نشریه آب و خاک (علوم و صنایع کشاورزی)
جلد ۲۶، شماره ۴، مهر – آبان ۱۳۹۱، ص. ۹۵۲-۹۴۳

Journal of Water and Soil
Vol. 26, No.4, Sept.-Oct. 2012, p. 943-952

تعیین ناحیه همگن ایستگاه فاقد آمار در آنالیز منطقه‌ای سیلاب به روش همبستگی متعارف

محمد حسین نوری قیداری[۱]

تاریخ دریافت:۹۰/۶/۱۳
تاریخ پذیرش:۹۰/۱۱/۲۳

چکیده

آنالیز منطقه‌ای سیلاب یک روش متداول برای برآورد سیلاب در ایستگاه فاقد آمار می‌باشد که دقت آن وابسته به میزان همگنی ایستگاه‌های هیدرومتری مورد استفاده در آنالیز منطقه سیلاب دارد. یکی از روش‌های معمول در ناحیه بندی ایستگاه‌ها روش خوشه‌بندی است که کارایی آن وابسته به شناخت کامل متغیرهای حوضه و وزن آنها می‌باشد که معمولاً میسر نمی‌گردد. در این تحقیق از همبستگی متعارف[۲] بین متغیرهای جریان و حوضه برای تعیین ناحیه مربوط به ایستگاه فاقد آمار استفاده شده است. در این روش با ایجاد زوج متغیرهای متعارف و به کمک رابطه بین آنها ایستگاه‌های مشابه به ایستگاه فاقد آمار شناسایی می‌گردد. این روش بطور موردی برای حوضه آبریز دریاچه ارومیه بکار برده شده که نتایج نشان می‌دهد ناحیه همگن ایستگاه فاقد آمار که به روش همبستگی متعارف تشکیل شده دارای تعداد ایستگاه زیادتری بوده و خطای برآورد منطقه‌ای چندک‌های سیلاب در ایستگاه فاقد آمار در مقایسه با روش خوشه‌ای کمتر می‌باشد.

واژه های کلیدی: برآورد منطقه‌ای، دریاچه ارومیه، متغیر متعارف، ناحیه همگن

مقدمه

سیلاب طرح یکی از پارامترهای مهم طراحی سازه‌های هیدرولیکی بوده که عدم برآورد صحیح آن می‌تواند خسارت جانی و مالی بسیاری به بار آورد. یکی از مشکلات اساسی در برآورد سیلاب طرح نبود و یا کمبود داده ثبت شده در محل مورد نظر می‌باشد. آنالیز فراوانی منطقه‌ای سیلاب یک روش مناسب جهت افزایش طول آماری بوده و عملکرد آن بهتر از برآورد نقطه‌ای می‌باشد و میزان این برتری به برقراری فرضیات اساسی آنالیز منطقه‌ای سیلاب یعنی همگنی ناحیه و استقلال داده‌ها بستگی دارد. وابسته بودن مکانی داده‌های سیلاب، خطای برآورد چندک‌های سیلاب را افزایش می‌دهد ولی میزان این افزایش کمتر از اریبی است که بخاطر ناهمگنی ناحیه به چندک‌های سیلاب اعمال می‌شود (۱۰، ۱۱، ۱۴ و ۲۰). روسبجرق و مادسن (۱۷) با بررسی عدم قطعیت مدل‌های تحلیل فراوانی سیلاب دریافتند بعضی از مدل ها شدیداً به همگنی ناحیه و مستقل بودن ایستگاه‌ها وابسته هستند. از طرفی مدل‌هایی که محدودیت‌های همگنی و مستقل بودن ایستگاه‌ها در آنها کمتر است دارای عدم قطعیت بالایی هستند. مطالعات گسترده‌ای بر روی اهمیت همگنی و استقلال ایستگاه‌ها صورت گرفته که از جمله می‌توان به هاسکینگ و والیس (۸)، لتینمایر و پوتر (۹)، گوس و همکاران (۷) و ویلتشایر (۱۹) اشاره کرد.

در آنالیز فراوانی سیلاب ابتدا ایستگاه‌ها براساس یک معیار مناسب، طبقه‌بندی و یا ناحیه‌بندی شده و سپس همگنی داخلی و ناهمگنی دوجانبه ناحیه‌ها بررسی می‌شود. معیار طبقه‌بندی و یا ناحیه‌بندی ایستگاه‌ها می‌تواند تشابه خصوصیات فیزیوگرافی حوضه و یا مشخصات جریان باشد (۶ و ۱۹). روش خوشه‌ای یکی از روش‌های متداول ناحیه‌بندی بوده که مشکل اساسی آن عدم وجود معیار مناسب جهت تعیین پارامترهای مؤثر بر ناحیه‌بندی و وزن آنها می‌باشد که به همین سبب ناحیه‌بندی صورت گرفته می‌تواند یگانه نباشد و با تغییر پارامترها و یا وزن آنها به ناحیه‌بندی دیگر دست یافت (۴). برن و کاندرلیک (۳) جهت طبقه‌بندی ایستگاه‌ها از اختلاف فراوانی نسبی ماهانه سیلاب‌ها استفاده کردند. نتایج آنها نشان داد، متغیر بودن فراوانی ماهانه ایستگاه‌ها نشان دهندهٔ تفاوت مکانیسم‌های سیلاب می‌باشد. مالکی نژاد و همکارانش (۱۲) برای ناحیه‌بندی حوضه دریاچه ارومیه جهت آنالیز فراوانی منطقه‌ای سیلاب از روش خوشه‌بندی استفاده کرد و برای تشخیص خصوصیات فیزیوگرافی

۱- استادیار گروه مهندسی عمران، دانشگاه آزاد اسلامی، واحد زنجان
Email: noori_mohammad2002@yahoo.com
2- Canonical Correlation

