



بررسی تأثیر مخرب نمونه خارج از رده در مطالعات آماری ژئوشیمیایی چند متغیره مطالعه موردی: برگه ۱:۵۰۰۰۰ مغانجق در شمال غربی ایران

یوسف قنبری*^۱، امیر حبیب‌نیا^۲، ایوب معمار^۱

۱- دفتر اکتشاف و استخراج، شرکت تولید مواد اولیه و سوخت هسته‌ای، سازمان انرژی اتمی ایران، صندوق پستی: ۱۳۳۹-۱۴۱۵۵، تهران- ایران

۲- دانشکده مهندسی معدن و متالورژی، دانشگاه صنعتی امیرکبیر، صندوق پستی: ۱۵۹۱۴، تهران- ایران

چکیده: در بررسی‌های اکتشافی ژئوشیمیایی رسوب‌های آبراهه‌ای منطقه مغانجق واقع در شمال غربی ایران تعداد ۱۵۲ نمونه رسوب آبراهه‌ای از منطقه برداشت شد. پس از بررسی مقادیر عناصر مختلف در این نمونه‌ها، نمونه‌ای مشخص شد که مقدار عناصر Co، As، Eu، Yb، Sc، Ni، Li در آن نسبت به مقدار همین عناصر در سایر نمونه‌ها بسیار بیشتر بود. پس از معرفی این نمونه به عنوان نمونه خارج از رده سعی شده است تأثیر این نمونه در بررسی‌های آماری چند متغیره بررسی شود. هدف از این مقاله نمایش گوشه‌ای از اثرهای مخرب نمونه‌های خارج از رده بر مطالعات آماری ژئوشیمیایی چند متغیره می‌باشد که باعث بروز خطای سیستماتیک در تفسیرها و تحلیل‌ها می‌شوند. جهت انجام مطالعات آماری چند متغیره، از ضریب همبستگی با روش‌های اسپیرمن و پیرسون و آنالیز خوشه‌ای و نمودار پراکنندگی عناصر به همراه معادله خط رگرسیون آنها در دو حالت استفاده شده است. در حالت اول نمونه خارج از رده در بررسی‌های چند متغیره حضور دارد و در حالت دوم نمونه خارج از رده از بررسی‌ها کنار گذاشته شده است، سپس نتایج حاصل از دو بررسی با هم مقایسه شده‌اند. پس از بررسی مطالعات آماری چند متغیره و مقایسه نتایج در دو حالت ذکر شده مشخص شد که با حضور نمونه خارج از رده در بین مجموعه از نمونه‌های برداشت شده از منطقه، ارتباط بین عناصر ممکن است به صورت‌های مختلف زیر باشد.

- ارتباط حقیقی بین دو عنصر در نمونه‌های برداشت شده وقتی که هیچکدام از آن دو عنصر دارای فراوانی غیرعادی در نمونه‌ها نباشند.

- ارتباط کاذب بین دو عنصری که یکی از آن دو دارای فراوانی غیرعادی در نمونه‌ها باشد.

- ارتباط شدیداً کاذب بین دو عنصر وقتی که هر دو دارای فراوانی غیرعادی در نمونه‌ها باشند.

واژه‌های کلیدی: نمونه خارج از رده، مقادیر خارج از رده، مطالعات آماری ژئوشیمیایی چند متغیره

The Outlier Sample Effects on Multivariate Statistical Data Processing in Geochemical Stream Sediment Survey (Moghangeh Region, NW of Iran)

Y. Ghanbari*¹, A. Habibnia², A. Memar¹

1- Deputy of Exploration, Exploration and Preparation of Raw Material of the Nuclear Industry Company, AEOL, P.O. Box: 14155-1339, Tehran-Iran
2- Mining and Metallurgy Engineering Department, Amirkabir University of Technology, P.O. Box: 15914, Tehran-Iran

Abstract: In geochemical stream sediment surveys in Moghangeh Region in north west of Iran, sheet 1:50,000, 152 samples were collected and after the analyze and processing of data, it revealed that Yb, Sc, Ni, Li, Eu, Cd, Co, as contents in one sample is far higher than other samples. After detecting this sample as an outlier sample, the effect of this sample on multivariate statistical data processing for destructive effects of outlier sample in geochemical exploration was investigated. Pearson and Spearman correlation coefficient methods and cluster analysis were used for multivariate studies and the scatter plot of some elements together the regression profiles are given in case of 152 and 151 samples and the results are compared. After investigation of multivariate statistical data processing results, it was realized that results of existence of outlier samples may appear as the following relations between elements:

- true relation between two elements, which have no outlier frequency in the outlier sample.
- false relation between two elements which one of them has outlier frequency in the outlier sample.
- complete false relation between two elements which both have outlier frequency in the outlier sample.

