مهندسی عمران شریف، (تابستان ۱۳۹۵)
دوره‌ی ۲ - ۳۲، شماره‌ی ۲/۱، ص. ۲۹-۳۷

Original Article

الگویابی داده‌های لرزه‌یی با استفاده از خوشه‌بندی به منظور پیش‌بینی زلزله

عادل معطی (دانشجوی کارشناسی ارشد)
محمدرضا امین‌ناصری* (دانشیار)
بخش مهندسی صنایع، دانشگاه تربیت مدرس
حمید زعفرانی (استادیار)
پژوهشگاه بین‌المللی زلزله‌شناسی و مهندسی زلزله

زلزله‌ها همواره به‌عنوان یکی از مخرب‌ترین بلایای طبیعی شناخته می‌شوند. به دلیل خسارت‌های اقتصادی و تلفات جانی بسیار بالا، پیش‌بینی زلزله امری ضروری به نظر می‌رسد. در این نوشتار، تغییرات زمانی پارامتر b از رابطه‌ی گوتنبرگ ریشتر قبل از زلزله‌هایی با بزرگای $M_w = ۶/۰$ و یا بالاتر از آن در ناحیه‌ی جنوبی ایران، منطقه‌ی قشم و اطراف آن مورد بررسی قرار گرفته است. از دو روش خوشه‌بندی K-Means و نقشه‌ی خود سازمان‌ده SOM، برای یافتن الگو از این نوع زلزله‌ها استفاده شده است. براساس دو سنجه‌ی سیلوئیت و دیویس بولدین، تعداد ۳ خوشه به‌عنوان تعداد بهینه‌ی خوشه برای هر دو روش مذکور به‌دست آمده است. قبل از تمامی زلزله‌های مورد بررسی، خوشه‌یی که معرف کاهش در مقدار b است، مشاهده شده است. به‌عنوان نتیجه‌ی نهایی، کاهش مقدار b در بازه‌ی زمانی مشخص به‌عنوان یک الگوی مشخص برای رخداد این زلزله‌های مخرب معرفی شده است.

واژگان کلیدی: پیش‌بینی زلزله، داده‌کاوی، خوشه‌بندی، الگویابی، رابطه‌ی گوتنبرگ ریشتر.

adel.moatti@modares.ac.ir
amin_nas@modares.ac.ir
h.zafarani@iiees.ac.ir

۱. مقدمه

پیش‌بینی زلزله به‌عنوان یک راهکار مؤثر در کاهش آثار این رخداد طبیعی از ۷۰ سال پیش توسط ایدیا و ایشیمیتو به کار گرفته شده است.[۱] مطالعات در این حوزه به دو بخش پیش‌بینی کوتاه‌مدت و پیش‌بینی بلندمدت تقسیم شده است. پیش‌بینی کوتاه‌مدت بر پایه‌ی مطالعات بر روی پیش‌نشانگرها از قبیل: پیش‌لرزه‌ها، سکوت لرزه‌یی، کاهش در غلظت گاز رادون، و دیگر تغییرات شیمیایی و ژئوفیزیکی انجام شده است.[۱و۲] در حوزه‌ی پیش‌بینی بلندمدت، اطلاعات تاریخی از رخداد زلزله‌ها، همان داده‌های لرزه‌یی برای یافتن الگویی از رخداد زلزله‌های بعدی در بازه‌یی بلندمدت مورد استفاده قرار گرفته‌اند. در واقع، در بسیاری از مناطق لرزه‌خیز دنیا، زمان و مکان رخداد زلزله‌ها و همچنین بزرگای زلزله‌ها تا حدودی از روندی مشخص پیروی می‌کنند. بنابراین استخراج الگوهای لرزه‌یی از پارامترهای زمین‌لرزه همانند: زمان، مکان، و بزرگای رخداد زلزله می‌تواند در پیش‌بینی بلندمدت زلزله مناسب باشد.[۳-۵] یکی از مهم‌ترین روابط که در بسیاری از مطالعات پیش‌بینی بلندمدت مورد استفاده قرار گرفته است، قانون تجربی گوتنبرگ ریشتر است. این قانون با معادله‌ی ۱ ارائه می‌شود و بیان می‌کند بزرگای زلزله‌ها با لگاریتم تعداد تکرارشوندگی آنها رابطه‌ی معکوس دارد:

$$\log(N(M)) = a - bM \qquad (۱)$$

که در آن، M بزرگای زلزله بر حسب مقیاس مشخص، a و b مقدار ثابت، و $N(M)$ تعداد زلزله‌هایی با بزرگای کمینه‌ی M است.[۶] تغییرات زمانی و مکانی مقدار b، به‌دلیل ارائه‌ی مشخصه‌های ژئوفیزیکی و خواص تکتونیکی از منطقه‌ی مورد بررسی می‌تواند به‌عنوان یک شاخص مناسب جهت پیش‌بینی زلزله‌های بزرگ باشد. این شاخص در مورد منطقه‌ی جغرافیایی بزرگ و یا بازه‌ی زمانی طولانی تقریباً برابر ۱ است، اما در مورد مناطق با وسعت کم و یا بازه‌ی زمانی کوتاه، بین مقادیری از ۰/۵ تا ۱/۵ تغییر می‌کند.[۷] مطالعات بر روی تغییرات زمانی و مکانی مقدار b، از سال ۱۹۶۸ توسط موگی و شولتز آغاز شده است.[۸و۹] در مطالعه‌یی در سال ۲۰۱۲، پژوهشگران به بررسی تغییرات زمانی قبل از رخداد ۱۵ زلزله‌ی بزرگ در منطقه‌ی جنوب شرقی آسیا پرداخته و تمامی ۱۵ زلزله‌ی مذکور با کاهش معنادار در مقدار b همراه بوده‌اند.[۱۰] در مطالعه‌ی دیگری برخی پژوهشگران اقدام به بررسی تغییرات زمانی و مکانی مقدار b، قبل از ۲ زلزله با بزرگای $M_w = ۶/۰$ بین سال‌های ۲۰۰۸ تا ۲۰۱۱ در جنوب ایران کردند و این نتیجه را به‌دست آوردند که مقدار b قبل از این ۲ زلزله، هم از منظر زمان و هم از منظر مکان کاهش معناداری را تجربه

---

* نویسنده مسئول
تاریخ: دریافت ۱۳۹۲/۸/۲۱، اصلاحیه ۱۳۹۳/۷/۲۸، پذیرش ۱۳۹۳/۸/۱۰.

کرده است. در پژوهشی دیگر، با بررسی زلزله‌هایی در شعاع ۲۰۰ کیلومتری مرکز یک زلزله با بزرگای $M_w = ۶/۴$، نشان داده شده است که قبل از رخداد این زلزله، مقدار b به‌صورت محسوس از ۱/۲ تا ۰/۷ کاهش یافته است.[۱۱] در پژوهش دیگری، پژوهشگران با به‌کارگیری روش پنجره‌ی زمانی و مکانی، به بررسی تغییرات زمانی و مکانی مقدار b، قبل از ۲ زلزله با بزرگای $M_w = ۶/۰$ و $M_w = ۹/۰$، ۲ افت شدید در مقدار b در راستای زمان و مقدار پایین b در مکان این دو زلزله را گزارش داده‌اند.[۱۲] با توجه به مرور مطالعات صورت‌گرفته در راستای زلزله، پیداکردن روند تغییرات مقدار b، همچنین زمان کاهش مقدار پارامتر مذکور، بسیار مهم به نظر می‌رسد. در این نوشتار با استفاده از روش خوشه‌بندی، سعی در تعیین روند تغییرات زمانی b، همچنین تعیین زمان کاهش با روشی متفاوت با مطالعات قبلی شده است. در این پژوهش با استفاده از داده‌های لرزه‌یی مربوط به جنوب ایران و جزیره‌ی قشم بین سال‌های ۲۰۰۵ تا اوایل ۲۰۱۲، الگوی تغییرات زمانی مقدار b، قبل از زلزله‌هایی با بزرگای $M_w = ۶/۰$ و یا بالاتر از آن براساس روش به‌کارگرفته‌شده در پژوهش مورالس ـ استبان و همکاران ارائه شده است.[۱۳]

در قسمت دوم این نوشتار، داده‌های لرزه‌یی (کاتالوگ لرزه‌یی) برای این منطقه‌ی جغرافیای معرفی شده است. در قسمت سوم، روش به‌کارگرفته‌شده در راستای برآورد مقدار b، همچنین روش خوشه‌بندی و تعیین تعداد خوشه‌های بهینه به منظور کشف الگو توضیح داده شده است. در آخر، با تحلیل نتایج به‌دست‌آمده از خوشه‌بندی، الگوهای به‌دست‌آمده ارائه شده‌اند.

