٤-١- پردازش داده های خام

٤-١-١- محاسبه پارامترهای آماری داده‌های خام

اولین مرحلهٔ پردازش داده‌های ژئوشیمیایی، بررسی پارامترهای آماری مربوط به تک‌تک عناصر جهت شناخت ماهیت توزیع هریک از آنها می‌باشد که با محاسبهٔ پارامترهای آماری از قبیل میانگین، انحراف معیار، چولگی، کشیدگی، واریانس و . . . می‌توان به این موضوع دست‌یافت. در این قسمت برای هر عنصر به عنوان یک متغیر آماری در یک جدول، تعداد نمونه‌ها، حداقل و حداکثر عیار، میانگین، میانه، انحراف معیار، چولگی و کشیدگی و نمودارهای هیستوگرام توزیع فراوانی محاسبه و ترسیم شده‌اند. در جدول ٤-١ پارامترهای آماری برای تمام عناصر آنالیز شده در محدوده ساردوئیه آمده است.

٤ -١- ٢- بررسی مقادیر خارج از رده : (Outliers)

هنگام بررسی مقادیر داده‌های خام به نمونه‌هایی برخورد می‌شود که در آستانه‌های بالا و پایین جامعه داده‌ها قرار گرفته و از جامعه اصلی جدا افتاده‌اند. اگر نمودار جعبه‌ای (Boxplot) آنها ترسیم شود این نمونه‌ها به نحو بارزی خودشان را از بقیه جدا می‌کنند. مقادیر خارج از رده به سه حالت مختلف زیر ممکن است بوجود آیند:

حالت اول) از یک خطای سیستماتیک به هنگام نمونه‌برداری، آماده‌سازی یا تجزیه شیمیایی نمونه‌ها ناشی شده باشند که باید از مرحله پردازش حذف یا اصلاح شوند.

حالت دوم) مشاهداتی که به صورت یک پدیده فوق‌العاده نمود پیدا می‌کنند که باید پس از بررسی اعتبار آنها در مورد حفظ یا حذف آنها تصمیم گرفت.

حالت سوم) مشاهدات فوق‌العاده‌ای که هیچ‌گونه توضیح مناسبی برای آنها وجود ندارد و کارشناس اگر احساس کند که آنها به عنوان گوشه‌ای از جامعه مورد بررسی هستند می‌تواند آنها را حفظ کند.

وجود مقادیر خارج از رده در جامعه نمونه‌ها موجب افزایش واریانس جامعه و نیز همبستگی بین متغیرها و همچنین افزایش چولگی در نمودار توزیع عناصر می‌شود. برای کاهش این تأثیر راههای مختلفی نظیر محاسبه ضریب همبستگی با استفاده از روشهای ناپارامتری مانند روش اسپیرمن (Spearman)، حذف یا جایگزین نمودن مقادیر استفاده می‌شود در این گزارش از روش جایگزین نمودن مقادیر خارج از رده استفاده شده است. جدول (٤-٢) نمونه‌های دارای مقادیر خارج از رده را نشان می‌دهد.

٤-١-٣- نرمال سازی داده‌های خام :

استفاده از برخی روشهای آماری منوط به نرمال بودن تابع توزیع متغیرهای مورد مطالعه است در حالیکه توابع توزیع از نوع لاگ نرمال است، به همین علت قبل از استفاده از این روشها داده‌های خام باید نرمال شوند. در این بخش از روش تبدیل کاکس و باکس (Cox & Box) جهت نرمال کردن تابع توزیع داده‌های خام استفاده شده است. این کار شرط لازم کاربرد برخی روشهای آماری مانند تعیین نمونه‌های آنومالی با استفاده از اضافه کردن ضرایبی از انحراف معیار به حد آستانه‌ای و یا محاسبه ضرایب همبستگی پیرسون می‌باشد.

تبدیل کاکس و باکس یکی از تبدیلات مهمی است که بسیاری از توابع توزیع را می پوشاند، لذا تا حدود زیادی می تواند داده ها را به توزیع نرمال نزدیک کند. شکل کلی این تبدیل به صورت زیر است:

$$z=\begin{cases}\dfrac{x^{\lambda}-1}{\lambda} & (\lambda\neq 0), x>0\\ ln(x) & (\lambda=0), x>0\end{cases}$$

که در آن:

x : مقدار داده هایی که باید نرمال شوند.

λ : یک مقدار حقیقی

Z : مقدار تبدیل یافته

این تبدیل اکثراً داده ها را به توزیع نرمال نزدیک می کند، حتی اگر این تبدیل، داده ها را به توزیع نرمال نزدیک نکند، حداقل داده ها را منظم می کند.

