

مدل سازی خوشه‌ای تراز آب زیرزمینی دشت تبریز با استفاده از مدل آریما

رسول جانی^{*}

۱- استادیار گروه عمران، دانشگاه آزاد اسلامی واحد تبریز
* نویسنده مسئول: jani@iaut.ac.ir

تاریخ پذیرش مقاله: ۱۳۹۸/۰۸/۰۷

تاریخ دریافت مقاله: ۱۳۹۸/۰۱/۳۱

چکیده

مطالعه‌ی تغییرات زمانی و مکانی و تخمین نوسانات تراز آب‌های زیرزمینی در مطالعات و برنامه‌ریزی مدیریت منابع آب برای استمرار و یا توسعه‌ی بهره‌برداری آن، از اهمیت بسیاری برخوردار است. نمونه‌ی پرکاربرد از مدل‌های تصادفی در بررسی سری‌های زمانی تراز آب زیرزمینی، مدل خود همبسته‌ی یکپارچه میانگین متحرک (آریما) می‌باشد. در تحقیق حاضر برای بررسی تغییرات زمانی و مکانی تراز آب‌های زیرزمینی دشت تبریز در مقیاس ماهانه، ۴۶ ایستگاه انتخاب شده است. در محدوده‌ی مورد مطالعه بین مقادیر سطح آب زیرزمینی چاه‌های مشاهداتی نزدیک به همدیگر همبستگی بیشتری نسبت به سایر چاه‌ها وجود دارد. از این‌رو برای جلوگیری از بررسی اثرات نوسانات چاه‌های نزدیک به همدیگر، در ابتدا کل ایستگاه‌ها با استفاده از فرآیند خوشه‌بندی به هفت بخش تقسیم شده و مطالعه‌ی مکانی بر روی چاه‌های منتخب واقع در هر کدام از این بخش‌ها، انجام شده است. نتایج، حاکی از آن است که خوشه‌های اول، سوم و هفتم به خاطر وجود داده‌های پرت، دقت قابل قبولی در شبیه‌سازی را نداشتند و چهار خوشه‌ی دیگر از دقت خوبی برخوردار بودند. همچنین، خوشه‌ی دوم و اول به ترتیب بهترین و بدترین مدل‌سازی را با دارا بودن بالاترین ضریب تعیین و کمترین ریشه‌ی میانگین مربعات خطا داشته‌اند. در ضمن دقت مدل آریما زمانی کاهش می‌یابد که سری‌های مورد بررسی با نوسان‌های بیش از حد در میانگین و واریانس با روند همراه باشند و موجب ناپیوستگی سری گردد. واژه‌های کلیدی: آریما، تغییرات زمانی و مکانی، خوشه‌بندی، نقشه حرارتی، دندروگرام.

مقدمه

و ساده بیان گردد. امروزه این روش در سطح وسیعی به کار گرفته می‌شود تا درک پیچیدگی‌های سیستم هیدرولوژیکی مورد مطالعه امکان‌پذیر گردد (عساکره، ۱۳۸۶ الف). در این میان، مدل‌هایی که بر پایه‌ی اصول آماری و احتمالی بنا گذاشته شده‌اند، از اهمیت ویژه و کاربرد پرشماری برخوردارند. در این تحقیق، تلاش بر این است که تراز آب زیرزمینی دشت تبریز در ایستگاه‌های منتخب به وسیله‌ی یکی از مدل‌های آماری موسوم به مدل خود همبسته‌ی یکپارچه‌ی میانگین متحرک^{۱۰} (آریما) مدل‌سازی شود. از این مدل، می‌توان برای

آب زیرزمینی مهم‌ترین منبع آب شیرین بر روی زمین است (ویلنیو، ۱۹۹۰) که مدیریت بهره‌برداری از این منبع در مناطق خشک و نیمه خشک جهان از جمله ایران اهمیت بیشتری می‌یابد. لذا، درک رفتار این مؤلفه‌ی مهم هیدرولوژیکی و پیش‌بینی آن کمک شایانی در تأمین آب مورد نیاز منطقه و توسعه پایدار آن از حیث فعالیت‌های اجتماعی و اقتصادی خواهد داشت. تهیه‌ی مدلی برای مطالعه‌ی سیستم هیدرولوژیکی، کمک می‌نماید تا مهم‌ترین ویژگی متغیر مورد نظر را به صورت کلی

¹⁰ Autoregressive Integrated Moving Average (ARIMA)

گرفتن تأخیر زمانی برای متغیرهای ورودی، باعث تخمین دقیق تر مقادیر سطح آب زیرزمینی شد. جعفرزاده و همکاران (۱۳۹۶) روش‌های پایه تحلیل سری‌های زمانی شامل مدل‌های خودبرگشتی، میانگین متحرک و ترکیبی از خودبرگشتی و میانگین متحرک (آریما) را بر روی داده‌های تراز آب زیرزمینی دشت اردبیل، اجرا نموده‌اند. با استفاده از آزمون‌های ضریب آکائیک و توابع خودهمبستگی بهترین مدل انتخاب و ملاحظه گردید که مدل $AR(2)$ نتایج بهتری نسبت به سایر مدل‌ها داشت. در نهایت، با استفاده از همین مدل، برای دوره‌ی زمانی سی ساله پیش‌بینی انجام شد. نتایج حاصل از مقادیر پیش‌بینی شده به وسیله‌ی سری‌زمانی، نشان می‌دهد که در صورت ثابت ماندن الگوی مصرف و هم‌چنین عدم تغییرات در روند تغذیه‌ی سفره در طی سی سال آینده، سطح سفره حدود ۱۱ متر نسبت به وضعیت فعلی کاهش خواهد یافت.

افروزی و زارع‌ابیانه (۱۳۹۶) نوسانات تراز آب زیرزمینی در دشت همدان را به روش مدل‌های سری زمانی مورد بررسی قرار دادند. پس از ویرایش و استفاده از چند ضلعی تیسن برای داده‌های خام، سری‌زمانی تراز آب زیرزمینی هر دشت تشکیل شد. آزمون من-کندال وجود روند قطعی در تمام سری‌های زمانی دشت‌های استان همدان را نشان داد. سپس آزمون ریشه‌ی واحد برای بررسی مانایی سری‌های زمانی انجام شد و با توجه به روش باکس-جنکینز، مدل‌های آریما فصلی روی داده‌های نمونه برازش داده شد و مناسب‌ترین آن‌ها انتخاب شدند. شهرکی و همکاران (۱۳۹۷) نوسانات سطح آب زیرزمینی دو پیرومتر واقع در دشت کرمان را با مدل‌های شبکه عصبی مصنوعی، سری زمانی آریما و رگرسیون خطی چند متغیره بررسی نمودند. مدل سری زمانی آریما (۱ و ۰) و (۲ و ۰) برای پیرومتر جنوب باغین و مدل سری زمانی آریما (۱ و ۱) و (۲ و ۰) برای پیرومتر اراضی فرودگاه بهترین برازش را با داده‌ها داشته است. در مدل‌سازی با رگرسیون خطی چند متغیره برای هر دو پیرومتر

پیش‌بینی تغییرات زمانی سطح آب زیرزمینی استفاده نمود. چرا که نوسانات سطح آب زیرزمینی در طول سال به‌صورت تناوبی عمل می‌کند و مدل خود همبسته‌ی میانگین متحرک^{۱۱} (آرما) برای چنین سری‌های زمانی دوره‌ای بهتر عمل می‌کند. لذا در این تحقیق سعی بر آن است تا با استفاده از مدل تصادفی آریما، تغییرات سطح آب زیرزمینی دشت تبریز بررسی گردد. علاوه بر این، به منظور کاهش حجم محاسبات در برآورد مقادیر تراز آب زیرزمینی در مقیاس ماهانه از روش خوشه‌بندی فازی بهره گرفته شده است تا با انتخاب ایستگاه‌های کافی درک درستی از تغییرات این مولفه‌ی مهم هیدروژئولوژیکی ایجاد شود.

صادق‌زاده و همکاران (۱۳۹۶) عوامل مؤثر بر آلودگی آبخوان دشت تبریز را با بهره‌گیری از شبکه مصنوعی، درخت تصمیم و زمین آمار بررسی نمودند و نتایج مطالعات نشان داد که کیفیت آبخوان دشت تبریز نسبت به سال‌های قبل نامطلوب‌تر شده است. بطوری‌که شوری چاه‌ها در طی سال‌های ۱۳۸۰-۱۳۹۰ حدود ۲۲ درصد افزایش پیدا کرده است. باباعلی و دهقانی (۱۳۹۶) تغییرات سطح ایستابی را با مدل‌های شبکه عصبی مصنوعی و موجکی از روی ورودی‌های همچون بارندگی، دما، دبی و تراز سطح آب ایستابی ماه قبل پیش‌بینی نمودند، مدل شبکه عصبی موجکی با بیشترین ضریب همبستگی و کمترین ریشه مربعات خطا بهترین عملکرد در پیش‌بینی را داشته است.

نیکبخت و نوری (۱۳۹۶) چاهک‌های مشاهده‌ای آبخوان دشت مراغه در استان آذربایجان شرقی را خوشه‌بندی و تراز آب زیرزمینی را به کمک شبکه‌های عصبی مصنوعی پیش‌بینی نمودند. به کمک خوشه‌بندی انجام گرفته، یک خوشه با سه زیرخوشه‌ی همگن، انتخاب و نماینده‌ی هر زیرخوشه، تعیین شد و در ادامه با استفاده از شبکه‌های عصبی مصنوعی، تراز آب زیرزمینی نماینده‌ی هر زیرخوشه، شبیه‌سازی گردید. نتایج نشان داد که در نظر گرفتن داده‌های دمای هوا به عنوان ورودی در شبکه‌های عصبی مصنوعی موجب اغتشاش شبکه و در نظر

¹¹ Autoregressive Moving Average (ARMA)

روش‌های موجود در این بخش را در دو گروه سلسله مراتبی و غیر سلسله مراتبی (افرازی) قرار می‌دهند.

