

شبیه‌سازی نگار ژئوشیمیایی با به کارگیری الگوریتم CACS مبتنی بر آنالیز خوشه‌بندی

زیبا حسینی*، علی کدخدائی و اصغر اصغری مقدم

۱- گروه زمین‌شناسی، دانشکده علوم طبیعی، دانشگاه تبریز، تبریز، ایران

تاریخ دریافت: ۹۳/۷/۷ تاریخ پذیرش: ۹۳/۱۱/۱۸

چکیده

یکی از نگارهای مهم ژئوشیمیایی در اکتشافات نفتی، مقدار کل کربن آلی (TOC) می‌باشد که جهت ارزیابی پتانسیل هیدروکربن‌زائی سنگ منشأ استفاده می‌شود. اندازه‌گیری TOC، مستلزم انجام آزمایشات ژئوشیمیایی پرهزینه و وقت‌گیر است، که بر روی نمونه‌های حاصل از حفاری چاهی شامل مغزه و کنده حفاری انجام می‌شود. از این جهت، این پژوهش به برآورد پارامتر ژئوشیمیایی TOC با استفاده از نگارهای پتروفیزیکی که امروزه از تمامی چاه‌ها تهیه شده و کم هزینه‌تر هستند، پرداخته و از روش سیستم کلونی مورچه مبتنی بر آنالیز خوشه‌بندی سلسله مراتبی (hierarchical)، جهت تخمین آن بهره گرفته است. برای تخمین داده‌های TOC به این روش، کد عددی با رویکردی دو مرحله‌ای طراحی شد. مرحله اول شامل خوشه‌بندی سلسله مراتبی و مرحله دوم شامل سیستم کلونی مورچه بوده که به صورت پیوسته با یک تابع توزیع احتمال نرمال طراحی گردید. سیستم کلونی مورچه پیوسته (CACS) ضرائب وزنی از تابع غیرخطی پیشنهادی را برای داده‌های پتروفیزیکی هر خوشه محاسبه کرد، نتایج نشان داد این الگوریتم در تلفیق با خوشه بندی جهت برآورد نگار TOC موفق عمل نموده است. این مطالعه بر روی داده‌های مربوط به سه چاه از میدان نفتی اهواز انجام گرفته و قابل تعمیم به سایر چاه‌های توسعه‌ای خواهد بود.

کلمات کلیدی: سیستم کلونی مورچه پیوسته (CACS)، مقدار کل کربن آلی، میدان نفتی اهواز، نگارهای پتروفیزیکی و نگار ژئوشیمیایی

مقدمه

بدون داشتن چنین اطلاعات ژئوشیمیایی اقدام به حفاری نمی‌کنند [۱]. گسترده‌ترین روش موجود برای تهیه TOC، پیرولیز Rock-Eval می‌باشد که به دلیل هزینه بالای آن و وجود ناپیوستگی در برداشت مغزه‌ها، تنها معدودی از نمونه‌ها با این روش آنالیز می‌شوند. در این راستا، توسعه یک روش سریع و کم هزینه برای تخمین داده ژئوشیمیایی با صحتی قابل قبول اهمیتی منحصر به فرد دارد.

تحلیل نگار ژئوشیمیایی مقدار کل کربن آلی^۱، یکی از مهمترین عوامل در ارزیابی پتانسیل تولید و بررسی خصوصیات ژئوشیمیایی واحدهای هیدروکربنی می‌باشد. اهمیت داده‌های این نگار تا حدی است که امروزه اکثر شرکت‌های نفتی تهیه آن را برای اکتشاف نفت و گاز الزامی دانسته و

1. Total Organic Carbon (TOC)

Hosseini@sadi.ut.ac.ir

*مسئول مکاتبات
آدرس الکترونیکی

نیز از آنالیز خوشه‌بندی در تلفیق با یکی دیگر از سیستم‌های هوشمند که نسبت به الگوریتم پس انتشار^۲ خطا در اکتشاف بهینه مطلق، جوابگوتر است، استفاده شده است.

روش سیستم کلونی مورچه پیوسته^۳ که تاکنون برای بررسی‌های ژئوشیمی اکتشافی و بهره‌برداری نفت بکارگرفته نشده، یکی از سیستم‌های هوشمند می‌باشد که در این پژوهش جهت بررسی تغییرات TOC در سه سازند پابده، گورپی و کژدمی در میدان نفتی اهواز (شکل ۱) از آن بهره گرفته شده است. برای این منظور، برآورد نگار ژئوشیمیایی از روی داده‌های پتروفیزیکی سه چاه ۵۵، ۶۷ و ۲۶۵ از این میدان انجام شد. عمق داده‌های پتروفیزیکی مورد استفاده در جدول ۱ نشان داده شده است. این مطالعه جهت شبیه‌سازی تکامل سنگ‌های مادر و بررسی تاریخچه نفت‌زایی آنها در میدان نفتی اهواز و همچنین جهت تعمیم به سایر چاه‌های توسعه‌ای این میدان کاربرد دارد.

روش تحقیق

نرم‌افزار MATLAB

در مطالعه حاضر از نرم‌افزار MATLAB نسخه ۷/۱۲ جهت خوشه‌بندی سلسله مراتبی و طراحی کد الگوریتم CACS استفاده شده است. نرم‌افزار MATLAB، یکی از نرم‌افزارهای پر قدرت برای انجام محاسبات طولانی، داده‌کاوی و طراحی انواع الگوریتم‌ها است. این نرم‌افزار یک زبان برنامه نویسی آسان با ویژگی‌های پیشرفته بوده که استفاده از آن ساده‌تر از سایر زبان‌های برنامه‌نویسی همانند PASCAL و یا C می‌باشد [۱۵]. پیش فرض MATLAB مفهوم ماتریسی برای اعداد است، بدین معنی که برای انجام خوشه‌بندی کلیه داده‌ها در قالب ماتریس در حافظه نرم‌افزار ثبت می‌شود.

