

خلاصه‌سازی عوارض خطی بر مبنای رگرسیون کمترین مربعات معمولی

جلیل جعفری^{۱*}، محمد سعدی مسگری^۲

^۱ کارشناس ارشد سیستم‌های اطلاعات مکانی - دانشکده مهندسی نقشه‌برداری - دانشگاه صنعتی خواجه نصیرالدین

طوسی

j.jafarikntu@yahoo.com

^۲ دانشیار دانشکده مهندسی نقشه‌برداری - دانشگاه صنعتی خواجه نصیرالدین طوسی

mesgari@kntu.ac.ir

(تاریخ دریافت اسفند ۱۳۹۹، تاریخ تصویب تیر ۱۴۰۰)

چکیده

خلاصه‌سازی عوارض برداری جهت تهیه نقشه‌های کوچک مقیاس کاغذی و برداری از نقشه‌های بزرگ مقیاس، ضروری می‌باشد. این فرآیند منجر به بهبود کیفیت نمایش نقشه‌ها، تسهیل آنالیز داده‌ها در سطوح مختلفی از جزئیات و کاهش حجم مورد نیاز برای ذخیره‌سازی آنها می‌گردد. روش‌های خلاصه‌سازی عوارض خطی و چند ضلعی با هدف حفظ هندسه و مساحت در عین تقلیل جزئیات آنها انجام می‌شود. مدل‌های گوناگونی از سوی محققان این حوزه مورد استفاده و ارزیابی قرار گرفته است؛ لیکن بیشتر آنها خلاصه‌سازی عوارض را با هدف انتخاب چند نقطه از آنها و حذف سایر نقاط انجام می‌دهند. حال آنکه ممکن هست نقاط حذف شده حاوی اطلاعات ارزشمندی برای آن عارضه بوده و حذف آنها منجر به نقصان در هندسه و مساحت آن گردد. در پژوهش حاضر، خلاصه‌سازی عوارض خطی با کمینه کردن فاصله عمودی از خط اصلی، به کمک رگرسیون کمترین مربعات معمولی انجام گرفت. برای بررسی مدل پیشنهادی، پس از پیاده‌سازی آن بر شکل‌های مختلف، حدود دریاچه ارومیه و برخی جزایر آن خلاصه‌سازی شده و نتایج مدل پیشنهادی با روشهای متداول داگلاس-پوکر و ویسوالینگم مورد مقایسه قرار گرفت. سپس برای ارزیابی نتایج از شاخص‌های اختلاف مساحت‌ها، شباهت انحنای متوسط، شباهت میزان تغییرات زاویه و فاصله هاسدرف میانه تصحیح شده استفاده شد. نتایج حاکی از برتری مدل پیشنهادی براساس سه شاخص اول، به ترتیب با میزان شباهت متوسط ۶۹،۹۱، ۶۶،۲۹ و ۶۰،۹۹ درصدی است. مدل پیشنهادی نسبت به داگلاس-پوکر و ویسوالینگم دارای برتری ۰،۱۶ و ۰،۲ درصدی براساس شاخص اختلاف مساحت، برتری ۷ و ۵ درصدی براساس شاخص انحنای متوسط و ۶ و ۲ درصدی براساس شاخص تغییرات تندی است. اما بر اساس شاخص فاصله هاسدرف میانه تصحیح شده، به طور متوسط در حدود ۲ متر بدتر از روشهای مذکور عمل کرده است که به دلیل عدم اتکا به نقاط اولیه عارضه می‌باشد.

واژگان کلیدی: خلاصه‌سازی، کمترین مربعات، داگلاس-پوکر، ویسوالینگم، رگرسیون

۱- مقدمه

همزمان با افزایش حجم داده‌ها و اطلاعات مکانی، ذخیره‌سازی، انتقال، پردازش و نمایش عوارض و اطلاعات مکانی به چالشی بزرگ و همه‌گیر در علم کارتوگرافی و به طور خاص، سیستم اطلاعات مکانی مبدل گشته است [۱]. از طرفی، واسطه‌های گرافیکی نرم‌افزارهای مکانی نمی‌توانند جهت نمایش نقشه‌های کوچک مقیاس کلیه جزئیات را نگه دارند و از طرف دیگر حجم زیاد اطلاعات مکانی، ذخیره‌سازی، انتقال، سرعت پردازش‌ها و گاهاً دقت پردازش‌ها را نیز به دلایلی مانند وجود نویز دچار مشکل می‌کند. راهکار عمومی که به منظور رفع این مشکلات در نظر گرفته می‌شود، خلاصه‌سازی (جنرالیزاسیون) عوارض و به دنبال آن تعمیم نقشه است. خلاصه‌سازی به دو دسته کلی مدل‌گرا و کارتوگرافی تمیز داده می‌شود [۱، ۲]. نوع کارتوگرافی آن عموماً بر وجه بصری‌سازی اطلاعات مکانی تمرکز دارد. درحالی‌که خلاصه‌سازی مدل‌گرا، که خود پیش‌پردازشی برای نوع کارتوگرافی آن به حساب می‌آید، کاهش تراکم و حجم و ساده‌سازی اطلاعات مکانی، براساس میزان جزئیات مورد نیاز را هدف قرار می‌دهد. خلاصه‌سازی مدل‌گرا نیز خود شامل خلاصه‌سازی روابط توپولوژی، هندسه عوارض، اطلاعات معنایی و دیتابیس می‌باشد [۳]. روشهای متنوعی برای هریک از اهداف مذکور توسعه داده شده است که از جمله آنها می‌توان به روشهای داگلاس-پوکر و ویسوالینگم اشاره کرد. روشهای داگلاس-پوکر و ویسوالینگم از منطق‌های مختلفی برای خلاصه‌سازی هندسی عوارض استفاده می‌کنند اما در نهایت هدف اصلی آنها، انتخاب تعدادی از نقاط خط اصلی و نگه داشتن آنهاست. به بیان دیگر، در این روشها، نقاط حذف شده به کلی نادیده گرفته می‌شوند و هیچ تاثیری در خط نهایی خلاصه شده ندارند. می‌توان اذعان داشت که وجود یا عدم وجود این نقاط بر نتیجه الگوریتم‌های یاد شده تاثیری ندارد. این در حالیست که این نقاط جزو عارضه اصلی بوده و ممکن است دارای اهمیت باشند و با نادیده گرفتن آنها، شباهت هندسی عارضه خلاصه‌شده به عارضه اصلی کاهش یابد [۴]. بنابراین مدل پیشنهادی این تحقیق، بر اساس رگرسیون کمترین مربعات می‌باشد. بدین ترتیب که جهت برازش خط، همه نقاط را در نظر می‌گیرد و در نتیجه تمام نقاط خطوط عارضه مورد نظر،

در برازش خط خلاصه‌شده تأثیر دارند. هدف این پژوهش ارائه رویکردی نوآورانه جهت خلاصه‌سازی هندسه عوارض خطی و مقایسه عملکرد آن نسبت به پرکاربردترین الگوریتم‌های مورد استفاده در این حوزه می‌باشد.

۲- پیشینه تحقیق

خلاصه‌سازی عوارض، عملیاتی هوشمند، بسیار پیچیده و فنی است. اهمیت و پیچیدگی آن متخصصان زیادی را در دهه‌های اخیر به ارائه رویکردها و الگوریتم‌های مختلف واداشته است. هدف تمامی این الگوریتم‌ها کاهش عناصر هندسی عوارض (مانند: نقطه و خط) می‌باشد [۵]. به نحوی که خروجی الگوریتم با داشتن حداقل تعداد گره (و یال)، حداکثر شباهت را به هندسه اصلی عارضه دارا باشد [۶]. در این حیطه، ابتدایی‌ترین رویکردها بر پایه روابط ریاضی و همسایگی خاصی نبودند. آن‌ها به طور مستقل گره‌هایی را انتخاب کرده و مابقی را کنار می‌گذاشتند. به عنوان نمونه، رویکرد random points پس از تقسیم عارضه به قطعات مختلف با گره‌های متوالی، از هر دسته به طور تصادفی گرهی را انتخاب می‌کرد [۶] و همچنین رویکرد n^{th} points از هر دسته گره ln ام را برمی‌گزید و در نهایت مجموعه گره‌های برگزیده، با رعایت ترتیب، هندسه ساده‌سازی شده را تشکیل می‌دادند [۷]. دسته دیگر از الگوریتم‌ها با یک پله پیشرفت، به صورت محلی رابطه میان هر دو یا سه گره متوالی را از نظر فاصله بررسی می‌کردند [۸]. "رویکرد فاصله میان نقاط" با در نظر گرفتن شعاع همسایگی، پردازش را از گره اول آغاز می‌کرد و برای هر نقطه (گره)، نقاط متوالی بعدی داخل شعاع همسایگی را کنار می‌گذاشت. "رویکرد فاصله عمودی" نیز هر نقطه را توسط یک خط، به ترتیب به دومین نقطه پس از خود متصل و فاصله عمودی نقطه میانی از خط مذکور را محاسبه می‌کرد. در صورتی که این فاصله از یک تلورانس از پیش تعیین شده بیشتر می‌بود، نقطه میانی نگه داشته می‌شد و در غیر این صورت حذف می‌گشت. روش‌های دیگری نیز مطرح شدند که مانند الگوریتم‌های مذکور، به صورت محلی پردازش‌ها را انجام می‌داند، با این تفاوت که گستره آن در محدوده هندسه عارضه می‌توانست تغییر کند و در هر مرحله، به صورت افزایشی نقاط بیشتری دخالت داده شوند. از میان آن‌ها

