

مقاله پژوهشی

کاربرد روش الگوریتم ژنتیک در برآورد پارامترهای سری زمانی خطی به منظور پیش‌بینی خشکسالی

عباس عباسی^۱، کیوان خلیلی^{۲*}، جواد بهمنش^۳ و اکبر شیرزاد^۴

تاریخ دریافت: ۱۳۹۷/۰۹/۱۹

تاریخ پذیرش: ۱۳۹۸/۰۹/۲۴

۱- دانش‌آموخته دکتری مهندسی منابع آب، گروه مهندسی آب، دانشکده کشاورزی، دانشگاه ارومیه، ارومیه

۲- استادیار گروه مهندسی آب، دانشکده کشاورزی، دانشگاه ارومیه، ارومیه

۳- استاد گروه مهندسی آب، دانشکده کشاورزی، دانشگاه ارومیه، ارومیه

۴- دانشیار گروه مهندسی عمران، دانشکده مهندسی عمران، دانشگاه صنعتی ارومیه، ارومیه

* نویسنده مسئول، پست الکترونیکی: Khalili2006@gmail.com

چکیده

به‌طور متداول برآورد پارامترهای سری زمانی خطی بر اساس روش‌های گرافیکی و تقریبی است. بنابراین استفاده از رویکردی جدید جهت افزایش سرعت و سهولت در دسترسی به بهترین مدل سری زمانی می‌تواند نقش مهمی در استفاده از این روش در پیش‌بینی وقایع هیدرولوژیک داشته باشد. در این تحقیق جهت تخمین پارامترهای سری زمانی آرما از رویکرد بهینه‌سازی بر مبنای الگوریتم ژنتیک استفاده شده است. در این مطالعه با استفاده از روش ترکیبی الگوریتم ژنتیک - آرما پیش‌بینی خشکسالی در سه ایستگاه منتخب حوضه آبریز دریاچه ارومیه شامل تبریز، سقز و ارومیه بر اساس شاخص خشکسالی SPEI مورد بررسی قرار گرفت. نتایج نشان داد که بر اساس آزمون BDS در هر سه ایستگاه و در همه مقیاس‌های زمانی سری قابلیت پیش‌بینی پذیری را دارد. همچنین به منظور بررسی میزان قابلیت اطمینان به مدل پیش‌بینی، از آماره Ljung-Box استفاده شد که مقادیر p-value آن در همه ایستگاه‌ها و مقیاس‌های زمانی بزرگتر از ۰/۰۵ می‌باشد که نشانگر تصادفی بودن باقی‌مانده‌های مدل و قابل اطمینان بودن آن است. همچنین بهترین مدل سری زمانی در مقیاس‌های زمانی مختلف محاسبه و بر اساس آن پیش‌بینی شاخص SPEI انجام گرفت. نتایج بخش پیش‌بینی نشان داد که روش ترکیبی الگوریتم ژنتیک - آرما در مقیاس‌های زمانی بلندمدت شاخص SPEI در همه ایستگاه‌ها از دقت مناسب برخوردار است، ولی در مقیاس‌های زمانی کوتاه‌مدت عملکرد آن مناسب نمی‌باشد.

واژه‌های کلیدی: الگوریتم ژنتیک، تخمین پارامتر، خشکسالی، دریاچه ارومیه، سری زمانی

Application of Genetic Algorithm in Estimation of Linear Time Series Parameters for the Purpose of Drought Prediction

A Abbasi¹, K Khalili^{2*}, J Behmanesh³ and A Shirzad⁴

Received: December 10, 2018 Accepted: December 15, 2019

¹Ph.D. Graduate, Dept. of Water Eng., Faculty of Agric., Univ. of Urmia, Urmia, Iran

² Assist. Prof., Dept. of Water Eng., Faculty of Agric., Univ. of Urmia, Urmia, Iran

³ Prof., Dept. of Water Eng., Faculty of Agric., Univ. of Urmia, Urmia, Iran

⁴ Assoc. Prof., Faculty of Civil Engineering, Urmia University of Technology, Urmia, Iran

* Corresponding Author, Email: khalili2006@gmail.com

Abstract

So far, the linear time series parameters are estimated, generally based on graphical and approximate methods. Therefore, the use of a new approach to increase the speed and ease of access to the best time series model can play an important role in using this method for predicting hydrological events. In this research, an optimization approach based on genetic algorithm has been used to estimate the ARMA time series parameters. A hybrid of Genetic Algorithm-ARMA method was used to drought prediction at three selected stations in the Urmia Lake basin, including Tabriz, Saqhez and Urmia, based on the SPEI drought index. The results showed that according to the BDS test, the model had the ability to predict the drought in all three stations and in all time scales. The Ljung-Box statistic was also used to evaluate the reliability of the prediction model. Its p-value at all stations and time-scales were greater than 0.05 which indicated the residuals of models were random and reliable. Also, the best time series model was calculated at different time scales and based on this, the SPEI index was predicted. The results of the prediction section showed that ARMA-GA hybrid method had a high accuracy at all long-term time scales of SPEI index at all the stations, but its performance was not suitable for short-term time scales.

Key words: Drought, Estimated parameters, Genetic Algorithm, Time series, Urmia Lake

متوالی بین مشاهدات وجود دارد که این از ویژگی‌های بارز مدل $ARIMA^1$ و $SARIMA^2$ است (طوفانی و همکاران ۲۰۱۱). شریفان و قهرمان (۲۰۰۷) مدل‌سازی پیش‌بینی بارندگی ماهانه استان گلستان را در ۴ ایستگاه انجام دادند و به این نتیجه رسیدند که روش $SARIMA$ مدل مناسبی به منظور پیش‌بینی بارندگی ماهانه این استان می‌باشد. معروفی و همکاران (۲۰۱۴) با استفاده از داده‌های ۳۵ ساله ۴ ایستگاه سینوپتیک در همدان از روش $SARIMA$ به منظور پیش‌بینی خشکسالی استفاده نمودند که نتایج این تحقیق نشان داد که مقادیر پیش‌بینی شده از همبستگی بالا با مقادیر مشاهداتی برخوردار هستند. معروفی و همکاران (۲۰۱۵) جهت برآورد متوسط دمای ماهانه در ایستگاه‌های سینوپتیک

مقدمه

یکی از روش‌های پیش‌بینی خشکسالی، استفاده از تحلیل سری‌های زمانی می‌باشد که از سال ۱۹۷۰ به بعد برای پیش‌بینی و تحلیل به سرعت توسعه پیدا کرده است. مدل‌های سری زمانی از دو بخش اصلی شامل بخش قطعی (مؤلفه جبری) و بخش تصادفی تشکیل شده‌اند که بخش جبری با استفاده از داده‌های مشاهداتی و بخش تصادفی با استفاده از روش‌های مختلف استوکاستیک به دست می‌آید. بنابراین، با توجه به ماهیت وقایع هیدرولوژیک، سری‌های زمانی می‌توانند در صورت انتخاب درست مدل و انجام محاسبات صحیح، سازگاری و مطابقت ویژه‌ای با سری‌های زمانی داده‌های هیدرولوژیک داشته باشند (سالاس ۱۹۹۳). در بسیاری از سری‌های زمانی همبستگی

¹ Autoregressive integrated moving average model

² Seasonal autoregressive integrated moving average model

قدیمی ایران از روش سری زمانی ARIMA استفاده

نمودند. نتایج نشان داد که مدل‌های غیرفصلی ARIMA جهت تعیین روند پارامتر دما در گستره ایران از قابلیت بالایی برخوردار می‌باشد و با استفاده از الگوهای سری زمانی به دست آمده می‌توان پارامتر دما را در دوره‌های ماهانه پیش‌بینی نمود. میشر و دسای (۲۰۰۵) و دوردو (۲۰۱۰) مدل ARIMA را برای پیش‌بینی خشکسالی از طریق برآورد مقادیر شاخص SPI^۳ توسعه دادند که نتایج تحقیق آنها نشان داد که به دلیل تأثیرات فصلی بر داده‌ها تنها این مدل می‌تواند شاخص SPI را برای یک ماه پیش‌بینی نماید. میشر و دسای (۲۰۰۶) به علت تصادفی بودن پارامترهای مؤثر در وقوع و شدت خشکسالی و امکان در نظر گرفتن این پدیده به عنوان یک مدل استوکاستیک، بررسی شدت خشکسالی را با استفاده از سری‌های زمانی انجام و مدل فصلی ARIMA را برای این منظور پیشنهاد نموده‌اند. میشر و همکاران (۲۰۰۷) با استفاده از مدل تلفیقی استوکاستیک خطی و شبکه عصبی مصنوعی غیرخطی نسبت به پیش‌بینی خشکسالی با استفاده از شاخص SPI در حوضه رودخانه کانسباتی هند پرداختند که نتایج این تحقیق نشان داد که مدل هیبرید در مقایسه با مدل‌های دیگر از دقت بالاتری برخوردار است. زارعی و همکاران (۲۰۱۶) با استفاده از شاخص RDI^۴ و مدل سری زمانی ARMA نسبت به پایش و پیش‌بینی خشکسالی در ایستگاه سینوپتیک تهران اقدام نمودند. نتایج این تحقیق نشان داد که مدل سری زمانی در پیش‌بینی خشکسالی از دقت قابل قبولی برخوردار است. شاخص‌های رایج تعیین خشکسالی مانند SPI صرفاً از میزان بارش روزانه و یا ماهانه استفاده می‌نمایند. اما صرفاً در نظر گرفتن عامل بارش در تعیین خشکسالی یک منطقه، از جمله انتقاداتی است که بر شاخص‌های خشکسالی مبتنی بر بارش خصوصاً شاخص SPI وارد

