

مقایسه کارآیی گشتاورهای معمولی، خطی و بیشینه درست‌نمائی در تحلیل منطقه‌ای سیل (مقاله پژوهشی)

۱- مؤده محمدی خوشنوی*، دکتری آبخیزداری دانشکده منابع طبیعی، دانشگاه یزد، یزد، ایران

m.mohammadi@stu.yazd.ac.ir

۲- محمدرضا اختصاصی، استاد دانشکده منابع طبیعی، دانشگاه یزد، یزد، ایران

۳- علی طالبی، استاد دانشکده منابع طبیعی، دانشگاه یزد، یزد، ایران

دریافت: ۱۳۹۹/۱۱/۲۰

پذیرش: ۱۳۹۹/۱۲/۱۸

چکیده

در تخمین احتمال وقوع سیلاب، انتخاب توزیع احتمالاتی مناسب برای برازش به داده‌های سیلاب ضروری است. روش‌های مختلفی برای برآورد مؤلفه‌های توابع توزیع احتمالاتی وجود دارد که در پژوهش حاضر عملکرد سه روش گشتاورهای خطی، گشتاورهای معمولی و بیشینه درست‌نمایی مورد بررسی قرار گرفته است. بدین منظور، تعداد ۱۶ ایستگاه آب‌سنجی در حوزه‌های گاوخونی، ابرکوه-سیرجان و دشت یزد-اردکان انتخاب و دبی حداکثر لحظه‌ای آن‌ها استخراج شد. با استفاده از روش خوشه‌بندی Ward ایستگاه‌های منطقه مورد مطالعه به دو منطقه همگن تقسیم و همگنی نواحی بدست آمده، با معیار ناهمگنی هاسکینگ و والیس بررسی شد. نتایج حاکی از ناهمگنی است. از این رو با حذف و انتقال ایستگاه‌های موجود، منطقه‌ای همگن از نظر فراوانی سیل به دست آمد. سپس مؤلفه‌های ۱۰ توزیع احتمالاتی در منطقه همگن، با سه روش گشتاورهای خطی، گشتاورهای معمولی و بیشینه درست‌نمایی به دست آمد. سه آزمون Z^{DIST} ، کولموگروف - اسمیرنوف و χ^2 برای آزمون نکویی برازش توزیع‌ها به داده‌های سیل به کار رفت. سپس عملکرد توزیع‌های احتمالاتی و سه روش برآورد مؤلفه‌ها، براساس آماره‌های $RRMSE$ و $RMSE$ مقایسه شد. بر اساس نتایج، توزیع پیرسون نوع ۳ و روش گشتاورهای خطی به ترتیب با داشتن کمترین و بیشترین مقدار $RMSE$ و R ، به عنوان توزیع منطقه‌ای انتخاب شد. نتایج آماره $RRMSE$ نیز نشان داد که روش گشتاورهای خطی در دوره بازگشت‌های ۱۰، ۲۵، ۵۰ و ۱۰۰ ساله برآوردهای دقیق‌تری را ارائه می‌دهد. در حالی که روش گشتاورهای معمولی در دوره بازگشت‌های ۲ و ۵ برازش بهتری نشان می‌دهد.

واژگان کلیدی: فراوانی سیل، همگنی، دبی حداکثر لحظه‌ای، توزیع منطقه‌ای.

مقدمه

به طور کلی از مهم‌ترین روش‌های تحلیل منطقه‌ای سیلاب می‌توان به روش رگرسیون چند متغیره [۳۸، ۳۵]، منطقه اثر [۷]، شاخص سیل [۹] و همبستگی کانونی [۸، ۴۱، ۳۷] اشاره کرد. در همه روش‌های تحلیل منطقه‌ای سیلاب ابتدا باید مهم‌ترین عوامل مؤثر بر دبی سیلاب تعیین شود. سپس به دسته‌بندی حوزه‌ها به یک یا چند منطقه همگن اقدام نمود که از نظر ویژگی‌های تولید سیل و به طوری کلی خصوصیات آب‌شناختی شباهت بیشتری

کمیود آمار و اطلاعات آب‌سنجی یکی از مشکلات اساسی در برآورد دبی سیلابی با دوره بازگشت‌های مختلف است. بنابراین، استفاده از روش‌هایی که بتوان طی آن‌ها با حداقل داده، برآورد قابل قبولی را ارائه نمود، ضروری به نظر می‌رسد. روش تحلیل منطقه‌ای یکی از این روش‌ها است. در طول دهه‌های اخیر تلاش‌های زیادی برای تحلیل و توسعه روش‌های منطقه‌ای صورت گرفته است، با این وجود، در حال حاضر هیچ روشی به صورت جهانی مورد قبول واقع نشده است.

1- Region Of Influence (ROI)

2- Index-Flood

3- Canonical Correlation Analysis (CCA)

معمولی برای دبی سیلاب در هند نشان داد که توزیع مقادیر حدی نوع اول و روش گشتاورهای خطی مناسب‌ترین روش در برآورد مؤلفه‌های توزیع احتمالاتی نقطه‌ای در منطقه مورد مطالعه است [۴۶]. این روش همچنین نتایج نزدیک‌تری به داده‌های مشاهداتی در دوره بازگشت‌های مورد نظر دارد [۵]. در پژوهشی دیگر دو توزیع نمایی و گاما با استفاده از روش بیشینه درست‌نمایی مناسب‌ترین توزیع برای برازش به داده‌های دبی طرح در هند گزارش شدند [۴۷].

در مناطق خشک و نیمه خشک ایران، بهترین توزیع احتمالاتی برای برازش به داده دبی حداکثر لحظه‌ای تابع توزیع پیرسون نوع ۳ و روش گشتاورهای خطی مشاهده شد [۲۳]. همچنین روش گشتاورهای خطی در ایستگاه‌های هواشناسی اصفهان، بوشهر، تهران، جاسک و مشهد روش مناسب‌تری برای تحلیل فراوانی بارندگی و برازش به داده‌های بارش در مقایسه با روش بیشینه درست‌نمایی و روش غیر پارامتری است. توزیع‌های برازش داده شده با روش گشتاور خطی برای تمام ایستگاه‌ها بر اساس کمترین متوسط انحراف نسبی، بهترین برازش را در مقایسه با روش‌های بیشینه درست‌نمایی و روش گشتاور معمولی دارند [۱۱].

مقایسه کارایی روش گشتاورهای خطی با روش‌های مرسوم برآورد دبی حداکثر لحظه‌ای سیلاب در حوزه آبخیز شیخ عثمان از توابع استان آذربایجان غربی نشان داد که روش گشتاورهای خطی و توزیع گمبل بهترین برازش را به داده‌های سیلاب دارند و دقت بالاتری نسبت به روش‌های گشتاور معمولی و بیشینه درست‌نمایی از خود نشان می‌دهند [۱۳]. در مطالعات تحلیل منطقه‌ای سیلاب در سراسر دنیا توزیع‌های مختلفی از جمله توزیع‌های مقادیر حدی تعمیم یافته [۳۲، ۴۰، ۴۸، ۳۵، ۲۴]، پیرسون نوع ۳ [۲۵، ۲۸، ۲۲]، لوگ پیرسون نوع ۳ [۲۶، ۲۱] و لجستیک تعمیم یافته [۱، ۴۴، ۱۵، ۳۳] مناسب‌ترین توزیع‌ها برای برازش به داده‌های دبی سیلاب گزارش شدند.

نتایج به دست آمده از پژوهش‌های انجام شده در زمینه تحلیل منطقه‌ای سیل نشان‌دهنده این نکته است که قطعیت در انتخاب یک تابع توزیع خاص برای یک منطقه یا کشور وجود ندارد. به همین منظور برای رسیدن

دارند. در ادامه انتخاب توزیع احتمالاتی مناسب برای هر منطقه همگن، تعیین مؤلفه‌های توزیع و در آخر تخمین سیلاب طرح، مراحل انجام گام به گام این روش‌ها است. دقت روش‌های تحلیل منطقه‌ای وابسته به میزان همگنی ایستگاه‌ها، روش‌های برآورد مؤلفه‌ها و تعیین توزیع احتمالاتی مورد استفاده است [۳۶]. برای تخمین احتمال وقوع و دوره برگشت سیلاب، انتخاب توزیع احتمالاتی مناسب برای برازش به داده‌های سیلاب ضروری است. روش‌های مختلفی برای برآورد مؤلفه‌های توزیع احتمالاتی وجود دارد که از جمله می‌توان به روش حداکثر درست‌نمایی^۱ [۴۷]، روش کمترین مربعات^۲ [۴۲]، روش گرافیکی^۳ [۳۶]، روش گشتاورهای معمولی^۴ [۲۳] و روش گشتاورهای خطی^۵ [۲۰] اشاره کرد.

بررسی‌های زیادی برای ارزیابی کیفیت روش‌های محاسبه مؤلفه‌های توزیع احتمالاتی انجام شده است. نتایج مطالعات نشان می‌دهد که با افزایش اندازه نمونه، کیفیت برآورد تابع توزیع احتمال بهتر شده و خطای محاسباتی برآورد مؤلفه‌های توزیع کمتر می‌گردد [۱۲]. برتری روش گشتاورهای خطی نسبت به سایر روش‌ها در عدم حساسیت این روش به مقادیر پرت در داده‌ها است که باعث می‌شود روش گشتاورهای خطی نتایج مطمئن‌تری برای یک تابع توزیع احتمال در مناطق با تعداد ایستگاه کم و تعداد سال آماری کوتاه نشان دهد [۴۳].

روش گشتاورهای معمولی و بعد از آن روش بیشینه درست‌نمایی تطابق بهتری با داده‌های دبی حداکثر سالانه در ترکیه برای برآورد مؤلفه‌های توزیع سه پارامتره پارتو تعمیم‌یافته نشان دادند [۴۲]. بررسی دو روش گشتاورهای خطی و بیشینه درست‌نمایی در برآورد دبی حداکثر سالانه سه تابع توزیع احتمالاتی پارتو، لجستیک و مقادیر حدی تعمیم یافته در پاکستان نشان داد که روش گشتاورهای خطی نسبت به روش بیشینه درست‌نمایی مناسب‌تر بوده و روش بیشینه درست‌نمایی تنها در ۵ درصد سایت‌ها برآورد مناسبی داشته است [۴]. همچنین، تخمین نقطه‌ای مؤلفه‌های تابع توزیع با استفاده از روش گشتاور خطی و

1- Maximum Likelihood Method (MLM)

2- Minimum Square Method

3- Graphical Method

4- Method of Moments (MOM)

5- Linear Moment Method (LMM)

است. در بخش سوم بهترین تابع توزیع منطقه‌ای با استفاده از آماره‌های Z^{DIST} ، دیاگرام گشتاورهای خطی، آماره کولموگروف-اسمیرنوف و آزمون χ^2 تعیین و عدم قطعیت آماره Z^{DIST} بررسی می‌شود. بخش چهارم، شامل ارزیابی عملکرد هر روش، توسط آماره‌های RMSE، RRMSE و R می‌شود.

آماده‌سازی داده و بررسی همگنی

در بخش اول، تعداد ۱۶ ایستگاه آب‌سنجی در محدوده مورد مطالعه انتخاب و ۱۷ عامل اقلیمی (حداکثر بارندگی ۲۴ ساعته با دوره بازگشت ۲ سال، درجه حرارت متوسط سالانه، بارندگی متوسط سالانه)، فیزیکی (زمان تمرکز، طول آبراهه اصلی، درصد شیب آبراهه اصلی، درصد شیب متوسط حوزه، مساحت، قطر دایره معادل، ارتفاع متوسط، ضریب گراولیوس و تراکم زهکشی)، هیدرولوژیکی (دبی متوسط سالانه، دبی حداکثر لحظه‌ای با دوره بازگشت ۲سال) و کاربری اراضی (کشاورزی، جنگل و مرتع) استخراج گردید.

هدف از تحلیل منطقه‌ای سیلاب برآورد دبی سیلاب و تناوب وقوع آن در یک منطقه است. وابسته بودن مکانی داده‌های سیلاب، خطای برآورد چندک‌های سیلاب را افزایش می‌دهد هرچند میزان این افزایش کمتر از اریبی است که به‌خاطر ناهمگنی ناحیه به چندک‌های سیلاب اعمال می‌شود [۳۹]. در همه روش‌های تحلیل منطقه‌ای سیلاب ابتدا باید مهم‌ترین عوامل موثر بر سیلاب را مشخص و سپس به تقسیم حوزه‌آبخیز به دو یا چند منطقه همگن اقدام نمود که از نظر خصوصیات تولید سیل و به‌طور کلی خصوصیات هیدرولوژیکی شباهت بیشتری دارند. از آن‌جا که تعداد مؤلفه‌ها بسیار زیاد هستند، به منظور کاهش تعداد آن‌ها از روش تجزیه و تحلیل عاملی استفاده شد تا مهمترین آن‌ها شناسایی شوند.

