

## تبیین تحولات رخساره‌های هیدروشیمیایی آبخوان سراب با استفاده از روش‌های خوشه‌بندی میانگین فازی و تحلیل خوشه سلسله‌مراتبی

میثم ودیعتی<sup>۱</sup>، اصغر اصغری مقدم<sup>۲\*</sup>، محمد نخعی<sup>۳</sup>

۱. دانشجوی دکتری، گروه علوم زمین، دانشکده علوم طبیعی، دانشگاه تبریز، تبریز

۲. استاد، دانشکده علوم طبیعی، دانشگاه تبریز، تبریز

۳. استاد، دانشکده علوم زمین، دانشگاه خوارزمی تهران، تهران

(تاریخ دریافت ۱۳۹۵/۱۱/۰۱؛ تاریخ تصویب ۱۳۹۵/۱۲/۲۵)

### چکیده

در پژوهش حاضر، خوشه‌بندی مجموعه‌ای از داده‌های هیدروشیمیایی آبخوان سراب با استفاده از روش‌های خوشه‌بندی C- میانگین فازی (FCM) و تحلیل خوشه سلسله‌مراتبی (HCA) انجام شده و کاربرد آنها در تبیین رخساره‌های هیدروشیمیایی مطالعه شد. خوشه‌های آماری شباهت مکانی را بررسی می‌کنند و نشان می‌دهند خوشه‌ها مطابقت هیدروژئولوژیکی با رخساره‌های هیدروشیمیایی آبخوان دارند. نمونه‌های آب زیرزمینی با استفاده از بهینه‌کردن تعداد خوشه و درجه فازی‌شدگی با استفاده از روش C- میانگین فازی به چهار خوشه طبقه‌بندی شدند. از داده‌های ۴۹ نمونه آب زیرزمینی و ۱۲ متغیر هیدروشیمیایی منطقه مطالعه شده استفاده شد. نتایج این دو روش مراکز خوشه را تولید می‌کند که در تشخیص فرایندهای فیزیکی و شیمیایی تغییرات هیدروشیمی منطقه مطالعه شده مؤثر است. روش FCM روشی مناسب در تحلیل داده‌ها در بیان توزیع رخساره‌های هیدروشیمیایی است. نتایج نشان داد رویه خوشه‌بندی برای تخصیص نمونه‌های شیمیایی آب زیرزمینی به گروه‌های همگن توسط روش FCM ابزاری مهم در تشخیص رخساره‌های هیدروشیمیایی آبخوان است و این روش در تحلیل داده‌های مرزی، نسبت به روش HCA که تغییراتی واضح و ناگهانی دارد؛ تواناتر است.

**کلیدواژگان:** آب زیرزمینی، خوشه‌بندی، دشت سراب، رخساره هیدروشیمیایی، منطق فازی.

## مقدمه

به‌طور کلی، فرایندهای زمین‌شناسی و هیدروژئولوژیکی پیچیده هستند و این پیچیدگی به‌طور مستقیم قوانین حاکم بر فرایند را نشان نمی‌دهد. از این‌رو، نمی‌توان با استفاده از روش‌های مرسوم کیفیت آب زیرزمینی در شرایط هیدروژئولوژیکی پیچیده را بررسی کرد و محققان 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. روش‌های آماری چندمتغیره برای تجزیه و تحلیل ویژگی‌های هیدروشیمیایی آبخوان مناسب‌اند [۱]. به‌صورت هم‌زمان در این روش‌ها از چندین پارامتر مؤثر استفاده می‌شود تا ارتباط و تأثیرپذیری بین متغیرها بهتر تحلیل شوند. خوشه‌بندی به‌منظور مجزا کردن داده‌های کلی به خوشه‌های معرف به‌کار می‌رود تا فرایندهای طبیعی تغییردهنده پارامترهای هیدروشیمیایی مشخص شوند. در این پژوهش با استفاده از داده‌های هیدروشیمیایی رخساره‌های آب زیرزمینی تعیین می‌شود. خوشه‌های معرف کمک شایانی به درک فرایندهای مؤثر بر آبخوان می‌کنند [۲].

روش‌های خوشه‌بندی به دو بخش طبقه‌بندی می‌شوند. روش‌های سخت یا کلاسیک که در آن هر عنصر فقط به یک گروه تعلق دارد و روش‌های نرم یا فازی که در آن هر عنصر با درجه عضویت‌های مختلف به همه خوشه‌ها تعلق دارد. هدف اصلی این روش‌ها چه سخت و چه فازی، تفکیک ماتریس داده با  $n$  نمونه و  $p$  متغیر به  $c$  زیرگروه همگن است. در نتیجه، نمونه‌های درون یک خوشه یکسان به‌وسیله درجه زیادی از تشابه مشخص می‌شوند؛ در حالی که نمونه‌های متعلق به خوشه‌های مختلف با درجه زیادی از عدم تشابه مشخص می‌شود. سپس این خوشه‌های همگن توسط ماتریس  $c \times p$  بعدی نمایش داده می‌شوند که متشکل از  $c$  نمونه شاخص نماینده هر خوشه است [۳].

ویژگی‌های فیزیکی و شیمیایی در سیستم‌های طبیعی اغلب به‌طور پیوسته تغییر می‌کنند. به‌واسطه این پیوستگی، خوشه‌های آماری جداکننده خوبی نیستند و باید توالی خوشه‌های همپوشانی شکل گیرد [۴].

منطق فازی<sup>۱</sup> برای اولین بار به همین نام توسط زاده

(۱۹۶۵) ارائه شد و بسیاری از مفاهیم، متغیرها و سیستم‌های بی‌دقت و 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 روش‌های پرکاربرد حل مسائل خوشه‌بندی فازی روش  $c$ - میانگین فازی است [۷].

## پیشینه تحقیق

در شاخه‌های مختلف علوم زمین، مطالعات متنوعی در زمینه خوشه‌بندی فازی انجام شده است. برای مثال، یوسفی راد و همکارانش (۱۳۹۰) از روش خوشه‌بندی فازی به‌منظور خوشه‌بندی توده سنگ استفاده کردند. نتیجه این تحقیق نشان می‌دهد که با روش خوشه‌بندی  $C$  میانگین فازی می‌توان انواع نمونه‌های سنگ را دسته‌بندی و بر اساس نیاز آنها را ارزیابی کرد. در مطالعه دیگری، عقیلی و همکارانش (۱۳۹۱) با استفاده از خوشه‌بندی فازی، ترکیب‌های مختلف با چهار داده هواشناسی شامل دمای هوا، رطوبت نسبی، سرعت باد و ساعات آفتابی روزانه، تبخیر از تشت را پیش‌بینی و بررسی کردند. در مطالعات آب زیرزمینی گولر و همکارانش (۲۰۰۲) از روش‌های آماری چندمتغیره به‌منظور طبقه‌بندی داده‌های شیمیایی کیفیت آب استفاده کردند و به بررسی کیفیت آب زیرزمینی منطقه‌ای در ترکیه با استفاده از این روش پرداختند [۳ و ۴]. در مطالعه دیگری گولر و تاین (۲۰۰۴) با استفاده از داده‌های هیدروشیمیایی آب زیرزمینی منطقه‌ای در جنوب کالیفرنیا نشان دادند روش خوشه‌بندی  $c$ - میانگین فازی (Fuzzy C-Mean Clustering (FCM)) نتایج بهتری نسبت به روش‌های دیگر دارد. در روش FCM تعداد خوشه‌ها بیشتر و مرز رخساره‌ها به‌صورت تدریجی است و ناگهانی نیست که با شرایط هیدروژئولوژیکی همخوانی بیشتری دارد. روش خوشه‌بندی  $c$ - میانگین فازی با موفقیت در دیگر علوم نیز به کار گرفته شده است [۱۰-۱۲]. مطالعات اغلب بر پایه بررسی کیفیت هیدروشیمیایی بوده است و تا کنون مطالعه کامل و جامعی در زمینه تغییر تیپ آب زیرزمینی و ترسیم آن در نمودارهای هیدروشیمیایی انجام نشده است که در پژوهش حاضر به آن پرداخته شده است.

