

تحلیل فراوانی منطقه‌ای سیلاب با استفاده از روش گشتاورهای خطی بر اساس تحلیل خوشه‌بندی در حوضه دریاچه ارومیه

فرشاد فتحیان^{۱*} و حجت احمدزاده^۲

* نویسنده مسئول، دانشجوی مقطع کارشناسی ارشد مهندسی منابع آب، دانشگاه تربیت مدرس farshad.fathian@gmail.com

۲- دانشجوی مقطع کارشناسی ارشد مهندسی منابع آب، دانشگاه تربیت مدرس

تاریخ پذیرش: ۹۱/۷/۲۲

تاریخ دریافت: ۹۰/۷/۹

چکیده

به منظور تعیین نواحی همگن و تخمین دبی‌های حداکثر لحظه‌ای با دوره بازگشت‌های مختلف در نقاط فاقد ایستگاه آب-سنجی، از روش خوشه‌بندی و تحلیل فراوانی منطقه‌ای به روش گشتاورهای خطی استفاده شده است. نتایج تحلیل خوشه‌بندی با روش K-means نشان داد پنج منطقه همگن در حوضه دریاچه ارومیه وجود دارد. با انجام آزمون ناجوری، ایستگاه‌های سه‌زب و بوکان ناجور بوده و پس از حذف به صورت ایستگاهی تحلیل فراوانی شدند. آزمون ناهمگنی نیز تمامی نواحی را همگن تشخیص داد و برای تحلیل فراوانی این نواحی از ۱۱ تابع توزیع احتمالی استفاده شد. بر اساس آزمون‌های نکونی برازش، توابع توزیع لجستیک تعمیم‌یافته برای نواحی A، B و C، لوگ نرمال سه پارامتری برای ناحیه D، مقادیر حدی تعمیم‌یافته برای ناحیه E و برای ایستگاه‌های سه‌زب و بوکان به ترتیب توابع لوگ نرمال دو و سه پارامتری انتخاب شدند. علاوه بر این، تابع توزیع مقادیر حدی تعمیم‌یافته برای تمامی نواحی همگن مناسب تشخیص داده شد.

کلید واژه‌ها: تحلیل منطقه‌ای سیلاب، تحلیل خوشه‌بندی، گشتاورهای خطی، حوضه دریاچه ارومیه.

مقدمه

مطالعات انجام شده در زمینه‌ی تحلیل خوشه بندی سابقه چندان طولانی ندارد که در ادامه به برخی از آنها اشاره می‌شود. موسلی^۲ از روش تحلیل خوشه‌ای برای تعیین نواحی همگن بارندگی نیوزیلند استفاده کرده است (۲۴). سانتوس و همکاران^۳ جهت ارزیابی الگوی مکانی و زمانی شاخص خشکسالی SPI، از روش‌های خوشه‌بندی k-means و تحلیل مولفه‌های اصلی^۴ استفاده کردند (۲۷). برای ناحیه‌ای کردن سیلاب‌ها، باسکر و اوکونور^۵ روش‌های تحلیل خوشه‌ای و باقیمانده‌ها را با هم مقایسه کردند که نتایج آنها کاملاً مشابه یکدیگر بود (۵). همچنین بورن^۶ جهت مشخص کردن ایستگاه‌ها در یک ناحیه برای تحلیل فراوانی سیل ناحیه‌ای، از تحلیل خوشه‌ای استفاده کرد (۶).

برای نخستین بار، کاربرد گشتاورهای وزنی احتمال^۷ در علم هیدرولوژی توسط گرین‌وود و همکاران^۸ ارائه شد و سپس توسط

تخمین دبی‌های حداکثر لحظه‌ای با دوره بازگشت مورد نظر در منطقه مورد مطالعه از جمله پارامترهای لازم در طراحی انواع سازه‌های آبی است. اما برآورد آنها در بیشتر پروژه‌های آبی به علت عدم وجود ایستگاه آب‌سنجی و یا ناقص بودن داده‌های آن در محل پروژه، با مشکل روبه‌رو شده است (۱۷). بدین منظور استفاده از روش‌های تحلیل فراوانی منطقه‌ای سیلاب در نواحی که از همگنی هیدرولوژیکی برخوردار هستند، می‌تواند کارساز باشد. به طور کلی تحلیل منطقه‌ای سیلاب دو هدف عمده را دنبال می‌کند: اول تعیین حوضه‌های آبریزی که از لحاظ هیدرولوژیکی همگن هستند. دوم بالا بردن دقت تحلیل فراوانی سیل در ساختگاه‌های دارای آمار و برآورد سیلاب در ساختگاه‌های فاقد آمار (۷).

تحلیل خوشه‌بندی^۱ از جمله روش‌های مفید و جدید جهت تعیین نواحی همگن به منظور تحلیل منطقه‌ای سیلاب در یک حوضه آبریز می‌باشد. در این روش، قرار دادن ایستگاه‌های آب‌سنجی حوضه در دسته‌های متمایز، بر اساس میزان مشابهت بین آنها (معیار فاصله) صورت می‌گیرد.

2 - Mosley

3 - Santos et al.

4- Principal component analysis

5- Bhaskar and Oconnor

6- Burn

7-Probability weighted moments

8- Greenwood et al.

1-Clustering

لجستیک تعمیم‌یافته به عنوان بهترین توزیع منطقه‌ای برای هر دو شاخص انتخاب گردید (۹ و ۱۰).

هدف از مطالعه حاضر تلفیق نتایج حاصل از تحلیل خوشه‌ای و روش گشتاورهای خطی، ارزیابی روش‌های تحلیل خوشه‌ای در تعیین نواحی همگن هیدرولوژیک و تحلیل فراوانی منطقه‌ای دبی سیلاب این نواحی با تئوری گشتاورهای خطی در حوضه دریاچه ارومیه است.

مواد و روش‌ها

تحلیل خوشه‌بندی

یکی از تکنیک‌هایی که در مبحث داده کاوی جهت استخراج الگوهای موجود در حجم عظیمی از داده‌ها مورد استفاده قرار می‌گیرد، خوشه‌بندی است (۱۳). روش‌های خوشه‌بندی به دو گروه کلی سلسله مراتبی و تجزیه‌ای تفکیک می‌شوند. در این مطالعه، جهت تعیین نواحی همگن حوضه از روش‌های خوشه‌بندی تجزیه‌ای K-means و K-medoids استفاده شد و در ادامه به شرح آنها پرداخته می‌شود.

روش خوشه‌بندی K-means: در این روش ابتدا اشیاء موردنظر به صورت تصادفی در خوشه‌هایی که تعداد آنها را خود مشخص کرده‌ایم، قرار می‌گیرند و سپس براساس مشابهت آنها با مراکز خوشه‌ها، کار اختصاص دادن آنها به خوشه‌ها به انجام می‌رسد و این حلقه تا زمانی ادامه می‌یابد که شرط همگرایی محقق شود (۱۲). شرط همگرایی زمانی محقق می‌شود که خطای مجذور که با استفاده از رابطه (۱) به دست می‌آید به حداقل مقدار خود رسیده باشد:

$$J(X; V) = \sum_{i=1}^c \sum_{k=i}^k \|x_k^{(i)} - v_i\| \quad (1)$$

در این رابطه c تعداد خوشه‌ها، $X_k^{(i)}$ فاصله متغیر k ام از خوشه i ام از مبدا مختصات بوده و V_i فاصله مرکز جرم خوشه i ام از مبدا مختصات می‌باشد.

روش خوشه‌بندی K-medoids: یکی دیگر از روش‌های خوشه‌بندی تجزیه‌ای، روش K-medoids است. در این روش، ابتدا مشاهدات به K خوشه تقسیم می‌شوند و بعد از آن برای هر خوشه‌ی t ($t=1, 2, \dots, k$) یک نمایانگر (m_t) انتخاب می‌شود. مشاهده‌ای در خوشه t به عنوان نمایانگر انتخاب می‌شود که عبارت $\sum_i d(c_i, m_t)$ را حداقل کند. $d(c_i, m_t)$ فاصله اقلیدسی بین متغیر i ام در خوشه t و نمایانگر خوشه‌ی t ام بوده و C_i مشاهده‌ی i ام می‌باشد (۲۱).