می‌توان به الگوریتم Reumann-Witkam اشاره کرد. در این روش، الگوریتم با شروع از نقطه اول، با اتصال هر نقطه به نقطه پس از خود تشکیل یک امتداد می‌داد. سپس باندی با پهنای دلخواه و در راستای امتداد مذکور در نظر گرفته می‌شد. با بررسی متوالی نقاط، نقطه قبل از اولین نقطه‌ای که خارج از این باند قرار می‌گرفت نگه داشته می‌شد و مابقی کنار گذاشته می‌شدند. پردازش به طور مجدد از نقطه نگه داشته شده شروع و تا آخرین نقطه از هندسه اولیه عارضه ادامه می‌یافت [۹]. روش دیگر از این دست، الگوریتم sleeve-fitting است. این الگوریتم مشابه Reumann-Witkam عمل می‌کند. با این تفاوت که در هر مرحله از تشکیل باند، امتداد آن را توسط اولین و آخرین نقطه مشخص می‌کند و هر نقطه‌ای که بیرون از باند بیوفتد، آخرین نقطه از مرحله قبل انتخاب شده و به جز نقطه اول مابقی کنار گذاشته می‌شوند [۱۰]. Opheim (۱۹۸۱ و ۱۹۸۲) در روش خود، با تعریف مقدار کمینه و بیشینه برای طول باند در الگوریتم Reumann-Witkam، حداقل طول باند و همچنین میزان گسترش طول آن را محدود کرد [۱۱]. Lang (۱۹۶۹) نیز محدوده جست‌وجو را بر اساس فاصله عمودی خط اتصال دو نقطه از هندسه عارضه اولیه، با نقاط بین آن‌ها را تنظیم ساخت. در این روش نیز محدوده جست‌وجو به تعداد ثابتی از نقاط محدود گشت (برای مثال n نقطه). جست‌وجو از n نقطه اول آغاز می‌گردد، به این صورت که نقطه اول و انتهای بازه تعیین شده توسط یک خط به یکدیگر متصل می‌شوند و فاصله عمودی نقاط میانی از آن خط بررسی می‌گردد. اگر همگی دارای فاصله کمتر از تلورانس باشند هیچیک از آن‌ها نگه داشته نمی‌شوند، در غیر این صورت محدوده جست‌وجو کوچک تر شده و عملیات تکرار می‌گردد. اولین باری که نقطه‌ای کنار گذاشته شد و یا تعداد نقاط میانی به صفر رسید، نوبت به بررسی n نقطه‌ی بعدی می‌شود. به طوری که اولین نقطه آن‌ها، آخرین نقطه نگه داشته شده از جست‌وجوهای قبل باشد [۱۲]. در سال ۱۹۷۳ داگلاس و پوکر رویکردی را مطرح ساختند که به یکی از رایج‌ترین الگوریتم‌های خلاصه‌سازی هندسه عوارض میدل گشت. ۱۹ سال بعد (۱۹۹۲) Hershberger و Snoeyink آن را بهبود بخشیدند [۱۳]. وجه تمایز این الگوریتم و رویکردهای مشابه، سراسری بودن ناحیه جست‌وجوی آن است. بدین صورت که در ابتدای کار، اولین و آخرین نقطه

توسط خطی فرضی به یکدیگر متصل می‌گردند. اگر فاصله تمامی نقاط تا خط مفروض کمتر از میزان تلورانس باشد، تمامی نقاط میانی کنار گذاشته شده و پردازش به اتمام می‌رسد. در غیر این صورت، نقطه‌ای که دارای بیشترین فاصله از خط مفروض است نگه داشته شده و در مرحله بعد ناحیه جست‌وجو، از آن نقطه به دو قسمت تقسیم می‌شود و هریک مجدداً به صورت مجزا بررسی می‌شوند. این عملیات تاجایی که تمامی نقاط در نواحی خود دارای فاصله‌ی عمودی کمتر از تلورانس مشخص باشند، تکرار می‌گردد. مطابق آنچه بیان گشت، الگوریتم داگلاس-پوکر بر اساس فاصله عمودی عمل می‌کند و مساحت شکل را در نظر نمی‌گیرد. برای رفع این مشکل، Visvalingam و Whyatt (۱۹۹۳) روشی را مطرح کردند که بر پایه مساحت موثر به خلاصه‌سازی می‌پرداخت. طبق این روش، در هر تکرار مساحت موثر برای هر گره (مساحت مثلث متشکل از هر گره با دو همسایه خود) محاسبه می‌گردد و گره دارای کمترین مساحت موثر، حذف می‌گردد. این روند تا جایی که به تعداد گره دلخواه رسیده شود ادامه می‌یابد [۱۴]. همچنین می‌توان در روش Visvalingam و Whyatt از حد آستانه مساحت استفاده کرد [۱۵]. به طوری که پس از اعمال الگوریتم، هیچ مساحت موثر کمتر از حد آستانه تعیین شده موجود نباشد. در سال‌های اخیر نیز محققان رویکردهای خلاقانه‌ای را ارائه دادند. به عنوان نمونه Yilang Shen و همکاران [۱۶] (۲۰۱۸) با استفاده از تکنیک‌های پردازش تصویر، ابتدا خط را به فضای رستر بردند. به طوری که پیکسل‌هایی که خط از آن‌ها عبور می‌کرد دارای ارزش یک و دیگر پیکسل‌ها صفر بودند. حال در کانال رستری به دست آمده (پیکسل‌های دارای ارزش یک) چندضلعی دارای کمترین طول ممکن را با بهم متصل کردن تعدادی خطوط مستقیم به دست آوردند. Tinghua Ai و همکاران (۲۰۱۷) نیز [۱۷] در روشی مبتکرانه ابتدا روی گره‌های هندسه اصلی عارضه مثلث‌بندی انجام دادند. سپس با توجه به تعداد همسایگی‌های مثلث‌های به دست آمده، به آن‌ها نوعی درجه اهمیت اطلاق کردند و از این طریق نهایتاً به گره‌های مهم می‌رسیدند. Türkay Gökgöz و همکاران (۲۰۱۵) نیز با این هدف که خروجی الگوریتم باید در تلورانسی از هندسه اصلی عارضه باشد، یک بافر به عنوان باند خطا (error band) در نظر گرفتند و با تغییر

زوایای انحراف، خطوط پشت سرهم را طوری ساده کردند که از محدوده باند خطا خارج نشوند [۱۸]. آنچه که بیان شد، پاره‌ای از رویکردهای مطرح شده توسط محققان مختلف سراسر دنیا بود. بدیهی است که نام بردن تمامی آن‌ها خارج از مجال این مقاله می‌باشد. لذا سعی شد پاره‌ای از سیر تاریخی این الگوریتم‌ها، روش‌های محبوب و پرستفاده و همچنین تعدادی از رویکردهای خلاقانه سال‌های اخیر بیان شوند تا یک دید کلی از آن‌ها به دست آید. با توجه به آنچه بیان گشت، اکثر الگوریتم‌های مطرح شده سعی دارند با ارائه روش‌هایی، گره‌های پراهمیت از هندسه عارضه را شناسایی کرده و نگه دارند. در این روش‌ها برخی گره‌ها به کلی نادیده گرفته می‌شوند، در حالی که بخشی از ماهیت هندسه عارضه‌ی هدف هستند و نادیده گرفتن آن‌ها می‌تواند باعث از دست رفتن اطلاعات هندسی و در نتیجه افت کیفیت خلاصه‌سازی گردد.

