



نسخه از دور

GIS ایران



سنجش از دور و GIS ایران
Iranian Remote Sensing & GIS

سال نهم، شماره سوم، پاییز ۱۳۹۶
Vol.9, No. 3, Autumn 2017

۷۰-۴۵

پیش‌بینی آلودگی هوای شهر تهران با استفاده از ترکیب شبکه عصبی فازی- تطبیقی و آنالیز مؤلفه اصلی

زینب قائمی^{۱*}، محمد طالعی^۲، مهدی فرنقی^۳، قاسم جوادی^۴

۱. دانشجوی دکتری سیستم اطلاعات مکانی، دانشکده مهندسی ژئودزی و ژئوماتیک دانشگاه صنعتی خواجه نصیرالدین طوسی
۲. دانشیار دانشکده مهندسی ژئودزی و ژئوماتیک، دانشگاه صنعتی خواجه نصیرالدین طوسی
۳. استادیار دانشکده مهندسی ژئودزی و ژئوماتیک، دانشگاه صنعتی خواجه نصیرالدین طوسی
۴. عضو هیئت علمی گروه مهندسی عمران، نقشه برداری، دانشکده فنی و مهندسی، دانشگاه بجنورد

تاریخ پذیرش مقاله: ۱۳۹۶/۸/۱۳

تاریخ دریافت مقاله: ۱۳۹۶/۰۳/۷

چکیده

رشد و توسعه شهری و افزایش استفاده از وسایل نقلیه در سال‌های اخیر منجر به افزایش آلودگی هوا، به‌ویژه در شهرهای بزرگ و صنعتی، شده است. با توجه به آثار نامطلوب آلودگی هوا در سلامت انسان‌ها و دیگر جانداران، پیش‌بینی و مدل‌سازی این پدیده پیچیده از دغدغه‌های اصلی محققان در سال‌های اخیر بوده است. هدف این تحقیق طراحی سیستمی است به‌منظور پیش‌بینی آلودگی هوا، طی ۲۴ ساعت آینده، تا با شناسایی مناطق آلوده، به مدیران و برنامه‌ریزان شهری برای کنترل و کاهش میزان آلاینده‌ها کمک کند. در سیستم طراحی شده، از ترکیب آنالیز مؤلفه اصلی و شبکه عصبی فازی- تطبیقی (PCA-ANFIS)، به‌منظور پیش‌بینی آلودگی هوا در فصل‌های متفاوت، استفاده شده است. در این سیستم، داده‌های هواشناسی و غلظت آلاینده‌ها در روزهای گذشته، برای پیش‌بینی آلودگی هوای شهر تهران در ۲۴ ساعت آینده، به‌کار رفته است. همچنین، از پارامترهای مکانی مانند ارتفاع، توپوگرافی سطح زمین و فاصله از جاده به‌منظور مدل‌سازی مکانی پراکندگی آلودگی هوا استفاده شده است. نتایج حاصل از مقایسه روش ترکیبی PCA-ANFIS با روش ANFIS دقت و سرعت بالاتر مدل ترکیبی طراحی شده را، به‌نسبت روش ANFIS در پیش‌بینی آلودگی هوا طی ۲۴ ساعت آینده، بیان می‌کند.

کلیدواژه‌ها: پیش‌بینی آلودگی هوای شهری، شبکه عصبی فازی تطبیقی، آنالیز مؤلفه اصلی، سامانه اطلاعات مکانی.

* نویسنده عهده‌دار مکاتبات: تهران، خیابان ولیعصر، تقاطع میرداماد، روبه‌روی ساختمان اسکان، دانشکده ژئودزی و ژئوماتیک دانشگاه صنعتی خواجه نصیرالدین طوسی. تلفن:

۱- مقدمه

توسعه شهرنشینی، گسترش صنایع و افزایش بی‌رویه استفاده از وسایل نقلیه منجر به گسترش آلودگی هوا، به‌ویژه در شهرهای بزرگ و صنعتی، شده است. آلودگی هوا، به‌دلیل آثار زیانبارش در سلامت انسان‌ها و سایر جانداران، اهمیت بسزایی دارد (Brunekreef and Holgate, 2002; Chan-Yeng, 2000; Combarro et al., 2013). تا کنون مطالعات بسیاری در زمینه بررسی ارتباط میان آلودگی هوا و بیماری‌های قلبی-عروقی و تنفسی صورت گرفته (Gorai et al., 2014; Brook et al., 2004) که بیانگر تأثیرات منفی آلودگی هوا در سلامت افراد است. این امر موجب شده است در سال‌های اخیر، دانشمندان به‌دنبال طراحی سیستم‌هایی برای پیش‌بینی آلودگی هوای شهری برآیند تا به‌کمک آن بتوانند میزان آلودگی هوا را پیش‌بینی کنند و نتایج را در اختیار مدیران و برنامه‌ریزان شهری و نیز عموم افراد قرار دهند.

آلودگی هوا از جمله پدیده‌های پیچیده‌ای است که به‌صورت زمانی و مکانی تغییر می‌کند و عوامل بسیاری در آن دخیل‌اند. همبستگی میان این پارامترها موجب می‌شود نتایج به‌دست‌آمده از مدل‌سازی و پیش‌بینی آلودگی هوا قابلیت اطمینان بالایی نداشته باشند (Sahoo et al., 2015; Nia et al., 2015). همچنین، تعداد بالای پارامترهای ورودی زمان پردازش‌ها را نیز افزایش می‌دهد. بنابراین، انتخاب مناسب عوامل ورودی نقش مهمی در زمان پردازش و دقت حاصل دارد. تا کنون، روش‌های گوناگونی برای انتخاب داده‌های ورودی مطرح شده است. یکی از این روش‌ها کاهش ابعاد داده‌های^۱ ورودی با استفاده از تحلیل مؤلفه اصلی است. با استفاده از تحلیل مؤلفه‌های اصلی می‌توان متغیرهای مستقل همبسته بسیاری را با تعداد محدودی متغیر مستقل جدید، که مؤلفه‌های اصلی نام دارند و به یکدیگر وابسته نیستند، جایگزین کرد (Statheropoulos et al., 1998).

پیچیدگی رفتار آلودگی هوا، از سوی دیگر، موجب می‌شود روش‌های آمار معمول قادر به مدل‌سازی این پدیده پیچیده و غیرخطی نباشند (Moussiopoulos et al., 1995; Boznar et al., 1993). به همین دلیل، دانشمندان در پی روش‌های نوینی برای مدل‌سازی این پدیده پیچیده و غیرخطی‌اند. در این میان، روش‌های مبتنی بر هوش مصنوعی^۲، به‌دلیل توانایی بالا در مدل‌سازی مسائل غیرخطی و پیچیده، در پیش‌بینی آلودگی هوا مورد استفاده گسترده‌ای قرار گرفته‌اند (Kolehmainen et al., 2000). برای نمونه، در تحقیقی، فرناندو و همکاران^۳ (۲۰۱۲) از شبکه عصبی برای طراحی سیستم هشدار آلودگی هوا استفاده کرده‌اند. در این سیستم، غلظت آلاینده PM₁₀، با استفاده از داده‌های هواشناسی و مقادیر PM₁₀ در روزهای گذشته، پیش‌بینی شده و نتایج حاصل با سیستم محلی موجود در منطقه مقایسه شده است که عملکرد بهتر شبکه عصبی در پیش‌بینی میزان غلظت PM₁₀ را بیان می‌کند. پرز و همکاران^۴ (۲۰۰۰) از داده‌های سرعت و جهت باد به‌همراه داده‌های مربوط به غلظت آلاینده‌ها در روزهای قبل، برای پیش‌بینی غلظت PM₁₀، با استفاده از شبکه عصبی بهره برده‌اند. در این تحقیق، نتایج سه روش ماندگاری^۵، رگرسیون خطی و شبکه عصبی چندلایه با یکدیگر مقایسه شده‌اند. نتایج نشان می‌دهد شبکه عصبی با سه لایه دقتی بیشتر از دو مدل دیگر حاصل می‌کند.

اما با وجود توانایی بالای شبکه‌های عصبی در مدل‌سازی مسائل غیرخطی، این روش‌ها همچنان با محدودیت‌هایی همچون ناهمگرایی به بهینه جهانی^۶ مواجه‌اند (Singh et al., 2013; Niska et al., 2004). این مشکل به‌دلیل استفاده از روش‌های سعی و خطا

- | | |
|--------------------|---------------------------------|
| 1. data reduction | 2. artificial intelligence (AI) |
| 3. Fernando et al. | 4. Pérez et al. |
| 5. persistence | 6. global minimum |

بنابراین، با انتخاب تعداد بهینه خوشه‌ها، تعداد بهینه قوانین فازی نیز مشخص خواهد شد. برای انتخاب تعداد خوشه‌ها، از یک روش پارتیشن‌بندی اکتشافی^۶ برای تقسیم‌بندی ورودی خروجی اولیه (چند ورودی و یک خروجی)^۷ به چند زیردسته استفاده می‌شود. بنابراین، هر چند ورودی و یک خروجی به چندین تک‌ورودی-تک‌خروجی^۸ تبدیل می‌شود. برای هر یک از این زیردسته‌ها، ارتباط میان ورودی و خروجی با استفاده از تکنیک fuzzy curve fitting به دست می‌آید و یک تابع خطی یک‌متغیره برای هر یک از زیردسته‌ها حاصل می‌شود. در میان این توابع، تابعی انتخاب می‌شود که به بهترین شکل خروجی را تخمین می‌زند. در نتیجه، دقت کلی مجموعه ورودی‌ها و خروجی به دست می‌آید. این کار آن قدر تکرار می‌شود تا بهترین دقت حاصل شود. زمانی که روش پارتیشن‌بندی اکتشافی به پایان رسید، تعداد زیردسته‌ها به‌منزله تعداد قوانین فازی بهینه برای ANFIS در نظر گرفته می‌شود. متغیرهای ورودی انتخاب‌شده برای پارتیشن‌بندی در روش اکتشافی گزینه‌ای برای متغیرهای ورودی در نظر گرفته می‌شوند و اهمیت نسبی آنها براساس تعداد دفعات تقسیم‌بندی محاسبه می‌شود (برای اطلاعات بیشتر در زمینه فرمول‌ها و روابط استفاده‌شده رک: Shahraini et al., 2015). نتایج این تحقیق نشان می‌دهد استفاده از این روش موجب افزایش ۱۰ درصدی دقت حاصل شده است. ساهو و همکاران (۲۰۱۵) از ANFIS برای پیش‌بینی شاخص کیفیت آب^۹ رودخانه‌ای در نزدیکی مناطق صنعتی استفاده کرده‌اند. همچنین در این تحقیق، از تحلیل آنالیز مؤلفه اصلی

برای انتخاب تعداد لایه‌های پنهان و گره‌ها ایجاد می‌شود. همچنین، شبکه‌های عصبی قادر به تفسیر و تحلیل اطلاعات زبانی^۱ نیستند (Nunnari, 2000). مدل‌های فازی تا حدودی توانسته‌اند این محدودیت شبکه‌های عصبی را برطرف کنند. این مدل‌ها، با استخراج دانش موجود در داده‌های ورودی و در نظر گرفتن عدم اطمینان، قادرند اطلاعات زبانی را تفسیر کنند (Fisher, 2003 McKone and). به همین دلیل، در سال‌های اخیر، مدل‌های ترکیبی که بر مبنای ترکیبی از شبکه‌های عصبی، روش‌های فازی و محاسبات تکاملی^۲ باشند برای حل و پیش‌بینی مسائل پیچیده و غیرخطی توسعه یافته‌اند. سیستم استنتاج عصبی فازی-تطبیقی (ANFIS)^۳، که از ترکیب سیستم‌های استنتاج فازی و شبکه‌های عصبی مصنوعی تشکیل شده است، ابزاری قدرتمند برای پیش‌بینی پدیده‌های پیچیده را فراهم می‌کند. این سیستم هم‌زمان از قابلیت‌های تفسیر زبانی مدل‌های فازی و قابلیت خودآموزی شبکه‌های عصبی بهره می‌برد (Jang, 1993). تا کنون تحقیقات متفاوتی در این زمینه انجام شده است (Jane and Khare, 2010; Morabito and Versaci, 2003; Taylan, 2017). در این میان، می‌توان به تحقیق خزاعی و همکاران (۲۰۱۳) اشاره کرد که از دو روش سوگونو و ممدانی برای استخراج قوانین فازی استفاده و نتایج حاصل از مدل ANFIS را برای این قوانین، در پیش‌بینی غلظت مونوکسید کربن، با یکدیگر مقایسه کرده‌اند. در تحقیقی دیگر، شهرایینی و همکاران (۲۰۱۵) ساختاری جدید برای پیش‌بینی آلودگی هوا روی ایستگاه‌های سنجش آلودگی به‌روش ANFIS عرضه کرده‌اند. در این تحقیق، از یک روش خوشه‌بندی^۴ برای انتخاب پارامترهای ورودی و نیز انتخاب تعداد قوانین فازی استفاده شده است. روش خوشه‌بندی فازی C-Mean^۵ برای خوشه‌بندی داده‌های آموزشی به‌کار رفته است. در این روش، هر خوشه به‌منزله یک قانون در نظر گرفته می‌شود.

