

ارزیابی کیفیت توده سنگ زیر سطحی با تلفیق نتایج وارون سازی داده های توموگرافی لرزه ای انکساری و مقاومت ویژه الکتریکی به روش خوشه بندی طیفی

بهمن محمدی^۱، ابوالقاسم کامکار روحانی^{۲*}

^۱ دانشجوی دکتری رشته مهندسی معدن - اکتشاف، دانشگاه صنعتی شاهرود، bahmanmohammadi2005@gmail.com

^۲ دانشیار، دانشگاه صنعتی شاهرود، دانشکده مهندسی معدن، نفت و ژئوفیزیک، kamkarr@yahoo.com

(دریافت ۱۶ آبان ۱۳۹۵، پذیرش ۱۹ اسفند ۱۳۹۶)

چکیده

هنگامی که ساختار داده ها غیرخطی باشد روش های کلاسیک خوشه بندی با شکست روبه رو می شوند. در این حالت، خوشه بندی طیفی روشی قدرتمند برای دسته بندی داده ها محسوب می شود. این تکنیک با تبدیل فضای ورودی، فضای جدیدی با قابلیت توصیف مناسب تر از داده ها را در اختیار قرار می دهد. در این مطالعه، ابتدا وارون سازی داده های مقاومت ویژه الکتریکی با روش کمترین مربعات گوس - نیوتن و با استفاده از نرم افزار RES2DINV انجام شد و سپس اولین زمان رسیده ها با استفاده از نرم افزار PickWin محاسبه و وارون سازی داده های توموگرافی لرزه - ای انکساری با استفاده از نرم افزار Geotom CG انجام شد. بعد از محاسبه ماتریس شباهت، مقادیر ویژه و بردارهای ویژه به دست آمد و سپس داده ها با روش خوشه بندی طیفی (Spectral clustering) مورد ارزیابی قرار گرفتند. با استفاده شاخص دان (Dunn index) برای بهینه سازی تعداد خوشه ها عدد ۱۲ به دست آمد که با توجه به نقشه های به دست آمده برای مقاومت ویژه الکتریکی و توموگرافی لرزه ای انکساری، تعداد خوشه مناسبی است. نتایج روش خوشه بندی طیفی همبستگی خوبی با نتایج لاگ گمانه حفاری شده در کف سد نشان می دهد. محاسبات مربوط به الگوریتم خوشه بندی طیفی و شاخص دان با کدنویسی در محیط نرم افزار MATLAB انجام شده است.

کلمات کلیدی

وارون سازی، روش خوشه بندی طیفی، توموگرافی لرزه ای انکساری، مقاومت ویژه الکتریکی، شاخص دان

۱- مقدمه

خوشه‌بندی یکی از مهم‌ترین روش‌های یادگیری بدون نظارت در زمینه شناسایی الگو و یادگیری ماشین است. هدف این تکنیک، یافتن ساختار در مجموعه‌ای از داده‌های بدون برچسب^۱ و دسته‌بندی آن‌ها است به صورتی که بعد از خوشه‌بندی، داده‌هایی که به یک خوشه تعلق گرفته‌اند کاملاً مشابه با یکدیگر و متفاوت با داده‌های سایر خوشه‌ها باشند. وقتی که ساختار داده‌ها در فضای ورودی متعلق به نواحی محدب نباشد یا به بیان دیگر داده‌ها ساختار غیر خطی داشته باشند، نتیجه الگوریتم‌های کلاسیک در خوشه‌بندی داده‌ها ضعیف خواهد بود [1]. در چنین مواقعی الگوریتم‌های یادگیری زیرفضا به عنوان راه حل برگزیده مطرح‌اند. هدف این دسته از الگوریتم‌ها، بازنمایی داده‌ها در فضایی جدید و با ابعاد کمتر است. مزیت عمده این فضای جدید، بیانی مناسب‌تر برای توصیف داده‌ها است. از روش‌های موجود در این مورد می‌توان، یادگیری مانیفلد [2,3]، تحلیل طیفی [4,5] و روش‌های مبتنی بر کرنل [6,7] را نام برد. خوشه‌بندی طیفی یکی از قدرتمندترین و محبوب‌ترین روش‌ها است که اولین بار توسط دونات و هافمن [8] در سال ۱۹۷۳ ارائه شده است. این الگوریتم سه مرحله اصلی دارد: محاسبه ماتریس شباهت، محاسبه بردارهای ویژه این ماتریس و کلاسه‌بندی نتایج [9]. در مقایسه با الگوریتم‌های سنتی، خوشه‌بندی طیفی برتری‌های بنیادی دارد. این خوشه‌بندی، روش اجرایی ساده دارد و به طور موثر با روش‌های جبری خطی استاندارد حل می‌شود [10,11]. روش‌های خوشه‌بندی طیفی از مفهوم گراف طیفی برآمده‌اند و از ماتریس شباهت برای کاهش بعد داده‌ها در خوشه‌بندی استفاده می‌نمایند. برای داده‌های بزرگ این ماتریس به صورت پراکنده محاسبه می‌شود [12,13]. ایده اصلی در این دسته از روش‌ها ساخت یک گراف وزن‌دار بدون جهت بر اساس داده‌هاست به صورتی که رئوس گراف نشان دهنده مجموعه نقاط و هر یال وزن‌دار نشان دهنده میزان شباهت جفت داده متناظر آن است. الگوریتم‌های SLH (Scott, Longuet) [14] Higgins [14]، PF (Perona, Freeman) [15]، برش نرمال [16] و NJW (Ng, Jordan, Weiss) [1] از جمله این روش‌ها است. خوشه‌بندی طیفی، مانند دیگر روش‌های خوشه‌بندی می‌تواند در هر زمینه علمی که با داده‌های تجربی سر و کار دارد و زیرمجموعه‌های داده‌های مشابه را شناسایی

می‌کند، به کار برده شود. تحلیل خوشه‌ای طیفی در دهه اخیر محبوبیت زیادی پیدا کرده است و امروزه به طور گسترده‌ای در مسایل مختلف داده‌کاوی با کاربردهای متفاوت استفاده می‌شود. برای مثال، تقسیم‌بندی تصویر به قطعات مختلف، استخراج ماتریس‌های توصیف‌کننده ژن، تشخیص دست خط و صوت یا صدا از این نمونه است [16,17,1,18,10,19].

