

کلیدواژه‌ها: جنگل تصادفی، حوضه مارون، ماشین بردار پشتیبان، مدیریت سیلاب، وقوع سیلاب.

مقدمه

در سطح جهانی، سیل از جمله بلایای طبیعی است که باعث تخریب چشم‌گیر محیط‌زیست می‌شود، و آثاری از جمله خسارات اقتصادی، جراحات و مرگ و میر، تخریب زیرساخت‌ها و انتشار عوامل بیماری‌زا در مناطق تحت تأثیر خود دارد. دینگ و همکاران [۶]. سالانه حدود ۲۰۰ میلیون نفر تحت تأثیر سیل قرار می‌گیرند که باعث خسارت اقتصادی ۹۵ میلیارد دلاری در سراسر جهان می‌شود. مابوکو و همکاران [۱۴]. در ایران پدیده سیلاب روندی افزایشی داشته است، به طوری که در ۵۰ سال اخیر حدود ۳۷۰۰ واقعه سیل در کشور گزارش شده که حدود ۵۳ درصد آن در سال‌های اخیر روی داده است. شریفی و همکاران [۲۱]. جریان سیل به طور عمده ناشی از رواناب سطحی می‌باشد که حاصل خصوصیات بارش و حوزه آبخیز است. تلوری [۲۴]. وقوع سیلاب موجب آثار مخربی از جمله شست و تخریب خاک‌های تشکیل شده و در حال تشکیل، فرسایش سریع زمین و ناهموار کردن آن، کاهش حاصلخیزی زمین بر اثر انباشته شدن رسوب‌های درشت، تخریب قنوات، تلف شدن احشام و حیات وحش، تخریب راه، پل و منازل، پر شدن سریع دریاچه‌ها، هرز رفتن آب و نبود فرصت کافی برای بهره‌وری از آب می‌شود. معیری و انتظاری [۱۶]. فعالیت‌های انسانی نظیر تخریب پوشش گیاهی در مناطق مختلف حوزه آبخیز، کاربری غیراصولی اراضی، توسعه سطوح غیرقابل نفوذ و امثال آن پتانسیل سیل‌خیزی را افزایش داده است. شعبانی بازنشین و همکاران [۲۰]. تحلیل حساسیت و خطرپذیری سیل به عنوان یکی از رویکردهای مهم و کارآمد مطرح در زمینه ارزیابی خطر سیل است که امروزه به عنوان یک نقطه عطف در مدیریت بحران، مدیریت منابع آب و تحلیل مخاطرات طبیعی مطرح شده است. ما و همکاران [۱۳]. به علاوه، شناسایی مناطق سیل‌خیز یک گام اساسی برای مدیریت ریسک سیل است و منطقه‌بندی خطر سیلاب، اندازه‌گیری کمی از خطر سیل را فراهم می‌کند، در نتیجه امکان ارسال هشدارهای مناسب در مواقع خطر سیل و تسهیل عملیات امداد و نجات فراهم می‌شود. ولیزاده کامران و همکاران [۲۶]. نقشه پهنه‌بندی میزان خطر سیلاب ابزاری کارآمد در برنامه‌ریزی جهت کاهش خطر سیل در سکونتگاه‌های انسانی و توسعه زیرساخت‌های شهری است. بوشله و همکاران [۴].

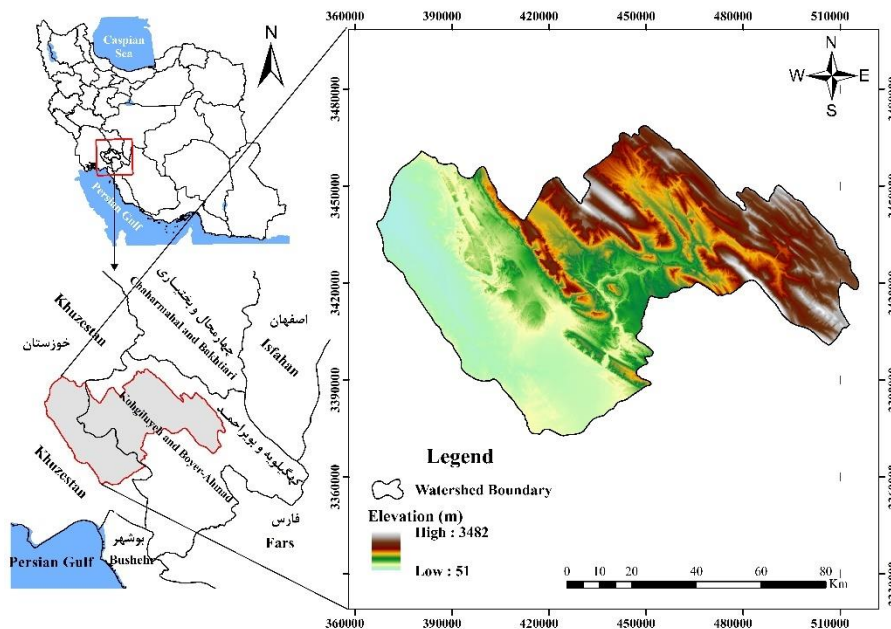
ارزیابی الگوریتم‌های ماشین یادگیرنده (SVM و RF) در تولید نقشه حساسیت سیلاب حوزه آبخیز مارون

محمد امین کیانی اصل^۱، بهزاد متشغ^۲ و سید حسین روشن^۳
تاریخ دریافت: ۱۴۰۱/۰۳/۲۶ تاریخ پذیرش: ۱۴۰۱/۰۵/۲۸

چکیده

عوامل متعدد اقلیمی، هیدرولوژیکی، ژئومورفولوژیکی و زمین‌شناسی در وقوع سیلاب نقش دارند. تجزیه و تحلیل سیلاب، مدیریت و کنترل آن می‌تواند با تهیه نقشه‌های پتانسیل سیل‌خیزی انجام شود. هدف این پژوهش تهیه نقشه پتانسیل سیل‌خیزی حوزه مارون با استفاده از روش‌های یادگیری ماشین جنگل تصادفی و ماشین بردار پشتیبان می‌باشد. به این منظور ۱۶ پارامتر مؤثر در وقوع سیلاب شامل طبقات ارتفاعی، مقدار و جهت شیب، انحنای زمین، سازندهای زمین‌شناسی، کاربری اراضی، شماره منحنی، بارندگی، درجه حرارت، شاخص قدرت جریان (SPI)، شاخص رطوبت توپوگرافی (TWI)، فاصله از آبراهه، تراکم آبراهه، فاصله از جاده، تراکم جاده و شاخص پوشش گیاهی NDVI در نظر گرفته شد. پارامترهای مذکور در محیط نرم‌افزارهای ArcGIS 10.8، ENVI 5.3 و GIS SAGA 7.2 تهیه شدند. سپس به فرمت خوانا برای محیط نرم‌افزار R به منظور اجرای مدل ماشین بردار پشتیبان و جنگل تصادفی تبدیل شدند. در نهایت با استفاده از بسته مدل‌های SVM و RF اجرا شدند و با استفاده از منحنی تشخیص عملکرد نسبی (ROC) مورد ارزیابی قرار گرفتند. نتایج نشان داد که مدل‌های SVM و RF بترتیب با دقت ۰/۹۹۷ و ۰/۹۴۷ درصد نقشه سیل‌خیزی حوضه مارون را پیش‌بینی کردند.

- ۱- دانشجوی کارشناسی ارشد مهندسی آبخیزداری، گروه مرتع و آبخیزداری، دانشکده منابع طبیعی، دانشگاه صنعتی خاتم الانبیا بهبهان، ایران.
- ۲- نویسنده مسئول و استادیار، گروه مرتع و آبخیزداری، دانشکده منابع طبیعی، دانشگاه صنعتی خاتم الانبیا بهبهان، ایران، رایانامه: bmoteshaffeh@gmail.com
- ۳- دانش‌آموخته دکتری، علوم و مهندسی آبخیزداری، گروه مهندسی آبخیزداری، دانشکده منابع طبیعی، دانشگاه علوم کشاورزی و منابع طبیعی ساری، ایران.



