

ارزیابی خطر رخداد سیل در مناطق سیل‌خیز با بکارگیری مدل‌های احتمالاتی و یادگیری ماشین

مطالعه موردی: حوزه آبخیز قره‌سو و گرگانرود استان گلستان

محمدعلی صدقی^۱

تاریخ دریافت: ۱۴۰۲/۰۸/۱۱ تاریخ پذیرش: ۱۴۰۳/۰۳/۱۱

چکیده

سیل یکی از مخرب‌ترین بلایای طبیعی با پیامدهای اجتماعی، اقتصادی و زیست محیطی است که امروزه روش‌های یادگیری ماشین و احتمالاتی به منظور مدل‌سازی و پیش‌بینی آن توسعه داده شده‌اند. اهداف این پژوهش اولویت‌بندی عوامل مؤثر، پهنه‌بندی و پیش‌بینی حساسیت نسبت به رخداد سیل با استفاده از مدل‌های رگرسیون بردار پشتیبان (SVR) و وزن شاهد بیزین (WOE) و معرفی مناسب‌ترین آن‌ها در حوزه‌ی قره‌سو و گرگانرود استان گلستان می‌باشد. نقشه پراکنش رخداد سیل بر اساس ۳۶۸ مکان با تعداد تکرار ۷۸۲ رخداد سیل بر اساس اطلاعات موجود، بازدید میدانی و بررسی تصاویر گوگل ارث در بازه زمانی قبل و بعد از سیلاب تهیه شد. از این تعداد ۷۰ درصد (۲۵۸ نقطه) به عنوان داده‌های آموزشی و ۳۰ درصد (۱۱۰ نقطه) به عنوان داده‌های آزمایشی به ترتیب برای اجرا و اعتبارسنجی مدل به طور تصادفی انتخاب و تقسیم شدند. در مرحله بعد از ۱۵ عامل زمین‌شناسی، هیدرولوژیک، مورفومتریک حوزه به همراه داده‌های اقلیم (به عنوان متغیر مستقل) و به منظور پهنه‌بندی و مدل‌سازی استفاده شد. ارتباط بین عوامل مؤثر و رخدادهای سیل با استفاده از نسبت فراوانی، کمی و وزن‌دار شدند. به منظور بررسی همپوشانی اطلاعاتی عوامل مؤثر، با استفاده از تحلیل رگرسیون چند متغیره خطی، استقلال داده‌ها مورد آزمون قرار گرفت. نقشه‌های پهنه‌بندی حساسیت حاصل از اجرای دو مدل با استفاده از سطح زیر منحنی ویژگی عملکرد گیرنده

^۱ کارشناس ارشد مهندسی نقشه برداری گرایش سیستم اطلاعات جغرافیایی (GIS)

(AUC-ROC) مورد ارزیابی، اعتبارسنجی و مقایسه قرار گرفتند. نتایج بررسی عوامل موثر در هر دو مدل متفقاً نشان داد، عوامل کاربری اراضی، ارتفاع و پوشش گیاهی به ترتیب تاثیر به‌سزایی در رخداد سیلاب داشته‌اند و بخش زیادی (بیش از ۸۰٪) از سیل‌های رخ داده در کلاس حساسیت خیلی زیاد و زیاد قرار می‌گیرد. نتایج ارزیابی طبقه‌بندی و اعتبارسنجی مدل‌ها نشان داد AUC-ROC برای میزان موفقیت مدل‌های رگرسیون بردار پشتیبان و وزن شاهد به ترتیب ۰/۹۲ و ۰/۸۸ است. بنابراین، نتایج مدل رگرسیون بردار پشتیبان نسبت به مدل وزن شاهد دقت بیشتری دارد. نقشه‌های پهنه‌بندی حساسیت به سیل حاصل از این پژوهش می‌تواند مبنای برنامه‌ریزی و مدیریت بحران ناشی از رخداد سیل قرار گیرد.

واژه‌گان کلیدی: پهنه‌بندی، سیل، رگرسیون بردار پشتیبان، قره‌سو و گرگان‌رود، وزن شاهد بیزین.

مقدمه

سیلاب‌ها از جمله مخاطراتی هستند که همیشه احتمال رخداد آنها به عنوان یک مخاطره طبیعی برای انسان‌ها وجود داشته است. بنابراین در راهکارهای کاهش اثرات بلایای طبیعی هدف اصلی به سوی آگاه‌سازی عمومی و تاکید بر امر اطلاع رسانی به مردم در راستای اقدامات پیشگیرانه قبل از وقوع حوادث ناهنجار تغییرکرده که این روش تغییر، خود از اهداف اساسی برنامه کاهش اثرات بلایای طبیعی در دهه گذشته بوده است. متأسفانه موضوع سیل و مدیریت و کاهش خسارات آن در کشور مورد توجه جدی قرار نگرفته و فقط زمانی که سیلاب مخربی جاری می‌شود و فاجعه‌ای بوجود می‌آید، توجه مسئولین و متخصصین به آن جلب می‌گردد. اگرچه بررسی طرح‌های مهار سیلاب که به صورت محدود و پراکنده در سطح کشور مطالعه و اجرا شده‌اند نشان می‌دهد یک راه حل مشخص و مطمئن برای کلیه مناطق سیل‌گیر وجود ندارد اما بدیهی است پدیده سیل علیرغم همه پیچیدگی‌هایش قابل بررسی و مطالعه بوده و می‌توان در جهت مهار و کاهش خسارات آن و حتی بهره‌برداری اقتصادی از سیل راه‌حل‌های مناسبی جستجو کرد. بر این اساس برنامه‌ریزی و انجام اقدامات

جامع جهت پیشگیری و کاهش خسارات سیل در قالب طرح‌های مطالعاتی و اجرایی از اهمیت بسزایی در راستای دستیابی به اهداف توسعه پایدار برخوردار می‌باشد. در استان گلستان و منطقه گرگان‌رود، طغیان رودخانه‌ها و بروز سیل، هر ساله خسارات و تلفات قابل ملاحظه‌ای را برجای می‌گذارد. بخش عمده‌ای از این خسارات، تلفات و معضلات در سطح جوامع روستایی ساکن در حاشیه رودخانه گرگان‌رود و در سطح زیرحوضه‌های مختلف آن به وقوع پیوسته است. آمارهای رسمی نشان می‌دهد که از سه دهه پیش تاکنون این استان شاهد وقوع سیلاب‌های متعددی بوده که خسارت‌های سنگینی بر بخش زیرساخت‌ها، ابنیه، تأسیسات زیربنایی و واحدهای مسکونی برجای گذاشته‌اند.

