

## ارزیابی روش‌های آزمون گاما، تحلیل خوشه‌ای، تحلیل تابع تشخیص و منحنی‌های اندرو در تفکیک آبخیزهای همگن برای آنالیز منطقه‌ای رسوب معلق

حسین خیرفام<sup>۱</sup> و مهدی وفاخواه<sup>۲\*</sup>

(۱) دانشجوی دکتری؛ گروه مهندسی آبخیزداری؛ دانشکده منابع طبیعی؛ دانشگاه تربیت مدرس؛ نور؛ مازندران؛ ایران  
(۲) دانشیار؛ گروه مهندسی آبخیزداری؛ دانشکده منابع طبیعی؛ دانشگاه تربیت مدرس؛ نور؛ مازندران؛ ایران  
\*نویسنده مسئول مکاتبات: [vafakhah@modares.ac.ir](mailto:vafakhah@modares.ac.ir)

تاریخ دریافت: ۱۳۹۲/۰۷/۱۸ تاریخ پذیرش: ۹۳/۱۱/۱۳

### چکیده

رسوب ناشی از فرسایش خاک به‌عنوان مهم‌ترین نمایه‌های تخریب اراضی، چالشی مهم در بحث توسعه پایدار و تهدیدی بر زیست‌بوم‌ها تلقی می‌شود. لذا تخمین رسوب خروجی از آبخیزها بسیار حائز اهمیت می‌باشد. گستردگی آبخیزها و کمبود ایستگاه‌های سنجش رسوب باعث شده تا از روش‌های غیرمستقیم از قبیل مدل‌های تهیه شده در آبخیزهای مشابه جهت تخمین رسوب در آبخیز فاقد آمار استفاده نمایند. در این پژوهش، ۴۲ ایستگاه رسوب‌سنجی جنوب و جنوب شرقی دریای خزر با دوره‌های آماری بیش از ۲۰ سال انتخاب گردید. با برقراری ارتباط بین مقادیر دبی و رسوب معلق در مواقع سیلابی اقدام به تخمین مقادیر روزانه رسوب معلق از روی دبی‌های روزانه شده و مقادیر رسوب متوسط سالانه محاسبه گردید. با استفاده از آزمون گاما، ۱۳ عامل موثر بر تولید رسوب به ۵ عامل اصلی کاهش داده شد و ایستگاه‌های مورد مطالعه با به‌کارگیری تحلیل خوشه‌ای، تحلیل تابع تشخیص و منحنی‌های اندرو در گروه‌های همگن قرار گرفتند. برای هر گروه همگن حاصل از هر یکی از روش‌های همگن‌سازی مذکور و برای کل ایستگاه‌ها به‌عنوان یک گروه با ۵ عامل اصلی انتخاب شده مدل رگرسیونی برای تخمین میزان رسوب معلق متوسط سالانه تهیه شد و مقادیر خطا و کارایی مدل‌های تهیه شده با استفاده از نمایه‌های آماری RE، RRMSE، RBIAS براساس داده‌های مشاهداتی محاسبه گردید. نتایج نشان داد که هر ۳ روش همگن‌سازی نتایج بهتری نسبت به مدل کلی داشته و روش منحنی‌های اندرو به‌ترتیب با مقادیر خطای نسبی (RE)، ریشه نسبی مجذور میانگین مربعات خطا (RRMSE) و انحراف نسبی (RBIAS) ۳۸/۱۲ و ۵۳/۱۶ درصد و ۰/۰۱- در مرحله واسنجی و ۴۵/۹۱ و ۳۳/۱۱ درصد و ۰/۰۱ در مرحله اعتبارسنجی نسبت به دو روش تحلیل خوشه‌ای و تحلیل تابع تشخیص عملکرد بهتری در همگن‌سازی حوزه‌های آبخیز بر اساس تولید رسوب داشته است. همچنین دبی اوج با دوره بازگشت دو ساله نیز بیش‌ترین تأثیر در تغییرات میزان رسوب معلق متوسط سالانه را دارد.

**کلید واژه‌ها:** تخمین رسوب معلق؛ مدل‌سازی رسوب؛ منحنی سنجه رسوب؛ همگن‌سازی آبخیزها

### مقدمه

اهمیت می‌باشد (Alp and Cigizoglu, 2007)، رسوب خروجی از آبخیز به‌وسیله فرآیندهای فیزیکی برداشت، انتقال و ته‌نشست صورت می‌گیرد و غلظت رسوب متاثر از شدت و حجم بارش، توپوگرافی، پوشش سطح زمین، فعالیت‌های مدیریتی، قابلیت جدایش و بافت خاک می‌باشد (Sadeghi and Singh, 2005; Melesse, 2011).

روند تخریب اراضی و به‌تبع آن پیامدهایی از قبیل سیل و رسوب‌زایی در سطح جهان، به‌ویژه در کشورهای در حال توسعه، رو به افزایش می‌باشد (Sadeghi et al., 2008). لذا فرسایش خاک و رسوب تولیدی ناشی از آن در علوم حفاظت خاک و مهندسی هیدرولیک حائز

رسوب معلق (Tramblay et al., 2007) رابطه معنی‌داری وجود دارد. نبود اطلاعات رسوب خروجی از حوضه‌ها به دلیل خرابی یا عدم وجود ایستگاه رسوب‌سنجی، مشکلاتی را برای برنامه‌ریزی ایجاد کرده است. برای رفع چنین مشکلاتی استفاده از آمار ایستگاه‌های مشابه از لحاظ جغرافیایی را توصیه کرده‌اند (خیرفام و همکاران، ۱۳۹۱ب)، که البته مجاورت جغرافیایی نیز دلیل بر تشابه بین حوضه‌ها نمی‌تواند باشد. روش‌های متعددی برای تعیین تشابه و همگنی حوضه‌ها ارائه شده‌اند که می‌توان از منحنی‌های اندرو<sup>۱</sup> (Moustafa, 2011)، تحلیل خوشه‌ای (Wilson, 2011) و تحلیل تابع تشخیص<sup>۲</sup> (Zhang et al., 2011; Detenbeck et al., 2005, Carroll et al., 2009; Lin and Chen, 2009) نام برد. بررسی کارایی مشخصات اقلیمی و فیزیکی آبخیز به منظور تخمین مقدار رسوب خروجی از حوضه‌های دارای ایستگاه رسوب‌سنجی و تعمیم روابط بدست آمده بین ویژگی‌های فیزیکی و اقلیمی و مقدار رسوب خروجی از حوضه، به حوضه‌های مشابه و همگن می‌تواند راه‌کاری برای تخمین رسوب در حوضه‌های فاقد ایستگاه باشد (Riggs, 1973; Caratti et al., 2004; Tramblay et al., 2010). از طرفی تعدد عوامل فیزیکی و اقلیمی موثر بر میزان رسوب تولیدی منجر به عدم امکان برآورد و تعیین تمام عوامل شده و یا توجیه علمی و اقتصادی ندارد، بنابراین باید با کاربرد روش‌هایی تعداد فاکتورهای موثر در تولید رسوب را به تعدادی کاهش داد که بیش‌ترین تاثیر را در تولید رسوب دارند (Lin and Wang, 2006; Noori et al., 2011). روش آزمون گاما<sup>۳</sup> از جمله روش‌هایی می‌باشد که امروزه به‌منظور تعیین عوامل اصلی مؤثر بر متغیر وابسته مورد توجه قرار گرفته است (Corcoran et al., 2003; Moghaddamnia et al., 2009). در نهایت با استفاده از روش‌های کاهش تعداد عوامل موثر در تولید رسوب و

با این حال حمل رسوب توسط رودخانه‌ها مشکلاتی از قبیل پر شدن مخازن سدها، تغییر در مسیر رودخانه‌ها یا کاهش ظرفیت حمل آن‌ها، تخریب تأسیسات مسیر رودخانه‌ها، کاهش کیفیت آب، اختلال در آبی‌پروری را در پی دارد (Cigizoglu, 2004). به این دلیل تخمین صحیح رسوبات خروجی از حوزه‌های آبخیز اهمیت بسزایی در مدیریت منابع دارد (Ki i, 2010). با وجود اهمیت تخمین صحیح رسوب خروجی آبخیز، به دلیل گستردگی و تعداد بالای آبخیزها استقرار ایستگاه‌های رسوب‌سنجی در تمام حوضه‌ها میسر نمی‌باشد و توجیه اقتصادی ندارد. بنابراین برآورد مقادیر رسوب خروجی از آبخیزها اغلب با استفاده از روش منحنی سنجه که مبتنی بر رابطه تجربی بین غلظت رسوب و دبی می‌باشد، صورت می‌گیرد (Tramblay et al., 2010). پژوهش‌ها نشان می‌دهد مدل‌های مبتنی بر منحنی سنجه برآورد بیش‌تری از مقدار رسوب دارند (خیرفام، ۱۳۹۱؛ Asselman, 2000) و با اصلاح انحراف مدل‌ها نیز بیش‌تخمینی وجود خواهد داشت (خیرفام و صادقی، ۱۳۹۱؛ Walling and Webb, 1988). بعضی پژوهش‌ها رابطه‌ی ضعیف بین رسوب و دبی خروجی آبخیزها را نشان داده‌اند. Tramblay و همکاران (۲۰۰۸) با بررسی رابطه‌ی بین دبی رودخانه و رسوب معلق خروجی ۲۰۸ ایستگاه شمال آمریکا، اعلام نمود که تنها در ۹۸ رودخانه رابطه‌ی مناسبی بین دبی و رسوب وجود دارد که حاکی از عدم جامع بودن روش منحنی سنجه رسوب جهت تخمین میزان رسوب معلق حوضه‌ها می‌باشد. از سوی دیگر پژوهش‌های متعددی نشان دادند که بین ویژگی‌های فیزیکی حوضه‌ها از قبیل توپوگرافی، زمین‌شناسی و کاربری اراضی و هم‌چنین ویژگی‌های اقلیمی از قبیل شدت و میزان بارندگی با میزان بارمعلق (Bray and Xie, 1993; Ludwig and Probs, 1998; Restrepo et al., 2006) میانگین غلظت رسوب معلق (Jarvie et al., 2002; Siakeu et al., 2004; Syvitski and Milliman, 2007; Nadal-Romero et al., 2011) یا حداکثر غلظت

<sup>1</sup> Andrews Curves

<sup>2</sup> Discriminant Function Analysis

<sup>3</sup> Gamma Test

(۲۰۱۲)، از خصوصیات حوزه آبخیز برای استخراج مدل‌های رگرسیونی چندمتغیره رسوب معلق سالانه در شرق آمریکا استفاده نمودند. نتایج نشان داد که متوسط رسوب سالانه با متغیرهای مساحت حوزه آبخیز، الگوی کاربری اراضی، بارش فصلی، خصوصیات خاک و اصلاح هیدرولوژیکی ارتباط دارد. Heng و Suetsugi (۲۰۱۴)، از خصوصیات فیزیکی حوزه آبخیز برای انتقال منحنی سنجه رسوب از چهار شاخص متفاوت همگن‌سازی در ۱۶ زیرحوزه آبخیز واقع در حوزه آبخیز مکونگ در جنوب شرقی آسیا استفاده نمودند. نتایج نشان داد که روش استفاده از فاصله اقلیدسی که توسط مقیاس چندوجهی، خصوصیات حوزه آبخیز (توپوگرافی، مورفولوژی، هیدرولوژی، پوشش زمین، خاک) را نرمال‌سازی می‌کند جواب مناسب‌تری نسبت به روش‌های دیگر ارائه می‌نماید. استفاده از آزمون گاما نیز توسط پژوهش‌گران متعددی به منظور تعیین موثرترین عوامل فیزیکی هیدرولوژیکی و اقلیمی موثر در پاسخگویی هیدرولوژیکی آبخیزها استفاده نموده‌اند. Corcoran و همکاران (۲۰۰۳) برای پیش‌بینی دبی، Moghaddamnia و همکاران (۲۰۰۹) برای پیش‌بینی تبخیر و Noori و همکاران (۲۰۱۱) به منظور کاهش تعداد عوامل موثر بر پاسخ‌های هیدرولوژیکی حوضه جهت پیش‌بینی جریان‌های ماهانه استفاده کرده‌اند. در ایران نیز وفاخواه (۱۳۸۵) با استفاده از تحلیل عاملی به منظور کاهش تعداد پارامترهای تاثیرگذار در تولید رسوب و تحلیل خوشه‌ای برای همگن‌بندی حوضه‌های ساحلی جنوب دریای خزر استفاده کرد که حوضه‌ها در دو گروه همگن طبقه‌بندی شدند. پژوهش‌های صورت گرفته با استفاده از روش‌های مذکور بیشتر در زمینه‌ی تخمین جریان‌های رودخانه‌ای بوده و به عنوان ابزار کار استفاده گردیده است. ولی به بررسی کارایی و انتخاب و تعیین روش برتر پرداخته نشده است. در این پژوهش سعی بر این است تا با استفاده روش آزمون گاما از بین عوامل مؤثر بر تولید رسوب معلق،

