



مدیریت آب و آبادانی

دوره ۱۱ ■ شماره ۴ ■ زمستان ۱۴۰۰

صفحه‌های ۸۸۵-۹۰۴

DOI: 10.22059/jwim.2022.332875.934

مقاله پژوهشی:

پیش‌بینی مناطق مستعد وقوع سیل با استفاده از مدل‌های پیشرفته یادگیری ماشین (دشت بیرجند)

سید احمد اسلامی نژاد^۱، ممین افخاری^۲، محمد اکبری^{۳*}، علی حاجی‌الیاسی^۴، هادی فرهادیان^۵

۱. دانش‌آموخته کارشناسی ارشد، گروه مهندسی نقشه‌برداری، دانشکده مهندسی نقشه‌برداری و اطلاعات مکانی، دانشگاه تهران، تهران، ایران.
۲. کارشناس ارشد مهندسی عمران آب و سازه‌های هیدرولیکی، عضو باشگاه پژوهشگران جوان و نخبگان، واحد مشهد، دانشگاه آزاد اسلامی، مشهد، ایران.
۳. دانشیار، گروه مهندسی عمران، دانشکده مهندسی، دانشگاه بیرجند، بیرجند، ایران.
۴. دانش‌آموخته کارشناسی ارشد، گروه آب و سازه‌های هیدرولیکی، دانشگاه صنعتی خواجه نصیرالدین طوسی، تهران، ایران.
۵. استادیار، گروه معدن، دانشکده مهندسی، دانشگاه بیرجند، بیرجند، ایران.

تاریخ دریافت مقاله: ۱۴۰۰/۰۸/۰۲
تاریخ پذیرش مقاله: ۱۴۰۰/۱۰/۱۶

چکیده

پژوهش‌ها در مورد مدل‌های پیش‌بینی سیل، یکی از اقدام‌های اولیه در کاهش خسارت سیل و مدیریت سیل‌های آینده در حوضه‌های آبریز است. هدف از مطالعه حاضر، ارزیابی حساسیت سیل در حوضه آبریز دشت بیرجند از طریق چهار مدل یادگیری ماشین شامل ماشین بردار پشتیبان (SVM)، درخت تصمیم J48 جنگل تصادفی (RF) و سیستم استنتاج عصبی فازی (ANFIS) است. لذا جهت پیاده‌سازی و اعتبارسنجی مدل‌های ذکر شده، فهرستی از مناطق مستعد سیل در منطقه موردمطالعه تهیه شد (۴۲ موقعیت سیل خیز). علاوه بر این، ۱۹ معیار هیدرولوژیکی، توپوگرافی، زمین‌شناسی و محیطی مؤثر بر وقوع سیل در مطالعه استخراج شدند تا برای پیش‌بینی نقشه حساسیت سیل مورداستفاده قرار گیرند. نتایج نشان داد که بالاترین دقت مریبوط به مدل RF (۰/۸۴۵) و کمترین دقت مریبوط به مدل SVM (۰/۷۹۱) بود. علاوه بر این، اعتبارسنجی نتایج با استفاده از منحنی ROC نشان داد که دقیق‌ترین مقادیر حساسیت سیل نیز به مدل RF اختصاص دارد (AUC = ۰/۹۵۸). نتایج این مطالعه می‌توانند به منظور مدیریت مناطق آسیب‌پذیر و کاهش خسارات سیل استفاده شود.

کلیدواژه‌ها: حساسیت سیل، دشت بیرجند، سیستم اطلاعات مکانی (GIS)، یادگیری ماشین.

Predicting flood prone areas using advanced machine learning models (Birjand Plain)

Seyed Ahmad Eslaminezhad¹, Mobin Eftekhar², Mohammad Akbari^{3*}, Ali Hajji Elyasi⁴, Hadi Frahadian⁵

1. M.Sc. Graduate, Department of Surveying Engineering, Faculty of Surveying Engineering and Spatial Information, University of Tehran, Tehran, Iran.

2. M.Sc., Civil Engineering, Water and Hydraulic Structures, Young Researchers and Elite Club, Mashhad Branch, Islamic Azad University, Mashhad, Iran.

3. Associate Professor, Department of Civil Engineering, Faculty of Engineering, University of Birjand, Birjand, Iran.

4. M.Sc. Graduate, Department of Water and Hydraulic Structure, K. N. Toosi University of Technology, Tehran, Iran.

5. Assistant Professor, Department of Mining Engineering, Faculty of Engineering, University of Birjand, Birjand, Iran.

Received: October, 24, 2021

Accepted: January, 06, 2022

Abstract

Research on flood predicting models is one of the first steps in reducing flood damage and managing future floods in catchments. The aim of this study was to evaluate flood susceptibility in Birjand plain catchment through four machine learning models including support vector machine (SVM), J48 decision tree, random forest (RF) and Adaptive neuro fuzzy inference system (ANFIS). Therefore, in order to implement and validate the mentioned models, a list of flood-prone areas in the study area was prepared (42 flood-prone locations). In addition, 19 hydrogeological, topographical, geological and environmental criteria affecting flood occurrence in the study area were extracted to be used to predict flood susceptibility map. The results showed that the highest accuracy was related to the RF model (0.845) and the lowest accuracy was related to the SVM model (0.791). In addition, the validation of the results using the ROC curve showed that the most accurate values of flood susceptibility belong to the RF model (AUC = 0.958). The results of this study can be used to manage vulnerable areas and reduce flood damage.

Keywords: Birjand plain, flood susceptibility, Geospatial information system (GIS), machine learning.

مقدمه

به اطلاعات دقیق در مورد مدل رقومی ارتفاعی (DEM^۱) می‌باشد (Dat *et al.*, 2019). امروزه با توسعه فناوری‌های مرتبط با سیستم اطلاعات مکانی (GIS^۲) و سنجش از دور (RS^۳)، تعیین مناطق مستعد سیل از طریق روش‌های یادگیری ماشین، بسیار سریع‌تر است و داده‌های موردنیاز برای مدل‌سازی اغلب آزادانه در دسترس هستند (Azareh *et al.*, 2019; Arabameri *et al.*, 2019) استفاده از مدل‌های یادگیری ماشین مدرن، کسب نتایج نهایی با دقت بالا را تضمین می‌کند. براساس پژوهش‌های پیشین، از دو رویکرد ادغام دانش‌محور^۴ و داده‌محور^۵ جهت پنهان‌بندی مناطق حساس به سیل استفاده می‌شود (Wang & Liu, 2019; Khosravi *et al.*, 2016). روش‌های داده‌محور در مناطق شناخته‌شده یا مناطقی که از لحاظ آماری تعداد شواهد شناخته شده کافی می‌باشند، کارایی بالایی دارند، درحالی‌که روش‌های دانش‌محور در محیط‌های کمتر شناخته شده‌اند و یا تعداد کمی از اهداف موردنظر در محدوده وجود دارند، کارآمد هستند. به تازگی مطالعات متعددی برای شناسایی مناطق مستعد سیل توسط رویکردهای مبتنی بر روش‌های ادغام دانش‌محور و داده‌محور انجام شده است که در ادامه به برخی از آن‌ها اشاره شده است.

Zeraatkar & Hassanpour (2016) رفتار سیلاب درون‌شهری بیرون‌جند واقع در حوضه شاهروود را موردمطالعه قرار دادند. در این راستا در ابتدا با توجه به تصاویر ماهواره‌ای، کاربری اراضی و بازدیدهای میدانی شماره منحنی رواناب تعیین شد. در ادامه پس از برداشت مشخصات هندسی رودخانه اصلی و سازه‌های هیدرولیکی، با استفاده از مدل HEC-HMS دنی سیلاب حوضه شبیه‌سازی شد. نتایج حاصل از شبیه‌سازی نشان داد که دوره بازگشت‌های ۲۵ و ۱۰۰ سال به ترتیب منجر به وقوع دبی‌های اوج سیلاب به میزان ۲۰۷/۶۳ و ۴۶۳/۵۶

سالیانه مخاطرات طبیعی باعث خسارت‌های مالی و انسانی گسترده به انسان‌ها در سراسر جهان می‌شود (Zhang & Wang, 2019; Alam *et al.*, 2020) بسیاری از نقاط جهان، تغییرات اقلیم، کاهش نفوذپذیری و جنگل‌زدایی‌های کترول‌نشده منجر به تقویت شدت و فراوانی بارش‌های سیل‌آسا می‌شوند که عامل اصلی ایجاد Shahid *et al.*, 2016; Markus *et al.*, 2016; Giang *et al.*, 2018; Nachappa *et al.*, 2021 (2020)، خسارت‌های اقتصادی ناشی از سیل‌های جهانی تا سال ۲۰۵۰ به حدود ۳۰۰ میلیارد دلار خواهد رسید. بررسی و آمارها نشان می‌دهد سیل در پنج دهه اخیر بیش از نیمی از کل خسارات ناشی از بلایای طبیعی در دنیا را به خود اختصاص داده است (Nachappa *et al.*, 2021; Saedi *et al.*, 2020; Arabameri *et al.*, 2019) سیل در کشور ایران مختص به منطقه و موقعیت خاصی نبوده و تمام کشور از این پدیده متأثر است، اما با توجه به ویژگی‌های هر منطقه، تفاوت در نوع سیل و میزان خسارت‌های واردشده وجود دارد (Saedi *et al.*, 2020; Giang *et al.*, 2020). با وجود تلاش‌های متخصصان، تصمیم‌گیران، ذی‌نفعان و اداره‌های دولتی در دهه‌های اخیر برای کاهش اثرات سیل، تعداد حوادث و تلفات اقتصادی و انسانی مرتبط با آن، در سراسر جهان در حال افزایش است (Johann & Leismann, 2017; Kocaman, 2017; Chen *et al.*, 2020). با توجه به افزایش وقوع سیل به ویژه در سطح شهرها و به وجود آمدن خطرهای جانی، مالی و محیط زیستی ناشی از آن، شناسایی معیارهای مؤثر بر وقوع سیل از اهمیت بالایی برخوردار می‌باشد (Hong *et al.*, 2018). مدل‌سازی‌های MIKE یا هیدرولیکی اجرشده توسط نرم‌افزارهایی مانند HEC-RAS بسیار وقت‌گیر است و برای پیاده‌سازی نیاز

