

مربوط به هر خطر یک نقشه‌ی واحد و جامع که نشان‌دهنده‌ی مناطق مستعد هر دو خطر باشد به دست آمد. نقشه نهایی می‌تواند به‌عنوان ابزاری باارزش برای برنامه‌ریزی پایدار برای کاربری اراضی در مناطق مستعد چند خطر مورد استفاده قرار گیرد.

کلیدواژه‌ها: پیش‌بینی مکانی، مخاطرات طبیعی، مدل‌های یادگیری ماشینی، منحنی تشخیص عملکرد نسبی، نقشه دو خطر.

مقدمه

زمین‌لغزش‌ها یکی از مخاطرات طبیعی هستند که پس از زلزله و سیل، بیشترین خسارات را به انسان وارد می‌کنند [۹]. همچنین فرسایش‌هایی از جمله فرسایش خندقی (آبکندی) دارای بیش‌ترین توان تولید رسوب در اکثر حوزه‌های آبخیز کشور است که توجه به آن از ارکان اصلی حفاظت خاک و مدیریت پایدار منابع طبیعی محسوب می‌شود [۲۰]. تبیین و به‌کارگیری مدل‌های یادگیری ماشینی و داده‌کاوی برای پیش‌بینی مکانی مناطق مستعد وقوع خطر فرسایش آبکندی و زمین‌لغزش به‌صورت منفرد توسط پژوهشگران مختلفی در سراسر جهان مورد بررسی قرار گرفته‌است، از جمله [۸، ۱۹، ۲۳]. در ایران هم محققان زیادی با استفاده از مدل‌های مختلف یادگیری ماشینی به مطالعه و شناسایی مناطق مستعد این نوع خطرات پرداخته‌اند [۱۴، ۱۵، ۱۷، ۲۲]. مدل‌های جنگل تصادفی (RF) و ماشین بردار پشتیبان (SVM) از مدل‌های شناخته‌شده و پیشرفته داده‌کاوی هستند که به دلیل مزایا و قابلیت‌های آن تحقیقات مختلفی [۱۰، ۲۴] برای پیش‌بینی و مدل‌سازی زمین‌لغزش و فرسایش خندقی مورد استفاده قرار گرفته‌اند. در بیشتر مطالعات انجام‌شده برای کنترل و مدیریت خطرات طبیعی حوزه‌های آبخیز تنها یک خطر به‌صورت منفرد مورد بررسی قرار گرفته است و این سبب عدم شناخت صحیح از وضعیت آسیب‌پذیری منطقه می‌شود. تهیه نقشه احتمال و حساسیت خطرات چندگانه یکی از راهکارهای مؤثر جهت شناسایی و پیش‌بینی مناطق آسیب‌پذیر است [۱۳]. ارزیابی چند خطر و شناسایی مناطق مستعد آن‌ها می‌تواند به آمایش منطقه و مدیریت جامع حوزه‌های آبخیز کمک شایانی کند و برای برنامه‌ریزی درست در راستای توسعه پایدار شناخت همه‌جانبه از عوامل مؤثر محیطی امری لازم و ضروری است. در سال‌های اخیر مطالعاتی انجام شده

شناسایی مناطق مستعد فرسایش خندقی و زمین‌لغزش در قالب نقشه دو خطر با استفاده از مدل‌های یادگیری ماشینی در حوزه آبخیز گرگانرود

نرگس جاویدان^۱، عطااله کاویان^۲، سجاد رجبی^۳، حمیدرضا پورقاسمی^۴
و زینب جعفریان^۵

تاریخ دریافت: ۱۴۰۰/۰۸/۰۴ تاریخ پذیرش: ۱۴۰۱/۰۱/۳۱

DOI: 10.22034/17.62.75

چکیده

پژوهش حاضر با هدف شناسایی مناطق مستعد فرسایش خندقی و زمین‌لغزش به‌صورت نقشه‌ی دو خطر در قالب یک نقشه‌ی واحد در حوزه آبخیز گرگانرود انجام شد. علاوه‌برآن، استفاده از دو مدل یادگیری ماشینی SVM و RF برای تعیین ارتباط میان این خطرات و متغیرهای زمین - محیطی و همچنین تهیه نقشه مناطق مستعد وقوع هر خطر به‌صورت منفرد به‌عنوان هدف ثانویه مدنظر قرار گرفت. برای اعتبارسنجی نقشه‌های استعداد وقوع هر خطر، از منحنی ROC استفاده شد، و بهترین مدل برای تعیین نقشه حساسیت‌پذیری هر خطر با بالاترین دقت انتخاب شد. در نهایت با ادغام نقشه حساسیت مربوط به خطرات زمین‌لغزش و فرسایش خندقی با ترکیب مدل‌های مختلف و مدل برتر، نقشه حساسیت دو خطر به دست آمد. نتایج نشان داد مدل RF با $AUC=82/9$ برای خطر زمین‌لغزش و $AUC=96/9$ برای فرسایش خندقی نسبت به مدل SVM با مقدار $AUC=77/0$ برای خطر زمین‌لغزش و $AUC=93/9$ برای فرسایش خندقی دارای دقت بالاتری است. در نهایت با ادغام نقشه حساسیت

- ۱- دانش‌آموخته دکتری علوم و مهندسی آبخیزداری، دانشکده منابع طبیعی، دانشگاه علوم کشاورزی و منابع طبیعی ساری، ایران
- ۲- استاد گروه مرتع و آبخیزداری، دانشکده منابع طبیعی، دانشگاه علوم کشاورزی و منابع طبیعی ساری، ایران نویسنده مسئول: ataollah.kavian@gmail.com
- ۳- دانش‌آموخته کارشناسی ارشد آبخیزداری، دانشکده منابع طبیعی، دانشگاه علوم کشاورزی و منابع طبیعی ساری
- ۴- استاد گروه مهندسی منابع طبیعی و محیط زیست، دانشکده کشاورزی، دانشگاه شیراز، ایران
- ۵- دانشیار گروه مرتع و آبخیزداری، دانشکده منابع طبیعی، دانشگاه علوم کشاورزی و منابع طبیعی ساری، ایران

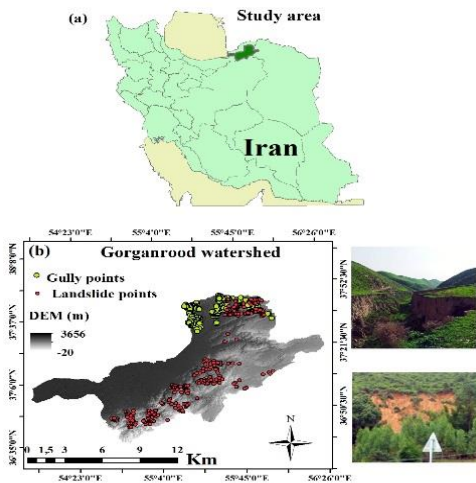
است که به بررسی دو یا چند خطر در قالب یک نقشه‌ی واحد پرداخته‌اند: [۵، ۱۶]. حوزه آبخیز گرگانرود همانند بسیاری از مناطق کشور از وقوع مخاطرات مستثنا نیست. به دلیل وجود اراضی شیب‌دار موجود در منطقه مورد مطالعه، پدیده زمین‌لغزش از حوادث مخرب منطقه بوده است. از طرفی دیگر، می‌توان مشاهده نمود که فرسایش خندقی در پایین‌دست حوزه آبخیز گرگانرود استان گلستان توسعه زیادی یافته‌اند و یکی از مشکلات اساسی حوزه‌های آبخیز استان گلستان را نرخ بالای فرسایش و تخریب شامل می‌شود. فرسایش خاک در این استان سالانه در مناطق جنگلی ۵-۶ تن در هکتار است، در حالی که این رقم در خاک‌های لسی (خاک‌های حاصلخیز و دانه‌ریز) به‌ویژه در شهرستان‌های کلاله و مراوه‌تپه به مرز ۳۵ تن در هکتار نیز می‌رسد (گزارشی از اداره کل منابع طبیعی و آبخیزداری گلستان).

