ارائه یک ماژول محاسباتی جدید برای سیستم تحلیل دیجیتالی خط ساحلی، جهت شناسایی داده‌های مشکوک به خطا در داده‌های تغییرات خط ساحلی

کیومرث محمودی[۱]، مصباح سایبانی[۲*]، عباس مرادی[۳]

[۱] کارشناس ارشد مهندسی سواحل، دانشگاه صنعتی امیرکبیر؛ kumarsmahmoodi@aut.ac.ir
[۲] استادیار دانشکده مهندسی دریا، دانشگاه صنعتی امیرکبیر؛ msaybani@aut.ac.ir
[۳] استادیار برنامه‌ریزی و مدیریت مناطق ساحلی، دانشگاه هرمزگان؛ abbas.moradi46@gmail.com

اطلاعات مقاله

تاریخچه مقاله:
تاریخ دریافت مقاله: ۱۳۹۳/۰۸/۱۰
تاریخ پذیرش مقاله: ۱۳۹۴/۰۵/۱۲
تاریخ انتشار مقاله: ۱۳۹۴/۰۶/۳۱

کلمات کلیدی:
تغییرات خط ساحلی
منطقه سوزا
سیستم تحلیل رقومی خط ساحلی
داده مشکوک به خطا
روش‌های آماری شناسایی داده پرت

چکیده

سیستم تحلیل رقومی خط ساحلی (Digital Shoreline Analysis System) که به اختصار DSAS نامیده می‌شود، پرکاربردترین ابزاری است که توسط پژوهشگران و متخصصان برای اندازه‌گیری نرخ تغییرات خط ساحلی مورد استفاده قرار می‌گیرد. عوامل متعددی ممکن است سبب بروز داده‌های خطا در مقادیر اندازه‌گیری شده توسط این ابزار شود. شناخت این خطاها و در صورت امکان بر طرف کردن آنها سبب افزایش میزان دقت نتایج حاصله می‌شود. ابزار DSAS فاقد چنین قابلیتی است. هدف از این تحقیق، ارائه یک ماژول محاسباتی جدید برای DSAS به منظور شناسایی داده‌های مشکوک به خطا است. این ماژول یک فایل اجرایی است که با استفاده از نرم‌افزار MATLAB نوشته شده و از طریق فایل-های XML با DSAS ارتباط برقرار می‌کند. شناسایی داده‌های مشکوک به خطا با استفاده از روش‌های آماری نمودار جعبه‌ای، آزمون هامپل، آزمون انحراف مطلق میانه و آزمون Z-Score انجام می‌شود. نتایج حاکی از عملکرد رضایت بخش این ماژول در شناسایی خطاهای احتمالی موجود در داده‌ها است.

Provide a New Computational Module for Digital Shoreline Analysis System to Detect Uncertain Data in the Shoreline Change Data

Kumars Mahmoodi[1], Mesbah Sayehbani[2*], Abbas Moradi[3]

[1] M.Sc. Coastal Engineering, Amirkabir University of Technology; kumarsmahmoodi@aut.ac.ir
[2] Assistant Professor, Faculty of Marine Engineering, Amirkabir University of Technology; msaybani@aut.ac.ir
[3] Assistant Professor, Hormozgan University; abbas.moradi46@gmail.com

ARTICLE INFO

Article History:
Received: 1 Nov. 2014
Accepted: 3 Aug. 2015
Available online: 22 Sep. 2015

Keywords:
Shoreline changes
Souza area
Digital Shoreline Analysis System (DSAS)
Suspected data error
Statistical outlier detection methods

ABSTRACT

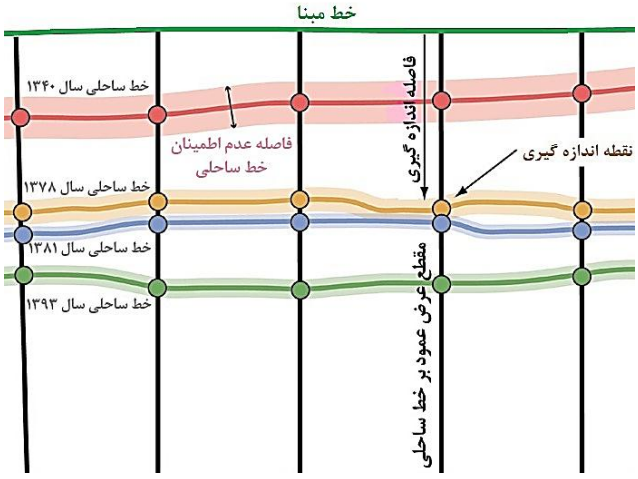
Digital Shoreline Analysis System (DSAS) is the most widely practical tool that is used by researchers and experts for measuring shoreline change rate. Numerical factors may cause data errors in the measured values by this tool. Understanding these errors and also fixing them in case of it is possible have improved the accuracy of results. The DSAS does not have this capability. The purpose of this paper is presenting a new computational module for DSAS to identify suspected data errors. This module is an executable file that is written using MATLAB software and communicates with DSAS through XML files. Suspected data errors detection is performed by using statistical methods include: Box plot, Hampel's test, Median Absolute Deviation (MAD) and Z-score test. The results show satisfactory performance of this module in identifying possible data errors.

[Downloaded from marine-eng.ir on 2026-11-12]
[DOR: 20.1001.1.17357608.1394.11.21.7.1]

۱- مقدمه

خط ساحلی یکی از مهمترین اجزای سیستم ساحلی به شمار می‌آید. کمیته بین‌المللی داده‌های جغرافیایی، خط ساحلی را به عنوان یکی از ۲۷ پدیده مهم سطح کره زمین تعریف کرده است. این پدیده تحت تاثیر فرایندهای طبیعی و فعالیت‌های انسانی در مقیاس‌های زمانی-مکانی دائماً در حال تغییر است. تعیین، پایش، کمی‌سازی و پیش‌بینی خط ساحلی در زندگی و فعالیت‌های روزانه مردم دارای اهمیت ویژه‌ای است. این فعالیت‌ها شامل کشتیرانی ساحلی، مدیریت منابع ساحلی، حفاظت محیط ساحلی، برنامه‌ریزی و توسعه پایدار مناطق ساحلی، کاربری اراضی و ایجاد خطوط حریم توسعه ساحلی است. این موارد سبب شده است مطالعه تغییرات خط ساحلی مشغله ذهنی عمده دانشمندان، مهندسین و مدیران ساحلی شود.

برای مطالعه خط ساحلی و کمی‌سازی آن از ابزارهای مختلفی استفاده می‌شود. یکی از پرکاربردترین ابزارهایی که به این منظور استفاده می‌شود، سیستم تحلیل رقومی خط ساحلی[۱] (DSAS) است [۱، ۲]. این ابزار به صورت یک افزونه[۲] روی نرم‌افزار ArcGIS نصب شده و با ایجاد مقاطع عرضی[۳] عمود بر خطوط ساحلی، نرخ تغییرات خط ساحلی را حول مقاطع ایجاد شده، نسبت به یک مبنای[۴] مشخص محاسبه می‌کند (شکل ۱). با استفاده از فاصله هر نقطه اندازه‌گیری تا خط مبنا و تاریخ خط ساحلی برداشت شده، نرخ تغییرات اندازه‌گیری می‌شود. DSAS با استفاده از چندین روش آماری نرخ تغییرات خط ساحلی را محاسبه می‌کند؛ برخی از این روش‌ها عبارتند از: نقطه نهایی[۵] (EPR)، رگرسیون خطی[۶] (LRR)، رگرسیون خطی وزن دار[۷] (WLR) و فاصله اطمینان نقطه نهایی[۸] (ECI).



شکل ۱- نحوه عملکرد DSAS برای محاسبه نرخ تغییرات خط ساحلی

داده‌های حاصل شده از محاسبات آماری انجام گرفته توسط DSAS، به عنوان پایه اساسی تجزیه و تحلیل خطوط ساحلی و سایر کاربردهای مرتبط به شمار می‌آیند. از اینرو صحت داده‌ها و عاری بودن آنها از هر گونه خطای احتمالی ضروری است. عوامل مختلفی ممکن است سبب بروز داده‌های خطا در داده‌های حاصل شده از محاسبات انجام شده توسط DSAS شود. برخی از این عوامل عبارتند از:

- خطاهای متعدد در سطوح مختلف استخراج خط ساحلی از داده‌های فضایی (مانند عکس‌های هوایی، نقشه‌ها، تصاویر ماهواره‌ای و غیره)؛
- خطاهای مربوط به داده‌ها: به عنوان مثال عکس‌های هوایی رایج‌ترین منبع استخراج خطوط ساحلی به شمار می‌آیند. این عکس‌ها خود دارای منابع مختلف تولید خط می‌باشند؛ به عنوان مثال عواملی نظیر اعوجاج شعاعی یا محوری و انحراف زاویه سبب بروز داده‌های خطا می‌شوند؛
- خطاهای مربوط به روش‌های اندازه‌گیری؛
- خطای مربوط به نقاط کنترل زمینی؛
- خطای دستگاه‌های اندازه‌گیری؛
- خطاهای انسانی.

