سال ۱۰، شماره ۳۸، تابستان ۱۳۹۵، صفحات ۸۲- ۷۱

برتـری روش خوشـه‌بندی C- میانگین فازی در بیان توزیع رخساره‌های هیدروشیمی آب زیرزمینی دشت ورامین

محمد نخعی[۱و*]، مهدی تلخابی[۲] و میثم ودیعتی[۳]

۱. دانشیار هیدروژئولوژی، دانشکده علوم زمین، دانشگاه خوارزمی تهران

۲. دانش آموخته هیدروژئولوژی، دانشکده علوم زمین، دانشگاه خوارزمی تهران

۳. دانشجوی دکتری هیدروژئولوژی، دانشکده علوم زمین، دانشگاه تبریز

تاریخ دریافت: ۹۳/۷/۵

تاریخ پذیرش: ۹۳/۱۲/۲۸

چکیده

در پژوهش حاضر، خوشـه‌بندی مجموعه‌ای از داده‌های هیدروشیمی دشت ورامین با استفاده از روش‌های خوشـه‌بندی C- میانگین فازی (FCM) و تحلیل خوشه سلسله مراتبی (HCA) انجام شده و کاربرد آن‌ها در تغییرات رخساره‌های هیدروشیمی بحث گردید. نمونه‌های آب زیرزمینی با استفاده از بهینه کردن تعداد خوشه و درجه فازی‌شدگی با استفاده از روش C- میانگین فازی به سه گروه طبقه‌بندی شدند. از داده‌های آب زیرزمینی ۹۰ نمونه چاه عمیق و نیمه عمیق و ۹ متغیر هیدروشـیمی منطقه موردمطالعه استفاده شد. نتایج این دو روش، مراکز خوشـه را تولید می‌کند که در تشـخیص فرایندهای فیزیکی و شیمیایی تغییرات هیدروشیمی منطقه موردمطالعه مؤثر است. در روش FCM تعداد خوشه بهینه توسط توابع بهینه‌یابی تعیین می‌شود اما در روش HCA براساس تجربه کاربر و سعی و خطا تعیین می‌شود. روش FCM روشی مناسب در تحلیل داده اکتشافی در بیان توزیع رخساره‌های هیدروشیمی است و زمانی‌که خوشه‌های پیوسته یا دارای هم‌پوشـانی وجود دارند، ابزار بهتری نسـبت به HCA برای خوشه‌بندی است. با ترسیم خطوط تراز مقدار عضویت هر خوشـه که بر روی نقشـه به‌صورت مکانی و پیوسته نشان داده شده، خوشه‌های نمونه‌های آب زیرزمینی به‌خوبی مشخص‌شـده است. نتایج نشـان داد؛ روش FCM در تحلیل داده‌های مرزی، نسبت به روش HCA که تغییراتی واضح و ناگهانی دارد؛ تواناتر است.

واژه‌های کلیدی: آب زیرزمینی، رخساره هیدروشیمی، دشت ورامین، خوشه‌بندی، منطق فازی.

مقدمه

روش‌های آماری چند متغیره روشی مناسب برای تجزیه و تحلیل نمونه‌های آب زیرزمینی و ویژگی‌های هیدروشیمی است (Singh et al., 2004). این روش‌ها به‌صورت هم‌زمان از چندین پارامتر اسـتفاده می‌کنند تا تحلیل مناسب‌تری از روابط بین متغیرها داشته باشند. در مطالعات هیدروشیمی، گروه‌بندی به‌منظور مجزا کردن داده‌ها به خوشه‌های معرف (کـه به‌عنوان گروه‌های آب یا رخسـاره‌های هیدروشـیمی نیز شـناخته می‌شـوند) به‌کار می‌رود تا فرایندهای طبیعی

* نویسنده مرتبط nakhaei@khu.ac.ir

تغییردهنده پارامترهای هیدروشیمی تشخیص داده شود. این خوشه‌های معرف که به تعیین روندهای اصلی شیمیایی کمک می‌کنند؛ می‌توانند برای بازسازی فرایندهای اصلی مورداستفاده قرار گیرند (Barbieri et al., 2001).

روش‌های خوشه‌بندی به دو شاخه اصلی تقسیم می‌شوند. روش‌های سخت[1] یا کلاسیک که در آن یک عنصر به‌طور کامل به یک گروه تعلق دارد و روش‌های نرم[2] یا فازی که در آن یک عنصر به تمام خوشه‌ها با درجه عضویت‌های مختلف تعلق دارد. هدف اصلی این روش‌ها چه سخت و چه فازی تفکیک یک ماتریس داده با n نمونه و p متغیر به c زیرگروه همگن به‌وسیله دسته‌بندی دقیق نمونه‌های مرتبط به خوشه‌های مشخص است. در نتیجه تقسیم‌بندی، نمونه‌های درون یک خوشه یکسان به‌وسیله درجه بالایی از تشابه مشخص می‌شوند، درحالی‌که نمونه‌های متعلق به خوشه‌های مختلف با درجه بالایی از عدم تشابه مشخص می‌شوند. سپس این خوشه‌های همگن توسط ماتریس c×p بعدی نمایش داده می‌شوند که متشکل از c نمونه شاخص نماینده هر خوشه است (Guler and Thyne, 2004).

فرآیندهای اصلی، همیشه خروجی‌های دقیقی تولید نمی‌کنند و در سیستم‌های هیدروشیمی و هیدرولوژیکی پیچیده به‌سختی تشخیص داده می‌شوند. خصوصیات فیزیکی و شیمیایی در سیستم‌های طبیعی اغلب نه ناگهانی بلکه به‌طور پیوسته تغییر می‌کنند. به‌واسطه این پیوستگی، خوشه‌های آماری نمی‌توانند جداکننده خوبی باشند و بایستی توالی خوشه‌های هم‌پوشانی شکل گیرد (Guler et al., 2002; Goyal and Gupta, 2014).

منطق فازی[3] برای اولین بار به همین نام توسط Zadeh (1965) ارائه شد و بسیاری از مفاهیم، متغیرها و سیستم‌های مبهم را صورت‌بندی ریاضی بخشید. بدین ترتیب زمینه را برای استدلال، استنتاج، کنترل و تصمیم‌گیری در شرایط عدم اطمینان فراهم آورد (طاهری، ۱۳۷۸). یکی از پرکاربردترین روش‌های حل مسائل خوشه‌بندی فازی روش c- میانگین فازی است (Bezdek, 1981).

قوانین فازی، قوانین ساده و آشکار برای توصیف پاسخ‌دهی مطلوب سیستم به‌جای فرمول‌های ریاضی از متغیرهای زبانی استفاده می‌کنند. جالب اینجاست که اگرچه سیستم‌های فازی، پدیده‌های غیردقیق و نامشخص را توصیف می‌کنند؛ با این‌حال تئوری فازی یک تئوری دقیق است. متغیر زبانی[4]، متغیری است که مقادیرش کلمات یا جملات یک‌زبان طبیعی باشد. به‌عنوان نمونه سن یک فرد را در نظر بگیرید؛ اگر مقادیری را که سن اختیار می‌کند با کلماتی مثل نونهال، نوجوان، جوان، مسن و پیر نشان دهیم؛ متغیر سن یک متغیر زبانی است (آذر و فرجی، ۱۳۸۶).

