۰۰۹). با افزایش فاکتور هموارساز مقدار ضریب همبستگی در داده‌های آموزش به آرامی کاهش ولی برای داده‌های آزمون به آرامی افزایش می‌یابد (مرادزاده و همکاران، ۱۳۹۰: ۵۸). طبق (جدول ۲) مقدار فاکتور هموارساز بهینه با توجه به مقدار ضریب همبستگی و میزان خطای مناسب در مرحله آموزش و مرحله آزمون، مقدار ۰/۳۵ انتخاب گردید.

نتایج آموزش مدل GRNN در پیش‌بینی غلظت ذرات معلق $PM_{2.5}$ در شهر سنندج در (شکل ۵) نشان داده شده است.

در این بخش در قسمت آموزش مقدار $R^2=0/81$, $RMSE=6/9468$ و $MAE=5/725$ بود.

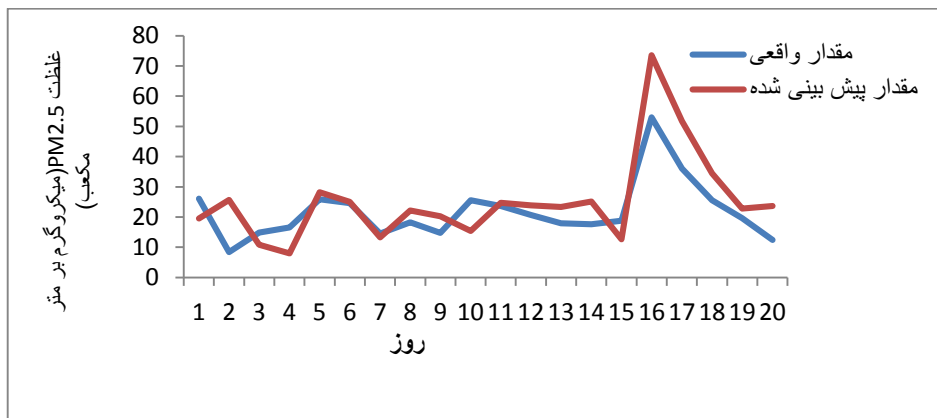


شکل ۵: نتایج مدل GRNN در مرحله آموزش

میزان RMSE در مدل ARIMA با $9/4137$ بیشترین و در GRNN با مقدار $6/9468$ کمترین مقدار، MAE در ARIMA با $6/767$ بیشترین و در GRNN با مقدار $5/0725$ کمترین را دارد. R^2 در مدل ARIMA کمترین و در GRNN بیشترین مقدار ($0/63$ و $0/81$ به ترتیب) را داراست.

- نتایج مرحله آزمون مدل MLR در پیش بینی ذرات معلق $PM_{2.5}$ برای شهر سنندج

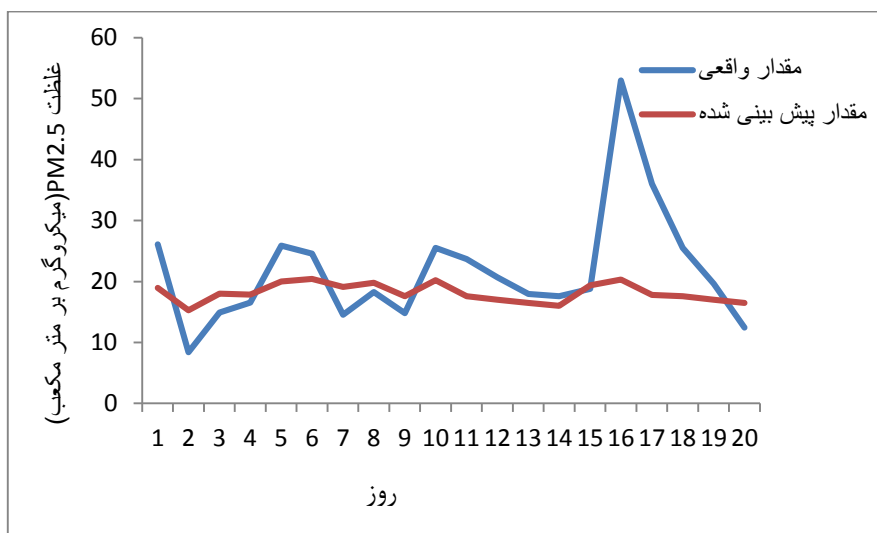
(شکل ۶) نتایج حاصل از مرحله آزمون مدل MLR را نشان می دهد. در این مرحله مقدار $R^2 = 0/70$ ، $RMSE = 8/9609$ و $MAE = 7/1385$ می باشد.