برمبنای گزارش اداره کل منابع طبیعی و آبخیزداری استان گلستان، تعداد وقایع زمانی سیل در استان از تاریخ ۱۳۷۰/۱/۱ الی ۱۳۹۷/۱/۱ به میزان ۱۲۷ بارش بوده و از نظر مکانی در ۷۰۴ مورد از حوزه‌های آبخیز استان رخ داده است. به عبارت دیگر در طی این مدت ۲۷ ساله در نتیجه ۱۲۷ مورد ریزش باران در ۷۰۴ مورد مکانی از حوزه‌های آبخیز استان سیل اتفاق افتاده است. که مجموع خسارات مالی استان ناشی از سیل ۸۳۵۷/۳۴ میلیارد ریال در طی دوره ۲۷ ساله بوده است (اداره آبخیزداری و حفاظت خاک استان گلستان، ۱۳۹۷). لذا پهنه‌بندی مناطق سیل خیز امکان اقدامات پیش‌گیرانه و مهار سیلاب را فراهم کرده و ضمن کاهش ریسک خطر سیل، هزینه‌های ناشی از احداث سازه‌های مربوطه در مناطق غیر لازم را کمتر می‌کند. نقشه‌های پهنه‌بندی حساسیت به سیل را می‌توان به عنوان ارزیابی کمی یا کیفی طبقه‌بندی، مساحت و توزیع فضایی سیل، که وجود دارد یا ممکن است در یک منطقه رخ دهد، تعریف کرد. کاربرد روش‌های یادگیری ماشین در مطالعات پهنه‌بندی خطر سیل توسط بسیاری از محققان انجام شده است (مریت، ۲۰۱۰). به طور مثال SAHA و همکاران از روش SVR به منظور مدل‌سازی استفاده کرده‌اند و نتایج خوبی به همراه داشته است (ساحا،

۲۰۲۱). در مطالعه‌ای دیگر که هدف شناسایی مناطق مستعد سیل با استفاده از ماشین بردار پشتیبان (SVM) در حوضه نکارود واقع در استان گلستان بود، نتایج ارزیابی نشان داد که مدل SVM از دقت خوبی در شناسایی مناطق مستعد سیل در منطقه مورد مطالعه برخوردار است. علاوه بر این، نتایج این مطالعه نشان داد که مناطق حساس به سیل در مناطق شمالی و شمال غربی حوضه و در بخش‌هایی که غلظت سکونتگاه‌های انسانی بیشتر است، بیشتر است، در حالی که مناطق مرکزی حوضه با پوشش گیاهی متراکم از حساسیت کمی برخوردار هستند (چراغی و همکاران، ۱۳۹۹).

در سوی دیگر مدل‌های احتمالاتی با دخالت دادن تمامی عوامل مؤثر بر رخداد یک پدیده وزنی مشخص برای هر کدام در نظر می‌گیرد و نشان می‌دهد کدام عامل‌کارایی بیشتری داشته و می‌تواند نقش هر عامل در وقوع سیل را به تفکیک نشان دهد. رحمتی و همکاران (۲۰۱۶) با بکارگیری مدل‌های نسبت فراوانی (FR) و وزن‌های شواهد (WofE) اقدام به تهیه نقشه حساسیت سیل در استان گلستان کرد. نتایج نهایی نشان داد که مدل‌های $AUC = FR$ و 76.47% و $WofE (AUC = 74.74\%)$ نتایج تقریباً مشابه و معقولی دارند. بنابراین، این نقشه‌های حساسیت به سیل توانستند برای محققین و برنامه ریزان در استراتژی‌های کاهش سیل مفید باشد. فرامرزی و همکاران (۱۳۹۸) به منظور شناسایی متغیرهای مهم در ایجاد مناطق سیل گیر و ارائه‌ی پتانسیل مخاطره‌ی سیل پارک ملی گلستان، از تکنیک‌های یادگیری ماشین شامل آنتروپی بیشینه، مدل جنگل تصادفی و درخت رگرسیون تقویت‌شده استفاده کردند. نتایج به‌دست‌آمده از ارزیابی صحت مدل‌ها با استفاده از ۳۰ درصد از داده‌های وقوع سیل که در مدلسازی وارد نشده بود نیز دقت زیاد مدل درخت رگرسیون تقویت‌شده و جنگل تصادفی را با مقدار ROC ۰/۹۹ و دقت مناسب آنتروپی بیشینه را با مقدار ROC ۰/۸۹، نشان داد، به‌طوری که نقشه‌های به‌دست‌آمده از این مدل‌ها به طور مشترک ۴۵۰۰

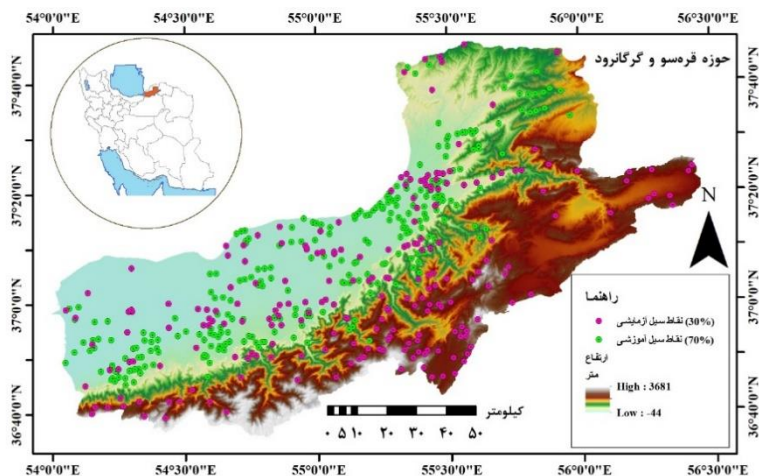
هکتر از مساحت پارک را دارای احتمال زیاد خطر سیل برآورد کردند. با توجه به اهمیت منطقه استان گلستان در مهار سیل و بکارگیری روش‌های جدید در تهیه نقشه‌های پهنه‌بندی خطر سیل، هدف این پژوهش ارزیابی حساسیت به رخداد سیل با استفاده از مدل‌های احتمالاتی و یادگیری ماشین با تأکید بر دو روش وزن شاهد (WOE) و رگرسیون بردار پشتیبان (SVR) در حوزه آبخیز قره‌سو و گرگانرود می‌باشد.

مواد و روش‌ها

منطقه مورد مطالعه

منطقه مورد مطالعه (حوضه آبخیز گرگانرود و قره‌سو) در سه استان گلستان، خراسان شمالی و سمنان واقع شده است. به طوری که ۴۳ درصد سطح استان گلستان را شامل می‌شود. به عبارتی ۱۳ درصد آب سطحی استان یعنی ۳۲۳ میلیون متر مکعب در این حوضه جریان دارد (شرکت آب منطقه‌ای استان گلستان). مساحت این حوضه تقریباً معادل ۱۳۰۵۰ کیلومتر مربع است که در بخش جنوب شرقی دریای خزر قرار دارد و در تقسیم‌بندی کلی هیدرولوژی ایران، جزئی از آبریز دریای خزر به شمار می‌رود. این حوضه در محدوده طول جغرافیایی ۵۳ درجه تا ۵۶ درجه و ۱۵ دقیقه شرقی و عرض جغرافیایی ۳۶ درجه و ۳۵ دقیقه تا ۳۷ درجه و ۴۳ دقیقه شمالی واقع شده است. این منطقه بین حدود ارتفاعی حداکثر ۳۶۸۱ و حداقل ۱۴- متر پایین‌تر از سطح دریا قرار دارد. رود اترک، گرگانرود و رود قره‌سو، سه رود مهم استان گلستان هستند و گرگانرود و قره‌سو رودخانه‌های اصلی حوضه آبریز گرگانرود- قره‌سو به شمار می‌روند (بردی شیخ، ۲۰۰۹).

