

## کاربرد مدل سری های زمانی در پیش بینی بارندگی ماهیانه و سالانه مشهد

فیروز احمدی - بیژن قهرمان - کامران داوری - حسین علی نیرومند<sup>۱</sup>

تاریخ دریافت: ۸۴/۷/۲۶

### چکیده

پیش بینی بارندگی و برآورد میزان نزولات جوی برای هر منطقه و حوضه آبریز یکی از مهمترین پارامترهای اقلیمی می باشد. لذا پیش بینی آن در استفاده بهینه از منابع آب و نقش آن در برنامه ریزی های خرد و کلان از اهمیت ویژه ای برخوردار است. الگوسازی و شبیه سازی رفتار بارندگی، به خصوص در سال های اخیر، مورد توجه محافل علمی بوده است. یکی از روش های دست یابی به این مقصود، مدل سازی بارندگی ماهیانه بر اساس الگوهای فصلی- ضربی باکس و جنکینس است. در این گونه مدل ها، میزان بارندگی در هر ماه بر اساس میزان بارندگی در همان ماه و ماه های قبل و بعد از آن در سال های پیشین و همچنین مؤلفه های تصادفی همان ماه و ماه های قبل و بعد از آن در سال های قبل از آن بیان می شود. در این تحقیق آمار درازمدت بارندگی ماهیانه در ایستگاه سینوپتیک مشهد به عنوان یک سری زمانی در نظر گرفته شد. الگوی مناسب برای این آمار بر اساس توابع خودهمبستگی (SAC) و خودهمبستگی جزئی (SPAC) و بررسی تمام الگوهای احتمالی به لحاظ نرمال بودن و تصادفی بودن باقی مانده ها، همبستگی پارامترهای مدل و نتایج آزمون پورت- مانتو، به صورت  $_{12}^{0.01}$  الى  $_{2003}^{2001}$  با دقت قابل قبولی پیش بینی گردید. براساس الگوی انتخاب شده، میزان بارندگی ماهیانه برای سال های  $_{2001}^{2003}$  با دقت نتایج نشان داد که بین سه دوره زمانی متفاوت  $_{40}^{50}$  و  $_{30}^{40}$  ساله، مدل تغییر نمی کند و برآوردها تقریباً یکسان است. با این وجود طول دوره های آماری کوتاه مدت تر تفاوت هایی با آمار واقعی نشان دادند.

واژه های کلیدی: سری های زمانی، خودهمبستگی، خودهمبستگی جزئی، مدل، ایستایی

### مقدمه

مشاهدات آن است و بنابراین، ترتیب مشاهدات دارای اهمیت است. وقتی مشاهدات متوالی نامستقل اند، مقادیر آینده را می توان از مشاهدات گذشته، پیش بینی کرد. اگر یک سری زمانی را بتوان کاملاً پیش بینی کرد، آن را غیرتصادفی گویند. لیکن بیشتر سری های زمانی تصادفی هستند که در آن ها آینده فقط تا حدی به وسیله ای مقادیر گذشته تعیین می شود. برای سری های تصادفی، پیش بینی های کامل غیرممکن بوده و باستی این ایده به جای آن قرار گیرد که برای یک توزیع احتمالی، پیش بینی مقادیر آینده مقید به آگاهی از مقادیر گذشته ای آن است (۹).

باکس و جنکینس<sup>۲</sup> در فراهم آوردن یک روش کلی برای پیش بینی سری های زمانی سهم بسزایی داشته اند. آن ها با تأکید بر تفاضلی کردن، توانستند الگوهایی را برای سری های نایستا

پیش بینی بارندگی و برآورد میزان آن در هر ماه (و یا سال) و برای هر منطقه و حوضه آبریز، به عنوان یکی از مهم ترین پارامترهای جوی و منبع اصلی تامین کننده ای آب های سطحی و زیرزمینی، برای برنامه ریزی در امورات مختلف زندگی اجتماعی انسان، از اهمیت ویژه ای در استفاده ای بهینه از منابع آب برخوردار است.

سری زمانی<sup>۱</sup>، مجموعه ای مشاهداتی است که بر حسب زمان و در فواصل مساوی از آن مرتب شده باشند، هر چند مرتب شدن داده ها ممکن است با توجه به ابعاد دیگری چون فاصله نیز باشد (۱۰ و ۱۱).

طبیعت ذاتی یک سری زمانی، وابسته یا هم بسته بودن

۱- به ترتیب دانشجوی سابق کارشناسی ارشد آبیاری، دانشیار، استادیار گروه آبیاری و زهکشی دانشگاه فردوسی مشهد و استاد ریاضی دانشگاه فردوسی مشهد.

1- Time Series

2- Box and Jenkins

شده است.

### مواد و روش ها

موقعیت منطقه و آمار. شهرستان مشهد (عرض جغرافیایی ۳۶ درجه و ۱۷ دقیقه شمالی، طول جغرافیایی ۵۹ درجه و ۳۸ دقیقه شرقی، ارتفاع ۹۴۶ متر از سطح دریا)، مرکز استان خراسان رضوی و در شمال شرقی کشور ایران قرار گرفته است. آمار بارندگی ماهیانه ایستگاه سینوپتیک مشهد طی دوره آماری ۱۹۵۱-۲۰۰۰ از سایت سازمان هواشناسی کشور (<http://www.weather.ir/farsi/amar/default.asp>) و آمار (پیش‌بینی باکس و جنکینس. مراحل اصلی درساختن الگوی پیش‌بینی باکس و جنکینس در شکل ۲ خلاصه شده است.

مدل‌های پیش‌بینی باکس و جنکینس با سری‌های زمانی ایستادروکار دارند. نایستایی<sup>۱</sup> در واریانس را با استفاده از تبدیل توانی باکس و کاکس (رابطه ۱) و نایستایی در میانگین را با استفاده از تفضیلی کردن می‌توان برطرف کرد.

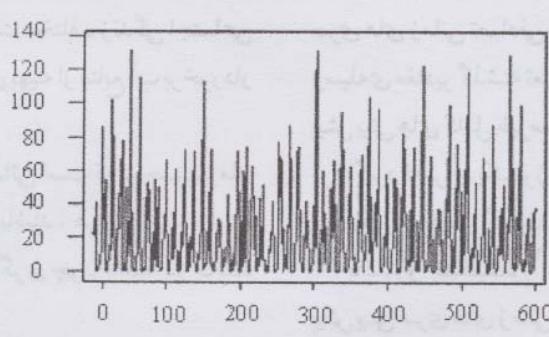
$$T(Z_i) = Z_i^{(\lambda)} - \frac{Z_i^\lambda - 1}{\lambda} \quad (1)$$

که در آن  $Z_i$  مقدار اولیه واقعه در زمان  $i$ ،  $\lambda$  پارامتر تبدیل

بسازند و رده کلی الگوهای<sup>۱</sup> ARIMA به آن‌ها نسبت داده می‌شود<sup>(۹)</sup>. اخیراً پژوهشگران به منظور تحلیل تغییرات پارامترهای اقلیمی تلاش نموده اند تا این پارامترها را الگوسازی و سپس شبیه‌سازی نمایند. الگوسازی در خانواده‌های آرما (ARMA)، آریما (ARIMA) و آریمای فصلی ضربی (SARIMA<sup>۳</sup>)، یکی از شیوه‌های مهم و معتر در شبیه‌سازی فراسنج‌های اقلیمی است<sup>(۱۱)</sup>.

