



Research Article

Vol. 37, No. 6, Feb.-Mar., 2024, p. 841-853

Flood Susceptibility Mapping of the Famnat Watershed, Gilan Province

F. Mirchooli¹, I. Gholami², M. Boroughani^{3*}

1- Ph.D. in Watershed Management, Expert in Watershed Management Department, Sari University of Agricultural Sciences and Natural Resources, Sari, Iran

2- Master Graduate in Watershed Management, Department of Watershed Management, Faculty of Natural Resources and Marine Sciences, Tarbiat Modares University, Noor, Iran

3- Assistant Professor, Research Center for Geosciences and Social Studies, Hakim Sabzevari University, Sabzevar, Iran

(* - Corresponding Author Email: m.boroughani@hsu.ac.ir)

Received: 29-08-2023
Revised: 13-11-2023
Accepted: 03-12-2023
Available Online: 03-12-2023

How to cite this article:

Mirchooli, F., Gholami, I., & Boroughani, M. (2024). Flood susceptibility mapping of the Famnat watershed, Gilan province. *Journal of Water and Soil*, 37(6), 841-853. (In Persian with English abstract). <https://doi.org/10.22067/jsw.2023.84146.1328>

Introduction

Flood is one of the most destructive natural disasters that has a negative impact on social, economic and environmental dimensions. Floods usually occur following a prolonged period of rain or snowmelt in combination with unfavorable conditions. In this regard, all over the world, the occurrence of floods has intensified by 40% in the last two decades. In Asia, almost 90% of all human casualties caused by natural disasters are due to floods. The increase in flooding is usually due to increased environmental degradation such as urbanization, increased population growth, and deforestation. Periodic and regular occurrences of floods over a certain timeframe significantly amplify the detrimental impacts on living organisms. Urban areas in close proximity to rivers bear the brunt of these damages, owing to high population density, economic infrastructure, and transportation networks. However, these consequences can be alleviated through meticulous vulnerability analysis. One of the primary objectives pursued by researchers and policymakers is the precise modeling and zoning of floods to mitigate associated risks. Consequently, a myriad of methods and approaches have been devised for flood risk modeling and zoning to address this pressing issue. Among them, hydrological methods such as rainfall-runoff modeling and data-based techniques, which are unable to comprehensively analyze rivers and flood zones due to their one-dimensional nature. This is despite the fact that the morphology of the river is not stable and due to its high erosion potential, it also has a dynamic characteristic. In addition, these methods require fieldwork and large budgets for data collection. Hence, comprehensive flood management is necessary to reduce these effects. Therefore, this study was conducted with the aim of identifying areas sensitive to the risk of flooding in Famnat watershed located in Gilan province. Fomanat watershed is located in Gilan province and is considered a part of the first grade watershed of the Central Plateau. This area is located in the range of 36.89 to 37.57 degrees north latitude and 48.77 to 49.69 degrees east longitude. This region has an area of 3595 square kilometers, the highest point of which is 3088 meters and the lowest point is -69 meters.

Materials and Methods

To carry out the current research, firstly, by reviewing the sources and history of the research, as well as knowing the region, a map and layers of information related to the factors affecting flood susceptibility zoning were prepared. These layers include land use map, slope degree, geology, distance from waterway, digital map of height, direction,



©2023 The author(s). This is an open access article distributed under [Creative Commons Attribution 4.0 International License \(CC BY 4.0\)](https://creativecommons.org/licenses/by/4.0/), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source.

<https://doi.org/10.22067/jsw.2023.84146.1328>

shape of land curvature, land curvature profile, rainfall and topographic humidity index, which are created using the collected data and also various additions in the environment. Geographic information system (Arcgis 10.4) was prepared. In this regard, machine learning models such as generalized linear model (GLM), multivariate adaptive regression model (MARS) and classification and regression tree model (CART) were used to zone the sensitivity of the watershed to floods. Also, among 100 flood events, 70% (70) were considered for training and 30% (30) for validation. In the following, using field survey and review of previous studies, 10 factors influencing the occurrence of floods in the watershed area were identified and used. Finally, the area under the ROC curve and the TSS index were used to evaluate the models.

Results and Discussion

The results of the evaluation of the most important factors affecting the sensitivity of the watershed to floods indicated that the distance from the river, the height and the curvature profile had the greatest impact on the sensitivity of the region, and on the other hand, the factors of slope, geology and topographic humidity index had the least impact. Based on the obtained results, the areas covered by very low, low, medium, high and very high classes in the CART model were 26.6, 17.6, 21.2, 0.1 and 34.0%, respectively. These results for the GLM model were 13.6, 12.7, 16.2, 25.1 and 32.4 percent, respectively. Based on the obtained results, the CART model performed better than other models, so that AUC for MARS model was equal to 0.76, CART model was equal to 0.9 and GLM model was equal to 0.84. Also, the better performance of CART model compared to other models was confirmed by other indicators. So, based on TSS, MARS model equal to 0.52, CART model equal to 0.77 and GLM model equal to 0.66 were obtained.

Conclusion

Implementing the findings of this study can facilitate the adoption of effective management strategies to mitigate losses and casualties. Moreover, in developing nations grappling with restricted access to hydrogeological and soil data, the utilization of geographic information systems (GIS) and data mining techniques assumes a pivotal role in conducting comprehensive studies. These technologies offer valuable insights and support decision-making processes, enabling proactive measures to address flood risks and enhance disaster resilience in vulnerable regions.

Keywords: Flood management, Flood mapping, Machine learning models, ROC curve

مقاله پژوهشی

جلد ۳۷، شماره ۶، بهمن-اسفند ۱۴۰۲، ص. ۸۵۳-۸۴۱

پهنه‌بندی حساسیت به سیل در حوزه آبخیز فامانات، استان گیلان

فهیمه میرچولی^۱ - عیسی غلامی^۲ - مهدی بروغنی^۳*

تاریخ دریافت: ۱۴۰۲/۰۶/۰۷

تاریخ پذیرش: ۱۴۰۲/۰۹/۱۲

چکیده

سیل یکی از مخرب‌ترین بلایای طبیعی است که تأثیر منفی بر ابعاد اجتماعی، اقتصادی و محیط‌زیستی دارد. افزایش سیل معمولاً به دلیل افزایش تخریب محیط‌زیست مانند توسعه شهرنشینی، افزایش رشد جمعیت و جنگل‌زدایی است. از این رو، مدیریت همه‌جانبه سیل برای کاهش این اثرات ضروری است. بنابراین، این مطالعه با هدف شناسایی مناطق حساس به خطر وقوع سیل در حوزه آبخیز فامانات واقع در استان گیلان انجام شد. در این راستا، از مدل‌های یادگیری ماشین از جمله مدل خطی تعمیم‌یافته (GLM)، مدل رگرسیون تطبیقی چندمتغیره (MARS) و مدل درخت‌های طبقه‌بندی و رگرسیون (CART) برای پهنه‌بندی حساسیت حوزه آبخیز به سیلاب استفاده شد. همچنین، از میان ۱۰۰ رخداد سیلاب، ۷۰ درصد آن (۷۰) برای آموزش و ۳۰ درصد (۳۰) برای اعتبارسنجی در نظر گرفته شد. در ادامه، با استفاده از پیمایش صحرایی و نیز مرور مطالعات پیشین، ۱۰ عامل تأثیرگذار بر وقوع سیل در حوزه آبخیز شناسایی و مورد استفاده قرار گرفت. در نهایت برای ارزیابی مدل‌ها، از سطح زیر منحنی ROC و نیز شاخص TSS استفاده شد. نتایج حاصل از مطالعه نشان داد که از میان متغیرهای تأثیرگذار، عامل ارتفاع و فاصله از رودخانه مؤثرترین عوامل در حوزه آبخیز مطالعاتی می‌باشند. به علاوه، سطح زیر منحنی در مدل MARS معادل ۰/۷۶، مدل GLM معادل ۰/۸۳ و در مدل CART معادل ۰/۹ بوده است که نشان‌دهنده عملکرد بهتر مدل CART در مقایسه با سایر مدل‌ها است. با به‌کارگیری نتایج حاصل از این مطالعه می‌توان اقدامات مدیریتی مناسبی جهت کاهش خسارات و تلفات اتخاذ کرد. به علاوه، در کشورهای در حال توسعه که با محدودیت دسترسی به اطلاعات هیدروژئولوژیکی و ادافیکی مواجه هستند، استفاده از سیستم اطلاعات جغرافیایی و نیز تکنیک‌های داده‌کاوی نقش مهمی در انجام مطالعات ایفا می‌کنند.

