

## مقایسه دقت مدل رگرسیون بردار پشتیبان با دو روش رایج شبکه عصبی مصنوعی و شبکه عصبی فازی - تطبیقی در پیش بینی غلظت آلاینده PM<sub>10</sub>

سید مصطفی میرزاده<sup>۱</sup>؛ فرهاد نژاد کورکی<sup>۲\*</sup>؛ وحید موسوی<sup>۳</sup> و سید ابوالقاسم میر حسینی<sup>۴</sup>

۱ - دانشجوی دکترای آلودگی محیط زیست دانشکده فنی مهندسی دانشگاه آزاد اسلامی یزد

۲ - دانشیار محیط زیست گروه محیط زیست دانشگاه یزد

۳ - استادیار گروه مهندسی آبخیزداری دانشکده منابع طبیعی دانشگاه تربیت مدرس تهران

۴ - دانشیار محیط زیست دانشکده فنی مهندسی دانشگاه آزاد اسلامی یزد

(تاریخ دریافت ۹۹/۰۵/۱۳ - تاریخ پذیرش ۹۹/۱۰/۱۳)

### چکیده:

شهر یزد واقع در مرکز ایران می باشد که در اکثر ایام سال طوفان های گرد و غبار را تجربه می کند، و همچنین طی دو دهه گذشته شاهد رشد سریع صنعتی بخصوص در صنایع با توان غبار زایی بالا (کاشی، فولاد و ..) بوده است، لذا پیش بینی غلظت آلاینده ذرات از طریق بکارگیری سیستم های پیش آگاهی دقیق در زمان طوفان های گرد و غبار و آلودگی های صنعتی برای حفظ بهداشت و سلامت شهروندان بسیار حیاتی است. در این مطالعه بطور همزمان از مدل شبکه عصبی فازی تطبیقی (ANFIS)، رگرسیون بردار پشتیبان (SVR) و شبکه عصبی مصنوعی (ANN) در پیش بینی میزان غلظت آلاینده PM<sub>10</sub> برای ۲۴ ساعت آینده و معرفی دقیق ترین روش استفاده شد. به همین منظور داده های PM<sub>10</sub> شهر یزد به همراه پارامتر هواشناسی شهر در بازه زمانی سال های ۱۳۹۴ تا ۱۳۹۸ مورد استفاده قرار گرفت. نتایج این تحقیق نشان داد که مدل ANFIS با (R<sup>2</sup>=0.989) و دقتی در حدود ۹۹ درصدی بهتر از سایر مدل ها در این حالت است و بعد از آن به ترتیب مدل ANN با (R<sup>2</sup>=0.978) و SVR با (R<sup>2</sup>=0.957) دارای بهترین دقت بودند. بنابراین می توان گفت، با توجه به صحت مدل، از این مدل می توان برای پیش بینی غلظت آلاینده PM<sub>10</sub> استفاده کرد و این امر می تواند مسئولین را در تصمیم گیری های به موقع در جهت حفظ سلامت عمومی قبل از شروع وضعیت هشدار غلظت آلاینده PM<sub>10</sub> کمک کند.

**کلید واژگان:** PM<sub>10</sub>، پیش بینی ۲۴ ساعته، شبکه عصبی فازی تطبیقی

## ۱. مقدمه

این روزها آلاینده های هوا موضوع مهم و تاثیرگذار بر سلامت بشر و محیط زیست بوده و به شدت جوامع بشری را تهدید می نماید (Liu et al., 2019). آلودگی در کشورهای در حال توسعه به طور پیوسته در حال افزایش و اثرات نامطلوبی بر کودکان، سالمندان و بیماران تنفسی و قلبی-عروقی می گذارد (Maleki et al., 2019; World Health, 2016 & 2017). در بین آلاینده های جوی، ذرات (PM) نقش مهمی در سلامت انسان و همچنین تغییرات آب و هوایی دارند (Forster et al., 2007; H. S. Kim et al., 2019). نوع آلودگی هوا محسوب می شود (Wasley et al., 2019). ذرات قابل استنشاق با قطر کمتر از ۱۰ میکرومتر تحت عنوان PM<sub>10</sub> طبقه بندی می شوند (EPA, 2016)، که منابع اصلی تولید کننده این ذرات بویژه (PM<sub>10</sub>) در هوا شامل وقایع طبیعی مانند طوفان های گرد و غبار، و وقایع انسان ساز مانند احتراق سوخت های فسیلی در فرآیندهای صنعتی و وسائل نقلیه موتوری هستند (Omidvarborna et al., 2015). در سال های اخیر، مطالعات اپیدمیولوژیک بسیاری ارتباط نزدیک و افزایشی بین مرگ و میر انسان و عوارضی همچون بیماری های قلبی عروقی، تنفسی با غلظت آلاینده PM<sub>10</sub> را نشان می دهد (Hamanaka et al., 2018; J. Kim, 2019). امروزه یکی از نیازهای مهم که نقش حیاتی بر سلامت افراد جامعه دارد، تدوین راهکارهای مناسب جهت اطلاع شهروندان از وضعیت آلاینده PM<sub>10</sub> برای انجام اقدامات احتیاطی است. از مدل های پیش بینی به عنوان یکی از گزینه های

کارآمد برای کمک به مدیریت کیفیت هوای شهرها در سراسر جهان استفاده شده است (Pérez et al., 2010; Schlink et al., 2003). ثابت شده است که تکنیک های مدل سازی خطی در شرایط غیر خطی دقت کمتری را در پیش بینی نشان می دهد، و این در حالی است که داده های غلظت PM<sub>10</sub> که از نظر تغییرات زمانی اندازه گیری می شوند، دارای ویژگی های غیر خطی می باشند و پیش بینی با این مدل ها با دقت مناسب و کافی همراه نیست (Zhang, 2003). از این رو به نظر می رسد که مدل های غیر خطی بسیار مهم هستند (Paschalidou et al., 2011). از مدل های غیرخطی که غالباً در پیش بینی آلودگی هوا استفاده می شود، می توان به مدل های شبکه عصبی مصنوعی، سیستم استنتاج فازی مبتنی بر شبکه تطبیقی (ANFIS) و ماشین بردار پشتیبان (SVM) اشاره نمود (Azadeh et al., 2011; Buragohain et al., 2008). ماشین های بردار پشتیبان یک نوع شبکه عصبی ویژه ای هستند که بر خلاف سایر انواع شبکه های عصبی خطرهای عملیاتی را کاهش و ابزاری بسیار قدرتمند در زمینه پیش بینی آلودگی ذرات معلق محسوب می شود (Scholkopf et al., 1997; Vapnik, 2013).