1. linguistic information
2. evolutionary
3. adaptive nero fuzzy inference system
4. clustering
5. fuzzy C-mean clustering
6. heuristic partitioning method
7. multi inputs-single output
8. single input-single output
9. water quality index

استفاده شده است. طراحی این سیستم از سه مرحله اصلی تشکیل شده است: در مرحله اول، ابعاد داده‌های ورودی که شامل یازده متغیر ورودی است، با استفاده از آنالیز مؤلفه اصلی به پنج متغیر کاهش پیدا می‌کند. کاهش ابعاد داده‌های ورودی موجب کاهش زمان پردازش‌ها می‌شود. سپس از سیستم استنتاج عصبی-فازی تطبیقی برای پیش‌بینی کلاس آلودگی هوا برای ۲۴ ساعت آینده استفاده می‌شود. به‌منظور ارزیابی کارایی روش طراحی شده، نتایج حاصل از این روش و همچنین، روش ANFIS با مقادیر واقعی مقایسه می‌شوند تا بررسی شود نتایج کدام روش به واقعیت نزدیک‌تر است. در نهایت، نتایج حاصل از سیستم استنتاج عصبی-فازی تطبیقی جهت پهنه‌بندی آلودگی هوا در سطح شهر تهران به‌کار می‌رود. همچنین به‌منظور مدل‌سازی دقیق‌تر آلودگی هوا و با توجه به رفتار متفاوت آلودگی هوا در فصل‌های گوناگون، مدل‌سازی برای فصل‌های متفاوت انجام و نتایج به‌صورت فصلی عرضه می‌شود.

در ادامه، ساختار مقاله به این صورت خواهد بود: در بخش ۲، منطقه مطالعاتی و داده‌های استفاده‌شده معرفی می‌شود. همچنین شیوه نرمال‌سازی داده‌ها و مبانی نظری تحلیل مؤلفه اصلی و سیستم استنتاج عصبی-فازی تطبیقی در این بخش توضیح داده شده است. سیستم طراحی شده برای پیش‌بینی آلودگی هوای شهری در بخش ۳ تشریح شده است. نتایج حاصل از اجرای مدل ارائه‌شده در بخش ۴ بحث می‌شود و در بخش ۵ نتیجه‌گیری و پیشنهادات مطرح می‌شود.

۲- مواد و روش‌ها

در این بخش منطقه مورد مطالعه و داده‌های

برای حذف همبستگی میان داده‌های ورودی و کاهش ابعاد داده‌های ورودی استفاده شده است. جین و همکارش^۱ نیز (۲۰۱۰) روش عصبی-فازی را برای پیش‌بینی یک‌ساعته غلظت مونوکسید کربن در دو منطقه از شهر دهلی هند استفاده کردند. در این تحقیق، از داده‌های دو سال گذشته هواشناسی، به‌همراه داده‌های ترافیکی، برای پیش‌بینی غلظت این گاز استفاده شده است. در این تحقیق سه روش دلفی، ماتریس همبستگی و تحلیل مؤلفه اصلی، برای استخراج پارامترهای مؤثر در غلظت مونوکسید کربن، به‌کار رفته که نتایج هر سه روش تقریباً مشابه بوده است. نوع و تعداد توابع فازی در این تحقیق به‌صورت تجربی و به‌روش سعی و خطا انتخاب و در نهایت، چهار ورودی با سه تابع فازی در نظر گرفته شده و همچنین، از تابع گوسین استفاده شده است. آموزش مدل روی یک ایستگاه برای هریک از فصل‌ها انجام شده و سپس مدل آموزش‌دیده روی ایستگاه دیگری سنجیده شده است. نتایج حاصل عملکرد مناسب مدل استفاده‌شده را نشان می‌دهد. نیز، مورابیتو و ورساچی^۲ (۲۰۰۳) یک سیستم عصبی-فازی برای پیش‌بینی میزان غلظت هیدروکربن توسعه دادند. در این تحقیق، از داده‌های توپوگرافی و ترافیک در کنار داده‌های غلظت آلاینده‌ها و داده‌های هواشناسی برای بهبود مدل استفاده شده است. در این تحقیق، ابتدا از شبکه عصبی و روش پرسپترون چندلایه^۳ برای پیش‌بینی غلظت هیدروکربن استفاده شده است. سپس از سیستم استنتاج فازی^۴ برای استخراج قوانین و تفسیر ارتباط میان ورودی‌ها بهره گرفته شده است. در نهایت، خروجی این مراحل برای ساخت یک سیستم خبره برای پیش‌بینی غلظت هیدروکربن‌ها به‌کار گرفته شده است.

هدف از این تحقیق طراحی سیستمی برای پیش‌بینی آلودگی هوای شهر تهران در ۲۴ ساعت آینده است. در این سیستم از ترکیب آنالیز مؤلفه اصلی و سیستم استنتاج عصبی-فازی تطبیقی برای ایجاد مدل ترکیبی PCA-ANFIS برای پیش‌بینی آلودگی هوا

1. Jain and Khare

2. Morabito and Versaci

3. multi-layer perceptron

4. fuzzy inference system

مدیران قرار دهد. همچنین، با آگاه‌سازی عموم و به‌ویژه افراد بیمار، از گسترش بیماری‌های ناشی از آلودگی هوا جلوگیری می‌شود.

شکل ۱ موقعیت جغرافیایی شهر تهران را به‌همراه شیوه توزیع ایستگاه‌های سنجش آلودگی هوا و هواشناسی در سطح شهر نمایش می‌دهد.

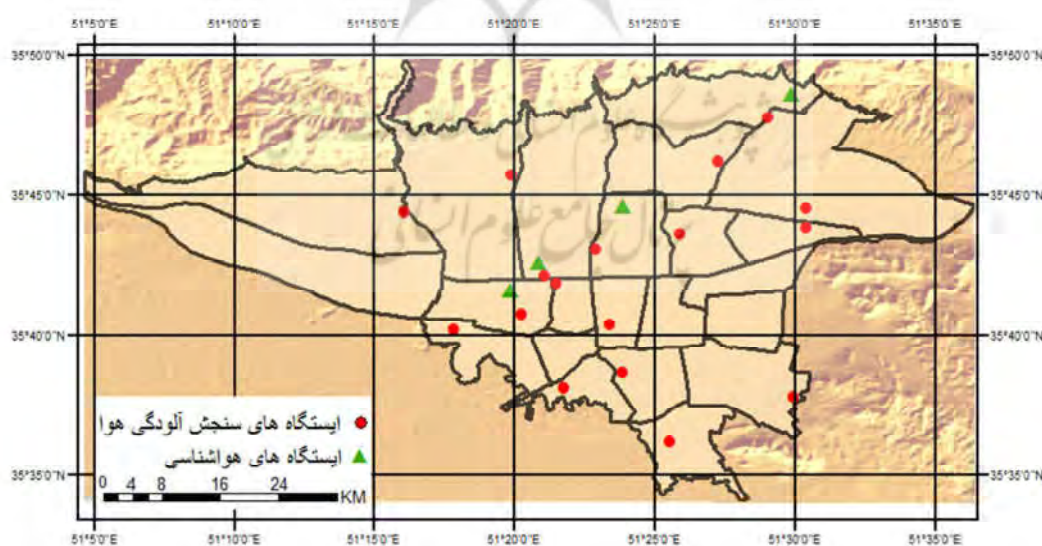
۲-۲- شاخص آلودگی هوا^۱

شاخص کیفیت هوا از جمله شاخص‌های اصلی در بیان میزان آلودگی هواست که امکان تفسیر آلودگی هوا را، با استفاده از مقداری عددی، فراهم می‌کند. به همین منظور، در این تحقیق، از این شاخص برای پیش‌بینی آلودگی هوای شهر تهران استفاده شده است. داده‌های ساعتی مربوط به غلظت آلاینده‌های SO_2 ، NO_2 ، PM_{10} ، O_3 و CO مربوط به ۲۱ ایستگاه سنجش آلودگی هوا از شرکت کنترل کیفیت هوای تهران در سال ۱۳۹۳ دریافت شده است. شاخص کیفیت هوا با استفاده از غلظت این آلاینده‌ها و با استفاده از رابطه (۱) محاسبه شده است (Mintz, 2012). پس از محاسبه شاخص

استفاده‌شده معرفی می‌شوند. همچنین، مبانی نظری روش‌های استفاده‌شده در این تحقیق در این بخش تشریح خواهد شد.

۲-۱- منطقه مطالعاتی و داده‌های مورد استفاده

شهر تهران مرکز سیاسی و تجاری و پایتخت کشور ایران است که بالغ بر ۸.۵ میلیون نفر در آن ساکن‌اند. تهران از سمت شمال با رشته‌کوه‌های البرز و از سمت جنوب با دشت کویر احاطه شده است. محصوربودن در بین کوه‌ها از سه طرف، که مانع خروج آلودگی‌ها از شهر می‌شود، از یک‌سو، و افزایش بی‌رویه استفاده از وسایل نقلیه و گسترش صنایع، از سوی دیگر، از عوامل اصلی آلودگی هوا در شهر تهران به‌شمار می‌رود. این نکته باعث شده در برخی از روزهای سال، میزان آلودگی از حد استاندارد بالاتر رود و موجب تعطیلی کامل شهر شود؛ بنابراین، ضرورت طراحی سیستمی برای پیش‌بینی آلودگی هوای شهر تهران نمایان شده است. طراحی چنین سیستمی می‌تواند اطلاعات لازم را برای اقدامات مورد نیاز جهت کنترل آلودگی در اختیار



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

1. air quality index (AQI)

است (Ghaemi et al., 2016).

۲-۴- نرمال‌سازی داده‌ها

داده‌های آلودگی هوا، هواشناسی و داده‌های مکانی پیش از اینکه در جایگاه ورودی به مدل معرفی شوند، نرمال می‌شوند تا مرحله آموزش مدل بهتر انجام شود (Kumar and Goyal, 2011). نرمال‌سازی داده‌ها تمامی داده‌ها را بی‌واحد و به بازه ۰ تا ۱ نگاشت می‌کند. به‌منظور نرمال کردن داده‌ها، از رابطه (۲) استفاده شده است:

$$X_{Normal} = \frac{(x_{max}-x)}{(x_{max}-x_{min})} \quad (2)$$

در این رابطه، x داده ورودی و x_{min} و x_{max} به ترتیب، بیشترین و کمترین مقدار داده‌اند.

۲-۵- آزمون بارتلت^۱ و KMO

دو آزمون بارتلت و KMO^۲، برای بررسی مناسب بودن داده‌ها برای تحلیل مؤلفه اصلی، استفاده می‌شود (Williams et al., 2010). از آزمون بارتلت برای بررسی کافی بودن نمونه‌ها در تحلیل عاملی اکتشافی^۳ استفاده می‌شود. تحلیل عاملی روشی ریاضی برای کاهش ابعاد داده‌هاست. در این آزمون، فرض صفر بدین معناست که نمونه‌های گرفته‌شده کافی نیستند و در مقابل، فرض یک بر این اساس است که نمونه‌های گرفته‌شده برای تحلیل عاملی متغیرهای مورد نظر کافی‌اند. در خروجی این آزمون، اگر میزان sig کمتر از ۵ درصد باشد، می‌توان نتیجه گرفت در سطح خطای ۵ درصد و یا سطح اطمینان ۹۵ درصد فرض صفر تأیید نشده و بنابراین، فرض یک، یعنی کافی بودن نمونه‌ها، پذیرفته می‌شود (Zhang et al., 2014). آزمون KMO نیز معیاری برای کافی بودن اندازه نمونه، برای تحلیل عاملی اکتشافی، در نظر گرفته شده است. ضریب آزمون

کیفیت هوا این شاخص، براساس جدول استاندارد طبقه‌بندی آلودگی هوا، به شش کلاس طبقه‌بندی می‌شود. در نهایت، این کلاس به‌منزله ورودی مدل برای پیش‌بینی آلودگی هوا به کار می‌رود.

$$I = \frac{I_{high}-I_{low}}{C_{high}-C_{low}} (C - C_{low}) + I_{low} \quad (1)$$

I شاخص کیفیت هوا، C غلظت هریک از آلاینده‌های هوا، C_{high} ، C_{low} ، I_{high} و I_{low} نقاط انفصال مربوط به آلاینده محسوب می‌شوند. منظور از نقاط انفصال نقاط جداسازی بازه‌های در نظر گرفته‌شده برای هر کلاس از آلودگی است که از جدول‌هایی که سازمان حفاظت محیط‌زیست در اختیار قرار داده است استخراج می‌شوند (Mintz, 2012).