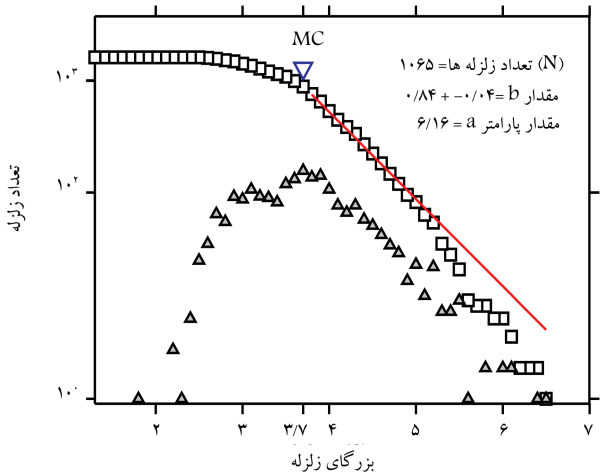
۲. خواص تکتونیکی

رشته کوه‌های زاگرس در جنوب کشور ایران، به‌عنوان بلندترین رشته کوه این کشور شناخته می‌شود. این رشته کوه از غرب (زاگرس مرکزی) تا جنوب شرقی ایران کشیده شده است و حاصل یک سری فعالیت تکتونیکی از برخورد دو صفحه‌ی تکتونیکی اوراسیا و عربستان است.[۱۴] تنش‌های فشاری که حاصل بازشدگی دریای سرخ و حرکت صفحه‌ی آفریقا ـ عربستان در راستای شمال و شمال شرقی، همچنین حرکت صفحه‌ی هند در راستای شمال و شمال غربی، موجب جابه‌جایی‌های متفاوت در پوسته‌ها و قطعات گوناگون قاره‌یی و اقیانوسی ایران شده است. از سوی دیگر، به دلیل حضور گسل‌های فعال جنبا و کواترنر، رشته کوه‌های زاگرس به عنوان یکی از لرزه‌خیزترین منطقه‌ی ایران شناخته می‌شود. مطالعه‌ی تاریخی زلزله‌های رخداده در این منطقه، خطرپذیری بالای این منطقه را از منظر رخداد زمین‌لرزه‌های مخرب تأیید می‌کند.

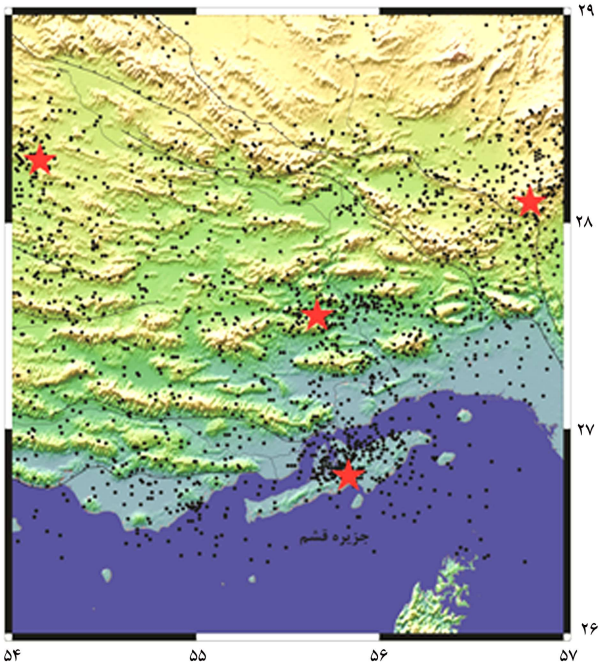
۳. کاتالوگ لرزه‌یی

در این مطالعه، مجموعه داده‌های لرزه‌یی از مرکز لرزه‌شناسی بین‌المللی (ISC) استخراج شده‌اند. منطقه‌ی جغرافیای محدودشده به طول جغرافیایی ۵۴° تا ۵۷/۵° درجه‌ی شرقی و عرض جغرافیایی ۲۶/۵° تا ۳۰° در نظرگرفته شده است. از ابتدای سال ۲۰۰۵ تا تاریخ ۲۰۱۲/۰۶/۱۹، به تعداد ۲۰۴۶ رکورد از زلزله‌ها اتفاق‌افتاده در این منطقه و ثبت‌شده در پایگاه ISC، در نظر گرفته شده‌اند. کمینه‌ی بزرگای زلزله‌ها برابر با $M_w = ۱/۸$ و بیشینه‌ی $M_w = ۶/۱$ ثبت شده است. کمترین بزرگای قابل اندازه‌گیری در یک منطقه به‌عنوان بزرگای آستانه و یا به اختصار M_c معرفی و در تمامی مطالعات لرزه‌خیزی، به‌عنوان یک پارامتر مهم شناخته می‌شود.

هر چه مقدار بزرگای آستانه (M_c) کمتر باشد، تعداد داده‌های لرزه‌یی در مطالعه بیشتر و نتایج استوارتر خواهد بود. در این نوشتار، برای برآورد M_c از روش برآورد بیشینه‌ی درست‌نمایی با بازه‌ی اطمینان ۹۰٪ استفاده شده است. براساس روش مذکور، $M_c = ۳/۷$ برآورد و تمامی زلزله‌ها با بزرگای کمتر از $M_w = ۳/۷$، از کاتالوگ لرزه‌یی حذف شده‌اند. در نهایت از ۲۰۴۶ زلزله‌ی ثبت‌شده، تعداد ۱۰۶۵ زلزله باقی مانده است. شکل ۱، توزیع فراوانی ـ بزرگای داده‌های لرزه‌یی رخ داده‌شده در بازه‌ی زمانی موردنظر را نشان می‌دهد. مقادیر ثابت a و b در رابطه‌ی گوتنبرگ ریشتر به ترتیب ۰/۸۹ و ۶/۶۱ برآورد شده‌اند. شکل ۲، زلزله‌های رخ داده‌ی ناحیه‌ی جنوب ایران، که شمال جزیره‌ی قشم نیز است، در تاریخ مورد بررسی را نشان می‌دهد. در شکل ذکرشده، زلزله‌های با بزرگای بالاتر از ۶/۰ در مقیاس گشتاوری



شکل ۱. توزیع فراوانی ـ بزرگا زلزله‌های رخداده‌شده در بازه‌ی زمانی مشخص.



شکل ۲. توزیع زلزله‌های رخداده در بازه‌ی زمانی سال‌های ۱۹۹۵ الی ۲۰۱۲/۰۶/۱۹ نقاط به شکل ستاره، نشان‌دهنده‌ی زلزله‌ها با بزرگای بالاتر از $M_w = ۶/۰$ است.

که مورد بررسی قرار گرفته‌اند، با ستاره نشان داده شده‌اند. این تذکر لازم است که به دلیل اینکه پیش‌لرزه‌ها و پس‌لرزه‌ها همگی در رخداد زلزله‌های بزرگ تأثیر دارند، از کاتالوگ لرزه‌یی مورد بررسی حذف نشده‌اند.[۱۴]

۴. روش‌شناسی

۱.۴. برآورد زمانی مقدار b

در بسیاری از مطالعات پیشین، محاسبه‌ی زمانی مقدار b از رابطه‌ی گوتنبرگ ریشتر و با استفاده از روش پنجره‌ی زمانی صورت گرفته است.[۴،۹و۱۵] پنجره با طول ثابت و پنجره با تعداد زلزله‌های ثابت به‌عنوان ۲ راهکار در این روش شناخته شده است. به دلیل تفاوت بالا در تعداد زلزله‌ها در پنجره‌های مختلف در روش پنجره‌ی زمانی با طول زمانی ثابت، در این نوشتار از پنجره‌هایی با تعداد زلزله‌های ثابت استفاده شده است.[۹] در هر پنجره مقدار b با استفاده از رابطه‌ی ۲، روش برآورد بیشینه‌ی درست‌نمایی ارائه‌شده توسط Aki (۱۹۶۵) برآورد شده است:[۱۶]