Keywords: Outlier Sample, Multivariate Statistical Data Processing, Geochemical Stream Sediment Survey

*email: ysf_ghn12@yahoo.com

تاریخ دریافت مقاله: ۸۶/۲/۲۴ تاریخ پذیرش مقاله: ۸۷/۴/۲۲



۱- مقدمه

هدف اصلی از بررسی‌های اکتشافی کوچک مقیاس به طور کلی تعیین نواحی امیدبخش برای مراحل بعدی اکتشاف می‌باشد. روش ژئوشیمیایی رسوب‌های آبراهه‌ای در تعیین مناطق امیدبخش در ایران روشی مفید جهت پیجویی است. به طور کلی در پروژه‌های اکتشافی ژئوشیمیایی رسوب‌های آبراهه‌ای در مقیاس ناحیه‌ای، چگالی نمونه‌برداری پیشنهادی جهت طراحی اولیه، بر مبنای یک نمونه آبراهه برای هر ۲/۵ کیلومتر مربع می‌باشد که در منطقه مورد مطالعه (برگه ۱:۵۰۰۰۰ مغانجق) نیز بدین روش نمونه‌برداری شده است. منطقه مغانجق در شمال غربی ایران و دارای مناطق غیرجنگلی است؛ آب و هوای منطقه از نوع سرد و کوهستانی می‌باشد و وسعت هوازدگی در منطقه زیاد است. مجموعاً ۱۵۲ نمونه از منطقه مورد مطالعه برداشت شد که پس از طی مراحل مختلف آماده‌سازی، نمونه‌های برداشت‌شده جهت انجام آنالیز شیمیایی به آزمایشگاه سازمان زمین‌شناسی کشور ارسال گردیده سپس توسط دستگاه ICP، مورد تجزیه و تحلیل قرار گرفتند. روش‌های آماری جهت تعیین آنومالی‌ها به دو زیر روش تقسیم می‌شوند: روش‌های آماری تک متغیره و روش‌های آماری چند متغیره. در روش آماری تک متغیره جهت تعیین آنومالی‌ها فقط یک متغیر (عنصر) وجود دارد ولی در روش‌های آماری چند متغیره باید به ارتباط عناصر با یکدیگر پی برد و به عناصری اهمیت داد که بتوانند بیشترین تغییرپذیری (بالای ۰/۷۵) از کل جامعه را نشان دهند. تجربه نشان داده است که اگر ترکیبی از مقادیر یک گروه از عناصر معرف به جای مقدار یک عنصر خاص به کار گرفته شوند، هاله‌های ژئوشیمیایی در اطراف توده‌های کانساری بهتر مشخص می‌گردند و اثر خطاهای تصادفی در آنها کمتر می‌شود [۱]. شناخت بستگی‌های ژنتیکی که در بین عناصر وجود دارد، اطلاعات لازم را جهت تفسیر هرچه صحیح‌تر داده‌های ژئوشیمیایی در اختیار می‌گذارد. لذا آمار چند متغیره می‌تواند جوابگوی مسایل فوق باشد. در روش‌های چند متغیره وجود نمونه‌های خارج از رده باعث بروز خطای سیستماتیک و ایجاد یک ارتباط کاذب بین عناصر می‌گردد، بنابراین شناخت این نوع نمونه‌ها در بررسی‌های چند متغیره بسیار حائز اهمیت است. در این مقاله پس از معرفی مقادیر خارج از رده و انواع آن، نمونه‌ای که نسبت به ۸ عنصر دارای

مقادیر خارج از رده می‌باشد معرفی شده، سپس تأثیر آن بر سایر نمونه‌ها در مطالعات آماری چندمتغیره، بررسی شده است. جهت انجام مطالعات آماری چند متغیره از روش ضریب همبستگی پیرسون و آنالیز خوشه‌ای استفاده شده است، سپس نمودار پراکندگی برخی از عناصر به همراه معادله خط رگرسیون آنها در دو حالت ۱۵۲ نمونه و ۱۵۱ نمونه آورده شده است.