روش سلسله مراتبی مانند یک درخت عمل می‌کند، به این صورت که هر شاخه‌ی کوچک‌تر جزئی از شاخه‌ی بزرگ‌تر است و در نهایت کل مجموعه به صورت سلسله مراتبی به تنه‌ی درخت وصل می‌شوند و اشیاء به صورت نمودار درختی و به فرم بازگشتی در خوشه‌های کوچک و کوچک‌تر قرار می‌گیرند که اصطلاحاً دندروگرام^{۱۲} گفته می‌شود. مزیت این روش این است که از طریق آن می‌توان رابطه‌ی سلسله مراتبی بین داده‌ها را کشف کرد و نیز از قبل نباید تعداد خوشه‌ها را تعیین کرد. البته، ایراد وارد بر این روش پیچیدگی و طولانی بودن تحلیل و تفسیر خوشه‌ها می‌باشد. در روش سلسله مراتبی، خوشه‌بندی با دو رویکرد اصلی تراکمی یا تجمیعی (پایین به بالا) و تقسیمی (بالا به پایین) انجام می‌شود. در روش غیر سلسله مراتبی (افرازی) برخلاف روش سلسله مراتبی، از قبل باید تعداد خوشه‌ها را مشخص و سپس اقدام به خوشه‌بندی کرد. همچنین، آنچه که از این روش به دست می‌آید به جای ایجاد نمودار درختی (دندروگرام)، خوشه‌های مجزا از یکدیگر است. به عبارت دیگر، حاصل این روش، خوشه‌هایی است که بین اعضای آن ارتباطی وجود ندارد. برای ایجاد این گونه خوشه‌ها، باید برخی از اشیاء را به عنوان نماینده برگزید و شباهت اشیای دیگر با آنها را سنجید. بر این اساس، هر شیء در خوشه‌ی مربوط به خود جای می‌گیرد و از این رو، اشیای نماینده عملاً مراکز خوشه‌ها را تشکیل می‌دهند (موی و سارسدت، ۲۰۱۱).

از سوی دیگر، در علم داده‌کاوی، نقشه‌ی حرارتی^{۱۳} ابزاری برای تشخیص داده‌های با رفتار مشابه با همدیگر می‌باشد. همین طور، از این نقشه برای تمایز سری‌های با رفتار مختلف استفاده می‌شود. به عبارت دیگر، سری‌ها به صورت عناصر دودویی با هم مورد مقایسه قرار می‌گیرند. به این منظور، طیف رنگی برای تمایز میان سری‌های با رفتار مختلف تعریف می‌گردد. به طوری که

بهترین روابط مدل بدست آمد به نحوی که عمق آب زیرزمینی ماه فعلی تابعی از عمق آب زیرزمینی ماه قبل است. نتایج آزمون معیار خطا بین سه مدل حاکی از پیش‌بینی خوب رگرسیون خطی چند متغیره از مدل‌های شبکه عصبی و مدل سری زمانی آریمای باشد.

رضایی‌بنفشه و همکاران (۱۳۹۷) پژوهشی با هدف پیش‌بینی تغییرات سطح آب زیرزمینی حوضه تسوج را با الگوی سری زمانی انجام دادند. برای این منظور داده‌های تراز آب زیرزمینی در مقیاس ماهانه بین سال‌های ۱۳۹۱-۱۳۷۳ اخذ شده و پس از برازش الگوهای مختلف بر داده‌ها، الگوی آریمای (۲،۰،۳) با داشتن کمترین مقدار آکائیک نسبت به سایر الگوها مناسب تشخیص داده شد. لذا، افت سطح آب زیرزمینی از سال ۱۳۹۲ تا سال ۱۴۰۱ نسبت به سال پایه (۱۳۷۳) با الگوی به دست آمده برازش و تغییرات پیش‌بینی گردید. نتایج نشان داد با فرض ادامه یافتن شرایط بهره‌برداری موجود و روند کنونی الگوی مصرف آب زیرزمینی در حوضه، میانگین افت تجمعی سطح آب زیرزمینی در سال دهم پیش‌بینی (۱۴۰۱) نسبت به سال اول پیش‌بینی (۱۳۹۲) ۷/۲۵ متر خواهد بود.

با نگاهی به مطالعات صورت گرفته بر روی تراز آب زیرزمینی غالباً یک یا چند چاه مشاهداتی با استفاده از مدل‌های تصادفی بررسی شده است، در حالی که در این تحقیق بررسی بر روی تمام چاه‌های منتخب می‌باشد.

مواد و روش

روش‌های خوشه‌بندی

از دیدگاه سنتی روش‌های خوشه‌بندی به دو دسته کلی سلسله مراتبی و افرازی تقسیم می‌شوند و در سال‌های اخیر روش‌های نوین مثل الگوریتم‌های کی مینز، کی مدویدز و خوشه‌بندی فازی به آن‌ها افزوده شده‌اند. ویژگی اصلی روش‌های سنتی در این است که هر شی تنها در یک خوشه قرار می‌گیرد و به طور کلی

¹³ Heatmap

¹² Dendrogram

پیش‌بینی شرط معکوس‌پذیری فرآیند $\{W_t\}$ نیز لازم است. شرط کافی برای معکوس‌پذیر بودن فرآیند $\{W_t\}$ این است که قدر مطلق هر یک از ریشه‌های معادله $\theta_p(B) = 0$ بزرگ‌تر از واحد باشد.

در رابطه (۳) مؤلفه‌های رابطه به قرار زیر است:

$$\mu_w = E(W_t), a_t \sim WN(0, \sigma^2) \quad (4)$$

رابطه (۴) حاکی از این است که فرآیند $\{a_t\}$ دنباله‌ای از متغیرهای تصادفی هم‌توزیع و دو به دو ناهمبسته (نوفه‌ی سفید^{۱۴}) با امید ریاضی صفر و واریانس ثابت (σ_a^2) است. معمولاً فرض می‌شود که a_t دارای توزیع نرمال است. در این صورت $\{a_t\}$ دنباله‌ای از متغیرهای تصادفی مستقل با توزیع $N(0, \sigma_a^2)$ است و به اختصار مطابق رابطه (۵) نوشته می‌شود.

$$a_t \sim WN(0, \sigma^2) \quad (5)$$

عملگرهای به کار رفته در رابطه (۵) مطابق روابط (۶) و (۷) تعریف می‌شوند.

$$\phi_p(B) = 1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p \quad (6)$$

$$\theta_q(B) = 1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q \quad (7)$$

در رابطه (۷) B عملگری است که مطابق معادله (۸) تعریف می‌شود.

$$B^j Z_t = Z_{t-j} \quad (8)$$

که در رابطه (۸)، J یک عدد صحیح نامنفی است.

اگر سری تفاضلی $\{Z_t\}$ یعنی $\{W_t\}$ یک وقوع به اندازه کافی طولانی از فرآیند $\{W_t\}$ باشد و اگر فرضیه‌های حاکم بر مدل صحیحی باشد، فاصله $\left(\bar{W} \pm 1.96 \frac{s_w}{n}\right)$ یک فاصله اطمینان ۹۵ درصد برای μ_w است که در آن $\bar{W}$ ، s_w ، n به ترتیب میانگین، انحراف معیار و طول سری زمانی $\{W_t\}$ است. با استفاده از این فاصله‌ی اطمینان می‌توان فرض صفر $\mu_w = 0$ را در مقابل $\mu_w \neq 0$ در سطح ۵ درصد خطا آزمود. میانگین سری زمانی $\{W_t\}$ یعنی $\bar{W}$ برآورد مناسبی از μ_w می‌باشد. در عمل می‌توان سری زمانی $\{W_t - \bar{W}\}$ را مورد مدل‌سازی قرار داد و سپس میانگین

رنگ تعریفی حاکی از مشابهت رفتار یا اختلاف بین آن دو سری است.

در این تحقیق، برای بررسی مشابهت رفتارهای سری زمانی از ترکیب دندروگرام و نقشه حرارتی استفاده شده است.

مدل خود همبسته‌ی یکپارچه‌ی میانگین متحرک (مدل آریما)

یکی از مدل‌های مناسب برای برازش بر یکی از فرآیندهای حاکم بر سیستم هیدروژئولوژیکی مورد مطالعه مدل خود همبسته‌ی یکپارچه‌ی میانگین متحرک است که در ادامه برای سادگی، عنوان آریما برای این مدل به کار گرفته می‌شود. مدل آریما برای شناخت انواع مختلفی از سری‌های زمانی قابل کاربرد است و مطابق با رفتار سری، نوع مدل نیز تغییر می‌یابد. از جمله داده‌های قابل کاربرد شامل: خطی (مدل‌های آریمای خطی)، غیرخطی (مدل‌های آریما غیرخطی) در داده‌های تناوبی (مدل‌های اس‌آریما با آریمای فصلی) و غیرتناوبی (مدل‌های آریمای غیرفصلی) می‌باشد.

در این صورت فرآیند تصادفی $\{Z_t\}$ را فرآیند آریما با درجه (p, d, q) گفته و به صورت نمادی به فرم معادله (۱) نشان داده می‌شود.

$$\{Z_t\} \sim \text{ARIMA}(p, d, q) \quad (1)$$

شرط نخست برای برازش مدل آریما بر فرآیند تصادفی $\{Z_t\}$ مانایی (ایستایی) است که منظور از مانایی ضعیف مرتبه دوم است که آن را مانایی کواریانس نیز می‌نامند. اگر این شرط احراز نشود لازم است قبل از برازش مدل بر روی سری زمانی مورد مطالعه به منظور احراز مانایی عملگر تفاضلی مطابق رابطه (۲) اعمال شود.

$$W_t = (1 - B)^d Z_t \quad (2)$$

$$\phi_p(B)(1 - B)^d(W_t - \mu_w) = \theta_q(B) a_t \quad (3)$$

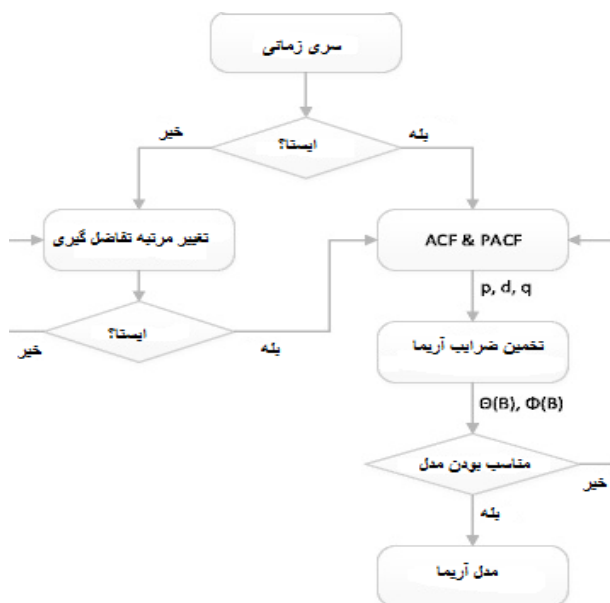
شرط کافی برای اینکه فرآیند $\{W_t\}$ مانا باشد این است که قدر مطلق هر یک از ریشه‌های معادله $\phi_p(B) = 0$ بزرگ‌تر از واحد باشند. بایستی توجه شود، برای قابل استفاده بودن مدل در امر

¹⁴ White noise

حذف شده را اضافه نمود. روش دیگر به دست آوردن میانگین سری $\{W_t\}$ نوشتن مدل به فرم معادله (۹) است:

$$\varphi_{P(B)}(1-B)^d Z_t = \theta_0 + \theta_q(B)a_t \quad (9)$$

مراحل کار مدل آریما را می‌توان مطابق شکل ۱ نشان داد.