برای تخمین مقدار λ می توان از روش حداقل چولگی استفاده کرد. در تبدیل کاکس و باکس یک مقدار ثابتی (λ) وجود دارد که باید طوری تعیین شود که توریع داده ها را تا حد ممکن به توزیع نرمال نزدیک کند. یعنی به تابع توزیعی نزدیک کند که چولگی آن صفر و کشیدگی آن ۳ باشد.

می توان تابعی مانند F را تعریف کرد که در آن اثر چولگی و کشیدگی تابع توزیع منعکس شده باشد :

$$F = W.|S| + |3 - K|$$

که در آن S چولگی، K کشیدگی و w وزن نسبی تأثیر چولگی در تابع F است. با مینیمم کردن این تابع می توان تا حد ممکن چولگی (به صفر) و کشیدگی (به ۳) را توأماً به حالت توزیع نرمال نزدیک کرد. بنابراین به ازاء مقادیر مختلف λ در توزیع کاکس و باکس می توان مقدار تابع F را محاسبه کرد. در جدول ٤-۳ پارامترهای آماری برای تمام عناصر نرمال شده در محدوده ساردوئیه آمده است.

٤-١-٤- بررسی های آماری تک متغیره

با توجه به این پارامترهای آماری می‌توان دریافت که مقادیر چولگی و کشیدگی متغیرها در مقایسه با مقادیر متناظر مربوط به داده‌های خام نرمال نشده تا چه اندازه کاهش یافته و منحنی توزیع تجمعی آنها به صورت یک خط راست که بیانگر توزیع نرمال می‌باشد، ظاهر شده است. هیستوگرام مقادیر نرمال شده نسبت به هیستوگرام مقادیر نرمال نشده نیز بیانگر مطلب فوق می‌باشد. همچنین از دیگر پارامترهای آماری تک متغیره می توان به میانگین، میانه، ماکسیمم ومینیمم جامعه آماری اشاره نمود. پارامترهای آماری و هیستوگرام‌های ترسیم شده برای داده‌های نرمال عناصر مهم مرتبط با منطقه اکتشافی در اشکال (٤-١) تا (٤-١٠) آورده شده است.

٤-١-٥- بررسی های آماری دو متغیره

هر تجزیه و تحلیل دو متغیره که بر روی دو متغیر انجام گیرد، می‌تواند در قالب آنالیزهای دو متغیره بیان شود. غالب تکنیکهای دو متغیره در اصل بسط و توسعه آنالیزهای تک متغیره می‌باشد.

٤-١-٦- تعیین ضریب همبستگی

برای تعیین اینکه آیا ارتباط معنی‌داری میان تغییرات متغیرهای آماری وجود دارد، ضرایب همبستگی میان آنها محاسبه می‌شود. این عمل به دو منظور کشف همبستگی بین متغیرها و تخمین مقدار یک یا چند متغیر دیگر صورت می‌گیرد. برای بررسی، دو نوع ضریب همبستگی پیرسون و اسپیرمن به صورت ماتریس ضرایب همبستگی محاسبه شده‌اند که در جداول (٤-٤) و (٤-٥) آمده‌است شرط محاسبه ضریب همبستگی پیرسون، نرمال بودن تابع توزیع متغیرها می‌باشد. دراین جداول، Sig(٢-Tailed) میزان معنی‌دار بودن ضرایب همبستگی طبق آزمون فرض مساوی صفر بودن ضریب همبستگی می‌باشد.

برای محاسبه ضریب همبستگی پیرسون به علت تأثیرپذیری این پارامتر از آستانه‌های بالا و پایین حتماً باید داده‌های خام نرمال شوند تا ضریب همبستگی محاسبه شوند. جدول (٤-٤) مقادیر این ضرایب را نشان می‌دهد.

همانطور که ملاحظه می شود بر پایه روش پیرسون ، بیشترین همبستگی مثبت بین عناصرلانتانیم و سریم به مقدار(٠/٩٦٦) می باشد. بعد از آن به ترتیب عناصر منگنز و اورانیم(٠/٩٠٧) ، کروم و

نیکل(۰/۸۹۰) وانادیم و آهن (۰/۸٤۷)، آهن و بیسموت(۰/۸۰۳) روی و منگنز (۰/۷٤۳) ، تلور وجیوه(۰/۷٦۱) و ، دارای بیشترین همبستگی های مثبت می باشند.

عنصر طلا (Au) ضرایب همبستگی متوسطی با عناصر داشته، بطوریکه بالاترین میزان همبستگی آن با عناصر مس، آنتیموان ، تنگستن و کبالت می باشد.

برای محاسبه ضریب همبستگی اسپیرمن از داده‌های خام استفاده شده است و همانطور که مشاهده می‌شود، در بعضی مواقع وضعیت متفاوتی نسبت به ضریب همبستگی پیرسون دارد. این اختلاف بیشتر زمانی بروز می‌کند که مقدار داده‌های خارج از رده زیاد باشد. اما مقایسه دقیق آنها، این نکته را بیان می‌کند که اختلاف این دو ضریب همبستگی خیلی زیاد نیست ، این امر نشان دهنده تاثیرپذیری کم داده‌ها از مقادیر خارج از رده است. جدول (٤-٥) مقادیر این ضرایب را نشان می‌دهد.