شاخص Silhouette: این شاخص جهت ارزیابی نحوه‌ی خوشه‌بندی با استفاده از روش‌های مختلف خوشه‌بندی، مورد استفاده قرار می‌گیرد. این شاخص با استفاده از رابطه‌ی (۲) محاسبه می‌شود:

هاسکینگ و والیس توسعه و تکامل یافت (۱۱). برتری این روش نسبت به سایر گشتاورها، در عدم حساسیت به مقادیر پرت در داده‌ها است. این امر باعث شده است که روش گشتاورهای وزنی احتمال نتایج مطمئن‌تری از نمونه‌های کوچک بر اساس یک توزیع احتمال ارائه دهند (۱۰، ۱۱، ۱۵، ۱۴ و ۱۷).

گشتاور خطی نیز یک ترکیب خطی از داده‌های مشاهده‌ای تصادفی است که بر اساس گشتاورهای وزنی احتمال بوده و برای اولین بار در سال ۱۹۹۰ توسط هاسکینگ^۱ معرفی شد. این روش به دلیل قدرت بالایی که در آزمون همگنی نواحی، تحلیل منطقه‌ای و انتخاب بهترین توزیع ممکن برای برازش داده‌های هواشناسی و هیدرولوژیکی نظیر بارندگی‌ها، سیلاب‌های حداقل و حداکثر ماهانه یا سالانه دارد، کاربردهای وسیعی در علوم هیدرولوژی و منابع آب یافته است (۱۵).

مطالعات در زمینه تحلیل فراوانی منطقه‌ای متغیرهای هیدرولوژیکی از دهه ۱۹۸۰ آغاز شد و کلمس^۲ در مطالعه خود روش‌ها و مزایای تحلیل فراوانی سیل منطقه‌ای را بیان کرد (۲۲). آداموسکی^۳ به مقایسه روش‌های غیرپارامتری و گشتاورهای خطی در تحلیل منطقه‌ای سیلاب حوضه‌های واقع در استان‌های انتاریو و کبک کانادا پرداخت. وی در پژوهش خود، از داده‌های حداکثر سالیانه و سری جزئی سیلاب استفاده کرد و منطقه مورد مطالعه به نه ناحیه همگن تقسیم نمود به طوری که نواحی همگن مکانیزم تولید سیلاب مشابه داشتند (۴). کومار و کاتارجی^۴ با استفاده از روش گشتاورهای خطی به تحلیل فراوانی سیلاب منطقه‌ای پرداختند. نتایج نشان داد که توزیع مقادیر حدی تعمیم یافته بهترین برازش را به سیلاب حوضه آبریز براهاماپاتوی شمالی هند دارد (۲۳).

در ایران نیز تحقیقاتی در این زمینه صورت گرفته که به برخی از آنها اشاره می‌شود. شامکوئیان و همکاران به تحلیل فراوانی سیلاب منطقه‌ای با روش گشتاورهای خطی و سیلاب نمایه در حوضه‌های آبریز استان خراسان پرداختند. نتایج نشان داد که توزیع مقادیر حدی تعمیم‌یافته بر نواحی هفت گانه همگن، مناسب تشخیص داده شده است (۲). مطالعاتی نیز در این زمینه برای بارندگی و شاخص‌های خشکسالی انجام شده است. اسلامیان و فیضی، تحلیل فراوانی منطقه‌ای حداکثر بارش ماهانه را با روش گشتاور خطی در ۱۸ ایستگاه حوضه زاینده‌رود مورد بررسی قرار دادند. مطالعات آنها نشان داد که توزیع‌های مقادیر حدی تعمیم‌یافته و پیرسون نوع سه بهترین برازش را به داده‌ها داشته است (۸). در تحقیق دیگری نیز اسلامیان و همکاران به تحلیل فراوانی منطقه‌ای شاخص‌های خشکسالی کمبود بارش تجمعی و حداکثر کمبود بارش در استان اصفهان پرداختند و توزیع

- 1- Hosking
- 2- Kleme
- 3- Adamowski
- 4- Kumar and Chatterjee

که در آن τ معیار پراکنش خطی (LC_V) ، τ_3 معیار چولگی خطی (LC_S) و τ_4 معیار کشیدگی خطی (LC_K) می‌باشد. در روش گشتاورهای خطی، برآورد پارامترهای توزیع از طریق معادل قرار دادن چند گشتاور خطی اول داده‌ها با گشتاورهای نظیر مربوط به توزیع انجام می‌پذیرد (۱۱).

تحلیل فراوانی سیلاب منطقه‌ای

تحلیل منطقه‌ای سیلاب شامل مراحل غربال داده‌ها، تعیین نواحی همگن، انتخاب تابع توزیع فراوانی منطقه‌ای، تخمین پارامترهای توابع توزیع هر ناحیه همگن و تخمین دبی سیلاب اوج لحظه‌ای با دوره بازگشت‌های مختلف در نواحی همگن می‌باشد. در ادامه به شرح آزمون‌های هر مرحله پرداخته شده است.

آزمون ناجوری^۱: اندازه ناجوری، D_i که برحسب گشتاور خطی توسط هاسکینگ و والیس^۲ تعریف شده است، برای حذف ایستگاه‌های غیر عادی و همچنین جهت ارزیابی داده‌ها از لحاظ مناسب بودن برای تحلیل فراوانی منطقه‌ای مورد استفاده قرار می‌گیرد (۱۷). فرض کنید $u_i = [\tau_{i1}, \tau_{i3}, \tau_{i4}]^T$ برداری متشکل از نسبت‌های گشتاور خطی برای ایستگاه i باشد، در این صورت میانگین گروه \bar{u} و ماتریس کواریانس نمونه S است و اندازه ناجوری برای ایستگاه i با استفاده از رابطه زیر به دست می‌آید:

$$D_i = 3^{-1} (u_i - \bar{u})^T S^{-1} (u_i - \bar{u}) \quad (6)$$

که n تعداد کل ایستگاه‌ها است (۸). اگر آماره D_i برای یک ایستگاه بزرگتر از ۳ باشد، داده‌های مشاهداتی آن ایستگاه ناهماهنگ با دیگر ایستگاه‌های منطقه در نظر گرفته می‌شود (۸) و (۹).

آزمون ناهمگنی منطقه‌ای^۳: آزمون ناهمگنی گشتاورهای خطی با هدف برآورد درجه ناهمگنی در یک گروه از ایستگاه‌ها و تشخیص اینکه آیا آنها می‌توانند به طور معقول به عنوان یک گروه همگن در نظر گرفته شوند یا نه، انجام می‌شود (۱۶ و ۱۸).

برای آزمون همگنی، سه آزمون مختلف وجود دارد: آزمون اول بر پایه معیار پراکنش خطی است. آزمون دوم بر پایه پراکنش و چولگی خطی بوده و آزمون سوم بر پایه چولگی و کشیدگی خطی است. این آزمون‌ها، آماره V نامیده شده و روابط آنها در منابع ذکر شده تشریح شده‌اند. در نتیجه میزان ناهمگنی با توجه به رابطه (۷) محاسبه می‌شود:

$$s(i) = \frac{b(i) - \text{dist}(i; k)}{\max(b(i), \text{dist}(i; k))} \quad (2)$$

$$b(i) = \min(\text{dist}(i; G), G \neq k)$$

در این رابطه، $s(i)$ مقدار شاخص silhouette برای واحد i ام بوده و $\text{dist}(i; k)$ متوسط فاصله واحد i ام واقع در خوشه k ام از بقیه اعضای واقع در همین خوشه می‌باشد. مقادیر این شاخص از ۱ تا -۱ تغییر می‌کند. اگر مقدار این شاخص مثبت باشد خوشه بندی به خوبی انجام گرفته و چنانچه مقدار آن منفی باشد این امکان وجود دارد که واحد موردنظر به طور اشتباهی در خوشه قرار گرفته باشد (۲۵).