$$b = \frac{\log e}{M_{\text{Mean}} - M_{Min}} \qquad (۲)$$

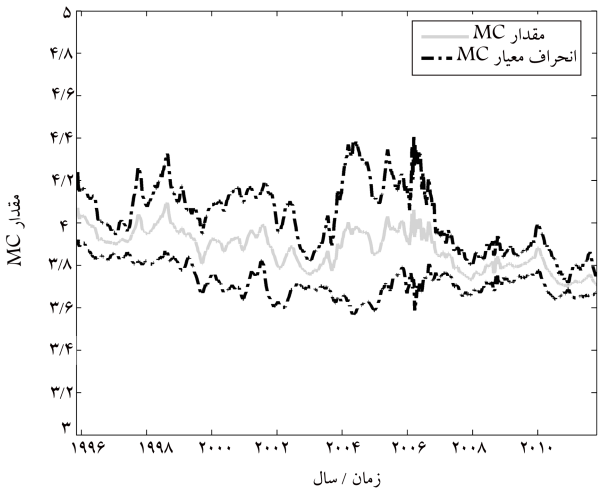
که در آن، M_{Mean} میانگین بزرگای زلزله‌های رخ داده در هر پنجره‌ی زمانی، و M_{Min} نیز کمترین بزرگای زلزله‌ی ثبت‌شده در هر پنجره‌ی زمانی است، که از رابطه‌ی ۳ به‌دست می‌آید، که در آن، M_c بزرگای آستانه‌یی است که در هر پنجره به‌صورت جداگانه محاسبه می‌شود و ΔM، به‌عنوان اختلاف بزرگا با مقدار ۰٫۱ در نظر گرفته شده است:

$$M_{Min} = M_c - \Delta M/۲ \qquad (۳)$$

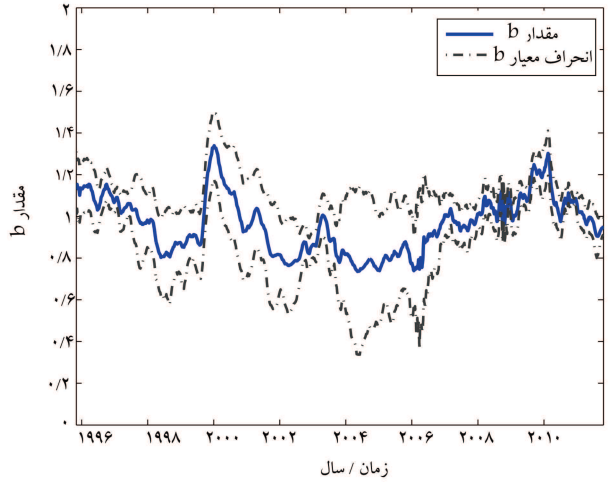
تعداد ۴۰، ۵۰، ۶۰، ۷۰، ۸۰ تا ۱۰۰ زلزله در هر پنجره‌ی زمانی برای رسیدن به بهترین و قابل رؤیت‌ترین برآورد از مقدار تغییرات b مورد بررسی قرار گرفته است. در نهایت، تعداد ۷۰ زلزله به‌عنوان تعداد هر زلزله در پنجره‌ی زمانی انتخاب شده است. با حرکت در راستای زمان به اندازه‌ی یک زلزله، مقادیر M_c و b با برآوردگر بیشینه‌ی درست‌نمایی از رابطه‌ی ۲ به‌دست آمده‌اند. نمودار ۱، نشان‌دهنده‌ی تغییرات زمانی مقدار M_c در بازه‌ی زمانی مورد مطالعه است. در نمودار ۱ مشاهده می‌شود که کمترین مقدار M_c، ۳٫۷ و بالاترین مقدار برابر با $M_w = ۴٫۲$ است. انحراف معیار نیز با استفاده از روش بوت استرپ به‌دست آمده و با خط‌چین مشکی نشان داده شده است. مقدار b در راستای زمان نیز با استفاده از پنجره‌ی زمانی در نمودار ۲ نشان داده شده است. می‌توان مشاهده کرد که مقدار b، از ۰٫۸ تا ۱٫۴ تغییر کرده است. مقادیر بالای b، نشان‌دهنده‌ی تنش پایین و پوسته‌ی ناهمگون در منطقه‌ی جغرافیایی مورد مطالعه و مقدار پایین b، نشان‌دهنده‌ی تنش بالای آن منطقه است.

۲.۴. مجموعه داده‌ی جدید

مقادیر برآوردشده‌ی b در راستای زمان، استخراج‌شده در بخش ۱.۴.، برای تشکیل مجموعه داده‌یی جدید به‌کارگرفته شده‌اند. هر زلزله درکاتالوگ لرزه‌یی با ۳ مشخصه‌ی b_i ضریب رابطه‌ی گوتنبرگ ریشتر، T_i زمان رخداد زلزله‌ی iاُم، و در نهایت M_i بزرگای هر زلزله در نظر گرفته شده‌اند. کاتالوگ لرزه‌یی در دسته‌های ۵تایی به ترتیب زمانی دسته‌بندی شده‌اند. با استفاده از روش به‌کارگرفته‌شده در مطالعه‌ی مورالس ـ استبان، مجموعه داده‌ی جدید با ۳ مشخصه معرفی می‌شود.[۱۳]



نمودار ۱. تغییرات زمانی مقدار M_c، بازه‌ی زمانی سال‌های ۱۹۹۵ الی ۲۰۱۲/۰۶/۱۹.



نمودار ۲. تغییرات زمانی مقدار b در راستای زمان در منطقه مورد بررسی در جنوب ایران.

مجموعه‌ی A_i، متشکل از تفاوت زمان رخداد زلزله‌ی اول و زلزله‌ی آخر در هر دسته‌ی ۵تایی زلزله‌ها، ΔT_i میانگین بزرگای زلزله‌ها، $\overline{M}_i$ و Δb_i اختلاف مقادیر b تخصیص داده‌شده به زلزله‌ی اول و زلزله‌ی ۵اُم در هر دسته‌ی ۵تایی تشکیل شده است (رابطه‌ی ۴):

$$A_i = (\overline{M}_i, \Delta b_i, \Delta T_i) \qquad i = ۱, \ldots, [N/۵] \qquad (۴)$$

مقادیر مجموعه‌ی A_i، از رابطه‌های ۵ الی ۷ به‌دست آمده‌اند:

$$\overline{M}_i = \sum_{k=j-۴}^{j} M_k/۵ \qquad j = ۵i \qquad (۵)$$

$$\Delta b_i = b_j - b_{j-۴} \qquad j = ۵i \qquad (۶)$$

$$\Delta T_i = T_j - T_{j-۴} \qquad j = ۵i \qquad (۷)$$

در روابط مذکور، N تعداد زلزله‌های کاتالوگ لرزه‌یی است. در نهایت مجموعه‌یی از تمامی A_iها به ترتیب زمانی تشکیل داده شده است. مجموعه‌ی ND، به فرم

رابطه‌ی ۸ تشکیل شده است:

$$ND = \{A_1, A_2, ..., A_{[N/5]}\} \quad (8)$$

در نهایت مجموعه‌ی ND، به‌عنوان مجموعه داده‌ی نهایی برای خوشه‌بندی و کشف الگو از زلزله‌هایی با بزرگای بالاتر از $M_w \geq 6/0$، در نظر گرفته شده است.

۳.۴. خوشه‌بندی

۱.۳.۴. خوشه‌بندی K-Means

دسته‌بندی عناصر براساس میزان تشابه آنها با توجه به یک تابع همسایگی در چند خوشه را خوشه‌بندی می‌نامند. خوشه‌بندی K-Means به‌عنوان یکی از کاربردی‌ترین خوشه‌بندها، که توسط بسیاری از پژوهشگران مورد استفاده قرار گرفته است، اولین بار توسط مک کوئین در سال ۱۹۶۸ ارائه شده است.[۱۷] در این روش، ابتدا K عنصر به‌صورت تصادفی به‌عنوان مراکز خوشه‌ها انتخاب می‌شوند. با استفاده از تابع همسایگی یا همان تابع فاصله بین ۲ عنصر، همانند توابع اقلیدسی یا منهتن، ماتریس فواصل استخراج و براساس نزدیکی و یا دوری فواصل عناصر از هر یک از این K عنصر، داده‌ها خوشه‌بندی می‌شوند. در این نوشتار از تابع فاصله اقلیدسی استفاده شده است. روند ذکرشده تا جایی ادامه می‌یابد که عناصر داخل یک خوشه، کمترین فاصله را با هم و بیشترین فاصله را با عناصر دیگر خوشه‌ها داشته باشند. با توجه به مسئله‌ی ذکرشده، برای کاهش فاصله‌ی عناصر داخل یک خوشه، معیار مربع خطا با توجه به رابطه‌ی ۹ تعریف می‌شود:

$$SSE = \sum_{i=1}^{k} \sum_{x_j \in C_j} |x_j - m_j|^2 \quad (9)$$

که در آن، k تعداد خوشه‌ها، x_j به‌عنوان jاُمین عنصر، m_j مرکز ثقل jاُمین خوشه و C_j نیز خوشه‌ی jاُم است. در این نوشتار، به‌منظور فرار از به دام‌افتادن در بهینه‌ی محلی، خوشه‌بندی ۵۰۰ بار تکرار شده است.