۲- مواد و روش‌ها

۲-۱- وارون‌سازی داده‌های توموگرافی لرزه‌ای

مدل اولیه مربوط به ساختار مورد بررسی، با استفاده از شواهد زمین‌شناسی و داده‌های ژئوفیزیک به دست می‌آید و زمان سیر محاسبه شده و مشاهده شده با یکدیگر مقایسه می‌شوند. هدف از وارون‌سازی به حداقل رساندن اختلاف زمان سیر حاصل از مدل اولیه و زمان سیر مربوط به داده‌های واقعی (برداشت شده) طی یک فرایند تکراری است [20,21]. زمان سیر از رابطه ۱ به دست می‌آید:

$$t_i = \int_{l_i(z)} s \cdot dl_i(z) \quad (1)$$

که در آن:

t_i زمان سیر مربوط به پرتو i ام
 s مقدار کندی یا عکس سرعت
 l طول پرتو i ام (با عمق z تغییر می‌کند).

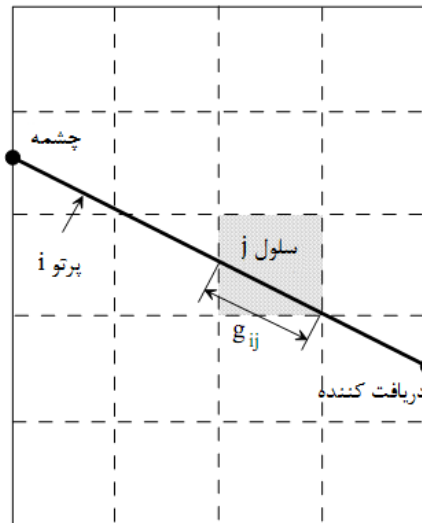
۲-۱-۱- توموگرافی اولین زمان رسید

توموگرافی اولین زمان رسید با استفاده از اولین زمان رسیده‌های ثبت شده انجام می‌شود. در این روش مدل اولیه به تعدادی سلول تقسیم می‌شود و در نهایت حل مساله وارون با استفاده از یک الگوریتم مانند الگوریتم بازسازی تکرار همزمان یا SIRT^۲ و با رابطه ۲ انجام می‌شود.

$$dt_i = M \cdot ds \quad (2)$$

روش توموگرافی لرزه‌ای قادر به تصویرسازی از ساختمان داخلی زمین بر اساس سرعت سیر امواج است که می‌تواند به صورت گالری به گالری، گمانه به گمانه و سطح زمین به گمانه باشد. شکل ۱ نمایش شماتیک از یک مقطع توموگرافی شامل چشمه

و گیرنده امواج لرزه‌ای، سلول‌بندی مقطع و پرتو لرزه‌ای است.



شکل ۱- نمایش شماتیک از یک مقطع توموگرافی شامل چشمه و گیرنده امواج لرزه‌ای، سلول‌بندی مقطع و پرتو لرزه‌ای [22]. محور افقی و عمودی، بسته به نوع برداشت می‌تواند XY ، XZ و یا YZ بر حسب متر انتخاب شود.

می‌شود. مراحل معمول در فرآیند وارون‌سازی شامل قرائت اولیه داده‌های مقاومت ویژه، تصحیح پارامترهای وارون‌سازی در صورت لزوم و انجام وارون‌سازی است. در مدلسازی قبل از هر چیز به یک مدل فرضی اولیه برای انجام مدلسازی نیاز است [23]. روش مدلسازی وارون معمولاً از یک تکنیک بهینه‌سازی غیرخطی برای بهبود بخشیدن به یک مدل ساده اولیه استفاده می‌کند [24].

در این تحقیق، برای مدلسازی دو بعدی (2D) داده‌های مقاومت ویژه، از نرم‌افزار RES2DINV استفاده شده است. در این نرم‌افزار، مدل مورد استفاده به عنوان مدل اولیه شامل یک سری از بلوک‌های مستطیلی دو بعدی همان‌طور که در شکل ۲ نشان داده شده، در نظر گرفته می‌شود. در این مدل تعداد بلوک‌ها با تعداد نقاط داده‌ها در شبه مقطع مقاومت ویژه ظاهری برابر است و به یک شیوه مناسب مرتب شده‌اند. همچنین عمق مراکز بلوک‌ها در عمق میانگین تحت بررسی برای بازه‌های الکترونی مختلف استفاده شده قرار گرفته‌اند [23].