موثر بر ناحیه‌بندی از روش آنالیز مولفه‌های اصلی استفاده کرد. مدرس و سرحدی (۱۳) از روش‌های آماری چند متغیره برای ناحیه‌بندی ایران جهت آنالیز فراوانی بارش حداکثر سالانه استفاده کرد و ایران را به ۸ ناحیه همگن تقسیم بندی کرد. آداموفسکی (۱) از تاریخ وقوع و شکل فراوانی منحنی سیلاب‌ها جهت طبقه‌بندی ایستگاه‌ها استفاده کرد. یکی از روش‌های ناحیه‌بندی روش گرافیکی است که کاملترین آنها منحنی‌های اندرو می‌باشد که در آن چندین متغیر حوضه در یک نمودار دو بعدی ترسیم می‌گردد (۲). در این منحنی شبیه بودن و نزدیک بودن منحنی اندرو ایستگاه‌ها نشان دهنده تشابه بیشتر ایستگاه‌ها می باشد. از روش‌های نوین ناحیه‌بندی می‌توان به روش ناحیه‌بندی چند مقیاسه پارامترها[1] اشاره کرده که عملکرد آن در مقایسه روش‌های متداول پیشین بهتر می‌باشد. بطور نمونه سامانیقو و همکاران (۱۸) یک روش نوین ناحیه‌بندی به نام روش نمایه جهت برآورد سیلاب در ایستگاه فاقد آمار ارائه نمودند. در همه روش‌های ذکر شده ناحیه‌بندی تنها بر اساس خصوصیات فیزیوگرافی صورت می‌گیرد که نمی‌تواند بهینه باشند. زیرا نه تنها برای شناسایی خصوصیات فیزیوگرافی موثر بر ناحیه‌بندی معیار دقیقی وجود ندارد بلکه در انتخاب و محاسبه وزن آنها نیز روش کاملاً صحیحی وجود ندارد (۳).

یکی از مهمترین کاربردهای آنالیز فراوانی منطقه‌ای سیلاب، برآورد سیلاب طرح در ایستگاه فاقد آمار می‌باشد که دقت آن وابسته به میزان همگنی ایستگاه‌های مورد استفاده می‌باشد. اگر همانند روش‌های خوشه‌ای تنها از متغیرهای حوضه برای تشکیل ناحیه همگن استفاده شود، دقت برآورد سیلاب طرح در ایستگاه فاقد آمار کاهش خواهد یافت. زیرا متغیرهای حوضه به تنهایی قادر به توصیف تشابه هیدرولوژیکی ایستگاه‌ها که فرض لازم آنالیز منطقه‌ای سیلاب است، نمی باشد (۳). کاوادیس و همکاران (۴) و اوادیا و همکاران (۱۶) جهت افزایش تاثیر متغیرهای مربوط به جریان در ناحیه‌بندی ایستگاه‌ها، از یک روش آماری چند متغیره به نام همبستگی متعارف استفاده کردند. در این روش علاوه بر خصوصیات فیزیوگرافی از مشخصه‌های جریان نیز برای ناحیه‌بندی استفاده می‌شود. همبستگی متعارف کاربرد وسیعی در هیدرولوژی دارد بطور نمونه هودچا و همکاران (۹) از روش همبستگی متعارف برای برآورد پارامترهای مدل‌های بارش و رواناب در ایستگاه‌های فاقد آمار و بررسی توزیع مکانی آنها استفاده کردند.

در این مقاله هدف استفاده از روش همبستگی متعارف جهت تعیین ناحیه مربوط به ایستگاه فاقد آمار می‌باشد. برای این منظور ابتدا از رگرسیون گام به گام برای تعیین مشخصه‌های جریان و خصوصیات فیزیوگرافی موثر بر ناحیه‌بندی استفاده شده (۱۵) و سپس با همبستگی متعارف که بین مشخصه‌های جریان و خصوصیات فیزیوگرافی موثر تعریف می‌شود، ناحیه مربوط به ایستگاه فاقد آمار تعیین می‌گردد. برای مقایسه دقت روش همبستگی متعارف، از روش خوشه‌بندی (به نام ماتریس فاصله) نیز برای تعیین ناحیه‌بندی استفاده گردید. روش معرفی شده بصورت مطالعه موردی برای ناحیه‌بندی حوضه آبریز دریاچه ارومیه بکار رفته است.

مواد و روش‌ها

تعیین متغیرهای حوضه موثر بر ناحیه بندی: در این مقاله انتخاب متغیرهای موثر حوضه و وزن مربوط به آنها با واقعیت تطابق بیشتری دارد. در این مقاله متغیرهای حوضه که موثر بر ناحیه‌بندی هستند، از طریق رگرسیون‌گیری گام به گام تعیین گردید. برای این منظور، یک رابطه توانی بین متغیرهای حوضه و چندک سیلاب بصورت زیر تشکیل می‌گردد:

$$Q_T = aX_1^{b_1} X_2^{b_2} \ldots X_M^{b_M} \quad (۱)$$

که در آن X_i: متغیرهای حوضه، b_i: ضریب رگرسیونی برای متغیر i حوضه، Q_T: دبی سیلاب با دورهٔ بازگشت T ، M: تعداد متغیرهای حوضه و a: ضریب ثابت می‌باشد. رابطه رگرسیونی ۱ برای دوره بازگشت‌های ۲، ۵، ۱۰ و ۲۵ انجام شده و چندک‌های مورد استفاده همان داده‌های ثبت شده هستند که دوره بازگشت آنها با رابطه ویبول بدست می‌آید. از آنجا که که طول دوره آماری بیش از ۲۵ سال است بنابراین می‌توان چندک‌های زیر ۲۵ سال را مستقیماً از داده‌های ثبت شده استخراج کرد.

بعد از برآورد ضرایب رابطه رگرسیونی ۱، آماره t^* برای همه خصوصیات فیزیوگرافی از رابطه زیر محاسبه می‌گردد (۱۵):

$$t_k^* = \frac{b_k}{S(b_k)} \quad (۲)$$

که در آن b_k و $S(b_k)$ به ترتیب ضریب رگرسیونی و انحراف معیار مربوط به متغیر kام حوضه (X_k) می‌باشد. آماره t_k^* دارای توزیع t استیودنت با درجه آزادی N-M-1 می باشد که N تعداد ایستگاه‌ها می‌باشد. از بین متغیرهای حوضه متغیری که کمترین آماره t^* را دارد و از مقدار بحرانی t که از جدول t استودنت در سطح معنی‌داری مورد نظر (که در اینجا ۵ درصد می‌باشد) استخراج شده کمتر باشد، حذف شده و دوباره مراحل فوق برای بقیه متغیرها تکرار می‌شود تا هیچ متغیر حوضه دارای آماره t^* کمتر از t بحرانی نباشد. این رگرسیون‌گیری گام به گام برای دورهٔ بازگشت‌های ۲، ۵، ۱۰ و ۲۵ سال انجام می‌شود و متغیرهای مؤثر حوضه در برآورد هر دبی مشخص شده و متغیرهایی که در اکثر دوره بازگشت‌ها موثر شناخته شوند به عنوان متغیرهای موثر بر ناحیه بندی تلقی می‌گردد.