در نظریه کلاسیک، تابع عضویت مجموعه‌ای است که عدد یک در مرز آن و صفر بیرون آن قرار می‌گیرد. برای یک مجموعه فازی، تابع عضویت به‌صورتی تعریف می‌گردد که در فاصله بین صفر تا یک است (رابطه ۱) (تشنه‌لب و همکاران، ۱۳۸۷).

$$\mu A: X \rightarrow [0, 1] \quad \text{رابطه (۱)}$$

مجموعه A برحسب تابع عضویت عبارت است از رابطه (۲):

$$A = \{(\mu A(X)), x \in X, \mu A(X) \in [0, 1]\} \quad \text{رابطه (۲)}$$

مهم‌ترین ویژگی منطق فازی در مقایسه با منطق کلاسیک این است که دانش و تجربه بشر را می‌تواند در قالب روابط ریاضی بیان نماید. این مهم باعث شده است که مسائل دنیای واقعی را به‌خوبی بتوان با استفاده از آن مدل‌سازی کرد (تشنه‌لب و همکاران، ۱۳۸۷).

در روش‌های خوشه‌بندی، تعداد و نوع معیارها ثابت و مشخص نیست، ازاین‌رو تلاش خواهد شد تا با تغییر آن‌ها به بهترین حالت ممکن دست پیدا کرد. به‌طورکلی خوشه‌بندی عبارت است از این‌که یک مجموعه مرجع مانند X شامل n داده به c دسته یا خوشه تقسیم‌بندی شود به‌طوری‌که دسته‌های مزبور دارای خصوصیات مشابهی باشند ($2 \leq c \leq n$). خوشه‌بندی را می‌توان به‌صورت فازی و غیر فازی انجام داد. الگوریتم‌های مختلفی برای خوشه‌بندی فازی و غیرفازی توسط محققین مختلف پیشنهاد شده است (Bezdek, 1981).

1. Hard
2. Soft
3. Fuzzy logic
4. Linguistic term

در خوشــه‌بندی قطعی، هر نمونــه ورودی متعلق به یک و فقط یک خوشــه است؛ بنابراین نمی‌تواند عضوی از چندین خوشه باشــد؛ به‌عبارت‌دیگر خوشه‌ها همپوشــانی ندارند. درحالی‌که در خوشــه‌بندی فازی یک نمونه می‌تواند متعلق به بیش از یک خوشــه باشــد. خوشــه‌بندی فازی به کشف مــدل فازی داده‌های موجود می‌پردازد. روش خوشــه‌بندی فازی توســط Bezdek (1981) پیشــنهاد شــد که روشی کارآمد برای خوشــه‌بندی فازی داده‌هاست و در واقع شکل گسترش‌یافته‌ای از خوشه‌بندی c- میانگین سخت است.

در شــاخه‌های مختلف علوم زمین، مطالعات متنوعی در زمینه خوشــه‌بندی فازی انجام شــده است. برای مثال شــکاری و باقرنــژاد (۱۳۸۴) به‌منظــور طبقه‌بندی خاک از روش فازی اســتفاده کردند. در این پژوهش، تعداد خوشــه بهینه تعیین شــد و نتایج آن بررســی و مقایســه شد. این پژوهش به‌صورت مطالعه موردی برای منطقه چشمه سفید کرمانشــاه انجام شــد. در مطالعه دیگری سلطانی محمدی و همکاران (۱۳۸۶) در خنثی‌ســازی اثر سنژنتیک (مربوط به تغییرات لیتوژئوشیمیایی) داده‌های ژئوشیمیایی رسوبات آبراهه‌ای در منطقه شــمال تکاب از روش خوشه‌بندی فازی استفاده کردند. در زمینه مدل‌سازی نیاز اکسیژن شیمیایی توسط روش ترکیبی شــبکه عصبی مصنوعی، عصبی فازی و خوشــه‌بندی k میانگین نیز مطالعاتی صورت گرفته است (Ay and Kisi، 2014)

Guler و همــکاران (۲۰۰۲) از روش‌هــای آمــاری چند متغیره به‌منظــور طبقه‌بندی داده‌های شــیمیایی کیفیت آب اســتفاده نمودنــد و به بررســی کیفیــت آب زیرزمینی منطقه‌ای در ترکیه با اســتفاده از ایــن روش پرداختند. در پژوهش دیگری در زمینه خوشه‌بندی C- میانگین فازی، به تخمین زمین‌لغزش ناشی از بارش با استفاده ترکیبی از روش شــبکه عصبی و خوشــه‌بندی C- میانگین فازی پرداختند (Alimohammadlou et al.، 2014).

Guler and Thyne (2004) بــا اســتفاده از داده‌هــای هیدروشیمی آب زیرزمینی منطقه‌ای در جنوب کالیفرنیا نشان دادند که روش خوشــه‌بندی c- میانگین فازی[1] (FCM) نتایج بهتری نســبت به روش‌های دیگــر دارد. در روش FCM تعداد خوشه‌ها بیشتر و مرز رخساره‌ها به‌صورت تدریجی و نه ناگهانی است که با واقعیت هماهنگی بیشتری دارد. این روش با موفقیت در دیگر علوم از جمله زیست‌شناسی (Zhang et al.، 1995)، در هواشناســی (Moore and McBratney، 1985)، در زمین‌شناسی (Rantitsch، 2004؛ Burrough et al.، 2000).

