

پهنه‌بندی پتانسیل سیل با استفاده از یادگیری ماشین و تصاویر Sentinel-1 SAR، مطالعه موردی: استان گلستان

الهه ابراهیمی^{۱*}، مهدی آخوندزاده^۲

^۱ کارشناس مهندسی نقشه‌برداری، دانشکده مهندسی نقشه‌برداری و اطلاعات مکانی، پردیس دانشکده‌های فنی،

دانشگاه تهران

ebrahimi.elah@ut.ac.ir

^۲ دانشیار دانشکده مهندسی نقشه‌برداری و اطلاعات مکانی، پردیس دانشکده‌های فنی، دانشگاه تهران

makhonz@ut.ac.ir

(دریافت: مرداد ۱۴۰۲، تصویب: مهر ۱۴۰۴)

چکیده

سیل به عنوان یک پدیده فاجعه بار، تأثیر زیادی بر اکوسیستم‌های طبیعی و زندگی انسان‌ها دارد. با توجه به افزایش احتمال فراوانی و شدت سیل در آینده به دلیل تغییرات اقلیمی، پیش‌بینی و برنامه‌ریزی زود هنگام در مناطق سیل‌خیز برای کاهش خسارات ضروری است. در ایران، در چند سال اخیر، استان گلستان سیل‌های متعددی را تجربه کرده‌است. در این مطالعه از روش‌های سنجش از دور پیشرفته و یادگیری ماشینی برای شناسایی و پیش‌بینی مناطق دارای پتانسیل سیل در استان گلستان استفاده شد. بطور خاص، با استفاده از تصاویر رادار Sentinel-1 (SAR) مناطق وقوع سیل بین سال‌های ۲۰۱۵ تا ۲۰۲۲ شناسایی شدند. متغیرهای مدل رقومی ارتفاع، شاخص نرمال شده تفاوت پوشش گیاهی، شیب توپوگرافی، شاخص رطوبت گیاهی، تجمع جریان، فاصله جریان، بارش ماهیانه، جهت جریان که نتایج آزمون VIF هیچ نشانه‌ای از هم خطی بودن بین آن‌ها را نشان نداد، بعنوان متغیرهای محیطی تأثیرگذار بر سیل انتخاب شدند. سپس با هفت الگوریتم یادگیری ماشینی شامل GLM، GAM، BRT، RF، MARS، FDA و CART مناطق مستعد سیل پیش‌بینی شدند. سپس با ترکیب مدل‌های منفرد، عدم اطمینان در نتایج کاهش یافت. عملکرد مدل‌ها با استفاده از چهار شاخص AUC، COR، TSS و Deviance ارزیابی شد. نتایج حاصل از این مطالعه نشان داد متغیرهای مدل رقومی ارتفاعی، تجمع جریان و بارندگی مهم‌ترین متغیرها در پیش‌بینی پتانسیل سیل در این ناحیه هستند. تمام مدل‌های منفرد عملکرد بالایی نشان دادند و در میان مدل‌ها RF بهترین عملکرد را داشت. بر اساس مدل ترکیبی، شهرستان گمیشان، آق قلا، بندر ترکمن، بخش‌های غربی گنبد کاووس، بخش شمالی بندرگز و بخش کوچکی در شمال شهرستان رامیان در استان گلستان بیشترین پتانسیل وقوع سیل را دارند. نتایج این مطالعه به برنامه‌ریزان، تصمیم‌گیرندگان و مدیران استان گلستان کمک می‌کند تا اقدامات مناسبی را برای پیشگیری و کاهش وقوع سیل در موقعیت‌های مکانی پیش‌بینی شده، انجام دهند.

واژگان کلیدی: پیش‌بینی سیل، سنجش از دور، یادگیری ماشینی، مدل گروهی، استان گلستان

۱- مقدمه

در میان مخاطرات مختلف آب و هوایی، سیل به عنوان مخرب ترین مخاطره در جهان شناخته می شود [۱]. سیل اغلب باعث تلفات شدید جانی و آسیب های اجتماعی-اقتصادی می شود [۲]. طبق داده های آماری دفتر کاهش خطر بلایا (UNISDR) سازمان ملل متحد، حوادث سیل منجر به ۱۱.۱٪ از تلفات جهانی بلایای طبیعی می شوند [۳] و [۴]. علاوه بر این، با توجه به پیش بینی های تغییرات اقلیمی، تغییرات الگوی کاربری اراضی و افزایش جمعیت نیز، پیش بینی می شود که میزان وقوع و شدت سیل ها احتمالاً تا سال ۲۰۵۰ تشدید شود که ممکن است خسارات زیادی را به همراه داشته باشد [۵]. لذا کنترل این بلای طبیعی و اقدامات پیشگیری از آن به دلیل خسارات بالقوه زیاد به منابع طبیعی، کشاورزی، حمل و نقل، پل ها و بسیاری از جنبه های دیگر زیرساخت های شهری، نیاز به فوریت دارد [۶] و [۷]. بنابراین، مطالعات بر روی مدل سازی و پیش بینی وقوع سیلاب که به صورت مکانی انجام می شود برای اتخاذ تدابیر انطباق و کاهش اثرات مخرب سیل، ضروری است [۸] و [۹]. سیل به عوامل مختلفی بستگی دارد، از جمله نوع استفاده از زمین، فاصله از رودخانه، پوشش گیاهی، خاک، ارتفاع و شیب [۱۰]. بنابراین، لازم است که تأثیر هر یک از این عوامل بر رخداد سیل در هر حوضه آبریز را درک و تعیین کنیم. امروزه، داده های ماهواره ای بهترین منبع اطلاعاتی برای تشخیص تغییرات هستند که به دلیل پوشش مناطق صعب العبور و ایجاد داده های سری زمانی، یک منبع اطلاعاتی موثر در پایش و بررسی تغییرات محیط زیستی بشمار می روند. در طول دو دهه گذشته، پیشرفت قابل توجهی در فناوری ماهواره ای حاصل شده است. با این حال، ظهور رادار دیافراگم مصنوعی (SAR) ظرفیت سنجش از راه دور برای نقشه برداری و پایش سیل را متحول کرده است [۱۱]. حسگرهای SAR می توانند مستقل از شرایط آب و هوایی کار کنند و مزیت جمع آوری داده ها را حتی از طریق پوشش ابری متراکم فراهم می کنند. توسعه اخیر ماموریت Sentinel-1 آژانس فضایی اروپا (ESA) تعداد زیادی از تصاویر SAR را به صورت رایگان در دسترس عموم قرار داده است، که نقطه عطف جدیدی را در ارزیابی سیل و کاهش خطر ایجاد کرده است [۱۲]. بهره گیری از این پیشرفت ها فرصت های جدیدی را برای درک مکان های مستعد وقوع سیل

جهت مدیریت در پیشگیری از بحران فراهم کرده است [۴] و [۱۳].

برای شناسایی و ارزیابی مناطق حساس به سیل، تا کنون از روش های متفاوتی استفاده شده است [۱۴]. برای نمونه، بعضی از تحقیقات از روش های تحلیل تصمیم چند معیاره (MCDA) مثل فرآیند تحلیل سلسله مراتبی (AHP) برای سنجش حساسیت سیل بهره گرفته اند [۱۵] که البته این روش ها بر پایه دانش تخصصی هستند که ممکن است با قضاوت های نامعین و عدم قطعیت منجر به نتایج نامطمئن شوند [۱۶]. دیگر تحقیقات از روش های آماری مثل نسبت فرکانس (FR) و رگرسیون لجستیک برای نقشه برداری سیل و مناطق مستعد خطر به کار برده اند [۸]. این روش ها وابسته به متغیرهای پیش بینی کننده هستند که بر اساس روابط پارامترهای توضیح دهنده مختلف تعیین می شوند و به شدت به اندازه داده ها وابسته هستند [۱۷]. مدل های فیزیکی مثل HEC-RAS و MIKE11 نیز برای مطالعه سیل ها کاربرد دارند [۱۸] اما نیاز به داده های زیاد، قدرت پردازش زیاد و پارامترهای زمانبر دارند [۱۹]. روش های جدیدتر شامل الگوریتم های یادگیری ماشین مثل جنگل تصادفی [۲۰]، شبکه های عصبی مصنوعی [۲۱]، ماشین های بردار پشتیبان [۱۰]، رگرسیون لجستیک [۲۲]، نسبت فرکانس درختان رگرسیون تقویت شده [۲۳]، مدل خطی تعمیم یافته [۲۰] هستند که قادر به شناسایی مناطق حساس به سیل هستند و نقشه های حساسیت سیل تولید می کنند [۱۰]. همگی این رویکردها، ضعف های خاصی را از نظر عدم قطعیت مدل های منفرد نشان می دهند که با یک رویکرد مدل سازی گروهی قابل بهبود است [۲۴]. در واقع مدل سازی مجموعه یک فرآیندی است که پیش بینی های مدل های منفرد را در یک مدل یکپارچه ترکیب می کند تا دقت پیش بینی را افزایش دهد [۲۵] و [۲۶] و [۲۷].