1- Multiscale Parameter Regionalization

برای تعیین اهمیت نسبی متغیرهای حوضه انتخاب شده، از مقدار استاندارد شده ضرایب رگرسیونی رابطه ۱ استفاده می‌گردد. از آنجا که هر کدام از متغیرهای حوضه بکار رفته واحدهای مختلفی دارند، برای وزن دهی به متغیرهای حوضه باید ضرایب رگرسیون بدست آمده را با رابطهٔ زیر استاندارد کرد (۱۵):

$$W_i = b_i \frac{S_{X_i}}{S_{Q_T}} \qquad (۳)$$

که در آن W_i: ضریب استاندارد متغیر X_i در رابطه ۱ بوده که اهمیت و یا وزن این متغیر می‌باشد, b_i: ضریب توانی رگرسیون برای متغیر X_i, S_{X_i}: انحراف معیار متغیر X_i, S_{Q_T}: انحراف معیار دبی سیلاب با دورهٔ بازگشت T می‌باشد. تعیین اهمیت نسبی متغیرهای حوضه برای دبی‌های مختلف انجام می شود و با متوسط‌گیری از ضرایب استاندارد هر متغیر که برای دبی‌های مختلف محاسبه شده مقدار وزن متغیر مورد نظر مشخص می‌شود.

ناحیه‌بندی به روش ماتریس فاصله : یکی از متداول‌ترین معیار ناحیه‌بندی ایستگاه‌ها ماتریس فاصله می‌باشد (۳). با استفاده از این روش فاصله متغیرهای حوضه هر ایستگاه نسبت به ایستگاه فاقد آمار با رابطه زیر سنجیده می‌شود:

$$D_{jk} = \left[\sum_{i=1}^{M} W_i (X_{j,i} - X_{k,i})^2 \right] \qquad (۴)$$

که در آن D_{jk} فاصله ایستگاه j از ایستگاه فاقد آمار k , M تعداد متغیرهای حوضه بکار برده شده, W_i وزن مربوط به پارامتر و یا متغیر i, و $X_{j,i}$ مقدار استاندارد متغیر i برای ایستگاه j می‌باشد. برای تشکیل ناحیه مربوط به ایستگاه فاقد آمار، فاصله ایستگاه‌ها تا ایستگاه فاقد آمار محاسبه می‌گردد و از آنجا که اولویت افزوده شدن یک ایستگاه به ناحیه همگن بر اساس حداکثر تشابه و یا حداقل فاصله می‌باشد، ایستگاه‌ها بر حسب فاصله تا ایستگاه فاقد آمار به صورت صعودی مرتب شده و به ترتیب به ناحیه همگن مربوط به ایستگاه فاقد آمار اضافه می‌گردد و همزمان آزمون همگنی انجام می‌شود. اگر ایستگاهی باعث ناهمگن شدن ناحیه گردد آن ایستگاه از ناحیه ایستگاه فاقد آمار خارج شده و سپس روند فوق برای بقیه ایستگاه‌ها ادامه می‌یابد. در این تحقیق برای آزمون همگنی از روش متداول گشتاورهای خطی استفاده می‌گردد (۸).

ناحیه‌بندی به روش همبستگی متعارف: در این تحقیق هدف نشان دادن قابلیت بالای روش همبستگی متعارف جهت تشکیل ناحیه مربوط با ایستگاه فاقد آمار می‌باشد. با استفاده از روش همبستگی متعارف دو متغیر جدید که یکی ترکیب خطی متغیرهای حوضه (یا خصوصیات فیزیوگرافی) بوده و دیگر ترکیب خطی از متغیرهای جریان می‌باشد. این دو متغیر دارای بیشترین همبستگی می‌باشند. متغیر جدیدی که ترکیب خطی از خصوصیات فیزیوگرافی می‌باشد، برای تمامی ایستگاه‌ها از جمله ایستگاه فاقد آمار قابل محاسبه است. اما متغیر جدید دیگری که ترکیب خطی از متغیرهای جریان می‌باشد برای ایستگاه فاقد آمار از روی رابطه رگرسیونی بین دو این متغیر جدید قابل محاسبه است. حال هرچقدر مقدار متغیر جدید جریان در ایستگاه فاقد آمار، نزدیک به متغیر جریان در ایستگاه‌های مجاور باشد آنگاه تشابه بین ایستگاه‌ها بیشتر خواهد بود. برای استخراج این متغیرهای جدید، داده‌های M متغیر حوضه و داده‌های P متغیر جریان که در N ایستگاه اندازه‌گیری شده‌اند به ترتیب در قالب ماتریس‌های Z با بعد $N \times M$ و Y با بعد $N \times P$ جاسازی شده و S_{ZZ}، S_{ZY} و S_{YY} که به ترتیب ماتریس کوواریانس متغیرهای حوضه، ماتریس کوواریانس متغیرهای جریان و ماتریس کوواریانس بین متغیرهای جریان و متغیرهای حوضه می‌باشد، محاسبه می‌گردد. همچنین فرض شود $\rho_1^2 \geq \rho_2^2 \geq \ldots \geq \rho_p^2$ مقادیر ویژه مرتب شده برای ماتریس $S_{ZZ}^{-\frac{1}{2}} S_{ZY} S_{YY}^{-1} S_{YZ} S_{ZZ}^{-\frac{1}{2}}$ با بردارهای ویژه $e_1, e_2, \ldots, e_p$ بوده و $S_{YY}^{-\frac{1}{2}} S_{YZ} S_{ZZ}^{-1} S_{ZY} S_{YY}^{-\frac{1}{2}}$ بردارهای ویژه ماتریس $f_1, f_2, \ldots, f_p$ می‌باشد. زوج متغیر متعارف kام برای متغیرهای حوضه (V_k) و متغیرهای جریان (U_k) عبارتند از (۴):

$$U_k = e_k' S_{ii}^{-\frac{1}{2}} Z \quad , \quad V_k = f_k' S_{jj}^{-\frac{1}{2}} Y \qquad (۵)$$