تئوری گشتاورهای خطی

هاسکینگ گشتاورهای خطی را معرفی کرد که توابعی از گشتاورهای وزنی احتمال می‌باشند. گشتاورهای خطی مناسب‌تر و آسان‌تر از گشتاورهای وزنی احتمال هستند، چون می‌توانند مستقیماً به عنوان اندازه‌هایی از مقیاس و شکل توزیع‌های احتمال تشریح شوند. بنابراین کاربرد آنها برای تحلیل فراوانی سیلاب (تعیین تابع توزیع مناسب و تخمین پارامترهای آن) مناسب است (۱۵ و ۱۷). گشتاورهای وزنی احتمالی که توسط گرینوود و همکاران (۱۱) تعریف شدند با توجه به رابطه (۳) تعیین می‌شوند:

$$\beta_r = E \left[x \{F(x)\}^r \right] \quad (3)$$

که $F=F(x)$ تابع توزیع تجمعی (CDF) برای متغیر x می‌باشد و $x(F)$ معکوس CDF می‌باشد. r نیز اعداد صحیح نامنفی می‌باشند. زمانی که $r = 0$ باشد، β_0 برابر با میانگین توزیع، $\mu = E[x]$ می‌باشد. گشتاورهای خطی که توسط هاسکینگ بر حسب گشتاورهای وزنی احتمال β مشخص شده‌اند به صورت رابطه (۴) محاسبه می‌شود:

$$\lambda_{r+1} = \sum_{k=0}^r \beta_k (-1)^{r-k} \binom{r}{k} \binom{r+k}{k} \quad (4)$$

طبق رابطه فوق، چهار گشتاور خطی نخست برای r صفر تا سه به دست می‌آید و نسبت گشتاورهای خطی به صورت روابط زیر بیان می‌شوند:

$$\begin{aligned} L - CV &= \tau = \lambda_2 / \lambda_1 \\ L - skew &= \tau_3 = \lambda_3 / \lambda_2 \\ L - kurtosis &= \tau_4 = \lambda_4 / \lambda_2 \end{aligned} \quad (5)$$

1-Discordancy
2- Hosking and Walls
3- Heterogeneity

طبق اظهار هاسکینگ، برازشی را می‌توان مناسب و صحیح دانست اگر مقدار Z^{Dist} به اندازه کافی به صفر نزدیک باشد. معیار معقول و مناسب برای این کار این است که $|Z^{Dist}| \leq 1.64$ باشد. زمانی که چند توزیع طبق معیار Z مناسب ارزیابی شدند، بهترین توزیع، توزیعی است که کمترین Z^{Dist} را دارا باشد (۱۹).

تخمین پارامترهای توزیع فراوانی منطقه‌ای: برای محاسبه گشتاورهای خطی استاندارد منطقه‌ای با توجه به مقادیر وزنی گشتاورهای خطی بی بعد، از رابطه زیر استفاده می‌شود:

$$\lambda_r^R = \frac{\sum_{i=1}^N n_i \lambda_r^{(i)}}{\sum_{i=1}^N n_i} \quad (10)$$

که $\lambda_r^{(i)}$ گشتاورهای خطی استاندارد منطقه‌ای با مرتبه r ، λ_r^R گشتاورهای خطی استاندارد شده با مرتبه r در ایستگاه i ، n_i تعداد سال داده‌ها در ایستگاه i و N تعداد ایستگاه‌ها در هر منطقه همگن می‌باشد. در نتیجه، پارامترهای بهترین توزیع منطقه‌ای تخمین زده شده و کمیت‌های متناظر با دوره‌های بازگشت مختلف برای متغیر مورد نظر به عنوان کمیت منطقه‌ای محاسبه می‌شود. تخمین کمیت‌های ایستگاهی با استفاده از رابطه زیر به دست می‌آید:

$$Q_i(F) = \lambda_1^{(i)} q(F) \quad (11)$$

که $Q_i(F)$ و $q(F)$ به ترتیب کمیت ایستگاهی در ایستگاه i و کمیت منطقه‌ای با احتمال عدم تجاوز F می‌باشد (۲۶).

نرم افزارهای مورد استفاده

در این تحقیق به منظور تحلیل خوشه‌بندی از نرم‌افزار RapidMiner5 استفاده شد. برنامه مذکور توانایی اجرای تحلیل خوشه‌ای و تعیین نواحی همگن بر اساس متغیرهای مورد نظر را دارد. جهت اجرای آزمون‌های ناچوری، ناهمگنی و تحلیل فراوانی

$$H_i = \frac{(V_i - \mu_{V_i})}{\sigma_{V_i}}; i=1,2,3 \quad (7)$$

که μ_{V_i} و σ_{V_i} به ترتیب میانگین و انحراف معیار مقادیر شبیه‌سازی شده V_i می‌باشند. شبیه‌سازی تعداد زیادی نواحی از طریق برازش توزیع کاپای چهار پارامتری بر مجموعه داده‌های منطقه‌ای انجام می‌شود و هاسکینگ و والیس (۱۹۹۳) تعداد نواحی شبیه‌سازی را حداقل ۵۰۰ ناحیه تشخیص دادند (۱۸). یک ناحیه، ناهمگن تلقی می‌شود اگر H_1 به اندازه کافی بزرگ باشد. هاسکینگ و والیس پیشنهاد کردند که اگر $H < 1$ باشد آنگاه ناحیه مورد نظر همگن خواهد بود و اگر $1 \leq H < 2$ ناحیه مورد نظر نسبتاً ناهمگن است. زمانی که $H \geq 2$ ناحیه کاملاً ناهمگن تلقی می‌شود. همچنین آنها مشاهده کردند که آماره H_1 بر اساس معیار V_1 شاخص بهتری برای تشخیص نواحی همگن است. بنابراین آماره H_1 به عنوان شاخص اساسی برای تشخیص ناهمگنی توصیه می‌شود (۱۶).

آزمون نکویی برازش منطقه‌ای: هنگامی که ایستگاه‌های موجود در یک ناحیه همگن بوده و از یک توزیع خاص پیروی کنند، در این صورت آزمون نکویی برازش مبتنی بر گشتاورهای خطی انجام می‌گیرد تا یکی از توزیع‌های رایج انتخاب و پارامترهای آن تخمین زده شود. معیار نکویی برازش برای هر توزیعی بر اساس آماره Z تعیین می‌شود:

$$Z^{Dist} = (\tau_4^{Dist} - \tau_4^R + B_4) / \sigma_4 \quad (8)$$

که τ_4^{Dist} میانگین کشیدگی خطی محاسبه شده با استفاده از شبیه‌سازی برای توزیع برازش داده شده منتخب و τ_4^R میانگین کشیدگی خطی محاسبه شده از داده‌های واقعی منطقه می‌باشند. B_4 و σ_4 به ترتیب اریب و انحراف استاندارد مقادیر کشیدگی هستند که از طریق شبیه‌سازی محاسبه می‌شوند:

$$B_4 = (N_{sim})^{-1} \sum_{m=1}^{N_{sim}} (\tau_4^{[m]} - \tau_4^R)$$

$$\sigma_4 = \left[(N_{sim})^{-1} \left\{ \sum (\tau_4^{[m]} - \tau_4^R)^2 - N_{sim} B_4^2 \right\} \right]^{0.5} \quad (9)$$

که $\tau_4^{[m]}$ میانگین کشیدگی منطقه‌ای می‌باشد که برای m امین منطقه شبیه‌سازی شده محاسبه شده است. N_{sim} تعداد منطقه شبیه‌سازی شده می‌باشد که با استفاده از توزیع کاپا تولید می‌گردد.