شکل ۱- موقعیت جغرافیایی حوزه آبخیز مارون در ایران
Fig 1. Geographical location of Maroon watershed in the Iran

مواد و روش‌ها منطقه مورد مطالعه

حوزه آبخیز مارون به عنوان یکی از حوضه‌های پرآب کشور در تامین منابع آب شرب، کشاورزی و صنعتی در جنوب غرب ایران و در استان‌های کهگیلویه و بویراحمد و خوزستان گستره دارد. این حوضه یکی از زیرحوضه‌های، حوضه بزرگ مارون - جراحی است. حوزه آبخیز مارون در موقعیت جغرافیایی ۴۹ درجه و ۳۵ دقیقه تا ۵۱ درجه و ۰۹ دقیقه طول شرقی و ۳۰ درجه و ۲۸ دقیقه تا ۳۱ درجه و ۲۱ دقیقه عرض شمالی قرار دارد (شکل ۱). رودخانه اصلی این حوضه مارون نام دارد که از ارتفاعات زاگرس میانی در استان کهگیلویه و بویراحمد سرچشمه می‌گیرد. حوضه رودخانه مارون بدلیل شیب نسبتاً زیاد و پوشش گیاهی اندک و همچنین سازندهای بسیار سست رسوبات تبخیری و مارن از فرسایش پذیری بالایی برخوردار است. بیشترین ارتفاع این حوضه ۳۴۸۱ متر و کمترین ارتفاع این حوضه ۵۳ متر می‌باشد. مساحت این حوضه نیز ۷۳۱۶/۱۶ کیلومترمربع می‌باشد.

روش تحقیق

با مطالعه تحقیقات گذشته و اسناد علمی مهم‌ترین عوامل مؤثر در وقوع سیلاب در قالب ۱۶ پارامتر شامل موارد زیر در این تحقیق انتخاب شدند.

طبقات ارتفاعی

وقوع سیل در مناطق مرتفع تقریباً غیرممکن بوده در صورتیکه مناطق مسطح پتانسیل بالایی برای وقوع سیل دارند بوتزن و همکاران [۳]. کیا و همکاران [۱۰] بیان کردند که توپوگرافی یک عامل تقویت کننده است که نقش بسزایی در وقوع سیل دارد. برای تهیه لایه طبقات ارتفاعی ابتدا مدل رقومی ارتفاع (DEM) حوضه از پایگاه

ارائه روشی دقیق برای بررسی خسارت سیلاب همواره مورد توجه پژوهشگران در اقصی نقاط جهان بوده است. یوسف و همکاران [۲۷] به بررسی مقایسه‌ای الگوریتم‌های شبکه عصبی کانولوشن (CNN) و ماشین بردار پشتیبان (SVM) در تهیه نقشه حساسیت سیل در رأس قریب، دریای سرخ پرداختند. آنها سیزده عامل مرتبط و مؤثر در وقوع سیل را با ArcGIS تهیه و با منحنی‌های مشخصه عملکرد (ROC) ارزیابی کردند. نتایج آنها نشان داد که شبکه عصبی کانولوشن^۱ (CNN) دقت پیش‌بینی ۸۶/۵ درصد (عملکرد بالا) را ارائه می‌دهد، در حالی که عملکرد الگوریتم ماشین بردار پشتیبان (SVM) ۷۱/۶ درصد (کارایی متوسط) می‌باشد. لیو و همکاران [۱۲] با استفاده از الگوریتم‌های ماشین بردار پشتیبان (SVM) و رگرسیون لجستیک نقشه حساسیت سیل را در مناطق جاده‌ای پیش‌بینی کردند. آنها ۱۱ عامل مؤثر در وقوع سیل را در نظر گرفتند. نتایج آنها بیانگر آنست که الگوریتم‌های مذکور با دقت بالایی (AUC=۰/۹۳) نقشه نهایی را تولید می‌کنند. حوضه مارون یکی از حوضه‌های مهم در تامین منابع آبی جنوب کشور بوده و در سال‌های اخیر شاهد وقوع سیلاب‌های متعدد در این حوضه هستیم، مطالعه پتانسیل سیل خیزی و بررسی مناطق مختلف از نظر خطر سیلاب ضروری به نظر می‌رسد همچنین علیرغم نتایج مطلوب استفاده از روش‌های یادگیری ماشین در مناطق مختلف دنیا و همچنین ایران، اما تاکنون از این روش‌ها در آنالیز حساسیت سیل خیزی حوزه مارون استفاده نشده است. بنابراین هدف این مطالعه بررسی پتانسیل سیل خیزی در حوزه آبخیز مارون با استفاده از روش‌های یادگیری ماشین جنگل تصادفی (RF) و ماشین بردار پشتیبان (SVM) می‌باشد.

1. Convolutional Neural Network

سازمان زمین‌شناسی ایالات متحده (USGS) بر اساس سنجنده ASTER اخذ و سپس طبقات ارتفاعی در نرم‌افزار ArcMap 10.8 بدست آمد.

مقدار و جهت شیب

شیب زمین در حوزه‌های آبخیز مهم‌ترین عامل در تعیین سرعت جریان است. هنگامی که شیب در سطح حوضه زیاد شود، فرصت لازم برای نفوذ کاهش یافته و همچنین با افزایش شیب حوضه، زمان تمرکز کاهش می‌یابد فرناندز و لوتز [۷]. جهت شیب پارامتری وابسته به روند جغرافیای طبیعی و الگوی رطوبت خاک است و بر روی شرایط هیدرولوژیک تأثیرگذار است. جهت شیب شدیداً بر روی فرآیندهای هیدرولوژی با جهت بارش رو به جلو تأثیر دارد سیدله و اوچیایی [۲۲].

انحنای زمین

انحنای زمین یکی از پارامترهای مکانی مورد استفاده در تهیه نقشه‌های پتانسیل سیل حوزه‌های آبخیز می‌باشد. نقشه انحنای زمین بر اساس لایه DEM در نرم‌افزار ArcMap 10.8 تهیه شد.

زمین‌شناسی (واحدهای سنگی)

سنگ‌شناسی یک عامل اساسی در تغییرات مکانی-زمانی پارامترهای هیدرولوژیکی و تولید رسوب در حوزه آبخیز است سربو استاوا و همکاران [۲۳]. واحدهای سنگی قابل نفوذ، شرایط نفوذ آب به داخل خاک را فراهم می‌کند و تخلیه آنرا به داخل آبراهه اصلی حوضه به تأخیر می‌اندازد که منجر به کاهش رواناب سطحی و در نتیجه کاهش وقوع سیلاب می‌شود. از طرفی، مناطقی از حوضه که دارای سنگ بستر یا واحدهای سنگی غیرقابل نفوذ هستند، موجب می‌شوند که حجم بالایی از رواناب سطحی تولید شده و وقوع سیلاب اجتناب‌ناپذیر باشد گراد [۹]. نقشه واحدهای سنگی با کمک نقشه زمین‌شناسی منطقه با مقیاس ۱:۱۰۰۰۰۰ در نرم‌افزار ArcMap 10.8 بدست آمد.

کاربری اراضی

کاربری اراضی به عنوان یک پارامتر دینامیک بر فرآیندهای هیدرولوژیکی مانند نفوذ، رواناب سطحی، تبخیر و تبخیر و تعرق تأثیر می‌گذارد بیون و کرکی [۲]. بین وقوع سیل و تراکم پوشش گیاهی رابطه معکوس وجود دارد. مناطق بدون پوشش گیاهی نسبت به مناطق جنگلی بارش را سریع‌تر به رواناب سطحی تبدیل می‌کنند

لی و همکاران [۱۱]. نقشه کاربری اراضی با استفاده از تصاویر ماهواره‌ای لندست ۸ سنجنده OLI تهیه گردید (جدول ۱).