کوه‌های آغلان‌داغ و سبلان در شمال احاطه شده است. بیشتر سنگ‌های تشکیل‌دهنده قسمت شمالی و جنوبی منطقه مطالعه شده سنگ‌هایی با سن ائوسن از جنس آندزیت، لاتیت، ایگنیمبریت و بازالت هستند.

در این پژوهش، از میان متغیرهای اندازه‌گیری شده ۴۹ نمونه منابع آب زیرزمینی دشت سراب، متغیرهای بی‌کربنات، سولفات، کلراید، کلسیم، منیزیم، سدیم، پتاسیم، فلوراید، نیترات، سیلیس، اسیدیت و هدایت الکتریکی استفاده شده است. خطای موازنه بار یونی داده‌ها که بیان‌کننده دقت آنالیز نمونه‌های مطالعه شده است؛ کمتر از ۵ درصد است که برای ارزیابی کیفیت آب زیرزمینی مناسب است. ماتریس داده‌های ورودی برای روش خوشه‌بندی سلسله‌مراتبی روشی مناسب برای تحلیل داده‌های نمونه‌های آبی و گروه‌بندی هیدروشیمیایی نمونه‌ها هستند و از نظر آماری در مطالعات هیدرولوژی و زمین‌شناسی اهمیت خاصی دارند [۱۷]. در این روش مشابه‌ترین عناصر به هم وصل می‌شوند و به‌طور متوالی بر اساس شباهت‌ها، عناصر دیگر نیز به یکدیگر وصل می‌شوند. سپس جفت نمونه‌های با بیشترین شباهت با یکدیگر ادغام می‌شوند و دندروگرام را به‌وجود می‌آورند [۱۸]. در این پژوهش از روش وارد<sup>۱</sup> و فاصله اقلیدسی به‌عنوان معیار فاصله در خوشه‌بندی تجمعی استفاده شد. انتخاب روش وارد به دو دلیل صورت می‌گیرد: اول اینکه داده‌های استفاده شده در این مطالعه از نوع داده‌های 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 سبب می‌شود اختلافات ساختاری دو روش تأثیری در مقایسه خروجی‌ها نداشته باشد.

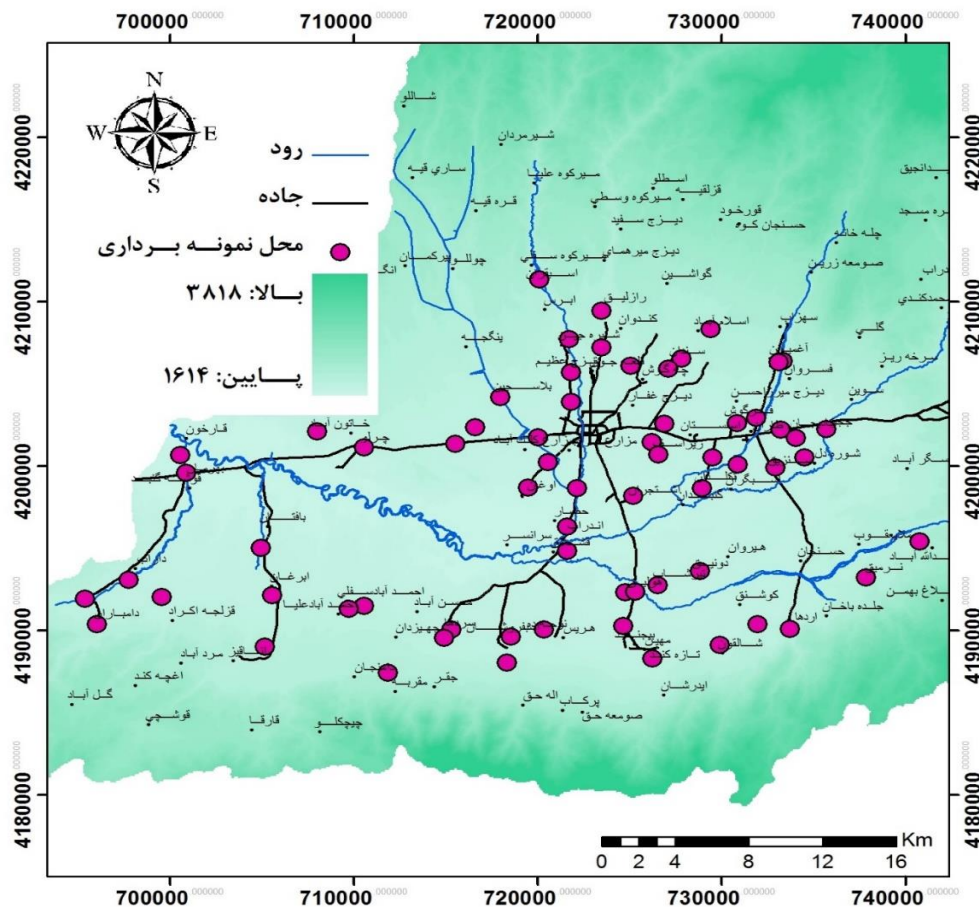
روش FCM در محدوده صفر تا یک استاندارد شدند و بدین ترتیب مشکل مقایسه متغیرها با واحدهای مختلف مرتفع شد. تئوری مجموعه‌های فازی، ابزار قدرتمند و انعطاف‌پذیر در مدل‌سازی عدم قطعیت‌های موجود در دنیای واقعی است. قوانین فازی، قوانین ساده و آشکار برای توصیف پاسخ‌دهی مطلوب سیستم، به‌جای فرمول‌های ریاضی از متغیرهای زبانی استفاده می‌کنند [۱۳].

در این پژوهش، ۴۹ نمونه آب زیرزمینی شامل چاه، چشمه و قنات استفاده شد که در خردادماه ۱۳۹۳ از دشت سراب نمونه‌برداری شده بودند و در آزمایشگاه آب‌شناسی دانشگاه تبریز آنالیز شده‌اند. با توجه به اینکه تغییرات مقدار عناصر شیمیایی موجود در آب‌های زیرزمینی پیوسته و تدریجی است؛ استفاده از روشی که روند تغییرات تدریجی کیفیت آب زیرزمینی و میزان عضویت هر نمونه را به گروه‌های مختلف بیان کند؛ ضروری به‌نظر می‌رسد.