در این آزمایش زمان سیر موج و مختصات مکانی چشمه‌ها و گیرنده‌های امواج به عنوان داده‌های اولیه در نظر گرفته می‌شود و مقطع تصویر شده نحوه توزیع سرعت را مشخص می‌کند. روش محاسباتی مبتنی بر حل معادله ماتریسی (رابطه ۳) است [22]:

$$t_j = \sum a_{ij} \cdot x_i \quad \begin{matrix} i = 1, 2, \dots, n \\ j = 1, 2, \dots, m \end{matrix} \quad (3)$$

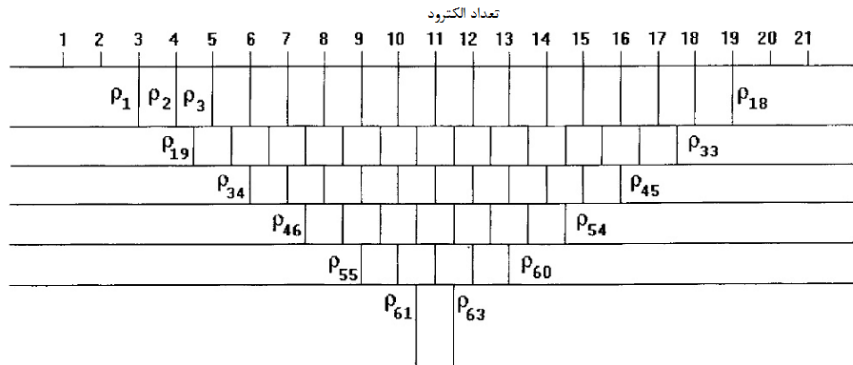
که در آن:

a_{ij} طول i امین پرتو در i امین سلول
 x_i عکس سرعت موج ($1/V$) در i امین سلول
 t_j زمان سیر j امین پرتو

برای حل این معادله ماتریسی چندین روش وجود دارد که بهترین روش استفاده از الگوریتم تکرار است.

۲-۲- مدلسازی وارون

در روش وارون‌سازی بر خلاف مدلسازی مستقیم یا پیشرو با استفاده از داده‌های به دست آمده، مدل زمین تخمین زده



شکل ۲- ترتیب بلوک‌های مستطیلی مورد استفاده در مدل‌سازی دوبعدی [24].

و d_{\max} بزرگ‌ترین فاصله D_k یا به عبارتی ماکزیمم فاصله (معمولا اقلیدوسی) بین نقاط در خوشه‌های مختلف است (رابطه ۷):

$$d_{\max} = \max_{1 \leq k \leq K} D_k \quad (7)$$

شاخص دان با رابطه ۸ تعریف می‌شود.

$$c = \frac{d_{\min}}{d_{\max}} \quad (8)$$

۲-۴- الگوریتم خوشه‌بندی طیفی

الگوریتم‌های خوشه‌بندی طیفی ارتباط نزدیکی با تئوری گراف طیفی [5] دارند. این روش‌ها مبتنی بر افراز داده‌ها بر اساس تجزیه ویژه ماتریس شباهت و به دست آوردن بردارها و مقادیر ویژه آن است. در مجموعه‌ای با N داده d بعدی که $X = \{x_1, x_2, \dots, x_N\}$ است، می‌توان برای این مجموعه گراف وزن دار و بدون جهت $G(V, A)$ را ساخت به صورتی که رئوس گراف $V = \{v_1, v_2, \dots, v_N\}$ نشان دهنده N داده و یال‌ها که ماتریس شباهت $N \times N$ را تشکیل می‌دهند بیانگر میزان شباهت بین هر جفت داده متناظر است. ماتریس شباهت با رابطه ۹ تعریف می‌شود:

$$A_{ij} = \begin{cases} h(x_i, x_j) & i \neq j \\ 0 & \text{otherwise} \end{cases} \quad (9)$$

تابع h میزان شباهت بین دو داده را اندازه می‌گیرد و می‌تواند یک تابع گوسی به شکل رابطه ۱۰ باشد.

پس از تخمین مدل اولیه نوبت به تصحیح این مدل با هدف رسیدن به انطباق قابل قبول با داده‌های صحرایی می‌رسد. برای این منظور دو روش مورد استفاده قرار می‌گیرد که عبارتند از روش معمول کمترین مربعات گوس- نیوتن و روش شبه نیوتن

۲-۳- روش شاخص دان

روش شاخص دان یکی از روش‌های دقیق کمی برای محاسبه تعداد خوشه‌های بهینه در الگوریتم‌های مختلف خوشه‌بندی است. فاصله بین خوشه‌های C_k و $C_{k'}$ با محاسبه فاصله بین نزدیک‌ترین نقاط آن‌ها اندازه‌گیری می‌شوند (رابطه ۴) [25].

$$d_{kk'} = \min_{\substack{i \in I_k \\ j \in I_{k'}}} \|M_i^{(k)} - M_j^{(k')}\| \quad (4)$$

و d_{\min} کوچک‌ترین فاصله $d_{kk'}$ یا به عبارتی مینیمم فاصله (معمولا اقلیدوسی) بین نقاط در خوشه‌های مختلف است (رابطه ۵):

$$d_{\min} = \min_{k \neq k'} d_{kk'} \quad (5)$$

برای هر خوشه C_k ، رابطه D_k به عنوان بزرگ‌ترین فاصله نقاط مشخص در خوشه (قطر خوشه هم نامیده می‌شود)، تعریف می‌شود (رابطه ۶):

$$D_k = \max_{\substack{i, j \in I_k \\ i \neq j}} \|M_i^{(k)} - M_j^{(k)}\| \quad (6)$$

۵- خوشه‌بندی مجموعه داده بازنمایی شده Y با استفاده از K -Means

ابزار اصلی برای خوشه‌بندی طیفی، ماتریس لاپلاسیین گراف است. در تئوری گراف طیفی، لاپلاسیین‌های متفاوت گراف وجود دارد که در خوشه‌بندی طیفی استفاده می‌کنند.