ضریب سیلوئیت: در روش‌های خوشه‌بندی بدون نظارت، همانند خوشه‌بندی K-Means، تعیین تعداد بهینه‌ی خوشه‌ها همواره به‌عنوان یک شاخص برای تعیین معتبربودن خوشه‌بندی ضروری است. در واقع به‌منظور اعتبارسنجی خوشه‌بندی و همچنین تعیین تعداد بهینه‌ی خوشه‌ها، سنجه‌های گوناگونی مورد استفاده قرار می‌گیرد. سنجه‌ی ضریب سایه‌روشن یا سیلوئیت به‌عنوان یک معیار برای اعتبارسنجی خوشه‌بندی در این نوشتار استفاده شده است. در این روش با محاسبه‌ی میانگین پهنه‌ی سایه‌روشن برای تمامی عناصر داخل یک مجموعه‌داده، تعداد بهینه‌ی خوشه‌ها مشخص می‌شود. رابطه‌ی ۱۰، پهنه‌ی سایه‌روشن را برای هر یک از عناصر محاسبه می‌کند:[۱۸]

$$S_i = \frac{b_i - a_i}{\max(a_i, b_j)} \quad (10)$$

که در آن، a_i میانگین فاصله‌ی عنصر iاُم از تمامی عناصری است که با این عنصر در یک خوشه قرار گرفته‌اند. b_i میانگین فاصله‌ی عنصر iاُم از تمامی عناصر در خوشه‌های دیگر و S_i پهنه‌ی سایه‌روشن عنصر iاُم است. میانگین تمامی پهنه‌های سایه‌روشن کل عناصر یک مجموعه داده را که در k خوشه دسته‌بندی شده‌اند، را با S_k نمایش می‌دهند. از تمامی مقادیر S_k، برای تعیین تعداد بهینه‌ی خوشه‌ها استفاده می‌شود. در این نوشتار، خوشه‌بندی برای تعداد ۲ تا ۸ خوشه تمامی مقادیر پهنه‌های سایه‌روشن محاسبه شده است. خوشه‌بندی با بالاترین مقدار S_k، خوشه‌بندی با مناسب‌ترین تعداد k است. شکل ۳، نمایش‌دهنده‌ی مقادیر به‌دست‌آمده به ازاء خوشه‌بندی با تعداد خوشه‌های متفاوت است. هر چه مقدار S_k به ۱ نزدیک‌تر باشد، اعتبار خوشه‌بندی بالاتر خواهد بود.

در این نوشتار، خوشه‌بندی با تعداد ۳ خوشه و ضریب سایه‌روشن ۰/۵۹۳۸، بهترین خوشه‌بندی معرفی شده است.[۱۹] براساس نتایج به‌دست‌آمده از ضرایب سیلوئیت، خوشه‌بندی ۵۰۰ بار بر روی مجموعه‌ی ND انجام شده است. جدول ۱، نشان‌دهنده‌ی مراکز خوشه‌های به‌دست‌آمده است.

در جدول ۱، خوشه‌ی ۱، بیانگر کاهش مقادیر مقدار b با وقوع زلزله‌ها با بزرگای پایین، خوشه‌ی ۲، بیانگر عدم تغییر در مقدار b و خوشه‌ی ۳ نیز با زلزله‌های با بزرگای بالا، افزایش مقدار b را نشان می‌دهد. بازه‌ی زمانی تغییرات مقدار b در ۳ خوشه نیز تقریباً ۱ ماهه است.

۲.۳.۴. الگوریتم نقشه‌ی خود سازمان‌ده

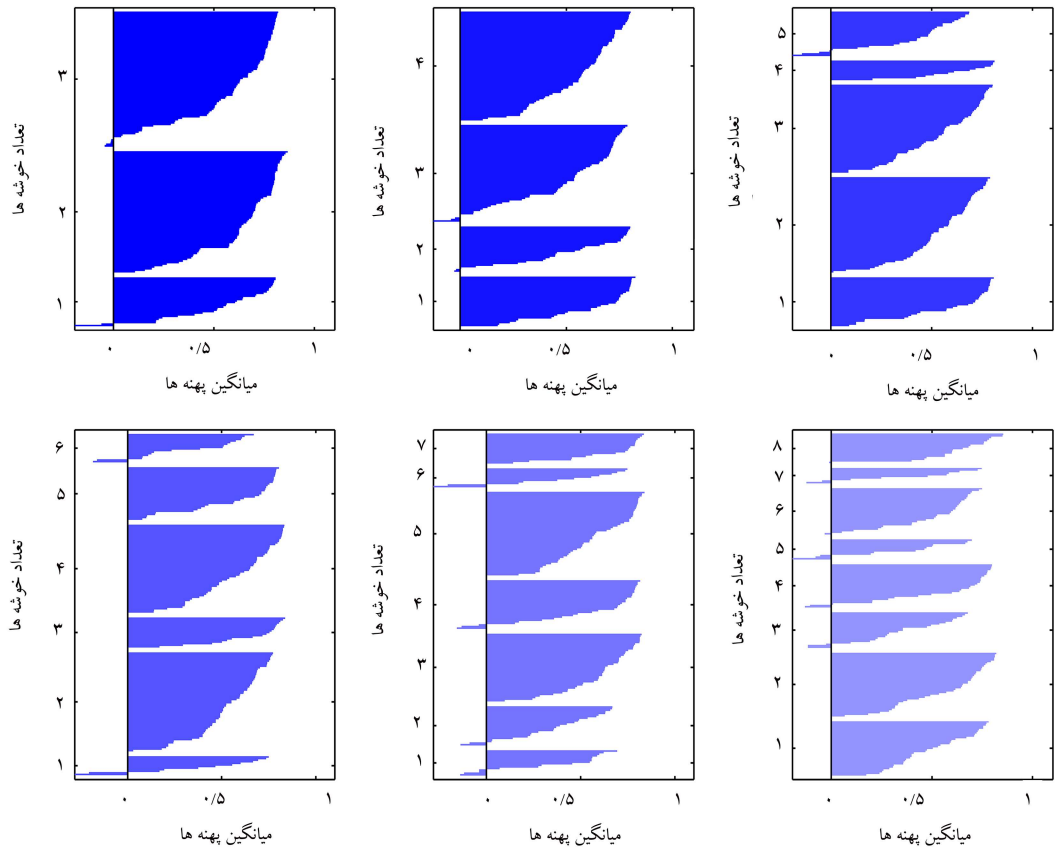
الگوریتم نقشه‌ی خود سازمان‌ده (SOM)، علاقه‌مندی زیادی را در میان پژوهشگران در زمینه‌های متنوعی ایجاد کرده است. الگوریتم SOM بسیار تجزیه و تحلیل شده و گونه‌های مختلفی از آن ایجاد شده است و اتفاقاً بیشتر به‌صورت قابل توجه در زمینه‌های علوم مهندسی تا پزشکی، بیولوژی، و اقتصاد به‌کار برده شده‌است. فقط ۲۰۹۲ نوشتار علمی در طی سال‌های ۱۹۹۸ الی ۲۰۰۲ ارائه شده است، که به نوعی از الگوریتم SOM استفاده کرده‌اند.[۲۰]

الگوریتم SOM بستری کارا برای تبدیل داده‌های با ابعاد زیاد به صورتی قابل تجسم است. هر چند برای درک کامل محتوای مجموعه داده‌ها، کشف ساختار خوشه‌یی داده‌ها ضروری است. اگر ساختاری در داده‌ها وجود دارد، خوشه‌ها برای استفاده‌ی کامل از ویژگی‌های مجموعه داده‌ها در کنار تولید اطلاعات خلاصه، باید استخراج شوند. همچنین، SOM به طور خلاصه در زمینه‌ی خوشه‌بندی مزایای آشکاری دارد. اولاً، مجموعه داده‌های اصلی با استفاده از مجموعه‌یی کوچک‌تر از بردارهای پیش‌نمونه نشان داده می‌شوند، که امکان استفاده‌ی مؤثر از الگوریتم‌های خوشه‌بندی برای تقسیم‌بندی پیش‌نمونه‌ها به خوشه‌ها را می‌دهد. کاهش میزان محاسبات به‌ویژه برای الگوریتم‌های سلسله مراتبی همراه با خوشه‌بندی‌هایی با اندازه و شکل دلخواه اهمیت دارد. ثانیاً، شبکه‌ی دو بُعدی امکان نمایش و تفسیر شکل بی‌نظم و ترتیب خوشه‌بندی‌ها را می‌دهد.[۲۱] شکل ۴، نیز روند نمایش داده‌ها در یک شبکه‌ی عصبی نقشه‌ی خود سازمان‌ده را نشان می‌دهد.