۲-۳- انتخاب داده‌های ورودی

مطالعات گذشته نشان می‌دهد غلظت آلاینده‌ها در هر روز با توجه به شرایط آب‌وهوایی (داده‌های هواشناسی) و میزان آلودگی هوا در روزهای گذشته مشخص می‌شود (Ziomas et al., 1995; Robeson and Steyn, 1990). شهر تهران پنج ایستگاه هواشناسی دارد که داده‌های مربوط به وضعیت آب‌وهوا را گرد می‌آورند و این داده‌ها، از طریق اداره کل هواشناسی استان تهران، در دسترس قرار می‌گیرند. در این مطالعه، داده‌های دما، فشار، سرعت وزش باد، میزان پوشش ابر و رطوبت به‌منزله پارامترهای تأثیرگذار در آلودگی هوا به کار رفته‌اند. این داده‌ها در بازه زمانی سه‌ساعته برداشت شده و به سال ۱۳۹۳ متعلق‌اند. همچنین، مطالعات گذشته نشان می‌دهد میزان آلودگی در سه روز گذشته بیشترین تأثیر را در میزان آلودگی هوا در ۲۴ ساعت آینده دارد (Ghaemi et al., 2016). به همین منظور، داده‌های مربوط به شاخص کیفیت هوا در سه روز گذشته نیز به‌منزله ورودی به مدل معرفی می‌شوند. مکان نیز تأثیر بسیار زیادی در توزیع آلاینده‌ها دارد (Zheng et al., 2013). بنابراین، برای مدل‌سازی توزیع مکانی آلاینده‌ها، از پارامترهای مکانی همچون ارتفاع، توپوگرافی سطح زمین و فاصله از جاده استفاده شده

1. Bartlett
2. Kaiser-Meyer-Olkin
3. exploratory factor analysis

$$PC_2 = C_{21}X_1 + C_{22}X_2 + C_{23}X_3 + \dots +$$

$$C_{2p}X_p \quad (\text{محور } Y_2)$$

...

$$PC_p = C_{p1}X_1 + C_{p2}X_2 + C_{p3}X_3 + \dots +$$

$$C_{pp}X_p \quad (\text{محور } Y_p)$$

$C_{a,b}$ ضریب مؤلفه متغیر b روی مؤلفه اصلی محور Y_a است. آنالیز مؤلفه اصلی مجموعه‌ای از متغیرهای (X_1, X_2, \dots, X_p) را به متغیرهای جدید (Y_1, Y_2, \dots, Y_p) ، که با یکدیگر همبستگی ندارند، تبدیل می‌کند.

۲-۷- سیستم استنتاج عصبی- فازی تطبیقی

در این تحقیق، برای پیش‌بینی آلودگی هوا، از سیستم استنتاج عصبی- فازی تطبیقی (ANFIS) استفاده شده است. در این سیستم استنتاجی، ترکیب شبکه عصبی و منطق فازی به منظور نگاشت غیرخطی بین داده‌های ورودی و خروجی به کار می‌رود که ابزاری قدرتمند برای پیش‌بینی پدیده‌های پیچیده را فراهم می‌کند (Jang, 2005; Güler and Übeyli, 1992). برای نمایش ساختار ANFIS، دو قانون اگر- آن‌گاه فازی مربوط به مدل سوگنو درجه یک، در ادامه، بیان شده است:

$$\text{If } X \in \mu_{A_1}(x) \text{ and } Y \text{ is } \in \mu_{B_1}(y) \text{ then } f_1 =$$

$$p_1X + q_1Y + r_1$$

$$\text{If } X \in \mu_{A_2}(x) \text{ and } Y \text{ is } \in \mu_{B_2}(y) \text{ then } f_2 =$$

$$p_2X + q_2Y + r_2$$

$$\mu_{A_i}(x) \text{ و } \mu_{B_i}(y) \text{ توابع عضویت و } p_1, q_1, r_1, p_2, q_2, r_2$$

و r_2 مقادیر ثابت‌اند.

همان‌طور که در شکل ۲ نشان داده شده، یک سیستم استنتاج عصبی- فازی تطبیقی از پنج لایه اصلی تشکیل شده است. در این شکل، لایه‌های ثابت با دایره و لایه‌های تطبیقی با مربع نمایش داده شده‌اند.

KMO همواره بین ۰ و ۱ در نوسان است. در صورتی که مقدار آن کمتر از ۰/۵ باشد، داده‌ها برای تحلیل عاملی مناسب نخواهند بود و در صورتی که مقدار آن بزرگ‌تر از ۰/۷ باشد، همبستگی‌های موجود در بین داده‌ها برای تحلیل مناسب خواهند بود. مقادیر بین ۰/۵ تا ۰/۶۹ نیز بیانگر میانگین میزان تناسب داده‌ها برای تحلیل است (Mostafaeipour et al., 2017). بنابراین، پیش از تحلیل مؤلفه اصلی، با انجام دادن این دو آزمون باید بررسی شود که آیا داده‌ها برای تحلیل مؤلفه اصلی مناسب‌اند یا خیر.

۲-۶- تحلیل مؤلفه اصلی^۱

زمانی که تعداد پارامترهای ورودی اندک است، مدل به راحتی و بدون صرف زمان زیادی اجرا می‌شود. اما به محض اینکه تعداد متغیرهای ورودی افزایش یابد، پیچیدگی مدل بیشتر می‌شود و زمان انجام دادن پردازش‌ها نیز به شدت افزایش می‌یابد. برای این منظور، می‌توان از تکنیک‌های کاهش داده^۲، مانند تحلیل مؤلفه اصلی برای استخراج مؤلفه‌های اصلی که همبستگی ندارند، استفاده کرد (Alizadeh et al., 2012; Sindelar and Babuska, 2004). با استفاده از تحلیل مؤلفه‌های اصلی، می‌توان متغیرهای مستقل همبسته بسیاری را، با تعداد محدودی متغیر مستقل جدید که مؤلفه‌های اصلی نامیده می‌شوند و با یکدیگر وابسته نیستند، جایگزین کرد. مؤلفه‌های غیرهمبسته به دست آمده مؤلفه‌های اصلی (PC) نامیده می‌شوند که از بردارهای ویژه ماتریس کوواریانس یا ماتریس همبستگی متغیرهای اصلی به دست می‌آیند. مزیت اصلی این روش از بین بردن هم‌خطی در مدل، به علت تعداد زیاد متغیرهای مؤثر در مدل است. در حالت کلی، آنالیز مؤلفه اصلی بدین صورت نوشته می‌شود (Ul-Saufie et al., 2013):

$$PC_1 = C_{11}X_1 + C_{12}X_2 + C_{13}X_3 + \dots +$$

$$C_{1p}X_p \quad (\text{محور } Y_1)$$

1. principal component analysis (PCA)

2. data reduction

شکل ۲. معماری یک سیستم استنتاج عصبی- فازی تطبیقی
منبع: Sahoo et al., 2015

خروجی لایه اول درجه توابع عضویت^۱ ورودی‌هاست:

رابطه (۳) $O_i^1 = \mu A_i(x), i = 1,2$

رابطه (۴) $O_i^1 = \mu B_i(y), i = 1,2$

در این روابط، μ وزن محاسبه‌شده براساس توابع عضویت فازی است. برای نمونه، اگر از تابع عضویت زنگوله‌ای استفاده شود، مقدار $\mu_{A_i}(x)$ از رابطه (۵) به دست می‌آید:

رابطه (۵) $\mu_{A_i}(x) = \frac{1}{1 + \left(\frac{x-c_i}{a_i}\right)^{2b_i}}$

که در آن a_i, b_i, c_i پارامترهای تابع عضویت‌اند. پارامترهای توابع عضویت در این لایه پارامترهای مقدم^۲ شناخته می‌شوند.

گره‌ها، در لایه دوم، گره‌های ثابت‌اند. در این لایه، توابع عضویت استخراج‌شده از لایه نخست در یکدیگر ضرب می‌شوند و خروجی لایه دوم از رابطه (۶) حاصل می‌شود. درواقع، با ضرب مقایر ورودی به هر گره در یکدیگر، وزن هر قانون به دست می‌آید.

رابطه (۶) $O_i^2 = w_i = \mu A_i(x) = \mu B_i(y)$

گره‌ها، در لایه سوم نیز، گره‌های ثابت‌اند. در این لایه، مقادیر w_i ها به صورت زیر نرمال می‌شود:

رابطه (۷) $O_i^3 = \bar{w}_i = \frac{w_i}{w_1+w_2} \quad i = 1,2$

گره‌های لایه چهارم گره‌های تطبیقی^۳ به‌شمار می‌روند. خروجی لایه چهارم نیز بدین صورت است:

رابطه (۸) $O_i^4 = \bar{w}_i f_i = \bar{w}_i(p_i x + q_i y + r_i), i = 1,2$

خروجی هر گره، در این لایه، محصولی از خروجی لایه سوم و رابطه پلی‌نومینال درجه یک است (برای مدل سوگنو درجه یک).

در نهایت، لایه پنجم، که تنها یک گره ثابت است، خروجی نهایی را محاسبه می‌کند و هدف آن جمع‌بندی تمامی ورودی‌ها به این لایه است. خروجی نهایی از رابطه (۹) حاصل می‌شود:

رابطه (۹) $O_i^5 = \sum_{i=1}^2 \bar{w}_i f_i = \frac{\sum_{i=1}^2 w_i f_i}{w_1+w_2}$

همان‌طور که مشاهده شد، در ساختار ANFIS، دو لایه تطبیقی وجود دارد: لایه نخست و لایه چهارم. در لایه اول، سه پارامتر قابل تعدیل^۴ $\{a_i, b_i, c_i\}$ وجود

1. membership functions
2. premise parameters
3. adaptive
4. modifiable

دارد که مربوط به تابع عضویت ورودی‌اند. این پارامترها «پارامترهای مقدم» نام دارند. در لایه چهارم، سه پارامتر قابل تعدیل $\{p_i, q_i, r_i\}$ وجود دارد که مربوط به رابطه پلی‌نومیال می‌شوند. این پارامترها «پارامترهای تالی»^۱ نامیده می‌شوند (Güler and Übeyli, 2005; Jang, 1992). هدف اصلی آموزش ANFIS تنظیم درست این پارامترهاست؛ به گونه‌ای که خروجی ANFIS با داده‌های آموزشی سازگار شود.

۳- طراحی سیستم پیش‌بینی آلودگی هوای

شهری با استفاده از مدل ترکیبی PCA-ANFIS

هدف این تحقیق مطرح کردن سیستمی برای پیش‌بینی آلودگی هوای شهر تهران طی ۲۴ ساعت آینده است. در این سیستم، ابتدا پیش‌بینی روی ایستگاه‌های سنجش آلودگی هوا انجام گرفته و سپس، از روش‌های درونیابی به منظور پهنه‌بندی آلودگی در سایر مناطق استفاده شده است. در این سیستم، مدل ترکیبی PCA-ANFIS برای پیش‌بینی آلودگی هوا به کار رفته است. از تحلیل مؤلفه اصلی برای کاهش ابعاد داده‌های ورودی استفاده شده که، افزون بر کاهش ابعاد داده‌های ورودی و در نتیجه، کاهش زمان محاسبات، همبستگی میان متغیرهای مستقل ورودی را نیز از بین می‌برد. داده‌های هواشناسی شامل دما، رطوبت، فشار، میزان پوشش ابر و سرعت باد، داده‌های مربوط به غلظت آلاینده‌ها در سه روز گذشته به همراه پارامترهای ارتفاع، توپوگرافی سطح زمین و فاصله از جاده، برای طراحی سیستم پیش‌بینی آلودگی هوا، به کار رفته‌اند. برای محاسبه فاصله از جاده، از داده‌های OSM استفاده شده است. جاده‌ها، براساس نوع و عرضشان، وزن‌دهی می‌شوند و با استفاده از تابع کرنل، یک رستر استخراج می‌شود که مقدار هر پیکسل بیانگر میزان اهمیت آن پیکسل به نسبت جاده است. نقشه ارتفاع با استفاده از داده‌های SRTM با قدرت تفکیک ۹۰ متری تهیه شده است. اثر توپوگرافی براساس میزان تحدب و تقعر زمین بررسی شده است. برای این منظور، از افزونه

$$L_{d+1}^h = f(AQI_{d+1}^h, AQI_{d-1}^h, AQI_{d-2}^h, T_{d+1}^h, RH_{d+1}^h, P_{d+1}^h, C_{d+1}^h, WS_{d+1}^h, DR_i, LH_i, SC_i)$$

رابطه (۱۰)

در این ساختار، اندیس‌های d و h به ترتیب روز و ساعت را مشخص می‌کنند. سه پارامتر ابتدایی شاخص آلودگی هوا در سه روز قبل و $P_{d+1}^h, RH_{d+1}^h, T_{d+1}^h$ و WS_{d+1}^h ، به ترتیب، دما، رطوبت نسبی، فشار و سرعت باد را نشان می‌دهند. سه پارامتر انتهایی نیز، به ترتیب، DR_i فاصله از جاده، LH_i میانگین ارتفاع منطقه و SC_i توپوگرافی سطح زمین را نمایش می‌دهند. i شماره ایستگاه سنجش آلودگی است. L_{d+1}^h نیز کلاس آلودگی هوای مربوط به این پارامترها در ۲۴ ساعت بعد را مشخص می‌کند.

