

کشف بی‌هنجاری‌های میدان ثقل مرتبط با وقوع زلزله‌های بزرگ در داده‌های ماهواره‌ای GRACE با استفاده از روش‌های تصمیم‌گیری جمعی

محسن شهریسوند^۱، مهدی آخوندزاده^{۲*} و یاسر جویباری مقدم^۱

۱. دانشجوی دکتری، گروه سنجش از دور، دانشکده مهندسی نقشه‌برداری و اطلاعات مکانی، پردیس دانشکده‌های فنی دانشگاه تهران، ایران

۲. استادیار، گروه سنجش از دور، دانشکده مهندسی نقشه‌برداری و اطلاعات مکانی، پردیس دانشکده‌های فنی دانشگاه تهران، ایران

(دریافت: ۹۴/۱۲/۲۶، پذیرش نهایی: ۹۵/۱۱/۵)

چکیده

هر چند سال هزاران نفر در سرتاسر جهان بر اثر وقوع زلزله جان خود را از دست می‌دهند. پژوهش‌های فراوانی درباره چگونگی وقوع زلزله انجام گرفته است که هدف نهایی همه آن‌ها به حداقل رساندن تأثیرات نامطلوب این پدیده است. در این مقاله نیز به منظور کشف ناهنجاری‌های تغییرات میدان ثقل قبل از وقوع زلزله‌های بزرگ، از داده‌های هفتگی پروژه ماهواره‌ای GRACE طی دوره هشت‌ساله استفاده شده است. در این مطالعه پس از تولید سری‌های زمانی تغییرات برخی از مؤلفه‌های تانسور مرتبه دوم پتانسیل ثقل، به منظور کشف رفتارهای غیرعادی میدان ثقل قبل از وقوع زلزله، از روش آماری میانه و دامنه بین چارکی، درخت تصمیم‌گیری و روش‌های تصمیم‌گیری جمعی بگینگ (Bagging)، بوستینگ (Boosting) و جنگل تصادفی (Random Forest) استفاده شد. مشتقات مرتبه دوم پتانسیل ثقل علاوه بر کاهش خطای نوارنوارشدگی موجود در داده‌های GRACE، مؤلفه‌های فرکانس بالای میدان ثقل را تقویت می‌کنند که موجب می‌شود تغییرات ناشی از پدیده‌های ژئوفیزیکی محلی مانند زلزله بهتر نشان داده شوند. مطابق با نتایج این پژوهش ۲ تا ۵ هفته قبل از وقوع زلزله، تغییرات ناگهانی در مؤلفه‌های ΔV_{xx} ، ΔV_{xz} از تانسور مرتبه دوم میدان ثقل دیده می‌شود.

واژه‌های کلیدی: بگینگ، بوستینگ، جنگل تصادفی، درخت تصمیم‌گیری، زلزله، میدان ثقل.

۱. مقدمه

تأثیرات فیزیکی و شیمیایی مختلفی در لیتوسفر، اتمسفر و یونسفر دارند و با عنوان پیش‌نشانگرهای زلزله شناخته می‌شوند. تاکنون ده‌ها پیش‌نشانگر مختلف زمین‌لرزه، شناسایی شده‌اند که با اندازه‌گیری آن‌ها با روش‌های مختلف می‌توان از آن‌ها به منظور هشدار و نجات مردم استفاده کرد. با وجود پیشرفت‌های صورت گرفته باید اذعان کرد که ماهیت پیش‌نشانگرهای زلزله به صورتی است که همه آن‌ها در تمامی زلزله‌ها وجود ندارند و نمی‌توان با تمرکز بر یک پیش‌نشانگر، درباره وقوع یا عدم وقوع زلزله در یک منطقه مشخص یا محدوده زمانی مشخص اظهار نظر کرد. اما هرچه تعداد پیش‌نشانگرها قبل از وقوع یک زلزله بیشتر باشد، می‌توان با اطمینان بیشتری پارامترهای آن زلزله را برآورد کرد.

تغییرات میدان ثقل ناشی از زلزله، نزدیک به دو دهه قبل با استفاده از روابط ریاضی بیان شد (سان و اکوبو، ۱۹۹۳). اولین مشاهدات تغییرات میدان ثقل مرتبط با

زلزله از مخرب‌ترین مخاطرات طبیعی است که بشر ناچار است همواره با آن روبه‌رو شود. تاکنون صدها هزار نفر در سرتاسر جهان به دلیل وقوع زلزله جان خود را از دست داده‌اند. بنابراین اگر اطلاعات جامعی درباره مکانیزم زلزله داشته باشیم، مسلماً به نحو بهتری می‌توانیم خودمان را در برابر تأثیرات مخرب این پدیده حفظ کنیم. پژوهش‌های فراوانی درباره چگونگی وقوع زلزله انجام گرفته که هدف نهایی همه آن‌ها رسیدن به جایگاهی است که بتوان تأثیرات ناشی از این پدیده را به حداقل رساند. در دهه‌های گذشته با پیشرفت تکنولوژی در علوم مختلف، محققان گام‌های بزرگی در زمینه مطالعات مربوط به تأثیرات فیزیکی و شیمیایی زلزله که به صورت پیش‌نشانگر بروز می‌کنند، برداشته‌اند.

در زمان شکل‌گیری زلزله فرآیندی دینامیک به صورت انتقال انرژی در اثر جابه‌جایی پوسته زمین بین منبع و محیط اتفاق می‌افتد. این تغییرات قبل از وقوع زلزله

وقوع زلزله توسط دستگاه‌های اندازه‌گیری شتاب ثقل زمینی بعد از زلزله Tokochi-Oki سال ۲۰۰۳ (Mw=8.0) در ژاپن مشاهده شد (ایمینیسی و همکاران، ۲۰۰۴). محدودیت تعداد ایستگاه‌های زمینی و عدم امکان اندازه‌گیری تغییرات میدان ثقل ناشی از زلزله، در مناطق پوشیده‌شده با آب از جمله مشکلات مربوط به مطالعه تغییرات میدان ثقل با استفاده از مشاهدات زمینی است. با ارسال پروژه‌های ماهواره‌ای ثقل سنجی فصل جدیدی در مطالعات مربوط به تغییرات زمانی میدان ثقل فراهم شد. پروژه ماهواره‌ای GRACE که با همکاری بین دو کشور آلمان و آمریکا در سال ۲۰۰۲ وارد مدار شد، از دو ماهواره مشابه و مجزا از هم تشکیل شده است که در ارتفاع ۵۰۰ کیلومتری از سطح زمین با فاصله ۲۲۰ کیلومتر از هم حرکت می‌کنند. با تغییر میدان ثقل محلی که ماهواره‌ها از آن عبور می‌کنند، فاصله بین آن‌ها دستخوش تغییر شده و با آنالیز این تغییرات فاصله، می‌توان به تغییرات میدان ثقل دست یافت. بنابراین GRACE می‌تواند تغییرات جرم در لایه سطحی زمین در مقیاس‌های بزرگ (مانند: تغییرات سطح آب‌های زیرزمینی، تغییرات جرم اقیانوس‌ها، بالا آمدن سطح دریاها و ذوب‌شدن صفحات یخ در مناطق قطبی) را به‌خوبی تخمین زند. طی سال‌های اخیر تحقیقات بسیاری در زمینه مشاهده تغییرات میدان ثقل ناشی از زلزله‌های بزرگ با استفاده از داده‌های جفت ماهواره GRACE انجام گرفته است. هان و همکاران (۲۰۰۶) (دو سال پس از وقوع زلزله مهیب سوماترای اندونزی Mw=9.1-9.3) برای اولین بار با استفاده از داده‌های سطح ۱ ماهواره GRACE توانستند تغییر شکل همزمان با وقوع زلزله سوماترا را کشف کنند. آن‌ها به منظور کاهش نویز موجود در داده‌های GRACE از فیلتر گوسین ۵۰۰ کیلومتر استفاده کردند. به علاوه به منظور ارزیابی نتایج به‌دست‌آمده از GRACE، از مدل‌های پیش‌بینی تغییرات میدان ثقل استفاده کردند. پس از انتشار مقاله هان و همکاران (۲۰۰۶)، مقالات متعددی در سال‌های بعد به منظور ارزیابی داده‌های GRACE در کشف سیگنال هم‌لرز و پس‌لرز ناشی از زلزله‌های ۲۰۰۴ سوماترا، ۲۰۱۰

شیلی و ۲۰۱۱ توهوکوی ژاپن منتشر شده، اما تمرکز اصلی همه مطالعات روی اندازه‌گیری تغییرات هم‌لرز و پس‌لرز ناشی از این زلزله‌ها بوده است. هر چند تلاش‌هایی برای مشاهده تغییرات میدان ثقل قبل از وقوع زلزله‌های بزرگ با استفاده از داده‌های این ماهواره انجام گرفته است. سیجی و ناکامورا (۲۰۱۳) به منظور کشف تغییرات ناگهانی قبل از وقوع زلزله توهوکوی ژاپن در سری‌های زمانی تغییرات میدان ثقل به‌دست‌آمده از داده‌های ماهواره GRACE تلاش‌هایی کردند. اما به دلیل استفاده از تابعک تغییرات شتاب ثقل، آن‌ها هیچ‌گونه آنامولی در داده‌های شتاب ثقل کشف نکردند. برای اولین بار شهریسوند و همکاران (۲۰۱۴) نشان دادند که می‌توان با استفاده از مشتقات مرتبه دوم میدان ثقل، تغییرات ناگهانی قبل از وقوع زلزله‌های بزرگ را در داده‌های ماهواره GRACE مشاهده کرد. آن‌ها برای اثبات این ادعا، سری زمانی تغییرات میدان ثقل را قبل و بعد از وقوع چهار زلزله بزرگ با استفاده از روش‌های IQR و شبکه‌های عصبی مصنوعی بررسی کردند و توانستند ۲ تا ۴ هفته قبل از وقوع زلزله آنامولی‌هایی را کشف کنند.

