

کاربرد الگوریتم رتبه‌بندی در تهیه نقشه آنومالی‌های چند عنصری ژئوشیمیایی ورقه ۱:۱۰۰۰۰۰۰ سه‌چنگی خراسان جنوبی

حمید گرانیان^{*}

۱- استادیار، گروه مهندسی معدن، دانشگاه صنعتی بیرجند، بیرجند

(دریافت ۱۳۹۸/۰۹/۲۶، پذیرش ۱۳۹۹/۰۲/۲۸)

چکیده

تلفیق لایه‌های اکتشافی به ویژه ترکیب داده‌های ژئوشیمیایی می‌تواند برای تعیین آنومالی‌های ژئوشیمیایی چند عنصری به کار رود که نشان‌دهنده مناطق دارای پتانسیل کانی‌زایی‌اند. هدف این مقاله معرفی الگوریتم‌های رتبه‌بندی به عنوان روشی دیگر در کنار روش‌های آماری، داده-کاوی و تصمیم‌گیری چند معیاره برای تلفیق داده‌های اکتشافی است. برای این منظور از ۳۶۲ نمونه رسوبات آبراه‌ای در ورقه سه‌چنگی در استان خراسان جنوبی استفاده شده که هر نمونه نیز برای ۲۳ عنصر آنالیز شده است. پیاده کردن شش الگوریتم رتبه‌بندی VSA، MA، CA، PA، LA و CAA بر روی مجموعه داده‌ها نشان می‌دهد که محدوده‌های مناطق مستعد کانی‌زایی به دست آمده همپوشانی و موقعیت‌های تقریباً مشابه دارند. این محدوده‌ها بر روی واحدهای توف، آندزیتی، ریولیتی و رسوبات کواترنری قرار گرفته‌اند. مقایسه‌های کمی (روش جایگشت و تحلیل مولفه‌های اصلی مقاوم) و کیفی (مقایسه با نقشه‌ی زمین‌شناسی و اندیس‌های معدنی) نشان‌دهنده برتری نسبی الگوریتم‌های CA و PA است. ترکیب شش نقشه به روش میانگین‌گیری وزن‌دار دو ناحیه یکی محدوده‌ای با پتانسیل بالاتر با مساحت حدود ۲۴ کیلومتر مربع و دیگری محدوده‌ای با پتانسیل پایین‌تر با مساحت حدود ۳۱۱ کیلومتر مربع را برای فاز اکتشافی بعدی در منطقه مطالعاتی پیشنهاد کرده است.

کلمات کلیدی

آنومالی چند عنصری، رتبه‌بندی نمونه‌ها، تحلیل ویژگی، پتانسیل‌یابی، ورقه سه‌چنگی.

۱- مقدمه

یادگیری ماشین تقسیم‌بندی می‌شوند [۱۷]. روش‌های بدون یادگیری شامل رتبه‌بندی داده‌ها مبتنی بر محتوا (استفاده شده در بازیابی اطلاعات سنتی)، مبتنی بر ساختار (استفاده شده در وب فعلی) و روش ترکیبی از دو روش یاد شده است. در روش‌های مبتنی بر محتوا مدل‌هایی همانند بولی، احتمالی و فضای برداری جهت رتبه‌بندی اسناد ارائه شده است. در حالی که در روش‌های ساختاری از معیار پیوندها استفاده می‌شود که بیان‌کننده کیفیت محتوای یک صفحه از منظر صفحات بیرونی است. به روش‌هایی که از هر دوی پیوند و محتوا برای رتبه‌بندی استفاده می‌کنند، روش‌های ترکیبی گفته می‌شود. این روش‌ها به دو گروه روش‌هایی که با کمک محتوا فرآیند آنالیز محتوا را بهبود می‌بخشند و روش‌های انتشار وابستگی که با کمک ساختار وب ساختار انتشار اطلاعات محتوا را بهبود می‌بخشند، تقسیم‌بندی می‌شوند [۱۹،۱۸].

روش‌های مبتنی بر یادگیری به سه بخش الگوریتم‌های مبتنی بر نقطه، مبتنی بر جفت و مبتنی بر لیست تقسیم‌بندی می‌شوند. روش‌های مبتنی بر نقطه و جفت، مسایل رتبه‌بندی را با تغییر در نحوه داده‌ها به مسایل دسته‌بندی تبدیل می‌کنند، در حالی که روش‌های مبتنی بر لیست، بدون تغییر در داده‌های ورودی، با استفاده از بهینه کردن یک تابع هدف سعی در رتبه‌بندی پرس‌وجوی جدید می‌کنند [۲۰،۲۱].

هدف از رتبه‌بندی در بازیابی اطلاعات پیدا کردن صفحاتی از وب است که بیشترین شباهت و نزدیکی را با یک پرس‌وجو (یا یک درخواست کاربر) داشته باشد. در حالی که هدف از رتبه‌بندی داده‌های اکتشافی می‌توان مرتب کردن داده‌های چند متغیره بر اساس بیشترین احتمال مرتبط بودن یک داده با یک فرآیند کانی‌زایی تعریف شود. بر این اساس و با اقتباس از روش‌های رتبه‌بندی در بازیابی اطلاعات الگوریتم‌های رتبه‌بندی زیر معرفی می‌شوند.

اگر یک مجموعه داده D دارای n نمونه و برای هر نمونه m متغیر اندازه‌گیری شده باشد، این مجموعه داده‌های چند بعدی به صورت رابطه ۱ تعریف خواهد شد:

$$D = \begin{bmatrix} x_{11} & \cdots & x_{1m} \\ \vdots & \ddots & \vdots \\ x_{n1} & \cdots & x_{nm} \end{bmatrix} \quad (1)$$

که در آن:

x_{ij} : نمونه i ام برای متغیر j ام است.

اولین مرحله برای رتبه‌بندی داده‌ها، استانداردسازی آنها به یک بازه معین برای از بین بردن اثر واحدهای اندازه‌گیری

اکتشاف مواد معدنی در مقیاس ناحیه‌ای یکی از پروژه‌های ملی هر کشور محسوب می‌شود. داده‌های زمین‌شناسی، ژئوشیمی، ژئوفیزیکی و دورسنجی لایه‌های اطلاعاتی اکتشافات ناحیه‌ای را تشکیل می‌دهد که هر کدام نیز زیر لایه‌هایی دارند. انتخاب روش تلفیق لایه‌ها و زیر لایه‌های اطلاعاتی و پیشنهاد محدوده‌های دارای پتانسیل بالاتر کانی‌زایی یکی از نکات کلیدی و حساس هر پروژه اکتشافی محسوب می‌شود [۲۰]. روش‌های آماری، داده‌کاوی و تصمیم‌گیری چند معیاره جزو مهم‌ترین روش‌های تلفیق‌اند [۴،۳]. هدف از این مقاله، پیشنهاد الگوریتم‌های رتبه‌بندی به عنوان یک دسته روش جدید در تلفیق لایه‌های اکتشافی است.

رتبه‌بندی به فرآیند مرتب کردن داده‌ها بر اساس نظر کاربر گفته می‌شود؛ یا فرآیندی که در آن کیفیت یا کمیت یک داده با الگوریتم‌های رتبه‌بندی تخمین زده شود [۵،۶]. بنابراین نمونه‌هایی (یا رکوردهایی) که در رتبه‌های بالاتر قرار می‌گیرند، به خواسته کاربر نزدیکتر خواهند بود (خواسته کاربر در یک پروژه اکتشافی ارتباط بالاتر با کانی‌زایی یا دارا بودن پتانسیل معدنی است). از آنجا که رتبه‌بندی یک نمونه به وسیله کلیه لایه‌ها و زیر لایه‌های اطلاعاتی آن انجام می‌گیرد، در نتیجه کمیت رتبه‌بندی ماحصل تلفیق همه داده‌ها خواهد بود. از کارهای انجام شده در مبحث تلفیق داده‌های اکتشافی به ویژه داده‌های ژئوشیمیایی می‌توان به روش تحلیل مولفه‌های اصلی [۷،۸]، روش خوشه‌بندی سلسله مراتبی [۹،۱۰]، روش خوشه‌بندی k - میانگین [۱۱،۱۲]، روش‌های تصمیم‌گیری چند معیاره از قبیل AHP [۱۳،۱۴]، TOPSIS [۱۵،۱۳] و VIKOR [۱۶] اشاره کرد.

از آنجا که روش‌های رتبه‌بندی تاکنون کمتر به عنوان یک روش تلفیق داده‌های اکتشافی به کار رفته است، هدف این مقاله امکان بررسی این موضوع خواهد بود. برای این منظور از داده‌های ژئوشیمیایی رسوبات آبراهه‌ای یکی از ورقه‌ها ۱/۱۰۰۰۰۰ یعنی ورقه سه‌چنگی در استان خراسان جنوبی استفاده خواهد شد. در ادامه ضمن معرفی شش روش رتبه‌بندی، کاربرد و مقایسه آنها مورد توجه و بررسی قرار می‌گیرد.

۲- معرفی الگوریتم‌های رتبه‌بندی

روش‌های رتبه‌بندی که بیشتر در بازیابی اطلاعات کاربرد دارند به دو دسته روش‌های بدون یادگیری و روش‌های با

$$\cos \theta_i = \frac{\vec{z}_i \cdot \vec{z}_{1d}}{|\vec{z}_i| \cdot |\vec{z}_{1d}|} = \frac{\sum_{j=1}^m z_{ij}}{\sqrt{\sum_{j=1}^m z_{ij}^2} \cdot \sqrt{m}} \quad (5)$$

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

هر چه زاویه بین بردار نمونه با بردار نمونه‌ای ایده‌آل کمتر باشد، نمونه مورد نظر به نمونه ایده‌آل نزدیکتر و رتبه‌ی آن بالاتر خواهد بود و بالعکس، بنابراین مقدار به دست آمده از رابطه ۵ مقدار رتبه‌بندی هر نمونه خواهد بود و با مرتب کردن این مقادیر به صورت نزولی، رتبه هر نمونه به دست می‌آید.

۲-۳- الگوریتم احتمالی^۶

هدف از الگوریتم احتمالی (PA) پیدا کردن احتمال وابستگی یک نمونه به نمونه ایده‌آل است، بنابراین هسته اصلی این الگوریتم بر مبنای تئوری احتمال است. اگر داده‌ها به وسیله توابع توزیع گوسی چند بعدی تولید شده باشند، احتمال وابستگی و یا احتمال اینکه نمونه‌ای مانند $z_i \in D_{St}$ با مدل توزیع گوسی داده ایده‌آل تولید شده باشد برابر است با رابطه ۶ [۲۵، ۲۶]:

$$p(z_i | \mu_{z_{1d}}, \Sigma_{z_{1d}}) = \frac{1}{\sqrt{(2\pi)^m |\Sigma_{z_{1d}}|}} \exp \left\{ -\frac{1}{2} (z_i - \mu_{z_{1d}}) \Sigma_{z_{1d}}^{-1} (z_i - \mu_{z_{1d}})^T \right\}$$

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

که در آن:

$\mu_{z_{1d}}$ و $\Sigma_{z_{1d}}$: به ترتیب ماتریس‌های میانگین و واریانس-

کواریانس مدل مولد توزیع داده ایده‌آل است.

ماتریس $\mu_{z_{1d}}$ یک ماتریس یکه $1 \times m$ بعدی یعنی برابر z_{1d} و ماتریس $\Sigma_{z_{1d}}$ نیز یک ماتریس واحد $m \times m$ بعدی است که مقادیر واریانس‌ها برابر یک و مقادیر کواریانس‌ها برابر صفر است، بنابراین دترمینان این ماتریس برابر یک خواهد بود. بالا بودن احتمال به دست آمده از رابطه ۶ نشان‌دهنده شباهت و نزدیکی بیشتر نمونه به نمونه ایده‌آل خواهد بود. در نتیجه رتبه نمونه بالاتر و احتمال وابستگی آن به کانی‌زایی بیشتر است، بنابراین احتمال به دست آمده مقدار رتبه‌بندی هر نمونه و با مرتب کردن این مقادیر به صورت نزولی جایگاه یا رتبه‌ی هر نمونه نیز به دست خواهد آمد.

مختلف است. برای این منظور می‌توان از رابطه ۲ برای انتقال هر داده به بازه صفر تا یک استفاده کرد:

$$z_{ij} = \frac{(x_{ij} - \min_j)}{(\max_j - \min_j)} \rightarrow D_{St.} = \begin{bmatrix} z_{11} & \dots & z_{1m} \\ \vdots & \ddots & \vdots \\ z_{n1} & \dots & z_{nm} \end{bmatrix}$$

از آنجا که با این انتقال کمترین مقدار هر داده صفر و بیشترین آن یک خواهد بود، بنابراین نمونه‌ای که تمامی متغیرهای آن برابر یک باشد، نمونه‌ای با بالاترین احتمال مرتبط با کانی‌زایی محسوب می‌شود. این نمونه که به صورت $Z_{1d} = [1 \dots 1]$ تعریف می‌شود و یک بردار $1 \times m$ بعدی است، نمونه ایده‌آل نامگذاری می‌شود.