مطالعات اخیر نشان داده اند که به کارگیری الگوریتم های پیشرفته یادگیری ماشین، به ویژه مدل های گروهی (ensemble) و شبکه های عصبی با ورودی چندگانه، می تواند دقت مدل های حساسیت سیل را به طور معناداری افزایش دهد [۲۸]. برای مثال، Maddah et al. (2025) با بهره گیری از شبکه کانولوشن دو ورودی و تصاویر Sentinel-1، دقت نقشه های حساسیت سیل در حوضه گرگان رود را افزایش داده اند [۲۹]. همچنین، Arabameri et al. (2020a) با ترکیب روش های Real AdaBoost، Random Subspace،

MultiBoosting و J48 عملکرد برتر مدل‌های گروهی را در پیش‌بینی سیل نشان داده‌اند [۳۰]. در پژوهشی دیگر، Arabameri et al. (2020b) بارزایی مدل‌هایی نظیر RF و Bagging در دشت مرد، به دقت بالای 0.90 دست یافته‌اند. به علاوه، Bashirgonbad et al. (2024) در حوضه گاماسیاب با استفاده از مدل‌های RF و SVM و شاخص‌هایی چون AUC و TSS، تأثیر مثبت متغیرهای محیطی بر تعمیم‌پذیری مدل‌های گروهی را تأیید کرده‌اند [۳۱]. این یافته‌ها مؤید ضرورت استفاده از داده‌های چندمنبعی و روش‌های ترکیبی برای بهبود پیش‌بینی سیلاب در مقیاس‌های منطقه‌ای هستند.

ایران در اثر تغییرات اقلیمی و بارش‌های زیاد در مناطق گوناگون با حوادث سیلابی فراوانی روبرو شده است [۳۲]. استان گلستان که در بخش شمالی ایران واقع شده است، در سال‌های اخیر با بارش‌های کوتاه مدت و در نتیجه سیل‌های فراوانی مواجه شده است که خسارات اجتماعی و اقتصادی زیادی را به همراه داشته است [۳۳]. این منطقه در سال ۱۳۸۰ با بارندگی شدید مواجه شد که باعث ایجاد سیل و تلفات انسانی و خسارات اقتصادی و اجتماعی فراوان شد [۳۴]. در سال ۱۳۹۷ به دلیل اینکه حجم ۵۰ تا ۷۰ درصد میانگین بارش سالانه در یک دوره پنج روزه رخ داد (این مقدار بارش در ۷۰ سال گذشته بدون سابقه بوده است) مجدداً سیل بخشی از این استان را در بر گرفت. در سال ۱۳۹۸ نیز سیل بار دیگر بخشی از این استان را تحت تأثیر قرار داد (برای مرور دقیق به [۳۳] مراجعه شود). بنابراین، همین موضوع لزوم مطالعات دقیق‌تر برای پیش‌بینی احتمال سیل در این استان و ایجاد استراتژی مدیریت و مهار سیلاب را تشدید می‌کند.

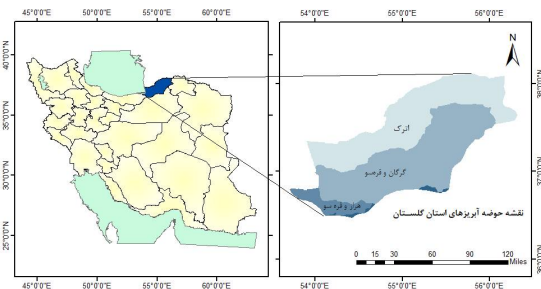
باتوجه به نگرانی‌های ذکر شده در بالا، هدف اصلی این مقاله استفاده از تصاویر پیشرفته Sentinel-1 (SAR) و روش‌های پیچیده یادگیری ماشینی است که می‌تواند رویکردی دقیق‌تر در نقشه برداری سیل و شناسایی عوامل کلی مؤثر بر وقوع سیل در این منطقه فراهم کند. از آنجا که مدل‌های یادگیری ماشینی بر اساس کمی کردن روابط بین دو گروه از داده‌ها شامل متغیرهای هدف و متغیرهای پیش‌بینی کننده کار می‌کنند، در این مطالعه داده‌های وقوع سیل از طریق تصاویر ماهواره‌ای Sentinel-1 و متغیرهای پیش‌بینی کننده بر اساس مرور مطالعات پیشین و با استفاده از مدل هیدرولوژیکی از طریق لایه‌ی مدل رقومی ارتفاعی تهیه شد. سپس برای بکارگیری الگوریتم‌های مختلف

یادگیری ماشینی از بسته آماری "sdm" استفاده شد. این بسته امکان بکارگیری رویکرد ترکیبی برای ترکیب نتایج مدل‌های فردی و ایجاد مدل‌های دقیق‌تر و پایدارتر را از طریق وزن‌دهی بیشتر به مدل‌هایی با عملکرد بالا را فراهم می‌کند. نتایج این مطالعه منجر به شناسایی محدوده‌های احتمال وقوع سیل و رتبه بندی اهمیت عوامل مختلف مؤثر و مدیران مقابله با بحران سیل کمک کند.

۲- مواد و روش‌ها

۲-۱- منطقه مورد مطالعه

استان گلستان با مساحت ۲۲,۰۳۳ کیلومتر مربع، ۱۰۷۵ درصد از مساحت کل کشور ایران را شامل می‌شود. این استان در قسمت شمال شرقی ایران قرار گرفته که عرض جغرافیایی آن بین ۳۶ درجه و ۳۰ دقیقه تا ۳۸ درجه و ۸ دقیقه شمالی و طول جغرافیایی آن بین ۵۳ درجه و ۵۷ دقیقه تا ۵۶ درجه و ۲۲ دقیقه شرقی است. جمعیت این استان بر اساس سرشماری سال ۱۳۹۰ بیش از ۱,۷۷۷,۰۰۰ نفر بوده است [۳۵]. منابع طبیعی، از جمله مراتع و مناطق طبیعی، ۷۰ درصد از مساحت استان را تشکیل می‌دهند. سه حوزه آبخیز مهم، گرگان‌رود و قره‌سو و اترک در این استان قرار دارد [۳۷]. در سال ۱۳۹۷، استان گلستان با سیل شدیدی در غرب خود مواجه شد که بین ۵۰ تا ۷۰ درصد از میانگین بارندگی سالانه خود را در یک دوره پنج روزه دریافت کرد و این میزان بارندگی در تمامی داده‌های هواشناسی، حداقل در ۷۰ سال گذشته در این منطقه بی‌سابقه بود [۳۷]. این سیلاب‌ها به شهرهایی مانند گنبدکاووس، بندرترکمان، آق‌قلا و مناطق دیگری که در حاشیه رودخانه واقع شده‌اند، آسیب زیادی وارد کرد [۳۸]. بیشترین آسیب‌ها در بخش ساخت و ساز، زیرساخت‌های شهری و کشاورزی بوده است [۳۹]. شکل ۱ محدوده مورد مطالعه در این پژوهش را نمایش می‌دهد.



شکل ۱- موقعیت جغرافیایی حوضه‌های آبریز استان گلستان

۲-۲- نقاط وقوع سیل، داده‌ها و مراحل پیش پردازش

برای شناسایی دقیق مناطق غرقابی ناشی از سیلاب، تصاویر راداری ماهواره Sentinel-1 در بازه زمانی ۲۰۱۵ تا ۲۰۲۲ از پایگاه داده آنلاین اتحادیه اروپا (eu.copernicus.scihub) استخراج و مورد پردازش قرار گرفتند. به منظور تعیین تاریخ دقیق وقوع سیلاب‌ها یا بارش‌های سنگین منجر به سیلاب ناگهانی، از داده‌های بارش ساعتی سه‌ساعته حاصل از مأموریت جهانی اندازه‌گیری بارش (GPM-IMERG) استفاده شد.

برای هر رویداد شناسایی‌شده، یک جفت تصویر Sentinel-1 از نوع Ground Range Detected (GRD) سطح ۱، شامل تصاویر قبل و بعد از سیلاب، با ویژگی‌های تصویربرداری مشابه انتخاب گردید. معیارهای انتخاب تصاویر شامل نزدیکی زمانی به تاریخ وقوع سیل، یکسان بودن مدار گذر (صعودی یا نزولی)، زاویه دید^۱ مشابه و استفاده از حالت تصویربرداری Interferometric Wide Swath (IW) بود.

