

ارزیابی دقت مدل‌های MLP، PCR، ARIMA و پیش‌بینی عمق نوری هواویزها

ساجده اکبری^۱، جمیل امان‌اللهی^{۲*}، محمد دارند^۳

sajdehakbari0524@yahoo.com

۱. کارشناسی‌ارشد محیط‌زیست، دانشکده منابع طبیعی، دانشگاه کردستان

۲. استادیار، گروه محیط‌زیست، دانشکده منابع طبیعی، دانشگاه کردستان.

m.darand@uok.ac.ir

۳. دانشیار، گروه آب‌وهوا شناسی، دانشکده منابع طبیعی، دانشگاه کردستان

تاریخ پذیرش مقاله: ۱۳۹۷/۰۹/۱۲

تاریخ وصول مقاله: ۱۳۹۷/۰۲/۱۲

چکیده

برآورد عمق نوری هواویزها (AOD) برای بررسی میزان ذرات معلق موجود در جو که یکی از آلاینده‌های هوا است استفاده می‌شود. در این پژوهش برای برآورد عمق نوری هواویزها در ایستگاه‌های فاقد تشعشع سنج یا برآورد یک‌ساله (اتورگرسیو) در ایستگاه‌های دارای تشعشع سنج از مدل‌های مختلف همچون مدل‌های رگرسیون چندگانه (MLR)، رگرسیون مؤلفه‌های مبنا (PCR)، خودرگرسیون میانگین متحرک انباشته (ARIMA) و نیز مدل شبکه عصبی مصنوعی (MLP)، استفاده شد. بدین منظور داده‌های خودرگرسیون نسبی، سرعت باد و ارتفاع لایه اتمسفری اخذ شده از پایگاه داده جهانی ECMWF در تراز ۸۵۰ هکتوپاسکال به‌عنوان متغیرهای مستقل و همچنین داده‌های تشعشع سنج خورشیدی اداره هواشناسی شهرستان سنندج در بازه زمانی ۲۰۰۵/۱/۱ تا ۲۰۱۶/۱۲/۳۱ به‌عنوان متغیر وابسته در نظر گرفته شدند. نتایج نشان داد که مدل ARIMA با دارا بودن مقادیر عددی $R^2=0/91$ ، $MAE=0/0501$ و $RMSE=0/033$ در مرحله آموزش مدل و نیز مقادیر $R^2=0/89$ ، $MAE=0/0374$ و $RMSE=0/0586$ در مرحله آزمون مدل دارای بهترین عملکرد در برآورد عمق نوری هواویزها در ایستگاه‌های فاقد تشعشع سنج است. همچنین نتایج مرحله اتورگرسیو نشان داد که مدل MLP با دارا بودن مقادیر عددی $R^2=0/96$ ، $MAE=0/028$ و $RMSE=0/0483$ بالاترین دقت را از میان مدل‌های فوق در برآورد عمق نوری هواویزها برای سال ۲۰۱۷ داشته است.

کلیدواژه

پایگاه داده ECMWF، پیش‌بینی، خودرگرسیون میانگین متحرک انباشته، شبکه عصبی مصنوعی، عمق نوری هواویزها.

۱. سرآغاز

ذرات معلق موجود در هوا نقش بسیار مهمی را در تابش انتقالی و فرایندهای شیمیایی ایفا می‌کنند که کنترل‌کننده اقلیم زمین هستند (Dickson, 1995; IPCC, 2007) و اثر منفی بر زندگی و سلامت انسان دارند (دانش جعفری و همکاران، ۱۳۹۴). آتروسول‌ها در لایه‌های پایینی جو یافت شده و قادرند مسافت‌های طولانی را به‌وسیله باد طی کنند، بنابراین اثر آن‌ها بر جو، بیانگر تغییرپذیری قوی در شرایط زمانی و مکانی است. فرایند اسکن و بررسی کردن آسمان

هواویزهای اتمسفری دارای منابع مختلفی هستند از جمله این منابع می‌توان به مواردی همچون فعالیت‌های آتش‌فشانی، ذرات گردوغبار، ذرات نمک موجود در دریاها و اقیانوس‌ها و ... اشاره کرد یا می‌توانند ناشی از فعالیت‌های انسانی باشند که می‌توان به فعالیت‌هایی نظیر فعالیت‌های صنعتی، حمل‌ونقل، مصارف سوختی و ... اشاره کرد (Seinfeld & Pandis, 1998). آتروسول‌ها یا همان

استفاده از مدل شبکه عصبی مصنوعی یا همان MLP صورت گرفته است. بدین ترتیب اولسس و همکاران به منظور سنجش و تعیین مقادیر از دست رفته عمق نوری هواویزها بر اساس شبکه‌های عصبی مصنوعی، به ارائه یک روش پرداختند. آن‌ها توانستند که مقادیر عمق نوری هواویزها را تحت شرایط ابری و نیز در موقعیت‌هایی تعیین کنند که داده‌های اندکی در دسترس است. برای رسیدن به این هدف، از مقادیر قبلی عمق نوری هواویزها در دو ایستگاه، مسیرهایی که توده هوا از آن عبور می‌کند و محاسبات شبکه عصبی استفاده کردند. این روش به منظور پیش‌بینی مقادیر عمق نوری هواویزها روی ۹ ایستگاه با طول موج 440nm در سواحل شرقی آمریکا و در طول سال‌های ۲۰۱۲_۱۹۹۹ اعمال شد و ضریب تعیین R^2 بین مقادیر محاسبه شده و مقادیر تخمین زده شده عمق نوری هواویزها به میزان $0/85$ حاصل شد که بیانگر عملکرد خوب این مدل بود. بنابراین آن‌ها به این نتیجه رسیدند که این روش می‌تواند به عنوان ابزاری قوی در پیش‌بینی مقادیر عمق نوری هواویزها در شرایط مختلف جوی و در طول موج‌های مختلف نیز به کار برده شود. همچنین می‌توان به کار تانجا^۳ و همکاران اشاره کرد. تانجا و همکاران در سال ۲۰۱۶ از مدل خودرگرسیون میانگین متحرک انباشته^۴ (ARIMA) به منظور شبیه‌سازی میانگین ماهیانه عمق نوری هواویزها استفاده کردند. آن‌ها میزان ضرایب تعیین به دست آمده را $0/53$ و $0/674$ گزارش دادند. تاکنون در ایران پژوهشی به منظور برآورد عمق نوری هواویزها با استفاده از مدل‌های مختلف همچون مدل‌های رگرسیون چندگانه، رگرسیون مؤلفه‌های مبنای^۵ (PCR)، شبکه عصبی مصنوعی و نیز مدل خودرگرسیون میانگین متحرک انباشته، صورت نگرفته است. از این رو در پژوهش حاضر برآورد عمق نوری هواویزها در دو حالت شامل برآورد برای مناطق فاقد ایستگاه تشعشع سنج خورشیدی و برآورد طولانی مدت برای آینده در ایستگاه‌های دارای تشعشع سنج خورشیدی آزموده شد.