هم‌چنین با قرار دادن آبخیزهایی که رفتار مشابهی در تولید رسوب دارند در گروه‌های همگن، می‌توان برای هر گروه مدلی ارائه داده و با انتخاب بهترین روش، برای آبخیزهای فاقد آمار و ایستگاه رسوب‌سنجی، با استفاده از عوامل موثر بر تولید رسوب، میزان رسوب هر آبخیز را تخمین زد. در این راستا Caratti و همکاران (۲۰۰۴) بر اساس داده‌های اقلیمی و مورفومتریک و با استفاده از روش‌های تحلیل تابع تشخیص و خوشه‌بندی اقدام به کلاس‌بندی زیرحوضه‌های آبخیز رودخانه‌ی کلمبیا بر اساس شناسه‌های محیط زیستی نمودند و اعلام نمودند که نتایج حاصل از دو روش مذکور اختلاف معنی‌داری با هم ندارند و این روش‌ها در انتخاب عوامل موثر به منظور کلاس‌بندی زیرحوضه‌ها در گروه‌های همگن، به صورت تصادفی عمل کرده و باید کاربرد چنین روش‌هایی با احتیاط صورت گیرد. Lin و Wang (۲۰۰۶) به منظور بررسی کارایی روش‌های تحلیل تابع تشخیص و تحلیل خوشه‌ای<sup>۱</sup> اقدام به همگن‌بندی حوضه‌های جنوب تایوان به منظور تخمین جریان‌های حداقل نمودند که نتایج حاکی از کارایی بالای روش تحلیل تابع تشخیص نسبت به روش تحلیل خوشه‌ای بوده و با ترکیب این دو روش نتایج بهتری حاصل شد. Tramblay و همکاران (۲۰۱۰)، به‌منظور تخمین حداکثر غلظت رسوب سالانه ۷۲ رودخانه در آمریکا و کانادا، با استفاده از مدل‌های منطقه-ای که برای پیش‌بینی سیلاب تهیه شده بودند و با قرار دادن حوضه‌ها در گروه‌های همگن، یک رابطه رگرسیونی چند متغیره بین ویژگی‌های فیزیکی و اقلیمی حوضه با مقدار غلظت رسوب معلق برقرار نمودند. نتایج نشان داد که تخمین رسوب با استفاده از مدل‌های منطقه‌ای کارایی بیشتری نسبت به یک مدل رگرسیونی کلی برای تمام حوضه‌ها دارد. درصد رس خاک، درصد پوشش جنگلی و شدت بارندگی را تاثیرگذارترین عوامل در میزان تولید رسوب معلق گزارش نمودند. Roman و همکاران

<sup>1</sup>Cluster Analysis

شده در شکل ۴) برای استفاده در پژوهش حاضر انتخاب گردید.

#### محاسبه رسوب معلق سالانه

با توجه به این که رسوب برداشتی در ایستگاه‌های رسوب‌سنجی به صورت رخدادی می‌باشد، ابتدا رسوب معلق برداشتی و دبی متناظر آن‌ها در طول دوره‌ی آماری مرتب گردید و از برازش داده‌های لگاریتمی دبی و رسوب (وفاخواه، ۱۳۸۵) بر مبنای رابطه‌ی (۱) (Sadeghi and Mahdavi, 2004) رابطه‌ی رگرسیونی برقرار شد (De Vente et al., 2007). با قرار دادن دبی‌های روزانه در رابطه‌ی تهیه شده، رسوب معلق روزانه به دست آمده و رسوب سالانه (Vanmaercke et al., 2011) نیز محاسبه گردید.

$$Q_s = aQ_w^b \quad (1)$$

#### عوامل موثر بر مقدار رسوب معلق سالانه

در هر یک از حوضه‌های بالادست ایستگاه‌های رسوب‌سنجی مورد بررسی، ۱۴ عامل موثر بر مقدار رسوب سالانه شامل: مساحت، محیط، ارتفاع متوسط وزنی حوضه، طول و شیب آبراهه‌ی اصلی، تراکم زهکشی، قطر دایره هم‌سطح، ضریب شکل هورتون، ضریب گراولوس و زمان تمرکز با روش کریچ با استفاده از نقشه‌های توپوگرافی ۱:۵۰۰۰۰ به عنوان ویژگی‌های فیزیکی حوضه و متوسط بارندگی سالانه، دبی حداکثر لحظه‌ای با دوره برگشت دو ساله (متوسط دبی حداکثر لحظه‌ای) و دبی حداکثر روزانه با دوره بازگشت دو ساله (متوسط دبی حداکثر روزانه) به سبب سادگی در استخراج، قابلیت دسترسی در سازمان‌های مربوطه، جامع بودن اطلاعات و اندازه‌گیری و یا موجودیت اطلاعات در اکثر آبخیزهای کشور و همچنین بر اساس مطالعات صورت گرفته و روابط و روش‌های رایج در منابع معتبر تعیین و محاسبه شدند (Chow et al., 2007).

متغیرهای مستقل انتخاب گردد و سپس با استفاده از سه روش منحنی‌های اندرو، تحلیل خوشه‌ای و تحلیل تابع تشخیص به عنوان روش‌های گرافیکی و غیرعددی اقدام به تعیین حوزه‌های آبخیز همگن واقع در جنوب و جنوب شرقی دریای خزر شده و براساس آن مدل‌های برآورد رسوب معلق در هر یک از گروه‌های همگن تهیه شود. در نهایت با مقایسه‌ی نتایج هر یک از مدل‌ها با استفاده از معیارهای آماری به عنوان روشی آماری و عددی، بهترین روش همگن‌سازی حوزه‌های آبخیز به منظور برآورد رسوب معلق انتخاب گردید.

#### مواد و روش‌ها

##### انتخاب ایستگاه‌های مناسب

برای انتخاب ایستگاه‌های رسوب‌سنجی مناسب، آمار دبی و رسوب معلق تمام ایستگاه‌های موجود در منطقه جنوب و جنوب شرق دریای خزر از شرکت تحقیقات و منابع آب (تماب) تهیه گردید. سپس برای تمام ایستگاه‌های موجود داده‌های رسوب معلق و دبی متناظر آن‌ها و دبی روزانه از اولین سال احداث ایستگاه رسوب‌سنجی و دبی‌سنجی تا آخرین داده‌های گزارش شده مرتب شد. به طوری که طول دوره‌ی آماری ۲۵ ساله طی سال‌های ۶۱-۱۳۶۰ تا ۸۶-۱۳۸۵ به عنوان طول دوره‌ی آماری مشترک در همه‌ی ایستگاه‌ها در نظر گرفته شد. در این دوره‌ی آماری ۲۵ ساله تعداد نمونه‌ی رسوب معلق و دبی اندازه‌گیری متناظر آن از حداقل ۵۲ تا حداکثر ۶۱۵۰ و به طور متوسط ۶۸۴ تعداد نمونه می‌باشد. برای انتخاب ایستگاه‌ها، ایستگاه‌هایی که دارای آمار ۲۵ ساله، تعداد نمونه‌ی مناسب و جدید بودند و شرط عدم وجود بند یا سد انحرافی در بالادست و عدم قرارگیری ایستگاه در مناطق دشتی را داشتند انتخاب گردیدند. با توجه به محدودیت‌های مذکور، در مجموع داده‌های ۴۲ ایستگاه رسوب‌سنجی (مندرج در جدول ۱ و نمایش داده

جدول ۱. شماره، کد و ایستگاه‌های انتخاب شده در منطقه مورد مطالعه

شماره	کد	ایستگاه	شماره	کد	ایستگاه	شماره	کد	ایستگاه
۱	۱۲-۰۰۱	دوغ	۱۵	۱۲-۰۴۵	شصت کلاته	۲۹	۱۴-۰۰۱	شیرگاه
۲	۱۲-۰۰۵	تمر	۱۶	۱۲-۰۴۹	پل جاده	۳۰	۱۴-۰۰۵	کسیلیان
۳	۱۲-۰۰۷	گالیکش	۱۷	۱۲-۰۵۳	وطنا	۳۱	۱۴-۰۰۷	کیکلا
۴	۱۲-۰۰۹	قلی تپه	۱۸	۱۲-۰۷۱	زرینگل	۳۲	۱۴-۰۱۱	قرآن تالار
۵	۱۲-۰۱۱	گنبد	۱۹	۱۲-۰۸۳	محمدآباد	۳۳	۱۴-۰۱۷	کشتارگاه
۶	۱۲-۰۱۳	لزوره	۲۰	۱۲-۰۸۵	پل اردوگاه	۳۴	۱۵-۰۱۵	رزن
۷	۱۲-۰۱۵	پس پشته	۲۱	۱۲-۰۹۷	سیاه آب	۳۵	۱۵-۰۱۷	کره سنگ
۸	۱۲-۰۱۷	نوده	۲۲	۱۳-۰۰۵	سفیدچاه	۳۶	۱۶-۰۰۳	تنگه لایوچ
۹	۱۲-۰۱۹	ارازکوسه	۲۳	۱۳-۰۰۶	نوذراآباد	۳۷	۱۶-۰۰۹	خیرود
۱۰	۱۲-۰۲۱	رامیان	۲۴	۱۳-۰۱۳	آبلو	۳۸	۱۶-۰۱۱	نوشهر
۱۱	۱۲-۰۲۳	قزاقلی	۲۵	۱۳-۰۱۹	سلیمان تنگه	۳۹	۱۶-۰۴۱	هرات‌بر
۱۲	۱۲-۰۳۳	تقی‌آباد	۲۶	۱۳-۰۲۳	ورند	۴۰	۱۶-۰۴۹	گانگسر
۱۳	۱۲-۰۳۵	امامزاده	۲۷	۱۳-۰۲۳	ریگ چشمه	۴۱	۱۶-۰۵۱	رامسر
۱۴	۱۲-۰۴۳	نهارخوران	۲۸	۱۳-۰۲۹	کردخیل	۴۲	۱۶-۰۸۹	دیناسرا

## آزمون گاما

آزمون گاما یک ابزار مدل‌سازی غیر خطی است که به کمک آن می‌توان ترکیب مناسبی از متغیرهای ورودی برای مدل‌سازی داده‌های خروجی و تهیه یک مدل هموار را بررسی نمود (Moghaddamnia et al., 2009). آزمون گاما اولین بار توسط (Koncar, 1997) و Agalbjorn و همکاران (1997) گزارش شده و بعدها توسط پژوهش‌گران دیگری هم‌چون Tsui و همکاران (2002) و Durrant (2001) بکار برده شده است. مدل آزمون گاما توسط Durrant (2001) به صورت بسته‌ی نرم‌افزاری معرفی گردید. هنگامی که عوامل موثر بر پدیده‌ای به‌طور قابل ملاحظه‌ای زیاد باشد با استفاده از آزمون گاما می‌توان ترتیب میزان اهمیت متغیرهای ورودی و بهترین ترکیب از میان تمام ترکیب‌های ممکن را به‌دست آورد. به‌منظور تهیه‌ی مدل هموار، از آزمون M برای انتخاب تعداد داده‌های ورودی مناسب استفاده می‌شود (Moghaddamnia et al., 2009).

