Archive of SID.ir

پژوهش های آبخیزداری
شاپا: ۲۰۳۸-۲۹۸۱
سازمان تحقیقات، آموزش و ترویج کشاورزی
مرکز تحقیقات و آموزش کشاورزی و منابع طبیعی فارس

پیش‌بینی مکانی مناطق مستعد سیل در آبخیز کرخه‌ی استان لرستان با استفاده از مدل ترکیبی جنگل تصادفی– وزن شواهد

اسلام ستارزاده[۱]، امیرپویا صراف[۲*]، هومن حاجی‌کندی[۳]، محمدصادق صادقیان[۴]

۱ – دانشجوی دکتری مهندسی عمران، گروه مهندسی عمران، واحد تهران مرکزی، دانشگاه آزاد اسلامی، تهران، ایران
۲ – استادیار گروه مهندسی عمران، واحد رودهن، دانشگاه آزاد اسلامی، رودهن، ایران
۳ و ۴ – استادیار گروه مهندسی عمران، واحد تهران مرکزی، دانشگاه آزاد اسلامی، تهران، ایران

چکیده‌ی مبسوط

مقدمه و هدف

پدیده‌ی سیل یکی از پدیده‌های پرتکرار دهه‌ی گذشته در ایران است که خسارت‌های مالی و تلفات جانی بسیاری را سبب شده است. یکی از سه مخاطرهٔ طبیعی و اصلی ایران، رخداد سیل است و بدون شک حداقل در سال در یک نقطه از این کشور سیلاب بزرگی رخ می دهد. سیل به‌عنوان یک تهدید بزرگ برای زندگی بشر (با آسیب‌زدن یا مرگ انسان و حیوان‌ها) و به طور خاص برای ساختمان و خانه، زمین کشاورزی و تولید محصول، زیرساخت‌های شهری، پل‌ها و جاده‌ها به شمار می‌آید. سیل در ایران خسارت‌های بسیاری از نظر اقتصادی، نابودی محیط زیست، منابع‌طبیعی و مسکونی و تلفات جانی وارد می‌نماید. در سال‌های گذشته حدود ۷۰٪ اعتبارهای سالانه‌ی طرح کاهش اثر بلاهای طبیعی و ستاد حوادث غیرمترقبه صرف جبران خسارت‌های ناشی از سیل شده است. در این پژوهش، با استفاده از داده‌های موقعیت مکانی سیل و به‌کارگیری مدل‌های یادگیری ماشینی و داده‌کاوی جنگل تصادفی و وزن و ترکیب آن‌ها به پیش‌بینی مکانی مناطق مستعد سیل پرداخته می‌شود.

مواد و روش‌ها

در این پژوهش از مدل‌های ترکیبی و ۱۱ متغیر پیش‌بینی کننده‌ی احتمال سیلاب در آبخیز کرخه واقع در استان لرستان استفاده شده است. این متغیرها شامل نقشه‌ی شاخص‌های سنجش شکل زمین از جمله شاخص رطوبت

نوع مقاله: پژوهشی
***مسئول مکاتبات، پست الکترونیکی:** sarraf@riau.ac.ir
استناد: ستارزاده، ا.، صراف، ا.پ.، حاجی‌کندی، ه.، صادقیان، م.ص. ۱۴۰۲. پیش‌بینی مکانی مناطق مستعد سیل در آبخیز کرخه‌ی استان لرستان با استفاده از مدل ترکیبی جنگل تصادفی– وزن شواهد. پژوهش‌های آبخیزداری، ۳۶ (۲): ۱۰۳–۸۷.
شناسه‌ی دیجیتال: 10.22092/wmrj.2022.359264.1483
تاریخ دریافت: ۱۴۰۱/۰۴/۱۹، **تاریخ بازنگری:** ۱۴۰۱/۰۶/۰۳، **تاریخ پذیرش:** ۱۴۰۱/۰۶/۳۰، **تاریخ انتشار:** ۱۴۰۲/۰۴/۰۱
پژوهش‌های آبخیزداری، سال ۱۴۰۲، دوره‌ی۳۶، شماره‌ی ۲، شماره‌ی پیاپی ۱۳۹، تابستان ۱۴۰۲، صفحه‌های۸۷ تا ۱۰۳
ناشر: مرکز تحقیقات و آموزش کشاورزی و منابع طبیعی استان فارس
©نویسندگان
CC BY

Archive of SID.ir

Archive of SID.ir

پستی بلندی، موقعیت شیب نسبی و شاخص موقعیت پستی بلندی، نقشه‌های آب‌شناختی شامل: تراکم زهکشی و فاصله از شبکه‌ی زهکشی است. به‌این منظور ابتدا مدل‌های داده‌کاوی برای تجزیه و تحلیل اولیه‌ی رابطه‌ی بین متغیرهای محیطی و رخداد سیل‌های گذشته استفاده شد و نتایج آن‌ها به عنوان اطلاعات ورودی مدل‌های یادگیری ماشینی استفاده شد. داده‌های سیل به شکل تصادفی به دو گروه آموزش ۷۰٪ و اعتبارسنجی ۳۰٪ تقسیم شدند. دقت پیش‌بینی با استفاده از روش منحنی مشخصه‌ی عملکرد (ROC) بررسی شد.

نتایج و بحث

براساس نتایج، دقت مدل جنگل تصادفی ۰/۹۰۴، دقت مدل وزن شواهد ۰/۸۸۶ و دقت مدل ترکیبی جنگل تصادفی وزن شواهد ۰/۹۷۸ بود. براساس مدل ترکیبی جنگل تصادفی و وزن شواهد به‌عنوان مدل برتر، ۲۰/۴۹٪ سطح ظرفیتی بیش از اندازه‌ی متوسط داشت. براساس مدل جنگل تصادفی، عوامل تراکم زهکشی، فاصله از آبراهه، بلندی و کاربری زمین مهم‌ترین عوامل مؤثر بر ظرفیت سیل بودند.

واژگان کلیدی: حساسیت سیل، سیل، عوامل محیطی، مدل‌سازی

مقدمه

در سال‌های کنونی یکی از بلاهای طبیعی که سبب خسارت‌های بسیاری شده سیل است. سیل به‌عنوان یک تهدید بزرگ برای زندگی بشر (با آسیب زدن یا مرگ انسان و حیوان‌ها)، و به‌طور خاص برای ساختمان و خانه، زمین کشاورزی، زیرساخت‌های راه‌آهن، زیرساخت‌های شهری، پل‌ها و جاده‌ها به شمار می‌آید (سامپسون و همکاران ۲۰۱۵). در سال‌های گذشته حدود ۷۰٪ اعتبارهای سالانه‌ی طرح کاهش اثر بلاهای طبیعی و ستاد حوادث غیرمترقبه صرف جبران خسارت‌های ناشی از سیل شده است. نابودی پوشش گیاهی در عرصه‌های آبخیز، کاربری غیراصولی زمین، افزایش سطوح غیرقابل نفوذ و سایر دخالت‌های انسان در یک منطقه خطر بروز سیل را افزایش می‌دهد (صراف و همکاران ۲۰۲۱). به‌عبارت دیگر تشدید سیل ناشی از فعالیت‌های غیرمسئولانه‌ی انسان است (مقدم و همکاران ۲۰۱۹). فرآیند رخداد سیل بسیار پیچیده است و عوامل مختلفی در آن نقش دارند (صراف و همکاران ۲۰۱۸). از این رو پیش‌بینی آن نیز دشوار است. در سال‌های کنونی، در جامعه‌های بین‌المللی ارزیابی حساسیت به موضوع مهم پژوهشی تبدیل شده و استفاده از مدل‌های مختلف برای پیش‌بینی مکانی مناطق مستعد خطر به‌وسیله‌ی پژوهشگران مختلفی بررسی‌شده است. در پژوهشی در کره‌ی جنوبی لی و همکاران (۲۰۱۷) با استفاده از مدل‌های درخت وایازی (رگرسیونی) تقویت‌شده و جنگل تصادفی به پیش‌بینی مکانی نقشه‌ی حساسیت سیلاب پرداختند. پس از محاسبه‌ی اهمیت عوامل نقشه‌های حساسیت سیل، فاصله از رودخانه، سنگ‌شناسی و مدل رقومی - ارتفاعی از مهم‌ترین ویژگی‌ها شناسایی شدند. براساس الگوی‌های وایازی و طبقه‌بندی، ارزیابی مدل جنگل تصادفی، به‌ترتیب دقت ۷۸/۷۸ و ۷۹/۱۸٪ و مدل وایازی تقویت‌شده به ترتیب، دقت ۷۷/۵۵ و ۷۷/۲۶٪ را نشان دادند. در پژوهشی در منطقه‌ای از چین برای مدل‌سازی حساسیت‌پذیری سیل چن و همکاران (۲۰۲۰) به ارزیابی و مقایسه‌ی قابلیت پیش‌بینی مدل‌های درخت بایس ساده (NBTree)، درخت تصمیم‌گیری متناوب (ADTree) و مدل جنگل تصادفی(RF) پرداختند. نتایج براساس مدل‌سازی حساسیت‌پذیری سیل نشان داد که مدل جنگل تصادفی یک الگوی کارآمد و قابل اعتماد بوده و بیشترین اندازه‌های AUC را دارد. در استان گلستان رحمتی و پورقاسمی (۲۰۱۷) برای مدل‌سازی مکانی مناطق مستعد سیل، از سه مدل تابع شاهدهای قطعی (EBF)، جنگل تصادفی (RF) و درخت وایازی پیشرفته (BRT) استفاده کردند. نتایج نشان داد که هر سه مدل قابلیت زیادی در تجزیه و تحلیل رابطه‌ی میان عامل‌های محیطی و رخدادهای سیل داشتند. همچنین نتایج ارزیابی دقت مدل‌ها نشان داد که مدل تابع شواهد قطعی دقت زیادی نسبت به دو مدل دیگر داشت. تهیه‌ی نقشه‌ی مناطق مستعد سیل، برای پیشگیری و مدیریت سیل بسیار مهم است (صراف و همکاران ۲۰۲۲). با پیشرفت دانش فنی و امکانات سنجش از دور و سامانه‌ی اطلاعات جغرافیایی، برای آبخیز امکان تهیه ی بخش وسیعی از اطلاعات لازم برای مدل‌ها بدون بازدیدهای محلی فراهم‌شده است. پیش‌بینی مکانی خطرهای طبیعی با کاربرد مدل‌هایی که به‌وسیله‌ی داده‌های مکانی طراحی‌شده است، سرانجام منجر به تهیه‌ی نقشه‌های حساسیت‌پذیری به پدیده‌های طبیعی از جمله سیل می شود، که در مناطق مختلف برای برنامه‌ریزان از راهکارهای مناسب مدیریت

