



ارائه روش برازش فرآیند گوسی در شناسایی گسل در مقاطع لرزه‌نگاری

مریم نوری^۱؛ حسین حسینی^{۲*}؛ عبدالرحیم جواهریان^۳؛ سید علی معلمی^۴

۱. دانشجوی دکتری؛ دانشکده مهندسی نفت، دانشگاه صنعتی امیرکبیر

۲. دانشیار؛ دانشکده مهندسی معدن و متالورژی، دانشگاه صنعتی امیرکبیر

۳. استاد بازنشسته موسسه ژئوفیزیک دانشگاه تهران و استاد دانشکده مهندسی نفت، دانشگاه صنعتی امیرکبیر

۴. استادیار، پژوهشکده ازدیاد برداشت از مخازن نفت و گاز، شرکت ملی نفت ایران

دریافت مقاله: ۱۳۹۸/۰۱/۰۹ پذیرش مقاله: ۱۳۹۸/۰۳/۱۹

DOI:10.22107/jpg.2019.177272.1088

واژگان کلیدی	چکیده
گسل، فرآیند گوسی، ناهنجاری عمومی، تعامل نرمال، بارز شدن محل گسل خوردگی	برازش فرآیند گوسی، به عنوان یک مدل احتمالاتی غیر پارامتری مبتنی بر آمار بیزین، در مطالعات ویژگی‌های پراکنده مانند ناهنجاری‌های عمومی بسیار توانمند است. توانمندی بارز این فرآیند، شناسایی رفتارهای ناهنجار از رفتارهای هنجار است. می‌توان این فرآیند را به عنوان یک روش در شناسایی لبه‌های گسلی در اطلاعات لرزه‌نگاری معرفی کرد. در این مطالعه، شناسایی ناهنجاری مبتنی بر برازش فرآیند گوسی بر روی داده‌های لرزه‌های مصنوعی و داده‌های واقعی به کار برده شده است. به منظور شناسایی لبه‌های گسلی، لایه‌های زمین‌شناسی به عنوان تعامل‌های نرمال در نظر گرفته شده‌اند. لبه‌های گسلی به عنوان ناهنجاری‌های عمومی در نظر گرفته شده‌اند که منجر به برهم ریختگی رفتار نرمال بازتابنده‌های لرزه‌ای می‌شود. در نهایت خطای برازش برای جدا کردن گسل‌ها از سایر ناهنجاری‌ها استفاده شده است. جهت بررسی عملکرد روش برازش فرآیند گوسی، این روش بر روی یکسری مجموعه داده مصنوعی و یک مقطع دو بعدی از بلوک $F3$ بخش هلندی دریای شمال حاوی گسل اعمال شده است. نتایج نشان‌دهنده توانمندی این روش در بارز کردن محل گسل است.

۱. پیش‌گفتار

بر مبنای نشانگرهای لرزه‌ای و شبکه‌های عصبی مانند روش ردیابی مورچه و روش بردار ردیابی گسترش داده شده‌اند. این تحقیقات برای ردیابی خودکار گسل بدون دخالت انسان توسعه داده شده‌اند. این روش‌ها در مرحله آغازین نیازمند دخالت انسان برای نشان کردن محل گسل خوردگی هستند. تاکنون روش‌های بسیاری برای شناسایی گسل ارائه شده است. اما حضور یک نیروی انسانی تقریباً در همه‌ی آن‌ها امری ضروری است. در پیشینه پژوهش مربوط به شناسایی گسل روش‌های خودکار وجود دارند، اما هنوز کاملاً خودکار نیستند و در بخش‌هایی به حضور انسان نیاز دارند. پیشینه‌ی تحقیقات در راستای شناسایی گسل به بیش از ۴ دهه‌ی قبل باز می‌گردد. به طور متداول در تفسیر مقاطع لرزه‌نگاری، گسل‌ها با استفاده از نشانگرهای لرزه‌ای مانند تشابه (Farmer & Bahorich, 1995)، همدوسی (Marfurt et al., 1998)،

شناسایی و بخش‌بندی مدل ساختمانی مخازن هیدروکربوری به شدت می‌تواند تحت تأثیر گسل باشد. اثر شکستگی‌های اطراف گسل در جریان سیال هیدروکربونی، عملکرد گسل در تغییر سطوح تماس، و همچنین عملکرد آن در ایجاد کردن تله‌های هیدروکربونی، شناسایی و استخراج گسل را در حجم داده‌های لرزه‌نگاری به امری لازم در علوم مهندسی نفت تبدیل کرده است. پدیده‌ی گسل خوردگی، یک پدیده‌ی مولد ناهنجاری معمولاً شدیدی است که سبب می‌شود سیر هموار لایه‌های زمین‌شناسی برهم بخورد. تفسیر و شناسایی گسل به صورت دستی فرآیندی زمان‌بر و خسته‌کننده است به خصوص زمانی که حجم عظیمی از داده‌های لرزه‌نگاری وجود داشته باشد. هر چند تنها تکیه بر نیروی انسانی خیلی کارآمد نیست؛ در سال‌های اخیر برخی روش‌های شناسایی گسل

واریناس، شیب، سمت و انحنا (Roberts, 2001) شناسایی می‌شوند. به این دسته از نشانگرهای لرزه‌ای نشانگرهای هندسی متداول می‌گویند زیرا اغلب ماهیت هندسی انعکاس-های لرزه‌ای را تعیین می‌کنند. نشانگرهای لرزه‌ای تنها به نشانگرهای هندسی محدود نمی‌شوند. دسته‌ای دیگر از آن‌ها نشانگرهای غیرمتداول نامیده می‌شوند که در حال گسترش هستند. این نشانگرها عبارت‌اند از: مجاورت شکاف، تراکم شکستگی، گرادیان شکستگی، سمت شکستگی، نشانگرهای جهتی، شدت وضوح و شدوارگی گسل (Jaglan et al., 2015). در برخی مطالعات پیشین جهت شناسایی گسل از مفهوم محیط‌های ناهمسانگرد (Rueger, 1997) استفاده شده است. از جمله نشانگرهای توسعه یافته بر مبنای این مفهوم می‌توان به تغییرات سرعت با آزمون، تغییرات دامنه با آزمون (Chopra et al., 2000; Zheng & Gray, 2002)، تغییرات زاویه در برابر دامنه (Aleardi & Mazzotti, 2014) و تغییرات فاکتور کیفیت میرایی با آزمون (Wang et al., 2015) اشاره کرد.

در این مقاله سعی بر آن شده است تا با دیدگاه جدیدی به شناسایی محل گسل خوردگی در داده‌های لرزه‌نگاری بپردازیم که اصول کلیدی آن مبتنی بر شناسایی نقاط ناهنجاری عمومی^۱ در پردازش فیلم است. در داده کلوی، تشخیص ناهنجاری به شناسایی اقلام یا حوادثی اشاره شده- است که با یک الگوی مورد انتظار یا با سایر اقلام موجود در یک مجموعه داده مطابقت ندارد. به عبارتی ناهنجاری‌ها رفتار غیر نرمالی نسبت به الگوهای نرمال موجود در مجموعه داده دارند (Chandola, 2009). شناسایی ناهنجاری‌های عمومی کاربرد متفاوتی می‌تواند داشته باشد. از جمله این کاربردها می‌توان به شناسایی داده‌های پرت (Grubbs, 1969; Basu et al., 2007)، اشیا متحرک عظیم^۲ (Li et al., 2006)، ترافیک و حمل و نقل دریایی (Lane et al., 2010); شناسایی ناهنجاری در مسیر کشتی (Mascaro et al., 2014)، ناهنجاری‌های ویدئویی و ردیابی مسیر ترافیک (Cheng et al., 2015; Liao et al., 2016) اشاره کرد.