مراحل طراحی سیستم نیز در شکل ۶ نمایش داده شده است. پس از شناسایی پارامترهای مؤثر در پیش‌بینی آلودگی هوا، داده‌های گردآوری شده، براساس آنچه در بخش ۴-۲ گفته شد، نرمال‌سازی می‌شوند. سپس تحلیل مؤلفه اصلی، به منظور استخراج مؤلفه‌های اصلی، انجام می‌گیرد. پس از استخراج و انتخاب تعداد مؤلفه‌های اصلی و نگاشت داده‌ها به فضای مؤلفه‌ها، داده‌های ورودی به دو دسته آموزش و آزمون تقسیم می‌شوند. ۸۰٪ از داده‌های ورودی برای آموزش و ۲۰٪ نیز برای آزمایش الگوریتم به کار می‌روند. داده‌های آموزشی برای آموزش ANFIS و داده‌های آزمون به منظور ارزیابی الگوریتم آموزش دیده استفاده می‌شوند.

1. consequent parameters

پس از تقسیم داده‌ها، تمامی داده‌های آموزشی و

ورودی و یک خروجی در شکل ۴ نمایش داده شده است. از آنجا که هر قانون یک خروجی گسسته^۱ دارد، خروجی کلی از میانگین وزن دار^۲ هر یک از خروجی‌ها به دست می‌آید. استفاده از این روش موجب کاهش زمان پردازش می‌شود. پس از آموزش، داده‌های تست برای بررسی کارایی مدل به کار می‌روند. در نهایت، نتایج حاصل از پیش‌بینی برای پهنه‌بندی آلودگی هوا در سطح شهر تهران استفاده می‌شود.

تست در ابزار ANFIS نرم‌افزار متلب، نسخه ۲۰۱۳، بارگذاری می‌شوند. داده‌های آموزشی ابتدا وارد سیستم می‌شوند. با تعیین نوع و تعداد توابع عضویت ورودی و نوع توابع عضویت خروجی، یک سیستم استنتاج فازی از نوع سوگنو با یک خروجی از مدل ایجاد می‌شود. همچنین، تابع خطی به‌منزله تابع عضویت خروجی انتخاب شده است. یک سیستم فازی سوگنو با دو



شکل ۳. ساختار PCA-ANFIS برای پیش‌بینی آلودگی هوای شهری

شکل ۴. یک سیستم فازی سوگنو با دو ورودی و یک خروجی

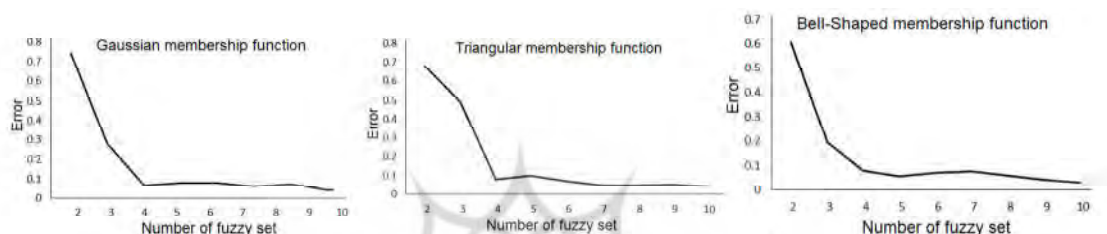
منبع: Riahi-Madvar et al., 2009

-
1. crisp output
 2. weighted average

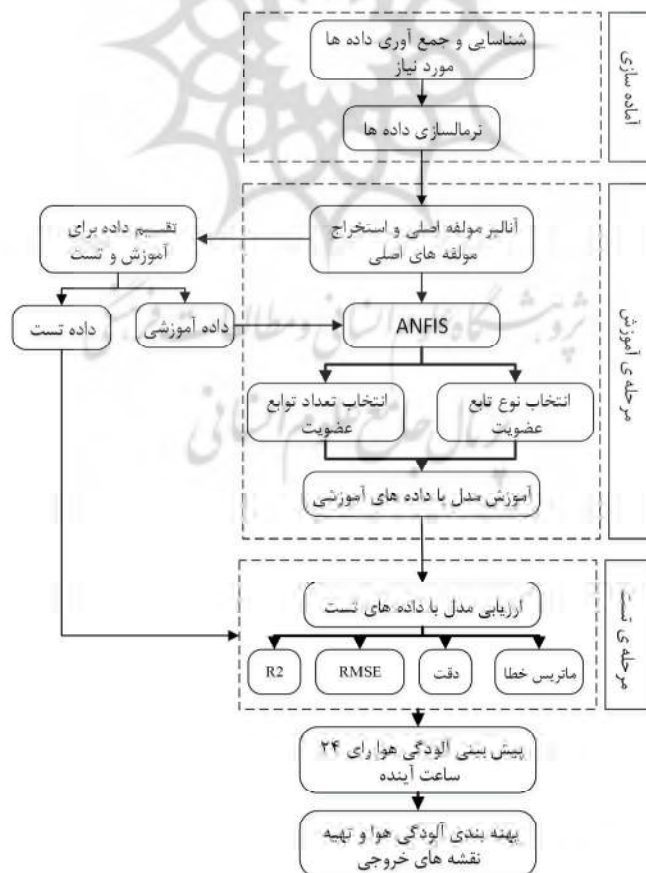
۳-۱- انتخاب تعداد و نوع تابع عضویت ورودی

به منظور تعیین بهترین تابع عضویت و تعداد توابع عضویت، به دلیل زمان پر بودن محاسبات، بخشی از داده‌ها برای نمونه انتخاب شدند. سپس سه تابع عضویت پر کاربرد مثلثی، زنگوله‌ای و گوسی، با تعداد توابع عضویت بین ۲ تا ۱۰، مورد بررسی قرار گرفتند. نتایج حاصل در شکل ۵ نمایش داده شده است. همان‌طور که در این شکل مشاهده می‌کنید، هر سه تابع، با چهار تابع عضویت، به دقت مناسبی می‌رسند. از آنجاکه با افزایش تعداد توابع

عضویت، تعداد قوانین به صورت نمایی افزایش می‌یابد و موجب افزایش زمان محاسبات می‌شود و از سویی دیگر، چون با افزایش تعداد این توابع، دقت حاصل تغییر چشمگیری ندارد؛ به منظور کاهش زمان محاسبات، از چهار تابع عضویت برای آموزش شبکه عصبی-فازی تطبیقی استفاده شده است. همچنین، از میان این سه تابع عضویت بررسی شده، تابع گوسین دقت بهتری به دست داده است. بنابراین، از این تابع برای اجرای مدل استفاده می‌شود.



شکل ۵. مقایسه خطای حاصل از توابع فازی با تغییر تعداد توابع عضویت ورودی



شکل ۶. ساختار سیستم طراحی شده به منظور پیش‌بینی آلودگی هوای شهر تهران برای ۲۴ ساعت آینده

جدول ۱ ماتریس خطایی با دو کلاس را نمایش می‌دهد. مقادیر این ماتریس را به صورت زیر می‌توان تفسیر کرد (Athanasiadis et al., 2003):

- a: تعداد نمونه‌های به‌درستی پیش‌بینی‌شده، زمانی که نمونه مورد نظر منفی است.
- b: تعداد نمونه‌های منفی که به اشتباه در کلاس مثبت پیش‌بینی شده‌اند.
- c: تعداد نمونه‌هایی که مثبت‌اند اما به اشتباه در کلاس منفی پیش‌بینی شده‌اند.
- d: تعداد نمونه‌های مثبتی که به‌درستی پیش‌بینی شده‌اند.

جدول ۱. ماتریس خطا

پیش‌بینی		مشاهداتی
مثبت	منفی	
b	a	منفی
d	c	مثبت

۴- نتایج و تحلیل

در این بخش، نتایج حاصل از اجرای مدل ترکیبی PCA-ANFIS برای پیش‌بینی آلودگی هوای شهر تهران، طی ۲۴ ساعت آینده، تشریح می‌شود. از آنجا که آلودگی هوا در فصل‌های گوناگون رفتاری متفاوت دارد، آموزش و سنجش مدل برای چهار فصل جداگانه صورت می‌گیرد و نتایج حاصل بیان می‌شوند.

۴-۱- ماتریس همبستگی

جدول ۲ میزان همبستگی شاخص کیفیت هوا را با پارامترهای هواشناسی، میزان آلودگی در روزهای گذشته و پارامترهای مکانی نشان می‌دهد. همان‌طور که مشاهده می‌کنید، شاخص کیفیت هوا با پارامترهای توپوگرافی، دما، فشار، سرعت باد و پوشش ابر همبستگی منفی و با پارامترهای غلظت آلاینده‌ها در روزهای قبل، ارتفاع، فاصله از جاده و رطوبت همبستگی مثبت دارد.

1. accuracy
2. root mean square error (RMSE)
3. Rsquared
4. confusion matrix

در این تحقیق، با تعیین نوع و تعداد توابع عضویت ورودی، یک سیستم استنتاج فازی از نوع سوگنو ایجاد شده است. ورودی این سیستم مؤلفه‌های استخراج‌شده از مرحله تحلیل مؤلفه اصلی‌اند و خروجی آن کلاس آلودگی هوا برای ۲۴ ساعت آینده محسوب می‌شوند.

۳-۲- ارزیابی نتایج

سیستم استنتاج عصبی فازی- تطبیقی با اختصاص وزن‌های تصادفی شروع به آموزش می‌کند. بنابراین، خروجی شبکه ممکن است با مقدار مشاهداتی اختلاف زیادی داشته باشد. زمانی که آموزش ادامه می‌یابد، وزن‌های شبکه به تدریج تعدیل می‌شوند و خروجی سیستم به مقادیر مشاهداتی نزدیک می‌شود. اختلاف مقادیر مشاهداتی و محاسبه‌شده میزان خطای سیستم را مشخص می‌کند که به روش‌های گوناگونی می‌توان آن را ارزیابی کرد. در این تحقیق، از پارامترهای دقت طبقه‌بندی^۱، خطای جذر میانگین مربعات^۲، ضریب تعیین^۳ و ماتریس خطا^۴ به منظور ارزیابی نتایج استفاده شده است. روابط (۱۱)، (۱۲) و (۱۳)، به ترتیب، شیوه محاسبه پارامترهای دقت، خطای مجذور میانگین مربعات و ضریب تعیین را بیان می‌کنند:

دقت = تعداد نمونه‌های تست که به‌درستی پیش‌بینی شده‌اند / تعداد کل نمونه‌های تست

رابطه (۱۱)

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n |Y_i - Y_i^*|^2}$$

رابطه (۱۲)

$$R^2 = 1 - \frac{\sum_{i=1}^n (Y_i^* - \bar{Y})^2}{\sum_{i=1}^n (Y_i - \bar{Y})^2}$$

رابطه (۱۳)

در معادلات (۱۲) و (۱۳)، Y_i^* مقدار پیش‌بینی‌شده و Y_i مقدار مشاهداتی‌اند و در رابطه (۱۳)، \bar{Y} میانگین مقادیر مشاهداتی است.

ماتریس خطا شامل اطلاعاتی درمورد مقادیر مشاهده‌شده و پیش‌بینی‌شده از طریق سیستم‌های طبقه‌بندی است. میزان کارایی سیستم، با توجه به اطلاعات موجود در این ماتریس، قابل ارزیابی است.

۴-۲- آزمون بارتلت و KMO

میزان Sig حاصل از نتایج آزمون بارتلت (جدول ۳) نشان می‌دهد فرضیه صفر (وجودنداشتن همبستگی) برای مجموعه داده‌های ورودی رد می‌شود. میزان sig برای هر چهار فصل کمتر از ۰.۵٪ است که به معنای ارتباط قوی بین متغیرهاست.

همچنین، در این جدول‌ها، شاخص KMO نیز نمایش داده شده است. میزان عددی بالای ۰.۶ برای این شاخص، در تمامی فصل‌ها، نشان از کفایت مناسب نمونه‌گیری برای تحلیل مؤلفه اصلی دارد.

