

مقدمه

متوسط بارندگی در ایران حدود ۲۵۰ میلی متر یعنی کمتر از یک سوم متوسط بارندگی در جهان می باشد که این میزان نیز از توزیع زمانی و مکانی نامناسبی برخوردار است به طوری که ۷۳ درصد از سطح کشور بارندگی کمتر از ۲۵۰ میلی متر، ۲۶ درصد بیشتر از ۲۵۰ میلی متر و سطح باقی مانده یک درصد نیز بالای ۱۰۰۰ میلی متر بارندگی دارد. بارندگی از نظر توزیع زمانی نیز به گونه ای می باشد که اغلب بارندگی های مناطق خشک و نیمه خشک به صورت رگبار در چند روز محدود از سال می بارد و دامنه نوسانات آن در طول یک دوره آماری بسیار بالاست، [۳]. به همین دلیل در چند سال اخیر شاهد بحران کمبود آب و چالش های مربوط به آن در اکثر نقاط کشور از جمله موضوع خشک شدن زاینده رود، خشک شدن دریاچه ارومیه و مطرح شدن مباحثی چون جیره بندی آب در تهران و غیره بوده ایم.

بارندگی به عنوان یکی از مهمترین داده های ورودی به سیستم های هیدرولوژیک به شمار می رود و مطالعه و اندازه گیری آن نیز در اکثر موارد برای هر منطقه یا حوزه آبخیز جهت مطالعات رواناب، آب های زیرزمینی، سیلاب، رسوب و غیره لازم می باشد، [۲]. در دهه های گذشته، رویکردهای مختلفی برای مطالعه و پیش بینی بارندگی فصلی و ماهانه استفاده شده است. این روش ها را می توان به دو گروه کلی تقسیم بندی نمود: یکی روش های تجربی و دیگری روش های دینامیکی. روش های تجربی بر اساس روابط مشاهده ای متغیر وابسته با چندین متغیر مستقل می باشد. روش مرسوم در مدل سازی تجربی فرض فرمی از یک تابع برازش با پارامترهای نامشخص است که باید این پارامترها با روش هایی چون رگرسیون یا سایر روش های بهینه سازی با داده ها تعیین شوند. رویکرد رگرسیونی در این گروه جای می گیرد. رویکردهای دینامیکی نیز با استفاده از روش عددی پیش بینی بارندگی انجام می شوند، [۴، ۱]. سری های زمانی هیدرولوژیک از جمله بارندگی می توانند ترکیبی از اجزاء استوکاستیک (تصادفی) و محض (معین) در نظر گرفته شوند، [۸]. زمانی که جزء استوکاستیک آن رفع شد جزء معین یا محض آن به سهولت می تواند مدل سازی شود، [۱۰]. آنالیز طیف تکین یکی از روش های موجود برای این منظور می باشد که کاربرد آن در هیدرولوژی وارد شده است. به عنوان مثال، لسی و همکاران [۶] در استخراج مؤلفه های مهم از سری های زمانی شاخص نوسانات جنوبی، از آنالیز طیف تکین استفاده کردند. سیواپراگاسام و همکاران

پیش بینی بارندگی ماهانه ایستگاه بارانسنجی پونل با استفاده از رگرسیون خطی و آنالیز طیف تکین