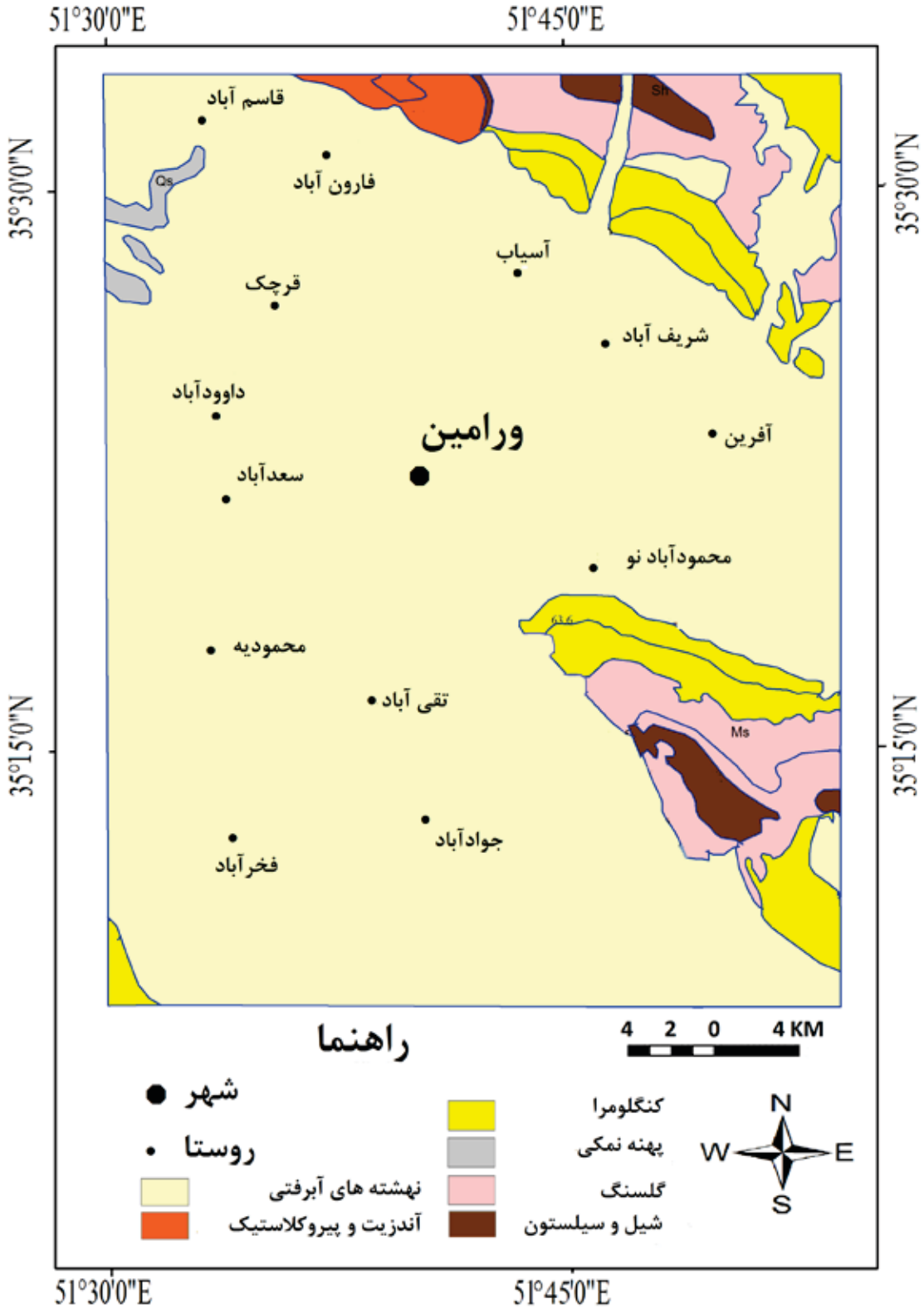
در ایــن پژوهش داده‌های برگرفته از آنالیز نمونه آب‌های زیرزمینی آبخوان ورامین با استفاده از روش‌های خوشه‌بندی C- میانگیــن و آنالیز خوشــه سلســله مراتبــی[2] (HCA) دسته‌بندی می‌شوند که کاربرد این روش‌ها در شرح تحولات هیدروشــیمی مورد بحث و بررسی قرار می‌گیرد. با توجه به این‌که تغییرات در میزان عناصر شیمیایی آب‌های زیرزمینی به‌طور پیوسته می‌باشند، استفاده از روشی که این تغییرات را به‌خوبــی نمایش دهد و میزان عضویت هر نمونه آب را به گروه‌های مختلف بیان کند؛ ضروری به‌نظر می‌رسد.

مواد و روش تحقیق

منطقه مورد مطالعه در اســتان تهران واقع شــده و از شمال به ادامه سلسله البرز و لواسانات، از شرق به ایوانکی و گرمسار، از جنوب به تپه ماهورهای جنوبی دشت ورامین و همچنیــن دریاچه نمک و از مغرب به دشــت تهران-کرج محدود می‌شود. بیشترین شیب دشــت برابر ۹/۳۲ درصد است. بیشــترین ارتفاع در بخش شمالی آن برابر ۱۰۲۰ متر بوده و کمترین آن در جنوب و جنوب شــرقی دشت و ۸۲۰ متر از ســطح دریا است. از کل منطقه موردمطالعه، حدود ۲۱/۴ کیلومترمربع را نمــک‌زار و حدود ۰/۶۸ کیلومترمربع را زمین‌های کشــاورزی تشــکیل می‌دهنــد. میانگین ۲۵ ســاله درجه حرارت سالانه، ۱۶ درجه سانتی‌گراد و متوسط بارندگی ســالانه، ۱۷۲ میلی‌متر در ایســتگاه ســینوپتیک شهرســتان ورامین است. از نظر زمین‌شناسی، این منطقه به‌واســطه فروافتادگــی دوران اول و دوم ایجــاد شــده که داخل آن را رســوبات دوران ســنوزوئیک و رسوبات جدید رودخانه جاجرود پر کرده است (آقانباتی، ۱۳۸۵). واحدهای تشکیل‌دهنده منطقه موردمطالعه به‌ترتیب سن از جدید به قدیم شــامل رســوبات دوران کواترنری، کنگلومرا، شیل،

1. Fuzzy C-Mean clustering (FCM)
2. Hierarchal clustering analysis

سیلتســتون، گل‌ســنگ، آهک، مارن، گدازه‌های آندزیتی و گدازه‌های پیروکلاســتی می‌شــوند. در شــکل ۱، نقشــه سنگ‌شــناختی و موقعیت جغرافیایی منطقه مورد مطالعه آورده شده است.



شکل ۱. نقشه سنگ‌شناختی منطقه موردمطالعه (برگرفته از سازمان زمین‌شناسی کشور، نقشه یک صدهزارم ورامین، ۱۳۸۵)

در ایـن پژوهش، از میان متغیرهای اندازه‌گیری شـده در ۹۰ نمونـه چاه آب زیرزمینی دشـت ورامین که توسـط آب منطقه‌ای اسـتان تهران انجام شـده است، متغیرهای بی‌کربنات، سـولفات، کلراید، کلسـیم، منیزیم، سـدیم، پتاسیم، اسیدیته و هدایت الکتریکی استفاده شدند. خطای موازنـه بار یونی داده‌ها کمتر از ۵٪ اسـت کـه برای ارزیابی کیفیت آب زیرزمینی مناسـب اسـت. ماتریـس داده‌های ورودی بـرای روش FCM در محدوده ۳+ تا ۳- با میانگین صفر استاندارد شـدند و مشکل مقایسه متغیرها با واحدها مختلف رفع شد.

الگوریتم خوشـه‌بندی FCM روش تحلیـل داده چند متغیره اسـت. این الگوریتم، مجموعه‌ای از داده‌ها (فضای چندبعدی اقلیدسی (p) را به خوشه فازی (C) تقسیم می‌کند که به‌وسیله مراکز خوشه‌هایشان (Vi) تعیین می‌شوند. جدا کردن داده‌ها به خوشـه‌های فازی با بهینه کردن تابع هدف در رابطه ۳ به‌دست می‌آید:

رابطه (۳) $J_{FCM}=(M,C)=\sum_{i=1}^{c}\sum_{K=1}^{n}U^{m}_{iK}\| X_K-V_i\|^2$

در رابطـه ۳، M ماتریـس اعضـاء، C ماتریـس مراکز خوشه‌ها، c تعداد خوشـه‌ها، n تعداد داده‌ها و U^{m}_{ik} درجه عضویت داده k در خوشـه i است. اگر فاصله اقلیدسی برای هریک از نمونه‌های آب زیرزمینی (فاصله خط مستقیم بین دو نقطـه در فضـای p بعدی که به‌وسـیله p متغیر تعریف می‌شـود) بین داده Xk که شامل ماتریسـی از داده‌های ۹ متغیر آب زیرزمینی نمونه‌برداری شـده است و مرکز خوشه Vi بالا باشد؛ J_{FCM} کم می‌شود (کوره‌پزان دزفولی، ۱۳۸۷). اگر فاصله کم باشـد؛ مقدار عضویت بـه یک میل می‌کند. پارامتر $m\in(1,\infty)$ پارامتر وزنی اسـت که درجه فازی بودن نتایـج طبقه‌بنـدی را کنترل می‌کند که درجه هم‌پوشـانی بین خوشـه‌ها است. با حداقل مقدار معنی‌دار m=1 پاسخ، تقسـیم‌بندی از نوع غیرفازی اسـت. با میـل m به مقدار بی‌نهایت، پاسـخ به بالاترین درجه فازی‌شدگی میل می‌کند (Bezdek، 1981). انتخاب مقدار m=2 در بسیاری حالات به‌عنوان انتخاب مناسـب پارامتر فازی‌شدگی پذیرفته شده اسـت (Hathaway and Bezdek، 2001). ماتریس M به اجزاء [0،1] محدود می‌شود (رابطه ۴):

رابطه (۴) $\sum_{i=1}^{c}U^{m}_{iK}=1\ ,1\le k\le n$

محدودیـت دوم در ماتریـس M در این اسـت که مجموع ضرایب مرکز هر خوشه باید کمتر از تعداد اجزاء باشد (رابطه ۵):

رابطه (۵) $\sum_{k=1}^{n}U_{iK}<n\ ,1\le i\le c$

در نتایـج FCM ممکن اسـت نمونه‌ها ۱۰۰ درصد عضو خوشـه نباشـند؛ در عوض عضویت نمونه‌ها بین خوشه‌ها تقسیم می‌شود.

یکی از مشـکلات روش‌های خوشه‌بندی، تعیین تعداد خوشـه‌ها است. بعضی از روش‌های اجتناب از این مشکل، اسـتفاده از دانش فردی یا روش‌های آماری اسـت. در این پژوهش، تعداد خوشـه‌های بهینه با استفاده از روابط ۶ و ۷ حاصل می‌شود. شاخص FPI درجه (رابطه ۶) فازی‌شدگی را برای تعداد خوشه‌های مشخص نشان می‌دهد (کوره‌پزان دزفولی، ۱۳۸۷):

رابطه (۶) $F=\frac{1}{n}\sum_{i=1}^{c}\sum_{k=1}^{n}(U_{ik})^2 \quad FPI=1-\frac{cF-1}{c-1}$

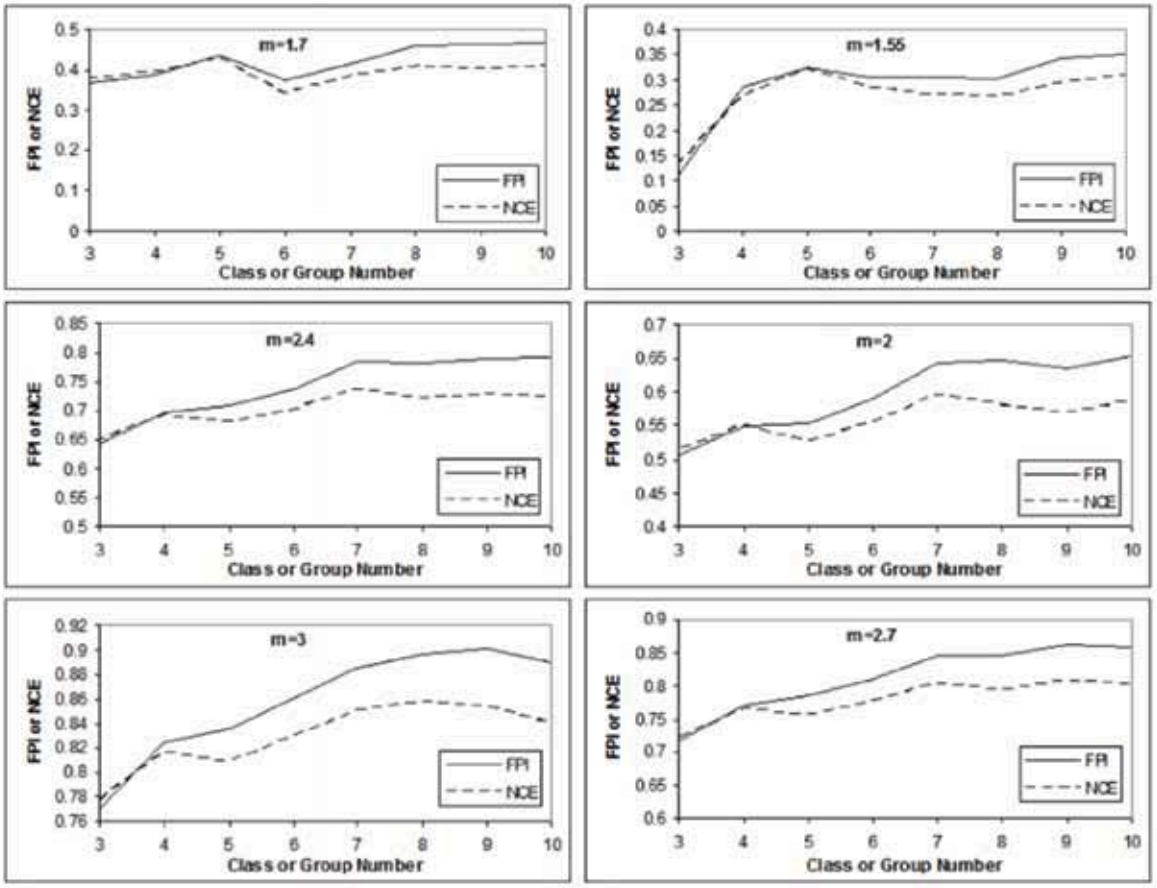
شـاخص مهم دیگری که درجـه به‌هم‌ریختگی را برای تعداد گروه‌های مشخص نشـان می‌دهد؛ NCE (رابطه ۷) است (کوره‌پزان دزفولی، ۱۳۸۷):

رابطه(۷) $F=\frac{1}{n}\sum_{i=1}^{c}\sum_{k=1}^{n}U_{ik}\times\log(U_{ik}) \quad NCE=\frac{H}{\log c}$

که در آن، H تابع انتروپی[1] اسـت. تعداد خوشه بهینه با کمینه کردن این دو شاخص به‌دست می‌آید.

در خوشه‌بندی به روش فازی، تعداد خوشه، نوع فاصله، معیار توقف و پارامتر فازی‌شـدگی بایستی در شروع تحلیل معین شد. به همین منظور، فاصله اقلیدسی به‌عنوان فاصله انتخاب شـد؛ داده‌ها اسـتاندارد شـدند تا وزن یکسانی به تمام متغیرها تعلق گیرد و مقدار توقف ۰/۰۰۰۱ تعیین شـد. خوشـه‌بندی برای c خوشـه (بین ۳ و ۱۰) و با اسـتفاده از مقادیـر m (۳، .....، ۱۰/۱۵، ۱/۱) و بـرای ترکیب‌های c و m درمجموع ۳۱۲ بار اجرا شـد. بـرای انتخاب تعداد گروه بهینه از توابع (FPI، NCE) اسـتفاده می‌شود. برخی نتایج در شـکل ۲ آمده اسـت که اغلب نمودارها الگوی همسانی را نشـان می‌دهند. تعداد خوشه بهینه در اغلب نمودارها ۳ خوشه به‌دست‌آمده است.

