



تعیین بهترین روش همگن‌بندی سلسله مراتبی برای تحلیل منطقه‌ای شاخص جریان پایه در حوزه‌های آبخیز استان کرمان

رجیم کاظمی^۱ و فرود شریفی^۲

^۱- استادیار، پژوهشکده حفاظت خاک و آبخیزداری، سازمان تحقیقات، آموزش و ترویج کشاورزی، تهران، (تویستنده مسؤول: ra_hkazemi@yahoo.com)

^۲- دانشیار، پژوهشکده حفاظت خاک و آبخیزداری، سازمان تحقیقات، آموزش و ترویج کشاورزی، تهران

تاریخ دریافت: ۹۵/۵/۱۰ تاریخ پذیرش: ۹۷/۱۰/۲۳

صفحه ۹۵ تا ۱۰۵

چکیده

فقدان پوشش کامل داده‌های هیدرولوژی، هیدرولوژیست‌ها را مجبور به استفاده از روش همگن‌بندی در تحلیل منطقه‌ای می‌کند. در این پژوهش، به منظور انتخاب بهترین روش خوشه‌بندی سلسله مراتبی در گروه‌بندی موضع‌ها، با استفاده از داده‌های جریان روزانه رودخانه، جریان پایه و شاخص آن از روش فیلتر رقومی برگشتی دو پارامتره در ۴۳ ایستگاه آب‌سنگی مربوط به حوضه‌های در درجه چهار استان کرمان، استخراج شد. پارامترهای هیدرولوژیکی و اقلیمی، محاسبه و عوامل فیزیوگرافی در محیط رقومی استخراج شد. با استفاده از روش آزمون تحلیل عاملی از بین ۱۴ پارامتر مؤثر در شاخص جریان پایه، شش عامل ثابت افت منحنی هیدرولوگراف، بارش متوسط سالانه، تبخر و تعقیق واقعی، نفوذپذیری، توبوگرافی (شبیب و ارتفاع) و شماره منحنی، به عنوان عوامل مستقل انتخاب شدند. همگن‌بندی حوضه‌ها با استفاده از عوامل مستقل انتخابی به روش‌های مختلف سلسله مراتبی شامل: نزدیک‌ترین همسایه، دورترین همسایه، فاصله از مرکز ثقل و روش وارد، انجام شد. سپس روابط منطقه‌ای به روش رگرسیون خطی، در سطوح معنی‌داری کمتر از یک درصد برای گروه‌های همگن تمام روش‌ها به دست آمد. ارزیابی صحت و کارایی مدل‌های برآورده، به روش‌های آزمون استقلال خطاهای، نرمال بودن توزیع خطاهای و همخطی انجام شد. سپس دقت مدل‌های استخراج شده با استفاده از ایستگاه‌های شاهد محاسبه و با یکدیگر مقایسه شد. نتایج نشان داد که دقت مدل‌های استخراج شده در حوضه‌های همگن انتخاب شده به روش "فاصله از مرکز ثقل" به دلیل حداقل خطای برآورد نسبی به میزان ۱۳٪/۲۷٪ برای منطقه یک و ۸۹٪/۳۶٪ برای منطقه همگن دو، حداقل خطای برآورد استاندارد و ضریب همبستگی بالا، به عنوان بهترین روش همگن‌بندی سلسله مراتبی برای برآورد منطقه‌ای شاخص جریان پایه در منطقه پژوهش تشخیص داده شد.

واژه‌های کلیدی: تحلیل عاملی، تحلیل منطقه‌ای، خوشه‌بندی سلسله مراتبی، شاخص جریان پایه، کارایی مدل

(۲۳) برای تعیین عرصه‌های مناسب پخش سیالاب، روش‌های همگن‌بندی سلسله مراتبی و مدل‌های آنالیز خوشه‌بندی سلسله مراتبی را مورد تجزیه و تحلیل قرار دادند. ایشان نتیجه گرفتند که روش آنالیز خوشه‌بندی خاکستری نسبت به روش تحلیل سلسله مراتبی از دقت بیشتری برخوردار است. جلسن و اسمیتر (۲۰) در آفریقای جنوبی برای تجزیه و تحلیل منطقه‌ای سیالاب از روش همگن‌بندی سلسله مراتبی استفاده کردند. موسوی و سپاسخواه (۲۵) برای برآورد دبی بیشینه روزانه استان فارس از روش آزمون همگنی لانگبین برای همگن‌بندی استفاده کردند. روحانی (۳۰) و عباسی‌زاده (۱) از چهار روش نزدیک‌ترین همسایه، دورترین همسایه، وارد و متوسط گروه استفاده کردند و روش وارد را برای منطقه مورد پژوهش، بهینه تشخیص داده‌اند. خسروی و همکاران (۱۸) در تحقیقی در حوزه‌های آبخیز غرب کشور، روش‌های سلسله مراتبی و آزمون همگنی لانگبین را مورد بررسی و مقایسه قرار داده‌اند و بهترین روش همگن‌بندی برای تحلیل منطقه‌ای سیالاب را روش سلسله مراتبی تشخیص داده‌اند. تحقیق مشابهی توسط آهنی و همکاران (۲) در حوضه سفیدرود به انجام رسیده و این روش را برای تحلیل منطقه‌ای سیالاب توصیه نموده‌اند. بیابانکی و اسلامیان (۶) و عطایی و شیران (۴) از روش تحلیل خوشه‌ای برای تعیین مناطق همگن بر اساس ویژگی‌های هیدرولوژیک استفاده کردند. رویکرد منطقه‌ای در برآورد در تحلیل منطقه‌ای منحنی تداوم جریان

مقدمه

مدیریت پایدار منابع آب، نیازمند داده‌های مناسب است که ما را در توسعه استراتژی مدیریت کمی و کیفی منابع آب یاری کند. کمود داده‌های مناسب در حوضه مربوط به مطالعات منابع آب، منجر به عدم قطعیت در بخش‌های مربوط به مدیریت این کمود می‌شود. توجه به این مشکل منجر به نام‌گذاری دهه ۲۰۱۲-۲۰۰۳ به پیش‌بینی در حوزه‌های آبخیز فاقد آمار، توسط انجمن بین‌المللی هیدرولوژی شد (۱۲). با توجه به محدودیت داده‌های اندازه‌گیری شده، تعیین داده‌ها به حوضه‌های فاقد آمار نیازمند پژوهش و بررسی جامع در خصوص ارتباط ویژگی‌های ثابت و پایدار حوضه‌ها با داده‌های جریان در حوضه‌های همگن برای تعیین روابط به حوضه‌های فاقد آمار است (۲۲). تاکنون پژوهش‌های مختلف برای تجزیه و تحلیل منطقه‌ای مؤلفه‌های هیدرولوژیکی به انجام رسیده است که در هریک از این پژوهش‌ها، روش‌های مختلف همگن‌بندی بر اساس توصیه محققین به کاربرده شده است. روش مورد استفاده در این پژوهش، خوشه‌بندی سلسله مراتبی داده‌هاست؛ که یکی از پرکاربردترین ابزارهای تشخیص الگو به منظور طبقه‌بندی نمونه، با توجه به ویژگی‌های برداشت شده به گروه‌ها یا خوشه‌های مشخص است. از روش‌های خوشه‌بندی در علوم مختلف و از جمله موارد مرتبط با مدیریت منابع آب، استفاده شده است. از جمله این تحقیقات می‌توان به موارد ذیل اشاره کرد. ملکیان و همکاران

به وسیله محققین مختلفی از جمله رحیمی و همکاران (۲۹)، قبریور (۱۵)، تیموری و همکاران (۳۳) و ولدرافائل (۳۴) توصیه شده است.