واژه‌های کلیدی: پهنه‌بندی سیل، مدل‌های یادگیری ماشین، مدیریت سیل، منحنی ROC

مقدمه

در سراسر جهان، وقوع سیل طی دو دهه گذشته ۴۰ درصد تشدید شده است؛ به طوری که در آسیا، تقریباً ۹۰ درصد از کل تلفات انسانی ناشی از بلایای طبیعی، به دلیل سیل است (Pourghasemi et al., 2020). افزایش سیل معمولاً به دلیل افزایش تخریب محیط‌زیست مانند توسعه شهرنشینی، افزایش رشد جمعیت و جنگل‌زدایی است. اگر سیل در یک بازه زمانی به‌طور متناوب و منظم رخ دهد، خسارات وارده نیز با شدت بیشتری بر زندگی موجودات زنده تأثیر می‌گذارد. این خسارات به خصوص در مناطق شهری نزدیک به رودخانه به دلیل تراکم جمعیت،

بلایای طبیعی عامل اصلی خسارات جبران‌ناپذیر در سراسر جهان هستند (Tehrany et al., 2014). سیل یکی از مخاطرات طبیعی خطرناکی است که بیش از هر رویداد طبیعی دیگری، مردم جهان را تحت تأثیر قرار می‌دهد (Rahmati et al., 2016). سیل معمولاً به دنبال یک دوره طولانی بارندگی یا ذوب برف در ترکیب با شرایط نامطلوب اتفاق می‌افتد (Nachappa et al., 2020). در همین راستا،

۱- دکتری آبخیزداری، گروه آبخیزداری، دانشگاه علوم کشاورزی و منابع طبیعی ساری، ساری، ایران

۲- دانش‌آموخته کارشناسی ارشد آبخیزداری، گروه آبخیزداری، دانشکده منابع طبیعی و علوم دریایی، دانشگاه تربیت مدرس، نور، ایران

۳- استادیار مرکز پژوهشی علوم جغرافیا و مطالعات اجتماعی، دانشگاه حکیم سبزواری، سبزوار، ایران

*- نویسنده مسئول: (Email: m.boroughani@hsu.ac.ir)

فاکتورهای مورد بررسی بوده و تأثیر سیل بر هدررفت گندم و برنج از شمال حوضه به جنوب، به تدریج کاهش می‌یابد. ناچایا و همکاران (Nachappa et al., 2020)، نیز با استفاده از الگوریتم‌های یادگیری ماشین یعنی جنگل تصادفی و ماشین بردار پشتیبان به پهنه‌بندی خطرات سیل و زمین‌لغزش پرداختند. نتایج حاصل از مطالعه ایشان نشان داد که مدل جنگل تصادفی برای هر دو نوع خطر، بهترین دقت را داشته است. آوند و همکاران (Avand et al., 2021) در مطالعه خود به تهیه نقشه احتمال سیل‌خیزی در حوزه آبخیز تچن پرداختند. ایشان در مطالعه خود، از الگوریتم‌های تحلیل ممیزی انعطاف‌پذیر و نیز شبکه عصبی مصنوعی استفاده کردند. نتیجه حاصل از مطالعه ایشان نشان داد که مدل تحلیل ممیزی انعطاف‌پذیر دقت بالاتری در تهیه این نقشه داشته است. آوند و همکاران (Avand et al., 2022)، به ارزیابی اثر دقت مکانی مدل رقومی ارتفاع (DEM) بر پهنه‌بندی خطر سیل پرداختند. نتایج پژوهش ایشان نشان داد که دقت مکانی DEM به تنهایی نمی‌تواند تأثیر قابل توجهی بر دقت پهنه‌بندی خطر سیل داشته باشد. سیدی و همکاران (Seydi et al., 2022)، به پهنه‌بندی آسیب‌پذیری سیل با استفاده از مدل جنگل آبخیزی و نیز مقایسه دقت آن با سایر مدل‌ها از جمله مدل‌های جنگل تصادفی، الگوریتم تقویت گرادیان سبک، درخت تصمیم‌گیری، ماشین بردار پشتیبان، شبکه عصبی مصنوعی، درخت‌های تصمیم تقویت‌شده با گرادیان پرداختند. نتایج نشان داد که مدل جنگل آبخیزی دارای بیشترین دقت می‌باشد. حسن‌الزمان و همکاران (Hasanuzzaman et al., 2022) به پهنه‌بندی آسیب‌پذیری در حوزه آبخیز Silabati واقع در کشور هند پرداختند. ایشان در مطالعه خود از مدل‌های جنگل تصادفی، بیز ساده و درخت‌های تصمیم تقویت‌شده با گرادیان استفاده کردند و به این نتیجه رسیدند که مدل جنگل تصادفی دارای بیشترین دقت در پهنه‌بندی سیل داشته است.

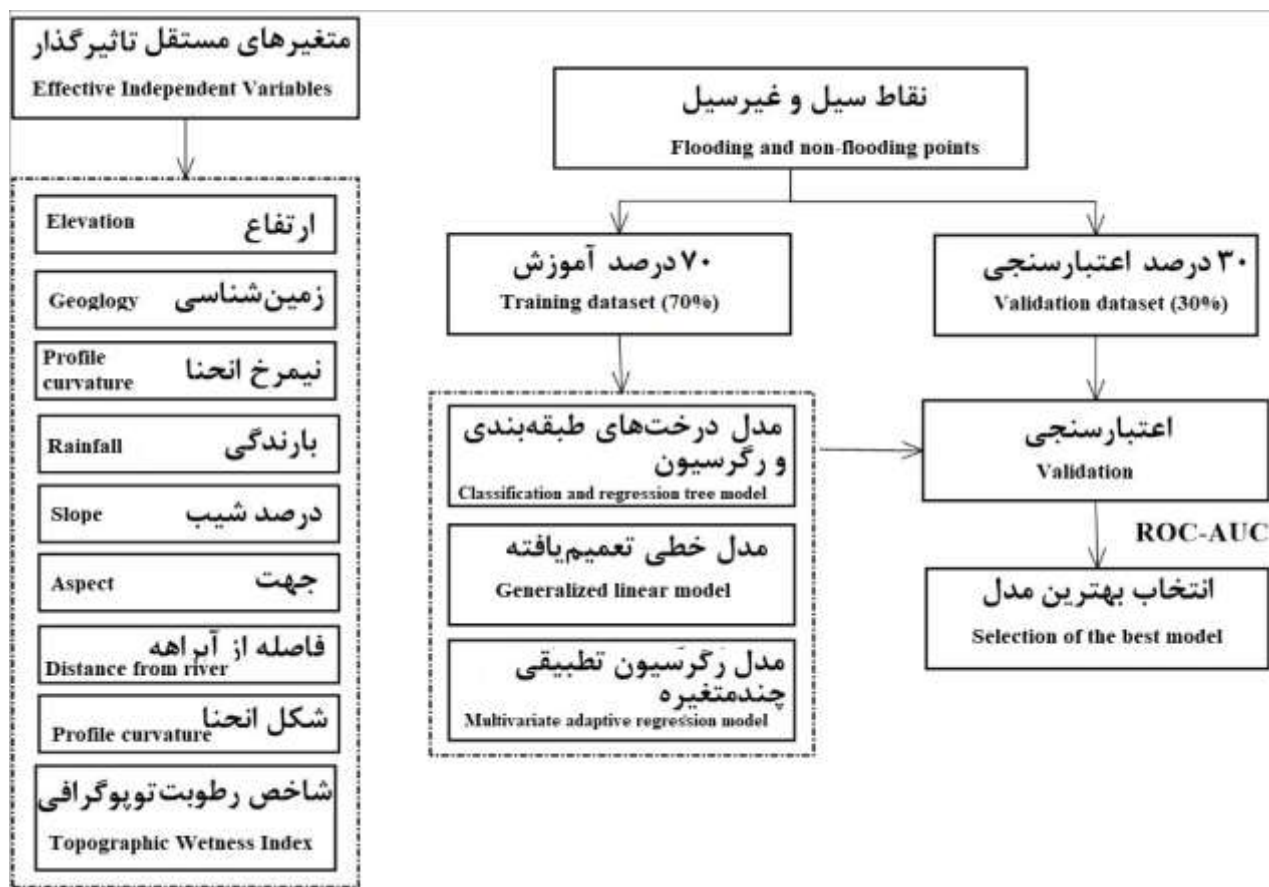
پژوهش حاضر، با هدف پهنه‌بندی خطر سیل در حوزه آبخیز فومنات واقع در استان گیلان، چارچوب‌بندی شده است. اهداف این پژوهش شامل (۱) تعیین مهم‌ترین عوامل محیط‌زیستی مؤثر بر خطر سیل (۲) به‌کارگیری الگوریتم‌های مدل خطی تعمیم‌یافته (GLM)، مدل رگرسیون تطبیقی چندمتغیره (MARS) در پهنه‌بندی خطر سیل و تعیین دقیق‌ترین مدل (۳) محاسبه مساحت تحت پوشش کلاس‌های مختلف خطر سیلاب در حوزه آبخیز مطالعاتی، می‌باشد.

