



دانشگاه گوارز و منابع آب

نشریه پژوهش‌های حفاظت آب و خاک
جلد بیست و یکم، شماره اول، ۱۳۹۳
<http://jwsc.gau.ac.ir>

تلفیق روش‌های خوشه‌بندی فازی و گشتاورهای خطی در تحلیل فراوانی منطقه‌ای سیل (مطالعه موردی: خراسان رضوی)

*اسماعیل دودانگه^۱، محمدباقر ناطقی^۲، هادی رهنما^۳ و ابوالقاسم دادرسی سبزواری^۴

^۱ دانشجوی دکتری گروه مهندسی آبخیزداری، دانشگاه علوم کشاورزی و منابع طبیعی ساری،

^۲ دانش‌آموخته کارشناسی ارشد گروه مهندسی منابع آب، وزارت دفاع و پشتیبانی نیروهای مسلح،

^۳ دانش‌آموخته کارشناسی ارشد گروه مهندسی آبخیزداری، دانشگاه علوم کشاورزی و منابع طبیعی ساری،

^۴ دانش‌آموخته کارشناسی ارشد بخش حفاظت خاک و آبخیزداری، مرکز تحقیقات کشاورزی و منابع طبیعی خراسان رضوی

تاریخ دریافت: ۹۱/۸/۶؛ تاریخ پذیرش: ۹۱/۱۰/۲۹

چکیده

تعیین مناطق همگن هیدرولوژیک یکی از اساسی‌ترین گام‌ها در انجام تحلیل فراوانی منطقه‌ای است. در این مطالعه روش گشتاورهای خطی و تحلیل خوشه‌ای فازی برای تعیین مناطق همگن هیدرولوژیک و تحلیل فراوانی منطقه‌ای سیل در شمال‌شرق کشور به‌کار گرفته شد. در ابتدا با استفاده از آزمون همگنی گشتاورهای خطی (H)، کل منطقه مطالعه از نظر هیدرولوژیک ناهمگن تشخیص داده شد. با توجه به این‌که با حذف ایستگاه‌های ناجور همگنی قابل قبول برای منطقه موردنظر به‌دست نیامد بنابراین از روش تحلیل خوشه‌ای فازی برای دستیابی به مناطق همگن کوچک‌تر بهره گرفته شد. خوشه‌بندی براساس مشخصات فیزیوگرافیک و آماری حوزه‌ها انجام شد. نتایج نشان داد مشخصات فیزیوگرافیک حوزه‌ها رابطه معناداری با گشتاورهای وزنی احتمال ندارد و خوشه‌های ایجاد شده ناهمگن می‌باشند. با استفاده از شاخص‌های فازی DI ، PC ، SC ، S ، XB و ADI الگوی پراکنش گشتاورهای خطی در منطقه مورد بررسی قرار گرفت و در نهایت با استفاده از مقادیر این شاخص‌ها منطقه به سه بخش شمالی، میانی و جنوبی تقسیم گردید. آزمون نکوئی برازش با استفاده از آماره Z^{Dist} نشان داد دو توزیع پارتوی تعمیم‌یافته در بخش شمالی و میانی و توزیع لجستیک تعمیم‌یافته در بخش جنوبی به‌عنوان بهترین توابع توزیع منطقه‌ای مشخص گردیدند. با استفاده از توابع توزیع منطقه‌ای منحنی‌های رشد منطقه‌ای ترسیم گردید. نتایج نشان داد شیب منحنی‌های رشد منطقه‌ای از شمال به جنوب افزایش پیدا می‌کند.

واژه‌های کلیدی: همگنی هیدرولوژیک، شاخص‌های خوشه‌بندی فازی، گشتاورهای خطی، تابع توزیع منطقه‌ای

* مسئول مکاتبه: smaeel.dodangeh@gmail.com

مقدمه

خسارت‌های ناشی از سیلاب‌ها از جنبه‌های مختلف مانند از دسترس خارج کردن خاک سطحی و حاصل‌خیز، تخریب و از بین بردن اراضی زراعی و مناطق مسکونی، نابودی احشام و حتی گاهی انسان زندگی انسان در کشور بی‌سابقه نبوده است. برآورد پتانسیل وقوع سیل در یک منطقه، نیاز ضروری بسیاری از طرح‌های مرتبط با عملیات مهندسی رودخانه و برنامه‌ریزی‌های شهری می‌باشد. نیاز اساسی بسیاری از این طرح‌ها برآورد قابل اعتماد از مقادیر سیل^۱ با بهره‌گیری از تکنیک تحلیل فراوانی می‌باشد. تحلیل فراوانی سیل پایه و اساس طراحی‌های مهندسی و تحلیل اقتصادی بسیاری از پروژه‌های مرتبط با کنترل سیل می‌باشد (آنلی و همکاران، ۲۰۰۷). تحلیل فراوانی و برآورد سیلاب در حوزه‌هایی که آمار موجود است کار چندان دشواری نیست (مدرس، ۲۰۰۸) با این وجود بسیاری از مواقع طول دوره آماری داده‌های در دسترس در یک نقطه بسیار کوتاه‌تر از یک دوره بازگشت خاص و یا اصلاً موجود نمی‌باشد. در چنین مواردی تحلیل فراوانی منطقه‌ای می‌تواند برای دستیابی به مقادیر قابل اعتماد از سیل مورد استفاده قرار گیرد (شیو و اواردا، ۲۰۰۸). با استفاده از تحلیل فراوانی منطقه‌ای، برآوردها در یک نقطه، از طریق تلفیق^۲ داده‌های ثبت شده در آن نقطه با داده‌های ناحیه‌ای (داده‌های ثبت شده در ایستگاه‌های مجاور با طول دوره آماری بالا) که دارای توزیع مشابه با نقطه مزبور می‌باشند، بهبود می‌یابد (صاف، ۲۰۰۹). در این راستا، گرین‌وود و همکاران (۱۹۷۹) مفهوم گشتاورهای وزنی احتمال (PWM)^۳ را معرفی کردند. هاسکینگ (۱۹۹۰) گشتاورهای خطی را به‌عنوان ترکیب‌های خطی PWM تعریف کرد. هاسکینگ و والیس (۱۹۹۷) کاربرد گشتاورهای خطی را گسترش دادند و آماره‌های مفیدی را برای تحلیل فراوانی منطقه‌ای برای محاسبه ناهمگونی، همگنی ناحیه‌ای و نکوئی برازش به‌کار بردند و اکنون این روش به فراوانی توسط هیدرولوژیست‌ها در مطالعات مربوط به سیلاب‌ها در نقاط مختلف دنیا از جمله ایالات متحده (ووگل و ویلسون، ۱۹۹۶) نیوزلند (پیرسون، ۱۹۹۱؛ مادسن و همکاران، ۱۹۹۷)، هند (پاریدا و همکاران، ۱۹۹۸؛ کومار و همکاران، ۲۰۰۳)، مالزی (لیم و لای، ۲۰۰۳) و ترکیه (صاف، ۲۰۰۹) به‌کار گرفته می‌شود. در ایران نیز اسلامیان و بروجنی (۲۰۰۳)، از روش گشتاورهای خطی برای انجام تحلیل فراوانی منطقه‌ای سیل در حوزه آبخیز مرکزی ایران به‌کار گرفتند. ملکی‌نژاد و همکاران (۲۰۱۱) نیز با استفاده از این روش به

