

مقاله پژوهشی

شناسایی الگوهای ژئوشیمیایی در محدوده کانی زایی خوینرود به روش تحلیل تطبیقی و خوشه بندی

حمید گرانیان^{*۱}

۱. استادیار گروه مهندسی معدن، دانشگاه صنعتی بیرجند، h.geranian@birjandut.ac.ir

(دریافت: ۱۳۹۹/۰۴/۲۰ - پذیرش: ۱۴۰۰/۰۲/۱۴)

چکیده

شناسایی الگوهای پراکندگی عناصر و مرتبط کردن آن‌ها با آنومالی‌های ژئوشیمیایی یکی از ابزارهای اکتشافی به ویژه در فازهای نیمه تفصیلی و تفصیلی محسوب می‌شود. روش‌هایی که بتواند نمونه‌ها و عناصر شیمیایی را به طور همزمان مورد تحلیل قرار دهد، برای این منظور پیشنهاد می‌شود. در این مقاله از دو روش تحلیل تطبیقی خوشه‌ای (CCA) و الگوریتم خوشه‌بندی بر اساس توابع توزیع چگالی (DENCLUE) و داده‌های ژئوشیمیایی در فاز اکتشافی نیمه تفصیلی منطقه خوینرود استفاده شده است. خوشه‌بندی ۱۶۵ نمونه برداشت از محیط خاکی به همراه نتایج آنالیز ۷ عنصر مرتبط با کانی‌زایی مس - طلا، پورفیری یعنی $As, Au, Cu, Hg, Pb, S, Zn$ نشان می‌دهد که چهار محدوده A, B, C و D مستعد کانی‌زایی در منطقه قابل مشاهده است. در روش CCA داده‌ها به ۶ خوشه تفکیک شده که خوشه اول شامل عنصر گوگرد به همراه ۵۷ نمونه، خوشه دوم شامل ۶۱ نمونه، خوشه سوم شامل عنصر سرب به همراه ۱۶ نمونه، خوشه چهارم شامل عنصر مس به همراه ۸ نمونه، خوشه پنجم شامل عنصر آرسنیک به همراه ۷ نمونه و خوشه ششم شامل عناصر طلا، جیوه و روی به همراه ۱۷ نمونه است. در حالی که نتایج خوشه‌بندی داده‌ها به روش DENCLUE شامل ۵ خوشه به ترتیب به صورت گوگرد با ۶۶ نمونه، ۴۳ نمونه، سرب و روی با ۳۸ نمونه، طلا و مس با ۱۰ نمونه و آرسنیک و جیوه با ۸ نمونه است. بخش C2 از محدوده C و بخش D2 از محدوده D به عنوان بهترین محدوده‌ها با احتمال کانی‌زایی پورفیری و محدوده A به احتمال کانی‌زایی رگه‌ای هیدروترمال پیشنهاد می‌شود. محدوده‌های B و D1 نیز با احتمال کانی‌زایی رگه‌ای با نیاز به اکتشاف تکمیلی، پیشنهادات بعدی‌اند. همچنین نتایج نشان‌دهنده خوشه‌بندی بهتر عناصر، انطباق بهتر محدوده‌های پیشنهادی برای کانی‌زایی با آنومالی‌های ژئوشیمیایی عناصر و شرایط زمین‌شناسی منطقه مورد مطالعه از برتری‌های الگوریتم DENCLUE است. بنابراین برای ارتباط الگوی پراکندگی عناصر با آنومالی ژئوشیمیایی آن‌ها می‌توان از این الگوریتم استفاده کرد.

کلمات کلیدی

تحلیل تطبیقی، خوشه‌بندی DENCLUE، پتانسیل‌یابی معدنی، آنومالی ژئوشیمیایی چند عنصری، محدوده کانی‌زایی خوینرود.

* نویسنده مسئول مکاتبات.

۱- مقدمه

برای شناسایی عناصر و مناطق مستعد کانی‌زایی با خوشه‌بندی داده‌ها در دو مد R و Q [۳۱ و ۳۳] و شناسایی الگوهای ژئوشیمیایی و زمین‌شناسی در تشکیل خاک‌های متفاوت [۳۹] نیز به کار رفته‌اند. روش خوشه‌بندی اولین بار برای تفکیک کانسارهای مشابه بر اساس متغیرهای زمین‌شناسی به کار رفته است [۱۶]. از این روش برای تفکیک عناصر مرتبط با کانی‌زایی [۴۵]، جداسازی جوامع سنگی برای حذف مولفه سنژنتیک در داده‌های رسوبات آبراه‌های [۹] و دسته‌بندی عناصر بر اساس رفتار آماری آن‌ها [۱۰] نیز استفاده شده است. روش‌های خوشه‌بندی به پنج گروه روش‌های مبتنی بر گرانیگاه، روش‌های بر پایه اتصال، روش‌های بر پایه توزیع داده‌ها، روش‌های مبتنی بر چگالی و روش‌های مبتنی بر گرید تقسیم‌بندی می‌شوند. از مهم‌ترین الگوریتم‌های خوشه‌بندی استفاده شده در تحلیل داده‌های ژئوشیمیایی می‌توان به الگوریتم‌های سلسله‌مراتبی [۱۸] و [۵۱]، فازی [۲۱ و ۵۱]، k- میانگین [۱۵] و مدل- ترکیبی [۵۳] اشاره کرد.

از آنجا که روش‌های مرسوم آماری تعیین آنومالی‌های ژئوشیمیایی، معایب حساس بودن به وجود داده‌های خارج از ردیف، لزوم تبعیت از توزیع نرمال و عدم امکان اختصاص آنومالی به چند متغیر معین را دارند؛ در این مقاله ضمن معرفی مجدد روش تحلیل تطبیقی، از این روش برای ترکیب داده‌ها در دو مد R و Q، استفاده شده و خوشه‌بندی هم‌زمان داده‌ها انجام شده است. همچنین در ادامه، الگوریتم خوشه‌بندی بر اساس توابع توزیع چگالی^۲ برای متخصصان علوم زمین معرفی شده و نشان داده شده است که این الگوریتم به دلیل مقاوم بودن، خوشه‌بندی داده‌ها به صورت مستقیم و با شکل‌های دلخواه و به کارگیری چگالی تابع توزیع داده‌ها چگونه می‌تواند در شناسایی الگوهای ژئوشیمیایی کمک کند. برای این منظور از داده‌های ژئوشیمیایی منطقه خوینرود استفاده شده که کلیه فازهای اکتشافی آن انجام گرفته است.

۲- محدوده مورد مطالعه

منطقه کانی‌زایی خوینرود در استان آذربایجان شرقی، در ۲۵ کیلومتری شمال شهر ورزقان و ۷۵ کیلومتری شمال شرقی شهرستان اهر قرار دارد. این منطقه از نظر موقعیت طبیعی در بخش میانی رشته کوه‌های قره‌داغ قرار گرفته و از نظر ژئومورفولوژی از دو بخش کوهستانی و کوهپایه‌ای تشکیل شده است. بر اساس نقشه پهنه‌های رسوبی- ساختاری ایران (شکل ۱) این منطقه در کمان ماگمایی ارومیه- دختر و در ایالت

شناسایی و برقراری ارتباط بین آنومالی‌های ژئوشیمیایی و الگوهای پراکندگی عناصر می‌تواند باعث تفکیک آنومالی‌های مرتبط با کانی‌زایی از نوع ضعیف و بی‌اهمیت شود. پراکندگی عناصر معمولاً ناشی از پدیده‌های فیزیکی از قبیل فرآیندهای مرتبط با تزریق ماگما و سیالات هیدروترمال و یا پدیده‌های شیمیایی و بیوشیمیایی است [۴۱ و ۵۵]. تشخیص الگوهای پراکندگی که باعث تجمع و غنی‌شدگی سنگ‌های یک منطقه از عناصر کانه‌ساز می‌شود، مستلزم تعیین مقدار زمینه و آنومالی در آن مناطق است [۵۵]. روش‌های آماری و الگوریتم‌های شناسایی الگو، به ویژه روش‌هایی که ارتباط بین نمونه‌ها و متغیرها را برقرار می‌کنند، می‌توانند در این راستا قابل کاربرد باشند. با توجه به ماهیت چند عنصری بودن داده‌های ژئوشیمیایی و غنی‌شدگی هم‌زمان محیط‌های ژئوشیمیایی برای چند عنصر هم پارژنز، استفاده از روش‌های چند متغیره بر تک متغیره برتری خواهند داشت [۸ و ۲۹].

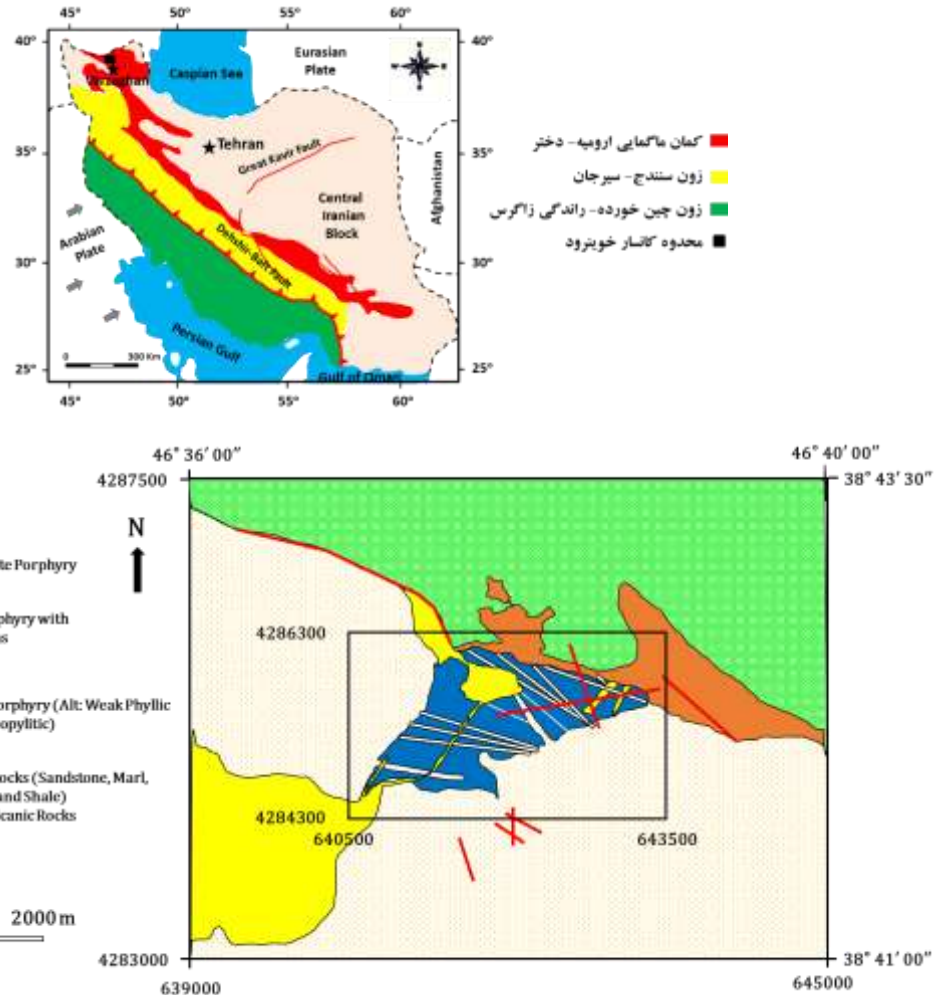
برای مشخص کردن یک آنومالی تک متغیره می‌توان از پارامترهای آماری، نمودارها، آماره‌ها و روش‌های فرکتالی استفاده کرد [۱۴، ۱۷ و ۴۸]. در حالی که برای تعیین آنومالی ژئوشیمیایی چند متغیره سه راه حل تعیین آنومالی‌های ژئوشیمیایی تک متغیره و سپس تلفیق بصری یا ریاضی آن‌ها، تبدیل داده‌های چند متغیره به یک مقدار با استفاده از روش هاله‌های ژئوشیمیایی مرکب جمعی یا ضربی و یا روش‌های آمار چند متغیره و سپس استفاده از روش تعیین آنومالی تک مقداره و استفاده از روش‌های خوشه‌بندی برای تفکیک جوامع آماری چند متغیره پیشنهاد شده است [۳۴، ۱۳ و ۲۳]. روش‌های آماری کاهش بعد از قبیل تحلیل مولفه‌های اصلی و تحلیل فاکتوری مهم‌ترین روش‌های تعیین آنومالی‌های ژئوشیمیایی چندعنصری محسوب می‌شوند [۷، ۴۹ و ۵۰]. اما وجود مقادیر خارج از ردیف چند متغیره و غیرنرمال بودن توزیع داده‌های از مهم‌ترین معایب کاربرد این روش‌ها است [۱۵]. در حالی که روش‌های تحلیل تطبیقی^۱ و خوشه‌بندی فاقد چنین معایبی است، ضمن آن که، این روش‌ها می‌توان ارتباط مستقیمی را بین نمونه‌ها و متغیرها برقرار کند. اولین استفاده از روش تحلیل تطبیقی بر روی داده‌های ژئوشیمیایی، کاربرد آن در ارتباط دادن عناصر اصلی و کمیاب با واحدهای سنگی حاوی کانی‌زایی است [۳۷]. این روش برای ارتباط دادن متغیرها با نمونه‌های حاوی مقادیر آنومال [۳، ۴، ۵۱ و ۵۴]، تفسیر داده‌های ژئوشیمیایی در مقیاس ناحیه‌ای

1-Correspondence Analysis (CA)

2-DENsity-based CLUstEring (DENCLUE)

گسل زرینه رود شده است [۲]. مطابق شکل ۱، از نظر زمین‌شناسی، سنگ‌های منطقه را می‌توان به چهار بخش زیر خلاصه کرد [۶]:

متالورژی ارسباران واقع شده است [۱ و ۲]. زمین‌شناسی ناحیه متأثر از دو دسته پدیده است، یکی پدیده‌های تکتونیکی دونین که عامل اصلی تشکیل گسل تبریز بوده و دیگری حرکات مهم تریاس بالایی که باعث شکافته شدن پلاتفرم آذربایجان و تشکیل



شکل ۱- موقعیت محدوده مورد مطالعه در نقشه زون‌های ساختاری ایران (شکل بالا) و نقشه ساده شده زمین‌شناسی منطقه کانی‌زایی خوینرود (برگرفته از نقشه زمین‌شناسی - اکتشافی ۱:۵۰۰۰ منطقه) به همراه موقعیت شکل ۴ (محدوده مستطیلی شکل) (شکل پایین) [۲۴].