شکل ۱- فلوچارت مراحل کار مدل آریما (دمیر و کولری ارگن، ۲۰۱۶)

آزمون ریشه واحد یا دکی فولر

برای بررسی ایستایی متغیرهای سری زمانی، آزمونی به نام ریشه واحد شهرت یافته است که توسط شخصی به نام دکی فولر ارائه شده است. در بررسی ایستایی یک مدل اگر مدل مورد نظر ما دارای ریشه واحد باشد، یعنی ضریب عبارت y_{t-1} در معادله $y_t = \alpha y_{t-1} + \varepsilon_t$ که یک معادله رگرسی شده‌ی y_t بر y_{t-1} است، برابر یک باشد. در این صورت متغیر سری زمانی غیر ایستاست و y_t دارای ریشه واحد است. برای ایستا نمودن از متغیر سری زمانی تفاضل‌گیری می‌شود. اگر با یک بار تفاضل‌گیری متغیر سری زمانی ایستا شد در آن صورت آن متغیر را انباشته از مرتبه یک گویند در غیر این صورت تفاضل‌گیری تا مرتبه‌های بالاتر ادامه داده می‌شود تا متغیر ایستا شود. ایستایی یا غیر ایستایی متغیرها را می‌توان با استفاده از آماره‌ی دکی فولر تعمیم یافته نیز بررسی کرد. بدین صورت که پس از

محاسبه‌ی این آماره، آن را در سطوح مختلف با مقادیر بحرانی جدول مربوط به آن مقایسه می‌نمایند. اگر قدر مطلق این آماره بزرگ‌تر از قدر مطلق مقادیر بحرانی شود، فرضیه صفر مبنی بر غیر ساکن بودن سری رد می‌شود و متغیرهای سری زمانی ما ساکن یا ایستا خواهند بود. در غیر این صورت اگر قدر مطلق آماره دکی فولر تعمیم یافته کوچک‌تر از قدر مطلق مقادیر بحرانی گردد، متغیرهای سری زمان ما نا ایستا خواهند بود و باید با عمل تفاضل‌گیری ایستا شوند (چان فنگ و همکاران، ۲۰۱۷). در این تحقیق به منظور مدل‌سازی تراز آب زیرزمینی در ایستگاه‌های منتخب دشت تبریز با استفاده از مدل آریما از جعبه ابزارهای موجود در نرم‌افزار R استفاده شده و مراحل در نظر گرفته شده برای این مدل مطابق گام‌های بیان شده است.

در بخش مدل‌سازی توسط روش آریما داده‌ها به دو بخش آموزش و آزمون تقسیم گردیده است؛ داده‌های بخش آموزش شامل ۸۰٪ کل داده‌ها بوده و بقیه تحت عنوان داده‌های آزمون نامیده

جنوب شرقی به سمت جنوب غربی و شمال غربی و نیز از غرب به سمت شرق است. بیشترین ژرفای سطح آب زیرزمینی در بخش شرقی و جنوبی، و کمترین ژرفای سطح آب زیرزمینی در بخش غربی دشت دیده می شود (اصغری مقدم و برزگر، ۱۳۹۳).

برای درک رفتار سیستم دینامیکی لازم است علاوه بر بُعد زمان از لحاظ بُعد مکان نیز مدل سازی انجام گیرد. به این منظور، در این پژوهش میزان مؤلفه تراز آب زیرزمینی در سطح حوضه و به

ازای ایستگاه های مختلف مدل سازی گردیده است.

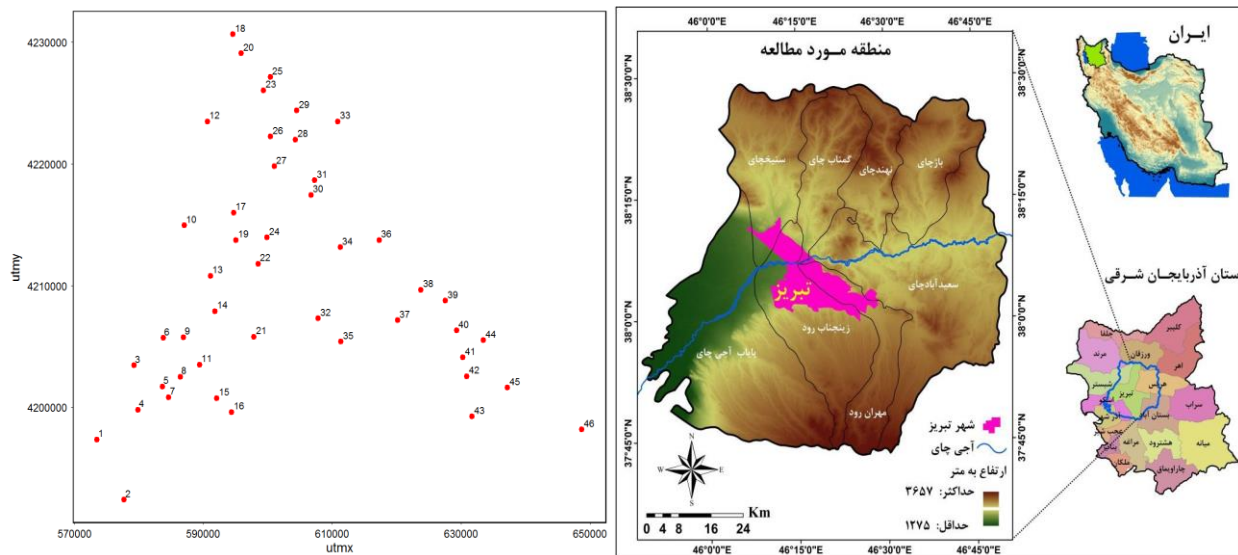
با وجود کوتاهی بُعد زمان سری های زمانی برای هر کدام از ایستگاه های موجود در سطح حوضه، تعداد ایستگاه ها قابل ملاحظه بوده، به این منظور در گام اول ایستگاه هایی انتخاب گردید که فصل مشترک زمانی آن ها طولانی بوده و علاوه بر این دارای کمترین میزان داده های اندازه گیری نشده باشد. پس با ملاک قرار دادن دو حالت فوق، طولانی بودن سطح هم پوشانی طول زمانی سری و کمترین میزان داده اندازه گیری نشده، ۴۶ ایستگاه انتخاب شد.

در شکل ۲ موقعیت قرارگیری ایستگاه های منتخب نیز در سطح دشت تبریز مشخص شده اند.

شده اند. هدف از مدل آریمای این است که با معرفی تاریخچه عملکرد سیستم دینامیکی، مدل آموزش یافته و نحوه عملکرد سیستم با معادلاتی تعریف گردد و در مواردی که مدل قبلاً با آن مواجه نشده است، مورد استفاده قرار گیرد و رفتار سیستم دینامیکی با استفاده از معادلات به دست آمده برای دوره های آتی پیش بینی گردد.

معرفی منطقه و داده های مورد استفاده

منطقه مورد مطالعه در این تحقیق، دشت تبریز واقع در استان آذربایجان شرقی است. منطقه مطالعه شده، اقلیم نیمه خشک دارد و میانگین دمای سالانه این دشت ۱۲/۸ درجه سانتی گراد، میانگین بارش سالانه آن ۳۰۰ میلی متر است. بیشترین ارتفاع منطقه ۳۶۵۷ متر در ارتفاعات کوه سهند و کمترین ارتفاع در بستر رودخانه آجی چای برابر ۱۲۷۵ متر از سطح دریاست (رضائی مقدم و همکاران، ۱۳۹۵). دشت تبریز شکل شبه مثلثی دارد رودخانه آجی چای مهم ترین رودخانه منطقه از ضلع شمال غربی این مثلث وارد دشت شده و در جهت های غربی و جنوب غربی به مسیر خود ادامه داده و در نهایت به دریاچه ارومیه تخلیه می شود (شکل ۲). جهت جریان عمومی آب زیرزمینی به پیروی از توپوگرافی منطقه، بیشتر از شمال شرقی و



شکل ۲- نمایی از منطقه مورد مطالعه و موقعیت ایستگاه های اندازه گیری تراز آب زیرزمینی منتخب در محدوده شهر تبریز

مشخصات آماری مربوط به داده‌های مورد استفاده در جدول ۱ ارائه شده است. در این تحقیق، داده‌های ایستگاه‌های منتخب از دشت تبریز به دو مجموعه آزمون و آموزش تقسیم شده و سپس در فرآیند مدل‌سازی مورد استفاده قرار گرفته‌اند. هشتاد درصد داده‌های ابتدایی برای آموزش و ۲۰ درصد انتهایی داده‌ها برای آزمون انتخاب شد. علاوه بر این، داده‌های هر ایستگاه به ازای ترکیبات مختلفی در فرآیند مدل‌سازی به کار گرفته می‌شوند که این ترکیب‌ها در مدل آریمای به فرم تأخیرهای مختلف داده‌های ورودی است. داده‌های ورودی به عنوان متغیرهای مستقل یا پیش‌بینی کننده‌هایی بوده و داده خروجی به عنوان متغیر وابسته یا پیش‌بینی شونده^{۱۶} می‌باشد.