Amanollahi و همکاران (۲۰۲۰) به منظور پیش بینی غلظت PM<sub>10</sub> در هوای تهران، از مدل های مختلفی استفاده از جمله شبکه عصبی مصنوعی استفاده نمودند. تجزیه و تحلیل داده ها نشان داد که در مقایسه با نتایج MLR و MLP، دقیق ترین نتایج را ANFIS دارا می باشد و مدل های ترکیبی از دقت بالاتری در پیش بینی غلظت PM<sub>10</sub> برخوردار هستند. Kim و همکاران

## ۲. مواد و روش ها

### ۲-۱. منطقه مورد مطالعه

شهر یزد مرکز استان یزد در حاشیه کویرهای مرکزی و حاشیه شمالی رشته کوه های شیرکوه واقع شده و در محدوده جغرافیایی ۳۱ درجه و ۵۱ دقیقه تا ۳۱ درجه و ۴۷ دقیقه شمالی و طول جغرافیایی ۵۴ درجه و ۱۸ دقیقه تا ۵۴ درجه و ۲۲ دقیقه شرقی از نصف النهار مبدأ قرار گرفته است. شهر یزد در دشتی پهناور گسترده شده و جمعیت آن طبق آخرین سرشماری در سال ۱۳۹۵ در حدود ۶۵۶ هزار نفر می باشد. در طی دو دهه گذشته شاهد رشد و توسعه سریع صنعتی و معدنی در اطراف و بخصوص در سمت باد غالب شهر هستیم صنایعی چون کاشی، فولاد و... و از سوی دیگر این شهر از دیرباز در معرض توفان های گرد و خاک بوده و وقوع پدیده ریزگردها در آن، بسیار قابل توجه می باشد.

### ۲-۲. داده های مورد استفاده

با توجه به مدل های انتخابی، در این تحقیق از داده های ساعتی پارامترهای هواشناسی (دما، جهت باد، سرعت باد، رطوبت، میزان بارش، فشار اتمسفر و ابرناکی و دید افقی) مربوط به ایستگاه هواشناسی شهر یزد و داده های ساعتی پارامتر آلاینده  $PM_{10}$  شهر مربوط به سه ایستگاه سنجش آلودگی هوای اداره کل حفاظت محیط زیست استان، برای بازه زمانی ۴ ساله (۱۳۹۵ تا ۱۳۹۸) استفاده شد.

### ۲-۳. روش کار

با توجه به اینکه داده های میزان آلودگی ذرات معلق کمتر از ۱۰ میکرون ( $PM_{10}$ ) شهر یزد، یک سری زمانی را تشکیل می دهند، لذا برای پیش بینی میزان آلودگی،

(۲۰۱۹) طی مطالعه ای در شهر سئول کره جنوبی با استفاده از شبکه عصبی مصنوعی اقدام به پیش بینی میزان روزانه ذرات ۲,۵ و ۱۰ میکرون نمودند و نتایج حاکی از عملکرد بالای مدل ANN در پیش بینی میزان ذرات بود. Ghsemi و همکاران (۲۰۱۹) که از دو مدل ANN و ANFIS به منظور پیش بینی غلظت  $PM_{10}$  در دو بازه زمانی امروز و فردا در هوای شهر کرمانشاه استفاده کردند و مدل عصبی - فازی به عنوان مدل بهینه انتخاب شد. Maleki و همکاران (۲۰۱۹) طی مطالعه ای از مدل های شبکه عصبی مصنوعی (ANN) در پیش بینی شاخص کیفیت هوا و همچنین غلظت آلاینده هوا به صورت ساعتی استفاده نمودند. این مطالعه نشان داد که ANN در پیش بینی کیفیت هوا با هدف جلوگیری از اثرات آلودگی هوا، در شهرهایی از جمله اهواز دارای کارایی بالایی است. Kaboodvandpour و همکاران (۲۰۱۵) به منظور پیش بینی ذرات معلق حاصل از طوفان های گرد و غبار در سنندج عملکرد مدل ANFIS را با ANN و MLR مقایسه نمودند، نتایج حاکی از برتری و دقت بالاتر مدل ANFIS در مقایسه با دو مدل دیگر بود. با توجه به تاثیر انکار ناپذیر و بالای  $PM_{10}$  بر روی سلامت بهداشت جامعه در این مقاله سعی شده است تا برای اولین بار به طور همزمان با استفاده از ۳ مدل کاربرد و دقیق (ANN, SVR, ANFIS) در بازه زمانی (۱۳۹۴ تا ۱۳۹۸) اقدام به پیش بینی غلظت آلاینده ( $PM_{10}$ ) برای ۲۴ ساعت آینده پرداخته و کارایی، دقت و عملکرد این مدل ها را با هم مقایسه، و بهترین و دقیق ترین مدل پیش بینی را با توجه به داده های ورودی برای شهر یزد معرفی نماییم.

های ورودی  $I_i$  دریافت می‌کند. هر یک از این گره‌ها ی ورودی، قبل از اینکه وارد هسته اصلی نورون شود، وزن دار می‌شود. یعنی مقدار هر ورودی در  $W_{i,m}$  ضرب می‌شود. سپس این مقادیر در بخش اول پردازشگر با هم جمع شده و مجموع کل ورودی به نورون تعیین می‌شود. در برخی موارد، یک مقدار ثابت در هر نورون بنام وزن اریب (Bias) به مقدار کل ورودی افزوده می‌شود. مقدار این ورودی یک و وزن آن  $b$  است. مقدار کل ورودی به نورون از رابطه ۲ محاسبه می‌شود؛ در مرحله بعد، یک تابع موسوم به تابع انتقال (Activation Function) که معمولاً تابعی غیرخطی است، روی این مقدار حاصل جمع عمل کرده و مقدار خروجی نورون از رابطه ۳ تعیین می‌گردد (Haykin, 1994).

$$U_m = \sum_{i=1}^n I_i W_{i,m} + b_m \quad (2)$$

$$O_m = f(\sum_{i=1}^n I_i W_{i,m}) + b_m \quad (3)$$

## ۲-۵. شبکه عصبی-فازی تطبیقی (Adaptive

### (Neuro Fuzzy Inference System (ANFIS)

در سال‌های اخیر از ترکیب منطق فازی با شبکه‌های عصبی مصنوعی، سیستم‌های فازی-عصبی به وجود آمده‌اند. یک شبکه تطبیقی، یک ساختار پیشرو چند لایه است که رفتار کلی خروجی آن به وسیله مقدار یک مجموعه از پارامترهای قابل اصلاح تعیین می‌گردد. به عبارت دیگر ساختار شبکه تطبیقی شامل یک مجموعه از گره‌های متصل به هم است که به طور مستقیم به هم مرتبط شده‌اند و در آن هر گره یک واحد پردازش محسوب می‌شود. ANFIS یکی از متداول‌ترین سیستم‌های عصبی فازی است که یک سیستم فازی سوگنو (Sugeno) را در یک ساختار عصبی اجرا می‌کند صفحه ۱۷۰