بر پایه این جدول ضریب همبستگی مشاهده شده بین عناصر در سطح اعتماد ۹۹٪ می‌باشد که بیشترین ارتباط همبستگی بین عناصر La,Ce (۰/۹۵۳) وجود دارد. ضریب همبستگی بین جفت متغیرها به روش پیرسون و اسپیرمن بیانگر اختلاف تقریبا کم بین ضرایب همبستگی عناصر متناظر می‌باشد که حکایت از توزیع نسبتاً نرمال عناصر و همین طور عدم تأثیر نمونه‌های دور افتاده دارد.

یکی دیگر از راههای بررسی ارتباط تغییرات عناصر با یکدیگر، رسم نمودار پراکنش (Scatter Plot)می‌باشد. زوج مرتب‌هایی از مقادیر دو متغیر که دارای توزیع دو متغیره یکسان باشند بر روی نمودار دو بعدی ترسیم می‌گردند. هر چه پراکندگی نقاط در نمودارهای پراکنش بیشتر باشد پیوند بین متغیرها ضعیف‌تر است. شکل (٤-۱۱) پراکنش مقادیر داده‌های خام نرمال

La,Ce بیشترین همبستگی را با یکدیگر نشان می‌دهد.

٤-١-٧- بررسی‌های آماری چند متغیره

هر تجزیه و تحلیل چند متغیره که بر روی بیش از دو متغیر انجام گیرد، می‌تواند در قالب آنالیزهای چند متغیره بیان شود. غالب تکنیکهای چند متغیره در اصل بسط و توسعه آنالیزهای تک متغیره می‌باشند و البته بعضی از روشهای چند متغیره تنها برای پاسخگویی به مقاصد چند متغیره طراحی شده‌اند که از جمله این روشها می‌توان به آنالیز فاکتوری اشاره کرد.

تجربه نشان داده است که چنانچه ترکیبی از متغیرها به جای یک متغیر به کار گرفته شوند و از نتایج ترکیبی آنها استفاده شود امکان تشخیص هاله‌های مرکب ژئوشیمیایی در اطراف توده‌های کانساری به مراتب افزایش می‌یابد و از طرفی اثرات خطاهای تصادفی در بکارگیری ترکیبی متغیرها نسبتاً کاهش می‌یابد. از دیگر مزایای استفاده از روشهای چند متغیره، کاهش تعداد متغیرها در مباحث داده‌پردازی و درنتیجه کاستن از تعداد نقشه می باشد.

دراین گزارش از روشهای چند متغیره مانند روشهای آنالیز خوشه‌ای و آنالیز فاکتوری و . . . استفاده شده است.

٤-١-٨- آنالیز خوشه‌ای و تفسیر آن

به دلیل اینکه هر گروه از عناصر نسبت به یکسری از شرایط محیطی کم و بیش به طور مشابه حساسیت نشان می‌دهند، شناخت ارتباط و همبستگی ژنتیکی متقابل بین عناصر مختلف می‌تواند در شناخت دقیق‌تر تغییرات موجود در محیطهای ژئوشیمیایی به کار گرفته شود. ضمناً تجمع ژنتیکی بعضی از عناصر ممکن است به عنوان راهنمای مستقیم در تفسیر نوع نهشته‌ای که احتمالاً در ناحیه وجود دارد، به کار رود. آنالیز خوشه‌ای یک روش آماری چند متغیره است که عناصر را

البته باید توجه داشت که آنالیز خوشه‌ای می‌تواند گروه‌های غیر قابل انتظاری را نیز ایجاد نماید که بیانگر روابط جدیدی خواهند بود و باید مورد بررسی قرار گیرند. در روش آنالیز خوشه‌ای از داده‌های خام نرمال (۲۷ عنصر اصلی) استفاده شده است تا اثر مقادیر غیر همساز از جامعه اصلی و نیز اثر تغییر مقیاس داده‌ها از میان برود. نتایج حاصل از آنالیز خوشه‌ای عناصر مورد مطالعه در شکل (٤-۱۲) آورده شده است. با توجه به شکل می‌توان چهار گروه اصلی را جدا نمود که بیانگر ارتباط پاراژنزی بین متغیرها باشد.

گروه اول: شامل عناصر Ce,La,As

گروه دوم: شامل عناصر Mo, S, Ba, Sn , Ni, Cr

گروه سوم: شامل عناصر Mn ,W, Zn, U,Cd,Ag

گروه چهارم: شامل عناصر Ti, V, Fe´,Sr, Te, Bi, Pb, Hg, Sb, Co, Cu,Au

حضور عناصر Te, Bi, Pb, Hg, Sb,Au با هم می تواند بیانگر حضور برخی از فعالیتهای هیدرو ترمالی ودر نتیجه تجمع این عناصر باشد.