مواد و روش‌ها

همانطور که در قسمت قبل اشاره شد، این مطالعه با استفاده از الگوریتم‌های یادگیری ماشین انجام شد. مراحل کلی این مطالعه در شکل ۱ نشان داده شده است.

زیرساخت‌های اقتصادی و حمل و نقل شدیدتر می‌باشد (Tehrany et al., 2015). این پیامدها را می‌توان از طریق تجزیه و تحلیل دقیق آسیب‌پذیری کاهش داد. از این‌رو، مدل‌سازی و پهنه‌بندی دقیق سیل یکی از اهداف اصلی محققان و سیاست‌گذاران برای کاهش خطر آن است.

در همین راستا، روش‌ها و رویکردهای مختلفی برای مدل‌سازی و پهنه‌بندی خطر سیل توسعه داده شده است. از جمله، روش‌های هیدرولوژیکی مانند مدل‌سازی بارش-رواناب و تکنیک‌های مبتنی بر داده‌ها که به دلیل تک‌بعدی بودن قادر به تجزیه و تحلیل جامع رودخانه‌ها و مناطق طغیان آب نیستند. این در حالی است که مورفولوژی رودخانه پایدار نیست و به دلیل پتانسیل فرسایشی بالا دارای ویژگی‌های پویایی نیز می‌باشد. به‌علاوه، این روش‌ها به کار میدانی و بودجه قابل توجهی برای جمع‌آوری داده‌ها نیاز دارند (Tehrany et al., 2014). تکنیک‌های سنجش از دور (RS) همراه با ابزار GIS می‌تواند برای تعیین مناطق خطر سیل، بستر مناسبی برای ترکیب، تجزیه و تحلیل اطلاعات فراهم کند. روش‌های مختلفی برای پهنه‌بندی خطر سیل با استفاده از GIS توسعه یافته است. از این میان، به‌کارگیری روش‌های مبتنی بر GIS و نیز یادگیری ماشین^۱ که از تجزیه و تحلیل انواع مختلف داده‌ها پشتیبانی می‌کنند، برای توسعه مدل‌های مخاطرات طبیعی و ایجاد مبنایی برای پهنه‌بندی آسیب‌پذیری و ریسک مورد توجه بسیاری از پژوهشگران قرار گرفته است (Skilodimou et al., 2019). از جمله مدل‌های یادگیری ماشین می‌توان به جنگل تصادفی (RF)، شبکه عصبی مصنوعی (ANN)، مدل‌های خطی تعمیم‌یافته (GLM)، رگرسیون تطبیقی چندمتغیره (MARS) اشاره کرد که علاوه بر مطالعات پهنه‌بندی سیل، در زمین‌لغزش، فرسایش، آب زیرزمینی نیز مورد استفاده قرار گرفته است. از جمله مطالعات مربوط به کاربرد این الگوریتم‌ها در خطر سیل می‌توان به موسوی و همکاران (Mosavi et al., 2022) اشاره کرد. ایشان به منظور اولویت‌بندی زیرحوضه‌های موجود در حوزه آبخیز تالار واقع در استان مازندران از نظر سیل و فرسایش، از مدل‌های خطی تعمیم‌یافته (GLM)، تحلیل‌های متمایز انعطاف‌پذیر (FDA)، رگرسیون تطبیقی چندمتغیره (MARS) و جنگل تصادفی (RF) استفاده کردند. نتایج آنها نشان داد که مدل‌های GLM، FDA، MARS و RF به ترتیب دارای بیشترین دقت مدل‌سازی بوده و زیرحوضه‌های سه و پنج دارای بیشترین حساسیت به خطر سیل و فرسایش می‌باشند. محمدی و همکاران (Mohammadi et al., 2020) با استفاده از GIS و برخی عوامل مؤثر مانند تجمع جریان، شیب، کاربری اراضی، شدت بارش، ارتفاع و زمین‌شناسی به پهنه‌بندی خطر سیل و هدررفت محصولات کشاورزی ناشی از آن پرداختند. نتایج حاصل از مطالعه ایشان نشان داد که ارتفاع مؤثرترین عامل از میان



شکل ۱- فلوجارت مراحل انجام مطالعه حاضر
Figure 1- Flowchart of the steps of the current study

2016). این لایه‌ها شامل نقشه کاربری اراضی، درجه شیب، زمین‌شناسی، فاصله از آبراهه، نقشه رقومی ارتفاع، جهت، شکل انحنا، زمین، نیمرخ انحنا، زمین، بارندگی و شاخص رطوبت توپوگرافی^۱ (TWI) می‌باشند. لایه مربوط به آبراهه، جهت، درصد شیب، شکل انحنا، زمین و نیمرخ انحنا، زمین در محیط Arcgis و با استفاده از نقشه رقومی ارتفاع (DEM) به دست آمد. نقشه کاربری اراضی از اداره منابع طبیعی تهیه شد.

با در نظر گرفتن تأثیر ارتفاع بر میزان رواناب، این متغیر به‌عنوان فاکتور مستقل و تأثیرگذار استفاده شد. ارتفاع تأثیر زیادی بر وقوع سیل دارد به طوری که طبقات ارتفاعی کمتر پتانسیل بیشتری برای وقوع سیل دارند. به علاوه، با توجه تأثیر مستقیم شیب بر رواناب سطحی و فرصت نفوذ، این فاکتور یکی از عوامل مؤثر بر وقوع سیل حوزه آبخیز در نظر گرفته می‌شود. فاصله از رودخانه یکی از مهم‌ترین عوامل در وقوع سیل به خصوص در اراضی مجاور می‌باشد.