۲- مقادیر خارج از رده

مقادیر خارج از رده همواره نمی‌توانند به صورت گروه جداگانه، یا حاوی ویژگی‌های مفید و مسأله‌ساز تلقی شوند. اما باید در قالب تجزیه و تحلیل به آنها نگریت و ارزیابی آنها بر اساس اطلاعات گوناگون صورت گیرد (Rousseeuw, 1987) مقادیر خارج از رده ممکن است متأثر از سه حالت زیر باشند:

حالت اول: این گونه مقادیر خارج از رده، ممکن است از یک خطای نظام‌دار نشأت گرفته باشند و در هنگام پردازش داده‌ها باید آنها را شناسایی کرده و از فرایند داده‌پردازی حذف و یا اصلاح گردند.

حالت دوم: دسته دوم مقادیر خارج از رده مشاهداتی هستند که به صورت یک پدیده غیرعادی نمود می‌یابند و کارشناس باید توجه داشته باشد که آیا این مقادیر، معتبرند یا خیر؟

حالت سوم: این گروه به مشاهدات فوق‌العاده‌ای تعلق می‌گیرد که کارشناس هیچگونه توضیح مناسبی برای آنها ندارد گرچه به نظر می‌رسد که این دسته از مقادیر خارج از رده بایستی حذف شوند، اما اگر کارشناس احساس کند که آنها به عنوان گوشه‌ای از جامعه مورد بررسی قرار می‌گیرند می‌تواند آنها را حفظ کند. آنومالی‌هایی که در این گروه قرار دارند هیچ ارتباطی با شواهد زمین‌شناسی ندارند و ممکن است ناشی از آلودگی‌های شیمیایی، صنعتی، کشاورزی و یا پدیده‌های خاص زمین‌شناسی باشند [۱].

در داده‌های مربوط به منطقه مغانجق در بین ۱۵۲ نمونه برداشت شده، یک نمونه وجود دارد که نسبت به بسیاری از عناصر دارای مقدار خارج از رده می‌باشد و مقادیر عناصر As, Co, Cd, Eu, Li, Ni, Sc, Yb در این نمونه نسبت به مقدار همین عناصر در سایر نمونه‌ها بسیار بالاتر است. این نمونه خارج



همچنین تجمع ژنتیکی بعضی از عناصر ممکن است به عنوان راهنمای مستقیمی در تفسیر نوع نهشته‌ای که می‌تواند در ناحیه‌ای وجود داشته باشد، به کار رود [۱]. در این مقاله برای تشخیص ارتباط ژنتیکی بین عناصر از جدول ضرایب همبستگی عناصر مختلف با یکدیگر و نمودار خوشه‌ای عناصر نسبت به هم استفاده شده است. در جدول همبستگی، ارتباط میان هر عنصر با عناصر دیگر به صورت عددی بین -۱ تا +۱ نشان داده می‌شود. در جدول محاسبه شده علاوه بر عدد مربوط به ضریب همبستگی، سطح اعتماد اعداد ذکر شده توسط نرم‌افزار SPSS اعلام می‌شود. معمولاً سطح اعتماد ۹۵٪ برای کارهای مهندسی مناسب است. بر این اساس در صورتی که نرم‌افزار سطح اعتماد کمتری را برای برخی ضرایب نشان دهد به آن اعداد اهمیتی نداده و جهت تفسیر و بررسی جدول‌ها استفاده نمی‌شود [۱]. در جدول ۱ ضرایب همبستگی بین عناصر در دو حالت ۱۵۲ نمونه (A) و ۱۵۱ نمونه (B) با دو روش اسپیرمن (I) و پیرسون (II) تعیین شده است. در حالت (A) نمونه خارج از رده وجود دارد ولی در حالت (B) نمونه خارج از رده کنار گذاشته شده است. خانه‌های هاشور خورده در هر جدول بیانگر بالا بودن سطح اعتماد و اهمیت دادن به اعداد به دست آمده در آن خانه‌ها است.