با انجام تجزیه و تحلیل عاملی مشخص شد که اطلاعات با اختصاص دادن حول ۴ محور می‌توانند خلاصه شوند (جدول ۲). دلیل بر ۴ محور، بالا بودن ریشه پنهان ماتریس همبستگی بیش از یک است. بنابراین از بین ۱۷ مؤلفه مذکور، چهار مؤلفه اصلی (مساحت، طول آبراهه اصلی، متوسط بارندگی سالانه و حداکثر بارش ۲۴ ساعته)

به نتایج قابل اعتماد بایستی از آزمون‌هایی مانند کولموگروف - اسمیرنوف^۱، خی‌دو^۲، دیاگرام گشتاورهای خطی و آماره Z^{DIST} بهترین تابع توزیع تعیین و نتایج آن مورد بررسی قرار گیرد [۴۷،۴]. به همین منظور در پژوهش حاضر علاوه بر استفاده از آماره‌های مختلف، عدم قطعیت آماره Z^{DIST} برای تعیین بهترین تابع توزیع منطقه‌ای نیز مورد بررسی قرار گرفت. همچنین پژوهش حاضر به مقایسه عملکرد سه روش گشتاورهای خطی، گشتاور معمولی و روش بیشینه درست‌نمایی در برآورد مؤلفه‌ها و چندک‌های توزیع‌های احتمالاتی در حوزه‌های گاوخونی، ابرکوه-سیرجان و دشت- یزد اردکان به منظور تعیین بهترین تابع توزیع برای برازش به دبی‌های حداکثر لحظه‌ای در منطقه مورد مطالعه پرداخته است.

مواد و روش‌ها

منطقه مورد مطالعه

حوزه‌های مورد مطالعه در پژوهش حاضر شامل، حوزه آبریز گاوخونی، حوزه کویر ابرکوه- سیرجان و دشت یزد- اردکان به عنوان حوزه‌های آبریز فلات مرکزی ایران هستند. وسعت حوزه‌ی گاوخونی و حوزه‌ی کویر ابرکوه- سیرجان به ترتیب $۴۱۵۵۲/۳$ و $۵۷۱۲۵/۳$ کیلومتر مربع است. حوزه دشت یزد-اردکان، یکی از زیر حوزه‌های حوزه‌ی کویر سیاه با مساحت $۱۵۹۵/۰۷$ کیلومترمربع در بخش شمالی استان یزد واقع شده است. حوزه گاوخونی دارای اقلیم فراخشک تا بسیار مرطوب، حوزه ابرکوه- سیرجان دارای اقلیمی با ویژگی بارندگی کم، دامنه نوسان حرارتی کم و بیش زیاده خشک و حوزه دشت یزد-اردکان دارای اقلیم فراخشک است (شکل ۱).

پژوهش حاضر از چهار بخش اصلی تشکیل شده است. بخش اول، آماده سازی داده‌ها، تعیین همگنی اولیه ناحیه مورد مطالعه با روش خوشه‌بندی وارد و بررسی همگنی اولیه با استفاده از دو آزمون ناهمنوائی^۳ و ناهمگنی^۴ هاسکینگ و والیس [۳۸] است. بخش دوم، برآورد مؤلفه‌های توزیع‌های احتمالاتی با استفاده از سه روش LMM، MOM و MLM و محاسبه چندک‌های سیلاب

6- Kolmogorov- Smirnov test (KS)

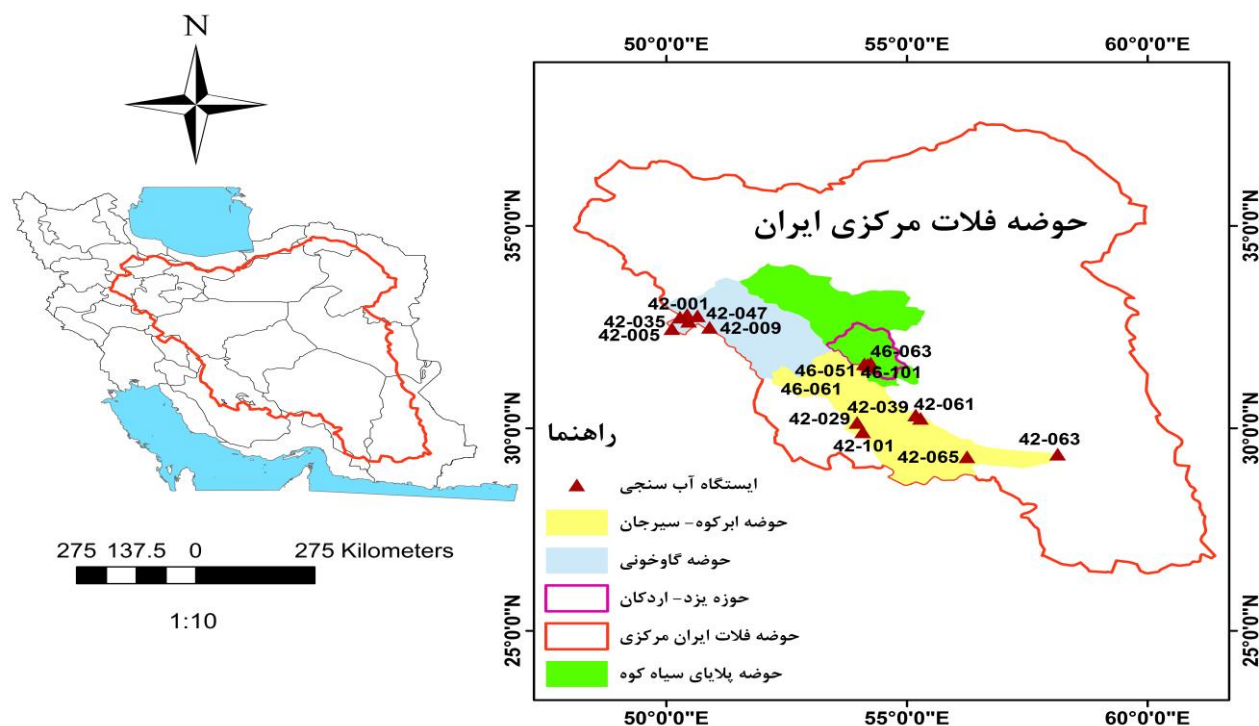
7- Chi-square test (X^2)

1- Discordancy, Di

2- Heterogeneity, H

دوم شامل ایستگاه‌های قلعه‌شاهرخ، اسکندری، دره، مندرجان، ده‌بالا، طزرجان، فخرآباد، رمرزج، بندپایین، مدوار، آبدرمیان، سوچ و کرخنگان است (شکل ۲). سپس همگنی مناطق به دست آمده از خوشه‌بندی وارد با دو آزمون ناهمنوائی و ناهمگنی هاسکینگ و والیس مورد بررسی قرار گرفت.

که در مجموع ۸۶/۷ درصد از واریانس داده‌ها را تبیین می‌کنند و به عنوان عامل‌های اصلی مورد بررسی قرار گرفت. سپس همگنی اولیه با استفاده از نتایج تجزیه و تحلیل عاملی، به روش خوشه‌بندی وارد انجام و ۱۶ ایستگاه آب‌سنجی واقع در ناحیه مورد مطالعه به دو منطقه همگن تقسیم شد. منطقه همگن اول شامل ایستگاه‌های چهلگرد، دره‌در و زمان‌خان و منطقه همگن



شکل ۱- موقعیت ایستگاه‌های آب‌سنجی ناحیه مورد مطالعه

$$D_i = \frac{1}{\sqrt{p}} (u_i - \bar{u})^T s^{-1} (u_i - \bar{u}) \quad (4)$$

در واقع رابطه (۱) ترانهاده ماتریس متشکل از $lcvi$ ، τ_3 و τ_4 به ترتیب ضریب تغییرات^۱، ضریب چولگی^۲ خطی و ضریب کشیدگی گشتاورهای خطی^۳ است. برای محاسبه میزان ناسازگاری، \bar{u} میانگین‌های گروه S ماتریس کوواریانس نمونه به صورت رابطه (۲) و (۳) محاسبه شد. در رابطه‌ها n تعداد کل ایستگاه‌ها، $(u_i - \bar{u})^T$ ترانهاده حاصل از اختلاف ماتریس ایستگاه مورد نظر و ماتریس منطقه‌ای است. اگر آماره D_i مربوط

آزمون ناهمنوائی هاسکینگ و والیس

آزمون D_i با استفاده از رابطه (۴) برای حذف ایستگاه‌های ناسازگار که در نمودار گشتاورهای خطی در محدوده فضای دو بعدی τ_3 و τ_4 قرار ندارند، انجام گرفت. فرض شد که تابع u_i (رابطه ۱) برداری دربرگیرنده نسبت‌های گشتاورهای خطی برای ایستگاه i بود. بنابراین:

$$u_i = [lcvi, \tau_3, \tau_4]^T \quad (1)$$

$$\bar{u} = \frac{1}{n} \sum_{i=1}^n u_i \quad (2)$$

$$s = \frac{1}{n} \sum_{i=1}^n (u_i - \bar{u})(u_i - \bar{u})^T \quad (3)$$

1- Linear- coefficient variation, L-cv
2- Linear- coefficient skewness, L-cs
3- Linear- coefficient kurtosis, L-c

سپس تغییر پذیری گشتاورهای خطی ناحیه واقعی با گشتاورهای خطی سری شبیه سازی شده مقایسه شد.

$$F(x) = \left\{ -h \left[1 - k(x - \xi) / \alpha \right]^{1/k} \right\}^{1/h} \quad (5)$$

که در آن $F(X)$ تابع توزیع تجمعی برای متغیر X ، ξ مؤلفه موقعیت، α مؤلفه مقیاس و h و k مؤلفه‌های شکل هستند.

به یک ایستگاه بیش از ۳ باشد، داده مشاهداتی آن ایستگاه ناهمگن و ناسازگار با دیگر ایستگاه‌های منطقه در نظر گرفته می‌شود و بایستی از سایر مراحل مطالعاتی کنار گذاشته شود [۲۰، ۲].

آزمون ناهمگنی هاسکینگ

در آزمون ناهمگنی گشتاورهای خطی یک توزیع چهار پارامتری کاپا (رابطه ۵) به سری داده‌های مشاهداتی منطقه برازش داده شد و به وسیله شبیه سازی عددی، یک دسته ۵۰۰ تایی از داده‌های منطقه‌ای تولید گردید.

جدول ۱- موقعیت جغرافیایی و ویژگی‌های ایستگاه‌های آب‌سنجی منتخب

کد ایستگاه	نام ایستگاه	نام رودخانه	مشخصات جغرافیایی		مساحت (km ²)	طول آبراهه اصلی (km)	حداکثر بارندگی ۲۴ ساعته با دوره بازگشت ۲ سال	ضریب گراولوس
			عرض	طول				
۰۰۱-۴۲	چهلگرد	تونل اول کوه‌رنگ	۲۸-۳۲	۰۷-۵۰	۴۳۰	۳۲/۷۵	۹۴/۸۲	۰/۹۷۱
۰۰۳-۴۲	قلعه شاهرخ	زاینده رود	۳۹-۳۲	۲۸-۵۰	۱۹۱۷/۵	۶۶/۵	۵۷/۵۲	۱/۱۶۷
۰۰۵-۴۲	اسکندری	پلاسجان	۴۹-۳۲	۲۶-۵۰	۱۶۰۰	۴۱/۵	۵۸/۸۳	۰/۹۹۹
۰۰۹-۴۲	زمانخان	زاینده رود	۳۰-۳۲	۵۴-۵۰	۵۷۶۶	۱۳۰/۵	۴۳/۸۲	۱/۳۳۸
۰۲۹-۴۲	بندپایین	رودخانه اعظم	۵۵-۲۹	۰۴-۵۴	۱۰۱۱/۵	۳۵	۴۳/۱۶	۱/۰۶
۰۳۵-۴۲	دره در	تونل دوم	۴۵-۳۲	۱۷-۵۰	۳۲	۹/۲۵	۹۲/۵۶	۰/۷۱
۰۳۹-۴۲	تاجو مدوار	سردرودخانه	۲۰-۳۰	۱۱-۵۵	۴۷/۷	۱۴	۴۷/۲۴	۱/۲۴۵
۰۴۷-۴۲	مندرجان	سمندگان	۴۷-۳۲	۳۹-۵۰	۱۱۹/۵	۱۶/۵	۴۶/۰۶	۱/۳۱۳
۰۶۱-۴۲	آبدرمیان	آبدرمیان	۱۵-۳۰	۱۷-۵۵	۴۶/۳	۱۲/۷۵	۴۶/۲۳	۱/۳۲۶
۰۶۳-۴۲	رمزج	اسطور	۲۲-۲۹	۰۸-۵۸	۱۲۰	۷	۲۸/۳۶	۰/۴۸۳
۰۶۵-۴۲	سوچ	سوچ	۱۸-۲۹	۱۵-۵۶	۲۱/۰۲	۲۵/۲۵	۳۱/۹۷	۱/۲۸۴
۱۰۱-۴۲	کرخنگان	کرخنگان	۰۹-۳۰	۵۸-۵۳	۹/۱	۵/۵	۵۴/۸۷	۱/۲۱۳
۰۵۱-۴۶	باغستان	ده بالا	۳۶-۳۱	۰۷-۵۴	۶۶/۲۲	۱۲/۷	۵۴/۲	۱/۲۳
۰۶۱-۴۶	طزرجان	طزرجان	۳۵-۳۱	۱۱-۵۴	۲۹/۸۵	۸/۲	۴۰/۸۲	۰/۳۷۶
۰۶۳-۴۶	دره	منشاد	۳۲-۳۱	۱۳-۵۴	۵۲/۵۷	۱۶	۵۶/۵	۱/۴۵
۱۰۱-۴۶	فخرآباد	فخرآباد	۳۷-۳۱	۱۵-۵۴	۲۰۵/۲	۲۴/۵	۲۵/۱۴	۱/۲