با استفاده از رادیومترها و تابش‌سنج‌های طیفی صورت می‌گیرد. رادیومترها عمل اندازه‌گیری تشعشع طیفی آسمان را در فواصل زاویه‌ای شناخته‌شده‌ای نسبت به خورشید انجام می‌دهند و آن را به آئروسول‌ها بسط می‌دهند. یکی از این منابع مهم شبکه AERONET است (Holben et al., 1998) که شامل شبکه بین‌المللی متحد و همگون از رادیومترها است و به طور گسترده‌ای برای بازیابی و جمع‌آوری اطلاعات در خصوص خواص نوری آئروسول محلی و نیز به منظور بررسی هواویزها و آلاینده‌های اتمسفری همچون ذرات معلق، معمولاً از پارامتری خاص تحت عنوان عمق نوری هواویزها^۱ (AOD) استفاده می‌شود (Rahul et al., 2014) که نمایانگر میزان عبوردهی پرتوهای نوری در اتمسفر است و بیان‌کننده نقش هواویزها در جذب و پراکنده‌سازی این پرتوهای نوری است و کمیتی بدون بعد است (Wmo, 1995). این شبکه قادر است عمل اندازه‌گیری رادیانس‌های ۸ باند طیفی مربوط به محدوده مرئی و مادون قرمز نزدیک را انجام دهد که در طول موج‌های $0/34$ تا $1/20$ میکرون واقع شده‌اند (Holben et al., 1998). از آنجایی که شهرستان سنندج یکی از شهرهای آلوده کشور شناخته شده است، بنابراین پیش‌بینی انواع آلودگی‌ها به‌ویژه ذرات معلق موجود در هوا اهمیت فراوانی دارد. یکی از روش‌های خطی مورد استفاده در پیش‌بینی انواع آلودگی‌های هوا، مدل رگرسیونی چندگانه است که به منظور بررسی ارتباطات خطی موجود میان یک متغیر وابسته در ارتباط با یک یا چندین متغیر مستقل می‌پردازد. (Kuzmanovski & Aleksovsk, 2003). از دیگر روش‌های پیش‌بینی می‌توان به روش شبکه‌های عصبی اشاره کرد که به منظور بررسی داده‌هایی که دارای روند غیرخطی هستند، دارای کارایی بالایی خواهند بود (حجتی و همکاران، ۱۳۹۵). همچنین می‌توان با استفاده از متغیر وابسته موجود، عمل پیش‌بینی یا همان اتورگرسیون در مدل‌های مذکور را انجام داد. از جمله کارهای انجام شده در سطح بین‌المللی در این زمینه مربوط به اولسس^۲ و همکاران در سال ۲۰۱۵ است که بر پایه

فاقد تشعشع سنج پیشنهاد داده شد بدین صورت که در این ایستگاه‌ها فقط از داده‌های مستقل به‌عنوان ورودی مدل ساخته شده بر اساس داده‌های ایستگاه سنندج استفاده شود و خروجی آن‌که همان مقادیر عمق نوری هواویزها است حاصل شود. در برآورد عمق نوری هواویزها برای یک سال بعد در ایستگاه‌های دارای تشعشع سنج از داده‌های وابسته (عمق نوری هواویزها) در بازه زمانی ۱۱ ساله ذکر شده به‌عنوان ورودی مدل استفاده شد و در مرحله اتورگرسیون مدل‌ها بر اساس این داده‌ها، عمق نوری هواویزها برای یک سال بعد پیش‌بینی شد.

مدل رگرسیون چندگانه^۷ (MLR)

در مدل MLR به بررسی رابطه موجود بین یک متغیر وابسته و تعدادی متغیر مستقل از طریق رابطه منتج شده برای آن‌ها در محیط نرم‌افزاری SPSS پرداخته می‌شود (McBratney et al., 2002). محققان زیادی از مدل MLR برای بررسی آلودگی‌های محیط‌زیست استفاده کردند (نادری و همکاران، ۱۳۹۶: احسان‌زاده و همکاران، ۱۳۹۵)، فرمول کلی مدل MLR به‌صورت رابطه ۱ است (Tiryaki et al., 2014).

$$Y = \beta_0 + \beta_1 x_1 + \dots + \beta_n x_n + \varepsilon \quad (1)$$

در این رابطه منظور از Y همان متغیر وابسته است. X_1, \dots, X_n بیانگر متغیرهای مستقل هستند، همچنین β_0, \dots, β_n ثابت‌های معلوم را گزارش می‌دهند. ε نیز نشان‌دهنده مقادیر باقیمانده است.

مدل رگرسیون مؤلفه‌های مبنا (PCR)

مدل PCR، ترکیبی از آنالیز اجزای اصلی^۸ (PCA) و MLR است. این محاسبات به‌صورت رابطه ۲ است:

$$Y = \varphi \beta_{PCR} + e \quad (2)$$

منظور از φ ماتریس مؤلفه‌های مبنا است که به‌صورت $n \times k$ به دست می‌آید و β_{PCR} نشان‌دهنده نخستین عدد از

۲. مواد و روش‌ها

منطقه مورد مطالعه

شهرستان سنندج مرکز استان کردستان است. به لحاظ موقعیت جغرافیایی این شهرستان، در محدوده ۳۵ درجه و ۲۰ دقیقه عرض شمالی و ۴۷ درجه طول شرقی از نصف‌النهار گرینویچ و در ارتفاع ۱۳۷۳/۴ متری از سطح دریا واقع شده است. متوسط دمای سالانه سنندج ۱۴/۳ درجه سانتی‌گراد است و نرخ تابش کلی سالانه سنندج برابر با ۳۹۳ کالری بر سانتی‌متر مربع در روز است (خلیلی و رضایی صدر، ۱۳۷۶).

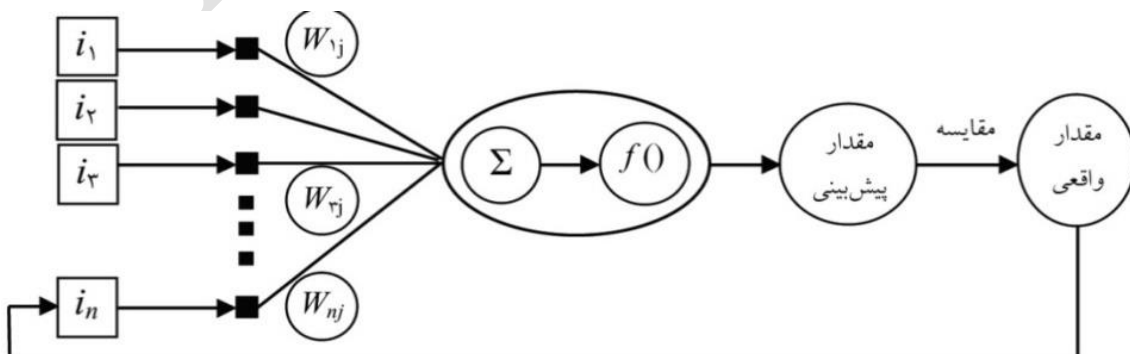
داده‌ها

در این مطالعه داده‌های مربوط به تشعشع سنج خورشیدی در بازه زمانی ۲۰۰۵/۰۱/۰۱ تا ۲۰۱۶/۱۲/۳۱ برای منطقه مورد مطالعه از طریق اداره هواشناسی در مرکز استان فوق جمع‌آوری شد. بدین ترتیب تعداد کل داده‌های موجود برای بازه زمانی ذکر شده، ۴۰۱۷ داده در منطقه مورد مطالعه بوده است و چون برای برخی از روزهای سال تشعشع سنج خورشیدی فاقد داده بوده است، بنابراین تعداد کل داده‌های مورد استفاده برای شهرستان سنندج، به ۳۹۵۶ تقلیل یافت. همچنین داده‌های مربوط به متغیرهای هواشناسی نظیر دما، رطوبت نسبی، سرعت باد و ارتفاع لایه اتمسفری در سطح تراز ۸۵۰ هکتوپاسکال، طی بازه زمانی ذکر شده و نیز با توجه به زمان‌هایی از سال که تشعشع سنج خورشیدی قادر به ارائه داده بوده است، تهیه شد. به‌منظور دستیابی به این داده‌ها از داده‌های نسخه ERA-Interim پایگاه داده جهانی مرکز پیش‌بینی میان مدت جوی اروپایی^۹ (ECMWF) استفاده شد. برای برآورد عمق نوری هواویزها در مناطق فاقد ایستگاه تشعشع سنج داده‌های مستقل و داده وابسته مربوط به شهرستان سنندج در مدل‌های ذکر شده مورد آموزش و آزمون قرار گرفتند و بعد از مقایسه نتایج، بهترین مدل برای استفاده در مناطق

