

کاربرد تحلیل خوشه‌ای داده‌های ژئوشیمی نفت برای تعیین پیوستگی مخزن سروک در یکی از میادین نفتی جنوب غرب ایران

میلاد زهره وند^۱، علی شکاری فرد^{۲*} و وحید توکلی^۱

۱- دانشکده زمین‌شناسی، پردیس علوم، دانشگاه تهران، ایران

۲- گروه مهندسی نفت و گاز، دانشکده مهندسی شیمی، پردیس دانشکده‌های فنی دانشگاه تهران، ایران

۳- انستیتو مهندسی نفت، پردیس دانشکده‌های فنی دانشگاه تهران، ایران

تاریخ دریافت: ۱۳۹۸/۴/۱۳ تاریخ پذیرش: ۱۳۹۸/۶/۲۷

چکیده

در این مقاله از تحلیل خوشه‌ای داده‌های ژئوشیمی مولکولی نفت جهت تعیین پیوستگی افق‌های مخزنی سروک در یکی از میادین نفتی ایران، استفاده شد. این روش‌ها شامل تحلیل مؤلفه‌های اصلی (Principal Component Analysis; PCA)، تحلیل خوشه‌ای سلسله‌مراتبی (Hierarchical Cluster Analysis; HCA) و K- میانگین است. بررسی‌های آماری بر روی نسبت پیک‌های حاصل از کروماتوگرافی گازی با تفکیک بالا (High-resolution GC; HRGC) و همچنین داده‌های بایومارکری حاصل از کروماتوگرافی گازی- طیف‌سنج جرمی (GC-MS)، صورت گرفت. نتایج این بررسی نشان داد که نمودار ترسیم شده PC1 در مقابل PC2 برای نسبت زوج پیک‌های کروماتوگرافی گازی و نمودار ترسیم شده برای داده‌های بایومارکری شباهت بسیار زیادی با یکدیگر دارند. از سوی دیگر، تحلیل خوشه‌ای به روش HCA نیز برای نشانه‌گرهای زیستی و پیک‌های کروماتوگرافی گازی انجام شد که نمونه‌های نفتی را دقیقاً شبیه به روش PCA جدا کرده است. پس از ترسیم چندین پارامتر بلوغ بایومارکری در مقابل محور PC1، مشخص شد که رابطه خطی و مستقیمی بین آنها وجود دارد و از طرفی، گروه‌های نفتی را نیز به درستی از یکدیگر جدا کرده‌اند. این روابط حاکی از کارایی بالای روش تحلیل مؤلفه‌های اصلی (PCA) برای مطالعه نمونه‌های نفت مخزن سروک در میدان مورد مطالعه است. نتایج نشان می‌دهند که اساساً بلوغ حرارتی عامل اصلی خوشه‌بندی داده‌ها در این مطالعه بوده است. تشخیص این موضوع پس از انجام روش PCA بر روی داده‌ها صورت گرفت. ترسیم نمودار نشانه‌گرهای زیستی معرف بلوغ در مقابل محور PC1 نشان داد که اختلاف در سطح بلوغ نفت‌ها منجر به ایجاد گروه‌های نفتی متفاوت شده است. از آنجا که خوشه‌بندی داده‌های HRGC دقیقاً مشابه با خوشه‌بندی نشانه‌گرهای زیستی است و با در نظر گرفتن سایر پارامترهای به‌دست آمده، وجود گروه‌های مختلف نفتی می‌تواند بیانگر پیوستگی ضعیف/عدم پیوستگی مخزن سروک در میدان مورد مطالعه باشد.

کلمات کلیدی: ژئوشیمی نفت، تحلیل خوشه‌ای، تحلیل مؤلفه‌های اصلی، پیوستگی مخزنی، مخزن

سروک، ایران

*مسئول مکاتبات

ashekary@ut.ac.ir

آدرس الکترونیکی

شناسه دیجیتال: (DOI: 10.22078/pr.2019.3842.2750)

مقدمه

امروزه علم ژئوشیمی مخزن به‌عنوان یک علم نوین، به‌دلیل دقت بالا، صرفه جویی در وقت و هزینه، کاربردهای فراوانی در صنعت نفت دارد. یکی از کاربردهای مهم آن تشخیص پیوستگی مخازن است که ابزار مهمی برای تعیین موقعیت سطح تماس سیالات [۱]، کاربرد آن در زمینه ازدیاد برداشت و مسائل مربوط به تولید، ارزیابی و توسعه یک میدان نفتی است [۷-۲].

طبیعتاً پیوسته بودن سیالات در مخازن مجاور یکدیگر نشان‌دهنده یکپارچه بودن سازند مخزنی است. اما بیشتر اوقات این ارتباط بین سیالات در اثر برخی عوارض زمین‌شناسی و سدهای جریان سیال مانند گسل‌ها، لایه‌های قیر، قرار گرفتن لایه‌های تراوا در مقابل لایه‌های ناتراوا (شیل) تحت فعالیت شدید تکتونیکی منطقه، وجود فرآیندهای دیاژنزی (سیمانی شدن و از بین رفتن تخلخل و تراوایی) و مواردی از این قبیل، از بین رفته و منجر به ایجاد ناهمگنی سیالات و تشکیل افق‌های مخزنی مجزا می‌شوند [۸]. وقتی یک مخزن در حال پر شدن باشد، در هر مرحله، هیدروکربن‌های دریافت شده از سنگ منشأ سطح بلوغ و ترکیب متفاوتی را خواهند داشت. اگر این مخزن شامل افق‌های پیوسته باشد، تولید نهایی شامل مخلوطی از تمامی این نفت‌ها در طول تاریخچه پر شدن مخزن است اما در یک مخزن با بیش از یک افق مخزنی^۳، تاریخچه پر شدن هر بخش متفاوت خواهد بود. در نتیجه، تفاوت‌هایی جزئی در هیدروکربن‌های هر بخش قابل مشاهده است. این اختلافات می‌توانند به‌خوبی توسط داده‌های ژئوشیمی نفت تشخیص داده شوند [۹].

تاکنون مطالعات گسترده‌ای جهت ارزیابی پیوستگی مخازن و تشخیص افق‌های مخزنی مجزا در نقاط مختلف دنیا صورت گرفته است که می‌توان به برخی از این مطالعات نظیر بررسی پیوستگی سیالات مخزنی در میداین حوضه دریای شمال نظیر میدان نفتی فورتیز^۴، به‌کمک داده‌های PVT (فشار نقطه

حباب) و داده‌های ژئوشیمیایی نفت [۱۰] و مقایسه داده‌های اثر انگشت نگاری نفت از کروماتوگرافی گازی مربوط به چند حلقه چاه در خلیج مکزیک اشاره نمود [۲]. در ایران نیز مطالعه پیوستگی بین مخازن کرتاسه (سروک، فهلیان، گدوان و کژدمی) در جهت قائم براساس تطابق نفت با نفت به‌روش اثر انگشت نگاری در میداین نفتی جنوب غرب ایران (دشت آبادان) انجام شده است [۱۱]. سازند کربناتی سروک (آلبین-تورونین) پس از آسماری (الیگومیوسن) به‌عنوان دومین سازند مخزنی در میان میداین نفتی جنوب غرب ایران به‌شمار می‌رود که بخش قابل توجهی از ذخایر هیدروکربنی را در خود جای داده است. بنابر این، مطالعه این مخزن حائز اهمیت است. در این مقاله با بهره‌گیری از ژئوشیمی مخزن، مطالعاتی در مورد ارزیابی پیوستگی افق‌های مخزنی سروک انجام شده است. هدف اصلی این مقاله، تحلیل پیوستگی مخزن سروک به‌کمک مقایسه نتایج حاصل از اثر انگشت نگاری نفت و داده‌های بایومارکری، جهت تشخیص افق‌های مخزنی و گروه‌های نفتی مجزا است. این مقایسه توسط چند نوع روش تحلیل خوشه‌ای صورت گرفته است که تمرکز اصلی این مطالعه برروی تحلیل مؤلفه‌های اصلی^۵ (PCA) است. تحلیل مؤلفه‌های اصلی نخستین بار توسط کارل پیرسون به منظور کاهش ابعاد مجموعه داده‌ها ارائه شد. این روش، مؤلفه‌های اصلی را شناسایی کرده و به ما کمک می‌کند تا به‌جای اینکه تمامی ویژگی‌ها را مورد بررسی قرار دهیم، برخی از ویژگی‌ها که ارزش بیشتری دارند را تحلیل کنیم. در واقع، PCA ویژگی‌هایی که ارزش بیشتری فراهم می‌کنند را برای ما استخراج می‌کند.

زمین‌شناسی منطقه مورد مطالعه

ناحیه مورد مطالعه واقع در دشت آبادان به‌عنوان یکی از

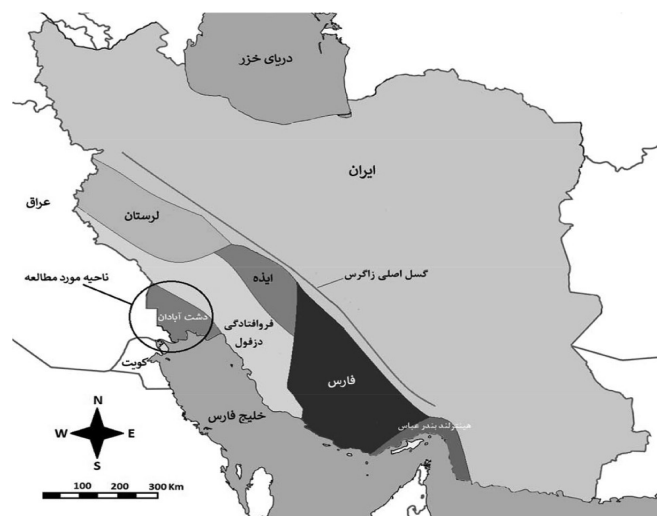
1. Reservoir Connectivity
2. Tar Mats
3. Compartment
4. Forties
5. Principal Component Analysis

هم‌شیب برروی سازند کژدمی (سنگ منشأ) قرار دارد و مرز بالایی آن شیل‌های لافان بوده (در فارس ساحلی و خوزستان) که به‌عنوان پوشش سنگ مخزن سروک شناخته شده‌اند [۱۳ و ۱۲]. سازند سروک طی پیشروی دریا در بازه زمانی آلبین پسین-تورونین پیشین رسوب کرده و لیتولوژی آن شامل سنگ آهک توده‌ای نرم تا نیمه سخت رودیست دار به‌رنگ سفید و کرم با میان لایه‌های آهک رسی به‌رنگ سفید و خاکستری است. شکل ۲ ستون چینه‌شناسی از توالی رسوبی در ناحیه مورد مطالعه را نشان می‌دهد. از جمله سازندهای مخزنی مهم به سن کرتاسه در ناحیه مورد مطالعه علاوه‌بر سازند سروک، می‌توان به سازندهای فهلیان، گدوان، داریان و ایلام نیز اشاره نمود [۱۳].

