

مقایسه روش‌های خوشه‌بندی در تهیه نقشه پتانسیل معدنی در بی‌هنجاری باریکا، آذربایجان غربی

محسن شادمان^۱ و بهزاد تخم‌چی^۲

^۱ کارشناسی ارشد، شرکت مهندسی مشاور کاوشگران، معدن سه‌چاهون، یزد، ایران
^۲ استادیار، دانشکده مهندسی معدن، نفت و ژئوفیزیک، دانشگاه صنعتی شاهرود، شاهرود، ایران

تاریخ دریافت: ۱۳۹۰/۰۹/۰۶ تاریخ پذیرش: ۱۳۹۱/۰۷/۳۰

چکیده

امروزه تلفیق داده‌های زمین مرجع با استفاده از فناوری سیستم اطلاعات جغرافیایی (GIS)، به صورت گسترده‌ای به کار گرفته می‌شود. در این روش‌ها، تلفیق لایه‌های اطلاعاتی با توجه به کیفیت و کمیت داده‌ها و به کارگیری روش‌های وزن‌دهی مناسب صورت می‌گیرد. در پایان با استفاده از روش‌های تقسیم‌بندی مانند روش فواصل هندسی، نقشه پتانسیل معدنی تهیه می‌شود. روش‌های خوشه‌بندی که استفاده از آنها در پردازش داده‌ها در حال افزایش است، با توجه به میزان شباهت میان داده‌ای، نمونه‌ها را میان گروه‌های مشخص تقسیم می‌کند. در این پژوهش روش تهیه نقشه پتانسیل معدنی با هدف اکتشاف طلا با استفاده از روش خوشه‌بندی میانگین‌های K- و میانگین‌های K- فازی بررسی می‌شود. پس از آن نتایج این روش‌ها با نقشه پتانسیل معدنی که در محیط GIS با استفاده از مدل تلفیق همپوشانی اندیس‌ها برای منطقه باریکا در برکه ۱:۱۰۰۰۰۰ آلوت در جنوب آذربایجان غربی تهیه شده، مقایسه می‌شود. در منطقه باریکا به دلیل در دسترس نبودن نقاط حفاری در دسترس نمی‌توان تعداد کلاس‌ها و مرز هر کلاس برای تعیین امتیازات نقشه نهایی را با روش‌های متداول مشخص کرد، ولی در روش‌های خوشه‌بندی، تعیین آستانه برای تعداد کلاس مفروض به صورت خودکار و بر پایه میزان شباهت داده‌ها به خوبی انجام می‌شود. افزون بر این با استفاده از روش‌های خوشه‌بندی می‌توان شمار کلاس بهینه برای نمایش امتیازات نقشه پتانسیل را نیز تعیین کرد که در این پژوهش انجام شده است. نتایج نشان می‌دهد که هر دو روش، خوشه‌بندی و همپوشانی اندیس‌ها، پرتانسبل‌ترین محدوده را تقریباً مشابه یکدیگر به دست آورده‌اند با این تفاوت که روش خوشه‌بندی به اطلاعات کمتری نسبت به منطقه مورد بررسی نیاز دارد در حالی که در روش همپوشانی اندیس‌ها، وزن‌دهی داده‌ها و تعیین آستانه برای تعیین مرز امتیازات پایانی به اطلاعات بیشتری نیاز دارد.

کلیدواژه‌ها: سیستم اطلاعات جغرافیایی، خوشه‌بندی فازی، میانگین‌های-K، نقشه پتانسیل معدنی، باریکا.

*نویسنده مسئول: محسن شادمان

E-mail: mohseneshademan@yahoo.com

۱- پیش‌گفتار

با توجه به پتانسیل بالای کانی‌سازی طلا و سولفید در برکه ۱:۱۰۰۰۰ باریکا که جزئی از برکه ۱:۱۰۰۰۰۰ آلوت است، پروژه‌های اکتشافی زیادی مانند برداشت‌های هوابرد مغناطیس، الکترومغناطیس و رادیومتری با فاصله خطوط پرواز ۲۰۰ متر، برداشت لیتوژئوشیمیایی با شبکه‌های مربعی ۵۰ متر، ژئوشیمیایی سیستماتیک و تهیه نقشه زمین‌شناسی در مقیاس ۱:۵۰۰۰ در این منطقه انجام شده است. وجود لایه‌های اطلاعاتی گوناگون برای اکتشاف ذخایر معدنی موجود در منطقه، نیاز به استفاده از روش‌های تلفیق به‌منظور تهیه نقشه پتانسیل معدنی (MPM) را آشکار می‌کند. با استفاده از این نقشه و با توجه به تمام اطلاعات به نتیجه‌ای با درصد خطا و ریسک کمتری دست یافت. نقشه‌های پتانسیل معدنی ابزار مفیدی است که به‌صورت بهینه در مورد احتمال گسترش یک عنصر معدنی با توجه به همه اطلاعات موجود نظر می‌دهد. تهیه نقشه پتانسیل معدنی در سامانه اطلاعات جغرافیایی، انجام می‌شود. انتخاب صحیح مبنای مختصات به کار رفته، روش وزن‌دهی و تلفیق از جمله تصمیماتی است که در مسیر تهیه نقشه باید به آن توجه شود. از جمله روش‌های مؤثر در سامانه اطلاعات جغرافیایی می‌توان به روش وزن‌های نشانگر (Weights of Evidence) (Bonham & Agterberg, 1989) و رگرسیون منطقی (Logistic Regression) (Caranza et al., 2008) اشاره کرد.