هدف این مقاله، بررسی رفتار تغییرات تابعک‌های میدان ثقل (که از داده‌های ماهواره‌ای به دست آمده است) به منظور کشف ناهنجاری میدان ثقل قبل از وقوع زلزله، با تکیه بر قابلیت‌های داده‌های ماهواره‌ای GRACE است. در صورتی که این ناهنجاری‌ها در داده‌های ماهواره GRACE قابل مشاهده باشد، می‌توانند در کنار سایر پیش‌نشانگرها در کاهش عدم قطعیت پیش‌بینی زلزله کمک شایانی کنند. برای رسیدن به این هدف سری‌های زمانی برخی از تابعک‌های میدان ثقل در بازه‌های زمانی طولانی‌مدت، تولید و با استفاده از پنج روش مختلف تجزیه و تحلیل شده است. با توجه به اینکه داده‌های ماهواره GRACE فقط توانایی کشف سیگنال تغییر شکل ناشی از زلزله‌هایی با بزرگای بیش از ۷/۵ ریشتر را دارند، فقط این نوع زلزله‌ها را می‌توان بررسی کرد. قابل مشاهده بودن ناهنجاری‌های میدان ثقل با استفاده از داده‌های ماهواره‌ای علاوه بر اینکه به

است (چنگ و تپلی، ۲۰۰۴). سپس اختلاف ضرایب هر مدل هفتگی از یک مدل میانگین که از میانگین‌گیری سال‌های ۲۰۰۵ تا ۲۰۱۳ به دست آمده بود محاسبه شد. به این ترتیب برای هر هفته تغییرات ضرایب هارمونیک‌های کروی نسبت به مدل میانگین مطرح‌شده به دست می‌آید. با استفاده از تغییرات ضرایب هارمونیک‌های کروی محاسبه‌شده برای هر هفته می‌توان تغییرات تابعک‌های مختلف میدان ثقل را محاسبه کرد. در این مطالعه برخلاف مطالعات قبلی از برخی مؤلفه‌های تانسور مرتبه دوم به‌جای تغییرات شتاب ثقل استفاده شده است. زیرا این مؤلفه‌ها علاوه بر تقویت فرکانس‌های بالای میدان ثقل، در مقایسه با سایر تابعک‌های میدان ثقل، به خطای نوارنواری‌شدگی داده‌های GRACE حساسیت کمتری دارند. از ۶ مؤلفه تانسور مرتبه دوم پتانسیل ثقل، دو مؤلفه ΔV_{xx} , ΔV_{zz} به خطای نوارنواری‌شدگی موجود در داده‌های GRACE حساسیت کمتری دارند. همان‌طور که در شکل ۱ دیده می‌شود نوارهای شمالی - جنوبی موجود در داده‌های GRACE سبب به‌وجود آمدن نوسانات شدید تابع (تغییرات میدان ثقل) در راستای شرقی - غربی می‌شود. اما این نوارها هیچ تأثیری در راستای شمالی - جنوبی ندارند. بنابراین در صورت مشتق‌گیری نسبت به راستای شمالی - جنوبی فقط تغییرات در این راستا باقی خواهد ماند و نوارها کاهش قابل‌توجهی خواهند یافت. این واقعیت برای زلزله توهوکوی ژاپن در شکل ۲ نشان داده شده است. در به‌دست‌آوردن شکل ۲ هیچ‌گونه فیلتری روی داده‌های GRACE اعمال نشده است، اما همان‌طور که دیده می‌شود خطای نوارنواری‌شدگی موجود در داده‌های GRACE کاهش قابل‌توجهی پیدا کرده است. به‌علاوه از شکل ۲ می‌توان نتیجه گرفت که پدیده‌های غیرپریودیك هیدرولوژی، مانند بازگشت پوسته زمین ناشی از عصر یخبندان در کانادا و ذوب توده‌های یخ در گرینلند نیز حذف شده‌اند، زیرا پدیده‌های مذکور با فرکانس پایین در حیطه مکان رخ می‌دهند. بنابراین وقتی مشتق نسبت به راستای شمالی - جنوبی گرفته می‌شود، لبه‌های آنامولی جرم ناشی از این

چگونگی وقوع زلزله و نوع گسل وابسته است، به حذف کلیه پارامترهای غیرژئوفیزیکی مانند سیگنال‌های هیدرولوژی نیز وابسته است. در روند پیاده‌سازی و بررسی زلزله‌های بزرگ در این مقاله و مقالات قبلی چاپ شده از نویسندگان، زلزله سراوان نیز بررسی شده است، اما با توجه به اینکه زلزله سراوان در عمق تقریباً ۷۰ کیلومتری از سطح زمین رخ داده است، داده‌های این ماهواره توانایی کشف سیگنال تغییر شکل ناشی از این زلزله را نداشتند.

بنا بر آنچه گفته شد، ساختار مقاله بدین گونه است: بخش اول شامل مقدمه و پیشینه تحقیق است. در بخش دوم پیش‌پردازش‌های انجام‌گرفته روی داده‌های GRACE بررسی خواهند شد. بخش سوم شامل معرفی روش‌های استفاده‌شده در این تحقیق جهت کشف بی‌هنجاری‌های میدان ثقل است. بخش چهارم نیز به بررسی نتایج و تحلیل آن‌ها می‌پردازد. در انتها و بخش آخر نیز نتیجه‌گیری مقاله آمده است.

۲. پیش‌پردازش داده‌های GRACE

در این تحقیق به منظور بررسی آشفتگی‌های میدان ثقل قبل از وقوع زلزله از داده‌های هفتگی مرکز GFZ (Release 05) که توسط فیلتر DDK3 هموار شده‌اند، استفاده شده است (کوشه و همکاران، ۲۰۰۹). این داده‌ها مدل‌های هفتگی تغییرات میدان ثقل زمین را تا درجه و مرتبه ۳۰ ضرایب هارمونیک‌های کروی در اختیار قرار می‌دهند. به‌طور کلی از ۴۴۸ مدل هفتگی که بازه زمانی هفته اول ژانویه ۲۰۰۵ تا هفته سوم فوریه ۲۰۱۳ را پوشش می‌دهند استفاده شد. دلیل استفاده از داده‌های هفتگی به‌جای استفاده از داده‌های ماهیانه، قدرت بیشتر تفکیک زمانی در داده‌های هفتگی است که می‌تواند در بررسی تغییرات غیرعادی میدان ثقل اهمیت زیادی داشته باشد. زیرا ممکن است رفتار غیرعادی به‌وجودآمده قبل از وقوع زلزله در بازه زمانی کوتاه‌تر از یک ماه اتفاق بیفتد. به‌دلیل دقت کم در به‌دست‌آوردن ضریب C_{20} از مشاهدات GRACE، این ضریب با استفاده از مشاهدات SLR جایگزین شده

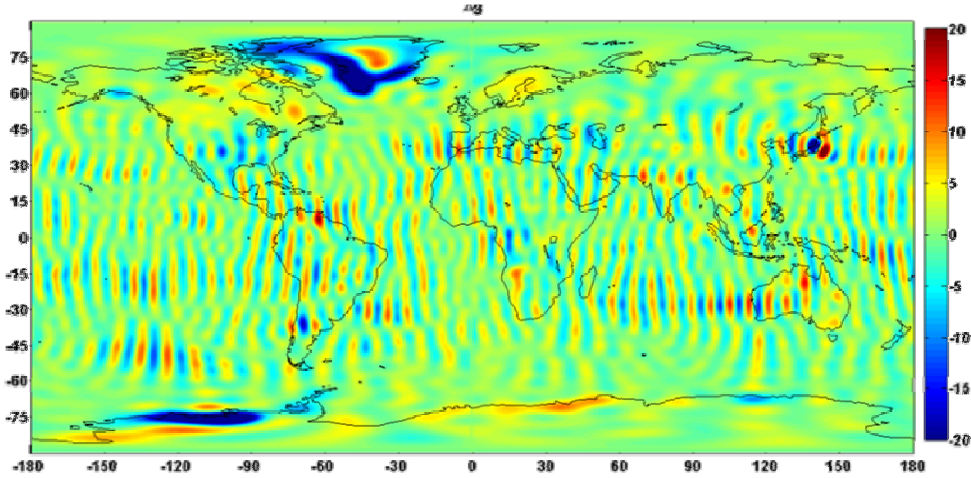
مؤلفه‌های $\Delta V_{r\theta}$ و $\Delta V_{\theta\theta}$ را با استفاده از داده‌های هفتگی تغییرات ضرایب هارمونیک‌های کروی در سیستم مختصات کروی نشان می‌دهند (ونگ و همکاران، ۲۰۱۲).

با توجه به اینکه هدف بررسی تغییرات محلی مؤلفه‌های تانسور مرتبه دوم پتانسیل ثقل است، این مؤلفه‌ها پس از به‌دست آمدن در سیستم مختصات کروی، با استفاده از تبدیل بین سیستم‌های مختصات، به سیستم مختصات محلی برده می‌شوند. در سیستم مختصات محلی در هر نقطه محور x به سمت شمال، محور y به سمت شرق و محور z به سمت داخل است. پس از اینکه در هر هفته تغییرات دو تابعک گفته شده محاسبه شد، برای هر نقطه می‌توان سری زمانی تغییرات میدان ثقل را محاسبه کرد.

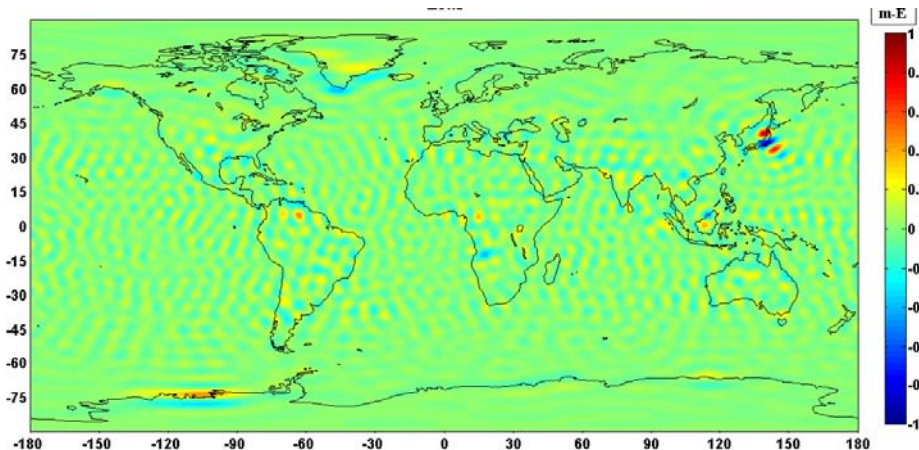
پدیده‌ها به قدری ضعیف هستند که در مقابل سیگنال تغییر شکل ناشی از زلزله دیده نمی‌شوند. با توجه به اینکه الگوی مکانی تغییرات به‌وجود آمده ناشی از زلزله با فرکانس بالا در حیطه مکان رخ می‌دهد؛ بنابراین انتظار نیز می‌رود که تغییرات به‌وجود آمده ناشی از زلزله با دامنه بالایی در مقایسه با پدیده‌های فرکانس پایین دیده شوند. با توجه به مطالب گفته‌شده مؤلفه‌هایی از تانسور مرتبه دوم پتانسیل ثقل که با مشتق‌گیری نسبت به راستای شمالی- جنوبی به دست می‌آیند $(\Delta V_{xx}, \Delta V_{xz})$ خطای نوآرشدگی موجود در داده‌های GRACE را کاهش می‌دهند. بنابراین در این تحقیق برای بررسی رفتارهای غیرعادی میدان ثقل قبل از وقوع زلزله فقط از دو مؤلفه $\Delta V_{xx}, \Delta V_{xz}$ استفاده شده است. روابط ۱ و ۲ چگونگی محاسبه تغییرات