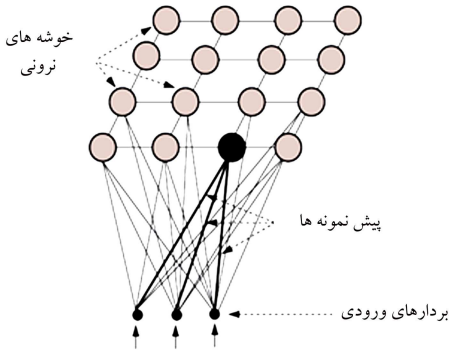
الگوریتم خوشه‌بندی SOM مجموعه داده‌ها را به تعدادی خوشه با کمینه‌سازی معیارهای نرخ یادگیری و فاصله‌ی همسایگی تقسیم می‌کند. تعداد خوشه‌ها معمولاً از پیش تعیین می‌شوند. الگوریتم با وقوع یکی از ۳ معیار پایان تعداد دوره‌های آموزشی، پایان مدت زمان آموزش، و رسیدن به کمینه‌ی خطا متوقف می‌شود. اگر تعداد خوشه‌ها نامعلوم باشد، الگوریتم SOM می‌تواند برای مجموعه‌یی از تعداد خوشه‌های مختلف، نوعاً از ۲ تا $\sqrt{N}$ تکرار شود، که N تعداد نمونه‌ها در مجموعه‌ی

جدول ۱. مرکز خوشه‌های حاصل از خوشه‌بندی K-Means.

شماره‌ی خوشه	$\overline{M}$	ΔT	Δb
۱	۳/۹۸۷	۰/۰۶۲	−۰/۰۲۳
۲	۴/۲۳۵	۰/۰۶۷	۰/۰۰۳
۳	۴/۵۳۱	۰/۰۶۲	+۰/۰۳۹



شکل ۳. ضرایب سیلوئیت برای تعداد ۳ تا ۸ خوشه به منظور تعیین تعداد بهینه‌ی خوشه‌ها.



شکل ۴. شبکه‌ی عصبی نقشه‌ی خود سازمان‌ده.

داده‌هاست. الگوریتم SOM مورد استفاده، تابع خطا (رابطه‌ی ۱۱) را کمینه می‌سازد:

$$E = \sum_{k=1}^{C} \sum_{x \in Q_k} ||x - c_k||^2 \qquad (11)$$

که در آن، C تعداد خوشه‌ها و c_k مرکز خوشه‌ی k است. الگوریتم SOM معمولاً متشکل از قاعده‌ی دوبُعدی (2-D) شبکه‌یی از واحدهای نقشه است. هر واحد i توسط یک بردار پیش‌نمونه $m_i = [m_{i1}, ..., m_{id}]$ نمایش داده می‌شود، که d بُعد بردار ورودی است. واحدها توسط یک رابطه‌ی همسایگی به مجاورت آن‌ها متصل می‌شوند. تعداد واحدهای نقشه که نوعاً از یک دوجین تا چندین هزار تغییر می‌کند، دقت و قابلیت تعمیم SOM را تعیین می‌کند. در طول آموزش، SOM یک شبکه‌ی ارتجاعی تشکیل می‌دهد که توده‌ی انبوه تشکیل‌شده از داده‌های ورودی را در هم می‌آمیزد. نقاط داده‌های واقع‌شده نزدیک یکدیگر در فضای ورودی، به‌صورت واحدهای نقشه نزدیک به هم ترسیم می‌شوند. بنابراین SOM می‌تواند همچون یک نقشه‌ی توپولوژی از فضای ورودی به شبکه‌ی دوبُعدی واحدهای نقشه تعبیر شود.

الگوریتم SOM به‌صورت تکراری آموزش داده می‌شود. در هر مرحله‌ی آموزش، یک بردار نمونه‌ی x به‌صورت تصادفی از مجموعه داده‌های ورودی انتخاب و فواصل بین x و همه‌ی بردارهای پیش‌نمونه محاسبه می‌شود. بهترین میزان تطبیق (BMU) که در اینجا توسط b مشخص می‌شود، واحد نقشه با نزدیک‌ترین پیش‌نمونه به x است (رابطه‌ی ۱۲):

$$||x - m_b|| = \min\{||x - m_i||\} \qquad (12)$$

سپس بردارهای پیش‌نمونه به‌روز می‌شوند. بردار BMU و همسایه‌های توپولوژیکی آن به نزدیکی بردار ورودی در فضای ورودی حرکت داده می‌شوند. قاعده‌ی به‌روزرسانی برای بردار وزنی واحد i به‌صورت رابطه‌ی ۱۳ است:

$$M_i(t+1) = m_i(t) + \propto(t) h_{b_i}(t)[x - m_i(t)] \qquad (13)$$

که در آن، t بیانگر زمان، $\propto(t)$ ضریب تطبیق، $h_{b_i}(t)$ هسته‌ی همسایگی متمرکزشده روی واحد برنده است، که از رابطه‌ی ۱۴ استخراج می‌شود:

$$h_{b_i}(t) = \exp\left(-\frac{||r_b - r_i||^2}{2\sigma^2(t)}\right) \qquad (14)$$

این تذکر لازم است که در رابطه‌ی ۱۴، r_b و r_i موقعیت‌های نرون‌های b و I بر روی شبکه‌ی SOM هستند. هر دو پارامتر $\sigma(t)$ و $(t) \propto$ به‌صورت یکنواخت با زمان کاهش می‌یابند.

مجموعه داده‌ی ND به‌عنوان داده‌ی ورودی همانند روند ذکرشده انتخاب و سپس بیشینه‌ی تعداد خوشه‌ها برای تقسیم‌بندی بردارها محاسبه می‌شود. بیشینه‌ی تعداد خوشه‌ها از جذر ۱۹۹ داده‌ی موجود در مجموعه داده‌ی ND حاصل می‌شود، که پس از محاسبه، مقدار ۱۴٫۱ به‌دست می‌آید. بنابراین خوشه‌بندی داده‌های لرزه‌یی باید برای ۱۴ خوشه امتحان شود. پیچیدگی محاسباتی برای خوشه‌بندی‌ها با تعداد مختلف، به‌صورت رابطه‌ی ۱۵ است:

$$\sum_{k=۲}^{۱۴} ۱۹۹k = ۱۹۹ * (۲ + ۳ + ... + ۱۴) = ۲۰۶۹۶ \qquad (۱۵)$$

به همین ترتیب تا تعداد ۲ خوشه، رابطه‌ی ۱۵ انجام می‌شود (رابطه‌ی ۱۶):

$$\sum_{k=۲}^{۲} ۱۹۹k = ۱۹۹ * (۲) = ۳۹۸ \qquad (۱۶)$$

همان‌طور که مشاهده می‌شود، با افزایش تعداد خوشه‌ها بر پیچیدگی مسئله افزوده می‌شود. بنابراین باید تعداد خوشه‌یی انتخاب کرد که علاوه‌بر تأمین اهداف کاهش میزان محاسبات و کاهش نویز، میزان پیچیدگی مسئله نیز کمینه شود. واضح است که تعداد خوشه‌های بهینه‌ی مسئله عملاً از بیشینه‌ی تعداد خوشه‌های به‌دست‌آمده کمتر است. در همین راستا تعداد خوشه‌ها از ۲ تا ۱۰ خوشه در نظر گرفته می‌شود. تقسیم‌بندی بردارهای ورودی می‌توانند توپولوژی‌های مختلفی داشته باشند. به‌منظور ارزیابی اولیه از توپولوژی موردنظر، ۲ خطای عددی و خطای توپولوژی مورد استفاده قرار گرفته‌اند. در خطای عددی‌سازی، میانگین فاصله‌ی هر بردار داده‌ی آموزش با بهترین واحد تطبیق‌دهنده‌ی خودش محاسبه می‌شود و خطای توپولوژی، نسبت به تمامی بردارهاست، که در اولین و دومین واحد تطبیق‌دهنده مجاور نیستند. در همین راستا نقشه‌ی توپولوژی با ۷۲ نرون به‌صورت ۴ × ۱۸ انتخاب شده است. میزان خطای عددی‌سازی در این نقشه برابر ۰٫۰۴۳ و میزان خطای توپولوژی برابر ۰٫۰۳۰ به‌دست آمده است. نهایتاً همانند خوشه‌بندی K-Means معیاری به منظور ارائه‌ی بهینه‌ترین خوشه‌بندی ارائه می‌شود.