شکل ۱ نشان‌دهنده محدوده مورد مطالعه می‌باشد:



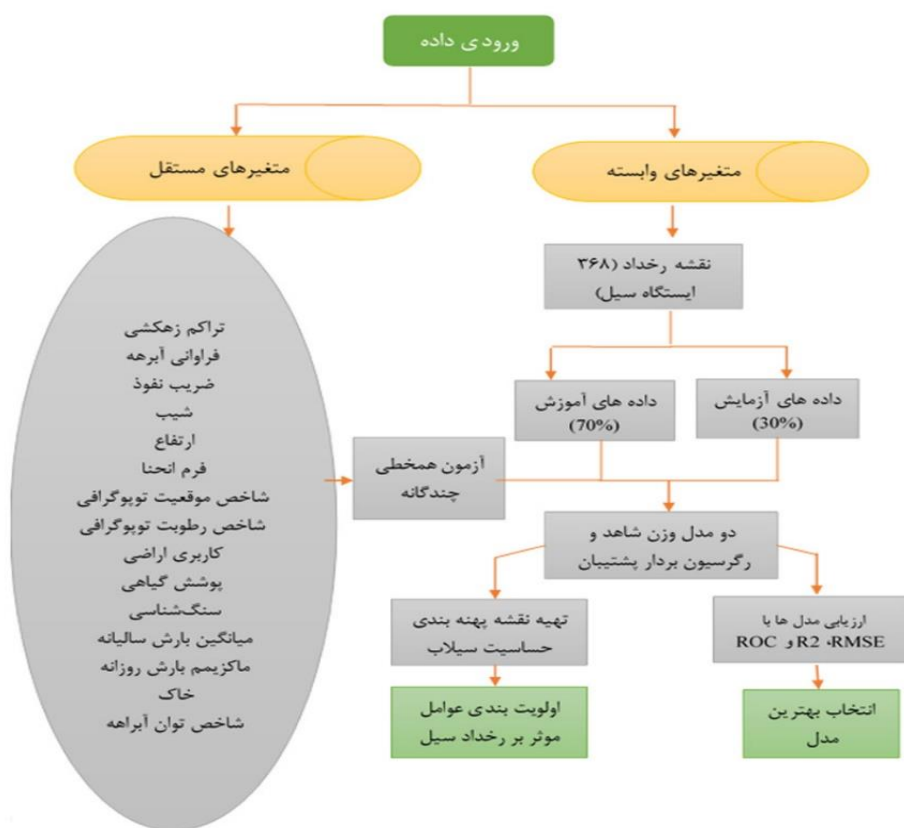
شکل ۱. موقعیت جغرافیایی منطقه

جمع‌آوری اطلاعات و تهیه عوامل مؤثر بر سیل

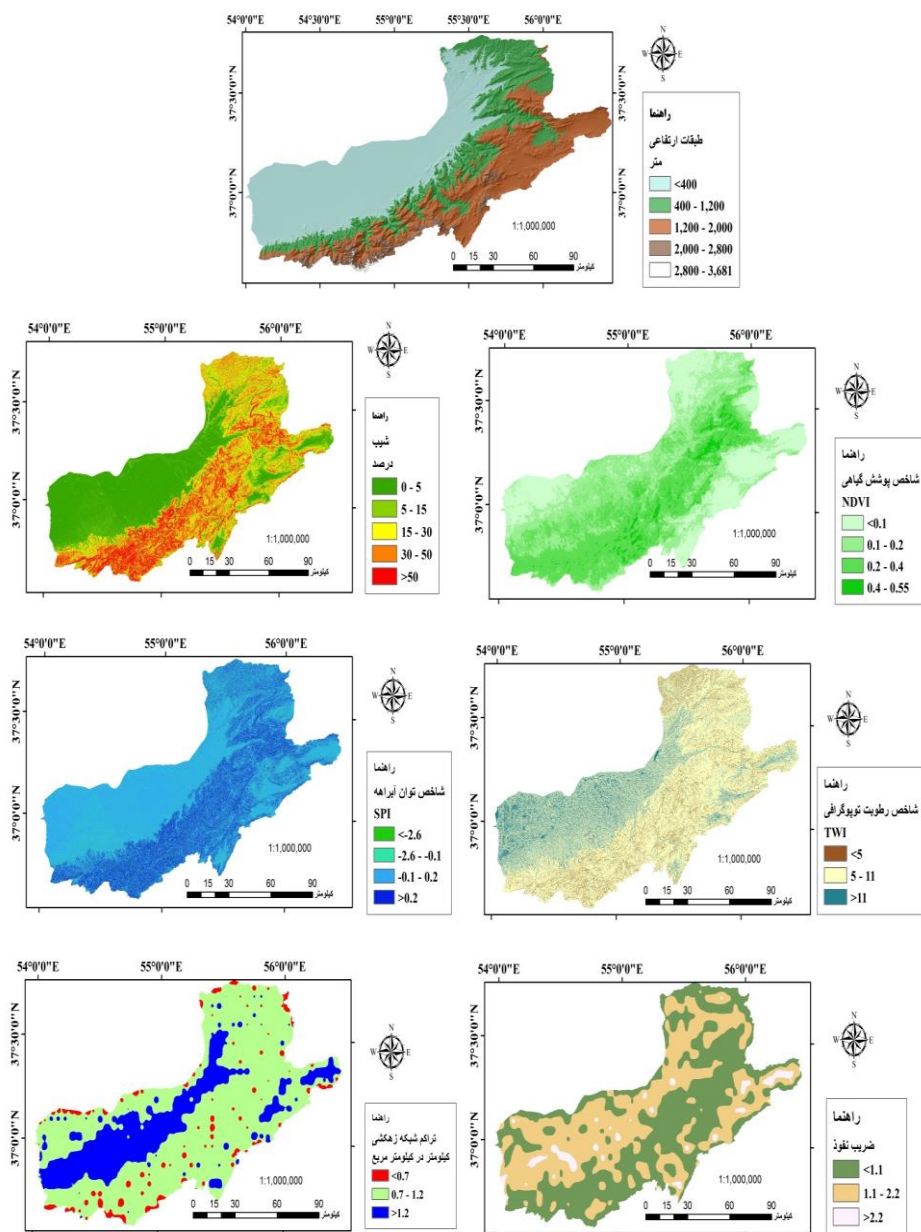
براساس سابقه تحقیق (رحمتی، ۲۰۲۰)، منابع موجود و تجارب متخصصان ۱۵ عامل تأثیرگذار شامل تراکم شبکه زهکشی، فراوانی آبراهه، ضریب نفوذ، شیب، ارتفاع، فرم انحنا، شاخص موقعیت توپوگرافی، شاخص رطوبت توپوگرافی، کاربری اراضی، پوشش گیاهی، سنگ شناسی، میانگین بارش سالیانه، ماکزیمم بارش روزانه، خاک و شاخص توان آبراهه برای تهیه نقشه حساسیت به رخداد سیل در نظر گرفته شد.

