

# پهنه‌بندی شاخص کیفی سنگ در ساختگاه سد سمیلان، براساس گسل‌ها و شبکه عصبی خودسازمانده

امین حسین‌مرشدی<sup>۱\*</sup> و حسین معاریان<sup>۱</sup>

<sup>۱</sup>دانشکده فنی، دانشگاه تهران، تهران، ایران.

تاریخ دریافت: ۱۳۸۹/۰۸/۲۳ تاریخ پذیرش: ۱۳۸۹/۱۲/۱۱

## چکیده

پهنه‌بندی، یکی از مباحث مهم، در علوم زمین به شمار می‌آید. در پهنه‌بندی، یک محدوده به چند بخش یا پهنه مجزا تقسیم و در نهایت از ترکیب نتایج پهنه‌های مختلف، یک مدل واحد حاصل می‌شود. در این مطالعه، از روش‌های خوشه‌بندی برای پهنه‌بندی سد سمیلان استفاده شده است. تعداد پهنه‌های خوشه‌ها براساس متغیر ژئوتکنیکی (لوژون و شاخص کیفی سنگ) و مقادیر شاخص گذاری شده سنگ‌شناسی و درجه اهمیت ساختگاه سد سمیلان تعیین شده است. براساس رتبه‌بندی ۷ شاخص اعتبارسنجی خوشه‌بندی، تعداد پهنه خوشه‌ها برابر با ۴ به دست آمد. در این مقاله، از روش خوشه‌بندی براساس شبکه عصبی خودسازمانده، به همراه نحوه قرارگیری گسل‌ها، برای پهنه‌بندی استفاده شده است. در حالت اول، براساس گسل‌ها، پهنه به ۴ بخش تقسیم شد که نوعی پهنه‌بندی دوبعدی است که این پهنه‌بندی در ژرفاهای مختلف، یکسان و مستقل از بعد سوم (ژرفا) بوده و هر گمانه به یک پهنه اختصاص داده شد. در حالت دوم، از روش نقشه خودسازمانده (SOM) که نوعی شبکه عصبی با قابلیت خوشه‌بندی است، استفاده شد. داده‌های ورودی این شبکه، شامل ۳ متغیر جهتی (X, Y, Z)، متغیر ژئوتکنیکی (لوژون و شاخص کیفی سنگ)، مقادیر شاخص گذاری سنگ‌شناسی و درجه اهمیت ساختگاه سد سمیلان است. در مرحله بعد، ۷ متغیر ورودی به صورت بهنجار شده (در دامنه ۰ تا ۱) برای آموزش وارد شبکه شده و خروجی شبکه براساس شاخص‌های اعتبارسنجی خوشه‌ها، داده‌ها به چهار پهنه (خوشه) مجزا اختصاص داده شد. سپس برای تعیین توزیع فضایی شاخص کیفی سنگ، اطلاعات مرتبط به واروی گرافی و ناهمسانگردی در هر چهار پهنه برای دو حالت محاسبه شد. مبنای خوشه‌بندی صحیح، بیشترین تفاوت بین خوشه‌های و بیشترین تشابه درون خوشه‌ای است. برای بررسی و اعتبارسنجی دو حالت خوشه‌بندی، از شاخص کیفیت خوشه‌بندی استفاده شد. شاخص کیفیت خوشه‌بندی به صورت مجموع اختلاف میانگین بین هر دو خوشه تقسیم بر مجموع انحراف معیار خوشه‌ها تعریف شد که مقادیر بیشینه این شاخص، معرف کیفیت بهتر خوشه‌بندی است. در این بررسی مشخص شد که خوشه‌بندی با استفاده از شبکه عصبی خودسازمانده، کیفیت بالاتری نسبت به خوشه‌بندی بر اساس گسل‌ها دارد.

**کلیدواژه‌ها:** پهنه‌بندی، اعتبارسنجی خوشه‌بندی، شاخص کیفی سنگ، شبکه عصبی خودسازمانده، ناهمسانگردی، سد سمیلان، ایران.

\*نویسنده مسئول: امین حسین‌مرشدی

E-mail: morshedy@ut.ac.ir

## ۱- مقدمه

روش‌های اکتشاف ژئوتکنیکی معمولاً پرهزینه و زمان‌بر هستند. از این رو، تمام این روش‌ها، به‌ویژه گمانه‌ها یا گالری‌های اکتشافی، باید به گونه‌ای تعیین محل شوند که در کمترین زمان و با کمترین هزینه، بیشترین اطلاعات را به دست دهند. در سال‌های اخیر، روش‌های نوین متعددی برای بهینه‌سازی کسب اطلاعات از زمین در سطح بین‌المللی ارائه شده است، که برخی از آنها با موفقیت در فرایند اکتشاف ساختگاه‌های بزرگ کشور، به کار گرفته شده‌اند. از آن جمله، روشی برای تعیین محل حفاری‌های جدید، بر اساس کریجینگ شاخص در ساختگاه سد سمیلان (حسین‌مرشدی و معاریان، ۱۳۸۷) و طراحی مرحله‌ای شبکه نمونه‌برداری، براساس متغیرهای ژئوتکنیکی و ویژگی‌های کیفی ساختگاه، با استفاده از کریجینگ و شبکه عصبی، در همین سد (حسین‌مرشدی و معاریان، ۱۳۸۹) است. در بررسی حاضر، پهنه‌بندی ساختگاه سد سمیلان با استفاده از گسل‌های موجود و الگوریتم نقشه خودسازمانده بررسی شده است.

روش‌های اکتشاف ژئوتکنیکی معمولاً پرهزینه و زمان‌بر هستند. از این رو، تمام این روش‌ها، به‌ویژه گمانه‌ها یا گالری‌های اکتشافی، باید به گونه‌ای تعیین محل شوند که در کمترین زمان و با کمترین هزینه، بیشترین اطلاعات را به دست دهند. در سال‌های اخیر، روش‌های نوین متعددی برای بهینه‌سازی کسب اطلاعات از زمین در سطح بین‌المللی ارائه شده است، که برخی از آنها با موفقیت در فرایند اکتشاف ساختگاه‌های بزرگ کشور، به کار گرفته شده‌اند. از آن جمله، روشی برای تعیین محل حفاری‌های جدید، بر اساس کریجینگ شاخص در ساختگاه سد سمیلان (حسین‌مرشدی و معاریان، ۱۳۸۷) و طراحی مرحله‌ای شبکه نمونه‌برداری، براساس متغیرهای ژئوتکنیکی و ویژگی‌های کیفی ساختگاه، با استفاده از کریجینگ و شبکه عصبی، در همین سد (حسین‌مرشدی و معاریان، ۱۳۸۹) است. در بررسی حاضر، پهنه‌بندی ساختگاه سد سمیلان با استفاده از گسل‌های موجود و الگوریتم نقشه خودسازمانده بررسی شده است.

در مطالعات علوم زمین، منطقه مورد مطالعه معمولاً به صورت پیوسته و یکپارچه بررسی می‌شود. این در حالی است که ممکن است عواملی مانند ناپیوستگی، گسل و تغییر سازند باعث تقسیم منطقه به پهنه‌های مختلف، با ویژگی‌های متفاوت شده باشد. بنابراین برای دستیابی به نتایج معتبرتر، ابتدا باید منطقه مورد مطالعه را به چند منطقه (خوشه) تقسیم کرد. بر اساس تعریف، در یک منطقه (خوشه)، متغیرهای مورد نظر باید با هم بیشترین شباهت و با مناطق دیگر، بیشترین تفاوت را داشته باشند. یکی از مسایل کلیدی در خوشه‌بندی، تعیین مرز بین دو منطقه یا خوشه است، که بیانگر ناحیه گذار و نحوه تغییر متغیر مورد مطالعه بین دو پهنه است (Wingle, 1997; Wingle & Poeter, 1996). در بسیاری از موارد، برای ساده‌سازی مسئله، میانگین رفتار مناطق مختلف محاسبه و به صورت یک منطقه واحد در نظر گرفته می‌شود. این روش، ممکن است باعث تعدیل دروغین در عملیات برآورد شود که دارای خطای زیادی است. بنابراین در این موارد، تقسیم منطقه به پهنه‌های مختلف پیشنهاد می‌شود.

الگوریتم نقشه خودسازمانده (SOM)، نوعی شبکه عصبی است که در سال‌های اخیر کاربردهای فراوانی را در زمینه‌های علوم مهندسی تا پزشکی، زیست‌شناسی و اقتصاد در میان پژوهشگران پیدا کرده است. الگوریتم نقشه خودسازمانده از روش یادگیری رقابتی برای آموزش استفاده می‌کند و بر مبنای ویژگی‌های خاصی از مغز انسان توسعه یافته است. نقشه خودسازمانده ابزار بسیار خوبی برای خوشه‌بندی داده‌هاست و می‌تواند روابط آماری غیرخطی بین داده‌های ورودی را به روابط هندسی ساده تبدیل کند. محاسبات این روش یک فرایند رگرسیونی بازگشتی نامتغیری (ناپارامتری) است که رگرسیون یک مجموعه مشخص از بردارهای مدل را به فضای بردارهای قابل مشاهده در قالب الگوریتمی تبدیل می‌کند. خوشه‌ها در یک فرایند یادگیری رقابتی نسبت به متغیرهای ورودی منظم می‌شوند.

## ۲- ساختگاه مورد مطالعه

سد سمیلان در جنوب ایران و در شهرستان رودان استان هرمزگان واقع شده است. با در نظر گرفتن شرایط ریخت‌شناسی و زمین‌شناسی مهندسی گستره طرح، ساختگاه

حفاری شده در همان منطقه، صورت گرفته است. در ساختگاه مورد مطالعه، جسم سد (بدنه اصلی سد) در درجه اول اهمیت، تأسیسات تخلیه آب (تونل‌ها، دریچه‌ها و سرریزها) در درجه دوم و مناطق اطراف بدنه سد، مانند حوضچه آرامش و مخزن سد در درجه سوم اهمیت، در نظر گرفته شده است. به این ترتیب، با تعیین درجه اهمیت هر بخش (سازه)، مقدار شاخص مربوط به گمانه‌های موجود در همان بخش، تخصیص داده می‌شود. برای گمانه‌هایی که بین دو سازه با اهمیت متفاوت، مشترک هستند، شاخص دارای اهمیت بالاتر، در نظر گرفته شده است. براساس شاخص گذاری سازه‌ها، گمانه‌های موجود، با توجه به اهمیتی که دارند، به چهار شاخص تقسیم شدند. گمانه‌های ۷، ۸ و ۹، که دارای اهمیتی بسیار کم و خارج از ساختگاه هستند، مطالعه نشده‌اند (جدول ۱).

### ۳-۴. شاخص گذاری سنگ‌شناسی

محدوده ساختگاه در مجموعه دگرگونی بچگان قرار گرفته است. افزون بر آن، نهشته‌های رودخانه‌ای و پادگانه‌های آبرفتی در بستر رودخانه و واریزه‌های پوششی در دامنه‌ها دیده می‌شود. در بررسی حاضر، شاخص‌بندی بر اساس چهار نوع جنس سنگ‌شناسی روباره و هوازده، کنگلومرا، شیست و متادیوریت، صورت گرفته است که عبارتند از:

– **روباره:** شامل پادگانه‌های آبرفتی، واریزه‌های دامنه‌ای و نهشته‌های آبرفتی و سیلابی است. پادگانه‌های آبرفتی، در بالادست محور و در ساحل راست رودخانه با اختلاف تراز نزدیک به ۵ متر نسبت به بستر رودخانه، دیده می‌شود که سبترای این واحد در حد ۲ تا ۳ متر است. پوشش‌های واریزه‌ای، حاصل تخریب مکانیکی سنگ‌های شیستی ساختگاه هستند که سبترای کمتر از ۱ متر دارند. نهشته‌های آبرفتی و سیلابی، بیشتر درشت‌دانه هستند که ناشی از حمل رودخانه و نیز از آوردهای سیلابی هستند و سبترای آنها غالباً کمتر از ۳ متر است.

