

## الگوسازی ARIMA برای بارش سالانه مشهد

آتوسا خجسته

دانشجوی دکتری گروه جغرافیای دانشگاه زنجان، زنجان، ایران

حسین عساکره

دانشیار گروه جغرافیا دانشگاه زنجان، زنجان، ایران (نویسنده مسئول)

### چکیده

گرمایش زمین، الگوی گردش عمومی و الگوی زمانی- مکانی بارش را تغییر داده است. کاربرد روش‌های آماری در تشریح تغییرات، ابزاری مفید به شمار می‌آیند. از آن جا که بارش به عنوان یک عنصر اقلیمی؛ رفتاری غیرخطی داشته، از توزیع نرمال تبعیت نمی‌کند. در نتیجه بهره‌گیری از مدل‌سازی برای شناخت رفتار عمومی بارش امری بسیار کارآمد است. روش‌های مختلفی برای مدل‌سازی تغییرات بارش وجود دارد که از آن جمله می‌توان به الگوسازی خطی (کمینه قدر مطلق انحرافات، روش کمترین مربعات خطا و...) و غیر خطی (الگوی چند جمله‌ای، شبکه عصبی مصنوعی) اشاره نمود. برای انجام پژوهش حاضر از داده‌های بارش ایستگاه فرودگاهی مشهد طی دوره آماری ۵۷ ساله (۱۹۶۰-۲۰۱۶) برای شناسایی رفتار دراز مدت بارش این ایستگاه بهره گرفته شده است. به منظور بررسی رفتار عمومی بارش مشهد تلاش شد با استفاده از نرم افزار مینی تب (MINITAB) به برازش الگوی مناسب از خانواده چند جمله‌ای و الگوسازی ARIMA پرداخته شود. نتیجه الگوسازی خانواده چند جمله‌ای حاکی از یک روند درجه دو یا سهمی گونه در بارش مشهد است. از سوی دیگر در خانواده الگوی ARIMA، الگوی (۱،۱،۴) ARIMA که از میان بقیه الگوها به نسبت بهتر بود به عنوان الگوی مناسب تعیین شد. همچنین با انجام تبدیل باکس-کاکس مجدداً به ارزیابی بارش‌های تبدیل یافته، مدل‌های کاندید و پیش بینی آنها پرداخته شد و نهایتاً مدل گزینش شده برای بارش مشهد الگوی (۱،۱،۱) ARIMA بوده است. لازم به ذکر است که براساس نتایج حاصل از مدل‌سازی و خروجی‌های پیش‌بینی مدل می‌توان استنباط نمود که ARIMA خطی برای برازش بر بارش مشهد الگوی مناسبی نیست. از سوی دیگر با توجه به عدم برازنده بودن این الگو برای بارش مشهد مدل‌های گزینش شده نیز با روند موجود در مشاهدات همخوانی ندارند و یا به عبارت دیگر مرتبه تفاضل‌گیری مدل‌ها برابر یک است که حاکی از روند خطی و نه سهمی در مشاهدات است.

**کلیدواژه:** مدل‌سازی بارش، ARIMA، الگوسازی چند جمله‌ای، روند درجه دو، مشهد.

## مقدمه

بررسی پیش‌بینی و روند بارش در مقیاس فصلی و سالانه یک از اقدامات مفید و موثر جهت برنامه‌ریزی‌های کلان و خرد اقتصادی- اجتماعی می‌باشد. در صورتی که برآوردهای قابل اعتمادی از وضعیت اقلیمی چند ماه یا چند سال آینده در دسترس باشد، برنامه‌ریزان، مدیران و اقشار مختلف اجتماعی می‌توانند ضمن آمادگی جهت مقابله با حوادث نامطلوب، نسبت به بهره‌برداری بهینه از منابع آب و خاک و نیروی انسانی اقدام نمایند (ناظم السادات و همکاران، ۱۳۸۵: ۱). در این میان روش‌های پیش‌بینی مختلفی سعی در تعیین ارتباط بین متغیرهای مستقل و وابسته داشته‌اند و مدل‌های مفهومی و آماری زیادی برای پیش‌بینی و پیش‌آگاهی متغیرهای اقلیمی استفاده شده است. پژوهشگران با اتخاذ شیوه‌های مختلف درصد برآمده اند تا نتایج قابل قبولی را در این زمینه برای نقاط مختلف دنیا به دست آورند. تجزیه و تحلیل سری‌های زمانی معمولاً دو هدف را دنبال می‌کند: درک یا مدل کردن مکانیسم تصادفی که به مشاهده سری منجر می‌شود و پیش‌بینی مقادیر آینده سری که بر مبنای گذشته آن صورت می‌گیرد (گلابی و همکاران، ۱۳۹۳: ۶۲).

پژوهش‌هایی در ایران و جهان در زمینه تحلیل روند انجام گرفته است که در ذیل به بیان برخی از آن‌ها می‌پردازیم. طی بررسی پژوهش‌های گذشته، دانشمندان به منظور تحلیل تغییرات پارامترهای اقلیمی تلاش نموده‌اند تا این پارامترها را الگوسازی و سپس شبیه‌سازی نمایند. در قاره آفریقا، چوبوکی و همکاران<sup>۱</sup> (۱۹۹۵: ۱۷۲۵) با تأکید بر اهمیت تعداد نمونه در تعیین تغییرات اقلیمی نظیر دما، روش مدل‌سازی آریما را از معتبرترین روش‌های بررسی تغییرات اقلیمی دانسته‌است. پادیلا و همکاران<sup>۲</sup> (۱۹۹۶: ۹۱۸) مدل‌های استوکستیک ARIMA را برای تحلیل سری‌های زمانی دبی سه - چشمه کارستی، اسپانیا و فرانسه به کار گرفتند نتایج مطالعه آن‌ها نشان داد که می‌توان از مدل‌های استوکستیک به منظور پیش‌بینی دبی استفاده کرد، همچنین ARIMA را بهترین مدل معرفی کردند. بورلاند و مونتانا<sup>۳</sup> (۱۹۹۶: ۱۹۹) از مدل‌های ARIMA جهت پیش‌بینی بارندگی‌های ساعتی در زمان وقوع آن‌ها استفاده کردند و مقادیر به دست آمده را با داده‌های باران‌سنجی مقایسه کردند. بوهادو و همکاران<sup>۴</sup> (۱۹۹۷: ۲۳) با استفاده از روش ARIMA به شبیه‌سازی پارامترهای آب و هوایی مانند دمای محیط و رطوبت پرداختند. هالتینر و ولیام<sup>۵</sup> (۱۹۹۸: ۴۴۷) از یک مدل (۱ و ۱) ARIMA فصلی در مدل‌سازی دو متغیره جریان ماهانه در رودخانه‌های یامپا و وایت<sup>۶</sup> در شمال غربی کالیفرنیا استفاده کردند. فراستو و

۱ - Chbouki, et al

۲ - Padilla, et al

۳ - Borland & Montana

۴- Bouhaddou et al

۵- Haltiner & Wiliams

۶ - Yampa & White rivers

همکاران<sup>۱</sup> (۲۰۰۳: ۱۴۸) نشان دادند که با استفاده از اتو رگرسیون (AR) و ARIMA می‌توان درجه حرارت هوا را در داخل وسایل گرمازا توصیف نمود. آگیلرا و همکاران<sup>۲</sup> (۲۰۰۷: ۳۱۸۷) برای پیش بینی داده‌های دو طرف طول جغرافیایی مدلی از تلفیق مدل ARIMA و مولفه‌های اصلی (PCA) به کار بردند که این مدل را برای پیش بینی خطر پدیده ال‌نینو شایسته دانستند. آن‌ها در تحقیق خود به این نتیجه رسیدند که با افزایش مدت دوام بارندگی، پیش‌بینی‌ها روند دقیق‌تری داشتند و با کوتاه‌تر شدن دوام بارندگی، اختلاف میزان باران پیش‌بینی از مقدار واقعی متناظر خود بیشتر می‌شود. دردو<sup>۳</sup> (۲۰۱۰: ۵۸۶) از شبکه عصبی مصنوعی و آریمای پیش‌بینی کیفیت آب رودخانه‌ای در ترکیه استفاده کرد و نشان داد که ترکیب این دو مدل با هم نتیجه بهتری می‌دهد.

خردمندنیسا و عساکره (۱۳۸۰: ۱۲۱) به الگوسازی ARIMA برای متوسط درجه‌ی سالانه‌ی دما در ایستگاه جاسک پرداخته‌اند. آن‌ها با استفاده از روش‌های آماری خودهمبستگی و خودهمبستگی جزئی و بازبینی بهنجار بودن باقیمانده‌ها نتیجه‌گیری کردند که متوسط دمای سالانه با شبیه‌سازی در خانواده چندجمله‌ای و ARIMA حاکی از یک روند افزایشی معنی‌دار بوده است. جهانبخش اصل و باباپورباصر (۱۳۸۲: ۳۶) با استفاده از مدل آریمای باکس-جنکینز متوسط دمای ماهانه ایستگاه تبریز را مورد بررسی قرار داده‌اند. پس از مقایسه معیار آکائیک الگوی آزمایشی مدل (۱، ۱، ۰) به عنوان مدل محاسباتی انتخاب گردید و براساس آن تغییرات متوسط دمای ماهانه ایستگاه تبریز تا سال ۲۰۱۰ پیش‌بینی شده است. عساکره (۱۳۸۶: ۱۹) با استفاده از روش‌های آماری به تشریح تغییرات دمای سالانه شهر تبریز پرداخته است، بررسی‌های انجام شده نشان می‌دهد که روند خطی در سطح ۰/۰۵ خطا معنی‌دار بوده و از روند افزایشی برخوردار است. عبدالله نژاد (۱۳۹۴: ۱۵) با بهره‌گیری از مدل ARIMA به بررسی و پیش‌بینی بارش ماهانه در گرگان پرداخت. ایشان مطرح نمودند که الگو آریمای به درستی نتوانسته است بارش را تصویر کند تا بتوان برای آینده پیش‌بینی مناسبی را انتظار داشت با این وجود (عدم اطمینان) می‌تواند نقش تأثیرگذاری در مدیریت منابع آب داشته باشد. پرتوی منش، عطایی (۱۳۹۵: ۲) ایشان به منظور مدل‌سازی بارش سالانه سنندج با بهره‌گیری از روش‌های آماری همچون تحلیل طیفی و ARIMA به تحلیل پرداختند. بر طبق نتایج این پژوهش بارش سالانه سنندج دارای روند معناداری می‌باشد.