۲- مبانی نظری

روشهای مختلفی برای ساده‌سازی، خلاصه‌سازی و فشرده‌سازی عوارض مختلف وجود دارد. با توجه به مدل پیشنهادی این تحقیق و مقایسه آن با الگوریتم‌های متداول، مفاهیم مربوط به کمترین مربعات^۱، الگوریتم داگلاس-پوکر^۲ و ویسوالینگم^۳ در این بخش شرح داده شده است.

۲-۱- الگوریتم داگلاس-پوکر

الگوریتم داگلاس-پوکر یکی از شناخته شده‌ترین الگوریتم‌های خلاصه‌سازی می‌باشد. این الگوریتم با کاهش تعداد نقاط خطوط، به ساده‌سازی آنها می‌پردازد. به عبارت بهتر در این روش، تمام نقاط خط روی پاره‌خط فرضی تصویر می‌شود و تنها با توجه به ویژگی مکانی نقاط، خلاصه‌سازی انجام می‌شود [۱۹]. در این روش، ابتدا توسط گره اول و آخر، یک خط تشکیل می‌شود. سپس فاصله تمامی گره‌ها از آن خط محاسبه می‌گردد. اگر دورترین گره از خط مذکور، دارای فاصله‌ای بیشتر از تلورانس از پیش تعیین شده باشد، آن گره نگه داشته

می‌شود و همین مراحل برای گره‌های قبل و بعد از آن مجدداً تکرار می‌گردد. این عمل تا جایی که تمامی گره‌ها در تلورانس تعیین شده باشند، تکرار می‌گردد. حال آنکه در مسائل مختلف، ویژگی‌های دیگری نظیر مساحت، زمان و توپولوژی نیز اهمیت فراوانی دارند. برای ارزیابی دقت این روش، فاصله اقلیدسی قائم هر نقطه نسبت به خط محاسبه می‌شود، سپس بر اساس میانگین یا مجموع این فواصل، دقت برآورد می‌گردد. لذا خط جدید بر اساس نقاط موجود خط اصلی ساخته می‌شود و انحراف مکانی نقاط نسبت به خط اصلی، به عنوان معیار ارزیابی عملکرد الگوریتم داگلاس-پوکر بکار می‌رود. حال آنکه این انتخاب هیچ تضمینی جهت بهترین انتخاب بودن ندارد و در مسائل مختلف با ماهیت‌های گوناگون، ممکن است به هیچ وجه قابل قبول و بهینه نباشد [۲۰].

۲-۲- الگوریتم ویسوالینگم

الگوریتم ویسوالینگم نیز روشی برای ساده‌سازی خطوط با منطق کاهش تعداد نقاط خط اصلی است. در این روش، خط خلاصه شده بر پایه نقاط خط اصلی ساخته می‌شود و فقط برای ساده‌سازی، نقاطی که در فاصله کمتری از حد آستانه مورد نظر قرار دارد، حذف می‌گردد [۲۱]. امتیاز^۴ هر نقطه، بر اساس سطح مثلثی که به خط مورد نظر اضافه می‌کند، محاسبه می‌شود. به عبارت دیگر، این روش برای نخستین بار از مفهوم مساحت موثر^۵ برای ساده‌سازی استفاده کرد. بر اساس این روش، ابتدا تمامی مثلث‌های ممکن متشکل از گره‌های عارضه خطی مد نظر تشکیل داده می‌شود. سپس از مثلث دارای کمترین مساحت، بررسی مساحت‌ها شروع می‌گردد. اگر مساحت مثلثی کمتر از تلورانس از پیش تعیین شده باشد، گره میانی آن مثلث حذف می‌شود. این عملیات تا جایی که مساحت تمامی مثلث‌ها بیشتر از تلورانس باشد، ادامه می‌یابد. همچنین محققان اذعان داشتند برای ساده‌سازی علاوه بر مساحت موثر، میزان جابجایی نواحی^۶، تکرار آنها، فاصله و زاویه را نیز می‌توان به عنوان امتیاز در نظر گرفت [۲۲، ۲۱]. با در نظر گرفتن هر کدام از این معیارها، خروجی ساده‌سازی می‌تواند بسیار متفاوت باشد که دو

^۴ Rank

^۵ Effective Area (EA)

^۶ Areal Displacement (AD)

^۱ Least Squares

^۲ Douglas-Peucker algorithm

^۳ Viswalingam

$$y = ax + b; v = ax + b - y \quad (1)$$

به تعداد نقاط مشاهده شده، معادله نوشته می‌شود. سپس برای حل این معادله‌ها به روش کمترین مربعات بایستی ضرایب مجهولات (a و b) و مشاهدات (y) از معادله-های رابطه (۱) محاسبه شود. ماتریس ضرایب مجهولات با A و ضرایب مشاهدات با L نشان داده می‌شود و از رابطه (۲) مقدار مجهولات ($\hat{X}$) محاسبه می‌گردد [۲۴].

$$\hat{X} = (A^T p A)^{-1} A^T p L \quad (2)$$

در رابطه (۲)، p ماتریس وزن نقاط می‌باشد که در صورتیکه نقاط مختلف، دارای دقت یا اهمیت‌های متفاوتی باشند، اعمال می‌گردد و در غیر اینصورت، بفرم ماتریس همانی فرض می‌گردد [۲۴، ۲۵]. مدل کمترین مربعات به دو فرم خطی و غیرخطی بوده و مقدار مربع خطای y یا مربع تصحیحات (v) را کمینه می‌نماید.

مدل کمترین مربعات معمولی^۲ نوعی از روشهای مدل کمترین مربعات خطی می‌باشد که برای تخمین پارامترهای مجهول در رگرسیون خطی مورد استفاده قرار می‌گیرد. مدل OLS مناسبترین تخمینگر خطی نالاریب می‌باشد که در واقع یک الگوی جهانی است و بر فرض ارتباطات ثابت در فضای نمونه گیری استوار است. این مدل، مجهولات را با استفاده از نقاط موجود بر روی خط تخمین میزند. رابطه (۳) نحوه محاسبه پارامتری مجهول یا وابسته (y) را به کمک پارامترهای مشاهده شده یا مستقل (Xi) نشان داده است [۲۶، ۲۷].

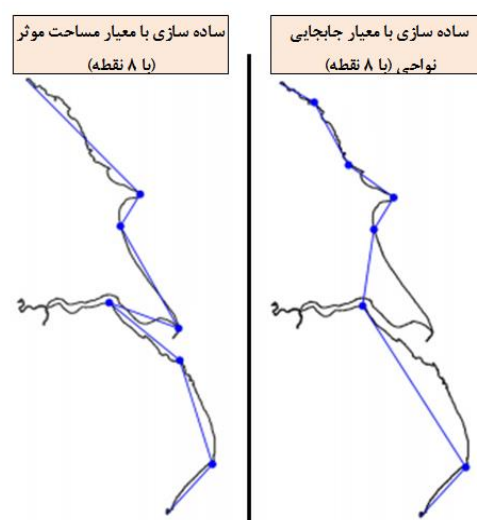
$$y = B_0(u_i + v_i) + \sum_{i=1}^p B_i(u_i + v_i) X_i + E \quad (3)$$

که در آن ui و vi مختصات نقطه i ام، $B_0(u_i + v_i)$ محل تقاطع موقعیت i ام، $B_i(u_i + v_i)$ ضریب محلی متغیر مستقل i ام و E نیز خطای تصادفی با فرض نرمال بودن داده‌ها می‌باشد [۲۷].

۳- مدل پیشنهادی

برای پیاده‌سازی مدل پیشنهادی، خط ساحلی و برخی جزایر دریاچه ارومیه انتخاب شد. دلیل اصلی این انتخاب

مورد آن در شکل ۱ نشان داده شده است. همانگونه که از شکل ۱ مشخص هست، با تغییر معیار ساده‌سازی، نتایج به شدت تغییر می‌یابد. لذا بایستی با توجه به هدف کار، معیار مناسب را انتخاب نمود و صرفاً متداول بودن یک روش، دلیلی بر بهینگی آن نیست. این الگوریتم، روشی ساده برای خلاصه‌سازی بوده و قابل تعمیم به ابعاد بالاتر است زیرا که فقط از همسایگان هر نقطه برای محاسبه امتیاز آن نقطه استفاده می‌کند. این مسئله سبب می‌شود کل طول خطوط و منحنی‌ها را بطور یکسان ساده نماید و نقاط مهم را نادیده بگیرد [۴].