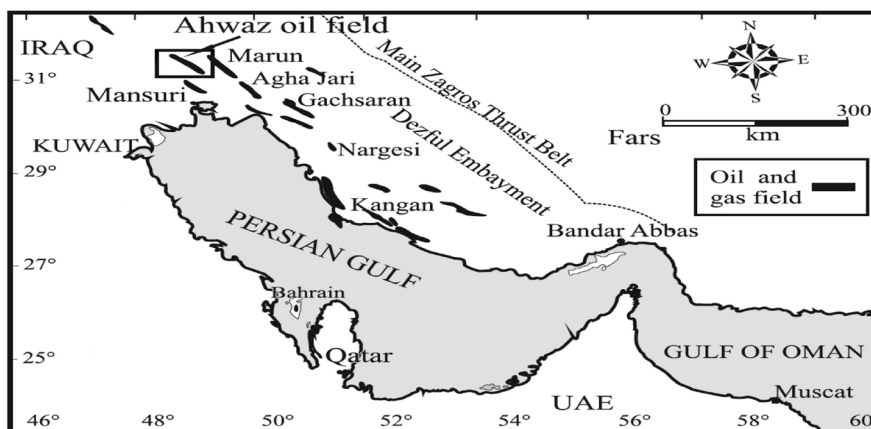
داده‌های پتروفیزیکی از دیگر اطلاعات رایج و با ارزش در صنایع نفت می‌باشند که نسبت به داده‌های ژئوشیمیایی کم هزینه بوده و برای بیشتر چاه‌ها قابل دسترس هستند. بنابراین استحصال داده‌های ژئوشیمیایی با استفاده از داده‌های پتروفیزیکی امری ارزشمند است که باعث صرفه‌جویی در زمان و هزینه خواهد شد. برای رسیدن به این هدف، تاکنون محققان زیادی سعی در برقراری ارتباط کیفی و کمی بین داده‌های TOC با داده‌های حاصل از نگارهای پتروفیزیکی نموده‌اند. از جمله این افراد می‌توان به مراجع [۲-۶] اشاره نمود که از نگار طیفی GR برای شناسایی سنگ‌های غنی از ماده آلی استفاده کردند. دلنباخ و همکارانش در سال ۱۹۸۳ و همچنین هوسین در سال ۱۹۸۷ به‌وسیله منحنی‌های زمان عبور صوت و اشعه گاما روشی ابداع کردند که پارامتر حاصل از آن با میزان غنی‌شدگی ماده آلی ارتباطی خطی داشت [۷] و [۸]. پس از آنها، پسی و همکارانش در سال ۱۹۹۰ تکنیک "ΔlogR" را ابداع نموده و با این روش TOC را محاسبه کردند [۹]. کمالی و میرشادی نیز در سال ۲۰۰۴ از روش نروفازی جهت تخمین TOC از داده‌های پتروفیزیکی بهره گرفته و نتایج حاصل از آن را با تکنیک "ΔlogR" مقایسه نمودند [۱۰].

کخدائی و همکارانش در طی سال‌های ۲۰۰۸ و ۲۰۰۹ با ابداع یک ماشین حاصل از عملکرد سیستم‌های هوشمند به تخمین TOC پرداختند [۱۱ و ۱۲]. سرانجام علیزاده و همکارانش در سال ۲۰۱۱ و سفیداری و همکارانش در سال ۲۰۱۲ از ترکیب آنالیز خوشه‌بندی^۱ با شبکه عصبی مصنوعی برای رسیدن به این هدف بهره گرفتند [۱۳ و ۱۴]. مطالعات آنها نشان داد که خوشه‌بندی باعث ایجاد خوشه‌هایی با سنگ‌شناسی تقریباً یکنواخت می‌شود که در نتیجه آن سیستم‌های هوشمند در مقایسه با قبل از خوشه‌بندی (سنگ‌شناسی غیر یکنواخت) با میانگین مربعات خطای پایین‌تر قادر به برآورد TOC می‌باشند. بنابراین، در این پژوهش

1. Clustering Analysis

2. Back Propagation (BP)

3. Continuous Ant Colony System (CACS)



شکل ۱- موقعیت میدان نفتی اهواز در فروافتادگی دزفول [۱۵]

جدول ۱- عمق داده‌های پتروفیزیکی در چاه ۵۵، ۲۶۵ و ۶۷ میدان نفتی اهواز

عمق داده‌های پتروفیزیکی (m)	سازند	نام چاه
۳۱۵۳-۳۲۱۵	پابده	۵۵
۲۹۳۱-۳۲۶۰	پابده و گورپی	۲۶۵
۴۲۲۱-۴۵۳۲	کزدمی	۶۷

هستند [۱۸]. اندازه‌گیری میزان این شباهت بستگی به نوع کاربرد آن دارد. آنالیز خوشه‌ای به روش‌های مختلفی صورت می‌گیرد، از جمله آن، خوشه‌بندی به روش سلسله مراتبی است، که در این پژوهش از آن استفاده شد. خوشه‌بندی سلسله مراتبی یک ابزار برای تجزیه و تحلیل داده در مقیاس‌های مختلف است که به طور گسترده مورد استفاده قرار می‌گیرد. این خوشه‌بندی با شروع شباهت دوتایی از داده آغاز می‌شود که پی در پی داده‌های مشابه را در خود قرار می‌دهند. این عمل همانند شاخه‌هایی هستند که کم کم به هم می‌پیوندند و تجسمی از یک درخت را می‌سازند [۱۹]. تصور این درخت در خوشه‌بندی سلسله مراتبی با نمایش یک دندروگرام که خلاصه‌ای مفید از این نوع خوشه‌بندی را فراهم می‌آورد به تحقق می‌پیوندد. دندروگرام به ما اجازه می‌دهد که تصمیم بگیریم کدام سطح یا مقیاس از خوشه‌سازی برای موضوع مورد نظر مناسبتر است که سپس تجزیه و تحلیل داده‌ها بر اساس آن سطح انجام گیرد.

پس از انجام خوشه‌بندی نیز، داده‌های پتروفیزیکی هر خوشه در قالب یک ماتریس به عنوان ورودی و TOC حاصل از آنالیز در همان خوشه به عنوان ماتریس خروجی مطلوب شبکه تعیین و برای اجرای شبیه‌سازی اعمال می‌شود. توانایی نرم‌افزار MATLAB در برخورداری از قابلیت گرافیکی قوی، تعریف شدن اکثر توابع ریاضی و داشتن ابزارهایی برای انجام برنامه‌های مختلف، آن را به یک برنامه کاربردی و هوشمند برای حل مسائل پیچیده تبدیل کرده است [۱۷]. در بحث و بررسی به چگونگی اعمال الگوریتم CACS مبتنی بر خوشه‌بندی سلسله مراتبی در این نرم‌افزار، پرداخته شده است.

خوشه‌بندی سلسله مراتبی

آنالیز خوشه‌بندی از جمله کارآمدترین روش‌ها در خوشه‌بندی و آنالیز آماری داده‌ها است. هدف آنالیز خوشه‌ای اندازه‌گیری شباهت و عدم شباهت بین داده‌ها و قرار دادن آنها در گروه‌هایی است که از لحاظ داخلی همگن و از لحاظ خارجی ناهمگن

انتخاب شدن تقریباً یکسانی دارند. پس از ارزیابی تابع هدف و بروزرسانی فرومون در طول هر تکرار به تدریج این رفتار کاوشی الگوریتم کاهش، و رفتار استخراجی الگوریتم باتوجه به تجمع فرومون حول بهترین ناحیه که مرکز این تابع است، افزایش می‌یابد. افزایش فرومون سبب افزایش احتمال انتخاب مسیر بهینه توسط مورچه می‌گردد و بهترین جواب مسئله محاسبه می‌شود. مزیت اصلی CACS، سادگی در ساختار و داشتن یک پارامتر کنترلی (فقط تعداد مورچه) می‌باشد که به این پارامتر هم حساسیت زیادی نشان نداده و در بازه متنوعی از آن می‌تواند دقت نسبتاً خوبی را ارائه نماید. البته مقادیر کم مورچه برای CACS خیلی قوی و پاسخگو نیست، چراکه احتمال گیرافتادن در بهینه محلی افزایش می‌یابد و به عبارتی پراکندگی تابع توزیع فرومون کاهش یافته و دیگر پرش به یک مینیمم محلی بهتر مقدور نیست و کمینه‌سازی متوقف می‌گردد [۲۶].