Archive of SID.ir

Archive of SID.ir

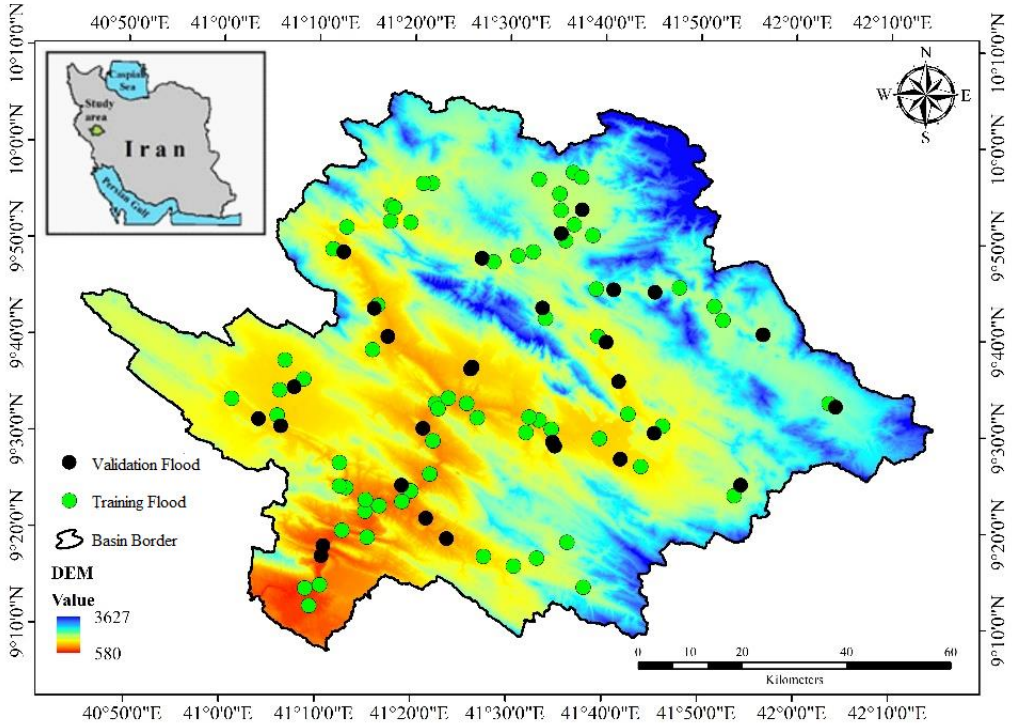
زمین‌ها برای پیشگیری از رخداد این پدیده‌ها است (صراف و همکاران ۲۰۱۹). پیش‌بینی سیل در آبخیزها و تعیین مناطق مستعد آب‌گرفتگی به دلیل کافی نبودن آمار و اطلاعات ایستگاه‌های آب‌سنجی دشوار است و این مسئله در کشورهای در حال توسعه به‌وفور مشاهده می‌شود. در آبخیز کرخه در استان لرستان رخداد سیل و اثرهای آن، یکی از مانع‌های اصلی برنامه‌های توسعه‌ای به‌شمار می‌آید. بنابراین برنامه‌ریزی‌های مهارکردن بلای سیل در این آبخیز به وسیله‌ی تهیه‌ی نقشه‌ی مناطق مستعد سیل و آب‌گرفتگی مفید و آسان‌تر شده است. مهارکردن بلای سیلاب در آبخیز کرخه‌ی استان لرستان یکی از مهم-ترین اهداف مدیریت این آبخیز بوده که با توجه به پیشینه‌ی رخداد سیلاب‌های ویران‌گر بسیار ضروری می‌باشد. برنامه‌ریزی و انجام اقدام‌های جامع برای پیشگیری و کاهش خسارت‌های سیل و توسعه‌ی مناسب و پایدار مدیریت آبخیز در قالب طرح‌های پژوهشی و اجرایی اهمیت به سزایی دارد (صراف و همکاران ۲۰۲۱). لازمه‌ی اقدام‌ها و برنامه‌ریزی‌ها برای اجرای طرح‌های مهارکردن و کاهش خسارت‌های سیل تهیه‌ی نقشه‌های ظرفیت و استعدادیابی تولید سیل است. از جمله روش‌های تهیه‌ی این نقشه‌ها که تعیین‌کننده‌ی مناطق سیل‌گیر است، به‌کارگیری مدل‌های داده‌کاوی و یادگیری ماشینی است. تهیه‌ی این نقشه‌ها در تعیین و ارزیابی نقاط حادثه‌خیز لازم است و ضرورت اجرای آن با توجه به خسارت‌ها و مشکلات پرشماری که به‌وسیله‌ی سیل ایجاد می‌شود، بسیار اهمیت دارد و پیشگیری از چنین رویدادهایی بهتر از جبران خسارت است. در بحث سیل در نقاط مختلف ایران بررسی‌های گوناگونی با استفاده از مدل‌های مختلف انجام‌شده است، اما ترکیب مدل‌های یادگیری ماشینی و داده‌کاوی تاکنون برای مدل‌سازی مناطق مستعد سیل به شکل محدود استفاده شده است. در این پژوهش کارایی مدل‌های ترکیبی براساس رخدادهای گذشته مقایسه می‌شود. علاوه بر این، تعیین اندازه‌ی مشارکت عامل‌های محیطی در پدیده‌ی سیل و تهیه‌ی نقشه‌ی پهنه‌بندی سیل بر اساس مدل‌های ترکیبی از دیگر جنبه‌های نوآوری این پژوهش است. مدل‌های ترکیبی می‌توانند مناطق مختلف یک آبخیز را از نظر استعداد سیل‌خیزی رتبه‌بندی کنند که این امر در بحث مدیریت سیلاب و اجرای اقدامات آبخیزداری اهمیت بسزایی دارد.

مواد و روش‌ها

آبخیز کرخه در استان لرستان، با مختصات ′۵۹ °۳۴ ′۰۱ °۳۳ و ′۰۳ °۴۹ ′۱۵ °۴۷ در غرب ایران واقع شده است. مساحت منطقه‌ی پژوهش ۹۴۴ هکتار بوده و از نظر بلندی، در محدوده‌ی ۵۸۰ تا ۳۶۲۷ قرار دارد. میانگین اندازه‌ی بارش سالانه ۹۶۵ میلی‌متر و دمای منطقه بین ۹- تا ۳۳/۳°C است. بخش بزرگی از کاربری زمین‌های منطقه، کشاورزی دیم است که ۳۵/۳۹٪ از منطقه را تشکیل می‌دهد. از نظر زمین‌شناسی، بخش بسیاری از منطقه، مربوط به سازند کواترنر (دوره‌ی چهارم) است. اینسپتی‌سول‌ها با منشاء نسبتاً جدید و ظاهر ضعیف افق‌ها به دلیل تجمع کم رس و مواد آلی، فراوان‌ترین نوع خاک در منطقه می‌باشند. طول شبکه‌ی زهکشی منطقه‌ی بررسی‌شده ۳۲۹۶/۹ کیلومتر می‌باشد. منطقه‌ی مورد مطالعه از نظر اقلیمی، آب و هوای سرد و خشک دارد. چهار شهر اصلی الشتر، کوهدشت، خرم‌آباد و پل‌دختر در چهارگوشه‌ی محدوده قرار دارد. محدوده‌ی بررسی‌شده و نقاط اعتبارسنجی و آموزش سیل در شکل شماره ۱ نشان داده شده است.

Archive of SID.ir

Archive of SID.ir



شکل ۱- منطقه‌ی بررسی‌شده در پژوهش.

Figure 1- The Scope of the Study area in the Current Research

در آبخیز کرخه واقع در استان لرستان از مدل‌های منفرد و ترکیبی و ۱۲ عامل مؤثر بر احتمال سیلاب شامل بلندی، شیب، شاخص ناهمواری‌های زمین (TRI)[1]، شاخص قدرت جریان (SPI)[2]، شاخص حمل و نقل رسوب (STI)[3]،شاخص‌رطوبت‌پستی‌بلندی(TWI)[4]،موقعیت شیب نسبی (RSP)[5] و شاخص موقعیت پستی بلندی (TPI)[6]، تراکم زهکشی و فاصله از شبکه‌ی زهکشی، نوع خاک و کاربری زمین استفاده شده است. مرحله‌ی مهم برای تهیه‌ی نقشه‌ی حساسیت‌پذیری، نقشه‌ی پراکنش نقاط رخداد سیل است (صراف و همکاران ۲۰۲۲). براساس گزارش‌ها و داده‌های به دست‌آمده از اداره‌ی کل منابع‌طبیعی و آبخیزداری استان لرستان، ۹۶ نقطه‌ی سیلاب شناسایی شد. به اندازه ی تعداد نقاط رخداد سیل، نقاط بدون رخداد سیل نیز انتخاب شدند. تعادل حضور و بدون حضور (۱:۱) می‌تواند منجر به پراکنش کمتری در نتایج شود (صراف و همکاران ۲۰۲۲). با استفاده از الگوی تقسیم‌بندی تصادفی نقاط رخداد سیل به کار برده‌شده برای آموزش مدل و اعتبارسنجی به نسبت ۷۰ به ۳۰ تقسیم شدند. روندنمای مراحل جریانی پژوهش در شکل ۲ نشان داده شده است.