شکل ۱- مقایسه ساده‌سازی بر اساس مساحت موثر و جابجایی نواحی

۲-۳- کمترین مربعات

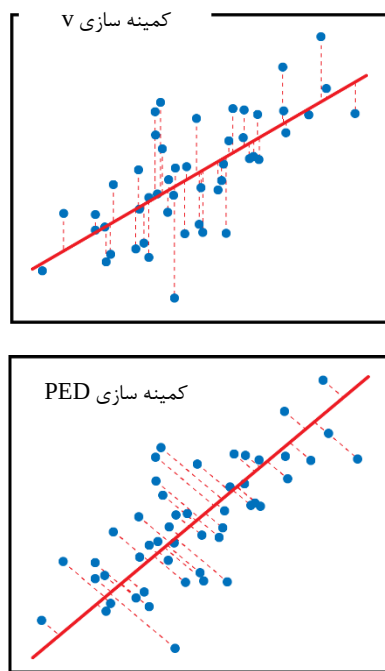
یکی از روشهای رایج تحلیل رگرسیون، روش کمترین مربعات^۱ می‌باشد که برای حل دستگاه معادلاتی به کار می‌رود که در آنها تعداد معادله مشاهده‌ها بیشتر از تعداد مجهول‌ها است. با توجه به اینکه یکی از مهم‌ترین کاربردهای روش کمترین مربعات، برازش منحنی بر داده‌ها می‌باشد، لذا در این تحقیق برای برازش خط به هریک از قطعات خطوط، مورد استفاده قرار گرفت. در صورتیکه معادله خط مطابق رابطه (۱) فرض شود که در آن x و y مقادیر مختصات نقاط هر قطعه از خط (مشاهده‌ها)، v مقدار تصحیحات، a و b ضرایب مربوط به خط (مجهول‌ها) می‌باشند که بایستی برآورد گردند و معادله خط تشکیل شود [۲۳، ۲۴].

^۲ Ordinary Least Squares (OLS)

^۱ Least Squares (LS)

پیچیدگی خطوط ساحلی و تنوع چینش نقطه و یال در آنها می باشد که در تحقیقات مختلف جهت ارزیابی روشهای خلاصه سازی بکار می روند.

برای بررسی کارایی رویکرد پیشنهادی این تحقیق، نخست تابع هدف کمینه سازی مورد بررسی قرار گرفت. سپس مراحل کار بر روی تعدادی از گره های یک عارضه خطی ساده شرح داده شد و در نهایت بر روی منطقه مطالعاتی تحقیق، پیاده شده و دقت مدل پیشنهادی با الگوریتم ویسوالینگم و داگلاس- پوکر مقایسه گردید. در شکل ۲، خط برازش داده شده به کمک کمترین مربعات بر تعدادی نقطه نشان داده شده است که برپایه معادله رابطه (۳) تشکیل یافته و با هدف کمینه سازی χ^2 اجرا شده است. اما از آنجا که در ساده سازی هندسه عوارض، فواصل اقلیدسی عمودی^۱ از اهمیت بیشتری برخوردارند [۲۹، ۲۸]، لذا در این پژوهش از OLS با هدف کمینه کردن PED به جای LS استفاده شده است که در شکل ۲ مقایسه گشته.



شکل ۲- مقایسه کمینه سازی برپایه تصحیحات χ^2 و PED

مراحل پیاده سازی مدل پیشنهادی این تحقیق به شرح ذیل می باشد.

(۱) ابتدا تلورانسی (T) بر مبنای فاصله اقلیدسی تعیین می گردد.

(۲) توسط OLS، خط (یال) L_1 بر سه گره ابتدایی عارضه خطی مورد نظر برازش داده می شود و پارامترهای a_1 و b_1 محاسبه می گردند. این سه گره، اعضای مجموعه فرضی A_1 در نظر گرفته می شوند ($A_1 = \{n_1, n_2, n_3\}$).

(۳) اگر فاصله یی یال برازش داده شده تا تک تک گره های مجموعه A_1 کمتر از مقدار T باشد، گره بعدی به مجموعه A_1 اضافه گشته و همین عملیات مجدد تکرار می گردد. در غیر این صورت آخرین گره از مجموعه A_1 حذف گشته و پارامترهای a_1 و b_1 حاصل از برازش یال، توسط OLS، بر گره های مجموعه A_1 به عنوان پارامترهای اولین یال ذخیره می شوند.

(۴) پس از ذخیره سازی پارامترهای اولین یال، آخرین گره افزوده شده به مجموعه A_1 و دو گره بعدی عارضه خطی، اعضای مجموعه فرضی A_2 در نظر گرفته می شوند.

(۵) توسط OLS، یال L_2 بر گره های مجموعه A_2 برازش داده می شود.

(۶) اگر فاصله یال برازش داده شده تا تک تک گره های مجموعه A_2 کمتر از مقدار T باشد، گره بعدی به مجموعه A_2 اضافه گشته و همین عملیات مجدد تکرار می گردد. در غیر این صورت آخرین گره از مجموعه A_2 حذف گشته و پارامترهای a_2 و b_2 حاصل از برازش یال بر نقاط مجموعه A_2 به عنوان پارامترهای دومین خط ذخیره می شوند.

(۷) پس از به دست آمدن دومین یال، معادله آن با معادله یال اول برابر قرار داده می شود تا اولین گره (N_1) از عارضه خطی هدف محاسبه گردد.

(۸) عملیات های فوق ادامه خواهند یافت و هر بار یال جدیدی ذخیره می گردد که می توان از برابر قرار دادن معادله آن با معادله یال قبل از خودش، گره جدیدی از عارضه خطی هدف را به دست آورد.

(۹) این روند تا جایی ادامه خواهد یافت که جهت محاسبه پارامترهای یال L_k ، گره آخر (به دلیل بسته بودن عارضه خطی، همان گرهی است که مختصاتی برابر با گره اول از عارضه خطی اولیه دارد) و بلافاصله گره N_1 در مجموعه فرضی A_k قرار داده شود. در این حال: الف - اگر تمامی گره های مجموعه A_k دارای فاصله کمتر از T نسبت به یال برازش داده شده بر آنها باشند، آن یال به عنوان آخرین یال عارضه خطی ساده سازی شده ذخیره می شود و گره N_1 و یال L_1 حذف

^۱ Perpendicular Euclidean Distance (PED)

این روند ادامه می‌یابد و یال‌ها و گره‌های عارضه خطی هدف پشت سر هم محاسبه می‌گردند (اشکال ۳-ج، ۳-چ، ۳-د، ۳-ر، ۳-ز، ۳-س). تا جایی که محاسبات به آخرین گره از عارضه خطی اولیه (n_{13}) و بلافاصله پس از آن به اولین گره از عارضه خطی هدف برسد (اشکال ۳-ص، ۳-و). زمانی که محاسبات به اولین گره (N_1) می‌رسد، دو حالت پیش می‌آید: اگر فاصله تک‌تک گره‌های آخرین مجموعه فرضی A (در اینجا مطابق شکل ۳-و $A_4 = \{N_1, n_{13}, n_{12}, n_{11}, n_{10}\}$ می‌باشد) کمتر از مقدار T باشد، آنگاه یال برازش داده شده بر گره‌های مجموعه A_4 جایگزین اولین یال (L_1) و محل تقاطع آن با L_2 ، جایگزین اولین گره (N_1) خواهد شد (اشکال ۳-ه، ۳-ی). نقطه برخورد یال جایگزین شده و یال L_3 نیز سومین گره عارضه خطی هدف خواهد بود (شکل ۳-ی). در غیر این صورت اولین یال و گره در عارضه خطی هدف باقی خواهند ماند و N_1 از مجموعه A_4 حذف می‌گردد. حال یال برازش داده شده بر گره‌های مجموعه $A_4 = \{n_{13}, n_{12}, n_{11}, n_{10}\}$ به عنوان L_3 ، محل برخورد آن با L_2 به عنوان N_3 (سومین گره عارضه خطی هدف) و محل برخورد آن با L_1 به عنوان N_4 ذخیره خواهد شد. بدین صورت عملیات ساده سازی در این مثال به پایان خواهد رسید و عارضه خطی موجود در شکل ۳-ی برگردانده خواهد شد.

می‌گردند. سپس از برابر قرار دادن معادله یال‌های L_2 و L_k ، گره N_k به عنوان آخرین گره (همچنین چون عارضه خطی مورد بحث، بسته است به عنوان اولین گره) به دست می‌آید.