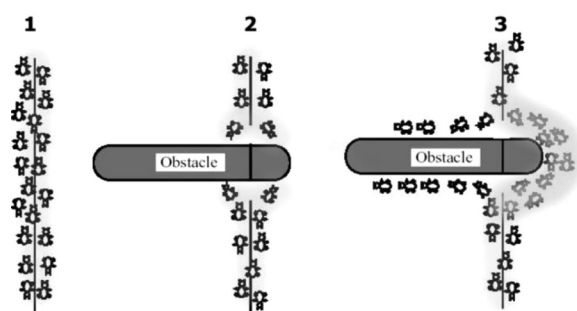
بحث و بررسی

منابع داده

به منظور شبیه‌سازی نگار ژئوشیمیایی با سیستم هوشمند در یک میدان نفتی، تهیه یک کاتالوگ داده‌ای مناسب شامل اطلاعات داده ژئوشیمیایی حاصل از آنالیز مغزه و همچنین داده‌های پتروفیزیکی کلیه چاه‌های مورد مطالعه اولین و اساسی‌ترین مرحله است. کاتالوگ داده‌ای پژوهش حاضر از شرکت ملی مناطق نفت‌خیز جنوب از سه چاه ۵۵، ۶۷ و ۲۶۵ در میدان نفتی اهواز تهیه گردید. هنگامی که با مجموعه‌های بزرگ داده جهت تحلیل‌های ژئوشیمیایی روبرو هستیم به سادگی نمی‌توانیم تکنیک‌ها و الگوریتم‌ها را روی آنها پیاده کنیم چرا که داده‌های اولیه دارای مشکلات بسیاری می‌باشد. از طرفی، ارتباط قوی بین داده‌های ورودی و خروجی، می‌تواند پیشگویی دقیق‌تری را نسبت به داده‌های

الگوریتم CACS

الگوریتم کلونی مورچه^۱ (ACO) [۲۰]، اولین الگوریتم مبتنی بر رفتار مورچه‌ها برای حل مسائل بهینه‌سازی ترکیبی بوده است. پس از آن در دهه‌های اخیر انواع مختلفی از آن بهبود یافته و به صورت کاربردی ارائه شدند، که سیستم کلونی مورچه پیوسته^۲ (الگوریتم CACS) برای تطبیق الگوریتم ACO به فضای جستجوی پیوسته اولین بار توسط پورتاکدوست و نوبهاری در سال ۲۰۰۴ ارائه گردید [۲۱]. تمامی انواع الگوریتم ACO بر شیوه‌ای استوار هستند که در آن مورچه‌ها به دنبال غذا می‌گردند. در طول سفر مورچه‌ها، یک اثر شیمیایی به نام فرومون^۳ بر زمین گذاشته می‌شود، که نقش آن، راهنمایی دیگر مورچه‌ها به سوی نقطه مطلوب (محل غذا) است. برای یک مورچه، میزان احتمال انتخاب هر مسیر، طبق اندازه و کمیت فرومون آن مسیر انجام می‌گیرد (شکل ۲). همان‌گونه که در شکل ۲ به تصویر کشیده شده است، مسیر کوتاه‌تر نیاز به زمان سفر کمتری دارد و در یک بازه زمانی، با سطح بالاتری از فرومون خاتمه یافته و به این طریق، این الگوریتم‌ها بهترین جواب مسأله را به دست می‌دهند [۲۲].



شکل ۲- تأثیر فرومون روی انتخاب مسیر بهینه توسط مورچه‌ها [۲۳]

الگوریتم CACS برای توزیع و پیاده‌سازی این مدل فرومون در فضای پیوسته از تابع توزیع احتمال نرمال^۴ که به آسانی توسط روش‌های Box-Muller و الگوریتم Ziggurat عنوان گردیده است، استفاده می‌نماید [۲۴ و ۲۵]. در مراحل اولیه اجرای الگوریتم، تمام نقاط فضای جستجو احتمال

1. Ant Colony Optimization
2. Continuous Ant Colony System
3. Pheromone
4. Normal Probability Distribution Function (pdf)

نوترون، گاما و صوتی) که با استفاده از خوشه‌بندی سلسله مراتبی به ۴ خوشه مجزا تقسیم شده‌اند، بهره گرفته شد. داده‌های پتروفیزیکی انعکاسی از عمق و دیگر خصوصیات هریک از سازندها را ارائه می‌دهند، تأثیر مستقیم داده‌های عمق باعث ایجاد خوشه‌های بهینه‌تری می‌گردد [۲۷]. بنابراین جهت خوشه‌بندی سلسله مراتبی داده‌های پتروفیزیکی به همراه عمق معادلشان لحاظ گردیده است. دندروگرام حاصل از کل داده‌های این خوشه‌بندی در شکل ۳ نشان داده شده است.

مجموعه داده‌ها همانطور که گفته شد شامل ۱۴۳ داده آموزشی جهت ساخت مدل و ۶۰ داده تست جهت صحت‌سنجی آن می‌باشد، که این مجموعه داده‌ها از سه چاه ۵۵، ۶۷ و ۲۶۵ در میدان نفتی اهواز تهیه شده‌اند. در جدول ۲ مقدار حداکثر و حداقل مجموعه داده‌های ورودی نشان داده شده است. برای ارزیابی و مقایسه بهتر، کلیه داده‌های ورودی و خروجی به روشی که لاروس مطرح نموده است، مطابق فرمول‌های زیر، در بازه [L, H] که در این پژوهش معادل [۰, ۱] انتخاب گردید، نرمالایز شدند [۲۸]. بازه انتخاب شده برای نرمال‌سازی داده‌ها مرسوم‌تر است.

$$X^* = mX_i + b \quad (1)$$

$$m = \frac{H-L}{\text{Max}(X) - \text{Min}(X)} \quad (2)$$

جدول ۲- مقادیر حداکثر و حداقل مجموعه داده‌های مورد استفاده در چاه‌های مورد مطالعه

Inputs	Min.	Max.
^۲ GR	۵/۳۲	۸۵/۲۳
^۳ DT	۵۱/۹۵	۱۱۸/۸۹
^۴ CNL	-۲/۴۷	۴۲/۷۸
^۵ LLD	۰/۱۸	۱/۷۹
Measured TOC	۰/۱۷	۴/۳۷

1. Hydrogen Index
2. Delta Gamma-Ray
3. Transit Time (ΔT)
4. Compensated Neutron Log
5. DLL (Depth Laterolog)

با تطابق ضعیف‌تر ایجاد کند. به همین جهت در این پژوهش، جهت بالا بردن کیفیت داده‌های اولیه و بروز آشکارتر روابط بین آنها اقداماتی انجام شده است که از جمله آن می‌توان به موارد زیر اشاره کرد:

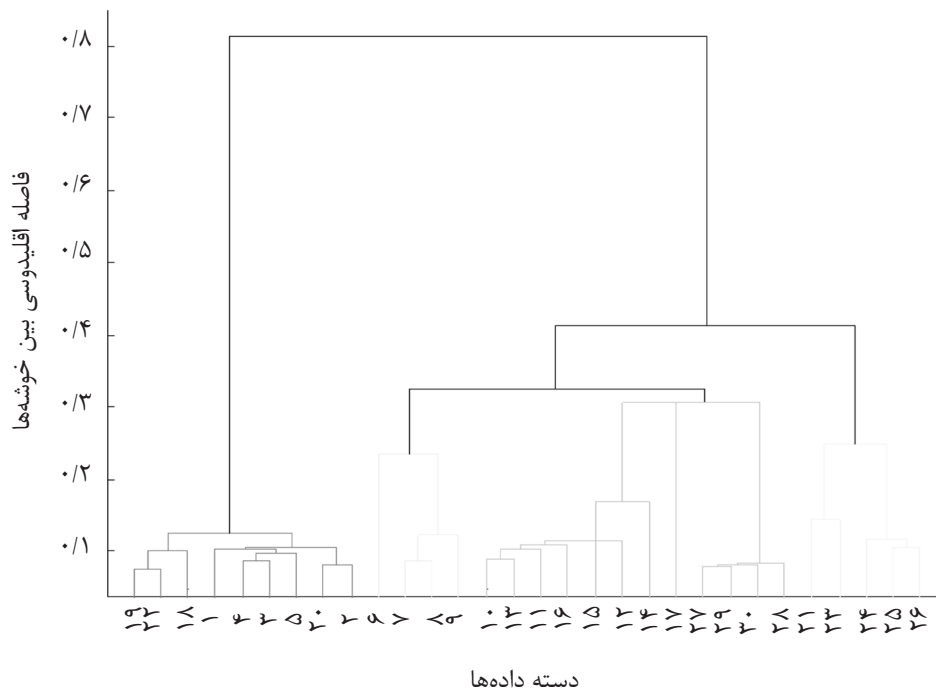
۱- بررسی تطبیق صحیح عمق داده‌های پتروفیزیکی و ژئوشیمیایی بر اساس روابط بین آنها
 ۲- تصحیح مقدارهای ناصحیح و بی‌معنی در محاسبات
 ۳- حذف داده‌های تکراری

۴- حذف پارامترهای غیر ضروری پتروفیزیکی برای جلوگیری از Over parametering. پس از انجام مراحل بالا جهت پاک‌سازی، داده‌های پتروفیزیکی معادل با ۲۰۳ داده آنالیز شده برای اندازه‌گیری TOC مرتب گردید که شامل نگار مقاومت، نوترون، گاما و صوتی است. ۱۴۳ داده از این مجموعه جهت ساخت مدل و ۶۰ داده نیز جهت صحت‌سنجی به صورت تصادفی برگزیده شد. در بخش زیر توجهی مبنی بر ارتباط بین این چهار نگار پتروفیزیکی و داده ژئوشیمیایی TOC آمده است:

لایه‌های دارای مواد آلی به خصوص زمانی که کروژن بالغ بوده و هیدروکربن تولید شده فضای خالی و شکستگی‌ها را پر می‌کند، مقاومت بیشتری نسبت به دیگر سنگ‌ها دارند و در واقع نگار نوترون اندازه‌گیری‌هایی از غلظت هیدروژن است که فاز کاهش سرعت آن به شدت وابسته به غلظت هیدروژن و شاخص هیدروژن^۱ (HI) در سازند است. نگار گاما نیز اندازه‌گیری‌هایی از تراکم مواد رادیو اکتیو (پتاسیم، توریم و اورانیوم) است که در سنگ‌های غنی از مواد آلی افزایش می‌یابد. افزایش مقدار ماده آلی همچنین سبب افزایش مدت زمان عبور صوت شده و بنابراین نگار صوتی نیز افزایش می‌یابد [۱۰].

تعمیم و اجرای الگوریتم

در این پژوهش از الگوریتم CACS برای تولید محتوی کربن آلی، از چهار نگار پتروفیزیکی (مقاومت،



شکل ۳- دندروگرام حاصل از کل داده‌های پتروفیزیکی با به‌کارگیری خوشه‌بندی سلسله‌مراتبی

شده توسط سیستم کلونی مورچه پیوسته و N تعداد نمونه‌ها می‌باشند.

این الگوریتم جهت پیاده‌سازی مدل فرومون پیوسته از تابع توزیع احتمال نرمال، طبق معادله ۶ استفاده می‌نماید:

$$\tau(x) = e^{-\frac{(x-x_{\min})^2}{2\sigma^2}} \quad (6)$$

در این معادله، x_{\min} بهترین نقطه یافت شده در بازه جستجو و σ یک شاخص از پراکندگی مورچه‌ها دور مینیمم فعلی است که به صورت معادله ۷ محاسبه می‌گردد:

$$\sigma^2 = \frac{\sum_{j=1}^k \frac{1}{(f_j - f_{\min})} (x_j - x_{\min})^2}{\sum_{j=1}^k \frac{1}{(f_j - f_{\min})}} \quad (7)$$

در این معادله، k تعداد کل مورچه‌ها، f همان تابع هدف و f_j مقدار تابع هدف برای مورچه شماره j ، بهترین مقدار برای تابع هدف، x_j جواب مسئله برای مورچه j ام و x_{\min} بهترین جواب مسئله در هر تکرار می‌باشد. تابع توزیع احتمال نرمال (معادله ۶) شبیه یک زنگوله است که نوک زنگوله (x_{\min}) دارای حداکثر مقدار فرومون است. برای به

$$b = \frac{\text{Max}(X)L + \text{Min}(X)H}{\text{Max}(X) - \text{Min}(X)} \quad (3)$$

در این رابطه X^* متغیر نرمال‌سازی شده و X_i متغیر اصلی است. طراحی کد الگوریتم CACS نیازمند پیاده‌سازی یک مدل ریاضی صحیح است که با ارائه مجموعه‌ای از معادلات محقق می‌شود. بنابراین برای این الگوریتم کد عددی به‌کارگرفته شد، که جهت پیش‌بینی مقدار کل کربن آلی هر خوشه با استفاده از داده‌های پتروفیزیکی از معادله زیر استفاده می‌نماید:

$$\text{TOC}_{\text{CACS}} = C_1 \cdot \text{GR}^{C_2} + C_3 \cdot \text{DT}^{C_4} + C_5 \cdot \text{LLD}^{C_6} + C_7 \cdot \text{CNL}^{C_8} + C_9 \quad (4)$$

وزنی و نمایی برای ورودی‌های GR، DT، LLD، CNL هستند و C_9 مقدار ثابت معادله است.

تابع هدف نیز طبق فرمول زیر توسط سیستم کلونی مورچه پیوسته برای هر خوشه مینیمم می‌گردد:

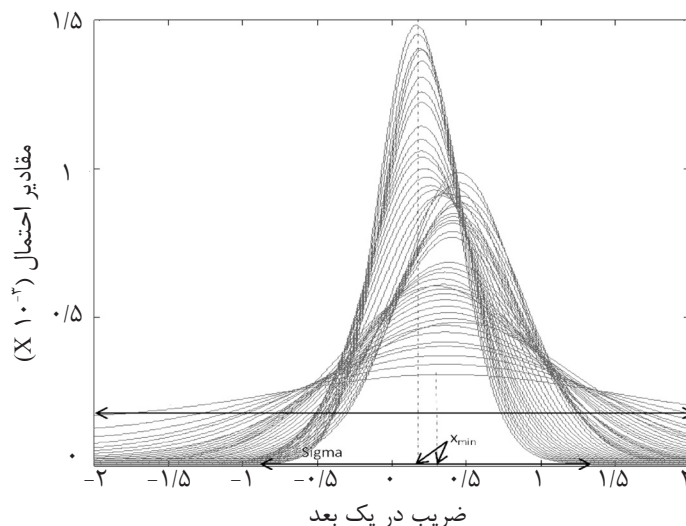
$$\text{MSE} = \sum_{i=1}^N \frac{1}{N} (\text{TOC}_{\text{real}} - \text{TOC}_{\text{CACS}})^2 \quad (5)$$

