

بررسی و پیش‌بینی متوسط بارندگی ماهانه با به‌کارگیری تکنیک SARIMA (مطالعه موردی: ایستگاه باران‌سنجی بابامان بجنورد)

محبوبه حاجی بیگلر^{۱*}، عباسعلی قزل سافلر^۲ و حسین علمیرزایی^۳

* نویسنده مسئول، عضو گروه پژوهشی کاوش، شرکت کاوش پی مشهد، Hajibigloo_m@yahoo.com

۲- استادیار گروه عمران، دانشگاه آزاد اسلامی مشهد

۳- معاونت طرح و توسعه شرکت آب منطقه ای خراسان شمالی

تاریخ پذیرش: ۹۱/۹/۲۷

تاریخ دریافت: ۹۱/۲/۲

چکیده

محدودیت منابع آبی جهت تأمین آب مورد نیاز کشاورزی و غیرکشاورزی موجب بروز مشکلات عمده‌ای شده است و باران به عنوان یکی از مهمترین منابع آبی موجود محسوب می‌شود. بنابراین، پیش‌بینی بارندگی و برآورد میزان آن در هر ماه و یا سال و برای هر منطقه و حوضه آبریز، به عنوان یکی از مهم‌ترین پارامترهای جوی، از اهمیت ویژه‌ای در استفاده بهینه از منابع آبی برخوردار است. برای پیش‌بینی باران می‌توان از سری‌های زمانی استفاده نمود. هدف از انجام این تحقیق بررسی مناسب‌ترین مدل برآورد باران می‌باشد، به طوری که با به‌کارگیری باران‌های ماهانه دوره ۳۰ ساله (۸۰-۱۳۵۰) و پس از تعیین پارامترهای مدل غیرفصلی و فصلی مدل SARIMA و با استفاده از نرم‌افزار Minitab مقادیر باران ماهانه ده سال آخر دوره آماری دوره (۹۰-۱۳۸۱) در ایستگاه باران‌سنجی بابامان واقع در استان خراسان شمالی تخمین زده شد. سپس مقادیر باران ماهانه پیش‌بینی شده بر اساس بهترین توزیع آماری، محاسبه گردید. با مقایسه مقادیر برآوردی باران ماهانه با مقادیر واقعی متناظر مشاهده شد برای الگوهایی که دارای پارامتر بیشتری هستند و مرتبه‌های اتورگرسیو یا میانگین متحرک آنها بیشتر از یک می‌باشد، برای سال‌های بعد مقادیر متفاوتی را نشان می‌دهد، اما این تفاوت‌ها نیز محدود به چند سال است که حداکثر از تعداد پارامترهای الگو تجاوز می‌نماید. ضریب همبستگی بین مقادیر واقعی و پیش‌بینی شده در ایستگاه مورد مطالعه ۰/۶۴+ به دست آمد که از رابطه رگرسیونی به دست آمده می‌توان برای تصحیح مقادیر متوسط بارندگی پیش‌بینی شده در ایستگاه استفاده نمود.

کلید واژه‌ها: سری زمانی، پیش‌بینی باران، مدل SARIMA، آزمون آکائیک.

Discussion and Forecast Monthly Average Rainfall Techniques Using SARIMA (Case study: Pluviometry Station Babaaman Bojnourd)

M. Hajibigloo¹, A. A. Ghezelsofloo² and H. alimirzai³

1- Member of Kavosh Research Group on Water Resources Management

2- Assistant Professor, IAUM, Mashhad, Iran

3- Research and Development Assistance in North Khorassan Regional Water Company

Received: 22.April.2012

Accepted: 18.December.2012

Abstract

Limited water resources needed for agricultural and non-agricultural water supply has caused major problems. Rain is considered as one of the available water resources. Therefore, to predict and estimate the amount of rainfall in any month or year and for each catchment area as one of the most important atmospheric parameters, of particular importance is the efficient use of water resources. For predict rain can be used of the time series. The aim of this study is the most appropriate model to estimate the rain, so that using the 30-year (1971-2001) monthly rainfall and after determining the model parameters and seasonal and non-seasonal SARIMA model and using the statistical software Minitab end of the period of ten years of monthly rainfall amounts (2002-2011) in the rain stations - located in North Khorasan Province Babaaman survey were estimated. The monthly rainfall amounts predicted by the statistical distribution, was calculated. By comparing the estimated values with actual values corresponding monthly

rainfall was result of models with more are parameters the order autoregressive or moving average is more than 1 shows different values for the following years. But these differences are also limited to a few years to exceed the maximum number of model parameters. The correlation coefficients between actual and predicted values at station 0.64 were studied. The regression equation obtained can be used to correct moderate amounts of rainfall stations used in forecasting.

Keywords: Time series, Rain forecast, SARIMA model, AIC.

مقدمه

متوسط بارندگی کشور، حدود ۲۵۰ میلی‌متر یعنی کمتر از یک سوم متوسط بارندگی جهان می‌باشد که از نظر پراکنش زمانی و مکانی نیز از شرایط نامطلوبی برخوردار است. از نظر توزیع مکانی، میانگین بارندگی در نقاط مختلف کشور بین ۵۰ میلی‌متر در بخش وسیعی از حاشیه‌ی کویر مرکزی و نواحی جنوب شرقی تا ۱۶۰۰ میلی‌متر در نواحی شمال و ارتفاعات غرب کشور متفاوت است، به طوری که ۷۳ درصد سطح کشور بارندگی کمتر از ۲۵۰ میلی‌متر، ۲۶ درصد بالای ۲۵۰ میلی‌متر و یک درصد سطح کشور بالای ۱۰۰۰ میلی‌متر، بارندگی دارد. از نظر توزیع زمانی نیز وضعیت مطلوب نیست به طوری که اغلب بارش‌های نواحی خشک و نیمه خشک به صورت رگبار در چند روز محدود از سال می‌بارد و دامنه‌ی نوسانات بارندگی در طول یک دوره آماری بسیار بالاست، به همین دلیل تقریباً همه ساله در مناطقی از کشور شاهد کم آبی و چالش‌های ناشی از خشکسالی هستیم (احمدی، ۱۳۸۴). پیش‌بینی بارندگی و برآورد میزان آن در هر ماه و یا سال و برای هر منطقه و حوضه آبریز، به عنوان یکی از مهم‌ترین پارامترهای جوی، از اهمیت ویژه‌ای در استفاده بهینه از منابع آبی برخوردار است. آگاهی از مقدار بارندگی در هر منطقه، به عنوان منبع اصلی تأمین‌کننده آب‌های سطحی و زیرزمینی، برای برنامه‌ریزی در امور مختلف زندگی اجتماعی انسان حائز اهمیت است. از جمله این امور، می‌توان به تأمین آب برای مصارف شرب و بهداشتی شهرها و روستاها، تأمین آب برای آبیاری مزارع و باغها در فعالیت‌های کشاورزی و تأمین آب برای واحدهای صنعتی و تجاری اشاره کرد (نیرومند، ۱۳۷۸).

سری زمانی به مجموعه‌ای از دیده‌بانی‌ها و یا مقادیر ثبت شده از یک متغیر گفته می‌شود که برحسب زمان مرتب شده باشد. هدف از سری زمانی، تعیین قانون‌مندی و شناسایی رفتار آن جهت پیش‌بینی در آینده می‌باشد. در مباحث آب‌شناختی همانند بارندگی و رواناب با فرآیندهای تصادفی سر و کار داریم. مجموعه متغیرهای فرآیند تصادفی ممکن است وابسته و یا مستقل از هم باشند. اگر فقط مقادیر غیر صفر مدنظر باشد، سری از نوع غیر متناوب و در غیر این صورت متناوب می‌باشد. باران‌های ساعتی، روزانه، ماهانه و یا سالانه از این نوع سری هستند. از طرف دیگر اگر قوانین احتمال حاکم بر فرآیند در طول زمان تغییر نکند، سری از نوع ایستا می‌باشد و این ایستایی می‌تواند در میانگین، واریانس و کوواریانس تعریف شود (شریفان و قهرمان، ۱۳۸۶). به طور نمونه در مباحث هیدرولوژی، تغییرات جوی زمین در یک دوره

نامعلوم مانند وقوع سال‌های پر باران و به دنبال آن خشکسالی‌های پی در پی موجب می‌شود تا سری زمانی بارندگی نا ایستا شود. یکی از شرایط اولیه استفاده از داده‌ها در مباحث سری زمانی، ایستا بودن آن‌هاست، در غیر این صورت باید نالیستایی رفع شود. برای ایستا نمودن داده‌ها در میانگین از روش تفاضلی و برای تبدیل پایداری در واریانس از روش باکس-کاکس (۱۹۶۴) می‌توان استفاده نمود. جهت مدل‌سازی با سری‌های زمانی، در اغلب موارد متخصصان از انواع متنوع مدل‌های ریاضی و آماری بهره می‌گیرند به طوری که همستگی‌های موجود ما بین زمان و مشاهدات مد نظر می‌باشد (نیرومند و بزرگ نیا، ۱۳۷۲). مدل‌های سری زمانی عبارتند از:

۱- مدل تصادفی خود همبسته AR(p): اساس این مدل بر پایه زنجیره مارکف در زنجیره زمانی بنا شده است. یکسری زمانی از زنجیره مارکف تبعیت می‌کند، اگر هر داده با زمان قبل و یا زمان t ثبت شده سری زمانی در زمان بعد از خود مرتبط باشد.