۲-۱- الگوریتم‌های جمعی و ضربی^۴

در این الگوریتم‌ها که از روش هاله‌های ژئوشیمیایی مرکب اقتباس شده است [۲۲، ۲۳]، متغیرهای استاندارد شده هر نمونه با هم جمع و یا در هم ضرب می‌شوند. ماتریس مقدار رتبه‌بندی مجموعه داده‌ها در الگوریتم جمعی (CA) از رابطه ۳ و در الگوریتم ضربی (MA) از رابطه ۴ و به صورت زیر به دست می‌آید:

$$R_{Cumulative} = \begin{bmatrix} \sum_{j=1}^m z_{1j} \\ \vdots \\ \sum_{j=1}^m z_{nj} \end{bmatrix} \quad (3)$$

$$R_{Multiplicative} = \begin{bmatrix} \prod_{j=1}^m z_{1j} \\ \vdots \\ \prod_{j=1}^m z_{nj} \end{bmatrix} \quad (4)$$

در مرحله بعد با مرتب کردن مقادیر به دست آمده به صورت نزولی رتبه هر نمونه به دست می‌آید.

۲-۲- الگوریتم فضای برداری^۵

در الگوریتم فضای برداری (VSA)، رتبه هر نمونه از مجموعه D_{St} به صورت زاویه بین نمونه مورد نظر با نمونه ایده‌آل تعریف می‌شود. برای محاسبه این زاویه از رابطه ۵ می‌توان استفاده کرد [۲۴]:

۴-۲- الگوریتم پیوندی^۷

سپس وزن هر یک از ویژگی‌ها به یکی از دو روش زیر محاسبه می‌شود.

I روش ریشه مربع: در این حالت وزن هر ویژگی (w) از رابطه ۱۰ به دست می‌آید [۲۸]:

$$w_j = [m^{-1} \sum_{k=1}^m s_{jk}^2]^{1/2} \quad j = 1, 2, \dots, m \quad (10)$$

II روش مولفه‌های اصلی: اگر λ_j آمین مقدار ویژه ماتریس S باشد که $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_m \geq 0$ و A_j بردار ویژه λ_j ماتریس S باشد، ماتریس ضرایب تطبیقی را می‌توان به روش رابطه ۱۱ تجزیه کرد [۲۸]:

$$S = \lambda_1 A_1 A_1^T + \lambda_2 A_2 A_2^T + \dots + \lambda_m A_m A_m^T \quad (11)$$

اگر مولفه اول به اندازه کافی بزرگ باشد، به نحوی که بخش اعظم تغییرپذیری (به طور مثال حدود ۸۰ درصد) با این مولفه قابل توجیه باشد، خواهیم داشت (رابطه ۱۲):

$$S \approx \lambda_1 A_1 A_1^T \quad (12)$$

بنابراین بردار وزن ویژگی‌ها برابر بردار مولفه اول یعنی $W = (a_1, a_2, \dots, a_m)^T$ خواهد بود. در صورت منظم بودن بردارها رابطه ۱۳ نیز بین وزن ویژگی‌ها برقرار است:

$$\sum_{j=1}^m w_j = 1 \quad \text{and} \quad 0 \leq w_j \leq 1 \quad (13)$$

وزن‌های بالاتر نشان‌دهنده معرفت‌تر بودن ویژگی‌ها (یا مستقل‌تر بودن آنها) و وزن‌های پایین‌تر نشان‌دهنده غیرمعرفت‌تر بودن ویژگی‌ها (یا وابسته‌تر بودن آنها) برای توجیه پدیده مرتبط با تحلیل ویژگی خواهد بود [۲۹، ۲۸]. بنابراین وزن‌های به دست آمده برابر کمیت رتبه‌بندی هر ویژگی یا نمونه خواهد بود که با مرتب کردن وزن‌ها به ترتیب نزولی، ویژگی‌ها یا نمونه‌ها رتبه‌بندی می‌شوند.

سادگی محاسبات مهم‌ترین مزیت الگوریتم‌های جمعی و ضربی محسوب می‌شود. در حالی که تاثیر مقادیر متغیرهای بزرگتر در نتیجه الگوریتم ضربی و دامنه کوچکتر مقادیر رتبه‌بندی در الگوریتم جمعی نیز تفاوت این دو الگوریتم خواهد بود. سادگی و انعطاف‌پذیری نسبت به تعداد متغیرها مزیت الگوریتم فضای برداری و عدم نمایش ارتباط بین رتبه‌ها از معایب آن است. همچنین محاسبه شباهت قطعی بین

در الگوریتم پیوندی (LA)، رتبه‌بندی نمونه‌ها مستقل از نمونه ایده‌آل و بر اساس مقایسه هر نمونه با سایر نمونه‌ها انجام می‌گیرد. بر این منظور، ابتدا ماتریس رتبه هر نمونه به صورت تک متغیره تشکیل می‌شود، بنابراین خواهیم داشت (رابطه ۷):

$$R_{Univariate} = \begin{bmatrix} r_{11} & \dots & r_{1m} \\ \vdots & \ddots & \vdots \\ r_{n1} & \dots & r_{nm} \end{bmatrix} \quad (7)$$

که در آن:

r_{ij} : رتبه‌ی نمونه‌ی i ام در متغیر j ام است.
با جمع کلیه رتبه‌های یک نمونه، مقدار رتبه‌بندی چند متغیره نمونه به وسیله الگوریتم پیوندی به دست می‌آید. یعنی رابطه ۸:

$$R_{Linkage} = \begin{bmatrix} \sum_{j=1}^m r_{1j} / (n-1) \times m \\ \vdots \\ \sum_{j=1}^m r_{nj} / (n-1) \times m \end{bmatrix} \quad (8)$$

در این رابطه، برای نرمالیزه کردن مقدار رتبه‌بندی هر نمونه، مقدار آن بر حداکثر مقدار ممکن رتبه یک نمونه (یعنی) تقسیم شده است، بنابراین با این الگوریتم، مقدار رتبه‌بندی نمونه ایده‌آل برابر یک و مقدار رتبه‌بندی سایر نمونه‌ها، عددی بین صفر تا یک خواهد بود. با مرتب کردن داده‌های ماتریس به دست آمده از رابطه ۸ به صورت نزولی رتبه هر نمونه نیز به دست می‌آید.

۵-۲- الگوریتم تحلیل ویژگی^۸

الگوریتم تحلیل ویژگی (CAA) به عنوان یک روش تحلیل آماری چند متغیره، اولین بار توسط بوتبال در سال ۱۹۷۱ پیشنهاد گردید [۲۷]. برای استفاده از تحلیل ویژگی داده‌ها باید به صورت باینری (مقدار صفر به معنی نامطلوب یا زمینه و مقدار یک به معنی مطلوب یا آنومالی) و یا استاندارد شده باشند. در صورت استفاده از داده‌های استاندارد شده ماتریس ضرایب تطبیقی بین ویژگی‌ها (که می‌تواند متغیرها یا نمونه‌ها باشد) به صورت رابطه ۹ تعریف می‌شود [۲۸]:

$$S = (s_{jk}) = n^{-1} D_{St} D_{St}^T = n^{-1} \sum_{i=1}^n z_{ij} z_{ik} \quad (9)$$

$j, k = 1, 2, \dots, m$

۱- ب). جنس این واحد سنگی از آندزیت تا آندزیت بازالیت متغیر و میزبان تعداد از کانی‌زایی‌های فلزی است. سنگ‌های ترشیری داسیت- ریوداسیت با مساحت حدود ۴ کیلومتر مربع به صورت یک تپه خیلی کوچک و در بخش جنوبی منطقه مطالعاتی قرار گرفته‌اند. واحدهای سنگی دیوریتی و ریولیتی نیز سن ترشیری دارند و به ترتیب حدود ۱۶ و ۶ کیلومتر مربع از مساحت منطقه را پوشش می‌دهند. بخشی از سنگ‌های رسوبی منطقه به صورت خیلی محدود دچار دگرگونی شده که کمتر از ۰/۲ کیلومتر مربع وسعت دارند و در بخش جنوبی منطقه مطالعاتی قابل مشاهده است (شکل ۱- ب). واحد سنگی رسوبی- آذرین با جنس توف‌های پیروکلاستیک و توف سبز به صورت پهنه‌های هم‌راستا با سنگ‌های آندزیتی در منطقه مشاهده می‌شوند. این واحد سنگی با سن ترشیری نیز حدود ۱۳/۳ درصد از مساحت منطقه مطالعاتی را پوشش می‌دهد. بزرگترین واحد سنگی با پوشش حدود ۶۹ درصد از منطقه مطالعاتی شامل رسوبات آبرفتی و پادگانه کواترنری است که در بعضی بخش‌ها به صورت دشت و کویر قابل مشاهده است. فعالیت‌های ماگمایی از نوع کالک آلکان پتاسیم متوسط تا بالا در ائوسن میانی تا اولیگوسن پایینی همراه با فرآیندهای تکنونیک بافت تشکیل کانی‌زایی‌های فلزی در منطقه مطالعاتی شده است [۳۶،۳۰]. سه کانسار سرب- روی \pm مس از نوع رگه‌ای مرتبط با فرآیندهای هیدروترمال ولکانیکی- ساب ولکانیکی همراه با آلتراسیون‌های آرژیلیکی، سریسیتی و کربناتی در منطقه تشکیل شده‌اند [۳۰]. همچنین ۶ اندیس کانی‌زایی چندفلزی مرتبط با فرآیندهای هیدروترمال شامل طلا، نقره، مس، سرب و روی و ۶ محدوده حاوی کانی‌زایی آهن پلاسری از نوع مگنتیت- هماتیت همراه با کانی‌های تیتانیم‌دار در منطقه وجود دارد (شکل ۱- ب).

۴- داده‌های ژئوشیمیایی

در محدوده ورقه سه‌چنگی به وسعت حدود ۳۰۰۰ کیلومتر مربع، ۳۶۲ نمونه ژئوشیمیایی از رسوبات آبراهه‌ای به وسیله سازمان زمین‌شناسی و اکتشافات معدنی ایران برداشت شده است. پراکندگی و موقعیت این نمونه‌ها در شکل ۲ نشان داده شده است. به دلیل قرارگیری بخش شمال‌غربی و شرق محدوده در رسوبات کویری از این نواحی نمونه‌برداری انجام نگرفته است. بخش کوچکتر از ۸۰ مش نمونه‌ها برای آنالیز شیمیایی به روش ICP-OES انتخاب شده است. نمونه‌ها در آزمایشگاه سازمان زمین‌شناسی و اکتشافات معدنی ایران برای ۴۴ عنصر