است و بسیاری از محققان بر اهمیت درجه حرارت بر شرایط خشکسالی تأکید نموده‌اند که منجر به افزایش تقاضای آب در اثر افزایش تبخیر- تعرق می‌گردد (کمپس و همکاران ۲۰۰۸). با این اوصاف اثر افزایش دما در تشدید خشک شدن به مراتب بیشتر از کاهش بارندگی است و این می‌تواند نقش مؤثر دما در تقویت یا تضعیف خشکسالی را تأیید کند. بر همین اساس، می‌توان شاخص خشکسالی استاندارد شده بارش و تبخیر- تعرق (SPEI^۵) را که ویسنت سرانو و همکاران (۲۰۱۰) عرضه کرد به عنوان شاخصی مناسب در تعیین خشکسالی مورد استفاده قرار داد. در این شاخص سه متغیر بارش، دما و تبخیر- تعرق پتانسیل (PET^۶) لحاظ شده است و براساس میزان بارندگی و درجه حرارت و محاسبه تبخیر- تعرق و بیلان آب می‌تواند اثرات تغییرات درجه حرارت را در ارزیابی خشکسالی لحاظ نماید. حسن‌زاده و همکاران (۲۰۱۲) جهت پایش خشکسالی در ایستگاه سینوپتیک تبریز از روش‌های فراکاوشی الگوریتم ژنتیک و الگوریتم شبیه‌سازی حرارتی استفاده کردند. نتایج این تحقیق نشان داد که روش الگوریتم ژنتیک در تحلیل شاخص SPI نتایج مطلوبی ارائه می‌دهد. محاسبه پارامترهای سری زمانی ARMA عموماً با استفاده از روش‌های نموداری و تقریبی و استفاده از روش حداکثر درستی انجام می‌گیرد که استفاده از این روش در تعیین بهترین سری زمانی که بر داده‌ها مطابقت مناسب داشته باشد عموماً وقت‌گیر و تقریبی می‌باشد. اخیراً از روش الگوریتم ژنتیک به عنوان یک ابزار بهینه‌سازی جهت تعیین ضرایب و پارامترهای سری زمانی ARMA در برخی مطالعات استفاده شده است. برای مثال، بلیگانیس و همکاران (۲۰۰۲)، پنگ و چن (۲۰۰۳)، ابوهامور و همکاران (۲۰۱۲) و ارورال و همکاران (۲۰۱۶)، کیم و کیم (۲۰۱۷) جهت پیش‌بینی و تخمین پارامترها و مرتبه

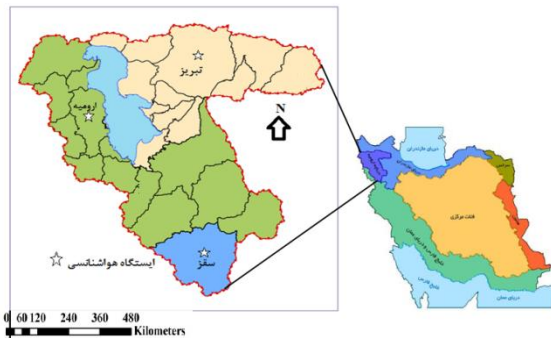
^۵ Standard precipitation and evapotranspiration

^۶ Potential evapotranspiration

^۳ Standardized Precipitation Index

^۴ Reconnaissance Drought Indices

حوضه آبریز دریاچه ارومیه واقع در شمال غرب ایران با مساحت ۵۱۸۷۶ کیلومتر مربع یکی از شش حوضه آبریز اصلی کشور است. دریاچه ارومیه در سال‌های اخیر با تنش آبی شدیدی روبرو بوده است و حجم و سطح آن کاهش قابل ملاحظه‌ای داشته است به نحوی که به عنوان یکی از نموده‌های بارز بحران آب در ایران از آن یاد می‌شود. پژوهشگران یکی از دلایل اصلی خشک شدن دریاچه را به وقوع خشکسالی‌های مستمر در سطح این حوضه نسبت می‌دهند که خشک شدن آن می‌تواند بر بخش‌های مختلف اقتصادی، اجتماعی، سلامت و ... این منطقه تبعات منفی بسیاری تحمیل نماید. لذا در این تحقیق، جهت بررسی و پیش‌بینی خشکسالی از داده‌های دما و بارش سه ایستگاه سینوپتیک تبریز، ارومیه و سقز که به عنوان شاخص‌ترین ایستگاه‌های دارای آمار بلندمدت واقع در سه بخش شرقی، غربی و جنوبی دریاچه ارومیه استفاده گردید. موقعیت این ایستگاه‌ها در شکل ۱ و مشخصات آنها در جدول ۱ ارائه شده است.



شکل ۱- موقعیت جغرافیایی ایستگاه‌های منتخب و منطقه مورد مطالعه.

شاخص SPEI

این شاخص مبتنی بر محاسبه بیلان آب است و برای محاسبه آن به داده‌های ماهانه بارندگی و دمای هوا نیاز است. معادله بیلان آب به صورت رابطه ۱ تعریف شده است.

مدل سری زمانی از روش الگوریتم ژنتیک در موضوعات مختلف، استفاده کرده‌اند که نتایج این تحقیقات نشانگر بهبود نتایج سری زمانی ARMA با استفاده از تلفیق آن با روش الگوریتم ژنتیک می‌باشد. به دلیل رشد جمعیت جهان و توسعه بخش‌های کشاورزی، انرژی و صنعت، نیاز به آب در دهه‌های اخیر به طور فزاینده‌ای افزایش یافته است، آن چنان که کمبود آب تقریباً هر ساله در بسیاری از بخش‌های کره زمین اتفاق می‌افتد که یکی از علل مهم کمبود آب، رخداد خشکسالی است. با توجه به قرارگیری ایران در یک منطقه خشک و نیمه‌خشک تا کنون خشکسالی‌های پی در پی اثرات زیان‌باری بر معیشت و زندگی مردم به وجود آورده که نمونه بارز آن وضعیت حاکم بر دریاچه ارومیه در سال‌های اخیر می‌باشد. بنابراین، پیش‌بینی خشکسالی برای مدیران و تصمیم‌گیران بخش آب و سایر بخش‌های اقتصادی این امکان را فراهم می‌کند که با این بلای طبیعی به صورت بهتر مقابله کنند. با توجه به اینکه عموماً از روش‌های گرافیکی و تقریبی برای تخمین بهترین مدل سری زمانی استفاده می‌گردد که این روش‌ها زمان‌بر و تقریبی می‌باشد، لذا استفاده از رویکردی مبتنی بر بهینه‌سازی جهت دستیابی به بهترین مدل سری زمانی می‌تواند نقش مهمی در افزایش دقت مدل سری زمانی و کاهش زمان دستیابی به مدل برتر گردد. بر این اساس، با توجه باینکه تا کنون در پیش‌بینی خشکسالی از روش ترکیبی سری زمانی الگوریتم ژنتیک استفاده نگردیده است، لذا هدف این پژوهش برآورد پارامترهای سری زمانی ARMA از روش بهینه‌سازی الگوریتم ژنتیک به منظور پیش‌بینی شاخص خشکسالی SPEI در گام‌های زمانی مختلف در سه ایستگاه شاخص در حوضه آبریز دریاچه ارومیه می‌باشد.

مواد و روش‌ها

منطقه مورد مطالعه

تغرق پتانسیل، D اختلاف آنها و i شماره ماه مورد نظر است.

$$D_i = P_i - PET_i \quad [1]$$

که در این رابطه، P و PET به ترتیب بارندگی و تبخیر -

جدول ۱- مشخصات ایستگاه‌های مورد مطالعه.