1. Entropy function

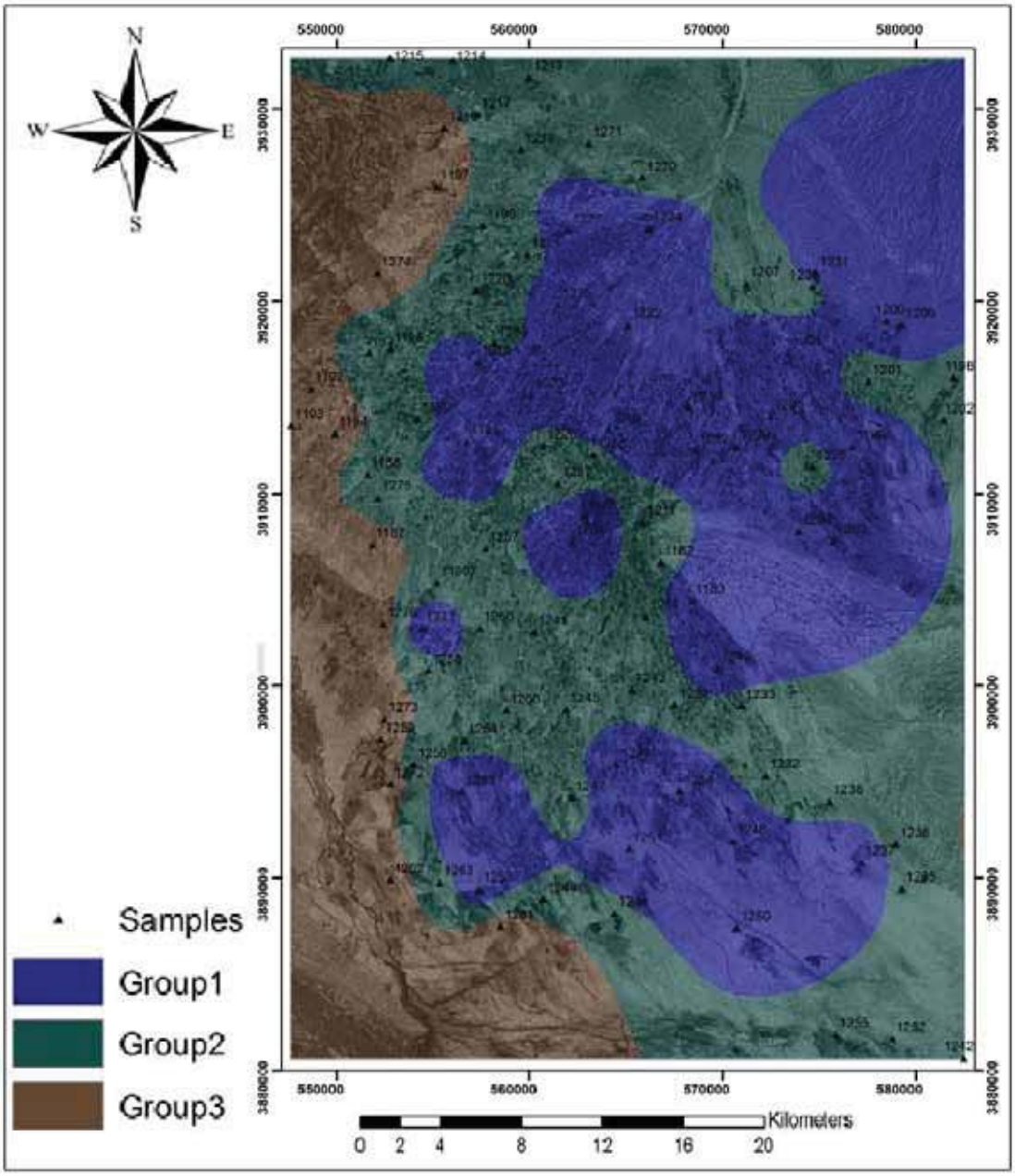


شکل ۲. نمودار مقادیر شاخص FPI و NCE در برابر تعداد خوشه

نتایج و بحث

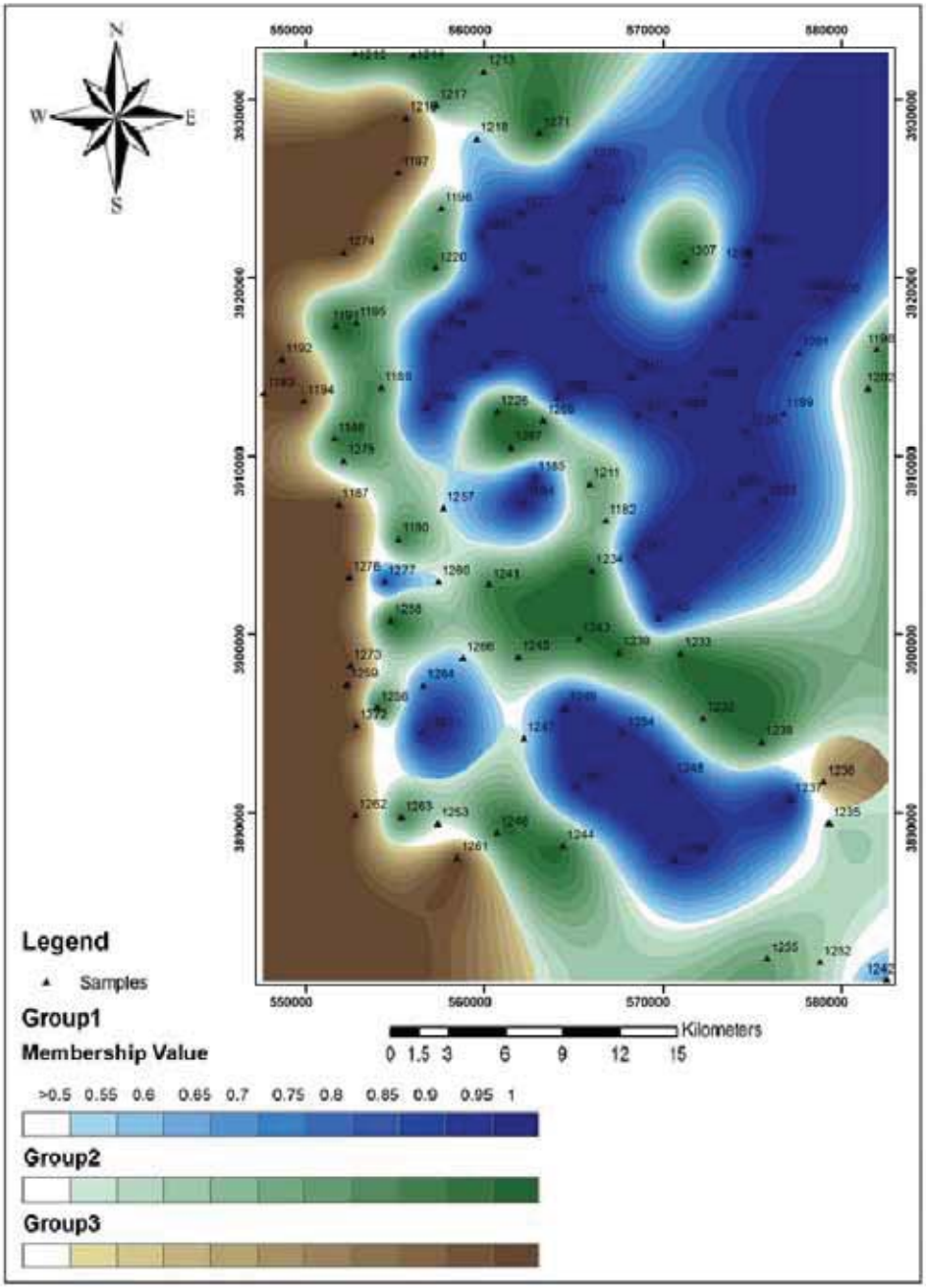
فرآیندهای هیدروژئولوژیک، بسیار پیچیده هستند، از این‌رو نمی‌توان با استفاده از روش‌های مرسوم به بررسی کیفیت آب زیرزمینی در شرایط هیدروژئولوژیک پیچیده پرداخت و محققین در مطالعات هیدرولوژی و زمین‌شناسی به روش‌های جدید و دقیق‌تری روی آورده‌اند. یکی از روش‌های پرکاربرد در زمینه ارزیابی کیفیت آب و بررسی روند تغییرات آب در آبخوان، روش‌های آماری چندمتغیره و روش‌های مختلف خوشه‌بندی است. خوشه‌بندی تجمعی برای این پژوهش انتخاب شد و روش وارد[1] به دو دلیل برای انتخاب خوشه‌ها و فاصله اقلیدسی به‌عنوان معیار فاصله انتخاب شد. اول این‌که داده‌های مورد استفاده در این مطالعه از نوع داده‌هایی با مقیاس فاصله‌ای است (داده‌های استاندارد شده) و دلیل دوم این است که روش وارد به لحاظ ساختار خوشه‌بندی شبیه به روش c- میانگین است و این موضوع باعث می‌شود، اختلافات ساختاری دو روش تأثیری در مقایسه خروجی‌ها نداشته باشند. شکل ۳ نتایج روش HCA را نشان می‌دهد تا روند تغییرات بهتر نمایش داده شود.