علوم زمین استفاده‌های فراوانی می‌شود؛ مانند تشخیص شکستگی‌ها از روی نگارهای چاه‌های نفتی (Tokhmechi et al., 2009)، تهیه نقشه شبه زمین‌شناسی (Geology Map -Pseudo) از روی داده‌های رادیومتری (Martelet et al., 2006) و پردازش داده‌های مقاومت سنجی (Song & Meng, 2010) است. البته کاربرد این روش در تلفیق لایه‌های اطلاعاتی، ایده به‌نسبت جدیدی است که در این پژوهش ارزیابی می‌شود. به‌دلیل غیر نظارتی (Unsupervised) بودن روش‌های خوشه‌بندی، این روش‌ها در مناطقی با اطلاعات کم بسیار مفید است (Song & Meng, 2010). افزون بر این، روش‌های خوشه‌بندی مستقل از ابعاد شبکه هستند و به‌عبارتی تلفیق داده‌های مختلف کمتر تحت تأثیر تباین مکانی داده‌ها مانند رسوبات رودخانه و داده‌های هوابرد می‌شوند (Song & Meng, 2010). در این پژوهش از دو روش پرکاربرد خوشه‌بندی، میانگین‌های-K- و میانگین‌های-K- فازی (K- means and fuzzy K- means clustering methods) به‌منظور تهیه نقشه پتانسیل معدنی با هدف طلا در بخشی از ورقه ۱:۱۰۰۰۰۰ آلوت در استان آذربایجان غربی استفاده شده است. در پایان به‌منظور تعیین میزان موفقیت این روش، نتایج روش همپوشانی اندیسی (Index Overlapping) نیز تهیه شده است که یکی از مرسوم‌ترین روش‌ها در سیستم اطلاعات جغرافیایی به‌شمار می‌رود (Theobald, 2007) و با نتایج به‌دست آمده از خوشه‌بندی داده‌ها مقایسه می‌شود.

۲- منطقه مورد مطالعه

برکه ۱:۱۰۰۰۰ باریکا واقع در برکه ۱:۱۰۰,۰۰۰ آلوت، در شمال بانه، میان طول‌های جغرافیایی ۳۷°۴۵' و ۴۱°۴۵' و عرض‌های جغرافیایی ۳۶°۱۲' و ۳۶°۱۲'، در

در خوشه‌بندی با روش‌های فاصله-پایه میانگین‌های K و میانگین‌های-K فازی، شمار خوشه‌ها (K) مفروض در نظر گرفته می‌شود و توسط الگوریتمی بهینه تعلق داده‌ها به خوشه‌های مفروض، صورت می‌گیرد. سپس با تغییر شمار خوشه‌های مفروض، عملیات فوق تکرار شده و نسبت به بهینه‌سازی شمار خوشه‌ها اقدام می‌شود. نقطه قوت این روش‌ها، راندمان محاسباتی بالای آن‌ها است. از این رو نسبت به دیگر روش‌های خوشه‌بندی مانند روش‌های مرتبه‌ای (Hierarchical clustering methods) برای داده‌های با حجم بالا بسیار مفید هستند (Velumrgan, 2010).

تنها تفاوت منطق خوشه‌بندی میانگین‌های K فازی، با میانگین‌های K در این است که در این روش، تعلق داده‌ها به خوشه‌های مختلف به صورت مطلق نبوده و فازی است. بنابراین برای هر داده یک احتمال تعلق به خوشه‌های مختلف تعریف می‌شود. در این روش به طور معمول، احتمال تعلق تابعی از فاصله داده از مرکز خوشه و میزان نزدیکی آن به خوشه‌های مجاور خواهد بود. هر چه یک داده به حاشیه‌های خوشه نزدیک‌تر شود، احتمال فازی تعلق آن به خوشه مربوطه اش کوچک‌تر خواهد شد (Velumrgan, 2010).

در این الگوریتم‌ها داده‌ها پس از این که به فرم استاندارد در آمدند، به صورت تصادفی در K خوشه مفروض پخش می‌شوند. سپس تابع هزینه‌ای برای بهینه‌سازی تعلق داده‌ها به هر خوشه با رابطه ۱ تعریف می‌شود:

$$J = \sum_{i=1}^k \sum_{x \in W_i} \|x - \mu_i\|^2 \quad (1)$$

که در آن l : تابع هزینه، k : شمار خوشه‌های مفروض، W_i : خوشه نام، x : بردار ویژگی‌های داده‌ها و μ_i : بردار میانگین ویژگی‌های داده‌ها در خوشه نام است. μ_i به صورت زیر تعریف می‌شود:

$$\mu_i = \frac{1}{N_i} \sum_{x \in W_i} x \quad (2)$$

که در آن N_i : شمار داده‌های واقع در خوشه نام است. تابع هزینه (J) معرف مجموع فاصله داده‌های تمام خوشه‌ها با میانگین آن‌ها در خوشه مربوطه‌شان است. بنابراین در صورتی که داده‌ها به درستی در خوشه خودشان طبقه‌بندی شده باشند، J کمینه خواهد شد. بدیهی است در غیر این صورت، باید داده‌ها را از خوشه‌ای به خوشه دیگر انتقال داد، با این شرط که J کوچک‌تر شود. بنابراین در کل روش کار به صورت زیر خواهد بود (Davis & Bouldin, 1979):

الف) داده‌ها تبدیل به حالت استاندارد می‌شوند.

ب) در ابتدا داده‌ها به صورت تصادفی در K خوشه طبقه‌بندی می‌شوند. به این ترتیب یک خوشه‌بندی اولیه صورت گرفته است.

ج) با توجه به معلوم بودن داده‌های هر خوشه، بردار میانگین ویژگی‌های داده‌های هر خوشه (μ_i)، محاسبه می‌شود. توجه شود که از هر داده شامل چندین ویژگی می‌تواند باشد و منظور از بردار میانگین ویژگی‌های داده‌ها، میانگین‌گیری از همه ویژگی‌های داده‌های هر خوشه است.