با توجه به مطالب ذکر شده می‌توان گفت اندازه‌گیری نرخ تغییرات خط ساحلی بدون هیچ‌گونه خطایی در عمل امکان‌پذیر نیست. با این وجود محققان ساحلی در تلاش هستند که از طریق بهبود کیفیت داده‌ها و بهبود روش‌های اندازه‌گیری و استخراج خط ساحلی، میزان خطاها را به حداقل برسانند.

شرکت توسعه دهنده ابزار DSAS، موسسه تحقیقاتی سیستم‌های محیطی[۹] (ESRI) آمریکا است [۳]. در حال حاضر این ابزار فقط دارای امکاناتی برای محاسبه تغییرات خط ساحلی است و نمی‌تواند خطاهای احتمالی موجود در مقادیر اندازه‌گیری شده را محاسبه کند. با توجه به اهمیت بسیار زیاد شناسایی خطاهای احتمالی، افزودن قابلیت شناسایی داده‌های مشکوک به خطا به این ابزار امری ضروری است. هدف از این مقاله، ارائه یک ماژول[۱۰] محاسباتی جدید به ابزار DSAS است که قابلیت شناسایی داده‌های مشکوک به خطا موجود در مقادیر اندازه‌گیری شده با استفاده از روش‌های اندازه‌گیری نرخ تغییرات را داشته باشد. شناسایی داده‌های مشکوک به خطا با استفاده از روش‌های آماری نمودار جعبه‌ای[۱۱]، آزمون هامپل[۱۲]، آزمون انحراف مطلق میانه[۱۳] و آزمون Z-Score انجام می‌شود. در بخش ۲ جزئیات روش‌های آماری مذکور، و در بخش ۳ جزئیات ماژول تهیه شده تشریح شده است. همچنین در بخش ۴، کارائی این ماژول روی یک مجموعه داده اندازه‌گیری نرخ تغییرات خط ساحلی بررسی شده است. مجموعه داده مورد بررسی، داده‌های حاصل از کمی‌سازی نرخ تغییرات خط ساحلی منطقه سوزا

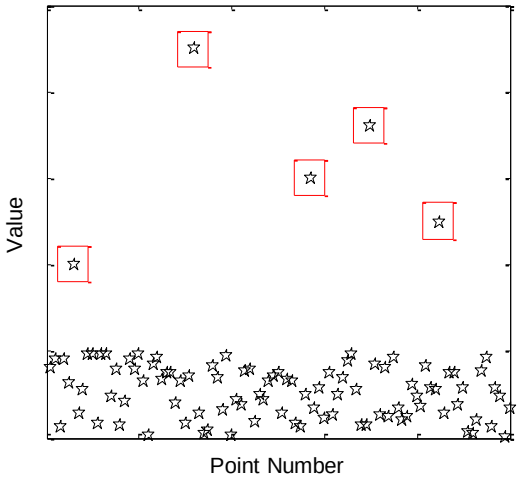
[Downloaded from marine-eng.ir on 2026-11-12]

[DOR: 20.1001.1.17357608.1394.11.21.7.1]

(واقع در جزیره قشم) با استفاده از روش‌های EPR [۴] و LRR [۵] است.

۲- معرفی روش‌های شناسایی داده‌های مشکوک به خطا

روش‌های مورد بحث در این تحقیق برای شناسایی داده‌های مشکوک به خطا، شامل شماری از الگوریتم‌های آماری است که از آنها به منظور شناسایی داده‌های متناقض در یک مجموعه داده استفاده می‌شود. معمولاً داده‌هایی که به صورت متناقض با سایر اعضا ایجاد می‌شوند، می‌توانند کاندیدای داده خطا باشند. داده‌های مشکوک به خطا به طور معمول در فاصله دورتری نسبت به سایر اعضای مجموعه داده واقع می‌شوند، به گونه‌ای که این گمان به وجود می‌آید که با یک روند متفاوت تولید شده‌اند [۶]. به عنوان مثال در شکل ۲، داده‌هایی که با یک مربع به دور آنها مشخص شده‌اند، می‌توانند بیانگر داده‌های مشکوک به خطا باشند. همانطور که در این شکل مشخص است، این داده‌ها در فاصله دورتری نسبت به سایر اعضای مجموعه داده قرار گرفته‌اند.

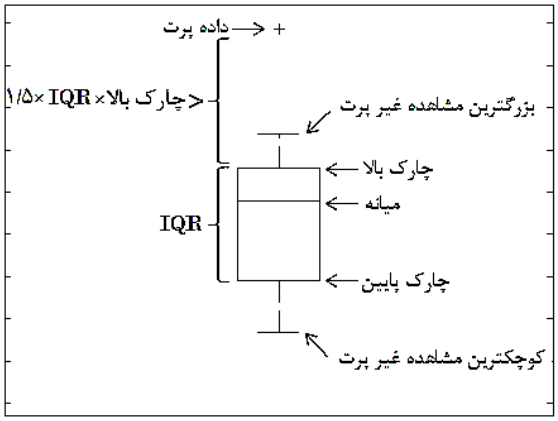


شکل ۲- یک مجموعه داده تصادفی با تعداد ۵ داده مشکوک به خطا

محققان تا کنون روش‌های متنوعی را برای شناسایی داده‌های مشکوک به خطا ارائه کرده‌اند [۷]. روش‌های آماری [۶، ۸، ۹] طیف وسیعی از این روش‌ها را به خود اختصاص می‌دهند. این روش‌ها به تکنیک‌های آماری وابسته بوده و توزیع و یا احتمالی از مجموعه داده‌ها را در نظر می‌گیرند. یکی از مزایای روش‌های آماری سادگی کاربرد آنها است، اما بیشتر آنها با مجموعه داده‌های با ابعاد بزرگ قابل استفاده نیستند. مجموعه داده‌های ایجاد شده با DSAS تک متغیره بوده و دارای ابعاد بزرگ نیستند، به همین دلیل می‌توان خطاهای احتمالی موجود در آنها را با استفاده از روش‌های آماری شناسایی کرد. روش‌های مورد بحث در این مقاله شامل چهار روش آماری است که در ادامه معرفی شده‌اند.

۲-۱- نمودار جعبه‌ای

نمودار جعبه‌ای [۱۰، ۱۱] یک روش گرافیکی تک متغیره است که توزیع داده‌ها را با استفاده از پنج مشخصه اصلی محاسبه می‌کند: کوچکترین مشاهده غیر پرت، چارک پایین ($Q1$)، میانه، چارک بالا (Q3) و بزرگترین مشاهده غیر پرت. مقدار $Q3 - Q1$ محدوده چارک میانی[۱۴] (IQR) را مشخص می‌کند. با استفاده از این پارامتر می‌توان پرت (مشکوک به خطا) و یا غیر پرت (نرمال) بودن یک داده را مشخص کرد؛ داده‌ای که $1.5 \times IQR$ بار کوچکتر از $Q1$، و یا $1.5 \times IQR$ بار بزرگتر از Q3 باشد، می‌تواند به عنوان کاندیدای مشکوک به خطا در نظر گرفته شود. مفاهیم گفته شده در شکل ۳ نشان داده شده است.



شکل ۳- نمودار جعبه‌ای و مفاهیم آن

۲-۲- آزمون هامپل

آزمون هامپل [۱۲، ۱۳] یکی از روش‌های آماری شناسایی داده‌های مشکوک به خطا در مجموعه داده‌های تک متغیره است. این آزمون از لحاظ تئوری روشی پایدار است، به این معنا که به داده‌های مشکوک به خطا حساس نیست (تعداد و مقادیر داده‌های مشکوک به خطا بر قابلیت آن اثرگذار نیست). از دیگر مزایای این روش این است که اندازه مجموعه داده بر عملکرد آن اثر گذار نیست. برای انجام آزمون هامپل گام‌های زیر ضروری است:

۱. محاسبه میانه داده‌ها (Med)

۲. محاسبه انحراف اعضای مجموعه داده از میانه (r_i)

$$r_i = (x_i - Med) \quad (۱)$$

داده x_i عضوی از مجموعه داده است.