شکل ۶: نتایج مدل MLR در مرحله آزمون

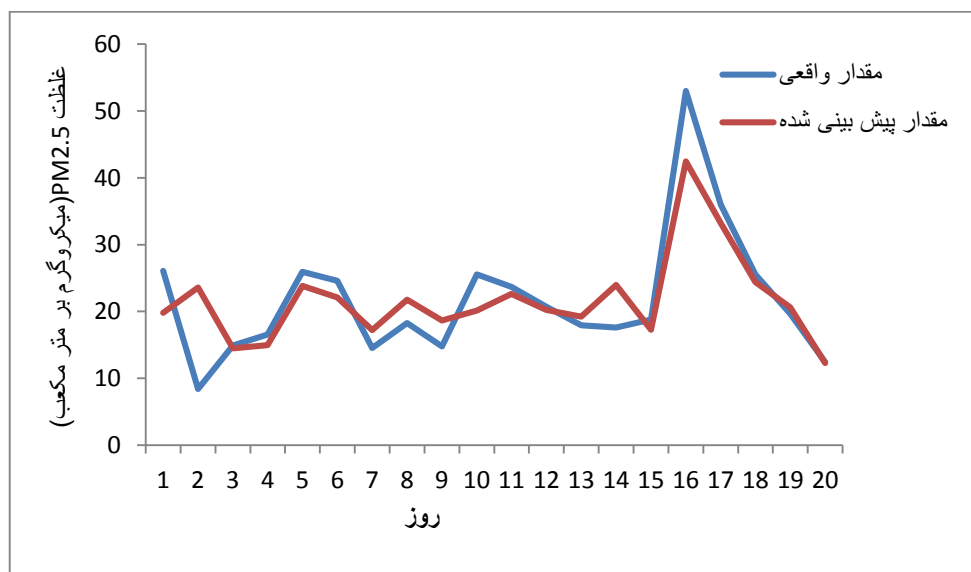
- نتایج مرحله آزمون مدل ARIMA در پیش بینی ذرات معلق $PM_{2.5}$ برای شهر سنندج

(شکل ۷) نتایج حاصل از مرحله آزمون مدل ARIMA را نشان می‌دهد. در این مرحله مقدار $R^2=0/30$ و $RMSE=11/3925$ و $MAE=9/0863$ می‌باشد.



شکل ۷: نتایج مدل ARIMA در مرحله آزمون مدل

- نتایج مرحله آزمون مدل GRNN در پیش‌بینی ذرات معلق $PM_{2.5}$ برای شهر (شکل ۸) نتایج حاصل از مرحله آزمون مدل GRNN را نشان می‌دهد. در این مرحله مقدار $R^2=0/74$ و $RMSE=5/0725$ و $MAE=3/4874$ می‌باشد.



شکل ۸: نتایج GRNN در مرحله آزمون مدل

مدل ARIMA بیشترین مقدار خطا ($MAE=3/4874$ ، $RMSE=5/0725$) را دارد. مقدار R^2 در ARIMA کمترین مقدار ($0/30$) و در GRNN بیشترین مقدار ($0/74$) می باشد.

یافته ها و بحث

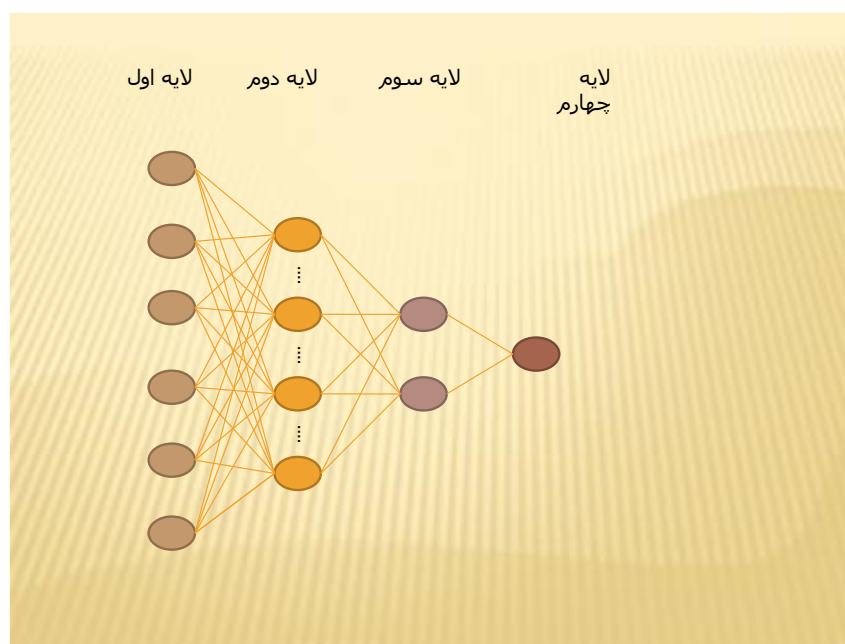
مدل رگرسیون چندگانه خطی (MLR)