حضور عناصر Ni و Cr ارتباط این عناصر را با واحدهای مافیک و التررا مافیک منطقه مشخص می کند.

٤-۲- پردازش داده های غنی شدگی

٤-۲-۱- محاسبه شاخص غنی‌شدگی و همگن‌سازی جوامع

پس از دسته‌بندی جوامع سنگی به منظور همگن‌سازی جوامع مختلف، مقدار زمینه محلی عناصر را در هر یک از جوامع سنگی محاسبه می‌گردد. بدین منظور از میانگین و یا میانه استفاده می‌شود. بدلیل اینکه میانگین خود متأثر از مقادیر حدی در تابع احتمال است و از طرفی توزیع اکثر عناصر، چولگی مثبت نشان می‌دهد، از مقوله میانه که مستقل از مقادیر می‌باشد استفاده شده است.

بنا به تعریف شاخص غنی‌شدگی یک عنصر خاص در یک نمونه معین عبارت است از نسبت غلظت آن عنصر در آن نمونه به غلظت میانگین یا میانه همان عنصر در جامعه‌ای که نمونه مربوطه متعلق به آن است.

شاخص غنی‌شدگی یک عنصر خاص در یک نمونه معین به مقدار غلظت آن عنصر در نمونه مربوطه و فراوانی همان عنصر در کل جامعه نمونه‌برداری بستگی دارد. بنابراین اگر فراوانی نقطه‌ای و منطقه‌ای یک عنصر هر دو با شیب ثابتی افزایش یا کاهش یابند آنچه که ثابت باقی خواهد ماند، شاخص غنی‌شدگی است. زیرا صورت و مخرج این کسر به یک نسبت افزایش و یا کاهش می‌یابند. بدین ترتیب شاخص غنی‌شدگی تا حدود زیادی مستقل از فاکتور لیتولوژی و یا مولفه سن‌ژنتیک فراوانی یک عنصر در ناحیه منشاء رسوبات آبراهه‌ای می‌باشد. بطور خلاصه می‌توان گفت شاخص غنی‌شدگی نشان دهنده نسبت غنی‌شدگی یا تهی‌شدگی یک عنصر در هر نمونه است. بدیهی است عناصری که مقدار شاخص غنی‌شدگی‌شان بیشتر از واحد باشد غنی شدگی و آنهایی که کمتر از واحد باشد تهی‌شدگی تلقی می‌شود.

شاخص غنی‌شدگی از رابطه زیر بدست می‌آید:

$$EI = \frac{Cj}{(Cmed)\,j}$$

در این رابطه EI شاخص غنی‌شدگی، Cj مقدار فراوانی عنصر j در نمونه معین و (cmed) j میانه مقادیر عنصر j درجامعه مربوط به آن نمونه می‌باشد. پس از جایگزینی مقادیر شاخص غنی‌شدگی به جای داده‌های خام یک جامعه کلی حاصل می‌شود که آن را جامعه شاخص غنی‌شدگی می‌نامند.

٤-٢-٢- محاسبه پارامترهای آماری شاخص غنی‌شدگی

حال با تشکیل جامعه شاخص غنی‌شدگی و محاسبه پارامترهای آماری و رسم هیستوگرام تجمعی فراوانی این داده‌ها و مقایسه آنها با محاسبات وهیستوگرامهای خام به نظر می‌رسد که اثرات ناهمگنی که به صورت جوامع آماری مختلف در هیستوگرام بروز کرده بود تا اندازه‌ای از بین رفته و شکل تابع توزیع همگن‌تر شده است، ولی همچنان حالت لاگ نرمال در شکل تابع توزیع مقادیر مشاهده می‌شود.(البته در این برگه محاسبه پارامترهای آماری بیشتر برای عناصر اصلی انجام گرفته است). در جدول ٤-٦ پارامترهای آماری برای شاخص غنی‌شدگی تمام عناصر آنالیز شده در محدوده ساردوئیه آمده است.

٤-۲-۳- نرمال سازی شاخص‌های غنی‌شدگی:

استفاده از برخی روشهای آماری منوط به نرمال بودن تابع توزیع متغیرهای مورد مطالعه است در حالیکه توابع توزیع از نوع لاگ نرمال است، به همین علت قبل از استفاده از این روشها داده‌های شاخص‌های غنی‌شدگی باید نرمال شوند. در این بخش از روش تبدیل کاکس و باکس (Cox & Box) جهت نرمال کردن تابع توزیع داده‌های شاخص‌های غنی‌شدگی استفاده شده است. این کار شرط لازم کاربرد برخی روشهای آماری مانند تعیین نمونه‌های آنومالی با استفاده از اضافه کردن ضرایبی از انحراف معیار به حد آستانه‌ای و یا محاسبه ضرایب همبستگی پیرسون می‌باشد.

