

مقایسه دو روش مدل سازی با استفاده از شبکه عصبی - فازی در پیش بینی غلظت آلاینده مونوکسید کربن

اله خزاعی^{۱*}، علی اصغر آل شیخ^۲، محمد کریمی^۳، محمد حسن وحیدنیا^۴

۱-دانشجوی کارشناسی ارشد GIS، دانشکده مهندسی نقشهبرداری، دانشگاه صنعتی خواجه نصیرالدین طوسی

۲- دانشیار، دانشکده مهندسی نقشهبرداری، دانشگاه صنعتی خواجه نصیرالدین طوسی Alesheikh@kntu.ac.ir

۳- استادیار، دانشکده مهندسی نقشهبرداری، دانشگاه صنعتی خواجه نصیرالدین طوسی Mkarimi@kntu.ac.ir

۴- دانشجوی دکتری GIS، دانشکده مهندسی نقشهبرداری، دانشگاه صنعتی خواجه نصیرالدین طوسی Vahidnia84@gmail.com

تاریخ دریافت: ۹۰/۱۱/۹ تاریخ پذیرش: ۹۱/۴/۱۷

چکیده

پایش و پیش بینی مشخصه های کیفیت هوا در مناطق شهری یکی از چالش های محیط زیست انسانی محسوب می شود. این مهم وابسته به عوامل متعددی مانند توپوگرافی، اقلیم، جمعیت و شبکه حمل و نقل است که نحوه تعامل این عوامل مکانی به عنوان پدیده ای دینامیک، غیر خطی و دارای ابهام عنوان شده است. در این تحقیق به منظور پیش بینی و مدل سازی میزان آلاینده مونوکسید کربن از شبکه عصبی- فازی و GIS در قالب دو مدل متفاوت استفاده شده است. در مدل اول داده های آموزشی با استفاده از روش کریجینگ ایجاد و برای هر ایستگاه ناحیه ای در نظر گرفته شد که از داده های موجود در آن ناحیه، برای آموزش آن استفاده شد. برای هر ناحیه قانون های فازی استخراج شده و به هر پیکسل آن ناحیه اعمال و میزان آلاینده تخمین زده شد. در مدل دوم هر ایستگاه به طور جداگانه با استفاده از داده های خود آموزش داده شد. برای هر ایستگاه قانون های فازی آن استخراج شده و غلظت آلاینده تخمین زده شد. به علت این که پیش بینی در ایستگاهها صورت می گیرد، در نهایت برای مدل سازی مکانی غلظت در محدوده مورد مطالعه از روش کریجینگ استفاده شده است. برای انجام تست عملی این تحقیق، از داده های هواشناسی ایستگاههای واقع در سطح شهر تهران استفاده شد. میانگین RMSE مجموعه ایستگاهها در مدل اول با قانون های سوگنو، ۱/۶۱۳ ppm و با قانون های ممدانی، ۱/۴۸۴ ppm و در مدل دوم با قانون های سوگنو، ۱/۴۴۵ ppm و با قانون های ممدانی، ۱/۳۷۴ ppm به دست آمد. نتایج حاکی از آن است که هر دو مدل به خوبی میزان آلاینده را پیش بینی می کنند.

کلید واژه

آلودگی هوا، سیستم اطلاعات مکانی، شبکه عصبی- فازی، سیستم استنتاج فازی ممدانی، کریجینگ

سراغاز

شود که عوامل متعددی را در نظر بگیریم باز هم عواملی از جمله افزایش تولیدات ناگهانی خودرو، معماری، شهر سازی و ... هستند که باعث نادرست شدن مدل و پیش بینی انجام شده می شوند. به دلیل وجود عدم قطعیت در اکثر موجودیت های مکانی نیاز به روشی است که با در نظر گرفتن عدم قطعیت توانایی یادگیری و استنتاج را داشته باشد. سیستم فازی قادر به تجزیه و تحلیل اطلاعات زبانی و اجرای پردازشها با نمایش دانش و استدلال با در نظر گرفتن عدم قطعیت است. یکی از اجزای اصلی سیستم استنتاج فازی، پایگاه دانش است که متشکل از قانون های اگر-آنگاه هستند. سیستم فازی با استفاده از قواعد استنتاج فازی قادر به ارائه تصمیم گیری و پاسخ

آلودگی هوا یکی از مشکلات قرن حاضر است که اکثر شهرهای بزرگ جهان با این پدیده روبه رو هستند. در واقع یکی از مهم ترین مسائل زیست محیطی به شمار می رود که به خودی خود می تواند زمینه ساز بسیاری از خطرهای زیست محیطی دیگر باشد. آلودگی هوا فرایندی بسیار پیچیده بوده که وابسته به عوامل متعددی از جمله توپوگرافی، اقلیم، جمعیت، شبکه حمل و نقل، صنعت و... است، بنابراین پیش بینی این گونه داده ها که دارای دینامیک غیر خطی اند بسیار مشکل است و می باید چگونگی پراکندگی و انتشار مواد آلاینده را در جو مشخص کرد. اگر برای مدل سازی سعی

سیستم فرمول ریاضی بوده و چارچوبی برای نمایش دانش بشری فراهم نمی‌کند، بنابراین در این تحقیق قانون‌ها در هر دو سیستم استنتاج فازی ممدانی و سوگنو ایجاد شد که درک و تفسیر پذیری آن آسان باشد.

به علت اهمیت پیش‌بینی و مدل‌سازی مکانی مشخصه‌های کیفیت هوا، در این تحقیق با بهره‌گیری از شبکه فازی-عصبی و GIS، دانش حاکم بر محیط را در قالب قانون‌های فازی، از داده‌ها استخراج کرده و با استفاده از این قانون‌ها، مشخصه آلودگی هوا پیش‌بینی و مدل‌سازی شد. برای این منظور دو مدل بررسی شد، که در مدل اول قانون‌های حاکم بر ناحیه پوشش ایستگاه و در مدل دوم قانون‌های موجود بر ایستگاه‌های پایش، بر اساس مشخصه‌های هواشناسی استخراج شد. در ادامه، چگونگی انجام این کار شرح داده شده است. ابتدا به مروری بر نظریه‌های مورد استفاده و سپس روش کار بیان شده است.

مواد و روش بررسی

در این بخش ابتدا با تشریح ویژگی‌های منطقه مورد مطالعه، نقش مشخصه‌های هواشناسی مؤثر در میزان آلودگی هوا و نظریه شبکه عصبی - فازی و روش کریجینگ، به بیان مراحل اجرا و عملی کردن مدل‌های توسعه داده شده پرداخته می‌شود.

منطقه مورد مطالعه

شهر تهران از آلوده‌ترین شهرهای جهان به شمار می‌رود. عوامل متعددی در آلودگی شهر تهران مؤثرند که در بین آنها عوامل جغرافیایی و هواشناسی دارای اهمیت‌اند. کوه‌های اطراف بخصوص کوه‌های البرز در شمال که تهران را احاطه کرده‌اند مانند سدی عمل می‌کنند که باعث تجمع آلاینده‌ها در سطح شهر می‌شوند. عوامل هواشناسی از جمله وارونگی‌های دمایی و استقرار مداوم سامانه‌های پر فشار همراه با هوای پایدار در دوره‌های سرد سال باعث افزایش آلاینده‌های هوا می‌شوند (صفوی، علیجانی، ۱۳۸۵). همچنین مطالعات نشان می‌دهد که بیش از ۷۰ درصد از مونوکسیدکربن منتشر شده در هوا در جریان عملیات حمل‌ونقل و حرکت خودروها تولید می‌شود (افیونی، ۱۳۸۸). گاز مونوکسیدکربن در هوای آزاد و به مقدار کم، زندگی بیماران قلبی و ریوی را به خطر می‌اندازد و در افراد سالم باعث سردرد، سرگیجه، خستگی زیاد و تحریک اعصاب می‌شود. بنابراین به علت اهمیت این آلاینده بر سلامتی و زیاد بودن مقدار این آلاینده در هوای تهران، آلاینده مونوکسید کربن برای مدل‌سازی انتخاب شد. ایستگاه‌های پیش

است. تعیین فرم مناسب این قواعد می‌تواند به سیستم توانایی لازم برای تصمیم‌گیری کمک کند.

در سالهای اخیر تکنیک جدید هوش مصنوعی با عنوان محاسبات نرم توسعه داده شده است که به تلفیق روش‌های هوش مصنوعی از قبیل شبکه عصبی و سیستم استنتاج فازی کمک کرده است. منطق فازی مکانیسم استنتاج را با عدم قطعیت شناخت انجام می‌دهد و شبکه عصبی نیز توانایی‌هایی از قبیل یادگیری، انطباق، پردازش موازی را داراست. سیستم ترکیبی شبکه فازی عصبی نامیده می‌شود (Fuller, 1995).

تحقیقات زیادی برای تلفیق توانایی یادگیری شبکه عصبی با سیستم استنتاج فازی به منظور استخراج قانون‌های اولیه سیستم فازی و بهینه کردن توابع عضویت صورت گرفته است (Lin, Lee, 1991; Horikawa, Furuhashi, et al., 1992; Jang, 1993; Ishibuchi, et al., 1995). از مطالعاتی که در شبکه عصبی - فازی به منظور پیش‌بینی غلظت آلاینده‌های موجود در هوا به کار برده شده می‌توان به موارد زیر اشاره کرد: استفاده از سیستم عصبی - فازی برای مدل کردن و پیش‌بینی غلظت هیدرو کربنات در هوای شهری (Versaci Morabito, 2003)، پیش‌بینی ماکزیمم غلظت روزانه ازن (Heo, Kim, 2004)، تخمین میزان آلودگی SO₂ و ذرات معلق (TSP)¹ در ناحیه شهری بر اساس آثار شاخص‌های هواشناسی (Yildirim, Bayramoglu, 2006)، پیش‌بینی متوسط یک ساعته غلظت مونوکسیدکربن (Jain, Khare, 2010).

سیستم اطلاعات جغرافیایی به عنوان ابزاری توانمند و کارآمد در طراحی و ایجاد پایگاه اطلاعات آلودگی هوا به منظور جمع‌آوری، ذخیره، بازیابی و تجزیه و تحلیل آلودگی هوا معرفی می‌شود. از روش‌های مدل‌سازی آلودگی هوا با استفاده از GIS می‌توان به درون‌یابی مکانی (Goovaerts, et al., 2006; Janssen, et al., 2008)، مدل‌های مربوط به انتشار و پراکندگی (Xiwen, 2010; Elbir, et al., 2010; Wenjun, 2010) و مدل رگرسیون کاربری (Smith, et al., 2006; Brauer, et al., 2003; Briggs, et al., 2005; Jerrett, et al., 2000) اشاره کرد.

