

استفاده از شبکه بازگشتی NAR برای پیش بینی غلظت مونوکسید کربن

مهرداد رفیع پور^{۱*}

mehrdadrafiepour@gmail.com

علی اصغر آل شیخ^۲

عباس علیمحمدی^۳

ابوالقاسم صادقی نیارکی^۴

تاریخ دریافت: ۹۲/۵/۸

تاریخ پذیرش: ۹۲/۱۰/۱۸

چکیده

زمینه و هدف: آلودگی هوا یکی از مشکلات مهم شهرهای بزرگ محسوب می‌شود. یکی از اهداف مسئولین شهری آگاهی از میزان کیفیت هوا در آینده است؛ برای پیشبینی کیفیت هوا، باید غلظت هریک از آلاینده‌ها مدلسازی شده و با استفاده از مدل ایجاد شده، نسبت به پیشبینی مقادیر هریک از آلاینده‌ها اقدام شود. با توجه به اینکه مونوکسید کربن یکی از آلاینده‌های مهم هوا است، و تاثیرات زیانباری بر سلامت انسان دارد.

روش بررسی: در این مقاله به مدلسازی و پیشبینی ۲۴ ساعته غلظت مونوکسید کربن با استفاده از شبکه‌های عصبی بازگشتی NAR و مدل آماری ARMA پرداخته شده و سپس نتایج این دو روش با یکدیگر مقایسه شده است. برای این منظور از داده‌های سال ۲۰۰۹ از ۲۹ نوامبر تا ۳۱ دسامبر، مربوط به غلظت آلاینده مونوکسید کربن اندازه‌گیری شده در ایستگاه آزادی از ایستگاه‌های پایش کیفیت هوا متعلق به سازمان محیط زیست استان تهران استفاده شده است.

یافته‌ها: نتایج مدلسازی نشان می‌دهد که شبکه عصبی NAR دارای دقت بهتری نسبت به روش ARMA برای پیشبینی و مدلسازی غلظت مونوکسید کربن است. شبکه عصبی NAR با MSE کمتر از ۱/۶ دقت مناسبی در پیشبینی غلظت آلاینده مونوکسید کربن داشت. همچنین همبستگی بین مقادیر پیشبینی شده و مقادیر واقعی برای شبکه عصبی NAR، ۸۴ درصد می‌باشد. در حالی که مدل ARMA دارای MSE برابر ۵/۴۶ و ضریب همبستگی ۷۲ درصد می‌باشد.

۱- دانشجوی کارشناسی ارشد رشته سیستم اطلاعات مکانی دانشگاه صنعتی خواجه نصیرالدین طوسی* (مسوول مکاتبات).

۲- استاد رشته سیستم اطلاعات مکانی دانشگاه صنعتی خواجه نصیرالدین طوسی

۳- دانشیار رشته سیستم اطلاعات مکانی دانشگاه صنعتی خواجه نصیرالدین طوسی

۴- استادیار رشته سیستم اطلاعات مکانی دانشگاه صنعتی خواجه نصیرالدین طوسی

نتیجه‌گیری: می‌توان نتایج پیشبینی را جهت آگاه‌سازی عمومی در اینترنت و شبکه‌های جمعی منتشر کرد. همچنین نتایج مدل‌سازی و پیشبینی می‌تواند برای مدیریت بهتر آلودگی هوا توسط مدیران مورد استفاده قرار گیرد. نتایج این تحقیق نشان می‌دهد که شبکه عصبی NAR قابلیت بسیار بالایی در پیشبینی سری زمانی غاظت مونوکسیدکربن دارد.

واژه‌های کلیدی: آلودگی هوا، مونوکسیدکربن، شبکه‌های عصبی، مدل‌سازی، پیشبینی.

Using dynamic recurrent neural network NAR for predicting monoxide carbon concentration

Mehrdad Rafiepour^{1*}

mehrdadrafiepour@gmail.com

Ali Asghar Alesheikh²

Abbas Alimohammad³

Abolghasem Sadeghi Niaraki⁴

Abstract

Background and Objective: Air pollution is one of the most important problems in big cities. One of the goals of urban managers is their awareness on air pollution in the future. For prediction of air quality, air pollutant must be modeled first. Carbon monoxide is one of the most toxic air pollutants that has harmful effect on human health.

Method: In this paper, modeling carbon monoxide concentration and 24-h prediction by ARMA and NAR neural network have been studied. Then, the results of the two methods are compared. For this purpose, data is collected on 29 November until 31 December 2009 in Azadi air quality monitoring station: belonged to Tehran department of environment.

Findings: The results of the two methods showed that, NAR is more accurate than ARMA for modeling and prediction of carbon monoxide. NAR neural network had MSE=1.6 and a correlation coefficient of 0.84 while ARMA had MSE=5.46 and correlation coefficient=0.72 for 24 hours prediction.

Discussion and Conclusion: Finally, the predicted values can be used and published in internet for public awareness. Also urban managers can use the results of modeling and prediction for a better management. Result of this paper showed NAR neural network has sufficient ability to model and predict time series of monoxide carbon.

Key words: air pollution, monoxide carbon, neural network, modeling and prediction.

1- Student of GIS, K.N.Toosi University of Technology*(Corresponding Author).

2- Full Professor, K.N.Toosi University of Technology

3- Associate Professor, K.N.Toosi University of Technology

4- Assistant Professor, K.N.Toosi University of Technology

مقدمه

آلاینده‌های مهم هوا که می‌تواند تاثیرات بدی بر سلامت انسان داشته باشد منوکسیدکربن است. منوکسیدکربن گازی بی رنگ و بی بو ولی بسیار سمی است. منبع اصلی تولید این گاز اتومبیل‌ها هستند. گاز منوکسید کربن 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 (۲). منبع اصلی تولید این گاز اتومبیل‌ها هستند (۶). استنشاق این گاز در محیط‌های در بسته و سقف دار باعث خفگی و مرگ می‌شود. مواد موجود در خاک می‌توانند مقداری از CO را جذب نموده و مقداری نیز در اثر فعل و انفعالات به دی اکسید کربن تبدیل می‌گردد. به هر حال در صورتی که میزان منوکسیکربن از ۷۵۰ ppm تجاوز کند سبب مرگ خواهد شد (۷). بنابراین مدل‌سازی غلظت منوکسید کربن و پیشبینی با استفاده از روش مناسب ضروری به نظر میرسد.