از نوع پورفیری، اسکارنی و رگه‌ای ایجاد کرده است.

- توده هورنبلند دیوریت پورفیری (HDP) و دایک‌های وابسته که فاقد آلتراسیون است.

از نظر تکتونیکی نیز دو دسته گسل در منطقه قابل مشاهده است؛ یکی گسل‌های امتداد لغز در شمال محدوده مربوط به کوهزایی آلپی پسین و دیگری گسل‌های غالباً شمالی- جنوبی مربوط به حرکات زمین‌ساختی بعد از ائوسن است که گسل‌های قدیمی‌تر را قطع می‌کند [۶].

در زون ارسباران سه فاز ماگمایی مرتبط با کانه‌زایی شامل ائوسن میانی تا اواخر الیگوسن، اواخر الیگوسن تا میوسن زیرین و میوسن فوقانی تا پلیوسن رخ داده است. فاز ماگمایی منطقه خوینرود

- سنگ میزبان منطقه شامل سازندهای رسوبی با جنس آهکی، ماسه‌سنگی، شیلی و رخساره فلیشی (FLY) و سنگ‌های آتشفشانی از نوع بازالتی تا آندزیت- بازالت (BAS) متعلق به کرتاسه بالایی تا پالئوسن است.

- توده نفوذی کوارتز مونزونیتی خوینرود با بافت پورفیری (KP) که به شکل یک استوک کوچک در وسط محدوده قابل مشاهده است. نفوذ این توده باعث دگرگونی مجاورتی در نهشت‌های فلیشی و ایجاد آلتراسیون‌های هیدروترمال در سنگ‌ها شده است.

- توده‌های نفوذی گرانودیوریتی تا کوارتزدیوریتی پورفیری (DP) که سنگ‌های گروه اول را قطع کرده و کانی‌زایی‌هایی

حال دو ماتریس قطری [R] با بعد N×N و [C] با بعد d×d را به شکل رابطه ۵ و ۶ تعریف می‌شود [۳۳ و ۵۱]:

$$[R] = \text{diag} (r_1, r_2, \dots, r_N) \quad (5)$$

$$[C] = \text{diag} (c_1, c_2, \dots, c_d) \quad (6)$$

که diag: نشان‌دهنده قطری بودن ماتریس است. سپس ماتریس تبدیل (W) و ماتریس شباهت (H) بین نمونه‌ها و متغیرها از رابطه ۷ و ۸ بدست می‌آید [۳۱]:

$$[W] = [R]^{-\frac{1}{2}}[D][C]^{-\frac{1}{2}} \quad (7)$$

$$[H] = [w]^T [w] = [C]^{-\frac{1}{2}}[D]^T [R]^{-1}[D][C]^{-\frac{1}{2}} \quad (8)$$

که در این روابط توان $-\frac{1}{2}$ بدان معنی است که اولاً عناصر ماتریس به توان $\frac{1}{2}$ می‌رسند و سپس معکوس ماتریس محاسبه می‌شود و T به معنی ترنهاد ماتریس است. ماتریس [H] یک ماتریس d×d بعدی خواهد بود. در مرحله بعد مقادیر ویژه و بردارهای ویژه این ماتریس محاسبه می‌شود. مقادیر ویژه این ماتریس d عدد خواهند بود که بین صفر و یک قرار دارند. با حذف مقدار ویژه یک، اغلب p مقدار ویژه (p<d) انتخاب می‌شود، بنابراین رابطه ۹ به صورت زیر است [۳۱]:

$$0 \leq \lambda_p \leq \dots \leq \lambda_2 \leq \lambda_1 \leq 1 \quad (9)$$

اگر λ_j مقدار ویژه و $[a_j]$ بردار ویژه منظر با آن باشد، سپس دو ماتریس ۱۰ و ۱۱ را می‌توان تعریف کرد [۳۳ و ۳۱]:

$$[A]_{(p \times p)} = \text{diag} (\lambda_1, \lambda_2, \dots, \lambda_p) \quad (10)$$

$$[A]_{(m \times p)} = [[a_1], [a_2], \dots, [a_p]] \quad (11)$$

که در آن:

[A]: یک ماتریس قطری است که عناصر روی قطر اصلی آن مقادیر ویژه یاد شده‌اند.

[A]: ماتریسی است که هر ستون آن را یکی از بردارهای ویژه تشکیل می‌دهد.

سپس دو ماتریس مهم زیر که نتیجه تحلیل تطبیقی است، از روابط ۱۲ و ۱۳ بدست می‌آید [۳۳ و ۳۱]:

$$[V] = [R]^{-\frac{1}{2}}[W][A] \quad (12)$$

$$[U] = [C]^{-\frac{1}{2}}[A][A]^{-\frac{1}{2}} \quad (13)$$

ماتریس [V] یک ماتریس N×p بعدی است که روابط بین نمونه‌ها (Q مد) را بیان می‌کند و ماتریس [U] یک ماتریس d×p

متعلق به اواخر میوسن تا پلیوسن است که دارای ترکیب کوارتز دیوریتی و کوارتز مونزونیتی است [۵]. سنگ‌های مونزونیتی منطقه به صورت یک توده نفوذی با بافت پورفیری است که آلتراسیون‌های پروپلیتیک، فلیک و آرژیلیک دارد. این توده پورفیری با دایک‌های دیوریتی قطع شده که کانی‌زایی بیشتر در این دایک‌ها انجام گرفته است. کانی‌زایی در منطقه، از نوع مس-طلا پورفیری و رگه‌های اپی‌ترمال کوارتز-طلا دار است [۵]. مهم‌ترین کانی‌های مرتبط با کانی‌زایی در منطقه نیز شامل پیریت، کوپریت، مالاکیت، کالکوپریت، گالن، اسفالریت، سروزیت، آرسنوپریت، استینیت، مولیبدنیت، کانی‌های جیوه و اکسیدهای آهن است. همچنین طلا به صورت دانه‌های آزاد و درون کانی‌های پیریت و آرسنوپریت قابل مشاهده است [۵ و ۶].

۳- روش‌ها

۳-۱- روش تحلیل تطبیقی

روش تحلیل تطبیقی (CA) اولین بار توسط بنزسری^۱ در سال ۱۹۶۴ ارائه شده است [۱۱]. در این روش، در یک دستگاه مختصات و تحت یک سیستم فاکتوری موقعیت متغیرهای نسبت به هم و توام با موقعیت نمونه‌های نمایش داده می‌شود. اگر مجموعه داده‌های چندمتغیری (چند بعدی) به صورت رابطه ۱ تعریف شود:

$$X = \{x_1, x_2, \dots, x_N\} \quad (1)$$

که در آن:

$x_i \in R^d$: یک نمونه یا داده.

N: تعداد نمونه‌ها.

d: تعداد متغیرها یا بعد داده‌ها.

اولین مرحله در پردازش داده‌ها به روش تحلیل تطبیقی، محاسبه ماتریس فراوانی نسبی یا استانداردسازی داده‌ها به صورت رابطه ۲ است [۱۲ و ۵۱]:

$$D = X / \sum_{i=1}^N \sum_{j=1}^d x_{ij} \quad (2)$$

سپس بردار ستونی وزن هر نمونه [r] و بردار سطری وزن هر متغیر [c] با جمع مقادیر به ترتیب سطرها و ستون‌ها ماتریس D به صورت ۳ و ۴ محاسبه می‌شود:

$$r_i = d_{i1} + d_{i2} + \dots + d_{id} \quad i = 1, 2, \dots, N \quad (3)$$

$$c_j = d_{1j} + d_{2j} + \dots + d_{Nj} \quad j = 1, 2, \dots, d \quad (4)$$

کرنل، با استفاده از تابع چگالی احتمالی مدل آماری مجموعه داده‌ها از رابطه ۱۶ محاسبه می‌شود [۳۵ و ۵۶]:

$$\hat{f}_\sigma(x) = \frac{1}{N} \sum_{i=1}^N \frac{1}{\sigma(x_i)^d} K\left(\frac{x-x_i}{\sigma(x_i)}\right) \quad (16)$$

که در آن:

$\sigma(x_i)$: پهنای باند بر روی نمونه x_i است و معمولاً برای ساده‌سازی برابر مقدار ثابتی در نظر گرفته می‌شود (یعنی $\sigma(x_i) = \sigma$).
 σ : یک پارامتر هموارساز.

$K(u)$ تابع کرنل انتگرال‌پذیر با مقدار واقعی و غیرمنفی است که برای تمام مقادیر u دارای دو شرط زیر است [۲۵]:

$$\int_{-\infty}^{+\infty} k(u) du = 1 \quad \text{and} \quad K(-u) = K(u) \quad (17)$$

تابع گوسی چندمتغیره استاندارد با میانگین صفر و واریانس ۱ بیشترین کاربرد را به عنوان تابع کرنل دارد (رابطه ۱۸) [۳۵]:

$$K(u) = \frac{1}{(2\pi)^{\frac{d}{2}}} \exp\left(-\frac{\|u\|^2}{2}\right) \quad (18)$$

با ترکیب روابط ۱۶ و ۱۸، تقریب چگالی کرنل برای نمونه x_j از رابطه ۱۹ بدست می‌آید:

$$\hat{f}(x_j) = \frac{1}{N} \sum_{i=1}^N \frac{1}{(2\pi)^{\frac{d}{2}} \sigma^d} \exp\left(-\frac{\|x_j - x_i\|^2}{2\sigma^2}\right) \quad (19)$$

$$j = 1, 2, \dots, N$$

مقدار بدست آمده از رابطه ۱۹ تابع چگالی هر نمونه خواهد بود که متاثر از تابع تاثیر کلیه نمونه‌ها است. برای پیدا کردن مرکز خوشه‌ها در الگوریتم DENCLUE از تعریف نقاطی به نام رباینده چگالی^۵ استفاده می‌شود. نقطه x_* هنگامی یک رباینده چگالی خوانده می‌شود که دارای یک حداکثر محلی از تابع چگالی باشد. برای اجتناب از حداکثر محلی جزئی، در این الگوریتم از یک حد آستانه نویز مانند ξ نیز استفاده می‌شود، بنابراین تنها آن رباینده‌های چگالی به عنوان مراکز خوشه در نظر گرفته می‌شوند که دارای شرط $\hat{f}(x_*) \geq \xi$ باشند [۲۷ و ۳۰].

در مرحله بعدی، نمونه‌ها از طریق رباینده‌های چگالی و با کمک یک فرآیند تپه‌نوردی^۶ به خوشه‌ها انتساب داده می‌شوند. در این فرآیند برای هر نمونه x رویه تپه‌نوردی کار خود را از

بعدی است که روابط یا فاکتورهای بین متغیرها (R مد) را نشان می‌دهد. ماتریس نهایی فاکتورها از ترکیب دو ماتریس قبلی و به صورت رابطه ۱۴ محاسبه می‌شود:

$$[F] = \begin{bmatrix} [V] \\ [U] \end{bmatrix} \quad (14)$$

ماتریس حاصل یک ماتریس $N+d$ سطری و p ستونی خواهد بود. N سطر اول معرف نمونه‌ها و d سطر بعدی معرف متغیرها خواهد بود. ستون‌های این ماتریس نیز معرف p فاکتور است [۱۲]. نتایج حاصل از تحلیل تطبیقی را می‌توان به دو صورت زیر ارائه داد:

- در یک یا چند نمودار پراکندگی متغیرها و نمونه‌ها را به نمایش گذاشت که در این صورت به آن تحلیل تطبیقی فاکتوری^۱ گفته می‌شود. برای بررسی فاکتورهای تطبیقی، دو به دو ستون‌های ماتریس $[F]$ را که همان فاکتورها است، در یک نمودار پراکندگی رسم می‌شود. بنابراین تعداد $\frac{p(p-1)}{2}$ نمودار بدست می‌آید. در این نمودارها متغیرهای جدا شده نشان‌دهنده عناصر دارای احتمال آنومالی و نمونه‌های جدا شده از بقیه نشان‌دهنده نمونه‌های دارای مقادیر آنومال خواهند بود.
- می‌توان نتایج را به صورت دندوگرام ارائه کرد که در این حالت به آن تحلیل تطبیقی خوشه‌ای^۲ گفته می‌شود. تحلیل خوشه‌ای در این حالت می‌تواند در دو حالت R و Q مد انجام شود. برای این منظور می‌توان از فاصله اقلیدسی به عنوان معیار شباهت و از رابطه ۱۵ استفاده کرد [۲۹ و ۳۰]:

$$d_{ij} = \sqrt{\sum_{k=1}^p (f_{ik} - f_{jk})^2} \quad i, j = 1, 2, \dots, N + d \quad (15)$$

در این رابطه f_{ij} عناصر ماتریس $[F]$ است.