۳. محاسبه میانه انحرافات اعضای مجموعه داده از میانه داده‌ها ($Med_{|r_i|}$)

۴. بررسی شرط $|r_i| \geq 4.5Me_{|r_i|}$. اگر برای داده x_i شرط فوق برقرار باشد، می‌توان گفت x_i کاندیدای یک داده خطا است.

[Downloaded from marine-eng.ir on 2026-11-12]

[DOR: 20.1001.1.17357608.1394.11.21.7.1]

۲-۳- آزمون انحراف مطلق میانه

آزمون انحراف مطلق میانه [۱۳] یکی از روش‌های پایه‌ای و در عین حال قوی شناسایی داده‌های مشکوک به خطا در مجموعه داده‌های تک متغیره است. این آزمون یک روش پایدار بوده و تحت تاثیر داده‌های مشکوک به خطا موجود در مجموعه داده قرار نمی‌گیرد. آزمون انحراف مطلق میانه به شکل‌های زیر بیان می‌شود:

$$2\ MAD_e\ Method = Median \pm 2MAD_e \quad (۲)$$

$$3\ MAD_e\ Method = Median \pm 3MAD_e \quad (۳)$$

برای داده‌های نرمال رابطه زیر برقرار است:

$$MAD_e = 1.483 \times MAD \quad (۴)$$

که

$$MAD = Median(|x_i - Median(x)|) \quad (۵)$$

طبق روابط (۲) و (۳) داده‌هایی که خارج از این بازه‌ها قرار گیرند، ممکن است کاندیدای داده خطا باشند.

۲-۴- آزمون Z-Score

آزمون Z-Score [۱۴، ۱۵] یک روش آماری است که می‌توان از آن برای شناسایی داده‌های مشکوک به خطا در مجموعه داده‌های تک متغیره استفاده کرد. این روش از میانگین حسابی و انحراف معیار برای شناسایی داده‌های مشکوک به خطا استفاده می‌کند، بنابراین تحت تاثیر اعضای مجموعه داده است که این عامل ممکن است بر عملکرد آن اثر گذار باشد. آزمون Z-Score به شکل زیر بیان می‌شود:

$$Z_i = \frac{x_i - \bar{x}}{SD} \quad (۶)$$

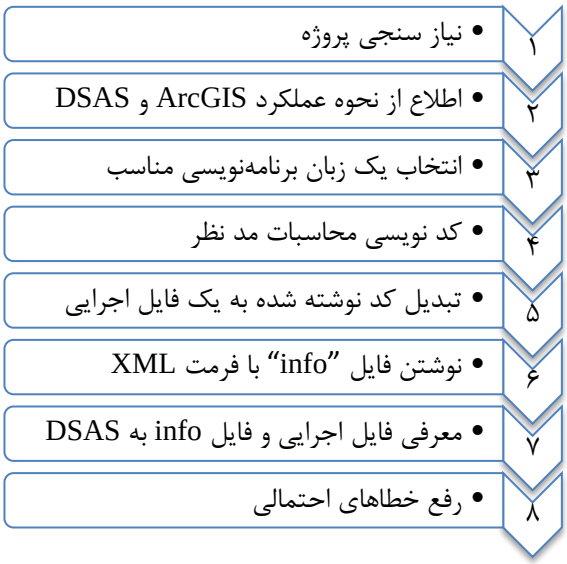
در رابطه (۶)، x_i عضوی از مجموعه داده، $\bar{x}$ میانگین حسابی و SD انحراف معیار داده‌ها است. مقادیر Z_i هایی که قدرمطلق آنها از ۳ تجاوز کند، به عنوان کاندیدای داده مشکوک به خطا محسوب می‌شوند.

۳- معرفی ماژول نوشته شده

در این بخش نحوه طراحی ماژول نوشته شده برای شناسایی داده‌های مشکوک به خطا و شیوه عملکرد آن تشریح شده است. آخرین نگارش DSAS نسخه ۴.۳ بوده که در آپریل ۲۰۱۲ عرضه شده است [۱]. این نسخه فقط با ArcGIS v.10 به بالا سازگاری داشته و قابل اجرا روی سیستم عامل‌های Windows XP، Vista و Windows 7 است. اما ماژول نوشته شده در این تحقیق کلی بوده و قابل بکارگیری با نسخه‌های پایین‌تر ArcGIS است. برای اجرای DSAS 4.3 روی یک سیستم عامل قابل پشتیبانی باید ملزومات زیر فراهم باشد، در غیر این صورت DSAS اجرا نشده و در نتیجه ماژول نوشته شده نیز غیر قابل اجرا خواهد بود:

۱. Microsoft .NET Framework 3.5 یا نسخه‌های بالاتر.
۲. ArcGIS Desktop 10. افزونه DSAS با نسخه‌های ArcGIS v.9.x سازگاری ندارد.
۳. شاخص‌های پشتیبانی .NET. برای ArcGIS
۴. کامپایلر زمان اجرای نرم‌افزار[۱۵] MATLAB (MCR).

تمامی محاسبات آماری قابل انجام توسط DSAS با استفاده از یک ماژول خارجی انجام می‌شود. این ویژگی به توسعه‌دهندگان آن این قابلیت را فراهم می‌آورد که بتوانند ایده‌های خود برای انجام محاسبات جدید توسط DSAS را عملی سازند. در شکل ۴ مراحل کلی طراحی یک ماژول برای DSAS نشان داده شده است.



شکل ۴- مراحل ایجاد یک ماژول برای DSAS

۱- نیاز سنجی پروژه: اولین مرحله در انجام یک پروژه، نیاز سنجی اجرای آن است. در مرحله نیاز سنجی، ضرورت تهیه ماژول جدید بررسی می‌شود. ماژول جدید باید به توانایی‌های قبلی ابزار DSAS افزوده و محاسبات را بهبود بخشد. در صورتی که این ماژول فاقد این امکان باشد و یا یک سری محاسبات غیر ضروری را انجام دهد، نوشتن آن مقرون به صرفه نبوده و عملاً یک کار بدون کاربرد انجام شده است.

۲- اطلاع از نحوه عملکرد ArcGIS و DSAS: قبل از تهیه یک ماژول برای یک نرم‌افزار خاص، باید از نحوه عملکرد آن نرم‌افزار مطلع شد. این مورد شامل یادگیری کلی کار با نرم‌افزار و همچنین آگاهی در خصوص نحوه توسعه آن است. در خصوص توسعه DSAS نیز این موارد برقرار است.

۳- انتخاب یک زبان برنامه‌نویسی مناسب: برای نوشتن یک ماژول باید از زبان‌های برنامه‌نویسی مناسب استفاده شود. با استفاده زبان برنامه‌نویسی می‌توان نحوه انجام محاسبات مورد نیاز را به رایانه

[Downloaded from marine-eng.ir on 2026-11-12]
[DOR: 20.1001.1.17357608.1394.11.21.7.1]

آموخت. ماژول‌هایی که برای توسعه DSAS نوشته می‌شوند را می‌توان با هر زبان برنامه‌نویسی که قابلیت تولید یک فایل اجرایی را داشته باشد، ایجاد کرد.

۴- کدنویسی محاسبات مد نظر: پس از انتخاب زبان برنامه‌نویسی، با استفاده از دستورات آن زبان می‌توان محاسباتی را که قرار است ماژول انجام دهد، نوشت.

۵- تبدیل کد نوشته شده به یک فایل اجرایی: پس از تهیه فایل کد منبع، باید آن را تبدیل به یک فایل اجرایی[۱۶] کرد. تمامی ماژول‌های طراحی شده برای DSAS، باید به شکل یک فایل اجرایی قابل اجرا توسط سیستم عامل ویندوز باشند. این فایل می‌تواند به صوت یک فایل دودویی[۱۷] قابل اجرا (در واقع با پسوند exe.) و یا هر فایل با یک تابع ریجیستر شده بکارگیرنده[۱۸] که قابلیت اجرای آن را داشته باشد (مثل فایل‌های Python با پسوند py.، فایل‌های Perl با پسوند pl.، و غیره)، تولید شود. قرارداد نام‌گذاری ماژول‌های اجرایی DSAS به صورت زیر است:

[module_name].calc. [file_extension]

که module_name نام ماژول، و file_extension پسوند آن (مثلاً exe.) است.