در تحقیقات انجام شده، شبکه عصبی-فازی برای پیش‌بینی غلظت آلاینده‌ها در یک یا چند ایستگاه استفاده شده است ولی برای تخمین میزان آلاینده در ناحیه به کار نرفته است. همچنین در این تحقیقات قانون‌های فازی با استفاده از سیستم استنتاج فازی سوگنو توسعه داده شده است. از آنجایی که بخش تالی قانون‌های این

—رطوبت نسبی: این مشخصه نشان‌دهنده میزان آب موجود در هوا به میزان آب در حالت اشباع است. از آثار رطوبت می‌توان میزان چسبندگی بیشتر در ذرات و ایجاد ذرات درشت‌تر را نام برد. در نتیجه افزایش رطوبت سرعت ته‌نشینی ذرات طبق قانون استوکس بالا می‌رود و پیش‌بینی می‌شود آلودگی کاهش می‌یابد.

—دما: توزیع افقی و توزیع عمودی درجه حرارت اثر مستقیم در ایجاد جریان‌های هوا و در نتیجه انتقال آلودگی هوا دارد. همچنین وجود دما باعث مجموعه‌ای از تغییرات شیمیایی می‌شود که می‌تواند آلودگی هوا را تشدید کند (افیونی، ۱۳۸۸).

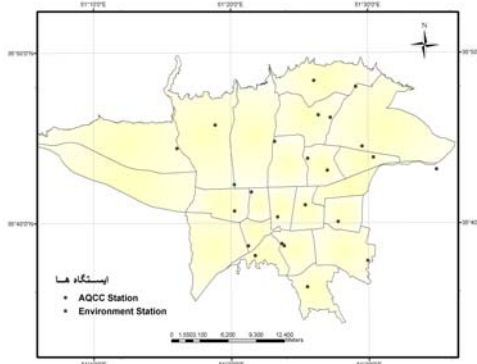
—سرعت و جهت باد: باد یکی از عوامل مهم در پخش و انتقال عوامل آلوده‌کننده محسوب می‌شود. درجه پراکندگی مواد به جریان‌های افقی و عمودی باد بستگی دارد. بنابراین میزان انتشار و انتقال هوای آلوده به سرعت و جهت باد بستگی دارد. معمولاً باد در سرعت‌های بیش از ۳۰ کیلومتر در ساعت در انتقال و انتشار آلودگی مؤثر است. جریان‌های عمودی باد از طریق بالا بردن آلاینده‌ها سبب دور شدن آنها از سطح زمین می‌شود. با توجه به افزایش سرعت جریان افقی باد در طبقات بالای جو، انتقال آلودگی‌ها به قسمت‌های فوقانی سبب پراکندگی بهتر آنها می‌شود (افیونی، ۱۳۸۸).

—نزولات جوی: نزولات جوی مانند برف و باران تغییرات مهمی در وضعیت آلودگی هوا به وجود می‌آورند. باران بسیاری از گازها و ذرات قابل انحلال را در خود حل می‌کند. مقدار جذب تابعی از قطر قطرات باران، قطر ذرات آلوده‌کننده، سرعت سقوط، سرعت آنها نسبت به یکدیگر و تراکم ذرات موجود در هواست (افیونی، ۱۳۸۸).

پدیده‌های دیگر هواشناسی نیز مانند وارونگی دما، فشار هوا، شدت نور خورشید، اوقات ابری و ساعات آفتابی بر آلودگی هوا مؤثرند (افیونی، ۱۳۸۸).

به علت این که در شبکه عصبی-فازی، تعداد زیاد ورودی‌ها نه فقط از شفافیت مدل می‌کاهند بلکه پیچیدگی محاسبات را زیاد می‌کند (Jang, 1996)، مشخصه‌هایی که تأثیرپذیری بیشتری بر غلظت آلاینده مورد نظر دارند، انتخاب شد. ابتدا برای انتخاب مشخصه‌های مناسب با توجه به مشخصه‌های موجود برای هر ایستگاه و با استفاده از ماتریس همبستگی، میزان همبستگی CO به مشخصه‌های دما (Temp)^۲، سرعت باد (WS)^۳، جهت باد (WD) (۱)، رطوبت نسبی (Hum)^۵ و فشار (Press)^۶ سنجیده شد، که جدول شماره (۱) این مقادیر را برای ایستگاه امام خمینی نشان می‌دهد.

مشخصه‌های کیفیت هوا واقع در شهر تهران زیر نظر سازمان حفاظت محیط زیست و شرکت کنترل کیفیت هوا وابسته به شهرداری تهران است. ایستگاه‌های سازمان حفاظت محیط زیست علاوه بر مشخصه‌های کیفیت هوا به طور همزمان مشخصه‌های هواشناسی را نیز پایش می‌کنند. برای انجام تحقیق، داده‌های هواشناسی (دما، رطوبت، فشار، سرعت باد، جهت باد) و غلظت آلاینده‌ی مونوکسیدکربن ایستگاه‌های واقع در شهر تهران برای فصل تابستان برای چهار سال متوالی (۸۶ تا ۸۹) از سازمان حفاظت محیط زیست و همچنین داده‌های غلظت آلاینده مونوکسیدکربن از شرکت کنترل کیفیت هوا تهیه شد. شکل شماره (۱) موقعیت ایستگاه‌ها را نشان می‌دهد. داده‌های جمع‌آوری شده با فاصله زمانی یک ساعته بودند و به علت این که در این تحقیق متوسط ۸ ساعته غلظت آلاینده پیش‌بینی می‌شود، میانگین ۸ ساعته داده‌ها محاسبه شد، بدین ترتیب که هر روز سه نوبت ۸ ساعته از ۱۲ شب تا ۸ صبح، از ۸ صبح تا ساعت ۱۶ بعد از ظهر و از ساعت ۱۶ تا ۲۴ در نظر گرفته شد.



شکل شماره (۱): موقعیت ایستگاه‌های پایش واقع در تهران

مشخصه‌های هواشناسی مؤثر در میزان آلودگی هوا

کیفیت هوا به طور روزانه در تغییر است؛ حتی در مواقعی که مقدار ورود آلاینده‌ها به هوا ثابت است عوامل تعیین‌کننده تغییرات آب و هوایی مانند سرعت باد، جهت باد، نيمرخ حرارتی توده‌های هوا، مقدار انرژی خورشیدی به منظور انجام واکنش‌های فتوشیمیایی، مدت زمان دوام باد، یا بارندگی به طور ویژه‌ای کیفیت هوا را تغییر می‌دهند (افیونی، ۱۳۸۸). در این قسمت ویژگی‌های کلی مشخصه‌های هواشناسی مؤثر در میزان آلودگی هوا ارائه می‌شود.

جدول شماره (۱): ماتریس همبستگی

| Hum | Press | WS | WD | Temp | CO | |
|-----|--------|--------|--------|--------|--------|-------|
| | | | | | ۱ | CO |
| | | | | ۱ | ۰/۱۷۷ | Temp |
| | | | ۱ | ۰/۳۳۷ | -۰/۲۱۶ | WD |
| | | ۱ | ۰/۲۶۹ | ۰/۲۹۱ | -۰/۴۷۸ | WS |
| | ۱ | -۰/۰۳۴ | -۰/۰۷۳ | -۰/۱۶۶ | ۰/۰۱۹ | Press |
| ۱ | -۰/۱۸۸ | -۰/۳۳۷ | -۰/۱۱۳ | -۰/۷۴۲ | -۰/۱۱۸ | Hum |

لایه ۳: هر گره در این لایه عمل محاسبه وزن نسبی قانون‌ها را انجام می‌دهد.

لایه ۴: هر گره در این لایه دارای تابع گره است که به تمام ورودی‌ها و یک گره در لایه سوم متصل است.

لایه ۵: گره در این لایه گره ثابت است و وظیفه جمع بندی تمام خروجی قانون‌ها را دارد.

روشهای مختلفی برای تغییر و بهینه کردن مشخصه‌های این مدل وجود دارد. برای نمونه استفاده از روش پس انتشار خطا برای برآورد کلیه مشخصه‌ها، روش پس انتشار خطا و یک عبور روش کمترین مربعات، ترکیب روش پس انتشار خطا و کمترین مربعات^۹ و همچنین روش کمترین مربعات ترکیبی (Jang, 1993).

سیستم استنتاج عصبی - فازی تطبیقی مددانی (M-ANFIS)^{۱۰}

این سیستم را در سال ۲۰۰۹ Yuanyuan Cha و همکاران معرفی کردند. این مدل تعمیمی از سیستم استنتاج عصبی - فازی تطبیقی است که برای اجرای سیستم فازی مددانی در ساختار شبکه عصبی استفاده می‌شود و دارای ۵ لایه است که به طور خلاصه می‌توان عملکرد لایه‌ها را به صورت زیر بیان کرد:

لایه ۱: این لایه وظیفه فازی سازی متغیرهای ورودی را بر عهده دارد.

لایه ۲: هر گره در این لایه به گره‌هایی که نمایش دهنده پیش فرض‌های قانون مورد نظر است وصل می‌شود. وزن قانون‌ها توسط عملگر ضرب ایجاد می‌شود.

لایه ۳: عملگر استلزام^{۱۱} در این لایه ضرب است.

لایه ۴: این لایه وظیفه جمع آوری نتایج حاصل از لایه قبل را دارا است، که عملگر جمع این کار را انجام می‌دهد.

لایه ۵: غیرفازی‌ساز مرکز ثقل عمل غیرفازی‌سازی را انجام می‌دهد. همانند مدل قبل روشهای مختلفی برای تغییر و بهینه کردن مشخصه‌ها این مدل وجود دارد. انتخاب روش مناسب بستگی

با توجه به مقادیر به دست آمده از ماتریس همبستگی، در بعضی از ایستگاهها همبستگی فشار، نسبت به جهت باد به آلاینده CO بیشتر بود و برعکس. با مقایسه ماتریس همبستگی در تمام ایستگاهها، معلوم شد که در ایستگاههای بیشتری همبستگی برای جهت باد، نسبت به فشار به آلاینده CO بیشتر است، بنابراین مشخصه‌های دما، سرعت باد و جهت باد برای تخمین آلاینده CO انتخاب شد.