با جمع‌بندی پژوهش‌های انجام‌شده، مشخص است که استفاده از روش همگن‌بندی در پژوهش‌های گذشته عمدتاً بر اساس توصیه سایر محققین بوده است. همچنان، اکثر تحقیقاتی که در خصوص برسی و ارزیابی روش‌ها انجام شده است، روش‌ها را با شاخص‌های آماری مرتبط با نوع و ماهیت ریاضی روش همگن‌بندی مورد ارزیابی قرار داده‌اند و به موضوع برسی دقت روابط منطقه‌ای حاصل از روش‌های مختلف نپرداخته‌اند. لذا با برسی‌های به عمل آمده تاکنون مطالعات جامعی در خصوص برسی روش‌های همگن‌بندی سلسه مراتبی به منظور تحلیل منطقه‌ای شاخص جریان پایه در سطح منطقه‌ای و استانی در کشور ایران مشاهده نشد. با توجه به اینکه روش‌های مختلف منجر به دسته‌بندی‌های متفاوتی از حوضه‌های همگن خواهد شد، لذا پژوهش در خصوص شناخت مناسب‌ترین روش همگن‌بندی، موجب دستیابی به اطلاعاتی برای استفاده در تحلیل شرایط هیدرولوژیکی و مدیریت بهینه منابع آب حوضه در منطقه پژوهش خواهد شد. هدف از این پژوهش، برسی و انتخاب مناسب‌ترین روش همگن‌بندی سلسه‌مرتبی در حوزه‌های آبخیز استان کرمان است.

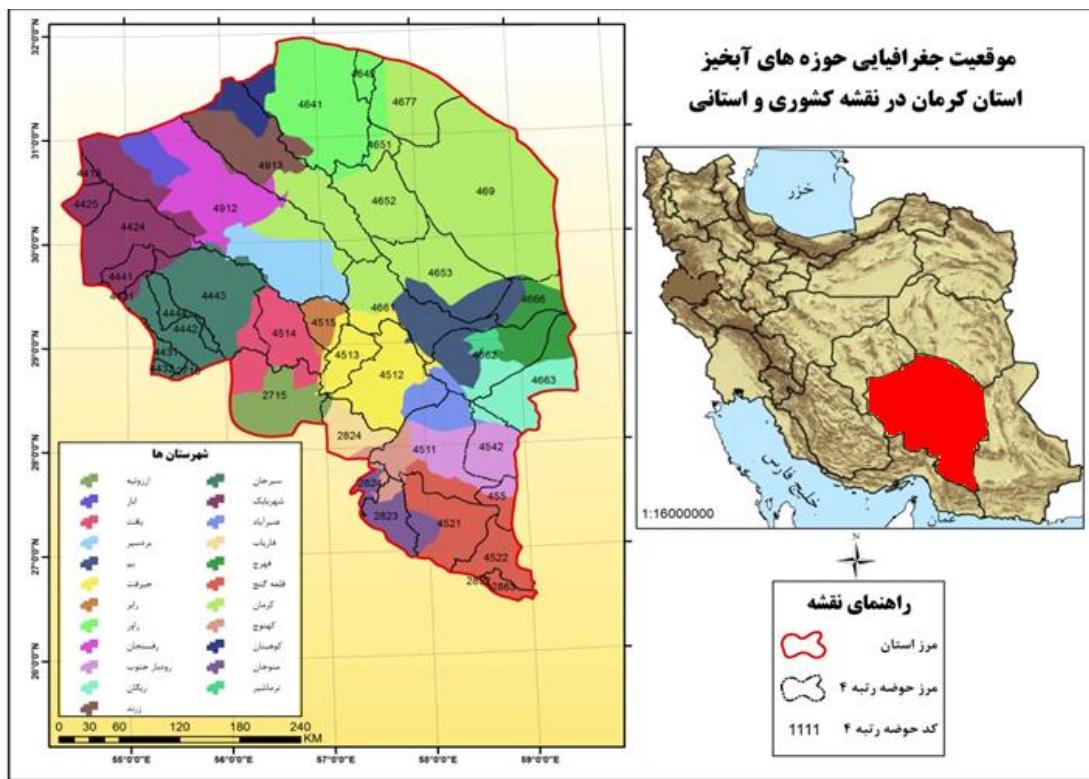
مواد و روش‌ها

منطقه مورد پژوهش

استان کرمان با مساحتی در حدود ۱۸۱۷۳۷ کیلومترمربع بین ۲۵ درجه و ۵۵ دقیقه تا ۳۲ درجه عرض شمالی و ۵۳ درجه و ۲۶ دقیقه تا ۵۹ درجه و ۲۹ دقیقه طول شرقی قرارگرفته است. حد شمالی استان کرمان به استان‌های خراسان، یزد و حد جنوبی به استان هرمزگان، حد شرقی به استان سیستان و بلوچستان و حد غربی آن به استان فارس محدود است. استان کرمان با در برگرفتن حدود ۱۱ درصد از مساحت کشور بخش وسیعی از پیکر جنوب شرقی فلات مرکزی ایران را می‌پوشاند. اقلیم استان کرمان از نظر تقسیمه‌بندی‌های اقلیمی، جزء اقلیم نیمه‌خشک بیانی گرم محسوب گردیده است. امتداد سلسه جبال زاگرس و کوههای مرکزی، استان پهناور کرمان را به دو بخش متایز خشک کویری و معتدل کوهپایه‌ای تقسیک نموده که در تلاقی با یکدیگر از لحاظ اقلیمی سه منطقه کویری و حاشیه کویری، گرمسیری و سردسیری و کوهستانی را شکل داده‌اند. میزان متوسط بارندگی سالیانه استان کرمان ۱۴۵ میلی‌متر و میزان نزولات جوی واردہ بر پهنه استان نزدیک به ۲۷ میلیارد مترمکعب برآورد گردیده است. مشخصه‌های پایه حوضه‌های مورد پژوهش در جدول (۱) و نقشه منطقه مورد پژوهش در شکل (۱) ارائه شده است.

توسط خسروپیگی و وفاخواه (۱۹) در حوزه آبخیز دریاچه نمک مورد بررسی قرار گرفت. ایشان نتیجه گرفتند که این رویکرد منجر به بالاتر رفتن دقت برآورد پارامترهای هیدرولوژیکی می‌شود.

یکی از مهم‌ترین مشخصه‌های هیدرولوژیکی که در پژوهش‌های مختلف بر آن تأکید شده، شاخص جریان پایه (BFI) است. شاخص جریان پایه، یک نسبت بدون بعد است که از نسبت دبی پایه به کل رواناب برای هر مقطع زمانی یا کل دوره آماری به دست می‌آید. مشخصه‌های هیدرولوژیکی حوزه‌های آبخیز مانند جریان پایه، یا سایر مؤلفه‌های جریان توسط پارامترهای هیدرولوژیکی و فیزیکی مختلف و به روش همگن‌بندی، قابل برآورد است (۲۶). مشخصه‌های هیدرولوژیکی مانند جریان پایه، دما، تبخیر و تعرق است و عوامل فیزیکی شامل پارامترهای فیزیوگرافی و ثابت حوضه بوده و ممکن است پارامترهای زمین‌شناسی و خاک‌شناسی از دیدگاه هیدرولوژی نیز مورد استفاده قرار گیرد (۲۷، ۲۸). اطلاع از میزان جریان پایه برای مقاصد متعددی مانند مدیریت منابع آب، حفاظت اکوسیستم‌های آبی، تولید انرژی، حمل و نقل و پیش‌بینی جریان‌های کمینه اهمیت دارد (۱۰، ۱۴). چنین اطلاعاتی در خصوص حوضه‌های فاقد آمار در دسترس نیست، بنابراین فرآیند آنالیز منطقه‌ای برای انتقال روابط پارامترهای حوضه‌های دارای آمار به حوضه‌های فاقد آمار مورد نیاز است (۷). پژوهش‌های متعددی مبتنی بر آنالیز منطقه‌ای برای ایجاد مدل‌های برآورد شاخص جریان پایه با استفاده از مشخصه‌های، هیدرولوژیکی، کلیماتولوژی و فیزیوگرافی حوزه‌های آبخیز انجام شده است (۳، ۵، ۹، ۲۱، ۲۷، ۲۴، ۲۲، ۲۱، ۳۱). روش‌های متعددی برای جداسازی جریان پایه از جریان رودخانه توسعه بافته است که مبتنی بر ردیاب‌های شیمیایی قابل طبقبندی است (۱۱). جداسازی جریان پایه به روش گرافیکی اغلب زمان بر و غیردقیق است و نتایج بدست آمده مبتنی بر ردیاب‌ها دقیق و مختلف یکسان نیست. روش‌های مبتنی بر ردیاب‌ها صرف وقت و هزینه قابل اعتماد است ولی زمان بر و مستلزم صرف وقت و هزینه زیادی است. روش‌های مبتنی بر الگوریتم‌های رقومی و فیلتر، علاوه بر سهولت و قابلیت تشخیص مناسب در تعیین دبی پایه، حساسیت بالایی نسبت به پارامترها دارد. همچنین به دلیل قابلیت خودکار کردن، مشکلات ناشی از عدم همخوانی نتایج تا حدودی برطرف شده است. به دلیل پیچیدگی و نامشخص بودن میزان واقعی مشارکت دبی پایه در رواناب و همچنین هزینه بر بودن روش‌های مبتنی بر ردیاب‌ها، محققین بسیاری نسبت به ارزیابی و مقایسه روش‌های مختلف تفکیک هیدرولوگراف جریان اقدام نموده‌اند و الگوریتم‌های رقومی مختلفی را پیشنهاد داده‌اند. در این پژوهش از روش فیلتر رقومی برگشتی دوپارامتره برای استخراج شاخص جریان پایه استفاده شده است. این روش