عموماً روش‌های مختلف و پرشماری برای مدل‌سازی (الگوسازی) رفتار عمومی پارامترهای اقلیمی وجود دارد. از جمله این روش می‌توان به روش‌های الگوی چند جمله‌ای و ARIMA اشاره نمود. در پژوهش حاضر به منظور شناسایی رفتار عمومی سری زمانی بارش مشهد، بارش سالانه مشهد طی دوره آماری ۱۹۶۰-۲۰۱۶ (۵۷ سال) با دو روش مطروحه مورد ارزیابی قرار

۱ - Frausto, et al

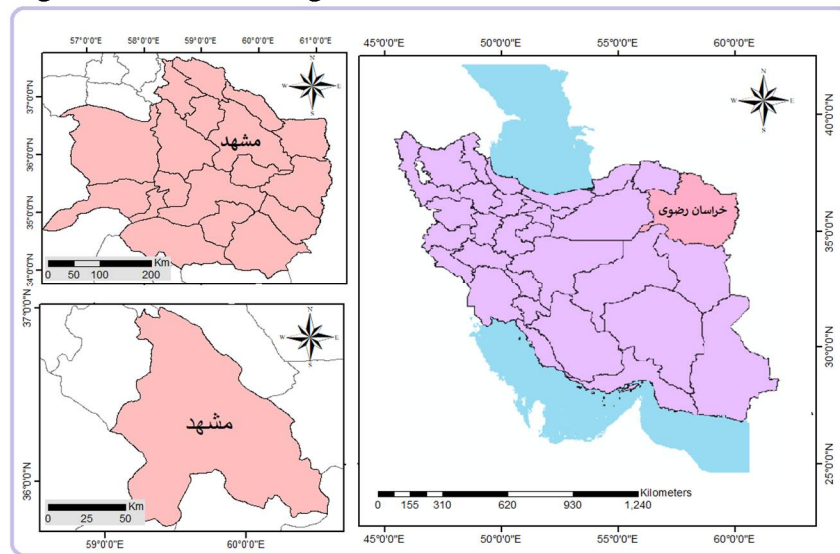
۲ - Aguilera, et al

۳ - Durdu

گرفت. پژوهش حاضر با بهره‌گیری از این دو روش، تلاش بر مدل‌سازی و پیش‌بینی بارش مشهد نمود. ضرورت انجام این پژوهش به این دلیل است که در میان پژوهش‌های پرشماری که بر مناطق مختلف ایران با بهره‌گیری از روش ARIMA انجام گرفته است ایستگاه مشهد کمتر مورد توجه قرار گرفته است.

## مواد و روش منطقه مورد مطالعه

مشهد کلانشهری در شمال شرقی ایران و مرکز استان خراسان رضوی است. این استان درحوزه رودخانه‌ی کشفروود و در دشت مشهد بین کوه‌های هزارمسجد و بینالود گسترده شده است. این شهر دارای آب و هوای متغیر، اما معتدل و متمایل به سرد و خشک است و از تابستان‌های گرم و خشک و زمستان‌های سرد و مرطوب برخوردار است (اعلمی، ۱۳۹۱: ۱۲۸). ایستگاه فرودگاهی مشهد در طول جغرافیایی ۵۹ درجه و ۱۵ دقیقه تا ۶۰ درجه و ۳۶ دقیقه و عرض جغرافیایی ۳۵ درجه و ۴۳ دقیقه تا ۳۷ درجه و ۸ دقیقه قرار گرفته است. ارتفاع کمینه این ایستگاه ۹۵۰ متر و ارتفاع بیشینه‌ی این ایستگاه ۱۱۵۰ متر می‌باشد (هواشناسی مشهد).



شکل (۱) منطقه مورد مطالعه (منبع: نگارنده)

## داده‌ها

در پژوهش حاضر جهت مدل‌سازی سری زمانی بارش ایستگاه مشهد داده سری زمانی بارش روزانه این ایستگاه در دوره زمانی ۱۹۶۰-۲۰۱۶ (۱۳۳۹-۱۳۹۵) از سازمان هواشناسی کشور اخذ گردید. در این مطالعه سعی شده است با بهره‌گیری از روش‌های متداول آماری به بررسی رفتار عمومی بارش سالانه مشهد پرداخته شود. به‌منظور سازماندهی و تجزیه و تحلیل داده‌ها از نرم افزار MINITAB استفاده شده است.

## روش

در سری‌های زمانی چند مشخصه وجود دارد، یکی از این مشخصه‌ها و اجزا روند نام دارد. برای تحلیل روند، باید یک سری زمانی را در دراز مدت بررسی نمود، بسیاری از مؤلفین کلمه روند را برای تغییر تدریجی

تابع میانگین، مانند یک تابع خطی یا درجه دوم به کار می‌برند. روش‌های آماری کلاسیک در تجزیه و تحلیل رگرسیون به سهولت می‌تواند پارامترهای مدل را بدون وجود مقدار ثابت مورد ارزیابی قرار دهد (کرایر، ۱۳۷۱: ۵۰). در مطالعه حاضر مفیدترین آن‌ها یعنی خطی و سهمی مورد سنجش قرار گرفت. ابتدا باید تشخیص داده شود که شیب خط از نوع خطی است یا سهمی؟ یکی از روش‌های پرکاربرد برای پاسخگویی به این سوال الگوسازی در خانواده چندجمله‌ای‌هاست (خردمندنیا و عساکره، ۱۳۸۰: ۱۲۵). الگوی چند جمله‌ای درجه  $K$  به صورت زیر تعریف می‌شود (یاراحمدی و میریان، ۱۳۹۳: ۱۶۹):

$$Y_t = \beta_0 + \beta_1 t + \beta_2 t^2 + \dots + \beta_k t^k + e_t \quad \text{رابطه ۱}$$

در رابطه فوق  $Y_t$  متغیر پاسخ،  $\beta$ ها فراسنج‌های مجهول و  $e_t$  دنباله‌ای از متغیرهای تصادفی مستقل نرمال با امید ریاضی صفر و واریانس ثابت هستند الگوی خط و الگوی سهمی به ترتیب متناظر با  $k=1$  و  $k=2$  می‌باشد. پس از حصول روند موجود در داده‌ها به منظور ایستاسازی مشاهدات تبدیل باکس کاکس بر مشاهدات بارش مشهد اعمال گردید. این تبدیل از جمله تبدیلات غیر خطی است که بر مشاهدات پیاده‌سازی می‌شود در حقیقت تبدیل باکس\_کاکس برای تغییر توزیع مشاهدات و دستیابی به توزیع مورد نظر (بوسیله به توان رساندن) اعمال می‌گردد به عبارت دیگر این تبدیل بر داده‌های غیر نرمال، دارای چولگی، دارای ناپایداری پرارش انجام می‌گیرد و تلاش می‌شود تا مسائل مطرح شده مرتفع گردد (این تبدیل به تبدیل شبه نرمال نیز معروف و از انواع تبدیلات توانی است). نکته حائز اهمیت در این تبدیل آن است که باید بر مشاهدات نامنفی پیاده سازی شود. این تبدیل توسط رابطه زیر محاسبه می‌گردد (عساکره، ۱۳۹۱: ۱۵۲):

$$Y = \frac{(X^{\lambda}-1)}{\lambda} \quad \text{رابطه ۲}$$

در رابطه ۲،  $X$  هریک از مشاهدات،  $\lambda$  مقادیر تبدیلات توانی است. هدف غایی از مدل‌سازی یک سری زمانی، یافتن تغییرات دراز مدت و پیش بینی آینده است. منظور از پیش‌بینی سری زمانی، تخمین مقادیری از مجموعه داده‌ها می‌باشد که در زمان اجرای تحلیل، مقدار آن مجهول است. یکی از فنون پیش‌بینی رفتار سری زمانی، روش باکس\_جنکیز است. گروه عمومی مدل‌ها برای یک سری زمانی در روش‌شناسی باکس\_جنکینز مدل‌های تلفیقی اتورگرسیو و میانگین متحرک است که در آمار به مدل‌های ARIMA معروف اند (علیجانی و رمضانی، ۱۳۸۱: ۱۵۸). مدل آریما یک مدل کلی است که توانایی نمایندگی طبقه گسترده‌ای از سری‌های زمانی ناپایستا را دارد، فرایند تلفیقی خود همبسته\_میانگین متحرک با درجه  $(p,d,q)$  است. با توجه به اینکه بیشتر سری‌های زمانی ناپایستا هستند، لذا این رده از فرایندها کاربرد گسترده‌ای دارند.

فرایند تصادفی  $\{Z_t\}$  یک فرایند ARIMA است که با درجه  $(p,d,q)$  در نظر گرفته و به صورت  $Z_t \sim ARIMA(p,d,q)$  نوشته می‌شود. هرگاه فرایند تفاضلی  $\{w_t = (1-B)^d Z_t\}$  یک فرایند ایستا باشد به شکل زیر بیان می‌گردد (خردمندنیا و عساکره، ۱۳۸۰: ۱۲۶):

$$\Phi_p(B)(1-B)^d(w_t - u_w) = \theta_q(B)a_t \quad \text{رابطه ۳}$$

که در رابطه فوق  $(w_t - u_w) = E(w_t)$ ،  $a_t \sim WN(0, \sigma_a^2)$ ، یعنی فرایند  $\{a_t\}$  دنباله ای از متغیرهای تصادفی با ویژگی‌هایی همچون هم‌توزیع، دو به دو ناهمبسته (نوفه سفید) با امید ریاضی صفر و واریانس  $\sigma_a^2$  می‌باشد. عملگرهای بکار رفته در رابطه ۳ به صورت زیر تعریف می‌شوند (وئی، ۱۳۷۶: ۹۹)

$$\Phi_p(B) = 1 - \Phi_1(B) - \Phi_2(B)^2 - \dots - \Phi_p(B)^p \quad \text{رابطه ۴}$$

$$\theta_q(B) = 1 - \theta_1(B) - \theta_2(B)^2 - \dots - \theta_q(B)^q \quad \text{رابطه ۵}$$

که در آن B عملگری است (به عملگرپسرو معروف است) که به صورت  $B^j Z_t = Z_{t-j}$  تعریف می‌شود (l عدد صحیح نامنفی است).

