

مدل سازی درجه حرارت ماهیانه، مطالعه موردی: ایستگاه هیدرومتری شاهمختار

محمد رضا میرزایی^۱، استادیار، دانشکده کشاورزی، دانشگاه یاسوج

تاریخ دریافت: ۱۳۹۱/۰۶/۲۰

تاریخ پذیرش: ۱۳۹۱/۱۱/۳۰

چکیده

در این مطالعه، میانگین درجه حرارت ماهانه در ایستگاه هیدرومتری شاهمختار در استان کهگیلویه و بویراحمد برای یک دوره ۳۹ ساله بین سال‌های ۱۳۴۹ تا ۱۳۸۸ مورد بررسی قرار گرفت. جهت مدل‌سازی درجه حرارت در این ایستگاه از تکنیک تحلیل کاوشگرانه داده‌ها (EDA) استفاده شد. روش‌های گرافیکی متعددی، مانند استفاده از نمودارهای تناوب‌نگار، خود هم‌بستگی و خود هم‌بستگی جزئی برای بررسی داده‌ها از جمله وجود چرخه (S) و مناسب بودن استفاده از تحلیل سری‌های زمانی جهت مدل‌سازی باقی‌مانده‌ها (R) تعیین شد. مدل‌سازی چرخه موجود در داده‌ها با استفاده از مدل سینوسی و با تعیین دقیق پارامترهای آن با استفاده از تکنیک‌های EDA انجام شد. باقی‌مانده‌ها نیز با استفاده از Modelexpert در نرم‌افزار SPSS 19 جهت تحلیل سری‌های زمانی مورد بررسی قرار گرفت. بهترین برازش با توجه به کمترین مقدار برای معیار NBIC، مدل ARIMA(0,1,2) تعیین شد. جهت تایید مدل، بررسی تصادفی بودن مقادیر خطاهای مدل (E) و نرمال بودن توزیع آن‌ها در سطح پنج درصد به ترتیب با استفاده از آزمون Ljung-Box (با مقدار آماره برابر ۲۴/۳) و Kolmogorov-Smirnov (با مقدار آماره برابر ۰/۴۰۸۸) بررسی و مورد تایید قرار گرفت.

واژه‌های کلیدی: خود هم‌بستگی، دما، سری‌های زمانی، EDA، ARIMA

مقدمه

نوسانات اقلیمی از جمله عواملی هستند که بر چرخه طبیعی آب در اکوسیستم تاثیر می‌گذارند. در دو دهه اخیر تغییرات حاصله در اقلیم جهانی از جمله موضوعات مهم مطرح در بین کارشناسان منابع آب در سطح دنیا بوده است. پیش‌بینی‌های بارش و دما به‌طور کارآمدی در تصمیم‌گیری و استفاده بهینه از منابع آب می‌تواند به کار گرفته شود. دما به‌عنوان یکی از پارامترهای اصلی هواشناسی تاثیر مستقیمی بر روی پدیده‌هایی مانند تبخیر و تعرق که پدید آورنده ابرها و بارش می‌باشد و در نتیجه چرخه هیدرولوژی است، دارد. زیرا دما نمایه‌ای از شدت گرما و انرژی تابشی مستقیم و غیرمستقیم نور خورشید است که موتور چرخاننده چرخه هیدرولوژی می‌باشد. در واقع درجه حرارت به‌عنوان عمده‌ترین عامل در تعیین و پراکندگی سایر عوامل اقلیمی شناخته شده است که بر روی مصارف آب، چرخه تولید محصولات زراعی و نیاز آبی گیاهان تاثیر تعیین کننده‌ای دارد. مطالعه و شناخت وضعیت کنونی و تغییرات احتمالی پارامترهایی از جمله دما در دراز مدت اهمیت زیادی در بررسی و پیش‌بینی تغییرات اقلیمی دارد و در زمینه‌هایی از قبیل بررسی تغییرات دبی رودخانه (Li و همکاران، ۲۰۱۱؛ Fan و همکاران، ۲۰۱۱)، بارش (Tularam و Ilahee، ۲۰۱۰)، مصارف آب (حمادی و همکاران، ۱۳۸۵) و ... کارهای متعددی انجام شده است.

در این راستا تجزیه و تحلیل پدیده‌ها با استفاده از مدل‌های آماری در توصیف، کنترل و پیش‌بینی فرآیندها یاریگر است. مطالعه این فرآیندها جهت درک یا به مدل در آوردن مکانیسم تصادفی آن‌ها و همچنین، پیش‌بینی وضعیت آینده به کمک گذشته‌ی آن‌ها می‌باشد. یکی از روش‌های بسیار معمول که به کمک آن می‌توان سیر تحولات دما و

^۱ نویسنده مسئول: mmirzaei@mail.yu.ac.ir

به‌طور کلی متغیرهای اقلیمی را در گذشته و حال بررسی کرد آنالیز سری‌های زمانی در مقیاس‌های مختلف زمانی می‌باشد (Jha و Machiwal، ۲۰۱۲). یک سری زمانی مجموعه‌ای از مقادیر عددی یک متغیر است که با زمان تغییر می‌کند. به‌طور سنتی مدل‌های $ARMA^1$ به‌صورت گسترده‌ای در مدل‌سازی سری‌های زمانی منابع آب مورد استفاده قرار گرفته است. استفاده از این مدل‌ها نسبت به مدل‌های مفهومی کاربردی‌تر است، زیرا در این مدل‌سازی نیاز به دانستن ساختار داخلی فرآیندهای فیزیکی مورد نظر نیست (Rao و همکاران، ۲۰۱۱).

در کار با سری‌های زمانی طبیعی، شامل سری‌های زمانی هیدرولوژیکی، اقلیمی و محیطی، برخورد با داده‌هایی که فرضیاتی همچون نرمال بودن، یکسانی، تصادفی بودن، بدون پیوند یا چرخه‌ای بودن و خصوصیات ایستایی در مورد آن‌ها صدق کند، یک استثنا است (Rao و همکاران، ۲۰۰۳) و عملاً در طبیعت برخورد با چنین داده‌هایی به‌ندرت پیش می‌آید. به‌همین‌علت، همیشه بایستی در تمامی مطالعات منابع آب شامل استفاده از داده‌های سری زمانی هیدرولوژیکی، تجزیه و تحلیل آماری مقدماتی صورت گیرد (Montaseri و Adeloey، ۲۰۰۲)، در غیر این‌صورت مدل‌سازی انجام شده ممکن است نتایج اشتباهی را ارائه دهد. مرور مطالعات گذشته به‌صورت فشرده توسط Jha و Machiwal (۲۰۰۶) نشان داد که تاکنون در هیچ تحقیقی موارد بالا به‌درستی مورد بررسی قرار نگرفته است. در مطالعات گذشته تنها به بررسی و تحلیل روند خطی موجود در داده‌ها پرداخته شده و از بقیه موارد آن صرف نظر شده است. جدول ۱ مروری بر برخی از مطالعات صورت گرفته با تاکید بر منابع منتشر شده در داخل ایران می‌باشد. همان‌طور که مشاهده می‌شود، تقریباً در اغلب قریب به اتفاق موارد، آزمون فرضیات صورت نگرفته است و یا تنها یک بخش آن یعنی وجود و یا عدم وجود روند مورد بررسی قرار گرفته است.

اکثر محققین هدف از مطالعه خود را بررسی منطقه‌ای روند دما ذکر کرده‌اند و در این بین از تکنیک‌های مختلفی سود جست‌ه‌اند؛ که اغلب آنها بیان می‌دارند که دما روندی افزایشی داشته است (براتی و طرقي، ۱۳۷۸؛ حمادی و همکاران، ۱۳۸۵؛ سعیدآبادی و همکاران، ۱۳۸۲؛ کاویانی و عساکره، ۱۳۸۰). گاهی بین نتایج محققین تعارض دیده می‌شود، به‌طور مثال نتایج سه تحقیق انجام شده در منطقه مشهد در جدول ۱ قابل مقایسه است. براتی و طرقي (۱۳۷۸) بیان کرده‌اند که روند افزایشی وجود دارد و سال‌های ۱۹۵۱-۱۹۴۴ را به سه دوره تقسیم کرده‌اند که در دوره اول روند افزایشی، در دوره دوم روند کاهشی و از سال ۱۹۷۴ به بعد یک روند افزایشی با شدت بیشتر وجود دارد. ابراهیمی و همکاران (۱۳۸۴) با مطالعه میانگین‌های ماهانه دما بیان داشته‌اند که در اغلب ماه‌ها درجه حرارت روند افزایشی داشته و بعضی ماه‌ها روند کاهشی بوده است و روند کلی را افزایشی تشخیص داده‌اند. بر خلاف دو مورد پیشین، غیور و عساکره (۱۳۸۴) با بررسی صد سال آمار بیان داشته‌اند که روند خاص معنی‌داری برای درجه حرارت وجود ندارد. منشا این تعارض‌ها شاید تعداد سال آماری و یا تکنیک‌های مورد استفاده آن‌ها جهت ارزیابی تغییرات دما باشد (جدول ۱). برخی محققین نیز با مدل‌سازی درجه حرارت، پیش‌بینی درجه حرارت را انجام داده‌اند که برخی از این مدل‌ها کارایی لازم برای پیش‌بینی درجه حرارت در طولانی مدت را ندارد (صراف و جامعی، ۱۳۸۱). در این پژوهش سعی شد، بررسی میانگین درجه حرارت ماهانه ایستگاه شاه‌مختار در حوزه آبخیز رودخانه بشار، با استفاده از یک روش مدل‌سازی گام به گام و با تحلیل کامل فرضیات آماری ارائه شود.

