

تعیین حوزه‌های آبخیز همگن جهت برآورد فرسایش و رسوب با استفاده از روش‌های مختلف تحلیل خوشه‌ای (مطالعه موردی: دامنه شمالی البرز)

- ❖ نسیم آرمان؛ استادیار، دانشکده منابع طبیعی دانشگاه صنعتی خاتم الانبیا بهبهان، ایران.
- ❖ علی سلاجقه*؛ استاد، دانشکده منابع طبیعی، دانشگاه تهران، ایران.
- ❖ سادات فیض‌نیا؛ استاد، دانشکده منابع طبیعی، دانشگاه تهران، ایران.
- ❖ حسن احمدی؛ استاد، دانشکده منابع طبیعی، دانشگاه تهران، ایران.
- ❖ جمال قدوسی؛ استادیار، گروه آبخیزداری دانشگاه آزاد واحد علوم و تحقیقات تهران، ایران.
- ❖ علی کیانی راد؛ استادیار، مؤسسه پژوهش‌های اقتصادی وزارت جهاد کشاورزی، ایران.

چکیده

شناسایی زیر حوزه‌های آبخیز همگن اولین گام ضروری تعمیم نتایج مطالعات محیطی می باشد. بدین منظور در این پژوهش ابتدا با استفاده از روش تجزیه و تحلیل عاملی (روش تحلیل مؤلفه‌های اصلی) از بین ۲۱ متغیر مؤثر بر تولید رسوب در ۲۷ ایستگاه هیدرومتری البرز شمالی، متغیرهای طول مستطیل معادل، میزان حساسیت سنگ به فرسایش، بارش متوسط سالانه، جهت و تراکم زهکشی که در مجموع ۸۰/۷۲ درصد واریانس داده‌ها را توجیه می‌کنند، انتخاب شدند ($KMO = 0/516$). سپس با استفاده از روش‌های مختلف تحلیل خوشه‌ای (سلسله‌مراتبی، دومرحله‌ای و میانگین K)، سه منطقه همگن به دست آمد. آنالیز تشخیص، صحت تحلیل خوشه‌ای در مناطق همگن را تأیید نمود. از سوی دیگر بر اساس این پنج عامل، یک تابع متمایزکننده تعریف شد که مقادیر همبستگی کانونیک، کای اسکور و آماره ویلکس لامبدا نیز تأیید کننده مجزا بودن این سه گروه منطقه همگن می‌باشند.

واژگان کلیدی: تحلیل عاملی، تحلیل خوشه‌ای، همگن‌بندی، آنالیز تشخیص، فرسایش و رسوب.

۱. مقدمه

مؤلفه‌های اصلی در ۱۲ حوزه آبخیز در سطح کشور، اقدام به تعیین مؤثرترین عامل‌ها شد. سپس بر پایه این عامل‌ها سه منطقه همگن با استفاده از تحلیل خوشه‌ای جداسازی گردید شد. نتایج نشانگر تأثیر برتر سه عامل شاخص خشکی دومازتن، شاخص پوشش گیاهی و درصد گستره سازندهای حساس به فرسایش در میزان تولید رسوب بوده است.

برای تعیین مناطق همگن بر اساس عوامل مساحت حوضه، تراکم آبراهه، شاخص بارندگی، طول آبراهه اصلی، شیب و عامل کمبود رطوبت خاک در ۱۶۸ حوضه مطالعاتی واقع در اسکاتلند تحقیقی صورت گرفته است که از روش تحلیل خوشه‌ای وارد استفاده نمودند و از بین پنج گروه همگن، چهار گروه قابل قبول و یک گروه قادر به توضیح تغییرات دبی نبوده است [۱].

در تحقیقی [۳] روش تحلیل خوشه‌ای را برای تقسیم‌بندی مناطق همگن به کار گرفتند. نتایج نشان داده که در گروه‌های همگن دارای ایستگاه هیدرومتری و یا فاقد ایستگاه می‌توان با دقت بیشتری جریان را به دست آورد.

به‌طور کلی همگن‌بندی زیر حوضه‌ها می‌تواند تعمیم نتایج مطالعات و بررسی‌ها در زمینه استفاده از مدل‌های منطقه‌ای برآورد فرسایش و رسوب، اولویت‌بندی زیر حوضه‌ها، اولویت‌بندی متغیرهای تأثیرگذار در تولید فرسایش و رسوب با آزمون تجزیه و تحلیل عاملی و کاربرد راهکارهای اجرائی به‌منظور کاهش و یا کنترل آن در مناطق همگن را فراهم نماید.

از آنجاکه استقرار دستگاه باران ساز جهت برآورد فرسایش و رسوب در ۲۷ حوزه آبخیز به علت وسعت منطقه، صرف هزینه و زمان بر بودن آن مقدور نیست، لذا هدف از آزمون همگن‌بندی، تعیین خوشه‌های همگن و انتخاب ایستگاه‌های معرف در هر خوشه می‌باشد. در این تحقیق، روش‌های مختلف خوشه‌بندی بررسی و بهترین روش به‌منظور تعیین ایستگاه‌های معرف استان‌های گیلان و مازندران جهت برآورد فرسایش و رسوب با دستگاه باران ساز انتخاب شد.

امروزه روند افزایشی فرسایش خاک و تولید رسوب یکی از نگرانی‌های زیست‌محیطی اصلی کشور محسوب می‌شود. بنابراین به‌منظور اجرای برنامه‌های حفاظت و مهار فرسایش خاک و کاهش تولید رسوبات ضرورت دارد که حجم کل بار رسوبی و شدت فرسایش در آن‌ها ارزیابی و برآورد گردد که لازمه آن نیز شناسایی عوامل مؤثر بر فرسایش و تولید رسوب است. اما عدم وجود و یا کمبود بسیار زیاد آمار و اطلاعات در این زمینه امکان بررسی مناسب را برای محاسبه مقادیر فرسایش و رسوب غیرممکن ساخته است. از جمله اقدامات مهم در رابطه با تعمیم اطلاعات در دسترس حوضه‌های دارای آمار به حوضه‌های بدون آمار، گروه‌بندی بر اساس همگنی حوضه‌ها می‌باشد. با توجه به اینکه عملاً لحاظ کردن کلیه عوامل دخیل در همگنی حوضه‌ها امکان‌پذیر نمی‌باشد، بدین خاطر محققین مختلف برخی از عوامل مهم را به‌عنوان مبنای همگنی حوضه‌ها لحاظ نموده‌اند.