۶- نوشتن فایل "info" با فرمت XML[۱۹]: تمامی ماژول‌های نوشته شده برای DSAS باید با یک فایل با نام info و با فرمت XML همراه باشد. این فایل بیانگر قابلیت‌های ماژول تهیه شده بوده و لیستی از فیلدهایی که به عنوان نتیجه محاسبات ماژول تولید می‌شود، و همچنین اطلاعات دسته‌ای که DSAS برای سازماندهی و قرار دادن لیست محاسبات موجود در پنجره محاسبات آماری خود (DSAS calculate statistic window) بکار می‌برد، را مشخص می‌کند. قرارداد نام‌گذاری فایل info در ماژول‌های DSAS به شکل زیر است:

[module_name].info.xml

که module_name نام ماژول نوشته شده است. به عنوان مثال در شکل ۵، یک فایل info نشان داده شده است.

```
<?xml version="1.0" encoding="UTF-8"
standalone="true"?>
<DSAS_CALCULATIONS parallel="Yes"
dynamic="No"><FIELD category="Distance
Measurement" length="" type="Double"
alias="Shoreline Change Envelope"
name="SCE"/><FIELD category="Distance
Measurement" length="" type="Double"
alias="Net Shoreline Movement"
name="NSM"/><FIELD category="Point
Change" length="" type="Double" alias="End
Point Rate" name="EPR"/><FIELD
category="Regression Statistics" length=""
type="Double" alias="Linear Regression Rate"
name="LRR"/><FIELD length=""
type="Double" alias="R-squared LRR"
name="LR2"/><FIELD length=""
type="Double" alias="Standard Error LRR"
name="LSE"/><FIELD length=""
type="Double" alias="Confidence Interval
LRR" name="LCI{CI}"/><FIELD
category="Regression Statistics" length=""
type="Double" alias="Weighted Linear
Regression" name="WLR"/><FIELD
length="" type="Double" alias="R-squared
WLR" name="WR2"/><FIELD length=""
type="Double" alias="Standard Error WLR"
name="WSE"/><FIELD length=""
type="Double" alias="Confidence Interval
WLR" name="WCI{CI}"/><FIELD
category="Advanced Statistics" length=""
type="Double" alias="Least Median of
Squares" name="LMS"/>
<!-- AE 20081001 Disabled until NLR
implementation is completed. <FIELD
name="NLR" alias="Neighborhood Linear
Regression Rate" type="Double" length=""
category="Advanced Statistics" /> <FIELD
name="NCI{CI}" alias="Confidence Interval
NLR" type="Double" length="" />-->
</DSAS_CALCULATIONS>
```

شکل ۵- یک نمونه فایل info برای ماژول محاسباتی DSAS

۷- معرفی فایل اجرایی و فایل info به DSAS: پس از ایجاد فایل اجرایی و فایل info، باید برای اجرا آن را به DSAS معرفی نمود. به این منظور باید این دو فایل را در مسیری که DSAS ماژول‌ها را از آنجا فراخوانی می‌کند، قرار داد. این مسیر در DSAS 4.3 به طور پیش‌فرض به صورت زیر است:

C:\Program Files (x86)\USGS\DSAS\template\calc

در صورتی که هنگام نصب DSAS یک مسیر دیگر توسط کاربر تعریف شده باشد، باید فایل اجرایی و فایل info را در پوشه‌ای با نام calc قرار داد. این پوشه به طور پیش‌فرض وجود داشته و نیازی به ایجاد آن به صورت دستی نیست.

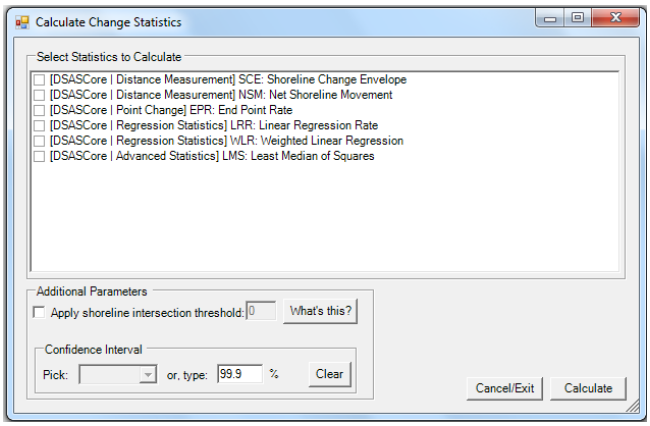
۸- رفع خطاهای احتمالی: پس از ایجاد ماژول‌ها، ممکن است خطاهایی حین مراحل مختلف کار رخ دهد، مثلاً ماژول توسط DSAS شناسایی نشود، داده‌ها به درستی در جداول قرار نگیرند، مقادیر اندازه‌گیری شده غیر واقع و نادرست باشد و غیره. بنابراین باید علت بروز خطا را شناسایی کرده و نسبت به رفع آن اقدام کرد.

پس از اجرای مراحل فوق، نوبت به اجرای ماژول می‌رسد. وقتی کاربر دکمه Calculate را از پنجره محاسبه نرخ تغییرات فشار می‌دهد (شکل ۶)، DSAS از مسیر مربوطه ماژول را جستجو کرده و در صورت وجود، آن را اجرا می‌کند. ورودی تمامی ماژول‌ها یک فایل XML است که محتوای آن، اطلاعات و داده‌های مقاطع عرضی ایجاد شده توسط DSAS است. به عنوان مثال در شکل ۷، بخش کوچکی از یک فایل XML ایجاد شده توسط DSAS که در واقع ورودی ماژول‌ها است، نشان داده شده است. خروجی ماژول‌ها

[Downloaded from marine-eng.ir on 2026-11-12]

[DOR: 20.1001.1.17357608.1394.11.21.7.1]

نیز باید یک فایل XML باشد، که محتوای آن حاصل محاسبات انجام شده توسط ماژول است. پس از تولید فایل خروجی XML توسط ماژول، DSAS آن را از ورودی دریافت کرده و فیلدهای محاسبه شده توسط ماژول را به جدول خروجی نرخ، با نام [transect_layer_name]_rates_[YYYYMMDD_hhmmss] الحاق می‌کند. [transect_layer_name] نام لایه مقاطع عرضی انتخاب شده توسط کاربر در منوی DSAS، در مرحله شروع محاسبات است. و YYYY بیانگر سال، MM بیانگر ماه، DD بیانگر روز، hh بیانگر ساعت، mm بیانگر دقیقه و ss بیانگر ثانیه است. با توجه به مطالب بیان شده، در شکل ۸ به طور خلاصه نحوه تعامل DSAS و ماژول برای انجام محاسبات مورد نیاز ارائه شده است.



شکل ۶- پنجره محاسبه نرخ تغییرات افزونه DSAS

```
<?xml version="1.0" encoding="UTF-8"?>
<!--DSAS output file-->
<DSAS_OUT>
<Header>
<CI>99.9</CI>
</Header>
<IntersectTable>
<IntersectList>
<Intersect>
<TransectID>1</TransectID>
<BaselineID>1</BaselineID>
<BaselineDistance>150</BaselineDistance>
<TotalCumulativeDistance>150</TotalCumulativeDistance
>
<Group>1</Group>
<GroupDistance>150</GroupDistance>
<Date>1956.00273790926</Date>
<DistanceToBaseline>82.41</DistanceToBaseline>
<M>NaN</M>
<MeasurementUncertainty>2</MeasurementUncertainty>
<Bias>NaN</Bias>
<BiasUncertainty>NaN</BiasUncertainty>
<X>428569.52</X>
<Y>2981170.62</Y>
</Intersect>
</Intersect>
...
</Intersect>
```

```
...
</IntersectList>
</IntersectTable>
</DSAS_OUT>
```

شکل ۷- یک نمونه فایل XML تولیدی توسط DSAS



شکل ۸- مراحل تعامل DSAS و ماژول

یک ماژول برای اجرای صحیح و تعامل با DSAS باید دارای شرایط زیر باشد:

- ماژول نوشته شده باید فاقد هرگونه درخواست ورودی از کاربر باشد؛ چون ماژول از طریق DSAS اجرا شده و هیچ‌گونه خروجی و یا پیامی به کاربر نشان داده نمی‌شود.
- خروجی ماژول همواره باید یک فایل XML معتبر و با شکل مطلوب باشد، حتی اگر ورودی‌ها نامعتبر بوده و یا وقتی خطایی رخ می‌دهد.
- ماژول باید توانایی دریافت فایل ورودی با ابعاد بزرگ را داشته باشد.