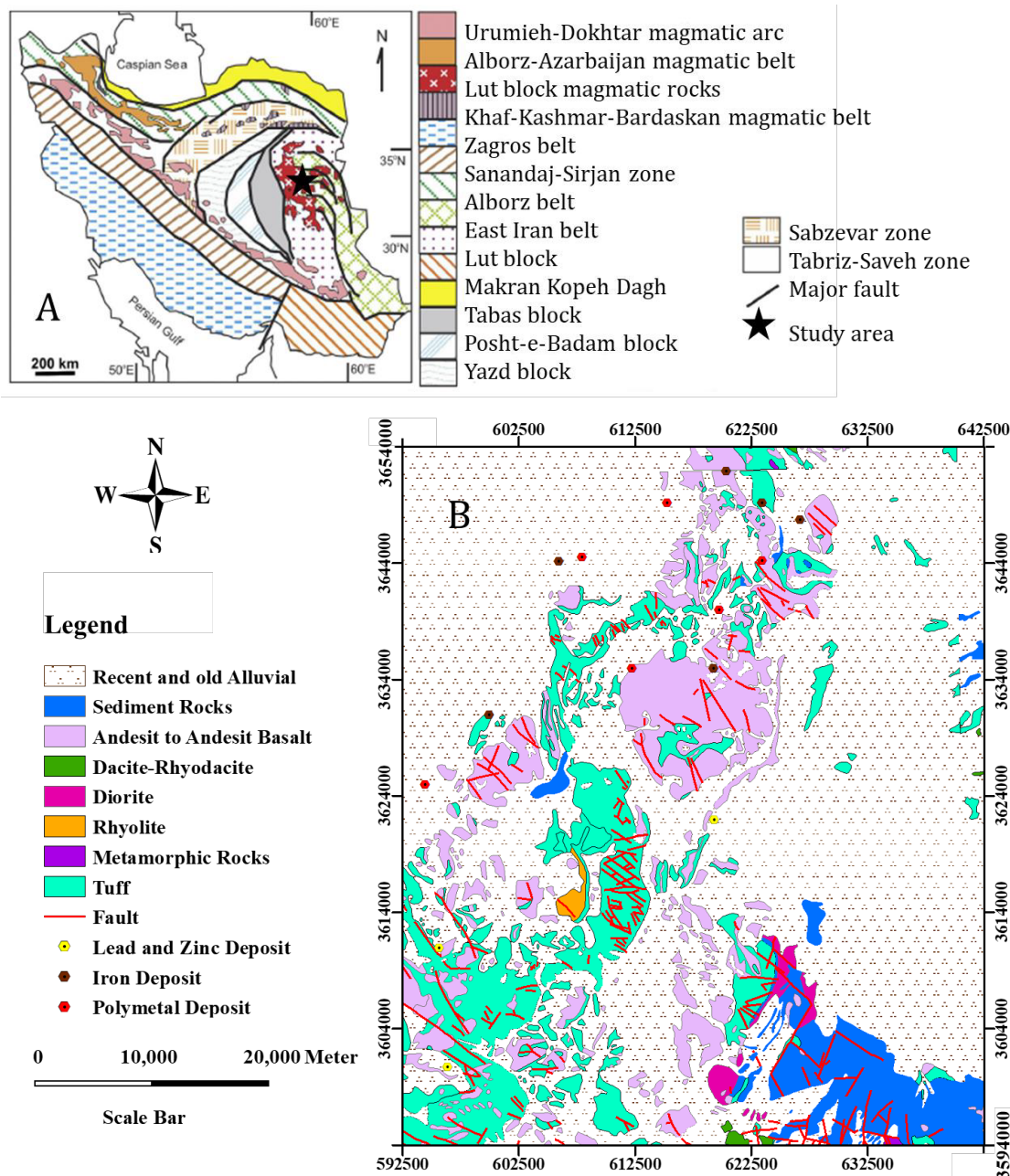
نمونه‌ها با نمونه ایده‌آل و دامنه محدوده نتایج رتبه‌بندی که باعث تفسیر راحت‌تر نتایج خواهد شد، نیز از مزایای الگوریتم احتمالی است. سادگی محاسبات و رتبه‌بندی نمونه‌ها به کمک اتصال و پیوند بین رتبه هر متغیر از مزیت‌های الگوریتم پیوندی است. در صورتی که یکسان فرض کردن تاثیر متغیرها در نتیجه نهایی رتبه‌بندی که باعث به دست آمدن نتایج یکسان و تفسیر مشکل‌تر خواهد شد نیز، از معایب مهم این الگوریتم است. پوشش داده مقادیر کم یک متغیر به وسیله متغیرهای دیگر و سادگی محاسبات، مزیت و تفسیر سخت‌تر نیز از معایب الگوریتم تحلیل ویژگی است.

نکات یاد شده مهم‌ترین مزایا و معایب الگوریتم‌های پیشنهادی از جنبه ریاضی محسوب می‌شود. در حالی که از جنبه کاربردی، تاکنون فقط الگوریتم‌های جمعی، ضربی و تحلیل ویژگی در مباحث اکتشافی استفاده شده است، بنابراین در این مقاله کاربرد کلیه الگوریتم‌ها مورد بررسی و تحلیل قرار خواهد گرفت.

۳- زمین‌شناسی منطقه مطالعاتی

منطقه مطالعاتی شامل ورقه زمین‌شناسی ۱/۱۰۰۰۰۰ سه‌چنگی است که در مختصات جغرافیایی $00^{\circ} 58'$ تا $30'$ 58° طول شرقی و $30' 32''$ تا $33^{\circ} 00'$ عرض شمالی قرار دارد (شکل ۱- الف). این محدوده به لحاظ جغرافیایی در شرق ایران در استان خراسان جنوبی و به لحاظ تقسیم‌بندی‌های زمین‌شناسی- ساختمانی در بخش مرکزی بلوک لوت قرار دارد [۳۰]. بلوک لوت یکی از خرد قاره‌های ایران مرکزی محسوب می‌شود که وسعتی حدود ۹۰۰ کیلومتر در راستای شمالی- جنوبی و حدود ۲۰۰ کیلومتر در راستای شرقی- غربی دارد. فعالیت گسل‌های امتداد لغز دو طرف بلوک لوت به همراه فروانش بلوک افغان به زیر بلوک لوت باعث تشکیل سنگ‌های کالک- آلکان در بخش‌های وسیعی از نواحی مرکزی و شمالی این بلوک شده است [۳۱-۳۴].

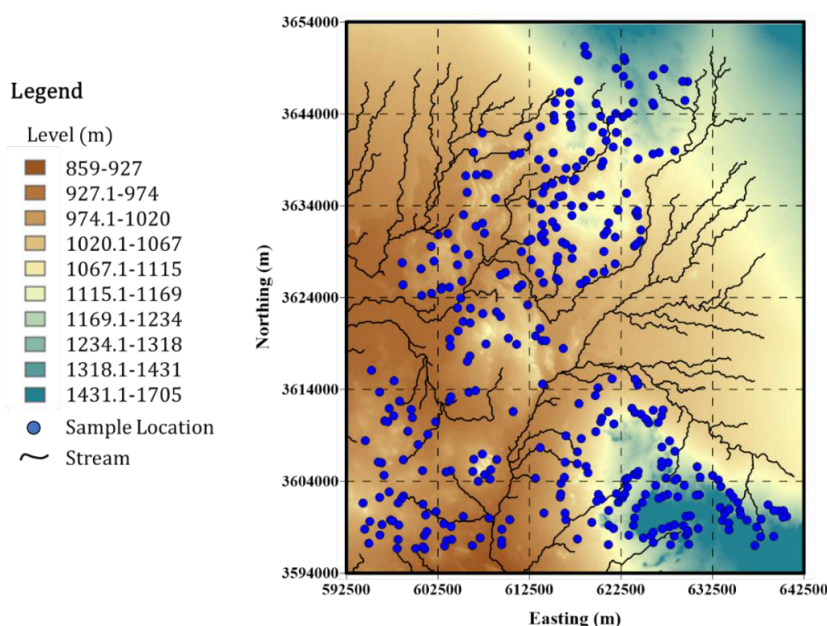
سنگ‌های رسوبی که بیشتر در بخش جنوب شرقی منطقه مطالعاتی مشاهده می‌شوند (شکل ۱- ب)، شامل کنگومرا، شیل، ماسه‌سنگ، مارن، سنگ آهک و ژئوپس است. این واحد سنگی حدود ۴/۹ درصد از مساحت منطقه مطالعاتی را پوشش می‌دهد و دارای سن تریاس بالایی تا ترشیری است. واحد سنگی آندزیتی با سن ترشیری به صورت یک پهنه با راستای شمال شرقی- جنوب غربی حدود ۱۱/۸ درصد از مساحت منطقه مطالعاتی را به خود اختصاص داده است (شکل



شکل ۱: موقعیت منطقه مطالعاتی بر روی نقشه واحدهای زمین‌شناسی - ساختاری ایران [۳۵،۳۰] و نقشه ساده شده زمین‌شناسی منطقه مطالعاتی (اقتباس از نقشه زمین‌شناسی ۱/۱۰۰۰۰۰ سه‌چنگی، سازمان زمین‌شناسی و اکتشافات معدنی ایران)

آنالیز شده است. برای ارزیابی دقت نمونه‌برداری‌ها و آنالیزها، ۱۴ نمونه تکراری نیز برداشت شده است. محاسبه انحراف معیار نسبی (RSD) به روش تامسون-هاوارث (۱۹۷۶) نشان می‌دهد که این پارامتر برای عناصر اصلی کمتر از ۵ درصد و برای سایر عناصر کمتر از ۱۰ درصد است. در این مقاله از نتایج آنالیز شیمیایی ۲۳ عنصر فلزی

استفاده شده است. مرتبط بودن این عناصر با کانسارهای فلزی هیدروترمال و داشتن مقادیر سنسورد کمتر ملاک انتخاب این عناصر است. جدول ۱ پارامترهای آماری نتایج این آنالیزها را بعد از جایگزینی مقادیر سنسورد با ۳/۴ حد حساسیت دستگاه نشان می‌دهد. مقایسه میانگین داده‌ها با مقادیر فراوانی این عناصر در پوسته زمین (عدد کلارک) نشان‌دهنده غنی‌شدگی



شکل ۲: موقعیت و پراکندگی نمونه‌های ژئوشیمیایی بر روی نقشه ارتفاع رقومی منطقه مطالعاتی

به کمک الگوریتم‌های رتبه‌بندی اشاره شده در بخش قبل با ترکیب ۲۳ بعد هر نمونه، یک مقدار برای هر نمونه به دست خواهد آمد که نشان‌دهنده کمیتی برای رتبه‌بندی آن نمونه است، بنابراین خروجی الگوریتم‌های رتبه‌بندی ماتریس 1×362 خواهد بود. هرچه رتبه نمونه یا کمیت معادل آن بالاتر باشد، احتمال ارتباط نمونه با کانی‌زایی بیشتر خواهد بود. شکل ۳ کمیت به دست آمده از الگوریتم‌های رتبه‌بندی را به صورت نقشه کنتوری نشان می‌دهد. برای مقایسه راحت‌تر نتایج به دست آمده نقشه‌های کنتوری به دو صورت با مقیاس پیوسته و کلاسه‌بندی شده (به روش فرکتالی عیار-مساحت) نمایش داده شده است. برای پیاده کردن این روش کمیت رتبه‌بندی برابر عیار فرض شده است. با توجه به آستانه‌های به دست آمده با روش فرکتالی عیار-مساحت، محدوده مطالعاتی به سه بخش محدوده دارای نمونه‌هایی با رتبه‌بندی بالا (مستعدترین محدوده دارای احتمال کانی‌زایی)، محدوده دارای نمونه‌هایی با رتبه‌بندی متوسط (با احتمال کانی‌زایی متوسط) و محدوده دارای نمونه‌هایی با رتبه‌بندی پایین (با احتمال کانی‌زایی کم) تقسیم شده است. تفکیک مقادیر رتبه‌بندی به سه بخش بر اساس بهترین نقاط شکست منحنی، حداقل بودن تعداد آنها و برازش نقاط با سه خط مستقیم با بالاترین ضریب تبیین انجام گرفته است.

نقشه‌های کنتوری با مقیاس پیوسته در شکل ۳ نشان می‌دهد که اکثر نمونه‌ها در الگوریتم‌های MA و PA رتبه‌های

برای عناصر Sb, Pb, Mo, Fe, Cr, Co, Bi, Ba, Au, As, Sn, Sc و V و تهی‌شدگی برای عناصر Hg, Cu, Cd, Ag, Zn, W, Ti, Sr, Ni, Mn حداکثر هر متغیر از میانه به علاوه دو برابر انحراف مطلق از میانه بیشتر است. همچنین پارامترهای چولگی و کشیدگی داده‌ها نیز نشان‌دهنده غیرنرمال بودن توزیع داده‌ها است، بنابراین داده‌های جدول ۱ نشان‌دهنده مقادیر آنومالی در داده‌ها یا چند جامعه‌ای بودن توزیع داده‌ها است.

با توجه به ماهیت بسته بودن داده‌های حاصل از نتایج آنالیزهای شیمیایی و برای کاهش رابطه کاذب بین متغیرها، نیاز به باز کردن سیستم داده‌ها قبل از هرگونه آنالیز آماری چند متغیره است [۳۸، ۳۷]. روش‌های انتقال لگاریتمی بهترین انتخاب هستند که در این مقاله از روش انتقال لگاریتمی میان مرکز (Clf) به دلیل حفظ ابعاد ماتریس داده‌ها استفاده شده است [۳۹، ۱۲]. با توجه به عدم نیاز به نرمال‌سازی توزیع داده‌ها در الگوریتم‌های رتبه‌بندی، داده‌های انتقال یافته Clf برای حذف اثر واحدهای اندازه‌گیری، استاندارد شده و به بازه صفر تا یک انتقال یافته است [۱۲]. در نهایت، الگوریتم‌های رتبه‌بندی بر روی مجموعه داده‌های استاندارد شده پیاده شده است.

۵- بحث

ماتریس داده‌های استاندارد شده دارای ابعاد 23×362 است.