در یک تحقیق طی تحقیقی بر روی در حوزه‌های آبخیز شمال کشور با انتخاب پارامترهای مساحت، طول آبراهه اصلی، شیب متوسط وزنی، ارتفاع متوسط، تراکم زهکشی و بارندگی متوسط سالیانه مربوط به ۳۱ حوزه آبخیز ناحیه خزر شرقی و ۲۳ حوزه آبخیز خزر غربی، با روش خوشه‌بندی سلسله‌مراتبی همگن‌بندی انجام گردید نمودند. در نهایت دو ناحیه فوق به چهار گروه همگن تقسیم‌بندی شدند [۱۰].

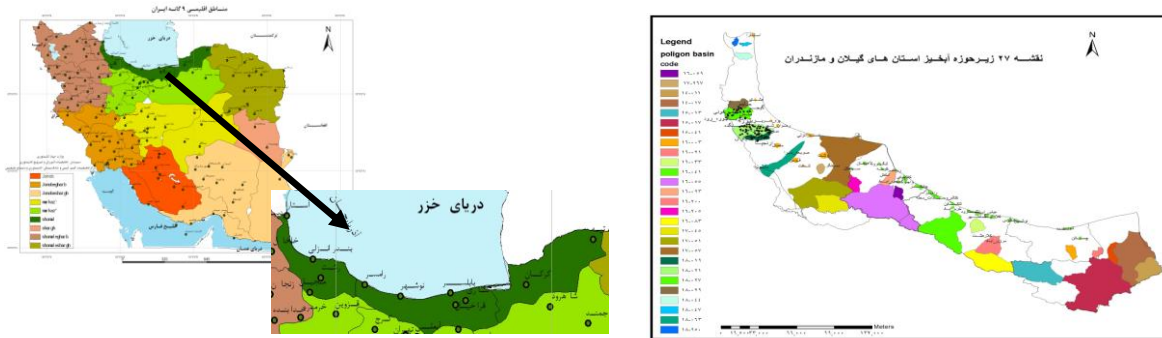
در پژوهشی دیگر همچنین در تحقیقی [۸] با استفاده از چهار عامل مساحت، بارش متوسط سالیانه، متوسط وزنی نفوذپذیری خیلی کم و شیب متوسط حوضه که ۳۶/۹۹ درصد از تغییرات داده‌های اصلی را توضیح می‌دهند، با تجزیه و تحلیل خوشه‌ای به روش واردز، دو منطقه همگن را بدست آوردند ($KMO = 0/629$).

بر اساس یک پژوهش در تحقیق [۵] با روش تحلیل

۲. روش‌شناسی تحقیق

در این تحقیق آمار ایستگاه‌های رسوب‌سنجی و آب‌سنجی ۲۷ زیرحوزه آبخیز استان‌های گیلان و مازندران (جدول ۱) که دارای حداقل ۳۰ سال آمار می‌باشند

استفاده می‌گردد (شکل ۱). ویژگی‌های ۲۷ حوزه آبخیز البرز شمالی تا محل ایستگاه‌ها در جدول ۲ لیست شده است. مراحل انجام پژوهش به شرح زیر می‌باشد.



شکل ۱. نقشه منطقه مطالعاتی (البرز شمالی)

جدول ۱. نام و مشخصات ایستگاه‌های استان‌های گیلان و مازندران

ردیف	ایستگاه	رودخانه	کد	طول	عرض	ارتفاع (متر)	مساحت (کیلومتر مربع)
۱	درازلات (طولوات)	پیلرود	۱۶-۰۵۵	۵۰-۱۷	۳۶-۵۹	۱۲۰۰	۱۶۳۴
۲	مرات بر (سموش)	سموش	۱۶-۰۵۹	۵۰-۱۸	۳۶-۵۹	۱۲۰۰	۶۱
۳	کل چال	شلمانرود	۱۶-۰۹۳	۵۰-۱۷	۳۷-۰۳	۱۵۰	۱۳۵
۴	توتکی	شمروود	۱۶-۰۲۵	۴۹-۵۲	۳۷-۰۴	۱۸۰	۱۳۱
۵	توتکا بن	توتکا بن	۱۷-۰۴۵	۴۹-۳۱	۳۶-۵۳	۱۳۰	۴۵۴
۶	شهریچار	زیلیکی رود	۱۷-۰۵۱	۴۹-۴۰	۳۷-۰۱	۴۰	۲۳۳
۷	آستانه (سفیدرود)	سفیدرود	۱۷-۰۵۷	۴۹-۵۶	۳۷-۱۷	-۱۰	۵۷۸۷۰
۸	یونل	شمارود	۱۸-۰۲۱	۴۹-۰۵	۳۷-۳۲	۵۰	۳۴۹
۹	ماشین خانه	گرگانرود	۱۸-۰۲۹	۴۸-۴۸	۳۷-۴۸	۱۳۰	۳۷۸
۱۰	باش محله	نوندویل	۱۸-۰۴۷	۴۸-۵۱	۳۸-۲۰	-۱۹	۵۰
۱۱	کمدول	ماسوله رودخان	۱۸-۰۶۳	۴۹-۰۳	۳۷-۰۷	۲۴۰	۲۱۳
۱۲	بهارستان	بهارستان	۱۸-۰۹۵	۴۸-۴۲	۳۸-۲۴	۲۰۰	۳۴
۱۳	رودبارسرا	چاقه رود	۱۸-۰۱۹	۴۹-۰۵	۳۷-۲۹	۱۳۵	۱۲۶
۱۴	لاکان	کومرود	۱۷-۰۶۷	۴۹-۳۴	۳۷-۱۱	۴۰	۲۹
۱۵	قربان محله	نمیر	۱۸-۰۴۴	۴۸-۵۱	۳۸-۱۵	۵۰	۵۱
۱۶	خرچگیل (اسالم)	ناورود	۱۸-۰۲۷	۴۸-۵۳	۳۷-۴۱	۶۵	۲۷۹
۱۷	قران طالار	بایلرود	۱۴-۰۱۱	۵۲-۴۷	۳۶-۱۷	۱۵۰	۴۰۳
۱۸	بایل (کشتارگاه)	بایلرود	۱۴-۰۱۷	۵۲-۳۹	۳۶-۳۲	۰	۱۴۳۰
۱۹	یلده	نور	۱۵-۰۱۳	۵۱-۴۸	۳۶-۱۲	۱۹۰۰	۷۵۴
۲۰	تنگه لایبج آغوزکتا	لایبج	۱۶-۰۰۳	۵۲-۰۳	۳۶-۲۴	۵۰۰	۱۰۴
۲۱	پل ذغال	چانوس	۱۶-۰۲۱	۵۱-۲۰	۳۶-۳۰	۳۰۰	۱۵۴۴
۲۲	ماشالله آباد	کاظم رود	۱۶-۰۲۳	۵۱-۰۶	۳۶-۴۰	۷۰	۱۸۸
۲۳	بلیبران	گرمروود	۱۵-۰۴۱	۵۲-۲۵	۳۶-۲۵	۳۰۰	۸۰
۲۴	اسکومحله	آلیش رود	۱۶-۰۲۰	۵۲-۱۸	۳۶-۲۳	۸۱	۸۱
۲۵	مرات بر (چشمه کیله)	چشمه کیله	۱۶-۰۴۱	۵۰-۵۰	۳۶-۴۵	۱۱۰	۷۶۸
۲۶	کره سنگ (۲۵ شهریور)	مرات	۱۵-۰۱۷	۵۲-۲۲	۳۶-۱۶	۲۲۰	۴۰۲۳
۲۷	ممین دره (آبشار)	چانوس	۱۶-۰۸۳	۵۱-۱۵	۳۶-۲۰	۸۵۰	۵۷۹