۲- مدل میانگین متحرک MA(q): در این مدل متغیر در زمان t از روی مقدار تصادفی همان لحظه به علاوه q برابر مقدار تصادفی مربوط به زمان‌های قبل از t برآورد می‌شود.

۳- مدل خود همبسته - میانگین متحرک ARMA(p,q): هرگاه دو مدل قبلی در یکدیگر ادغام شوند، مدل آرما^۱ با مرتبه‌های p و q تصادفی مربوط به زمان‌های قبل از t برآورد می‌شود (معادله ۱):

$$Z_t = \mu + \sum_{i=1}^p (z_{t-i} - \mu) - \sum_{i=1}^q \theta_i a_{t-i} + a_t \quad (1)$$

که در آن Z_t : مقدار پیش‌بینی شده، z_{t-i} : اطلاعات مربوط به گذشته سری، μ : میانگین سری، θ : ضریب مدل، a_{t-i} : به ترتیب اغتشاش در حال و گذشته می‌باشد.

۴- مدل خودهمبسته - میانگین متحرک تلفیق شده ARIMA(p,d,q): از آنجا که برای استفاده از مدل‌های فوق باید فرآیند ایستایی برقرار باشد، از این رو باکس و همکاران^۲ (۱۹۹۴) در شرایط نالیستایی، مدل آرما^۱ را با در نظر گرفتن مرتبه تفاضلی d ارائه نمودند.

1- Box-Cox
2- ARMA
3- Box et al.
4- ARIMA

رسولی (۱۳۸۱) برای پیش‌بینی دمای ماهانه شهر تبریز اشاره نمود.

خردمندنیا و عساکره (۱۳۸۰)، برای پیش‌بینی درجه حرارت متوسط ماهانه منطقه جاسک از مدل‌سازی ساریما^۹ استفاده نمودند. آشگرطوسی (۱۳۸۲)، با استفاده از سری‌های زمانی، خشکسالی منطقه شیروان را در استان خراسان پیش‌بینی نمود و براساس نتایج به دست آمده، بهترین الگوی کشت را پیشنهاد کرد. احمدی (۱۳۸۴)، در تحقیقی مقدار باران سالانه ایستگاه‌های هواشناسی استان خراسان را با استفاده از سری‌های زمانی و مدل ساریما پیش‌بینی نمود.

در این تحقیق سعی شده است با استفاده از آمار بارندگی سال‌های قبل و به‌کارگیری روش سری‌های زمانی، مناسب‌ترین روش برای پیش‌بینی بارندگی ماهانه در ایستگاه مورد مطالعه ارائه گردد. بنابراین هدف از این تحقیق ارائه مناسب‌ترین مدل براساس ضریب آکائیک^{۱۰} جهت پیش‌بینی بارش‌های ماهانه ایستگاه مورد نظر با دو نگرش برآورد مستقیم و برآورد بارش ماهانه می‌باشد.

مواد و روش‌ها

در این تحقیق به منظور پیش‌بینی بارندگی از داده‌های بارندگی ماهانه ایستگاه باران‌سنجی بابامان در طول دوره آماری ۵۰-۱۳۴۹ الی ۹۰-۱۳۸۹ استفاده شد. این ایستگاه در موقعیت جغرافیایی ۵۷ درجه، ۲۸ دقیقه طول شرقی و ۳۷ درجه، ۲۸ دقیقه عرض شمالی و ارتفاع ۱۰۲۰ متر از سطح دریاهای آزاد قرار گرفته است. در شکل (۱) موقعیت ایستگاه باران‌سنجی بابامان در خروجی حوضه شاه‌جوب در استان خراسان شمالی نشان داده شده است.

پس از اطمینان از صحت داده‌های بارندگی ماهانه و انجام آزمون‌های آماری لازم، به پیش‌بینی بارندگی پرداخته شد. برای این منظور از ۴۹۰ داده موجود ۳۷۰ داده برای آموزش سیستم و پیدا کردن مدل مناسب استفاده شد و ۱۲۰ داده بعدی برای آزمون مدل به‌دست آمده به‌کار گرفته شد.

مراحل ساخت سری زمانی: ساخت مدل‌های سری زمانی شامل چهار مرحله می‌باشد که به‌صورت تکراری انجام شد (ساپلیوگلو و همکاران، ۲۰۱۰):

۱- **مرحله شناسایی الگو:** در این مرحله با رسم نمودارهای خودهمبستگی^{۱۱} و خودهمبستگی جزئی^{۱۲} و آزمون بارتلت-لون^{۱۳} ایستایی در میانگین و واریانس داده‌ها مورد ارزیابی قرار گرفت. در صورت ناپایداری ابتدا سری مورد نظر با استفاده از

۵- مدل خودهمبسته - میانگین متحرک تلفیق شده فصلی SARIMA(p,d,q)(P,D,Q): هر گاه در یک سری بعد از هر فاصله زمانی مشخص (s)، شباهت‌هایی پیدا شود، سری دارای رفتار فصلی یا تناوبی با دوره تناوب S می‌شود. برای ساخت این مدل‌ها، چهار مرحله شناسایی مدل، برازش الگو، تشخیص درستی الگو و پیش‌بینی باید انجام گیرد.

در زمینه استفاده از سری‌های زمانی در مدل‌سازی پارامترهای هیدرولوژیکی همچون بارندگی، دما و جریان‌های رودخانه پژوهش‌های متعددی صورت گرفته است (آشگر طوسی، ۱۳۸۲). طی بررسی تحقیقات گذشته، دانشمندان به منظور تحلیل تغییرات پارامترهای اقلیمی تلاش نموده‌اند تا این پارامترها را الگوسازی و سپس شبیه‌سازی نمایند. الگوسازی در خانواده آرما، آریمای فصلی یکی از شیوه‌های مهم و معتبر در شبیه‌سازی پارامترهای اقلیمی است (باکس و کاکس، ۱۹۶۴).

پیش‌بینی بارندگی و شناسایی الگوی آن نسبت به زمان در برنامه‌ریزی و مدیریت آن نقش بزرگی داشته و محققین مختلفی با استفاده از روش‌های پیش‌بینی مختلف به پیش‌بینی آن پرداخته‌اند که می‌توان به تحقیقات ساپلیوگلو و همکاران^۱ (۲۰۱۰)، کارآموز و همکاران (۲۰۱۰)، کی کیم و همکاران^۲ (۲۰۰۴)، احمدی (۱۳۸۴)، هانسن و لیدف^۳ (۱۹۸۸)، بلوفیلد و نیچکا^۴ (۱۹۹۲) و فولاند^۵ (۱۹۹۰) اشاره نمود.

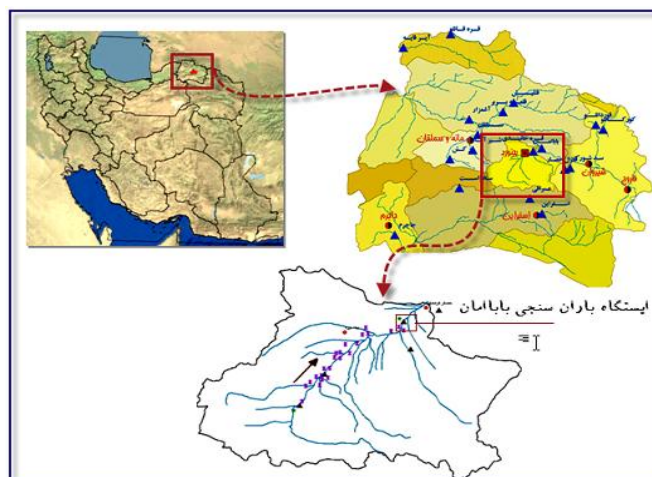
بورلاندو و همکاران^۶ (۱۹۹۶) از مدل‌های آریمای جهت پیش‌بینی همزمان بارندگی‌های ساعتی استفاده کرده و مقادیر به دست آمده را با داده‌های باران‌سنجی مقایسه کردند. آنها در تحقیق خود به این نتیجه رسیدند که با افزایش تداوم بارندگی، پیش‌بینی‌ها روند دقیق‌تری داشته و با کوتاه‌تر شدن دوام بارندگی، اختلاف میزان باران پیش‌بینی از مقدار واقعی متناظر خود بیشتر می‌شود.

هالتینر و سالاس^۷ (۱۹۸۸) از یک مدل فصلی (ARMA(1,1)) در مدل‌سازی دو متغیره جریان ماهانه در رودخانه‌های یامپا و وایت^۸ در شمال غربی کالیفرنیا استفاده کردند. آنها پارامترهای مدل را از دو روش حداکثر درست‌مایی و گشتاورها به دست آورده و با یکدیگر مقایسه کردند.