جدول ۱- فهرست ایستگاه‌های آب‌سنجی مورد استفاده واقع در حوضه ارومیه

ردیف	ایستگاه	رودخانه	کد ایستگاه	طول دوره آماری	ردیف	ایستگاه	رودخانه	کد ایستگاه	طول دوره آماری
۱	سهزاب	آغمیون چای	۳۱-۰۰۱	۳۳	۱۹	نظام آباد	زرینه رود	۳۳-۹۱۷	۱۴
۲	سرانسر	آجی چای	۳۱-۰۰۵	۲۲	۲۰	کوتر	مه‌باد چای	۳۴-۰۰۳	۲۳
۳	ونیار	آجی چای	۳۱-۰۱۵	۳۶	۲۱	پی قلعه	گادارچای	۳۴-۰۱۱	۳۸
۴	آخوله	آجی چای	۳۱-۰۴۵	۲۴	۲۲	نقده	گادارچای	۳۴-۰۱۹	۴۲
۵	سدعلویان	صوفی چای	۳۲-۰۰۷	۳۴	۲۳	پل بهراملو	گادارچای	۳۴-۰۲۱	۴۲
۶	چکان	چکان چای	۳۲-۰۱۱	۳۳	۲۴	قاسملو	بالانچ چای	۳۵-۰۰۱	۳۴
۷	ینگجه	قلعه چای	۳۲-۰۱۵	۲۴	۲۵	بابارود	باراندوز	۳۵-۰۰۷	۴۲
۸	شیشوان	قلعه چای	۳۲-۰۲۱	۲۴	۲۶	میرآباد	شهرچای	۳۵-۰۰۹	۳۱
۹	قشلاق امیر	مردوق چای	۳۳-۰۰۳	۳۳	۲۷	بند ارومیه	شهرچای	۳۵-۰۱۱	۴۲
۱۰	شیرین کند	لیلان چای	۳۳-۰۰۵	۳۴	۲۸	گوبجلی	نازلوچای	۳۵-۰۲۶	۲۵
۱۱	دریان	دریان چای	۳۸-۰۰۱	۲۵	۲۹	تپیک	نازلوچای	۳۵-۰۳۱	۴۲
۱۲	قبیلو	سقرچای	۳۳-۰۰۷	۲۳	۳۰	آباجالوسفلی	نازلوچای	۳۵-۰۳۳	۴۳
۱۳	پل آتیان	جیغاتوچای	۳۳-۰۱۵	۲۲	۳۱	کلهور	روضه چای	۳۵-۰۳۹	۳۱
۱۴	صفا خانه	ساروق چای	۳۳-۰۲۱	۲۰	۳۲	چهریق علیا	زولاچای	۳۶-۰۰۱	۲۳
۱۵	سنته	خرخره چای	۳۳-۹۱۹	۲۰	۳۳	نظر آباد	دریک	۳۶-۰۰۳	۲۱
۱۶	ساری قمیش	زرینه رود	۳۳-۰۲۳	۳۶	۳۴	تمر	خرخره	۳۶-۰۰۹	۲۹
۱۷	بوکان	سیمینه رود	۳۳-۰۳۵	۴۲	۳۵	یالفوز آغاج	زولاچای	۳۶-۰۱۱	۳۳
۱۸	تازه کند	سیمینه رود	۳۳-۰۳۹	۲۲	-	--	--	--	--

ایستگاه‌های ۱ تا ۱۱ در آذربایجان شرقی، ۱۲ تا ۱۵ در شمال کردستان و بقیه ایستگاه‌ها در آذربایجان غربی قرار دارند. به طور کلی طول دوره‌های آماری ایستگاه‌های مورد بررسی متفاوت بوده و بین ۱۴ تا ۴۳ سال متغیر می‌باشد.

نتایج و بحث

در این بخش ابتدا، نتایج حاصل از تحلیل خوشه‌بندی بر روی ایستگاه‌های آب‌سنجی جهت تعیین بهترین نواحی همگن بیان شده و سپس نتایج حاصل از تحلیل فراوانی منطقه‌ای دبی سیلاب حوضه آورده شده است.

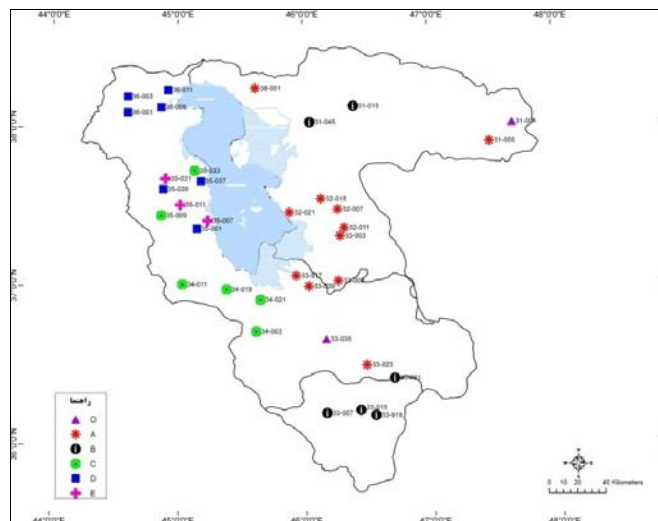
مقایسه روش‌های تحلیل خوشه‌بندی

در این تحقیق برای تعیین نواحی همگن با توجه به ویژگی مورد نظر، از روش‌های K-means و K-medoids استفاده شد. در هر کدام از روش‌ها، خوشه‌بندی حوضه با تعداد خوشه‌های سه، چهار و پنج مورد آزمایش قرار گرفت و نتایج آنها با توجه به شاخص Silhouette ارزیابی شدند. مطابق جدول (۲)، هر دو روش مقدار شاخص محاسبه شده آنها مثبت است و این مقدار شاخص در روش K-means در مقایسه با روش K-medoids برای تمام خوشه‌ها بیشتر می‌باشد. در نتیجه، تعداد خوشه پنج با روش K-means به عنوان مناسب‌ترین روش برای خوشه‌بندی نواحی همگن از لحاظ رفتار دبی حداکثر لحظه‌ای انتخاب شد. همچنین جدول (۳) و شکل (۱) به ترتیب شماره ایستگاه‌های واقع

منطقه‌ای از نرم‌افزارهای xfit.exe، xtest.exe و freq.exe که تحت کد فرترن^۱ نوشته شده‌اند، استفاده شد.

منطقه مورد مطالعه و داده‌ها

حوضه دریاچه ارومیه با عرض جغرافیایی $38^{\circ}30'$ و $35^{\circ}40'$ شمالی و طول جغرافیایی $44^{\circ}07'$ و $47^{\circ}53'$ شرقی، در شمال غرب کشور قرار گرفته است. این حوضه که در سه استان آذربایجان شرقی، غربی و کردستان واقع شده، دارای وسعت 51801 کیلومترمربع بوده و وسعت دریاچه نیز حدوداً 5750 کیلومترمربع می‌باشد. کوه‌هایی از غرب، جنوب و شرق دریاچه را احاطه نموده‌اند، ولی از سمت شمال ارتفاعات مهمی وجود ندارد. حداکثر ارتفاع در غرب 3608 متر، در شرق 3173 متر و در جنوب 4811 متر از سطح آب‌های آزاد می‌باشد. وضعیت افت تراز آب دریاچه و فرآیند تدریجی خشک شدن آن در سال‌های اخیر مشکلات مهم اقتصادی، زیست محیطی و اجتماعی برای منطقه و از طرفی نگرانی‌های ملی و بین‌المللی را نیز به همراه داشته است. در این حوضه رودخانه‌های زیادی جریان داشته که از مهمترین آنها می‌توان به زرینه رود، سیمینه رود و آجی‌چای اشاره کرد. در این مطالعه از آمار و اطلاعات 35 ایستگاه آب‌سنجی که از قدمت کافی و پراکنش مکانی مناسب در سطح کل حوضه برای تحلیل‌های مورد نظر برخوردار بودند، استفاده گردید. جدول (۱) فهرست ایستگاه‌های منتخب را نشان می‌دهد. مطابق این جدول



شکل ۱- موقعیت ایستگاه‌های آب‌سنجی منتخب در نواحی پنج‌گانه

جدول ۲- متوسط مقادیر شاخص Silhouette

تعداد خوشه‌ها			روش خوشه‌بندی
۵	۴	۳	
۰/۳۹	۰/۳۰	۰/۳۷	k-means
۰/۳۷	۰/۳۰	۰/۳۰	k-medoids

جدول ۳- ایستگاه‌های خوشه بندی شده هر ناحیه با استفاده از تحلیل خوشه‌ای

شماره ایستگاه‌ها*	تعداد ایستگاه‌ها	ناحیه
۱۹ و ۱۸، ۱۶، ۱۱، ۵، ۲	۱۱	A
۱۵ و ۴، ۳	۶	B
۲۰ و ۲۳، ۲۶ و ۳۰	۶	C
۳۵ و ۳۱، ۲۸، ۲۴	۷	D
۲۹ و ۲۷، ۲۵	۳	E

*بر اساس جدول شماره (۱)

شده در هر ناحیه همگن و موقعیت آنها را در سطح حوضه نشان می‌دهند. مطابق شکل (۱)، ایستگاه‌های مثلثی شکل سه‌زب و بوکان با استفاده از هر دو روش خوشه‌بندی در هیچ‌کدام از نواحی همگن تعیین شده قرار نگرفتند. همچنین ایستگاه‌های واقع در بخش‌های شرق، جنوب و شمال کردستان تراکم بیشتری دارند، اما ایستگاه‌های غرب و شمال‌غرب دریاچه میزان تراکم آنها کمتر است.