برای به کار بردن خوشه‌بندی طیفی، بر اساس n مدل ورودی، یک گراف غیر جهت‌دار تولید می‌شود [28] و با ماتریس شباهت که ناهمگنی ساختاری همه مدل‌های ورودی را نشان می‌دهد، توصیف می‌شود. به طور پی در پی یک ماتریس لاپلاسیین گراف از گراف غیر جهت‌دار تعیین می‌شود، سپس مقادیر ویژه (طیف) و متناظر آن بردارهای ویژه از ماتریس لاپلاسیین گراف محاسبه می‌شود. بزرگ‌ترین مقادیر ویژه طیف تعیین و نظیر به نظیر آن، بردارهای ویژه به عنوان ورودی برای تحلیل خوشه‌ای انتخاب می‌شوند. سرانجام، اطلاعات عضویت خوشه در محدوده مدل اصلی دوباره به تصویر درآورده می‌شود تا یک مدل چند پارامتری منطقه‌بندی شده به دست آید [27].

تشابه‌های جفت جفت بین نمونه‌ها در یک فضای n بعدی، معمولاً با گراف‌های متشابه [28] با استفاده از گره‌ها (رئوس) و لبه‌های (اضلاع) مرتبط با دو گره، مدل می‌شود. تعداد گره‌ها در یک گراف برابر تعداد نمونه‌های موجود یا سلول مدل‌ها در فضای پارامتری یا مدل‌های ورودی است. تعداد لبه‌ها ممکن است تغییر کند، اما حداکثر مقدار ممکن از لبه‌ها موقعی است که هر گره به همه گره‌های دیگر متصل باشد. در صورت نبود وابستگی جهتی در تشابه‌های جفت جفت، گراف غیر جهت‌دار می‌شود. تشابه‌های جفت جفت (رابطه ۱۰) در گراف‌های غیر جهت‌دار، عناصری از ماتریس شباهت A است که در واقع وزن‌دار شده ماتریس مجاورتی گراف‌اند. این یک ماتریس مربعی متقارن است که بعد آن برابر تعداد گره‌های گراف است [27].

۳- تحلیل نتایج

۳-۱- وارون‌سازی داده‌های ورودی

داده‌های توموگرافی لرزه‌ای و مقاومت ویژه الکتریکی از ساختگاه یکی از پروژه‌های سد در حال اجرا، در ایران برداشت شده است. با توجه به نتایج حاصل از وارون‌سازی داده‌های مقاومت ویژه الکتریکی، لایه‌های زیر سطحی مطابق شکل ۳ به دست آمد. در شکل ۳ همچنین نتایج حاصل از حفاری نیز آورده شده است. وارون‌سازی با روش کمترین مربعات گوس-

$$h(x_i, x_j) = \exp(-d^2(x_i, x_j) / \sigma^2) \quad (10)$$

که d فاصله بین دو نمونه را نشان می‌دهد. فاصله اقلیدوسی بین دو نمونه با پارامتر σ وزن‌دار می‌شود که این، کاهش تشابه را با افزایش فاصله نمونه (سرعت کاهش تابع h با افزایش فاصله بین دو نمونه) در فضای پارامتری، کنترل می‌کند. پارامتر σ به صورت تجربی تعیین می‌شود [26]. اگر بین ۱ تا ۱۰۰ انتخاب شود، یک اثر ضعیف روی نتایج خوشه‌بندی داده‌های نرمال شده دارد. زمانی که داده‌ها غیر نرمال هستند، انتخاب σ خیلی مشکل است [27]. در ادامه الگوریتم استفاده شده در این مطالعه آورده شده است.

۲-۴-۱- روش NJW

ایده الگوریتم (NJW (Ng, Jordan, Weiss) [1] استفاده از اولین k بردار ویژه متناظر با بزرگ‌ترین k مقدار ویژه ماتریس لاپلاسیین است. مراحل این الگوریتم به صورت زیر است:

۱- ساخت ماتریس شباهت با استفاده از رابطه ۹

۲- محاسبه ماتریس درجه D (رابطه ۱۱):

$$D_{ii} = \sum_{j=1}^n A_{ij} \quad (11)$$

و ماتریس لاپلاسیین (رابطه ۱۲):

$$L_N = D^{-\frac{1}{2}} A D^{-\frac{1}{2}} \quad (12)$$

۳- به دست آوردن اولین k بردار ویژه v^1, v^2, \dots, v^k متناظر با اولین k بزرگ‌ترین مقدار ویژه ماتریس L_N و تشکیل ماتریس ستونی (رابطه ۱۳):

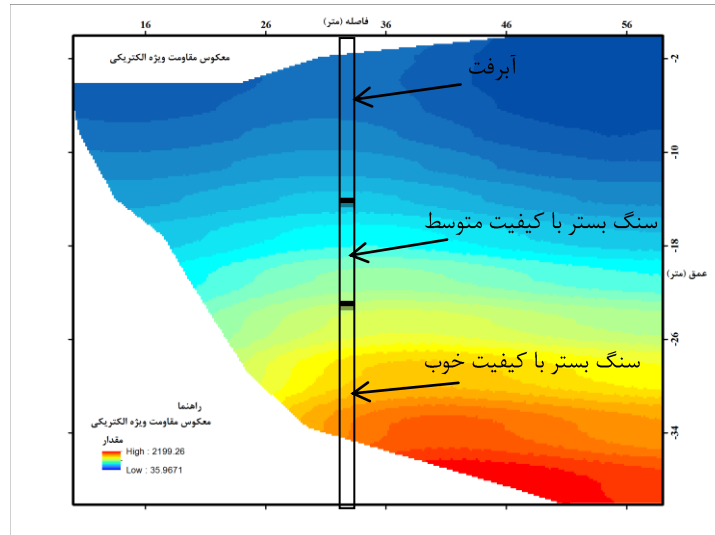
$$V = [v^1, v^2, \dots, v^k] \in R^{n \times k} \quad (13)$$

۴- نرمال‌سازی مجدد V و تشکیل Y به طوری که همه سطرهای آن طول واحد داشته باشند (رابطه ۱۴):

$$Y_{ij} = V_{ij} / \sqrt{\sum_j V_{ij}^2} \quad (14)$$

متر در واقعیت است. لایه‌های به‌دست آمده نشان می‌دهد که مقاومت‌ویژه به طرف عمق به تدریج افزایش می‌یابد.