$$\Delta V_{r\theta}(r, \theta, \lambda) = \frac{GM}{R^2} \sum_{l=2}^{\infty} \sum_{m=0}^l (l+1) \left(\frac{R}{r}\right)^{l+2} (\Delta \bar{C}_{lm} \cos m\lambda + \Delta \bar{S}_{lm} \sin m\lambda) \bar{P}'_{lm}(\cos \theta) \sin \theta \quad (1)$$

$$\Delta V_{\theta\theta}(r, \theta, \lambda) = \frac{GM}{R} \sum_{l=2}^{\infty} \sum_{m=0}^l \left(\frac{R}{r}\right)^{l+1} (\Delta \bar{C}_{lm} \cos m\lambda + \Delta \bar{S}_{lm} \sin m\lambda) (\bar{P}''_{lm}(\cos \theta) \sin^2 \theta - \bar{P}'_{lm}(\cos \theta) \cos \theta) \quad (2)$$



شکل ۱. تغییرات شتاب ثقل ناشی از زلزله ۱۱ مارس ۲۰۱۱ توهوکو بدون اعمال هیچ‌گونه فیلتری (واحد: میکروگال).



شکل ۲. تغییرات مشتق شمالی شتاب ثقل ناشی از زلزله ۱۱ مارس ۲۰۱۱ توهوکو بدون اعمال هیچ‌گونه فیلتری (واحد: میلی اتوش).

نشان می‌دهد (چن و همکاران، ۲۰۰۷). بنابراین یک ترکیب خطی از این فرکانس‌ها و یک روند خطی را به سری زمانی تغییرات میدان ثقل می‌توان برآزش داد و فرکانس‌های مورد نظر را از سری زمانی حذف کرد. به علاوه اثر جزر و مدی S_2 در مشاهدات GRACE قابل توجه است و باید در نظر گرفته شود (دی لینچ و همکاران، ۲۰۰۹). بنابراین از ترکیب خطی زیر برای حذف تأثیرات غیر ژئوفیزیکی در این تحقیق استفاده شده است:

$$y(t, t_0) = \begin{cases} A + Bt + \sum_{i=1}^3 a_i \cos(\omega_i t + \phi_i) & \text{if } t \leq t_0 \\ A' + B't + \sum_{i=1}^3 a_{i+3} \cos(\omega_i t + \phi_{i+3}) & \text{if } t \geq t_0 \end{cases} \quad (۳)$$

که در این رابطه t_0 تاریخ وقوع زلزله است و سایر پارامترهای مدل عبارت‌اند از:

(۱) $a_1, \phi_1, a_2, \phi_2, a_3, \phi_3$ به ترتیب فاز و دامنه سیگنال‌های هیدرولوژی نیم‌سالانه، سالیانه و فرکانس جزر و مدی S_2 قبل از وقوع زلزله هستند.

(۲) $a_4, \phi_4, a_5, \phi_5, a_6, \phi_6$ به ترتیب فاز و دامنه سیگنال‌های هیدرولوژی نیم‌سالانه و سالیانه و فرکانس جزر و مدی S_2 پس از وقوع زلزله هستند.

(۳) B, B' ترند خطی قبل و بعد از وقوع زلزله هستند.

(۴) A, A' میانگین سیگنال قبل و بعد از وقوع زلزله هستند.

و W در این روابط مشخص‌کننده فرکانس سیگنال‌های هیدرولوژی مورد نظر است. به عنوان مثال برای به دست آوردن فرکانس سالیانه مقدار ω برابر با $\frac{2\pi}{365}$ است.

در صورتی که هدف بررسی تغییرات میدان ثقل قبل از وقوع زلزله باشد، قسمت دوم از تابع پیشنهاد شده حذف می‌شود. اما با توجه به اینکه سری زمانی موجود برای بررسی در این مقاله شامل داده‌های پس از زلزله نیز هست، نمی‌توان فقط از قسمت اول استفاده کرد و به ناچار قسمت دوم نیز اضافه شده است. به علاوه با توجه به اینکه مقدار میانگین سیگنال قبل و بعد از وقوع زلزله یکسان نیست، نمی‌توان از یک تابع یکسان برای حذف

به منظور به دست آوردن تغییرات میدان ثقل ناشی از زلزله باید تأثیرات غیر ژئوفیزیکی از سیگنال تغییرات میدان ثقل حذف شود. از جمله عواملی که موجب تغییرات میدان ثقل می‌شود، می‌توان تغییراتی را نام برد که به وسیله عوامل هیدرولوژیکی ایجاد می‌شود. اثر تغییرات هیدرولوژیکی می‌تواند به قدری قوی باشد که بر سیگنال تغییر شکل ناشی از زلزله غالب شود. در یک چرخه هیدرولوژی در طول یک سال در فصل زمستان به علت افزایش نزولات جوی، افزایش جرم و چگالی را در منطقه مورد نظر خواهیم داشت؛ در حالی که در فصل تابستان به علت کمبود بارش و افزایش تبخیر سطحی کاهش جرم و چگالی را در منطقه مورد نظر انتظار داریم. بنابراین طی یک دوره زمانی طولانی مدت در صورتی که در یک منطقه زلزله یا سایر مخاطرات طبیعی رخ ندهد، تغییرات میدان ثقل روند تناوبی خواهد داشت (چن و همکاران، ۲۰۰۹).

با توجه به مطالب گفته شده این تغییرات هیدرولوژی دارای دامنه قابل توجهی است و نادیده گرفتن آن‌ها موجب عدم اطمینان در نتایج برآورد شده خواهند شد. روش‌های مختلفی توسط محققان جهت حذف تأثیرات هیدرولوژی معرفی شده است. یکی از روش‌های مرسوم در حذف اثر پارامترهای هیدرولوژی در یک منطقه استفاده از مدل‌های جهانی هیدرولوژی است، اما مشکل عمده و اصلی در استفاده از چنین مدل‌های هیدرولوژی نداشتن داده در مناطق پوشیده شده با آب می‌باشد. بیشتر زلزله‌های بزرگ در دنیا در مناطقی رخ می‌دهند که محل اتصال یک صفحه اقیانوسی و یک صفحه قاره‌ای است؛ بنابراین قسمت عمده‌ای از دگرشکلی ایجاد شده ناشی از زلزله در آب است؛ لذا امکان استفاده از این مدل‌ها برای حذف تأثیرات هیدرولوژی در این مطالعه وجود ندارد.

با توجه به مطالب گفته شده درباره تناوب بودن تغییرات هیدرولوژی می‌توان سیگنال هیدرولوژی را یک سیگنال سینوسی با پیوند دوره تناوب آن در نظر گرفت. تغییرات هیدرولوژی معمولاً با پیوندهای یک‌ساله و شش ماهه در یک منطقه خود را در داده‌های GRACE

این پدیده چنان پیچیده است که طی سالیان طولانی مطالعه درباره مکانیزم وقوع زلزله، هنوز به طور کامل شناخته نشده است. در این مطالعه تلاش شده است با بررسی روش‌های گوناگون کلاسیک و پیشرفته و غیرخطی، روشی مناسب جهت تشخیص بی‌هنجاری‌های میدان ثقل پیشنهاد شود.

۳-۱. روش میانه و دامنه بین چارکی

متداول‌ترین روش برای تشخیص آنامولی در سری‌های زمانی به دست آمده از تغییرات هر پیش‌نشانگر زلزله، استفاده از میانگین و انحراف معیار داده‌هاست. برای استفاده از این روش برای استخراج آنامولی از سری زمانی لازم است که داده‌ها از توزیع نرمال تبعیت کنند. تغییرات میدان ثقل به پارامترهای بسیار زیادی بستگی دارد. از جمله این پارامترهای تأثیرگذار می‌توان به زمان، موقعیت جغرافیایی، جریان‌های دریایی، جزر و مد و ... اشاره کرد؛ بنابراین این تغییرات غیرخطی مانع استفاده از روش‌های مبتنی بر توزیع نرمال داده‌ها می‌شود. برخی محققان استفاده از مقادیر میانه و دامنه بین چارکی داده‌ها را پیشنهاد می‌کنند (لیو و همکاران، ۲۰۰۴). در این تحقیق نیز به با توجه به غیرخطی بودن سری‌های زمانی بررسی شده از این روش استفاده شده است.

محدوده‌های مجاز بالا و پایین با استفاده از روابط زیر به دست می‌آیند:

$$x_{high} = M + k \cdot IQR \quad (4)$$

$$x_{low} = M - k \cdot IQR \quad (5)$$

$$x_{low} < x < x_{high} \Rightarrow -k < \frac{x - M}{IQR} < k ; \quad (6)$$

$$Dx = \frac{x - M}{IQR}$$

در این روابط x ، x_{low} ، x_{high} ، M و IQR در ترتیب مقدار پارامتر، مرز بالایی، مرز پایینی، مقدار میانه، دامنه بین چارکی و انحراف پارامتر x می‌باشند. مطابق با رابطه ۶ اگر مقدار مطلق پارامتر Dx بزرگ‌تر از k ($|Dx| > k$) باشد، رفتار پارامتر مورد نظر (x)

تأثیرات هیدرولوژی استفاده کرد (دی لینیچ و همکاران، ۲۰۰۹).

با به دست آوردن دامنه تغییرات ناشی از پارامترهای هیدرولوژی با استفاده از معادله وابسته به زمان بالا، می‌توان اثر آن‌ها را حذف کرد. برای به دست آوردن ضرایب معادله بالا از روش سرشکنی کمترین مربعات استفاده می‌شود. پس از حذف این تأثیرات فقط تغییرات میدان ثقل ناشی از مخاطرات طبیعی یا همان زلزله باقی می‌ماند.

