

مدل سازی و پیش بینی وضعیت آلاینده های هوای شهر تهران

کاربرد مدل خود رگرسیونی با ویژگی حافظه بلندمدت

رضا اخباری*^۱

rakhbary@gmail.com

حمید آماده^۲

تاریخ پذیرش: ۹۳/۱۲/۲۷

تاریخ دریافت: ۹۳/۸/۱۹

چکیده

زمینه و هدف: مدل سازی آلاینده های زیست محیطی یکی از نیازهای اساسی در زمینه پایش کیفیت هوا محسوب می شود که با بهره گیری از نتایج حاصله می توان اقدامات پیشگیرانه ای جهت بهبود شرایط آبی اتخاذ کرد. ادبیات موجود در زمینه الگوسازی آلاینده های زیست محیطی را می توان به دو دسته کلی تقسیم کرد، دسته اول شامل مطالعاتی می شود که علاوه بر داده های مربوط به آلاینده ها با وارد کردن عوامل محیطی از قبیل دمای هوا، جهت وزش باد، سرعت وزش باد و میزان رطوبت، وضعیت انتشار را مورد بررسی قرار داده اند. دسته دوم مطالعات -که تحقیق حاضر در این دسته می گنجد- با استفاده از الگوهای رگرسیون سری های زمانی و غالباً با استفاده از داده های موجود هر آلاینده، پیش بینی وضعیت آبی آن را مد نظر قرار داده اند.

روش بررسی: در این مقاله با استفاده از سه الگوی (ARIMA(AutoRegressive Integrated Moving Average) ، (ARFIMA(AutoRegressive Fractionally Integrated Moving Average) و (ARIMA-GARCH(Generalized AutoRegressive Conditional Heteroskedasticity) و رویکرد باکس-جنکینز وضعیت آبی آلاینده های NO_2 ، PM_{10} ، CO ، SO_2 ، O_3 و $PM_{2.5}$ در شهر تهران پیش بینی شد و در مورد هر آلاینده بهترین مدل بر اساس معیارهای MSE (Mean Squared Error)، $MAPE$ (Mean Absolute Error) و MAE (Mean Absolute Error)، $RMSE$ (Root Mean Squared Error)، $Error$ (Percent Error) معرفی گردید.

یافته ها: آن چه این مطالعه را از مطالعات قبلی متمایز می سازد، مد نظر قرار دادن ویژگی حافظه بلندمدت و مقایسه دقت خروجی مدل مربوطه با الگوهای رایج خود رگرسیونی است. نتایج نشان می دهد که فرض وجود حافظه بلندمدت پذیرفته خواهد شد، ولی این که بهترین پیش بینی ها همواره توسط مدل ARFIMA ارائه می شود، رد می شود.

۱- کارشناس ارشد اقتصاد محیط زیست، دانشگاه علامه طباطبایی* (مسوول مکاتبات)

۲- استادیار گروه اقتصاد کشاورزی و محیط زیست دانشکده اقتصاد دانشگاه علامه طباطبایی

بحث و نتیجه گیری: این مطالعه کاربرد مدل‌های اقتصادسنجی را برای پیش‌بینی وضعیت آلاینده‌ها اثبات می‌کند. براین اساس توصیه می‌شود با توجه به هزینه‌های اجتماعی بالای انتشار آلاینده‌ها، با بکارگیری این الگوها، آلاینده‌های تأثیرگذار بر آینده هوای شهر شناسایی و در جهت کاستن از سطح انتشار آن‌ها طرح‌های کارآمدی پیاده شود.

واژه های کلیدی: پیش بینی، حافظه بلندمدت، مدل خودرگرسیون هم انباشته کسری (ARFIMA)، آزمون GPH، آزمون R/S اصلاح شده.

Modeling and Forecasting Air Pollution of Tehran Application of Autoregressive Model with Long Memory Properties

Reza Akhbari^{1*}

rakhbary@gmail.com

Hamid Amadeh²

Admission Date: March 18, 2015

Date Received: November 10, 2014

Abstract

Background and Objective: Environmental pollution modeling is one of the essential requirements in the field of air quality monitoring which with using the output of the model, improvement of future situation can be possible. The existing literature of the modeling of environmental pollution – especially air pollutants- could be divided to two whole categories. First, those researches that in addition of pollutants data, they used some factors such as temperature, wind direction, wind speed and humidity. The second one –which this study belong to- with using time series regression models and by usage of the existing data about each pollutant, the future situation was forecasted.

Method: In this study, we forecast future pollutants (CO, PM₁₀, NO₂, SO₂, O₃, PM_{2.5}) status with ARIMA, ARFIMA and ARIMA-GARCH models with Box-Jenkins approach, then the best model is determined with MSE, RMSE, MAE and MAPE.

Findings: Results indicate that the assumption of existence of long-memory is acceptable but the hypothesis that always ARFIMA models prepare the best forecast is rejected.

Discussion and Conclusion: This study proves the application of econometric models to predict the pollutants state. Based on the high social costs of pollutant emissions, it is recommended that using these models, identify the pollutants affecting the future of the city and reduce the level of their dissemination of efficiency plans.

Keywords: Forecasting, Long-Memory, ARFIMA, GPH test, Modified R/S Test.

1- MSc of Environmental economics, University of Allameh Tabataba’I, Tehran, Iran.* (Corresponding Author)

2- Department of agricultural and environmental Economics, Faculty of economics, University of Allameh Tabataba’I, Tehran, Iran.

مقدمه

مدل‌سازی آلاینده‌های زیست محیطی یکی از نیازهای اساسی در زمینه پایش کیفیت هوا محسوب می‌شود که با بهره‌گیری از نتایج حاصله می‌توان اقدامات پیشگیرانه‌ای جهت بهبود شرایط آتی اتخاذ کرد. در صورتی که بتوان الگویی قابل اطمینان برای پیش‌بینی وضعیت آبی آلاینده‌های موجود در هوای شهر ارابه داد، می‌توان با شناخت دقیق تری که از روند آینده آن‌ها به دست می‌آید، راهکاری مناسب با کارایی بالا برای مقابله با هر یک از این آلاینده‌ها، طراحی کرد و از صرف هزینه‌های سنگین پیش‌رو که در صورت عدم توجه لازم به وضع موجود بروز خواهد کرد، جلوگیری نمود.

آمار و ارقام موجود در زمینه آلودگی هوا نشان می‌دهد شهر تهران در سال ۹۰، ۲۱۸ روز هوای ناسالم داشته، در حالی که این رقم برای سال ۸۹ معادل ۱۰۳ روز و در سال ۸۸، ۴۰ روز بوده است (۱). با توجه به گزارش کیفیت هوای شهر تهران در سال ۹۰، تمامی شواهد حاکی از افزایش سطوح آلاینده‌های هوا و ظهور پیامدهای منفی ناشی از آن می‌باشد. در صورتی که بتوان الگویی قابل اطمینان برای پیش‌بینی وضعیت آبی آلاینده‌های موجود در هوای شهر ارابه داد می‌توان با شناخت دقیق‌تری که از روند آینده آن‌ها به دست می‌آید، راهکاری متناسب برای مقابله با هر یک از این آلاینده‌ها طراحی نمود و از صرف هزینه‌های سنگینی جلوگیری نمود که در صورت عدم توجه لازم به وضع موجود به واسطه پیامدهای منفی آلودگی هوا به شکل هزینه‌های سنگین درمان، خسارات ناشی از تعطیلی اجباری ادارات و مدارس و مؤسسات مالی، اختلالات رادیویی و مخابراتی و ... بروز خواهد کرد.