کاربرد سری‌های زمانی از دیرباز در منابع علمی جهان در حال انجام بوده است<sup>(۱۲-۱۵)</sup>. با این وجود کشور ما به لحاظ سابقه کوتاه در بهره‌گیری از روش‌های اندازه‌گیری مستقیم داده‌های جوی از این نظر کمتر مورد مطالعه قرار گرفته است<sup>(۴)</sup>. در این میان می‌توان به تحقیقات جمشیدی<sup>(۳)</sup> در مدل بندی دما-بارش تهران، مالکی<sup>(۷)</sup> در الگوسازی دما و بارش غرب کشور، مشکاتی<sup>(۸)</sup> در تحلیل استوکاستیکی عوامل اقلیمی شهر تهران، رحیم‌زاده<sup>(۵)</sup> در مدل سازی فشار جو در ایستگاه تهران-مهرآباد، رسولی<sup>(۶)</sup> در تحلیل مقدماتی سری‌های زمانی دمای هوای تبریز، خردمندی و عساکره<sup>(۴)</sup> در مدل SARIMA برای متوسط درجه حرارت ماهانه جاسک و آشگر طوسی<sup>(۲)</sup> در پیش‌بینی وقوع خشک‌سالی در استان خراسان و بهینه‌سازی الگوی کشت جهت سازگاری با آن اشاره نمود.

در این مقاله با استفاده از روش باکس و جنکینس و استفاده از الگوی آریمای فصلی ضربی (SARIMA)، رفتار بارندگی ماهیانه ایستگاه مشهد طی دوره آماری ۱۹۵۱-۲۰۰۳<sup>۲</sup> الگوسازی

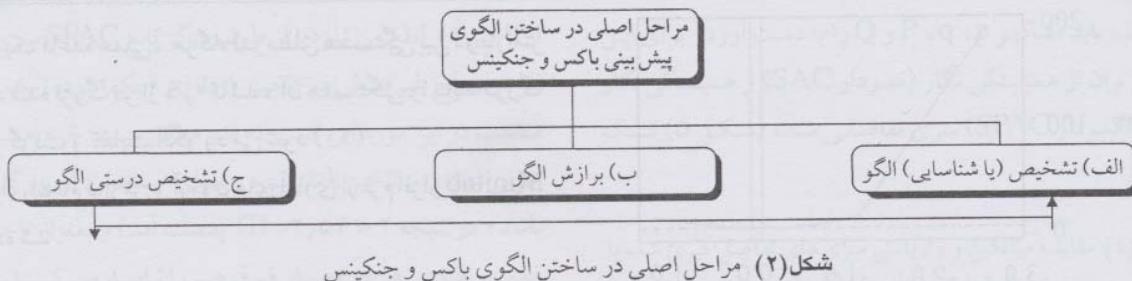


شکل (۱) نمودار سری زمانی بارندگی ماهیانه ایستگاه مشهد- ۱۹۵۱- ۲۰۰۰- ۱۹۵۱- ۲۰۰۳

1) Autoregressive Integrated Moving Average

3) Seasonal Autoregressive Integrated Moving Average

2) Autoregressive Moving Average



شکل (۲) مراحل اصلی در ساختن الگوی باکس و جنکینس

الگوی عمومی باکس و جنکینس از مرتبه  $p,q,d,D,P,Q$  به صورت زیر نوشته می شود:

$$\phi_p(B)\phi_p(B^L)\nabla_L^D\nabla^d y_t^* = \delta + \theta_q(B)\theta_Q(B^L)a_t \quad (2)$$

که به مدل «آریمای فصلی-ضریبی باکس و جنکینس» مشهور است و در آن:

$$\phi_p(B) = (1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p) \text{ عملگر اتورگرسیو غیر فصلی از مرتبه } p,$$

$$\theta_q(B) = (1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q) \text{ عملگر میانگین اتورگرسیو فصلی از مرتبه } q,$$

$$\theta_Q(B^L) = (1 - \theta_{1,L} B^L - \theta_{2,L} B^{2L} - \dots - \theta_{Q,L} B^{QL}) \text{ عملگر میانگین متوجه غیر فصلی از مرتبه } Q,$$

$$\mu\phi_p(B)\phi_p(B^L) = \delta \text{ و مقدار ثابت مدل (}\mu\text{ میانگین واقعی سری زمانی است)،}$$

$$\phi_1, \phi_2, \dots, \phi_p, \phi_{1,L}, \phi_{2,L}, \dots, \phi_{Q,L}, \theta_1, \theta_2, \dots, \theta_q, \theta_{1,L}, \theta_{2,L}, \dots, \theta_{Q,L} \text{ پارامترهای نامعلوم و مجھول هستند که بایستی از داده های نمونه برآورد گردند، } a_1, a_{t-1}, \dots, a_t \text{ جملات تصادفی اغتشاش تابع توزیع نرمال، } B \text{ عملگر پس رو}$$

$$\nabla_L^D y_t^* = (\nabla - B^L)^k y_t \text{ عملگر غیر فصلی (}\nabla - B^L)^k \text{ می باشد.}$$

پارامترهای برآورده شده در مدل باکس و جنکینس همیشه با هم مقداری همبستگی دارند، اما همبستگی زیاد بین آنها نشان دهنده کیفیت پایین و عدم کفايت مدل می باشد (۱۰). به

$Z^{(A)}$  مقدار تبدیل یافته نظری است. در تبدیل توانی باکس و کاکس مقدار  $\lambda = 0$  متناظر با تبدیل لگاریتمی است. آزمون بارتلت و لون یک روش متداول برای تشخیص نایستایی در واریانس به شمار می آید.

تبدیل باکس و کاکس برای رفع نایستایی در واریانس نیاز به وجود اطلاعات غیر صفر دارد. بارندگی ماهیانه مشهد در بسیاری از مواقع صفر است. برای رفع مشکل، عدد ثابتی مثل  $c$  را می توان به تمامی داده ها اضافه کرد. این عمل تغییری در واریانس ایجاد نخواهد کرد.

یک روش مناسب برای تعیین مقادیر  $D$  و  $d$  (به ترتیب مرتبه تفاضل گیری فصلی و غیر فصلی)، کمینه کردن واریانس سری تفاضلی حاصل ( $\sigma^2$ ) است. بنابراین با اختیار نمودن مقادیر مختلف برای  $D$  و  $d$  و مقایسه واریانس سری تفاضلی شده به ازای مقادیر مذبور، می توان درجه ای تفاضلی را برای مرتب فصلی و غیر فصلی تعیین نمود. به عبارت دیگر مقادیر  $D$  و  $d$  طوری انتخاب می شوند که واریانس سری تفاضلی مقدار خود باشد (۴).