تئوری شبکه عصبی - فازی

تمام تکنیک‌های ترکیبی شبکه عصبی و سیستم‌های فازی، شبکه عصبی- فازی نامیده می‌شود. ترکیب‌های مختلف این تکنیک‌ها، سیستم‌های متفاوتی از قبیل سیستم‌های عصبی- فازی Cooperative, Concurrent و Hybrid را به وجود آورده است (VIEIRA, 2004). در ادامه، دو روش معمول توسعه سیستم‌های عصبی فازی Hybrid که در این تحقیق مورد استفاده قرار می‌گیرند، ارائه می‌شوند.

سیستم استنتاج عصبی - فازی تطبیقی (ANFIS)^۷

Jang این سیستم را در سال 1993 معرفی نمود که برای اجرای سیستم فازی سوگنو در ساختار شبکه عصبی استفاده می‌شود. قانون‌ها در این شبکه‌ها باید مشخص باشد و تنها توابع عضویت برای مقدم و مؤخر قانون را مشخص می‌کند. این مدل دارای ۵ لایه است که عملکرد این لایه‌ها را مختصراً می‌توان به صورت زیر بیان کرد:

لایه ۱: توابع عضویت متغیرهای ورودی به عنوان تابع گره استفاده می‌شود.

لایه ۲: هر گره در این لایه به گره‌هایی که نمایش دهنده پیش فرض‌های قانون مورد نظر است وصل می‌شود. هر گره در این لایه گره ثابت است که در آن سیگنال‌های ورودی در هم ضرب می‌شوند و وزن^۸ یک قانون را بیرون می‌فرستد.

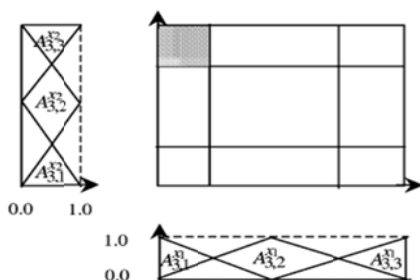
نقطه‌ای که دارای بالاترین پتانسیل است به عنوان مرکز خوشه انتخاب می‌شود. پس از آن اندازه پتانسیل هر نقطه تصحیح می‌شود، به طوری که پتانسیل داده‌هایی که نزدیک اولین مرکز خوشه هستند به طور چشمگیری کاهش می‌یابد (Hammouda, 2000).

پس از تصحیح پتانسیل، مرکز خوشه بعدی با دارا بودن بیشترین پتانسیل انتخاب می‌شود. روند اخذ خوشه جدید و تصحیح پتانسیل تا وقتی ادامه پیدا می‌کند که پتانسیل تمام نقاط زیر کسری از پتانسیل مرکز اولین خوشه باشد. حاصل خوشه‌بندی فازی k مرکز خوشه به ابعاد مجموع تعداد متغیرهای ورودی و خروجی است. برای هر خوشه یک قانون تعریف می‌شود که رفتار سیستم را نشان می‌دهد (Chiu, 1994).

روش تقسیم بندی گریدی^{۱۳}

روش تقسیم بندی گریدی دارای ساختار گریدی است. تفسیرپذیری آن آسان است و به طور گسترده برای ایجاد قانون‌های فازی استفاده می‌شود. این روش برای تقسیم بندی فضای ورودی مورد استفاده قرار می‌گیرد که در سیستم فازی سوگنو کاربرد دارد.

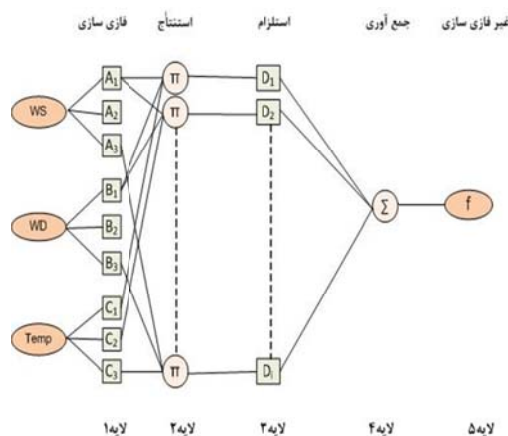
مشکل اساسی این است که با افزایش تعداد متغیرهای ورودی، تعداد قانون‌ها به صورت نمایی افزایش می‌یابد. برای هر قانون تعدادی مشخصه باید بهینه شود که با افزایش تعداد متغیرها تعداد مشخصه‌های زیادی در الگوریتم‌های یادگیری باید بهینه شود. تعداد زیاد قانون‌های فازی تفسیر پذیری سیستم را کاهش می‌دهد. بنابراین این روش برای مجموعه داده‌ها با ابعاد کم و پوشش خوب مناسب است. (Wu, et al., 2011) برای مثال شکل شماره (۳) ساختار آن را برای فضای ورودی دو متغیره و هر متغیر دارای ۳ تابع عضویت نمایش می‌دهد.



شکل شماره (۳): تقسیم بندی گریدی با دو متغیر و هر کدام سه

تابع عضویت (Hua, et al., 2003)

به رابطه بین پیچیدگی محاسبات و عملکرد نتایج دارد (Chai, 2009). با توجه به توضیحات داده شده، به طور کلی ساختار عمومی سیستم استنتاج عصبی-فازی تطبیقی را می‌توان به صورت شکل شماره (۲) نمایش داد. در شکل این ساختار برای سه متغیر ورودی و برای هر متغیر، سه تابع عضویت رسم شده است.



شکل شماره (۲): ساختار کلی سیستم استنتاج عصبی-فازی تطبیقی

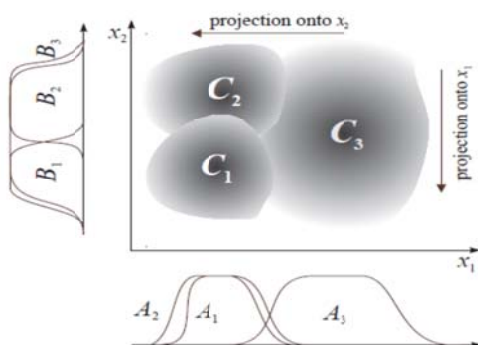
همان طور که بیان شد، در سیستم استنتاج عصبی-فازی تطبیقی قانون‌ها و توابع عضویت اولیه باید مشخص باشد و این سیستم فقط مشخصه‌های توابع عضویت را بهینه می‌کند. برای تعیین قانون‌ها و توابع عضویت اولیه از داده‌ها (بدون دانش کارشناسی) روش‌های متفاوتی وجود دارد که در این تحقیق از روش خوشه بندی کاهشی، تقسیم بندی گریدی و تقسیم بندی فضای ورودی - خروجی استفاده شده است.

همچنین در روش‌های خوشه بندی امکان ایجاد توابع عضویت با همپوشانی زیاد وجود دارد که با استفاده از روش کاهش تعداد توابع عضویت برای هر متغیر، تفسیرپذیری قانون‌ها بهبود یافت. در ادامه هر یک از موارد فوق به اختصار تشریح می‌شود.

روش خوشه بندی کاهشی^{۱۲}

خوشه بندی کاهشی یکی از روش‌های معمول ایجاد قانون‌های فازی اولیه است. در این الگوریتم هر نقطه به عنوان یک پتانسیل برای مرکز خوشه در نظر گرفته می‌شود.

پتانسیل تخصیص داده شده به مرکز هر خوشه به فاصله آن از نقاط دیگر وابسته است. بعد از محاسبه پتانسیل برای هر نقطه،



شکل شماره (۴): توابع عضویت اضافی که به وسیله خوشه‌بندی به وجود می‌آید (Setnes, et al., 1998)

درونیابی به روش کریجینگ

به طور کلی تخمین زمین آماری فرایندی است که طی آن می‌توان مقدار کمیت در نقاطی با مختصات معلوم را با استفاده از مقدار همان کمیت در نقاط دیگری با مختصات معلوم به دست آورد. کریجینگ روشی تخمینی است که بر منطق میانگین متحرک وزن دار استوار است و در مورد آن می‌توان گفت که بهترین تخمین گر خطی نا اریب است. از مهم‌ترین ویژگی‌های کریجینگ آن است که به ازای هر تخمینی خطای مرتبط با آن را می‌توان محاسبه کرد. این تخمین گر به صورت رابطه (۲) تعریف می‌شود (حسنی پاک، ۱۳۷۷):

$$\hat{z}(x_0) = \sum_{i=1}^n \lambda_i \cdot z(x_i) \quad (2)$$

که در آن $z(x_0)$ مشخصه تخمینی، λ_i وزن یا اهمیت کمیت وابسته به نمونه‌ی i ام و $z(x_i)$ مشخصه معلوم است. این نوع را کریجینگ خطی می‌نامند زیرا ترکیب خطی از n داده است. شرط استفاده از این تخمین گر آن است که متغیر مورد نظر توزیع نرمال داشته باشد. در غیر این صورت باید از کریجینگ غیرخطی استفاده کرد یا با بهره‌گیری از تبدیل‌های آماری، توزیع متغیر را نرمال کرد. این تخمین گر اولاً عاری از خطای نظاممند است. ثانیاً واریانس تخمین آن حداقل است. شرط نا اریب بودن این محدودیت را مجموع ضرایب کریجینگ باید معادل واحد باشد به وجود می‌آورد. رابطه (۳) این شرط را نشان می‌دهد (حسنی پاک، ۱۳۷۷).

$$\sum_{i=1}^n \lambda_i = 1 \quad (3)$$

روش تقسیم بندی فضای ورودی و خروجی

این روش را Wang و Mendel در سال ۱۹۹۲ ارائه دادند که با تقسیم بندی فضای ورودی و خروجی به استخراج توابع عضویت و قانون‌های فازی می‌پردازد. مراحل انجام این روش به طور مختصر عبارت است از:

- تقسیم دامنه متغیرها به m ناحیه فازی. (m می‌تواند از یک متغیر به متغیر دیگر متفاوت باشد)
- تعیین درجه عضویت هر داده آموزشی به ناحیه‌های مختلف و سپس برای هر متغیر تابع عضویتی که دارای درجه عضویت ماکزیمم است، انتخاب می‌شود. بنابراین برای هر داده آموزشی یک قانون ایجاد می‌شود.