مطالعات صورت گرفته در زمینه مدل‌سازی و پیش‌بینی آلاینده‌ها را می‌توان مطابق جدول زیر به شکلی جزئی‌تر تقسیم کرد.

مدل‌سازی آلاینده‌های زیست محیطی یکی از نیازهای اساسی در زمینه پایش کیفیت هوا محسوب می‌شود که با بهره‌گیری از نتایج حاصله می‌توان اقدامات پیشگیرانه‌ای جهت بهبود شرایط آبی اتخاذ کرد. در صورتی که بتوان الگویی قابل اطمینان برای پیش‌بینی وضعیت آبی آلاینده‌های موجود در هوای شهر ارابه داد، می‌توان با شناخت دقیق تری که از روند آینده آن‌ها به دست می‌آید، راهکاری مناسب با کارایی بالا برای مقابله با هر یک از این آلاینده‌ها، طراحی کرد و از صرف هزینه‌های سنگین پیش‌رو که در صورت عدم توجه لازم به وضع موجود بروز خواهد کرد، جلوگیری نمود.

آمار و ارقام موجود در زمینه آلودگی هوا نشان می‌دهد شهر تهران در سال ۹۰، ۲۱۸ روز هوای ناسالم داشته، در حالی که این رقم برای سال ۸۹ معادل ۱۰۳ روز و در سال ۸۸، ۴۰ روز بوده است (۱). با توجه به گزارش کیفیت هوای شهر تهران در سال ۹۰، تمامی شواهد حاکی از افزایش سطوح آلاینده‌های هوا و ظهور پیامدهای منفی ناشی از آن می‌باشد. در صورتی که بتوان الگویی قابل اطمینان برای پیش‌بینی وضعیت آبی آلاینده‌های موجود در هوای شهر ارابه داد می‌توان با شناخت دقیق‌تری که از روند آینده آن‌ها به دست می‌آید، راهکاری متناسب برای مقابله با هر یک از این آلاینده‌ها طراحی نمود و از صرف هزینه‌های سنگینی جلوگیری نمود که در صورت عدم توجه لازم به وضع موجود به واسطه پیامدهای منفی آلودگی هوا به شکل هزینه‌های سنگین درمان، خسارات ناشی از تعطیلی اجباری ادارات و مدارس و مؤسسات مالی، اختلالات رادیویی و مخابراتی و ... بروز خواهد کرد.

مطالعات صورت گرفته در زمینه مدل‌سازی و پیش‌بینی آلاینده‌ها را می‌توان مطابق جدول زیر به شکلی جزئی‌تر تقسیم کرد.

جدول ۱- طبقه بندی مطالعات

Table 1. classification of studies

روش	محقق
پیش بینی با ANN	Bodaghpur et al(2011)(۲) – Ibarra et al (2008)(۳)
پیش بینی با رگرسیون خطی	Pires et al (2008)(۴)
پیش بینی با مدل های خود رگرسیونی	Mohd Zamri et al (2009)(۵) – Marzuki et al (2011)(۶) - Kumar& Ridder (2010)(۷)
پیش بینی با مدل های ترکیبی	Diaz-Robles et al (2008)(۸) - Siew et al (2008)(۹)

ARFIMA مطالعه کردند و علاوه بر تشخیص وجود حافظه بلند مدت، روند آتی آن را نیز به خوبی پیش بینی نمودند (۱۵). بایلی به مرور مقالات نوشته شده در زمینه الگوی ARFIMA پرداخت و نمونه ای از کاربرد آن را در مدل سازی حلقه های تشکیل شده در تنه درختان نشان داد. وی ۵۴۰۵ داده را مابین سال های ۳۴۳۶ قبل از میلاد تا سال ۱۹۶۹ میلادی جمع آوری کرد و وجود حافظه بلندمدت را در این داده ها تشخیص داد (۱۶). فستیک و همکاران به تحلیل و بررسی سود بازار سهام با استفاده از رویکرد تفاضل گیری کسری پرداختند. نتایج، وجود ویژگی های حافظه بلندمدت در سود دارایی را اثبات کرد. فرض این مطالعه عدم وجود حافظه بلندمدت در بازار سهام بود که با انجام آزمون هایی از قبیل WOLS(Wavelet modified R/S, KPSS و برآوردگر WOLS(Ordinary Least Squares) مورد ارزیابی قرار گرفت و در نهایت فرض عدم وجود مورد پذیرش واقع نشد (۱۷). عرفانی و صمیمی نیز دو مدل ARFIMA و ARIMA را به لحاظ عملکرد و نتایج حاصل در بازار سهام با یک دیگر مقایسه کرده و نتیجه گرفتند که مدل ARFIMA در این زمینه نتیجه بهتری به دست می دهد (۱۸). از معدود مقالاتی که با استفاده از مدل ARFIMA وضعیت آلاینده های هوا را پیش بینی کردند می توان به موردی اشاره کرد که از دو روش ARIMA و ARFIMA برای پیش بینی API(Air Pollution Index) مربوط به آلاینده های SO_2 ، NO_2 ، PM_{10} ، CO و O_3 استفاده شده، است. در این تحقیق بیان می شود که از

استفاده از الگوی ARIMA برای پیش بینی روند سری های زمانی نسبت به دو مدل دیگر، چه در مطالعات داخلی و چه خارجی از پیشینه بیشتری برخوردار است. از جمله مطالعات داخلی علاوه بر موارد ذکر شده در شکل (۱)، می توان به مطالعه سلطان آبادی (۱۰) که برای پیش بینی کوتاه مدت قیمت نفت خام ایران در کنار مدل ARIMA از روش برنامه ریزی ژنتیک نیز استفاده کرده و به مقایسه نتایج حاصله پرداخت اشاره کرد. در تحقیق دیگری از دو مدل ARIMA و PCR(Principal Component Regression) و حالت ترکیبی این دو برای پیش بینی شاخص AQI(Air Quality Index) در هر فصل سال استفاده شد و نتایج حاصل از حالت ترکیبی با نتایج تک تک مدل ها مورد مقایسه قرار گرفت (۱۱). در مطالعه دیگری با استفاده از مدل ARIMA روند داده های سالانه آلاینده های CO ، PM_{10} ، NO_x ، SO_2 ، H_2S ، TSP و PB در بازه زمانی ۲۰۰۴-۱۹۹۲، الگوسازی شد و مورد پیش بینی قرار گرفت (۱۲).

روشی که نسبت به ARIMA جدیدتر محسوب می شود و در سالیان اخیر مورد استفاده قرار گرفته، مدل ARFIMA است. از این مدل برای سری های زمانی دارای ویژگی حافظه بلندمدت استفاده می شود. ابتدا هورست در مقاله ای با عنوان ظرفیت بلندمدت مخازن، به وجود حافظه بلندمدت اشاره کرد (۱۳). سپس والیس و مندلبرت وضعیت حافظه بلندمدت را در مسایل مربوط به هیدرولوژی مورد بررسی قرار دادند (۱۴). بایلی و بلرسلو روند مربوط به نرخ های برابری ارز را با استفاده از مدل

ادعا کرد که مطالعه حاضر از اولین پژوهش‌های منسجم در این زمینه محسوب می‌شود.