در مرحله بعد، به منظور آماده‌سازی داده‌ها برای تحلیل تغییرات، مجموعه‌ای از پیش‌پردازش‌های استاندارد بر روی تصاویر SAR طبق روش‌های ارائه‌شده در [۴۰] انجام شد. این مراحل شامل کالیبراسیون رادیومتریک^۲، فیلتر حذف نویز اسپکل، تصحیح هندسی با استفاده از مدل رقومی ارتفاعی، و هم‌راستاسازی^۳ تصاویر قبل و بعد از سیلاب جهت مقایسه پیکسل به پیکسل بود. تمامی مراحل پیش‌پردازش بر پایه پلاریزاسیون VV انجام شد، زیرا این پلاریزاسیون در مطالعات مرتبط با سیلاب، به دلیل حساسیت بیشتر نسبت به تغییرات رطوبت سطحی، عملکرد دقیق‌تری در شناسایی مناطق آب‌گرفته دارد.

در این نوع تصاویر ماهواره‌ای، نوعی اختلال وجود دارد که به آن نویز اسپکل گفته می‌شود [۴۱]. این نویز که وابسته به سیگنال است به دلیل نوسانات فاز سیگنال‌های بازگشتی امواج الکترومغناطیسی بوجود می‌آید که به صورت نقطه نقطه ظاهر می‌شود که باعث تداخل سیگنال‌های بازتابیده شده و ایجاد دانه‌های ریز و درشت در تصویر به دست آمده می‌شود

و به بافت فلفل نمکی مشهور است [۴۲]. برای رفع این مشکل از فیلتر حذف لکه استفاده شد که علاوه بر حفظ وضوح تصویر، با محاسبه ارزش پیکسل از پیکسل‌های اطراف، ارزش پیکسل مورد نظر به دست می‌آورد [۴۳]. علاوه بر نویزها، تصاویر ماهواره‌ای همواره با خطاهای سیستمی و غیرسیستمی همراه هستند که معمولاً به دلیل حرکات و سیستم سنجنده است، برای رفع این مشکلات از تصحیح هندسی استفاده شد [۴۴]. سپس از یک تکنیک آستانه‌بندی کالیبراسیون استفاده کردیم که مقادیر شدت بازتاب پراکنش در مناطق غرقابی و غیرغرقابی را مقایسه می‌کند. برای تمایز پیکسل‌های مربوط به آب‌های دائمی (مانند دریاچه‌ها)، از شاخص آب نرمال‌سازی شده تغییر یافته^۴ (فرمول ۱)

$$NDWI = \frac{GREEN\ BAND - NIR\ BAND}{GREEN\ BAND + NIR\ BAND} \quad (1)$$

حاصل از تصاویر چندطیفی Sentinel-2 سطح 1-C استفاده کردیم [۴۵] و آستانه‌ای تعیین کردیم که می‌تواند مناطق آب‌های دائمی و مناطق سیل زده را شناسایی کند. سپس تصاویر پس و پیش از رویداد سیل با استفاده از آستانه -dB- ۲۱ باینری شدند [۴۵] و از یک روش مقایسه نقشه باینری برای شناسایی مناطق غرقابی استفاده شد. همه فرآیندهای جمع‌آوری و پردازش داده‌های این بخش در پلتفرم SNAP از نرم‌افزار QGIS به نقاط وقوع سیل تبدیل شدند.

۲-۳- پردازش داده‌های محیطی

بررسی پژوهش‌های پیشین نشان می‌دهد با توجه به ویژگی‌های محلی، هشت عامل محیطی به طور قابل توجهی بر وقوع سیل تاثیر گذار هستند [۴۶] که شامل مدل رقومی ارتفاع^۵، شیب توپوگرافی، شاخص رطوبت توپوگرافی^۶، فاصله جریان، جهت جریان، تجمع جریان، شاخص گیاهی تفاوت نرمال شده^۷ و بارش (شکل ۲). لایه مدل رقومی ارتفاع با وضوح فضایی ۹۰ متر، از پایگاه داده مأموریت توپوگرافی رادار شاتل^۸ تهیه شده‌است. پر کردن پیکسل‌های سینک برای جلوگیری از هرگونه خطا در تحلیل‌های هیدرولوژیکی از قبل پردازش شد. عملیات پر کردن سینک یک مقدار ارتفاع را در

۵ Digital Elevation Model

۶ Topographic Wetness Index

۷ The Normalized Difference Vegetation Index

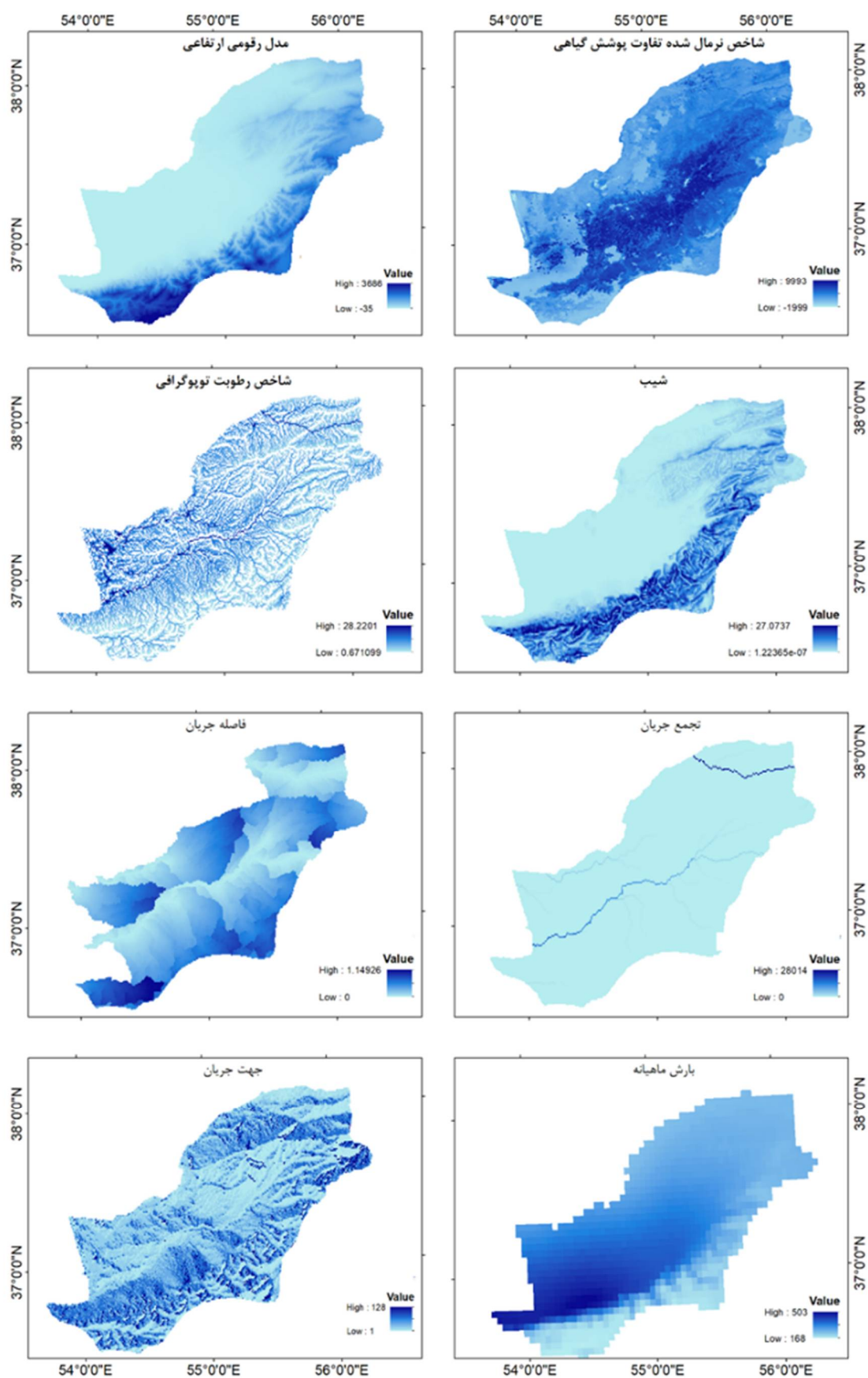
۸ NASA Shuttle Radar Topographic Mission

۱ incidence angle

۲ Radiometric calibration

۳ co-registration

۴ Normalized Difference Water Index



شکل ۲- پارامترهای محیطی موثر بر وقوع سیل

یک پیکسل تغییر می دهد که در مقایسه با ۸ پیکسل همسایه آن کوچکتر است [۴۷]. با استفاده از مدل رقومی ارتفاع به عنوان ورودی، شیب توپوگرافی، و جهت جریان در هر پیکسل تولید شد. جهت جریان در هر پیکسل با استفاده از الگوریتم D8 در ESRI ArcGIS Pro تولید شد که جهت جریان را در هر پیکسل به سمت شیب دارترین همسایه پایین شیب خود تعیین میدهد. سپس تجمع جریان به عنوان وزن انباشته تمام پیکسل های بالادستی که به هر پیکسل پایین شیب جریان می یابند، تعیین شد. بزرگی جریان ها به سمت پایین دست به شرایط دیگر درون پیکسل ها نیز بستگی دارد (به عنوان مثال، تراکم پوشش گیاهی)، بنابراین، شاخص گیاهی تفاوت نرمال شده بر اساس باندهای مادون قرمز تصاویر ماهواره ای، به دست آمده از Sentinel-2 در ده دوره سالانه (از ۱۳۹۱ تا ۱۴۰۱) محاسبه شد و خطر نسبی فاصله جریان به عنوان یکی دیگر از عوامل سیل در نظر گرفته شد. ابتدا، شبکه های جریان با اعمال آستانه ای از تجمع جریان استخراج شد به طوری که پیکسل هایی با تجمع جریان بیشتر از آستانه (۵۰۰) به عنوان جریان طبقه بندی شدند. سپس فواصل جریان افقی و عمودی تا نزدیک ترین جریان با استفاده از شبکه جریان و جهت جریان محاسبه شد و این مقادیر به همراه مقدار تجمع جریان به یک مقدار واحد از ریسک فاصله نسبی جریان بین ۰ و ۱ تبدیل شدند. دلیل این امر این است که فاصله کمتر با نهرا با تجمع جریان بیشتر، خطر سیل را افزایش می دهد.