1- Flood Quantiles

2- Pooling

3- Flood Quantiles

تحلیل فراوانی منطقه‌ای سیل در حوزه دریاچه نمک مرکزی ایران پرداختند. سرحدی و همکاران (۲۰۱۲) روش گشتاورهای خطی را برای تحلیل فراوانی منطقه‌ای دبی‌های پیک سیلابی در حوزه هلیل رود جیرفت به‌کار بستند و بهترین توابع توزیع منطقه‌ای را در این منطقه تعیین کردند. باید در نظر داشت برآوردهای منطقه‌ای که با استفاده از این روش استخراج می‌شوند در صورتی قابل اعتماد خواهند بود که مناطق همگن از نظر پارامتر مورد بررسی وجود داشته باشند. بنابراین تشخیص مناطق همگن مهم‌ترین و مشکل‌ترین مرحله در انجام تحلیل فراوانی منطقه‌ای است (دیکباس و همکاران، ۲۰۱۱). حوزه‌هایی که براساس مشخصات جغرافیایی در یک گروه قرار می‌گیرند به‌طور قطعی نمی‌توانند از نظر هیدرولوژیک نیز همگن باشند. در چنین مواقعی تعمیم اطلاعات هیدرولوژیک از یک ایستگاه به ایستگاه دیگر در یک گروه مشترک، که با استفاده از این روش گروه‌بندی شده‌اند از نظر اصولی صحیح نمی‌باشد. برای غلبه بر این مشکل تئوری استفاده از خوشه‌بندی فازی (FCM)^۱ توسط پژوهش‌گران ارایه شده است. شیو و برن (۲۰۰۴)، از این روش برای تشخیص مناطق همگن هیدرولوژیک برای برآورد بزرگی سیلاب بهره گرفتند. دیکباس و همکاران (۲۰۱۱) نیز از این روش برای گروه‌بندی سری‌های بارش و تعیین مناطق همگن بارشی در ترکیه بهره گرفتند. در خصوص استفاده از روش گشتاورهای خطی در تحلیل فراوانی منطقه‌ای سیل در ایران مطالعات اندکی صورت گرفته است (سرحدی و همکاران، ۲۰۱۲). در زمینه استفاده از روش تحلیل خوشه‌ای فازی در زمینه هیدرواقلیم نیز تاکنون مطالعه‌ای در کشور گزارش نشده است. بنابراین در این مطالعه سعی خواهد شد تا با تلفیق دو روش گشتاورهای خطی و خوشه‌بندی فازی به تعیین مناطق همگن هیدرولوژیک و تحلیل فراوانی منطقه‌ای سیل در منطقه مطالعه پرداخته شود.

مواد و روش‌ها

منطقه مطالعه: استان خراسان رضوی در شمال‌شرق کشور واقع گردیده است. میانگین بارندگی در سطح استان ۲۰۸ میلی‌متر می‌باشد و بیش‌تر آن در زمستان و اوایل بهار صورت می‌گیرد. مناطق پرباران در نواحی مرتفع قرار دارند که از شمال به جنوب نیز میزان بارش کاهش می‌یابد. بارش در نواحی مرتفع بیش‌تر به‌صورت برف است که در تامین منابع آب استان به‌ویژه آب‌های زیرزمینی نقش مهمی دارد. نواحی کوهستانی شامل دو بخش شمالی و جنوبی می‌باشند که در بخش شمالی قسمت

1- Fuzzy C-Means

قابل توجهی از دو رشته کوه کپه‌داغ- هزارمسجد و آلاداغ- بینالود قرار دارند. رشته‌کوه هزارمسجد با حدود ۳۰۴۰ متر ارتفاع در جهت شمال‌غربی- جنوب‌شرقی امتداد یافته است (فلاح‌قاهری و همکاران، ۲۰۱۰).

برای انجام تحلیل فراوانی منطقه‌ای و انتخاب بهترین تابع توزیع منطقه‌ای، اطلاعات سیلاب حوزه ۵۲ ایستگاه هیدرومتری با طول دوره آماری بالای ۱۵ سال جمع‌آوری گردید (شکل ۱). در هر ایستگاه سری زمانی حداکثر سالانه (AMS)^۱ با انتخاب دبی پیک لحظه‌ای در هر سال تشکیل گردید.



شکل ۱- موقعیت منطقه مطالعه و ایستگاه‌های هیدرومتری منتخب.

تعریف گشتاورهای خطی: گشتاورهای خطی ترکیبات خطی از گشتاورهای وزنی احتمال (PWMS) (گرین‌وود و همکاران، ۱۹۷۹) می‌باشند. هاسکینگ و والیس (۱۹۹۳) گشتاورهای خطی را به شکل زیر تعریف کردند:

1- Annual Maximum Series (AMS)

$$\lambda_1 = \beta, \quad (1)$$

$$\lambda_2 = 2\beta_1 - \beta, \quad (2)$$

$$\lambda_3 = 6\beta_2 - 6\beta_1 + \beta, \quad (3)$$

$$\lambda_4 = 20\beta_3 - 30\beta_2 + 12\beta_1 - \beta, \quad (4)$$

که در آن، λ_r : گشتاورهای خطی و β_r : گشتاورهای وزنی احتمال می‌باشند که به شکل زیر تعریف می‌شوند:

$$\beta_r = \int_{-\infty}^{\infty} X(F) F^r dF \quad (5)$$

که در آن، F : احتمال وقوع نشدن می‌باشد. نمودار نسبت گشتاورهای خطی به شکل $\tau = \lambda_r / \lambda_1$ و $\tau_r = \lambda_r / \lambda_1$ ($r \geq 3$) تعریف می‌شوند که λ ، τ ، τ_3 و τ_4 به ترتیب پارامترهای موقعیت، مقیاس برای تشخیص مناطق همگن است. رائو و حامد (۱۹۹۷) از این روش برای تشخیص مناطق همگن هیدرولوژیک در حوزه رودخانه واپاش استفاده کردند.

آزمون همگنی و ناهمگنی: علاوه بر نمودار نسبت گشتاورهای خطی، هاسکینگ و والیس دو آماره نبود توافق (D_i) و ناهمگنی (H_i) را برای تشخیص مناطق همگن معرفی کردند. برای تشخیص مناطق همگن ابتدا آماره ناهمگنی (H_i) مورد بررسی قرار می‌گیرد. آماره ناهمگنی شامل سه آماره H_1 ، H_2 و H_3 می‌باشد. آماره H_1 که براساس ضریب تغییرات (L-CV) محاسبه می‌شود بسیار قوی‌تر از دو آماره دیگر است (هاسکینگ و والیس، ۱۹۹۳)، بنابراین در این مطالعه نیز از این آماره برای بررسی همگنی منطقه‌ای بهره گرفته شده است. یک منطقه همگن است اگر H_1 کمتر از ۱ باشد، اگر بین ۱ و ۲ باشد تا حدودی ناهمگن است و اگر بیش از ۲ باشد کاملاً ناهمگن است. در صورتی که منطقه ناهمگن بود ($H_1 > 2$) در این صورت آماره ناجوری (D_i) در تک تک ایستگاه‌ها مورد محاسبه قرار می‌گیرد. ایستگاهی که در آن میزان این آماره بیش از ۳ باشد به‌عنوان ایستگاه ناجور شناخته می‌شود و بنابراین باید از مجموعه ایستگاه‌های انتخاب شده حذف شود. در صورتی که با حذف ایستگاه ناجور، همگنی

قابل قبول برای منطقه مورد نظر به دست آمد در این صورت با استفاده از آزمون نکوئی برازش تابع توزیع منطقه‌ای به داده‌ها برازش داده می‌شود.