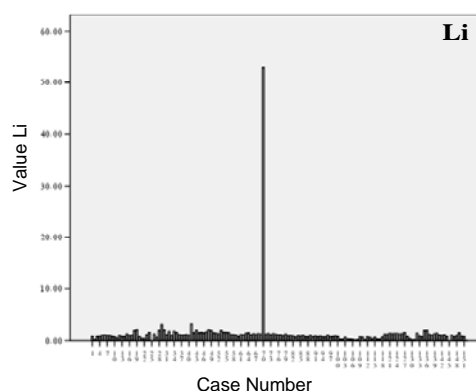
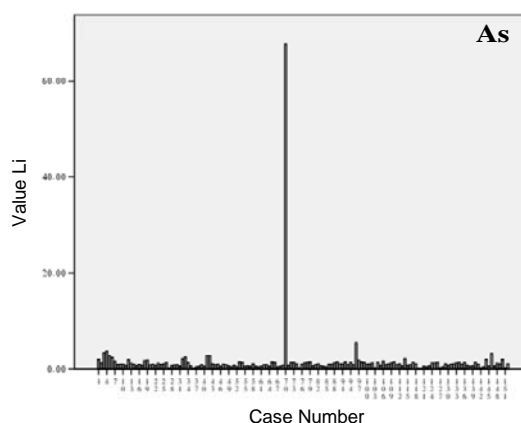
۴- رگرسیون

در تعیین رابطه بین متغیرها از ضریب همبستگی، معادله خط رگرسیون، روش آنالیز خوشه‌ای، روش آنالیز فاکتوری و برخی روش‌های دیگر استفاده می‌شود. ضریب همبستگی رابطه بین دو متغیر را نشان می‌دهد که هر دو تحت تأثیر عوامل مشترک، تغییرپذیری هم‌جهت و یا غیر هم‌جهت از خود بروز می‌دهند. برای تعیین رابطه علت و معلول بین تغییرات دو متغیر تصادفی که یکی از آنها به عنوان تابع و دیگری به عنوان متغیر می‌باشند از روش رگرسیون استفاده می‌شود [۳]. روش‌های مختلفی جهت تعیین معادله رگرسیون و برازش خط به نقاط وجود دارد که از آن جمله می‌توان روش‌های کمترین مربعات معمولی^(۱)، کمترین مربعات وزن‌دار^(۲)، کمترین انحراف مطلق^(۳)، کمترین مربعات نرمال^(۴)، مجموع میانه بازماندهای وزن‌دار^(۵)، کمترین مربعات میانه^(۶) را نام برد [۴]. در این