ایستگاه	طول جغرافیایی	عرض جغرافیایی	ارتفاع (m)	متوسط بارش سالانه (mm)	متوسط دمای سالانه (°C)	دوره آماری (Year)
تبریز	۴۶ درجه و ۲۸ دقیقه شرقی	۳۸ درجه و ۸ دقیقه شمالی	۱۳۶۱	۲۶۰/۱	۱۲/۵	۱۹۵۶ - ۲۰۰۹
سقز	۴۶ درجه و ۱۵ دقیقه شرقی	۳۶ درجه و ۱۵ دقیقه شمالی	۱۵۲۲	۴۶۴/۵	۹/۵	۱۹۶۱ - ۲۰۰۹
ارومیه	۴۵ درجه و ۵ دقیقه شرقی	۳۷ درجه و ۶۷ دقیقه شمالی	۱۳۲۸	۲۶۱/۴	۱۱/۱	۱۹۵۱ - ۲۰۰۹

را دارند گسترش یافته است، اما همچنان به دو دلیل عمده استفاده از مدل‌های خطی در کارهای تجربی رواج دارد؛ اول اینکه بسیاری از اوقات تقریب خطی مدل‌های غیرخطی ساده‌تر بوده و به طور معمول اطلاعات چندانی از دست نمی‌رود. دوم اینکه با کنار گذاشتن استفاده از تقریب خطی یک فرآیند غیرخطی، طیف وسیعی از الگوهای غیرخطی مطرح می‌شود که در بسیاری از اوقات تعیین بهترین الگوی غیرخطی کاری بسیار دشوار بوده و گاهاً استفاده از الگوی غیرخطی نامناسب عواقب بسیار ناگوارتری از تقریب خطی فرآیند مورد نظر دارد. هدف اصلی از مدل‌سازی سری‌های زمانی، پیش‌بینی مقادیر آتی سری است. معروف‌ترین مدل مورد استفاده در مدل‌سازی سری‌های زمانی، مدل خود همبسته با میانگین متحرک (ARMA) است. این مدل از ترکیب نمودن مدل‌های خود همبسته (AR) با مرتبه p و میانگین متحرک (MA) با مرتبه q ایجاد می‌شود و به صورت مدل ARMA با مرتبه (p, q) ارائه می‌شود. ساختار کلی این مدل به صورت رابطه ۲ می‌باشد (سالاس و همکاران ۱۹۸۰).

$$Z_t = \sum_{i=1}^p \phi_i Z_{t-1} + \varepsilon_t - \sum_{i=1}^q \theta_i \varepsilon_{t-i} \quad [2]$$

که در این رابطه، ϕ_i و θ_i به ترتیب ضرایب مدل AR و MA و ε_t سری تصادفی و مستقل از زمان باقیمانده مدل می‌باشد که از توزیع نرمال با میانگین صفر و واریانس σ_ε^2 تبعیت می‌کند. نمایش فرآیند ARMA به صورت ARMA(p,q) تا حدی می‌تواند

معادلات مختلفی برای محاسبه PET وجود دارد که در استفاده از هریک از آنها محدودیتی برای SPEI وجود ندارد. در نسخه نخست SPEI از روش ساده تورنت‌وایت (۱۹۴۸) استفاده شده، ولی در نسخه دوم آن (در صورت عدم محدودیت داده) توصیه شده است که از روش پنمن - مانتیث فائو (آلن و همکاران ۱۹۹۸) استفاده شود. پس از محاسبه مقادیر D ، از رهیافتی همانند شاخص SPI برای محاسبه SPEI استفاده می‌شود. برای این منظور، ابتدا مجموع مقادیر متوالی سری D در پنجره‌های زمانی مختلف محاسبه می‌شود. جزئیات نحوه محاسبه شاخص SPEI در منبع ویسته - سرانو و همکاران (۲۰۱۰) ارائه شده است. در این تحقیق از مقیاس‌های زمانی ۱، ۳، ۶، ۱۲، ۲۴ و ۴۸ ماهه شاخص SPEI جهت پیش‌بینی خشکسالی توسط مدل سری زمانی ARMA استفاده گردید.

سری زمانی ARMA

سری‌های زمانی در هیدرولوژی از اوایل دهه ۱۹۶۰ توسط توماس، فایرینگ و یوجویچ مورد استفاده قرار گرفت و در دهه ۱۹۷۰ توسط باکس و جنکینز توسعه داده شد (خلیلی و همکاران ۲۰۱۴). در این مدل فرض می‌شود که فرآیند تولید داده‌های سری زمانی مبتنی بر یک الگوی خطی است که ساده‌ترین نوع آن مدل‌های سری زمانی خودهمبسته (AR) می‌باشد که بر اساس زنجیره مارکف بنا نهاده شده است. اگر چه در سال‌های اخیر ارائه و بسط مدل‌هایی نظیر شبکه عصبی مصنوعی که قابلیت تخمین مدل‌های پیچیده و غیرخطی

انجام پیش‌بینی پرداخت. برای این منظور آزمون‌های متعددی مطرح گردیده است که در این تحقیق از آزمون BDS استفاده می‌گردد. آزمون BDS بروک و همکاران (۱۹۹۱) می‌تواند برای بررسی تبعیت یک سری زمانی از فرآیند اخلال خالص^۸ بودن آن مورد استفاده قرار گیرد. در واقع چنانچه استقلال عناصر یک سری زمانی از یکدیگر اثبات شود و به تعبیری هیچ‌گونه همبستگی با یکدیگر نداشته باشند، امکان مدل‌سازی و پیش‌بینی چنین سری‌ای وجود ندارد. این آزمون بر این اساس قرار دارد که چنانچه عناصر سری زمانی $\{Y_t = Y_1, Y_2, Y_3, \dots, Y_T\}$ به صورت مستقل و یکسان توزیع شده باشد (i.i.d)، احتمال اینکه فاصله بین هر جفت از مشاهدات (به عنوان مثال Y_i و Y_j) کمتر از مقدار از پیش تعیین شده ξ باشد، بایستی به ازای تمامی مقادیر i و j یکسان باشد. برای ایجاد مجموعه‌ای حاوی جفت‌هایی از مشاهدات سری زمانی، در ابتدا زیربردارهایی از $\{Y_t\}$ به صورت زیر تعریف می‌شود:

$$Y_t^m = \{Y_t, Y_{t-1}, \dots, Y_{t-m+1}\} \quad [۳]$$

که در این رابطه، m بیانگر تعداد جفت‌های متوالی است که بعد محاط خوانده می‌شود. $C_{m,\xi}$ بیانگر تعداد نقاطی است که در فضای m بعدی در درون کره‌ای به شعاع ξ قرار گرفته و محاط می‌شوند. بروک و همکاران (۱۹۹۶) نشان دادند که عبارت $C_m(\xi) - C_1^m(\xi)$ دارای توزیع نرمال با میانگین صفر و واریانس $\sigma_{m,\xi}$ به صورت رابطه ۴ است.

$$\sigma_{m,\xi}^2 = 4(k^m + 2 \sum_{j=1}^{m-1} k^{m-j} C_1^{2j} + (m-1)^2 C_1^{2m} - m^2 k C_1^{2m-2}) \quad [۴]$$

در نهایت آماره BDS که دارای دو پارامتر m و ξ می‌باشد به صورت رابطه زیر تعریف شد.

$$BDS_{m,\xi} = \sqrt{T_m} \frac{C_m(\xi) - C_1^m(\xi)}{\sigma_{m,\xi}} \quad [۵]$$

گمراه‌کننده باشد. به عنوان مثال، تصور عمومی در مورد فرآیند $ARMA(3,2)$ این است که این فرآیند به طور همزمان دارای سه جزء $AR(1)$ ، $AR(2)$ و $AR(3)$ و دو جزء میانگین متحرک $MA(1)$ و $MA(2)$ است. اما چنین تفسیری از این فرآیند این است که برای ارتباط یک متغیر با وقفه سوم خود، حتماً بایستی این متغیر با وقفه‌های اول و دوم خود نیز در ارتباط باشد که این نتیجه به وضوح غلط بوده و می‌توان در دنیای واقعی سری‌های زمانی متعددی یافت که در آن مقادیر یک سری زمانی با تأخیرهای چند دوره‌ای بر روی مقدار جاری خود تأثیرگذار باشد. به همین جهت، کاربردی‌تر است که یک فرآیند خودتوضیح میانگین متحرک^۷ به صورت $ARMA([p],[q])$ نمایش داده شود. به عنوان مثال، $ARMA([1,4],[2])$ بیانگر یک فرآیند $ARMA$ است که در آن مقدار جاری متغیر متأثر از وقفه اول و چهارم خود و وقفه دوم جزء تصادفی است. مدل‌سازی خطی سری‌های زمانی به معنای برازش یک مدل $ARMA$ بر روی داده‌های سری زمانی است. این برازش تنها بر نامانایی سری زمانی اولیه بایستی بر اساس روش‌های مرسوم به یک سری مانا دست پیدا کرد. از این رو چنانچه لازم باشد یک سری زمانی d بار تفاضل‌گیری شود تا مانا گردد و سپس آن را در قالب یک الگوی $ARMA(p,q)$ برآورد نمود، گفته می‌شود سری زمانی اولیه یک فرآیند خودتوضیح انباشته میانگین متحرک از مرتبه p ، d و q است و آن را با $ARIMA(p,d,q)$ نمایش می‌دهند.