آزمون نکوئی برازش: برای تشخیص بهترین تابع توزیع منطقه‌ای آزمون نکوئی برازش به کار گرفته می‌شود. آماره نکوئی برازش (Z^{Dist}) به شکل زیر تعریف می‌شود:

$$Z^{dist} = \frac{\bar{t}_\xi - \tau_\xi^{dist}}{\sigma_\xi} \quad (6)$$

که در آن، \bar{t}_ξ : متوسط منطقه‌ای ضریب کشیدگی خطی، τ_ξ^{dist} : ضریب کشیدگی خطی توزیع برازش داده شده به داده‌ها و σ_ξ : انحراف معیار ضریب کشیدگی خطی برازش داده شده است. برای انتخاب بهترین تابع توزیع منطقه‌ای، گشتاورهای توزیع‌های مختلف با گشتاورهای متوسط نمونه‌ای مورد مقایسه قرار می‌گیرد. تابع توزیعی که در آن $|Z^{dist}| \leq 1/64$ می‌تواند به عنوان توزیع منطقه‌ای انتخاب شود. در مواردی که بیش از یک توزیع شرایط بیان شده را داشته باشند در این صورت توزیعی که کم‌ترین مقدار (Z^{Dist}) را داشته باشد به عنوان بهترین توزیع منطقه‌ای انتخاب خواهد شد.

خوشه‌بندی فازی: الگوریتم خوشه‌بندی FCM اصلاح شده الگوریتم خوشه‌بندی K میانگین^۱ می‌باشد. این الگوریتم واریانس داخل خوشه را به حداقل می‌رساند (آیوازا و همکاران، ۲۰۰۷). فرض الگوریتم FCM بر این است که داده‌ها در یک فضای برداری هستند و هدف آن به حداقل رساندن مجموع واریانس داخل خوشه D_v می‌باشد.

$$D_v = \sum_{k=1}^K \sum_{x_j \in S_k} |x_j - c_k|^2 \quad (7)$$

که در آن، c_k : مرکز یا میانگین نقاط در خوشه k ، K : مجموع تعداد خوشه‌ها، S_k : مجموعه نقاط در کلاستر k ام و x_j : بردار استاندارد در نقطه j می‌باشد. در این روش ابتدا به طور تصادفی k خوشه انتخاب می‌شود، سپس میانگین هر خوشه به دست می‌آید. در ضمن FCM برای هر داده یک درجه عضویت تعیین می‌کند و با به‌هنگام کردن مراکز و درجه عضویت‌ها برای هر داده، FCM مراکز خوشه‌ها را تصحیح می‌کند. این تکرار بر پایه کمینه کردن تابع هدف می‌باشد که در این جا تابع هدف،

فاصله هر داده از مرکز خوشه می‌باشد، که البته این مراکز خوشه‌ها به وسیله درجه عضویت هر داده وزن‌دهی می‌شوند. بر خلاف روش خوشه‌بندی k میانگین که هر داده فقط به یک خوشه تعلق می‌گیرد، در روش FCM عضویت جزئی مطرح می‌باشد به این معنی که هر داده با درجه‌ای از عضویت به تمامی خوشه‌ها تعلق دارد. بنابراین داده‌ای که در مرز یک خوشه قرار می‌گیرد با درجه عضویت کم‌تری نسبت به داده‌ای که به مرکز خوشه نزدیک‌تر است مشخص می‌گردد. درجه عضویت ایستگاه i به خوشه k ام، با فاصله این ایستگاه تا مرکز خوشه نسبت عکس دارد:

$$b_k(i) = \frac{1}{d(c_{k,i})} \quad (8)$$

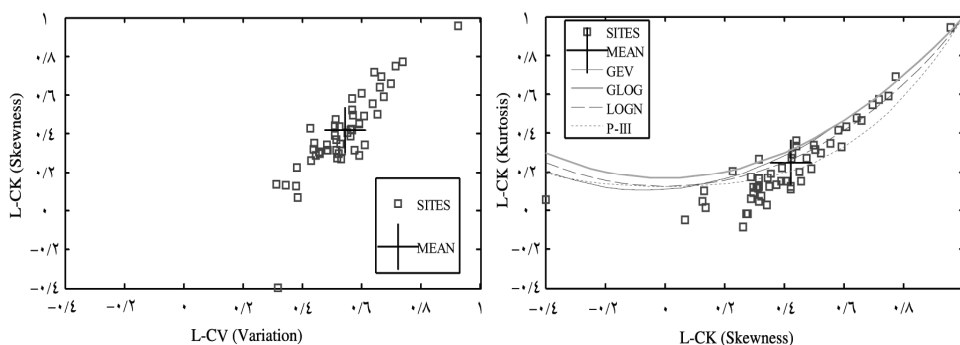
که در آن، $b_k(i)$: درجه عضویت ایستگاه i به خوشه k ام و $d(c_{k,i})$: فاصله ایستگاه i تا مرکز خوشه i می‌باشد. هر ایستگاه به خوشه‌ای تعلق می‌گیرد که بیش‌ترین درجه عضویت را به آن خوشه دارد. درجه‌های عضویت ضرایب نرمال شده هستند به طوری که مجموع درجه عضویت هر ایستگاه به تمام خوشه‌ها برابر با واحد خواهد بود.

$$\forall_i = \sum_{k=1}^K U_k(i) = 1 \quad (9)$$

که در آن، $U_k(i)$: درجه عضویت استاندارد ایستگاه i در خوشه k ام می‌باشد ($\in [0,1]$). در صورت ناهمگنی منطقه‌ای (براساس آماره همگنی هاسکینگ و والیس H)، با استفاده از درجه عضویت محاسبه شده برای ایستگاه‌ها و مقادیر گشتاورهای خطی محاسبه شده در هر ایستگاه می‌توان به طور دستی ایستگاه‌ها را بین خوشه‌ها از یک خوشه به خوشه دیگر منتقل کرد.