در این معادله MSE میانگین مربع خطا، TOC_{real} مقدار کل کربن آلی اندازه‌گیری شده توسط پیرولیز Rock-Eval، TOC_{CACS} مقدار کل کربن آلی پیش‌بینی

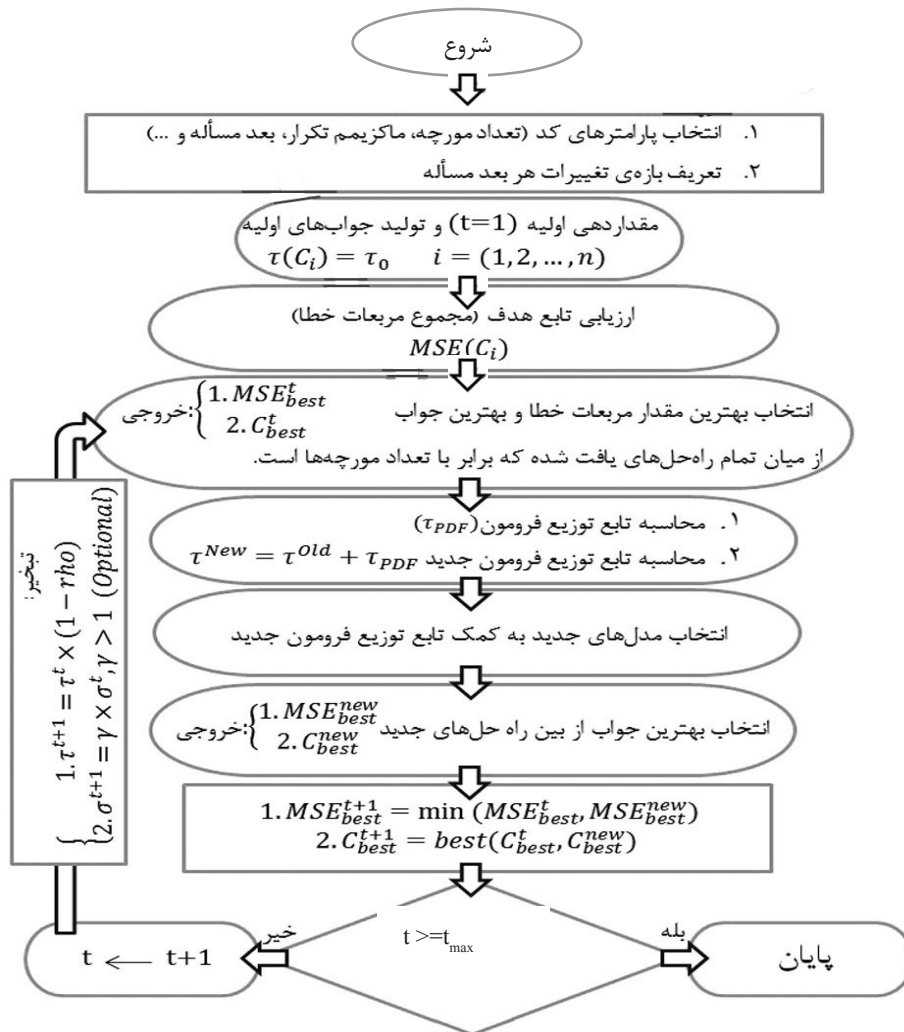
الگوریتم زیاد می‌باشد. در تکرارهای انتهایی که خاصیت استخراجی الگوریتم مد نظر قرار می‌گیرد مقدار σ به صفر نزدیک می‌شود. این نوع عملکرد سیستم کلونی مورچه پیوسته انعطاف زیادی در یافتن ضرایب بهینه معادله ۴ ایجاد می‌نماید و در نهایت بهترین جواب برای مسئله پیدا می‌شود. نمایشی از ۵۰ تکرار اول تابع توزیع احتمال نرمال برای یک بعد مسئله در شکل ۴ آورده شده است. همانطور که این شکل نیز نشان می‌دهد با افزایش تکرار، کم کم خاصیت استخراجی الگوریتم افزایش می‌یابد.

فلوچارت کد نوشته شده برای الگوریتم CACS، در شکل ۵ و شکل ۶ به تصویر کشیده شده است. در فلوچارت شکل ۵، t تعداد تکرار تا مرحله کنونی، i بعد مسئله، τ_0 مقدار فرومون در لحظه صفر برای تمام ابعاد مسئله و ρ ضریب تبخیر می‌باشد. γ نیز ضریبی ثابت است که در مرحله تبخیر آورده شده است. مفهوم اصلی تبخیر جلوگیری از تجمع فرومون در یک نقطه می‌باشد که از باریک شدن سریع زنگوله تابع توزیع احتمال نرمال در این الگوریتم، جلوگیری می‌نماید. بنابراین زمانیکه γ بزرگتر از ۱ باشد، به الگوریتم این فرصت را خواهد داد تا با تکرارهای بیشتر تمام نقاط دیگر بازه جستجو را ارزیابی کند. تنظیم شاخص‌های اصلی الگوریتم طراحی شده برای تخمین TOC نیز در جدول ۳ آورده شده است.

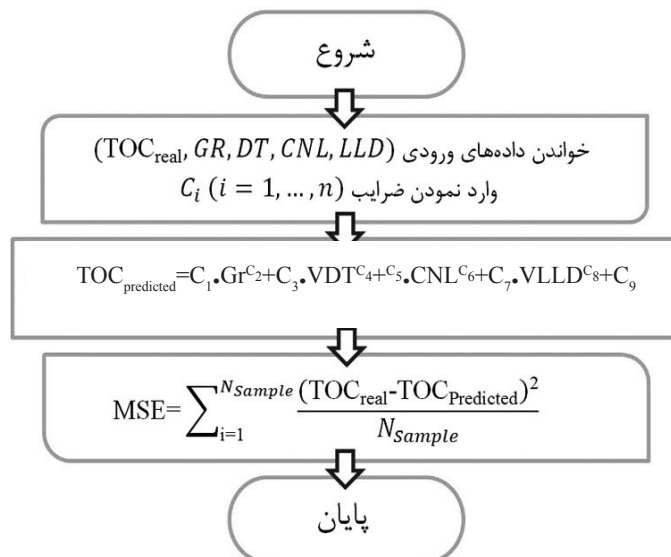
کارگیری این الگوریتم، کد به نحوی باید طراحی شود که در مرحله اول x_{min} در بازه جستجو به صورت تصادفی انتخاب گردد. در این پژوهش بازه جستجو برای تمامی ابعاد مسئله $[-2, 2]$ انتخاب شده است، چراکه بازه‌های خیلی بزرگ سرعت الگوریتم در همگرایی جواب را کاهش می‌دهند. تحت تأثیر x_{min} ثبت شده برای هر بعد مسئله در حافظه الگوریتم، در هر مرحله فرومون بر اساس تابع توزیع احتمال نرمال در مسیر بهینه افزایش و برعکس در مسیرهایی که سبب بهبود تابع هدف نشده‌اند تبخیر شده و کاهش می‌یابد. در کد طراحی شده این الگوریتم عمل تبخیر فرومون با دستور یک ضریب تبخیر مشخص در هر مرحله اعمال می‌شود. تعیین مقدار ضریب تبخیر فرومون در کد طراحی شده، دارای اهمیت خاصی است. زیرا زمانی که مقدار تبخیر خیلی بالا باشد، الگوریتم اغلب به مسیر بهینه محلی همگرا می‌گردد. بنابراین ضریب تبخیر فرومون در این پژوهش ۰/۱ در نظر گرفته شد که در اکثر موارد به بهینه‌ترین جواب مسئله می‌رسد. در حین اجرای الگوریتم، با بروزرسانی فرومون مقدار σ در تکرارهای بعدی کوچکتر شده که به منزله زنگوله باریکتر برای تابع توزیع احتمال نرمال است. به عبارت دیگر مقدار σ در طول تکرارها به نحو پویایی تغییر می‌کند. در تکرارهای اولیه مقدار آن بیشتر است که به این دلیل خاصیت اکتشافی