جدول ۲. ماتریس همبستگی میان پارامترهای ورودی

	AQI _d	AQI _{d-1}	AQI _{d-2}	LH	DR	C	T	P	RH	WS	CC
AQI _d	1.000										
AQI _{d-1}	.467	1.000									
AQI _{d-2}	.384	.485	1.000								
LH	.125	.127	.122	1.000							
DR	.068	.067	.069	.466	1.000						
SC	-.002	.000	.000	.167	.099	1.000					
T	-.029	-.022	-.042	-.001	.028	.021	1.000				
P	-.014	-.006	-.017	-.007	-.030	-.010	-.446	1.000			
RH	.012	-.008	.020	-.007	-.008	-.018	-.761	.278	1.000		
WS	-.002	.002	.001	-.002	.003	.004	.143	-.263	-.204	1.000	
C	-.031	-.051	-.031	-.002	.005	-.012	-.202	-.107	.368	.055	1.000

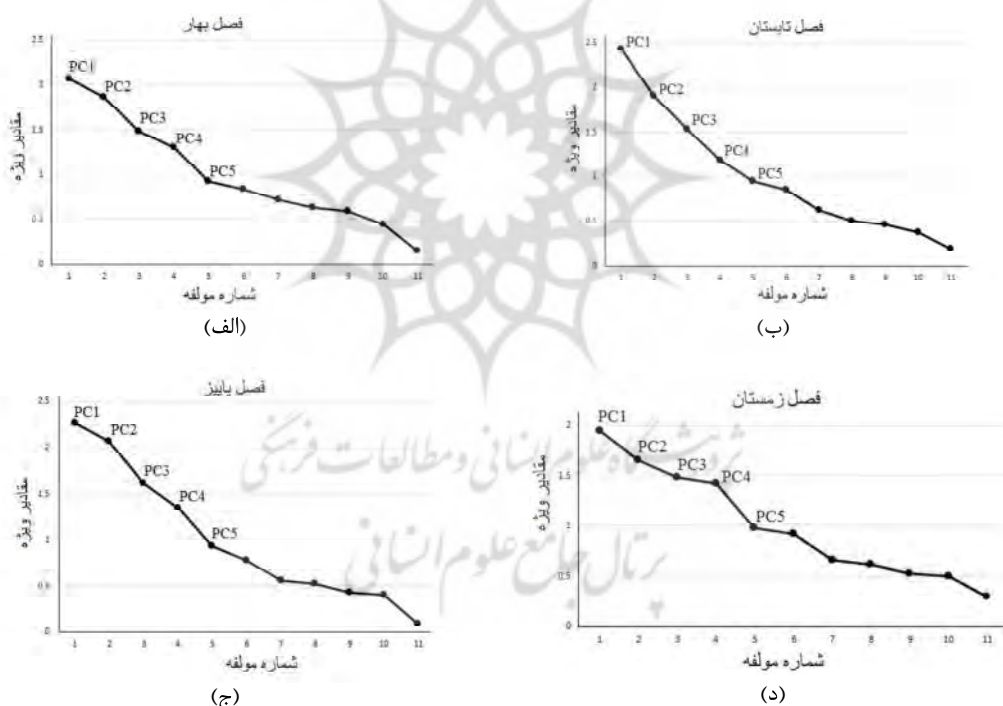
جدول ۳. تست بارتلت و KMO مربوط به داده‌های فصل‌های گوناگون

Spring	Kaiser-Meyer-Olkin Measure of Sampling Adequacy.	.659
	Bartlett's Test of Sphericity	Sig. .000
Summer	Kaiser-Meyer-Olkin Measure of Sampling Adequacy.	.671
	Bartlett's Test of Sphericity	Sig. .000
Fall	Kaiser-Meyer-Olkin Measure of Sampling Adequacy.	.685
	Bartlett's Test of Sphericity	.000
Winter	Kaiser-Meyer-Olkin Measure of Sampling Adequacy.	.618
	Bartlett's Test of Sphericity	Sig. .000

۳-۴- تخمین تعداد مؤلفه‌های اصلی

تعداد کل مؤلفه‌های استخراج‌شده با تعداد متغیرهای اولیه برابر است. اما برای اجرا، می‌توان تعداد مشخصی از این مؤلفه‌ها را انتخاب کرد. معمولاً دو یا سه مؤلفه اول مقدار چشمگیری از پراکندگی داده‌ها را در نظر می‌گیرد. در این تحقیق، یازده پارامتر ورودی وجود دارد؛ بنابراین، یازده مؤلفه اصلی خواهیم داشت. برای انتخاب مؤلفه‌ها، از مؤلفه‌هایی استفاده شده که مقادیر ویژه آنها بیش از یک است. همچنین، مؤلفه‌هایی که مقادیر ویژه 0.9 دارند نیز، به دلیل نزدیک بودن به مقدار یک، در نظر گرفته می‌شوند. سهم اطلاعات و متغیرها

در یازده مؤلفه اصلی به دست آمده برای فصل‌های گوناگون در شکل ۷ نمایش داده شده است. با توجه به جدول بردارهای ویژه عناصر در مؤلفه‌های اصلی، پنج مؤلفه اصلی نخست برای پیش‌بینی آلودگی هوا مناسب تشخیص داده شدند. همان‌طور که در جدول ۴ آمده است؛ این پنج مؤلفه، در مجموع، دربردارنده 70% تغییرات در فصل بهار، 73% در فصل تابستان، 75% در فصل پاییز و 68% تغییرات در فصل زمستان‌اند. همچنین، جدول ۵ درصد تغییرات مربوط به یازده مؤلفه اصلی استخراج‌شده را در فصل‌های گوناگون نمایش می‌دهد.



شکل ۷. مقادیر ویژه مؤلفه‌های اصلی برای (الف) فصل بهار؛ (ب) فصل تابستان؛ (ج) فصل پاییز؛ (د) فصل زمستان

جدول ۴. مجموع درصد تغییرات مربوط به پنج مؤلفه اول

فصل سال	بهار	تابستان	پاییز	زمستان
درصد تغییرات	70%	73%	75%	68%

جدول ۵. درصد تغییرات مربوط به هر مؤلفه اصلی برای فصل‌ها گوناگون سال

فصل	بهار	تابستان	پاییز	زمستان	شماره مؤلفه
	%۱۹	%۲۲	%۲۱	%۱۸	PC1
	%۱۷	%۱۷	%۱۹	%۱۵	PC2
	%۱۳	%۱۴	%۱۵	%۱۳	PC3
	%۱۲	%۱۱	%۱۲	%۱۳	PC4
	%۸	%۹	%۸	%۹	PC5
	%۸	%۸	%۷	%۸	PC6
	%۶	%۶	%۵	%۶	PC7
	%۶	%۵	%۵	%۶	PC8
	%۵	%۴	%۴	%۵	PC9
	%۴	%۳	%۴	%۵	PC10
	%۱	%۲	%۱	%۳	PC11

جدول ۶ تا جدول ۹ ماتریس مؤلفه‌ها برای فصل‌های گوناگون را نمایش می‌دهند. با استفاده از این ماتریس می‌توان مؤلفه‌ها را محاسبه کرد. برای نمونه، پنج مؤلفه اصلی نخست برای فصل بهار، با توجه به مقادیر جدول ۷، با استفاده از معادلات زیر محاسبه می‌شوند:

$$PC_1 = 0.859*Temp - 0.857*Hum + 0.198* AQI^{d-1} + 0.132* AQI^{d-2} + 0.217* AQI + 0.209*DR + 0.235*LH - 0.156*CC - 0.415*Press + 0.427*WS + 0.106*Curv \quad \text{رابطه (۱۴)}$$

$$PC_2 = -0.324*Temp + 0.269*Hum + 0.644* AQI^{d-1} + 0.618* AQI^{d-2} + 0.607* AQI + 0.353*DR + 0.544*LH - 0.093*CC + 0.053*Press - 0.128*WS + 0.262*Curv \quad \text{رابطه (۱۵)}$$

$$PC_3 = 0.004*Temp + 0.034*Hum - 0.398* AQI^{d-1} - 0.353* AQI^{d-2} - 0.377* AQI + 0.742*DR + 0.621*LH + 0.046*CC + 0.106*Press - 0.114*WS + 0.303*Curv \quad \text{رابطه (۱۶)}$$

$$PC_4 = -0.171*Temp + 0.283*Hum + 0.014* AQI^{d-1} - 0.103* AQI^{d-2} + 0.056* AQI + 0.034*DR + 0.037*LH + 0.719*CC - 0.644*Press - 0.489*WS + 0.076*Curv \quad \text{رابطه (۱۷)}$$

$$PC_5 = -0.04*Temp - 0.038*Hum - 0.026* AQI^{d-1} + 0.027* AQI^{d-2} - 0.052* AQI - 0.276*DR - 0.164*LH + 0.043*CC + 0.055*Press - 0.056*WS + 0.9*Curv \quad \text{رابطه (۱۸)}$$

جدول ۸. ماتریس پنج مؤلفه اصلی اول، فصل پاییز

Component	Component				
	1	2	3	4	5
AQI _{d-1}	0.734	-0.169	0.257	-0.183	0.011
AQI _d	0.682	-0.112	0.25	-0.161	-0.114
AQI _{d-2}	0.656	-0.178	0.268	-0.17	0.164
P	-0.125	-0.822	-0.132	0.009	-0.124
C	-0.023	0.713	0.378	-0.122	-0.329
H	-0.403	0.115	0.777	-0.279	0.102
T	0.485	0.454	-0.505	0.172	-0.387
LH	0.202	-0.071	0.328	0.708	-0.062
DR	-0.071	0.045	0.273	0.7	0.01
WS	0.216	0.395	-0.224	0.092	0.802
SC	0.096	-0.108	0.179	0.458	0.068

جدول ۶. ماتریس پنج مؤلفه اصلی اول، فصل تابستان

Component	Component				
	1	2	3	4	5
T	-0.702	0.585	-0.077	0.076	-0.057
AQI _{d-1}	0.696	0.498	-0.092	-0.008	0.092
AQI _{d-2}	0.665	0.49	-0.102	0.101	0.043
AQI _d	0.606	0.511	-0.085	0.1	0.033
RH	0.532	-0.569	0.147	0.385	0.001
P	0.466	-0.518	0.04	-0.164	0.137
DR	-0.064	0.11	0.841	0	-0.223
LH	0.15	0.309	0.792	-0.011	-0.112
CC	-0.115	-0.158	0.117	0.86	0.077
SC	-0.07	-0.067	0.364	-0.304	0.777
WS	-0.308	0.293	-0.02	0.381	0.493

جدول ۹. ماتریس پنج مؤلفه اصلی اول، فصل بهار

Component	Component				
	1	2	3	4	5
RH	-0.883	0.35	-0.035	0.044	-0.028
T	0.875	-0.33	0.013	0.131	0.016
AQI _{d-1}	0.341	0.756	-0.167	-0.051	0.001
AQI _{d-2}	0.308	0.728	-0.148	-0.008	-0.001
AQI _d	0.324	0.718	-0.152	-0.126	0.016
DR	0.045	0.086	0.85	0.037	-0.287
LH	0.112	0.331	0.808	0.012	-0.14
P	-0.062	-0.165	0.076	-0.704	-0.054
WS	0.342	-0.135	-0.023	0.695	-0.025
C	-0.515	0.223	-0.042	0.572	-0.105
SC	-0.071	0.096	0.396	0.059	0.903

جدول ۷. ماتریس پنج مؤلفه اصلی اول، فصل بهار

Component	Component				
	1	2	3	4	5
T	0.859	-0.324	0.004	-0.171	0.04
H	0.857	0.269	0.034	0.283	-0.038
AQI _{d-1}	0.198	0.644	-0.398	0.014	-0.026
AQI _{d-2}	0.132	0.618	-0.353	-0.103	0.027
AQI _d	0.217	0.607	-0.377	0.056	-0.052
DR	0.209	0.353	0.742	0.034	-0.276
LH	0.235	0.544	0.621	0.037	-0.164
C	-0.156	-0.093	0.046	0.719	0.043
P	-0.415	0.053	0.106	-0.644	0.055
WS	0.427	-0.128	-0.114	0.489	-0.056
SC	0.106	0.262	0.303	0.079	0.9

روش ANFIS مقایسه شد. پارامترهای دقت، خطای جذر میانگین مربعات و ضریب تعیین به‌دست‌آمده برای فصل‌های متفاوت، با استفاده از داده‌های تست برای این دو روش، در جدول ۱۰ آمده است. طبق این جدول، روش ترکیبی PCA-ANFIS توانسته دقت بالاتری را در پیش‌بینی آلودگی هوا در فصل‌های متفاوت حاصل کند. دلیل آن ممکن است حذف همبستگی میان پارامترهای ورودی باشد که در نتایج تأثیر دارد. شکل ۷ زمان پردازش دو روش ANFIS و PCA-ANFIS را، در فصل‌های

به همین ترتیب و براساس ماتریس مؤلفه‌های مربوط به هر فصل، مؤلفه‌های اصلی برای فصل‌های دیگر نیز به‌دست می‌آیند.

۴-۴- مقایسه نتایج روش ترکیبی PCA-ANFIS و ANFIS برای فصل‌های گوناگون

برای ارزیابی کارایی روش ترکیبی PCA-ANFIS، نتایج حاصل از این روش، از نظر دقت و زمان پردازش، با

خطای جذر میانگین مربعات را دارد؛ یعنی مدل ترکیبی توانسته است، در این فصل، آلودگی هوا را با دقت بالاتری پیش‌بینی کند. جدول ۱۱ تا جدول ۱۴، به ترتیب، ماتریس خطای حاصل برای فصل‌های بهار، پاییز، زمستان و زمستان را نشان می‌دهد. ستون آخر هر یک از این جدول‌ها دقت حاصل برای هر یک از کلاس‌های آلودگی هوا را بیان می‌کند. همچنین، جدول ۱۵ دقت به‌دست‌آمده برای کلاس‌های متفاوت در فصل‌های گوناگون را نمایش می‌دهد. مطابق این جدول، دقت به‌دست‌آمده برای کلاس‌های ۵ و ۶ در تابستان، پاییز و زمستان صفر و در بهار نیز بسیار پایین است. همچنین، با توجه به ماتریس‌های خطا در فصل‌های گوناگون و توجه به تعداد نمونه‌ها در هر کلاس، مشخص می‌شود که با کاهش تعداد نمونه‌ها در کلاس‌های متفاوت، دقت حاصل نیز کاهش می‌یابد. بنابراین، به دلیل تعداد پایین نمونه‌ها در کلاس‌های ۵ و ۶، مدل PCA-ANFIS نتوانسته به‌درستی این کلاس‌ها را پیش‌بینی کند. مقایسه جدول ۱۰ و جدول ۱۱ با یکدیگر نیز نشان می‌دهد، با وجود اینکه مدل PCA-ANFIS بالاترین دقت کلی را در پیش‌بینی آلودگی هوا در فصل زمستان داشته است، در فصل بهار توانسته تا حدودی کلاس‌های آلودگی ۵ و ۶ را نیز پیش‌بینی کند.