نتایج و بحث

گام اصلی در تحلیل فراوانی منطقه‌ای تشخیص مناطق همگن است. در این مطالعه برای بررسی همگنی منطقه موردنظر از نمودار نسبت گشتاورهای خطی و نیز آماره همگنی (H_1) استفاده شد. نمودار نسبت گشتاورهای خطی (شکل ۲) و نیز میزان آماره ناهمگنی هاسکینگ و والیس (۱۹۹۳) نشان‌دهنده ناهمگنی منطقه مورد مطالعه می‌باشد (جدول ۱). نتایج نشان داد پس از حذف دو ایستگاه ناجور کلاته‌منار و حسین‌آباد با مقادیر آماره ناجوری بیش از حد بحرانی ($D > 3$) به ترتیب برابر با ۱۱/۷۳ و ۶/۷۶ منطقه هم‌چنان ناهمگن باقی می‌ماند.



شکل ۲- نمودار نسبت گشتاورهای خطی L-Cs و L-Ck (سمت راست) و L-Cv و L-Cs (سمت چپ) در منطقه مورد مطالعه.

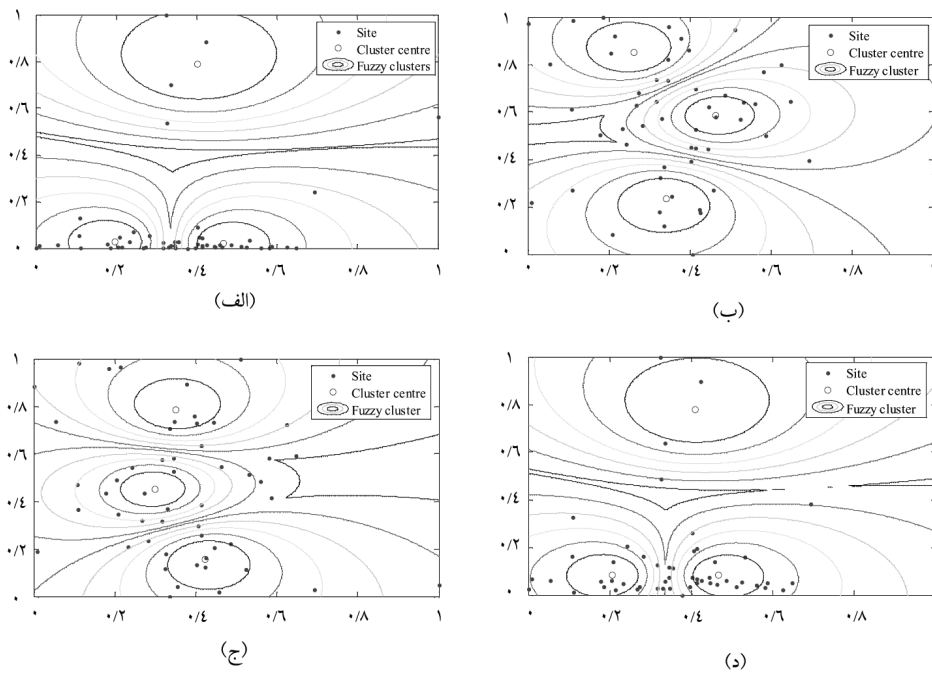
تعیین مناطق همگن با استفاده از خوشه‌بندی فازی: با توجه به ناهمگنی منطقه مطالعه سعی شد با استفاده از مؤلفه‌های فیزیکی^۱ (مساحت، ارتفاع ماکزیمم، ارتفاع متوسط، ارتفاع مینیمم، شیب متوسط، طول بلندترین آبراهه اصلی، ضریب شکل، ضریب شکل معکوس، ضریب فشردگی، طول مستطیل معادل، عرض مستطیل معادل، زمان تمرکز و بارش متوسط حوزه‌ها) و آماری^۲ (آماره‌های مربوط به گشتاورهای خطی سری‌های زمانی دبی اوج) حوزه‌ها و با به‌کارگیری روش تحلیل خوشه‌ای فازی FCM، کل منطقه مطالعه به مناطق کوچک‌تر و همگن تقسیم‌بندی شود.

ابتدا تعداد خوشه‌ها $c=3$ انتخاب گردید و کل منطقه مطالعه براساس ورودی‌های نام‌برده به ۳ بخش تقسیم‌بندی شد. از دیدگاه هاسکینگ و والیس (۱۹۹۷)، یک منطقه همگن شامل گروهی از ایستگاه‌ها می‌باشد که در آن داده‌های هیدرولوژیک در تمامی ایستگاه‌ها از توزیع واحدی پیروی می‌کنند. بنابراین دوباره همگنی خوشه‌های ایجاد شده با استفاده از آماره H_1 مورد بررسی قرار گرفت. نتایج نشان داد خوشه‌های ایجاد شده به‌طور کامل ناهمگن می‌باشد. با افزایش تعداد خوشه‌ها نیز میزان همگنی تغییر چندانی نداشت و مناطق ایجاد شده ناهمگن باقی ماندند. در تفسیر این نتیجه می‌توان گفت ارتباط مشخصی بین مشخصات فیزیوگرافیک حوزه‌ها و گشتاورهای وزنی احتمال وجود ندارد بنابراین گروه‌بندی مناطق براساس مشخصات فیزیوگرافیک ممکن است به ایجاد مناطق همگن هیدرولوژیک

1- Site Characteristics

2- At-Site Statistics

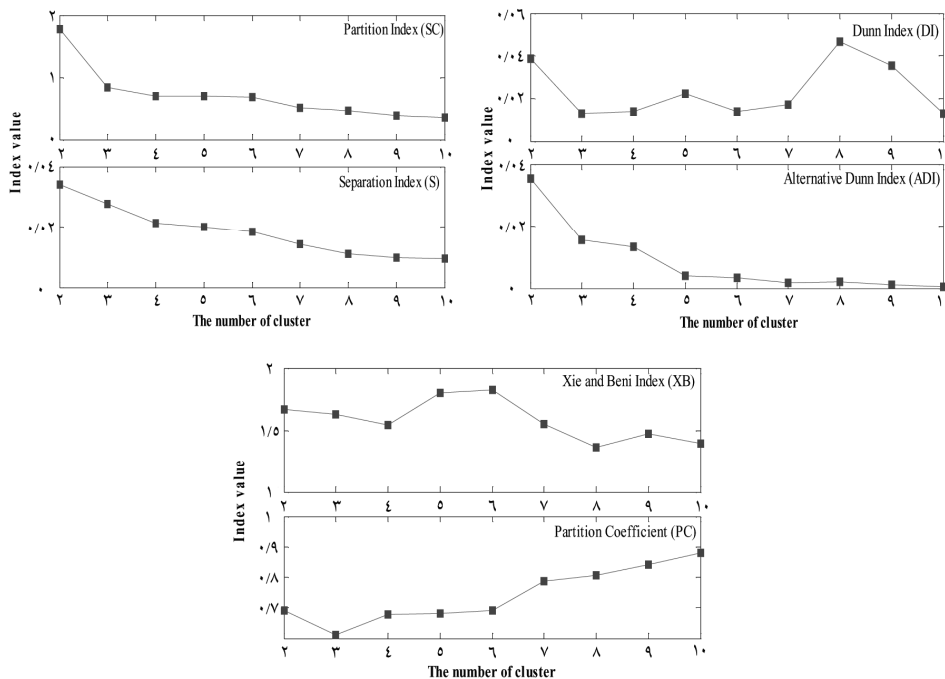
نینجامد به خصوص اگر تغییرپذیری خصوصیات هیدرولوژیکی منطقه زیاد باشد. دیکباس و همکاران (۲۰۱۱) بیان نمودند برای دستیابی به مقادیر قابل اعتماد در تحلیل فراوانی منطقه‌ای، گروه‌بندی مناطق باید براساس تشابهات هیدرولوژیک باشد. شکل ۳ ارتباط بین بعضی از مشخصات اصلی فیزیوگرافیک و اقلیمی حوزه‌ها و گشتاورهای خطی را در خوشه‌بندی فازی ایستگاه‌ها نشان می‌دهد. با توجه به شکل خوشه‌های ایجاد شده براساس مشخصات فیزیوگرافیک- هیدرولوژیک حوزه‌ها از ساختار مشخصی پیروی نمی‌کند بنابراین علت ناهمگنی هیدرولوژیک در خوشه‌های ایجاد شده مشخص می‌شود.