امان محمد کلتله^۱

تاریخ دریافت: ۱۳۹۳/۰۴/۱۱ تاریخ پذیرش: ۱۳۹۳/۰۶/۱۹

چکیده

بارندگی به عنوان یکی از مهمترین داده های ورودی به سیستم های هیدرولوژیک، از پارامترهای مهم مورد استفاده در مدل های هیدرولوژیک می باشد. در دهه های گذشته روش های مختلفی برای تجزیه و تحلیل و پیش بینی بارندگی ماهانه استفاده شده است. روش آماری رگرسیون خطی یکی از روش های مورد استفاده در مدل سازی فرایندهای مختلف هیدرولوژیک از جمله بارندگی می باشد. اخیراً نیز استفاده از آنالیز طیف تکین در تحقیقات منابع آب جهت رفع اجزاء تصادفی موجود در سری های زمانی هیدرولوژیک به طور گسترده ای شایع شده است. هدف اصلی از این تحقیق پیش بینی بارندگی ماهانه با استفاده از ترکیب رگرسیون خطی و آنالیز طیف تکین می باشد. بدین منظور داده های بارندگی ماهانه طی دوره آماری ۲۰ سال (سال آبی ۱۳۶۹-۷۰ تا سال آبی ۱۳۸۸-۸۹) در ایستگاه بارانسنجی پونل (رضوانشهر) در حوزه آبخیز سفارود در استان گیلان مورد استفاده قرار گرفت. برای مقایسه عملکرد مدل پیشنهادی، مدل رگرسیون خطی منفرد (بدون استفاده از آنالیز طیف تکین) نیز استفاده شد. نتایج این مطالعه نشان داد که مدل ترکیب رگرسیون خطی و آنالیز طیف تکین در مقایسه با مدل رگرسیون خطی منفرد عملکرد بسیار مناسبتری را ارائه می نماید.

واژه های کلیدی: بارندگی، پردازش اولیه داده، رگرسیون خطی، آنالیز طیف تکین

۱- نویسنده مسئول و استادیار دانشکده منابع طبیعی دانشگاه گیلان پست الکترونیک: amanmkaltah@gmail.com

[۹] از آنالیز طیف تکین برای پردازش اولیه داده‌های بارندگی و رواناب استفاده کردند. نتایج این محققین نشان داد که استفاده از آنالیز طیف تکین باعث عملکرد بسیار بهتر مدل‌های پیش‌بینی بارندگی و رواناب گردید. وو و همکاران [۱۰] نیز از این روش برای پردازش اولیه داده‌های بارندگی استفاده کردند و نتیجه‌گیری نمودند که پردازش اولیه داده‌ها با این روش، عملکرد مدل‌های مورد استفاده آن‌ها را به‌طور قابل ملاحظه‌ای بهبود بخشید.

هدف اصلی در این تحقیق معرفی و ارائه یک رویکرد جدید برای پیش‌بینی بارندگی ماهانه مبتنی بر ترکیب روش‌های رگرسیون خطی و آنالیز طیف تکین جهت بهبود عملکرد مدل رگرسیون خطی می‌باشد. جهت ارائه عملکرد این رویکرد، بارندگی ماهانه ایستگاه بارانسنجی پونل (رضوانشهر) در حوزه آبخیز سفارود در استان گیلان مورد بررسی قرار گرفت. در این‌جا نقش آنالیز طیف تکین، فیلترکردن داده‌های ورودی قبل از مدل‌سازی و استفاده از این داده‌های فیلترشده یا رفع جزء استوکاستیک شده در مدل‌های رگرسیون خطی جهت ارتقاء در عملکرد این مدل‌ها می‌باشند.

مواد و روش‌ها

منطقه مورد مطالعه و داده‌ها

در این مطالعه داده‌های بارندگی ماهانه ایستگاه بارانسنجی پونل (رضوانشهر) واقع در حوزه آبخیز سفارود در استان گیلان مورد استفاده قرار گرفت. این ایستگاه بارانسنجی توسط آب منطقه‌ای استان گیلان اداره می‌شود و داده‌های بارندگی ماهانه موردنیاز برای انجام این مطالعه نیز از این اداره تهیه شد. برای ایستگاه مورد مطالعه داده‌های بارندگی ماهانه ۲۰ سال که از سال‌آبی ۱۳۶۹-۷۰ شروع و به سال‌آبی ۱۳۸۸-۸۹ ختم می‌شوند برای توسعه مدل‌ها استفاده شد. ۷۰ درصد اول از کل داده‌ها (۱۴ سال اول) برای کالیبراسیون مدل‌ها و ۳۰ درصد مابقی داده‌ها نیز (۶ سال باقی‌مانده) برای اعتباریابی مدل‌ها مورد استفاده قرار گرفت.