### روش کار (مواد و روش)

منطقه مطالعه شده در شمال غربی ایران در حوضه آبریز ارومیه با وسعتی حدود ۴۲۰ کیلومتر مربع و در فاصله ۱۳۰ کیلومتری شرقی شهرستان تبریز واقع شده است. دشت سراب خود بخشی از حوضه آبریز رودخانه آجی‌چای، وناق‌چای و بیوک‌چای به حساب می‌آید. این دشت از دو سفره آزاد و تحت فشار تشکیل شده است. سفره آب زیرزمینی در نواحی شمالی و در بعضی مناطق از جمله اطراف رودخانه آجی‌چای از نوع تحت فشار است. تغییرات سطح تراز آب زیرزمینی در دشت سراب مربوط به تغذیه دشت، ناشی از نزولات جوئی، جریان‌های زیرزمینی، سیلاب‌ها و آب‌های برگشتی کشاورزی است که سبب بالا آمدن سطح آب زیرزمینی می‌شود و تخلیه آب زیرزمینی دشت به دلیل بهره‌برداری از چاه‌ها، تخلیه قنات و چشمه‌هاست که سبب پایین رفتن سطح آب زیرزمینی در سفره می‌شود؛ بنابراین تغییرات کیفیت آب زیرزمینی در منطقه مطالعه شده چشمگیر است. در شکل ۱ نقشه محل نمونه‌برداری از منابع آب زیرزمینی دشت سراب آورده شده است. در این پژوهش ۴۹ نمونه آب زیرزمینی شامل چاه، چشمه و قنات در خردادماه ۱۳۹۳ از دشت سراب نمونه‌برداری و استفاده شد. متغیرهای بی‌کربنات، سولفات، کلراید، کلسیم، منیزیم، سدیم، پتاسیم، فلوراید، نیترات، سیلیس، اسیدیت و هدایت الکتریکی که در صحرا و آزمایشگاه آب‌شناسی دانشگاه تبریز آنالیز شده بودند؛ مطالعه شد.

منطقه مطالعه شده امتداد شرقی- غربی دارد و می‌توان نسبت به ارتفاعات شمالی و جنوبی به‌صورت فروزمینی در نظر گرفت که به‌وسیله کوه‌های بزقوش در جنوب و



شکل ۱. نقشه محل نمونه برداری از منابع آب زیرزمینی دشت سراب

پیشنهاد شده است. در خوشه بندی قطعی، هر نمونه ورودی فقط متعلق به یک خوشه است و نمی تواند با دیگر خوشه ها هم پوشانی داشته باشد. در حالی که در خوشه بندی فازی هر نمونه می تواند متعلق به بیش از یک خوشه باشد. روش خوشه بندی فازی توسط بزدک ۱۹۸۱ پیشنهاد شد که روشی کارآمد برای خوشه بندی فازی داده ها است و در واقع شکل گسترش یافته ای از خوشه بندی c- میانگین سخت است.

الگوریتم خوشه بندی FCM مجموعه ای از داده ها (فضای چندبعدي اقلیدسی (p)) را به خوشه فازی (C) تقسیم می کند که به وسیله مراکز خوشه هایشان (V<sub>i</sub>) تعیین می شوند. ابزار داده ها به خوشه های فازی با کمینه کردن تابع هدف زیر به دست می آید [۱۴]:

$$J_{FCM}(M, C) = \sum_{i=1}^c \sum_{k=1}^n U_{ik}^m \|X_k - V_i\|^r \quad (۳)$$

مجموعه های کلاسیک، مرز دقیق و صریح دارند. هر عضو از مجموعه مرجع (X) یا به مجموعه A تعلق دارد (عدد یک) یا ندارد (عدد صفر). در مقابل مجموعه های فازی مرز مبهم دارند و انتقال از تعلق کامل به یک مجموعه به عدم تعلق به آن مجموعه به صورت تدریجی است [۱۳]. برای یک مجموعه فازی، تابع عضویت به صورتی تعریف می شود که در فاصله بین صفر تا یک است.

$$\mu_A : X \rightarrow [0, 1] \quad (۱)$$

مجموعه A برحسب تابع عضویت عبارت است از:

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

به طور کلی، خوشه بندی عبارت است از تقسیم بندی مجموعه مرجع مانند X شامل n داده به c دسته یا خوشه، به طوری که این دسته ها دارای ویژگی های مشابهی باشند (2 ≤ c ≤ n). خوشه بندی را می توان به دو صورت فازی و غیرفازی انجام داد که برای آن الگوریتم های مختلفی

فازی‌شدگی باید تعیین شوند. بدین‌منظور، فاصله اقلیدسی به‌عنوان فاصله انتخاب شد. داده‌ها استاندارد شدند تا وزن یکسانی به همه متغیرها تعلق گیرد و مقدار توقف  $0/0001$  تعیین شد. خوشه‌بندی برای  $c$  خوشه (بین ۳ و ۱۰) و با استفاده از مقادیر مختلف درجه فازی‌شدگی بررسی شدند تا تعداد خوشه بهینه مشخص شود. در شکل ۲، تعداد خوشه در برابر شاخص‌های FPI و NCE و برای مقادیر مختلف درجه فازی‌شدگی ( $m$ ) نشان داده شده است. اغلب نمودارها الگوی یکسانی را نشان می‌دهند. هرچه شاخص‌های FPI و NCE به هم نزدیک‌تر باشند، تعداد خوشه بهینه‌تر است. تعداد خوشه بهینه در اغلب نمودارها چهار خوشه است.

#### یافته‌ها

الگوریتم خوشه‌بندی روشی مؤثر برای یافتن ساختار خوشه‌بندی داده‌های پیچیده همپوشانی و مقادیر مبهم دارد [۴]. توزیع مکانی نتایج خوشه‌بندی به روش FCM هماهنگی معناداری با فرایندهای فیزیکوشیمیایی رخساره‌های هیدروشیمیایی آب نشان می‌دهد. هر دو روش HCA و FCM می‌توانند اندازه‌گیری‌های فیزیکی و شیمیایی را استفاده کنند، اما روش FCM می‌تواند داده‌های کیفی را با داده‌های کمی استفاده‌شده در این مطالعه تطبیق دهد. در روش HCA تعداد خوشه به‌صورت سعی و خطا تعیین می‌شود؛ در حالی که در روش FCM توسط توابع هدف مشخص می‌شود. توزیع مکانی خوشه‌های FCM ارتباط نزدیکی با رخساره‌های هیدروشیمیایی دارند و روند تغییرات کیفیت هیدروشیمیایی را بهتر نشان می‌دهند. روش FCM ماتریس داده‌های اولیه (۴۹ ردیف که همان تعداد نمونه است و ۱۲ ستون که برابر تعداد متغیر مطالعه‌شده است) را به ماتریس توابع عضویت ۴۹ در ۴ کاهش داد. ماتریس توابع عضویت در جدول ۱ تعلق هر نمونه به خوشه‌ها را نشان می‌دهد. برای مثال در نمونه ۱ (جدول ۱)، بیشترین مقدار عضویت ( $u$ ) به خوشه ۴ و برابر  $0/63416$  است و تابع هدف روش FCM نمونه ۱ را در خوشه ۴ قرار داده است. مجموع مقادیر عضویت برای هر نمونه برابر یک خواهد بود.