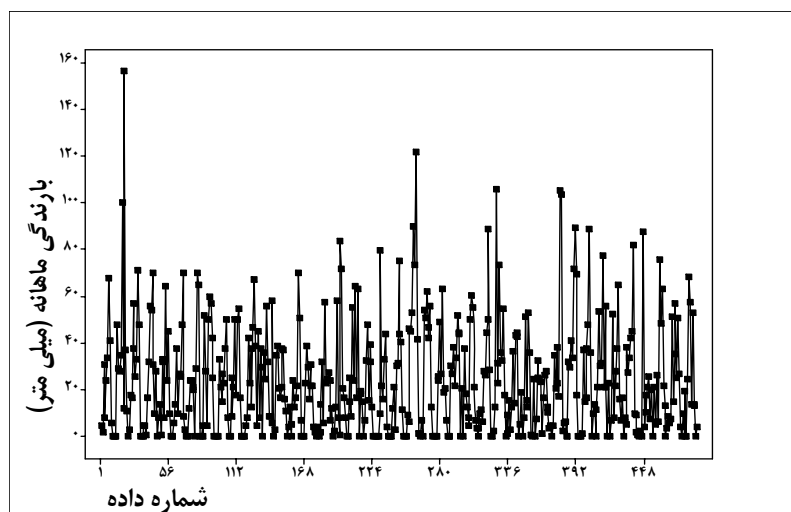
در کشور ایران به لحاظ سابقه کوتاه در بهره‌گیری از روش‌های اندازه‌گیری مستقیم داده‌های جوی، از این روش کمتر استفاده شده است (خردمندنیا و عساکره، ۱۳۸۰). در این میان می‌توان به کار جمشیدی (۱۳۶۸) در مدل دما-بارش ایستگاه تهران، مالکی (۱۳۶۸) در الگوسازی دما و بارش غرب کشور و

- 1- Saplioglu et al.
- 2- Ki kim et al.
- 3- Hansen and Lebedeff
- 4- Bloomfield and Nychka
- 5- Folland
- 6- Burlando et al.
- 7- Haltiner and Salas
- 8- Yampa and White rivers

- 9- SARIMA
- 10- AIC Coefficient
- 11- SAC Chart
- 12- SPAC Chart
- 13- Bartlett and lowen Test



شکل ۱- موقعیت ایستگاه باران‌سنجی بابامان در خروجی حوضه شاهجوب



شکل ۲- نمودار سری زمانی مقادیر بارندگی ماهانه در ایستگاه باران‌سنجی بابامان (۹۰-۱۳۴۹)

(SSE) و میانگین مربعات خطا^۲ (MSE) کوچکتری بود، انتخاب گردید.

۳- تشخیص درستی الگو: جهت بررسی درستی مدل، نمودار باقیمانده‌ها از نظر نرمال بودن و ایستایی مورد ارزیابی قرار گرفت.

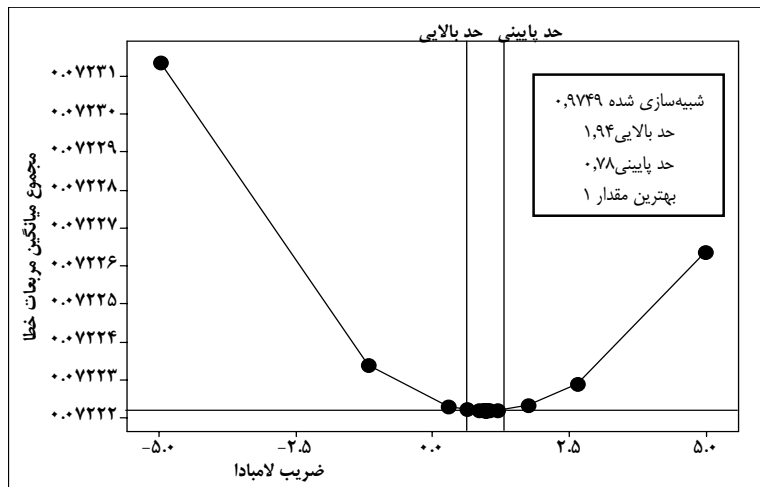
۴- پیش‌بینی: با استفاده از تبدیل باکس-کاکس، مقادیر سری داده‌های پیش‌بینی شده به مقادیر باران تصحیح شدند. نتایج به دست آمده به عنوان داده‌های نهایی باران پیش‌بینی شده برای سال‌های مورد نظر ارزیابی شد. جهت مدل‌سازی داده‌های بارندگی ایستگاه‌های فوق از نرم‌افزار Minitab استفاده گردید که اساس

سری تفاضلی مناسب و تبدیل داده‌ها از روش باکس - کاکس در میانگین و واریانس پایدار شده، سپس سری ایستا گردید. بنابراین، در این مرحله با تجزیه واریانس داده‌های تفاضلی شده، مرتبه پارامترهای d و D برای مدل طوری انتخاب شد که دارای حداقل واریانس باشد. از سوی دیگر با استفاده از نمودارهای خود-همبستگی جزئی و خودهمبستگی مراتب p ، q ، P و Q مشخص گردید (هالینتر و سالاس، ۱۹۸۸).

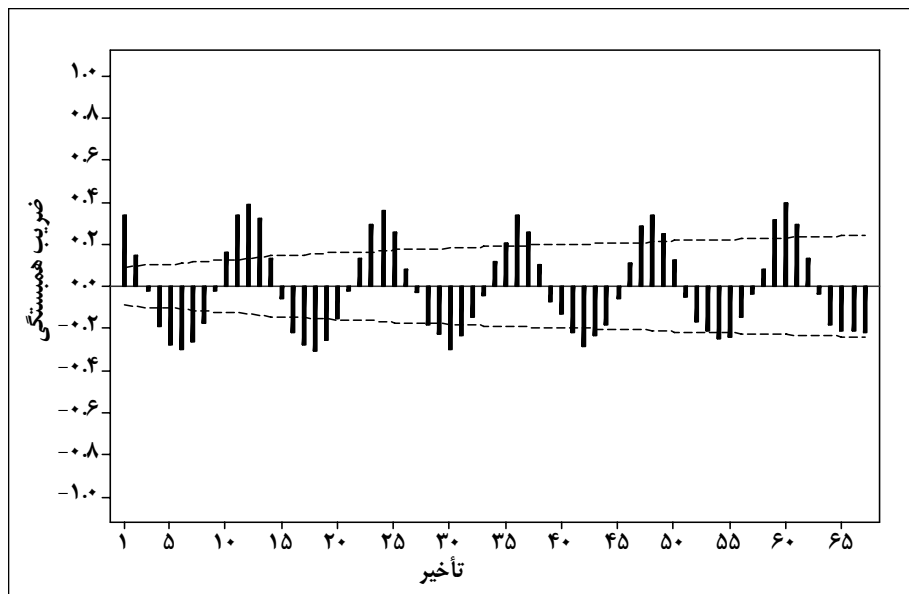
۲- برازش الگو (برآورد پارامترها): در این مرحله با شناسایی الگوهای مناسب در مرحله قبل، معنی‌دار بودن مقادیر t و p متناظر بررسی شد و مدلی که دارای مجموع مربعات خطا^۱ (SSE)

2- Mean of Squares Error

1- Sum of Squares Error



شکل ۳- نتایج تبدیل باکس-کاکس بر روی سری زمانی Y_t



شکل ۴- همبستگی نگار سری Y_t (داده‌های اصلی)

به منظور بررسی همگنی واریانس داده‌های سری زمانی از این آزمون استفاده می‌شود. اگر مقدار شاخص تصحیح آزمون بارتلت و لون^۲ کمتر از پنج درصد باشد داده‌ها ناپایستا در واریانس و غیرنرمال می‌باشند. داده‌های فوق نشان از ایستایی داده‌ها در واریانس و نرمال بودن داده‌ها دارد.

تبدیل باکس کاکس

در این تحقیق مقدار بهترین شاخص عملکرد تقریباً معادل یک می‌باشد که بیانگر ایستا بودن داده‌های اصلی در واریانس می‌باشد. در شکل (۳) تبدیل باکس کاکس را بر روی داده‌های اصلی نشان داده شده است.

کار آن همان شیوه تکراری باکس - جنکینز^۱ می‌باشد (فولاند، ۱۹۹۰).

مدل‌سازی و پیش‌بینی متوسط بارندگی ماهانه ایستگاه بابامان

رژیم بارندگی

در شکل (۲) نمودار سری زمانی بارندگی ماهانه در ایستگاه باران‌سنجی بابامان نشان داده شده است. نمودار فوق نشان می‌دهد که داده‌ها دارای روند کاهشی (ناایستا) با تغییرات فصلی هستند (در هر فصل از سال تقریباً رفتاری مشابه دیده می‌شود).