معیار دیویس بولدین: این شاخص هم فاصله‌ی درون خوشه‌یی و هم فاصله‌ی بین خوشه‌یی را هنگام ارزیابی خوشه‌بندی حاصل‌شده مورد ارزیابی قرار می‌دهد. بر طبق شاخص اعتبارسنجی دیویس ـ بولدین، بهترین خوشه‌بندی، رابطه‌ی ۱۷ را کمینه می‌سازد:

$$\frac{۱}{C}\sum_{k=۱}^{c} \max_{l \neq k} \left\{ \frac{S_c(Q_k) + S_c(Q_l)}{d_{ce}(Q_k, Q_l)} \right\} \qquad (۱۷)$$

که در آن، S_c را برای فاصله‌ی درون خوشه‌یی و d_{ce} را برای فاصله‌ی بین خوشه‌یی استفاده می‌کند.

الگوریتم SOM در دو فاز ترتیب و تنظیم آموزش داده شده‌اند. آموزش دو فاز با تعداد دوره‌های آموزشی مختلف (۱۰۰۰، ۲۰۰۰، ۳۰۰۰، ۴۰۰۰ و ۵۰۰۰) آزمایش شده‌اند و نرخ‌های یادگیری اولیه برای فاز ترتیب ۰٫۹ و برای فاز تنظیم ۰٫۰۲ در نظر گرفته شده‌اند. نرخ یادگیری به صورت خطی در طول آموزش با میل به صفر کاهش پیدا می‌کند. تعداد مراحل در فاز ترتیب، ۱۰۰۰ مرحله طول می‌کشد و فاصله‌ی همسایگی در فاز تنظیم، در فاصله‌ی نزدیک ۱ باقی می‌ماند. برای به‌کارگیری بیشینه‌ی نرخ یادگیری در زمان آموزش از رابطه‌ی ۱۸ استفاده شده است:

$$\text{Maxlinlr} = ۰٫۹۹۹ * P' * P \qquad (۱۸)$$

همچنین تابع یادگیری نقشه‌ی خود سازمان‌ده، تغییروزنی (dw) را برای نرون تعیین‌شده از ورودی نرون P، فعال‌ساز $a۲$ و نرخ یادگیری lr محاسبه می‌کند (رابطه‌ی ۱۹):

$$dw = lr * a۲ * (P' - w) \qquad (۱۹)$$

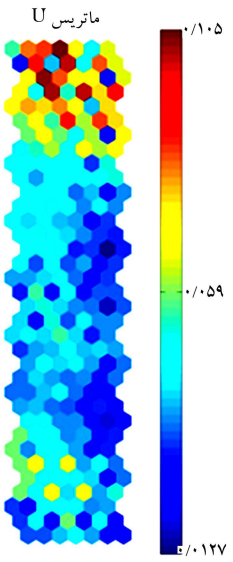
که فعال‌ساز $a۲$ از خروجی لایه‌ی $a۲$ و فواصل نرونی d از اندازه‌ی همسایگی فعلی nd یافته می‌شود (رابطه‌ی ۲۰):

$$\begin{cases} a۲(i,q) = ۱, & \text{if } a(i,q) = ۱ \\ a۲(i,q) = ۰٫۵, & \text{if } a(j,q) = ۱ \text{ and } D(i,j) \leq nd \\ a۲(i,q) = ۰ & \text{Otherwise} \end{cases} \qquad (۲۰)$$

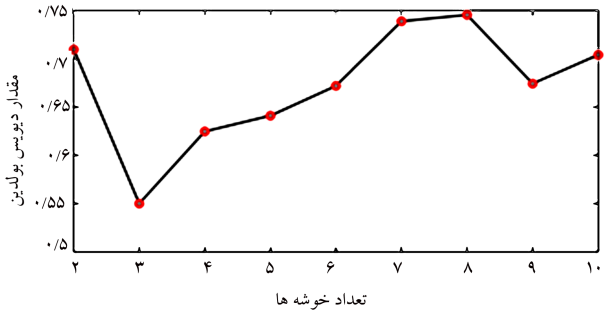
با شیوه‌ی ذکرشده، نرون‌ها یا خوشه‌ها مجموعه داده‌ها را به خوبی پوشش می‌دهند و اثر بیشتر برون‌هشته‌ها در مجموعه داده‌ها را حذف می‌کنند.

با توجه به روند توضیح داده‌شده، الگوریتم نقشه‌ی خود سازمان‌ده با پارامترها آموزش بالا بر روی مجموعه داده‌ی ND انجام شده است. در شکل ۵، ماتریس U حاصل از اجرای الگوریتم SOM بر روی مجموعه‌ی ND نشان داده شده است.

در شکل ۵، هر چه رنگ‌ها به سمت قرمز می‌روند، فاصله‌ی نرون‌ها از یکدیگر بیشتر می‌شود. به عبارت دیگر، می‌توان گفت مرز خوشه‌ها را می‌توان براساس رنگ‌بندی خوشه‌ها در نظر گرفت. همچنین هر چه خوشه‌ها به هم نزدیک‌تر و مشابهت بیشتری داشته باشند، رنگ‌بندی به سمت آبی می‌رود. همان‌طور که در شکل ۵ مشاهده می‌شود، توزیع مشخص و قابل تفکیکی از نقشه‌ی خود سازمان‌ده ارائه نشده است. به عبارت دیگر، به‌صورت چشمی نمی‌توان خوشه‌های ایجادشده در نقشه‌ی SOM را تولید کرد. با توجه به این امر با کمک خوشه‌بندی دیگر، نقشه‌ی تولیدشده دسته‌بندی می‌شود. در ادامه، ابتدا با استفاده از شاخص دیویس بولدین تعداد بهینه‌ی خوشه‌ها به‌دست آمده است، سپس با استفاده از خوشه‌بندی K-Means نقشه‌ی ایجادشده توسط SOM دوباره خوشه‌بندی می‌شود.



شکل ۵. ماتریس U براساس الگوریتم SOM.



شکل ۶. تعداد بهینه‌ی خوشه‌ها براساس شاخص دیویس ـ بولدین.

با توجه به رابطه‌ی ۱۷، تعداد ۳ خوشه به‌عنوان تعداد بهینه‌ی خوشه‌بندی مشخص می‌شود(شکل ۶). خوشه‌بندی K-Means با پیش‌فرض ۳ خوشه بر روی نقشه‌ی تولیدشده از SOM اعمال شده و خوشه‌ها براساس جدول ۲ به‌دست آمده‌اند.

در جدول ۲، خوشه‌ی ۱ بیانگر عدم تغییر در مقدار b، خوشه‌ی ۲ بیانگر افزایش در مقدار b در بازه‌ی زمانی تقریبی ۱ ماهه، و خوشه‌ی ۳، نمایانگر کاهش در مقدار b در بازه‌ی زمانی ۲۱ روزه است.

۵. الگویابی

براساس خوشه‌های به‌دست‌آمده از خوشه‌بندی K-Means و SOM، دسته‌های ۵تایی از زلزله‌ها در مجموعه داده‌ی تشکیل‌شده در قسمت ۲.۴. در راستای زمان براساس شماره‌ی خوشه‌یی که به آن تعلق دارند، رسم می‌شوند. در ابتدا الگوهای به‌دست‌آمده از K-Means ارائه شده است. همان‌گونه که در شکل ۲ مشاهده می‌شود، ۴ زلزله با بزرگای بیش از $M_w = ۶/۰$ بین سال‌های ۱۹۹۵ تا ۲۰۱۲/۰۶/۱۹ رخ داده است. جدول ۳، بزرگا و مکان جغرافیایی و زمان رخداد این ۴ زلزله را نشان می‌دهد.

۱.۵. الگویابی K-Means

با توجه به خوشه‌بندی مشاهده می‌شود که تمامی زلزله‌ها با بزرگای $M_w \geq ۶/۰$، در خوشه‌ی ۳ قرارگرفته‌اند. با افزایش تعداد پیش‌لرزه‌ها و پس‌لرزه‌ها، نرخ لرزه‌خیزی

جدول ۲. مرکز خوشه‌های حاصل از خوشه‌بندی SOM.

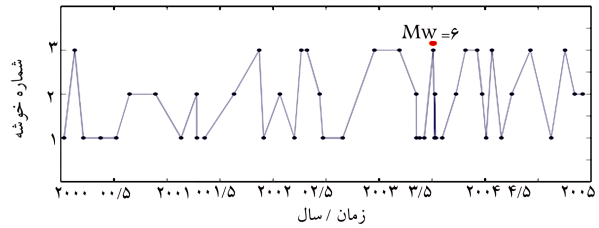
شماره‌ی خوشه	$\overline{M}$	ΔT	Δb
۱	۴/۲۱	۰/۰۷	۰
۲	۴/۴۷	۰/۰۹	۰/۰۲۸
۳	۴	۰/۰۶	−۰/۰۲۴

جدول ۳. مشخصات زلزله‌هایی با بزرگای بالاتر از $M_w = ۶/۰$ در کاتالوگ مورد بررسی.