۱- روش شناسی

اولین نکته‌ای که باید بررسی شود این است که آیا داده‌های مربوط به هر یک از آلاینده‌ها دارای حافظه بلندمدت است یا خیر؟ نکته مهمی که در قالب این سؤال مطرح می‌شود این است که اگر پاسخ منفی باشد، استفاده از الگوی ARFIMA توجیه‌پذیر نیست و در این صورت باید بهترین مدل پیش‌بینی کننده را پس از دریافت نتایج از میان دو مدل ARIMA و ARIMA-GARCH، انتخاب کرد. هم چنین بررسی می‌شود که آیا پیش‌بینی حاصل از رویکرد ARFIMA دارای پسماند کوچکتری نسبت به دو مدل دیگر است یا خیر. در زمینه الگوی ARIMA-GARCH نیز وجود اثرات واریانس ناهمسانی در داده‌ها مورد سؤال قرار می‌گیرد و پاسخ مثبت به این پرسش نیز استفاده از مدل ARIMA-GARCH را توجیه می‌کند.

با توجه به پیشینه تحقیقات صورت گرفته با مدل ARFIMA در زمینه الگوسازی برای وضعیت آبی آلاینده‌ها، می‌توان فرض کرد که اولاً داده‌های مربوط به غلظت آلاینده‌ها دارای حافظه بلندمدت است و ثانياً هم چنان که نتایج مطالعات پیشین نشان از بالاتر بودن دقت مدل ARFIMA نسبت به دیگر مدل‌ها دارد، فرض دوم این است که دقت مدلی که حافظه بلندمدت را در نظر می‌گیرد بیشتر از مدل‌هایی است که توجهی به این ویژگی ندارند.

برای پاسخ‌گویی به سؤالات فوق مسیر خاصی دنبال شده که در شکل زیر ملاحظه می‌شود. بر این اساس پیش از تخمین الگو (تمامی محاسبات با استفاده از نرم افزار STATA12 صورت گرفته است) باید آزمون مانایی در مورد تک تک سری‌ها صورت گیرد. اگر متغیری مانا باشد، شوک‌های وارده به آن حتی اگر در کوتاه‌مدت سطح متغیر را تغییر دهند در بلندمدت بر متغیر مذکور کاملاً بی‌اثرند و این متغیر حرکت خود را در حوالی روند ثابت بلندمدت ادامه خواهد داد. در این مقاله با دو نوع شوک مواجه هستیم، شوک‌های طبیعی از قبیل وارونگی هوا، بارندگی

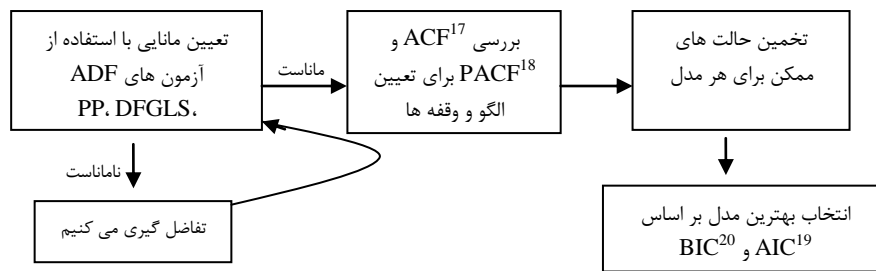
پیش‌بینی روند این آلاینده‌ها می‌توان برای مدیریت و سیاست‌گذاری در زمینه کنترل و بهبود شرایط بهره برد. در این مطالعه از داده‌های ۷۰ ماه استفاده شد و برای تعیین بهترین مدل شاخص‌های MAE، RMSE و MAPE، مد نظر قرار گرفت. نتایج مطمئن‌تر بودن پیش‌بینی مدل ARFIMA را اثبات کرد (۹).

در زمینه مطالعات صورت گرفته‌ای که از مدل GARCH برای الگوسازی نوسانات استفاده شده می‌توان به برسلو که به معرفی شکل تعمیم یافته ARCH پرداخت، اشاره کرد (۱۹). از مدل‌های ARCH (AutoRegressive Conditional Heteroskedasticity) به طور گسترده در الگوسازی نوسانات بازار سهام و ریسک مالی و بازارهای نفت و فرآورده‌های آن استفاده شده است. از جمله این مطالعات می‌توان به مصدق اشاره کرد که نوسانات معاملات آبی نفت خام سبک در نایمکس (New York Mercantile Exchange) را برای دوره ۱۹۹۸-۲۰۰۹ با استفاده از خانواده مدل‌های ARCH، الگو سازی کرد (۲۰). عالم شیفر با کمک مدل ARIMA-GARCH به توضیح نوسانات لگاریتم بازده قیمتی محصولات کشاورزی پرداخت. به دلیل نوسانی بودن تغییرات واریانس، مدل‌های ARCH و GARCH به عنوان مناسب‌ترین الگو در نظر گرفته شد (۲۱).

رویکرد این مقاله را می‌توان تا حدودی در زمره دسته سوم - توجه به جدول ۱ مطالعات صورت گرفته قرار داد. در این مقاله از سه روش ARIMA، ARFIMA و ARIMA-GARCH برای پیش‌بینی وضعیت آلاینده‌های PM_{10} ، CO ، NO_2 ، SO_2 و O_3 استفاده شده است. در زمینه استفاده از الگوی ARFIMA، با مروری که بر مطالعات گذشته انجام گرفت می‌توان ادعا کرد که تاکنون از این مدل در ایران برای پیش‌بینی آلاینده‌ها استفاده نشده - در سایر موضوعات نیز به طور انگشت شمار به کار گرفته شده - و این مطالعه اولین پژوهش در زمینه استفاده از مدل خودرگرسیون هم انباشت کسری (ARFIMA) برای پیش‌بینی روند آلاینده‌های مذکور است. در مورد دو روش دیگر هم می‌توان

DFGLS(Dickey-Fuller Generalized Least Squares) ،
 Squares) و PP(Phillips-Perron) صورت گرفت. در
 آزمون ADF از روش هلدن و پرمن به روایت نوفرستی استفاده
 شده و هم چنین با بهره گیری از متدولوژی باکس-جنکینز
 بهترین الگو در هر یک از مدل ها مورد شناسایی قرار
 گرفت(۲۲).

اندک، پایین بودن سرعت وزش باد و به نوعی پایداری جوی در
 بازه زمانی بلندمدت و نیز شوک های غیرطبیعی مانند تعطیلات
 اضطراری به علت آلودگی هوا که باعث کاسته شدن از حجم
 آلاینده ها می شود. مانایی یا نامانایی در این وضعیت می تواند بر
 اثرات بلندمدت این قبیل شوک ها تأثیر گذار باشد.
 تعیین وضعیت مانایی سری های زمانی با سه آزمون
 ADF(Augmented Dickey-Fuller test)