مدیریت آب و آسیاری

^{۱۷}) با الگوریتم‌های فرابتکاری شامل الگوریتم بهینه‌سازی کلونی مورچه‌ها (ACO)^{۱۸}، الگوریتم ژنتیک (GA)^{۱۹} و الگوریتم بهینه‌سازی ازدحام ذرات (PSO)^{۲۰} جهت پیش‌بینی حساسیت سیل در شهرستان جهرم استان فارس استفاده کردند. بهمنظور پیاده‌سازی روش‌های موردنظر، نه معیار مؤثر بر وقوع سیل استخراج شدند. سطح زیر منحنی (AUC) به دست آمده از ROC، دقت ۹۱/۸، ۹۲/۶ و ۹۴/۵ درصد را به ترتیب برای مدل‌های ANFIS-GA، ANFIS-ACO و ANFIS-PSO نشان داد. بنابراین مدل ترکیبی ANFIS-PSO به عنوان مدل برتر در تهیه نقشه حساسیت سیل منطقه مورد مطالعه معرفی شد. Ziaian Firouz Abadi *et al.* (2020) به شناسایی مناطق مستعد پخش سیلاب از منظر سازندگان زمین‌شناسی در حوضه آبریز بیرجند پرداختند. تمامی فرایندها و تجزیه‌وتحلیل داده‌ها در محیط GIS به کمک سه روش فرایند تحلیلی سلسله مراتبی (AHP)^{۲۱}، مجموع ساده وزنی (SAW)^{۲۲} و منطق بولین برای وزن‌دار کردن و تلفیق لایه‌هایی همچون (شب، فرسایش، کاربری اراضی، شبکه زهکشی، خاک، پراکنش گسل، واحد زمین‌شناسی، سطح آب‌های زیرزمینی و تصاویر ماهواره‌ای) استفاده شد و مناطق پرمخاطره آن مشخص شد. با توجه به نقشه نهایی به دست آمده، ۲۷۶۵ هکتار از مساحت تعیین شده در حوضه آبریز بیرجند در اولویت اول (مناسب برای پخش سیلاب) و ۱۹۴۵ هکتار در اولویت دوم (نامناسب برای پخش سیلاب) قرار گرفت. بهترین روشی که توانست، ارزیابی و نتیجه‌گیری بهتری را در این مورد ارائه دهد، روش منطق بولین در مکان‌یابی پخش سیلاب بود. به دلیل توانایی شناخت دقیق‌تر و قدرت پیش‌بینی بالاتر، پیش‌بینی نواحی مستعد سیل با استفاده از روش‌های ترکیبی جدید برای مطالعات سیل بسیار مهم است. اگرچه در پژوهش‌های پیشین از الگوریتم‌های یادگیری ماشین برای

متزمکعب بر ثانیه می‌شوند. نتایج شبیه‌سازی سیلاب با دوره بازگشت ۲۵ ساله نشان می‌دهد که این سیلاب ماکریم جریانی است که مقاطع و سازه‌های هیدرولیکی رودخانه، قادر به عبور آن می‌باشند، بنابراین در راستای کاهش خطرات ناشی از سیلاب در شهر بیرجند، توجه بیش‌تر به افزایش پتانسیل عبور سیلاب رودخانه ضروری می‌باشد. Rahmati & Phourghasemi (2017) از سه مدل یادگیری ماشین تابع باور شواهد (EBF)^{۲۳}، جنگل تصادفی (RF)^{۲۴} و درخت رگرسیونی تقویت‌شده (BRT)^{۲۵} جهت پیش‌بینی مناطق حساس به سیل در استان گلستان استفاده کردند. نتایج نشان داد که هر سه مدل قابلیت بالایی در تجزیه و تحلیل ارتباط میان عوامل محیطی و رخدادهای سیل دارند. هم‌چنین نتایج ارزیابی دقت مدل‌ها نشان داد که مدل EBF دقت بالاتری را نسبت به دو مدل دیگر در پیش‌بینی سیل منطقه مورد مطالعه دارد. Siahkamari *et al.* (2018) در پژوهشی به بررسی کاربرد مدل‌های نسبت فراوانی (FR)^{۲۶} و بیشینه آنتروپی (ME)^{۲۷} در پیش‌بینی حساسیت سیل در حوزه آبریز مادرسو در استان گلستان پرداختند. منحنی مشخصه عملکرد سیستم (ROC)^{۲۸}، سطح زیر منحنی را برای مدل‌های FR و ME به ترتیب ۷۴/۳ و ۹۲/۶ درصد تخمین زد که نشان داد مدل ME منجر به نتایج بهتری برای ارزیابی پتانسیل وقوع سیل در منطقه مورد مطالعه شده است. Hong *et al.* (2018) از ترکیب روش‌های رگرسیون لجستیک (LR)^{۲۹}، جنگل تصادفی و ماشین بردار پشتیبان (SVM)^{۳۰} با روش وزن شواهد (WOE)^{۳۱} جهت تهیه نقشه حساسیت سیل در منطقه پویانگ^{۳۲} کشور چین استفاده کردند. نتایج نشان داد که مدل ترکیبی SVM-WOE بالاترین مقدار عملکرد پیش‌بینی (AUC)^{۳۳} را نسبت به سایر مدل‌های ترکیبی Razavi Temeh *et al.* (2018) از ترکیب مدل سیستم استنتاج عصی فازی

مدیریت آب و آسیاری

آن براساس هدف کاربردی است. مبانی تئوریک براساس مطالعات اسنادی و کتابخانه‌ای انجام گرفته است. تمامی پردازش‌های مربوط به داده‌ها در محیط‌های نرم‌افزاری Google earth engine و SAGA GIS 7.9، QGIS 3.1 و Matlab انجام گرفته است. طبق شکل (۱)، روند اجرایی این مطالعه شامل پنج مرحله اساسی می‌باشد:

- ۱- تهیه نقشه مرجع سیل (شناسایی ۴۲ موقعیت سیل خیز در حوضه آبریز منطقه مورد مطالعه).
- ۲- استخراج معیارهای مکانی مؤثر بر سیل در منطقه مورد مطالعه (۱۹ معیار هیدرولوژیکی، توپوگرافی، زمین‌شناسی و محیطی).
- ۳- انجام آزمایش هم خطی چندگانه جهت بررسی مستقل بودن معیارهای مکانی مؤثر بر سیل در منطقه مورد مطالعه.
- ۴- پیش‌بینی حساسیت سیل براساس رویکردهای یادگیری ماشین شامل مدل‌های SVM، RF، J48 و ANFIS.
- ۵- ارزیابی و مقایسه عملکرد رویکردهای یادگیری ماشین توسعه داده شده توسط شاخص‌های آماری مناسب.

پیش‌بینی نواحی مستعد سیل استفاده شده است، اما هیچ روشی در پیش‌بینی دقیق نواحی مستعد سیل کامل نیست. تهیه نقشه حساسیت به سیل از جمله روش‌های جدیدی می‌باشد که در سال‌های اخیر مورد توجه پژوهش‌گران قرار گرفته است. لذا نوآوری این پژوهش استفاده از چهار مدل ANFIS، RF، J48 و SVM مبتنی بر یادگیری ماشین یعنی الگوریتم‌ها می‌باشد که در نتایج پژوهش‌های اخیر کارایی این الگوریتم‌ها به اثبات رسیده است. لذا تصمیم بر آن شد در این پژوهش از الگوریتم‌های ذکر شده به‌طور همزمان جهت تعیین مدل دقیق‌تر در پیش‌بینی حساسیت به سیل در حوضه آبریز دشت بی‌جهد واقع در استان خراسان جنوبی استفاده و مشخص شود کدام یک از این روش‌ها توانایی بالاتری را در پیش‌بینی مناطق مستعد وقوع سیل در منطقه مورد مطالعه دارند. در نهایت، معیار سطح زیر منحنی (AUC) و سایر معیارهای آماری مناسب جهت اعتبارسنجی مدل‌های پیشنهادی در منطقه مورد مطالعه مورد استفاده قرار گرفتند.

مواد و روش‌ها

روش انجام پژوهش حاضر، توصیفی-تحلیلی بوده و نوع

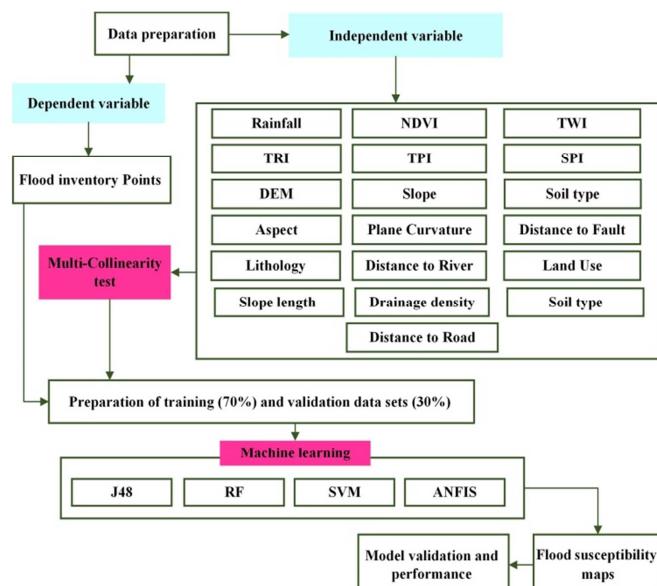


Figure 1. The flowchart of the research

مدیریت آب و آسیاری

دوره ۱۱ ■ شماره ۴ ■ زمستان ۱۴۰۰

منطقه مورد مطالعه

حداکثر ارتفاع حوضه از سطح دریا ۲۷۲۹ متر در ارتفاعات باقران و حداقل ارتفاع نیز ۱۱۷۲ متر در خروجی دشت (منطقه فدشک) می‌باشد. بیرجند به دلیل دوری از مسیر بادهای باران‌آور، عدم وجود ارتفاعات مهمی که بتواند در تعديل آب‌وهوا و جذب نزولات جوی مؤثر واقع شود و مجاورت با کویر لوت، در خشک‌ترین تقسیم‌بندی اقلیمی قرار دارد. میانگین درجه حرارت در همه نقاط شهرستان به یک اندازه نیست. هنگامی‌که شرایط اقلیمی شهر بیرجند در وضعیت وسیع‌تری و در قالب پنهانهای جغرافیایی گسترش‌تر مورد بررسی قرار می‌گیرد، مشخص می‌شود که این محدوده، بخشی از حوضه آبریز کویر مرکزی است و از نظر اقلیمی در محدوده جغرافیایی اقلیمی خشک و نیمه‌خشک واقع شده است. منشأ اصلی بارش بیرجند، متأثر از سیلیکون‌هایی است که به دو صورت تشکیل می‌شود؛ آن‌ها ممکن است در خارج از کشور تشکیل شده و رطوبت خود را تا شهر بیرجند حفظ کنند و یا در دامنه غربی کوه‌های زاگرس به وجود آمده و با کسب رطوبت از خلیج فارس و دریای عمان و با حرکت به سمت محدوده مطالعه بارش‌هایی را در شهر به وجود بیاورند.

استان خراسان جنوبی در یک اقلیم خشک واقع شده است و تراکم و تمرکز مناطق روستایی آن بیشتر در مناطق کوهستانی و مرکز استان واقع شده است. این استان در دو زیر‌حوضه درجه یک کشور به نام‌های فلاٹ مرکزی و مرزی شرقی کشور واقع است. همچنین براساس زیر‌حوضه‌های درجه دو در چهار زیر‌حوضه کویر مرکزی، نمکزار خوف، کویر لوت و هامون هیرمند قرار دارد. دشت بیرجند، یکی از محدوده‌های مطالعاتی حوضه آبریز کویر لوت می‌باشد که مختصات جغرافیایی آن بین طول‌های 58° و 59° تا 41° و 42° شرقی و عرض‌های 32° و 35° تا 33° و $8'$ شمالی قرار گرفته است (شکل ۲). مساحت کلی محدوده مطالعاتی $3406/72 \text{ km}^2$ است که از این مقدار سهم دشت و ارتفاعات به ترتیب $900/63 \text{ km}^2$ و $2506/09 \text{ km}^2$ می‌باشد. رودخانه شاهروд به طول 108 کیلومتر از شرق به غرب، دشت را زهکشی کرده و به کویر لوت می‌ریزد. این رودخانه در تمام سال به جز موضع سیلابی خشک است، لذا اکثر آب موردنیاز کشاورزی، صنعت و شرب شهرستان بیرجند از سفره آب زیرزمینی تأمین می‌شود (Eftekhari *et al.*, 2021a).