- نسبت دادن درجه به هر قانون که از ضرب درجه عضویت متغیرهای ورودی و خروجی هر قانون به دست می‌آید.
- انتخاب قانون با ماکزیمم درجه از بین قانون‌هایی که دارای ناسازگاری با یکدیگرند. (منظور از قانون‌های ناسازگار قانون‌هایی است که دارای مقدمه یکسان ولی نتیجه متفاوت‌اند).
بدین ترتیب قانون‌های فازی ایجاد می‌شود (Wang, Mendel, 1992؛ تشنه لب، ۱۳۸۸).

کاهش تعداد توابع عضویت برای هر متغیر و ساده سازی قانون‌ها

در روش‌های خوشه‌بندی امکان ایجاد توابع عضویت با همپوشانی زیاد برای یک متغیر وجود دارد که این توابع عضویت تفسیر پذیری قانون‌ها را کاهش می‌دهند (شکل شماره ۴)، بنابراین برای حذف این توابع عضویت از رابطه ۱ استفاده می‌شود (Setnes, et al., 1998; Chen, et al., 2004):

$$S(A_{ij}, A_{kj}) = \frac{\sum_{l=1}^L \min\{u_{ij}(x_{jl}), u_{kj}(x_{jl})\}}{\sum_{l=1}^L \max\{u_{ij}(x_{jl}), u_{kj}(x_{jl})\}} \quad (1)$$

در این رابطه z نشان دهنده هر متغیر، i و k نشان دهنده تابع عضویت آن متغیر و l نمایانگر داده‌های آن متغیر است. اگر مقدار $S > \lambda$ باشد، آن دو تابع عضویت دارای همپوشانی زیادی هستند و می‌باید تابع عضویت جدیدی با میانگین‌گیری مشخصه‌های آن دو تابع عضویت، جایگزین آن دو مجموعه فازی کرد. λ را مقداری بین ۰/۶ تا ۰/۸۵ تعریف می‌کنند.

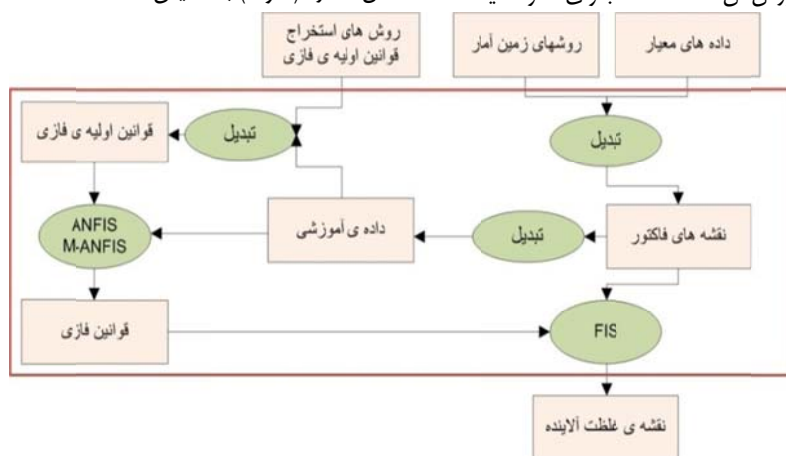
بدین ترتیب بدون این‌که دقت تغییر محسوسی کند، تفسیرپذیری قانون افزایش می‌یابد (Chen, et al., 2004).

روش کار

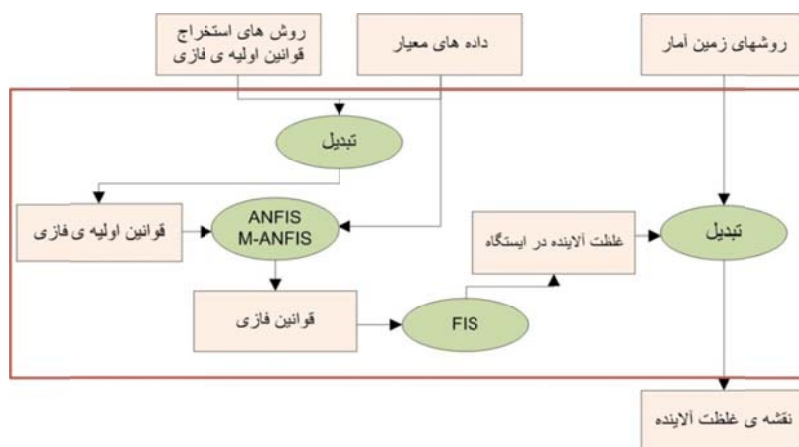
در این قسمت مراحل اجرا و عملی کردن دو مدل ارائه شده در این تحقیق تشریح می شود. در این تحقیق به منظور پیش بینی و مدل سازی میزان آلاینده مونوکسیدکربن از شبکه عصبی- فازی و GIS در قالب دو مدل متفاوت استفاده شده است.

در مدل اول داده های آموزشی با استفاده از روش کریجینگ ایجاد و برای هر ایستگاه ناحیه ای در نظر گرفته شد که از داده های موجود در آن ناحیه، برای آموزش آن استفاده شد. برای هر ناحیه

قانون های فازی استخراج شده و به هر پیکسل آن ناحیه اعمال و میزان آلاینده تخمین زده شد. در مدل دوم هر ایستگاه به طور جداگانه با استفاده از داده های خود آموزش داده شد. برای هر ایستگاه قانون های فازی آن استخراج و غلظت آلاینده تخمین زده شد. به علت این که پیش بینی در ایستگاهها صورت می گیرد، در پایان برای مدل سازی مکانی غلظت در محدوده مورد مطالعه از روش کریجینگ استفاده شده است. چارچوب کلی این دو مدل در شکل شماره (۵ و ۶) به نمایش گذاشته شده است.



شکل شماره (۵): مدل مفهومی روش اول



شکل شماره (۶): مدل مفهومی روش دوم

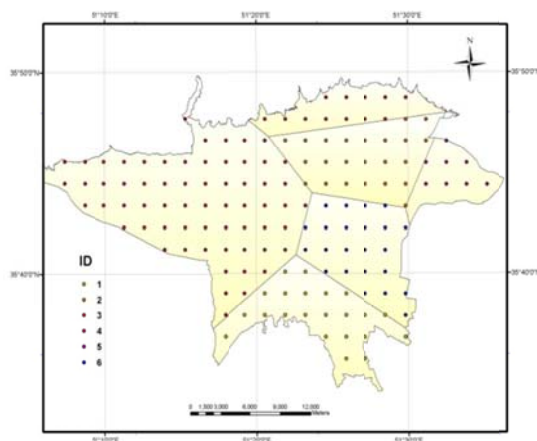
پیاده سازی مدل اول

مراحل انجام این مدل شامل موارد ذیل است.

- تعیین منطقه پوشی هر ایستگاه

برای تعیین منطقه نفوذ هر ایستگاه از چندضلعی های تیسسن استفاده شد. چندضلعی های تیسسن سعی در تعریف هندسی منطقه

نفوذ نقطه به صورت سطحی، یا منطقه ای دارند. در این روش تقسیم بندی کاملاً وابسته به موقعیت نقاط مشاهده است. این موضوع می تواند چندضلعی هایی را با شکل هایی کاملاً متفاوت با پدیده مورد نظر است ایجاد کند.



شکل شماره (۸): گرید ایجاد شده در سطح شهر

آموزش و تست شبکه ANFIS و M-ANFIS

به منظور آموزش شبکه ANFIS و M-ANFIS از نقاط ایجاد شده در ناحیه هر ایستگاه استفاده شد. برای این کار در هر مدل داده‌ها به سه دسته داده‌های آموزشی (۷۰٪ داده‌ها)، داده‌های چک (۲۰٪ داده‌ها) و داده‌های تست (۱۰٪ داده‌ها) تقسیم شدند. برای آموزش شبکه از نرم افزار Matlab 7.10 استفاده شد. در شبکه ANFIS آموزش در دو روش پس انتشارخطا و روش ترکیبی پس انتشارخطا و کمترین مربعات استفاده شد. برای ایجاد قانون‌های اولیه از دو روش تقسیم‌بندی گریدی و خوشه‌بندی کاهشی استفاده شد، که در روش اول تعداد توابع عضویت و نوع آن برای مشخصه‌های دما، سرعت باد و جهت باد از روش صحیح و خطا به دست آمد.

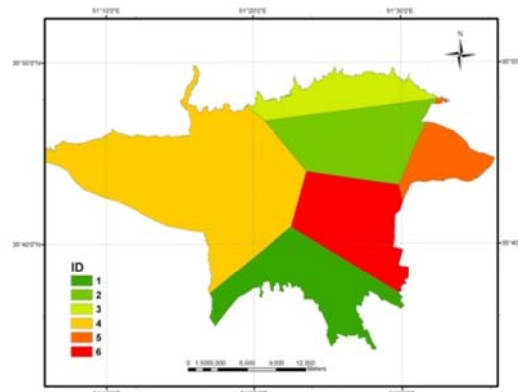
نتیجه نشان داد که با در نظر گرفتن دو تابع عضویت برای مشخصه سرعت باد و سه تابع عضویت برای مشخصه‌های دما و جهت باد و استفاده از تابع عضویت گوسین بهترین حالت ایجاد می‌شود. در خوشه‌بندی فاصله‌ی مؤثر ۰/۴۵ در نظر گرفته شد. برای تولید قانون‌ها از تابع Genfis نرم افزار Matlab استفاده شد. همان طور که ذکر شد قانون‌های فازی سوگو برای به کار بردن در این شبکه ایجاد می‌شود.

در شبکه M-ANFIS از روش پس انتشار خطا برای بهبود قانون‌های فازی استفاده شد. قانون‌های فازی اولیه از روش خوشه‌بندی کاهشی با فاصله تأثیر ۰/۴۵ و روش تقسیم‌بندی فضای ورودی- خروجی ایجاد شد. مدل‌های مختلفی با توابع

در شکل شماره (۷) منطقه پوشش ایستگاههای سنجش آلودگی هوای سازمان حفاظت محیط زیست با مشخصه بهمن (۱)، قلهک (۲)، تجریش (۳)، آزادی (۴)، سرخه حصار (۵) و امام خمینی (۶) نشان داده شده است.

- ایجاد داده‌های آموزشی

در این مرحله با توجه به تعداد و پراکندگی بیشتر ایستگاههای سنجش غلظت آلاینده‌ها نسبت به ایستگاههای هواشناسی، برای تعمیم شبکه در سطح ناحیه، بهبود آموزش و به دست آوردن قانون‌ها برای ناحیه تحت پوشش هر ایستگاه، نسبت به تولید داده‌های آموزشی اقدام شد.