رگرسیون خطی

رگرسیون خطی روشی برای مدل‌سازی رابطه بین یک متغیر وابسته و یک یا چند متغیر مستقل می‌باشد و کاملاً روش شناخته شده‌ای بین محققین شاخه‌های مختلف از جمله متخصصین منابع طبیعی می‌باشد. درحالتی که یک متغیر مستقل در مدل‌سازی استفاده شود به آن رگرسیون خطی ساده گفته می‌شود و در حالتی که بیش از یک متغیر مستقل در مدل‌سازی مورد استفاده قرارگیرد به آن رگرسیون خطی چندگانه گفته می‌شود. در این تحقیق رگرسیون خطی هم به صورت ساده و هم به صورت چندگانه مورد استفاده قرار گرفت. با توجه به این‌که بیشتر مدل‌های مورد استفاده در این مطالعه دارای بیش از یک متغیر مستقل بودند شرح مختصری درباره رگرسیون خطی چندگانه در ذیل آورده می‌شود. برای حصول نتیجه مطلوب از رگرسیون خطی چندگانه نه‌تنها

نمونه‌ها باید زیاد و دقیق باشند بلکه برای استفاده از این روش متغیرها دارای توزیع نرمال بوده و تغییر آن‌ها نیز از یک رابطه خطی پیروی نماید، [۵]. همان‌گونه که قبلاً اشاره شد رگرسیون خطی چندگانه در حقیقت برقراری رابطه بین یک سری از متغیرهای مستقل $(x_1, x_2, \dots, x_n)$ با یک متغیر مستقل (y) می‌باشد به‌صورت زیر:

$$y = a_1x_1 + a_2x_2 + \dots + a_nx_n + e \quad (۱)$$

که در رابطه فوق $a_1, a_2, \dots, a_n$ ضرایب رگرسیون نامیده می‌شوند. برای انجام رگرسیون خطی به صورت فوق از نرم‌افزار متلب استفاده شد.

آنالیز طیف تکین

جهت استخراج اطلاعات از سری‌های زمانی کوتاه و نوفه‌ای آنالیز طیف تکین طراحی شد و آنالیز طیف تکین را در چهار گام می‌توان به‌صورت زیر خلاصه نمود، [۷]. گام اول تشکیل بردارهای M بُعدی متوالی از سری و قرارگیری در دنباله $\{\tilde{x}_t\}$ است.

$$\begin{aligned} \tilde{x}_t &= (x_t, x_{t+1}, \dots, x_{t+M-1}) \\ t &= 1, \dots, N', \quad N' = N - M + 1 \end{aligned} \quad (۲)$$

با قراردادن بردارهای بالا در ستون‌های کنار هم، ماتریس $D_{N' \times M}$ تشکیل می‌گردد که به صورت ماتریس هنکل درخواهد آمد. در گام بعدی، ماتریس $M \times M$ کوواریانس C_x از فرمول زیر به‌دست می‌آید:

$$C_x = \frac{1}{N'} D^T D \quad (۳)$$

عناصر ویژه C_x با تفکیک مقادیر تکین تعیین می‌شوند:

$$C_x = U \sum V^T; U^T U = I, V^T V = I \quad (۴)$$

که عناصر قطری ماتریس $\sum = [\text{diag}(\sigma_1, \dots, \sigma_M)]$ همان مقادیر تکین ماتریس D و برابر ریشه دوم مقادیر ویژه C_x می‌باشند. عناصر ویژه ماتریس C_x ، $\{\lambda_k, \rho_k\}; k=1, \dots, M\}$ از فرمول زیر به‌دست می‌آیند:

$$C_x \rho_k = \lambda_k \rho_k \quad (۵)$$

هر مقدار ویژه λ_k واریانس جزئی در جهت ρ_k را مشخص می‌نماید و مجموع مقادیر ویژه نیز واریانس سری زمانی را مشخص می‌کند.