## روش منحنی‌های اندرو

منحنی اندرو توسط Andrews (1971) برای تشخیص بقایای فسیل انسان از میمون بسط داده شد. منحنی اندرو

یک روش غیر عددی مصورسازی داده‌های پیوسته چند متغیره می‌باشد که به آشکارسازی ساختارهای پنهان مجموعه داده‌های نسبتاً کوچک کمک می‌کند. این ساختارها ممکن است به صورت لایه‌های خروجی، خوشه‌ها و میزان همبستگی باشد (Moustafa, 2011). منحنی اندرو هر نقطه را در یک محیط سه بعدی به صورت دو بعدی و سه بعدی نشان داده و مطابق با رابطه ۲ از سری‌های فوریه تبعیت می‌کند (Horhota and Atiken, 2006).

$$y_{it} = \frac{x_{i1}}{\sqrt{2}} + x_{i2} \cos(1t) + x_{i3} \sin(1t) + x_{i4} \cos(2t) + x_{i5} \sin(2t) + \dots \quad (2)$$

که در آن  $x_{i1}$ ،  $x_{i2}$ ،  $x_{i3}$  و  $x_{i5}$  هر یک از عوامل استفاده شده برای نمایش ویژگی‌های حوضه را نمایش می‌دهند که به ترتیب اهمیتی که دارا می‌باشند بیش‌ترین همبستگی را با متغیر وابسته دارند (Moustafa, 2011). فاصله اقلیدوسی بین صفات متغیرها، اختلاف بین دو منحنی را نشان می‌دهد. اگر متغیرهای n بعدی در فضای اقلیدوسی نزدیک هم‌دیگر باشند، منحنی‌های مشابهی دارند (وفاخواه، ۱۳۸۵).

## روش تحلیل خوشه‌ای

تعیین گروه‌های همگن بر اساس عوامل انتخابی از طریق هر یک از روش‌های کاهش تعداد متغیرها، برای هر یک از روش‌های مذکور، به‌صورت جداگانه روش‌های منحنی اندرو، تحلیل خوشه‌ای و تحلیل تابع تشخیص استفاده گردید.

### معیارهای آماری

به‌منظور بررسی کارایی و میزان دقت و خطای هر یک از روش‌های به‌کار برده شده در همگن‌سازی آبخیزها در مناطق مورد مطالعه، از معیارهای خطای نسبی (رابطه ۳) (خیرفام و همکاران، ۱۳۹۱ الف)، انحراف نسبی مدل (رابطه ۴) (خیرفام و صادقی، ۱۳۹۱) و ریشه نسبی مجذور میانگین مربعات خطا (رابطه ۵) (Tramblay et al., 2010) استفاده گردید.

$$RE = \left| \frac{Y_o - Y_m}{Y_o} \right| \times 100 \quad (3)$$

$$RBIAS = \frac{1}{N} \left( \frac{Y_o - Y_m}{Y_o} \right) \quad (4)$$

$$RRMSE = \sqrt{\frac{\sum_{i=1}^n (Y_o - Y_m)^2}{n}} \times 100 \quad (5)$$

در آن،  $Y_o$  داده‌های مشاهداتی رسوب متوسط سالانه،  $Y_m$  داده‌های تخمینی رسوب متوسط سالانه توسط مدل و  $n$  تعداد مشاهدات رسوب متوسط سالانه است. هر چه مقادیر معیارهای آماری  $RE$ ،  $RBIAS$  و  $RRMSE$  کم‌تر باشد نشان از کارایی بالای معیارهای مذکور خواهد بود.

### نتایج و بحث

نتایج کاهش تعداد عوامل با استفاده از آزمون گام‌عوامل زیادی در میزان تولید رسوب موثر و دخیل می‌باشند و استفاده از تمام این عوامل در تخمین رسوب علاوه بر کاهش دقت مدل ارائه شده برای حوضه، زمان‌بر بوده و همچنین کاربرد این مدل در حوضه‌های مشابه که فاقد اطلاعات تمام عوامل موثر می‌باشند غیرممکن خواهد بود. با استفاده از آزمون گاما می‌توان تا حد زیادی حجم

اصطلاح تحلیل خوشه‌ای اولین بار توسط Tryon در سال ۱۹۳۹ برای گروه‌بندی اشیائی که با هم شبیه بود استفاده شد. هدف از خوشه‌بندی تقسیم داده‌ها در گروه‌های مشابه می‌باشد، به‌طوری که مشاهدات هر گروه بیش‌ترین شباهت و مشاهدات گروه‌های مختلف کم‌ترین شباهت را با هم داشته باشند (Jobson, 1992; Lin and Wang, 2006). الگوریتم‌های زیادی برای تحلیل خوشه‌ای پیشنهاد شده است که به دو گروه روش‌های سلسله مراتبی و تفکیک‌کننده‌ی تکراری تقسیم می‌شوند. بیش‌تر پژوهشگران روش سلسله مراتبی را مناسب‌تر می‌دانند (Ramos, 2001). در روش سلسله مراتبی با دو روش تجمعی و تقسیم، گروه‌ها تشکیل می‌شوند. در روش تجمعی ابتدا هر فرد یک گروه مجزا را تشکیل می‌دهد و سپس گروه‌های نزدیک به هم به تدریج ترکیب می‌شوند تا در نهایت کلیه افراد در یک گروه واقع می‌شوند. در روش تقسیم نیز عکس روش تجمعی گروه‌بندی انجام می‌گیرد. خروجی روش سلسله مراتبی به صورت دندوگرامی بوده و روابط طبقاتی چند فرد بر اساس شباهت آن‌ها نشان داده می‌شود (Tryon, 1939).

### روش تحلیل تابع تشخیص

تحلیل تابع تشخیص روش تحلیل آماری چند متغیره است که در آن مجموعه داده‌ها حاوی  $X$  متغیر است که با استفاده از ترکیب خطی از تجزیه و تحلیل متغیرها به تعدادی از گروه‌های از پیش تعریف شده، تقسیم می‌شوند (Carroll et al., 2009). تحلیل تابع تشخیص این امکان را بوجود می‌آورد تا افراد و متغیرها به صورت فضایی در یک محیط دو بعدی از افراد گروه دیگر قابل تمیز باشند (Wilson, 2002).

در پژوهش حاضر پس از تعیین رسوب سالانه هر یک از ایستگاه‌های مورد مطالعه، با استفاده از روش آزمون گاما از بین ۱۴ عامل موثر موجود بر تولید رسوب، تاثیرگذارترین عوامل انتخاب گردید. سپس به منظور

فاکتور شکل حوضه بود که با حذف هر کدام از آن‌ها از ورودی‌های مدل، مقدار ضریب گامای بیش‌تری را داشت. در جدول ۲ مقادیر ضریب گاما با حذف هر کدام از عوامل ارائه شده است.

پس از کاهش تعداد عوامل توسط آزمون گاما، اقدام به همگن‌بندی حوزه‌های آبخیز بر اساس عوامل انتخاب شده و با استفاده از سه روش تحلیل خوشه‌ای، تحلیل تابع تشخیص و منحنی‌های اندرو گردید. نتایج همگن‌بندی حوزه‌های آبخیز جنوب و جنوب شرقی دریای خزر بر اساس تولید رسوب و تعیین میزان دقت هر یک از روش‌های مذکور و مقایسه‌ی بین آن‌ها به تشریح ذیل ارائه گردیده است.

کار تهیه‌ی مدل و تعداد داده‌های ورودی مدل را قبل از تهیه و توسعه‌ی آن کاهش داد (Tsui et al., 2002). در واقع، آزمون گاما قادر است بهترین میانگین مربع خطای به‌دست آمده توسط مدل‌های غیرخطی را ارائه دهد. در این مطالعه، در هر مرحله از اجرای مدل آزمون گاما یکی از داده‌های ورودی با جایگزینی دوباره به‌منظور ارزیابی تاثیر آن‌ها بر میزان رسوب خروجی حوضه وارد مدل نشد. در نهایت با خروج هر یک از مدل‌ها، ضریب گامای مدل مورد بررسی قرار گرفت و با خروج هر یک از عواملی که بیش‌ترین مقدار ضریب گاما را داشتند به عنوان عوامل موثر در تولید رسوب انتخاب گردیدند (Noori et al., 2010). عوامل مذکور شامل قطر هم‌سطح حوضه، دبی حداکثر روزانه با دوره بازگشت دو ساله، شیب رودخانه، طول حوزه‌ی آبخیز به کیلومتر و ضریب

جدول ۲. مقادیر ضریب گاما در استفاده از آزمون گاما در طی فرآیند انتخاب عوامل مؤثر بر تولید رسوب

مقدار عدد گاما ( $V_{ratio}$ )	متغیر ورودی	مقدار عدد گاما ( $V_{ratio}$ )	متغیر ورودی
۰/۱۳۱۷	All-Dd	۰/۱۳۱۶	همه متغیرها
۰/۱۳۱۲	All-Pw	۰/۶۱۷۴	All-Qp
۰/۱۳۱۲	All-Pr	۰/۱۳۵۶	All-Di
۰/۱۳۱۱	All-Area	۰/۱۳۵۴	All-Sr
۰/۱۳۱۱	All-Lr	۰/۱۳۳۲	All-Lw
۰/۱۲۴۵	All-Qm	۰/۱۳۲۷	All-Cc
۰/۱۲۰۹	All-Sw	۰/۱۳۲۶	All-H
		۰/۱۳۲۵	All-Tc

\*\* متغیرها و مقادیری که زیر آن‌ها خط کشیده شده عوامل اصلی انتخاب شده و مقادیر متناظر آن‌ها می‌باشند.

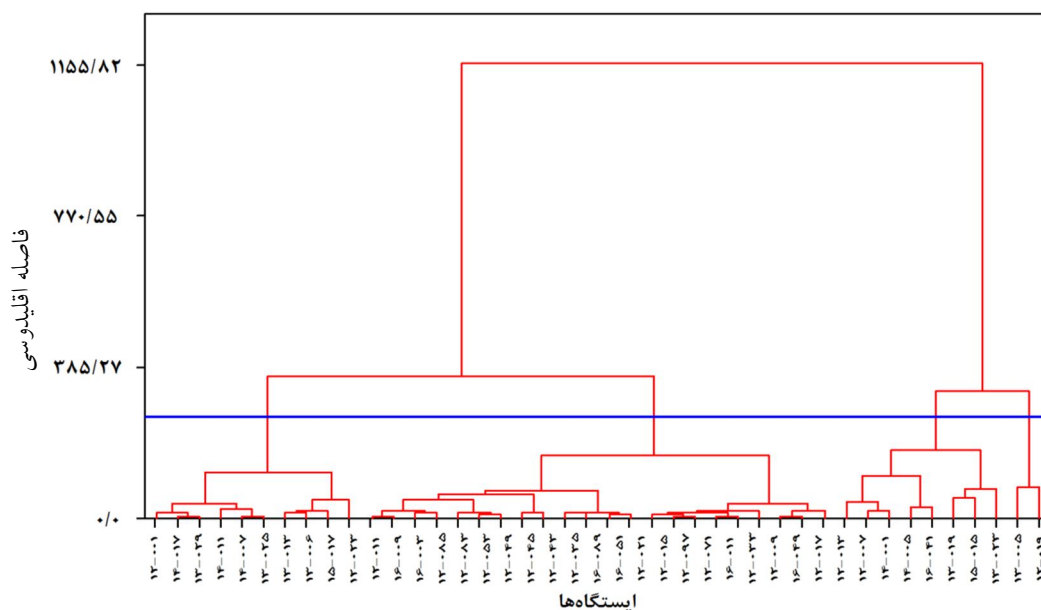
\*\* Qp (دبی حداکثر لحظه‌ای با دوره بازگشت دو ساله بر حسب متر مکعب بر ثانیه)، Di (قطر هم‌سطح حوضه بر حسب کیلومتر)، Sr (شیب رودخانه بر حسب درصد)، Lw (طول حوزه‌ی آبخیز بر حسب کیلومتر)، Cc (ضریب فاکتور شکل حوضه)، H (ارتفاع متوسط وزنی حوضه بر حسب متر)، Tc (زمان تمرکز حوضه بر حسب ساعت)، Dd (تراکم زهکشی حوضه بر حسب کیلومتر بر کیلومتر مربع)، Pw (محیط حوضه بر حسب کیلومتر)، Pr (بارش بر حسب میلی‌متر)، Area (مساحت حوضه بر حسب کیلومتر مربع)، Lr (طول آبراهه‌ی اصلی بر حسب متر)، Qm (دبی حداکثر روزانه با دوره بازگشت دو ساله بر حسب متر مکعب بر ثانیه)، Sw (شیب متوسط حوضه بر حسب درصد).