۳-۲- الگوریتم خوشه‌بندی بر اساس توابع توزیع چگالی

۳-۲-۱- معرفی الگوریتم

الگوریتم خوشه‌بندی بر اساس توابع توزیع چگالی (DENCLUE) یکی از روش‌های خوشه‌بندی بر اساس چگالی است که برای اولین بار توسط کیم و هینبرگ^۳ در سال ۱۹۹۸ پیشنهاد شده است [۲۷]. در این روش چگالی اطراف هر نمونه به کمک روش برآورد چگالی کرنل^۴ محاسبه می‌شود که یک رویکرد برآورد چگالی بدون پارامتر آماری محسوب می‌شود. تقریب چگالی

4-Kernel Density Estimation (KDE)

5-Density Attractor

6-Hill climbing procedure

1-Correspondence Factor Analysis (CFA)

2-Correspondence Cluster Analysis (CCA)

3-Keim and Hinneburg

مجموعه داده‌ها.

- شناسایی سلول‌های (ابرمکعب‌های) دارای نمونه.
- محاسبه میانگین سلول‌های دارای نمونه.
- اتصال سلول‌هایی که فاصله میانگین دو سلول دارای نمونه کمتر از σ^4 باشد.
- در نظر گرفتن سلول‌های متصل شده برای تشکیل خوشه‌ها.
- پیدا کردن رباینده‌های چگالی به کمک فرآیند تپه‌نوردی.
- انتخاب تصادفی یک نمونه (i) .
- محاسبه تابع چگالی محلی نمونه i تا نمونه‌هایی با فاصله σ^4 .
- رفتن به نمونه نزدیک بعدی $(i+1)$ که تابع چگالی آن قبلاً محاسبه شده است..
- ادامه فرآیند تپه‌نوردی، اگر تابع چگالی نمونه i کوچکتر از تابع چگالی نمونه $i+1$ باشد.
- قراردادن نمونه‌های مسیر تپه‌نوردی در یک خوشه.
- اتصال رباینده‌های چگالی و تشکیل خوشه‌ها.

بنابراین برای خوشه‌بندی داده‌ها به کمک الگوریتم DENCLUE علاوه بر ماتریس داده‌ها نیاز به ورود دو پارامتر σ و ξ توسط کاربر نیز است. پارامتر σ نشان‌دهنده تاثیر یک نمونه در فضای داده‌ها و پارامتر ξ نشان می‌دهد که چه وقت یک رباینده چگالی اهمیت دارد. شکل ۲ تاثیر این دو پارامتر در خوشه‌بندی را نشان می‌دهد. مطابق این شکل، افزایش مقدار σ باعث کاهش تعداد خوشه‌ها و کاهش مقدار ξ نیز باعث افزایش تعداد خوشه‌ها و تغییر شکل خوشه‌ها می‌شود [۲۷ و ۴۰]. در نتیجه انتخاب دقیق مقادیر این دو پارامتر در نتیجه خوشه‌بندی موثر خواهد بود.

برای برآورد مقادیر بهینه برای این پارامترها روش‌های مختلفی ارایه شده است. Obthong و Sriphum در سال ۲۰۱۱ استفاده از الگوریتم کلنی مورچه‌گان را پیشنهاد کرده‌اند. در این الگوریتم مقدار بهینه پارامتر σ در نقطه‌ای خواهد بود که چگالی آنتروپی (H) محاسبه شده از رابطه ۲۲ مینیمم باشد [۳۸]:

$$H = - \sum_{i=1}^N \frac{\hat{f}(x_i)}{Z} \log \left(\frac{\hat{f}(x_i)}{Z} \right) \quad (22)$$

که $0 \leq H \leq \log(N)$ و $\sigma = [0, +\infty]$ ، $Z = \sum_{i=1}^N \hat{f}(x_i)$ است. یکی دیگر از روش‌های محاسبه مقدار بهینه پارامتر σ استفاده از مقادیر پراکندگی در مجموعه داده‌ها است. این مقدار از رابطه ۲۳ بدست می‌آید [۳۰ و ۵۲]:

همان نقطه شروع می‌کند و با کمک تغییر تابع چگالی به طرف رباینده‌های چگالی راهنمایی می‌شود. در حالتی که از تابع کرنل گوسی استفاده شود، مقدار تپه‌نوردی برابر گرادیان تابع چگالی خواهد بود (رابطه ۲۰) [۲۶ و ۲۷]:

$$\nabla \hat{f}(x) = \frac{\partial \hat{f}(x)}{\partial x} = \frac{1}{N\sigma^{d+2}} \sum_{i=1}^N K\left(\frac{x-x_i}{\sigma}\right)(x-x_i) \quad (20)$$

و مقدار تغییر تابع چگالی به کمک رابطه ۲۱ محاسبه می‌شود:

$$x_0 = x, x_{i+1} = x_i + \delta \cdot \frac{\nabla \hat{f}(x_i)}{\|\nabla \hat{f}(x_i)\|} \quad (21)$$

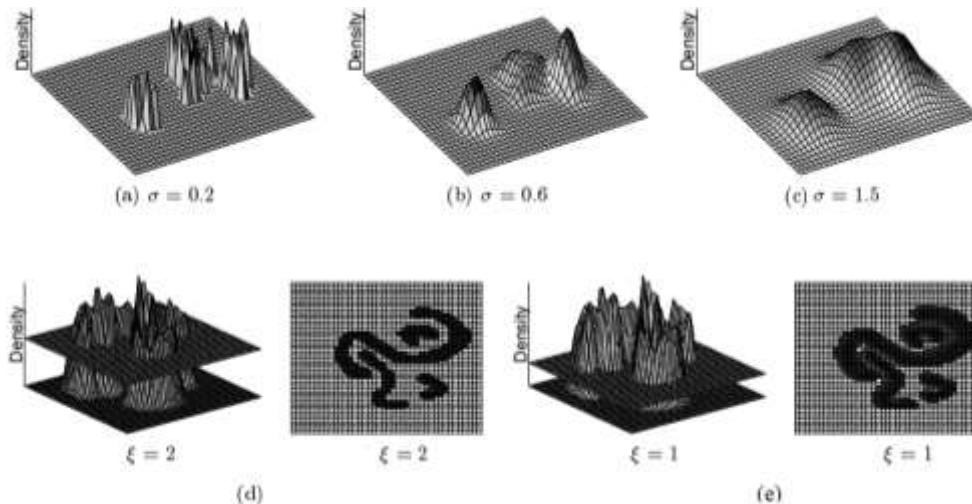
فرآیند تپه‌نوردی در گام k ام خود که شرط $\hat{f}(x_{k+1}) < \hat{f}(x_k)$ برقرار شود، متوقف شده و نمونه x را به رباینده چگالی $x_* = x_k$ تخصیص می‌دهد. بدین ترتیب نمونه‌ها به خوشه‌ها نسبت داده شده و خوشه‌ها تشکیل می‌شوند. رباینده چگالی که دارای یک حداکثر محلی جزیبی بوده است و شرط $\hat{f}(x_*) < \xi$ برقرار باشد نیز یک داده پرت یا نویز محسوب می‌شود [۲۶ و ۲۷].

الگوریتم DENCLUE می‌تواند با کمک چندین رباینده چگالی متصل به هم و از طریق مسیرهای بین هر زوج رباینده، خوشه‌هایی با شکل دلخواه را پیدا کند. پایه ریاضی قوی، عملکرد خیلی خوب بر روی مجموعه داده با نویز یا داده پرت بالا، سرعت عمل بالاتر در خوشه‌بندی داده‌ها به ویژه داده‌ها با ابعاد بالا و پیچیدگی محاسباتی کمتر نسبت به سایر روش‌های خوشه‌بندی بر پایه چگالی از مهم‌ترین مزیت‌های این الگوریتم محسوب می‌شود [۴۳، ۴۴ و ۵۰].

۲-۲-۳- مراحل الگوریتم

به طور کلی، الگوریتم DENCLUE دارای دو مرحله پیش‌خوشه‌بندی و مرحله خوشه‌بندی است. در مرحله اول یک نقشه از ابرچهارگوشه‌ها^۱ بر روی مجموعه داده‌ها ساخته شده و سپس ابرمکعب‌های^۲ (ابرچهارگوشه‌های) دارای نمونه‌ها زیاد شناسایی می‌شوند. در مرحله دوم به کمک فرآیند تپه‌نوردی، رباینده‌های چگالی مشخص شده و با اتصال رباینده‌های چگالی دارای مسیر مشابه و نمونه‌های متصل شده به آن‌ها، خوشه‌ها تشکیل می‌شوند [۲۷ و ۴۶]. الگوریتم DENCLUE دارای گام‌های اجرایی زیر است [۲۷ و ۵۰]:

- تهیه یک نقشه شامل ابرچهارگوشه‌هایی با پهنا σ^2 بر روی



شکل ۲- تاثیر مقادیر مختلف پارامتر σ (شکل‌های a, b, c) و پارامتر ξ (شکل‌های d و e) در نتیجه خوشه‌بندی [۲۷].

در یک خوشه قرار داده می‌شود [۳۶].

در رابطه ۲۱ پارامتر δ کنترل‌کننده سرعت همگرایی برای رسیدن به نقطه‌ای با تابع چگالی حداکثر محلی است. بنابراین زمان رسیدن و مکان احتمالی این نقطه متأثر از این پارامتر خواهد بود. برای رفع این مشکل هینبرگ و گابریل^۲ در سال ۲۰۰۷ مدل DENCLUE 2.0 را پیشنهاد کردند. در این مدل مقدار تغییر تابع چگالی از رابطه ۲۵ محاسبه می‌شود که این رابطه با مشتق‌گیری از رابطه ۲۱ بدست می‌آید [۲۶]:

$$x_{i+1} = \frac{\sum_{i=1}^N K\left(\frac{x-x_i}{\sigma}\right)x_i}{\sum_{i=1}^N K\left(\frac{x-x_i}{\sigma}\right)} \quad (25)$$

رسیدن به نقطه ماکزیمم محلی با گام‌های تکراری کمتر، همگرایی دقیق‌تر به طرف نقطه ماکزیمم محلی و مطابقت بیشتر فرآیند تپه‌نوردی با شکل دلخواه خوشه (شکل ۳) از مزیت‌های مدل DENCLUE 2.0 نسبت به مدل DENCLUE 1.0 است [۲۶].

استفاده از الگوریتم ژنتیک^۳ و الگوریتم تبرید شبیه‌سازی‌شده^۴ منجر به ارایه دو مدل دیگر از این الگوریتم خوشه‌بندی به ترتیب با عنوان‌های DENCLUE-GA و DENCLUE-SA توسط Idrissi و همکارانش در سال ۲۰۱۵ شده است. این الگوریتم به لحاظ زمان اجرا و کیفیت خوشه‌ها نسبت به الگوریتم اولیه برتری دارند [۲۸]. یکی دیگر از مدل‌های بهبود یافته، مدل DENCLUE-IM است که در سال ۲۰۱۶ توسط Rehioui و همکارانش ارایه شده است. تفاوت مدل ارایه شده با مدل اولیه، در محاسبه تابع چگالی برای هر نقطه توسط نقاط موجود در داخل هر ابرمکعب ($f_X(x_{Hcube})$) به جای محاسبه تابع چگالی توسط کلیه نقاط موجود در مجموعه داده‌ها

$$\sigma = \sqrt{\prod_{i=1}^d \sigma_i^2} \quad (23)$$

$$\sigma_i = \left(\frac{4}{3N}\right)^{1/5} \times \min\left\{s_i, \frac{Q_{i3} - Q_{i1}}{1.34}\right\}$$

$$i = 1, 2, \dots, d$$

که در آن:

s_i : انحراف معیار متغیر i ام.

Q_{i1} و Q_{i3} : نیز به ترتیب چارک سوم و اول متغیر i ام.

در این حالت مقدار بهینه پارامتر ξ برابر با رابطه ۲۴ است [۲۷ و ۲۲]:

$$\xi = \|C_N\| \cdot c \cdot \sqrt{2\pi\sigma^2} \quad (24)$$

که در آن:

$\|C_N\|$: تعداد نمونه‌های نویز یا پرت در داده‌ها.

c : مقدار ثابت $0 < c \leq 1$.

