



Flood susceptibility zoning using machine learning improved by genetic algorithm

Peyman Karami¹ | Seyed Ahmad Eslaminezhad² | Mobin Eftekhari³ |
 Mohammad Akbari⁴ | Melika Rastgoo⁵

1. Department of Surveying Engineering, Faculty of Surveying Engineering and Spatial Information, University of Tehran, Tehran, Iran. E-mail: peyman.karami@ut.ac.ir
2. Department of Surveying Engineering, Faculty of Surveying Engineering and Spatial Information, University of Tehran, Tehran, Iran. E-mail: peyman.karami@ut.ac.ir
2. Corresponding Author, Department of Water Engineering and Hydraulic Structures, Faculty of Engineering, Mashhad Branch, Islamic Azad University, Mashhad, Iran. E-mail: mobineftekhari@yahoo.com
3. Department of of Civil Engineering, Faculty of Engineering, University of Birjand, Birjand, Iran. E-mail: moakbari@birjand.ac.ir
4. Department of Engineering and Water Resources Management, Faculty of Civil and Environmental Engineering, Tarbiat Madras University, Tehran, Iran. E-mail: melika_rastgoo@modares.ac.ir

Article Info

Article type:

Research Article

Article history:

Received 20 October 2022

Received in revised form 9

December 2022

Accepted 18 December 2022

Published online 5 May 2023

Keywords:

Optimization,

Flood,

Random forest,

Support vector machine.

ABSTRACT

Due to the increase in the risk of floods, especially in the cities, and the emergence of human, financial, and environmental risks due to its increase, the flood zoning areas are of great importance. Therefore, in this study, flood susceptible areas in Birjand plain were tried to be zoned with the help of effective criteria. In this regard, the data-driven methods of support vector machine (SVM) and random forest (RF) were used in combination with genetic algorithm to zoning flood susceptible areas. Therefore, in order to implement and validate the mentioned models, 42 flood prone locations in the study area were extracted. In addition, 19 hydrogeological, topographical, geological and environmental criteria affecting flood susceptibility in the study area were extracted to be used to predict flood susceptibility map. Area under the curve (AUC) and a variety of other statistical indicators including coefficient of determination (R^2) and Root mean square error (RMSE) were used to evaluate the performances of the models. The values of R^2 , RMSE and AUC obtained from the SVM-GA method were 0.9032, 0.2751 and 0.931, respectively, and the RF-GA method were 0.9823, 0.2321 and 0.914, respectively, which indicate the compatibility and The RF model is more accurate than the SVM model. The results also showed that the susceptibility of flooding in the central areas of the study area, due to lower altitude and slope angle, is higher than other areas.

Cite this article: Karami, M., Eslaminezhad, S.A., Eftekhari, M., Akbari, M., & Rastgoo, M. (2023). Flood susceptibility zoning using machine learning improved by genetic algorithm. *Journal of Natural Environment*, 76 (1), 43-60. DOI: <http://doi.org/10.22059/jne.2022.350170.2485>



پهنه‌بندی حساسیت سیلاب با استفاده از روش‌های یادگیری ماشین بهبودیافته توسط الگوریتم ژنتیک

پیمان کرمی^۱ | سید احمد اسلامی نژاد^۲ | مبین افتخاری^۳ | محمد اکبری^۴ | ملیکا راستگو^۵

۱. گروه مهندسی نقشه برداری، دانشکده مهندسی نقشه برداری و اطلاعات مکانی، دانشگاه تهران، تهران، ایران. رایانامه: peyman.karami@ut.ac.ir
۲. گروه مهندسی نقشه برداری، دانشکده مهندسی نقشه برداری و اطلاعات مکانی، دانشگاه تهران، تهران، ایران. ahmad.eslami73@ut.ac.ir
۳. نویسنده مسئول، گروه مهندسی عمران آب و سازه‌های هیدرولیکی، دانشکده فنی مهندسی، دانشگاه تهران، تهران، ایران. رایانامه: mobineftekhari@yahoo.com
۴. گروه مهندسی عمران، دانشکده مهندسی، دانشگاه بیرجند، بیرجند، ایران. رایانامه: moakbari@birjand.ac.ir
۵. گروه مهندسی و مدیریت منابع آب، دانشکده مهندسی عمران و محیط زیست، دانشگاه تربیت مدرس، تهران، ایران. رایانامه: melika_rastgoo@modares.ac.ir

اطلاعات مقاله	چکیده
<p>نوع مقاله: مقاله پژوهشی</p> <p>تاریخ دریافت: ۱۴۰۱/۰۷/۲۸</p> <p>تاریخ بازنگری: ۱۴۰۱/۰۹/۱۸</p> <p>تاریخ پذیرش: ۱۴۰۱/۰۹/۲۷</p> <p>تاریخ انتشار: ۱۴۰۲/۰۲/۱۵</p> <p>کلیدواژه‌ها: بهبودسازی، سیل، جنگل تصادفی، ماشین بردار پشتیبان.</p>	<p>با توجه به بالا رفتن خطر وقوع سیلاب خصوصاً در سطح شهرها و به وجود آمدن خطرات جانی، مالی و محیط زیستی ناشی از افزایش آن، پهنه‌بندی مناطق سیل خیز از اهمیت بالایی برخوردار است. بنابراین در این مطالعه سعی شد مناطق حساس به سیلاب در دشت بیرجند با استفاده از معیارهای مؤثر پهنه‌بندی شود. در این راستا از روش‌های داده‌محور ماشین بردار پشتیبان (SVM) و جنگل تصادفی (RF) در ترکیب با الگوریتم ژنتیک جهت پهنه‌بندی مناطق حساس به سیل استفاده شد. بنابراین برای پیاده‌سازی و اعتبارسنجی مدل‌های ذکر شده، ۴۲ موقعیت سیل خیز در منطقه مورد مطالعه استخراج شد. علاوه بر این، ۱۹ معیار هیدروژئولوژیکی، توپوگرافی، زمین‌شناسی و محیطی مؤثر بر حساسیت سیلاب در منطقه مورد مطالعه استخراج شدند تا برای پیش‌بینی نقشه حساسیت سیل مورد استفاده قرار گیرند. سطح زیر منحنی (AUC) و انواع شاخص‌های آماری دیگر از جمله ضریب تشخیص (R^2) و ریشه میانگین خطای مربعات (RMSE) برای ارزیابی عملکرد مدل‌ها استفاده شد. مقادیر R^2، RMSE و AUC حاصل از روش SVM-GA به ترتیب ۰/۹۰۳۳، ۰/۲۷۵۱ و ۰/۹۳۱ و روش RF-GA به ترتیب ۰/۹۸۲۳، ۰/۲۳۲۱ و ۰/۹۱۴ به دست آمد که نشان‌دهنده سازگاری و دقت بالای مدل RF نسبت به مدل SVM است. هم‌چنین نتایج نشان داد که حساسیت سیل به دلیل ارتفاع و زاویه شیب کمتر در مناطق مرکزی منطقه مطالعاتی بیشتر از سایر مناطق است. نتایج این مطالعه می‌تواند به منظور مدیریت مناطق آسیب‌پذیر و کاهش خسارت‌های سیل مورد استفاده قرار گیرد.</p>

استناد: کرمی، پیمان؛ اسلامی نژاد، سید احمد؛ افتخاری، مبین؛ اکبری، محمد؛ و راستگو، ملیکا (۱۴۰۲). م پهنه‌بندی حساسیت سیلاب با استفاده از روش‌های یادگیری

ماشین بهبودیافته توسط الگوریتم ژنتیک. محیط زیست طبیعی، ۷۶ (۱)، ۴۳-۶۰.

DOI: <http://doi.org/10.22059/jne.2022.350170.2485>



مقدمه

هر جریان سطحی آب در صورتی که در مقطع خاصی از رودخانه از جریان عادی رودخانه بیشتر باشد، تداوم زمانی آن کم و از بستر طبیعی رودخانه تجاوز کند سیل تلقی می‌شود (Papaioannou, 2015; Eini *et al.*, 2020; Hudson and Botzen, 2020). در کشور ایران تمامی نواحی از پدیده سیل متأثر است و این رخداد محدود به موقعیت و منطقه خاصی نمی‌باشد. نوع سیل و میزان خسارات وارد شده با توجه به خصوصیات هر منطقه، متفاوت است (Khosravi *et al.*, 2016; Eslaminezhad *et al.*, 2021a). در دهه‌های اخیر با توجه به تلاش‌های تصمیم‌گیران و متخصصان در راستای کاهش خطرات ناشی از سیل، تعداد حوادث و تلفات اقتصادی و انسانی هم‌چنان در سراسر جهان تلفاتی از این قبیل در حال افزایش است (Johann and Leismann, 2017; Kocaman *et al.*, 2020).

پهنه‌بندی حساسیت سیلاب یکی از مراحل حساس در مدیریت رودخانه‌ها می‌باشد. یکی از روش‌ها برای تهیه نقشه‌های پهنه‌بندی حساسیت سیلاب، استفاده از سیستم اطلاعات مکانی (GIS) است (Pourghasemi *et al.*, 2020; Eslaminezhad *et al.*, 2022a). امروزه با توسعه فن‌آوری‌های مرتبط با سیستم اطلاعات مکانی (GIS) و سنسور از دور^۲ (RS)، تعیین مناطق مستعد سیل از طریق روش‌های یادگیری ماشین، بسیار سریع‌تر است، کسب نتایج نهایی با دقت بالا را تضمین می‌کند و داده‌های موردنیاز برای مدل‌سازی اغلب آزادانه در دسترس‌اند (Eslaminezhad *et al.*, 2021a). با توجه به تحقیقات پیشین، پهنه‌بندی سیلاب توسط روش‌های ادغام در دو دسته کلی دانش‌محور و داده‌محور تقسیم‌بندی می‌شوند (Guevara *et al.*, 2018; Wang and Liu, 2019). معمولاً در مناطقی که به لحاظ آماری تعداد شواهد شناخته‌شده در آن کافی باشند، روش‌های داده‌محور کارایی بالایی دارند، در حالی که روش‌های دانش‌محور در مناطقی که کمتر شناخته شده‌اند و یا تعداد محدودی از اهداف مورد نظر در محدوده وجود دارند، کارآمد هستند. مطالعات متعددی در خصوص پهنه‌بندی حساسیت سیلاب با رویکردهای مبتنی بر روش‌های ادغام دانش‌محور و داده‌محور صورت گرفته است که می‌توان به موارد زیر اشاره نمود:

Zeraatkar و Hassanpour (۲۰۱۶) در حوضه شاهرود، رفتار سیلاب درون شهری بیرجند را مورد مطالعه قرار دادند. در این راستا، ابتدا شماره منحنی رواناب تعیین و سپس دبی سیلاب حوضه با استفاده از مدل HEC-HMS شبیه‌سازی گردید. نتایج نشان داد که دوره بازگشت‌های ۲۵ ساله منجر به وقوع دبی اوج ۲۰۷/۶۳ مترمکعب بر ثانیه و ۱۰۰ ساله منجر به وقوع دبی اوج ۴۶۳/۵۶ مترمکعب بر ثانیه می‌شوند. حداکثر جریانی که سازه‌های هیدرولیکی و مقاطع رودخانه قادر به عبور آن می‌باشند، به سیلاب با دوره بازگشت ۲۵ ساله مرتبط بود. بنابراین توجه بیشتر به افزایش پتانسیل عبور سیلاب رودخانه در راستای کاهش خطرات ناشی از سیل ضروری است.