معمولا غلظت این آلاینده توسط سنجنده‌های پایش کیفیت هوا بصورت ساعتی ثبت و جمع‌آوری می‌شود؛ این داده‌ها در کنار یکدیگر یک سری زمانی را تشکیل می‌دهند. تا قبل از دهه ۱۹۸۰ سری‌های زمانی به کمک معادلات پارامتریک (Auto Regressive) AR، (Auto Regressive Moving-Average) ARMA یا (Average) پیشبینی می‌شدند. این معادلات خطی هستند و قادر به پیشبینی صحیح فرآیندهای غیر خطی نیستند (۸). از شبکه‌های عصبی برای پیشبینی سری‌های زمانی علی‌الخصوص در زمانی که شرایط برای استفاده از تکنیک‌های کلاسیک فراهم نیست و سری‌های زمانی پیچیده می‌باشند بسیار استفاده می‌شود (۹). از طرفی استفاده از شبکه‌های عصبی به خصوص شبکه‌های عصبی بازگشتی و پویا می‌تواند در مدل‌سازی و پیشبینی سری زمانی آلودگی هوا مفید باشد (۷)؛ زیرا قابلیت خوبی در پیشبینی سری‌های زمانی پیچیده با ماهیت غیرخطی به خصوص سری زمانی مربوط به غلظت آلاینده‌های هوا را دارند (۱۰).

آلودگی هوا یکی از مسائل مهم در کشورهای توسعه یافته و در حال توسعه است. آلودگی هوا به معنای مخلوط شدن هوا با گازها، قطرات و ذراتی است که کیفیت هوا را کاهش می‌دهند. هوای آلوده، پدیده ایست که از ترکیب یا اختلاط هوا با مواد یا ذرات خاصی، در مدت زمان معینی تولید می‌شود و در صورت تداوم، بیماری‌ها و اختلالاتی برای انسان، حیوانات و گیاهان ایجاد می‌کند و به میزان قابل ملاحظه ای زندگی بشر را دچار مخاطره می‌کند (۱). آلودگی هوا از منابع مختلف سرچشمه می‌گیرد. گسترش شهرنشینی، توسعه شهرها، افزایش جمعیت، توسعه فعالیت‌های صنعتی، افزایش مصرف سوخت‌های فسیلی، کمبود سیستم حمل و نقل عمومی کارا، کیفیت کم سوخت و تراکم ترافیک باعث شده که روزانه حجم وسیعی از آلاینده‌های غیرسازگار با مکانیسم‌های طبیعی در هوا تخلیه گردد. این مساله یکی از معضله‌های جهانی رو به رشد است و از شاخصه‌های مهم آن این است که در یک لحظه ده‌ها هزار نفر یکجا از آن آسیب می‌بینند؛ بنابراین پیشبینی آلودگی هوا یکی از راهکارهای مدیریتی برای جلوگیری و یا کاهش پیامدهای مخرب آن می‌باشد (۲). آلودگی هوای تنفسی در شهرهای بزرگ در بعضی از موارد از حد استاندارد هم بالاتر می‌رود و این موضوع وضعیت نگران کننده‌ای را به وجود می‌آورد؛ این امر باعث شده که در بسیاری از شهرهای دنیا از آلودگی هوا به عنوان یک معضل یاد شود لذا بررسی آلودگی هوا به عنوان یک بحران و اقدامات لازم جهت مدیریت آن ضروری است. تعطیلی‌های اخیر شهر تهران نشان از اهمیت این مساله دارد. حساسیت عمومی نسبت به این مساله مسئولین را وادار به وضع قوانینی برای جلوگیری از این آلودگی کرده است (۳). یکی از اهداف مسئولین آماده‌سازی اطلاعات برای شهروندان جهت آگاهی از میزان کیفیت هوا است (۴)؛ همچنین آگاهی عمومی از میزان کیفیت هوا در روزهای آینده با توجه به روند موجود، می‌تواند در کنترل این آلودگی بسیار موثر باشد. برای دستیابی به این هدف باید ابزاری مناسب جهت مدل‌سازی و پیشبینی آلودگی هوا به کار گرفته شود (۵). از طرفی یکی از

برخی تحقیقات نیز پیشبینی آلودگی هوا با استفاده از مدل‌های خاص آماری را بررسی کرده‌اند. پتلین و همکاران برای پیشبینی غلظت ازن به صورت ۲۴ ساعته از مدل فرآیند گوسی با مرتبه یک و مرتبه بالاتر در شهر بلغاریا بلغارستان استفاده کرده‌اند. نتایج آن‌ها نشان داد که مدل مرتبه ۲ از بهترین دقت برخوردار است (۱۶). شلینک و همکاران ۱۵ روش آماری مختلف را برای مجموعه داده‌ای شامل ۱۰ منطقه اروپایی به کار گرفتند. نتایج آنها نشان داد که بهترین پیشبینی در مناطق روستایی و زیرمجموعه شهری در اروپای مرکزی اتفاق می‌افتد. همچنین مدل‌های GAM (Generalised Additive Model) و شبکه عصبی از بیش‌ترین دقت برخوردار هستند (۱۷). لاوسون و همکاران سری زمانی ساختار یافته را برای پیشبینی نیتروژن دی‌اکسید و نیتروژن اکسید در شهر دوبلین اتریش بکار گرفته‌اند. نتایج آن‌ها نشان داد زمانی که پارامترهای هواشناسی و حجم ترافیک در دسترس نیست، سری زمانی ساختار یافته ابزاری مناسب جهت پیشبینی آلودگی هوا می‌باشد (۱۸). این مطالعات معمولاً با حجم زیادی از داده‌ها (داده‌های غلظت ساعتی به صورت حداقل ۲ ساله) انجام شده‌اند که حجم محاسباتی بالایی را طلب می‌کنند.