شکل ۱- مسیر انتخاب مدل

Figure 1. model selection procedure

معرفی مدل ها

مدل ARIMA و ARFIMA

مدل ARFIMA توسط گرنجر و ژوکس (۲۳) و هاسکینگ
 (۲۴) معرفی شد. بایلی (۱۶) و بایلی و چونگ (۲۵) مقدماتی را
 در جهت معرفی کاربردهای این مدل فراهم آوردند. اساس مدل
 $ARFIMA(p,d,q)$ همانند مدل $ARIMA(p,d,q)$
 می باشد، با این تفاوت که درجه هم انباشتگی (d) در این فرایند
 عدد صحیح نیست و اصولاً مقادیری بین ۱ و -۱ را اختیار
 می کند. اگر $X_t (t=1, \dots, T)$ از یک فرایند $ARMA(p,q)$
 تبعیت کند، می توان حالت عمومی این فرایند را به صورت زیر
 نوشت:

$$X_t = \phi_1 X_{t-1} + \phi_2 X_{t-2} + \dots + \phi_p X_{t-p} + \varepsilon_t - \psi_1 \varepsilon_{t-1} - \psi_2 \varepsilon_{t-2} - \dots - \psi_q \varepsilon_{t-q} \quad (1)$$

$$\nabla X_t = X_t - X_{t-1} = (1-L)X_t \quad (3)$$

در این صورت رابطه (۲) به صورت زیر نوشته می شود:

$$\phi(L)\nabla^d X_t = \psi(L)\varepsilon_t \quad (4)$$

برای مدل ARFIMA پیش از انجام مراحل فوق باید وجود
 حافظه بلندمدت با آزمون هایی که در ادامه معرفی می شود مورد
 ارزیابی قرار گیرد. هم چنین در مورد مدل ARIMA-
 GARCH، لازم است پس از انجام آزمونی که وجود اثرات
 ناهمسانی واریانس را اثبات می کند، روند گام به گام فوق انجام
 پذیرد. پس از برازش الگوها با مقایسه نتایج و تعیین بهترین
 مدل از میان الگوهای مطرح، بهترین مدلی که دارای کمترین
 خطا در پیش بینی باشد معرفی شده و مقادیر پیش بینی در کنار
 مقادیر واقعی نشان داده خواهد شد.

فرآیند فوق را می توان به صورت زیر بازنویسی کرد :

$$\phi(L)X_t = \psi(L)\varepsilon_t, \quad \varepsilon_t \sim (0, \sigma_\varepsilon^2) \quad (2)$$

حال اگر فرایند X_t ، فرآیندی مانا نبوده و دارای ریشه واحد
 باشد با تفاضل گیری مانا خواهد شد.

در برآورد مدل ARIMA-GARCH این گونه عمل می شود که با فرض وجود اثرات ARCH (یا همان واریانس ناهمسانی) که با آزمون LM شناسایی شده و با در نظر گرفتن اینکه جمله اختلال در مدل ARIMA به صورت زیر تعریف شود:

$$U_t = \sigma_t \varepsilon_t \quad (۸)$$

در معادله (۸) مقدار U_t مشخص می شود که جمله اختلال فرآیند ARIMA را تشکیل می دهد. بر این مبنا مدل ARIMA-GARCH برآورد خواهد شد.

نتایج و بحث

توصیف داده ها

داده های مورد استفاده در این مقاله شامل آلاینده های CO، PM₁₀، NO₂، SO₂، O₃ و PM_{2.5} در دوره زمانی ۹۰/۱/۱۳ تا ۹۱/۶/۳۰ (شامل ۵۳۶ داده روزانه) می باشد که اطلاعات مربوط به آن از تجمیع و شاخص سازی داده های حاصل از ایستگاه های سنجش آلاینده هوا در شهر تهران توسط سازمان حفاظت محیط زیست تهران حاصل شده است.

مانایی هر یک از سری ها با استفاده از آزمون های ADF، DFGLS و PP مورد بررسی قرار گرفت. نتایج نشان داد که تمامی آلاینده ها در سطح مانا هستند.

برای الگوسازی سری های زمانی و برازش الگوی ARMA از رویکرد باکس-جنکینز استفاده شد که شامل سه مرحله تشخیص، تخمین و آزمون دقت است. بر این اساس در مرحله تشخیص توابع خودهمبستگی (ACF) و خود همبستگی جزئی (PACF) مورد بررسی قرار گرفتند و برای انتخاب وقفه بهینه، از شاخص های آکاییک (AIC) و بیزین (BIC) استفاده شد.

برازش الگوی ARMA

مدل های زیر بهترین برازش ممکن را بر اساس رویکرد باکس - جنکینز ارائه می کنند:

در اینجا d مرتبه تفاضل گیری لازم برای رسیدن به فرآیندی مانا است.

در مدل ARFIMA(p,d,q)، مقدار d عددی حقیقی در بازه ۱ تا -۱ است. در این حالت عبارت ∇X_t در معادله (۴) را با یک فرآیند خود توضیح از درجه بی نهایت جایگزین می کنیم.

عملگر تفاضل گیری کسری $\nabla^d X_t$ عبارت خواهد بود از:

$$\nabla^d X_t = (1-L)^d X_t = \sum_{k=0}^{\infty} (-1)^k \binom{d}{k} L^k X_t \quad (۵)$$

با جای گذاری عبارت (۵) در معادله (۴)، مدل ARFIMA به صورت زیر تبدیل می شود:

$$\phi(L)(1-L)^d X_t = \psi(L)\varepsilon_t \quad (۶)$$

در این حالت می بایست d ، شرط زیر را داشته باشد:

$$|d| < 0.5$$

برای تخمین مقدار d در این فرآیند سری زمانی، تا کنون روش های متعددی معرفی شده اند که تحلیل R/S و برآوردگرهای GPH از جمله این روش ها هستند و توسط اقتصاددانانی از قبیل هورست (۱۳)، واروتسوس و دیویدف (۲۶) و جوک و همکاران (۲۷) معرفی شدند.

مدل ARIMA-GARCH

مدل GARCH توسط بلسلو (۱۹) معرفی شد. اگر برای یک مدل ARCH فرآیندی شبیه به مدل ARMA در نظر بگیریم، مدل GARCH پدید خواهد آمد. در این حالت P درجه جمله GARCH (یعنی σ^2) و q درجه جمله ARCH (یعنی ε^2) خواهد بود. فرم کلی مدل GARCH(p,q) به صورت زیر می باشد:

$$\sigma_t^2 = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \dots + \alpha_q \varepsilon_{t-q}^2 + \beta_1 \sigma_{t-1}^2 + \dots + \beta_p \sigma_{t-p}^2 \quad (۷)$$

$$CO = 37/0.82 + 0.231(CO_{t-2}) + 0.141(CO_{t-3}) + 0.091(CO_{t-5}) + 0.094(CO_{t-8}) + 0.172(CO_{t-35}) + 0.112(CO_{t-41}) - 0.173(CO_{t-45}) + 0.082(CO_{t-51}) + 0.061(CO_{t-66}) + 0.406(U_{t-1}) - 0.181(U_{t-3}) + 0.073(U_{t-34})$$

AIC : ۳۵۶۳/۳۷۲ , BIC: ۳۶۲۳/۴۰۲ , loglikelihood = -۱۷۶۷/۶۸۶

$$NO_2 = 66/1.42 - 0.029(trend) + 0.62(NO_{2(t-1)}) + 0.346(NO_{2(t-3)}) - 0.606(NO_{2(t-4)}) + 0.671(NO_{2(t-5)}) - 0.17(NO_{2(t-6)}) - 0.521(U_{t-2}) + 0.222(U_{t-4}) - 0.468(U_{t-5}) + 0.072(U_{t-29})$$

AIC : ۳۷۵۰/۱۹ , BIC : ۳۸۰۱/۶۴۴ , loglikelihood = -۱۸۶۳/۰۹۵

$$O_3 = 38/6.93 + 0.381(O_{3(t-1)}) + 0.095(O_{3(t-28)}) - 0.338(U_{t-1}) + 0.184(U_{t-17}) + 0.143(U_{t-26}) + 0.096(U_{t-42})$$