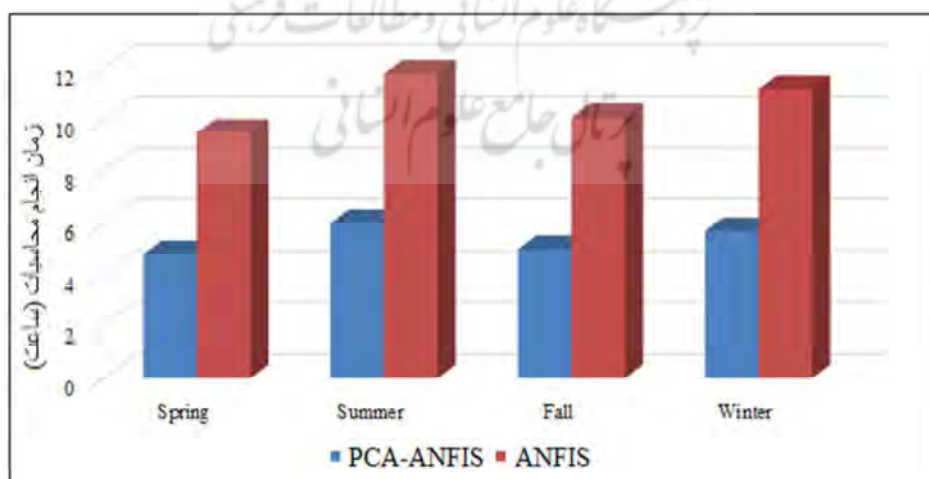
گوناگون، با یکدیگر مقایسه می‌کند. همان‌گونه که مشاهده می‌کنید، وقتی از روش ANFIS استفاده شده است، به دلیل افزایش تعداد پارامترهای ورودی، زمان پردازش تا حدود دو برابر افزایش یافته است. بنابراین، روش ترکیبی PCA-ANFIS افزون بر افزایش دقت توانسته است با سرعت بالاتری محاسبات را انجام دهد.

جدول ۱۰. پارامترهای ارزیابی به‌دست‌آمده برای داده‌های تست در فصل‌های گوناگون، در دو روش ANFIS و PCA-ANFIS

فصل	ANFIS			PCA-ANFIS		
	RMSE	R2	دقت	RMSE	R2	دقت
بهار	۰/۱۶۸	۰/۱۵۵	۰/۱۶۵	۰/۱۵۷	۰/۱۶۱	۰/۱۷۳
تابستان	۰/۱۶۳	۰/۱۶۰	۰/۱۶۷	۰/۱۵۴	۰/۱۶۷	۰/۱۷۴
پاییز	۰/۱۶۲	۰/۱۵۳	۰/۱۵۹	۰/۱۵۸	۰/۱۶۸	۰/۱۷
زمستان	۰/۱۵۲	۰/۱۶۶	۰/۱۶۴	۰/۱۴۶	۰/۱۸۷	۰/۱۷۸

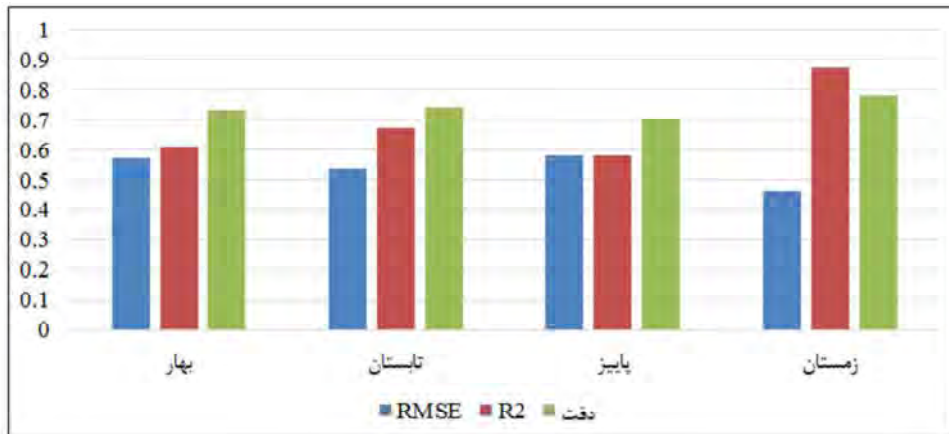
۴-۴-۱- تحلیل نتایج حاصل از روش ترکیبی PCA-ANFIS

دقت، خطای جذر میانگین مربعات و ضریب تعیین به‌دست‌آمده برای فصل‌های گوناگون، در روش PCA-ANFIS، در شکل ۹ آمده است. مشخص است که فصل زمستان بالاترین دقت و ضریب تعیین و کمترین میزان



شکل ۸. مقایسه زمان پردازش برای دو روش ANFIS و PCA-ANFIS

پیش‌بینی آلودگی هوای شهر تهران با استفاده از ترکیب ...



شکل ۹. مقایسه دقت، ضریب تعیین و خطای جذر میانگین مربعات در فصل‌های گوناگون، برای روش PCA-ANFIS

جدول ۱۱. ماتریس خطای فصل بهار

مشاهداتی	پیش‌بینی					
	۱	۲	۳	۴	۵	۶
تعداد نمونه در هر کلاس	۱۱۰۲	۴۹۸	۱۲	۱	۰	۰
دقت کلاس	۰/۱۶۹	۰/۸۳	۰/۶۰	۰/۳۹	۰/۲۱	۰/۰۵

جدول ۱۲. ماتریس خطای فصل تابستان

مشاهداتی	پیش‌بینی					
	۱	۲	۳	۴	۵	۶
تعداد نمونه در هر کلاس	۱۲۹۳	۷۶۵	۱۰	۰	۰	۰
دقت کلاس	۰/۶۲	۰/۸۸	۰/۳۹	۰/۳۲	۰	۰

زینب قائمی و همکاران

جدول ۱۳. ماتریس خطای فصل پاییز

مشاهداتی	پیش بینی						
	۱	۲	۳	۴	۵	۶	تعداد نمونه در هر کلاس
۱	۱۴۰۵	۶۹۶	۳۷	۰	۰	۰	۲۱۳۸
۲	۴۸۶	۲۳۶۸	۱۸۱	۳	۰	۰	۳۰۳۸
۳	۱۱	۲۵۳	۳۹۲	۵۴	۰	۰	۷۱۰
۴	۲	۲۲	۳۳	۳۱	۰	۰	۸۸
۵	۲	۲	۱۱	۲	۰	۰	۱۷
۶	۰	۲	۵	۰	۰	۰	۷

جدول ۱۴. ماتریس خطای فصل زمستان

مشاهداتی	پیش بینی						
	۱	۲	۳	۴	۵	۶	تعداد نمونه در هر کلاس
۱	۱۸۰۶	۶۹۲	۱۳	۵	۱	۰	۲۵۱۷
۲	۳۸۷	۳۱۸۶	۱۳۵	۷	۳	۰	۳۷۱۸
۳	۲۶	۱۸۴	۴۶۰	۶	۱	۰	۶۷۷
۴	۳	۲۰	۱۹	۲۶	۱	۰	۶۹
۵	۰	۵	۵	۰	۰	۰	۱۰
۶	۰	۳	۳	۰	۶	۰	۶

جدول ۱۵. دقت به دست آمده برای هر کلاس در فصل های گوناگون

فصل	کلاس					
	۱	۲	۳	۴	۵	۶
بهار	۰/۶۹	۰/۸۳	۰/۶	۰/۳۹	۰/۲۱	۰/۰۵
تابستان	۰/۶۲	۰/۸۸	۰/۳۹	۰/۳۲	۰	۰
پاییز	۰/۶۶	۰/۷۸	۰/۵۵	۰/۳۵	۰	۰
زمستان	۰/۷۱	۰/۸۵	۰/۶۷	۰/۳۲	۰	۰

می‌توان نواحی آلودگی را، در روزهای مورد نظر، شناسایی کرد و راهکارهایی را برای کاهش آلودگی هوا در این مناطق پیشنهاد داد. افراد بیمار و گروه‌های حساس نیز می‌توانند، با مراجعه به نقشه‌های خروجی، فعالیت‌های خود را در مناطق آلوده کاهش دهند.

شکل ۱۳ تا شکل ۱۶، به ترتیب، مقایسه مقادیر پیش‌بینی شده را، با استفاده از دو روش PCA-ANFIS و ANFIS، با مقادیر واقعی برای بهار، تابستان، پاییز و زمستان نمایش می‌دهد. برای مقایسه مقادیر پیش‌بینی شده با مقادیر واقعی، مجموع اختلاف کلاس‌های پیش‌بینی شده با مقادیر واقعی، برای هر دو روش در فصل‌های متفاوت، محاسبه شده است. نتایج حاصل در جدول ۱۶ ارائه شده‌اند. همان‌طور که دیده می‌شود، اختلاف مقادیر مشاهده شده با استفاده از روش PCA-ANFIS با مقادیر واقعی در سه فصل بهار، تابستان و پاییز کمتر از اختلاف مقادیر پیش‌بینی شده به روش ANFIS با مقادیر واقعی است و در فصل زمستان، اختلاف مقادیر پیش‌بینی شده با روش با مقادیر واقعی برابر است.

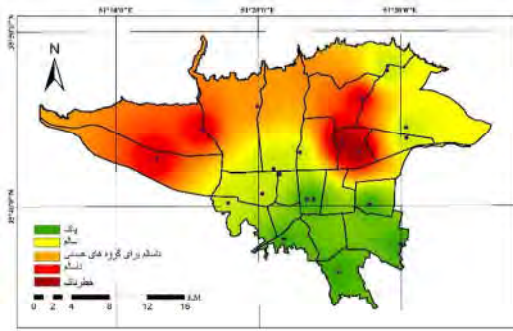
جدول ۱۶. اختلاف میان مقادیر واقعی کلاس آلودگی هوا با مقادیر پیش‌بینی شده، با استفاده از دو روش PCA-ANFIS و ANFI

ANFI				
فصل سال	بهار	تابستان	پاییز	زمستان
PCA-ANFIS	۷	۸	۸	۶
ANFIS	۸	۹	۱۱	۶

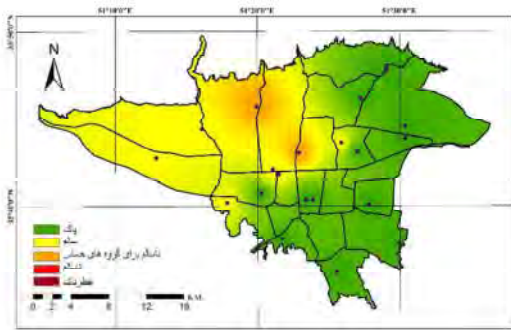
1. inverse distance weighting (IDW)

پس از آموزش مدل و پیش‌بینی آلودگی هوای تهران در ۲۴ ساعت آینده، با استفاده از مدل PCA-ANFIS، می‌توان از نتایج به‌دست‌آمده برای پهنه‌بندی آلودگی هوا در مناطق گوناگون استفاده کرد. بدین‌منظور، پس از پیش‌بینی روی ایستگاه‌های سنجش آلودگی هوا، برای به‌دست‌آوردن میزان آلودگی هوا در مناطقی غیر از ایستگاه‌های سنجش آلودگی هوا، از روش درونیابی معکوس فاصله وزن‌دار^۱ استفاده شده است. شکل ۱۱ نتایج حاصل از پیش‌بینی آلودگی هوا با استفاده از روش ترکیبی PCA-ANFIS و شکل ۱۲ نیز نتایج حاصل از پیش‌بینی آلودگی هوا در تاریخ‌های مشخص‌شده را با استفاده از روش ANFIS، در فصل‌های گوناگون، نمایش می‌دهد. مقادیر واقعی آلودگی هوا برای این تاریخ‌ها در شکل ۱۲ نمایش داده شده است. با مقایسه نقشه‌های خروجی با مقادیر واقعی می‌توان به این نتیجه رسید که نقشه‌های حاصل از خروجی روش PCA-ANFIS به نقشه‌های مقادیر واقعی نزدیک‌تر است. مسئله دیگری که می‌توان به آن پرداخت علت بالابودن آلودگی هوا در برخی مناطق شهر است. این امر می‌تواند به عوامل متفاوتی، از جمله شدت و جهت وزش باد، میزان ترافیک و محل صنایع آلوده‌کننده ربط داشته باشد. برای نمونه، بررسی پارامترهای ورودی برای تاریخ ۹۳/۰۶/۰۲ نشان داد، در این تاریخ، جهت باد از سمت جنوب بوده که موجب شده آلودگی به بخش مرکزی شهر منتقل شود. همچنین، تمرکز بیشتر جاده‌ها در بخش مرکزی و وجود اتوبان و صنایع آلوده‌کننده در بخش‌های غربی شهر موجب تمرکز بیشتر آلودگی در این نواحی شده است. بدین‌ترتیب، با بررسی پارامترهای ورودی، می‌توان ارتباط میان آلودگی هوا در مناطق گوناگون با این پارامترها را استخراج کرد و برنامه‌ریزی‌های لازم را، به‌منظور کاهش میزان آلودگی در این مناطق، اتخاذ کرد. همچنین، با استفاده از نقشه‌های تولیدشده