شکل ۴- نمایشی از ۵۰ تکرار اول تابع توزیع احتمال نرمال در یک بعد مسئله



شکل ۵- فلوجارت کد طراحی شده برای سیستم کلونی مورچه پیوسته



شکل ۶- فلوجارت نحوه ارزیابی تابع هدف برای هر مورچه

جدول ۳- تنظیم شاخص‌های اصلی سیستم کلونی مورچه پیوسته

تنظیمات	شاخص‌های الگوریتم CACS
۹	بعد مسئله
$[-۲,۲]$	بازه جستجو در هر بعد
Rolette-Wheel	الگوریتم احتمال
Normal PDF	تابع توزیع احتمال
۱۰۰۰۰	تعداد مورچه
۱۲۰	ماکزیمم تکرار
۰/۱	ضریب تبخیر

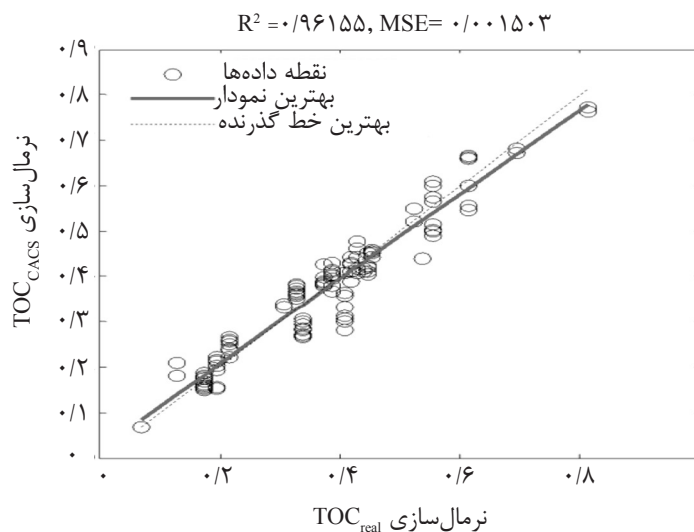
مربوط به هر خوشه، n_i تعداد نمونه در هر خوشه و k تعداد کل خوشه‌ها می‌باشد.

مقدار میانگین مربع خطا و ضریب همبستگی بین TOC واقعی نرمال و TOC برآورد شده توسط سیستم کلونی مورچه پیوسته بعد از خوشه‌بندی، برای داده‌های آموزشی (۱۴۳ نمونه) به ترتیب، $۰/۰۰۱۵۰۳۳$ و $۰/۹۶۱۵۵$ محاسبه شد (شکل ۷). این مقادیر برای داده‌های تست (۶۰ نمونه) نیز به ترتیب $۰/۰۰۲۰۸۰$ و $۰/۹۵۸۲$ به دست آمده است (شکل ۸). بنابراین این مدل می‌تواند به عنوان یک مدل بهینه برای برآورد نگار TOC از داده‌های پتروفیزیکی اهمیت بسزایی داشته باشد.

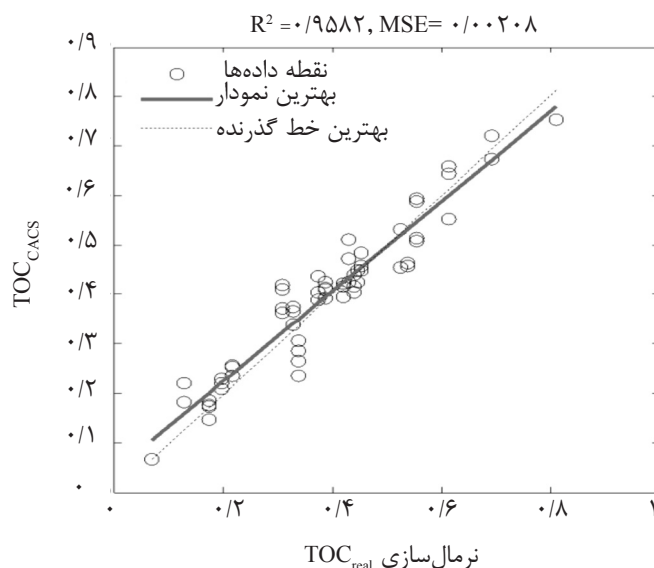
پس از اجرای برنامه برای داده‌های آموزشی، الگوریتم CACS بهترین ضرایب را برای هر خوشه برآورد می‌کند. سپس این ضرایب بر روی داده‌های تست که فقط شامل داده‌های پتروفیزیکی می‌باشند اعمال شده و TOC پیش‌بینی می‌شود. در انتها نتایج خوشه‌ها با هم ترکیب شده و مقدار میانگین مربع کل خطا (MSE_t) طبق معادله ۸ و همچنین ضریب همبستگی^۱ برای هر خوشه‌بندی به عنوان معیاری از کارایی مدل، محاسبه می‌گردد.

$$MSE_t = \frac{\sum_{i=1}^k MSE_i \times n_i}{\sum_{i=1}^k n_i} \quad (۸)$$

در این معادله، MSE_i مقدار میانگین مربع خطا



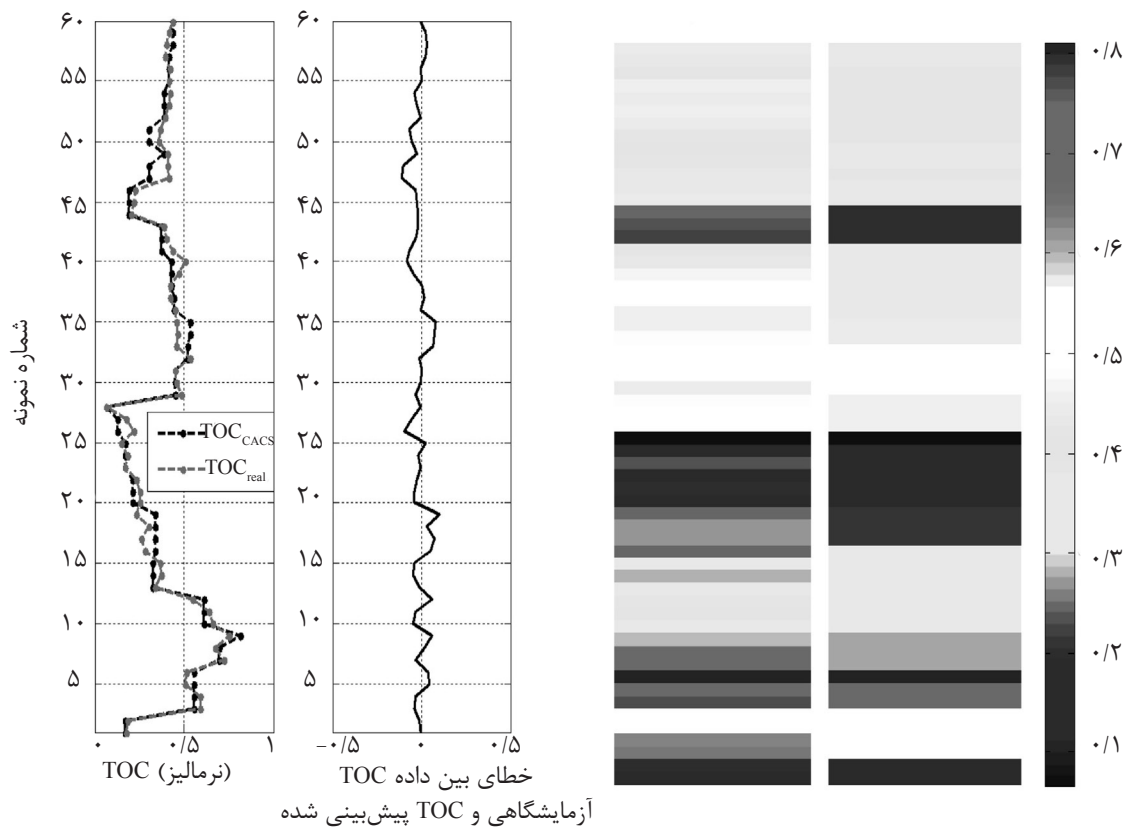
شکل ۷- ضریب همبستگی بین TOC حاصل از سیستم کلونی مورچه پیوسته و TOC واقعی در داده‌های آموزشی