آزمون BDS

مدل‌سازی خطی سری‌های زمانی مستلزم بررسی پیش‌بینی‌پذیری داده‌ها می‌باشد که در صورت احراز آن می‌توان با برازش الگوی مناسب بر روی داده‌ها، پارامترهای مدل را برآورد نموده و در نهایت به

^۸ Independent and identically distributed (i.i.d)

^۷ Autoregressive Moving Average

جزئی (PACF^{۱۰}) به ترتیب برای شناسایی مرتبه وقفه-های MA و AR مدل بهینه می‌باشد. اما بسیاری از سری‌های زمانی دارای ترکیب‌های پیچیده‌ای از وقفه-های MA و AR هستند که در این صورت ترسیم ACF و PACF نمی‌تواند به طور دقیق وقفه‌های بهینه را مشخص کند. روش دیگری که برای شناسایی مدل‌های ARMA (یا ARIMA) مطرح است استفاده از معیارهای اطلاعات نظیر آکایکه، حنان-کوئین و شوارتز (آیالو و همکاران ۲۰۱۲) می‌باشد که علیرغم اینکه هر یک از آنها دارای یک تابع هزینه مشخص برای تعیین مدل بهینه هستند، با مشکل فضای جستجوی بسیار بزرگ مواجه-اند. با افزایش تعداد وقفه‌های MA و AR فضای جستجو یا تعداد کل مدل‌های بالقوه با نرخ رشدی نمایی افزایش می‌یابد، به طوری که چنانچه p_{max} و q_{max} به ترتیب، بیانگر حداکثر وقفه‌های مورد نظر برای AR و MA باشند، فضای جستجو برابر است با $(2^{p_{max}+q_{max}} - 1)$ که به طور مشخص تخمین کلیه حالت‌ها، حتی در وقفه-های کوچک نیز به لحاظ زمانی بسیار هزینه‌بر خواهد بود. روش‌شناسی مدل ARMA-GA مشابه روش-شناسی باکس-جنکینز در مدل‌سازی سری‌های زمانی می‌باشد، با این تفاوت که به جای بررسی کلیه وقفه‌های AR و MA (به میزان $(2^{p_{max}+q_{max}} - 1)$)، با استفاده از الگوریتم ژنتیک به جستجوی هدفمند بخشی از فضای بالقوه پرداخته تا به یکی از بهترین مدل‌ها (و نه لزوماً بهترین مدل) با توجه به یکی از معیارهای اطلاعات دست یافته شود. در این تحقیق، مدل بهینه از بین مدل‌هایی با حداکثر ۵ وقفه AR و ۵ وقفه MA انتخاب شد. از این رو هر جواب در مدل ARMA-GA به مانند کروموزومی با طول ۱۰ است که ۵ ژن اول آن معرف عناصر AR(1) تا AR(5) و ژن‌های ششم تا دهم معرف عناصر MA(1) تا MA(5) می‌باشد (جدول ۲).

این آماره دارای توزیع نرمال با میانگین صفر و انحراف معیار یک است. بنابراین، فرض صفر مبنی بر i.i.d بودن آزمون BDS در سطح معنی‌داری ۹۵٪ رد می‌شد چنانچه شرط زیر ارضا می‌گردید.

$$|BDS_{m,\xi}| \geq 1.96 \quad [6]$$

روش الگوریتم ژنتیک

الگوریتم ژنتیک یک تکنیک جستجوی سراسری است که بر اساس مفاهیم ژنتیک طبیعی و فرضیه تکامل داروین شکل گرفته است. این الگوریتم نیز مانند سایر الگوریتم‌ها، دارای یک تابع هدف می‌باشد که بایستی حداقل یا حداکثر شود. در این تحقیق، از الگوریتم ژنتیک دودویی استفاده می‌شود. بدین نحو که انتخاب وقفه‌های مختلف AR و MA می‌تواند با در نظر گرفتن عدد یک برای وجود وقفه مورد نظر و عدد صفر برای عدم وجود آن انجام پذیرد. جهت آگاهی از جزئیات مربوط به الگوریتم ژنتیک می‌توان به هولاند (۱۹۶۲) و میخالویچ (۱۹۹۹) رجوع گردد.

روش ARMA-GA

در پیش‌بینی سری‌های زمانی برای دستیابی به یک مدل ARMA بهینه بایستی سه مرحله شناسایی، برآورد و بازبینی مدل طی شوند. در سال‌های گذشته تمرکز بسیاری از تحقیقات در زمینه سری‌های زمانی منوط به ارائه روش‌های نوین در برآورد مدل و یا بهبود نتایج پیش‌بینی صورت گرفته است؛ در حالی که بهبود فرآیند شناسایی مدل نیز می‌تواند بر سرعت دستیابی به مدل بهینه مؤثر باشد، زیرا انتخاب مدل نادرست باعث افزایش هزینه زمان دستیابی به مدل بهینه می‌شود. روش الگوریتم ژنتیک می‌تواند سبب سرعت در جستجوی فضای مدل‌های ARMA بالقوه برای دستیابی به مدل بهینه گردد. اولین روش شناسایی مدل که توسط باکس و جنکینز (۱۹۷۶) مطرح شد، استفاده از توابع خودهمبستگی (ACF^۹) و خودهمبستگی

¹⁰ Partial Autocorrelation Function

⁹ Autocorrelation Function

جدول ۲- مشخصات کروموزم مورد استفاده جهت مدل‌سازی مدل ARMA.

جواب نام	AR(1)	AR(2)	AR(3)	AR(4)	AR(5)	MA(1)	MA(2)	MA(3)	MA(4)	MA(5)
	X_{j1}	X_{j2}	X_{j3}	X_{j4}	X_{j5}	X_{j6}	X_{j7}	X_{j8}	X_{j9}	X_{j10}

استفاده شد.

که در آن، $X_{i,j} \in \{0,1\}$ می‌باشد (هر کروموزوم فقط عدد صفر یا یک را اختیار می‌نماید). انتخاب تعداد وقفه‌های بهینه بسیار مهم می‌باشد، چرا که اگر تعداد وقفه‌ها بیش از اندازه کم باشد، باقی‌مانده‌های حاصل از تخمین مدل رگرسیونی دارای رفتار یک الگوی اخلال خالص نبوده و دچار خود همبستگی می‌باشد. از سوی دیگر با افزایش تعداد وقفه‌ها، تعداد پارامترهای اضافی در مدل افزایش و در نتیجه درجه آزادی و قدرت آزمون ریشه واحد کاهش خواهد یافت. از این رو، برای انتخاب وقفه‌های بهینه به طور معمول از معیارهایی تحت عنوان معیارهای اطلاعات استفاده می‌شود. این معیارها در حالت کلی شامل دو جزء هستند که جزء اول تابعی از مجموع مجزورات باقیمانده‌های مدل اولیه برآزش شده و جزء دوم حاوی جمله‌ای است که زیان ناشی از کاهش درجه آزادی را در صورت افزودن وقفه‌های مختلف به مدل اولیه اندازه‌گیری می‌کند. سه معیار آکایکه، شوارتز و حنان - کوئین از مهم‌ترین معیارهای اطلاعات هستند که در این تحقیق از معیار شوارتز در مدل ARMA-GA استفاده شده است که رابطه آن به صورت رابطه ۷ می‌باشد.

$$SC = \frac{-2\ln(L)}{T} + \frac{\ln(T)k}{T} \quad [V]$$

که در این رابطه، T تعداد مشاهدات، k تعداد پارامترها و L بیانگر تابع درستنمایی است. در روش تلفیقی ARMA-GA، هدف کمینه کردن معیار شوارتز جهت انتخاب مدل برتر می‌باشد.

معیارهای ارزیابی

در این تحقیق، به منظور ارزیابی مدل‌های مورد نظر از معیارهای ضریب همبستگی، میانگین قدر مطلق خطا و جذر میانگین مربعات خطا به صورت زیر

$$r = \frac{\sum_{i=1}^n (Q_i - \bar{Q}_i)(\hat{Q}_i - \bar{\hat{Q}}_i)}{\sum_{i=1}^n (Q_i - \bar{Q}_i)^2 \sum_{i=1}^n (\hat{Q}_i - \bar{\hat{Q}}_i)^2} \quad [8]$$

$$MAE = \frac{1}{n} \sum_{i=1}^n |Q_i - \hat{Q}_i| \quad [9]$$

$$RMSE = \left(\frac{\sum_{i=1}^n (Q_i - \hat{Q}_i)^2}{n} \right)^{0.5} \quad [10]$$

که در این روابط Q_i مقدار داده مشاهداتی در گام زمانی i ام، \hat{Q}_i مقدار داده پیش‌بینی شده در همان زمان، n تعداد داده‌ها، \bar{Q}_i میانگین مقادیر مشاهداتی و $\bar{\hat{Q}}_i$ میانگین مقادیر پیش‌بینی شده می‌باشد. هر چه مقادیر R به یک نزدیکتر باشد و $RMSE$ و MAE به صفر نزدیکتر باشد، نتایج دقیق‌تر و قابل اعتمادتر می‌باشد. جهت پیش‌بینی شاخص SPEI در مقیاس‌های زمانی مختلف با استفاده از روش ARMA، ابتدا داده‌های شاخص مطابق روابط ۱۱ و ۱۲ نرمال گردید.

$$Y_i = \frac{X_{oi}}{X_{o \max}}, X_{oi} \geq 0 \quad [11]$$

$$Y_i = \frac{X_{oi}}{|X_{o \min}|}, X_{oi} < 0 \quad [12]$$

که در این روابط، Y_i مقادیر استاندارد شده شاخص، X_{oi} مقادیر شاخص مشاهداتی، $X_{o \max}$ حداکثر مقدار شاخص مشاهداتی و $X_{o \min}$ حداقل مقدار شاخص مشاهداتی می‌باشد. سپس ۸۰ درصد داده‌ها برای آموزش و ۲۰ درصد باقی‌مانده آن برای صحت‌سنجی انتخاب گردید.