۳. کشف بی‌هنجاری در سری‌های زمانی تغییرات

میدان ثقل

طبق تحقیقات انجام گرفته، سایر پیش‌نشانگرها (مانند تغییرات مقدار الکترون‌های موجود در جو و دما) زلزله دقیقاً در تصویر قائم مرکز رخ نمی‌دهد، بلکه در اطراف این نقطه اتفاق می‌افتد (آخوندزاده، ۱۳۹۰). حتی سیگنال تغییر شکل هم‌لرز نیز در رومرکز زلزله بسیار ضعیف است و با یک شیفت مکانی خود را در اطراف رومرکز نشان می‌دهد. بنابراین نقطه‌ای که برای محاسبه سری زمانی انتخاب می‌شود، بسیار مهم است که به درستی انتخاب شود. طبق تحقیقات انجام گرفته آنامولی تغییرات میدان ثقل در رومرکز زلزله رخ نمی‌دهد. به علاوه این آنامولی در محلی که بیشترین جابه‌جایی هم‌لرز رخ می‌دهد هم اتفاق نمی‌افتد و یک شیفت مکانی نسبت به آن دارد (سیجی و ناکامورا، ۲۰۱۳). بنابراین در هر مورد مطالعاتی، یک شبکه منظم 1×1 درجه در اطراف رومرکز زلزله تشکیل و تغییرات میدان ثقل در کلیه نقاط این شبکه، به منظور کشف بی‌هنجاری بررسی شد. انتخاب ابعاد شبکه با توجه به قدرت تفکیک مکانی داده‌های GRACE بوده است.

پس از محاسبه سری زمانی، تشخیص بی‌هنجاری یکی از مراحل مهم و کلیدی برای بررسی امکان استفاده از آن پیش‌نشانگر است. انتخاب روش‌های استخراج بی‌هنجاری به خطی یا غیرخطی بودن رفتار سری‌های زمانی بستگی دارد. زلزله مانند سایر پدیده‌های ژئوفیزیکی دارای فرآیندهای غیرمنظم و پیچیده است.

۳. الگوریتم‌های ID3, C4, C5

(Induction of Decision Tree)

۴. الگوریتم QUEST

(Quick Unbiased Efficient Statistical Trees)

۵. الگوریتم CRUISE

(Classification Rule with Unbiased Interaction Selection and Estimation)

در این مطالعه برای دستیابی به الگوریتم بهینه، با بررسی میانگین، انحراف معیار و توزیع مقادیر خطا در همه روش‌ها، الگوریتم C5 به عنوان روش انتخابی جهت پیش‌بینی سری زمانی تغییرات میدان ثقل انتخاب شد. با توجه به اینکه الگوریتم C5 از معروف‌ترین و کارآمدترین روش‌های تولید درخت تصمیم‌گیری است، در این مقاله نیز از این الگوریتم استفاده شد و سایر الگوریتم‌ها مورد بررسی قرار نگرفت.

۳-۳. روش‌های یادگیری دسته‌جمعی

همواره ادغام در سطوح مختلف موجب افزایش صحت و دقت نتایج یک الگوریتم می‌شود؛ چرا که روش‌های مختلف رگرسیون و پیش‌بینی هر یک نقاط ضعف و قوتی دارند و با ترکیب و ادغام نتایج آن‌ها می‌توان نتیجه‌ای به دست آورد که علاوه بر اینکه نقاط ضعف این روش‌ها را پوشش دهد شامل نقاط قوت همه روش‌ها باشد. ادغام در سطوح مختلفی انجام می‌گیرد که بالاترین سطح آن ادغام در سطح تصمیم است. یکی از مهم‌ترین نکات در استفاده از روش‌های پیش‌بینی سری‌های زمانی که دارای رفتاری بسیار پیچیده و غیرخطی هستند، ناکارآمد بودن استفاده از یک الگوریتم پیش‌بینی، به تنهایی است. از آنجا که هر الگوریتم بر اساس ویژگی‌های خاصش نتایج متفاوتی را تولید می‌کند، استفاده همزمان از مجموعه‌ای از پیش‌بینی‌کننده‌ها امید بیشتری به ارائه نتیجه‌ای قابل اعتماد ایجاد می‌کند. این روش‌ها که تحت عنوان روش‌های یادگیری دسته‌جمعی شناخته می‌شوند از مجموعه‌ای از چند روش یا جایگشت‌های مختلفی از داده‌ها محاسبات را اجرا می‌کند و بعد از تلفیق نتایج آن‌ها، تصمیم نهایی را می‌گیرد. از جمله این روش‌ها

غیرعادی تشخیص داده می‌شود. همچنین بر اساس رابطه $p = \pm 100 \times (|Dx| - k) / k$ درصد انحراف پارامتر از وضعیت عادی نشان داده می‌شود. با فرض اینکه داده‌ها دارای توزیع نرمال با میانگین μ و انحراف معیار σ باشند، مقادیر میانه و دامنه بین چارکی به ترتیب با μ و $1.34 \times \sigma$ برابر است (آخوندزاده، ۱۳۹۰). اگر مقدار پارامتر مشاهده شده بیرون محدوده بالایی یا پایینی قرار گیرد، با سطح اطمینان ۸۰ تا ۸۵ درصد می‌توان بیان کرد که مقدار غیرعادی پارامتر تشخیص داده شده است (لیو و همکاران، ۲۰۰۴). مقدار پارامتر k در این رابطه، با توجه به بزرگای زلزله، محل جغرافیایی آن، ساختار ژئوفیزیکی و بسیار پارامترهای دیگر به دست می‌آید.

۳-۲. روش درخت تصمیم‌گیری

روش درخت تصمیم‌گیری از جمله روش‌های غیرپارامتریک است که با توجه به نوع متغیر وابسته، به دو دسته طبقه‌بندی درختی (classification Tree) و رگرسیون درختی (regression Tree) برای متغیر پیوسته تقسیم‌بندی می‌شود. الگوی این روش از ساختاری بسیار ساده و قابل درک برای تصمیم‌گیری استفاده می‌کند؛ البته با وجود اینکه این روش از الگوهای ساده برای تصمیم‌گیری استفاده می‌کند اما در زمینه تشخیص و پیشگویی می‌تواند به‌خوبی روش‌های پیچیده‌ای نظیر شبکه‌های عصبی عمل کند (خسروی و آخوندزاده، ۱۳۹۳).

در این الگوریتم از مجموعه‌ای از شرط‌های منطقی به صورت یک الگوریتم با ساختار درختی برای طبقه‌بندی یا پیش‌بینی (رگرسیون) یک پیامد استفاده می‌شود. با توجه به اینکه شاخص‌های متنوع و روش‌های گوناگونی برای تعیین و تولید یک درخت تصمیم‌گیری معرفی شده است، الگوریتم‌های متنوعی نیز ارائه شده است که عبارتند از:

۱. الگوریتم CHAID

(Chi-squared Automatic Interaction Detector)

۲. الگوریتم CART

(Classification and regression Tree)

بهرتر از یک حدس تصادفی عمل می‌کند- می‌تواند به یک طبقه‌بندی کننده قوی تبدیل شود که قادر است به‌طور صحیح بخش زیادی از نمونه‌ها را طبقه‌بندی کند. بگینگ در برخی موارد مشابه با بوستینگ است چون هر دو روش مجموعه‌ای از طبقه‌بندی کننده‌ها را طراحی کرده و نتایجشان را با رأی‌گیری حداکثری یا میانگین‌گیری ترکیب می‌کنند. به‌هرحال، این دو روش با یکدیگر تفاوت‌هایی دارند، از جمله آن‌که بگینگ همواره از نمونه‌برداری مجدد به‌جای وزن‌دهی مجدد استفاده می‌کند. ضمن آن‌که این روش توزیع نمونه‌ها را تغییر نمی‌دهد و همه کلاس‌ها دارای وزن یکسان در طول فرآیند میانگین‌گیری هستند. به‌علاوه که بگینگ می‌تواند به‌صورت موازی انجام گیرد، یعنی همه bagها به‌صورت یک مرتبه طراحی می‌شوند. درمقابل، بوستینگ همواره به‌صورت پشت‌سرم و دنباله‌ای انجام می‌گیرد و هر مجموعه نمونه مبتنی بر آخرین وزنش است (فریدام و همکاران، ۲۰۰۱).

۳-۳-۳. جنگل تصادفی

جنگل تصادفی یک مدل بهبودیافته از روش بگینگ است و علت نامگذاری آن، این است که از درخت‌های تصمیم ساخته می‌شود. یک جنگل تصادفی می‌تواند از درخت‌های تصمیم تکی ساخته شود که پارامترهای آموزشی مخصوص آن به‌طور تصادفی تغییر می‌کند. همانند روش بگینگ، چنین پارامترهایی می‌تواند جای‌گذاری‌های بوت استرپ شده داده‌های آموزشی باشند؛ درعین حال همانند روش‌های جای‌گذاری تصادفی، آن‌ها می‌توانند زیرمجموعه‌های ویژگی مختلف نیز باشند. جنگل تصادفی نشان داده است که از لحاظ دقت قابل مقایسه با بوستینگ است؛ بدون آنکه آسیب‌های آن از جمله حساسیت به نویز و بار محاسباتی بالا را داشته باشد (فریدام و همکاران، ۲۰۰۱).

۴. پیاده‌سازی و تحلیل نتایج

زلزله مورد مطالعه در این تحقیق زلزله ۸/۶ ریشتری سال ۲۰۱۲ اقیانوس هند است. در جدول ۱ مشخصات این

می‌توان به بگینگ، بوستینگ و جنگل تصادفی که مبتنی بر دستکاری نمونه‌های آموزشی هستند، اشاره کرد. به‌طور کلی می‌توان روش‌های یادگیری دسته‌جمعی را به چهار دسته تقسیم‌بندی کرد:

۱. روش‌های با الگوریتم‌های مختلف (الگوریتم‌های مختلفی بر روی داده یکسانی اعمال شده و در پایان نتایج به دست آمده از هر کدام از این الگوریتم‌ها با هم تلفیق می‌شوند)

۲. الگوریتم‌هایی با انتخاب متفاوت پارامترها

۳. مجموعه داده‌ها با فضای ویژگی متفاوت (جنگل تصادفی)

۴. مجموعه داده‌ها شامل زیرمجموعه‌های مختلف (بگینگ و بوستینگ)

در این مقاله به منظور بررسی کارایی و کسب نتایج هرچه بهتر هر سه روش بگینگ، بوستینگ و جنگل تصادفی پیاده‌سازی شده است.