عامل بعدی شاخص رطوبت توپوگرافی است (فرمول ۲) که یکی از مهم ترین پارامترهای هیدرولوژیکی است که میزان پتانسیل تجمع رطوبت در سطح زمین را بر اساس موقعیت توپوگرافی مشخص می کند [۲۰]:

$$TWI = \ln(\alpha / \tan \beta) \quad (2)$$

که پارامتر α نشان دهنده سطح حوضه آبریز با واحد متر مربع است و پارامتر β نشان دهنده شیب سطح زمین با واحد درجه است.

مقادیر بالاتر TWI نشان دهنده مناطقی هستند که آب در آنها تجمع بیشتری دارد و در نتیجه مستعد اشباع شدن و وقوع سیلاب اند، در حالی که مقادیر پایین تر TWI معمولاً در مناطق مرتفع یا شیب دار دیده می شود که آب به سرعت از آنها تخلیه می شود.

۳- مدلسازی پتانسیل وقوع سیل با تکنیک های یادگیری ماشینی

در این پژوهش از بسته نرم افزاری sdm در نرم افزار R استفاده شد [۲۴]. این بسته آماری یک الگوریتم یادگیری ماشینی شامل مجموعه ای از مدل ها و چندین روش جهت ارزیابی و پیش بینی مکانی است [۴۸]. دو نوع داده در مطالعات پیش بینی سیل شامل لایه های محیطی، که برگرفته از ویژگی های طبیعی و رفتاری منطقه مورد مطالعه و نقاط وقوع سیل، امکان پیش بینی سیل را با استفاده از این روش فراهم می سازد. در این روش ۱۰۰۰ مکان تصادفی در سراسر منطقه مورد مطالعه به عنوان پس زمینه یا شبه فقدان رویداد سیل تولید شد، که برای آموزش مدل های یادگیری ماشین، زمانی که یک متغیر پاسخ دوجمله ای است ضروری است [۴۹]. با ترکیب هفت الگوریتم یادگیری ماشین (جدول ۱) شامل مدل خطی تعمیم یافته (GLM^۱)، مدل افزایشی تعمیم یافته (GAM^۲)، رگرسیون درختی پیشرفته (BRT^۳)، مدل جنگل تصادفی (RF^۴)، مدل رگرسیون سازشی چندمتغیره (MARS^۵)، مدل آنالیز تفکیکی انعطاف پذیر (FDA^۶)، درختان طبقه بندی و رگرسیون (CART^۷)، مدل ترکیبی (ensemble) ایجاد شد. واضح است هر یک از این مدل ها به تنهایی دارای عدم قطعیتی هستند که قابل چشم پوشی نیست، به همین علت بهترین راهکار جهت کاهش عدم قطعیت در پیش بینی ها، ترکیب مدل ها و استفاده از روش ترکیبی است [۵۰]. قبل از آموزش مدل ها، ما مشکل چند خطی در متغیرهای پیش بینی کننده را با محاسبه ضریب تورم واریانس (VIF) از طریق یک روش گام به گام در بسته usdm در محیط R بررسی کردیم [۵۱].

۶ Flexible Discriminant Analysis

۷ Classification and Regression Trees

۸ Area Under the Curve

۹ Correlation

۱۰ True Skill Statistic

۱ Generalized Linear Model

۲ Generalized Additive Model

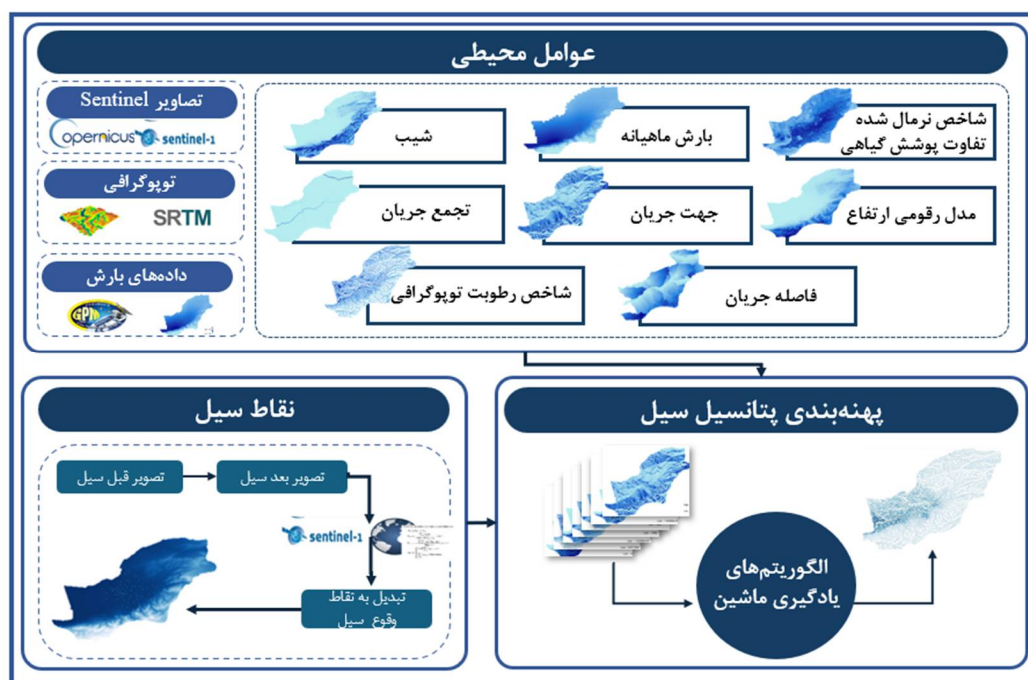
۳ Boosted Regression Trees

۴ Random Forest

۵ Multivariate Adaptive Regression Splines

جدول ۱- مدل‌های مورد استفاده در پژوهش و شرح عملکرد هر یک از آنها

مدل	اختصار	کارکرد مدل
خطی تعمیم یافته	GLM	تابع رگرسیون مورد استفاده برای مدل‌سازی، داده‌های حضور و عدم حضور است و از تابع رگرسیون (LR) برای ارتباط بین نقاط وقوع سیل با پارامترهای محیطی استفاده شده. در این مدل احتمال وقوع سیل و اهمیت متغیرهای محیطی در هر سلول نشان داده می‌شود.
افزایشی تعمیم یافته	GAM	الگوریتم یادگیری ماشین که با اضافه کردن توابع اسپلاین به مدل خطی تعمیم یافته، یک مدل انعطاف‌پذیر و ناپارامتری برای پیش‌بینی متغیرهای پاسخ از روی متغیرهای توضیحی ارائه می‌کند. GAM مزایایی مانند قابلیت تفسیر، کنترل بر روی بیش برازش و توانایی مدل‌سازی روابط غیرخطی و تعاملات را دارد.
رگرسیون درختی پیشرفته	BRT	الگوریتم یادگیری ماشین که با استفاده از تکنیک بوستینگ (boosting)، چندین درخت تصمیم را به صورت ترتیبی و تکاملی ساخته و ترکیب می‌کند. مزایایی مانند تطبیق پذیری بالا، تفسیرپذیری قوی، کنترل خوب بر روی بیش برازش و توانایی مدل‌سازی روابط غیرخطی و تعاملات را دارد.
جنگل تصادفی	RF	الگوریتم یادگیری ماشین که با ساختن و رای گیری بین چندین درخت تصمیم دسته‌بندی، رگرسیون و دیگر وظایف را انجام می‌دهد. این روش مقاوم، قابل تفسیر، کارآمد و انعطاف‌پذیر است.
رگرسیون سازشی چند متغیره	MARS	الگوریتم یادگیری ماشین که با استفاده از توابع پایه‌ای شکسته (broken-stick functions)، یک مدل خطی قطعه قطعه را برای پیش‌بینی متغیر وابسته از متغیرهای مستقل ایجاد می‌کند. این روش می‌تواند با داده‌های ناهمسان، ناپویسته و نامتوازن کار کند و تعاملات بین متغیرها را به صورت خودکار شناسایی کند. این روش با استفاده از دو مرحله اصلی، گسترش و هرس، مدل را ساخته و بهینه می‌کند.
آنالیز تفکیکی انعطاف پذیر	FDA	الگوریتم یادگیری ماشین که بین ورودی‌ها و خروجی‌های دستمای رابطه برقرار می‌کند، سعی می‌کند تفاوت‌های بین گروه‌های مختلف را بشکند. این مدل معمولاً با استفاده از روش حداقل مربعات جزئی (PLS) یا روش تحلیل عاملی (FA) آموزش داده می‌شود.
درختان طبقه بندی و رگرسیون	CART	الگوریتم یادگیری ماشین که با استفاده از ساختار درختی، داده‌ها را بر اساس ویژگی‌های مختلف به زیر گروه‌های همگن تقسیم می‌کند. این روش می‌تواند برای حل مسائل طبقه‌بندی یا رگرسیون استفاده شود. این روش با استفاده از شاخص ناخالصی جینی به دنبال بهترین ویژگی و مقدار آستانه برای تقسیم هر گره است.