همگن‌بندی حوزه‌های آبخیز با استفاده از تحلیل خوشه‌ای پنج عاملی که توسط آزمون گاما به‌عنوان مؤثرترین عوامل دخیل در تولید رسوب انتخاب شده‌اند برای همگن‌سازی آبخیزهای مورد مطالعه توسط روش تحلیل خوشه‌ای مورد استفاده قرار گرفتند. در مرحله‌ی دوم برای تقسیم‌بندی آبخیزهای متشابه بر اساس شاخص تشابه از الگوریتم ارتباطی Ward که تمایل به تشکیل خوشه‌های گروهی مساوی دارد استفاده گردید (Nathan and McMahon, 1990; Ouarda et al., 2006). در مرحله‌ی نهایی که انتخاب فاصله‌ی مناسب برای انقطاع خوشه‌ها برای تعیین گروه‌های همگن می‌باشد پژوهشگران روش‌های متعددی را بر اساس هدف مطالعه و اندازه‌ی گروه‌ها معرفی کرده‌اند. با این حال روش نظر کارشناسی و روش‌های سعی و خطا بر اساس مدل‌سازی رگرسیونی، شبکه عصبی و منطق فازی بیش‌تر مورد توجه قرار گرفته است (Raju and Kumar, 2011). در این پژوهش نیز بر

همگن‌بندی حوزه‌های آبخیز با استفاده از تحلیل خوشه‌ای پنج عاملی که توسط آزمون گاما به‌عنوان مؤثرترین عوامل دخیل در تولید رسوب انتخاب شده‌اند برای همگن‌سازی آبخیزهای مورد مطالعه توسط روش تحلیل خوشه‌ای مورد استفاده قرار گرفتند. در مرحله‌ی دوم برای تقسیم‌بندی آبخیزهای متشابه بر اساس شاخص تشابه از الگوریتم ارتباطی Ward که تمایل به تشکیل خوشه‌های گروهی مساوی دارد استفاده گردید (Nathan and

اساس روش سعی و خطا از چندین نقطه، انقطاع صورت گرفت و پس از تعیین گروه‌های همگن با استفاده از مدل‌سازی رگرسیونی و تعیین دقت و خطای متوسط هر یک از مدل‌های ارائه شده برای گروه‌های همگن، در نهایت بهترین فاصله‌ی انقطاع خوشه‌ها بدست آمد

اساس روش سعی و خطا از چندین نقطه، انقطاع صورت گرفت و پس از تعیین گروه‌های همگن با استفاده از مدل‌سازی رگرسیونی و تعیین دقت و خطای متوسط هر یک از مدل‌های ارائه شده برای گروه‌های همگن، در نهایت بهترین فاصله‌ی انقطاع خوشه‌ها بدست آمد



شکل ۱. نحوه جدایش هر یک از حوضه‌ها در گروه‌های همگن با استفاده از تحلیل خوشه‌ای

تاثیرگذار در تولید رسوب به عنوان متغیرهای مستقل اقدام به مدل‌سازی رگرسیونی شد (Trambly et al., 2010). هم‌چنین یک مدل هم برای کل ایستگاه‌ها بدون جدایش ایستگاه‌ها در گروه‌های همگن ساخته شد. قبل از ساخت مدل، به طور تصادفی و با رعایت توزیع مناسب ۷۰ درصد داده‌ها برای مدل‌سازی و ۳۰ درصد برای اعتبارسنجی مدل انتخاب شدند (صادقی و همکاران، ۱۳۸۴) مدل‌های ساخته شده برای هر یک از گروه‌ها در جدول ۳ ارائه گردیده است. در جدول ۴ نیز میزان خطا و کارایی هر یک از روش‌ها با استفاده از معیارهای آماری ارائه شد. پس از برقراری رابطه‌ی رگرسیونی بین مقدار رسوب متوسط سالانه (تن بر سال) به عنوان متغیر وابسته و ۵ عامل انتخاب شده توسط آزمون گاما به عنوان متغیرهای مستقل در هر یک از گروه‌های همگن، نتایج

آب‌خیزهای گروه‌بندی شده توسط روش تحلیل خوشه‌ای به ترتیب در گروه‌های همگن ذیل قرار گرفتند: گروه همگن ۱: کسبیلان، رزن، هرات‌بر، دوغ، تمر، ارازکوسه، سفیدچاه، سلیمان تنگه، ورنند و شیرگاه. گروه همگن ۲: گالیکش، قلی تپه، لزوره، پس پشته، رامیان، نوده، تقی آباد، امام زاده، نهارخوران، پل جاده، وطناء، محمد آباد، شصت کلات، پل اردوگاه، رامسر، تنگه لاویج، سیاه آب، خیرود، نوشهر، دیناسرا، گانگسر و زرینگل. گروه همگن ۳: نوذر آباد، آبدلی، ریگ چشمه، کیکلا، کره سنگ، قرآن تالار، قزاقلی و گنبد. گروه همگن ۴: کشتارگاه، و کردخیل. در هر یک از گروه‌ها، برای تهیه‌ی مدل به‌منظور تخمین متوسط رسوب سالانه خروجی از ایستگاه‌ها به عنوان متغیر وابسته و با استفاده از ۵ متغیر انتخاب شده توسط آزمون گاما به عنوان عامل‌های اصلی



بنابراین در جمع‌بندی می‌توان این‌گونه بیان کرد که روش تحلیل خوشه‌ای به دلیل ماهیت تا حدودی تصادفی بودن آن (Caratti *et al.*, 2004) توانایی بالایی در تفکیک ایستگاه‌ها بر اساس تفاوت‌ها و شباهت‌های جزئی را ندارد که در نهایت به تقسیم‌بندی کلی‌تر و با دقت کم‌تر منجر خواهد شد. نتایج کارایی سنجی مدل‌سازی رگرسیونی و مقایسه مقادیر RE، RBIAS و RRMSE نیز نشان می‌دهد که در گروه همگن ۱ و ۲ هم با این‌که مدل‌ها معنی‌دار بوده ولی خطای مدل‌ها بالا می‌باشد که ناهمگونی و تفاوت زیاد بین مقادیر متغیرهای مستقل دلیل بر این امر می‌تواند باشد (کیانی هرچگانی و صادقی، ۱۳۹۲؛ Sadeghi and Saeidi, 2010). هرچند هر دو مدل نسبت به مدل کلی نتایج مطلوب‌تر و قابل قبول ارائه داده‌اند.

بررسی مشخصات فیزیکی، اقلیمی و هیدرولوژیکی حوضه‌های بالادست این ایستگاه‌ها حاکی از آن بود که فقط در گروه‌های همگن ۱ و ۲ آبخیزها از نظر مقادیر دبی حداکثر لحظه‌ای با دوره بازگشت دو ساله و ضریب فاکتور شکل حوضه تا حدودی مشابه به هم می‌باشند که دلیل قرار گرفتن این ایستگاه‌ها در یک گروه همگن با روش تحلیل خوشه‌ای نیز همین امر می‌تواند باشد. در حالی‌که در دو گروه همگن دیگر هیچ نوع تشابه محسوسی در متغیرهای فیزیکی، اقلیمی و هیدرولوژیکی مشاهده نمی‌گردد. به‌گونه‌ای که برای دو گروه همگن ۳ و ۴ مدل معنی‌داری ارائه نشد. هم‌چنین بررسی موقعیت جغرافیایی ایستگاه‌ها نشان می‌دهد که در این گروه‌بندی ایستگاه‌های قرار گرفته در یک گروه همگن از نزدیکی جغرافیایی برخوردار بودند. مطمئناً تشابه جغرافیایی باعث تشابه در متغیرهای اقلیمی را در پی خواهد داشت.

جدول ۳. مدل‌های تهیه شده برای هر یک از گروه‌های همگن انتخاب شده با تحلیل خوشه‌ای و برای کل منطقه

مدل‌های ارائه شده برای هر یک از گروه‌ها	
گروه ۱	$104.01 Qp + 54.098 Di + 34.714 Sr + 145.729 Lw - 2288 Cc - 9359.821$
گروه ۲	$-4.995Qp - 24.87 Di - 61.97 Sr + 22.891 Lw - 236.855 Cc + 1250.726$
گروه ۳	مدل معنی‌داری ارائه نشد
گروه ۴	کافی نبودن تعداد ایستگاه برای مدل‌سازی
مدل کلی	مدل معنی‌داری ارائه نشد

\* Qp (دبی حداکثر لحظه‌ای با دوره بازگشت دو ساله بر حسب متر مکعب بر ثانیه)، Di (قطر هم‌سطح حوضه بر حسب کیلومتر)، Sf (شیب رودخانه بر حسب کیلومتر)، Lw (طول حوضه آبخیز بر حسب کیلومتر) و Cc (ضریب فاکتور شکل حوضه).

جدول ۴. نتایج میزان دقت و خطای هر یک از مدل‌های ارائه شده در مرحله واسنجی و اعتبارسنجی

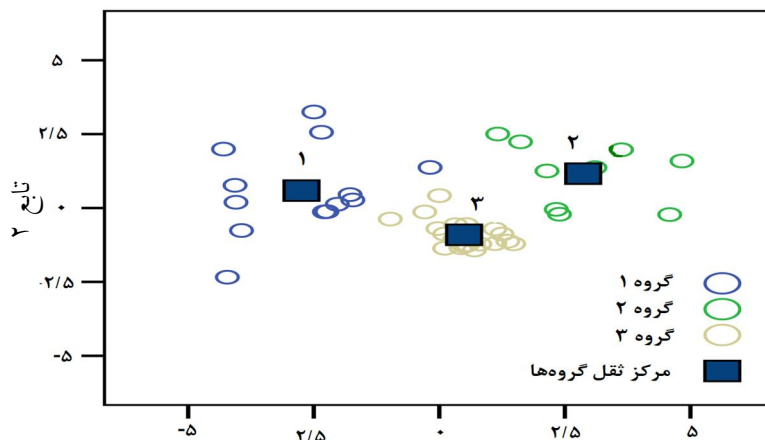
r	RRMSE (%)		RBIAS		RE (%)		
	اعتبارسنجی	واسنجی	اعتبارسنجی	واسنجی	اعتبارسنجی	واسنجی	
۰/۷۷	۱۱۹/۰۱	۷۳/۸۰	۰/۱۷	۰/۱۹	۱۰۴/۹۵	۶۴/۸۴	گروه ۱
۰/۵۶	۳۷/۲۲	۷۶/۶۲	۰/۱۶	-۰/۲۱	۳۴/۷۰	۵۹/۳۵	گروه ۲
			-----				گروه ۳
			-----				گروه ۴
			-----				مدل کلی

با استفاده از عوامل انتخاب شده توسط آزمون گاما از روش تحلیل تابع تشخیص به منظور تعیین حوضه‌های

همگن‌بندی حوضه‌های آبخیز با استفاده از تحلیل تابع تشخیص

محمد آباد، سفیدچاه، کشتارگاه، کردخیل، ریگ چشمه و هم‌چنین در گروه همگن ۳: کسلیان، هرات‌بر، سلیمان تنگه، گالیکش، قلی تپه، لزوره، پس پشته، رامیان، نوده، تقی آباد، امام زاده، پل جاده، شصت کلات، پل اردوگاه، رامسر، تنگه لایوچ، سیاه آب، خیرود، نوشهر، دیناسرا، گانگسر قرار داد.