برای انتخاب مدل‌های کاندید (آزمایشی) به شرح زیر عمل می‌شود:

**مرحله اول:** پیش از انجام هر کار در ابتدا باید مرتبه تفاضل گیری به منظور ایستاسازی مشاهدات صورت گیرد. اگر سری زمانی  $\{Z_t\}$  حول یک محور افقی نوسان کند مقدار تفاضل‌گیری را  $d=0$ ، اگر سری زمانی حول یک خط مستقیم (رو به بالا یا پایین) نوسان کند  $d=1$  و اگر حول یک سهمی نوسان کند  $d=2$  قرار می‌دهیم. بندرت پیش می‌آید یک سری زمانی شیب درجه سوم  $d=3$  داشته باشد. به بیان دیگر؛ برای تعیین مقدار  $d$  یا مرتبه تفاضل گیری باید به نحوی عمل نمود که سری تفاضل‌گیری شده  $\{w_t\} = (1-B)^d Z_t$  حاکی از ایستایی یا مانایی سری مولد  $\{w_t\}$  باشد (بدان معنا که اگر سری زمانی حول یک محور افقی و با واریانس ثابت نوسان نماید در حقیقت سری مورد بررسی ایستا شده است).

**مرحله دوم:** پس از تعیین مرتبه تفاضل گیری با بهره‌گیری از همبستگی نگار  $\{w_t\}$  مقادیر  $p$  و  $q$  را حاصل می‌کنیم. در حقیقت پس از تعیین  $d$ ، با توجه به شاخک‌های معنی‌دار نمودار همبستگی نگار سری  $\{w_t\}$  می‌توان مدل اول کاندید به صورت  $ARIMA(0, q)$  و با توجه به شاخک‌های خودهمبستگی نگار جزئی  $\{w_t\}$  می‌توان مدل کاندید دوم  $ARIMA(p, 0)$  را تعیین نمود (در حقیقت همبستگی نگار بیان کننده مقدار  $q$  و خودهمبستگی نگار جزئی مقدار  $p$  را نمایان می‌کند).

پس از استخراج مدل‌های کاندید باید با روش زیاد برآزاندن و با آزمایش و خطا مدل‌های کاندید تعدیل شده، و به الگویی دست یابیم که اضافه و حذف کردن هیچ پارامتری آن را به شکل معناداری بهتر یا بدتر نکند.

به سبب اینکه باقی مانده های هر مدل باید دارای خصوصیات از جمله نرمال بودن (هم توزیع بودن) و دو به دو ناهمبسته بودن (نوفه سفید)، باشد در نتیجه پس از برآزش هر الگو باقی‌مانده‌های آن، از نظر نرمال بودن و استقلال مورد سنجش قرار گرفت. برای این منظور علاوه نمودارهای معمول همچون خودهمبستگی نگار، احتمال نرمال و هیستوگرام از برخی

آزمون‌های دیگر همچون آزمون نقاط عطف و علامت (برای آگاهی از نحوه محاسبه و فرضیات آماری این آماره‌ها به عساکره ۱۳۹۰ رجوع گردد) و آزمون خی دو (کای دو) برای بررسی استقلال داده‌ها بهره‌گرفته شد. آزمون خی دو، از جمله آزمون‌های ناپارامتریکی است که می‌توان در چندین حالت از آن بهره گرفت؛ ۱- زمانی که قصد داریم سری مورد بررسی را از نظر نیکویی برازش ارزیابی کنیم، ۲- زمانی که هدف بررسی همگنی مشاهدات است، ۳- سومین استفاده از این نوع آزمون زمانی است که، قصد داریم استقلال یا عدم استقلال سری را مورد ارزیابی قرار دهیم (برای آگاهی از فرضیات هر سه حالت ممکن و نحوه محاسبه‌ی آن به بلبوکا و همکاران، ۲۰۱۱: ۵۳۰ مراجعه نمایید). همچنین برای بررسی نرمال بودن مشاهدات از روش‌ها و آزمون‌های کلموگروف-اسمیرنوف، رایان جویئر و اندرسون دارلینگ بهره گرفته شد؛

آماره کلموگروف-اسمیرنوف به دو صورت مورد استفاده قرار می‌گیرد؛ ۱- زمانی که دو سری زمانی از دو جامعه متفاوت با یکدیگر از نظر توزیعی خاص مورد قیاس قرار می‌گیرند، ۲- زمانی که یک توزیع نمونه از نظر توزیعی خاص (نرمال، گاما و...) مورد بررسی و تحلیل قرار می‌گیرد که با عنوان آزمون نیکویی برازش نیز معروف می‌باشد. این آماره توسط رابطه ذیل برآورد می‌گردد (عساکره، ۱۳۹۰: ۵۱۱)؛

$$D_n = \max |F_n - F_E|$$

رابطه ۶

براساس رابطه (۶)  $F_n$  بیانگر تابع تجربی (حاصل از مشاهدات)،  $F_E$  تابع توزیع تجمعی نظری است که حاصل می‌گردد.

آماره رایان جویئر نیز روش دیگری است که برای بررسی توزیع مورد نظر در داده‌ها مورد استفاده قرار گرفت. برای محاسبه این آماره از رابطه زیر بهره برده می‌شود (چنتاسورن، ۲۰۱۱: ۱۰)؛

$$R_p = \frac{\sum y_i b_i}{\sqrt{s^2(n-1) \sum b_i^2}}$$

رابطه ۷

آزمون اندرسون دارلینگ بررسی می‌کند که آیا داده‌ها از توزیع مشخص شده پیروی می‌کنند یا خیر. آماره‌ی این آزمون به شرح رابطه زیر است (رومه، ۲۰۰۳: ۲)؛

$$AD = \sum_{i=1}^n \frac{1-i}{n} \{ \ln(F_n[Z_{(i)}]) + \ln(1 - F_n[Z_{(n+1-i)}]) \} - n$$

رابطه ۸

در رابطه فوق  $F_n$  توزیع مورد نظر (نرمال)،  $Z_{(i)}$  پارامتر مورد بررسی (طبقه بندی یا استاندارد شده)،  $n$  اندازه نمونه،  $\ln$  لگاریتم طبیعی (پایه  $e$ )، اندیکس  $i$  از ۱ تا  $n$  را شامل می‌گردد.

گاهی این امکان وجود دارد که چندین مدل برازنده برای الگوسازی انتخاب گردد برای انتخاب بهترین از میان آن‌ها باید از ضریب اطلاع آکائیک (AIC) استفاده شود که براساس رابطه زیر حاصل می‌گردد (خردمندیا و عساکره، ۱۳۸۰: ۱۲۶)

$$AIC(M) = n \ln(S_a^2) + 2m$$

رابطه ۹

$S_a^2$  برآورد حداکثر درست‌نمایی  $\sigma_a^2$ ،  $n$  طول سری  $w_t$  و  $m$  تعداد پارامترهای مدل می‌باشد. برای یک مدل ARIMA که دارای پارامتر  $\theta$  می‌باشد  $m = p+q+1$  بصورت نمایش داده می‌شود. با در نظر گرفتن معیار AIC اگر قصد داشته باشیم مدلی را گزینش نماییم باید مدلی با میزان ضریب آکائیک کمینه انتخاب گردد.

## یافته‌ها

### الگوی خانواده چند جمله‌ای

شکل (۱) میانگین بارش سالانه مشهد را طی دوره آماری ۱۹۶۰-۲۰۱۶ نشان می‌دهد. در این شکل بصورت واضحی مشخص است که رفتار عمومی سری زمانی بارش مشهد رفتاری حول یک خط (رو به بالا یا پایین) نیست در نتیجه این احتمال می‌رود که رفتار این سری زمانی رفتاری سهمی گونه داشته باشد. برای کسب نتیجه علمی و مبتنی بر دلیل و مدرک نیاز است که هردو (درجه یک و دو) مورد ارزیابی قرار گیرد.

پس از بررسی‌های انجام گرفته بر روند درجه یک و دو در سری زمانی بارش مشهد، وجود روند درجه یک (نوسان مقادیر حول یک خط صاف روبه بالا یا پایین) نقض و وجود روند درجه دو (ابتدا افزایش مقادیر و سپس کاهش مقادیر و عکس آن که به رفتار سهمی معروف است) تأیید گردید.

الگوی خطی برازش داده به سری زمانی به صورت زیر نمایان شده است؛

$$Y_t = 254.4 - 0.207t \quad (0.36) \quad (13.39)$$

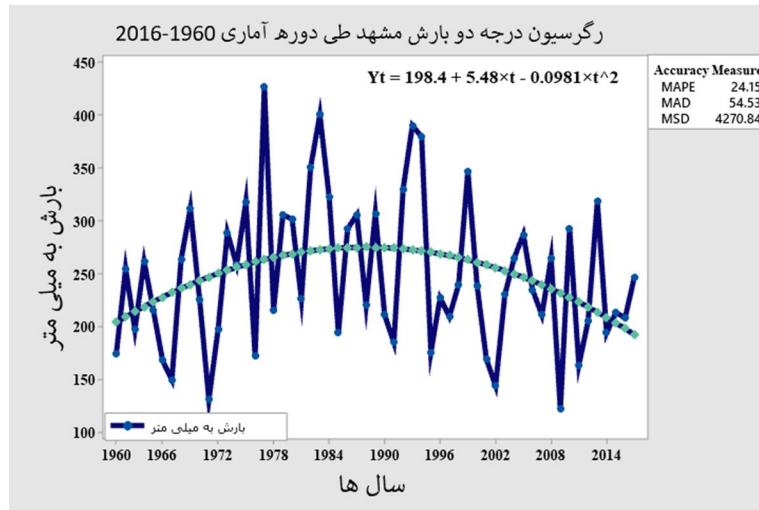
با اضافه نمودن متغیر توضیحی  $t^2$  به الگو فوق تلاش شد تا وجود روند سهمی در سری زمانی مورد ارزیابی قرار گیرد:

$$Y_t = 198.4 + 0.48t - 0.0981t^2 \quad (7.18) \quad (2.49) \quad (2.67)$$

اعدادی که در زیر هریک از پارامترهای مدل ثبت شده است نمایش دهنده مقادیر  $t$  (t-student) است. این مقدار برای ارزیابی معناداری یا عدم معناداری ضرایب مورد استفاده قرار گرفت. اگر قدرمطلق مقادیر  $t$  برای آزمون دو دامنه بزرگتر یا مساوی ۲ و برای آزمون یک دامنه بزرگتر یا مساوی ۱/۹۶ باشد می‌توان در سطح ۹۵ درصد اطمینان قضاوت نمود که پارامتر مورد نظر تفاوت معناداری را با صفر نشان می‌دهد. در غیر اینصورت عملاً پارامتر مورد نظر را می‌توان مساوی صفر (عدم معناداری) در نظر گرفت.

بر این اساس می‌توان مشاهده نمود مقدار  $\beta_1 = 0$  به سبب کوچکتر بودن این مقدار از ۲ می‌توان این پارامتر را از مدل حذف نمود. این در حالی است که با اضافه نمودن متغیر توضیحی، تمامی پارامترهای مدل، تفاوت معناداری را با صفر نشان می‌دهند در نتیجه می‌توان وجود روند سهمی را پذیرفت. لازم به ذکر است که وجود روند درجه سه و چهار نیز در بارش مشهد مورد ارزیابی قرار گرفت و ضرایب معنادار نشد در نتیجه ترجیح با روند سهمی است.