در عمل به دلیل کم بودن تعداد نمونه‌های نویز، می‌توان این تعداد را بر اساس هیستوگرام توزیع داده‌ها و برابر تعداد بهینه ستون‌های این هیستوگرام (یعنی $L = [10 \cdot \log_{10}(N)]$) برآورد کرد [۲۲].

۳-۲-۳- مدل‌های الگوریتم

اولین مدل تعمیم‌یافته برای الگوریتم DENCLUE توسط لی^۱ و همکارانش در سال ۲۰۰۳ ارایه شده است. در این مدل که با عنوان DENCLUE-M شناخته می‌شود، برای جلوگیری از اشتغال بیشتر حافظه و کاهش محاسبات، از تفکیک فضای داده‌ها به شبکه‌ای با پهنای سلول σ/\sqrt{d} و محاسبه چگالی هر سلول به جای محاسبه چگالی هر نمونه استفاده می‌شود. سپس نمونه‌های موجود در هر سلول که چگالی سلول آن بیشتر از آستانه باشد

3-Genetic algorithm
4-Simulated annealing algorithm

1-Li
2-Hinneburg and Gabriel

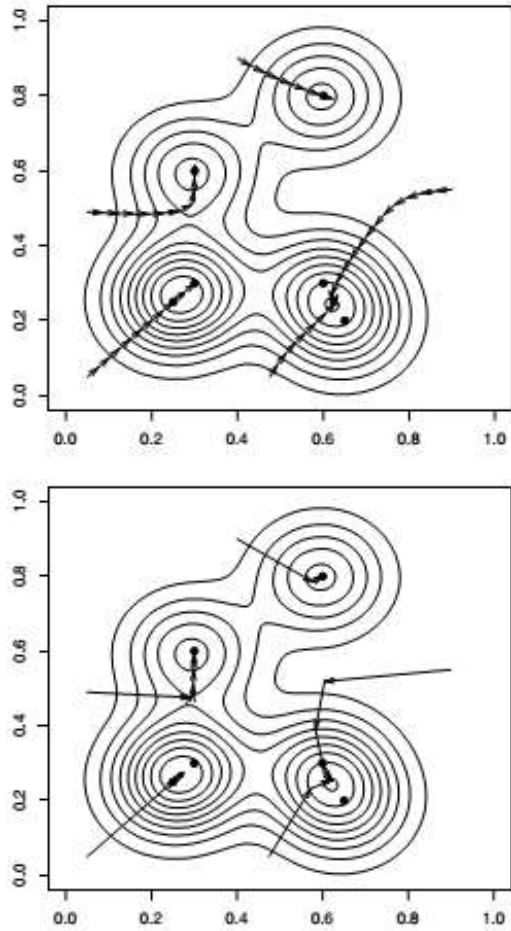
$(\hat{f}_X(x))$ است (رابطه ۲۶) [۴۶]:

$$\forall x \in C_p \hat{f}_X(x) \leq \hat{f}_X(x_{Hcube}) \quad (26)$$

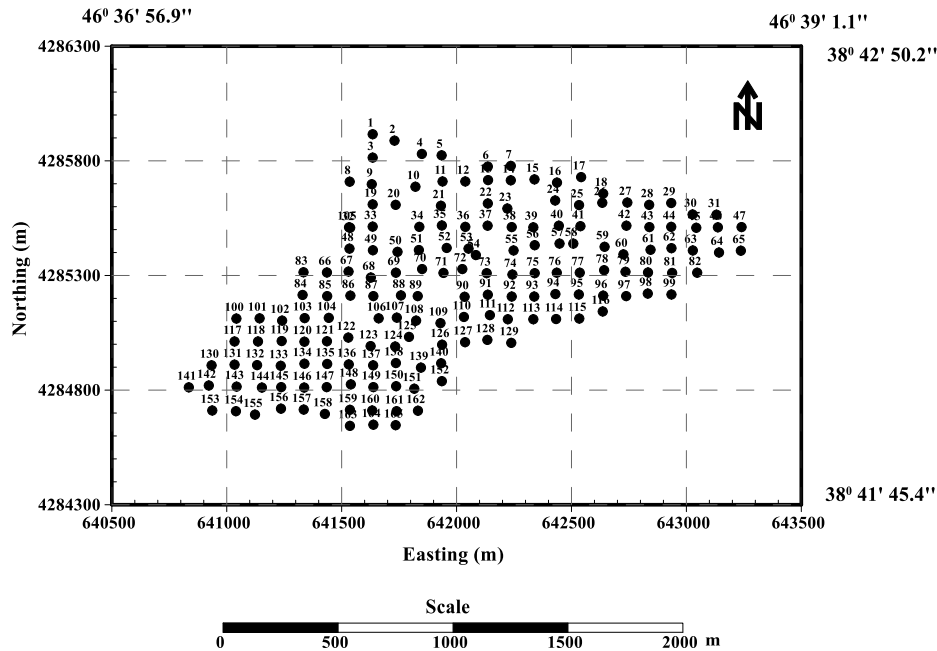
که C_p سلول‌ها یا ابرمکعب‌هایی‌اند که حاوی نمونه‌اند. سپس هر ابرمکعب خوشه اولیه را تشکیل می‌دهد که با x_{Hcube} نشان داده می‌شود. در نهایت خوشه‌های اولیه x_{Hcube} در هم ادغام شده و خوشه‌های نهایی را تشکیل خواهند داد اگر دارای مسیر یکسان (دارای اتصال) باشند [۴۶].

۴- داده‌های ژئوشیمیایی

داده‌های ژئوشیمیایی محدوده مورد مطالعه مربوط به فاز اکتشافی نیمه تفصیلی فلزات پایه و گرانبها در منطقه کانی‌زایی خوینرود و در مقیاس ۱:۵۰۰۰ است. در این فاز اکتشافی ۱۶۵ نمونه از افق B خاک در یک شبکه مربعی 100×100 متری از محدوده‌ای به ابعاد حدودی 2000×3000 متر برداشت شده است (شکل ۴). نمونه‌ها از عمق ۳۰ تا ۴۰ سانتی‌متری سطح زمین و از رئوس یک مثلث به مرکز نقطه نمونه‌برداری انجام شده است. سپس بخش زیر ۵۰ مش (به دلیل تمرکز فلزات در این بخش) به منظور تجزیه شیمیایی برای ۲۹ عنصر (۲۸ عنصر به روش ICP-MS و عنصر طلا به روش Fire Assay انتخاب شده است) (نمونه‌برداری‌ها و تجزیه شیمیایی توسط شرکت اسپیر انجام گرفته است). برای ارزیابی صحت و دقت نمونه‌برداری و تجزیه، ۲۴ نمونه تکراری نیز برداشت شده است که مقدار انحراف معیار نسبی (RSD) به روش تامپسون و هوارث برای کلیه عناصر کمتر از ۱۰ درصد است.



شکل ۳- گرادیان تپه‌نوردی با استفاده از مدل DENCLUE 1.0 (شکل بالا) و اندازه گام‌ها مطابق با فرآیند تپه‌نوردی در مدل DENCLUE 2.0 (شکل پایین) [۲۶].



شکل ۴- موقعیت نمونه‌های ژئوشیمیایی (دایره‌های سیاه) برداشت شده از محدوده مورد مطالعه به همراه شماره نمونه‌ها.

آهن و منیزیم در بخش میانی محدوده مورد مطالعه کمتر از بخش شرقی و غربی آن است، در حالی که توزیع سدیم و پتاسیم نسبت معکوس با توزیع آهن و منیزیم دارند و تمرکز در بخش میانی محدوده دارند. برای شناسایی محدوده مستعد فاز بعدی اکتشافی، از نتایج تجزیه ۷ عنصر مرتبط با کانی‌زایی مس- طلا استفاده شده است که دارای آنومالی شاخص‌اند. جدول ۱ پارامترهای آماری این عناصر را نشان می‌دهد.

پردازش اولیه آماری داده‌های ژئوشیمیایی پس از جایگزینی داده‌های سنسورد به روش درست‌نمایی کوهن و تعدیل داده‌های خارج از ردیف، نشان می‌دهد که عناصری همچون Mg , Ni , Ti , Ba , K و Al دارای توزیع نرمال‌اند و در مقابل عناصر P , Na , Sr و Mn , Au , S , Pb و Zn و As , Hg , Cu جامعه دو مدی از خود نشان می‌دهند. همچنین عناصر W و Sb , Cd , Bi , Ag به دلیل تعداد بالای داده‌های سنسورد توزیع خاصی را ارایه نمی‌دهند. توزیع مقادیر

جدول ۱- پارامترهای آمار توصیفی عناصر انتخابی در محدوده مورد مطالعه.

متغیر	میانگین	انحراف معیار	حداقل	میانه	حداکثر	چولگی	کشیدگی
As (ppm)	۱۴,۸۸	۷,۰۲	۵	۱۵	۴۳	۱,۷	۳,۹
Au (ppb)	۶۲,۵۴	۴۸,۳۲	۲۸	۴۸	۴۰۸	۴,۱	۲۲,۷
Cu (ppm)	۲۳,۹۰	۴۱,۴۸	۱	۸	۲۸۵	۴,۰	۱۹,۲
Hg (ppb)	۴۵,۳۶	۵۰,۹۸	۱۰	۳۰	۳۴۰	۴,۰	۱۷,۲
Pb (ppm)	۳۳,۶۷	۱۲,۸۸	۱۴	۳۰	۸۱	۱,۳	۱,۸
S (%)	۰,۰۵۸	۰,۰۲۵	۰,۰۲	۰,۰۶	۰,۱۴	۱,۰	۱,۳
Zn (ppm)	۹۸,۳۵	۲۲,۳۶	۶۴	۹۳	۱۶۲	۰,۹	۰,۲

نسبت به عیار و به صورت رابطه ۲۷ است:

$$A(\geq \rho) = C\rho_i^{-D} \quad (27)$$

که در آن:

$A(\geq \rho)$ مساحت.

ρ_i عیار.

در شکل ۵ نقشه کنتوری عناصر انتخابی در محدوده مورد مطالعه با دو مقیاس پیوسته و کلاسه‌بندی شده نشان داده شده است. کلاسه‌بندی داده‌ها به روش فرکتالی عیار- مساحت انجام گرفته است. برای استفاده از این روش باید نقشه کنتوری عیار را ترسیم کرد. اگر A مساحت مقادیر عیار بزرگتر از کنتور ρ در یک نقشه کنتوری باشد، این مساحت دارای یک تابع نمایی کاهش

D: بعد فرکتال.

C: ضریب ثابت.

جیوه خواهد بود.