مواد و روش‌ها

ایستگاه هیدرومتری شاه‌مختار با مختصات جغرافیایی، به‌طول $31^{\circ} 20'$ و به‌عرض $31^{\circ} 40' 02''$ در سال ۱۳۴۹ بر روی رودخانه بشار در ارتفاع ۱۷۴۰ متری از سطح دریا تاسیس شده است که با بارندگی متوسط سالیانه ۸۵۰ میلی‌متر بخشی از حوزه آبخیز کارون به وسعت ۱۱۸۷ کیلومتر مربع می‌باشد (شکل ۱). آمار توصیفی آورده شده در جدول ۲ نشان می‌دهد که میانگین درجه حرارت ماهانه در طول دوره ۳۹ ساله این ایستگاه ۱۴/۲۳ سانتی‌گراد بوده و دامنه تغییرات درجه حرارت بین حداقل ۴- و حداکثر ۲۷/۴ با انحراف معیار ۸/۰۳ سانتی‌گراد می‌باشد.

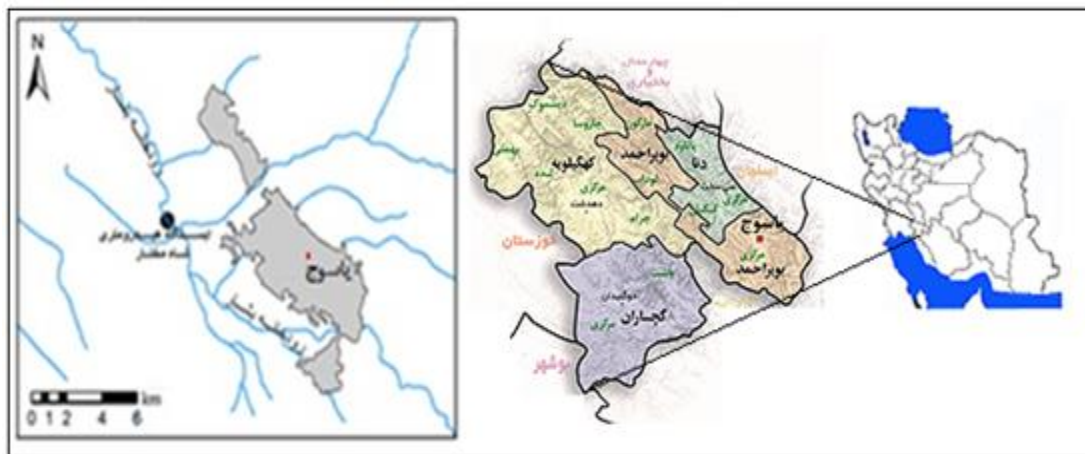
¹ AutoRegressive Moving Average

جدول ۱- مروری بر برخی تحقیقات انجام شده با تاکید بر تحقیقات داخلی

| محقق | منطقه مورد مطالعه | تعداد ایستگاه | تعداد سال آزمون | فرصیات | پارامترهای اقلیمی | هدف | تکیک | تغییرات روند |
|---------------------------|----------------------|---------------|-----------------|----------|---|------------------------------------|---|--|
| مسیبی و غیور (۱۳۷۴) | گلپایگان | ۷ | ۲۵ | خیز | دما و پایدگی | اقلیم‌شناسی | برخی از آمار توصیفی | اشاره نشده |
| مقصودی (۱۳۷۷) | سواحل شمال خلیج فارس | ۴ | حدود ۳۷ | خیز | متوسط دمای سالانه | تغییرات دما | برازش خطی، میانگین متحرک پنج‌ساله، نیمه متوسط‌ها | دو ایستگاه کاهشی دو ایستگاه افزایشی |
| براتی و طرقي (۱۳۷۸) | مشهد | ۱ | ۴۴ | خیز | دما و پایدگی | تغییرات روند بارش و دما | برازش خطی، میانگین متحرک میانگین متحرک پنج‌ساله، ضریب همبستگی و آمار توصیفی | دما و بارش هر دو افزایشی |
| مقیبی (۱۳۷۸) | تهران | ۱ | ۳۰ | خیز | رطوبت نسبی جازدگی | رابطه رطوبت نسبی با بارش | مقایسه اختلاف میانگین دو نیمه آماري، شیب خط، هموارسازی | ----- |
| کاویانی و عساکره (۱۳۸۰) | جاسک | ۱ | ۱۰۴ | خیز | دما | تغییرات دما | آمار توصیفی، SARIMA | افزایشی |
| صراف و جامعی (۱۳۸۱) | سندج، میاندوآب | ۲ | ۱۹ | کم | دما | مدل‌سازی و پیش‌بینی | آمار توصیفی، SARIMA و مدل‌های سری زمانی، ARIMA | بدون روند مشخص، افزایشی |
| سیدآبادی و همکاران (۱۳۸۲) | تبریز | ۱ | ۵۰ | خیز | پایدگی-دما (ماهانه و فصلی) | مدل‌سازی تغییرات دما | SARIMA به‌صورت ماهانه | افزایشی |
| جهانبخش و ترائی (۱۳۸۳) | کشور ایران | ۴۱ | ----- | خیز | حداقل و حداکثر دمای ماهانه، حداقل و حداکثر زمانی پایدگی و دما | تغییرات زمانی پایدگی و دما | SARIMA | عدم وجود روند افزایشی یا کاهشی معنی‌دار |
| ایراهیمی و همکاران (۱۳۸۴) | مشهد | ۲ | ۵۰ | خیز | حداقل، حداکثر، میانگین دما | تعیین روند تغییرات دما | آزمون من کدال، آزمون لئن‌مایر برازش خطی، نمایی و دوجمله‌ای | به‌طور کلی افزایشی |
| غیور و عساکره (۱۳۸۴) | مشهد | ۱ | ۱۰۶ | خیز | دمای متوسط ماهانه | مدل‌سازی، پیش‌بینی | تناوب، مدل قوریه (سینوسی-کسینوسی)، نمودار دوره‌نگار | عدم وجود روند دما |
| موحدی و همکاران (۱۳۸۴) | حوزه مارون | نامشخص | نیم سده | خیز | دمای ماهانه | تغییرات زمانی و مکانی دما | رگرسیون خطی، نقشه‌های دما | در ماه‌های مختلف متفاوت، روند کلی افزایشی |
| حمادی و همکاران، (۱۳۸۵) | اهواز | ۱ | ۴۷ | خیز | دمای ماهانه، سالانه | تأثیر تغییرات دما بر مصرف آب | میانگین پنج‌ساله، مؤلفه تغییرپذیری، مؤلفه روند | به‌صورت صعودی |
| عزیزی و همکاران (۱۳۸۷) | نیمه غربی ایران | ۱۲ | ۵۰ | خیز | ۱۶ متغیر اقلیمی | تغییر اقلیم | تحلیل آماری چند متغیر، مدل‌های پکس | دارای روند متفاوت در مکان‌ها و پارامترهای مختلف |
| خرابی و همکاران (۱۳۹۰) | کهنگیویه و پویز احمد | ۵ | ۳۹ تا ۲۶ | خیز | دبی، بارش، دما و تبخیر | تغییرات | من کدال در مقیاس‌های زمانی مختلف و روش‌های SDI و SPI | در متغیرهای بارش و دبی روند منفی در دما و تبخیر روند مثبت |
| Jalaei و Tularam (۲۰۱۰) | استراليا | ۴ | ۵۰ | خیز | دما و بارش | روند و رابطه متقابل | نمودار طیفی، ARIMA، رگرسیون خطی | رابطه معکوس باران با مقدار دما، روند خاصی مشاهده نشده است. |
| Hughes و همکاران (۲۰۰۶) | قطب جنوب | ۱ | ۵۴ | تا حدودی | حداقل و حداکثر دما ماهانه | روند تغییر دما | معادله رگرسیون چند متغیر با خطاهای غیرنرمال، ARIMA | روند افزایشی به‌علت افزایش حداقل‌های ماهانه |
| Fan و همکاران (۲۰۱۱) | چین | ۴ | ۵۰ | خیز | پایدگی، درجه حرارت | تأثیر پارامترهای اقلیمی بر روی دبی | تحلیل همبستگی و همبستگی جزئی، ARIMA | گرم‌تر و مرطوب‌تر شدن منطقه |