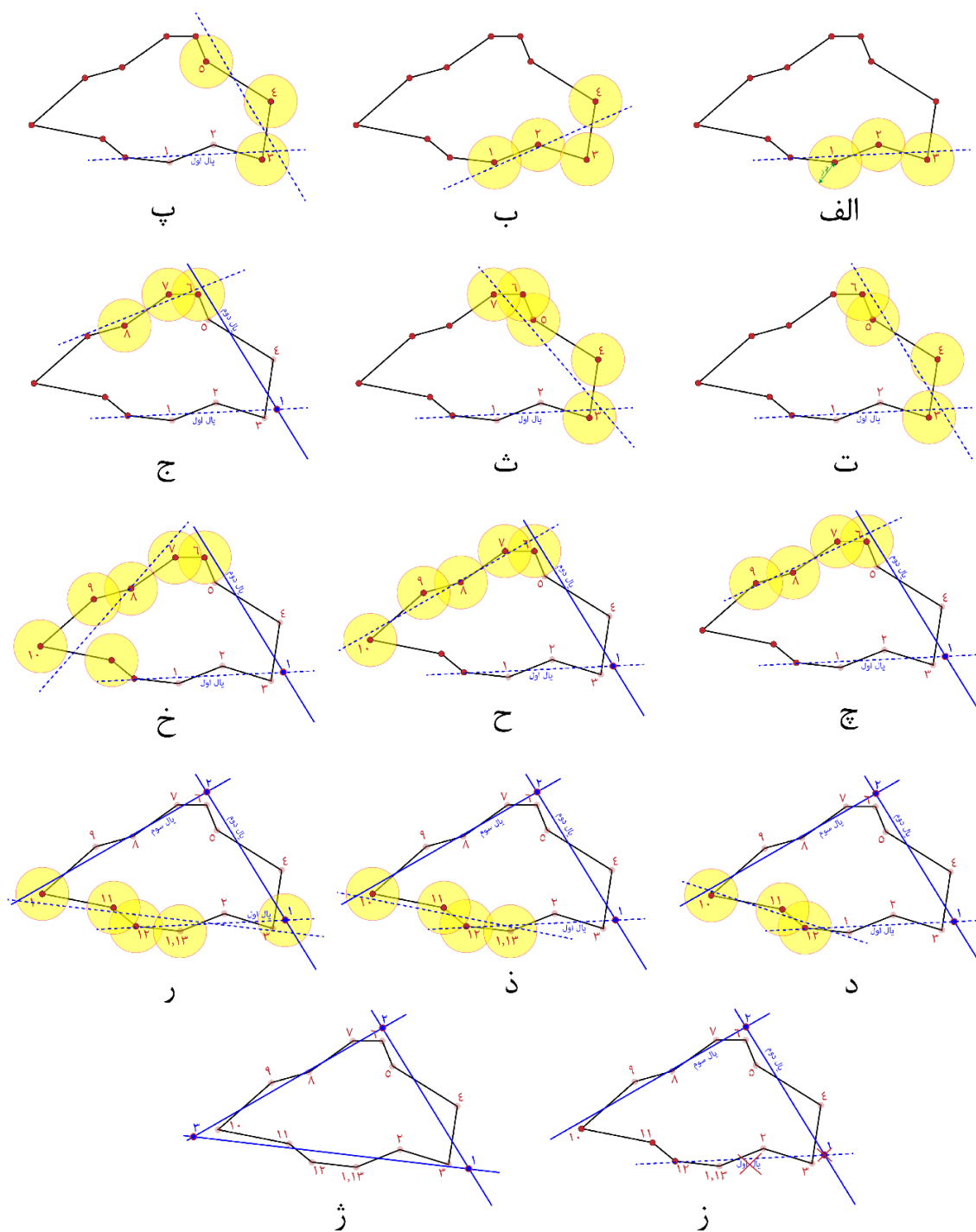
۱۰ ب - در غیر این صورت N_1 به A_k اضافه نمی‌گردد و آخرین یال از برازش بر گره‌های A_k به دست می‌آید. در نهایت با برابر قرار دادن معادله یال‌های L_1 و L_k ، گره N_k به عنوان آخرین گره به دست می‌آید.

برای بیان بهتر، ۹ مرحله بیان شده در شکل ۳ نشان داده شد. پس از در نظر گرفتن تلورانس (T) که همان شعاع دایره‌های زرد رنگ موجود در تصاویر است، یالی توسط OLS، بر سه گره ابتدایی (n_1, n_2, n_3) عارضه خطی که اعضای مجموعه فرضی A_1 در نظر گرفته می‌شوند، برازش داده می‌شود (شکل ۳-الف).

از آنجا که این یال از محدوده T واحدی هریک از سه گره ابتدایی عبور می‌کند (به عبارت دیگر فاصله آن تا هریک از سه گره n_1, n_2, n_3 کمتر از T است)، گره چهارم (n_4) به مجموعه فرضی A_1 افزوده شده و مجدداً برازش یال روی این گره‌ها انجام می‌گیرد (شکل ۳-ب).

اینبار یال برازش داده شده، در تلورانس T از تمامی نقاط قرار ندارد (خارج از تلورانس T واحدی گره سوم می‌باشد). بنابراین n_4 از مجموعه A_1 حذف گشته و یال به دست آمده از مرحله قبل (یال موجود در شکل ۳-الف) به عنوان اولین یال (L_1) ذخیره می‌گردد.

حال مجموعه فرضی A_2 با سه گره n_3, n_4, n_5 در نظر گرفته و یال L_2 بر این گره‌ها برازش داده می‌شود (شکل ۳-پ). با توجه به اینکه فاصله تک‌تک گره‌های مجموعه A_2 از یال L_2 کمتر از T است، گره n_6 به مجموعه A_2 اضافه شده و این فرایند مجدد تکرار می‌گردد (شکل ۳-ت). پس از افزودن گره n_7 و انجام برازش، یال به دست آمده خارج از تلورانس گره چهارم قرار می‌گیرد. بنابراین گره هفتم از مجموعه A_2 حذف گشته و یال به دست آمده از مرحله قبل (شکل ۳-ت) به عنوان دومین یال از عارضه خطی هدف ذخیره می‌گردد. از تقاطع یال اول و دوم، اولین گره از عارضه خطی هدف (N_1) به دست می‌آید. نکته حائز اهمیت این است که اولین گره و اولین یال به دست آمده، مجدداً در انتهای کار مورد بررسی قرار می‌گیرند.