۱.۲. روش تجزیه و تحلیل عاملی

عمده‌ترین هدف استفاده از تحلیل عاملی، کاهش حجم داده‌ها و تعیین مهم‌ترین متغیرهای مؤثر در شکل‌گیری پدیده‌هاست. قبل از تجزیه و تحلیل خوشه‌ای، به منظور کاهش متغیرهای موجود (۲۱ متغیر مؤثر در فرایند فرسایش و تولید رسوب) از روش تجزیه و تحلیل عاملی استفاده شد و تأثیرگذارترین متغیرها به ترتیب اولویت انتخاب و در فرایند تجزیه خوشه‌ای مورد استفاده قرار گرفتند.

در این روش، نخست برای پرهیز از تأثیر واحدهای اندازه‌گیری، متغیرهای مورد استفاده استاندارد گردیدند. برای اطمینان از مناسب بودن داده‌ها برای تحلیل عاملی از ضریب $KMO < 0.5$ باشد، داده‌ها برای تحلیل عاملی مناسب صفر و یک در نوسان است (در صورتی که مقدار از همواره بین $0.5 < KMO$ باشد، داده‌ها برای تحلیل عاملی مناسب نخواهند بود). سپس با توجه به مقدار ارزش ویژه^۱ به دست آمده بر اساس روش مؤلفه‌های اصلی^۲، تمام عامل‌هایی که دارای ارزش ویژه بزرگ‌تر از یک بودند، به عنوان عامل‌های معنی‌دار انتخاب شدند. در گام بعدی، برای نشان دادن اهمیت نسبی هر عامل از نسبت واریانس منظور شده برای هر عامل چرخش شده (چرخش واریماکس^۳) استفاده شد. نهایتاً هر جا که متغیر، بیشترین مقدار مطلق بار عاملی را داشت و از نظر آماری نیز معنی‌دار بود، به عنوان متغیر تأثیرگذار انتخاب گردید.

۲.۲. روش تجزیه و تحلیل خوشه‌ای

در خوشه‌بندی مشاهدات هر گروه بیشترین شباهت و مشاهدات گروه‌های مختلف کمترین شباهت را باهم دارند. در این روش آن خصوصیتی که با عوامل دیگر وابستگی دارند، نباید توأمأً به عنوان خصوصیات متمایزکننده مورد استفاده قرار گیرند. ورود داده‌های خام، به خاطر اختلاف در مقیاس و تغییرپذیری دامنه داده‌ها، گروه‌های نادرستی را به وجود می‌آورند، بنابراین قبل از

محاسبه فواصل افراد، داده‌ها استاندارد شده و هم‌وزن گردیدند. فواصل بین افراد با استفاده از شاخص مربع فاصله اقلیدسی تعیین و در نهایت با استفاده از سه روش زیر، آزمون خوشه‌بندی انجام شد.

۱- آنالیز خوشه‌ای دمرحله‌ای^۴ که قابلیت خوشه‌بندی بر اساس متغیرهای گسسته و پیوسته، انتخاب خودکار تعداد خوشه‌ها و قابلیت تحلیل فایل داده‌های بسیار بزرگ را دارد.

۲- آنالیز خوشه‌ای میانگین^۵ که محدود به متغیرهای قابل اندازه‌گیری بوده، مناسب برای کار با داده‌های بزرگ می‌باشد و امکان ذخیره‌سازی فاصله‌ها از مرکز خوشه را دارد.

۳- آنالیز خوشه‌ای سلسله مراتبی^۶ که به خوشه‌های نهایی بر اساس میزان عمومیت آن‌ها ساختاری سلسله‌مراتبی، به صورت درختی (دندوگرام^۷) نسبت داده می‌شود. در اینجا از روش وارد^۸ برای خوشه‌بندی استفاده شد.