د) به صورت تصادفی یک داده از مجموعه داده‌ها انتخاب شده و فاصله آن از میانگین داده‌های خوشه‌های مختلف محاسبه می‌شود. اگر فاصله این داده از میانگین خوشه خودش کمتر از فاصله آن تا میانگین دیگر خوشه‌ها بود، آن داده در خوشه اولیه‌اش باقی می‌ماند؛ در غیر این صورت، آن داده به خوشه دیگری که کوچک‌ترین فاصله را با میانگینش دارد، انتقال می‌یابد.

ه) داده‌های دیگر به ترتیب و به صورت تصادفی انتخاب شده و نسبت به تعیین تکلیف تعلق آنها به خوشه‌های مختلف اقدام می‌شود.

و) در پایان مقدار تابع J و بردار میانگین خوشه‌ها به ترتیب از روابط ۱ و ۲ محاسبه

استان آذربایجان غربی قرار دارد. همان‌گونه که در شکل ۱ دیده می‌شود واحد متاولکانیک کرتاسه، بیشتر از نوع تراکی آندزیت و توف، سرتاسر منطقه را پوشانده است. در نقشه زمین‌شناسی ۱:۱۰۰۰۰ به ترتیب محدوده‌های به رنگ سبز تیره، واحد متاولکانیک بدون دگرسانی، سبز کم‌رنگ، متاولکانیک با دگرسانی کم و تا حدودی تحت تأثیر پهنه برشی، محدوده زرد رنگ به عنوان پهنه برشی با میزان سولفید بالا و آبی رنگ، پهنه فیلیتی بدون کانی‌زایی است. معدن باریت فعالی که در مرکز نقشه وجود دارد در برداشت‌های اکتشافی به عنوان یک بی‌هنجاری درجه اول طلا و سولفید سرب با تیپ ماسیو سولفید ولکانوژن (Massive Sulfide Volcanogenic) شناخته شده است. منطقه باریکا تحت تأثیر پهنه برشی با امتداد شمال باختر- جنوب خاوری قرار گرفته است و محل مناسبی را برای ته‌نشست سولفیدها بوجود آورده است. در این منطقه دگرسانی آرژیلیتی (با سولفید بالا) و دگرسانی پروپلیتی (با سولفید کم) وجود دارد (یارمحمدی و همکاران، ۱۳۸۷).

۳- داده‌های به کار رفته در روش خوشه‌بندی

در طول مطالعات زمین‌شناسی در مقیاس بزرگ، امکان تعیین بهتر ویژگی‌های کانسار و تیپ کانی‌سازی فراهم شده است. بنابراین در پردازش داده‌ها تنها از لایه‌های اطلاعاتی استفاده شد که در شناسایی مناطق امیدبخش طلا نتایج بهتری ارائه می‌کردند. این لایه‌های اطلاعاتی بر حسب اولویت از بالا به پایین عبارت است از:

- واحد متاولکانیک کرتاسه
- پهنه‌های برشی سولفیددار
- نقشه‌های لیتوژئوشیمیایی طلا، نقره، سرب، آرسنیک و سولفور
- بی‌هنجاری منفی مقاومت ویژه الکتریکی (برداشت‌شده از داده‌های الکترومغناطیس هوابرد)
- بی‌هنجاری مثبت پتاسیم در مطالعات طیف‌سنجی رادیومتری
- نزدیکی به گسل‌ها.

پیش از پردازش داده‌ها، لازم بود تا لایه‌های اطلاعات آماده شوند. برای این منظور همه نقشه‌ها با ابعاد پیکسل ۵۰ متر راقومی شدند؛ سپس با توجه به نیاز، محدوده‌هایی در پیرامون برخی لایه‌های اطلاعاتی مانند گسل‌ها و پهنه‌های برشی به عنوان بافر (Buffer) با شعاع مشخص در نظر گرفته شد. از آنجا که در خوشه‌بندی، بازه عددی داده‌ها در هر تیپ متفاوت است، لازم بود تا داده‌ها بهنجار شوند. بهنجار کردن داده‌ها می‌تواند به روش‌های مختلفی صورت پذیرد که در این پژوهش تمام سلول‌ها در لایه‌های اطلاعاتی با توجه به ارزشی که دارند میان ۱ تا ۱۰ نمره گرفته‌اند. این ارزش‌ها به صورت بازه‌های عددی بر پایه نظر کارشناس و یا میزان تأثیر گسل مربوط در کانی‌سازی تعریف می‌شوند. در پایان، نقشه‌های رستر به‌عارضه‌هایی از نوع نقطه‌ای تبدیل شدند. هر نقطه در هر لایه اطلاعاتی با سه ویژگی معرفی می‌شود؛ مختصات و عددی میان ۱ تا ۱۰. از فایل‌های داده‌ها یک ماتریس 12×18719 تهیه شد که سطرها، شماره نمونه‌ها و ستون‌ها، ویژگی‌های هر سلول به همراه مختصات مربوط هستند. در شکل ۲ نحوه برخی از مهم‌ترین پارامترهای دخیل در کانی‌سازی طلا و سولفید در منطقه باریکا دیده می‌شود.

۴- الگوریتم خوشه‌بندی میانگین‌های K و میانگین‌های-K فازی

هدف الگوریتم‌های خوشه‌بندی، به حداقل رساندن عدم تشابه نمونه‌ها در یک گروه است. بنابراین نمونه‌هایی با تشابه بالا در یک خوشه کلاسه‌بندی می‌شوند (Nugraha, 2011). از همین رو، با توجه به معیارهای تشابه/عدم تشابه، الگوهای خوشه‌بندی به دو رهیافت، بر پایه فاصله و بر پایه مدل تقسیم می‌شوند. در روش فاصله‌ای، الگوریتم خوشه‌بندی اعضای یک خوشه را بر پایه فاصله میان نمونه تعیین می‌کند و در روش بر مبنای مدل، انتخاب اعضا به وسیله تعیین یک مدل پراکنندگی مناسب میان نمونه‌ها صورت می‌گیرد (Reimann et al., 2008).