متغیرهای ورودی به مدل های یادگیری ماشین، میزان ذرات معلق کمتر از ۱۰ میکرون ( $PM_{10}$ ) شهر یزد در روزهای پیشین یا همان گام‌های پیشین است. به این روند ایجاد تاخیر زمانی بر روی پارامترهای ورودی گفته می‌شود. لذا به منظور پیش بینی میزان آلودگی ذرات شهر یزد در ۲۴ ساعت آینده، در ابتدا تاخیر زمانی بر روی متغیر ( $PM_{10}$ ) شهر یزد لحاظ می‌شود که به صورت  $X - a$  مشخص می‌شود که  $X$  پارامتر ورودی و  $a$  تعداد روز تاخیر می‌باشد. در این تحقیق ورودی‌های مدل های یادگیری ماشین، میزان متغیر ( $PM_{10}$ ) شهر یزد در ۱، ۲ و ۳ روز قبل یعنی  $X-1$ ,  $X-2$ ,  $X-3$  است، به عبارتی میزان تاخیرها را ۱، ۲ و ۳ روز قبل در نظر گرفته‌ایم و متغیر خروجی میزان متغیر ( $PM_{10}$ ) شهر یزد در یک روز بعد یا همان ۲۴ ساعت آینده یعنی  $X+1$  است که قصد داریم آن را پیش بینی کنیم. به عبارت دیگر میزان متغیر ( $PM_{10}$ ) شهر یزد در یک روز آینده را تابعی از میزان این متغیر در ۱، ۲ و ۳ روز قبل در نظر گرفته‌ایم که این عبارت در رابطه **Error!** **Reference source not found.** نمایش داده شده است.

$$X+1 = f(X-1, X-2, X-3) \quad (1)$$

## ۲-۴. شبکه عصبی مصنوعی

یک شبکه عصبی مصنوعی، شبکه‌ای از عناصر به هم مرتبط است. شبکه‌های عصبی را می‌توان با اغماض زیاد، مدل‌هایی الکترونیکی از ساختار عصبی مغز انسان نامید و پایه مدل‌های تحلیل عصبی بر شبیه‌سازی فعالیت‌های یک سلول عصبی نورون استوار است (Lawrence, 1994). در یک شبکه عصبی مصنوعی، نورون  $m$  ام اطلاعات ورودی خود را از طریق گره

(Jang, 1993).

حاشیه را از مقادیر آموزشی  $y$  داشته باشد ( Smola *et al.*, 2004; Vapnik, 2013).

## ۲-۶. رگرسیون بردار پشتیبان

ماشین بردار پشتیبان ( Support Vector Machine: SVM) یک نوع الگوریتم یادگیری ماشین نظارت شده است که برای طبقه‌بندی و برای تخمین و برآورد تابع برازش داده‌ها در مسائل رگرسیون نیز به کار می‌رود به طوری که کمترین خطا در طبقه‌بندی داده‌ها یا تابع برازش رخ دهد. این روش بر مبنای تئوری یادگیری آماری استوار است که از اصل کمینه سازی خطای ساختاری (Structural Risk Minimization: SRM) بهره می‌برد و موجب یک جواب بهینه سراسری می‌شود (Vapnik, 2013). هدف رگرسیون بردار پشتیبان (Support Vector Regression: SVR) که مدل رگرسیونی SVM می‌باشد این است که تابع  $F(x)$  را برای الگوهای آموزشی  $X$  طوری تشخیص دهد که بیشینه

## ۳. نتایج

۳-۱. پیش بینی ( $PM_{10}$ ) برای ۲۴ ساعت آینده با استفاده از شبکه عصبی

در این پژوهش برای پیش بینی آلودگی ذرات معلق کمتر از ۱۰ میکرون ( $PM_{10}$ ) شهر یزد در ۲۴ ساعت آینده، از شبکه عصبی MLP با شیوهی آموزش با نظارت با استفاده از الگوریتم لونیگ-مارکوات (LM) استفاده شده است، جزئیات شبکه عصبی MLP استفاده شده در **Error! Reference source not found.** نمایش داده شده است. معیارهای ارزیابی برای مراحل آموزش و آزمایش مدل ANN برای پیش بینی  $PM_{10}$  شهر یزد برای ۲۴ ساعت آینده، به شرح **Error! Reference source not found.** می‌باشند.

جدول ۱- جزئیات شبکه عصبی MLP استفاده شده در پیش بینی آلودگی ( $PM_{10}$ ) شهر یزد برای ۲۴ ساعت آینده

پارامتر	مقدار
ساختار شبکه عصبی	$3 \times 8 \times 1$
داده‌های آموزش و آزمون	۷۰ درصد ابتدایی داده‌ها و ۳۰ درصد انتهایی داده‌ها
تعداد تکرارها	۵۰۰
الگوریتم آموزش	LM
مقدار هدف	$1e-8$

جدول ۲- معیارهای ارزیابی مدل ANN در پیش بینی متغیر  $PM_{10}$  شهر یزد برای ۲۴ ساعت آینده

متغیر هدف	مرحله	$R^2$	RMSE	MSE	MAE	RSE	EVS
$PM_{10}$ شهر یزد	آموزش	۰/۹۴۲	۰/۰۴۵	۰/۰۰۲	۰/۰۳۳	۰/۰۵۷	۰/۷۶
آزمایش		۰/۹۷۸	۰/۰۲۶	۰/۰۰۱	۰/۰۲۴	۰/۰۳۲	۰/۸۵

۲-۳. پیش بینی ( $PM_{10}$ ) برای ۲۴ ساعت آینده با

استفاده از شبکه عصبی-فازی تطبیقی

در این پژوهش برای پیش بینی آلودگی ذرات معلق کمتر از ۱۰ میکرون ( $PM_{10}$ ) شهر یزد، از شبکه عصبی-فازی تطبیقی ANFIS با توابع عضویت از نوع گوسی

(Gaussmf) به تعداد ۵ عدد برای هر ورودی و داده های آموزش (۷۰ درصد) و داده های آزمون (۳۰ درصد) استفاده شده است. معیارهای ارزیابی برای مراحل آموزش و آزمایش مدل ANFIS برای پیش بینی  $PM_{10}$  شهر یزد برای ۲۴ ساعت آینده، به شرح **Error!** **Reference source not found.** می باشند.