AIC : ۳۵۸۶/۱۹۵ , BIC : ۳۶۲۰/۴۹۸ , loglikelihood = -۱۷۸۵/۰۹۸

$$PM_{2.5} = 112/0.18 - 0.33(trend) + 0.636(PM_{2.5(t-1)}) + 0.086(PM_{2.5(t-21)}) - 0.083(PM_{2.5(t-28)}) - 0.079(PM_{2.5(t-41)}) + 0.059(PM_{2.5(t-51)})$$

AIC : ۴۶۴۵/۷۲۹ , BIC : ۴۶۸۰ , loglikelihood = -۲۳۱۴/۸۶۵

$$PM_{10} = 66/6.05 + 0.45(PM_{10(t-1)}) + 0.169(PM_{10(t-51)}) + 0.075(PM_{10(t-60)}) + 0.141(U_{t-9})$$

AIC : ۴۷۰۲/۴۹۸ , BIC : ۴۷۲۸/۲۲۶ , loglikelihood = -۲۳۴۵/۲۴۹

$$SO_2 = 33/7.62 + 0.279(SO_{2(t-3)}) + 0.173(SO_{2(t-7)}) + 0.15(SO_{2(t-11)}) + 0.133(SO_{2(t-17)}) + 0.458(U_{t-1}) + 0.3(U_{t-2})$$

AIC : ۲۹۹۹/۳۵۱ , BIC : ۳۰۳۳/۶۵۳ , loglikelihood = -۱۴۹۱/۶۷۵

برازش الگوی ARFIMA

نتیجه داد. البته آماره های R/S, GPH, KPSS اصلاح شده و نیز وجود حافظه بلندمدت را برای این آلاینده تأیید می کنند. می توان از آزمون KPSS برای شناسایی حافظه بلندمدت و کوتاه مدت استفاده کرد (۲۹). این آزمون غالباً برای بررسی مانایی استفاده می شود، اما در این مورد فرضیه صفر عبارت است از وجود حافظه کوتاه مدت که در مقابل فرضیه حافظه بلندمدت قرار می گیرد. تعداد وقفه های بهینه در این آزمون با استفاده از معیار شوارتز تعیین شد.

در این مقاله از چهار روش مشاهده، آماره R/S, GPH اصلاح شده و آماره KPSS برای آزمون وجود حافظه بلندمدت در هر سری از داده ها استفاده شد. بر مبنای روش مشاهده، نمودار تابع خود همبستگی (ACF) باید دارای روندی نزولی به شکل هذلولی باشد. این ویژگی توسط هورست مطرح گردید (۱۳)(۲۸) و در فرآیندهای سری زمانی مختلف مورد استفاده قرار گرفت. استفاده از این روش فقط در مورد آلاینده O₃

جدول ۲- نتایج آزمون وجود حافظه بلند مدت

Table 2. results of the long memory test

آلاینده	GPH	R/S	KPSS
CO	0.50 , 0.34= Power	** 2/0.1	$\hat{\eta}_u = 0.706^{**}$ $\hat{\eta}_r = 0.726^*$
NO ₂	0.65 , 0.34= Power	* 2/27	$\hat{\eta}_u = 3/27^*$ $\hat{\eta}_r = 0.357^*$

O ₃	۰/۷۵ ، ۰/۴۵= Power	*** ۱/۸۵	$\hat{\eta}_u = ۱/۲۷^*$ $\hat{\eta}_r = ۱/۲^*$
PM _{2.5}	۰/۵۵ ، ۰/۳۲= Power	۱/۱۸	$\hat{\eta}_u = ۰/۸۱۸^*$ $\hat{\eta}_r = ۰/۰۸$
PM ₁₀	۰/۵۰ ، ۰/۱۵= Power	۱/۴۸	$\hat{\eta}_u = ۰/۵۱۲^{***}$ $\hat{\eta}_r = ۰/۱۹۴^{**}$
SO ₂	۰/۶۵ ، ۰/۳۷= Power	* ۲/۲۶	$\hat{\eta}_u = ۰/۷۰۸^{**}$ $\hat{\eta}_r = ۰/۷۲۶^*$

منبع: محاسبات محقق

• $\hat{\eta}_r$: روند مانا (trend stationarity)، $\hat{\eta}_u$: سطح

مانا (level stationarity)

• فرض صفر در آزمون kpss وجود حافظه کوتاه مدت و

فرض مقابل وجود حافظه بلند مدت در سری است. در

*۹۹٪ اطمینان، **۵/۹۷٪ اطمینان، ***۹۵٪ اطمینان،

اطمینان ۹۰٪، H₀ رد می‌شود. در غیر این صورت

فرض صفر را نمی‌توان رد کرد.

با این فرض که فرم کلی مدل ARFIMA عبارت است از:

• به طور پیش فرض، تعداد وقفه های آزمون R/S در

نرم افزار STATA 12 با استفاده از معیار اطلاعات

اندروز محاسبه می‌شود. فرضیه صفر در این آزمون

عدم وجود رابطه بلندمدت است. در مورد آزمون R/S

H₀ با *۹۹٪ اطمینان، **۹۵٪ اطمینان، ***۹۰٪

اطمینان، رد می‌شود. در غیر این صورت فرض صفر را

نمی‌توان رد کرد.

• برای وجود حافظه بلندمدت باید آماره GPH مابین

۵/۰ تا ۵/۰- قرار گیرد. البته این مقدار می‌تواند به

ازای هر توانی در بازه (۰/۴ تا ۰/۷۵) قرار گیرد.

$$(1 - \phi_1 L - \phi_2 L^2 - \dots - \phi_p L^p)(1 - L)^d x_t = (1 + \theta_1 L + \theta_2 L^2 + \dots + \theta_q L^q) \varepsilon_t$$

مدل های برازش شده ARFIMA برای آلاینده های مورد نظر

عبارت خواهند بود از:

$$(1 - 0.154 L^{35} - 0.127 L^{41} + 0.151 L^{45} - 0.078 L^{51})(CO) = \varepsilon_t + 37/224 \quad d=0.35$$

$$AIC: 3571/921, \quad BIC: 3601/936, \quad \loglikelihood = -1778/96.5$$

$$(1 - 0.399 L + 0.104 L^4 + 0.116 L^6 - 0.123 L^{29} - 0.079 L^{56} + 0.082 L^{68})(NO_2) = \varepsilon_t + 66/17 - 0.29(\text{trend})$$

$$d=0.19 \quad AIC: 3747/985, \quad BIC: 3790/863, \quad \loglikelihood = -1863/9924$$

$$(1 - 0.583 L^3 - 0.099 L^{17} - 0.122 L^{28} - 0.101 L^{33} + 0.086 L^{45})(O_3) = (1 - 0.615 L^7) \varepsilon_t + 39/445 \quad d=0.46$$

$$AIC: 3584/183, \quad BIC: 3622/774, \quad \loglikelihood = -1783/0.916$$

$$(1 - 0.446 L - 0.093 L^{11} + 0.105 L^{28} + 0.08 L^{41} + 0.102 L^{44} - 0.079 L^{51})(PM_{2.5}) = \varepsilon_t + 111/555 - 0.32(\text{trend})$$

$$d=0.19 \quad AIC: 4640/136, \quad BIC: 4683/14, \quad \loglikelihood = -2310/0.678$$

$$(1 - 0.304 L - 0.137 L^9 + 0.107 L^{28} - 0.208 L^{51})(PM_{10}) = (1 - 0.118 L^{42}) \varepsilon_t + 66/534 \quad d=0.13$$

AIC : ۶۴۹۳/۷۲ , BIC : ۴۷۳۰/۰۲۳ , loglikelihood = -۲۳۳۹/۸۶۰

$$(1 + 0.106L - 0.104L^2)(SO_2) = \varepsilon_t + 33/573 \quad d=0.37$$

AIC : ۳۱۸۵/۱۵۴ , BIC : ۳۲۰۶/۵۹۴ , loglikelihood = -۱۵۸۷/۵۷۷۲

دارای میانگین و واریانس ثابت هستند. طبق جدول زیر آماره مربوط به هر دو آزمون، عدم تفاوت ضرایب هم بستگی پسماندها تا وقفه صدم با صفر را تأیید می کند.