تبدیل کاکس و باکس یکی از تبدیلات مهمی است که بسیاری از توابع توزیع را می پوشاند، لذا تا حدود زیادی می تواند داده ها را به توزیع نرمال نزدیک کند. شکل کلی این تبدیل به صورت زیر است:

$$z=\begin{cases}\dfrac{x^{\lambda}-1}{\lambda} & (\lambda\neq 0), x>0\\ ln(x) & (\lambda=0), x>0\end{cases}$$

که در آن:

x : مقدار داده هایی که باید نرمال شوند.

λ : یک مقدار حقیقی

Z : مقدار تبدیل یافته

این تبدیل اکثراً داده ها را به توزیع نرمال نزدیک می کند، حتی اگر این تبدیل، داده ها را به توزیع نرمال نزدیک نکند، حداقل داده ها را منظم می کند.

برای تخمین مقدار λ می توان از روش حداقل چولگی استفاده کرد. در تبدیل کاکس و باکس یک مقدار ثابتی (λ) وجود دارد که باید طوری تعیین شود که توریع داده ها را تا حد ممکن به توزیع نرمال نزدیک کند. یعنی به تابع توزیعی نزدیک کند که چولگی آن صفر و کشیدگی آن ۳ باشد.

می توان تابعی مانند F را تعریف کرد که در آن اثر چولگی و کشیدگی تابع توزیع منعکس شده باشد :

$$F = W.|S| + |3 - K|$$

که در آن S چولگی، K کشیدگی و w وزن نسبی تأثیر چولگی در تابع F است. با مینیمم کردن این تابع می توان تا حد ممکن چولگی (به صفر) و کشیدگی (به ۳) را توأماً به حالت توزیع نرمال نزدیک کرد. بنابراین به ازاء مقادیر مختلف λ در توزیع کاکس و باکس می توان مقدار تابع F را محاسبه کرد. در جدول ٤-۷ پارامترهای آماری برای تمام داده‌های شاخص‌های غنی‌شدگی عناصر نرمال شده در محدوده ساردوئیه آمده است.

٤-٢-٤- بررسی های آماری تک متغیره

با توجه به این پارامترهای آماری می‌توان دریافت که مقادیر چولگی و کشیدگی متغیرها در مقایسه با مقادیر متناظر مربوط به داده‌های شاخص‌های غنی‌شدگی نرمال نشده تا چه اندازه کاهش یافته و منحنی توزیع تجمعی آنها به صورت یک خط راست که بیانگر توزیع نرمال می‌باشد، ظاهر شده است. هیستوگرام مقادیر نرمال شده نسبت به هیستوگرام مقادیر نرمال نشده نیز بیانگر مطلب فوق می‌باشد. همچنین از دیگر پارامترهای آماری تک متغیره می توان به میانگین، میانه، ماکسیمم ومینیمم جامعه آماری اشاره نمود. پارامترهای آماری و هیستوگرام‌های ترسیم شده برای داده‌های نرمال عناصر مهم مرتبط با منطقه اکتشافی در اشکال (٤-١٣) تا (٤-٢٢) آورده شده است.

٤-٢-٥- بررسی های آماری دو متغیره

٤-٢-٦- تعیین ضریب همبستگی

برای تعیین اینکه آیا ارتباط معنی‌داری میان تغییرات متغیرهای آماری وجود دارد، ضرایب همبستگی میان آنها محاسبه می‌شود. این عمل به دو منظور کشف همبستگی بین متغیرها و تخمین مقدار یک یا چند متغیر دیگر صورت می‌گیرد. برای بررسی، دو نوع ضریب همبستگی اسپیرمن و پیرسون به صورت ماتریس ضرایب همبستگی محاسبه شده‌اند که در جداول (٤-٨) و (٤-٩) آمده است. شرط محاسبه ضریب همبستگی پیرسون، نرمال بودن تابع توزیع متغیرها می‌باشد. دراین جداول، Sig(٢-Tailed) میزان معنی‌دار بودن ضرایب همبستگی طبق آزمون فرض مساوی صفر بودن ضریب همبستگی می‌باشد.

برای محاسبه ضریب همبستگی پیرسون به علت تأثیرپذیری این پارامتر از آستانه‌های بالا و پایین حتماً باید داده‌های شاخص غنی‌شدگی نرمال شوند تا ضریب همبستگی محاسبه شوند. جدول (٤-٨) مقادیر این ضرایب را نشان می‌دهد.