شکل ۳- فلوچارت کلی روند اجرا

سپس برای ارزیابی عملکرد مدل‌ها از چهار شاخص مختلف شامل AUC^h ، COR^h ، TSS^h و Deviance استفاده شد که در میان این شاخص‌ها، شاخص‌های TSS و Deviance وابسته به آستانه و شاخص‌های AUC و COR مستقل از آستانه هستند. مقادیر AUC ، Deviance، COR ، بین ۱ تا ۰ متغیر است و مقادیر TSS بین -۱ تا +۱ متغیر است. افزایش میزان عددی شاخص‌های TSS ، COR ، AUC نشان دهنده افزایش کارایی مدل‌ها و افزایش میزان عددی Deviance نشان دهنده کاهش کارایی مدل‌ها است [۵۲].

۴- نتایج

۴-۱- نقاط وقوع سیل، داده‌ها و مراحل پیش پردازش

در این مطالعه با استفاده از داده‌های راداری Sentinel-1 و داده‌های بارندگی ساعتی در بازه زمانی ۲۰۱۵ تا ۲۰۲۲، نواحی متأثر از رخداد‌های سیل شناسایی و استخراج شدند تا مورد تحلیل قرار گیرد. در مجموع ۳۷ نقطه وقوع سیلاب با استفاده از چندضلعی‌های حاصل از رویدادهای سیل مشخص گردید و به عنوان متغیر وابسته در فرایند مدل‌سازی خطر سیلاب مورد استفاده قرار گرفتند.

در این تحقیق، هشت متغیر مستقل (پیش‌بینی کننده) شامل پارامترهایی چون شیب، جهت شیب، ارتفاع، فاصله از رودخانه، کاربری اراضی، شاخص پوشش گیاهی، بارندگی و تراکم زهکشی وارد مدل‌سازی شدند. برای بررسی عدم وجود هم‌خطی بین متغیرها، از آزمون ضریب تورم واریانس (VIF) استفاده شد. نتایج این آزمون نشان داد که تمامی متغیرهای مستقل دارای مقادیر کمتر از آستانه ۱۰ بودند، که این امر بیانگر نبود هم‌خطی شدید بین متغیرها و مناسب بودن آن‌ها برای ورود به مدل است (جدول ۲).

جدول ۲- ارزیابی عملکرد مدل‌های پیش‌بینی توزیع سیل

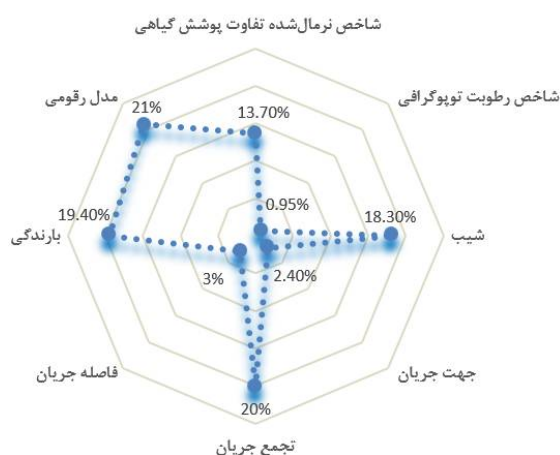
میزان ضریب تورم واریانس	متغیر	میزان ضریب تورم واریانس	متغیر
۶.۳	تجمع جریان	۴.۲	مدل رقومی ارتفاع
۴.۵	فاصله جریان	۳.۸	شاخص نرمال شده تفاوت پوشش گیاهی
۳.۹	بارش ماهیانه	۴.۱	شیب توپوگرافی
۳.۷	جهت جریان	۵.۹	شاخص رطوبت گیاهی

۴-۲- اعتبار سنجی مدل‌ها و اهمیت متغیرهای محیطی

جدول شماره ۳ نتایج صحت‌سنجی مدل‌ها با شاخص‌های AUC ، COR ، TSS ، Deviance را نشان می‌دهد که در تمام مدل‌ها هر سه شاخص AUC ، TSS ، COR مقداری بالاتر از ۰.۷۵ دارند که نمایانگر عملکرد مورد قبول برای تمام مدل‌هاست. با این حال بنظر می‌رسد شاخص COR و TSS وضعیت بهتری برای مدل‌های مورد استفاده نشان می‌دهند اما شاخص‌های AUC و Deviance شاخص‌های سختگیرانه‌تری در صحت‌سنجی مدل‌ها، به‌خصوص برای مدل CART هستند. بر اساس این نتایج، در بین مدل‌های منفرد، مدل RF بهترین عملکرد را نشان داده است. از طرف دیگر نتایج تجزیه و تحلیل اهمیت متغیرهای محیطی نشان داد مدل رقومی ارتفاعی، تجمع جریان و بارندگی و شیب با اهمیت نسبی به ترتیب ۲۱٪ و ۲۰٪ و ۱۹.۴۰٪ و ۱۸.۳۰٪ مهم‌ترین متغیرها در پیش‌بینی وقوع سیل هستند که بر مدل نهایی تاثیر بیشتری داشته‌اند (شکل ۳).

جدول ۳- ارزیابی عملکرد مدل‌های پیش‌بینی توزیع سیل

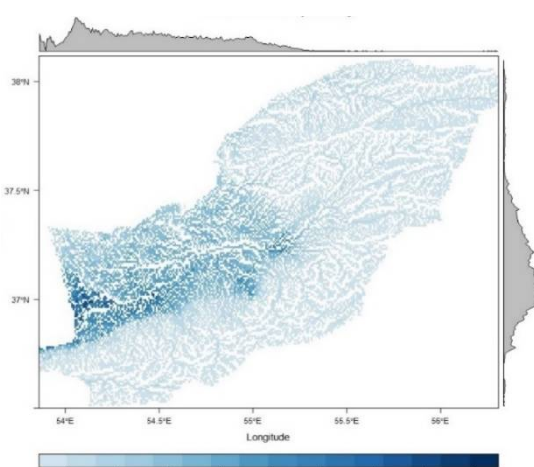
	AUC	COR	TSS	Deviance
GAM	۰.۸۹	۰.۹۴	۰.۸۷	۰.۲۶
GLM	۰.۷۸	۰.۸۴	۰.۸۴	۰.۱۵
RF	۰.۹۴	۰.۹۷	۰.۹۶	۰.۱۴
MARS	۰.۸۸	۰.۸۹	۰.۹۱	۰.۲۱
CART	۰.۷۵	۰.۸۶	۰.۸۸	۰.۲۷
BRT	۰.۸۶	۰.۸۲	۰.۹۴	۰.۱۶
FDA	۰.۸۹	۰.۹۳	۰.۷۸	۰.۱۱



شکل ۳- اهمیت هر یک از متغیرهای محیطی در پیش‌بینی سیل

۳-۴- پیش بینی وقوع سیل بر اساس مدل ترکیبی

بر اساس مدل ترکیبی، شهرستان گمیشان، آق‌قلا، بندر ترکمن، بخش‌های غربی گنبدکاووس، بخش شمالی بندرگز و بخش کوچکی در شمال شهرستان رامیان در استان گلستان بیشترین پتانسیل وقوع سیل را نشان داده است. بر جنوب غربی گمیشان و شمال غربی بندر ترکمن نشان داده است. بطور کلی پیش بینی این مطالعه نشان می‌دهد ناحیه‌ی شمال غربی استان گلستان پتانسیل بالاتری برای وقوع سیل دارد (شکل ۴).