غالباً سرمایه‌های زیادی از کشور صرف زیرساخت‌های حمل‌ونقل زمینی، خطوط انتقال نیرو و سدهای مخزنی شده است که در صورت عدم شناسایی مناطق مستعد این خطرات و مدیریت منطقه، قطعاً خسارت‌های جانی و مالی متعددی در اثر تخریب زیرساخت‌ها و تولید رسوب و پر شدن مخازن سدها به بار می‌آید؛ بنابراین تهیه نقشه حساسیت بخش‌های مختلف حوزه آبخیز در قالب یک نقشه و تعیین عوامل مؤثر بر آن‌ها بسیار ضروری است. نقشه‌های حساسیت این خطرات به‌منظور توسعه پایدار برای مدیریت یکپارچه حوزه آبخیز لازم است. همچنین نیاز به داشتن یک رویکرد دقیق و قابل‌اعتماد برای شناسایی مناطق مستعد این خطرات به‌صورت جامع و یکپارچه در بیشتر مناطق کشور احساس می‌شود.

مواد و روش‌ها

معرفی منطقه مطالعاتی

حوزه آبخیز گرگانرود در استان گلستان در بخش شمال شرقی ایران واقع شده است. مساحت این منطقه ۱۰۱۹۷ کیلومتر مربع است که در محدوده طول جغرافیایی ۳۶ درجه و ۳۴ دقیقه تا ۳۸ درجه و ۱۵ دقیقه شمالی و عرض جغرافیایی ۵۴ درجه و ۵ دقیقه تا ۵۶ درجه و ۸ دقیقه شرقی قرار گرفته است (شکل ۱). از نظر توپوگرافیکی، میانگین ارتفاع این منطقه بین ۲۰- تا ۳۶۵۲ متر است. میانگین بارش سالانه درازمدت این استان ۶۰۴ میلی‌متر است. همچنین میانگین بیشینه و کمینه دمایی حوضه به ترتیب ۱۸/۵ و ۱۱ درجه سانتی‌گراد است. بر طبق گزارش اداره منابع طبیعی و آبخیزداری استان گلستان، منطقه مورد مطالعه با برخی از سازندهای سنگ‌شناسی حساس مانند رسوبات کواترنری پوشیده شده است. وقوع بارش‌های شدید به همراه تشکیلات زمین‌شناسی حساس به فرسایش و تغییر شدید کاربری اراضی از جنگل به اراضی دیم، از جمله مشکلات این منطقه است [۷].



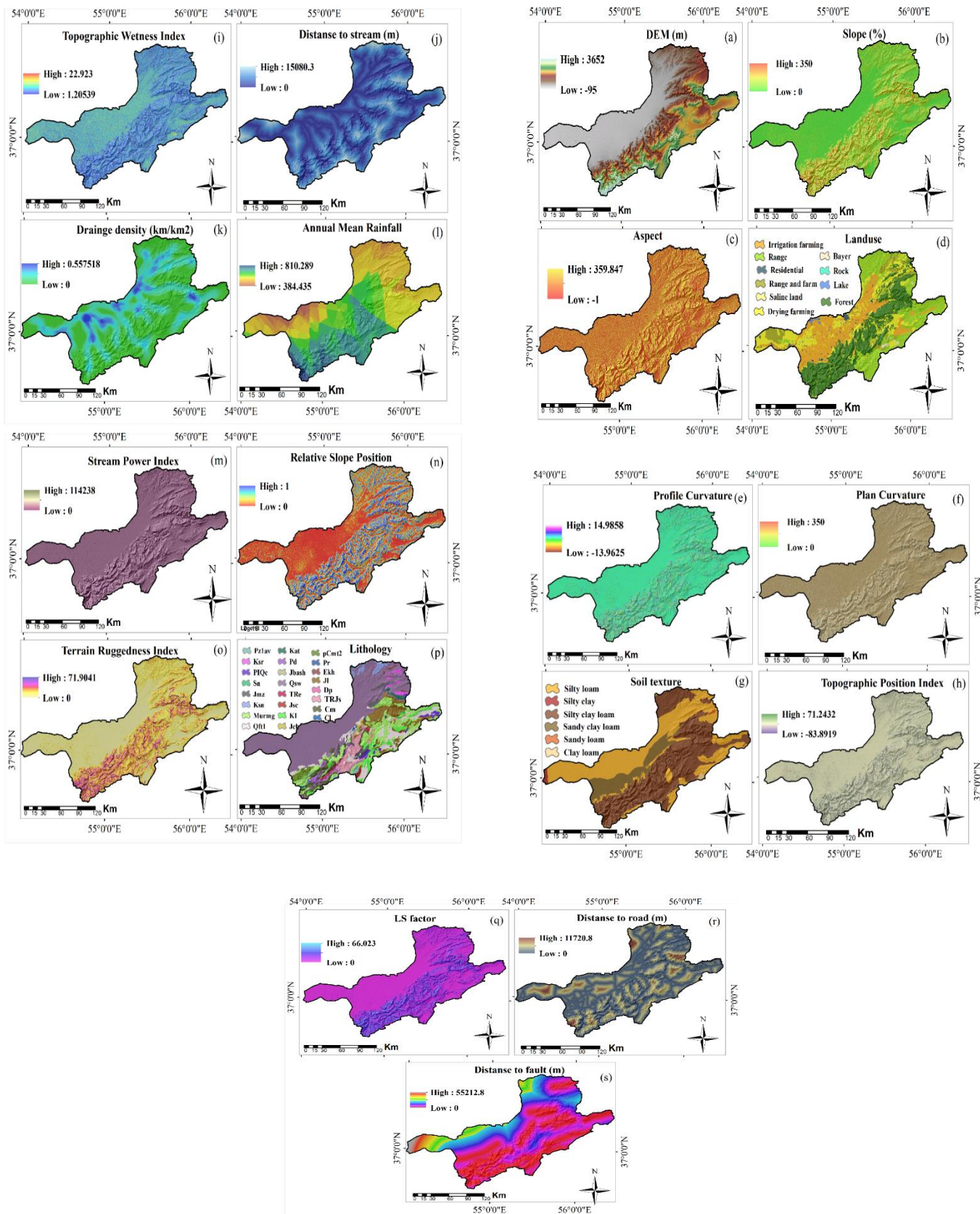
شکل ۱- موقعیت جغرافیایی منطقه مورد مطالعه در ایران
Fig1. Geographical location of the study area

روش تحقیق

آماده کردن نقشه پراکنش نقاط وقوع و تهیه فاکتورهای پیش‌بینی کننده

تهیه نقشه پراکنش خطر یک مرحله کلیدی در تهیه نقشه‌ی حساسیت‌پذیری است [۱۲]. موقعیت‌های وقوع خطرات طبیعی در مدل‌سازی به‌عنوان متغیر وابسته و عوامل محیطی مؤثر به‌عنوان متغیرهای مستقل مورد تجزیه و تحلیل قرار می‌گیرد [۳]. در این پژوهش، موقعیت مکانی ۲۸۳ نقطه فرسایش خندقی و ۳۵۱ نقطه زمین‌لغزش ثبت شده و بر اساس آن نقشه پراکنش موقعیت‌های این خطرات در نرم‌افزار ArcGIS 10.5 تهیه شد. الگوریتم تقسیم‌بندی تصادفی^۲ [۱۴] برای جدا کردن نقاط آموزشی و نقاط اعتبارسنجی مورد استفاده قرار گرفت؛ لذا از میان نقاط ثبت شده ۷۰ درصد به‌عنوان گروه آموزش و ۳۰ درصد باقی‌مانده به‌عنوان گروه اعتبارسنجی به‌طور تصادفی انتخاب گردید.