همچنین تحقیقات زیادی در مورد پیشبینی سری زمانی حاصل از مقادیر ساعتی اندازه‌گیری شده در ایستگاه‌های پایش کیفیت هوا با استفاده از شبکه‌های عصبی انجام شده است. دیاز روبلس و همکاران از روشی ترکیبی شامل مدل ARIMA و شبکه عصبی برای پیشبینی میزان ذرات کوچک‌تر از ۱۰ میکرون به صورت ۲۴ ساعته در شهر تموکو شیلی استفاده کرده و نتایج را با Multi Linear (MLR) Regression مقایسه کرده‌اند. نتایج آنها نشان می‌دهد که روش ترکیبی عملکرد خوبی در پیشبینی ذرات کوچک‌تر از ۱۰ میکرون دارد (۱۹). پسیونی و همکاران مدل NARX را برای پیشبینی مقدار پیک ازن بکار گرفته‌اند. نتایج آن‌ها نشان داد که این نوع شبکه عصبی در پیشبینی سری زمانی غلظت ازن در شهر میلان عملکرد خوبی داشته است (۲۰). نیسکا و همکاران مدلی بر پایه شبکه عصبی و الگوریتم ژنتیک برای

در این مقاله با هدف مدلسازی و پیشبینی زمانی غلظت مونوکسیدکربن از شبکه عصبی NAR (None Linear Auto regressive) و مدل آماری ARMA استفاده شده است. این دو روش قادر به پیشبینی فرآیندهای اتورگرسیو هستند. فرآیند اتورگرسیو فرآیندی است که به کمک داده‌های پیشین قادر به پیشبینی مقادیر در آینده است (۸). در نهایت قابلیت‌های شبکه عصبی NAR که یک شبکه عصبی پویا با قدرت تشخیص فرآیندهای غیرخطی است؛ با مدل آماری ARMA که یک مدل پارامتریک خطی است، در مدلسازی و پیشبینی غلظت مونوکسید کربن با استفاده از پارامترهای آماری مختلف شامل R (ضریب همبستگی) و MSE (Mean Square Error) مقایسه شده است.

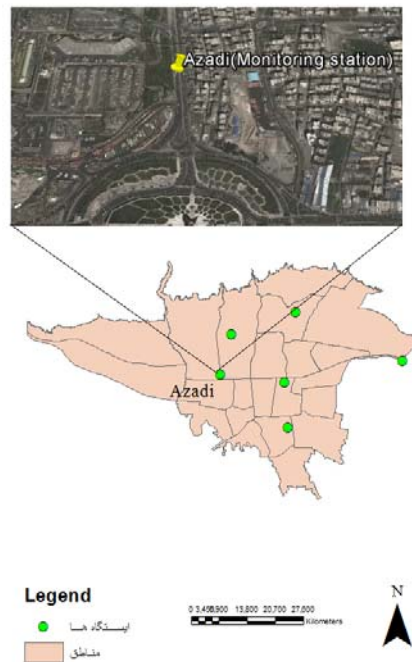
تحقیقات متعددی در زمینه پیشبینی کیفیت هوا با استفاده از شبکه‌های عصبی انجام شده است. این تحقیقات نشان می‌دهند که شبکه‌های عصبی ابزاری مناسب جهت مدلسازی زمانی پارامترهای کیفیت هوا هستند (۴، ۱۱-۱۳)؛ بعضی مطالعات به پیشبینی غلظت آلاینده با استفاده از تحلیل اجزای اصلی پرداخته‌اند. العلوی و همکاران با هدف مدلسازی و پیشبینی غلظت ازن به صورت ۲۴ ساعته در کویت روشی ترکیبی شامل رگرسیون چندمتغیره، آنالیز اجزاء اصلی و شبکه عصبی مصنوعی را به کار گرفته است (۱۴). در تحقیقی مشابه سوزا و همکاران به توسعه مدلی با استفاده از شبکه عصبی بر پایه آنالیز اجزاء اصلی برای پیشبینی غلظت ازن به صورت ۲۴ ساعته در شهر اسپورتو پرتغال پرداخته‌اند و نتایج حاصل از آن را با رگرسیون چند متغیره و شبکه عصبی مقایسه کردند (۱۵). سینگ و همکاران انواع روش‌های خطی و غیر خطی را برای پیشبینی غلظت آلاینده‌های بکار بردند؛ و به این نتیجه رسیدند که شبکه عصبی GRNN (Generalized Regression Neural Network) دارای بهترین کارایی در پیشبینی غلظت آلاینده‌ها است (۴). اکثر مدل‌های ذکر شده نیاز به داده‌های اضافی علاوه بر داده‌های غلظت آلاینده مورد نظر، شامل پارامترهای هواشناسی و یا ترافیک دارند؛ در حالی که ممکن است این داده‌ها در دسترس نباشند.

شده است. این در حالی است که در مقالات مشابه از میانه به جای میانگین استفاده شده است (۲۱). داده‌های غلظت آلاینده مونوکسیدکربن از ایستگاه آزادی محیط زیست استان تهران اخذ شده و مربوط به ۲۹ نوامبر تا ۳۱ دسامبر ۲۰۰۹ می باشد. داده‌های این ایستگاه به این علت مورد استفاده قرار گرفت که آزادی یکی از شلوغ‌ترین مناطق شهر تهران از لحاظ ترافیک و وجود آلاینده‌های مضر هوا است. شکل ۱ موقعیت ایستگاه‌های سنجش کیفیت هوا را نشان می‌دهد.

پیشبینی سری زمانی نیتروژن دی اکسید در شهر هلسینکی یونان بکار گرفته‌اند؛ نتایج آن‌ها نشان داد که الگوریتم ژنتیک قادر به رفع محدودیت‌های شبکه عصبی است (۱۳).