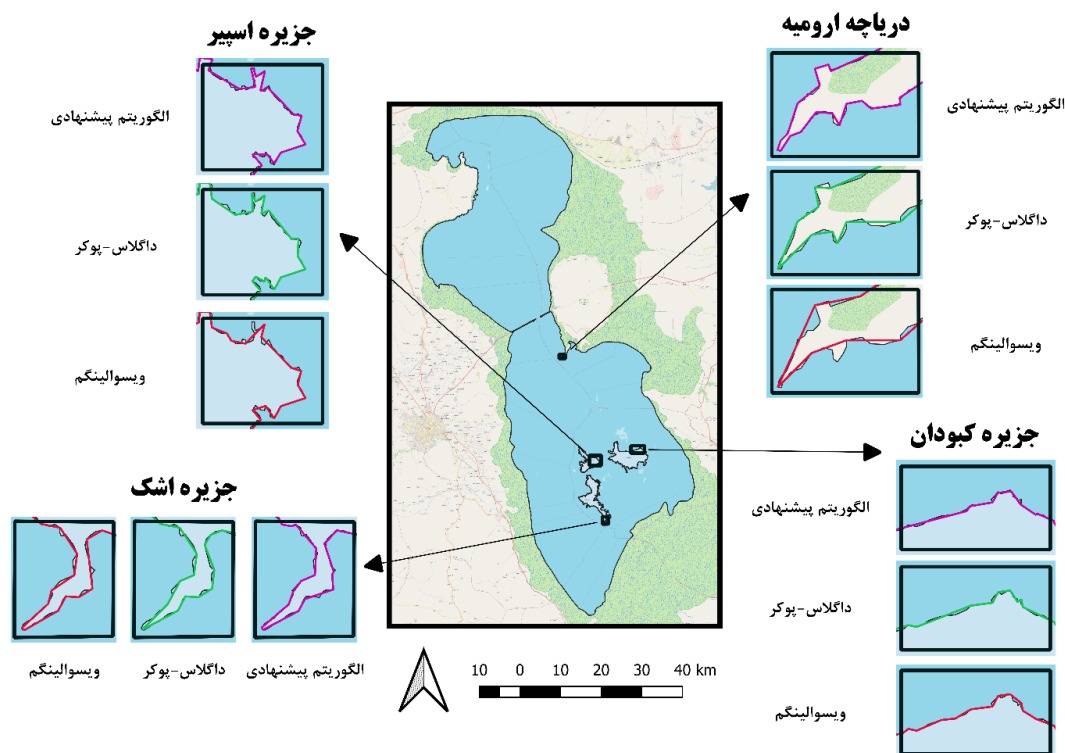


شکل ۳- مراحل کار مدل پیشنهادی تحقیق در ساده‌سازی عوارض خطی

۴- بحث و بررسی نتایج

مدل پیشنهادی مطابق روند بیان شده، برای دریاچه ارومیه، جزیره کبودان، جزیره اشک و جزیره اسپیر پیاده شد و برای بررسی بهتر، نتایج مدل پیشنهادی در قیاس با جوابهای روش داگلاس- پوکر و ویسوالینگم در شکل ۴ ارائه گردید. مقایسه‌های مذکور در حالی انجام گرفته‌اند که درصد کاهش حجم خروجی به نسبت هندسه عارضه اولیه

برای هر سه الگوریتم یکسان باشد. مقدار کاهش حجم به ترتیب برای دریاچه ارومیه و جزایر کبودان، اسپیر و اشک، به ترتیب حدوداً برابر با ۲۰، ۹، ۱۳ و ۲۲ درصد در نظر گرفته شد. همانطور که در شکل مشاهده می‌شود، اختلاف میان حدود اصلی عوارض و شکل ساده‌سازی شده آن‌ها، برای الگوریتم پیشنهادی کمتر از سایر روش‌ها است و توانسته است تا حد امکان شکل اصلی عارضه را حفظ کند.



شکل ۴- نتایج مدل پیشنهادی و روشهای داگلاس- پوکر و ویسوالینگم برای دریاچه ارومیه و جزایر آن

برای ارزیابی نتایج روشهای ساده سازی عوارض شاخص های متنوعی وجود دارد. در این پژوهش چهار شاخص فاصله هاسدرف میانه تصحیح شده، اختلاف مساحت ها، شباهت میزان تغییرات زاویه و شباهت انحنا متوسط برای ارزیابی کارایی مدل پیشنهادی و روشهای پر کاربرد ساده سازی بکار گرفته شد [۳۱]. این شاخص ها به صورت زیر محاسبه می گردند:

- فاصله هاسدرف میانه تصحیح شده: بر پایه کمینه فواصل هر نقطه از تمامی نقاط خط محاسبه می شود [۳۰].
- اختلاف مساحت ها: اختلاف مساحت عارضه خطی ساده شده با شکل اولیه آن (برای چند خطی ها)
- میزان تغییرات زاویه ای: مجموع زوایای عارضه خطی ساده شده بر مجموع زوایای اولیه آن.
- انحنا متوسط: مجموع تغییرات زاویه ای در واحد طول.

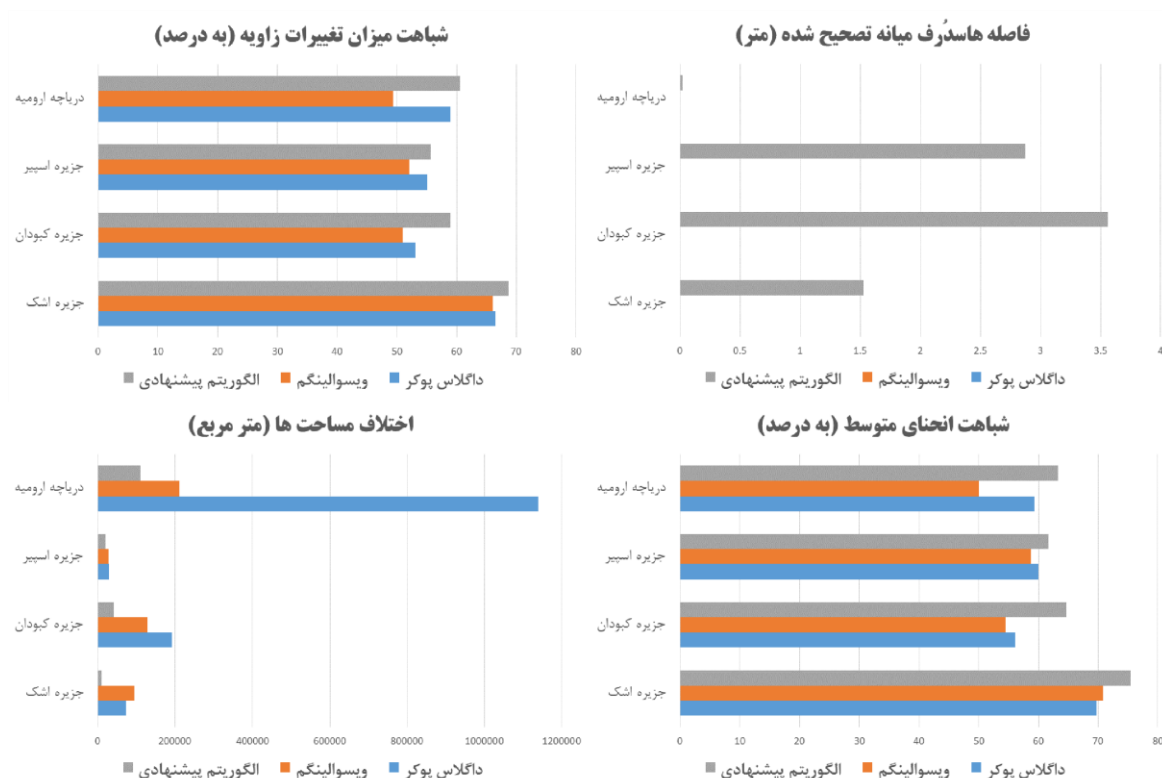
نتایج ارزیابی در شکل ۵ ارائه گردیده است.

با توجه به اینکه روشهای داگلاس پوکر و ویسوالینگم از نقاط خط اولیه برای برازش استفاده میکنند، در نتیجه کمینه فاصله برای آنها صفر خواهد بود در حالی که برای مدل پیشنهادی این تحقیق، به دلیل عدم استفاده از نقاط و برازش به کل آنها صفر نخواهد بود. لذا بر اساس این معیار، دو روش مذکور بهتر از مدل پیشنهادی عمل می -

کنند. لیکن در قیاس با معیارهای پراهمیتی نظیر اختلاف مساحت ها، شباهت میزان تغییرات زاویه و شباهت انحنا متوسط، مدل پیشنهادی کارایی بهتری داشته است [۳۲،۳۱]. درصد شباهت تغییرات زاویه و انحنا متوسط بطور میانگین ۲ الی ۸ درصد بهتر از سایر روشها عمل کرده است. در خصوص معیار مساحت نیز، مدل پیشنهادی نسبت به شکل اولیه عارضه، مساحت را بهتر حفظ نموده است. همچنین اختلاف زیادی بین مساحت عارضه اولیه و عارضه خلاصه شده در روشهای داگلاس پوکر و ویسوالینگم دیده می شود که به دلیل تاکید این روشها بر حفظ گره عارضه بوده و سبب اعوجاج زیادی در مساحت می گردد. درصد شباهت مساحت، تغییرات زاویه و انحنا عارضه ساده شده نسبت به عارضه اصلی در جدول ۱ آورده شده است. همچنین میزان فاصله نقاط از خطوط، برحسب متر نیز در شاخص فاصله هاسدرف میانه تصحیح شده ذکر شده است. نظر به اینکه تلورانس الگوریتم پیشنهادی از نوع فاصله است و با توجه به اینکه در نهایت هیچ قسمتی از خروجی، خارج از تلورانس تعیین شده از عارضه خطی ورودی نخواهد بود، بنابراین تعیین تلورانس کاملاً وابسته به هدف مورد نظر است. مقدار آن تابعی از ابعاد عارضه و سطح جزئیات مورد نظر می باشد و با توجه به اینکه تعیین آن نیازمند شناخت کاربر از منطقه مطالعاتی، هدف مورد

نظر و دقت داده‌های ورودی است، تعیین آن جهت حفظ

انعطاف مدل به کاربر واگذار میگردد.