تا اینجا کلیات طراحی یک ماژول برای DSAS مورد بحث قرار گرفت. در ادامه ماژول طراحی شده در این تحقیق، معرفی شده است.

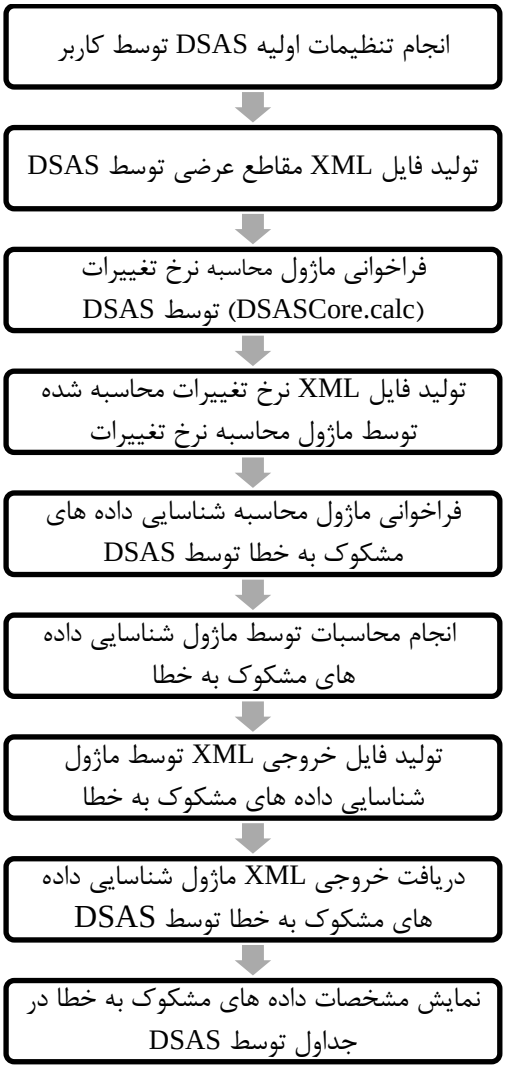
۳-۱- معرفی ماژول شناسایی داده‌های مشکوک به خطا

ماژول مورد بحث در این مقاله با استفاده از روش‌های آماری تشریح شده در بخش ۲، به شناسایی داده‌های مشکوک به خطا در مقادیر

[Downloaded from marine-eng.ir on 2026-11-12]

[DOR: 20.1001.1.17357608.1394.11.21.7.1]

نرخ تغییرات خط ساحلی محاسبه شده با استفاده از روش‌های موجود در DSAS می‌پردازد. بنابراین ورودی این ماژول، مقادیر نرخ تغییرات خط ساحلی محاسبه شده با استفاده از نوع آمار انتخابی کاربر است. و خروجی آن مقادیر مشکوک به خطا و اندیس آنها است. ماژول پس از دریافت داده‌ها از ورودی (با فرمت XML)، با استفاده از روش‌های آماری ذکر شده در بخش ۲ داده‌های مشکوک به خطا را شناسایی کرده و آنها را در قالب یک فایل XML همراه با مشخصات مجموعه داده مورد بررسی و نام روش شناسایی داده مشکوک به خطا، ذخیره می‌کند. سپس DSAS فایل خروجی ماژول را دریافت کرده و در یک جدول نام مجموعه داده[1]، نام روش شناسایی داده مشکوک به خطا، اندیس و مقادیر داده‌های مشکوک به خطا را ذخیره می‌کند. با توجه به مطالب بیان شده، نحوه عملکرد کلی ماژول تهیه شده برای شناسایی داده‌های مشکوک به خطا در شکل ۹ ارائه شده است.



شکل ۹- مراحل عملکرد ماژول برای شناسایی داده‌های مشکوک به خطا

[1] نام مجموعه داده همان نام روش محاسبه نرخ تغییرات در نظر گرفته شده است، مثلاً EPR، LRR و غیره.

همانطور که از شکل ۹ مشخص است، عملکرد ماژول شناسایی داده‌های مشکوک به خطا وابسته به عملکرد صحیح یک ماژول دیگر است. نام این ماژول، ماژول محاسبه نرخ تغییرات می‌باشد که نام آن در DSAS به طور پیش‌فرض DSASCore.calc است. این ماژول هسته محاسباتی DSAS بوده و وظیفه محاسبه نرخ تغییرات خط ساحلی را با استفاده از آمار موجود بر عهده دارد. در واقع خروجی ماژول DSASCore.calc، ورودی ماژول شناسایی داده‌های مشکوک به خطا است. به همین دلیل اگر این ماژول نتواند کار خود را به درستی انجام دهد، ماژول شناسایی داده‌های مشکوک به خطا نیز نمی‌تواند به درستی عمل کند.

ماژول DSASCore.calc با استفاده از نرم‌افزار MATLAB نوشته شده است. علت انتخاب MATLAB امکانات بسیار زیاد محاسباتی آن است که به برنامه‌نویسان این امکان را فراهم می‌آورد که بتوانند ایده‌های خود را با کمترین پیچیدگی و بیشترین کارائی عملی سازند. ایجاد فایل‌های اجرایی از کدهای MATLAB نیازمند همراه کردن کامپایلر زمان اجرای آن است. این کامپایلر MCR نام دارد و شامل توابع کتابخانه‌ای زمان اجرای MATLAB است. بنابراین اگر پس از ایجاد فایل اجرایی روی رایانه مقصد MCR نصب نباشد، فایل اجرایی ایجاد شده قابل اجرا نخواهد بود. برای استفاده از کامپایلر MATLAB، نیازی به نصب نسخه کامل MATLAB نیست و فقط کافی است کامپایلر به صورت جداگانه روی سیستم نصب شود[2].

در این تحقیق نیز ماژول شناسایی داده‌های مشکوک به خطا با استفاده از نرم‌افزار MATLAB نوشته شده است. نام این ماژول OutlierDetection.calc، و نام فایل info آن OutlierDetection.info است. پس از معرفی ماژول به DSAS، لیست محاسباتی که انجام می‌دهد به منوی انجام محاسبات آماری DSAS اضافه شد. در شکل ۱۰ این گزینه‌ها نشان داده شده‌اند. در زیر برخی از ویژگی‌های ماژول تشریح شده است:

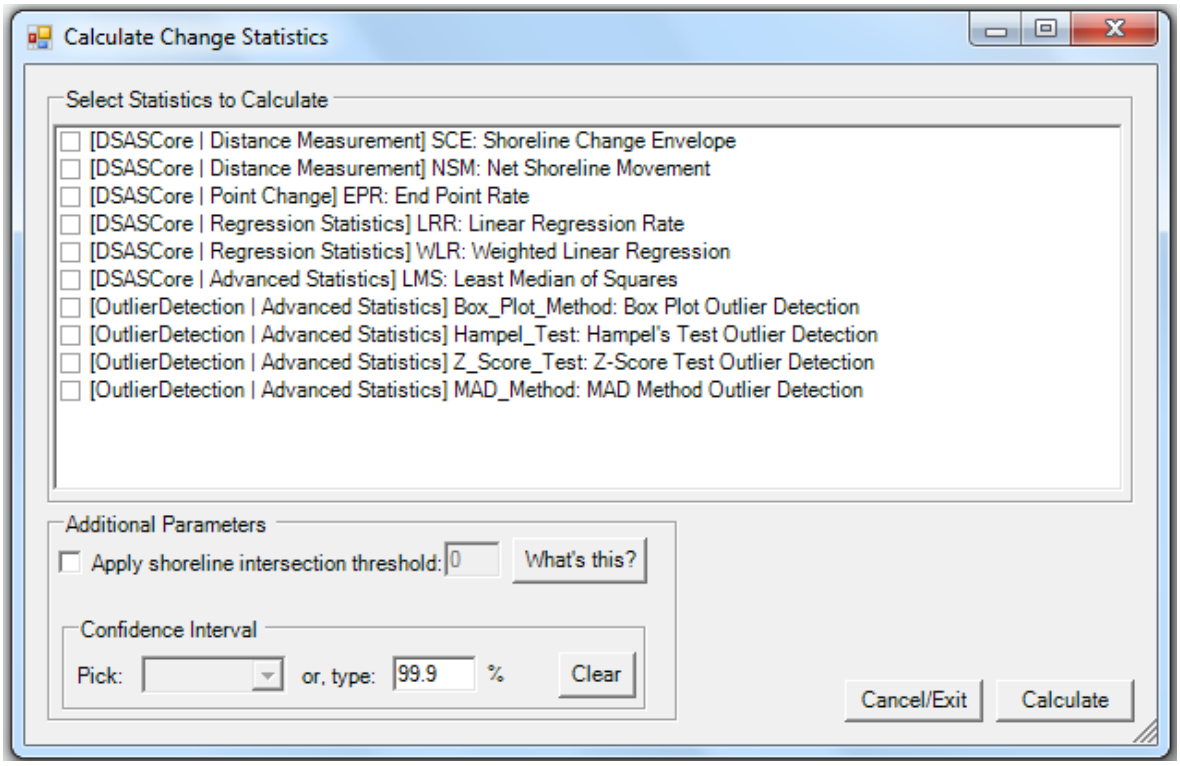
- فایل XML ورودی را به صورت فقط خواندنی[20] از ورودی دریافت می‌کند. این ویژگی یکی از قابلیت‌های ماژول محسوب می‌شود، چون این قابلیت سبب می‌شود در صورتی که سایر ماژول‌ها به این فایل نیاز داشته باشند، بتوانند بدون هیچ‌گونه مشکی به آن دسترسی پیدا کنند.
- پویا[21] نیست. در واقع مقادیر محاسبه شده را برای استفاده‌های آتی ذخیره نمی‌کند. چون با توجه به ساختار ماژول نیازی به پویا بودن آن نیست.

[2] این کامپایلر به صورت رایگان در سایت MathWorks در دسترس بوده و از آدرس زیر قابل دانلود است:
http://www.mathworks.com/products/compiler/mcr

[Downloaded from marine-eng.ir on 2026-11-12]
[DOR: 20.1001.1.17357608.1394.11.21.7.1]

- به موازات و بدون هیچ‌گونه ناسازگاری با سایر برنامه‌ها بوسیله DSAS قابل اجرا است.
- توانایی ذخیره و پردازش مجموعه داده‌های با ابعاد بزرگ را دارا است.
- سازگاری با نسخه‌های مختلف ArcGIS و DSAS.
- یک فایل اجرایی با پسوند exe است.
- کامپایل شده با MCR نسخه ۸.۱
- برای اجرای آن باید MCR نسخه ۸.۱ نصب باشد.

علاوه بر تولید فایل XML، مشخصات داده‌های مشکوک به خطا شناسایی شده توسط روش‌ها را در قالب یک فایل اکسل به صورت جداگانه و خارج از محیط ArcGIS در اختیار کاربر قرار می‌دهد.



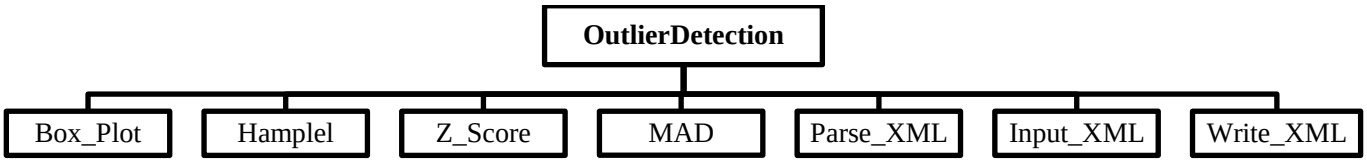
شکل ۱۰- پنجره محاسبات آماری DSAS و گزینه‌های اضافه شده توسط ماژول به آن

برای کد نویسی ماژول در MATLAB از چندین تابع استفاده شده است، که هر کدام از این توابع وظیفه‌ای بر عهده دارند و بخشی از وظیفه کلی ماژول را انجام می‌دهند. در شکل ۱۱ نام این توابع به همراه سلسله مراتب آنها نشان داده شده است. وظیفه هر یک از این توابع به شرح زیر است:

- تابع OutlierDetection: تابع اصلی برنامه بوده که فراخوانی سایر توابع با استفاده از آن انجام می‌شود.
- تابع Input_XML: فایل XML تولیدی ماژول DSASCore.calc را از مسیر تعیین شده می‌خواند.
- تابع Write_XML: حاصل محاسبات ماژول را به فرم XML تبدیل کرده و در مسیر تعیین شده ذخیره می‌کند.
- تابع Parse_XML: فایل XML ورودی را به نوع داده ساختار[۲۲] با فیلدهای Name، Attributes، Data و Children تجزیه می‌کند. علت ایجاد این تابع، استخراج مقادیر موجود در فایل XML برای استفاده در MATLAB است.
- تابع Box_Plot: داده‌های مشکوک به خطا موجود در مجموعه داده ورودی را با استفاده از روش نمودار جعبه‌ای شناسایی می‌کند. خروجی‌های این تابع مقادیر داده‌های مشکوک به خطا و اندیس آنها است.
- تابع Hampel: داده‌های مشکوک به خطا موجود در مجموعه داده ورودی را با استفاده از روش هامپل شناسایی می‌کند. خروجی‌های این تابع مقادیر داده‌های مشکوک به خطا و اندیس آنها است.
- تابع Z_Score: داده‌های مشکوک به خطا موجود در مجموعه داده ورودی را با استفاده از روش Z-Score شناسایی می‌کند. خروجی‌های این تابع مقادیر داده‌های مشکوک به خطا و اندیس آنها است.

[Downloaded from marine-eng.ir on 2026-11-12]

[DOR: 20.1001.1.17357608.1394.11.21.7.1]



شکل ۱۱- توابع بکار گرفته شده در کد نویسی ماژول

- تابع MAD: داده‌های مشکوک به خطا موجود در مجموعه داده ورودی را با استفاده از روش MAD شناسایی می‌کند. خروجی‌های این تابع مقادیر داده‌های مشکوک به خطا و اندیس آنها است

۴- نتایج

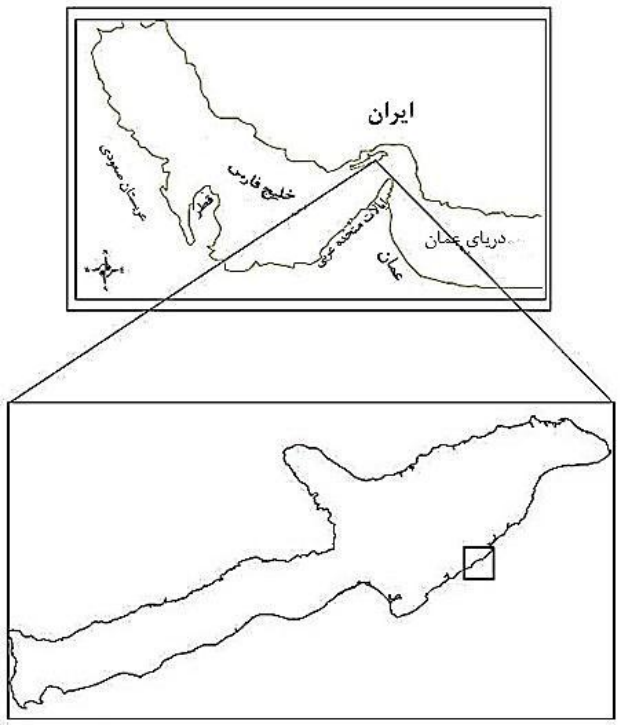
در این بخش کارائی ماژول طراحی شده روی داده‌های مربوط به نرخ تغییرات خط ساحلی یک مجموعه داده واقعی بررسی شده است. مجموعه داده مورد بررسی حاصل کمی‌سازی نرخ تغییرات خط ساحلی منطقه سوزا (واقع در جزیره قشم) است. در شکل ۱۲ موقعیت این منطقه روی نقشه نشان داده شده است. همانطور که از این شکل مشخص است، این ناحیه ساحلی در قسمت جنوبی جزیره قشم واقع شده و بوسیله دو ساحل صخره‌ای (تراس‌های دریایی بالارونده) از سایر سواحل شنی مجزا شده است. شهر ساحلی سوزا یکی از مناطق پر جمعیت بوده که در سال‌های اخیر با توسعه امکانات، جمعیت زیادی را به سمت خود جذب کرده است. در جدول ۱ مشخصات ناحیه ساحلی انتخاب شده برای بررسی نرخ تغییرات خط ساحلی آن، ارائه شده است.