مدل اول استفاده شده، رگرسیون چندگانه خطی (MLR) است که کاربرد زیادی در پیش بینی داشته و در مطالعات زیادی از آن استفاده شده است (فرناندز و همکاران^{۳۹}، ۲۰۱۲: ۳۵۲۸). در MLR از روش گام به گام استفاده گردید. بنابراین از متغیرهایی که دارای اثر معنی داری بر متغیر وابسته بود استفاده شد. از میان متغیرهای مستقل وارد شده در مدل، متغیرهای SO_2 ، PM_{10} ، $PM_{2.5}$ روز قبل، O_3 ، MAXT و WS در مدل باقی مانده و در نتیجه میزان غلظت این متغیرها بر میزان غلظت $PM_{2.5}$ شهر سنندج مؤثر بوده است. طبق معادله به دست آمده در MLR پارامترهای PM_{10} ، $PM_{2.5}$ روز قبل، O_3 و MAXT با ذرات معلق $PM_{2.5}$ شهر سنندج رابطه مثبت و پارامترهای SO_2 و WS با ذرات معلق $PM_{2.5}$ شهر سنندج رابطه منفی دارند. این مدل در مقایسه با مدل غیرخطی شبکه عصبی کارایی کمتری دارد به طوری که میزان ضریب تعیین (R^2) آن در مرحله آموزش $0/76$ و میزان RMSE و MAE ($7/3539$ ، $5/2555$) به ترتیب کمتر و بیش تر از مدل غیرخطی می باشند. این نتیجه با نتایج تیریاکیو همکاران (۲۰۱۴) و ولی و همکاران در سال ۱۳۹۰ همخوانی دارد. ولی و همکاران به مقایسه کارایی مدل های غیر خطی نظیر شبکه عصبی و MLR برای پیش بینی پرداختند. نتایج نشان داد که مدل غیرخطی با ضریب تعیین $0/86$ و مجذور میانگین مربعات خطای $1/815$ بهتر از مدل خطی چندگانه با ضریب تعیین $0/81$ با معیار خطای معادل $2/142$ است. استفاده از مدل MLR فرض های زیادی دارد که استفاده از آن در مسائل علمی مشکل است به طوری که در بسیاری از تحقیقات صورت گرفته، به دلیل این که، این فرض ها مباحث آماری پیچیده ای اند، از دید محققان مخفی مانده و به آن توجهی نشده است. در نتیجه مدل ارائه شده ای آن ها دارای دقت لازم نیست (نوری و همکاران، ۱۳۸۷: ۱۵۰).

- خود رگرسیون میانگین متحرک انباشته (ARIMA)

نتایج مدل ARIMA در پیش بینی ذرات معلق $PM_{2.5}$ در مقایسه با نتایج مدل های دیگر به کار برده شده ضعیف تر است و در رتبه آخر قرار دارد که به نظر می رسد پایین بودن میزان همبستگی به دلیل عدم رعایت ترتیب مشاهدات