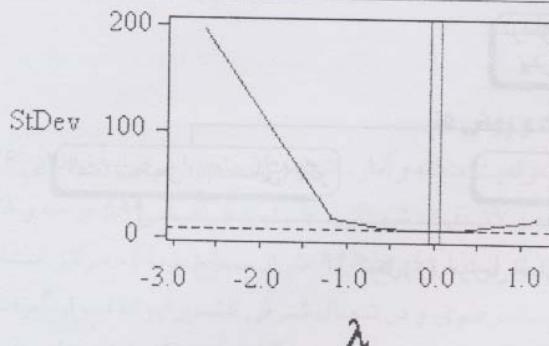
نایستایی در میانگین را به روش هایی چون تنزل آرام یک موج سینوسی در همبستگی نگار سری زمانی (SAC<sup>1</sup>) می توان تشخیص داد.

تعیین مرتبه های مدل  $(p, q, P, Q)$  از روی همبستگی نگار (SAC) و همبستگی نگار نمونه ای (PSAC<sup>2</sup>) قابل تشخیص است. روش های کلاسیک در کتب سری های زمانی (نظیر ۹ و ۱۰) آورده شده است.

1) Non- Stationarity

2) Sample Autocorrelation

3) Partial Sample Autocorrelation



شکل(۳) نتایج تبدیل باکس و کاکس بر روی سری زمانی داده های بارندگی مشهد بعد از اضافه کردن مقدار ثابت  $c = 1$

به ازای مقادیر مختلف  $D$  و  $d$  سری های تفاضلی مربوطه پارامترهای آنها محاسبه گردید. در جدول ۱ میانگین ( $\bar{w}_t$ ) و واریانس ( $s^2$ ) سری تفاضلی شده ارائه شده است.  $n$  تعداد داده هایی است که در محاسبه پارامترهای فوق به کار رفته است. همان طور که در جدول ۱ ملاحظه می شود، مقدار واریانس سری تفاضلی شده  $y_t^d = (1 - B^{12})^D (1 - B)^d y_t^*$  به ازای  $D=1$  و  $d=0$  برابر با  $1/1933$  و به عنوان مرتبه های تفاضل گیری در نظر گرفته می شود. سری تفاضلی شده حاصل از  $D=1$  و  $d=0$  به شرح زیر با  $\{Z_t\}$  نمایش داده می شود:

$$z_t = \nabla_{12}^1 \nabla^0 y_t^* = (1 - B^{12})^1 (1 - B)^0 y_t^* = y_t^* - B^{12} y_{t-12}^* \quad (3)$$

#### انتخاب ساختار مدل

جهت شناسایی الگوی مناسب فصلی ضربی باکس جنکینس که به صورت  $SARIMA(p, d, q)(P, D, Q)_{12}$  نمایش داده

عنوان یک قاعده عملی، هرگاه قدر مطلق همبستگی بین دو پارامتر برآورد شده بزرگ تر از  $0.9$  باشد، آن همبستگی را باید بزرگ دانظر گرفت و کنایت الگوریتمی شود (۱۰).

برای انجام برآذش و آزمون های آماری از نرم افزار Minitab استفاده شد.

#### نتایج و بحث

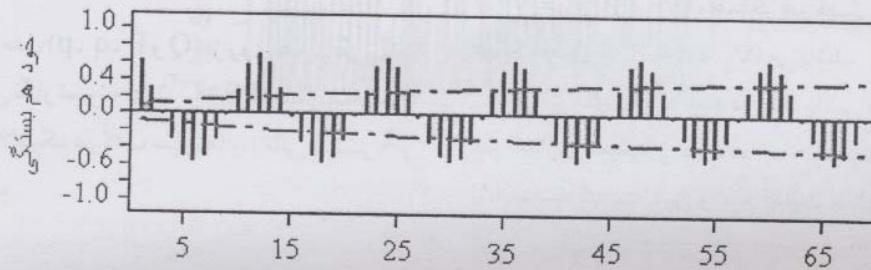
الف) بررسی پایداری واریانس داده های اصلی بارندگی ایستگاه مشهد و انجام تبدیل باکس و کاکس

نتایج آزمون بارتلت و لون روی داده های اصلی بارندگی ایستگاه مشهد، نشان داد که داده های اصلی نایستا در واریانس هستند. در اینجا عدد ثابت یک به کلیه اطلاعات اضافه شد تا از مشکل صفر بودن پرهیز شود. بر اساس شکل ۳، مقدار  $\lambda = 0$  مناسب ترین پارامتر تبدیل برای تبدیل باکس و کاکس به نظر می رسد. از این پس سری تبدیل یافته  $y_t^* = \ln(y_t + 1)$  دارای پایداری واریانس می باشد.

ب) بررسی ایستایی سری زمانی  $\{y_t^*\}$  و تعیین مرتبه های تفاضلی فصلی و غیرفصلی

در شکل (۴) همبستگی نگار سری زمانی  $\{y_t^*\}$  که میان یک سری نایستا است ارائه شده است. همان طور که مشاهده می گردد، نمودار SAC سری  $\{y_t^*\}$  به شکل یک موج سینوسی و به آرامی تنزل می یابد و با توجه به آهنگ کند کاهش خود همبستگی ها، سری مورد مطالعه را باید نایستا در نظر گرفت.

به منظور به دست آوردن سری تفاضلی مناسب، برای تبدیل نمودن این سری نایستا به سری ایستا، باید مقادیر مناسبی را برای  $D$  (مرتبه تفاضل گیری فصلی) و  $d$  (مرتبه تفاضل گیری غیر فصلی) تعیین نمود.



شکل(۴) همبستگی نگار سری  $\{y_t^*\}$

از تاخیر ۱، قطع می شود. از طرف دیگر تابع  $Z_t$  سری SPAC نیز در سطح غیرفصلی در تاخیر ۱ دارای «پیک» بوده (یک معیار مناسب در این مورد این است که قدر مطلق نسبت خود همبستگی جزئی نمونه در تاخیر  $k$  تقسیم بر خطای استاندارد آن بزرگتر از ۲ باشد؛ در نتیجه  $2 > 2,54 > |T|$  بحسب آمد) و بعد از تاخیر ۱، قطع می شود. هر دو معیار فوق هم را تایید می کنند. پس می بایست یکی از عملگرهای میانگین متوجه غیرفصلی از مرتبه  $q = 1$  (یعنی  $(1 - \theta_1 B)^{-1} = (B - \theta_1)$ ) و یا اتورگرسیو از مرتبه  $q = 1$  (یعنی  $(1 - \phi_1 B)^{-1} = (B - \phi_1)$ ) در مدل منظور شود. با توجه به این که هر دو تابع  $SAC$  و  $SPAC$  در سطح غیرفصلی تقریباً با یک شدت قطع می شوند، بنابراین یک بار عملگر میانگین متوجه غیرفصلی از مرتبه ۱ و یک بار عملگر اتورگرسیو از مرتبه ۱ در مدل منظور می شود و هر کدام که نتیجه بهتری را نشان داد به عنوان گزینه نهایی انتخاب می شود. بدین ترتیب در مدل پیشنهادی، یکی از گزینه های  $(1 - \theta_1 B)^{-1}$  و  $(1 - \phi_1 B)^{-1}$  (یا  $p = 0$  و  $q = 0$ ) به عنوان انتخاب برتر در نظر گرفته می شود.

