نشریه علمی

علوم و مهندسی آبخیزداری ایران
Iran-Watershed Management
Science & Engineering

سال هفدهم- شماره ٦٣- زمستان ١٤٠٢

Vol. 17, No. 63, Winter 2024

ارزیابی نقشه حساسیت زمین‌لغزش با استفاده از الگوریتم یادگیری ماشین و اولویت‌بندی عوامل مؤثر بر وقوع زمین‌لغزش

علی دسترنج*[1]، حمزه نور[2]، فرزانه وکیلی تجره[3]

تاریخ دریافت: ۱۴۰۱/۰۳/۱۷ تاریخ پذیرش: ۱۴۰۱/۰۹/۱۷

DOI: 10.22034/17.63.71

چکیده

هدف از مطالعه پیش رو، مدل‌سازی مکانی حساسیت وقوع زمین‌لغزش با استفاده از الگوریتم یادگیری ماشین جنگل تصادفی و اولویت‌بندی عوامل مؤثر بر وقوع زمین‌لغزش در حوزه آبخیز بار نیشابور، استان خراسان رضوی است. الگوریتم جنگل تصادفی مبتنی بر دسته‌ای از درخت‌های تصمیم است و در حال حاضر یکی از بهترین الگوریتم‌های یادگیری ماشین است. برای این منظور، لایه نقشه پراکنش زمین‌لغزش‌های منطقه شامل ۷۳ زمین‌لغزش تهیه و به دو دسته برای آموزش مدل (۷۰ درصد) و اعتبارسنجی مدل (۳۰ درصد) به‌صورت تصادفی تقسیم شدند. همچنین، ١٦ عامل مؤثر بر وقوع زمین‌لغزش در منطقه موردمطالعه با توجه به‌مرور منابع گسترده شناسایی و لایه‌های رقومی در سامانه اطلاعات جغرافیایی تهیه شد. به‌منظور ارزیابی قدرت پیش‌بینی مدل از مساحت زیر منحنی تشخیص عملکرد نسبی (ROC) برای دو مرحله آموزش و اعتبارسنجی مدل استفاده شد. نتایج ارزیابی مدل نشان داد که مدل جنگل تصادفی با مقادیر سطح زیر منحنی ۰/۹ دارای دقت عالی در مرحله آموزش و ۰/۸۹ دارای دقت خیلی خوب در مرحله اعتبارسنجی است. نتایج اولویت‌بندی عوامل مؤثر بر وقوع زمین‌لغزش در منطقه موردمطالعه نشان داد که عوامل طول شیب و شیب دارای بیشترین اهمیت هستند. بر اساس نتایج مدل جنگل تصادفی، ۲۳/۷ درصد منطقه موردمطالعه در پهنه حساسیت خیلی زیاد و زیاد واقع‌شده است.