شکل ۴- نقشه پیش بینی وقوع سیل در استان گلستان بر اساس مدل ترکیبی

۵- بحث و نتیجه‌گیری

شناسایی و پهنه بندی مناطق سیل خیز در یک حوزه آبخیز برای کاهش خسارات ناشی از سیل ضروری است [۵۳] و [۵۴] تاکنون چندین متدولوژی توسط محققان در سراسر جهان برای پیش‌بینی سیل توسعه داده شده است [۵۵] که هر کدام دارای اشکالات و مزایای هستند. سنجش از دور (RS) نتایج بهتری نسبت به مدل‌های منفرد ارائه می‌دهند [۵۶]. از این مدل‌ها می‌توان در توسعه و برنامه‌ریزی‌های مدیریتی استفاده کرد [۵۷]. در این مطالعه ما برای تولید نتایجی با قابل اطمینان بالا برای پیش‌بینی مناطق با پتانسیل وقوع سیل در استان گلستان، با ادغام تصاویر پیشرفته سنجش از دور و مدل‌های یادگیری ماشینی و در نهایت به کارگیری از یک مدل گروهی نسبت به پیش‌بینی مدل مکانی سیل پرداختیم. در نتیجه می‌توان اذعان داشت، این مطالعه به ارائه

بینشی در مورد مکان‌های با پتانسیل بالای سیل‌خیزی در استان گلستان می‌پردازد.

عملکرد بالای مدل جنگل تصادفی در این مطالعه با مطالعات قبلی در این زمینه [۵۸] و [۵۹] همخوانی دارد. RF مدلی است که با استفاده از تعدادی درخت رگرسیون ساده، یک مدل پیچیده و دقیق ایجاد می‌کند. برای افزایش کارایی مدل، RF به صورت تصادفی از بخش‌های مختلف داده‌ها و متغیرها برای ساخت هر درخت استفاده می‌کند. این کار باعث کاهش واریانس و خطای مدل و جلوگیری از برازش بیش از حد می‌شود [۶۰] و [۱۴]. از همین رو در مطالعات پیشین نیز در میان مدل‌های منفرد، مدل RF بیشترین عملکرد را نشان داده است.

شرط موفقیت و عملکرد یک الگوریتم یادگیری ماشین در مرحله آموزش، بستگی به انتخاب دقیق و صحیح عوامل مؤثر دارد [۶۱]. بر اساس مرور منابع، مهم‌ترین عوامل پیش‌بینی‌کننده سیل در مطالعه لحاظ شد و در بین هشت پارامتر تأثیرگذار محیطی (شکل ۲)، مدل رقومی ارتفاعی، تجمع جریان و بارندگی و شیب تأثیرگذاری بیشتری در ایجاد سیل منطقه داشتند که با مطالعات [۶۲] تطابق دارد. مدل رقومی ارتفاع یک منبع اطلاعاتی مهم برای شناسایی توپوگرافی حوضه آبخیز و هندسه رودخانه است. توپوگرافی حوضه آبخیز تأثیر بسزایی بر جریان آب و فرآیندهای هیدرولوژیک دارد [۴۷]. در واقع مدل رقومی ارتفاع، شکل، شیب، جهت و طول حوضه آبخیز، مقدار و توزیع بارندگی، نحوه جذب و رسوب آب، سرعت و حجم جریان و خطر سیلاب را تحت تأثیر قرار می‌دهد [۶۲]. تجمع جریان نشان دهنده قابلیت حیاتی یک حوضه آبریز در تولید جریان است. به عبارت دیگر، هرچه قدر تجمع جریان در یک نقطه بیشتر باشد، خروج جریان در آن نقطه بیشتر خواهد بود [۶۳]. این پارامتر یک عامل مهم در تعیین مناطق سیل‌خیز و مسیرهای جریان سیل است که می‌تواند به عنوان یک شاخص برای اولویت‌بندی مناطق حساس به سیل و انتخاب مکان‌های مناسب برای احداث سازه‌های کنترل سیل مورد استفاده قرار گیرد [۳۶]. بارندگی به دلیل اینکه باعث افزایش حجم آب در رودخانه‌ها و دشت‌های سیلابی می‌شود، یکی از مهم‌ترین پارامترهای مؤثر بر سیل در منطقه است [۴۷]. همچنین پارامتر شیب باعث تأثیرگذاری سرعت جریان آب بر روی سطح زمین می‌شود. با افزایش شیب، سرعت جریان آب بالاتر و امکان سیل‌خیزی بیشتر می‌شود [۲۴].

نتایج مطالعه ما همخوانی بالایی با مطالعات دیگر در مورد پتانسیل سیل خیزی در استان گلستان دارد. به عنوان مثال، [۶۴] با استفاده از روش‌های آمار دو متغیره، تصمیم‌گیری چند معیاره و یادگیری ماشین، نقشه پتانسیل سیل در استان گلستان را تهیه کرده‌اند. آن‌ها نشان داده‌اند که مناطق شمال غرب استان گلستان دارای پتانسیل سطح بالای سیل هستند. علاوه بر این [۶۵]، با استفاده از سامانه اطلاعات جغرافیایی (GIS)، ارزیابی خطر سیل را در استان گلستان انجام داده‌اند. آن‌ها نیز مناطق شمال غربی استان را به عنوان مناطق با خطر سیلابی بالا شناسایی کردند. همچنین [۳۶]، با استفاده از GIS، فرآیند اسکان پس از سیل را در استان گلستان مورد بررسی قرار داده‌اند. آن‌ها نشان داده‌اند که مناطق شمال غربی استان گلستان به شدت تحت تأثیر سیل قرار گرفته و نیاز به بهبود فرآیند اسکان دارند.

بطور کلی نتایج حاصل از این مطالعه برای بررسی و مدیریت بحران در جهت کاهش تخریب‌های ناشی از سیل، ایجاد نقشه‌های دقیق پتانسیل وقوع سیل در منطقه مورد مطالعه که اخیراً در معرض سیل‌های متعدد بوده است، ضروری بنظر می‌رسید. در این مطالعه هفت مدل یادگیری ماشینی و آماری به کار برده شده تا مناطق مستعد سیل مورد مطالعه قرار گیرند و پیش‌بینی سیل در این مناطق برای آینده انجام شود. تنوع در بین پیش‌بینی مدل‌های فردی، نشانگر آن است که هر مدل در مناطق خاصی پیش‌بینی دقیق‌تری داشته‌اند، اما ترکیب مدل‌های فردی در غالب رویکرد همادی یک راه‌حل نوین با قابلیت پیش‌بینی بالا به‌عنوان یک رویکرد نو در این پژوهش معرفی گردید. براساس نتایج به دست آمده شهرستان گمیشان، آق‌قلا، بندرترکمن، بخش‌های غربی گنبد کاووس، بخش شمالی بندرگز و بخش کوچکی در شمال شهرستان رامیان در استان گلستان بیشترین احتمال وقوع سیل را دارند که بهره‌گیری از این نتایج سبب کاهش خسارات

و بهبود عملکرد مدیریتی می‌شود. توجه به نتایج مطالعات شناسایی مناطق دارای پتانسیل وقوع سیل، از جمله این مطالعه، می‌تواند زیان مالی و اقتصادی را در آینده به میزان قابل توجهی کاهش دهد و سبب جلوگیری از توسعه مناطق مسکونی یا تجاری در مناطق آسیب پذیر شود. در کنار یافته‌های این مطالعه، وجه تمایز آن با پژوهش‌های پیشین در سه محور اساسی قابل بیان است. نخست، بهره‌گیری از داده‌های چندمنبعی شامل تصاویر راداری Sentinel-1 برای استخراج واقعی نواحی غرقابی، تصاویر نوری Sentinel-2 برای حذف آب‌های دائمی و محاسبه شاخص NDVI، و داده‌های بارش ساعتی ماهواره‌ای GPM برای تعیین زمان دقیق وقوع سیل، به همراه لایه‌های محیطی متنوع و با دقت بالا، پایگاه داده‌ای جامع‌تر و کاربردی‌تر از بیشتر مطالعات مشابه فراهم کرده است. دوم، این پژوهش با رویکردی عملیاتی و در مقیاسی منطقه‌ای، برای نخستین‌بار در سطح استان گلستان - به عنوان یکی از مناطق حساس به سیل در کشور - با استفاده از تصاویر چندساله (۲۰۱۵ تا ۲۰۲۲) و به‌کارگیری مدل‌های ترکیبی یادگیری ماشین اجرا شده که نسبت به مطالعات قبلی که عموماً در حوضه‌های محدود و با مدل‌های منفرد انجام شده‌اند، جامع‌تر است. سوم، استفاده از هفت الگوریتم یادگیری ماشین و یک مدل ترکیبی همادی به همراه ارزیابی دقیق عملکرد مدل‌ها با شاخص‌های مختلف (AUC)، (COR)، (TSS)، (Deviance)، بررسی هم‌خطی با آزمون VIF، و تحلیل وزن نسبی متغیرها، باعث افزایش دقت پیش‌بینی و قابلیت اعتماد در نتایج نهایی شده‌است. علاوه بر این، استفاده از نقاط واقعی وقوع سیل به‌عنوان متغیر پاسخ در آموزش مدل‌ها، دقت نتایج را نسبت به مدل‌هایی که صرفاً بر پایه احتمال یا شاخص‌های محیطی عمل می‌کنند، افزایش داده و کاربردپذیری آن را برای برنامه‌ریزی‌های مدیریت بحران و توسعه شهری ارتقاء می‌دهد.