آماده کردن فاکتورهای مؤثر برای تهیه نقشه‌ی حساسیت‌پذیری هر خطر، ضروری است [۶]. در این پژوهش بر اساس مرور منابع و با بررسی‌های انجام‌شده عوامل مؤثر بر وقوع هر خطر در منطقه مورد مطالعه تهیه و به لایه‌های رستری با تفکیک مکانی ۳۰ متر در ۳۰ متر تبدیل شد. عوامل پیش‌بینی کننده که به‌طور مشترک برای هر دو خطر زمین‌لغزش و فرسایش خندقی مورد استفاده قرار گرفت عبارت‌اند از: مدل رقومی ارتفاعی، درصد شیب، جهت شیب کاربری اراضی، انحنای طولی دامنه، انحنای عرضی دامنه، بافت خاک، شاخص موقعیت توپوگرافی، شاخص رطوبت توپوگرافیکی، فاصله از شبکه آبراهه، تراکم زهکشی، بارش سالانه، شاخص قدرت جریان، موقعیت شیب نسبی، شاخص زبری سطح، سنگ‌شناسی، فاکتور طول شیب، فاصله از جاده و فاصله از گسل. این لایه‌ها



شکل ۲- لایه‌های پیش‌بینی کننده حساسیت پذیری: الف) مدل رقومی ارتفاعی؛ ب) درصد شیب؛ پ) جهت شیب؛ ت) کاربری اراضی (ث) انحنای طولی دامنه؛ ج) انحنای عرضی دامنه؛ چ) بافت خاک؛ ح) شاخص موقعیت توپوگرافی؛ خ) شاخص رطوبت توپوگرافی؛ د) فاصله از شبکه آبراهه؛ ذ) تراکم زهکشی؛ ر) بارش سالانه؛ ز) شاخص قدرت جریان؛ ژ) موقعیت شیب نسبی؛ س) شاخص زبری سطح؛ ش) سنگ‌شناسی؛ ص) فاکتور طول شیب؛ ض) فاصله از جاده؛ ط) فاصله از گسل

Fig 2. Sensitivity predictor layers. (a) Digital Elevation Model (DEM) (m), (b) Slope (%), (c) Aspect, (d) Land use Profile curvature, (e) Plan curvature, (f), (g) Soil texture, (h)

Topographic Position Index, (i) Topographic position Index, (j) Distance to stream (m), (k) Drainage density, (l) Rainfall (mm), (m) Stream Power Index, (n) Relative slope position, (o) Terrain ruggedness index, (p) lithology, (q) LS factor, (r) Distance to roads (m), (s) Distance to fault.

ترکیبی جهت تخمین صحیح استعداد وقوع خطر (متغیر وابسته) مورد استفاده قرار می‌گیرد [۶]. تعیین تعداد بهینه درختان، ابتدا یک تعداد درخت اولیه جهت تولید گراف تغییرات میانگین مربعات خطا (MSE) در مقابل تعداد خاص درختان نمونه آموزشی و ارزیابی است، مورد استفاده قرار می‌گیرد. یکی از پارامترهای اصلی در اجرای مدل جنگل تصادفی، پیش‌بینی کننده K (متغیر مستقل) در هر گره جهت پیش‌بینی مقدار وابسته (پاسخ) است. ساده‌ترین راه جهت تعیین مقدار K ، محاسبه جذر تعداد کل متغیرهای مستقل است [۱۸]. به منظور به کارگیری این مدل در مدل‌سازی خطرات طبیعی از نرم‌افزار R و پکیج Random Forest استفاده می‌شود.

مدل ماشین بردار پشتیبان (SVM)

مدل ماشین بردار پشتیبان یکی از مدل‌هایی ماشینی یادگیری نظارت شده است که جهت طبقه‌بندی تفکیک گروه‌ها به کار می‌رود [۲۶]. به عبارت دیگر، پس از مشخص شدن عوامل محیطی اثرگذار (متغیرهای مستقل) و وقوع خطرات (متغیر وابسته)، مدل ماشین بردار پشتیبان به تجزیه و تحلیل روابط میان آن‌ها پرداخته و داده‌ها را به گروه‌های متمایزی تقسیم می‌کند. ماشین بردار پشتیبان، الگوریتمی پیشرفته است که بر اساس شناسایی شرایط مختلف و تعیین توابع مختلف نا پارامتری، حداکثر تفکیک را بین داده‌های محیطی ایجاد می‌کند؛ به گونه‌ای که بتوان روابط مختلف میان عوامل محیطی و وقوع خطرات ایجاد نمود. اگر نقاط آموزشی را به صورت $[x_i, y_i]$ و بردار ورودی $X_i \in \{-1, 1\}$ تعریف کنیم و $i = 1, \dots, l$ آنگاه در حالتی که داده‌ها به صورت خطی قابل تفکیک هستند، قواعد تصمیم‌گیری که تعریف می‌شود و توسط یک صفحه بهینه که طبقات تصمیم‌گیری باینری را تفکیک می‌کند، به صورت معادله زیر خواهد بود:

$$Y = \text{Sign}(\sum_{i=1}^N y_i a_i (X, X_i) + b) \quad (1)$$

که در آن، Y نتیجه معادله، y_i ارزش طبقه‌ی نمونه‌ی آموزشی است. همچنین X و X_i نیز به ترتیب نشان‌دهنده بردار داده‌های ورودی و بردارهای پشتیبان هستند. در این معادله پارامترهای α و b تعیین‌کننده ابر صفحه هستند.

در شکل (۲) ارائه شده است. به منظور شناخت خصوصیات زمین‌شناسی منطقه مورد مطالعه و تهیه نقشه سنگ‌شناسی از چندین نقشه زمین‌شناسی با مقیاس ۱:۱۰۰۰۰۰ استفاده شد. نقشه کاربری اراضی منطقه مورد مطالعه از اداره آبخیزداری استان گلستان تهیه شد که اساس تهیه این نقشه استفاده از تصاویر ماهواره‌ای لندست بوده است. نقشه مزبور پس از انجام مطالعات صحرایی و تطبیق آن با واقعیات کاربری اراضی منطقه، تصاویر گوگل ارث و تصاویر ماهواره‌ای اصلاح گردید. مدل رقومی ارتفاعی منطقه مورد مطالعه از اداره منابع طبیعی و آبخیزداری استان گلستان که با استفاده از نقشه‌های توپوگرافی در مقیاس ۱:۲۵۰۰۰ به دست آمده بود، تهیه شد. همچنین نقشه بافت خاک از اداره منابع طبیعی استان گلستان تهیه شد.

با توجه به این که هم‌خطی یکی از دلایل افزایش خطا و کاهش کارایی مدل است و ممکن است به پیش‌بینی‌های خارج از دامنه مورد انتظار منجر شود، از این رو پیش از فرایند مدل‌سازی این موضوع می‌بایست بررسی شود. بعد از آماده‌سازی لایه‌ها، برای آنالیز این داده‌ها و بررسی هم‌خطی آن‌ها از نرم‌افزار SPSS استفاده شد. شرط هم‌خطی بودن لایه‌ها این است که عامل تورم واریانس^۱ (VIF) بیشتر از ۵ و آستانه تحمل^۲ کمتر از ۰/۱ باشد [۱۱]. با برقرار بودن این شرط لایه‌های داری همبستگی حذف و دوباره مدل بودن آن لایه اجرا می‌شود.