ARIMA قادر است که ویژگی‌ها و روندها تغییرات در متغیرهای موجود در سری‌های زمانی را پیدا کند و از این طریق پیش‌بینی‌های مربوط به سری‌های زمانی را ارائه دهد. این مدل روش پیش‌بینی با استفاده از نظریه‌های آماری است و به دلیل دارا بودن مزایایی نظیر دقت بالا و نیز قابلیت انطباق‌پذیری قوی، قادر است، در بسیاری از زمینه‌ها دارای کاربرد گسترده‌ای باشد (Soni et al., 2014; Díaz-Robles et al., 2008).

مدل شبکه‌های عصبی مصنوعی^۱ (ANN)

شبکه پرسپترون چند لایه^۱ (MLP)، شناخته‌شده‌ترین و اغلب مورد استفاده‌ترین نوع در میان انواع مختلف شبکه‌های عصبی هستند و در بیشتر موارد به‌عنوان سیگنال‌هایی عمل می‌کنند که کار انتقال را از ورودی‌ها به خروجی‌ها، در داخل شبکه بر عهده دارند. در این نوع از شبکه‌های چند لایه، لایه‌ها به ترتیب به یکدیگر متصل‌اند، به‌نحوی که خروجی‌های حاصل شده از لایه اول به‌عنوان ورودی‌های لایه دوم عمل می‌کنند و خروجی‌های به‌دست‌آمده از لایه دوم نیز به‌عنوان ورودی‌های لایه سوم در نظر گرفته می‌شوند و این روال تا زمانی که خروجی‌های لایه آخر به‌عنوان خروجی‌های اصلی و پاسخ‌های قطعی و واقعی، مطرح هستند، ادامه می‌یابد.



شکل ۱. ساختار MLP

نمره مؤلفه‌های K است. منظور از بردار e نیز که به‌صورت $n \times 1$ تعریف می‌شود، خطاهای تصادفی است. نمرات و امتیازهای مؤلفه‌های مبنا بر اساس نسخه اصلی از روش OLS به دست می‌آید که به‌صورت رابطه ۳ است:

(۳)

$$\beta_{PCR} = (\varphi' \varphi)^{-1} \varphi' y = (L^2)^{-1} \varphi' y$$

در اینجا منظور از L^2 میزان اریب ماتریس است که بر اساس پارامتر K^{th} حاصل می‌شود و آن نیز نمایانگر میزان اریب پارامتر λ_k است (K^{th} بزرگ‌ترین ارزش و مقدار ویژه از $X'X$ است). برآورد دامنه ضرایب رگرسیونی مورد استفاده در مدل PCR مشابه با روش OLS است. که از رابطه ۴ به دست می‌آید (Jeffers, 1967; Krzanowski & Marriott, 1994)

$$\beta_{PCR} = \sum_{k=1}^K \frac{v_k u_k'}{d_k} y, \quad K < \min(n, p) \quad (۴)$$

مدل خودرگرسیون میانگین متحرک انباشته (ARIMA)

مدل ARIMA، یکی از روش‌های مهم در پیش‌بینی سری‌های زمانی است که به‌وسیله باکس و جنکینز در سال ۱۹۷۰ ارائه شد (Zhang et al., 2013). مدل ARIMA مدلی داده محور است و این بدان معناست که مدل مذکور از ساختار خود داده‌ها استفاده می‌کند و اگر داده‌ها دارای هر گونه روابط غیرخطی معنادار باشند، این مدل، با محدودیت مواجه خواهد شد (Pai & Lin., 2005). مدل

۴. نتایج

نتایج حاصل از مدل MLR در مرحله‌های آموزش و آزمون در سطح تراز ۸۵۰ هکتوپاسکال در شهرستان سنندج نتایج حاصل از شبیه‌سازی MLR در سطح تراز ۸۵۰ هکتوپاسکال در مرحله آموزش مدل مذکور در شکل (۲-۲) الف) آورده شده است. معادله ارائه شده برای آن در رابطه ۸ آمده است:

(۸)

$$AOD = 0.458 + 0.397T_{850} -$$

$$0.127RH_{850} + 0.21Speed_{850} - 0.064BLH$$

در این قسمت مقادیر $R^2 = 0.071$ ، $RMSE = 0.1698$ ، $MAE = 0.1498$ حاصل شد. نتایج حاصل از مرحله آزمون این مدل نیز در شکل (۲-۲) ب) نشان داده شده است که در این مرحله مقادیر $R^2 = 0.096$ ، $RMSE = 0.1703$ ، $MAE = 0.1494$ به دست آمد.

نتایج مرحله‌های آموزش و آزمون مدل MLR نشان‌دهنده دقت ضعیف این مدل در پیش‌بینی عمق نوری هواویزها در شهرستان سنندج است که عدم همخوانی روند تغییرات داده‌های موجود و داده‌های پیش‌بینی شده در شکل‌های (۲-الف) و (۲-ب) نیز بیانگر این موضوع است. از این رو نتایج نشان می‌دهد که مدل MLR، دارای دقت لازم برای پیش‌بینی مقادیر هواویزها در مناطق دارای دستگاه تشعشع سنج نیست.

پرسپترون از یک لایه ورودی، یک لایه خروجی و چندین لایه میانی تشکیل شده است. هر نرون حاصلضرب ورودی‌ها و وزن‌ها را با مقداری بایاس جمع زده و خروجی باعث فعالیت یک تابع انتقال می‌شود (رابطه ۵).

$$y_k = \varphi\left(\sum_{i=1}^n (w_{kj}x_j + \theta_k)\right) \quad (5)$$

در رابطه فوق x_j ، ورودی زام به نرون؛ k ام (که ممکن است ورودی شبکه یا خروجی لایه قبل باشد) است. w_{kj} ، شدت ارتباطات بین این نرون و ورودی زام است. y_k ، خروجی شبکه و θ_k بایاس ثابت است.