شکل ۱- منطقه مورد پژوهش
Figure 1. Study area

جدول ۱- مشخصه‌های حوزه‌های آبخیز مورد پژوهش

Table 1. Watershed Characteristics of study area

مساحت (km²)	نسبت (%)	أعلى طول أينما (km)	أقصى ارتفاع متوسط (m)	كثافة سكانية (نسمة/km²)	نوع (جغرافية)	مساحت (km²)	نسبت (%)	أعلى طول أينما (km)	أقصى ارتفاع متوسط (m)	كثافة سكانية (نسمة/km²)	رتبة
٣٠/٩٢	٣٥٣/٣٧	٣٥٣	١٥٨٠/١٨	٤٥١٢	٢٣	١١٣٨/٠٠	٢١/٨٣	٤٨٠	١٤٨٨/٠١	٤٥٥	١
٣٠/٩٥	٢٩٢/٢٨	٢٩٢	١٤٤٥/٠٨	٤٥١٣	٢٤	١٨٧٩/٢٤	٢٧/٥٩	٨١٧	١٨١٣/٠٠	٤٦٩	٢
٤٠/٥٨	٤٥٧/٤٦	٤٥٧	١٩١٣/٣٧	٤٥١٤	٢٥	١٦٩١/١٦	١٨/٤	٦٦٦	٢٢٤٤/٤٥	٢٧١٥	٣
٤٢/٤١	١٧٧/٨٣	١٧٨	١٩٦٣/٥٨	٤٥١٥	٢٦	١٦٧٣/٨	١٦/٢٦	٦٨١	٢٢٠٦/٤٥	٢٧١٦	٤
٣٣/٨٨	١٠٧/٣٣	١٠٧	١٨٤٣/٠٠	٤٥٢١	٢٧	١٦٠١/١٢	٢٥/٩٣	١٢٠	١٢٨٧/٣٢	٢٨١٥	٥
٤٢/٨٢	١٩٩/٨٢	٢٠٠	٢٠٨٦/٧٤	٤٥٢٢	٢٨	٣٥٣/٣	٢٠/٠٣	٣٦	٣٨٢/٧٣	٢٨٢٣	٦
٢٠/٧٣	٤٥٧/٩٢	٤٥٨	١٤٧٦/٤٠	٤٥٤١	٢٩	١٠٦٨/٠٤	٢٨/٤١	٥١٠	١٤٤٤/٠٠	٢٨٢٤	٧
٣٩/٦٠	٦٢١/٦٣	٦٢١	١٤٤٥/٨١	٤٥٤٢	٣٠	٧٥٢/٧٥	٩/٩٠	٤٠٨	٣٨٤/٤٥	٢٨٣١	٨
٣١/٠٣	٢٦٤/٥٦	٤٨٢	١٩١٢/٩٩	٤٦٢٣	٣١	٧٣٥/٨	١٩/١٤	٣٩٤	١٠٠/١٦	٢٨٤٣	٩
٢٥/٩١	٤٨١/٩٢	٢٧٧	١٨٧٨/١٢	٤٦٤١	٣٢	١١٤٩/٧٢	٣٥/٤	٥٧٢	١٥٦٥/٧٣	٢٨٥٢	١٠
٣٥/٥٨	٣٣٧/٠١	١٧٤٣	١٩٤٤/٤٥	٤٦٤٢	٣٣	١٢٣٩/٠٨	٣٨/٢	٥٣٤	١٨٤٢/٤٢	٢٨٤٣	١١
١٩/٥٠	١٧٦٢/٧٠	٥٩١	١٨٥٢/٨٢	٤٦٥١	٣٤	١٨٩٩/١٦	١٥/٤١	٨٥٩	١٠٥١/٩٢	٤٤١٣	١٢
١٨/٤١	٦٩١/١٨	٦٨٧	٢٤١٠/١٠	٤٦٥٢	٣٥	٦٨٨/٦٨	١٥/٣٩	٤٠٠	٩٤٧/٤٠	٤٤٢٤	١٣
١٠/٣٩	٦٨٧/١٠	٢٠١	١٧٤٢/٩٢	٤٦٥٣	٣٦	١٦٤٢/٣٤	٩/٨٤	١٧٦	٢١٦٤/٣٢	٤٤٢٥	١٤
٩/٧٨	٢٠١/٣٨	٣١	١٨١٥/٤٨	٤٦٦١	٣٧	١٩٢٤/٣٢	٢٢/٢٥	٧٢٢	٧٦٤/٩١	٤٤٢٦	١٥
١٢/٤١	٣١٠/٣٧	٢٥٨	٢٣٧/١٨	٤٦٦٢	٣٨	٢١٣٣/٤٨	١٤/٦	٩٨١	٥٨١/٥٧	٤٤٢١	١٦
١١/٧٧	٢٥٨/١٥	١٥٥٧	٢٠٣٥/٢١	٤٦٦٣	٣٩	٣٠٠٥/٣٢	١٣/٦	١٠٣٠	١١٩٩/٧٩	٤٤٣٢	١٧
٢٠/١٦	١٥٥٧/٠٨	٢٣٨	٢٢٨٤/٣٨	٤٦٦٤	٤٠	٢٢٠٩/٢٤	١٢/٩٥	١٥٢٢	٤٨٩/٨٨	٤٤٤١	١٨
٣٦/٣٣	٢٣٨/٠٩	٣٥٣	٢٧٠٩/٠٤	٤٦٧٧	٤١	١٤١٨/٥٢	٥/٦٣	١٤١٣	١٣٩٣/٠٠	٤٤٤٢	١٩
٢١/٧٩	٧١١/٧٩	٢٩٢	٢١٦٦/٥٥	٤٩١٢	٤٢	٣٧٠٢٠	٣٠/٣٤	٢٠٢	١٦٧٠/٥٠	٤٤٤٣	٢٠
١٣/٧٤	٢٩٥/٨٩	٤٥٧	١٧٠٣/٤٣	٤٩١٣	٤٣	٦٧٣/٤٩	٢٩/٧	٣٢٣	١٨٤٨/٣٩	٤٤٤٤	٢١
						٤٤٩/٣٤	٢٨/٠٧	٣٠٨	١٦٩٠/٥٧	٤٥١١	٢٢

شاخه نزولی در برابر زمان به خط راست تبدیل می‌شود. روش‌های مختلفی به منظور تفکیک هیدروگراف جریان توسعه یافته و در پژوهش‌های گذشته مورد استفاده قرار گرفته است. در این پژوهش، تفکیک هیدروگراف جریان و استخراج جریان پایه و محاسبه شاخص مربوطه با استفاده از الگوریتم فیلتر رقومی برگشتی دو پارامتره و پس از بررسی و آماده‌سازی داده‌های روزانه جریان و کالیبراسیون مدل و به دست آوردن مقادیر صحیح پارامترهای مدل، انجام شد.