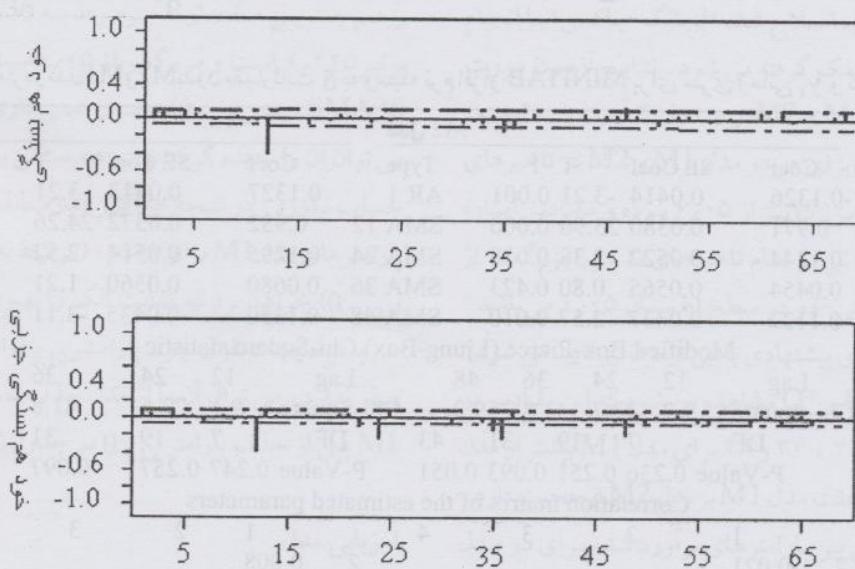
تابع  $Z_t$  سری  $SAC$  در سطح فصلی در تاخیرهای ۱۲، ۳۵، ۳۶ و ۴۷ دارای «پیک» بوده (قدرمطلق  $T$  به ترتیب  $14, 10, 16, 3, 16, 1, 88, 2, 72$  و  $1, 6$  بوده که متناظراً از اعداد استاندارد  $2,54 > |T| > 1, 6$  بزرگتر از ۱ باشد؛ در نتیجه  $2,54 > |T| > 1, 6$  بحسب آمد) (اطلاعات بیشتر در احمدی، ۱۳۸۴ آورده شده است) و بعد تابع  $Z_t$  سری  $SPAC$  نیز در سطح فصلی، نزولی است (شکل ۵).

می شود، باید مقادیر  $p, q, P$  و  $Q$  را به دست آورد. برای این کار می توان از همبستگی نگار ( $SAC$ ) و همبستگی نگار جزئی ( $SPAC$ ) سری تفاضلی شده (شکل ۵) کمک گرفت.

جدول (۱) مقایسه میانگین و واریانس سری های تفاضل گیری شده با ترکیبات مختلفی از  $D$  و  $d$

$\sigma^2$	$w_t$	n	D	d
۱,۳۷۲۷	۰,۰۰۰۸	۵۹۹	۰	۱
۱,۱۹۳۳	۰,۰۰۱۶	۵۸۸	۱	۰
۱,۵۹۴۸	۰,۰۰۱۴	۵۸۷	۱	۱
۱,۸۶۶۷	۰,۰۰۱۵	۵۹۸	۰	۲
۲,۰۲۸۱	-۰,۰۱۰۳	۵۷۶	۲	۰
۱,۶۴۳۲	۰,۰۰۰۲۷	۵۸۶	۱	۲
۲,۷۳۹	۰,۰۱۱	۵۷۵	۲	۱

تابع  $Z_t$  سری  $SAC$  در سطح غیرفصلی در تاخیر ۱ دارای «پیک» بوده (یک معیار مناسب در این مورد این است که قدر مطلق نسبت خود همبستگی نمونه در تاخیر  $k$  تقسیم بر خطای استاندارد آن بزرگتر از ۱ باشد؛ در نتیجه  $2,54 > |T| > 1, 6$  بحسب آمد) (اطلاعات بیشتر در احمدی، ۱۳۸۴ آورده شده است) و بعد



شکل (۵) همبستگی نگار و همبستگی نگار جزئی سری زمانی  $Z_t$

در حالت کلی، این مدل‌ها با حداقل پارامترهای ممکن (منجمله مقدار ثابت $\delta$ ) نوشته شده‌اند. با این وجود بسته به تعداد پارامترهای معنی‌دار آن‌ها، می‌توانند به مدل‌هایی با پارامترهای معنی‌دار کمتری تبدیل شوند.

#### برآورد الگو (برآورد پارامترها)

نتایج نشان داد (احمدی، ۱۳۸۴) که مقدار آماره $t$  (مقدار $T$ ) برای دو مدل M1 و M2، به ترتیب برابر ۱/۸۲ و ۱/۶۴ بوده و هردو از عدد ۲ کوچک‌تر هستند. در نتیجه می‌توان نتیجه گرفت که مقدار ثابت $\delta$  در هر دو مدل تفاوت معنی‌داری با صفر نداشته و می‌توان آن را از مدل حذف نمود. حال با توجه به حذف مقدار ثابت $\delta$  از مدل‌های M1 و M2، پارامترهای مدل مجدداً برآورده شد (جدول ۲).

#### تشخیص درستی الگو

همبستگی نگار (SAC) باقی مانده‌ها، همبستگی نگار جزئی (SPAC) باقی مانده‌ها، بافت نگار باقی مانده‌ها به همراه منحنی نرمال و نمودار باقی مانده‌ها در طول زمان، برای الگوهای M1 و M2 بررسی گردید (این نمودارها برای مدل M1 در شکل ۶ نشان داده شده است؛ نتایج برای مدل M2 در احمدی، ۱۳۸۴ یافت می‌شود). نمودار باقی مانده‌ها در طول زمان برای هر دو مدل، در اطراف سطح افقی صفر تقریباً پراکنده‌گی مستطیلی بدون روندی

بنابراین باید عملگر میانگین متعدد فصلی از مرتبه ۴ (معادله ۴) در مدل منظور شود. بدین ترتیب در مدل پیشنهادی،  $Q=4$  و  $p=0$ .

$$\theta_4(B^{12}) = 1 - \theta_{1,12}B^{12} - \theta_{2,12}B^{24} - \theta_{3,12}B^{36} - \theta_{4,12}B^{48} \quad (4)$$