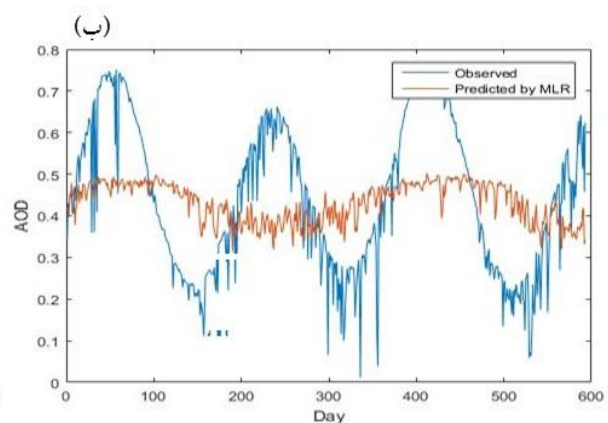
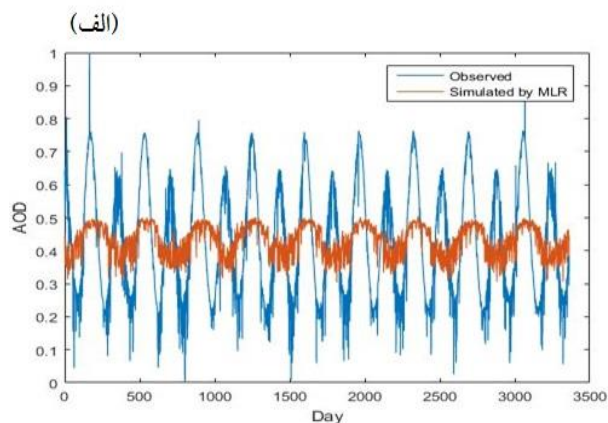
۳. نرمال‌سازی داده‌ها

قبل از استفاده از داده‌های ورودی در مدل‌های مذکور بایستی عمل استانداردسازی داده‌ها صورت گیرد. بدین ترتیب با استفاده از عمل نرمال‌سازی، داده‌ها محدوده‌ای عددی بین صفر تا یک را به خود اختصاص خواهند داد. به منظور نرمال‌سازی داده‌ها می‌توان از روابط ۶ و ۷ استفاده کرد: (دستورانی و همکاران، ۱۳۹۱).

$$X_n = 2 * \frac{X - X_{min}}{X_{max} - X_{min}} - 1 \quad (6)$$

$$X_n = \frac{X - X_{min}}{X_{max} - X_{min}} \quad (7)$$

در این روابط منظور از x_n داده نرمال شده است. x بیانگر داده مشاهده شده است. x_{min} و x_{max} نیز به ترتیب مقادیر حداکثر و حداقل از داده‌های مشاهده شده را به خود اختصاص می‌دهند.



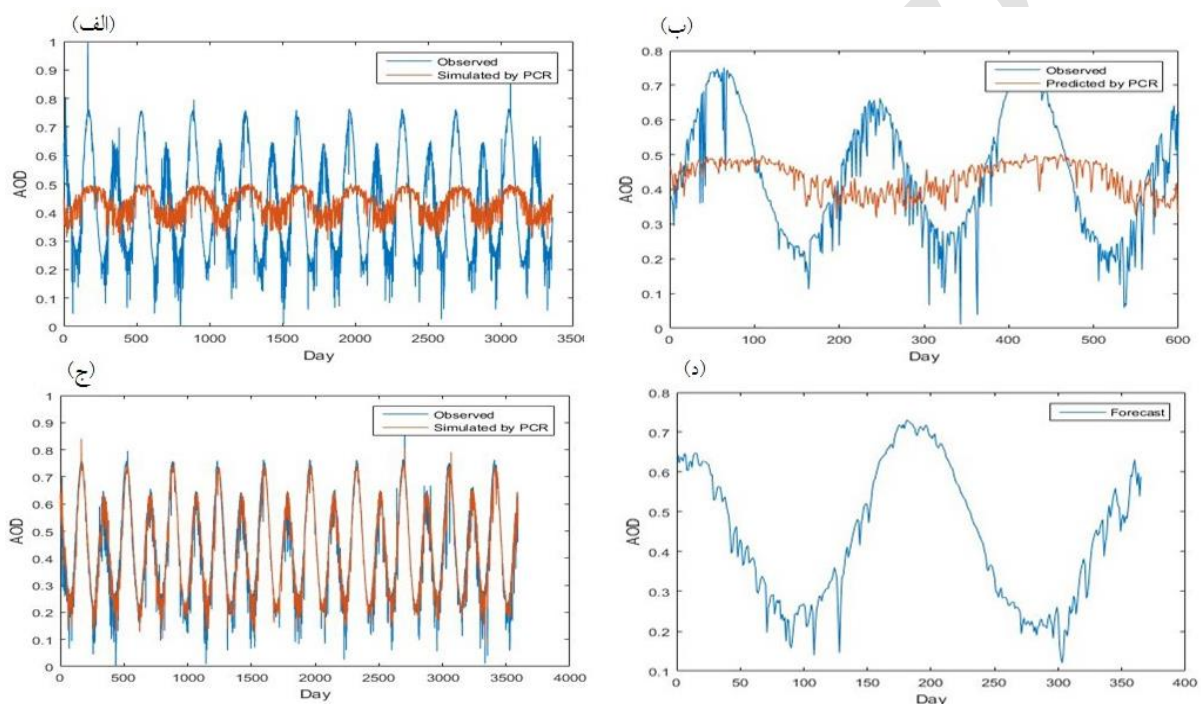
شکل ۲. الف) نتایج مدل MLR در مرحله آموزش مدل، ب) نتایج مدل MLR در مرحله آزمون مدل

$0.126RH_{850} + 0.021Speed_{850} - 0.065BLH$
 ارزش مقادیر عددی در هر کدام از پارامترهای 0.071
 $R^2 = 0.1699$ ، $RMSE = 0.15$ ، $MAE = 0.15$ به دست آمد. در
 شکل (۳-ب) نتایج حاصل از مرحله آزمون مدل PCR در
 شهرستان سنندج نشان داده شده است. در اینجا مقادیر
 عددی $R^2 = 0.069$ ، $RMSE = 0.1694$ ، $MAE = 0.1484$
 حاصل شد.

نتایج حاصل از مدل PCR در مرحله‌های آموزش، آزمون و اتورگرسیو در سطح تراز ۸۵۰ هکتوپاسکال در شهرستان سنندج

در ادامه نتایج حاصل از شبیه‌سازی مدل PCR در سطح
 تراز فوق برای شهرستان سنندج و نیز در مرحله آموزش
 مدل مذکور در شکل (۳-الف) ارائه شده است. معادله
 استخراج شده برای آن به صورت رابطه ۹ است:

$$AOD = 0.457 + 0.041T_{850}^- \quad (9)$$



شکل ۳. (الف) و (ب) به ترتیب نتایج مدل PCR در مرحله آموزش و آزمون مدل، (ج) و (د) به ترتیب
 نتایج مرحله آموزش و اتورگرسیو مدل PCR

طی بازه زمانی یک‌ساله یا به عبارتی دیگر همان سال ۲۰۱۷
 پرداخته شده است. شکل (۳-ج) هماهنگی بین روند
 افزایش و کاهش داده‌های موجود و داده‌های پیش‌بینی
 شده بازه ۱۰ ساله برای مرحله اتورگرسیو در شهرستان
 سنندج را نشان می‌دهد. شکل (۳-د) نیز بیانگر نتایج
 اتورگرسیو مدل برای سال ۲۰۱۷ است که مقادیر $R^2 = 0.9$ ،
 $RMSE = 0.0548$ ، $MAE = 0.0378$ به دست آمد. این
 نتایج نشان می‌دهد که مدل PCR در مرحله آزمون که برای
 پیش‌بینی عمق نوری هواویزها در مناطق فاقد دستگاه