۳-۳-۱. بگینگ

در یادگیری ماشین، بگینگ مخفف کلمه جمع‌آوری خودراه‌انداز (Bootstrap Aggregating) است. در بگینگ، زیرمجموعه‌های آموزشی مختلف (bagها)، به‌طور تصادفی و با جای‌گذاری از کل داده‌های آموزشی انتخاب می‌شود. سپس، روش‌های پیش‌بینی (رگرسیون) مختلف از یک نوع یکسان توسط هر زیرمجموعه، آموزش داده می‌شود. این پیش‌بینی کننده‌های تکی با به‌کارگیری یک روش میانگین‌گیری از تصمیماتشان با یکدیگر ترکیب می‌شوند. برای یک نمونه آزمایشی، مقدار پیش‌بینی برابر خواهد بود با مقداری که از میانگین‌گیری همه پیش‌بینی کننده‌ها به دست آمده است.

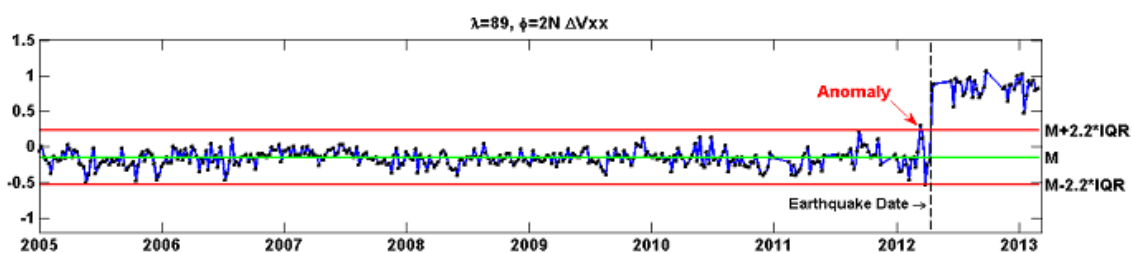
۳-۳-۲. بوستینگ

بوستینگ یک روش عمومی و مشهور در سیستم‌های تصمیم‌گیری جمعی است که برای افزایش دقت هر پیش‌بینی کننده استفاده می‌شود. اسپچاپیر (۱۹۹۰) ثابت کرد که یک طبقه‌بندی کننده ضعیف -الگوریتمی که

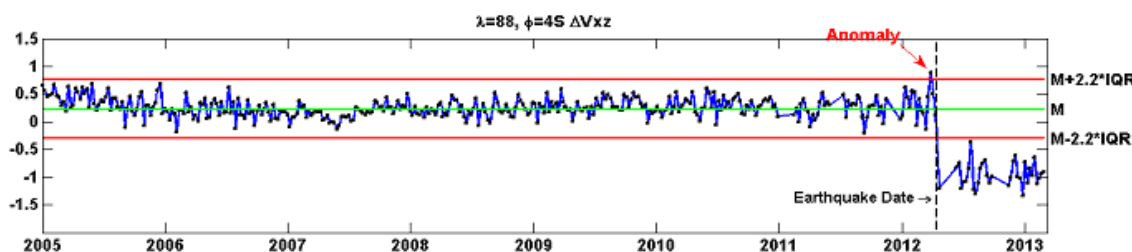
شده‌اند. با اعمال روش IQR برای استخراج آنامولی از سری زمانی ΔV_{xx} حد آستانه‌های بالا و پایین به ترتیب $M + 2.2 \times IQR$ و $M - 2.2 \times IQR$ انتخاب شدند. خط‌چین مشکی تاریخ وقوع زلزله و محور افقی بیان‌گر تاریخ محاسبه تابعک مورد نظر است. میانه و حد آستانه‌های بالا و پایین به ترتیب خطوط سبز و قرمز رنگ هستند. پنج هفته قبل از وقوع زلزله مقدار ΔV_{xx} از حد آستانه‌های بالا با مقدار ۱۸٪ عبور می‌کند. در سری زمانی ΔV_{xz} نیز پنج هفته قبل از وقوع زلزله تغییرات ΔV_{xz} با مقدار ۲۴٪ از حد آستانه بالایی تعریف شده عبور می‌کند؛ البته سه هفته قبل از وقوع زلزله نیز مورد مشکوک دیگری وجود دارد که توسط روش IQR قابل شناسایی نبوده است. به علاوه در تاریخ وقوع زلزله نیز پرش قابل توجهی در سری زمانی تغییرات هر دو مؤلفه دیده می‌شود که ناشی از دگرشکلی هم‌لرز این زلزله است. این پرش در سری زمانی توانایی داده‌های هفتگی در کشف دگرشکلی هم‌لرز ناشی از زلزله‌های بزرگ را نشان می‌دهد.

جدول ۱. مشخصات زلزله اقیانوس هند.

منطقه مطالعاتی	زمان	طول جغرافیایی	عرض جغرافیایی	بزرگا	عمق کانونی
اقیانوس هند	2012 Apr 11	93.063°E	2.327°N	8.6 MW	20 km



(الف)



(ب)

شکل ۳. سری زمانی تغییرات مؤلفه‌های ΔV_{xx} و ΔV_{xz} به ترتیب (الف) و (ب) به دست آمده از داده‌های هفتگی پروژه ماهواره‌ای GRACE که با استفاده از فیلتر DDK3 هموار شده‌اند (کوشه، ۲۰۰۷). فرکانس‌های سالیانه، نیم‌سالانه، فرکانس جزر و مدی S_2 و ترند خطی با استفاده از سرشکنی کمترین مربعات حذف شده است. خط‌چین تیره زمان وقوع زلزله را نشان می‌دهد و خطوط قرمز حد آستانه‌های بالا و پایین را نشان می‌دهد. خط سبز مقدار میانه را نشان می‌دهد. پارامتر λ معرفی‌کننده طول جغرافیایی است.

زمین‌لرزه آمده است. این زلزله که در ۵۰۰ کیلومتری زلزله ۲۰۰۴ سوماترای اندونزی (با بزرگای ۹/۱ ریشتر) اتفاق افتاد، بزرگ‌ترین زلزله‌ای است که در گسل‌های راستالغز به ثبت رسیده است. دو ساعت پس از وقوع این زلزله در ۱۲۰ کیلومتری غربی مرکز زلزله، زلزله دیگری به بزرگی ۸/۲ ریشتر در همان روز اتفاق افتاد (شریواستاوا و ردی، ۲۰۱۳). همان‌طور که گفته شد زلزله ۲۰۱۲ اقیانوس هند در نزدیکی زلزله ۲۰۰۴ سوماترا رخ داده است. بنابراین دگرشکلی ناشی از زلزله سوماترا در سری زمانی مربوط به زلزله اقیانوس هند دیده خواهد شد. بنابراین سری‌های زمانی استفاده شده در مورد این زلزله از ابتدای سال ۲۰۰۵ (یک ماه پس از زلزله سوماترا) هستند. به منظور ارزیابی و مقایسه چهار روش پیشنهادی، در این مقاله برای پیش‌بینی سری‌های زمانی تغییرات میدان ثقل، دو مؤلفه ΔV_{xx} ، ΔV_{xz} از تانسور مرتبه دوم پتانسیل ثقل انتخاب شدند.

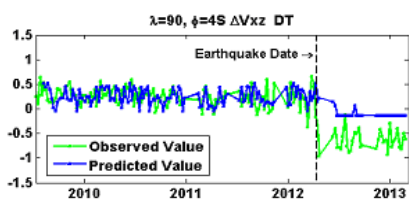
در قسمت الف و ب شکل ۳ تغییرات مؤلفه‌های ΔV_{xx} و ΔV_{xz} در اطراف رومرکز این زلزله نشان داده

۴-۱. روش درخت تصمیم‌گیری

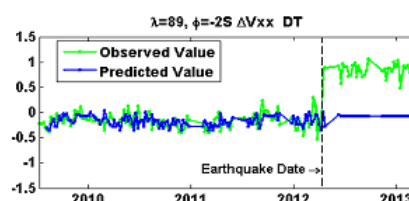
به منظور اجرای روش درخت تصمیم‌گیری جهت پیش‌بینی سری زمانی تغییرات میدان ثقل، از کل داده‌های موجود در سری زمانی ۶۰ درصد داده‌ها به عنوان داده‌های آموزشی و ۴۰ درصد به عنوان داده تست در نظر گرفته شد. داده‌های آموزشی به منظور تشکیل درخت تصمیم و انتخاب ویژگی‌های مورد نظر استفاده می‌شوند. سپس با استفاده از درخت به دست آمده از مرحله آموزش، سایر داده‌ها به منظور ارزیابی روش پیشنهادی استفاده می‌شوند.

شکل ۴ (الف و ب) مقادیر مشاهدات به دست آمده از داده‌های ماهواره GRACE و مقادیر پیش‌بینی شده به وسیله روش درخت تصمیم‌گیری را نشان می‌دهد. همان‌طور که در این شکل دیده می‌شود تا قبل از وقوع زلزله روش درخت تصمیم‌گیری با دقت قابل قبولی توانسته است رفتار تغییرات تابعک‌های میدان ثقل را پیش‌بینی کند. اما بعد از وقوع زلزله به دلیل تغییر ناگهانی (پرش در سری زمانی) در نمودار و تغییر رفتار تابع، روش درخت تصمیم‌گیری نتوانسته است رفتار تابع را پیش‌بینی کند. علت این امر را می‌توان به ساختار نسبتاً ساده درخت تشکیل شده با استفاده از داده‌های آموزشی نسبت داد؛ چرا که در طول روند تشکیل درخت تصمیم از داده‌های تاریخ قبل از وقوع زلزله استفاده شده و قوانین تصمیم نیز بر مبنای رفتار تابع در آن بازه زمانی است. شکل ۴ (ج و د) اختلاف بین مقادیر پیش‌بینی