Mohamadinia و Ghavami (۲۰۱۷) به پیش‌بینی مکانی مناطق سیل‌خیز استان گیلان با استفاده از GIS و هفت معیار دما، رطوبت، بارش، ارتفاع، شیب، پوشش گیاهی و فاصله از رودخانه پرداختند. در این مطالعه از دو روش رگرسیون وزن‌دار جغرافیایی و رگرسیون خطی استفاده کردند. رگرسیون وزن‌دار جغرافیایی دقت بالاتری را در پیش‌بینی مناطق سیل‌خیز نشان داد (Ghavami and Mohamadinia, 2017). Xiao و همکاران (۲۰۱۷) یک چارچوب تجزیه و تحلیل چندمعیاره با استفاده از GIS، فرآیند سلسله‌مراتب تحلیلی فازی^۳ (AHP) و روش میانگین وزن‌دار مرتبه‌یافته^۴ (OWA) را برای ارزیابی حساسیت سیلاب تهیه کردند. اهمیت نسبی معیارها با روش AHP فازی تعیین شد. از روش OWA برای تحلیل اثرات نگرش‌های مختلف میزان ریسک تصمیم بر نتایج ارزیابی، استفاده کردند. همچنین روش میانگین محاسبه‌شده را با استفاده از مزایای فضای متغیر در محیط GIS برای ادغام معیارها اجرا کردند (Xiao *et al.*, 2017). Hong و همکاران (۲۰۱۸) از ترکیب روش‌های رگرسیون لجستیک (LR)^۵، جنگل

^۱Geographic information system

^۲Remote sensing

^۳Analytic hierarchy process

^۴Ordered weighted averaging

^۵Logistic regression

تصادفی (RF)^۶ و ماشین بردار پشتیبان (SVM)^۷ با روش وزن شواهد^۸ (WOE) جهت تهیه نقشه حساسیت سیل در منطقه پویانگ^۹ کشور چین استفاده کردند. نتایج نشان داد که مدل ترکیبی SVM-WOE بالاترین مقدار عملکرد پیش‌بینی^{۱۰} (AUC) را نسبت به سایر مدل‌های ترکیبی RF-WOE و LR-WOE دارد (Hong et al., 2018).

Razavi Temeh و همکاران (۲۰۱۸) از ترکیب مدل سیستم استنتاج عصبی فازی (ANFIS)^{۱۱} با الگوریتم‌های فرا ابتکاری شامل الگوریتم ژنتیک (GA)^{۱۲}، الگوریتم بهینه‌سازی کلونی مورچه‌ها (ACO)^{۱۳} و الگوریتم بهینه‌سازی ازدحام ذرات (PSO)^{۱۴} برای پیش‌بینی حساسیت سیل در شهرستان جهرم استان فارس استفاده کردند. به‌منظور پیاده‌سازی روش‌های مورد نظر، نه معیار مؤثر بر وقوع سیل استخراج شدند. سطح زیر منحنی (AUC) به‌دست آمده از ROC دقت ۹۴/۵ درصد را برای مدل ترکیبی ANFIS-PSO نشان داد (Razavi Temeh et al., 2018).

Ziaian Firouz Abadi و همکاران (۲۰۲۰) در حوضه آبریز بیرجند به شناسایی مناطق مستعد پخش سیلاب براساس سازندهای زمین‌شناسی پرداختند. کلیه عملیات مربوط به تجزیه و تحلیل داده‌ها در محیط ArcGIS به کمک سه روش منطق بولین، سلسله مراتبی (AHP) و مجموع ساده وزنی^{۱۵} (SAW) برای وزن‌دار کردن و تلفیق لایه‌ها (شیب، فرسایش، کاربری اراضی، شبکه زهکشی، خاک، پراکنش گسل، واحد زمین‌شناسی، سطح آب‌های زیرزمینی و تصاویر ماهواره‌ای) استفاده گردید و مناطق پرمخاطره آن مشخص شد. منطق بولین در مکان‌یابی پخش سیلاب کارایی بهتری را نشان داد.

در تمام مطالعات ذکر شده، محققان مدل‌های دانش‌محور و داده‌محور را برای انتخاب بهترین مدل در راستای پیش‌بینی حساسیت سیلاب مقایسه کرده‌اند. شناسایی روش‌های ترکیبی جدید با حداقل داده‌های میدانی جهت پیش‌بینی حساسیت سیلاب و شناسایی نواحی مستعد سیل در مطالعات حوزه سیل مهم است.

اگرچه در تحقیقات پیشین از الگوریتم‌های یادگیری ماشین برای پیش‌بینی نواحی مستعد سیل استفاده شده است، اما هیچ روشی در پیش‌بینی دقیق نواحی مستعد سیل کامل نیست. در سال‌های اخیر تهیه نقشه حساسیت به سیل از جمله روش‌های جدیدی است که توجه محققان را به خود جلب کرده است. بنابراین نوآوری این تحقیق بهینه‌سازی دو مدل مبتنی بر یادگیری ماشین یعنی SVM و RF توسط الگوریتم ژنتیک است که در نتایج تحقیقات اخیر کارایی این الگوریتم‌ها به اثبات رسیده است. در تحقیق حاضر از الگوریتم‌های ذکر شده به‌طور همزمان جهت تعیین مدل دقیق‌تر در پیش‌بینی مناطق مستعد سیل و شناسایی معیارهای مؤثر در پیش‌بینی حساسیت به سیل در حوضه آبریز دشت بیرجند واقع در استان خراسان جنوبی استفاده شد. در نهایت، معیار سطح زیر منحنی (AUC) و سایر معیارهای آماری مناسب جهت اعتبارسنجی مدل‌های پیشنهادی در منطقه مورد مطالعه استفاده قرار شدند.

روش‌شناسی پژوهش

روش انجام پژوهش حاضر، توصیفی-تحلیلی بوده و نوع آن بر اساس هدف کاربردی است. تمامی پردازش‌های مربوط به داده‌ها در محیط‌های نرم‌افزاری QGIS 3.1، SAGA GIS 7.9 و Google earth engine و محاسبات کمی و توسعه روش‌ها در محیط برنامه‌نویسی Matlab انجام شد. براساس شکل ۱، روند اجرایی این مطالعه شامل شش مرحله اساسی است: ۱) تهیه نقشه مرجع سیل از طریق شناسایی ۴۲ موقعیت سیل خیز در حوضه آبریز منطقه مورد مطالعه، ۲) استخراج معیارهای مکانی مؤثر بر سیل در

⁶Random forest

⁷Support vector machine

⁸Weight of Evidence

⁹Poyang

¹⁰Area under the curve

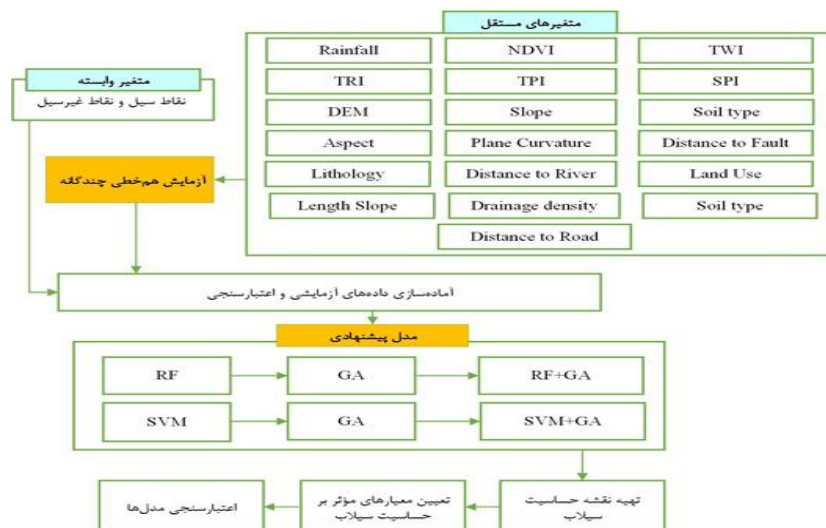
¹¹Adaptive neuro-fuzzy inference system

¹²Genetic algorithm

¹³Ant colony optimization

¹⁴Particle swarm optimization

¹⁵Simple weight additive



شکل ۱. روند اجرایی تحقیق

منطقه مورد مطالعه شامل ۱۹ معیار هیدروژئولوژیکی، توپوگرافی، زمین‌شناسی و محیطی، (۳) انجام آزمایش هم‌خطی چندگانه جهت بررسی مستقل بودن معیارهای مکانی مؤثر بر سیل در منطقه مورد مطالعه، (۴) پیش‌بینی حساسیت وقوع سیلاب توسط الگوریتم‌های SVM-GA و RF-GA و تهیه نقشه حساسیت سیلاب به روش کریجینگ، (۵) تعیین ترکیب بهینه معیارهای مؤثر در پیش‌بینی حساسیت سیل منطقه مورد مطالعه و (۶) ارزیابی و مقایسه عملکرد رویکردهای بکار برده شده توسط شاخص‌های آماری مناسب.