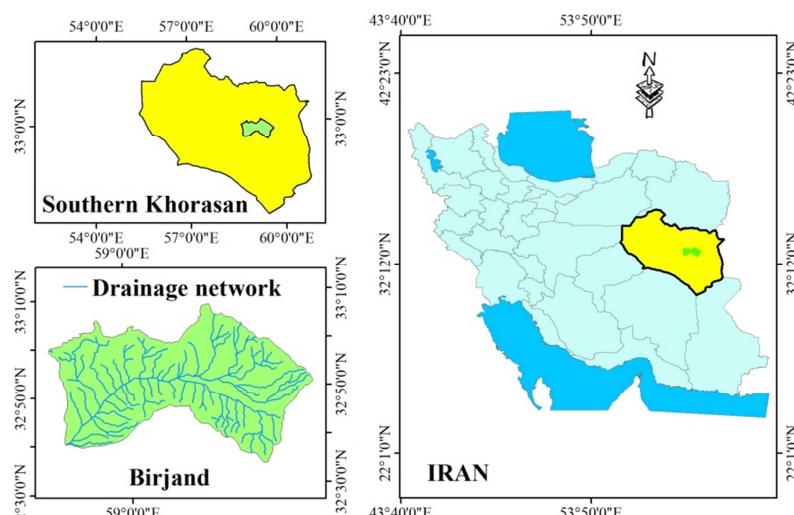


Figure 2. Geographical location of the study area

مدیریت آب و آسیاری

دوره ۱۱ ■ شماره ۴ ■ زمستان ۱۴۰۰

موقعیت سیل خیز) استفاده شدند. همچنین براساس نتایج پژوهش‌های قبلی و به منظور نزدیکتر شدن نتایج پژوهش به واقعیت، ۴۲ موقعیت غیرسیل (مقدار صفر) نیز با استفاده از Google earth بازدیدهای میدانی، نقشه توپوگرافی و نرم‌افزار بازدیدهای در مناطقی مانند تپه‌ها و کوه‌ها که سیلاب قادر به پیشروی در آنجا نیست، ایجاد شد (Chen *et al.*, 2018). این نسبت بین مجموعه داده‌های آموزشی و مجموعه داده‌های اعتبارسنجی توسط اکثر محققان مرکز بر ارزیابی حساسیت مخاطرات طبیعی توصیه شده است (Razavi Termeh *et al.*, 2018; Chen *et al.*, 2018; Nachappa *et al.*, 2021). به منظور صحبت‌سنجی نقاط سیل موجود در منطقه موردمطالعه، اقدام به مطالعه صحرا ای حوضه آبریز شد (شکل ۳).

معیارهای مؤثر بر وقوع سیل

براساس بررسی پژوهش‌های پیشین و در دسترس بودن داده‌ها، ۱۹ معیار هیدرورژئولوژیکی، توپوگرافی، زمین‌شناسی و محیطی شامل بارندگی، شاخص نرمال شده تفاوت پوشش گیاهی (NDVI)^{۳۳}، شاخص رطوبت توپوگرافی (TWI)^{۴۴}، شاخص توان آبراهه (SPI)^{۲۵}، شاخص نامهواری زمین (TRI)^{۳۶}، شاخص موقعیت توپوگرافی (TPI)^{۷۷}، مدل رقومی ارتفاعی، زاویه شیب، جریان تجمعی، جهت شیب، انحنای سطح، فاصله تا گسل، فاصله تا جاده، فاصله تا رودخانه، سنجش‌نامه، جنس خاک، کاربری اراضی، تراکم زهکشی و طول شیب برای پیش‌بینی حساسیت سیل در منطقه موردمطالعه استخراج شدند (Ahmadlou *et al.*, 2018; Chapi *et al.*, 2017; Chen *et al.*, 2019; Razavi Termeh *et al.*, 2018; Tien Bui *et al.*, 2016; Mojaddadi *et al.*, 2017). هرکدام از این معیارها در قالب نقشه رستری با اندازه ۳۰ متر تهیه شدند.

متوسط بارندگی منطقه به صورت سالیانه $133/5$ میلی متر می باشد. منطقه دارای پوشش گیاهی یکسانی نمی باشد و علت آن عوامل آب و هوایی، جنس خاک، ارتفاعات، اختلافات دما و وجود خرده اقلیم هایی در منطقه می باشد.

پوشش گیاهی شهرستان را به طور کلی به سه دسته اصلی پوشش گیاهی نیمه بیابانی و حاشیه کویری شامل گز، تاغ و درمنه، پوشش گیاهی نیمه بیابانی و کوهپایه ای شامل کنگره، گون، کاهو و جارو علفی و پوشش گیاهی کوهستانی شامل بنه، پسته بنه، بادام کوهی و انجیر کوهی تقسیم کرد.

بررسی ها نشان می دهد حدود 60 درصد دشت در محلوده شیب صفر تا 10 درصد قرار دارد که این شیب در دشت به طور عمده در قسمت های مرکزی قرار گرفته اند.

۱۰۵

مناطق تاریخی سیل

برای برآورد سیل های آینده در منطقه موردمطالعه، تجزیه و تحلیل رخداد گذشته سیلاب ها ضروری می باشد. نقاط سیل یک سطح مهم از رابطه بین رخداد سیل و معیارهای به وجود آورنده آن می باشد. رویدادهای تاریخی سیل به عنوان مبنایی برای پیش بینی وقوع سیل در آینده به حساب می آیند، به طوری که مناطق نزدیک به رخدادهای گذشته حساسیت بالایی به سیل گیری دارند (Rahmati *et al.*, 2016).

روش های مختلف مانند سوابق تاریخی (بانک اطلاعاتی از وقوع سیل های گذشته) و تفسیر عکس های ماهواره ای برای به دست آوردن یک نقشه مرجع سیل وجود دارد (Chen *et al.*, 2019). در مجموع، ۴۲ موقعیت سیل خیز (مقدار یک) در منطقه موردمطالعه ثبت شده است که مربوط به بانک اطلاعاتی شرکت آب منطقه ای استان خراسان جنوبی در سال ۱۳۹۹ می باشد (شکل ۳). در این مطالعه، به طور تصادفی ۷۰ درصد از این نقاط برای آموزش مدل ها (۳۰ موقعیت سیل خیز) و ۳۰ درصد مانع برای اعتبارسنجی مدل ها (۱۲

مدرسہ آب و آسارتی

پیش‌بینی مناطق مستعد وقوع سیل با استفاده از مدل‌های پیشرفته یادگیری ماشین (دشت بیرجند)

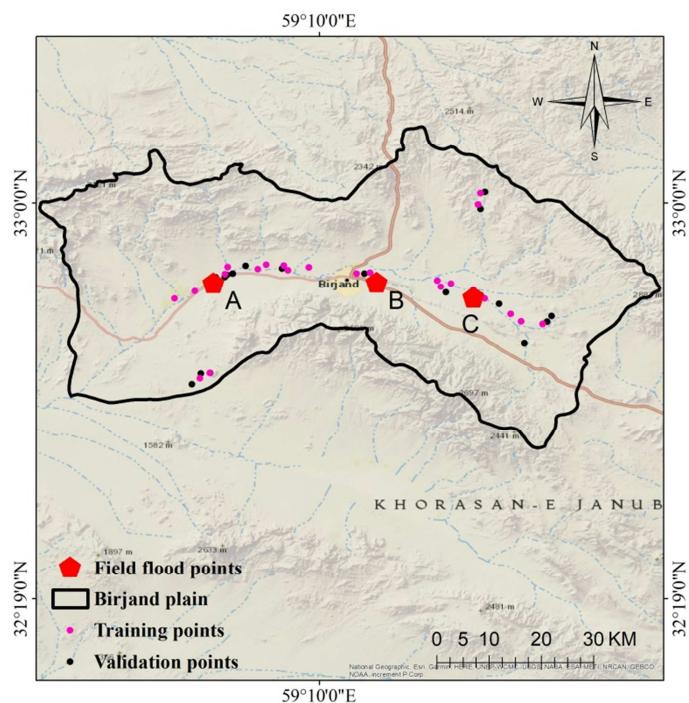


Figure 3. Parts of the floodplain areas in the study area

رابطه (۳) محاسبه می‌شود (Kanani-Sadat *et al.*, 2019):

$$TPI = Z_0 - \frac{\sum_{i=1}^n Z_i}{n} \quad (3)$$

که Z_0 ارتفاع سلول موردنظر و $\frac{\sum_{i=1}^n Z_i}{n}$ میانگین ارتفاع سلول‌های همسایه است. شاخص TRI، در واقع نشان‌دهنده اختلاف ارتفاع بین سلول‌های مجاور در یک شبکه ارتفاعی می‌باشد که طبق رابطه (۴) محاسبه می‌شود (Kalantari *et al.*, 2017):

$$TRI = \sqrt{Abs(max^2 - min^2)} \quad (4)$$

که max و min به ترتیب بزرگ‌ترین و کوچک‌ترین مقادیر ارتفاعی سلول‌های همسایه هستند. معیار طول شیب (LS) ^{۳۰} میکاری از ظرفیت انتقال رسوب برای جریان سطحی است که از رابطه (۵) محاسبه می‌شود (Khosravi *et al.*, 2016):

$$LS = \left(\frac{A_s}{22.13} \right)^{0.6} \left(\frac{\sin\beta}{0.0896} \right)^{0.06} \quad (5)$$

که A_s مساحت یک حوضه مشخص (m^2/m) و β زاویه شیب بر حسب درجه است. تراکم زهکشی نشان‌دهنده شرایط تغذیه آب‌های زیرزمینی است که برای Manap اندازه‌گیری نفوذ ناشی از آب‌های سطحی است (Kanani-Sadat *et al.*, 2014). نقشه تراکم زهکشی از روی نقشه DEM محاسبه شد. بافت خاک به عنوان یک مکانیسم مهم کنترل‌کننده نفوذ، تولید رواناب و در نتیجه رخداد خطر طبیعی شناخته شده است (Kanani-Sadat *et al.*, 2019). نقشه جنس خاک منطقه موردمطالعه در مقیاس ۱:۲۵۰۰۰۰ نقشه جنس نرمافزار Google earth engine اخذ شد. از طریق نرمافزار سنگ‌شناسی متغروت، اختلاف هم‌چنین واحدهای سنگ‌شناسی متفاوت، اختلاف چشم‌گیری در ناپایداری دامنه نشان می‌دهد. نقشه سنگ‌شناسی منطقه موردمطالعه با استفاده از نقشه زمین‌شناسی کشور ایران در مقیاس ۱:۱۰۰۰۰۰ (۳۹ سنگ‌شناسی) تهیه شد.