کلیدواژه‌ها: *الگوریتم جنگل تصادفی، حوزه بار، ROC*

مقدمه

زمین‌لغزش از پر خسارت‌ترین حرکت‌های دامنه‌ای است که با دست‌کاری انسان در سامانه‌های طبیعی در دهه‌های اخیر شتاب افزاینده‌ای گرفته است [۹]. زمین‌لغزش درواقع یک پدیده طبیعی است که سالیانه در نقاط مختلف ایران و جهان رخ‌داده و خسارت‌های جانی، مالی و زیست‌محیطی فراوانی به بار آورده است [٤]. زمین‌لغزش نوعی حرکت دامنه‌ای است که مواد در امتداد یک سطح مشخص روی دامنه لغزنده به سمت پایین حرکت می‌کنند [٦]. به طور عمده در اثر نیروی ثقل، عوامل طبیعی مثل زلزله، بارش شدید و دخالت‌های انسانی مانند تخریب پوشش گیاهی و جاده‌سازی به وقوع می‌پیوندد [۲]. زمین‌لغزش به طور معمول شکل زمین را به‌طور ناگهانی برهم میزند و خسارت‌های فراوانی به مناطق مسکونی، جاده‌ها و زمین‌های کشاورزی وارد می‌کند [۹، ۲۰]. کارآمدی پروژه‌های عمرانی همچون احداث راه‌های اصلی و فرعی کوهستانی، بزرگراه‌ها، طرح‌های توسعه جنگل و منابع طبیعی، انتخاب محل احداث سدهای خاکی و بتونی و هرگونه مناطق مسکونی و توسعه معدنی در گروه مطالعات پایداری دامنه‌های طبیعی منطقه است [۷]. با توجه به تأثیرات نامطلوب پدیده زمین‌لغزش بر سیستم‌های اقتصادی، اجتماعی و طبیعی، شناخت نواحی مستعد وقوع زمین‌لغزش در کشور لازم و ضروری هست. با تهیه نقشه پهنه‌بندی حساسیت وقوع زمین‌لغزش می‌توان مناطق حساس به زمین‌لغزش را شناسایی و درباره برنامه‌های موردنظر تصمیم‌گیری کرد. به علت پیچیدگی سیستم زمین‌لغزش‌ها، دستورالعمل واحدی در بین محققین برای انتخاب عوامل مؤثر در تهیه نقشه پهنه‌بندی حساسیت زمین‌لغزش وجود ندارد و نیاز به داده‌های مختلف طبیعی و انسانی دارد [۳۵، ۳۰]. تهیه نقشه پهنه‌بندی مستلزم بررسی روابط بین شرایط حاکم بر حوضه و عوامل مؤثر بر وقوع زمین‌لغزش است [۲]. روش‌های مختلفی برای پهنه‌بندی حساسیت وقوع زمین‌لغزش ارائه‌شده ولی هیچ‌کدام قطعیت لازم را ندارند و برای کاربرد آن‌ها در مناطق مختلف نیاز به اصلاحات لازم دارند [۲۵]. بر پایه وضعیت منطقه موردمطالعه و عوامل مؤثر در رخداد زمین‌لغزش، دقت و

۱- *نویسنده مسئول: بخش تحقیقات حفاظت خاک و آبخیزداری، مرکز تحقیقات و آموزش کشاورزی و منابع‌طبیعی خراسان‌رضوی، سازمان تحقیقات، آموزش و ترویج کشاورزی، مشهد. Dastranj66@gmail.com
۲- بخش تحقیقات حفاظت خاک و آبخیزداری، مرکز تحقیقات و آموزش کشاورزی و منابع‌طبیعی خراسان‌رضوی، سازمان تحقیقات، آموزش و ترویج کشاورزی، مشهد.
۳- دانشگاه تهران، گروه احیاء مناطق خشک و کوهستانی، دانشکده منابع طبیعی، دانشجویی دکتری رشته علوم و مهندسی آبخیزداری

انعطاف هر یک از این روش‌ها متفاوت خواهد بود. این روش‌ها به‌طورکلی به دو دسته کمی و کیفی طبقه‌بندی می‌شوند. مدل‌های کیفی بر پایه قضاوت متخصص یا گروهی از متخصصان قرار دارد و برای تخمین خطر از ارزیابی متخصص استفاده می‌شود (مثل روش تحلیلی سلسله مراتبی و فاکتور اطمینان) در حالی که در روش‌های کمی بر پایه ریاضیاتی و به دور از قضاوت متخصصان است (مانند رگرسیون چند متغییره و لجستیک و منطق فازی) [۲٤، ۳۲]. در سال‌های اخیر تکنیک‌های داده‌کاوی مانند روش‌های یادگیری ماشین به دلیل دقت بالا مورداستفاده قرارگرفته‌اند. ازآنجایی‌که روش‌های یادگیری ماشین شامل: شبکه عصبی مصنوعی [۱۰، ۳۱]، رگرسیون لجستیک [۱، ۱۱]، ماشین‌بردار پشتیبان [۱۳] و جنگل تصادفی [۲٦، ۲۷]، دارای قابلیت پردازش قوی اطلاعات هستند به‌منظور تهیه نقشه حساسیت زمین‌لغزش مورداستفاده قرارگرفته‌اند. الگوریتم جنگل تصادفی یکی از این روش‌های یادگیری ماشین هست. الگوریتم جنگل تصادفی یکی از روش‌های مدل‌سازی درخت تصمیم است که از تلفیق نتیجه‌های به‌دست‌آمده از درخت‌های مختلف، ارتباط میان وقوع زمین‌لغزش و پارامترهای محیطی را برآورد می‌کند [۱۲].