که آماره D آن از D بحرانی بزرگتر بود به عنوان ایستگاه ناجور تلقی گردید و در تحلیل منطقه‌ای کنار گذاشته شد و بقیه ایستگاه‌ها مورد آزمون همگنی قرار داده شدند. نتایج نشان داد که ایستگاه‌های سه‌زب و بوکان با آماره ناجوری ۳ و ۳/۱۹ نسبت به بقیه ایستگاه‌ها ناجور بوده و از گروه ایستگاه‌ها حذف شدند. در انتها آزمون‌های ناجوری برای هر منطقه همگن که با توجه به تحلیل خوشه‌بندی تعیین شده بودند، تکرار شد و مشاهده شد که ایستگاه ناجور در منطقه جدید وجود ندارد (جدول (۴)).

جدول (۵) نیز نتایج آزمون‌های H را در هر یک از نواحی نشان می‌دهد. مقادیر بزرگتر قدر مطلق H_1 در مقایسه با H_2 و H_3 نشان می‌دهد که غالباً چنانچه آماره H_1 همگنی نواحی را تأیید نماید، دو آماره دیگر نیز این مطلب را تأیید می‌کنند. این مورد در منابع دیگر توسط هاسکینگ و والیس نیز اشاره شده است و مقدار منفی آماره

شده در هر ناحیه همگن و موقعیت آنها را در سطح حوضه نشان می‌دهند. مطابق شکل (۱)، ایستگاه‌های مثلثی شکل سه‌زب و بوکان با استفاده از هر دو روش خوشه‌بندی در هیچ‌کدام از نواحی همگن تعیین شده قرار نگرفتند. همچنین ایستگاه‌های واقع در بخش‌های شرق، جنوب و شمال کردستان تراکم بیشتری دارند، اما ایستگاه‌های غرب و شمال‌غرب دریاچه میزان تراکم آنها کمتر است.

نتایج تحلیل فراوانی منطقه‌ای دبی سیلاب
آزمون‌های ناجوری و ناهمگنی

برای شناسایی ایستگاه‌های ناجور، آزمون D هاسکینگ و والیس به کار برده شد (۱۵). برای این منظور آماره D برای هر ایستگاه محاسبه و با مقدار بحرانی مقایسه گردید. هر ایستگاهی

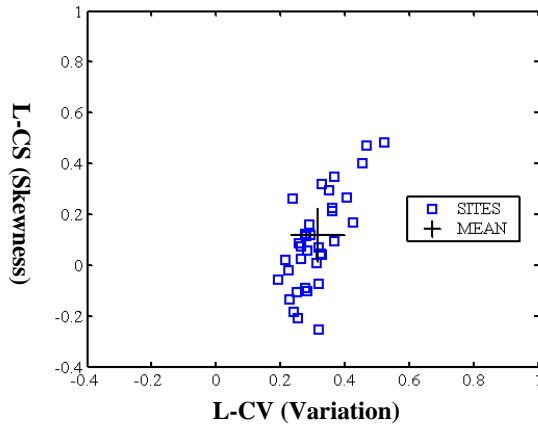
می‌باشد و از بین توزیع‌های برازش یافته، بهترین توزیع، توزیعی است که کمترین آماره Z را داشته باشد. بنابراین برای نواحی A، B و C تابع لجستیک تعمیم یافته و برای ناحیه D و E بترتیب توابع لوگ نرمال ۳ پارامتری و مقادیر حدی تعمیم یافته به عنوان مناسب‌ترین توابع برازش داده شده، انتخاب شدند. علاوه بر توابع توزیع مناسب برای هر ناحیه، تابع مقادیر حدی تعمیم یافته به دلیل اینکه مقادیر Z هاسکینگ آن برای تمامی نواحی کمتر از $1/64$ می‌باشد و بر تمام نواحی برازش مناسب یافته است. همچنین مطابق شکل (۳)، که بیان کننده نسبت گشتاورهای خطی (LCS) در برابر (LCK) است، مناسب بودن توزیع مقادیر حدی تعمیم یافته را برای کل منطقه را تأیید می‌کند. به عبارت دیگر می‌توان برای

H_1 نشان می‌دهد که در نواحی همگن، میزان پراکندگی داده‌ها کمتر از مقدار مورد انتظار می‌باشد (۱۷). لازم به ذکر است که تمامی مراحل

آزمون ناهمگنی در هر یک از نواحی پنج‌گانه به طور مجزا اجرا شد و هر پنج منطقه از لحاظ دبی سیلاب، همگن بودند. همچنین شکل (۲)، که بیان کننده نسبت گشتاورهای خطی (L-CV) در برابر (L-CS) است، همگنی را تأیید می‌کند.

آزمون نکویی برازش برای تعیین مناسب‌ترین تابع توزیع در نواحی همگن

نتایج حاصل از آزمون نکویی برازش Z در جدول (۶) نشان داده شده است. بر اساس این جدول هر توزیع احتمالی که آماره Z آن کوچکتر از $1/64$ باشد به داده‌های منطقه‌ای قابل برازش



شکل ۲- نمودار نسبت گشتاورهای خطی LCv در برابر LCS

جدول ۴- نسبت‌های گشتاورهای خطی و آزمون ناجوری (D_i) برای دبی سیلاب نواحی پنج‌گانه

D_i	L-Kurt	L-Skew	L-CV	ایستگاه	ناحیه	D_i	L-Kurt	L-Skew	L-CV	ایستگاه	ناحیه
۱/۱۵	۰/۱۸۴	۰/۲۳۰	۰/۳۰۸	کو تر	C	۱/۷۶	۰/۴۴۹	۰/۵۲	۰/۳۵۹	سرانسر	A
۰/۸۹	۰/۰۵۵	۰/۰۹۹	۰/۲۳۷	پی قلعه		۱/۲۵	۰/۳۷۸	۰/۲	۰/۲۷۷	سدعلویان	
۱/۵۱	۰/۳۱۸	۰/۳۸۵	۰/۲۴۹	نقده		۰/۸۳	۰/۲۶۸	۰/۳۳۴	۰/۳۸۳	چکان	
۱/۳۹	۰/۱۸۶	۰/۰۶۶	۰/۲۰۱	پل بهراملو		۰/۷	۰/۱۶۹	۰/۰۲۱	۰/۲۶۱	ینگجه	
۰/۴۷	۰/۲۱۸	۰/۱۸۴	۰/۲۶۷	آباجالوسفلی		۰/۴۱	۰/۱۵۲	۰/۰۵۷	۰/۲۷۵	شیشوان	
۰/۶۰	۰/۱۲۳	۰/۱۸۱	۰/۲۲۹	میرآباد		۱/۳۵	۰/۲۲۲	۰/۲۸۶	۰/۳۰۵	قشلاق‌امیر	
۱/۲۴	۰/۲۷۲	۰/۳۱۸	۰/۳۸۱	قاسملو	D	۱/۰	۰/۰۵۳	۰/۱۸۸	۰/۳۸۷	شیرین کند	B
۰/۴۷	۰/۱۵۵	۰/۱۴۷	۰/۳۳۵	گویجعلی اصلان		۰/۳۴	۰/۲۶۰	۰/۱۹۴	۰/۲۸۳	دریان	
۱/۵۷	۰/۰۹۰	۰/۲۶۴	۰/۴۵۱	کلهور		۱/۶۴	۰/۱۶۲	۰/۰۷۱	۰/۳۴۹	تازه کند	
۰/۱۸	۰/۱۲۱	۰/۱۹۸	۰/۳۵۵	چهریق علیا		۱/۱	-۰/۰۲۹	۰/۰۵۴	۰/۳۶۶	نظام آباد	
۱/۳۱	۰/۱۹۶	۰/۴۳۸	۰/۴۴۳	نظر آباد		۰/۶۱	۰/۰۵۳	۰/۰۰۵	۰/۲۹۱	ساری قمیش	
۱/۶۴	۰/۰۹۸	۰/۲۴۰	۰/۳۱۱	یالقوز آغاج		۰/۴۵	۰/۲۱۰	۰/۲۳۰	۰/۲۹۵	ونیار	
۰/۵۹	۰/۱۷۶	۰/۱۵۷	۰/۳۵۱	تمر	E	۱/۷۱	۰/۰۲۲	۰/۰۴۳	۰/۳۰۷	آخوله	
۱/۰	۰/۳۱۵	۰/۴۳۹	۰/۳۸۲	بابارود		۰/۸۷	۰/۳۱۱	۰/۳۱۳	۰/۳۰۱	قبقلو	
۱/۰	۰/۲۶۶	۰/۴۸۷	۰/۴۷۴	بندارومیه		۱/۵۹	۰/۱۹۴	۰/۰۶۵	۰/۲۴۷	پل آنیان	
۱/۰	۰/۲۶۰	۰/۳۸۶	۰/۳۶۵	تپیک		۱/۵۳	۰/۲۸۷	۰/۲۵۶	۰/۳۵۰	صفاخانه	
						۰/۳۸	۰/۱۰۳	۰/۰۹۹	۰/۲۷۶	سنه	