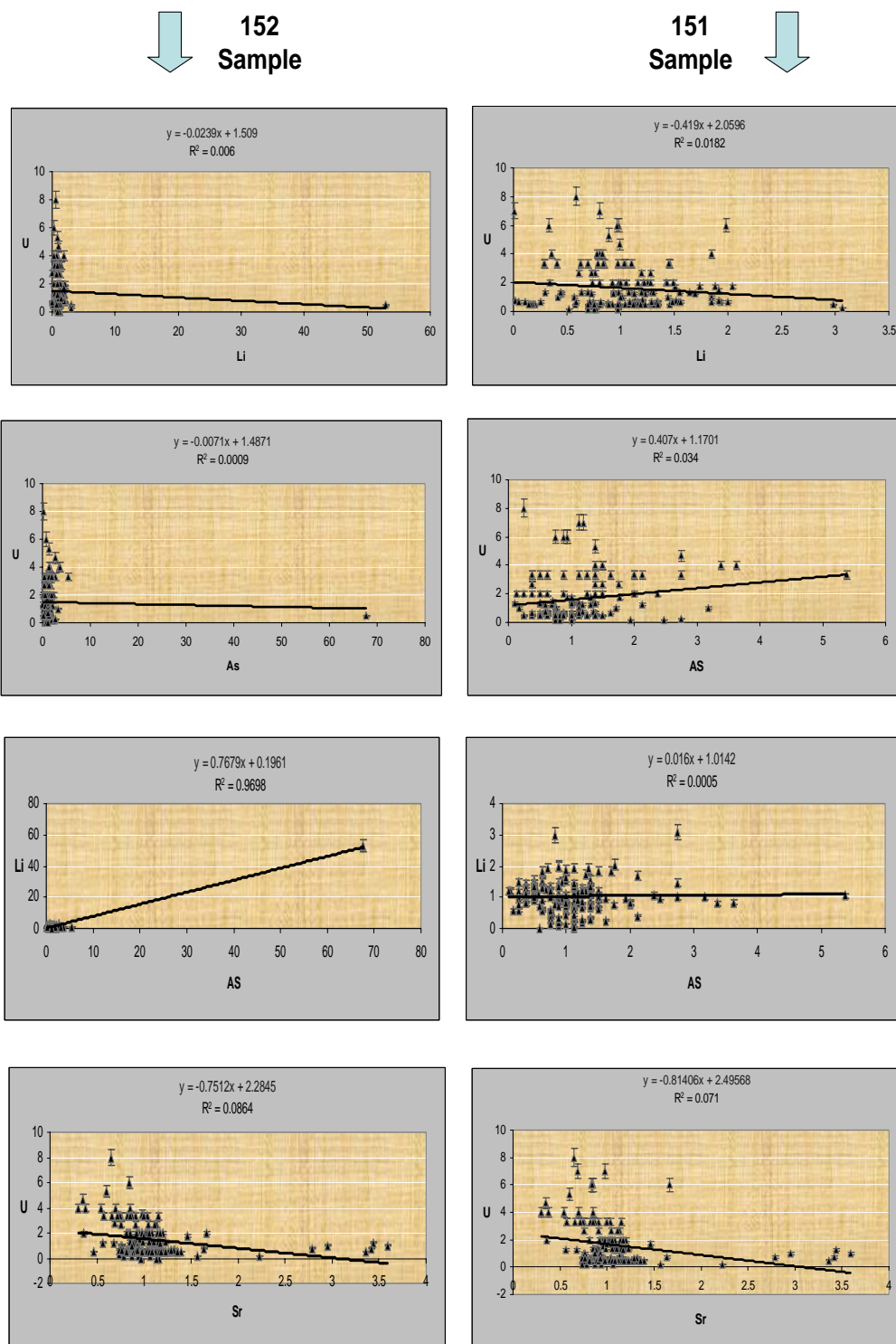
از رده از نوع اول تلقی می‌گردد و هیچ توضیح مناسبی نمی‌توان برای آن در نظر گرفت. در شکل ۱ نمودار فراوانی دو عنصر آورده شده است. نمودارهای میله‌ای فراوانی عناصر Li و As نشان‌دهنده فراوانی غیرعادی این عناصر در نمونه موردنظر می‌باشند. محور افقی مشخص‌کننده شماره نمونه‌ها است و محور قائم میزان مقادیر عناصر را در هر نمونه نشان می‌دهد.

۳- ضریب همبستگی

هر گروه معین از عناصر نسبت به یک سری از شرایط محیطی، حساسیت کم و بیش مشابهی نشان می‌دهد [۲]. شناخت ارتباط و وابستگی‌های ژنتیکی متقابل موجود بین عناصر مختلف ممکن است در شناخت دقیق‌تر تغییرات موجود در محیط‌های ژئوشیمیایی به کار گرفته شود،