در این مطالعه گسل به عنوان یک ناهنجاری عمومی در نظر گرفته شده است که تعامل نرمال^۵ لایه‌های زمین‌شناسی را برهم می‌زند. محل این ناهنجاری‌های پراکنده با استفاده از برازش فرآیند گوسی (GPR^۶) به دست آمده است. روش GPR فرآیندی کاملاً غیر پارامتری بر مبنای تئوری بیزین است که با بزرگ شدن داده، تعداد پارامترهایش تغییر نمی‌کند. این روش ابزار قدرتمندی برای داده‌های دارای نوفه است (Cheng et al., 2015). در این مطالعه، محل گسل به طور خودکار به صورت یکسری نقاط حاصل از خطای برازش فرآیند گوسی در تصویر حاصل از اعمال GPR ظاهر می‌شود. نیازی نیست مفسر مانند سایر روش‌های لبه‌یابی (مانند واریناس) گسل‌های ظاهر شده را از لبه‌هایی حاصل از مرز واحدهای سنگی جدا نماید. زیرا لبه‌های مربوط به واحدهای سنگی (بازتابنده‌های لرزه‌ای) در روش GPR آشکار نمی‌شوند. اعمال پس‌پردازش مورفولوژی در این مقاله به کیفیت تصویر حاصل از روش GPR کمک می‌کند و خروجی را به صورت صفر و یک باینری درمی‌آورد. این خروجی باینری، ورودی واضح و شفاف‌تری نسبت به سایر روش‌ها برای استخراج سطوح گسل است.

¹ Global anomaly

² Massive moving objects

³ Principal component analysis (PCA)

⁴ Support vector machine (SVM)

⁵ Normal interaction

⁶ Gaussian process regression

۲. تئوری و روش تحقیق

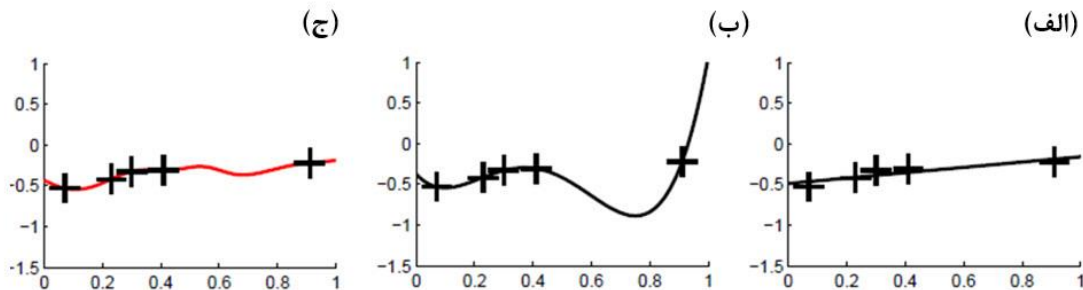
۱.۲ برازش فرآیند گوسی

به صورت متداول یک مدل برازش خطی به صورت رابطه (۱) است. که در آن ε میزان خطا است که به صورت یک توزیع نرمال است که مطابق رابطه (۲) بیان می‌شود. σ^2 پراش (واریانس) خطا و β مقادیر ضرایب هستند که از روی داده به دست می‌آید. این نوع برازش‌ها به عنوان برازش‌های پارامتری معروف هستند. در این روش‌ها تعداد پارامترهای مدل از قبل مشخص می‌شود. الگوریتم‌های بهینه‌ی متنوعی برای این روش‌ها ارائه گردیده‌اند. برخلاف سادگی و سرعت بیشتر این روش‌ها نسبت به روش‌های دسته‌ی دیگر دو اتفاق در استفاده

از این روش‌ها ضعف اصلی آن‌هاست. اولین اتفاق کم‌برازش^۱ است که در صورت ناکافی بودن تعداد پارامترهای مدل به وجود می‌آید (شکل الف) و در آن مدل از توصیف داده عاجز می‌ماند. حالت دوم بیش‌برازش^۲ است که در این حالت مدل به حدی پیچیده است که داده‌ی آموزش را به‌طور کامل توصیف می‌کند اما بر روی داده‌ی جدید خطای بسیار زیادی از خود نشان می‌دهد. علت این موضوع تلاش نابجای مدل برای توصیف نوفه ذاتی موجود در داده است (شکل ب). (ب).

$$y = \mathbf{x}^T \beta + \varepsilon \quad (1)$$

$$\varepsilon = N(0, \sigma^2). \quad (2)$$



شکل ۱. مدل‌های برازش تابع بر روی داده مشاهده شده. (الف) برازش ساده (کم‌برازش). (ب) برازش پیچیده (بیش‌برازش). (ج) مدل واقعیت.

و هر داده آموزشی زیرمجموعه‌ی محدودی از این ابعاد پارامتری در نظر گرفته می‌شود. این کار با حاشیه‌گیری روی ابعاد استفاده نشده انجام می‌شود.

داده‌ی مجموعه آموزشی $\{(\mathbf{x}_i, \mathbf{y}_i); i = 1, 2, \dots, n\}$ را در نظر بگیرد، که در آن $\mathbf{x}_i \in R^d$ و $\mathbf{y}_i \in R$ از یک توزیع نامشخص حاصل شده‌اند. مدل GPR با در دست داشتن بردار جدید \mathbf{x}_{new} ورودی و داده‌ی آموزشی به سؤال مربوط به مقادیر پیش‌بینی یک متغیر پاسخ \mathbf{y}_{new} جواب می‌دهد (شکل ۳). در توصیف مشاهداتی که فرض می‌شود از یک مجموعه متغیرهای تصادفی انتخاب شده‌اند، برازش فرآیند گوسی بر روی مجموعه داده‌های مشاهده‌ای به‌صورت رابطه

در مقابل سایر برازش‌ها مدل‌های GPR (شکل ۲) مدل‌های احتمالاتی مبتنی بر کرنل غیرپارامتری هستند. به این معنی که، از خود داده‌ی آموزش برای انتخاب مدل کمک گرفته می‌شود؛ بدین‌صورت که تعدادی مدل ساده پایه در نظر گرفته شده و هر بار در صورت ناتوان بودن مدل کلی برای توصیف داده‌ی ورودی جدید، یک مدل ساده‌ی دیگر برای کمک به مدل موجود اضافه می‌شود. بدین ترتیب مدل‌ها تنها در صورت نیاز پیچیده می‌شوند. این کار موجب می‌شود که بدون نیاز به پردازش‌های اضافه، از بیش‌برازش جلوگیری شود. در حقیقت مدل غیر پارامتری بی‌زی یک مدل پارامتری بی‌زی است که فضای پارامتری آن بی‌نهایت بُعدی است ولی برای هر مسئله

¹ Under fitting

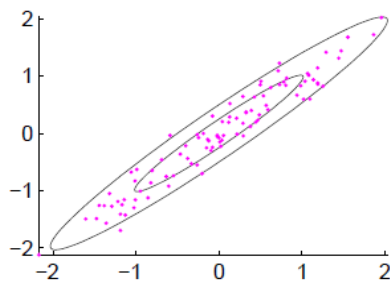
² Over fitting

مشاهده‌شده γ متناظر کند. فرض کنید که از پیش می‌دانیم مقدار میانگین مشاهداتمان (γ) از تابع میانگین صفر تبعیت می‌کند و تابع کوواریانس آن‌ها از تابع مربع نمایی^۲ (رابطه (۴)) تبعیت می‌کند. توزیع پیش‌بینی شده از نقاط ورودی دارای یک توزیع چندمتغیره گوسی است که به صورت رابطه (۵) نمایش داده می‌شود.