در این تحقیق، ایستگاه‌های منتخب از دشت تبریز برای بازه‌ی داده‌برداری فروردین سال ۱۳۸۰ لغایت اسفند ۱۳۸۸ مورد مطالعه قرار گرفته است. داده‌های در دسترس پس از بررسی و انجام اصلاحات مورد نیاز از جمله تعیین فصل مشترک بازه زمانی داده‌ها و نیز تعیین مقادیر اندازه‌گیری نشده^{۱۵} برای تحلیل به کار گرفته شده‌اند. مقادیر اندازه‌گیری نشده به صورت میانگین مقادیر موجود پر شده تا سری کامل گردد. بازه زمانی مورد استفاده به نحوی انتخاب گردیده است تا در این بازه کمترین مقدار داده‌های اندازه‌گیری نشده در بازه مورد مطالعه قرار بگیرد تا تنها با استفاده از میانگین‌گیری مقادیر ناموجود را بدست آورد. داده‌های سال‌های اخیر به دلیل شرایط فوق قابل کاربرد نبودند. در ادامه

جدول ۱- مشخصات آماری داده‌های مورد استفاده

مشخصات آماری						نام ایستگاه	خوشه
ضریب چولگی	انحراف معیار	واریانس	میانگین	مینیمم	ماکزیمم		
-۸/۶۹	۳/۲۱	۱۰/۳۳	۸۵/۷۲	۵۴/۶۰	۸۸/۶۰	۳۵	۱
-۰/۰۳	۰/۴۱	۰/۱۷	۲۵/۴۷	۲۴/۶۱	۲۶/۳۶	۲۹	۲
۰/۷۴	۰/۹۸	۰/۹۶	۷/۳۴	۰/۶۹	۱۴/۵۰	۳۶	۳
-۰/۱۵	۰/۴۷	۰/۲۲	۱۱/۱۹	۱۰/۰۷	۱۲/۲۴	۴۰	۴
-۰/۴۷	۰/۶۶	۰/۴۳	۳/۳۱	۱/۵۵	۴/۷۱	۱۷	۵
-۰/۲۶	۳/۳۹	۱۱/۴۸	۱۲۸/۵۰	۱۱۷/۳۵	۱۳۴/۰۷	۴۳	۶
۰/۱۰	۰/۷۷	۰/۵۹	۴/۱۳	۲/۴۰	۵/۸۶	۶	۷

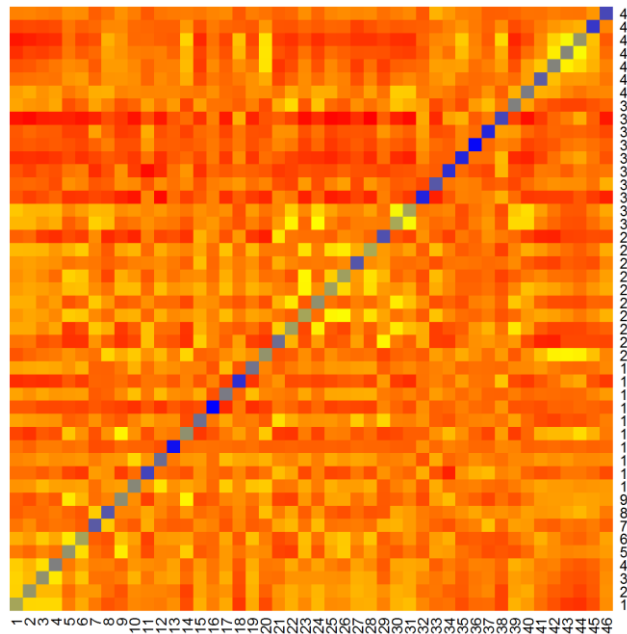
نتایج و بحث

رفتاری بین دو سری مورد بررسی بوده، که این انطباق باتوجه به شکل برای المان‌های واقع در قطر اصلی ماتریس قابل مشاهده است. رنگ قرمز نشان‌دهنده‌ی اختلاف یا تمایز میان دو سری است و در نهایت رنگ زرد حالت بینابینی میان دو رنگ آبی و قرمز را معرفی می‌کند.

در این تحقیق نمودار حرارتی به‌عنوان یکی از روش‌های ارتباط سری زمانی باهمدیگر بکاربرده شده است. نمودار حرارتی برای پارامتر تراز آب زیرزمینی اندازه‌گیری شده در ایستگاه‌های منتخب دشت تبریز به صورت شکل ۳ می‌باشد. مطابق شکل ۳ رنگ‌های تعریفی به این قرار است که رنگ آبی حاکی از مطابقت

¹⁶ Predictant

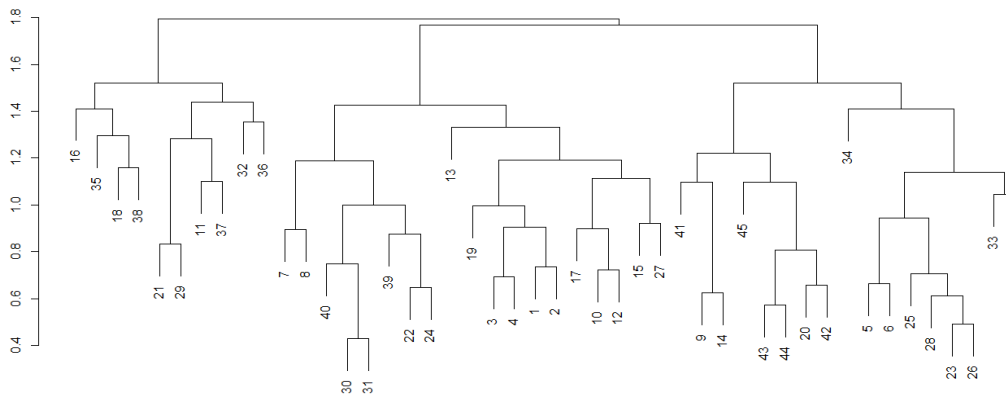
¹⁵ Missing data



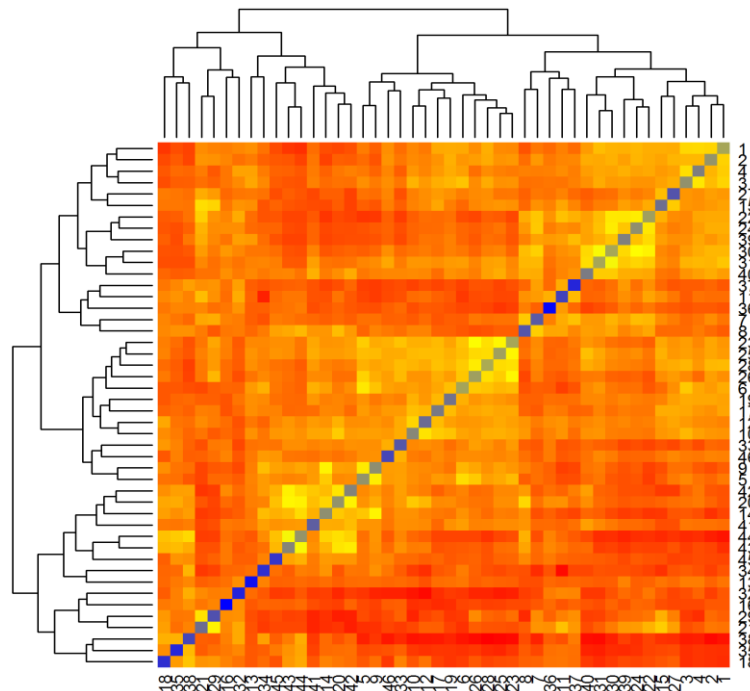
شکل ۳- نمودار حرارتی پارامتر تراز آب زیرزمینی میان ایستگاه‌های منتخب در دشت تبریز

در ادامه به منظور تشخیص بهتر رفتار سری‌های مورد بررسی از نمودار دندروگرام استفاده شد. تقسیم سری زمانی هر کدام از ایستگاه‌ها به خوشه‌های مجزا از هم با استفاده از نمودار دندروگرام امکان‌پذیر می‌باشد و برای حالتی بیشتر توصیه می‌شود که تعداد نقاط مورد بررسی زیاد است. این نمودار، سری‌های مشابه با همدیگر را در داخل خوشه یکسان قرار می‌دهد و انتخاب تعداد مناسب خوشه براساس تجربه پژوهشگر و دقت مورد نیاز در فرآیند مدل‌سازی است. نمودار دندروگرام برای سری‌های زمانی منتخب مولفه‌ی تراز آب زیرزمینی در دشت تبریز در شکل ۴ ارائه شده است.

با توجه به شکل ۴ قرارگیری ایستگاه‌های ۳۰ و ۳۱ در کنار همدیگر حاکی از مشابهت رفتاری سری‌های این دو ایستگاه می‌باشد که این مشابهت با توجه به شکل (۴) نیز قابل تایید است. به طوری که رنگ نشان داده شده در شکل ۳ برای این دو ایستگاه زرد می‌باشد. در ادامه برای مقایسه هر چه بهتر نتایج به دست آمده از نمودارهای نقشه حرارتی و دندروگرام ارائه شده در شکل‌های ۳ و ۴، این دو نمودار با همدیگر ترسیم و در شکل ۵ ارائه گردید. مطابق شکل ۵ نواحی زرد حاکی از مشابهت سری‌ها نسبت به همدیگر است و رنگ قرمز تفاوت میان سری‌ها را نشان می‌دهد.



شکل ۴- نمودار دندروگرام سری زمانی مولفه‌ی تراز آب زیرزمینی میان ایستگاه‌های دشت تبریز



شکل ۵- نمودار حرارتی به همراه دندروگرام سری زمانی تراز آب زیرزمینی ایستگاه‌های دشت تبریز

در فرآیند خوشه‌بندی هر چقدر تعداد خوشه‌ها زیاد در نظر گرفته شود، مدل‌سازی از دقت زیادتری برخوردار خواهد بود در مقابل، این دقت با صرف هزینه زیاد در فرآیند مدل‌سازی توأم است. خوشه‌های در نظر گرفته شده برای ایستگاه‌های منتخب در دشت تبریز در جدول ۲ ارائه شده است. مطابق جدول ۲ هفت خوشه تعریف شده است که این تعداد خوشه براساس جذر تعداد ایستگاه‌های منتخب در سطح حوضه تبریز می‌باشد.