منطقه مورد مطالعه: استان خراسان جنوبی بر اساس اقلیم نمای دومارتن در یک اقلیم خشک واقع شده‌است. مناطق روستایی آن بیشتر در مناطق کوهستانی و مرکز استان متمرکزند. این استان بر اساس زیرحوضه‌های درجه یک، در دو زیر حوضه مرزی شرق و فلات مرکزی و براساس زیرحوضه‌های درجه دو در چهار زیرحوضه کویر مرکزی، کویر لوت، نمکزار خواف و هامون هیرمند واقع شده‌است. دشت بیرجند، یکی از محدوده‌های مطالعاتی حوضه آبریز کویر لوت است که مختصات جغرافیایی آن بین طول‌های ۵۸° و ۴۱' تا ۵۹° و ۴۶' شرقی و عرض‌های ۳۲° تا ۳۵' و ۳۳° و ۸' شمالی واقع شده است (شکل ۲). مساحت کلی محدوده مطالعاتی ۳۴۰۶/۷۲ km² است که از این مقدار سهم دشت و ارتفاعات به ترتیب ۹۰۰/۶۳ km² و ۲۵۰۶/۰۹ km² می‌باشد. رودخانه شاهرود به طول ۱۰۸ کیلومتر زهکش اصلی دشت بوده که در اغلب مواقع سال به‌جز مواقع سیلابی خشک است، بنابراین اکثر آب موردنیاز بخش‌های شرب، کشاورزی و صنعت منطقه توسط سفره آب زیرزمینی تأمین می‌شود (Eftekhari et al., 2021).

روش تحقیق

تهیه لایه‌های اطلاعاتی منطقه مورد مطالعه

معیارهای مکانی مؤثر بر سیلاب: براساس بررسی مطالعات و در دسترس بودن داده‌ها، ۱۹ معیار هیدروژئولوژیکی، توپوگرافی، زمین‌شناسی و محیطی شامل بارندگی، شاخص نرمال شده تفاوت پوشش گیاهی^{۱۶} (NDVI)، شاخص رطوبت توپوگرافی^{۱۷} (TWI)، شاخص ناهمواری زمین^{۱۸} (TRI)، شاخص موقعیت توپوگرافی^{۱۹} (TPI)، شاخص توان آبراهه^{۲۰} (SPI)، مدل رقومی ارتفاعی^{۲۱} (DEM)، زاویه شیب، جریان تجمعی، جهت شیب، انحنای سطح، فاصله تا گسل، فاصله تا جاده، فاصله تا رودخانه، سنگ‌شناسی، جنس خاک، کاربری اراضی، تراکم زهکشی و طول شیب به‌عنوان متغیرهای مستقل برای پیش‌بینی حساسیت سیل در منطقه مورد مطالعه استخراج شدند (Tien Bui et al., 2016; Chapi et al., 2017; Mojaddadi et al., 2017; Ahmadlou et al., 2018; Razavi Termeh et al., 2018; Chen et al., 2019).

¹⁶Normalized difference vegetation index

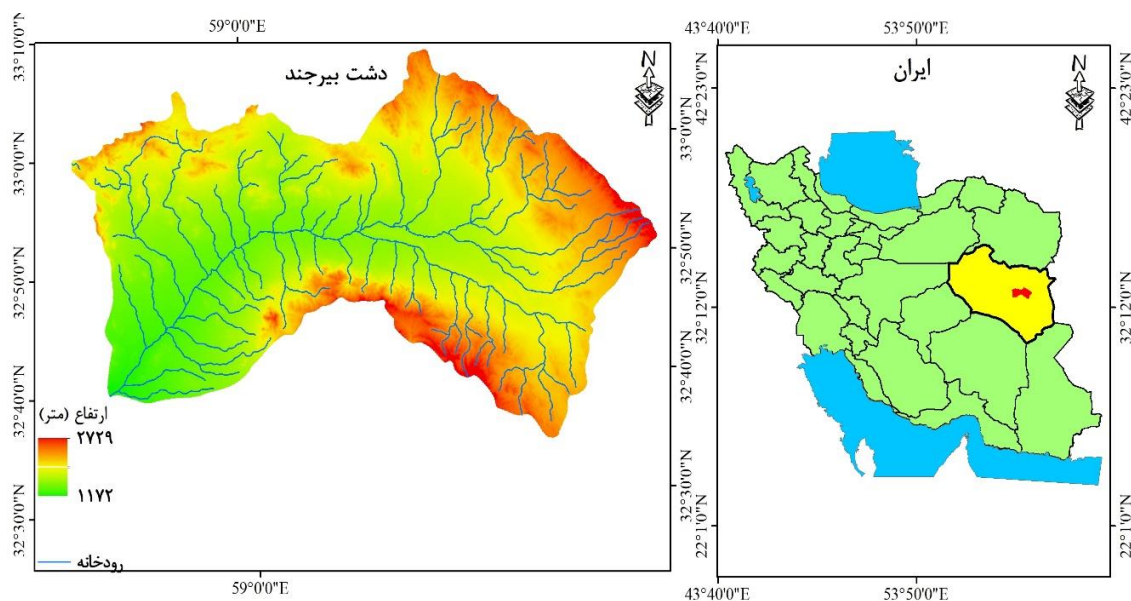
¹⁷Topographic Wetness Index

¹⁸Terrain ruggedness index

¹⁹Topographic Position Index

²⁰Stream power index

²¹Digital Elevation model



شکل ۲. محدوده منطقه مطالعاتی

نقشه رستری هر یک از معیارها با اندازه پیکسل ۳۰ متر تهیه شد. به منظور تهیه لایه DEM از تصاویر ماهواره‌ای SRTM^{۲۲} با قدرت تفکیک ۳۰ متر استفاده شد (Tien Bui *et al.*, 2016; Razavi Termeh *et al.*, 2018). در ادامه لایه‌های زاویه شیب، جهت شیب و انحنای سطح از لایه DEM ایجاد شدند. نقشه توپوگرافی منطقه مورد مطالعه از سازمان نقشه‌برداری کشور ایران در مقیاس ۱:۵۰۰۰۰۰ اخذ و تهیه شد. لایه‌های فاصله تا رودخانه و جاده با استفاده از نقشه‌های توپوگرافی ۱:۵۰۰۰۰۰ تهیه شدند. همچنین، لایه فاصله تا گسل نیز براساس نقشه‌های توپوگرافی ۱:۱۰۰۰۰۰۰ تهیه و استخراج شدند. لایه رستری کاربری اراضی منطقه مورد مطالعه از طریق تصاویر ماهواره‌ای لندست ۸ در ۱۰ کلاس (شهری، صخره، کشاورزی/آبش، مرتع، آبش، باغ، جنگل، دیم، کشاورزی و بایر) استخراج گردید. همچنین، از باند چهار و پنج لندست ۸، برای تولید لایه NDVI استفاده شد. هر چه مقدار NDVI کمتر باشد، به دلیل کاهش تراکم پوشش گیاهی، حساسیت بیشتر سیل را به همراه دارد (Eslaminezhad *et al.*, 2020). بارندگی به‌عنوان یکی از معیارهای محرک سیل به حساب می‌آید. بنابراین نقشه میانگین بارندگی با استفاده از شش ایستگاه سینوپتیک در منطقه مورد مطالعه محاسبه شد. همچنین شاخص TWI کاربرد زیادی در تعیین میزان تأثیر توپوگرافی به محل، اندازه ناحیه اشباع شده و تولید رواناب دارد که طبق رابطه ۱ محاسبه می‌شود (Jancewicz *et al.*, 2019):

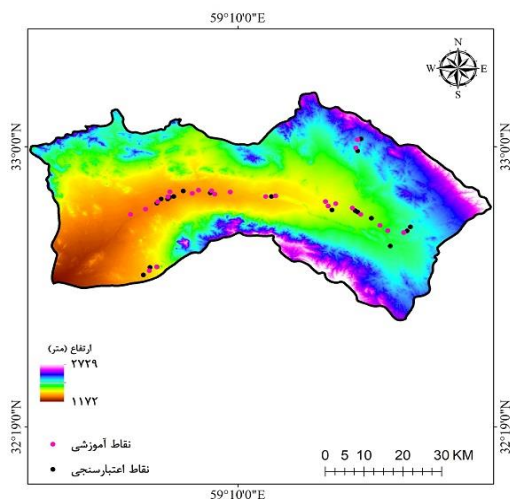
$$TWI = \ln\left(\frac{A_s}{\tan \beta}\right) \quad \text{رابطه ۱}$$

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

$$SPI = A_s \times \tan \beta \quad \text{رابطه ۲}$$

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

²²Shuttle Radar Topography Mission



شکل ۳. دسته‌بندی نقاط سیل و غیر سیل در منطقه مورد مطالعه

$$\text{TPI} = Z_0 - \frac{\sum_{i=1}^n Z_i}{n} \quad \text{رابطه ۳}$$

که Z_0 : ارتفاع سلول مورد نظر 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 است که طبق رابطه ۴ محاسبه می‌گردد (Kalantari et al., 2017):

$$\text{TRI} = \sqrt{\text{Abs}(\text{max}^2 - \text{min}^2)} \quad \text{رابطه ۴}$$

که max و min به ترتیب بزرگترین و کوچکترین مقادیر ارتفاعی سلول‌های همسایه هستند. معیار طول شیب LS^{33} (LS) مقیاسی از ظرفیت انتقال رسوب برای جریان سطحی است که از رابطه ۵ محاسبه می‌شود (Khosravi et al., 2016):

$$LS = \left(\frac{A_s}{22.13}\right)^{0.6} \left(\frac{\sin\beta}{0.0896}\right)^{0.06} \quad \text{رابطه ۵}$$

که A_s : مساحت یک حوضه مشخص (m^2/m) و β : زاویه شیب بر حسب درجه است. تراکم زهکشی بر اساس میزان نفوذ ناشی از آب‌های سطحی، شرایط تغذیه آب‌های زیرزمینی را نشان می‌دهد. نقشه تراکم زهکشی از روی نقشه DEM محاسبه شد. بافت خاک یکی از عوامل و اجزا مهم در کنترل نفوذ، میزان رواناب و در نتیجه رخداد خطر طبیعی می‌باشد (Kanani-Sadat et al., 2019). نقشه جنس خاک منطقه مورد مطالعه در مقیاس ۱:۲۵۰۰۰۰ از طریق نرم‌افزار Google earth engine اخذ گردید. همچنین واحدهای سنگ‌شناسی متفاوت، اختلاف چشم‌گیری در ناپایداری دامنه نشان می‌دهد. با استفاده از نقشه زمین‌شناسی کشور ایران با مقیاس ۱:۱۰۰۰۰۰، نقشه سنگ‌شناسی منطقه مورد مطالعه تهیه شد.