نیوتن و با استفاده از نرم‌افزار RES2DINV به‌دست آمده است. برای یکی کردن دو نقشه مقاومت ویژه و توموگرافی لرزه-ای انکساری، داده‌ها از عمق ۱۴ متر به پایین و زیر سطح ایستایی در نظر گرفته شدند. عمق صفر در نقشه برابر عمق ۱۴



شکل ۳- نتایج وارون‌سازی مقاومت‌ویژه الکتریکی برای اعماق پایین‌تر از ۱۴ متری سطح زمین

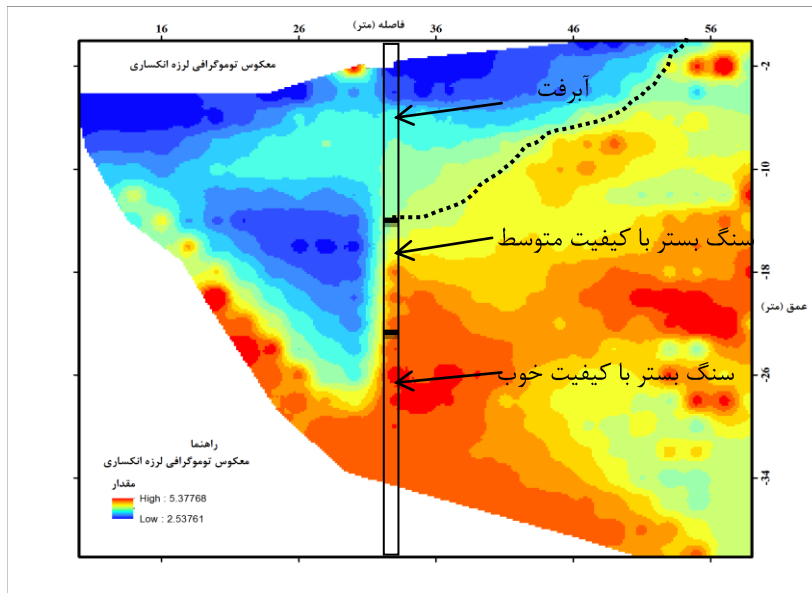
حفاری محاسبه می‌شود، درصد مغزه بازیافتی^۱ است. به این منظور، مجموع طول مغزه‌های حاصل از حفاری را به طول معین حفاری تقسیم می‌کنند و درصد آن را به‌دست می‌آورند. شاخص Fracture per m، نیز به تعداد درزه و شکاف در یک متر اشاره دارد.

گمانه KD3 در کف سد (در فاصله ۳۳ متری مقطع به دست آمده شکل ۴) حفاری شده و اطلاعات مغزه‌های آن مورد بررسی قرار گرفت. نتایج نشان می‌دهد که روش‌های ژئوفیزیکی مرز لایه آبرفت و سنگ بستر را به خوبی تفکیک کرده‌اند و تغییر خصوصیات در سنگ‌ها موید این مطلب است. با تلفیق لاگ حفاری شده و نقشه شکل ۴، مناطق آبرفت، سنگ بستر با کیفیت متوسط و سنگ بستر با کیفیت خوب ارزیابی شده است (بر اساس اطلاعات لاگ حفاری، ضخامت آبرفت ۳۰ متر است و با توجه به اینکه مقطع، از عمق ۱۴ متری به پایین است، بنابراین، در عمق ۱۶ متری (شکل ۴) مرز آبرفت مشخص شده است).

روش توموگرافی لرزه‌ای انکساری توانایی قابل توجهی در تشخیص ویژگی‌های لرزه‌شناسی لایه‌های زمین دارد. ابتدا با استفاده از نرم‌افزار PickWin اولین زمان رسیده‌ها به دست آمد و بعد با استفاده از نرم‌افزار Geotom CG وارون‌سازی داده‌های توموگرافی لرزه‌ای انکساری انجام شد. با توجه به نتایج به‌دست آمده از این مطالعه، لایه آبرفتی و هوازده سرعتی معادل ۳ کیلومتر بر ثانیه و سنگ بستر سرعتی بالای ۴ کیلومتر بر ثانیه دارد. با کیفیت‌ترین قسمت‌های توده‌سنگ با سرعت‌های بالاتر از ۴۰۰۰ متر بر ثانیه‌اند. قسمت‌هایی با سرعت پایین، نواحی با کیفیت کم و یا آبرفتی‌اند (شکل ۴).