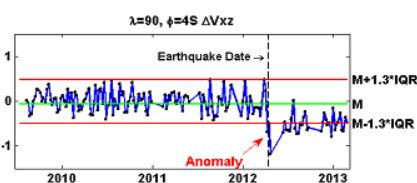
شده و مقادیر مشاهده شده در شکل ۴ (الف و ب) است. همان‌طور که در این شکل دیده می‌شود روش درخت تصمیم توانسته است به خوبی تغییرات ناگهانی میدان ثقل را سه هفته قبل از وقوع زلزله برای تابعک ΔV_{xx} و یک هفته قبل از وقوع زلزله برای تابعک ΔV_{xz} کشف کند. اختلاف مقادیر بین مقادیر پیش‌بینی و مشاهده شده تابعک ΔV_{xx} با مقدار ۱۲ درصد از حد آستانه پایینی عبور کرده است. حد آستانه‌های انتخاب شده با استفاده از روش میانه و فاصله بین چارکی انتخاب شده‌اند. این روش یک روش کلاسیک و متداول در حوزه بررسی پیش‌نشانگرهای زلزله است (آخوندزاده و همکاران، ۲۰۱۰). در شکل ۴-د نیز مشابه شکل ۴-ج اختلاف مقادیر بین مقادیر پیش‌بینی و مشاهده شده تابعک ΔV_{xz} با مقدار ۱۲ درصد از حد آستانه پایینی عبور کرده است؛ البته بررسی دقیق‌تر شکل ۴-ج نشان می‌دهد که پنج هفته قبل از وقوع زلزله نیز بی‌هنجاری در تغییرات میدان ثقل دیده می‌شود که به علت مقدار کم، این بی‌هنجاری در این تحقیق در نظر گرفته نشده است. به طور کلی در بررسی و کشف بی‌هنجاری‌های سری زمانی در صورتی که اختلاف بین مقادیر پیش‌بینی شده و مقادیر مشاهده از حد آستانه‌های روش میانه و فاصله بین چارکی بیشتر شود، این مقادیر به عنوان بی‌هنجاری در نظر گرفته می‌شود. چگونگی به دست آوردن حد آستانه‌ها در روش میانه و فاصله بین چارکی در بخش ۴ مقاله به تفصیل توضیح داده شده است.



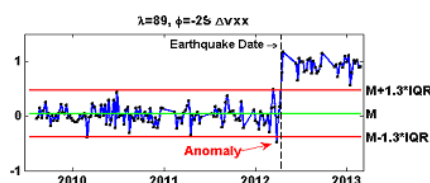
(ب)



(الف)



(د)



(ج)

شکل ۴. تغییرات مشاهده شده (منحنی سبز رنگ) و پیش‌بینی شده (منحنی آبی رنگ) مؤلفه‌های ΔV_{xx} (الف) و ΔV_{xz} (ب) با استفاده از الگوریتم درخت تصمیم‌گیری؛ (ج) و (د) اختلاف بین مقادیر مشاهده شده و پیش‌بینی شده در شکل‌های (الف) و (ب).

۴-۲. روش بگینگ

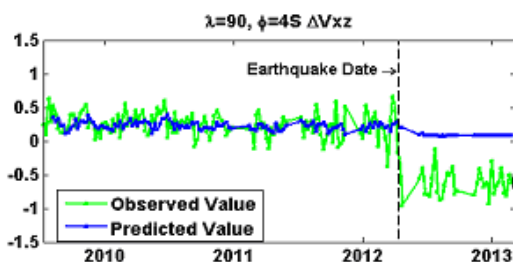
مشابه روش درخت تصمیم به منظور اجرای الگوریتم بگینگ جهت پیش‌بینی سری زمانی تغییرات میدان ثقل از کل داده‌های موجود در سری زمانی ۶۰ درصد داده‌ها به عنوان داده‌های آموزشی و ۴۰ درصد به عنوان داده تست در نظر گرفته شد. شکل ۵ نتایج پیاده‌سازی این الگوریتم را نشان می‌دهد. لازم به ذکر است که داده‌های آموزشی استفاده شده، به زمان قبل از وقوع زلزله مربوط هستند.

شکل ۵ (الف و ب) مقادیر مشاهدات به دست آمده از داده‌های ماهواره GRACE و مقادیر پیش‌بینی شده را نشان می‌دهد. همان‌طور که در این شکل دیده می‌شود تا قبل از وقوع زلزله الگوریتم بگینگ با دقت قابل قبولی توانسته است رفتار تغییرات تابعک‌های میدان ثقل را پیش‌بینی کند، اما بعد از وقوع زلزله به دلیل تغییر ناگهانی (پرش در سری زمانی) در نمودار و تغییر رفتار تابع، این روش نتوانسته است رفتار تابع را پیش‌بینی کند. البته با توجه به ساختار الگوریتم بگینگ این پیش‌بینی دور از انتظار نیز نیست؛ چرا که این الگوریتم نیز عملکردی مشابه عملکرد درخت تصمیم‌گیری دارد و فقط انحراف معیار خطای پیش‌بینی را در مقایسه با حالت استفاده از یک درخت تصمیم کاهش می‌دهد. بنابراین هنگامی که پرش درخور توجهی در سری زمانی تغییرات میدان ثقل پس از وقوع زلزله رخ می‌دهد، سبب به وجود آمدن بایاس در سری زمانی می‌شود که توسط الگوریتم بگینگ قابل پیش‌بینی نیست. با مقایسه شکل ۵ (الف و ب) (نتایج مربوط به

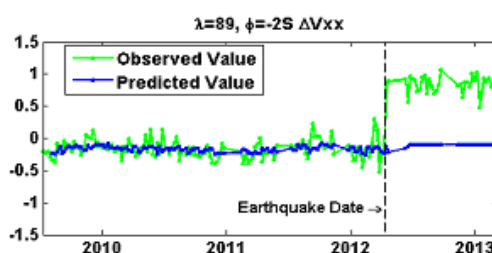
درخت تصمیم) و شکل ۵ (الف و ب) (نتایج بگینگ) نیز این امر به وضوح قابل تشخیص است. همان‌طور که در شکل ۵ (الف و ب) می‌بینیم نوسانات تابع در مقایسه با نتایج درخت تصمیم‌گیری کمتر است، چرا که این نتایج از میانگین نتیجه پیش‌بینی ۱۰۰ درخت تصمیم‌گیری به دست آمده است.

شکل ۵ (ج و د) اختلاف بین مقادیر پیش‌بینی شده و مقادیر مشاهده شده در شکل ۵ (الف و ب) است. همان‌طور که در شکل (ج) دیده می‌شود روش بگینگ توانسته است تغییرات ناگهانی میدان ثقل را سه و پنج هفته قبل از وقوع زلزله کشف کند. در این شکل اختلاف بین مقادیر پیش‌بینی شده و مشاهده شده برای تابعک ΔV_{xx} سه هفته قبل از وقوع زلزله با مقدار ۱۲ درصد از حد آستانه پایین و پنج هفته قبل از وقوع زلزله با مقدار ۲۸ درصد از حد آستانه بالایی عبور کرده است. شایان ذکر است که روش میانه و فاصله بین چارکی و روش درخت تصمیم‌گیری تنها توانسته بودند یک بی‌هنجاری را به ترتیب سه و پنج هفته قبل از وقوع زلزله کشف کنند. در مقابل الگوریتم بگینگ هر دو آنامولی کشف شده توسط این دو روش را شناسایی کرده است.

شکل ۵ (د) نیز مشابه شکل ۵ (ج) اختلاف مقادیر بین مقادیر پیش‌بینی و مشاهده شده تابعک ΔV_{xz} را نشان می‌دهد. همان‌طور که در این شکل دیده می‌شود اختلاف بین مقادیر پیش‌بینی شده و مشاهده شده، پنج هفته قبل از وقوع زلزله با مقدار ۵۰ درصد از حد آستانه پایینی تعریف شده عبور کرده است.

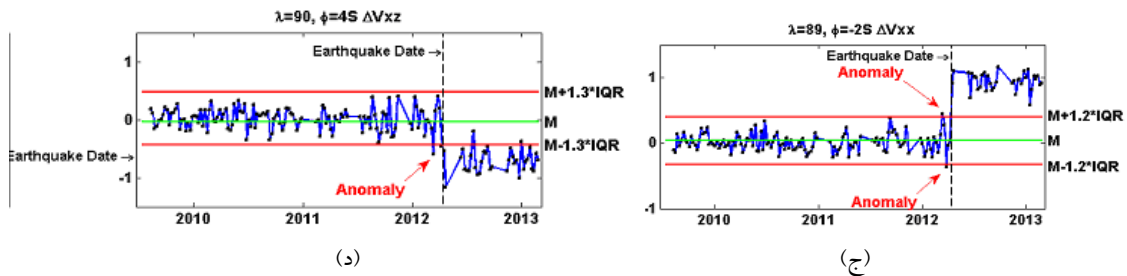


(ب)



(الف)

شکل ۵. تغییرات مشاهده شده (منحنی سبزرنگ) و پیش‌بینی شده (منحنی آبی‌رنگ) مؤلفه‌های ΔV_{xx} (الف) و ΔV_{xz} (ب) با استفاده از الگوریتم بگینگ؛ (ج) و (د) اختلاف بین مقادیر مشاهده شده و پیش‌بینی شده در شکل‌های (الف) و (ب).



ادامه شکل ۵.

پیش‌بینی که استفاده شود توانایی مدل کردن کلیه عوامل مؤثر در تغییرات به وجود آمده در سری زمانی را ندارد، یکسری هشدارهای اشتباه نیز به وجود خواهد آمد؛ در حوزه بررسی پیش‌نشانگرهای زلزله به منظور جلوگیری از نتیجه‌گیری‌های غلط ناشی از این هشدارهای اشتباه، باید از سایر پیش‌نشانگرها نیز به منظور کاهش عدم قطعیت استفاده کرد؛ البته این عدم توانایی در کشف آنامولی در تغییرات سری زمانی ΔV_{xz} به نحوی بیانگر توانایی بالای روش در پیش‌بینی سری زمانی می‌باشد، زیرا الگوریتم بوستینگ در واقع توانسته است رفتارهایی را که از دید ما به عنوان رفتارهای غیرعادی سری زمانی در نظر گرفته شده است، پیش‌بینی کند. بنابراین با توجه به رفتار سری زمانی در هفته‌های قبلی این تغییرات آنامولی نبوده و در صورتی که روش پیش‌بینی روشی قدرتمند باشد باید بتواند آن‌ها را مدله کند. همان‌طور که در سری زمانی دیده می‌شود در تاریخ وقوع زلزله یک پرش ناگهانی ناشی از تغییرات میدان ثقل دیده می‌شود. الگوریتم بوستینگ توانسته است این پرش ناگهانی را به خوبی پیش‌بینی کند که این امر نمایانگر توانمندی بسیار زیاد روش است. در سایر الگوریتم‌های به کار رفته در این پژوهش الگوریتم‌ها در تاریخ وقوع زلزله نتوانسته‌اند به خوبی پیش‌بینی کنند و بعد از چند هفته الگوریتم خودش را با سری زمانی منطبق کرده است.