در رابطه ۳،  $M$  ماتریس اعضا،  $C$  ماتریس مراکز خوشه‌ها،  $c$  تعداد خوشه‌ها،  $n$  تعداد داده‌ها و  $U_{ik}$  درجه عضویت داده  $k$  در خوشه  $i$  است. اگر فاصله اقلیدسی (فاصله خط مستقیم بین دو نقطه در فضای  $p$  بعدی که به‌وسیله  $p$  متغیر تعریف می‌شود) بین داده  $X_k$  و مرکز خوشه  $V_i$  بالا باشد؛  $J_{FCM}$  کم می‌شود. اگر فاصله کم باشد؛ مقدار عضویت به سمت یک میل می‌کند. پارامتر  $m \in (1, \infty)$  پارامتر وزنی است که درجه فازی‌بودن نتایج طبقه‌بندی را کنترل می‌کند که درجه همپوشانی بین خوشه‌هاست. با حداقل مقدار معنادار  $m=1$  پاسخ، تقسیم‌بندی از نوع غیرفازی است. با میل  $m$  به مقدار بی‌نهایت، پاسخ به بیشترین درجه فازی‌شدگی میل می‌کند [۷]. انتخاب مقدار  $m=2$  در بسیاری حالات به‌عنوان انتخاب مناسب پارامتر فازی‌شدگی پذیرفته شده است [۱۶]. ماتریس  $M$  به اجزای یک و صفر محدود می‌شود:

$$\sum_{i=1}^c U_{ik} = 1, \quad 1 \leq k \leq n \quad (4)$$

محدودیت دوم در ماتریس  $M$  در این است که مجموع ضرایب مرکز هر خوشه باید کمتر از تعداد اجزا باشد یا:

$$\sum_{k=1}^n U_{ik} < n, \quad 1 \leq i \leq c \quad (5)$$

در نتایج FCM ممکن است نمونه‌ها ۱۰۰ درصد عضو خوشه نباشند؛ در عوض عضویت نمونه‌ها بین خوشه‌ها تقسیم می‌شود.

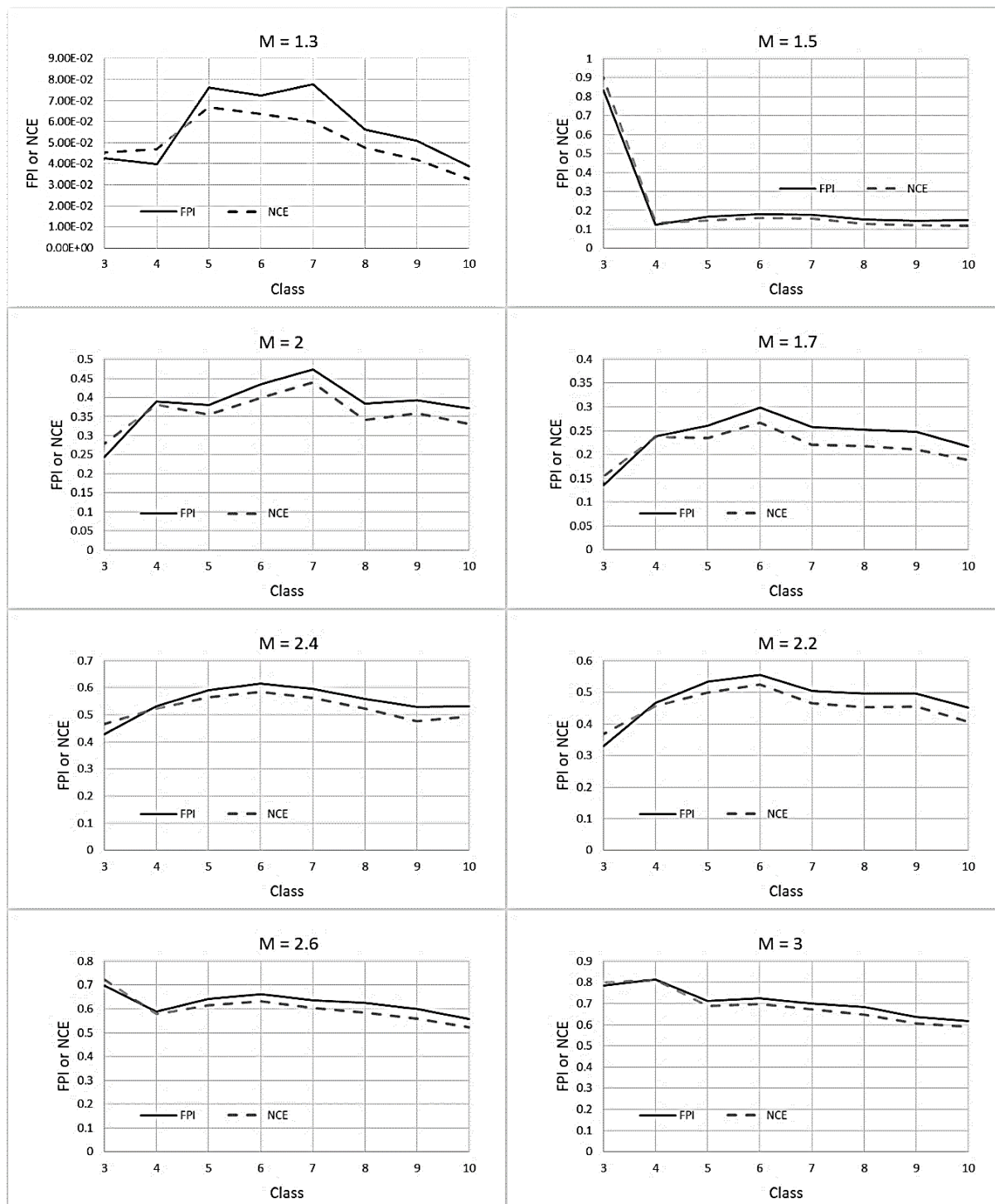
یکی از محدودیت‌های روش‌های خوشه‌بندی، تعیین تعداد خوشه‌هاست. در برخی از روش‌ها از دانش فردی یا روش‌های آماری استفاده می‌شود. در این پژوهش تعداد خوشه‌های بهینه با استفاده از رابطه‌های ۶ و ۷ به‌دست می‌آید. شاخص FPI درجه فازی‌شدگی را برای تعداد خوشه‌های مشخص نشان می‌دهد:

$$F = \frac{1}{n} \sum_{i=1}^c \sum_{k=1}^n (U_{ik})^2 \quad FPI = 1 - \frac{cF-1}{c-1} \quad (6)$$

شاخص NCE درجه به‌هم‌ریختگی را برای تعداد گروه‌های مشخص نشان می‌دهد؛

$$F = \frac{1}{n} \sum_{i=1}^c \sum_{k=1}^n U_{ik} \times \log(U_{ik}) \quad NCE = \frac{H}{\log c} \quad (7)$$

