

بررسی تأثیر ویژگی زمین‌شناسی در تحلیل فراوانی سیلاب منطقه‌ای حوضه‌های آبریز شمال شرق ایران

طیبه طاهرپور^۱ - بیژن قهرمان^{۲*} - کامران داوری^۳

تاریخ دریافت: ۱۳۹۲/۱۰/۰۳

تاریخ پذیرش: ۱۳۹۶/۰۵/۰۸

چکیده

تحلیل فراوانی منطقه‌ای سیلاب، نیازمند شناسایی حوضه‌های مشابه از نظر مکانیزم تولید سیلاب می‌باشد. تشابه حوضه‌ها در تولید سیلاب تابع عواملی نظیر ویژگی‌های فیزیوگرافی و هواشناسی حوضه، موقعیت جغرافیایی و زمین‌شناسی می‌باشد. در این مطالعه به منظور تعیین اثر شاخص زمین‌شناسی در تعیین مناطق همگن هیدرولوژیکی، ۷۳ ایستگاه آب‌سنجی واقع در شمال شرق ایران که دارای اقلیم خشک تا نیمه‌خشک است با میانگین طول دوره آماری ۲۹ سال، با در نظر گرفتن شاخص زمین‌شناسی و نیز بدون در نظر گرفتن این شاخص با استفاده از الگوریتم خوشه‌بندی C-میانگین فازی به ۶ ناحیه‌ی همگن تقسیم شدند. جهت تعیین تعداد بهینه‌ی خوشه‌ها از سه شاخص صحت‌سنجی خوشه‌بندی فازی وون، ژی-بنی و فوکویاما-سوگنو استفاده شد. نواحی حاصل از دو روش با استفاده از آماره‌های همگنی برپایه‌ی گشتاورهای خطی همگن تشخیص داده شد. توابع توزیع ناحیه‌ها با استفاده از آزمون‌های نیکوئی برازش Z و کولموگروف-اسمیرنوف انتخاب شدند. با مقایسه توزیع ایستگاه‌ها و همچنین دو آماره ارزیابی میانه خطای نسبی و نسبت دبی پیش‌بینی شده به دبی برآورد شده به ازاء ۵ دوره بازگشت مختلف (۵، ۱۰، ۲۰، ۵۰ و ۱۰۰ سال)، هر دو روش تحلیل، نتایجی قابل قبول داشته و اضافه کردن شاخص زمین‌شناسی منجر به بهبود نتایج گردید. با در نظر گرفتن ویژگی زمین‌شناسی در ناحیه‌بندی حوضه‌ها، علاوه بر توابع توزیع برتر، تابع توزیع لوگ نرمال سه پارامتری برای تمامی نواحی مناسب تشخیص داده شد، که از این نقطه نظر استفاده از ویژگی زمین‌شناسی مناسب بود.

واژه‌های کلیدی: تحلیل خوشه‌ای فازی، سیلاب نمایه، حوضه‌های آبریز فاقد آمار، شاخص نفوذ

مقدمه

(۳۰). تحلیل فراوانی سیلاب منطقه‌ای با روش‌های زیادی در بسیاری از کشورهای مختلف جهان بکار رفته که یکی از روش‌های بسیار رایج، روش سیلاب نمایه است. از جمله کاربردهای این روش در نقاط مختلف جهان می‌توان به مطالعات دال‌ریمپل (۴)، هاسکینگ و والیس (۱۲)، سف (۲۵ و ۲۶)، دس و کیونین (۵)، سوین‌سن و همکاران (۲۸)، زمان (۳۰) و کومار (۱۶) اشاره کرد.

گام نخست در تحلیل فراوانی منطقه‌ای، تعیین نواحی همگن هیدرولوژیکی می‌باشد. برای رسیدن به این هدف می‌بایست ایستگاه‌ها را به گروه‌هایی جدا از هم تقسیم‌بندی نمود. باتوجه به پژوهش‌های انجام شده برای تعیین منطقه‌های همگن، می‌توان از معیارها و روش‌های مختلفی استفاده نمود. تحلیل خوشه‌ای، به عنوان نوعی تحلیل چند متغیره، ابزاری آماری است که در آن داده‌ها و مشاهدات بر اساس متغیرهای تعیین شده توسط محقق طبقه‌بندی می‌شوند. روش‌های تفکیکی به دو بخش خوشه‌بندی سخت (مثلاً K-میانگین، گوئل (۸)) و خوشه‌بندی فازی (مثلاً C-میانگین، جین‌گی و هال (۱۴)) تقسیم می‌شوند. در خوشه‌بندی سخت، هر داده به یک و

تخمین دبی سیلاب با دوره‌های بازگشت مختلف در برنامه‌ریزی و مدیریت منابع آب مانند مطالعات زیست‌محیطی رودخانه، سازه‌های هیدرولیکی، ذخیره سیلاب و مدیریت دشت سیلابی از اهمیت به سزایی برخوردار است. تحلیل فراوانی نقطه‌ای سیل مستقیم‌ترین روش برآورد دبی طراحی است که نیازمند دوره طولانی مدت از داده‌های جریان رودخانه می‌باشد. با این حال در بسیاری از موارد در محل مورد نظر داده‌های ثبت شده جریان محدود بوده یا به‌طور کامل وجود ندارد. در این شرایط جهت برآورد دبی طراحی از تحلیل فراوانی سیلاب منطقه‌ای استفاده می‌شود که در آن ویژگی‌های جریان از حوضه‌های دارای آمار جریان به حوضه‌های فاقد آمار انتقال می‌یابد

۱ و ۲ و ۳- به ترتیب دانش آموخته کارشناسی ارشد منابع آب و استادان گروه مهندسی آب، دانشکده کشاورزی، دانشگاه فردوسی مشهد

(Email: bijangh@um.ac.ir)

(*- نویسنده مسئول)

نحوه کاربری، که در سال‌های مختلف متفاوت می‌باشد، بستگی داشته و در کل، نااطمینانی‌های زیادی برای تعیین آن وجود دارد.

برای یافتن نمایه‌ای مناسب، در مطالعه حاضر برای در نظر گرفتن نقش زمین‌شناسی، به شاخصی از نفوذپذیری حوضه که به طور معکوس بر روی شدت سیل‌خیزی حوضه تأثیر دارد، توجه شد. تأثیر جنس زمین در واحدهای هیدرولوژیک یکسان نمی‌باشد. وجود درز و شکاف در زمین‌های آهکی یا کارستی که خود سبب نفوذ آب به درون زمین می‌شود، از شدت سیلاب‌ها کاسته و آن را تعدیل می‌کند این در حالی است که در زمین‌های ماری و رسی عکس آن است. بنابراین با استفاده از نقشه‌های زمین‌شناسی و توجه به لیتولوژی سنگ‌ها، اقدام به برآورد وضعیت نفوذپذیری حوزه‌ها گردید.

از جمله مطالعاتی که در منطقه شمال شرق کشور صورت گرفته است می‌توان به مطالعات شامکوئیان و همکاران (۲۷) اشاره کرد که در آن با انتخاب چهار متغیر مساحت، طول بلندترین آبراهه، طول عرض جغرافیایی مرکز ثقل حوضه، جهت ناحیه‌ای کردن حوضه‌های آبریز استان‌های خراسان از روش‌های تحلیل حوضه‌ای کلاسیک استفاده و ۷ ناحیه همگن حاصل شد. موشخیان و همکاران (۱۹) نیز جهت توسعه‌ی مدل‌های منطقه‌ای برآورد سیلاب در حوضه‌های آبریز استان خراسان رضوی از تحلیل حوضه‌ای سلسله‌مراتبی وارد استفاده کردند. در مطالعات صورت گرفته در منطقه تنها از روش کلاسیک برای حوضه‌بندی استفاده شده و همچنین ویژگی زمین‌شناسی به عنوان عاملی مؤثر در تحلیل حوضه‌ای لحاظ نشده است. هدف از این مطالعه، تعیین مناطق همگن هیدرولوژیکی حوضه‌های آبریز شمال شرق کشور با استفاده از الگوریتم حوضه بندی C- میانگین فازی و همچنین ارزیابی میزان تأثیر استفاده از شاخص زمین‌شناسی در تحلیل فراوانی سیلاب منطقه‌ای می‌باشد.