نقشه توپوگرافی منطقه موردمطالعه در مقیاس ۱:۵۰۰۰۰ (منبع: سازمان نقشه‌برداری کشور ایران) تهیه شد. به منظور تهیه لایه DEM از تصاویر ماهواره‌ای SRTM ^{۲۸} با قدرت تفکیک ۳۰ متر استفاده شد و سپس لایه‌های زاویه شیب، جهت شیب و انحنای سطح از لایه DEM ایجاد شدند. لایه‌های فاصله تا رودخانه و جاده با استفاده از نقشه‌های توپوگرافی ۱:۵۰۰۰۰ تهیه شد. هم‌چنین، لایه فاصله تا گسل نیز با استفاده از نقشه‌های توپوگرافی ۱:۱۰۰۰۰ تهیه و استخراج شد. لایه کاربری اراضی منطقه موردمطالعه نیز از طریق تصاویر ماهواره‌ای لندست هشت سنجنده OLI ^{۲۹} استخراج گردید. هم‌چنین، از باند چهار و پنج تصاویر ماهواره‌ای لندست هشت جهت تولید لایه NDVI استفاده شد. هرچه مقدار NDVI کم‌تر باشد، حساسیت بیشتر سیل را به همراه دارد. بارندگی به عنوان یکی از معیارهای محرک سیل به حساب می‌آید. لذا نقشه میانگین بارندگی با استفاده از شش ایستگاه سینوپتیک در منطقه موردمطالعه محاسبه شد. هم‌چنین شاخص TWI در تعیین میزان اثر توپوگرافی بر محل، اندازه منطقه اشباع‌شده و تولید رواناب کاربرد زیادی دارد که طبق رابطه (۱) محاسبه می‌شود (Jancewicz *et al.*, 2019):

$$TWI = \ln\left(\frac{A_s}{Tan\beta}\right) \quad (1)$$

که A_s مساحت یک حوضه مشخص (m^2/m) و β زاویه شیب بر حسب درجه است. شاخص SPI نیز بیان‌گر توان فرسایندگی جریان بوده که با درجه شیب و مساحت حوضه آبریز رابطه مستقیم دارد و مطابق رابطه (۲) محاسبه می‌شود (Hong *et al.*, 2018):

$$SPI = A_s \times Tan\beta \quad (2)$$

که A_s و β به ترتیب نشان‌دهنده مساحت یک حوضه مشخص (m^2/m) و زاویه شیب بر حسب درجه است. شاخص TPI نیز اختلاف ارتفاع هر سلول با میانگین ارتفاعات سلول‌های همسایه آن است که طبق

مدیریت آب و آسیاری

روش‌ها

ماشین بردار پشتیبانی (SVM)

ماشین بردار پشتیبانی (SVM)، یک روش یادگیری ماشین نسبتاً جدید و یک الگوریتم یادگیری ماشین ناظارت شده است (Arabgol *et al.*, 2016). SVM یکی از مقاعد کننده‌ترین روش‌های پیش‌بینی است که براساس روش حداقل سازی ریسک ساختاری می‌باشد. در مقابل، بیشتر مدل‌های هوش مصنوعی مانند شبکه‌های عصبی مصنوعی، از تکنیک‌های به حداقل رساندن ریسک تجربی استفاده می‌کنند. بنابراین، روش SVM می‌تواند خطای تجربی را کاهش دهد، پیچیدگی را مدل کند و احتمال را بیش از حد تنظیم کند (Ghorbanzadeh *et al.*, 2019). هدف SVM پیدا کردن ابر صفحه جداساز بهینه است که بتواند حاشیه را بین کلاس‌های مختلف مشخص کرده و فاصله یک کلاس را به حداقل برساند. در بیشتر شرایط، ابر صفحه توسط یک سطح غیر خطی تعریف خواهد شد (شکل ۴). در این مورد، عبارت ریاضیاتی زیر برای طبقه‌بندی مجموعه داده‌ها به کار گرفته خواهد شد (Arabgol *et al.*, 2016):

$$f(x) = \sum_{i=1}^n (a_i - a_i^*) K(x_i, x) + b \quad (4)$$

که در آن a_i و a_i^* ضرایب لایکرانز، K تابع کرنل و b انحراف ابرصفحه از مبدأ است.

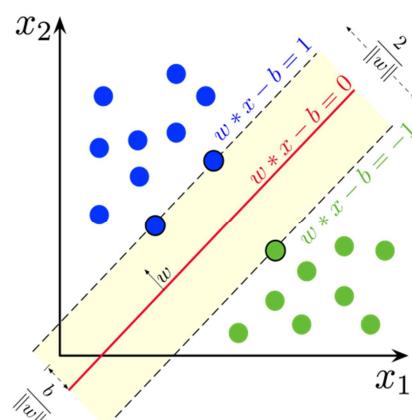


Figure 4. SVM algorithm process

مدیریت آب و آسیاری

الگوریتم درخت تصمیم 48، که شامل گره ریشه، گره‌های داخلی و گره‌های برگ است، در پژوهش حاضر برای ایجاد یک طبقه‌بندی دودویی (پیکسل‌های سیل و غیرسیل) استفاده خواهد شد (Tien Bui *et al.*, 2012). گره ریشه شامل تمام داده‌های ورودی خواهد بود، گره‌های داخلی کاربرد تابع تصمیم را تضمین خواهند کرد، در حالی که گره‌های برگ با داده‌های خروجی مرتبط هستند (Pham *et al.*, 2017). فرایند آموزش مدل 48 در دو مرحله اصلی انجام می‌شود: مرحله اول) ساخت درخت طبقه‌بندی؛ مرحله دوم) هرس درخت طبقه‌بندی. در مرحله اول، در گره ریشه، داده‌های ورودی دارای نسبت افزایش با بالاترین مقدار تعیین خواهد شد. براساس مقادیر گره ریشه، روش تقسیم در مورد مجموعه داده آموزشی برای ایجاد زیرگره‌ها اجرا خواهد شد. مرحله دوم شامل تولید نسبت بهره فردی برای هر زیرگره است. در نهایت طبقه‌بندی پیش‌بینی پیکسل‌های سیل و غیرسیل با درنظر گرفتن نسبت افزایش مربوط به هر زیرگره انجام خواهد شد (Pham *et al.*, 2017). فرایند الگوریتم درخت تصمیم J48 در شکل (۵) نشان داده شده است.

سیستم استنباط عصبی فازی (ANFIS)

ANFIS یک الگوریتم ترکیبی است که توسط Jang (1993) پیشنهاد شده و از طریق شبکه عصبی مصنوعی و منطق فازی تولید می‌شود. ANFIS یک الگوریتم داده محور می‌باشد که با یکتابع یادگیری خودکار مشخص می‌شود و می‌تواند بدون شرایط ورودی اجرا شود (Gao *et al.*, 2019). قانون اگر-آنگاه فازی روشی را برای توصیف یک مسئله خاص در اصطلاحات زبانی نشان می‌دهد. قوانین اگر-آنگاه شکل زیر را دارند و متعلق به تاکاگی و سوگنو هستند (Ghorbanzadeh *et al.*, 2020).

(et al., 2018; Quiroz et al., 2018). مدل پیش‌بینی‌کننده RF براساس میانگین‌گیری از نتایج حاصل از تمامی درخت‌های تصمیم مربوطه استوار است و برای بسیاری از مجموعه داده‌ها، طبقه‌بندی را با صحت بالایی انجام می‌دهد (de Santana et al., 2018). چهار مرحله زیر فرایند RF را مشخص می‌کنند؛ ۱- تعریف و بازنمونه‌گیری داده‌های آموزشی، ۲- انتخاب مجموعه ویژگی‌های تصادفی مربوط به هر نمونه مجدد، ۳- اختصاص یک درخت تصمیم برای هر کدام از آن‌ها به مجموعه ویژگی‌های تصادفی و گستردۀ، ۴- ایجاد یک درخت تصمیم واحد از طریق تجمعیع درخت تصمیم اختصاص‌داده شده به هر مثال. فرایند الگوریتم RF در شکل (۷) نشان داده شده است.

Rule1: if x is A₁ and y is B₁, then f₁ (۷)
 $= p_1x + q_1y + r_1$
Rule1: if x is A₂ and y is B₂, then f₂
 $= p_2x + q_2y + r_2$
 که در آن A₁, A₂, B₁, B₂ توابع عضویت برای ورودی‌های x و y هستند و p_i , q_i , r_i (i, j=1,2) پارامترهای برآیند هستند. فرایند الگوریتم ANFIS در شکل (۶) نشان داده شده است.

جنگل تصادفی (RF)

الگوریتم جنگل تصادفی یکی از رایج‌ترین الگوریتم‌های به کار رفته برای بررسی مشکلات طبقه‌بندی و پیش‌بینی چندگانه است که حساسیت کمی به چندخطی‌بودن دارد و نتایج آن از نظر داده‌های از de Santana دست‌رفته و نامتعادل نسبتاً پایدار است

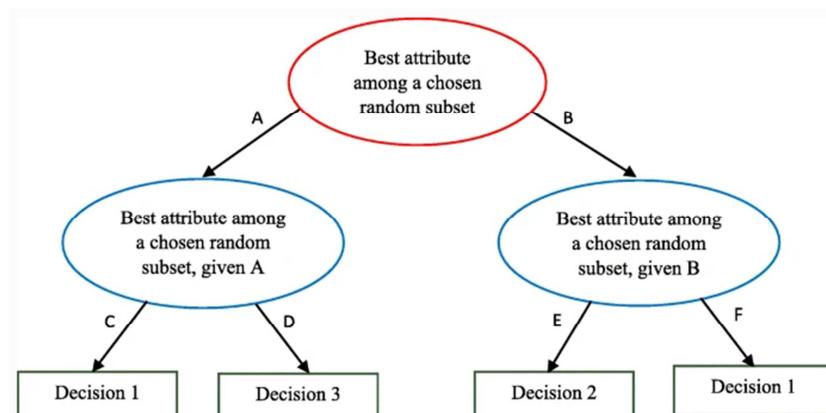


Figure 5. J48 decision tree algorithm process

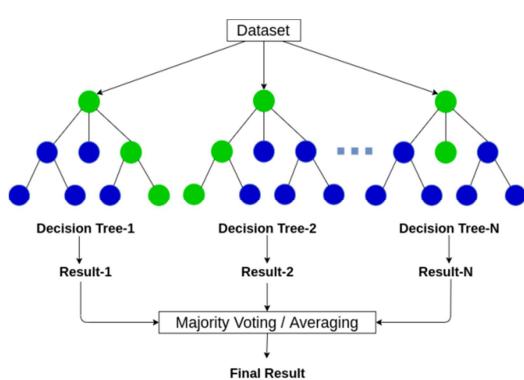


Figure 7. RF algorithm process

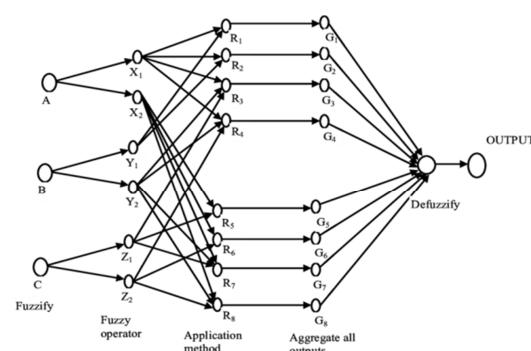


Figure 6. ANFIS algorithm process

قبلی و بهمنظور نزدیکتر شدن نتایج پژوهش به واقعیت، ۴۲ موقعیت غیرسیل (مقدار صفر) نیز با استفاده از بازدیدهای میدانی، نقشه توپوگرافی و نرمافزار Google earth در مناطقی مانند تپه‌ها و کوهها که سیلاب قادر به پیشروی در آنجا نیست، ایجاد شد. خروجی هرکدام از روش‌های یادگیری ماشین دارای مقداری مربوط به حساسیت سیل برای این نقاط می‌باشد که جهت محاسبه مقادیر TP، FP، TN و FN در نظر گرفته می‌شوند.