در پایان عدم وجود همبستگی بین پسماندهای هر مدل مورد بررسی قرار می گیرد. برای این منظور از دو آزمون Q و بارتلت استفاده شد. در آزمون بارتلت فرضیه صفر این است که داده ها از یک فرآیند نوفه سفید بدون همبستگی، ایجاد شده اند و

جدول ۳- آماره های Q و بارتلت

Table 3- Q and Bartlet stats

آلاینده	مدل	آماره Q	Prob>chi2(100)	آماره بارتلت	Prob>B
CO	ARMA	۵۷/۵۰۵	۰/۹۹	۰/۳۴۴	۰/۹۹
	ARFIMA	۵۸/۶۳۶	۰/۹۹	۰/۸۳۴	۰/۴۸
O ₃	ARMA	۸۸/۲۹۶	۰/۷۹	۰/۵۱۹	۰/۹۴
	ARFIMA	۸۰/۴۲۶	۰/۹۲	۰/۵۸۹	۰/۸۷
NO ₂	ARMA	۶۸/۴۹۲	۰/۹۹	۰/۶۳۸	۰/۸۰
	ARFIMA	۷۰/۶۴۰	۰/۹۸	۰/۹۸۳	۰/۲۸
PM _{2.5}	ARMA	۹۰/۴۸۰	۰/۷۴	۰/۷۵۶	۰/۶۱
	ARFIMA	۷۲/۰۲۰	۰/۹۸	۰/۶۴۷	۰/۷۹
PM ₁₀	ARMA	۶۳/۹۰۸	۰/۹۹	۰/۶۳۲	۰/۸۱
	ARFIMA	۶۴/۳۱۰	۰/۹۹	۰/۶۲۴	۰/۸۳
SO ₂	ARMA	۹۷/۹۰۰	۰/۵۴	۰/۳۹۴	۰/۹۹
	ARFIMA	۱۰۶/۶۰۸	۱/۳۰	۰/۶۴۵	۰/۷۹

منبع: محاسبات محقق

برازش الگوی ARMA-GARCH

اثرات ARCH در تمام آلاینده ها تأیید می شود و در نتیجه استفاده از مدل های ARIMA-GARCH می تواند توجیه پذیر باشد. الگوهای برازش شده با در نظر گرفتن رابطه ۳۵ به صورت زیر خواهند بود:

برای به کارگیری مدل GARCH با استفاده از آزمون LM وجود اثرات ARCH در پسماند مدل ها بررسی شد. ابتدا مقدار هر آلاینده بر عرض از مبدأ رگرس و سپس پسماندهای آن محاسبه شد. آزمون LM روی پسماندها می تواند وجود یا عدم وجود اثرات ARCH را نشان دهد. بر این اساس وجود

$$(1 - 0.263L - 0.126L^2 - 0.123L^3 - 0.103L^4 - 0.142L^5 + 0.151L^6)CO = 36/933 + (1 + 0.441L + 0.075L^2)U_t$$

$$\sigma^2 = 21/0.36 + 0.138U_{t-1}^2 + 0.381\sigma_{t-1}^2$$

$$AIC: 3481/284, \quad BIC: 3532/422, \quad \text{loglikelihood} = -1728/642$$

$$(1 - 0.139L - 0.157L^2 - 0.136L^3)NO_2 = 65/48 - 0.26(\text{trend}) + (1 + 0.587L + 0.324L^2 + 0.179L^3 + 0.109L^4)U_t$$

$$\sigma^2 = 3/0.52 + 0.58U_{t-1}^2 + 0.89\sigma_{t-1}^2$$

$$AIC: 3715/269, \quad BIC: 3766/723, \quad \text{loglikelihood} = -1845/634$$

$$(1 - 0.477L - 0.174L^2 - 0.096L^3 - 0.105L^4 + 0.11L^5)O_3 = 37/415 + (1 + 0.519L + 0.417L^2 + 0.315L^3 + 0.234L^4 - 0.243L^5 + 0.099L^6)U_t$$

$$\sigma^2 = 35/53 + 0.186U_{t-1}^2$$

$$AIC: 3566/902, \quad BIC: 3626/932, \quad \text{loglikelihood} = -1769/45$$

$$(1 - 0.618L - 0.11L^2 + 0.122L^3 + 0.064L^4 + 0.057L^5 - 0.052L^6)PM_{2.5} = 115/809 - 0.45(\text{trend}) + U_t$$

$$\sigma^2 = 68/383 + 0.281U_{t-1}^2 + 0.535\sigma_{t-2}^2$$

$$AIC: 4609/344, \quad BIC: 4656/51, \quad \text{loglikelihood} = 4656/51$$

$$(1 - 0.642L)PM_{10} = 64/0.84 + (1 - 0.094L - 0.19L^2 + 0.148L^3 + 0.214L^4 + 0.08L^5 - 0.066L^6)U_t$$

$$\sigma^2 = 259/358 + 0.326U_{t-1}^2$$

$$AIC: 4636/412, \quad BIC: 4679/291, \quad \text{loglikelihood} = -2308/206$$

$$(1 - 0.303L - 0.082L^2 - 0.111L^3 - 0.131L^4 - 0.122L^5)SO_2 = 33/543 + (1 + 0.421L + 0.286L^2)U_t$$

$$\sigma^2 = 0.23 + 0.065U_{t-1}^2 + 0.926\sigma_{t-1}^2$$

$$AIC: 2938/201, \quad BIC: 2985/367, \quad \text{loglikelihood} = -1458/1$$

شرایط اقلیمی و جغرافیای شهر تهران، انتظار انباشت آلاینده ها دور از ذهن نیست. با این اوصاف غلظت هر آلاینده در هر لحظه از زمان می تواند از غلظت آن آلاینده در هفته های گذشته تبعیت نماید.

نتایج آزمون های Q و بارتلت از صفر بودن ضرایب خود همبستگی در سری مربوط به پسماند مدل های فوق حکایت دارد.