شکل شماره (۷): منطقه پوشش ایستگاهها

در این راستا، نقشه‌های شاخص دما، سرعت باد، جهت باد را با داده‌های ۶ ایستگاه محیط زیست و نقشه شاخص غلظت CO را با داده‌های ۶ ایستگاه سازمان حفاظت محیط زیست و ۱۶ ایستگاه شرکت کنترل کیفیت هوا با روش کریجینگ برای هر تاریخ و ساعت تولید شد. در استفاده از روش کریجینگ، ابتدا می‌باید داده‌ها را با استفاده از روش لاگ نرمال، نرمال کرد.

برای ایجاد داده، یک گرید ۲۰۰۰ متری در منطقه مورد مطالعه لحاظ شد (شکل شماره ۸). اکنون بر اساس مختصات، مقدار دما، سرعت باد، جهت باد و غلظت مونوکسید کربن از نقشه‌های شاخص تولید شده استخراج شد و از آنجایی که دقت مشخصه‌های تولید شده به علت کم بودن تعداد نقاط در کل ناحیه، در بعضی نقاط کم است با کنترل مقدار خطای استاندارد برای هر نقطه، نقاط دارای خطای زیاد حذف شد.

عضویت متفاوت قابل بررسی است که ۳ مدل زیر به عنوان نمونه در نظر گرفته شد:

در یک مدل مشخصه های توابع عضویت مقدم و نتیجه گوسین، در مدل دیگر مشخصه های توابع عضویت مقدم گوسین و مشخصه های نتیجه دوزنقه ای و همچنین مدلی با مشخصه های توابع عضویت مقدم زنگوله ای و مشخصه های نتیجه دوزنقه ای مورد بررسی قرار گرفت. پس از به دست آوردن قانون ها برای افزایش تفسیرپذیری قانون ها از روش کاهش تعداد توابع عضویت فازی با در نظر گرفتن $\lambda = 0.8$ استفاده شد.

تعمیم قانون ها به ناحیه پوشی ایستگاه

قانون های تولید شده در ناحیه هر ایستگاه، به ناحیه تحت پوشش آن ایستگاه اعمال می شود. در این روش ابتدا منطقه به چندضلعی های تیسین تقسیم و سپس با فوری به اندازه نصف قطر پیکسل پیرامون هر ناحیه رسم می شود تا مقدار غلظت CO در محل های وصل تصاویر مقدار مناسبی باشد و تصویر نرم شود. پیکسل های واقع در هر ناحیه وارد مدل فازی همان ناحیه شده و مقدار غلظت مونوکسید کربن به دست می آید و در آخر تصاویر ایجاد شده با هم ادغام می شود.

پیاده سازی مدل دوم

از آنجایی که کلیات پیاده سازی مدل دوم شبیه مدل اول است، از ذکر دوباره چگونگی انجام کار، خودداری می شود. تفاوت مرحله آموزش و تست شبکه ANFIS و M-ANFIS با مدل اول در این است که برای هر ایستگاه از داده های هواشناسی و غلظت آلاینده مونوکسید کربن خود ایستگاهها استفاده می شود.

غلظت CO در نقاط ایستگاهها با استفاده از قانون های ممدانی و سوگنو تولید شده برای فاصله زمانی ۸ ساعته، پیش بینی شده و با اجرای روش کریجینگ بر روی ایستگاهها، نقشه مکانی غلظت CO تهیه شد.

ارزیابی دقت

برای ارزیابی مدل پیشنهادی به دو طریق عمل شد:

- قانون های تولید شده در هر ایستگاه یا در هر ناحیه با داده های تست همان ایستگاه یا ناحیه با روشهای آماری ارزیابی شد. دقت هر ایستگاه به طور جداگانه بررسی شد.
- دقت پیش بینی غلظت آلاینده تمام ایستگاهها به طور همزمان، برای چند روز متفاوت در محل ایستگاههای شرکت کنترل کیفیت هوا که در سطح شهر پراکنده اند

(این نقاط در آموزش وارد نشده اند) با روشهای آماری

بررسی شد. به علت نداشتن داده در نقاط بیشتری، برای

ارزیابی دقت آن به همین نقاط بسنده می کنیم.

در زیر روشهای آماری استفاده شده برای ارزیابی دقت را به طور

مختصر شرح می دهیم.

روشهای آماری ارزیابی دقت

از روشهای آماری می توان به جذر میانگین مربع خطا

(RMSE)¹⁴، (IA)¹⁵، میانگین خطای مطلق (MAE)¹⁶، میانگین

بایاس خطا (MBE)¹⁷، (FB)¹⁸ اشاره کرد. در زیر به توضیح این

روشها می پردازیم:

- روش IA با استفاده از فرمول (۴) به دست می آید:

$$d=1-\frac{\sum_{i=1}^n(P_i - O_i)^2}{\sum_{i=1}^n(|P_i - \bar{O}| + |O_i - \bar{O}|)^2} \quad (4)$$

که در آن P_i مقدار پیش بینی شده و O_i مقدار مشاهده شده و

\bar{O} میانگین مشاهدات است. مقدار آن بین صفر (عدم پذیرش) و یک

(پذیرش کامل سری زمانی) است. (Jain, et al., 2010)

- RMSE اندازه واقعی خطای تولید شده را نشان می دهد و از

فرمول (۵) به دست می آید.

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^n (P_i - O_i)^2} \quad (5)$$

یک مدل خوب باید RMSE متمایل به صفر داشته باشد.

- میانگین خطای بایاس (MBE) درجه تطابق بین میانگین

پیش بینی و میانگین مشاهدات را نشان می دهد. هر چه مقدار آن

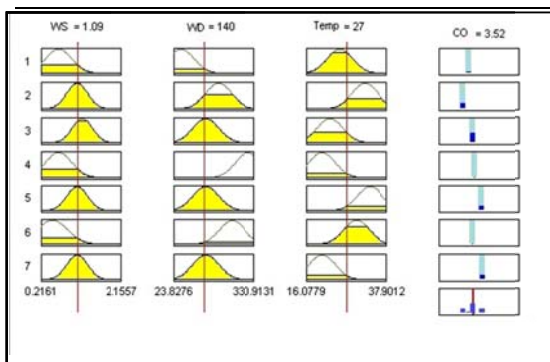
کتر باشد بهتر است. رابطه (۶) آن را بیان می کند.

$$MBE = \frac{1}{N} \sum_{i=1}^n (P_i - O_i) \quad (6)$$

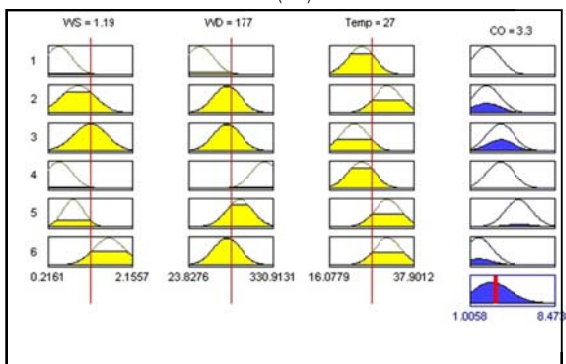
- میانگین خطای مطلق (MAE) به وسیله معادله (۷) محاسبه

می شود.

$$MAE = \frac{1}{N} \sum_{i=1}^n |P_i - O_i| \quad (7)$$



(الف)



(ب)

شکل شماره (۹): الف-قانون های سوگنو. ب- قانون های ممدانی

جدول شماره (۲ و ۳) مقادیر RMSE، IA، FB، MBE و MAE برای حالت های منتخب برای ناحیه هر ایستگاه را به دو روش ANFIS و M-ANFIS برای داده های تست نمایش می دهد.

جدول شماره (۲): دقت روش های انتخاب شده برای ایجاد قانون های سوگنو در ناحیه هر ایستگاه

| ایستگاه | ۱ | ۲ | ۳ | ۴ | ۵ | ۶ |
|------------|------------------------------------|---------|---------|---------|---------|---------|
| روش | Hyb ^a -sub ^b | Hyb-sub | Hyb-sub | Hyb-sub | Hyb-sub | Hyb-sub |
| RMSE (ppm) | ۰/۷۰۵ | ۰/۸۷۰ | ۰/۶۵۹ | ۱/۳۳۸ | ۰/۴۱۲ | ۰/۹۴۳ |
| IA | ۰/۸۹۷ | ۰/۷۹۰ | ۰/۶۸۰ | ۰/۷۱۹ | ۰/۴۶۸ | ۰/۸۵۳ |
| FB | -۰/۰۳۷ | ۰/۰۴۵ | -۰/۰۲۶ | ۰/۰۰۳ | ۰/۰۰۸ | ۰/۰۱۷ |
| MBE (ppm) | ۰/۱۲۲ | -۰/۱۴۱ | ۰/۰۷۹ | -۰/۰۱۰ | -۰/۰۰۸ | -۰/۰۶۳ |
| MAE (ppm) | ۰/۵۲۹ | ۰/۶۷۴ | ۰/۴۷۵ | ۱/۰۳۱ | ۰/۳۱۷ | ۰/۷۸۹ |

Hyb^a: آموزش با استفاده از روش ترکیبی

Sub^b: ایجاد قانون ها با استفاده از روش خوشه بندی کاشی

FB - مقدار کم و زیاد پیش بینی را نشان می دهد. مقدار مثبت FB نشان دهنده مقدار پیش بینی شده کمتر از مشاهده شده است و مقدار منفی آن نشان دهنده این است که مقدار پیش بینی شده بیشتر از مشاهده شده است. FB=0 نشان دهنده وضعیت مطلوب است. رابطه (۸) آن را بیان می کند.

$$FB = \frac{2(\bar{O} - \bar{P})}{(\bar{O} + \bar{P})} \quad (8)$$

که در این رابطه \bar{O} میانگین مشاهدات و \bar{P} میانگین مقادیر پیش بینی شده است.

نتایج

در این قسمت نتایج حاصل از دو مدل پیاده سازی به شرح زیر بیان می شود:

نتایج مدل اول

همان طور که توضیح داده شد قانون های فازی در هر دو سیستم سوگنو و ممدانی با روش شبکه عصبی-فازی ایجاد شد. شکل شماره (۹) توابع عضویت به دست آمده و قانون های ایجاد شده در هر دو سیستم را به عنوان نمونه برای ناحیه ایستگاه قلهدک نشان می دهد، که قانون های ممدانی نشان داده شده با روش خوشه بندی کاشی و توابع عضویت ورودی و خروجی گوسین و قانون های سوگنو با روش خوشه بندی کاشی و توابع عضویت ورودی گوسین و با استفاده از روش آموزش ترکیبی ایجاد شده است.