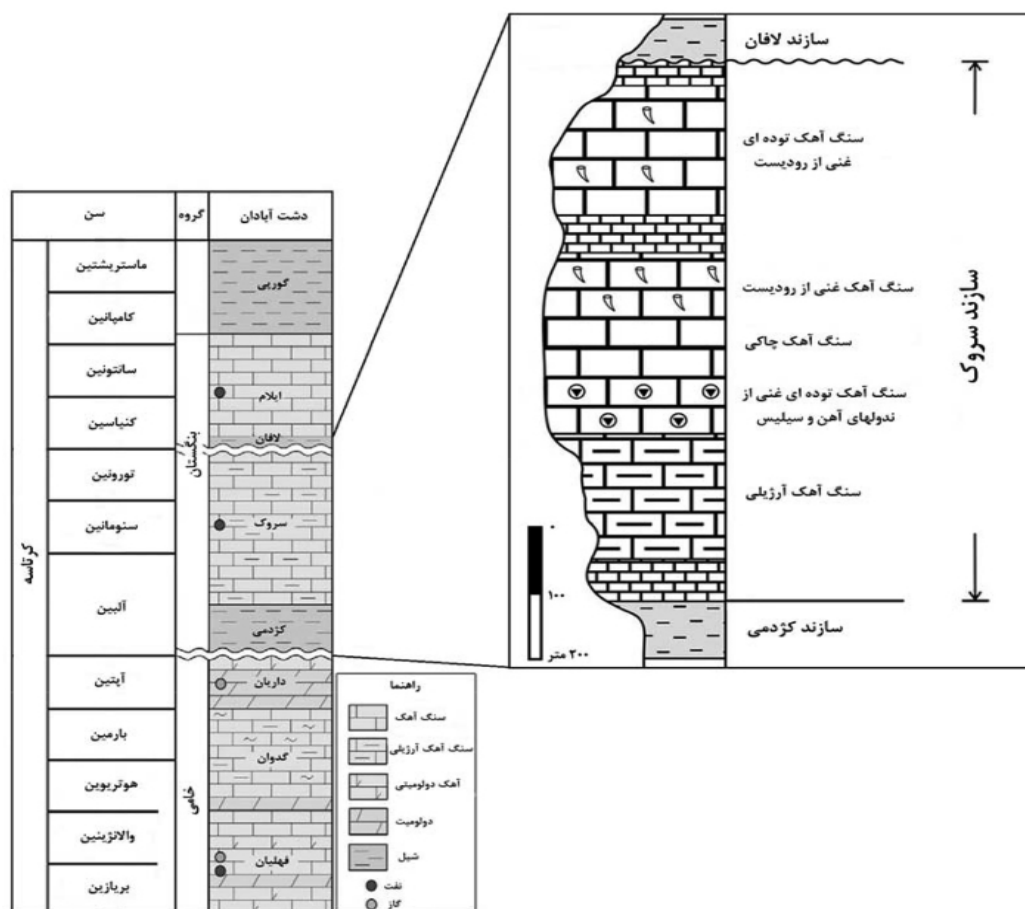
مواد و روش‌ها

یازده نمونه نفت خام از مخزن سروک در یکی از میداین نفتی جنوب غرب ایران توسط آزمایش ساق مته (DST) و آزمایش مجدد سازند (RFT) از ۵ حلقه چاه در در افق‌های مخزنی متفاوت جمع‌آوری شدند

زیرپهنه‌های زاگرس بوده که در بخش شمال شرق صفحه عربستان (حاشیه فعال صفحه عربستان) قرار گرفته است. این ناحیه در جنوب غربی کمربند چین‌خورده-تراستی زاگرس و فروافتادگی دزفول قرار دارد که از جنوب به خلیج فارس منتهی می‌شود (شکل ۱). در زمان کرتاسه میانی، فعال شدن دوباره گسل‌های پی‌سنگی و حرکات لایه‌های نمکی مربوط به سری هرمز، باعث ایجاد پستی و بلندی در ناحیه فروافتادگی دزفول و دشت آبادان شده است [۱۶-۱۲]. سازند سروک در مرز سنومانین-تورونین با خروج از آب، تحت تأثیر دیاژنز جوی قرار گرفته است که انحلال و کارستی شدن را در بخش بالایی سازند سروک گسترش داده و منجر به بالا رفتن کیفیت مخزنی شده است. علاوه‌بر انحلال و کارستی شدن، عوامل دیگری همچون دولومیتی شدن، شکستگی، تراکم، میکرایتی شدن، تبلور مجدد، استیلولیت و سیمانی شدن، کم و بیش بر کیفیت مخزنی سازند سروک تأثیر گذار بوده است [۱۳]. محیط رسوبی این سازند به‌صورت یک رمپ کربناتی و شامل رخساره‌های کم عمق (لاگون و ریف) و عمیق است. سازند سروک (متعلق به گروه بنگستان) به‌صورت



شکل ۱ تقسیمات ساختمانی زاگرس و موقعیت جغرافیایی منطقه مورد مطالعه (دشت آبادان)



شکل ۲ ستون چینه‌شناسی کر تاسه مربوط به ناحیه دشت آبادان و توالی رسوبی از سازند سروک در این ناحیه

بخش آسفالتین به کمک محلول نرمال هگزان ته نشین و سپس از نفت خام جدا شد. سایر بخش‌های نفت از جمله اشباع، آروماتیک و پولار (NSO) با استفاده از کروماتوگرافی ستونی به ترتیب توسط محلول‌های نرمال هگزان، بنزن و متانول جداسازی و برای آنالیزهای GC و GC-MS آماده شدند. از آنجاکه سری آلکان‌ها در کروماتوگرافی گازی مربوط به نمونه‌ها اختلاف کمی با یکدیگر داشتند، برای دقت در مقایسه آنها، از روش کروماتوگرافی گازی با قدرت تفکیک بالا^۱ (HRGC) استفاده شد و سپس نسبت ارتفاع زوج پیک‌های مجاور یکدیگر در محدوده ۹ تا ۲۰ کربنی محاسبه شدند. آنالیز کروماتوگرافی گازی توسط دستگاه Mega Series HRGC 5160 (ساخت کشور ایتالیا)، دارای ستون موئین به طول ۳۰ m، با قطر داخلی ۰/۲۵ mm، در محدوده دمایی ۴۰ تا ۲۲۰ °C با نرخ افزایش دمای ۵ °C/min انجام شد. آنالیز GC-

MS نیز با استفاده از دستگاه طیف‌سنج جرمی^۲ از نوع Micromass ProSpec (ساخت کشور آمریکا) دارای ستون موئین ۶۰ m با قطر داخلی ۰/۲۵ mm در دمای ۵۰ تا ۱۲۰ °C با نرخ افزایش دمای ۲ °C/min انجام شد و نشانه‌گرهای زیستی شاخص مورد نیاز مانند سری استران‌ها، ترپان‌ها و ایزوپرنوئیدها به دست آمدند. گاز حامل در تمامی دستگاه‌های مورد استفاده، هلیوم بوده است. در این پژوهش، برای انجام تمامی کارهای آماری، محاسباتی و ترسیم نمودارها از نرم‌افزارهای XLSTAT، پایتون و گرافر ۱۲ استفاده شده است. به دلیل اهمیت روش آماری PCA در این مطالعه، مراحل و روند کار این روش در زیر به‌طور مختصر شرح داده شده است: در این روش، داده‌ها به یک دستگاه مختصات جدید برده می‌شوند.

1. High Resolution GC
2. Mass Spectrometer

$$\begin{pmatrix} \text{var}_{(x)} & \text{cov}_{(x,y)} & \text{cov}_{(x,z)} \\ \text{cov}_{(y,x)} & \text{var}_{(y)} & \text{cov}_{(y,z)} \\ \text{cov}_{(z,x)} & \text{cov}_{(z,y)} & \text{var}_{(z)} \end{pmatrix}$$

در ماتریس کوواریانس، دامنه تغییرات دو متغیر نسبت به یکدیگر سنجیده می‌شود. بنابراین، کوواریانس هر متغیر نسبت به خودش برابر با واریانس آن متغیر است. به همین دلیل قطر اصلی ماتریس کوواریانس، واریانس متغیرها است. مرحله بعد به دست آوردن مقادیر ویژه^۱ و بردارهای ویژه^۲ ماتریس کوواریانس داده‌ها است که به صورت یک رابطه پایه مانند معادله ۴ قابل محاسبه است: در واقع بردارهای ویژه می‌توانند جهت توزیع‌شدگی داده‌ها را مشخص کنند.

$$AX = \lambda X \quad (۴)$$

در اینجا A یک ماتریس مربعی، X بردار ویژه و λ مقدار ویژه ماتریس کوواریانس داده‌ها است. از آنجاکه معادله ۴ یک معادله غیرخطی است، نمی‌توان مجهول‌ها را به راحتی پیدا کرد. بنابراین، باید X و λ از هم جدا شوند که می‌توان آنها را توسط ماتریس همانی به نام I (ماتریسی که قطر اصلی آن یک و بقیه درایه‌های آن صفر است)، طبق رابطه ۵ جدا کرد.

$$AX = \lambda IX \rightarrow (A - \lambda I)X = 0 \quad (۵)$$

برای اینکه این معادله جواب غیر صفر داشته باشد باید وارون‌پذیر باشد یعنی دترمینان آن صفر شود:

$$|A - \lambda I| = 0$$

ریشه‌های این معادله مقادیر ویژه ماتریس A خواهند بود. با یافتن مقادیر ویژه، حال می‌توان معادله ۵ را حل کرد و بردارهای ویژه را به دست آورد. مرحله بعد انتخاب اجزای اصلی^۳ و ساخت بردار ویژگی^۴ (FV) است. کاهش ابعاد داده‌ها در اینجا رخ می‌دهد یعنی می‌توان مقادیر ویژه کوچک را حذف کرد.

به این صورت که بزرگ‌ترین واریانس داده‌ها بر روی اولین محور مختصات، بزرگ‌ترین واریانس دوم بر روی دومین محور مختصات قرار می‌گیرند و به همین ترتیب برای بقیه محورها (در جهت کاهش واریانس). در این حالت گونه‌های جدیدی از داده‌ها پدید می‌آیند که ویژگی هیچ‌کدام از متغیرهای اولیه را نداشته و برخلاف متغیرهای اولیه که همبستگی کامل داشتند، متغیرهای ساخته شده هیچ‌گونه همبستگی با یکدیگر ندارند. مؤلفه‌های اصلی (PC) در روش PCA در حقیقت یکسری از بردارها هستند که از ماتریس کوواریانس داده‌ها محاسبه می‌شوند که بردارهای ویژه نام دارند. همچنین از مجموع مجذورات بردارهای ویژه، اعدادی به دست می‌آیند که مقادیر ویژه نامیده می‌شوند. بیشترین واریانس داده‌ها در راستایی قرار می‌گیرد که بردار ویژه متناظر با بزرگ‌ترین مقدار ویژه، در آن راستا قرار دارد. به همین ترتیب هر چه مقدار ویژه کوچک‌تر شود، واریانس داده‌ها در راستای بردار ویژه متناظر با آن نیز کمتر می‌شود [۱۸ و ۱۷]. برای محاسبه تحلیل مؤلفه‌های اصلی ابتدا باید داده‌های ماتریس استاندارد شوند. پس از میانگین‌گیری از داده‌ها، با استفاده از محاسبه انحراف معیار طبق رابطه ۱، می‌توان واریانس داده‌ها را طبق رابطه ۲ و سپس کوواریانس را طبق رابطه ۳ محاسبه کرد:

$$SD = \sqrt{\frac{\sum_{i=1}^n (X_i - \bar{X})^2}{(n-1)}} \quad (۱)$$

$$(S^2) \text{Var} = \frac{\sum_{i=1}^n (X_i - \bar{X})}{n-1} \quad (۲)$$

$$\text{COV}_{(x,y)} = \frac{\sum_{i=1}^n (X_i - \bar{X}) \cdot (Y_i - \bar{Y})}{n-1} \quad (۳)$$