فیلتر رقومی برگشتی دو پارامتره

روش فیلتر رقومی برگشتی برای تحلیل و پردازش سیگنال‌های با فرکانس بالا از فرکانس پایین معروفی شده است، فیلتر رقومی برگشتی دو پارامتره با الگوریتم ذیل، اولین بار توسط بوتون^(۸) و با رابطه^(۳) معروفی شد.

$$q_{b(i)} = \frac{K}{1+C} q_{b(i-1)} + \frac{C}{1+C} q_{(i)} \quad (3)$$

$$q_{b(i)} \leq q_{(i)}$$

K : پارامتر فیلتر، قابل تعیین به وسیله ثابت افت منحنی است. این پارامتر برابر است با شبیه شاخه نزولی منحنی هیدروگراف، که در این روش مقدار 0.975 به دست آمده است. پارامتر C : پارامتری است که با تغییر آن، شکل جداسازی تغییر می‌کند و قابلیت کالیبراسیون مدل را تأمین می‌کند. با استفاده از قابلیت نمایش گرافیکی نرم‌افزار و به کار بردن مقادیر مختلف پارامتر و کالیبراسیون آن، مقدار پارامتر بهینه مورد نیاز $C=0.5$ انتخاب شده است. $q_{b(i-1)}$: جریان پایه فیلتر شده برای زمان قبل از i : $q_{(i)}$ جریان اصلی رودخانه برای زمان i : $q_{b(i)}$: جریان پایه فیلتر شده، برای زمان i .

همگن‌بندی حوضه‌ها

یکی از روش‌هایی که می‌توان به وسیله آن آمار جریان را به سایر حوضه‌ها تعمیم داد، تعیین همگن‌بندی حوزه‌های آبخیز می‌باشد. در این پژوهش به منظور همگن‌بندی مناطق موردنظر، از روش‌های مختلف تحلیل خوشهای سلسله مراتبی شامل: نزدیک‌ترین همسایه، دورترین همسایه، فاصله از مرکز نقل و روش وارد، استفاده شد. تحلیل خوشهای به عنوان نوعی تحلیل چندمتغیره، ابزاری آماری است که در آن داده‌ها و مشاهدات بر اساس متغیرهای تعیین شده به وسیله محقق طبقه‌بندی می‌شوند. در این روش، حوضه‌ها از نظر خصوصیات فیزیکی دو به دو با هم مقایسه می‌شوند و باید از مقایسه‌های کمی استفاده کرد تا بتوان بیشترین شیاهت یا کمترین فاصله بین داده‌ها و مشاهدات را اندازه‌گیری نمود.

خوشه‌بندی سلسله مراتبی

خوشه‌بندی سلسله مراتبی، روشی است که در آن به خوشه‌های نهایی بر اساس میزان عمومیت آن‌ها ساختاری سلسله مراتبی، نسبت داده می‌شود.

روش نزدیک‌ترین همسایه

این روش یکی از قدیمی‌ترین و ساده‌ترین روش‌های خوشه‌بندی است و جزو روش‌های خوشه‌بندی سلسله مراتبی پایین به بالا محسوب می‌شود. در این روش برای محاسبه شیاهت بین دو خوشة A و B از رابطه^(۴) استفاده می‌شود.

روش پژوهش

ابتدا با بررسی داده‌های دبی روزانه ایستگاه‌های آب‌سنگی، تعداد ۴۳ ایستگاه مربوط به حوضه‌های رتبه چهار با آمار مناسب و دوره مشترک آماری سال‌های آبی ۱۳۵۵-۱۳۸۰ انتخاب گردید. با استفاده از نقشه توپوگرافی با مقیاس ۱:۵۰۰۰۰ و تعیین موقعیت ایستگاه‌ها، محدوده مورد پژوهش مشخص و پارامترهای اولیه حوضه با استفاده از سیستم اطلاعات جغرافیایی، استخراج گردید. سپس شاخص جریان پایه با استفاده از داده‌های روزانه جریان و پس از کنترل سری زمانی و آماده‌سازی اطلاعات، با استفاده از نرم‌افزار Hydro Office 2012 به روش فیلتر رقومی برگشتی دو پارامتره^(۱)، استخراج شد. تجزیه و تحلیل برای شناخت عوامل مستقل موثر بر شاخص جریان پایه انجام شد. همگن‌بندی حوضه‌ها با استفاده از عوامل مستقل انتخابی به روش‌های مختلف سلسله مراتبی شامل: نزدیک‌ترین همسایه، دورترین همسایه^(۲)، فاصله از مرکز نقل^(۳) و روش وارد^(۴) انجام شد.

بررسی و آماده‌سازی اطلاعات

در بسیاری از مطالعات هیدرولوژیکی، آگاهی از میزان جریان پایه و شاخص آن از موارد ضروری است، این در حالی است که در اکثر موارد با فقدان و یا کمبود ایستگاه‌های آب‌سنگی مواجه بوده و نیز عوامل بسیاری وجود دارد که باعث از دست رفتن آمار برخی ایستگاه‌ها می‌شوند، ازین‌ron در ابتدا اقدام به جمع‌آوری، اصلاح، و بازسازی آمار دبی در منطقه پژوهش گردید. به منظور شناسایی داده‌های پرت، از آزمون داده‌های بالا و پایین، انجام گردید که برای تعیین آستانه داده‌های پرت بالا از رابطه^(۱) و برای تعیین آستانه داده‌های پرت پایین از رابطه^(۲) استفاده شد.

$$Y_H = Y + K_n S_y \quad (1)$$

$$Y_L = Y + K_n S_y \quad (2)$$

Y_H : آستانه داده‌های پرت بالا به صورت لگاریتمی Y_L : آستانه داده‌های پرت پایین به صورت لگاریتمی K_n : ضریبی است که از جدول مربوطه به تعداد داده‌ها انتخاب می‌شود. S_y انحراف از معیار داده‌ها می‌باشد. در سال‌هایی که داده‌های دبی ثبت نگردیده بود، با استفاده از ماتریس همیستگی و رابطه همیستگی با ایستگاه‌های دیگر نواقص دی کلیه ایستگاه‌ها تا ۲۵ سال آماری مشترک تکمیل گردید.

جداسازی جریان آب‌پایه

جداسازی جریان پایه، طی فرآیند تفکیک هیدروگراف جریان و با شناسایی نقطه شروع و خاتمه رواناب مستقریم قابل انجام است. نقطه شروع نقطه‌ای است که جریان، روند صعودی پیدا می‌کند و نقطه خاتمه، زمانی است که لگاریتم

مساحت حوضه، ارتفاع متوسط حوضه، طول آبراهه اصلی، شبیب، درصد عدم تجاوز ۹۰ و ۷۵ درصد و متوسط CN حوضه است. محورهای عاملی با روش واریمکس^۳ که روش مناسب و استاندارد برای ساده کردن ساختار عامل‌ها و به حداقل رساندن رابطه همبستگی بین متغیرها می‌باشد، دوران یافته. سپس با استفاده از روش مدل رگرسیونی، ماتریس امتیازات عاملی استخراج شد. برای تعیین متغیرهایی که خروج آن‌ها از مدل تجزیه عاملی به روند بهتر شدن کیفیت تجزیه عاملی کمک می‌کند، از ماتریس همبستگی Anti-image و آماره MSA^۴ استفاده شد. عناصر روی قطر ماتریس همبستگی، مقادیر اندازه دقت نمونه‌گیری با نماد MSA نامیده می‌شود. این معیار که برای هر متغیر جداگانه تحلیل می‌شود، بیان می‌کند که آیا متغیر خاصی از جهت وارد شدن به مدل تجزیه عاملی صلاحیت دارد یا نه؟ مقادیر بزرگ MSA در جهت تأیید ورود متغیر به مدل حرکت می‌کنند. پس از انتخاب متغیرهای مهم، تجزیه عاملی بر اساس این متغیرها صورت گرفته و مهمنترین پارامترها به عنوان عوامل مستقل مشخص شد. سپس روابط بین عامل وابسته شاخص جریان پایه (BFI) با سایر عوامل مستقل، با استفاده از روش رگرسیونی (رابطه ۷) مورد بررسی قرار گرفت.

$$BFI = aMr + bHc + C \quad (7)$$

که در آن پارامترهای Mr: فاکتور مورفومتری، Hc: فاکتور هیدرواقلیمی و پارامترهای a و b و c ضرایب ثابت رابطه می‌باشند. مدل‌های استخراج شده پس از بررسی میزان دقت و نیز خطای برآورده آن‌ها بر اساس معیارهای ضربیت همبستگی، خطای برآورد استاندارد و خطای نسی، و تأمین اعتبار مدل‌ها به روش‌های آزمون استقلال خطاهای بررسی نرمال بودن توزیع خطاهای و بررسی همخطی مورد استفاده قرار گرفت. سرانجام به لحاظ آماری با توجه به سطوح معنی‌داری کمتر از یک درصد مناسب‌ترین مدل‌ها انتخاب شدند.