جدول ۲- مقادیر ریشه پنهان ماتریس همبستگی

مؤلفه‌های اصلی	مقادیر ویژه	واریانس %	واریانس تجمعی %
۱	۵/۱۳	۳۳/۷	۳۳/۷
۲	۴/۰۳	۲۷	۶۰/۷
۳	۲/۴۵	۱۶/۴	۷۷/۱
۴	۱/۰۸۱	۹/۶	۸۶/۷

ضریب چولگی خطی و ضریب کشیدگی خطی به کار گرفته شد [۳۸]. هر یک از آماره‌های H دارای شکل عمومی بر اساس رابطه ۶ است [۲۰]:

در بررسی تغییرپذیری سه آماره ناهمگنی H_1 ، H_2 و H_3 مختلف استفاده شد. آماره H_1 برای بررسی ضریب تغییرات خطی، آماره H_2 برای بررسی ضریب تغییرات خطی و ضریب چولگی خطی و آماره H_3 برای بررسی

$$V_{\tau} = \sum_{i=1}^N \left(n_i \left[(\tau_{\tau_i} - \bar{\tau}_{\tau})^2 + (\tau_{\tau_i} - \bar{\tau}_{\tau})^2 \right] \right)^{1/2} / \sum_{i=1}^N n_i \quad (9)$$

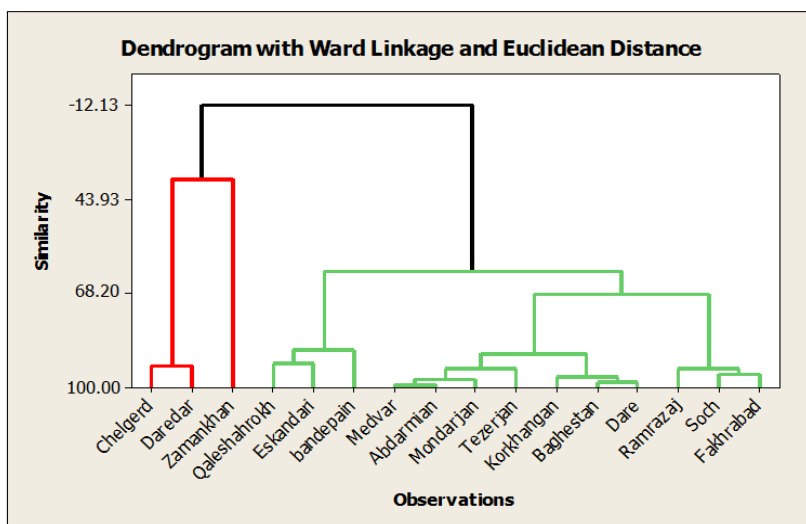
$$H = (V_{obs} - \mu_V) / \sigma_V \quad (6)$$

که در آن‌ها: n تعداد ایستگاه‌ها، Lcv_i ، \bar{Lcv} ، τ_{τ_i} ، $\bar{\tau}_{\tau}$ ، τ_{τ_i} و $\bar{\tau}_{\tau}$ به ترتیب ضریب تغییرات در ایستگاه مورد نظر، ضریب تغییرات منطقه‌ای، ضریب چولگی خطی در ایستگاه مورد نظر، ضریب چولگی میانگین منطقه‌ای، ضریب کشیدگی خطی در ایستگاه مورد نظر و ضریب کشیدگی میانگین منطقه‌ای است. در صورتی که H کوچکتر از یک باشد منطقه همگن و اگر بین ۱ و ۲ باشد، منطقه نسبتاً ناهمگن و اگر بزرگتر از ۲ باشد، منطقه ناهمگن است. بنابراین، یک مجموعه از ایستگاه‌های مورد بررسی باید دارای پارامتر H کمتر از ۱ باشند تا به عنوان منطقه‌ی همگن در نظر گرفته شود [۲۰].

که در آن: μ_V و σ_V به ترتیب میانگین و انحراف معیار مقادیر شبیه سازی شده متغیر مورد نظر و متغیر V_{obs} مقادیر محاسبه شده متغیر مورد نظر با استفاده از داده‌های منطقه‌ای و مبتنی بر آماره V است. برای هر یک از آماره‌های H به ترتیب V_1 ، V_2 ، V_3 به صورت رابطه‌های (۷) تا (۹) محاسبه شدند [۲۹].

$$V_1 = \sum_{i=1}^N \left(n_i (Lcv_i - \bar{Lcv})^2 \right) / \sum_{i=1}^N n_i \quad (7)$$

$$V_2 = \sum_{i=1}^N \left(n_i \left[(Lcv_i - \bar{Lcv})^2 + (\tau_{\tau_i} - \bar{\tau}_{\tau})^2 \right] \right)^{1/2} / \sum_{i=1}^N n_i \quad (8)$$



شکل ۲- نمودار روش خوشه‌بندی Ward برای ایستگاه‌های ناحیه مورد مطالعه

و بهترین توزیع یعنی توزیعی دارای $|Z^{DIST}| < 1.64$ انتخاب شد.

$$|Z^{DIST}| = \frac{(\tau_{\tau}^{DIST} - \tau_{\tau}^R + \beta_{\tau})}{\sigma_{\tau}} \quad (10)$$

τ_{τ}^{DIST} ضریب کشیدگی میانگین منطقه همگن و τ_{τ}^R ضریب کشیدگی مربوط به هر تابع توزیع محاسبه شده است. β_{τ} و σ_{τ} به ترتیب بیانگر مقدار اریب و انحراف

تعیین بهترین تابع توزیع احتمالاتی

در پژوهش حاضر با استفاده از آماره‌های Z^{DIST} ، کولموگروف-اسمیرنوف، دیاگرام گشتاورهای خطی و آزمون χ^2 [۴۷] بهترین تابع توزیع منطقه‌ای تعیین شد.

آماره Z^{DIST} و بررسی عدم قطعیت آن

به منظور انتخاب بهترین تابع توزیع منطقه‌ای مقدار آماره Z^{DIST} (رابطه ۱۰) برای توزیع‌های احتمالاتی محاسبه

شبیه‌سازی بررسی می‌شود [۱۸]. دقت تخمین چندک های منطقه‌ای هر تابع توزیع بر اساس معیار مجذور میانگین مربعات خطای نسبی RMSE به صورت منطقه ای، با استفاده از نسبت گشتاورهای خطی حاصل از ۱۰۰۰۰ بار تکرار شبیه‌سازی بررسی شد. بر اساس این معیار تابع توزیعی دارای بهترین برآزش خواهد بود که کمترین مقدار RMSE نسبی منطقه‌ای را برای چندک های تخمینی تابع توزیع دارا باشد.

$$R^R(F) = \sum_{i=1}^N R_i(F) / N \quad (14)$$

$$R_i(F) = \left(M^{-1} \sum_{m=1}^M \left\{ \hat{q}_i^{(m)}(F) - q_i(F) / q_i(F) \right\}^2 \right)^{1/2} \quad (15)$$

که در آن‌ها: R^R مجذور میانگین مربعات خطای نسبی منطقه‌ای، $R_i(F)$ مجذور میانگین مربعات خطای نسبی در ایستگاه i ، چندک بدون بعد در ایستگاه i ، $\hat{q}_i^{(m)}(F)$ کوانتیل بدون بعد در m امین تکرار شبیه سازی، M تعداد دفعات شبیه سازی و N تعداد ایستگاه‌های داخل منطقه‌ای همگن است [۲۰].

آماره‌ی دو

در آزمون خی دو با درجه آزادی $K-1$ ، ابتدا داده‌ها به k فاصله رده‌ای تقسیم شد. این آماره به صورت رابطه (۱۶) تعریف و به طور مجانبی توزیع شده است:

$$\chi^2 = \frac{k}{n} \sum_{j=1}^k O_j^2 - n \quad (16)$$

که در آن: O_j مقدار عددی مشاهده شده در فاصله رده‌ای، n اندازه نمونه، k تعداد اختیاری رده‌هایی است که داده‌های مشاهده شده به آن تقسیم شدند. فواصل رده‌ای با استفاده از معکوس تابع توزیع مطابق با مقادیر مختلف احتمال F مشابه با چندک‌های برآورد شده محاسبه شد. برای آماره خی دو لازم است که χ^2 محاسبه شده از توزیع نمونه کمتر از مقدار χ^2 جدول در سطح اعتماد لازم باشد [۱۶].

معیار مربوط به τ_f^R هستند که از طریق شبیه‌سازی (رابطه ۱۱) محاسبه شد:

$$\beta_f = N^{-1} \sum_{m=1}^{N_{sim}} (\tau_f^m - \tau_f^R) \quad (11)$$

τ_f^m ضریب کشیدگی میانگین منطقه‌ای برای m امین منطقه شبیه‌سازی شده و N_{sim} تعداد منطقه شبیه‌سازی شده است که با استفاده از توزیع کاپا تولید گردید [۲]. برای ارزیابی آزمون نکویی Z^{DIST} در هنگام برآزش توابع توزیع مختلف به یک منطقه‌ی همگن، چندک‌های هر تابع توزیع برای تمام ایستگاه‌ها و چندک منطقه‌ای هر تابع با استفاده از داده‌های واقعی و شبیه‌سازی تخمین زده می‌شود. چندک منطقه‌ای حاصل از شبیه‌سازی با تقسیم چندک‌های ایستگاه‌های داخل گروه همگن به میانگین هر ایستگاه (μ) مقیاس می‌شود. در این حالت یک عدد بدون بعد برای چندک منطقه‌ای خواهیم داشت.

$$q_i(F) = Q_i(F) / \mu_i \quad (12)$$

که در آن: $q_i(F)$ چندک بدون بعد در ایستگاه i ، $Q_i(F)$ چندک ایستگاه i و μ_i میانگین ایستگاه است. توسط مقادیر $q_i(F)$ در هر ایستگاه چندک منطقه‌ای مربوط به هر تابع توزیع قابل محاسبه است [۱۹]. برای ارزیابی معیار Z^{DIST} و انتخاب بهترین تابع توزیع، تعیین یک محدوده‌ی خطا در سطح ۹۰ درصد برای مقادیر چندک منطقه‌ای مربوط به هر تابع توزیع است (رابطه ۱۳). در این محدوده خطا حد پایین و بالا به صورت زیر بیان می‌شود.

$$\begin{cases} U_{0.05}(F) = 1.05 \hat{q}(F) / q(F) \\ L_{0.05}(F) = 0.95 \hat{q}(F) / q(F) \end{cases} \quad (13)$$

که در آن: $\hat{q}(F)$ مقدار چندک منطقه‌ای حاصل از شبیه‌سازی و $q(F)$ چندک منطقه‌ای واقعی مربوط به هر تابع توزیع است. همچنین دقت تخمین چندک منطقه‌ای هر تابع توزیع بر اساس معیار مجذور میانگین مربعات خطای نسبی (RMSE) به صورت منطقه‌ای، با استفاده از نسبت گشتاورهای خطی حاصل از ۱۰۰۰۰ بار تکرار

آماره کولموگروف-اسمیرنوف

در این آماره انحراف‌های تابع نمونه $F_N(x)$ از توزیع فرضی پیوسته کاملاً معلوم $F_0(x)$ محاسبه شد. آماره این آزمون D_N به صورت رابطه (۱۷) است.

$$D_N = \max |F_N(x) - F_0(x)| \quad (17)$$

مقادیر $F_N(x)$ به صورت N_j/N برآورد شد که در آن N_j تعداد نمونه تجمعی مشاهدات در هر رده z می‌باشد. $F_0(x)$ نیز برابر $1/k, 2/k, \dots$ است که در آن k تعداد فواصل رده‌ای است. حدود رده مشابه آنچه که در آزمون خی‌دو به دست می‌آید. مقدار D_N باید کمتر از مقدار D_N موجود در جداول آماری در سطح اعتماد مورد نظر برای توزیع مورد قبول باشد [۴۷].

نمودار گشتاورهای خطی

نمودار گشتاورهای خطی بر اساس روابط بین نسبت گشتاورهای خطی است. دیاگرام ضریب تغییرات خطی در مقابل ضریب چولگی خطی، روشی کیفی در بررسی همگنی منطقه است. نمودار ضریب چولگی در مقابل ضریب کشیدگی خطی روشی کیفی در تعیین بهترین تابع توزیع احتمالاتی است. به طوری که در این روش منطقه‌ای همگن و توزیعی به عنوان توزیع منتخب انتخاب می‌شود که متوسط و پراکنش داده‌ها در اطراف آن به صورت خوشه‌ای و متمرکز گسترش یافتند [۴۳].