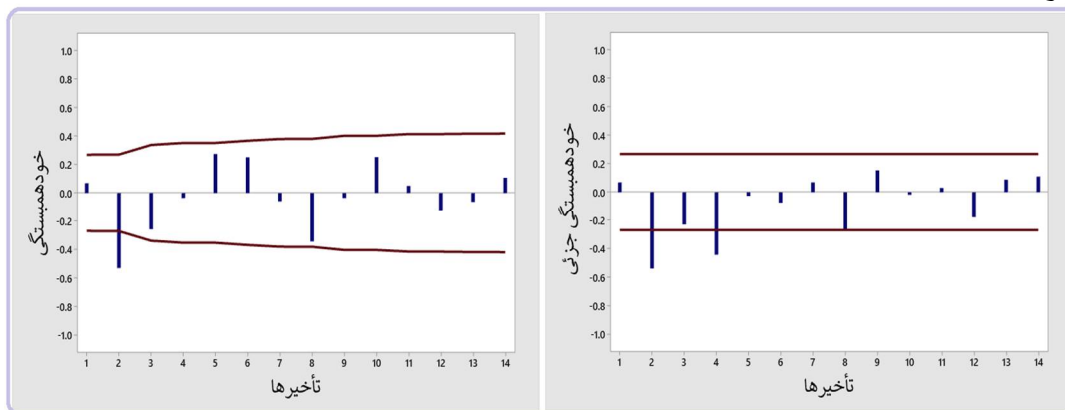




شکل (۱) میانگین بارش مشهد طی دوره آماری ۱۹۶۰-۲۰۱۶

### ۳-۲ تعیین مدل ARIMA برای بارش مشهد

با توجه به شکل (۱) و بررسی‌های صورت گرفته این مسئله تأیید گردید که روند موجود در بارش مشهد روندی است سهمی و مقادیر نیز بیانگر علت ترجیح مدل سهمی به مدل خطی (درجه سه یا چهار) است. بدین ترتیب مقدار تفاضل‌گیری با  $d=2$  برابری می‌کند. حال باید با بهره‌گیری از نمودارهایی که پیشتر از آن‌ها صحبت به میان آمد مقادیر  $p$  و  $q$  را حاصل نمود (شکل ۲).



شکل (۲) خودهمبستگی و خودهمبستگی نگار جزئی سری تفاضلی مجموع بارش سالانه مشهد ( $d=2$ )

شکل ۲ نشان دهنده مدل‌های کاندید (آزمایشی) است که باید مورد ارزیابی قرار گیرد. مدل کاندید اول که براساس نمودار خودهمبستگی به آن دست یافتیم  $ARIMA(1,2,0)$  می‌باشد به سبب آنکه تنها یک شاخک از مرز معناداری خارج شده است در نتیجه مقدار  $q=1$  خواهد بود. پس از اضافه نمودن  $\theta$  مقدار  $t$  برابر با  $0/80$  حاصل گردید که به سبب آنکه مقدار آن از ۲ کوچکتر است بنابراین مقدار با صفر تفاوت معناداری ندارد و می‌توان این ضریب را از مدل حذف نمود. در نتیجه الگوی برازش داده شده بصورت زیر خواهد بود:

$$w_t = 1.004 a_{t-1} + a_t$$

(۴۴.۲۵)

اکنون باید ارزیابی نمود که نقش ضریب  $\Phi_1$  در مدل مورد بررسی به چه نحو است و در صورت حضور معنادار باید به عنوان پارامتر به مدل اضافه گردد. با توجه به مقادیر حاصل از  $t$  می‌توان مشاهده نمود که قدر مطلق  $t$  پارامتر مورد نظر از ۲ بزرگتر است در نتیجه این پارامتر به مدل اضافه می‌گردد و مدل به  $ARIMA(1,2,1)$  ارتقاء می‌یابد نهایتاً پارامترهای مدل برآزش یافته به شرح زیر می‌باشد:

$$w_t = (-0.443)w_{t-1} - (0.9920)a_{t-1} + a_t$$

(۳.۴۳)                      (۱۵,۹۵)

در فرایند زیاد برآزاندن با اضافه نمودن  $\theta_p$  و دیگر پارامترها به مدل بار دیگر مدل مورد ارزیابی قرارگرفت و با توجه به عدم معناداری پارامترها در این پروسه در نتیجه مدل ارتقاء یافته هیچ رقیب دیگری ندارد.

در این مرحله با توجه به مقدار واریانس باقی‌مانده‌ها برای مدل‌های برآزش داده شده می‌توان تعیین نمود که کدام مدل به عنوان مدلی مطلوب جهت انجام بررسی‌های بیشتر است. در میان مقادیر واریانس‌های هر مدل، الگوی دوم (الگوی ارتقاء یافته)، الگوی مناسبی برای برآورد بارش مشهد تشخیص داده شد. برای کسب داوری علمی‌تر بین دو مدل معنادار از ضریب اطلاع آکائیک بهره‌گرفته شد تا بتوان مدل بهینه را در میان دو مدل شناسایی شده استخراج نمود. با توجه به مقادیر AIC حاصل شده برای دو مدل می‌توان استنباط نمود که بین مدل‌ها، مدل ارتقاء یافته می‌تواند کاندید مناسبی برای مطالعه باشد.

جدول (۱) مقادیر واریانس باقی‌مانده‌ها

مدل‌ها	الگوها	واریانس باقی‌مانده‌ها ( $\sigma_a^2$ )	ضریب اطلاع (AIC) آکائیک
مدل پایه	ARIMA (۱,۲,۰)	۹۱۷۶/۷۰	۵۰۳/۸۴۲۰
مدل ارتقاء یافته	ARIMA (۱,۲,۱)	۷۵۴۶/۲۷	۴۹۵/۰۵۸۴

پیشتر با بهره‌گیری از نمودار (PACF)<sup>۱</sup> توانستیم مقدار P را برای ساخت مدل کاندید دوم (مدل پایه دوم) حاصل کنیم ( $p=7$ ). در نتیجه می‌توان استنباط نمود که مدل بصورت  $ARIMA(7,2,0)$  به عنوان مدل کاندید دوم برگزیده شد. با توجه به محدودیت نرم افزار مینیتب که اجازه نمی‌دهد مقادیر  $p$  یا  $q$  بیش از ۵ باشد در نتیجه مدل را با  $p=5$  آغاز می‌کنیم  $ARIMA(5,2,0)$ . پس از اضافه کردن مقدار  $\theta$  مشاهده شد که مقدار  $t=0/09$  حاصل گردید با توجه به اینکه قدر مطلق  $0/09$  از ۲ کوچکتر است در نتیجه وجود این ضریب نقش معناداری را در مدل ایفا نمی‌کند بنابراین از آن صرف نظر می‌شود. الگو برآزش یافته به شرح زیر است:

<sup>۱</sup> -Partial autocorrelation function

$$W_t = (-1.390)w_{t-1} - 1.369w_{t-2} - 1.230w_{t-3} - 0.973w_{t-4} - 0.399w_{t-5} \quad (3.01)$$

(10.69)                      (7.10)                      (5.80)                      (4.97)

در پروسه زیاد برازاندن پارامتر  $\theta_1$  نیز به مدل اضافه و مدل ارتقاء یافته و بصورت  $ARIMA(5,2,1)$  حاصل گردید و پارامترهای مدل به شرح زیر است:

$$W_t = 2.12w_{t-1} - 2.14w_{t-2} - 1.76w_{t-3} - 1.21w_{t-4} - 0.42w_{t-5} - 0.96a_{t-1} + a_t \quad (3.15)$$

(15.55)                      (7.44)                      (5.20)                      (4.21)                      (3.15)

(15.48)

در فرایند زیاد برازاندن با اضافه نمودن دیگر پارامترها به مدل، بار دیگر الگوها مورد ارزیابی قرارگرفت و با توجه به عدم معناداری پارامترها در این پروسه در نتیجه مدل ارتقاء یافته هیچ رقیب دیگری ندارد.

براساس جدول (2) می‌توان مقادیر واریانس باقی مانده‌ها (ستون سوم جدول) و مقادیر ضریب اطلاع آکائیک (ستون آخر جدول) را مشاهده نمود. از میان الگوی ارتقاء یافته و مدل کاندید می‌توان استنباط نمود که ترجیح با مدل پایه است چراکه هم مقدار واریانس آن (نوسانات حول میانگین) مقدار کمتری (8039/99) را نمایش می‌دهد و هم مقدار ضریب آکائیک آن (504/5655) نسبت به الگوی دیگر کمینه‌تر است.

**جدول (2) مقادیر واریانس باقی‌مانده‌های الگوی کاندید دوم**

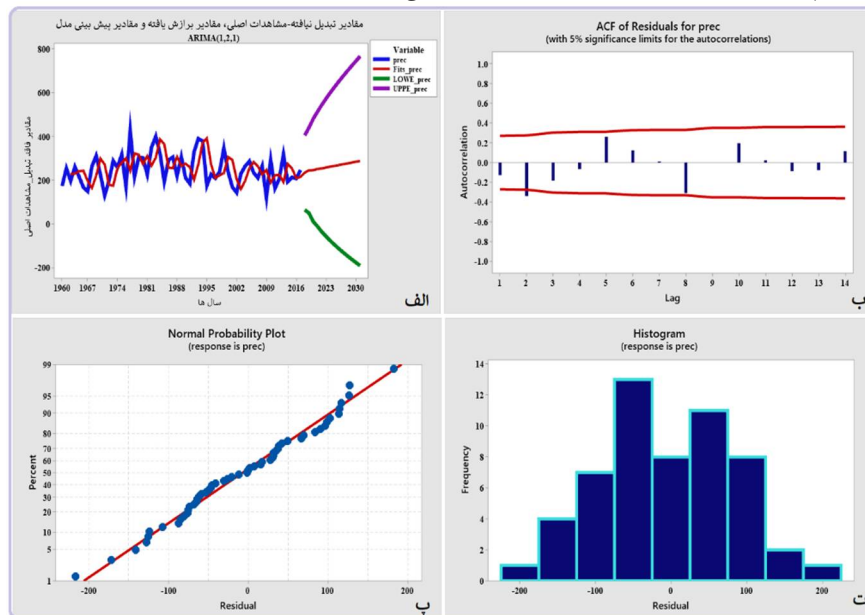
مدلها	الگوها	واریانس باقی مانده‌ها ( $\sigma_a^2$ )	ضریب اطلاع آکائیک (AIC)
مدل پایه	ARIMA (0,2,5)	8039/99	504/5655
مدل ارتقاء یافته	ARIMA (1,2,5)	9459/72	515/5085

با توجه به نتایج حاصل شده از برازش و زیاد برازاندن مدل‌های کاندید اول و دوم می‌توان استنباط نمود که تنها یک مدل از میان مدل‌های مطرح شده و مورد ارزیابی قرار گرفته می‌تواند کاندید مناسبی باشد به سبب اینکه هم مقدار واریانس باقی مانده‌ها و هم مقدار ضریب اطلاع آکائیک آن نسبت به دیگر الگوها، مقداری کمینه است این الگو، مدل  $ARIMA(1,2,1)$  می‌باشد. لازم به ذکر است که به سبب عدم پذیرش ضریب ثابت در هیچ یک از مدل‌ها در نتیجه نمی‌توان وجود یا عدم وجود این ضریب را ملاکی برای پذیرش یا نقض مدل در نظر گرفت (وجود ضریب ثابت از قطعیت وجود روند در داده‌ها حکایت دارد درحالی که مدل‌های برازش یافته هیچ یک دارای ضریب ثابت نمی‌باشند).