این تصاویر از پایگاه USGU اخذ و با نرم‌افزار ENVI 5.3 پردازش‌های اولیه شامل تصحیحات رادیومتریکی و اتمسفری روی آنها انجام گردید. سپس از روش طبقه‌بندی نظارت شده (حداکثر احتمال) لایه کاربری اراضی حوضه استخراج گردید. ارزیابی نقشه تولید شده با ضریب کاپا و صحت کلی انجام گردید. شماره منحنی (CN)

هر چه شماره منحنی در یک مکان بیشتر باشد، میزان رواناب تولیدی بیشتری داشته که می‌تواند نقش مهمی در وقوع سیلاب داشته باشد. جهت تولید نقشه CN ابتدا لایه گروه‌های هیدرولوژیکی خاک از نقشه خاک FAO استخراج و با لایه کاربری اراضی در الحاقیه ArcCN-Runoff از نرم‌افزار ArcMap 10.8 ترکیب گردید. در نهایت با استفاده از جداول مربوط به روش SCS-CN مقدار شماره منحنی برای نقشه تلفیق شده، بدست آمد.

بارندگی

به طور کلی بارندگی عامل اصلی وقوع سیلاب است. از طرفی ویژگی‌های بارش از جمله مدت و شدت می‌توانند در وقوع سیلاب نقش مهمی داشته باشند. به منظور تهیه نقشه توزیع مکانی بارش در سطح حوضه از آمار ۲۷ ایستگاه هواشناسی استفاده شد. ابتدا این آمار از اداره هواشناسی استان کهگیلویه و بویراحمد اخذ و صحت داده‌ها از لحاظ همگنی در نرم‌افزار SPSS بررسی گردید. سپس میانگین بارش سالانه برای هر ایستگاه محاسبه و با روش درونیابی وزن‌دهی فاصله معکوس (IDW) نقشه توزیع مکانی بارش در سطح حوضه بدست آمد.

درجه حرارت

درجه حرارت بطور غیرمستقیم بر روی شرایط سیلابی یک حوضه تأثیر می‌گذارد. به طوری که با بالا رفتن درجه حرارت میزان ذوب برف و یخ در حوضه زیاد می‌شود در نتیجه شرایط برای وقوع سیلاب ناگهانی ایجاد می‌شود. به منظور تهیه نقشه توزیع مکانی درجه حرارت در سطح حوضه از آمار ایستگاه‌های سینوپتیک استفاده گردید. این نقشه نیز با روش IDW در نرم‌افزار ArcMap 10.8 تولید گردید.

شاخص قدرت جریان (SPI)

جدول ۱- مشخصات تصاویر ماهواره‌ای مورد استفاده برای تهیه لایه کاربری اراضی

Table 1. Specifications of satellite images used to prepare land use layer

ردیف	نوع ماهواره و سنجنده	تاریخ تصویر	ردیف	گذر
Row	Satellite and Sensor Type	Date Acquired	Row	Path
1		15-04-2021	38	164
2	لندست ۸، سنجنده OLI	30-03-2021	39	164
3	Landsat 8, OLI Sensor	06-04-2021	39	165

شاخص تفاوت پوشش گیاهی نرمال شده (NDVI) پوشش گیاهی نقش مهمی در نفوذ بارش به خاک و جلوگیری از ایجاد رواناب و وقوع سیلاب دارد. پوشش گیاهی اغلب با شاخص گیاهی بنام پوشش گیاهی تفاضلی نرمال شده (NDVI) قابل اندازه‌گیری می‌باشد محمودزاده و باکویی [۱۵]. محدوده این شاخص بین ۱ و -۱ است و مقادیر بالای این شاخص نشان دهنده افزایش تراکم پوشش گیاهی است. این شاخص با استفاده از رابطه زیر قابل محاسبه است.

$$NDVI = \frac{NIR - R}{NIR + R} \quad (3)$$

NIR: بازتاب طول موج مادون قرمز نزدیک، R: بازتاب طول موج قرمز

به منظور تولید نقشه شاخص NDVI از تصاویر ماهواره جدول ۱ بر اساس رابطه ۳، نسبت بانندی بین باند ۵ و ۴ لندست ۸ در نرم‌افزار ENVI 5.3 اقدام گردید.

پس از تهیه لایه رستری پارامترهای فوق، همه لایه‌ها در نرم‌افزار ArcGIS 10.8 به فرمت ASCII تبدیل و به نرم‌افزار R فراخوانی شدند. سپس با استفاده از لایه آموزشی که حاصل نمونه‌برداری از سیلاب‌های گذشته و مناطق سیل‌گیر در حوضه بوده، تحت بسته SDM و الگوریتم‌های جنگل تصادفی (RF) و ماشین بردار پشتیبان (SVM) نقشه پتانسیل سیل‌خیزی برای حوضه پیش‌بینی گردید. الگوریتم‌های مذکور با استفاده از منحنی‌های ROC مورد ارزیابی قرار گرفتند.

نتایج

تهیه نقشه حساسیت به سیل به عنوان یک گام ضروری برای جلوگیری و مدیریت سیل آینده شناخته می‌شود. با این حال، سیل شرایط پیچیده‌ای دارد که پیش‌بینی قابل اعتماد را دشوار می‌کند. با انتخاب ۱۶ پارامتر مؤثر در وقوع سیلاب، نقشه رستری هر کدام از این عوامل تهیه گردید (شکل ۲).

مور و همکاران [۱۷] این شاخص را پیشنهاد دادند. شاخص SPI بیان‌کننده قدرت جریان آب به لحاظ فرسایش در یک نقطه مشخص در سطح توپوگرافیک است. چنانچه سطح حوزه آبخیز و شیب آن افزایش پیدا کند، مقدار آب مشارکت‌کننده در مناطق بالادست شیب‌ها و سرعت آن بالا می‌رود، از این رو، شاخص SPI و میزان خطر فرسایش زیاد می‌شود فلورینسکی [۸]. توان آبراهه بیانگر قدرت فرساینده‌گی جریان بوده که با درجه شیب و مساحت حوضه رابطه مستقیم دارد. این شاخص با استفاده از رابطه زیر قابل محاسبه است مور و همکاران [۱۷].

$$SPI = A_s \tan \beta \quad (1)$$

نقشه شاخص SPI بر پایه لایه DEM در نرم‌افزار SAGA GIS 7.2 بدست آمد.

شاخص رطوبت توپوگرافی (TWI)

شاخص TWI اولین بار توسط بیون و کرکبی [۲] در سال ۱۹۷۹ در مدل بارش - رواناب TopModel معرفی شد که می‌تواند تأثیر توپوگرافی بر روی تولید رواناب را به صورت کمی بیان کند و به عنوان یک شاخص فیزیکی، مکان مناطق اشباع سطحی و توزیع مکانی رطوبت خاک را تقریب بزند شین و همکاران [۱۸]. شاخص TWI یک پارامترسازی ساده ریاضی از وضعیت رطوبت خاک است که به صورت رابطه زیر تعریف می‌شود:

$$TWI = Ln\left(\frac{\alpha}{\tan \beta}\right) \quad (2)$$

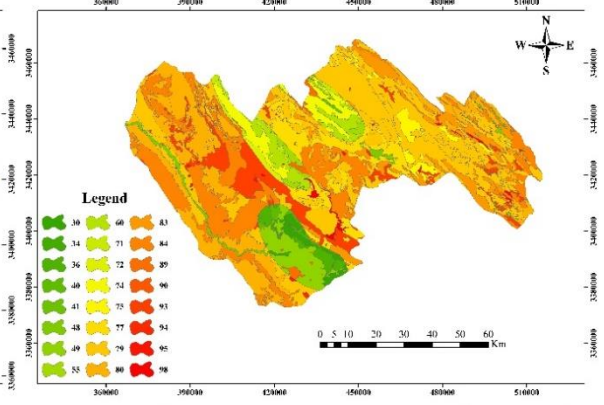
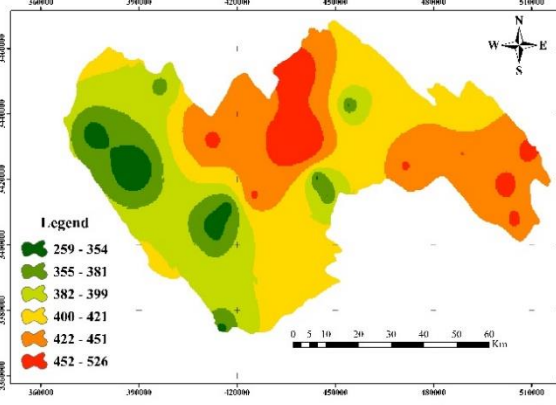
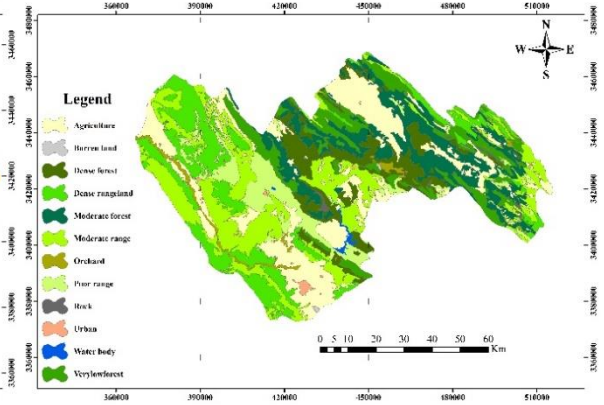
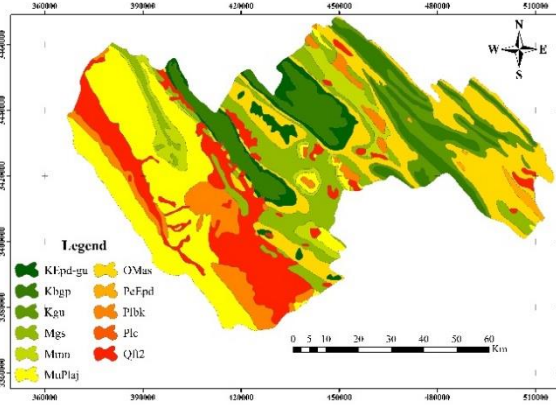
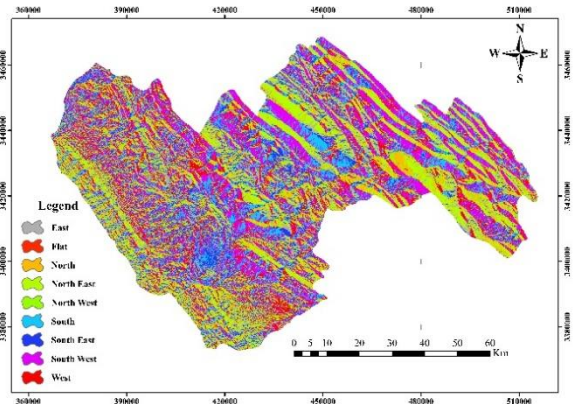
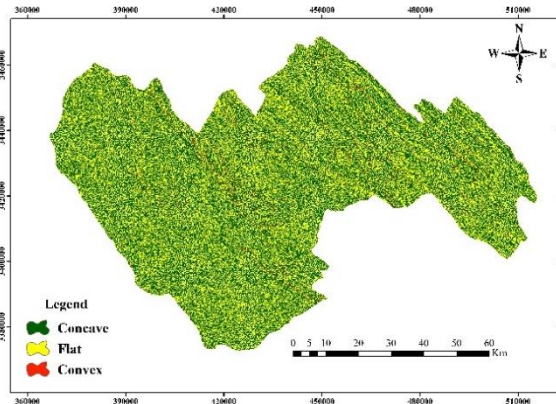
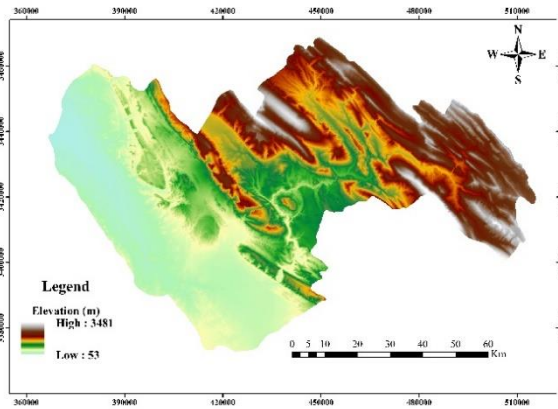
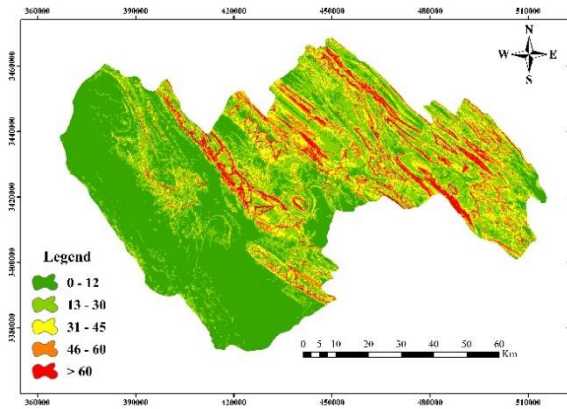
که در آن α : مساحت بالادست در واحد طول خط تراز یا مساحت ویژه حوضه و $\beta \tan$ شیب زمین (برحسب درجه) شین و همکاران [۱۸]. مقادیر بالای TWI مربوط به مناطقی است که انباشت آب و رواناب زیاد را دوست دارند. مقادیر TWI بالا به معنی زهکشی کم است که نشان دهنده زمین‌های اشباع بیشتر است که ممکن است باعث سیل ناگهانی شود فرناندز و لوتز [۷]. نقشه شاخص TWI بر پایه لایه DEM در نرم‌افزار SAGA GIS 7.2 تولید گردید.