تأمین اعتبار رگرسیون‌ها

انجام هر رگرسیونی، مستلزم فرضیاتی از قبیل، مستقل بودن خطاهای از یکدیگر، نرمال بودن توزیع خطاهای و پایین بودن میزان همخطی بین متغیرهای مستقل است و تا زمانی که این مفروضات برقرار نباشد، اعتبار رگرسیون خدشه‌دار خواهد بود.

آزمون استقلال خطاهای

خطاهای عبارت است از تفاوت بین مقادیر واقعی و مقادیر پیش‌بینی شده. برای بررسی استقلال خطاهای از همدیگر از آزمون دوربین-واتسون استفاده شد. اگر همبستگی بین خطاهای را به P نشان دهیم در این صورت آماره دوربین-واتسون به کمک رابطه (۸) محاسبه می‌شود.

$$DW = 2(1 - P) \quad (8)$$

اگر $P=0$ باشد آنگاه مقدار آماره $DW=2$ خواهد بود که نشانه عدم همبستگی خطاهای می‌باشد. اگر $P=1$ باشد آنگاه آماره $DW=0$ خواهد بود که نشان می‌دهد خطاهای دارای خود همبستگی مثبت هستند. اگر $P=-1$ باشد آنگاه مقدار آماره $DW=4$ خواهد بود که نشانه همبستگی منفی بین خطاهای است. از آنجایی که فرض H_0 بر عدم همبستگی بین خطاهای

$$d_{AB} = \min d_{ij} \quad (4)$$

$i = A, j = B$ که A یک نمونه داده متعلق به خوشه A و B یک نمونه داده متعلق به خوشه B می‌باشد. در واقع در این روش شباهت بین دو خوشه، کمترین فاصله بین یک عضو از یکی با یک عضو از دیگری است.

روش دورترین همسایه

این روش نیز جزء روش‌های خوشه‌بندی سلسله مراتبی پایین به بالا محسوب می‌شود. در این روش برای محاسبه شباهت بین دو خوشه A و B از رابطه (۵) استفاده می‌شود.

$$d_{AB} = \max d_{ij} \quad (5)$$

$i = A, j = B$ که A یک نمونه داده متعلق به خوشه A و B یک نمونه داده متعلق به خوشه B است. در واقع در این روش شباهت بین دو خوشه بیشترین فاصله بین یک عضو از یکی با یک عضو از دیگری است.

روش فاصله از مرکز ثقل

این روش نیز جزء روش‌های خوشه‌بندی سلسله مراتبی پایین به بالا محسوب می‌شود. در این روش برای محاسبه شباهت بین دو خوشه A و B از رابطه (۶) استفاده می‌شود.

$$d_{AB} = d\left(\frac{\sum_{i=A} X_i}{N_A}, \frac{\sum_{i=B} X_j}{N_B}\right) \quad (6)$$

که A یک نمونه داده متعلق به خوشه A، X_i یک نمونه داده متعلق به خوشه B، N_A تعداد اعضاء خوشه A و N_B تعداد اعضاء خوشه B است. در واقع در این روش، شباهت بین دو خوشه فاصله بین بردار میانگین تمام اعضاء یکی با بردار میانگین تمام اعضاء دیگری است.

روش وارد

در این روش خوشه‌بندی برای کاهش تلفات ناشی داده‌های دورافتاده از معیاری جدید برای محاسبه عدم شباهت بین خوشه‌ها استفاده می‌شود. در این روش از مجموع مربعات تفاضل هر داده از یک خوشه با بردار میانگین آن خوشه، به عنوان معیاری برای سنجش یک خوشه استفاده می‌شود. الگوریتم مربوطه بدین صورت است که ابتدا هر داده به عنوان یک خوشه در نظر گرفته می‌شود. به ازاء تمام جفت خوشه‌های ممکن از مجموعه خوشه‌ها آن دو خوشه‌ای که مجموع مربعات تفاضل داده‌های خوشه حاصل از اجتماع آن‌ها با بردار میانگین خوشه حاصل کمینه باشد، انتخاب می‌شوند. دو خوشه انتخاب شده باهم ترکیب می‌شوند و تا زمانی که تعداد خوشه‌ها به تعداد مورد نظر نرسیده است، مراحل تکرار می‌شوند.

تجزیه و تحلیل برای شناخت عوامل مستقل موثر بر شاخص جریان پایه

روش تجزیه و تحلیل عاملی به روش (PCA)^۵ با استفاده از نرم‌افزار SPSS برای ۱۴ متغیر فیزیوگرافی، هیدرولوژیکی و اقلیمی در حوضه‌های منتخب به کار برده شد. متغیرهای مورداستفاده، شامل ویژگی‌های مختلف حوضه از جمله، شاخص جریان پایه، ثابت افت منحنی، میزان نفوذپذیری، ظرفیت ذخیره سطحی، بارش، تبخیر تعرق واقعی، دما،

در محدوده مورد نیاز هر روش و با توجه به بهترین پاسخ داده شده به تغییرات ضرایب و با استفاده از قابلیت نمایش گرافیکی نرمافزار Hydro Office و مقایسه چشمی انتخاب و نتایج برای استفاده در تحلیل عاملی مورد استفاده قرار گرفت. میانگین سالیانه شاخص جریان پایه در منطقه تحقیق برابر با $0/50$ ، حداکثر برابر یک و حداقل آن برابر $0/11$ به دست آمد. انحراف معیار این شاخص در کل منطقه تحقیق برابر $0/21$ بوده است. شاخص جریان پایه در بیش از 75 درصد از حوضه‌های مورد بررسی بیشتر $0/40$ بوده است، که انعکاسی از خشک بودن منطقه و بالا بودن سهم مشارکت آب‌های زیرزمینی در جریان رودخانه‌ها است.

آزمون تحلیل عاملی نشان داد که کلیه متغیرها در قالب شش عامل طبقه‌بندی می‌شوند که درمجموع $78/72$ درصد از واریانس داده‌ها را توضیح می‌دهند، (جدول ۲). پارامترهای ضریب نگهداشت حوضه و ثابت افت هیدروگراف به عنوان عامل اول، پارامترهای بارش متوسط سالانه، درصدهای عدم تجاوز Q_{90} در عامل دوم و پارامترهای تبخیر و تعرق واقعی و دمای سالانه در عامل سوم، پارامتر نفوذپذیری به عنوان عامل چهارم، پارامتر شیب و طول رودخانه اصلی به عنوان عامل پنجم و پارامترهای مساحت، ارتفاع متوسط حوضه و شماره منحنی در عامل ششم به عنوان مؤثرترین پارامترها شناخته شدند. تصویر گرافیکی مقادیر ویژه در هریک از عامل‌های استخراج شده در شکل (۲) نمایش داده شده است. همان‌طور که مشخص است، مقدار واریانس توجیه شده (مقدار ویژه) با استخراج عامل‌های ششم به بعد سریعاً افت می‌کند. همگن‌بندی به روش‌های موردنظر انجام و سپس جهت اعتباریابی گروههای همگن، روش تابع تشخیص مورد استفاده قرار گرفت و تعویض گروه‌ها بر مبنای ارزیابی مدل تا زمانی که میزان دقت مدل برابر 100 درصد شود، ادامه و نتایج در (جدول ۳) ارائه شده است. همان‌طور که مشخص است برای روش نزدیک‌ترین همسایه، یک منطقه همگن و برای سایر روش‌ها دو منطقه همگن در فاصله اقلیدسی صفر تا 10 تشخیص داده شد.