در مورد هر سه مدل ارائه شده، در نظر گرفتن حجم داده ها - ۵۳۶ داده برای هر آلاینده - و ویژگی انباشت آلاینده ها در محیط زیست و به خصوص ساختار اقلیمی تهران، استفاده از وقفه های طولانی را توجیه می کند. در مطالعات سری زمانی با بازه های زمانی بزرگ استفاده از وقفه های طولانی همان طور که در دیگر مطالعات آمده، توجیه پذیر است (۱۲)(۳۰). از سوی دیگر برخی آلاینده ها مانند ذرات معلق، در شرایط پایدار جوی، قابلیت ماندگاری بالایی در محیط دارند که با لحاظ کردن

جدول ۴- آماره های Q و بارتلت

Table 4. Q and Bartlet stats

آلاینده	مقدار آماره آزمون Q	Prob>chi2(100)	مقدار آماره آزمون بارتلت	Prob>B
CO	۷۴/۰۹۷	۰/۹۷	۰/۵۲۸	۰/۹۴
NO ₂	۷۲/۷۳۳	۰/۹۸	۰/۳۸۶	۰/۹۹
O ₃	۸۸/۳۹۱	۰/۷۹	۱/۰۴	۰/۲۲۸
PM _{2.5}	۸۴/۴۸	۰/۸۶۷	۰/۸۵۷	۰/۴۵۴
PM ₁₀	۶۳/۴۱۴	۰/۹۹	۰/۹۵۸	۰/۳۱۷
SO ₂	۱۱۱/۵۱	۰/۲۰	۰/۸۸	۰/۴۱۲

منبع: محاسبات محقق

نتایج حاصل از هر سه مدل با استفاده از چهار معیار MSE

، RMSE، APE و MAPE، به منظور تعیین مدلی با

کمترین خطا در پیش بینی با یک دیگر مقایسه می شوند.

جدول ۵- مقایسه دقت مدل ها

Table 5. comparison of models accuracy

آلاینده	مدل	MSE	RMSE	MAE	MAPE
CO	ARMA	۴۱/۷۵۵*	۶/۴۶۱*	۵/۱۹۲*	۱۴/۵۸۳*
	ARFIMA	۴۲/۷۸۹	۶/۵۴۱	۵/۲۵۱	۱۴/۷۲۳
	ARMA-GARCH	۴۳/۱۸۱	۶/۵۷۱	۵/۲۹۰	۱۴/۸۲۱
O ₃	ARMA	۴۴/۵۶۲	۶/۶۷۵	۵/۰۱۹	۱۵/۰۷۲
	ARFIMA	۴۴/۱۳۵	۶/۶۴۳	۴/۹۶۲*	۱۴/۷۹۸*
	ARMA-GARCH	۴۳/۶۰۷*	۶/۶۰۳*	۴/۹۸۵	۱۴/۹۲۵
NO ₂	ARMA	۵۹/۵۷۹*	۷/۷۱۸*	۵/۶۸۵	۱۰/۱۷۶
	ARFIMA	۵۹/۷۶۵	۷/۷۳۰	۵/۶۳۷*	۱۰/۱۰۲*
	ARMA-GARCH	۶۱/۱۰۶	۷/۸۱۷	۵/۶۶۷	۱۰/۱۸۱
PM _{2.5}	ARMA	۳۲۱/۸۴۶	۱۷/۹۴۰	۱۲/۹۵۲	۱۳/۳۸۵
	ARFIMA	۳۱۶/۶۹۹*	۱۷/۷۹۶*	۱۲/۸۱۰*	۱۳/۱۸۳*
	ARMA-GARCH	۳۱۹/۷۷۹	۱۷/۸۸۲	۱۲/۹۹۸	۱۳/۵۲۰
PM ₁₀	ARMA	۳۶۱/۶۰۹	۱۹/۰۱۶	۱۰/۷۰۹	۱۵/۹۸۱

	ARFIMA	۳۵۷/۲۶۱*	۱۸/۹۰۱*	۱۰/۷۸۰	۱۶/۱۲۴
	ARMA-GARCH	۳۵۹/۵۱۴	۱۸/۹۶۰	۱۰/۵۵۸*	۱۵/۴۵۵*
SO ₂	ARMA	۱۴/۹۷۱*	۳/۸۶۹*	۲/۸۳۳	۸/۳۰۸
	ARFIMA	۱۴/۹۸۷	۳/۸۷۱	۲/۸۳۳*	۸/۲۴۷*
	ARMA-GARCH	۱۵/۲۷۱	۳/۹۰۷	۲/۸۵۲	۸/۳۵۷

منبع: محاسبات محقق

علامت * نشانگر مقدار کمینه در میان مقادیر مربوط به هر مدل است.

اکنون براساس الگوی برآورد شده برای هر آلاینده، پیش‌بینی روز آتی انجام می‌شود که نتیجه آن در کنار مقادیر واقعی در جدول زیر ارائه شده است.

مقادیر کمینه برای هر معیار به صورت ستاره دار مشخص شده اند. بر اساس جدول فوق، در مورد آلاینده‌های CO و PM_{2.5} تمامی معیارهای خطای پیش‌بینی فقط برای یک الگو حداقل شده‌اند. اما در مورد سایر آلاینده‌ها، کمینه معیارهای دقت در پیش‌بینی، الگوی واحدی را پیشنهاد نمی‌کنند.

جدول ۶- مقادیر پیش‌بینی شده در هر مدل

Table 6. predictions of the models

آلاینده	CO			NO ₂			O ₃			PM _{2.5}			PM ₁₀			SO ₂		
مدل	ARMA	ARMA-GARCH	ARFIMA	ARMA	ARMA-GARCH	ARFIMA	ARMA	ARMA-GARCH	ARFIMA	ARMA	ARMA-GARCH	ARFIMA	ARMA	ARMA-GARCH	ARFIMA	ARMA	ARMA-GARCH	ARFIMA
پیش‌بینی	۴۰/۹	۴۴/۳	۴۲/۳	۵۲/۵	۵۵/۱	۵۴/۴	۸۴/۴	۴۲/۸	۴۲/۳	۱۰/۳/۵	۱۰/۲/۳	۱۰/۳/۷	۶۶/۲	۶۶/۴	۶۸/۳	۳۲/۷	۳۲/۵	۳۲/۴
مقدار واقعی	۲۵			۴۸			۴۴			۹۵			۶۹			۳۴		

منبع: محاسبات محقق

نتیجه‌گیری و پیشنهادات

پیش‌بینی‌های بهتری از وضعیت آتی به دست آید، اما مباحث مربوط به این واقعیت خارجی که برخی از آلاینده‌ها در هوا و محیط زیست دارای ماندگاری طولانی تری است، با کاربرد مفهوم حافظه بلند مدت به طور دقیق تری می‌تواند مورد بررسی قرار گیرد.