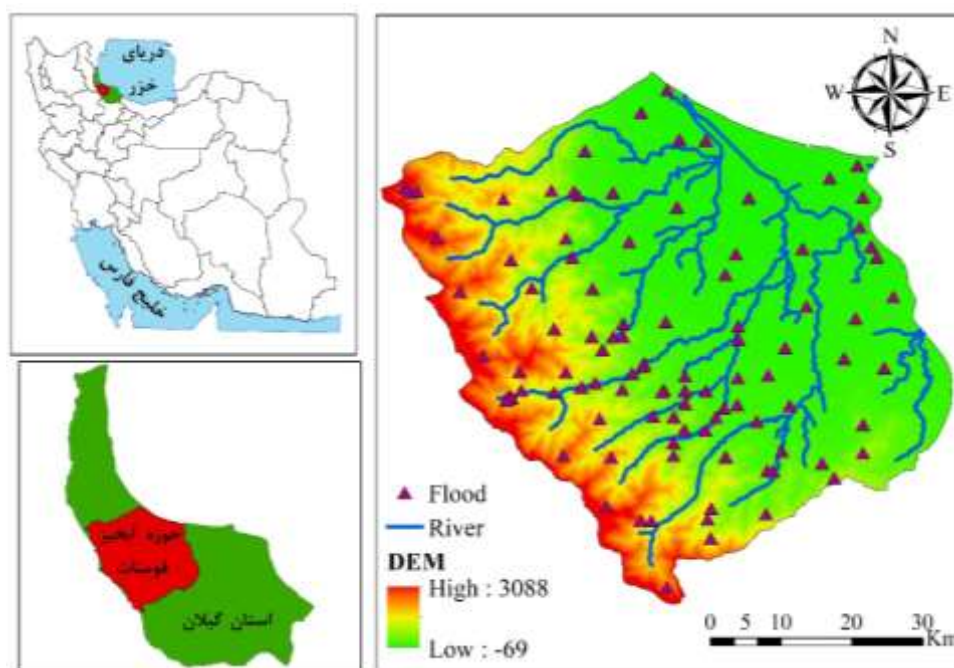
منطقه مطالعاتی

حوزه آبخیز فومنات در استان گیلان واقع شده و بخشی از حوزه آبخیز درجه یک فلات مرکزی محسوب می‌شود. این منطقه در محدوده شرقی واقع شده است. این منطقه دارای مساحتی معادل ۳۵۹۵ کیلومتر مربع است که مرتفع‌ترین نقطه آن، ۳۰۸۸ متر و کم‌ارتفاع‌ترین آن، ۶۹- متر می‌باشد. میانگین بارش سالانه در این حوضه، ۷۸۰ میلی‌متر است. شکل ۲ موقعیت جغرافیایی حوزه آبخیز مطالعاتی را در استان گیلان و کشور نشان می‌دهد.

روش پژوهش

برای انجام پژوهش حاضر، ابتدا با استفاده از مرور منابع و سابقه پژوهش و نیز شناخت منطقه، نقشه و لایه‌های اطلاعاتی مرتبط با عوامل مؤثر بر پهنه‌بندی حساسیت به سیل تهیه شدند (Khosravi *et al.*, 2016; Kourgialas & Karatzas, 2011; Rahmati *et al.*,

1- Topographic Wetness Index (TWI)



شکل ۲- موقعیت جغرافیایی حوزه آبخیز فومنات در استان گیلان و کشور ایران
Figure 2- Geographical location of Fomanat watershed in Gilan province and Iran

باران‌سنجی موجود در داخل و خارج از حوزه آبخیز استفاده شد. در ادامه، هم‌خطی بین فاکتورهای مورد مطالعه با استفاده از فاکتور تورم واریانس^۱ (VIF)، بررسی شد. هم‌خطی در واقع وضعیتی است که در آن یک فاکتور توصیفی با یک یا چند فاکتور دیگر رابطه خطی داشته به طوری که می‌توان آن را به صورت ترکیب خطی از دیگر فاکتورهای مورد بررسی، در نظر گرفت. زمانی که هم‌خطی بین فاکتورها در یک مدل وجود داشته باشد، ضرایب مدل معتبر نخواهد بود.

به‌علاوه، رخدادهای تاریخی سیل به‌عنوان مبنایی برای پیش‌بینی وقوع سیل مورد استفاده قرار می‌گیرد. به طوری که مناطق نزدیک به وقایع سیلاب گذشته، پتانسیل بیشتری در برابر سیل‌خیزی دارند. لذا در این مطالعه، نقاط سیل از سازمان آب منطقه‌ای استان گیلان دریافت شد. در همین راستا، نقاط غیرسیل با استفاده از نقشه توپوگرافی، بازدید میدانی و نرم‌افزار Google Earth انتخاب شد. در ادامه، این نقاط به عنوان متغیر وابسته به‌طور تصادفی به ۷۰٪ (برای آموزش مدل) و ۳۰٪ (برای اعتبارسنجی) تقسیم شد. ابتدا، مدل‌سازی با استفاده از ۷۰٪ داده انجام و سپس برای اعتبارسنجی از ۳۰٪ باقیمانده استفاده می‌شود.

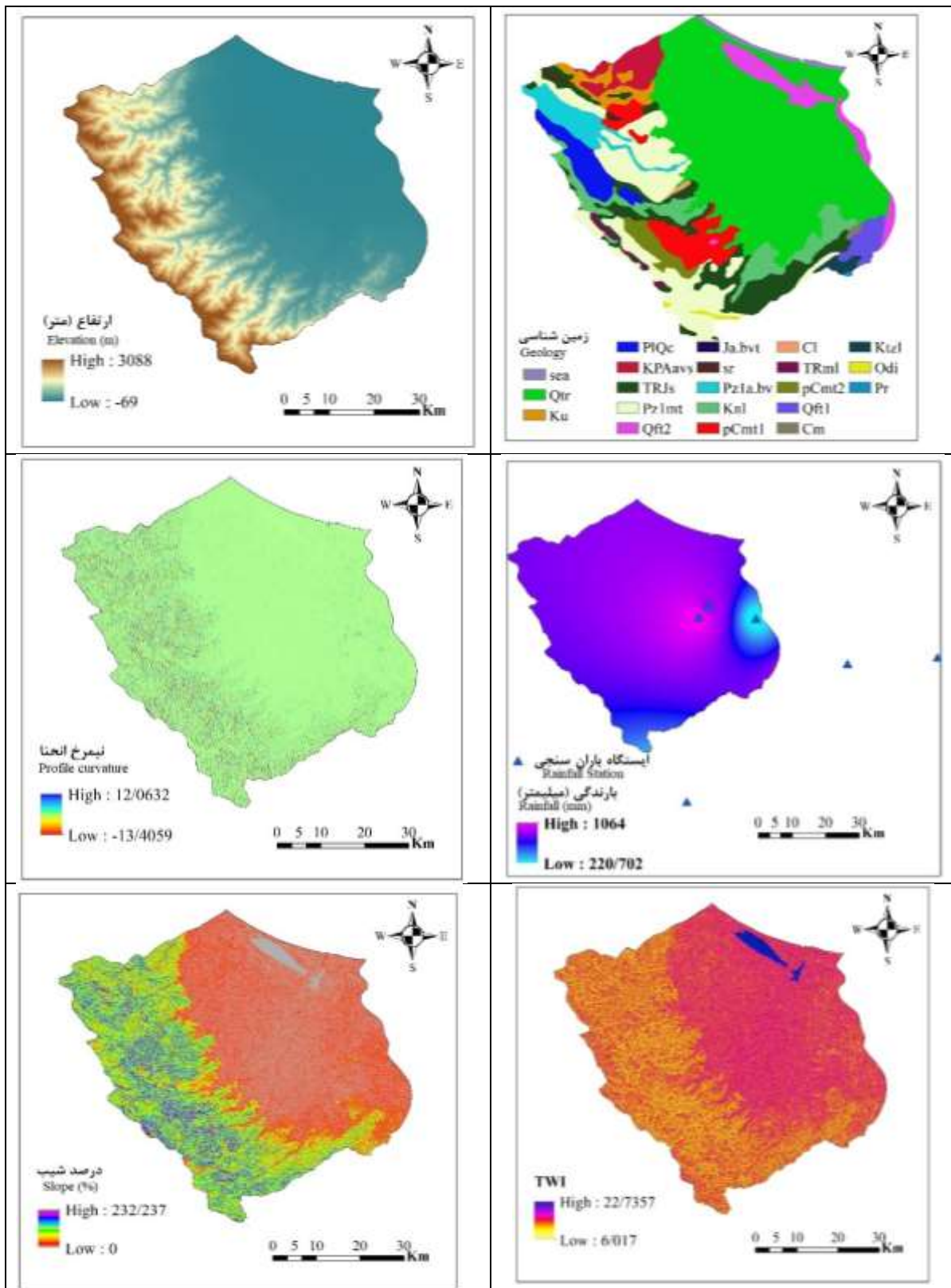
بعد از تعیین فاکتورها و نقشه‌های مذکور (شکل ۳) و تهیه بانک اطلاعاتی بر اساس مرور منابع، اقدام به تهیه نقشه پهنه‌بندی حساسیت به سیلاب با استفاده از مدل خطی تعمیم‌یافته، مدل رگرسیون تطبیقی چندمتغیره و مدل درخت‌های طبقه‌بندی و رگرسیون شد.