۳. نتایج

۱.۳. خروجی‌های تحلیل عاملی

با توجه به جدول ۳ مقدار $KMO > 0.5$ بوده، در نتیجه داده‌ها برای تحلیل عاملی مناسب خواهند بود. نتایج آزمون کرویت بارتلت نیز معنی‌دار است، به این مفهوم که فرض مخالف تأیید می‌شود، یعنی بین متغیرها همبستگی معنی‌دار وجود دارد. معنی‌دار بودن مربع کای نیز، بیانگر حداقل شرایط لازم برای اجرای تحلیل عاملی است.

¹ Eigen value

² PCA

³ Varimax

⁴ Two Step Cluster Analysis

⁵ K-means Cluster Analysis

⁶ Hierarchical Cluster Analysis

⁷ Dendogram

⁸ Ward

جدول ۳. آماره KMO و نتایج آزمون کرویوت بارتلت

KMO and Bartlett's Test		
Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		
		۰/۵۱۶
Bartlett's Test of Sphericity	Approx. Chi-Square	۸۳۱/۴۰۰
	Df	۲۱۶
	Sig.	۰/۰۰۰

عوامل به ترتیب ۳۴/۷۷۲، ۱۶/۸۵۸، ۱۳/۴۷۷، ۸/۵۴۱۸ و ۷/۰۵۷ و در مجموع ۸۰/۷۲۴ درصد از واریانس را در بردارند.

با توجه به جدول (۴) ستون Rotation Sums of Squared Loadings مشاهده می شود پنج عامل قابلیت تبیین واریانسها را دارند. این

جدول ۴. درصد واریانس و مقادیر ویژه عاملهای مختلف

Component	Total Variance Explained								
	Initial Eigenvalues			Extraction Sums of Squared Loadings			Rotation Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
۱	۹,۰۰۰	۳۷,۵۰۰	۳۷,۵۰۰	۹,۰۰۰	۳۷,۵۰۰	۳۷,۵۰۰	۸,۳۴۵	۳۴,۷۷۲	۳۴,۷۷۲
۲	۵,۳۱۲	۲۲,۱۳۲	۵۹,۶۳۲	۵,۳۱۲	۲۲,۱۳۲	۵۹,۶۳۲	۴,۰۴۶	۱۶,۸۵۸	۵۱,۶۳۰
۳	۲,۱۳۵	۸,۸۹۶	۶۸,۵۲۸	۲,۱۳۵	۸,۸۹۶	۶۸,۵۲۸	۳,۲۳۴	۱۳,۴۷۷	۶۵,۱۰۷
۴	۱,۶۳۹	۶,۸۲۷	۷۵,۳۵۵	۱,۶۳۹	۶,۸۲۷	۷۵,۳۵۵	۲,۰۵۰	۸,۵۴۱	۷۳,۶۴۸
۵	۱,۲۸۹	۵,۳۶۹	۸۰,۷۲۴	۱,۲۸۹	۵,۳۶۹	۸۰,۷۲۴	۱,۶۹۸	۷,۰۷۵	۸۰,۷۲۴
۶	۰,۹۶۸	۴,۰۳۴	۸۴,۷۵۸						
۷	۰,۸۲۲	۳,۴۲۶	۸۸,۱۸۵						
۸	۰,۷۰۹	۲,۹۵۵	۹۱,۱۴۰						
۹	۰,۵۶۷	۲,۳۶۴	۹۳,۵۰۴						
۱۰	۰,۴۳۵	۱,۸۱۳	۹۵,۳۱۸						
۱۱	۰,۳۵۵	۱,۴۷۹	۹۶,۷۹۶						
۱۲	۰,۲۳۳	۰,۹۶۹	۹۷,۷۶۵						
۱۳	۰,۱۹۴	۰,۸۰۸	۹۸,۵۷۴						
۱۴	۰,۱۰۵	۰,۴۳۶	۹۹,۰۱۰						
۱۵	۰,۰۸۸	۰,۳۶۵	۹۹,۳۷۵						
۱۶	۰,۰۵۳	۰,۲۲۱	۹۹,۵۹۶						
۱۷	۰,۰۴۹	۰,۲۰۴	۹۹,۸۰۰						
۱۸	۰,۰۲۰	۰,۰۸۳	۹۹,۸۸۳						
۱۹	۰,۰۱۵	۰,۰۶۴	۹۹,۹۴۷						
۲۰	۰,۰۰۸	۰,۰۳۱	۹۹,۹۷۸						
۲۱	۰,۰۰۲	۰,۰۰۹	۹۹,۹۸۸						
۲۲	۰,۰۰۱	۰,۰۰۶	۹۹,۹۹۳						
۲۳	۰,۰۰۱	۰,۰۰۵	۹۹,۹۹۸						
۲۴	۰,۰۰۱	۰,۰۰۲	۱۰۰,۰۰۰						

در ۲۷ ایستگاه هیدرومتری دامنه‌های البرز شمالی، پنج متغیر طول مستطیل معادل، میزان حساسیت سنگ به فرسایش، بارش متوسط سالانه، جهت و تراکم زهکشی که در مجموع ۸۰/۷۲ درصد واریانس داده‌ها را توجیه

با توجه به جدول (۵) (ماتریس عاملی دوران یافته به روش واریماکس) در شناسایی زیر حوزه‌های آبخیز همگن با استفاده از روش تجزیه و تحلیل عاملی (روش تحلیل مؤلفه‌های اصلی) از بین ۲۱ متغیر مؤثر بر تولید رسوب

کار می‌رود. با توجه به این نمودار مشاهده می‌شود که از عامل پنجم به بعد تغییرات مقدار ویژه کم می‌شود، پس می‌توان پنج عامل مذکور در جدول (۵) را به‌عنوان عوامل مهم که بیشترین نقش را در تبیین واریانس داده‌ها دارند، استخراج کرد.