برای اعتبارسنجی مدل‌های به دست آمده، از اطلاعات یا نتایج لاگ برداشت شده از گمانه استفاده می‌شود (جدول‌های ۱ و ۲). مقدار RQD به ناپیوستگی موجود در توده سنگ وابسته است. به منظور محاسبه این شاخص، برای طول معینی از حفاری، طول مغزه‌های سنگ بزرگ‌تر از ۱۰ سانتی‌متر را اندازه‌گیری و مجموع آن‌ها را به طول حفاری معین شده تقسیم کرده و شاخص RQD را بر حسب درصد محاسبه می‌کنند و سپس توده سنگ رده‌بندی می‌شود. شاخص دیگری که هم زمان با محاسبه درصد RQD در محل گمانه در حال

1- Core recovery



شکل ۴- نتایج وارون سازی توموگرافی لرزه ای انکساری برای اعماق پایین تر از ۱۴ متری سطح زمین

جدول ۱- لاگ گمانه حفاری شده در کف سد (مشخصات رنگ ها در جدول ۲ آورده شده است)

گمانه حفاری	KD 3
ارتفاع	2156
عمق (متر)	
10	
15	
20	
25	
30	
35	سنگ
40	
45	
50	
55	

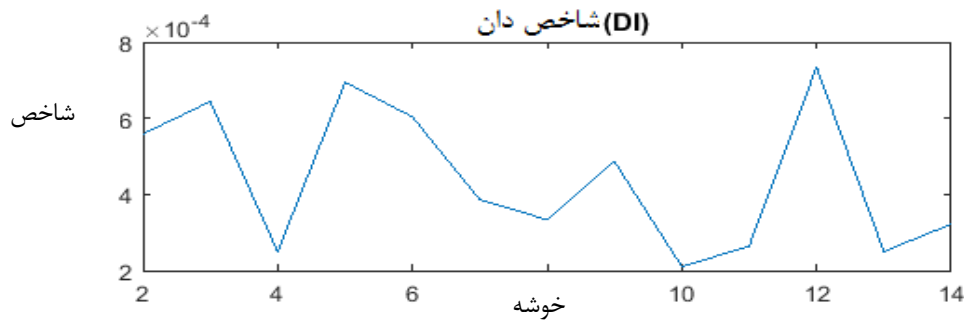
جدول ۲- راهنمای لاگ گمانه حفاری شده

شکستگی ها در هر متر	شاخص کیفیت سنگ (%)	بازیابی مغزه (%)	طبقه بندی مغزه
0-1	90-100	90-100	عالی
2-4	75-89	75-89	خوب
5-9	50-74	50-74	متوسط
> 10	25-49	25-49	ضعیف
Fracture zone	0-24	0-24	خیلی ضعیف

۳-۲- خوشه‌بندی طیفی داده‌های صحرایی

سنگ بستر است. با بکار بردن فرمول‌های الگوریتم خوشه‌بندی طیفی و شاخص دان (در بخش مواد و روش‌ها آمده است) در نرم‌افزار Matlab، کد محاسباتی برای آن‌ها نوشته شده و نقشه مربوطه ترسیم شد (شکل‌های ۵ و ۶).

با استفاده از روش شاخص دان برای بهینه‌سازی تعداد خوشه‌ها عدد ۱۲ به دست آمد (شکل ۵) که با توجه به نقشه‌های مقاومت ویژه الکتریکی و توموگرافی لرزه‌ای انکساری (شکل ۳ و ۴)، تعداد خوشه مناسبی از لحاظ تفکیک مناطق آبرفتی و

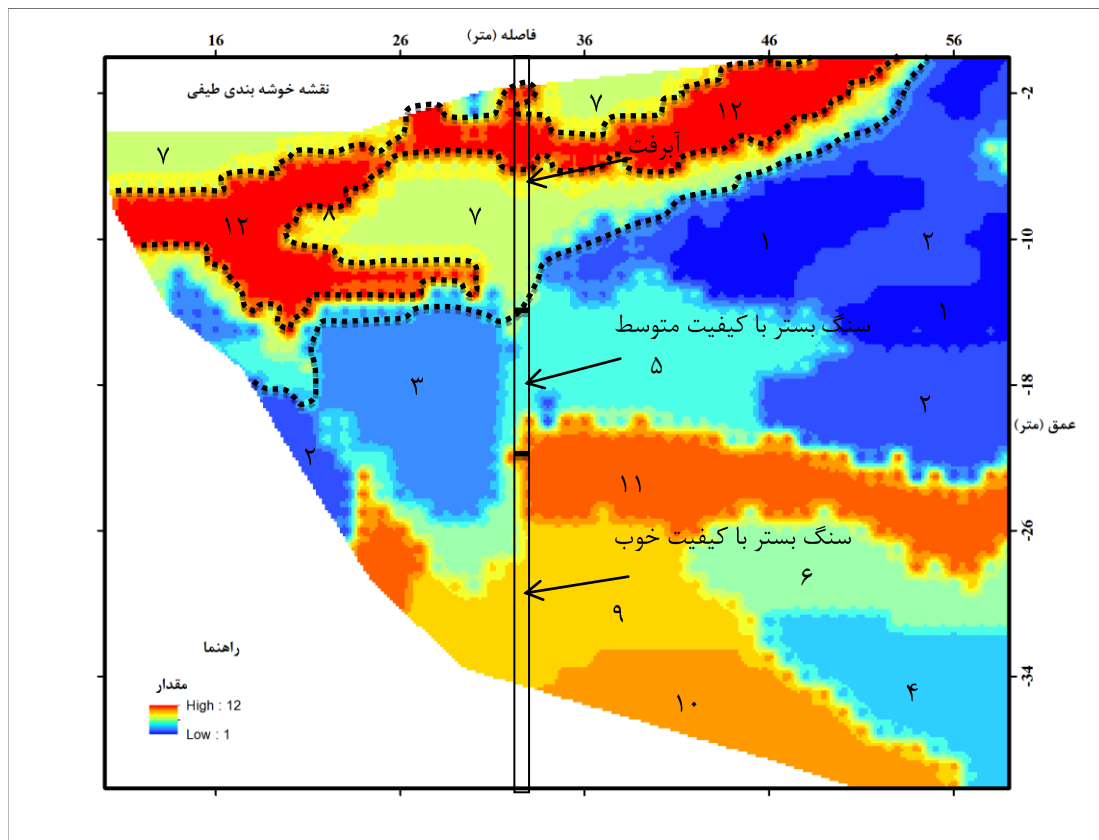


شکل ۵- شاخص دان که برای محاسبه تعداد خوشه‌های بهینه مورد استفاده قرار گرفته است. در این شاخص حداکثر مقدار به دست آمده برای خوشه‌ها، به عنوان تعداد خوشه بهینه در نظر گرفته می‌شود. محور افقی تعداد خوشه‌ها و محور عمودی مقدار شاخص را نشان می‌دهد.