شکل ۶-د اختلاف بین مقادیر پیش‌بینی شده و مقادیر مشاهده شده در سری زمانی تغییرات ΔV_{xz} را نشان می‌دهد. با توجه به این شکل سه و پنج هفته قبل از وقوع زلزله مقدار ΔV_{xz} از حد آستانه‌های بالا و پایین با مقادیر ۱۰۶٪ و ۹۲٪ عبور کرده است. با توجه به اینکه

۴-۳. روش بوستینگ

در اجرای الگوریتم بوستینگ نیز مشابه دو روش قبلی جهت پیش‌بینی سری زمانی تغییرات میدان ثقل از کل داده‌های موجود در سری زمانی ۶۰ درصد داده‌ها به عنوان داده‌های آموزشی و ۴۰ درصد به عنوان داده تست در نظر گرفته شد. شکل ۶ نتایج پیاده‌سازی این الگوریتم را نشان می‌دهد. لازم به ذکر است که داده‌های آموزشی استفاده شده مربوط به زمان قبل از وقوع زلزله هستند.

شکل ۶ (الف و ب) مقادیر مشاهدات به دست آمده از داده‌های ماهواره GRACE و مقادیر پیش‌بینی شده را نشان می‌دهد. همان‌طور که در این شکل دیده می‌شود بر خلاف دو روش قبلی استفاده شده (درخت تصمیم‌گیری و بگینگ)، الگوریتم بوستینگ پس از وقوع زلزله به خوبی توانسته است رفتار تابع را دنبال کند و پرش ناگهانی در سری زمانی تأثیری در روند پیش‌بینی الگوریتم ندارد. علت این امر این است که بوستینگ در مقایسه با بگینگ علاوه بر کاهش واریانس پیش‌بینی، بایاس را نیز کاهش می‌دهد (فریدام و همکاران، ۲۰۰۱). بنابراین نتایج به دست آمده از این الگوریتم نشان می‌دهد که الگوی بسیار پیچیده موجود در سری زمانی تغییرات تابعک‌های میدان ثقل با استفاده از این روش قابل شناسایی است.

شکل ۶ (ج و د) اختلاف بین مقادیر پیش‌بینی شده و مقادیر مشاهده شده در شکل ۶ (الف و ب) می‌باشد. همان‌طور که در قسمت (ج) دیده می‌شود، الگوریتم بوستینگ نه تنها آنامولی را قبل از وقوع زلزله کشف نکرده است، بلکه به اشتباه آنامولی را در سال ۲۰۱۱ نشان می‌دهد. با توجه به اینکه هر روشی که به منظور

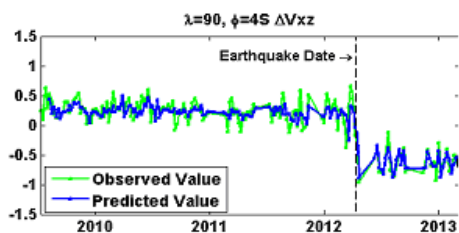
تصادفی در انتخاب ویژگی‌ها برای هر درخت است. با توجه به اینکه در روش جنگل تصادفی از ویژگی‌هایی یکسان در هر درخت استفاده نمی‌شود؛ بنابراین امکان اینکه پاسخ کلی مجموعه، پاسخی مطمئن‌تر و قابل اعتمادتر باشد، بیشتر است. این مطلب خودش را در پیش‌بینی تغییرات میدان ثقل پس از وقوع زلزله نشان می‌دهد.

شکل ۷ (ج و د) اختلاف بین مقادیر پیش‌بینی شده و مقادیر مشاهده شده در شکل ۷ (الف و ب) است. همان‌طور که در قسمت (ج) دیده می‌شود الگوریتم جنگل تصادفی توانسته است سه و پنج هفته قبل از وقوع زلزله آنامولی‌هایی در تغییرات مؤلفه ΔV_{xx} تانسور پتانسیل ثقل کشف کند. مقدار مؤلفه ΔV_{xx} سه هفته قبل از وقوع زلزله با مقدار ۶۸ درصد از حدآستانه پایینی و پنج هفته قبل از وقوع زلزله ۵۶ درصد از حدآستانه بالایی عبور می‌کند. شکل ۷ (قسمت د) اختلاف بین مقادیر پیش‌بینی شده و مشاهده شده برای مؤلفه ΔV_{xz} را نشان می‌دهد. مطابق با نتایج به دست آمده مقدار مؤلفه ΔV_{xz} سه هفته قبل از وقوع زلزله با مقدار ۲۶ درصد از حدآستانه بالایی و پنج هفته قبل از وقوع زلزله با مقدار ۵۰ درصد از حدآستانه پایینی عبور می‌کند.

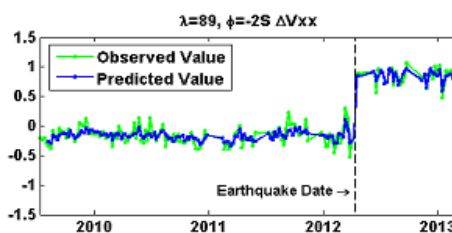
مؤلفه ΔV_{xz} در مقایسه با مؤلفه ΔV_{xx} تغییرات جابه‌جایی توده‌های جرم ناشی از زلزله را بیشتر نمایان می‌کند، دور از انتظار نیست که شاهد رفتارهای غیرعادی قوی‌تری در سری‌های زمانی این مؤلفه باشیم (شهریسوند و همکاران، ۲۰۱۴).

۴-۴. روش جنگل تصادفی

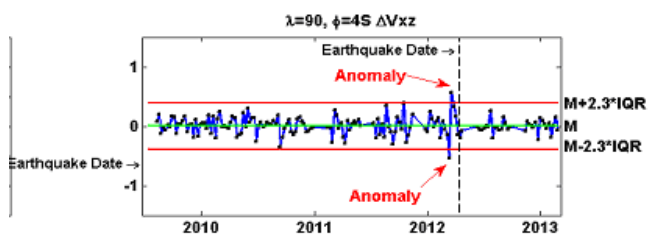
در اجرای الگوریتم جنگل تصادفی نیز مشابه همه روش‌های قبلی جهت پیش‌بینی سری زمانی تغییرات میدان ثقل از کل داده‌های موجود در سری زمانی، ۶۰ درصد به عنوان داده‌های آموزشی و ۴۰ درصد به عنوان داده تست در نظر گرفته شد (شهریسوند و همکاران ۲۰۱۴). شکل ۷ نتایج پیاده‌سازی این الگوریتم را نشان می‌دهد. لازم به ذکر است که داده‌های آموزشی استفاده‌شده مربوط به زمان قبل از وقوع زلزله می‌باشند. شکل ۷ (الف و ب) مقادیر مشاهدات به دست آمده از داده‌های ماهواره GRACE و مقادیر پیش‌بینی شده را نشان می‌دهد. همان‌طور که در این شکل دیده می‌شود الگوریتم جنگل تصادفی توانسته است پس از وقوع زلزله با تأخیر زمانی نسبتاً کوتاهی خودش را با رفتار سیگنال وقف دهد. همان‌طور که در قسمت مواد و روش‌ها گفته شد، تفاوت روش بگینگ و جنگل



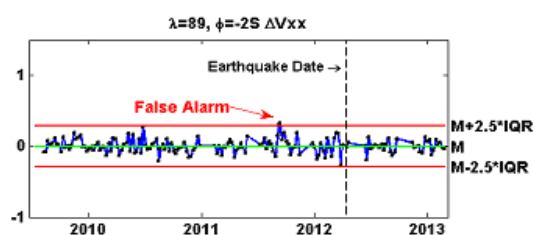
(ب)



(الف)

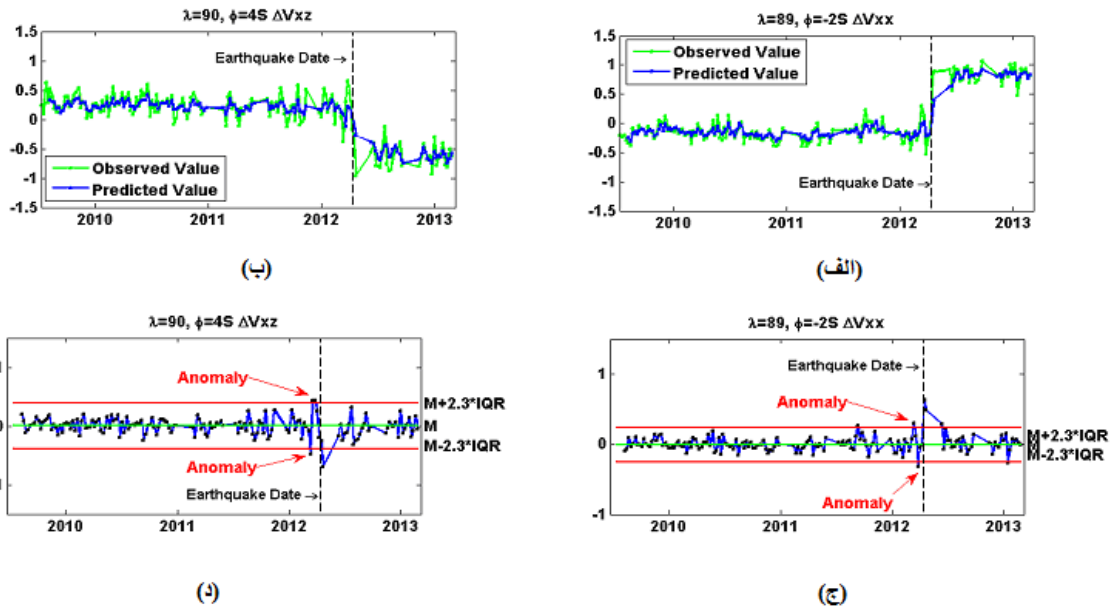


(د)



(ع)

شکل ۶. تغییرات مشاهده شده (منحنی سبزرنگ) و پیش‌بینی شده (منحنی آبی‌رنگ) مؤلفه‌های ΔV_{xx} (الف) و ΔV_{xz} (ب) با استفاده از الگوریتم بوستینگ. (ج) و (د) اختلاف بین مقادیر مشاهده‌شده و پیش‌بینی شده در شکل‌های (الف) و (ب).