همانطور که ملاحظه می شود بر پایه روش پیرسون ، بیشترین همبستگی مثبت بین عناصرلانتانیم و سریم به مقدار(۰/۹٦٤) می باشد. بعد از آن به ترتیب عناصر منگنز و اورانیم(۰/۹۰۰) ، کروم و نیکل(۰/۸۸۱) وانادیم و آهن (۰/۸۳۳)، آهن و بیسموت(۰/۸۱۷) روی و منگنز (۰/۷۳٦) ، تلور وجیوه(۰/۷۳۵) و ، دارای بیشترین همبستگی های مثبت می باشند.

عنصر طلا (Au) ضرایب همبستگی متوسطی با عناصر داشته، بطوریکه بالاترین میزان همبستگی آن با عناصر مس، آنتیموان ، تنگستن و کبالت می باشد.

برای محاسبه ضریب همبستگی اسپیرمن از داده‌های شاخص غنی‌شدگی استفاده شده است و همانطور که مشاهده می‌شود، در بعضی مواقع وضعیت متفاوتی نسبت به ضریب همبستگی پیرسون دارد. این اختلاف بیشتر زمانی بروز می‌کند که مقدار داده‌های خارج از رده زیاد باشد. اما مقایسه دقیق آنها، این نکته را بیان می‌کند که اختلاف این دو ضریب همبستگی خیلی زیاد نیست ، این امر نشان دهنده تاثیرپذیری کم داده‌ها از مقادیر خارج از رده است. جدول (٤-٩) مقادیر این ضرایب را نشان می‌دهد.

بر پایه این جدول ضریب همبستگی مشاهده شده بین عناصر در سطح اعتماد ۹۹٪ می‌باشد که بیشترین ارتباط همبستگی بین عناصر La,Ce(۰/۹۵۱) وجود دارد. ضریب همبستگی بین جفت متغیرها به روش پیرسون و اسپیرمن بیانگر اختلاف تقریبا کم بین ضرایب همبستگی عناصر متناظر می‌باشد که حکایت از توزیع نسبتاً نرمال عناصر و همین طور عدم تأثیر نمونه‌های دور افتاده دارد.

یکی دیگر از راههای بررسی ارتباط تغییرات عناصر با یکدیگر، رسم نمودار پراکنش (Scatter Plot) می‌باشد. زوج مرتب‌هایی از مقادیر دو متغیر که دارای توزیع دو متغیره یکسان باشند بر روی نمودار دو بعدی ترسیم می‌گردند. هر چه پراکندگی نقاط در نمودارهای پراکنش بیشتر باشد پیوند بین متغیرها ضعیف‌تر است. شکل (٤-۲۳) پراکنش مقادیر داده‌های داده‌های شاخص غنی‌شدگی نرمال شده برای چند زوج عنصری است که بیشترین ارتباط را نشان می‌دهد. دراین نمودارها زوج عنصر La,Ce بیشترین همبستگی را با یکدیگر نشان می‌دهد.

٤-٢-٧- بررسی‌های آماری چند متغیره

هر تجزیه و تحلیل چند متغیره که بر روی بیش از دو متغیر انجام گیرد، می‌تواند در قالب آنالیزهای چند متغیره بیان شود. غالب تکنیک‌های چند متغیره در اصل بسط و توسعه آنالیزهای تک متغیره می‌باشند و البته بعضی از روشهای چند متغیره تنها برای پاسخگویی به مقاصد چند متغیره طراحی شده‌اند که از جمله این روشها می‌توان به آنالیز فاکتوری اشاره کرد. تجربه نشان داده است که چنانچه ترکیبی از متغیرها به جای یک متغیر به کار گرفته شوند و از نتایج ترکیبی آنها استفاده شود امکان تشخیص هاله‌های مرکب ژئوشیمیایی در اطراف توده‌های کانساری به مراتب افزایش می‌یابد و از طرفی اثرات خطاهای تصادفی در بکارگیری ترکیبی متغیرها نسبتاً کاهش می‌یابد. از دیگر مزایای استفاده از روشهای چند متغیره، کاهش تعداد متغیرها در مباحث داده‌پردازی و درنتیجه کاستن از تعداد نقشه‌هاست. با استفاده از این روشها امکان مقایسه متغیرها و کسب نتایج راحت‌تر خواهد بود. البته استفاده بهینه از روشهای چند متغیره در حالتی صادق خواهد بود که در پردازش داده‌ها با تعداد زیادی متغیر روبرو باشیم و تا حدودی امکان اخذ نتیجه از متغیرها به گونه منفرد غیر ممکن و یا توام با خطای زیاد باشد. دراین گزارش از روشهای چند متغیره مانند روشهای آنالیز خوشه‌ای و آنالیز فاکتوری استفاده شده است.