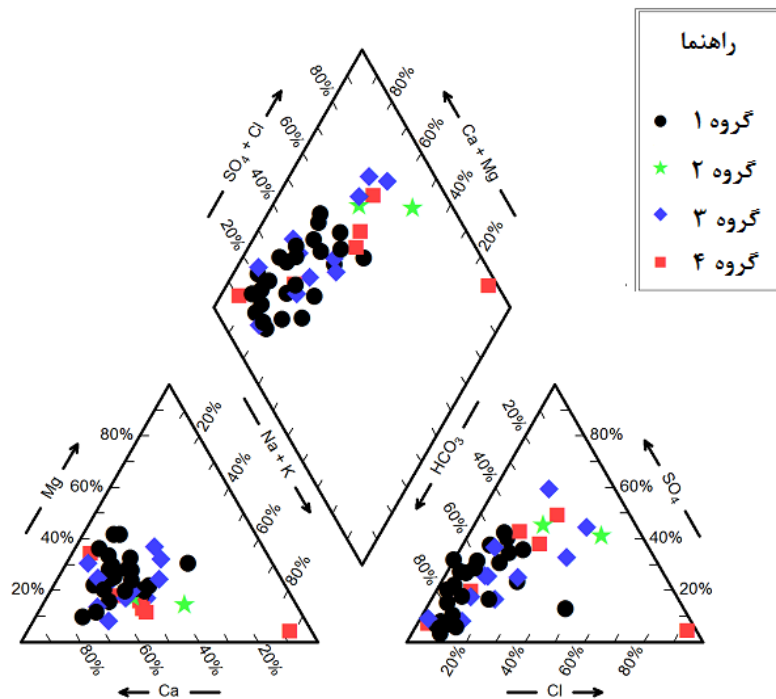
درنهایت، تعداد خوشه بهینه با کمینه‌کردن شاخص FPI و شاخص NCE به‌دست می‌آید. در خوشه‌بندی به روش فازی، تعداد خوشه، نوع فاصله، معیار توقف و پارامتر



شکل ۲. نمودار مقادیر شاخص FPI و NCE در برابر تعداد خوشه و برای مقادیر مختلف درجه فازی شدگی

می تواند در سنجش عملکرد روش های خوشه بندی به عنوان شاخصی مهم استفاده شود. در ادامه رفتار دو روش خوشه بندی در تعلق داده ها به گروه های هیدروشیمیایی، توسط نمودار پایپر ارزیابی شد. در شکل ۳ عملکرد روش خوشه بندی HCA در نمودار پایپر نمایش داده شده است.

استفاده از نمودار پایپر به عنوان اصلی ترین نمودار در شناسایی رخساره ها یا گروه های هیدروشیمیایی در این پژوهش، امری اجتناب ناپذیر به نظر می رسد. در نمودار پایپر، مقایسه هم زمان تعداد زیادی نمونه کیفیت آب زیرزمینی بر اساس غلظت نسبی امکان پذیر است [۱۹]. نمودار پایپر



شکل ۳. نمودار پایپر برای داده‌های خوشه‌بندی شده به روش HCA

جدول ۱. ماتریس توابع عضویت به دست آمده از روش خوشه‌بندی میانگین C فازی

مقدار عضویت (u)				خوشه	شماره نمونه
۱	۲	۳	۴		
۰/۰۸۷۸۸	۰/۰۴۴۴	۰/۲۳۳۵۴	۰/۶۳۴۱۶	۴	۱
۰/۳۵۰۷۸	۰/۰۱۸۲۹	۰/۴۷۱۳۷	۰/۱۵۹۵۵	۳	۲
۰/۲۸۲۳۸	۰/۳۶۵۱۵	۰/۴۵۰۰۹	۰/۲۳۱۰۰	۳	۳
۰/۱۴۶۰۵	۰/۱۲۶۳۲	۰/۲۶۶۱۵	۰/۴۶۱۴۷	۴	۴
۰/۱۳۷۲۹	۰/۰۵۴۴۷	۰/۳۲۳۸۹۳	۰/۴۸۴۳۴	۲	۵
۰/۱۷۱۱۶	۰/۰۷۱۸۷	۰/۳۱۱۴۰	۰/۴۴۵۵۵	۴	۶
۰/۱۲۴۲۴	۰/۵۲۱۵۸	۰/۱۵۹۶۵	۰/۱۹۴۵۱	۲	۷
۰/۸۹۸۷۳	۰/۰۰۶۴۳	۰/۰۶۶۷۹	۰/۰۲۸۰۳	۱	۸
۰/۹۱۴۴۳	۰/۰۰۵۱۵	۰/۰۵۶۸۰	۰/۰۲۳۶۱	۱	۹
۰/۸۰۰۱۶	۰/۰۱۵۲۸	۰/۱۲۳۹۴	۰/۰۶۰۶۱	۱	۱۰
۰/۸۸۶۴۵۸	۰/۰۰۷۴۶	۰/۰۷۴۳۷	۰/۰۳۱۷۱	۱	۱۱
۰/۳۱۵۹۷	۰/۰۴۷۳۴	۰/۳۹۴۸۱	۰/۲۴۱۸۷	۳	۱۲
۰/۰۷۸۲۳	۰/۶۴۳۳۹	۰/۱۰۹۱۴	۰/۱۶۹۲۲	۲	۱۳
۰/۱۰۳۰۳	۰/۴۳۲۳۴	۰/۱۷۵۰۴	۰/۲۸۹۵۷	۲	۱۴
۰/۱۲۲۵۹	۰/۰۶۴۱۲۸	۰/۲۸۰۶۵	۰/۵۳۲۶۱	۴	۱۵
۰/۱۳۳۸۳	۰/۱۹۸۱۳	۰/۲۳۰۰۹	۰/۴۳۷۹۴	۴	۱۶
۰/۲۷۲۵	۰/۱۲۵۴۵	۰/۳۱۷۷۹	۰/۲۸۴۲۲	۳	۱۷
۰/۰۸۰۳۹	۰/۰۶۸۹۲۵	۰/۲۰۳۵۵	۰/۶۴۷۱۲	۴	۱۸
۰/۲۶۷۴	۰/۱۲۹۶۷	۰/۳۲۶۲۰	۰/۲۷۶۶۸	۵	۱۹
۰/۱۸۶۸۷	۰/۳۵۲۸۶۷	۰/۲۲۰۲۳	۰/۲۴۰۰۲	۲	۲۰
۰/۱۰۴۳۱	۰/۰۱۰۰۰۳	۰/۷۸۷۰۷	۰/۰۹۸۶۰	۳	۲۱