همان گونه که بیان شد، حالت های مختلفی برای ایجاد قانون های اولیه فازی و آموزش شبکه بررسی شد. برای هر ناحیه مدلی که بهترین جواب را می دهد، انتخاب شد.

جدول شماره (۳): دقت روشهای انتخاب شده برای ایجاد قانون های ممدانی در ناحیه هر ایستگاه

| ایستگاه | ۱ | ۲ | ۳ | ۴ | ۵ | ۶ |
|------------|--|-------------------------------|---------------|---------------|---------------|-----------------|
| روش | Sub ^a -bell ^b -trap ^c | Sub-gauss ^d -gauss | Sub-bell-trap | Sub-bell-trap | Sub-bell-trap | sub-gauss-gauss |
| RMSE (ppm) | ۱/۱۴۵ | ۰/۹۳۲ | ۰/۷۳۵ | ۱/۴۸۲ | ۰/۴۰۹ | ۱/۰۹۴ |
| IA | ۰/۷۴۱ | ۰/۶۸۸ | ۰/۵۷۶ | ۰/۶۰۵ | ۰/۵۳۱ | ۰/۵۵۸ |
| FB | -۰/۰۷۷ | ۰/۰۰۴ | -۰/۰۸۲ | ۰/۰۰۱ | -۰/۰۲۶ | -۰/۰۰۵ |
| MBE (ppm) | ۰/۲۶۷ | -۰/۰۱۳ | ۰/۲۵۱ | -۰/۰۰۲ | ۰/۰۲۹ | ۰/۰۱۸ |
| MAE (ppm) | ۰/۸۱۱ | ۰/۷۹۹ | ۰/۵۲۱ | ۰/۹۷۷ | ۰/۲۹۹ | ۰/۸۷۲ |

Sub^a: ایجاد قانون ها با استفاده از روش خوشه بندی کاهشی Bell^b: تابع عضویت زنگوله ای Trap^c: تابع عضویت دوزنقه ای Gauss^d: تابع عضویت گوسین

از داده های هواشناسی ۶ ایستگاه سازمان حفاظت محیط زیست کریچینگ انجام شد و مقادیر دما، سرعت باد و جهت باد برای ایستگاههای سازمان کنترل کیفیت هوا استخراج شد و این مقادیر با توجه به مکان ایستگاهها و ناحیه ای که در آن قرار دارند، وارد قانون های فازی ناحیه مربوطه شده و مقدار غلظت مونوکسید کربن پیش بینی می شود و این مقدار پیش بینی شده با مقدار واقعی غلظت مونوکسید کربن در آن نقطه از طریق روشهای ارزیابی دقت ذکر شده، مقایسه می شود. جدول شماره (۴) این مقادیر را برای روش ANFIS و جدول شماره (۵) این مقادیر را برای روش M_ANFIS نشان می دهد.

جدول شماره (۴) ارزیابی دقت قانون های به دست آمده از روش ANFIS برای ناحیه هر ایستگاه (واحد اعداد ppm است)

| ایستگاه | ۱ | ۲ | ۳ | ۴ | ۵ | ۶ |
|---------|--------|--------|--------|-------|--------|-------|
| RMSE | ۱/۸۵۴ | ۲/۱۲۰ | -۰/۹۱۲ | ۱/۷۹۲ | ۱/۲۴۸ | ۱/۷۵۵ |
| MBE | -۰/۵۹۲ | -۰/۰۹۴ | -۰/۴۵۹ | ۱/۱۶۶ | -۱/۰۷۰ | ۱/۲۸۹ |
| MAE | ۱/۴۶۲ | ۱/۴۶۹ | -۰/۸۳۲ | ۱/۴۳۷ | ۱/۰۷۰ | ۱/۳۱۴ |

جدول شماره (۵): ارزیابی دقت قانون های به دست آمده از روش M-ANFIS برای ناحیه هر ایستگاه

روش M-ANFIS برای ناحیه هر ایستگاه

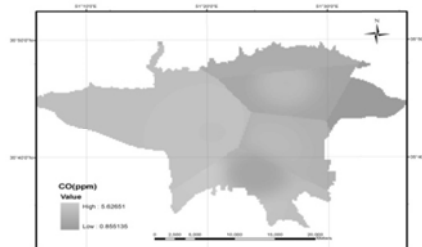
(واحد اعداد ppm است)

| ایستگاه | ۱ | ۲ | ۳ | ۴ | ۵ | ۶ |
|---------|--------|--------|--------|-------|--------|-------|
| RMSE | ۱/۸۴۴ | ۱/۹۵۳ | ۱/۰۹۹ | ۱/۵۱۸ | ۰/۶۹۴ | ۱/۶۷۹ |
| MBE | -۰/۶۵۰ | -۰/۵۰۵ | -۰/۸۷۰ | ۱/۱۱۷ | -۰/۵۲۱ | ۱/۴۰۱ |
| MAE | ۱/۴۴۶ | ۱/۵۰۱ | -۰/۹۶۶ | ۱/۲۲۹ | ۰/۶۱۳ | ۱/۴۰۲ |

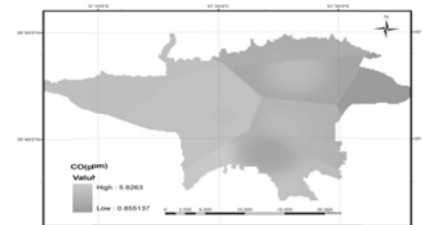
نتایج مدل دوم

قانون های فازی در هر دو سیستم سوگنو و ممدانی ایجاد شد. شکل شماره (۱۱) توابع عضویت و قانون های ایجاد شده در هر دو سیستم را به عنوان نمونه برای ایستگاه قلهک نشان می دهد، که قانون های ممدانی نشان داده شده با روش خوشه بندی کاهشی و توابع عضویت ورودی زنگوله ای و خروجی دوزنقه و قانون های

برای نمونه در جدول شماره (۳ و ۴) مشاهده می شود که برای ایستگاه بهمین (۱) در سیستم ANFIS، روش ایجاد قانون های اولیه با روش خوشه بندی کاهشی و آموزش شبکه به روش ترکیبی و در سیستم M-ANFIS، ایجاد قانون های اولیه به روش خوشه بندی کاهشی و توابع عضویت متغیرهای ورودی زنگوله ای و متغیر خروجی دوزنقه با آموزش شبکه به روش پس انتشار خطا انتخاب شد. برای نمونه تاریخ ۱۵ تیر ۱۳۸۸ در نظر گرفته شد. شکل شماره (۱۰) نقشه به دست آمده از قانون های فازی ممدانی و سوگنو تولید شده را برای بازه زمانی ساعت ۸ تا ۱۶ را نمایش می دهد.



الف

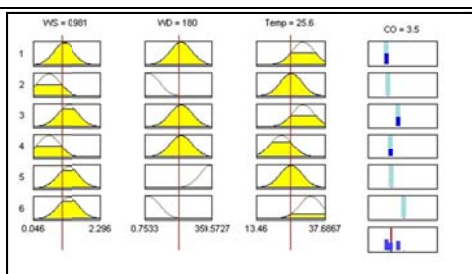


ب

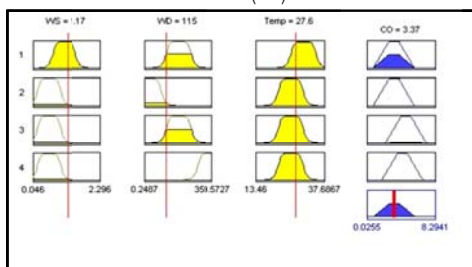
شکل شماره (۱۰): مدل سازی انجام شده الف: استفاده از

قانون های سوگنو. ب: استفاده از قانون های ممدانی

برای ارزیابی دقت مدل پیاده سازی شده، از داده های ایستگاههای شرکت کنترل کیفیت هوا برای ۱۰ روز که به طور تصادفی انتخاب شده است، استفاده شد. بدین صورت که با استفاده



(الف)



(ب)

سوگنو با روش خوشه بندی کاهشی و توابع عضویت ورودی گوسین و با استفاده از روش آموزش پس انتشارخطا ایجاد شده است. حالت های مختلف ایجاد قانون های اولیه فازی و آموزش شبکه بررسی شد. برای هر ایستگاه مدلی که بهترین جواب را می دهد انتخاب شد. جداول شماره (۶ و ۷) مقادیر RMSE, IA, MBE, FB و MAE برای حالت های انتخاب شده برای هر ایستگاه را به دو روش ANFIS و M-ANFIS برای داده های تست نمایش می دهد. برای نمونه همان طور که در جدول شماره (۶ و ۷) مشاهده می شود برای ایستگاه بهمن (۱) در سیستم ANFIS، روش ایجاد قانون های اولیه با روش خوشه بندی کاهشی و آموزش شبکه به روش ترکیبی و در سیستم M-ANFIS، ایجاد قانون های اولیه با استفاده از روش تقسیم بندی فضای ورودی-خروجی و توابع عضویت متغیرهای ورودی زنگوله ای و متغیر خروجی دوزنقه با آموزش شبکه به روش پس انتشار خطا انتخاب شد.