مواد و روش‌ها

منطقه‌ی مورد بررسی و داده‌ها

منطقه‌ی مورد بررسی شامل زیرحوضه‌های سه حوضه آبریز اصلی کشور، اترک، کشف‌رود (قره‌قوم) و کال‌شور، واقع در استان‌های خراسان شمالی و رضوی می‌باشد. از ۷۳ ایستگاه آب‌سنجی واقع در این حوضه‌ها با حداقل طول دوره آماری ۵ سال (در ایستگاه ازغند) و حداکثر طول دوره آماری ۴۶ سال (در ایستگاه آق‌دریوند) استفاده شد. میانگین طول دوره‌ی آماری ایستگاه‌ها ۲۷ سال بوده و طول دوره‌ی آماری مختوم به سال آبی ۱۳۸۹-۱۳۹۰ می‌باشد. طول دوره‌ی آماری تنها در ۱۵ درصد ایستگاه‌ها کم‌تر از ۱۵ سال بود. جهت تحلیل حوضه‌ای از ویژگی‌های متداول فیزیوگرافی حوضه‌ها شامل مساحت، شیب، ارتفاع متوسط، طول و عرض مرکز ثقل حوضه، طول آبراهه‌ی اصلی و همچنین ویژگی زمین‌شناسی حوضه استفاده شد. برای در

تنها یک حوضه نسبت داده می‌شود، در حالی که در حوضه‌بندی فازی، تفکیک فازی صورت می‌گیرد. در شرایط واقعی، حوضه‌بندی فازی بسیار طبیعی‌تر از حوضه‌بندی سخت است، چون داده‌های موجود در مرز حوضه‌های مختلف مجبور به تعلق کامل به یکی از حوضه‌ها نبوده و با درجه تعلق بین صفر تا یک که نشان دهنده تعلق نسبی آن‌ها است تفکیک می‌شوند. عضویت جهت شناسایی حوضه‌های واقع شده در محدوده همپوشانی ناحیه‌ها، آگاهی از این توزیع مفید است. بنابراین حوضه‌بندی فازی اطلاعات بیش‌تری در مورد ساختار داده‌ها نسبت به حوضه‌بندی سخت فراهم کرده و جهت تحلیل فراوانی سیل مناسب‌تر است (۲۲).

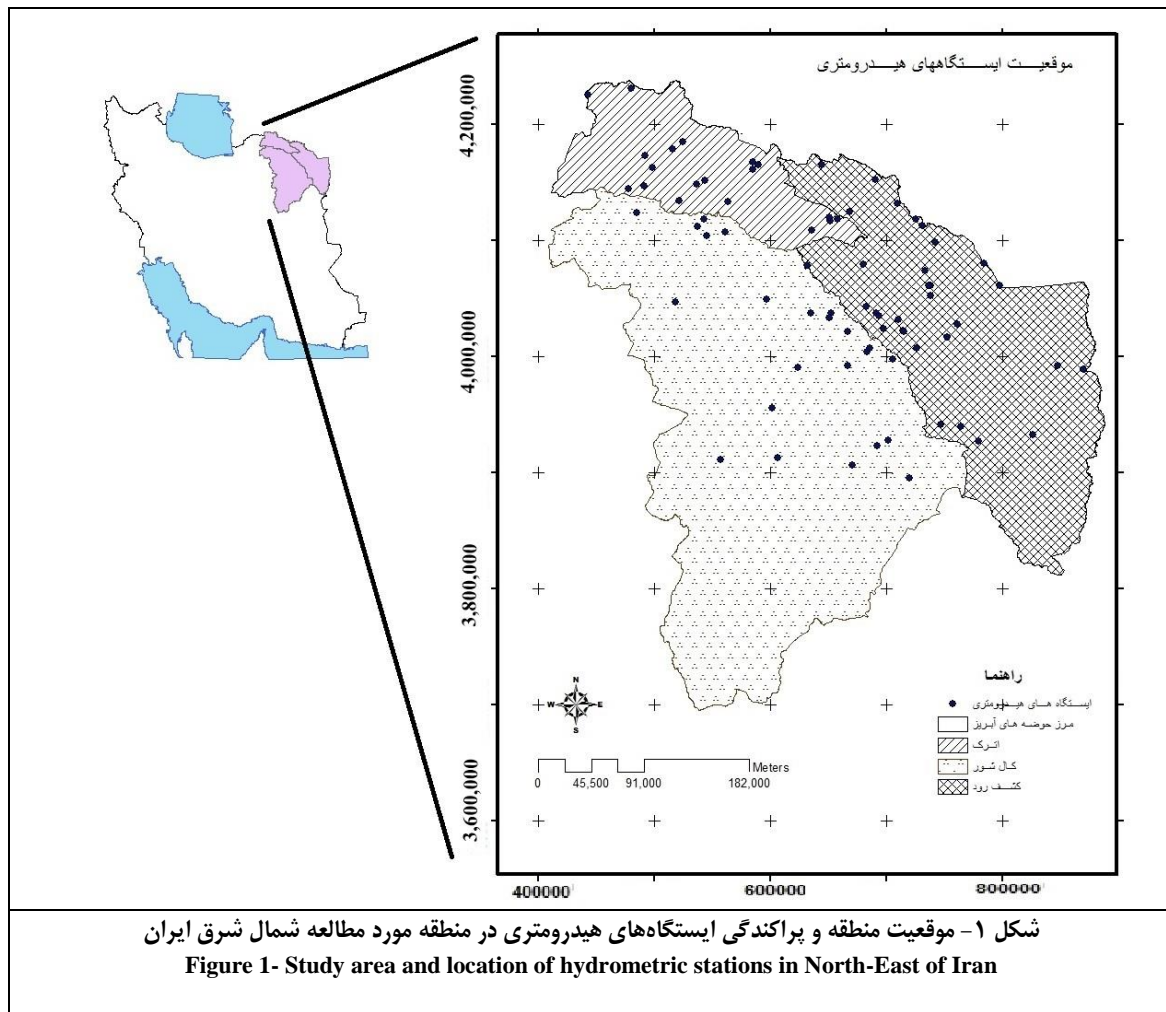
مطالعات انجام شده از جمله: هال و مینس (۱۰)، جین‌گی و هال (۱۴)، راتو و سری‌نیواس (۲۲)، نشان داده که جهت حاصل شدن حوضه‌بندی با کیفیت بالا استفاده از تابع حوضه‌بندی فازی ضروری بوده و ناحیه‌های حاصل از حوضه‌بندی فازی همگنی فیزیکی بهتری را نسبت به حوضه‌بندی کلاسیک دارند.