در این مقاله از مدل ARFIMA برای اولین بار در زمینه پیش‌بینی آلاینده‌های هوا در ایران استفاده شد و وجود حافظه بلندمدت در سری‌های مربوط به آلاینده‌های مختلف اثبات گردید و در نتیجه فرضیه اول تحقیق اثبات شد. البته وجود حافظه بلند مدت صرفاً یک بحث فنی است که باعث می‌شود

2. Bodaghpour, S., Charkestanti, A., 2011. Prediction of gas pollutant's concentration by means of artificial neural network in Tehran urban air. [Journal of environmental science and technology](#) Volume 13 , Number 1 (48); 1-10.
3. Ibarra-Berastegi, G., Elias, A., Barona, A., Saenz, J., Ezcurra, A., Diaz de Argandona, J., 2008. From diagnosis to prognosis for forecasting air pollution using neural networks: Air pollution monitoring in Bilbao. *Environmental modeling and software*, 23, 622-637.
4. Pires, J. C. M., Martins, F. G., Sousa, S. I. V., Alvim-Ferraz, M. C. M., Pereira, M. C., 2008. Prediction of the daily mean PM₁₀ concentrations using linear models. *American Journal of Environmental Sciences*, 4(5), 445-453.
5. Ibrahim, M. Z., Zailan, R., Marzuki, I., Lola, M. S., 2009. Forecasting and time series analysis of air pollutants in several area of Malaysia. *American Journal of Environmental Sciences*, 5(5), 625-632.
6. Marzuki, I., Ibrahim, M. Z., Ibrahim, A., Ahmad Makmon, A., 2011. Time series analysis of surface ozone monitoring records in Kemaman, Malaysia. *Sains Malaysiana*, 40(5), 411-417.
7. Kumar, U., Ridder, K. D., 2010. GARCH modelling in association with FFT-ARIMA to forecast ozone episodes. *Atmospheric Environment*, 44, 4252-4265.
8. Diaz-Robles, L. A., Juan, C. O., Joshua, S. Fu., Gregory, D. R., Judith, C. C., Watson, J. G., Moncada-Herrea, J. A., 2008. A hybrid ARIMA and

شکل (۶) نشان می دهد که برای همه آلاینده ها بر خلاف فرضیه دوم تحقیق، نمی توان نتایج مدل ARFIMA را بهتر از دو مدل دیگر دانست، زیرا شواهد نشان می دهد که در مورد آلاینده های CO، NO₂ و PM_{2.5}، پیش بینی یک روز توسط دو مدل دیگر دارای خطای کمتری نسبت به پیش بینی مدل ARFIMA است. بنابراین فرض دوم این تحقیق مبنی بر بالاتر بودن دقت مدل ARFIMA نسبت به دیگر مدل های خود رگرسیون در شرایط وجود حافظه بلند مدت در سری زمانی رد می شود.

در مجموع می توان مدل ARMA در مورد آلاینده CO و مدل ARFIMA در مورد آلاینده PM_{2.5} را به عنوان الگوهایی که دارای کمترین خطا در پیش بینی هستند، معرفی کرد، اما در مورد دیگر آلاینده ها، با توجه به مقادیر به دست آمده برای چهار معیار MAPE، APE، RMSE، MSE، نمی توان به طور قطعی مدلی را معرفی کنیم که کمترین خطا در پیش بینی را دارا باشد، ولی به طور کلی تر می توان پیش بینی حاصل از مدل های ARFIMA و ARMA-GARCH را برای O₃ و PM₁₀ بهتر از مدل ARMA دانست. برای آلاینده های NO₂ و SO₂ نیز دو مدل ARMA و ARFIMA پیش بینی های بهتری نسبت به مدل ARMA-GARCH ارائه می کنند.

مطالعات فوق کاربرد مدل های اقتصادسنجی را برای پیش بینی وضعیت آلاینده ها اثبات می کند. براین اساس توصیه می شود سیاست گذاران مسوول در حوزه کیفیت هوا با توجه به هزینه های اجتماعی بالایی که در اثر انتشار آلاینده های گازی بر جامعه تحمیل می شود، با به کارگیری این الگوها، آلاینده های تأثیرگذار بر آینده هوای شهر را شناسایی و در جهت کاستن از سطح انتشار آن ها طرح های کارآمدی را پیاده نمایند.

منابع

1. Tehran Air Quality Annual Report, 2012. Technical Report of the Air Quality Control Company, No. QM92/03/03/(U)/01. (In Persian)

17. Festic, M., Kavkler, A., Dajcman, S., 2012. Long memory in the Croatian and Hungurian stock market returns. *Zb. Rad. Ekon. Fak. Rij*, 30(1), 115-139.
18. Erfani, A., Jafari samimi, A., 2009. Long memory forecasting of stock price index using a fractionally differenced ARMA model. *Journal of Applied sciences research*, 5(10), 1721-1731.
19. Bollerslev, T., 1986. Generalized autoregressive conditional Hetrokedasticity. *Journal of Econometrics*, 31, 307-328.
20. Mussadiq, T., 2012. Modeling and forecasting the volatility of oil futures using the ARCH family models. *The Lahore Journal of Business*, 1, 79-108.
21. Alem shiferaw, Y., 2012. Modeling volatility of price of some selected Agricultural products in Ethiopia: ARIMA-GARCH applications. Available at SSRN: <http://ssrn.com/abstract=2125712> or <http://dx.doi.org/10.2139/ssrn.2125712>.
22. Noferesti, M., 2010. Unit root and cointegration in econometrics, 3ed. Rasa publication. (In Persian).
23. Granger, C. W. J., Joyeux, R., 1980. An introduction to long memory time series models and fractional differencing. *Journal of time series analysis*, 1(1), 15-29.
24. Hosking, J. R. M., 1981. Fractional differencing. *Biometrika*, 68(1), 165-176.
25. Baillie, R. T., Chung, S-K., 2012. Modeling and forecasting from trend-stationary long memory models with applications to climatology. artificial neural networks model to forecast particulate mater in urban areas: The case of Temuco. Chile. *Atmospheric Environment*, 42, 8331-8340
9. Siew, L. Y., Chin, L. Y., Wee, P. M. J., 2008. ARIMA and integrated ARFIMA models for forecasting air pollution index in Shah Alam, Selangor. *The Malaysian Journal of Analytical Sciences*, 12(1), 257-263.
10. Bagheri Soltanabadi, H., 2009. comparison of genetic planning and ARIMA time series models in Short Term Forecasting of Crude Oil Prices in Iran. MSc thesis. Faculty of economics, Ferdowsi University of Mashhad, Iran. (In Persian).
11. Kumar, A., Goyal, P., 2011. Forecasting of daily air quality index in Delhi. *Science of the total environment*, 409, 5517-5523.
12. Saffarini, G., Odat, S., 2008. Time series analysis of air pollution in AL-Hashimeya Town Zarqa, Jordan. *Jordan Journal of Earth and Environmental Sciences*, 1(2), 63-72.
13. Hurst, H. E., 1951. Long-term storage capacity of reservoirs. *Transactions of the American society of civil engineers*, 116, 770-779.
14. Mandelbrot, B. B., Wallis, J., 1968. Noah, Joseph and operational hydrology. *Water resources research*, 4(5), 909-918.
15. Baillie, R. T., Bollerslev, T., 1994. The long memory of the forward premium. *Journal of international money and finance*, 13(5), 565-571.
16. Baillie, R. T., 1996. Long memory processes and fractional integration in econometrics. *Journal of Econometrics*, 73: 5-59.

29. Lee, D., Schmidt, P., 1996. on the power of the KPSS test of stationary against fractionally-integrated alternatives. *Journal of Econometrics*, 73, 285-302.
30. Baillie, R. T., Chung, C-F., Tieslau, M. A., 1996. Analysing inflation by the fractionally integrated ARFIMA-GARCH model. *Journal of applied econometrics*, 11, 23-40.
- International Journal of Forecasting, 18, 215-226.
26. Varotsos, C., Kirk-Davidoff, D., 2006. long-memory processes in ozone and temperature variations at the region 60 S-60 N. *Atmospheric Chemistry and Physics*, 6(12), 4093-4100.
27. Geweke, J., Porter-Hudak, S., 1983. The estimation and application of long memory time series models. *Journal of time series analysis*, 4(4), 221-238.
28. Hurst, H. E., 1957. A suggested statistical model of some time series that occur in nature. *Nature*, 180-494.