تأکید دارد، چنانچه این آماره در بازه $1/5$ تا $2/5$ قرار گیرد، فرض H_0 پذیرفته می‌شود و در غیر این صورت فرض H_0 رد می‌شود، یعنی بین خطاهای همبستگی وجود دارد.

بررسی نرمال بودن خطاهای
به منظور بررسی نرمال بودن خطاهای مقادیر استاندارد خطاهای محاسبه و نمودار هیستوگرام مقادیر خطاهای به همراه منحنی توزیع نرمال آن‌ها رسم می‌شود. توجه به میزان چولگی، میانگین و انحراف معیار خطاهای برای بررسی اعتبار روابط لازم است.

آزمون همخطی

همخطی، وضعیتی است که نشان می‌دهد یک متغیر مستقل، تابعی خطی از بین سایر متغیرهای مستقل است. اگر همخطی در یک معادله رگرسیون بالا باشد، بدین معنی است که بین متغیرهای مستقل همبستگی بالایی وجود دارد و ممکن است با وجود بالا بودن ضریب تعیین، مدل دارای اعتبار بالایی نباشد. به عبارت دیگر با وجود آن که مدل خوب به نظر می‌رسد، ولی دارای متغیرهای مستقل معنی داری نمی‌باشد. در جدول همبستگی داخلی، هر چقدر تولرانس کم باشد، اطلاعات مربوط به متغیرها کم بوده و شکل‌لایی در استفاده از رگرسیون ایجاد و باعث می‌شود، واریانس ضرایب رگرسیون، افزایش یافته و رگرسیون را برای پیش‌بینی نامناسب می‌سازد. همچنین مقادیر ویژه نزدیک به صفر نشان می‌دهد، همبستگی داخلی پیش‌بینی‌ها زیاد است و تغییرات کوچک در مقادیر داده‌ها به تغییرات بزرگ در براورد ضرایب معادله رگرسیون منجر می‌شود. از طرف دیگر شاخص وضعیت با مقدار بیشتر از 15 نشان‌دهنده احتمال همخطی بین متغیرهای مستقل است و مقدار بیشتر از 30 بیانگر مشکل جدی در استفاده از رگرسیون است.

نتایج و بحث

جریان پایه و شاخص مربوطه در ایستگاه‌های مورد نظر برای سال‌های آبی $1380-1385$ به روش فیلتر رقومی برگشتی دو پارامتر محاسبه گردید. ضرایب فیلتر استفاده شده

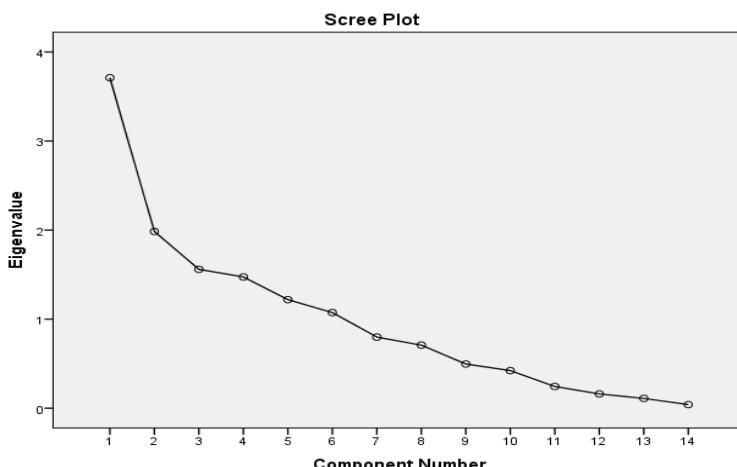
جدول ۲- مقادیر کل درصد از واریانس توضیح داده شده بر اساس عوامل مستقل انتخابی در آزمون تحلیل عاملی
Table2. Total percentages of the variance explained on the basis of independent factors in the factor analysis

عامل	درصد از واریانس	درصد از واریانس تجمعی
۱	$26/509$	$26/509$
۲	$143/173$	$40/682$
۳	$11/140$	$51/822$
۴	$10/527$	$62/349$
۵	$8/708$	$71/057$
۶	$7/671$	$78/728$

جدول ۳- زیر حوضه های انتخابی واقع در هر یک از زیر گروه ها بر اساس روش همگن بندی
Table 3. Selected sub-basins located in each of the sub-groups based on the homogeneity method

یک درصد دارد. مدل‌های رگرسیونی چند متغیره عامل شاخص جریان پایه، نسبت به متغیرهای مستقل در مناطق همگن به روش‌های مختلف، استخراج و در (جدول ۴) ارائه شده است.

روابط رگرسیونی برآورد شاخص جریان پایه نتایج روابط استخراجی با دخالت تمامی زیرخوشه‌ها، در مناطق همگن به روش‌های مختلف، نشان داد که مقادیر شاخص جریان پایه در برابر پارامترهای هیدرولوژیکی و فیزیوگرافی منتخب، همبستگی معنی‌داری را در سطح کمتر از



شکل ۲- مقادیر ویژه عامل‌های استخراج شده

٤- دا بات گ سون، در مناطق همگن:

Table 4. Regression relations in homogeneous regions

روش / شماره خوشه	خطای برآورد نسبی مدل (%)	رابطه
وارد (۱)	۰/۳۱	$BFI = -0.007S - 0.000058H + 0.001E - 0.00002P - 0.488K - +0.492$
وارد (۲)	۴۱/۷	$BFI = 0.01CN + 0.005F + 0.007S.A - 0.016S + 0.0000558H - 0.042T - 0.001E + 0.001P - 0.171K - 0.004Q90 + 1.032$
نزدیکترین همسایه (۱)	۳۷/۵۲	$BFI = S \times (-0.005) + LR \times (0.00004829) + H \times (-0.0000246) + T \times (-0.012) + Q_{90} \times (0.007) + Q_{75} \times (-0.01) + 0.518$
دورترین همسایه (۱)	۴۴/۰۲	$BFI = S \times (-0.002) + LR \times (9.503) + H \times (-0.0000443) + T \times (-0.005) + K \times (0.376) + 0.338$
دورترین همسایه (۲)	۶۰/۸۵	$BFI = S \times (-0.020) + H \times (-0.001) + T \times (-0.049) + H \times (-0.006) + P \times (-0.002) + K \times (-0.706) + Q_{90} \times (0.706) + 4.139$
فاصله از مرکز ثقل (۱)	۲۷/۱۳	$BFI = S \times (-0.002) + LR \times (0.000095) + H \times (0.0000443) - T \times (0.005) + K \times (0.376) + Q_{75} \times (0.003)$
فاصله از مرکز ثقل (۲)	۳۶/۸۹	$BFI = T \times (-0.105) + 1.233$

اعتبارسنجی مدل‌های رگرسیونی

نتایج اعتبارسنجی روابط رگرسیونی با استفاده از روش آزمون استقلال خطاهای (روش دوربین-واتسون)، برای مناطق همگن به روش‌های مختلف، نشان داد که مقدار آماره دوربین-واتسون در این پژوهش در مناطق همگن در فاصله ۱/۵ تا ۲/۵ قرار دارد و نشان‌دهنده این است که فرض استقلال خطاهای پذیرفته است و معادلات رگرسیون قابل استفاده است (جدول ۵).

پارامترهای مورداستفاده در روابط فوق

H: ارتفاع متوسط حوضه (متر)، S: شبب متوسط (درصد)، E: تبخیر و تعرق واقعی (میلی‌متر)، P: بارش متوسط سالانه (میلی‌متر)، K: ثابت افت منحنی هیدروگراف، Q₉₀ و Q₇₅: درصد عدم تجاوز، S_A: ضریب نگهداشت حوضه، F: نفوذپذیری (سانتی‌متر در ساعت)، CN: شماره منحنی، T: دمای متوسط سالانه (سانتی‌گراد)، LR: طول رودخانه اصلی.