جدول ۵- نتایج آزمون همگنی، Hi، برای دبی سیلاب نواحی پنج گانه

میزان ناهمگنی			ناحیه
H ₃	H ₂	H ₁	
۲/۱۵	۱/۳۳	-۰/۱۷	A
۰/۲۶	۰/۰۱	-۰/۹۷	B
۰/۷۸	۰/۷۴	۰/۱۸	C
-۰/۱۰	۰/۳۱	۰/۶۲	D
-۱/۰۲	-۰/۵۸	۰/۱۹	E

جدول ۶- مقادیر آماره Z در نواحی پنج گانه

ناحیه	لجستیک تعمیم یافته	مقادیر حدی تعمیم یافته	لوگ نرمال سه پارامتری		پارتو تعمیم یافته
			پیرسون نوع سه	لوگ نرمال سه	
A	-۰/۷۳*	-۱/۱۹	-۲/۴۲	-۲/۹۸	-۵/۴۴
B	-۰/۰۹*	-۱/۱۲	-۱/۲۶	-۱/۶۱	-۳/۳۹
C	۰/۱۷*	-۱/۰۲	-۱/۲۵	-۱/۷۶	-۳/۷۱
D	۱/۶۰	۰/۷۶	۰/۲۸*	-۰/۵۹	-۱/۸۰
E	۰/۳۸	۰/۱۷*	-۰/۳۹	-۱/۳۶	-۰/۶۷

* مناسبترین تابع توزیع برازش یافته برای هر ناحیه

جدول ۷- پارامترهای منطقه‌ای و توابع توزیع تعیین شده برای نواحی پنج گانه

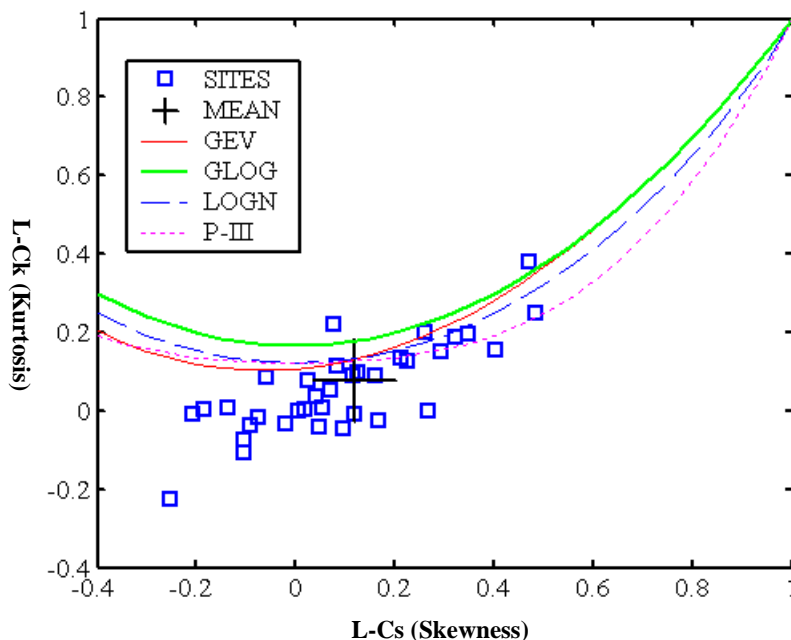
ناحیه	پارامترهای توزیع			تابع چندک بدون بعد ناحیه‌ای
	موقعیت	مقیاس	شکل	
A	۰/۹۰۶	۰/۳۰۴	-۰/۱۸۱	$Q_T = 0.906 - 1.679[1 - (T - 1)^{0.181}]$
B	۰/۹۱۷	۰/۲۸۲	-۰/۱۷۳	$Q_T = 0.917 - 1.63[1 - (T - 1)^{0.173}]$
C	۰/۹۲۵	۰/۲۳۱	-۰/۱۹۰	$Q_T = 0.925 - 1.215[1 - (T - 1)^{0.19}]$
D	۰/۸۳۵	۰/۵۹۳	-۰/۵۱۹	$Q_T = 0.835 + \exp(0.593 - 0.519 u)$
E	۰/۵۸۴	۰/۳۵۶	-۰/۳۷۹	$Q_T = 0.584 - 0.939 [1 - \{-\ln(1-1/T)\}^{-0.379}]$

A، B و C که بر توزیع یکسان برازش یافته‌اند، تقریباً مقادیر آنها نزدیک به یکدیگر است.

در نهایت مقادیر سیلاب با دوره بازگشت‌های مختلف برای توابع توزیع مناسبی که بر روی هر ایستگاه و ناحیه برازش یافتند، تخمین زده شد. مطابق جدول (۹)، دبی سیلاب ایستگاه‌های سه‌زاب و بوکان که به صورت ایستگاهی تحلیل فراوانی شدند، ذکر شده است. همچنین در جدول (۱۰)، دبی سیلاب نواحی پنج‌گانه به ازای دوره بازگشت‌های مختلف بیان شده است. مطابق این جدول حدود ۸۰ درصد ایستگاه‌های ناحیه A و تمام ایستگاه‌های نواحی B، C، D و E از نظر مکانیزم تولید سیلاب در هر ناحیه، یکسان هستند که این نتایج با مطالعات آداموسکی (۴) مطابقت دارد. به طور کلی در تمامی نواحی مورد مطالعه مقادیر

نواحی مورد مطالعه، تابعی واحد اما با پارامترهای مختلف پیشنهاد کرد. ایستگاه‌های سه‌زاب و بوکان نیز به دلیل ناچور بودن ایستگاه سه‌زاب توزیع لوگ نرمال دو پارامتری و برای ایستگاه بوکان تابع لوگ نرمال سه پارامتری به عنوان مناسب‌ترین تابع توزیع بر اساس آزمون کلموگروف-اسمیرنوف، انتخاب گردید. برآورد پارامترهای توابع توزیع نواحی همگن، با استفاده از روش پیشنهادی هاسکینگ صورت گرفت. پس از برآورد پارامترهای توابع توزیع، تابع چندک هر یک از نواحی پنج‌گانه و مقادیر کمی منحنی رشد (کمیت‌های منطقه‌ای) برای دوره بازگشت‌های مختلف مطابق جدول‌های (۷) و (۸) محاسبه شد. همچنین مطابق این جداول، پارامترهای توزیع و مقادیر کمی منحنی رشد نواحی

برآورد شده سیلاب اوج لحظه‌ای در دوره بازگشت‌های کم از نزدیکی خوبی با آمار واقعی برخوردار است.