می‌باشد. ARIMA با استفاده از داده‌های قبلی مقادیر آینده را پیش‌بینی می‌کند در حالی که در دیگر روش‌های مدل‌سازی اغلب با استفاده از متغیرهای مستقل دیگر به پیش‌بینی متغیر مورد نظر پرداخته می‌شود. معمولاً قدرت این مدل‌ها در پیش‌بینی کم‌تر است، اما به دلیل این‌که به اطلاعات جانبی کم‌تری نیاز دارد تمایل استفاده از آن زیاد است. طبیعت ذاتی یک سری زمانی، وابسته یا هم‌بسته بودن مشاهدات آن است و بنابراین ترتیب مشاهدات دارای اهمیت است. اگر داده‌ها مستقل و تصادفی باشند برای پیش‌بینی مناسب نخواهند بود و مقادیر گذشته تأثیری روی مقادیر حال و آینده نخواهد داشت. در سری زمانی غلظت PM_{2.5} مورد مطالعه، فقدان داده وجود داشت و به همین دلیل ترتیب مشاهدات رعایت نشده و ARIMA نتوانسته در پیش‌بینی عملکرد خوبی داشته باشد. این نتایج با نتایج سینگ و میشر^۹ همخوانی دارد. سینگ و میشر در سال ۲۰۱۵ از مدل‌های شبکه عصبی و ARIMA برای پیش‌بینی سری‌های زمانی مالی استفاده کردند، مدل ARIMA به دلیل بی‌نظمی داده‌ها نتایج ضعیف‌تری در پیش‌بینی سری‌های زمانی مالی نسبت به شبکه عصبی داشته است.

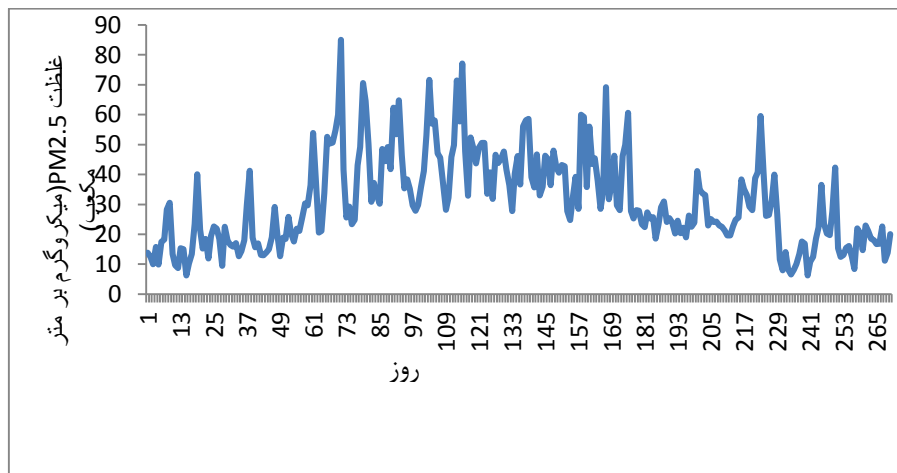


شکل ۹: ساختار GRNN استفاده شده

- شبکه عصبی رگرسیون عمومی (GRNN)

مدل GRNN مورد استفاده شده برای برآورد میزان ذرات معلق PM_{2.5} شهر سنندج در مقایسه با دیگر مدل‌ها در رتبه اول قرار دارد. این شبکه دارای یک لایه ورودی با شش نرون، یک لایه پنهان با ۲۷۰ نرون و تابع محرک شعاعی و یک لایه خروجی با یک نرون و تابع محرک خطی در آن است (شکل ۹). مزیت GRNN آموزش سریع و عملکرد

مناسب برای تعداد داده های کم است. با تغییر مقدار فاکتور هموارساز در GRNN مقدار خروجی تغییر کرده که می توان با آزمون و خطا بهترین مدل با مناسب ترین بازدهی را به دست آورد. مقدار فاکتور هموارساز با توجه به مقدار ضریب همبستگی و میزان خطای مناسب در هر دو مرحله آموزش و آزمون ۰/۳۵ انتخاب گردید. این شبکه در مقایسه با مدل های خطی عملکرد بهتری داشته و این چنین برمی آید که مدل GRNN در پیش بینی ذرات معلق PM_{2.5} شهر سنندج مناسب می باشد.