بررسی ایستایی واریانس در داده‌ها

2- Bartlet and Lowen

1- Box and Jenkins

تعیین درجه تفاضل فصلی و غیرفصلی

برای بررسی ایستایی داده‌ها و تعیین مرتبه‌های تفاضلی فصلی و غیر فصلی از نمودار خود همبستگی و خود همبستگی جزئی استفاده می‌شود. در رسم این نمودارها خط نقطه‌چین حاکی از همبستگی بین داده‌های سری زمانی در تأخیرهای زمانی مختلف است ($\pm 2/\sqrt{N}$) که آن را معمولاً با τ_k نشان می‌دهند که مقدار آن همیشه بین -1 و 1 می‌باشد که مقادیر نزدیک به 1 و -1 حاکی از وابستگی خطی زیاد بین مشاهدات و مقادیر

نزدیک به صفر نشان‌دهنده همبستگی ضعیف یا عدم همبستگی و (تصادفی بودن، ایستا و نرمال بودن) می‌باشد. ملاحظه می‌شود که این خط به آرامی به صفر نزدیک می‌شود که نشان از ناپیوستایی داده‌ها دارد. همچنین خط نقطه‌چین دارای نوساناتی می‌باشد که این نوسانات نشان از تغییرات فصلی در داده‌ها می‌باشد (باکس و کاکس، ۱۹۶۴). پس داده‌ها هم ناپیوستا و هم دارای تغییرات فصلی می‌باشند. شکل (۴) نمودار همبستگی-نگار سری داده‌های اصلی را نشان می‌دهد.

برای تعیین درجه تفاضل فصلی و غیرفصلی دو روش مهم را می‌توان انجام داد (احمدی، ۱۳۸۴):

۱- انجام تفاضل‌های (فصلی و غیرفصلی) مرتبه اول و دوم و ... و رسم نمودارهای خودهمبستگی

برای این منظور از نمودار خودهمبستگی استفاده می‌شود. با رسم این نمودار یکسری اطلاعات نیز به دست می‌آید که یکی از آنها آماره T است که به صورت زیر تعریف می‌شود:

تابع خودهمبسته/خطای استاندارد $T(t_{rk}) =$

$$t_{rk} = \tau_k / S_{\tau k} \text{ (for ACF)}$$

$$t_{rkk} = \Gamma_{kk} / S_{\Gamma kk} \text{ (for PACF)}$$

اگر تابع ACF سری زمانی در سطح غیرفصلی و در سطح فصلی قطع شود و یا به سرعت تنزل کند بایستی مقدار سری را ایستا در نظر گرفت. اگر در تأخیرهای بزرگ‌تر از k در ACF حداکثر وجود نداشته باشد، گفته می‌شود که ACF بعد از تأخیر k قطع می‌شود. برای تشخیص حداکثر در تأخیر k از آماره T به صورت زیر استفاده می‌شود:

برای سری‌های غیرفصلی (d) در تابع ACF

$$\text{For } k=1, k=2, k=3 \quad |t_{rk}| > 1.6$$

اگر برقرار نباشد تأخیرهای زیر بررسی می‌شود:

$$\text{For } k = 3, \dots, 9 \quad |t_{rk}| > 2$$

برای سری‌های فصلی (D) در تابع ACF

$$\text{For } k = L, 2L, 3L, 4L \quad |t_{rk}| > 1.25$$

اگر برقرار نباشد تأخیرهای زیر بررسی می‌شود:

$$\text{For } k = L \pm 2, L \pm 1, L, \dots \quad |t_{rk}| > 1.6$$

۲- روش دیگر برای تعیین مقادیر D, d کمینه کردن واریانس سری تفاضلی حاصل است. بنابراین با اختیار نمودن مقادیر مختلف برای D و d مقایسه برای $D=1$ و $d=1$ (الف) هر دو ضریب p و q در نظر گرفته می‌شود اگر ACF و PACF هر دو تنزل پیدا کنند. (ب) برای تعیین ضرایب P و Q در حالت فصلی هم موارد بالا بایستی رعایت شود.

$$Z_t = \nabla_L^D \nabla^d y_t = (1 - B^L)^D (1 - B)^d y_t = y_t - y_{t-1} - y_{t-L} + y_{t-L-1} \quad (2)$$

برای سری‌های فصلی و غیرفصلی PACF دارای حداکثر است اگر:

$$\text{For } k = 1, \dots, 9 \quad |t_{rk}| > 2$$

$$\text{For } k = L, 2L, 3L, 4L \quad |t_{rk}| > 2$$

تعیین مرتبه‌های p, q, P, Q

برای این منظور از نمودار ACF و PACF که حاصل سری داده‌های تفاضلی شده است استفاده می‌شود. از نمودار ACF برای تعیین درجه q و Q و از نمودار PACF برای تعیین درجه p و P استفاده می‌شود. در نتایج نمودارهای فوق بایستی موارد مطرح شده را بررسی نمود (خردمندنیو و عساکره، ۱۳۸۰).

۱- الگوی میانگین متحرک غیرفصلی از مرتبه q : ACF در تأخیرهای $1, 2, \dots, q$ دارای حداکثر است و بعد از تأخیر q قطع می‌شود و PACF نزولی است.

۲- الگوی اتورگرسیو غیرفصلی از مرتبه p : PACF در تأخیرهای $1, 2, \dots, p$ دارای حداکثر است و بعد از تأخیر p قطع می‌شود و ACF نزولی است.

۳- الگوی میانگین متحرک غیرفصلی از مرتبه q یا الگوی اتورگرسیو غیرفصلی از مرتبه p :

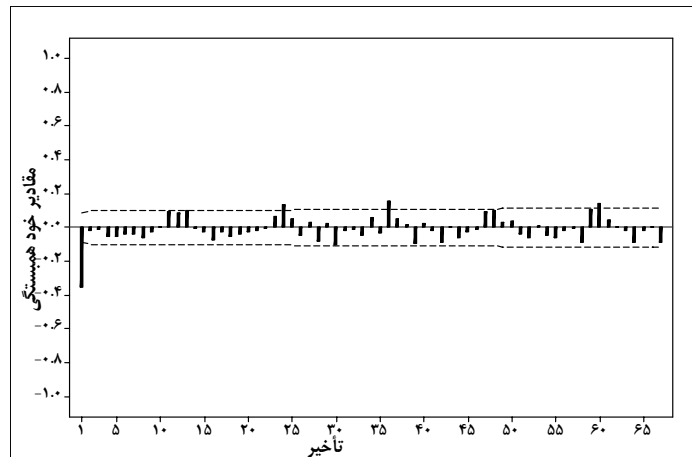
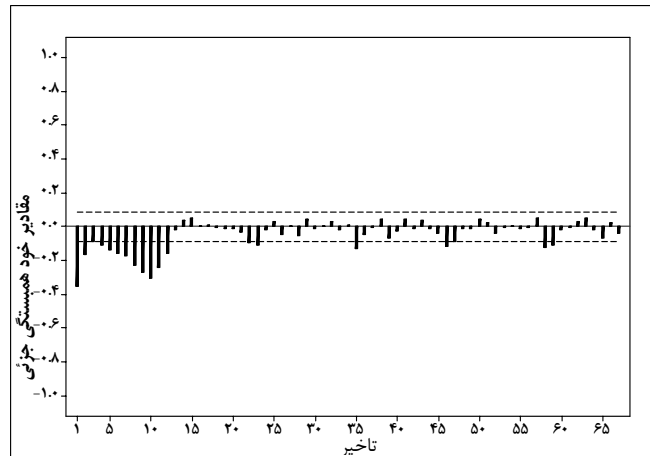
واریانس سری تفاضلی شده به ازای مقادیر مزبور می‌توان درجه تفاضلی را برای مراتب فصلی و غیرفصلی مشخص کرد (نیرومند، ۱۳۷۸).

با توجه به درجه‌های تفاضلی فصلی و غیرفصلی الگوی عمومی باکس و جنکینز به صورت زیر خواهد بود (باکس و کاکس، ۱۹۶۴):

$$Z_t = \nabla_L^D \nabla^d y_t = (1 - B^L)^D (1 - B)^d y_t$$

$$\{ \nabla = 1 - B, \nabla^L (1 - B^L) \} \quad (3)$$

در این حالت مانند دو روش فوق درجه q و p تعیین می‌شود. اگر ACF با سرعت و شدت بیشتری نسبت به PACF قطع شود از q و در حالت برعکس از p استفاده می‌شود. در صورتی که سرعت و شدت قطع‌شدگی هر دو یکسان باشد یکبار از q و یکبار از p در

شکل ۵- همبستگی نگار سری زمانی Z_1 (سری تفاضلی)شکل ۶- همبستگی نگار جزئی سری زمانی Z_1 (سری تفاضلی)

غیرفصلی از مرتبه $q=2$ و یا اتورگرسیون از مرتبه $p=2$ در مدل منظور شود. چون شدت قطع‌شدگی (ACF و PACF) هر دو یکسان است بنابراین یک بار از عملگر q و یک بار هم از عملگر p در مدل استفاده می‌شود.