شکل شماره (۱۱): الف: قانون های سوگنو. ب: قانون های

ممدانی

جدول شماره (۶): دقت روش های انتخاب شده برای ایجاد قانون های سوگنو هر ایستگاه

| ایستگاه | ۱ | ۲ | ۳ | ۴ | ۵ | ۶ |
|------------|------------------------------------|----------------------|------------------------|--------|-----------|---------|
| روش | Hyb ^a -sub ^b | BP ^c -sub | Hyb-gridp ^d | Bp-sub | Hyb-gridp | Hyb-sub |
| RMSE (ppm) | ۰/۸۳۱ | ۰/۹۷۴ | ۰/۹۰۹ | ۱/۵۶۴ | ۰/۳۷۲ | ۰/۸۱۹ |
| IA | ۰/۸۳۲ | ۰/۵۹۳ | ۰/۶۹۴ | ۰/۵۹۷ | ۰/۵۳۵ | ۰/۸۴۷ |
| FB | -۰/۰۰۲ | ۰/۰۲۳ | ۰/۰۱۱ | -۰/۰۲۸ | -۰/۰۳۳ | -۰/۰۱۴ |
| MBE (ppm) | ۰/۰۰۷ | -۰/۰۸۲ | -۰/۰۳۳ | ۰/۱۲۵ | ۰/۰۳۴ | ۰/۰۵۶ |
| MAE (ppm) | ۰/۶۱۵ | ۰/۷۸۲ | ۰/۷۱۲ | ۱/۲۳۲ | ۰/۲۸۴ | ۰/۶۳۹ |

Hyb^a: آموزش با استفاده از روش ترکیبی
 BP^c: آموزش با استفاده از روش پس انتشارخطا
 Sub^b: ایجاد قانون ها با استفاده از روش خوشه بندی کاهشی
 gridp^d: ایجاد قانون ها با استفاده از روش تقسیم بندی گریدی

جدول شماره (۷): دقت روش های انتخاب شده برای ایجاد قانون های ممدانی هر ایستگاه

| ایستگاه | ۱ | ۲ | ۳ | ۴ | ۵ | ۶ |
|------------|---|-----------------------------|--|----------------|-----------------|-----------------|
| روش | Mandel ^a -bell ^b -trap ^c | Sub ^d -bell-trap | Mandel ^a -gauss ^e -gauss | Sub-gauss-trap | Sub-gauss-gauss | Sub-gauss-gauss |
| RMSE (ppm) | ۱/۰۴۸ | ۱/۲۵۴ | ۰/۹۵۳ | ۱/۷۱۹ | ۰/۴۱۷ | ۱/۰۰۴ |
| IA | ۰/۶۰۶ | ۰/۴۱۴ | ۰/۵۰۲ | ۰/۴۸۸ | ۰/۴۸۳ | ۰/۶۰۸ |
| FB | -۰/۰۴۵ | -۰/۱۰۸ | -۰/۰۳۴ | -۰/۰۳۶ | -۰/۰۲۹ | ۰/۰۱۷ |
| MBE (ppm) | ۰/۱۳۸ | ۰/۳۸۱ | ۰/۱۱۱ | ۰/۱۶۱ | ۰/۰۳۱ | -۰/۰۶۷ |
| MAE (ppm) | ۰/۷۹۰ | ۰/۹۵۹ | ۰/۷۷۵ | ۱/۳۹۴ | ۰/۳۰۳ | ۰/۷۸۳ |

Mandel^a: ایجاد قانون ها با استفاده از روش تقسیم بندی فضای ورودی و خروجی
 Trap^c: تابع عضویت دوزنقه ای
 Sub^d: ایجاد قانون ها با استفاده از روش خوشه بندی کاهشی
 Gauss^e: تابع عضویت گوسین
 Bell^b: تابع عضویت زنگوله ای

اکنون برای نمونه تاریخ ۱۵ تیر ۱۳۸۸ در نظر گرفته شد. غلظت CO در نقاط ایستگاهها با استفاده از قانون های ممدانی و فازی تولید شده برای فاصله زمانی ۸ ساعته پیش بینی کرده و سپس کریجینگ را انجام داده و بدین ترتیب نقشه مکانی غلظت CO به دست می آید (شکل شماره ۱۲). برای ارزیابی دقت مدل پیاده سازی شده از داده های ایستگاههای سازمان کنترل کیفیت هوا برای ۱۰ روز که به طور تصادفی انتخاب شده است، استفاده شد.

جدول شماره (۹): ارزیابی دقت قانون های به دست آمده از روش M-ANFIS برای هر ایستگاه (واحد اعداد ppm است)

| ایستگاه | ۱ | ۲ | ۳ | ۴ | ۵ | ۶ |
|---------|--------|--------|-------|-------|-------|-------|
| RMSE | ۱/۷۹۱ | ۱/۹۲۵ | ۰/۸۷۵ | ۱/۱۷۰ | ۱/۴۳۳ | ۱/۰۵۶ |
| MBE | -۰/۲۰۵ | -۰/۲۸۲ | ۰/۴۶۹ | ۰/۵۹۹ | ۱/۳۳۶ | ۰/۶۵۰ |
| MAE | ۱/۴۹۵ | ۱/۵۱۲ | ۰/۶۷۵ | ۰/۹۷۴ | ۱/۳۳۶ | ۰/۹۱۷ |

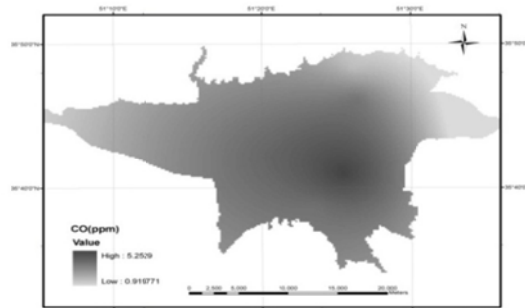
بحث نتایج

همان گونه که ذکر شد، در این تحقیق دو مدل عملی شد. در مدل اول برای ایجاد داده های آموزشی از کریجینگ استفاده شد. با توجه به کم بودن تعداد نقاطی (حداکثر ۷ ایستگاه) که مشخصه های معیار را پایش می کردند، دقت داده های تولید شده در نواحی دور از ایستگاه های پایش، پایین بودند، که این داده ها با توجه به شرط خطای استاندارد، حذف شدند؛ بنابراین داده های آموزشی ورودی سیستم تمام سطح این ناحیه را پوشش نمی دهند. پس قانون های ایجاد شده در این ناحیه به طور کامل قادر به شناسایی رفتار حاکم بر محیط نیست.

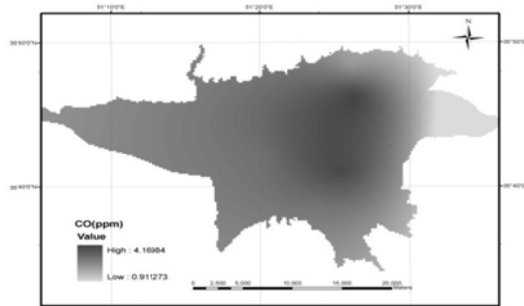
می توان برای بیان بهتر این مطلب به نقشه غلظت مونوکسید کربن ایجاد شده مراجعه کرد. در ناحیه ایستگاه آزادی غلظت مونوکسید کربن تقریباً یکنواخت بود که این به دلیل گستردگی ناحیه ایستگاه و نبود ایستگاه در نواحی آن است. همچنین در نقشه مشاهده می شود که گستردگی در محل های اتصال ناحیه ها وجود دارد، که علت آن را نیز می توان به تعداد کم ایستگاه های پایش و گستردگی ناحیه ها نسبت داد.

در این مدل میانگین RMSE کلیه ایستگاهها در روش ANFIS، ۱/۶۱۳ ppm و در روش M-ANFIS، ۱/۴۸۴ ppm به دست آمد. در مدل دوم چون از داده های موجود در هر ایستگاه برای آموزش آن ایستگاه استفاده کردیم، در صورت اطمینان به داده ها می توان گفت که قادر است با دقت مناسب غلظت را در نقاط ایستگاه پیش بینی کند ولی چون مدل سازی با استفاده از کریجینگ صورت گرفت، در این مدل نیز با مشکل تعداد نقاط کم (۶ ایستگاه) برای درون یابی مواجه هستیم.

در این مدل میانگین RMSE کلیه ایستگاهها در روش ANFIS، ۱/۴۴۵ ppm و با روش M-ANFIS، ۱/۳۷۴ ppm به دست آمد.



(الف)



(ب)

شکل شماره (۱۲): مدل سازی با الف: قانون های سوگنو - ب:

قانون های ممدانی

بدین صورت که از با استفاده از داده های هواشناسی ۶ ایستگاه سازمان حفاظت محیط زیست، مقدار غلظت مونوکسید کربن را در این ایستگاهها با استفاده از قانون های فازی ایجاد شده مربوطه پیش بینی کرده و سپس با استفاده از روش کریجینگ مقدار غلظت مونوکسید کربن را برای ایستگاه های شرکت کنترل کیفیت هوا به دست می آوریم. این مقدار پیش بینی شده با مقدار واقعی غلظت مونوکسید کربن در آن نقطه، از طریق روشهای ارزیابی دقت ذکر شده، مقایسه می شود. جدول شماره (۸) این مقادیر را برای روش ANFIS و جدول شماره (۹) این مقادیر را برای روش M-ANFIS نشان می دهد.

جدول شماره (۸): ارزیابی دقت قانون های به دست آمده از روش ANFIS برای هر ایستگاه (واحد اعداد ppm است)

| ایستگاه | ۱ | ۲ | ۳ | ۴ | ۵ | ۶ |
|---------|--------|--------|-------|-------|-------|-------|
| RMSE | ۱/۸۴۵ | ۱/۷۷۴ | ۱/۱۰۲ | ۱/۱۲۶ | ۱/۵۸۳ | ۱/۲۴۳ |
| MBE | -۰/۰۶۸ | -۰/۰۰۸ | ۰/۴۲۷ | ۰/۶۸۲ | ۱/۳۸۲ | ۰/۸۰۰ |
| MAE | ۱/۵۰۵ | ۱/۲۹۶ | ۰/۷۵۴ | ۰/۹۱۷ | ۱/۳۸۲ | ۰/۹۵۷ |

نتیجه گیری

پیش‌بینی آلودگی هوا یکی از راهکارهای مدیریتی برای جلوگیری، و یا کاهش پیامدهای مخرب آن است. در این تحقیق با بهره‌گیری از شبکه عصبی-فازی و GIS، قانون‌های ممدانی و سوگنو حاکم بر ایستگاههای پایش و ناحیه پوشش آنها، برای پیش‌بینی ۸ ساعته غلظت مونوکسید کربن با استفاده از مشخصه‌های هواشناسی به دست آمد و نقشه پیش‌بینی غلظت مونوکسید کربن مدل‌سازی شد.

دو مدل برای پیش‌بینی و مدل‌سازی مکانی غلظت آلاینده مونوکسید کربن بررسی شد.