شکل ۱۰: غلظت سری زمانی PM_{2.5}

شبکه های عصبی برخلاف روش های خطی، دارای پیش فرض های بسیار کمی است و به عبارتی دیگر به صورت مدل آزاد هستند. این شبکه ها از مثال ها درس می آموزند و روابط تبعی میان داده ها را در صورت ناشناخته بودن روابط یا دشوار بودن توصیف آن ها کشف می کنند. از این رو در صورتی که داده های کافی وجود نداشته باشد این شبکه ابزار بسیار مفیدی است. شبکه های عصبی قابلیت تعمیم بسیار بالایی دارند و اگر داده ها دارای نویز باشند نسبت به دیگر مدل ها بهتر عمل می کنند (روبلس و همکاران^{۴۱}، ۲۰۰۸: ۸۳۳۲) و طبق نظر اسچاپ و همکاران^{۴۲} در سال ۱۹۹۸ در طراحی شبکه های عصبی نوعی خاص از معادلات لازم نیست و با ایجاد رابطه مناسب بین داده های ورودی و خروجی می توان به نتایج مناسب دست پیدا کرد. برتری دیگر شبکه عصبی در مقایسه با مدل های خطی حساس نبودن آن به وجود تعداد محدودی خطا در داده های آماری است. تنها محدودیت این شبکه ها در غیر قابل تفسیر بودن نتایج حاصل از مدل سازی از نظر میزان تأثیر هر متغیر ورودی بر خروجی مورد نظر است (نورانی و سیاح فرد، ۱۳۹۲: ۹۷). برتری شبکه عصبی نسبت به MLR و ARIMA توسط محققان زیادی ثابت شده است.

41- Robles et al

42- Schaap et al

تقوی و همکاران (۱۳۹۱) به مقایسه کارایی شبکه عصبی و MLR در پیش‌بینی غلظت روزانه منواکسیدکربن پرداختند. نتایج حاصل از اجرای این دو مدل نشان داد که مدل شبکه عصبی مصنوعی توانایی بیشتری نسبت به روش MLR در پیش‌بینی غلظت روزانه منواکسیدکربن دارد. به طوری که ضریب همبستگی و میانگین مجموع مربعات خطا در شبکه عصبی به ترتیب برابر با ۰/۸۱ و ۰/۰۶۹ و برای مدل رگرسیونی به ترتیب ۰/۶۱ و ۰/۱ به دست آمد. به نظر می‌رسد با توجه به سری زمانی غلظت $PM_{2.5}$ (شکل ۱۰) که دارای نوسانات زیادی است، مدل‌های خطی توانایی ارائه مدلی که بتواند این نوسانات را در نظر بگیرد ندارد و در این موارد استفاده از شبکه عصبی که قادر به پیش‌بینی روابط غیرخطی بین ورودی‌ها و خروجی‌ها باشد راه حل مناسبی برای جایگزین مدل‌های خطی می‌باشد.

نتیجه‌گیری

از آنجایی که سنندج به‌عنوان یکی از شهرهای آلوده‌ی کشور به حساب می‌آید، پیش‌بینی انواع آلودگی‌های آن به‌ویژه پیش‌بینی ذرات معلق $PM_{2.5}$ که عامل بسیاری از بیماری‌ها می‌باشد، می‌تواند در اعلام به موقع و قبل از افزایش میزان $PM_{2.5}$ کمک شایانی به سلامت جامعه کند. در ابتدا به بررسی رابطه بین تغییرات غلظت NO_2 , SO_2 , PM_{10} , CO , O_3 , $MaxT$, $MinT$, AP , PR , RH , WS و $PM_{2.5}$ روز قبل موجود در هوای شهر سنندج بر روی میزان غلظت $PM_{2.5}$ در شهر سنندج پرداخته شد. با توجه به روش گام‌به‌گام مورد استفاده تنها متغیرهای SO_2 , PM_{10} , $PM_{2.5}$ روز قبل، O_3 , $MaxT$ و WS در مدل باقی‌مانده و در نتیجه میزان غلظت این متغیرها بر میزان $PM_{2.5}$ شهر سنندج مؤثر بوده است. سپس به ارزیابی دقت مدل‌های MLR , $ARIMA$ و $GRNN$ در پیش‌بینی غلظت ذرات معلق $PM_{2.5}$ شهر سنندج پرداخته شد. در MLR مقادیر یک متغیر وابسته از روی مقادیر دو یا چند متغیر مستقل دیگر برآورد می‌شود. در $ARIMA$ مشاهدات گذشته یک متغیر وابسته از روی مقادیر دو یا چند متغیر مستقل بین مشاهدات و تعیین یک مدل توصیفی، مورد تجزیه و تحلیل قرار می‌گیرد. به منظور تخمین روابط رگرسیونی سری‌های زمانی، نوع خاصی از شبکه‌های عصبی به نام $GRNN$ مطرح شدند که قادرند با دقت بالایی پارامترهای روابط رگرسیونی حاکم بر سری‌های زمانی را تخمین زنند. نتایج نشان داد که مدل $GRNN$ با مقدار $R^2=0/81$ و $MAE=0/069$ در مرحله آموزش و مقدار $R^2=0/74$ و $RMSE=0/075$ و $MAE=0/0725$ در مرحله آزمون بهترین عملکرد را نسبت به مدل‌های خطی در پیش‌بینی ذرات معلق $PM_{2.5}$ در شهر سنندج داشته است.