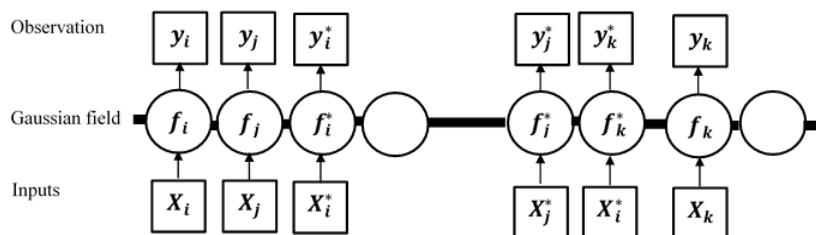
$$k(\mathbf{x}, \mathbf{x}') = \sigma_f^2 \exp\left(-\frac{1}{2l}(\mathbf{x} - \mathbf{x}')^2\right). \quad (4)$$

$$f_{\mathbf{x}}(\mathbf{x}) = \frac{\exp\left(-\frac{1}{2}(\mathbf{x} - \boldsymbol{\mu})\boldsymbol{\Sigma}^{-1}(\mathbf{x} - \boldsymbol{\mu})\right)}{\sqrt{(2\pi)^k |\boldsymbol{\Sigma}|}}. \quad (5)$$

مشاهده‌می‌شود که کوواریانس خروجی‌ها بر اساس تابعی از ورودی‌ها بیان می‌شود. در تابع کوواریانس $(\mathbf{x} - \mathbf{x}')$ ، فاصله اقلیدسی بین دو بردار \mathbf{x} و \mathbf{x}' است. به σ_f و l فرامؤلفه‌های مدل گفته می‌شود. درنهایت بعد از آموزش مدل GP بر اساس اطلاعات و مشاهدات (داده‌های لرنه‌نگاری)، داده‌ی آموزشی و فرامؤلفه‌های مدل به دست می‌آیند.



شکل ۲. نمونه یک مدل برازش گوسی چندمتغیره بر روی یک مجموعه داده مشاهده‌شده.



شکل ۳. مدل گرافیکی (گراف زنجیره‌ای) برای برازش فرآیند گوسی. مربع نشان دهنده‌ی متغیرهای مشاهده‌شده و دایره‌ها توابع و روابط مجهول هستند. خط افقی ضخیم نشان دهنده مجموعه‌ای از گره‌های کاملاً متصل هستند. متغیرهایی ستاره‌دار داده‌های GPR را برای مرحله استنتاج نشان می‌دهند، و اجزای غیر ستاره‌دار متغیرها در مرحله یادگیری هستند (Cheng et al., 2015; Rasmussen & Williams, 2006).

¹ Hyper parameters

² Squared exponential

۱.۱.۲ GPR و شناسایی ناهنجاری عمومی

موضوع مهم در شناسایی ناهنجاری‌ها شناسایی نوعی از ناهنجاری است که تأثیر عمومی بر کل داده داشته باشد و تعامل‌های نرمال کل مجموعه را برهم زند. به عبارتی رفتار غیر نرمالی در رخداد‌های نرمال ایجاد کند درست مانند گسل-ها در داده‌های لرزه‌نگاری. برخلاف ناهنجاری‌های محلی^۱ که با استفاده از روابط هندسی ساده به دست می‌آیند، تفسیر ناهنجاری‌های عمومی بسیار دشوار است.

۲.۲ شناسایی محل گسل با GPR

استفاده از فرآیند گوسی برای تشخیص ناهنجاری عمومی پیش‌تر در شماری از پژوهش‌ها بررسی شده‌اند. برای مثال اسمیت و همکاران (Smith et al., 2014) برای توصیف ترافیک دریایی، لایو و همکاران (Liao et al., 2015) برای مدل کردن فعالیت‌های^۲ ناهنجر در یک مجموعه فعالیت و چنگ و همکاران (Cheng et al., 2015) برای توصیف ناهنجاری در ویدئویی‌های مربوط به ترافیک و تشخیص تصادف و یا سایر ناهنجاری که در سطح وسیع‌تری اثر می‌گذارد. مراحل کلی این روش در شکل ۴ نشان داده شده است. پایه‌ی کار در این روش‌ها که ایده اولیه در این مقاله را تشکیل داده است به صورت زیر است:

۱. نقاط ناهنجاری محلی به دست می‌آید (شکل ۴(الف))، در این مرحله هر نقطه‌ی احتمالی در نظر

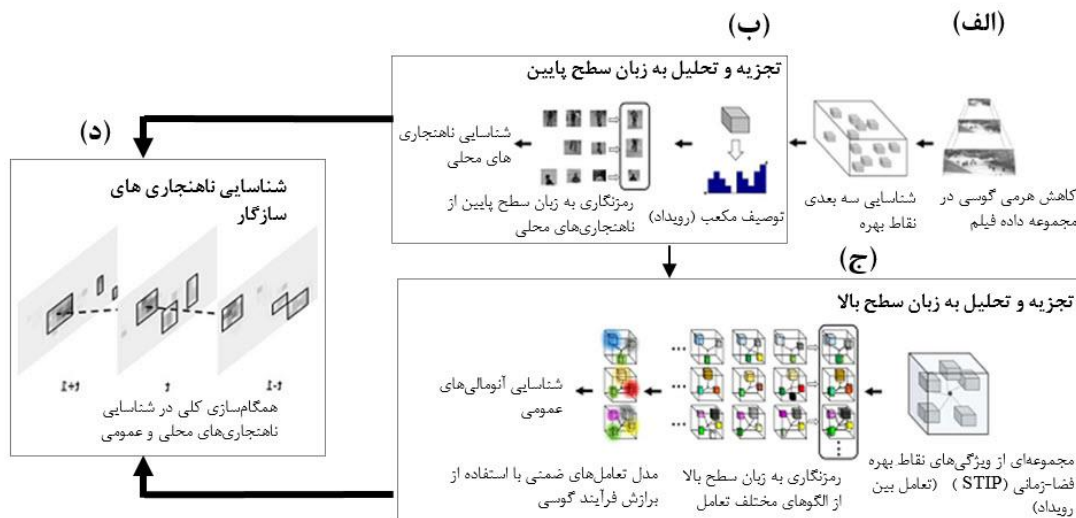
- گرفته می‌شود. برای مثال اگر برای یافتن ناهنجاری از لبه‌یابی^۳ استفاده می‌شود، تمامی لبه‌های ضعیف را نیز در نظر می‌گیرند.
۲. با استفاده از یک الگوریتم خوشه‌بندی^۴ مراکز هر ناهنجاری را به دست آورده و به‌عنوان نماینده‌ی کل ناهنجاری در نظر گرفته می‌شود. برای مثال در تصویر ۴(ب) نقاط نامزد ناهنجاری نمایش داده شده است.
۳. رابطه‌ی موجود بین ناهنجاری‌های نماینده با یک مدل فرآیند گوسی توصیف می‌شود.
۴. نقاطی که در اثر پیش‌بینی مدل با این فرآیند گوسی خطای زیادی از خود نشان می‌دهند به عنوان نقاط ناهنجاری عمومی در نظر گرفته می‌شود. به عنوان مثال در شکل ۵(ب) نقاط به رنگ قرمز نمایانگر نقاط ناهنجاری عمومی می‌باشند.
۵. حال نقاط ناهنجاری محلی موجود در خوشه‌ی نقاط ناهنجاری عمومی را در نظر گرفته و از بقیه نقاط صرف نظر می‌شود. به شکل ۵(ج) رجوع شود
۶. با مقداری پس‌پردازش ساده مثل اعمال مورفولوژی یا سایر فرایندهای بهبود تصویر محل نقاط ناهنجاری را به‌طور دقیق مشخص می‌شود (شکل ۴(د) و شکل ۵(ج)).

¹ Local anomaly

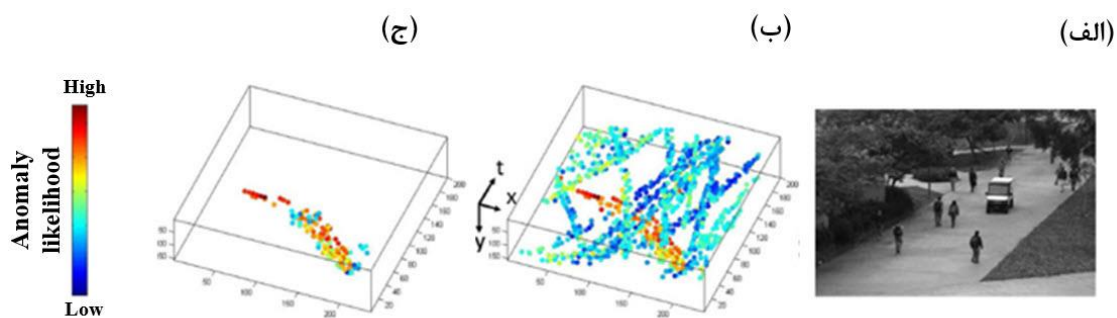
² Activity

³ Edge Detection

⁴ Clustering



شکل ۴. چارچوب کلی روش پیشنهادی برای تشخیص ناهنجاری در ویدئو (Cheng et al., 2015).