– **کنگلومرای چند آمیختگی:** گسترش این واحد، به‌ویژه در تکیه‌گاه چپ و به‌سوی بالادست، چشم‌گیر است و بخش زیادی از گستره دریاچه در شاخه کندر را این واحد تشکیل می‌دهد. در محدوده دایک‌های تکیه‌گاه چپ، سبترای این واحد در گمانه‌های اکتشافی تا ۲۵ متر ثبت شده است. این کنگلومرا از نوع چند آمیختگی است، که به‌طور عمده از خرده‌سنگ‌های دگرگونی تشکیل شده است که با سیمان کربناتی به هم متصل شده‌اند.

– **شیست:** محدوده ساختگاه در درون مجموعه دگرگونی بچگان واقع شده است که شیست‌های بچگان به‌عنوان سنگ بستر اصلی ساختگاه مطرح هستند. در محدوده محور سد، سنگ پی را شیست‌های پلیتی تشکیل می‌دهند و در واریزه‌های دامنه‌ای قطعات شیست دارای گرونا قابل مشاهده است. ترکیب چیره شیست‌های ساختگاه شامل آمفیبول-اپیدوت شیست، اپیدوت-کلریت شیست و کلریت شیست آهنکی است.

– **متادیوریت:** در پی مرکزی ساختگاه، توده سنگ‌هایی آذرین با ترکیبی در حد دیوریت، که آثار دگرگونی در آن مشهود است (متادیوریت)، دیده می‌شود. در این توده سنگ‌ها، ضمن حفظ بافت اولیه آذرین، نوعی جهت‌یافتگی حاصل از تأثیر فرایند دگرگونی، دیده می‌شود. این سنگ‌ها، در گمانه‌های حفار شده در پی مرکزی، تا ژرفای نزدیک به ۳۰ متر مشاهده می‌شوند.

به‌طور کلی می‌توان گفت که عدم قطعیت و ریسک‌پذیری از سنگ‌های روباره و هوازده به سمت متادیوریت کاهش می‌یابد. این شاخص به‌طور کلی بیانگر تغییر سنگ‌شناسی سطح به ژرفاست. بر اساس شاخص گذاری جنس سنگ‌ها، بخش روباره و هوازده در شاخص ۱، سنگ کنگلومرا در شاخص ۲، سنگ شیست در شاخص ۳ و سنگ متادیوریت در شاخص ۴ پهنه‌بندی شدند.

### ۴- شاخص‌های اعتبار سنجی تعداد خوشه‌ها

محاسبه درستی اعتبار خوشه‌بندی، به‌صورت نسبی بر اساس متغیرهای مختلف

سد در ۲/۵ کیلومتری پایین دست پیوستگاه رودخانه‌های کندر و دزدی با مختصات "۱۹' ۲۱' ۵۷" طول خاوری و "۴۱' ۳۷' ۲۷" عرض شمالی انتخاب شده است.

محدوده ساختگاه در مجموعه سنگ‌های دگرگونی بچگان قرار گرفته است. این مجموعه به‌طور عمده از شیست پلیتی، سنگ آهک‌های بازتبلور یافته و سنگ‌های آذرین دگرگون، تشکیل شده است. در محدوده ساختگاه سه گسل وجود دارد. دو گسل به نام‌های F1 و F2، روند شمال خاوری - جنوب باختری دارند. گسل F1 از نوع عادی (نرمال)، با طول تقریبی ۲۰۰ متر و دارای شیب حدود ۵۰ درجه به سمت جنوب خاوری است. گسل F2 دارای مؤلفه امتدادلغز و چپ‌گرد است که طول آن حدود ۱۵۰ متر است. شیب این گسل ۵۸ درجه به سمت جنوب خاوری اندازه‌گیری شده است. گسل F3 دارای شیب نزدیک به قائم و از نوع عادی است. در شکل ۱ موقعیت گسل‌های منطقه، محل گمانه‌ها و سازه‌های مختلف سد، نمایش داده شده است (مهندسین مشاور مهتاب قدس، ۱۳۸۷).

### ۳- داده‌های مورد مطالعه

در این بررسی، تعداد گمانه‌های مورد مطالعه ۲۴ عدد است که به‌صورت شبکه نامنظم در منطقه توزیع شده‌اند. تراکم شبکه نمونه‌برداری در محدوده بدنه سد و تونل‌ها چگال‌تر است. تمام گمانه‌ها به‌صورت قائم حفاری شده‌اند و ژرفای گمانه‌ها بین ۱۲۰-۲۵ متر متغیر است. برای آگاهی از شرایط زیرسطحی پی، در تمامی گمانه‌های حفار شده، آزمایش‌های لازم درون‌چاهی (نفوذپذیری به‌روش‌های لوژون و لوفران و نفوذ استاندارد به روش SPT) انجام و برای ارزیابی شاخص کیفی توده سنگ همزمان با حفاری گمانه‌های اکتشافی، شاخص کیفی سنگ (RQD) برای همه گمانه‌ها اندازه‌گیری شده است. در این مطالعه، از مقادیر RQD و لوژون به‌عنوان متغیرهای ژئوتکنیکی سنگ در تعیین محل بهینه حفاری استفاده می‌شود (مهندسین مشاور مهتاب قدس، ۱۳۸۷).

### ۳-۱- متغیر لوژون

در این ساختگاه آزمایش تراوایی لوژون برای تعیین تراوایی توده سنگ، در درون گمانه‌ها انجام شده است. مقدار لوژون بزرگ‌تر از ۳۰، نمایانگر قابلیت تراوایی بالا و مقادیر لوژون کمتر از ۵، نمایشگر تراوایی پایین در نظر گرفته شده است (Clayton et al., 1995). شیست‌های بچگان به‌عنوان سنگ بستر اصلی ساختگاه، تراوایی پایینی دارند. در محور سد در تکیه‌گاه چپ، تراوایی قطعات غالباً زیر ۳ لوژون ثبت شده است. برای قطعاتی که در ژرفاهای کمتر از ۴۵ متر قرار دارند، مقدار تراوایی در حد متوسط تا بالاست. اما در قطعات واقع در ژرفاهای بیشتر از ۴۵ متر، تراوایی عموماً پایین است. به‌طور کلی، با افزایش ژرفا، مقدار تراوایی و لوژون کاهش می‌یابد.

### ۳-۲- متغیر RQD

شاخص کیفی سنگ (RQD) بر اساس تعیین کیفیت توده سنگ، بر اساس مغزه بازیافتی از علمیات حفاری توسط مته الماسی، استوار است. در تکیه‌گاه چپ سد سمیلان، میانگین شاخص کیفی سنگ در حدود ۶۳ درصد است که در بخش میانی سنگ‌های رده متوسط قرار می‌گیرد. در امتداد محور سد، در ژرفاهای ۵۰ و ۷۸ متر، میانگین شاخص کیفی برابر ۶۲ درصد است، اما در گمانه واقع در رقوم تاج که ژرفای آن ۱۲۰ متر بوده است، این شاخص نزدیک به ۸۰ درصد است. در تکیه‌گاه راست میانگین کل شاخص کیفی در حدود ۵۴ درصد است که در مقایسه با تکیه‌گاه چپ، در درجه پایین‌تری قرار می‌گیرد. به‌طور کلی می‌توان گفت با افزایش ژرفا، میانگین این شاخص افزایش می‌یابد.

### ۳-۳. شاخص گذاری اهمیت سازه‌ها

یکی از کاربردهای شاخص گذاری، تبدیل داده‌های کیفی به کمی، به منظور مدل‌سازی و مطالعات عددی است (Vann & Geoval, 2003). در بررسی حاضر، براساس درجه اهمیت و حساسیت سازه‌های مختلف، شاخص گذاری برای گمانه‌های

خوشه‌بندی و تعداد خوشه‌ها صورت می‌گیرد. اعتبارسنجی خوشه‌بندی، یکی از مسائل بسیار مهم در تحلیل خوشه‌بندی است، به‌علت اینکه نتایج خوشه‌بندی باید معتبر باشد، در بیشتر الگوریتم‌های خوشه‌بندی، تعداد خوشه‌ها از عوامل مهم قابل بررسی است. روش‌های زیادی برای تعیین تعداد بهینه خوشه‌ها وجود دارد که بر اساس بیشترین تفاوت بین خوشه‌های و بیشترین تشابه درون خوشه‌ای تعیین می‌شود (Topchy & Punch, 2003). تعداد بهینه خوشه‌ها براساس متغیر ژئوتکنیکی لوژون و شاخص کیفی سنگ و مقادیر شاخص‌گذاری شده سنگ‌شناسی و درجه اهمیت ساختگاه سد سمیلان انجام شده است.

#### ۴-۱. شاخص اعتبارسنجی دان

روش دان، بر پایه ایده‌ای که درون خوشه‌ها به هم نزدیک‌تر و خوشه‌ها به‌خوبی از هم جدا باشند، بیان شده است. برای هر بخش از خوشه‌ها،  $c_i$  نماینده هر یک از خوشه‌هاست. شاخص اعتبارسنجی دان ( $D$ ) توسط رابطه زیر محاسبه می‌شود (Dunn, 1974):

$$D = \min_{1 \leq i \leq n} \left\{ \min_{\substack{1 \leq j \leq n \\ i \neq j}} \left\{ \frac{d(c_i, c_j)}{\max_{1 \leq k \leq n} \{d'(c_k)\}} \right\} \right\} \quad (1)$$

که در آن  $d(c_i, c_j)$  فاصله بین خوشه‌های  $c_i$  و  $c_j$  (فاصله بین خوشه‌های)،  $d'(c_k)$  فاصله درون خوشه‌ای، خوشه  $c_k$  و  $n$  تعداد خوشه‌هاست. هدف اصلی این شاخص، بیشینه‌سازی فاصله بین خوشه‌ها و کمینه‌سازی فاصله درون خوشه‌ای است که بر اساس راهبرد کمینه‌سازی تعداد خوشه‌ها صورت می‌گیرد. بنابراین تعداد بهینه خوشه‌ها براساس بیشینه‌سازی شاخص اعتبارسنجی دان صورت می‌پذیرد. تعداد خوشه‌های بهینه براساس شاخص دان، ۴ خوشه است (شکل ۲- الف).

#### ۴-۲. شاخص اعتبارسنجی دیویس-بولدین

شاخص اعتبارسنجی دیویس-بولدین، تابعی از نسبت مجموع پرش درون خوشه به فاصله بین خوشه‌هاست که در ادامه به رابطه آن اشاره شده است

$$DB = \frac{1}{n} \sum_{i=1}^n \max_{i \neq j} \left\{ \frac{S_n(Q_i) + S_n(Q_j)}{S(Q_i, Q_j)} \right\} \quad (2)$$

که  $n$  تعداد خوشه‌ها،  $S_n$  میانگین فاصله داده‌های خوشه از مرکز خوشه و  $S(Q_i, Q_j)$  فاصله بین مراکز خوشه‌هاست. بنابراین هنگامی که درون خوشه به هم نزدیک و خوشه‌ها از یکدیگر دور باشند، این نسبت کوچک می‌شود. مقدار کوچک شاخص اعتبارسنجی دیویس-بولدین، نمایش خوشه‌بندی معتبر است. تعداد خوشه‌های بهینه براساس شاخص دیویس-بولدین، ۹ خوشه است (شکل ۲- ب).

#### ۴-۳. شاخص اعتبارسنجی سیلوهته (Silhouette)

روش اعتبارسنجی سیلوهته، براساس میانگین فاصله هر یک از نمونه‌های یک خوشه با تمام نمونه‌های موجود در همان خوشه و میانگین فاصله کل نمونه موجود در خوشه‌های دیگر با یک خوشه مشخص تعریف می‌شود. براساس این دیدگاه، برای هر خوشه میزان پراکندگی و همبستگی داده‌ها تعیین می‌شود، که مقادیر بیشینه این شاخص برای تعیین تعداد بهینه خوشه به کار می‌رود (Rousseeuw, 1987):

$$S(i) = \frac{b(i) - a(i)}{\max\{a(i), b(i)\}} \quad (3)$$

که  $a(i)$  نمایانگر تشابه نداشتن یک نمونه با نمونه‌های دیگر در یک خوشه و  $b(i)$  نمایانگر تشابه نداشتن یک نمونه نسبت به همه نمونه‌های موجود در خوشه‌های دیگر (نزدیک‌ترین خوشه‌ها) است.