$$Cov(U_k, V_l) = Cov(U_k, U_l) = Cov(V_k, V_l) = 0 \quad k \neq l \qquad (۶)$$

$$r_{U_k, V_k} = \rho_k \qquad (۷)$$

در رابطه فوق U_k و V_k بترتیب k امین ترکیب خطی از متغیرهای جریان و متغیرهای حوضه می‌باشند. همبستگی بین این دو متغیر (U_k, V_k)، k امین همبستگی متعارف نامیده می‌شود که در روابط فوق با r_{U_k, V_k} نمایش داده شده است. از آنجا که ρ_1^2 بزرگترین مقدار ویژه می‌باشد، اولین زوج متغیر متعارف (U_1, V_1) دارای بیشترین همبستگی بوده و دومین زوج متغیر متعارف (U_2, V_2) از نظر همبستگی در مرتبه دوم قرار دارند و در مقابل بقیه زوج متغیرهای متعارف در مقایسه با دو زوج متغیر متعارف اول دارای همبستگی ناچیزی می‌باشند. در این تحقیق از دو زوج متغیر متعارف اول جهت تعیین ناحیه مربوط به ایستگاه همگن استفاده می‌گردد. از آنجا که در ایستگاه فاقد آمار مشخصه‌های جریان وجود ندارد مقدار متغیرهای متعارف جریان U_1 و U_2 قابل محاسبه نبوده و باید برآورد گردد. می‌توان مقادیر $\hat{U}_1$ را در محل ایستگاه فاقد آمار از رابطه رگرسیون خطی بین V_1 و U_1 برآورد کرد. به همین ترتیب

مقدار متغیر $\hat{U}_2$ از رابطه رگرسیونی بین V_2 و U_2 برآورد می‌گردد. در مرحله بعد فاصله نقاط دو متغیر متعارف اول جریان (U_2 ، U_1) در ایستگاه‌های مختلف نسبت به نقطه ($\hat{U}_2$، $\hat{U}_1$) در ایستگاه فاقد آمار محاسبه شده و به ترتیب صعودی مرتب می‌شود که به ترتیب به ناحیه همگن اضافه شده و آزمون همگنی با روش متداول گشتاورهای خطی انجام می‌گردد (۸). اگر ایستگاهی باعث ناهمگن شدن ناحیه گردد آن ایستگاه از ناحیه ایستگاه فاقد آمار خارج شده و سپس روند فوق برای بقیه ایستگاه‌ها ادامه می‌یابد.

برای بررسی بیشتر دقت ناحیه‌بندی دو روش بکارگرفته شده، در هر تکرار یکی از ایستگاه‌های هیدرومتری به عنوان ایستگاه فاقد آمار در نظر گرفته و با دو روش همبستگی متعارف و ماتریس فاصله ناحیه همگن شناسایی و سپس چندک‌های سیلاب در دوره بازگشت‌های ۲، ۵، ۱۰ و ۲۵ سال با آنالیز منطقه‌ای سیلاب (روش نمایه سیلاب) محاسبه می‌شود (۵). بعد از آن که چندک‌های سیلاب در دوره بازگشت‌های مختلف برای تمامی ایستگاه‌ها برآورد گردید، می‌توان متوسط خطای نسبی برآورد چندک‌های سیلاب برای هر کدام از روش‌های ناحیه‌بندی از رابطه زیر بدست آورد:

$$error(Q_T)=\frac{1}{N}\sum_{i=1}^{N}\frac{\left|\hat{Q}_{T,i}-Q_{T,i}\right|}{Q_{T,i}} \qquad T=2,5,10,25 \qquad (۸)$$

در رابطه فوق T دوره بازگشت، N تعداد ایستگاه‌های هیدرومتری، $\hat{Q}_{T,i}$ مقدار سیلاب برآورد شده برای دوره بازگشت T در ایستگاه i (که به عنوان ایستگاه فاقد آمار فرض شده بود) با آنالیز منطقه‌ای سیلاب ، $Q_{T,i}$ مقدار سیلاب ثبت شده در ایستگاه i با دوره بازگشت T و $error(Q_T)$ مقدار متوسط خطای نسبی برآورد Q_T می‌باشد. این خطا برای دوره بازگشت‌های ۲، ۵، ۱۰ و ۲۵ سال برآورد گردید. در هر روشی که مقدار خطای نسبی برآورد چندک‌های سیلاب کم باشد به عنوان روش ناحیه بندی بهتر شناسایی می‌گردد.

منطقه مورد مطالعه: روش ذکر شده بصورت مطالعهٔ موردی برای ناحیه بندی حوزهٔ دریاچهٔ ارومیه استفاده شده است.حوزهٔ آبریز دریاچه ارومیه، منطقه ای کوهستانی است که در تقاطع دو رشته کوه البرز و زاگرس درشمال غربی کشور، بین طول‌های ′۰۷ ، °۴۴ تا ′۵۳ ، °۴۷ شرقی وعرض‌های ′۴۰ ، °۳۵ تا ′۳۰ ، °۳۸ شمالی قرار گرفته است. جریان سطحی حوزهٔ آبریز این دریاچه از طریق ۲۰ رودخانه اصلی بزرگ و کوچک به دریاچه شور ارومیه واقع در بخش شمالی حوضه می‌ریزد. موقعیت حوضه آبریز دریاچه ارومیه و شبکه آبراهه آن در شکل۱ نشان داده شده است. در این مطالعهٔ ۳۴ ایستگاه هیدرومتری این حوزه که غیر تنظیمی بوده انتخاب گردیده که طول آماری آنها از ۲۶ تا ۴۰ سال متغیر است.