مدل‌سازی خطرات طبیعی با استفاده از مدل‌های یادگیری

ماشین

مدل جنگل تصادفی^۳ (RF)

مدل جنگل تصادفی یکی از روش‌های مدل‌سازی درخت تصمیم است [۱۸] به گونه‌ای که از تلفیق نتایج به دست آمده از درخت‌های مختلف، نتیجه یا برآورد نهایی را راجع به ارتباط میان خطرات طبیعی و عوامل محیطی ارائه می‌دهد. در مسائل مربوط به رگرسیون، جنگل تصادفی تعداد اختیاری از درختان ناآزموده (زیرمجموعه‌ای از متغیرهای مستقل) است که نتایج واکنش این درختان به صورت

1. Variance Inflation Factor (VIF)
2. Tolerance
3. Random forest

2. Support vector machine

اگر به صورت خطی قابل تفکیک نباشند، معادله فوق به صورت معادله شماره ۲ تغییر می‌یابد.

$$Y = \text{Sign}(\sum_{i=1}^N y_i a_i K(X, X_i) + b) \quad (2)$$

که تابع $K(X, X_i)$ تابع کرنلی است که برای ایجاد ماشین‌هایی با انواع مختلفی از سطوح تصمیم‌گیری غیرخطی در فضای داده‌ها، ضرب‌های داخلی تولید می‌کند [۲۵]. برای اجرای مدل ماشین بردار پشتیبان، تمام لایه‌های رستری ساخته شده وارد نرم‌افزار R می‌شود و مدل ماشین بردار پشتیبان اجرا می‌شود.

شیوه اعتبارسنجی نتایج حاصل از مدل‌سازی خطرات زمین‌لغزش و فرسایش خندقی

در این مرحله، نتایج حاصل از مدل‌سازی خطرات زمین‌لغزش و فرسایش خندقی، بر اساس وقوع خطرات گروه اعتبارسنجی (۳۰ درصد وقایع گذشته) و در قالب روش منحنی مشخصه عامل گیرنده (ROC) برای مدل‌های مورد استفاده در این پژوهش اعتبارسنجی می‌گردد [۱۲]. هرگاه مقدار مساحت زیر منحنی بیش‌تر از ۷۰ درصد باشد، دقت مدل قابل قبول خواهد بود [۵۰]. به‌طور کلی، تقسیم ۰/۹-۰/۱ عالی، ۰/۸-۰/۹ بسیار خوب، ۰/۷-۰/۸ خوب، ۰/۶-۰/۷ متوسط و ۰/۶-۰/۵ ضعیف، را برای همبستگی کیفی و کمی سطح زیر منحنی و ارزیابی تخمین ارائه کرده‌اند [۲۷]. در این مرحله کارایی مدل‌های به‌کاررفته (جنگل تصادفی و ماشین بردار پشتیبان) در زمینه تعیین مناطق مساعد وقوع این خطرات با یکدیگر مقایسه می‌شود. اگرچه در مدل‌سازی‌های خطرات طبیعی بایستی از معیارهای ارزیابی مختلفی برای اعتبارسنجی دقت مدل‌ها استفاده کرد، اما با توجه با پیشینه تحقیق در این زمینه، روش منحنی مشخصه عامل گیرنده به‌عنوان جامع‌ترین روش در نظر گرفته می‌شود [۲۶]. منحنی ROC بر اساس معیارهای حساسیت و تشخیص به دست می‌آید.

مقایسه نتایج مدل‌ها برای تعیین بهترین مدل برای هر خطر و تهیه نقشه واحد دو خطر

در این مرحله بعد از تعیین کارایی و دقت نقشه هر خطر با استفاده از مدل‌های مختلف، بهترین مدل برای تعیین نقشه هر خطر انتخاب گردید، تا مشخص گردد در منطقه مورد مطالعه کدام مدل برای تهیه نقشه خطر زمین‌لغزش و یا هر فرسایش خندقی به‌صورت منفرد مناسب است، سپس روش‌های برتر هر خطر باهم ترکیب شد. با ادغام نقشه حساسیت مربوط به خطرات زمین‌لغزش و فرسایش خندقی، نقشه حساسیت دو خطر با استفاده از ابزار ArcGIS به دست آمد. ابتدا نقشه نهایی هر مدل برای هر خطر با استفاده از روش شکستن طبیعی^۲ طبقه‌بندی شد. برای استفاده از قانون بولین (صفر و یک) نقشه به دو کلاس (بدون خطر و دارای خطر) تقسیم

1. Receiver operating characteristic
2. Natural break

می‌شود [۱۳]. در نهایت با ترکیب این دو نقشه و اجرای دستور Combine این دو نقشه به یک نقشه واحد با چهار کلاس تبدیل شد. اگر در مورد متغیرهای محیطی منطقه اطلاعات کافی موجود باشد، این نقشه دارای دقت بالایی خواهد بود. همچنین این نقشه می‌تواند مناطق حساس را به‌صورت واقعی‌تری نسبت به ارزیابی حساسیت به‌صورت منفرد نشان دهد. زیرا این نقشه به‌صورت جامع و همه‌جانبه بوده و نشان‌دهنده این است که در کدام قسمت از حوزه مورد مطالعه فرسایش و در چه قسمت‌هایی زمین‌لغزش اتفاق می‌افتد.

نتایج

نتایج بررسی هم‌خطی^۳ بین عوامل مؤثر بر خطرات (زمین‌لغزش و فرسایش خندقی) در منطقه مورد مطالعه با برقرار بودن شرط هم‌خطی لایه‌های داری همبستگی حذف و دوباره مدل بودن آن لایه اجرا می‌شود. با انجام این آزمون برای هر دو خطر زمین‌لغزش و فرسایش خندقی لایه شاخص زبری زمین (TRI) و همچنین فاکتور LS برای زمین‌لغزش حذف شد.

نقشه پیش‌بینی مناطق مستعد خطر با استفاده از مدل جنگل تصادفی

نقشه استعداد وقوع برای خطر زمین‌لغزش و فرسایش خندقی با استفاده از مدل جنگل تصادفی (RF) در شکل (۳) نشان داده شده است. همان‌گونه که مشاهده می‌گردد، مناطق شیب‌دار و مرتفع دارای استعداد وقوع زیاد و کم به ترتیب برای وقوع زمین‌لغزش و فرسایش خندقی هستند. فراوانی نسبی کلاس‌های کم، متوسط، زیاد و خیلی زیاد نقشه‌های مزبور در شکل (۴) محاسبه شده است. جدول (۱) ویژگی‌های آماری مقادیر احتمال به‌دست آمده از مدل جنگل تصادفی را برای هر خطر نشان می‌دهد.

جدول ۱- ویژگی‌های آماری مقادیر احتمال به‌دست آمده از مدل جنگل تصادفی

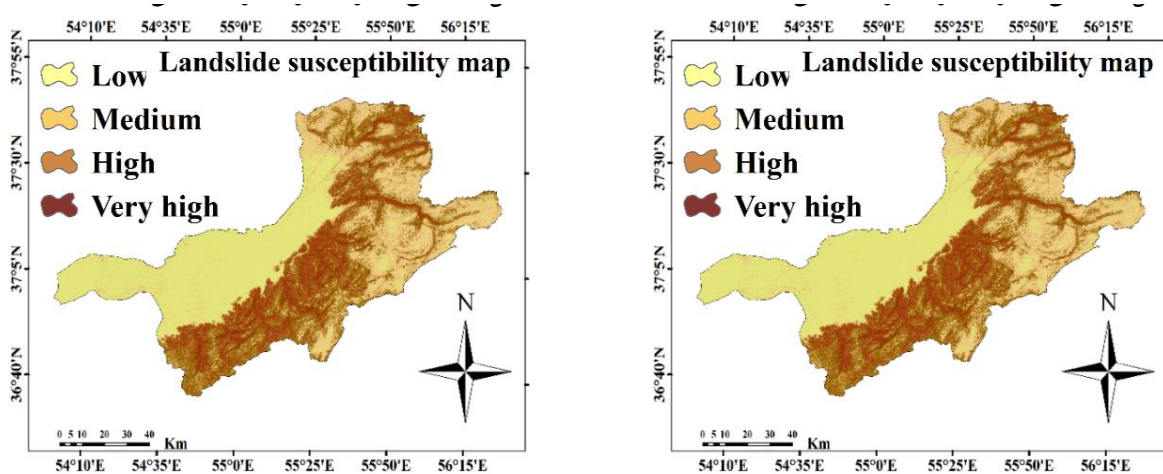
Table 1. Statistical Characteristics of Probability Values Obtained from Random Forest Model