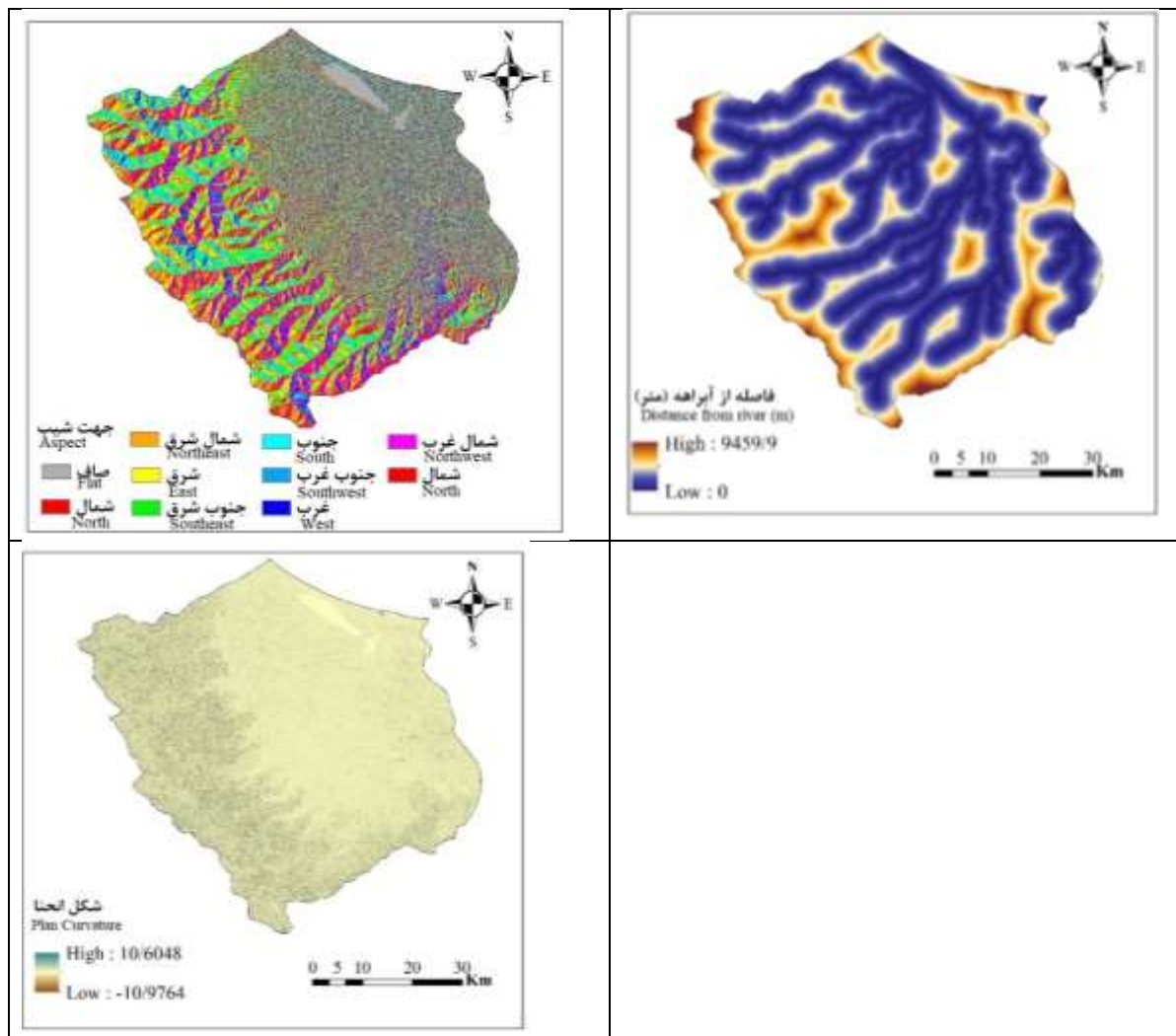
نقشه فاصله از رودخانه با استفاده از لایه رقومی شبکه جریان رودخانه حوزه آبخیز فومنات در محیط نرم‌افزار Arcgis تهیه شد. کاربری اراضی نیز نتیجه اثر متقابل فاکتورهای اجتماعی و اقتصادی و نیز توان بالقوه زمین است. تغییر کاربری اراضی می‌تواند تأثیر چشمگیری در سیل‌گیری منطقه داشته باشد (Tehrany et al., 2015). زمین‌شناسی به‌دلیل تأثیر بر میزان نفوذپذیری و در نتیجه رواناب سطحی، به‌عنوان یکی از فاکتورهای تأثیرگذار در نظر گرفته شد. شکل انحنای زمین که نشان‌دهنده همگرایی جریان و نیز انحراف آن است به‌عنوان فاکتور تأثیرگذار بر خطر سیل استفاده شد (Avand et al., 2021). نیمرخ انحنای زمین که تغییر شیب را در راستای خط جریان نشان می‌دهد (Avand et al., 2021)، به‌عنوان دیگر فاکتور مؤثر در نظر گرفته شد. این نقشه‌ها، با استفاده از مدل رقومی ارتفاع و در محیط نرم‌افزار Arcgis برای حوزه آبخیز مطالعاتی تهیه شد. شاخص رطوبت توپوگرافی نیز یکی دیگر از فاکتورهای دیگر مؤثر بر سیل بوده که در محیط SAGA GIS و با استفاده از معادله ۱ محاسبه می‌شود:

$$TWI = \frac{A_s}{\tan \beta} \quad (1)$$

که در آن A_s سطح ویژه حوزه آبخیز و β شیب زمین (به درصد) می‌باشد.

به‌علاوه، بارش فاکتور مؤثر دیگری بر سیلاب می‌باشد. در این پژوهش، به‌منظور تهیه لایه بارش از داده‌های ۲۰ ساله ایستگاه‌های





شکل ۳- نقشه متغیرهای مستقل و تاثیرگذار بر پهنه‌بندی خطر سیلاب در حوزه آبخیز فومنات، استان گیلان
 Figure 3- Map of independent and influencing variables on flood risk zoning in the Fomanat watershed, Gilan province

استفاده از آن توسط پژوهشگران به‌طور فزاینده‌ای در حال انجام است (Atkinson *et al.*, 1998). به‌علاوه، محققان از این مدل، برای داده‌ها با ساختارهای واریانس خطی و داده‌های غیرخطی و به‌خصوص داده‌های مکانی استفاده می‌کنند (Marmion *et al.*, 2009; Mirchooli *et al.*, 2019). در مدل GLM، ارتباط بین متغیرهای پیش‌بینی کننده و تابع توزیع با استفاده از معادله ۲ محاسبه می‌شود:

$$l = a + b_1x_1 + b_2x_2 + \dots + b_kx_k \quad (2)$$

که در آن، l پیش‌بینی کننده خطی است که با u در معادله زیر مرتبط است:

$$f(u) = l \quad (3)$$

الگوریتم‌های یادگیری ماشین

برای پهنه‌بندی خطر سیل در منطقه مطالعاتی، مدل‌سازی لازم با استفاده از الگوریتم‌های یادگیری ماشین انجام شد. در این مطالعه، الگوریتم‌های زیر مورد استفاده قرار گرفت:

مدل خطی تعمیم‌یافته (GLM)

مدل خطی تعمیم‌یافته، روشی برای برقراری ارتباط بین متغیرهای وابسته و مستقل به حساب می‌آید. این مدل یکی از الگوریتم‌های یادگیری ماشین است که به دلیل طبیعت رگرسیونی آن، برای متغیرهای عددی مناسب می‌باشد (Kalantar *et al.*, 2020). با توجه به امکان به‌کارگیری این مدل، برای توزیع‌های مختلف و نیز برای انواع متفاوت داده‌ها، از جمله انواع داده‌های پیوسته و طبقه‌ای یا ترکیبی از آن‌ها،

باشد. در این روش، ارزیابی مدل با استفاده از داده‌های وقوع و عدم وقوع انجام می‌شود. در واقع، این منحنی حساسیت مدل نسبت به سلول‌هایی که به درستی توسط مدل پیش‌بینی شده‌اند، در مقابل درصد سلول‌های پیش‌بینی شده نسبت به کل را نشان می‌دهد. این مقادیر، نشان‌دهنده توانایی مدل برای تمایز صحیح بین مشاهدات مثبت و منفی اعتبارسنجی می‌باشد. به طوری که حساسیت زیاد بیان‌کننده تعداد زیاد مثبت‌های واقعی یا پیش‌بینی‌های درست بوده و ویژگی بالا (اختلاف 1- Specificity) بیانگر عدد کم مثبت کاذب می‌باشد. در نمودار مربوطه، مقدار مثبت کاذب (1- Specificity) در محور افقی و مقدار مثبت حقیقی (Sensitivity) در محور عمودی نشان داده می‌شود.