نتایج و بحث

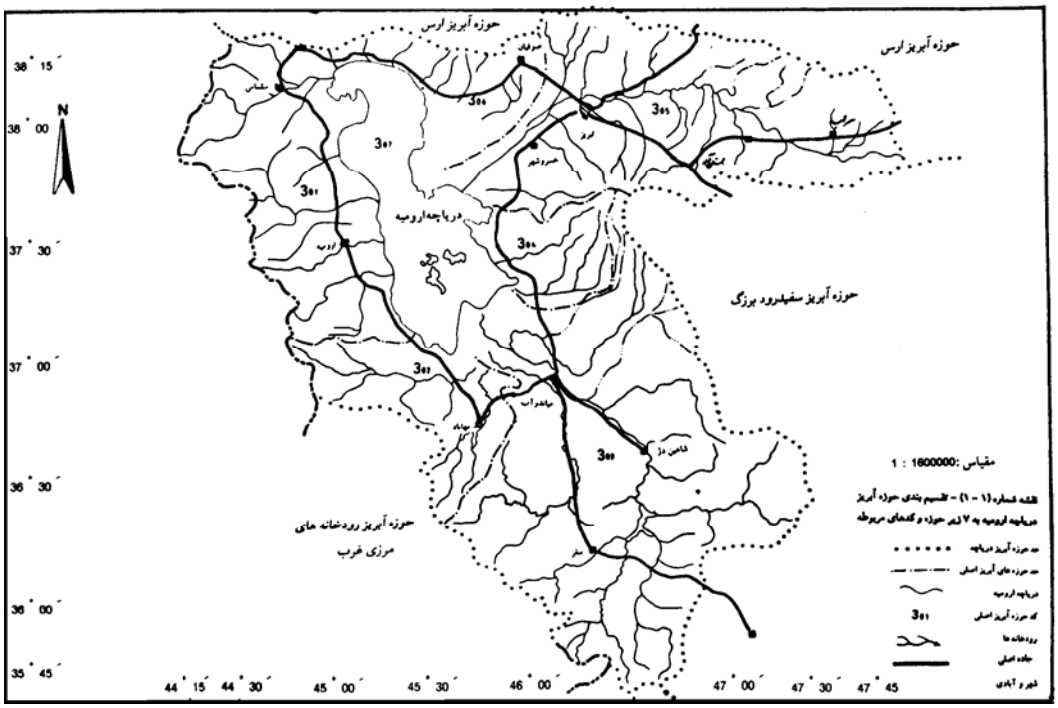
متغیرهای فیزیوگرافی مورد استفاده در این تحقیق شامل مساحت حوزه (km^2)، بارش متوسط سالانه (mm) در محل ایستگاه، محیط حوزه (km)، ارتفاع متوسط حوزه (m) از سطح دریا، زمان تمرکز (hr)، شیب آبراهه اصلی (درصد)، شیب متوسط حوضه (درصد)، عرض و طول جغرافیایی ایستگاه (درجه)، ضریب گراولیوس و ضریب شکل می‌باشد که از بین آنها با کمک روش رگرسیون‌گیری گام به گام متغیرهای مؤثر بر ناحیه‌بندی مشخص گردید که در جدول ۱ نمایش داده شده است. از آنجا که بزرگی متغیرهای حوضه کاملاً متفاوت هستند، جهت کاهش بزرگی و دامنه تغییرات، از لگاریتم متغیرها استفاده می شود (۴). با استفاده از ضرایب رگرسیونی استاندارد شده که در بخش مواد و روش ها آورده شده است، اهمیت نسبی هر متغیر حوضه مشخص گردید که در جدول ۱ رتبه آنها ارائه شده است. بطوری که متغیری که بیشترین ضریب استاندارد (یا اهمیت) را دارد رتبهٔ ۱ و به متغیری که کمترین ضریب استاندارد را دارد رتبهٔ ۵ دارند.

در جدول ۱ متغیرهای جریان نیز ارئه شده عبارتند از حداکثر سیلاب سالانه با دوره بازگشت ۲ سال (Q_2) و نسبت حداکثر سیلاب سالانه با دوره بازگشت ۱۰۰ سال به حداکثر سیلاب سالانه با دوره بازگشت ۲ سال ($Q_{100/2}$). متغیر Q_2 متوسط جریان بوده و متغیر $Q_{100/2}$ که نسبت دبی ۱۰۰ ساله به دبی ۲ ساله است، معیاری از دامنه تغییرات یا انحراف چارکی سیلاب می‌باشد که مشخصه‌ای از منحنی توزیع احتمالاتی جریان‌های حداکثر می‌باشد. به عبارت دیگر این دو متغیر یکی پارامتر موقعیت و دیگری پارامتر شکل بوده و اگر این دو متغیر در دو ایستگاه نزدیک هم باشند آنگاه تشابه هیدرولوژیکی بیشتر خواهد شد.

در نرم افزار Matlab7 با استفاده از دستور ([U,V] = canoncorr(Z,Y زوج متغیرهای متعارف که یکی ترکیب خطی متغیرهای جریان (V) و دیگری تابع خطی از متغیرهای حوضه می‌باشد (U)، استخراج گردید که در شکل ۲ اولین زوج متغیر متعارف (U_1 ,V_1) همراه با معادله رگرسیونی بین آنها نمایش داده شده است. این زوج متغیر متعارف نسبت به دیگر متغیرهای متعارف دارای بیشترین ضریب همبستگی (R^2 =۰/۷۱) می‌باشد. در شکل ۳ دومین زوج متغیر متعارف (U_2 ,V_2) برای ۳۴ ایستگاه هیدرومتری نمایش داده شده که ضریب همبستگی آن R^2 =۰/۵۴ بوده و در سطح معنی داری ۵ درصد معنی‌دار می‌باشد. دو متغیر متعارف حوضه یعنی V_1 و V_2 در شکل ۴ برای ایستگاه‌های مورد مطالعه نمایش داده شده که در آن موقعیت ایستگاه شماره ۱۴ با Z مشخص گردیده که به عنوان ایستگاه فاقد آمار در نظر گرفته شده است.

جدول ۱- متغیرهای فیزیوگرافی موثر بر ناحیه بندی به ترتیب اهمیت همراه با متغیرهای جریان در جدول زیر آمده است که در این تحقیق لگاریتم آنها مورد استفاده قرار می‌گیرد.

	نماد	متغیر
متغیر های حوضه	X_1	محیط حوزه (km)
	X_2	ارتفاع متوسط حوزه (m) از سطح دریا
	X_3	بارش متوسط سالانه (mm) در محل ایستگاه
	X_4	شیب آبراهه اصلی (٪)
	X_5	مساحت حوزه (km^2)
متغیر های جریان	Q_2	حداکثر سیلاب سالانه با دوره بازگشت ۲ سال (m^3/s)
	$Q_{100/2}$	نسبت حداکثر سیلاب سالانه با دوره بازگشت ۱۰۰ سال به حداکثر سیلاب سالانه با دوره بازگشت ۲ سال