نقشه‌های پایه مورد استفاده در این پژوهش شامل نقشه‌های زمین‌شناسی با مقیاس ۱:۱۰۰۰۰۰، عکس‌های هوایی در مقیاس ۱:۴۰۰۰۰، نقشه توپوگرافی با مقیاس ۱:۵۰۰۰۰، نقشه کاربری زمین با مقیاس ۱:۱۰۰۰۰۰، تصاویر ماهواره‌ای لندست ۸ (به تاریخ یازدهم و نوزدهم ماه آگست سال ۲۰۲۱) و نقشه خطوط هم‌باران (بدست آمده از آمار

ایستگاه‌های سازمان هواشناسی کل کشور) می‌باشند. شایان ذکر است مقیاس تمامی نقشه‌ها در محیط GIS براساس کوچک مقیاس‌ترین نقشه عامل مورد استفاده یکسان‌سازی و نرمال گردید و دقت نقشه‌های خروجی براساس مقیاس ۱/۱۰۰۰۰۰۰۰ مورد تحلیل قرار گرفتند. معیارها بر اساس مقیاس، موقعیت منطقه مورد نظر و روش پهنه‌بندی مورد استفاده به‌منظور پهنه‌بندی حساسیت نسبت به سیلاب مورد ارزیابی قرار گرفتند. مختصات ۵۹ نقطه سیلابی حوضه با دستگاه GPS توسط اطلاعات دریافتی از اداره کل منابع طبیعی و آبخیزداری استان گلستان، ثبت و وارد ArcGIS®10.5 گردید و نقشه پراکنش سیلاب از این طریق تهیه شد.



شکل ۲. نمودار روش تحقیق



شکل ۳. نقشه عوامل مؤثر بر سیل

داده‌های کافی به‌منظور برآورد اهمیت نسبی موضوعات مستدل از طریق میانگین‌های آماری مد نظر باشد، استفاده می‌گردد (لی و چوی، ۲۰۰۴). احتمال قبلی یا اولیه در واقع احتمال یک واقعه یا رخداد (رخداد سیلاب) است که توسط انواع وقایع مشابه که در گذشته در یک دوره زمانی معین رخ می‌دهند مشخص و تعیین می‌گردد. برای مثال احتمال آتی سیل یک سطح واحد (یا پیکسل) از زمین براساس فراوانی یا تناوب وقوع سیل سطح واحد (یا پیکسل) در گذشته می‌تواند تخمین زده شود. به این منظور ابتدا با استفاده از رابطه‌های (۱) و (۲) وزن‌های محتمل هر عامل و طبقه‌های مربوط به آن محاسبه می‌گردد.

$$W_i^+ = \log_e [P(B_i|S)/P(B_i|\bar{S})] \quad (۱)$$

$$W_i^- = \log_e [P(\bar{B}_i|S)/P(\bar{B}_i|\bar{S})] \quad (۲)$$

که در آن: W_i^+ وزن شاهد عامل در وقوع سیل در صورت حضور رده عامل مؤثر (شرطی)، W_i^- وزن شاهد عامل در وقوع سیل در صورت عدم حضور رده عامل مؤثر (شرطی)، $\log_e$ لگاریتم طبیعی، P احتمال، B_i وجود (سطح) رده عامل مؤثر (شرطی) در وقوع سیل بالقوه، S وجود (سطح) سیل، $\bar{B}_i$ عدم وجود (سطح) رده عامل مؤثر (شرطی) در وقوع سیل بالقوه و $\bar{S}$ عدم وجود (سطح) سیل می‌باشند. به‌منظور درک بهتر روابط فوق و سهولت کار، به‌ویژه در محیط GIS مدل براساس جدول (۱) و به‌صورت رابطه‌های (۳) و (۴) ارائه می‌گردد (ون و بتر، ۲۰۰۴). و نهایتاً وزن‌نهایی براساس روابط (۵) و (۶) برای هر عامل به‌منظور پهنه‌بندی خطر سیلاب مورد استفاده قرار می‌گیرد (رمضانی و ابراهیمی، ۱۳۸۸)

$$W_i^+ = \log_e \left[\left(N_{\text{pix}1} / (N_{\text{pix}1} + N_{\text{pix}2}) \right) / \left(N_{\text{pix}3} / (N_{\text{pix}3} + N_{\text{pix}4}) \right) \right] \quad (۳)$$

$$W_i^- = \log_e \left[\left(N_{pix2} / (N_{pix1} + N_{pix2}) \right) / \left(N_{pix4} / (N_{pix3} + N_{pix4}) \right) \right] \quad (4)$$

جدول ۱. چهار حالت پتانسیل ایجاد سیل

	Causes of Floods	
Flood	Presence	absence
Presence	N_{pix1}	N_{pix2}
absence	N_{pix3}	N_{pix4}

که در آن: N_{pix1} برابر با تعداد پیکسل‌های سیلابی در یک کلاس، N_{pix2} برابر با تعداد کل پیکسل‌های سیلابی در یک نقشه منهای تعداد پیکسل‌های سیلابی در یک کلاس، N_{pix3} برابر با تعداد پیکسل‌های یک کلاس منهای تعداد پیکسل‌های سیلابی در یک کلاس و N_{pix4} برابر با تعداد کل پیکسل‌های یک نقشه منهای تعداد کل پیکسل‌های سیلابی یک نقشه منهای تعداد پیکسل‌های یک کلاس می‌باشد.

$$C = [(W^+) + (W^-)] \quad (5)$$

$$W_{final} = C / S_C \quad (6)$$

که در آن: C برآیند وزن‌های مثبت و منفی، W_{final} وزن نهایی استاندارد شده و S_C انحراف معیار برآیند وزن‌های شاهد مثبت و منفی است که برابر با جذر واریانس هریک از وزن‌های مثبت و منفی می‌باشد.

مدل رگرسیون بردار پشتیبان

روش رگرسیون بردار پشتیبان (SVR) مستقیماً از نظریه یادگیری آماری وپنیک به نام ماشین‌های بردار پشتیبان استخراج شده است. ماشین‌های بردار پشتیبان برای مسائل طبقه‌بندی به کار می‌روند. بعدها الگوریتم آن‌ها برای کار با مسائل رگرسیون یا تخمین داده‌ها توسعه

یافت. این الگوریتم جدید رگرسیون بردار پشتیبان نامیده شد (اسمولا و شالکوف^۱، ۱۹۹۸). رگرسیون بردار پشتیبان روشی است که با اتکا بر مبنای ماشین بردار پشتیبان می‌تواند به عنوان رگرسیون در مسائل مهندسی عمل کند. این روش که جزء روش‌های آموزش نظارت شده است، بین داده‌های ورودی و مقدار پارامتر وابسته، بر اساس کمینه سازی ریسک ساختاری ارتباط برقرار می‌کند. SVR روشی برای تخمین یک تابع است که بر اساس داده‌های آموزشی از یک شیء ورودی به یک عدد واقعی نگاشت می‌شود.