جدول ۲- خوشه‌بندی ایستگاه‌های مورد بررسی در دشت تبریز

ایستگاه منتخب	ایستگاه‌ها	خوشه
۳۵	۱۶،۳۵،۱۸،۳۸	۱
۲۹	۲۱،۲۹،۱۱،۳۷	۲
۳۶	۳۲،۳۶	۳
۴۰	۷،۸،۴۰،۳۰،۳۱،۳۹،۲۲،۲۴	۴
۱۷	۱۳،۱۹،۳،۴،۱۰،۲،۱۷،۱۰،۱۲،۱۵،۲۷	۵
۴۳	۴۱،۹،۱۴،۴۵،۴۳،۴۴،۲۰،۴۲	۶
۶	۳۴،۵،۶،۲۵،۲۸،۲۳،۲۶،۳۳،۴۶	۷

انتخاب ایستگاه‌ها در هر کدام از خوشه‌ها با استفاده از نمودار دندروگرام صورت گرفته به این صورت که خط افقی مجازی تراز از نمودار را قطع نموده تا ۷ خوشه ایجاد گردد. علاوه بر این، ایستگاه منتخب در هر کدام از خوشه‌ها به نحوی در نظر گرفته شده‌اند تا از لحاظ مکانی بیشترین سطح حوضه را پوشش داده و نیز تفاوت کامل در شرایط توپولوژی داشته باشند. همچنین بیشترین تفاوت در الگوی رفتاری سری زمانی داشته باشند تا قابلیت مدل آریما در مدل‌سازی انواع مختلف سری‌های

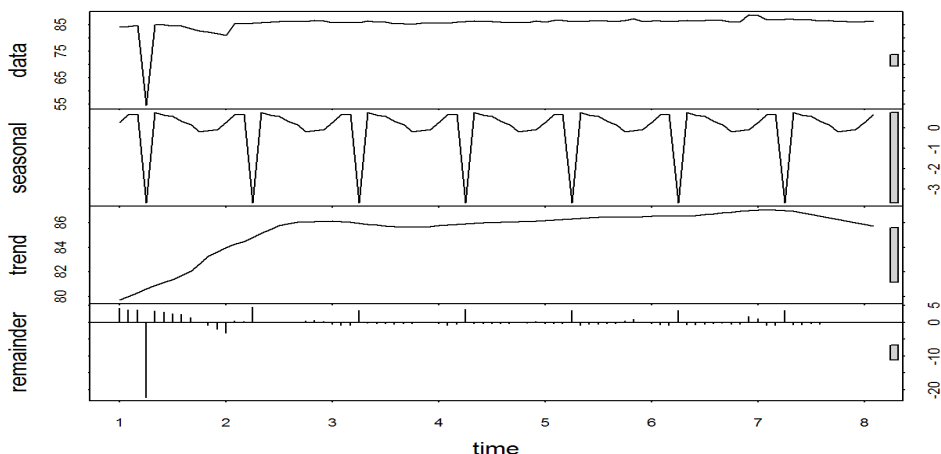
به عنوان ورودی مدل مورد استفاده قرار گرفت. مراحل مدل سازی توسط مدل آریما شامل واسنجی مدل، کنترل ایستایی، انتخاب مقادیر ضرایب مدل، ارزیابی مدل برای داده های بخش آموزش و آزمون توسط نمودارهای همبستگی نگار، همبستگی نگار جزئی و سایر آزمون های مربوطه و پیش بینی در صورت نیاز می باشد.

در شکل ۶، اجزاء تشکیل یافته سری زمانی مربوط به ایستگاه منتخب از خوشه ی اول نشان داده شده است. همان طوری که در شکل فوق قابل مشاهده است، سری زمانی مربوط به بخش آموزش از مولفه های فصلی^{۱۷}، روند^{۱۸}، و باقی مانده^{۱۹} تشکیل یافته است که مجموع این مولفه ها داده اصلی را تشکیل می دهد.

زمانی مشخص شود. با توجه به جدول فوق از میان ایستگاه های ۱۶، ۳۵، ۱۸ و ۳۸ در خوشه اول، ایستگاه انتخابی ۳۵ است، به همین صورت از میان ایستگاه های تعریفی ۳۴، ۵۶، ۲۵، ۲۸، ۲۳، ۲۶، ۳۳، ۴۶ در خوشه ی هفتم، ایستگاه شماره ی ۶ انتخاب شده است.

نتایج مدل آریما

به منظور مدل سازی میزان تراز آب زیرزمینی به کمک مدل های تصادفی از روش آریما استفاده شد. در تحقیق حاضر برای اجرای مدل آریما، هر کدام از سری های زمانی به ازای ایستگاه های منتخب حاصل شده از روش خوشه بندی در سطح دشت تبریز



شکل ۶- اجزاء تشکیل دهنده ی سری زمانی تراز آب زیرزمینی ایستگاه منتخب خوشه اول

متناظر با آن به ازای فرض صفر^{۲۲} نایستا برابر با ۰/۰۱ است که با فرض سطح اطمینان ۵ درصد، مقدار پی کوچک تر از ۰/۰۵ می باشد. در نتیجه دلایل لازم برای رد فرض صفر موجود بوده و با رد آن، فرض ثانویه قبول می شود. در نتیجه، سری جدید بدون مولفه فصلی ایستا می باشد.

لازم است در ابتدا، مولفه ی فصلی سری از داده ی اصلی حذف و مدل سازی بر روی آن انجام گیرد. به این منظور با حذف مولفه ی فصلی و تشکیل سری جدید ایستایی، مورد آزمون قرار می گیرد. برای انجام این کنترل از آزمون دیکی-فولر^{۲۰} بهره گرفته شد. با انجام آزمون مزبور بر روی داده های بخش فصلی سری مورد مطالعه، مقدار آماره ی آزمون ۴/۹۰۹۷- بوده و مقدار پی^{۲۱}

²⁰ Dicky-Fuller test

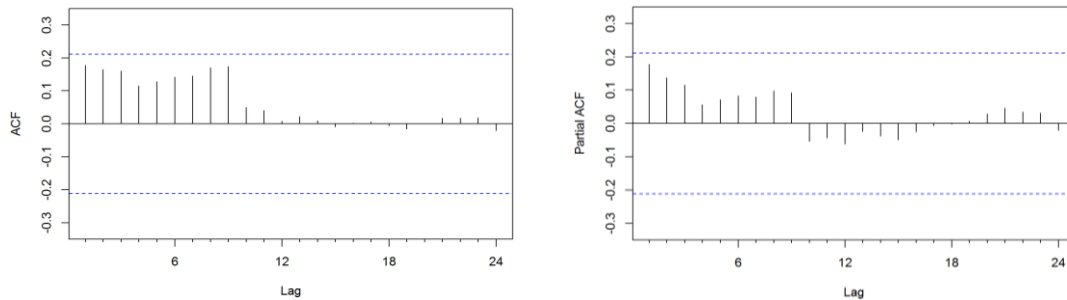
²¹ P-value

²² Null hypothesis

¹⁷ Seasonal

¹⁸ Trend

¹⁹ Remainder



شکل ۷- توابع همبستگی نگار و همبستگی نگار جزئی داده‌های خوشه‌ی اول بدون مولفه‌ی فصلی

نمودارهای توابع همبستگی نگار و همبستگی نگار جزئی (شکل ۷) برای سری زمانی ایستگاه خوشه‌ی اول بدون مولفه‌ی فصلی نشان می‌دهند که سری زمانی تقریباً ایستا است. البته خود سری زمانی با توجه به عدم وجود روند، نوسان و غیره حاکی از

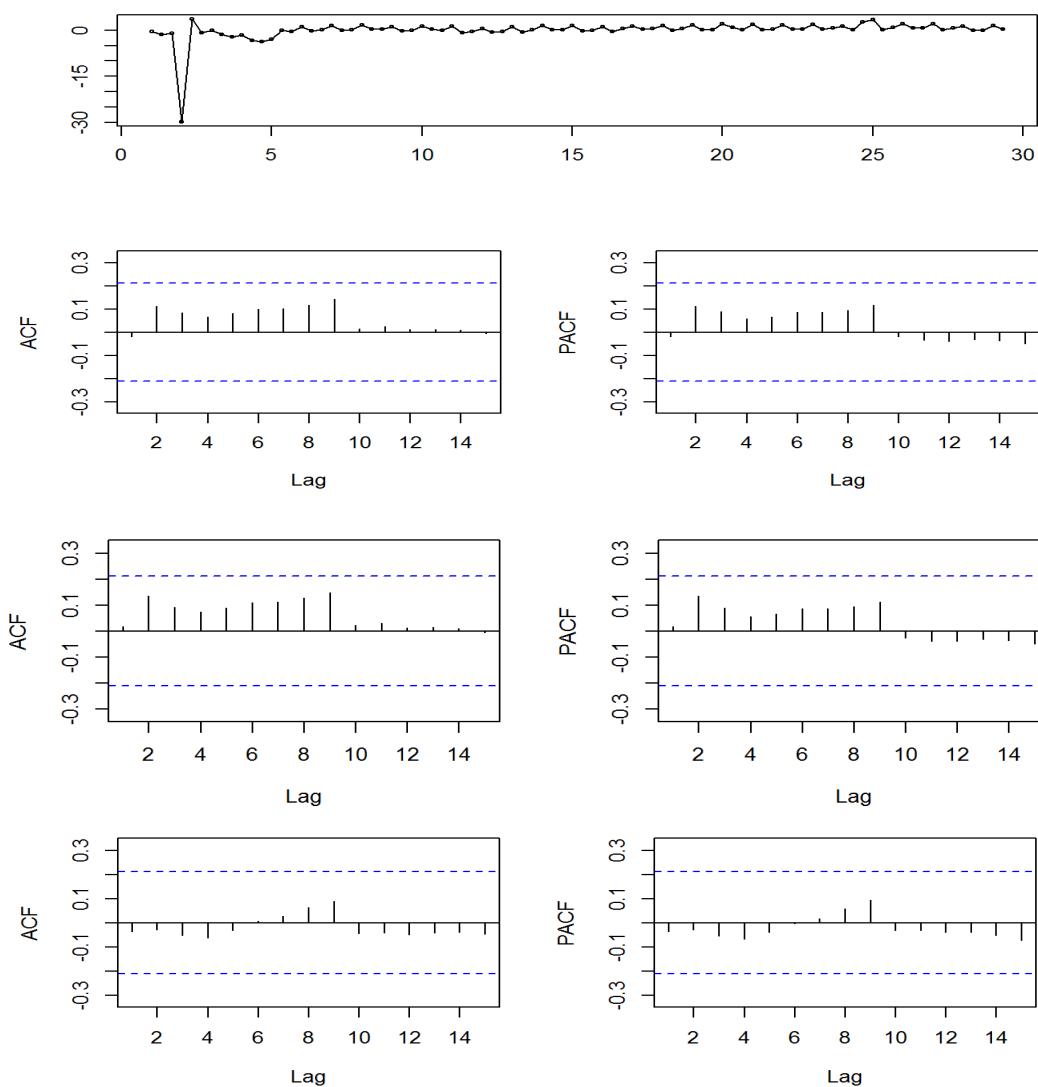
ایستابودن داده‌ها بوده، لذا مقدار d صفر در نظر گرفته می‌شود. پارامترهای مدل‌های پیشنهادی در جدول ۳ و نمودارهای توابع همبستگی نگار و همبستگی نگار جزئی در شکل ۸ قابل مشاهده است.