به علت پیچیدگی در اطلاع دقیق عوامل مؤثر در تولید سیل، رهنمود مشخصی جهت شناسایی نواحی همگن وجود ندارد. لذا اولین مسئله مشترک در همه روش‌های ناحیه‌ای کردن، انتخاب متغیرهای حوضه‌بندی است (۱۸). برای مشخص کردن نواحی همگن حوضه‌های آبریز، از دو گروه متغیر (الف) آماره‌های ایستگاهی و (ب) ویژگی‌های ایستگاهی استفاده می‌شود. به عنوان مثال در نیوزیلند، موس‌لی (۲۰) به کمک تحلیل حوضه‌ای و با استفاده از آمار دبی متوسط سالانه و ضریب تغییرات حوضه‌ها، به گروه‌بندی مناطق مذکور پرداخت. عطایی و شیران (۲) جهت ناحیه‌ای کردن حوضه‌های آبریز دشت کرون از چهار گروه متغیر ژئومورفولوژی حوضه (شاخص‌های تراکم آبراهه‌ها، ژئومتری حوضه، متغیرهای ارتفاعی و اقلیمی حوضه)، استفاده کردند. جریان سیل بطور عمده ناشی از رواناب سطحی می‌باشد که حاصل مقدار و خصوصیات بارش و خصوصیات حوضه آبریز آن است. بزرگی سیلاب رودخانه‌ها وابستگی مستقیم به جنس و نفوذپذیری زمین دارد. از این رو، الوانکار (۱) نشان داد علاوه بر مساحت، ارتفاع متوسط و شاخص پوشش گیاهی حوضه‌های آبریز، اطلاعات دیجیتالی ماهواره لندست که به جنس زمین وابسته بوده و بیانگر نفوذپذیری زمین می‌باشد نیز در بزرگی سیلاب حوضه‌های تأثیر فراوانی دارد. مشکل استفاده از تصاویر ماهواره‌ای این است که این تصاویر در دسترس نبوده و فن‌آوری آن هنوز همگانی نشده و برای انتخاب تصاویر مناسب و قابل اطمینان مشکلات فراوانی وجود دارد. میشرای و همکاران (۱۸) جهت ناحیه‌ای کردن حوضه‌های آبریز نیپال از شماره منحنی رواناب (CN) که تابع نوع خاک از لحاظ نفوذپذیری، پوشش سطح حوضه و شیب حوضه می‌باشد، به عنوان روشی مؤثر در ناحیه‌ای کردن استفاده کرد. اما شماره منحنی به نوع کاربری بستگی داشته که خود عاملی پویا بوده و به فصل سال و

منطقه و پراکندگی ایستگاه‌های هیدرومتری در سه حوضه آورده شده است.

جهت پرهیز از تأثیر واحدهای اندازه‌گیری و عدم وابستگی به آن‌ها، داده‌های استخراج شده استاندارد می‌شوند. در این صورت، متغیرها تأثیر غیرمتجانسی بر اندازه‌گیری فاصله نخواهند داشت. از متداول-ترین روش‌های استاندارد کردن داده‌ها، روش نرمال‌سازی (کسر کردن مقدار متغیر تصادفی از میانگین داده‌ها و سپس تقسیم آن به انحراف استاندارد داده‌ها) است.

نظر گرفتن نقش زمین‌شناسی، به شاخصی از نفوذپذیری حوضه توجه شد. این عامل به‌طور معکوس بر روی شدت سیل‌خیزی حوضه تأثیر دارد. با استفاده از نقشه‌های زمین‌شناسی، اقدام به برآورد وضعیت نفوذپذیری حوزه‌ها گردید. با توجه به لیتولوژی سنگ‌ها، نفوذپذیری به‌طور کیفی در ۹ دسته به صورت بسیار ناچیز، ناچیز، خیلی کم، کم، متوسط تا کم، متوسط، متوسط تا زیاد، زیاد و خیلی زیاد در نظر گرفته شد. به هر کدام از این دسته‌ها به ترتیب اعداد (۰، ۰/۵، ۱، ۲، ۲/۵، ۳، ۴، ۵) اختصاص داده شد. سپس میزان نفوذپذیری نسبی در حوزه‌ها توسط میانگین وزنی (نسبت به مساحت) محاسبه و تحت عنوان شاخص زمین‌شناسی حوضه لحاظ شد. در شکل (۱) موقعیت



$$J(U,V : X) = \sum_{i=1}^c \sum_{k=1}^N (u_{ik})^m d^2(x_k, v_i) \quad (1)$$

که در آن، N تعداد نمونه و u_{ik} میزان تعلق نمونه i ام در خوشه k ام را نشان می‌دهد. m توان وزنی است که عددی حقیقی بزرگ‌تر از ۱

الگوریتم خوشه‌بندی فازی

کاربردی‌ترین الگوریتم خوشه‌بندی فازی، الگوریتم C-میانگین می‌باشد (۲). این الگوریتم بر اساس بهینه‌سازی تکراری تابع هدف فازی (رابطه‌ی ۱) می‌باشد.

هیدرولوژیکی همگن نیستند لذا جهت بهبود همگنی این ناحیه‌ها را باید تعدیل کرد. در تحلیل خوشه‌ای فازی آگاهی از توزیع عضویت یک حوضه در میان ناحیه‌های فازی در تعدیل ناحیه‌ها جهت بهبود همگنی آن‌ها مفید است (۲۲).

دو ابزار آماری مفید (اندازه‌ی ناهم‌نوایی برای مشخص کردن ایستگاه‌های ناهماهنگ در منطقه و اندازه‌ی ناهماهنگی جهت حصول اطمینان از همگن بودن مناطق) در تحلیل فراوانی سیلاب منطقه‌ای بر اساس گشتاورهای خطی وجود دارد (۱۲).

آزمون ناهمگنی تغییرات بین ایستگاهی گشتاورهای خطی نمونه برای گروهی از ایستگاه‌ها با مقدار مورد انتظار برای یک ناحیه همگن مقایسه می‌شود. عموماً سه آماره‌ی ناهمگنی H_1 برای بررسی ضریب تغییرات خطی (L_{CV})، آماره‌ی H_2 برای ترکیبی از ضریب تغییرات خطی و ضریب چولگی (L_{CS}) و آماره‌ی H_3 برای ترکیبی از ضریب چولگی خطی و ضریب کشیدگی خطی (L_{CK}) در بررسی تغییرپذیری سه آماره‌ی خطی مختلف به کار برده می‌شود. جزئیات بیشتر در هاسکینگ و والیس (۱۲ و ۱۳) آورده شده است.

برآورد دبی با استفاده از تحلیل فراوانی منطقه‌ای

در روش سیلاب‌نامه، گشتاورهای خطی جهت برآورد دبی با دوره بازگشت T سال ($Q_i(T)$) از رابطه‌ی زیر استفاده می‌شود.

$$Q_i(T) = \bar{Q}_i q(T), i = 1, \dots, N. \quad (5)$$

که در آن، \bar{Q}_i سیلاب نمایه یا شاخص سیلاب بوده، که می‌تواند سیل میانگین در نظر گرفته شود. در مناطق فاقد آمار، میانگین مورد نظر قابل محاسبه نیست. در چنین شرایطی، رابطه‌ی همبستگی بین مقادیر میانگین دبی اوج لحظه‌ای سالانه مناطق دارای آمار و ویژگی‌های فیزیوگرافی و اقلیمی آن‌ها برقرار می‌شود.

$$\bar{Q}_i = \theta_0 + \sum_{j=1}^k \theta_j z_j^{(i)} + u_i \quad (6)$$

در این رابطه، \bar{Q}_i متوسط دبی اوج لحظه‌ای (سیلاب نمایه)، $z_1^{(i)}, \dots, z_k^{(i)}$ ویژگی‌های منتخب فیزیوگرافی و اقلیمی ایستگاه i ام، $\theta_j, j = 0, 1, \dots, k$ پارامترهای مدل و u_i پارامتر خطا می‌باشد.

از آزمون نیکویی برآزش Z (۱۳) و آزمون کولموگروف-اسمیرنوف جهت انتخاب بهترین تابع توزیع استفاده گردید.