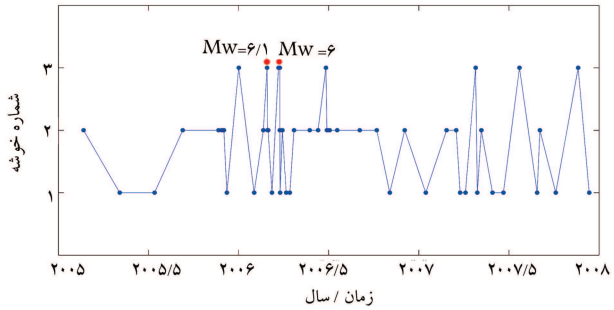
رخداد	عرض جغرافیایی	طول جغرافیایی	زمان رخداد	بزرگا
۱	۵۴/۱۵	۲۸/۳۰	۲۰۰۳/۰۷/۱۰	۶/۰
۲	۵۶/۸۲	۲۸/۱۰	۲۰۰۶/۰۲/۲۸	۶/۱
۳	۵۵/۶۶	۲۷/۵۵	۲۰۰۶/۰۳/۲۵	۶/۰
۴	۵۵/۸۳	۲۶/۷۷	۲۰۰۸/۰۹/۱۰	۶/۱

نیز افزایش محسوس داشته است. روند تغییرات خوشه‌ها تمامی دسته‌های ۵تایی از زلزله‌ها با برچسب خوشه‌یی که در آن قرار دارند، در شکل‌های ۷ الی ۹ نمایش داده شده‌اند. شکل ۷، روند تغییرات خوشه‌ها بین سال‌های ۲۰۰۰ الی ۲۰۰۵؛ شکل ۸، روند تغییرات خوشه‌ها بین سال‌های ۲۰۰۵ الی ۲۰۰۸؛ و شکل ۹، روند تغییرات از سال ۲۰۰۸ تا انتهای زمان مورد بررسی را نشان می‌دهند. دایره‌های آبی‌رنگ، نشان‌دهنده‌ی هر دسته‌ی ۵تایی از زلزله‌ها و نقاط قرمزرنگ نشان‌دهنده‌ی زلزله‌هایی با بزرگای بیش از $M_w = ۶/۰$ هستند. در بررسی زلزله‌ی رخ داده در سال ۲۰۰۳ که بزرگایی برابر با $M_w = ۶/۰$ دارد، همان‌طورکه روند تغییرات خوشه‌ها مشاهده می‌شود، قبل از رخداد این زلزله، دسته‌های ۵تایی از زلزله قرار دارند که در خوشه‌ی ۱ قرارگرفته‌اند. این به معنای افزایش تنش در پوسته و رخداد زلزله‌های با بزرگای پایین در منطقه‌ی مورد بررسی است. به عبارت دیگر، با کاهش مقدار b، احتمال رخداد زلزله‌های بزرگ افزایش می‌یابد. بازه‌ی زمانی رخداد خوشه‌های قبل از زلزله‌ی مورد بررسی نیز نشان می‌دهد که کاهش مقدار b، در بازه‌ی زمانی تقریباً ۳ ماهه رخ داده است.

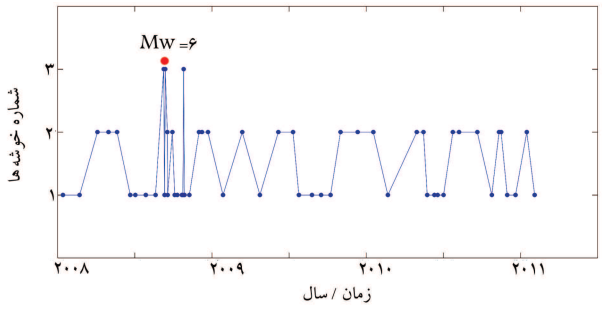
شکل ۸، تغییرات خوشه‌ها در بازه‌ی زمانی ۲۰۰۵ الی ۲۰۰۸ را نشان می‌دهد. در این بازه‌ی زمانی دو زلزله با بزرگای ۶/۱ و ۶/۰ در مقیاس گشتاوری در اواسط سال ۲۰۰۶ رخ داده‌اند. روند تغییرات مقدار b پیش از این دو زلزله، نشان‌دهنده‌ی کاهش مقدار b، قبل از آنهاست. به عبارت دیگر، با رخداد دسته‌های ۵تایی از زلزله‌هایی که



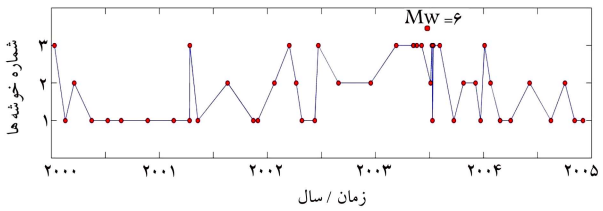
شکل ۷. روند تغییرات خوشه‌ها بین سال‌های ۲۰۰۰ الی ۲۰۰۵.



شکل ۸. روند تغییرات خوشه‌ها در بازه‌ی زمانی سال ۲۰۰۵ الی ۲۰۰۸.



شکل ۹. روند تغییرات خوشه‌ها در بازه‌ی زمانی ۲۰۰۸ الی ۲۰۱۲.



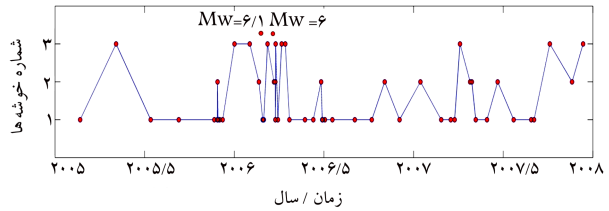
شکل ۱۰. روند تغییرات خوشه‌ها در بازه‌ی زمانی ۲۰۰۰ الی ۲۰۰۵.

در خوشه‌ی ۱ قرارگرفته‌اند و نمایانگر کاهش در مقدار b است، می‌توان روند ذکرشده را مشاهده کرد. از طرفی نیز روند رخداد خوشه‌ها از ۳ به ۲ و به ۱ معرف کاهش مقدار b، افزایش تنش و افزایش رخداد زلزله‌هایی با بزرگای بالاتر را دارد. تغییرات خوشه‌ها به منظور الگویابی در بازه‌ی زمانی سال ۲۰۰۸ تا انتهای کاتالوگ مورد بررسی نیز در شکل ۹، بررسی شده‌اند. زلزله‌یی با بزرگای $M_w = ۶/۱$ در سال ۲۰۰۸ در منطقه‌ی جنوبی ایران رخ داده است. روند تغییرات خوشه‌ی دسته‌های ۵تایی قبل از این زلزله در شکل ۹، نیز همانند ۳ زلزله‌ی بزرگ قبلی بیانگر کاهش در مقدار b است. رخداد دسته‌های ۵تایی که در خوشه‌ی ۱ قرارگرفته‌اند، نشان می‌دهد که در بازه‌ی زمانی ۳ ماهه کاهش مقدار b به‌صورت محسوس مشاهده می‌شود.

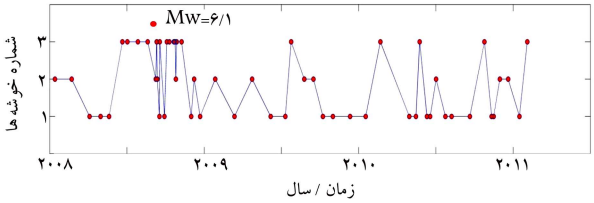
۲.۵. الگویابی نقشه‌ی خود سازمان‌ده

همانند روند ذکرشده در بخش پیشین، به منظور کشف الگو، دسته‌های ۵تایی از خوشه‌های به‌دست‌آمده از SOM نیز در راستای زمان نگاشت می‌شوند. برخلاف خوشه‌های به‌دست‌آمده از K-Means، این بار خوشه‌ی ۳، نماینده‌ی کاهش مقدار b است. خوشه‌ی ۱، نماینده‌ی عدم تغییر محسوس در مقدار b و خوشه‌ی ۲، بیانگر افزایش در مقدار b است. شکل ۱۰، روند تغییرات خوشه‌ها بین سال‌های ۲۰۰۰ الی ۲۰۰۵ را نشان می‌دهد. همان‌گونه که در شکل ۱۰ مشاهده می‌شود، روند تغییر خوشه‌ها قبل از رخداد زلزله‌ی موردنظر نیز رفتاری همانند الگوی خوشه‌بندی K-Means دارد. به عبارتی، روند تغییر خوشه‌ها قبل از رخداد زلزله‌ی موردنظر از ۲ به ۳ و ادامه پیداکردن همین روند تا رخداد زلزله، نمایانگر کاهش در مقدار b است. همان‌طورکه در جدول ۲ توضیح داده شده است، خوشه‌ی ۳ نشان از کاهش مقدار b دارد. به همین دلیل می‌توان روند کاهش مقدار b را همان‌طور که از خوشه‌بندی K-Means استخراج شده است، به‌عنوان الگوی به‌دست‌آمده از خوشه‌بندی SOM ارائه کرد.