جدول ۱: پارامترهای آمار توصیفی نتایج آنالیزهای شیمیایی

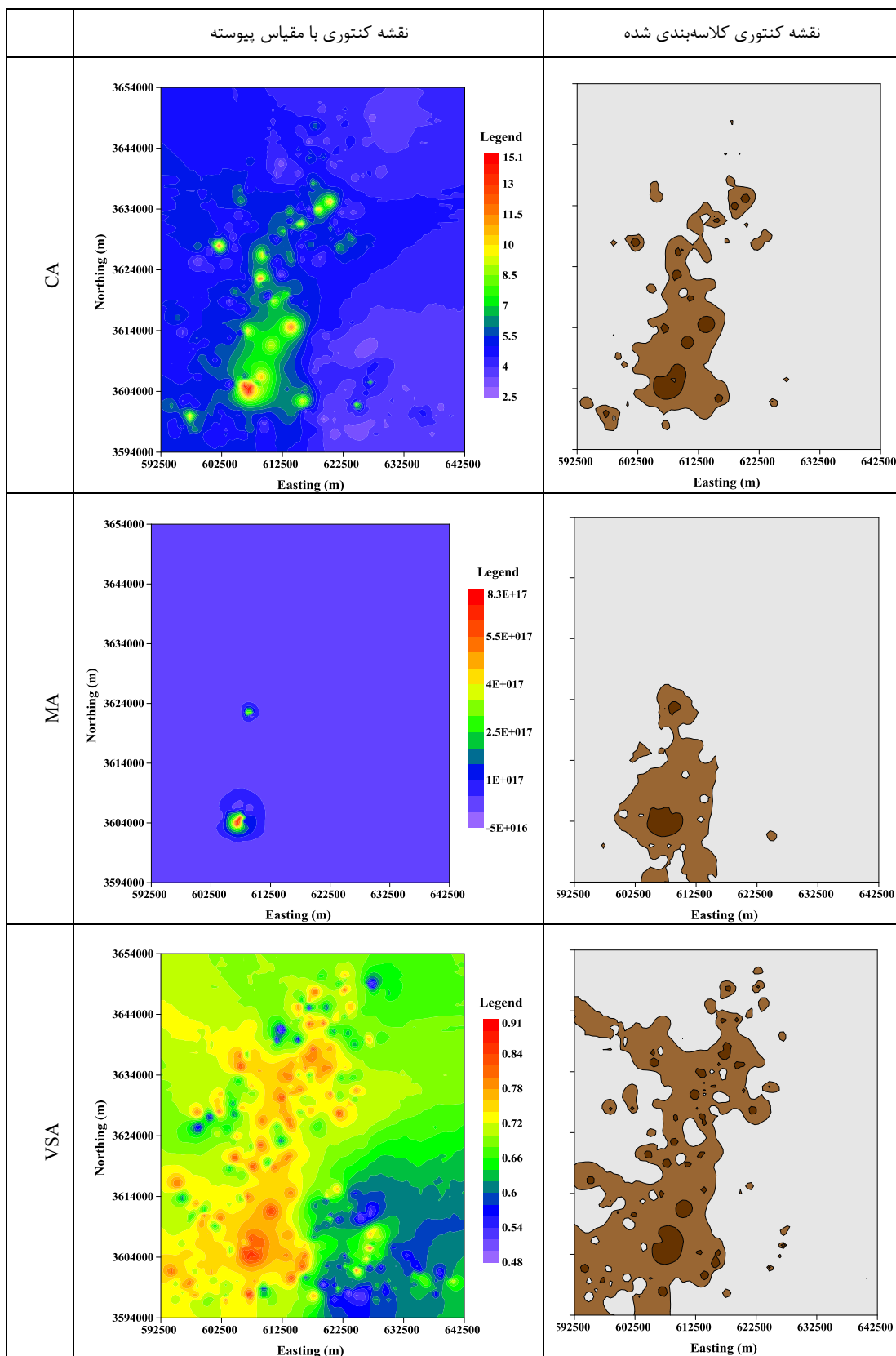
Variable (unit)	Mean	St.Dev.	Minimum	Median	Maximum	Skewness	Kurtosis	MAD
Ag (ppm)	0.056	0.008	0.044	0.054	0.147	5.35	50.15	0.003
As (ppm)	9.29	2.92	5.59	8.65	18.81	1.78	3.24	1.40
Au (ppb)	13.17	3.68	0.88	1.2	3.8	2.11	8.55	0.2
Ba (ppm)	408.07	60.41	32.10	423.20	475.59	-2.81	10.57	23.42
Bi (ppm)	0.202	0.064	0.139	0.186	0.687	3.70	18.77	0.021
Cd (ppm)	0.093	0.076	0.025	0.08	1.278	11.03	165.22	0.02
Co (ppm)	22.35	16.11	9.90	17.12	101.09	3.17	11.29	4.45
Cr (ppm)	121.99	94.07	45.31	90.56	502.41	2.64	7.41	29.23
Cu (ppm)	28.13	3.46	14.91	27.85	54.61	1.44	10.40	1.78
Fe (%)	6.026	1.426	4.460	5.62	12.502	2.39	6.94	0.60
Hg (ppm)	0.0164	0.0022	0.011	0.016	0.044	5.74	74.32	0.001
Mn (%)	0.077	0.022	0.055	0.071	0.165	2.22	5.71	0.009
Mo (ppm)	1.26	0.34	0.67	1.21	3.42	2.30	9.91	0.18
Ni (ppm)	37.75	36.18	13.25	26.00	225.03	3.58	14.13	7.55
Pb (ppm)	23.03	33.24	0.14	19.32	322.45	8.34	72.05	3.79
Sb (ppm)	0.75	0.19	0.27	0.741	3.09	5.02	59.14	0.09
Sc (ppm)	16.72	9.64	9.22	13.67	61.21	3.02	10.01	2.76
Sn (ppm)	3.59	0.18	2.75	3.61	4.06	-0.76	1.40	0.11
Sr (ppm)	173.47	19.05	143.40	170.09	396.00	5.33	53.87	8.30
Ti (ppm)	1.23	0.27	0.88	1.16	2.33	2.44	6.75	0.10
V (ppm)	160.81	112.88	64.40	123.18	616.02	2.62	7.31	36.66
W (ppm)	0.79	0.28	0.44	0.76	3.86	4.28	40.51	0.12
Zn (ppm)	67.69	29.47	35.27	62.00	334.18	4.63	30.07	9.57

تقریباً یکسان است. اختلاف به دست آمده بیشتر در وسعت محدوده‌ها است. به طوری که بیشترین وسعت محدوده بالاترین احتمال کانی‌زایی متعلق به روش VSA و کمترین وسعت متعلق به روش PA است. همچنین روش VSA نیز بیشترین وسعت را برای محدوده با احتمال کانی‌زایی متوسط و روش MA کمترین وسعت را نشان می‌دهد (البته بالا بودن وسعت آنومالی در بعضی شرایط نیز ممکن است نشان‌دهنده آنومالی کاذب باشد). در جدول ۲ وسعت هر کدام از این محدوده‌ها و در روش‌های مختلف رتبه‌بندی ذکر شده است. مقایسه شکل ۳ با نقشه زمین‌شناسی محدوده (شکل ۱) نشان می‌دهد که محدوده‌های دارای احتمال کانی‌زایی بالا و متوسط بیشتر در واحدهای سنگی آندزیت تا آندزیت-بازالت و توف‌های داسیتی اتفاق افتاده است.

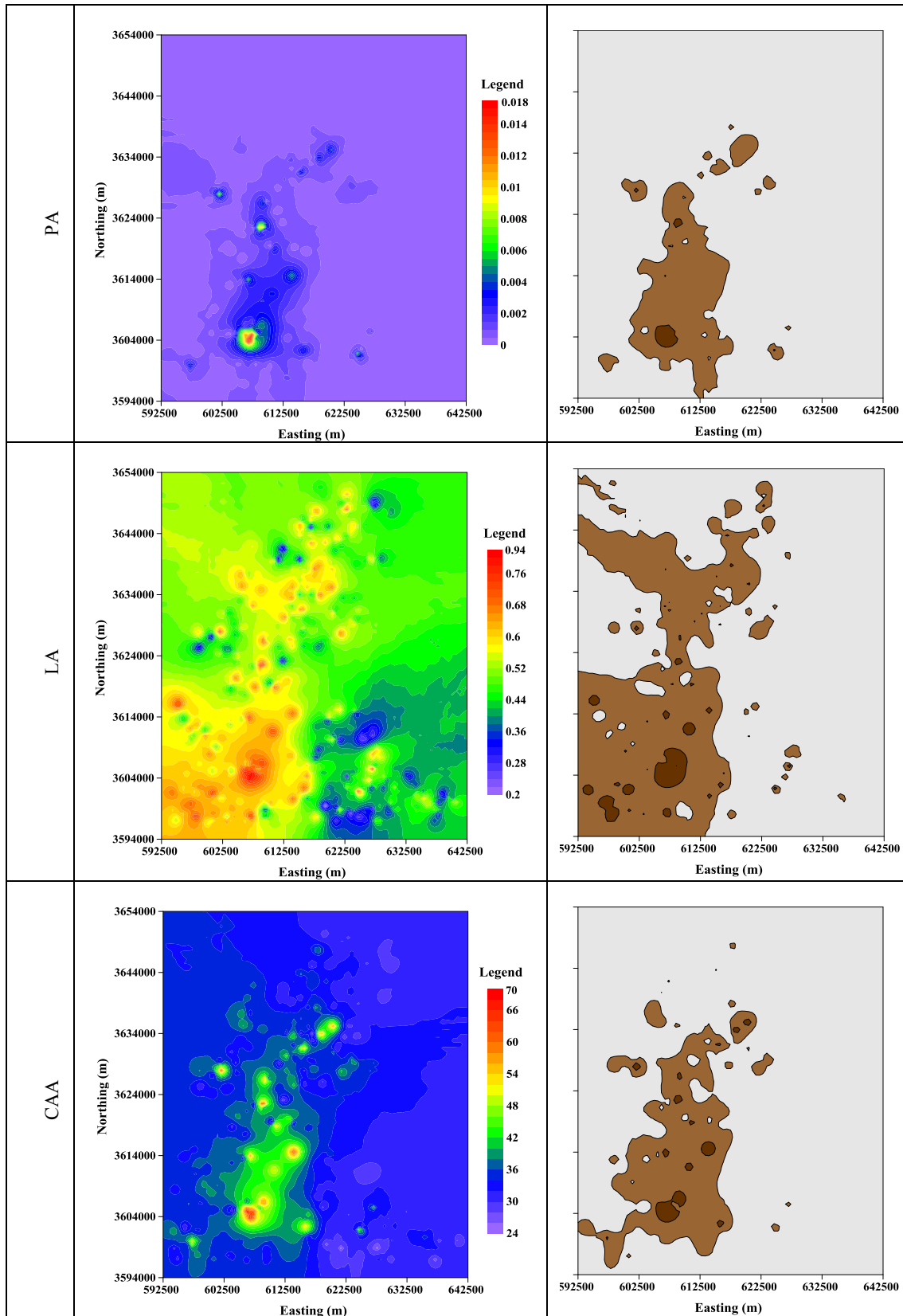
مقایسه نتایج به دست آمده از الگوریتم‌های رتبه‌بندی در شکل ۳ از دو جنبه کمی (ریاضی) و کیفی (زمین‌شناسی) امکان‌پذیر است که در ادامه جزئیات آنها مورد بررسی قرار می‌گیرد.

پایین دارند. در حالی که برعکس در الگوریتم‌های VSA و LA اکثر نمونه‌ها رتبه‌های بالا دارند. این نکته به ماهیت ریاضی الگوریتم‌ها در محاسبه کمیت رتبه‌ی نمونه‌ها برمی‌گردد. به طوری که ضرب کردن مقادیر متغیرهای استاندارد شده و تقسیم کردن احتمال وابستگی بر $\sqrt{(2\pi)^{23}}$ به ترتیب در الگوریتم‌های MA و PA مقدار کمیت رتبه‌بندی را کاهش و کسینوس گرفتن از زوایایی کوچک و جمع کردن رتبه‌های یک نمونه به ترتیب در الگوریتم‌های VSA و LA مقدار کمیت رتبه‌بندی را افزایش می‌دهد. مقادیر کمیت رتبه‌بندی در الگوریتم‌های MA و CAA نیز بین این دو دسته قرار می‌گیرند. نقشه کنتروری با مقیاس پیوسته این الگوریتم‌ها در شکل ۳ نیز این نکته را به خوبی نشان می‌دهد.