جدول ۱- مشخصات ناحیه ساحلی مورد بررسی

جنس	شنی
طول انتخابی	۱/۶۳ کیلومتر
محل	قسمت جنوبی جزیره قشم
شکل مورفولوژی	پلاژهای ماسه ای کم شیب[۲۳]

اندازه‌گیری نرخ تغییرات خط ساحلی به کمک ابزار DSAS و در محیط نرم‌افزار ArcGIS انجام شده است. منابع داده‌ای که خط ساحلی از آنها استخراج شده، عبارتند از: عکس‌های هوایی اسکن شده با کیفیت DPI[۲۴] ۵۰۰، عکس‌های ماهواره‌ای با دقت بالا شامل GeoEye، QuickBird و panchromatic IRS. تمامی این داده‌ها زمین مرجع شده و خطای تخمینی آنها محاسبه شده است. برای زمین مرجع کردن عکس‌های هوایی و تصاویر ماهواره‌ای از نقاط کنترل زمینی استفاده می‌شود. امروزه یکی از رایجترین روش‌ها برای تصحیح و زمین مرجع نمودن داده‌ها استفاده از سیستم موقعیت‌یاب جهانی یا GPS است. بدون شک عملیات میدانی با دستگاه GPS دقیق‌ترین و قابل اعتمادترین نقاط کنترلی را می‌تواند فراهم کند. در این تحقیق برای زمین مرجع کردن داده-ها از فن‌آوری GPS با دقت ۳± متر استفاده شده است.

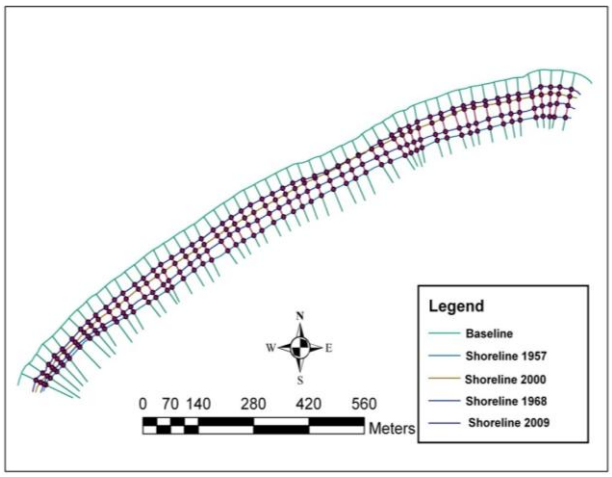
در شکل ۱۳مقاطع عرضی ایجاد شده توسط DSAS، در محیط ArcGIS نشان داده شده است. فاصله بین تمامی مقاطع ۳۰ متر در نظر گرفته شده که در مجموع باعث تولید ۵۴ مقطع شده است. با استفاده از داده‌های موجود، در مجموع ۴ خط ساحلی استخراج شده که متعلق به سال‌های ۱۹۵۷، ۱۹۶۸، ۲۰۰۰ و ۲۰۰۹ میلادی است. بنابراین نرخ تغییرات خط ساحلی طی ۵۲ سال اندازه‌گیری شده است. پس از ایجاد مقاطع، نرخ تغییرات خط ساحلی حول مقاطع ایجاد شده نسبت به خط مبنا، با استفاده از روش‌های EPR و LRR محاسبه شد. در شکل ۱۴ نرخ‌های اندازه‌گیری شده نشان داده شده است.



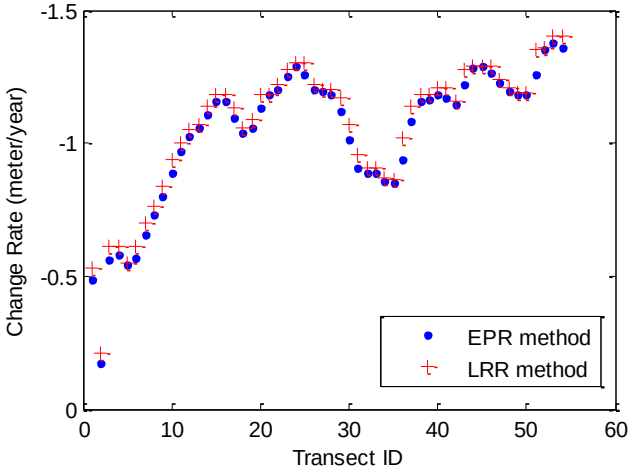
شکل ۱۲- موقعیت منطقه سوزا روی نقشه (علامت مربع)

[Downloaded from marine-eng.ir on 2026-11-12]

[DOR: 20.1001.1.17357608.1394.11.21.7.1]



شکل ۱۳- مقاطع عرضی ایجاد شده توسط DSAS، در ناحیه خط ساحلی سوزا



شکل ۱۴- نرخ تغییرات خط ساحلی سوزا به روش EPR و LRR

در جدول ۲ حاصل اجرای ماژول روی مجموعه داده EPR، و در جدول ۳ حاصل اجرای آن روی مجموعه داده LRR ارائه شده است. در این جداول، مقدار و اندیس داده‌های مشکوک به خطا شناسایی شده با استفاده از روش‌های آماری مورد بحث، قرار داده شده است.

جدول ۲- داده‌های مشکوک به خطا در مجموعه داده EPR

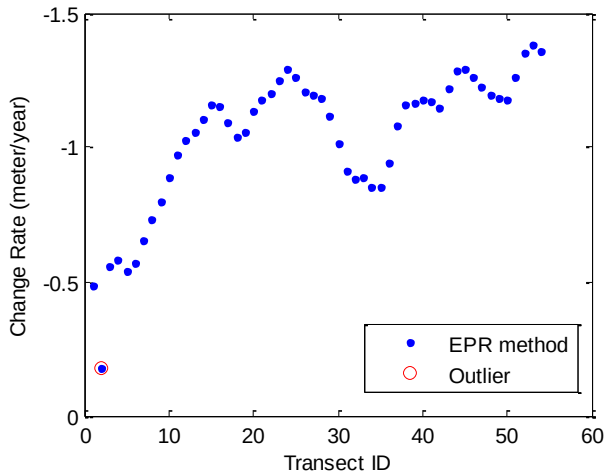
روش	اندیس	مقدار
نمودار جعبه‌ای	۲	-۰/۱۷
آزمون هامپل	۱، ۲، ۵	-۰/۴۸، -۰/۱۷، -۰/۵۴
آزمون Z-Score	۲	-۰/۱۷
آزمون انحراف مطلق میانه	۲	-۰/۱۷

جدول ۳- داده‌های مشکوک به خطا در مجموعه داده LRR

روش	اندیس	مقدار
نمودار جعبه‌ای	۲	-۰/۲۱
آزمون هامپل	۱، ۲، ۵	-۰/۵۳، -۰/۲۱، -۰/۵۵
آزمون Z-Score	۲	-۰/۲۱
آزمون انحراف مطلق میانه	۲	-۰/۲۱

با توجه به نتایج موجود در جداول ۲ و ۳ می‌توان به موارد زیر اشاره کرد:

- حاصل اجرای ماژول روی هر دو مجموعه داده کاملاً یکسان بوده که علت آن، تفاوت بسیار ناچیز مقادیر موجود در هر مجموعه داده است. به عنوان مثال، میانگین قدرمطلق اختلاف دو به دوی اعضای مجموعه داده ۰/۰۲۹ می‌باشد. همانطور که مشخص است، مقدار این اختلاف ناچیز است.
- هر چهار روش، در هر دو مجموعه داده به مشکوک به خطا بودن داده با اندیس ۲ اتفاق نظر دارند. این داده در شکل ۱۵ با یک دایره به دور آن در مجموعه داده EPR نشان داده شده است. همانطور که از شکل نیز مشخص است، این داده به طرز مشکوکی ایجاد شده و اختلاف آن با سایر اعضای همسایه آن زیاد است. بنابراین می‌تواند یک داده خطا باشد. در این نمونه پس از بررسی‌های انجام شده مشخص شد که علت ایجاد آن، مقطع عرضی نامناسب بوده است. در واقع مقطع ایجاد شده برای این قسمت توسط DSAS، به درستی عمود بر خطوط ساحلی نبوده است. پس از تصحیح مقطع مورد نظر و محاسبه نرخ تغییرات از ابتدا، خطای این داده بر طرف شد. مقدار جدید نرخ تغییر خط ساحلی حول این مقطع به روش EPR برابر ۰/۴۷۵-، و به روش LRR برابر ۰/۵۱- محاسبه شد، که اختلاف این مقادیر با همسایه‌هایشان بسیار ناچیز است.