نتایج و بحث

نتایج حاصل از تجزیه و تحلیل VIF بین فاکتورهای مؤثر در نظر گرفته شده، در جدول ۱ نشان داده شده است. مقدار آستانه VIF معمولاً ۱۰ در نظر گرفته می‌شود. به عبارت دیگر، اگر مقدار VIF محاسبه شده بیشتر از این حد آستانه باشد، نشان‌دهنده وجود هم‌خطی بین فاکتورهای مورد بررسی است. با توجه به جدول ۱، VIF تمامی فاکتورها کمتر از ده است. لذا همه فاکتورهای برای ادامه تجزیه و تحلیل، وارد فرآیند مدل‌سازی خواهند شد.

در ادامه مدل‌سازی، حساسیت حوزه آبخیز به سیل با استفاده از مدل‌های CART، GLM و RF در نرم‌افزار R و پکیج SDM انجام شد. در این مطالعه، ورودی مدل شامل ده متغیر مستقل و نیز نقاط سیل و عدم وجود سیل بوده که به مدل‌های مورد استفاده معرفی شدند. در ادامه، خروجی این مدل‌ها، نقشه پهنه‌بندی حساسیت به سیلاب می‌باشد.

به علاوه، اهمیت فاکتورهای مؤثر بر حساسیت حوزه آبخیز به سیل با استفاده از کد getVarIm در پکیج مذکور به دست آمد. نتایج حاصل از ارزیابی این فاکتورها در جدول ۲ نشان داده شده است. براساس نتایج حاصل، متغیرهای فاصله از رودخانه، ارتفاع و نیمرخ انحنای بیش‌ترین تأثیر را بر حساسیت منطقه به سیل داشته و در مقابل، فاکتورهای شیب، زمین‌شناسی و شاخص رطوبت توپوگرافی کم‌ترین تأثیر را داشته‌اند. این نتایج با مطالعه آوند و همکاران (Avand et al., 2021) در حوزه آبخیز تجن و نیز یوسفی و همکاران (Yousefi et al., 2022) در حوزه آبخیز کشکان مینی بر اهمیت زیاد فاصله از رودخانه، هم‌راستا می‌باشد. به علاوه، این نتایج با نتایج مطالعه پورقاسمی و همکاران، (Pourghasemi et al., 2015) و نیز خسروی و همکاران، (Khosravi et al., 2016) مطابقت دارد.

مدل درخت‌های طبقه‌بندی و رگرسیون (CART)

همانطور از نام این روش مشخص است، CART ساخت درختان رگرسیونی دودویی را نیز تأمین می‌نماید. درختان رگرسیونی از درختان طبقه‌بندی ساده‌ترند زیرا معیار رشد و هرس نمودن آن‌ها مشابه است. ساختار درخت رگرسیونی مشابه درخت طبقه‌بندی است فقط هر برگ یک مقدار حقیقی را پیش‌بینی می‌کند. الگوریتم CART یک پروسه افزار بازگشتی دودویی است که گره‌های والدین را دقیقاً به دو گره فرزند منسب می‌کند و به طور بازگشتی منسب کردن را تا زمانی که به یک معیار همگن برسد یا معیار توقف دیگری وجود داشته باشد (مثل تمام روش‌های رشد درخت) ادامه می‌دهد. توسعه‌ها با پرسیدن یک سوال با جواب بلی یا خیر تعیین می‌شوند. CART داده‌ها را به دو زیر مجموعه افزار می‌کند بطوریکه رکوردهای درون هر زیر مجموعه نسبت به زیر مجموعه قبلی همگن‌ترند. یک فیلد پیش‌بینی کننده چندین مرتبه در چندین سطح در درخت تکرار می‌شود. این الگوریتم از جداکننده جانشین بهره می‌گیرد تا بهترین استفاده از داده با مقادیر گمشده را داشته باشد (Rutkowski et al., 2014).

مدل رگرسیون تطبیقی چندمتغیره (MARS)

این مدل، یکی از بهترین مدل‌های مبتنی بر رگرسیون می‌باشد که می‌تواند پیش‌بینی را براساس روابط خطی و غیرخطی بین متغیرهای وابسته و مستقل انجام دهد. این مدل دارای انعطاف مناسبی برای پیش‌بینی و پهنه‌بندی مخاطرات و رویدادها براساس متغیر وابسته به متغیرهای مستقل ورودی است (Adnan et al., 2019). یکی از مهم‌ترین ویژگی‌های این مدل آن است که روابط بین متغیرهای مستقل و وابسته را با استفاده از مجموعه‌ای از ضرایب تعریف می‌کند به طوری که اثر هر یک از آن‌ها را به طور جداگانه محاسبه می‌کند. بر این اساس، این مدل رگرسیونی به عنوان روشی مطرح می‌باشد که توابع اصلی را با فواصل مختلف از متغیرهای مستقل تعریف می‌کند (Lazarus & Constantine, 2013).

ارزیابی مدل‌های تهیه نقشه آسیب‌پذیری

به منظور ارزیابی مدل‌های تهیه شده، از منحنی تشخیص عملکرد نسبی (ROC) استفاده شد. منحنی یک نمایش گرافیکی از موازنه بین نرخ خطای منفی و مثبت برای هر مقدار احتمالی از برش‌ها است. سطح زیر منحنی (AUC)، بیانگر مقدار پیش‌بینی سیستم از طریق توصیف توانایی آن در تخمین درست وقایع رخ داده و عدم وقوع رخداد آن است. ایده‌آل‌ترین مدل، بیش‌ترین سطح زیر منحنی را دارد و از ۰/۵ تا ۱ متغیر می‌باشد (Mirchooli et al., 2019). به عبارت دیگر، هرچه AUC بیشتر باشد، نقشه پهنه‌بندی تهیه شده دارای دقت بیشتری می‌

جدول ۱- نتایج آنالیز هم خطی چندگانه بین متغیرهای مستقل

Table 1- Results of multiple collinearity analysis between independent variables

فاکتورها Factors	VIF	فاکتورها Factors	VIF
شاخص رطوبت توپوگرافی Topographic Wetness Index	1.28	ارتفاع Elevation	2.43
زمین شناسی Geology	1.33	جهت شیب Aspect	1.01
کاربری اراضی Land use	1.16	شیب Slope	2.23
بارش Rainfall	1.15	شکل انحنای Plan curvature	1.37
فاصله از رودخانه Distance to river	1.19	نیمرخ انحنای Profile curvature	1.27