نقشه‌های کنتروری کلاسه‌بندی شده در شکل ۳ نشان می‌دهد که موقعیت محدوده‌ها با رتبه‌بندی بالا و متوسط (محدوده‌ها با احتمال کانی‌زایی بالا و متوسط) در تمامی الگوریتم‌های رتبه‌بندی



شکل ۳: نقشه‌های کنتوری پیوسته (سمت چپ) و کلاسه‌بندی شده (سمت راست) رتبه‌ی نمونه‌ها به روش‌های مختلف (راهنمای نقشه‌های کلاسه‌بندی شده: محدوده دارای رتبه‌بندی پایین، محدوده دارای رتبه‌بندی متوسط و محدوده دارای رتبه‌بندی بالا)



شکل ۳ (ادامه): نقشه‌های کنتوری پیوسته (سمت چپ) و کلاسه‌بندی شده (سمت راست) رتبه‌ی نمونه‌ها به روش‌های مختلف (راهنمای نقشه‌های کلاسه‌بندی شده: محدوده دارای رتبه‌بندی پایین، محدوده دارای رتبه‌بندی متوسط و محدوده دارای رتبه‌بندی بالا)

جدول ۲: مساحت محدوده‌ها با رتبه‌بندی مختلف

الگوریتم رتبه‌بندی	مساحت (کیلومتر مربع)		
CA	۲۶۴۹٫۸	۳۵۰٫۲	۴۰٫۹
MA	۲۷۱۳٫۱	۲۸۶٫۹	۲۴٫۶
VSA	۲۱۸۰٫۲	۸۱۹٫۸	۶۳٫۸
PA	۲۶۳۹٫۰	۳۶۱٫۰	۱۲٫۰
LA	۲۰۴۸٫۶	۹۵۱٫۴	۵۰٫۷
CAA	۲۵۱۳٫۱	۴۸۶٫۹	۲۶٫۰

جدول ۳: مقادیر معیار R_i در ۶ روش جایگشتی رتبه‌بندی

CA	MA	VSA	PA	LA	CAA
۲۲۷۰٫۶	۲۳۱۷۱	۲۲۴۵۹	۲۳۱۸۰	۲۲۵۸۸	۱۸۱۳۲

ابتدا قدر مطلق اختلاف رتبه هر نمونه در دو الگوریتم محاسبه و سپس مجموع قدر مطلق کلیه نمونه‌ها با یکدیگر جمع می‌شود. هرچه میزان این مجموع کمتر باشد، نتایج رتبه‌بندی دو الگوریتم به هم نزدیکتر خواهد بود. جدول ۴ مقادیر اختلاف رتبه‌بندی نمونه‌ها را نشان می‌دهد، که کمترین مقدار متعلق به مقایسه الگوریتم CA با PA است، بنابراین نتایج این دو الگوریتم بیشترین شباهت را به یکدیگر دارد که شکل ۳ نیز این نکته را تایید می‌کند. بیشترین تفاوت نیز بین دو الگوریتم LA و CAA است.

جدول ۴: مقدار اختلاف رتبه‌بندی نمونه‌ها در الگوریتم‌های مختلف رتبه‌بندی

	CA	MA	VSA	PA	LA	CAA
CA	۰					
MA	۵۶۲۲	۰				
VSA	۹۷۰۶	۶۸۸۸	۰			
PA	۳۰۸۸	۴۰۱۲	۶۷۶۲	۰		
LA	۱۰۱۹۲	۶۱۷۶	۱۱۱۱۲	۹۶۰۰	۰	
CAA	۱۱۴۳۲	۱۵۹۴۶	۲۰۳۷۲	۱۴۱۵۰	۱۸۵۵۵۸	۰

استفاده از نقشه کنترلی امتیازات مولفه‌های اصلی یکی از روش‌های شناخته شده ترسیم آنومالی‌های چند عنصری است [۸،۷]. برای این منظور از روش تحلیل مولفه‌های اصلی مقاوم پیشنهادی توسط Filzmoser و همکارانش استفاده شده است [۳۸]. یعنی ابتدا داده‌ها به روش انتقال لگاریتمی ایزومتریک (ilr) از سیستم بسته به سیستم باز تبدیل شده و سپس تحلیل مولفه‌های اصلی مقاوم بر روی مجموعه داده‌های باز انجام شده است. در نهایت نیز ماتریس امتیازات مولفه‌های اصلی از سیستم عددی ilr به سیستم عددی clr برگشت داده می‌شود. شکل ۴ بردار ویژه سه مولفه اول را نشان می‌دهد که ۹۰٪ درصد تغییرپذیری را توجیه می‌کند. علیرغم بالا بودن سهم مولفه اول در تغییرپذیری کل داده‌ها (حدود ۷۳٫۶ درصد)، به دلیل پایین بودن مقادیر بردارهای ویژه اکثر متغیرها در این مولفه و عدم ارتباط ژنتیکی بین عناصر وانادیم، نیکل، کروم و کبالت با سنگ‌های رسوبی این منطقه، نمی‌توان از

۵-۱- جنبه ریاضی

استفاده از معیار R_i در روش جایگشت یکی از روش‌های تست نتایج رتبه‌بندی است. در روش جایگشت تعداد حالت‌های رتبه‌بندی برای n نمونه برابر $n!$ است. یعنی در این مجموعه داده‌ها، ما با $362!$ حالت برای رتبه‌بندی نمونه مواجه هستیم که فقط ۶ حالت آن در شکل ۳ نشان داده شده است. معیار R_i از رابطه ۱۴ به دست می‌آید [۴۰]:

$$(14)$$

$$R_i = \sum_{j \in C'_{kl}} w_j - \sum_{j \in D'_{kl}} w_j ; i = 1, 2, \dots, n!$$

$$\begin{cases} C'_{kl} = \{j | z_{kj} \geq z_{lj} \\ D'_{kl} = \{j | z_{kj} \leq z_{lj} \end{cases} \quad k, l = 1, 2, \dots, n ; k \neq l$$

که در آن:

C'_{kl} و D'_{kl} : به ترتیب مجموعه‌های هماهنگ و ناهماهنگ با نتایج رتبه‌بندی

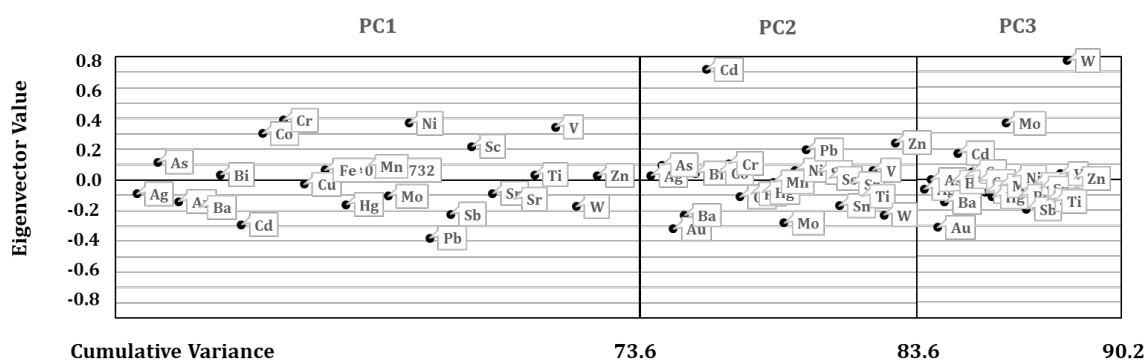
w_j : برابر وزن هر متغیر است [۴۰].

در این تحقیق فرض شده است که همه متغیرها وزن برابر و معادل $\frac{1}{23}$ دارند. در نهایت الگوریتمی که دارای معیار R_i بالاتری باشد، مناسب‌ترین جایگشت برای رتبه‌بندی خواهد بود [۴۱،۴۰]. برای مقایسه نتایج الگوریتم‌های نشان داده شده در شکل ۳، معیار R_i از رابطه ۱۴ محاسبه شده که در جدول ۳ مقادیر آنها ذکر شده است. مطابق داده‌های این جدول، الگوریتم MA بهترین عملکرد، الگوریتم PA در اولویت دوم و الگوریتم CAA در اولویت آخر قرار دارد.

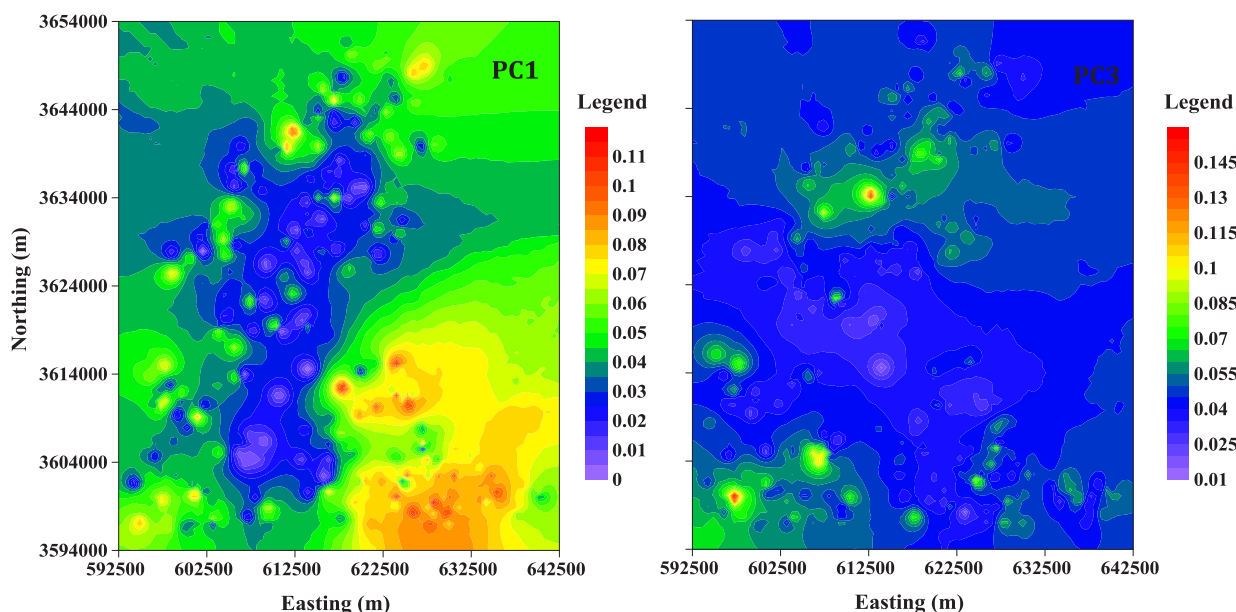
یکی دیگر از روش‌های مقایسه نتایج رتبه‌بندی، استفاده از تفاوت بین نتایج دو الگوریتم رتبه‌بندی است. در این روش

پوشش می‌دهد دارای بارهای PC بالا برای عناصری از قبیل Cd, Pb, Zn, Au, Ba و Mo است (شکل ۴). شکل ۶ نقشه کنتوری امتیازات مولفه اصلی دوم را نشان می‌دهد که شباهت زیادی با نقشه‌های به دست آمده از الگوریتم‌های رتبه‌بندی در شکل ۳ دارد. در شکل ۳-ب نقشه کلاسه‌بندی این امتیازات به روش فرکتالی عیار-مساحت برای مقایسه بهتر نشان داده شده است. مساحت محدوده‌های دارای امتیازات PC بالا و متوسط به ترتیب حدود ۴۴٫۴ و ۲۷۶٫۶ کیلومتر مربع است. در ادامه برای مقایسه شکل ۶ با نتایج به دست آمده در شکل ۳ از روش آنالیز تصویر استفاده می‌کنیم. برای این منظور شکل‌ها به تصویر رستری تبدیل شده و نوع کلاسه‌بندی هر سلول با

داده‌های این مولفه در تشخیص مناطق مستعد کانی‌زایی استفاده کرد. نقشه کنتوری امتیازات مولفه اول در شکل ۵ این نکته را نشان می‌دهد. به طوری که این نقشه نشان‌دهنده محدوده فعالیت‌های مرتبط با تشکیل سنگ (به ویژه سنگ‌های رسوبی) در منطقه مطالعاتی است. پایین بودن مقادیر بردارهای ویژه اکثر متغیرها (به استثناء تنگستن و مولیبدن) در PC3 (شکل ۴) و عدم همخوانی نقشه کنتوری امتیازات این فاکتور (شکل ۵) با نتایج الگوریتم‌های رتبه‌بندی نشان می‌دهد که PC3 بیشتر ممکن است نشان‌دهنده کانی‌زایی‌های مرتبط با سنگ‌های رسوبی-آذرین باشد. مولفه اصلی دوم که ۱۰ درصد تغییرپذیری داده‌ها را



شکل ۴: مقادیر ویژه سه مولفه اول به همراه درصد تغییرپذیری تجمعی مولفه‌ها



شکل ۵: نقشه کنتوری امتیازات مولفه اصلی اول و سوم در منطقه مطالعاتی

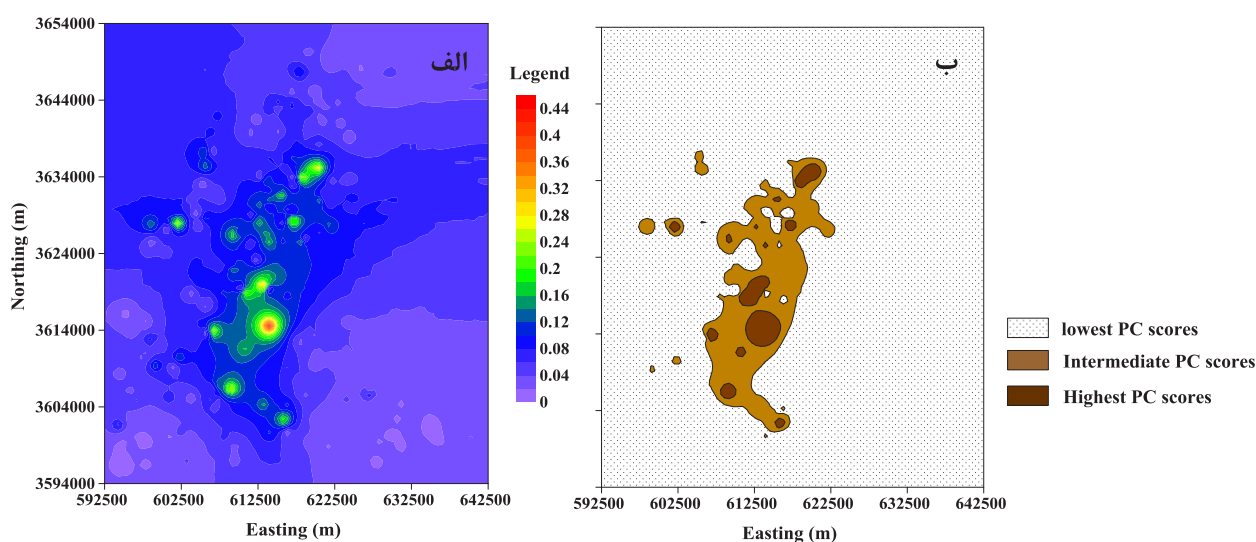
به الگوریتم‌های PA و CA است. همچنین بیشترین درصد پوششی این محدوده‌ها به وسیله واحد سنگی آندزیتی نیز در الگوریتم‌های VSA و CA دیده می‌شود. واحد سنگی ریولیتی بیشتر در محدوده‌های مستعد کانی‌زایی به دست آمده با الگوریتم‌های CA، PA و CAA خود را نشان می‌دهد. یکی دیگر از نکات جالب توجه نتایج رتبه‌بندی، قرارگیری این محدوده‌های مستعد پیشنهادی بر روی مناطق تکتونیزه شده به ویژه در مرکز منطقه مطالعاتی است.