ارزیابی عملکرد و دقت مدل‌ها

معیارهای آماری

عملکرد و دقت پیش‌بینی مدل‌های ترکیبی توسعه‌یافته باید با استفاده از معیارهای آماری مختلف ارزیابی شود. در این پژوهش مجموعه‌ای از معیارهای آماری شامل نرخ پیش‌بینی مثبت (PPR)^{۳۱}، نرخ پیش‌بینی منفی (NPR)^{۳۲}، حساسیت (SST)^{۳۳}، ویژگی (SPC)^{۳۴} و دقت (ACC)^{۳۵} برای ارزیابی و Eslaminezhad *et al.* (2021a; Khosravi *et al.*, 2016) اعتبارسنجی تمام مدل‌ها استفاده شدند. همه این معیارها براساس ماتریس درهم‌ریختنگی^{۳۶} محاسبه می‌شوند که متشکل از چهار نوع پیامد احتمالی شامل مثبت واقعی (TP)، منفی واقعی (TN)^{۳۸}، مثبت کاذب (FP)^{۳۹} و منفی کاذب (FN)^{۴۰} برای محاسبه معیارهای آماری است. معیار TP تعداد مکان‌هایی است که به درستی به عنوان یک مکان با پتانسیل بالای سیل طبقه‌بندی می‌شوند. در حالی که معیار FP تعداد مکان‌هایی است که به طور نادرست به عنوان یک مکان با پتانسیل بالای سیل طبقه‌بندی می‌شوند. علاوه بر این، معیارهای TN و FN تعداد مکان‌هایی هستند که به طور صحیح و نادرست به عنوان یک مکان بدون پتانسیل بالای سیل طبقه‌بندی می‌شوند. این معیارهای آماری با استفاده از Chapi *et al.*, 2017; Khosravi (2016) روابط زیر محاسبه می‌شوند:

: (Chapi *et al.*, 2017; Khosravi (2016)

$$PPR = \frac{TP}{TP + FP} \quad (8)$$

$$NPR = \frac{TN}{TN + FN} \quad (9)$$

$$SST = \frac{TP}{TP + FN} \quad (10)$$

$$SPC = \frac{TN}{TN + FP} \quad (11)$$

$$ACC = \frac{TP + TN}{TP + FP + TN + FN} \quad (12)$$

همان‌طور که ذکر شد، در مجموع، ۴۲ موقعیت سیل خیز (مقدار یک) در منطقه مورد مطالعه ثبت شده است که مربوط به بانک اطلاعاتی شرکت آب منطقه‌ای استان خراسان جنوبی در سال ۱۳۹۹ می‌باشد. هم‌چنین براساس نتایج پژوهش‌های

نتایج آماده‌سازی داده‌ها

همان‌طور که ذکر شد، در این مطالعه از ۱۹ معیار هیدرولوژیکی، توپوگرافی، زمین‌شناسی و محیطی شامل بارندگی، شاخص نرمال‌شده تفاوت پوشش گیاهی، شاخص

مدیریت آب و آسیاری

حساسیت سیل در حوضه آبریز دشت بیرجند استفاده شد. برای تولید نقشه هر کدام از معیارها (شکل ۸)، از درون یابی Eftekhari *et al.*, 2021b; (Eslaminezhad *et al.*, 2021b).

روطوبت توپوگرافی، مدل رقومی ارتفاعی، زاویه شبیب، حریان تجمعی، جهت شبیب، انحنای سطح، فاصله تا جاده، فاصله تا گسل، فاصله تا رودخانه، کاربری اراضی، جنس خاک، سنگ‌شناسی، تراکم زهکشی و طول شبیب پیش‌بینی

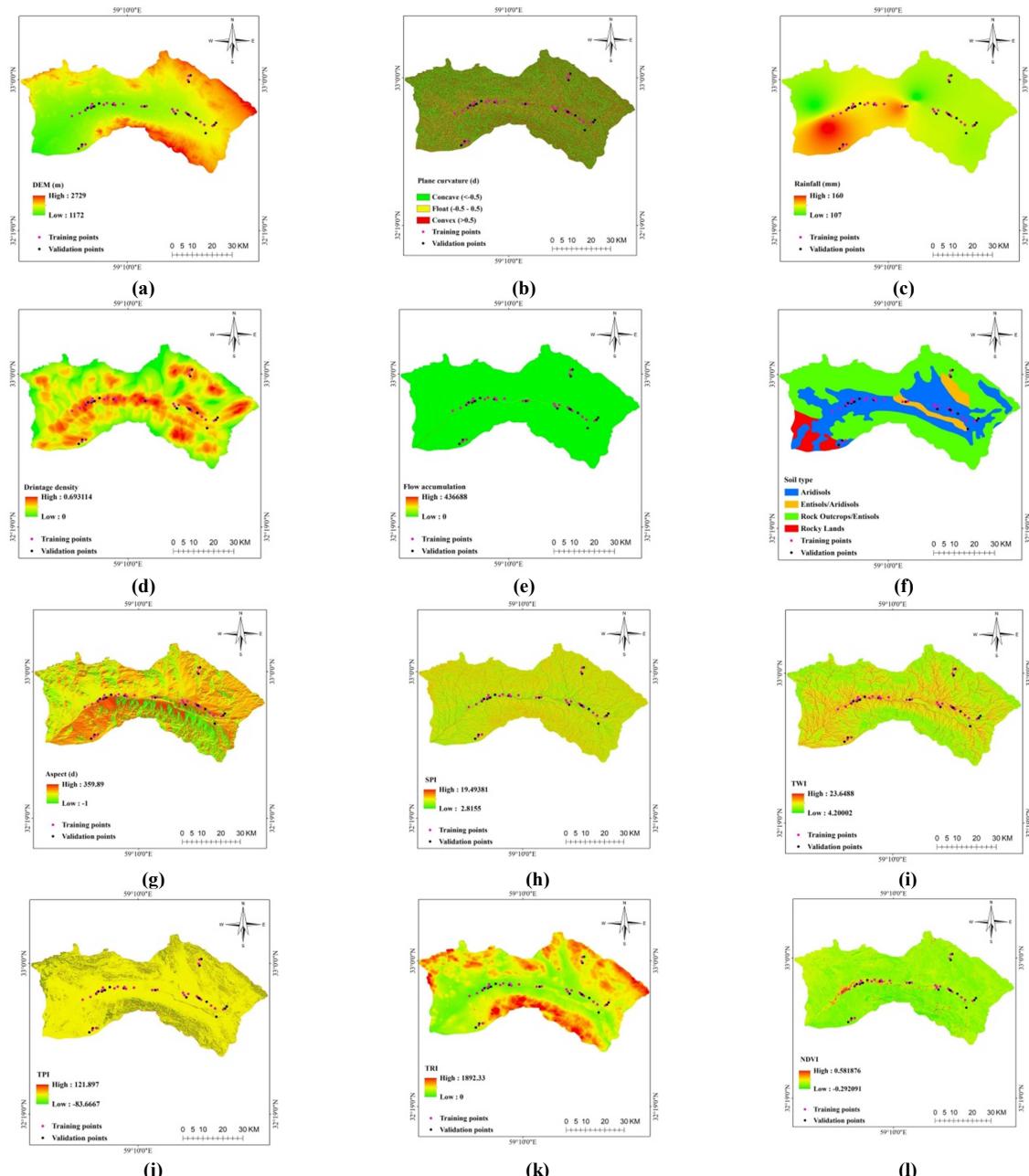
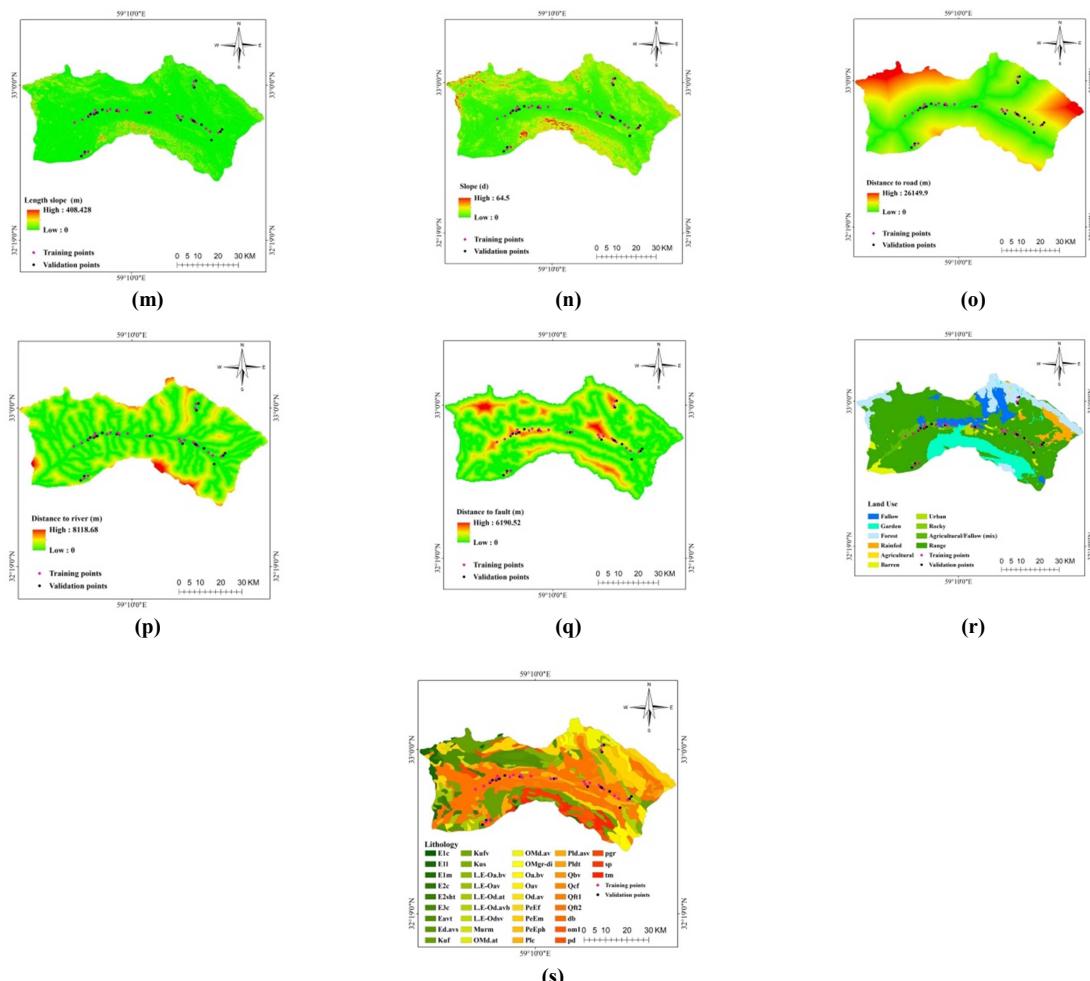