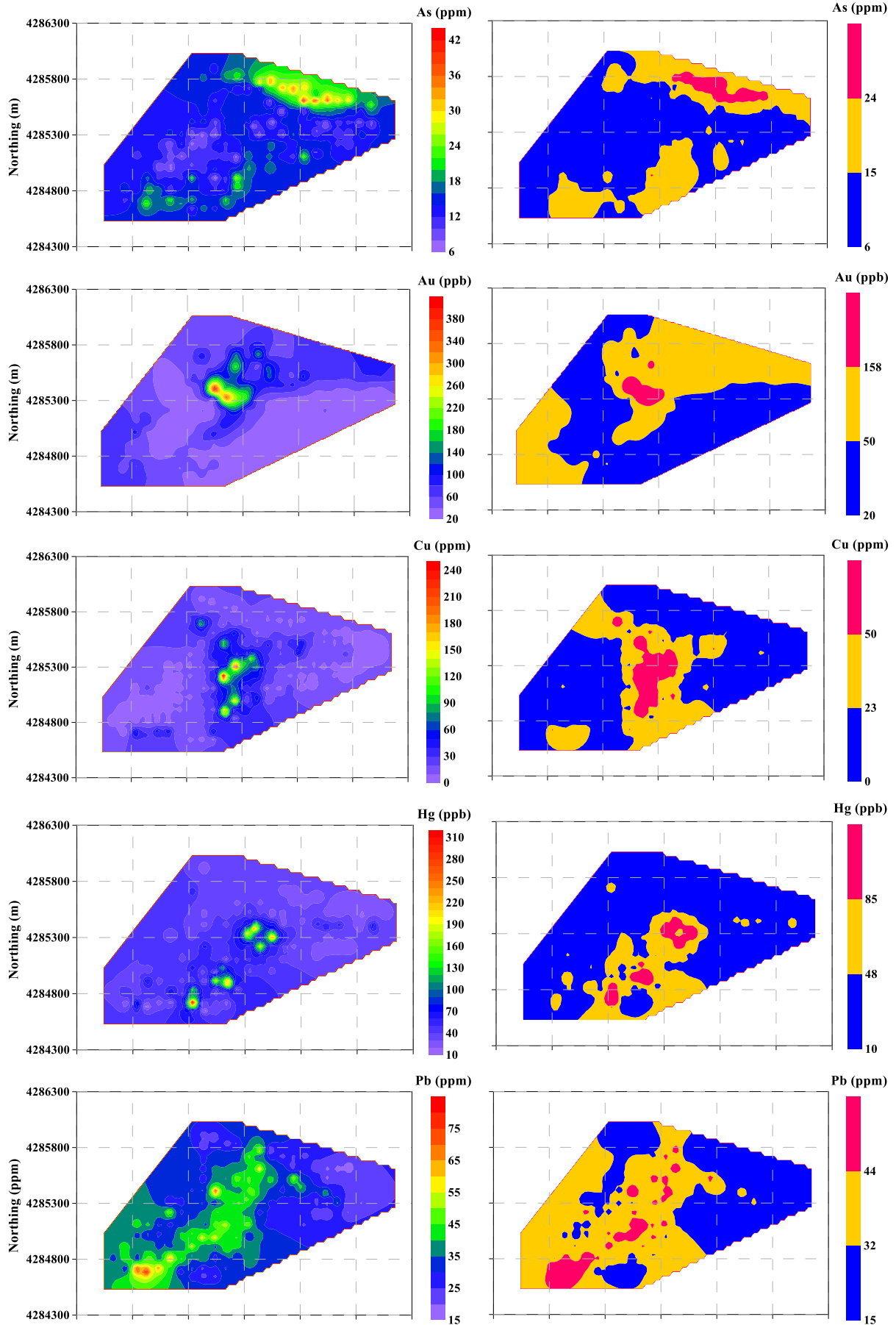
۵- تحلیل داده‌ها

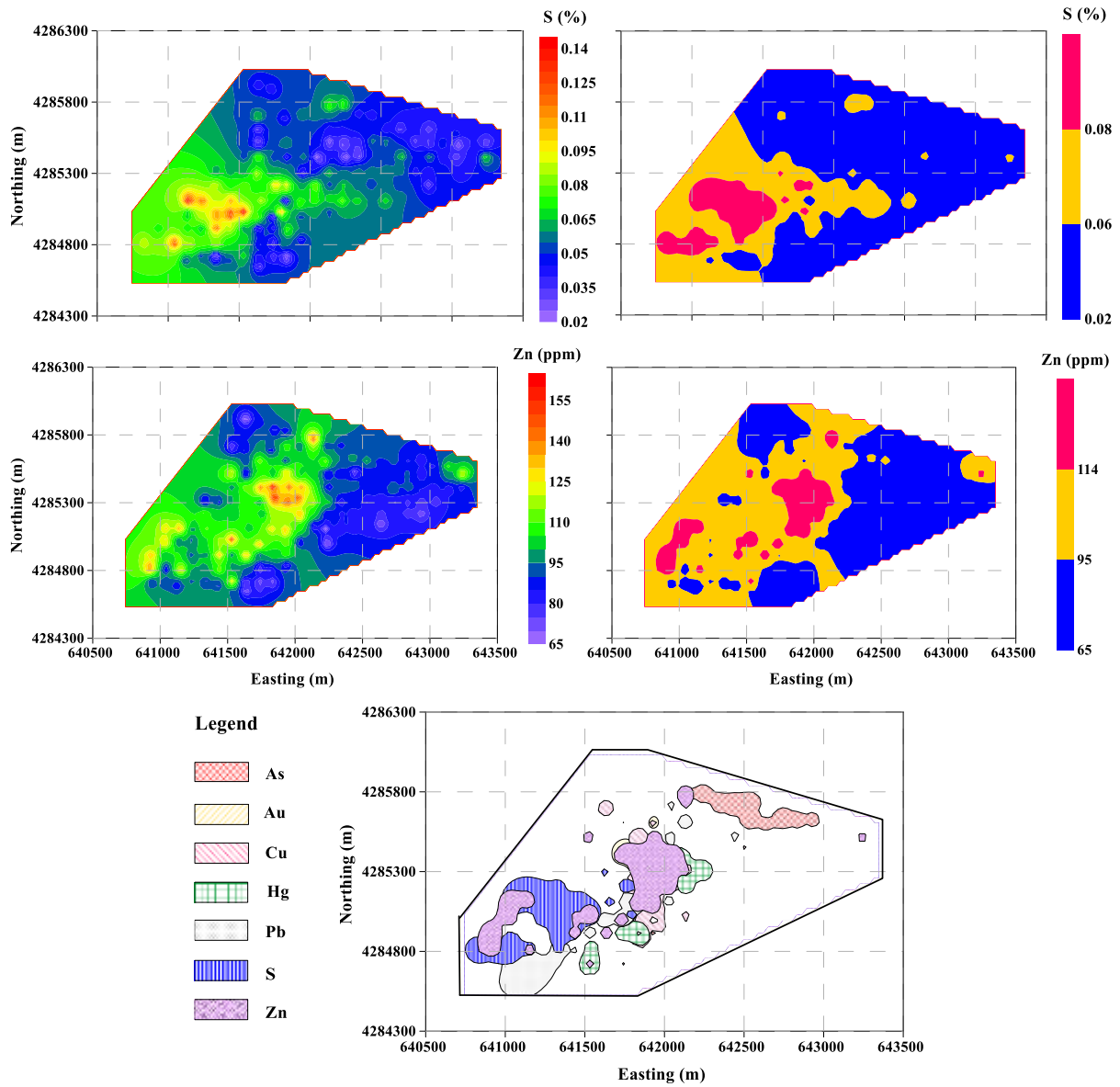
در ابتدا، ماتریس داده‌های اولیه با ابعاد 165×7 با روش تحلیل تطبیقی مورد پردازش قرار گرفت که ماتریس نهایی بدست آمده دارای ابعاد 172×6 است. برای این منظور از برنامه‌نویسی در نرم‌افزار متلب استفاده شده است. بنابراین به روش تحلیل تطبیقی فاکتوری، می‌توان ۱۵ نمودار پراکندگی رسم کرد تا تغییرات فاکتورهای نمونه‌ها و متغیرها به طور همزمان به نمایش گذاشته شود. در شکل ۷ چهار عدد از این نمودارها نشان داده شده است.

مطابق شکل ۷-الف عناصر طلا و جیوه خود را از سایر نمونه‌ها و متغیرها جدا کرده‌اند. همچنین نمونه‌های ۳۶، ۵۰، ۱۰۷، ۱۵۴ و ۱۵۵ حاوی مقادیر آنومال سرب است.

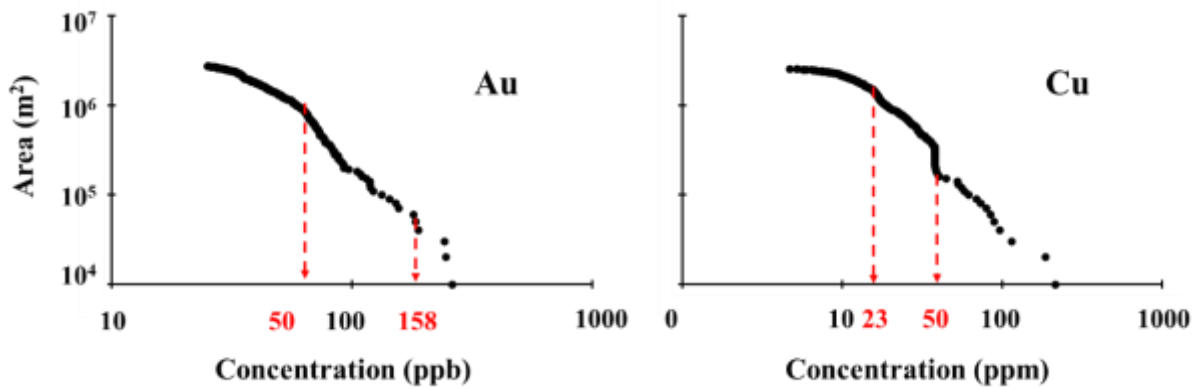
رابطه ۲۷ نشان می‌دهد که اگر مقادیر مساحت بر حسب عیار در یک نمودار تمام لگاریتمی ترسیم شود، نقاط شکست منحنی را می‌توان به عنوان آستانه تفکیک جوامع آماری در نظر گرفت [۲۰ و ۵۷]. برای نمونه در شکل ۶ نمودار فرکتالی دو عنصر طلا و مس نشان داده شده است.

مطابق شکل ۵، آنومالی عنصر آرسنیک در بخش شمال‌شرقی، طلا و مس در بخش میانی، جیوه و سرب در بخش میانی و جنوبی، گوگرد در بخش جنوب‌غربی و روی در بخش میانی و جنوب‌غربی منطقه قرار دارد. در جدول ۲ مساحت هر کدام از این آنومالی‌ها و زمینه‌ها به تفکیک عنوان شده است. اگر محدوده نشان داده شده به رنگ قرمز را به عنوان محدوده آنومالی هر عنصر در نظر گرفته شود؛ بزرگترین آنومالی متعلق به عنصر گوگرد و سپس روی و کوچکترین آنومالی متعلق به طلا و سپس





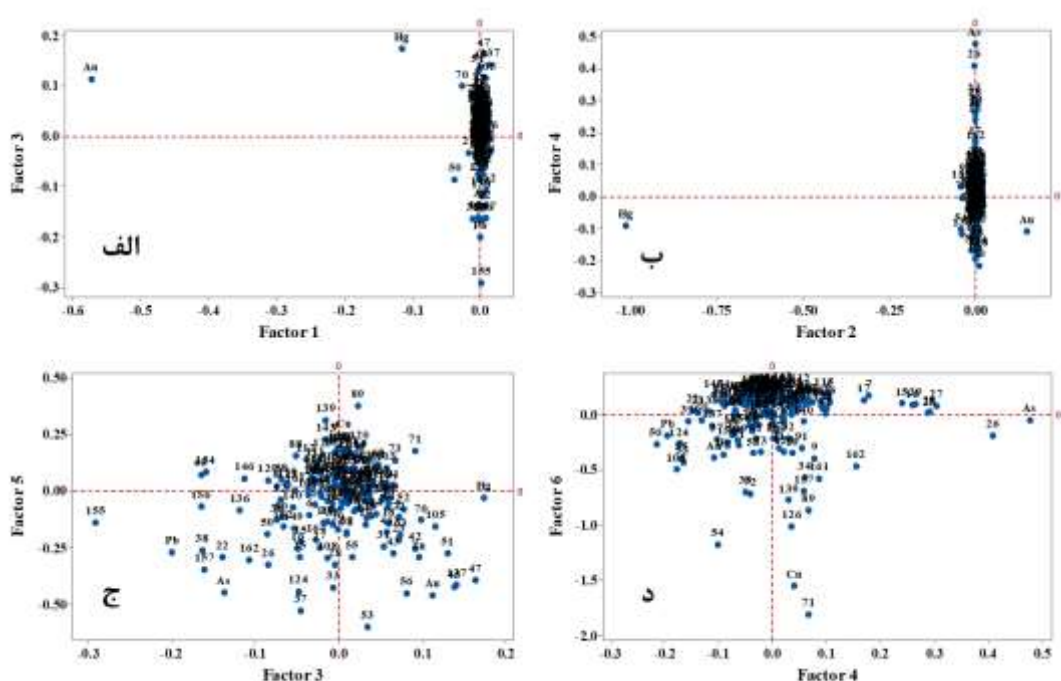
شکل ۵- نقشه کنتوری عناصر انتخابی با مقیاس پیوسته (نقشه‌های سمت چپ) و کلاسه‌بندی شده (نقشه‌های سمت راست) در محدوده مورد مطالعه به همراه نقشه آنومالی‌های ژئوشیمیایی تک عنصری (آخرین نقشه).



شکل ۶- نمودار فرکتالی عیار- مساحت برای عناصر طلا و مس.

جدول ۲- مساحت جوامع آماری کلاسه‌بندی شده برای هر کدام از عناصر انتخابی

عناصر	مساحت جوامع آماری کلاسه‌بندی شده (به هزار متر مربع)		
As	۱۷۰۸	۸۱۵	۱۰۲
Au	۱۳۰۵	۱۲۵۸	۶۲
Cu	۱۷۶۸	۶۸۷	۱۷۰
Hg	۲۰۱۹	۵۰۷	۹۹
Pb	۱۱۳۸	۱۳۰۴	۱۸۳
S	۱۶۵۹	۷۳۱	۲۳۵
Zn	۱۱۸۱	۱۱۹۴	۲۵۰



شکل ۷- نمودار تحلیل تطبیقی نمونه‌ها و متغیرها در دستگاه مختصات فاکتوری.

نمودار پراکندگی خود را نشان می‌دهند. نمونه‌های ۹، ۸۹، ۱۳۹ و ۱۴۳ نیز دارای مقادیر آنومال مس‌اند. در شکل ۷-د مقادیر آنومال نمونه‌های ۵۰، ۵۳، ۱۰۷ و ۱۲۴ برای سرب، نمونه‌های ۹، ۲۱، ۳۴، ۸۹، ۱۲۶، ۱۳۹، ۱۵۷، ۱۶۱ و ۱۶۲ برای مس، نمونه‌های ۷، ۱۵، ۱۷، ۲۶، ۲۷، ۲۸ و ۲۹ برای آرسنیک و نمونه‌های ۳۷، ۵۱ و ۱۵۷ برای طلا و جیوه مشاهده می‌شود.