منابع

- احمدی، ف؛ قهرمان، ب؛ داوری، ک؛ نیرومند، ح. ع (۱۳۸۶)، «کاربرد مدل سری های زمانی در پیش بینی بارندگی ماهیانه و سالانه مشهد»، *مجله علوم و صنایع کشاورزی ویژه خاک، آب و هوا*، شماره ۲، صص ۲۱-۳۳.
- بوداقپور، س؛ چرخستانی، ا (۱۳۹۰)، «پیش بینی میزان غلظت آلاینده های هوای تهران با استفاده از شبکه عصبی مصنوعی»، *علوم و تکنولوژی محیط زیست*، شماره ۱، صص ۱۰-۱.
- تقوی، ه؛ دانش، ش؛ مساعدی، ا (۱۳۹۱)، «مقایسه کارایی مدل های شبکه عصبی مصنوعی و رگرسیون خطی در پیش بینی غلظت روزانه منواکسیدکربن براساس پارامترهای هواشناسی»، *همایش ملی جریان و آلودگی هوا*، ۲۵-۱۴ آبان ۱۳۸۹، دانشگاه تهران، ایران.
- خاشعی، م؛ بیجاری، م (۱۳۸۷)، «بهبود عملکرد پیش بینی های مالی با ترکیب مدل های خطی و غیرخطی خودرگرسیون میانگین متحرک انباشته و شبکه های عصبی مصنوعی»، *فصلنامه پژوهش های اقتصادی*، شماره ۸، صص ۸۳-۱۰۰.
- خاشعی، م؛ بیجاری، م؛ مخاطب رفیعی، ف (۱۳۹۲)، «پیش بینی نرخ ارز با به کارگیری مدل های ترکیبی پرسپترون های چند لایه (MLPs) و طبقه بندی کننده های عصبی احتمالی (PNNs)»، *روش های عددی در مهندسی*، شماره ۳۲، صص ۱-۱۴.
- خرمی، م؛ بزرگ نیا، ا (۱۳۸۶)، «تجزیه و تحلیل سری های زمانی با نرم افزار MINITAB 14»، تهران، انتشارات سخن گستر.
- زراءنژاد، م؛ خداپناه، م؛ کیانی، پ؛ ابراهیمی، ص (۱۳۹۲)، «ارزیابی و مقایسه عملکرد مدل رگرسیون خودبازگشتی میانگین متحرک انباشته فازی و شبکه عصبی فازی در پیش بینی رشد اقتصادی ایران»، *فصلنامه مطالعات اقتصادی کاربردی ایران*، شماره ۸، صص ۵۱-۳۳.
- سرمیدیان، ف؛ تقی زاده، ر؛ مهرجردی، ح؛ عسگری، م؛ اکبرزاده، ع (۱۳۸۹)، «مقایسه روش های نروفازی، شبکه عصبی و رگرسیون چندمتغیره در پیش بینی برخی خصوصیات خاک (مطالعه موردی: استان گلستان)»، *مجله تحقیقات آب و خاک ایران*، شماره ۴۱، صص ۲۲۰-۲۱۱.