شکل ۸- ضریب همبستگی بین TOC حاصل از سیستم کلونی مورچه پیوسته و TOC واقعی در داده‌های صحت‌سنجی

بهبود برآوردها شود [۳۰]. نتایج حاصل از تفسیر TOC اندازه‌گیری شده توسط پیرولیز Rock-Eval برای سنگ‌های متعلق به سازندهای کژدمی، گورپی و پابده در محدوده میدان نفتی اهواز بیانگر این است که سازند پابده نسبت به سایر سازندها پتانسیل بیشتری برای زایش نفت از خود نشان می‌دهد، ولی به علت پائین بودن درجه بلوغ وارد مرحله زایش هیدروکربن نشده است. سازند کژدمی در منطقه با توجه به درجه بلوغ و میزان مواد آلی از نقش بیشتری در زایش نفت برخوردار است و تصور می‌شود این سازند به عنوان یک سنگ منشأ مهمترین نقش را در شارژ مخازن این میدان (آسماری و بنگستان) به عهده داشته باشد. با توجه به تطبیق بالای TOC برآورد شده توسط الگوریتم CACS بعد از خوشه‌بندی سلسله مراتبی در این پژوهش، که به صورت گرافیکی و نموداری در شکل ۹ در مقایسه با مقادیر TOC واقعی نرمالایز شده در داده‌های تست به تصویر کشیده شده است، چنین تفاسیری از نگار TOC برآورد شده توسط این مدل برای چاه‌هایی که فاقد داده‌های TOC کافی هستند هم می‌تواند صورت بگیرد.

طبق نگارهای پتروفیزیکی و تفاسیر سرچاهی زمین‌شناسان، نمونه‌های آنالیز شده در این مطالعه دارای بازه تغییرات سنگ‌شناسی کربناته-مارنی تا شیلی می‌باشند. همان‌طور که اشاره شد، خوشه‌بندی می‌تواند ضخامت‌هایی از سازندها با خصوصیات مشابه را در یک خوشه قرار دهد. بنابراین، در این پژوهش بازه‌های دارای سنگ‌شناسی یکنواخت به‌طور جداگانه مورد بررسی قرار گرفتند و نتایج آنها بهتر از قبل از تأثیر خوشه‌بندی بوده است. مقدار میانگین مربع خطا و ضریب همبستگی حاصل از عملکرد الگوریتم CACS قبل از خوشه‌بندی برای داده‌های تست ۰/۰۰۲۵ و ۰/۹۴ بوده است [۲۹]. اگرچه الگوریتم CACS در بازه نسبتاً وسیعی از تغییرات سنگ‌شناسی نیز موفق عمل نموده است اما در برآورد صحیح مقادیر TOC جهت تحلیل و ارزیابی پتانسیل هیدروکربن‌زایی کاهش میانگین مربع خطا حتی در حد کم از اهمیت بسزایی برخوردار است. بنابراین تلفیق این روش با خوشه‌بندی می‌تواند به عنوان مدلی بهینه‌تر شناخته شود. لازم به ذکر است که افزایش تعداد خوشه در حد معقول و در صورتی که با کمبود داده مواجه نباشیم نیز می‌تواند سبب



شکل ۹- تطابق گرافیکی و نموداری TOC واقعی نرمال و TOC برآورد شده توسط الگوریتم CACS بعد از خوشه‌بندی آزمایشگاهی و TOC پیش‌بینی شده

ساده دارد و تنها دارای یک پارامتر کنترلی اصلی، یعنی تعداد مورچه است. این الگوریتم دارای دقت و سرعت بالاتری نسبت به سایر الگوریتم‌ها بر پایه جمعیت همچون الگوریتم ژنتیک می‌باشد و در مورد مسائل بسیار پیچیده مفید واقع می‌شود، چراکه نیازی به دانستن هیچ قاعده‌ای در ارتباط با مسئله مورد حل ندارد و تنها با قواعد داخلی خودش عمل می‌کند. این مدل کم هزینه است و در زمانی کوتاه قادر به تخمین نگار TOC از داده‌های پتروفیزیکی می‌باشد. با توجه به دقت بالای الگوریتم CACS در برآورد نگار TOC، تفاسیر ژئوشیمیایی با ارزشی می‌تواند با تعمیم این مدل برای چاه‌هایی که فاقد داده‌های TOC کافی هستند صورت بگیرد. همچنین این پژوهش افزایش تعداد خوشه در حد معقول را پیشنهاد می‌دهد چرا که در صورت عدم کمبود داده این نکته سبب بهبود برآوردها می‌شود. به امید آنکه این پژوهش گامی کوچک در پیشرفت صنعت ملی نفت و بهره‌وری از ذخایر و پتانسیل‌های هیدروکربنی باشد.