با محاسبه کوواریانس داده‌ها می‌توان ماتریس کوواریانس را به صورت زیر نوشت:

$$\begin{pmatrix} \text{COV}_{(x,x)} & \text{COV}_{(x,y)} & \text{COV}_{(x,z)} \\ \text{COV}_{(y,x)} & \text{COV}_{(y,y)} & \text{COV}_{(y,z)} \\ \text{COV}_{(z,x)} & \text{COV}_{(z,y)} & \text{COV}_{(z,z)} \end{pmatrix} =$$

1. Eigenvalues
2. Eigenvectors
3. Components
4. Feature Vector

با تفکیک بالا را نشان می‌دهد. روش آنالیز HRGC باعث می‌شود که پیک‌های غیر قابل رؤیت در کروماتوگرام‌های گازی، با وضوح بالاتر نمایان شوند و دقت کار افزایش یابد. در این مطالعه پیک‌های انتخاب شده، آلکان‌های نرمال ۹ تا ۲۰ کربنی است. این بازه از هیدروکربن‌ها جزء بخش سبک نفت بوده و در بیشتر کروماتوگراف‌های گازی به خوبی ظاهر می‌شوند. همان‌طور که مشاهده می‌شود پیک‌های بسیار کوچک، با بزرگ‌نمایی بالاتر نمایان شده‌اند. نتایج حاصل از آنالیز HRGC و GC-MS به ترتیب در جداول ۱ و ۲ نشان داده شده است. در این مطالعه، ابتدا مقایسه نسبت ارتفاع زوج پیک‌ها و نشانه‌گرهای زیستی به کمک دیاگرام‌های ستاره‌ای مورد تجزیه و تحلیل قرار گرفت (شکل ۴ الف و ب). در این نمودارها، به دلیل وجود اختلاف بسیار کم بین داده‌ها و زیاد بودن تعداد متغیرها، منحنی‌ها در بسیاری از نقاط همپوشانی زیادی با یکدیگر داشته که داده‌ها به خوبی از هم جدا نمی‌شوند. بنابراین، کیفیت مقایسه دقیق‌تر داده‌ها بر روی چنین نموداری چندان مطلوب نیست.

در واقع بردار ویژه‌ای که بیشترین مقدار ویژه را داشته باشد به عنوان اجزای اصلی (PC) مجموعه داده‌ها انتخاب می‌شود. بردار ویژگی، شامل بردارهای ویژه‌ای است که از لیست بردارهای ویژه انتخاب شده‌اند (همان بردارهایی که بیشترین واریانس مجموعه داده‌ها را دارند). این بردار (بردار ویژگی)، ماتریسی است که ستون‌هایی از بردارهای ویژه را شامل شده است (رابطه ۶).

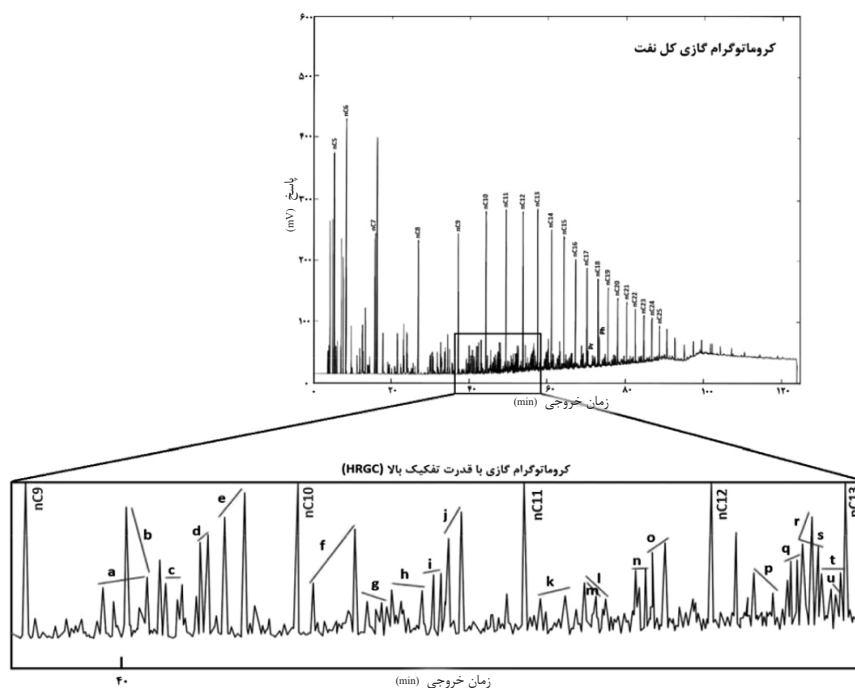
$$FV = (\text{eig}_1, \text{eig}_2, \text{eig}_3, \dots, \text{eig}_n) \quad (6)$$

مرحله پایانی، محاسبه مجموعه داده‌های جدید است که به صورت زیر محاسبه می‌شود: داده‌های جدید = داده‌های تعدیل شده ترانهاده^۱ x بردارهای ویژه ترانهاده^۲.

در نهایت، داده‌های جدید به دست آمده برای نمایش خوشه‌ها به فضای مختصاتی جدید منتقل می‌شوند.

نتایج و بحث

شکل ۳ نمونه‌ای از آنالیز یک کروماتوگرافی گازی



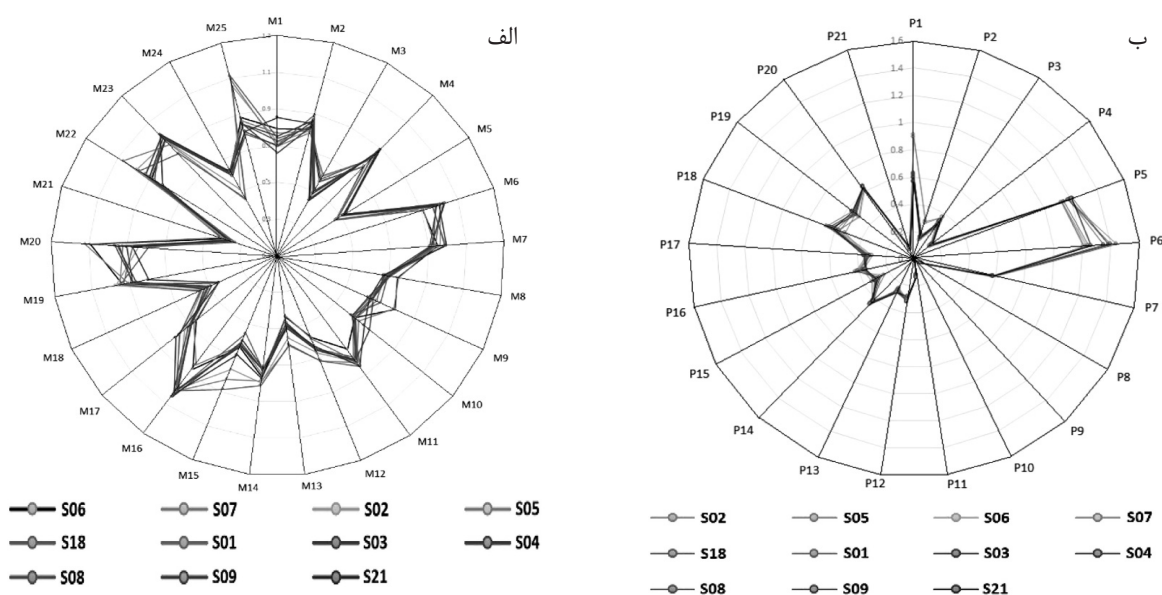
شکل ۳ نمونه‌ای از طیف خروجی کروماتوگرافی گازی با وضوح بالا

جدول ۱ نسبت ارتفاع زوج پیک‌های حاصل از HRGC برای نمونه‌های نفت مورد مطالعه (نسبت ارتفاع پیک‌ها با M نمایش داده شده است، همچنین شماره هر M نیز ترتیب شماره‌گذاری زوج پیک‌ها بر روی کروماتوگرام‌های گازی است)

نسبت پیک‌ها	نمونه‌های نفت خام										
	S06	S07	S02	S05	S18	S01	S03	S04	S08	S09	S21
M1	۰/۷۹	۰/۷۷	۰/۷۲	۰/۷	۰/۷۴	۰/۷	۰/۷	۰/۷۵	۰/۷۲	۰/۶۶	۰/۸۵
M2	۰/۸۱	۰/۸۱	۰/۸	۰/۷۹	۰/۸۲	۰/۸۹	۰/۸۶	۰/۸۶	۰/۸۳	۰/۸۳	۰/۸۴
M3	۰/۵۶	۰/۵۶	۰/۵۹	۰/۶	۰/۵۸	۰/۴۹	۰/۵۱	۰/۴۶	۰/۴۹	۰/۴۹	۰/۴۴
M4	۰/۷۷	۰/۷۷	۰/۸۲	۰/۸۲	۰/۸۳	۰/۸۶	۰/۸۷	۰/۸۹	۰/۸۸	۰/۸۶	۰/۸۹
M5	۰/۴۷	۰/۴۶	۰/۴۷	۰/۴۷	۰/۴۷	۰/۵۲	۰/۵۲	۰/۵۲	۰/۵۱	۰/۵۱	۰/۵۱
M6	۰/۹۱	۰/۹۱	۰/۹۶	۱	۰/۹۶	۱/۰۱	۱/۰۲	۰/۹۸	۰/۹۸	۱/۰۳	۱
M7	۰/۹۵	۰/۹۳	۰/۹۹	۰/۹۶	۰/۹۹	۰/۹۴	۰/۹۱	۰/۹۱	۰/۹۳	۰/۹۱	۰/۹۹
M8	۰/۷۴	۰/۷۵	۰/۶۸	۰/۶۷	۰/۶۹	۰/۷	۰/۷	۰/۶۸	۰/۶۹	۰/۷	۰/۶۷
M9	۰/۷۹	۰/۷۸	۰/۶۸	۰/۶۸	۰/۷	۰/۶۴	۰/۶۶	۰/۶۶	۰/۶۵	۰/۶۳	۰/۶۴
M10	۰/۶۳	۰/۶۶	۰/۶۳	۰/۶۴	۰/۶۱	۰/۶۱	۰/۶۳	۰/۷۳	۰/۶۲	۰/۶۱	۰/۶۹
M11	۰/۷۳	۰/۸۱	۰/۸۲	۰/۸۲	۰/۸۵	۰/۷۸	۰/۸	۰/۸	۰/۸۲	۰/۸۱	۰/۸۴
M12	۰/۵۸	۰/۵۹	۰/۶۷	۰/۷۲	۰/۷۲	۰/۶۴	۰/۶۵	۰/۶۵	۰/۶۴	۰/۶۵	۰/۶۳
M13	۰/۴۴	۰/۴۳	۰/۵۵	۰/۵۹	۰/۵۹	۰/۴۷	۰/۴۷	۰/۴۸	۰/۴۹	۰/۵	۰/۴۸
M14	۰/۷۷	۰/۷۵	۰/۸	۰/۷۹	۰/۸۱	۰/۷۲	۰/۷۲	۰/۷۳	۰/۷۳	۰/۷۳	۰/۷۷
M15	۰/۶۴	۰/۶۳	۰/۷۷	۰/۸۲	۰/۸۸	۰/۵۶	۰/۶	۰/۶۳	۰/۶۲	۰/۶۳	۰/۶۸
M16	۰/۸۶	۰/۸۳	۱	۱/۰۱	۱/۰۱	۱/۰۲	۱/۰۵	۰/۹۹	۱	۱/۰۱	۱/۰۳
M17	۰/۶۶	۰/۶۸	۰/۷۴	۰/۷۷	۰/۷۸	۰/۷	۰/۷۱	۰/۷	۰/۶۹	۰/۶۹	۰/۸
M18	۰/۴۵	۰/۴۵	۰/۵۴	۰/۵۶	۰/۵۵	۰/۵۱	۰/۵۱	۰/۴۹	۰/۵۱	۰/۵۳	۰/۴۴
M19	۰/۹۳	۰/۹	۰/۹۷	۰/۹۵	۰/۹۵	۰/۹۳	۰/۸۸	۰/۸۸	۰/۸۸	۰/۸۶	۰/۷۹
M20	۱/۰۹	۱/۱۲	۰/۸۴	۰/۹۲	۰/۷۹	۰/۸۷	۰/۹	۰/۹۶	۰/۹۴	۰/۹۵	۰/۸۷
M21	۰/۴	۰/۴۲	۰/۳۹	۰/۳۶	۰/۳۷	۰/۳۳	۰/۳۷	۰/۳۶	۰/۳۷	۰/۳۵	۰/۳۸
M22	۰/۸۹	۰/۹۳	۰/۹۸	۱/۰۷	۰/۹۶	۰/۹	۰/۸۹	۰/۹	۰/۸۷	۰/۹۲	۰/۸۲
M23	۰/۹۷	۰/۹۹	۰/۹۶	۰/۹۳	۰/۸۶	۰/۹۷	۰/۹۷	۰/۹۸	۱	۱	۱/۰۱
M24	۰/۶۲	۰/۶۳	۰/۵۱	۰/۴۵	۰/۴۶	۰/۶۱	۰/۶	۰/۶۳	۰/۶۲	۰/۶۱	۰/۶۲
M25	۰/۸۸	۰/۸۶	۱/۰۵	۱/۰۹	۱/۱۱	۰/۷۸	۰/۸	۰/۸۴	۰/۸۴	۰/۸۴	۰/۸