نتایج و بحث

جهت بررسی پیش‌بینی پذیر بودن سری زمانی در گام‌های زمانی مختلف از آزمون BDS استفاده گردید.

جدول ۳- نتایج آزمون BDS در سه ایستگاه مورد مطالعه.

تبریز			ارومیه			سقز			گام زمانی SPEI (ماه)
آماره BDS	خطای استاندارد	احتمال	آماره BDS	خطای استاندارد	احتمال	آماره BDS	خطای استاندارد	احتمال	
۰/۰۰۲	۰/۰۰۱	۰/۵۶	۰/۰۰۳	۰/۰۰۱	۰/۶۶۹	۰/۰۰۲	۰/۰۰۵	۰/۰۴۲	۲
۰/۰۰۳	۰/۰۰۵	۰/۱۴۸	۰/۰۰۴	۰/۰۰۱	۰/۸۱۴	۰/۰۰۴	۰/۰۱۱	۰/۰۰۶	۳
۰/۰۰۴	۰/۰۰۶	۰/۱۳۷	۰/۰۰۵	۰/۰۰۳	۰/۵۰۵	۰/۰۰۵	۰/۰۰۹	۰/۰۴۲	۴
۰/۰۰۴	۰/۰۰۴	۰/۲۶۸	۰/۰۰۵	۰/۰۰۳	۰/۵۰۳	۰/۰۰۵	۰/۰۱۱	۰/۰۴۲	۵
۰/۰۰۴	۰/۰۰۴	۰/۲۹۸	۰/۰۰۵	۰/۰۰۳	۰/۵۵۱	۰/۰۰۵	۰/۰۰۸	۰/۰۸۶	۶
۰/۰۰۲	۰/۰۸۱	.	۰/۰۰۲	۰/۰۸۲	.	۰/۰۰۲	۰/۰۶۸	.	۲
۰/۰۰۳	۰/۱۲۲	.	۰/۰۰۴	۰/۱۲۵	.	۰/۰۰۴	۰/۱۰۳	.	۳
۰/۰۰۴	۰/۱۳۸	.	۰/۰۰۵	۰/۱۴۵	.	۰/۰۰۵	۰/۱۱۷	.	۴
۰/۰۰۴	۰/۱۳۹	.	۰/۰۰۵	۰/۱۴۷	.	۰/۰۰۵	۰/۱۱۷	.	۵
۰/۰۰۴	۰/۱۳۳	.	۰/۰۰۵	۰/۱۳۹	.	۰/۰۰۵	۰/۱۱۲	.	۶
۰/۰۰۲	۰/۱۲۹	.	۰/۰۰۳	۰/۱۲۵	.	۰/۰۰۳	۰/۱۱۶	.	۲
۰/۰۰۳	۰/۲۰۸	.	۰/۰۰۴	۰/۲۰۳	.	۰/۰۰۴	۰/۱۸۸	.	۳
۰/۰۰۴	۰/۲۵۱	.	۰/۰۰۵	۰/۲۴۴	.	۰/۰۰۵	۰/۲۲۷	.	۴
۰/۰۰۴	۰/۲۶۸	.	۰/۰۰۵	۰/۲۶	.	۰/۰۰۵	۰/۲۴۲	.	۵
۰/۰۰۴	۰/۲۶۸	.	۰/۰۰۵	۰/۲۶۲	.	۰/۰۰۵	۰/۲۴۴	.	۶
۰/۰۰۲	۰/۱۶	.	۰/۰۰۳	۰/۱۵۹	.	۰/۰۰۲	۰/۱۴۹	.	۲
۰/۰۰۳	۰/۲۶۹	.	۰/۰۰۴	۰/۲۶۷	.	۰/۰۰۳	۰/۲۴۹	.	۳
۰/۰۰۳	۰/۳۳۷	.	۰/۰۰۵	۰/۳۳۴	.	۰/۰۰۴	۰/۳۱	.	۴
۰/۰۰۴	۰/۳۷۷	.		۰/۳۷۳	.	۰/۰۰۴	۰/۳۴۳	.	۵
۰/۰۰۳	۰/۳۹۸	.	۰/۰۰۵	۰/۳۹۴	.	۰/۰۰۴	۰/۳۵۸	.	۶
۰/۰۰۲	۰/۱۸	.	۰/۰۰۳	۰/۱۷۸	.	۰/۰۰۳	۰/۱۶۹	.	۲
۰/۰۰۳	۰/۳۰۴	.	۰/۰۰۴	۰/۳۰۲	.	۰/۰۰۴	۰/۲۸۴	.	۳
۰/۰۰۳	۰/۳۸۸	.	۰/۰۰۵	۰/۳۸۴	.	۰/۰۰۵	۰/۳۶	.	۴
۰/۰۰۴	۰/۴۴	.	۰/۰۰۶	۰/۴۳۷	.	۰/۰۰۵	۰/۴۰۷	.	۵
۰/۰۰۳	۰/۴۷۶	.	۰/۰۰۵	۰/۴۷	.	۰/۰۰۵	۰/۴۳۵	.	۶
۰/۰۰۲	۰/۱۸۹	.	۰/۰۰۲	۰/۱۹۲	.	۰/۰۰۳	۰/۱۸۲	.	۲
۰/۰۰۳	۰/۳۱۹	.	۰/۰۰۴	۰/۳۲۶	.	۰/۰۰۵	۰/۳۰۸	.	۳
۰/۰۰۳	۰/۴۰۸	.	۰/۰۰۵	۰/۴۱۷	.	۰/۰۰۶	۰/۳۹۲	.	۴
۰/۰۰۳	۰/۴۶۸	.	۰/۰۰۵	۰/۴۷۸	.	۰/۰۰۶	۰/۴۴۷	.	۵
۰/۰۰۳	۰/۵۰۷	.	۰/۰۰۵	۰/۵۱۸	.	۰/۰۰۶	۰/۴۸۳	.	۶

بودن سری قابل رد نبوده و سری از فرآیند اخلال خالص تبعیت می‌کند که نشانگر قابلیت پیش‌بینی‌پذیر بودن سری زمانی می‌باشد. از آنجا که در حالت معمول جهت شناسایی الگو از رسم نمودارهای خودهمبستگی و خودهمبستگی جزئی استفاده می‌شود و بسیاری از

نتایج این آزمون در گام‌های زمانی مختلف برای هر سه ایستگاه در جدول ۳ ارائه شده است و بر این اساس به ازای بعدهای محاط ۲ تا ۶ در همه‌ی گام‌های زمانی شاخص SPEI کلیه مقادیر BDS کوچکتر از ۱/۹۶ بوده و از این رو، فرضیه صفر اخلال خالص

بودن معیار شوارتز می‌باشد، نسبت به ایجاد مدل تلفیقی ARMA-GA اقدام گردید. نتایج بهینه‌سازی برای این تابع هدف برای مقیاس‌های زمانی مختلف شاخص SPEI و رابطه برازش شده بر داده‌ها در جدول ۴ ارائه شده است.

سری‌های زمانی دارای ترکیبات پیچیده‌ای از وقفه‌های AR و MA می‌باشند که در این صورت با ترسیم ACF و PACF نمی‌توان به صورت دقیق وقفه‌های بهینه مدل را تشخیص داد، لذا با استفاده از کدنویسی در محیط نرم افزار MATLAB بر اساس تابع هدف مدل که کمینه

جدول ۴- نتایج مدل تلفیقی ARMA-GA در پیش بینی شاخص SPEI.