مقدار شاخص اعتبارسنجی سیلوهته بین ۱- تا ۱ متغیر است. در صورتی که این شاخص برابر ۱ باشد، نمایانگر این است که خوشه‌بندی به‌درستی صورت گرفته است. اگر مقدار شاخص نزدیک صفر باشد، این بدین معنی است که نمونه را می‌توان به یک خوشه نزدیک‌تر نسبت داد و نحوه قرارگیری نمونه از دو خوشه، به یک اندازه دور است. در صورتی که این شاخص برابر ۱- باشد، این بدین معنی است که

#### ۴-۴. شاخص C

این شاخص بر اساس مقایسه فاصله‌های جفت نمونه‌های موجود در یک خوشه و تمام خوشه‌ها نسبت به هم سنجیده می‌شود و به‌صورت یک رابطه بهنجارسازی تعریف می‌شود (Hubert & Schultz, 1976):

$$C = \frac{S - S_{\min}}{S_{\max} - S_{\min}} \quad (4)$$

که  $S$  مجموع فاصله هر یک از جفت‌های تشکیل شده از یک خوشه معین است،  $S_{\min}$  مجموع کوچک‌ترین فواصل بین جفت نمونه‌های موجود در خوشه‌های مختلف و  $S_{\max}$  مجموع بزرگ‌ترین فاصله‌های بین جفت نمونه‌های موجود در خوشه‌های مختلف است. بنابراین مقادیر کوچک شاخص  $C$  نمایانگر خوشه‌بندی صحیح است. تعداد خوشه‌های بهینه براساس شاخص  $C$ ، ۹ خوشه است (شکل ۲- د).

#### ۴-۵. شاخص‌های اعتبارسنجی کالینسکی-هاراباسز، هارتیگان و کرزانوسکی-لای

در این سه روش از مجموع مربعات پرش بین خوشه‌ها و درون خوشه‌ای استفاده می‌شود. که شاخص‌های اعتبارسنجی کالینسکی-هاراباسز و هارتیگان به‌صورت زیر تعریف می‌شود (Calinski & Harabasz, 1974; Hartigan, 1975):

$$CH_k = \frac{tr(B_k)/(k-1)}{tr(W_k)/(n-1)} = \frac{(B_k)(k-1)}{(W_k)(n-k)} \quad (5)$$

$$H_k = \left( \frac{tr(W_k)}{tr(W_{k+1})} \right) (n-k-1) \quad (6)$$

که  $tr(B_k)$  مجموع قطر ماتریس مجموع مربعات بین خوشه‌ها،  $tr(W_k)$  مجموع قطر ماتریس مجموع مربعات درون خوشه‌ای،  $k$  تعداد خوشه‌های موجود و  $n$  تعداد مجموعه داده‌ها است. تعداد خوشه‌های بهینه براساس شاخص کالینسکی-هاراباسز و هارتیگان، ۴ خوشه است (شکل‌های ۲- ه و و).

برای محاسبه شاخص کرزانوسکی-لای، تابع تفاوت بر اساس ماتریس مجموع مربعات درون خوشه‌ای تعریف و نسبت تابع تفاوت برای خوشه‌ها دو به دو بررسی می‌شود که  $m$ ، بعد داده‌های موجود است (Krzanowski & Lai, 1985):

$$diff_k = (k-1)^{2/m} tr(W_{k-1}) - k^{2/m} tr(W_k) \quad (7)$$

$$KL_k = \frac{|diff_k|}{|diff_{k+1}|}$$

مقادیر کوچک شاخص هارتیگان و مقادیر بزرگ شاخص‌های کالینسکی-هاراباسز و کرزانوسکی-لای نمایانگر خوشه‌بندی مناسب است. تعداد خوشه‌های بهینه براساس شاخص کرزانوسکی-لای، ۳ خوشه است (شکل ۲- ز).

#### ۵- تعیین تعداد بهینه خوشه

در مرحله اول، ارتباط مستقیم یا غیرمستقیم بین شاخص‌های اعتبارسنجی خوشه‌بندی و تعداد خوشه‌ها بررسی می‌شود. مقادیر بیشینه شاخص‌های کالینسکی-هاراباسز، کرزانوسکی-لای، دان و سیلوهته و در مقابل مقادیر کمینه شاخص‌های دیویس-بولدین، هارتیگان و  $C$  تعداد بهینه خوشه‌ها را مشخص می‌کنند (شکل ۲- ح و ط). برای نمایش شاخص‌های متفاوت در کنار هم ابتدا می‌بایست مقدار شاخص‌ها را بهنجارسازی کرد (Wang et al., 2009).

براساس ۷ شاخص، اعتبارسنجی خوشه‌بندی و رابطه مستقیم یا غیرمستقیم با تعداد خوشه‌ها رتبه‌بندی شدند. در این رتبه‌بندی، ۴ شاخص از ۷ شاخص، تعداد بهینه خوشه را ۲، ۴ شاخص از ۷ شاخص، تعداد بهینه خوشه را ۹ و ۱ شاخص از ۷ شاخص، تعداد بهینه خوشه را ۳ نمایش می‌دهند (شکل ۲- ی و جدول ۲).

## ۶- تعیین مناطق و ارتباط بین آنها

تعیین مناطق، نخستین قدم در فرایند پهنه‌بندی است. این فرایند برای استفاده مناطق مختلف تکرار می‌شود و بیشترین عدم قطعیت مرتبط به موقعیت مرزهاست. ارتباط بین مرزهای مناطق ممکن است به چند صورت مختلف باشد که در مرز ناگهانی مناطق کاملاً از هم مجزا هستند، مرز تدریجی، یک پهنه به‌طور نامحدود با پهنه دیگر ادغام می‌شود و مرز فازی، یک منطقه به‌صورت تدریجی در فاصله محدود به هم رسیده و از هم جدا می‌شود. در مرزهای ناگهانی، دو واحد یا سازند در تماس با یکدیگر هستند، یا به‌عبارت دیگر دو بخش از هم مجزا هستند. برای مثال، می‌توان به گسل‌ها و سازندهای زمین‌شناسی اشاره کرد. در این وضعیت، نمی‌توان از داده‌های یک بخش برای برآورد توزیع فضایی در بخش دیگر استفاده کرد (Wingle, 1997). در بعضی از محیط‌ها، طبقه‌بندی در یک منطقه با منطقه دیگر مجزا است، اما این تغییرات به‌صورت تدریجی است. برای مثال نوعی نهشته‌های ساحلی که به‌صورت ماسه‌های ساحلی به‌داخل رس و شیل‌های دریایی طبقه‌بندی شده‌اند. ارتباط بین مناطق فازی شبیه ارتباط تدریجی است، اما فاصله منطقه گذار از نظر گسترش محدود باشد (جدول ۳). در مطالعه حاضر پهنه‌بندی به دو صورت انجام شده است. در حالت اول گسل‌های منطقه به چهار پهنه تقسیم شده و در حالت دوم الگوریتم شبکه نقشه خودسازمانده عامل پهنه‌بندی است. این دو روش براساس داده‌های شاخص کیفی سنگ با هم مقایسه شده‌اند.

### ۶-۱. پهنه‌بندی براساس گسل‌های منطقه

پهنه‌بندی بر اساس گسل‌های موجود در منطقه که نوعی مرزبندی ناگهانی است، منطقه را به چهار پهنه تقسیم کرد. پهنه اول، بین گسل F1 و بخش پایینی گسل F3 واقع شده است که شامل شش گمانه است. پهنه دوم، بین گسل F1 و بخش بالایی گسل F3 واقع شده است که شامل شش گمانه است. زون سوم، بین گسل F2 و بخش پایینی گسل F3 واقع شده است که شامل چهار گمانه است. پهنه چهارم، بین گسل F2 و بخش بالایی گسل F3 واقع شده است که شامل هشت گمانه است (شکل ۳). پهنه‌بندی انجام‌شده توسط گسل‌ها، یک نوع پهنه‌بندی دو بعدی است که مرزهای با همین شکل و راستا در ژرفا امتداد یافته‌اند. در این حالت هر گمانه به یک منطقه اختصاص داده می‌شود (شکل ۴). همان‌گونه که پیش‌تر اشاره شد از داده‌های شاخص کیفی سنگ برای مطالعه خوشه‌بندی استفاده شده است. به‌منظور نمایش توزیع آماری مقادیر شاخص کیفی سنگ، از نمودار ستونی استفاده شده است. توزیع مقادیر شاخص کیفی سنگ نزدیک به صفر در هر چهار پهنه، بیشترین فراوانی را دارد که مقادیر زیاد شاخص کیفی سنگ مرتبط با مناطق با خردشدگی پایین واقع در ژرفاست است و در مقابل مقادیر کم شاخص کیفی سنگ مرتبط با مناطق با خردشدگی بالا و هوازده واقع در سطح است (شکل ۵).

### ۷- کلیاتی در مورد شبکه نقشه خودسازمانده

الگوریتم نقشه خودسازمانده (Self Organization Map (SOM)) یک رابطه رگرسیونی بازگشتی نامتغیری است که یک مجموعه از بردارهای  $m \in R^n$  را به فضای بردارهای  $x \in R^n$  از طریق مراحل زیر نگاشت می‌کند:

در هر مرحله از آموزش، یک بردار نمونه  $x$  از مجموعه داده‌های ورودی به‌صورت تصادفی انتخاب می‌شود و فاصله‌های بین  $x$  و همه بردارهای پیش‌نمونه محاسبه می‌شوند. براساس کمینه‌سازی فاصله یک نمونه با نمونه‌های دیگر، بهترین میزان تطبیق (Best Matching Unit (BMU)) توسط رابطه ۸ محاسبه می‌شود (Vesanto et al., 2000) (شکل ۶):

$$\|x - m_b\| = \min \{\|x - m_i\|\} \quad (8)$$

در مرحله بعد، بردارهای پیش‌نمونه به‌روز رسانی می‌شوند که بهترین میزان تطبیق

و همسایه‌های توپولوژیک آن به نزدیکی بردار ورودی در فضای ورودی منتقل می‌شوند. برای به‌روزرسانی برای بردار پیش‌نمونه واحد  $i$  از رابطه ۹ استفاده می‌شود (Vesanto et al., 2000):

$$m_i(t+1) = m_i(t) + \alpha(t) h_{bi}(t) [x(t) - m_i(t)] \quad (9)$$

که  $t$  نمایانگر زمان است که شبکه خودسازمانده به صورت یک فرایند بازگشتی آموزش داده می‌شود،  $\alpha(t)$  آهنگ یادگیری و آموزش که بیانگر میزان تطبیق است و به‌صورت یکنواخت با مراحل رگرسیونی (زمان) کاهش می‌یابد و  $h_{bi}(t)$  هسته همسایگی است که یک تابع کاهشی از فاصله بین مدل‌های  $\lambda$ مین و  $\lambda$ مین بر روی شبکه نقشه است و بر روی واحد برنده متمرکز می‌شود. تابع همسایگی به‌صورت رابطه ۱۰ در نظر گرفته می‌شود (Vesanto & Alhoniemi, 2000).

$$h_{bi}(t) = \exp\left(-\frac{\|r_i - r_b\|^2}{2\sigma^2(t)}\right) \quad (10)$$

که  $\sigma^2(t)$  برابر با شعاع گسترش تابع همسایگی و  $r_i \in R^2$  و  $r_b \in R^2$  مکان نرون‌های  $\lambda$ مین و  $\lambda$ مین بر روی شبکه خودسازمانده هستند که در این حالت نیز همراه با گذشت زمان و مراحل رگرسیونی، شعاع گسترش کاهش می‌یابد. برای تعیین تعداد خوشه‌ها رویکرد مشخصی وجود ندارد و فقط قاعده‌ای کلی از ۲ تا  $\sqrt{N}$  برای تعیین تعداد خوشه‌ها ارائه شده است که  $N$  تعداد نمونه‌ها در مجموعه داده‌هاست. الگوریتم خودسازمانده، تابع خطا در رابطه ۱۱ را کمینه می‌کند (Vesanto & Alhoniemi, 2000):

$$E = \sum_{i=1}^N \sum_{j=1}^C h_{bi} \|x_i - m_j\|^2 \quad (11)$$

که  $C$  تعداد خوشه‌ها است، هسته همسایگی  $h_{bi}(t)$  در واحد  $b$  تمرکز یافته که بیانگر بهترین میزان تطبیق بردار  $x_i$  است و برای واحد  $j$  بررسی می‌شود. با توجه به رابطه ۱۱، SOM برای خطاهای بزرگ (فواصل بیشتر) امتیاز منفی بیشتری را در نظر می‌گیرد. داده‌های ورودی SOM از بردارهایی با  $n$  عنصر تشکیل شده‌اند. برای خوشه‌بندی این بردارهای ورودی، ترکیب آرایه‌ای مختلفی را می‌توان در نظر گرفت. همان‌گونه که پیش از این بیان شد، رویکرد از پیش تعیین شده‌ای برای تعیین تعداد دسته‌ها وجود ندارد. گفتنی است که تعداد نرون‌ها در هر آرایه از ضرب تعداد دسته‌های دو عنصر آرایه در یکدیگر به‌دست می‌آید. همچنین، برای مقایسه فواصل درون خوشه‌ای و بین خوشه‌ای از فاصله اقلیدسی استفاده می‌شود (Schatzmann, 2003).