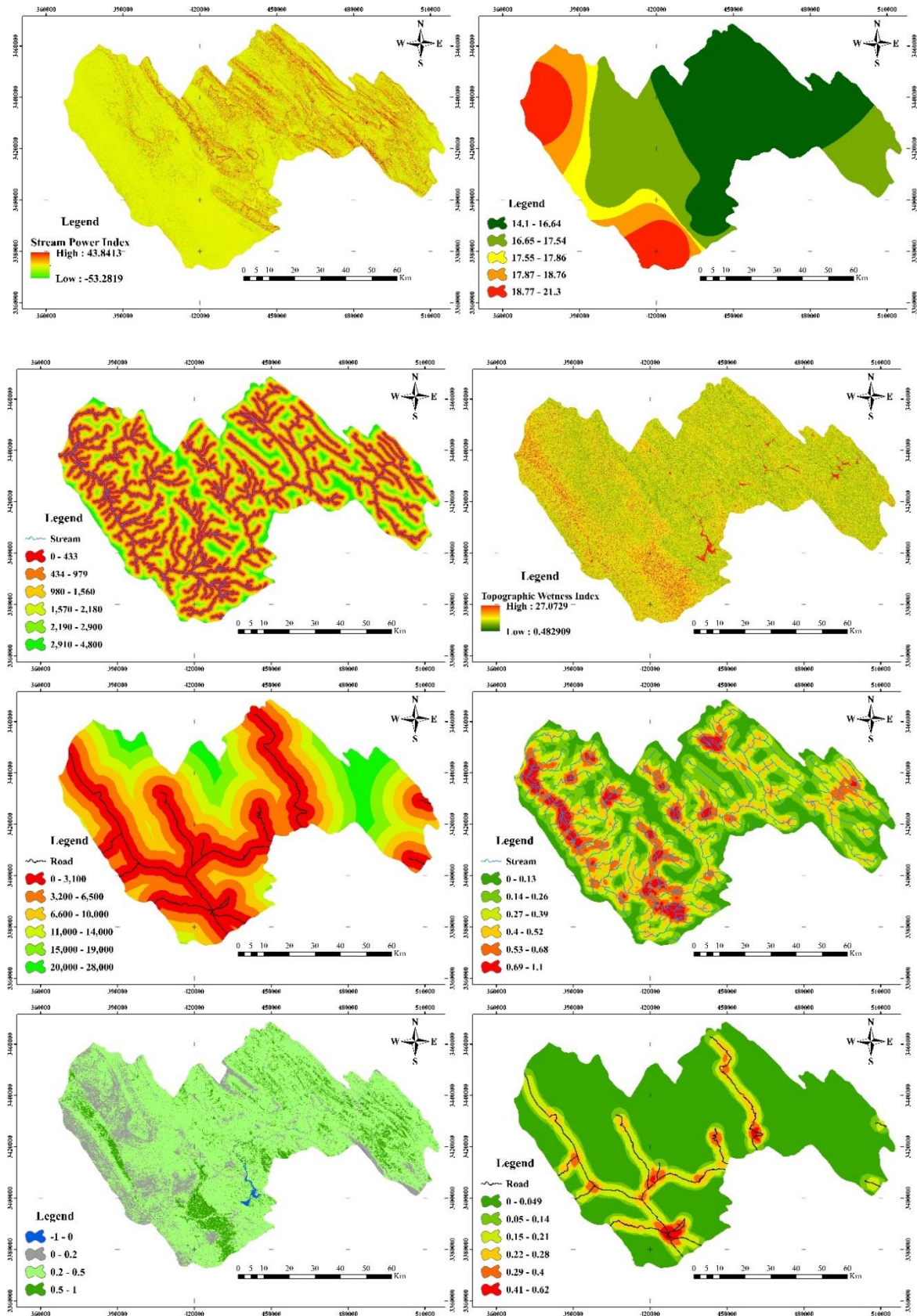
فاصله از آبراهه و تراکم آبراهه

اکثر مناطق مستعد وقوع سیلاب در حاشیه رودخانه‌ها قرار دارند. نقشه شبکه آبراهه حوضه از لایه DEM و با ابزار Hydrology در ArcMap 10.8 تولید گردید. جهت تولید نقشه فاصله از آبراهه از ابزار Euclidean Distance و تولید نقشه تراکم آبراهه از ابزار Line Density در نرم‌افزار ArcMap 10.8 استفاده گردید.

فاصله از جاده و تراکم جاده

جاده‌ها به عنوان سطوح با ضریب نفوذپذیری خیلی کم در حوزه‌های آبخیز نقش مهمی در ایجاد رواناب دارند. وجود مناطق با ضریب نفوذپذیری زیاد در سطح حوضه تبدیل بارش به رواناب را تسهیل کرده و شرایط را برای تشکیل جریان‌های سیلابی مهیا می‌سازند روشن [۱۹]. همچنین تراکم و گذر جاده‌ها در حاشیه رودخانه‌ها در هنگام وقوع سیلاب افزایش آسیب‌پذیری را به دنبال دارد. برای تهیه لایه‌های فاصله از جاده و تراکم آن از نقشه توپوگرافی با مقیاس ۱:۵۰۰۰۰ استفاده گردید.

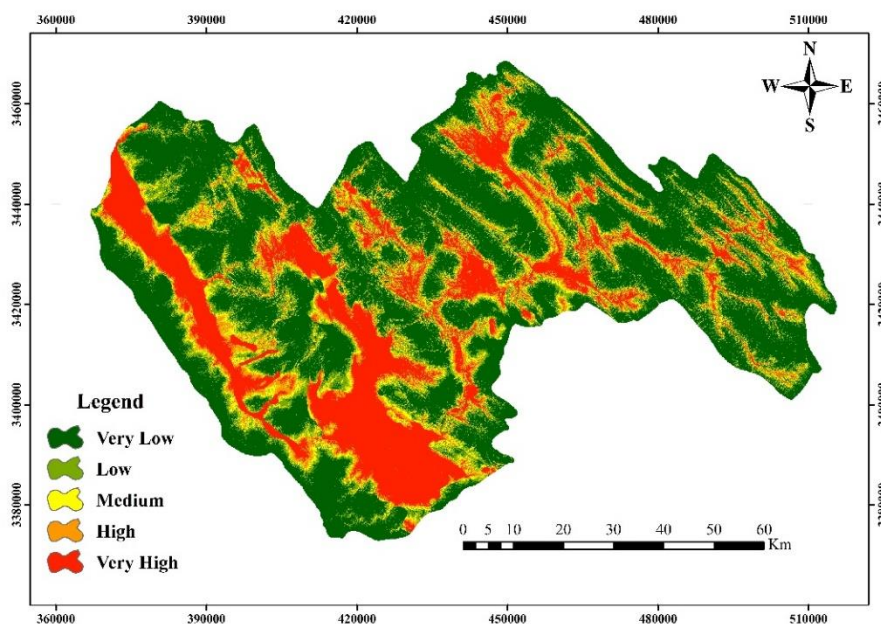




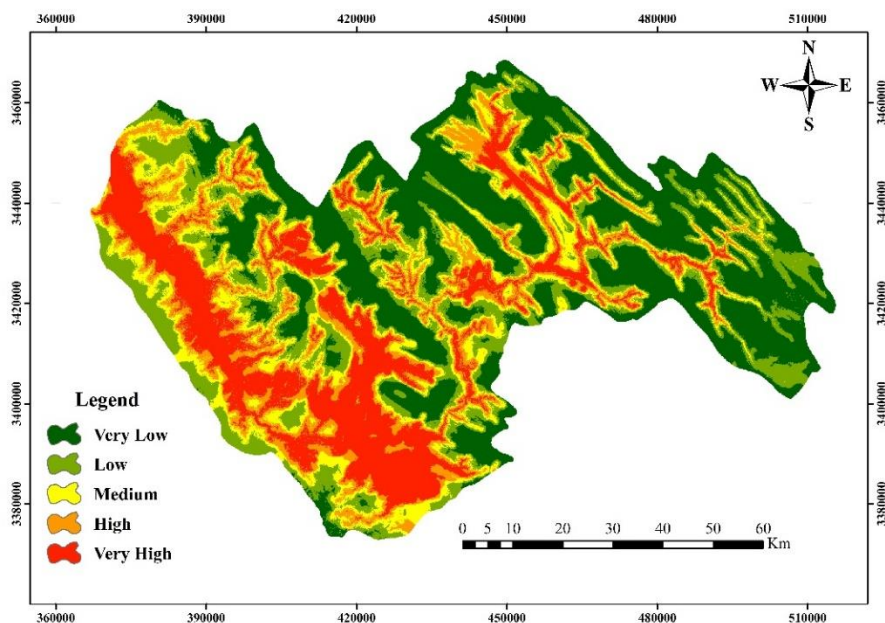
شکل ۲- پارامترهای مؤثر در وقوع سیل
 Fig 2. Effective Parameters In Flood Occurrence

(SVM) اجرا شدند. نتایج نهایی تولید نقشه پتانسیل سیل خیزی با استفاده از این دو الگوریتم در شکل ۳ و ۴ قابل مشاهده است. ارزیابی الگوریتم‌های SVM و RF به منظور ارزیابی مدل‌های بکار گرفته شده در تولید نقشه پتانسیل سیل خیزی حوضه از منحنی مشخصه عملکرد (منحنی ROC) استفاده شد که نتایج آن در شکل‌های ۵ و ۶ قابل مشاهده است.

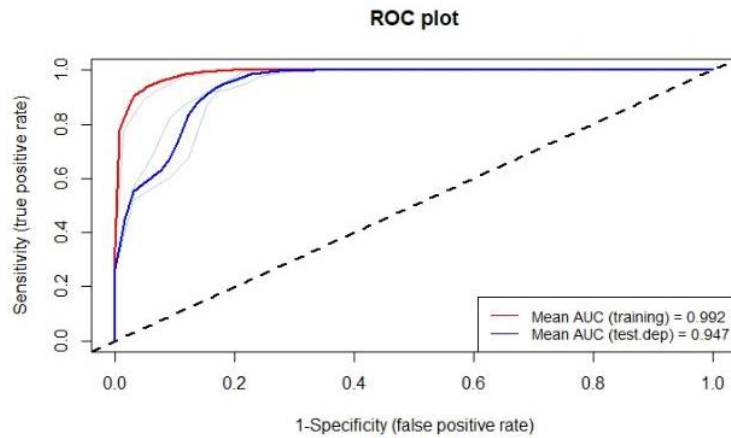
اجرای الگوریتم‌های جنگل تصادفی (RF) و ماشین بردار پشتیبان (SVM) بعد از تهیه لایه‌های اطلاعاتی مؤثر در وقوع سیلاب در محیط نرم‌افزارهای مختلف، همه این لایه‌ها به فرمت قابل خوانا برای نرم‌افزار R تبدیل شدند. در محیط نرم‌افزار R با استفاده از بسته SDM، الگوریتم‌های جنگل تصادفی (RF) و ماشین بردار پشتیبان