شکل ۳- ارتباط بین متغیرهای فیزیوگرافیک حوزه‌ها (الف) مساحت (ب) ارتفاع متوسط (ج) شیب متوسط (د) زمان تمرکز (به روش کریچ) و گشتاورهای خطی در خوشه‌بندی فازی ایستگاه‌ها.

برای تعیین مناطق همگن هیدرولوژیک در منطقه مطالعه، سعی گردید گشتاورهای خطی محاسبه شده از سری زمانی داده‌های دبی اوج در هر یک از ایستگاه‌ها با استفاده از الگوریتم خوشه‌بندی FCM، به گروه‌های مشابه تقسیم‌بندی گردد. در منابع مختلف در خصوص تعیین تعداد خوشه مشخص در ارتباط با تعداد ایستگاه‌ها در یک منطقه فرض مشخصی وجود ندارد. به عبارت دیگر معیار مشخصی برای تعیین ارتباط بین تعداد ایستگاه‌ها و تعداد خوشه‌ها در یک منطقه در منابع مختلف بیان نشده است. با این حال در تحلیل فراوانی منطقه‌ای هدف تعیین تعداد مشخصی از خوشه‌ها (مناطق) می‌باشد که در آن‌ها توزیع فراوانی در منطقه از ایستگاهی به ایستگاه دیگر کم‌ترین تغییرات را داشته باشد (صدری و برون، ۲۰۱۱). در این مطالعه از شاخص‌های فازی برای بررسی الگوی ساختاری داده‌ها و پراکنش گشتاورهای خطی در منطقه برای تعیین تعداد بهینه خوشه‌ها استفاده گردید. شاخص‌های فازی XB ، DI ، ADI ، PC ، SC و S^1 به این منظور به کار گرفته شدند. برای ارزیابی عملکرد بهینه خوشه‌بندی فازی، مقادیر شاخص‌های فازی در $c=2:10$ محاسبه گردید. بهترین عملکرد خوشه‌بندی با کم‌ترین میزان این شاخص‌ها مشخص می‌شود (جزئیات بیشتر در خصوص محاسبه‌های مربوط به این شاخص‌ها در راثو و اسرینیواس ۲۰۰۶ آورده شده است). شکل ۴ تغییرات این شاخص‌ها را در ارتباط با تعداد خوشه‌ها در خوشه‌بندی گشتاورهای خطی نشان می‌دهد. با توجه به شکل مشخص می‌شود مقادیر شاخص‌های SC و S ، با افزایش تعداد خوشه‌ها کاهش پیدا می‌کند. با این حال شیب تغییرات شاخص SC بعد از $c=3$ محسوس نیست و می‌توان گفت تغییرات شاخص با افزایش تعداد خوشه‌ها از $c \geq 3$ قابل توجه نمی‌باشد. براساس شاخص‌های PC و DI کم‌ترین میزان شاخص در $c=3$ دیده می‌شود. براساس شاخص ADI میزان شاخص در $c \geq 5$ تقریباً یکنواخت و به کم‌ترین مقدار خود رسیده است. براساس شاخص XB در $c \geq 4$ افزایش و در $c \geq 6$ کاهش پیدا کرده است و در $c=8$ به کم‌ترین میزان خود رسیده است. در تحلیل فراوانی منطقه‌ای هدف دستیابی به مناطق همگنی می‌باشد که به اندازه کافی بزرگ باشند (صدری و برون، ۲۰۱۱) چرا که با افزایش تعداد ایستگاه‌های شرکت‌کننده در محاسبه‌های دقت برآوردها بیشتر و ارزیابی کاهش پیدا می‌کند. بنابراین همگنی خوشه‌ها در سه سناریوی $c=3, 4, 5$ مورد بررسی قرار گرفت.

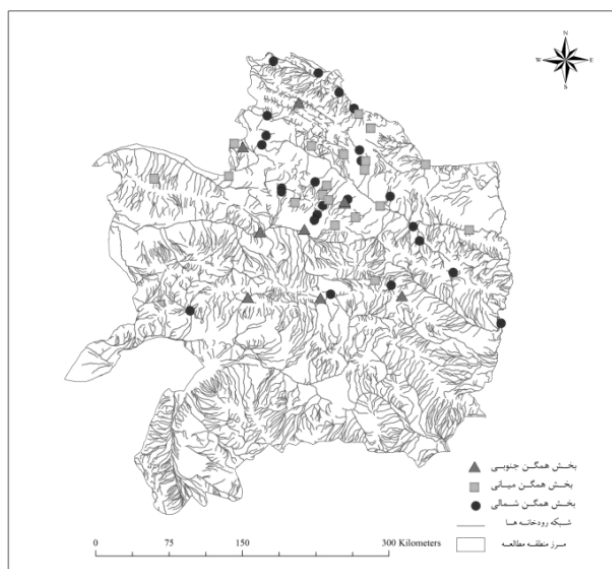
1- Xie-Beni (XB), Dunn Index (DI), Alternative Dunn Index (ADI), Partition Coefficient (PC), Partition Index (SC), Separation Index (S)



شکل ۴- تغییرات مقادیر شاخص‌های فازی با افزایش تعداد خوشه‌ها.

براساس آماره همگنی روش گشتاورهای خطی (H) مشخص گردید بهترین نتیجه زمانی به دست می‌آید که تعداد خوشه‌ها برابر با ۳ خوشه انتخاب می‌شود (جدول ۱). بنابراین کل منطقه مطالعه به ۳ بخش همگن هیدرولوژیک (شمالی، میانی و جنوبی) تقسیم‌بندی شد. شکل ۵ موقعیت ایستگاه‌های هر خوشه را بر روی شبکه هیدروگرافی منطقه نشان می‌دهد.

جدول ۱ مقادیر آماره همگنی و متوسط گشتاورهای خطی نمونه را در خوشه‌های ایجاد شده نشان می‌دهد. خوشه‌های ایجاد شده از نظر همگنی سری‌های زمانی سیل از همگنی قابل‌قبولی برخوردار می‌باشند. با توجه به نتایج مندرج در جدول مشخص می‌شود متوسط گشتاورهای خطی از شمال به جنوب افزایش پیدا می‌کند (جدول ۱) که این خود ناشی از کاهش متوسط بارش از شمال به جنوب استان می‌باشد (فلاح‌قاله‌ری و همکاران، ۲۰۱۰) که باعث افزایش تغییرپذیری میزان بارش و در نتیجه افزایش ضریب تغییرات سیل در این منطقه گردیده است.