آماره‌های ارزیابی

جهت ارزیابی کلی دو روش تحلیل منطقه‌ای یک روش صحت سنجی به کار رفته است. در این روش یک ایستگاه را انتخاب و سپس با استفاده از روابط تحلیل منطقه‌ای (سیلاب نمایه و منحنی رشد

است. زمانی که $m = \infty$ باشد، داده‌ها تمایل دارند عضویت برابر در همه خوشه‌ها داشته باشند (۲۲). محدوده‌ی کلی m توسط راس (۲۴) در بازه $[2 - 1/25]$ و توسط پال و بزدک (۲۱) در بازه‌ی $[1/5 - 2/5]$ پیشنهاد شده است. x_k نمونه k ام است و v_i نماینده یا مرکز خوشه i ام است. $d^2(x_k, v_i)$ میزان تشابه (فاصله) نمونه k -ام با (از) مرکز خوشه i -ام می‌باشد که می‌توان از هر تابعی (مثلاً اقلیدسی) که بیانگر تشابه نمونه و مرکز خوشه باشد را استفاده کرد. توضیحات بیشتر در کوره پزان دزفولی (۱۵) و هاسکینگ و والیس (۲۲) آورده شده است. خوشه‌های فازی به وسیله‌ی اختصاص ایستگاه‌هایی که عضویت آن‌ها در خوشه از مقدار آستانه‌ای تجاوز کند تشکیل می‌شود. به‌طور کلی انتخاب مقدار آستانه برای ایجاد خوشه‌های فازی با قضاوت شخصی انجام می‌شود. در فازی‌ترین افراز، عضویت داده‌ها به تمامی خوشه‌ها مساوی $1/c$ است (۲۲).

یکی از مهم‌ترین مسایل در خوشه‌بندی، انتخاب تعداد خوشه‌های مناسب (حداکثر فشردگی خوشه‌ها و حداکثر جدایی بین آن‌ها) می‌باشد. برای مشخص کردن تعداد درست خوشه‌ها، ۳ تابع به صورت شاخص فوکویاما-سوگنو (۷) (رابطه‌ی ۲)، ژی-بِنی (۲۹) (رابطه‌ی ۳) و وون (۱۷) (رابطه‌ی ۴) در منابع وجود دارد که می‌توان با استفاده از آن‌ها تعداد خوشه‌ها را مشخص کرد.

$$V_{FS,m}(U, V; X) = \sum_{i=1}^c \sum_{k=1}^N u_{ik}^m (\|x_k - v_i\| - \|v_i - \bar{v}\|_A)^2 \quad (2)$$

$$V_{XB}(U; V; X) = \frac{c \sum_{i=1}^c \sum_{k=1}^N u_{ik}^m \|x_k - v_i\|^2}{n(\min\{v_i - v_j\})} \quad (3)$$

$$V_K(U, X) = \frac{\sum_{i=1}^c \sum_{k=1}^N (u_{ik})^m \|v_i - x_k\|^2 + \frac{1}{c} \sum_{i=1}^c \|v_i - \bar{v}\|}{\min_{i \neq k} \|v_i - v_k\|^2} \quad (4)$$

در رابطه‌ی ۲، \bar{v} میانگین کل نمونه‌ها است. جمله‌ی اول در تابع فوق معیاری برای فشردگی خوشه‌ها و جمله‌ی دوم معیاری برای جدایی خوشه‌ها از هم می‌باشد. بنابراین حداقل کردن تابع فوق معیار مناسبی برای ارزیابی خوشه‌بندی و تعداد خوشه‌ها می‌باشد. لذا کم‌ترین مقدار برای V_{FS} بیانگر بهترین افراز بوده که هم ارز خوشه‌های متراکم و دور از هم خواهد بود (۲۲).

هدف شاخص ژی و بنی، به کمیت درآوردن نسبت مجموع واریانس درون خوشه‌ها و میزان جداشدگی خوشه‌ها است. انتخاب تعداد خوشه‌های مناسب با حداقل کردن رابطه‌ی ۳ به‌دست می‌آید. مقدار شاخص ژی-بِنی به‌طور یکنواخت با افزایش تعداد خوشه‌ها کاهش می‌یابد (۲۲). جهت حل این مسئله، وون (۱۷) شاخص جدیدی را ارائه داد (رابطه ۴).

آزمون‌های گشتاورهای خطی

ناحیه‌های حاصل از الگوریتم‌های خوشه‌بندی به‌طور کلی از لحاظ

نظر از مقدار آستانه‌ای مدنظر بیش‌تر بود به ناحیه‌ی مربوطه تعلق یافت. ناحیه‌های حاصل از روش اول (با در نظر گرفتن شاخص زمین شناسی) و روش دوم (بدون در نظر گرفتن شاخص زمین‌شناسی) در جدول (۲) ارایه شده است. همان‌گونه که از این دو جدول پیداست، نواحی A و A' و نیز ناحیه‌های C و C' کاملاً مشابه بوده ولی ایستگاه‌های تعلق یافته به ۴ ناحیه‌ی باقی‌مانده در دو روش به مقدار جزئی متفاوت می‌باشند. حوضه‌های تعلق یافته به ناحیه A شامل حوضه‌های آبریزی است که از لحاظ مساحت بسیار کوچک‌اند، در حالی که حوضه‌های تعلق یافته به ناحیه‌ی C شامل حوضه‌های بسیار بزرگی است که از لحاظ مساحت با سایر حوضه‌های ناحیه همخوانی ندارند بنابراین در هر دو روش، این ۵ حوضه در یک دسته‌ی واحد واقع شده‌اند. همان‌گونه که از جدول (۲) پیداست اندازه‌ی ناحیه‌ها (تعداد ایستگاه‌های واقع شده در ناحیه) در روش دوم که از شاخص زمین‌شناسی در دسته‌بندی ایستگاه‌ها استفاده نشده بزرگتر است که این خود بیانگر همپوشانی بیشتر ناحیه‌ها می‌باشد.

آزمون‌های گشتاورهای خطی

آماره‌ی ناهم‌نوابی برای همه‌ی ایستگاه‌ها محاسبه گردید. بر پایه‌ی نتایج، ایستگاه ازغند (که دارای کمترین طول دوره آماری، ۵ سال، می‌باشد) ناهم‌نوا بوده و از مجموعه ایستگاه‌ها حذف شد. در مطالعه‌ی شامکوئیان و همکاران (۲۷) نیز ایستگاهی که از طول دوره‌ی آماری کوتاه‌تری (۱۰ سال) نسبت به سایر ایستگاه‌ها برخوردار بود، ناهم‌نوا گزارش شده است.

نتایج آزمون همگنی (جدول ۴)، حاکی از این است که تمام نواحی حاصل از روش اول (با در نظر گرفتن شاخص زمین‌شناسی) همگن است، هم‌چنین تمام نواحی حاصل از روش دوم (بدون در نظر گرفتن شاخص زمین‌شناسی) بجز ناحیه B' که نسبتاً همگن است نیز همگن می‌باشند لذا جهت تحلیل فراوانی سیلاب مناسب می‌باشند.

روابط چندک بی‌بعد ناحیه‌ای

جهت تعیین توابع توزیعی که بهترین برازش را با داده‌های هر یک از نواحی مذکور دارند، از آزمون‌های نیکوئی برازش Z و کولموگروف-اسمیرنوف به‌طور توأم استفاده شد. در جدول (۴) توابع توزیع منتخب هر ناحیه و نیز پارامترهای آن‌ها ارائه شده است. در مطالعه‌ی شامکوئیان و همکاران (۲۷) علاوه بر توابع انتخابی در هر ناحیه به عنوان بهترین حالت ممکن، تابع توزیع مقادیر حدی تعمیم یافته نیز برای تمامی نواحی در سطح $\alpha=0/2$ معنی‌دار بود. در این مطالعه در روش اول (در نظر گرفتن شاخص زمین‌شناسی) تابع توزیع لوگ نرمال سه پارامتری برای تمامی نواحی در سطح $\alpha=0/2$ معنی‌دار است ولی برای نواحی حاصل از روش دوم (بدون در نظر گرفتن

ناحیه‌ای) مقدار سیلاب با دوره بازگشت‌های مختلف برای ایستگاه مورد نظر محاسبه شده و مقدار خطا محاسبه می‌شود و این رویه برای تمام ایستگاه‌ها تکرار می‌شود. میزان کارایی روش تحلیل منطقه‌ای توسط میانه خطای نسبی (RE) ارزیابی می‌شود.

$$RE(\%) = \text{mediane}_{i=1}^n \left[\text{abs} \left(\frac{Q_{pred,i} - Q_{obs,i}}{Q_{obs,i}} \right) \right] \times 100 \quad (7)$$

$Q_{obs,i}$ مقدار سیلاب حاصل از تحلیل فراوانی نقطه‌ای است و $Q_{pred,i}$ دبی محاسباتی با استفاده از روابط منطقه‌ای (روابط ۵ و ۶) و n تعداد ایستگاه‌های موجود در ناحیه می‌باشد. آماره‌ی ارزیابی دیگر، مقدار $\frac{Q_{pred,i}}{Q_{obs,i}}$ که اگر این مقدار به‌طور قراردادی در بازه‌ی ۰/۵-۲ باشد مناسب است و برای مقادیر کم‌تر از ۰/۵ و بیش‌تر از ۲، به‌ترتیب برآورد خیلی کم و خیلی زیاد، روابط منطقه‌ای مفید وجود خواهد داشت (۳۰).