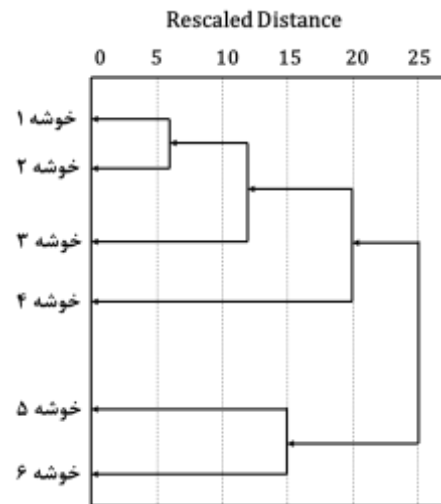
برای تفکیک کلیه نمونه‌ها و متغیرها باید از روش تحلیل تطبیقی خوشه‌ای استفاده کرد. برای این منظور، از رابطه ۱۵

در شکل ۷-ب علاوه بر جدایش شاخص عناصر جیوه و طلا از سایر نقاط، مقادیر آنومال آرسنیک برای نمونه‌های ۱۵، ۲۶ و ۲۷ قابل مشاهده است. سایر نمونه‌ها و متغیرها در این دو نمودار در اطراف مبدا مختصات پراکنده بوده و جدایشی را نشان نمی‌دهند. در حالی که در شکل ۷-ج پراکندگی وسیعتری از نقاط دیده می‌شود. در این شکل عناصر طلا، روی و جیوه به همراه نمونه‌های ۴۲، ۴۳، ۴۵، ۴۷، ۵۱، ۵۶، ۷۰، ۱۰۵ و ۱۳۷ در یک محدوده قرار دارند. همچنین عناصر سرب و آرسنیک به همراه نمونه‌های ۲۲، ۲۶، ۲۸، ۳۳، ۳۷، ۳۸، ۵۰، ۱۰۸، ۱۲۴، ۱۳۶، ۱۴۰، ۱۴۹، ۱۵۵، ۱۵۶ و ۱۵۷ در محدوده دیگری از

مونزونیت، آندزیت و دایک‌های دیوریتی است. بنابراین این محدوده می‌تواند به عنوان مستعدترین بخش منطقه مورد مطالعه برای کانی‌زایی پیشنهاد شود. محدوده D به عنوان بزرگترین محدوده بدست آمده، بیشتر برای عنصر S (عنصر خوشه ۱) دارای نمونه‌های آنومال است. البته در بخش جنوب و جنوب‌شرقی تا شمال‌شرقی این محدوده نمونه‌های آنومال عناصر Hg, Ag, Cu و Zn (عناصر خوشه ۴ و ۶) نیز وجود دارد. انطباق موقعیت این محدوده با آنومالی‌های ژئوشیمیایی همان عناصر در شکل ۵ و وجود سنگ بستر مونزونیتی میزبان دایک‌های دیوریتی، مستعد بودن این محدوده را برای کانی‌زایی به ویژه در نوار بخش جنوب‌شرقی تا شمال‌شرقی محدوده نشان می‌دهد.

در مرحله بعدی ماتریس داده‌های 172×6 بعدی با الگوریتم معرفی شده در بخش ۳-۲ خوشه‌بندی می‌شود. برای این منظور از کدهای نوشته شده برای الگوریتم DENCLUE 2.0 به زبان Python استفاده شده است. مقدار بهینه پارامترهای الگوریتم نیز به ترتیب $\sigma = 0.1$ و $\xi = 3.36$ بدست آمده است. جدول ۴ نتایج خوشه‌بندی داده‌ها به وسیله این الگوریتم را نشان می‌دهد. متغیر S به همراه ۶۷ نمونه در خوشه اول، ۴۲ نمونه در خوشه دوم، ۳۸ نمونه به همراه Pb و Zn در خوشه سوم، ۱۰ نمونه به همراه Au و Cu در خوشه چهارم و متغیرهای As و Hg به همراه ۸ نمونه در خوشه پنجم قرار گرفته‌اند. شکل ۱۰ موقعیت این خوشه‌ها در منطقه مورد مطالعه را نشان می‌دهد. در این شکل نیز همانند نتایج روش CCA چهار محدوده مستعد کانی‌زایی قابل مشاهده است. موقعیت و وسعت محدوده A در شکل‌های ۹ و ۱۰ تقریباً یکسان است. این محدوده بر روی سنگ‌های کوارتز مونزونیتی و در مجاورت توده آندزیتی-بازالتی قرار دارد. بخشی از غنی‌شدگی نمونه در بخش شمال‌غربی این محدوده به عنصر آرسنیک را می‌توان به سنگ‌های آندزیتی-بازالتی و غنی‌شدگی نمونه‌ها به عناصر سرب و روی در بخش جنوب‌شرقی را به دایک‌های هورنبلند دیوریت پورفیری نسبت داد. از آنجا که عناصر این محدوده (As, Hg, Pb و Zn) بیشتر در فازهای نهایی ماگماتیسم، در درجه حرارت پایین و بصورت کمپلکس‌های بی‌سولفیدی و کلروری در فاصله نسبتاً دوری از منشا ماگمای اولیه تشکیل می‌شوند و با توجه به سنگ بستر آن، کانی‌زایی در این محدوده را می‌توان به نوع کانی‌سازی‌های هیدروترمال رگه‌ای نسبت داد [۳۲ و ۴۱].

به عنوان معیار شباهت و روش اتصال وارد^۱ برای اتصال سلسله مراتبی خوشه‌ها استفاده شد. شکل ۸ نمودار دندروگرام ساده شده تحلیل تطبیقی خوشه‌ای را نشان می‌دهد و در جدول ۳ عناصر هر خوشه شامل متغیرها و نمونه‌ها آورده شده است. برای تحلیل نتایج بدست آمده و امکان مقایسه آن با نقشه‌های کنتوری شکل ۵، موقعیت نمونه‌های هر خوشه در شکل ۹ به نمایش گذاشته شده است. در این شکل چهار محدوده مستعد کانی‌زایی (با نام‌های A, B, C و D) معرفی شده است. محدوده A حاوی نمونه‌های آنومال عناصر خوشه‌های ۵ و ۶ یعنی As, Au, Hg, Zn است. مقایسه موقعیت این محدوده با آنومالی ژئوشیمیایی در شکل ۵، نشان‌دهنده انطباق خوب این محدوده با آنومالی‌های عناصر As و Au و تا حدودی Zn است. همچنین قرارگیری این محدوده بر روی سنگ‌های آندزیتی حاوی دایک‌های دیوریتی مستعد بودن آن را برای کانی‌زایی نشان می‌دهد.

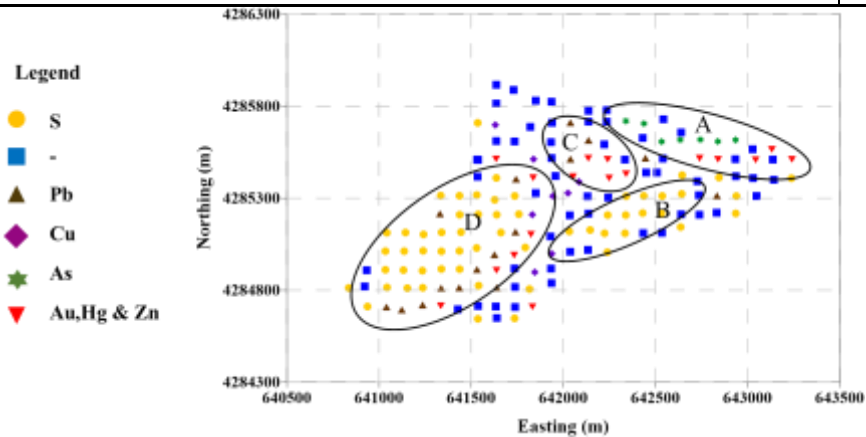


شکل ۸- دندروگرام ساده شده تحلیل تطبیقی خوشه‌ای داده‌های ماتریس F

نمونه‌های قرار گرفته در محدوده B بیشتر دارای آنومال عنصر گوگردند (خوشه اول). نقشه کنتوری این عنصر در شکل ۵ هیچگونه آنومالی ژئوشیمیایی را در این محدوده نشان نمی‌دهد. البته با توجه به سنگ میزبان این محدوده و وجود آلتراسیون‌های هیدروترمال در کل منطقه مورد مطالعه، مستعد بودن کانی‌زایی در این محدوده نیاز به بررسی‌های دقیق‌تر صحرائی خواهد داشت. محدوده C حاوی نمونه‌های آنومال عناصر Au, Pb, Hg و Zn است. موقعیت این محدوده با آنومالی‌های ژئوشیمیایی این عناصر در شکل ۵، انطباق خوبی را نشان می‌دهد. همچنین جنس سنگ‌های این محدوده نیز

جدول ۳- خوشه‌بندی نمونه‌ها و متغیرهای به روش CCA

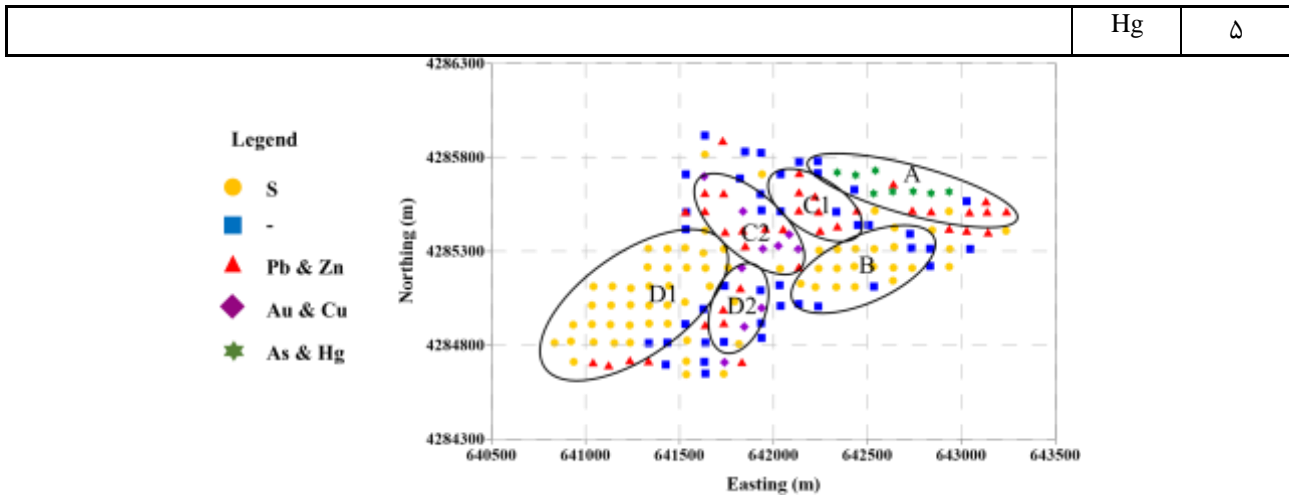
عناصر هر خوشه		متغیرها	عنوان خوشه
شماره‌ی نمونه‌ها			
۸۱، ۸۲، ۸۳، ۸۴، ۸۵، ۸۶، ۸۷، ۸۸، ۹۲، ۹۳، ۹۴، ۹۵، ۹۹، ۱۰۰، ۱۰۱، ۱۰۲، ۱۰۳، ۱۰۴، ۱۰۶، ۱۱۰، ۱۱۱، ۱۱۲، ۱۱۳، ۱۱۴، ۱۱۷، ۱۱۸، ۱۱۹، ۱۲۰، ۱۲۱، ۱۲۲، ۱۲۵، ۱۲۹، ۱۳۱، ۱۳۲، ۱۳۳، ۱۳۴، ۱۳۵، ۱۴۱، ۱۴۳، ۱۴۴، ۱۴۵، ۱۴۸، ۱۵۱، ۱۵۳، ۱۶۳ و ۱۶۵		S	خوشه ۱
۱، ۲، ۳، ۴، ۵، ۶، ۷، ۱۰، ۱۱، ۱۲، ۱۳، ۱۴، ۱۷، ۱۸، ۱۹، ۲۰، ۲۱، ۲۳، ۲۴، ۳۰، ۳۲، ۳۵، ۳۹، ۴۱، ۴۴، ۴۶، ۴۸، ۵۲، ۵۷، ۵۸، ۶۰، ۶۲، ۶۳، ۶۴، ۷۰، ۷۳، ۷۴، ۷۹، ۸۲، ۹۰، ۹۱، ۹۶، ۹۷، ۹۸، ۱۰۵، ۱۰۹، ۱۱۴، ۱۱۵، ۱۲۷، ۱۲۸، ۱۳۰، ۱۳۸، ۱۴۰، ۱۴۲، ۱۵۰، ۱۵۲، ۱۵۸، ۱۵۹، ۱۶۰، ۱۶۱ و ۱۶۴		-	خوشه ۲
۱۲، ۲۲، ۳۶، ۴۰، ۵۰، ۸۰، ۸۴، ۱۰۷، ۱۲۳، ۱۳۶، ۱۴۶، ۱۴۷، ۱۴۹، ۱۵۴، ۱۵۵ و ۱۵۶		Pb	خوشه ۳
۹، ۳۴، ۵۴، ۷۱، ۷۲، ۸۹، ۱۲۶ و ۱۳۹		Cu	خوشه ۴
۱۵، ۱۶، ۲۵، ۲۶، ۲۷، ۲۸ و ۲۹		As	خوشه ۵
۳۱، ۳۳، ۳۷، ۳۸، ۴۲، ۴۳، ۴۵، ۴۶، ۵۱، ۵۳، ۵۵، ۵۶، ۱۰۸، ۱۲۴، ۱۳۷، ۱۵۷ و ۱۶۲		Hg, Au و Zn	خوشه ۶



شکل ۹- موقعیت نمونه‌های هر خوشه بر اساس نوع متغیر به روش CCA.