جدول ۲ داده‌های مربوط به نشانه‌گرهای زیستی حاصل از آنالیز GC-MS برای نمونه‌های نفت مورد مطالعه

نماد	نشانه‌گرهای زیستی	نمونه‌های نفت خام										
		S02	S05	S06	S07	S18	S01	S03	S04	S08	S09	S21
P1	pri/ph	۰/۶۲۳	۰/۵۹۹	۰/۹۰۸	۰/۸۹	۰/۶۳	۰/۵۹	۰/۵۶۶	۰/۶۱۸	۰/۵۹۱	۰/۵۹۷	۰/۵۷۲
P2	pri/nC17	۰/۱۵۷	۰/۱۴۵	۰/۲۸۷	۰/۲۷۸	۰/۱۶	۰/۱۷۸	۰/۱۷۱	۰/۱۷۹	۰/۱۷۶	۰/۱۶۵	۰/۱۶۲
P3	ph/nC18	۰/۳	۰/۲۸۳	۰/۳۷۲	۰/۳۶۷	۰/۲۹۴	۰/۳۳۲	۰/۳۳۹	۰/۳۲۴	۰/۳۴۲	۰/۳۰۶	۰/۳۳۳
P4	Ts/Ts+Tm	۰/۱۴۶	۰/۱۴۷	۰/۱۴۸	۰/۱۴۹	۰/۱۶	۰/۱۷۳	۰/۱۷۲	۰/۱۷	۰/۱۷	۰/۱۶۲	۰/۱۶۷
P5	C29/C30 hopanes	۱/۱۳۳	۱/۱۳۱	۱/۱۲۵	۱/۱۲۷	۱/۱۴۸	۱/۱۷۷	۱/۱۸۴	۱/۱۹۲	۱/۲۰۸	۱/۱۸۱	۱/۲۰۶
P6	C35/C34 hopanes	۱/۲۳۴	۱/۲۱۴	۱/۲۱۳	۱/۲۱۱	۱/۲۴۳	۱/۴۴	۱/۳۵۵	۱/۳۸۶	۱/۳۹	۱/۴	۱/۲۶۲
P7	C32 22S/22S+R	۰/۵۷۳	۰/۵۷۵	۰/۵۷۴	۰/۵۷۴	۰/۵۷۲	۰/۵۸۱	۰/۵۷۷	۰/۵۷۸	۰/۵۷۷	۰/۵۷۸	۰/۵۷۷
P8	C30 moretane/ moreα+βhopane	۰/۰۵۷	۰/۰۵۷	۰/۰۵۸	۰/۰۵۸	۰/۰۵۷	۰/۰۶۱	۰/۰۶۱	۰/۰۶۲	۰/۰۶۲	۰/۰۶۱	۰/۰۵۹
P9	C19-C29 Tricyclics/ C29-3517α hopanes	۰/۰۰۵	۰/۰۰۴	۰/۰۰۶	۰/۰۰۶	۰/۰۰۵	۰/۰۰۶	۰/۰۰۵	۰/۰۰۶	۰/۰۰۶	۰/۰۰۵	۰/۰۰۵
P10	Gammacerane / C30 hopane	۰/۰۱۳	۰/۰۱۳	۰/۰۱۲	۰/۰۱۲	۰/۰۱۲	۰/۰۱۳	۰/۰۱۳	۰/۰۱۳	۰/۰۱۳	۰/۰۱۳	۰/۰۱
P11	Extended Tri- cyclic Terpane Ratio	۰/۱۲	۰/۱۲	۰/۱۲۳	۰/۱۲۳	۰/۱۱۳	۰/۱۳۳	۰/۱۳	۰/۱۳۳	۰/۱۳۱	۰/۱۲۸	۰/۱۴
P12	C27 sterane /% C27+28+29	۰/۳۱۹	۰/۳۰۱	۰/۲۹۷	۰/۳۲۱	۰/۳۱۴	۰/۳۱۲	۰/۳۰۹	۰/۲۹۷	۰/۳۲۴	۰/۳۱۶	۰/۲۷۸
P13	C28 sterane /% C27+28+29	۰/۲۶۴	۰/۲۴	۰/۲۵۸	۰/۲۴۵	۰/۲۴۷	۰/۲۷۴	۰/۲۷	۰/۲۷۱	۰/۲۵۲	۰/۲۶۱	۰/۲۷۴
P14	C29 sterane /% C27+28+29	۰/۴۱۷	۰/۴۵۹	۰/۴۴۵	۰/۴۳۳	۰/۴۳۸	۰/۴۱۴	۰/۴۲۱	۰/۴۳۱	۰/۴۲۴	۰/۴۲۳	۰/۴۴۸
P15	C27 diast /% C27+28+29	۰/۲۶۶	۰/۲۴۲	۰/۲۹۲	۰/۳۳۵	۰/۳۰۴	۰/۳۰۹	۰/۳۰۸	۰/۲۸۶	۰/۳۱۹	۰/۳۱۸	۰/۲۹۵
P16	C28 diast /% C27+28+29	۰/۴۱	۰/۴۲۲	۰/۴۰۱	۰/۳۷	۰/۳۴۹	۰/۳۶۲	۰/۳۶۵	۰/۴۰۲	۰/۳۴۱	۰/۳۵۶	۰/۳۵۱
P17	C29 diast /% C27+28+29	۰/۳۲۴	۰/۳۳۷	۰/۳۰۷	۰/۲۹۴	۰/۳۴۷	۰/۳۲۹	۰/۳۲۸	۰/۳۱۳	۰/۳۴	۰/۳۲۶	۰/۳۵۴
P18	C28/C29 Steranes	۰/۶۳۵	۰/۵۲۲	۰/۵۸۱	۰/۵۶۶	۰/۵۶۴	۰/۶۶	۰/۶۴۱	۰/۶۲۹	۰/۵۹۴	۰/۶۱۶	۰/۶۱۲
P19	20S/20S+R ααα C29 sterane	۰/۵۳۶	۰/۵۳	۰/۴۸۹	۰/۵۰۱	۰/۵۴۶	۰/۵۵۶	۰/۵۳۷	۰/۵۲۳	۰/۵۳۳	۰/۵۳	۰/۵۵۳
P20	ββ/αα+ ββ C29 sterane	۰/۶۳۳	۰/۶۲۴	۰/۶۱۳	۰/۶۲۵	۰/۶۳۹	۰/۶۳۷	۰/۶۴۲	۰/۶۴۱	۰/۶۴	۰/۶۳۷	۰/۶۴
P21	diasteranes / st+dia	۰/۰۷۴	۰/۰۷۵	۰/۰۹۵	۰/۰۹۲	۰/۰۶۹	۰/۰۹۳	۰/۰۹۶	۰/۰۹۴	۰/۰۹۴	۰/۰۷۹	۰/۰۸۲



شکل ۴ الف) نمودار ستاره‌ای مربوط به نسبت زوج پیک‌های کروماتوگرافی گازی و ب) نمودار ستاره‌ای مربوط به داده‌های بایومارکری برای نمونه‌های نفت مورد مطالعه از سازند سروک

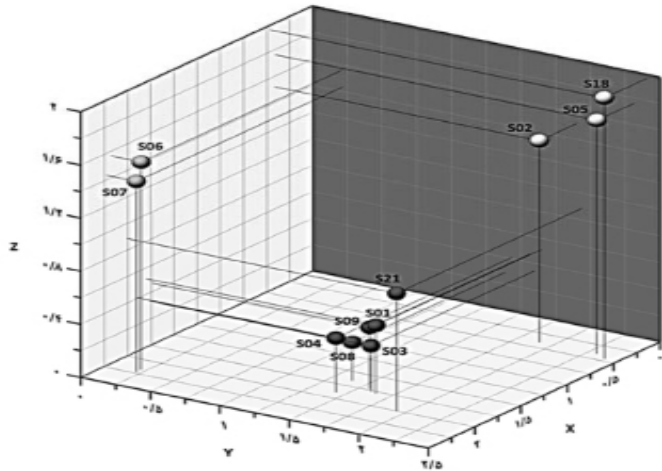
خود جدا شده و وارد مجموعه دیگری شده است. بنابراین، روش K- میانگین دقت نسبتاً پایین‌تری برای داده‌های بایومارکری در این مطالعه دارد. شکل ۶، موقعیت نمونه‌های نفت به دست آمده از مخزن سروک در افق‌های مخزنی مختلف از چندین چاه در این میدان را نشان می‌دهد. لازم به ذکر است که از چاه O نمونه‌ای دریافت نشده است. در الگوریتم‌های خوشه‌بندی سلسله‌مراتبی (HCA) که به نمودار حاصل از آن دندوگرام^۱ می‌گویند، در بهترین حالت تمام داده‌ها به‌عنوان یک خوشه در نظر گرفته می‌شوند و سپس، در طی یک فرآیند تکراری در هر مرحله داده‌هایی که شباهت کمتری به یکدیگر دارند به خوشه‌های مجزا شکسته می‌شوند. برای تعیین تعداد خوشه می‌توان دندوگرام را در یک نقطه مناسب برش داد. در این نمودار آنچه که اهمیت دارد، ارتفاع است. هرچه خوشه‌های تشکیل شده در ارتفاع پایین‌تری ایجاد شده باشند، خوشه‌ها یا مشاهدات به یکدیگر شبیه‌تر هستند و بالعکس. در شکل ۷ الف و ب، دندوگرام‌های مربوط به نسبت زوج پیک‌های کروماتوگرافی گازی و نشانه‌گرهای زیستی نشان داده شده است.