در مسائل رگرسیون، بردارهای ورودی به یک فضای چندبعدی، نگاشت می‌شوند، سپس یک ابر صفحه ساخته می‌شود که بردارهای ورودی را با بیش‌ترین فاصله ممکن، از یکدیگر جدا می‌کند. برای رفع مشکل انجام عملیات در فضای با ابعاد زیاد، از یک تابع کرنل استفاده می‌شود. در این صورت عملیات می‌تواند با همان سرعت فضای داده‌های ورودی انجام شود. درواقع با استفاده از تابع کرنل، مشکل چندبعدی و غیرخطی بودن نگاشت، مرتفع می‌شود. فرآیند بهینه‌سازی باید با یک تابع تلفات اصلاح‌شده همراه شود، تا سنجش فاصله را در برگیرد. در حقیقت هدف SVR، تخمین پارامترهای وزن و بایاس تابعی است که بهترین انطباق را بر داده‌ها داشته باشد (بالوین^۲ و همکاران، ۲۰۱۱).

بحث

نتایج حاصل از همپوشانی نقشه پراکنش تعداد وقوع نقاط سیل و هریک از ۱۵ نقشه طبقه‌بندی شده عوامل موثر بر رخداد سیلاب، نشان داد که با افزایش تراکم زهکشی، میزان رخداد سیل نیز افزایش می‌یابد که از این نظر با مطالعات آرابامری (۲۰۲۰) مطابقت دارد. شدت بارندگی که در این پژوهش بصورت بیشترین بارش روزانه بوده است، با افزایش مقدار

^۱ Smola and Schölkopf

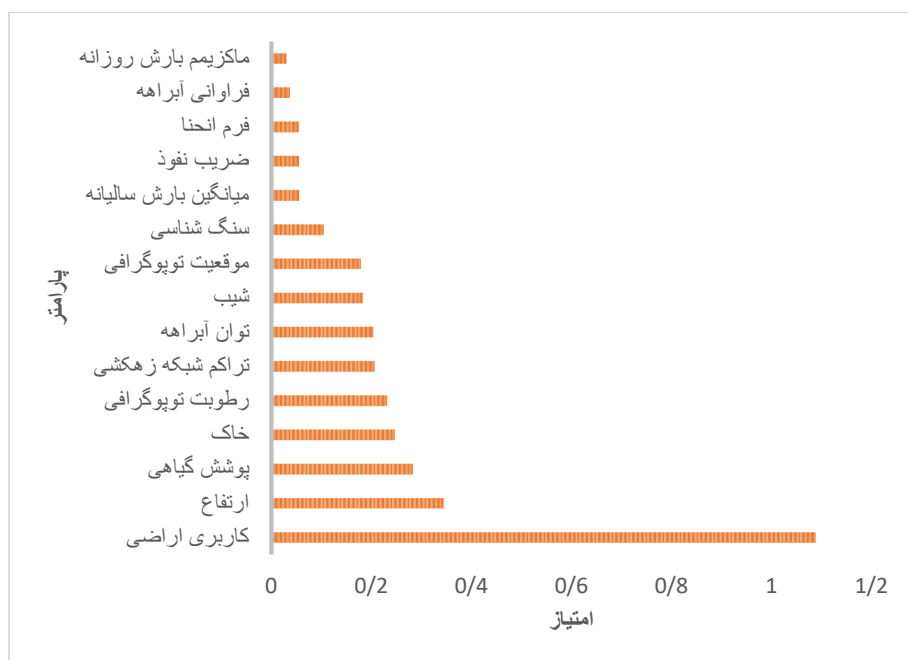
^۲ Balabin and Lomakina

آن میزان رخداد سیل نیز افزایش می‌یابد و سیلاب‌هایی با حجم بیشتری باعث می‌گردد. میانگین بارش سالیانه با توجه به وزن رده‌های آن، در ابتدا روند افزایشی و سپس کاهشی داشته است و این بخاطر نوع پوشش گیاهی منطقه و کاربری زمین در مناطق مورد نظر می‌باشد. ولی به‌طور کلی بیشترین میزان رخداد سیل با میزان بارش سالیانه خیلی کم مواجه بوده است که حاکی از فقدان پوشش گیاهی و یا بارش‌های رگباری در آن منطقه می‌باشد و با نتایج مطالعه رحمتی (۲۰۱۶) مطابقت دارد. شاخص رطوبت توپوگرافی یک خصوصیت توپوگرافی ثانویه است که توزیع مکانی از شرایط رطوبت را نشان می‌دهد که زیاد بودن این عامل باعث حجم زیاد سیلاب می‌شود که نتایج با مطالعات رحمتی و همکاران (۲۰۱۶) و ایسلام (۲۰۲۱) مطابقت دارد. طبق نتایج این پژوهش و مطالعه اسلامی‌نژاد (۲۰۲۰)، بین شاخص پوشش گیاهی و وقوع سیل‌گیری در منطقه رابطه منفی وجود دارد. به گونه‌ای که با افزایش تراکم پوشش گیاهی، حساسیت مناطق نسبت به وقوع سیل‌گیری کاهش پیدا می‌کند و برعکس. با توجه به معیارهای شاخص موقعیت توپوگرافی و کاربری اراضی بیشترین میزان سیل مربوط قسمت‌های سکونت‌گاهی و زراعی می‌باشد که این به دلیل شیب ملایم زمین و در صدد آن هدایت رواناب ایجاد شده در ارتفاعات به این مناطق می‌باشد و همچنین نداشتن پوشش گیاهی مناسب نیز به چشم می‌خورد. عامل سنگ‌شناسی در واقع وضعیت زمین‌شناسی منطقه با تأکید بر ویژگی‌های ساختار زمین‌شناسی را نشان می‌دهد. هرچه خاک سطحی‌تر و نفوذپذیری سنگ ضعیف‌تر باشد، رواناب سطحی و به تبع آن خطر وقوع سیل نیز افزایش می‌یابد. با افزایش فراوانی آبراهه پتانسیل سیل‌خیزی نیز افزایش پیدا می‌کند. طبق نتایج مطالعه رحمتی (۲۰۱۶)، بیشترین میزان سیل مربوط به طبقات دارای ارتفاع کم و هموار بودن زمین می‌باشد که دلیل آن جاری شدن سیل در مناطق با ارتفاع کمتر می‌باشد. اگرچه با افزایش شاخص توان آبراهه سرعت سطحی جریان نیز افزایش می‌یابد ولی در این پژوهش میزان

وقوع سیل روند کاهشی دارد و این به دلیل وجود پوشش گیاهی قوی در مناطق پایین دست حوضه می‌باشد. رده‌های عامل خاک تأثیرات مختلفی در ایجاد سیلاب دارد به طوری که تخته‌های نمکی به دلیل نفوذپذیری بالا میزان وقوع سیل را به شدت کاهش می‌دهد و برعکس گروه مولی‌سول‌ها بیشترین میزان سیل را به خود اختصاص داده‌اند که این با مطالعه رحمتی (۲۰۱۶) مطابقت دارد. در نواحی پرشیب میزان وقوع سیل کاهش یافته است چون در نواحی با شیب کم، رواناب زمان بیشتری برای نفوذ در اختیار دارد و موجب بروز سیل می‌شود (رحمتی، ۲۰۱۶). همچنین با توجه به نتایج حاصل از اجرای آزمون چندخطی در این پژوهش نشان داد که تمامی عوامل به دلیل عدم همپوشانی اطلاعاتی (میزان تحمل بیش‌تر از ۰/۱ و شاخص تورم کمتر از ۱۰) در نظر گرفته شدند. بنابراین اجرای دو مدل وزن شاهد و رگرسیون بردار پشتیبان با ۱۵ عامل صورت پذیرفت.