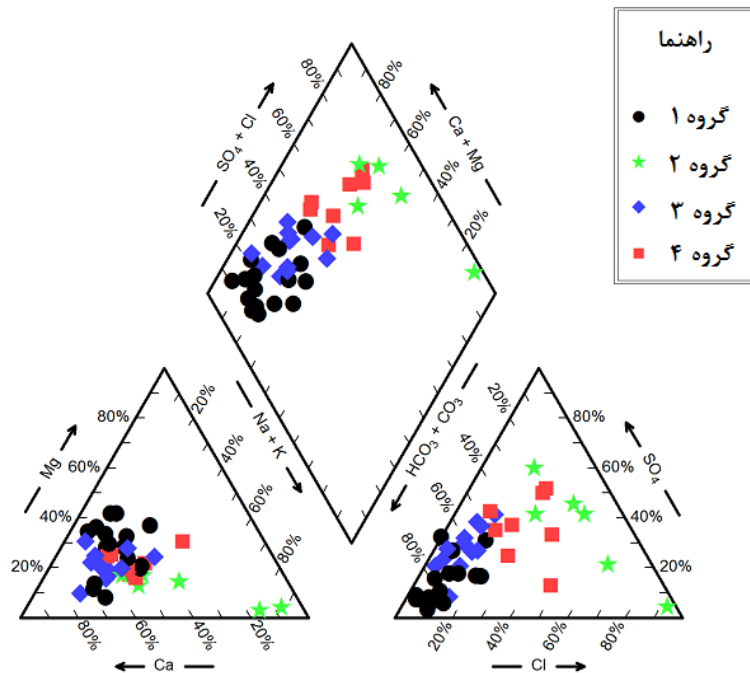
ادامه جدول ۱. ماتریس توابع عضویت به دست آمده از روش خوشه‌بندی میانگین C فازی

شماره نمونه	خوشه	مقدار عضویت (u)			
		۱	۲	۳	۴
۲۲	۳	۰/۲۳۹۳	۰/۰۳۶۰۱	۰/۴۹۷۸۵	۰/۲۲۶۸۳
۲۳	۳	۰/۳۱۶۸۵	۰/۰۱۲۵۱	۰/۳۱۰۳۲	۰/۳۴۷۶۴۳
۲۴	۱	۰/۴۹۸۱۰۴	۰/۰۱۴۳۸۹	۰/۳۸۳۰۱	۰/۱۰۴۴۹۷
۲۵	۱	۰/۶۲۷۵۸	۰/۰۱۸۳۸	۰/۲۴۱۷۹	۰/۱۱۲۲۴
۲۶	۱	۰/۸۰۹۳۶	۰/۰۱۰۳۹	۰/۱۲۷۶۵	۰/۰۵۲۵۹
۲۷	۱	۰/۸۸۳۳۶	۰/۰۰۷۷۸۱	۰/۰۷۵۹۶	۰/۰۳۲۸۸
۲۸	۴	۰/۱۶۶۱۵۳	۰/۰۵۸۴۷	۰/۳۵۱۹۷	۰/۴۲۳۳۹
۲۹	۲	۰/۱۰۷۸۵۳	۰/۴۱۷۶۸	۰/۱۶۶۷۱	۰/۳۰۷۷۴
۳۰	۲	۰/۱۸۶۹۵۸	۰/۳۳۵۳۵	۰/۲۱۸۲۲	۰/۲۵۹۴۶
۳۱	۱	۰/۳۹۸۵۳	۰/۳۱۴۹۸	۰/۳۸۵۳۲	۰/۱۸۴۶۵۲
۳۲	۴	۰/۰۴۴۳۶۱	۰/۰۲۹۱۲	۰/۱۱۹۶۱	۰/۸۰۶۹
۳۳	۱	۰/۸۲۷۷۶	۰/۰۱۰۵۳۸	۰/۱۱۴۲۵	۰/۰۴۷۴۴
۳۴	۱	۰/۸۵۶۸	۰/۰۰۹۴۸	۰/۰۹۲۹۸	۰/۰۴۰۶۸
۳۵	۱	۰/۶۷۳۶۳	۰/۰۱۵۸۶۹	۰/۲۲۶۱	۰/۰۸۴۴۰
۳۶	۴	۰/۰۸۸۲۹	۰/۰۲۲۹۰	۰/۳۸۴۲۱	۰/۵۰۴۶۵
۳۷	۳	۰/۰۴۳۲۵	۰/۰۴۳۵۲	۰/۸۹۸۵۴	۰/۰۵۳۸۴
۳۸	۳	۰/۱۳۲۶۵	۰/۰۱۹۶۴	۰/۵۲۱۸۲	۰/۳۲۵۸۷
۳۹	۴	۰/۱۰۵۵۳	۰/۰۲۹۸۸	۰/۲۶۲۷۱	۰/۶۰۱۸۶
۴۰	۳	۰/۳۶۸۱۱	۰/۰۳۲۹۳۱	۰/۳۸۱۹۳۷	۰/۲۱۷۰۱
۴۱	۱	۰/۶۲۷۲۰	۰/۰۱۵۹۰	۰/۲۶۳۶۳	۰/۰۹۳۲۶
۴۲	۱	۰/۶۶۴۴۶	۰/۰۱۶۴۴	۰/۲۳۰۱۹	۰/۰۸۸۸۹
۴۳	۳	۰/۲۵۷۶۷	۰/۰۵۴۷۳	۰/۴۷۰۲۷	۰/۲۱۷۳۱
۴۴	۱	۰/۵۰۱۲۶	۰/۰۲۰۸۵	۰/۳۶۸۹۱	۰/۱۰۸۹۷
۴۵	۳	۰/۱۵۰۴۹	۰/۰۱۹۴۹	۰/۶۶۲۶۱۲	۰/۱۶۷۳۹
۴۶	۱	۰/۸۲۴۹۳	۰/۰۱۱۶۵۸	۰/۱۱۳۷۰۷	۰/۰۴۹۷۳
۴۷	۱	۰/۵۱۸۲۰	۰/۰۲۴۲۱	۰/۳۴۴۵۵	۰/۱۱۳۰۲
۴۸	۳	۰/۳۴۴۴۲	۰/۰۸۰۰۹۹	۰/۳۶۳۵۳	۰/۲۱۱۹۴
۴۹	۳	۰/۳۰۲۶۴	۰/۱۱۶۸۹	۰/۳۴۱۶۸	۰/۲۳۸۷۷

۴ در منطقه‌ای بین تیپ بی کربناته کلسیک و سولفات‌ها منیزیک قرار دارند. جهت جریان آب زیرزمینی در دشت سراب از شرق به غرب است. بنابراین، شرق دشت منطقه تغذیه و غرب آن منطقه تخلیه است. گروه ۳ نزدیک به منطقه تغذیه و گروه ۴ نزدیک به منطقه خروجی دشت است. به‌طور کلی، قرار گرفتن نمونه‌ها در گروه ۳ و ۴ نشان‌دهنده آن است که روند تکامل هیدروشیمیایی کامل طی نشده است. در جدول ۲ مقادیر متوسط خوشه‌ها برای هر دو روش FCM و HCA آمده است.