مطالعات محققان نشان داده است که کانی‌زایی‌های سرب، روی و مس در منطقه مطالعاتی بیشتر به شکل رگه‌ای و پورفیری مرتبط با فعالیت‌های هیدروترمال انجام شده است [۳۶،۳۰]. این کانی‌زایی‌ها بیشتر در سنگ‌های آندزیتی و توف‌های آندزیتی متعلق به ائوسن پسین تا میانی دیده می‌شود. کانی‌های گالن، اسفالریت، کالکوپیریت و پیریت در بخش هیپوژن و کانی‌های سرروزیت، آنگلوژیت، کولیت، کالکوسیت، ملاکیت، آزرویت، کوپریت، هماتیت، گوتیت و باریت در بخش سوپرژن این محدوده‌ها قابل مشاهده است. همچنین سنگ‌های آتشفشانی ترشیری منطقه مطالعاتی با ترکیب حد

هم مقایسه می‌شود. نتایج این مقایسه در جدول ۵ نشان داده شده است. مطابق داده‌های این جدول بیشترین شباهت بین نقشه کنتوری امتیازات مولفه دوم (شکل ۶) با نتایج رتبه‌بندی متعلق به الگوریتم CA و بعد از آن به الگوریتم‌های PA و CAA است

۵-۲- جنبه زمین‌شناسی

مقایسه موقعیت محدوده‌های دارای رتبه‌بندی متوسط و بالا در شکل ۳ با نقشه زمین‌شناسی منطقه مطالعاتی (شکل ۱) نشان می‌دهد که این محدوده‌ها بر روی واحدهای توف، آندزیتی، ریولیتی و رسوبات کواترنری قرار گرفته‌اند. اگرچه بیشترین درصد مساحت این محدوده‌ها را رسوبات کواترنری می‌پوشاند که آن هم به دلیل گسترش قابل ملاحظه این واحد سنگی در محدوده مطالعاتی است، کمترین درصد پوششی مربوط به الگوریتم‌های PA و CA است که می‌تواند به عنوان یک نکته مثبت در نظر گرفته شود. سنگ‌های توف به لحاظ درصد تشکیل سطح این محدوده‌ها در جایگاه بعدی قرار دارد. بیشترین درصد پوششی این واحد سنگی متعلق

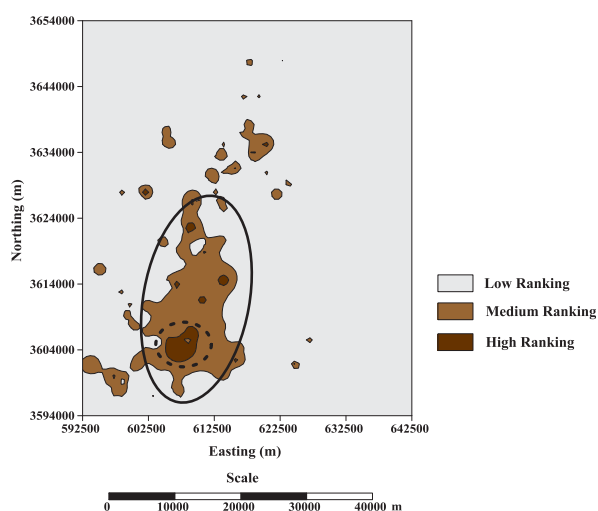


شکل ۶: نقشه کنتوری امتیازات مولفه اصلی دوم، (الف) با مقیاس پیوسته و (ب) با مقیاس کلاسه‌بندی شده در منطقه مطالعاتی

جدول ۵: نتایج آنالیز تصویر بین امتیازات مولفه اصلی دوم با تصاویر نتایج رتبه‌بندی

		CA	MA	VSA	PA	LA	CAA
درصد هم‌پوشانی		۴۷٫۴	۰	۱۹٫۳	۸	۱۷٫۴	۵۵٫۵
		۵۴٫۴	۳۴٫۴	۲۵٫۴	۴۸٫۵	۲۰	۴۵٫۹

شکل ۷ با اندیس‌های معدنی و اکتشافی محدوده مورد مطالعه می‌توان از روش نمودار نرخ پیش‌بینی- مساحت استفاده کرد [۴۴]. برای رسم این نمودار علاوه بر ۱۵ کانسار نشان داده شده در شکل ۱، از ۷۶ اندیس اکتشافی فلزی دیگری که در این محدوده به ثبت رسیده است، نیز استفاده شده است. شکل ۸ نمودار نرخ پیش‌بینی- مساحت و نقطه تقاطع دو منحنی را نشان می‌دهد. مطابق شکل، حدود ۶۰ درصد از

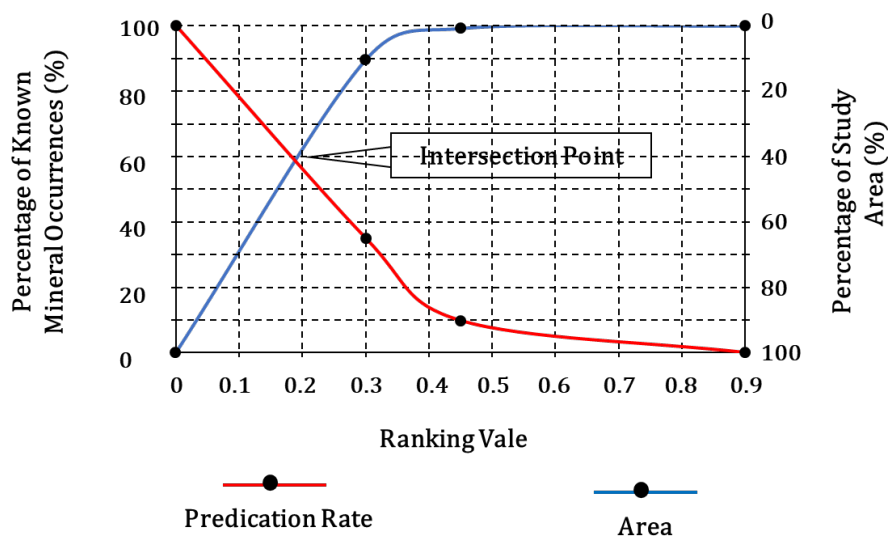


شکل ۷: نقشه محدوده‌های دارای پتانسیل کانی‌زایی در منطقه مطالعاتی به دست آمده از تلفیق نتایج الگوریتم‌های رتبه‌بندی (محدوده خط‌چین با اولویت بالا و محدوده خط‌پر با اولویت پایین‌تر اکتشافی)

واسط تا اسیدی از قبیل ریولیت‌ها نیز میزبان کانی‌زایی مس و پرلیت‌اند [۴۲]. پیش‌بینی وجود کانی‌زایی‌های طلا، نقره، مس، مولیبدن، سرب و روی در سنگ‌های آندزیتی تا بازالتی از دیگر شواهد کانساری در منطقه است [۴۳]. کانی‌زایی آهن و تیتانیم به ترتیب با ترکیب هماتیت و مگنتیت و ایلمنیت و رتیل نیز در رسوبات کواترنری به صورت پلاسری دیده شده است، بنابراین تشکیل محدوده‌های دارای رتبه‌بندی متوسط تا بالا در سنگ‌های یاد شده، ممکن است نشان‌دهنده شناسایی قابل قبول این محدوده‌ها به عنوان مناطق مستعد کانی‌زایی در منطقه مطالعاتی باشد.

با توجه به نتایج مطلوب به دست آمده از الگوریتم‌های رتبه‌بندی و برای تهیه نقشه نهایی پیشنهادی برای مناطق مستعد کانی‌زایی، از روش میانگین‌گیری برای تلفیق استفاده شده است. برای میانگین‌گیری به نتایج الگوریتم‌های MA، VSA، LA و CAA ضریب یک و به نتایج الگوریتم‌های PA و (به علت عملکرد بهتر کمی و کیفی این الگوریتم‌ها) داده شده است. شکل ۸ نقشه تلفیق را نشان می‌دهد. در این شکل دو منطقه یکی با اولویت بالاتر (محدوده خط‌چین در شکل ۷) و دیگری با اولویت پایین‌تر (محدوده خط‌پر در شکل ۷) برای اکتشاف بعدی پیشنهاد شده است. محدوده با پتانسیل بالاتر مساحتی حدود ۲۴ کیلومتر مربع و محدوده با پتانسیل پایین‌تر مساحتی در حدود ۳۱۱ کیلومتر مربع دارد.

برای تعیین ارتباط فضایی بین محدوده‌های پیشنهادی در



شکل ۸: نمودار نرخ پیش‌بینی- مساحت برای مدل به دست آمده از روش تلفیق الگوریتم‌های رتبه‌بندی

۷- سپاس‌گزاری

از سازمان زمین‌شناسی و اکتشافات معدنی کشور به دلیل امکان استفاده از داده‌های ژئوشیمیایی و نقشه‌های زمین‌شناسی و از سازمان صنعت، معدن و تجارت استان خراسان جنوبی به دلیل اطلاعات کانسارهای محدوده مطالعاتی تشکر و قدردانی می‌شود.

۸- مراجع

- [۱] حسنی‌پاک، ع.؛ ۱۳۹۴؛ "طراحی پروژه‌های اکتشافی (ژئوشیمیایی، ژئوفیزیکی و حفاری)". انتشارات دانشگاه تهران، ۴۷۹ صفحه.
- [۲] یوسفی، م.، کامکارروحانی، ا.؛ ۱۳۸۹؛ "اصول روش‌های مدل‌سازی پتانسیل معدنی در محیط سیستم اطلاعات جغرافیایی". جهاد دانشگاهی (دانشگاه صنعتی امیرکبیر)، ۲۲۶ صفحه.
- [۳] قدیری صوفی، ا.، یوسفی، م.؛ ۱۳۹۴؛ "ترکیب روش‌های فازی داده‌محور و دانش‌محور در مدل‌سازی پتانسیل معدنی به منظور تولید نواحی اهداف اکتشاف". مجله علوم زمین، شماره ۹۸، ص ۱۱-۱۸.
- [4] Awange, J. L., Paláncz, B., Lewis, R. H., and Völgyesi, L. (2018). "Mathematical geosciences; hybrid symbolic-numeric methods". Springer, pp. 615.
- [5] Alvo, M., and Yu, P. L. H. (2014). "Statistical methods for ranking data". Springer, New York, pp. 276.
- [6] Corain, L., Arboretti, R., and Bonnini, S. (2016). "Ranking of Multivariate Populations, A Permutation Approach with Applications". CRC Press, pp. 330.
- [7] Carranza, E. J. M. (2010). "Mapping of anomalies in continuous and discrete fields of stream sediment geochemical landscapes". *Geochemistry: Exploration, Environment, Analysis*, 10: 171-187.
- [8] Sadeghi, M., Morris, G. A., Carranza, E. J. M., Ladenberger, A., and Andersson, M. (2013). "Rare earth element distribution and mineralization in Sweden: An application of principle component analysis to FOREGS soil geochemistry". *Journal of Geochemical Exploration*, 133: 160-175.
- [9] Morrison, J. M., Goldhaber, M. B., Ellefsen, K. J., and Mills, C. T. (2011). "Cluster analysis of a regional-scale soil geochemical dataset in northern California". *Applied Geochemistry*, 26: 105-107.
- [10] Ellefsen, K. J., and Smith, D. B. (2016). "Manual hierarchical clustering of regional geochemical data using a Bayesian finite mixture model". *Applied Geochemistry*, 75: 200-210.