نتیجه‌گیری

در این مطالعه، از الگوریتم مبتنی بر آنالیز خوشه‌بندی سلسله مراتبی جهت تخمین داده‌های TOC از داده‌های پتروفیزیکی در میدان نفتی اهواز استفاده شد. داده‌های پتروفیزیکی (مقدار مقاومت، زمان عبور صوت، نوترون و همچنین گاما) به عنوان ورودی و داده‌های TOC حاصل از آنالیز ژئوشیمیایی به عنوان خروجی مطلوب شبکه تعیین گردید. خوشه‌بندی سلسله مراتبی با توجه به کمترین فاصله اقلیدسی نمونه‌ها با هم ۴ خوشه را ایجاد نمود. سپس الگوریتم CACS ضرایب وزنی تابع پیشنهاد شده را برای هر خوشه به دست آورد. با اعمال این ضرایب، نتیجه تلفیق صحت‌سنجی شبیه‌سازی خوشه‌ها مقادیر مقدار میانگین مربع خطا و ضریب همبستگی را برای داده‌های تست به ترتیب ۰/۰۲۰۸ و ۰/۹۵۸۲ ارائه نمود. نتایج بیانگر این است که با اعمال خوشه‌بندی ضریب اطمینان حاصل از تخمین TOC افزایش می‌یابد. الگوریتم CACS که در این پژوهش نقش اصلی را دارا بود، ساختاری

مراجع

- [۱]. علیزاده ب.، مرادی م.، قلاوند ه.، مطالعه ژئوشیمیایی سازند پابده در میدان نفتی اهواز، نهمین همایش انجمن زمین‌شناسی ایران، دانشگاه تربیت معلم تهران، ایران، ۱۳۸۴.
- [2]. Beers R. F., "Radioactivity and organic content of some Paleozoic shales", AAPG Bulletin, Vol. 26, pp. 1 – 22, 1945.
- [3]. Swanson V. E., "Oil yield and uranium content of black shales", United States Geological Survey Professional Paper, Vol. 356-A, pp.1-44, 1960.
- [4]. Schmoker J. W., "Determination of organic-matter content of Appalachian Devonian shales from gamma-ray logs", AAPG Bulletin, Vol. 65, pp. 2165–2174, 1981.
- [5]. Fertle H., and George V., "Total organic carbon content determined from well logs", SPE, Vol. 3, No. 2, pp. 407-419, 1988.
- [6]. Hertzog R., Colson L., Seeman B., O'Brian M., Scott H., Mckeen D., Wraight P., Grau J., Schweitzer J. and Herron M., "Geochemical logging with spectrometry tools", SPE Formation Evaluation, Vol. 4, pp. 153–162, 1989.
- [7]. Dellenbach J., Espitalie J. and Lebreton F., *Source rock logging*, Trans. 8th European SPWLA Symposium, paper D, 1983.
- [8]. Hussain F. A., "Source rock identification in the state of Kuwait using wireline logs", SPE 15747, pp. 477– 488, 1987.
- [9]. Passey Q. R., Creaney S., Kulla B., Moretti F. J., and Stroud J. D., "A practical model for organic richness from porosity and resistivity logs", AAPG, Vol. 74, No. 12, pp. 1777-1794, 1990.
- [10]. Kamali M. R., and Mirshady A. A., "Total organic carbon content determined from well logs using $\Delta\log R$ and neuro fuzzy techniques", J. Petrol. Sci. Eng., Vol. 45, pp. 141–148, 2004.
- [11]. Kadkhodaie-Ilkhchi A., Rahimpour-Bonab H. and Rezaee M. R., "A committee machine with intelligent systems for estimation of total organic carbon content from petrophysical data", Comput. Geosci., Vol. 35, pp. 457-474, 2008.
- [12]. Kadkhodaie-Ilkhchi A., Rezaee M. R. A., and Rahimpour-Bonab H., "A committee neural network for prediction of normalized oil content from well log data: An example from South Pars Gas Field", Persian Gulf. J. Petrol. Sci. Eng., Vol. 65, pp. 23-32, 2009.
- [13]. Alizadeh B., Najjari S., and Kadkhodaie-Ilkhchi A., "Artificial neural network modeling and cluster analysis for organic facies and burial history estimation using well log data: A case study of the South Pars Gas Field", Persian Gulf, Iran. Comput. Geosci., Vol. 45, pp. 261–269, 2011.
- [14]. Sefidari E., Kadkhodaie-Ilkhchi A., and Najjari S., "Comparison of intelligent and statistical clustering approaches to predicting total organic carbon using intelligent systems", J. Petrol. Sci. Eng., Vol. 86-87, pp. 190-205, 2012.
- [۱۵]. جباریه ع.، مرجع کامل MATLAB، انتشارات ارس رایانه، تهران، ۱۳۷۷.
- [16]. Jafarzadeh M., and Hosseini-Barzi M., "Petrography and geochemistry of Ahwaz Sandstone Member of Asmari Formation, Zagros, Iran: implications on provenance and tectonic setting", Rev. Mex. Cienc. Geol., Vol. 25,

No. 2, pp. 247-260, 2008.

[۱۷]. گودرزوند چگینی ف.، فرخنده کواکی م.، گوگل م.، محمدی فرد م.، محاسبات تحلیل پینچ برای انتگراسیون حرارتی با استفاده از نرم افزار MATLAB، پژوهش نفت، دوره ۱۴، شماره ۴۹، صفحات ۴۷-۵۳، ۱۳۸۳.

[18]. MATLAB Users Guide, Version 7.12, Statistics Toolbox, The MathWorks Inc, 2011.

[19]. Blei D. M., *Hierarchical clustering*, Lecture slides. published by: Princeton University Available at: www.cs.princeton.edu/courses/archive/spr08/cos424_clustering-2.pdf, 2008.

[20]. Dorigo M., "*Optimization, Learning and Natural Algorithms*", PhD thesis, Dipartimento di Elettronica, Politecnico di Milano, Milan, pp. 140, 1992.

[21]. Pourtakdoust S. H., and Nobahari H., "*An Extension of Ant Colony System to Continuous Optimization Problems*", ACO-SI, Vol. 3172, pp. 294-301, 2004.

[22]. Dorigo M., Birattari M., and Stutzle T., *Ant Colony Optimization: Artificial Ants as a Computational Intelligence Technique*, Available at: www.cimat.mx/~jortega/MaterialDidactico/modestol10/.../ACO1.pdf, 2006.

[23]. Kuhn L., "*Ant Colony Optimization for Continuous Spaces*", Bachelor Thesis, University of Queensland, pp. 52, 2002.

[24]. Box G. E. P., and Muller M. A., "*A note on the generation of random normal deviates*", Annals. Math. Stat., Vol. 29, pp. 610-611, 1958.

[25]. Marsaglia G., Tsang W. W., "*The ziggurat method for generating random variables*", J. Stat. Softw., Vol. 5, No. 8, pp. 1-7, 2000.

[26]. Karimi A., Nobahari H., and Siarry P., "*Continuous Ant Colony System and Tabu Search Algorithms Hybridized for Global Minimization of Continuous Multi-minima Functions*", Comput. Optim. Appl., Vol. 45, pp. 639-661, 2008.

[۲۷] حسینی ز.، کدخدائی ع.، طباطبائی س.م.ا.، برآورد نگار ژئوشیمیایی با استفاده از الگوریتم ژنتیک مبتنی بر آنالیز خوشه بندی، دومین همایش مهندسی مخازن هیدروکربوری و صنایع بالادستی، مرکز همایش های صدا و سیما، ایران، ۱۳۹۲.

[28]. Larose D. T., "*Discovering knowledge in data: an introduction to data mining*", Jhon Wiley & Sons Inc, pp. 240, 2005.

[۲۹] حسینی ز.، کدخدائی ع.، طباطبائی س. م. ا.، استفاده از بهینه سازی کلونی مورچه جهت تخمین نگار ژئوشیمیایی از نگارهای پتروفیزیکی، مطالعه موردی در میدان نفتی اهواز، اولین کنفرانس و نمایشگاه تخصصی نفت، پژوهشگاه صنعت نفت، ایران، ۱۳۹۲.

[۳۰] حسینی ز.، سنتز لاگ ژئوشیمیایی از داده های پتروفیزیکی با استفاده از بهینه سازی کلونی مورچه در میدان نفتی اهواز، پایان نامه کارشناسی ارشد، دانشگاه تبریز، ایران، ۱۳۹۲.