شکل ۳- نمودار نسبت گشتاورهای خطی LCs در برابر LCK

جدول ۸- مقادیر کمی منحنی رشد بهترین توزیع انتخاب شده برای نواحی پنج گانه

ناحیه	نوع توزیع	مقادیر کمی منحنی رشد				
		۰/۹۹۹	۰/۹۹	۰/۹۵	۰/۹	۰/۵
A	لجستیک تعمیم یافته	۵/۰۸۴	۳/۰۸۲	۲/۰۸۷	۱/۷۲۵	۰/۹۰۶
B	لجستیک تعمیم یافته	۴/۶۷۱	۲/۸۹۶	۱/۹۹۹	۱/۶۷	۰/۹۱۷
C	لجستیک تعمیم یافته	۴/۲۲۳	۲/۶۱۸	۱/۸۳۵	۱/۵۵۴	۰/۹۲۵
D	لوگ نرمال ۳ پارامتری	۵/۳۷۲	۳/۵۱۳	۲/۳۷۵	۱/۹۱۴	۰/۸۳۵
E	مقادیر حدی تعمیم یافته	۱۲/۵۳۷	۵/۰۱۹	۲/۵۴۱	۱/۸۴۹	۰/۷۲۴

جدول ۹- دبی سیلاب (مترمکعب بر ثانیه) با دوره‌های بازگشت مختلف در ایستگاه سهزاب و بوکان

ایستگاه	دوره بازگشت					
	۱۰۰	۵۰	۲۰	۱۰	۵	۲
سهزاب	۷۷/۸۹	۵۸/۶۸	۳۸/۴۲	۲۶/۴۱	۱۶/۸۳	۷/۲۲
بوکان	۱۷۱۱/۶۷	۱۳۴۷/۴۹	۹۴۱/۲۱	۶۸۴/۲۶	۴۶۵/۱۱	۲۲۲/۲۲

جدول ۱۰- دبی سیلاب (مترمکعب بر ثانیه) با دوره‌های بازگشت مختلف در نواحی پنج‌گانه

ناحیه	ایستگاه	دوره بازگشت					
		۱۰۰	۵۰	۲۰	۱۰	۵	
A	سراتسر	۱۷۱/۶۳	۱۴۷/۰۶	۱۱۸/۱۳	۹۸/۲۶	۷۹/۲۷	۵۲/۱
	سدعلویان	۱۷۲/۱۲	۱۴۷/۴۸	۱۱۸/۴۷	۹۸/۵۵	۷۹/۵	۵۲/۲۵
	چکان	۳۹/۳۲	۳۳/۶۹	۲۷/۰۶	۲۲/۵۱	۱۸/۱۶	۱۱/۹۴۰
	ینگجه	۱۵۲/۰۴	۱۳۰/۲۸	۱۰۴/۶۵	۸۷/۰۵	۷۰/۲۳	۴۶/۱۶
	شیشوان	۲۰۵/۴۱	۱۷۶/۰۱	۱۴۱/۳۸	۱۱۷/۶	۹۴/۸۸	۶۲/۳۶
	قشلاق-امیر	۱۳۴/۴۵	۱۱۵/۲	۹۲/۵۴	۷۶/۹۸	۶۲/۱	۴۰/۸۲
	شیرین-کند	۱۶۳/۹۳	۱۴۰/۴۷	۱۱۲/۸۳	۹۳/۸۶	۷۵/۷۲	۴۹/۷۷
	دریان	۱۷/۷۳	۱۵/۱۹	۱۲/۲۰۰	۱۰/۱۵۰	۸/۱۹۰	۵/۳۸۰
	تازه-کند	۱۵۴/۵۱	۱۳۲/۴	۱۰۶/۳۵	۸۸/۴۶	۷۱/۳۷	۴۶/۹۱
	نظام-آباد	۱۰۸۶/۸۹	۹۳۱/۳۲	۷۴۸/۰۹	۶۲۲/۲۹	۵۰۲/۰۴	۳۲۹/۹۷
B	ساری-قمیش	۱۱۶۰/۰۱	۹۹۳/۹۸	۷۹۸/۴۲	۶۶۴/۱۵	۵۳۵/۸۱	۳۵۲/۱۷
	ونیار	۴۰۹/۹۸	۳۵۱/۴۳	۲۸۳/۰۴	۲۳۶/۴۸	۱۹۲/۳۲	۱۲۹/۸۱
	آخوله	۳۴۱/۶۷	۲۹۲/۸۷	۲۳۵/۸۸	۱۹۷/۰۷	۱۶۰/۲۷	۱۰۸/۱۸
	قیقلو	۳۸۴/۵۴	۳۲۹/۶۲	۲۶۵/۴۷	۲۲۱/۸	۱۸۰/۳۸	۱۲۱/۷۵
	پل-آنیان	۶۲۹/۲۶	۵۳۹/۴	۴۳۴/۴۲	۳۶۲/۹۶	۲۹۵/۱۸	۱۹۹/۲۴
	صفاخانه	۳۰۴/۶۷	۲۶۱/۱۶	۲۱۰/۳۴	۱۷۵/۷۳	۱۴۲/۹۲	۹۶/۴۷
	سنه	۲۹۶/۱۴	۲۵۳/۸۵	۲۰۴/۴۵	۱۷۰/۸۱	۱۳۸/۹۲	۹۳/۷۶
	کوتر	۳۱۱/۵۱	۲۶۸/۲۱	۲۱۸/۳۴	۱۸۴/۸۷	۱۵۳/۵۵	۱۱۰/۰۱
	پی-قلعه	۲۱۷/۶۴	۱۸۷/۳۹	۱۵۲/۵۴	۱۲۹/۱۶	۱۰۷/۲۸	۷۶/۸۶
	نقده	۲۶۲/۴۴	۲۲۵/۹۶	۱۸۳/۹۴	۱۵۵/۷۵	۱۲۹/۳۶	۹۲/۶۸
C	پل بهراملو	۲۶۰/۰۲	۲۲۳/۸۸	۱۸۲/۲۵	۱۵۴/۳۱	۱۲۸/۱۷	۹۱/۸۳
	آباجالوسفلی	۲۰۷/۰۷	۱۷۸/۲۹	۱۴۵/۱۴	۱۲۲/۸۹	۱۰۲/۰۷	۷۳/۱۳
	میرآباد	۱۱۱/۵۸	۹۶/۰۷	۷۸/۲۱	۶۶/۲۲	۵۵	۳۹/۴۱
	قاسملو	۶۳/۵۸	۵۴/۴۶	۴۲/۹۸	۳۴/۶۴	۲۶/۴۴	۱۵/۱۲
	گویجلی-اصلان	۶۱/۱۶	۵۲/۳۹	۴۱/۳۵	۳۳/۳۲	۲۵/۴۳	۱۴/۵۴
	کلهور	۱۳۵/۱۴	۱۱۵/۷۶	۹۱/۳۶	۷۳/۶۳	۵۶/۱۹	۳۲/۱۳
	چهریق-علیا	۲۲۸/۰۷	۱۹۵/۳۵	۱۵۴/۱۹	۱۲۴/۲۷	۹۴/۸۴	۵۴/۲۳
	نظرآباد	۷۸/۵	۶۷/۲۴	۵۳/۰۷	۴۲/۷۷	۳۲/۶۴	۱۸/۶۶
	یالقوزآغاج	۱۰۵/۵۸	۹۰/۴۳	۷۱/۳۸	۵۷/۵۳	۴۳/۹	۲۵/۱
	نمر	۶۳/۵۴	۵۴/۴۳	۴۲/۹۶	۳۴/۶۲	۲۶/۴۲	۱۵/۱۱
E	بابارود	۴۵۱/۵۲	۳۳۹/۰۵	۲۲۸/۶۱	۱۶۶/۳۶	۱۱۷/۲۳	۶۵/۰۹
	بندارومیه	۴۴۴/۸۶	۳۳۴/۰۵	۲۲۵/۲۴	۱۶۳/۹۱	۱۱۵/۵۱	۶۴/۱۴
	تپیک	۵۴۷/۳۳	۴۱۰/۹۹	۲۷۷/۱۲	۲۰۱/۶۶	۱۴۲/۱۱	۷۸/۹۱

مورد استفاده قرار گرفت و اطلاعات آنها برای انجام خوشه‌بندی و تحلیل فراوانی نواحی همگن مورد ارزیابی قرار گرفت. نتایج زیر از این تحقیق قابل ارائه می‌باشد:

نتیجه‌گیری

تحقیق حاضر تلاشی بود برای تحلیل فراوانی منطقه‌ای دبی سیلاب حداکثر در نواحی پنج‌گانه حوضه دریاچه ارومیه که به صورت همگن تقسیم شدند. بدین منظور ۳۵ ایستگاه آب‌سنجی

مناسب‌ترین تابع توزیع احتمالاتی برای تخمین دبی سیلاب اوج لحظه‌ای با دوره بازگشت‌های مختلف، انتخاب شدند.