جدول ۳- معیارهای ارزیابی مدل ANFIS در پیش بینی متغیر  $PM_{10}$  شهر یزد برای ۲۴ ساعت آینده

متغیر هدف	مرحله	$R^2$	RMSE	MSE	MAE	RSE	EVS
$PM_{10}$ شهر یزد	آموزش	۰/۹۷۹	۰/۰۲۷	۰/۰۰۱۱	۰/۰۲۴	۰/۰۳۲	۰/۸۵۵
	آزمایش	۰/۹۸۹	۰/۰۱۸	۰/۰۰۰۵	۰/۰۱۷	۰/۰۱۵	۰/۸۹۹

۳-۳. پیش بینی ( $PM_{10}$ ) برای ۲۴ ساعت آینده با

استفاده از رگرسیون بردار پشتیبان

در این پژوهش برای پیش بینی آلودگی ذرات معلق کمتر از ۱۰ میکرون ( $PM_{10}$ ) شهر یزد، از رگرسیون بردار پشتیبان SVR با تابع کرنل خطی استفاده شده است. جزئیات رگرسیون بردار پشتیبان SVR استفاده شده

مانند پارامتر جریمه C، پارامتر حاشیه خطا  $\epsilon$  و پارامتر مقیاس کرنل  $\sigma$  در **Error! Reference source not found.** نمایش داده شده است. معیارهای ارزیابی برای مراحل آموزش و آزمایش مدل SVR برای پیش بینی  $PM_{10}$  شهر یزد برای ۲۴ ساعت آینده، به شرح **Error!** **Reference source not found.** می باشند.

جدول ۴- جزئیات رگرسیون بردار پشتیبان SVR استفاده شده در پیش بینی آلودگی ( $PM_{10}$ ) شهر یزد برای ۲۴ ساعت آینده

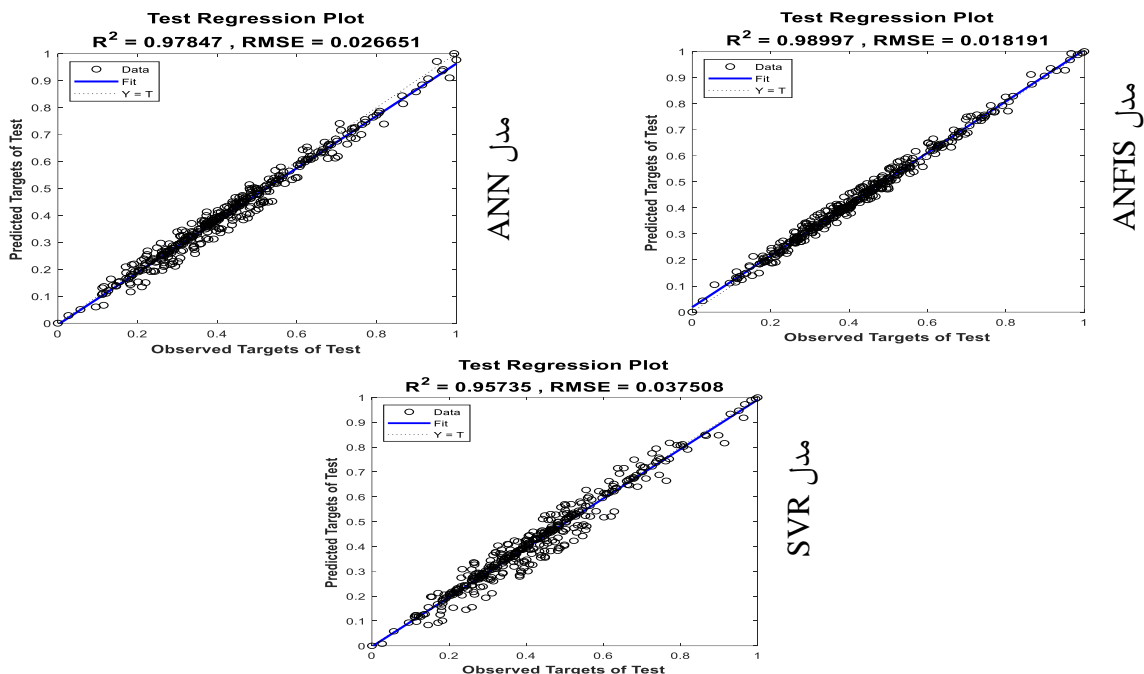
پارامتر	مقدار
تعداد ورودی ها و خروجی های SVR	۳ و ۱
پارامتر جریمه C، حاشیه خطا $\epsilon^6$ ، مقیاس کرنل $\sigma$	۱، ۰/۰۱۹۱۷۴، ۰/۰۱
داده های آموزش و آزمون	۷۰ درصد ابتدایی داده ها و ۳۰ درصد انتهایی داده ها

جدول ۵- معیارهای ارزیابی مدل SVR در پیش بینی متغیر  $PM_{10}$  شهر یزد برای ۲۴ ساعت آینده

متغیر هدف	مرحله	$R^2$	RMSE	MSE	MAE	RSE	EVS
$PM_{10}$ شهر یزد	آموزش	۰/۹۰۷	۰/۰۵۷	۰/۰۰۶۲	۰/۰۶۸	۰/۱۷۴	۰/۶۸۶
	آزمایش	۰/۹۵۷	۰/۰۳۷	۰/۰۰۱۴	۰/۰۲۵	۰/۰۴۵	۰/۷۹

۳-۴. نمودارهای پراکنش پیش بینی ( $PM_{10}$ ) شهر یزد برای ۲۴ ساعت آینده با روش های (ANN, ANFIS)

and SVR) در فاز آزمایش



شکل ۱- سه نمودار پراکنش میان مقادیر حقیقی و پیش بینی شده متغیر  $PM_{10}$  شهر یزد برای ۲۴ ساعت آینده توسط مدل SVR (فاز آموزش)

برای ۲۴ ساعت آینده با روش های (ANN, ANFIS and SVR) در فاز آزمایش

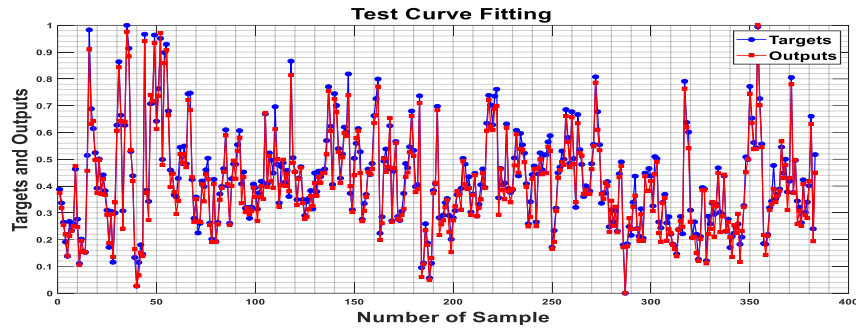
ANFIS and SVR) در فاز آزمایش

Error! Reference source not found.