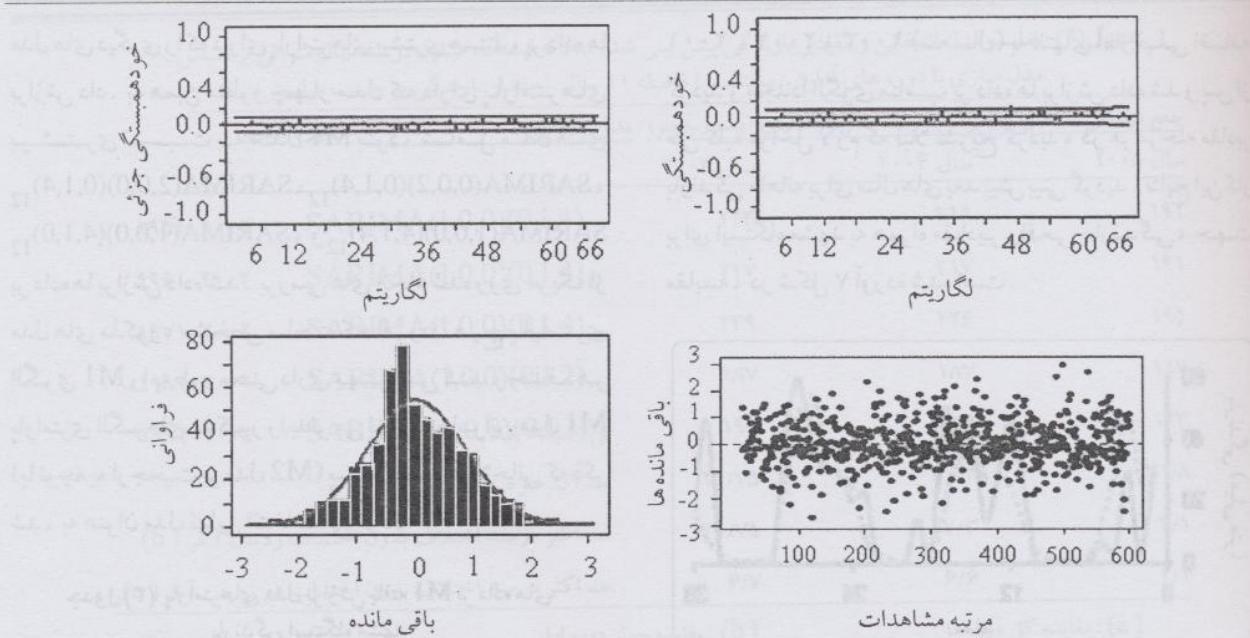
از تلفیق عملگرهای انتخابی فوق با مدل عمومی باکس و جنکینس، در این مرحله دو مدل پیشنهادی به دست می‌آید. مدل با انتخاب گزینه ۱،  $p=0$ ،  $q=0$ ،  $Q=0$  و شکل M1 SARIMA(1,0,0)(0,1,4)<sub>12</sub> و مدل M2 با انتخاب گزینه ۲،  $p=0$ ،  $q=1$ ،  $Q=0$  و شکل M2 SARIMA(0,0,1)(0,1,4)<sub>12</sub> نمایش داده می‌شود. این مدل‌ها به ترتیب از تلفیق عملگر  $(1 - \phi_1 B)$  برای مدل M1 و از تلفیق عملگر  $(1 - \theta_1 B)$  برای مدل M2 با عملگر ارائه شده در معادله ۴ به دست می‌آیند و به شکل زیر نوشته می‌شوند:

$$(1 - \phi_1 B)z_t = \delta + (1 - \theta_{1,12}B^{12} - \theta_{2,12}B^{24} - \theta_{3,12}B^{36} - \theta_{4,12}B^{48})a_t \quad M1 \quad (5)$$

$$z_t = \delta + (1 - \theta_1 B)(1 - \theta_{1,12}B^{12} - \theta_{2,12}B^{24} - \theta_{3,12}B^{36} - \theta_{4,12}B^{48})a_t \quad M2 \quad (6)$$

جدول (۲) برآورد پارامترهای M1 و M2 بدون مقدار ثابت $\delta$  به وسیله نرم افزار MINITAB برای سری زمانی $y^*$  و نتایج آزمون پورت مانتو

M2 مدل				M1 مدل					
Type	Coef	SE Coef	T	P	Type	Coef	SE Coef	T	P
MA 1	-0.1326	0.0414	-3.21	0.001	AR 1	0.1327	0.0413	3.21	0.001
SMA 12	0.971	0.0380	23.90	0.000	SMA 12	0.932	0.0372	24.26	0.000
SMA 24	0.1244	0.0522	2.38	0.018	SMA 24	0.1295	0.0514	2.52	0.012
SMA 36	0.0454	0.0565	0.80	0.423	SMA 36	0.0680	0.0560	1.21	0.225
SMA 48	-0.1122	0.0437	-2.57	0.010	SMA 48	-0.1350	0.0435	-3.11	0.002
Modified Box-Pierce (Ljung-Box) Chi-Square statistic									
Lag	12	24	36	48	Lag	12	24	36	48
Chi-Square	9.2	22.7	41.8	59.2	Chi-Square	9.1	22.6	41.6	59.1
DF	7	19	31	43	DF	7	19	31	43
P-Value	0.236	0.251	0.093	0.051	P-Value	0.247	0.257	0.097	0.052
Correlation matrix of the estimated parameters									
1	2	3	4		1	2	3	4	
2	-0.021				2	0.008			
3	-0.003	-0.591			3	-0.005	-0.577		



شکل (۶) همبستگی نگار (بالا راست)، همبستگی نگار جزئی (بالا چپ)، نمودار باقی مانده ها (پایین راست) و توزیع نرمال باقی مانده ها (پایین چپ) برای مدل M1

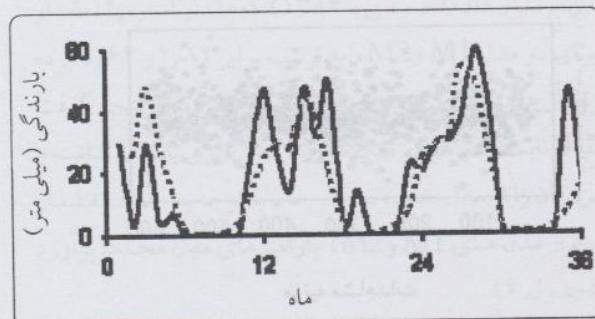
مشاهده می شود که حد اکثر قدر مطلق همبستگی بین پارامتر های مدل M1 برابر  $0,638$  (مربوط به همبستگی بین  $\theta_{3,12}$  و  $\theta_{4,12}$ ) می باشد و کوچک تر از  $0,9$  است و بنابراین با توجه به همبستگی بین پارامتر ها نمی توان کفایت مدل M1 را در نمود. همچنین مشاهده می شود که مقدار قدر مطلق T مربوط به SMA36 (یعنی مقدار  $|t|$ ) در مدل M1 برابر  $1,21$  است و کمتر از  $2$  می باشد و هم چنین مقدار P مربوط به SMA36 در مدل M1 برابر  $0,225$  است و بزرگ تر از  $0,05$  است. بنابراین پارامتر SMA36 (یعنی  $\theta_{3,12}$ ) حضور معنی داری در مدل نداشته و می توان آن را نادیده گرفت و در همبستگی بین پارامتر ها هم به حساب نیارود. در نتیجه حد اکثر قدر مطلق همبستگی بین پارامتر های مدل M1 برابر  $0,577$  (مربوط به همبستگی بین  $\theta_{1,12}$  و  $\theta_{2,12}$ ) خواهد بود که کفایت الگو را بهبود می بخشد.