در سطح فصلی

تابع ACF در تأخیر ۱۲ ($|T|=1.91 > 1.25$) و ۴۷ ($|T|=2.52 > 1.6$) دارای حداکثر بوده و سپس قطع می‌شود و همچنین تابع PACF در تأخیرهای ۱۲، ۲۴، ۳۶، ۴۸ دارای حداکثر بوده و تابع به صورت ($T=3.54, 2.46, 2.92$) و سپس قطع می‌شود بنابراین مانند حالت قبل ($2 > 2.23$) و سپس قطع می‌شود بنابراین مانند حالت قبل $P=4$ و $Q=4$ از هر دو عملگر در مدل استفاده می‌شود و در نهایت هر کدام که جواب بهتری داشت به عنوان مدل نهایی استفاده می‌شود.

مدل استفاده می‌شود هر کدام که جواب مناسب‌تری داشت مورد استفاده قرار می‌گیرد.

۴- هیچ نوع عملگر غیرفصلی نیاز نیست در صورتی که ACF و PACF هیچ‌کدام در هیچ تأخیری دارای حداکثر نیستند بلکه دارای مقادیر کوچک‌تر خود همبستگی هستند.

همبستگی نگار و همبستگی نگار جزئی سری تفاضلی شده (Z_1) حاصل تفاضل $D=1$ و $d=1$ بر

روی سری (Y_t) به ترتیب در شکل‌های (۵) و (۶) نمایش داده شده است.

در سطح غیرفصلی

تابع ACF در تأخیر ۲ دارای حداکثر بوده و سپس قطع می‌شود ($|T|=7.85, 2.31 > 1.6$) و همچنین تابع PACF در تأخیر ۲ دارای حداکثر بوده و سپس قطع می‌شود ($T=7.85$)، بنابراین باید یکی از عملگرهای میانگین متحرک ($3.57 > 3$)،

$$AIC = n \ln(MSE) + 2M \quad (۴)$$

در رابطه فوق M : تعداد پارامترهای مدل، MSE : مجموع مربعات خطاها و n : تعداد مشاهدات می‌باشد.

چنانچه با استفاده از معیار آکائیک مدلی از بین چند مدل رقیب انتخاب شود، مقدار AIC متناظر مدل انتخابی باید کمینه باشد. مدل آزمایشی اولیه به هر شکلی که تعیین شود باید به این نکته توجه داشت که مدل تعیین شده، مورد کنترل‌های همه جانبه قرار گیرد. این کنترل‌ها شامل نرمال و مستقل بودن و برازش مقادیر می‌باشد.

بررسی مناسبت مدل

از دو روش که مکمل یکدیگر هستند استفاده می‌شود (احمدی، ۱۳۸۴):

۱- تجزیه و تحلیل باقیمانده‌های مدل برازش داده شده (تصادفی یا ناهمبسته بودن باقیمانده‌ها اثبات می‌شود).

۲- تجزیه و تحلیل مدل‌هایی که پارامتر بیشتری دارند.

در صورتی که چند مدل مناسب تشخیص داده شده باشد از معیار آکائیک استفاده می‌شود.

در تجزیه و تحلیل باقی‌مانده‌های مدل برازش داده شده، فرض نرمال بودن باقیمانده‌ها، فرض ثابت بودن واریانس باقیمانده‌ها، فرض استقلال باقیمانده‌ها، نمودار باقیمانده‌ها در برابر زمان، آزمون پرت - مانتو^۱ برای هر یک از مدل‌ها مورد بررسی قرار می‌گیرد.

فرض نرمال بودن باقی مانده‌ها در صورتی پذیرفته می‌شود که نقاط تقریباً در امتداد یک خط راست قرار گرفته و از توزیع یکنواختی برخوردار باشد.

به عنوان یک روش رسمی‌تر برای آزمون فرضیه ناهمبسته بودن باقی‌مانده‌ها از آزمون پرت مانتو که بر مبنای آماره اصلاح شده باکس - پیروسن^۲ می‌باشد، استفاده می‌شود.

آزمون پرت مانتو به صورت زیر می‌باشد (جمشیدی، ۱۳۶۸):

$$Q(LBQ) = n(n+2) \sum_{h=1}^k (n-h)^{-1} \rho_h^2 \quad (۵)$$

n : تعداد مشاهدات، Q : آماره آزمون که اصلاح شده آن LBQ لجانگ باکس^۳ است. تحت فرض H_0 تقریباً دارای توزیع کیدو^۴ است.

شرط اول: هرگاه مقدار آماره Q از مقدار متناظر در جدول کیدو بیشتر باشد فرض H_0 رد می‌شود یعنی داده‌ها همبسته هستند.

شرط دوم: همچنین بایستی مقدار شاخص تصحیح از مقدار α که معمولاً برابر پنج درصد است بزرگ‌تر باشد. نتایج مربوط به آزمون

نتایج و بحث

در این تحقیق، برای انتخاب مدل مناسب جهت پیش‌بینی داده‌های ماهیانه آینده، در اولین مرحله، ایستا کردن مقادیر سری زمانی با توجه به نمودار خودهمبستگی نمونه مربوط به مشاهدات صورت گرفت. برای از بین بردن تغییرات زمانی موجود در سطح مینا، با استفاده از نرم افزار Minitab عمل تفاضل‌گیری با مرتبه d انجام گرفت تا بتوان مدل را با تبدیل‌هایی ایستا نمود. مقدار d مرتبه تعداد دفعاتی را که برای ایستا کردن سری زمانی لازم است را نشان می‌دهد.

پس از ایستا کردن سری زمانی با دو بار تفاضل‌گیری مدل آزمایشی می‌توان از نمودار همبستگی نگار و همبستگی نگار جزئی استفاده کرد

اگر فرض شود که $Y_{(k)}$ تابع خودهمبستگی نمونه مربوط به سری Y_t باشد، مقادیر p و q را می‌توان طوری تعیین کرد که $r(s) = r(1), r(2), r(3), \dots, r(d)$ به طور معقول منطبق بر تابع خود همبستگی نگار فرایند (d, p) باشند. سپس نمودار تابع خود همبستگی نمونه مربوط به سری (Y_t) در تأخیرهایی که مضرری از S هستند، اعمال می‌شود. مقادیر Q و p طوری تعیین می‌شوند که $r(s), r(2s), r(3s), \dots$ به طور معقول منطبق بر تابع خودهمبستگی نگار فرایند (Q, P) باشند تا بدین ترتیب با روش فوق یک مدل ترکیب شده از دو بخش غیر فصلی و فصلی $SARIMA(p,d,q)(P,D,Q)_{12}$ به طور آزمایشی انتخاب شود. در واقع در این مرحله مدل آزمایشی از لحاظ درستی فرض‌ها و مناسبت آن از طریق نمودار همبستگی نگار و همبستگی نگار جزئی مورد کنترل قرار می‌گیرد.

الگوسازی

مرتبه‌های $D=4$ و $d=2$ با داشتن کمترین مقدار واریانس سری، تفاضل‌گیری و انتخاب گردیده و با ترسیم خودهمبستگی خودهمبستگی جزئی سری زمانی، با تفاضل‌گیری فصلی $(D=4)$ ، مقادیر $(4,0,1)$ و فصلی $(0,1,2)$ ، به عنوان الگوی اولیه مشخص گردیده، سپس با برازش و اصلاح الگوی اولیه، مدل‌های نهایی آریمای فصلی به صورت زیر تعریف شده است:

$SARIMA(2,1,0)*(4,1,0)_{12}$ (الگوی M_1)

$SARIMA(0,1,2)*(4,1,0)_{12}$ (الگوی M_2)

$SARIMA(2,1,0)*(0,1,4)_{12}$ (الگوی M_3)

$SARIMA(0,1,2)*(0,1,4)_{12}$ (الگوی M_4)

مدل آزمایشی در مرحله بعد باید از لحاظ درستی فرض‌ها و مناسبت آن مورد کنترل قرار گیرد.

تعیین رسته مدل با استفاده از آزمون معیار آکائیک صورت گرفته است. این آزمون که در سال ۱۹۷۴ توسط آکائیک ارائه شده بر این اصل استوار است که هر رسته AIC کمتری داشته باشد، برازش بهتری بر سری مزبور خواهد داشت. مقدار AIC برای مدل مورد نظر از رابطه زیر به دست می‌آید:

1- Pert-Manto test

2- Box-Pierson

3-Box Logang

4- χ^2 Distribution

جدول (۱) آورده شده است. بنابراین فرضیه ثابت بودن واریانس باقیمانده‌ها را می‌توان پذیرفت. (شکل ۷-د) نیز طرح خاصی را نشان نمی‌دهد و دارای پراکندگی افقی حول صفر با واریانس ثابت می‌باشد. بنابراین به نظر می‌رسد باقیمانده‌های حاصل از برازش مدل مشابه یک فرآیند اغتشاش خالص رفتار می‌کنند که نشان‌دهنده مناسبت مدل برازش شده است.