داده‌های گذشته سیلاب: به منظور برآورد سیل‌های آینده در دشت بیرجند، تجزیه و تحلیل سیلاب‌های گذشته ضروری است. نقاط سیل، یک سطح مهم از رابطه بین رخداد سیل و معیارهای به‌وجود آورنده آن است. رویدادهای گذشته سیل به‌عنوان مبنایی برای پیش‌بینی وقوع سیل در آینده به حساب می‌آیند، به طوری که مناطق نزدیک به رخداد‌های گذشته حساسیت بالایی به سیل‌گیری دارند (Rahmati et al., 2016). روش‌های مختلفی مانند پایگاه‌های اطلاعاتی سیل‌های گذشته و تفسیر تصاویر ماهواره‌ای برای به‌دست آوردن نقشه‌های مرجع سیل وجود دارد (Chen et al., 2019).

در مجموع، ۴۲ موقعیت سیل خیز (مقدار ۱) در منطقه مورد مطالعه در بازه زمانی ۱۰ ساله بین سال‌های ۱۳۸۸ تا ۱۳۹۸ ثبت شده است که مربوط به بانک اطلاعاتی شرکت آب منطقه‌ای استان خراسان جنوبی در سال ۱۳۹۹ است (شکل ۳). در این مطالعه، به‌طور تصادفی ۷۰ درصد از این نقاط برای آموزش مدل‌ها و ۳۰ درصد باقی‌مانده برای اعتبارسنجی مدل‌ها استفاده شدند. همچنین براساس نتایج مطالعات قبلی و به‌منظور نزدیک‌تر شدن نتایج تحقیق به واقعیت، ۴۲ موقعیت غیرسیل (مقدار ۰) نیز با استفاده از نقشه توپوگرافی، بازدیدهای میدانی و نرم‌افزار Google earth در مناطقی که سیلاب قادر به پیشروی در آنجا نیست مانند کوه‌ها

³³Length slope

و تپه‌ها ایجاد گردید (Khosravi et al., 2016). این نسبت بین مجموعه داده‌های آموزشی و اعتبارسنجی توسط بیشتر محققان متمرکز بر ارزیابی حساسیت مخاطرات طبیعی، توصیه شده است (Razavi Termeh et al., 2018; Nachappa et al., 2021). **روش ماشین بردار پشتیبان (SVM):** یک روش یادگیری ماشین نسبتاً جدید و یک الگوریتم یادگیری ماشین نظارت شده است (Arabgol et al., 2016). 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. در این مورد، رابطه ۶ برای طبقه‌بندی مجموعه داده‌ها استفاده خواهد شد (Arabgol et al., 2016):

$$f(x) = \sum_{i=1}^n (a_i - a_i^*)K(x_i, x) + b \quad \text{رابطه ۶}$$

که در آن α_i و α_i^* ضرایب لاگرانژ، K : تابع کرنل و b : انحراف ابر صفحه از مبدأ است.

جنگل تصادفی (RF): الگوریتم جنگل تصادفی یکی از پرکاربردترین الگوریتم‌ها برای بررسی مسائل طبقه‌بندی و پیش‌بینی چندگانه است که حساسیت کمتری به چندخطی بودن دارد و از نقطه نظر داده‌های از دست‌رفته و نامتعادل منجر به نتایج نسبتاً پایداری می‌شود (de Santana et al., 2018; Quiroz et al., 2018; Eslaminezhad et al., 2022b). چهار مرحله زیر فرآیند RF را مشخص می‌کنند: (۱) تعریف و باز نمونه‌گیری 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 گسترده و (۴) ایجاد یک درخت تصمیم واحد با تجمع درخت تصمیم اختصاص داده شده به هر مثال.

الگوریتم ژنتیک (GA): الگوریتم ژنتیک را می‌توان به‌عنوان یک روش جستجوی کلی در نظر گرفت که بر قوانین تکامل زیستی استوار است و از آن تقلید می‌کند (Mirjalili, 2019). در الگوریتم‌های ژنتیک از عملگرهای ژنتیکی در هنگام تولید مثل استفاده می‌شود. از طریق تأثیر این عملگرها بر جمعیت، نسل بعدی آن جمعیت تولید می‌شود. عملگرهای انتخاب^{۲۴}، ترکیب^{۲۵} و جهش^{۲۶} اغلب در الگوریتم‌های ژنتیک استفاده می‌شوند (Mirjalili, 2019). گام‌های اجرای یک مدل بهینه‌سازی براساس الگوریتم ژنتیک به‌صورت ذیل می‌باشد (Sun et al., 2020):

(۱) ایجاد و ارزیابی جمعیت تصادفی، (۲) انتخاب والدین و ادغام آن‌ها برای ایجاد جمعیت اولیه فرزندان، (۳) انتخاب اعضای جمعیت برای اعمال جهش و ایجاد جمعیت جهش‌یافته، (۴) ترکیب یا ادغام جمعیت اصلی، فرزندان و جهش‌یافته‌ها و ایجاد جمعیت اصلی جدید، (۵) تکرار مرحله ۲ در صورت عدم تحقق شرایط خاتمه و (۶) پایان. در این تحقیق هر کروموزوم دارای ۱۷ ژن (معیار) است که به‌صورت باینری مقداردهی شده است، به گونه‌ای که هر بار تعدادی ژن، برای ترکیب انتخاب و مقدار ۱ به ژن‌های منتخب و مقدار ۰ به ژن‌های انتخاب نشده اختصاص می‌یابد.

آزمایش هم‌خطی چندگانه: به‌منظور تشخیص هم‌خطی چندگانه در میان عوامل مختلف، آماره تحمل^{۲۷} (TOL) و عامل تورم واریانس^{۲۸} (VIF) دو پارامتر آماری رایج هستند که طبق رابطه‌های ۷ و ۸ محاسبه می‌شوند (Choubin et al., 2019):

$$TOL = 1 - R_j^2 \quad \text{رابطه ۷}$$

$$VIF = \frac{1}{TOL} \quad \text{رابطه ۸}$$

که در آن R_j : مقدار مربع R رگرسیون با استفاده از پارامتر رگرسیون Z در موارد دیگر است. هنگامی که مقدار TOL کوچکتر از ۰/۱ و مقدار VIF بزرگتر از ۵ باشد، هم‌خطی چندگانه بالایی در میان متغیرهای پیش‌بینی‌کننده وجود دارد (Choubin et al., 2019).

²⁴Selection

²⁵Crossover

²⁶Mutation

²⁷Tolerance

²⁸Variance Inflation Factor

ارزیابی عملکرد و دقت مدل‌ها

شاخص‌های آماری: خروجی مدل‌های داده‌محور پارامترهای متعددی را شامل می‌شود. پارامتر ضریب تشخیص R^2 برای سنجش مناسبت برازش مدل و پارامتر خطای جذر میانگین مربعات (RMSE) جهت سنجش توزیع باقی‌مانده‌های مدل بکار می‌روند که به ترتیب، براساس رابطه‌های ۹ و ۱۰ محاسبه می‌شوند (Fotheringham and Oshan, 2016; Oshan et al., 2019):

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad \text{رابطه ۹}$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad \text{رابطه ۱۰}$$

که n : تعداد مشاهدات، y_i : مشاهده i ام، \hat{y}_i : مقدار پیش‌بینی شده مشاهده i ام و \bar{y} : میانگین مشاهدات است. با توجه به این که یکی از مهم‌ترین پارامترهای ارزیابی روش‌های چندمعیاره، پارامتر RMSE است، از این رو تابع برازش الگوریتم ژنتیک، کمینه کردن مقدار RMSE انتخاب شده است تا میزان سنجش توزیع باقی‌مانده‌های مدل و در واقع دقت مدل در پیش‌بینی حساسیت سیلاب بررسی شود (Fotheringham and Oshan, 2016; Oshan et al., 2019).

منحنی مشخصه عملکرد سیستم (ROC): در این مطالعه برای ارزیابی عملکرد مدل‌ها، منحنی مشخصه عملکرد سیستم ROC^{39} به ترتیب با نرخ‌های مثبت واقعی و نرخ مثبت کاذب بر روی محور Y و محور X استفاده شده است (Razavi Termeh et al., 2018). معیار نرخ مثبت واقعی معرف تعداد مکان‌هایی است که به‌عنوان یک مکان با پتانسیل بالای سیل به درستی طبقه‌بندی می‌شوند. در مقابل معیار نرخ مثبت کاذب معرف تعداد مکان‌هایی است که به‌طور نادرست به‌عنوان یک مکان با پتانسیل بالای سیل طبقه‌بندی می‌شوند. منطقه تحت منحنی ROC یعنی AUC عملکرد مدل را به‌صورت کمی تعیین می‌کند (Khosravi et al., 2016). مقادیر بالاتر AUC و نزدیک به ۱ نشان‌دهنده عملکرد بهتر مدل‌ها است (Khosravi et al., 2016). میزان تناسب یا قابلیت یادگیری مدل با استفاده از منحنی ROC براساس مجموعه داده‌های آموزشی تعیین می‌شود؛ در حالیکه مهارت پیش‌بینی مدل را مجموعه داده‌های اعتبارسنجی نشان می‌دهند (Chapi et al., 2017).