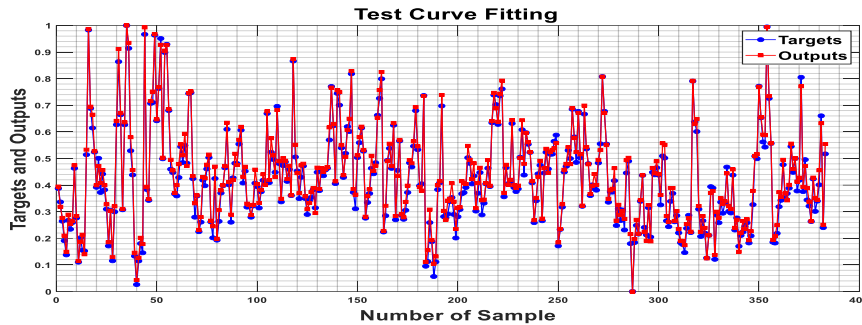
نشان دهنده میزان تطبیق و سازگاری مربوط به مقادیر مشاهداتی و پیش بینی شده متغیر هدف ( $PM_{10}$ ) شهر یزد برای ۲۴ ساعت آینده) در مرحله آزمایش توسط ۳ مدل می باشد. به خوبی از این نمودار ها مشخص است که هر ۳ مدل در تخمین مقادیر هدف با دقت بالا عمل کرده است، به طوری که این مقادیر بسیار نزدیک به مقدار واقعی آنها تخمین زده شده است و این مقادیر با یکدیگر تطابق بسیار بالایی دارند.

همانطور که از این نمودار های شکل ۱ مشخص است، خط آبی رنگ، نیمساز ربع اول در محور عمودی و افقی است. مقادیر متغیر هدف ( $PM_{10}$  شهر یزد) اندازه گیری شده بر روی محور X و مقادیر متغیر هدف پیش بینی شده توسط مدل های مذکور بر روی محور Y جانمایی شده اند. هر چقدر این مقادیر به نیمساز ربع اول (خط آبی رنگ) نزدیک تر باشند، بیانگر برازش عالی و پیش بینی دقیق تر متغیر هدف خواهد بود. به خوبی از این شکل ها مشخص است که درصد بالایی از مقادیر متغیر هدف به خط نیمساز ربع اول نزدیک می باشند.

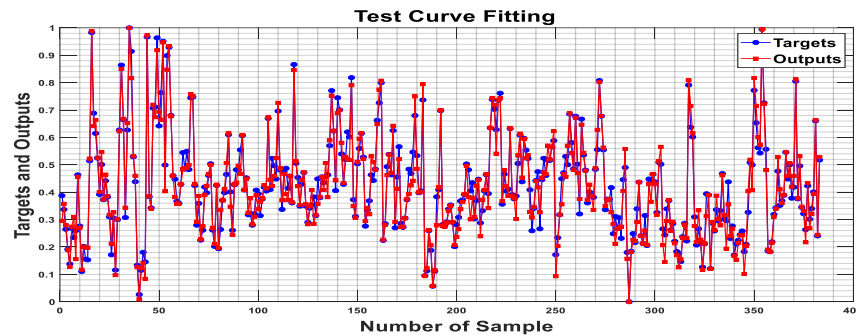
۳-۵. نمودار های تطابق پیش بینی ( $PM_{10}$ ) شهر یزد



مدل ANN



مدل ANFIS



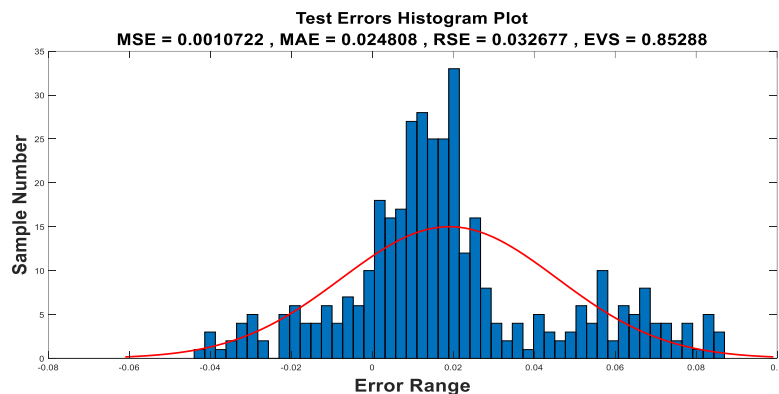
مدل SVR

شکل ۲. نمودار تطابق میان مقادیر حقیقی و پیش بینی شده متغیر  $PM_{10}$  شهر یزد برای ۲۴ ساعت آینده توسط مدل ANFIS (فاز آزمایش)

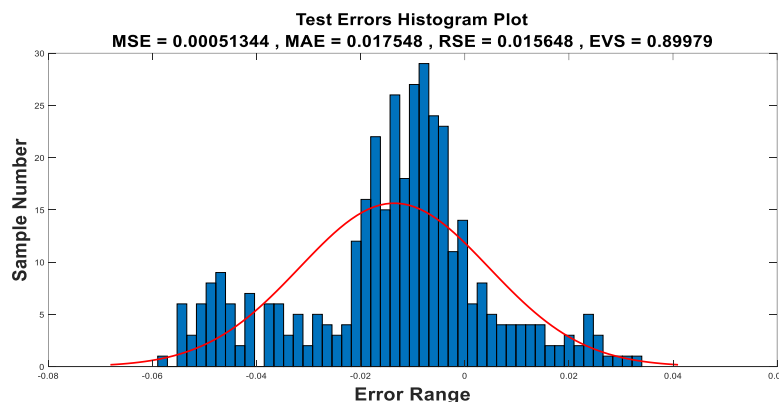
$X$  است. بازه‌ی خطاهای پیش بینی متغیر هدف توسط مدل ANN، بر روی محور  $X$  و تعداد نمونه‌های دارای این خطاها بر روی محور  $Y$  جانمایی شده‌اند. به عبارت دیگر از این نمودارها مشخص است مقدار این خطا برای تقریباً کل نمونه‌ها بسیار نزدیک به صفر می باشد که حاکی از دقت بالای ۳ مدل در پیش بینی متغیر هدف است.