در مدل اول قانون‌های فازی (ممدانی و سوگنو) برای ناحیه تحت پوشش هر ایستگاه با استفاده از داده‌های ایجاد شده با روش زمین آمار (کریجینگ) استخراج شد. با توجه به نتایج می‌توان استنباط کرد که روش ایجاد داده آموزشی به منظور استخراج قانون‌های حاکم بر ناحیه تحت پوشش ایستگاه، برای نواحی‌ای که دارای گستردگی زیاد و به دور از ایستگاههای پایش هستند، کارا نیست. با بررسی نتایج ارزیابی آماری دقت به دست آمده با استفاده از روش ANFIS و M-ANFIS مشاهده می‌شود که هر دو روش تقریباً دارای دقت یکسان و مطلوبی است.

ولی از آنجایی که بخش تالی قانون‌های سوگنو یک فرمول ریاضی بوده و چارچوبی برای نمایش دانش بشری فراهم نمی‌کند، بنابراین می‌توان قانون‌های ممدانی استخراج شده به عنوان قانون‌های دارای کارایی بیشتر، معرفی کرد.

همچنین این روش برای تهیه نقشه پیش‌بینی آلاینده‌هایی که مقدار آن وابسته به مشخصه‌های هواشناسی بوده ولی در نقاطی که این مشخصه‌ها پایش می‌شوند، اندازه‌گیری نمی‌شوند، روشی مناسب است. به طور کلی می‌توان از این روش در مواقعی که با داده‌های کم و پراکنده روبرو هستیم، برای به دست آوردن قانون‌های فازی ممدانی و سوگنو مناسب بر پدیده مکانی با استفاده از داده‌ها (بدون دانش کارشناسی) بهره بگیریم.

در مدل دوم از شبکه‌های ANFIS و M-ANFIS به منظور ایجاد قانون‌های مناسب برای هر ایستگاه پایش استفاده شد. از آنجایی که مدل‌سازی با استفاده از کریجینگ صورت گرفت در حالتی که تعداد نقاط پایش بیشتر باشد، با استفاده از توانایی‌های سیستم فازی-عصبی و کریجینگ قادر به مدل‌سازی مناسب غلظت آلاینده هستیم. با بررسی نتایج ارزیابی آماری دقت به دست آمده با

استفاده از روش ANFIS و M-ANFIS مشاهده می‌شود که هر دو روش تقریباً دارای دقت یکسان و مطلوبی است. مانند مدل قبل قانون‌های استخراج شده از روش M-ANFIS، به علت تفسیرپذیر و قابل درک بودن آن و دقت مطلوبش، مناسب برای مدل‌سازی رفتار سیستم معرفی می‌شود.

مدل اول با هدف به دست آوردن قانون‌ها برای ناحیه توسعه داده شد. در این مدل گسستگی در نقشه غلظت آلاینده مشاهده می‌شود که این به دلیل کمبود ایستگاههای پایش و گستردگی ناحیه‌هاست، ولی توانست با وجود کمبود ایستگاههای پایش هدف اصلی این تحقیق که اکتشاف دانش است، برای ناحیه مرتفع سازد. در مدل دوم که قانون‌های هر ایستگاه به دست آمد، به این علت که مدل‌سازی توسط کریجینگ صورت گرفته است، نقشه غلظت آلاینده پیوسته است.

با توجه به مقادیر RMSE و همچنین نقشه غلظت پیش‌بینی شده و همچنین به دلیل این که تعداد ایستگاههای پایش کم هستند، مدل دوم مدلی مناسب‌تر از نظر مدل‌سازی است.

یادداشت‌ها

- 1- Total Suspended Particles
- 2- Temperature
- 3- Wind Speed
- 4- Wind Direction
- 5- Humidity
- 6- Pressure
- 7- Adaptive Neuro Fuzzy Inference System
- 8- Firing Strength
- 9- Least Square Estimation (LSE)
- 10- Mamdani- Adaptive Neuro Fuzzy Inference System
- 11- Implication
- 12- Subtractive Clustering
- 13- Grid Partitioning
- 14- Root Mean Square Error(RMSE)
- 15- Index of Agreement(IA)
- 16- Mean Absolute Error(MAE)
- 17- Mean Bias Error(MBE)
- 18- Fractional Bias(FB)

تشکر و قدردانی

از سازمان حفاظت محیط زیست استان تهران و شرکت کنترل کیفیت هوا وابسته به شهرداری تهران، که داده‌های مورد نیاز برای این تحقیق را در اختیار ما قرار دادند، متشکریم.

منابع مورد استفاده

- افیونی، م، عرفان منش، م. ۱۳۸۸. آلودگی محیط زیست: آب، خاک و هوا. انتشارات ارکان دانش. ۹۷۸-۹۶۴-۲۵۹۱-۸۶-۲.
- تشنه لب، م. ۱۳۸۸. سیستم‌های فازی و کنترل فازی. دانشگاه خواجه نصیرالدین طوسی.
- حسنی پاک، ع. ۱۳۷۷. زمین آمار (ژئواستاتستیک). مؤسسه انتشارات و چاپ دانشگاه تهران.
- صفوی، ی، علیجانی، ب. ۱۳۸۵. بررسی عوامل جغرافیایی در آلودگی هوای تهران. پژوهش‌های جغرافیایی، جلد ۵۸، ص. ۱۱۲ تا ۹۹.
- Brauer, M., et al. 2003. Estimating long-term average particulate air pollution concentrations: Application of traffic indicators and geographic information systems. *Epidemiology*. Vol. 14. pp. 228-239.
- Briggs, D.J., et al. 2000. A regression-based method for mapping traffic-related air pollution: Application and testing in four contrasting urban environments. *Sci Total Environ*. Vol. 253. pp. 151-167.
- Chai, Y., L., Jia, Z., Zhang. 2009. Mamdani Model based Adaptive Neural Fuzzy Inference System and its Application. *International Journal of Information and Mathematical Sciences* 5:1.
- Chen, M.Y., D.A., Linkens. 2004. Rule-base self-generation and simplification for data-driven. *Fuzzy Sets and Systems*, Vol. 142, pp. 243 - 265.
- Chiu. 1994. fuzzy model identification based on cluster estimation. *Journal of Intelligent & Fuzzy Systems*, Vol. 2, pp. 267-278.
- Elbir, T., et al. 2010. Development of a GIS-based decision support system for urban air quality management in the city of Istanbul. 4, *Atmospheric Environment*, Vol. 44, pp. 441-454.
- Fuller, R. 1995. *Neural Fuzzy Systems*.
- Goovaerts, P., A., Auchincloss, A.V., Diez-Roux. 2006. Performance comparison of spatial and space-time interpolation techniques for prediction of air pollutant concentrations in the Los Angeles area. *Society for Mathematical Geology XIth International Congress*, S13-11.
- Hammouda, K. 2000. A comparative study of data clustering techniques. *Citeseer*, Vol. 625.
- Heo, J.S., D.S., Kim. 2004. A new method of ozone forecasting using fuzzy expert and neural network systems. *Science of the Total Environment*, Vol. 325, pp. 221-237.
- Horikawa, S., T., Furuhashi, Y., Uchikawa. 1992. On fuzzy modelling using fuzzy neural networks with the back-propagation algorithm. 5, *IEEE Trans Neural Netw*, Vol. 3, pp. 801-806. doi:10.1109/72.159069.
- Hua, Y.C., R.S., Chena, T G., H.zengb. 2003. Discovering fuzzy association rules using fuzzy partition methods. *Knowledge-Based Systems*, Vol. 16, pp. 137-147.
- Jang, J.S.R. 1993. ANFIS: Adaptive-Neural-Network-Based Fuzzy Inference Systems. *IEEE Transaction on Systems, Man, and Cybernetics* 23, pp. 665-685

Jain,S. , M.,Khare .2010. Adaptive neuro-fuzzy modeling for prediction of ambient CO concentration at urban intersections and roadways. *Air Qual Atmos Health*, Vol. 3, pp. 203–212.

Janssen,S., et al .2008. Spatial interpolation of air pollution measurements using CORINE land cover data. *Atmos Environ*, Vol. 42, pp. 4884–4903.

Jang,J.S.R. 1996. Input selection for ANFIS learning. *IEEE*, pp. 1493-1499. 1098-7584, 978-1-4244-3596-8.

Jerrett,M., et al .2005 .A review and evaluation of intraurban air pollution exposure models. *J Expos Anal Environ Epidemiol*.Vol. 15. pp. 185–204.

Lin, C.T., C.S.G.,Lee .1991. Neural-network-based fuzzy logic control and decision system. 12, *IEEE Trans Comput C*, Vol. 40, pp. 1320–1336.

Ishibuchi,H., K.,Kwon, H.,Tanaka .1995. A learning algorithm of fuzzy neural networks with triangular fuzzy weights. *s.l. : Fuzzy Sets Syst*, Vol. 71, pp. 277–293. doi:10.1016/0165-0114(94)00281-B.

Morabito,F., M.,Versaci .2003. Fuzzy neural identification and forecasting techniques to process experimental urban air pollution data. *Neural Networks*, Vol. 16, pp. 493–506.

Smith,L. , et al .2006 .Use of GIS and ancillary variables to predict volatile organic compound and nitrogen dioxide levels at unmonitored locations. 40, *Atmospheric Environment*, pp. 3773–3787.

Setnes,M., et al .1998. Similarity Measures in Fuzzy Rule Base Simplification. 3, *IEEE Transactions on systems, man, and cybernetics part b cybernetics*,vol. 28.

Vieira,J. , M.,Dias, F.A.,Mota .2004. Neuro-Fuzzy Systems: A Survey. 5th WSEAS NNA International Conference on Neural Networks and Applications.

Wang, L.X.; J.M.,Mendel .1992. Generating fuzzy rules by learning from examples. 6, *IEEE Systems, Man, and Cybernetics Society*, Vol. 22. 0018-9472.

Wenjun,C. 2010. Application of GIS technology in the emergency monitoring of sudden air pollution accident. *IEEE. 2nd International Conference Information Science and Engineering (ICISE)*, pp. 3550-3555.

Wu,Y., et al .2011. Fuzzy Logic and Neuro-fuzzy Systems: A Systematic Introduction. 2, *International Journal of Artificial Intelligence and Expert Systems (IJAE)*, Vol. 2.

Xiwen,W.2010.The Research of Urban air pollution forecast Base on GIS Technology.*IEEE. 3rd International Conference on Advanced Computer Theory and Engineering(ICACTE)*. Vol. 4, pp. 200-202

Yildirim,Y. , M.,Bayramoglu .2006. Adaptive neuro-fuzzy based modelling for prediction of air pollution daily levels in city of Zonguldak. *Chemosphere*, Vol. 63, pp. 1575–1582.