مواد و روش‌ها

در این مقاله داده‌های ساعتی غلظت آلاینده مونوکسیدکربن متعلق به محیط زیست استان تهران مورد استفاده قرار گرفته است. سری زمانی حاصل از غلظت مونوکسید کربن مورد بررسی قرار گرفته و برای پیدا کردن مقادیر استثنایی (Outlier) از میانگین و انحراف معیار استفاده



شکل ۱- موقعیت ایستگاه‌های سنجش کیفیت هوا

Fig 1- Location of air pollution monitoring stations

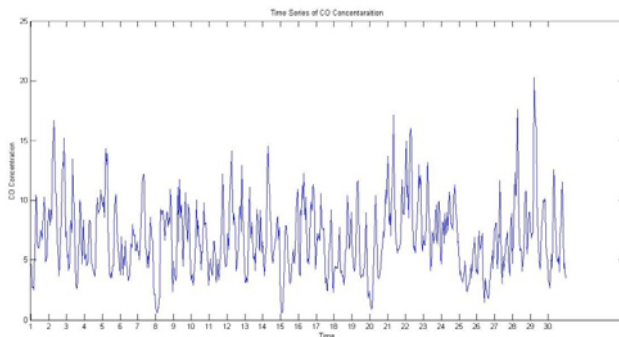
$$\frac{\text{حجم یگانه از شمارهای آلوده}}{10^6 \text{ حجم هوا و شمارهای آلوده}} = 1 \text{ ppm} \quad ۱$$

$$\frac{\text{حجم یگانه از شمارهای آلوده}}{10^9 \text{ حجم هوا و شمارهای آلوده}} = 1 \text{ ppb} \quad ۲$$

معمولاً میزان آلاینده موجود در هوا بر حسب تعداد در میلیون یعنی ppm (Part Per Million) و یا تعداد در بیلیون یعنی ppb (Part Per Billion) و مطابق با روابط ۱ و ۲ بیان می‌شود (۲۲).

می‌شود غلظت ازن در ساعت مختلف دارای آشفتگی زیادی است.

به عنوان آنالیز اولیه سری زمانی بررسی گراف حاصل از سری زمانی می‌تواند مفید باشد (۲۳). شکل ۲ سری زمانی غلظت مونوکسید کربن را نشان می‌دهد. همانطور که در شکل دیده



شکل ۲ - سری زمانی غلظت مونوکسید کربن

Fig 2- Time Series of carbon monoxide concentration

همچنین بررسی مشخصات آماری سری زمانی مانند بیشینه، کمینه و میانگین می‌تواند مفید باشد. مشخصات آماری غلظت آلاینده مونوکسید کربن در جدول ۱ آمده است.

جدول ۱- پارامترهای آماری سری زمانی غلظت مونوکسید کربن در ایستگاه آزادی

در سال ۲۰۰۹ از ۲۹ نوامبر تا ۳۰ دسامبر

Table1- Statistical parameters of carbon monoxide concentration time series in Azadi station from 29 Nov 2009 to 30 Dec 2009

پارامتر آماری	غلظت (ppm)
میانگین	۶/۹۰۶
بیشینه	۲۰/۲۶
کمینه	۰/۵۷۴
میانه	۶/۵۷۵
انحراف معیار	۳/۰۵۳

عدد نزدیک به ۱ باشد، نشان‌دهنده همبستگی بیشتری بین خروجی شبکه و مقادیر مورد انتظار است (رابطه ۳).

$$Corr(t, y) = \frac{Cov(t, y)}{\sqrt{Var(t)Var(y)}} \quad ۳$$

$Cov(t, y)$ نشان‌دهنده کوواریانس بین مقادیر مقایره واقعی t و مقدار پیش‌بینی شده y است؛ $Var(t)$ و $Var(y)$ به ترتیب

برای نمایش میزان خطای روش به کار گرفته شده از R و MSE (Mean Squared Error) استفاده شده است. معمولاً برای نشان دادن اجرای درست روش مدلسازی از ضریب همبستگی بین مقادیر خروجی و مقادیر مورد نظر استفاده می‌شود. R تغییرات بین مقادیر بدست آمده از مدل و مقادیر واقعی را نشان می‌دهد. این عدد بین صفر و ۱ است، هر چه این

H شاخص هرست است؛ و پایایی را برای سری زمانی مشخص نشان می‌دهد (۲۶). این شاخص برای یک سری زمانی عددی بین ۰ و ۱ را مشخص می‌کند. $H=0.5$ نشان‌دهنده این است که سری زمانی مورد مطالعه حاصل یک فرآیند تصادفی است. $0.5 < H < 1$ نشان می‌دهد که اگر سیستم در یک دوره دارای روند افزایشی باشد احتمال اینکه در آینده هم روند افزایش داشته باشد وجود دارد. اگر $0 < H < 0.5$ باشد در این حالت سیستم روند کاهشی داشته و احتمال روند افزایشی در آینده وجود دارد (۲۷).

مدل آماری ARMA

برای پیشبینی و مدلسازی سری‌های زمانی روش‌های مختلفی وجود دارد. یکی از روش‌های تحلیلی برای پیشبینی و مدلسازی سری‌های زمانی مدل ARMA است. در این تحقیق از مدل آماری ARMA که یک مدل آماری خطی است، برای مدلسازی و پیشبینی غلظت مونوکسید کربن استفاده شده است. مدلی که سری زمانی y_n را با استفاده از ترکیب خطی مشاهدات پیشین و نویز v_{n-i} توصیف می‌کند، مدل ARMA نامیده می‌شود.

$$y_n = \sum_{i=1}^m a_i y_{n-i} + v_n - \sum_{i=1}^{\ell} b_i v_{n-i}$$

در رابطه فوق M و a_i به ترتیب درجه اتورگرسیو و ضرایب اتورگرسیو نامیده می‌شوند؛ و به صورت مشابه l و b_i نیز به ترتیب درجه میانگین متحرک و ضرایب میانگین متحرک نامیده می‌شوند. در واقع ترم مهم این معادله قسمت AR است که به کمک داده‌های پیشین مقادیر را می‌تواند پیشبینی کند (۲۳).

$$y_n = \sum_{i=1}^m a_i y_{n-i} + v_n$$

همچنین به این فرآیند یعنی پیشبینی مقادیر اخیر به کمک داده‌های پیشین اتورگرسیو گفته می‌شود. قابل ذکر است که برای تشخیص درجه ARMA از معیار BIC (Bayesian Information Criterion) استفاده شده است، به این ترتیب

واریانس t و y است؛ در حالی که $\text{cov}(t, y)$ همبستگی بین مقادیر واقعی و پیشبینی شده را نشان می‌دهد.