گام بعدی نیز تصویرکردن سری زمانی در جهت هر یک از بردارهای ویژه می‌باشد که با این عمل اجزای اصلی سری زمانی به‌دست می‌آیند. برای هر جزء اصلی داریم:

$$PC_k(t) = \sum_{j=1}^M X(t+j-1) \rho_k(j) \quad (۶)$$

در گام چهارم یا آخر، سری با ترکیب اجزای اصلی متناظر بازسازی می‌شود:

$$R_k(t) = \frac{1}{M_t} \sum_{k \in K} \sum_{j=L_t}^{U_t} P_k(t-j+1) p_k(j) \quad (7)$$

که در آن فاکتور هنجارسازی (M_t) و حد بالا (U_t) و حد پایین (L_t) برای مرکز و شروع و پایان سری متفاوت می‌باشد:

$$(M_t, L_t, U_t) = \begin{cases} (t, 1, t), & \text{when } 1 \leq t \leq M-1 \\ (M, 1, M), & \text{when } M \leq t \leq N' \\ (\min(t, N-t+1), t-N+M, M), & \text{when } N'+1 \leq t \leq N \end{cases} \quad (8)$$

لازم به ذکر است که چنانچه تمامی اجزاء برای بازسازی به کار روند، سری اولیه حاصل خواهد شد.

نتایج

برای پیش‌بینی بارندگی ماهانه (y) در ایستگاه بارانسنجی پونل (رضوانشهر) سه مدل رگرسیون خطی با متغیرهای مستقل مختلف طراحی شد: در مدل اول بارندگی ماه قبل (x_1) به عنوان متغیر مستقل در نظر گرفته شد، در مدل دوم علاوه بر بارندگی ماه قبل، بارندگی دو ماه قبل (x_2) نیز به عنوان متغیرهای مستقل در مدل وارد شدند و در مدل سوم نیز علاوه بر بارندگی‌های ماه قبل و دو ماه قبل، بارندگی‌های سه ماه قبل (x_3) نیز به عنوان متغیرهای مستقل در نظر گرفته شدند. معادلات رگرسیونی حاصل برای سه حالت فوق به صورت زیر می‌باشند:

$$y = 0.701x_1 \quad (9)$$

$$y = 0.518x_1 + 0.262x_2 \quad (10)$$

$$y = 0.489x_1 + 0.199x_2 + 0.117x_3 \quad (11)$$

نتایج رگرسیون خطی برای این سه مدل با توجه به معیارهای

ارزیابی ضریب همبستگی (R) و میانگین مجذور مربعات خطا ($RMSE$) در جدول (۱) نشان داده شده است. همان‌گونه که در جدول (۱) مشاهده می‌شود عملکرد مدل اول ضریب همبستگی ۰/۱۹ و میانگین مجذور مربعات خطا برابر ۹۶/۹۰ را برای دوره کالیبراسیون مدل و مقادیر ضریب همبستگی ۰/۳۲ و میانگین مجذور مربعات خطا ۸۰/۵۳ را برای دوره اعتباریابی نشان می‌دهد. مدل دوم نیز برای دوره کالیبراسیون دارای عملکردی با ضریب همبستگی ۰/۱۵ و میانگین مجذور مربعات خطا برابر ۹۳/۵۹ می‌باشد و برای دوره اعتباریابی نیز این مدل دارای ضریب همبستگی ۰/۳۰ و میانگین مجذور مربعات خطا ۷۷/۱۱ می‌باشد. در نهایت عملکرد مدل سوم نیز با ضریب همبستگی ۰/۱۲ و میانگین مجذور مربعات خطا ۹۲/۹۶ برای دوره کالیبراسیون و با ضریب همبستگی ۰/۲۷ و میانگین مجذور مربعات خطا ۷۶/۹۳ برای دوره اعتباریابی می‌باشد. آنچه که قابل توجه می‌باشد این است که وارد کردن متغیرهای مستقل بیشتر به مدل‌های رگرسیون خطی خطای عملکرد این مدل‌ها را تا اندازه‌ای کاهش داده ولی ضریب همبستگی بین مقادیر مشاهده شده و مقادیر واقعی را نیز کم نموده است. به عبارتی این موضوع نشان‌دهنده این است که این متغیرهای مستقل اضافه شده به مدل‌های رگرسیون خطی فوق، دارای اطلاعات ارزشمندی نبوده و یا این‌که به مدل‌های رگرسیون خطی اختلال تحمیل می‌نمایند و در مجموع این مدل‌ها با توجه به عملکرد بسیار ضعیف آن‌ها، قابل قبول جهت پیش‌بینی بارندگی ماهانه در ایستگاه بارانسنجی مورد مطالعه نمی‌باشند. آنالیز طیف تکین یکی از روش‌های موجود برای پردازش اولیه داده‌ها یا فیلترکردن داده‌ها می‌باشد. در این مطالعه، برای فیلترکردن