برآورد مؤلفه‌ها و چندک‌های توزیع احتمالاتی

مؤلفه‌ها و چندک‌های تابع توزیع‌های احتمالاتی به سه روش گشتاورهای معمولی، گشتاورهای خطی و بیشینه درست‌نمایی برآورد شد.

روش گشتاورهای معمولی

قدیمی‌ترین و گسترده‌ترین روش شناخته شده برای برازش توزیع‌های فراوانی مشاهداتی روش گشتاورهاست. در روش گشتاورها معادله کلی برای محاسبه گشتاور مرتبه r حول مبدأ یک توزیع، $f(x)$ به کار می‌رود (رابطه

(۱۸) یا معادله متناظر برای گشتاورهای مرکزی توزیع مورد استفاده قرار می‌گیرد (رابطه ۱۹)

$$\mu'_r = \int_{-\infty}^{+\infty} x^r f(x) dx \quad r = 0, 1, 2, \dots \quad (18)$$

$$\mu_r = \int_{-\infty}^{+\infty} (x - \mu)^r \cdot f(x) dx \quad r = 0, 1, 2, \dots \quad (19)$$

که در آن μ'_r ، μ_r امین گشتاور جامعه حول مبدا مختصات و μ_r گشتاور حول میانگین برای جامعه است. r مرتبه گشتاور، $f(x)$ تابع چگالی احتمال توزیع آماری و μ میانگین داده‌ها است. روش گشتاورها، رابطه بین گشتاورها و مؤلفه‌های توزیع را مشخص می‌سازد. عموماً روش گشتاورها بوسیله ایجاد معادله گشتاورهای تئوریک قابل کاربرد است [۱۶].

روش گشتاورهای خطی

روش گشتاورهای خطی از جمله روش‌هایی است که برای انتخاب توزیع‌های مناسب آماری، تخمین مؤلفه‌های توزیع و منطقه‌ای کردن داده‌های هواشناسی و هیدرولوژیکی به کار می‌رود. گشتاورهای خطی به صورت ترکیب خطی گشتاورهای وزن دار احتمالی [۱۴] تعریف می‌شوند. گشتاورهای وزن دار احتمالی را می‌توان به صورت زیر نمایش داد [۴۶، ۱۸]:

$$\beta_r = E\{X[F(X)]^r\} \quad (20)$$

که در آن: $F(X)$ تابع توزیع تجمعی برای متغیر X و $X(F)$ معکوس تابع توزیع تجمعی از X محاسبه شده برای احتمال F و $r=0, 1, 2, \dots$ یک عدد صحیح غیر منفی است که برای $r=0$ ، β_r برابر با میانگین توزیع است. گشتاورهای خطی با استفاده از رابطه‌های (۲۱) تا (۲۴) برآورد شد.

$$\lambda_1 = \beta_1 \quad (21)$$

$$\lambda_2 = 2\beta_1 - \beta_0 \quad (22)$$

مشتق جزئی گرفتن از لگاریتم L نسبت به هر یک از مؤلفه‌ها و مساوی صفر قرار دادن نتایج با صفر، تعدادی معادله n مجهولی به دست می‌آید که برای حل معادلات فوق از روش‌های مبتنی بر تکرار نظیر نیوتن-رافسون، شبه نیوتن و بردار گرادیان استفاده می‌شود [۳].

ارزیابی عملکرد سه روش LMM، MOM و MLM
برای بررسی عملکرد سه روش LMM، MOM و MLM در برآورد مؤلفه‌ها، ریشه میانگین مربعات خطا^۱، ضریب تعیین^۲ و ریشه میانگین مربعات خطای نسبی^۳ چندک‌های به دست آمده از هر یک از توزیع‌های احتمالاتی به دست آمد.

$$RMSE = \sqrt{\left(\frac{1}{n} \sum_{i=1}^n [(Q_i) - (P_i)]^2\right)} \quad (26)$$

$$R = \sqrt{1 - \frac{\sum_{i=1}^n (Q_i - P_i)^2}{\sum_{i=1}^n (Q_i - \bar{Q})^2}} \quad (27)$$

$$RRMSE = \sqrt{\left(\frac{1}{n} \sum_{i=1}^n [(Q_i) - (P_i)] / (Q_i)\right)^2} \quad (28)$$

که در آن‌ها: P_i مقدار حداکثر دبی لحظه‌ای پیش‌بینی شده از تحلیل فراوانی منطقه‌ای، Q_i مقدار حداکثر دبی لحظه‌ای مشاهداتی در دوره بازگشت‌های مختلف با روش تعیین احتمال وقوع ویبول، n تعداد داده‌ها و \bar{Q} میانگین مقادیر حداکثر دبی لحظه‌ای مشاهداتی هستند.

نتایج و بحث

بررسی همگنی منطقه

شرط اصلی تحلیل منطقه‌ای، همگن بودن منطقه‌ی مورد مطالعه است. به همین منظور آزمون ناهمنوائی و ناهمگنی هاسکینگ و والیس برای دو منطقه همگن به دست آمده از خوشه‌بندی Ward اجرا شد. در ابتدا نسبت گشتاورهای خطی در هریک از ایستگاه‌ها با استفاده از برنامه نوشته شده در محیط Matlab، محاسبه گردید

$$\lambda_3 = 6\beta_2 - 6\beta_1 + \beta_0 \quad (23)$$

$$\lambda_4 = 20\beta_3 - 30\beta_2 + 12\beta_1 - \beta_0 \quad (24)$$

که در آن‌ها: β تا β_3 گشتاورهای وزنی احتمال، λ_1 یا میانگین گشتاورهای خطی، معیاری از تمایل به مرکز و λ_4 انحراف معیار گشتاورهای خطی معیاری از پراکندگی، λ_3 گشتاور خطی سوم داده‌ها و λ_4 گشتاور خطی چهارم داده‌ها است. نسبت‌های گشتاورهای خطی از نسبت λ_1 به λ_2 ، λ_3 به λ_4 ، ضریب تغییرات گشتاورهای خطی، از نسبت λ_3 به λ_4 ، ضریب چولگی گشتاور خطی و از نسبت λ_4 به λ_2 ، ضریب کشیدگی گشتاورهای خطی برآورد شدند [۴۵، ۱۸].

روش بیشینه درست‌نمایی

از میان روش‌های برآورد مؤلفه‌های توابع توزیع آماری، روش بیشینه درست‌نمایی به نظر می‌رسد که یکی از کارآمدترین روش‌ها باشد، چون کمترین واریانس نمونه‌گیری از مؤلفه‌های برآورد شده را ارائه می‌کند و از این‌رو چندک‌های برآورد شده آن قابل اعتماد هستند [۱۶]. اصل بیشینه درست‌نمایی بیان می‌کند که برای یک توزیع یا تابع چگالی احتمال $p(\alpha, \beta, \dots)$ که در آن α و β مؤلفه‌های توزیع بوده و باید برآورد شوند احتمال به دست آوردن مقداری از X مانند x_i متناسب با $p(x_i; \alpha, \beta, \dots)$ و احتمال توأم، L ، برای به دست آوردن نمونه‌ای از n مقدار X_1, X_2, \dots, X_n که متناسب است با حاصلضرب متغیرهای زیر که آن را تابع درست‌نمایی می‌گویند [۴۷]:

$$L = \prod_{i=1}^n p(x_i; \alpha, \beta, \dots) \quad (25)$$

تابع درست‌نمایی به دلیل سهولت در محاسبه، اغلب به صورت لگاریتمی نشان داده شده و با توجه به نوع توزیع احتمال متغیر وابسته نیز به شکل‌های مختلفی بیان می‌شود [۱۷]. روش بیشینه درست‌نمایی مؤلفه‌های α و β را طوری برآورد می‌کند که L حداکثر شود. این حداکثر از مشتقات نسبی L نسبت به هر یک از مؤلفه‌ها که مساوی صفر قرار گرفته باشند به دست می‌آیند. با

1- Root Mean Square Error (RMSE)

2- Coefficient of Determination (R)

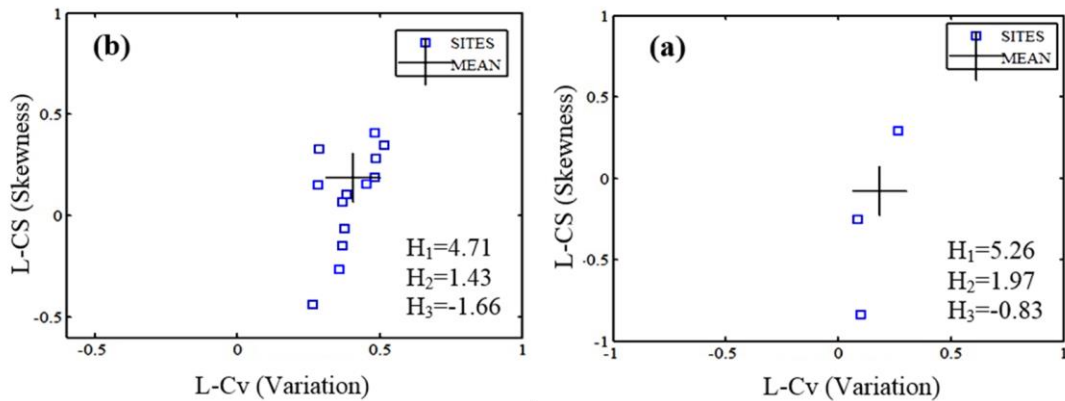
3- Relative Root Mean Square Error (RRMSE)

این مسئله در نمودارهای نسبت گشتاورهای خطی LCV-LCS نیز به خوبی دیده می‌شود و نقاط داده‌ها به طور گسترده‌ای در نمودار LCV-LCS پراکنده و فاصله زیادی از میانگین منطقه‌ای دارند (شکل ۳).

(جدول ۳). معیار ناهمگنی هاسکینگ و والیس با برآزش توزیع کاپا به داده‌های دو منطقه همگن به دست آمده از خوشه‌بندی وارد محاسبه شد. نتایج نشان داد که آماره H_1 در هر دو منطقه یک (۵/۲۶) و دو (۵/۰۲) بیشتر از یک می‌باشد که نشان دهنده‌ی ناهمگن بودن دو منطقه است.

جدول ۳- مقادیر نسبت گشتاورهای خطی و آماره ناهمنوائی هاسکینگ کل ایستگاه‌های منطقه مورد مطالعه

کد ایستگاه	نام ایستگاه	تعداد سال آماری	میانگین	L-CV	L-CS	L-CK	آماره غیرهمگنی (D_i)
۰۰۱-۴۲	چهلگرد	۲۹	۲۶/۶۹	۰/۱۰۷	۰/۳۵۳۱	۰/۴۵۹	۲/۶۴
۰۰۳-۴۲	قلعه شاهرخ	۴۶	۳۲۴/۷	۰/۳۰۰	۰/۲۷۰۷	۰/۱۹۷	۰/۶۳
۰۰۵-۴۲	اسکندری	۴۶	۳۹/۶۴	۰/۵۰۳	۰/۴۹۱۵	۰/۳۳۹	۰/۵۲
۰۰۹-۴۲	زمانخان	۳۰	۱۰۳/۱	۰/۲۷۶	۰/۴۲۶۳	۰/۳۴۹	۱/۲۰
۰۲۹-۴۲	بندپابین	۳۰	۳۱/۹۱	۰/۵۲۲	۰/۴۲۴۷	۰/۲۴۴	۰/۱۷
۰۳۵-۴۲	دره در	۲۴	۵۸/۷۴	۰/۱۳۱	-۰/۰۷۶۰	۰/۳۲۹	*۳/۲۱
۰۳۹-۴۲	مدوار	۲۲	۲۶/۴۸	۰/۴۵۳	۰/۳۹۴۰	۰/۴۰۷	۰/۲۷
۰۴۷-۴۲	مندرجان	۲۸	۱۰۴/۵	۰/۷۰۵	۰/۵۸۸۴	۰/۳۰۴	۱/۸۴
۰۶۱-۴۲	آبدرمیان	۲۲	۲۷/۴۸	۰/۴۳۳	۰/۳۶۱۶	۰/۴۰۷	۰/۲۶
۰۶۳-۴۲	رمزرج	۲۳	۱۵/۵	۰/۵۳۹	۰/۳۹۳۰	۰/۲۰۳	۰/۲۲
۰۶۵-۴۲	سوچ	۲۱	۴۳/۰۱	۰/۵۱۶	۰/۴۰۲۶	۰/۰۹۰	۱/۰۳
۱۰۱-۴۲	کرخنگان	۱۴	۴/۹۹	۰/۴۷۹	۰/۳۶۳۵	۰/۱۵۹	۰/۲۵
۰۵۱-۴۶	باغستان	۲۲	۱۲/۵۵	-۰/۳۱۹۹	-۰/۰۲۴۱	۰/۷۳۹	۱/۹۵
۰۶۱-۴۶	طرزجان	۱۵	۵/۷۰	۰/۴۶۴۲	۰/۲۵۲۶	۰/۰۳۱۸	۱/۲۸
۰۶۳-۴۶	دره	۱۵	۱۴/۳۳	۰/۴۵۷	۰/۳۳۱۰	۰/۱۶۴	۰/۱۷
۱۰۱-۴۶	فخرآباد	۲۸	۱۰/۶	۰/۵۴۶	۰/۴۵۱۸	۰/۲۶۹	۰/۳۵



شکل ۳- نسبت گشتاورهای خطی LCV-LCS منطقه همگن اول (سمت راست) و منطقه همگن دوم (سمت چپ)

در منطقه مورد مطالعه نیز غیر از ایستگاه‌های قلعه شاهرخ، زمان‌خان، اسکندری و بندپابین، در مابقی ایستگاه‌ها طول دوره آماری کمتر از ۳۰ سال و تعداد ایستگاه‌های داخل منطقه کم هستند. از این رو آماره H_1 شاخصی اساسی برای ناهمگنی بوده و آماره H_2 و H_3 حالت کاذبی از همگنی را ارائه می‌دهند.