- صالحی، ص؛ قائمی‌اصل، ز (۱۳۹۲)، «بررسی رابطه آموزش زیست‌محیطی و رفتارهای حفاظت از محیط‌زیست (مورد مطالعه: دانش‌آموزان دبیرستان‌های دخترانه شهر بابل)»، *فصلنامه آموزش محیط‌زیست و توسعه پایدار*، شماره ۳، صص ۷۹-۶۷.
- صدرموسوی، م؛ رحیمی، ا (۱۳۸۸)، «مقایسه نتایج شبکه‌های عصبی پرسپترون چند لایه با رگرسیون چندگانه در پیش‌بینی غلظت ازن در شهر تبریز»، *پژوهش‌های جغرافیایی طبیعی*، شماره ۷۱، صص ۷۲-۶۵.
- عزیزی‌فر، م؛ نادافی، ک؛ محمدیان، م؛ صفدری، م؛ خزایی، م (۱۳۹۰)، «بررسی شاخص کیفیت هوا و غلظت ذرات معلق با قطر آئرودینامیکی در هوای شهر قم»، *مجله دانشگاه علوم پزشکی قم*، شماره ۲، صص ۶۳-۵۹.
- قوامی، س (۱۳۹۲)، «پیش‌بینی وقوع پدیده‌های گردوغبار شهر سنج» پایان‌نامه کارشناسی ارشد، دانشکده منابع طبیعی، دانشگاه کردستان.
- کهنسال، م. ر؛ پرمه، ز (۱۳۹۱)، «پیش‌بینی قیمت تخم‌مرغ با استفاده از ARIMA، شبکه عصبی مصنوعی و هموارسازی هالت-وینترز»، *فصلنامه پژوهشنامه بازرگانی*، شماره ۶۲، صفحات ۷۲-۴۹.
- مرادزاده، ع؛ دولتی‌ارده‌جانی، ف؛ روکی، ر؛ رحیمی، م (۱۳۹۰)، «برآورد نفوذپذیری میدان گازی پارس جنوبی با استفاده از شبکه عصبی رگرسیون عمومی»، *مجله فیزیک زمین و فضا*، شماره ۳۷، صص ۶۵-۵۱.
- منصوری، ف؛ خانجانی، ن؛ راننده‌کلانکس، ل؛ پورموسوی، ر (۱۳۹۲)، «بررسی و پیش‌بینی وضع آلاینده‌های هوای شهر کرمان با مدل سری‌های زمانی»، *مجله دانشکده بهداشت و انستیتو تحقیقات بهداشتی*، شماره ۱۱، صص ۸۶-۷۵.
- نورانی، و؛ سیاح‌فرد، م (۱۳۹۲)، «آنالیز حساسیت داه‌های ورودی به شبکه عصبی مصنوعی به منظور برآورد مقدار تبخیر روزانه»، *مجله آب و فاضلاب*، شماره ۳، صص ۱۰۰-۸۰.
- نوری، ر. ا؛ اشرفی، خ؛ اژدرپور، ا (۱۳۸۷)، «مقایسه کاربرد روش‌های شبکه عصبی مصنوعی و رگرسیون خطی چندمتغیره براساس تحلیل مولفه‌های اصلی برای پیش‌بینی غلظت میانگین کربن‌منواکسید: بررسی موردی: شهر تهران»، *مجله فیزیک زمین و فضا*، شماره ۳۴، صص ۱۵۲-۱۳۵.
- ولی، ع. ع؛ رامشت، م. ح؛ سیف، ع؛ قضاوی، ر (۱۳۹۰)، «مقایسه کارایی مدل‌های شبکه عصبی مصنوعی و رگرسیون برای پیش‌بینی بار رسوب جریان مطالعه موردی حوضه آبخیز سمندگان»، *مجله جغرافیا و برنامه‌ریزی محیطی*، شماره ۴۴، صفحات ۳۴-۱۹.

- Adamson, I., Prieditis, H., Vincent, R., (1999), "Pulmonary toxicity of an atmospheric particulate sample is due to the soluble fraction", *Toxicology and Applied Pharmacology*, 157: 43–50.
- Baldasano, V. E., Jiménez, P., (2003), "Air quality data from large cities", *Sci Total Environ*, 307: 141-65.
- Barcenas, O., Olivas, E., Guerrero, J. D., Valls, G., Rodriguez, C., Tascon, S., (2005), "Unbiased sensitivity analysis and pruning techniques in neural networks for surface ozone modeling", *Ecol Model*, 182: 149–58.
- Basurko, E., Berastegi, G., Madariaga, I., (2006), "Regression and multilayer perceptron-based models to forecast hourly O₃ and NO₂ levels in the Bilbao area", *Environ Model Softw*, 21: 430–46.
- Cai, M., Yin, Y., Xie, M., (2009), "Prediction of hourly air pollutant concentrations near urban arterials using artificial neural network approach", *Transportation Research Part D*, 14: 32-4.
- Cobourn, W. G., (2010), "An enhanced PM_{2.5} air quality forecast model based on nonlinear regression and back-trajectory concentrations", *Atmos Environ*, 44: 3015–3023.
- Chaloulakou, A., Grivas, G., Spyrellis, N., (2003), "Neural network and multiple regression model for PM₁₀ prediction in Athens: A comparative assessment", *Journal of the Air & Waste Management Association*, 53 (10): 1183-1190.
- Demuth, H., Beale, M., (2002), "Neural network toolbox for use with matlab", users guide version 4. [online]: https://www.mathworks.com/help/pdf_doc/nnet/nnet Ug.pdf
- Elkamel, A., Abdul-Wahab, S., Bouhamra, W., Alper, E., (2001), "Measurement and prediction of ozone levels around a heavily industrialized area: A neural network approach", *Advances in Environmental Research*, 5: 47–59.
- Fernandez, F. G., Palacios, P., Esteban, L. G., Garcia-Iruela, A., Rodrigo, B. G., Menasalvas, E., (2012), "Prediction of MOR and MOE of structural plywood board using an artificial neural network and comparison with a multivariate regression model", *Composites Part B*, 43: 3528– 3533.
- Gheyas, I. A., Smith, L. S., (2009), "A neural network approach to time series forecasting", proceeding of the world congress on engineering, July 1-3, London, U.K. pp 1-5.
- Gupta, P., Sundar, C. A., (2008), "An evaluation of Terra-MODIS sampling for monthly and annual particulate matter air quality assessment over the Southeastern United States", *Atmospheric Environment*, 42: 6465-6471.
- Hassan. J., (2014), "ARIMA and regression models for prediction of daily and monthly clearness index", *Renewable Energy*, 68: 421- 427.
- Kaufman, Y. J., Tanre, D., Boucher, O., (2002), "A satellite view of aerosols in climate systems", *Nature*, 419: 215–223.
- Kuzmanovski, I., Aleksovska, S., (2003), "Optimization of artificial neural networks for prediction of the unit cell parameters in orthorhombic perovskites; Comparison with multiple linear regression", *Chemometrics and Intelligent Laboratory Systems*, 67: 167– 174.