نتایج و بحث

در این مطالعه جهت ارزیابی میزان تأثیر استفاده از شاخص زمین‌شناسی در تحلیل فراوانی سیلاب منطقه‌ای مراحل تحقیق طی دو روش اجرا شد:

روش اول: تحلیل فراوانی سیلاب منطقه‌ای با در نظر گرفتن شاخص زمین‌شناسی
روش دوم: تحلیل فراوانی سیلاب منطقه‌ای بدون در نظر گرفتن شاخص زمین‌شناسی

نتایج تحلیل خوشه‌ای

پس از استاندارد کردن مشخصه‌های حوضه‌ها، با اجرای الگوریتم C-میانگین فازی به ازاء مقادیر مختلف m (توان وزنی) با فواصل ۰/۱ و تغییر مقدار c (تعداد خوشه)، مقادیر سه شاخص فوکویاما-سوگونو (V_{FS})، ژی-بنی ($V_{XB,m}$) و وون (V_K) محاسبه شد. همان‌گونه که از جدول (۱) پیداست مقدار شاخص فوکویاما-سوگونو با افزایش تعداد خوشه‌ها به‌طور یکنواخت افزایش می‌یابد، بنابراین در این جا کارایی لازم را ندارد. مقدار بهینه‌ی دو شاخص ژی-بنی و وون نشان می‌دهد که $c=9$ بهترین تعداد برای خوشه‌هاست، زمانی که $m=1/4-1/7$ و زمانی که $m=1/8-2/2$ ، $c=6$ بهترین تعداد خوشه‌ها می‌باشد. هم‌چنین به ازاء مقادیر $m < 1/9$ مقدار دو شاخص به‌طور معنی‌دار افزایش می‌یابد، لذا تعداد ۶ خوشه با توان وزنی ۱/۹ که حداقل مقدار شاخص‌ها را به خود اختصاص داده است، به‌عنوان افزاز بهینه، انتخاب می‌شود. جهت استخراج ناحیه‌های هیدرولوژیکی، با در نظر گرفتن مقدار آستانه‌ای ۱/c، ایستگاه‌هایی که درجه عضویت آن‌ها به خوشه‌ی مورد

شاخص زمین‌شناسی) هیچ تابع توزیعی در تمامی نواحی، در هیچ کاربردی‌تر و مؤثرتر واقع شود، بنابراین در نظر گرفتن شاخص زمین سطح احتمالی (0/05, 0/1, 0/1, 0/2) معنی‌دار نبود. استفاده از یک نوع تابع توزیع برای تمامی نواحی مورد مطالعه، می‌تواند

جدول ۱- مقادیر سه شاخص صحت‌سنجی وون (V_K)، زی-بنی ($V_{XB,m}$) و فوکویاما-سوگنو (V_{FS}). C تعداد خوشه‌ها است
 Table 1- Values of three validation criteria of Kwon (V_K), Xie and Beni ($V_{XB,m}$), and Fukuyama and Sugeno (V_{FS}). C is number of clusters

$V_{XB,M}$	V_{FS}	V_K	m	$V_{XB,M}$	V_{FS}	V_K	m	$V_{XB,M}$	V_{FS}	V_K	m	C
0.76	194	55.69		0.75	217	54.69		0.74	235	54.09		2
0.59	91.5	43.54		0.49	40.9	36.89		0.55	25	41.23		3
0.57	5.84	43.76		0.58	-28	44.21		0.48	-62	36.51		4
0.94	-57.44	74.82		0.51	-77	44.28		0.45	-123	34.28		5
0.35	-92	28.5	2	0.41	-116	32.52	1.7	0.51	-163	38.84	1.4	6
0.78	-107	65		0.45	-139	36.28		0.52	-188	37.41		7
0.88	-121	75		0.44	-168	30.9		0.44	-209	51.36		8
0.73	-127	61.8		0.36	-175	30.6		0.43	-236	32.51		9
0.51	-142	56.8		0.50	-209	42		0.38	-244	28.1		10
0.77	186	56.2		0.75	210	54.95		0.74	230	54.3		2
0.85	92.8	62.7		0.47	52.4	34.98		0.53	27.73	40		3
0.56	14.25	42.36		0.6	-18	45.5		0.51	-51.6	38.6		4
1.15	-53.8	93.7		0.54	-67.39	50.54		0.47	-108	36.1		5
0.37	-82.7	30.8	2.1	0.36	-107	29.1	1.8	0.48	-145	37.7	1.5	6
0.70	-99.7	60		0.43	-128	39.6		0.41	-165	40		7
0.72	-105	66.8		0.44	-149	34.6		0.45	-215	34.77		8
0.86	-112	73.8		0.54	-169	52.7		0.38	-240	34		9
1.47	-120	1376		0.45	-187	60.88		0.35	-253	44.9		10
0.77	178	56.77		0.75	202	55.28		0.74	224	54.48		2
1.15	90	84.63		0.44	68	32.53		0.51	32.78	38.5		3
0.54	20	41.28		0.59	-6	45.18		0.54	-40	41.46		4
1.33	-5.82	103.4		0.76	-61.24	60.04		0.49	-91	37.6		5
1.32	-72.52	200	2.2	0.34	-99.8	27.2	1.9	0.45	-127	35.35	1.6	6
595	-83.66	29817		0.81	-114	66.37		0.49	-149	38.45		7
2.4	-13.91	2966710		0.83	-130	54.69		0.44	-174	34.9		8
185	-96.92	155.67		0.75	-148	65.6		0.38	-213	30.4		9
114	-101	2119		0.46	-156	69.56		0.54	-236	44.4		10

جدول ۲- توزیع حوضه‌ها در نواحی. حروف انگلیسی ساده مربوط به در نظر گرفتن ویژگی زمین‌شناسی و حروف انگلیسی با علامت پریم، مربوط به در نظر نگرفتن شاخص زمین‌شناسی است

Table 2- Distribution of watersheds in regions. For plain English letters, geological feature is included, while for the other case (with prime) this feature is not included