شکل ۵- مقایسه شاخص‌های ارزیابی برای مدل پیشنهادی و روشهای داکلاس- پوکر و ویسوالینگم

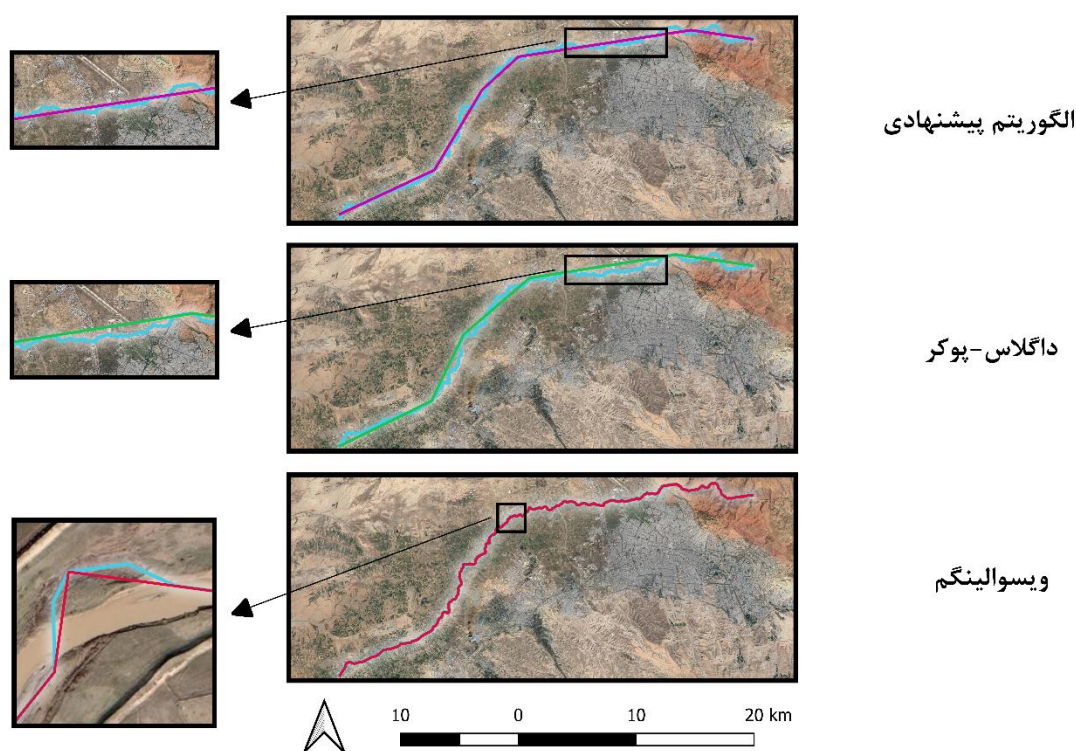
جدول ۱- مقایسه شبهات عوارض خطی ساده شده نسبت به شکل اصلی آنها توسط شاخص‌های ارزیابی

کمترین مربعات (بر پایه PED)		ویسوالینگم		داگلاس-پوکر	
جزیره اشک		۹۹,۹۶۲۵	۹۹,۶۱۷۴	۹۹,۶۹۷۶	
جزیره کبودان		۹۹,۸۷۰۵	۹۹,۶۰۱۰	۹۹,۴۰۷۹	
جزیره اسپیر		۹۹,۸۳۴۳	۹۹,۷۸۸۷	۹۹,۷۶۹۹	
دریاچه ارومیه		۹۹,۹۹۶۷	۹۹,۹۹۳۷	۹۹,۹۶۶۳	
میانگین		۹۹,۹۱۶۰	۹۹,۷۵۰۲	۹۹,۷۱۰۴	
جزیره اشک		۷۵,۴۸۰۲	۷۰,۸۱۹۲	۶۹,۷۶۷۵	
جزیره کبودان		۶۴,۷۰۱۵	۵۴,۵۲۲۹	۵۶,۱۲۸۵	
جزیره اسپیر		۶۱,۶۷۱۹	۵۸,۷۸۰۳	۵۹,۹۶۳۷	
دریاچه ارومیه		۶۳,۳۱۰۱	۵۰,۰۰۹۴	۵۹,۳۲۰۵	
میانگین		۶۶,۲۹۰۹	۵۸,۵۳۲۹	۶۱,۲۹۵۱	
جزیره اشک		۶۸,۶۸۹۲	۶۶,۰۵۰۰	۶۶,۴۷۸۸	
جزیره کبودان		۵۸,۹۵۸۸	۵۰,۹۱۱۳	۵۳,۱۱۷۰	
جزیره اسپیر		۵۵,۶۹۴۷	۵۲,۱۳۹۹	۵۵,۰۹۹۴	
دریاچه ارومیه		۶۰,۶۰۱۹	۴۹,۳۴۱۱	۵۸,۹۳۷۶	
میانگین		۶۰,۹۸۶۲	۵۴,۶۱۰۶	۵۸,۴۰۸۲	
جزیره اشک		۱,۵۲۸۰	.	.	
جزیره کبودان		۳,۵۶۴۱	.	.	
جزیره اسپیر		۲,۸۷۶۷	.	.	
دریاچه ارومیه		۰,۰۱۹۱	.	.	
میانگین		۱,۹۹۷۰	.	.	

الگوریتم پیشنهادی بسیار نزدیک تر به هندسه اولیه مسیر رود می باشد. الگوریتم ویسوالینگم نیز، که تلورانس از جنس مساحت دارد، از ۶۰۸ گره عارضه اولیه، تنها ۲۰۷ گره را حذف نموده است. برای اینکه معیار مساحت نیز قابل محاسبه باشد، از اتصال گره اول به گره آخر توسط یک یال فرضی استفاده شد. باتوجه به نتایج عددی نشان داده شده از ارزیابی ها، الگوریتم داگلاس-پوکر و الگوریتم پیشنهادی این پژوهش، با تلورانس ۱۰۰۰ متری، کاهش حجم یکسانی را نتیجه داده اند. این درحالیست که به جز شاخص فاصله تصحیح شده هاسدرف، بر اساس دیگر شاخص ها، الگوریتم پیشنهادی بر پایه کمترین مربعات معمولی شباهت بیشتری با هندسه اصلی عارضه دارد.

همچنین عملیات ساده سازی توسط الگوریتم های مذکور، روی رودخانه آجی چای نیز اعمال شد. این رودخانه با داشتن حالت کلی منحنی شکل و همچنین پیچ و تاب فراوان در طول مسیر خود، دارای هندسه مناسبی جهت ارزیابی کارایی الگوریتم ها می باشد. اینبار بجای مقایسه خروجی الگوریتم ها با ثابت نگه داشتن میزان کاهش حجم، مقدار تلورانس ثابت و برابر با ۱۰۰۰ واحد نگه داشته شد. که در اینجا برای الگوریتم پیشنهادی و داگلاس-پوکر، واحد تلورانس برابر با متر و برای ویسوالینگم برابر با مترمربع می باشد. با توجه به شکل ۶، الگوریتم پیشنهادی و داگلاس-پوکر، با حد آستانه ۱۰۰۰ متری، به میزان یکسانی دچار کاهش حجم شده اند (در هر دو ۶ گره باقی مانده است). این درحالی است که خروجی

رودخانه آجی چای



شکل ۶- نتایج مدل پیشنهادی و روشهای داگلاس-پوکر و ویسوالینگم برای رودخانه آجی چای

جدول ۱- مقایسه شباهت مسیر رودخانه آجی چای ساده شده نسبت به شکل اصلی آن توسط شاخص های ارزیابی

کمترین مربعات (بر پایه PED)	ویسوالینگم	داگلاس-پوکر	
۹۸,۰۲	۹۹,۹۹	۹۴,۴۷	شباهت مساحت (به درصد)
۶۹,۸۰	۵,۳۸	۳,۰۳	شباهت انحنای متوسط (به درصد)
۶۹,۲۳	۲,۷۴	۲,۵۷	شباهت تغییرات تندی خط (به درصد)
۷,۵۲	۰	۰	اختلاف از نظر فاصله تصحیح شده هاسدرف (به متر)
۹۹,۰۱	۳۴,۰۴	۹۹,۰۱	کاهش حجم (به درصد)