٤-٢-٨- آنالیز خوشه‌ای و تفسیر آن:

به دلیل اینکه هر گروه از عناصر نسبت به یکسری از شرایط محیطی کم و بیش به طور مشابه حساسیت نشان می‌دهند، شناخت ارتباط و همبستگی ژنتیکی متقابل بین عناصر مختلف می‌تواند در شناخت دقیق‌تر تغییرات موجود در محیطهای ژئوشیمیایی به کار گرفته شود. ضمناً تجمع

آنالیز خوشه‌ای یک روش آماری چند متغیره است که عناصر را بر اساس شباهت تغییرپذیری بین آنها در قالب دسته‌ها یا گروههایی طبقه‌بندی می‌کند. دلایل زیادی برای ارزشمند بودن آنالیز خوشه‌ای وجود دارد، از جمله اینکه آنالیز خوشه‌ای می‌تواند در یافتن گروههای واقعی کمک کند و همچنین باعث کاهش تراکم داده‌ها شود. البته باید توجه داشت که آنالیز خوشه‌ای می‌تواند گروههای غیر قابل انتظاری را نیز ایجاد نماید که بیانگر روابط جدیدی خواهند بود و باید مورد بررسی قرار گیرند. در روش آنالیز خوشه‌ای از داده‌های شاخص غنی شدگی نرمال شده استفاده شده است تا اثر مقادیر غیر همساز از جامعه اصلی و نیز اثر تغییر مقیاس داده‌ها از میان برود. نتایج حاصل از آنالیز خوشه‌ای عناصر مورد مطالعه در شکل (٤-٢٤) آورده شده است. با توجه به شکل می‌توان سه گروه اصلی را جدا نمود که بیانگر ارتباط پاراژنزی بین متغیرها باشد.

گروه اول: شامل عناصر Ce,La,As

گروه دوم: شامل عناصر Mo, S, Ba, Sn , Ni, Cr

گروه سوم: شامل عناصر Mn ,W, Zn, U,Cd,Ag,Hg,Pb,Cu

گروه چهارم: شامل عناصر Ti, V, Fe´,Sr, Te, Bi, Sb, Co, Au

با مقایسه نتایج حاصل از آنالیز خوشه‌ای داده های خام و شاخص غنی شدگی مشاهده می شود تغییرات زیادی ندارند و آن به دلیل تنوع کم واحدهای لیتولوژیکی در منطقه مورد مطالعه می باشد.

٤-۲-۹- آنالیز فاکتوری

آنالیز آماری نیز یک روش دیگر برای بررسی و مطالعه همزمان تغییرات متغیرهای مورد بررسی در یک نقطه و انعکاس نحوه تغییرات آنها و در نتیجه روشی برای کاهش تعداد متغیرهای مورد بررسی است. به این ترتیب که بر اساس مدل خاصی به‌نام فاکتور ارتباط پیچیده بین متغیرها تعیین می‌گردد. آنالیز فاکتوری شامل محاسبه ماتریس ضرایب همبستگی بین متغیرها، تعیین متغیرهایی که به نظر می‌رسد وابستگی ضعیفی باسایر متغیرها دارند (با استخراج فاکتورها)، تعیین تعداد فاکتورها و روش محاسبه آنها و بالاخره دوران و اعمال تبدیلاتی خاص بر روی فاکتورها می‌باشد. مهمترین مسئله در آنالیز فاکتوری اصل بیان همبستگی بین مقادیر غلظت عناصر به منظور نمایش الگوی تغییرات همزمان آنها در یک مکان است. بدین منظور در جهت کاستن از تعداد داده‌ها از آنالیز فاکتوری استفاده گردیده است. هدف از به کار گیری آنالیز فاکتوری عبارت است از :

۱) تشخیص و تعیین فاکتورها (تجزیه)

۲) تعیین سهم نسبتی هر یک از فاکتورها در بوجود آمدن تغییرات توزیع عناصر در واقع هدف از تجزیه و تحلیل فاکتوری تشخیص اصلی‌ترین متغیرهای کنترل شده از متغیرهایی با نقش کمتر است. در این صورت می‌توان با حداقل تعداد متغیرهای فاکتوری، حداکثر تغییرپذیری بین داده‌ها را توجیه کرد و سهم نسبی هر یک از متغیرهای فاکتوری را در توجیه تغییرپذیری مشخص نمود.

به تجربه ثابت شده است که آنالیز فاکتوری تفکیک مناسبی برای کاهش داده‌ها در اکتشافات ژئوشیمیایی است به طوری که با استفاده از امتیازات فاکتوری به جای متغیرهای اولیه می‌توان مشاهدات صحرائی و کل تمرکز آنومالیها را تغییر داد.

بدین منظور ابتدا باید میزان اعتبار آنالیز فاکتوری بر روی مقادیر شاخص غنی‌شدگی نرمال بررسی شود. در این راه از آزمونهای KMO , Bartlett بهره‌گرفته می‌شود. هر چه مقدار KMO به عدد یک نزدیکتر باشد، دلالت بر تأیید بیشتر آنالیز فاکتوری دارد (به طور استاندارد KMO باید از (۰/٦ بیشتر باشد) که با توجه به جدول (٤-۱۰) مقدار KMO معادل ۰/۷۷۲ حد مناسبی می‌باشد که آنجام آنالیز فاکتوری را تأیید می‌نماید.