جدول ۲- آمار توصیفی میانگین درجه حرارت ماهانه در ایستگاه شاهمختار (C°)

| آزمون تصادفی بودن | پارامترهای پراکندگی داده‌ها | | | پارامترهای مکانی: میانگین | | |
|-------------------|-----------------------------|--------|-------|---------------------------|-------|---------|
| | ضریب خودهمبستگی پایه ۱ | حداکثر | حداقل | انحراف معیار | دامنه | میانگین |
| ۰/۸۴ | ۲۷/۴ | -۴ | ۸/۰۳ | ۳۱/۴ | ۱۳/۷۵ | ۱۴/۲۳ |



شکل ۱- موقعیت ایستگاه هیدرومتری شاهمختار بر روی رودخانه بشار

تکنیک آماری مورد استفاده: در ابتدا مدل انتخاب شده براساس رابطه (۱) که یک مدل سازی خطی و ساده می‌باشد، انتخاب ده و در ادامه به بهبود و توسعه مدل پرداخته شد:

$$Y_i = C + E_i \quad (1)$$

که در آن، Y_i نشان دهنده درجه حرارت متوسط ماهانه در زمان t, i مقادیر مدل سازی و قسمت غیر تصادفی مدل و E_i نیز خطای حاصل از مدل می‌باشد که دارای میانگین صفر و توزیع کاملاً تصادفی است. در این پژوهش از تکنیک تحلیل کاوشگرانه داده‌ها^۱ (EDA) استفاده شد. در این تکنیک مرحله اول شناسایی مساله، جمع‌آوری داده‌های مورد نیاز و تجزیه و تحلیل داده‌ها است که بیشتر به روش گرافیکی می‌باشد، سپس تعیین و تایید مدل و در نهایت نتیجه‌گیری است (NIST، ۲۰۱۲). البته تعیین مدل فرآیندی همراه با تجزیه و تحلیل داده‌ها می‌باشد. در صورتی که در آمار کلاسیک ابتدا نوع مدل سازی معین شده و سپس به تجزیه و تحلیل داده‌ها پرداخته می‌شود.

به طور مثال در مدل سازی کلاسیک ابتدا مشخص می‌شود که از سری‌های زمانی برای مدل سازی استفاده شود در حالی که در تکنیک EDA با آنالیز داده‌ها مشخص می‌شود که کدام نوع مدل سازی برای این داده‌ها مناسب خواهد بود. شاید بتوان مهم‌ترین مزیت این روش نسبت به روش‌های دیگر را سادگی درک مفاهیم و مراحل مدل سازی به دور از دانستن ریز مباحث ریاضی و سنگین آن دانست. مراحل کار با توجه به تکنیک EDA به صورت گام به گام در زیر آورده شده است. لازم به ذکر است از آنجایی که مدل‌ها در تکنیک EDA در قسمت تجزیه و تحلیل داده‌ها تعیین می‌شود، جزئیات بیشتر در بخش نتایج آورده شده است.

- ۱- شناسایی مساله: مدل سازی درجه حرارت ماهیانه (نوع مدل سازی هنوز مشخص نمی‌باشد).
 - ۲- جمع‌آوری داده‌ها: داده‌های درجه حرارت ماهیانه ایستگاه هیدرومتری شاهمختار در بازه زمانی ۱۳۴۹/۷ الی ۱۳۸۸/۶ (۳۹ سال آبی) تهیه شده به وسیله اداره کل امور آب استان کهگیلویه و بویراحمد مورد استفاده قرار گرفت.
 - ۳- تجزیه و تحلیل داده‌ها و تعیین مدل: در پنج گام صورت گرفت که به ترتیب ذکر می‌شوند.
- گام اول: آزمایش فرضیات موجود در داده‌ها می‌باشد. فرض‌های اصلی در تمامی فرآیندهای اندازه‌گیری به‌ویژه برای اینکه داده‌ها رفتاری مشابه داشته باشند، بایستی قبل از انجام هرگونه تجزیه و تحلیل مورد بررسی قرار گیرند. این

¹ Exploratory Data Analysis

فرض‌ها عبارتند از تصادفی بودن داده‌ها که اصلی‌ترین فرض در تحلیل آماری می‌باشد، توزیع آماری (نرمال) و ایستایی داده‌ها به این معنا که میانگین و واریانس ثابت در طول زمان داشته باشند. از چهار نمودار زیر برای بررسی این فرضیات استفاده شده است.

الف) نمودار ترتیبی^۱: نموداری است که داده‌ها در توالی زمانی نمایش داده می‌شوند (مقادیر درجه حرارت ماهیانه Y_i در مقابل شماره ترتیب زمانی داده‌ها N). این نمودار جهت پاسخ به الف: آیا داده‌ها دارای میانگین ثابت و یا دارای روند خاصی می‌باشند؟ ب: آیا واریانس داده‌ها تغییر یافته است و یا خیر؟ استفاده می‌شود.

ب) نمودار تاخیر^۲ با پایه ۱ (مقادیر Y_i در مقابل Y_{i-1}): با کمک این نمودار به سوالات زیر پاسخ داده می‌شود: ب-۱- آیا داده‌ها تصادفی هستند؟ در واقع اگر داده‌ها تصادفی باشند، داده‌ها به صورت پراکنده و بدون هیچ‌گونه رابطه خاصی بر روی این نمودار ظاهر می‌شوند در غیر این صورت ب-۲- آیا داده‌ها از ساختار توالی خاصی پیروی می‌کنند؟ ب-۳- مدل مناسب برای برازش داده‌ها چیست؟ ب-۴- آیا در بین داده‌ها داده پرت نیز موجود می‌باشد یا خیر؟
ج) هیستوگرام داده‌ها: توزیع داده‌ها را نمایش می‌دهد.

د) نمودار احتمال نرمال^۳: آیا داده‌ها از توزیع نرمال تبعیت می‌کنند یا خیر؟

گام دوم: آزمون تعیین روند: مدل‌سازی روند، برازش مدل خطی

گام سوم: آزمون وجود چرخه در داخل داده‌ها و مدل‌سازی آن: این عمل با توجه به نمودار تاخیر، تجزیه و تحلیل نمودار طیفی^۴ و نمودار دوره‌نگار^۵ صورت می‌گیرد.

گام چهارم: مدل‌سازی چرخه موجود در داخل داده‌ها: با توجه به نتایج به دست آمده از مدل سینوسی استفاده شد.

گام پنجم: تعیین مدل آماری مناسب جهت باقی‌مانده‌ها با توجه به ساختار داده‌ها و استفاده از دو نمودار خودهم‌بستگی^۶ و نمودار خودهم‌بستگی جزئی^۷

۴- معرفی مدل: فرمول نهایی مدل به دست آمده در این قسمت ارائه شده است.

۵- تایید مدل: خطاهای مدل بایستی دارای توزیع نرمال و پراکنش کاملاً تصادفی باشد. در این قسمت از آزمون آماری Ljung-Box جهت تصادفی بودن خطاها و کفایت مدل استفاده شد. البته از آزمون‌های Kolmogorov-Smirnov، Anderson-Darling و Chi-Squared جهت آزمون نرمال بودن خطاها نیز استفاده شد.

نتایج و بحث

نتایج با توجه به گام‌های آورده شده در بخش مواد و روش‌ها به ترتیب ارائه می‌شود. در این قسمت داده‌ها با توجه به زمان آن‌ها به ترتیب شده و به وسیله نرم‌افزار Dataplot V2، ۴ نمودار (شکل ۲) رسم شد. نتایج زیر از تفسیر این ۴ نمودار به دست آمد.

نمودار ترتیبی: نتایج به صورت زیر می‌باشد (شکل ۲ الف):

- واریانس داده‌ها در طول زمان تغییر نکرده است، همان‌طور که دیده می‌شود داده‌ها بین دو خط فرضی موازی با محور زمان قرار گرفته‌اند. با استفاده از آزمون Levene test نیز به صورت کمی برابر بودن واریانس‌ها مورد بررسی قرار گرفت. به این ترتیب که کل داده‌ها به چهار گروه مساوی تقسیم شدند. در فرض H_0 واریانس چهار گروه برابر هستند و در فرض H_1 واریانس حداقل یکی از گروه‌ها با بقیه متفاوت است. با توجه به مقدار F محاسبه شده (۰/۵۰۱) و مقایسه آن با مقدار $F_{0.95,3,464} = ۲/۶۲$ در سطح پنج درصد مورد قبول بوده و واریانس گروه‌ها با هم برابر می‌باشد.