دقت مدل نیز معمولاً با استفاده از MSE سنجیده می‌شود، که میانگین مربعات اختلاف بین مقادیر واقعی و مقادیر بدست آمده از شبکه عصبی را نشان می‌دهد.

$$MSE = \frac{1}{N} \sum_{i=1}^N (e_i)^2 = \frac{1}{N} \sum_{i=1}^N (t_i - y_i)^2 \quad ۴$$

در این رابطه ۴ ، N تعداد نمونه‌ها (مقادیر پیشبینی شده) است.

شاخص هرست (Hurst Index)

داده‌های غلظت آلاینده‌ها معمولاً توسط ایستگاه‌های پایش کیفیت هوا در شهرهای بزرگ به صورت ساعتی اندازه‌گیری می‌شوند (۲۴). به این گونه اطلاعات که راجع به پدیده‌ای خاص، که به طور نامنظم در طول زمان تغییر می‌کند، سری زمانی گویند (۲۵). برای تحلیل رفتار سری زمانی و قابلیت پیشبینی آن می‌توان تحلیل R/S را به کار گرفت. آنالیز R/S توسط هرست به منظور تعیین میزان پایداری سری زمانی مورد استفاده قرار گرفت. روش محاسبه به این صورت است که (۷):

سری زمانی X_t به d بردار n تایی تقسیم می‌شود.

برای هر بردار Z_m ، میانگین μ_m و انحراف معیار S_m محاسبه می‌شود.

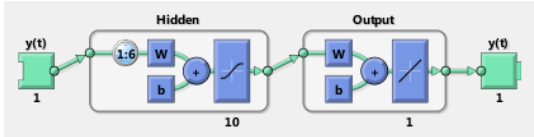
$R_m = \max - \min$ نیز محاسبه می‌گردد.

با بدست آوردن مقیاس بندی مجدد R_m/S_m به ازای هر m ، مقدار متوسط مقیاس بندی مجدد بردارهای n بعدی بصورت زیر خواهد بود:

$$\sum_{m=1}^d \frac{R_m}{S_m} \left(\frac{R}{S}\right)_n = \frac{1}{d} \quad ۵$$

بر پایه تحلیل R/S تشخیص یک سری زمانی تصادفی از یک سری زمانی غیر تصادفی امکان‌پذیر است. همچنین بر اساس یک قانون تجربی بیان می‌گردد که:

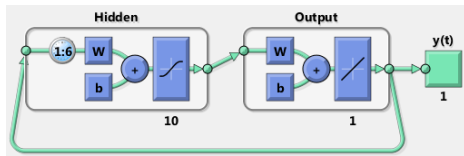
$$\left(\frac{R}{S}\right)_n = (n)^H \quad ۶$$



شکل ۴- معماری شبکه عصبی به کار گرفته شده در این

تحقیق

Fig 4- The architecture of applied neural network بعد از آموزش شبکه، شبکه برای پیشبینی آینده می‌تواند مورد استفاده قرار گیرد. برای این کار باید شبکه به صورت حلقوی مورد استفاده قرار گیرد. شکل زیر منظور از حالت حلقوی را در شبکه عصبی مورد استفاده نشان می‌دهد.



شکل ۵- شبکه عصبی بسته

Fig 5- A closed loop neural network

نتایج و یافته‌ها

شاخص هرست

همانطور که ذکر شد نمای هرست برای هر سری زمانی مقداری بین صفر تا ۱ است؛ مقدار H برای سری زمانی مورد استفاده در این تحقیق نیز محاسبه شد و مقدار آن برابر با 0.80 بدست آمد؛ که نشان‌دهنده رفتار پایا برای سری زمانی مورد استفاده در این تحقیق است.

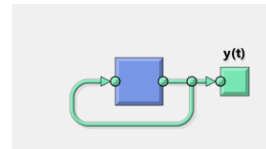
مدلسازی و پیشبینی غلظت مونوکسید کربن

برای اجرای مدل آماری ARMA نیاز به محاسبه تعدادی ضرایب می‌باشد. برای این کار از معیار BIC استفاده شد. به این ترتیب که مدل‌های با تعداد درجات مختلف AR و MA برای داده‌های یک ماه ایجاد شده و مقدار معیار BIC برای هر کدام از آن‌ها محاسبه شد، که ARMA(5,5) از میان همه مدل‌ها دارای BIC کمتر بود؛ در نتیجه این مدل انتخاب شد. مقدار BIC برای ضرایب ۱ تا ۶ در جدول ۲ آمده است.

که نتایج چندین مدل با درجات مختلف امحاسبه شد، سپس مدلی که کمترین مقدار را از نظر میزان این معیار دارا بود به عنوان مدل مناسب برگزیده شد.

شبکه عصبی NAR

در مقابل روش‌های آماری، شبکه‌های عصبی مصنوعی است که نیاز به مهارت زیادی در استفاده از آن‌ها وجود ندارد. در این تحقیق شبکه عصبی NAR که قابلیت پیشبینی فرآیندهای اتورگرسیو را داراست (۲۸)، نیز مورد استفاده قرار گرفته است. شبکه عصبی NAR یک شبکه عصبی پویا بازگشتی است که دارای چندین لایه است. شکل زیر منظور از بازگشتی بودن این نوع شبکه عصبی را نشان می‌دهد.



شکل ۳- شبکه عصبی بازگشتی

Fig 3- A recurrent neural network

این شبکه با داشتن مقادیر پیشین سری زمانی $x(t-1), \dots, x(t-d)$ قادر به پیشبینی مقادیر آینده $x(t)$ است (۲۹).

$$x(t) = f(x(t-1), \dots, x(t-d)) \quad 9$$

شبکه عصبی استفاده شده در این تحقیق دارای ۱۰ نرون در لایه میانی و تاخیر زمانی ۱:۶ است؛ بعد از چندین بار آموزش با تعداد نرون‌های مختلف و تاخیرهای زمانی متفاوت تعداد ۱۰ نرون در لایه میانی و تاخیر ۱:۶ وضعیت بهتری داشت و به همین دلیل مورد استفاده قرار گرفت. الگوریتم مورد استفاده برای آموزش شبکه، LM (Levenberg Marquardt) می‌باشد. شکل ۴ معماری شبکه مورد استفاده در این مقاله را نشان می‌دهد.