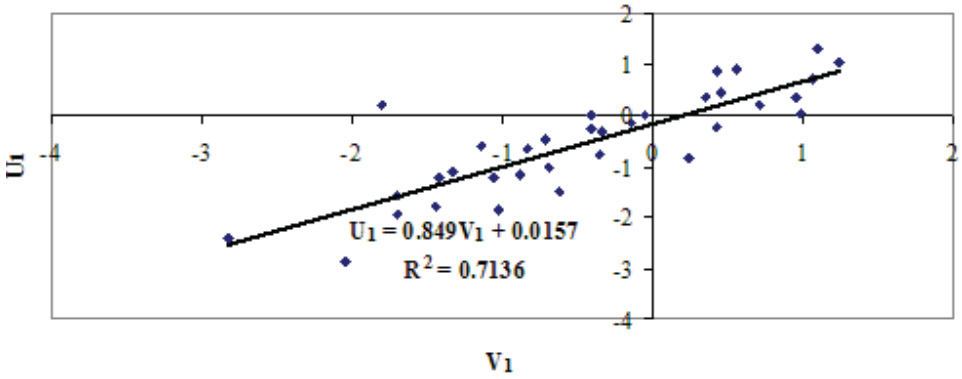


شکل ۱- موقعیت حوزه آبریز دریاچه ارومیه و شبکه آبراهه

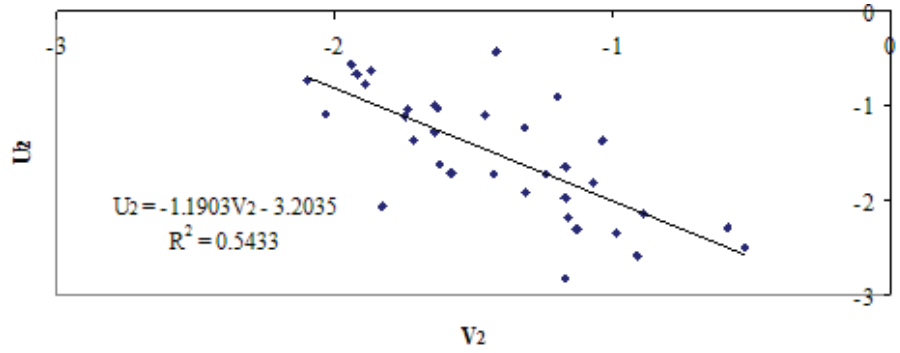
با استفاده از روابط رگرسیونی ارائه شده در شکل‌های ۲ و ۳، دو متغیر متعارف جریان U_1 و U_2 برای ایستگاه فاقد آمار برآورد گردیده که در شکل ۵ موقعیت آن با $\check{Z}$ مشخص شده است. برای تعیین ناحیه مربوط به ایستگاه فاقد آمار، با توجه به شکل ۵ بر حسب نزدیکی فاصله ایستگاه‌ها تا ایستگاه فاقد آمار، ناحیه همگن شناسایی می‌گردد. این ناحیه همگن که با آزمون همگنی گشتاورهای خطی مورد تایید قرار گرفته دارای ۱۴ ایستگاه می باشد که به ترتیب اولویت عبارتند ایستگاه‌های ۱۱، ۲۷، ۱۶، ۲۲، ۵، ۳۰، ۴، ۲۵، ۲۱، ۱۸، ۲۳، ۱۷، ۱۵ و ۸. اگر ناحیه‌بندی تنها بر اساس متغیرهای حوضه با روش ماتریس فاصله صورت می‌گرفت آنگاه ناحیه همگن مربوط به ایستگاه فاقد آمار دارای ۱۰ ایستگاه به شماره‌های ۲۱، ۱۷، ۲۵، ۸، ۲۸، ۲۰، ۱۶، ۲۷، ۱۸ و ۱۱ خواهد بود. با مقایسه ایستگاه‌های ناحیه‌بندی شده به دو روش، می‌توان گفت تعداد ایستگاه‌های ناحیه همگن در روش همبستگی متعارف که علاوه بر متغیرهای حوضه از جریان استفاده می‌کند، زیاد بوده که این باعث افزایش طول آماری در آنالیز منطقه‌ای سیلاب شده و سبب بالا رفتن اعتمادپذیری می‌گردد. در ناحیه‌بندی اضافه کردن یک ایستگاه ممکن است باعث سلب عضویت یک ایستگاه دیگر گردد به عبارت دیگر ترتیب افزودن ایستگاه به ناحیه

همگن بسیار مهم است. از آنجا که در ناحیه بندی هدف دسته‌بندی بر اساس شباهت هیدرولوژیکی می‌باشد، با توجه به نتایج به نظر می‌رسد همبستگی متعارف در تعیین شباهت هیدرولوژیکی ایستگاه‌ها عملکرد بهتری در مقایسه با روش ماتریس فاصله (که تنها بر اساس متغیرهای حوضه می‌باشد) دارد. از آنجا که ابزار شناسایی شباهت ایستگاه‌ها در این دو روش ناحیه‌بندی متفاوت است باعث شده ایستگاه‌های دو ناحیه‌بندی کاملاً یکسان نباشد و در چند ایستگاه دو ناحیه‌بندی متفاوت باشد.

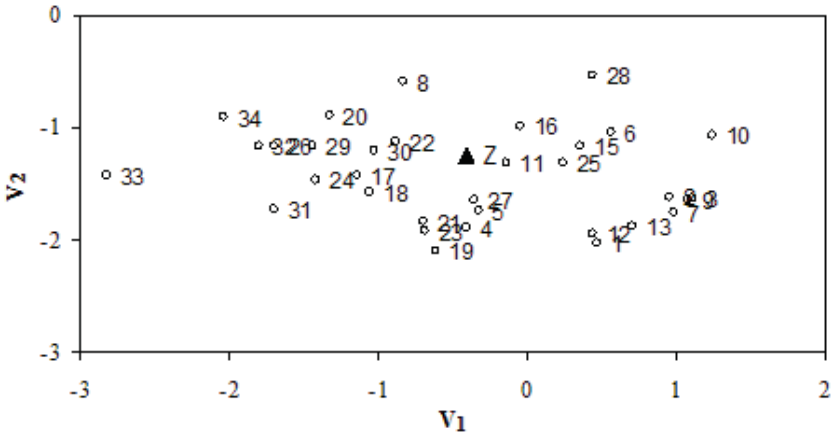
برای بررسی بیشتر دقت ناحیه‌بندی دو روش بکارگرفته شده، متوسط درصد خطای نسبی برآورد چندک‌های سیلاب برای هر کدام از روش‌های ناحیه‌بندی برای دوره بازگشت‌های ۲، ۵، ۱۰ و ۲۵ سال برآورد گردید که نتایج آن در شکل ۶ ارائه شده است.