بمنظور داوری دقیق‌تر در زمینه مدل  $ARIMA(1,2,1)$  و دیگر مدل‌ها، وضعیت باقی مانده‌های مدل و همچنین نحوه پیش بینی مدل‌ها مورد ارزیابی قرار گرفت. باتوجه به این مسئله که پیش‌بینی همواره بر  $\frac{n}{4}$  طول دوره آماری صورت می‌گیرد در نتیجه پیش بینی‌ها برای 14 سال آینده (2017-2030) انجام پذیرفت.

بررسی‌ها بر دو فاکتور مطروحه نیز که بصورت آماری و ترسیمی انجام گرفت و حاکی از آن است که مدل گزینش شده براساس ضریب آکائیک و

واریانس باقی مانده‌ها همچنان نسبت به مدل‌های دیگر بهترین گزینه برای مدل‌سازی بارش ایستگاه مشهد می‌باشد. خودهمبستگی نگار باقی‌مانده‌های این الگو، با وجود معناداری یک شاخک در میان دیگر مدل‌ها به نسبت بهتر است و نمودار احتمال نرمال با فاصله اطمینان ۹۵ درصد نیز شرایط نرمال بودن باقی‌مانده‌ها را تأیید می‌کند. از سوی دیگر نمودار هیستوگرام نیز تأییدگر نرمال بودن باقی‌مانده‌هاست. به عبارت دیگر شرایط مورد نیاز برای پذیرش یک الگو، باقی‌مانده‌های نرمال تأیید و دو به دو ناهمبسته توسط این مدل نسبت به مدل‌های دیگر بیشتر رعایت شده است (اما همچنان دو به دو همبسته هستند). از سوی دیگر پیش‌بینی انجام شده براساس این مدل برآورد مطلوبی نمی‌باشد به سبب اینکه بین حد میانی، حد بالا و پایین پیش‌بینی فاصله زیادی وجود دارد در نتیجه این انتظار می‌رود که، فرایند پیش‌بینی ۱۴ سال آینده توأمان با خطا و عدم اطمینان صورت پذیرد. همچنین حد میانی پیش‌بینی در راستای تغییرات بارش این ایستگاه نمی‌باشد چرا که براساس معادله سهمی که در بخش ۳-۱ مشاهده شد انتظار می‌رود که ابتدا بارش در این ایستگاه تا نیمه اول دوره روندی صعودی و پس از آن روندی نزولی را تجربه نماید.



شکل (۳) الف: پیش‌بینی مدل برگزیده (۱، ۲، ۱)، مقادیر برازش یافته و مشاهدات اصلی، ب: خودهمبستگی نگار باقیمانده‌ها، پ: احتمال نرمال باقیمانده‌ها، ت: هیستوگرام باقیمانده‌ها

چنان که دیدیم، با اضافه و کم کردن فراسنج‌ها در هریک از الگوها توانستیم الگوی مناسب را اختیار کنیم. این تکنیک بخشی از عملیات زیاد برازاندن است. در بخش دیگر، از زیاد برازاندن با تغییر مقدار الگوهای پیشنهادی را آزمودیم. در حقیقت تمامی مراحل فوق برای شرایط  $d=0$  و  $d=2$  محاسبه گردید.

جدول (۳) مدل‌های کاندید حاصل از زیاد برآزاندن براساس درجات تفاضلی ۱ و ۳

شماره الگو	الگو	واریانس	معیار AIC
$M_1$	(۱،۱،۰) ARIMA	۰۰۳۹/۹۰	۴۷۹/۴۰۵۶
$M_2$	(۱،۱،۴) ARIMA	۰۸۰۸/۲۸	۴۸۷/۱۰۸۰
$M_3$	(۲،۳،۰) ARIMA	۱۳۰۸۸/۷	۵۱۷/۹۱۷۵
$M_4$	(۰،۳،۰) ARIMA	۱۶۷۳۰/۵	۵۳۵/۱۴۴۶

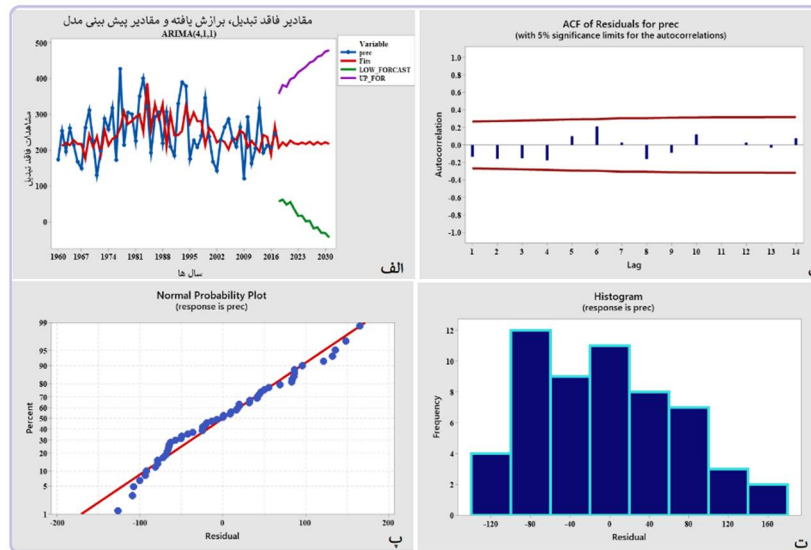
براساس جدول (۳) همان‌گونه که مشاهده می‌شود بجز مدل  $M_2$  بقیه مدل‌ها از نظر مقدار واریانس باقی مانده‌ها و مقدار ضریب آکائیک نسبت به مدل برگزیده‌ی اول، مدل‌های مناسبی به شمار نمی‌آیند. بدین ترتیب تنها یک مدل می‌تواند به عنوان رقیب برای الگوی اول باشد. پارامترهای این الگو بصورت زیر حاصل شده است:

$$W_t = a_t - (-1.073)w_{t-1} - (-1.104)w_{t-2} - (-0.808)w_{t-3} - (-0.411)w_{t-4} - (-0.9415)a_{t-1}$$

(۱۱/۸۳)                      (۴/۸۳)                      (۳/۵۵)                      (۳/۱۸)                      (۱۶/۳۸)

لازم به ذکر است که تمامی مدل‌های ارائه گشته در جدول ۳ از نظر باقی مانده‌ها و پیش‌بینی نیز مورد ارزیابی قرار گرفته و نهایتاً مدل  $M_2$  از میان دیگر مدل‌ها مورد تأیید قرار گرفت.

با توجه به شکل (۴) می‌توان مشاهده نمود که پیش‌بینی مدل برگزیده دوم ( $M_2$ ) از دقت مناسبی برخوردار نیست. براساس نمودار واضح است که بین حد متوسط پیش‌بینی و حد بالا و پایین پیش‌بینی فاصله زیادی وجود دارد از سوی دیگر به دلیل رفتار متغیر در راستای پیش‌بینی می‌توان نتیجه گرفت که احتمال این مسئله که به تناوب زیادی و سپس کمی برآورد (یابالعکس) رخ دهد دور از انتظار نیست، در نتیجه پیش‌بینی از عدم اطمینان برخوردار می‌باشد. از سوی دیگر، افزایش فواصل پیش‌بینی در انتهای دوره‌ی مورد برآورد، حاکی از عدم قطعیت برآورد برای سال‌های انتهایی است. به سبب این که مهم‌ترین رکن مناسب بودن یک مدل، پیش‌بینی مناسب آن است براین اساس می‌توان نتیجه گرفت که این مدل در مقایسه با الگوی گزینش شده اول الگوی مناسبی برای بارش مشهد از نظر پیش‌بینی نیست. مقادیر باقی مانده‌های مدل، مستقل (براساس نمودار خود همبستگی نگار)، و از نظر توزیع دارای توزیع نرمال می‌باشند. با این وجود نمودار هیستوگرام در الگوی اول قابل قبول‌تر از الگوی دوم است. نکته حائز اهمیت آن است که به سبب اینکه مهم‌ترین و اولیه‌ترین ویژگی در پذیرش یک مدل آن است که باقی مانده‌های مدل باید مستقل باشند با توجه به این مسئله که باقی مانده‌های مدل (۱،۲،۱) ARIMA مستقل نمی‌باشد در نتیجه می‌توان این مدل را نپذیرفت. از سوی دیگر مدل (۴،۱،۱) ARIMA با برقراری این اصل در باقی مانده‌ها و همچنین کمینه بودن ضریب آکائیک در آن نسبت به مدل (۱،۲،۱) ARIMA مورد تأیید قرار می‌گیرد.



شکل (۴) الف: پیش بینی مدل برگزیده (۱، ۱، ۴)، مقادیر برازش یافته و مشاهدات اصلی، ب: خودهمبستگی نگار باقیمانده ها، پ: احتمال نرمال باقیمانده ها، ت: هیستوگرام باقیمانده ها

پس از شناسایی مدل بهینه (۱، ۱، ۴) از میان دیگر مدل‌ها که بر مشاهدات اصلی بارش مشهد پیاده سازی شد. با انجام تبدیل باکس-کاکس باردیگر با گذراندن بارش مشهد از یک فیلتر شرایط و مدل‌های کاننید مجدداً مورد ارزیابی قرار گرفتند. با انجام تبدیل باکس-کاکس بر مشاهدات مورد بررسی لاندا برابر با ۰/۳۸ حاصل شد و با اعمال لاندا و تبدیل باکس-کاکس بر داده‌ها، سری زمانی جدید تبدیل یافته حاصل گردید. با انجام تمامی مراحل مدل سازی ARIMA بر مشاهدات تبدیل یافته مشهد مدل‌های کاننید و ارتقاء یافته حاصل شدند. مدل‌ها به شرح جدول ۴ می‌باشد.