¹ Run Sequence Plot

² lag plot

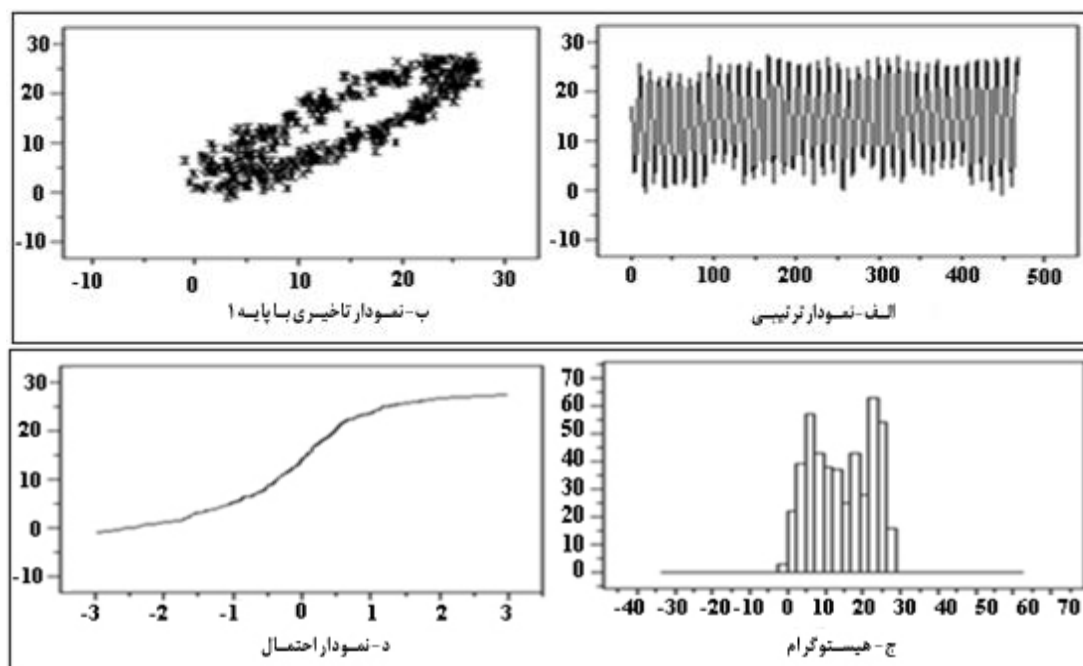
³ normal probability plot

⁴ Spectral Plot

⁵ Periodogram Plot

⁶ Autocorrelation plots

⁷ Partial autocorrelation plots



شکل ۲- نمودارهای به کار رفته در گام نخست جهت آزمون فرضیات اصلی داده‌های آماری

- داده‌ها تقریباً بدون روند می‌باشند و میانگین درجه حرارت ماهیانه تقریباً در طول دوره ثابت می‌باشد. این عمل به صورت کمی نیز با برازش یک خط با فرمول $Y_i = a + b * T_i$ به کل داده‌ها (جدول ۳) مورد آزمون قرار گرفت. در این آزمون فرض H_0 عدم تفاوت شیب خط (b) با صفر می‌باشد. با توجه به اینکه مقدار آماره t شیب محاسبه شده (۱/۹۳) کوچک‌تر از مقدار معیار $t_{0.975,466} = 1/96$ در سطح پنج درصد می‌باشد، فرض H_0 مورد تایید قرار می‌گیرد.

جدول ۳- نتایج برازش خطی با استفاده از میانگین درجه حرارت‌های ماهانه در ایستگاه شاه‌مختار

| ضریب | مقدار تخمینی | خطای استاندارد | آماره t |
|------------|--------------|----------------|---------|
| ثابت (a) | ۱۳/۰۱ | ۰/۷۴۲ | ۱۷/۵۷ |
| شیب خط (b) | ۰/۰۰۵۳ | ۰/۰۰۲۷ | ۱/۹۳ |

- وجود حالت فصلی در داده‌ها به طوری که حداقل مقادیر در ماه‌های فصل زمستان و حداکثر در تابستان قرار می‌گیرد. نمودار تاخیر با پایه ۱: این نمودار (شکل ۲ ب) برای تعیین بخش غیرتصادفی موجود در داده‌ها می‌باشد، همان‌طور که دیده می‌شود:

- داده‌ها دارای یک روند افزایشی با ضریب خودهم‌بستگی ۰/۸۴ (جدول ۲) را نشان می‌دهند که نشان دهنده وجود ساختار زمانی بین داده‌ها می‌باشد.

- شکل حلقوی به وجود آمده نشان دهنده الف: وجود تناوب فصلی در بین داده‌ها ب: مناسب بودن مدل سینوسی جهت مدل‌سازی این چرخه می‌باشد.

- داده‌ها تصادفی نیستند زیرا همان‌طور که دیده شد دارای مولفه فصلی و همچنین، ساختار زمانی هستند. از این قسمت نتیجه می‌شود که استفاده از مدل خطی کارآیی لازم را برای مدل‌سازی ندارد. در این‌طور مواقع مدل‌های سری زمانی مورد توجه محققین می‌باشند. همان‌طور که در معادله (۲) دیده می‌شود مدل‌های سری زمانی

دارای چهار مولفه می‌باشند (Jha و Machiwal، ۲۰۱۲). یک سری زمانی ممکن است دارای روند (T)، خاصیت چرخه فصلی (S)، باقی‌مانده تصادفی وابسته^۱ (E) و باقی‌مانده مستقل^۲ (η) باشد:

$$Y_i = T_i + S_i + \varepsilon_i + \eta_i \quad (2)$$

اگر مقادیر روند، فصلی بودن را از مقادیر مشاهداتی درجه حرارت کم کنیم مقادیر باقی‌مانده (R_i) به دست می‌آید. این مقادیر باقی‌مانده طبق رابطه (۳) خود شامل دو بخش E و η است: که بخش η یعنی باقی‌مانده تصادفی وابسته را می‌توان پس از آزمون تصادفی بودن با مدل‌های سری زمانی ARIMA مدل‌سازی کرد.

$$R_i = \varepsilon_i + \eta_i \quad (3)$$

همان‌طور که دیده می‌شود در این پژوهش مولفه روند وجود ندارد. اما مولفه فصلی بودن (S) وجود دارد که بایستی مدل‌سازی می‌شد. در نهایت مقادیر خطای مستقل (η) را می‌توان برای بررسی کفایت مدل به کار برد. مدل مورد بررسی در این پژوهش به صورت رابطه (۴) نوشته می‌شود.

$$Y = S (\text{Sinusidal Model}) + \eta (\text{ARIMA}) + E (\text{Model Validation}) \quad (4)$$

هیستوگرام و نمودار احتمال نرمال: با توجه به اینکه داده‌ها تصادفی نیستند، نیازی به بررسی دقیق این بخش نیست، زیرا اصلی‌ترین فرض در این بخش تصادفی بودن داده‌ها می‌باشد. بنابراین ابتدا قسمت غیرتصادفی یعنی چرخه از داده‌ها کسر می‌شود تا بخش تصادفی داده‌ها یعنی (R) به دست آید.

مدل‌سازی چرخه: برای مدل‌سازی چرخه با توجه به نمودار تاخیر، مدل سینوسی طبق رابطه (۵) انتخاب شد که تابعی از زمان (T_i) است:

$$Y_i = a + \alpha \sin(2\pi\omega T_i + \phi) + R_i \quad (5)$$

که در آن، Y_i درجه حرارت میانگین ماهیانه، a پارامتر ثابتی که نشان دهنده سطح میانگین داده‌ها است، α دامنه تابع سینوس، ω فرکانس و ϕ نشان دهنده فاز و R_i خطا و باقی‌مانده حاصل برای هر یک از نقاط دمایی می‌باشند. از آنجایی که این مدل حساس به انتخاب مقادیر اولیه مناسب است، بنابراین مقادیر اولیه بایستی با دقت و با روش‌های مناسب انتخاب شوند. در زیر به شرح نحوه به دست آوردن این پارامترها می‌پردازیم:

پارامتر a : از آنجایی که داده‌ها دارای روند خطی به خصوصی نیستند، مقدار میانگین می‌تواند تقریب خوبی برای a باشد. بنابراین میانگین درجه حرارت ماهیانه در این ایستگاه یعنی $14/23$ برای آن در نظر گرفته شد.