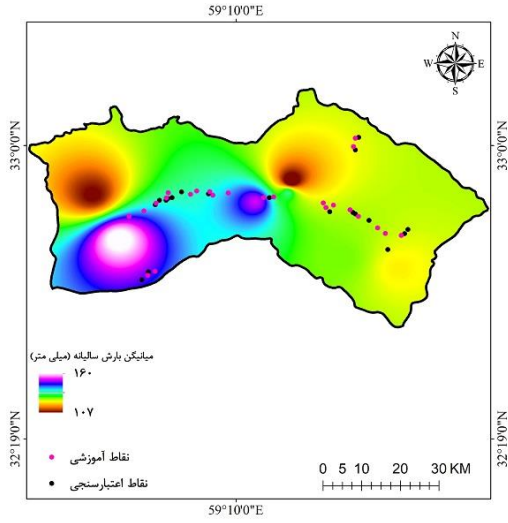
یافته‌های پژوهش

آماده‌سازی داده‌ها: جهت پیاده‌سازی، هر کدام از معیارها در قالب نقشه رستری با اندازه پیکسل ۳۰ متر از روش درون‌یابی کریجینگ استفاده شد که در شکل ۴ نشان داده شده است (Eslaminezhad et al., 2021b).

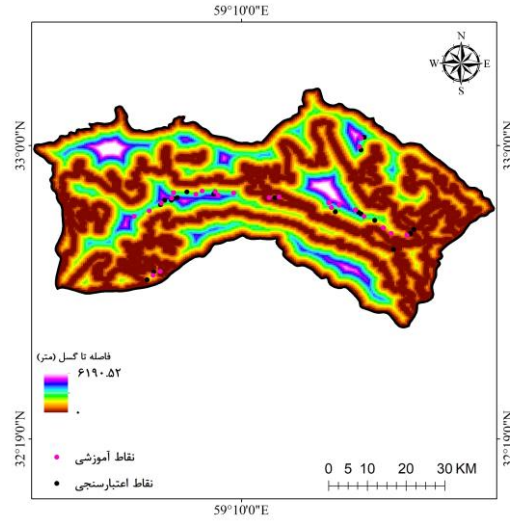
انتخاب معیارهای مستقل مؤثر بر سیل با استفاده از آزمایش هم‌خطی چندگانه: براساس جدول ۱، نتایج نشان داد که معیارهای LS و TPI شروط هم‌خطی چندگانه را ($TOL > 0.1$ and $VIF < 5$) رد کرده‌اند، بنابراین این معیارها را نمی‌توان به‌عنوان ورودی در مدل‌سازی استفاده کرد و در نتیجه از بین معیارهای این تحقیق حذف می‌شود. به‌جز معیارهای طول شیب و شاخص موقعیت توپوگرافی، هیچ هم‌خطی چندگانه بالایی در میان سایر معیارهای انتخاب‌شده مشاهده نشد. بنابراین ۱۷ معیار دیگر جهت پیادسازی مدل‌های این تحقیق بکار برده می‌شوند.

تعیین ترکیب بهینه معیارهای مؤثر در پیش‌بینی حساسیت سیلاب: معیارها با همان ترتیبی که در جدول ۱ ارائه شده است، کروموزوم‌های الگوریتم ژنتیک را تشکیل داده‌اند. پس از اجرای الگوریتم SVM-GA، بهترین مقدار تابع برازش برابر با ۰/۲۰۷ به‌دست آمد و بر این اساس، ۱۲ معیار بارش، فاصله تا جاده، شاخص نرمال شده تفاوت پوشش گیاهی، مدل رقومی ارتفاعی، جهت شیب، زاویه شیب، سنگ‌شناسی، شاخص توان آبراهه، فاصله تا رودخانه، شاخص رطوبت توپوگرافی، جنس خاک و کاربری زمین به‌عنوان معیارهای مؤثر در پیش‌بینی حساسیت سیلاب منطقه مورد مطالعه شناخته شدند. هم‌چنین برای الگوریتم RF-GA، بهترین مقدار تابع برازش برابر با ۰/۱۵۳ به‌دست آمد و بر این اساس، ۱۴ معیار بارش، فاصله تا جاده، شاخص نرمال شده تفاوت پوشش گیاهی، مدل رقومی ارتفاعی، جریان تجمعی، جهت شیب، زاویه شیب، سنگ‌شناسی، شاخص توان آبراهه، فاصله تا رودخانه، تراکم

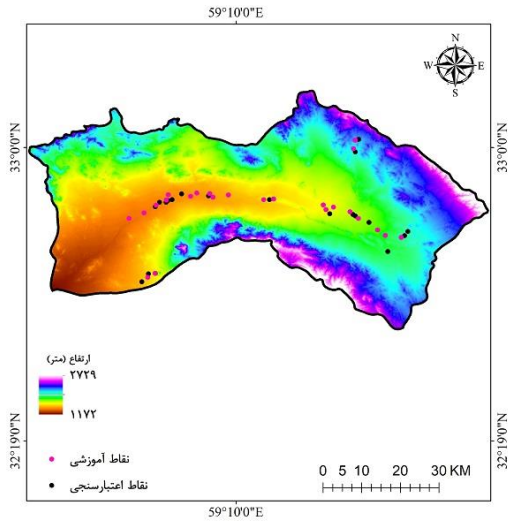
²⁹Receiver operating characteristics curve