ایستگاه‌های واقع در هر ناحیه Stations in the region	تعداد ایستگاه Number of stations	ناحیه Region
۱۱،۱۲،۱۳،۱۴،۲۸،۳۳،۳۴،۳۵،۳۷،۴۷،۴۸،۵۴،۶۹،۷۱	14	A' و A
۱،۳،۵،۱۶،۱۷،۳۲،۳۸،۳۹،۴۳،۴۵،۴۶،۴۹،۵۰،۵۱،۵۳،۶۰،۷۰	17	B
۱،۳،۵،۱۶،۱۷،۳۱،۳۲،۳۶،۳۸،۳۹،۴۳،۴۵،۴۶،۴۹،۵۰،۵۱،۵۳،۶۰،۷۰	19	B'
۱۵،۱۶،۴۳،۵۹،۶۳	5	C' و C
۴۶،۷۹،۱۹،۲۰،۲۱،۲۲،۲۳،۲۴،۴۰،۴۲،۵۷،۶۲،۶۳،۶۵،۶۶،۶۸،۷۲،۷۳	20	D
۲،۴،۶،۸،۹،۱۰،۲۰،۲۱،۲۲،۲۳،۲۴،۲۵،۳۹،۴۲،۴۴،۴۶،۴۷،۴۹،۵۲،۵۵،۵۶،۵۷،۵۸،۶۱،۶۶،۶۷،۶۸	27	D'
۱،۲،۳،۸،۱۰،۱۷،۲۵،۲۶،۲۷،۲۸،۲۹،۳۰،۳۱،۳۲،۳۶،۳۷،۳۹،۴۴،۴۷،۴۹،۵۱،۵۲،۵۴،۵۵،۵۷،۶۰،۶۷،۶۸،۷۰	29	E
۱،۲،۳،۵،۸،۹،۱۰،۱۳،۱۷،۲۵،۲۶،۲۷،۲۸،۲۹،۳۰،۳۱،۳۲،۳۶،۳۷،۳۹،۴۴،۴۷،۴۹،۵۴،۵۵،۶۰،۶۷،۶۸،۷۰،۷۱،۷۲	31	E'
۱۸،۱۹،۳۹،۴۰،۴۱،۴۲،۴۶،۵۲،۵۶،۵۷،۵۸،۶۱،۶۳،۶۴،۶۵	15	F
۷،۱۸،۱۹،۲۰،۴۰،۴۱،۴۲،۴۶،۶۱،۶۲،۶۳،۶۴،۶۵،۶۶،۷۲،۷۳	16	F'

جدول ۳- نتایج آزمون همگنی
Table 3- Results of homogeneity test

وضعیت همگنی Homogeneity	H_3	H_2	H_1	ناحیه Region	وضعیت همگنی Homogeneity	H_3	H_2	H_1	ناحیه Region
همگن	-0.492	-0.313	0.122	A'	همگن	-0.492	-0.313	0.122	A
نسبتاً همگن	-0.154	-0.255	-1.02	B'	همگن	-0.118	-0.075	-0.859	B
همگن	-0.123	-0.078	-0.912	C'	همگن	-0.123	-0.078	-0.912	C
همگن	-0.073	-0.061	-0.082	D'	همگن	-0.253	-0.442	-0.795	D
همگن	-0.193	-0.707	-0.505	E'	همگن	-0.425	-0.311	-0.72	E
همگن	-0.217	-0.417	-0.733	F'	همگن	-0.36	-0.254	-0.391	F

جدول ۴- توزیع‌های آماری تعیین شده برای نواحی با استفاده از آزمون‌های نکویی برازش
Table 4- Distribution functions for regions by using goodness of fit tests

Region ناحیه	Distribution function توزیع آماری انتخابی	ξ	α or μ	k or σ
A and A'	Three parameter log-Normal لوگ نرمال سه پارامتری	-0.008	-0.4664	0.9741
B	Generalized Pareto پارتو تعمیم یافته	-0.0028	0.8409	-0.1631
C and C'	Three parameter log-Normal لوگ نرمال سه پارامتری	-0.0872	-0.3299	0.9145
D	Three parameter log-Normal لوگ نرمال سه پارامتری	-0.0658	-0.4493	0.0129
E	Three parameter log-Normal لوگ نرمال سه پارامتری	-0.0629	-0.4397	1.001
F	Generalized Extreme Value	0.4025	-0.0338	-0.4694
B'	Three parameter log-Normal لوگ نرمال سه پارامتری	-0.1632	-0.2441	0.891
D'	Three parameter log-Normal لوگ نرمال سه پارامتری	-0.0316	-0.5169	0.9925
E'	Three parameter log-Normal لوگ نرمال سه پارامتری	-0.0931	-0.3517	0.9388
F'	Generalized Logestic لجستیک تعمیم یافته	0.0065	-0.3298	-0.5192

$$\text{Three parameter log-Normal } x(F) = \xi + \exp(\mu + \sigma \Phi^{-1}(F)),$$

$$\text{Generalized Extreme Value } x(F) = \xi + \frac{\alpha}{k} \left\{ 1 - (-\log F)^k \right\},$$

$$\text{Generalized Logestic } x(F) = \xi + \frac{\alpha}{k} \left(1 - \left(\frac{1-F}{F} \right)^k \right),$$

$$\text{Generalized Pareto } x(F) = \xi + \frac{\alpha}{k} \left[1 - (1-F)^k \right]$$

نهایت مدلی با ویژگی‌های آماری مناسب‌تر (ضریب همبستگی بیشتر و متوسط خطای استاندارد کمتر) انتخاب شد. مدل لگاریتمی چهار متغیره با استفاده از مقادیر متوسط دبی سیلاب اوج لحظه‌ای بر حسب متر مکعب بر ثانیه (μ) به‌عنوان متغیر وابسته و سایر پارامترهای

تخمین متوسط دبی سیلاب لحظه‌ای (سیلاب‌نمایه)

به منظور برآورد سیلاب‌نمایه (رابطه ۶) برای هر یک از نواحی تحت مطالعه، ترکیبات مختلفی از متغیرها مستقل، در نظر گرفته شد و مدل لگاریتمی آن در هر ناحیه مورد بررسی و آزمون قرار گرفت، و در

ارائه شده است. مقدار ضریب همبستگی مدل در ناحیه F برابر ۰/۲۹ به‌دست آمد. جهت افزایش همبستگی در ناحیه‌ی مورد نظر و به منظور افزایش دقت برآورد سیلاب‌نمایه، ناحیه‌ی مورد نظر با ناحیه‌های دیگر (A, B, C, D, E)، دو به دو ترکیب شد. زمانی که از مجموع داده‌های این ناحیه و ناحیه E استفاده شد، مقدار ضریب همبستگی مدل بیشترین مقدار (۰/۵۵) را داشت که از این مدل برای برآورد سیلاب نمایه تنها در ناحیه‌ی F استفاده شد.

فیزیوگرافی همچون شیب متوسط (S)، ارتفاع متوسط برحسب متر (H)، مساحت برحسب متر مربع (A) و طول بلندترین آبراهه برحسب متر به‌عنوان متغیرهای مستقل حوضه‌های آبریز در هر ناحیه استخراج شد. شکل کلی این روابط به صورت زیر است:

$$\text{Log}(\mu) = \theta_0 + \theta_1 \text{Log}(H) + \theta_2 \text{Log}(A) + \theta_3 \text{Log}(S) + \theta_4 \text{Log}(L) \quad (8)$$

در جدول (۵) ضرائب مدل لگاریتمی پیشنهادی و مقادیر متوسط خطای استاندارد (MSE) و آماره‌ی R مربع برای هر یک از نواحی

جدول ۵- ضرائب مدل لگاریتمی چهار متغیره پیشنهادی رابطه (۸) برای هر یک از نواحی
Table 5- Coefficients of 4-parameters logarithmic model (Eq. 8) for all regions

R ²	MSE	θ_0	θ_1	θ_2	θ_3	θ_4	Region ناحیه
0.7	0.27	-19	2.76	1.55	-1/2	-2	A
0.67	0.17	17.5	-2	0.71	-0.5	-1.1	B
1	0	-25	2.14	-1.2	3.31	1.12	C
0.54	0.39	3.68	-0.8	0.72	-0.1	-0.7	D
0.61	0.2	5.62	-1.3	0.64	1.01	-0.7	E
0.55	0.3	5.05	-1.24	0.61	0.55	-0.52	F
0.66	0.16	16.59	-1.86	0.61	-0.37	-0.97	B'
0.49	0.36	8.53	-1.93	0.84	0.88	-0.9	D'
0.56	0.19	6.99	-1.23	0.54	0.45	-0.58	E'
0.42	0.45	18.1	-2.38	0.22	-0.14	-0.05	F'

در حالی که زمان و همکاران (۳۰) برای مناطق خشک و نیمه‌خشک استرالیا مقدار این آماره را در بازه‌ی ۵۷-۶۹ درصد گزارش دادند. به دلیل تغییرپذیری بیشتر رخدادهای هیدرولوژیک در مناطق خشک و نیمه خشک نسبت به مناطق مرطوب- ساحلی، خطای برآورد چارک‌ها تا ۱/۵-۲ برابر بیشتر از مناطق مرطوب- ساحلی، قابل قبول است (۳۰). مقادیر محاسبه شده میان‌ه‌ی خطای نسبی در مطالعه‌ی حاضر نشان می‌دهد که این مقادیر نسبت به مناطق خشک و نیمه‌خشک استرالیا خیلی کمتر است که این به دلیل تفکیک منطقه به نواحی

همگن می‌باشد. همچنین مقادیر آماره‌ی $\frac{Q_{pred,i}}{Q_{obs,i}}$ در جدول (۷)

نشان می‌دهد که در روش اول ۷۵-۹۵ درصد ایستگاه‌ها و در روش دوم ۷۸-۹۷ درصد ایستگاه‌ها برآورد قابل قبولی را دارا هستند.