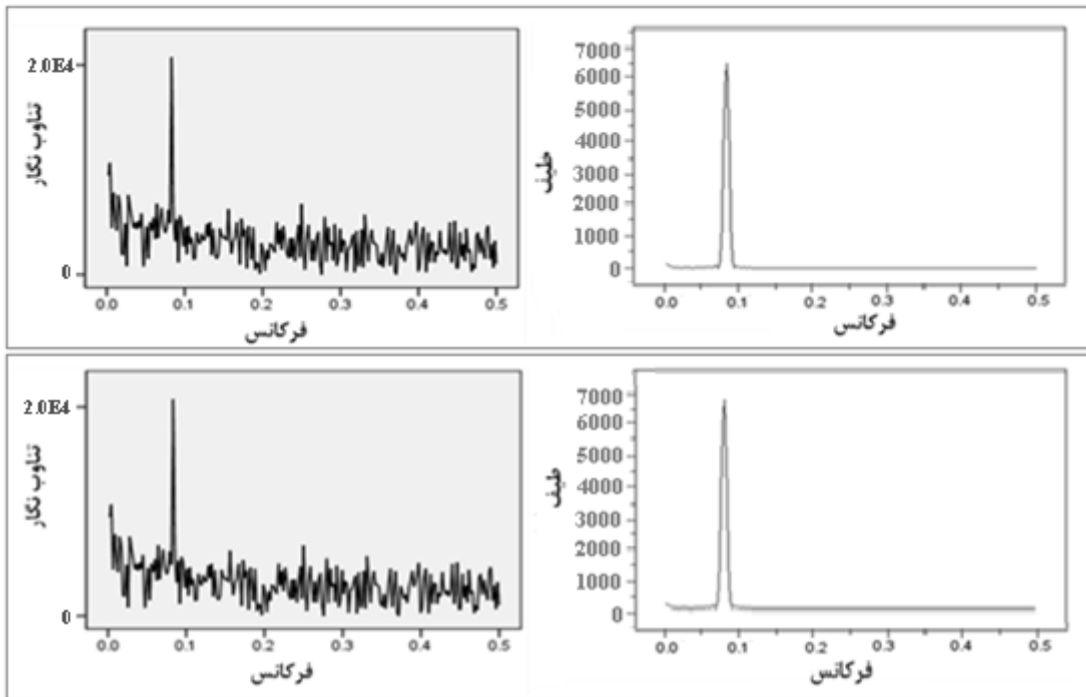
پارامتر ω : در مدل بالا ω نشان دهنده فرکانس غالب می‌باشد که استفاده از نمودار دوره‌نگار (نرم‌افزار SPSS 19) و یا طیف‌نگار (با کمک نرم‌افزار Dataplot 2) می‌تواند روش مناسبی برای تخمین این پارامتر باشد. بنابراین نمودار طیفی و نمودار دوره‌نگار به ترتیب رسم شدند. همان‌طور که در شکل ۳ دیده می‌شود، نمودار طیفی تغییرات را کمتر کرده و به خوبی نشان دهنده وجود یک تناوب غالب با فرکانس $0/083$ (محور افقی نمودار) یعنی همان عدد یک‌دوازدهم می‌باشد. این امر نشان دهنده وجود یک چرخه با فرکانس 12 ماهه (داده‌ها ماهانه می‌باشند) می‌باشد. بنابراین مقدار اولیه ω برابر $0/0833$ در نظر گرفته شد. برای تقریب بهتر مقدار فرکانس می‌توان از نمودار Complex Demodulation Phase (با کمک نرم‌افزار Dataplot 2) استفاده کرد. به‌طور مثال هشت گام قبل و بعد از عدد $0/0833$ را انتخاب می‌کنیم و بعد از رسم نمودارها و تفسیر آنها عدد مناسب‌تر را انتخاب می‌کنیم. مقدار گام‌های عددی $0/0011$ انتخاب شدند. با توجه به نمودار آورده شده در شکل ۴ بهترین مقدار پارامتر زمانی است که شیب خط مربوط به فازهای تخمینی تقریباً برابر صفر شود یعنی موازی محور زمان شود. بنابراین بهترین تخمین برای عدد اولیه برای این پارامتر برابر $0/08333$ است یعنی منحنی که در ستون سوم و ردیف سوم از چپ در شکل ۴ است.

پارامتر α : برای تعیین مقدار اولیه این پارامتر نمودار Complex Demodulation Amplitude (Granger و Hatanaka، ۱۹۶۴) با استفاده از نرم‌افزار Dataplot2 رسم شد. پارامتر α می‌تواند در طول زمان تغییر کند و یک عدد ثابت

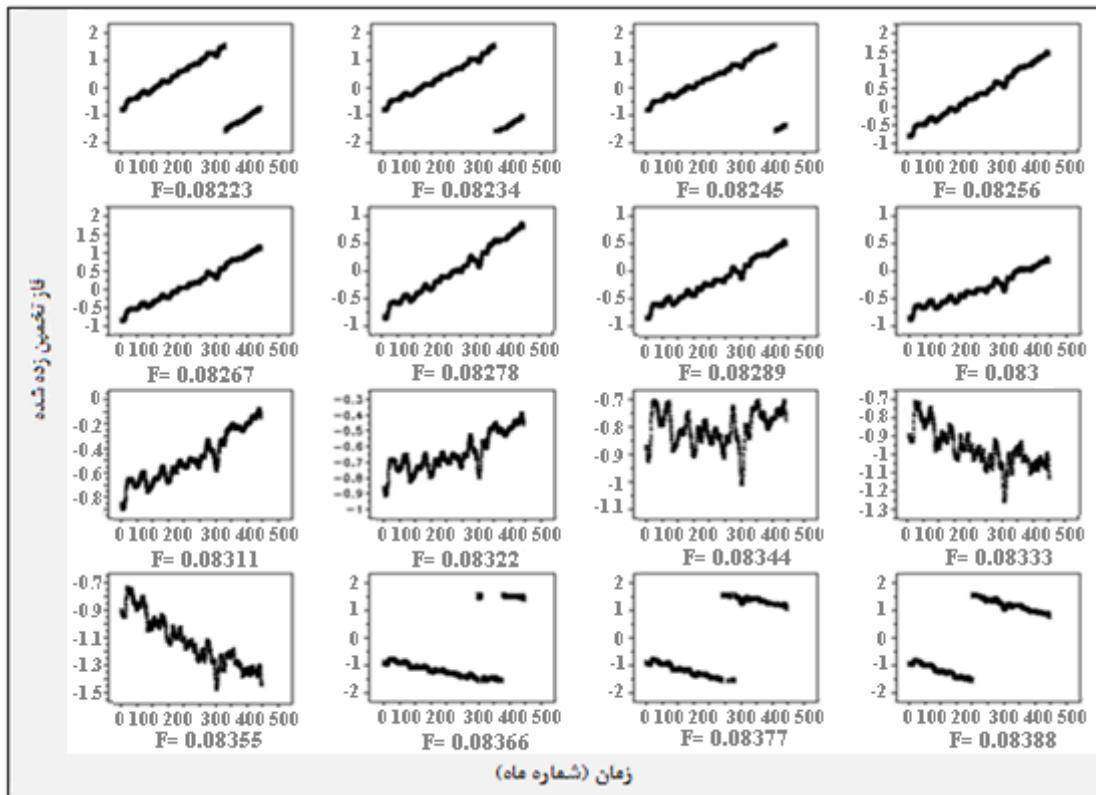
¹ Dependent stochastic component

² Independent residual component

نداشته باشد. این منحنی نشان می‌دهد که آیا این پارامتر در طول زمان ثابت است یا خیر؟ در شکل ۵ مقدار این پارامتر بین ۹/۴۵ و ۱۲/۵ با میانگین ۱۱/۰۲ قرار دارد و نشان‌دهنده تغییر در طول زمان می‌باشد.

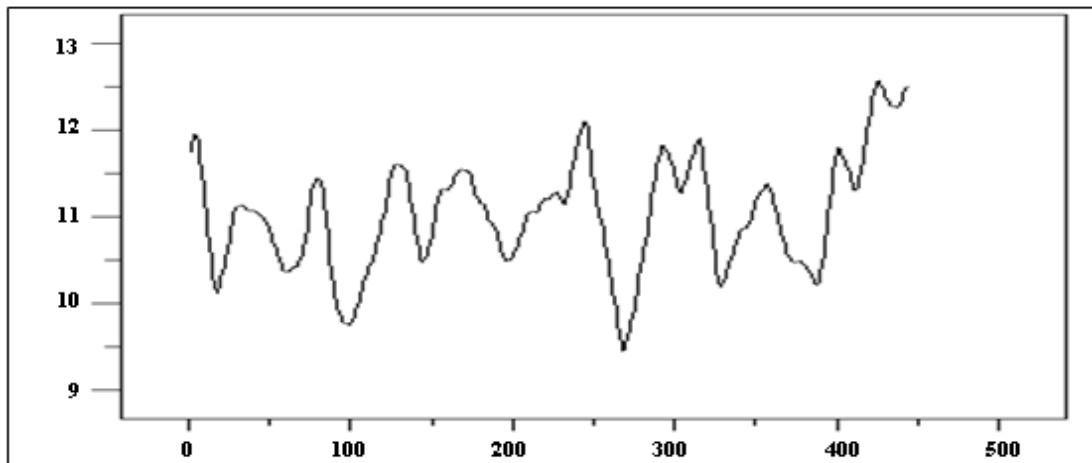


شکل ۳- نمودارهای دوره‌نگار و طیف‌نگار داده‌های درجه حرارت میانگین ماهانه ایستگاه شاه‌مختار



شکل ۴- نمودار Complex Demodulation Phase جهت تعیین مقدار اولیه پارامتر فرکانس (ω)