نکته قابل توجه در نظر داشتن آماره H_1 به عنوان معیاری برای همگنی است. آماره H_1 توان خیلی بهتری برای تفکیک مناطق همگن از ناهمگن دارد و به عنوان شاخصی اساسی برای ناهمگنی به ویژه، در مواردی که تعداد ایستگاه‌های داخل منطقه کم باشد توصیه شده است [۲۰].

تغییرات ضریب چولگی و یا کشیدگی بر روی واریانس چندک‌های تخمینی دارد.

نتایج آزمون ناهمنوائی هاسکینگ و والیس برای همه ایستگاه‌های مورد بررسی، ایستگاه دره‌در را به دلیل داشتن D_i بیشتر از ۳ ناسازگار نشان داد و به همین خاطر از سایر مراحل پژوهش حذف شد (جدول ۳)؛ اما باز نتایج حاکی از ناهمگن بودن منطقه مورد مطالعه است (جدول ۴).

از دیگر دلایل برتری معیار H_1 در تشخیص ناهمگنی این است که معیار H_1 معیاری از انحراف از معیار پراکندگی ایستگاه‌های داخل یک منطقه است. به بیان دیگر، یک معیار ساده پراکندگی نسبت گشتاورهای خطی ایستگاه‌های داخل یک منطقه می‌باشد. برای محاسبه این معیار از ضریب تغییرات استفاده می‌شود زیرا که تغییرات بین ایستگاه‌های داخل یک منطقه تاثیر بیشتری نسبت به

جدول ۴- معیار همگنی برای منطقه مورد مطالعه

H_3	H_2	H_1	
-۲/۸۶	-۰/۸۳	۵/۲۶	منطقه همگن اول
-۰/۹۲	۱/۶۳	۵/۰۲	منطقه همگن دوم
-۱/۲۵	۱/۳۸	۵/۲۳	کل منطقه بدون ایستگاه دره در
-۰/۹۸	۱/۴۴	۴/۸۶	کل منطقه بدون ایستگاه‌های دره‌در و کرخگان
-۱/۲۱	۱/۰۳	۴/۳۲	کل منطقه بدون ایستگاه‌های دره‌در، کرخگان و طرزجان
-۱/۴۳	۰/۶۴	۴/۱۱	کل منطقه بدون ایستگاه‌های دره‌در، کرخگان، طرزجان و ده‌بالا
-۱/۶۴	-۰/۸۸	۳/۵۶	کل منطقه بدون ایستگاه‌های دره‌در، طرزجان، ده‌بالا و چهلگرد
-۱/۹۷	-۱/۱۷	۰/۹۲	کل منطقه بدون ایستگاه‌های دره‌در، طرزجان، ده‌بالا، چهلگرد و زمانخان = منطقه همگن جدید

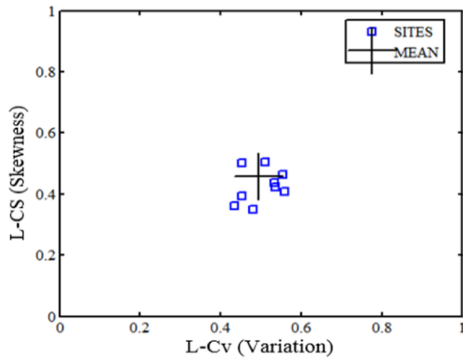
قابلیت کافی در تعیین نواحی ناسازگار و ناهماهنگ در منطقه مورد مطالعه را ندارد چرا که تنها ایستگاه دره‌در با دارا بودن مقدار عددی ۳/۳۱ به عنوان ایستگاه ناسازگار شناخته شد. در حالی که در حوزه زاینده‌رود علاوه بر ایستگاه دره‌در، ایستگاه‌های چهلگرد، قلعه شاهرخ، زمانخان و اسکندری قرار دارند که دارای شرایط اقلیمی و اکولوژیکی نسبتاً مشابهی می‌باشند؛ اما در محدوده مجاز D_i واقع شدند. بنابراین، تنها با استفاده از نتایج آماره D_i نمی‌توان به ناحیه همگن جدید دست یافت.

مقادیر آماره D_i در سه ایستگاه قلعه شاهرخ، چهلگرد و زمانخان به ترتیب ۰/۶۳، ۲/۶۴ و ۱/۲۰ است. این در حالی است که نتایج محاسبه آماره همگنی نشان داد این سه ایستگاه بیشترین تاثیر را در میزان آماره ناهمگنی منطقه دارند و با حذف این ایستگاه‌ها میزان ناهمگنی از ۴/۳۲ به ۰/۹۲ کاهش می‌یابد. پس می‌توان گفت آماره D_i برای مطالعه نواحی ناهماهنگ در منطقه مورد مطالعه چندان موثر نمی‌باشد. نتایج معیار ناهمگنی هاسکینگ برای داده های این ۹ ایستگاه نشان داد که در این منطقه آماره ناهمگنی H_1 کمتر از یک بوده و منطقه همگن است (جدول ۴).

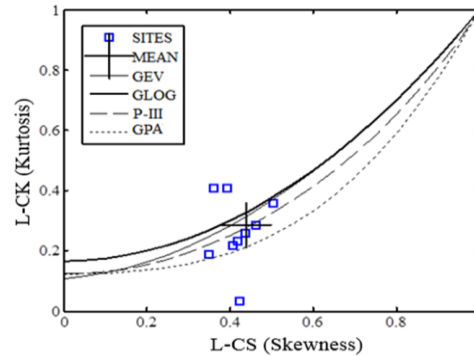
در واقع، به دلیل نبود معیار مناسب برای تعیین مؤلفه‌های مؤثر بر ناحیه‌بندی و وزن آن‌ها، معمولاً مناطق شکل گرفته برای تحلیل فراوانی منطقه‌ای سیلاب همگن نخواهند بود و نیاز به تعدیل برای بهبود همگنی آن وجود دارد [۴۳]. برای تعدیل مناطق ناهمگن گزینه‌هایی ارائه شده است، شامل: حذف یک یا چند ایستگاه، انتقال یک یا چند ایستگاه از یک منطقه به مناطق دیگر، تقسیم یک منطقه به دو یا چند منطقه جدید، به اشتراک گذاشتن یک ایستگاه در چند منطقه، حذف یک منطقه با انتقال ایستگاه‌های آن به مناطق دیگر، ادغام کردن یک منطقه به مناطق دیگر، ادغام دو یا چند منطقه و تعیین دوباره مناطق و استفاده از بردارهای مشخصه‌ی بیشتر و تعیین دوباره مناطق هستند [۲۰]. از این رو برای به دست آوردن ناحیه‌ای همگن، اقدام به حذف ایستگاه‌ها در منطقه مورد مطالعه گردید و هر بار آماره‌های ناهمگنی برای منطقه مورد مطالعه محاسبه شد (جدول ۴).

با حذف ایستگاه‌های ناسازگار در آخر، ناحیه‌ای همگن حاصل شد که شامل ۹ ایستگاه به نام‌های اسکندری، بندپایین، مدوار، مندرجان، آبدرمیان، رمرزج، سوچ، دره و فخرآباد است. نتایج این بررسی نشان داد که آماره D_i

در حقیقت این فرضیه که دبی‌ها در ایستگاه‌های مختلف دارای توزیع اصلی مشابه هستند، برای منطقه همگن به دست آمده مورد قبول است که ناشی از نزدیک بودن مؤلفه‌های توزیع برای منطقه همگن است.



شکل ۴- نسبت گشتاورهای خطی LCV-LCS و LCS-LCK منطقه همگن



انتخابی توسط Z^{DIST} در هر برازش برای ایستگاه‌های داخل منطقه‌ی همگن محاسبه شد و با مقیاس کردن چندک‌ها بر اساس میانگین آن‌ها، مقادیر چندک منطقه‌ای $(q(F))$ هر تابع توزیع در منطقه‌ی همگن به دست آمد که یک عدد بدون بعد است.

مقادیر چندک منطقه‌ای $q(F)$ و شش احتمال عدم وقوع F از ۰/۰۱ تا ۰/۹۹۹، به همراه محدوده‌های خطای ۹۰ درصد، برای چهار تابع توزیع در منطقه‌ی همگن در مقابل هم رسم شدند. این نمودارها که منحنی رشد نامیده می‌شوند برای تعیین دقیق‌ترین تخمین توابع توزیع کاربرد دارند. مقادیر احتمال F در نمودارهای منحنی رشد برای بهتر شدن مقیاس نمودار توسط متغیر گامبل [۲۰] و با استفاده از رابطه‌ی $\log(\log(1/F))$ تبدیل شدند که برای چهار تابع توزیع آمده است (شکل ۵).

بر پایه نتایج به دست آمده، توزیع پیرسون نوع سوم به عنوان مناسب‌ترین توزیع منطقه‌ای می‌باشد. چرا که محدوده‌ی خطا در این تابع به چندک‌های منطقه‌ای توزیع نزدیک‌تر است و از این رو دقت تخمین‌های آن بیشتر می‌باشد. برای بررسی عدم قطعیت نتایج آماره Z^{DIST} همچنین، مقادیر RMSE نسبی در تمام ایستگاه‌های منطقه‌ی همگن برای چهار تابع توزیع محاسبه و با هم مقایسه شد. همان‌طور که از نتایج مشخص است، تابع

نمودار نسبت گشتاورهای خطی LCV-LCS منطقه نیز تأیید کننده همگنی منطقه به دست آمده (ناحیه با ۹ ایستگاه) بوده و نقاط در اطراف میانگین منطقه‌ای تجمع یافته‌اند (شکل ۴).

انتخاب بهترین تابع توزیع احتمالاتی در منطقه

بعد از تعیین مناطق همگن با استفاده از آزمون نکوئی برازش Z^{DIST} ، از میان ۱۰ توزیع احتمالاتی لوگ نرمال ۳ پارامتری^۱، نرمال^۲، ویبول^۳، نمایی^۴، گاما^۵، گمبل^۶، توزیع لجستیک تعمیم یافته^۷، مقادیر حد تعمیم یافته^۸، پیرسون نوع ۳^۹ و پارتو تعمیم یافته^{۱۰} بهترین تابع توزیع منطقه‌ای در منطقه همگن انتخاب گردید.

نتایج آماره Z^{DIST} چهار توزیع GEV ، GLO ، $PE3$ و GPA را به ترتیب با مقادیر ۰/۳۶، ۰/۶۱، ۰/۳۶ و ۱/۰۱ و ۰/۹۷- با مقدار آماره کمتر از ۱/۶۴ به عنوان بهترین توزیع منطقه‌ای سیلاب نشان دادند (جدول ۵). با توجه به این که نتایج حاصل از تحلیل‌های آماری دارای عدم قطعیت هستند، برای استفاده بهتر از این نتایج باید مقدار عدم قطعیت ارزیابی شود. به همین منظور نتایج حاصل از آماره Z^{DIST} در انتخاب تابع توزیع در منطقه همگن جدید ارزیابی شد. برای ارزیابی نتایج این آماره، با استفاده از الگوریتم مونت-کارلو ۱۰۰۰۰ بار شبیه‌سازی اجرا شد. با استفاده از داده‌های شبیه‌سازی شده چندک‌های مربوط به توابع توزیع

- 1- 3-parameters Lognormal (LN3)
- 2- Normal (N)
- 3- Weibull (Wei)
- 4- Exponential (EXP)
- 5- Gamma (Gam)
- 6- Extreme-value type 1 (EV1)
- 7- Generalized Logistic (GLO)
- 8- Generalized Extreme-Value (GEV)
- 9- Pearson type-3 (PE3)
- 10- Generalized Pareto (GPA)

می‌شود برای رسیدن به نتایج قابل اعتماد در انتخاب بهترین تابع توزیع منطقه‌ای، از دو آزمون کولموگروف-اسمیرنوف و χ^2 و همچنین دیاگرام گشتاورهای خطی نیز استفاده شد.