آماره‌های ارزیابی

جهت ارزیابی کلی دو روش، مقادیر دبی‌های برآوردی با استفاده از روابط منطقه‌ای (Q_{pred}) برای تمام ایستگاه‌ها به‌ازاء ۵ دوره بازگشت (۵، ۱۰، ۲۰، ۵۰ و ۱۰۰ سال) با مقادیر دبی‌های برآوردی به روش تحلیل نقطه‌ای (Q_{obs}) مقایسه می‌شود. برای محاسبه‌ی Q_{obs} توابع توزیعی که بهترین برازش را با داده‌های جریان هر ایستگاه دارند با استفاده از دو آزمون نیکویی برازش کولموگروف- اسمیرنوف و مجموع مربعات باقیمانده انتخاب شدند. میان‌ه‌ی خطای نسبی با استفاده از رابطه‌ی (۷) محاسبه شد (جدول ۶). با مقایسه‌ی مقادیر آماره‌ی ارزیابی ناحیه‌های دو روش (بدون در نظر گرفتن ناحیه‌هایی که در دو روش متشابهند) مقادیر میان‌ه‌ی خطای نسبی در روش اول در بازه‌ی ۰/۲۱-۰/۴۵ درصد و در روش دوم در بازه‌ی ۰/۲۱-۰/۴۹ درصد می‌باشد. حداد و رحمان (۹) برای نواحی مرطوب- ساحلی استرالیا میان‌ه‌ی خطای نسبی را در بازه‌ی ۳۰-۴۰ درصد گزارش دادند

جدول ۶- مقادیر آماره ارزیابی میانه خطای نسبی ناحیه‌ها به ازاء ۵ دوره بازگشت

Table 6- Median of relative error for regions corresponding to 5 return periods

Median of relative error (RE)					Region ناحیه
میانۀ خطای نسبی (RE)					
T=100	T=50	T=25	T=10	T=5	
0.53	0.45	0.35	0.29	0.16	A
0.23	0.27	0.35	0.37	0.33	B
0.14	0.01	0.04	0.01	0.04	C
0.39	0.37	0.34	0.36	0.3	D
0.43	0.31	0.21	0.24	0.23	E
0.36	0.45	0.33	0.36	0.39	F
0.37	0.32	0.29	0.31	0.29	B'
0.44	0.39	0.31	0.26	0.26	D'
0.38	0.33	0.24	0.21	0.25	E'
0.28	0.26	0.35	0.43	0.49	F'

$$\frac{Q_{pred,i}}{Q_{obs,i}}$$

جدول ۷- مقادیر آماره ارزیابی

Table 7- Evaluation criterion of $\frac{Q_{pred,i}}{Q_{obs,i}}$

$\frac{Q_{pred,i}}{Q_{obs,i}}$			Region ناحیه
<0.5(%)	2-0.5(%)	>2(%)	
9	76	16	A
2.5	95	2.5	B
0	100	0	C
9	75	16	D
3	90	7	E
7	75	8	F
1	97	2	B'
12	85	3	D'
4	89	6	E'
3	78	20	F'

نگرفتن آن)، تفاوت زیادی در توزیع حوضه‌ها در نواحی مشاهده نشد به‌علاوه، برآورد سیلاب در هر دو روش نتایج قابل قبولی به همراه داشت و می‌توان گفت که هر دو روش خوب بودند و اضافه کردن شاخص زمین‌شناسی منجر به بهبود معنی‌دار نتایج نگردید. زمانی که از شاخص زمین‌شناسی نیز جهت ناحیه‌بندی استفاده شد علاوه بر توابع انتخابی در هر ناحیه به‌عنوان بهترین حالت ممکن، تابع توزیع لوگ نرمال سه پارامتری برای تمامی نواحی در سطح $\alpha=0/2$ ، معنی‌دار بود. استفاده از یک نوع تابع توزیع برای تمامی نواحی مورد مطالعه، می‌تواند کاربردی‌تر و مؤثرتر واقع شود، بنابراین در نظر گرفتن شاخص زمین‌شناسی از این نقطه نظر مناسب است.

با مقایسه‌ی توزیع ایستگاه‌ها و همچنین دو آماره‌ی ارزیابی میانه‌ی خطای نسبی (RE) و نسبت Q_{pred}/Q_{obs} مشاهده شد که هر دو روش خوب بودند و اضافه کردن شاخص زمین‌شناسی نتایج را به میزان اندک بهبود بخشید.

نتیجه‌گیری

نتایج تحلیل خوشه‌ای فازی حوضه‌های آبریز شمال شرق کشور نشان داد که تحلیل خوشه‌ای فازی روشی مناسب جهت تعیین ناحیه‌های همگن منطقه مورد نظر می‌باشد. همچنین با مقایسه‌ی دو روش تحلیل منطقه‌ای سیلاب (در نظر گرفتن شاخص زمین‌شناسی و در نظر

منابع

- 1- Alvankar S.R. 2011. Estimating maximum flood discharge by using Landsat satellite data. The 4th Conference on