شکل ۱۱، روند تغییرات خوشه‌ها را بین سال‌های ۲۰۰۵ الی ۲۰۰۸ نشان می‌دهد. همان‌طورکه قبلاً ذکر شده است، در این بازه ۲ زلزله با بزرگای ۶/۱ و ۶/۰ رخ داده است. براساس مشاهده‌ی روند تغییرات خوشه‌ها، رخداد اولین زلزله با بزرگای ۶/۱ با وقوع دسته‌های ۵تایی از زلزله همراه است، که در خوشه‌ی ۳ قرار دارند. درخصوص زلزله‌ی بعدی که با بزرگای ۶/۰ است، نیز می‌توان روند تغییر خوشه‌ها از



شکل ۱۱. روند تغییرات خوشه‌ها در بازه‌ی زمانی ۲۰۰۵ الی ۲۰۰۸.



شکل ۱۲. روند تغییرات خوشه‌ها در بازه‌ی زمانی ۲۰۰۸ الی ۲۰۱۲.

۲ به ۱ و سپس به ۳ را مشاهده کرد. بعد از این روند، که نشان‌دهنده‌ی کاهش مقدار b است، زلزله‌ی مذکور رخ داده است. بنابراین می‌توان الگوی کاهش مقدار b در بازه‌ی زمانی ۲ تا ۳ ماهه را به‌عنوان الگوی رخداد این دو زلزله معرفی کرد. شکل ۱۲، روند تغییرات خوشه‌ها بین سال‌های ۲۰۰۸ الی ۲۰۱۲ را نشان می‌دهد. همان‌طور که قبلاً ذکر شده است، در این بازه زلزله‌یی با بزرگای ۶/۱ در مقیاس گشتاوری در سال ۲۰۰۸ رخ داده است. براساس شکل مذکور، می‌توان رفتار دسته‌های ۵تایی از زلزله را قبل از رخداد زلزله‌یی با بزرگای ۶/۱ مشاهده کرد. همان‌گونه که در این شکل مشاهده می‌شود، روند تغییرات خوشه‌ها مجدداً حکایت از کاهش مقدار b در بازه‌ی تقریباً ۳ ماهه دارد. در شکل ۱۲، قبل از رخداد زلزله‌ی موردنظر دسته‌های ۵تایی از زلزله با برچسب ۳ مشاهده می‌شوند، که براساس جدول ۲ بیانگر کاهش مقدار b است.

۶. نتیجه‌گیری

در این نوشتار، از رابطه‌ی گوتنبرگ ـ ریشتر و مفاهیم داده‌کاوی برای الگویابی رخداد زلزله‌های با بزرگای بیش از ۶/۰ در مقیاس استفاده شده است. از دو روش خوشه‌بندی K-Means و نقشه‌ی خود سازمان‌ده به منظور کشف الگو از رخداد ۴ زلزله‌ی مخرب استفاده شده است. هر دو روش، ارائه‌یی تقریباً مشابه از روند رخداد زلزله‌های موردنظر داشته‌اند. روند به‌دست‌آمده از الگوریتم SOM با دقت بالاتری الگوهای رخداد را بیان می‌کند. این ادعا با توجه به نمودارهای ارائه‌شده بیان می‌شود، به نحوی که روند پیش‌بینی‌شده‌ی قبل از رخداد زلزله‌های مورد بررسی در مورد زلزله‌های دیگر با بزرگای پایین به‌دست نیامده است. در انتها می‌توان گفت که هر دو روش، بیانگر یک جواب به‌صورت واحد هستند و آن کاهش معنادار مقدار ضریب گوتنبرگ ریشتر قبل از رخداد زلزله‌هایی با بزرگای بیش از ۶/۰ در مقیاس گشتاوری است.

منابع (References)

1. Panakkat, A. and Adeli, H. "Recent efforts in earthquake prediction (1990-2007)", *Natural Hazards Review*, **9**(2), pp. 70-80 (2008).

2. Wyss, M. and Booth, D.C. "The IASPEI procedure for the evaluation of earthquake precursors", *Geophysical*

Journal International, **131**(3), pp. 423-424 (1997).

3. Wyss, M., Pacchiani, F., Deschamps, A. and Patau, G. "Mean magnitude variations of earthquakes as a function of depth: Different crustal stress distribution depending on tectonic setting", *Geophysical research letters*, **35**(1), p. L01307 (2008).
4. Chen, B., Bai, T. and Li, B. "The b-value and earthquake occurrence period", *Chinese Journal of Geophysics*, **46**(4), pp. 736-749 (2003).
5. Rani, V.S., Srivastava, K., Srivastava, K. and Dimri, V.P. "Spatial and temporal variations of b-value and fractal analysis for the Makran Region", *Marine Geodesy*, **34**(1), pp. 77-82 (2011).
6. Gutenberg, B. and Richter, C.F., *Seismicity of the Earth and Associated Phenomena*, Hafner, New York (1965).
7. Chan, C.H., Wu, Y.M., Tseng, T.L., Lin, T.L. and Chen, C.C. "Spatial and temporal evolution of b-values before large earthquakes in Taiwan", *Tectonophysics*, **532-535**, pp. 215-222 (2012).
8. Scholz, C. "The frequency-magnitude relation of microfracturing in rock and its relation to earthquakes", *Bulletin of the Seismological Society of America*, **58**(1), pp. 399-415 (1968).
9. Kulhanek, O. "Seminar on b-value", Dept. of Geophysics, Charles University, Prague (2005).
10. Nuannin, P., Kulhánek, O. and Persson, L. "Variations of b-values preceding large earthquakes in the Andaman-Sumatra subduction zone", *Journal of Asian Earth Sciences*, **61**, pp. 237-242 (2012).
11. Tsukakoshi, Y. and Shimazaki, K. "Decreased b-value prior to the M 6.2 Northern Miyagi, Japan, earthquake of 26 July 2003", *Earth Planets and Space (EPS)*, **60**(9), pp. 915-924 (2008).
12. Nuannin, P., Kulhánek, O. and Persson, L. "Spatial and temporal b value anomalies preceding the devastating off coast of NW Sumatra earthquake of December 26, 2004", *Geophysical Research Letters*, **32**(11), p. L11307 (2005).
13. Morales-Esteban, A., MartÍnez-Álvarez, F., Troncoso, A., Justo, J.L. and Rubio-Escudero, C. "Pattern recognition to forecast seismic time series", *Expert Systems with Applications*, **37**(12), pp. 8333-8342 (2010).
14. Sorbi, M.R. and et.al. "Seismicity patterns associated with the September 10th, 2008 Qeshm earthquake, South Iran", *International Journal of Earth Sciences*, **101**(8), pp. 2215-2223 (2012).
15. Bender, B. "Maximum likelihood estimation of b values for magnitude grouped data", *Bulletin of the Seismological Society of America*, **73**, pp. 831-851 (1983).
16. Aki, K. "Maximum likelihood estimate of b in the formula $\log(N) = a - bM$ and its confidence limits", *Bull. Earthq. Res. Inst. Tokyo Univ.*, **43**, pp. 237-239 (1965).
17. MacQueen, J. "Some methods for classification and analysis of multivariate observations", *in Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability*, p. 14 (1967).
18. Rousseeuw, P.J. "Silhouettes: A graphical aid to the interpretation and validation of cluster analysis", *Journal of Computational and Applied Mathematics*, **20**(1), pp. 53-65 (1987).
19. Ansari, Z., Azeem, M.F., Ahmed, W. and Babu, A.V. "Quantitative evaluation of performance and validity indices for clustering the web navigational sessions", *World of Computer Science and Information Technology Journal*, **1**(5), pp. 217-226 (2011).
20. Oja, M., Kaski, S. and Kohonen, T. "Bibliography of self-organizing map (SOM) papers: 1998-2001 addendum", *Neural Computing Surveys*, **3**, pp. 1-156 (2003).
21. Vesanto, J. and Alhoniemi, E. "Clustering of the self-organizing map", *IEEE Transactions Neural Networks*, **11**(3), pp. 586-600 (2000).