1. Ward method



شکل ۳. توزیع مکانی رخساره‌های هیدروشیمی به روش HCA

همان‌طورکـه در شـکل ۳ دیده می‌شـود؛ مـرز بین گروه‌هـا در این روش کاملاً واضح و ناگهانی تغییر می‌کند، به‌طوری‌کـه یک نمونه یا به یک گروه تعلق دارد یا خیر. در شکل ۴ نتایج روش FCM نشان داده شده تا روند تغییرات بهتر نمایش داده شـود. تغییرات مقادیر عضویت در بازه ۱ و ۰/۵ در خوشـه‌های مختلف نشـان داده شده است. در این شـکل برای نمایش بهتر تغییرات عضویت در خوشه‌ها از نقشـه منطقه صرف‌نظر شده است. مناطقی که با رنگ سـفید در شکل مشخص‌شده اسـت مربوط به نمونه‌هایی اسـت که متعلق به هیچ یک از خوشـه‌ها نیستند؛ یعنی مناطقـی که مقدار عضویت آن‌ها به هریک از گروه‌ها کمتر از ۰/۵ است.



شکل ۴- توزیع مکانی رخساره‌های هیدروشیمی به روش FCM

الگوریتم خوشه‌بندی FCM، روشی مؤثر برای یافتن ساختار خوشه‌بندی داده‌های پیچیده دارای هم‌پوشانی و مقادیر مبهم است (Guler and Thyne، 2004). توزیع مکانی نتایج خوشه‌بندی به روش FCM، هماهنگی معناداری با فرآیندهای فیزیکوشیمیایی رخساره‌های هیدروشیمی آب نشان می‌دهد. با مقایسه دو روش مورد استفاده در این پژوهش مشخص می‌شود که انطباق خوبی بین تفسیر کمی و کیفی داده‌های آب زیرزمینی در روش FCM نسبت به روش HCA وجود دارد. همچنین در روش HCA تعداد خوشه به‌صورت سعی و خطا تعیین می‌شود؛ درحالی‌که در روش FCM توسط توابع هدف (روابط ۶ و ۷) مشخص می‌گردد. همان‌طورکه در شکل ۴ مشخص است، توزیع مکانی خوشه‌های FCM ارتباط نزدیکی با رخساره‌های هیدروشیمی دارد. نتایج حاصل از خوشه‌بندی داده‌ها سه

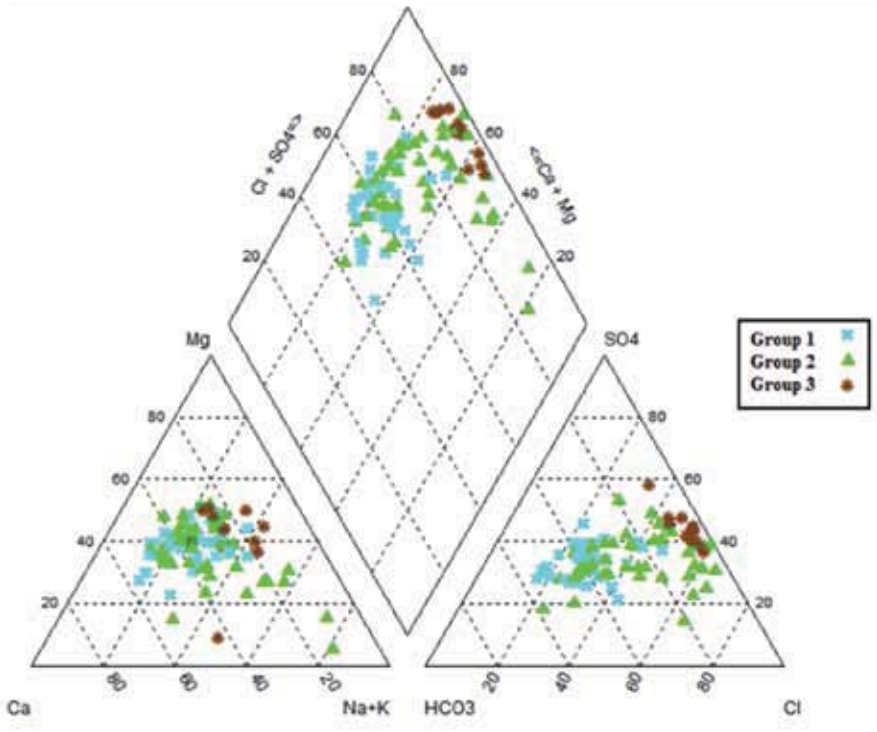
گروه معرف را نشان می‌دهد که نشان‌دهنده وجود سه تیپ آب زیرزمینی مختلف در ایجاد این گروه‌ها است.