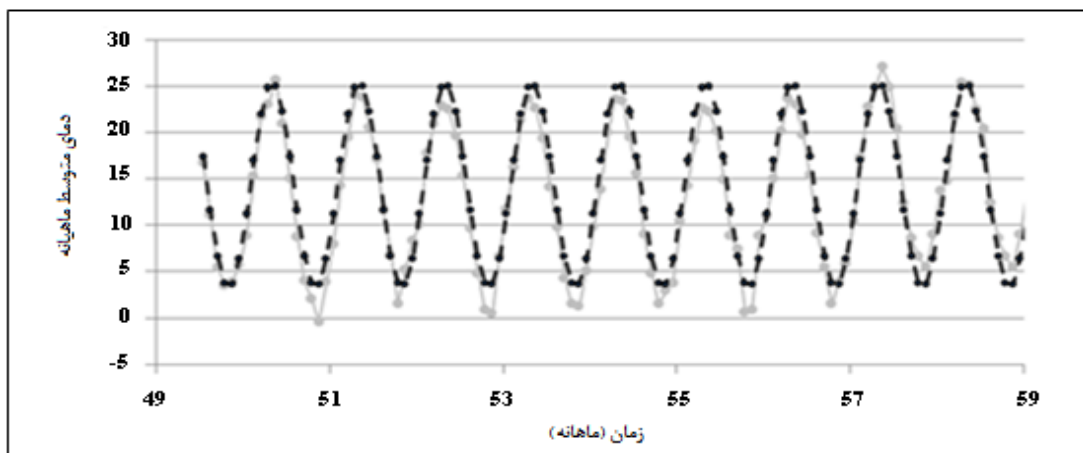
بهترین روش استفاده از این پارامتر مدل سازی این پارامتر به صورت تابعی از زمان است گرچه احتمال دارد دقت مدل را بالاتر ببرد ولی منجر به پیچیده تر شدن مدل سازی و مدل به دست آمده خواهد شد. در این پژوهش از مدل سازی این پارامتر خودداری شد و مقدار میانگین به عنوان مقدار اولیه در نظر گرفته شد. سپس با استفاده از Solver نرم افزار Excel، و قرار دادن این مقادیر اولیه برای سه پارامتر بالا و با استفاده از روش حداقل مجذور مربعات تمامی پارامترها بهینه سازی شدند که مقادیر آنها در جدول ۴ و نمودار برازش داده شده از مدل بر روی داده ها در شکل ۶ ارائه شده است.



شکل ۵- نمودار Complex Demodulation Amplitude. نشان دهنده تغییرات پارامتر دامنه (X) در طول زمان

جدول ۴- مقادیر پارامترهای رابطه سینوسی جهت مدل سازی فصلی بودن داده های درجه حرارت میانگین ماهانه

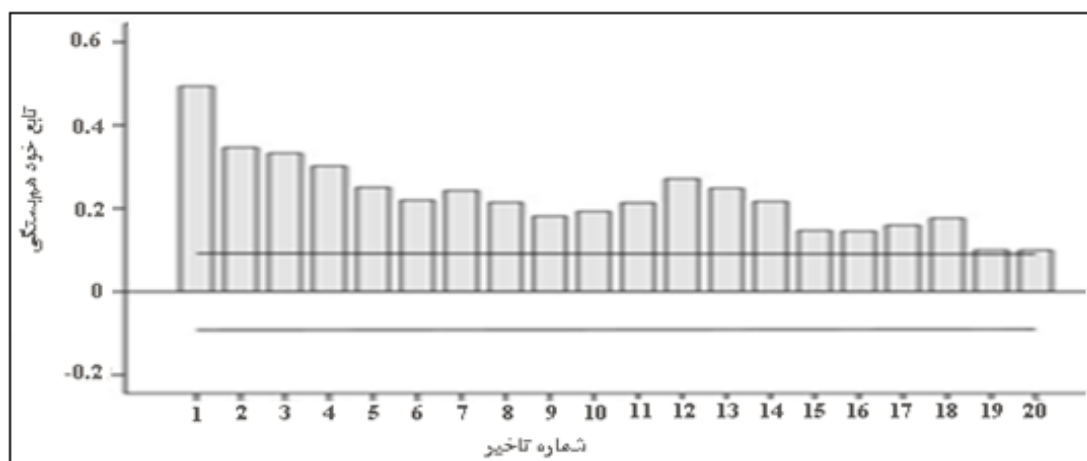
| مدل نهایی | فاز (θ) | فرکانس غالب (ω) | دامنه فرکانس (α) | سطح میانگین (a) |
|-----------|------------------|--------------------------|---------------------------|-----------------|
| Y_i | ۲/۳۳۹۳۸ | ۰/۰۸۳۳۳۱ | ۱۱/۰۱ | ۱۴/۳۱ |



شکل ۶- چرخه مدل سازی شده توسط مدل سینوسی: مشاهده ای (خط چین) و مدل سازی شده (خط خاکستری)

تعیین مدل آماری مناسب برای باقی مانده ها: در این مرحله مقدار باقی مانده های (R) حاصل از تفاضل داده های اصلی و مدل سازی چرخه، مدل سازی می شوند. در ابتدا ۴ نمودار برای باقی مانده ها رسم شدند که تنها به نتیجه آن در این قسمت پرداخته می شود. نمودار ترتیبی روند خاصی را در داده ها نمایش نداد، اما نمودار تاخیری هم بستگی نسبتا بالایی نشان داد. همچنین، توزیع داده ها نیز نرمال نبود. هم بستگی بالا نشان دهنده مناسب بودن مدل های سری زمانی

برای مدل سازی است به همین علت اقدام به رسم نمودار تابع خودهم بستگی داده ها یا ACF^1 (شکل ۷) در سطح ۹۵ درصد شد که نشان دهنده: ۱- وجود هم بستگی مثبت بالا می باشد. ۲- این هم بستگی با سرعت کمی رو به کاهش می گذارد که موید این مطلب است که این فرآیند، می تواند یک فرآیند میانگین متحرک باشد. ۳- از طرف دیگر این هم بستگی بالا می تواند نشان دهنده غیر ایستا بودن این داده ها باشد. ۴- داده ها تصادفی نبوده و دارای ساختاری زمانی هستند که مویدی دیگر بر استفاده از مدل های ARIMA هستند. جهت استفاده از مدل $(ARIMA(p,d,q))$ بایستی رتبه های مربوط به تابع خودهم بستگی (p) ، رتبه تفاضل گیری (d) و رتبه میانگین متحرک (q) برای مدل تعیین شود و الگوی سیستماتیک داده ها یا باقی مانده های وابسته یعنی (ϵ) در رابطه (۳) به دست آید.



شکل ۷- نمودار تابع خودهم بستگی داده های باقی مانده (R)

در ابتدا بایستی داده های باقی مانده ایستا شوند. روش های مختلفی جهت ایستا کردن داده ها وجود دارد که معمول ترین آن ها تفاضل گیری با پایه های مختلف است. به همین علت یک بار از داده تفاضل گیری با پایه $d=1$ (که $\text{diff}1$ نامیده می شود) صورت گرفت. سپس با رسم ۴ نمودار و با تفسیر آن ها مشاهده شد داده ها تا حدودی ایستا شدند. با رسم نمودار ACF برای داده های $\text{diff}1$ در شکل ۸ و با مقایسه آن با نمودار ACF باقی مانده ها (R) در شکل ۷ مقدار زیادی از میزان خودهم بستگی زمانی داده ها کاسته شد و تنها در پایه ۱ هم بستگی $0.43-$ را نشان می دهد. براساس نمودار ACF برای داده های $\text{diff}1$ (شکل ۸)، مدل میانگین متحرک^۲ با رتبه $q=1$ و با توجه به نمودار PACF در شکل ۸ که به صورت توانی تغییر نکرده و با سرعتی کم کاهشی-افزایشی است رتبه $q=2$ برای داده های $\text{diff}1$ پیشنهاد می شود. به عبارت دیگر مدل های $ARIMA(0,0,1)$ و یا $ARIMA(0,0,2)$.

از طرف دیگر چون شکل سینوسی و یا کاهش ناگهانی در ACF دیده نمی شود، مدل از الگوی تابع خودهم بستگی^۳ استفاده نمی کند. اما برای اطمینان با توجه به اینکه تا حدود پایه ۴ دارای خودهم بستگی معنی داری است، مدل های زیر پیشنهاد شدند: $ARIMA(1,0,0)$ ، $ARIMA(2,0,0)$ ، $ARIMA(3,0,0)$ ، $ARIMA(4,0,0)$. جهت انتخاب مدل برتر، نیاز به استفاده از یک معیار می باشد که معیارهای مختلفی برای این کار از جمله AIC^4 ، BIC^5 می باشد. گرچه در اغلب مطالعات از معیار AIC استفاده شده است اما هیچ دلیل خاصی برای برتری یکی از این دو ارائه نشده است (Jha و Machiwal، ۲۰۱۲). با توجه به اینکه نرم افزار SPSS معیار AIC را ارائه نمی دهد از معیار BIC استفاده شد. هر کدام از مدل ها که کمترین مقدار نرمال شده معیار BIC را داشته باشند به عنوان مدل برتر براساس اصل امساک