۵- نتیجه‌گیری و پیشنهادات

مدل پیشنهادی این تحقیق برای ساده‌سازی عوارض مبتنی بر کمترین مربعات معمولی با هدف کمینه‌سازی فاصله عمودی از خطوط می‌باشد. رگرسیون انجام شده بر پایه کمترین مربعات، برای برازش یک خط بوده است. لذا از معادله درجه یک استفاده شده است. با توجه به اینکه تمام عارضه‌های مورد مطالعه تک خط نبوده و به فرم عارضه خطی یا منحنی نیز می‌باشند، تقریب سایر خطوط نیز استخراج شده و با اتصال قطعات خط ساده شده، عارضه خطی نهایی ایجاد گردید. همچنین اکثر روشهای خلاصه‌سازی با اعمال حد‌آستانه‌هایی نظیر زاویه، فاصله اقلیدسی و فاصله اقلیدسی synchronous سعی بر حفظ شکل نهایی در محدوده‌ای از شکل اصلی دارند. مشابه پژوهش‌های پیشین، در این تحقیق نیز علاوه بر تقریب

خطوط برپایه کمترین مربعات، مدل پیشنهادی سعی بر حفظ خروجی خلاصه شده در حد آستانه‌ای از هندسه شکل اصلی دارد. نتایج پیاده‌سازی بر خط ساحلی دریاچه ارومیه و برخی جزایر آن، حاکی از بهینگی نتایج مدل پیشنهادی به لحاظ معیارهای حفظ مساحت، شباهت تغییرات زاویه و انحنای متوسط بوده است که شباهت بیشتری نسبت به عارضه اولیه داشته‌اند. لیکن در این تحقیق امکان مقایسه با روشهای بیشتری فراهم نشد. لذا پیشنهاد می‌گردد سایر روشهای رگرسیون در حیطه برازش چندخطی‌ها و ساده‌سازی عوارض مورد بررسی قرار گیرد. همچنین تقریب همزمان زاویه و طول نیز می‌تواند جهت ساده‌سازی عوارض مورد بررسی قرار گیرد. بهبودهایی نظیر جلوگیری از خودمقاطع شدن نیز میتواند در مطالعات آتی در دستور کار قرار گیرد.

مراجع

- [1] Filippovska, Y., Walter, V., Fritsch, D., "Quality evaluation of generalization algorithms", ISPRS Beijing, 2008.
- [2] He, X., Zhang, X., Yan, J., " Progressive Amalgamation of Building Clusters for Map Generalization Based on Scaling Subgroups", ISPRS Int. J. Geo-Inf. , 7, 116, 2018.
- [3] Ying, S., Li, L., "A FRAMEWORK OF MODEL-ORIENTED MAP GENERALIZATION AND ITS IMPLEMENTATION". International Cartographic Conference (ICC) Durban, South Africa, 2003.
- [4] Visvalingam, M., "The Visvalingam Algorithm: Metrics, Measures and Heuristics", The Cartographic Journal, pp. 1-11, 2016.
- [5] Ma, D., Zhao, Zh., Zheng, Y., Guo, R., Zhu, W., "PolySimp: A Tool for Polygon Simplification Based on the Underlying Scaling Hierarchy", ISPRS Int. J. Geo-Inf., 9(10), 594, 2020.
- [6] Shi, W., Cheung, Ch., " Performance Evaluation of Line Simplification Algorithms for Vector Generalization", The Cartographic Journal, The World of Mapping, 43(1), 2006.
- [7] White, E. R., "Assessment of Line-Generalization Algorithms Using Characteristic Points", The American Cartographer, 12(1), 1985.
- [8] Clayton, V., H., "A review of feature simplification and systematic point elimination algorithms". NOAA technical report NOS 112, 1985.
- [9] Reum, K., Witkam, A., P., M., "Optimizing Curve Segmentation in Computer Graphics". International Computing Symposium. Amsterdam, North Holland, 467- 472, 1974.
- [10] Zhao, Z., Saalfeld, A., "Linear-time sleeve-fitting polyline". in Autocarto 13, ACSM/ASPRS'97 Technical Papers, Seattle, Washington, 5, 214-223, 1997.
- [11] Opheim, H., "Fast data reduction of a digitized curve", Geo-Processing, 2, 33-40, 1982.
- [12] Lang, T., "Rules for robot draughtsmen", Geographical Magazine, 42, 50-51, 1969.
- [13] Hershberger, J., Snoeyink, J., "Speeding up the Douglas-Peucker line-simplification algorithm", in Proceedings of the Fifth International Symposium on Spatial Data Handling, 1, 134-143, 1992.
- [14] Visvalingam, M., Whyatt, J. D., "Line generalization by repeated elimination of points", The Cartographic Journal, 30, 46-51, 1993.
- [15] Guo, Q., Wang, H., He, J., Zhou, Ch., Liu, Y., Xing, B., Jia, Zh., Li, M., "Graphic Simplification and Intelligent Adjustment Methods of Road Networks for Navigation with Reduced Precision". ISPRS Int. J. Geo-Inf., 9(8), 490, 2020.

- [16] Shen, Y., Ai, T., He, Y., "A New Approach to Line Simplification Based on Image Processing: Case Study of Water Area Boundaries", *ISPRS Int. J. Geo-Inf.*, 7(2), 41, 2018.
- [17] Ai, T., Ke, Sh., Yang, M., Li, J., "Envelope generation and simplification of polylines using Delaunay triangulation". *International Journal of Geographical Information Science*, 31(2), 2017.
- [18] Gökğöz, T., Sen, A., Memduhoglu, A., Hacı, M., "A New Algorithm for Cartographic Simplification of Streams and Lakes Using Deviation Angles and Error Bands". *ISPRS Int. J. Geo-Inf.*, 4(4), 2185-2204, 2015.
- [19] Douglas, D., Peucker, T., "Algorithms for the Reduction of the Number of Points Required to Represent a Line or its Caricature," *Cartographica*, vol. 10, no. 2, pp. 112-122, 1973.
- [20] Aghel Shahneshtin, S., Sadat Mirvahami, S., Abbaspour, R.A., "An algorithm for compression of a spatio-temporal trajectory preserving its semantic nature", *Journal of Geospatial Information Technology* Vol.3, No.4, Winter 2016.
- [21] Visvalingam, M. and Whyatt, J. D., "Cartographic algorithms: problems of implementation and evaluation and the impact of digitising errors", *Computer Graphics Forum*, 10(3), pp. 225-235, 1990.
- [22] Visvalingam, M. and Whyatt, J. D., "Line generalisation by repeated elimination of points", *Cartographic Journal*, 30(1), pp. 46-51, 1993.
- [23] Otto, B., "Linear Algebra With Applications", 3rd ed., Upper Saddle River NJ: Prentice Hall, 1995.
- [24] Ghilani, Ch. D., Wolf, P. R., "Adjustment Computations: Spatial Data Analysis", Wiley, Fourth Edition, 2006.
- [25] Farahanipooya, A., Abbaspour, R. A., Alavipanah, S. K., Jafarbeiglou, M., "Roads matching in a multi-scale spatial database using a least square line", *JGST*. 3 (1) :87-104, 2013.
- [26] Tamari, S., WoEsten, J.H.M., and Ruiz-SuaÁrez, J.C., "Testing an artificial neural network for predicting soil hydraulic conductivity". *Soil Sci. Soc. Amer. J.* 60: 1732-1741, 1996.
- [27] Barikloo, A., Servati, M., Oliaei, M.S., "Performance of Ordinary least Squares (OLS) and Bayesian network (BN) in Exchange sodium percentage prediction based on Sodium adsorption Ratio", *J. of Water and Soil Conservation*, Vol. 26(2), 2019.
- [28] Shi, W., Cheung, Ch., "Performance Evaluation of Line Simplification Algorithms for Vector Generalization", *the cartographic journal*, vol. 43, no. 1, pp. 27-44, 2006.
- [29] Lin, X., Jiang, J., Ma, Sh., Zuo, Y., Hu, Ch., "Onepass trajectory simplification using the synchronous Euclidean distance". *The VLDB Journal* vol. 28, no. 6, pp. 897-921, 2019.
- [30] Tonga, X., Dan Lianga, D., Jina, Y., "A linear road object matching method for conflation based on optimization and logistic regression". *International Journal of Geographical Information Science*, Vol. 28, No. 4, 824-846, 2014.
- [31] Saberinasab, M., Karimi, M., "Comparison and performance evaluation of generalization algorithms of linear features in topographic maps". *The 2nd National Conference on Geospatial Information Technology (NCGIT)*, 18 January 2017.
- [32] Chamani M., Ali Abbaspour, R., Chehreghan, A.R., "Comparison of Polygon Matching Methods in Multi-represented Vector Datasets". *GEJ*. 9 (2) :73-84, 2018.