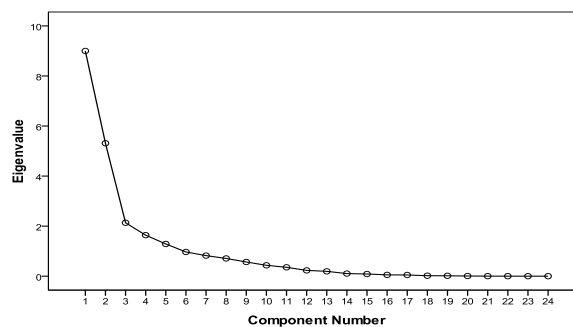
می‌کنند، انتخاب شدند. عامل شیب متوسط حوضه نیز در ستون دوم پس از متغیر میزان حساسیت سنگ به فرسایش در اولویت بعدی قرار می‌گیرد. شکل (۲) تغییرات مقادیر ویژه را در ارتباط با عامل‌ها نشان می‌دهد. این نمودار برای تعیین تعداد بهینه مؤلفه‌ها به

جدول ۵. ماتریس عاملی دوران یافته

Rotated Component Matrix ^a					
	Component				
	۱	۲	۳	۴	۵
متوسط بارش سالانه	-.۰۶۹	-.۲۵۵	.۸۰۵	.۳۲۶	-.۱۵۱
متوسط دبی رسوب سالانه	.۹۰۶	-.۱۷۴	.۱۱۱	-.۰۴۳	.۰۰۷
متوسط دبی آب سالانه	.۸۸۰	-.۲۰۵	.۱۵۵	.۰۱۶	.۰۳۸
ارتفاع از سطح دریا	-.۰۶۴	.۳۹۴	-.۲۲۸	.۱۲۲	.۰۹۷
طول ابراهه های کل حوضه	.۹۳۳	-.۱۲۳	.۰۰۴	.۱۴۱	.۲۱۲
جهت	-.۰۷۸	.۰۹۴	-.۰۱۴	.۷۸۹	.۲۶۹
تراکم زهکشی	.۲۸۳	-.۳۸۹	.۱۹۲	.۰۶۰	.۷۱۸
ضریب فشردگی (کراپوس)	.۸۷۵	-.۱۳۴	-.۰۳۹	-.۳۱۵	.۰۷۱
ضریب فرم حوضه (هورتون)	.۱۲۶	.۴۲۰	-.۲۲۴	.۲۵۸	.۰۵۷
ضریب کشیدگی حوضه (شیوم)	-.۶۶۰	-.۰۶۶	.۳۲۷	-.۳۲۸	-.۳۵۸
ضریب گردی (میلر)	-.۸۱۵	.۰۱۶	.۱۷۸	.۳۲۸	-.۰۲۳
عرض مستطیل معادل	.۶۲۹	.۲۶۸	-.۳۵۰	.۵۶۹	.۰۹۰
طول مستطیل معادل	.۹۶۸	-.۰۶۰	-.۰۶۲	.۰۶۶	.۱۵۱
طول حوزه به کیلومتر	.۹۴۶	-.۰۸۸	-.۱۶۲	.۱۶۵	-.۰۸۲
اختلاف ارتفاع KM	.۳۰۸	.۵۸۵	-.۲۳۸	.۵۱۲	-.۲۵۴
ارتفاع متوسط حوزه	.۰۳۵	.۸۲۵	-.۲۶۶	.۳۴۱	.۱۹۹
طول شاخه اصلی رودخانه	.۹۰۵	.۰۰۳	-.۱۹۲	.۲۰۱	-.۱۱۷
شیب متوسط حوزه %	-.۳۸۱	.۸۳۹	.۱۷۵	.۱۶۰	.۰۳۶
شیب خالص شاخه اصلی %	-.۶۷۱	.۱۲۹	.۴۲۴	.۱۱۶	-.۰۱۵
شدت فرسایش	.۰۳۲	.۱۱۰	-.۸۰۶	.۱۰۲	-.۰۵۸
میزان حساسیت سنگ	-.۱۴۳	.۸۴۵	-.۱۲۷	-.۱۲۳	.۲۱۰
شاخص شدت فرسایش دهندگی	.۰۲۲	-.۳۰۹	.۷۴۳	-.۲۱۹	-.۰۸۸

Extraction Method: Principal Component Analysis.
 Rotation Method: Varimax with Kaiser Normalization.
 a. Rotation converged in ۷ iterations.

Scree Plot



شکل ۲. نمودار اسکری گراف برای تعیین تعداد عامل‌ها

۲.۳. آنالیز تشخیص

پس از مشخص شدن گروه‌های همگن باید به دو مسئله اساسی توجه داشت. یکی اینکه آیا گروه‌های ایجادشده کاملاً مجزا از یکدیگر هستند؟ دیگر اینکه وضعیت حوضه‌ای که در تجزیه و تحلیل شرکت نکرده و یا حوضه فاقد آمار را چگونه می‌توان مشخص نمود که به کدام گروه تعلق دارد؟

با انجام آنالیز تشخیص یک تابع یا مجموعه‌ای از توابع ساخته می‌شود برای K گروه، $K-1$ تابع تشخیص ساخته می‌شود. اولین تابع بهترین ترکیب خطی برای پیش‌بینی عضویت در گروه‌ها به دست می‌دهد. برای تعیین بهترین تابع از شاخص لامبدای ویلکس استفاده می‌شود. مقدار این شاخص بین صفر و یک متغیر است. هر چه مقدار برای یک تابع کوچک‌تر باشد، آن تابع تفکیک‌کننده خوبی است. از آنجاکه توزیع این شاخص به کای اسکور شبیه است، از این رو از طریق این آماره نیز تعبیر می‌شود.