نتایج مرحله آموزش و آزمون مدل بیانگر دقت ضعیف
 مدل PCR در پیش‌بینی مقادیر عمق نوری هواویزها در
 شهرستان سنندج است که این امر به وضوح در شکل‌های
 (۳-الف) و (۳-ب) نیز نشان داده شده است. همان‌طور که
 این شکل‌ها نشان می‌دهد هماهنگی کمی بین روند افزایش
 و کاهش داده‌های موجود و داده‌های مدل‌سازی و آزمون
 شده دیده می‌شود. در ادامه با استفاده از داده‌های عمق
 نوری هواویزها در بازه زمانی $2005/1/1$ تا $2016/12/31$ ،
 به پیش‌بینی عمق نوری هواویزها در شهرستان سنندج در

نشان‌دهنده نتایج مدل ARIMA در مرحله آزمون مدل است. در این مرحله نیز مقادیر عددی $R^2=0/89$ ، $MAE=0/374$ ، $RMSE=0/586$ حاصل شد. ارزش مقادیر عددی به دست آمده در این مدل در مرحله‌های آموزش و آزمون مدل نشان‌دهنده دقت بالاتر این مدل در پیش‌بینی مقادیر عمق نوری هواویزها در مناطق فاقد ایستگاه تشعشع سنج بر اساس مدل شبیه‌سازی شده در ایستگاه شهرستان سنندج است. این امر به وضوح در شکل‌های (۴-الف) و (۴-ب) نیز نشان داده شده است. در این شکل‌ها همخوانی بالایی بین روند تغییرات داده‌های موجود عمق نوری هواویزها و داده‌های پیش‌بینی شده توسط مدل دیده می‌شود. پیش‌بینی عمق نوری هواویزها در بازه یک‌ساله (۲۰۱۷)، با استفاده از مدل مذکور نیز در شکل‌های (۴-ج) و (۴-د) نشان داده شده است. مقادیر عددی $R^2=0/91$ ، $RMSE=0/517$ ، $MAE=0/338$ برای مرحله اتورگرسیون نشان داده شده در شکل (۴-ج) به دست آمد.

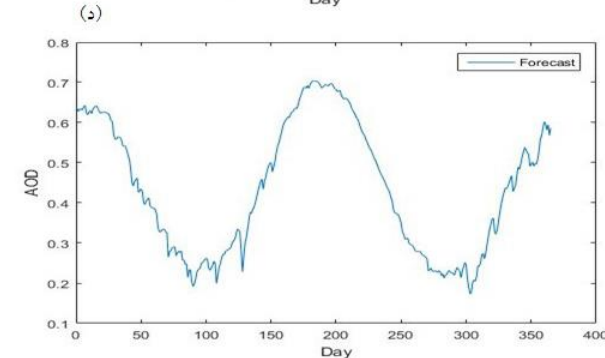
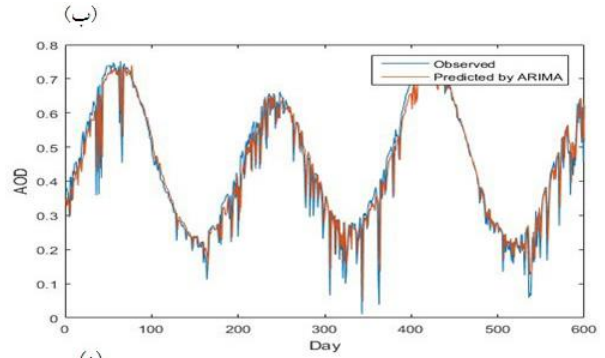
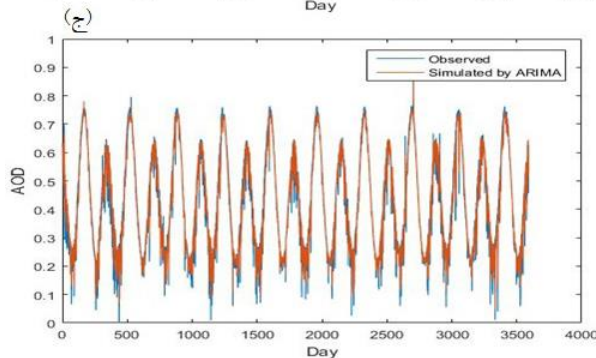
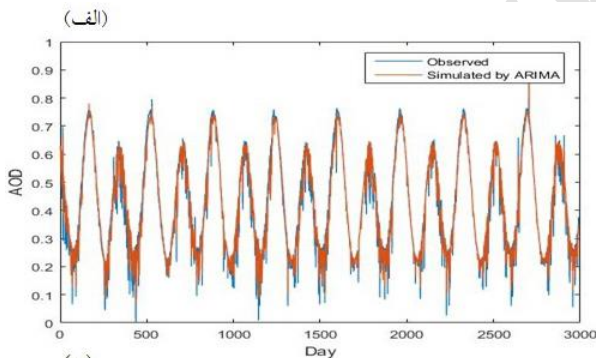
تشعشع سنج بر اساس مدل ساخته شده در ایستگاه سنندج به کار می‌رود فاقد دقت مناسب است ولی در مرحله اتورگرسیون که برای پیش‌بینی طولانی مدت در مناطق دارای دستگاه تشعشع سنج به کار می‌رود دارای دقت قابل قبولی است.

نتایج حاصل از مدل ARIMA در مرحله‌های آموزش، آزمون و اتورگرسیون در سطح تراز ۸۵۰ هکتوپاسکال در شهرستان سنندج

نتایج حاصل از شبیه‌سازی مدل ARIMA در سطح تراز ذکر شده برای شهرستان سنندج و در مرحله آموزش مدل مذکور در شکل (۴-الف) نشان داده شده است. معادله ارائه شده برای آن در رابطه ۱۰ آمده است:

$$AOD = 0/0061 + 0/7084y_{t-1} + 0/572y_{t-2} + 0/2189y_{t-3} \quad (10)$$

در این مرحله از مدل مقادیر عددی $R^2=0/91$ ، $MAE=0/33$ ، $RMSE=0/501$ حاصل شد. شکل (۴-ب)