عامل پستی بلندی نقش مهمی در مهارکردن سیلاب‌ها بر اساس شرایط آب‌شناختی و رطوبت خاک دارد. سیلاب با بلندی، رابطه‌ی عکس دارد و مناطق با بلندی کم، بیشتر در معرض رخداد سیلاب قرار دارند (صراف و همکاران ۲۰۲۰). این لایه با تفکیک مکانی ۳۰×۳۰ متر در محیط ArcGIS تهیه شده است. شیب به عنوان یک شاخص مؤثر سطحی برای تعیین مناطق مستعد سیل، اندازه‌ی رواناب سطحی و فرصت نفوذ در نظر گرفته شده است. همچنین این عامل به عنوان مهم‌ترین عامل در تعیین سرعت جریان بر روی تراکم زهکشی، رطوبت خاک و فرایندهای ریخت‌شناسی زمین تأثیرگذار است (دارابی و همکاران ۲۰۲۲). شاخص موقعیت پستی‌بلندی الگویی برای اندازه‌گیری موقعیت‌های شیب پستی‌بلندی و به‌طور کلی برای طبقه‌بندی شکل زمین استفاده می‌شود و به شکل تفاوت بین پیکسل مرکزی و میانگین سلول‌های اطراف آن تعریف می‌شود. شاخص ناهمواری‌های زمین یکی از عامل‌های مهم تأثیرگذار در حوادث سیل است که به پستی‌بلندی

1 - Terrain Roughness Index
2 - Stream Power Index
3 - Sediment Transport Index
4 - Topographuc Wetness Index
5 - Relative Slope Position
6 - Topographic Position Index

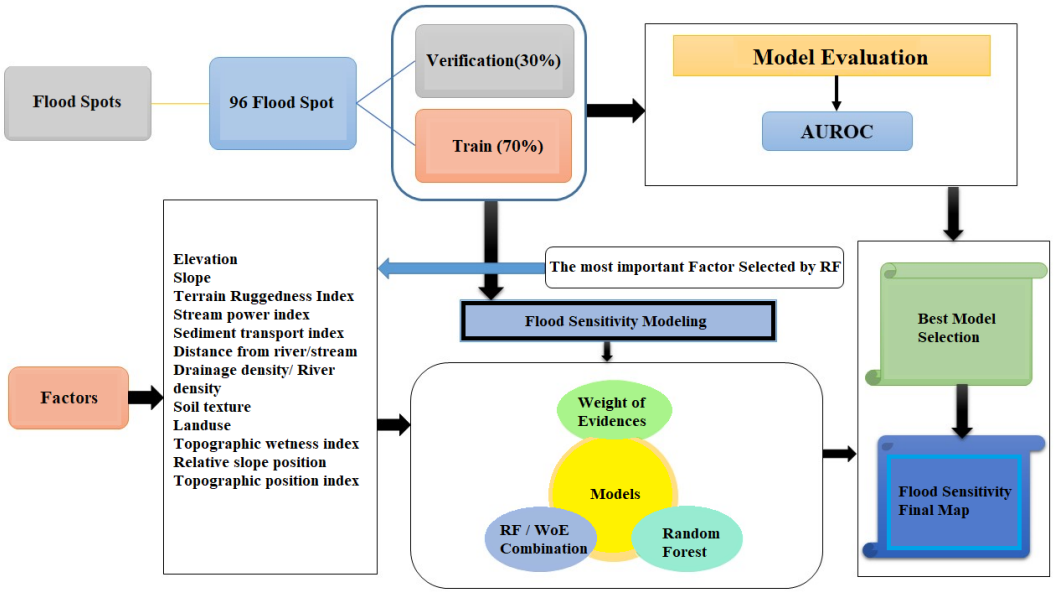
Archive of SID.ir

Archive of SID.ir

محلی آبخیز پژوهش شـــده بستگی داشت و برای بیان تفاوت بلندی بین یک سلول و میانگین ماتریس هشت سلولی سلول‌های اطراف ایجاد شد. ازاین‌رو هنگامی که احتمال رخداد سیلاب بیشتر باشد، اندازه‌ی TRI کمتر اســـت (صراف و همکاران ۲۰۱۷). این شاخص از روی نقشه‌ی مدل رقومی ارتفاعی در نرم‌افزار SAGA تهیه شد. شـــاخص رطوبت پستی‌بلندی برای کمی‌کردن اثر پستی‌بلندی بر روی سطح اشباع‌شده‌ی جریان استفاده می‌شود که یکی از عامل‌های ضروری در ارزیابی ظرفیت سیلاب برای نشان دادن توزیع مکانی و اختلاف شرایط رطوبت در آبخیزها است (دهقانیان و همکاران ۲۰۲۰). این شاخص، اندازه‌ی آب موجود در هر اندازه‌ی پیکسل منطقه را نشان می‌دهد (آرنال و همکاران ۲۰۱۶). یکی دیگر ازشـــاخص‌های مهم در ظرفیت سیلاب‌هایی که در آبخیزها رخ می‌دهد، شـــاخص قـــدرت جریان بوده که تأثیر اساســـی بر روی ســـامانه‌ی رودخانه دارد. این شاخص بیانگر قدرت فرسایندگی جریان بوده و بر روی ناپایداری دامنه‌ها اثرگذار اســـت (پویـــدال و همکاران ۲۰۱۰). این عامل نیز براســـاس نقشـــه‌ی DEM در محیط نرم‌افزار SAGA تهیه شد. شاخص حمل و نقل رسوب به‌عنوان عامل دیگری برای ایجاد سیل، می‌تواند فرکانس سیل را افزایش دهد که منجر به آسیب ساخت و سازها می‌شـــود. این شـــاخص از روی نقشه‌ی مدل رقومی ارتفاعی تهیه شده است. همچنین STI براساس ویژگی‌های آب و هوایی و زمین ریخت‌شناسی منطقه، به‌دست می‌آید. بیشتر مناطق تحت تأثیر طغیان سیل معمولاً در مجاورت رودخانه قرار دارند. فاصله از رودخانه یک عامل اساسی برای شناسایی مناطق آسیب‌پذیر در برابر طغیان سیلاب است (دهقانیان و همکاران ۲۰۲۰). همچنیـــن فاصله از رودخانه، جریان آن را مهار می‌کند (رحمتی و همکاران ۲۰۱۵). این عامل یکی از مهم‌ترین لایه‌هایی است که بر روی حجم، سرعت و پخش سیلاب تأثیر بسیاری دارد. هرچه فاصله از رودخانه بیشتر باشد، احتمال رخداد سیل کمتر خواهد بود. در مقیاس آبخیز، ســـیل منطقه مربوط به ذخیره‌ی آب زیرزمینی اســـت (دنگ و همکاران ۲۰۲۰). این عامل با اســـتفاده از تابع فاصله‌ی اقلیدسی در نرم‌افزار ArcGIS تهیه می‌شود. این عامل با اســـتفاده از دستور «Line density» در محیط نرم‌افزار ArcGIS به‌دست می‌آید و این کمیت قابل‌اندازه‌گیری از روی نقشه‌ی شبکه‌ی آبراهه به‌دست آمد. عموماً انـــدازه‌ی کمتر این عامل نشـــان‌دهنده‌ی نفوذپذیـــری زیـــاد، تراکم پوشـــش گیاهی فـــراوان و هموار‌بودن منطقه اســـت. نوع خاک یکی از عامل‌های مهم تأثیرگذار و مهارکننده‌ی ســـاختار بارندگی رواناب است (پُرهمت ۲۰۱۶). خصوصیت‌های خاک مستقیماً بـــر روی نفـــوذ آب و تولید رواناب تأثیر گذار اســـت و افزون بر آن، ســـایر عامل‌ها مانند شـــرایط آب و هوای محلی و فرایندهای فرسایشـــی سبب مهارکردن رواناب می‌شـــود (کاظمی و همـــکاران ۲۰۲۱). کاربری زمین جریان‌های سطحی را مهار می‌کند، ازاین‌رو نقش مهمی در فرایندهای آب‌شـــناختی دارد. ایـــن عامل بر رواناب ســـطحی و حمل و نقل رسوب و در نتیجه، مستقیماً بر فراوانی سیل اثرگذار است. زیرا پوشش زمین مستقیماً تولید رواناب ســـطحی و نفوذ را مهار می‌کند. در نتیجه در مناطقی که پوشـــش گیاهی وضعیت مناسبی دارد ســـیلاب کمتری رخ می‌دهد. کاربـــری زمین‌های این منطقه شـــامل: زمین‌های زراعی خشک، باغ، دریاچه، مناطق مســـکونی، جنگل، مرتع و زمین‌های کشاورزی است.

Archive of SID.ir

Archive of SID.ir



شکل ۲- نمودار مراحل جریانی پژوهش.

Figure 2- Flow Diagram of Current Research.