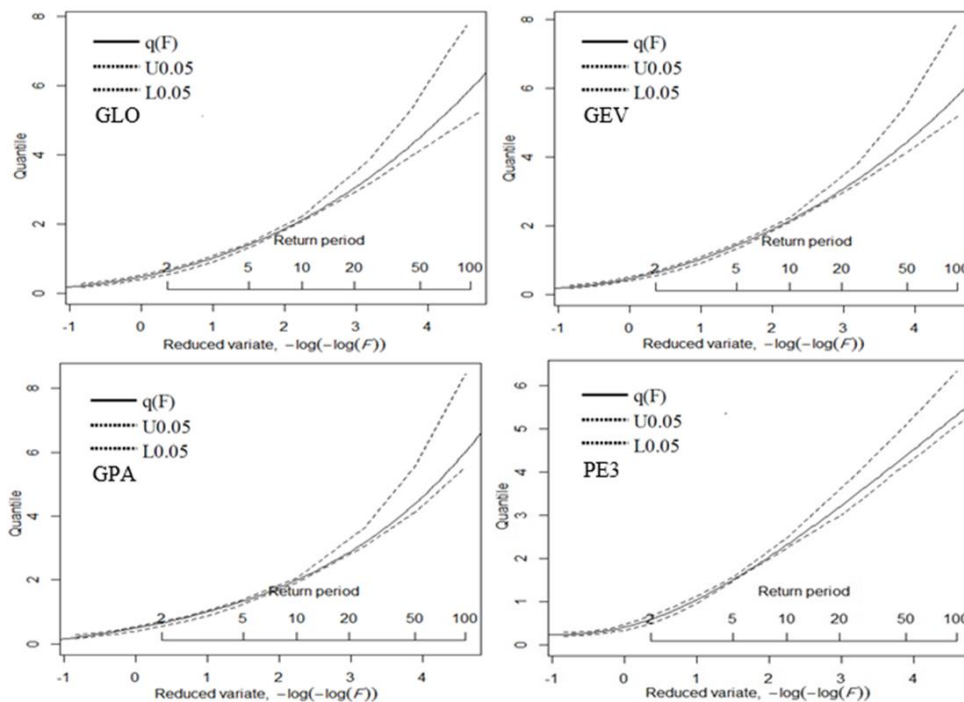
توزیع پیرسون نوع سوم، RMSE نسبی منطقه‌ای کمتری نسبت به سایر توابع توزیع نشان می‌دهد. بنابراین دقت تخمین منطقه‌ای تابع توزیع پیرسون نوع سوم به مراتب بالاتر از سه توزیع دیگر است (شکل ۶). از آنجایی که آماره Z^{Dist} بر اساس نسبت گشتاورهای خطی محاسبه

جدول ۵- مقادیر RMSE و R برای هر یک از توابع توزیع احتمال

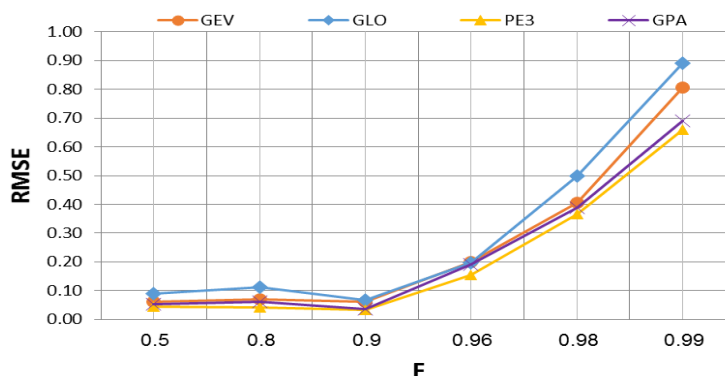
روش	توزیع احتمالاتی	آبدرمیان	اسکندری	فخرآباد	مدوار	بندپایین	دره	مندرجان	سوج	رمزج	Sum rank
MLM	LN3	RMSE	۱۰/۸	۳۴/۸	۱۰/۶	۱۲/۲	۲۰/۱	۵/۴	۱۳/۱	۳۴	۱۰/۵
	R		۰/۸۹	۰/۸۴	۰/۸۱	۰/۸۶	۰/۸۱	۰/۸۸	۰/۸۳	۰/۷۱	۰/۸۱
MLM	EV1	RMSE	۱۵	۵۱	۱۵/۳	۱۶/۳	۲۴/۶	۵/۴	۱۰/۶/۶	۳۷/۲	۱۳
	R		۰/۷۷	۰/۵۹	۰/۵۳	۰/۷۴	۰/۷۱	۰/۸۷	۰/۷۷	۰/۶۸	۰/۶۸
MLM	GPA	RMSE	۹/۳۵	۳۴/۳	۱۰/۶	۱۷/۷	۲/۵۷	۹۵/۴	۲۲/۸	۹/۶۷	۹۷/۱۰
	R		۰/۹۵	۰/۹۱	۰/۷۰	۰/۹۴	۰/۹۱	۰/۹۶	۰/۹۱	۰/۸۹	۰/۸۹
MLM	N	RMSE	۱۱/۹	۴۳/۵	۱۳/۲	۱۲/۵	۱۹/۳	۴/۱	۹۴	۳۳	۱۱
	R		۰/۸۷	۰/۷۳	۰/۶۸	۰/۸۶	۰/۸۳	۰/۹۳	۰/۸۱	۰/۷۹	۰/۷۹
MLM	Exponential	RMSE	۹/۴	۳۵/۱	۱۰/۸	۱۰/۷	۱۷/۸	۳/۲	۹۸/۷	۳۳	۹/۷
	R		۰/۹۲	۰/۸۳	۰/۸۰	۰/۹۰	۰/۸۶	۰/۹۵	۰/۸۹	۰/۸۰	۰/۸۴
MLM	GAM	RMSE	۱۱	۴۰/۲	۱۲/۳	۱۲/۳	۲۰/۳	۳	۹۹/۲	۳۲	۱۰/۸
	R		۰/۸۹	۰/۷۷	۰/۷۳	۰/۸۶	۰/۸۲	۰/۹۴	۰/۸۶	۰/۷۶	۰/۸۰
MLM	GEV	RMSE	۹/۸۸	۳۶/۳	۶۰/۰۳	۱۱/۱	۱۸/۴	۳/۲۹	۹۸/۶	۲۶/۵	۱۰/۱
	R		۰/۹۴	۰/۹۰	۰/۸۸	۰/۹۳	۰/۹۱	۰/۹۷	۰/۹۷	۰/۹۰	۰/۸۸
MLM	PE3	RMSE	۸/۱	۳۳/۲	۱۰/۲	۱۰/۱	۱۷/۱	۲/۲۵	۹۵/۶	۲۴/۵	۷۹/۸
	R		۰/۹۱	۰/۹۲	۰/۹۱	۰/۹۵	۰/۹۲	۰/۹۷	۰/۹۰	۰/۹۰	۰/۹۰
MLM	LN3	RMSE	۹/۸	۳۳/۹	۱۱/۲	۱۱	۱۸/۳	۳/۱	۹۸/۸	۳۲/۱	۱۰
	R		۰/۹۱	۰/۹۰	۰/۷۹	۰/۸۹	۰/۸۵	۰/۹۶	۰/۸۸	۰/۸۰	۰/۸۳
MLM	GLO	RMSE	۱۰/۷۹	۳۹	۱۱/۸	۱۲	۱۹/۶	۲/۴۸	۱۰/۱	۲۶/۵	۱۰/۷
	R		۰/۹۳	۰/۸۹	۰/۸۷	۰/۹۲	۰/۸۹	۰/۹۶	۰/۹۶	۰/۸۷	۰/۸۷
MLM	EV1	RMSE	۹/۹	۳۴	۱۱/۵	۱۰/۸	۱۷/۶	۳/۱	۹۸/۱۱	۳۱/۹	۱۰
	R		۰/۹۱	۰/۸۹	۰/۷۷	۰/۸۹	۰/۸۶	۰/۹۷	۰/۸۷	۰/۸۴	۰/۸۳
MLM	GPA	RMSE	۸/۱	۳۲/۹	۱۰/۲	۹/۹۹	۱۶/۸	۲/۲۹	۹۶/۶	۲۵/۶	۹/۳۱
	R		۰/۹۱	۰/۹۰	۰/۹۱	۰/۹۵	۰/۹۲	۰/۹۷	۰/۹۶	۰/۹۰	۰/۹۰
MLM	Weibull	RMSE	۹	۳۴/۳	۱۰/۵	۱۰/۲	۱۶/۹	۳/۱	۹۷/۳	۳۱	۹/۵
	R		۰/۹۳	۰/۹۲	۰/۸۱	۰/۹۱	۰/۸۷	۰/۹۷	۰/۹۰	۰/۸۲	۰/۸۵
MLM	N	RMSE	۱۱/۹	۳۴/۴	۱۳/۲	۱۲/۵	۱۹/۳	۳/۶	۹۹/۷	۳۱	۱۱
	R		۰/۸۷	۰/۸۴	۰/۶۸	۰/۸۶	۰/۸۳	۰/۹۳	۰/۸۱	۰/۸۴	۰/۷۹
MLM	Exponential	RMSE	۹/۱	۳۴/۲	۱۰/۵	۱۰/۳	۱۷/۱	۳	۹۸/۵	۳۱/۷	۹/۵
	R		۰/۹۲	۰/۹۲	۰/۸۱	۰/۹۱	۰/۸۷	۰/۹۶	۰/۹۰	۰/۸۱	۰/۸۵
MLM	GAM	RMSE	۹/۲	۳۴	۱۰/۶	۱۰/۴	۱۷/۴	۳/۱	۹۸/۲	۳۱/۸	۹/۶
	R		۰/۹۲	۰/۹۲	۰/۸۱	۰/۹۰	۰/۸۷	۰/۹۶	۰/۹۰	۰/۸۴	۰/۸۴
MLM	GEV	RMSE	۱/۷۲	۹/۲	۱/۷۶	۱/۸۶	۵/۷	۱۶/۱	۳۸/۰۵	۱۵/۴	۴/۳۴
	R		۰/۹۹	۰/۹۴	۰/۹۷	۰/۹۹	۰/۹۶	۰/۸۵	۰/۹۵	۰/۸۴	۰/۹۰
MLM	PE3	RMSE	۴/۳۶	۵/۶۱	۱/۳۵	۴/۴۹	۳/۴۷	۱۷/۴۶	۴۱/۳	۱۳/۷۷	۲/۹۸
	R		۰/۹۳	۰/۹۸	۰/۹۸	۰/۹۶	۰/۹۹	۰/۹۳	۰/۹۲	۰/۹۷	۰/۹۵
MLM	LN3	RMSE	۳/۳	۷/۷۸	۱/۳۵	۲/۶۸	۴/۸	۱۷	۴۱/۵	۱۴/۹۵	۳/۷
	R		۰/۹۸	۰/۹۶	۰/۹۸	۰/۹۸	۰/۹۷	۰/۹۰	۰/۹۱	۰/۹۳	۰/۹۳
MLM	GLO	RMSE	۱/۴۳	۹/۷۶	۱/۹۶	۱/۹۳	۶/۱۱	۱۵/۸۲	۳۶/۳	۱۵/۷	۴/۶۲
	R		۰/۹۹	۰/۹۳	۰/۹۶	۰/۹۹	۰/۹۶	۰/۸۵	۰/۹۰	۰/۸۵	۰/۸۸
MLM	EV1	RMSE	۵	۱۵/۰۷	۳/۶۳	۴/۸۴	۸/۵۳	۱۴/۵۵	۱۹/۱۴	۱۱	۴/۶
	R		۰/۹۴	۰/۸۴	۰/۸۵	۰/۹۴	۰/۹۱	۰/۹۲	۰/۹۴	۰/۸۸	۰/۸۸
MLM	GPA	RMSE	۳/۲۵	۷/۴	۱/۴۳	۲/۸۲	۴/۶۲	۱۷	۳۹/۹۴	۱۴/۵۳	۳/۶۴
	R		۰/۹۷	۰/۹۶	۰/۹۸	۰/۹۸	۰/۹۸	۰/۸۸	۰/۹۴	۰/۹۱	۰/۹۳
MLM	Weibull	RMSE	۵/۱۲	۵/۹۵	۱/۲۹۵	۴/۱	۳/۸۸	۱۸	۴۳/۳	۱۴/۲	۳/۱۶
	R		۰/۹۵	۰/۹۸	۰/۹۸	۰/۹۶	۰/۹۸	۰/۸۷	۰/۹۲	۰/۸۸	۰/۹۵
MLM	N	RMSE	۷/۱	۲۱/۷۵	۴/۹۸	۷	۱۳/۰۵	۱۲/۹۶	۲۰/۶	۱۲/۱	۵/۹
	R		۰/۸۷	۰/۶۹	۰/۸۷	۰/۷۸	۰/۷۸	۰/۷۹	۰/۹۶	۰/۷۹	۰/۸۰
MLM	Exponential	RMSE	۴/۵۴	۹/۵	۲/۴	۴/۱	۴/۸	۱۶/۱	۲۸/۱	۱۱	۳/۵
	R		۰/۹۵	۰/۹۴	۰/۹۳	۰/۹۶	۰/۹۷	۰/۸۳	۰/۹۵	۰/۸۹	۰/۹۴
MLM	GAM	RMSE	۴/۵۴	۹/۷۶	۴/۲	۴/۱	۴/۸	۱۶/۱	۲۸	۱۱	۳/۵۲
	R		۰/۹۵	۰/۹۳	۰/۹۳	۰/۹۶	۰/۹۷	۰/۸۳	۰/۹۳	۰/۸۸	۰/۹۳

عنوان توزیع منطقه‌ای در نظر گرفته می‌شود. این نتایج نشان می‌دهد که در میان آزمون‌های برازش نکوئی، روش Z^{Dist} نتایج مطمئن‌تری در انتخاب بهترین تابع توزیع منطقه‌ای ارائه می‌دهند. اما برای رسیدن به نتایج مطمئن‌تر، RMSE چندک‌های به دست آمده از ۱۰ توزیع احتمالاتی مورد بررسی قرار گرفت. به همین منظور، به ترتیب مؤلفه‌ها و چندک‌های توزیع‌های احتمالاتی به سه روش گشتاورهای خطی، معمولی و بیشینه درست‌نمایی برآورد شدند (شکل ۷).