Probabilistic prediction values		
SD	Mean	Hazards (خطرات)
0.229	0.313	Landslide (زمین‌لغزش)
0.207	0.137	Gully erosion (فرسایش خندقی)

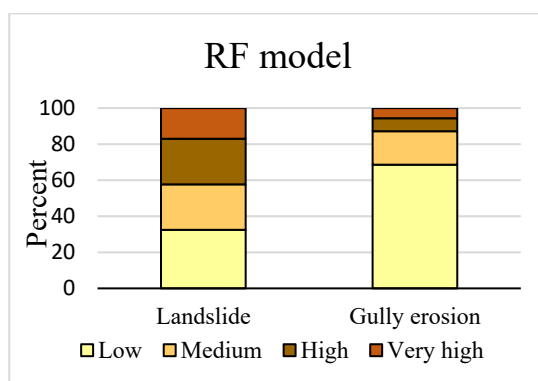
نقشه پیش‌بینی مناطق مستعد خطر با استفاده از مدل ماشین بردار پشتیبان

شکل (۵) به ترتیب نقشه‌های حساسیت وقوع زمین‌لغزش و

3. Multi-collinearity test



شکل ۳ (الف) - نقشه حساسیت وقوع زمین لغزش و وقوع فرسایش خندقی بر اساس مدل جنگل تصادفی
 Fig 3. (a) Landslide sensitivity map (b) Gully erosion sensitivity map based on Random Forest model



شکل ۴ - درصد فراوانی نسبی کلاس‌های استعداد وقوع خطر زمین لغزش و فرسایش خندقی در مدل جنگل تصادفی
 Fig 4. Relative distributions of the landslide and gully susceptibility classes.

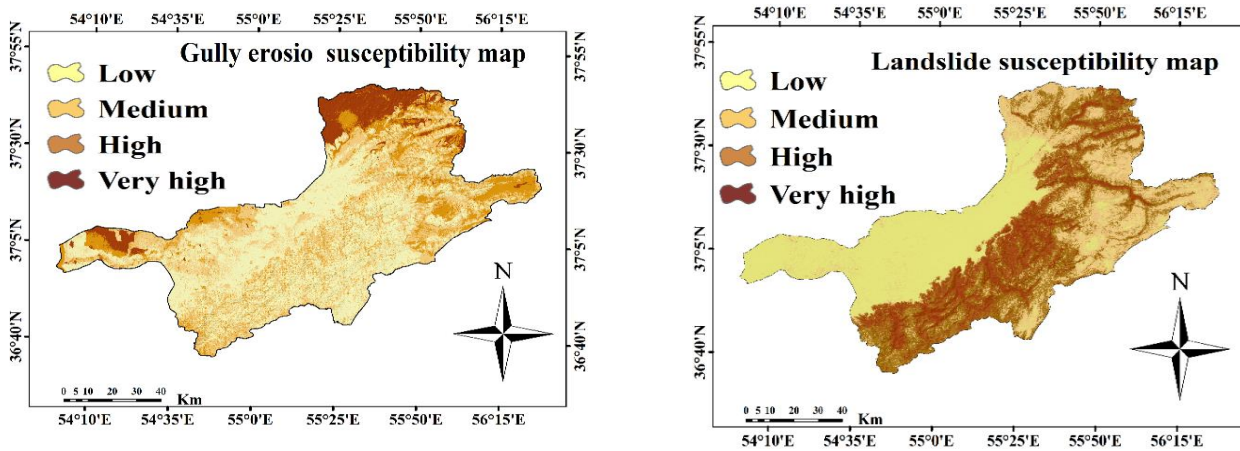
جدول ۲ - ویژگی‌های آماری مقادیر احتمال به دست آمده از مدل ماشین بردار پشتیبان

Table 2. Statistical Characteristics of Probability Values Obtained from Random Forest Model

Probabilistic prediction values		
SD	Mean	Hazards (خطرات)
0.327	0.321	Landslide (زمین لغزش)
0.269	0.174	Gully erosion (فرسایش خندقی)

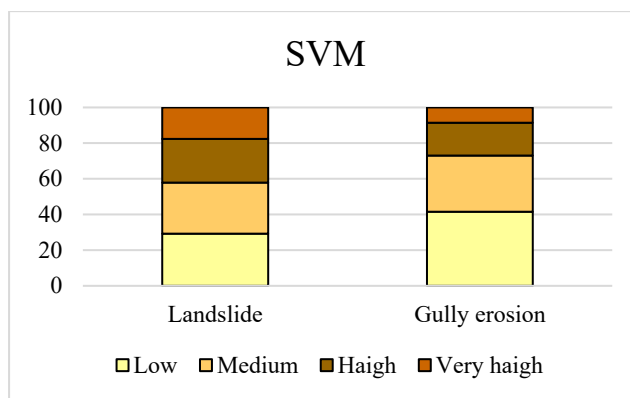
اعتبارسنجی نقشه‌های پیش‌بینی مناطق مستعد خطرات زمین لغزش و فرسایش خندقی
 نقشه پیش‌بینی مناطق مستعد هر خطر بر اساس منحنی مشخصه عامل گیرنده (ROC) اعتبارسنجی گردید شکل (۸). همچنین مساحت زیر این منحنی‌ها (AUC) برای هر یک از مدل‌ها محاسبه شد که نتایج آن در جدول ۳ به ترتیب برای خطرات زمین لغزش و فرسایش خندقی ارائه داده شده است که نتایج آن در جدول ۳ به ترتیب برای خطرات زمین لغزش و فرسایش خندقی ارائه داده شده است.

فرسایش خندقی با استفاده از مدل ماشین بردار پشتیبان را نشان می‌دهد. فراوانی نسبی کلاس‌های کم، متوسط، زیاد و خیلی زیاد نقشه‌های مزبور در شکل (۶) محاسبه شده است. جدول (۲) ویژگی‌های آماری مقادیر احتمال به دست آمده از این مدل را برای هر خطر نشان می‌دهد. شکل ۷، مقایسه فراوانی نسبی کلاس‌های استعداد وقوع خطر زمین لغزش و فرسایش خندقی را در هر دو مدل جنگل تصادفی و ماشین بردار پشتیبان نشان می‌دهد.



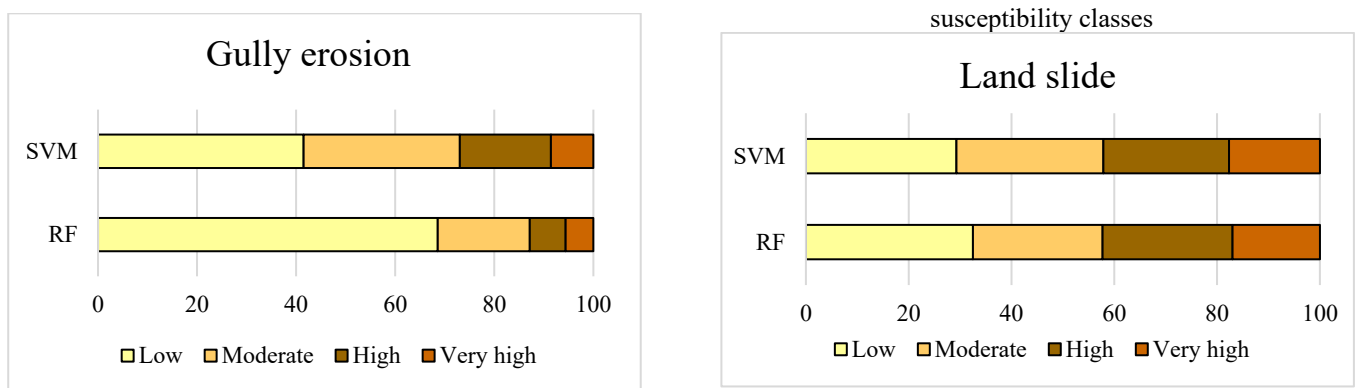
شکل ۵- (الف) نقشه حساسیت وقوع زمین لغزش (ب) نقشه حساسیت وقوع فرسایش خندقی بر اساس مدل ماشین بردار پشتیبان