با اعمال روابط (۱) تا (۶) مدل وزن شاهد با ۱۵ عامل اجرا شد و در مجموع براساس اوزان محاسبه شده، به ترتیب عوامل کاربری اراضی، ارتفاع، شاخص پوشش گیاهی، خاک، شاخص رطوبت توپوگرافی، تراکم شبکه زهکشی، شاخص توان آبراهه، شیب، شاخص موقعیت توپوگرافی، سنگ‌شناسی، میانگین بارش سالیانه، ضریب نفوذ، فرم انحنا، فراوانی آبراهه و حداکثر بارش روزانه در وقوع سیل مؤثر می‌باشند (شکل ۵). نقشه نهایی حساسیت به سیل Y_{woe} با استفاده از مدل وزن شاهد از مجموع حاصل ضرب وزن استاندارد شده هر رده پارامتر C/Sc در وزن خود رده پارامتر C طبق رابطه زیر براساس نتایج موجود در جدول (۱) محاسبه شده است. i تا n همان پارامترها یا عوامل مؤثر در سیل می‌باشند.

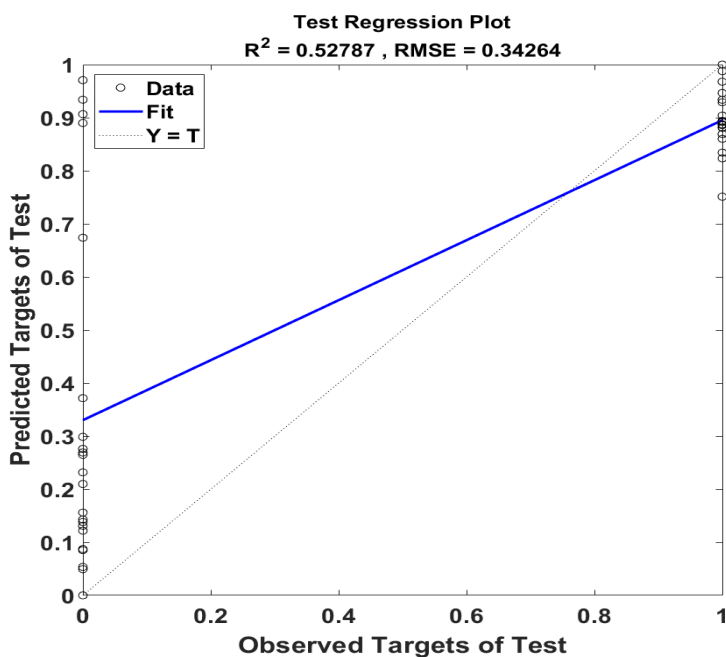
$$Y_{woe} = \sum_{i=1}^n C/Sc \times C \quad (7)$$



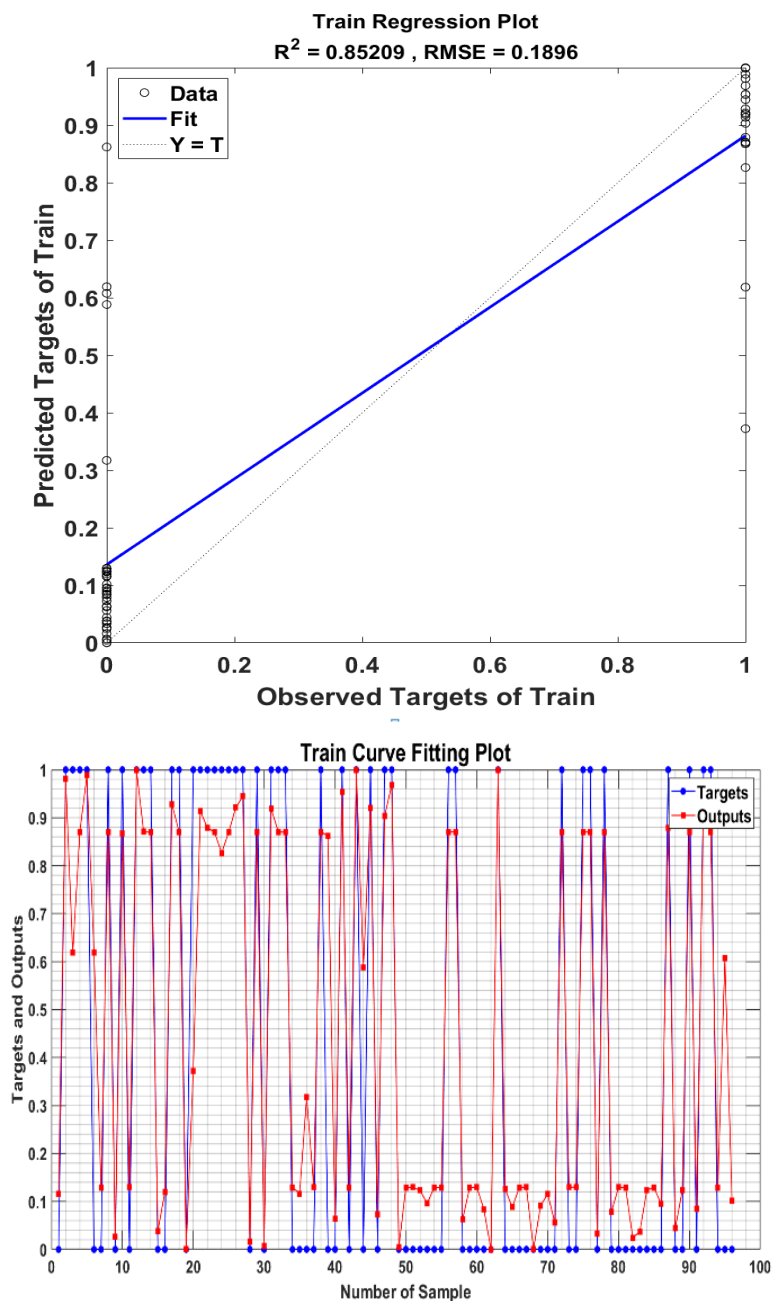
شکل ۵. اهمیت نسبی معیارهای تاثیرگذار بر رخداد سیل در مدل وزن شاهد

مدل رگرسیون بردار پشتیبان نیز در دو فاز آموزش و آزمایش به ترتیب برای ۷۰ و ۳۰ درصد نقاط سیلابی انجام شد (شکل ۶ و ۷). در شکل ۶ نمودار پراکنش مربوط به پیش‌بینی متغیر هدف در مدل رگرسیون بردار پشتیبان SVR در فاز آموزش و آزمایش نشان داده شده است. همانطور که از دو این نمودار مشخص است، خط تیره مشکی، نیمساز ربع اول در محور عمودی و افقی است. مقادیر هدف اندازه‌گیری شده بر روی محور X و مقادیر هدف پیش‌بینی شده بر روی محور Y جانمایی شده‌اند. هر چقدر این مقادیر به نیمساز ربع اول (خط تیره مشکی) نزدیک‌تر باشند، بیانگر برازش عالی و برآورد دقیق خروجی‌ها (پیش‌بینی دقیق‌تر متغیر هدف) خواهد بود. به خوبی از این شکل مشخص است که در مرحله‌ی آموزش مقادیر متغیر هدف نزدیک خط نیمساز ربع اول می‌باشند و خطای کمتری دارند. شکل ۷

نشان‌دهنده میزان تطبیق و سازگاری مربوط به مقادیر مشاهداتی و پیش‌بینی شده متغیر هدف در مرحله آموزش و آزمایش توسط مدل SVR است. به خوبی از این دو نمودار مشخص است که این مدل در مرحله آموزش در تخمین مقادیر متغیر هدف با دقت بالا عمل کرده است، به طوری که این مقادیر بسیار نزدیک به مقدار واقعی آنها تخمین زده شده است.