در این الگوریتم در هر مرتبه آموزش، همه بردارهای آموزشی هر کدام یک بار به‌صورت تصادفی به شبکه اعمال و ضرایب وزنی و بایاس پس از ارائه هر بردار به‌روز رسانی می‌شوند. شبکه نقشه خودسازمانده با به‌کارگیری الگوریتم آموزش افزایشی تصادفی (Random order incremental training) برای خوشه‌بندی مجموعه داده‌های ورودی، آموزش داده شده است. آموزش، زمانی متوقف می‌شود که یکی از معیارهای بیشترین تعداد دوره‌های آموزش، دستیابی به کمترین خطا یا رسیدن به بیشترین مقدار زمان تحقق یابد. سپس شبکه، نرون برنده را تعیین می‌کند و وزن‌های نرون برنده و نرون‌های مجاور نرون برنده در هر مرحله یادگیری به بردار ورودی نزدیک‌تر می‌شوند. وزن نرون‌های برنده و وزن نرون‌های در همسایگی آن متناسب با آهنگ یادگیری تغییر داده می‌شوند. آهنگ یادگیری و فاصله همسایگی در قالب در دو مرحله ترتیب و تنظیم به‌روز رسانی می‌شوند.

در مرحله ترتیب، آهنگ یادگیری از یک مقدار اولیه آغاز می‌شود و کاهش می‌یابد و فاصله همسایگی از بیشینه فاصله نرونی تا ۱ کاهش داده می‌شود. در مرحله ترتیب انتظار می‌رود که وزن‌های نرونی خودشان را در فضای ورودی با موقعیت‌های نرونی هم‌بسته سازگار و با گام‌های بزرگ یک آرایش کلی در وزن همه نرون‌ها ایجاد کنند، از این رو، متغیر آهنگ یادگیری مقدار قابل توجهی است و با انجام تعداد گام‌های معین، مرحله ترتیب متوقف می‌شود. در طول مرحله تنظیم، بر خلاف

رنگ تیره تر، فاصله بین نرون‌ها بیشتر و هر چه رنگ روشن تر فاصله بین نرون‌ها کمتر است. در صورت مشاهده یک نوار تیره در ماتریس فاصله وزنی، می‌توان نرون‌ها را به دو بخش تقسیم کرد (Demuth et al., 2008) (شکل ۹).

یکی دیگر از قابلیت‌های این روش، تعیین تعداد نمونه‌های مرتبط با هر نرون است. نمونه‌ها به ترتیب ۲۷٪، ۲۶٪، ۲۵٪ و ۲۲٪ از کل نمونه‌ها بین ۴ خوشه تقسیم شده است (شکل ۱۰). خروجی مدل شبکه عصبی یادشده، هر نمونه را به یکی از چهار زون نسبت داده و چهار پهنه مجزا را تشکیل می‌دهد که تقریباً مرزبندی زون‌ها، ناگهانی است (شکل ۱۱).

#### ۸-۱. اختصاص هر نمونه به یک زون

بر اساس مدل پهنه‌بندی خروجی SOM، هر نمونه را باید به یک منطقه اختصاص داد. به دلیل اینکه در هر پهنه تغییرات ژرفا وجود دارد و یک مدل سه‌بعدی واقعی است، احتمال دارد بخش بالایی گمانه، مربوط به یک پهنه و بخش پایینی گمانه، مربوط به پهنه دیگر باشد (شکل ۱۲).

در حالت پهنه‌بندی با استفاده از SOM توزیع آماری مقادیر شاخص کیفی سنگ توسط نمودار ستونی نمایش داده شده است. توزیع مقادیر شاخص کیفی سنگ همانند حالت پهنه‌بندی گسل‌ها مقادیر نزدیک به صفر در هر چهار پهنه، بیشترین فراوانی را دارد و با این تفاوت که مقادیر بیشینه شاخص کیفی سنگ (صد) از نظر فراوانی در درجه دوم هستند که در حالت پهنه‌بندی براساس گسل‌ها چنین نیست. در شکل ۱۳ توزیع آماری مقادیر شاخص کیفی سنگ در چهار پهنه انتخابی ساختگاه سد سیمان نمایش داده شده است.

#### ۸-۲. واریوگرافی

همسایگی نقاط، بر اساس ارتباط بین مناطق انتخاب می‌شوند. همچنان که پیش تر گفته شد، اگر مرز به صورت ناگهانی باشد، هیچ نقطه‌ای از دیگر سوی مرز منطقه، مورد استفاده قرار نمی‌گیرد. در مقابل، اگر مرز از نوع تدریجی باشد، نزدیک ترین نقاط، صرف نظر از پهنه‌شان قابل استفاده هستند. در صورتی که ناحیه‌گذار از نوع فازی باشد، نقاط از منطقه همسایه انتخاب می‌شود، نقطه‌ای که در یک فاصله مشخص هستند، مورد استفاده قرار می‌گیرند. در این مقاله، مرزها از نوع تغییرات ناگهانی بوده که در حالت اول براساس گسل‌ها و در حالت دوم براساس خوشه‌بندی SOM به دست آمده است و داده‌های هر پهنه در برآورد پهنه دیگر شرکت نمی‌کند.

واریانس مقدار عناصر بین نقاطی به فاصله  $h$  از یکدیگر، می‌تواند همبستگی متقابل مقدار دو نقطه به فاصله  $h$  را بیان کند. در صورت وجود ساختار فضایی، طبیعی است که وابستگی مقادیر نقاط نزدیک به هم بیشتر از وابستگی مقادیر نقاط دور از هم است. بنابراین چنین واریانسی می‌تواند معیاری برای نمایش تأثیرگذاری و یا تأثیرپذیری مقدار یک نمونه روی مقادیر محیط مجاور خود باشد. این واریانس وابسته به فاصله را واریوگرام می‌نامند و با نماد  $2\gamma(h)$  نشان می‌دهند. تحلیل ساختاری برای مدل‌سازی ساختار فضایی، توسط واریوگرام انجام می‌گیرد و خروجی مدل واریوگرام در برآورد متغیرها و توزیع خطای مربوط به آن به کار می‌رود (Myers & Journel, 1990; Webster & Margaret, 2007).

$$2\gamma(h) = \frac{1}{N(h)} \sum_{i=1}^{N(h)} [Z(x+h) - Z(x)]^2 \quad (12)$$

در رابطه بالا،  $h$  فاصله در جهت مشخص بین موقعیت  $x$ ،  $x+h$ ،  $N(h)$  تعداد جفت نمونه‌ها به فاصله  $h$  از یکدیگر،  $\gamma(h)$  مقدار واریوگرام برای فاصله  $h$ ،  $Z(x)$  مقدار نمونه در نقطه  $x$ ،  $Z(x+h)$  مقدار نمونه در نقطه  $x+h$  است.

برای هر یک از چهار پهنه با در نظر گرفتن گسل‌ها، ویژگی‌های مدل برازش شده واریوگرام، متغیر شاخص کیفی سنگ استخراج شد که مدل نمایی برای پهنه‌های اول، دوم و سوم و مدل کروی برای پهنه چهارم انتخاب شد و

مرحله ترتیب آهنگ یادگیری به کندی کاهش می‌یابد و با تغییرات کوچک در وزن‌ها، به نوعی تنظیم نهایی و دقیق‌تری در وزن‌ها می‌رسد که در نهایت منجر به همگرایی می‌شود. در مرحله تنظیم انتظار می‌رود ضمن حفظ نظم توپولوژیک در مرحله ترتیب، وزن‌ها در کل فضای ورودی به طور تصادفی پراکنده شوند. در مرحله تنظیم که مرحله همگرایی است، آهنگ آموزش مقدار کوچک‌تری دارد تا با تغییرات کوچک در وزن‌ها به نوعی تنظیم نهایی و دقیق‌تری در وزن‌ها برسد. بدین ترتیب، نقشه‌های ویژگی در طی فراگیری خوشه‌بندی ورودی‌ها، توپولوژی و توزیع ورودی‌ها را نیز فرا می‌گیرند (Demuth et al., 2008).

#### ۷-۱. کاربرد شبکه نقشه خودسازمانده در خوشه‌بندی

امروزه، الگوریتم نقشه خودسازمانده (SOM) کاربردهای گسترده‌ای را در عرصه‌های مختلف علوم مهندسی، بیولوژی، اقتصاد و ... در میان پژوهشگران یافته است. الگوریتم نقشه خودسازمانده ابزار مناسبی برای خوشه‌بندی داده‌هاست و قادر است روابط آماری غیرخطی بین داده‌های ورودی را به روابط هندسی ساده تبدیل کند که یک رابطه رگرسیونی بازگشتی نامتغیری است، چنان‌که با ارائه هر نمونه، رگرسیون به صورت بازگشتی اجرا می‌شود. بر این اساس می‌توان ادعا کرد که SOM می‌تواند میزان خطاها را اعمال کند. خوشه‌بندی  $Q$ ، به معنای افزایش مجموعه‌ای از داده‌ها، به مجموعه‌ای از خوشه‌های  $Q, i=1, \dots, C$  است و خوشه‌بندی پهنه، افزایش است که فواصل بین نمونه‌های داخلی را کمینه و فواصل بین خوشه‌ها را بیشینه کند (Vesanto & Alhoniemi, 2000).

الگوریتم نقشه خودسازمانده شامل روابط دو بعدی، شبکه‌ای از واحدهای نقشه‌ای است که واحدها توسط یک رابطه همسایگی به مجاورت آنها متصل می‌شوند. تعداد واحدهای نقشه که معمولاً از چند تا چندین هزار تغییر می‌کند، دقت و قابلیت تعمیم SOM را مشخص می‌کنند. هر واحد  $i$  توسط یک بردار پیش‌نمونه  $m_i = [m_{i1}, \dots, m_{id}]$  نمایش داده می‌شود که  $d$  بعد بردار ورودی است. در مرحله آموزش، SOM یک شبکه شبه فزنی به وجود آورده است که توده تشکیل شده از داده‌های ورودی را با هم در می‌آمیزد. داده‌هایی که نزدیک به هم در فضای ورودی قرار گرفته‌اند، به صورت واحدهای نقشه نگاشت می‌شوند. بنابراین SOM همچون یک نقشه توپولوژی است که امکان نمایش، تفسیر و ترتیب خوشه‌بندی را فراهم می‌کند و قادر است از درجه فضای داده‌های ورودی به شبکه دو بعدی واحدهای نقشه نگاشت شود (شکل ۷).