جدول ۲- مقدار معیار BIC برای هر کدام از مدل‌ها با ضرایب مختلف

(مدل با درجه $MA=5$ و $AR=5$ بهتر از بقیه است)Table 2- The value of BIC criteria for each model (The Model with $AR=5$ and $MA=5$ is the best)

	MA6	MA5	MA4	MA3	MA2	MA1
AR1	۲۸۹۳/۷۴	۲۸۹۷/۷۶	۲۸۹۱/۱۴	۲۸۸۶/۵۴	۲۸۹۰/۸۹	۲۸۹۰/۵۸
AR2	۲۸۹۸/۳۲	۲۸۳۰/۶۹	۲۸۹۳/۹۷	۲۸۹۱/۰۹	۲۸۸۷/۸۱	۲۸۸۳/۳۵
AR3	۲۸۳۰/۹۵	۲۹۰۰/۵۱	۲۸۹۹/۲۲	۲۸۹۵/۱۱	۲۸۹۱/۶۴	۲۸۸۷/۸۵
AR4	۲۸۲۸/۵۹	۲۹۰۰/۵۴	۲۸۹۶/۶۲	۲۸۹۲/۵۲	۲۸۸۸/۰۳	۲۸۹۰/۱۹
AR5	۲۸۲۶/۱	۲۸۱۴/۸۹	۲۹۰۱/۲۲	۲۸۹۳/۷۶	۲۸۹۲/۵۴	۲۸۹۳/۹۸
AR6	۲۸۲۲/۸	۲۹۰۰/۶۴	۲۸۲۰/۶۳	۲۹۰۰/۹۷	۲۹۰۰/۴۴	۲۸۹۷/۸۹

برای مدلسازی غلظت مونوکسید از داده‌های یک ماه استفاده شده است؛ و طبق آخرین ماه سال ۲۰۰۹ ضرایب مربوط به مدل به دست آمد. ضرایب این مدل طبق جدول ۳ است.

جدول ۳- ضرایب و مقادیر بدست آمده از مدل $ARMA(5,5)$

برازش داده شده به داده‌های یک ماه

Table 3- The coefficient and values for $ARMA(5,5)$ fitted to one month data.

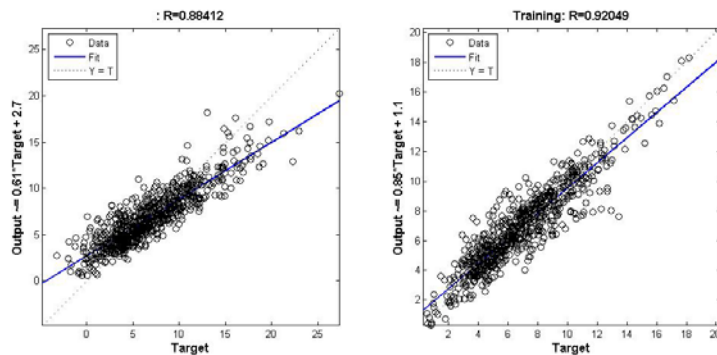
Parameter	Value	Parameter	Value
Constant	۰/۰۴۶۷۳	MA1	۰/۷۴۸۵
AR1	۰/۰۸۰۷	MA2	-۰/۳۶۷
AR2	۰/۸۱۶۳	MA3	-۰/۴۵۳
AR3	۰/۱۵۵۵	MA4	۰/۲۷۴۳
AR4	-۰/۵۱۷	MA5	-۰/۰۱۵
AR5	۰/۲۵۰۲	Variance	۴/۶۶۷

شده است. برای شبکه عصبی مورد نظر شرط توقف الگوریتم رسیدن به ۴۰۰۰۰ تکرار در نظر گرفته شد.

شکل ۷، ۶ آنالیز رگرسیونی بین پاسخ مدل و مقادیر واقعی را برای داده‌های آموزشی در شبکه عصبی NAR و مدل ARMA را نشان می‌دهد.

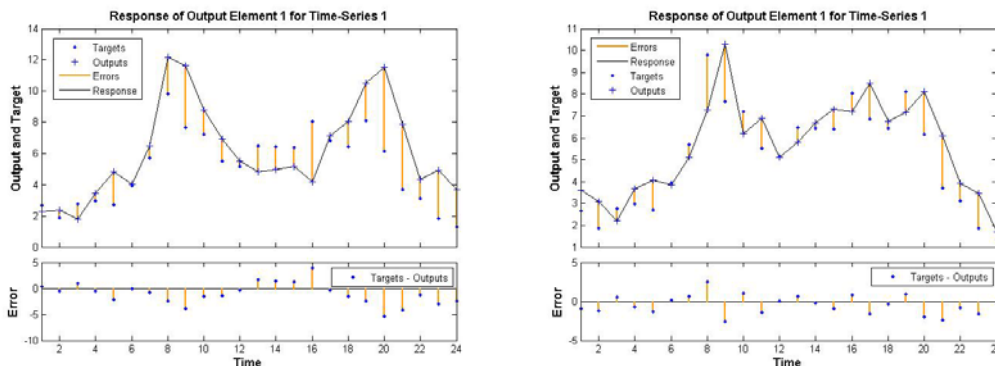
بعد از بدست آوردن ضرایب می‌توان دقت مدل را برای داده‌های مورد استفاده محاسبه کرد.

برای مدلسازی غلظت مونوکسید کربن و آموزش شبکه عصبی همچنین از داده‌های یک ماه (۲۹ نوامبر تا ۳۰ دسامبر) استفاده



شکل ۶- راست: نتایج حاصل از شبکه عصبی NAR چپ: نتایج حاصل از مدل ARMA برای داده‌های مدل‌سازی (محور افقی مقادیر واقعی و محور عمودی مقادیر پیش‌بینی شده را نشان می‌دهند).