شکل ۵- مناطق همگن ایجاد شده با استفاده از روش تحلیل خوشه‌ای براساس گشتاورهای خطی.

جدول ۱- مقادیر آماره‌های همگنی برای کل منطقه مورد مطالعه و بخش‌های همگن.

متوسط منطقه‌ای گشتاورهای خطی			همگنی منطقه‌ای	تعداد زیرحوزه	منطقه
L-Ck	L-Cs	L-Cv	H_1		
۰/۲۳	۰/۴۱	۰/۵۳	۵/۰۹	۵۲	کل منطقه مطالعه
۰/۱۱	۰/۲۶	۰/۴۴	۰/۹۲	۲۴	بخش همگن شمالی
۰/۲۳	۰/۴۳	۰/۵۴	-۱/۵	۱۹	بخش همگن میانی
۰/۴۷	۰/۶۴	۰/۶۵	۰/۵۸	۹	بخش همگن جنوبی

انتخاب تابع توزیع منطقه‌ای: پس از بررسی همگنی منطقه مطالعه از نظر هیدرولوژیک با استفاده از آماره نکوئی برازش (Z^{Dist}) (هاسکینگ و والیس، ۱۹۹۳)، توابع توزیع پارتوی تعمیم‌یافته^۱ در بخش شمالی و میانی و توزیع لجستیک تعمیم‌یافته^۲ در بخش جنوبی به ترتیب عنوان بهترین توابع توزیع

- 1- Generalized Pareto (GPAR)
- 2- Generalized Logistic (GLOG)

منطقه‌ای شناخته شدند. نتایج جدول ۲، بیانگر مقادیر این آماره برای توابع توزیع مختلف می‌باشد. با توجه به نتایج درج شده در جدول ۲ مشخص می‌شود توزیع پارتوی تعمیم‌یافته در تمامی نواحی همگن هیدرولوژیک و نیز در کل منطقه مطالعه برازش قابل‌قبولی به داده‌های دبی سیل داشته است. توابع توزیع منتخب در این مطالعه (توزیع پارتوی تعمیم‌یافته و توزیع لجستیک تعمیم‌یافته) در مطالعات پژوهش‌گران دیگر نیز به‌عنوان توابع توزیع مناسب در برازش به داده‌های منطقه‌ای سیل مشخص گردیده‌اند به‌عنوان مثال اسلامیان و چاوشی (۲۰۰۳) نشان دادند توابع توزیع لجستیک تعمیم‌یافته و پارتوی تعمیم‌یافته به‌عنوان مناسب‌ترین توابع توزیع در برازش به داده‌های منطقه‌ای سیل در بخش‌هایی از حوزه‌های آبخیز مرکزی ایران بوده‌اند. ملکی‌نژاد و همکاران (۲۰۱۳) نیز توزیع لجستیک تعمیم‌یافته را به‌عنوان مناسب‌ترین توزیع منطقه‌ای در برازش به داده‌های سیل در بخشی از حوزه دریاچه نمک معرفی کردند. بنابراین توابع توزیع منتخب در این مطالعه به‌عنوان توابع توزیع مناسب در برازش به داده‌های منطقه‌ای سیل می‌باشند.

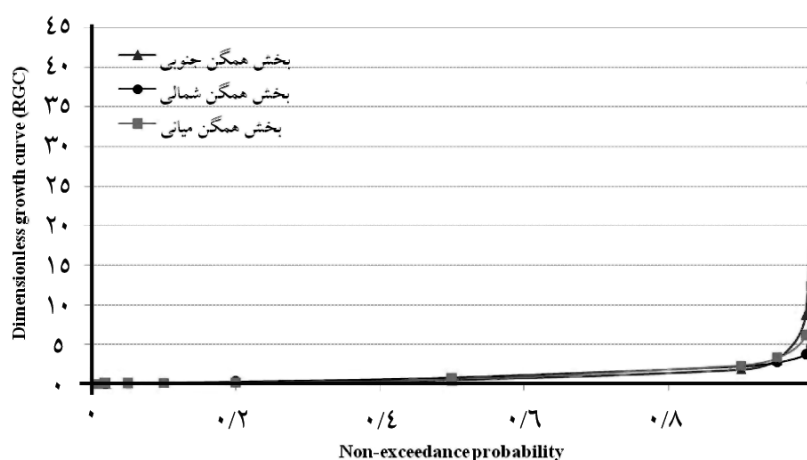
جدول ۲- آزمون نکویی برازش برای انتخاب بهترین تابع توزیع منطقه‌ای.

تابع توزیع منطقه‌ای					منطقه
GPD	P111	GNO	EVI	GLOG	
-۱/۲۳*	-۲/۹۷	-۰/۰۳*	۱/۷۰	۲/۵۱	کل منطقه مطالعه
۰/۳۱*	۱/۷۴	۳/۱۳	۳/۹۲	۵/۳۴	بخش همگن شمالی
۰/۲۷*	۱/۱۶*	۰/۸۲*	۱/۹۸	۲/۴۱	بخش همگن میانی
-۱/۵۸*	-۳/۱۵	-۱/۹۱	-۱/۱۲*	-۱/۱۱*	بخش همگن جنوبی

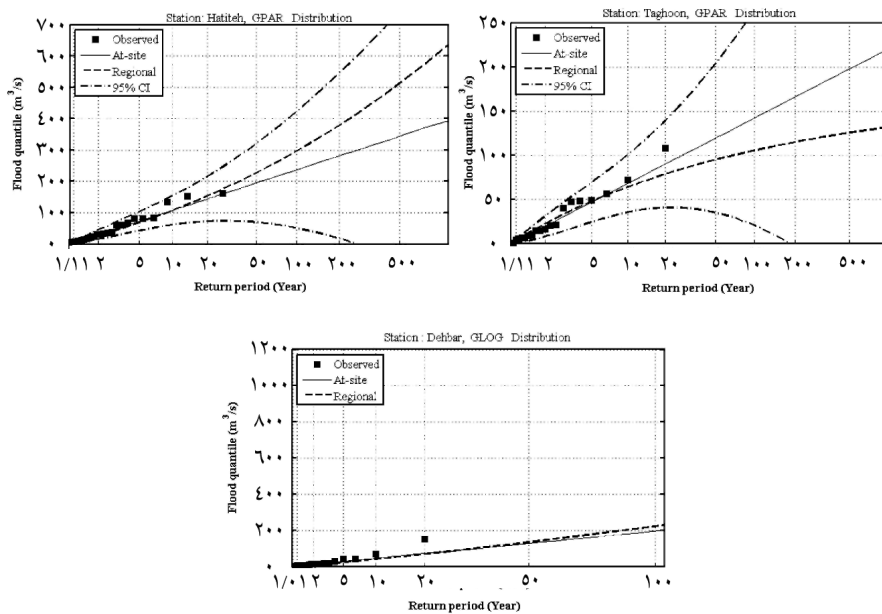
* نشان‌دهنده توزیع قابل‌قبول.