همگن استفاده شد و ۳ گروه همگن توسط این روش بر اساس میزان تولید رسوب انتخاب گردید. شکل ۲ این گروه‌ها را نشان می‌دهد. روش تحلیل تابع تشخیص با استفاده از ۵ عامل موثری که توسط آزمون گاما انتخاب شده بودند حوزه‌های آبخیز را به ترتیب در ۳ گروه همگن شامل گروه همگن ۱: دوغ، تمر، ارازکوسه، گنبد، قزاقلی، نهارخوران، ورنه، نوذر آباد، آبدلی، رزن، کیکلا، کره سنگ، گروه همگن ۲ شامل: قرآن تالار، شیرگاه، زرینگل، وطن،



شکل ۲. نحوه جدایش و قرارگیری گروه‌های همگن با استفاده از تحلیل تابع تشخیص

به مراتب بیشتر می‌باشد به گونه‌ای که شکل پراکنش موقعیت ایستگاه‌های این دو گروه در شکل ۲ و هم‌چنین مقادیر معیارهای آماری در جدول ۶ حاکی از این تشابه و همگنی می‌باشد. اما مقایسه‌ی نحوه‌ی پراکنش ایستگاه‌های گروه همگن ۱ و ۲ در تحلیل تابع تشخیص و مقادیر معیارهای آماری به ویژه در مرحله اعتبارسنجی باعث بروز تناقض‌هایی در نتایج شده است. بر اساس شکل ۲ گروه همگن ۳ به مراتب دارای نزدیکی بیشتر ایستگاه‌ها به همدیگر و تجمع بیشتر در حول مرکز ثقل نسبت به گروه همگن ۲ می‌باشند ولی مقادیر معیارهای آماری حاکی از آن است که خطای مدل‌های ارائه شده در مرحله اعتبارسنجی برای گروه همگن ۳ بیش‌تر از گروه همگن ۲ می‌باشد. در تحلیل این تناقض و نتایج می‌توان این‌گونه بیان کرد که در مرحله تقسیم‌بندی داده‌ها به دو

دقت در مقادیر متغیرهای مستقل مؤثر و حتی با تأثیرگذاری کم‌تر در تولید رسوب ایستگاه‌های قرار گرفته در گروه همگن ۱ نشان می‌دهد که این آبخیزها از نظر مقادیر متغیرها در مجموع با هم شباهت زیادی ندارند. اما بررسی‌ها نشان داد که در این گروه تشابه و نزدیکی در مقادیر بعضی متغیرها در بعضی آبخیزها وجود دارد و در عین حال در متغیرهای دیگر تفاوت زیادی بین مقادیر متغیرها وجود دارد. مذاقه در شکل ۲ و هم‌چنین مقادیر معیارهای آماری نیز نشان می‌دهد که تشابه و همگنی کم‌تری در ایستگاه‌های این گروه نسبت به دو گروه دیگر وجود دارد. لذا در مجموع می‌توان گفت تحلیل تابع تشخیص در تقسیم‌بندی ایستگاه‌های گروه مذکور از الگوی خاصی نتوانسته پیروی کند. اما در دو گروه همگن ۲ و ۳ تشابه و نزدیکی بین مقادیر متغیرهای مستقل

قرار گرفتن دو ایستگاه شیرگاه و قرآن تالار با تفاوت زیاد بین مقادیر مستقل و وابسته باعث افزایش خطای مرحله‌ی اعتبارسنجی شده است. بنابراین استفاده از الگو و قانون مشخص در تقسیم‌بندی داده‌ها به دو مرحله‌ی واسنجی و اعتبارسنجی امری ضروری به‌نظر می‌رسد. مدل‌های ارائه شده برای هر یک از گروه‌های همگن نیز در جدول ۵ ارائه شده است.

قسمت واسنجی (مدل‌سازی) و اعتبارسنجی به دلیل انتخاب تصادفی داده‌ها و عدم پیروی از قانون و الگوی مشخص و همچنین تعداد داده‌ی کم در مرحله‌ی اعتبارسنجی (Murad, 2005; Trucano *et al.*, 2006) امکان قرارگیری دو ایستگاه با متغیرهای مستقل و وابسته‌ی با ناهمگونی بیش‌تر بالا بوده و باعث بالا بودن خطای مدل در مرحله‌ی اعتبارسنجی می‌شود (Steyerberg *et al.*, 2010). در گروه همگن ۳ نیز به‌دلیل

جدول ۵. مدل‌های گروه‌های همگن انتخاب شده با تحلیل تابع تشخیص و مدل کلی

مدل‌های ارائه شده برای هر یک از گروه‌ها	
$-7.301Qp + 3.427 Di + 44.593 Sr + 1.78 Lw - 171.6 Cc + 550.071$	گروه ۱
$-20.944 Qp + 3.014 Di + 32.237 Sr + 35.195 Lw + 190.279 Cc + 11.086$	گروه ۲
$85.913 Qp + 91.867 Di + 364.556 Sr + 52.205 Lw + 1.302 Cc - 8600.483$	گروه ۳
مدل معنی‌داری ارائه نشد	مدل کلی

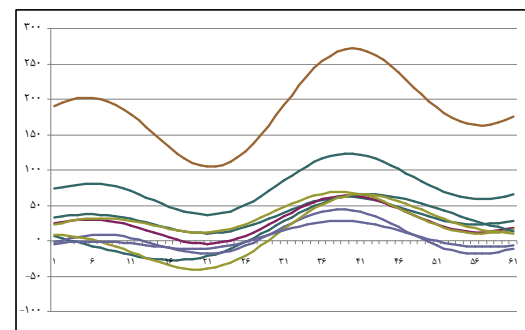
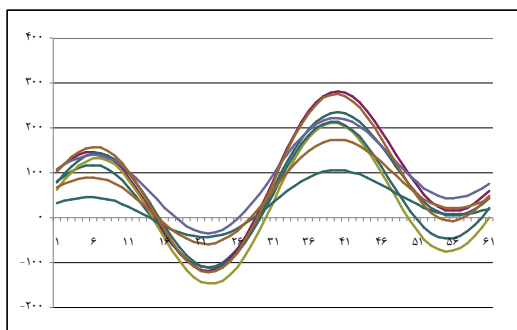
\* Qp (دبی حداکثر لحظه‌ای با دوره بازگشت دو ساله بر حسب متر مکعب بر ثانیه)، Di (قطر هم‌سطح حوضه بر حسب کیلومتر)، Sr (شیب رودخانه بر حسب کیلومتر)، Lw (طول حوزه‌ی آبخیز بر حسب کیلومتر) و Cc (ضریب فاکتور شکل حوضه).

جدول ۶. نتایج میزان دقت و خطای هر یک از مدل‌های ارائه شده در مرحله واسنجی و اعتبارسنجی

مدل‌های ارائه شده برای هر یک از گروه‌ها	
$-7.301Qp + 3.427 Di + 44.593 Sr + 1.78 Lw - 171.6 Cc + 550.071$	گروه ۱
$-20.944 Qp + 3.014 Di + 32.237 Sr + 35.195 Lw + 190.279 Cc + 11.086$	گروه ۲
$85.913 Qp + 91.867 Di + 364.556 Sr + 52.205 Lw + 1.302 Cc - 8600.483$	گروه ۳
-----	مدل کلی

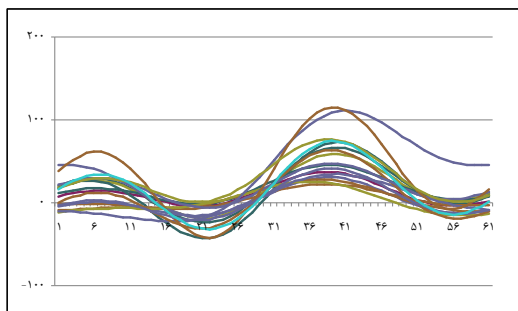
آبخیز شد که در نهایت منحنی‌های اندرو این آبخیزها را در ۴ گروه همگن قرار داد که هر یک از گروه‌ها در شکل ۳ آورده شده است.

همگن‌بندی حوزه‌های آبخیز با استفاده از منحنی‌های اندرو بر اساس ۵ عامل اصلی تاثیرگذار بر میزان تولید رسوب که با استفاده از آزمون گاما تعیین شده بودند توسط منحنی‌های اندرو اقدام به همگن‌سازی حوزه‌های

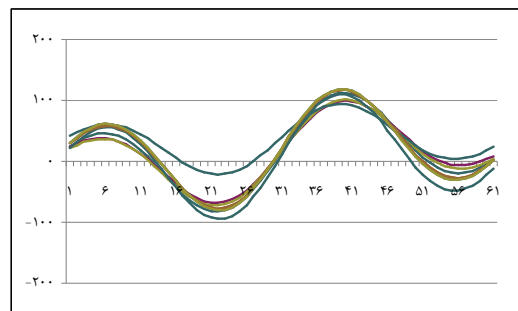


گروه همگن ۲

گروه همگن ۱



گروه همگن ۳



گروه همگن ۴

شکل ۳. منحنی‌های اندروی تهیه شده برای گروه‌های همگن

این گروه شد. گروه همگن ۳: گالیکش، قلی تپه، لزوره، نوده، تقی آباد، امامزاده، نهارخوران، پل جاده، وطن، محمد آباد، شیرگاه، خیرود، نوشهر، هرات‌بر، دیناسرا، پل اردوگاه و گانگسر. این گروه نسبت به سایر گروه‌ها تعداد بیشتری از ایستگاه‌ها را شامل شد و با بررسی مشخصات فیزیکی حوضه‌های بالادست این ایستگاه‌ها مشخص شد که حوضه‌های این گروه دارای مساحت‌های کم، دبی حداکثر لحظه‌ای با دوره بازگشت دو ساله کم، زمان تمرکز کم و شیب رودخانه‌ی اصلی زیاد و به جزء ایستگاه پل جاده، مقادیر رسوب تولیدی این ایستگاه‌ها نیز زیاد می‌باشد. گروه همگن ۴: دوغ، تمر، ارازکوسه، سفیدچاه، کسلیان، رزن، ورنند. در این گروه تعداد کمی از ایستگاه‌ها قرار گرفتند ولی به جزء ایستگاه کسلیان سایر ایستگاه‌ها دارای متغیرهای با مقادیر مشابه بودند. هم‌چنین زمان تمرکز این ایستگاه‌ها زیاد بوده و مقادیر دبی حداکثر لحظه‌ای با دوره بازگشت دوساله‌ی بالایی نیز داشتند.

با توجه به این که تعدادی از ایستگاه‌هایی که از لحاظ جغرافیایی در مجاورت هم قرار نداشتند ولی در یک گروه همگن قرار گرفتند نشان می‌دهد که تشابه هیدرولوژیکی حوضه‌ها می‌تواند عاملی جدا از عامل جغرافیایی برای تشخیص حوضه‌های همگن بوده که سهیلی و وفاخواه (۱۳۸۴) نیز چنین بیان کرده‌اند.

مطابق جدول ۷ برای گروه همگن ۱ مدل ارائه شده از معنی‌داری قابل قبولی برخوردار نبود و برای این گروه

پس از ترسیم منحنی‌های اندرو، مناطق همگن بر مبنای تشابه در روند به چهار گروه همگن تقسیم شدند که در گروه همگن ۱، ایستگاه‌های گنبد، قزاقلی، نوذر آباد، آبدلی، سلیمان تنگه، ریگ چشمه، کیکلا، کره سنگ و کردخیل قرار گرفتند. در این گروه ایستگاه‌هایی قرار گرفتند که دارای مساحت بیشتری بوده و مقادیر دبی حداکثر لحظه‌ای با دوره بازگشت دوساله و قطر هم‌سطح حوزه بیشتر از سایر گروه‌ها می‌باشد. علت عدم ارائه‌ی مدل برای این گروه را می‌توان به قرار گرفتن حوضه‌هایی با مقادیر دبی حداکثر لحظه‌ای با دوره بازگشت دو ساله، قطر هم‌سطح حوزه و مساحت متفاوت از هم ارتباط داد که باعث کاهش ارتباط بین متغیرها می‌شود. گروه همگن ۲: پس پشته، رامیان، رامسر، زرینگل، قرآن تالار، کشتارگاه، تنگه لایچ، سیاه آب و شصت کلات. ایستگاه‌های این گروه نیز دارای متغیرهای با مقادیر مشابهی بوده ولی وجود ایستگاه کشتارگاه با مقادیر دبی حداکثر لحظه‌ای با دوره بازگشت دو ساله و رسوب تولیدی بسیار بالا نسبت به سایر ایستگاه‌ها باعث شده که همگنی بین ایستگاه‌های این گروه کاهش پیدا کند و مقادیر خطای مدل‌ها افزایش داشته باشد. با اینکه مقادیر دبی حداکثر لحظه‌ای با دوره بازگشت دو ساله و رسوب تولیدی در ایستگاه نسبت به سایر ایستگاه‌ها بسیار بالا و ناهمگون بود ولی سایر متغیرهای اصلی این ایستگاه مشابه با بقیه‌ی ایستگاه‌ها بود که باعث فرارگیری این ایستگاه در

پل جاده با داشتن مقادیر مشابه متغیرهای مستقل با متغیرهای مستقل سایر ایستگاه‌ها در مرحله واسنجی، مقدار رسوب تولیدی بسیار کمی نسبت به سایر ایستگاه‌ها داشت و پس از ساخت مدل براساس ایستگاه‌های این مرحله مقدار تخمینی رسوب متوسط سالانه برای ایستگاه مذکور بسیار زیاد بود که باعث افزایش خطای مدل برای مرحله واسنجی مدل شد که با نتایج پژوهش‌های Kim و Kaluarachchi (۲۰۰۸) هم‌خوانی دارد. هم‌چنین مدلی برای کل منطقه بدون همگن‌بندی آبخیزها به دلیل عدم معنی‌داری تهیه نشد که ضرورت تهیه‌ی مدل جداگانه برای ایستگاه‌های مشابه را بیان می‌نماید. هم‌چنین براساس نمایه‌ی آماری RBias که انحراف مدل را در تخمین مقادیر متغیر وابسته بر اساس داده‌های مشاهداتی را نشان می‌دهد، مدل‌های تهیه شده برای گروه‌های همگن برای مرحله‌ی اعتبارسنجی دارای حدود ۱۲ تا ۲۰ درصد انحراف بوده که مثبت بودن مقادیر آن برای گروه‌ها بیان کننده‌ی این امر است که مدل‌های تهیه شده مقادیر رسوب متوسط سالانه را بیش از مقادیر مشاهده‌ای تخمین زده‌اند که گواه بیش تخمین بودن مدل‌هاست. نقشه‌ی پراکنش گروه‌های همگن در شکل ۴ آورده شده است.