Figure 8. The maps of spatial criteria affecting the flood The maps of spatial criteria affecting the flood (a) DEM (b) Plane curvature (c) Rainfall (d) Drainage density (e) Flow accumulation (f) Soil type (g) Aspect (h) SPI (i) TWI (j) TPI (k) TRI (l) NDVI (m) Length slope (n) Slope (o) Distance to road (p) Distance to river (q) Distance to fault (r) Land Use (s) Lithology

پیش‌بینی مناطق مستعد وقوع سیل با استفاده از مدل‌های پیشرفته یادگیری ماشین (دشت بیرجند)



Continued figure 8. The maps of spatial criteria affecting the flood The maps of spatial criteria affecting the flood (a) DEM (b) Plane curvature (c) Rainfall (d) Drainage density (e) Flow accumulation (f) Soil type (g) Aspect (h) SPI (i) TWI (j) TPI (k) TRI (l) NDVI (m) Length slope (n) Slope (o) Distance to road (p) Distance to river (q) Distance to fault (r) Land Use (s) Lithology

نمی‌توان به عنوان ورودی در مدل‌سازی استفاده کرد و در نتیجه از بین معیارهای این پژوهش حذف می‌شود. به جز معیارهای LS و TPI، هیچ هم‌خطی چندگانه بالایی در میان سایر معیارهای انتخاب شده مشاهده نشد.

عملکرد مدل‌های حساسیت سیل

شکل (۹) قابلیت‌های صحت‌سنگی و پیش‌بینی مدل‌های به کار گرفته شده برای پیش‌بینی حساسیت سیل با استفاده از مجموعه داده‌های آموزشی و داده‌های اعتبارسنجی را نشان می‌دهد.

انتخاب معیارهای مستقل مؤثر بر سیل با استفاده از آزمایش هم‌خطی چندگانه
تجزیه و تحلیل همبستگی نشان می‌دهد که یک رابطه بین دو یا چند متغیر ورودی ممکن است باعث انحراف شود. بهمنظور تشخیص هم‌خطی چندگانه در میان عوامل مختلف، آماره تحمل (TOL)^۱ و عامل تورم واریانس (VIF)^۲ دو پارامتر آماری رایج هستند (Choubin *et al.*, 2019; Eftekhari *et al.*, 2021c). براساس جدول (۱)، نتایج نشان داد که معیارهای LS و TPI شرط هم‌خطی چندگانه را رد کرده‌اند، لذا این معیارها را ($TOL > 0.1$ and $VIF < 5$)

آموزشی و داده‌های اعتبارسنجی مدل‌های J48، SVM و RF و ANFIS نشان می‌دهد. طبق شکل (۱۰)، برای مجموعه داده‌های آموزشی، مدل RF دارای بالاترین مقدار J48 (۰/۹۷۲) AUC می‌باشد و پس از آن مدل‌های SVM (AUC=۰/۹۴۲) و ANFIS (AUC=۰/۹۵۷) مجموعه داده‌های آموزشی، مدل RF دارای بالاترین مقدار J48 (۰/۹۷۲) AUC می‌باشد و پس از آن مدل‌های SVM (AUC=۰/۹۴۲) و ANFIS (AUC=۰/۹۵۷) مجموعه داده‌های اعتبارسنجی، مدل RF قوی‌تر بوده و دارای دقت بالاتری (AUC=۰/۹۵۸) نسبت به سایر مدل‌ها می‌باشد. بنابراین، می‌توان نتیجه گرفت که اگرچه تمام مدل‌های به کاررفته قدرت پیش‌بینی خوبی دارند، اما عملکرد و دقت مدل RF برای پیش‌بینی حساسیت سیل در منطقه موردمطالعه بهتر است که این نتایج با یافته‌های پژوهش (2016) Hong *et al.* و Khosravi *et al.* (2018) مبنی برتری روش RF مطابقت دارد.

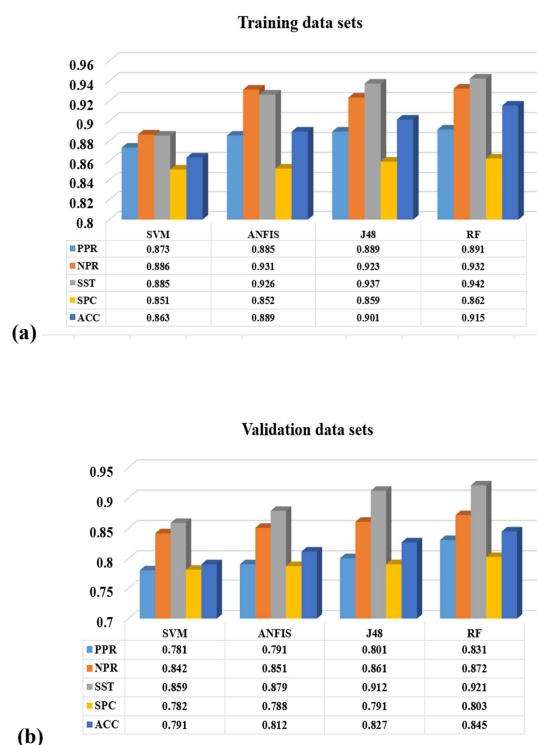


Figure 9. Performance of models using (a) training and (b) validation data sets

Order	Criteria	TOL	VIF
1	Rainfall	0.78	2.25
2	Distance to road	0.23	1.95
3	NDVI	0.78	2.65
4	LS	0.095	5.82
5	Distance to fault	0.45	3.23
6	DEM	0.64	2.85
7	Plane curvature	0.11	3.23
8	Flow Accumulation	0.19	2.27
9	Aspect	0.95	1.58
10	Slope	0.85	2.39
11	Lithology	0.75	4.73
12	SPI	0.25	4.23
13	TRI	0.85	1.99
14	Distance to river	0.19	2.85
15	Drainage density	0.49	3.85
16	TWI	0.55	2.78
17	TPI	0.025	6.23
18	Soil type	0.44	2.75
19	Land use	0.55	1.09

برای طبقه‌بندی پیکسل‌های سیل، روش RF بیشترین مقدار SST را برای مجموعه داده‌های آموزشی و اعتبارسنجی نشان می‌دهد (۰/۹۴۲ و ۰/۹۲۱). همچنین برای طبقه‌بندی پیکسل‌های غیرسیلی، روش RF بهترین مقدار SPC را نیز برای مجموعه داده‌های آموزشی و اعتبارسنجی نشان می‌دهد (۰/۸۶۲ و ۰/۸۰۳). علاوه بر این، مدل RF دارای بالاترین مقدار نرخ پیش‌بینی مشت، نرخ پیش‌بینی منفی و دقت برای مجموعه داده‌های آموزشی و اعتبارسنجی است که پس از آن مدل‌های J48 و SVM و ANFIS قرار گرفته‌اند. نتایج مجموعه داده‌های اعتبارسنجی نشان داد که مدل RF قابلیت بهتری را نسبت به مدل‌های J48 و SVM در پیش‌بینی حساسیت سیل در منطقه موردمطالعه دارد. عملکرد بهتر مدل RF می‌تواند به دلیل توانایی آن در مدل‌سازی پایگاه داده‌های بزرگ و توانایی ادغام متغیرهای ورودی زیاد بدون تغییر متغیر باشد (Rahmati *et al.*, 2016) و (Chen *et al.* 2018) Zhao *et al.* (2020) مطالعات نشان می‌دهند که مدل RF توانایی بسیاری زیادی در تهیی نقشه مناطق حساس به سیل دارد.

شکل (۱۰) منحنی ROC را برای مجموعه داده‌های

پیش‌بینی مناطق مستعد وقوع سیل با استفاده از مدل‌های پیشرفته یادگیری ماشین (دشت بیرجند)

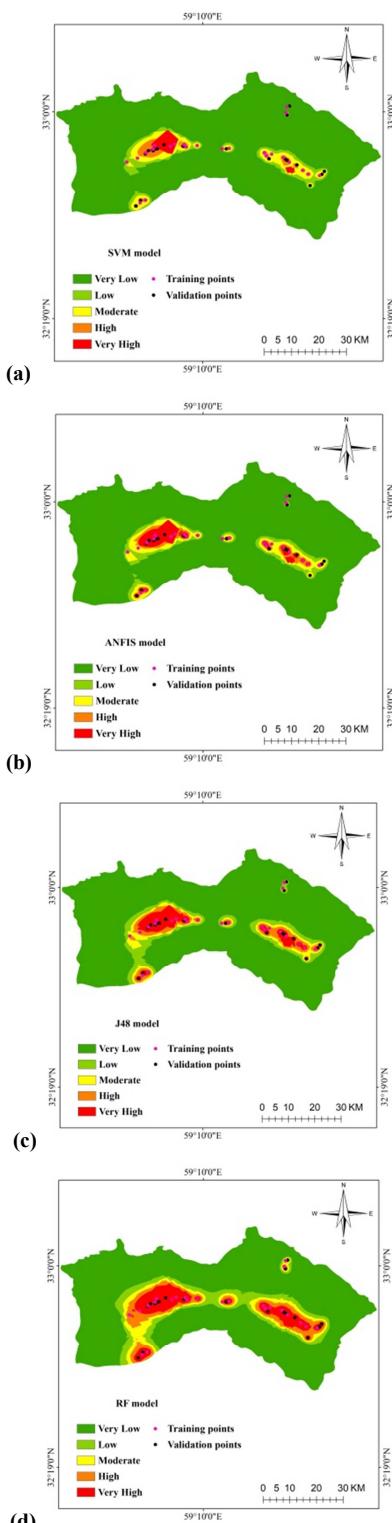


Figure 11. Flood susceptibility prediction maps based on machine learning models (a) SVM (b) ANFIS (c) J48 (d) RF