در روش K- میانگین، ابتدا K نقطه به‌عنوان نقاط مراکز خوشه‌ها انتخاب می‌شوند و هر نمونه داده، به خوشه‌ای که مرکز آن کمترین فاصله تا آن داده را دارد، نسبت داده می‌شود. از لحاظ ریاضی، روش K- میانگین مجموع مربعات درون خوشه‌ها را به حداقل می‌رساند تا میانگین مربعات فاصله بین خوشه‌ها به حداکثر برسد [۱۷]. در این روش، ابتدا باید تعداد خوشه‌ها مشخص شود. برای این منظور قبل از انجام روش K- میانگین، خوشه‌بندی توسط روش‌های PCA و HCA انجام شد و سپس، تعداد خوشه‌های به دست آمده از این دو روش، به‌عنوان تعداد خوشه‌ها برای انجام روش K- میانگین انتخاب شدند. با انجام روش K- میانگین بر روی داده‌های کروماتوگرافی گازی مشخص شد که جداسازی خانواده‌های نفتی دقیقاً مشابه با خوشه‌بندی توسط روش PCA و HCA است اما برای نشانه‌گرهای زیستی کمی اختلاف وجود دارد که همین امر منجر به ناکارآمدی این روش برای نشانه‌گرهای زیستی در این مطالعه شده است (شکل ۵ الف و ب). همان‌طور که در شکل ۵ (الف و ب) مشاهده می‌شود، نمونه‌ها به سه گروه نفتی تقسیم شده‌اند. اما در شکل ۵ ب، نمونه S21 از چاه E با اختلاف زیادی از مجموعه

1. Dendrogram

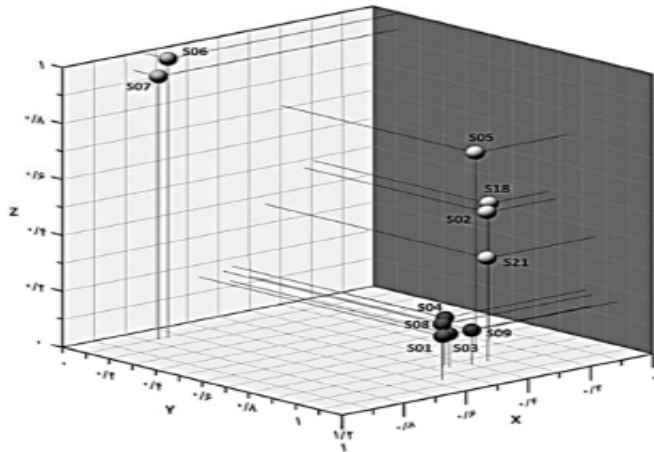
الف)

نمونه‌ها	K=۳			عوشه‌ها
	۱	۲	۳	
S02	-۲۱۸۷۵	۱۵۶۵۳۸	۲۱۷۵۱	۱
S05	-۲۱۸۷۵	۱۲۳۳۱۵	۲۲۲۶۸۷	۱
S06	۱۹۰۲۵۵	۱۵۲۰۷۲	-۲۰۷۹۷	۳
S07	۲۲۱۵۵۵	۱۷۶۰۹۸	-۲۰۳۳۳	۳
S18	۲۲۰۸۷۵	۱۹۶۲۴۵	-۲۲۳۳۳	۳
S01	۱۶۲۸۷۵	-۵۱۶۳۸	۱۷۸۸۶۷	۲
S03	۱۶۰۰۰۵	-۲۳۵۰۵	۱۷۷۸۷۳	۲
S04	۱۲۳۳۴۵	-۲۱۱۳۵	۱۹۰۰۶	۲
S08	۱۲۸۹۵۵	-۲۹۰۰۵	۱۶۵۴۳۳	۲
S09	۱۵۶۶۷۵	-۲۶۵۴۲	۱۷۳۳۳۳	۲
S21	۱۹۳۹۰۵	-۸۸۸۱۵	۲۰۰۰۰۷	۲

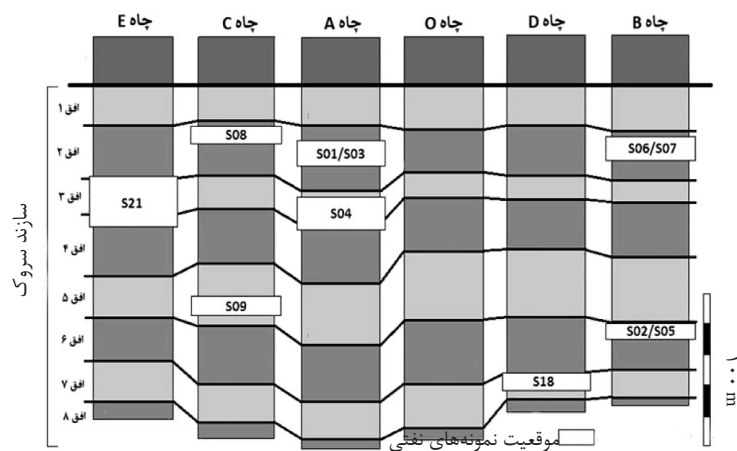


ب)

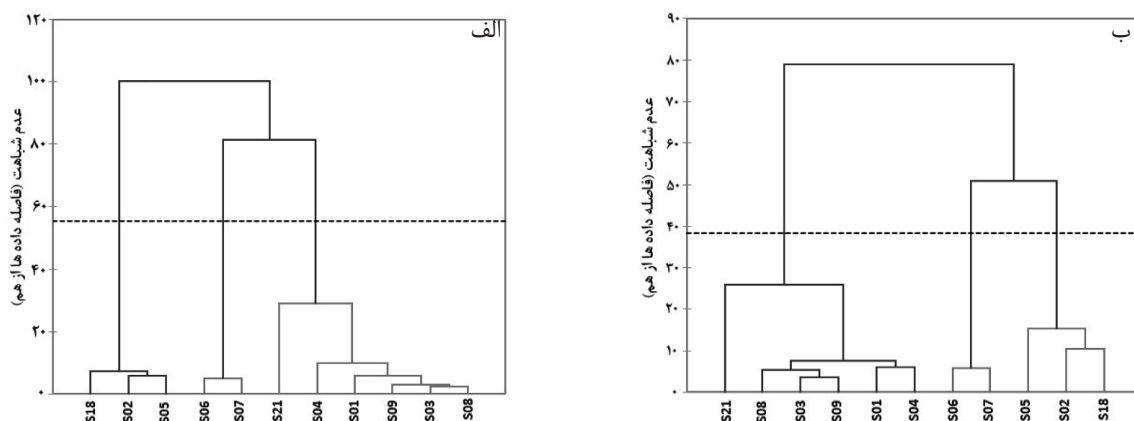
نمونه‌ها	K=۳			عوشه‌ها
	۱	۲	۳	
S02	-۷۸۵۵	-۲۷۱۶	-۲۲۲	۳
S05	-۸۳۳۵	-۷۰۹۸	-۲۹۶	۳
S06	-۱۰۷۵	-۹۹۵۸	-۷۴۱	۱
S07	-۱۰۷۵	-۹۳۹۶	-۷۷۱۵	۱
S18	-۷۵۳۵	-۲۹۰۲	-۱۹۱۵	۳
S01	۱۰۲۸۵	-۱۵۶۲	-۵۶۲۵	۲
S03	-۹۳۹۵	-۱۲۰۲	-۴۶۱	۲
S04	-۹۱۲۵	-۱۷۰۶	-۴۵۳	۲
S08	-۹۳۹۵	-۱۵۹۲	-۴۸۲۵	۲
S09	-۹۵۵۵	-۱۲۳	-۳۹۹	۲
S21	-۹۵۲۵	-۳۶۹۲	-۳۴۸۵	۳



شکل ۵ الف) K- میانگین محاسبه شده برای داده‌های مربوط به نسبت زوج پیک‌های کروماتوگرافی گازی همراه با نمایش آن بر روی نمودار سه‌بعدی ب) K- میانگین محاسبه شده برای داده‌های مربوط به نشانه‌گرهای زیستی همراه با نمایش آن بر روی نمودار سه‌بعدی (محورها برحسب فاصله نمونه‌ها از یکدیگر هستند)



شکل ۶ موقعیت نمونه‌های نفت به‌دست آمده از اعماق و افق‌های تولیدی مختلف از مخزن سروک



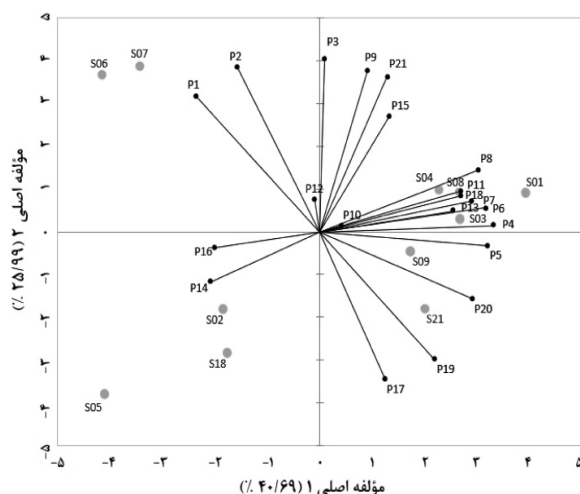
شکل ۷ تحلیل خوشه‌ای سلسله مراتبی (HCA). الف) دندوگرام ترسیم شده برای نسبت زوج پیک‌های مربوط به کروماتوگرافی گازی ب) دندوگرام مربوط به نشانه‌گرهای زیستی برای نمونه‌های نفت مخزن سروک

شامل ۲۵/۹۹٪ واریانس کل داده‌ها است. محورهای PC1 و PC2 شامل ۶۶/۶۸٪ کل واریانس مجموعه داده‌ها است. برای نشان دادن ارتباط بین متغیرها و نمونه‌ها در روش تحلیل مؤلفه‌های اصلی می‌توان از نمودار بای پلات^۲ استفاده کرد. در این نمودار، هر متغیر با یک بردار نشان داده می‌شود که نسبت به متغیر دیگر با یک زاویه خاص قرار گرفته است. افزایش این زاویه ارتباط کم بین متغیرها و همبستگی منفی بین آنها را نشان می‌دهد (با افزایش یکی، دیگری کاهش می‌یابد و برعکس). همچنین، کاهش این زاویه ارتباط نزدیک بین متغیرها و همبستگی مثبت آنها را بیان می‌کند. شکل ۹ نمونه‌ای از بای پلات تحلیل مؤلفه‌های اصلی مربوط به نشانه‌گرهای زیستی در این مطالعه است. در این نمودار تأثیر متغیرها بر خوشه‌بندی نمونه‌ها با نزدیک شدن بردار متغیرها به محدوده نمونه‌ها نمایان می‌شود. به‌عنوان مثال نمونه‌های نفتی S06 و S07 برای جدا شدن از سایر نمونه‌ها، بیشترین تأثیر را از نشانه‌گرهای زیستی P1 و P2 گرفته‌اند. همچنین، متغیرهایی که از نمونه‌ها فاصله گرفته و نزدیک به محور مؤلفه‌های اصلی قرار دارند مانند نشانه‌گرهای زیستی P3، P9، P15، P21، نقش در جداسازی داده‌ها نداشته‌اند [۱۹].