خدا به گونه‌ای مشخص کرد که نقاط حفاری بر پایه اهمیت آن در کلاس مشخصی قرار گیرد (Paasche & Eberle, 2009).

به‌منظور تعیین شمار بهینه کلاس‌ها، برای هر دو روش خوشه‌بندی، مقدار هزینه برای ۲ تا ۱۰ کلاس محاسبه می‌شود. این عملیات ۱۰ مرتبه تکرار شده و سپس منحنی میانگین مجذور هزینه‌ها رسم می‌شود (Paasche & Eberle, 2009).

با توجه به شکل ۵ تعداد بهینه کلاس برای روش میانگین‌های-K، ۶ کلاس و برای روش میانگین‌های-K فازی ۹ کلاس است. از این روش نقشه پتانسیل معدنی این تعداد کلاس‌ها برای هر روش در شکل ۶ آورده شده است.

در شکل ۶ محل ماده معدنی با علامت به‌علاوه مشکلی پر مشخص شده است. با توجه به محل ماده معدنی و مشاهده لایه‌های اطلاعاتی به‌صورت تک به تک، دیگر پهنه‌های با پتانسیل بالای شناسایی شده از روش خوشه‌بندی، به‌ترتیب با رنگ‌های سبز فسفری و قرمز رنگ نمایش داده شده‌اند. چنانکه در این شکل دیده می‌شود، مکان پتانسیل‌های مطلوب در حالی که شمار خوشه‌ها متفاوت در نظر گرفته شده‌اند، جابه‌جا شده‌اند. بنابراین می‌توان نتیجه‌گیری کرد که بهینه‌سازی شمار خوشه‌ها از اهمیت بسیاری در شناسایی و اولویت‌بندی پتانسیل‌های معدنی برخوردارند.

۷- تهیه نقشه پتانسیل معدنی باریگا با روش همپوشانی اندیسی‌ها

روش همپوشانی اندیسی‌ها، روش ساده و کاربردی است که در این روش مقدار یک پیکسل در موقعیت مکانی مشخص در تمام لایه‌های اطلاعاتی در وزن مربوط به خود ضرب و سپس با یکدیگر جمع بسته می‌شوند. وزن هر یک از لایه‌های اطلاعاتی بر پایه نوع کانی‌سازی و نظر کارشناسان مربوطه در جدول ۱ آورده شده است. امتیازات در نقشه پتانسیل معدنی با استفاده روش فاصله هندسی برای ۶، ۷ و ۸ کلاس در شکل ۷ نشان داده شده است. به‌نظر می‌رسد که با توجه به روندهای خطی و شکل محدوده‌های تعیین شده، نقشه‌های پتانسیل معدنی با ۶ و ۸ کلاس با استفاده از همپوشانی اندیسی‌ها به نقشه‌های پتانسیل معدنی با روش‌های خوشه‌بندی (شکل‌های ۶ و ۷) شباهت بسیاری دارند و با توجه به محل معدن باریت می‌توان در هر دو راهکار محدوده با اولویت بالا را شناسایی کرد.

۸- نتیجه‌گیری

روش‌های خوشه‌بندی این پژوهش، با استفاده از شباهت داده‌ها، نمونه‌ها را در خوشه‌هایی مشخص قرار می‌دهد. وظیفه اصلی کاربرد این روش‌ها، تعیین بهینه شمار خوشه است که می‌توان با تعریف یک تابع هزینه این پارامتر را تعیین کرد. به‌عنوان مثال چنان که در شکل ۶ قابل ملاحظه است، با تغییر تعداد خوشه‌ها، مکان پتانسیل‌های مطلوب جابه‌جا شده‌اند. از همین رو این روش‌ها را می‌توان مستقل از دانش شخصی روی مجموعه‌ای از داده اجرا کرد. بنابراین در بسیاری از موارد، مانند مراحل اکتشافات مقدماتی که اطلاعات از محیط نمونه برداری کم است و بهتر است از روش‌های مستقل از نظر کارشناس، مانند روش‌های خوشه‌بندی استفاده کرد. در کل به‌نظر می‌رسد که حداقل در شناسایی پتانسیل‌های پلی‌متال، استفاده از روش‌های خوشه‌بندی پاسخ بهتری نسبت به دیگر روش‌های شناسایی پتانسیل‌های مطلوب دارد.

سپاسگزاری

از مسئولین محترم سازمان زمین‌شناسی و اکتشافات معدنی کشور، به‌ویژه مهندس فتوتی و مهندس خیرالهی که در تهیه داده‌ها همکاری خوبی داشتند، سپاسگزاری می‌شود.

شده و عملیات بالا تا زمانی که شرط توقف ارضا شود ادامه می‌یابد. شرط توقف می‌تواند کوچک‌تر شدن شیب کاهش ل، از یک حد آستانه مشخص باشد. این دو روش، در محیط نرم‌افزار Matlab گد نویسی شده و برای داده‌های این مقاله مورد استفاده قرار گرفته‌اند.

۵- تهیه نقشه پتانسیل معدنی باریگا با استفاده از خوشه‌بندی

یکی از کاربردهای تلفیق لایه‌های اطلاعاتی با استفاده از روش‌های خوشه‌بندی، تهیه نقشه‌های شبه زمین‌شناسی با استفاده از داده‌های ژئوفیزیک هابرد است. در این نقشه‌های شبه زمین‌شناسی، پیکسل‌های موجود روی نقشه‌ها به گروه‌هایی تقسیم می‌شوند که هر یک می‌تواند معرف یک واحد زمین‌شناسی روی سطح زمین باشند (شادمان و همکاران، ۱۳۹۱؛ Eberle et al., 2009; Paasche & Eberle, 2010).