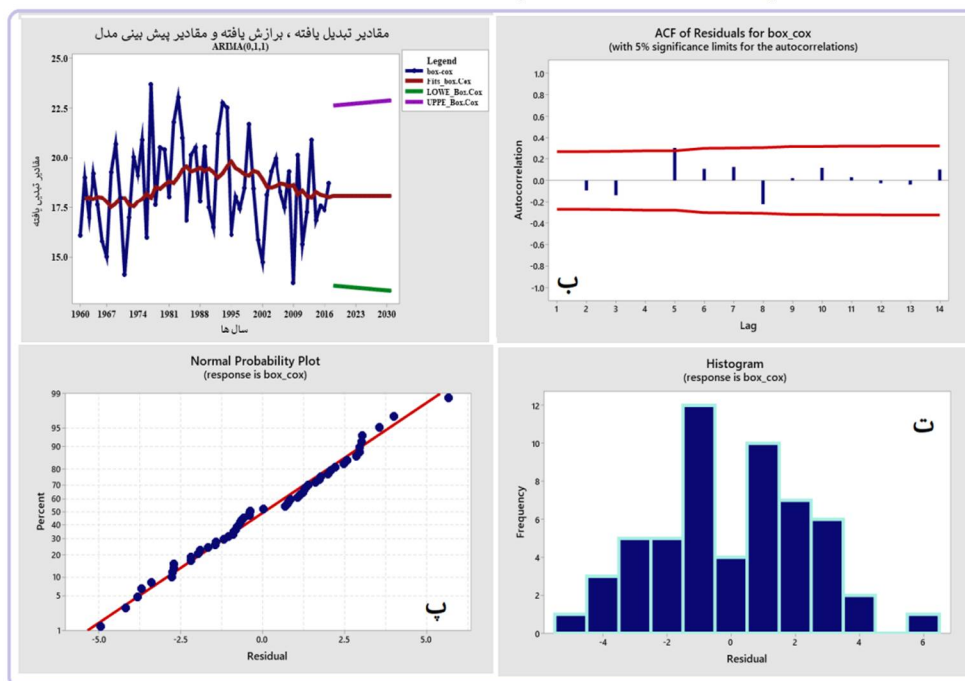
جدول (۴) مدل‌های کاننید حاصل از زیاد برازاندن براساس درجات تفاضلی ۱، ۲ و ۳

رتبه تفاضل گیری	شماره الگو	مدل	واریانس	ضریب AIC
d=۲	$M_1$	ARIMA(۱، ۲، ۱)	۸/۱۱۰۶۵	۱۱۹/۱۲۰۵
	$M_2$	ARIMA(۴، ۲، ۰)	۹/۶۸۳۷۶	۱۳۲/۸۷۴۷
d=۱	$M_3$	ARIMA(۰، ۱، ۱)	۵/۳۶۵۰۷	۹۶/۰۷۷۴
	$M_4$	ARIMA(۴، ۱، ۱)	۴/۹۸۳۹۴	۱۰۳/۰۸۳۲
d=۳	$M_5$	ARIMA(۱، ۳، ۲)	۱۴/۹۵۰۹	۱۵۲/۰۵۳۸
	$M_6$	(۴، ۳، ۱) con ARIMA	۱۰/۹۶۹۸	۱۴۱/۳۳۵۴

براساس جدول فوق و باتوجه به این مسئله که گزینش مدل مناسب در مقاله حاضر به سبب عدم برازنده بودن مدل خطی بر مشاهدات (عدم رعایت تمام خصوصیات لازم برای باقی‌مانده‌ها) علاوه بر وضعیت باقی مانده‌ها ضریب آکائیک مد نظر قرار گرفت. می‌توان استنباط نمود که در میان مدل‌های کاننید و یا ارتقاء یافته برای مشاهدات دارای تبدیل تنها مدلی که از نظر ضریب آکائیک و شرایط باقی مانده‌ها شرایط بهتری را

نسبت به مدل های دیگر احراز نموده است مدل  $M_3$  با مرتبه تفاضل گیری ۱ می باشد.

با توجه به شکل (۵) میتوان مشاهده نموده که حد متوسط پیش بینی در راستای مشاهدات نیست و با توجه به آنکه بازه های برآوردی بسیار بزرگ هستند در نتیجه انتظار می رود پیش بینی توأمان با خطا باشد اما نکته حائز اهمیت آن است که به سبب یکسانی بازه های برآوردی می توان نتیجه گرفت که میزان خطا از ابتدا تا به انتهای دوره ۱۴ ساله مورد پیش بینی یکسان خواهد بود. نمودار احتمال نرمال باقی مانده ها (پ) نیز گویای توزیع نرمال باقی مانده هاست (نوسان حول خط قرمز رنگ است). از سوی دیگر هیستوگرام و نمودار خودهمبستگی (ت و ث) به ترتیب حاکی از آن است که باقی مانده های شبه نرمال و دو به دو همبسته یا به عبارتی استقلال باقی مانده ها رعایت نشده است



شکل (۵) الف: پیش بینی مدل برگزیده (۱، ۱، ۰)، مقادیر برازش یافته و مشاهدات اصلی، ب: خودهمبستگی نگار باقیمانده ها ، پ: احتمال نرمال باقیمانده ها ، ت: هیستوگرام باقیمانده ها

پس از گزینش دو مدل از میان داده های تبدیل یافته (۱، ۱، ۰) و فاقد تبدیل (۴، ۱، ۱) برای کسب نتایج دقیق تر در زمینه باقی مانده ها ابتدا استقلال آن ها و سپس نرمال بودن آن ها مورد ارزیابی قرار گرفت. نتایج حاصل از بررسی استقلال به شرح جدول ۵ است؛

جدول (۵) آزمون های استقلال باقی مانده های تبدیل یافته و فاقد تبدیل بارش

مشهد

	داده های تبدیل یافته	داده های فاقد تبدیل
آزمون نقاط عطف	۰/۶۵۵۵	۰/۴۹۵۳
آزمون علامت	۰/۸۹۳۹	۰/۹۹۵۰
آزمون خی دو	۰/۷۳۶۷	۰/۴۳۱۰

با توجه به نتیجه حاصل شده از آماره‌های استنباطی همگی از ۰/۰۵ بزرگتر می‌باشند در نتیجه شواهد کافی برای رد کردن فرض صفر وجود ندارد. به عبارت دیگر اگر فرض صفر که حاکی از استقلال مشاهدات و تصادفی بودن آن هادر دو آزمون اول می‌باشد را رد کنیم ۶۵/۵، ۸۹/۳۹، ۴۹/۵۳ و ۹۹/۵ درصد خطا خواهیم کرد. همچنین شرایط آزمون خی دو نیز بیانگر آن است که اگر فرض صفر را ( $H_0$ ) متغییر کیفی مورد بررسی (مستقل است) رد کنیم به ترتیب مرتکب ۷۳/۶۷ و ۴۳/۳۶ درصد خطا خواهیم شد. این درحالی است که تنها امکان تا ۲/۵ (اگر یک دامنه باشد ۵) درصد خطا را داریم. در نتیجه شواهد کافی برای رد فرض صفر وجود ندارد. این بدان معناست که برای باقی مانده‌های دو مدل تبدیل یافته و فاقد تبدیل می‌توانیم استقلال را بپذیریم. با توجه به اینکه روش‌های پارامتریک (خودهمبستگی نگار) در مقایسه با روش‌های ناپارامتریک (آزمون‌های علامت، نقاط عطف و آزمون خی دو) از اعتبار علمی بیشتری برخوردار هستند در نتیجه برای تصمیم‌گیری نهایی نمودارهای خودهمبستگی را ملاک ارزیابی قرار می‌دهیم (به سبب اینکه این روش‌ها معمولاً بر رتبه‌ها و یا علامت و... کار می‌کنند و نه بر روی مشاهدات اصلی در نتیجه برای دریافت نتیجه دقیق‌تر بهتر است نتیجه حاصل شده از مشاهدات اصلی را مد نظر قرار دهیم).

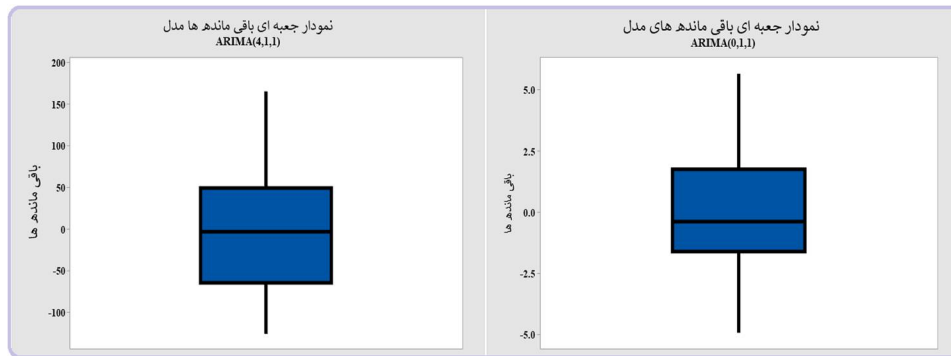
نتایج حاصل از بررسی سه آزمون نرمال بودن (جدول ۶)، بدین صورت حاصل گردید که تمامی روش‌های سه گانه تأییدگر نرمال بودن باقی‌مانده‌های مدل تبدیل یافته و فاقد تبدیل می‌باشد. مقادیر  $P\_value$  برای مشاهدات تبدیل یافته عبارتند از ۰/۵۵۸ برای آندرسون دارلینگ، ۰/۱۰ برای رایان جویئر و ۰/۱۵۰ برای کلموگروف- اسمیرنوف می‌باشد همچنین برای مقادیر تبدیل نیافته (مشاهدات اصلی) مقادیر برابر با ۰/۱۸۴ برای آندرسون دارلینگ، ۰/۱۰ برای رایان جویئر و ۰/۱۵۰ برای کلموگروف- اسمیرنوف است. با توجه به مقادیر حاصل گشته از  $P\_value$  برای هر یک از آزمون‌ها و در نظر گرفتن این مسئله که اگر مقدار  $P\_value$  از ۰/۰۵ کوچکتر باشد شرایط برای پذیرش فرض مقابل فراهم می‌گردد (یا به عبارتی معنادار خواهد بود) و اگر از ۰/۰۵ بزرگتر باشد شواهد کافی برای رد فرض صفر وجود ندارد در نتیجه می‌توان پذیرفت که باقی مانده‌های مدل‌های تبدیل یافته و فاقد تبدیل از توزیع نرمال تبعیت می‌نمایند.