با توضیحات فوق چون پارامتر  $\theta_{3,12}$  اختلاف معنی داری با صفر ندارد، می توان آن را از مدل حذف نمود. پارامتر های مدل M1 پس از حذف پارامتر  $\theta_{3,12}$  در جدول ۳ ارائه شده است.

جهت اطمینان از صحت انتخاب مدل نهایی، می توان ارزیابی مدل

رانشان می دهنده و هم چنین بافت نگار باقی مانده ها نیز نرمال بودن آن ها را رد نمی کند. بررسی رفتار توابع RSAC و RSPAC برای هر دو مدل M1 و M2 نشان می دهد که توابع مذکور رفتاری شبیه متغیر های تصادفی داشته و هیچ طرحی را تشکیل نمی دهند. بنابراین تا این مرحله هر دو مدل پیشنهادی M1 و M2 مناسب به نظر می رسدند. جهت بررسی بیشتر و انتخاب مدل برتر از میان دو مدل پیشنهادی، از آماره اصلاح شده «لجانگ-باکس» با انجام آزمون «پورت مانتو» کمک گرفته می شود. نتایج آزمون پورت مانتو برای مدل های M1 و M2 در جدول ۲ آورده شده است. مقادیر P در این جدول برای هر دو مدل M1 و M2 در تاخیر های ۱۲، ۲۴، ۳۶ و ۴۸ بزرگ تر از  $0,05$  بوده و مقادیر آماره لجانگ-باکس ( $Q^*$ ) نیز در تمام تاخیر های مذکور بزرگ تر از مقادیر بحرانی توزیع کی-دو در سطح اطمینان  $0,05$  می باشد. بنابراین کفایت الگوهای پیشنهادی با این آزمون نیز تایید می شود. اما با توجه به مقادیر P و  $Q^*$  مشاهده می شود که این مقادیر در کلیه تاخیر های ۱۲، ۲۴، ۳۶ و ۴۸ برای مدل M1 بیشتر از مدل M2 می باشد و به این علت مدل M1 بر مدل M2 ارجحیت دارد. ماتریس همبستگی بین پارامتر های برآورد شده برای دو مدل M1 و M2 در جدول ۲ نشان داده شده است. با توجه به این جدول

۲۰۰۱، ۲۰۰۲ و ۲۰۰۳ (سه سال) به انتهای آمار قبلی اضافه گردید و مجدداً الگوی مناسب بر داده‌ها برآش داده شد و پس از طی کلیه مراحل لازم که قبل از تشریح گردید، در هر مرحله مقادیر بارندگی ماهانه برای سال‌های بعد پیش‌بینی گردید. نتایج این کار برای ایستگاه مشهد به همراه مقادیر واقعی بارندگی، جهت مقایسه، در شکل ۷ آورده شده است.



شکل(۷) نمودار بارندگی ماهیانه واقعی (خط ممتدا) و پیش‌بینی شده (خط منقطع) در ایستگاه مشهد برای سه سال ۲۰۰۱ الی ۲۰۰۳

جهت تعیین نقش تعداد سال‌های آماری در انتخاب مدل باکس و جنکینس برای ایستگاه سینوپتیک مشهد، با توجه به آمار موجود (۱۹۵۱-۲۰۰۳)، دوره‌های آماری متفاوت (۱۲ الی ۵۰ ساله) با فاصله ۱۰ سال و در نظر گرفتن سال ۲۰۰۲ به عنوان پایان هر دوره آماری، تشکیل گردید و براساس آنها مدل مناسب باکس جنکینس برای هر دوره انتخاب شد. سپس برای ۳ سال از انتهای دوره آماری، یعنی سال ۲۰۰۳ الی ۲۰۰۵ مقادیر بارندگی ماهانه پیش‌بینی گردید. حداقل آمار در دوره آماری اول برابر ۵۲ سال و حداقل آمار در دوره آماری آخر برابر ۱۲ سال در نظر گرفته شد. دوره‌های آماری انتخابی، مدل مناسب باکس جنکینس و مقادیر بارندگی سالانه پیش‌بینی شده با استفاده از مدل‌های مناسب در هر دوره آماری برای سال‌های ۲۰۰۳ الی ۲۰۰۵، در جدول ۴ و نتایج پیش‌بینی بارندگی ماهیانه و سالانه در دوره‌های مختلف آماری در جدول ۵ با هم مقایسه شده‌اند. بر اساس این جداول، مدل‌های انتخابی در سه دوره آماری اول با طول دوره ۵۲، ۴۲ و ۳۲ سال، یکسان هستند ولی مدل انتخابی برای دو دوره آخر با طول دوره ۲۲ و ۱۲ سال با بقیه متفاوت و درصد اختلاف مقادیر بارندگی پیش‌بینی شده با دو مدل اخیر، بیشتر از سه دوره اول می‌باشد. طول دوره آماری در دو مورد آخر کوتاه است و لذا نتایج

مدل‌های دیگری را که دارای پارامترهای بیشتری هستند، بر داده‌ها برآش داد. به همین منظور چهار مدل که دارای پارامترهای بیشتری نسبت به مدل M1 بود، شامل مدل‌های SARIMA(2,0,0)(0,1,4)<sub>12</sub>، SARIMA(0,0,2)(0,1,4)<sub>12</sub>، SARIMA(1,0,0)(4,1,0)<sub>12</sub> و SARIMA(1,0,0)(4,1,4)<sub>12</sub> بر داده‌ها برآش داده شد. بررسی‌های انجام شده روی هریک از مدل‌های مذکور، مشخص ساخت که افزودن هیچ پارامتری الگوی M1 را به طور معنی داری بهتر نمی‌کند و حذف هر پارامتری الگوهای مذکور را بدتر می‌نماید. بنابراین مدل M1 (باتوجه به ارجحیت بر مدل M2) پس از انجام اصلاحاتی که ذکر شد، به عنوان مدل نهایی انتخاب می‌گردد.