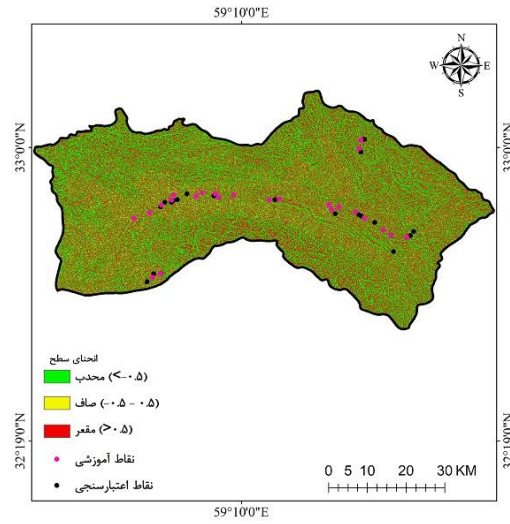
الف



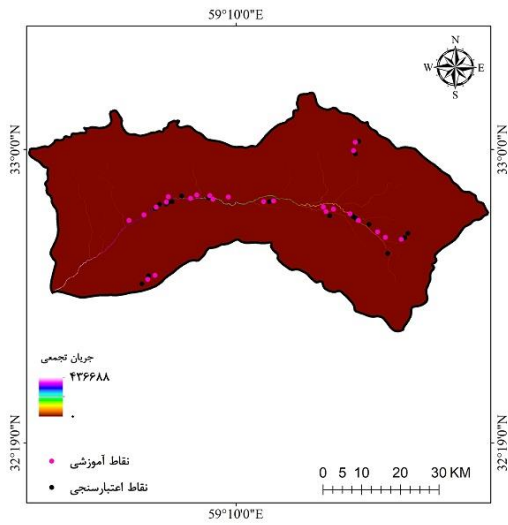
ب



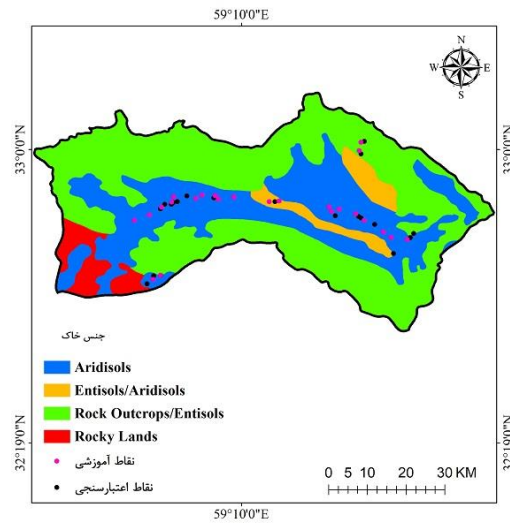
ج



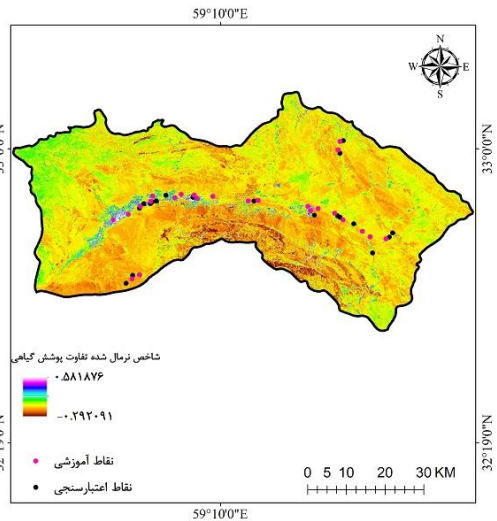
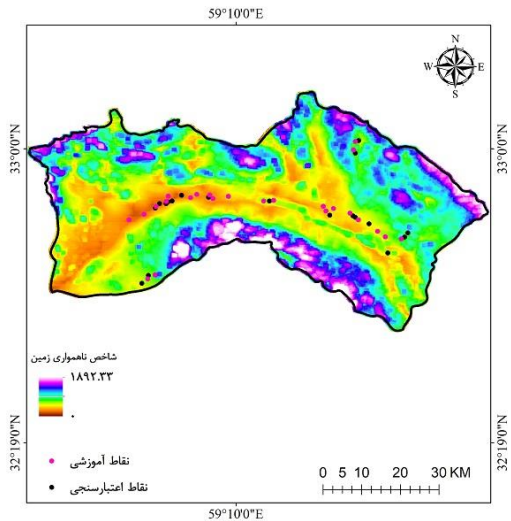
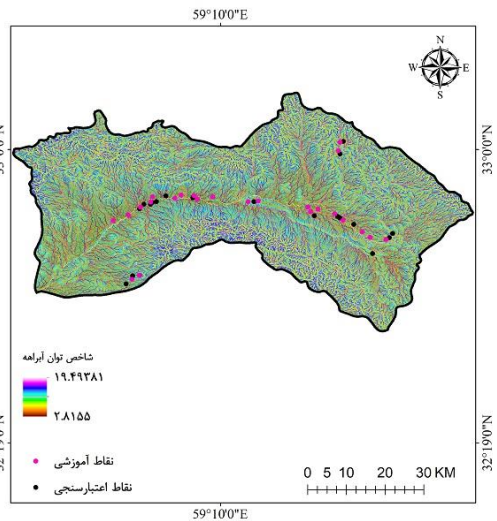
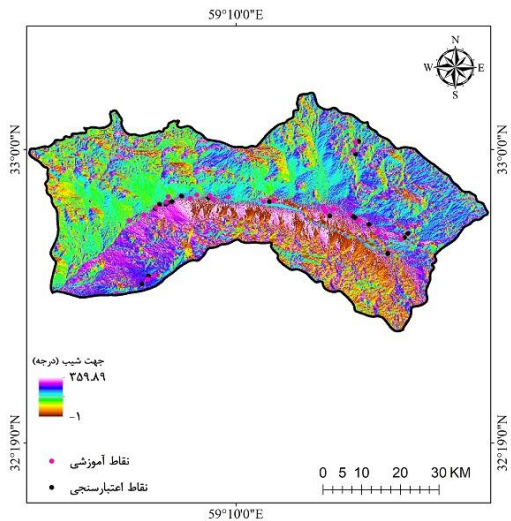
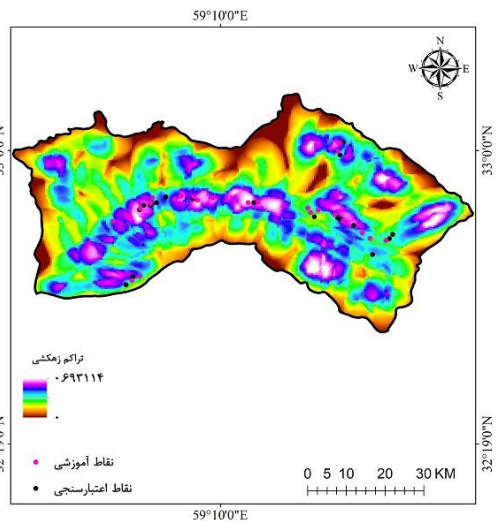
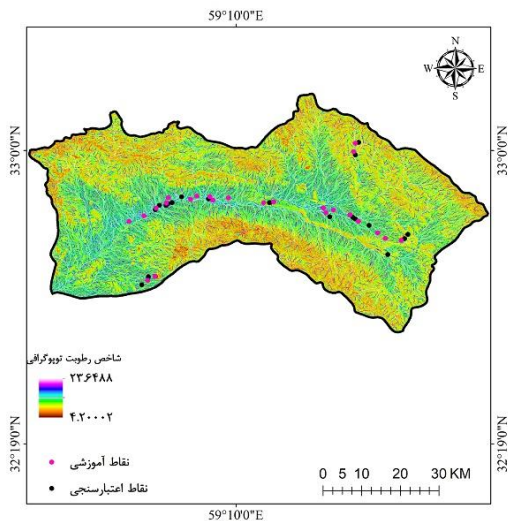
د

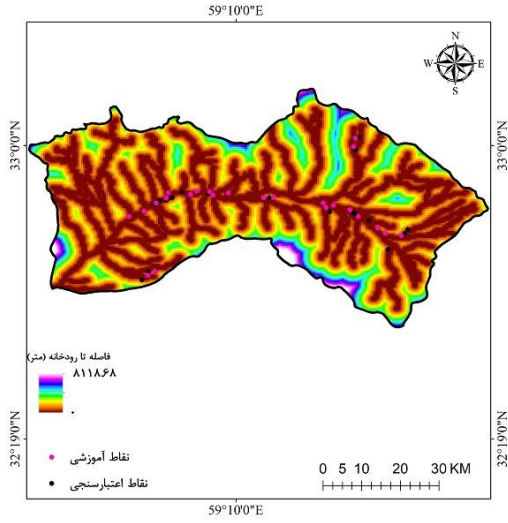


ه

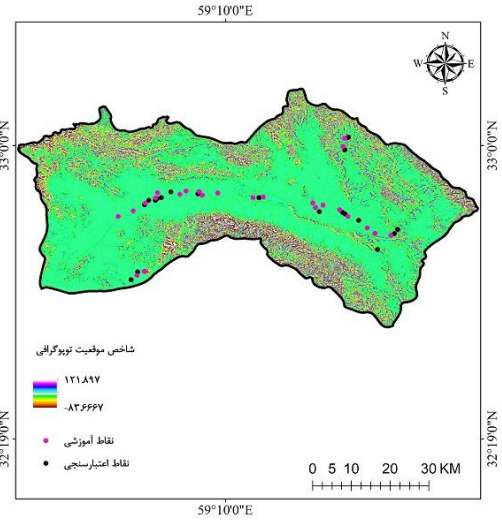


و

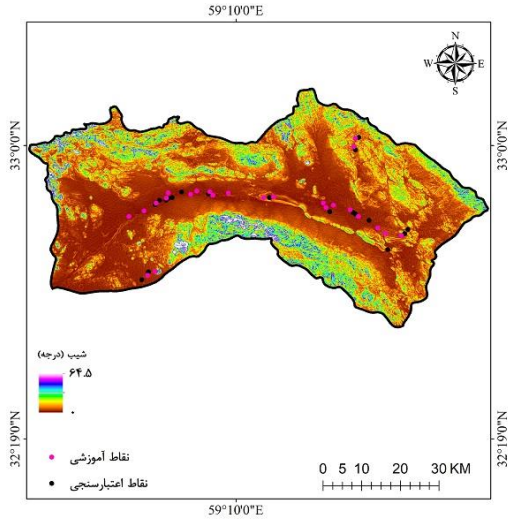




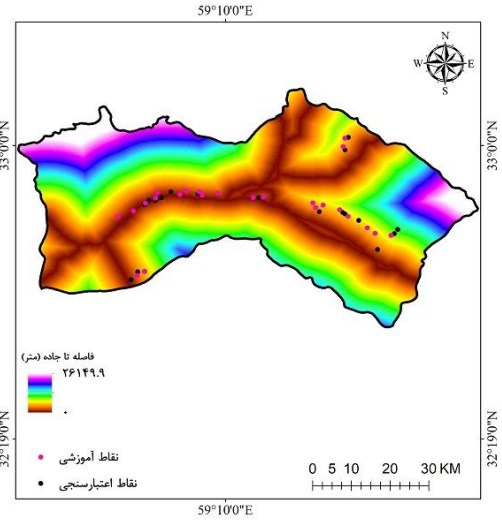
ژ



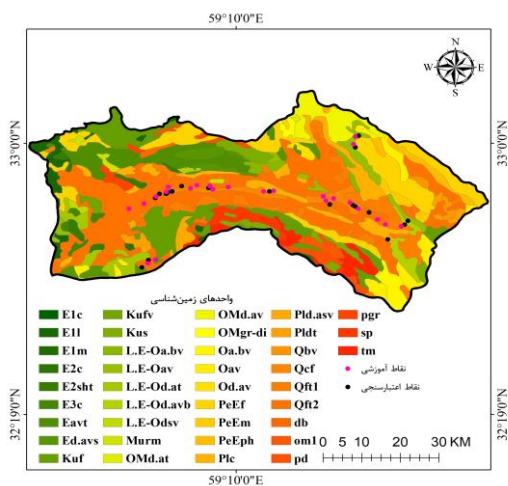
ز



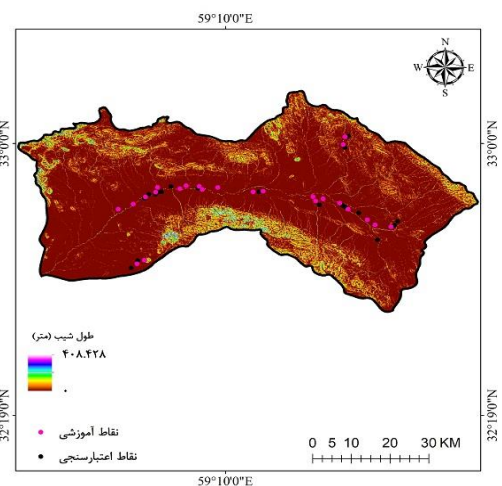
ش



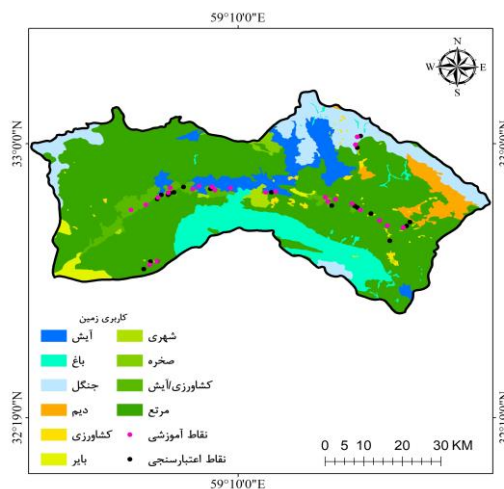
س



ض



ص



ط

شکل ۴. (الف) فاصله تا گسل، (ب) بارش، (پ) انحنای سطح، (ت) ارتفاع، (ث) جنس خاک، (ج) جریان تجمعی، (چ) تراکم زهکشی، (ح) TWI، (خ) SPI، (د) جهت شیب، (ذ) NDVI، (ر) TRI، (ز) TPI (ژ) فاصله تا رودخانه، (س) فاصله تا جاده، (ش) شیب، (ص) طول شیب، (ض) واحدهای زمین‌شناسی، (ط) کاربری اراضی