شکل ۵. استخراج ناهنجاری عمومی در فیلم کنترل ترافیک. (الف) عبور ماشین حمل‌ونقل از پیاده‌رو. (ب) استخراج ناهنجاری‌های ممکن در فیلم با استفاده از آموزش ماشین توسط فرآیند گوسی. (ج) استخراج محل ناهنجاری عمومی (محل عبور ماشین حمل و نقل در پیاده‌رو) در فیلم کنترل ترافیک (Cheng et al., 2015).

۲.۲.۱ محل گسل خوردگی به عنوان ناهنجاری عمومی

شده معمولاً رفتاری هموار و نرمالی در جهت افق نشان می‌دهند، اگرچه با توجه به عوامل محیطی و تصادفی مقداری از سیر خطی خود خارج شده و حالتی هموار ولی غیرخطی به خود می‌گیرند. در صورت بروز تنش‌های تکتونیکی در منطقه و ایجاد گسل، این سیر هموار برهم می‌خورد. در واقع گسل خوردگی یک پدیده‌ی مولد ناهنجاری معمولاً شدیدی است که سبب می‌شود سیر هموار لایه‌های زمین‌شناسی برهم بخورد. با توجه به مطالب بیان شده در این مطالعه فرض شده است که هر لایه‌ی زمین‌شناسی که در آن یک پدیده‌ی مولد

لایه‌های مختلف زمین‌شناسی معمولاً به صورت هموار در زیرزمین قرار گرفته‌اند. هر لایه با توجه به تاریخچه تدفین و تحولات زمین‌شناسی که طی کرده است خصوصیات متفاوتی را از خود بروز می‌دهد. علت این موضوع این است که میزان فشار، دما و نیز موجودات زنده و مواد شیمیایی در عمق‌های مختلف حوضه رسوبی متغیر بوده است. در صورتی که شرایط شیمیایی و فیزیکی محیط رسوب‌گذاری در طی نهشته شدن رسوبات تغییرات زیادی نداشته باشد و منطقه دچار تنش‌های تکتونیکی شدیدی نشده باشد، لایه‌های زمین‌شناسی نهشته

بازتاب لرزه‌ای) لایه‌های زمین‌شناسی خواهد بود. این متغیر تصادفی را می‌توان به صورت یک فرآیند گوسی بر روی مختصات ورودی مدل کرد. در این مقاله با استفاده از کتابخانه GPML¹ این فرآیند گوسی برای داده‌ی ورودی یادگرفته شده شده است. در تعریف فرآیند گوسی برای داده‌های لرزه‌نگاری

$X = \{ \mathbf{x}_i \in \mathbb{R}^3 \mid i=1..n \}$ مجموعه موقعیت فضایی نقاط ورودی، و Y را معادل مقدار دامنه‌ی لرزه‌ای (شدت رنگ) در آن نقاط در نظر بگیرید. همانگونه که پیش‌تر گفته شد فرآیند گوسی با تابع میانگین، تابع کوواریانس و یک مجموعه از فرموله تعریف می‌شود (Rasmussen & Williams, 2006). فرض شده است که مقدار میانگین مشاهداتمان (Y) (در اینجا دامنه لرزه‌ای) از تابع میانگین صفر و تابع کوواریانس آن‌ها از تابع مربع نمایی (رابطه (۴)) تبعیت می‌کند. فرض میانگین صفر برای داده‌های لرزه‌ای به دور از واقع نیست چون اغلب دامنه این داده‌ها بین \pm یک مقداری قرار می‌گیرند. پس از یادگیری مدل برازش فرآیند گوسی بر روی داده‌های لرزه‌نگاری، با آنالیز خطای حاصل از مدل برازش و داده‌های مشاهده شده، محل گسل خوردگی بدست می‌آید. یعنی در مکان‌هایی که فرآیند گوسی به خوبی نمی‌تواند داده را توصیف نماید در آنجا احتمال وجود ناهنجاری عمومی (گسل) وجود دارد. محل‌هایی با خطای بالا (شکل ۶(ب) بیضی مشکی رنگ) به عنوان نقاط گسل شناسایی و بارز می‌شوند. به عبارتی پس از یادگیری فرآیند گوسی بر روی داده لرزه‌ای، میزان احتمال شدت رنگ در داده توسط فرآیند به دست می‌آید.

میزان احتمال شدت رنگ که بر مبنای مدل فرآیند گوسی فراگرفته شده است، در نقاط گسلی کم خواهد بود، و فرآیند گوسی در این نقاط بیشتر دچار خطای پیش‌بینی درست می‌شود. الگوریتم یادگیری فرآیند گوسی و شناسایی نقاط گسلی با استفاده از روش پیشنهادی در جدول (۱) آمده است. در انتها با استفاده از الگوریتم‌های اتصال نقاط (همچون تبدیل هاف و یا اعمال مورفولوژی) که در پردازش تصاویر بسیار مرسوم هستند (Dougherty, 1992) نقاط گسلی نزدیک را به هم متصل ساخته و ساختار متمرکزی از نقاط گسل قابل استخراج است.

ناهنجاری رخ نداده باشد، به صورت یک تابع هموار قابل مدل‌سازی است و هر عاملی که این سیر هموار را برهم زند به عنوان یک ناهنجاری عمومی می‌تواند قابل ردیابی باشد. جهت توصیف فرآیندهای هموار و نرمال معمولاً از توزیع گوسی یا یکی از تعمیم‌های آن مانند گوسی چند بُعدی، و یا مدل‌های ترکیبی گوسی (فرآیند گوسی) به کار برده می‌شود. علت این موضوع علاوه بر پیروی بیش‌تر پدیده‌های طبیعی از فرآیند گوسی، در نظر گرفتن کمترین اطلاعات آماری اضافه در مورد داده است، زیرا توزیع گوسی شامل کمترین میزان اطلاعات در بین توزیع‌های با میانگین و واریانس ثابت است (Bishop, 2006). با توجه به موارد پیش‌گفته، استفاده از فرآیند گوسی برای توصیف لایه‌های زمین‌شناسی مناسب به نظر می‌رسد. در این مطالعه تلاش گردیده لایه‌های زمین‌شناسی با فرآیند گوسی مدل شوند تا بدین صورت بتوانیم ردیابی محل ناهنجاری عمومی (محل گسل خوردگی) در داده‌ی ژئوفیزیک با استفاده از خطای حاصل از برازش فرآیند گوسی توصیف شود. با توجه به تاریخچه تدفین لایه در شرایط متداول می‌توان گفت شدت رنگ در هر بازتاب (دامنه بازتابی حاصل از بازتاب موج از سطح لایه‌های زمین‌شناسی) حول میانگین آن لایه به صورت یک توزیع گوسی پخش شده است.

از آنجا که تنها میانگین شدت رنگ در هر بازتابنده برای ما حائز اهمیت است، به همین دلیل در مرحله‌ی اول با استفاده از فیلتر میانگین، شدت رنگ هر نقطه را با میانگین یک همسایگی حول آن جایگزین می‌نماییم. از آنجا که در بین عبور بین لایه‌ها این میانگین‌گیری به خراب شدن اطلاعات می‌انجامد؛ به همین دلیل برای کمینه کردن میزان از دست رفتن اطلاعات، از یک پنجره افقی برای میانگین‌گیری استفاده می‌نماییم و از طرفی ورودی را همراه با یک نوفه گوسی در نظر می‌گیریم. بدین صورت تفاوت شدت رنگ در یک بازتابنده بر حسب نوفه ورودی گذاشته می‌شود و در نتیجه با کوچک گرفتن اندازه‌ی پنجره‌ی میانگین و در نظر گرفتن نوفه می‌توان از تفاوت‌های شدت رنگ در یک لایه صرف‌نظر نمود. بدین منظور در داده‌ی دوبعدی راستای میانگین‌گیری در راستای Strike line می‌باشد. چیزی که مکعب لرزه‌ای پس از فیلتر شدن نمایش می‌دهد، تقریبی از متغیر تصادفی (دامنه

¹ Gaussian Process for Machine Learning

جدول ۱. الگوریتم یادگیری فرایند گوسی و شناسایی نقاط گسلی با استفاده از روش پیشنهادی.