شکل ۳- نقشه پتانسیل سیل خیزی حوضه مارون با مدل SVM
Fig 3. Flood potential map of Maroon Basin with SVM model

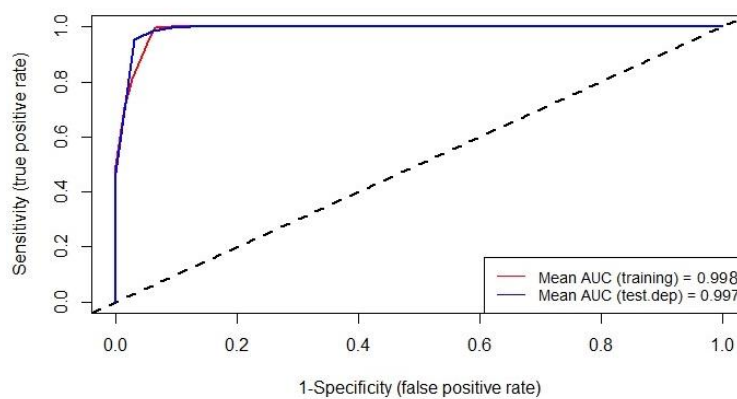


شکل ۴- نقشه پتانسیل سیل خیزی حوضه مارون با مدل RF
Fig 4. Flood potential map of Maroon Basin with RF model



شکل ۵- نتایج ارزیابی مدل SVM با استفاده از منحنی ROC

Fig 5. Evaluation results of SVM model using ROC curve



شکل ۶- نتایج ارزیابی مدل RF با استفاده از منحنی ROC

Fig 6. Results of evaluation of RF model using ROC curve

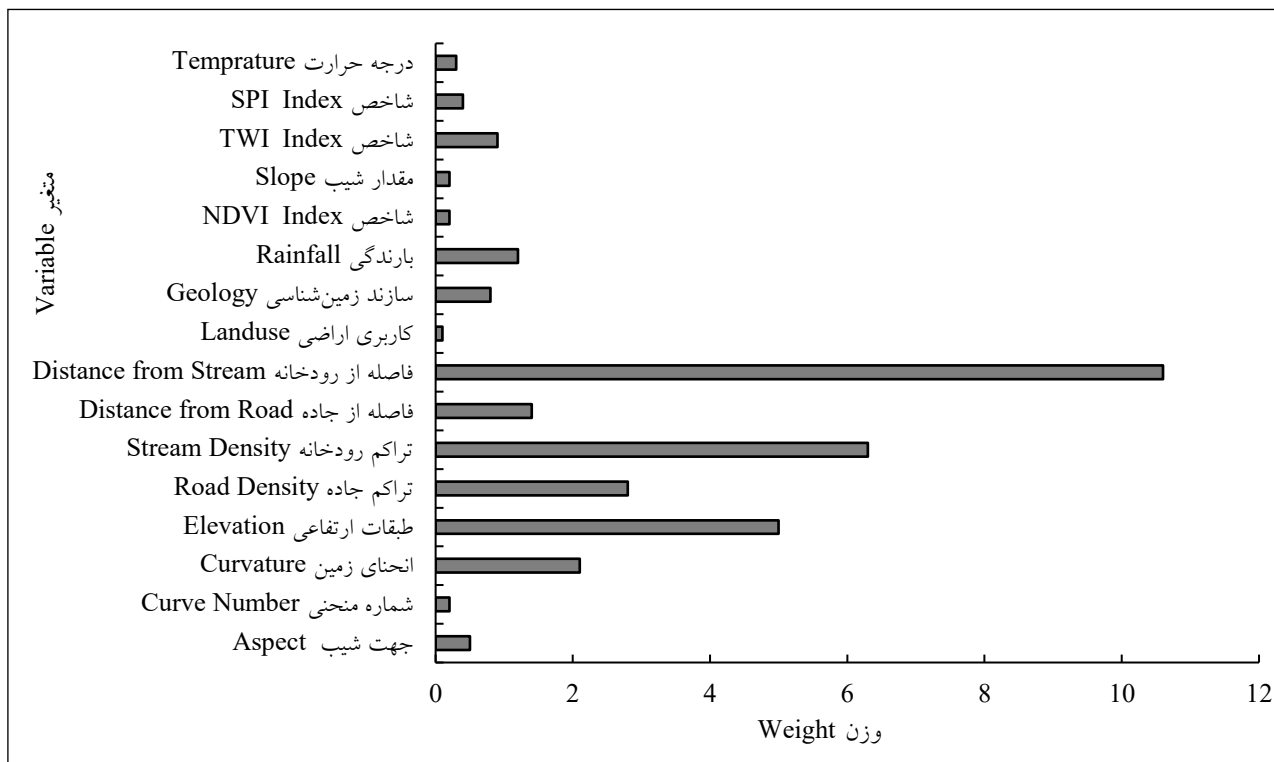
جدول ۲- نتایج ارزیابی مدل‌های RF و SVM با استفاده از منحنی ROC

Table 2. Evaluation results of RF and SVM models using ROC curve

ROC-AUC		مدل Model
Testing	Training	
0.997	0.998	جنگل تصادفی Random Forest
0.947	0.992	ماشین بردار پشتیبان Support Vector Machine

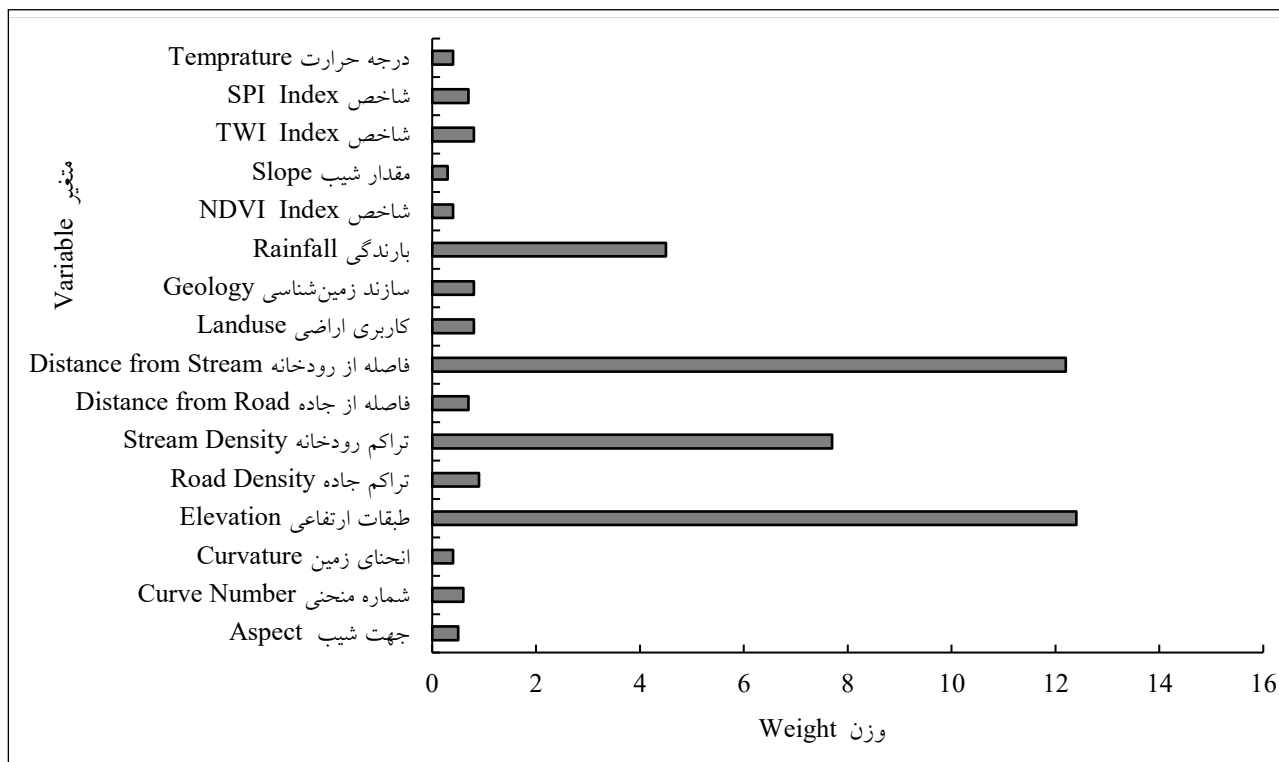
مهم‌ترین پارامترهای مؤثر در وقوع سیلاب بر اساس اولویت تأثیر با استفاده از دو مدل ماشین بردار پشتیبان (SVM) و جنگل تصادفی (RF) نیز تعیین گردید (شکل‌های ۷ و ۸).