۳-۶. نمودارهای هیستوگرام خطای پیش بینی  $(PM_{10})$  شهر یزد برای ۲۴ ساعت آینده با روش های  $(ANN, ANFIS \text{ and } SVR)$  در فاز آزمایش همانطور که از این نمودارها شکل ۳ مشخص است، مستطیل‌های آبی رنگ، نشان دهنده‌ی تعداد نمونه‌های دارای خطای پیش بینی نشان داده شده بر روی محور

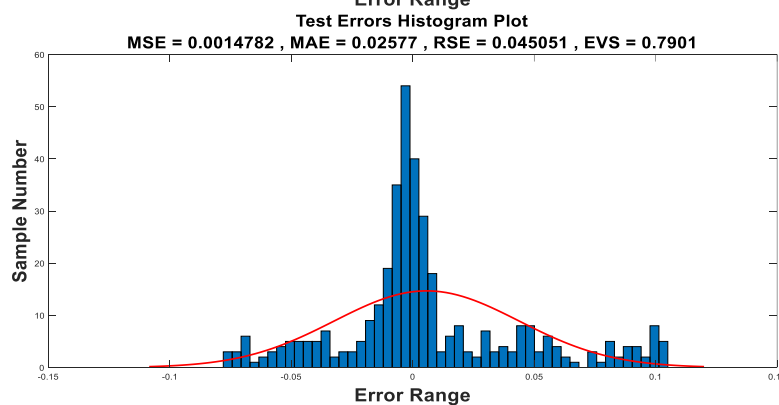




مدل ANN



مدل ANFIS



مدل SVR

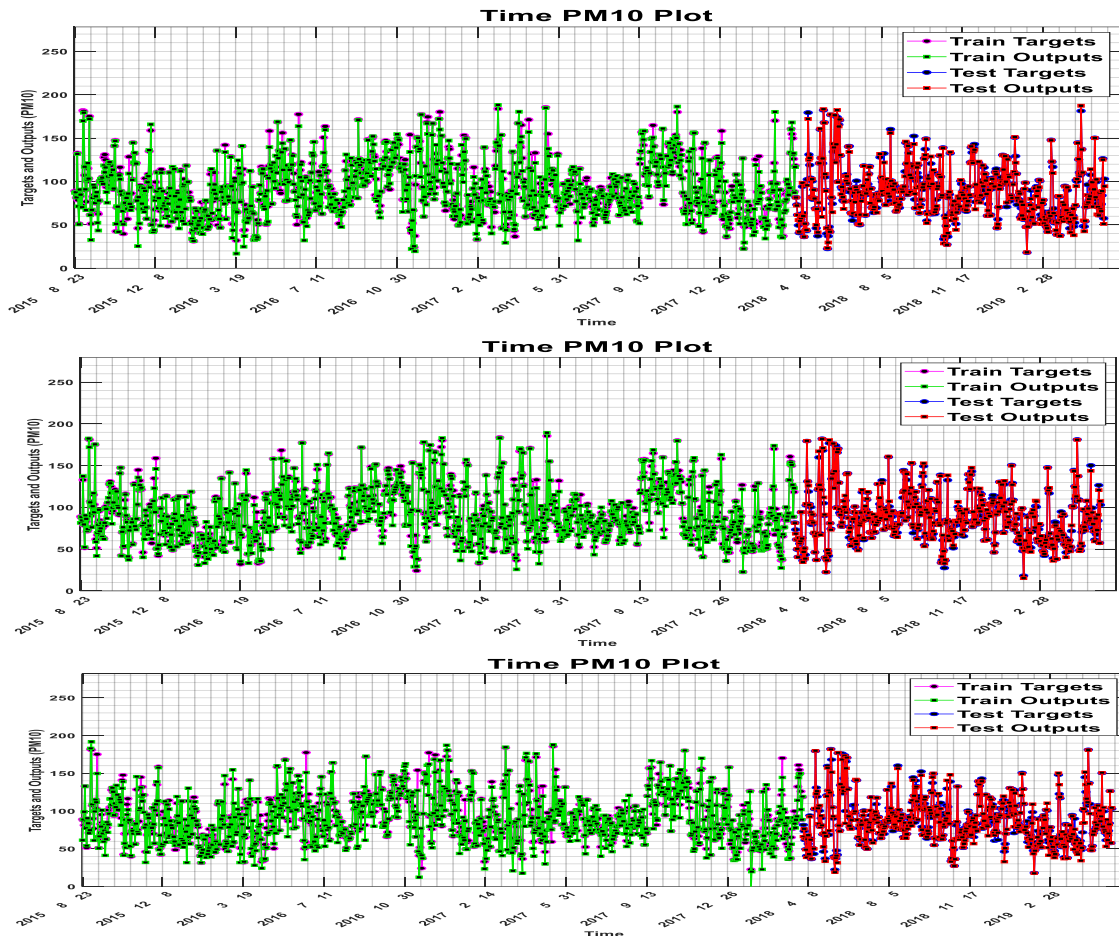
شکل ۳. هیستوگرام خطاهای بین مقادیر حقیقی و پیش بینی شده متغیر  $PM_{10}$  شهر یزد برای ۲۴ ساعت آینده توسط مدل ANN (فاز آزمایش)

شکل ۴ نشان دهنده میزان تطبیق و سازگاری مربوط به مقادیر واقعی متغیر ( $PM_{10}$ ) شهر یزد و مقادیر پیش بینی شده متغیر هدف ( $PM_{10}$  شهر یزد) توسط مدل ANN است که بر حسب زمان ترسیم شده است. ۷۰ درصد از ابتدای داده‌ها به عنوان مجموعه داده‌ی آموزش

۳-۷. نمودارهای ( $PM_{10}$ ) شهر یزد بر حسب زمان و مقایسه مقادیر واقعی و پیش بینی شده با استفاده از روش‌های (ANN, ANFIS and SVR) در فاز آزمایش

طوری که این مقادیر بسیار نزدیک به مقدار واقعی آنها تخمین زده شده است و این مقادیر با یکدیگر تطابق بسیار بالایی دارند.

و ۳۰ درصد انتهایی داده‌ها به عنوان مجموعه داده‌ی آزمایش در نظر گرفته شده است. به خوبی از این نمودار مشخص است که هر ۳ مدل در تخمین مقادیر هدف در فاز آموزش و آزمایش با دقت بالا عمل کرده است، به



شکل ۴- نمودار PM<sub>10</sub> شهر یزد بر حسب زمان و مقایسه مقادیر واقعی و پیش بینی شده توسط مدل ANN

پیش بینی (PM<sub>10</sub>) شهر یزد برای ۲۴ ساعت آینده به شرح جدول ۷ می باشند. از این جداول میتوان دریافت که نتایج پیش بینی در حالت آزمایش و آموزش بهم نزدیک و داری دقت بالایی هستند.