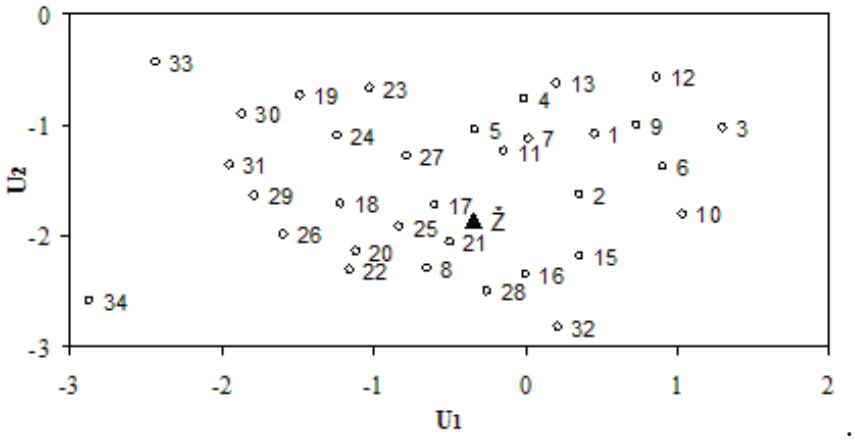
شکل ۲- مقدار اولین زوج متغیر متعارف برای منطقه مورد مطالعه همراه با رابطه رگرسیونی آنها



شکل ۳- مقدار دومین زوج متغیر متعارف برای منطقه مورد مطالعه همراه با رابطه رگرسیونی آنها

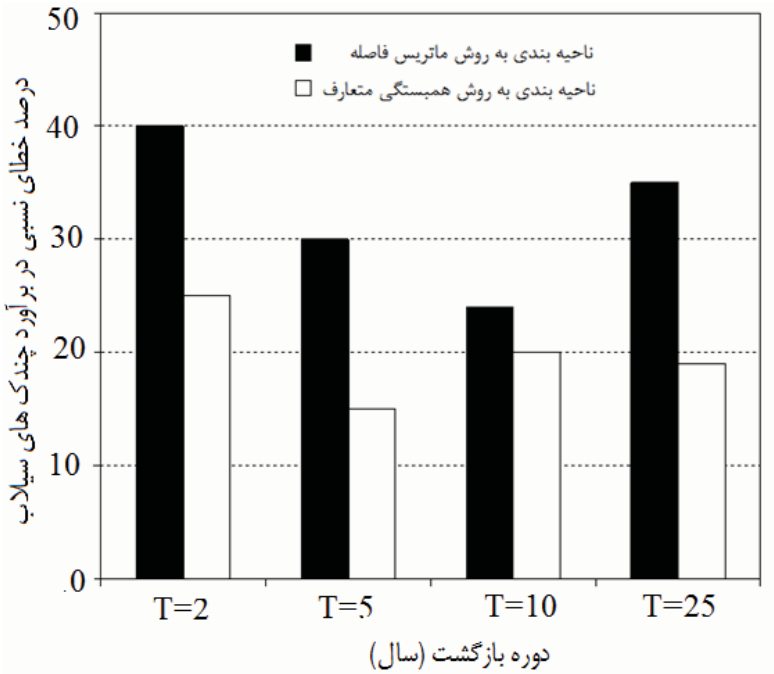


شکل ٤- مقدار دو متغیر اول متعارف حوضه که حاصل ترکیب خطی از متغیرهای حوضه می‌باشند



شکل ۵- مقدار دو متغیر اول متعارف جریان که حاصل ترکیب خطی از متغیرهای جریان می‌باشند

همانطور که در این شکل نشان داده شده است مقدار درصد خطای برآورد سیلاب در حالتی که ناحیه‌بندی با روش همبستگی متعارف انجام شده کمتر از حالتی است که ناحیه‌بندی با روش ماتریس فاصله انجام شده است. در این تحقیق از آزمون t استیودنت در سطح معنی داری ۵ درصد برای مقایسه میانگین خطا در دو ناحیه بندی استفاده گردید (۱۵) که نتایج نشان داد که تفاوت میانگین دو جامعه در دوربازگشت های مختلف معنی دار می باشد. بنابراین می‌توان گفت روش همبستگی متعارف قابلیت مناسبی برای تشکیل ناحیه مربوط به ایستگاه فاقد آمار را دارد که علت آن استفاده توام متغیرهای جریان و حوضه می‌باشد. کاوادیس و همکاران (۴) با مقایسه عملکرد روش همبستگی متعارف و آنالیز خوشه در ناحیه‌بندی ایستگاه‌های هیدرومتری دریافت که عملکرد روش همبستگی متعارف در دوره بازگشت بالا (بیشتر ۱۰) بهتر بوده و روش خوشه‌بندی برای دوره بازگشت پایین موثرتر است. اما در این تحقیق عملکرد همبستگی متعارف در تمای دوره بازگشت‌ها بهتر بوده که می‌تواند ناشی از آن باشد که در این تحقیق، شناسایی خصوصیات فیزیوگرافی موثر بر ناحیه بندی و تعیین اهمیت نسبی آنها بیشتر با واقعیت منطبق می‌باشد در حالی که در تحقیق کاوادیس و همکاران (۴) انتخاب خصوصیات فیزیوگرافی سلیقه‌ای بوده و اهمیت نسبی همه خصوصیات فیزیوگرافی یکسان فرض شده است.



شکل ٦- متوسط درصد خطای نسبی برآورد چندک‌های سیلاب در ایستگاه فاقد آمار در دو روش ناحیه‌بندی ماتریس فاصله و همبستگی متعارف