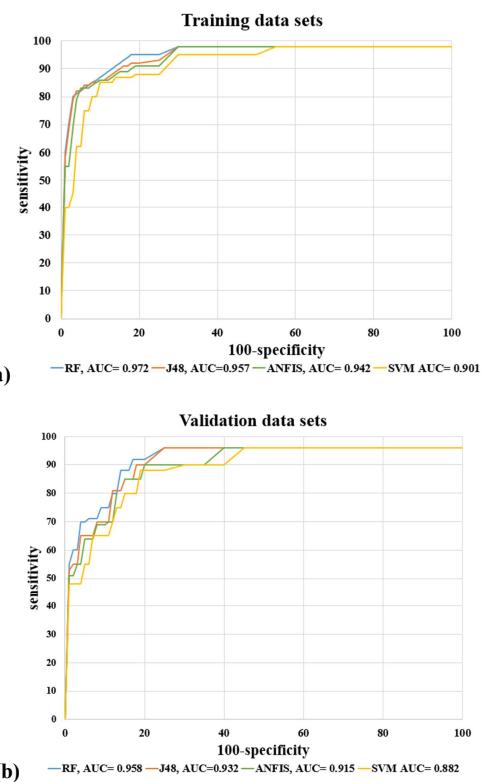


Figure 10. ROC curve and AUC values for (a) training data sets and (b) validation data sets

تئیه نقشه پیش‌بینی حساسیت سیل

نقشه میزان حساسیت سیل پیش‌بینی شده با استفاده از روش طبقه‌بندی شکست طبیعی^{۴۳} به کلاس‌های مختلف در نرم‌افزار ArcGIS طبقه‌بندی شد. بر این اساس، میزان حساسیت سیل پیش‌بینی شده، به پنج کلاس شامل حساسیت خیلی کم، کم، متوسط، زیاد و خیلی زیاد طبقه‌بندی شد. شکل (۱۱) نقشه‌های میزان حساسیت سیل در منطقه موردمطالعه را نشان می‌دهد که با استفاده از مدل‌های موردنظر پیش‌بینی شده‌اند.

نقشه حساسیت سیل منطقه موردمطالعه نشان می‌دهد که احتمال وقوع سیل در مناطق مرکزی منطقه موردمطالعه، به دلیل ارتفاع و زاویه شبکه کمتر، بیشتر از سایر مناطق است که مطابق با نتایج پژوهش‌های Kanani-Sadat *et al.* (2019) Chen *et al.* (2019) و (2019) است.

ایران، ارزیابی حساسیت سیل در مناطق آسیب‌پذیر کشور اقدامی است که نمی‌توان آن را حذف و به تعویق انداخت. الگوریتم‌های یادگیری ماشین، ارزیابی بسیار دقیقی از حساسیت سیل ارائه می‌دهند و بنابراین، استفاده از مدل‌های مدرن و پیشرفته ضروری است. لذا در این مطالعه، چهار مدل یادگیری ماشین شامل RF J48، SVM و ANFIS برای دستیابی به نقشه‌های دقیق حساسیت سیل در حوضه آبریز دشت بیرجند آموزش داده شدند و در نهایت با استفاده از یک سری شاخص‌های آماری و مقادیر AUC مورد ارزیابی و مقایسه قرار گرفتند. لذا جهت پیاده‌سازی، موقعیت ۴۲ موقعیت سیل خیز در منطقه موردمطالعه و ۱۹ معیار هیدرولوژیکی، توپوگرافی، زمین‌شناسی و محیطی استخراج شدند. نتایج نهایی نشان داد که مدل RF نسبت به سه مدل دیگر برای پیش‌بینی حساسیت سیل عملکرد بهتری دارد و قابلیت برتر مدل RF توسط پژوهش‌های مرتبط دیگر تأیید شده‌است (Choubin *et al.*, 2019; Zhao *et al.*, 2018).

با مدل‌های SVM، J48 و ANFIS یک الگوریتم ترکیبی است که می‌تواند نتایج رضایت‌بخش‌تری با نرخ‌های دقیق بالاتر و واریانس‌ها و سوگیری‌های کمتر ایجاد کند (Chen *et al.*, 2019; Hong *et al.*, 2018). هر چهار مدل مورداستفاده در این مطالعه می‌توانند نتایج قابل قبولی در پیش‌بینی حساسیت سیل ایجاد کنند. برای معیار فاصله، در کل با افزایش فاصله از رودخانه احتمال وقوع سیل کمتر می‌شود. انتخاب ارتفاع از سطح دریا به عنوان یکی از مهم‌ترین معیارهای مؤثر در حساسیت سیل است که سیل در ارتفاعات پایین‌تر از سطح دریا اتفاق می‌افتد و در ارتفاعات بالا احتمال وقوع سیل بسیار کم است. با بررسی شاخص بارش می‌توان به این نتیجه رسید که با افزایش ارتفاع، احتمال بارندگی بیش‌تر می‌گردد و از طرف دیگر احتمال وقوع سیل در ارتفاعات بالاتر کمتر است. کاربری

هم‌چنین، در مناطق مرکزی منطقه موردمطالعه به‌دلیل کاهش سطح اراضی جنگلی و تغییر کاربری‌های شدید صورت گرفته در چند سال اخیر که سبب کاهش اراضی جنگلی و افزایش اراضی دیم و شهری شده، احتمال وقوع خسارت‌های را افزایش داده است. هم‌چنین، در شکل (۱۲)، درصد هر یک از کلاس‌های حساسیت سیل پیش‌بینی شده توسط مدل‌های RF، SVM و ANFIS نشان داده شده است. نتایج نشان داد که در مدل RF، چهار کلاس با حساسیت خیلی زیاد، حساسیت زیاد، متوسط و کم درصد بیش‌تری را در منطقه موردمطالعه نسبت به کلاس‌های مشابه در مدل‌های SVM و J48 تحت پوشش قرار می‌دهد. علاوه بر این در مدل SVM، کلاس با حساسیت خیلی کم، درصد بیش‌تری را در منطقه موردمطالعه نسبت به کلاس‌های RF، RF-BPSO J48 و ANFIS دربر گرفته است در واقع می‌توان نتیجه گرفت که به ترتیب ۷۵/۷۵ درصد و ۷/۶۲ درصد از منطقه موردمطالعه در کلاس با حساسیت سیل خیلی کم و حساسیت خیلی زیاد قرار گرفته است.

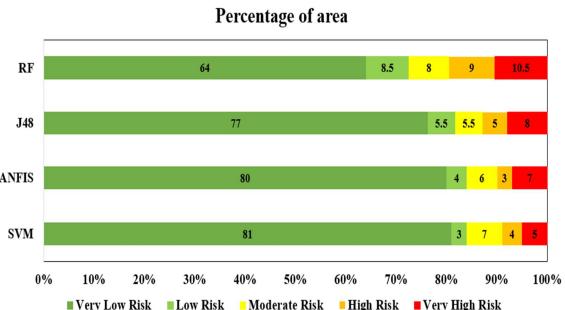


Figure 12. Percentage of each flood susceptible class in the study area

نتیجه‌گیری

برنامه‌ریزی و مدیریت خطر سیل نیازمند دانش مناسب و کافی از نواحی در معرض وقایع مهم سیل است. با توجه به وقوع سیلاب‌های شدید در سال‌های اخیر در کشور

برآورد آسیب‌پذیری سیل در مناطق پژوهشی مختلف استفاده کنند.

اراضی نیز از دیگر معیارهای مؤثر بر وقوع سیل می‌باشد، زیرا نوع کاربری اراضی بر روی نفوذ رواناب و یا تشدید سرعت آن مؤثر است. بهدلیل تغییر معیار کاربری اراضی توسط فعالیت‌های انسانی، توان آبراهه نیز کم شده و در نتیجه به عنوان معیار مهمی در وقوع سیل دخالت دارد. این نتایج با نتایج پژوهش‌های Khosravi *et al.*, Razavi Termeh *et al.*, Rahmati *et al.*, Liu *et al.* (2016) و (2018) مطابقت دارد. با توجه به این‌که روش‌های ذکر شده داده‌محور می‌باشند و در این مطالعه در حوضه آبریز دشت بیرجند مورداستفاده قرار گرفته است، بنابراین پیشنهاد می‌شود این الگوریتم‌ها برای سایر مناطق با ویژگی‌های جغرافیایی مشابه و متفاوت نیز به کار گرفته شوند و نتایج مربوط به دقت و ارزیابی روش‌ها در پیش‌بینی حساسیت سیل مشخص شود و بررسی شود تا چه اندازه‌ای نوع داده‌ها و منطقه موردمطالعه می‌تواند روی دقت و خروجی مدل‌ها اثر بگذارد. همچنین با توجه به تعداد بالای معیارهای به کار گرفته شده در این پژوهش، ممکن است برخی از معیارها تأثیر کمتری نسبت به مجموعه معیارهای دیگر داشته باشند و این مسئله بر روی دقت نهایی اثرگذار باشد. بنابراین پیشنهاد می‌شود در پژوهش‌های آینده از الگوریتم‌های فرالبتکاری مانند الگوریتم زنتیک، بهینه‌سازی انبوه ذرات و ... در ترکیب با الگوریتم‌های یادگیری ماشین استفاده شود تا بهترین ترکیب معیارها براساس بهینه‌سازی مقدار تابع برآش و افزایش دقت مدل‌ها تعیین شوند. تجزیه و تحلیل نقشه‌های پیش‌بینی حساسیت سیل با استفاده از روش‌های نوین یادگیری ماشین می‌تواند به برنامه‌ریزان و مدیران برای جلوگیری و کاهش خسارت‌های ناشی از سیل در مناطق مستعد سیل کمک کنند. الگوریتم‌هایی با دقت بسیار بالا همچنین می‌توانند تصمیم‌گیری برای پژوهش‌گران دیگر باشند که از آن‌ها برای

تعارض منافع

هیچ‌گونه تعارض منافع توسط نویسنده‌گان وجود ندارد.