منحنی‌های رشد منطقه‌ای سیلاب: منحنی رشد منطقه‌ای نسبت بزرگی سیلاب با دوره بازگشت T سال، به مقدار میانگین سیلاب در منطقه می‌باشد. در برآورد منحنی‌های رشد منطقه‌ای فرض بر این است که این نسبت برای تمامی ایستگاه‌های منطقه مطالعه یکسان است (دین‌پژوه و همکاران، ۲۰۰۴) منحنی‌های رشد منطقه‌ای سیل کل منطقه مورد مطالعه و نیز بخش‌های همگن شمالی، میانی و جنوبی در شکل ۶ نشان داده شده است. با افزایش شیب منحنی‌ها، به دلیل افزایش ضریب تغییرات داده‌ها،

دقت پیش‌بینی‌ها پایین‌تر خواهد بود (دین‌پژوه و همکاران، ۲۰۰۴). منحنی رشد منطقه‌ای سیل در بخش جنوبی منطقه بیش‌ترین شیب را نشان می‌دهد. افزایش شیب منحنی در بخش همگن جنوبی نسبت به بخش میانی و شمالی می‌تواند ناشی از افزایش نسبی در متوسط منطقه‌ای نسبت گشتاورهای خطی و به‌ویژه ضریب تغییرات خطی در این منطقه باشد (جدول ۱) که این خود ناشی از کاهش متوسط بارش از شمال به جنوب استان می‌باشد (فلاح‌قاله‌ری و همکاران، ۲۰۱۰) که باعث افزایش تغییرپذیری میزان بارش و در نتیجه افزایش ضریب تغییرات سیل در این منطقه گردیده است. در بخش جنوبی حوزه که منحنی رشد منطقه‌ای سیل شیب تندتری را نشان می‌دهد، شاهد وقوع سیلاب‌هایی بسیار بیش‌تر (و یا کم‌تر) از مقدار میانگین آن خواهیم بود و دقت پیش‌بینی‌ها در تحلیل فراوانی به‌خصوص در دوره‌های بازگشت بالا، پایین خواهد بود. شکل ۷ تحلیل فراوانی سیل با استفاده از توابع منطقه‌ای سیل را در سه ایستگاه معرف طاغون (معرف بخش همگن شمالی)، حطیبه (معرف بخش همگن میانی) و دهبار (معرف بخش همگن جنوبی) نشان می‌دهد.



شکل ۶- منحنی‌های رشد منطقه‌ای سیل در بخش‌های مختلف منطقه.



شکل ۷- برآورد مقادیر سیل با استفاده از توابع توزیع منطقه‌ای GPAR در بخش شمالی (ایستگاه معرف طاغون) و میانی (ایستگاه معرف حطیبه) و GLOG در بخش جنوبی (ایستگاه معرف دهبار).

نتیجه‌گیری

این پژوهش به بررسی تحلیل فراوانی منطقه‌ای سیلاب در استان خراسان رضوی با تلفیق تئوری گشتاورهای خطی و روش تحلیل خوشه‌ای فازی می‌پردازد. در ابتدا با استفاده از نمودار نسبت گشتاورهای خطی و آماره همگنی (H)، مشخص گردید کل منطقه مورد مطالعه از نظر هیدرولوژیک کاملاً ناهمگن می‌باشد. با حذف ایستگاه‌های ناجورکلاته منار و حسین‌آباد از مجموعه ایستگاه‌های منطقه، هم‌چنین ناهمگنی بسیار بالا می‌ماند. در ادامه برای دستیابی به مناطق همگن هیدرولوژیک با استفاده از مؤلفه‌های فیزیکی و آماری زیرحوزه‌ها با استفاده از الگوریتم خوشه‌ای فازی منطقه به مناطق کوچک‌تر تقسیم‌بندی شد. نتایج نشان داد مناطق ایجاد شده براساس آماره همگنی H کاملاً ناهمگن می‌باشند و با افزایش تعداد خوشه‌ها تغییری در میزان همگنی خوشه‌ها به دست نمی‌آید. دلیل این امر نبود رابطه منطقی بین مشخصات فیزیوگرافیک و هیدرولوژیک در منطقه می‌باشد. این نتیجه تأییدی بر غیراصولی بودن استفاده از مشخصات فیزیکی حوزه‌ها برای تعیین مناطق همگن بدون در نظر گرفتن

مشخصات هیدرولوژیک آن‌ها می‌باشد. برای دست‌یابی به مناطق همگن، با استفاده از شاخص‌های فازی ساختار داده‌ای و الگوی پراکنش گشتاورهای خطی در منطقه مورد تجزیه و تحلیل قرار گرفت و تعداد بهینه خوشه‌ها با توجه مقادیر شاخص‌ها تعیین گردید. با استفاده از الگوریتم فازی FCM منطقه مورد مطالعه به ۳ بخش شمالی، میانی و جنوبی تقسیم‌بندی گردید. شاخص همگنی H نشان داد خوشه‌های ایجاد شده از همگنی قابل‌قبول برخوردارند. با استفاده از آزمون نکوئی برازش مشخص گردید تابع توزیع پارتوی تعمیم‌یافته در بخش‌های شمالی و میانی و تابع توزیع لجستیک تعمیم‌یافته در بخش جنوبی بهترین برازش را به داده‌های دبی سیل منطقه‌ای داشته است. با استفاده از توابع توزیع منطقه‌ای اقدام به تهیه منحنی‌های رشد منطقه‌ای در بخش‌های مختلف منطقه گردید. نتایج نشان‌دار شیب منحنی‌ها از شمال به جنوب منطقه افزایش می‌یابد که نشان‌دهنده ضریب تغییرات بالای سیل در این منطقه می‌باشد که باعث می‌شود از قابلیت اعتماد پیش‌بینی‌ها در تحلیل فراوانی سیل کاسته شود که این مسأله باید در مطالعات مربوط به اقدامات پیش‌گیرانه و کنترل سیل و تبدیل مدیریت بحران به مدیریت ریسک در نظر گرفته شود. نتایج این مطالعه نشان می‌دهد تلفیق روش‌های خوشه‌بندی فازی و گشتاورهای خطی می‌تواند رویکرد مناسبی برای تعیین مناطق همگن هیدرولوژیک به‌خصوص در مناطق وسیع جغرافیایی با تنوع توپوگرافی، اقلیمی و هیدرولوژیک باشد که منجر به افزایش دقت پیش‌بینی‌ها در تحلیل فراوانی می‌شود.

منابع

1. Anli, A.S., Apaydin, H., and Öztürk, F. 2007. Regional flood frequency estimation for the Göksu river basin through L-moments, International River Basin Management Conference, State Hydraulic Works. Gloria Golf Resort Hotel, Belek, Antalya, Pp: 424-438.
2. Ayvaza, M.T., Karahana, H., and Aral, M.M. 2007. Aquifer parameter and zone structure estimation using kernel-based Fuzzy C-Means clustering and genetic algorithm. J. Hydrol. 343: 3-4. 240-253.
3. Dikbas, F., Firat, M., Cem, K.A., and Gungor, M. 2011. Classification of precipitation series using fuzzy cluster method. Int. J. Climatol. DOI: 10.1002/joc.2350.
4. Dinpashoh, Y., Fakheri-Fard, A., Moghaddam, M., Jahanbakhsh, S., and Mirnia, M. 2004. Selection of variables for the purpose of regionalization of Iran's precipitation climate using multivariate methods. J. Hydrol. 297: 109-123.