#### ۸- استفاده از SOM در پهنه‌بندی داده‌های شاخص کیفی سنگ

در بررسی حاضر، داده‌های ژئوتکنیکی ساختگاه سد سیمان برای خوشه‌بندی SOM استفاده شده است. داده‌های ورودی SOM مرتبط با سه متغیر جهتی  $(X, Y, Z)$ ، متغیر ژئوتکنیکی لوژون و شاخص کیفی سنگ و در پایان مقادیر شاخص گذاری شده سنگ‌شناسی و درجه اهمیت ساختگاه سد سیمان است. در مرحله بعد،  $\gamma$  متغیر ورودی به صورت بهنجارسازی شده (در دامنه ۰ تا ۱) برای آموزش وارد شبکه شده و خروجی شبکه براساس شاخص‌های اعتبارسنجی خوشه‌ها، تقسیم داده‌ها به چهار زون (خوشه) مجزا است. بنابراین فضای ورودی یک فضای ۷ بعدی است (شکل ۸). برای آموزش SOM، بردار وزن مربوط به هر نرون برای نزدیک شدن به مرکز خوشه بردارهای ورودی، مکان خود را تغییر می‌دهند. همچنین، نرون‌هایی که براساس توپولوژی در مجاورت یکدیگر قرار دارند، برای نزدیک شدن به هم، مکان خود را تغییر می‌دهند. در این بررسی شبکه توپولوژی به صورت شش ضلعی است. هر شش ضلعی نمایانگر یک نرون است و از یک شبکه  $2 \times 2$  استفاده شده است که مجموعاً از ۴ نرون تشکیل شده است. به علت اینکه SOM دارای توپولوژی دو بعدی است، ارتباط بین مراکز خوشه ۷ بعدی، به صورت ۲ بعدی نمایش داده می‌شود. یکی از ابزارهای نمایش SOM، ماتریس فاصله وزنی (Weight distance matrix) (ماتریس U) است. در ماتریس U از خط برای ارتباط نرون‌های مجاور استفاده می‌شود. هر چه

### ۱۰- نتیجه‌گیری

- براساس رتبه‌بندی شاخص‌های اعتبارسنجی خوشه‌بندی، تعداد پهنه خوشه‌ها با استفاده از متغیر ژئوتکنیکی (لوژون و شاخص کیفی سنگ) و مقادیر شاخص گذاری شده سنگ‌شناسی و درجه اهمیت ساختگاه سد سمیلان، ۴ خوشه تعیین شد. - از روش نقشه خودسازمانده (SOM)، برای پهنه‌بندی استفاده شده است. داده‌های ورودی SOM، سه متغیر جهتی (X,Y,Z)، متغیر ژئوتکنیکی (لوژون و شاخص کیفی سنگ)، مقادیر شاخص گذاری شده سنگ‌شناسی و درجه اهمیت ساختگاه سد سمیلان هستند و خروجی شبکه براساس شاخص‌های اعتبارسنجی خوشه‌ها، تقسیم داده‌ها به چهار پهنه (خوشه) مجزا است.

- برای بررسی توزیع فضایی شاخص کیفی سنگ، ابتدا هر نمونه از گمانه را به یک پهنه اختصاص داده، سپس اطلاعات مرتبط به واریوگرافی و ناهمسانگردی در هر چهار پهنه براساس گسل‌ها و شبکه خود سازمانده تعیین شد.

- شاخص کیفیت خوشه‌بندی به صورت مجموع اختلاف میانگین بین هر دو خوشه تقسیم بر مجموع انحراف معیار خوشه‌ها تعریف شد. شاخص کیفیت خوشه‌بندی برای هر دو حالت پهنه‌بندی خودسازمانده و پهنه‌بندی با گسل‌ها بررسی شد که مقدار این شاخص در حالت پهنه‌بندی خود سازمانده برابر ۰/۶۷ و در حالت پهنه‌بندی با گسل‌ها برابر ۰/۵۶ است که بیانگر کیفیت بهتر خوشه‌بندی پهنه‌بندی روش خود سازمانده است.

- در مطالعه حاضر مشخص شد که روش نقشه خودسازمانده انعطاف و کارایی قابل توجهی برای پهنه‌بندی دارد، قابلیت در نظر گرفتن متغیرهای مختلف را داراست و از یادگیری رقابتی برای خوشه‌بندی استفاده می‌کند.

بیشترین شعاع تأثیر به ترتیب متعلق به پهنه چهارم، دوم، اول و سوم است (شکل ۱۴). برای هر یک از چهار پهنه با در نظر گرفتن پهنه‌بندی SOM، ویژگی‌های مدل برازش شده واریوگرام متغیر شاخص کیفی سنگ استخراج شد که مدل کروی برای پهنه‌های اول و سوم و مدل نمایی برای پهنه دوم و چهارم انتخاب شد و بیشترین شعاع تأثیر به ترتیب متعلق به زون سوم، دوم، اول و چهارم است (شکل ۱۵).

### ۸-۳. ناهمسانگردی

ناهمسانگردی به تغییرات دامنه تأثیر و یا سقف واریوگرام در جهات مختلف گفته می‌شود و به همین دلیل واریوگرام ابزار بسیار مفیدی برای تشخیص ناهمگنی‌هایی است که موجب بروز ناهمسانگردی می‌شود. به‌طور کلی، دو نوع ناهمسانگردی هندسی و منطقه‌ای وجود دارد. برای شناخت ناهمسانگردی در یک محیط خاص ابتدا واریوگرام‌ها در جهات مختلف رسم شد. سپس به دامنه‌های تأثیر در فضای سه‌بعدی، یک بیضوی (بیضوی ناهمسانگردی (Anisotropic ellipsoid)) برازش داده شد تا دامنه تأثیر را به عنوان تابعی از جهت نشان دهد. به این منظور، ابتدا لازم است موقعیت بیضوی در فضا مشخص شود. در این صورت باید هم از متغیرهای بعدی و هم جهتی استفاده شود (Myers & Journel, 1990). برای نمایش بیضوی، معین کردن دو زاویه و سه دامنه تأثیر، یا نسبت قطر بزرگ بیضی ناهمسانگردی به قطر کوچک آن و نسبت قطر بزرگ بیضی ناهمسانگردی به قطر متوسط آن ضروری است. در همین رابطه، ویژگی‌های بیضی ناهمسانگردی چهار پهنه برای دو حالت به‌دست آمد (جدول ۴).

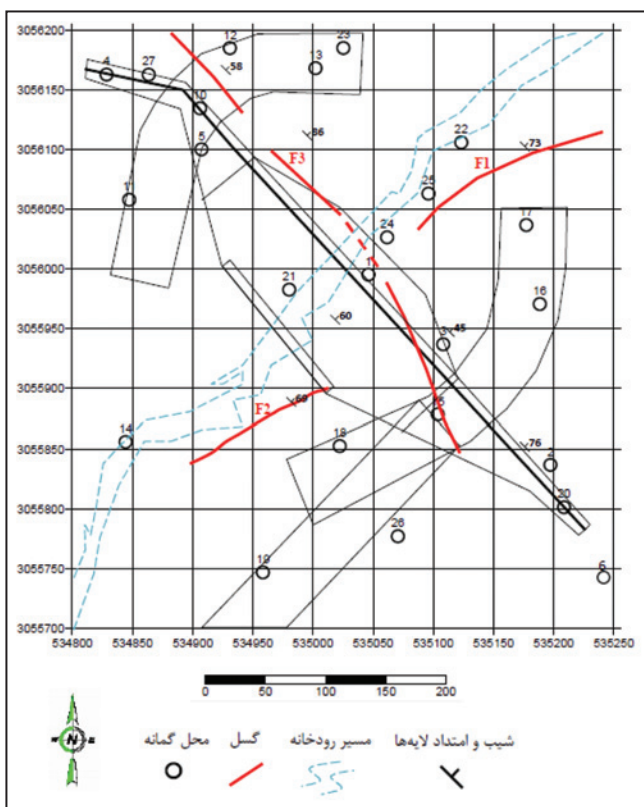
بنابراین، بیضی ناهمسانگردی را می‌توان با معلوم کردن زاویه قطر بزرگ آن نسبت به شمال، طول قطر بزرگ، و ضریب ناهمسانگردی هم مشخص کرد. در حالت سه بعدی بیضوی ناهمسانگردی را با معلوم کردن طول سه محور آن و دو زاویه که یکی زاویه محور بزرگ آن نسبت به شمال و دیگری زاویه محور بزرگ آن نسبت به افق است، می‌توان رسم کرد. برای ساختگاه مورد مطالعه، برای هر یک از چهار پهنه، بیضوی ناهمسانگردی رسم شد (شکل ۱۶).

### ۹- بحث و تعیین شاخص کیفیت خوشه‌بندی در دو حالت

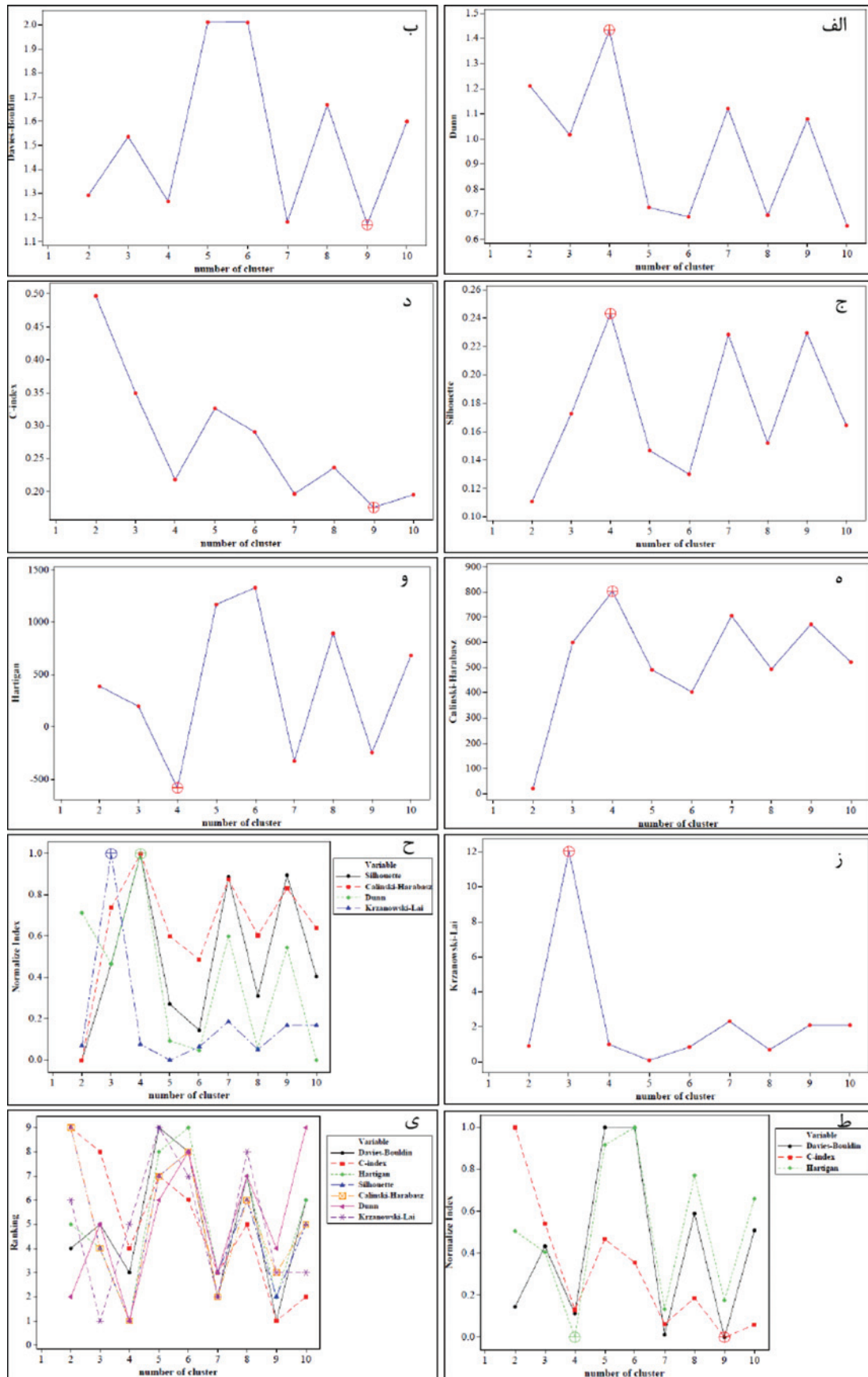
در این مطالعه، در حالت دوم از منطق خوشه‌بندی خود سازمانده برای پهنه‌بندی استفاده شده است. یکی از روش‌های شناخت خوشه‌بندی صحیح، این است که بیشترین تفاوت بین خوشه‌ای و بیشترین تشابه درون‌خوشه‌ای وجود داشته باشد. برای اعتبارسنجی دو حالت خوشه‌بندی باید یک شاخص مناسب تعریف کرد. برای تعریف شاخص کیفیت خوشه‌بندی، به میانگین و انحراف معیار مقادیر شاخص کیفی سنگ هر یک از خوشه‌ها در دو حالت مورد مطالعه نیاز است. در این صورت، شاخص مورد نظر با مجموع اختلاف میانگین بین هر دو خوشه (ملاک بیشترین تفاوت بین خوشه‌ای) ارتباط مستقیم و با مجموع انحراف معیار خوشه‌ها (ملاک کمترین تشابه درون‌خوشه‌ای) رابطه عکس دارد و مقادیر بیشینه این شاخص، معرف کیفیت بهتر خوشه‌بندی است. شاخص کیفیت خوشه‌بندی به صورت مجموع اختلاف میانگین بین هر دو خوشه تقسیم بر مجموع انحراف معیار خوشه‌ها تعریف می‌شود که  $\bar{X}$  میانگین شاخص کیفی سنگ در هر خوشه،  $\sigma$  انحراف معیار شاخص کیفی سنگ در هر خوشه و  $n$  تعداد خوشه‌هاست.