الگوریتم بارز کردن محل گسل با استفاده از GPR

این الگوریتم روند یادگیری و آشکارسازی نقاط گسلی را با استفاده از مدل پیشنهادی بیان می‌دارد.

ورودی:

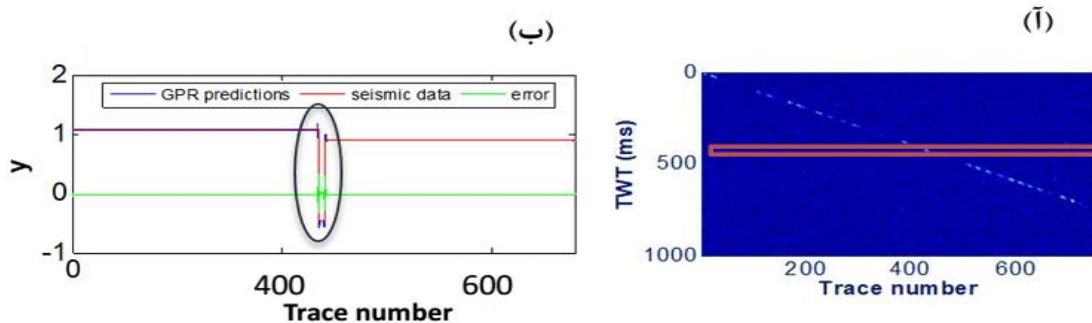
داده لرزه‌ای

خروجی:

نقاط گسلی محتمل در داده لرزه‌ای

مراحل الگوریتم:

۱. ابتدا یک فیلتر میانگین برای هموارتر نمودن دامنه‌ها و حذف برخی نوفه‌های پراکنده احتمالی استفاده کن.
۲. حال بر روی مکعب فیلتر شده یک فرایند گوسی با ورودی مختصات مکعب و خروجی شدت رنگ یاد بگیر.
۳. اینک بر اساس فرایند گوسی یادگرفته شده احتمال شدت رنگ را برای هر مختصات حساب کن.
۴. در نقاطی که احتمال فرایند گوسی کمتر از حد یک آستانه است را به عنوان نقطه احتمالی گسل با احتمال مکمل احتمال حاصل از فرایند گوسی اعلام کن.
۵. از نقاط گسلی با بیشترین احتمال گسل بودن شروع کن و نقاط گسلی همسایه را با در نظر گرفتن میزان احتمال طوری به هم متصل کن که نقاط گسلی یکپارچه‌تری حاصل شود.



شکل ۶. (الف) خطای حاصل از نتایج GPR بر روی مقطع لرزه‌نگاری مصنوعی. مستطیل قرمز رنگ محدوده‌ی انتخابی در زمان ۴۷۵ میلی‌ثانیه جهت بررسی میزان خطای GPR نشان می‌دهد. (ب) بررسی میزان خطای حاصل شده پس از پیش‌بینی مدل مصنوعی توسط فرآیند گوسی یاد گرفته شده بر روی داده مصنوعی. بیضی مشکی رنگ محل شدت خطا (محل خطای حاصل از برازش فرآیند گوسی) در نزدیکی گسل را نشان می‌دهد.

۳. مجموعه داده‌ها

در راستای بررسی روش پیشنهاد شده دو مجموعه داده‌های مصنوعی و واقعی انتخاب شدند. داده‌ی مصنوعی و واقعی حاوی گسل به تشریح در ادامه آورده شده است. در مورد داده‌های واقعی نیاز به داده‌ی پیش‌پردازش شده به صورت داده‌ی مهاجرت شده پس از برآینارش سه بعدی است. در بخش‌های ۱.۳ و ۲.۳ به معرفی و نتیجه حاصل از اعمال روش GPR بر روی داده‌های مصنوعی و واقعی استفاده شده می‌پردازیم.

۱.۳ داده مصنوعی

در راستای ارزیابی روش ارائه شده مجموعه داده مصنوعی لرزه‌ای حاوی سطوح بازتابی و گسل‌های نرمال ساخته شده است. در این مجموعه داده مصنوعی بازتابنده‌ها به صورت افقی در نظر گرفته شده‌اند. ضریب بازتاب لایه‌ها به صورت یک سری نرمال تصادفی در نظر گرفته شده است (Robinson, 1985). که با موجک ریگر (فاز صفر) هم‌آمیخت شده است تا دامنه بازتاب حاصل گردد. در نهایت بازتابنده‌ها توسط گسل‌های نرمال با

گرفته شدند که توسط یک گسل قاشقی جایجا شده‌اند. سپس مجموعه‌ای از بازتابنده افقی بر روی آن‌ها قرار داده شد تا بتوان سطح ناپیوستگی را ایجاد نمود. در نهایت مقطعی با سطح نوفه متفاوت با استفاده از رابطه (۶) و (۷) ساخته شد، به گونه‌ای که نسبت $SNR = 2, 4, 8$ ایجاد گردد.

شکل ۸(الف) این مقاطع مصنوعی دوبعدی را نشان می‌دهد که به ترتیب از راست به چپ دارای نسبت سیگنال به نوفه ۸، ۴ و ۲ می‌باشند. مشاهده می‌شود که الگوریتم پیشنهاد شده به خوبی توانسته است در حضور سطوح مختلف نوفه محل گسل را بارز نماید. همانگونه که در مطالعات پیشین شناسایی ناهنجاری عمومی بیان شده است (Cheng et al., 2015) این روش ابزار قدرتمندی در برابر حضور نوفه است. سطح ناپیوستگی در این مطالعه توسط روش GPR بارز نشده است.

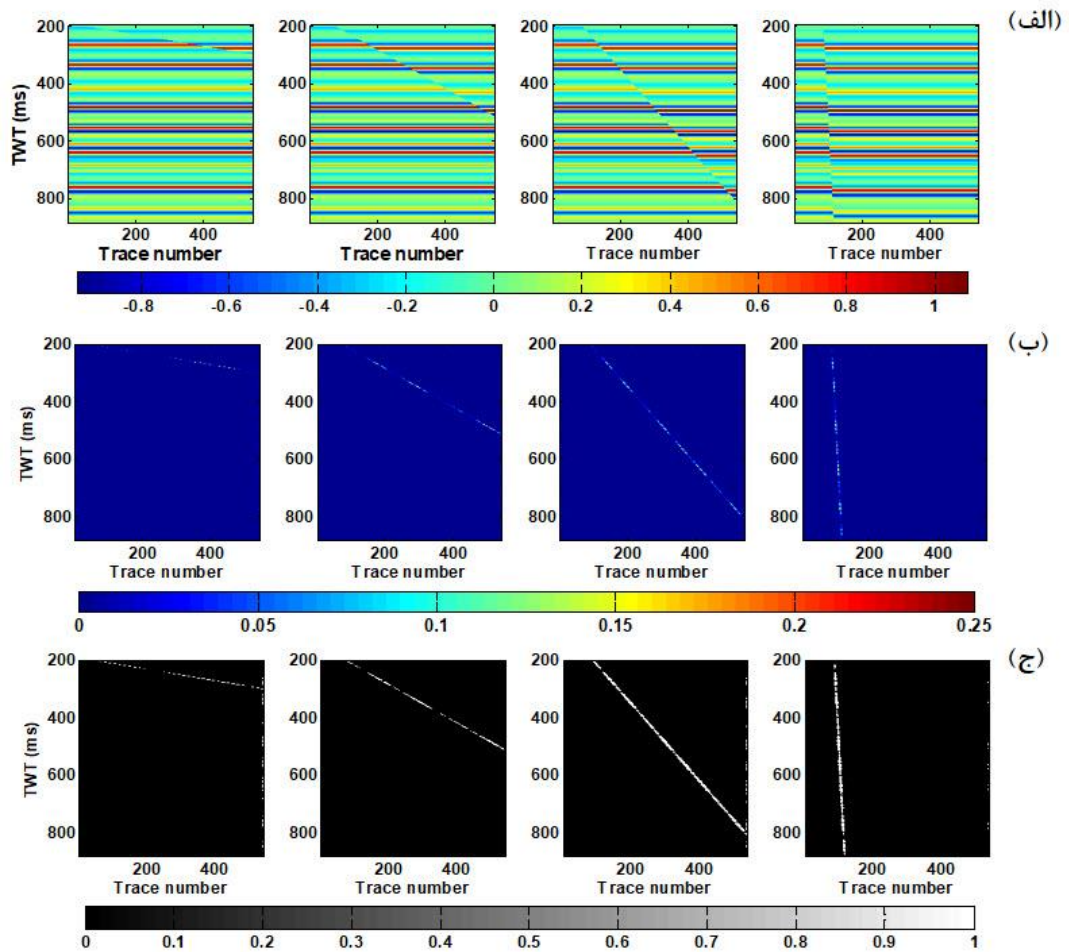
$$\text{Section}_{ij} = \mathbf{S}_{ij} + \mathbf{N}_{ij} \quad (6)$$