جدول ۲- اهمیت فاکتورهای مستقل مورد استفاده در حساسیت حوزه آبخیز به سیل

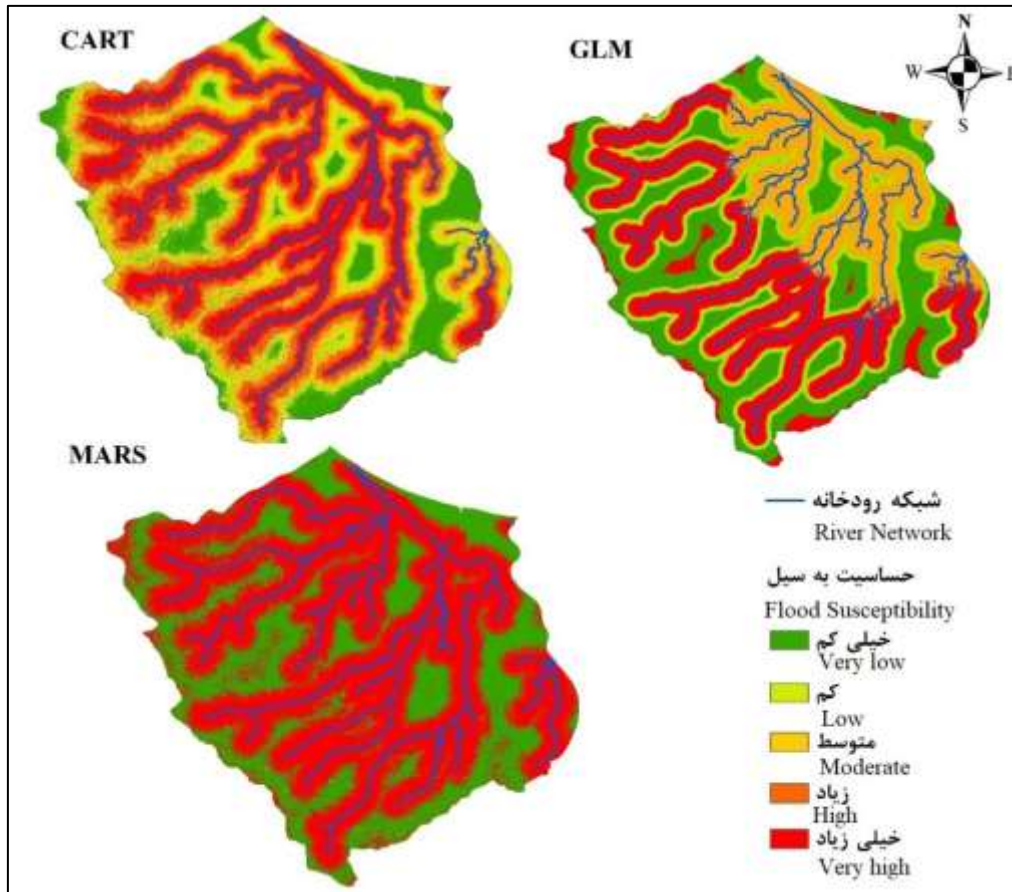
Table 2- The importance of independent factors used in the sensitivity of the watershed to floods

فاکتورها Factors	MARS	GLM	CART
ارتفاع Elevation	1.9	13.0	25.9
جهت شیب Aspect	0.0	1.6	0.0
شیب Slope	2.5	3.9	0.0
شکل انحنای Plan curvature	0.4	5.6	0.0
نیمرخ انحنای Profile curvature	0.0	19.6	0.0
شاخص رطوبت توپوگرافی Topographic Wetness Index (TWI)	0.0	0.0	0.0
زمین شناسی Geology	2.2	4.9	0.0
کاربری اراضی Land use	0.6	6.3	0.0
بارش Rainfall	0.0	5.6	0.0
فاصله از رودخانه Distance to river	30.8	56.6	63.1

جدول ۳- مساحت (درصد) مناطق موجود در کلاس‌های مختلف خطر سیلاب در حوزه آبخیز فومنات، ایران

Table 3- The area (percentage) of the areas in different flood risk classes in the Fomanat watershed, Iran

	خیلی زیاد Very high	زیاد High	متوسط Moderate	کم Low	خیلی کم Very low
CART	34.0	0.1	21.2	17.6	26.6
GLM	32.4	25.1	16.2	12.7	13.6
MARS	67.2	0.5	0.1	0.1	32.1



شکل ۴- نقشه پهنه‌بندی حساسیت به سیل با استفاده از مدل‌های GLM، CART، MARS
 Figure 4- Zoning map of flood susceptibility using GLM, CART, and MARS models

۰/۷۶، مدل CART معادل ۰/۹ و مدل GLM معادل ۰/۸۴ می‌باشد. همچنین عملکرد بهتر مدل CART در مقایسه با مدل‌های دیگر توسط شاخص‌های دیگر تأیید شد. به طوری که براساس TSS، مدل MARS معادل ۰/۵۲، مدل CART برابر ۰/۷۷ و مدل GLM معادل ۰/۶۶ به دست آمد. در این راستا، لازم به ذکر است که در بسیاری از مطالعات پیشین آزادی و همکاران (Azadi et al., 2020) و نیز یوسفی و همکاران (Yousefi et al., 2021) تنها از AUC برای ارزیابی دقت مدل‌ها، استفاده کردند؛ اما در مطالعه حاضر از هر دو شاخص AUC و TSS استفاده شده است.

نتیجه‌گیری

سیل یکی از وقایع مخاطره آمیز در نقاط مختلف ایران به خصوص مناطق شمالی کشور به حساب می‌آید. بنابراین، لازم است به منظور کاهش خسارات جانی و مالی ناشی از وقوع آن مطالعات لازم جهت پهنه‌بندی و تعیین اراضی تحت تأثیر با بیش‌ترین میزان خطر صورت گیرد. در این راستا، یکی از ابزارهای مناسب برای این‌گونه مطالعات،

نتایج این مطالعه مبنی بر اهمیت عامل ارتفاع بر حساسیت به وقوع سیل در حوزه مطالعاتی نیز با سایر پژوهش‌های انجام شده از جمله چاپی و همکاران، (Chapi et al., 2017)، یوسفی و همکاران (Yousefi et al., 2021) همخوانی دارد. تأثیر عامل ارتفاع بر وقوع سیل به این صورت است که با افزایش ارتفاع، حساسیت حوزه آبخیز به وقوع سیل کاهش می‌یابد. این موضوع به این دلیل است که در طبقات ارتفاعی پایین‌تر، به دلیل تجمع آب باران، احتمال وقوع سیل افزایش می‌یابد.

در نهایت، نقشه حساسیت به سیل تهیه شده، در نرم‌افزار R به پنج کلاس حساسیت خیلی کم، کم، متوسط، زیاد و خیلی زیاد طبقه‌بندی شدند. براساس نتایج به دست آمده (جدول ۳ و شکل ۴)، مساحت‌های تحت پوشش کلاس‌های خیلی کم، کم، متوسط، زیاد و خیلی زیاد در مدل CART به ترتیب ۲۶/۶، ۱۷/۶، ۲۱/۲، ۰/۱ و ۳۴/۰ درصد می‌باشد. این نتایج برای مدل GLM به ترتیب ۱۳/۶، ۱۲/۷، ۱۶/۲، ۲۵/۱ و ۳۲/۴ درصد می‌باشد.

براساس نتایج به دست آمده، مدل CART عملکرد بهتری نسبت به سایر مدل‌ها دارد. به طوری که AUC در مورد مدل MARS معادل

همچنین نتایج حاصل از مدل‌سازی نشان داد که مدل CART کارایی بالاتری نسبت به سایر مدل‌های مورد استفاده، داشته است. با به‌کارگیری نتایج حاصل از این مطالعه می‌توان اقدامات مدیریتی مناسبی جهت کاهش خسارات و تلفات اتخاذ کرد. به‌علاوه، در کشورهای در حال توسعه که با محدودیت دسترسی به اطلاعات هیدروژئولوژیکی و اداکیکی مواجه هستند، استفاده از سیستم اطلاعات جغرافیایی و نیز تکنیک‌های داده‌کاوی نقش مهمی در انجام مطالعات ایفا می‌کنند.

الگوریتم‌های یادگیری ماشین همراه با شناسایی فاکتورهای مؤثر بر وقوع سیل از طریق بررسی مطالعات پیشین و نیز بازبدهی‌های میدانی است. لذا، پژوهش حاضر با هدف ارزیابی دقت الگوریتم‌های یادگیری ماشین شامل مدل خطی تعمیم‌یافته، مدل رگرسیون تطبیقی چندمتغیره و مدل درخت‌های طبقه‌بندی و رگرسیون در تعیین مناطق حساس به سیل، انجام شد. بدین منظور، ۱۰ فاکتور مؤثر محیط‌زیستی و نیز ۱۰۰ نقطه سیل استفاده شد. نتایج حاصل از این بررسی نشان داد که متغیر فاصله از آبراهه نقش تعیین‌کننده‌ای در وقوع سیل‌ها داشته و بعد از متغیرهایی مانند ارتفاع و نیمرخ طولی نقش زیادی ایفا می‌کنند.