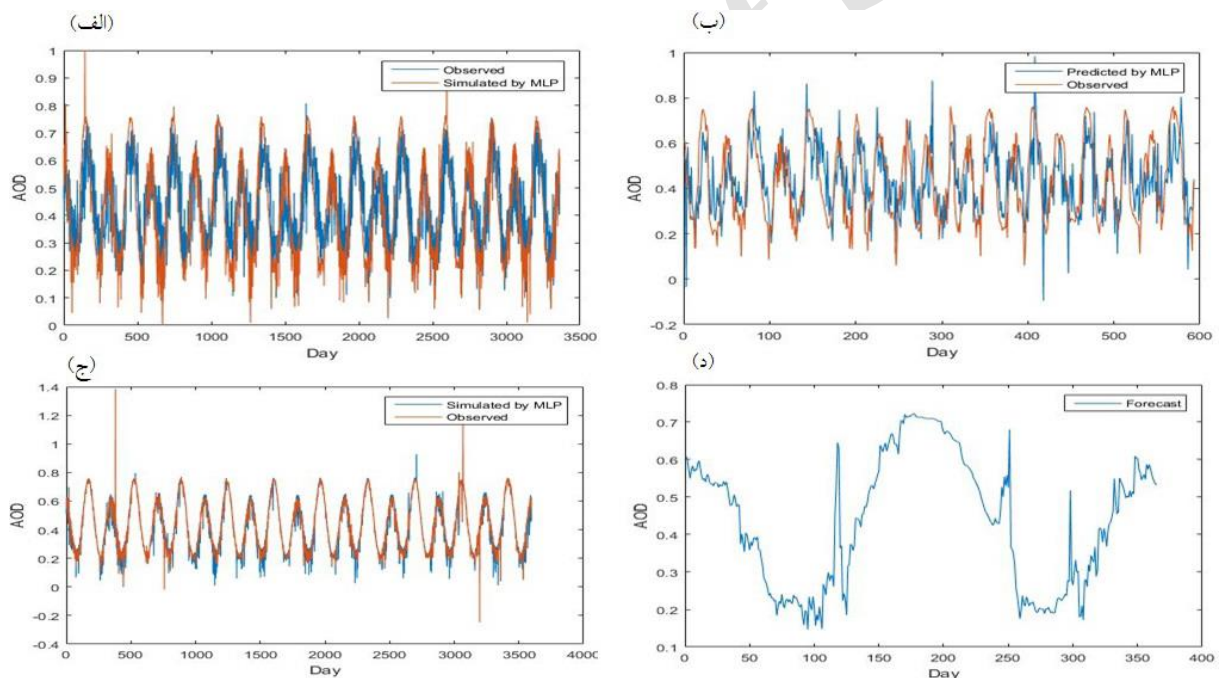
شکل ۴. (الف) و (ب) به ترتیب نتایج مدل ARIMA در مرحله‌های آموزش و آزمون مدل، (ج) و (د) به ترتیب نتایج مدل ARIMA در مرحله‌های آموزش و اتورگرسیون

استفاده شده است. از داده‌های بازه زمانی مورد استفاده در این پژوهش، ۷۵ درصد به منظور آموزش، ۱۵ درصد برای اعتبارسنجی و ۱۰ درصد باقیمانده از کل داده‌ها نیز به منظور تست استفاده شد. نتایج حاصل از شبیه‌سازی مرحله آموزش در شکل (۵-الف) نشان داده شده است که مقادیر عددی $R^2=0.75$ ، $RMSE=0.1162$ ، $MAE=0.0921$ در این مرحله به دست آمد. نتایج حاصل از مرحله آزمون مدل MLP نیز در شکل (۵-ب) نشان داده شده است. که برای این مرحله نیز مقادیر عددی $R^2=0.63$ ، $MAE=0.14$ ، $RMSE=0.113$ حاصل شد.

این نتایج نشان می‌دهد که مدل ARIMA علاوه بر توانایی بالا در پیش‌بینی مقادیر عمق نوری هواویزها در مناطق فاقد دستگاه تشعشع سنج دارای دقت بالایی در پیش‌بینی مقادیر عمق نوری هواویزها در مناطق دارای دستگاه تشعشع سنج برای مدت طولانی است که این امر به روشنی در همخوانی بالای داده‌های موجود و داده‌های مدل‌سازی شده در شکل (۴-ج) بیان شده است.

نتایج حاصل از مدل MLP در مرحله‌های آموزش، آزمون و اتورگرسیو در سطح تراز ۸۵۰ هکتوپاسکال در شهرستان سنندج

در مدل MLP از دولایه پنهان با تعداد ۳۳ و ۲۴ نرون



شکل ۵. (الف) و (ب) به ترتیب نتایج مدل MLP در مرحله‌های آموزش و آزمون مدل، (ج) و (د) به ترتیب نتایج مدل MLP در مرحله‌های آموزش و اتورگرسیو

نوری هواویزها و داده‌های پیش‌بینی شده مدل نشان داده شده در شکل‌های (۵-الف) و (۵-ب) نیز بیانگر این میزان دقت است. نتایج حاصل از مرحله اتورگرسیو مقادیر عمق نوری هواویزها برای سال ۲۰۱۷ با استفاده از مدل MLP در شکل‌های (۵-ج) و (۵-د) نیز آمده است در این مرحله

طبق نتایج به دست آمده، پیش‌بینی مقادیر عمق نوری هواویزها برای مناطق فاقد دستگاه تشعشع سنج با استفاده از مدل MLP مدل‌سازی شده براساس داده‌های ایستگاه شهرستان سنندج دارای دقت متوسطی است. روند همخوانی تغییرات افزایشی و کاهش‌ی داده‌های موجود عمق

نتیجه دست یافتند که مدل MLP با دارا بودن ضریب تعیین ۰/۹۴، بهترین مدل با بالاترین میزان دقت است و مدل ANFIS^{۱۳} با دارا بودن ضریب تعیین ۰/۸۶ در رتبه دوم قرار دارد و مدل MLR همواره با دارا بودن کمترین میزان ضریب تعیین، دارای کمترین میزان دقت بوده است.

دومین مدل استفاده شده در پژوهش حاضر مدل PCR بود. در این مدل نیز میزان AOD با پارامترهای دما و سرعت باد رابطه‌ای مستقیم دارد ولی با پارامترهای دیگر نظیر رطوبت نسبی و ارتفاع لایه اتمسفری در سطح تراز فوق، رابطه‌ای منفی را نشان می‌دهد. مدل PCR در صورت وجود هم‌خطی بین متغیرهای مستقل، باعث از بین بردن این هم‌خطی‌ها می‌شود (Zhou et al., 2014). طبق نتایج به‌دست آمده، خروجی‌های منتج شده از طریق مدل‌های MLR و PCR در برآورد عمق نوری هواویزها برای مناطق فاقد ایستگاه تشعشع سنج دارای نتایجی نزدیک به هم هستند. این نتیجه با کار سیفی و همکاران همخوانی ندارد. سیفی و همکاران (۱۳۸۹) به منظور پیش‌بینی میزان تبخیر و تعرق به مقایسه دقت مدل‌های خطی MLR و PCR پرداختند. آن‌ها با توجه به ضرایب تعیین به‌دست آمده، دقت مدل PCR را بالاتر از دقت مدل MLR ارزیابی کردند که عدم همخوانی نتایج می‌تواند به دلیل همبستگی نه چندان بالای متغیرهای مستقل و وابسته مورد بررسی در پژوهش حاضر باشد. سومین مدل استفاده شده، مدل ARIMA بود. این مدل دارای بهترین عملکرد در برآورد AOD در ایستگاه‌های دارای تشعشع سنج بود هر چند توانایی بالایی نیز در برآورد عمق نوری هواویزها در ایستگاه‌های فاقد تشعشع سنج داشت. نتایج به‌دست آمده با نتایج تانجا و همکاران همخوانی دارد. تانجا و همکاران (۲۰۱۶) از مدل ARIMA به منظور شبیه‌سازی میانگین ماهیانه عمق نوری هواویزها استفاده کردند، آن‌ها میزان ضرایب تعیین به‌دست آمده را ۰/۵۳ و ۰/۶۷۴ گزارش دادند. نتایج این بخش با کار اوسطی و همکاران همخوانی ندارد. اوسطی و همکاران (۲۰۱۵) به پیش‌بینی غلظت ذرات معلق PM_{2.5} در

مقادیر عددی $MAE=0/028$ ، $RMSE=0/0483$ ، $R^2=0/96$ به دست آمد. نتایج مرحله اتورگرسیون مدل MLP نشان می‌دهد که با وجود اینکه این مدل توانایی کمی در استفاده در مناطق فاقد دستگاه تشعشع سنج را دارد ولی توانایی بالایی در پیش‌بینی مقادیر عمق نوری هواویزها برای مدت طولانی در مناطق دارای دستگاه تشعشع سنج را دارد.