۳-۸. مقایسه نتایج پیش بینی (PM<sub>10</sub>) شهر یزد برای ۲۴ ساعت آینده با روش های ANN, ANFIS, SVR معیارهای ارزیابی برای مراحل آموزش و آزمایش برای

جدول ۷- مقایسه دقت مدل های یادگیری ماشین در پیش بینی (PM<sub>10</sub>) شهر یزد در ۲۴ ساعت آینده

متغیر هدف	نوع مدل	مرحله	R <sup>2</sup>	RMSE	MSE	MAE	RSE	EVS
PM <sub>10</sub> شهر یزد در ۲۴ ساعت آینده	ANN	آموزش	۰/۹۴۲	۰/۰۴۵	۰/۰۰۲	۰/۰۳۳	۰/۰۵۷	۰/۷۶
		آزمایش	۰/۹۷۸	۰/۰۲۶	۰/۰۰۱	۰/۰۲۴	۰/۰۳۲	۰/۸۵
	ANFIS	آموزش	۰/۹۷۹	۰/۰۲۷	۰/۰۰۱۱	۰/۰۲۴	۰/۰۳۲	۰/۸۵۵
		آزمایش	۰/۹۸۹	۰/۰۱۸	۰/۰۰۰۵	۰/۰۱۷	۰/۰۱۵	۰/۸۹۹
	SVR	آموزش	۰/۹۰۷	۰/۰۵۷	۰/۰۰۶۲	۰/۰۶۸	۰/۱۷۴	۰/۶۸۶
		آزمایش	۰/۹۵۷	۰/۰۳۷	۰/۰۰۱۴	۰/۰۲۵	۰/۰۴۵	۰/۷۹

#### ۴. بحث و نتیجه گیری

بر اساس گزارش سازمان بهداشت جهانی، ذرات معلق یکی از مهمترین آلاینده هوا در شهرهاست و بیش از هر آلاینده دیگر افراد را تحت تأثیر قرار می دهد (World Health, 2016). شواهد محکمی مبنی بر تاثیرگذاری شدید ذرات PM<sub>10</sub> بر روی آب، خاک، گیاهان و سلامت انسان وجود دارد (EPA, 2016) و می توانند باعث انواع وسیعی از اثرات بهداشتی مثل برونشیت، آسم، سرطان ریه و بیماری های قلبی و عروقی شوند (Adar et al., 2014). همانطور که اشاره شد برای پیش بینی غلظت آلاینده PM<sub>10</sub> از روش ها و ترکیب های مختلفی به همراه داده های مختلف از جمله هواشناسی استفاده شده است، مانند Amanollahi و همکاران (۲۰۲۰) از ۳ مدل MLR ، MLP و ANFIS به منظور پیش بینی غلظت PM<sub>10</sub> در هوای تهران استفاده نمودند و تجزیه و تحلیل داده ها نشان داد که در مقایسه با نتایج MLR و MLP، دقیق ترین نتایج را ANFIS دارا می باشد و یا Ghsemi و همکاران (۲۰۱۹) که از دو مدل ANFIS و ANN به منظور پیش بینی غلظت PM<sub>10</sub> در دو بازه

زمانی امروز و فردا در هوای شهر کرمانشاه استفاده کردند و مدل عصبی - فازی به عنوان مدل بهینه انتخاب شد و یا مطالعه Kaboodvandpour و همکاران (۲۰۱۵) که عملکرد مدل ANFIS را با MLR و ANN مقایسه نمودند و نتایج حاکی از برتری و دقت بالاتر مدل ANFIS در مقایسه با دو مدل دیگر بود. در کل اکثر مطالعات حاکی از برتری و دقت بالاتر روش های هوش مصنوعی نسبت به روش های خطی مانند رگرسیون بوده و در بین روش های هوش مصنوعی روش ANFIS دارای بهترین عملکرد می باشد. در این مطالعه سعی کردیم برای بررسی دقت پیش بینی آلاینده (PM<sub>10</sub>)، روش ماشین های بردار پشتیبان (SVR) که یک نوع روش هوش مصنوعی و ابزاری بسیار قدرتمند در زمینه پیش بینی محسوب می شود را با دو روش پرکاربرد ANFIS و ANN مقایسه کنیم. نتیجه گیری بر اساس نتایج این مطالعه براساس ۶ معیار ارزشیابی و اعتبارسنجی نشان می دهد که:

۱- مدل ANFIS در حالت پیش بینی ۲۴ ساعته با (MAE=0.017, RMSE=0.018, R2=0.989) ، (EVS=0.899 و MSE=0/0005, RSE=0.015)

۳- نتایج کلی نشان داد که با توجه به داده های ورودی هر سه مدل با دقتی بالای ۹۵ درصد توانایی بسیار خوبی در پیش بینی غلظت آلاینده ذرات PM<sub>10</sub> را برای ۲۴ ساعت آینده دارا می باشند و در مقیاس عملیاتی مدیران شهر می توانند بر اساس مدل ها و پیش بینی های با دقت بالای ارائه شده تصمیمات مناسب و به موقعی را در جهت حفظ سلامت عمومی قبل از شروع وضعیت هشدار غلظت آلاینده PM<sub>10</sub> بگیرند.

دارای بالاترین دقت (در حدود ۹۹ درصد)، کمترین خطا و بهترین شبیه سازی در پیش بینی از سایر مدل ها است و در مقایسه با مدل SVR دقت بهتر و بالاتری را دارا می باشد.

۲- بعد از مدل ANFIS به ترتیب مدل ANN با (R<sup>2</sup> = 0.978, MSE = 0/001, RMSE = 0.026, RSE SVR = 0.032, MAE = 0.024, EVS = 0.85) و برای پیش بینی ۲۴ ساعت بعد میزان غلظت آلاینده PM<sub>10</sub> بهترین کارایی را دارند.