جدول ۱- نتایج رگرسیون خطی و ترکیب رگرسیون خطی و آنالیز طیف تکین در پیش‌بینی بارندگی ماهانه ایستگاه بارانسنجی پونل (رضوانشهر)

اعتباریابی	کالیبراسیون		مدل
	<i>RMSE</i>	<i>R</i>	
رگرسیون خطی			
۰/۳۲	۸۰/۵۳	۰/۱۹	۹۶/۹۰
۰/۳۰	۷۷/۱۱	۰/۱۵	۹۳/۵۹
۰/۲۷	۷۶/۹۳	۰/۱۲	۹۲/۹۶
ترکیب رگرسیون خطی و آنالیز طیف تکین			
۰/۵۷	۶۰/۴۰	۰/۵۵	۶۹/۸۸
۰/۷۲	۵۲/۵۶	۰/۷۰	۶۱/۲۰
۰/۸۰	۴۳/۲۶	۰/۷۶	۵۳/۶۸

رگرسیون خطی و آنالیز طیف تکین) در نظر گرفته شدند. اگر نتایج حاصل از آنالیز طیف تکین برای متغیرهای مستقل اصلی x_1, x_2, x_3 با x_1, x_2, x_3 نشان داده شوند، معادلات رگرسیونی حاصل برای سه حالت مدل‌سازی به صورت زیر می باشند:

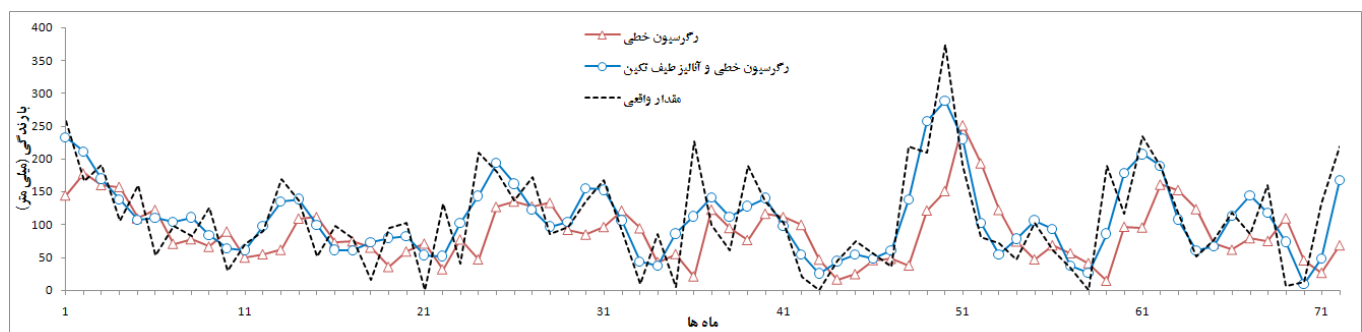
$$y = 0.629 x_1' \quad (12)$$

$$y = 1.137 x_1' - 0.542 x_2' \quad (13)$$

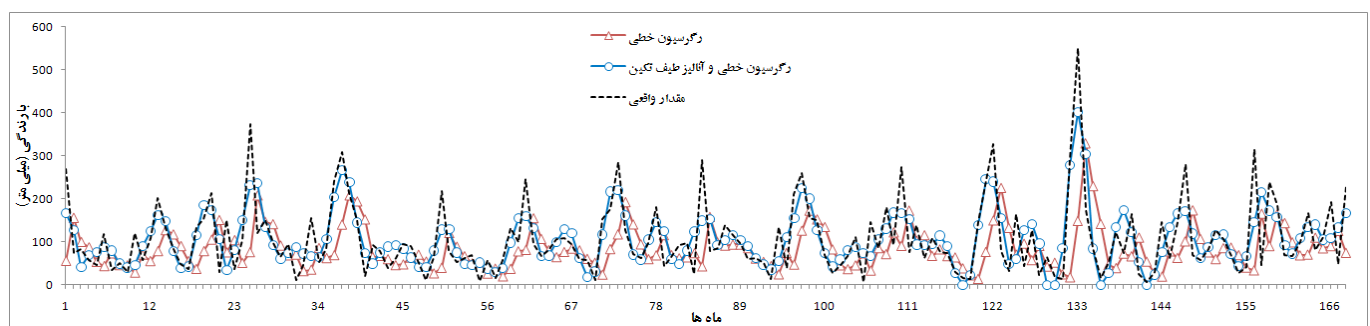
$$y = 1.558 x_1' - 1.524 x_2' + 0.626 x_3' \quad (14)$$

نتایج ترکیب رگرسیون خطی و آنالیز طیف تکین برای سه مدل فوق با توجه به معیارهای ارزیابی ضریب همبستگی (R) و میانگین مجذور مربعات خطا ($RMSE$) در جدول (۱) نشان داده شده است. همان‌گونه که در جدول (۱) نیز مشاهده می‌شود عملکرد مدل اول ضریب همبستگی ۰/۵۵ و میانگین مجذور مربعات خطا برابر ۶۹/۸۸ را برای دوره کالیبراسیون مدل و مقادیر ضریب همبستگی ۰/۵۷ و میانگین مجذور مربعات خطا ۶۰/۴۰ را برای دوره اعتباریابی نشان می‌دهد. مدل دوم نیز دارای عملکرد برای دوره کالیبراسیون با ضریب همبستگی ۰/۷۰ و میانگین مجذور مربعات خطا ۶۱/۲۰ و برای دوره اعتباریابی دارای ضریب همبستگی ۰/۷۲ و میانگین

داده‌های ورودی به مدل‌ها یعنی بارندگی، این داده‌ها با کمک آنالیز طیف تکین به جزءهایی تجزیه می‌شوند. در تجزیه به کمک آنالیز طیف تکین دو پارامتر بایستی توسط کاربر تعیین شوند که یکی از آن‌ها تعیین M و دیگری انتخاب گروه‌های/مؤلفه‌های اصلی از تعداد M گروه موجود می‌باشد. در این مطالعه M برابر ۱۰ انتخاب شد و این مقدار برای انجام مدل‌سازی بارندگی ماهانه مناسب تشخیص داده شد. در این تحقیق تعداد اجزاء اصلی نیز بر اساس سهم هر جزء در واریانس داده مورد نظر تعیین گردید. هر چقدر این سهم‌ها اندک باشند نشان‌دهنده این است که اجزاء مورد نظر مرتبط با جزء تصادفی داده‌ها می‌باشند و در نتیجه اطلاعات با ارزشی را در مدل‌سازی فراهم نمی‌کنند و اضافه نمودن آن‌ها به مدل نه تنها عملکرد مدل‌ها را بهبود نخواهد داد بلکه با وارد نمودن اطلاعات نادرست به مدل یا ایجاد اختلال در مدل باعث عملکرد نامناسب مدل‌ها نیز می‌گردند. در مطالعه حاضر از ۱۰ جزء موجود، جزء اول ۶۶/۷۴ درصد، جزء دوم ۶/۸۱ درصد، جزء سوم ۶/۱۵ درصد و جزء چهارم نیز ۴/۹۰ درصد از واریانس داده را به خود اختصاص می‌دادند، لذا ۴ جزء اول به عنوان مؤلفه‌های اصلی جهت بازسازی سری‌های زمانی انتخاب شدند. سپس سری‌های بازسازی شده حاصل از آنالیز طیف تکین مربوط به هریک از متغیرهای مستقل (به جای متغیرهای مستقل اصلی)، به عنوان ورودی‌های مدل‌های رگرسیون خطی (ترکیب