در مطالعه حاضر، میانگین و انحراف معیار مقادیر شاخص کیفی سنگ هر یک از خوشه‌ها محاسبه و شاخص کیفیت خوشه‌بندی برای هر دو حالت بررسی شد که مقدار بیشتر این شاخص در حالت پهنه‌بندی خودسازمانده، بیانگر کیفیت بهتر خوشه‌بندی این روش است (جدول ۵).

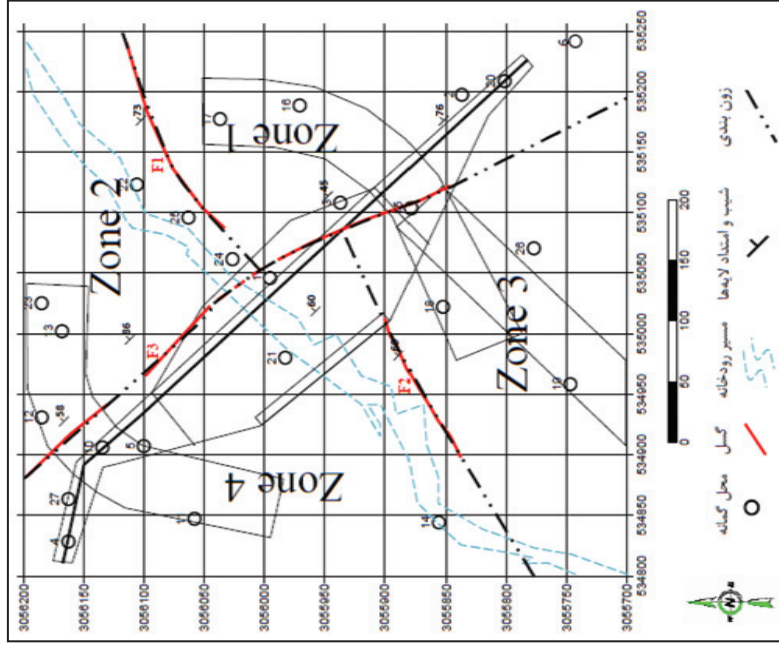
$$Index = \frac{\sum_{i=1}^n (\bar{X}_i - \bar{X})}{\sum_{i=1}^n \sigma_i} \quad (13)$$



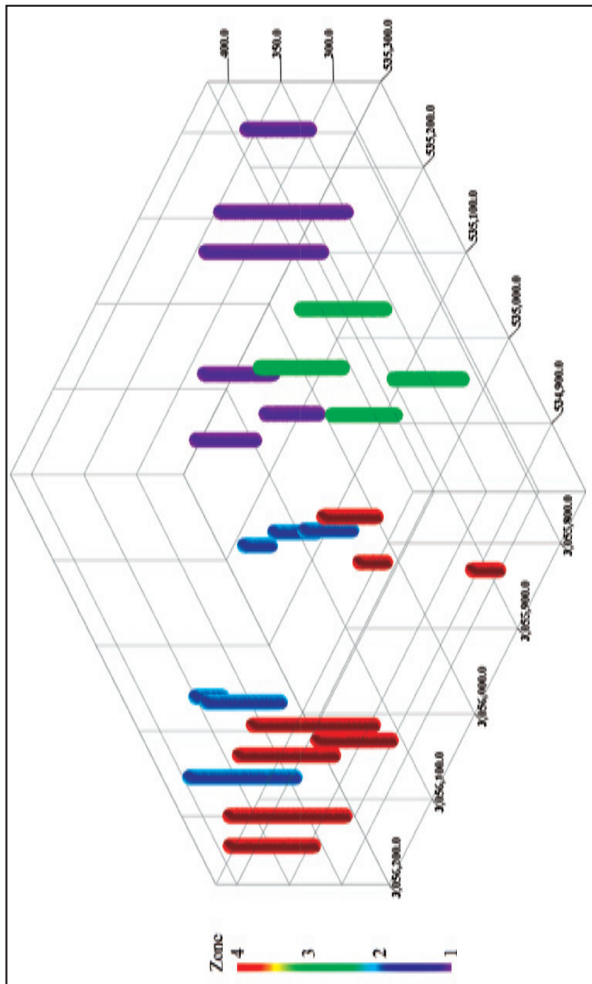
شکل ۱- نقشه خلاصه شده موقعیت سد سمیلان و شبکه گمانه‌های اکتشافی.



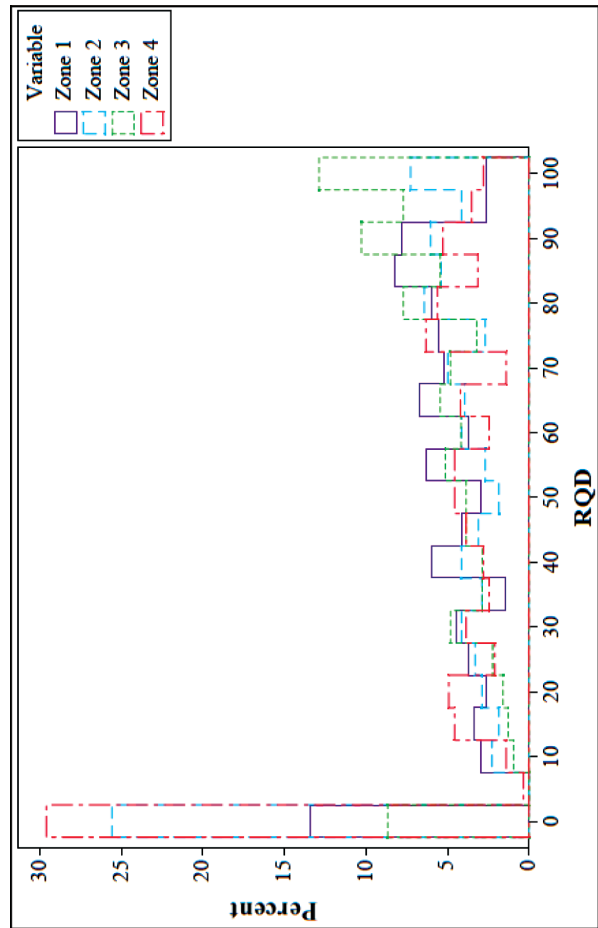
شکل ۲- تعداد بهینه خوشه‌ها براساس شاخص‌های اعتبارسنجی (الف دان، ب دیویس-بولدین، ج سیلهوت، د شاخص C، ه کالینسکی-هاراباز، و هارتگان، ز کرزنوسکی-لای، ح) مقادیر بهنجارسازی شده شاخص‌هایی که با تعداد خوشه‌ها رابطه مستقیم دارند، ط) مقادیر بهنجارسازی شده شاخص‌هایی که با تعداد خوشه‌ها رابطه غیرمستقیم دارند و ی) نمایش رتبه‌بندی تعداد خوشه‌های بهینه براساس شاخص‌های مختلف.



شکل ۳- موقعیت گسل‌های پهنه، محل گمانه‌ها و ساختارهای مختلف و نحوه پهنه‌بندی پهنه سد سمیلان.

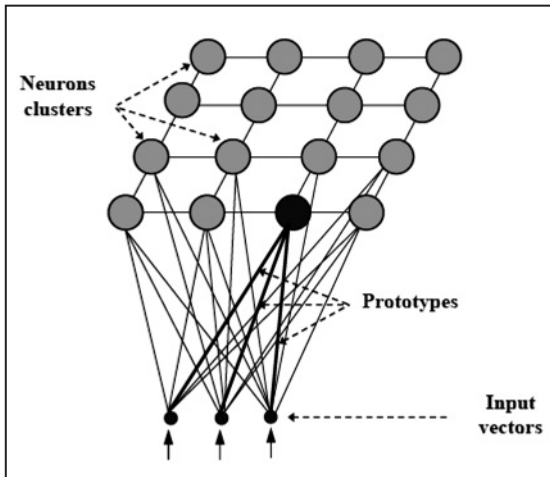


شکل ۴- موقعیت گمانه‌ها براساس پهنه‌بندی چهار گانه سد سمیلان براساس پهنه‌بندی گسل‌ها.

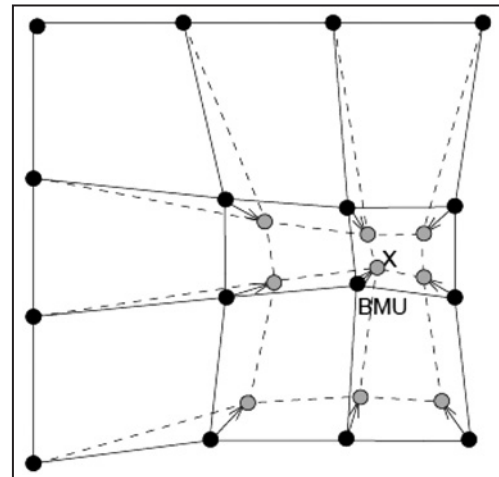


شکل ۵- نمایش توزیع آماری متغیر لوزون مناطق چهار گانه براساس پهنه‌بندی گسل‌ها.

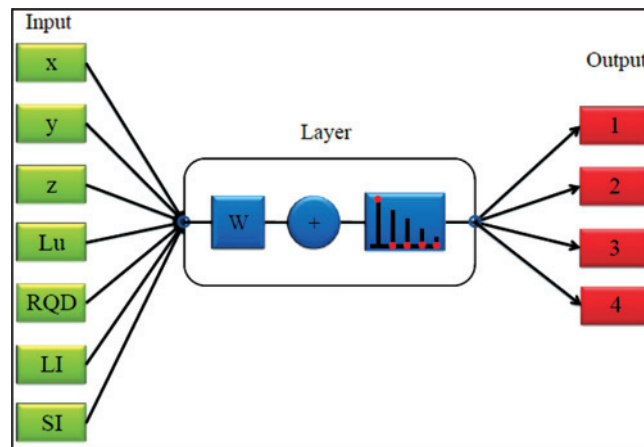




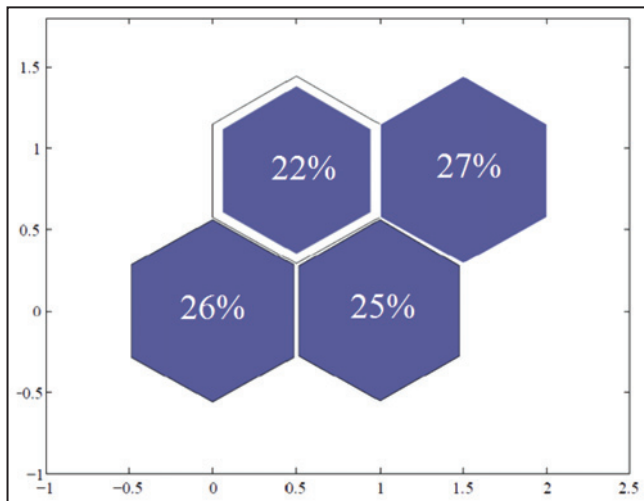
شکل ۷- نمایش نمادین نقشه خودسازمانده.