- Legates, D. R., McCabe, G. J., (1999), "Evaluating the use of "Goodness-of-fit" Measures in hydrologic and hydroclimatic model validation", *Water Resour. Res.*, 35: 233-241.
- Leung, M. T., Daoock, H., Chen, A., (2000), "Acomparison of classification and level estimation models", *International Journal of Forecasting*, 16: 173-190.
- Nasrollhi, Z., Ghaffari Goolak, M., (2010), "Air pollution and its effective factors", *Faslnameh Pazhooresh Eghtesadi*, 3: 375-395.
- Pastuszka, J. S., Kozłowska, W. R., Zubek, E, Z., (2010), "Characteriazation of PM₁₀ and PM_{2.5} and associated heavy metals at the crossroads and urban background site in Zabrze, upper Silesia, Poland, during the smog episodes", *Environmntal Monitoring and Assessment*, 168: 613- 627.
- Pérez, P., Trier, A., Reyes, J., (2000), "Prediction of PM_{2.5} concentrations several hours in advance using neural networks in Santiago, Chile", *Atmos Environ*, 34: 1189-96.
- Robles, L. A., Ortega, J. C., Fu, J. S., Reed, G. D., Chow, J. C., Watson, J. G., Herrera, J. A., (2008), "A hybrid ARIMA and artificial neural networks model to forecast particulate matter in urban areas: The case of Temuco, Chile", *Atmospheric Environment*, 42: 8331- 8340.
- Sandstrom, T.FB., (2008), "Desert dust, an unrecognized source of dangerous air pollution", *Epidemiology*, 19: 808-9.
- Schaap, M. G., (1998), "Using neural network to predict soil water retention and soil hydraulic conductivity", *Soil and Till Research*, 47: 37- 42.
- Singh, A., Mishra, G. C., (2015), "Application of box- jenkins method and artificial neural network procedure for time series forecasting of prices", *Statistics in Transition new series*, 16 (1): 83- 96.
- Specht, D. F., (1991), "A general regression neural network", *IEEE TransnNeural Networks*, 2: 568-76.
- Timmreck, C., Lorenz, M. P., Crowley, T. J., Kinne, S., Raddatz, T. J., Thomas, M., A., (2009), "Limited temprature response to the very large AD 1258 volcanic eruption", *Geographical Research Letters*, 36: 14-29.
- Tiryaki, S., Özşahin, Ş., Yıldırım, I., (2014), "Comparison ofartificial neural network and multiple linear regression models to predict optimum bonding strength of heat treated woods", *International Journal of Adhesion and Adhesives*, 55: 29- 36.
- Zhou, Q., Jiang, H., Wang, J., Zhou, J., (2014), "A hybridmodel for PM_{2.5} forecasting based on ensemble empirical mode decomposition and a general regression neural nrtwork", *Science of the Total Environment*, 496: 264-274.