جدول(۳) پارامترهای مدل برآش یافته M1 بر داده‌های بارندگی ایستگاه مشهد

پارامتر	$\theta_{4,12}$	$\theta_{2,12}$	$\theta_{1,12}$	$\phi_1$
برآورد	-۰,۱۳۵۰	۰,۹۰۳۲	۰,۱۳۲۷	
مقدار	-۳,۱۱	۲,۵۲	۲۴,۲۶	۳,۲۱
مقدار	۰,۰۰۲	۰,۰۱۲	۰,۰۰۰	۰,۰۰۱

با یادآوری روابط (۳) و (۵) و استفاده از جدول ۳ مدل نهایی را می‌توان به شکل زیر نوشت:

$$y_t^* = y_{t-12}^* + 0.1327y_{t-1}^* - 0.1327y_{t-13}^* + a_t - 0.9032a_{t-12} - 0.1295a_{t-24} + 0.135a_{t-48} \quad (V)$$

که در آن  $y_t^* = \ln(y_t' + 1)$  و  $y_t' = y_t^*$  و  $a_t$  جملات اغتشاش خالص (با میانگین  $0,0293$  و پراش  $0,0293$ ) می‌باشد ( $a_t$  سری باقی مانده‌های مدل M1 است).

#### پیش‌بینی

براساس مدل نهایی (رابطه ۷)، مقادیر پیش‌بینی شده سری  $y_t'$  با استفاده از نرم افزار Minitab به دست آمد. جهت ارزیابی و بررسی کارآیی مدل انتخابی در پیش‌بینی بارندگی ماهیانه و سالانه، آمار واقعی بارندگی ایستگاه سینوپتیک مشهد از سال ۲۰۰۱ الی ۲۰۰۳ تهیه گردید. با استفاده از آمار قبلی و آمار جدید، برای ایستگاه مشهد، در سه مرحله به ترتیب آمار سال ۱ (یک سال)، آمار سال‌های ۲۰۰۱ و ۲۰۰۲ (دو سال) و آمار سال‌های

جدول (۴) مقایسه مدل انتخابی و مقادیر بارندگی پیش بینی شده سالانه ایستگاه مشهد در سه سال متوالی بر اساس مدل سازی با دوره های آماری مختلف با فاصله ۱۰ سال (مقدار بارندگی بر حسب میلیمتر)

دوره آماری	جمع بارندگی سالانه پیش بینی شده توسط مدل در سالهای آینده (میلیمتر)	مدل انتخابی	سال		
			۲۰۰۳	۲۰۰۴	۲۰۰۵
۱۹۵۱-۲۰۰۲	SARIMA(1,0,0)(0,1,4) <sub>12</sub>		۲۲۳	۲۱۱	۱۹۱
۱۹۶۱-۲۰۰۲	SARIMA(1,0,0)(0,1,4) <sub>12</sub>		۲۲۷	۲۱۷	۱۹۱
۱۹۷۱-۲۰۰۲	SARIMA(1,0,0)(0,1,4) <sub>12</sub>		۲۳۹	۲۲۶	۱۹۵
۱۹۸۱-۲۰۰۲	SARIMA(1,0,0)(0,1,2) <sub>12</sub>		۱۸۷	۱۸۷	۱۸۷
۱۹۹۱-۲۰۰۲	SARIMA(0,0,0)(0,1,3) <sub>12</sub>		۱۹۰	۱۳۵	۱۳۳
حد اکثر درصد اختلاف (a)			۲۱/۷۵	۴۰/۳	۳۱/۸
حد اکثر درصد اختلاف بدون احتساب ردیف آخر (b)			۲۱/۷۵	۱۷/۳	۴/۱
حد اکثر درصد اختلاف بدون احتساب دو ردیف آخر (c)			۶/۷	۶/۶	۲/۱
(b) : مقایسه چهار دوره اول					(a) : مقایسه کل دوره ها
(c) : مقایسه سه دوره اول					

نتایج نشان داد که تفاوت مقادیر آزمون شده، اختلاف معنی داری با صفر ندارند و به عبارت دیگر مقادیر میانگین واقعی بارندگی ماهیانه با مقادیر پیش بینی شده آنها در سطح اطمینان ۹۵٪ با هم تفاوت معنی داری ندارند. مقایسه میانگین درازمدت بارندگی سالانه واقعی و پیش بینی شده در ایستگاه مشهد نشان داد که مقدار میانگین بارندگی سالانه پیش بینی شده به میزان ۲۱،۲۸ درصد میانگین بارندگی سالانه پیش بینی شده آن است (احمدی، ۱۳۸۴). نتایج کمتر از مقدار واقعی آن است. نتایج الگوسازی مجدد با دوره آماری طولانی تر (۱۱ تا ۳ سال در انتهای دوره آماری نسبت به آمار اولیه) برای ایستگاه سینوپتیک مشهد، تغییری در الگوی انتخابی ایجاد نکرد و همان الگوی قبلی به عنوان بهترین الگو تشخیص داده شد. با وجود این مطلب عمومیت ندارد و ممکن است با افزایش طول دوره آماری و ماهیت داده ها، مدل انتخابی جدید با مدل اولیه تفاوت داشته باشد. اگر در ایستگاه مورد مطالعه جهت پیش بینی بارندگی سالیانه (بدون اینکه داده های واقعی در هر سال به انتهای آمار اضافه شود و مجدد الگوسازی گردد) از مدل تعیین شده استفاده شود، به طور متوسط ۴ درصد خطای بدست می آید.

براساس نتایج بدست آمده در این تحقیق، حداقل آمار لازم برای الگوسازی سری های زمانی فصلی ( $L=12$ ) جهت دستیابی به نتایج قابل قبول، ۳۰ سال تعیین می گردد و مدل های باکس-جنکینس که با استفاده از دوره های آماری بالای ۳۰ سال به دست

آنها خیلی قابل اعتماد نمی باشد. آشگر طوسی (۲) در تحقیقی مشابه، سری زمانی باران ماهیانه مشهد را تحلیل کرده است. نامبرده برای دوره آماری ۲۰۰۲-۱۹۷۰ مدل SARIMA(2,0,2)(2,1,0)<sub>12</sub> را گزارش نمود در حالی که مدل تعیین شده در این تحقیق، SARIMA(1,0,0)(0,1,4)<sub>12</sub> می باشد. در جدول ۶ مقایسه ای بین نتایج برآورده با مدل نامبرده و مدل بدست آمده در این تحقیق ارائه شده است. گرچه در مقیاس ماهیانه تفاوت های زیاد است، ولی این تفاوت ها در مقیاس سالانه کمتر می باشد. دلیل احتمالی تفاوت در نوع مدل احتمالاً مربوط به این می شود که آشگر طوسی (۲) از داده های اصلی برای مدل سازی استفاده کرده و تبدیل پایداری واریانس روی داده ها انجام نداده است.

### نتیجه گیری

در این تحقیق اطلاعات بارندگی ماهیانه مشهد با مدل سری زمانی ARIMA برازش داده شد. در دوره پنجماه ساله ۲۰۰۰-۱۹۵۱ دو مدل مدنظر قرار گرفت و نشان داده شد که مدل M1 (رابطه ۷) به عنوان مدل نهایی بر داده های سری زمانی بارندگی ماهیانه ایستگاه مشهد قابل برازش است. جهت بررسی کارآیی مدل انتخابی، مقادیر پیش بینی شده بارندگی ماهیانه در مقابل مقادیر واقعی آنها با استفاده از آزمون  $t$  مورد آزمون واقع شدند.