با این حال در این پژوهش، هدف تهیه نقشه شبه زمین‌شناسی نیست؛ بلکه تعیین محدوده‌های روی منطقه مورد مطالعه است که همگی تقریباً در واحد متاولکانیک کوتاه قرار دارند و هیچ واحد سنگی دیگری نیز وجود ندارد و این محدوده‌ها هرکدام امتیاز متفاوتی از احتمال حضور کانی‌سازی دارند.

برای تهیه نقشه پتانسیل معدنی با استفاده از روش‌های خوشه‌بندی نیاز است تا دو مرحله پردازش روی داده‌ها انجام شود که به‌ترتیب عبارتند از آماده‌سازی داده‌ها و خوشه‌بندی در صورتی که هدف تهیه نقشه شبه زمین‌شناسی است باید پس از تهیه نتایج خوشه‌بندی، میزان درستی نتایج بررسی شود. برای تعیین درستی نتایج می‌توان از ماتریس خطا و ضریب کاپا (Kappa Coefficient) استفاده کرد. در این صورت پیش از تعیین درستی باید خروجی خوشه‌بندی درون‌یابی، رسترسازی (Rasterizig) و کلاسه‌بندی دوباره شود (Nugraha, 2011).

نقشه پتانسیل معدنی با استفاده از دو روش خوشه‌بندی بیان شده، برای شمار کلاس‌های ۴، ۵ و ۶ در شکل‌های ۳ و ۴ نشان داده شده است. با توجه به تراکم نقاط مورد استفاده، از درون‌یابی استفاده نشد و تنها با انتخاب مناسب ابعاد پیکسل این نقشه‌ها به‌دست آمده است. از آنجا که خروجی این الگوریتم‌ها در هر بار اجرا برای یک شمار خوشه مشخص، به دلیل مقادیر اولیه تصادفی متفاوت است بنابراین با استفاده از معیار خوشه مشخص (Davis & Bouldin, 1979) برای روش میانگین‌های-K (فرمول ۳) و روش ضرائب افزایش (Partition Coefficient (PC)) (Lucieer & Lucieer, 2009) (فرمول ۴) برای روش میانگین‌های-K فازی بهترین نتیجه در چند بار تکرار برنامه انتخاب شده است. با توجه به راندمان بالای محاسبه هر کدام از الگوریتم‌ها برای شمار کلاس مشخص حداقل ۱۰ بار تکرار شد. مقدار پارامتر فازی شدگی برای روش فازی برابر ۱/۵ در نظر گرفته شد.

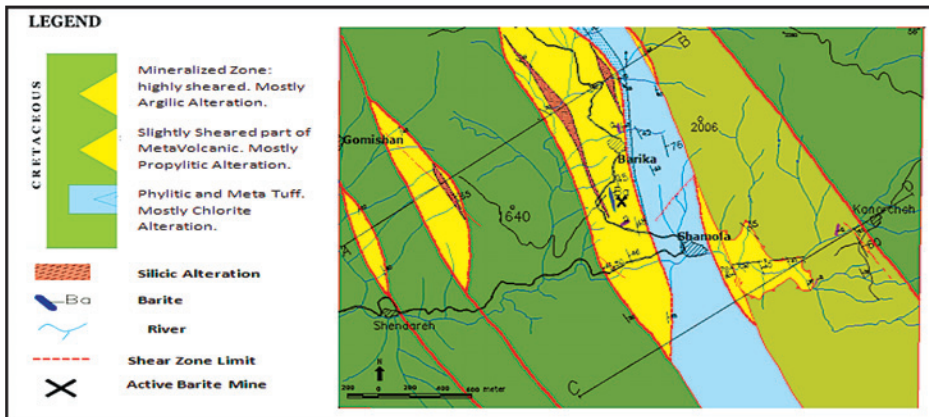
$$D(\underline{u}) = \frac{\min_{i=1}^c \min_{j=1}^c d(\omega_i, \omega_j)}{\max_{i=1}^c d(\omega_i, \omega_1)}, \text{ for } i \neq j \quad (3)$$