هر دو نمودار شباهت زیادی از خوشه‌بندی گروه‌های نفتی را نمایش می‌دهند که مانند روش PCA نمونه‌ها را به سه گروه نفتی مجزا طبقه‌بندی کرده‌اند. برای داده‌های این پژوهش، فاصله اقلیدسی به‌عنوان سنجش فاصله‌ای یا سنجش شباهت بین نقاط به‌کار گرفته شده است که یکی از پرکاربردترین آنها نیز دیاگرام پایپر است. در دیاگرام‌های شکل ۷ الف و ب، خط برش نمودار^۱ به ترتیب در فواصل ۴۸ و ۳۸ به‌صورت اتوماتیک توسط نرم‌افزار XLSTAT ترسیم شده است. شکل ۸ نمودار ترسیم شده حاصل از روش تحلیل مؤلفه‌های اصلی (PCA) برای نشانه‌گرهای زیستی است. همان‌طور که مشاهده می‌شود، نمونه‌های نفتی به خوبی از هم تفکیک شده‌اند (هر خوشه یا یک خانواده نفتی، توسط گروهی از نمونه‌ها به‌کمک یک دایره احاطه شده است). به‌کمک این تحلیل می‌توان گفت که نمونه‌های نفتی S06 و S07 از چاه B در یک خانواده نفتی مجزا، نمونه S18 از چاه D و نمونه‌های S02 و S05 از چاه B مربوط به یک خانواده نفتی دیگر، همچنین نمونه S21 از چاه E، نمونه‌های S08 و S09 از چاه C و نمونه‌های S01، S03 و S04 مربوط به چاه A، در یک خانواده نفتی دیگر قرار گرفته‌اند. در این نمودار بیشترین واریانس داده‌ها بر روی محور PC1، شامل ۴۰/۶۹٪ واریانس کل داده‌ها و محور PC2 بزرگ‌ترین واریانس دوم و

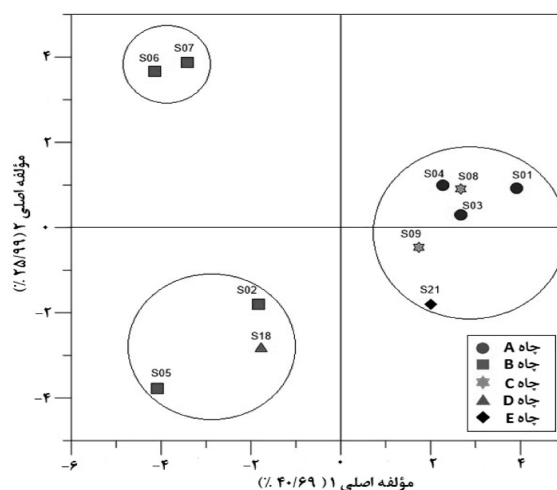
1. Cut off

2. Biplot



شکل ۹ بای پلات حاصل از تحلیل مؤلفه‌های اصلی مربوط به نشانه‌گرهای زیستی

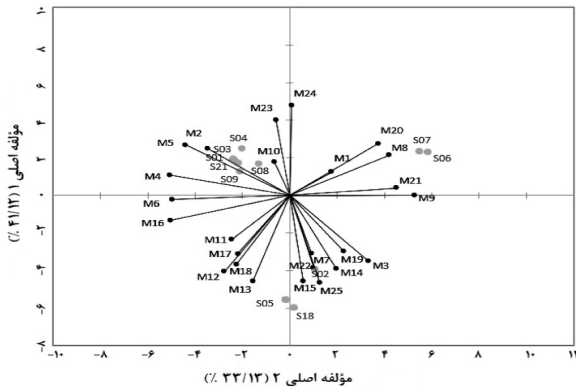
تعدادی از این نمودارها شبیه به خوشه‌بندی داده‌ها در روش‌های آماری عمل کرده و نمونه‌های نفتی را در سه گروه با سطوح بلوغ متفاوت جدا کرده‌اند (شکل ۱۳ الف، ب و ج). این ارتباط خطی و مستقیم بین نشانه‌گرهای زیستی شاخص بلوغ و محور PC1 نشان می‌دهد که اختلاف در بلوغ نفت‌ها منجر به خوشه‌بندی شده است. به این معنی که هر گروه نفتی سطح بلوغ خود را داشته و از گروه دیگر مجزا است. جهت اطمینان بیشتر از نمودارهای شکل ۱۲ (الف و ب)، از ترسیم داده‌های بایومارکری حساس به بلوغ حرارتی در مقابل یکدیگر مانند $Ts/Ts+Tm$ در برابر $\beta\beta/\alpha\alpha+\beta\beta$ C29 sterane (شکل ۱۴) و C32 Homohopane 22S/22S+22R برابر $\beta\beta/\alpha\alpha+\beta\beta$ C29 sterane (شکل ۱۵)، برای نشان دادن روند بلوغ نمونه‌های نفت استفاده شد [۲۰]. همان‌طور که در این دو نمودار مشاهده می‌شود، روند بلوغ نفت‌ها بسیار مشابه با نمودارهای شکل ۱۲ است. شکل ۱۶، ترسیم دو پارامتر Ph/n-C18 در مقابل Pr/n-C17، برای نشانه‌گرهای زیستی است [۲۱]. بر روی این نمودار اطلاعات مفیدی در مورد سنگ منشأ مولد نفت شامل احیایی یا اکسیدی بودن آن، نوع کروژن، میزان بلوغ و تجزیه زیستی نفت، قابل استخراج است.



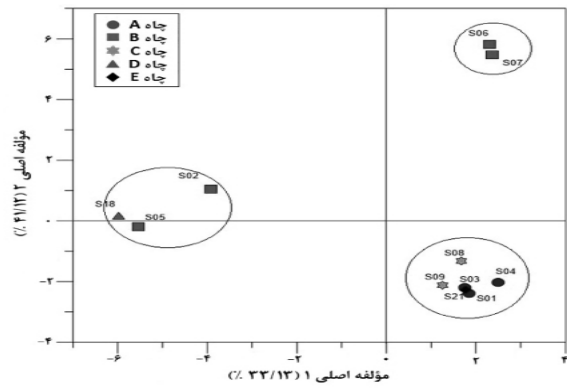
شکل ۸ جداسازی داده‌های مربوط به نشانه‌گرهای زیستی به کمک تحلیل مؤلفه‌های اصلی توسط نمودار PC1 در مقابل PC2

شکل ۱۰ نمودار ترسیم شده PC1 در مقابل PC2 مربوط به نسبت زوج پیک‌های کروماتوگرافی گازی را نشان می‌دهد. همان‌طور که مشاهده می‌شود، در این نمودار، نمونه‌های نفتی همان‌گونه که در تحلیل مؤلفه‌های اصلی برای نشانه‌گرهای زیستی دسته‌بندی شدند (شکل ۸)، در اینجا نیز دسته‌بندی نمونه‌ها به همان شکل صورت گرفته است. نمودار شکل ۱۱ نیز بای پلات متغیرها (نسبت زوج پیک‌ها) و نمونه‌های نفتی مربوط به تحلیل مؤلفه‌های اصلی است. در اینجا همانند شکل ۹، نزدیک شدن بردار مربوط به هر نسبت زوج پیک (M) به هر نمونه، نشان می‌دهد که آن نسبت زوج پیک بیشترین تأثیر خود را بر جداسازی آن نمونه نسبت به سایر نمونه‌ها داشته است.

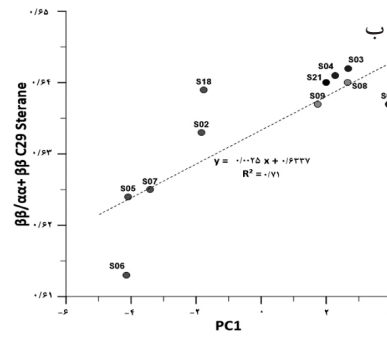
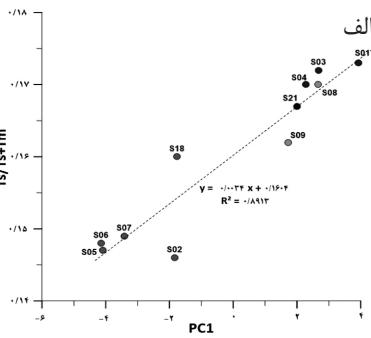
پس از انجام روش‌های آماری بر روی داده‌ها، برای کاهش عدم قطعیت در تعیین گروه‌های نفتی و نشان دادن کارایی بالای روش PCA در این مطالعه، برخی از نشانه‌گرهای زیستی شاخص بلوغ در مقابل محور PC1 که بیشترین واریانس داده‌ها را دارد، ترسیم شدند [۱۹]. نتایج حاصل از آن نشان داد که بین این دو پارامتر رابطه خطی و مستقیم وجود دارد به گونه‌ای که در هر دو نمودار با افزایش مقدار PC1، نشانه‌گرهای زیستی شاخص بلوغ نیز افزایش می‌یابند (شکل ۱۲ الف و ب). همچنین،



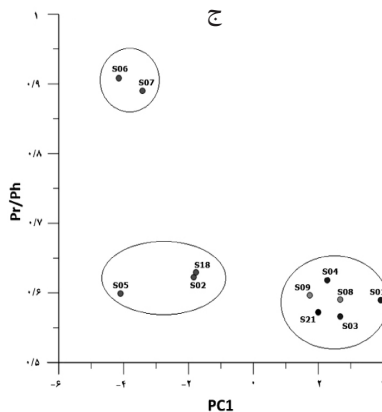
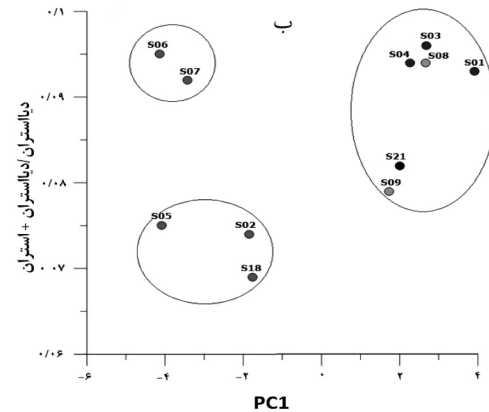
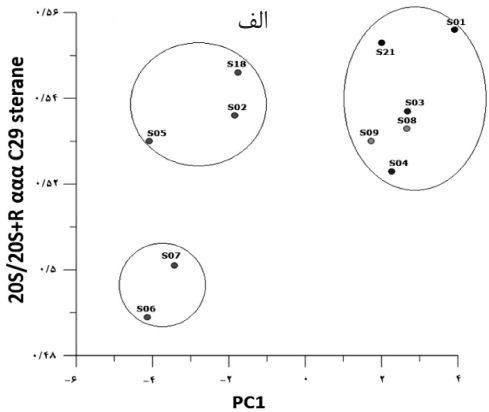
شکل ۱۱ بای پلات حاصل از تحلیل مؤلفه‌های اصلی مربوط به نسبت زوج پیک‌های کروماتوگرافی گازی