مدل جنگل تصادفی (RF)

مدل جنگل تصادفی یکی از روش‌های مدل‌سازی درخت تصمیم است و روشی است که از ترکیب نتایج به دست‌آمده از درخت‌های مختلف، نتیجه یا برآورد نهایی رابطه‌ی بین خطر طبیعی و عامل‌های محیطی را پیشنهاد می‌دهد (صراف و همکاران ۲۰۲۲). در مدل جنگل تصادفی به جای رشد یک درخت، صدها یا هزاران درخت طبقه‌بندی تولید می‌شود (صراف و همکاران ۲۰۲۲). نمونه‌گیری تصادفی از متغیرها، به‌شکل تکرارشونده و خودکار برای تولید بسیاری از درختان وایازی انجام می‌شود. برای تعیین تعداد بهینه‌ی درختان، یک تعداد درخت اولیه برای تولید نمودار تغییرهای میانگین مربعات خطا (MSE) در مقابل تعداد خاص درختان نمونه‌ی آموزشی و اعتبارسنجی استفاده می‌شود. این روش یک ابزار تحلیلی بسیار توانمند برای جستجوی اطلاعات و اصلاح تعداد بهینه‌ی درختان در مدل است. یکی از مهم‌ترین سنجه‌های اجرای مدل جنگل تصادفی، پیش‌بینی‌کننده‌ی K (متغیر مستقل) در هر گره برای پیش‌بینی اندازه‌ی وابسته (پاسخ) است. محاسبه‌ی جذر تعداد کل متغیرهای مستقل، آسان‌ترین راه برای تعیین اندازه‌ی K است (شفاپورتهرانی و همکاران ۲۰۱۹). به منظور استفاده از این روش پیشرفته برای مدل‌سازی حساسیت‌پذیری سیلاب از نرم‌افزار R و پکیج Random Forest استفاده شده است.

مدل وزن شواهد (WOE)

این روش بر پایه‌ی انحراف معیار و ترکیب اطلاعات تعریف‌شده و در شرایطی که تعداد رخدادهای سیل کم نباشد، ابتدا اهمیت نسبی متغیرهای محیطی را مشخص کرده و بر پایه‌ی آن پیش‌بینی رخداد سیل را انجام می‌دهد (کورسین و همکاران ۲۰۰۹). یکی از روش‌های دومتغیره، مدل احتمالاتی وزن شواهد است که با استفاده از شکل ورود به لگاریتم خطی و از روش احتمالاتی بیزین، اهمیت نسبی بین سنجه‌های اثرگذار را تعیین می‌کند. این مدل براساس تعیین وزن مثبت W+ و وزن منفی W- است. در این پژوهش مدل وزن شواهد براساس معادله‌های ۱ و ۲ بررسی وزن طبقه‌های هریک از متغیرهای محیطی را انجام می‌دهد.

$$W_i^+ = \ln \frac{P\{B/A\}}{P\{B/\overline{A}\}} \quad (۱)$$

$$W_i^- = \ln \frac{P\{\overline{B}/A\}}{P\{\overline{B}/\overline{A}\}} \quad (۲)$$

B و به‌ترتیب مشارکت و مشارکت نداشتن طبقه مزبور در رخداد سیل است. همچنین، سنجه‌های A و به ترتیب رخدادن و رخندادن سیل در طبقه مزبور است. وزن مثبت، نشان دهنده‌ی بودن سنجه‌ی اثرگذار در نقاط سیل‌خیز است و اندازه‌ی آن نشانه‌ای از همبستگی مثبت بین حضور متغیر مستقل و سیل است. وزن منفی نشان‌دهنده‌ی نبودن متغیر اثرگذار است و اندازه‌ی آن همبستگی منفی را نشان می‌دهد. سنجه‌ی C اختلاف بین وزن مثبت و منفی است و نشان‌دهنده‌ی بزرگی

Archive of SID.ir

Archive of SID.ir

رابطه‌ی مکانی عامل علت و معلول است. در صورتی که رابطه‌ی طبقه سنجه‌ی مزبور (شیب زمین، کاربری زمین، سنگ‌شناسی و غیره) با رخداد سیل مثبت باشد اندازه‌ی این ضریب نیز مثبت است و در حالاتی که رابطه‌ی همبستگی طبقه عامل مزبور با رخداد سیل منفی باشد، اندازه‌ی آن منفی است. بر اساس پراکنش وزن نهایی، انحراف معیار از رابطه‌ی ۳ به‌دست می‌آید.

$$S(C)=\sqrt{S^{2}(W^{+})+S^{2}(W^{-})} \quad (۳)$$

و پراکنش وزن مثبت و منفی است. مؤلفه‌های و به‌ترتیب پراکنش وزن‌های مثبت و منفی است که به‌شکل معادله‌های ۴ و ۵ به‌دست می‌آید.

$$S^{2}(W)^{+}=\frac{1}{N\{B\cap A\}}+\frac{1}{B\cap A} \quad (۴)$$

$$S^{2}(W)^{-}=\frac{1}{N\{B\cap A\}}+\frac{1}{B\cap A} \quad (۵)$$

وزن نهایی (Ψ) طبقه‌های هریک از عامل‌های محیطی و نحوه‌ی رابطه‌ی آن‌ها با رخدادهای سیل گذشته از رابطه‌ی ۶ به‌دست می‌آید.

$$\psi=\left(\frac{C}{S(C)}\right) \quad (۶)$$

در شرایطی که اندازه‌ی وزن نهایی به‌دست‌آمده مثبت باشد $(\Psi<۰)$، طبقه مزبور رابطه‌ی مستقیم و مؤثر با رخداد سیل دارد. این در حالی است که اندازه‌های منفی وزن نهایی $(\Psi>۰)$ نشان‌دهنده‌ی رابطه‌ی معکوس طبقه مزبور با رخدادهای سیل است. این وزن به‌عنوان شاخص استعداد رخداد سیل در نظر گرفته شده است.

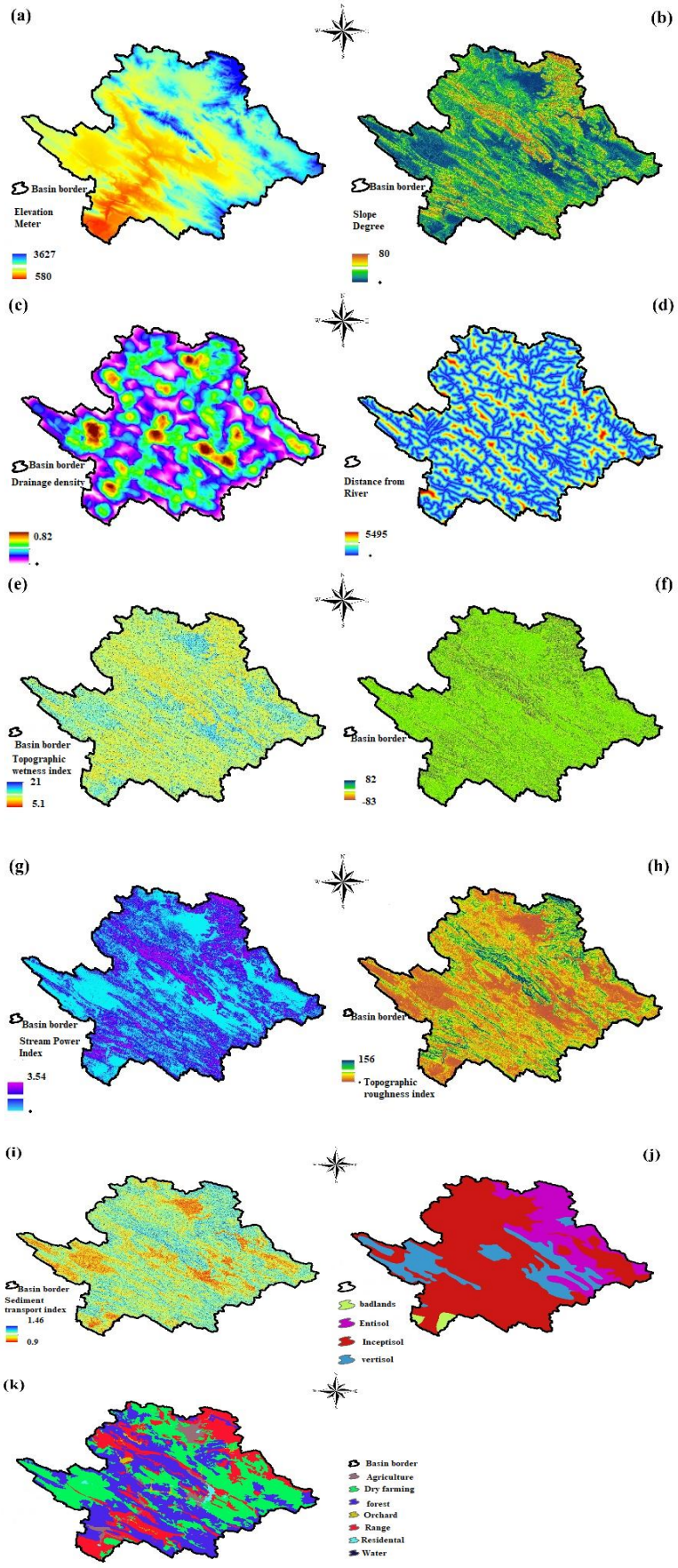
نتایج حاصل از مدل‌سازی احتمال رخداد سیلاب، براساس نقاط رخداد (۳۰٪ رخدادهای سیل گذشته) و در قالب روش منحنی مشخصه‌ی عامل گیرنده (ROC) اعتبارسنجی شد و مساحت زیر این منحنی (AUC) به‌عنوان معیار کمی برای اعتبارسنجی استفاده شده است (فیلیسیمو و همکاران ۲۰۱۳). به‌دلیل در نظر گرفتن همزمان دو معیار حساسیت و تشخیص و با توجه با پیشینه‌ی پژوهش در این زمینه، روش منحنی مشخصه‌ی عامل گیرنده به‌عنوان جامع‌ترین روش ارزیابی دقت مدل در نظر گرفته شد.