شکل ۷. تغییرات مشاهده شده (منحنی سبز رنگ) و پیش‌بینی شده (منحنی آبی رنگ) مؤلفه‌های ΔV_{xx} (الف) و ΔV_{xz} (ب) با استفاده از الگوریتم جنگل تصادفی. (ج) و (د) اختلاف بین مقادیر مشاهده شده و پیش‌بینی شده در شکل‌های (الف) و (ب).

۵. نتیجه‌گیری

می‌رود که بی‌هنجاری‌های میدان ثقل ناشی از زلزله دقیقاً در رومرکز زمین‌لرزه نباشند. زیرا عمده تغییرات توزیع چگالی جرم در اطراف رومرکز زلزله رخ می‌دهد.

به علاوه با توجه به روش‌های مختلف معرفی شده در این مقاله به منظور کشف بی‌هنجاری‌های میدان ثقل، نتایج نشان داد که روش‌های تصمیم‌گیری جمعی در مقایسه با روش‌های میانه و فاصله بین چارکی و درخت تصمیم‌گیری توانایی بیشتری در کشف بی‌هنجاری‌های میدان ثقل دارند.

مراجع

آخوندزاده، م.، ۱۳۹۰، مطالعه تطبیقی پیش‌نشانگرهای زلزله به دست آمده از داده‌های ماهواره‌ای. رساله دکتری، دانشگاه تهران، دانشکده فنی، گروه مهندسی نقشه‌برداری و ژئوماتیک.

خسروی، ا. و آخوندزاده، م.، ۱۳۹۳، استفاده از سیستم‌های طبقه‌بندی چندگانه به منظور بهبود دقت طبقه‌بندی تصاویر تمام پلاریمتریک راداری با فضای ویژگی ابعاد بالا، مهندسی فناوری اطلاعات مکانی،

در این مطالعه پس از نمایش رفتار تغییرات میدان ثقل به دست آمده از داده‌های ماهواره GRACE به صورت سری‌های زمانی طولانی‌مدت، با استفاده از قابلیت‌های روش میانه و فاصله بین چارکی، درخت تصمیم‌گیری، بگینگ، بوستینگ و جنگل تصادفی آنامولی‌هایی در تغییرات میدان ثقل قبل از وقوع زلزله‌های بزرگ کشف شد. از نتایج قابل توجه این تحقیق مشاهده آنامولی به وسیله همه روش‌های استفاده‌شده در این مقاله است. به دلیل حذف تأثیرات هیدرولوژی با استفاده از سرشکنی کمترین مربعات از سری‌های زمانی تغییرات میدان ثقل، بی‌هنجاری‌های کشف شده می‌توانند به عنوان تغییرات غیرعادی میدان ثقل قبل از وقوع زلزله در نظر گرفته شوند.

نتایج این تحقیق نشان می‌دهد که بی‌هنجاری‌های کشف شده قبل از وقوع زلزله برای هر یک از تابعک‌های میدان ثقل می‌تواند مثبت یا منفی باشد. به علاوه مناطقی که بی‌هنجاری‌های تغییرات میدان ثقل قبل از وقوع زلزله در آن‌ها کشف شدند، دقیقاً در رومرکز زلزله نبودند و نسبت به آن یک شیفت مکانی داشتند. با توجه به توزیع مکانی نشان داده شده در شکل ۲ انتظار نیز

- Akhoondzadeh, M., Michel Parrot and Saradjian, M. R., 2010, "Electron and ion density variations before strong earthquakes ($M > 6.0$) using DEMETER and GPS data." *Natural Hazards and Earth System Sciences* 10.1, 7-18.
- Cheng, M. and Tapley, B. D., 2004, Variations in the Earth's oblateness during the past 28 years, *Journal of Geophysical Research*, 109, B09402.
- Chen, J. L., Wilson, C. R., Tapley, B. D. and Grand, S., 2007, GRACE detects coseismic and postseismic deformation from the Sumatra-Andaman earthquake, *Geophysical Research Letters*, 34, L13302.
- Chen, J. L., Wilson, C. R., Tapley, B. D., Yang, Z. L. and Niu, G., 2009, 2005 drought event in the Amazon River basin as measured by GRACE and estimated by climate models, *Journal of Geophysical Research*, 114, B05404.
- De Linage, C., Rivera, L., Hinderer, J., Boy, J. P., Rogister, Y., Lambotte, S. and Biancale, R., 2009, Separation of coseismic and postseismic gravity changes for the 2004 Sumatran earthquake from 4.6 yr of GRACE observations and modelling of the coseismic change by normal mode summation, *Geophysical Journal International*, 176, 695-714.
- Friedman, J., Hastie, T. and Tibshirani, R., 2001, *The elements of statistical learning*, Springer, Berlin: Springer series in statistics, Vol. 1.
- Kusche, J., 2007, Approximatedecorrelationand non-isotropic smoothing of time-variable GRACE-type gravity field models. *Journal of Geodesy*, 81, 733-749.
- Han, S. C., Shum, C. K., Bevis, M., Ji, C. and Kuo, C. Y., 2006, Crustal dilatation observed by GRACE after the 2004 Sumatra Andaman earthquake, *Science*, 313, 658-662.
- Imanishi, Y., Sato, T., Higashi, T., Sun, W. and Okubo, S., 2004, A network of superconducting gravimeters detects submicrogal coseismic gravity changes, *Science*, 306, 476-478.
- Kusche, J., Schmidt, R., Petrovic, S. and Rietbroek, Ro., 2009, Decorrelated GRACE time-variable gravity solutions by GFZ, and their validation using a hydrological model, *Journal of geodesy*, 83, 903-913.
- Liu, J. Y., Chuo, Y. J., Shan, S. J., Tsai, Y. B., Pulinets, S. A. and Yu, S. B., 2004, Pre-earthquake ionospheric anomalies registered by continuous GPS TEC, *Annales Geophysicae*, 22, 1585-1593.
- Schapiro, R. E., 1990, The strength of weak learnability, *Machine learning*, 5, 197-227.
- Seiji, T. and Nakamura, T., 2013, Sea surface gravity changes observed prior to March 11, 2011 Tohoku earthquake, *Physics of the Earth and Planetary Interiors*, 221, 60-65.
- Shrivastava, M. N. and Reddy, C. D., 2013, The M_w 8.6 Indian Ocean earthquake on 11 April 2012: Coseismic displacement, Coulomb stress change and aftershocks pattern, *Journal of the Geological Society of India*, 81, 813-820.
- Shahrisvand, M., Akhoondzadeh, M. and Sharifi, M. A., 2014, Detection of gravity changes before powerful earthquakes in GRACE satellite observations, *Annals of Geophysics*, 57(5), A0543.
- Sun, W. and Okubo, S., 1993, Surface potential and gravity changes due to internal dislocations in a spherical earth—I. Theory for a point dislocation, *Geophysical Journal International*, 114(3), 569-592.
- Wang, L., Shum, C. K. and Christopher J., 2012, Gravitational gradient changes following the 2004 December 26 Sumatra-Andaman Earthquake inferred from GRACE. *Geophysical Journal International* 191.3, 1109-1118.

Detection of gravity-field anomalies associated with great earthquakes in GRACE satellite data using ensemble methods

Shahrisvand, M.¹, Akhoondzadeh, M.^{2*} and Jouybari Moghadam, Y.¹

1. Ph.D. Student, Remote Sensing Department, School of Surveying and Geospatial Engineering, College of Engineering, University of Tehran, Iran
2. Assistant Professor, Remote Sensing Department, School of Surveying and Geospatial Engineering, College of Engineering, University of Tehran, Iran

(Received: 16 Mar 2016, Accepted: 24 Jan 2017)

Summary

In recent years, thousands of people around the world are affected by earthquake. There are many prospects of doing research on earthquake, that the ultimate goal all the researchers want to achieve is the reduction effects caused by this phenomenon. Activities in recent decades in reducing the effects of natural disasters such as earthquake, cause attention on earthquake precursors. Since satellite data have global coverage, suitable temporal resolution and low cost, they are useful for monitoring earthquake precursors. By launching GRACE mission in 2002, the possibility of measuring gravity field variations in weekly temporal resolution is provided. In this paper, 8 years GRACE Level 2 weekly data (have been smoothed by DDK3 filter) have been analyzed in order to detect abnormal gravity field behavior before large earthquakes. We replaced the Earth's oblateness values (C20) with those from Satellite Laser Ranging because of their poor accuracy. We know that GRACE stripe errors elongated in north-south direction, hence these strips generate fluctuations in east-west direction. Therefore by taking x-axis (north direction) derivative the of these, variations are dramatically suppressed. So independence of these components of gravitational gradient tensor to GRACE stripy errors, cause increase signal to noise ratio. By this consideration we used just $\Delta V_{xx}, \Delta V_{xz}$ components of gravitational gradient tensor for anomaly detection. However, we must note that horizontal derivative operator shifts the phase of the original anomaly distribution in spatial domain. So the positions of time series computation of two selected components are different. In addition second derivative of gravitational potential amplify high-frequency components of the earth gravity field and hence the gravitational gradient changes delineate more clearly in the rupture line, revealing refined mass redistribution features caused by the earthquake. In order to suppress seasonal variations and isolate seismic effects, we removed seasonal variations (annual and semiannual and S2 tidal wave) from time series using least squares analysis. The time of earthquakes are excluded in the least squares fit. Since a large part of the deformation is in the ocean, the hydrological model (e.g. GLDAS) cannot be used to remove seasonal variations. By considering fact that other preseismic anomaly (e.g. ionosphere precursors) does not occur in the vertical projection of earthquake epicenter, we test outskirts of each epicenter in order to detect the anomaly. In order to search for earthquake anomaly from time series a reasonable range of gravitational gradient variations must be determined. We used median and Inter-Quartile Range (IQR) of data as the first method for anomaly detection in time series. Afterwards, Bagging, Boosting and Random forest models has been proposed in the detection process of prominent gravity field anomalies prior the earthquakes. Gravity field depends of many parameters such as location, tidal force, oceanic variation, etc. So distribution of gravity field variation time series is not normally used. By consideration this fact we cannot use mean and standard deviation of data for anomaly detection. According to obtained results gravity field anomalies occur within time interval of 2-5 weeks before earthquakes. The results in this study indicate that in each case study, the unusual variations of gravity field have had different sign but the signs of two selected components of gravitational gradient tensor for each case study are the same.

Keywords: Gravity, GRACE, earthquake precursor, Bagging, Boosting, Random Forest, anomaly detection.