- Iran Water Resources Management. Amirkabir Industrial University. May 2-3, 2011. Tehran. 12 pages (in Persian).
- 2- Ataee H., and Shiran M. 2011. Identifying homogeneous hydrologic watersheds based on geomorphologic feature effective on flood by using cluster analysis (case study: Kroon plain). *Journal of Geography and environmental planning*, 42, 79-98 (in Persian).
 - 3- Bezdek J.C. 1981. *Pattern Recognition with Fuzzy Objective Function Algorithms*. Plenum Press, New York.
 - 4- Dalrymple T. 1960. *Flood frequency analysis*. US Geological Survey Water Supply Paper 1543-A.
 - 5- Das S., and Cunnane C. 2011. Examination of homogeneity of selected Irish pooling groups. *Hydrology and Earth System Sciences*, 15: 819-830.
 - 6- Dunn J.C. 1974. A fuzzy relative of the ISODATA process and its use in detecting compact, well-separated clusters. *Journal of Cybernetics*, 3 (3): 32-57.
 - 7- Fukuyama Y., and Sugeno M. 1989. A new method of choosing the number of clusters for the fuzzy c-means method. *Proceedings of Fifth Fuzzy Systems Symposium*, pp. 247-250 (in Japanese).
 - 8- Goel G.N. 2000. The formation of groups for regional flood frequency analysis. *Hydrological Sciences-Journal-des, Sciences Hydrologiques*; 45(1) 97-112.
 - 9- Haddad K., and Rahman A. 2012. Regional flood frequency analysis in eastern Australia: Bayesian GLS regression-based methods within fixed region and ROI framework: Quantile regression vs. parameter regression technique. *Journal of Hydrology*. 430-431, 142-161.
 - 10- Hall M.J., and Minns A.W. 1999. The classification of hydrologically homogeneous regions. *Hydrological Sciences Journal*, 44 (5): 693-704.
 - 11- Hosking J.R.M. 1990. L-Moments: Analysis and estimation of distributions using linear combinations of order statistics. *Journal of Royal Statistical Society B*, 52: 105-124.
 - 12- Hosking J.R.M., and Wallis J.R. 1993. *Some statistics useful in regional frequency analysis*, Res. Rep. RC 17096, IBM Research Division, York town Heights, NY 10598.
 - 13- Hosking J.R.M., and Wallis J.R. 1997. *Regional Frequency Analysis (An Approach Based on Linear Moments)*. Cambridge University Press.
 - 14- Jingyi Z., and Hall M.J. 2004. Regional flood frequency analysis for the Gan-Ming River basin in China. *Journal of Hydrology*, 296: 98-117.
 - 15- Koorehpazan Dezfooli A. 2008. *Theoretical Principles of Fuzzy sets and its Applications in Problems of Water Engineering Modeling*. Amirkabir Industrial University Publication, 8737. Tehran. 261 pp (in Persian).
 - 16- 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 Resources Management*, 17: 243-257.
 - 17- Kwon S.H. 1998. Cluster validity index for fuzzy clustering. *Electronics Letters* 34 (22), 2176-2177.
 - 18- Mishra B.K., Takara K., and Tachikawa Y. 2008. Regionalization of Nepalese river basins for flood frequency analysis. *Annual Journal of Hydraulic Engineering, JSCE*, 52: 91-96.
 - 19- Mooshkhan Y., Ownagh M., Bardy Shykh V., Mosaedi A., and Saad-o-ddin A. 2012. Developing regional models for estimating flood discharge in selected watersheds of Khorasan Razavi watersheds. 8th National Meeting on Science and Engineering of Watershed Management. May 1-17, 2012. Lorrestan University (in Persian)
 - 20- Mosley M.P. 1981. Delimitation of New Zealand hydrologic regions. *Journal of Hydrology*, 49: 173-192.
 - 21- Pal N.R., and Bezdek J.C. 1995. On cluster validity for the fuzzy c-means model. *IEEE Transactions on Fuzzy systems*, 3 (3): 370-379.
 - 22- Rao A.R., and Srinivas V.V. 2006. Regionalization of watersheds by fuzzy cluster analysis. *Journal of Hydrology*, 318: 57-79.
 - 23- Roa A.R., and Srinivas V.V. 2006. Regionalization of watersheds by Hybrid-Cluster analysis. *Journal of Hydrology*, 318: 37-56.
 - 24- Ross T.J. 1995. *Fuzzy Logic with Engineering Applications*. McGraw-Hill, New York.
 - 25- Saf B., Dikbas F., and Yasar M. 2007. Determination of regional frequency distributions of floods in West Mediterranean river basins in Turkey. *Fresenius Environmental Bulletin*, 16 (10): 1300-1308.
 - 26- Saf B. 2010. Assessment of the effects of discordant sites on regional flood frequency analysis. *Journal of Hydrology*, 380: 362-375.
 - 27- Shamkooian H., Ghahraman B., Davary K., and Sarmad M. 2009. Regional flood frequency analysis by using Linear Moment theory and flood index method in Khorasan Province watersheds. *Journal of Soil and Water (Agricultural Science and Industry)*. 23(1), 31-43 (in Persian).
 - 28- Sveinsson O.G.B., Boes D.C., and Salas J.D. 2001. Population index flood method for regional frequency analysis. *Water Resources Research*, 37 (11): 2733-2748.
 - 29- Xie X.L., and Beni G. 1991. A validity measure for fuzzy clustering. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 13 (8): 841-847.
 - 30- Zaman M.A., Rahman A., and Haddad K. 2012. Regional flood frequency analysis in arid regions: A case study for Australia. *Journal of Hydrology*, 475: 74-83.

Effect of Geological Feature on Regional Flood Frequency of Watersheds of North-East of Iran

T. Taherpoor¹- B. Ghahraman^{2*}- K. Davary³

Received: 24-12-2013

Accepted: 30-07-2017

Introduction: Finding out homogeneous watersheds based on their flood potential mechanisms, is needed for conducting regional flood frequency analysis. Similarity of watersheds based on flood potential severity depends on many factors such as physiographic and meteorological features of the watershed, geographical location and geological features. These criteria although are sound ones, they suffer from this concept that there is no attention to hydrological losses of runoff into the soil. As a result, current literature lacks for considering geological features into delineating homogeneous regions. The primary contribution of this paper is to include one geological criterion on flood regionalization. In a previous study we made a homogeneous classification for Khorasan Province of Iran without taking into consideration of infiltration features of the region. So, by taking geological features there may provide a sound comparison to regionalization issue.

Materials and Methods: To find out the effect of geological feature on delineation of homogeneous regions, 73 hydrometric stations at North-East of Iran with arid and semi-arid climate covering an average of 29 years of record length were considered. Initially, all data were normalized. Watersheds were clustered in homogeneous regions adopting Fuzzy c-mean algorithm and two different scenarios, considering and not considering a criterion for geological feature. Three validation criteria for fuzzy clustering, Kwon, Xie-Beni, and Fukuyama-Sugeno, were used to learn the optimum cluster numbers. Homogeneity approval was done based on linear moment's algorithm for both methods. We adopted 4 common distributions of three parameter log-Normal, generalized Pareto, generalized extreme value, and generalized logistic. Index flood was correlated to physiographic and geographic data for all regions separately. To model index flood, we considered different parameters of geographical and physiological features of all watersheds. These features should be easily-determined, as far as practical issues are concerned. Cumulative distribution functions for all regions were chosen through goodness of fit tests of Z and Kolmogorov-Smirnov.

Results and Discussion: Watersheds were clustered to 6 homogenous regions adopting Fuzzy c-mean algorithm, in which fuzziness parameter was 1.9, under the two different scenarios, considering and not considering a criterion for geological feature. Homogeneity was approved based on linear moment's algorithm for both methods, although one discordant station with the lowest data was found. For the case with inclusion of geologic feature, 3-parameter lognormal distribution was selected for all regions, which is a highly practical result. On the other hand, for not considering this feature there were no unique distribution for all regions, which fails for practical usages. As far as index flood estimation is concerned, a logarithmic model with 4 variables of average watershed slope, average altitude, watershed area, and the longest river of the watershed was found the best predicting equation to model average flood discharge. Determination coefficient for one of the regions was low. For this region, however, we merged this region to other regions so that reasonable determination coefficient was found; the resulting equation was used only for that specific region, however. By comparing the distributions of stations and also two evaluation statistics of median relative error and predicted discharge to estimated discharge ration corresponding to 5 different return periods (5, 10, 20, 50, and 100 years). Both perspectives showed acceptable results, and including geological feature was effective for flood frequency studies. With considering the geological feature for regionalization, Besides, Log normal 3 parameters distribution was found appropriate for all of the regions. From this point of view, geological feature was useful. Median of relative error was lower for small return periods and gradually increased as return period was increased. Median of relative error was between 0.21 to 00.45 percentages for the first method, while for the second method it varied between 0.21 to 0.49 percentages. These errors are quite smaller than those reported in literature under the same climatic region of arid and semi-arid. The probable reason may due to the fact that we made a satisfactory regionalization via fuzzy logic algorithm., We considered another mathematical criterion of "predicted discharge to the observed discharge". The optimum range for this criterion is between 0.5 and 2. While under-estimation and over-estimation are found if this criterion is lower than 0.5 and higher than 2,

1, 2 and 3- Former Graduate Student of Water Resources and Professors of Water Engineering Department, Ferdowsi University of Mashhad, Mashhad, Iran, Respectively
(*- Corresponding Author Email: bijangh@um.ac.ir)

respectively. Based on this premise, 75 to 95 percentages of stations were categorized as good estimation under the first method of analysis. On the other hand, 78 to 97 percentages of stations were considered good for the second approach.

Keywords: Flood index, Fuzzy cluster analysis, Infiltration measure, Non-data watersheds