نتایج

در این پژوهش عامل‌های استفاده‌شده در شکل ۳ نشان‌داده شده است. براساس این شکل بلندی منطقه از ۵۸۰ متر تا ۳۶۲۷ متر متغیر بود. بیشترین بلندی منطقه در بخش‌های شمال‌شرقی و میانی و کمترین بلندی در جنوب غربی قرار گرفت. درجه‌ی شیب زمین در این منطقه از صفر تا ۸۰/۳ درجه متغیر بود. بر اساس تحلیل‌های شیب زمین، بیشترین درجه‌ی شیب مربوط به بخش‌های میانی و شمال‌شرقی منطقه‌ی پژوهش بود. براساس نقشه‌ی شاخص موقعیت پستی‌بلندی منطقه‌ی پژوهش، اندازه‌ی این شاخص از ۸۳/۲- تا ۸۲/۶ متغیر بود. بخش‌هایی که اندازه‌ی شاخص موقعیت پستی‌بلندی بسیار داشت در بخش‌های میانی منطقه قرار گرفت. بیشترین و کمترین اندازه‌ی شاخص ناهمواری به ترتیب ۱۵۶ و صفر به دست آمد. بیشترین ضریب ناهمواری مربوط به بخش‌های میانی و شمال‌شرقی آبخیز بود. اندازه‌ی شاخص رطوبت پستی بلندی منطقه از ۳/۱ تا ۲۱/۱ متغیر بود. بیشترین اندازه‌ی این شاخص مربوط به بخش‌هایی از منطقه با مساحت بالادست بیشتر و شیب زمین کمتر بود. براساس نقشه‌ی شاخص توان جریان، اندازه‌ی این شاخص در منطقه‌ی پژوهش از ۰/۰۱ تا ۳/۵۴ متغیر بود و بیشترین اندازهی این شاخص در بخش‌های میانی، شمال‌شرقی و جنوب‌غربی مشاهده شد. با توجه به نقشه‌ی شاخص توان جریان در مناطق هموار اندازه‌ی شاخص توان جریان کمتر بود. اندازه‌ی شاخص انتقال رسوب منطقه‌ی پژوهش از ۰/۹ تا ۱/۴۶ متغیر بود. مناطق با شاخص انتقال رسوب زیاد در بخش‌های میانی منطقه قرار گرفتند. براساس محاسبه‌های فاصله‌ی اقلیدسی پیکسل‌ها از آبراهه، اندازه‌ی فاصله از آبراهه در این منطقه از صفر تا ۵۴۹۵ متر متغیر بود. بخشی از آبخیز که فاصله کمی تا آبراهه یا رودخانه داشت، ظرفیت بیشتری نیز برای سیلاب و آب‌گرفتگی داشت. اندازه‌ی تراکم زهکشی آبخیز پژوهش‌شده بین صفر تا ۰/۸۲ متغیر بود. بیشترین اندازه‌ی تراکم زهکشی در بخش‌های میانی آبخیز بود. بخش‌هایی که تراکم زهکشی زیاد می‌باشد، طول آبراهه در یک واحد سطح مشخص بیشتر است و بنابراین ظرفیت رخداد سیل نیز بیشتر خواهد بود. از نظر خصوصیت‌های خاک‌شناسی چهار طبقه مختلف در منطقه قابل تفکیک بود. این چهار طبقه نوع خاک براساس نقشه‌ی خاک‌شناسی ثبت‌شده به‌وسیله‌ی سازمان منابع‌طبیعی و آبخیزداری کشور شامل انتی سولز، انسپتی سولز، ورتی سولز و زمین‌های بدخیم طبقه انسپتی سولز با بیشترین مساحت در بخش‌های میانی و شرقی منطقه قرار گرفته است. در آبخیز پژوهش‌شده هفت نوع کاربری زمین شامل کشاورزی آبی، کشاورزی دیم، جنگل، باغ، مرتع، مسکونی و پهنه‌های آبی بود. همچنین کاربری‌های جنگل و کشاورزی دیم، بیشترین مساحت و کاربری‌های پهنه‌ی آبی و مسکونی، کمترین مساحت را داشتند.

Archive of SID.ir

Archive of SID.ir



شکل ۳- عامل‌های استفاده‌شده در پژوهش.

Figure 3- The Factors Used in Current Research.

Archive of SID.ir

Archive of SID.ir

حساسیت سیل براساس مدل وزن شواهد، جنگل تصادفی و ترکیب مدل‌ها

نتایج مدل وزن شواهد در جدول ۱ نشان‌داده شده است. وزن‌های مثبت نشان‌دهنده‌ی رابطه‌ی مستقیم طبقه مربوطه و ظرفیت سیل‌گیری است، درحالی‌که وزن‌های منفی نشان‌دهنده‌ی رابطه‌ی معکوس و عدد صفر نشان‌دهنده‌ی نبودن رابطه‌ی معنی‌دار است. براساس تحلیل‌های انجام‌شده، بیشترین وزن شاخص ناهمواری (۲/۱۲۲) در طبقه اول بود و این طبقه رابطه‌ی مستقیمی با ظرفیت سیل‌گیری داشت. بیشترین وزن شاخص رطوبت پستی‌بلندی (۱/۶۹۷) مربوط به طبقه پنجم بود که در این مناطق به‌دلیل رواناب زیاد و سطح هموار، احتمال سیل‌گیری بیشترین بود. وزن منفی مناطقی که شاخص رطوبت پستی‌بلندی آن‌ها کم است نشان‌دهنده‌ی رابطه‌ی معکوس بین این مناطق و ظرفیت سیل‌گیری بود. وزن منفی به‌دست آمده در طبقه‌هایی با تراکم زهکشی کم، نشان‌دهنده‌ی کم استعدادبودن این مناطق برای سیل‌گیری است؛ در حالی که وزن‌های زیاد (۶/۱۶۲ و ۵/۷۶۴) مناطق با تراکم زهکشی بیشتر، نشان‌دهنده‌ی احتمال زیاد رخداد سیل‌گیری در این بخش‌ها است. مناطق با بلندی کم بیشترین وزن را داشتند (۱/۲۷۳ و ۶/۱۸۷) و ظرفیت سیل‌گیری در این بخش‌ها زیاد بود. نتایج نشان داد که کاربری باغ (۶/۹۶) و کشاورزی آبی (۴/۷۶) بیشترین رخداد سیل و همچنین بیشترین وزن را داشتند. نتایج به‌دست آمده‌ی عامل فاصله از آبراهه نشان داد که دو طبقه اول بیشترین وزن و طبقه‌های چهارم و پنجم کمترین وزن را داشتند. طبقه‌های اول (۰-۵۰) و دوم (۵۰-۱۰۰) اگرچه سطح کمی داشتند، اما بیشترین تجربه‌ی رخداد سیل‌گیری را داشتند و براساس محاسبه‌ی مدل وزن شواهد بیشترین وزن را به خود اختصاص دادند. مناطق کم شیب تعداد رخداد سیل‌گیری بیشتری را تجربه کردند و بیشترین وزن (۳/۵۱۵) را در میان طبقه‌های عامل شیب زمین به‌دست آوردند. از دیدگاه خاک‌شناسی، طبقه خاک انسپتی‌سول بیشترین وزن (۱/۷۲۲) را داشت. براساس نتایج مدل وزن شواهد، در مناطقی که شاخص توان جریان کم بود، جریان انرژی نیز کم و سیل‌گیری اتفاق افتاد و طبقه اول این عامل بیشترین وزن (۲/۹۰۷) را به دست آورد. شاخص انتقال رسوب نیز مانند شاخص توان جریان به گونه‌ای نشان‌دهنده‌ی اندازه‌ی انرژی جریان برای حمل مواد است که در این پژوهش مناطق با شاخص انتقال رسوب کم (۰/۰ - ۰/۰۰۰۰۱۶)، وزن زیادی داشتند. نتایج مدل وزن شواهد برای شاخص موقعیت ناهمواری نشان داد در مناطقی که اندازه‌ی این شاخص کم است، رخدادهای سیل بیشتر بوده و لذامعادله‌های مدل وزن شواهد، بیشترین وزن را به این طبقه اختصاص داده‌اند (۰/۵۵۱).

Archive of SID.ir

Archive of SID.ir

جدول ۱- وزن طبقه‌های عامل‌های محیطی براساس مدل وزن شواهد.

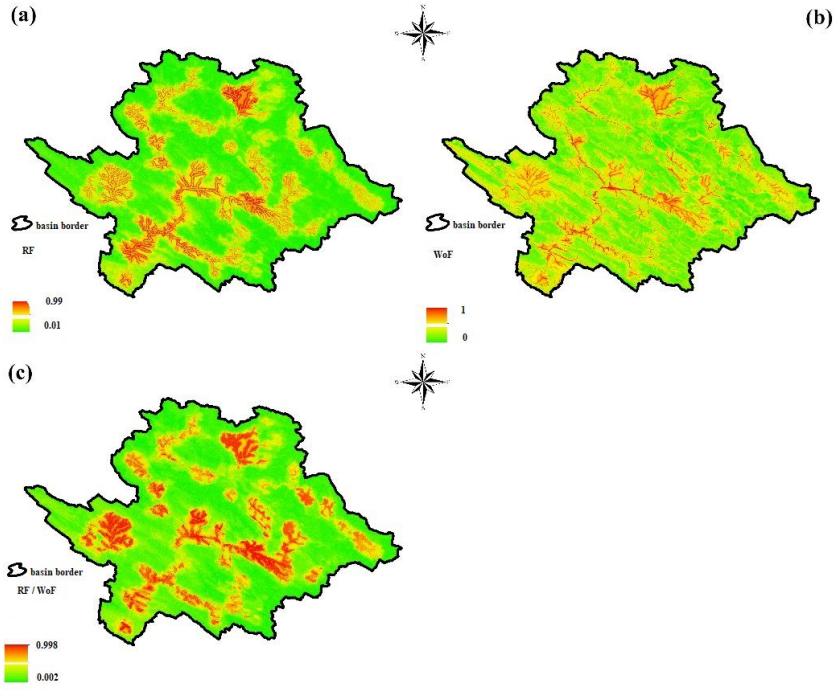
Table 1- The Weight of Classes of Environmental Factors Based on the Weight of Evidence Model.