مناطق به خوبی عمل کرده و نسبت به روش‌های سنتی دقیق‌تر است. تحلیل خوشه‌ای طیفی در فرآیند تلفیق مدل، ناهمگنی فضایی موجود در همه مدل‌های توموگرافی را در برمی‌گیرد. در خوشه‌های ۶ و ۴ که در شکل ۶ نشان داده شده است، سرعت موج P با توجه به نقشه‌های شکل (۳ و ۴) نسبت به لایه‌های بالاتر کاهش و مقاومت ویژه افزایش را نشان می‌دهد که این مناطق در خوشه‌بندی طیفی به خوبی تفکیک شده‌اند. از لحاظ زمین‌شناسی این قسمت سنگ آهک متخلخل با تخلخل موثر پایین است. با تلفیق نتایج وارون‌سازی داده‌های توموگرافی لرزه‌ای انکساری و مقاومت ویژه الکتریکی با استفاده از روش خوشه‌بندی طیفی کیفیت توده سنگ از ضعیف به خوب، مشخص و از ۱ تا ۱۲ در شکل ۶ ارزیابی شده است. تحلیل خوشه‌ای طیفی (شکل ۶) ناهمگنی موجود در دو مدل توموگرافی لرزه‌ای و مقاومت ویژه (شکل ۳ و ۴) را در بر گرفته است. مناطق آبرفتی (خوشه‌های ۷، ۸، ۱۲) به دلیل از هم گسستگی ذرات، سرعت لرزه‌ای پایین دارند و نیز به دلیل رسانش بالاتر در اثر وجود رس و تخلخل، مقاومت ویژه الکتریکی پایین است. در محدوده سنگ بستر به طور نسبی به طرف عمق، سرعت لرزه‌ای و مقاومت ویژه الکتریکی افزایش می‌یابد. با توجه به نتایج حفاری و تحلیل مغزه‌ها و تطابق آن با خوشه‌بندی طیفی (شکل ۶)، نتیجه گرفته شد که روش

مدل‌های شکل‌های ۳ و ۴ نرمال شده و تحت تحلیل خوشه‌ای طیفی قرار گرفتند. با محاسبه تشابه‌های جفت جفت گوسین با $\sigma = 2$ (پارامتر σ در فرمول ۱۰ معرفی شده است) ماتریس شباهت به دست آمد. محاسبه تشابه‌های جفت نیازمند همه مدل‌های مورد نظر است که باید به طور کامل در کنار هم قرار گیرند و به صورت یکسان از هم تفکیک شوند. بردارهای ویژه مطابق با بزرگ‌ترین مقادیر ویژه لاپلاسیان گراف مورد مطالعه قرار گرفتند (به دلیل بزرگ بودن ماتریس (۱۸۹۱) در (۱۸۹۱)، از آوردن مقادیر و بردارهای ویژه خودداری می‌شود). شکل ۶ مدل منطقه‌بندی شده نهایی را با یک جواب ۱۲ خوشه‌ای نشان می‌دهد. ساختارهای اصلی در مدل‌های لرزه‌ای و مقاومت ویژه (شکل‌های ۳ و ۴) به عنوان اطلاعات ترکیب شده با خوشه‌ها، نشان داده شده‌اند. هر زون می‌تواند با مقادیر سرعت موج P و مقاومت ویژه الکتریکی توصیف شود. در ناحیه آبرفتی سرعت موج P به دلیل از هم گسستگی ذرات و مقاومت ویژه به دلیل رسانش بالاتر، پایین است. با حرکت از سطح به عمق، سرعت موج P و مقاومت ویژه به طور نسبی افزایش نشان می‌دهد. نقشه ارایه شده در شکل ۶ که سه منطقه آبرفتی با کیفیت پایین توده سنگ و ۹ منطقه سنگ بستر با کیفیت بالای توده سنگ را به صورت زون‌بندی نمایش می‌دهد و نیز با توجه به نتایج اعتبارسنجی، خوشه‌بندی طیفی در تفکیک

خوشه‌بندی طیفی، لایه‌های مختلف زمین را به خوبی از هم تفکیک کرده و در خوشه‌های ۱ تا ۱۲ جای داده است.



شکل ۶- مدل منطقه‌بندی به دست آمده با الگوریتم خوشه‌بندی طیفی از مدل‌های وارون نشان داده شده در شکل‌های ۳ و ۴ (زون‌ها (طبقه‌ها) با رنگ‌های آبی تا قرمز مشخص شده‌اند).

۴- نتیجه‌گیری

خوشه‌بندی طیفی یکی از قدرتمندترین و محبوب‌ترین الگوریتم‌های خوشه‌بندی است و در مقایسه با الگوریتم‌های خوشه‌بندی سنتی، برتری‌های بنیادی دارد. خوشه‌بندی طیفی روش اجرای ساده‌ای دارد و به طور موثر با روش‌های جبری خطی استاندارد حل می‌شود. با توجه به نتایج به دست آمده از ۱۲ کلاس مورد بررسی (که با استفاده از روش شاخص دان محاسبه شد)، سه کلاس در منطقه آبرفتی با کیفیت پایین توده سنگ و ۹ کلاس در منطقه سنگ بستر با کیفیت بالای توده سنگ واقع شده‌اند. با تطابق نتایج حفاری و تحلیل مغزه‌ها با خوشه‌بندی طیفی نشان داده شد که خوشه‌بندی طیفی در تفکیک مناطق عمقی به خوبی عمل کرده و نسبت به روش‌های خوشه‌بندی سنتی دقیق‌تر است. در فرآیند تلفیق مدل، تحلیل خوشه‌ای طیفی، ناهمگنی فضایی موجود در دو مدل توموگرافی لرزه‌ای و مقاومت ویژه را در بر می‌گیرد.