۵. بحث و نتیجه‌گیری

همان‌طور که یاد شد برای انجام این پژوهش از مدل‌های مختلفی بهره گرفته شد. نخستین مدل، مدل MLR که طبق نتایج حاصل شده برای این مدل، میزان AOD موجود در شهرستان مورد مطالعه، با پارامترهای دما و سرعت باد در سطح تراز ۸۵۰ هکتوپاسکال، رابطه‌ای مستقیم دارد اما با پارامترهای رطوبت نسبی و ارتفاع لایه اتمسفری دارای رابطه‌ای عکس است. همچنین مقدار ضریب تعیین به‌دست آمده از طریق این مدل، مقدار عددی کمی را به خود اختصاص داده است که آن نیز به دلیل ساختار خطی در داده‌های مورد استفاده است. این نتیجه با نتایج تیریایی^{۱۱} و همکاران در سال ۲۰۱۴ و کبودوندپور^{۱۲} و همکاران در سال ۲۰۱۵ همخوانی دارد. کبودوندپور و همکاران در سال ۲۰۱۵ با ارزیابی دقت مدل‌های MLR و مدل MLP و نیز مدل سیستم استنتاج تطبیقی عصبی - فازی در پیش‌بینی وقوع طوفان‌های گردوغبار در غرب کشور (سندج) پرداختند. آن‌ها به این نتیجه دست یافتند که مدل MLR با دارا بودن ضریب تعیین ۰/۴۱ دارای دقت کمتری نسبت به دو مدل دیگر با دارا بودن ضریب تعیین ۰/۸۶ برای مدل MLP و ضریب تعیین ۰/۶۴ برای مدل سیستم استنتاج تطبیقی عصبی - فازی هستند. همچنین این نتیجه با نتایج خاشعی و همکاران نیز همخوانی دارد. خاشعی و همکاران (۲۰۱۳) با استفاده از روش‌های کلاسیک آماری و سیستم‌های خبره (شبکه عصبی مصنوعی و سیستم‌های استنتاج فازی عصبی) به پیش‌بینی سطح آب‌های زیرزمینی در دشت نیشابور پرداختند. آن‌ها به این

ذرات معلق با استفاده AOD پرداخته شده است. زانگ و همکاران (۲۰۱۸) با استفاده از مدل غیرخطی ترکیبی (PCA-GRNN) ذرات معلق PM_{10} را در چین پیش‌بینی کردند. نتایج آن‌ها نشان داد که این مدل دارای دقت بالایی در پیش‌بینی ذرات معلق PM_{10} است که این امر نیز بیانگر برتری مدل‌های غیرخطی بر مدل‌های خطی است. با توجه به نتایج این تحقیق می‌توان نتیجه گرفت که بهتر است در برآورد AOD برای مناطق دارای ایستگاه تشعشع سنج به‌جای استفاده از مدل ساخته شده برای شهر سنج از مدل‌سازی (آموزش) و متعاقب آن اتورگرسیو مدل بر اساس داده‌های همان ایستگاه استفاده شود. چون همان‌طور که بیان شد، داده مورد نیاز مرحله اتورگرسیو فقط داده عمق نوری هواویزها است که در ایستگاه دارای دستگاه تشعشع سنج وجود دارد. در کل از نتایج این تحقیق می‌توان استنباط کرد که استفاده از مدل‌های مختلف و کارآمد می‌تواند راه‌حلی مناسب برای برآورد عمق نوری هواویزها برای مناطق دارای ایستگاه تشعشع سنج و همچنین مناطق فاقد ایستگاه تشعشع سنج باشد.

یادداشت

1. Aerosol Optical Depth
2. Olcese
3. Taneja
4. Autoregressive integrated moving average model
5. Principal Component Regression
6. European Centre for Medium-Range Weather Forecasts
7. Multiple linear regression
8. Principal Component Analysis
9. Artificial Neural Networks
10. Multi Layer Perceptron
11. Tiryaki
12. Kaboodvandpour
13. Adaptive neuro fuzzy inference system
14. Zaman
15. Soni
16. Zang

شهر سنج پرداختند. برای این کار چندین مدل مختلف را استفاده کردند. آن‌ها در مقایسه دقت مدل‌های PCR و ARIMA به منظور پیش‌بینی آینده، دقت مدل ARIMA را با بودن ضریب تعیین $0/63$ را کمتر از مدل PCR ارزیابی کردند. که به نظر می‌رسد دلیل عدم همخوانی نتایج به ساختار خطی در داده‌های مورد استفاده در کار آن‌ها بر می‌گردد. چهارمین مدل مورد استفاده در پژوهش فوق مدل MLP بود. در مدل MLP مورد استفاده در این تحقیق از دولایه پنهان استفاده شد. تعداد بهینه نرون‌ها برای منطقه مورد مطالعه نیز با توجه به تعداد داده‌های موجود متفاوت بود. تعداد نرون‌های بهینه برای شهرستان سنج ۳۳ و ۲۴ نرون در دولایه پنهان تعیین شد. این مدل دارای بهترین عملکرد در برآورد عمق نوری هواویزها در مدت‌های طولانی (یک سال) در ایستگاه‌های فاقد تشعشع سنج است نتایج این تحقیق با نتایجی که اولس و همکاران به دست آوردند، همخوانی دارد. اولس و همکاران (۲۰۱۵) به برآورد مقادیر AOD با استفاده از مدل MLP پرداختند. بدین منظور آن‌ها از داده‌های بازه زمانی ۱۹۹۹-۲۰۱۲ استفاده کردند. آن‌ها در مقایسه مقدارهای محاسبه و تخمین زده شده AOD، ضریب تعیین $0/85$ را گزارش دادند. به صورت مقایسه کلی نتایج این تحقیق با نتایج سونی و همکاران و زمان^{۱۴} و همکاران همخوانی دارد. سونی^{۱۵} و همکاران (۲۰۱۸) با استفاده از داده‌های هواشناسی و AOD به پیش‌بینی ذرات معلق پرداختند نتایج آن‌ها نیز نشان داد که مدل رگرسیون غیرخطی دارای عملکرد بهتری در مقایسه با مدل رگرسیون خطی بود. زمان و همکاران (۲۰۱۷) با استفاده از داده‌های AOD منتجه از ماهواره و مدل‌های شبکه عصبی مصنوعی و رگرسیون چندگانه به پیش‌بینی ذرات معلق در جو پرداختند نتایج آن‌ها نشان داد که شبکه عصبی مصنوعی دارای عملکرد بهتری در مقایسه با رگرسیون چندگانه در پیش‌بینی ذرات معلق موجود در جو مالزی است. در بعضی مطالعات دیگر مانند زانگ^{۱۶} و همکاران به تنهای از مدل‌های غیرخطی برای پیش‌بینی