همان‌طور که مشاهده می‌شود بی‌نظمی شایان توجهی در گروه‌بندی داده‌ها به روش HCA دیده می‌شود. این موضوع در تفکیک گروه‌های ۱ و ۳ بهتر مشخص است. در شکل ۴ داده‌های خوشه‌بندی شده به روش FCM نشان داده شده است. همان‌طور که انتظار می‌رود، در این روش جداسازی بهتری در نمونه‌های ترسیم شده روی نمودار پایپر دیده می‌شود. با توجه به شکل ۴ نمونه‌های گروه ۱ بیشتر در تیپ بی کربناته کلسیک قرار دارند. نمونه‌های گروه ۲ بیشتر در تیپ سولفات‌ها منیزیک قرار می‌گیرند. نمونه‌های گروه ۳ و





شکل ۴. نمودار پایپر برای داده‌های خوشه‌بندی شده به روش FCM

جدول ۲. مقادیر متوسط خوشه‌ها برای هر دو روش HCA و FCM

EC	SiO <sub>2</sub>	NO <sub>3</sub>	F	SO <sub>4</sub>	K	Na	Cl	Mg	Ca	HCO <sub>3</sub>	تعداد نمونه	خوشه	روش
۶۳۹	۴۵/۲۱	۸/۰۵	۰/۱۷	۸۲	۳/۶	۳۹/۶	۳۸/۶	۲۰/۷	۷۶/۹	۲۴۸	۲۵	۱	HCA
۱۳۴۰	۶۷/۱۷	۱۰/۱۵	۰/۳۸	۲۵۳/۴	۵/۹	۱۳۲/۸	۱۶۰/۴	۲۹/۸	۱۲۸	۲۲۷	۴	۲	
۱۳۵۲	۶۲/۴	۲۰/۱۳	۰/۱۹	۲۱۳/۲	۱۱/۷	۱۳۴/۹	۱۵۶	۳۰/۷	۱۴۰/۴	۳۳۹/۸	۱۳	۳	
۱۱۰۷	۷۲/۱۴	۱۶/۱۲	۰/۱۹	۲۰۳	۸/۴	۵۳۴/۷	۷۵۱	۲۹/۳	۱۳۸/۹	۲۴۸/۴	۷	۴	
۴۰۳	۳۱/۷۷	۳۹/۴	۰/۰۷	۳۶/۷	۳	۲۱/۸	۱۶/۹	۱۳/۷	۵۱/۹	۱۹۰/۶	۱۷	۱	FCM
۲۳۳۳	۱۰۷/۸	۴/۳۳	۰/۳۳	۴۲۶/۸	۱۹/۶	۷۱۸	۹۵۵	۵۱	۲۲۴/۷	۴۰۲	۷	۲	
۶۸۰	۷۷/۶	۹/۶۴	۰/۲۸	۹۷/۸	۴/۷	۳۵/۸	۲۸/۴	۱۸	۸۹/۵	۲۵۶/۷	۱۴	۳	
۱۲۶۹	۳۰/۳	۱۵/۷	۰/۲۱	۲۰۷	۶/۴	۱۰۲	۱۳۷/۹	۳۶/۱	۱۳۸/۸	۳۲۸/۴	۱۱	۴	

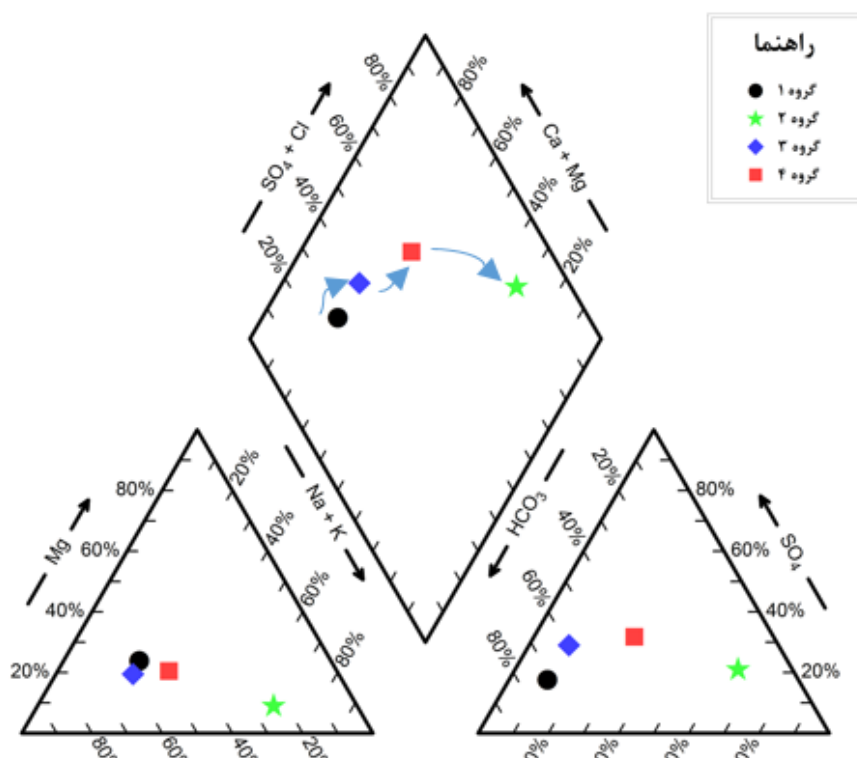
(همه غلظت‌ها برحسب میلی‌گرم بر لیتر، اسیدیته بدون واحد و هدایت الکتریکی میکرو زیمنس بر سانتی‌متر)

الکتریکی متوسط ۲۳۳ میکرو زیمنس بر سانتی‌متر دارند. بیشتر تیپ سولفات دارند و تشکیل این گروه را می‌توان متأثر از رسوبات سولفات در انتهای دشت و غرب محدوده مطالعه شده دانست.

گروه سوم که با رنگ آبی در شکل‌ها مشخص شده است و گروه چهارم با رنگ قرمز و به شکل مربع نشان داده شده است. متوسط هدایت الکتریکی گروه سوم ۶۸۰ و گروه چهارم ۱۲۶۹ میکرو زیمنس بر سانتی‌متر است. این دو گروه تیپ غالب ندارند و تیپ اختلاطی دارند.

نتایج به دست آمده از خوشه‌بندی داده‌ها چهار گروه معرف را نشان می‌دهد. گروه نخست که با رنگ سیاه و دایره پر در تصاویر نشان داده شده است، آب‌هایی با کیفیت زیاد، املاح کم و با هدایت الکتریکی متوسط ۴۰۳ میکرو زیمنس بر سانتی‌متر است. این گروه در طبقه‌بندی آب‌ها در گروه آب شیرین قرار می‌گیرد که بیشتر مربوط به منطقه تغذیه دشت است.