References

- Adnan, R.M., Liang, Z., & Heddani, S. (2019). Least square support vector machine and multivariate adaptive regression splines for streamflow prediction in mountainous basin using hydro-meteorological data as inputs. *Journal of Hydrology*, (586), 124371. <https://doi.org/10.1016/J.JHYDROL.2019.124371>
- Atkinson, P., Jiskoot, H., Massari, R., & Murray, T. (1998). Generalized linear modelling in geomorphology. *Earth Surface Processes and Landforms. The Journal of the British Geomorphological Group*, 23(13), 1185-1195.
- Avand, M., & Moradi, H. (2021). Using machine learning models, remote sensing, and GIS to investigate the effects of changing climates and land uses on flood probability. *Journal of Hydrology*, 595, 125663. <https://doi.org/10.1016/j.jhydrol.2020.125663>
- Avand, M. T., Moradi, M., Ramazanzadeh Lasbooyee, Mehdi. (2021). Spatial prediction of future flood risk: an approach to the effects of climate change. *Geosciences*, 11(1), 1-20. <https://doi.org/10.3390/geosciences11010025>
- Avand, M.T., Moradi, H.R., & Ramzanzadeh Lesboi, M. (2020). Preparation of flood sensitivity map using Bayesian random forest and linear generalized machine learning models. *Environment and Water Engineering*, 6(1), 83-95.
- Avand, M., Kuriqi, A., Khazaei, M., & Ghorbanzadeh, O. (2022). DEM resolution effects on machine learning performance for flood probability mapping. *Journal of Hydro-Environment Research*, 40, 1-16. <https://doi.org/10.1016/j.jher.2021.10.002>
- Azadi, F., Sadouq, S.H., Ghahrodi, M., & Shahabi, H. (2020). Flood risk sensitivity zoning in Kashkan watershed using WOE and EBF models. *Scientific Journal of Geography and Environmental Hazards*, 9(1), 45-60.
- Chapi, K., Singh, V.P., Shirzadi, A., Shahabi, H., Bui, D.T., Pham, B.T., & Khosravi, K. (2017). A novel hybrid artificial intelligence approach for flood susceptibility assessment. *Environmental Modelling & Software*, 95, 229-245. <https://doi.org/10.1016/j.envsoft.2017.06.012>
- Hasanuzzaman, M., Islam, A., Bera, B., & Shit, P.K. (2022). A comparison of performance measures of three machine learning algorithms for flood susceptibility mapping of river Silabati (tropical river, India). *Physics and Chemistry of the Earth, Parts A/B/C*, 127, 103198. <https://doi.org/10.1016/j.pce.2022.103198>
- Kalantar, B., Ueda, N., Saeidi, V., Ahmadi, K., Halin, A. A., & Shabani, F. (2020). Landslide susceptibility mapping: Machine and ensemble learning based on remote sensing big data. *Remote Sensing*, 12(11), 1737. <https://doi.org/10.3390/rs12111737>
- 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, 947-987. <https://doi.org/10.1007/s11069-016-2357-2>
- Kourgialas, N.N., Karatzas, G.P. (2011). Flood management and a GIS modelling method to assess flood-hazard areas—a case study. *Hydrological Sciences Journal*, 56, 212-225. <https://doi.org/10.1080/02626667.2011.555836>
- Lazarus, E.D., Constantine, J.A. (2013). Generic theory for channel sinuosity. *Proceedings of the National Academy of Sciences USA*, 110, 8447-8452. <https://doi.org/10.1073/pnas.1214074110>
- Marmion, M., Luoto, M., Heikkinen, R.K., & Thuiller, W. (2009). The performance of state-of-the-art modelling techniques depends on geographical distribution of species. *Ecological Modelling*, 220(24), 3512-3520. <https://doi.org/10.1016/j.ecolmodel.2008.10.019>
- Mosavi, A., Golshan, M., Janzadeh, S., Choubin, B., Melesse, A.M., & Dineva, A.A. (2020). Ensemble models of GLM, FDA, MARS, and RF for flood and erosion susceptibility mapping: a priority assessment of sub-basins. *Geocarto International*, 37(9): 1-20. <https://doi.org/10.1080/10106049.2020.1829101>
- Mirchooli, F., Motevalli, A., Pourghasemi, H.R., Mohammadi, M., Bhattacharya, P., Maghsood, F.F., & Tiefenbacher, J.P. (2019). How do data-mining models consider arsenic contamination in sediments and variables importance? *Environmental Monitoring and Assessment*, 191. <https://doi.org/10.1007/s10661-019-7979-x>

17. Mohammadi, M., Darabi, H., Mirchooli, F., Bakhshae, A., & Torabi Haghghi, A. (2021). Flood risk mapping and crop-water loss modeling using water footprint analysis in agricultural watershed, northern Iran. *Natural Hazards*, 105, 2007-2025. <https://doi.org/10.1007/s11069-020-04387-w>
18. Nachappa, T.G., Ghorbanzadeh, O., & Gholamnia, K. (2020). Multi-Hazard Exposure Mapping Using Machine Learning for the State of Salzburg, Austria. *Remote Sensing*, 12(17), 2757. <https://doi.org/10.3390/rs12172757>
19. Mosavi, A., Golshan, M., Janizadeh, S., Choubin, B., Melesse, A.M., & Dineva, A.A. (2022). Ensemble models of GLM, FDA, MARS, and RF for flood and erosion susceptibility mapping: a priority assessment of sub-basins. *Geocarto International*, 37(9), 2541-2560. <https://doi.org/10.1080/10106049.2020.1829101>
20. Pourghasemi, H.R., & Beheshtirad, M. (2015). Assessment of a data-driven evidential belief function model and GIS for groundwater potential mapping in the Koohrang Watershed, Iran. *Geocarto International*, 30, 662-685. <https://doi.org/10.1080/10106049.2014.966161>
21. Pourghasemi, H.R., Kariminejad, N., Amiri, M., Edalat, M., Zarafshar, M., Blaschke, T., & Cerda, A. (2020). Assessing and mapping multi-hazard risk susceptibility using a machine learning technique. *Scientific Reports*, 10, 1-11. <https://doi.org/10.1038/s41598-020-60191-3>
22. 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, 42-70. <https://doi.org/10.1080/10106049.2015.1041559>
23. Rutkowski, L., Jaworski, M., Pietruczuk, L., & Duda, P., (2014). The CART decision tree for mining data streams. *Information Sciences*, 266, 1-15. <https://doi.org/10.1016/j.ins.2013.12.060>
24. Seydi, S. T., Kanani-Sadat, Y., Hasanlou, M., Sahraei, R., Chanussot, J., & Amani, M. (2022). Comparison of machine learning algorithms for flood susceptibility mapping. *Remote Sensing*, 15(1), 192
25. Skilodimou, H.D., Bathrellos, G.D., Chousianitis, K., Youssef, A.M., Pradhan, B. (2019). Multi-hazard assessment modeling via multi-criteria analysis and GIS: a case study. *Environmental Earth Sciences*, 78, 42. <https://doi.org/10.1007/s12665-018-8003-4>
26. Tehrany, M.S., Pradhan, B., & Jebur, M.N. (2015). Flood susceptibility analysis and its verification using a novel ensemble support vector machine and frequency ratio method. *Stochastic Environmental Research and Risk Assessment*, 29, 1149-1165. <https://doi.org/10.1007/s00477-015-1021-9>
27. Tehrany, M.S., Pradhan, B., & Jebur, M.N. (2014). Flood susceptibility mapping using a novel ensemble weights-of-evidence and support vector machine models in GIS. *Journal of Hydrology*, 512, 332-343. <https://doi.org/10.1016/j.jhydrol.2014.03.008>
28. Tehrany Shafapour, M., Pradhan, B., Mansor, S.H., & Noordin, A. (2015). Flood susceptibility assessment using GIS-based support vector machine model with different kernel types. *Catena*, 125, 91-101.
29. Yousefi, H., Yonsei, H. A., Davoudi-Moghadam, D., Arshiya, A., & Shamsi, Z. (2022). Determination of flood potential using CART, GLM and GAM machine learning models (case study: Kashkan Basin). *Scientific Research Journal of Irrigation and Water Engineering of Iran*, 12(48), 84-105.
30. Yousefi, H., Yonsei, H.A., Arshiya, A., Yarahamdi, Y., & Guderzi, A. (2021). Determining flood-prone areas with models to reduce flood risks (case study: Kashkan watershed). *Ecohydrology*, 8(1), 307-319.