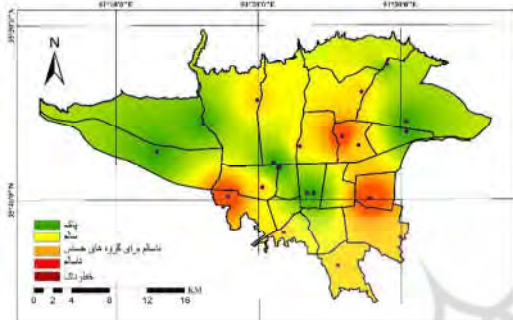
زینب قائمی و همکاران



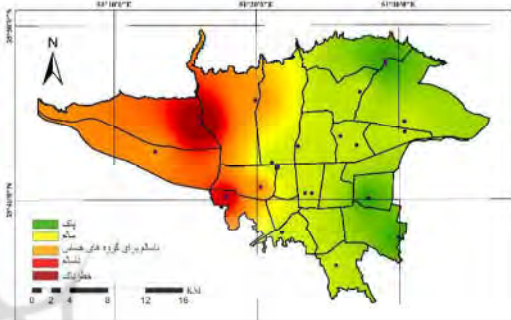
ب) مقادیر مشاهداتی برای تاریخ ۱۳۹۳/۶/۰۲



الف) مقادیر مشاهداتی برای تاریخ ۱۳۹۳/۲/۱۲

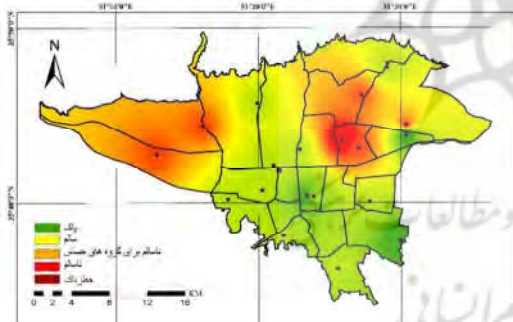


د) مقادیر مشاهداتی برای تاریخ ۱۳۹۳/۱۱/۲۶

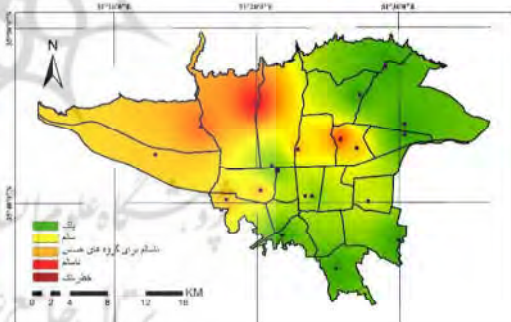


ج) مقادیر مشاهداتی برای تاریخ ۱۳۹۳/۰۷/۰۶

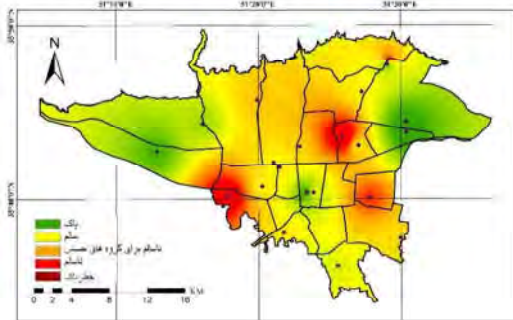
شکل ۱۰. مقادیر مشاهداتی آلودگی هوا برای فصل بهار (الف)؛ فصل تابستان (ب)؛ فصل پاییز (ج)؛ فصل زمستان (د)



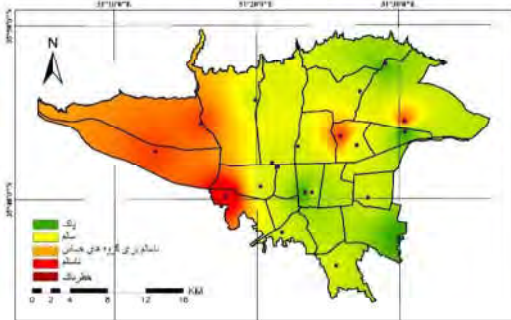
ب) پیش‌بینی برای تاریخ ۱۳۹۳/۶/۰۲



الف) پیش‌بینی برای تاریخ ۱۳۹۳/۲/۱۲



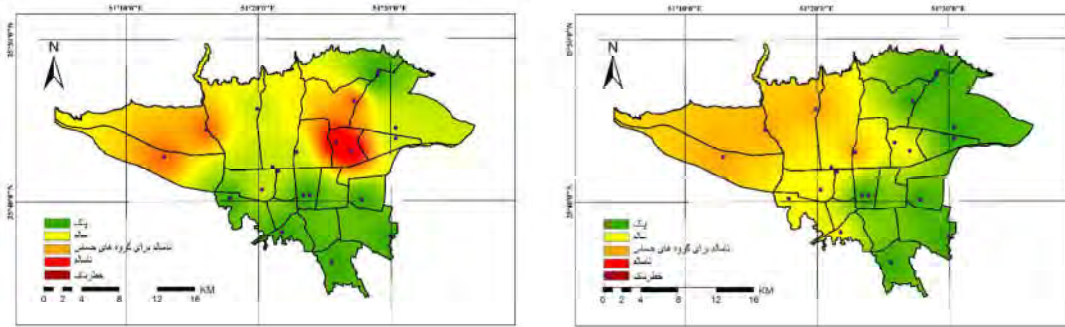
د) پیش‌بینی برای تاریخ ۱۳۹۳/۱۱/۲۶



ج) پیش‌بینی برای تاریخ ۱۳۹۳/۰۷/۰۶

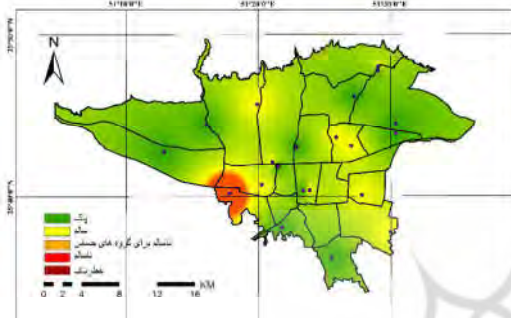
شکل ۱۱. پیش‌بینی آلودگی هوا برای فصل بهار (الف)؛ فصل تابستان (ب)؛ فصل پاییز (ج)؛ و فصل زمستان (د) با استفاده از روش PCA-ANFIS

پیش‌بینی آلودگی هوای شهر تهران با استفاده از ترکیب ...

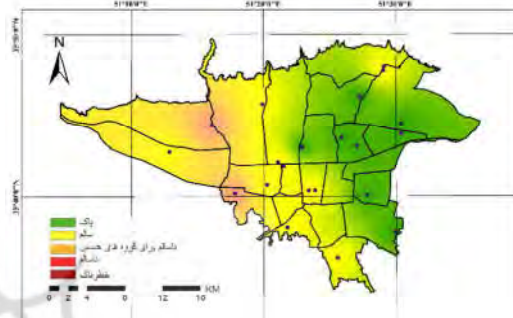


ب) پیش‌بینی برای تاریخ ۱۳۹۳/۶/۰۲

الف) پیش‌بینی برای تاریخ ۱۳۹۳/۲/۱۲



د) پیش‌بینی برای تاریخ ۱۳۹۳/۱۱/۲۶



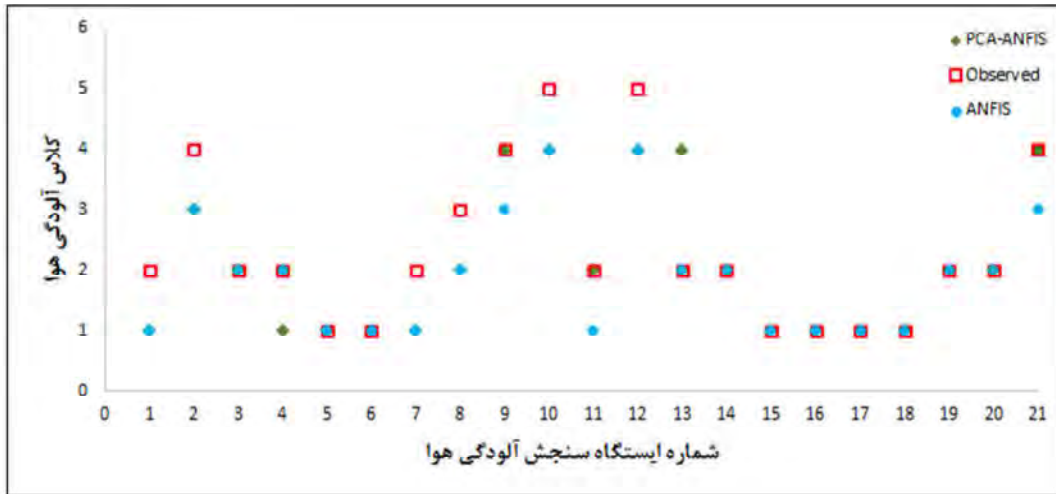
ج) پیش‌بینی برای تاریخ ۱۳۹۳/۰۷/۰۶

شکل ۱۲. پیش‌بینی آلودگی هوا برای فصل بهار (الف)؛ فصل تابستان (ب)؛ فصل پاییز (ج)؛ فصل زمستان (د) با استفاده از روش ANFIS