جدول ۴- خوشه‌بندی نمونه‌ها و متغیرهای به روش DENCLUE

عناصر هر خوشه		متغیرها	عنوان خوشه
شماره‌ی نمونه‌ها			
۳، ۱۱، ۴۱، ۴۴، ۴۹، ۵۹، ۶۱، ۶۵، ۶۶، ۶۷، ۶۸، ۶۹، ۷۴، ۷۵، ۷۶، ۷۷، ۷۸، ۸۱، ۸۳، ۸۴، ۸۵، ۸۶، ۸۷، ۸۸، ۹۰، ۹۲، ۹۳، ۹۴، ۹۵، ۹۶، ۹۷، ۹۹، ۱۰۰، ۱۰۱، ۱۰۲، ۱۰۳، ۱۰۴، ۱۰۶، ۱۱۱، ۱۱۲، ۱۱۳، ۱۱۴، ۱۱۶، ۱۱۷، ۱۱۸، ۱۱۹، ۱۲۰، ۱۲۱، ۱۲۲، ۱۲۳، ۱۲۵، ۱۲۶، ۱۳۱، ۱۳۲، ۱۳۳، ۱۳۴، ۱۳۵، ۱۴۱، ۱۴۲، ۱۴۳، ۱۴۴، ۱۴۵، ۱۵۱، ۱۵۳، ۱۵۹، ۱۶۳ و ۱۶۵		S	خوشه ۱
۱، ۲، ۳، ۴، ۵، ۶، ۷، ۸، ۱۰، ۱۲، ۱۴، ۲۱، ۲۴، ۳۰، ۳۲، ۳۵، ۳۶، ۳۹، ۴۸، ۵۷، ۵۸، ۶۰، ۷۹، ۸۰، ۸۲، ۹۸، ۱۰۷، ۱۰۹، ۱۱۰، ۱۱۱، ۱۱۵، ۱۲۳، ۱۲۸، ۱۲۹، ۱۳۶، ۱۴۰، ۱۴۶، ۱۴۷، ۱۴۹، ۱۵۰، ۱۵۲، ۱۵۸، ۱۶۰ و ۱۶۴		-	خوشه ۲
۲، ۱۳، ۱۸، ۱۹، ۲۰، ۲۲، ۲۳، ۳۱، ۳۳، ۳۷، ۳۸، ۴۰، ۴۲، ۴۳، ۴۵، ۴۶، ۴۷، ۵۰، ۵۱، ۵۲، ۵۳، ۵۵، ۵۶، ۶۲، ۶۳، ۶۴، ۷۰، ۹۱، ۱۰۵، ۱۰۸، ۱۲۴، ۱۳۷، ۱۳۸، ۱۵۴، ۱۵۵، ۱۵۶، ۱۵۷ و ۱۶۲		Pb و Zn	خوشه ۳
۹، ۳۴، ۵۴، ۷۱، ۷۲، ۷۳، ۸۹، ۱۲۶ و ۱۳۹		Au و Cu	خوشه ۴
۱۵، ۱۶، ۱۷، ۲۵، ۲۶، ۲۷، ۲۸ و ۲۹		As و	خوشه ۵



شکل ۱۰- موقعیت نمونه‌های هر خوشه بر اساس نوع متغیر به روش DENCLUE.

کانی‌زایی رگه‌ای را در این بخش تقویت می‌کند [۶ و ۴۲]. همچنین در قسمت جنوبی این بخش در مجاورت توده هورنبلند دیوریت پورفیری، نمونه‌های حاوی مقادیر آنومال سرب و روی نیز مشاهده می‌شود. بنابراین احتمال کانی‌زایی در این بخش همانند محدوده B نیاز به مطالعات تکمیلی خواهد داشت.

بنابراین، با توجه به نتایج بدست آمده از خوشه‌بندی داده‌ها به روش CCA (شکل ۹) و روش DENCLUE (شکل ۱۰)، می‌توان خوشه‌بندی بهتر عناصر (قرار دادن عناصر هم‌پارازنز در یک خوشه)، قرار دادن دو عنصر کانی‌زایی منطقه مورد مطالعه یعنی طلا و مس در یک خوشه، انطباق بهتر محدوده‌های پیشنهادی برای کانی‌زایی با آنومالی‌های ژئوشیمیایی عناصر و شرایط زمین‌شناسی منطقه مورد مطالعه و امکان پیش‌بینی بهتر نوع کانی‌زایی محدوده‌ها را به عنوان برتری روش DENCLUE نسبت به روش CCA عنوان کرد.

۶- نتیجه‌گیری

شناسایی الگوهای پراکندگی عناصر و مرتبط کردن آن‌ها با آنومالی‌های ژئوشیمیایی یکی از ابزارهای اکتشافی به ویژه در فازهای نیمه‌تفصیلی و تفصیلی محسوب می‌شود. در این مقاله از دو روش تحلیل تطبیقی خوشه‌ای و الگوریتم خوشه‌بندی بر اساس توابع توزیع چگالی استفاده شد که می‌توانند نمونه‌ها و عناصر شیمیایی را به طور همزمان مورد تحلیل قرار دهند. منطقه اکتشافی خوینرود به دلیل بیرون‌زدگی توده‌های نفوذی اسیدی تا متوسط در داخل سنگ ولکانیکی و فلیشی مستعد کانی‌زایی‌های هیدروترمال و پورفیری بوده و با توجه به انجام فازهای اکتشافی مقدماتی، نیمه‌تفصیلی و تفصیلی برای شناسایی الگوهای ژئوشیمیایی مناسب است. نتایج خوشه‌بندی داده‌های ژئوشیمیایی این منطقه به دو روش یاد شده، نشان داد که:

محدوده کانی‌زایی B بر روی توده نفوذی کوارتزمونزونیتی قرار دارد که با دایک‌های از جنس دیوریت پورفیری قطع شده است. همانطور که قبلاً نیز اشاره شد، عدم تشکیل آنومالی ژئوشیمیایی و قرارگیری فقط عنصر گوگرد در خوشه‌بندی نمونه‌های این محدوده، پیش‌بینی مستعد بودن کانی‌زایی اقتصادی در آن را به مطالعات اکتشافی تکمیلی وابسته می‌کند. محدوده کانی‌زایی C در شکل ۱۰ به دو بخش تقسیم شده است، بخش C1 که نمونه‌های آن در خوشه ۳ همراه با مقادیر آنومال عناصر Zn و Pb و بخش C2 که نمونه‌های آن در خوشه ۳ و ۵ همراه با مقادیر آنومال Au, Zn, Pb و Cu قرار گرفته‌اند. سنگ بستر این محدوده توده نفوذی کوارتزمونزونیتی خوینرود به همراه هورنبلند دیوریت پورفیری است. آنومالی‌های ژئوشیمیایی (شکل ۵) نیز با این محدوده همپوشانی دارند. همچنین سنگ‌های این محدوده دارای آلتراسیون‌های پروپلیتیک، فلیک و آرژیلیک‌اند. بنابراین با توجه به نکات یاد شده و نبود آلتراسیون پتاسیک در سطح، کانی‌زایی در این محدوده را می‌توان از نوع پورفیری و در عمق در نظر گرفت [۶، ۲۴ و ۳۲]. بدیهی است، کانی‌زایی پورفیری در محدوده C2 از احتمال بالاتری برخوردار است (فاز اکتشاف تفصیلی توسط شرکت اسپیر در این بخش انجام گرفته است).

در شکل ۱۰، محدوده کانی‌زایی D نیز به دو بخش تقسیم شده است. نمونه‌های بخش D1 بیشتر همراه با مقادیر آنومال S و بخش D2 حاوی مقادیر آنومال Au, Zn, Pb و Cu است. شرایط زمین‌شناسی بخش D2 همانند C2 بوده، بنابراین مستعد کانی‌زایی پورفیری‌اند. بخش D1 را سنگ‌های کوارتزمونزونیتی پوشش می‌دهند که در مجاورت توده دیوریتی پورفیری قرار دارد. وجود دایک دیوریتی همراه با رگه‌های سیلیسی احتمال

قدردانی می‌شود. همچنین از شرکت جیوه خویناوی به دلیل استفاده از گزارش محدوده مینرالیزه خویناوی تشکر می‌شود.

منابع

۱. آقاباتی، سیدعلی؛ ۱۳۸۳؛ «زمین‌شناسی ایران»، تهران، سازمان زمین‌شناسی کشور.
۲. برنا، بهروز؛ ۱۳۹۱؛ «اکتشاف ژئوشیمیایی سیستماتیک در ورقه یکصد هزار ورزقان»، سازمان زمین‌شناسی و اکتشافات معدنی کشور، طرح زمین‌شناسی عمومی (پروژه ژئوشیمیایی).
۳. دارابی گلستان، فرشاد؛ قوامی ربابی، رضا؛ مجلسی، محمد جواد؛ معمارزاده، مجتبی؛ اسدی هارونی، هوشنگ؛ ۱۳۹۱؛ «شناسایی و تفکیک متغیرهای آنومال با روش‌های آنالیز تطبیقی و تفریقی در منطقه دالی شمالی»، نشریه علمی- پژوهشی روشهای تحلیلی و عددی در مهندسی معدن، شماره ۳، صفحه ۳۵ تا ۴۳.
۴. دارابی گلستان، فرشاد؛ هزارخانی، اردشیر؛ سیف‌پناهی شعبانی، کیومرث؛ ۱۳۹۴؛ «به کارگیری آنالیز تطبیقی و هندسه فراکتال جهت شناسایی عناصر ناهنجار در خاک و تعیین محدوده آن در ورقه ۱:۱۰۰۰۰۰ خوسف»، نشریه علمی- ترویجی یافته‌های نوین زمین‌شناسی کاربردی، دوره ۹، شماره ۱۸، صفحه ۹ تا ۱۷.
۵. سهرابی، قهرمان؛ ۱۳۹۴؛ «بررسی متالورژی و ژئوشیمی ذخایر مولیبدن در نوار قره داغ- شیورداغ، آذربایجان شرقی، شمال غرب ایران»، رساله دکتری، دانشگاه تبریز.
۶. شرکت اسپیر؛ ۱۳۸۵؛ «گزارش اکتشاف محدوده مینرالیزه خوینه‌رود (شمال شهرستان ورزقان، استان آذربایجان شرقی) با هدف اکتشاف فلزات پایه و گران‌بها»، گزارش فاز اکتشافات ۱:۵۰۰۰، سازمان زمین‌شناسی و اکتشافات معدنی کشور.
۷. عباس‌زاده، سمیه؛ رحیمی‌پور، غلامرضا؛ نجم‌الدینی، مجید؛ ۱۳۹۲؛ «شناسایی مناطق کانی‌زایی مس پورفیری با استفاده از تلفیق روش‌های تک متغیره و چند متغیره بر روی داده‌های ژئوشیمیایی آبراهه‌ای در منطقه قلعه عسکر، استان کرمان»، نشریه علمی- پژوهشی روش‌های تحلیلی و عددی در مهندسی معدن،

- چهار محدوده کانی‌زایی A، B، C و D در منطقه مورد مطالعه قابل شناسایی است.
 - با توجه به جنس سنگ‌ها (توده نفودی کوارتزومونونیت و سنگ‌های هورنبلند دیوریت پورفیری)، نوع آنومالی‌های ژئوشیمیایی (آنومالی‌های Au، Pb، Zn و Cu) و مطالعات زمین‌شناسی (وجود آلتراسیون‌های پروپلیتیک، فلیک و آرژیلیک)، بخش C2 از محدوده C و بخش D2 از محدوده D به عنوان مستعدترین محدوده‌ها با احتمال کانی‌زایی مس- طلا پورفیری پیشنهاد می‌شوند.
 - محدوده A به دلیل قرارگیری بر روی سنگ‌های کوارتزومونونیتی مجاور توده آندزیتی- بازالتی و آنومالی‌های ژئوشیمیایی As، Hg، Pb و Zn با احتمال کانی‌زایی رگه‌ای هیدروترمال پیشنهاد بعدی است.
 - محدوده B، بخش C1 از محدوده C و بخش D1 از محدوده D که بیشتر بروی سنگ‌های دیوریت پورفیری قرار دارند نیز با احتمال کانی‌زایی رگه‌ای برای اکتشاف تکمیلی از پیشنهادات دیگر این مقاله است.
- مقایسه نتایج خوشه‌بندی به روش CCA با روش DENCLUE نشان می‌دهد که روش دوم به دلیل خوشه‌بندی بهتر عناصر، قرار دادن دو عنصر اصلی کانی‌زایی در منطقه مورد مطالعه یعنی طلا و مس در یک خوشه، انطباق بهتر محدوده‌های پیشنهادی برای کانی‌زایی با آنومالی‌های ژئوشیمیایی عناصر و شرایط زمین‌شناسی منطقه مورد مطالعه و امکان پیش‌بینی بهتر نوع کانی‌زایی محدوده‌ها نسبت به روش اول برتری دارد. این نکته در ساختار این الگوریتم یعنی خوشه‌بندی داده‌ها با شکل‌های دلخواه و به کارگیری چگالی تابع توزیع داده‌ها نهفته است. بنابراین استفاده از الگوریتم‌های که نمونه‌ها و متغیرها را به صورت هم‌زمان خوشه‌بندی می‌کنند برای ارتباط دادن الگوهای پراکندگی عناصر با آنومالی‌های ژئوشیمیایی آنها پیشنهاد می‌شود. همچنین به عنوان انتخاب اول الگوریتم DENCLUE به دلیل پایه ریاضی قوی، عملکرد خیلی خوب بر روی مجموعه داده‌ی با نویز یا داده پرت بالا (مقاوم بودن)، سرعت عمل بالاتر در خوشه‌بندی داده‌ها به ویژه داده‌های با ابعاد بالا، پیچیدگی محاسباتی کمتر و خوشه‌بندی مستقیم داده‌ها معرفی می‌شود.