مدل ریاضی بسط داده شده	مدل برتر	بهترین کروموزوم	مقیاس زمانی (ماه)	ایستگاه
$SPEI_t = 0.00039 - 0.01976 SPEI_{t-3} + \epsilon_t$	ARIMA([3],1,[0])	۰۰۱۰۰۰۰۰۰۰	۱	تبریز
$SPEI_t = 0.00085 + 0.2534 SPEI_{t-1} - 0.921 SPEI_{t-2} - 0.147 SPEI_{t-5} + 0.90007 \epsilon_{t-2} + \epsilon_t$	ARIMA([1,2,5],1,[2])	۱۱۰۰۱۰۱۰۰۰	۳	
$SPEI_t = -0.0005 + 0.122 SPEI_{t-1} - 0.188 \epsilon_{t-3} + \epsilon_t$	ARIMA([1],0,[3])	۱۰۰۰۰۰۰۱۰۰	۶	
$SPEI_t = -0.0001 - 0.227 SPEI_{t-3} + 0.7379 SPEI_{t-5} - 0.7017 \epsilon_{t-5} + \epsilon_t$	ARIMA([3,5],1,[5])	۰۰۱۰۱۰۰۰۰۱	۱۲	
$SPEI_t = -0.00192 + 0.1589 \epsilon_{t-3} + \epsilon_t$	ARIMA([0],1,[1])	۰۰۰۰۰۱۰۰۰۰	۲۴	
$SPEI_t = -0.0019 + 0.1981 \epsilon_{t-3} + \epsilon_t$	ARIMA([0],1,[1])	۰۰۰۰۰۱۰۰۰۰	۴۸	
$SPEI_t = 0.0006 - 0.054 SPEI_{t-5} - 0.009 \epsilon_{t-2} - 0.121 \epsilon_{t-3} - 0.108 \epsilon_{t-4} + \epsilon_t$	ARIMA([5],1,[2,3,4])	۰۰۰۰۱۰۱۱۱۰	۱	سقز
$SPEI_t = 0.0001 - 0.132 \epsilon_{t-1} - 0.0806 \epsilon_{t-2} - 0.633 \epsilon_{t-3} + \epsilon_t$	ARIMA([0],1,[1,2,3])	۰۰۰۰۰۱۱۱۱۰	۳	
$SPEI_t = 0.00008 + 0.0297 SPEI_{t-3} - 0.628 SPEI_{t-4} - 0.234 \epsilon_{t-2} - 0.716 \epsilon_{t-4} + \epsilon_t$	ARIMA([3,4],1,[2,4])	۰۰۱۱۰۰۱۰۱۰	۶	
$SPEI_t = 0.0001 + 0.044 SPEI_{t-1} + 0.716 SPEI_{t-4} - 0.961 \epsilon_{t-4} - 0.0218 \epsilon_{t-5} + \epsilon_t$	ARIMA([1,4],1,[4,5])	۱۰۰۱۰۰۰۰۱۱	۱۲	
$SPEI_t = -0.0004 + 0.0808 SPEI_{t-5} + \epsilon_t$	ARIMA([5],1,[0])	۰۰۰۰۱۰۰۰۰۰	۲۴	
$SPEI_t = 0.0006 + 0.002 \epsilon_{t-3} - 0.038 \epsilon_{t-4} + 0.052 \epsilon_{t-5} + \epsilon_t$	ARIMA([0],1,[3,4,5])	۰۰۰۰۰۰۱۱۱	۴۸	
$SPEI_t = 0.001 + 0.007 SPEI_{t-3} - 0.032 SPEI_{t-4} - 0.02 SPEI_{t-5} + \epsilon_t$	ARIMA([3,4,5],1,[0])	۰۰۱۱۱۰۰۰۰۰	۱	ارومیه
$SPEI_t = 0.0007 - 0.544 \epsilon_{t-3} + \epsilon_t$	ARIMA([0],1,[3])	۰۰۰۰۰۰۰۱۰۰	۳	
$SPEI_t = 0.002 - 0.009 SPEI_{t-2} + 0.093 SPEI_{t-3} - 0.09 SPEI_{t-4} - 3 - 0.648 SPEI_{t-5} + 0.219 \epsilon_{t-1} + 0.738 \epsilon_{t-5} + \epsilon$	ARIMA([2,3,4,5],1,[1,5])	۰۱۱۱۱۱۰۰۰۱	۶	
$SPEI_t = -0.001 + 0.059 SPEI_{t-1} + 0.0228 SPEI_{t-5} - 0.077 \epsilon_{t-4} + \epsilon_t$	ARIMA([1,5],1,[4])	۱۰۰۰۱۰۰۰۱۰	۱۲	
$SPEI_t = -0.004 + 0.0359 SPEI_{t-5} + \epsilon_t$	ARIMA([5],1,[0])	۰۰۰۰۱۰۰۰۰	۲۴	
$SPEI_t = -0.008 + 0.0635 SPEI_{t-3} + 0.152 \epsilon_{t-1} + \epsilon_t$	ARIMA([3],1,[1])	۰۰۱۰۰۱۰۰۰۰	۴۸	

پیش‌بینی در مرحله آزمون در جدول ۵ ارائه شده است که بر اساس آماره‌های ارزیابی با افزایش مقیاس زمانی شاخص SPEI دقت پیش‌بینی مدل افزایش پیدا می‌کند، به نحوی که در مرحله آموزش ضریب همبستگی در SPEI1 تا SPEI48 از ۰/۱۹۵ به ۰/۹۶۴ در ایستگاه تبریز و از ۰/۲ به ۰/۹۸۳ در ایستگاه سقز و از ۰/۲۱۶ به ۰/۹۸۳ در ایستگاه ارومیه رسیده است که نشان از دقت

یکی از مزیت‌های روش تلفیقی ARMA-GA به دست آوردن مناسب‌ترین وقفه‌های مدل ARMA با استفاده از رویکرد بهینه‌سازی بر مبنای روش الگوریتم زنتیک می‌باشد که سبب تسریع در دستیابی به مدل بهینه می‌گردد. نتایج روش تلفیقی نشان داد که جهت ایستا نمودن داده‌ها باید از اولین تفاضل آنها در همه مقیاس‌های زمانی شاخص SPEI استفاده شود. نتایج

مناسب برای پیش‌بینی آن در مقیاس‌های کوتاه مدت را می‌توان در این مورد ارزیابی کرد. به منظور بررسی میزان قابلیت اطمینان به مدل پیش‌بینی داده‌ها از آماره Ljung-Box استفاده شد که مطابق جدول ۵ مقادیر p-value آن بزرگتر از ۰/۰۵ می‌باشد که این نتیجه نشان می‌دهد باقی‌مانده مدل به صورت تصادفی می‌باشد.

مناسب روش تلفیقی در مقیاس‌های زمانی بالای شاخص SPEI می‌باشد. این نتایج با نتایج تحقیق حسینی موعاری و عراقی‌نژاد (۲۰۱۶) و میشرا و دسای (۲۰۰۶) برای پیش‌بینی شاخص SPI مطابقت دارد. با توجه به اینکه شاخص SPEI در مقیاس زمانی کوتاه مدت تغییرات و پراکندگی بسیاری دارد، دلیل عدم نتایج

جدول ۵- آماره‌های ارزیابی مدل ARMA-GA در مقیاس‌های زمانی مختلف در ایستگاه‌های مورد مطالعه.

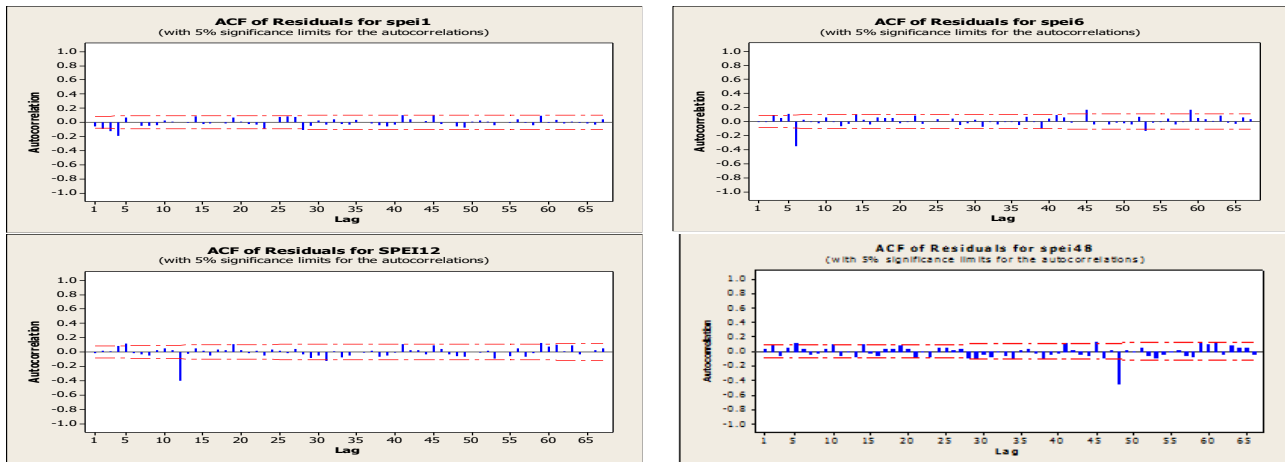
ایستگاه	مقیاس زمانی (ماه)	r	RMSE	MAE	Ljung-Box
سقز	۱	۰/۲	۰/۳۲۶	۰/۲۵۷	۰/۷۶۴
	۳	۰/۷۰۴	۰/۲۴۱	۰/۱۷۹	۰/۹۸۱
	۶	۰/۸	۰/۱۴۳	۰/۱۰۲	۰/۴۲۸
	۱۲	۰/۹۴۳	۰/۱۲۶	۰/۰۹	۰/۲
	۲۴	۰/۹۷۹	۰/۰۸۱	۰/۰۵۵	۰/۳۸۹
	۴۸	۰/۹۸۳	۰/۰۶۵	۰/۰۴۴	۰/۸۶۸
تبریز	۱	۰/۱۹۵	۰/۴۵۸	۰/۳۶۹	۰/۰۹
	۳	۰/۷۱۸	۰/۲۵۶	۰/۱۹۲	۰/۱۱
	۶	۰/۸۲۸	۰/۱۸۳	۰/۱۳۸	۰/۱۳
	۱۲	۰/۹۱	۰/۱۲۱	۰/۰۹۲	۰/۱۸
	۲۴	۰/۹۵۱	۰/۰۷	۰/۵۲	۰/۵۶۸
	۴۸	۰/۹۶۴	۰/۰۴۳	۰/۰۳	۰/۸۳۴
ارومیه	۱	۰/۲۱۶	۰/۳۹۶	۰/۲۵۸	۰/۷۴۶
	۳	۰/۶۷۲	۰/۳۲۹	۰/۱۸۸	۰/۱۲۹
	۶	۰/۸۹۲	۰/۲۰۲	۰/۱۴۱	۰/۷۰۳
	۱۲	۰/۹۱۴	۰/۱۷۹	۰/۱۲۲	۰/۸۲۵
	۲۴	۰/۹۳۲	۰/۱۵۱	۰/۰۷۷	۰/۵۰۸
	۴۸	۰/۹۸۳	۰/۰۴	۰/۰۲۷	۰/۳۳۴