$$SNR = \frac{P_{\text{signal}}}{P_{\text{noise}}} \quad (7)$$

شیب‌های متفاوت به ترتیب ۸۰، ۴۰، ۲۰ و ۱۰ درجه بازتابنده‌های افقی جایجا شده‌اند. شکل‌های ۷(الف)، این مقاطع لرزه‌نگاری مصنوعی را نشان می‌دهد. شکل‌های ۷(ب) به ترتیب از راست به چپ نتایج حاصل از خطای حاصل از برازش فرآیند گوسی را بر روی شکل‌های ۷(الف) نشان می‌دهند.

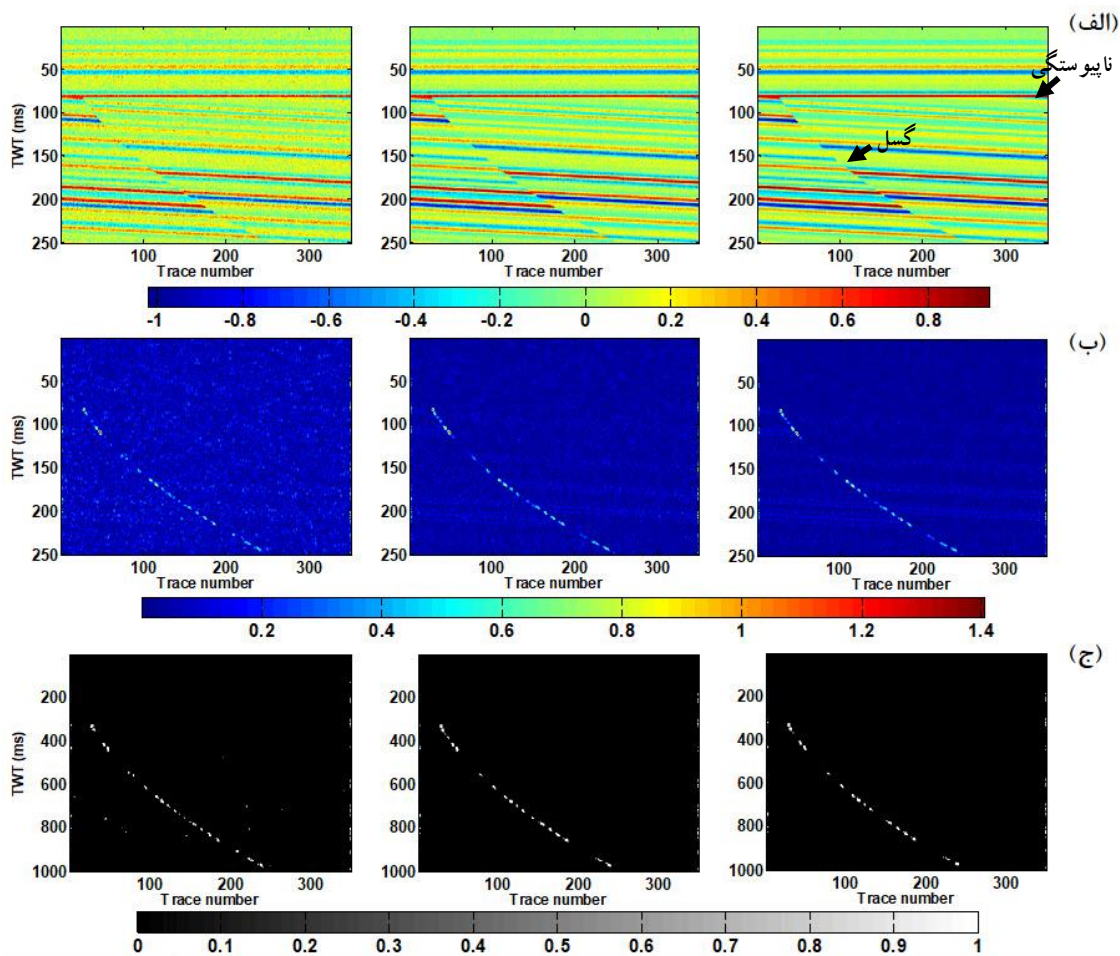
شکل‌های ۷(ج) نتیجه نهایی روش پیشنهاد شده را بعد از اعمال مورفولوژی بر روی این مقاطع نشان می‌دهد. همان‌گونه که مشاهده می‌شود این الگوریتم به خوبی توانسته است گسل‌های نرمال با شیب و جابجایی متفاوت را شناسایی و بارز نماید. هرچند گسل‌ها با میزان جابجایی و شیب متفاوت به خوبی شناسایی شده‌اند اما هرچه میزان جابجایی لایه توسط گسل بیشتر باشد اثر گسل بهتر و پیوسته‌تر توسط این الگوریتم شناسایی می‌شود. این امر به علت برهم ریختگی بیشتر حالت هموار لایه توسط گسل می‌باشد. علاوه بر بررسی شرایط شیب گسل، توانمندی روش در حضور سطوح مختلف نوفه و وجود دو نوع ناهنجاری عمومی، گسل قاشقی و ناپیوستگی، نیز مورد بررسی قرار گرفت. در این مجموعه داده، در ابتدا یک مجموعه بازتابنده شیب‌دار در نظر

¹ Signal noise ratio



شکل ۷. (الف) مقاطع لرزه‌نگاری مصنوعی پس از برانبارش شامل گسل‌های نرمال با شیب‌های مختلف برحسب درجه، به ترتیب از راست به چپ، ۸۰، ۴۰، ۲۰، ۱۰ درجه. مجموعه شکل‌های (ب) نتایج حاصل از الگوریتم GPR بر روی مقاطع لرزه‌ای مصنوعی (الف) قبل از اعمال مورفولوژی را به ترتیب نمایش می‌دهد. (د) نتایج حاصل از اعمال الگوریتم مورفولوژی بر روی نتایج (ب).

ارائه روش برازش فرآیند گوسی در شناسایی گسل در مقاطع لرزه‌نگاری



شکل ۸. (الف) مقاطع لرزه‌نگاری مصنوعی حاوی گسل قاشقی و ناپیوستگی. (ب) خطای حاصل از برازش فرآیند گوسی (پیش‌بینی مدل مصنوعی که توسط فرآیند گوسی یاد گرفته شده) بر روی داده مصنوعی. (ج) نتایج نهایی پس از بررسی خطای برازش و اعمال مورفولوژی.