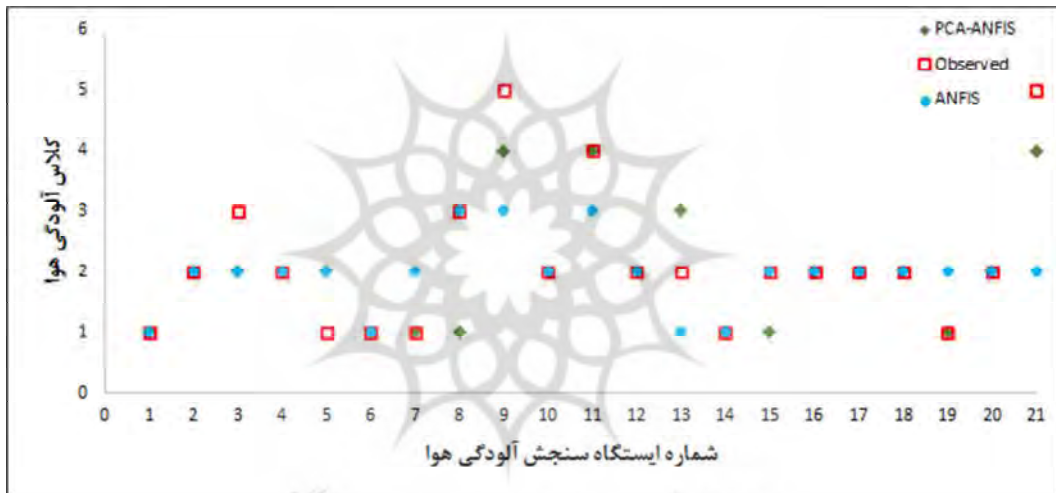


شکل ۱۳. مقایسه مقادیر مشاهداتی و پیش‌بینی‌شده برای تاریخ ۱۳۹۳/۲/۱۲

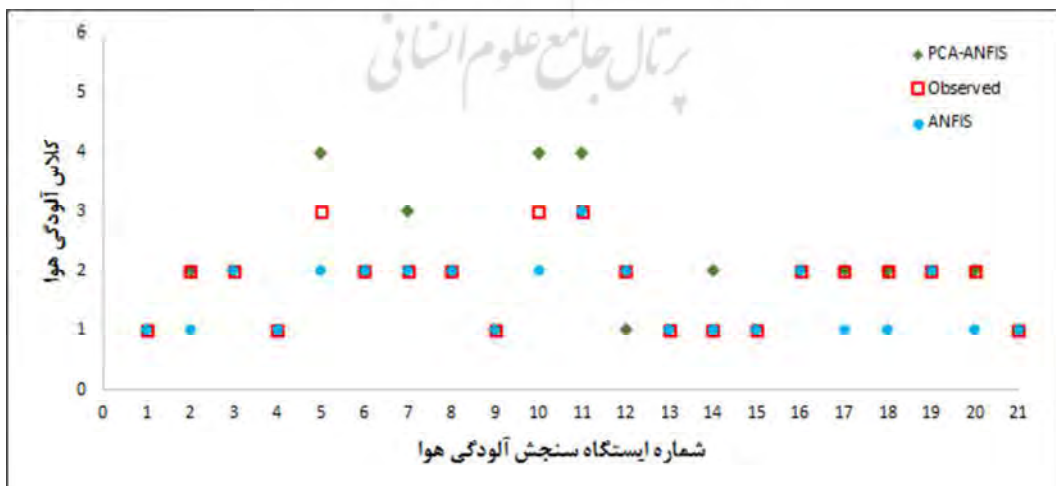
زینب قائمی و همکاران



شکل ۱۴. مقایسه مقادیر مشاهداتی و پیش‌بینی‌شده برای تاریخ ۱۳۹۳/۶/۰۲



شکل ۱۵. مقایسه مقادیر مشاهداتی و پیش‌بینی‌شده برای تاریخ ۱۳۹۳/۰۷/۰۶



شکل ۱۶. مقایسه مقادیر مشاهداتی و پیش‌بینی‌شده برای تاریخ ۱۳۹۳/۱۱/۲۶

۵- نتیجه‌گیری

هدف این پژوهش عرضه سیستمی برای پیش‌بینی آلودگی هوای شهر تهران طی ۲۴ ساعت آینده است. این سیستم می‌تواند داده‌های ایستگاه‌های سنجش آلودگی هوا، هواشناسی و داده‌های مکانی را دریافت و آلودگی هوا را روی ایستگاه‌های سنجش آلودگی هوا پیش‌بینی کند. با توجه به زیان آلودگی هوا برای سلامت انسان و میزان خساراتی که به طبیعت و جامعه وارد می‌کند، استفاده از این سیستم پیش‌بینی آلودگی هوا ضروری می‌نماید. به منظور پیش‌بینی، مدل ترکیبی PCA-ANFIS طراحی شده است. استفاده از تحلیل مؤلفه اصلی موجب حذف همبستگی میان پارامترهای ورودی و افزایش قابلیت اطمینان نتایج حاصل می‌شود. همچنین، با استفاده از این تحلیل، با وجود اینکه ابعاد داده‌های ورودی کاهش می‌یابد، مقدار معتنابهی از پراکندگی موجود در داده‌های ورودی حفظ می‌شود. کاهش ابعاد داده‌های ورودی زمان محاسبات را نیز کاهش می‌دهد. برای ارزیابی روش طراحی شده، این مدل با روش ANFIS در پیش‌بینی آلودگی هوا مقایسه شد که نتایج آن دقت بهتر و سرعت بالاتر روش ترکیبی PCA-ANFIS را، در مقایسه با روش ANFIS، نشان می‌دهد. همچنین، نتایج به‌دست‌آمده بیان می‌کند که در فصل زمستان، بالاترین دقت کلی را حاصل می‌کند. اما در فصل بهار، توانسته در پیش‌بینی هریک از کلاس‌ها بهتر عمل کند. یکی از مشکلات مدل توسعه‌یافته عملکرد ضعیف آن در پیش‌بینی کلاس‌هایی است که تعداد نمونه‌های اندکی دارند. استفاده از راهکارهایی با هدف حل این محدودیت گام بعدی است که برای این پژوهش در نظر گرفته شده است. استفاده از دیگر روش‌ها در پیش‌بینی آلودگی هوا و مقایسه نتایج حاصل از مدل مطرح‌شده با روش‌های دیگر نیز در دستور کار قرار دارد.

۶- منابع

خزاعی، ا.، آل‌شایخ، ع.ا.، کریمی، م.، وحیدنیا، م.ح.، ۲۰۱۳،

مقایسه دو روش مدل‌سازی با استفاده از شبکه

عصبی-فازی در پیش‌بینی غلظت آلاینده

مونوکسید کربن، محیط‌شناسی، دوره ۳۸، شماره ۴.

Alizadeh, M., Jolai, F., Aminnayeri, M. & Rada, R., 2012, **Comparison of Different Input Selection Algorithms in Neuro-Fuzzy Modeling**, Expert Systems with Applications, 39(1), PP. 1536–1544.

Athanasiadis, I.N., Kaburlasos, V.G., Mitkas, P.A. & Petridis, V., 2003, **Applying Machine Learning Techniques on Air Quality Data for Real-Time Decision Support**, First International NAISO Symposium on Information Technologies in Environmental Engineering (ITEE'2003), Gdansk, Poland.

Boznar, M., Lesjak, M. & Mlakar, P., 1993, **A Neural Network-Based Method for Short-Term Predictions of Ambient SO2 Concentrations in Highly Polluted Industrial Areas of Complex Terrain**, Atmospheric Environment, Part B. Urban Atmosphere, 27(2), PP. 221–230.

Brook, R.D., Franklin, B., Cascio, W., Hong, Y., Howard, G., Lipsett, M., Luepker, R., Mittleman, M., Samet, J. & Smith, S.C., 2004, **Air Pollution and Cardiovascular Disease**, Circulation 109(21), PP. 2655–2671.

Brunekreef, B. & Holgate, S.T., 2002, **Air Pollution and Health**, The Lancet 360(9341), PP. 1233–1242.

Chan-Yeung, M., 2000, **Air Pollution and Health**, Hong Kong Medical Journal 6(4), PP. 390–398.

Fernando, H.J., Mammarella, M., Grandoni, G., Fedele, P., Di Marco, R., Dimitrova, R. & Hyde, P., 2012, **Forecasting PM 10 in Metropolitan Areas: Efficacy of Neural Networks**, Environmental Pollution 163, PP. 62–67.

Fisher, B., 2003, **Fuzzy Environmental Decision-Making: Applications to Air Pollution**, Atmospheric Environment 37(14), PP. 1865–1877.

García Nieto, P.J., Combarro, E.F., del Coz Díaz, J.J. & Montañés, E., 2013, **A SVM-Based Regression Model to Study the Air Quality at Local Scale in Oviedo Urban Area (Northern Spain): A Case Study**, Applied Mathematics and Computation 219(17), PP. 8923–8937.

- Ghaemi, Z., Farnaghi, M. & Alimohammadi, A., 2016, **An Online Approach for Spatio-Temporal Prediction of Air Pollution in Tehran Using Support Vector Machine**, Journal of Geospatial Information Technology, 3(4): PP. 43–63.
- Gorai, A.K., Tuluri, F. & Tchounwou, P.B., 2014, **A GIS Based Approach for Assessing the Association between Air Pollution and Asthma in New York State, USA**, International Journal of Environmental Research and Public Health, 11(5), PP. 4845–4869.
- Güler, I. & Übeyli, E.D., 2005, **Adaptive Neuro-Fuzzy Inference System for Classification of EEG Signals Using Wavelet Coefficients**, Journal of Neuroscience Methods, 148(2), PP. 113–121.
- Jain, S. & Khare, M., 2010, **Adaptive Neuro-Fuzzy Modeling for Prediction of Ambient CO Concentration at Urban Intersections and Roadways**, Air Quality, Atmosphere & Health, 3(4), PP. 203–212.
- Jang, J.-S., 1992, **Self-Learning Fuzzy Controllers Based on Temporal Backpropagation**, IEEE Transactions on Neural Networks, 3(5), PP. 714–723.
- Jang, J.-S., 1993, **ANFIS: Adaptive-Network-Based Fuzzy Inference System**, IEEE Transactions on Systems, Man, and Cybernetics, 23(3), PP. 665–685.
- Kolehmainen, M., Martikainen, H., Hiltunen, T. & Ruuskanen, J., 2000, **Forecasting Air Quality Parameters Using Hybrid Neural Network Modelling**, Urban Air Quality: Measurement, Modelling and Management, Springer, PP. 277–286.
- Kumar, A. & Goyal, P., 2011, **Forecasting of Air Quality in Delhi Using Principal Component Regression Technique**, Atmospheric Pollution Research, 2(4), PP. 436–444.
- McKone, T.E. & Deshpande, A.W., 2005, **Can Fuzzy Logic Bring Complex Environmental Problems into Focus?**, ACS Publications.
- Mintz, D., 2012, **Technical Assistance Document for the Reporting of Daily Air Quality-the Air Quality Index (AQI)**, US Environmental Protection Agency, Office of Air Quality Planning and Standards.
- Morabito, F.C. & Versaci, M., 2003, **Fuzzy Neural Identification and Forecasting Techniques to Process Experimental Urban Air Pollution Data**, Neural Networks, 16(3), PP. 493–506.
- Mostafaeipour, A., Zarezade, M., Goudarzi, H., Rezaei-Shouroki, M. & Qolipour, M., 2017, **Investigating the Factors on Using the Solar Water Heaters for Dry Arid Regions: A Case Study**, Renewable and Sustainable Energy Reviews, 78, PP. 157–166.
- Moussiopoulos, N., Sahm, P. & Kessler, C., 1995, **Numerical Simulation of Photochemical Smog Formation in Athens, Greece—a Case Study**, Atmospheric Environment, 29(24), PP. 3619–3632.
- Nia, A.M., Alimohammadi, A., Habibi, R. & Shirzadi, M., 2015, **Spatial and Statistical Analysis of Leptospirosis in Guilan Province, Iran**, The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, 40(1), P. 497.
- Niska, H., Hiltunen, T., Karppinen, A., Ruuskanen, J. & Kolehmainen, M., 2004, **Evolving the Neural Network Model for Forecasting Air Pollution Time Series**, Engineering Applications of Artificial Intelligence, 17(2), PP. 159–167.
- Numari, G., 2000, **Simplified Fuzzy Modelling of Pollutant Time Series**, Neural Network World, 10(6), PP. 983–1000.
- Pérez, P., Trier, A. & Reyes, J., 2000, **Prediction of PM 2.5 Concentrations Several Hours in Advance Using Neural Networks in Santiago, Chile**, Atmospheric Environment, 34(8), PP. 1189–1196.
- Riahi-Madvar, H., Ayyoubzadeh, S.A., Khadangi, E. & Ebadzadeh, M.M., 2009, **An Expert System for Predicting Longitudinal Dispersion Coefficient in Natural Streams by Using ANFIS**, Expert Systems with Applications, 36(4), PP. 8589–8596.
- Robeson, S. & Steyn, D., 1990, **Evaluation and Comparison of Statistical Forecast Models for Daily Maximum Ozone Concentrations**, Atmospheric Environment, Part B, Urban Atmosphere, 24(2), PP. 303–312.
- Ross, T.J., 2009, **Fuzzy Logic with Engineering Applications**, John Wiley & Sons.
- Sahoo, M.M., Patra, K. & Khatua, K., 2015, **Inference of Water Quality Index Using ANFIA and PCA**, Aquatic Procedia, 4: PP. 1099–1106.
- Shahraiyini, H.T., Sodoudi, S., Kerschbaumer, A. & Cubasch, U., 2015, **A New Structure**

- Identification Scheme for ANFIS and its Application for the Simulation of Virtual Air Pollution Monitoring Stations in Urban Areas**, Engineering Applications of Artificial Intelligence, 41, PP. 175–182.
- Sindelar, R. & Babuska, R., 2004, **Input Selection for Nonlinear Regression Models**, IEEE Transactions on Fuzzy Systems, 12(5), PP. 688–696.
- Singh, K.P., Gupta, S. & Rai, P., 2013, **Identifying Pollution Sources and Predicting Urban Air Quality Using Ensemble Learning Methods**, Atmospheric Environment, 80, PP. 426–437.
- Statheropoulos, M., Vassiliadis, N. & Pappa, A., 1998, **Principal Component and Canonical Correlation Analysis for Examining Air Pollution and Meteorological Data**, Atmospheric Environment, 32(6), PP. 1087–1095.
- Taylan, O., 2017, **Modelling and Analysis of Ozone Concentration by Artificial Intelligent Techniques for Estimating Air Quality**, Atmospheric Environment, 150, PP. 356–365.
- Ul-Saufie, A.Z., Yahaya, A.Z., Ramli, N.A., Rosaida, N. & Hamid, H.A., 2013, **Future Daily PM 10 Concentrations Prediction by Combining Regression Models and Feedforward Backpropagation Models with Principle Component Analysis (PCA)**, Atmospheric Environment, 77, PP. 621–630.
- Williams, B., Onsmann, A. & Brown, T., 2010, **Exploratory Factor Analysis: A Five-Step Guide for Novices**, Australasian Journal of Paramedicine, 8(3).
- Zhang, J., Wang, C.-m., Liu, L., Guo, H., Liu, G.-d., Li, Y.-w. & Deng, S.-h., 2014, **Investigation of Carbon Dioxide Emission in China by Primary Component Analysis**, Science of The Total Environment, 472, PP. 239–247.
- Zheng, Y., Liu, F. & Hsieh, H.-P., 2013, **U-Air: When Urban Air Quality Inference Meets Big Data**, Proceedings of the 19th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, ACM.
- Ziomas, I.C., Melas, D., Zerefos, C.S., Bais, A.F. & Paliatsos, A.G., 1995, **Forecasting Peak Pollutant Levels from Meteorological Variables**, Atmospheric Environment, 29(24), PP. 3703–3711.