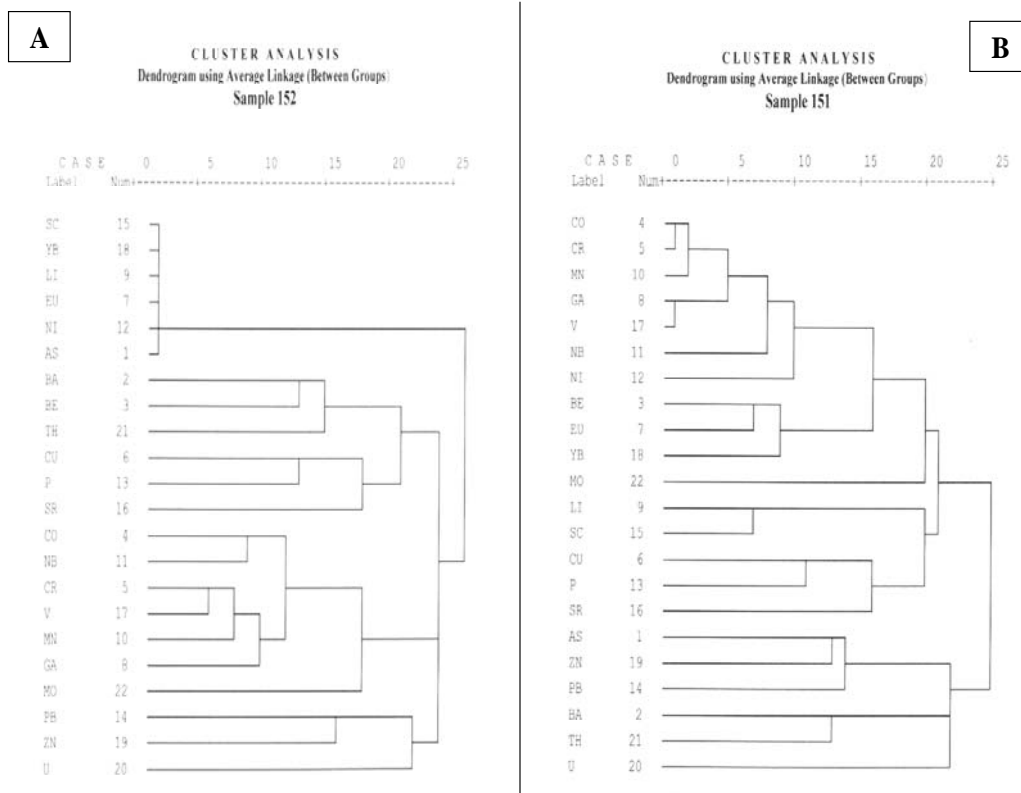
شکل ۱ - نمودار میله‌ای فراوانی عناصر Li و As که نشان‌دهنده فراوانی غیرعادی عناصر Li و As در یک نمونه از نمونه‌های برداشت شده هستند (محور افقی مشخص‌کننده شماره نمونه‌ها و محور قائم مشخص‌کننده میزان مقادیر عناصر در نمونه‌ها).



شکل ۲- نمودارهای پراکندگی برخی از عناصر نسبت به هم و خط رگرسیون برازش شده به آنها (روش کمترین مربعات معمولی).

استفاده شده است که شکل ۳ نمودار آنالیز خوشه‌ای عناصر نسبت به یکدیگر در دو حالت ۱۵۲ نمونه (A) و ۱۵۱ نمونه (B) را نشان می‌دهد. در حالت (A) نمونه خارج از رده وجود دارد ولی در حالت (B) نمونه خارج از رده کنار گذاشته شد.

کامل، اتصال به مرکز استفاده می‌شود و برای محاسبه فاصله بین داده‌ها، روشهای همبستگی پیرسون، فاصله اقلیدسی ساده، فاصله اقلیدسی توان دوم از مهم‌ترین روش‌ها می‌باشند [۶]. در این مقاله از روش فاصله اقلیدسی ساده جهت ترسیم نمودار آنالیز خوشه‌ای



شکل ۳- نمودار آنالیز خوشه‌ای عناصر نسبت به یکدیگر در دو حالت ۱۵۲ نمونه (A) و ۱۵۱ نمونه (B).

اعدادی در یک جدول بسیار نادر است. در جدول (A)، ۱۵ ضریب بین عناصر وجود دارد که بزرگتر از $0/98$ می‌باشند و این بیانگر یک ارتباط کاذب بین عناصر می‌باشد که با حضور نمونه خارج از رده در بررسی‌های آماری به وجود آمده است. در جدول B (روش پیرسون) ضرایب همبستگی به دست آمده بین عناصر پس از حذف نمونه خارج از رده تفاوت زیادی با جدول A دارد به طوری که اعداد واقع در جدول B (روش پیرسون) شباهت زیادی با اعداد بدست آمده از روش اسپیرمن دارند و عددی که بزرگتر از $0/98$ باشد در هیچکدام از جدول‌ها وجود ندارد. برای تعیین رابطه علت و معلول بین تغییرات دو متغیر تصادفی که یکی از آنها به عنوان متغیر مستقل و دیگری به عنوان متغیر وابسته شناخته می‌شود از روش رگرسیون استفاده شده است.