جدول (۵) مقایسه مقادیر بارگاهی پیش بینی شده ماهیانه و مساله استگاه مشهد در سه مساله متوجه (مقابل بارگاهی . ت حسب ملکتر )

a : حد اکر درصد اختلاف بین کیک گزینه ها  
b : حد اکر درصد اختلاف بین گزینه ها بدون احساب گزینه آخر (۱۹۹۱)  
c : حد اکر درصد اختلاف بین گزینه ها بدون احساب دو گزینه آخر (۱۹۹۱ و ۱۸۷۱)

جدول (۶) مقایسه مقادیر بارندگی پیش بینی شده ماهیانه ایستگاه مشهد در سه سال متالی بر اساس مدل سازی با دوره آماری ۱۹۷۰-۲۰۰۲ توسط احمدی و آشگر طوسی (مقدار بارندگی بر حسب میلیمتر)

	سال ۲۰۰۵			سال ۲۰۰۴			سال ۲۰۰۳			ماه	
	درصد	بررسی	درصد	بررسی	درصد	بررسی	درصد	بررسی			
			حاضر		مرجع ۲ اختلاف		حاضر				
۱۶/۶۷	۲۵	۳۰	۱۰	۲۷	۳۰	۲۸/۱	۲۳	۳۲	ژانویه		
۶۰	۱۴	۳۵	۷۸/۷	۱۰	۴۷	۶۱/۵	۱۵	۳۹	فوریه		
۱۶/۶۷	۴۰	۴۸	۲۸/۳	۲۸	۵۳	۳۷/۵	۳۵	۵۶	مارس		
۳۷/۵	۲۰	۳۲	۶۳/۶	۱۶	۴۴	۷۹/۶	۱۱	۵۴	آوریل		
۵۶/۷	۳۰	۱۳	۴۰	۲۵	۱۶	۴۲/۳	۱۵	۲۶	می		
.	۲	۲	۳۳/۳	۲	۳	۰	۲	۲	ژوئن		
۸۰۰	۸	۰	۶۰۰	۶	۰	۶۶/۷	۳	۱	ژوئیه		
۱۰۰	۰	۱	۱۰۰	۰	۱	۱۰۰	۰	۱	اوت		
۱۰۰	۰	۱	۱۰۰	۰	۱	۰	۱	۱	سپتامبر		
۵۰	۲	۴	۶۰	۲	۵	۲۵	۴	۳	اکتبر		
۶۶/۷	۲۴	۸	۷۰/۸	۲۴	۷	۷۶	۲۷	۷	نوامبر		
۳۰	۳۰	۲۱	۴۴	۳۴	۱۹	۵۴	۳۷	۱۷	دسامبر		
۰	۱۹۰	۱۹۰	۱۸/۶	۱۸۴	۲۲۶	۲۷/۶	۱۷۳	۲۳۹	جمع		

آماری، مدل های مناسب تغییر می کند و برآوردها ممکن است به مقدار بیشتری با واقعیت تفاوت داشته باشند.

آمده اند، نتایج مشابه و نزدیکی را به همراه دارند و اختلاف نتایج پیشینی ها ناچیز است. نشان داده شد که با کاهش طول دوره

#### منابع

۱. احمدی، فیروز، ۱۳۸۴، پیش بینی بارندگی سالانه استان خراسان با استفاده از سری های زمانی، پایان نامه کارشناسی ارشد آبیاری و زهکشی، دانشگاه فردوسی مشهد.
۲. آشگر طوسی، شادی، ۱۳۸۳، پیش بینی وقوع خشک سالی در استان خراسان و بهینه سازی الگوی کشت جهت سازگاری با آن، پایان نامه کارشناسی ارشد آبیاری و زهکشی، دانشگاه فردوسی مشهد.
۳. جمشیدی، وحید، ۱۳۶۸، تجزیه و تحلیل درجه حرارت میزان بارندگی شهر تهران به وسیله سری های زمانی، پایان نامه کارشناسی ارشد آمار، دانشکده علوم پایه، دانشگاه تربیت مدرس.
۴. خردمندی، منوچهر و عساکره، حسین، ۱۳۸۰، الگوسازی ARIMA برای متوسط درجه حرارت سالانه هوا در جاسک، سومین سمینار احتمال فرآیندهای تصادفی، دانشگاه اصفهان.
۵. رحیم زاده، فاطمه، ۱۳۷۲، تجزیه و تحلیل سری زمانی فشار در ایستگاه تهران-مهرآباد، نیوار، شماره ۱۷-۲۰، ص ۵۸-۶۸.
۶. رسولی، علی اکبر، ۱۳۸۱، مدل سازی از عناصر اقلیمی شمال غرب کشور، پیش بینی مقادیر درجه حرارت ماهانه شهر تبریز به

- روشن مدل آریما، دانشگاه تبریز، مجله دانشکده علوم انسانی و اجتماعی، شماره ۸.
۷. مالکی، محمد، ۱۳۶۸، تجزیه و تحلیل و مدل بندی درجه حرارت و بارندگی در غرب کشور، پایان نامه کارشناسی ارشد آمار، دانشگاه شهید بهشتی.
۸. مشکاتی، محمد رضا، ۱۳۷۱، مدل های استوکاستیک عوامل اقلیمی شهر تهران، مجله نیوار، شماره ۱۳ تا ۱۶، ص ۳۶-۴۶.
۹. نیرومند، حسینعلی و بزرگ نیا، ابوالقاسم (مترجمین)، ۱۳۷۲، مقدمه ای بر تحلیل سری های زمانی، تالیف سی چنفیلد، مشهد، دانشگاه فردوسی مشهد.
10. Bowerman, B.L. and O'Connel, R.T. 1987, "Time Series Forecasting", Second Edition , PWS Publisher.
11. Box, G.E.P. and Jenkins, G.M. and Reinsel, G.C., 1994 , "Time Series Analysis:Forcasting and Control", 3 Edition, Holden-Day.
12. Jones, P.D, R.S Bradley, H.F.Diaz, P.M.Kelly, and T.M.L. Wigley, 1986, "Northern hemisphere surface air temperature variations: 1851-1984", Journal of Climate and Applied Meteorroology 25: 161-179.
13. Jones, P.D., C.B. Raper and T. Wigley 1986: Southern hemisphere surface temperature variation, 1851-1984. Journal of Climate and Applied Meteorroology, 25, 1213, 1230.
14. Jones, P. D., T. M. L. Wigley, and P. B. Wright. 1986. Global temperature variation between 1861 and 1984. Nature 322: 430-434.
- 15 Hansen.J. and S. Lebedeff. 1988 , "Global Surface Air Temperature",Geo Lett. 15, 323-326.