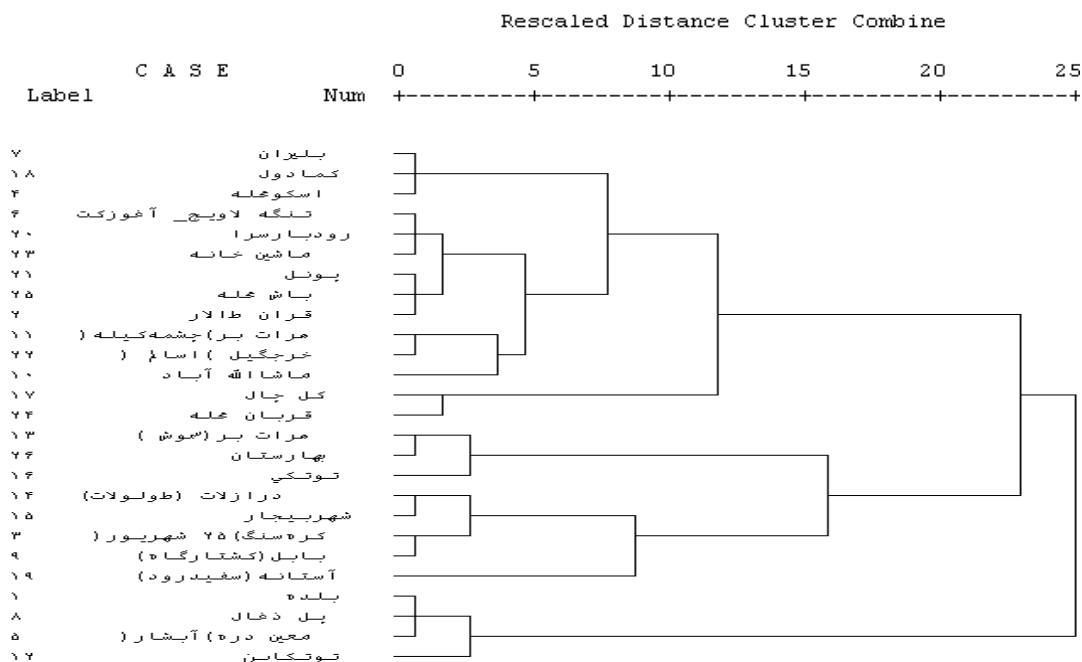
با توجه به جداول ۹، ۱۰ و ۱۱ آنالیز تشخیص، صحت تحلیل خوشه‌ای در مناطق همگن را تأیید می‌کند.

۳.۳. نتایج تحلیل خوشه‌ای به روش سلسله‌مراتبی

شکل ۳، دندروگرام حاصل از تحلیل خوشه‌ای سلسله‌مراتبی را نشان می‌دهد. تعداد مطلوب خوشه‌ها در روی نمودار دندروگرام جایی است که فاصله زیاد بین ادغام دو خوشه مشاهده می‌شود. از معادله ۱ برای تعیین تعداد خوشه استفاده می‌شود که در آن N ، تعداد کل افراد است. با توجه به این معادله $\sqrt{\frac{27}{2}}$ برابر $\frac{3}{6}$ می‌شود. در روش خوشه‌بندی سلسله‌مراتبی ۲۷ ایستگاه در فاصله اقلیدسی ۲۳ در ۳ گروه همگن قرار گرفته‌اند.

$$\sqrt{\frac{N}{2}} \quad (1) \quad \text{تعیین تعداد خوشه}$$

Dendrogram using Ward Method



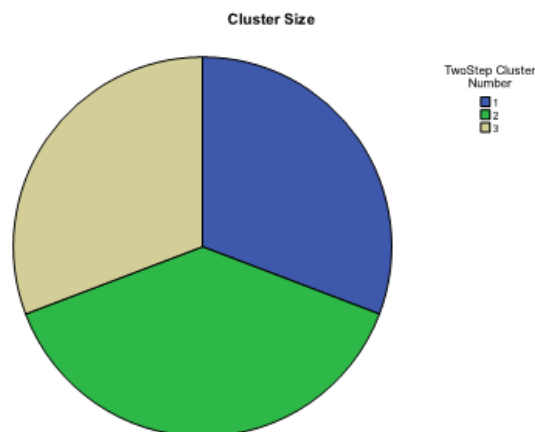
شکل ۳. دندروگرام حاصل از تحلیل خوشه‌ای سلسله‌مراتبی

می‌دهد، ایستگاه‌ها در ۳ منطقه همگن جای گرفته‌اند که در جدول (۷) نام ایستگاه‌های مربوط به هر خوشه همگن ارائه شده است.

۴.۳. نتایج تحلیل خوشه‌ای به روش دومرحله‌ای
جدول (۶) تعداد و درصد حضور ایستگاه‌ها را در خوشه‌های تعیین شده نشان می‌دهد. شکل (۴)، نمایی از اندازه خوشه‌ها با روش خوشه‌بندی دومرحله‌ای را نشان

جدول ۶. توزیع ایستگاه‌ها در خوشه‌ها

Cluster Distribution				
Cluster		N	% of Combined	% of Total
1		8	30.8%	29.6%
2		10	38.5%	37.0%
3		8	30.8%	29.6%
Combined		26	100.0%	96.3%
Excluded Cases		1		3.7%
Total		27		100.0%



شکل ۴. اندازه خوشه‌ها

جدول ۷. تعیین گروه‌های همگن به روش خوشه‌بندی دومرحله‌ای

ایستگاه	گروه همگن	ایستگاه	گروه همگن	ایستگاه	گروه همگن	ایستگاه	گروه همگن
بلده	۱	بلیران	۱	درازلات	۲	آستانه (سفیدرود)	۲
قران‌طلالار	۳	پل‌ذغال	۱	شهربیجار	۳	رودبارسرا	۲
کره‌سنگ	۳	بابل (کشتارگاه)	۱	توتکی	۲	پونل	۳
اسکومحله	۱	ماشالله‌آباد	۳	کل‌چال	۲	خرگیل (اسالم)	۲
معین‌دره (آبشار)	۱	هراتیر (چشمه‌کیله)	۲	لاکان		ماشین‌خانه	۲
تنگه‌لاویج (آغوزکشی)	۱	توتکابن	۲	کمدول	۱	قربان‌محله	۳
بهارستان	۳	هرات‌یر (سموش)	۲	باش‌محله	۳		

نشان می دهد.