در این مقاله نمودار پراکنش عناصر Li، As، Sr، Cd نسبت به عنصر U به همراه معادله خط رگرسیون مربوطه آورده شده است. در معادله رگرسیون As نسبت به اورانیوم، شیب خط معادله با حضور نمونه خارج از رده $-0/407$ است ولی پس از

۶- بحث و نتیجه‌گیری

برای تشخیص ارتباط ژنتیکی بین عناصر از جدول ضرایب همبستگی عناصر مختلف نسبت به یکدیگر استفاده شده است. در جدول همبستگی، ارتباط میان هر عنصر با سایر عناصر به صورت عددی بین -1 تا $+1$ نشان داده می‌شود. علامت مثبت بیانگر ارتباط مستقیم بین دو عنصر و علامت منفی بیانگر ارتباط معکوس بین دو عنصر می‌باشد. هرچه قدر مطلق ضریب به دست آمده بین دو عنصر به عدد ۱ نزدیک باشد بیانگر ارتباط قوی (مستقیم یا معکوس) بین آن دو عنصر می‌باشد. تمام اعداد واقع در خانه‌های هاشور خورده در سطح اعتماد 95% حساب شده‌اند. در روش اسپیرمن که به توزیع داده‌ها بستگی ندارد در جدول‌های A و B تفاوت چندانی بین ضرایب همبستگی عناصر وجود ندارد ولی در روش پیرسون حضور یک نمونه خارج از رده باعث از بین بردن همبستگی واقعی بین عناصر می‌شود. به عنوان مثال در جدول A (روش پیرسون) ضریب همبستگی بین عناصر Li-Sc، Ni-Yb، Yb-Sc و Eu-Yb به ترتیب برابر ۱، ۱، $0/99$ و $0/99$ می‌باشد که این ضرایب بسیار بزرگ بوده و احتمال حضور چنین



فلدسپار و حضور As در این گروه بیانگر وجود ترکیبات گوگردی (سولفیدی) در منطقه است و معمولاً در اندیس‌های هیدروترمال علاوه بر وجود کانی‌های Zn و Pb، کانی باریت و فلدسپارها و ترکیبات سولفیدی نیز یافت می‌شود. در دندروگرام A که نمونه خارج از رده در ترسیم این نمودار حضور دارد، عناصر Sc، Yb، Ni، As، Eu، Li در یک گروه قرار دارند به طوری که این عناصر تأثیر یکسانی بر یکدیگر و بر کل عناصر اندازه‌گیری شده در منطقه دارند و هیچ فرایند زمین‌شناسی نمی‌تواند در یک منطقه چنین ارتباطی را بین عناصر توجیه کند؛ در صورتی که در نمودار B عناصر Sc، Yb، Ni، As، Eu، Li در گروه‌های مختلف قرار دارند و وابستگی‌های مختلفی با یکدیگر و سایر عناصر اندازه‌گیری شده در منطقه نشان می‌دهند.

۷- نتیجه‌گیری

با توجه به بررسی‌های انجام گرفته مشاهده می‌شود نمونه‌هایی که نسبت به چندین عنصر دارای فراوانی غیرعادی باشند باعث ایجاد ارتباط کاذب بین عناصر می‌شوند و حضور این نوع نمونه‌ها باعث بروز خطای سیستماتیک در بررسی‌های آماری ژئوشیمیایی رسوبات آبراه‌های چندمتغیره می‌گردد. با توجه به مقایسه جدول‌ها همبستگی عناصر با دو روش اسپیرمن و پیرسون مشخص شد روش اسپیرمن، روش قدرتمندتری است و مستقل از توزیع داده‌ها می‌باشد به طوری که وجود نمونه خارج از رده تأثیر زیادی در نتایج به‌دست آمده با این روش نخواهد داشت.

به طور کلی با حضور نمونه‌های خارج از رده در بین مجموعه‌ای از نمونه‌های برداشت شده از یک منطقه، ارتباط بین عناصر به صورت‌های مختلف زیر می‌باشد:

- ارتباط حقیقی بین دو عنصر در نمونه‌های برداشت شده که هیچکدام از آن دو عنصر دارای فراوانی غیرعادی در نمونه‌ها نباشند.
- ارتباط کاذب بین دو عنصری که یکی از آن دو دارای فراوانی غیرعادی باشند.
- ارتباط شدیداً کاذب بین دو عنصری که هر دو دارای فراوانی غیرعادی باشند.