$$PC(\underline{u}) = \frac{\sum_{i=1}^K \sum_{j=1}^N u_{ij}^2}{N} \quad (4)$$

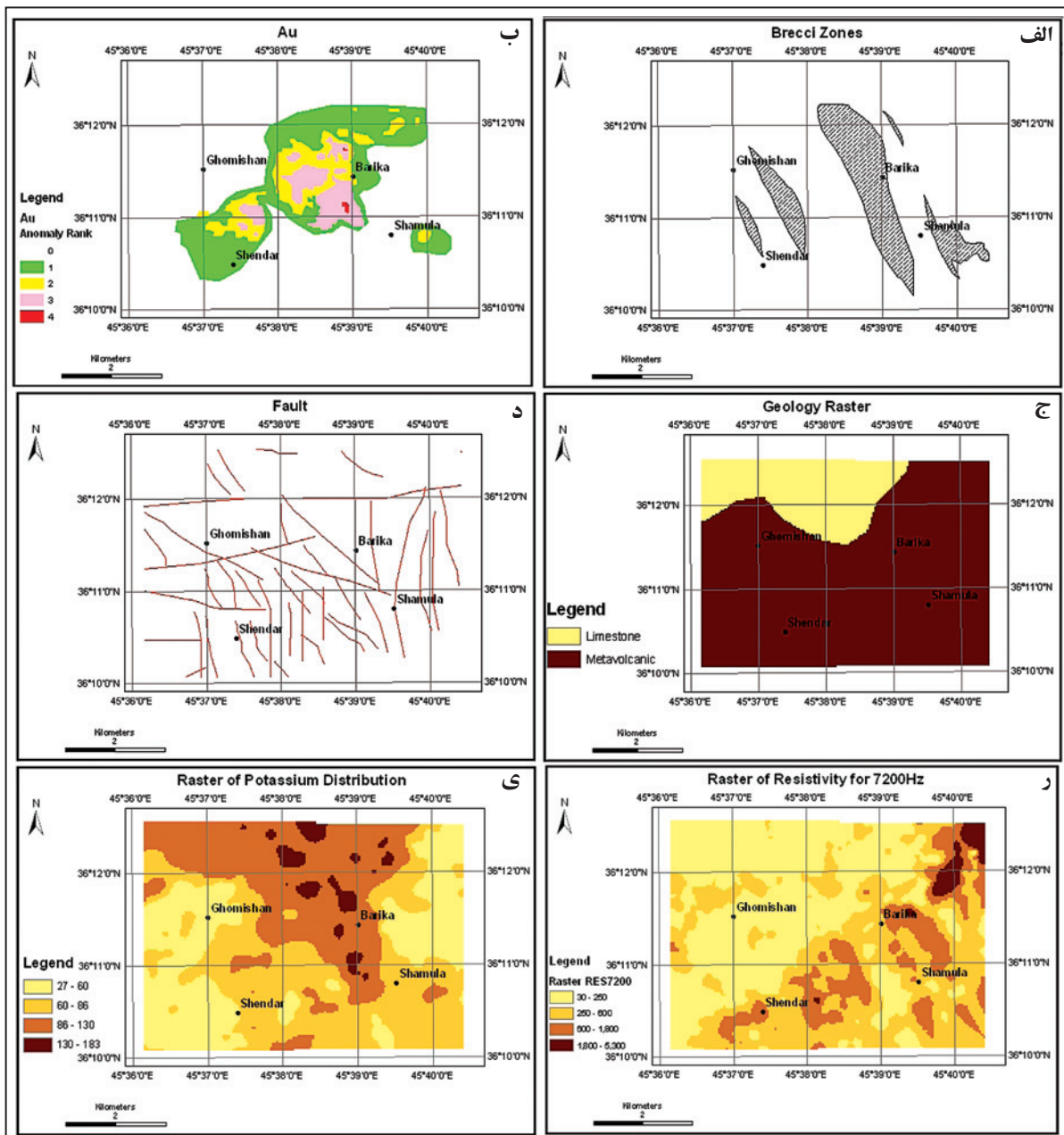
در روابط بالا، min تابع حداقل، ω_i نماد خوشه i ام و d نماد فاصله میان خوشه‌ای، K شمار کلاس مفروض، N شمار داده‌ها و u_{ij} درجه عضویت داده j در خوشه i است.

۶- انتخاب تعداد خوشه بهینه در خوشه‌بندی داده‌ها

همان‌گونه که بیان شد در سیستم مختصات جغرافیایی، پس از تعیین امتیاز هر پیکسل در نقشه نهایی پتانسیل معدنی، باید نقشه را به کلاس‌های مختلف تقسیم و محدوده‌ها را به عنوان بی‌هنجاری اول، بی‌هنجاری دوم، ... معرفی کرد. برای این کار نیاز به استفاده از داده‌هایی مانند نقاط حفاری است، سپس باید محدوده‌ها را با سعی و

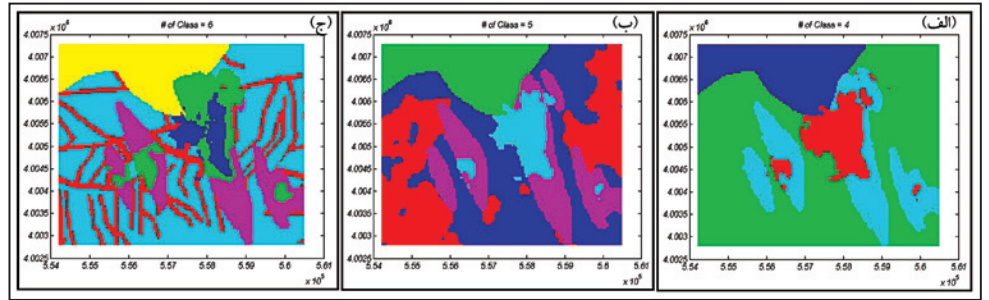


شکل ۱- نقشه زمین‌شناسی و ساختمانی ۱:۱۰۰۰۰ باریکا (مجله، ۱۳۸۳).

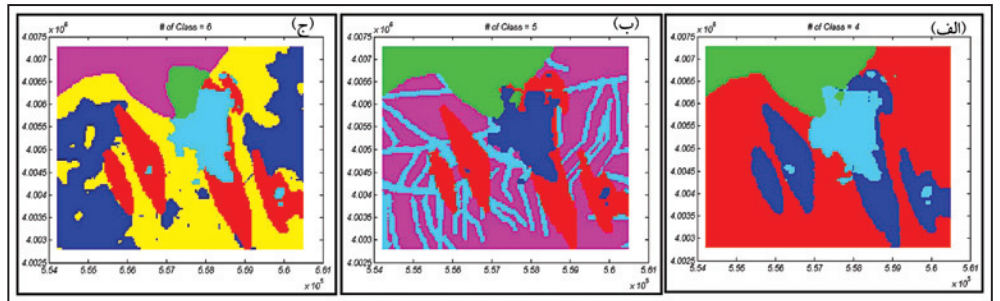


شکل ۲- نقشه پراکندگی پارامترهای مهم در کانی‌سازی طلای بی‌هنجاری باریکا؛ الف) بینه‌های بُرشی؛ ب) پراکندگی عنصر طلا در برداشت‌های لیتوزئوشیمیایی؛ ج) واحدهای آهکی کرتاسه (رنگ روشن) و متاولکانیک کرتاسه (رنگ تیره)؛ د) پراکندگی خطواره‌های به‌دست آمده از داده‌های مغناطیس‌سنجی؛ ر) پراکندگی مقاومت ویژه الکتریکی به‌دست آمده از داده‌های الکترومغناطیس هواپرد؛ ی) پراکندگی رادیو عنصر پتاسیم.