گروه دوم که با رنگ سبز و ستاره در تصاویر نشان داده شده است؛ آب‌هایی با کیفیت نسبتاً کم که هدایت



شکل ۵. نمودار پایپر برای متوسط گروه‌های خوشه‌بندی شده به روش FCM

خوشه‌بندی نسبت به روش HCA کمتر است و خوشه‌بندی فازی فرایندهای ایجادکننده تغییرات طبیعی پیوسته در منطقه را بهتر نشان می‌دهند. بنابراین، انتظار می‌رود که در روش خوشه‌بندی فازی، خطای خوشه‌بندی کمتر از روش‌های خوشه‌بندی سخت باشد. نمونه‌های آب زیرزمینی نزدیک به مرزهای خوشه در روش‌های خوشه‌بندی سخت مرسوم، در معرض خوشه‌بندی اشتباه هستند. در مقابل خوشه‌بندی فازی با قراردادن هر نمونه در چندین خوشه، خوشه‌بندی اشتباه را به کمترین حد می‌رساند. مقادیر عضویت هر خوشه که ترکیب آماری همه پارامترهای استفاده شده در خوشه‌بندی است؛ میزان تعلق هر نمونه به آن خوشه را نشان می‌دهد. در این پژوهش از روش فاصله اقلیدسی و ضریب فازی سازی  $m=2$  برای انتخاب تعداد بهینه خوشه استفاده شد. با ترسیم نمودار مقادیر شاخص FPI و NCE در برابر تعداد خوشه و برای مقادیر مختلف درجه فازی‌شدگی تعداد چهار خوشه به‌عنوان خوشه بهینه انتخاب شد. در روش FCM، ۱۷ نمونه آب زیرزمینی در خوشه ۱، هفت نمونه در خوشه ۲ و به ترتیب ۱۴ و ۱۱ نمونه در خوشه‌های ۳ و ۴ قرار گرفتند. در این پژوهش

در شکل ۵ نمودار پایپر برای متوسط گروه‌های خوشه‌بندی شده به روش FCM آمده است. همان‌طور که مشخص است تیپ آب زیرزمینی از بی‌کربناته کلسیک و منیزیک به سولفاته سدیک تبدیل شده است. گروه ۲ و ۳ نیز حد واسط این دو هستند. به‌طور خلاصه می‌توان میزان املاح محلول در آب زیرزمینی را بدین صورت خلاصه کرد. گروه اول > گروه سوم > گروه چهارم > گروه دوم. نمودار پایپر نیز این روند را به‌خوبی نشان می‌دهد و تحولات رخساره آب زیرزمینی در دشت سراب را به‌خوبی نشان داده است.

### نتیجه‌گیری

در این پژوهش تلاش شد تا با بررسی و مقایسه دو روش پرکاربرد از روش‌های خوشه‌بندی، روشی مناسب برای تعیین رخساره‌های هیدروشیمیایی انتخاب شود. الگوریتم خوشه‌بندی FCM روشی مؤثر برای یافتن ساختار خوشه‌بندی داده‌های پیچیده، دارای همپوشانی و مقادیر مبهم است. توزیع خوشه‌بندی در روش FCM هماهنگی معناداری با رخساره‌های هیدروشیمیایی آب زیرزمینی نشان می‌دهد. نتایج نشان می‌دهد در روش FCM خطای

- [7]. Bezdek JC. Pattern Recognition with Fuzzy Objective Algorithms. Plenum Press New; 1981.
- [8]. Yousefi Rad M, Ahmadi M, Shafiei Haghshenas S. Mass rock clustering using C-Mean clustering. 4<sup>th</sup> National Proceeding of Economic Geology [Persian].
- [9]. Aghili R, Boromand Tesb S, Kahe M. The application of Fuzzy modeling based on C-mean clustering in estimation from evapotranspiration pan (Case study: Khuzestan Province). Soil and Water Conservation Researches. 2012;19(2):81-97 [Persian].
- [10]. Rantitsch G. Application of fuzzy clusters to quantify lithological background concentrations in stream-sediment geochemistry. Journal of Geochemical Exploration. 2000;71(1):73-82.
- [11]. Burrough PA, van Gaans PF, MacMillan RA. High-resolution landform classification using fuzzy k-means. Fuzzy sets and systems. 2000;113(1):37-52.
- [12]. Zhang CT, Chou KC, Maggiora GM. Predicting protein structural classes from amino acid composition: application of fuzzy clustering. Protein engineering. 1995;8(5):425-35.
- [13]. Azar E, Faraji H. Science of Fuzzy management, Ketabe Mehraban Nashr Press; 2006 [Persian].
- [14]. Jang JS. ANFIS: adaptive-network-based fuzzy inference system. IEEE transactions on systems, man, and cybernetics. 1993;23(3):665-85.
- [15]. Teshne Lab M, Safar Pour N, Afyouni D. Fuzzy systems and fuzzy control, Khaje Nasiredin Tousi Universty Press; 2007 [Persian].
- [16]. Hathaway RJ, Bezdek JC. Fuzzy c-means clustering of incomplete data. IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics). 2001;31(5):735-44.
- [17]. Zhang B, Song X, Zhang Y, Han D, Tang C, Yu Y, Ma Y. Hydrochemical characteristics and water quality assessment of surface water and groundwater in Songnen plain, Northeast China. Water research. 2012;46(8):2737-48.
- [18]. Davis, J.C. Statistics and data analysis in geology, New York: John Wiley and sons; 1986.
- [19]. Fetter, C.W. Contaminant hydrogeology, second edition, Prentic Hall, Inc; 1999.

به‌منظور صحت‌سنجی و بررسی توزیع رخساره‌های هیدروشیمیایی آب زیرزمینی از نمودار پایپر استفاده شد. در نمودار پایپر ترسیم‌شده به روش HCA بی‌نظمی در گروه‌بندی داده‌ها دیده می‌شود. در تفکیک گروه‌های ۱ و ۳ بی‌نظمی بهتر مشخص شده است؛ اما در نمودار پایپر ترسیم‌شده به روش FCM گروه‌ها نظم بهتری دارند و نمونه‌هایی با رخساره هیدروشیمیایی یکسان، اغلب در یک گروه قرار گرفتند. به‌طور کلی، می‌توان نتیجه گرفت در روش HCA تغییرات در مرز بین گروه‌ها کاملاً واضح و ناگهانی است، اما در روش FCM، تغییرات تدریجی است و در بررسی توزیع رخساره‌های هیدروشیمیایی آب زیرزمینی که تغییرات پیوسته و تدریجی است؛ روش FCM توانایی زیادی دارد و می‌تواند شرایط هیدروژئوشیمیایی حاکم بر آبخوان را بهتر نشان دهد.

#### منابع

- [1]. Banoeng-Yakubo B, Yidana SM, Nti E. Hydrochemical analysis of groundwater using multivariate statistical methods—the Volta region, Ghana. KSCE Journal of Civil Engineering. 2009;13(1):55-63.
- [2]. Barbieri P, Adami G, Favretto A, Lutman A, Avoscan W, Reisenhofer E. Robust cluster analysis for detecting physico-chemical typologies of freshwater from wells of the plain of Friuli (northeastern Italy). Analytica Chimica Acta. 2001;440(2):161-70.
- [3]. Güler C, Thyne GD. Delineation of hydrochemical facies distribution in a regional groundwater system by means of fuzzy c-means clustering. Water Resources Research. 2004;40(12).
- [4]. Güler C, Thyne GD, McCray JE, Turner KA. Evaluation of graphical and multivariate statistical methods for classification of water chemistry data. Hydrogeology journal. 2002;10(4):455-74.
- [5]. Zadeh LA. Fuzzy sets. Information and control. 1965;8(3):338-53.
- [6]. Taheri M. Introduction to Fuzzy set theory, second edition, Mashahd Jahad Daneshgahi, Inc; 1999 [Persian].