شکل ۱ - مقایسه بارندگی ماهانه واقعی و نتایج رگرسیون خطی و ترکیب رگرسیون خطی و آنالیز طیف تکین برای دوره کالیبراسیون



شکل ۲ - مقایسه بارندگی ماهانه واقعی و نتایج رگرسیون خطی و ترکیب رگرسیون خطی و آنالیز طیف تکین برای دوره اعتباریابی

حرارت ماهانه و غیره وجود دارد. علاوه بر این، این مطالعه فقط داده‌های مربوط به یک ایستگاه بارانسنجی را مورد مطالعه قرار داد و مطالعات دقیق‌تری با استفاده از داده‌های ایستگاه‌های بارانسنجی از مناطق مختلف جهت تقویت این نتیجه‌گیری‌ها ضروری می‌باشند.

منابع

- 1- Ahmed, I. Menon, S. Nikitha, K.B. 2013. Rainfall prediction using multiple regression technique. International Journal of Applied Engineering Research 8 (19): 2267-2270.
- 2- Gholizade, M.H. Darand, M. 1389. Monthly precipitation forecasting using artificial neural networks. Physical Geography Research 71: 51-63.
- 3- Hajibigloo, M. Ghezelsolfloo, A.A. Alimirzaei, H. 1392. Discussion and forecast monthly average rainfall techniques using SARIMA (case study: pluviometry station Babaaman Bojnourd). Journal of Irrigation Science and Engineering 36 (3): 41-54.
- 4- He, X. Guan, H. Zhang, X. Simmons, C.T. 2014. A wavelet-based multiple linear regression model for forecasting monthly rainfall. International Journal of Climatology 34: 1898-1912.
- 5- Jalali Lichaei, H. Nabi-Bidhendi, M. 1385. Comparison between multiple linear regression and artificial neural networks for porosity and permeability estimation. Journal of Geosciences 61: 140-149.
- 6- Lisi, F. Nicolis, O. Sandri, M. 1995. Combining singular-spectrum analysis and neural networks for time series forecasting. Neural Processing Letters 2 (4): 6-10.
- 7- Mohammadi Gheimasi, H. Siahkoohi, H.R. Lucas, C. 1389. Improvement of temporal resolution of seismic data using singular spectrum analysis and autoregressive methods. Journal of the Earth and Space Physics 36(3): 87-99.
- 8- Salas, J.D. Delleur, J.W. Yevjevich, V. Lane, W.L. 1985. Applied modeling of hydrologic time series. Water Resources Publications, Littleton, Colorado.
- 9- Sivapragasam, C. Liong, S.Y. Pasha, M.F.K. 2001. Rainfall and runoff forecasting with SSA-SVM approach. Journal of Hydroinformatics 3 (7): 141-152.
- 10- Wu, C.L. Chau, K.W. Fan, C. 2010. Prediction of rainfall time series using modular artificial neural networks coupled with data-preprocessing techniques. Journal of Hydrology 389: 146-167.

مجذور مربعات خطا ۵۲/۵۶ می‌باشد. در نهایت عملکرد مدل سوّم ترکیب رگرسیون خطی و آنالیز طیف تکین نیز با ضریب همبستگی ۰/۷۶ و میانگین مجذور مربعات خطا ۵۳/۶۸ برای دوره کالیبراسیون و با ضریب همبستگی ۰/۸۰ و میانگین مجذور مربعات خطا ۴۳/۲۶ برای دوره اعتباریابی می‌باشد. مشاهده می‌شود که عملکرد مدل‌های ترکیب رگرسیون خطی و آنالیز طیف تکین در مقایسه با عملکرد مدل‌های رگرسیون خطی منفرد به میزان قابل‌توجهی افزایش یافته است. این موضوع نشان‌دهنده این است که آنالیز طیف تکین به‌خوبی داده‌ها را فیلتر نموده و داده‌های با اطلاعات ارزشمندی را به مدل‌های رگرسیونی فراهم نموده است. نتیجه این تحقیق با نتایج سایر محققین از جمله لیبی و همکاران [۶]، سیواپراگاسام و همکاران [۹] و نیز وو و همکاران [۱۰] همخوانی دارد از این جهت که مدل‌های مورد استفاده آن‌ها نیز پس از ترکیب با آنالیز طیف تکین، عملکرد بسیار مناسب‌تری را ارائه نمودند. شکل‌های (۱) و (۲) مقادیر واقعی بارندگی ماهانه و مقادیر پیش‌بینی شده از ترکیب رگرسیون خطی و آنالیز طیف تکین (مدل ترکیب رگرسیون خطی و آنالیز طیف تکین با بهترین عملکرد یعنی مدل سوّم) را به‌ترتیب برای دوره کالیبراسیون و اعتباریابی نشان می‌دهند. برای مقایسه، نتایج مدل سوّم رگرسیون خطی منفرد نیز در شکل‌های (۱) و (۲) آورده شده است. شکل‌های ارائه شده نیز گویای عملکرد بسیار مناسب‌تر مدل ترکیب رگرسیون خطی و آنالیز طیف تکین در مقایسه با مدل رگرسیون خطی منفرد - هم برای دوره کالیبراسیون و هم برای دوره اعتباریابی - می‌باشند.