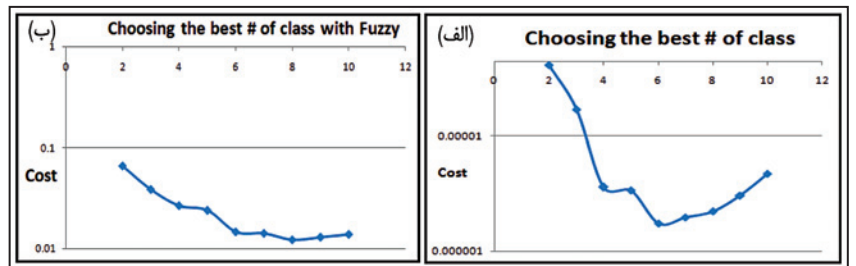
شکل ۳- نقشه پتانسیل معدنی با استفاده از خوشه بندی میانگین های-K، الف) ۴ کلاس؛ ب) ۵ کلاس و ج) ۶ کلاس.



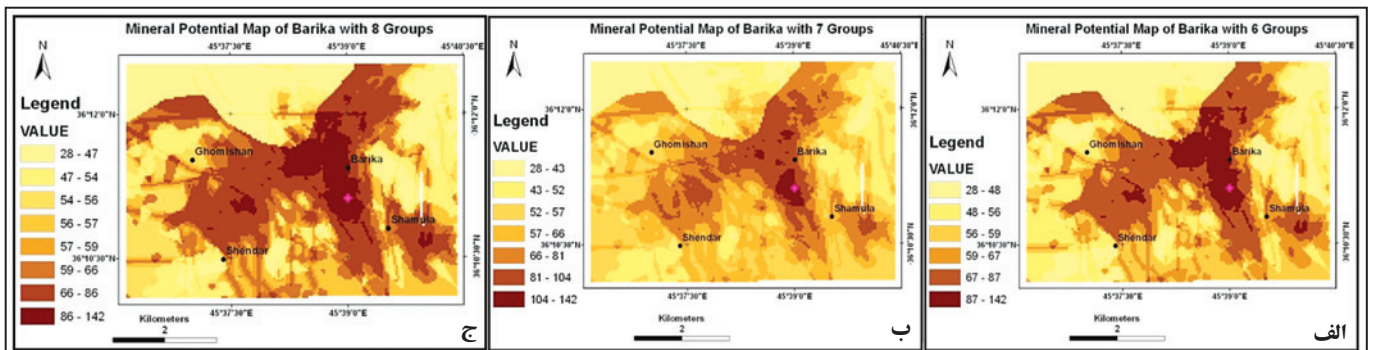
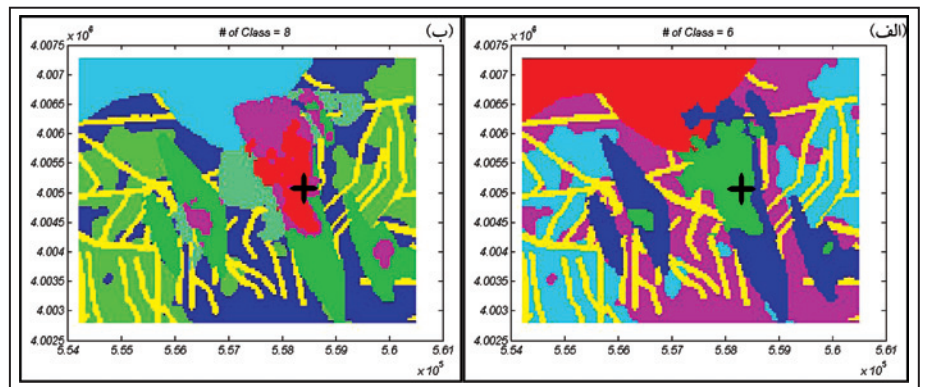
شکل ۴- نقشه پتانسیل معدنی با استفاده از خوشه بندی میانگین های-K فازی، الف) ۴ کلاس؛ ب) ۵ کلاس و ج) ۶ کلاس.



شکل ۵- منحنی تابع هزینه نتایج خوشه بندی به منظور پیدا کردن تعداد کلاس بهینه. الف) روش فازی و ب) غیر فازی.



شکل ۶- نقشه نهایی پتانسیل معدنی با استفاده از روش خوشه بندی برای تعداد کلاس بهینه. الف) میانگین های-K با ۶ کلاس و ب) میانگین های-K فازی با ۸ کلاس.



شکل ۷- نقشه های پتانسیل معدنی با استفاده از مدل تلفیق همپوشانی اندیس ها در نرم افزار ArcGIS برای الف) ۶ کلاس؛ ب) ۷ کلاس؛ ج) ۸ کلاس. محل معدن سولفید به صورت علامت به علاوه صورتی رنگ مشخص شده است.

جدول ۱- وزن‌های مربوط به هر یک از لایه‌های اطلاعاتی به کار رفته در تلفیق با روش همبوشانی اندیس‌ها.

لایه اطلاعاتی	نقره	آرسنیک	طلا	سرب	سولفور	زون پرشی	واحد آهک کراتاسه	گسل	مقاومت ویژه الکتریکی	طیف پتاسیم
وزن لایه	۰/۱۱	۰/۰۸۳	۰/۱۳	۰/۱۱۱	۰/۰۸۳	۰/۰۹۷	۰/۰۵۵	۰/۱۲	۰/۱۱۱	۰/۰۹۷

کتابنگاری

شادمان، م.، تخم‌چی، ب. و خیراللهی، ح.، ۱۳۹۱- ارائه رویکردی معتبر جهت تهیه نقشه‌های شبه زمین‌شناسی با استفاده از خوشه‌بندی داده‌های ژئوفیزیک هوابرد، مجله مهندسی معدن، شماره ۱۶، ص ۸-۱.