مراجع

- [۱] Towfiqul Islam, A.R.M., Talukdar, S., Mahato, S., Kundu, S., Eibek, K.U., Pham, Q.B., Kuriqi, A., Linh, N.T.T. (2021). "Flood susceptibility modelling using advanced ensemble machine learning models." Vol.Geoscience Frontiers, No.12(3), PP.101075.
- [۲] Costache, R. (2019). "Flash-flood Potential Index mapping using weights of evidence, decision Trees models and their novel hybrid integration." Vol.Stochastic Environmental Research and Risk Assessment, No.33(7), PP.1375-1402.
- [۳] Bubeck, P., Thielen, A.H. (2018). "What helps people recover from floods? Insights from a survey among flood-affected residents in Germany." Vol.Regional environmental change, No.18, PP.287-296.

- [۴] Hong, H., Panahi, M., Shirzadi, A., Ma, T., Liu, J., Zhu, A.X., Chen, W., Kougias, I., Kazakis, N. (2018a). "Flood susceptibility assessment in Hengfeng area coupling adaptive neuro-fuzzy inference system with genetic algorithm and differential evolution." Vol. Science of the total Environment, No.621, PP.1124–1141.
- [۵] Huang, K., Li, X., Liu, X., Seto, K.C. (2019). "Projecting global urban land expansion and heat island intensification through 2050." Vol.Environmental Research Letters, No.14(11), P.114037.
- [۶] Alvarado-Aguilar, D., Jiménez, J.A., Nicholls, R.J. (2012). "Flood hazard and damage assessment in the Ebro Delta (NW Mediterranean) to relative sea level rise." Vol.Natural Hazards, No.62, PP.1301–1321.
- [۷] Billa, L., Shattri, M., Mahmud, A.R., Ghazali, A.H. (2006). "Comprehensive planning and the role of SDSS in flood disaster management in Malaysia." Vol.Disaster Prevention and Management: An International Journal, No.15(2), PP.233–240.
- [۸] Siahkamari, S., Haghizadeh, A., Zeinivand, H., Tahmasebipour, N., Rahmati, O. (2018). "Spatial prediction of flood-susceptible areas using frequency ratio and maximum entropy models." Vol.Geocarto international, No.33(9), PP.927–941.
- [۹] Su, J., Lü, H., Zhu, Y., Cui, Y., Wang, X. (2019). "Evaluating the hydrological utility of latest IMERG products over the Upper Huaihe River Basin, China." Vol.Atmospheric Research, No.225, PP.17–29.
- [۱۰] Tehrany, M.S., Pradhan, B., Jebur, M.N. (2015a). "Flood susceptibility analysis and its verification using a novel ensemble support vector machine and frequency ratio method." Vol.Stochastic environmental research and risk assessment, No.29, PP.1149–1165.
- [۱۱] Li, Y., Martinis, S., Wieland, M. (2019). "Urban flood mapping with an active self-learning convolutional neural network based on TerraSAR-X intensity and interferometric coherence." Vol.ISPRS Journal of Photogrammetry and Remote Sensing, No.152, PP.178–191.
- [۱۲] Benzougagh, B., Frison, P.L., Meshram, S.G., Boudad, L., Dridri, A., Sadkaoui, D., Mimich, K., Khedher, K.M. (2022). "Flood Mapping Using Multi-temporal Sentinel-1 SAR Images: A Case Study—Inaouene Watershed from Northeast of Morocco." Vol.Iranian Journal of Science and Technology, Transactions of Civil Engineering, No.46, PP.1481–1490.
- [۱۳] Sachdeva, S., Kumar, B. (2022). "Flood susceptibility mapping using extremely randomized trees for Assam 2020 floods." Vol. Eco Inform, No.67, P.101498.
- [۱۴] Zhao, G., Pang, B., Xu, Z., Yue, J., Tu, T. (2018). "Mapping flood susceptibility in mountainous areas on a national scale in China." Vol. Science of the Total Environment, No.615, PP.1133–1142.
- [۱۵] Tang, Z., Zhang, H., Yi, S., Xiao, Y. (2018). "Assessment of flood susceptible areas using spatially explicit, probabilistic multi-criteria decision analysis." Vol.Journal of Hydrology, No.558, PP.144–158.
- [۱۶] Miles, R.E., Snow, C.C. (1984). "Designing strategic human resources systems." Vol.Organizational dynamics, No.13, PP.36–52.
- [۱۷] McLay, C.D.A., Dragten, R., Sparling, G., Selvarajah, N. (2001). "Predicting groundwater nitrate concentrations in a region of mixed agricultural land use: A comparison of three approaches. Environ." Vol.Environmental pollution, No.115, PP.191–204.
- [۱۸] Alaghmand, S., bin Abdullah, R., Abustan, I., & Eslamian, S. (2012). "Comparison between capabilities of HEC-RAS and MIKE11 hydraulic models in river flood risk modelling (a case study of Sungai Kayu Ara River basin, Malaysia)." Vol.International Journal of Hydrology Science and Technology, No.2(3), PP.270-291.
- [۱۹] Hong, H., Tsangaratos, P., Ilia, I., Liu, J., Zhu, A.X., Chen, W. (2018b). "Application of fuzzy weight of evidence and data mining techniques in construction of flood susceptibility map of Poyang County, China." Vol.Science of the total environment, No.625, PP.575–588.
- [۲۰] Chapi, K., Singh, V.P., Shirzadi, A., Shahabi, H., Bui, D.T., Pham, B.T., Khosravi, K. (2017). "A novel hybrid artificial intelligence approach for flood susceptibility assessment." Vol.Environmental modelling & software, No.95, PP.229–245.
- [۲۱] Zou, J., Han, Y., & So, S. S. (2009). "Overview of artificial neural networks." Vol. Artificial neural networks: Methods and applications, PP. 14-22.
- [۲۲] Tehrany, M.S., Pradhan, B., Jebur, M.N. (2013). "Spatial prediction of flood susceptible areas using rule based decision tree (DT) and a novel ensemble bivariate and multivariate statistical models in GIS." Vol.Journal of hydrology, No.504, PP.69–79.
- [۲۳] Rahmati, O., Pourghasemi, H.R. (2017). "Identification of Critical Flood Prone Areas in Data-Scarce and Ungauged Regions: A Comparison of Three Data Mining Models." Vol.Water resources management, No.31, PP.1473–1487.