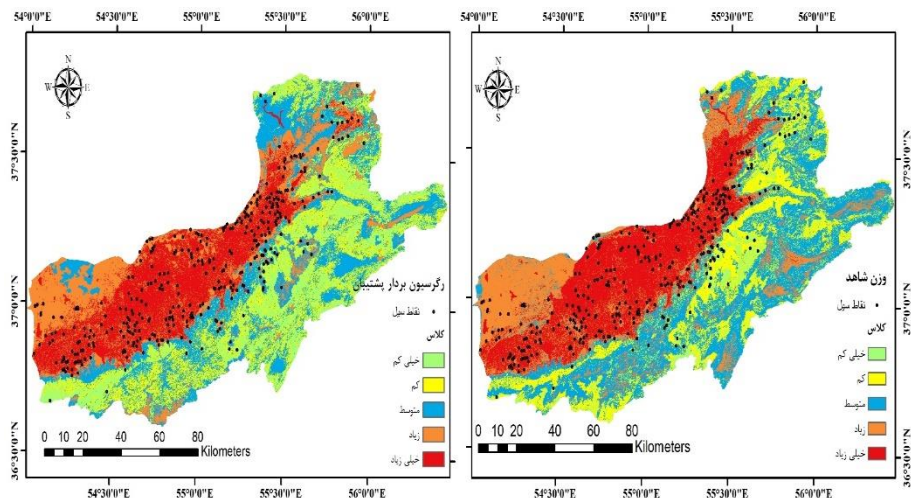


شکل ۶: نمودار پراکنش میان مقادیر اندازه‌گیری شده و پیش‌بینی شده متغیر هدف توسط مدل رگرسیون بردار پشتیبان SVR (فرآیند آموزش و آزمایش)



شکل ۷: مقادیر متغیر هدف اندازه گیری شده و پیش بینی شده توسط مدل مدل رگرسیون بردار پشتیبان SVR

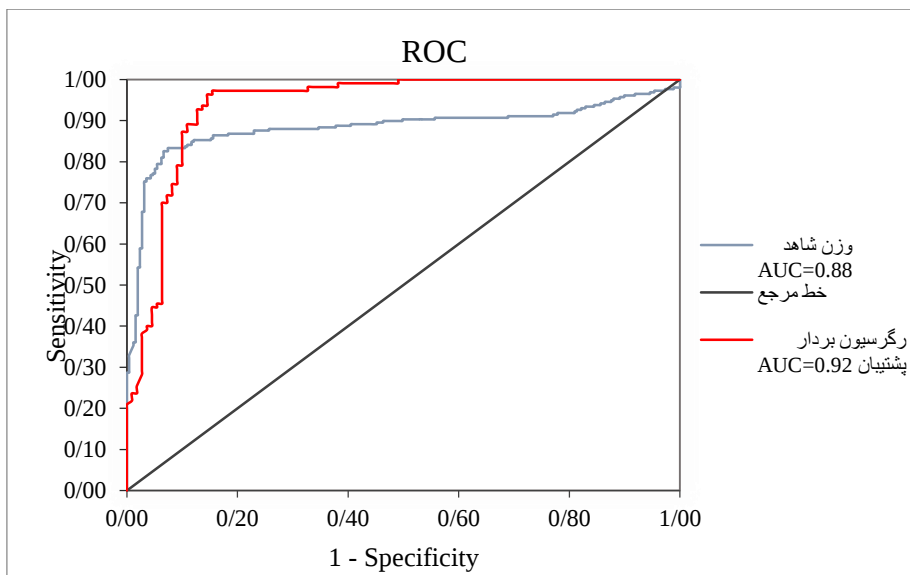
(فرآیند آموزش و آزمایش)



شکل ۸: نقشه‌های پهنه‌بندی حساسیت به رخداد سیل حاصل از اجرای دو مدل

از هم‌پوشانی نقشه پراکنش سیلاب با نقشه طبقه‌بندی پهنه‌بندی حساسیت نسبت به سیل مشخص گردید میزان سطح رخداد سیل با افزایش رده حساسیت از رده خیلی کم تا خیلی زیاد روند افزایشی داشته و در رده حساسیت خیلی زیاد بطور قابل توجهی افزایش یافته است (شکل ۸).

بر مبنای سطوح طبقه‌بندی دقت براساس منحنی ROC، هر دو مدل در طبقه خیلی خوب (۰/۸-۰/۹) قرار دارند و قابل توصیه جهت تهیه نقشه پهنه‌بندی حساسیت نسبت به سیل در امور زیر بنایی و توسعه‌ای می‌باشند. ولی در مقام مقایسه نسبت به همدیگر، نتایج حاصل از مقادیر مساحت زیر منحنی (AUC) مدل رگرسیون بردار پشتیبان نشان‌دهنده دقت امتیاز بالاتر (۰/۹۲) نسبت به مدل وزن شاهد (۰/۸۸) می‌باشد که بیانگر همبستگی بالا بین نقشه حساسیت نسبت به سیل تهیه شده و نقشه پراکنش تعداد وقوع سیل می‌باشد (شکل ۹).



شکل ۹: منحنی ROC

نتیجه‌گیری و پیشنهاد

سیل یکی از پدیده‌های مهم مصیبت‌بار و فاجعه‌انگیز در نواحی شمالی ایران محسوب می‌شود. از طریق آنالیز داده‌ها به کمک سیستم اطلاعات جغرافیایی می‌توان عوامل مؤثر بر سیلاب و مناطق حساس به سیل را شناسایی کرد و بنابراین از این طریق خسارت ناشی از آن را کاهش داد. در این پژوهش به منظور تهیه نقشه پتانسیل خطرپذیری سیل، از ۱۵ عامل مؤثر بر رخداد سیل استفاده شد که تمامی آنها دارای استقلال اطلاعاتی بودند. از ۳۶۸ نقطه سیل ۷۰٪ آن برای مرحله آموزش و ۳۰٪ آن برای مرحله آزمایش در نظر گرفته شد. در مرحله اجرای مدل‌ها هر دو مدل، در هر دو مرحله‌ی آموزش و آزمایش با خطاهای استاندارد خیلی کم برازش خوبی در مدل‌سازی کسب کردند. نتایج حاصل از هم‌پوشانی نقشه پراکنش سیلاب با نقشه‌های پهنه‌بندی حساسیت نسبت به رخداد سیل نشان داد رده‌های حساسیت خیلی کم تا کم درصد کمی از نقاط سیل را به خود اختصاص می‌دهند و برعکس رده‌های حساسیت زیاد تا خیلی زیاد بیشترین نقاط سیل را شامل شده است. همچنین در هر دو مدل،