تعداد Lag Time های خارج از محدوده صفر کمتر از ۵ درصد کل تعداد Lag Time ها می‌باشد که نشان از قابل اطمینان بودن مدل پیش‌بینی انجام شده می‌باشد. همچنین بر اساس این نمودارها مشاهده می‌گردد که باقی مانده‌ها هیچ‌گونه همبستگی با یکدیگر ندارند که نشان دهنده برازش خوب مدل تلفیقی بر روی داده‌ها و تصادفی بودن باقی مانده‌ها می‌باشد. این نتایج با تحقیق رضیئی (۲۰۱۷) و معروفی و همکاران (۲۰۱۴) که با استفاده از روش سری زمانی ARMA و SARIMA و

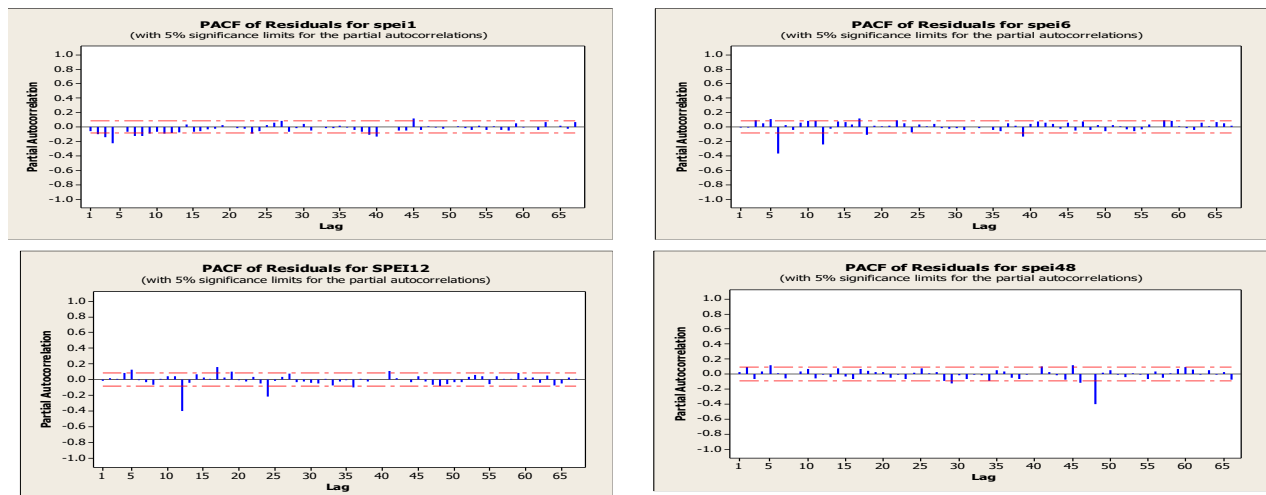
یکی از ملاک‌های تشخیص مناسب بودن مدل، مستقل بودن باقی‌مانده‌های حاصل از برازش آن بر روی سری زمانی داده‌های مشاهداتی می‌باشد که جهت بررسی این موضوع از نمودار خودهمبستگی باقی مانده‌های حاصل از برازش مدل استفاده می‌گردد. نمودارهای خودهمبستگی و خودهمبستگی جزئی مانده‌ها در مقیاس‌های زمانی ۱، ۶، ۱۲ و ۴۸ ماهه برای نمونه برای ایستگاه تبریز در شکل‌های ۲ و ۳ نشان داده شده است که بر اساس آن مشاهده می‌شود که

زنجیره مارکف نسبت به پیش‌بینی شاخص خشکسالی SPI در مقیاس‌های زمانی مختلف اقدام نمودند، مطابقت

دارد.



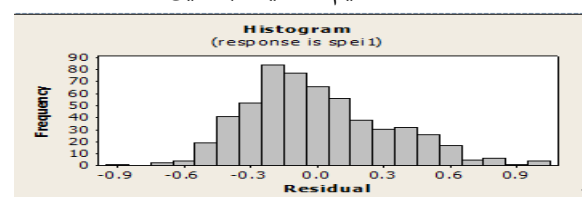
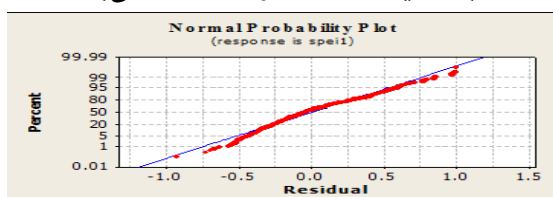
شکل ۲- نمودارهای خود همبستگی با احتمال ۹۵ درصد از مانده‌های شاخص SPEI در مقیاس‌های زمانی ۱، ۶، ۱۲، ۴۸ ماهه در ایستگاه تبریز.

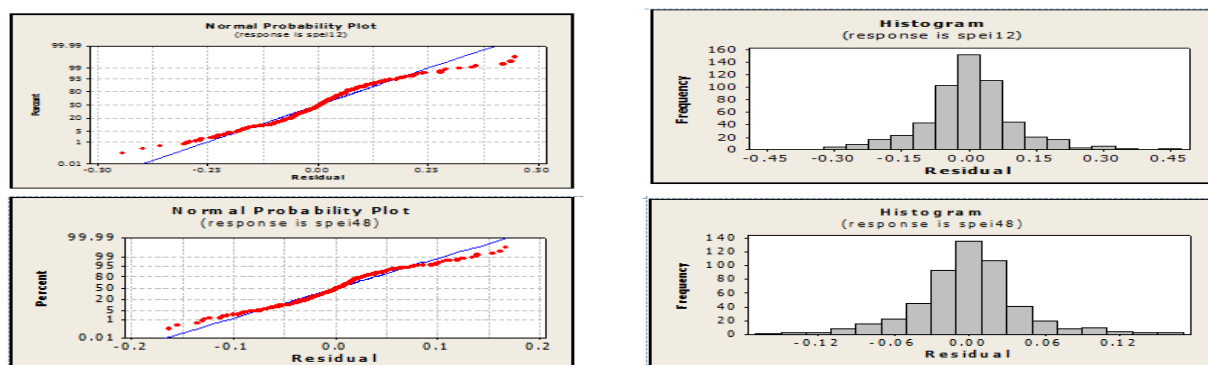


شکل ۳- نمودارهای خود همبستگی جزئی با احتمال ۹۵ درصد از مانده‌های شاخص SPEI در مقیاس‌های زمانی ۱، ۶، ۱۲، ۴۸ ماهه در ایستگاه تبریز.

تخطی خاصی از نرمال بودن و تصادفی بودن ندارند و باقی‌مانده‌ها کاملاً تصادفی توزیع شده و دارای توزیع نرمال می‌باشند که نشان از عملکرد بهینه مدل ARMA-GA برای ایستگاه‌های مورد مطالعه می‌باشد.

به منظور ارزیابی عملکرد مدل و درستی و مناسب بودن مدل، نمودارهای احتمال نرمال مانده‌ها و هیستوگرام مانده‌ها برای مقیاس‌های زمانی ۱، ۱۲ و ۴۸ ماهه در شکل ۴ ترسیم گردید. بر این اساس مانده‌ها





شکل ۴- نمودارهای احتمال نرمال مانده‌های شاخص SPEI در مقیاس‌های زمانی ۱، ۳، ۶، ۱۲ و ۴۸ ماهه در ایستگاه ارومیه.