منابع

1. Ahmadlou, M., Karimi, M., Alizadeh, S., Shirzadi, A., Parvinnejhad, D., Shahabi, H., & Panahi, M. (2018). Flood susceptibility assessment using integration of Adaptive Network-Based Fuzzy Inference System (ANFIS) and Biogeography-Based Optimization (BBO) and BAT Algorithms (BA). *Geocarto International*, 34(11), 1252-1272.
2. Alam, Z., Zhang, C., & Samali, B. (2020). Influence of seismic incident angle on response uncertainty and structural performance of tall asymmetric structure. *The Structural Design of Tall and Special Buildings*, 29(12), 1750.
3. Arabameri, A., Rezaei, K., Cerdá, A., Conoscenti, C., & Kalantari, Z. (2019). A comparison of statistical methods and multi-criteria decision making to map flood hazard susceptibility in Northern Iran. *Science of the Total Environment*, 660, 443-458.
4. Arabgol, R., Sartaj, M., & Asghari, K. (2016). Predicting nitrate concentration and its spatial distribution in groundwater resources using support vector machines (SVMs) model. *Environmental Modeling & Assessment*, 21(1), 71-82.
5. Azareh, A., Rafiei Sardooi, E., Choubin, B., Barkhori, S., Shahdadi, A., Adamowski, J., & Shamshirband, S. (2019). Incorporating multi-criteria decision-making and fuzzy-value functions for flood susceptibility assessment. *Geocarto International*, 1-21.
6. Chapi, K., Singh, V.P., Shirzadi, A., Shahabi, H., Tien Bui, D., Pham, B.T., & Khosravi, K. (2017). A novel hybrid artificial intelligence approach for flood susceptibility assessment. *Environmental Modelling & Software*, 95, 229-245.
7. Chen, W., Hong, H., Li, S., Shahabi, H., Wang, Y., Wang, X., & Ahmad, B.B. (2019). Flood susceptibility modelling using novel hybrid approach of reduced-error pruning trees with bagging and random subspace ensembles. *Journal of Hydrology*, 575, 864-873.
8. Chen, W., Li, Y., Xue, W., Shahabi, H., Li, S., Hong, H., & Ahmad, B.B. (2020). Modeling flood susceptibility using data-driven approaches of naïve bayes tree, alternating decision tree and random forest methods. *Science of the Total Environment*, 701, 134979.
9. Choubin, B., Moradi, E., Golshan, M., Adamowski, J., Sajedi-Hosseini, F., & Mosavi, A. (2019). An ensemble prediction of flood susceptibility using multivariate discriminant analysis classification and regression trees and support vector machines. *Science of the Total Environment*, 651(Pt2), 2087-2096.
10. Dat, T.T., Tri, D.Q., Truong, D.D., & Hoa, N.N. (2019). Application of mike flood model in inundation simulation with the dam-break scenarios: a case study of Dak-Drinh Reservoir in Vietnam. *International Journal of Earth Sciences*, 12, 60-70.
11. de Santana, F.B., de Souza, A.M., & Poppi, R.J. (2018). Visible and near infrared spectroscopy coupled to random forest to quantify some soil quality parameters. *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy*, 191, 454-462.
12. Eftekhari, M., Eslaminezhad, S. A., Haji Elyasi, A., & Akbari, M. (2021a). Predicting Groundwater Potential Areas Using Hybrid Artificial Intelligence Methods (Case study: Birjand Plain). *Iranian Journal of Soil and Water Research*. (In persian).
13. Eftekhari, M., Eslaminezhad, S. A., Akbari, M., DadrasAjirlou, Y., & Elyasi, A. H. (2021b). Assessment of the Potential of Groundwater Quality Indicators by Geostatistical Methods in Semi-arid Regions. *Journal of Chinese Soil and Water Conservation*, 52(3), 158-167.
14. Eftekhari, M., Eslaminezhad, S., Haji Elyasi, A., & Akbari, M. (2021c). Development of DRASTIC model using artificial intelligence on the potential of aquifer contamination in semi-arid regions. *Iranian Journal of Ecohydrology*, 8(3), 651-665.
15. Eslaminezhad, S., Eftekhari, M., Mahmoodizadeh, S., Akbari, M., & Haji Elyasi, A. (2021a). Evaluation of Tree-Based Artificial Intelligence Models to Predict Flood Risk using GIS. *Iran-Water Resources Research*, 17(2), 174-189. (In persian).
16. Eslaminezhad, S.A., Omarzadeh, D., Eftekhari, M., & Akbari, M. (2021b). Development af a Data-Driven Model to Predict Landslide Sensitive Areas. *Geographia Technica*, 16(1).
17. Gao, W., Moayedi, H., & Shahsavar, A. (2019). The feasibility of genetic programming and ANFIS in prediction energetic performance of a building integrated photovoltaic thermal (BIPVT) system. *Sol Energy*, 183, 293-305.
18. Ghorbanzadeh, O., Blaschke, T., Aryal, J., & Gholaminia, K. (2020). A new GIS-based technique using an adaptive neuro-fuzzy inference system for land subsidence susceptibility mapping. *Journal of Spatial Science*, 65(3), 401-417.

19. Ghorbanzadeh, O., Blaschke, T., Gholamnia, K., Meena, S.R., Tiede, D., & Aryal, J. (2019). Evaluation of different machine learning methods and deep-learning convolutional neural networks for landslide detection. *Remote Sensing*, 11(2), 196.
20. Giang, P.Q., Trang, N.T.M., Anh, T.T.H., & Binh, N.T. (2020). Prediction of economic loss of rice production due to flood inundation under climate change impacts using a modeling approach: A case study in Ha Tinh Province, Vietnam. *Climate Change*, 6, 52-63.
21. Hong, H., Tsangaratos, P., Ilia, I., Liu, J., Zhu, A.X., & Chen, W. (2018). Application of fuzzy weight of evidence and data mining techniques in construction of flood susceptibility map of Poyang County, China. *Science of the Total Environment*, 625, 575-588.
22. Jancewicz, K., Migoń, P., & Kasprzak, M. (2019). Connectivity patterns in contrasting types of tableland sandstone relief revealed by Topographic Wetness Index. *Science of the Total Environment*, 656, 1046-1062.
23. Jang, J.S. (1993). ANFIS: adaptive-network-based fuzzy inference system. *IEEE Transactions on Systems, Man, and Cybernetics*, 23(3), 665-685.
24. Johann, G., & Leismann, M. (2017). How to realise flood risk management plans efficiently in an urban area—the S eseke project. *Journal of Flood Risk Management*, 10(2), 173-181.
25. Kalantari, Z., Ferreira, C.S.S., Walsh, R.P.D., Ferreira, A.J.D., & Destouni, G. (2017). Urbanization development under climate change: hydrological responses in a peri-urban Mediterranean catchment. *Land Degradation & Development*, 28 (7), 2207-2221.
26. Kanani-Sadat, Y., Arabsheibani, R., Karimipour, F., & Nasseri, M. (2019). A new approach to flood susceptibility assessment in data-scarce and ungauged regions based on GIS-based hybrid multi criteria decision-making method. *Journal of Hydrology*, 572, 17-31.
27. 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. *Natural Hazards*, 83(2), 947-987.
28. Kocaman, S., Tavus, B., Nefeslioglu, H.A., Karakas, G., & Gokceoglu, C. (2020). Evaluation of floods and landslides triggered by a meteorological catastrophe (Ordu, Turkey, August 2018) using optical and radar data. *Geofluids*, 2020, 1-18.
29. Liu, R., Chen, Y., Wu, J., Gao, L., Barrett, D., Xu, T., Li, L., Huang, C., & Yu, J. (2016). Assessing spatial likelihood of flooding hazard using naïve Bayes and GIS: A case study in Bowen Basin, Australia. *Stochastic Environmental Research and Risk Assessment*, 30(6), 1575-1590.
30. Manap, M.A., Nampak, H., Pradhan, B., Lee, S., Sulaiman, W.N.A., & Ramli, M.F. (2014). Application of probabilistic-based frequency ratio model in groundwater potential mapping using remote sensing data and GIS. *Arabian Journal of Geosciences*, 7(2), 711-724.
31. Markus, M., Angel, J., Byard, G., McConkey, S., Zhang, C., Cai, X., Notaro, M., & Ashfaq, M. (2018). Communicating the impacts of projected climate change on heavy rainfall using a weighted ensemble approach. *Journal of Hydrologic Engineering*, 23(4), 04018004.
32. Mojaddadi, H., Pradhan, B., Nampak, H., Ahmad, N., & Ghazali, A. H. B. (2017). Ensemble machine-learning based geospatial approach for flood risk assessment using multi-sensor remote-sensing data and GIS. *Geomatics, Natural Hazards and Risk*, 8(2), 1080-1102.
33. Nachappa, T.G., Piralilou, S.T., Gholamnia, K., Ghorbanzadeh, O., Rahmati, O., & Blaschke, T. (2021). Flood susceptibility mapping with machine learning, multi-criteria decision analysis and ensemble using Dempster Shafer Theory. *Journal of hydrology*, 125275, 590.
34. Pham, B.T., Tien Bui, D., & Prakash, I. (2017). Landslide susceptibility assessment using bagging ensemble based alternating decision trees, logistic regression and J48 decision trees methods: a comparative study. *Geotechnical and Geological Engineering*, 35(6), 2597-2611.
35. Quiroz, J.C., Mariun, N., Mehrjou, M.R., Izadi, M., Misron, N., & Mohd Radzi, M.A. (2018). Fault detection of broken rotor bar in LS-PMSM using random forests. *Measurement*, 116, 273-280.
36. Rahmati, O., Pourghasemi, H.R., & Zeinivand, H. (2016). Flood susceptibility mapping using frequency ratio and weights-of-evidence models in the Golastan Province, Iran. *Geocarto International*, 31(1), 42-70.
37. 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. *Water Resources Management*, 31(5), 1473-1487.

38. Razavi Termeh, S.V., Kornejad, A., Pourghasemi, H.R., & Keesstra, S. (2018). Flood susceptibility mapping using novel ensembles of adaptive neuro fuzzy inference system and metaheuristic algorithms. *Science of the Total Environment*, 615, 438-451.
39. Saedi, A., Saghafian, B., & Moazami, S. (2020). Uncertainty of flood forecasts via ensemble precipitation forecasts of seven NWP models for Spring 2019 Golestan Flood. *Iran-Water Resources Research*, 16(1), 347-359. (In Persian).
40. Shahid, S., Wang, X.J., Harun, S.B., Shamsudin, S.B., Ismail, T., & Minhans, A. (2016). Climate variability and changes in the major cities of Bangladesh: observations, possible impacts and adaptation. *Regional Environmental Change*, 16(2), 459-471.
41. 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.
42. 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. *Journal of Hydrology*, 504, 69-79.
43. Tien Bui, D., Pradhan, B., Lofman, O., & Revhaug, I. (2012). Landslide susceptibility assessment in Vietnam using support vector machines, decision tree, and Naive Bayes Models. *Mathematical Problems in Engineering*, 2012, 1-26.
44. Tien Bui, D., Pradhan, B., Nampak, H., Bui, Q.T., Tran, Q.A., & Nguyen, Q.P. (2016). Hybrid artificial intelligence approach based on neural fuzzy inference model and metaheuristic optimization for flood susceptibility modeling in a high-frequency tropical cyclone area using GIS. *Journal of Hydrology*, 540, 317-330.
45. Wang, X., & Liu, H. (2019). A Knowledge-and Data-Driven Soft Sensor Based on Deep Learning for Predicting the Deformation of an Air Preheater Rotor. *IEEE Access*, 7, 159651-159660.
46. Zeraatkar, Z., & Hassanpour, F., (2016), Simulation of Birjand Urban Flood Using HEC-RAS and ARC-GIS. *Watershed Management Research Journal*, 29(3), 41-56. (In persian).
47. Ziaian Firouz Abadi, P., Badragh Nejad, A., Sarli, R., & Babaie, M. (2020). Measurement and identification of areas susceptible to flood spreading from the viewpoint of geological formations in Birjand watershed using RS / GIS. *researches in Geographical Sciences*, 20 (57), 1-24. (In persian).
48. Zhang, C., & Wang, H. (2019). Robustness of the active rotary inertia driver system for structural swing vibration control subjected to multi-type hazard excitations. *Applied Sciences*, 9(20), 4391.
49. Zhao, G., Pang, B., Xu, Z., Yue, J., & Tu, T. (2018). Mapping flood susceptibility in mountainous areas on a national scale in China. *Science of the Total Environment*, 615, 1133-1142.