میزان سطح رخداد سیل با افزایش رده حساسیت از رده خیلی کم تا خیلی زیاد روند افزایشی داشته و در رده حساسیت خیلی زیاد بطور قابل توجهی افزایش یافته است. با استفاده از سطح زیر منحنی ROC (AUC) مقادیر ۰/۹۲ و ۰/۸۸ به ترتیب برای مدل‌های رگرسیون بردار پشتیبان و وزن شاهد بدست آمد که مدل رگرسیون بردار پشتیبان با کسب درجه‌ی عالی (۱-۰/۹) مطلوبیت مدل‌سازی بالایی دارد. بر اساس برازش مدل‌ها و اولویت‌بندی عوامل با استفاده از آنها مشخص شد به ترتیب عوامل کاربری اراضی، ارتفاع و پوشش گیاهی دارای اهمیت و اولویت در رخداد سیل در منطقه مورد مطالعه می‌باشد. هر دو مدل در پهنه‌بندی و پیش‌بینی رخداد سیل در حوضه آبخیز قره‌سو و گرگانرود واقع در استان گلستان از کارآمدی مناسبی برخوردار می‌باشند. در مدل رگرسیون رده حساسیت خیلی زیاد با وسعت حدود ۲۲ درصد منطقه، حدود ۶۵ درصد رخداد سیل و در مدل وزن شاهد با وسعت حدود ۲۵ درصد منطقه، حدود ۶۳ درصد رخداد سیل را پوشش می‌دهد. در پایان با توجه به اینکه اجرای همزمان روش‌های یادگیری ماشین و احتمالاتی در این پژوهش دقت خوبی داشته است و همچنین ترکیب با روش‌های دیگر در مطالعات اخیر نتایج خوبی به همراه داشته است، بنابراین به پژوهشگران علاقه‌مند توصیه می‌شود روش‌های ترکیبی نوین در دیگر زیرحوضه‌های استان گلستان مورد بررسی و ارزیابی قرار گیرد.

منابع

- اسداللهی شهر، مهدی، اداره کل منابع طبیعی و آبخیزداری استان گلستان، معاونت فنی و آبخیزداری. بهار ۱۳۹۷.
- فرامرزی، ح.، و حسینی، س.، و پورقاسمی، ح.، و فرنقی، م. (۱۳۹۸). ارزیابی و پهنه بندی وقوع مخاطره سیلاب در پارک ملی گلستان. اکوهیدرولوژی، ۶(۴)، ۱۰۵۵-۱۰۶۸.
- چراغی قلعه سری، ع.، ر. حبیب نژاد و م. روشن. (۱۳۹۹). تهیه نقشه حساسیت سیلاب با استفاده از مدل ماشین بردار پشتیبان (SVM) و سیستم اطلاعات جغرافیایی (GIS). مخاطرات محیط طبیعی. ۹(۲۵): ۷۱-۸۰.

منابع انگلیسی

- Maier, H.R.; Jain, A.; Dandy, G.C.; Sudheer, K.P. Methods used for the development of neural networks for the prediction of water resource variables in river systems: Current status and future directions. *Environ. Model. Softw.* 2010, 25, 891–909.
- Saha, A., Pal, S. C., Arabameri, A., Blaschke, T., Panahi, S., Chowdhuri, I. ... & Arora, A. (2021). Flood susceptibility assessment using novel ensemble of hyperpipes and support vector regression algorithms. *Water*, 13(2), 241.
- Rahmati, O., A. Haghizadeh and S.Stefanidis. (2016a). Assessing the accuracy of GIS-based analytical hierarchy process for watershed prioritization; Gorganrood River Basin, Iran; *Water Resour. Manag.* 30(3) 1131–1150.
- Rahmati, O., Pourghasemi, H. R., & Zeinivand, H. (2016b). Flood susceptibility mapping using frequency ratio and weights-of-evidence models in the Golastan Province, Iran. *Geocarto International*, 31(1), 42-70.
- Bardi sheikh, V., Babaee, A. and Mosheikhian, Y. (2009). Assessment of the changes in precipitation regimes Gorganroud Basin. *Iran-Watershed Management Science and Engineering* 3(8), 29- 38.
- Arabameri, A., S.Saha, K. Mukherjee, T. Blaschke, W. Chen, W., Ngo, P. T. T., & S.S. Band. (2020). Modeling spatial flood using novel ensemble artificial intelligence approaches in northern Iran. *Remote Sensing*, 12(20), 3423.
- Haghizadeh, A., S. Siahkamari, A. H. Haghiabi, & O. Rahmati. 2017. Forecasting flood-prone areas using Shannon's entropy model. *Journal of Earth System*

Science, 126(3), 39.

Regmi, N. R., Giardino, J. R. and Vitek, J. D. (2010). modeling susceptibility to landslides using the weight of evidence approach: Western Colorado, USA, *Geomorphology*, 115, 172-187.

Van Western, C. J. (2002). Use of weights of evidence modeling for landslide susceptibility mapping 1-21.

Lee, S. and Choi, J. (2004). Landslide susceptibility mapping using GIS and the weight-of-evidence model. *Intl. J. Geograph. Inform. Sci.*, 18(8), 789-814.

Yesilnacar, E. K. (2005). The Application of Computational Intelligence to Landslide Susceptibility Mapping in Turkey, Ph.D Thesis. Department of Geomatics the University of Melbourne, 1-423.

Roering, J. J., Kirchner, J. W. and Dietrich, W. E. (2005). Characterizing Structural and Lithologi Controls on Deepseated Landsliding: Implications for Topographic Relief and Landscape Evolution in the Oregon Coast Range, *Geological Society of America Bulletin*, 17, 654-668.

Smola, Alex J., and Bernhard Schölkopf. 1998. "On a Kernel-Based Method for Pattern Recognition, Regression, Approximation, and Operator Inversion." *Algorithmica* 22(1-2):211-31.

Balabin, Roman M., and Ekaterina I. Lomakina. 2011. "Support Vector Machine Regression (SVR/LS-SVM)—an Alternative to Neural Networks (ANN) for Analytical Chemistry? Comparison of Nonlinear Methods on near Infrared (NIR) Spectroscopy Data." *Analyst* 136(8):1703-12.

Eslaminezhad, S. A., Eftekhari, M., & Akbari, M. (2020). GIS-Based Flood Risk Zoning Based On Data-Driven Models. *Journal of Hydraulic Structures*, 6(4), 75-98.

Islam, A. R. M. T., Talukdar, S., Mahato, S., Kundu, S., Eibek, K. U., Pham, Q. B., & Linh, N. T. T. (2021). Flood susceptibility modelling using advanced ensemble machine learning models. *Geoscience Frontiers*, 12(3), 101075.