جدول ۱. نتایج آزمایش هم‌خطی چندگانه بین معیارها

شماره	معیار	VIF	TOL	شماره	معیار	VIF	TOL
۱	بارش	۲/۲۵	۰/۷۸	۱۰	زاویه شیب	۲/۳۹	۰/۸۵
۲	فاصله تا جاده	۱/۹۵	۰/۲۳	۱۱	سنگ‌شناسی	۴/۷۳	۰/۷۵
۳	شاخص نرمال شده پوشش گیاهی	۲/۶۵	۰/۷۸	۱۲	شاخص توان آبراهه	۰/۲۳	۰/۲۵
۴	طول شیب	۵/۸۲	۰/۰۹	۱۳	شاخص ناهمواری زمین	۱/۹۹	۰/۸۵
۵	فاصله تا گسل	۳/۲۳	۰/۴۵	۱۴	فاصله تا رودخانه	۱/۸۵	۰/۱۹
۶	مدل رقومی ارتفاع	۲/۸۵	۰/۶۴	۱۵	تراکم زهکشی	۳/۸۵	۰/۴۹
۷	انحنای سطح	۳/۲۳	۰/۱۱	۱۶	شاخص رطوبت توپوگرافی	۲/۷۸	۰/۵۵
۸	جریان تجمعی	۲/۲۷	۰/۱۹	۱۷	شاخص موقعیت توپوگرافی	۶/۲۳	۰/۰۲۵
۹	جهت شیب	۱/۵۸	۰/۹۵	۱۸	جنس خاک	۲/۷۵	۰/۴۴

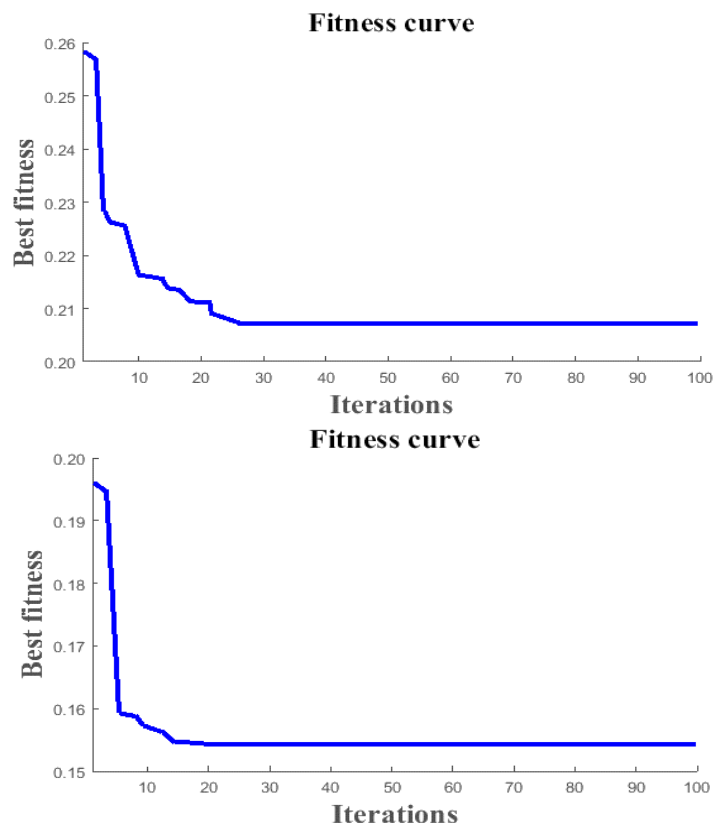
جدول ۲. پارامترهای مورد استفاده در الگوریتم ژنتیک

پارامتر	مقدار
اندازه جمعیت	۲۰
تعداد نسل‌ها (تکرار)	۱۰۰
نرخ ترکیب	۰/۸
نوع ترکیب	تک نقطه‌ای
نرخ جهش	۰/۷
نخبه‌گرایی	۱

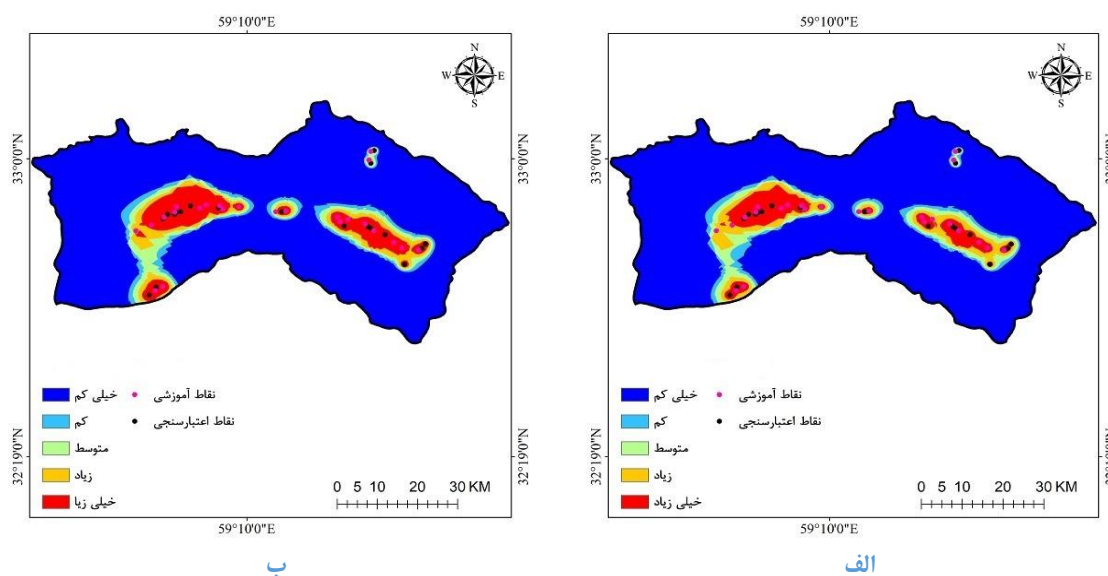
زهکشی، شاخص رطوبت توپوگرافی، جنس خاک و کاربری زمین به‌عنوان معیارهای مؤثر در پیش‌بینی حساسیت سیلاب شناخته شدند (شکل ۵)

تهیه نقشه حساسیت وقوع سیلاب: این نقشه‌ها در ۵ طبقه برابر براساس روش فاصله مساوی^{۳۰} در محدوده [۰ و ۱] ایجاد شد. شکل ۶ نشان‌دهنده نقشه حساسیت وقوع سیلاب در منطقه مورد مطالعه با استفاده از الگوریتم‌های پیشنهادی در ۵ طبقه

³⁰Equal Interval



شکل ۵. نمودار همگرایی (الف) SVM-GA (ب) RF-GA



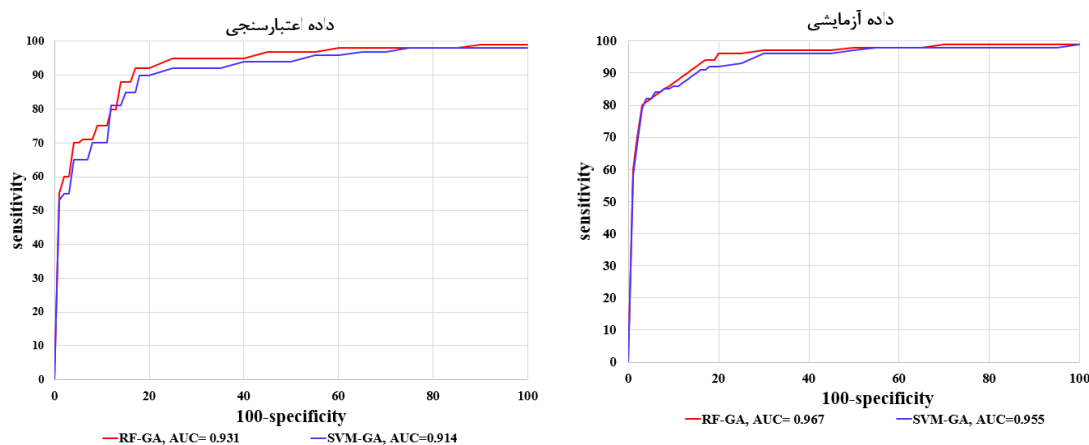
شکل ۶. نقشه حساسیت وقوع سیلاب (الف) SVM-GA (ب) RF-GA

برابر است. همچنین، در نواحی مرکزی منطقه مورد مطالعه احتمال وقوع خسارات به دلیل کاهش سطح اراضی جنگلی و تغییر کاربری‌های شدید صورت گرفته در چند سال اخیر که سبب افزایش اراضی دیم و شهری شده، را افزایش داده است.

عملکرد مدل‌های حساسیت سیل: مقادیر R^2 و RMSE حاصل از الگوریتم‌های پیشنهادی با استفاده از داده‌های اعتبارسنجی محاسبه شده‌اند که در جدول ۳ نشان داده شده است. براساس نتایج الگوریتم RF-GA دقت و سازگاری بالاتری را نسبت به هسته گوسین در پهنه‌بندی مناطق سیل خیز منطقه مورد مطالعه داشته است.

جدول ۳. مقادیر R^2 و RMSE حاصل از روش‌های SVM-GA و RF-GA

مدل	R^2	RMSE
SVM-GA	۰/۹۰۳۲	۰/۲۰۷
RF-GA	۰/۹۸۲۳	۰/۱۵۳



شکل ۷. منحنی ROC و مقدار AUC برای روش پیشنهادی تحقیق (الف) مجموعه داده‌های آموزشی (ب) مجموعه داده‌های اعتبارسنجی

شکل ۷ منحنی ROC را برای مجموعه داده‌های آموزشی و اعتبارسنجی مدل‌ها نشان می‌دهد. شکل ۷ (الف) نشان می‌دهد که برای مجموعه داده آموزشی مدل RF-GA دارای بالاترین مقدار AUC (۰/۹۶۷) و پس از آن مدل SVM-GA (۰/۹۵۵) قرار دارند. به طور کلی، مجموعه داده‌های اعتبارسنجی قابلیت مدل‌ها در پیش‌بینی حساسیت سیل را ارزیابی می‌کنند. بر این اساس شکل ۷ (ب) نشان می‌دهد که مدل RF-GA قوی‌تر بوده و دارای بالاترین دقت (AUC= ۰/۹۳۱) نسبت به مدل SVM-GA (AUC= ۰/۹۱۴) است. بنابراین، می‌توان نتیجه گرفت که اگر چه تمامی مدل‌های مورد استفاده از قدرت پیش‌بینی خوبی برخوردارند، اما عملکرد مدل RF-GA برای حساسیت سیل در منطقه مورد مطالعه مناسب‌تر است.