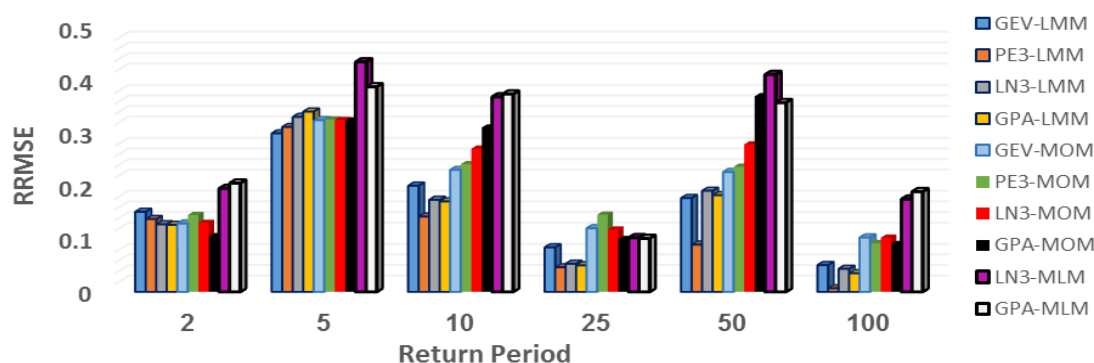
نتایج آزمون χ^2 در سطح معنی‌دار ۵ درصد نشان داد از میان ۱۰ توزیع احتمالاتی، تنها سه توزیع EV1، GAM و EXP برازش مناسبی با داده‌ها داشته و مقدار محاسباتی آماره χ^2 در آن‌ها کمتر از مقادیر جدول است. اما نتایج آزمون کولموگروف-اسمیرنوف نشان داد که تمام توزیع‌ها در سطح ۵ درصد معنی‌دار هستند. نتایج نمودار گشتاورهای خطی LCS-Lck نشان داد که میانگین منطقه ای و پراکنش داده‌ها در اطراف توزیع پیرسون نوع ۳ به صورت خوشه‌ای و مناسب گسترش دارند و از این رو به



شکل ۵- منحنی رشد حاصل از برازش توابع توزیع‌های احتمالاتی در منطقه همگن



شکل ۶- RMSE توابع منطقه‌ای حاصل از آماره Z^{Dist} در منطقه همگن



شکل ۷- RRMSE توزیع‌های PE_3 ، GEV، LN3 و GPA به سه روش LMM، MOM و MLM

نمرات یکسان داده شد و مجموع این نمرات برای هر توزیع و روش به دست آمد. بر این اساس روشی مناسب است که کمترین نمره را دارا است (جدول ۵). با توجه به نتایج این پژوهش در مورد دبی حداکثر لحظه‌ای و با مقایسه مقادیر نمرات RMSE و R مشخص می‌شود که توزیع پیرسون نوع ۳ و روش گشتاورهای خطی به ترتیب با دارا بودن نمره ۴۳ و ۲۵ کمترین نمره را در میان ۱۰ توزیع احتمالاتی و روش MOM و MLM دارد و به عنوان بهترین روش و توزیع احتمالاتی انتخاب می‌شود. توزیع پارتو تعمیم یافته و روش گشتاورهای خطی به ترتیب با دارا بودن نمره ۴۸ و ۲۸ در رتبه دوم و پس از آن توزیع لوگ نرمال سه پارامتری و روش گشتاورهای خطی به ترتیب با دارا بودن نمره ۴۹ و ۲۹ در رتبه سوم، توزیع مقادیر حد تعمیم یافته و روش گشتاورهای خطی به ترتیب با نمره ۵۱ و ۳۸ در رتبه چهارم قرار گرفته است (جدول ۷).

این مقایسه نشان از برتری روش گشتاورهای خطی در محاسبه مقادیر حدی جریان است. نتایج RRMSE نشان داد که روش MOM در دوره بازگشت‌های ۲ و ۵ ساله برآوردهای بهتری ارائه می‌دهد و در دوره بازگشت‌های ۱۰، ۲۵، ۵۰ و ۱۰۰ ساله، توزیع PE_3 و روش LMM مقادیر خطای محاسباتی کمتری داشته و مناسب‌تر است (شکل ۷). از طرفی برای مدیریت جامع منابع آب، طراحی سازه‌ها به طور عمده بر پایه دوره بازگشت‌های بالا انجام می‌شود. بنابراین با توجه به دقت روش LMM در برآورد توزیع منطقه‌ای پیرسون نوع ۳، می‌توان از این روش و توزیع احتمالاتی برای ناحیه مورد مطالعه استفاده کرد.

برآورد مؤلفه‌ها و چندک‌های توزیع‌های احتمالاتی

نتایج بررسی سه روش نشان داد که بر خلاف دو روش MOM و LMM، روش MLM از میان ۱۰ توزیع احتمالاتی مورد بررسی، کارایی لازم را برای برآورد مؤلفه‌های توزیع‌های آماری GEV، GLO، و بیبول و PE_3 ندارد. در بررسی صورت گرفته توسط [۴] نیز بیان شده است که برای نمونه‌های با طول دوره آماری کوتاه و تعداد کم ممکن است برآوردهای MLM امکان پذیر نباشد. این مسئله نشان از اهمیت حداقل طول دوره آماری ۳۰ سال در انجام مطالعات دارد.

ارزیابی عملکرد سه روش LMM، MOM و MLM

هدف پژوهش حاضر بررسی عملکرد سه روش LMM، MOM و MLM در برآورد مؤلفه‌های توزیع‌های احتمالاتی است. در ابتدا احتمال تجربی داده‌های مشاهداتی با استفاده از روش و بیبول به دست آمد و با استفاده از آماره‌های RMSE، RRMSE و R چندک‌های به دست آمده از هر یک از توزیع‌های احتمالاتی مورد بررسی قرار گرفت. بر این اساس، روش و توزیعی مناسب تشخیص داده می‌شود که، کمترین مقدار RMSE، RRMSE و بیشترین مقدار R را دارا باشند. به منظور مقایسه بهتر، مجموعه نمرات داده شده (رتبه‌ها) برای هر یک از توزیع‌های احتمالاتی و روش‌های LMM، MOM و MLM به دست آمد. در این روش به کمترین مقدار RMSE و بیشترین مقدار R، نمره یک و روشی که بیشترین مقدار RMSE و کمترین مقدار R را دارا است، بالاترین نمره و به مجموع مقدار RMSE و R مساوی

شناخته شد. در حالی که نتایج آماره H نشان از ناهمگنی منطقه مورد مطالعه دارد.

۲) بین سه روش LMM، MOM و MLM در برآورد مؤلفه و چندک‌های ۱۰ توزیع احتمالاتی برای دبی‌های حداکثر لحظه‌ای، توزیع PE3 و روش LMM در رتبه اول قرار دارد. رتبه‌های بعدی به ترتیب، توزیع GPA و روش LMM، توزیع LN3 و روش LMM و توزیع GEV و روش LMM به عنوان توزیع منطقه‌ای قرار گرفتند. این مقایسه حاکی از توانمندی بالای روش LMM در مقایسه با دو روش دیگر است.

۳) در بین سه آزمون نکوئی برآزش Z^{DIST} ، کولموگراف-اسمیرنوف و χ^2 ، آماره Z^{DIST} از توانایی بهتری در انتخاب بهترین تابع توزیع منطقه‌ای برخوردار است و می‌توان به نتایج آن اعتماد نمود. با استفاده از آزمون Z^{DIST} ، چهار توزیع آماری GLO، GEV، PE3 و GPA از میان ۱۰ توزیع احتمالاتی برای بررسی عملکرد روش‌های LMM، MOM و MLM در نظر گرفته شد که سه توزیع آن یعنی توزیع PE3، GPA و GEV جز توزیع‌های منطقه‌ای منتخب هستند. این در حالی است که آماره χ^2 ، تنها سه توزیع EV1، GAM و EXP را مناسب برای برآزش به داده‌ها نشان می‌دهد اما هیچ‌کدام از این سه توزیع جز توزیع‌های منتخب نیستند. نتایج آماره کولموگراف-اسمیرنوف نیز تمام توزیع‌ها را در سطح ۵ درصد معنی‌دار نشان می‌دهد.

۴) ارزیابی دقت آماره Z^{DIST} توسط شبیه‌سازی با الگوریتم مونت کارلو، راهی مناسب و قابل اعتماد برای انتخاب مناسب‌ترین تابع توزیع منطقه‌ای است. هم مقادیر RMSE منطقه‌ای و هم نمودارهای Growth Curve توزیع‌های احتمالاتی، مناسب‌ترین و دقیق‌ترین تخمین‌ها را مربوط به تابع توزیع PE3 نشان دادند.

۵) در پژوهش حاضر، روش LMM نسبت به روش MOM و MLM از توانایی بالاتری در برآورد مؤلفه‌های توزیع برای نمونه‌های کوچک برخوردار است. بنابراین با توجه به مشکل نبود و کوتاهی آماره دبی پیک در اغلب ایستگاه‌های آب‌سنجی محدوده ایران مرکزی، می‌توان روش LMM را به عنوان روش منطقه‌ای مناسب در برآورد دبی پیک در این مناطق به کار برد. در پایان باید

اسلامیان و چاووشی (۲۰۰۴) در حوزه آبخیز مرکزی، کشتکار و همکاران (۲۰۱۲) و مهدوی و همکاران (۲۰۰۶) در مناطق خشک و نیمه خشک ایران، Langat و همکاران (۲۰۱۹)، Khandi و همکاران (۲۰۰۲) و Jingyi و Hall (۲۰۰۴) نیز توزیع پیرسون نوع ۳ را به عنوان مناسب‌ترین تابع توزیع برای مقادیر حدی جریان بیان کردند. نتایج این بررسی با نتایج Alam و همکاران (۲۰۱۶)، غلامی و همکاران (۲۰۱۶)، Vivekananda (۲۰۱۵ الف) و Ahmad و همکاران (۲۰۱۵) و اسلامیان و همکاران (۲۰۱۲)، همخوانی داشته و می‌توان بیان داشت که روش LMM نسبت به روش MLM و روش MOM برآوردهای مناسب‌تری را ارائه می‌کند. در حالی که ÖZTEKİN (۲۰۰۵) نشان داد که روش MOM و بعد از آن روش MLM تطابق بهتری با داده‌های دبی حداکثر سالانه دارند. Vivekananda (۲۰۱۵ ب) روش MLM را مناسب‌ترین روش برآورد مؤلفه‌های توزیع احتمالاتی بیان کرد. این مسئله، ناشی از عدم وجود قطعیت در انتخاب یک تابع توزیع و روش خاص برای یک منطقه یا کشور است. چرا که وجود تنوع هیدرواقلمی منجر به تغییر در داده‌های هیدرولوژیکی و در نتیجه توزیع احتمالاتی حاکم بر آن می‌شود. ولی می‌توان توزیع و روش مناسب را بر اساس کمترین خطا و نزدیکی به داده‌های مشاهداتی منطقه مورد مطالعه انتخاب نمود. بنابراین می‌توان نتیجه گرفت که توزیع PE3 در منطقه مورد مطالعه بهترین تابع برای برآزش به داده‌های دبی حداکثر لحظه‌ای است و بعد از آن توزیع GPA و LN3 قرار دارد.

نتیجه‌گیری

روش تحلیل منطقه‌ای یک روش کارآمد برای برآورد دبی سیل در ایستگاه‌های بدون داده یا دارای آمار کوتاه مدت است. تحلیل منطقه‌ای شامل چهار مرحله، تهیه داده مشاهداتی، تعیین مناطق همگن، انتخاب یک توزیع تناوب منطقه‌ای و تخمین مؤلفه‌های آن و برآورد سیل است. نتایج این بررسی نشان داد:

۱) آماره D_i قابلیت کافی در تعیین نواحی ناهم‌هنگ در منطقه مورد مطالعه را ندارد. چرا که تنها ایستگاه دره در با دارا بودن مقدار عددی ۳/۳۱ به عنوان ایستگاه پرت

دقت روش LMM در برآورد توزیع منطقه‌ای پیرسون نوع ۳، می‌توان از این روش و توزیع احتمالاتی در منطقه مورد مطالعه استفاده کرد.