Fig 5. (a) Landslide Sensitivity Map (b) Gully erosion Sensitivity map based on SVM model



شکل ۶- فراوانی نسبی کلاس‌های استعداد وقوع خطر زمین لغزش و فرسایش خندقی در مدل ماشین بردار پشتیبان

Fig 6. Relative distributions of the landslide and gully susceptibility classes



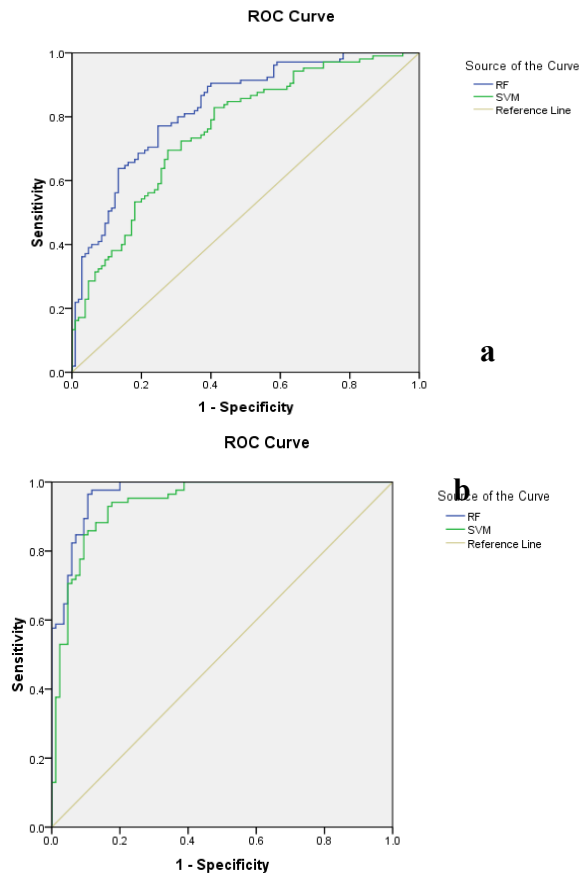
شکل ۷- مقایسه فراوانی نسبی کلاس‌های استعداد وقوع الف: خطر زمین لغزش، ب: فرسایش خندقی در مدل‌های مختلف

Fig 7. Comparison of the relative distributions of a: Landslide, b: Gully erosion in different models.

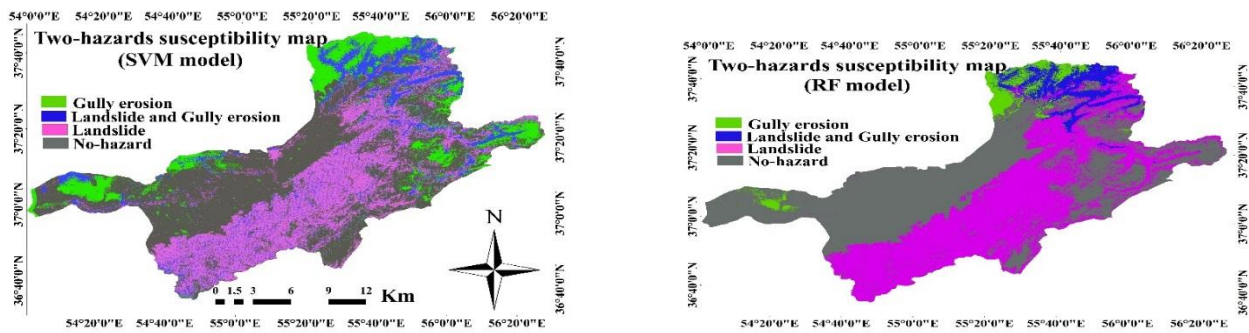
جدول ۳- نتایج اعتبارسنجی مدل‌های مختلف با استفاده از روش منحنی ROC برای خطر زمین لغزش و فرسایش خندقی

Table 3. Validation Results of Different Models Using ROC Curve Method for Landslide and Gully Erosion

	(SVM)	(RF)	AUC
	76.0	82.9	Landslide
	93.9	96.9	Gully erosion



شکل ۸- منحنی مشخصه عامل گیرنده (ROC) مدل‌های مختلف برای الف: خطر زمین لغزش، ب: فرسایش خندقی
 Fig 8. ROC curve of different models for a: landslide, b: Gully erosion



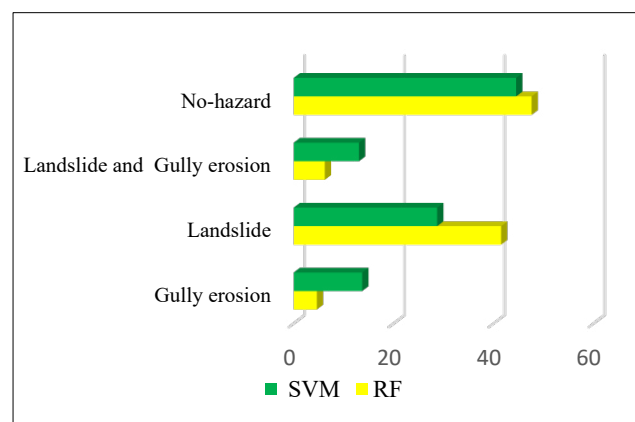
شکل ۹- نقشه حساسیت دو-خطر بر اساس الف: مدل جنگل تصادفی و ب: ماشین بردار پشتیبان
 Fig 9. Two- hazards Sensitivity map of a: Random forest model, b: Support vector machine

بالتر به دست آمد، لذا نقشه‌ی حساسیت دو خطر بر اساس این مدل برای حوزه گرگانرود مناسب‌تر است. شکل ۱۰ درصد خطرات را بر اساس نقشه‌ی دوخطر هر دو مدل در حوزه آبخیز گرگانرود نشان می‌دهد.

تهیه نقشه دو خطر

بعد از تعیین کارایی و دقت نقشه هر خطر با استفاده از مدل‌های جنگل تصادفی و ماشین بردار پشتیبان، نقشه‌ی حساسیت دو خطر بر اساس ArcGIS به دست آمد شکل (۹). از آنجایی که در منطقه مورد مطالعه دقت مدل جنگل تصادفی بر اساس جدول ۵ برای تهیه نقشه خطر زمین لغزش و یا هر فرسایش خندقی به صورت منفرد