مدلی پیشنهاد نشد ولی برای سایر گروه‌ها مدل‌هایی برخوردار از سطح معنی‌داری ۹۹ درصد تهیه شد. برای این‌که میزان خطای هر یک از گروه‌ها بدست آید از نمایه‌های آماری مذکور در پژوهش استفاده گردید. نتایج حاصل از کارایی‌سنجی مدل‌ها برای هر یک از گروه‌ها در جدول ۸ نشان داد که برای هر سه گروه همگن که مدلی برای آن‌ها ارائه شده بود از دقت قابل قبولی برخوردار بوده و مدل ارائه شده برای گروه همگن ۴ عملکرد مناسب‌تری نسبت به سایر مدل‌ها داشت. اما نکته قابل توجه این است که بر خلاف انتظار معمول در گروه‌های همگن ۲ و ۳ مقادیر معیارهای آماری RE، RBias و RRMSE در مرحله‌ی واسنجی بیش‌تر از مقادیر آن در مرحله‌ی اعتبارسنجی مدل می‌باشند که دلیل آن قرارگیری ایستگاه کشتارگاه با مقادیر بسیار بالایی از میزان رسوب متوسط سالانه بود که در مرحله جدا کردن ایستگاه‌ها برای مراحل واسنجی و اعتبارسنجی در مرحله‌ی واسنجی قرار گرفت که مدل تهیه شده بر اساس کل ایستگاه‌های مرحله‌ی واسنجی، مدل بسیار کم‌تر از مقادیر مشاهده‌ای برای این ایستگاه تخمین زد که باعث بالا رفتن خطای مدل برای این مرحله شد. برای گروه همگن ۳ نیز ایستگاه

جدول ۷. مدل‌های تهیه شده برای گروه‌های همگن حاصل از منحنی اندرو و مدل کلی

مدل‌های ارائه شده برای هر یک از گروه‌ها	
مدلی ارائه نشد	گروه ۱
$12.235 Qp + 43.203 Di + 180.336 Sr + 8 Lw - 177.923 Cc - 1800$	گروه ۲
$- 20.944 Qp + 3.014 Di + 32.237 Sr + 35.195 Lw + 190.279 Cc + 11.086$	گروه ۳
$85.913 Qp + 91.867 Di + 364.556 Sr + 52.205 Lw + 1.302 Cc - 8600.483$	گروه ۴
مدل معنی‌داری ارائه نشد	مدل کلی

\* Qp (دبی حداکثر لحظه‌ای با دوره بازگشت دو ساله بر حسب متر مکعب بر ثانیه)، Di (قطر هم‌سطح حوضه بر حسب کیلومتر)، Sr (شیب رودخانه بر حسب کیلومتر)، Lw (طول حوزه آبخیز بر حسب کیلومتر) و Cc (ضریب فاکتور شکل حوضه).

جدول ۸. نتایج میزان دقت و خطای هر یک از مدل‌های ارائه شده در مرحله واسنجی و اعتبارسنجی

R	RRMSE (%)		RBIAS		RE (%)		
	اعتبارسنجی	واسنجی	اعتبارسنجی	واسنجی	اعتبارسنجی	واسنجی	
							مدل معنی‌داری ارائه نشد
۰/۹۵	۳۸/۰۰	۶۴/۵۵	-۰/۰۱	۰/۰۱	۳۰/۸۲	۵۵/۹۰	گروه ۱
۰/۶۲	۳۶/۱۴	۵۹/۰۰	-۰/۱۷	-۰/۰۸	۲۹/۱۳	۴۵/۵۳	گروه ۲
۰/۸۲	۲۵/۲۰	۳۵/۹۳	۰/۱۵	-۰/۱۰	۱۹/۶۹	۲۹/۵۵	گروه ۳
							مدل معنی‌داری ارائه نشد
							مدل کلی

## انتخاب بهترین روش همگن‌سازی

پس از کاهش تعداد عوامل موثر در تولید رسوب به مؤثرترین عوامل با استفاده از روش آزمون گاما و استفاده از عوامل انتخاب شده و به‌کارگیری روش‌های همگن‌سازی تحلیل خوشه‌ای، تحلیل تابع تشخیص و منحنی‌های اندرو به‌منظور تعیین گروه‌های همگن، میزان خطای هر یک از روش‌ها ابتدا برای هر یک از گروه‌های

مدل‌سازی رگرسیونی با رعایت انتخاب بازه‌ی مرحله‌ی واسنجی و اعتبارسنجی صورت گرفت. سپس به‌منظور انتخاب بهترین روش همگن‌سازی متوسط خطا و کارایی هر یک از روش‌های همگن‌سازی حوضه‌ها با استفاده از معیارهای آماری استخراج شد که در جدول ۹ نشان داده شده‌اند.

جدول ۹. متوسط خطای مدل‌های هر یک از روش‌های همگن‌سازی آبخیزها

R	RRMSE (%)		RBIAS		RE (%)		
	اعتبارسنجی	واسنجی	اعتبارسنجی	واسنجی	اعتبارسنجی	واسنجی	
۰/۶۷	۷۸/۱۱	۷۵/۲۱	۰/۱۶	-۰/۰۱	۶۹/۸۲	۶۲/۹۵	تحلیل خوشه‌ای
۰/۶۱	۵۸/۴۵	۴۱/۰۷	۰/۳۰	۰/۰۲	۵۷/۳۹	۴۱/۳۱	تحلیل تابع تشخیص
۰/۸۰	۳۳/۱۱	۵۳/۱۶	-۰/۱	۰/۰۱	۴۵/۹۱	۳۸/۱۲	منحنی‌های اندرو
			----				مدل کلی

است. دلیل برتری روش منحنی‌های اندرو به‌عنوان یک روش غیرعددی نسبت به دو روش دیگر به‌عنوان روش‌های عددی را می‌توان به ماهیت و فلسفه‌ی ارائه هر یک از روش‌های مذکور نسبت داد. روش منحنی‌های اندرو با توجه به این‌که به هر یک از متغیرهای مستقل بر اساس میزان تأثیری که بر متغیر وابسته دارند وزن معینی در نظر می‌گیرد. لذا متغیر مستقل با تأثیرگذاری بیش‌تر نقش بیش‌تری را در ترسیم دوبعدی الگوی رفتاری هر یک از حوضه‌ها داشته و هر چقدر تعداد متغیر مستقل

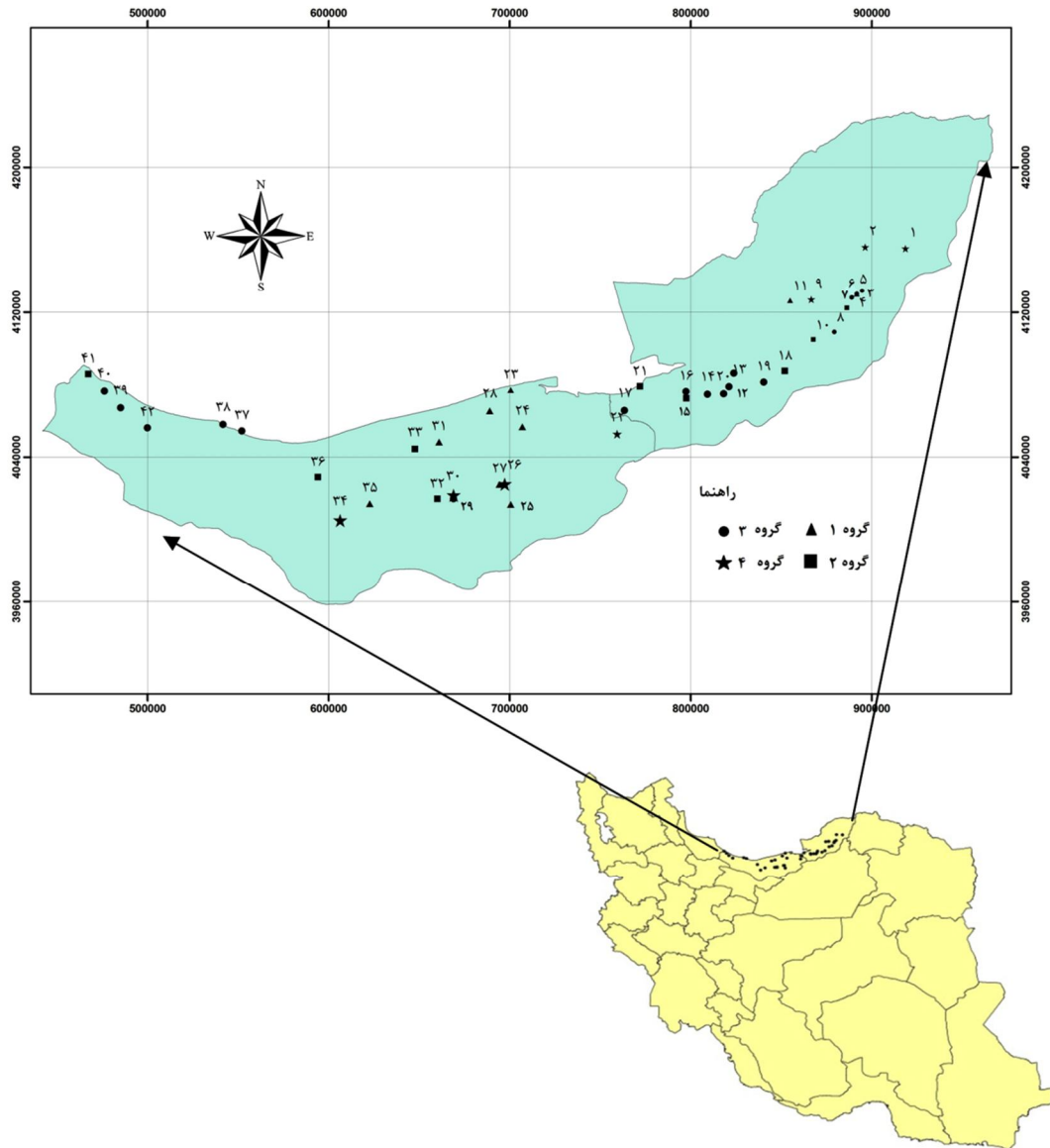
نتایج حاکی از آن بود که روش منحنی‌های اندرو به‌ترتیب با ۳۸/۱۲ و ۴۵/۹۱ درصد میزان خطای نسبی (RE)، ۵۳/۱۶ و ۳۳/۱۱ درصد ریشه نسبی مجذور میانگین مربعات خطا (RRMSE) و ۰/۰۱- و ۰/۰۱ میزان ریشه نسبی انحراف مدل (RBIAS) در مرحله‌ی واسنجی و اعتبارسنجی و با متوسط ضریب تعیین ۰/۸۰ نسبت به دو روش تحلیل خوشه‌ای و تحلیل تابع تشخیص عمل‌کرد بهتری در همگن‌سازی حوضه‌های آبخیز جنوب و جنوب شرقی دریای خزر بر اساس تولید رسوب داشته

ارائه نشده است. لذا روش تحلیل خوشه‌ای نیز هر چند نتایج مطلوبی نداشت ولی در قیاس با ارائه مدل کلی برای منطقه مناسب می‌باشد.