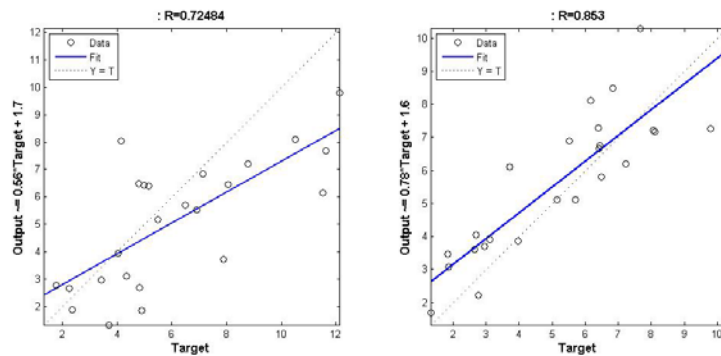
Fig 6- Right: The results of NAR neural networks Left: The results of ARMA model (Horizontal axis is the real data and vertical axis is the modeled value)



شکل ۷- نمودار راست: شبکه عصبی NAR چپ: مدل ARMA (رابطه میان غلظت پیش‌بینی شده به صورت ۲۴ ساعته با مقادیر واقعی)

Fig 7- Right: NAR neural network Left: ARMA model (The relation between 24h predicted and real values)

شکل ۸ ارتباط رگرسیونی بین مقادیر واقعی و مقادیر پیش‌بینی شده را نشان می‌دهد. همانطور که در شکل دیده می‌شود، شبکه عصبی NAR دارای دقت بهتری در پیش‌بینی ۲۴ ساعته غلظت مونوکسید کربن است.



شکل ۸- ارتباط رگرسیونی بین مقادیر واقعی و مقادیر پیشبینی شده توسط راست: شبکه عصبی NAR چپ:

توسط ARMA (محور افقی مقادیر واقعی و محور عمودی مقادیر پیشبینی شده را نشان می‌دهند).

Fig 8- The regression relation between real and predicted values Right: NAR neural network Left: ARMA model (Horizontal axis is real data and vertical axis is predicted value)

کربن می‌باشد. این شبکه عصبی با استفاده از داده‌های غلظت آلاینده مونوکسید کربن مربوط به گذشته می‌تواند برای پیشبینی مقدار غلظت این آلاینده برای ۲۴ ساعت آینده در حالی که داده‌های هواشناسی و ترافیک موجود نباشد، با دقت مطلوبی به کار گرفته شود. این شبکه دارای بار محاسباتی کمتری نسبت به روش‌های کلاسیک است. در حالی که روش‌های کلاسیک نیاز به مهارت و تجربه زیاد برای به کارگیری دارند؛ شبکه‌های عصبی راحت‌تر از روش‌های کلاسیک می‌توانند به کار گرفته شوند. در تحقیقات مشابه از داده‌های چند ساله و روش‌هایی استفاده شده که بار محاسباتی بالایی را تحمیل می‌کنند (۶). در بعضی تحقیقات دیگر نیز که از شبکه‌های عصبی استفاده شده است، از داده‌های همبسته و پارامترهای هواشناسی نیز استفاده شده است (۷). در صورتی که در این مقاله از داده‌های یک ماه غلظت آلاینده مونوکسید کربن بدون بهره‌گیری از پارامترهای همبسته دیگر برای مدل‌سازی استفاده شد؛ و دقت نسبتاً مطلوبی به دست آمد. پیشبینی انجام شده برای ۲۴ ساعت آینده می‌تواند برای کنترل و مدیریت آلودگی، و همچنین آگاهی عمومی مورد استفاده قرار گیرد.

نتایج پیشبینی ۲۴ ساعته با استفاده از مدل آماری ARMA، خطای $MSE=5.461$ را نشان می‌دهد، در حالی که میزان این خطا برای شبکه عصبی NAR، $MSE=1.752$ است. همچنین همبستگی بین مقادیر پیشبینی شده و مقادیر واقعی برای مدل ARMA، $R=0.72$ بدست آمد؛ مقدار این پارامتر برای شبکه عصبی NAR، $R=0.85$ بود؛ که نشان از برتری شبکه عصبی نسبت به این مدل آماری دارد.

بحث و نتیجه‌گیری

آلودگی هوا یکی از مشکلات بزرگ جوامع شهری است. یکی از اهداف مسئولین پیشبینی میزان آلودگی هوا جهت اطلاع رسانی عمومی است. برای پیشبینی میزان آلودگی هوا باید روشی مناسب جهت مدل‌سازی غلظت هریک از آلاینده‌ها به کار گرفته شود. در این مقاله به بررسی شبکه عصبی بازگشتی NAR و مدل آماری ARMA جهت پیشبینی ساعتی غلظت مونوکسید کربن برای ۲۴ ساعت آینده پرداخته شد. نتایج این تحقیق نشان داد که دقت مدل NAR ($MSE=$) $1.6685ppm^2$ و $R=0.84$) بهتر از مدل ARMA ($MSE=5.461$ و $R=0.72$) که یک روش کلاسیک و خطی محسوب می‌شود، در پیشبینی ۲۴ ساعته غلظت مونوکسید

فنی . گروه نقشه برداری . رشته نقشه برداری سیستم اطلاعات مکانی; ۱۳۸۷.

۷. هاشمی ف. بررسی و ارزیابی الگوریتم‌های مونت کارلو و شبکه عصبی برای پیش‌بینی آلودگی هوا در محیط یک سیستم اطلاعات مکانی زمانمند. تهران: دانشکده مهندسی نقشه برداری، دانشگاه صنعتی خواجه نصیرالدین طوسی، رشته نقشه برداری سیستم اطلاعات مکانی; ۱۳۹۰.

8. Diaconescu E. The use of NARX neural networks to predict chaotic time series. WSEAS Trans Comp Res. 2008; 3(3):182-91.

۹. منهای م.ب. مبانی شبکه‌های عصبی. تهران: مرکز نشر دانشگاه امیرکبیر; ۱۳۸۹.