Enviroment Affect	Class	W^+	W^-	C	$S2(W^+)$	$S2(W^-)$	WoF
Terrian Ruggness	0 - 4.16	0.267	-0.254	0.521	0.027	0.033	2.122
	4.16 - 8.55	-0.183	0.071	-0.253	0.059	0.020	-0.901
	8.55 - 14.26	-0.588	0.098	-0.686	0.143	0.017	-1.717
	14.26 - 24.86	-0.014	0.001	-0.015	0.200	0.016	-0.032
	24.86 - 156.08	0.505	-0.006	0.511	1.000	0.015	0.507
Topography Wetness Index	5.10 - 9.0	-0.018	0.006	-0.023	0.063	0.020	-0.081
	9.0 - 10.32	-0.002	0.001	-0.003	0.040	0.024	-0.012
	10.32 - 11.90	-0.229	0.064	-0.292	0.077	0.019	-0.947
	11.90 - 14.35	0.139	-0.017	0.157	0.125	0.017	0.416
	14.35 - 21.15	0.747	-0.042	0.789	0.200	0.016	1.697
Drainage Density	0 - 0.132	0	0.241	0	0	0.015	0.000
	0.132 - 0.238	-2.183	0.277	-2.460	0.500	0.015	-3.427
	0.238 - 0.345	-0.409	0.105	-0.514	0.091	0.018	-1.557
	0.345 - 0.480	0.967	-0.542	1.510	0.028	0.032	6.162
	0.480 - 0.819	1.348	-0.241	1.589	0.056	0.020	5.764
Elevation	<750	0.896	-0.018	0.914	0.500	0.015	1.273
	750-1000	1.552	-0.221	1.773	0.063	0.020	6.187
	1000-1500	0.191	-0.132	0.323	0.033	0.027	1.316
	1500-2000	-0.336	0.172	-0.507	0.053	0.021	-1.872
	>2000	0	0.187	0	0	0.015	0.000
Landuse	Agriculture	1.504	-0.128	1.632	0.100	0.018	4.760
	Dryfarming	-0.394	0.164	-0.558	0.063	0.020	-1.946
	Forest	-0.582	0.234	-0.816	0.071	0.019	-2.716
	Orchard	2.505	-0.117	2.622	0.125	0.017	6.960
	Range	0.189	-0.061	0.250	0.056	0.020	0.908
	Residental	0.801	-0.008	0.809	1.000	0.015	0.803
	Water	0	0.000	0	0	0.015	0.000
Distance form River	0-50	2.834	-0.948	3.782	0.024	0.040	14.973
	50-100	1.715	-0.180	1.895	0.077	0.019	6.134
	100-200	0.492	-0.044	0.537	0.143	0.017	1.343
	200-500	-1.810	0.171	-1.981	0.500	0.015	-2.760
	>500	-3.129	1.115	-4.244	0.500	0.015	-5.911
Slope	0-5	0.564	-0.300	0.864	0.033	0.027	3.515
	5-10	-0.524	0.113	-0.636	0.111	0.017	-1.776
	10-15	-0.266	0.048	-0.314	0.111	0.017	-0.878
	15-20	0.038	-0.006	0.044	0.111	0.017	0.122
	>20	-0.360	0.079	-0.439	0.100	0.018	-1.280
Soil Type	Badland	0	0.012	0	0	0.015	0.000
	Entisol	-0.724	0.110	-0.834	0.167	0.016	-1.950
	Inceptisol	0.146	-0.337	0.483	0.020	0.059	1.722
	Vertisol	0.034	-0.007	0.041	0.091	0.018	0.124
Stream Power	0.000016 - 0.264	0.392	-0.320	0.712	0.028	0.032	2.907
	0.264 - 0.708	-0.419	0.073	-0.492	0.125	0.017	-1.306
	0.708 - 1.125	-0.015	0.004	-0.018	0.077	0.019	-0.059
	1.125 - 1.508	-0.729	0.091	-0.820	0.200	0.016	-1.763
	1.508- 3.543	-0.331	0.032	-0.363	0.200	0.016	-0.781
Sediment Transfer	0.909 - 0.992	0.489	-0.106	0.595	0.067	0.019	2.029
	0.992 - 1.057	0.236	-0.111	0.347	0.042	0.023	1.362
	1.057 - 1.116	0.046	-0.019	0.065	0.050	0.021	0.242
	1.116 - 1.183	-1.092	0.174	-1.266	0.200	0.016	-2.722
	1.183 - 1.459	-0.479	0.029	-0.509	0.333	0.016	-0.861
Topography Position	(-83.22) - (-3.83)	0.386	-0.010	0.396	0.500	0.015	0.551
	(-3.83) - (-0.85)	-0.074	0.015	-0.089	0.091	0.018	-0.270
	(-0.85) - (0.83)	0.055	-0.092	0.147	0.023	0.042	0.575
	(0.83) - (2.85)	-0.256	0.040	-0.296	0.125	0.017	-0.787
	(2.85) - (82.66)	0.082	-0.004	0.086	0.333	0.016	0.145

Archive of SID.ir

Archive of SID.ir

نقشه‌ی ظرفیت رخداد سیل براساس روش وزن شاهدها در شکل ۴ نشان‌داده شده است. با توجه به لکه‌های این شکل در بخش‌های بسیاری از منطقه ظرفیت سیل زیاد است و در بخش‌های شمالی، غربی، مرکزی، شرقی و جنوبی قرار گرفته‌اند. اندازه‌های میانه و انحراف معیار ارزش پیکسل‌ها در نقشه‌ی ظرفیت سیل به‌دست آمده از مدل وزن شاهدها به‌ترتیب ۰/۳۰۵ و ۰/۱۵۱۵ بود. نقشه‌ی ظرفیت رخداد سیل حاصل از مدل جنگل تصادفی در شکل ۴ نشان داد که ظرفیت رخداد سیل منطقه‌ی پژوهش از اندازه‌ی ۰/۰۱ تا ۰/۹۹ متغیر است. الگوی پراکنش مناطق با ظرفیت زیاد رخداد سیل نشان داد که عامل‌های پرشماری در رخداد سیل در منطقه مؤثر بوده و در بخش‌های مختلف منطقه می توان دامنه‌ی وسیعی از تغییرهای ظرفیت رخداد سیل را مشاهده کرد. اندازه‌ی احتمال رخداد سیل‌گیری در مدل ترکیبی جنگل تصادفی و وزن شاهدها از ۰/۰۰۲ تا ۰/۹۹۸ در منطقه‌ی پژوهش متغیر بود (شکل ۴). براساس این مدل ترکیبی میانگین ظرفیت سیل منطقه‌ی پژوهش ۰/۲۸۳ به‌دست آمد. نتایج ظرفیت سیل‌گیری منطقه‌ی پژوهش در مدل ترکیبی نشان داد که لکه‌های بحرانی در بخش‌های مختلفی توزیع شده‌اند و این مدل توانست این بخش‌ها را از نظر وضعیت سیل‌گیری به شکل مناسبی تفکیک کند. همچنین، در بخش‌هایی که ظرفیت سیل گیری بسیار زیاد بود، به‌طور دقیق‌تر مشخص شد و توزیع اندازه‌های ظرفیت سیل‌گیری منطقه نیز به‌خوبی انجام شد.



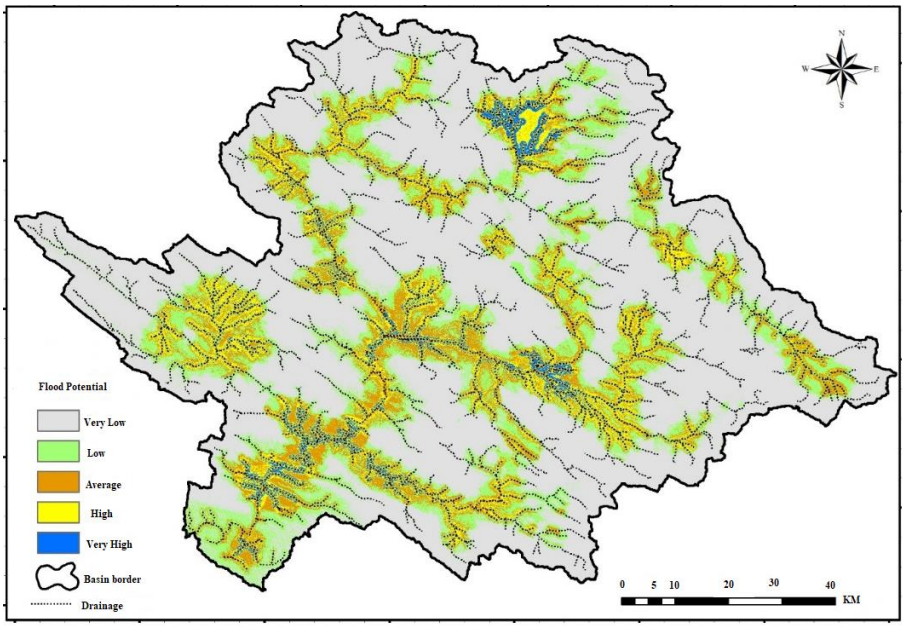
شکل ۴- نقشه‌ی ظرفیت سیل براساس مدل جنگل تصادفی، وزن شواهد و ترکیب دو مدل.

Figure 4- Flood Potential Map Based on Random Forest Model, Weight of Evidence and Combination of two Models.

مساحت هریک از طبقه‌های ظرفیت سیل (مدل ترکیبی جنگل تصادفی و وزن شواهد) در جدول ۲ و شکل ۵ نشان‌داده شده است. براساس نتایج به‌دست آمده، ۶۰/۲۴٪ مساحت منطقه در طبقه ظرفیت سیل خیلی کم، ۱۹/۲۵٪ در طبقه ظرفیت سیل کم، ۱۱/۷۴٪ در طبقه ظرفیت سیل متوسط، ۶/۸۸٪ در طبقه ظرفیت سیل زیاد و ۱/۸۷٪ در طبقه ظرفیت سیل خیلی زیاد قرار گرفتند. بنابراین، وضعیت ظرفیت سیل در بخش بسیاری از منطقه نیازمند توجه و مدیریت جدی است.