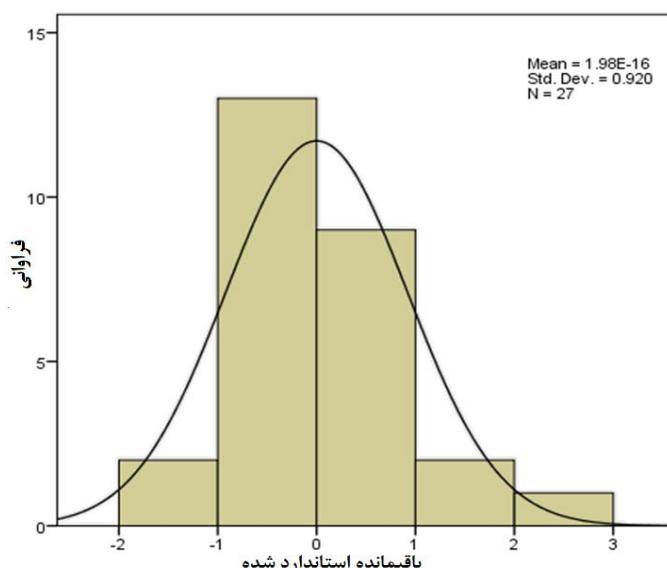
جدول ۵- خلاصه نتایج آزمون دوربین-واتسون

Table 5. Summary of Durbin Watson results

روش	منطقه همگن	ضریب همبستگی	ضریب تعیین	ضریب تعیین	خطای برآورد استاندارد مدل	ضریب دوربین-واتسون	ضریب دوربین-واتسون
فاصله از مرکز	۱	۰/۹۱۲	۰/۸۳۱	۰/۶۶۲	۰/۱۲۰۲۴	۰/۸۶۵	۰/۸۶۵
فاصله از مرکز	۲	۰/۸۰۰	۰/۷۷۵	۰/۷۲۵	۰/۱۲۹۵۰	۰/۵۹۲	۰/۵۹۲
دورترین همسایه	۱	۰/۹۱۲	۰/۸۳۱	۰/۶۶۲	۰/۱۲۰۲۴	۰/۷۱۸	۰/۷۱۸
دورترین همسایه	۲	۱	۱	۰	۰	۰/۹۲۲	۰/۹۲۲
نزدیکترین همسایه	۱	۰/۸۶۱	۰/۷۴۲	۰/۶۰۷	۰/۱۳۵۶۷	۰/۵۹۲	۰/۵۹۲
وارد	۱	۰/۹۰۲	۰/۸۱۳	۰/۳۲۷	۰/۱۶۶۴۳	۰/۵۹۵	۰/۵۹۵
وارد	۲	۰/۹۶۶	۰/۹۳۲	۰/۹۰۵	۰/۱۶۵۵۵	۰/۳۷۷	۰/۳۷۷
وارد	۳	۰/۸۰۰	۰/۶۴۰	۰/۵۷۵	۰/۱۶۱۸۹	۰/۹۰۷	۰/۹۰۷

رگرسیونی است، همچنین مقدار میانگین توزیع خطاهای در تمام مناطق همگن، نزدیک به صفر است و انحراف معیار آن‌ها نیز تقریباً یک است. شکل شماره (۳) نمونه‌ای از توزیع خطاهای برای شاخص BFI در منطقه همگن یک، به روش "فاصله از مرکز نقل" را نشان می‌دهد.

بررسی نرمال بودن خطاهای، مقادیر استاندارد خطاهای محاسبه و نمودار توزیع داده‌ها و نمودار نرمال آن‌ها رسم شد. با مقایسه نمودار توزیع فراوانی خطاهای و نمودار توزیع نرمال حاصل از آزمون نرمال بودن خطاهای مشاهده شد که توزیع خطاهای در تمامی مدل‌های استخراج شده، در مناطق همگن، تقریباً نرمال است که نشان‌دهنده قابلیت استفاده از روابط



شکل ۳- توزیع خطاهای برای شاخص BFI در منطقه همگن یک، به روش فاصله از مرکز نقل
Figure 3. Distribution of errors for the BFI in the first homogeneous region, by Centroid method

"فاصله از مرکز ثقل" به دلیل کمترین خطای نسبی به میزان ۲۷/۱۳۷ برای منطقه یک و ۳۶/۸۹ درصد برای منطقه همگن دو، حداقل خطای برآورد استاندارد و ضریب همبستگی بالا، به عنوان بهترین روش همگن‌بندی سلسله‌مراتبی برای برآورد منطقه‌ای شاخص جریان پایه در منطقه پژوهش تشخیص داده شد.

تشکر و قدردانی

این پژوهش با استفاده از امکانات پژوهشکده حفاظت خاک و آبخیزداری به انجام رسیده است. بدین‌وسیله از همراهی و مساعدت مسئولین محترم پژوهشکده تشکر می‌نماییم.

آزمون همبستگی داخلی

نتایج بررسی جدول همبستگی داخلی، نشان داد که شاخص وضعیت بعضی از عوامل مانند، عامل مساحت با میزان ۵۳/۶۹۷ و Q_{75} با میزان ۳۰/۰۴۹ و همچنین با دارا بودن مقادیر ویژه نزدیک به صفر (به ترتیب ۰/۰۰۳ و ۰/۰۱۰) نشانگر وجود احتمال همخطی بین متغیرهای مستقل در روش وارد است که هنگام انتخاب روابط به آن‌ها توجه شده است. برای بررسی میزان خطای نسبی برآورد روابط ارائه شده، مقادیر مربوط به پارامترهای حوضه‌های واقع در یک منطقه همگن به عنوان شاهد برای منطقه دیگر، در نظر گرفته شد و رابطه مربوطه برای پارامترهای حوضه‌های شاهد اجرا شد. سپس، میزان خطای نسبی برآورده برای روابط مناطق مختلف محاسبه و نتایج در جدول ۴ ارائه شده است. همان‌طور که از جدول مشخص است، روش همگن‌بندی

منابع

- Abbassizade, G.H. 2003. Regional flood frequency analysis using a hybrid method in South Alborz. MSc Thesis, Department of Natural Resources, Tehran University, Iran, 107 pp (In Persian).
- Ahani, A., S. Emamgholizadeh, S.S. M. Nadoushani and K. Azhdari. 2015. Regional Flood Frequency Analysis by Hybrid Cluster Analysis and L-moments, Journal of Watershed Management Research, 6(12): 11-20 (In Persian).
- Ahiablame, L., I. Chaubey, B. Engel, K. Cherkauer and V. Merwade. 2013. Estimation of annual base flow at ungauged sites in Indiana USA, Journal of Hydrology, 476: 13-27.
- Ataei, H. and M. Shiran. 2011. Identifying homogeneous hydrological basins based on effective geomorphologic variants on flood by cluster analysis. Journal of Geography and Environmental Planning, 42(2): 79-98 (In Persian).
- Beck, H.E., A.I.J.M. Van Dijk, D.G. Miralles, R.A.M. DeJeu, L.A. Bruijnzeel, T.R. McVicar and J. Schellekens. 2013. Global patterns in base flow index and recession based on stream flow observations from 3394 catchments, Journal of Water Resources Research, 49: 7843-7863.
- Biyabanaki, M. and S.S. Eslamiyan. 2005. Using cluster method in the determination of hydrologic homogeneous and evaluation by discriminant analysis and Andrew curves in Karkheh catchment, Journal of Agriculture, 2: 13-26 (In Persian).
- Blöschl, G., M. Sivapalan, T. Wagener, A. Viglione and H. Savenije. 2013. Runoff prediction in ungauged basins: Synthesis across Processes, Places and Scales, 1st edition, Cambridge Univ. Press, New York. 484 pp Hardback ISBN, 978-1-107-02818-0.
- Boughton, W.C. 1993. A hydrograph-based model for estimating the water yield of ungauged catchments. In: Hydrology and Water Resources Conference, Institution of Engineers, Australia, National Conference Publication no, 14: 317-324 pp.
- Brandes, D., J.G. Hoffmann and J.T. Mangarillo. 2005. Base flow recession rates, low flows, and hydrologic features of small watersheds in Pennsylvania, USA, Journal of The American Water Resources Association, 41: 1177-1186.
- Brauman, K.A., G.C. Daily, T.K. Duarte and H.A. Mooney. 2007. The nature and value of ecosystem services: An overview highlighting hydrologic services, Journal of Annual Review of Environment and Resources, 32: 67-98.
- Brodie, R.S. and S. Hostetle. 2005. A review of techniques for analyzing base-flow from stream hydrographs. Proceedings of the NZHS-IAH-NZSSS Conference, Auckland, New Zealand,
- Chahoki, A., A. Salajeghe, M. Mahdavi, S.H. Khaligi Sigarodi and A. sadi. 2013. Regional modeling of flow duration curve at ungauged catchments of arid region, Case study: Central of Iran. Journal of Range and Watershed Management, 66(2): 251-265 (In Persian).
- Chapman, T.G and A.I. Maxwell. 1996. Base flow Separation-Comparison of numerical methods with tracer experiments. Proceedings of the Hydrological and Water Resources Symposium. Water and the Environment; Preprints of Papers Institution of Engineers, 539-545 pp.
- Cyr, J., M. Landry and Y. Gagnon. 2011. Methodology for the large-scale assessment of small hydroelectric potential: Application to the province of New Brunswick (Canada), Journal of Renew. Energy, 36(11): 2940-2950.
- Ghanbarpor, M., M. Teymori and S.H.A. Gholami. 2008. Comparison of Base Flow Estimation Methods Based on Hydrograph Separation (Case study: Karun Basin). Journal of Science and Technology of Agriculture and Natural Resources, Water and Soil Science, 1: 1-10 (In Persian).