محجل، م.، ۱۳۸۳- گزارش زمین‌شناسی و ساختاری منطقه مطالعاتی پروژه آلوت، سازمان زمین‌شناسی و اکتشافات معدنی کشور.

یارمحمدی، ع.، راستاد، ا.، محجل، م. و شمس‌ا، م. ح.، ۱۳۸۷- رخداد طلای باریکا: کانه‌زایی تیپ ماسیوسولفید و لکانوژنیک غنی از طلای ایران، مجله علوم دانشگاه تهران، جلد ۳۴، شماره ۱، ص ۶۰-۴۷.

References

- Bonham-Carter, G. F. & Agterberg, F. P., 1989- Weights of evidence modeling: A new approach to mapping mineral potential; in Statistical Application in the Earth Sciences, Geological Survey of Canada, pp. 89-90 and 171-183.
- Caranza, E. J. M., Hale, M. & Faassen, C., 2008- Selection of coherent deposit-type locations and their application in data-driven mineral prospectively mapping, Ore Geology, Reviews 33, pp. 536-558.
- Davis, D. L. & Bouldin, D. W., 1979- A cluster separation measure, IEEE Transaction Pattern Analyzing Machine Intelligence, Vol. 1, pp. 224-227.
- Eberle, D., Tessema, A., Muiuane, E. A. & Daudi, E. X., 2009- Integration of Geophysical and satellite imaginary data from the Alto de Liogoha pegmatite fields, northern Mozambique: implications for the control of mineralization. 11th SAGA Biennial Technical Meeting and Exhibition Swaziland, pp. 21-26.
- Lucieer, V. & Lucieer, A., 2009- Fuzzy clustering for seafloor classification. Marine Geology, Vol. 264, pp. 230-241.
- Martelet, G., Truffert, C., Tourliere, B. & Ledru, P., 2006- classifying airborne radiometry data with agglomerative hierarchical clustering: A tool for geological mapping in context of rainforest (French Guiana), International Journal of Applied Earth Observations and Geo-information, Vol. 8, pp. 208-223.
- Nugraha, H. S., 2011- Integration of stream sediment geochemical and airborne gamma-ray data for surficial lithological mapping using clustering methods, Master of Science Thesis, Twente University, Netherland, Supervisor: E.J.M. Caranza.
- Paasche, H. & Eberle, D. G., 2009- Rapid Integration of large airborne geophysical data suites using a fuzzy partitioning cluster algorithm: a tool for geological mapping and mineral exploration targeting, Exploration Geophysics, Vol. 40, pp. 227-287.
- Paasche, H. & Eberle, D., 2010- Automated integration of large geophysical data using three partitioning cluster algorithms: A Comparison. 11th SAGA Biennial Technical Meeting and Exhibition Swaziland, pp. 286-291.
- Reimann, C., Filzmoser, P., Garrett, R. & Dutter, R., 2008- Statistical data analysis explained: Applied Environmental statistics with R, West Sussex, John Wiley Publication, pp. 341.
- Song, Y. C. & Meng, H. D., 2010- The application of Cluster Analysis in Geophysical data Interpretation, Computer Geosciences, Vol. 14, pp. 263-271.
- Theobald, D. M., 2007- GIS concepts and ArcGIS methods, 3rd edition, Conservation Planning Technology, Colorado State University, USA, pp. 247-256.
- Tokhmechi, B., Memarian, H., Ahmad Noubari, H. & Moshiri, B., 2009- A Novel Approach for fracture Zone detection using petrophysical logs, Journal of Geophysics and Engineering, Vol. 6, pp. 365-373.
- Velmurugan, T., 2010- Performance evaluation of K_Means and Fuzzy K_Means clustering algorithms for statistical distributions of input data points. European Journal of Scientific Research, Vol. 46, No. 3, pp. 320-330.

Comparison of Clustering Methods in Mineral Potential Mapping of Barika Anomaly, West Azerbaijan

M. Shademan ^{1*} & B. Tokhmechi ²

¹ M.Sc., Kavoshgaran Consulting Engineers, Sechahoon Mine, Yazd, Iran

² Assistant Professor, Department of Mining, Petroleum and Geophysics Engineering, Shahrood University of Technology, Shahrood, Iran

Received: 2011 November 27

Accepted: 2012 October 21

Abstract

Nowadays GIS techniques are used as a conventional tool for integrating geographic information datasets. In these methods, integration is done according to quality and quantity of datasets and using appropriate weighting approaches. Finally, with Classification methods like Geometric Distances, Mineral Potential Maps (MPM) is produced. By increasing application in data processing, clustering methods classify samples into groups by similarity between them. In this paper, K_ Means and Fuzzy K_ Means clustering methods are discussed for mapping potential zones of Gold mineralization, then the results are compared with GIS method, Index Overlaying, for the Barika area in 1:100,000 Alut Sheet in South of Azerbaijan. In the Barika area, information of drilling points aren't available, so it's not possible to determine number of classes and boundaries of each class for final score map, but in clustering methods, the optimum number of class for output map is done automatically and is tested for the Barika Anomaly data. The results show that clustering methods need no weighting and it's possible to use it with low information than GIS_ based method. The results also show that both of approaches, clustering methods and Index overlaying, display almost an equal area for the most potential zone, however clustering methods need low information for analyzing, but in Index overlaying, it is needed to have more information for weighting and determining threshold for classification of final scores.

Keywords: Geographic Information System, Fuzzy Clustering, K_ Means, Mineral Potential Map, Barika.

For Persian Version see pages 67 to 72

*Corresponding author: M. Shademan; E-mail: mohseneshademan@yahoo.com