اندیس‌های معدنی و اکتشافی در حدود ۴۰ درصد مساحت منطقه، با محدوده‌های پیشنهادی الگوریتم‌های رتبه‌بندی (شکل ۷) پیش‌بینی شده‌اند. با توجه به اینکه، محدوده‌های پیشنهادی در مقیاس ناحیه‌ای است و همچنین این محدوده فقط به وسیله یک لایه اکتشافی (داده ژئوشیمیایی) به دست آمده است، می‌توان نتیجه گرفت که الگوریتم‌های رتبه‌بندی در شناسایی مناطق مستعد کانی‌زایی عملکرد مناسبی داشته‌اند.

۶- نتیجه‌گیری

ماهیت اکتشاف در مقیاس ناحیه‌ای و آنالیز نمونه‌های ژئوشیمیایی رسوبات آبراه‌های برای تعداد زیادی از عناصر، لزوم در نظر گرفتن کلیه اطلاعات برای پیشنهاد مناطق مستعد کانی‌زایی را دیکته می‌کند. در این مقاله روش رتبه‌بندی نمونه‌ها برای این منظور پیشنهاد گردید. رتبه‌بندی ۳۶۲ نمونه آنالیز شده برای ۲۳ عنصر به شش روش VSA، MA، CA، PA، LA و CAA نشان داد که:

- ۱- بزرگترین محدوده‌های پیشنهادی دارای پتانسیل کانی‌زایی به وسیله الگوریتم VSA و کوچکترین آنها با الگوریتم‌های PA و MA به دست آمده است.
- ۲- مقایسه محدوده‌های پیشنهادی با نقشه زمین‌شناسی نشان داد که این محدوده‌ها بر روی واحدهای توف، آندزیتی، ریولیتی و رسوبات کواترنری قرار گرفته‌اند.
- ۳- مقایسه کمی و کیفی نتایج الگوریتم‌های رتبه‌بندی نشان‌دهنده برتری نسبی الگوریتم‌های CA و PA بود.
- ۴- در نقشه تلفیق نهایی محدوده‌های با پتانسیل بالاتر و پایین‌تر به ترتیب با مساحت حدود ۲۴ و ۳۱۱ کیلومتر مربع برای فاز اکتشافی بعدی در منطقه مطالعاتی پیشنهاد گردید.
- ۵- نمودار نرخ پیش‌بینی- مساحت نشان می‌دهد که حدود ۶۰ درصد از اندیس‌های معدنی و اکتشافی در حدود ۴۰ درصد مساحت منطقه، به وسیله محدوده‌های پیشنهادی الگوریتم‌های رتبه‌بندی پیش‌بینی شده‌اند.
- ۶- الگوریتم‌های رتبه‌بندی به عنوان روشی مستقیم برای تلفیق داده‌های اکتشافی و همچنین تهیه نقشه‌های آنومالی چند عنصری پیشنهاد می‌شود.
- ۷- در تهیه نقشه رتبه‌بندی متغیرها (لایه‌های اطلاعاتی) وزن یکسان داشتند، در حالی که به کمک روش‌های دانش-محور و داده-محور می‌توان به هر لایه وزن داد و نقشه‌ای با دقت و صحت بالاتری به دست آورد.

- [23] Hosseini-Dinani, H., Aftabi, A., Esmaili, A., and Rabbani, M. (2015). "Composite soil-geochemical halos delineating carbonate-hosted zinc-lead-barium mineralization in the Irankuh district, Isfahan, west-central Iran". *Journal of Geochemical Exploration*, 156: 114-130.
- [24] Singh, J. N., and Dwivedi, S. K. (2012). "Analysis of Vector Space Model in Information Retrieval". National Conference on Communication Technologies & its impact on Next Generation Computing CTNGC 2012.
- [25] Larsen, J., Szymkowiak, A., and Hansen, L. K. (2001). "Probabilistic Hierarchical Clustering with Labeled and Unlabeled Data". *International Journal of Knowledge-Based Intelligent Engineering Systems*, 6: 56-62.
- [26] Brauer, S. (2014). "A Probabilistic Expectation Maximization Algorithm for Multivariate Laplacian Mixtures". MS Thesis of Paderborn University, pp. 78.
- [27] Botbol, J. M. (1971). "An application of characteristic analysis to mineral exploration: Proe. 9th Ins. Syra. on Techniques for Decision-Making in the Mineral Industry". Canadian Inst. of Mineral and Metallurgy, Montreal, Canada, Special 12: 92-99.
- [28] Pan, G., and Harris, D. P. (1992). "Decomposed and weighted characteristic analysis for the quantitative estimation of mineral resources". *Mathematical Geology*, 24(7): 807-823.
- [29] Bridges, N. J., Hanley, J. T., and McCammon, R. B. (1985). "PREPRO: A Computer program for encoding regional exploration data for use in characteristic analysis". *Mathematical Geology*, 11(5): 513-519.
- [30] Malekzadeh Shafaroudi, A., and Karimpour, M. H. (2015). "Mineralogic, fluid inclusion, and sulfur isotope evidence for the genesis of Sechangi lead-zinc (-copper) deposit, Eastern Iran". *Journal of African Earth Sciences*, 107: 1-14.
- [31] Asadi, S., and Kolahdani, S. (2014). "Tectono-magmatic evolution of the Lut block, eastern Iran: A model for spatial localization of porphyry Cu mineralization". *Journal of Novel Applied Sciences*, 3: 1058-1069.
- [32] Esmaili, D., Bouchez, J. L., and Siqueira, R. (2007). "Magnetic fabrics and microstructures of the Jurassic Shah-Kuh granite pluton (Lut Block, Eastern Iran) and geodynamic inference". *Tectonophysics* 439: 149-170.
- [33] Mahmoudi, S., Masoudi, F., Corfu, F., and Mehrabi, B. (2010). "Magmatic and metamorphic history of the Deh-Salm metamorphic Complex, Eastern Lut block, (Eastern Iran), from U-Pb geochronology". *International Journal of Earth Science*. 99: 1153-1165.
- [34] Mazhari, S. A., and Safari, M. (2013). "High-K Calc-
- [11] Fatehi, M., and 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.
- [12] Templ, M., Filzmoser, P., and Reimann, C. (2008). "Cluster analysis applied to regional geochemical data: Problems and possibilities". *Applied Geochemistry*, 23(8): 2198-2213.
- [۱۳] قزلباش، ر.، مقصودی، ع.؛ ۱۳۹۷؛ "استفاده از روش ترکیبی TOPSIS-AHP برای مدلسازی پتانسیل کانی‌زایی مس پورفیری در ورقه ورزقان، شمال باختر ایران". *مجله علوم زمین*، شماره ۱۰۹، ص ۳۳-۴۲.
- [۱۴] یوسفی، ت.؛ ۱۳۹۷؛ "تلفیق داده‌های زمین‌شناسی، ژئوشیمی و دورسنجی به منظور شناسایی مناطق امید بخش اکتشافی در ورقه ۱:۱۰۰۰۰۰ دهسلم با استفاده از روش‌های تصمیم‌گیری چند معیاره در محیط GIS". پایان‌نامه کارشناسی ارشد، دانشگاه بیرجند.
- [۱۵] قاسمی، ر.؛ ۱۳۹۷؛ "تولید نقشه‌های پیشگوی اکتشافی بوسیله روش‌های تصمیم‌گیری چندمعیاره مکانی فازی و ارائه مدل عدم قطعیت، مطالعه موردی: ذخایر ماسیوسولفاید بوانات". رساله دکتری، دانشگاه شاهرود.
- [16] Abedi, M., Mohammadi, R., Norouzi, G. H., and Mir Mohammadi, M. S. (2016). "A comprehensive VIKOR method for integration of various exploratory data in mineral potential mapping". *Arabian Journal of Geoscience*, 9: 482.
- [17] Büttcher, S., Clarke, C. L. A., and Cormack, G. V. (2016). "Information Retrieval: Implementing and Evaluating Search Engines (2nd edition)". MIT, USA, pp. 632.
- [۱۸] زارع بیدکی، ع. م.؛ ۱۳۸۸؛ "رتبه‌بندی و خزش مؤثر در وب". رساله دکتری، دانشگاه تهران.
- [19] Wu, L., Hsieh, C. J., and Sharpnack, J. (2018). "SQL-Rank: A Listwise Approach to Collaborative Ranking". Proceedings of the 35th International Conference on Machine Learning, Stockholm, Sweden, PMLR 80.
- [۲۰] قنبری، ا.، شاکری، آ.؛ ۱۳۹۴؛ "ارائه الگوریتمی مبتنی بر یادگیری جمعی به منظور یادگیری رتبه‌بندی در بازیابی اطلاعات". فصلنامه فناوری اطلاعات و ارتباطات ایران، شماره ۲۵ و ۲۶، ص ۶۷-۸۶.
- [21] Liu, T. Y. (2011). "Learning to Rank for Information Retrieval". Now Publisher Inc., pp. 285.
- [22] Beus, A. A., and Grigorian, S. V. (1977). "Geochemical Exploration Methods for Mineral Deposits". Published by Applied Publishing Ltd, pp. 276.

- of solutions in terms of preferability technique". Technological and Economic Development, 14: 224-239.
- [۴۲] یوسف‌زاده، م. ح؛ ۱۳۹۲؛ "پتروگرافی سنگ‌های آتشفشانی ترشیری منطقه خور - سه‌چنگی (غرب خوسف، استان خراسان جنوبی) با نگرشی بر پرلیت‌زایی". ششمین همایش انجمن زمین‌شناسی اقتصادی ایران، دانشگاه سیستان و بلوچستان.
- [۴۳] روحبخش، پ، کریم‌پور، م. ح، ملک‌زاده شفارودی، آ.؛ ۱۳۹۴؛ "بررسی سنگ‌نگاری، کانی‌شناسی منطقه‌های دگرسانی و ژئوشیمی منطقه کانی‌سازی مس - طلای سرچاه، شرق ایران". مجله بلورشناسی و کانی‌شناسی ایران، سال بیست و سوم، شماره چهارم، ص ۷۷۵-۷۸۸.
- [44] Yousefi, M., and Carranza, E. J. M. (2015). "Prediction-area (P-A) plot and C-A fractal analysis to classify and evaluate evidential maps for mineral prospectivity modeling". Computers & Geosciences, 79: 69-81.
- [35] Alavi, M. (1996). "Tectonostratigraphy synthesis and structural style of the Alborz Mountain system in northern Iran". Journal of Geodynamics, 21: 1-33.
- [36] Malekzadeh Shafaroudi, A., and Karimpour, M. H. (2013). "Hydrothermal alteration mapping in northern Khur, Iran, using ASTER image processing: a new insight to the type of copper mineralization". Acta Geological Sinica, 87(3): 830-842.
- [37] Filzmoser, P., Hron, K., Reimann, C., and Garrett, R. (2009). "Robust factor analysis for compositional data". Computers & Geosciences, 35: 1854-1861.
- [38] Filzmoser, P., Hron, K., and Reimann, C. (2009). "Principal component analysis for compositional data with outliers". Environmetrics, 20: 621-632.
- [39] Zhou, S., Zhou, K., Wang, J., Yang, G., and Wang, S. (2017). "Application of cluster analysis to geochemical compositional data for identifying ore-related geochemical anomalies". Frontiers of Earth Science, 12(3): 491-505.
- [40] Karimi, H., and Rezaeinia, A. (2011). "Adjusted permutation method for multiple attribute decision making with meta-heuristic solution approaches". International Journal of Industrial Engineering Computations, 2: 369-384.
- [41] Turskis, Z. (2008). "Multi-attribute contractors ranking method by applying ordering of feasible alternatives

¹ Analytic Hierarchy Process

² Vlse Kriterijumsk Optimizacija Kompromisno Resenje

³ Technique for Order Preference by Similarity to Ideal Solution

⁴ Cumulative and Multiplicative Algorithms

⁵ Vector Space Algorithm

⁶ Probability Algorithm

⁷ Linkage Algorithm

⁸ Characteristic Analysis Algorithm



Application of Ranking Algorithms for Mapping of Multi-element Geochemical Anomalies at the 1:100,000 Scale Sechangi Sheet in the Southern Khorasan Province

Geranian H.^{1*}

1- Assistant Professor, Dept. of Mining Engineering, Birjand University of Technology, Birjand, Iran
h.geranian@birjandut.ac.ir

(Received: 17 Dec. 2019, Accepted: 17 May 2020)

Abstract: The integration of exploration layers can be used to determine the multi-element geochemical anomalies that represent areas with mineralization potentiality. This is especially true for the combination of geochemical data. This paper aims to introduce ranking algorithms as an alternative to statistical data mining and multi-criteria decision-making methods for integrating exploration data. For this purpose, 362 stream sediment samples were used from Sechangi map sheet of South Khorasan Province. Each sample was analyzed for 23 elements. The implementation of six ranking algorithms including CA, MA, VSA, PA, LA, and CAA on the dataset shows that the obtained mineralization zones have overlapping and almost similar locations. These zones are located on tuff, andesitic, rhyolitic, and quaternary sedimentary rock units. Quantitative comparisons such as permutation and robust principal component analysis methods reveal the relative superiority of CA and PA algorithms. This is also the case with qualitative comparisons that deal with comparing the results of the ranking algorithms with geological mapping and mineral indices of the same study area. Combining six maps through weighted averaging suggested two zones. While the first zone has higher mineralization potentiality of about 24 km², the second zone shows lower mineralization potentiality of approximately 311 km² for the next exploration phase in the study area.