16. Kazemi, R. and A.R. Eslami. 2013. Investigation on the role of geological formation and hydrological parameter on base flow index, case study: Khazar region, *Journal of Watershed Engineering and Management*, 5(2): 85-93 (In Persian).
17. Kazemi, R., H. Davoodi, M.J. Soltani and A. Sarreshtedari. 2013. Investigation of land use change on base flow index, case study: Taleghan catchment, *Journal of Watershed Engineering and Management*, 5(1): 1-8 (In Persian).
18. Khosravi, M., M. Mohseni Saravi, M. Vafakhah and G. Fathi. 2011. Determination of the best hydrometric clustering method for regional flood analysis in west catchments of country. *Journal of Iran-Watershed Management Science and Engineering*, 4(13): 15-20 (In Persian).
19. Khosrobeigi Bozcheloei, S. and M. Vafakhah. 2016. Regional Analysis of Flow Duration Curve in Namak Lake Basin, Iran, *Journal of Watershed Management Research*, 7(14): 228-236 (In Persian).
20. Kjeldsen, T.R. and J.C. Smithers. 2002. Regional flood frequency analysis in the KwaZulu-Natal provinces, South Africa, using the index-flood method. *Journal of Hydrology*, 255: 194-211.
21. Krakauer, N.Y. and M. Temimi. 2011. Stream recession curves and storage variability in small watersheds, *Journal of Hydrology and Earth System Sciences*, 15: 2377-2389.
22. Longobardi, A. and P. Villani. 2008. Base flow index regionalization analysis in a Mediterranean area and data scarcity context: Role of the catchment permeability index. *Journal of Hydrology*, 355: 63-75.
23. Malekiyan, A., S. Hasanzadeh and H. Faraji Sabokbar. 2013. Application and comparing of gray cluster analysis and hierarchical analysis in locating suitable areas for flood spreading. Case study: Garebayegan catchment. *Journal of Iran-Watershed Management Science and Engineering*, 7(20): 35-46 (In Persian).
24. Mazvimavi, D., A.M.J. Maijerink, H.H.G. Savenije and A. Stein. 2005. Prediction of flow characteristics using multiple regression and neural networks: A case study in Zimbabwe. *Journal of Physics and Chemistry of the Earth*, 30: 639-647.
25. Mosavi, H. and A. Sepaskhah. 1989. Estimation of maximum daily discharge of ungauged catchments in Fars province. Proceedings of First Conference on Hydrology, University Tehran, 284-316 (In Persian).
26. Nathan, R.J. and T.A. McMahon. 1990. Identification of homogeneous regions for the purposes of regionalization. *Journal of Hydrology*, 121: 217-238.
27. Peña-Arcibia, J.L., A.I.J.M. Van Dijk, M. Mulligan and L.A. Bruijnzeel. 2010. The role of climatic and terrain attributes in estimating base flow recession in tropical catchments, *Journal of Hydrology and Earth System Sciences*, 14: 2193-2205.
28. Porhemmat, J. and R. Kazemi. 2017. Regional Modeling and Evaluation of Runoff Coefficient in Karkhe Basin, *Journal of Watershed Management Research*, 8(15): 82-91 (In Persian).
29. Rahimi, L., A. Ahmad Dehghani and K. Ghorbani. 2016. Comparison of Total flow, Base flow and Water-Quality Characteristics Trend in Arazkuseh Hydrometric Station, *Journal of Watershed Management Research*, 7(13): 83-91 (In Persian).
30. Rohani, H. 2001. Regional flood frequency analysis using a hybrid method in arid and semi-arid area. A Case Study of Khorasan. MSc Thesis. Department of Natural Resources, Tehran University, 132pp (In Persian).
31. Santhi, C., P.M. Allen, R.S. Muttiah, J.G. Arnold and P. Tappad. 2008. Regional estimation of base flow for the conterminous United States by hydrologic landscape regions. *Journal of Hydrology*, 351: 139-153.
32. Smakhtin, V.Y. 2001. Low flow hydrology: a review. *Journal of Hydrology*, 240: 147-186.
33. Teimouri, M., M.R. Ghanbarpour, M. Bashirgongbad, M. Zolfaghari and S. Kazemikia. 2011. Comparison of Base flow Index in Hydrograph Separation with Different Methods in Some Rivers of West Azarbaijan Province. *Journal of Science and Technology of Agriculture and Natural Resources, Water and Soil Science*, 15: 219-228 (In Persian).
34. Welderfael, W.A. and Y.E. Woyessa. 2010. Stream flow analysis and comparison of base flow separation methods, Case Study of the Modder River Basin in central South Africa. *Journal of European Water*, 31: 3-12.

Determination of the Best Hierarchical Clustering Method for Regional Analysis of Base Flow Index in Kerman Province Catchments

Rahim Kazemi¹ and Forood Sharifi²

1- Assistant Professor, Soil Conservation and Watershed Management Research Institute, Agricultural Research, Education and Extension Organization, Tehran, Iran, (Corresponding author: ra_hkazemi@yahoo.com)

2- Associate Professor, Soil Conservation and Watershed Management Research Institute, Agricultural Research, Education and Extension Organization, Tehran, Iran

Received: July 31, 2016

Accepted: January 13, 2019

Abstract

The lack of complete coverage of hydrological data forces hydrologists to use the homogenization methods in regional analysis. In this research, in order to choose the best Hierarchical clustering method for regional analysis, base flow and related index were extracted from daily stream flow data using two parameter recursive digital filters in 43 hydrometric stations of the Kerman province. Physiographic, climatic and hydrological parameters were calculated. Using factor analysis of 14 parameters, six factors including: hydrograph recession constant, mean annual precipitation, actual evapotranspiration, permeability, topography (slope and height) and curve number were chosen as the most important independent factors. Hierarchical clustering of catchments using independent factors in different methods including: Nearest neighbor, Furthest neighbor, Centroid clustering and Ward were down. Then the regional equations using linear regression at 1% significant level for each homogeneity region were determined. To compare and evaluate the accuracy and efficiency of the models, Independence errors, Colinearity and normal distribution of errors were tested. The results showed that "Centroid clustering method" because of low relative error with the amount of 27.137% for zone A and 36.89% for zone B of homogeneous region, minimum standard error of the estimate and the high correlation coefficient, diagnosed as the best homogeneity method to estimate regional base flow index in the study area.

Keywords: Base Flow Index, Factor Analysis, Hierarchical Clustering, Model Efficiency, Regional Analysis