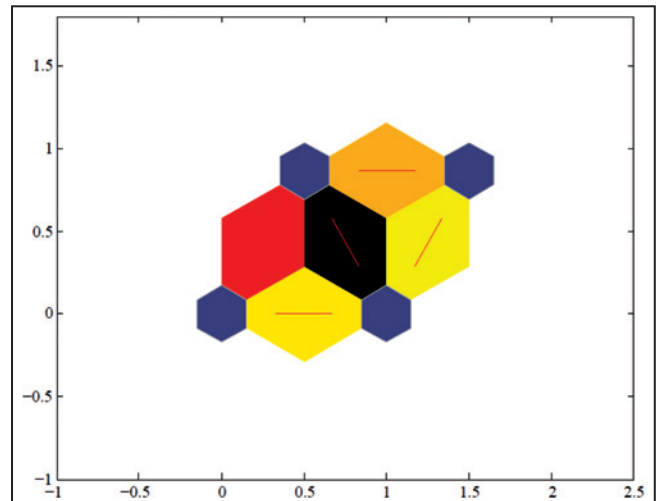
شکل ۶- نحوه انجام بهترین میزان تطبیق (Vesanto et al., 2000).



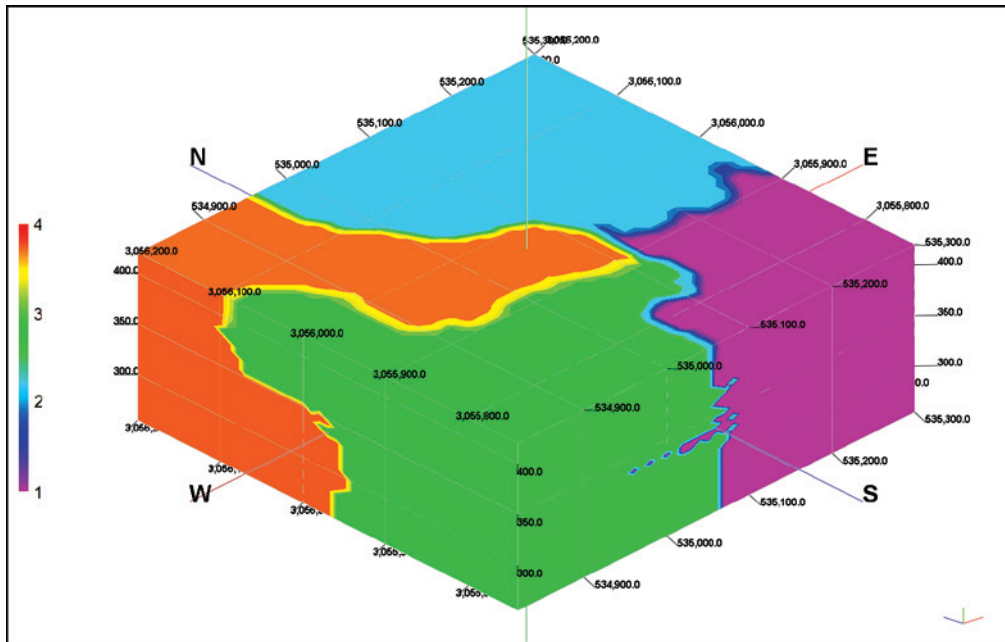
شکل ۸- نمایش ساختار شبکه عصبی خود سازمانده مورد مطالعه.



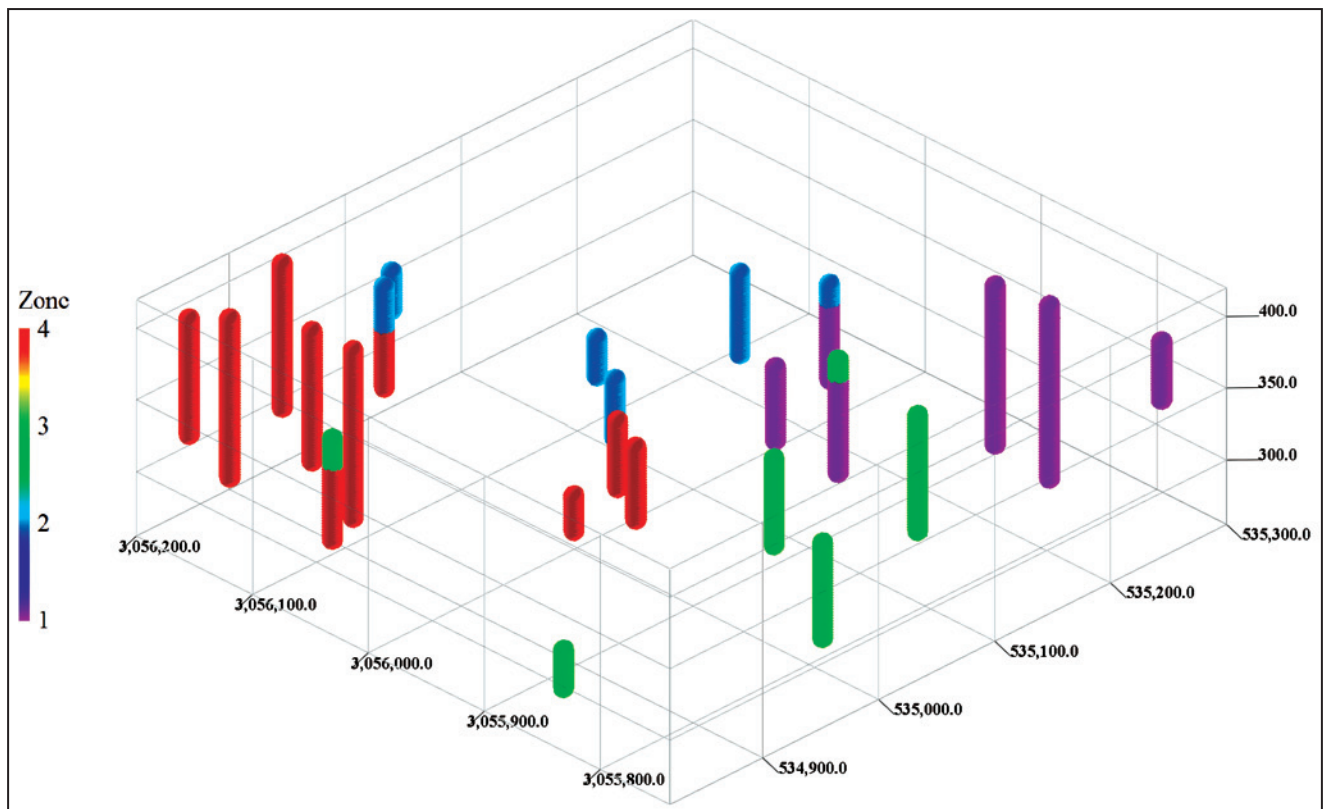
شکل ۱۰- تعیین تعداد نمونه‌های مربوط به هر خوشه.



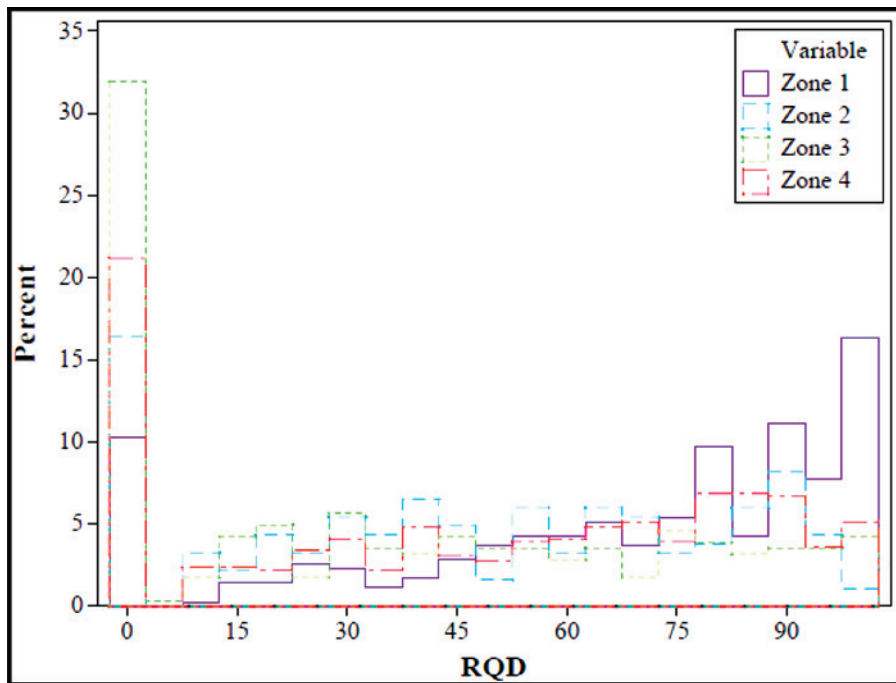
شکل ۹- نمایش نتایج ماتریس فاصله وزنی.



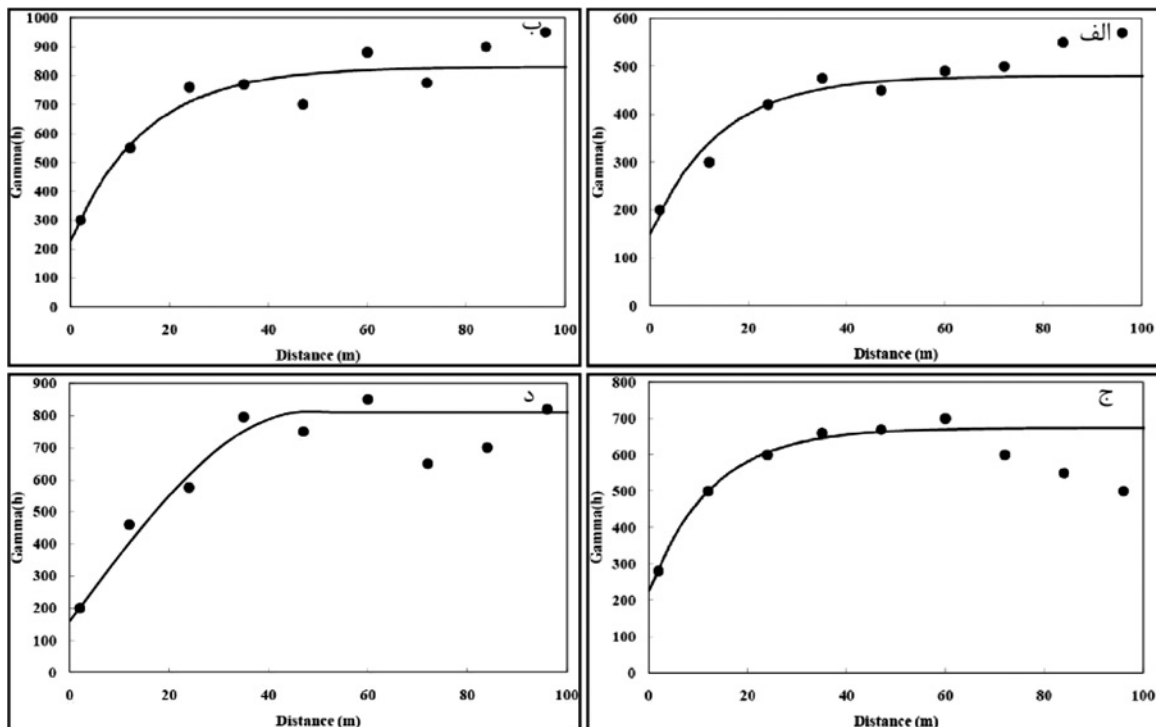
شکل ۱۱- نمایش پهنه‌بندی براساس SOM.