۵.۳. نتایج تحلیل خوشه‌ای به روش میانگین K

جدول (۸) عضویت خوشه‌ها در روش میانگین K را

جدول ۸. عضویت خوشه‌ها در روش میانگین K

ایستگاه	گروه همگن	ایستگاه	گروه همگن	ایستگاه	گروه همگن	ایستگاه	گروه همگن
بلده	۱	بلیران	۲	درازلات	۳	آستانه (سفیدرود)	۳
قران طالار	۲	پل ذغال	۱	شهربیجار	۳	رودبارسرا	۲
کره سنگ	۲	بابل (کشتارگاه)	۱	توتکی	۳	پونل	۳
اسکومحله	۱	ماشالله آباد	۳	کل چال	۳	خرگیل (اسالم)	۳
معین دره (آبشار)	۱	هراتیر (چشمه کیله)	۳	لاکان	۳	ماشین خانه	۲
تنگه لایوچ (آغوزکتی)	۲	توتکابن	۱	کمادول	۲	قربان محله	۳
بهارستان	۳	هراتیر (سموش)	۳	باش محله	۳		

جدول ۹. نتایج آنالیز تشخیص به روش تحلیل خوشه‌ای سلسله مراتبی

مقادیر ویژه	عامل	مقدار ویژه	درصد واریانس	درصد تجمعی	همبستگی کانونیکال
(روش سلسله مراتبی)	۱	۵/۵۸۴	۷۴	۷۴	۰/۹۲۱
	۲	۱/۹۶۶	۲۶	۱۰۰	۰/۸۱۴
ویلیکس لامبدا	آزمون عوامل	ویلیکس لامبدا	کای اسکور	درجه آزادی	سطح معناداری
	۱ بواسطه ۲	۰/۵۱	۶۲/۴۰۶	۱۰	۰/۰۰۰
	۲	۰/۳۳۷	۲۲/۸۲۹	۴	۰/۰۰۰
تعداد خوشه‌های روش سلسله مراتبی	اعضای گروه پیش‌بینی شده			مجموع	
	۱	۲	۳		
۱	۱۳	۱	۰	۱۴	
۲	۰	۸	۰	۸	
۳	۰	۰	۴	۴	
گروه‌بندی نشده	۱	۰	۰	۱	
واقعی	۱	۷/۱	۰	۱۰۰	
	۲	۰	۱۰۰	۱۰۰	
	۳	۰	۰	۱۰۰	
گروه‌بندی نشده	۱۰۰	۰	۰	۱۰۰	

تعداد خوشه‌های روش سلسله مراتبی

طبقه‌بندی نتایج

۹۶/۲٪ درست طبقه‌بندی شده است
96.2% of original grouped cases Correctly classified.

جدول ۱۰. نتایج آنالیز تشخیص به روش تحلیل خوشه‌ای دومرحله‌ای

عامل	مقدار ویژه	درصد واریانس	درصد تجمعی	همبستگی کانونیکال	مقادیر ویژه (روش دومرحله‌ای)
۱	۱/۹۹۵	۹۴/۲	۹۴/۲	۰/۸۱۶	
۲	۰/۱۲۳	۵/۸	۱۰۰	۰/۳۳۱	
آزمون عوامل	ویلکس لامبدا	کای اسکور	درجه آزادی	سطح معناداری	ویلکس لامبدا
۱ بواسطه ۲	۰/۲۹۷	۲۵/۴۷۲	۱۰	۰/۰۰۵	
۲	۰/۸۹۰	۲/۴۳۹	۴	۰/۶۵۶	
تعداد خوشه‌های روش دومرحله‌ای				اعضای گروه پیش‌بینی شده	
تعداد				مجموع	
۱				۳	
۲				۰	
۳				۰	
گروه‌بندی نشده				۱	
واقعی				۱۰۰	
۱				۰	
۲				۲۰	
۳				۶۲/۵	
گروه‌بندی نشده				۱۰۰	
				۰	
				۳۷/۵	
				۱۰۰	

۷۶/۹٪ درست طبقه بندی شده است
76.9% of original grouped cases Correctly classified.

جدول ۱۱. نتایج آنالیز تشخیص به روش تحلیل خوشه‌ای میانگین K

عامل	مقدار ویژه	درصد واریانس	درصد تجمعی	همبستگی کانونیکال	مقادیر ویژه (روش میانگین k)
۱	۹/۰۹۹	۹۴/۳	۹۴/۳	۰/۹۴۹	
۲	۰/۵۵۲	۵/۷	۱۰۰	۰/۵۹۶	
آزمون عوامل	ویلکس لامبدا	کای اسکور	درجه آزادی	سطح معناداری	ویلکس لامبدا
۱ بواسطه ۲	۰/۰۶۴	۵۷/۷۹۰	۱۰	۰/۰۰۰	
۲	۰/۶۴۴	۹/۲۳۰	۴	۰/۰۵۶	
تعداد خوشه‌های روش میانگین k				اعضای گروه پیش‌بینی شده	
تعداد				مجموع	
۱				۳	
۲				۰	
۳				۰	
گروه‌بندی نشده				۱	
واقعی				۱۰۰	
۱				۰	
۲				۱۰۰	
۳				۹۲/۳	
گروه‌بندی نشده				۱۰۰	
				۰	
				۷/۷	
				۱۰۰	

۹۶/۲٪ درست طبقه‌بندی شده است
96.2% of original grouped cases Correctly classified.

۴. بحث و نتیجه گیری

نتایج حاصل از تجزیه و تحلیل عاملی جهت تشخیص پارامترها و متغیرهای مستقل از میان ۲۱ متغیر مورد بررسی در ۲۷ ایستگاه هیدرومتری دامنه شمالی البرز به کمک نرم افزار SPSS حاکی از این است که پنج عامل طول مستطیل معادل، میزان حساسیت سنگ به فرسایش، بارش متوسط سالانه، جهت و تراکم زهکشی در مجموع ۸۰/۷۲ درصد واریانس داده‌ها را توجیه می‌کنند. سپس با استفاده از روش‌های مختلف تحلیل خوشه‌ای (سلسله‌مراتبی، دومرحله‌ای و میانگین K)، سه منطقه همگن به دست آمد. آنالیز تشخیص، صحت تحلیل خوشه‌ای در مناطق همگن را تأیید نمود. یک تابع متمایزکننده تعریف شد که مقادیر همبستگی کانونیک، کای اسکور و آماره ویلکس لامبدا نیز نشان‌دهنده مجزا بودن این سه گروه همگن می‌باشند که این نتایج با نتایج تحقیقات [۱، ۲، ۳، ۴، ۵، ۶، ۷، ۸، ۹، ۱۰، ۱۱] هم‌خوانی دارد.