همچنین عدم رد آزمون کرویت که به آزمون فرض ماتریس واحد بودن ماتریس ضرایب همبستگی اشاره می‌کند. به این معنی است که کلیه متغیرها مستقل از یکدیگر عمل می‌کنند. با توجه به جدول (٤-۱۰) عناصر Pb, Zn, Hg, Fe, Te, Ni, Mn, Cr از بیشترین ضرایب برخوردار بوده و بیشترین مشارکت را در این روش دارا می‌باشند.

در آنالیز فاکتوری به روش مولفه‌های اصلی (PCA) ، برآورد ماتریس ضرایب همبستگی بدست می‌آید. با محاسبه مقادیر ویژه این ماتریس مقادیر بزرگتر از یک جدا شده و برای آنها بردارهای ویژه محاسبه می‌گردد. در جدولی که تحت عنوان Total Variance Explained آمده است. مقادیر ویژه، درصد واریانس و درصد تجمعی واریانس متناظر با عوامل، محاسبه شده و سپس مقادیر بزرگتر از یک استخراج و دوران داده شده‌اند که با توجه به جداول (٤-۱۱) و(٤-۱۲) و(٤-۱۳) بیشترین تغییرپذیری محیط مربوط به مولفه‌های اول و دوم به ترتیب ۲٦/۳۸۲ و ۱٦/۸٤۳ می‌باشد. نمودار مقادیر ویژه که بر حسب اهمیت از بزرگترین تا کوچکترین مقادیر ردیف شده‌اند (Scree Plot) درشکل (٤-۲۵) آورده شده‌است.

از آنجا که اغلب یک یا چند عامل ویژه چند متغیره را کنترل می‌کنند، روشهایی بوجود آمده‌اند که بدون تغییر میزان اشتراک تفسیر عوامل را ساده‌تر می‌سازند. این روشها همان دوران عوامل هستند

Varimax

که دوران متعامد است بر روی ضرایب عاملی دوران صورت می‌گیرد. مقادیری با قدر مطلق نسبتاً بزرگ و یا صفر به ستونهای ماتریس ضرایب عاملها اختصاص یافتند. در نتیجه عواملی ایجاد شده‌اند که یا شدیداً به متغیرها وابسته‌اند یا مستقل از آنها هستند و سبب ساده شدن تفسیر عاملها می‌گردند. با استفاده از این روش می‌توان عناصری را که در هر عامل از اهمیت بیشتری برخوردارند تعیین کرد. با توجه به این جداول ۷ فاکتور جدا شده‌است.

فاکتور اول: این فاکتور بیشتر تحت تأثیر عناصر Ti, V, Fe, Te, Bi, Sb,Co می‌باشد. همراهی Fe و CoوV وTi امری عادی محسوب می‌شود ومی‌تواند نشانه عادی بودن وضعیت ژئوشیمیایی این عناصر باشد.که در ارتباط با واحدهای مافیکی و الترا مافیکی منطقه است.

فاکتور دوم: این فاکتور تحت تأثیر عناصر U, Zn, Mn, W, Agمی‌باشد. که در ارتباط با واحدهای اسیدی و حدواسط منطقه است.

فاکتور سوم: این فاکتور تحت تأثیر عناصر Ba, Sn می‌باشد.که حضور این عناصر با هم نشانگر این است که این عناصر توانایی قرارگیری در فاکتور های دیگر را نداشته و نمی توانند تشکیل گروه هایی با رفتارهای مشابه بدهند.

فاکتور چهارم: این فاکتور تحت تأثیر عناصر Ni و Cr می باشد که همراهی Ni و Cr امری عادی محسوب می‌شود که در ارتباط با واحدهای مافیکی و الترا مافیکی منطقه است.

فاکتور پنجم: این فاکتور تحت تأثیر عنصر Cu, S, Moمی‌باشد.حضور گوگرد در این فاکتور می تواند در ارتباط با آلتراسیون های منطقه باشدو همچنین حضور مس و مولیبدن و گوگرد می تواند موید وجود نهشته های پورفیری در منطقه باشد.

فاکتور ششم: این فاکتور تحت تأثیر عنصر Pb, Hgمی‌باشد.این فاکتور نشان دهنده فعالیت فاز هیدروترمالی هستند که خود می‌تواند یکی از ردیابهای عنصر طلا در منطقه باشند

فاکتور هفتم: این فاکتور تحت تأثیر عنصر AS, Cd می‌باشد.این فاکتور نشان دهنده فعالیت فاز هیدروترمالی هستند که خود می‌تواند یکی از ردیابهای عنصر طلا در منطقه باشند