۶- برای ایستگاه‌های سه‌زاب و بوکان که به صورت ایستگاهی تحلیل فراوانی شدند، به ترتیب توابع لوگ نرمال دو و سه پارامتری با توجه به آزمون‌های نکوئی برازش کلموگروف-اسمیرنوف و کای اسکوتر مناسب تشخیص داده شدند.

۷- برای تمام نواحی پنج‌گانه توابع توزیع سه پارامتری بهتر برازش یافته‌اند و تابع مقادیر حدی تعمیم‌یافته نیز برای تمامی نواحی در سطح حوضه، مناسب تشخیص داده شد. بنابراین با استفاده از نتایج به دست آمده و به کمک این مدل‌ها می‌توان در هر نقطه از حوضه آبریز به خصوص نواحی که فاقد آمار کافی جهت مطالعات هیدرولوژی هستند از روابط چندک‌های بی‌بعد ناحیه‌ای استفاده کرد. قابل ذکر است که استفاده از یک نوع تابع توزیع برای تمامی نواحی مورد مطالعه، کاربردی‌تر و مؤثرتر واقع شود.

۱- مقایسه بین دو روش خوشه‌بندی K-means و K-medoids با تعداد خوشه‌های سه، چهار و پنج نشان داد که مقدار شاخص Silhouette در روش K-means با هر تعداد خوشه بیشتر است، بنابراین دقت طبقه‌بندی آن بالاتر است.

۲- برای تشخیص نواحی همگن در حوضه مورد مطالعه، روش K-means با تعداد خوشه پنج، با توجه به شاخص Silhouette، مناسب‌ترین حالت خوشه‌بندی جهت منطقه‌بندی حوضه انتخاب شد.

۳- نواحی همگن حاصل از تحلیل خوشه‌ای با آزمون ناهمگنی هاسکینگ مورد بررسی قرار گرفتند و در نتیجه صحت همگنی این نواحی پنج‌گانه تأیید شد. همچنین، بخش شرقی دریاچه و شمال استان کردستان ایستگاه‌های همگن با تراکم بیشتر را نشان می‌دهند.

۴- از طرفی دیگر، کاربرد آزمون همگنی و آماره‌های آن بدون استفاده از دیاگرام نسبت گشتاورهای خطی LCV-LCS، ممکن است نتایج همگنی درستی را ارائه ندهند.

۵- با استفاده از آزمون نکوئی برازش هاسکینگ، توابع لجستیک تعمیم‌یافته در نواحی A، B و C، لوگ نرمال سه پارامتری در ناحیه D و مقادیر حدی تعمیم‌یافته در ناحیه E به عنوان

منابع

- ۱- اسلامیان، س.، مدرس، ر. و س. سلطانی. ۱۳۸۵. گروه‌بندی مکانی خشکسالی با استفاده از شاخص استاندارد بارش در استان اصفهان. یادداشت فنی. نشریه آب و فاضلاب، جلد ۱۷، شماره ۵۷، صفحات ۷۵-۷۲.
- ۲- شامکوئیان، ح.، قهرمان، ب.، داوری، ک. و م. سرمد. ۱۳۸۸. تحلیل فراوانی سیلاب منطقه‌ای با استفاده از تئوری گشتاورهای خطی و سیلاب نمایه در حوضه‌های آبریز استان‌های خراسان. مجله آب و خاک (علوم و صنایع کشاورزی)، جلد ۲۳، شماره ۱، صفحات ۴۳-۳۱.
- ۳- هاشمی، س. م. ۱۳۸۷. خوشه‌بندی زمانی و مکانی شبکه‌های آبیاری با استفاده از روش‌های کلاسیک و فازی (مطالعه موردی شبکه آبیاری قزوین). پایان‌نامه کارشناسی ارشد، دانشکده کشاورزی، دانشگاه تربیت مدرس.
- 4- Adamowski, K. 2000. Regional analysis of annual maximum and partial duration flood data by nonparametric and l-moment methods. *Journal of Hydrology*, 229: 219-231.
- 5- Bhaskar, N. R., and C. A. O'Connor. 1989. Comparison of method of residuals and cluster analysis for flood regionalization. *Journal of Water Resources Planning and Management*, ASCE. 115(6): 793-808.
- 6- Burn, D. H. 1989. Cluster analysis as applied to regional flood frequency. *Journal of Water Resources Planning and Management*, ASCE. 115(5): 567-582.
- 7- Cunnane, C. 1989. Statistical distributions for flood frequency analysis. *World Meteorological Organization Operational Hydrology. Report No. 33, WMO-No. 718, Geneva, Switzerland*. edn, Sussex : J.Wiley and Sons. England.
- 8- Eslamian, S. S. and H. Feizi. 2007. Maximum monthly rainfall analysis using l-moments for an arid region in Isfahan province, Iran. *Journal of Applied Meteorology and Climatology*. 46(4): 494-503.
- 9- Eslamian, S., Hassanzadeh, H., Abedi-Koupai, J. and M. Gheysari. 2011. Application of l-moments for regional frequency analysis of monthly drought Indexes. *Journal of Hydrologic Engineering*. 17(1): 32-42.

- 10- Gommes, R. and F. Petrassi. 1994. Rainfall variability and drought in Sub-Saharan Africa since 1960. Agrometeorology Series Working Paper (FAO). No. 9. Roma(Italia).
- 11- Greenwood, J. A., Landwehr, J. M. Matalas, N. C. and J. R. Wallis. 1979. Probability weighted moments: Definition and relation to parameters of several distributions expressible in inverse form. *Water Resour. Res.*, 15: 1049–1054.
- 12- Han, J. and M. Kamber. 2006. *Data mining: Concepts and techniques*. Morgan Kaufmann Publishers, 2nd Edn., New Delhi, ISBN: 978-81-312-0535-8.
- 13- Hoppner, F., klawonn, F., kruse, R. and R. Thomas. 1999. *Fuzzy cluster analysis: methods for classification, data analysis and image recognition*. John.
- 14- Hosking, J. R. M. 1986. The theory of probability weighted moments. Research. Rep. RC, 12210, 160. IBM Research Division, Yorktown Heights, New York.
- 15- Hosking, J. R. M. 1990. L-moments: Analysis and estimation of distributions using linear combinations of order statistics. *Journal of Royal Statistical Society. Series B*, 52(1): 105–124.
- 16- Hosking, J. R. M. 1991. Approximations for use in constructing L-moment ratio diagrams. Research. Rep. RC 16635, IBM Research Division, TJ Watson Research Center.
- 17- Hosking, J. R. M. and J. R. Wallis. 1993. Some statistics useful in regional frequency analysis. *Water Resour. Res.* 29(2): 271–281.
- 18- Hosking, J. R. and J. R. Wallis. 1997. *Regional Frequency Analysis: An Approach based on l-moments*. Cambridge University Press, Cambridge, UK, 244 p.
- 19- Kaufman, L. and P. Rousseeuw. 1990. *Finding groups in data: an introduction to cluster analysis*. John Wiley & Sons, Interscience Publication. 368 p.
- 20- Klemes, V. 1988. A hydrological perspective. *Journal of Hydrology*. 100(1): 3-28.
- 21- Kumar, R. and C. Chatterjee. 2005. Regional flood frequency analysis using L-Moments for North Brahmaputra region of India. *Journal of Hydrologic Engineering*, 10: 1-7.
- 22- Mosley, M. P. 1981. Delimitation of New Zealand hydrologic regions. *Journal of Hydrology*, 49: 173-192.
- 23- Pollard, K. S. and M. Van Der Laan. 2002. A method to identify significant clusters in gene expression data.
- 24- Rooy, M. and P. Van. 1965. A rainfall anomaly index independent of time and space. *Notos* 14: P. 43.
- 25- Santos, J. F., Pulido-Calvo, I. and M. Manuela Portela. 2010. Spatial and temporal variability of droughts in Portugal. *Water Resources Research*. 46(3), W03503.
- 26- Velmurugan, T. and T. Santhanam. 2010. Computational complexity between k-means and k-medoids clustering algorithms for normal and uniform distributions of data points. *Journal of Computer Science*. 6 (3): 363-368.