## References

Adar, Sara D., Filigrana, Paola A., Clements, Nicholas, & Peel, Jennifer L. (2014). Ambient Coarse Particulate Matter and Human Health: A Systematic Review and Meta-Analysis. *Current Environmental Health Reports*, 1 (3), 258-274. doi: 10.1007/s40572-014-0022-z

Amanollahi, Jamil, & Ausati, Shadi. (2020). Validation of linear, nonlinear, and hybrid models for predicting particulate matter concentration in Tehran, Iran. *Theoretical and Applied Climatology*, 140(1), 709-717. doi: 10.1007/s00704-020-03115-5

Azadeh, A., Saberi, M., Anvari, M., Azaron, A., & Mohammadi, M. (2011). An adaptive network based fuzzy inference system-genetic algorithm clustering ensemble algorithm for performance assessment and improvement of conventional power plants. *Expert Systems with Applications*, 38(3), 2224-2234. doi: https://doi.org/10.1016/j.eswa.2010.08.010

Buragohain, Mrinal, & Mahanta, Chitralakha. (2008). A novel approach for ANFIS modelling based on full factorial design. *Applied Soft Computing*, 8(1), 609-625. doi: https://doi.org/10.1016/j.asoc.2007.03.010

EPA, US. (2016). *Health and Environmental Effects of Particulate Matter (PM)*. Retrieved from https://www.epa.gov/pm-pollution/health-and-environmental-effects-particulate-matter-pm.

Forster, P., Ramaswamy, V., Artaxo, P., Berntsen, T., Betts, R., Fahey, D. W., . . . Van Dorland, R. (2007). *Changes in Atmospheric Constituents and in Radiative Forcing Chapter 2*. United Kingdom: Cambridge University Press.

Ghasemi, Afsaneh, & Amanollahi, Jamil. (2019). Integration of ANFIS model and forward selection method for air quality forecasting. *Air Quality, Atmosphere & Health*, 12(1), 59-72. doi: 10.1007/s11869-018-0630-0

Hamanaka, R. B., & Mutlu, G. M. (2018). Particulate Matter Air Pollution: Effects on the Cardiovascular System. *Front Endocrinol (Lausanne)*, 9, 680. doi: 10.3389/fendo.2018.00680

Haykin, Simon. (1994). *Neural networks: a comprehensive foundation*: Prentice Hall PTR.

Jang, J-SR. (1993). ANFIS: adaptive-network-based fuzzy inference system. *IEEE transactions on systems, man, and cybernetics*, 23(3), 665-685.

Kaboodvandpour, Shahram, Amanollahi, Jamil, Qhavami, Samira, & Mohammadi, Bakhtiyar. (2015). Assessing the accuracy of multiple regressions, ANFIS, and ANN models in predicting dust storm occurrences in Sanandaj, Iran. *Natural Hazards*, 78(2), 879-893.

Kim, H. S., Park, I., Song, C. H., Lee, K., Yun, J. W.,

- Kim, H. K., . . . Han, K. M. (2019). Development of a daily PM10 and PM2.5 prediction system using a deep long short-term memory neural network model. *Atmos. Chem. Phys.*, *19*(20), 12935-12951. doi: 10.5194/acp-19-12935-2019
- Kim, J. (2019). Particulate Matter Mortality Rates and Their Modification by Spatial Synoptic Classification. *Int J Environ Res Public Health*, *16*(11). doi: 10.3390/ijerph16111904
- Lawrence, Jeannette. (1994). Introduction to Neural Networks: Design, Theory, and Applications (California Scientific Software, Nevada City, CA).
- Liu, Huixiang, Li, Qing, Yu, Dongbing, & Gu, Yu. (2019). Air Quality Index and Air Pollutant Concentration Prediction Based on Machine Learning Algorithms. *Applied Sciences*, *9*, 4069. doi: 10.3390/app9194069
- Maleki, Heidar, Sorooshian, Armin, Goudarzi, Gholamreza, Baboli, Zeynab, Tahmasebi Birgani, Yaser & Rahmati, Mojtaba. (2019). Air pollution prediction by using an artificial neural network model. *Clean Technologies and Environmental Policy*. doi: 10.1007/s10098-019-01709-w
- Omidvarborna, Hamid, Kumar, Ashok, & Kim, Dong-Shik. (2015). Recent studies on soot modeling for diesel combustion. *Renewable and Sustainable Energy Reviews*, *48*, 635-647. doi: https://doi.org/10.1016/j.rser.2015.04.019
- Paschalidou, Anastasia K., Karakitsios, Spyridon, Kleanthous, Savvas, & Kassomenos, Pavlos A. (2011). Forecasting hourly PM10 concentration in Cyprus through artificial neural networks and multiple regression models: implications to local environmental management. *Environmental Science and Pollution Research*, *18*(2), 316-327. doi: 10.1007/s11356-010-0375-2
- Pérez, Noemí, Pey, Jorge, Cusack, Michael, Reche, Cristina, Querol, Xavier, Alastuey, Andrés, & Viana, Mar. (2010). Variability of Particle Number, Black Carbon, and PM10, PM2.5, and PM1 Levels and Speciation: Influence of Road Traffic Emissions on Urban Air Quality. *Aerosol Science and Technology*, *44*(7), 487-499. doi: 10.1080/02786821003758286
- Schlink, Uwe, Dorling, Stephen, Pelikan, Emil, Nunnari, Giuseppe, Cawley, Gavin, Junninen, Heikki, Doyle, Martin. (2003). A rigorous inter-comparison of ground-level ozone predictions. *Atmospheric Environment*, *37*(23), 3237-3253. doi: https://doi.org/10.1016/S1352-2310(03)00330-3
- Scholkopf, Bernhard, Sung, Kah-Kay, Burges, Christopher JC, Girosi, Federico, Niyogi, Partha, Poggio, Tomaso, & Vapnik, Vladimir. (1997). Comparing support vector machines with Gaussian kernels to radial basis function classifiers. *IEEE transactions on Signal Processing*, *45*(11), 2758-2765.
- Smola, Alex J, & Schölkopf, Bernhard. (2004). A tutorial on support vector regression. *Statistics and computing*, *14*(3), 199-222.
- Vapnik, Vladimir. (2013). *The nature of statistical learning theory*: Springer science & business media.
- Wasley, Andrew; , Heal, Alexandra; , Harvey, Fiona; , & Lainio, Mie (2019). Revealed: UK government failing to tackle rise of serious air pollutant. *The Guardian*.
- World Health, Organization. (2016). *Ambient air pollution: a global assessment of exposure and burden of disease*. Geneva: World Health Organization.
- World Health, Organization. (2017). (Evolution of WHO air quality guidelines: past, present and future (pp. 39): Copenhagen: WHO Regional Office for Europe.
- Yadav, V., & Nath, S. (2019). Novel hybrid model for daily prediction of PM10 using principal component analysis and artificial neural network. *International Journal of Environmental Science and Technology*, *16*(6), 2839-2848. doi: 10.1007/s13762-018-1999-x
- Zhang, G. Peter. (2003). Time series forecasting using a hybrid ARIMA and neural network model. *Neurocomputing*, *50*, 159-175. doi: https://doi.org/10.1016/S0925-2312(01)00702-0