5. Eslamian, S.S., and Boroojeni, S.C. 2003. Application of L-moment theory in flood frequency analysis of central catchments of Iran. *J. Sci. Tech. Agri. Natur. Resour.* 7: 1. 1-17. (In Persian)
6. Fallah Ghalhary, Gh.A., Habibi Nokhandan, M., and Khoshhal, J. 2010. Spring rainfall estimation of Khorasan Razavi province based on Tele-connection synoptically patterns using adaptative neuro-fuzzy inference system. *J. Natur. Resour. Iran.* 63: 1. 55-74. (In Persian)
7. Greenwood, J.A., Landwehr, J.M., Matalas, N.C., and Wallis, J.R. 1979. . Probanility weighted moments: definition and relation to parameters of several distributions expressible in inverse form. *Water Resour Res.* 15: 5. 1049-1054.
8. Hoskins, J.R.M. 1990. L-moments: analyzing and estimation of distributions using linear combinations of order statistics. *J. R. Stat. Soc. B.* 52: 105-124.
9. Hosking, J.R.M., and Wallis, J.R. 1993. Some statistics useful in regional flood frequency analysis. *Water Resour Res.* 23: 271-281.
10. Hoskins J.R.M., and Wallis, J.R. 1997. Regional frequency analysis: An approach based on l-moments. Cambridge, Angleterre.
11. Kumar, R., Chatterjee, C., Kumar, S., Lohani, A.K., and Singh, R.D. 2003. Development of Regional Flood Frequency Relationships Using L-moments for Middle Ganga Plains Subzone 1(f) of India. *Water Resour Manage.* 17: 4. 243-257.
12. Lim, Y.H., and Lye, L.M. 2003. Regional flood estimation for ungauged basins in Sarawak, Malaysia. *Hydrol. Sci. J.* 48: 1.
13. Madsen, M., Pearson, C.P., and Rosbjerg, D. 1997. Comparison of annual maximum series and partial duration methods for modelling extreme hydrologic events. *Water Resour. Res.* 33: 4. 759-769.
14. Malekinezhad, H., Nachtnebel, H.P., and Klik, A. 2013. Regionalization approach for extreme flood analysis using L-Moments. *J. Agri. Sci. Technol.* 13: 1183-1196.
15. Modarres, R. 2008. Regional Frequency Distribution Type of Low Flow in North of Iran by L-moments. *Water Resour. Manage.* 22: 7. 823-841.
16. Parida, B., Kachroo, R., and Shrestha, D. 1998. Regional flood frequency analysis of Mahi-Sabarmati basin (Subzon 3-a) using index flood procedure with L-moments. *Water Resour. Manage.* 12: 1-12.
17. Pearson, C.P. 1991. New Zealand regional flood frequency analysis using L-moments, The NewZealand hydrological society. *J. Hydrol.* 30: 2. 53-64.
18. Rao, A.R., and Hamed, K.H. 1997. Regional frequency analysis of Wabash river flood data by L-moments. *J. Hydrol. Eng.* 2: 4. 169-179.
19. Rao, A.R., and Srinivas, V.V. 2006. Regionalization of watersheds by fuzzy cluster analysis. *J. Hydrol.* 318: 1-4. 57-79.
20. Sadri, S., and Burn, D.H. 2011. A Fuzzy C-Means approach for regionalization using a bivariate homogeneity and discordancy approach. *J. Hydrol.* 401: 231-239.

21. Saf, B. 2009. Regional Flood Frequency Analysis Using L-Moments for the West Mediterranean Region of Turkey. *Water Resour Manage.* 23: 531-551.
22. Sarhadi, A., Soltani, S., and Modarres, R. 2012. Probabilistic flood inundation mapping of ungauged rivers: linking GIS techniques and frequency analysis. *J. Hydrol.* 458-459: 68-86.
23. Shu, C., and Burn, D.H. 2004. Homogeneous pooling group delineation for flood frequency analysis using a fuzzy expert system with genetic enhancement. *J. Hydrol.* 291: 132-149.
24. Shu, C., and Ouarda, T. 2008. Regional flood frequency analysis at ungauged sites using the adaptive neuro-fuzzy inference system. *J. Hydrol.* 349: 1-2. 31-43.
25. Vogel, R.M., and Wilson, I. 1996. Probability distribution of annual maximum, mean, and minimum streamflows in the United States. *J. Hydrol. Eng.* 1: 2. 69-76.



Gorgan University of Agricultural
Sciences and Natural Resources

J. of Water and Soil Conservation, Vol. 21(1), 2014
<http://jwsc.gau.ac.ir>

Regional flood frequency analysis by integrating L-moments and fuzzy clustering techniques (Case study: Khorasan Razavi)

***E. Dodangeh¹, M.B. Nateghi², H. Rahnema³ and A.Gh. Dadrasi Sabzevar⁴**

¹Ph.D. Student, Dept. of Watershed Management Engineering, Sari University of Agricultural Sciences and Natural Resources, ²M.Sc. Graduated, Dept. of Water Resources Engineering, Ministry of Defense and Armed Forces Logistics, ³M.Sc. Graduated, Dept. of Watershed Management Engineering, Sari University of Agricultural Sciences and Natural Resources, ⁴M.Sc. Graduated, Dept. of Soil and Watershed Conservation, Agriculture and Natural Resources Research Centre of Khorasan Razavi

Received: 10/27/2012; Accepted: 01/18/2013

Abstract

Regional flood frequency analysis is carried out by integrating L-moments and fuzzy cluster analysis in north east Iran. L-moments based homogeneity test revealed that entire region is absolutely heterogeneous. Removing discordant sites, homogeneity was not achieved for the study region, therefore Fuzzy C-Means algorithm is applied to gain acceptable homogeneous regions. Cluster analysis based on site characteristics and at-site statistics resulted in heterogeneous regions. It is because of not existing a significant relationship between site characteristics and probability weighted moments. Screening L-moments pattern by fuzzy cluster validation indexes, XB, S, SC, PC, DI and ADI and homogeneity measure H, the optimum number of regions ascertained to be $c=3$. By applying fuzzy C-Means algorithm, the whole study region is divided into three north, middle and south sub-regions. Based on the goodness-of-fit test, Z^{Dist} , GPAR distribution is determined as the best fitted regional distribution for the north and middle and GLOG distribution is determined for the south homogeneous sub-regions. Regional growth curves (RGC) are then constructed by the applied regional distributions. Because of increasing coefficient of variation (CV) of the peak flows, the regional growth curves tend to increase from north to the south.

Keywords: Hydrologic homogeneity, Fuzzy cluster validation index, L-moments, Regional distribution

* Corresponding Author; Email: smaeel.dodangeh@gmail.com