جدول (۶) ضرایب معناداری و آزمون‌های نرمال برای باقی مانده‌های مدل تبدیل یافته و فاقد تبدیل بارش مشهد

	ضرایب داده های تبدیل یافته	P_VALUE	ضرایب داده های فاقد تبدیل	P_VALUE
آندرسون دارلینگ	۰/۳۰۵	۰/۵۵۸	۰/۵۱۵	۰/۱۸۴
رایان جویئر	۰/۹۹۴	۰/۱۰	۰/۹۸۷	۰/۱۰
کلموگروف- اسمیرنوف	۰/۰۹۰	۰/۱۵۰	۰/۰۹۲	۰/۱۵۰

همچنین با ترسیم نمودار جعبه‌ای تلاش شد که نحوه توزیع مشاهدات و همچنین وجود یا عدم وجود داده‌ای پرت مورد ارزیابی قرار گیرد و نتیجه به صورت شکل ۶ حاصل گردید:





شکل (۶) نمودار جعبه‌ای دو مدل برگزیده؛ مدل فاقد تبدیل  $ARIMA(4,1,1)$  و مدل تبدیل یافته  $ARIMA(0,1,1)$

با توجه به شکل ۶ میتوان استنباط نمود که مدل  $(4,1,1)$  بسیار به شرایط نرمال نزدیک است به عبارت دیگر میتوان استنباط نمود که توزیع باقی مانده‌ها شبه نرمال است و داده‌ی پرتی در باقی‌مانده‌های این مدل دیده نمی‌شود. از سوی دیگر نمودار جعبه‌ای مقادیر تبدیل یافته  $(0,1,1)$  حاکی از چولگی در توزیع مشاهدات و عدم نرمال بودن باقی‌مانده‌هاست همچنین براساس شکل باقی مانده‌های این مدل نیز داده‌ی پرتی ندارند.

### نتیجه‌گیری

یکی از چالش‌هایی که بشر با آن روبرو است، مسأله تغییر اقلیم و مشکلات ناشی از آن می باشد. اگرچه علل تغییر هنوز کاملاً شناخته نشده، اما بحث در مورد تغییر آب و هوا مورد توجه بسیاری از پژوهشگران بوده و خواهد بود (ترابی پوده و امامقلی زاده، ۱۳۹۴: ۵۷). یکی از ابزارهای معمول مورد استفاده برای تشخیص تغییرات در سری‌های زمانی اقلیمی و هیدرولوژیکی تجزیه و تحلیل روند است (خلیق و همکاران، ۲۰۰۹: ۱۸۳). برای شناسایی روند موجود در سری‌های روش‌های مختلفی همچون روش‌های خطی و روش‌های غیر خطی وجود دارد. که هر یک از این روش‌های دارای زیرشاخه‌های مختلفی است. به عنوان مثال روش‌های خطی شامل: کمترین قدر مطلق انحرافات، روش کمترین مربعات خطا و... در شیوه روندیابی غیر خطی: الگوی چند جمله‌ای، شبکه عصبی مصنوعی و... را می‌توان نام برد. روش  $ARIMA$  که در پژوهش حاضر نیز مورد استفاده قرار گرفته است از جمله روش‌های مدل‌سازی خطی غیر تناوبی (سالانه) است.

در این مطالعه با استفاده از روش الگوسازی در خانواده چند جمله‌ای و الگوسازی  $ARIMA$  تلاش شد تا به شناسایی رفتار عمومی بارش ایستگاه مشهد طی دوره آماری ۱۹۶۰-۲۰۱۶ و همچنین پیش‌بینی بارش ۱۴ سال آینده مشهد پرداخته شود. نتایج نشان داد که بارش مشهد دارای روند معنادار درجه دو است، سهمی آن کمانی روبه بالاست در نتیجه انتظار می‌رود که تا نیمه دوره‌ی مورد مطالعه بارش افزایش یافته و پس از آن بارش رفتار نزولی داشته باشد. سرانجام برای پیش‌بینی بارش مشهد با بهره‌گیری از مدل  $ARIMA$  و بررسی مانده‌ها و توان پیش‌بینی بارش

آن‌ها بهترین مدل نسبت به دیگر مدل‌ها برای پیش‌بینی بارش ایستگاه مشهد، مدل  $(1,1,4)$  ARIMA تعیین گردید. از سوی دیگر به سبب عدم پذیرش ضریب  $\theta$  توسط مدل می‌توان نتیجه گرفت که روند موجود در بارش ایستگاه مشهد روندی قطعی نیست. با در نظر گرفتن این مسئله که پیش‌بینی در سال‌های ابتدایی مورد برآورد، دارای دقت بالاتر و در سال‌های انتهایی از دقت آن کاسته می‌شود در نتیجه انتظار می‌رود که در پیش‌بینی بارش مشهد نیز این مسئله تداعی یابد در حالی که به سبب بازه برآوردی زیاد و همچنین نوسانات موجود در حدود (حد بالایی، حد میانی و حد پایین) برآورد که زمینه ساز زیادی برآورد و کمی برآورد در راستای پیش‌بینی می‌شود در نتیجه این خصلت تحقق پیدا نکرده است. همچنین برای دستیابی به نتیجه‌ای دقیق‌تر بارش مشهد را از فیلتری به نام تبدیل باکس-کاکس نیز گذر دادیم و پس از آن تمامی فرایندهای مدل‌سازی ARIMA را مجدداً بر بارش تبدیل یافته اعمال نمودیم. از میان چندین مدل برازش یافته مدل  $(0,1,1)$  ARIMA که به نسبت بقیه‌ی مدل‌ها از شرایط بهتری در پیش‌بینی و... برخوردار بود، مورد تأیید قرار گرفت.

همچنین با انجام آزمون‌های ناپارامتریک جهت بررسی باقی‌مانده‌های مدل‌های برگزیده این نتیجه حاصل گردید که باقی‌مانده‌ها مستقل هستند این در حالی است که نمودار خودهمبستگی نگار حاکی از عدم استقلال باقی‌مانده‌هاست به سبب اینکه همواره روش‌های پارامتریک نسبت به روش‌های ناپارامتریک از اعتبار علمی بیشتری برخوردار هستند در نتیجه استقلال باقی‌مانده‌ها تنها در مدل  $(1,1,4)$  مورد تأیید است. همچنین با بهره‌گیری از سه آزمون کلموگروف-اسمیرنوف، اندرسون دارلینگ و رایان جویئر نرمال بودن باقی‌مانده‌های دو مدل ارزیابی گردید نتایج حاکی از نرمال بودن باقی‌مانده‌ها است.

غفوریان و همکاران (۱۳۹۲: ۵) در پژوهش خود با بهره‌گیری از مدل SARIMA و ARIMA به شبیه‌سازی بارش مشهد پرداختند و بیان نمودند که مقادیر پیش‌بینی شده بارش مشهد در مقادیر ماهانه و فصلی دارای دقت بیشتری بوده است، این در حالی است که در پژوهش حاضر تنها به بررسی بارش سالانه مشهد و مدل‌سازی آن پرداخته و از نظر طول دوره آماری پژوهش غفوریان و همکاران از ابتدای تأسیس ایستگاه مشهد تا سال ۲۰۱۰ یا به عبارتی ۵۹ سال را مورد ارزیابی قرار داده‌اند در نتیجه از نظر طول دوره آماری مورد بررسی بسیار نزدیک به پژوهش حاضر عمل شده است. در پژوهش جعفرپور و تفتی (۱۳۹۳: ۱) بیان شده است که مدل آریما در بازه‌های زمانی کوتاه مدت بهتر از دیگر مدل‌ها عمل می‌کند و پیش‌بینی بهتری را از خود تصویر می‌کند، این احتمال وجود دارد که اگر در پژوهش حاضر نیز دوره‌ی آماری کمی کوتاه‌تر و یا همچون پژوهش غفوریان و همکاران با استفاده از مدل SARIMA به ارزیابی پرداخته شود، به نتایج متقن‌تر و نزدیک‌تری به حال (به سبب اینکه مطالعه حاضر شرایط را تا سال ۲۰۱۶ مورد ارزیابی قرار داده است) می‌توان دست یافت. از سوی دیگر مطالعات صورت گرفته بر بارش در ایران بیشتر با بهره‌گیری از مدل‌های SARIMA و گاه‌ها ARIMA درکنار یکدیگر به نتایج دقیق‌تری دست یافته‌اند همچون پژوهش صلاحی و

مرشت (۱۳۹۴)، پروین و حجازی زاده (۱۳۸۶)، حلمی و همکاران (۱۳۹۹) و... را می‌توان نام برد.

از این رو برای پژوهش‌های آینده پیشنهاد می‌شود که، برای کسب دآوری علمی‌تر برای بارش مشهد از الگوی XARIMA (الگوی آریمای غیر خطی) به سبب آنکه در پژوهش حاضر این مسئله تأیید گردید که الگوی آریمای خطی قادر به پیش‌بینی مناسب برای بارش مشهد نیست و SARIMA بهره گرفته شود، همچنین با توجه به احتمالی که نویسنده می‌دهد در صورت طولانی بودن دوره آماری وجود روند سینوسی یا درجه سه در بارش مشهد تأیید خواهد گردید در نتیجه با در دست بودن دوره آماری طولانی‌تر از ۵۷ سال می‌توان وجود روند درجه سه را مورد ارزیابی قرار داد.