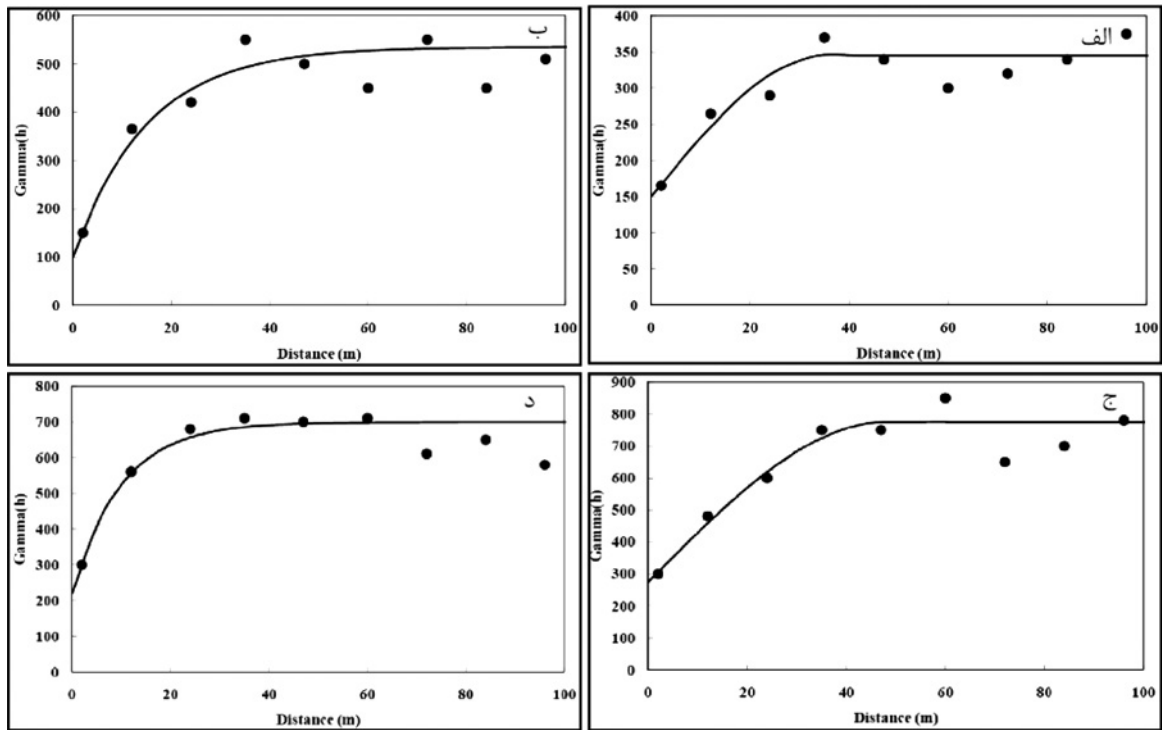
شکل ۱۲- اختصاص هر نمونه به یک پهنه براساس پهنه‌بندی SOM.



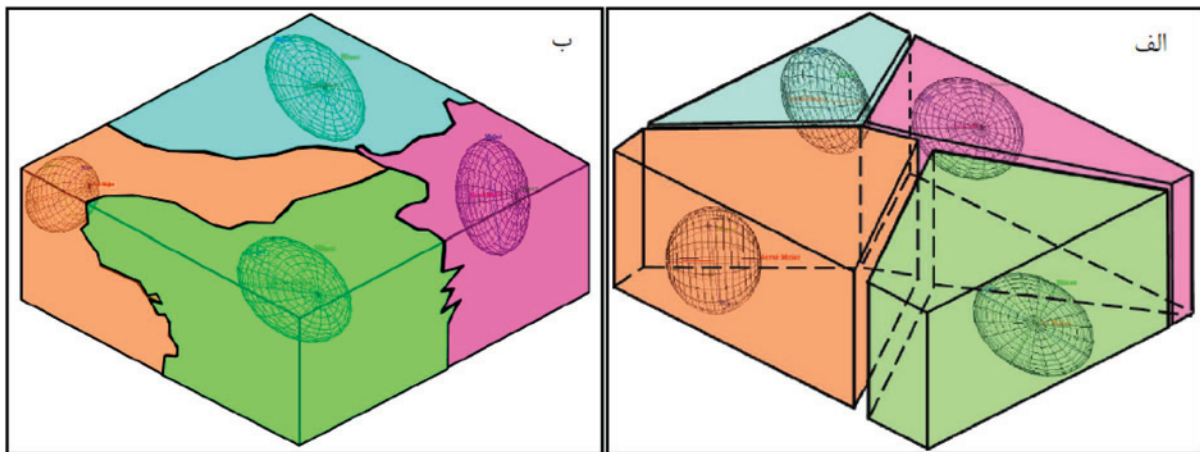
شکل ۱۳- نمایش توزیع آماری مقادیر شاخص کیفی سنگ برای چهار پهنه براساس پهنه‌بندی SOM.



شکل ۱۴- استخراج متغیرهای واریوگرام و واریوگرافی متغیر شاخص کیفی سنگ براساس پهنه‌بندی گسل‌ها الف) پهنه ۱، ب) پهنه ۲، ج) پهنه ۳ و د) پهنه ۴.



شکل ۱۵- استخراج متغیرهای واریوگرام و واریوگرافی متغیر شاخص کیفی سنگ براساس پهنه‌بندی SOM. الف) پهنه ۱، ب) پهنه ۲، ج) پهنه ۳ و د) پهنه ۴.



شکل ۱۶- نمایش بیضوی ناهمسانگردی برای هر پهنه. الف) براساس گسل‌ها، ب) براساس شبکه عصبی خودسازمانده.

جدول ۱- نحوه شاخص گذاری، بر اساس سازه‌ها و بخش‌های مختلف سد.

گمانه‌ها	شاخص	درجه اهمیت	نوع سازه
۱۳، ۴، ۵، ۱۰، ۱۵، ۲۰، ۲۱، ۲۲، ۲۷	۳	زیاد	بدنه سد
۱۱، ۱۲، ۱۳، ۱۶، ۱۷، ۱۸، ۱۹، ۲۳، ۲۴	۲	متوسط	تونل‌ها
۲۶، ۲۵، ۱۴، ۶، ۲	۱	کم	مناطق اطراف تونل‌ها و سد

جدول ۲- رتبه‌بندی بهینه‌سازی تعداد خوشه‌ها بر اساس شاخص‌های مختلف.

Silhouette	Davies-Bouldin	Calinski-Harabasz	Dunn	C-index	Krzanowski-Lai	Hartigan	شاخص	تعداد خوشه
۹	۴	۹	۲	۹	۶	۵		۲
۴	۵	۴	۵	۸	۱	۴		۳
۱	۳	۱	۱	۴	۵	۱		۴
۷	۹	۷	۶	۷	۹	۸		۵
۸	۸	۸	۸	۶	۷	۹		۶
۳	۲	۲	۳	۳	۲	۳		۷
۶	۷	۶	۷	۵	۸	۷		۸
۲	۱	۳	۴	۱	۳	۲		۹
۵	۶	۵	۹	۲	۳	۶		۱۰

جدول ۳- نمایش مرزبندی‌های مختلف و فضای برآورد در مناطق مختلف (Wingle, 1997).

نوع مرز بندی	نمایش نمادین مرزبندی سه گانه	همسایگی نمونه‌ها در برآورد بر اساس مرزبندی
ناگهانی		
تدریجی		
فازی		

جدول ۴- متغیرهای بیضوی ناهمسانگردی بر اساس پهنا واحد و پهنا بندی.

پلانچ	شیب	آزموت	نسبت محور اصلی به محور کوچک	نسبت محور اصلی به محور متوسط	پهنه مورد مطالعه	اساس پهنا بندی
۵۰	۰	۷۵	۱/۸۶	۱/۴۶	پهنه اول	گسل‌ها
۴۵	۱۰	۲۳۳	۱/۶۰۴	۱/۳۴	پهنه دوم	
۳۰	۵	۱۳۰	۱/۲۴۷	۱/۱۹۴	پهنه سوم	
۳۵	۸	۱۱۰	۱/۳۵	۱/۲۵	پهنه چهارم	
۸۰	۲/۵	۱۰۰	۱/۵۹	۱/۳۴	پهنه اول	سازمانده خود
۴۵	۰	۱۳۰	۱/۶۱	۱/۴۱	پهنه دوم	
۳۵	۵/۴	۱۱۰	۱/۵۵	۱/۳۶	پهنه سوم	
۳۰	۱۵	۲۶۲	۱/۲۴	۱/۰۸	پهنه چهارم	

جدول ۵- متغیرهای بیضوی ناهمسانگردی بر اساس پهنه واحد و پهنه‌بندی.

شاخص کیفیت خوشه‌بندی	انحراف معیار	میانگین	منطقه مورد مطالعه	اساس پهنه‌بندی
۰/۵۶۴	۳۱/۲۳	۵۱/۶۹	پهنه اول	گسل‌ها
	۳۶	۴۶/۶۲	پهنه دوم	
	۳۰/۹۹	۶۳/۳۴	پهنه سوم	
	۳۴/۳۸	۴۰/۱۱	پهنه چهارم	
۰/۶۶۸	۳۱/۷۴	۶۶/۲۵	پهنه اول	پهنه ساز مانده
	۳۱/۶۳	۴۷/۴	پهنه دوم	
	۳۴/۲۱	۳۷/۴۱	پهنه سوم	
	۳۴/۳۷	۴۹	پهنه چهارم	

## کتابنگاری

- حسین مرشدی، ا.، معماریان، ح.، ۱۳۸۷- پیشنهاد روشی برای تعیین محل حفاری‌های جدید، بر اساس کریجینگ شاخص. نشریه انجمن زمین‌شناسی مهندسی ایران، جلد اول، شماره سوم و چهارم، ص. ۴۵-۵۸.
- حسین مرشدی، ا.، معماریان، ح.، ۱۳۸۹- طراحی مرحله‌ای شبکه نمونه‌برداری، بر اساس متغیرهای ژئوتکنیکی و خصوصیات کیفی ساختگاه سد سمیلان، با استفاده از کریجینگ و شبکه عصبی. نشریه انجمن مهندسی معدن ایران، دوره پنجم، شماره دهم، ص. ۲۰-۱.
- مهندسین مشاور مهتاب قدس، ۱۳۸۷- زمین‌شناسی ساختگاه سد سمیلان.

## References

- Calinski, T. & Harabasz, J., 1974- A dendrite method for cluster analysis. *Communications in Statistics* 3, 1-27.
- Clayton, C. R., Matthews M. C. & Simons N. E., 1995- *Site investigation: a handbook for engineers*. Blackwell, Oxford.
- Davies, D. L. & Bouldin D. W., 1979- A cluster separation measure, *IEEE Trans. Pattern Anal. Machine Intell.* 1 (4). 224-227.
- Demuth, H., Beale, M. & Hagan, M., 2008- *Neural Network Toolbox (MATLAB)*, version 6, The MathWorks, Inc.
- Dunn J. C., 1974- Well separated clusters and optimal fuzzy partitions, *J. Cybern.* 4. 95-104.
- Hartigan, J. A., 1975- *Clustering Algorithms*. Wiley Series in Probability and Mathematical Statistics, John Wiley & Sons, Inc. New York.
- Hubert, L. & Schultz, J., 1976- Quadratic assignment as a general data-analysis strategy. *British Journal of Mathematical and Statistical Psychology* 29. 190-241.
- Krzanowski, W., Lai, Y., 1985- A criterion for determining the number of groups in a dataset using sum of squares clustering. *Biometrics* 44, 23-34.
- Myers, D. E. & Journel, A. G., 1990- Variograms and Zonal Anisotropies and Noninvertible Kriging Systems. *Mathematical Geology* 22(7), 779-785.
- Rousseeuw, P. J., 1987- Silhouettes: a graphical aid to the interpretation and validation of cluster analysis. *Journal of Computational and Applied Mathematics* 20, 53-65.
- Schatzmann, J., 2003- *Using Self-Organizing Maps to Visualize Clusters and Trends in Multidimensional Datasets*. Department of Computing Data Mining Group, Imperial College, London.
- Topchy, A. & Punch, J. W., 2003- Combining multiple weak clusterings *Proc. Third IEEE International Conference on Data Mining (ICDM'03)*, 331-338.
- Vann, J. & Geoval, D. G., 2003- Beyond Ordinary Kriging "An Overview of Non-linear Estimation". *Geostatistical Association of Australasia (GAA)*, 6-25.
- Vesanto, J. & Alhoniemi, E., 2000- Clustering of the Self-Organizing Map. *IEEE Transactions on Neural Networks* 11(3), 586-600.
- Vesanto, J., Himberg, J., Alhoniemi, E. & Parhankangas, J., 2000- *SOM Toolbox for Matlab 5*. Helsinki University of Technology.
- Wang, K., Wang, B. & Peng, L., 2009- CVAP: Validation for cluster analyses. *Data Science Journal* 8, 88-93.
- Webster, R. & Margaret, A., 2007- *Geostatistics for Environmental Scientists*, Wiley.
- Wingle, W. L., 1997- *Evaluating Subsurface Uncertainty Using Modified Geostatistical Techniques*. Degree of Doctor of Philosophy (Geological Engineer), Colorado School of Mines.
- Wingle, W. L. & Poeter, E. P., 1996- *Evaluating Subsurface Uncertainty Using Zonal Kriging Uncertainty*. '96 (ASCE), University of Wisconsin.