می‌دهد. همانگونه که مشاهده می‌شود محل گسل خوردگی به خوبی توسط این الگوریتم استخراج شده است. لبه گسل شناسایی شده توسط روش *GPR* به صورت یکسری نقاط آنومالی است. جهت مقایسه روش پیشنهاد شده با روش‌های متداول در نرم‌افزار پترل مقطع لرزه‌نگاری واقعی توسط دو روش انحنای و واریانس نیز مورد بررسی قرار گرفته است. جهت استخراج بهینه‌ترین جواب این دو نشانگر، این دو روش بر روی مکعب شیب راهنمای^۱ (Tingdahl and de Groot, 2003) حاصل از مقطع لرزه‌نگاری اعمال شده‌اند. به دست آوردن

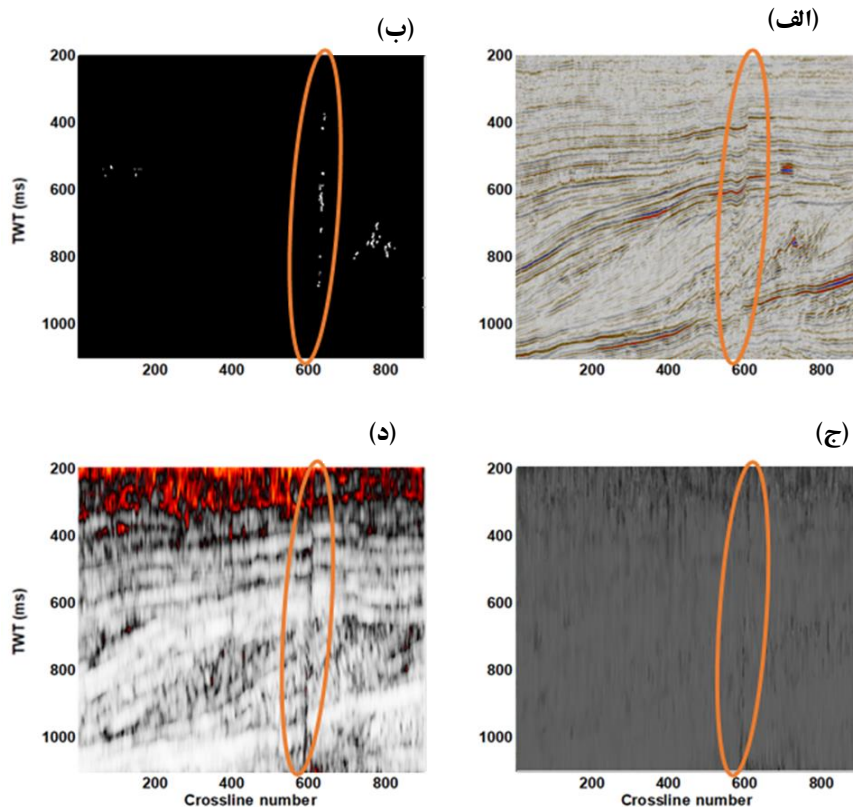
۲.۳ داده واقعی

مقطع لرزه‌نگاری مطالعه شده قسمتی از بلوک *F3* بخش هلندی دریای شمال است. ساختمان زیرسطحی دریای شمال اغلب نتیجه‌ی تکتونیک کششی اواخر ژوراسیک و اوایل کرتاسه است. این ناحیه نمونه‌ای از مخازن نفت و گاز محسوب می‌شود (Michatowicz, 2016). شکل ۹(الف) گسل را با طول ۹۰۰ میلی‌ثانیه (بیضی نارنجی) نشان می‌دهد که در سمت چپ مقطع لرزه‌نگاری واقع شده است. شکل ۹(ب) نتایج حاصل از الگوریتم *GPR* را بر روی این مقطع نشان

^۱ Dip steered cube

نقاط گسلی استخراج شده توسط *GPR* است. تنها ورودی روش *GPR* در این مطالعه داده‌ی لرزه‌ای است که یک فیلتر میانگین بر روی آن اعمال شده است. همچنین روش *GPR* به نوفه پس زمینه حساس نمی‌باشد و توانمندی فراگیری الگوی نوفه پس زمینه را دارد. بدین خاطر نیازی به پیش‌پردازش پیچیده‌ای جهت حذف نوفه‌های پس زمینه نیست. به همین خاطر در نتیجه نهایی حاصل از روش *GPR* در مقایسه با سایر روش‌ها نوفه پس زمینه‌ای دیده نمی‌شود. این خصوصیات توانمندی این روش را در جدا کردن محل گسل نسبت به سایر روش‌ها نشان می‌دهد.

مکعب شیب راهنما هزینه محاسباتی و زمانی در پی خواهد داشت. نتایج حاصل از اعمال این دو روش در شکل ۹ (ج)-(د) آورده شده است. همانگونه که مشاهده می‌شود روش استخراج لبه گسل با استفاده از واریانس توانسته است لبه گسل را بهتر از روش انحنا مشخص نماید. روش انحنا نیز محل گسل را بارز نموده است اما به وضوح روش واریانس و *GPR* نیست. نتایج حاصل از نشانگر واریانس در مقایسه با روش *GPR* به صورت لبه‌ای پیوسته است اما لبه‌های غیر گسلی (مرز بازتابنده‌ها) نیز بارز شده‌اند. اما در روش *GPR* اکثر نقاط آنومالی شناسایی شده مربوط به گسل هستند و لبه‌ی غیر گسلی در داده مشاهده نمی‌شود و این نشان دهنده‌ی میزان دقت و صحت



شکل ۹. اعمال روش شناسایی گسل مبتنی بر برآزش فرایند گوسی بر روی داده واقعی و مقایسه آن با روش‌های متداول شناسایی بخش هلندی دریای شمال که دارای گسل نرمال (بیضی نارنجی رنگ) است. (ب) نقشه *F3* گسل. (الف) مقطع دو بعدی از بلوک نقاط گسلی، شناسایی شده توسط پیشنهاد شده. (ج) اعمال نشانگر انحنا (بیشینه انحنا منفی) بر روی مکعب شیب راهنمای حاصل از داده لرزه‌نگاری. (د) اعمال روش استخراج لبه با استفاده از نشانگر واریانس بر روی مکعب شیب راهنمای حاصل از داده لرزه‌نگاری.

۴. نتیجه‌گیری

استفاده از فرآیند گوسی برای تحلیل ناهنجاری شاید کاری نادرست به نظر برسد، زیرا فرآیند گوسی در فرایندهای طبیعی هنجار به کار می‌رود. نکته‌ی اصلی استفاده از این روش‌ها همین است. فرآیند گوسی یک پدیده را طوری مدل می‌کند که طبیعی‌ترین رفتار را از خود نشان دهد. وجود ناهنجاری سبب می‌شود که فرآیند گوسی در نزدیکی ناهنجاری متحمل خطای بیشتری شود تا بتواند فرآیند ناهنجار را که احتمالاً بسیار متفاوت از یک پدیده‌ی طبیعی عمل می‌کند توصیف نماید. شاید این مسئله به نظر برسد که در این صورت فرآیند گوسی در مورد نوفه‌های داده نیز همین تلاش را می‌کند، گرچه این حرف تا حدی درست است ولی توجه به این نکته ضروری است که پدیده‌ی ناهنجار محلی مانند نوفه بسیار بیشتر از پدیده‌ی ناهنجار عمومی اتفاق می‌افتد. به همین دلیل تابع گوسی فرصت دارد که توابع پایه‌ای برای توصیف نوفه ایجاد نماید ولی به دلیل تُنک بودن ناهنجاری‌های عمومی، توابع پایه‌ی مناسب با آن‌ها یاد گرفته می‌شود.