جدول ۳- مقایسه مدل‌های پیشنهادی مختلف برای بخش غیرفصلی سری زمانی از خوشه اول

مدل پیشنهادی	ضرایب مدل آریما				AIC
	AR	d	MA	Intercept	
الف: (۱،۰،۰)	۰/۱۵۲۹	-	-	۸۵/۵۱۳۶	۴۶۳/۷۵
ب: (۰،۰،۱)	-	-	۰/۱۲۳۸	۸۵/۵۱۳۷	۴۶۴/۱۶
ج: (۱،۰،۱)	۰/۹۶۸۷	-	-۰/۸۶۳۶	۸۵/۰۱۵۸	۴۶۰/۰۶

مقدار آلفا برابر با ۰/۰۵، میزان پی فراتر از مقدار آلفا بوده که این فراتر بودن به معنای این است که شواهد لازم برای رد فرض صفر موجود نمی‌باشد. به عبارت دیگر فرض صفر با عنوان نایستایی سری مورد قبول می‌باشد. برای تبدیل سری نایستا به ایستا از روش تفاضل‌گیری استفاده شده که با اعمال آن بر روی سری و کنترل مجدد آزمون مقدار آماره‌ی آن $۸/۰۱۵۸$ - بدست آمد و میزان پی برابر با ۰/۰۱ تعیین شد که بصورت مجدد با فرض مقدار آلفا معادل ۰/۰۵، میزان پی کمتر از مقدار آلفا بوده یعنی فرض صفر نایستایی سری رد می‌شود. در ادامه نمودارهای توابع همبستگی نگار و همبستگی نگار جزئی برای بخش فصلی سری ارائه شده است (شکل ۹). توجه شود که مقادیر مشخص شده در این نمودارها در انتخاب مولفه‌های مناسب مدل آریما کمک شایانی می‌نمایند.

با بررسی شکل ۸ می‌توان نتیجه گرفت که هر سه مدل پیشنهادی برای مدل‌سازی سری مورد مطالعه قابل قبول می‌باشد چرا که توابع همبستگی نگار و همبستگی نگار جزئی به ازای لگ‌های مختلف در داخل بازه اطمینان مشخص شده واقع گردیده است. در نهایت برای تشخیص بهترین مدل با بررسی همزمان ضریب اطلاع آکائیک AIC و توابع همبستگی نگار و همبستگی نگار جزئی می‌توان مدل (۱،۰،۱) را با کمترین میزان AIC برای مدل‌سازی رفتار سری مناسب تشخیص داد. در ادامه، فرآیند مدل‌سازی بر روی بخش فصلی سری مورد مطالعه انجام گرفت. به منظور مطالعه رفتار بخش فصلی سری لازم است در ابتدا، ایستایی سری مورد آزمون قرار بگیرد. به این صورت که برای کنترل ایستایی از آزمون دی‌کی-فولر بهره گرفته می‌شود. مقدار آماره‌ی این آزمون برابر با $۲/۸۸۳۹$ - بود و میزان پی با فرض صفر نایستا برابر ۰/۲۱۳ تعیین شد که با در نظر گرفتن

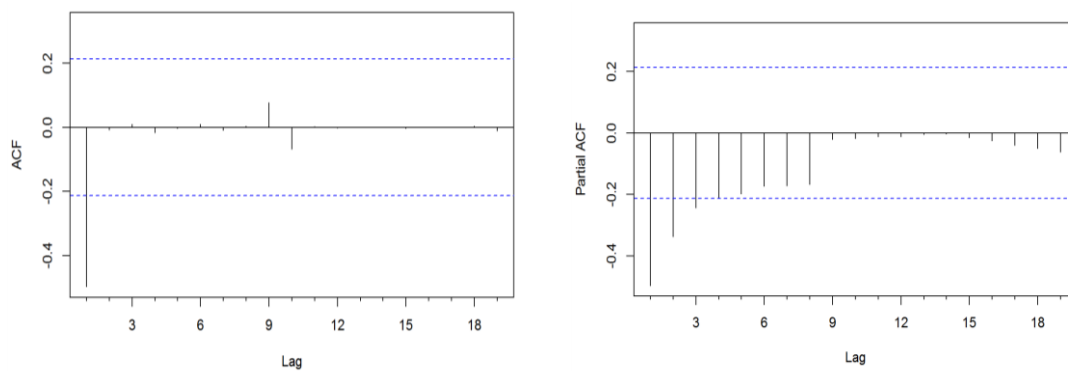


الف

ب

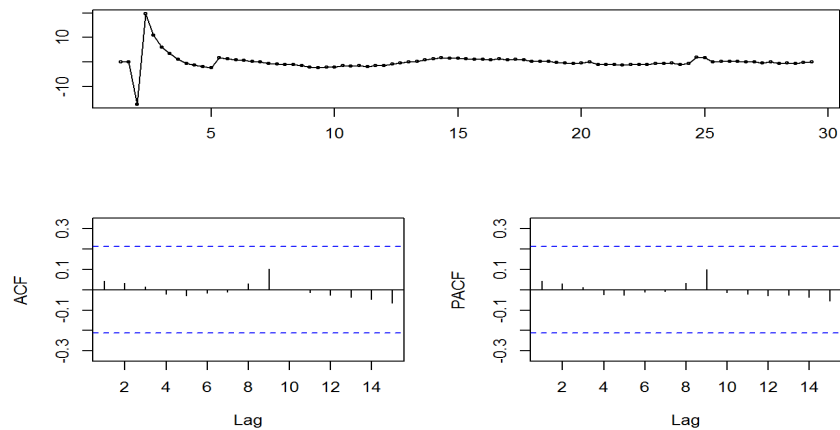
ج

شکل ۸- توابع همبستگی نگار و همبستگی نگار جزئی مدل های پیشنهادی الف: (۱،۰،۰)، ب: (۰،۰،۱) و ج: (۱،۰،۱)



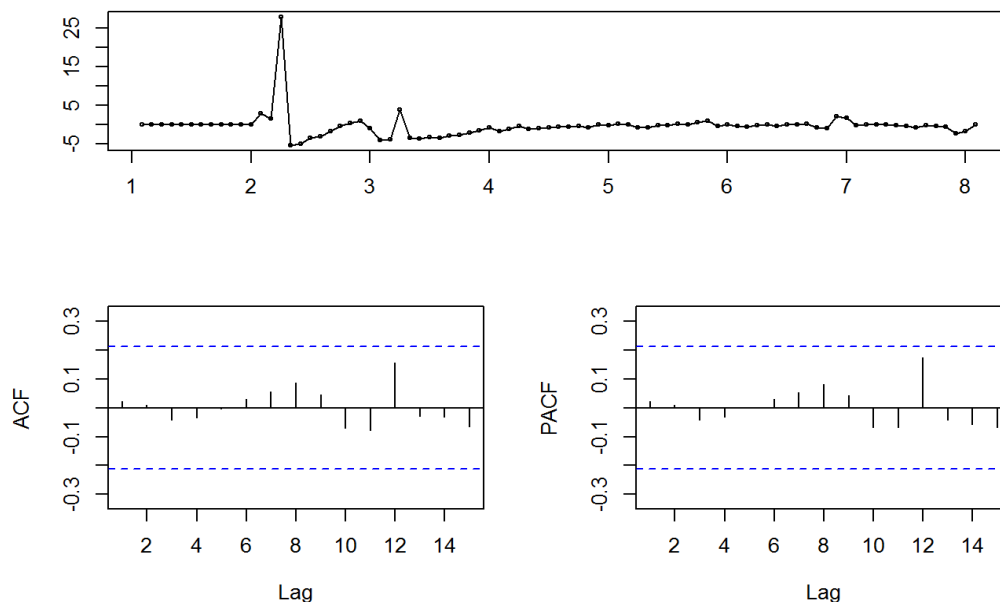
شکل ۹- توابع همبستگی نگار و همبستگی نگار جزئی مربوط به بخش فصلی سری تفاضلی مورد مطالعه از خوشه ی اول

مدل پیشنهادی براساس فرض نایستایی به دست آمده (۰،۱،۲) در نظر گرفته شده و درستی آن با استفاده از نمودارهای از آزمون دی‌کی فولر و نیز با توجه به نمودارهای توابع همبستگی‌نگار و همبستگی‌نگار جزئی مورد آزمون قرار همبستگی‌نگار و همبستگی‌نگار جزئی مطابق با شکل ۹ به فرم گرفته است (شکل ۱۰).

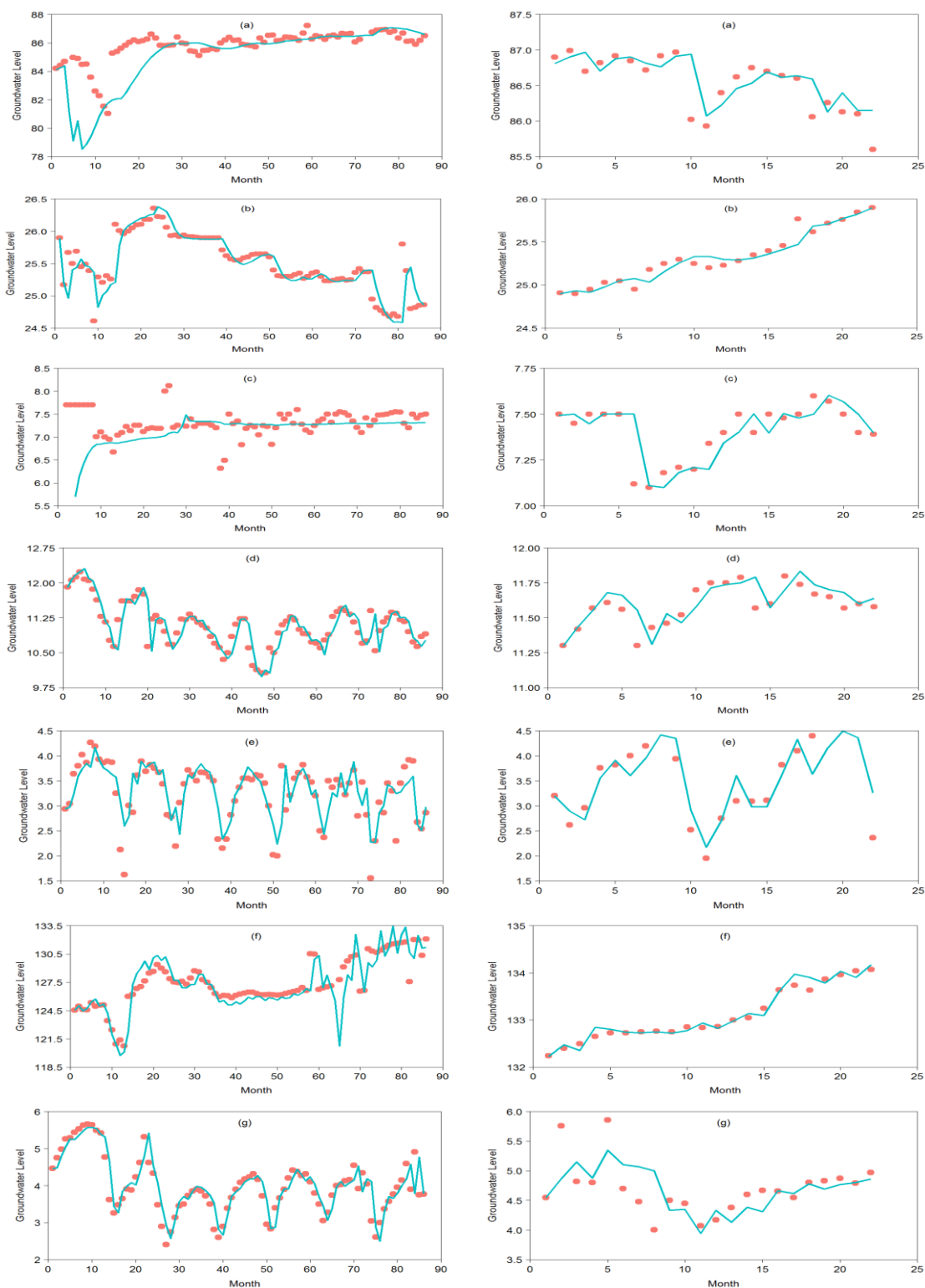


شکل ۱۰- نمودار سری به همراه توابع همبستگی همبستگی‌نگار و همبستگی‌نگار جزئی بخش فصلی سری مورد مطالعه از خوشه‌ی اول به ازای مدل آریما با مشخصات (۰،۱،۲)