نتایج حاصل از ارزیابی مدل‌های RF و SVM با استفاده از منحنی تشخیص عملکرد نسبی (ROC) و سطح زیر آن (AUC) در دو مرحله آموزش و تست در جدول ۲ نشان داده شده است. نتایج اهمیت و اولویت عوامل مؤثر در سیل‌خیزی حوضه مارون



شکل ۷- مهم ترین عوامل مؤثر در وقوع سیلاب بر اساس مدل SVM

Fig 7. The most important factors in the flood occurrence based on the SVM model



شکل ۸- مهم ترین عوامل مؤثر در وقوع سیلاب بر اساس مدل RF

Fig 8. The most important factors in the flood occurrence based on the RF model

بحث و نتیجه‌گیری

در صورت عدم برنامه‌ریزی برای کنترل و مدیریت سیلاب در این حوضه خسارات این بخش نیز بسیار بالا خواهد بود. مناطق با خطر پایین و بسیار پایین دارای کاربری از نوع جنگل و مرتع بوده که نشان می‌دهد این نوع کاربری در جلوگیری از تبدیل بارش به رواناب تأثیر زیادی داشته و موجب شده که بیشتر باران به داخل نفوذ کرده یا سرعت جریان رواناب را کاهش دهد. از نظر تأثیر واحدهای زمین‌شناسی و واحدهای لیتولوژیک نیز سازندهایی که معمولاً از نوع سخت و صخره‌ای هستند فرصت نفوذ بارش به رواناب را فراهم کرده و در سیل‌خیزی نقش مهمی دارند در صورتی که در مناطق پوشیده شده از واحدهای نفوذپذیری که اغلب واحدهای تراس‌های آبراه‌ای مربوط به دوران چهارم را شامل می‌شوند بر روی مناطق سیل‌گیر منطبق شده‌اند و دارای خطر بالا و بسیار پایین هستند که دلیل عمده قرار گیری این مناطق در حاشیه آبراه‌ها و رودخانه‌ها بوده و ارتفاع آنها نیز پایین و در محدوده واحد دشت قرار می‌گیرند. بر اساس نقشه‌های نهایی حاصل از دو مدل SVM و RF مناطق قابل تشخیص شاخص SPI بالایی دارند به عنوان مناطق مولد سیل قابل تشخیص هستند و مناطق دارای شاخص SPI پایین به عنوان مناطق سیل‌گیر قابل مشاهده هستند و در مورد شاخص TWI نیز به همین ترتیب است. با توجه به اینکه مناطق مؤثر در سیل‌خیزی حوضه بیشتر در مناطق بالادست و آبراه‌های با شماره کمتر بودند پیشنهاد می‌شود در این حوضه برای کاهش خسارات و کاهش سرعت سیلاب در پایین‌دست، از عملیات آبخیزداری در مناطق بالادست و سرشاخه‌های آبراه اصلی بهره‌گیری شود.

منابع

1. Bathrellos, G.D. Skilodimou, H.D. Soukis, K. and Koskeridou, E. 2018. Temporal and spatial analysis of flood occurrences in the drainage basin of pinios river (thessaly, central greece). Land. 7(3): 1-18.
2. Beven, K.J. and Kirkby, M.J.A. 1979. Physically based, variable contributing area model of basin hydrology/Un modèle à base physique de zone d'appel variable de l'hydrologie du bassin versant. Hydrological Sciences Bulletin. 24(1): 43-69.
3. Botzen, W. Aerts, J. and van den Bergh, J. 2013. Individual preferences for reducing flood risk to near zero through elevation. Mitigation and Adaptation Strategies for Global Change. 18(2): 229-244.
4. Büchele, B. Kreibich, H. Kron, A. Thielen, A. Ihringer, J. Oberle, P. Merz, B. and Nestmann, F. 2006. Flood-risk mapping: contributions towards an enhanced assessment of extreme events and associated risks. Natural Hazards and Earth System Sciences. 6(4): 485-503.
5. Fernandez, D.S. and Lutz, M.A. 2010. Urban flood hazard

در این تحقیق ۱۶ پارامتر مؤثر در وقوع سیلاب بصورت لایه اطلاعاتی در محیط نرم‌افزارهای ArcGIS 10.8, ENVI 5.3 و SAGA GIS 7.2 استخراج شدند. سپس با بکارگیری روش‌های ماشین یادگیرنده جنگل تصادفی (RF) و ماشین پشتیبان (SVM) و لایه آموزشی (training) حاصل از عملیات میدانی و برداشت نمونه‌ها در قالب ۱۸۵ نقطه (۱۳۰ وقوع سیلاب و ۵۵ نقطه در مناطق عدم وقوع سیلاب) نقشه نهایی حساسیت سیل‌خیزی حوضه تهیه گردید. نتایج حاصل از نقشه نهایی سیلاب بدست آمده از مدل ماشین بردار پشتیبان (SVM) نشان می‌دهد که سطح وسیعی از حوضه مارون در معرض سیلاب قرار دارد که اغلب شامل اراضی کم شیب و کمتر از ۱۲ درصد می‌باشد. بدلیل اختلاف توپوگرافی در این حوضه که از دشت‌های استان خوزستان، شهرستان بهبهان شروع شده تا ارتفاعات زاگرس جنوبی در استان کهگیلویه و بویراحمد را شامل می‌شود، بر طبق شکل ۳ اغلب مناطق با ارتفاعات بالا و شیب زیاد به عنوان عامل اصلی در سیل‌خیزی و مناطق پایین‌دست و با ارتفاع کمتر مستعدترین مناطق در سیل‌گیری بحساب می‌آیند. بر اساس نقشه کاربری اراضی نیز می‌توان نتیجه گرفت که محدوده شهری بهبهان در معرض سیلاب قرار دارد. همچنین در مناطق اراضی کشاورزی و حاشیه رودخانه‌ها نیز خطر سیلاب بسیار زیاد است. همچنین نقشه نهایی سیلاب حاصل از مدل جنگل تصادفی (RF) در شکل ۴ نیز بیانگر وقوع سیلاب بیشتر در مناطق پست‌تر است و نقش اراضی با ارتفاع و شیب بالا در سرعت بخشیدن به سرعت رواناب و نقش آنها در ایجاد سیلاب در اراضی پایین‌دست می‌باشد که با نتایج اول و ژانگ [۲۵] و باترلوس و همکاران [۱] همخوانی دارد. بر طبق این شکل نیز مناطق زیادی از حوضه در معرض سیلاب قرار دارند که در مطالعات مربوط به مدیریت و کنترل سیلاب به منظور کاهش خسارات باید در اولویت قرار بگیرند. بنابراین بر اساس نتیجه این دو مدل می‌توان نتیجه‌گیری کرد حوضه مارون از نظر حساسیت سیلاب در طبقه حساسیت بالا قرار دارد و لازم است مطالعات مدیریت سیلاب و کنترل آن بطور جدی انجام شود. بر اساس جدول ۲ نتایج ارزیابی مدل‌های جنگل تصادفی و ماشین بردار پشتیبان نشان می‌دهد که دقت مدل جنگل تصادفی نسبت به مدل ماشین بردار پشتیبان در پیش‌بینی نقشه پتانسیل سیل‌خیزی بیشتر بوده است. همچنین با توجه به نتایج دو مدل SVM و RF مهم‌ترین عوامل مؤثر در وقوع سیلاب در حوزه آبخیز مارون شامل فاصله از رودخانه، تراکم رودخانه و طبقات ارتفاعی (توپوگرافی زمین) می‌باشند نتایج این بخش با مطالعه ما و همکاران [۱۳] مطابقت دارد. از طرفی بر اساس نقشه نهایی حاصل از این دو مدل (شکل‌های ۳ و ۴) مناطق با خطر بسیار زیاد، تا حدودی منطبق بر پارامتر جاده در حوضه بوده که بیانگر اینست که این نوع کاربری در حوضه در معرض تخریب در اثر سیل قرار دارد. همچنین مناطق با خطر بالا و بسیار بالا بر روی اراضی کشاورزی منطبق هستند که

biological applications. *Hydrological processes*. 5(1): 3-30.