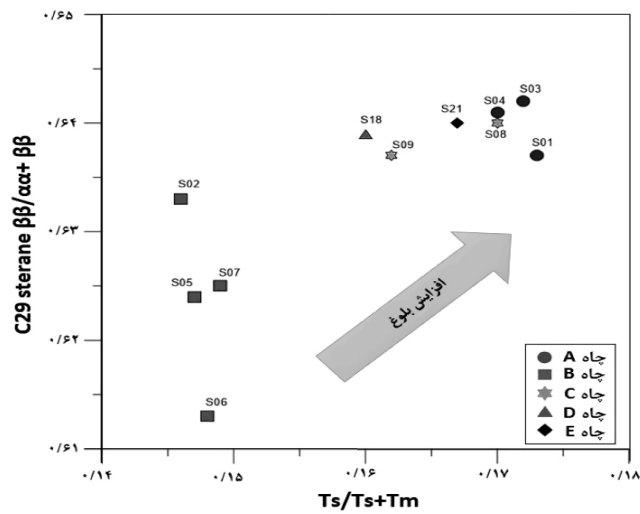
شکل ۱۰ خوشه‌بندی نسبت زوج پیک‌های کروماتوگرافی گازی به کمک تحلیل مؤلفه‌های اصلی توسط نمودار PC1 در مقابل PC2



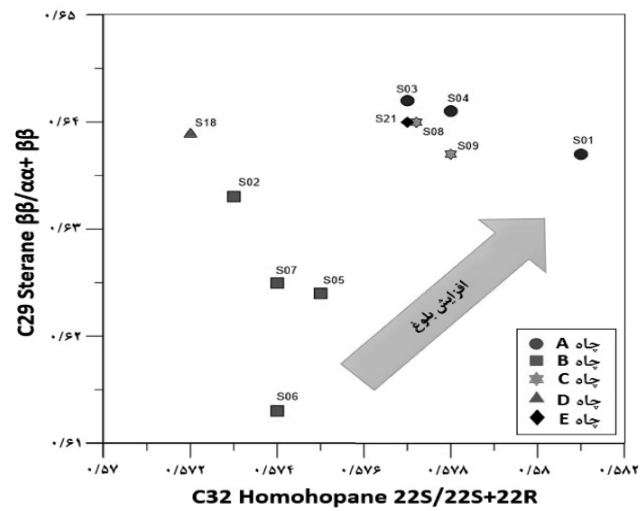
شکل ۱۲ الف) نمودار PC1 در مقابل Ts/Ts+Tm (ب) نمودار PC1 در مقابل $\beta\beta/\alpha\alpha + \beta\beta$ C29 sterane



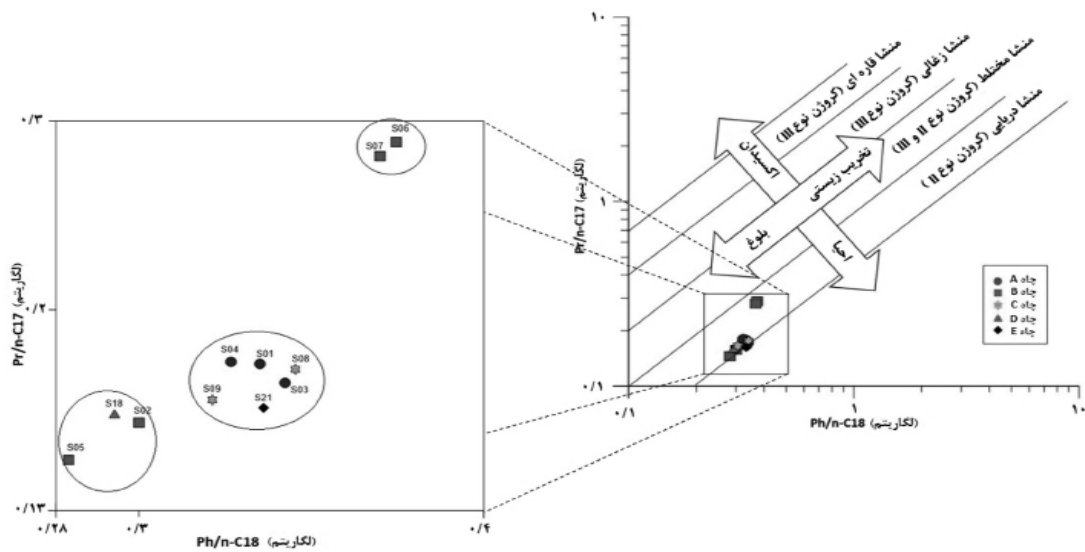
شکل ۱۳ الف) نمودار PC1 در مقابل 20S/20S+R $\alpha\alpha$ C29 sterane (ب) نمودار PC1 در مقابل نسبت دیاستران/دیاستران + استران (ج) نمودار PC1 در مقابل Pr/Ph



شکل ۱۴ نمودار $Ts/Ts+Tm$ در برابر C_{29} sterane $\beta/\alpha+\beta$ برای نشان دادن روند بلوغ نمونه‌های نفت مخزن سروک



شکل ۱۵ نمودار C_{32} Homohopane $22S/22S+22R$ در برابر C_{29} sterane $\beta/\alpha+\beta$ برای نشان دادن روند بلوغ حرارتی نمونه‌های نفت مخزن سروک



شکل ۱۶ نمودار $Ph/n-C18$ در مقابل $Pr/n-C17$ مربوط به داده‌های بایومارکری از نمونه‌های نفت مخزن سروک

خوشه‌بندی داده‌ها در این مطالعه هستند. ۳- کارآیی بیشتر روش PCA نسبت به سایر روش‌های آماری در این مطالعه اینگونه مشخص شد که (۱) با ترسیم مؤلفه PC1، که بیشترین واریانس داده‌ها را در خود جای داده است، در مقابل برخی پارامترهای بایومارکری حساس به بلوغ حرارتی، روندی خطی و مستقیم دارد. (۲) تعدادی از نمودارهای بایومارکری بلوغ حرارتی در مقابل محور PC1، گروه‌های نفتی را دقیقاً مانند روش PCA از یکدیگر تفکیک کرده‌اند که نشان می‌دهد سطح بلوغ هر خوشه (گروه نفتی) با دیگری متفاوت است.

۴- پس از انجام تحلیل‌های آماری بر روی داده‌ها مشخص شد که تمامی نمونه‌های نفتی مخزن سروک می‌توانند به سه گروه تقسیم شوند. نمونه‌های نفتی S06 و S07 از چاه B در یک گروه نفتی مجزا، نمونه S18 از چاه D و نمونه‌های S02 و S05 از چاه B مربوط به یک گروه نفتی دیگر، همچنین نمونه S21 از چاه E، نمونه‌های S08 و S09 از چاه C و نمونه‌های S03، S01 و S04 مربوط به چاه A، دارای سطح بلوغ مشابهی بوده که در گروه نفتی واحدی قرار می‌گیرند.

۵- در این مطالعه، اساساً تفاوت در بلوغ منجر به گروه‌بندی نفت‌ها شده است. این موضوع به کمک ترسیم نمودار نشانه‌گرهای زیستی معرف بلوغ در مقابل محور PC1 تشخیص داده شد. از طرفی خوشه‌بندی داده‌های HRGC، خوشه‌بندی داده‌های مربوط به نشانه‌گرهای زیستی را تایید می‌کند. به این معنی که نفت‌های با بلوغ متفاوت مخزن سروک را پر کرده‌اند. بنابراین می‌توان هر گروه نفتی را به یک افق مخزنی مجزا نسبت داد.

همان‌طور که در بخش با بزرگ‌نمایی بالاتر مشاهده می‌شود، هر گروه از نمونه‌های نفتی که با یک دایره احاطه شده است، سطح بلوغ و تجزیه زیستی متفاوتی را دارد و مهم‌تر از آن این است که مشابه با نمونه‌های نفتی موجود در هر کدام از این گروه‌ها، در خوشه‌های مربوط به روش PCA نیز یافت می‌شود. بنابراین، به کمک این نمودار با اطمینان بیشتری می‌توان گفت که روش آماری PCA یکی از روش‌های مناسب برای مطالعه مخزن سروک در این مقاله بوده است. در اینجا دو نمونه S06 و S07 که از سایر نمونه‌های مربوط به چاه B با اختلاف زیادی جدا شده است، احتمالاً تحت فرآیندهای ثانویه (مانند تجزیه زیستی) قرار گرفته‌اند.

نتیجه‌گیری

۱- در این مطالعه، از سه نوع روش آماری تحلیل مولفه‌های اصلی (PCA)، خوشه‌بندی سلسله‌مراتبی (HCA) و K- میانگین و همچنین، از دیاگرام‌های قطبی (ستاره‌ای)، برای تشخیص پیوستگی مخزن سروک و تعیین گروه‌های نفتی مجزا استفاده شده است. در نهایت، مشخص شد که روش PCA به‌عنوان یکی از بهترین روش‌های خوشه‌بندی داده‌ها در این مطالعه نسبت به سایر روش‌ها است.

۲- دلیل اصلی انتخاب این روش به‌عنوان بهترین روش آماری در این مطالعه این است که (۱) نمودار PCA ترسیم شده برای نشانه‌گرهای زیستی تشابه بسیار زیادی با نمودار PCA ترسیم شده برای نسبت زوج پیک‌های کروماتوگرافی گازی برای همان نمونه‌ها دارد. (۲) روش آماری K- میانگین و دیاگرام‌های ستاره‌ای دارای کاستی‌هایی در

مراجع

- [1]. Snedden J. W., Vrolijk P. J., Sumpter L. T., Sweet M. L., Barnes K. R., White E. and Farrell M. E., "Reservoir connectivity: Definitions, examples, and strategies," In International Petroleum Technology Conference, IPTC, Vol. 2, pp. 772-793, International Petroleum Technology Conference, 2007.

- [2]. Kaufman R. L., Ahmed A. S. and Elsinger R. J., "Gas chromatography as a development and production tool for fingerprinting oils from individual reservoirs: Applications in the Gulf of Mexico," In: Schumaker, D., Perkins, B.F. (eds.), Gulf coast oils and gases: Their characteristics, origin, distribution, and exploration and production significance, Proceedings of the 9th Annual Research Conference of the Society of Economic Paleontologists and Mineralogists (SEPM), New Orleans, pp. 263-282, 1990.
- [3]. Ahanjan A., Rabbani A. R., Khajooie S., "Assessing vertical compartmentalization within the KHM field, southwest of Iran: An integrated approach," Journal of Natural Gas Science and Engineering, Vol. 35, pp. 1277-1283, 2016.
- [4]. Larter S. R., Aplin A. C., "Reservoir geochemistry: methods, applications and opportunities," Geological Society, London, Special Publications, Vol.86, pp 5-32, 1995.
- [5]. Ghassal I., "Reservoir connectivity, water washing and oil to oil correlation: an integrated geochemical & petroleum Engineering Approach," Society of Petroleum Engineers (SPE)e-194957-MS, 2019.
- [6]. Vrolijk P., James B., Myers R., Maynard J., Sumpter L. and Sweet M., "Reservoir connectivity analysis - defining reservoir connections and plumbing," SPE-93577-PP., Society of Petroleum EngineersSourceSPE Middle East Oil and Gas Show and Conference, 12-15 March, Kingdom of Bahrain, 2005.
- [7]. Paez R. H., Lawrence J. J., Zhang M., "Compartmentalization or gravity segregation? Understanding and predicting characteristics of near-critical petroleum fluids," In: Jolley, S.J., Fisher, Q.A., Ainsworth, R.B., Vrolijk, P.J., Delisle, S. (Eds.), Reservoir Compartmentalization, Vol. 347. Geol. Soc. London Spec. Pub., pp. 43e53, 2010.
- [8]. Smalley P. C., England W. A., "Assessing reservoir compartmentalization during field appraisal: how geochemistry can help," SPE Paper No. 25005, pp. 423e443, 1992.
- [9]. Dembick Jr. Harry., 2017. "Practical Petroleum Geochemistry for Exploration and Production Petroleum Geochemistry for Exploration and Production," Elsevier, pp.,1-342 Oct. 6, 2016 .
- [10]. Smalley P. C., England W. A., "Reservoir compartmentalization assessed with fluid compositional data," SPE Reservoir Engineering, Vol. 8, pp. 175-180, 1994.
- [11]. Alizadeh B., Rashidi M., Zarasvandi A., Seyedali S. R. and Aliee M. H., "Hydrocarbon vertical continuity evaluation in the cretaceous reservoirs of Azadegan oilfield, Southwest of Iran: implications for reservoir geochemistry," Acta Geologica Sinica - English Edition, 2019.
- [۱۲]. مطیعی ه.، "زمین شناسی ایران - زمین شناسی نفت زاگرس"، سازمان زمین شناسی و اکتشافات معدنی کشور، ص. ۵۷۲، ۱۳۷۲.
- [۱۳]. آقانباتی س.ع.، "زمین شناسی ایران و کشورهای هم جوار"، سازمان زمین شناسی و اکتشافات معدنی کشور، ص. ۶۷۶، ۱۳۹۲.
- [14]. Murriss R. J., "Middle East: stratigraphic evolution and oil habitat," AAPG Bulletin, Vol. 64, No. 5, pp. 597-618, 1980.
- [15]. Beydoun Z. R., "Arabian plate hydrocarbon geology and potential: A plate tectonic approach," American Association of Petroleum Geologists (AAPG), Tulsa, 1991.
- [16]. Alavi M., "Regional stratigraphy of the Zagros fold-thrust belt of Iran and its proforeland evolution," American