نتایج حاصل از اعمال روش شناسایی ناهنجاری عمومی توسط برآزش فرآیند گوسی بیانگر این است که می‌توان با این روش به شناسایی محل گسل خوردگی در داده‌های پس از برانبارش پرداخت. در این روش نقاط گسل خوردگی همان محل ناهنجاری‌های عمومی هستند که خط و سیر روال طبیعی و هموار لایه‌های زمین‌شناسی (بازتابنده‌ها) را بر هم زده‌اند. هرچند با وجود گسل‌هایی با شیب و جابجایی‌های متفاوت روش پیشنهاد شده قادر به شناسایی محل گسل می‌باشد اما هرچه میزان جابجایی لایه توسط گسل بیشتر باشد اثر گسل بهتر توسط این الگوریتم شناسایی می‌شود. جهت شناسایی محل گسل با استفاده از روش GPR به پیش پردازش‌های پیچیده جهت حذف نوفه‌نیازی نیست و تنها یک فیلتر میانگین‌گیری در امتداد افق کفایت می‌کند. زیرا روش

GPR فرصت دارد که توابع پایه‌ای برای توصیف نوفه ایجاد نماید. نتایج حاصل از اعمال این روش بر داده واقعی نیز بیان می‌کند روش پیشنهاد شده همانند روش‌های موجود در نرم-افزارهای تجاری توانایی شناسایی محل گسل در داده واقعی را دارد و می‌تواند محل گسل را از سایر پدیده‌ها جدا نماید و در مقایسه با آن‌ها نتیجه‌نهایی دارای نوفه پس زمینه کمتری است و گسل‌ها به صورت واضح و شفاف با نوفه پس زمینه کمتر شناسایی می‌شوند. در مقایسه با سایر روش‌های شناسایی محل گسل، خروجی روش پیشنهاد شده می‌تواند ورودی بهتری برای استخراج صفحات گسل باشد. به عبارتی دیگر در روش ارائه شده نیاز به پس‌پردازش‌های پیچیده نیست و از پیچیدگی مسئله دست‌بندی نتایج حاصل از سایر روش‌ها به منظور استخراج صفحات گسل می‌کاهد.

در رابطه با شناسایی سایر ناهنجاری‌های عمومی موجود در داده لرزه‌نگاری مانند سطوح ناپیوستگی، کانال‌های ماسه‌سنگی و ... نیاز به بررسی و مطالعه بیشتری می‌باشد. هرچند برای یک نمونه خاص ناپیوستگی (ناپیوستگی زاویه-ای) بر روی داده مصنوعی این بررسی انجام شد. سطح ناپیوستگی توسط روش پیشنهاد شده شناسایی نشد. اما نیاز هست حالت‌های ممکن دیگر در مطالعه جداگانه‌ای بررسی شود. آن‌گونه که به نظر می‌رسد روش GPR بتواند علاوه بر گسل ناهنجاری‌های عمومی دیگری را که میزان پخش‌شدگی آن‌ها در مقایسه با کل مجموعه داده لرزه‌ای تُنک^۱ هستند، شناسایی نماید.

۵. سپاس‌گزاری

با سپاس از مدیریت پژوهش و فناوری شرکت ملی نفت ایران و پژوهشکده ازدیاد برداشت از مخازن نفت و گاز که امکانات سخت‌افزاری و نرم‌افزاری لازم جهت انجام این تحقیق را فراهم نمودند.

مراجع

Bahorich, M. S., and Farmer, S. L. (1995). 3D seismic coherency for faults and stratigraphic features. *The Leading Edge*, pp.1053-1058.

¹ Sparse

- Marfurt, K. J., Kirlin, R. L., Farmer, S. L., and Bahorich, M. S. (1998). 3D seismic attributes using a running window semblance based algorithm. *Geophysics*, Vol. 63, pp. 1150-1165.
- Roberts, A. (2001). Curvature attributes and their application to 3D interpreted horizons. *First Break*, Vol. 19(2), pp. 85-100.
- Jaglan H. Qayyum F. and Huck H. (2015). Unconventional seismic attributes for fracture characterization. *First Break*, Vol. 33, pp.101-109.
- Rueger, A., (1997). P-wave reflection coefficients for transversely isotropic models with vertical and horizontal axis of symmetry. *Geophysics*, Vol. 62, pp. 713-722.
- Chopra S. Sudhakar V. Larsen G. and Leong H. (2000). New methodology of the azimuthal variation of smaller faults and fractures. *Word Oil Magazine*, pp. 57-62.
- Zheng Y. and Gray D. (2002). SEG Technical Program Expanded Abstracts. *Geophysics*, pp. 1642-1645.
- Aleardi, M., and Mazzotti, A. (2014). A feasibility study on the expected seismic AVA signatures of deep fractured geothermal reservoirs in an intrusive basement. *Journal of Geophysics Engineering*, Vol. 11(6), Sinopec Geophysical Research Institute.
- Wang, Y., Chen, Sh., and Li, X.Y. (2015). Anisotropic characteristics of mesoscale fractures and applications to wide azimuth 3D P-wave seismic data. *Journal of Geophysics and Engineering*, Vol. 12(3), Sinopec Geophysical Research Institute.
- Chandola V., Banerjee A., and Kumar V. (2009). Anomaly detection. *A survey. ACM Computing Surveys*, Vol. 41(3), Article 15 . 58 pages.
- Grubbs F., (1969). Procedures for detecting outlying observations in samples. *Technometrics*, Vol. 11(1), 1-21.
- Basu S., and Meckesheimer M., (2007). Automatic outlier detection for time series. *an application to sensor data. Know. Inform. Syst*, Vol. 11(2), 137-154.
- Li X., Han J., Kim S., (2006). Motion-Alert: Automatic anomaly detection in massive moving objects. *IEEE Intelligence and Security Informatics Conference (ISI 2006)*, Springer, Berlin, P166-177.
- Laxhammar R., Falkman G., Sviestius E., (2009). Anomaly detection in sea traffic- A comparison of the Gaussian mixture model and kernel density estimator. *12th international conference on information, USA*, P756-763.
- Lane R. O., Nevell D. A., Hayward S. D., and Beaney T. W., (2010). Maritime anomaly detection and threat assessment. *13th International conference on information fusion, UK*.
- Mascaro S., Nicholson A., Korb K., (2014). Anomaly detection in vessel tracks using bayesian networks. *International Journal of Approximate Reasoning*, Vol. 55, 84-98.

- Li Y., Luo X., Qian Y., and Zhao X., (2015). Network-wide traffic anomaly detection and localization based on robust multivariate probabilistic calibration model. *Mathematical Problems in Engineering 2015*, (923792), 26 pages.
- Cheng K.W. Chen Y.T. and Fang W.H. (2015). Gaussian process regression-based video anomaly detection and localization, with hierarchical feature representation. *IEEE Transactions on Image Processing*, Vol. 24(12), pp. 5288–5301.
- Liao W., Rosenhahn B., and Yang M.Y., 2015. Gaussian process for activity modeling and anomaly detection. *ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences*, Vol. 1, 467- 474.
- Pawar A. M., and Mahindrakar M. S., (2015). A comprehensive survey on online anomaly detection. *International Journal of Computer Applications*, Vol. 119. (17).
- Qin K., Wang Y., and Wang B., (2018). Detecting anomalous trajectories using the dempster-Shafer evidence theory considering trajectory features from taxi GNSS data. *Information 9*, 258.
- Herlands W., McFowland E., Wilson A. G., Neil D. B., (2018). Gaussian Process Subset Scanning for Anomalous Pattern Detection in Non-iid Data. *21st International Conference on Artificial Intelligence and Statistics (AISTATS), Lanzarote, Spain*. PMLR: Vol. 84.
- Goldstein M., Uchida S., (2016). A comparative evaluation of unsupervised anomaly detection algorithms for multivariate data. *PLoS ONE* Vol.11(4).
- Rasmussen C.E. and Williams C. K. I. (2006). Gaussian Processes for Machine Learning. *MIT Press*, ISBN 0-262-18253-X.
- Smith M. Roberts S. Psorakis I. S. and Rezek I. (2014). Maritime abnormality detection using Gaussian processes. *Knowledge and Information Systems*, Vol. 38 (3), 717–741.
- Bishop C.M. (2006). Pattern Recognition and Machine Learning. *Springer*, ISBN 0-387-31073-8.
- Dougherty E. (1992). Mathematical Morphology in Image Processing. *Editor, Marcel-Dekker. NewYork*, pp. 255-288.
- Robinson, A. E. (1985). Seismic time invariant convolutional model. *Geophysics*, Vol. 50(12), pp. 2742-2752.
- Michatowicz, P. (2016). Three-dimensional model of the geological structure of the F3 block (Netherland), *9th Geosymposium of Young Researchers, kroczyce*.
- Tingdahl K., and de Groot P. (2003). Post-Stack Dip and Azimuth Processing. *Journal of Seismic Exploration*. Vol. 12, 113-126.