با توجه به شکل ۱۰ مقادیر ACF و PACF در بازه‌های اطمینان واقع شده که حاکی از قابل قبول بودن مقادیر تعریفی برای پارامترهای مدل آریما است. لذا مدل انتخابی برای بخش فصلی سری مورد مطالعه به صورت (۰،۱،۲) می‌باشد. در فرآیند مدل‌سازی سری مورد مطالعه از خوشه‌ی اول بخش‌های فصلی و غیرفصلی سری از هم تفکیک شده و پارامترهای مدل آریما به دست آمد که در ادامه این دو بخش با هم تلفیق و پارامترهای مدل آریما به طور همزمان بر روی سری به کار گرفته شده‌اند (شکل ۱۱).



شکل ۱۱- نمودار سری به همراه توابع همبستگی‌نگار و همبستگی‌نگار جزئی سری مربوط به خوشه‌ی اول از حوضه‌ی مورد مطالعه به ازای مدل آریما با مشخصات (۰،۱،۲)(۱،۰،۱)



شکل ۱۲- سری زمانی مقادیر مشاهداتی و محاسباتی نسبت به زمان و نمودار پراکنش برای بهترین ترکیب حاصل از مدل آریمای برای بخش آموزش (سمت چپ) و برای بخش آزمون (سمت راست) داده‌های تراز آب زیر زمینی در دشت تبریز به ترتیب برای هفت خوشه

مشاهده است که این شکل‌ها امکان مقایسه بصری را فراهم می‌نمایند. مشخصات بهترین مدل‌های انتخابی برای سری‌های زمانی منتخب هر کدام از خوشه‌های تعریفی در جدول ۴ ارائه شده است. این مشخصات شامل پارامترهای در نظر گرفته شده، ضرایب هر کدام از اجزاء مدل آریما شامل خودهمبستگی و میانگین متحرک، درجه تفاضل‌گیری، و نیز قسمت ثابت معادله حاصله از مدل آریما برای بخش‌های فصلی و غیرفصلی سری زمانی می‌باشند. نتایج حاصل از مدل‌سازی با استفاده از روش آریما برای سری زمانی منتخب برای هر کدام از خوشه‌های تعریفی مولفه‌ی تراز آب زیرزمینی واقع در دشت تبریز در جدول ۵ اعمال شده است.

کنترل انجام گرفته بر روی شکل ۱۱ حاکی از قابل قبول بودن پارامترهای انتخابی در مدل آریما می‌باشد، به نحوی که تغییرات توابع ACF و PACF در بازه‌ی اطمینان مدل قرار گرفته و نیز از الگوی مناسبی برخوردار است. فرآیند مدل‌سازی اشاره شده برای ایستگاه منتخب از خوشه‌ی اول برای تمام خوشه‌های تعریفی تکرار گردیده و مدل مناسب با هر کدام از ایستگاه‌ها مشخص شده است، که در ادامه تنها به نتایج به دست آمده از بهترین مدل به دست آمده بسنده می‌شود. نمودارهای سری زمانی مقادیر مشاهداتی به همراه مقادیر محاسباتی حاصل شده از بهترین مدل آریما به دست آمده و نیز نمودار پراکنش آنها برای داده‌های بخش آموزش و نیز بخش آزمون در شکل ۱۲ قابل

جدول ۴- مشخصات مدل‌های انتخابی برای سری‌های زمانی مربوط به خوشه‌های هفت‌گانه

خوشه	مدل پیشنهادی	بخش غیرفصلی				بخش فصلی			
		ضرایب مدل				ضرایب مدل			
		AR	d	MA	Intercept	AR	d	MA	Intercept
اول	(۱,۰,۱)	۰/۹۶۹	-	-۰/۸۶۴	۸۵/۰۱۶	-	۰	-۱/۹۲۵	۰/۰۶۲۹
	(۰,۱,۲) _s							۰/۹۵۴	۰/۰۶۴۰
دوم	(۰,۱,۱)	-	۱	-۱	۰/۰۳۱۸	-	۱	-۱	۰/۰۳۱۵-
	(۰,۱,۱) _s								
سوم	(۰,۰,۰) _s (۱,۱,۱)	-۰/۴۹۲	۱	-۱	۰/۱۱۴	-	-	-	۷/۳۱۹۴
					۰/۰۴۳۶				
چهارم	(۰,۱,۳)	-	۱	۱/۲۶۴۳	-	۰/۹۱۰۱	-	-	۱۱/۱۰۵۱
	(۳,۰,۰) _s			۰/۰۶۴۸		-۰/۰۰۴۳			
پنجم	(۴,۰,۴) _s (۴,۱,۴)	-۰/۹۱۶۹	۱	-۰/۵۶۱۴	-	۰/۲۱۲۸	-	۰/۵۰۷۳	۳/۲۳۷۴
		-۰/۶۵۴۱		-۰/۵۹۵۱		۰/۱۶۷۰		-۰/۰۱۲۷	
ششم	(۱,۳,۲) _s (۲,۱,۴)	۰/۲۲۷۳		-۰/۷۱۰۶	-	۰/۵۸۳۰		-۰/۷۲۰۰	
		-۰/۰۸۲۰		۰/۸۸۶۸		-۰/۶۷۳۸		۰/۰۷۵۲	
هفتم	(۱,۰,۳) _s (۱,۱,۱)	-۰/۹۹۰۵	۱	-۰/۴۱۲۶	-	۰/۳۸۲۵	۳	-۱/۸۱۶۰	۳/۹۷۶۳
		-۰/۷۵۲۶		-۰/۸۴۷۵				۰/۸۱۶۰	
				۰/۵۸۶۵					
								۰/۳۶۴۶	
						۰/۶۸۴۵	-	۰/۳۵۸۳	
								۰/۰۰۷۴	

جدول ۵- نتایج ارزیابی عملکرد مدل‌های منتخب مدل‌سازی تراز آب زیرزمینی در دشت تبریز با استفاده از روش آریما

خوشه	الگوی ورودی	خروجی	ساختار مدل	آموزش		آزمون	
				RMSE (m)	R ²	RMSE (m)	R ²
اول	(H) _{t-2} , (H) _{t-1} , ...	H _t	s(۱,۰,۱)(۰,۱,۲)	۳/۳۲۲	۰/۱۷	۰/۲۸۵	۰/۵۰
دوم	(H) _{t-2} , (H) _{t-1} , ...	H _t	s(۰,۱,۱)(۰,۱,۱)	۰/۲۴۴	۰/۷۰	۰/۰۹۱	۰/۹۲
سوم	(H) _{t-2} , (H) _{t-1} , ...	H _t	(۱,۱,۱)(۰,۰,۰) _s	۱/۰۸۹	۰/۲۲	۰/۱۰۴	۰/۵۵
چهارم	(H) _{t-2} , (H) _{t-1} , ...	H _t	(۰,۱,۳)(۳,۰,۰) _s	۰/۲۵۹	۰/۷۴	۰/۰۹۹	۰/۵۹
پنجم	(H) _{t-2} , (H) _{t-1} , ...	H _t	(۴,۱,۴)(۴,۰,۴) _s	۰/۳۶۵	۰/۶۳	۰/۳۵۹	۰/۸۱
ششم	(H) _{t-2} , (H) _{t-1} , ...	H _t	(۲,۱,۴)(۱,۳,۲) _s	۱/۷۰۵	۰/۶۵	۰/۰۸۸	۰/۹۸
هفتم	(H) _{t-2} , (H) _{t-1} , ...	H _t	(۱,۱,۱)(۱,۰,۳) _s	۰/۳۶۷	۰/۷۸	۰/۳۷۳	۰/۳۲

به چشم می‌خورد. وجود روند در سری به کمک فرآیند تفاضل-گیری به درستی برطرف شده است و به عبارتی مدل در خوشه‌ی دوم بسیار بهتر عمل کرده است. مدل آریما در مدل‌سازی سری خوشه‌ی سوم با توجه به وجود مقادیر پرت از دقت پایینی برخوردار است که با افزایش همگنی داده‌ها دقت مدل افزایش می‌یابد.

با بررسی سری منتخب از خوشه‌ی چهارم وجود همزمان نوسان و روند به چشم می‌خورد که روند حاکم برای داده‌های بخش آموزش به صورت نزولی و برای بخش آزمون به صورت صعودی است. برای بخش آموزش مدل با خطای بیشتری نسبت به بخش آزمون مواجه شده ولی در مقابل میزان ضریب تعیین برای بخش آموزش بیشتر است که این حالت می‌تواند به علت کوتاهی طول سری برای بخش آزمون باشد. سری انتخاب شده از خوشه‌ی پنجم برای بخش آموزش نوسانات بیشتر و بدون روند آن‌چنانی کاهشی و افزایشی بود. به عبارت دیگر، روند حاکم به صورت تغییرات در واریانس سری به چشم می‌خورد که سبب نالیستایی سری می‌شود؛ ولی برای بخش آزمون این نوسانات از شدت بیشتر ولی از حیث تعداد از نوسانات کمتری برخوردار است. به نظر نوسانات کمتر سری مورد مطالعه برای بخش آزمون سبب افزایش دقت مدل انتخاب شده برای داده‌های بخش آموزش شده است.