Archive of SID.ir

Archive of SID.ir



شکل ۵- نقشه‌ی طبقه‌بندی ظرفیت سیل با استفاده از مدل ترکیبی جنگل تصادفی و وزن شواهد.

Figure 5- Flood Potential Classification Map Using the Combined Random Forest Model and Weight of Evidence.

جدول ۲- مساحت طبقه‌های ظرفیت سیل.

Table 2- The Area of Flood Potential Classes.

Number	Area(٪)	Area (ha)	Flood Potential
1	60.24	567.226	Very low
2	19.25	181.308	low
3	11.74	110.645	Average
4	6.88	64.844	High
5	1.87	17.656	Very Hig

ارزیابی دقت مدل‌ها

نتایج مدل‌سازیاعتبارسنجی شده‌ی سیل و معیار کمی مساحت زیرمنحنی (AUC) ارزیابی شد. براساس تحلیل‌های به‌دست آمده از اعتبارسنجی، دقت مدل جنگل تصادفی ۰/۹۰۴، دقت مدل وزن شواهد ۰/۸۸۶ و دقت مدل جنگل تصادفی وزن شواهد ۰/۹۷۸ بود. براساس مقایسه‌ی نتایج اعتبارسنجی مدل‌های مختلف رویکرد ترکیب مدل‌های داده‌کاوی و ماشین یادگیری سبب ارتقاء عملکرد مدل و دقت پیش بینی ظرفیت سیل‌گیری شد (جدول ۳).

جدول ۳- نتایج ارزیابی مدل‌ها در بخش آموزش و اعتبارسنجی.

Table 3- Evaluation Results of Models in the Training and Validation Section.

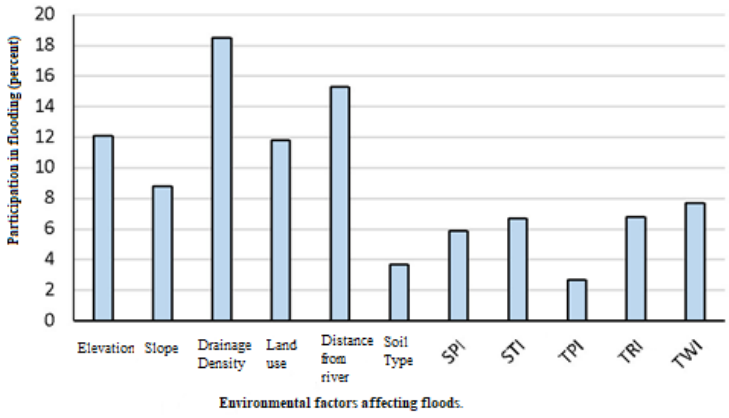
		AUC	
Models	**Model Accuracy**	**Validation**	**Training**
Random Forest	Exclent	0.904	1
Weight of Evidence	Very Good	0.886	0.966
Combination RF/WoF	Exclent	0.978	1

اندازه‌ی مشارکت عوامل محیطی در ظرفیت سیل

براساس نتایج مدل جنگل تصادفی اندازه‌ی مشارکت عامل‌های محیطی در سیل در شکل ۵ نشان‌داده شده است. در منطقه‌ی پژوهش‌شده تمامی عامل‌های محیطی در سیل نقش داشته و این دستاورد نشان داد که سیل یک پدیده‌ی پیچیده است که تحت تأثیر شرایط چندعامله در آبخیز اتفاق می‌افتد. براساس نتایج به‌دست آمده، عامل زهکشی بیشترین اهمیت را داشت و اندازه‌ی مشارکت آن در سیل منطقه‌ی پژوهش‌شده ۱۸/۵٪ تعیین شد. اندازه‌ی اهمیت دیگر عامل‌ها از جمله بلندی ۱۲/۱٪، شیب زمین ۸/۸٪، کاربری زمین ۱۱/۸٪، فاصله از آبراهه ۱۵/۳٪، نوع خاک ۳/۷٪، شاخص SPI ۵/۹٪، شاخص STI ۶/۷٪، شاخص TPI ۲/۷٪، شاخص TRI ۶/۸٪ و شاخص TWI ۷/۷٪ به‌دست آمد.

Archive of SID.ir

Archive of SID.ir



شکل ۵- اندازه‌ی مشارکت عوامل محیطی در سیل.

Figure 5- The Degree of Participation of Environmental Factors in the Occurrence of Floods.

بحث و نتیجه‌گیری

در این پژوهش، مدل‌سازی مکانی ظرفیت سیل با استفاده از موقعیت مکانی رخدادهای سیل‌های گذشته انجام شد. نتایج نشان داد که اطلاعات مکانی رخدادهای سیل می‌تواند کمک شایانی به مدل‌سازی سیل کند و در نهایت نقشه‌ی علمی و کاربردی ظرفیت سیل به‌دست آید. ازآنجایی‌که رخدادهای سیل تحت تأثیر عامل‌های محیطی مختلفی پدید می‌آید، استفاده از اطلاعات واقعی رخداد سیل که به‌عنوان رخدادهای تاریخی نیز شناخته می‌شوند، امکان نزدیک‌تر شدن نتایج مدل‌سازی به واقعیت را بیشتر می‌کند. سیاه‌کمری و همکاران (۲۰۱۸) نیز به اهمیت موقعیت مکانی رخدادهای سیل اشاره کرده‌اند. استفاده از موقعیت مکانی رخدادهای سیل گذشته می‌تواند تا حدودی شرایط سیل را نشان دهد. این موضوع برای کشورهای در حال توسعه که همواره با کمبود داده در پژوهش‌های مهارکردن سیل مواجه‌اند، مفید است.

براساس نتایج، بسیاری از مساحت منطقه‌ی پژوهش در بخش‌های جنوب‌غربی، مرکز، شمال و شرق ظرفیت سیل زیاد بود. این بخش‌ها همچنین به‌دلیل نوع کاربری و نوع پوشش گیاهی حساسیت بیشتری داشتند. این محدوده‌ها به‌شکل لکه‌ای بوده و با استفاده از نقشه‌ی ظرفیت سیل می‌توان برای هریک از آنها برنامه‌های جامعی پیشنهاد داد. از آنجایی‌که نقشه‌ی ظرفیت سیل به‌دست آمده از مدل‌سازی در این پژوهش با استفاده از داده‌های واقعی و مشاهده‌شده‌ی رخداد سیل (اعتبارسنجی شده) ارزیابی شد، نتایج آن قابل اطمینان بوده و می‌توان در مدیریت سیل آبخیز پژوهش به‌کار برد. اقدام‌های اجرایی (ساخت سازه‌های آبخیزداری، بندهای مهارکردن جریان، دیواره‌های سیل‌گیر در رودخانه‌ها و احیاء پوشش گیاهی در سطح آبخیز و غیره) و اقدام‌های مدیریتی (الگوی صحیح کاربری زمین، جلوگیری از تجاوز به حریم قانونی رودخانه‌ها، قانونمند کردن برداشت مصالح رودخانه‌ای، تعادل دام و مرتع و غیره) می‌تواند در هر بخش از آبخیز براساس نقشه‌ی ظرفیت سیل انجام شود. نتایج این پژوهش نشان داد که عامل تراکم زهکشی در پدیده‌ی سیل بیشترین اهمیت را دارد. دیگر عامل‌های محیطی نیز اهمیت معنی‌داری در رخداد پدیده‌ی سیل نشان دادند. عامل‌های فاصله از آبراهه، بلندی، کاربری زمین، شیب و شاخص رطوبت پستی‌بلندی نیز مشارکت بسیاری در پدیده‌ی سیل منطقه‌ی پژوهش‌شده داشتند. توجه به این نکته ضروری است که اندازه‌ی اهمیت متغیرهای محیطی از یک منطقه به منطقه‌ای دیگر متفاوت خواهد بود. بنابراین، اولویت‌بندی اهمیت عامل‌های محیطی یک منطقه را نمی توان برای منطقه‌ای دیگر استفاده کرد. از این‌رو، برنامه‌ریزی مدیریت سیل مبتنی بر مشارکت عامل‌های محیطی هر منطقه باید براساس اهمیت عامل‌های محیطی همان منطقه انجام شود. براساس نتایج این پژوهش، دقت مدل جنگل تصادفی ۰/۹۰۴، دقت مدل وزن شاهدها ۰/۸۸۶ و دقت مدل جنگل تصادفی وزن شاهدها معادل با ۰/۹۷۸، به‌دست آمد. براساس مقایسه‌ی نتایج اعتبارسنجی مدل‌های مختلف، رویکرد ترکیب مدل‌های داده‌کاوی و ماشین یادگیری موجب ارتقاء عملکرد مدل و دقت پیش‌بینی ظرفیت سیل‌گیری شد. در پژوهش چن و همکاران (۲۰۲۰) نیز عملکرد مناسب مدل‌های یادگیری ماشینی در زمینه‌ی پیش‌بینی مکانی ظرفیت سیل بررسی و تأییدشده است. اگرچه در برخی پژوهش‌ها (پرادهان و همکاران ۲۰۱۵) ترکیب مدل‌های داده‌کاوی برای پیش‌بینی مکانی ظرفیت سیل استفاده شده، اما ترکیب مدل‌های یادگیری ماشینی و مدل‌های داده‌کاوی کمتر استفاده‌شده است. در این پژوهش این موضوع عمیق و دقیق بررسی شد و نتایج ارزنده‌ای به‌دست آمد که می‌تواند برای ارتقاء کارایی دستگاه‌های اجرایی و پژوهشی در مهارکردن و مدیریت بهتر سیلاب‌ها به‌کار برده شود.

Archive of SID.ir

Archive of SID.ir

فهرست منابع

فهرست منابع

Arnal L, Ramos MH, Coughlan de Perez E, Louise Cloke H, Stephens E, Wetterhall F, Jan van Andel S, Pappenberger F. 2016. Willingness-to-pay for a probabilistic flood forecast: A risk-based decision-making game. Hydrology and Earth System Sciences, 20(8): 3109-3128.