Keywords: Multi-elements anomaly, Ranking samples, Characteristic analysis, Mineral potential mapping, Sechangi sheet.

INTRODUCTION

Mineral exploration at regional scale is one of the most important substructure projects for identifying mineral potential mapping. The integration of exploration layers, especially the combination of geochemical data, can be used to determine the multi-element geochemical anomalies. These anomalies represent areas with potential mineralization [1-2]. The current study set to introduce ranking algorithms as an alternative to statistical data mining and multi-criteria decision-making methods for integrating exploration data. For this purpose, 362 stream sediment samples were selected from Sechangi map sheet of South Khorasan Province. The results presented two zones for the next exploration phase in the study area.



METHODS

The following six ranking algorithms were used to determine the mineral potential zones:

1. Cumulative Algorithm (CA)
2. Multiplicative Algorithm (MA)
3. Vector Space Algorithm (VSA)
4. Probability Algorithm (PA)
5. Linkage Algorithm (LA)
6. Characteristic Analysis Algorithm (CAA)

Then, the obtained mineral potential maps were compared with the classification maps of the PCs scores. The Robust PCA method was employed to obtain PCs scores [4]. All maps were classified based on the concentration-area fractal algorithm [3].

FINDINGS AND ARGUMENTS

The data set includes 362 stream sediment samples, from Sechangi map sheet of South Khorasan province. Each sample was analyzed for 23 elements (Figure 1). Figure 2 shows the mineral potential zones obtained by ranking algorithms. These maps were classified into three classes based on the concentration-area fractal algorithm. High-, medium-, and low-ranking zones represent high, medium, and low possibility mineralization areas, respectively. The obtained mineralization zones have overlapping and almost similar locations.

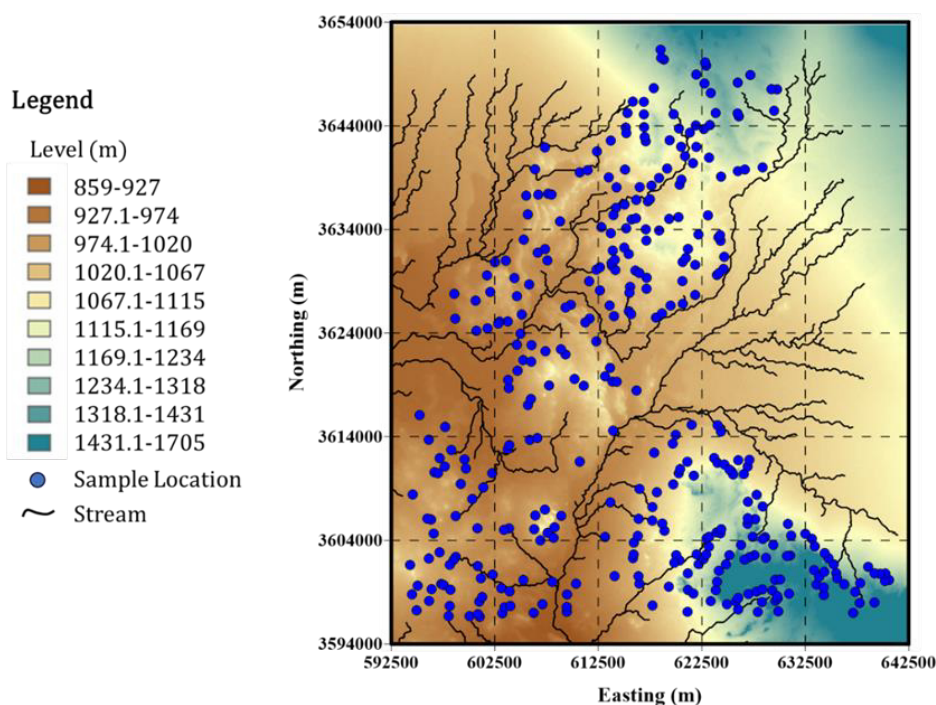


Figure 1. Locations of geochemical samples (blue dots) overlain on the topography map of the study area

The ranking algorithms maps were compared with the classification maps of the PCs scores. The Robust PCA was carried on \ln -transformed data set. Means and covariance matrices were back-transformed to clr space in accordance with the PCA for compositional data with outlier method [4]. Figure 3 shows the PC2 scores' contour map. It can be approximately equivalent to the results of the ranking algorithm maps. It is also possible to measure the similarity of images using image analysis (Table 1). The strongest similarity was observed between Figure 3 and the mineral potential map obtained from the CA algorithm.

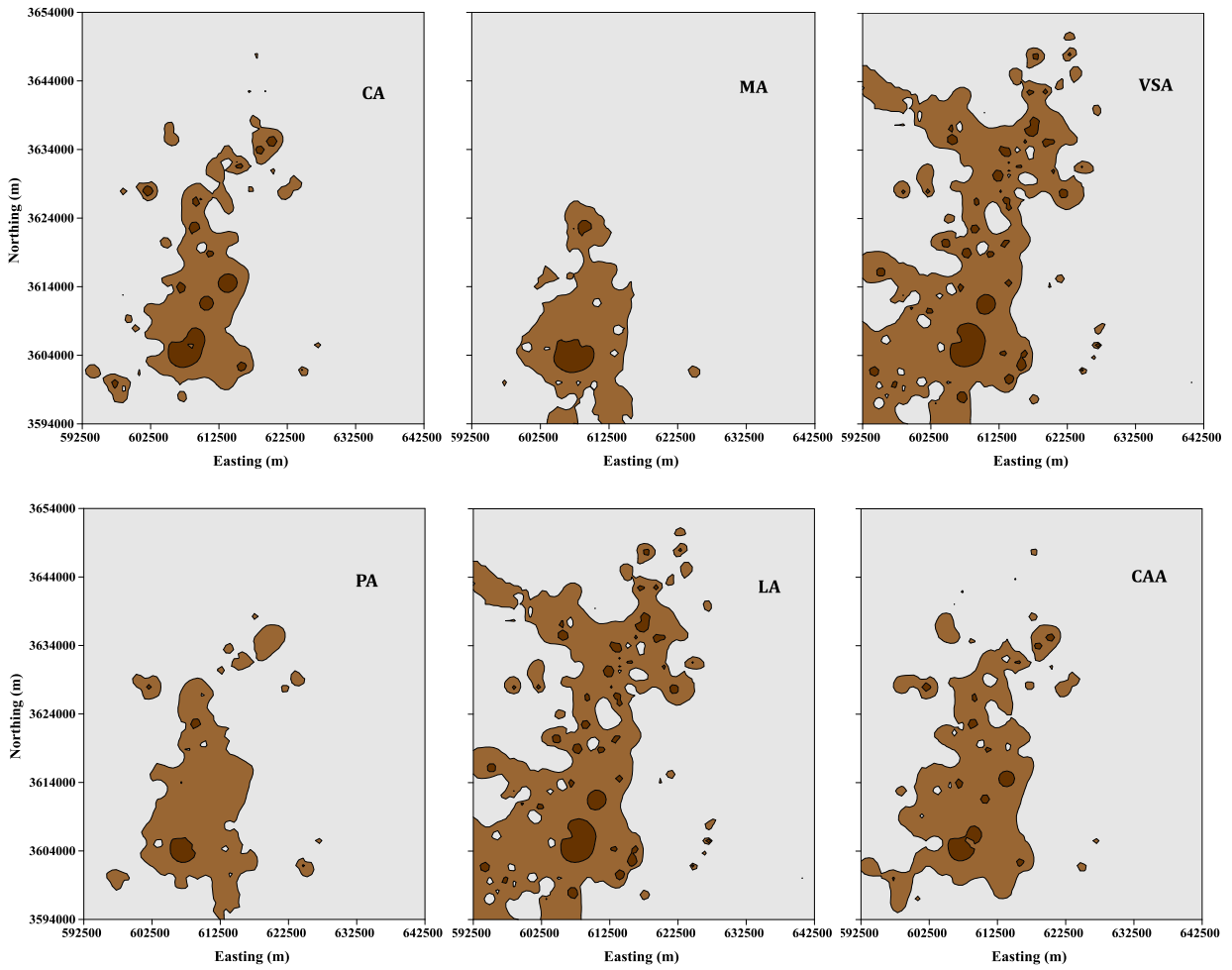


Figure 2. Classification maps of the ranking algorithms in the study area (■ high ranking, ■ medium ranking and ■ low ranking zones)

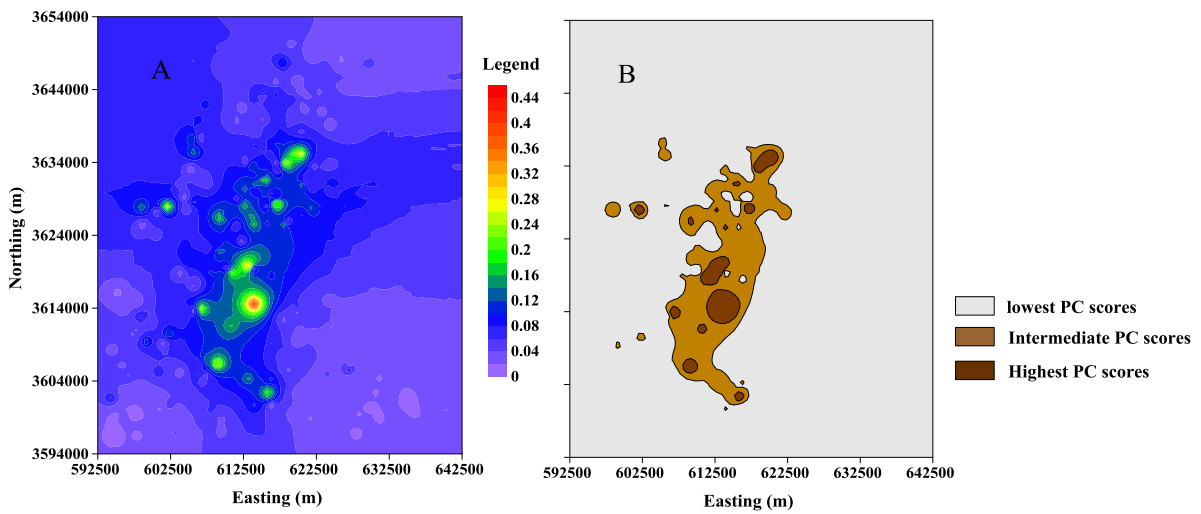


Figure 3. Contour map of the PC2 scores with continous scale A: and classified scale B: in the study area

Table 1. Overlapping results (%) between Figure 2 and Figure 3

	CA	MA	VSA	PA	LA	CAA
	47.4	0	19.3	8	17.4	55.5
	54.4	34.4	25.4	48.5	20	45.9

CONCLUSIONS

The obtained zones from ranking algorithms are located on tuff, andesitic, rhyolitic and quaternary sedimentary rock units. Quantitative comparisons such as permutation and robust principal component analysis methods demonstrated the relative superiority of CA and PA algorithms. This is also the case with qualitative comparisons that deal with comparing the results of the ranking algorithms with geological mapping and mineral indices of the same study area. Combining the six maps by weighted averaging presented two zones; the first zone had higher mineralization potentiality of about 24 km², but the second zone showed lower mineralization potentiality of approximately 311 km² for the next exploration phase in the study area. All six ranking algorithms provided strongly similar results which are also confirmed by PCA and rock units relating to hydrothermal processes. So, it is recommended to use the ranking methods as integration methods for identifying mineral potential mapping.

REFERENCES

- [1] Abedi, M., Mohammadi, R., Norouzi, G. H., and Mir Mohammadi, M. S. (2016). "A comprehensive VIKOR method for integration of various exploratory data in mineral potential mapping". *Arabian Journal of Geoscience*, 9: 482.
- [2] Awange, J. L., Paláncz, B., Lewis, R. H., and Völgyesi, L. (2018). "Mathematical geosciences; hybrid symbolic-numeric methods". Springer, pp. 615.
- [3] Zhou, S., Zhou, K., Wang, J., Yang, G., and Wang, S. (2017). "Application of cluster analysis to geochemical compositional data for identifying ore-related geochemical anomalies". *Frontiers of Earth Science*, 12(3): 491-505.
- [4] Filzmoser, P., Hron, K., and Reimann, C. (2009). "Principal component analysis for compositional data with outliers". *Environmetrics*, 20: 621-632.