ماتریس همبستگی

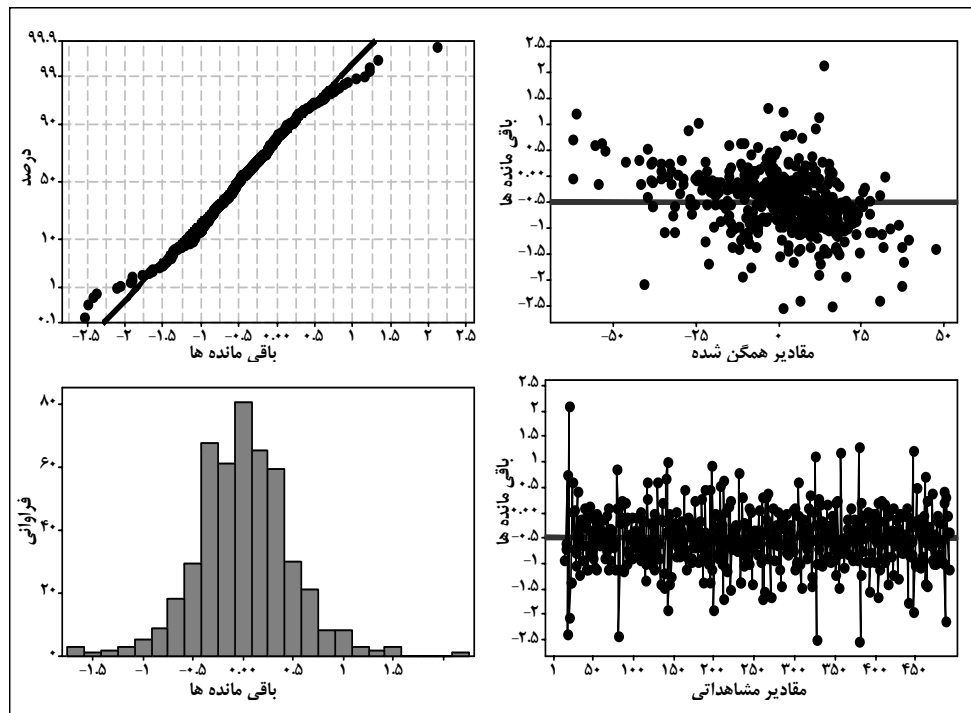
با توجه به اینکه برآورد پارامترهای مدل انتخابی از نتیجه انجام عملیات مختلف بر روی داده‌های سری مورد مطالعه به دست می‌آیند، بنابراین به طور طبیعی پارامترهای برآورد شده همیشه مقداری با هم همبستگی دارند. مقدار همبستگی بین پارامترهای مدل به وسیله نرم‌افزار Minitab محاسبه می‌گردد و نتایج به صورت یک ماتریس که به "ماتریس همبستگی بین پارامترها" مشهور است، نمایش داده می‌شود (باکس و کاکس، ۱۹۶۴).

پرت-مانتو برای هر یک از مدل‌ها انجام شد و نتایج مربوط در در الگوهای شناسایی شده برای کلیه تأخیرها مقدار شاخص تصحیح بزرگتر از پنج درصد بوده که فرضیه ناهمبسته بودن باقیمانده‌ها را مورد تأیید قرار می‌دهد. همچنین با درجه آزادی‌های مختلف مقدار آماره کلیه آزمون‌ها در سطح پنج درصد از کیدو جدولی برای الگوهای M_1 و M_2 بزرگتر است پس فرض صفر رد و کفایت مدل رد می‌شود و برای الگوهای M_3 و M_4 مقدار آماره کلیه آزمون‌ها در سطح پنج درصد از کیدو جدولی کوچکتر است پس فرض صفر پذیرفته و کفایت مدل پذیرفته می‌شود. در شکل (۷) نمودار باقیمانده‌های مدل برازش داده شده نشان داده شده است.

در نمودار احتمال نرمال (شکل ۷-الف) نقاط تقریباً در امتداد یک خط راست قرار دارند. در نمودار هیستوگرام (شکل ۷-ب) نیز باقیمانده‌ها به صورت نرمال توزیع شده‌اند. بنابراین می‌توان فرض نرمال بودن توزیع باقیمانده‌ها را پذیرفت. با توجه به اینکه (شکل ۷-ج) طرح خاصی را نشان نمی‌دهد (به شکل کیفی)

جدول ۱- نتایج آزمون پرت-مانتو بر روی الگوهای شناسایی شده

الگو	رابطه الگو	نتایج محاسبات آماری آزمون لجانگ-باکس
M_1	$(2,1,0)*(4,1,0)_{12}$	تأخیر ۱۲ ۲۴ ۳۶ ۴۸ کای اسکوتر ۴۴/۱ ۵۱/۷ ۷۴/۳ ۸۴/۱ درجه آزادی ۵ ۱۷ ۲۹ ۴۱ شاخص تصحیح ۰/۰۹۴ ۰/۰۷۳ ۰/۰۸۵ ۰/۰۷۲
M_2	$(0,1,2)*(4,1,0)_{12}$	تأخیر ۱۲ ۲۴ ۳۶ ۴۸ کای اسکوتر ۲۰/۹ ۳۴/۸ ۳۹/۶ ۵۹/۲ درجه آزادی ۵ ۱۷ ۲۹ ۴۱ شاخص تصحیح ۰/۰۶۰ ۰/۶۵۵ ۰/۷۹۵ ۰/۸۳۲
M_3	$(2,1,0)*(0,1,4)_{12}$	تأخیر ۱۲ ۲۴ ۳۶ ۴۸ کای اسکوتر ۷/۹ ۱۵/۷ ۱۸/۸ ۲۶/۸ درجه آزادی ۵ ۱۷ ۲۹ ۴۱ شاخص تصحیح ۰/۱۸۷ ۰/۷۱ ۰/۳۵۶ ۰/۴۴۵
M_4	$(0,1,2)*(0,1,4)_{12}$	تأخیر ۱۲ ۲۴ ۳۶ ۴۸ کای اسکوتر ۴/۰ ۸/۹ ۲۲/۲ ۲۹/۳ درجه آزادی ۵ ۱۷ ۲۹ ۴۱ شاخص تصحیح ۰/۵۶۱ ۰/۹۴۸ ۰/۹۱۱ ۰/۹۲۴



شکل ۷- نمودار باقی‌مانده‌های مدل برازش داده شده

سومین مدل SARIMA (2,1,0)*(0,1,4)12							چهارمین مدل SARIMA (0,1,2)*(0,1,4)12						
ماتریس همبستگی پارامترهای شبیه‌سازی شده							ماتریس همبستگی پارامترهای شبیه‌سازی شده						
	۱	۲	۳	۴	۵	۶		۱	۲	۳	۴	۵	۶
۲	-۰/۶۱۸						۲	۰/۰۸۲					
۳	-۰/۰۰۰	-۰/۰۳۲					۳	۰/۰۴	-۰/۳۳۲				
۴	-۰/۰۱۰	-۰/۰۲۵	-۰/۶۵۱				۴	۰/۰۲۵	۰/۲۲	۰/۵۴۳			
۵	-۰/۰۷۷	-۰/۰۴۵	-۰/۰۲۱	-۰/۵۶۴			۵	-۰/۰۱۲	-۰/۵۳	۰/۶۵	-۰/۲۶۵		
۶	-۰/۰۸۱	-۰/۰۵۱	-۰/۰۱۸	-۰/۱۰۱	-۰/۷۰۸		۶	۰/۰۴۱	۰/۲۰	-۰/۱۴	۰/۲۰۲	-۰/۵۰۶	
۷	-۰/۰۰۴	-۰/۰۰۴	-۰/۰۳۶	-۰/۰۱۲	-۰/۰۰۱۸	-۰/۰۰۸	۷	-۰/۰۲۴	-۰/۰۴	۰/۳۵	-۰/۱۰۴	۰/۰۳	۰/۰۰۹
شبیه‌سازی نهایی پارامترها							شبیه‌سازی نهایی پارامترها						
ضریب تصحیح تابع خود همبسته ضریب مجموع خطا ضریب همبستگی نوع مدل							ضریب تصحیح تابع خود همبسته ضریب مجموع خطا ضریب همبستگی نوع مدل						
AR 1	-۰/۹۴۳	۰/۰۴۳۵	-۱۴/۴۷	۰/۰۰۲			AR 1	۰/۸۹۵	۰/۰۹	۹/۸۱	۰/۰۰۱		
AR 2	-۰/۳۴۱	۰/۰۴۳۱	-۷/۶۸	۰/۰۰۱			AR 2	۰/۰۷۱	-۰/۰۱۹	۳/۹۵	-۰/۰۰۱		
SMA 12	۰/۹۸۷	۰/۰۴۰۵	۰/۴۴	۰/۰۷			SMA12	۰/۹۵۰	۰/۰۳۹۵	۱/۶۱	۰/۰۰۲		
SMA 24	۰/۰۱۶	۰/۰۴۹۵	۰/۱۵	۰/۰۷			SMA 24	۰/۰۳۹	-۰/۰۲۲۲	۰/۶۴	-۰/۵۲۶		
SMA 36	-۰/۰۳۳	۰/۰۵۶۶	-۰/۸۱	۰/۸۱۲			SMA 36	-۰/۰۱۶	۰/۰۱۴۵	-۰/۳۵	۰/۸۰۲		
SMA 48	-۰/۰۱۲	۰/۰۵۸۴	-۰/۳۶	۰/۷۹۷			SMA 48	-۰/۰۲۷	-۰/۰۱۹۴	-۰/۷۸	-۰/۴۴۲		