### نتیجه‌گیری

پژوهش حاضر با هدف ارزیابی میزان کارایی روش‌های همگن‌سازی حوزه‌های آبخیز در تشخیص آبخیزهای همگن جنوب و جنوب شرقی دریای خزر از لحاظ مقدار تولید رسوب صورت گرفت. برای این منظور از روش آزمون گاما برای کاهش تعداد عوامل موثر بر تولید رسوب در حوضه‌ها و انتخاب موثرترین عوامل استفاده گردید و روش‌های تحلیل خوشه‌ای، تحلیل تابع تشخیص و منحنی‌های اندرو به عنوان روش‌های غیر عددی و به عبارتی گرافیکی و قیاس چشمی برای قرار دادن حوزه‌های آبخیز در گروه‌های همگن به کار برده شد. در نهایت به منظور بررسی میزان کارایی هر یک از این روش‌ها در تخمین مقدار رسوب معلق اقدام به مدل‌سازی رگرسیونی در هر یک از گروه‌های همگن و هم‌چنین برای کل ایستگاه‌ها به عنوان یک مدل کلی برای منطقه‌ی مورد مطالعه شد. براساس نتایج بدست آمده روش‌های تجزیه‌ی خوشه‌ای و منحنی‌های اندرو، ۴۲ ایستگاه مورد مطالعه را در ۴ و تحلیل تابع تشخیص در ۳ گروه همگن قرار دادند. با توجه به عدم ارائه‌ی مدل مناسب برای برخی گروه‌ها و مقادیر خطای بالاتر مدل‌ها می‌توان به این نتیجه رسید که هر چه مقادیر متغیرهای موثر بر تولید رسوب در ایستگاه‌ها دارای تشابه بیشتری باشند همگنی گروه‌ها بیشتر خواهد بود. Tramblay و همکاران (۲۰۱۰) نیز در پژوهشی به منظور مدل‌سازی منطقه‌ای رسوب معلق با استفاده از متغیرهای فیزیکی حوضه‌ها به نتایج مشابهی دست یافتند.

کم‌تر شود به تبع آبخیزهای با الگوی رفتاری یکسان با احتمال بیشتری در یک گروه همگن قرار خواهند گرفت. هم‌چنین روش تحلیل تابع تشخیص نیز به ترتیب با ۴۱/۳۱ و ۵۷/۳۹ درصد میزان خطای نسبی (RE)، ۴۱/۰۷ و ۵۸/۴۵ درصد ریشه نسبی مجذور میانگین مربعات خطا (RRMSE) و ۰/۰۲ و ۰/۳۰ میزان ریشه نسبی انحراف مدل (RBIAS) در مرحله‌ی واسنجی و اعتبارسنجی و با متوسط ضریب تعیین ۰/۶۱ نتایج مطلوب و قابل قبولی را ارائه داد. اما با این حال روش تحلیل خوشه‌ای به ترتیب با ۶۳/۹۵ و ۶۹/۸۲ درصد میزان خطای نسبی (RE)، ۷۵/۲۱ و ۷۸/۱۱ درصد ریشه نسبی مجذور میانگین مربعات خطا (RRMSE) و ۰/۰۱- و ۰/۱۶ میزان ریشه نسبی انحراف مدل (RBIAS) در مرحله‌ی واسنجی و اعتبارسنجی و با متوسط ضریب تعیین ۰/۶۷ در همگن‌سازی حوزه‌های آبخیز نسبت به دو روش دیگر مورد تأیید قرار نگرفت. دقت در نتایج و کارایی هر یک از روش‌ها نشان می‌دهد که هرچه تعداد الگوها و الگوریتم‌های داخلی روش‌های مذکور بیشتر باشد، میزان دقت آن روش نیز کم‌تر خواهد بود. در روش تحلیل خوشه‌ای برای قرار دادن داده‌های با تشابه بیشتر در یک گروه از الگوریتم‌های متعددی استفاده می‌گردد که دلیل ارائه الگوریتم‌های متعدد می‌تواند به رفتار متفاوت روش تحلیل خوشه‌ای نسبت به داده‌های مختلف باشد. از طرفی پژوهشگران بیان می‌کنند که برای استفاده از تحلیل خوشه‌ای باید تعداد متغیرها و داده‌ها هر چه قدر بیشتر باشد نتایج خوشه‌بندی قابل اعتمادتر و دقیق‌تر خواهد بود که در پژوهش حاضر کم بودن تعداد متغیرها می‌تواند دلیلی بر کارایی ضعیف این روش نسبت به دو روش دیگر باشد. با این حال توجه به جدول‌های ارائه شده فوق حاکی از آن است که مدل معنی‌داری برای کل منطقه مورد مطالعه بدون انجام همگن‌سازی حوزه‌های آبخیز



شکل ۴. نقشه موقعیت ایستگاه‌های مورد مطالعه و پراکنش گروه‌های همگن در جنوب و جنوب شرقی دریای خزر

ایستگاه‌های منطقه‌ی مورد مطالعه نتایج مناسبی را ارائه نخواهد داد و در تخمین رسوب برای ایستگاه‌های دارای نقص آماری یا فاقد آمار استفاده از مدل‌های تهیه شده در ایستگاه‌های مشابه نتایج مطلوب‌تری را در پی خواهد داشت که Tramblay و همکاران (۲۰۱۰) نیز به چنین نتایجی دست یافتند. همچنین بر اساس بررسی‌های صورت گرفته در این پژوهش مقادیر دبی حداکثر لحظه‌ای با دوره بازگشت ۲ ساله بیش‌ترین تاثیر را در تولید میزان رسوب ایفا می‌کند که هم‌سو با نتایج Restrepo و

نتایج حاصل از گروه‌بندی حوضه‌ها و خطای مدل‌های ساخته شده نشان می‌دهد که استفاده از یک مدل واحد برای تخمین رسوب معلق حوضه‌ها با ویژگی‌های فیزیکی متفاوت هم‌چون مساحت، شیب رودخانه و همچنین متغیرهای هیدرولوژیکی از قبیل دبی حداکثر لحظه‌ای با دوره بازگشت دو ساله، علاوه بر معنی‌دار بودن مدل، تخمین غیرواقعی از مقدار حقیقی را خواهد داشت که با نتایج تحقیق Bray و Xie (۱۹۹۳) موافق می‌باشد. بنابراین استفاده از یک مدل واحد برای



تصاویر چرنف استفاده شود. از طرفی در تهیه مدل‌های تخمین متغیر مستقل در گروه‌های همگن نیز کاربرد روش‌های هوش مصنوعی و الگوریتم ژنتیک مطمئناً در ارائه نتایج جامع‌تر می‌تواند مؤثر باشد.

#### فهرست منابع

خیرفام، ح. ۱۳۹۱. تعیین و ارزیابی تغییرپذیری زمانی نسبت بار بستر به معلق در رودخانه‌ی کجور. پایان‌نامه کارشناسی ارشد، دانشگاه تربیت مدرس، ۷۳ ص.  
خیرفام ح. و صادقی س.ح.ر. ۱۳۹۱ الف. مدل‌سازی رفتار هیدرومئولوژی حوزه‌ی آبخیز آق‌قلا، همایش ملی جریان و آلودگی آب، ۳ و ۴ خرداد ۱۳۹۱، دانشگاه تهران، ۸ ص.

خیرفام، ح.، وفاخواه، م. و حسینی، ص. ۱۳۹۱ ب. تخمین جریان رودخانه‌ای در حوزه‌های بالادست سد مخزنی بوکان با استفاده از متغیرهای اقلیمی دما و بارش، فصلنامه علمی پژوهشی مهندسی آبیاری و آب، ۱۲(۷): ۱۱۲-۱۲۳.

سهیلی، الف. و وفاخواه، م. ۱۳۸۴. بررسی کارایی روش همگن‌سازی غیر عددی (منحنی‌های اندرو) در برآورد سیل. پژوهش و سازندگی (در منابع طبیعی)، ۶۹: ۷۳-۸۱

صادقی س.ح.ر.، مرادی ح.ر.، مزین م. و وفاخواه م. ۱۳۸۴. کارایی روش‌های مختلف تجزیه و تحلیل آماری در مدل‌سازی بارش و رواناب (مطالعه موردی: حوزه آبخیز کسلیان). مجله علوم کشاورزی و منابع طبیعی. ۱۲(۳): ۸۱-۹۰.

صادقی، س.ح.ر.، مزین، م. و مرادی، ح.د. ۱۳۸۶. تهیه هیدروگراف با استفاده از اجزای مختلف بارندگی در حوزه آبخیز کسلیان، مجله منابع طبیعی ایران، ۶۰ (۱): ۳۳-۴۴.

کیانی هرچگانی م.، صادقی س.ح.ر. ۱۳۹۲. تغییرات مکانی ارتباط بین انتقال فلزات سنگین و توزیع اندازه

همکاران (۲۰۰۶) می‌باشد. با دقت بر این موضوع که ایستگاه‌هایی که در یک گروه همگن قرار گرفته‌اند بعضاً تشابه جغرافیایی ندارند دلالت بر این امر دارد که تشابه حوضه‌ها فقط از نزدیکی جغرافیایی تبعیت نمی‌کند که با نتایج سهیلی و وفاخواه (۱۳۸۴) و Tramblay و همکاران (۲۰۱۰) مطابقت دارد. بر اساس جدول‌های معیارهای ارزیابی روش منحنی‌های اندرو مناسب‌ترین عمل‌کرد در تعیین آبخیزهای همگن را داشت و همچنین تحلیل تابع تشخیص نیز نتایج مطلوبی ارائه داد. ولی نتایج همگن‌سازی تحلیل خوشه‌ای نسبت به دو روش دیگر ضعیف‌تر بود، هرچند استفاده از این روش نسبت به ارائه‌ی مدل کلی برای منطقه توجیه‌پذیر است. از نتایج دیگر پژوهش حاضر لزوم تأکید بر پیروی از یک قانون و الگوی مشخص در تقسیم‌بندی داده‌ها به دو مرحله‌ی واسنجی و اعتبارسنجی می‌باشد. در کل نتایج این پژوهش نشان می‌دهد که کاهش تعداد عوامل مؤثر بر تولید رسوب به عوامل با تاثیرگذاری بیش‌تر در حوزه‌های آبخیز جنوب و جنوب شرقی دریای خزر با استفاده از آزمون گاما و گروه‌بندی این آبخیزها با استفاده از روش‌های همگن‌سازی گرافیکی و غیر عددی به‌ویژه منحنی‌های اندرو در گروه‌های همگن از لحاظ میزان تولید رسوب متوسط سالانه و همچنین بررسی میزان کارایی این روش‌ها با معیارهای عددی و آماری به‌منظور جمع‌بندی جامع و مطمئن‌تر، رویکرد مناسبی به‌منظور همگن‌سازی حوزه‌های آبخیز منطقه‌ی مورد مطالعه بوده که در نهایت با به‌کارگیری مدل‌های تهیه شده منتج به تخمین قابل قبولی از میزان رسوب معلق متوسط سالانه‌ی ایستگاه‌های فاقد آمار رسوب معلق و یا دارای نقص آماری مشابه با ایستگاه‌های هر یک از گروه‌های همگن خواهد شد. همچنین پیشنهاد می‌شود در پژوهش‌های آتی هم‌راستا با پژوهش حاضر از روش‌های دیگر کاهش عوامل مؤثر همچون تحلیل عاملی و همچنین از روش‌های دیگری در تفکیک آبخیزهای همگن از قبیل تصاویر ستاره‌ای و