گروه اول که با رنگ آبی در تصاویر نشان داده شده است، آب‌هایی با کیفیت بالا، املاح کم و با میانگین پایین مجموع مواد جامد محلول (۴۰۰ میلی‌گرم در لیتر) است. این گروه در طبقه‌بندی آب‌ها بر اساس TDS (Todd and Mays، 2005) در گروه آب شیرین قرار می‌گیرد که عمدتاً مربوط به منطقه تغذیه و مناطقی می‌باشد که تحت تأثیر آلودگی‌ها قرار نگرفته‌اند.

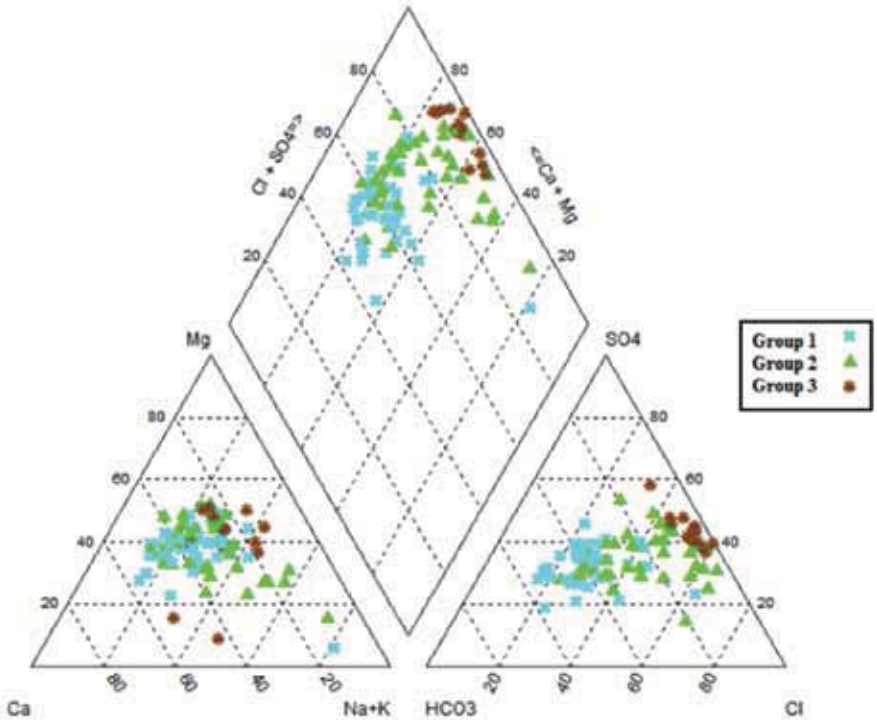
گروه دوم که با رنگ سبز در تصاویر نشان داده شده است؛ آب‌هایی با کیفیت نسبتاً پایین که میانگین مواد جامد محلول در آن ۱۰۰۰ میلی‌گرم در لیتر است. از نظر مقدار TDS، این نمونه‌ها در گروه آب‌های لب‌شور قرار دارند که برای مصارف کشاورزی و صنایعی که نیاز به آب‌های با کیفیت بالا ندارند، مناسب است. تشکیل این گروه را می‌توان نتیجه چند فرآیند دانست که به نظر می‌رسد به‌دلیل اختلاط آب‌های منطقه تغذیه و آب‌های شور در غرب محدوده موردمطالعه، ارتباط پساب صنعتی و کشاورزی در مناطقی که سطح آب زیرزمینی بالاست و برخورد آب زیرزمینی با پهنه‌های نمکی و سولفاته محلی (بر روی شکل ۱، با Qs نمایش داده شده) است.

گروه سوم که با رنگ قهوه‌ای در شکل‌ها مشخص‌شده است. آب‌هایی با کیفیت پایین هستند که میانگین مواد جامد محلول در آن ۴۶۰۰ میلی‌گرم در لیتر است. آب‌های این گروه از لحاظ مقدار TDS در گروه آب‌های شور و خیلی شور قرار می‌گیرند که منشأ آن سازندهای نمکی و سولفاته غرب محدوده موردمطالعه است.

استفاده از نمودار پایپر به‌عنوان اصلی‌ترین نمودار در شناسایی رخساره‌ها یا گروه‌های هیدروشیمی در این مطالعه، امری اجتناب‌ناپذیر به‌نظر می‌رسد. نمودار پایپر می‌تواند در سنجش عملکرد روش‌های خوشه‌بندی به‌عنوان یک شاخص مورد استفاده قرار گیرد. در ادامه، رفتار دو روش خوشه‌بندی در تعلق داده‌ها به گروه‌های هیدروشیمی توسط نمودار پایپر مورد ارزیابی قرار می‌گیرد. در شکل ۵ عملکرد روش خوشه‌بندی HCA در نمودار پایپر نمایش داده می‌شود.



شکل ۵. نمودار پایپر برای داده‌های خوشه‌بندی شده به‌روش HCA



شکل ۶. نمودار پایپر برای داده‌های خوشه‌بندی شده به روش FCM

همان‌طورکه مشاهده می‌شود بی‌نظمی قابل ملاحظه‌ای در گروه‌بنـدی داده‌هـا به روش HCA دیده می‌شـود. این موضوع در تفکیک گروه‌های۱ و ۲ بهتر مشخص است. شکل ۶ نمودار پایپر را برای داده‌های خوشـه‌بندی شده به روش FCM نشـان می‌دهد، چنانچه در شکل دیده می‌شود؛ در این روش جدایش بهتری در نمونه‌ها دیده می‌شود.

با توجه به شـکل ۶، نمونه‌هـای گروه یک به‌طور عمده در منطقه اختلاط و منطقه بدون تیپ غالب قرار می‌گیرند. همان‌طورکـه در شـکل شـماره ۴ مشـاهده می‌شـود در بخش‌های شمال شرقی منطقه موردمطالعه اغلب نمونه‌ها در گـروه یک قرار گرفته‌انـد. در گروه دو نیز اغلب، نمونه‌ها در منطقـه اختلاط قرار می‌گیرند اما در برخی نمونه‌ها تیپ کلراید سدیک نیز مشاهده می‌شود که به‌نظر می‌رسد به‌دلیل وجود پهنه‌های نمکی در منطقه موردمطالعه باشد. گروه سه نیز اغلب نمونه‌ها تیپ سـولفاته منیزیـک دارند که به‌دلیل انحلال سازندهای آبرفتی و پهنه‌های نمکی و ژیپسی در غرب منطقه موردمطالعه است. با بررسی روند تغییرات گروه‌های یک، دو و سـه در شکل شـماره ۴ با نمودار پایپر، مشخص می‌شـود، روش FCM در ارزیابـی تحولات هیدروشـیمی منطقه موردمطالعه نسبت به روش HCA برتری دارد. بدین ترتیب که روش FCM نسـبت به روش HCA روند تحولات هیدروشیمی منطقه موردمطالعه را بهتر نشان می‌دهد. در نمودار پایپر نیز می‌توان روند تغییرات هیدروشـیمی منطقه موردمطالعه را از گروه یک به گروه دو و در نهایت گروه سـه مشاهده نمود. به‌صورت کلی می‌توان گروه یک را نمونه‌های ابتدایی دشت، گروه دو را نمونه‌های میان دشتی و گروه سه را نمونه‌های خروجی و انتهایی خوشه‌بندی نمود.