شکل ۸- ماتریس همبستگی مدل‌های M₃ و M₄

جدول ۲- مقادیر ضریب آکائیک دو مدل شناسایی شده

الگو	رابطه الگو	ضریب آکائیک
M ₃	(2,1,0)*(0,1,4)12	-۲۷۵/۸۴
M ₄	(0,1,2)*(0,1,4)12	-۳۲۶/۱۹

آزمون آکائیک

اگر در یک مجموعه داده تعداد چند مدل قابل قبول وجود داشت معمولاً ملاک انتخاب مدل بهتر بر مبنای آماره‌های خلاصه شده‌ای است که از باقیمانده‌های مدل‌های برازش داده شده محاسبه می‌شوند. با استفاده از معیار آکائیک که بر مبنای باقیمانده‌های مدل برازش شده می‌باشد مدل مناسب با استفاده از این آزمون به دست آمده است. جدول (۲) مقادیر آزمون آکائیک برای دو مدل نهایی را نشان می‌دهد.

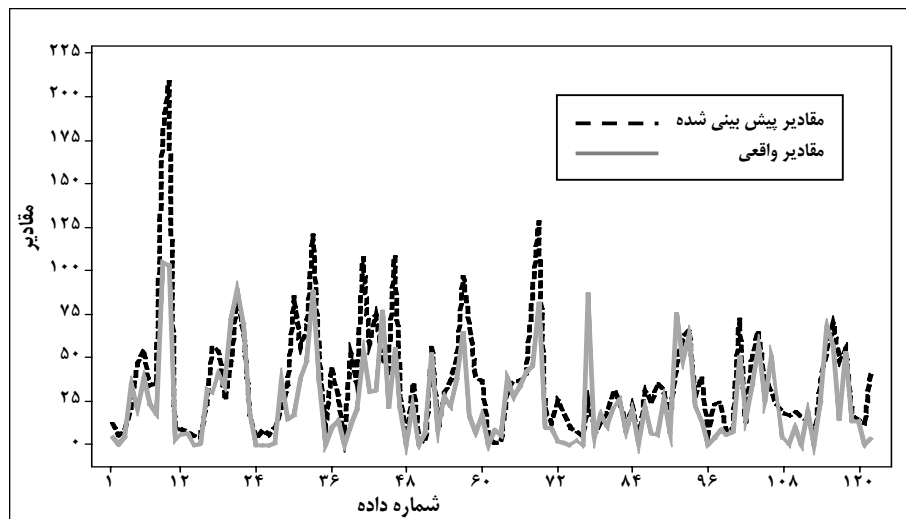
بر اساس نتایج جدول (۲)، مدل M₄ چهارمین مدل، معیار آکائیک کمتری داشته و به عنوان مدل برتر انتخاب می‌شود بهترین رابطه از بین روابط شناسایی شده رابطه‌ای است که دارای کمترین خطای استاندارد، بیشترین سطح معنی داری و بیشترین ضریب تعیین باشد. همان‌گونه که نتایج جدول (۳) نشان می‌دهد مقادیر خطای استاندارد چهارمین مدل کمتر از مقادیر مدل سوم شناسایی شده می‌باشد. از طرفی مدل فوق دارای سطح معنی‌دار بالاتری نیز نسبت به مدل سوم می‌باشد. لذا مدل SARIMA(0,1,2)*(0,1,4) به عنوان مدل نهایی جهت پیش بینی داده‌ها انتخاب شد.

مرحله آخر بعد از شناسایی بهترین الگو، عمل تولید داده‌ها انجام شده و فرآیند پیش‌بینی کامل می‌شود. در شکل (۹) نمودار سری زمانی مقادیر واقعی و پیش‌بینی شده نشان داده شده است.

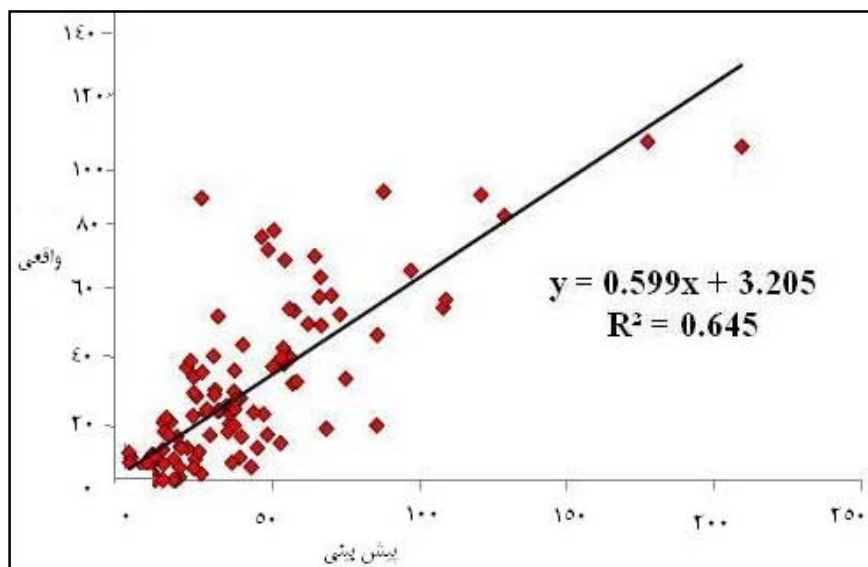
پارامترهای برآورده شده در مدل باکس و جنکینز همیشه با هم مقداری همبستگی دارند، اما همبستگی زیاد بین آنها نشان‌دهنده کیفیت پایین و عدم کفایت مدل می‌باشد. به عنوان یک قاعده عملی هرگاه قدرمطلق همبستگی بین دو پارامتر برآورد شده بزرگ‌تر از ۰/۹ باشد آن همبستگی را باید بزرگ در نظر گرفت و بایستی به مرحله شناسایی الگو برگشت و با استفاده از توابع خودهمبستگی و خودهمبستگی جزئی یک مدل دیگر را به عنوان جایگزین انتخاب کرد که همبستگی کمتری داشته باشد. در شکل (۸) ماتریس همبستگی برای مدل‌های M₃ و M₄ نشان داده شده است. در مدل سوم مقادیر شاخص تصحیح برای SMA12 به بعد بزرگ‌تر از پنج درصد بوده و همچنین آماره T نیز برای آنها کوچک‌تر از دو می‌باشد بنابراین SMA12 حضور معنی‌داری در مدل نداشته و می‌توان آن را نادیده گرفت و در همبستگی بین پارامترها هم به حساب نیاورد. بنابراین حداکثر قدرمطلق همبستگی بین پارامترهای سومین مدل ۰/۶۵۱ و چهارمین مدل ۰/۵۴۳ می‌باشد که کفایت الگوها را بهبود می‌بخشد چون هر دو کمتر از ۰/۹ می‌باشند.

جدول ۳- مقادیر خطای استاندارد دو مدل شناسایی شده

رابطه الگو	تأخیر	خطای استاندارد	شاخص تصحیح	کای اسکوتر
M ₃ (2,1,0)*(0,1,4)12	۱۲	۰/۰۴۰۵	۰/۱۸	۷/۹
	۲۴	۰/۰۴۹۵	۰/۲۷	۱۵/۷
	۳۶	۰/۰۵۶۶	۰/۳۵	۱۸/۸
	۴۸	۰/۰۵۸۴	۰/۴۴	۲۶/۸
M ₄ (0,1,2)*(0,1,4)12	۱۲	۰/۰۳۹۵	۰/۵۶	۴/۰
	۲۴	۰/۰۲۲۲	۰/۹۴	۸/۹
	۳۶	۰/۰۱۴۵	۰/۹۱	۲۲/۲
	۴۸	۰/۰۱۹۴	۰/۹۲	۲۹/۳



شکل ۹- نمودار سری زمانی مقادیر واقعی و پیش‌بینی شده



شکل ۱۰- رابطه رگرسیونی بین مقادیر واقعی و پیش‌بینی شده

احمدی (۱۳۸۴) از کار مشابهی به‌دست آورده است، تفاوت اساسی دارد. در تحقیق حاضر الگوی $SARIMA(0,1,2)*(0,1,4)12$ به عنوان مدل مناسب انتخاب شده است در حالی که در تحقیق احمدی (۱۳۸۴) الگوی $SARIMA(1,1,2)*(1,1,4)12$ به عنوان مدل مناسب انتخاب شده است. البته لازم به ذکر است که در این دو تحقیق تفاوت‌هایی در روش تحقیق مشاهده می‌شود که می‌تواند یکی از دلایل تفاوت در مدل‌های انتخابی باشد. در تحقیق انجام شده توسط احمدی (۱۳۸۴) دوره آماری برای این ایستگاه کوتاه در نظر گرفته شده بود در حالی که در تحقیق حاضر از دوره آماری ۳۰ ساله برای پیش‌بینی بهره گرفته شده است. در

بررسی میزان دقت و انطباق مدل پیش‌بینی

رابطه خط رگرسیون بین مقادیر پیش‌بینی شده و واقعی می‌تواند بیانگر دقت مدل انتخابی باشد به طوری که هر چقدر ضریب تعیین برای خط رگرسیون بیشتر باشد نشان از دقت بالای مدل انتخاب شده دارد. شکل (۱۰) رابطه رگرسیونی بین مقادیر واقعی و پیش‌بینی شده را نشان می‌دهد.