### منابع

- ۱- اعلمی، حمید. ۱۳۹۱. راهنمای جامع ایرانگردی. انتشارات اعلمی. ۱-۳۶۰.
- ۲- پرتوی منش، سحر. عطایی، هوشمند. ۱۳۹۵. تحلیل و مدلسازی بارش سالیانه سنندج با استفاده از مدل ARIMA. اولین کنفرانس بین المللی آب، محیط زیست و توسعه پایدار. اردبیل. ۱۵-۱.
- ۳- ترابی پوده، حسن. امامقلی‌زاده، صمد. ۱۳۹۴. تحلیل روند تغییرات آبدهی رودخانه های شمال حوضه رودخانه دز با استفاده از روش TFPW-MK. نشریه علمی پژوهش‌های حفاظت آب و خاک. ۳: ۳۹-۵۵.
- ۴- جهانبخش اصل، سعید. باباپورباصر، علی اکبر. (۱۳۸۲). بررسی و پیش‌بینی متوسط دمای ماهانه تبریز با استفاده از مدل آریما (ARIMA)، تحقیقات جغرافیایی، ۱۸: ۳۴-۴۶.
- ۵- جعفرپور، زین‌العابدین. فلاح تفتی، سحر. ۱۳۹۳. پیش‌بینی روند بارش با استفاده از مدل آریما (مطالعه موردی: استان فارس). دومین همایش ملی پژوهش‌های کاربردی در جغرافیا و گردشگری. تهران - دانشگاه جامع علمی کاربردی. ۱-۵.
- ۶- حجازی زاده، زهرا. پروین، نادر. ۱۳۸۶. مدل سازی بارش و پیش‌بینی آن با استفاده از مدل های SARIMA و پایش خشکسالی به کمک شاخص BMI و نمایه PDRI در حوضه آبریز دریاچه ارومیه. تحقیقات جغرافیایی، ۸۷: ۸۷-۹۷-۱۲۴.
- ۷- حلمی، مهتاب. بختیاری، بهرام. قادری، کوروش. ۱۳۹۹. مدلسازی و پیش‌بینی خشکسالی هواشناسی با استفاده از مدل سری زمانی SARIMA در نمونه‌های اقلیمی مختلف ایران. نشریه علمی آبیاری و زهکشی ایران. ۳: ۱۰۷۹-۱۰۹۰.
- ۸- خردمندنیا، منوچهر. عساکره، حسین. ۱۳۸۰. الگوسازی ARIMA برای متوسط درجه حرارت سالانه هوا در جاسک. سومین سمینار احتمال و فرایندهای تصادفی. انتشارات دانشگاه اصفهان. ۱۲۱-۱۳۲.
- ۹- صلاحی، برومند. ملکی مرشت، رقیه. ۱۳۹۴. پیش‌بینی و تحلیل تغییرات بارش‌های ماهانه شهرستان اردبیل با استفاده از مدل‌های آریما، اتورگرسیو و وینترز. نشریه آب و خاک (علوم و صنایع کشاورزی). ۵: ۱۳۹۱-۱۴۰۵.
- ۱۰- ناظم السادات، محمد جعفر. کامکارحقیقی، علی اکبر. شریف زاده، مریم. احمدوند، مصطفی. ۱۳۸۵. پذیرش پیش‌بینی‌های بلند مدت بارش (مورد مطالعه: گندمکاران استان فارس)، مجله علوم ترویج و آموزش کشاورزی ایران، جلد ۲۲، تهران. ۱۵-۱.

- ۱۱- عساکره، حسین. ۱۳۸۰. الگوسازی ARIMA برای متوسط درجه حرارت سالانه هوا در جاسک. سومین سمینار و فرایندهای تصادفی. دانشگاه اصفهان. ۱۲۱-۱۳۲.
- ۱۲- عساکره، حسین. ۱۳۸۶. بررسی آماری روند دمای سالانه تبریز، اندیشه جغرافیایی، ۱: ۹-۲۱.
- ۱۳- عساکره، حسین. ۱۳۹۰. مبانی اقلیم شناسی آماری. انتشارات دانشگاه زنجان. ۵۴۵.
- ۱۴- عبدالله نژاد، کامل. ۱۳۹۴. مدل های تصادفی سری زمانی در پیش بینی بارندگی ماهانه مطالعه موردی: ایستگاه هاشم آباد گرگان. مجله آمایش جغرافیایی فضا. ۱۷: ۱۵-۲۵.
- ۱۵- علیجانی، بهلول. رضائی، نبی الله. ۱۳۸۱. پیش بینی خشکسالی ها و ترسالی های استان مازندران با استفاده از مدل باکس- جنگیز. پژوهش های جغرافیایی. ۱۵۵-۱۷۰.
- ۱۶- غفوریان، هادی. ثنائی نژاد، سید حسین. جباری نوقابی، مهدی. محمد پور، محمدعلی. ۱۳۹۲. بررسی دقت تکنیک SARIMA در پیش بینی مقادیر بارش مطالعه موردی: شهر مشهد. پنجمین کنفرانس مدیریت منابع آب ایران. تهران - دانشگاه شهید بهشتی. ۱-۱۶.
- ۱۷- کرایر، جانان دی، ترجمه: حسینعلی نیرومند. ۱۳۷۱. تجزیه و تحلیل سری های زمانی؛ مشهد؛ دانشگاه فردوسی مشهد. ۴۰۸.
- ۱۸- شریفان، حسین. بیژن قهرمان. ۱۳۸۶. ارزیابی پیش بینی باران با بکارگیری تکنیک SARIMA در استان گلستان. مجله علوم کشاورزی و منابع طبیعی، جلد چهاردهم، شماره سوم. ۱-۱۵.
- ۱۹- گلابی، محمدرضا. آخوندعلی، علی محمد. رادمنش، فریدون. کاشفی پور، محمد. ۱۳۹۳. مقایسه دقت پیش بینی مدل های باکس جنکینز در مدلسازی بارندگی فصلی (مطالعه موردی: ایستگاه های منتخب استان خوزستان). فصلنامه تحقیقات جغرافیایی. ۳: ۶۱-۷۲.
- ۲۰- یاراحمدی، داریوش. میریان، مینا. ۱۳۹۴. بررسی تغییرات و مدلسازی دمای سالانه ایستگاه سینوپتیکی تهران با استفاده از مدل سری های زمانی. نشریه تحقیقات کاربردی علوم جغرافیایی. ۳۸: ۱۶۵-۱۸۶.
- ۲۱- وئی ویلیام دبلیو اس، ترجمه نیرومند حسنعلی. ۱۳۷۶. تحلیل سری های زمانی، چاپ اول، مشهد: انتشارات دانشگاه فردوسی، ۵۸۶ صفحه.
- ۲۲- Aguilera, A M. Escabias, M; Valderrama, M J. (۲۰۰۷). Forecasting binary longitudinal data by a functional PC-ARIMA model, computational statistical & data analysis, ELSEVER. ۳۱۸۷-۳۱۹۷.
- ۲۳- Bouhaddou, H. Hassani, M.M. Zeroual, A. and Wilkonson, A.J. (۱۹۹۷). Stochastic simulation of weather data using higher statistics, Renewable Energy, ۸۳ (۸): ۳۸-۸۱.
- ۲۴- Barry, R.G. Carleton, A. M. ۲۰۰۱. Synoptic and Dynamic Climatology. Routledge, London. ۶۲۰ pp.
- ۲۵- Borland, P., and Montana, A. ۱۹۹۶. Forecasting of storm rain full by combined use of rider, rain gages and linear models. Atmospheric research, ۴۲: ۱۹۹-۲۱۶.
- ۲۶- Bolboacă, Sorana D. Jäntschi, Lorentz. Sestraş, Adriana F. Sestraş, Radu E. Pamfil, Doru C. ۲۰۱۱. Pearson-Fisher Chi-Square Statistic Revisited. Information. ۲: ۵۲۸-۵۴۵.
- ۲۷- Chbouki, N. Stockton, CW. Myers, DE. ۱۹۹۵. Spatio - temporal patterns of Drought in Morocco. International journal of climatology. ۱۵: ۱۸۷-۲۰۵.
- ۲۸- Chantasorn, Umaporn. ۲۰۱۱. Efficiency Comparisons of Normality Test Using Statistical Packages. Thammasat. ۳: ۹-۲۵.
- ۲۹- Durdu, o. f. ۲۰۱۰. A hybrid neural network and ARIMA model for water quality time series prediction. Engineering Applications of Artificial Intelligence. ۲۳: ۵۸۶-۵۹۴

- ۳۰- Frausto, H.U., Pieters, J.G. and Deltour, J.M. (۲۰۰۳). Modelling greenhouse temperature by means of Auto Regressive Models, *Biosystems Engineering*, ۳۹(۳۴):۸۹۱-۱۱۱.
- ۳۱- Haltiner G.J and Williams R.T (۱۹۸۰) Numerical Prediction and dynamic meteorology ۲nd Edition. New York Wiley & Sons pp.۴۴۷.
- ۳۲- J.L, Romeu. ۲۰۰۳. Anderson- Darling: A Goodness of Fit Test for Small Samples Assumptions. *Cinii journal*. ۱۰:۱- ۶.
- ۳۳- Padilla, A., Puldo-Bosch, A., Cavache, M., and Vallejos, A. ۱۹۹۶. The ARMA model applied to the flow of Karst Spiring. *Water resources bulletin*, ۳۲: ۹۱۷-۹۲۸.
- ۳۴- Khaliq M.N., Ouarda T.B., and Gachon P. ۲۰۰۹. Identification of temporal trend in annual and seasonal low flows occurring in Canadian rivers: The effect of short-and long-term persistence. *Journal of Hydrology*, ۳۶۹: ۱۸۳-۱۹۷.
- ۳۵- Tokar, A.S., Santon,P.A.,(۱۹۹۹): Rainfall-Run off modeling using artificial neural networks, *journal of Hydrologic Engineering*,No.۳,pp.۲۳۲-۲۳۳.
- ۳۶-

### ARIMA modeling for annual precipitation in MashhadAtoosa Khojasteh

#### **Abstract:**

Following the global warming phenomenon, the general circulation pattern and the temporal-spatial pattern of precipitation have also changed. The use of statistical methods in describing changes is a useful and effective tool. Since precipitation as a climatic element; has non-linear behavior, does not follow the normal distribution. As a result, the usage of modeling to understand the general behavior of precipitation is very efficient. There are various methods for modeling precipitation changes, including linear modeling (least absolute of deviations, least squares error method, etc.) and nonlinear modeling (polynomial modeling, artificial neural network). For the present study, the precipitation data of Mashhad airport station during the statistical period of ۵۷ years (۱۹۶۰-۲۰۱۶) have been used to identify the long-term precipitation behavior of this station. In order to study the general behavior of Mashhad precipitation, it was tried to fit the appropriate pattern of the polynomial family and ARIMA modeling using MINITAB software. The result of modeling a polynomial family indicates a quadratic equation, parabolic pattern in Mashhad precipitation. On the other hand, in the ARIMA model family, the ARIMA  $(\varepsilon, 1, 1)$  model, which was relatively better than the other models, was determined as a suitable model. Also, by performing the box-cox conversion, the converted precipitations were evaluated again, the candidate models and their predictions were made, and finally, the model selected for Mashhad precipitation was the ARIMA  $(0, 1, 1)$  model. It should be noted that based on the results of modeling and model prediction outputs, it can be inferred that linear ARIMA is not a suitable model for fitting Mashhad precipitation. On the other hand, due to the unsuitability of this model for Mashhad precipitation, the selected models do not match the trend in the observations, in other words, the order of differencing of the models is equal to one, which indicates a linear trend and not a parabola in the observations.

**Key Words:** Precipitation modeling, Mashhad, ARIMA, polynomial modeling, second degree trend .