18. Qin, C.Z. Zhu, A.X. Pei, T. Li, B.L. Scholten, T. Behrens, T. and Zhou, C.H. 2011. An approach to computing topographic wetness index based on maximum downslope gradient. *Precision Agriculture*. 12(1): 32-43.

19. Roshan, S.H. 2020. Flood Analysis using Data Mining Techniqe and SWAT Model in Northern Karun. Ph.D. Thesis in Watershed Management and Sciences, Faculty of Natural Resources, Sari Agricultural Sciences and Natural Resources University. (In Persian)

20. Shaabani Bazneshin, A. Emadi, A. and Fazloula, R. 2017. Investigation the Flooding Potential of Basins and Determination Flood Producing Areas (Case Study: NEKA Basin). *Journal of Watershed Management Research*. 7(14): 28-20. (In Persian)

21. Sharifi, F. Saghafian, B. and Telvari, A. 2002. The Great 2001 Flood in Golestan Province. Iran: Causes and Consequences. International Conference on Flood Estimation, Berne, Switzerland, 263-271.

22. Sidle, R.C. and Ochiai, H. 2006. Landslides: processes, prediction, and landuse. American Geophysical Union, Washington, D.C. *Water Res Monograph*. 18(1): 312.

23. Srivastava, O.S. Denis, D. M. Srivastava, S. K. Kumar, M. and Kumar, N. 2014. Morphometric analysis of a Semi Urban Watershed, trans Yamuna, draining at Allahabad using Cartosat (DEM) data and GIS. *International Journal of Engineering Science*. 3(11): 71-79.

24. Telvari, A. 1997. Effecting Factor in Flood Occurrence or Sevirity and Damages. *River Flood Control Workshop*, 138-159. (In Persian)

25. Ullah, K. and Zhang, J. 2020. GIS-based flood hazard mapping using relative frequency ratio method: A case study of Panjkora River Basin, eastern Hindu Kush, Pakistan. *Plos one*. 15(3): e0229153.

26. Valizadeh Kamran, K. Delire Hasannia, R. and Azari Amghani, K. 2019. Flood zoning and its impact on land use in the surrounding area using unmanned aerial vehicles (UAV) images and GIS. *Journal of RS and GIS for Natural Resources*, 10(3): 59-75. (In Persian)

27. Youssef, A. M. Pradhan, B. Dikshit, A. and Mahdi, A. M. 2022. Comparative study of convolutional neural network (CNN) and support vector machine (SVM) for flood susceptibility mapping: a case study at Ras Gharib, Red Sea, Egypt. *Geocarto International*. 1-28.

zoning in Tucumán Province, Argentina, using GIS and multicriteria decision analysis. *Engineering Geology*. 111(1-4): 90-98.

6. Ding, L. Ma, L. Li, L. Liu, C. Li, N. Yang, Z. Yao, Y. and Lu, H. 2021. A Survey of Remote Sensing and Geographic Information System Applications for Flash Floods. *Remote Sensing*. 13(9): 1818.

7. Fernández, D.S. and Lutz, M.A. 2010. Urban flood hazard zoning in Tucumán Province, Argentina, using GIS and multicriteria decision analysis. *Engineering Geology*. 111(1-4): 90-98.

8. Florinsky, I. 2011. *Digital Terrain Analysis in Soil Science and Geology*, first ed. Academic Press, ISBN-13: 978-0123850362.

9. Garde, R.J. 2006. *River Morphology*, Published by New Age International (P) Ltd., Publishers, New Delhi.

10. Kia, M.B. Pirasteh, S. Pradhan, B. Mahmud, A.R. Sulaiman, W.N.A. and Moradi, A. 2012. An artificial neural network model for flood simulation using GIS: Johor River Basin, Malaysia. *Environment Earth Science*. 67(1): 251-264.

11. Lee, M.J. Kang, J.E. and Jeon, S. 2012. Application of frequency ratio model and validation for predictive flooded area susceptibility mapping using GIS. *IEEE International of Geoscience and Remote Sensing Symposium (IGARSS)*, 895–898.

12. Liu, J. Wang, J. Xiong, J. Cheng, W. Li, Y. Cao, Y. He, Y. Duan, Y. He, W. and Yang, G. 2022. Assessment of flood susceptibility mapping using support vector machine, logistic regression and their ensemble techniques in the Belt and Road region. *Geocarto International*, (just-accepted), 1-40.

13. Ma, M. Liu, C. Zhao, G. Xie, H. Jia, P. Wang, D. Wang, H. and Hong, Y. 2019. Flash flood risk analysis based on machine learning techniques in the Yunnan Province, China. *Remote Sensing*. 11(2): 170.

14. Mabuku, M.P. Senzanje, A. Mudhara, M. Jewitt, G. and Mulwafu, W. 2018. Rural households' flood preparedness and social determinants in Mwandia district of Zambia and Eastern Zambezi Region of Namibia. *International journal of disaster risk reduction*. 28: 284-297.

15. Mahmoudzadeh, H. and Bakooyi, M. 2019. Flood zoning using fuzzy analysis (case study: Sari city). *Journal of Natural Environmental Hazards*. 7(18): 51-68. (In Persian)

16. Moayeri, M. and Entezari, M. 2008. Floods And Reviow Floods In Province Of Esfahan. *Journal of Studies of Human Settlements Planning*. 3(6): 110-124. (In Persian)

17. Moore, I.D. Grayson, R.B. and Ladson, A.R. 1991. Digital terrain modeling: a review of hydrological, geomorphological and

Evaluation of Machine Learning Algorithms (RF and SVM) in Producing Flood Susceptibility Mapping Maroon Watershed

M.A. Kiani Asl¹, B. Motesaffeh² and S.H. Roshaan³

Received: 30-05-2018 Accepted: 10-10-2022

Abstract

Floods are one of the most important natural hazards that have caused economic and social damage in most areas. Numerous climatic, hydrological, geomorphological and geological factors are involved in the occurrence of floods. Flood analysis, management and control can be done by preparing flood potential maps. The purpose of this study is to map the flood potential of Maroon watershed using random forest and support vector machine machine learning methods. For this purpose, 16 effective parameters in flood occurrence including altitude, slope and aspect, curvature, geological formations, land use, curve number, rainfall, temperature, stream power index (SPI), topographic wetness index (TWI), distance From stream, stream density, road distance, road density and NDVI index were considered. The mentioned parameters were prepared in ArcGIS 10.8, ENVI 5.3 and SAGA GIS 7.2 software environments and then converted to readable format for R software environment in order to implement the support vector machine and random forest models. Finally, RF and SVM models were implemented using SDM packages and evaluated using receiver operating characteristic (ROC). The results showed that RF and SVM models correctly predicted the flooding map of Maroon Basin with 0.997 and 0.947 percent accuracy, respectively.

Keywords: *Flood management, Flood occurrence, Maroon basin, Random forest, Support vector machine*

1. Master of Watershed Engineering, Department of Rangeland and Watershed Management, Faculty of Natural Resources, Behbahan Khatam Alanbia University of Technology, Behbahan, Behbahan, Iran.
2. Corresponding Author and Assistant Professor, Department of Rangeland and Watershed Management, Faculty of Natural Resources, Behbahan Khatam Alanbia University of Technology, Behbahan, Iran, Email: bmoteshaffeh@gmail.com
3. PhD, Watershed Management Science and Engineering, Department of Watershed Management Engineering, Faculty of Natural Resources, Sari University of Agricultural Sciences and Natural Resources, Sari, Iran.