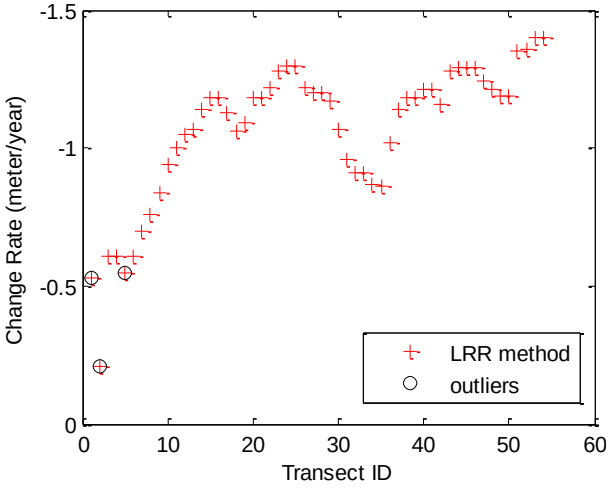


شکل ۱۵- داده مشکوک به خطا در مجموعه داده EPR به روش نمودار جعبه‌ای

[Downloaded from marine-eng.ir on 2026-11-12]

[DOR: 20.1001.1.17357608.1394.11.21.7.1]

- آزمون هامپل در هر دو مجموعه داده به غیر از داده با اندیس ۲، داده‌های با اندیس ۱ و ۵ را نیز به عنوان مشکوک به خطا شناسایی کرده است. این داده‌ها در شکل ۱۶ با یک دایره به دور آنها مشخص شده‌اند. همانطور که قبلاً توضیح داده شد، داده با اندیس ۲ یک داده خطا است اما دو داده دیگر خطا نیستند. چون همانطور که در شکل مشخص است، اختلاف این داده‌ها با همسایه‌هایشان بسیار ناچیز است و نمی‌توانند یک داده خطای واقعی باشند. بنابراین در این دو مورد تشخیص آزمون هامپل اشتباه بوده است.



شکل ۱۶- داده مشکوک به خطا در مجموعه داده LRR به روش هامپل

در حالت کلی با توجه به مطالب بیان شده می‌توان گفت همواره داده‌هایی که توسط روش‌ها به عنوان کاندیدای خطا انتخاب می‌شوند، بیانگر وجود خطا در داده‌ها نیستند و ممکن است بر اساس ماهیت عملکرد هر روش، به عنوان خطا تشخیص داده شده باشند. بنابراین پس از شناسایی داده‌های مشکوک به خطا، باید در خصوص علت ایجاد آنها تحقیق کرده و در صورتی که واقعاً بر اثر بروز خطایی رخ داده باشند، نسبت به تصحیح و یا حذف آنها اقدام کرد. همچنین پیشنهاد می‌شود بیشتر داده‌هایی مورد توجه قرار گیرند که توسط اکثر روش‌ها به عنوان داده خطا انتخاب شده‌اند. چرا که این داده‌ها معمولاً دارای پتانسیل بیشتری از جنبه خطا بودن، هستند. پس از برطرف کردن خطاهای احتمالی در داده‌ها، داده‌های نهایی از صحت بیشتری برخوردار بوده و نتایج حاصله از آنها دارای اعتبار بیشتری است.

۵- نتیجه‌گیری

خطا جز جدایی‌ناپذیر هر کمیت اندازه‌گیری است. اندازه‌گیری نرخ تغییرات خط ساحلی نیز از این قاعده مستثنا نیست. در این تحقیق یک ماژول جدید برای شناسایی داده‌های مشکوک به خطا در داده‌های تغییرات خط ساحلی ارائه شد. اساس عملکرد این ماژول، استفاده از روش‌های آماری شناسایی داده‌های مشکوک به خطا است. کارایی این ماژول روی یک مجموعه داده واقعی نرخ تغییرات خط ساحلی مورد بررسی قرار گرفت. نتایج حاکی از عملکرد مناسب این ماژول در شناسایی خطاهای احتمالی است. مولفین پیشنهاد می‌کنند قبل از استفاده از مقادیر اندازه‌گیری شده خط ساحلی برای تجزیه و تحلیل‌های آتی، از این ماژول برای شناسایی خطاهای احتمالی موجود در داده‌ها استفاده شود؛ چرا که این امر سبب افزایش دقت نتایج حاصله می‌شود.

۶- تشکر و قدردانی

نویسندگان مقاله مراتب تشکر و قدرانی خود را از جناب آقای دکتر محمد جواد کتابداری عضو هیئت علمی دانشکده مهندسی دریا دانشگاه صنعتی امیرکبیر بخاطر توصیه‌ها و راهنمایی‌های ارزشمند ایشان در روند نگارش مقاله اعلام می‌دارند.

کلید واژگان

1- Digital Shoreline Analysis System (DSAS)
2- Extension
3- Transect
4- Baseline
5- End Point Rate
6- Linear Regression Rate
7- Weighted Linear Regression Rate
8- Confidence of End Point Rate
9- Environmental Systems Research Institute (ESRI)
10- Module
11- Box Plot
12- Hampel's Test
13- Median Absolute Deviation (MAD)
14- Inter Quartile Range (IQR)
15- MATLAB Compiler Runtime (MCR)
16- Executable file
17- Binary
18- Registered Handler
19- Extensible Markup Language (XML)
20- Read only
21- Dynamic
22- Structure
23- Flat Sandy Beach (FSB)
24- Dot Per Inch (DPI)

۷- مراجع

1- Thieler, E.R., Himmelstoss, E.A., Zichichi, J.L., and Ergul, Ayhan, (2009), *Digital Shoreline Analysis System (DSAS) version 4.0 — An ArcGIS extension for calculating shoreline change*, U.S. Geological Survey Open-File Report 2008-1278.

2- Moore, L.J., (2000), *Shoreline mapping techniques*, *Journal of Coastal Research*, Vol. 16, p. 111-124.

[Downloaded from marine-eng.ir on 2026-11-12]

[DOR: 20.1001.1.17357608.1394.11.21.7.1]

3- http://www.usgs.gov/
4- Liu, J. K., (1998), *Developing Geographic Information System Applications in Analysis of Responses to Lake Erie Shoreline Changes*, M.S. thesis, The Ohio State University.
5- Galgano, F., and Douglas, B., (2000), *Shoreline Position Prediction: Methods and Errors*, *Environmental Geosciences*, Vol. 7 (1), p. 23-31.
6- Hawkins, D., (1980), *Identification of Outliers*, Chapman and Hall, London.
7- Singh, K., and Upadhyaya, S., (2012), *Outlier Detection: Applications And Techniques*, IJCSI International Journal of Computer Science Issues, Vol. 9, Issue 1, No 3, p. 307-323.
8- Hodge, V., and Austin, J., (2004), *A survey of outlier detection methodologies*, *Artificial Intelligence Review*, Vol. 22 (2), p. 85-126.
9- Barnett, V., and Lewis, T., (1992), *Outliers in Statistical Data*, New York, NY, John Wiley & Sons., 3rd edition.
10- Tukey, J.W., (1977), *Exploratory data analysis*, Addison-Wesley publication, p. 205-235.
11- McGill, R., J. W. Tukey., and W. A. Larsen., (1978), *Variations of Boxplots*, The American Statistician.Vol. 32, No. 1, p. 12–16.
12- Redman, T. C., (2001), *Data Quality*, The Field Guide, Boston Digital Press.
13- BEN-GAL, I., (2005), Outlier detection, *Data mining and knowledge discovery handbook: A complete guide for practitioners and researchers*, Kluwer Academic Publishers.
14- Schiffle, R.E., (1988), *Maximum Z Score and outliers*, The American Statistician, Vol. 42, No.1, p. 79-80.
15- Iglewicz, B., and Hoaglin, D., (1993), *How to detect and handle outliers*, ASQC Quality Press.

[Downloaded from marine-eng.ir on 2026-11-12]

[DOR: 20.1001.1.17357608.1394.11.21.7.1]