نتیجه‌گیری

آنالیز منطقه‌ای سیلاب یک روش مناسب برای برآورد سیلاب طرح در ایستگاه فاقد آمار می‌باشد که میزان دقت آن وابسته به همگنی ایستگاه‌های مورد استفاده در آنالیز منطقه‌ای سیلاب می‌باشد. در روش‌های متداول ناحیه مثل روش خوشه‌بندی، تنها از خصوصیات فیزیوگرافی (متغیرهای حوضه) استفاده می‌شود. در حالی که متغیرهای حوضه به تنهایی قابلیت کافی برای تشخیص تشابه ایستگاه‌ها را ندارد و از طرف دیگر انتخاب خصوصیات فیزیوگرافی موثر بر ناحیه بندی و وزن آنها معیار مشخصی ندارد. در این تحقیق از روش رگرسیون گام به گام برای شناسایی متغیرهای حوضه و تعیین اهمیت آنها استفاده شده که انطباق خوبی با واقعیت دارد. همچنین در این تحقیق برای تشکیل ناحیه مربوط به ایستگاه فاقد آمار علاوه بر روش ماتریس فاصله از روش آماری چندگانه به نام همبستگی متعارف استفاده شد که در آن علاوه بر متغیرهای حوضه از متغیرهای جریان نیز استفاده می‌شود. در روش همبستگی متعارف با استفاده از رابطه بین زوج متغرهای متعارف، ناحیه ایستگاه فاقد آمار شناسایی گردید. نتایج بدست آمده از حوضه دریاچه ارومیه نشان می‌دهد تعداد ایستگاه‌های مربوط به ناحیه بدست آمده از روش همبستگی متعارف نسبت به ماتریس فاصله زیاد می‌باشد که این باعث افزایش طول دوری آماری و سبب افزایش اعتمادپذیری می‌گردد. بررسی‌های بعمل آمده نشان داد که خطای برآورد سیلاب طرح در ایستگاه فاقد آمار در ناحیه‌بندی به روش همبستگی متعارف کمتر از ناحیه‌بندی به روش ماتریس فاصله می‌باشد. بنابراین می‌توان گفت روش همبستگی متعارف قابلیت کافی برای ناحیه‌بندی ایستگاه‌های هیدرومتری جهت آنالیز فراوانی منطقه سیلاب را دارد.

منابع

1- Adamowski K. 2000. Regional analysis of annual maximum and partial duration flood data by nonparametric and L-Moment methods. J. Hydrol., 229:219-231.
2- Andrews D. 1972. Plots of high-dimensional data, Biometrics 28: 125-136.
3- Burn D.B. 1988. Cluster analysis as applied to regional flood frequency. Water Resources Research Planning and Management, 115:567-582.
4- Cavadias G.S., Ouarda T.B.M., Bobee B., and Girard C. 2001. A canonical correlation approach to the determination of homogeneous region for regional flood estimation of ungauged basins. Scientific Report no. R-578, INRS-Eau, Ste-Foy, Canada.
5- Chandra R.A., and Hamed K.H. 2000. Flood frequency analysis. CRC Press.
6- Fill H.D., and Stedinger J.R. 1995. Homogeneity tests based upon Gumbel distribution and a critical appraisal of Dalrymple's test. J. Hydrol., 17(1):111-117.
7- Guse B., Castellarin A., Thieken A.H., and Merz B. 2009. Effects of intersite dependence of nested catchment structures on probabilistic regional envelope curves. Hydrolcal Earth System Sciences, 13:1699-1712.
8- Hosking J.R.M., and Wallis J.R. 1993. Some statistics useful in regional frequency analysis. Water Resour Res 29(2):271–281.
9- Hundecha Y., Ouarda T.B.M.J., and Bardossy A. 2008. Regional estimation of parameters of a rainfall-runoff model at ungauged watersheds using the spatial structures of the parameters within a canonical physiographic-climatic space. Water Resour. Res., 44, W01427, doi: 10.1029/2006WR005439.
10- Lettenmaier D.P., and Potter K.W. 1985. Testing flood frequency estimation methods using a regional flood generation model. Water Resources Research, 21(12):1903-1914.
11- Lettenmaier D.P., Wallis J.R., and Wood E.F. 1987. Effect of regional heterogeneity on flood frequency estimation. Water Resources Research, 23(2):313-323.
12- Malekinezhad H., Nachtnebel H.P., and Klik A. 2011. Regionalization approach for extreme flood analysis using L-moments. J.Agr. Sci. Tech., 13:1183-1196.
13- Modarres R., and Sarhadi A. 2011. Statistically-based regionalization of rainfall climates of Iran. Global and Planetary Change., 75:67-75.
14- Nathan R.J., and Mcmabon T.A. 1990. Identification of homogeneous regions for the purposes of regionalization. J. Hydrol., 121:217-238.
15- Neter J., and Wasserman W. 1975. Applied linear statistical models. Richard D.Irwin, INC.
16- Ouarda T.B.M.J., Hache M., and Bruneau P. 2000. Regional flood peak and volume estimation in northern Canadian basin. J. Cold Regiona Engng ASCE 14(4), 176-191.
17- Rosbjerg D., and Madsen H. 1995. Uncertainty measures of regional flood frequency estimators. J. Hydrol., 167:209-224.
18- Samaniego L., Kumar R., and Attinger S. 2010. Multiscale parameter regionalization of a grid-based hydrologic model at the mesoscale, Water Resour. Res., 46, W05523, doi: 10.1029/2008WR007327.

19- Waltshire S.E. 1986. Identification of homogeneous regions for flood frequency Analysis. J. Hydrol., 84:287-302.
20- Zrinji Z., and Burn H. 1994. Flood frequency analysis for ungauged sites using a region of influence approach. J. Hydrol., 153:1-21.

Journal of Water and Soil
Vol. 26, No.4, Sept.-Oct. 2012, p. 943-952

نشریه آب و خاک (علوم و صنایع کشاورزی)
جلد ۲۶، شماره ۴، مهر – آبان ۱۳۹۱، ص. ۹۵۲-۹۴۳

Identifying the Regional Homogeneity of Ungauged Station for Regional Flood Frequency Analysis Using Canonical Correlation Method

M.H. Noori Gheidari[1]
Received: 4-9-2011
Accepted: 12-2-2012

Abstract

Regional flood analysis is a conventional method for estimation of flood design in ungauged site which its accuracy relates the how much homogeneity is among the stations that used in regional flood frequency. Cluster method is commonly used to regionalization of stations which its accuracy is dependent on identifying the effective basin variables and their weight that it commonly can't accessible. In this research, Canonical Correlation between basin and flow variables was used to identify the region of ungauged site. In this method the pair canonical variables and its relations are able to determine the similarity of stations. This method was applied to Ouroumieh Lake basin and results shown that created homogeneous region of ungauged sites using Canonical Correlation method has more stations and the error of design flood estimation in ungauged sites is low in compare the cluster method.

Keywords: Regional estimation, Homogeneous region, Canonical variable, Ouroumieh Lake

1- Assistant Professor, Department of Civil Engineering, Zanjan Branch, Islamic Azad University, Zanjan, Iran
Email: noori_mohammad2002@yahoo.com