تشکر و قدردانی

از شرکت اسپیر به دلیل در اختیار قرار دادن داده‌های اکتشافی منطقه مورد مطالعه (نقشه زمین‌شناسی، داده‌های ژئوشیمیایی و گزارش‌های فازهای مختلف اکتشافی) تشکر و

- شماره ۶، صفحه ۶۹ تا ۸۲.
۸. گرائیان، حمید؛ ۱۳۹۶؛ «کاربرد روش‌های خوشه‌بندی در شناسایی آنومالی‌های مرکب در اکتشافات ژئوشیمیایی ناحیه‌ای در کردگان خراسان جنوبی»، نشریه علمی- پژوهشی مهندسی معدن، دوره دوازدهم، شماره ۳۷، صفحه ۸۱ تا ۹۴.
۹. محمدزاده، محمد جعفر؛ ناصری، اینور؛ محمودیان، امید؛ ۱۳۸۸؛ «مقایسه روش‌های جداسازی جوامع سنگی و خوشه‌بندی فازی میان مرکز برای حذف مولفه سنژنتیک در اکتشافات ناحیه‌ای رسوبات آبراهه‌ای قره چمن- آذربایجان شرقی»، نشریه علمی- پژوهشی مهندسی معدن، دوره چهارم، شماره هشتم، صفحه ۵۱ تا ۵۸.
۱۰. معینی، حمید؛ محمد تراب، فرهاد؛ کیخانی حسین‌پور، مجید؛ ۱۳۹۴؛ «بررسی کاربرد نگاشت‌های خودسازمانده در خوشه‌بندی داده‌های آبراهه‌ای و مقایسه آن با دندوگرام اکتشافی داده‌های ترکیبی»، نشریه علمی- پژوهشی مهندسی معدن، دوره دهم، شماره ۲۷، صفحه ۹۵ تا ۱۰۷.
11. Benzécri, J.P., 1964; “*Cours de linguistique mathématique*”, Publication Mimeo, Faculté des sciences, Rennes, France.
12. Beh, E.J., Lombardo, R., 2014; “*Correspondence Analysis: Theory, Practice and New Strategies*”, John Wiley & Sons, 593 p.
13. Carranza, E.J.M., Zuo, R., 2017; “*Introduction to the thematic issue: analysis of exploration geochemical data for mapping of anomalies*”, *Geochemistry: Exploration, Environment, Analysis* 17(3), 183–185.
14. Chen, D., Wei, J., Wang, W., Shi, W., Li, H., Zhan, X., 2019; “*Comparison of methods for determining the thresholds of geochemical anomalies and the prospecting direction—A case of gold deposits in the Gouli exploration area, Qinghai Province*”, *Minerals* 9, 368.
15. Clare, A.P., Cohen, D.R., 2001; “*A comparison of unsupervised neural networks and k-means clustering in the analysis of multi-element stream sediment data*”, *Geochemistry: Exploration, Environment, Analysis* 1, 119–134.
16. Collyer, P.L., Merriam, D.F., 1973; “*An application of cluster analysis in mineral exploration*”, *Mathematical Geosciences* 5(3), 213–223.
17. Daya, A.A., Boomeri, M., Mazraee, N., 2017; “*Identification of geochemical anomalies by the use of concentration-area (C-A) fractal model in Nakhilab region, SE Iran*”, *International Journal of Mining and Mineral Engineering* 8(1), 70-81.
18. Ellefsen, K.J., Smith, D.B., 2016; “*Manual hierarchical clustering of regional geochemical data using a Bayesian finite mixture model*”, *Applied Geochemistry* 75, 200–210.
19. Ellefsen, K.J., Smith, D.B., Horton, J.D., 2014; “*A modified procedure for mixture-model clustering of regional geochemical data*”, *Applied Geochemistry* 51, 315-326.
20. Farahmandfar, Z., Jafari, M., Afzal, P., Ashja-Ardalan, A., 2020; “*Description of gold and copper anomalies using fractal and stepwise factor analysis according to stream sediments in Lahrud 1:100,000 sheet, NW Iran*”, *Geopersia* 10 (1), 135-148.
21. Fatehi, M., Asadi, H.H., 2017; “*Application of semi-supervised fuzzy c-means method in clustering multivariate geochemical data, a case study from the Dalli Cu-Au porphyry deposit in central Iran*”, *Ore Geology Reviews* 81, 245–255.
22. Gan, Li, D., 2003; “*Optimal Choice of Parameters for a Density-Based Clustering Algorithm*”, In: G. Wang et al. (Eds.): *RSFDGrC 2003, LNAI 2639*, pp. 603–606.
23. Geranian, H., Khajeh Miry, Z., 2020; “*Application of probabilistic clustering algorithms to determine mineralization areas in regional-scale exploration studies*”, *Journal of Mining & Environment* 11(4), 1059-1078.
24. Ghezelbash, R., Maghsoudi, A., Daviran, M., 2019; “*Prospectivity modeling of porphyry copper deposits: recognition of efficient mono- and multi-element geochemical signatures in the Varzaghan district, NW Iran*”, *Acta Geochimica* 38, 31–144.
25. Han, J., Kamber, M., Pei, J., 2012; “*Data Mining: Concepts and Techniques*”, Morgan Kaufmann, 740 p.
26. Hinneburg, A., Gabriel, H.H., 2007; “*DENCLUE 2.0: Fast clustering based on kernel density estimation*”, In *Advances in*

- Heidelberg.
36. Li, C., Sun, Z., Song, Y., 2003; “*DENCLUE-M: Boosting DENCLUE algorithm by mean approximation on grids*”, International Conference on Web-Age Information Management, WAIM 2003, pp 202-213.
 37. Mellinger, M., 1984; “*Correspondence analysis in the study of lithochemical data: general strategy and the usefulness of various data-coding schemes*”, Journal of Geochemical Exploration 21, 455–469.
 38. Obthong, N., Sriphum, W., 2011; “*Optimal Choice of Parameters for DENCLUE-based and Ant Colony Clustering*”, International Conference on Modeling, Simulation and Control, IPCSIT vol.10, IACSIT Press, Singapore, pp 69-73.
 39. Patinha, C., Correia, E., Ferreira da Silva, E., Simões, A., Reis, P., Morgado, F., Cardoso Fonseca, E., 2008; “*Definition of geochemical patterns on the soil of Paul de Arzila using correspondence analysis*”, Journal of Geochemical Exploration 98, 34–42.
 40. Piorecký, M., Štrobl, J., Krajča, V., 2019; “*Automatic EEG Classification Using Density Based Algorithms DBSCAN and DENCLUE*”, Acta Polytechnica 59(5), 498–509.
 41. Pirajno, F., 2009; “*Hydrothermal Processes and Mineral Systems*”, Springer Publication, Australia, 1273 p.
 42. Pirajno, F., 2012; “*Hydrothermal Mineral Deposits: Principles and Fundamental Concepts for the Exploration Geologist*”, Springer-Verlag, 688 p.
 43. Prabahari, R., Thiagarasu, D.V., 2014; “*A comparative analysis of density-based clustering techniques for outlier mining*”, International Journal of Engineering Sciences & Research Technology 3(11), 132-136.
 44. Prabahari, R., Thiagarasu, D.V., 2014; “*Density Based Clustering Using Gaussian Estimation Technique*”, International Journal on Recent and Innovation Trends in Computing and Communication 2(12), 4078-4081.
 45. Rantitsch, G., 2000; “*Application of fuzzy clusters to quantify lithological background concentrations in stream-sediment geochemistry*”, Journal of Geochemical Exploration 71, 73–82.
 46. Rehioui, H., Idrissi, A., Abourezq, M., Zegrari, Intelligent Data Analysis VII, pp. 70–80. Springer Berlin.
 27. Hinneburg, A., Keim, D., 1998; “*An efficient approach to clustering in large multimedia databases with noise*”, In: Proceedings KDD’98, pp. 58–65.
 28. Idrissi, A., Rehioui, H., Laghrissi, A., Retal, S., 2015; “*An Improvement of DENCLUE Algorithm for the Data Clustering*”, International Conference on Information and Communication Technology and Accessibility, ICTA 2015. Marrakech, Morocco.
 29. Jang, G.H., Won, H.C., Hwang, B.H., Choi, C.M., 2020; “*Exploratory data analysis applied in mapping multi-element soil geochemical anomaly for drill target definition: A case study from the Unpha Layered non-magmatic hydrothermal Pb-Zn deposit, DPR Korea*”, Acta Geological Sinica, <https://doi.org/10.1111/1755-6724.14404>.
 30. Jin, H., Yu, W., Li, S., 2018; “*A clustering algorithm for determining community structure in complex networks*”, Physica A 492, 980-993.
 31. Ji, H., Zeng, D., Shi, Y., Wu, Y., Wu, X., 2007; “*Semi-hierarchical correspondence cluster analysis and regional geochemical pattern recognition*”, Journal of Geochemical Exploration 93, 109–119.
 32. John, D.A., 2010; “*Porphyry Copper Deposit Model*”, Scientific Investigations Report 2010–5070–B, U.S. Geological Survey, 186 p.
 33. Hongjin, J., Yongzheng, Z., Xisheng, W., 1995; “*Correspondence cluster analysis and its application in exploration geochemistry*”, Journal of Geochemical Exploration 55, 137-144.
 34. Karbalaei Ramezanali, A., Feizi, F., Jafarirad, A., Lotfi, M., 2020; “*Geochemical Anomaly and Mineral Prospectivity Mapping for Vein-Type Copper Mineralization, Kuhsiah-e-Urmak Area, Iran: Application of Sequential Gaussian Simulation and Multivariate Regression Analysis*”, Natural Resources Research 29, 41–70.
 35. Latecki, L.J., Lazarevic, A., Pokrajac, D., 2007; “*Outlier Detection with Kernel Density Functions*”, In: Perner P. (eds) Machine Learning and Data Mining in Pattern Recognition. MLDM 2007. Lecture Notes in Computer Science, vol 4571. Springer, Berlin,

52. Sreevani, R., Murthy, C.A., 2016; "On bandwidth selection using minimal spanning tree for kernel density estimation", Computational Statistics and Data Analysis 102, 67-84.
53. Templ, M., Filzmoser, P., Reimann, C., 2008; "Cluster analysis applied to regional geochemical data: Problems and possibilities", Applied Geochemistry 23(8), 2198-2213.
54. Tian, D., Sorooshian, S., Myers, D.E., 1993; "Correspondence analysis with MATLAB", Computers & Geosciences 19(7), 1007-1022.
55. White, W.M., 2013; "Geochemistry", Wiley-Blackwell Publications, 668 p.
56. Xu, Y., Xu, N., Feng, X., 2016; "A New Outlier Detection Algorithm Based on Kernel Density Estimation for ITS", The IEEE International Conference on Internet of Things, Green Computing and Communications, Cyber, Physical and Social Computing and Smart Data. Chengdu, China.
57. Zadmehr, F., Shahrokhi, S.V., 2019; "Separation of geochemical anomalies by concentration-area and concentration number methods in the Saqez 1:100,000 sheet, Kurdistan", Iranian Journal of Earth Sciences 11(3), 196-204.
- F., 2016; "DENCLUE-IM: A New Approach for Big Data Clustering", Procedia Computer Science 83, 560 - 567.
47. Reimann, C., Filzmoser, P., Garrett, R.G., 2002; "Factor analysis applied to regional geochemical data: problems and possibilities", Applied Geochemistry 17, 185-206.
48. Saadati, H., Afzal, P., Torshizian, H., Solgi, A., 2020; "Geochemical exploration for lithium in NE Iran using the geochemical mapping prospectivity index, staged factor analysis, and a fractal model", Geochemistry: Exploration, Environment, Analysis 20(4), 461-472.
49. Sadeghi, M., Morris, G.A., Carranza, E.J.M., Ladenberger, A., Andersson, M., 2013; "Rare earth element distribution and mineralization in Sweden: An application of principal component analysis to FOREGS soil geochemistry", Journal of Geochemical Exploration 33, 160-175.
50. Shah, H., Napanda, K., D'mello, L., 2015; "Density Based Clustering Algorithms", International Journal of Computer Sciences and Engineering 3(11), 54-57.
51. Shi, M., Carr, J.R., 2001; "A modified code for R-mode correspondence analysis of large-scale problems", Computers & Geosciences 27, 139-146.