و بسیار زیاد طبقه‌بندی شد [۱۵]. برای ارزیابی مدل، نسبت نمونه ۳۰/۷۰ درصد در نظر گرفته شد و مدل‌سازی برای هر دو خطر انجام شد. شکل ۸ (الف و ب) منحنی ROC مربوط به AUC محاسبه شده برای هر خطر را نشان می‌دهد. بر اساس نتایج به دست آمده در جدول (۳)، مدل جنگل تصادفی برای خطر زمین‌لغزش و فرسایش خندقی نسبت به مدل ماشین بردار پشتیبان دارای دقت بالاتری بود. از طرفی بر اساس تقسیم‌بندی [۲۷] می‌توان بیان نمود که مدل‌های جنگل تصادفی و ماشین بردار پشتیبان، به ترتیب دارای کارایی بسیار خوب و خوب در پیش‌بینی مناطق مستعد خطر زمین‌لغزش هستند. همچنین عملکرد این مدل‌ها برای پیش‌بینی مناطق مستعد فرسایش خندقی، عالی به دست آمد. در مطالعه‌ای که توسط گاین و همکاران [۴] در مناطقی از هند با استفاده از الگوریتم‌های مختلف یادگیری ماشینی انجام شد، مدل جنگل تصادفی با دقت ۹۶/۲ درصد بر اساس معیار AUC دارای بهترین عملکرد و بالاترین دقت پیش‌بینی بود، که نتایج حاصل از این پژوهش را تأیید می‌کند. از طرفی عملکرد عالی و بسیار خوب مدل‌های ماشین بردار پشتیبان و جنگل تصادفی برای پیش‌بینی زمین‌لغزش و فرسایش خندقی در مطالعات مختلفی [۱۰، ۲۴] تأیید شده است. از آنجایی که نقشه ارزیابی حساسیت چند خطره ابزاری ارزشمند برای برنامه‌ریزان و سیاست‌گذاران فراهم می‌کند [۱۳]. در این پژوهش با توجه به دقت هر مدل در تهیه نقشه منفرد حساسیت خطر، با ادغام نقشه حساسیت مربوط به خطرات یک نقشه‌ی واحد و جامع که نشان‌دهنده‌ی مناطق مستعد هر دو خطر باشد بر اساس هردو مدل با استفاده از ابزار ArcGIS به دست آمد (شکل ۹). این نقشه‌ی دارای چهار کلاس یعنی بدون خطر و کم‌خطر، زمین‌لغزش، فرسایش خندقی، زمین‌لغزش و فرسایش خندقی است. از آنجایی که دقت مدل جنگل تصادفی برای هر دو خطر بیشتر از مدل ماشین بردار پشتیبان به دست آمد بنابراین نقشه نهایی حاصل از ترکیب دو نقشه‌ی منفرد نیز دارای دقت بالاتری بوده و می‌تواند در برنامه‌های آمایش سرزمین برای مدیریت منطقه مورد استفاده قرار بگیرد. بر اساس این نقشه (شکل ۹-الف) بالاترین درصد خطر در این منطقه زمین‌لغزش است (۶۱/۵۱٪). حدود ۵ درصد مستعد فرسایش خندقی، ۶/۲۳ درصد در طبقه زمین‌لغزش و فرسایش خندقی و ۴۷/۶۲ درصد از منطقه در کلاس بی‌خطر و کم‌خطر قرار گرفت. با توجه به وسعت زیاد منطقه مورد مطالعه همین درصد مشترک برای هر دو خطر قابل چشم‌پوشی نبوده و باید در هنگام مطالعات آمایش سرزمین مورد توجه مدیران و برنامه‌ریزان قرار بگیرد. در مطالعه پورقاسمی و همکاران مختلفی [۱۳] که از همین روش برای تهیه نقشه چند خطره در استان لرستان استفاده شد، بالاترین کلاس خطر زمین‌لغزش و سیل بود که حدود ۳۳/۷۰ درصد از منطقه را به خود اختصاص داد. ایشان در پژوهش خود به این نتیجه رسیدند که مناطقی که در طبقه بیشتر از یک خطر قرار گرفتند برای زندگی پایدار بشر آسیب‌پذیرتر تلقی می‌شوند. بهاروند و همکاران [۱]، در تحقیقی به پهنه‌بندی مخاطرات زیست‌محیطی زمین‌لغزش، سیل،



شکل ۱۰- درصد خطرات چندگانه حوزه آبخیز گرگانرود
Fig 10. Percentage of multiple hazards in Gorganrood watershed

بحث و نتیجه‌گیری

زمین‌لغزش و فرسایش خندقی از جمله پدیده‌های مخرب و مخاطرات طبیعی هستند. کوهستانی بودن منطقه، وجود واحدهای لیتولوژی مستعد لغزش، اختلاف فصلی و شبانه‌روزی دما، باران‌های شدید، دخالت بی‌رویه انسان در محیط طبیعی مهم‌ترین عوامل وقوع این خطرات در حوزه آبخیز گرگانرود است. شناخت مکان‌هایی که پتانسیل بالایی به وقوع این خطرات را دارند، به مدیران و برنامه‌ریزان در برنامه‌های آمایش سرزمین و توسعه اراضی کمک شایانی می‌نماید. در این مطالعه، علاوه بر تهیه نقشه‌های پیش‌بینی مکانی خطرات زمین‌لغزش و فرسایش خندقی حوزه آبخیز گرگانرود با استفاده از دو الگوریتم شناخته‌شده و توانمند یادگیری ماشین به صورت منفرد، تهیه نقشه‌های دو خطره با ترکیب نقشه‌های حساسیت خروجی مدل‌ها از اهداف دیگر این پژوهش است. یالسن [۲۷] اظهار داشته‌اند که متغیرهای محیطی بسیاری هستند که به آسانی قابل تشخیص و جمع‌آوری نیستند و یا به عبارت دیگر قابلیت تبدیل به نقشه را ندارند. به این دلیل مدل‌های توسعه‌یافته از نظر مفهومی (Conceptual) ممکن است کامل نبوده و نتایج به دست آمده از مدل‌سازی با قطعیت مواجه نشوند. با این حال، بر اساس قابلیت دسترسی به داده‌ها در مناطق مختلف پژوهشگران اقدام به تهیه و جمع‌آوری عوامل در ارتباط با وقوع خطرات طبیعی نموده‌اند. با انجام آزمون همبستگی برای هر دو خطر زمین‌لغزش و فرسایش خندقی، لایه شاخص زبری زمین و همچنین فاکتور طول شیب برای زمین‌لغزش حذف شد. سپس با استفاده از موقعیت مکانی ۲۸۳ نقطه فرسایش خندقی و ۳۵۱ نقطه زمین‌لغزش در منطقه، نقشه‌های پیش‌بینی مناطق مستعد این خطرات به دست آمد. شکل ۳ و ۵ نقشه‌ی حساسیت به این خطرات را بر اساس دو مدل جنگل تصادفی و ماشین بردار پشتیبان نشان می‌دهد. نقشه‌ی منفرد پیش‌بینی شده با استفاده از فاصله شکستن طبیعی در چهار پهنه‌ی کم، متوسط، زیاد

M. 2018. Spatial modeling of flood inundation using GIS and data mining tools (A case study: the Gorganrood Watershed, Iran). In Третьи виноградовские чтения. Грани гидрологии (pp. 812-817).

8. Lee, S. Choi, J. 2004. Landslide susceptibility mapping in the Damrei Romel area, Cambodia using frequency ratio and logistic regression models. *Environmental Geology*. 50(6): 847-855.

9. Maraai, A. and Arayesh, S. 2009. Landslide Hazard Zoning in Tea Gardens (South of Lahijan and Langrud), M.Sc. Thesis, Faculty of Earth Sciences, Shahid Beheshti University. 150 p. (In Persian).

10. Mohamadi, M. and Pourqasemi, H.R. 2017. Prioritization of factors affecting the occurrence of landslides and preparation of its sensitivity map using a new random forest algorithm (Case study: part of Golestan). *Watershed Management Research Journal*, 8 (15): 161-170. (In Persian)

11. Ozdemir, A. 2011. Using a binary logistic regression method and GIS for evaluating and mapping the groundwater spring potential in the Sultan Mountains (Aksehir, Turkey). *J Hydrol*, 405:123-136.

12. Pourghasemi, H.R. Mohammady, M. and Pradhan, B. 2012. Landslide susceptibility mapping using index of entropy and conditional probability models in GS: Safarood Basin, Iran. *Catena*, 97: 71-84

13. Pourghasemi, H.R. Gayen, A. Panahi, M. Rezai, F. and Blaschke, T. 2019. Multi-hazard probability assessment and mapping in Iran. *Science of the total environment*, 692: 556-571.

14. Rahmati, O. Haghizadeh, A. Pourghasemi, H.R. and Noormohamadi, F. 2016. Gully erosion susceptibility mapping: the role of GIS-based bivariate statistical models and their comparison. *Natural Hazards*, 82(2): 1231-1258.

15. Rahmati, O. Tahmasebipour, N. Haghizadeh, A. Pourghasemi H.R. and Feizizadeh, B. 2017. Evaluating the influence of geo-environmental factors on gully erosion in a semi-arid region of Iran: An integrated framework. *Science of The Total Environment*, 579: 913-927

16. Rajabi, M. Hejazi, MA. Rostaei, sh. and Alii, N. 2018. Zoning of vulnerability of natural and geomorphological hazards of rural settlements in Saqez (Case study of floods and earthquakes). *Quantitative Geomorphological Research*, 7 (2): 183-195. (In Persian).