با نگاهی به جدول ۴ و ۵، نتایج بررسی خوشه‌ها نشان می‌دهد که در خوشه‌ی اول وجود مقدار پرت برای داده‌های بخش آموزش سبب شده دقت مدل آریما کاهش داشته باشد ولی دقت مدل برای داده‌های بخش آزمون حاکی از توانایی قابل قبول مدل آریما در مدل‌سازی و درک رفتار سری مورد مطالعه است. در این مطالعه منظور از داده‌های پرت، مقادیری است که سبب ناهموار شدن الگوی رفتاری سری داده‌ها شده است. از آنجایی که هدف از کاربرد مدل آریما، کنترل دقت آن در مدل‌سازی سری‌ها با الگوهای مختلف است، لذا از وجود همچون مقادیر صرف‌نظر نشده است. از آنجایی که معیار انتخاب و ارزیابی مدل به کار رفته بخش آزمون آن است، لذا نتایج نشان می‌دهد مدل آریما تا حدی توانسته برای داده‌های این خوشه قابل قبول باشد.

در خصوص داده‌های اندازه‌گیری نشده، وجود این حالت برای تعداد کم سعی گردید با میانگین‌گیری حول مقادیر موجود سری مورد مطالعه تکمیل گردد ولی در حالت مقادیر بسیار زیاد استفاده از همچون سری خودداری گردید. علاوه بر این طول سری زمانی به نحوی انتخاب شده است تا دارای کمترین میزان داده اندازه‌گیری نشده باشد. خوشه‌ی دوم دارای نوسانات قابل توجه برای داده‌های بخش آموزش بوده و برای بخش آزمون نوسان سری قابل اغماض بوده ولی روند مثبت در بخش آزمون

در تحقیق حاضر، سری زمانی ماهانه ۴۶ ایستگاه سنجش تراز آب زیرزمینی دشت تبریز با استفاده از نمودار دندروگرام و نقشه حرارتی به ۷ خوشه تقسیم‌بندی گردید. نتایج حاکی از آن است که خوشه‌های اول، سوم و هفتم شرایط خوبی ندارند و در بخش آموزش یا آزمون نتایج ضریب تعیین و ریشه میانگین مربعات خطا بسیار ضعیف است. در بقیه خوشه‌ها نتایج قابل قبول است. به طور کلی می‌توان این‌گونه بیان کرد که مدل آریما برای سری‌های زمانی کوتاه و با شرایط یکسان از لحاظ نوسان و روند دقت پایینی داشته که دقت پایین مدل بدون توجه به طول سری با وجود داده پرت همچنان برقرار است.

لذا، پیشنهاد می‌شود قبل از استفاده از مدل آریما بر روی داده‌های مورد مطالعه از پیش‌پردازش بهره گرفته تا مقادیر پرت و سایر مقادیر غیرقابل قبول در سری مورد مطالعه تعیین و برطرف گردد. علاوه بر این، مدل آریما برای داده‌های بسیار همگن توصیه نمی‌شود که منظور از سری بسیار همگن مقادیری است که چندان تغییری در طول زمان ندارد.

استفاده از مدل آریما بر روی ایستگاه‌های مختلف موجود در سطح حوضه تبریز نشان داد، دقت مدل زمانی کاهش می‌یابد که سری‌های مورد بررسی با میزان نوسان بیش از حد مواجه باشند و همچنین مقادیر میانگین و واریانس سری دارای روند باشد که سبب نایستایی سری می‌گردد. بنابراین مدل آریما برای سری‌های با نوسان معمول بدون روند در میانگین و واریانس مناسب هستند.

در نهایت توصیه می‌شود برای افزایش دقت مدل آریما از فرآیند پیش‌پردازش بهره گرفته شود. علاوه بر این، با وجود توانایی قابل توجه مدل آریما در مدل‌سازی سری‌های با روند افزایشی یا کاهشی، ولی این مدل در حین کاربرد بر روی سری‌های رونددار توأم با نوسان شدید دارای دقت پایین‌تری است، لذا برای چنین سری‌هایی لازم است در استفاده از مدل آریما بیشتر دقت نمود.

با بررسی سری مورد مطالعه از خوشه ششم روند صعودی در کل سری برای بخش‌های آموزش و آزمون مشاهده می‌شود که علاوه بر وجود روند، نوسانات در قسمت ابتدایی و همچنین، داده‌های پرت در قسمت انتهایی سری بخش آموزش به چشم می‌خورد که وجود نوسانات در داده‌ها یکی از عوامل مهم در کاهش دقت مدل به شمار می‌آید که نبود این دو عامل در داده‌های بخش آزمون سبب افزایش قابل توجه دقت مدل شده است، به طوری که مقدار ضریب تعیین از مقدار $0/65$ به $0/98$ افزایش، و میزان جذر میانگین مربعات خطا از $1/705$ به مقدار $0/088$ تقلیل یافته است.

خوشه هفتم دارای روند دوگانه کاهشی و افزایشی برای داده‌های بخش آموزش است که این روند توأم با نوسان سری می‌باشد. به طوری که در ابتدای سری بخش آموزش نوسانات از شدت بیشتر نیز همراه است که این افزایش نوسانات در ابتدای سری برای داده‌های بخش آزمون نیز صادق است.

برای داده‌های بخش آزمون، علاوه بر نوسان، روند افزایشی قابل مشاهده می‌باشد. نتایج به دست آمده برای داده‌های بخش آموزش نشان می‌دهد که با وجود روند دوگانه و نوسانات شدید قسمت ابتدایی سری، مدل آریما توانسته تا حد قابل قبولی سری مورد مطالعه را مدل‌سازی نماید ولی برای داده‌های بخش آزمون دقت مدل کاهش یافته که می‌تواند به علت کوتاهی طول سری زمانی این بخش باشد.

نتیجه‌گیری

در مطالعات هیدروژئولوژی معمولاً برای بررسی تغییرات زمانی و مکانی تراز آب زیرزمینی از یک یا چند ایستگاه استفاده می‌شود. در این مطالعه سعی شد با توجه به بزرگ بودن منطقه مورد مطالعه و تعداد چاه‌های بیشتر ابتدا ارتباط سری‌های زمانی با همدیگر بررسی شود و سپس بر اساس ارتباط آن‌ها، سری‌ها به خوشه‌های مناسب تقسیم شوند و یک سری زمانی از هر خوشه به عنوان سری زمانی منتخب مورد بررسی قرار گیرد.

منابع

عساکره، ح.، ۱۳۸۶ الف. بررسی آماری روند دمای سالانه تبریز. مجله اندیشه جغرافیایی، شماره ۱، اول، ۹-۲۱.

عساکره، ح.، ۱۳۸۸. الگوسازی ARIMA برای میانگین سالانه دمای شهر تبریز. فصلنامه تحقیقات جغرافیایی، شماره ۲ (پیاپی ۹۳)، ۱۶۶-۱۸۸.

نیکبخت، ج.، نوری، س.، ۱۳۹۶. خوشه‌بندی شبکه چاهک‌های مشاهده‌ای و پیش‌بینی سطح آب زیرزمینی به کمک شبکه‌های عصبی مصنوعی (مطالعه موردی: دشت مراغه). فصلنامه دانش آب و خاک، سال بیست و هفتم، دوره ۲۷، شماره ۱، ۲۸۱-۲۹۴.

Demir, U., Coleri Ergen, S. 2016. Arima-based time variation model for beneath the chassis UWB channel. Eurasip Journal on wireless communicationa and networking. 178:1-11.

Chun Feng, H., Jun, Q., Fang fang, L. 2017. Methodology for Analyzing and Predicting the Runoff and Sediment into a Reservoir. Water journal. 9:1-16.

Mooi, E., Sarstedt, M. 2011. Cluster analysis. Springer: Berlin Heidelberg.

Wei, W. 2006. Time series analysis: Univariate and Multivariate Methods (2nd Edition). Pearson, USA.

Box, G., Jenkins, G. 1970. Time Series Analysis: Forecasting and Control. (Holden-Day, San Francisco).

Villeneuve, J.P., Banton, O., Lafrance, P. 1990. A probabilistic approach for the groundwater vulnerability to contamination by pesticides: the VULPEST model. Ecological Modelling. 51: 47-58.

اصغری مقدم، ا.، برزگر، ر.، ۱۳۹۳. بررسی عوامل مؤثر بر غلظت بالای آرسنیک در آب زیرزمینی آبخوان‌های دشت تبریز. زمین شناسی مهندسی و محیط زیست، شماره ۲۴، ۱۷۷-۱۹۰.

افروزی، ع.، زارع‌ایبانه، ح.، ۱۳۹۶. مدل‌سازی و پیش‌بینی تراز آب زیرزمینی با کاربرد مدل‌های سری‌زمانی (مطالعه موردی: دشت‌های استان همدان). پژوهش‌نامه مدیریت حوزه آبخیز، سال هشتم، جلد ۸، شماره ۱۵، ۱۱۱-۱۱۲.

باباعلی، ح.، دهقانی، ر.، ۱۳۹۶. مقایسه مدل‌های شبکه عصبی موجک و شبکه عصبی مصنوعی در پیش‌بینی سطح آب زیرزمینی. مجله هیدروژئولوژی، سال دوم، شماره ۲، ۹۶-۱۰۸.

جعفرزاده، ج.، رستم‌زاده، ه.، اسدی، ا.، ۱۳۹۶. مدل‌سازی زمانی تراز آب زیرزمینی با استفاده از روش‌های پایه تحلیل سری‌های زمانی (مطالعه موردی: دشت اردبیل). فصلنامه دانش آب و خاک، سال بیست و هفتم، شماره ۴، ۱۸۵-۱۹۶.

رضایی‌بنفشه، م.، جلالی‌عنصرودی، ط.، حسن‌پور اقدم‌نگلو، م.، ۱۳۹۷. تحلیل و مدل‌سازی تغییرات سطح آب زیرزمینی حوضه آبریز تسوج با استفاده از فرآیند اتو رگرسیون میانگین متحرک. فصلنامه فضای جغرافیایی، سال هفدهم، شماره ۵۷، ۲۷۳-۲۸۷.

شهرکی، ن.، یونسی، م.، طاهری تیزرو، ع.، ۱۳۹۷. مقایسه‌های مدل‌های شبکه عصبی مصنوعی، سری زمانی آریمای و رگرسیون خطی چند متغیره در پیش‌بینی تغییرات سطح آب زیرزمینی. مجله هیدروژئولوژی، سال ۴، شماره ۱، ۱۲۶-۱۳۹.

صادق زاده سادات، م.، ناظمی، ا.م.، صدرالدینی، ع.، ۱۳۹۶. اثرات کیفیت آب‌های سطحی بر کیفیت آب‌های زیرزمینی (مطالعه موردی: دشت تبریز). فصلنامه دانش آب و خاک، جلد بیست و هفتم، شماره ۳، ۲۲۵-۲۳۷.