بحث

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 ایران، ارزیابی حساسیت سیل در مناطق آسیب‌پذیر کشور اقدامی است که نمی‌توان آن را حذف و به تعویق انداخت. با پهنه‌بندی مناطق حساس به سیل می‌توان در جهت پیشگیری هرچه بیشتر این پدیده و مقابله با معیارهای محرک افزایش میزان سیلاب گام برداشت. جهت رسیدن به هدف این تحقیق، از الگوریتم‌های یادگیری ماشین RF و SVM در ترکیب با الگوریتم ژنتیک جهت پهنه‌بندی حساسیت سیلاب در منطقه مورد مطالعه براساس تعیین معیارهای مؤثر استفاده شد. نتایج نشان داد که هسته مکعبی سه‌گانه نتایج بهتری را نسبت به هسته گوسین در اختیار قرار می‌دهد. مقدار R^2 حاصل از الگوریتم‌های RF-GA و SVM-GA به ترتیب برابر ۰/۹۰۳۲ و ۰/۹۸۲۳ به دست آمد که نشان‌دهنده سازگاری بالای الگوریتم RF با داده‌های این تحقیق می‌باشد. همچنین مقدار RMSE حاصل از الگوریتم‌های RF-GA و SVM-GA به ترتیب برابر ۰/۲۷۵۱ و ۰/۲۳۲۱ به دست آمد که نشان‌دهنده دقت بالای الگوریتم RF است. نتایج نهایی نشان داد که مدل RF نسبت به مدل SVM در پیش‌بینی حساسیت سیل عملکرد بهتری دارد و قابلیت بالای مدل RF توسط مطالعات مرتبط دیگر مورد تأیید است. (Choubin *et al.*, 2019). مطابق نقشه حساسیت سیل دشت بیرجند، احتمال وقوع سیل در مناطق مرکزی دشت، به دلیل ارتفاع و زاویه شیب کمتر، بیشتر از سایر مناطق است که مطابق با نتایج مطالعات Kanani-Sadat و همکاران (۲۰۱۹) و Chen و همکاران (۲۰۱۹) است. در واقع، در مقایسه با مدل SVM، مدل RF یک الگوریتم ترکیبی است که می‌تواند نتایج رضایت‌بخش‌تری را با نرخ‌های دقت بالاتر و واریانس‌ها و سوگیری‌های کمتر تولید کند (Hong *et al.*, 2018; Chen *et al.*, 2019).

(2019). در مورد معیار فاصله، به طور کلی با افزایش فاصله از رودخانه، احتمال وقوع سیل کاهش می یابد. ارتفاع از سطح دریا به عنوان یکی از مهم ترین معیارهای مؤثر در حساسیت سیل است، زیرا که سیل در ارتفاعات پایین تر از سطح دریا رخ می دهد و احتمال وقوع آن در ارتفاعات بالاتر بسیار کم است. با بررسی شاخص بارش می توان نتیجه گرفت که با افزایش ارتفاع، احتمال بارندگی افزایش می یابد درحالی که احتمال وقوع سیل در ارتفاعات کمتر است. کاربری اراضی نیز یکی از معیارهای مؤثر بر وقوع سیل است؛ زیرا نوع کاربری زمین در نفوذ رواناب و یا تسریع در سرعت آن مؤثر است. به دنبال تغییر معیار کاربری اراضی ناشی از فعالیت های انسانی، ظرفیت آبراهه ها نیز کاهش یافته و در نتیجه به عنوان معیار مهمی بر وقوع سیل دلالت دارد. این نتایج با نتایج پژوهش های Khosravi و همکاران (۲۰۱۶)، Liu و همکاران (۲۰۱۶)، Rahmati و همکاران (۲۰۱۸) و Razavi Termeh و همکاران (۲۰۱۸) مطابقت دارد. با توجه به این که روش های ذکر شده داده محور می باشند و در این بررسی حوضه آبریز دشت بیرجند مورد مطالعه قرار گرفته است، بنابراین پیشنهاد می شود این الگوریتم ها برای سایر مناطق با ویژگی های جغرافیایی مشابه و متفاوت نیز بکار گرفته شوند و نتایج مربوط به دقت و ارزیابی روش ها در پیش بینی حساسیت سیل مشخص شود و بررسی شود تا چه اندازه ای نوع داده ها و منطقه مورد مطالعه می تواند روی دقت و خروجی مدل ها اثر بگذارد. همان طور که ذکر شد، روش های ادغام به دو دسته کلی دانش محور و داده محور تقسیم بندی می شوند که در این تحقیق با توجه به در دسترس بودن داده های مرجع، از رویکرد ادغام داده محور جهت پهنه بندی حساسیت مناطق سیل خیز استفاده شد. بنابراین پیشنهاد می شود در تحقیقات آتی از رویکرد ادغام دانش محور جهت پهنه بندی مناطق سیل خیز استفاده گردد. همچنین در این تحقیق از ترکیب الگوریتم ژنتیک با الگوریتم RF و SVM برای شناسایی معیارهای تأثیرگذار در پهنه بندی حساسیت سیلاب استفاده شد. بنابراین برای تحقیقات آتی می توان توانایی سایر الگوریتم های تکاملی مانند الگوریتم انبوه ذرات و کلونی زنبور جهت شناسایی معیارهای تأثیرگذار مورد ارزیابی قرار داد.

References

- Ahmadi, M., Karimi, M., Alizadeh, S., Shirzadi, A., Parvinnejhad, D., Shahabi, H., Panahi, M., 2019. 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.
- 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.
- 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.
- 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.
- 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.
- 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.
- Eftekhari, M., Eslaminezhad, S., Haji Elyasi, A., Akbari, M., 2021. Predicting Groundwater Potential Areas Using Hybrid Artificial Intelligence Methods (Case Study: Birjand Plain). *Iranian Journal of Soil and Water Research* 52(9), 2383-2397. (In Persian)
- Eini, M., Kaboli, H.S., Rashidian, M., Hedayat, H., 2020. Hazard and vulnerability in urban flood risk mapping: Machine learning techniques and considering the role of urban districts. *International Journal of Disaster Risk Reduction* 50, 101687.
- Eslaminezhad, S.A., Eftekhari, M., Akbari, M., 2020. GIS-Based Flood Risk Zoning Based On Data-Driven Models. *Journal of Hydraulic Structures* 6(4), 75-98.

- 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)
- Eslaminezhad, S.A., Omarzadeh, D., Eftekhari, M., Akbari, M., 2021b. Development of a data-driven model to predict landslide sensitive areas. *Geographia Technica* 16(1).
- Eslaminezhad, S.A., Eftekhari, M., Azma, A., Kiyanfar, R., Akbari, M., 2022a. Assessment of flood susceptibility prediction based on optimized tree-based machine learning models. *Journal of Water and Climate Change* 13(6), 2353-2385.
- Eslaminezhad, S.A., Eftekhari, M., Akbari, M., Bayat, H., Barghi, W., 2022b. Using Boosted Regression Tree, Logistic Model Tree, and Random Forest Algorithms to Evaluate the Groundwater Potential. *Watershed Management Research Journal* 35(3), 44-59.
- Fotheringham, A.S., Oshan, T.M., 2016. Geographically weighted regression and multicollinearity: dispelling the myth. *Journal of Geographical Systems* 18(4), 303-329.
- Ghavami, Z., Mohamadnia, A., 2017. Spatial forecasting of flood-prone areas using Geographic Information System (GIS). 4th national conference on Application of GIS in Water and Electric Industries, Arak. 13 & 14 December 2017. (In Persian)
- 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.
- Guevara, J., Zadrozny, B., Buoro, A., Lu, L., Tolle, J., Limbeck, J., Wu, M., Hohl, D., 2018. A hybrid data-driven and knowledge-driven methodology for estimating the effect of completion parameters on the cumulative production of horizontal wells. In *SPE Annual Technical Conference and Exhibition*. OnePetro.
- 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.
- Hudson, P., Botzen, W.W., 2019. Cost-benefit analysis of flood-zoning policies: A review of current practice". *Wiley Interdisciplinary Reviews: Water* 6(6), e1387.
- 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.
- 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.
- 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.
- Kanani-Sadat, Y., Arabsheibani, R., Karimipour, F., Nasser, 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.
- Khosravi, K., Maroufinia, E., Nohani, E., Chapy, K., 2017. Evaluation of Logistic Regression Efficiency in Mapping Flood Susceptibility. *Journal of Range and Watershed Management* 69(4), 863-876. (In Persian)
- 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. *Nat. Hazards* 83(2), 947-987.
- 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.
- Mirjalili, S., 2019. Genetic algorithm. In *Evolutionary algorithms and neural networks*. pp. 43-55. Springer, Cham.
- Mojaddadi, H., Pradhan, B., Nampak, H., Ahmad, N., and 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.

- 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* 590, 125275.
- Papaoiannou, G., Vasiliades, L., Loukas, A., 2015. Multi-criteria analysis framework for potential flood prone areas mapping. *Water resources management* 29(2), 399-418.
- Pourghasemi, H.R., Razavi-Termeh, S.V., Kariminejad, N., Hong, H., Chen, W., 2020. An assessment of metaheuristic approaches for flood assessment. *Journal of Hydrology* 582, 124536.
- Quiroz, J.C., Mariun, N., Mehrjou, M.R., Izadi, M., Misron, N, and Mohd Radzi, M.A., 2018. Fault detection of broken rotor bar in LS-PMSM using random forests. *Measurement* 116, 273-280.
- 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.
- Razavi Termeh, S.V., Kornejady, 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.
- Sun, Y., Xue, B., Zhang, M., Yen, G.G., Lv, J., 2020. Automatically designing CNN architectures using the genetic algorithm for image classification. *IEEE Transactions on Cybernetics* 50(9), 3840-3854.
- 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.
- 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.
- Xiao, Y., Yi, S., Tang, Z., 2017. Integrated flood hazard assessment based on spatial ordered weighted averaging method considering spatial heterogeneity of risk preference. *Science of the Total Environment* 599, 1034-1046.
- Zeraatkar, Z., Hassanpour, F., 2016. Simulation of BirjandUrban FloodUsing HEC-RAS and ARC-GIS. *Watershed Management Research Journal* 29(3), 41-56. (In Persian)
- 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. *Scientific Journals Management System* 20 (57), 1-24. (In Persian)