به این نکته اشاره نمود که روش‌های به کار گرفته شده در تحلیل منطقه‌ای سیلاب بستگی به شرایط منطقه از نظر نوع و طول دوره آماری، دوره بازگشت مورد انتظار، شرایط اقلیمی دارد و تعمیم نتایج آن‌ها، نیازمند طول دوره آماری مناسب در منطقه مورد بررسی است. بنابراین با توجه به

References

- [1]. Abida. H. & Ellouze. M. (2007). Probability distribution of flood flow in Tunisia. *Hydrology and Earth System Sciences Discussions*. 4: 957-981.
- [2]. Adamowski. K. (2000). Regional analysis of annual maximum and partial flood data by nonparametric and L-moments methods. *Journal of Hydrology*. 229: 219-231.
- [3]. Afandizadeh. S., Ameri. M. & Mirabi Moghadaam. M. H. (2010). A new method for optimization of log likelihood function for the purpose of determining the coefficients of accident prediction models. *Journal of Transportation Research*. 7(3(24)):215-255 (in Farsi).
- [4]. Ahmad. I., Fawad. M. & Mahmood. I. (2015). At-Site Flood Frequency Analysis of Annual Maximum Stream Flows in Pakistan Using Robust Estimation Methods. *Polish Journal of Environmental Studies*. 24(6): 1-9.
- [5]. Alam. J., Muzzammil. M. & Khan. M. K. (2016). Regional flood frequency analysis: comparison of L-moment and conventional approaches for an Indian catchment. *ISH Journal of Hydraulic Engineering*. 22(3):1-7.
- [6]. Bhuyan. A., Borah. M. & Kumar. R. (2010). Regional flood frequency analysis of north-bank of the river Brahmaputra by using LH-moments. *Water Resources Management*. 24(9): 1779-1790.
- [7]. Burn. D. H. (1990). Evaluation of regional flood frequency analysis with a region of influence approach. *Water Resource Research*. 26(10): 2257-2265.
- [8]. Cavadias, G. (1990). The canonical correlation approach to regional flood estimation. *Regionalization in hydrology*, 191, 171-178.
- [9]. Dalrymple. T. (1960). Flood Frequency Analysis. US Geological Survey, Water Supply Paper, 1543 p.
- [10]. Eslamian. S. S. & Chavoshi. S. (2003). Regional flood frequency analysis using L-Moments in central basins of Iran. *Journal science and Technology Agricultural and Natural Resources*. 7:1.1-17 (in Farsi).
- [11]. Eslamian. S. S., Fathian. F. & Hasanzadeh. H. (2012). A comparative evaluation of L-moment method with maximum likelihood parametric method and kernel functions of nonparametric methods in five Iranian stations. *Watershed Engineering and Management*. 4(2): 63-72 (in Farsi).
- [12]. Faucher. D., Rasmussen. P. S. & Bobee. B. (2001). A distribution function based on bandwidth selection method for kernel quantile estimation. *Journal of Hydrology*. 250(1): 1-11.
- [13]. Gholami. A. (2016). Comparison efficiency of linear moment method with common methods for estimation of annual maximum peak flood (case study: Sheikh Osman watershed in Oshna Vieh area). *Journal of watershed management research*. 7(13): 119-127 (in Farsi).
- [14]. Greenwood. J. A., Landwehr. J. M., Matalas. N. C. & Wallis. J. R. (1979). Probability Weighted Moments: definition and relation to parameters of several distributions expressible in inverse form. *Water Resources Research*. 15(5): 1049-1054.
- [15]. Hailegeorgis. T. T. & Alfredsen. K. (2017). Regional flood frequency analysis and prediction in ungauged basins including estimation of major uncertainties for mid-Norway. *Journal of Hydrology: Regional Studies*. 9:104-126.
- [16]. Hamed. K. & Rao. A. R. (1999). Flood frequency analysis. CRC press, 59p.
- [17]. Hassanzadeh. Y. & Abdikordani. A. (2012). The four-parameter kappa distribution in prediction of standardized

- precipitation index. *Irrigation Sciences and Engineering*. 35(2):21-31 (in Farsi).
- [18]. Hosking. J. R. M. (1986). The theory of probability weighted moments. Research Report RC12210, IBM Research Division, Yorktown Heights, NY, 10598p.
- [19]. Hosking. J. R. M. (1990). L-moment: analysis and estimation of distributions using linear combinations of order statistics. *Journal of the Royal Statistical Society*. 52(2):105-124.
- [20]. Hosking. J. R. M. & Wallis. J. R. (1997). Regional Flood Frequency Analysis: An approach based on L-moment. Cambridge University Press, London, UK, 224p.
- [21]. Izinyon. O. C. & Ajumuka. H. N. (2013). Probability distribution models for flood prediction in Upper Benue River Basin–Part II. *Civil and Environmental Research*. 3(2): 62-74.
- [22]. Jingyi. Z. & Hall. M. J. (2004). Regional Flood Frequency Analysis for the Gan-Ming River Basin in China. *Journal of Hydrology*. 296:98-117.
- [23]. Keshtkar. A., Salajegheh. A. & Najafi Hajivar. M. (2012). Flood flow frequency model selection using L-moment method in arid and semi arid regions of Iran. *Desert*. 17(1): 41-48.
- [24]. Khalili. K., Nazeri. Tahrudi. M., Abbaszadeh Afshar. M. & Nazeri tahrudi. Z. (2014). Comparison of different peak flow frequency distribution functions (Case Study: Babolrood River). *Journal of Middle East Applied Science and Technology*. 7(4): 174-179.
- [25]. Khandi. S., Kachroo. H. & Gunasekara. R. K. (2002). Comparison of Annual Maximum Series and Partial Duration of Homogeneous Regions. *Hydrological Science Journal*. 45:437-447.
- [26]. Khosravi. G., Majidi. A. & Nohegar. A. (2012). Determination of suitable probability distribution for annual mean and peak discharges estimation (case study: Minab river-barantin gage, Iran). *International Journal of Probability Statistics*. 1(5): 160-163.
- [27]. Kjeldsen. T. R., Smithers. J. & Schulze. R. (2002). Regional flood frequency analysis in the KwaZulu-Natal province, South Africa, using the index-flood method. *Journal of hydrology*. 255(1-4): 194-211.
- [28]. Kroll. C. N. & Vogel. R. M. (2002). Probability distribution of low stream flow series in the United States. *Journal of Hydrologic Engineering*. 7(2): 137-146.
- [29]. Kumar. R. & Chatterjee. C. (2005). Regional flood frequency analysis using L-Moments for North Brahmaputra region of India. *Journal of Hydrologic Engineering*. 10: 1-7.
- [30]. Langat. P. K., Kumar. L. & Koech. R. (2019). Identification of the Most Suitable Probability Distribution Models for Maximum, Minimum, and Mean Streamflow. *Water*. 11(4): 1-24.
- [31]. Landwehr. J. M., Matalas. N. C. & Wallis. J. R. (1979). Probability weighted moments compared with some traditional techniques in estimating Gumbel parameters and quantiles. *Water Resources Research*. 15(5): 1055-1064.
- [32]. Leonardo. V. N. & Goffredo. L. L. (2008). Use of L-moments approach for regional flood frequency analysis in Sicily, Italy. *Water Resources Management*. 23: 2207-2229.
- [33]. Li. M., Li. X. & Ao. T. (2019). Comparative Study of Regional Frequency Analysis and Traditional At-Site Hydrological Frequency Analysis. *Water*. 11(3): 1-19.
- [34]. Mahdavi. M., Salajegheh. A., Keshtkar. A. R., Fahmi. H. & Sharifi. F. (2006). A Study of the Appropriate Probability Distributions for Annual Flow Series, Using L-Moment Method in Arid and Semi-arid Regions. *Iranian journal of natural resources*. 59(1): 74-87 (in Farsi).
- [35]. Malekinezhad. H., Nachtnebel. Hp. & Klik. A. (2011). Comparing the index flood and multiple regression methods using L-moment. *Journal of Physics and Chemistry of the Earth*. 36: 54-60.
- [36]. Mohammadi. M. (2008). Regional Flood Frequency Analysis using L-moments approach in some watersheds of Isfahan – Sirjan and Yazd - Ardakan basins. M.Sc. Thesis, Faculty of Natural Resources, Yazd University (in Farsi).
- [37]. Mohammadi. M., Malekinezhad. H. & Afkhami. H. (2017a). Determination of

- Hydrological Homogeneous Regions using Canonical Correlation Analysis. *Iranian Journal of Watershed Management Science and Engineering*. 11(38): 11-24 (in Farsi).
- [38]. Mohammadi. M., Malekinezhad. H. & Dastorani. M. T. (2017b). Comparing Regional Analysis methods of estimation of the peak flow in some watershed s of Isfahan- Sirjan and Yazd- Ardakan Basins. *Iranian Journal of Natural Resources*. 70(2): 515-529 (in Farsi).
- [39]. Noori Gheidari. M. H. (2012). Identifying the Regional of Ungauged Station for Regional Flood Frequency Analysis Using Canonical Correlation Method. *Journal of Water and Soil*. 26: 943-952 (in Farsi).
- [40]. Noto. L. V. & La Loggia. G. (2009). Use of L-moments approach for regional flood frequency analysis in Sicily, Italy. *Water Resources Management*. 23(11): 2207-2229.
- [41]. Ourada. TBMJ. Girard. C., Cavadias. G. S. & Bobee. B. (2001). Regional flood frequency estimation with canonical correlation analysis. *Journal of hydrology*. 254:157-173.
- [42]. ÖZTEKİN. T. (2005). Comparison of parameter estimation methods for the three-parameter generalized Pareto distribution. *Turkish journal of agriculture and forestry*. 29(6): 419-428.
- [43]. Rao. A. R. & Srinivas. V. (2006). Regionalization of watersheds by hybrid-cluster analysis. *Journal of Hydrology*. 318(1-4): 37-56.
- [44]. Saf. B. (2009). Regional flood frequency analysis using L-moments for the West Mediterranean region of Turkey. *Water Resources Management*. 23(3): 531-551.
- [45]. Sakarasubramaniam. A. & Srinivasan. K. (1999). Investigation and composition of sampling properties of L-Moment and conventional Moments. *Journal of hydrology*. 218:13-34.
- [46]. Vivekanandan. N. (2015a). Flood frequency analysis using method of moments and L-moments of probability distributions. *Cogent Engineering*. 2(1): 1-10.
- [47]. Vivekanandan. N. (2015b). Frequency analysis of annual maximum flood discharge using method of moments and maximum likelihood method of gamma and extreme value family of probability distributions. *International Journal of Math Computer Science*. 1(3): 141-146.
- [48]. Vogel. R. M., McMahon. T. A. & Chiew. FHS. (1993). Flood flow frequency model selection in Australia. *Journal of hydrology*. 146:421-449.

Comparison efficiency of moments, L-moments and maximum likelihood methods in regional flood analysis

1- Mojdeh Mohammadi*, Ph.D. of Watershed Management Engineering, Faculty of Natural Resources and Desert Studies, Yazd University, Yazd, Iran.

m.mohammadi@stu.yazd.ac.ir

2- Mohammad Reza Ekhtesasi, Professor, Faculty of Natural Resources and Desert Studies, Yazd University, Yazd, Iran.

3- Ali Talebi, Professor, Faculty of Natural Resources and Desert Studies, Yazd University, Yazd, Iran.

Received: 08 Feb 2021

Accepted: 08 Mar 2021

Abstract

It is necessary to determine the appropriate probability distribution for fitting the flood data. There are several methods for estimating the parameters of probability distribution functions. In the present study, L-moments approach is used for flood frequency analysis in Gavkhouni Basin, Abarkouh- Sirjan Basin and Yazd- Ardakan Basin. Some stations were selected and their data were analyzed to find out peak discharge. Then, Hosking and Wallis homogeneity criterion were used to identify the homogeneous region. Examining the homogeneity criterion closely shows the discordant region. Eliminating and transferring some selected stations, leads to determining a homogeneous region of flood frequency. Parameters of the regional frequency distribution were evaluated using L-moment, moments, and maximum likelihood methods. The results of this analysis were compared using RMSE, R, and RRMSE statistics. The goodness of fit tests such as Chi-square, Kolmogorov-Smirnov were applied for checking the adequacy of fitting of probability distributions to the recorded data. The study reveals that the Pearson Type-3 distribution (using LMM) is better suited amongst the ten distributions used in estimating the peak flood. In general, the L-moment method estimated the peak flood for return periods of 10, 25, 50, and 100-year, and moments method at return periods of 2 and 5-year.

Keywords: Flood frequency, Homogeneity, peak discharge, Regional distribution.