10. Pisoni E, Farina M, Pagani G, Piroddi L. Environmental Over-Threshold Event Forecasting using NARX Models. Preprints of the 18th IFAC World Congress; Milan 2011.

11. Kurt A, Gulbagci B, Karaca F, Alagha O. An online air pollution forecasting system using neural networks. Environment international. 2008; 34(5):592-8.

12. Karatzas KD, Kaltsatos S. Air pollution modelling with the aid of computational intelligence methods in Thessaloniki, Greece. Simulation Modelling Practice and Theory. 2007; 15(10):1310-9.

13. Niska H, Hiltunen T, Karppinen A, Ruuskanen J, Kolehmainen M. Evolving the neural network model for forecasting air pollution time series. Engineering Applications of Artificial Intelligence. 2004; 17(2):159-67.

14. Alalawi S, Abdulwahab S, Bakheit C. Combining principal component regression and artificial neural

در این مقاله فقط داده‌های مربوط به غلظت آلاینده مونوکسید کربن به کار گرفته شده است، اما با توجه به همبستگی پارامترهای هواشناسی با آلودگی هوا، اگر داده‌های هواشناسی اندازه‌گیری شده در همان ایستگاه‌ها و در مدلسازی دخالت داده شوند ممکن است پیش‌بینی به نتایج بهتری منجر شود؛ همچنین با توجه به همبستگی ترافیک با میزان غلظت مونوکسیدکربن، استفاده از داده‌های ترافیک ساعتی نیز می‌تواند به کارایی بهتر مدل کمک کند.

منابع

۱. ندافی کج، م.؛ حسنوند، م.ص.؛ نقی‌زاده، ع. آلودگی هوا منشأ و کنترل آن. تهران: انتشارات نص; ۱۳۹۰.

۲. خزایی ا. اکتشاف دانش به روش فازی-عصبی در محیط GIS (مطالعه موردی آلودگی هوا). تهران: دانشکده مهندسی نقشه برداری، دانشگاه صنعتی خواجه نصیرالدین طوسی، رشته نقشه برداری سیستم اطلاعات مکانی ۱۳۹۰.

3. Brunelli U, Piazza V, Pignato L, Sorbello F, Vitabile S. Three hours ahead prevision of SO2 pollutant concentration using an Elman neural based forecaster. Building and Environment. 2008; 43(3):304-14.

4. Singh KP, Gupta S, Kumar A, Shukla SP. Linear and nonlinear modeling approaches for urban air quality prediction. The Science of the total environment. 2012; 426:244-55.

5. Fernando HJ, Mammarella MC, Grandoni G, Fedele P, Di Marco R, Dimitrova R, et al. Forecasting PM10 in metropolitan areas: Efficacy of neural networks. Environmental pollution. 2012; 163:62-7.

۶. هاتفی‌افشار، ا. پیش‌بینی آلودگی هوا با استفاده از داده کاوی مکانی. تهران: دانشگاه تهران. دانشکده

- on air pollutants concentrations in the region of Annaba, Algeria. *Atmospheric Research*. 2012; 113:89-101.
۲۲. فتح‌تبار فیروزجایی س. ارزیابی روش‌های مختلف درون‌یابی در پهنه‌بندی آلاینده‌ها در شهر تهران. تهران: دانشگاه آزاد اسلامی واحد علوم تحقیقات، رشته سنجش از دور و سیستم اطلاعات مکانی؛ ۱۳۹۰.
23. Kitagawa G. *Introduction to Time Series Modeling*: Taylor & Francis; 2010.
 24. Barrero MA, Grimalt JO, Cantón L. Prediction of daily ozone concentration maxima in the urban atmosphere. *Chemometrics and Intelligent Laboratory Systems*. 2006; 80(1):67-76.
 25. Tsay RS. *Analysis of Financial Time Series*: Wiley; 2010.
 26. Reiter C. *Fractals Visualization and J*: Lulu.com; 2007.
 27. Alvarez-Ramirez J, Echeverria JC, Rodriguez E. Performance of a high-dimensional method for Hurst exponent estimation. *Physica A: Statistical Mechanics and its Applications*. 2008;387(26):6452-62.
 28. M. Abbaspour AMRaMT. Carbon monoxide prediction using novel intelligent network *International Journal of Environmental Science & Technology*. Winter 2005; Vol. 1(No. 4):257-64.
 29. Zemouri R, Gouriveau R, Zerhouni N. Defining and applying prediction performance metrics on a recurrent NARX time series model. *Neurocomputing*. 2010;73(13-15):2506-21.
 - networks for more accurate predictions of ground-level ozone. *Environmental Modelling & Software*. 2008; 23(4):396-403.
 15. Sousa S, Martins F, Alvimferraz M, Pereira M. Multiple linear regression and artificial neural networks based on principal components to predict ozone concentrations. *Environmental Modelling & Software*. 2007; 22(1):97-103.
 16. Petelin D, Grancharova A, Kocijan J. Evolving Gaussian process models for prediction of ozone concentration in the air. *Simulation Modelling Practice and Theory* 80-33: 68; 2013.
 17. Schlink U, Herbarth O, Richter M, Dorling S, Nunnari G, Cawley G, et al. Statistical models to assess the health effects and to forecast ground-level ozone. *Environmental Modelling & Software*. 2006; 21(4):547-58.
 18. Lawson AR, Ghosh B, Broderick B. Prediction of traffic-related nitrogen oxides concentrations using Structural Time-Series models. *Atmospheric Environment*. 2011; 45(27):4719-27.
 19. Díaz-Robles LA, Ortega JC, Fu JS, Reed GD, Chow JC, Watson JG, et al. A hybrid ARIMA and artificial neural networks model to forecast particulate matter in urban areas: The case of Temuco, Chile. *Atmospheric Environment*. 2008; 42(35):8331-40.
 20. Pisoni E, Farina M, Carnevale C, Piroddi L. Forecasting peak air pollution levels using NARX models. *Engineering Applications of Artificial Intelligence*. 2009; 22(4-5):593-602.
 21. Khedairia S, Khadir MT. Impact of clustered meteorological parameters