نتیجه‌گیری

این مقاله یک مدل ترکیبی با عنوان رگرسیون خطی و آنالیز طیف تکین را برای مدل‌سازی سری‌های زمانی بارندگی ماهانه معرفی می‌نماید. در این مدل، آنالیز طیف تکین برای پردازش اولیه داده‌ها و رفع جزء استوکاستیک موجود در این سری‌های زمانی مورد استفاده قرار می‌گیرد. این مدل پیشنهادی برای پیش‌بینی بارندگی در ایستگاه بارانسنجی پونل (رضوانشهر) به‌کار برده شد. سری‌های زمانی رفع جزء استوکاستیک شده یا فیلتر شده توسط آنالیز طیف تکین به‌عنوان متغیرهای مستقل مدل رگرسیون خطی و سری‌های زمانی واقعی نیز به‌عنوان متغیر وابسته مورد استفاده قرار گرفتند. ضرایب مدل رگرسیون خطی با استفاده از داده‌های ۱۴ سال اول بارندگی ماهانه کالیبره شدند و باقی‌مانده داده‌های بارندگی ماهانه نیز برای اعتباریابی مدل رگرسیون خطی به‌کار رفت. نتایج این تجزیه و تحلیل کارایی بهتر مدل ترکیبی رگرسیون خطی و آنالیز طیف تکین را نسبت به مدل منفرد رگرسیون خطی نشان داد.

آنچه لازم است اشاره شود این است که در این تحقیق فقط داده‌های بارندگی ماهانه به‌عنوان متغیر مستقل در مدل مدّنظر قرار گرفت، لذا امکان بهبود در عملکرد مدل‌ها با وارد کردن متغیرهای مستقل دیگر از جمله حداکثر درجه حرارت ماهانه و حداقل درجه

*Abstract*

Monthly Precipitation Forecasting in Ponel Raingauge Station Using Linear Regression and Singular Spectrum Analysis

A.M. Kalte¹

Recived: 2014.07.02 Accepted: 2014.09.10

Rainfall/ precipitation, as one of the most important inputs to hydrological systems, is one of the most significant parameters in many hydrological models. In the recent decades, different types of forecasting methods are employed for forecasting and analyzing monthly precipitation rates. Linear regression is one of the methods are being used for this purpose. Recently, the use of singular spectrum analysis in water resources studies for removing random components of hydrological series has extensively increased. The main objective of this study is to investigate the use of linear regression coupled with singular spectrum analysis for monthly precipitation forecasting. The monthly data of Ponel raingauge station which span the period from 1991 to 2010 (i.e. 20 years) were used to develop the proposed model. The proposed model was compared with regular linear regression and the results indicated the superiority of combined linear regression and singular spectrum analysis models.

Keywords: *Precipitation, Data preprocessing, Linear regression, Singular spectrum analysis*

1. Assistant Prof, Faculty of Natural Resources, University of Guilan, Corresponding Author Email: amanmkalteh@gmail.com