منابع

- احسان زاده، ع.، نژاد کورکی، ف. و طالبی، ع. ۱۳۹۵. مقایسه کاربرد شبکه عصبی مصنوعی، درخت تصمیم، رگرسیون مؤلفه‌های اصلی و رگرسیون خطی چندگانه در مدل‌سازی شاخص کیفیت هوای شهری، محیط‌شناسی، ۴۲ (۳): ۴۵۵-۴۷۳.
- اوسطی، ش.، امان‌اللهی، ج. و محمدی، ب. ۱۳۹۵. ارزیابی دقت مدل‌های رگرسیون چندگانه، ARIMA و GRNN در پیش‌بینی غلظت ذرات معلق $PM_{2.5}$ در شهر سنندج، فصلنامه علمی-پژوهشی فضای جغرافیایی، ۵۵: ۱۵۳-۱۷۵.
- حجتی، م.، درویش بلورانی، ع.، علوی پناه، ک.، کیاورز، م.، براق جمالی، ج. ۱۳۹۵. مدل مبتنی بر شبکه‌های عصبی مصنوعی به‌منظور محتوای PM_{10} توفان‌های گردوغبار به کمک نساویر ماهواره‌ای MODIS، محیط‌شناسی، ۴۲ (۴): ۸۲۳-۸۳۸.
- خلیلی، ع.، رضایی صدر، ح. ۱۳۷۶. تخمین تابش کلی خورشید در گستره ایران بر مبنای داده‌های اقلیمی، فصلنامه تحقیقات جغرافیایی، ۴۶: ۱۵-۳۵.
- دانش جعفری، د.، آماده، ح. و خون سیاوشان، ش. ۱۳۹۴. برآورد خسارات ناشی از پدیده ریزگردها بر سلامت افراد در ایران (مطالعه موردی: استان‌های خوزستان، کرمانشاه و کردستان)، محیط‌شناسی، ۴۱ (۳): ۵۷۳-۵۸۷.
- دستورانی، م.، عظیمی فشی، خ.، طالبی، ع. و اختصاصی، م. ۱۳۹۱. برآورد رسوبات معلق با استفاده از شبکه عصبی مصنوعی مطالعه موردی (حوضه آبخیز جامیشان استان کرمانشاه)، پژوهشنامه مدیریت حوضه آبخیز، ۳ (۶): ۶۱-۷۴.
- سیفی، ا.، میرلطیفی، م. و ریاحی، ح. ۱۳۸۹. توسعه مدل ترکیبی رگرسیون چندگانه-تحلیل مؤلفه‌ها و عامل‌های اصلی (MLR-PCA) در پیش‌بینی تبخیر-تعرق مرجع، مطالعه موردی: ایستگاه کرمان، نشریه آب و خاک، ۲۴ (۶): ۱۱۸۶-۱۱۹۶.
- نادری، آ.، دلاور، م. و کبودین، ب. ۱۳۹۶. مقایسه عملکرد مدل‌های خطی و هیبریدی در پیش‌بینی پراکنش عناصر سنگین با استفاده از سنجش از دور و تحلیل فضایی در شرق زنجان، محیط‌شناسی، ۴۳ (۱): ۸۷-۱۰۰.
- Díaz-Robles, L.A., Ortega, J.C., Fu, J.S., Reed, G.D., Chow, J.C., Watson, J.G. and Moncada-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.
- Dickson, D. 1995. Aerosols role simulated in new global warming model. *Nature*, 374: 487-487.
- Holben, B.N., Eck, T.F., Slutsker, I., Tanre, D., Buis, J.P., Setzer, A. and Lavenu, F. 1998. AERONET—A federated instrument network and data archive for aerosol characterization. *Remote sensing of environment*, 66: 1-16.
- Intergovernmental Panel on Climate Change (IPCC), Climate change: the Physical Science Basis. In: Solomon, S. (Ed.). (2007). *Climate change 2007-the physical science basis: Working group I contribution to the fourth assessment report of the IPCC (Vol. 4)*. Cambridge University Press.
- Jeffers, J.N.R. 1967. Two case studies in the application of principal component analysis. *Applied Statistics*, 225-236.
- Kaboodvandpour, S., Amanollahi, J., Qhavami, S. and Mohammadi, B. 2015. Assessing the accuracy of multiple regressions, ANFIS, and ANN models in predicting dust storm occurrences in Sanandaj, Iran. *Natural Hazards*, 78: 879-893.
- Krzanowski, W. J. and Marriott, F. H. C. 1994. *Multivariate Analysis*, E. Arnold, London.
- Kuzmanovski, I. and 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.
- McBratney, A.B., Minasny, B., Cattle, S.R. and Vervoort, R.W. 2002. From pedotransfer functions to soil inference systems. *Geoderma*, 109: 41-73.
- Olcese, L.E., Palancar, G.G. and Toselli, B.M. 2015. A method to estimate missing AERONET AOD values based on artificial neural networks. *Atmospheric Environment*, 113: 140-150.
- Pai, P.F. and Lin, C.S. 2005. A hybrid ARIMA and support vector machines model in stock price forecasting. *Omega*, 33: 497-505.

- Rahul, P.R.C., Sonbawne, S.M. and Devara, P.C. S. 2014. Unusual high values of aerosol optical depth evidenced in the Arctic during summer 2011. *Atmospheric Environment*, 94: 606-615.
- Seinfeld, J. H. and Pandis, S.N. 2016. *Atmospheric chemistry and physics: from air pollution to climate change*. John Wiley & Sons.
- Soni, K., Kapoor, S., Parmar, K.S. and Kaskaoutis, D.G. 2014. Statistical analysis of aerosols over the Gangetic–Himalayan region using ARIMA model based on long-term MODIS observations. *Atmospheric Research*, 149: 174-192.
- Soni, M., Payra, S., Verma, S. 2018. Particulate matter estimation over a semi arid region Jaipur, India using satellite AOD and meteorological parameters. *Atmospheric Pollution Research*, (in Press). <https://doi.org/10.1016/j.apr.2018.03.001>
- Taneja, K., Ahmad, S., Ahmad, K. and Attri, S.D. 2016. Time series analysis of aerosol optical depth over New Delhi using Box–Jenkins ARIMA modeling approach. *Atmospheric Pollution Research*, 7: 585-596.
- Tiryaki, S., Özşahin, Ş. and Yıldırım, İ. 2014. Comparison of artificial 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.
- World meteorological organization. (1995), (2001) edition. *Manual on Codes, Suppl. 6(VIII.2007)*, WMO, Geneva, Switzerland.
- Zaman, N.A.F.K., Kanniah, K.D. and Kaskaoutis, D.G. 2017. Estimating Particulate Matter using satellite based aerosol optical depth and meteorological variables in Malaysia. *Atmospheric Research*. 193: 142-162.
- Zang, L., Mao, F., Guo, J., Gong, W., Wang, W. and Pan, Z. 2018. Estimating hourly PM1 concentrations from Himawari-8 aerosol optical depth in China. *Environmental Pollution*. 241: 654-663.
- Zhang, X., Huang, Y., Zhu, W. and Rao, R. 2013. Aerosol characteristics during summer haze episodes from different source regions over the coast city of North China Plain. *Journal of Quantitative Spectroscopy and Radiative Transfer*, 122: 180-193.
- Zhou, Q., Jiang, H., Wang, J. and Zhou, J. 2014. A hybrid model for PM 2.5 forecasting based on ensemble empirical mode decomposition and a general regression neural network. *Science of the Total Environment*, 496: 264-272.

Archive of SID