¹ Autocorrelation Function

² Moving Average

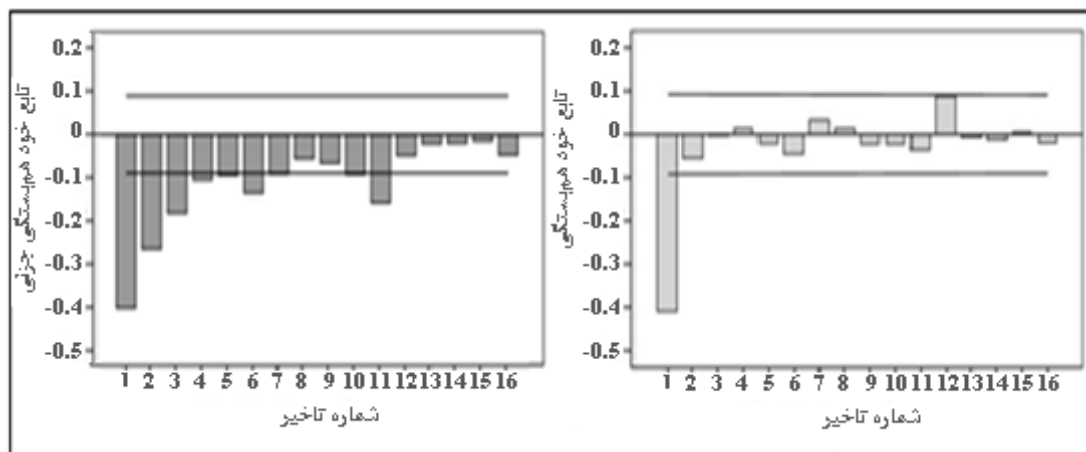
³ Autoregressive

⁴ Akaike Information Criterion

⁵ Bayesian Information Creterion

خواهند بود. با توجه به نتایج به دست آمده، مدل $ARIMA(0,0,2)$ به عنوان بهترین مدل انتخاب شد که مدل آن برابر رابطه (۶) می باشد.

$$diff1_i = a_i + 0.962 * a_{i-1} + 0.192 * a_{i-2} + 0.003 \quad (۶)$$



شکل ۸- تابع خودهمبستگی (ACF) در سمت راست و تابع خودهمبستگی جزئی (PACF) برای داده های $diff1$ در سمت چپ

از طرف دیگر چون شکل سینوسی و یا کاهش ناگهانی در ACF دیده نمی شود، مدل از الگوی تابع خودهمبستگی^۱ استفاده نمی کند. اما برای اطمینان با توجه به اینکه تا حدود پایه ۴ دارای خودهمبستگی معنی داری است، مدل های زیر پیشنهاد شدند: $ARIMA(1,0,0)$ ، $ARIMA(2,0,0)$ ، $ARIMA(3,0,0)$ ، $ARIMA(4,0,0)$.

جهت انتخاب مدل برتر، نیاز به استفاده از یک معیار می باشد که معیارهای مختلفی برای این کار از جمله AIC^2 ، BIC^3 می باشد. گرچه در اغلب مطالعات از معیار AIC استفاده شده است اما هیچ دلیل خاصی برای برتری یکی از این دو ارائه نشده است (Machiwal و Jha، ۲۰۱۲). با توجه به اینکه نرم افزار SPSS معیار AIC را ارائه نمی دهد از معیار BIC استفاده شد. هر کدام از مدل ها که کمترین مقدار نرمال شده معیار BIC را داشته باشند به عنوان مدل برتر براساس اصل امساک خواهند بود. با توجه به نتایج به دست آمده، مدل $ARIMA(0,0,2)$ به عنوان بهترین مدل انتخاب شد که مدل آن برابر رابطه (۶) می باشد.

$$diff1_i = a_i + 0.962 * a_{i-1} + 0.192 * a_{i-2} + 0.003 \quad (۶)$$

از آنجایی که یک بار از باقی مانده ها تفاضل گیری شده بود، بنابراین مدل نهایی ارائه شده به صورت مدل $ARIMA(0,1,2)$ می باشد. لازم به ذکر است پس از این مراحل با استفاده از Modelexpert و با فرآیند خودکار در نرم افزار SPSS19 مدل پیشنهادی نرم افزار نیز به صورت $ARIMA(0,1,2)$ به دست آمد. رابطه (۷) یک مدل ارائه شده برای $diff1$ را با توجه به تفاضل گیری انجام شده در سمت راست معادله بسط داده است:

$$diff1 = R_i - R_{i-1} = a_i + 0.962 * a_{i-1} + 0.192 * a_{i-2} + 0.003 + \eta_i \quad (۷)$$

که در آن، R_{i-1} مقدار باقی مانده در یک زمان قبل است. با جابه جایی در سمت راست رابطه بالا، رابطه (۸) به دست می آید:

$$R_i = R_{i-1} + a_i + 0.962 * a_{i-1} + 0.192 * a_{i-2} + 0.003 + \eta_i \quad (۸)$$

معرفی مدل نهایی: با جای گذاری پارامترهای جدول ۴ در رابطه (۵)، معادله سینوسی چرخه فصلی به دست آمد. مدل فصلی به دست آمده و مدل به دست آمده برای باقی مانده در رابطه (۲) جای گذاری شد که ترکیب مدل نهایی درجه حرارت ماهیانه ایستگاه شاهمختار به دست طبق رابطه (۹) به دست آمد.

¹ Autoregressive

² Akaike Information Criterion

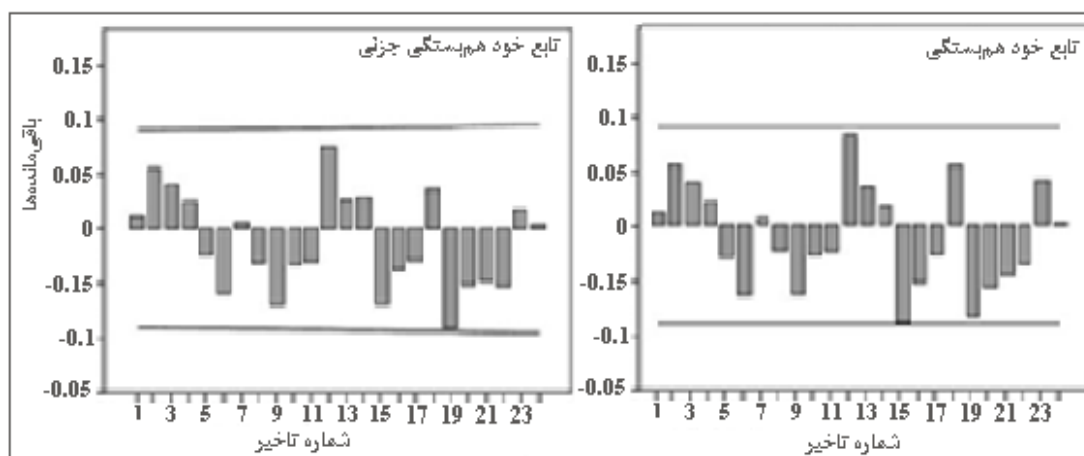
³ Bayesian Information Creterion

$$Y_i = [14.31 + 11.01 \sin(0.523585T_i + 2.33938)] + [R_{i-1} + a_i + 0.962 * a_{i-1} + 0.192 * a_{i-2} + 0.003] + \eta_i \quad (9)$$

که در آن، Y_i دمای میانگین ماه مورد نظر در سری زمان و در مکان i ، R_{i-1} مقدار باقی‌مانده در یک زمان قبل، a_i ، a_{i-1} و a_{i-2} مولفه‌های خطای تصادفی در زمان‌های گذشته i ، $i-1$ و $i-2$ و η_i خطاهای مدل می‌باشند. **تایید مدل:** با به‌دست آوردن خطاهای مدل (η_i) و با مشاهده نمودارهای ACF و PACF مربوطه به آن‌ها در شکل ۹، هم‌بستگی معنی‌داری در پایه‌های مختلف در سطح ۹۵ درصد دیده نشد. بنابراین می‌توان گفت که داده‌ها دارای ساختار خاصی نبوده و تصادفی می‌باشند.

برای کفایت مدل و تصادفی بودن خطاها از آزمون Ljung-Box استفاده شد. در این قسمت تا پایه ۲۴ خودهم‌بستگی مورد بررسی قرار گرفت. آماره آزمون Ljung-Box برابر ۲۴/۳ به‌دست آمد. اگر این مقدار کوچک‌تر از مقدار معیار کای-اسکور محاسبه شده باشد فرض تصادفی بودن داده‌ها رد خواهد شد. مقادیر کای-اسکور محاسبه شده در درصد نقاط ۰، ۵۰، ۷۵، ۹۵ و ۹۹ به‌ترتیب ۰، ۲۳/۳۳، ۲۸/۲۴، ۳۶/۴۱ و ۴۲/۹۸ است. همان‌طور که استنتاج می‌شود تصادفی بودن داده‌های η_i نه تنها در سطح پنج درصد بلکه در سطوح بالاتر نیز تایید می‌شود.