نتیجه‌گیری

در این پژوهش تلاش شـد تا با بررسـی و مقایسـه دو روش پرکاربرد از روش‌های خوشـه‌بندی، روشـی مناسـب بـرای تعیین رخسـاره‌های هیدروشـیمی انتخاب شـود. الگوریتم خوشه‌بندی FCM روشی مؤثر برای یافتن ساختار خوشـه‌بندی داده‌های پیچیده دارای هم‌پوشـانی و مقادیر

مبهم اســت. توزیــع مکانی نتایــج خوشــه‌بندی در روش FCM با فرایندهای فیزیکوشیمیایی هماهنگی معناداری با رخساره‌های هیدروشیمی آب نشان می‌دهد. با مقایسه توزیع مکانی رخســاره‌های هیدروشیمی و نمودار پایپر برای هر دو روش، نتایج نشــان می‌دهــد؛ در روش FCM توزیع مکانی و خطای خوشه‌بندی نســبت به روش HCA کمتر است و خوشه‌بندی فازی، فرایندهای ایجادکننده تغییرات طبیعی پیوسته در منطقه را بهتر نشان می‌دهد و مقدار زیادی داده را می‌تواند بررسی و تحلیل نماید. قرار دادن نمونه‌های آب در گروه‌هایی با مرزهای مشخص باعث ایجاد خطای طبقه‌بندی داده‌ها می‌شــود که از اشکالات روش HCA است. در روش خوشه‌بندی فازی با تخصیص هر نمونه به چند گروه، خطای طبقه‌بندی به حداقل می‌رســد. به‌منظور صحت‌سنجی، از نمودار پایپر اســتفاده شــد. در نمودار پایپر ترسیم شده به روش HCA بی‌نظمی در گروه‌بندی داده‌ها دیده می‌شــود. این موضوع در تفکیک گروه‌هــای یک و گروه دو بهتر دیده می‌شــود؛ اما در نمودار پایپر ترسیم شــده به روش FCM گروه‌ها نظم بهتری دارند و نمونه‌هایی با رخساره هیدروشیمی یکسان، با هم در یک گروه قرار گرفتند. به‌طورکلی می‌توان نتیجه گرفــت؛ در روش HCA تغییرات در مرز بین گروه‌ها کاملاً واضح و ناگهانی اســت امــا در روش FCM، تغییرات تدریجی است.

منابع

- آذر، ع. و فرجــی، ح.، ۱۳۸۶. علــم مدیریت فازی. انتشارات موسسه کتاب مهربان نشر.
- آقانباتی، ع.، ۱۳۸۵. زمین‌شناسی ایران. انتشارات سازمان زمین‌شناسی و اکتشافات معدنی کشور.
- تشــنه‌لب، م.، صفارپــور، ن. و افیونی، د.، ۱۳۸۷. سیستم‌های فازی و کنترل فازی. انتشارات دانشگاه صنعتی خواجه‌نصیرالدین طوسی.
- ســازمان زمیــن شناســی کشــور، ۱۳۸۵. نقشــه زمین‌شناسی یک صدهزارم ورامین، سازمان زمین‌شناسی و اکتشافات معدنی کشور.
- سلطانی محمدی، س.، علی نیا، ف. و آویانی، د.، ۱۳۸۶. خنثی‌ســازی اثر ســنژنتیک داده‌های ژئوشیمیایی رســوبات آبراهــه‌ای در منطقه شــمال تکاب با اســتفاده از خوشه‌بندی فازی. امیرکبیر، ۶۷، ۲۷-۳۴.
- شــکاری، پ. و باقر نژاد، م.، ۱۳۸۴. بررسی کاربرد روش فــازی در طبقه‌بندی خاک، مطالعه موردی: چشــمه سفید کرمانشــاه. علوم و فنون کشــاورزی و منابع طبیعی زمستان، ۹، ۵۵-۶۸.
- طاهری، م.، ۱۳۷۸. آشــنایی با نظریه مجموعه‌های فازی. انتشارات جهاددانشگاهی مشهد.
- کوره‌پــزان دزفولــی، ا.، ۱۳۸۷. اصــول تئــوری مجموعه‌هــای فــازی و کاربرد آن در مدل‌ســازی مســائل مهندســی آب. انتشــارات جهاددانشــگاهی واحد صنعتی امیرکبیر.

- Alimohammadlou, Y., Najafi, A. and Gokceoglu, C., 2014. Estimation of rainfall-induced landslides using ANN and fuzzy clustering methods: A case study in Saeen Slope, Azerbaijan province, Iran. Catena, 120, 149-162.
- Ay, M. and Kisi, O., 2014. Modelling of chemical oxygen demand by using ANNs, ANFIS and k means clustering techniques Journal of Hydrology, 511, 279-289.
- Barbieri, P., Adami, G., Favretto, A., Lutman, A., Avoscan, W. and Reisenhofer, E., 2001. Robust cluster analysis for detecting physicochemical typologies of freshwater from wells of the plain of Friuli (northeastern Italy). Analytica Chimica Acta, 440, 161-170.
- Bezdek, J.C., 1981. Pattern Recognition with Fuzzy Objective Function Algorithms Plenum, 256.
- Burrough, P.A., Van Gaans, P.M. and McMillan, R.A., 2000. High resolution landform classification using fuzzy k-means. Fuzzy Sets Systems, 113, 37-52.
- Goyal, M.K. and Gupta, V., 2014. Identification of Homogeneous Rainfall Regimes in Northeast Region of India using Fuzzy Cluster Analysis. Water Resources Management, 28, 4491-4511.
- Guler, C. and Thyne, G.D., 2004. De-

lineation of hydrochemical facies distribution in a regional groundwater system by means of fuzzy c-means clustering. Water Resource Research, 40, w12503. Doi: 10.1029/2004WR003299.

- Guler, C., Thyne, G.D., McCray, J.E. and Turner, A.K., 2002. Evaluation of graphical and multivariate statistical methods for classification of water chemistry data. Hydrogeology Journal, 10, 455- 474.

- Hathaway, R.J. and Bezdek, J.C., 2001. Fuzzy C-means clustering of incomplete data. IEEE Transactions System Man and Cybernetics, 31, 735- 744.

- McBratney, A.B. and Moore, A.W., 1985. Application of fuzzy sets to climatic classification. Agriculture and Forest Meteorology, 35, 165- 185.

- Rantitsch, G., 2000. Application of fuzzy clusters to quantify lithological background concentrations in stream-sediment geochemistry. Journal of Geochemical Exploration, 71, 73- 82.

- Singh, K., Malik, A., Mohan, D. and Sinha, S., 2004. Multivariate statistical techniques for the evaluation of spatial and temporal variations in water quality of Gomti River (India)-a case study. Water Research, 38, 3980-3992.

- Todd, K.D., and Mays, 2005. Groundwater Hydrology. John Wiley and Sons. 636.

- Zadeh, L.A., 1965. Fuzzy sets. Information Control, 8, 338- 353.

- Zhang, C.T., Chou, K.C. and Maggiora, G.M., 1995. Predicting protein structural classes from amino acid composition: Application of fuzzy clustering. Protein Engineering, 8, 425-435.