نتیجه‌گیری کلی

هدف از این تحقیق، ارائه مدلی جدید جهت تخمین پارامترهای سری زمانی ARMA با استفاده از رویکرد بهینه‌سازی و پیش‌بینی خشکسالی بر پایه شاخص SPEI در مقیاس‌های زمانی ۱، ۳، ۶، ۱۲، ۲۴ و ۴۸ ماهه بود که برای این منظور، از داده‌های دما و بارش ایستگاه‌های منتخب حوضه دریاچه ارومیه در ۳ بخش شرقی، غربی و جنوبی دریاچه شامل ایستگاه‌های تبریز، ارومیه و سقز استفاده گردید. جهت تخمین پارامترهای سری زمانی خطی ARMA از ترکیب آن با روش الگوریتم ژنتیک استفاده گردید. بر این اساس نتایج زیر از این تحقیق بدست آمد:

۱- بر اساس آزمون BDS داده‌ها در مقیاس‌های زمانی مختلف شاخص SPEI قابلیت پیش‌بینی دارند و مدل برتر انتخابی بر اساس روش ترکیبی ARMA-GA در مقیاس‌های زمانی مختلف انتخاب گردید.

۲- پیش‌بینی SPEI1 با دقت بسیار پایینی انجام می‌گیرد و با افزایش مقیاس محاسبه SPEI دقت پیش‌بینی مدل ARMA-GA افزایش می‌یابد به نحوی که در SPEI1 عملکرد نامناسب و در SPEI48 بهترین عملکرد را داشته است. همچنین بر اساس آماره‌های ارزیابی از مقیاس ۱۲ ماهه تا ۴۸ ماهه دقت مدل پیش‌بینی بهبود مناسبی پیدا کرده است.

۳- با توجه به احتمال آماره Ljung-Box در مقیاس‌های زمانی مختلف که بیش‌تر از ۰/۰۵ می‌باشد و همچنین بررسی نرمال بودن توزیع باقی‌مانده‌های مدل ARMA-GA می‌توان قابل اطمینان بودن پیش‌بینی را در هر سه ایستگاه استنباط نمود.

۴- پیشنهاد می‌گردد در مطالعات آتی، ترکیب سایر روش‌های بهینه‌سازی با سری زمانی جهت تخمین پارامترهای مدل سری زمانی نیز مورد بررسی قرار گیرد.

منابع مورد استفاده

- Abo-Hammour ZES, Alsmadi OM, Al-Smadi AM, Zaqout MI and Saraireh MS, 2012. ARMA model order and parameter estimation using genetic algorithms. *Mathematical and Computer Modelling of Dynamical Systems* 18(2): 201-221.
- Allen RG, Pereira LS, Raes D and Smith M, 1998. *Crop Evapotranspiration-Guidelines for Computing Crop Water Requirements-FAO Irrigation and drainage paper 56* FAO, Rome 300:D05109.
- Ayalew S, Babu MC and Rao LKM, 2012. Comparison of new approach criteria for estimating the order of autoregressive process. *IOSR Journal of Mathematics* 1(3): 10-20.
- Beligiannis GN, Demiris EN and Likothanassis SD, 2000. Self-adaptive evolution strategies for ARMA model identification. In *Signal Processing Conference, 2000 10th European* (pp. 1-4). IEEE.
- Box GE and Jenkins GM, 1976. *Time series analysis: forecasting and control*, revised ed. Holden-Day,
- Brock WE, Brock WA, Hsieh DA and LeBaron BD, 1991. *Nonlinear dynamics, chaos, and instability: statistical theory and economic evidence*. MIT press.
- Brock W.A, Scheinkman J.A, Dechert WD and LeBaron B, 1996. A test for independence based on the

- correlation dimension. *Econometric Reviews* 15: 197-235.
- Durdu O.F, 2010. Application of linear stochastic models for drought forecasting in the Buyuk Menderes river basin, western Turkey. *Stochastic Environmental Research and Risk Assessment* 24(8): 1145-1162.
- Ervural BC, Beyca OF and Zaim S, 2016. Model estimation of ARMA using genetic algorithms: A case study of forecasting natural gas consumption. *Procedia-Social and Behavioral Sciences* 235: 537-545.
- Hassanzadeh Y, Abdi Kordani A, Fakheri Fard A. 2012. Application of meta-heuristic methods in drought monitoring (Case study: Tabriz station). *Water and Soil Science- University of Tabriz*, 22(3): 29-46. (In Persian)
- Holland J.H, 1962. Outline for a logical theory of adaptive systems. *Journal of the ACM (JACM)* 9(3): 297-314.
- Hosseini-Moghari SM and Araghinejad S, 2016. Application of Statistical, Fuzzy and Perceptron Neural Networks in Drought Forecasting (Case Study: Gonbad-e Kavous Station). *Water and Soil (Agricultural Sciences and Technology)* 30(1): 247-259. (In Persian)
- Kempes C, Myers O, Breshears D and Ebersole J, 2008. Comparing response of *Pinus edulis* tree-ring growth to five alternate moisture indices using historic meteorological data. *Journal of Arid Environments* 72(4):350-357.
- Khalili K, Nazeri Tahrudi M, Abbaszadeh Afshar M and Nazeri Tahrudi Z, 2014. Modeling monthly mean air temperature using SAMS2007 (Case Study: Urmia synoptic station). *Journal of Middle East Applied Science and Technology (JMEAST)* 15: 578-583.
- Kim M and Kim J, 2017. GA-ARMA model for predicting IGS RTS corrections. *International Journal of Aerospace Engineering*, 2017: 1-7.
- Maerufi S, Khatar B, Sadeghifar M, Parsafar N and Eildurmi A, 2014. Drought prediction using SARIMA time series and SPI index in the central region of Hamedan province. *Water Research in Agriculture* 28(1): 213-225. (In Persian)
- Maerufi S, Saghaei S, Ershadfath F and Khatar B, 2015. Evaluating time series models to estimate monthly temperature of Iran's old synoptic stations during 1977-2005. *Water and Soil Science - University of Tabriz*, 24(4): 215-226. (In Persian)
- Michalewicz Z, 1999. *Genetic Algorithms + Data Structures = Evolution Programs*, Department of Computer Science University of North Carolina, USA. Springer Publishing 388p.
- Mishra A, Desai V, 2005. Drought forecasting using stochastic models. *Stochastic Environmental Research and Risk Assessment* 19(5): 326-339.
- Mishra A and Desai V, 2006. Drought forecasting using feed-forward recursive neural network. *Ecological Modelling* 198(1-2): 127-138.
- Mishra A, Desai V and Singh V, 2007. Drought forecasting using a hybrid stochastic and neural network model. *Journal of Hydrologic Engineering* 12(6): 626-638.
- Peng P and Chen Q, 2003. Improved genetic algorithm and application to ARMA modelling Pp. 7-7. In *SICE Annual Conference Program and Abstracts SICE Annual Conference 2003*. The Society of Instrument and Control Engineers. Fukui, Japan.
- Raziei T, 2017. Drought forecasting in eastern and central arid and semi-arid regions of Iran using time series and Markov chain models. *Watershed Engineering and Management* 8(4): 454-477. (In Persian)
- Salas JD, 1993. Analysis and modelling of hydrological time series Pp. 19.1-19.72. In: Maidment, D.R., Ed., *Handbook of Hydrology*, McGraw-Hill, New York.
- Salas JD, Delleur JW, Yevjevich V and Lane WL, 1980. *Applied Modeling of Hydrologic Time Series*. Water Resource Publications, P. O. Box 2841. Littleton, Colorado .80161, U.S.A. 484 P.
- Sharifan H. and Ghahraman B, 2007. Evaluation of rainfall forecasting in Golestan province using time series. *Agricultural Sciences and Natural Resources* 14(3): 196-209 (In Persian).
- Thornthwaite C.W, 1948. An approach toward a rational classification of climate. *Geographical Review* 38(1): 55-94.
- Toufani P, Mosaedi A and Fakheri Fard A, 2011. Prediction of precipitation applying wavelet network model (case study: Zarringol station, Golestan province, Iran). *Water and Soil (Agricultural Sciences and Technology)* 25(5): 1217-1226. (In Persian)

- Vicente-Serrano SM, Begueria S and Lopez-Moreno J.I, 2010. A multiscalar drought index sensitive to global warming: the standardized precipitation evapotranspiration index. *Journal of Climate* 23(7): 1696-1718.
- Zarei AR, Moghimi MM and Mahmodi M.R, 2016. Modeling and prediction of seasonal drought, using RDI index and time series models (Case study: Tehran synoptic station). *Journal of Desert Ecosystem Engineering*. 5(11): 105-116. (In Persian)