نتیجه‌گیری

نتایج به‌دست آمده از مدل‌سازی $SARIMA$ فصلی ضربی باکس و جنکینس برای ایستگاه باران‌سنجی بابامان با نتایجی که

استفاده از مدل‌های سری زمانی (حتی مدل آریم) در پیش‌بینی‌های کوتاه‌مدت قابل توصیه است. در تمامی تکرارها و تمامی مدل‌ها، نکته قابل توجه این است که عملکرد مناسب یک مدل در مرحله واسنجی، ضامن کیفیت عملکرد آن در مرحله اعتبارسنجی نمی‌باشد. لذا تکرار این پژوهش برای سایر مناطق اقلیمی کشور توصیه می‌گردد. یکی از معیارهایی که می‌تواند در تأیید روش انتخابی برای تجزیه و تحلیل مسأله مفید باشد، مقایسه روابط خط روندی که بر داده‌های ورودی و نیز خروجی‌های پیش‌بینی شده برازش داده شده است، می‌باشد که باید به لحاظ شیب خط روند با یکدیگر، انطباق تدریجی داشته باشند. بنابراین بررسی مؤلفه‌های روند و دوره‌ای در آغاز حل مسائل پیش‌بینی به کمک سری زمانی در خور توجه است.

پیشنهادها

مقادیر پیش‌بینی شده بارندگی سالیانه توسط مدل انتخابی برای ایستگاه مورد مطالعه، با استفاده از رابطه رگرسیونی به‌دست‌آمده، تصحیح گردد. جهت دستیابی به نتایج مطلوب، بهتر است هر ساله آخرین آمار واقعی بارندگی هر ایستگاه به داده‌های اصلی اضافه گردیده و الگوی تعیین‌شده، مجدداً برازش داده شود و در صورت لزوم تغییرات لازم در آن اعمال گردد و سپس برای الگوی جدید اقدام به پیش‌بینی گردد.

این تحقیق از داده‌های اصلی برای مدل‌سازی استفاده شده و تبدیل پایداری واریانس روی داده‌ها انجام نشده است.

به طور کلی می‌توان گفت الگوهایی که دارای پارامتر بیشتری هستند و مرتبه‌های اتورگرسیو یا میانگین متحرک آنها بیشتر از یک می‌باشد، برای سال‌های بعد مقادیر متفاوتی را نشان می‌دهد، اما این تفاوت‌ها نیز محدود به چند سال است که حداکثر از تعداد پارامترهای الگو تجاوز می‌نماید.

میانگین درازمدت بارندگی سالیانه پیش‌بینی شده با الگوی تعیین شده برای ایستگاه مورد مطالعه، بیشتر از میانگین درازمدت بارندگی سالیانه واقعی می‌باشد، به طوری که در حدود ۸۴ درصد از داده‌های ایستگاه مقادیر متوسط بارندگی پیش‌بینی شده بیشتر از مقدار واقعی آنها است.

نظر به تفاوت قابل ملاحظه بین میانگین درازمدت بارندگی سالیانه واقعی و پیش‌بینی شده و با توجه به رابطه خط رگرسیون بین میانگین واقعی بارندگی سالیانه و میانگین پیش‌بینی شده، این خط دارای ضریب همبستگی برابر $0.645+$ می‌باشد و از همبستگی نسبتاً خوبی برخوردار می‌باشد. بنابراین می‌توان از این رابطه برای تصحیح مقادیر متوسط بارندگی پیش‌بینی شده در ایستگاه مورد مطالعه استفاده نمود.

$$y = 0.599x + 3.205 \quad (6)$$

در این رابطه y : مقدار واقعی متوسط بارندگی سالانه و x : مقدار پیش‌بینی شده متوسط بارندگی سالانه در ایستگاه مورد نظر می‌باشد.

منابع

- ۱- احمدی، ف. ۱۳۸۴. بررسی و پیش‌بینی بارندگی سالانه استان خراسان براساس سری‌های زمانی. پایان‌نامه کارشناسی ارشد، گرایش مهندسی منابع آب، دانشکده کشاورزی، دانشگاه فردوسی مشهد، ۲۸۴ صفحه
- ۲- نیرومند، ح. ع. ۱۳۷۸. تجزیه و تحلیل سری‌های زمانی. انتشارات دانشگاه مشهد. چاپ دوم. ۳۵۵ صفحه
- ۳- شریفان، ح و ب. قهرمان. ۱۳۸۶. ارزیابی پیش‌بینی باران با بکارگیری تکنیک SARIMA در استان گلستان، مجله علوم کشاورزی و منابع طبیعی، جلد ۱۴ (۳): ۱۷۳-۱۵۸.
- ۴- نیرومند، ح و ا. بزرگ‌نیا. ۱۳۷۲. مقدمه‌ای بر تحلیل سری‌های زمانی. انتشارات دانشگاه فردوسی مشهد. ۱۸۵ صفحه
- ۵- آشگر طوسی، ا. ۱۳۸۲. پیش‌بینی خشکسالی در استان خراسان و بهینه‌سازی الگوی کشت مناسب. پایان‌نامه کارشناسی ارشد، گرایش زهکشی، دانشکده کشاورزی، دانشگاه فردوسی مشهد، ۱۲۰ صفحه
- ۶- جمشیدی، و. ۱۳۶۸. بررسی و ارزیابی درجه حرارت و بارندگی در شهر تهران با استفاده از سری زمانی. پایان‌نامه کارشناسی ارشد، گرایش آمار، دانشکده ریاضیات، دانشگاه تربیت مدرس تهران، ۲۶۳ صفحه
- ۷- خردمندنیا، م و ح. عساکره. ۱۳۸۰. تطبیق مدل آریم برای داده‌های متوسط درجه حرارت سالیانه در جاسک. سومین کنفرانس فرآیندهای تصادفی، دانشگاه اصفهان.

۸- مالکی، م. ۱۳۶۸. بررسی و مدلسازی دما و بارش در غرب کشور. پایان‌نامه کارشناسی ارشد، گرایش آمار، دانشکده ریاضیات، دانشگاه شهید بهشتی، ۱۸۴ صفحه

۹- رسولی، ا. ۱۳۸۱. پیش‌بینی درجه حرارت ماهانه شهر تبریز با استفاده از مدل آریم. مجله علوم جامعه‌شناسی، دانشگاه تبریز، جلد ۵(۸): ۶۸-۸۵

- 10- Bloomfield, P and D. Nychka. 1992. Climate spectra and detecting climate change. *Climate Change*, 21: 275-287.
- 11- Box, G. E. P and D. R. Cox. 1964. An analysis of transformations, *Journal of the Royal Statistical Society, Series B (Methodological)*, 26(2): 211-252.
- 12- Box, G. E. P and G. M. Jenkins., and G. C. Reinsel. 1994. *Time series analysis: Forecasting and control*. Third Edition, Holden-Day.
- 13- Burlando, P., Montana, A. and R. Raze. 1996. Forecasting of storm rainfall by combined use of radar, rain gages and linear models. *Atmospheric Research*, 42: 199-216.
- 14- Folland, C. K. 1990. Observed climatic variation and change, climate change. *The IPCC Scientific Assessment*, Cambridge University Press, 25:195-238.
- 15- Haltiner, J. P. and J. D.Salas. 1988. Development and testing of a multivariate, seasonal ARIMA(1,1) model, *Journal of Hydrology*, 104: 247-272.
- 16- Hansen, J. and S. Lebedeff. 1988. Global surface air temperature. *Geology Letter*, 15: 323-326.
- 17- Karamouz, M., Fallahi, M., Nazif, S., and M. Rahimi Farahani. 2010. Long lead rainfall prediction using statistical downscaling and artificial neural network modeling. *Civil Engineering*, 16(2): 165-172.
- 18- Ki kim, M., Sik kang, I., Kyupark, C. and K. Myong kim. 2004. Superensemble prediction of regional precipitation over Korea. *International Journal of Climatology*, 28: 158-169
- 19- Saplioglu, K., Cimen, M. and B. Akman. 2010. Daily precipitation prediction in Isparta station by artificial neural network: Ohrid, Republic of Macedonia.