- [۲۴] Naimi, B., Araújo, M.B. (2016). "sdm: a reproducible and extensible R platform for species distribution modelling." Vol.Ecography, No.39(4), PP.368–375.
- [۲۵] Ebrahimi, E., Sayahnia, R., Ranjbaran, Y., Vaissi, S., Ahmadzadeh, F. (2021). "Dynamics of threatened mammalian distribution in Iran ' s protected areas under climate change." Vol.Mammalian Biology, No.101(6), PP.759-774.
- [۲۶] Ebrahimi, E., Ahmadzadeh, F. (2022). "Dynamics of habitat changes as a result of climate change in Zagros Mountains Range (Iran), a case study on Amphibians." Vol.Nova Biologica Reperta, No.9(1), PP.29–39.
- [۲۷] Mohammadi, S., Ebrahimi, E., Shahriari Moghadam, M., Bosso, L. (2019). "Modelling current and future potential distributions of two desert jerboas under climate change in Iran." Vol.Ecological Informatics, No.52, PP.7–13.
- [۲۸] Maddah, S., Mojaradi, B., & Alizadeh, H. (2025). Improving deep learning-based flood susceptibility modeling by integrating data balancing technique and dual-input convolutional neural network. Natural Hazards, 1–23.
- [۲۹] Arabameri, A., Saha, S., Mukherjee, K., Blaschke, T., Chen, W., Ngo, P. T. T., & Band, S. S. (2020a). Modeling spatial flood using novel ensemble artificial intelligence approaches in northern Iran. Remote Sensing, 12(20), 3423.
- [۳۰] Arabameri, A., Saha, S., Mukherjee, K., Blaschke, T., Chen, W., Ngo, P. T. T., & Band, S. S. (2020b). Flood susceptibility modeling utilizing advanced ensemble machine learning techniques in the Marand Plain, East Azarbaijan, Iran. Remote Sensing, 12(20), 3423.
- [۳۱] Bashirgonbad, M., Farokhzadeh, B., & Gholami, V. (2024). Enhancing flood mapping through ensemble machine learning in the Gamasyab watershed, Western Iran. Environmental Science and Pollution Research, 31(38), 50427–50442.
- [۳۲] Khosravi, K., Nohani, E., Maroufinia, E., Pourghasemi, H.R. (2016). "A GIS-based flood susceptibility assessment and its mapping in Iran: a comparison between frequency ratio and weights-of-evidence bivariate statistical models with multi-criteria decision-making technique." Vol.Natural hazards, No.83, PP.947–987.
- [۳۳] Rajabizadeh, Y., Ayyoubzadeh, S. A., Zahir, A. (2019). "Flood Survey of Golestan Province in 2018-2019 and Providing Solutions for Its Control and Management in the Future." Vol.Iranian journal of Ecohydrology No.6(4), PP.921-942.
- [۳۴] Khalifeh Soltanian, F., Abbasi, M. and Riyahi Bakhtyari, H.R. (2019). "Flood monitoring using ndwi and mndwi spectral indices: A case study of aghqala flood-2019, Golestan Province, Iran." Vol.The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, No.42, PP.605-607.
- [۳۵] Safaripour, M., Monavari, M., Zare, M., Abedi, Z., Gharagozlou, A. (2012). "Flood risk assessment using GIS (case study: Golestan province, Iran)." Vol.Pol.J. Environ. Stud, No.21, PP.1817–1824.
- [۳۶] Omidvar, B., Khodaei, H. (2008). "Using value engineering to optimize flood forecasting and flood warning systems: Golestan and Golabdare watersheds in Iran as case studies." Vol.Natural hazards, No.47, PP.281–296.
- [۳۷] Abdollahipour, A., Ahmadi, H., Aminnejad, B. (2021). "Evaluating the reconstruction method of satellite-based monthly precipitation over Golestan province, Northern Iran." Vol.Acta Geophys. No.69(6), PP.2305–2323.
- [۳۸] Tourani, M., Çağlayan, A. (2021). "Example of the biggest flood disaster in Iranian history: Golestan province (NE Iran)." Geoscience for Society, Education and Environment.
- [۳۹] Safariana, A., & Cheshomi, A. (2020). "The impact of climate change on micro, small, and medium enterprises, and the factors influencing the intention to flood risk reduction and climate change adaptation: A case study in Agh ghala city and Agh ghala industrial park, Golestan province, Iran."
- [۴۰] Singha, M., Dong, J., Sarmah, S., You, N., Zhou, Y., Zhang, G., Doughty, R., Xiao, X. (2020). "Identifying floods and flood-affected paddy rice fields in Bangladesh based on Sentinel-1 imagery and Google Earth Engine. ISPRS J. Photogramm." Vol.Remote Sensing, No.166, PP.278–293.
- [۴۱] Mousavi, M., Amini, J. and Maghsudi, Y. (2015). "Proposal speckle reduction algorithm for SAR images." Vol.Journal of Geomatics Science and Technology, No.5(1), PP.189-202.
- [۴۲] Dodangeh, P., Ebadi, H., Kiani, A. (2021). "Deep Learning Network for Flood Extent Mapping based on the integration of Sentinel 2 and MODIS satellite imagery." Vol.Journal of Environmental Studies, No.47(2), PP.181-204
- [۴۳] Yommy, A.S., Liu, R. and Wu, S. (2015). "August. SAR image despeckling using refined Lee filter." Proc. of the 7th International Conference on Intelligent Human-Machine Systems and Cybernetics, Hangzhou, China, pp. 260-265.

- [۴۴] Esmailzadeh, M., Amini, J. (2016). "Geometric Calibration of SAR Images to Eliminate Earth's Surface Topography Distortions." Vol. Journal of Geomatics Science and Technology. No.5(4), PP.173-185.
- [۴۵] Du, Y., Zhang, Y., Ling, F., Wang, Q., Li, W., Li, X. (2016). "Water bodies' mapping from Sentinel-2 imagery with Modified Normalized Difference Water Index at 10-m spatial resolution produced by sharpening the swir band." Vol. Remote Sensing, No.8(4), P.354.
- [۴۶] Shafizadeh-Moghadam, H., Valavi, R., Shahabi, H., Chapi, K., Shirzadi, A. (2018). "Novel forecasting approaches using combination of machine learning and statistical models for flood susceptibility mapping." Vol. Journal of environmental management, No.217, PP.1-11.
- [۴۷] Alkhurajji, W. (2020). "GIS and DEM Based Watershed Characteristics of Wadi Al-Adaira Basin, KSA. Bull." Vol. la Société Géographique d'Egypte, No.93(1), PP.131-158.
- [۴۸] Naimi, B., Hamm, N.A., Groen, T.A., Skidmore, A.K. and Toxopeus, A.G. (2014). "Where is positional uncertainty a problem for species distribution modelling?." Vol. Ecology, No.37(2), PP.191-203.
- [۴۹] Ghane-Ameleh, S., Khosravi, M., Saberi-Pirooz, R., Ebrahimi, E., Aghbolaghi, M.A., Ahmadzadeh, F. (2021). "Mid-Pleistocene Transition as a trigger for diversification in the Irano-Anatolian region: Evidence revealed by phylogeography and distribution pattern of the eastern three-lined lizard." Vol. Ecology and Conservation, No.31, P.e01839.
- [۵۰] Ebrahimi, E., Ranjbaran, Y., Sayahnia, R., Ahmadzadeh, F. (2022). "Assessing the climate change effects on the distribution pattern of the Azerbaijan Mountain Newt (*Neurergus crocatus*)." Vol. Ecological Complexity, No.50, P.100997.
- [۵۱] Naimi, B. (2015). usdm: Uncertainty analysis for species distribution models (Version 1.1-12).
- [۵۲] Ahmadzadeh, F., Ebrahimi, E., & Naimi, B. (2017). "Species distribution potential of striped hyaena (*Hyaena hyaena*) in response to climate change in Iran." Vol. *Environmental Sciences*, No.15(4), PP.215-232.
- [۵۳] Arabameri, A., Saha, S., Mukherjee, K., Blaschke, T., Chen, W., Ngo, P.T.T., Band, S.S. (2020). "Modeling spatial flood using novel ensemble artificial intelligence approaches in northern Iran." Vol. Remote Sensing, No.12(20), PP.1-30.
- [۵۴] Park, N.W. (2015). "Using maximum entropy modeling for landslide susceptibility mapping with multiple geoenvironmental data sets." Vol. Environmental Earth Sciences, No.73, PP.937-949.
- [۵۵] Tehrany, M.S., Pradhan, B., Mansor, S., Ahmad, N. (2015b). "Flood susceptibility assessment using GIS-based support vector machine model with different kernel types." Vol. Catena, No.125, PP.91-101.
- [۵۶] Naimi, B., Capinha, C., Ribeiro, J., Rahbek, C., Strubbe, D., Reino, L., Araújo, M.B. (2022). "Potential for invasion of traded birds under climate and land-cover change." Vol. Global Change Biology, No.28(19), PP.5654-5666.
- [۵۷] Arabameri, A., Rezaei, K., Cerdà, A., Conoscenti, C., Kalantari, Z. (2019). "A comparison of statistical methods and multi-criteria decision making to map flood hazard susceptibility in Northern Iran." Vol. Science of the Total Environment, No.660, PP.443-458.
- [۵۸] Avand, M., Janizadeh, S., Naghibi, S.A., Pourghasemi, H.R., Khosrobeigi Bozchaloei, S. and Blaschke, T., 2019. A comparative assessment of random forest and k-nearest neighbor classifiers for gully erosion susceptibility mapping. Water, 11(10), p.2076.
- [۵۹] Taalab, K., Cheng, T., Zhang, Y. (2018). "Mapping landslide susceptibility and types using Random Forest." Vol. Big Earth Data, No.2(2), PP.159-178.
- [۶۰] Breiman, L. (2001). "Random forests." Vol. Machine learning, No.45, PP.5-32.
- [۶۱] Ghayoumi, R., Ebrahimi, E., Hosseini Tayefeh, F., & Keshtkar, M. (2019). "Predicting the effects of climate change on the distribution of mangrove forests in Iran using the maximum entropy model." Journal of RS and GIS for Natural Resources, No. 10(2), PP. 34-47.
- [۶۲] Rahmati, O., Karnejadi, A., Choobin, J., ... & Ata. (2023). "Evaluation of the Capability of the Flexible Discriminant Analysis Model in Predicting Flood Susceptibility in the Zarrineh River Watershed." Water and Soil Modeling and Management.
- [۶۳] Ebrahimi, E., Araújo, M. B., & Naimi, B. (2023). "Flood susceptibility mapping to improve models of species distributions." Vol. Ecological Indicators, No. 157, PP. 111-250.

- [۶۴] Shahiri Tabarestani, E., Hadian, S., Pham, Q.B., Ali, S.A., Phung, D.T. (2022). "Flood potential mapping by integrating the bivariate statistics, multi-criteria decision-making, and machine learning techniques." Vol. Stochastic Environmental Research and Risk Assessment, No.37, PP.1415–1430.
- [۶۵] Alavi, M.S., Fallahi, A., Mottaki, Z. and Aslani, F. (2021). "Post-disaster sheltering process after the 2019 floods, in Golestan province, Iran." Vol. International journal of disaster resilience in the built environment, No.13(5), PP.568-582.