17. Shadfar, S. 2014 Application of Fuzzy Logic in the Study of Trench Erosion Using Geographic Information System (GIS) (Case Study: Taroud Basin). *Sepehr Geographical Information*, (92) 23: 42-35. (In Persian)

فرسایش و زمین لرزه در حوزه وارک واقع در استان لرستان پرداختند. نتایج بررسی نقشه مخاطرات زیست محیطی منطقه نشان داد که به ترتیب زمین لغزش و سیلاب از نقش بیشتری در ایجاد پهنه‌های با خطر بالا برخوردار می‌باشند. به‌طور کلی نقشه‌های مخاطرات مختلف در مرحله اولیه برنامه‌ریزی شهری بسیار مهم و ضروری می‌باشند. اگرچه، تجزیه و تحلیل با تعداد زیادی نقشه مخاطرات با وضوح مختلف و پوشش مکانی برای برنامه‌ریزان ممکن است دشوار باشد؛ بنابراین، یک نقشه احتمال چند خطر که ترکیبی از مخاطرات مختلف است، از مهندسین و برنامه‌ریزان برای شروع برنامه‌ریزی شهری پایدار پشتیبانی می‌کند، زیرا این نقشه، اطلاعات همگن در مورد مخاطرات مختلف محیطی برای منطقه ذکر شده ارائه می‌دهد [۲۱]. بنابراین این نقشه مبنای مناسبی برای ارزیابی احتمال چند خطر برای برنامه‌ریزی پایدار کاربری اراضی است.

منابع

1. Baharvand, S. Suri, S. and the Rahnnavard, J. 2017. Zoning of environmental hazards of landslides, earthquakes, floods and erosion using fuzzy hierarchical method (Case study: Wark Basin). *Remote Sensing and GIS in Natural Resources (Application of Remote Sensing and GIS in Natural Resources Sciences*, 3 (28): 103-89).In Persian(

2. Baqerzadeh, A.F. Asghari, K.A. and Dabiri, R. 2013. Investigation of landslide and flood hazards in Ahar city and hazard zoning. The First National Conference on Geotechnical Engineering of Iran, Faculty of Engineering, Mohaghegh Ardabili University, p. 8).In Persian(

3. Dube, F. Nhapi, I. Murwira, A. Gumindoga, W. Goldin, J. and Mashauri D.A. 2014. Potential of weight of evidence modeling for gully erosion hazard assessment in Mbire District- Zimbabwe. *Physics and Chemistry of the Earth*, 67-69: 145-152

4. Gayen, A. Haque, S.M. and Saha, S. 2020. Modeling of Gully Erosion Based on Random Forest Using GIS and R. In *Gully Erosion Studies from India and Surrounding Regions*, Springer, Cham. Pp: 35-44.

5. Khodadadi, F. Entezari, M. and Sasanpour, F. 2019. Analysis and zoning of geomorphological hazards (landslides and floods) of Alborz province using VIKOR-AHP and FR models. *Natural Geography Research*, 51 (1): 199-183. (In Persian)

6. 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. *Environmental Earth Sciences*, 67(1): 251-264.

7. Kornejady, A. Pourghasemi, H.R. Nalivan, O.A. and Ownegh,

23. Trigila, A. Iadanza, C. Esposito, C. Scarascia-Mugnozza, G. 2015. Comparison of Logistic Regression and Random Forests techniques for shallow landslide susceptibility assessment in Giampilieri (NE Sicily, Italy). *Geomorphology*. 249: 119-136.
24. Talebi, A. Goodarzi, S. And Pourqasemi H.R. 2017. Investigating the possibility of preparing a landslide hazard map using a random forest algorithm (study area: Sardarabad watershed, Lorestan province). *Journal of Natural Hazards*: 7, doi: 10.22111 / jneh.2017.3213. (In Persian)
25. Vapnik, V. 1995. The Nature of Statistical Learning Theory, Springer- Verlag, New York.
26. Wang, L. 2005. Support Vector Machines: Theory and Applications. Springer- Verlag, New York.
27. Yesilnacar, EK. 2005. The Application of Computational Intelligence to Landslid Susceptibility Mapping in Turkey (Ph.D Thesis) Department of Geomatics the University of Melbourne, p. 423.
18. Shataee, S. Weinaker, H. and Babanejad, M. 2011. Plot- level Forest Volume Estimation Using Airborne Laser Scanner and TM Data, Comparison of Boosting and Random Forest Regression Algorithms. *Procedia Environmental Sciences*. 7: 68-73.
19. Shit, P.K. Paira, R. Bhunia, G. and Maiti, R. 2015. Modeling of potential gully erosion hazard using geo-spatial technology at Garbheta block, West Bengal in India. *Modeling Earth Systems and Environment*, 1(1-2): 1-16.
20. Servati, M.R. Ghahrodi Tali, M. Golkarami, A. and Njafi, E. 2014. Geomorphological thresholds for gully erosion in Kchick watershed, NE Golestan Province. *Applied researches in geographical sciences*, 32: 231-249 (in Persian).
21. Skilodimou, H.D. Bathrellos, G.D. Chousianitis, K. Youssef, A.M. and Pradhan, B. 2019. Multi-hazard assessment modeling via multi-criteria analysis and GIS: a case study. *Environmental earth sciences*, 78(2): 47.
22. Solaimani, K. Mousavi, SZ. Kaviani, A. 2013. Landslide susceptibility mapping based on frequency ratio and logistic regression models. *Arabian Journal of Geosciences*. 6(7): 2557-2569.

Identification the Areas Prone to Gully Erosion and Landslides in the form of Two-Hazards Map Using Machine Learning Models in Gorganrood Watershed

N. Javidan¹, A. Kavian², S. Rajabi³, H R. Pourghasemi⁴ and Z. Jafarian⁵

Received: 26-10-2021 Accepted: 09-05-2022

Abstract

The aim of this study is to identify the areas prone to gully erosion and landslide as a two-hazard map in the form of a single map in Gorganrood watershed. In addition, the use of two machine learning models such as RF and SVM to establishing the spatial relationship between these hazards and the GEFs and any hazard susceptibility mapping separately. In addition, the validation of the hazards susceptibility maps was conducted based on the ROC curve method, and the best model was chosen with the highest predictive performances. Finally, by combining the susceptibility landslides and gully maps, the two-hazard probability maps were produced, which were a combination of different models and the best model. The results showed a RF model with (AUC = 82.9) for landslide and (AUC = 96.9) for gully erosion have higher accuracy compared to the SVM model with a value of (AUC = 0.76) for landslide and (AUC = 93.9) for gully erosion. Finally, a single and comprehensive map was obtained by combining of the each hazard susceptibility map for identification the areas prone to both hazards based on both models. The final two-hazard map can be used as a valuable tool for sustainable land use planning in multi-hazard prone areas.

Keywords: *Machine learning models, Natural hazards, Relative receiver operating characteristic curve, Spatial prediction, Two- hazard map.*

1. Ph.D. of Watershed Management, Faculty of Natural Resources, Sari Agricultural Sciences and Natural Resources University, Iran

2. Corresponding author and Professor of Watershed Management Department, Faculty of Natural Resources, Sari Agricultural Sciences and Natural Resources University. ataollah.kavian@gmail.com

3. M.Sc. of Watershed Management, Faculty of Natural Resources, Sari Agricultural Sciences and Natural Resources University, Iran

4. Professor of Natural Resources and Environment Engineering Department r, Faculty of Agriculture, Shiraz University, Shiraz, Iran

5. Assistant Professor of Range Management Department, Faculty of Natural Resources, Sari Agricultural Sciences and Natural Resources University (SANRU)