حذف نمونه، شیب خط مثبت شده و به $+0/407$ تغییر می‌یابد به طوری که اختلاف شیب خطوط ($|b_2 - b_1|$) بین دو حالت نزدیک به $0/4$ است و این می‌تواند بیانگر یک ارتباط کاذب بین دو عنصر در حالتی باشد که نمونه خارج از رده برای یکی از دو عنصر دارای یک مقدار غیرعادی است. عنصری که مقدارش در نمونه خارج از رده غیرعادی نبود (Sr)، نسبت به عنصر U در دو حالت آورده شده است و مشخص گردید که معادلات تغییرات چندانی نسبت به هم ندارند و ارتباط چنین عنصری نسبت به عنصر U، یک ارتباط حقیقی است. با توجه به جدول معادلات و نمودار پراکندگی عناصر، دو عنصر Sr، U که مقدار این عناصر در نمونه خارج از رده به صورت غیرعادی نیست، ارتباطی که این عناصر نسبت به هم نشان می‌دهند در دو حالت ذکر شده شبیه به هم است به طوری که مشاهده می‌شود شیب معادله خط بین دو عنصر در دو حالت ۱۵۲ نمونه و ۱۵۱ نمونه به ترتیب $-0/751$ و $-0/814$ است که اختلاف آن دو اهمیت چندانی ندارد. همچنین مقدار ضریب همبستگی بین دو عنصر U و Sr در دو حالت مذکور به ترتیب $0/08$ و $0/07$ است که تقریباً مشابه هم هستند. Li و As دارای فراوانی غیرعادی در نمونه خارج از رده می‌باشند، نمودار پراکندگی و معادله خط رگرسیون آنها نسبت به هم تفاوت بسیار زیادی در دو حالت نشان می‌دهند. ارتباط واقعی بین دو عنصر As و Li پس از حذف نمونه خارج از رده مشخص می‌شود. تفاوت بین معادله خط رگرسیون دو عنصر As و Li در مقایسه با معادله خط رگرسیون سایر عناصر در شکل ۲ بیشتر است.

در آنالیز خوشه‌ای هدف دست یافتن به ملاکی جهت طبقه‌بندی هر چه مناسب‌تر متغیرها است. در این روش می‌توان متغیرها را با توجه به حداکثر تشابهی که نسبت به هم دارند گروه‌بندی کرد [۲]. در شکل ۳ نمودار B نشان‌دهنده هم‌گروهی عناصر Ni، Nb، V، Ga، Mn، Co، Cr با یکدیگر، همچنین وابستگی بیشتر این عناصر با عناصر فلزی کمیاب مانند Yb، Be، Eu نسبت به عناصر دیگر می‌باشد که با توجه به لیتولوژی منطقه، چنین هم‌گروهی و وابستگی چندان دور از انتظار نیست. از طرف دیگر در این دندروگرام عناصر U، Th، Ba، Pb، Zn، As در یک گروه قرار دارند که بیانگر وجود اندیس‌های هیدروترمال در منطقه است. حضور Ba در این گروه بیانگر وجود باریت و



References:

1. A.R.H Swan, M. Sndilands, P. Mccabe, "Introduction to geological data analysis," Backwill Science, p. 446 (1995).
2. ع.ا. حسنی پاک، "تحلیل داده‌های اکتشافی"، انتشارات دانشگاه تهران (۱۳۸۰).
3. B.R. Clarke, C.R. Heathcote, "Robust estimation of k-component univariate normal mixtures," Ann. Inst. Statist. Math. 46, 83-93 (1994).
4. C. Robert, Szava-Kovats, "Outlier-resistant errors-in-variables regression: anomaly recognition and grain-size correction in stream sediments," J. Applied Geochemistry. No. 17, pp.1149-1157 (2002).
5. P.J. Rousseeuw, A. Leroy, "Robust regression and outlier detection," John Wiley & Sons, Inc, New York (1987).
6. W.B. Size, "Use and abuse of statistical methods in the earth sciences," New York, Oxford (1986).

بنابراین تشخیص نمونه‌های خارج از رده، همچنین عناصری که در این نوع نمونه‌ها دارای مقادیر غیرعادی هستند یکی از مراحل مهم در بررسی‌های آماری ژئوشیمیایی است زیرا با وجود چنین نمونه‌هایی نمی‌توان ارتباط حقیقی بین عناصر را تشخیص داد.

پی‌نوشت‌ها:

- ۱- OLS: Ordinary Least Square
- ۲- WLS: Weighted Least Squares
- ۳- LAD: Least Absolute Deviation
- ۴- LNS: Least Normal Squares
- ۵- MSWR: Median Sum of Weighted Residuals
- ۶- LMS: Least Median Squares