- problem revisited. *Progress in Physical Geography*, 31: 155-178.
- Durrant, P.J. 2001. wingamma: A non-linear data analysis and modeling tool with applications to flood prediction. Ph.D thesis, Department of Computer Science, Cardiff University, Wales, UK.
- Heng, S., and Suetsugi, T. 2014. Development of a regional model for catchment-scale suspended sediment yield estimation in ungauged rivers of the Lower Mekong Basin. *Geoderma*, 235: 334-346.
- Horhota, S.T. and Aitken, C.L. 2006. Multivariate cluster analysis of pharmaceutical formulation data using Andrews plots. *Journal of Pharmaceutical Sciences*, 80: 85-90.
- Jarvie, H.P., Oguchi, T. and, Neal, C. 2002. Exploring the linkages between river water chemistry and watershed characteristics using GIS-based catchment and locality analysis. *Regional Environmental Change*, 36: 36-50.
- Jobson, J.D. 1992. *Applied multivariate data analysis, Vol. II: Categorical and multivariate methods*. Springer-Verlag, 731 pp.
- Kim, U. and Kaluarachchi, J.J. 2008. Application of parameter estimation and regionalization methodologies to ungauged basins of the Upper Blue Nile River Basin, Ethiopia. *Journal of Hydrology*, 362: 39-56.
- Ki i, Ö. 2010. River suspended sediment concentration modeling using a neural differential evolution approach. *Journal of Hydrology*, 389: 227-235
- Koncar, N. 1997. *Optimisation methodologies for direct inverse neurocontrol*. PhD thesis, Department of Computing, Imperial College of Science, Technology and Medicine, University of London.
- Lin, G.F. and Wang, Ch.M. 2006. Performing cluster analysis and discrimination analysis of hydrological factors in one step. *Advances in Water Resources*, 29:1573-1585.
- Lin, Sh.W. and Chen, Sh.Ch. 2009. PSOLDA: A particle swarm optimization approach for enhancing classification accuracy rate of linear discriminant analysis. *Applied Soft Computing*, 9: 1008-1015.
- Ludwig, W. and Probst, J.L. 1998. River sediment discharge to the oceans: present-day controls and global budgets. *American Journal of Science*, 298: 265-295.
- Melesse, A.M., Ahmad, S., McClain, M.E., Wang, X. and Limd, Y.H. 2011. Suspended sediment load prediction of river systems: An artificial neural network approach. *Agricultural Water Management*, 98: 855-866.
- Moghaddamnia, A., Ghafari-Gousheh, M., Piri, J., Amin, S. and Han, D. 2009. Evaporation ذرات رسوبات معلق، پژوهش‌های حفاظت آب و خاک، ۲۰(۱): ۱۶۹-۱۸۴.
- وفاخواه، م. ۱۳۸۵. تجزیه و تحلیل ناحیه‌ای رسوب در بخشی از حوزه‌های آبخیز ساحلی دریای خزر. *مجله علوم کشاورزی و منابع طبیعی*، ۱۳(۶): ۱۲۱-۱۳۱.
- Alp, M. and Cigizoglu H.K. 2007. Suspended sediment load simulation by two artificial neural network methods using hydrometeorological data. *Environmental Modelling & Software*, 22: 2- 13.
- Asselman, N.E.M. 2000. Fitting and interpretation of sediment rating curves. *Journal of Hydrology*, 234: 228-248.
- Agalbjörn, S., Koncar, N. and Jones, A.J. 1997. A note on the gamma test. *Neural Computing Applications*, 5: 131-133.
- Andrews D. 1972. Plots of high-dimensional data. *Biometrics*, 28: 125-136.
- Bray D.I. and Xie, H. 1993. A regression method for estimating suspended sediment yields for ungauged watershed in Atlantic Canada. *Canadian journal of Civil Engineering*, 20: 82-87.
- Caratti, J.F., Nesser, J.A. and Maynard, C.L. 2004. Watershed classification using canonical correspondence analysis and clustering techniques: A cautionary note. *Journal of the American Water Resources Association (JAWRA)*, 40(5):1257-1268.
- Carroll. S.P., Dawes. L., Hargreaves, M. and Goonetilleke, A. 2009. Fecal pollution source identification in an urbanizing catchment using antibiotic resistance profiling, discriminant analysis and partial least squares regression. *Water Research*, 43: 1237-1246.
- Chow, V.T., Maidment, D.R. and Mays, L.W. 2007. *Applied Hydrology*. McGraw-Hill. 588 pp.
- Cigizoglu, H.K. 2004. Estimation and forecasting of daily suspended sediment data by multi-layer perceptrons. *Advances in Water Resources*, 27: 185-195.
- Corcoran, J., Wilson, I. and Ware, J. 2003. Predicting the geo-temporal variation of crime and disorder. *International Journal of Forecasting*, 19: 623-634.
- Detenbeck, N.E., Brady, V.J., Taylor, D.L., Snarski, V.M. and Batterman, S.L. 2005. Relationship of stream flow regime in the western Lake Superior basin to watershed type characteristics. *Journal of Hydrology*, 309: 258-276.
- De Vente, J., Poesen, J., Arabkhedri, M. and Verstraeten, G. 2007: The sediment delivery

- suspended-sediment discharge for the eastern United States. *Journal of Hydrology*, 472: 53-62.
- Sadeghi, S.H.R. and Mahdavi, M. 2004. Applicability of SEDIMOT II model in flood and sediment yield estimation. *Journal of Agricultural Science and Technology*, 6: 147-154.
- Sadeghi, S.H.R., Mizuyama, T., Miyata, S., Gomi, T., Kosugi, K., Fukushima, T., Mizugaki, S. and Onda, Y. 2008. Determinant factors of sediment graphs and rating loops in a reforested watershed. *Journal of Hydrology*, 356: 271-282.
- Sadeghi, S.H.R. and Saeidi, P. 2010. Reliability of sediment rating curves for a deciduous forest watershed in Iran. *Hydrological Sciences Journal*, 55(5): 821- 831.
- Sadeghi, S.H.R. and Singh, J.K. 2005. Development of a synthetic sediment graph using hydrological data. *Journal of Agricultural Science and Technology*, 7: 69-77.
- Siakeu, J., Oguchi, T., Aoki, T., Esaki, Y. and Jarvie, H.P. 2004. Change in riverine suspended sediment concentration in central Japan in response to late 20th century human activities. *Catena*, 55: 231-254.
- Steyerberg, E.W., Vickers A.J., Cook, N.R., Gerds, Th., Gonen, M., Obuchowski, N., Pencina, M.J. and Kattan, M.W. 2010. Assessing the performance of prediction models: A framework for traditional and novel measures. *Epidemiology*, 21(10): 128-138.
- Syvitski, J.P.M. and Milliman, J.D. 2007. Geology, geography and humans battle for dominance over the delivery of fluvial sediment to the coastal ocean. *Journal of Geology*, 115: 1-19.
- Tramblay, Y., Ouarda, T.B.M.J., St-Hilaire, A. and Poulin, J. 2010. Regional estimation of extreme suspended sediment concentrations using watershed characteristics. *Journal of Hydrology*, 380: 305-317.
- Tramblay, Y., St-Hilaire, A. and Ouarda, T.B.M.J. 2007. Modelling extreme suspended sediment concentrations in North America: frequency analysis and correlations with watershed characteristics. In: *Water Quality and Sediment Behaviour of the Future: Predictions for the 21st Century*, Proceedings of Symposium HS2005 at IUGG2007, Perugia, Italy, July 2007. IAHS Publication no, 314: 20-27.
- Tramblay, Y., St-Hilaire, A. and Ouarda, T.B.M.J. 2008. Frequency analysis of maximum annual suspended sediment concentrations in North America. *Hydrological Science Journal*, 53(1): 236-252.
- Trucano, T.G. Swiler, L.P., Igusa, T., Oberkamp, W.L. and Pilc, M. 2006. Calibration, validation, and sensitivity analysis: What's what? *Reliability Engineering and System Safety*, 91: 1331-1357.
- Tryon, R.C. 1939. *Cluster analysis*. New York: McGraw-Hill.
- estimation using artificial neural networks and adaptive neuro-fuzzy inference system techniques. *Advances in Water Resources*, 32: 88-97.
- Mourad, M., Bertrand-Krajewski, J.L. and Chebbo, G., 2005. Calibration and validation of multiple regression models for stormwater quality prediction: data partitioning, effect of dataset size and characteristics. *Water Science and Technology*, 52(3): 45-52.
- Moustafa, R.E. 2011. *Andrews curves*. Advanced review. John Wiley & Sons, Inc. 3: 373-382.
- Nadal-Romero, E., Martínez-Murillo, J.F., Vanmaercke, M. and Poesen, J. 2011: Scale-dependency of sediment yield from badland areas in Mediterranean environments. *Progress in Physical Geography*, 35: 297- 332.
- Nathan, R.J. and McMahon, T.A. 1990. Identification of homogeneous regions for purposes of regionalization. *Journal of Hydrology*, 121: 217-238.
- Noori, R., Hoshyaripour, G., Ashrafi, K. and Nadjar-Araabi, B., 2010. Uncertainty analysis of developed ANN and ANFIS models in prediction of carbon monoxide daily concentration. *Atmospheric Environment*, 44: 476-482.
- Noori, R., Karbassi, A.R., Moghaddamnia, A., Han, D., Zokaei-Ashtiani, M.H., Farokhnia, A. and Ghafari G.M. 2011. Assessment of input variables determination on the SVM model performance using PCA, Gamma test, and forward selection techniques for monthly stream flow prediction. *Journal of Hydrology*, 401: 177-189.
- Ouarda, T.B.M.J., Cunderlik, J.M., St-Hilaire, A., Barbet, M., Bruneau, P. and Bobée, B. 2006. Data-based comparison of seasonality-based regional flood frequency methods. *Journal of Hydrology*, 330: 329-339.
- Raju, K.S. and Kumar, D.N. 2011. Classification of micro watersheds based on morphological characteristics. *Journal of Hydro-environment Research*, 5: 101-109.
- Ramos, M.C. 2001. Divisive and hierarchical clustering techniques to analyze variability of rainfall distribution patterns in a Mediterranean region. *Journal of Hydrology*, 57: 123-138.
- Restrepo, J., Kjerfe, B. and Hermelin, M. 2006. Factors controlling sediment yield in a major South American drainage basin: the Magdalena River, Colombia. *Journal of Hydrology*, 316: 213-232.
- Riggs, H.C. 1973. *Regional analysis of stream flow characteristics*. Techniques of Water Resources Investigations Book 4, Chapter B3. Washington, DC: USGS.
- Roman, D. C., Vogel, R. M., and Schwarz, G. E. 2012. Regional regression models of watershed

- December 1988. IAHS Publication no, 174: 337-350.
- Wilson, D.I. 2002. Derivation of the chalk superficial deposits of the North Downs, England: an application of discriminant analysis. *Geomorphology*, 42: 343–364.
- Zhang, Q.F., Wu, F.Q., Wang, L., Yuan, L. and Zhao, L. 2011. Application of PCA integrated with CA and GIS in eco-economic regionalization of Chinese Loess Plateau. *Ecological Economics*, 70: 1051–1056.
- Tsui, A.P.M., Jones, A.J. and De Oliveira, A.G. 2002. The construction of smooth models using irregular embeddings determined by a gamma test analysis. *Neural Computing Applications*, 10: 318–329.
- Vanmaercke, M., Poesen, J., Verstraeten, G., De Vente, J. and Ocakoglu, F. 2011. Sediment yield in Europe: Spatial patterns and scale dependency. *Geomorphology*, 130: 142–161.
- Walling, D.E. and Webb, B.W. 1988. The reliability of rating curve estimates of suspended yield: some further comments, sediment budgets. In: *Proceedings of the Porto Alegre Symposium*.



## Evaluation of gamma test, cluster analysis, discriminant function analysis and andrews curves methods to separate homogeneous watersheds for regional analysis of suspended sediment

Hossein Kheirfam<sup>1</sup> and Mehdi Vafakhah<sup>2\*</sup>

1) Ph.D Student, Department of Watershed Management Engineering, Faculty of Natural Resources, Tarbiat Modares University, Noor, Mazandaran, Iran

2\*) Associate Professor, Department of Watershed Management Engineering, Faculty of Natural Resources, Tarbiat Modares University, Noor, Mazandaran, Iran. \*Corresponding author email: [vafakhah@modares.ac.ir](mailto:vafakhah@modares.ac.ir)

Received: 21-06-2013

Accepted: 02-07-2014

### Abstract

Sediment yield resulting from soil erosion in the watersheds is the major limitation in achieving the sustainable development and major threat to ecosystems. Therefore estimation of output sediment from watersheds is very important. Extent of watershed and deficiency of sediment measuring stations have caused us to use different indirect methods to estimate sediment, such as the use of models provided in similar watersheds. In this study 42 sediment measuring stations in south and southeast of the Caspian Sea with over 20 year period were chosen. By relating suspended sediment load to stream discharge in the times of flood, daily suspended sediment was estimated by using the daily discharge and the average of annual sediment was calculated. By using the Gamma Test, 13 factors affecting sediment yield were reduced to 5 main factors and by using cluster analysis, discriminate function analysis and andrews curves, study stations were put in homogeneous groups. For each homogeneous group obtained from any one of the mentioned homogenization methods and by using five main factors selected, regression models were developed to estimate the average of annual suspended sediment. Error rates and accuracy of prepared models by using statistical indices of RE, RBIAS and RRMSE were calculated according to observed data. Results indicated that all three homogenous techniques had better results than those of the general model and Andrews Curves with 38.12 and 45.91% RE, 53.16 and 33.11% RRMSE and -0.01 and 0.01 RBIAS in calibration and validation stages, respectively and had better performance than those of two methods i.e. Cluster Analysis and Discriminate Function Analysis for homogenizing of south and southeaster Caspian Sea watersheds based on sediment yield. Also peak discharge ( $Q_p$ ) has the most impact on the average of annual suspended sediment changes.

**Keywords:** homogeneous watersheds; sediment modeling; sediment rating curve; suspended sediment estimation