Journal of Science, Vol. 304, No. 1, pp.1-20, 2004.

[17]. Gan G., Ma C., Wu J., "Data clustering: theory, algorithms, and applications," Philadelphia, PA: Society for Industrial and Applied Mathematics (SIAM), Vol. 20, pp. 1- 489 , 2007.

[18]. Jolliffe I., "Principal component analysis," 2nd ed. Springer Series in Statistics. New York: Springer-Verlag. pp. 583-588, 2002.

[19]. Bhullar A. G., Karlsen D. A., Holm K., Backer-Owe K. and Le Tran K., "Petroleum geochemistry of the Fray field and Rind discovery, Norwegian Continental Shelf. Implications for reservoir characterization, compartmentalization and basin scale hydrocarbon migration patterns in the region," Org. Geochem. Vol. 29, No. 1 3, pp. 735 768, 1998.

[20]. Hunt J. M., "Petroleum geochemistry and geology," 2nd ed. W. H. Freeman, p. 743, New York.

[21]. Talukdar S. C., B. De Toni, F. Marcano, J. Sweeney, and A. Rangel. 1993. "Upper cretaceous source rocks of northern South America," American Association of Petroleum Geologists, United States, Vol. 77, No. CONF-930306, 1993.



Petroleum Research

Petroleum Research, 2020(February-March), Vol. 29, No. 109, 23-29

DOI: 10.22078/pr.2019.3842.2750

Application of Cluster Analysis of Oil Geochemical Data to Determine the Connectivity of Sarvak Reservoir in One of the Fields of Southwest Iran

Milad Zohrevand¹, Ali Shekarifard^{2,3*} and Vahid Tavakoli¹

1. Department of Geology, Faculty of Science, University of Tehran, Iran

2. Institute of Petroleum Engineering, College of Engineering, University of Tehran, Iran

3. Department of Oil and Gas Engineering, School of Chemical Engineering, College of Engineering, University of Tehran, Iran

ashekary@ut.ac.ir

DOI: 10.22078/pr.2019.3842.2750

Received: July/4/2019

Accepted: September/18/2019

INTRODUCTION

Cluster analysis is a statistical method that divides a set of objects into homogeneous groups. The main focus of this study is on a statistical methodology that is called principal component analysis (PCA). The main purpose of principal component analysis (PCA) is to reduce the dimensionality of a high-dimensional data set comprising a large number of interrelated variables and at the same time to retain as much as possible of the variation present in the data set [1,2].

In this study, using reservoir geochemistry, studies have been conducted on the evaluation of Sarvak Carbonate Formation (Albian-Turonian) is the second largest reservoir formation in the southwest of Iran after the Asmari (Oligo-Miocene), which it has a significant portion of hydrocarbon reserves. The

Sarvak Formation consists of shallow ramp/low-gradient shelf facies. [3]

The main purpose of this study is reservoir connectivity analysis (RCA) of Sarvak by the oil fingerprinting and biomarkers data to detect compartments and separate petroleum groups. This comparison has been implemented by several methods of cluster analyses. The principal component analysis (PCA) method is more suitable and more effective in this study.

Methodology

This study is based on the analysis and study of eleven oil samples from 5 wells in one of the oil fields of southwest Iran. Due to the small differences between the alkanes in the gas chromatography of the samples, for the precision of their comparison, high resolution gas chromatography (HRGC) method was used

[4]. This will lead to that invisible peaks appearing in higher-resolution gas chromatographs and increasing the precision of study. In this study, the selected peaks are C9 to C20 normal alkanes. This range of hydrocarbons is part of the light oil sector and appears well in most gas chromatographs. Then, the ratio of the height of the pair of peaks was calculated. Also, using the mass spectrometer, the biomarkers achieved to oil-oil correlation such as the series of steranes, diasteranes, terpanes and isoprenoids.

RESULTS AND DISCUSSION

Before performing the PCA, HCA and K-mean methods on the data, first, the comparison

of the peak ratio and biomarkers with a star (polar) diagram was analyzed. In this plot, due to the large variables, the curves have many overlapping points in many places. On the other hand, by performing the K-means method on gas chromatography data, it was determined that the separation of oil families from each other is exactly the same as clustering by PCA and HCA. However, there is little difference for biomarker data, which the difference lead to the ineffectiveness of this method for Biomarkers in this study.

Figures 1 and 2 illustrate the PCA method for biomarkers and gas chromatography peak ratios respectively.

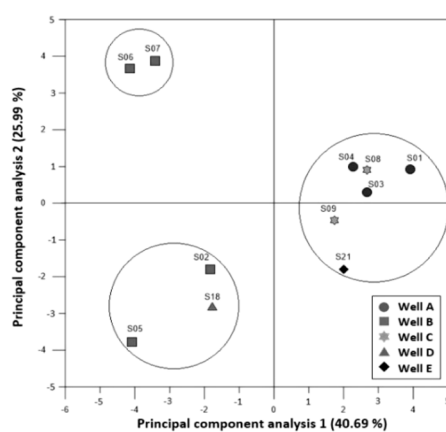


Figure 1: Principal Component Analysis for biomarker data from oil samples.

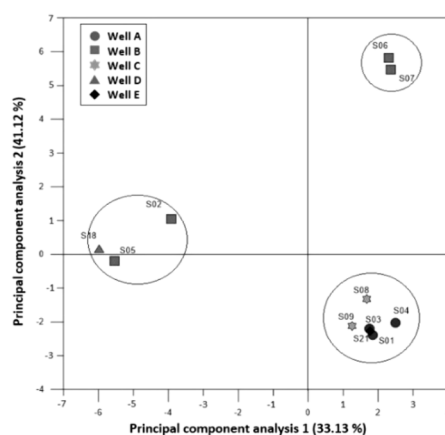


Figure 2: Principal component analysis for gas chromatography peak ratios from oil samples.

As shown in Figures 1 and 2, three oil groups are identifiable. But the PCA method surprisingly show that this clustering is based on different levels of maturity. Figure 3 (a and b) shows the direct relationship between PC1 axis and biomarker maturity parameters. Therefore, cluster analysis based on these data as well as evaluation of the Sarvak reservoir connectivity in this study is impossible.

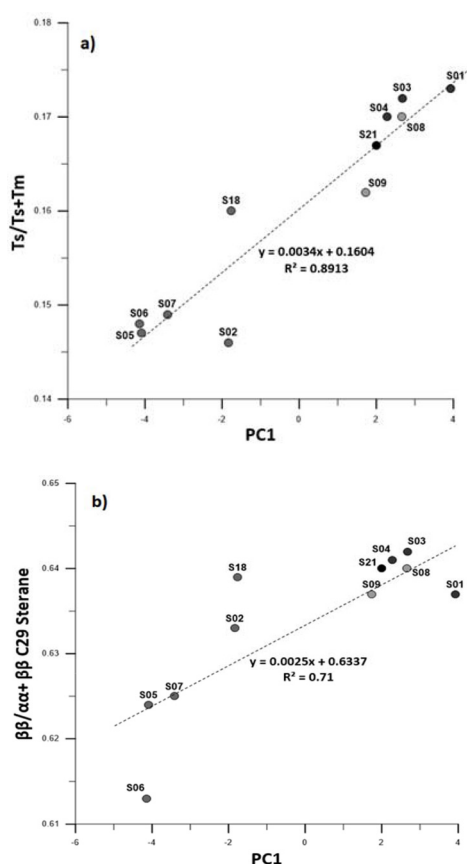


Figure 3: a: Cross-plot of PC1 axis against $Ts/Ts + Tm$. b: cross-plot of PC1 axis against $\beta\beta/\alpha\alpha + \beta\beta C29$ sterane.

CONCLUSIONS

In this study, three types of statistical methods such as principal component analysis (PCA), hierarchical cluster analysis (HCA), K-means in order to data clustering, detection of the connectivity of the Sarvak reservoir and the identification of separate petroleum groups have been used. Finally, PCA method was identified as one of the best methods for data clustering

in this study in comparison with other methods. The main reasons for choosing this method as the best statistical method in this study are as follows: (1) the PCA graphs plotted for biomarkers are very similar to the PCA graphs plotted for the peaks of gas chromatography from the same samples, and (2) some statistical methods such as star diagrams and K-means have low accuracy in data clustering.

After performing PCA method, although it was found out that the data for the cluster analysis in this study have been influenced by the maturation, however with considering the other parameters obtained, presence of different oil families might indicate no/weak connectivity between the Sarvak reservoirs in the studied oil filed.

REFERENCES

- [1]. Jolliffe I., "Principal component analysis," 2nd ed. Springer Series in Statistics. New York: Springer-Verlag. pp. 583-588, 2002.
- [2]. Gan G., Ma C., Wu J., "Data clustering: theory, algorithms, and applications," Philadelphia, PA: Society for Industrial and Applied Mathematics (SIAM), Vol. 20, pp. 1- 489 , 2007.
- [3]. Alizadeh B., Rashidi M., Zarasvandi A., Seyedali S. R. and Aliee M. H., "Hydrocarbon vertical continuity evaluation in the cretaceous reservoirs of Azadegan oilfield, Southwest of Iran: implications for reservoir geochemistry," Acta Geologica Sinica - English Edition, 2019.
- [4]. Kaufman R. L., Ahmed A. S. and Elsinger R. J., "Gas chromatography as a development and production tool for fingerprinting oils from individual reservoirs: Applications in the Gulf of Mexico," In: Schumaker, D., Perkins, B.F. (eds.),

Gulf coast oils and gases: Their characteristics, origin, distribution, and exploration and production significance, Proceedings of the 9th Annual Research Conference of the Society of Economic Paleontologists and Mineralogists (SEPM), New Orleans, pp. 263-282, 1990.