با استفاده از نرم‌افزار EasyFit5.5 pro که یک نرم‌افزار تخصصی جهت بررسی توزیع داده‌هاست و با انجام سه آزمون مختلف نرمال بودن داده‌ها یعنی Kolmogorov-Smirnov، Anderson-Darling و Chi-Squared در سطح پنج درصد تایید شد. فرض H_0 برابر است با نرمال بودن داده‌ها. حال اگر مقدار مقادیر معیار محاسبه شده در هر سطح معنی‌داری بزرگ‌تر از مقدار آماره P محاسبه برای خطاهای به‌دست آمده از مدل باشد، فرض H_0 رد می‌شود. نتایج آزمون نرمال Kolmogorov-Smirnov نشان داد که در سطوح معنی‌دار (α) ۰/۲، ۰/۱، ۰/۰۵، ۰/۰۲ و ۰/۰۱ مقادیر معیار محاسبه شده به‌ترتیب ۰/۰۴۹۶، ۰/۰۵۶۵، ۰/۰۶۲۷، ۰/۰۷۰۱ و ۰/۰۷۵۳ است که از مقدار آماره P (۰/۴۰۸۸) کمتر بوده، در نتیجه فرض H_0 تایید می‌شود و مقادیر خطا از توزیع نرمال تبعیت می‌کنند.



شکل ۹- تابع خودهم‌بستگی (PACF) در سمت راست و تابع خودهم‌بستگی (ACF) برای خطاهای (η_i) در سمت چپ

نتایج حاصل از این طرح را می‌توان به‌صورت زیر خلاصه کرد:

- روش EDA یک روش مناسب جهت مدل‌سازی گام به گام داده‌ها است. با استفاده از این روش و تفسیر نمودارهای گرافیکی موجود در این روش، به‌دور از محاسبات ریاضی پیچیده، به‌راحتی می‌توان اقدام به تفسیر و تعیین مدل مناسب کرد.
- رابطه سینوسی قابلیت مدل کردن بخش تناوب فصلی داده‌ها درجه حرارت ماهیانه را دارد. روش‌های مختلفی برای مدل‌سازی بخش تناوبی داده‌ها، علی‌الخصوص داده‌های میانگین دما وجود دارد. به‌طور مثال می‌توان به روش SARIMA (صراف و جامعی، ۱۳۸۱؛ سعیدآبادی و همکاران، ۱۳۸۲؛ جهانبخش و ترابی، ۱۳۸۳) و یا استفاده از مدل

فوریه (سینوسی-کسینوسی) توسط غیور و عساکره (۱۳۸۴) اشاره کرد اما با توجه به روش‌های ذکر شده، روش به کار رفته در این پژوهش از دیگر روش‌ها ساده‌تر می‌باشد.

- تحلیل سری‌های زمانی میانگین دمای ماهانه در ایستگاه شاه مختار جهت مدل‌سازی این پارامتر مناسب است.
- قابلیت بهبود و توسعه مدل نهایی ارائه شده در این پژوهش با در نظر گرفتن جایگزین‌های مناسب برای مدل‌سازی فصلی (مانند بسط فوریه)، استفاده از مدل‌سازی چند متغیره به همراه عوامل دیگر هواشناسی و یا برای مطالعات بعدی، پیشنهاد می‌شود مولفه دامنه تابع سینوس (X) در رابطه (۵) به عنوان یک متغیر در نظر گرفته شود و برای آن یک مدل جداگانه ایجاد شود.

منابع مورد استفاده

۱. ابراهیمی، ح.، ا. علیزاده و س. جوانمرد. ۱۳۸۴. بررسی وجود تغییر دما در دشت مشهد به عنوان نمایه تغییر اقلیم در منطقه. فصلنامه تحقیقات جغرافیایی، ۷۹: ۱۸-۵.
۲. براتی، غ و م. طرقی. ۱۳۷۸. تعیین روند تغییرات دما و بارش شهر مشهد طی دوره آماری ۹۴ - ۱۹۵۱. فصلنامه تحقیقات جغرافیایی، ۵۴: ۱۶۵-۱۵۱.
۳. جهانخش، س. و س. ترابی. ۱۳۸۳. بررسی و پیش بینی تغییرات دما و بارش در ایران. فصلنامه تحقیقات جغرافیایی، ۷۴: ۱۲۵-۱۰۴.
۴. حمادی، ک. م. جامعی و ن. شاهین‌زاده. ۱۳۸۵. تحلیل سری زمانی درجه حرارت متوسط سالانه و تاثیر آن بر مصرف آب کشاورزی. همایش ملی مدیریت شبکه‌های آبیاری و زهکشی، دانشگاه شهید چمران اهواز، دانشکده مهندسی علوم آب، ۱۲-۱۴ اردیبهشت.
۵. صراف ج. و ب. جامعی. ۱۳۸۱. مدل‌سازی و پیش‌بینی درجه حرارت ایستگاه‌های منتخب غرب کشور. مجله فضای جغرافیایی، ۶۸-۴۱.
۶. سعیدآبادی، ر.، ا. جهانخش و س. فاخری فرد. ۱۳۸۲. مدلی برای پیش‌بینی تغییرات دمایی با استفاده از مطالعات رژیم دمایی و روند تغییرات آن در طول زمان (مطالعه موردی ایستگاه تبریز). فصلنامه جغرافیا و برنامه‌ریزی، ۱۳: ۱۱۶-۹۵.
۷. غیور، ح. و ح. عساکره. ۱۳۸۴. کاربرد مدل‌های فوریه در برآورد دمای ماهانه و آینده‌نگری آن، مطالعه موردی: دمای مشهد. فصلنامه پژوهشی علوم انسانی دانشگاه اصفهان، ۷۷: ۹۹-۸۳.
۸. کاویانی، م.ر. و ح. عساکره. ۱۳۸۰. بررسی و مدل‌سازی روند دما طی سده گذشته، مطالعه موردی: ایستگاه جاسک. فصلنامه دانشکده ادبیات و علوم انسانی، ۲۶: ۳۸-۱۹.
9. Adeloye, A.J. and M. Montaseri. 2002. Preliminary streamflow data analyses prior to water resources planning study. *Hydrological Sciences Journal*, 47: 679-692.
10. Granger, C.W.J. and M. Hatanaka. 1964. Spectral analysis of economic time series. Princeton University Press, 299 pages.
11. Li, Y., B.M. Chen, Z.G. Wang and S.L. Peng. 2011. Effects of temperature change on water discharge, and sediment and nutrient loading in the lower Pearl River Basin based on SWAT modelling. *Hydrological Sciences Journal*, 56: 68-83.
12. Machiwal, D. and M.K. Jha. 2006. Time series analysis of hydrologic data for water resources planning and management: a review. *Journal of Hydrology and Hydromechanics*, 54: 237-257.
13. Machiwal, D. and M.K. Jha. 2012. Hydrologic time series analysis: Theory and Practice. Capital Publishing Co., New Delhi, India and Springer, Germany.
14. NIST. 2012. SEMATECH E-Handbook of Statistical Methods, <http://www.itl.nist.gov/div898/handbook/>, last update April 2012.
15. Rao, A.R., K.H. Hamed and H.L. Chen. 2003. Nonstationarities in hydrologic and environmental time series. Springer.
16. Tularam, G.A. and M. Ilahee. 2010. Time series analysis of rainfall and temperature interactions in Coastal Catchments. *Journal of Mathematics and Statistics*, 6: 372-380.
17. Rao, Y.R.S., B. Krishna and P.C Nayak. 2011. Time series modeling in water resources planning and management. *International Journal of Earth Sciences and Engineering*, 4: 247-253.
18. Fan, Y.T., Ch. YaNing, Li. WeiHong, W. HuaiJun and X. Li. 2011. Impacts of temperature and precipitation on runoff in the Tarim River during the past 50 years. *Journal of Arid Land*, 3: 220-230.

Modeling the average monthly temperature, case study: Shah-Mokhtar hydrometric station

Mohammad Reza Mirzaei¹, Assistant Professor, Faculty of Agriculture, Yasouj University, Iran

Received: 10 September 2012

Accepted: 18 February 2013

Abstract

In this study, the mean monthly air temperature data from Shah-Mokhtar hydrometric station in Kohgiluyeh and Boyer Ahmad province for a period of 39 years between 1970 and 2009 was investigated. Using different graphical EDA techniques such as spectral, autocorrelation and partial autocorrelation plots, i) existence of seasonal part and ii) suitability of using time series analysis to model the residuals were determined. The lag plot and autocorrelation plot of the original data showed that a sinusoidal model was appropriate to model the seasonal effect. So using sinusoidal model and determining its parameters precisely, the seasonal effect was modeled properly. Time series analysis was also used to model the residuals using ARIMA models. Among different models, ARIMA(0,1,2) model was selected as the best model using Normalized Bayesian Information Criterion (NBIC). Finally, null hypotheses for Kolmogorov-Smirnov and Ljung-Box tests were not to be rejected at 5% level for the obtained model which confirms the adequacy of the model.

Key words: ARIMA, Autocorrelation, EDA, Temperature, Time series

¹ Corresponding author: mmirzaei@mail.yu.ac.ir