با توجه به اینکه الگوریتم‌های زیادی برای تجزیه خوشه‌ای وجود دارند، ولی هیچ‌کدام از آن‌ها به‌عنوان بهترین روش پذیرفته نشده‌اند. الگوریتم‌های مختلف نتایج یکسانی را برای مجموعه معینی از داده‌ها به بار نمی‌آورد. معمولاً نظر پژوهشگر نقش مهمی در ارزیابی نتایج حاصل از هر روش خاص بازی می‌کند [۱۲]. در این مورد گفتن خوب‌ترین در زمینه الگوریتم‌های خوشه‌بندی صحیح نیست، از مشکلات خوشه‌بندی به روش میانگین k این است که محدود به متغیرهای قابل اندازه‌گیری است. تعداد خوشه‌ها از ابتدا مشخص است، اما معمولاً در کاربردهای زیادی تعداد خوشه‌ها مشخص نمی‌باشد. جواب نهایی به انتخاب خوشه‌های اولیه بستگی دارد. در این روش روالی مشخص برای محاسبه اولیه مراکز

خوشه‌ها وجود ندارد و اگر در تکراری از الگوریتم تعداد داده‌های متعلق به خوشه‌ای صفر شد، راهی برای تغییر و بهبود ادامه روش وجود ندارد. با توجه به اینکه روش تحلیل خوشه‌ای دومرحله‌ای هم‌زمان با متغیرهای پیوسته و گسسته به‌خوبی کار می‌کند و می‌تواند فایل داده‌های بسیار بزرگ را تحلیل نماید، همچنین تعداد خوشه‌ها هم به‌صورت خودکار و هم مطابق با نظر پژوهشگر تعیین می‌گردد، اما امروزه از روش تجزیه خوشه‌ای سلسله‌مراتبی به دلیل وجود روش‌های مختلف خوشه‌بندی (نزدیک‌ترین همسایه، دورترین همسایه، خوشه‌بندی متمرکز، خوشه‌بندی میانه و روش واردز)، انتخاب خودکار و عدم دخالت پژوهشگر در تعداد خوشه‌ها و نمایش خروجی به‌صورت شاخه درختی بیشتر استفاده می‌شود و به‌عنوان پرکاربردترین روش تجزیه خوشه‌ای به کار می‌رود.

با توجه به هدف این تحقیق در برآورد فرسایش و رسوب، با در نظر گرفتن وسعت استان‌های گیلان و مازندران، صرف‌هزینه و زمان بر بودن استقرار باران ساز در کل منطقه، تعیین ایستگاه‌های معرف منطقه و بالطبع آن ضرورت انجام تجزیه و تحلیل عاملی برای اولویت‌بندی ورودی‌های آنالیز خوشه‌ای به‌منظور تعیین حوزه‌های آبخیز همگن لازم و ضروری به نظر می‌رسد. با توجه به نتایج برآورد میزان رسوب و فرسایش در حوضه‌های معرف منطقه که از هر خوشه انتخاب می‌گردند، می‌توان نتایج را به حوضه‌های بدون آمار یا حوضه‌های دارای آمار ناقص تعمیم داد، چراکه حوضه‌های همگن انتخابی در هر خوشه، معرف کل حوضه‌های موجود در آن خوشه می‌باشند و در نهایت می‌توان سیاست‌ها، برنامه‌ریزی‌ها و تصمیم‌گیری‌های لازم را در زمینه کاهش میزان فرسایش و تولید رسوب در سطح منطقه مورد مطالعه اتخاذ کرد.

References

- [1] Acreman, M. C. and Sinclair, C. D. (1986). Classification of Drainage Basins According to Their Physical Characteristics for Flood Frequency Analysis in Scotland. *Journal of Hydrology*, 84, 365- 384.
- [2] Biabanaki, M. and Eslamian, S. (2004). Cluster analysis for determination of the Hydrologic homogeneity, assessment with discriminant analysis and Andrew curves in Karkheh basin. *Agriculture Journal*, 6(2), 13-26.
- [3] Burn D. H. and Goel, N.K. (2000). The Formation of Groups for Regional Floods Frequency Analysis. *Hydrological Sciences Journal*, 45, 97-112.
- [4] Kaufman, H. and Rousseeuw, P.J. (1990). Finding groups in data an introduction to cluster analysis. *Jahn willey and Sons Inc*. DOI: 10.1002/9780470316801.
- [5] Modarres, R., Feiznia, S., Nasri, M. and Najafi, A. (2010). Determination of Homogenous Regions Based on Some Effective Factors on Sediment Yield. *Journal of Range and Watershed Management. Iranian Journal of Natural Resources*, 63(2), 249-260.
- [6] Nathan, R.J. and McMahon, T.A. (1990). Identification of homogenous regions for purposes of regionalization. *Journal of Hydrology*, 121, 217-238.
- [7] Nosrati, K., Mohseni Saravi, M., and Mahdavi, M. (2004). Identification of Homogenous Regions for Low Flow Frequency Analysis. *Journal of Range and Watershed Management, Iranian Journal of Natural Resources*, 57(1), 45-58.
- [8] Samiee, M., Mahdavi, M., Saghafian, B., and Mohseni Saravi, M. (2005). Determining effective factors on low flow regime using component analysis. *Journal of Agricultural Sciences and Natural Resources*, 12(2), 118-127.
- [9] Tasker, G.D. (1982). Comparing Methods of Hydrologic Regionalization. *Water Resources Bulletin*, 18(6), 965-970.
- [10] Telvari, A. and Islami, A. (2005). Influence of Watershed basin homogeneity in *Regional flood frequency water and watershed journal*, 1(3), 39-48.
- [11] Telvari, A. and Samiee, M., (2009). Homogeneous watershed using statistical analysis and efficient factor on flood in Fars province. *Watershed Management Research (Pajouhesh & Sazandegi)*, 85, 2-10.
- [12] Zare Chahooki, M. A. (2010). *Data analysis in natural resources research using SPSS software*, first edition, university of Tehran press.