Corsini A, Cervi F, Ronchetti F. 2009. Weight of evidence and artificial neural networks for potential groundwater spring mapping: an application to the Mt. Modino area (Northern Apennines, Italy). Geomorphology, 111(1-2): 79–87.

Darabi H, Rahmati O, Naghibi SA, Mohammadi F, Ahmadisharaf E, Kalantari Z, Torabi Haghighi A, Soleimanpour SM, Tiefenbacher JP, Tien Bui D. 2022. Development of a novel hybrid multi-boosting neural network model for spatial prediction of urban flood. Geocarto International, 37(19): 5716-5741.

Dehghanian N, Mousavi Nadoushani SS, Saghafian B, Rayati Damavandi M. 2020. Evaluation of coupled ANN-GA model to prioritize flood source areas in ungauged watersheds. Hydrology Research, 51(3): 423-442.

Deng S, Liu S, Mo X. 2020. Assessment of Three Common Methods for Estimating Terrestrial Water Storage Change with Three Reanalysis Datasets. Journal of Climate, 33(2): 11–525.

Donyaii A, Sarraf A. P, Mardanifar M. 2020. Evaluation of Crisis Resolution Strategies for Groundwater Revival Plan Using Fuzzy Best – Worst Muti Criteria Decision Model. Journal of Hydraulic Structures. 6 (3): 21-44.

Donyaii A, Sarraf AP, Ahmadi H. 2021. Optimization of reservoir dam operation using gray wolf, crow search and whale algorithms based on the solution of the nonlinear programming model. JWSS-Isfahan University of Technology. 24(4): 159-175.

Donyaii A, Sarraf A. P, Ahmadi H. 2021. Multi-Objective Optimal Utilization Policy of Boostan Dam Reservoir Using Whale and NSGA-II Algorithms Based on Game Theory and Shannon Entropy Method. Iranian Water Researches Journal. 14(4): 99-111.

Felicísimo Á, Cuartero A, Remondo J, Quirós E. 2013. Mapping landslide susceptibility with logistic regression, multiple adaptive regression splines, classification and regression tress, and maximum entropy methods: A comparative study. Landslides, 10,(11): 175-189.

Kazemi R, Porhemmat J. 2021. An Analysis of the factors Affecting Flooding Severity in Iran. Watershed Management Research Journal, 34(1): 59-73.

Lee S, Kim JC, Jung HS, Lee MJ, Lee S. 2017. Spatial prediction of flood susceptibility using random-forest and boosted-tree models in Seoul metropolitan city, Korea.Geomatics, Natural Hazards and Risk, pp. 1-19.

Mirzaee N, Sarraf A. P. 2022. Application of data fusion models in river flow simulation using signals of large-scale climate, case study: Jiroft Dam Basin, Watershed Engineering and Management. 13(4): 672-689.

Moghaddam DD, Pourghasemi HR, Rahmati O. 2019. Assessment of the Contribution of Geo-environmental Factors to Flood Inundation in a Semi-arid Region of SW Iran: Comparison of Different Advanced Modeling Approaches. In Natural Hazards GIS-Based Spatial Modeling Using Data Mining Techniques. Springer, Cham, pp. 59-78.

Poudyal CP, Chang C, Oh HJ, lee S. 2010. Landslide susceptibility maps comparing frequency ratio and artificial neural networks: a case study from the Nepal Hima-

Archive of SID.ir

Archive of SID.ir

laya. Environmental Earth Sciences, 61(5): 1049-1064.

Porhemmat J. 2016. A model on investigation on flood hazard over watersheds of Iran. Iranian Journal of Watershed Management and Science, 10(34): 1-15.

Pradhan B, Neamah Jebur M, Zulhaidi Mohd Shafri H, Shafapour Tehrany M. 2015. Data fusion technique using wavelet transform and Taguchi methods for automatic landslide detection from airborne laser scanning data and quickbird satellite imagery, IEEE Transactions on Geoscience and remote sensing. 54(3): 1610-1622.

Rahmati O, Tahmasebipour N, Pourghasemi HR. 2015. Sub-watershed flooding prioritization using morphometric and correlation analysis (Case study: Golestan Watershed). Iranian journal of Ecohydrology, 2(2): 151-161.

Rahmati O, Pourghasemi HR. 2017. Identification of critical flood prone areas in data-scarce and ungauged regions: A comparison of three data mining models. Water Resources Management, 31(5): 1473-1487.

Sampson CC, Smith AM, Bates PD, Neal JC, Alfieri L, Freer JE. 2015. A high-resolution global flood hazard model. Water Resource Research, 51(9): 358-7381.

Sarraf A. P, Mohagheghzadeh G, Mohagheghzadeh N. 2017. Flood Hydraulic Modeling and its Deliniation Using Orthophoto Pictures in Shapourkhesht River located in Kazeroun City. Geographic Space, 17(57): 175-194.

Satarzadeh E, Sarraf A. P, Hajikandi H, Sadeghian MS. 2022. Flood hazard mapping in western Iran: assessment of deep learning vis à vis machine learning models. Natural Hazard, 111(2): 1355-1373.

Shafapour Tehrany M, Kumar L, Neamah Jebur M, Shabani F. 2019. Evaluating the application of the statistical index method in flood susceptibility mapping and its comparison with frequency ratio and logistic regression methods. Geomatics, Natural Hazards and Risk, 10(1): 79-101.

Siahkamari S, Haghizadeh A, Zeinivand H, Tahmasebipour N, Rahmati O. 2018. Spatial prediction of flood susceptible areas using frequency ratio and maximum entropy models.Geocarto International,33(9): 927-941.

Soleimanipour M, Sarraf A. P. 2020. Evaluating the effects of climate change on Lar Basin Water Resources Using SWAT Model and comparing its results with Bayesian Networks and Hybrid Intelligent Models. Natural Geography 13 (50): 61-79. (In Persian).

Zamanirad M, Sedghi H, Sarraf A. P, Saremi A, Rezaee P. 2018. Potential Impacts of Climate change on groundwater levels on the Kerdi-Shirazi plain, Iran. Environmental Earth Sciences. 77(11): 1-10.

Zamanirad M, Sarraf A. P, Sedghi H, Saremi A, Rezaee P. 2019. Modeling the influence of Groundwater Exploitation on Land Subsidence Susceptibility Using Machine Learning Algorithms. Natural Resources Research, 29 (1): 127-1141.

Archive of SID.ir

Archive of SID.ir

Fars Agricultural and Natural Resources Research and Education Center

Watershed Management Research

ISSN: 2981-2038

Agricultural Research, Education and Extension Organization

Spatial Prediction of Flood Susceptable Areas in Karkheh Watershed of Lorestan Province Using the Combined Random Forest – Weight of Evidence Model

Eslam Satarzadeh[1], Amirpouya Sarraf [*2], Houman Hajikandi[3], Mohammadsadegh Sadeghian[4]

1- Ph.D. Student, Department of Civil Engineering, Central Tehran Branch, Islamic Azad University, Tehran, Iran
2- Assistant Professor, Department of Civil Engineering, Roudehen Branch, Islamic Azad University, Roudehen, Iran
3 and 4- Assistant Professor, Department of Civil Engineering, Central Tehran Branch, Islamic Azad University, Tehran, Iran

Extended Abstract

Introduction and goal

The phenomenon of flood is one of the frequent hazards that have caused financial losses and many lives in Iran in the last decade. One of the three main natural hazards of Iran is the occurrence of floods, and it is safe to say that a large flood occurs at least once a year in some part of this country. Floods are considered as a great threat to human life by damaging or killing people and animals, especially buildings and houses, agricultural land and crop production, urban infrastructure, bridges and roads. Floods in Iran cause a lot of damage from an economic point of view; it causes destruction of the environment, natural and residential resources and loss of life. In the past years, about 70% of the annual credits of the plan to reduce the effects of natural disasters and the headquarters of unexpected events have been used to compensate for the damages caused by floods. In this research, by using flood location data and using machine learning models and random forest data mining and the weight of evidence and their combination, the spatial prediction of flood prone areas has been discussed.

Materials and methods

In this research, combined models and 11 predicting variables of flood probability in Karkheh

Article Type: Research Article

***Corresponding Author E-mail:** sarraf@riau.ac.ir

Citation: Satarzadeh, E., Sarraf, A. P., Hajikandi, H., Sadeghian, M.S. 2023. Spatial Prediction of Flood Susceptable Areas in Karkheh Watershed of Lorestan Province Using the Combined Random Forest – Weight of Evidence Model. Watershed Management Research. 36(2): 87-103.

DOI: 10.22092/wmrj.2022.359264.1483

Received: 10 July 2022, **Received in revised form:** 25 August 2022, **Accepted:** 21 September 2022, **Published online:** 22 June 2023

Watershed Management Research, VOL. 36, No. 2, Ser. No: 139, Summer 2023, pp. 87-103.

Publisher: Fars Agricultural and Natural Resources Research and Education Center ©Author(s)

Archive of SID.ir

Archive of SID.ir

watershed located in Lorestan province have been used. These variables include maps of geomorphometric indicators, including topographic humidity index, relative slope position and topographic position index, hydrological maps including: drainage density and distance from the drainage network, for this purpose, first, data mining models for the initial analysis of the relationships between environmental variables and events Past risks are used and their results are used as input data for machine learning models. Flood data were randomly divided into two groups: training (70%) and validation (30%). Prediction accuracy was evaluated using operating characteristic curve (ROC) method.

Results and discussion

According to the results of the random forest model, the accuracy was 0.904, the weight of evidence was 0.886, and the combined model of random forest - the weight of evidence was 0.978. Based on the combined model of random forest and the weight of evidence as the superior model, 20.49% of the surface has medium upward potential. Based on the random forest model, drainage density, distance from waterways, height and land use were the most important factors affecting flood potential.

Keywords: Environmental factors, flood, flood susceptibility, modeling

Archive of SID.ir