۵- مراجع

- [1] Ng, A.Y., Jordan, M., and Weiss, Y., 2002, *On spectral clustering: analysis and an algorithm*: Advances in Neural Information Processing Systems, 14.
- [2] Tenenbaum, J.B., deSilva, V., and Langford, J.C., 2000, *A global geometric framework for nonlinear dimensionality reduction*: Science, 290, p. 2319–2323.
- [3] Roweis, S.T and Saul, L.K., 2000, *Nonlinear dimensionality reduction by locally linear embedding*: Science, 290, p. 2323–2326.
- [4] Zhang, T., et al., 2008, *A unifying framework for spectral analysis based dimensionality reduction*: in IEEE World Congress on Computational Intelligence.
- [5] Chung, F.R.K., 1997, *Spectral Graph Theory*: CBMS Regional Conference Series in Mathematics, 92.

- [18] Zelnik-Manor, L., and Perona, P., 2005, *Self-tuning spectral clustering*: Saul, L. K., Weiss, Y., and Bottou, L., eds., *Advances in neural information processing systems 17*, MIT Press, 1601–1608.
- [19] Zhang, X., Jiao, L., Liu, F., Bo, L., and Gong, M., 2008, *Spectral clustering ensemble applied to SAR image segmentation*: *IEEE Transactions on Geoscience and Remote Sensing*, 46, 2126–2136.
- [20] Zhou, H., 1997, *Determination of velocities and interfaces by multi-scale tomography*: 67th Ann. Internat. Mtg, Society of Exploration Geophysicists, Expanded Abstracts, pp.1877-1880.
- [21] Zhou, H., 2003, *Multi-scale traveltime tomography*: *Geophysics*, 68, pp.1639-1649.
- [22] Lehmann, B., 2007, *Seismic travel time tomography for engineering and exploration applications*: EAGE Publications bv, pp.14.
- [23] Loke M.H., Barker, R.D., 1996a, *Rapid least-square inversion of apparent resistivity pseudosections using a quasi-Newton method*: *Geophysical Prospecting*, 44, 131-152
- [24] Loke, M.H., 2004, *Tutorial: 2-D and 3-D electrical imaging surveys*.
- [25] Desgraupes, B., 2013, *Clustering Indices*: University Paris Ouest, Lab Modal'X.
- [26] Kanaan, EL.B., Fadi, EL.F., Ashour, W., 2014, *Spectral Clustering Using Optimized Gaussian Kernel Function*: *International Journal of Artificial Intelligence and Applications for Smart Devices*, Vol.2 , No.1, pp. 41-56.
- [27] Hachmoler, B., Paasche, H., 2013, *Integration of surface-based tomographic models for zonation and multimodel guided extrapolation of sparsely known petrophysical parameters*: *GEOPHYSICS*, VOL.78, NO.4, PP43-53.
- [28] West, D., 2007, *Introduction to graph theory*: Prentice Hall.
- [6] Müller, K.R., et al., 2001, *An introduction to kernel-based learning algorithms*: *Neural Networks*, IEEE Transaction on, 12(2), p. 181–202.
- [7] Schölkopf, B. and A.J. Smola., 2002, *Learning with Kernels, Support Vector Machines, Regularization, Optimization, and Beyond*: Cambridge, MA, USA, MIT Press.
- [8] Donath, W.E and Hoffman, A.J., 1973, *Lower bounds for the partitioning of graphs*: *IBM J. Res. Develop*, Vol. 17, No. 5, pp. 420-425.
- [9] TREMBLAY, N., PUY, G., GRIBONVAL, R., VANDERGHEYNST, P., 2016, *Compressive Spectral Clustering*: *Proceedings of the 33 rd International Conference on Machine Learning*, New York, NY, USA, JMLR: W&CP volume 48.
- [10] von Luxburg, U., 2007, *A tutorial on spectral clustering: Statistics and Computing*, 17, 395–416.
- [11] Auffarth, B., 2007, *Spectral Graph Clustering*: Universitat de Barcelona, course report for Tècnicas Avanzadas de Aprenidzaje at Universitat Politècnica de Catalunya.
- [12] Hadjighasem, A., Karrasch, D., Teramoto, H., Haller, G., 2016, *Spectral-clustering approach to Lagrangian vortex detection*: *PHYSICAL REVIEW E* 93, 063107.
- [13] Wu, S., Feng, X., Zhou, W., 2014, *Spectral clustering of high-dimensional data exploiting sparse representation vectors*: *Neurocomputing* 135, 229–239.
- [14] Scott, G.L. and Longuet-Higgins, H.C., 1990, *Feature grouping by relocalisation of eigenvectors of the proximity matrix*: in *British Machine Vision Conference*, p. 103-108.
- [15] Perona, P. and Freeman, W.T., 1998, *A factorization approach to grouping*: in *ECCV*.
- [16] Shi, J., and Malik, J., 2000, *Normalized cuts and image segmentation*: *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22, 888–905.
- [17] Meila, M., and Shi, J., 2001, *Learning segmentation by random walks*: Leen, T.K., Dietterich, T. G., and Tresp, V., eds., *Advances in neural information processing systems 13*, MIT Press, 873–879.