عفیفی [۲] در مطالعه‌ای به بررسی خطر مکانی زمین‌لغزش با تأکید بر ژئومورفولوژیک با استفاده از مدل جنگل تصادفی پرداخت. نتایج ارزیابی دقت مدل با استفاده از ۳۰درصد نقاط لغزشی با کمک منحنی تشخیص عملکرد نسبی (ROC[1])، بیانگر دقت عالی مدل جنگل تصادفی با سطح زیر نمودار بیش از ۹۸ درصد بود. کرم و همکاران [۱٤]، ناپایداری دامنه‌های طبیعی را با استفاده از الگوریتم جنگل تصادفی موردبررسی قرار دادند. آن‌ها در این مطالعه پارامترهایی نظیر شیب، جهت شیب، ارتفاع، دامنه انحنا، زمین‌شناسی، بارش، کاربری اراضی، فاصله از گسل و فاصله از جاده را موردبررسی قراردادند. نتایج ارزیابی مدل با استفاده از منحنی عملکرد نسبی (ROC) با AUC[2]=0.93، بیانگر عملکرد عالی مدل است. همچنین نتایج اولویت‌بندی عوامل ورودی نشان‌دهنده اهمیت بیشتر عامل شیب و ارتفاع در پیش‌بینی نهایی است. کرنژادی و همکاران [۱۷] نیز به پیش‌بینی حساسیت زمین‌لغزش با استفاده از روش یادگیری ماشین جنگل تصادفی در حوزه آبخیز اوغان در استان گلستان پرداختند. برای این منظور از ۷۰درصد زمین‌لغزش‌های موجود برای مدل‌سازی و ۳۰ درصد باقیمانده به‌منظور اعتبارسنجی مدل استفاده شد. ارزیابی نتایج به‌دست آمده با در نظر گرفتن ۱۵ عامل مؤثر بر وقوع زمین‌لغزش با استفاده از منحنی ROC، با AUC=0/92 بیانگر توانایی خیلی خوب مدل در پیش‌بینی وقع زمین‌لغزش است. همچنین عوامل بارش و شاخص تفاضلی پوشش گیاهی (NDVI[3])، به‌عنوان مهم‌ترین عوامل مؤثر در وقوع زمین‌لغزش‌های منطقه موردمطالعه معرفی شدند.

ژین ژو و همکاران [۳٤] با استفاده از الگوریتم جنگل تصادفی اقدام به پیش‌بینی حساسیت زمین‌لغزش در کشور چین پرداختند. برای این منظور از ۲۲ عامل مؤثر بر وقوع زمین‌لغزش و نقشه زمین‌لغزش‌های موجود استفاده کردند. نتایج تأییدکننده کارایی بالای الگوریتم جنگل تصادفی در پیش‌بینی حساسیت زمین‌لغزش می‌باشد. سوگن و همکاران [۲٦] در مطالعه‌ای به مقایسه عملکرد مدل شبکه عصبی و الگوریتم جنگل تصادفی به‌منظور پیش‌بینی حساسیت زمین‌لغزش پرداختند. نتایج نشان داد که مدل جنگل تصادفی توانایی بالاتری در پیش‌بینی حساسیت زمین‌لغزش نسبت به مدل شبکه عصبی دارد. ژائو و همکاران [۳۳] با استفاده از مدل جنگل چرخشی (ROF[4]) و جنگل تصادفی (RF[5]) اقدام به پیش‌بینی حساسیت زمین‌لغزش در کشور چین کردند. نتایج نشان داد که هر دو مدل کارایی خوبی در پیش‌بینی حساسیت زمین‌لغزش دارند ولی با این‌وجود مدل الگوریتم جنگل تصادفی دقت بالاتری نسبت به مدل جنگل چرخشی، در پهنه‌بندی حساسیت زمین‌لغزش دارد. کنگ و همکاران [۱۸] در مطالعه‌ای حساسیت وقوع زمین‌لغزش در ژوآن پینگ در کشور چین را با استفاده سه روش یادگیری ماشین جنگل تصادفی (RF)، ماشین بردار پشتیبان (SVM[6]) و جنگل تصادفی بهینه‌سازی شده ازدحام ذرات (PSO-RF) را موردبررسی قراردادند. ارزیابی عملکرد سه مدل با استفاده از منحنی ROC نشان‌دهنده عملکرد خوب هر سه مدل است که نسبت به دو مدل دیگر مدل جنگل تصادفی بهینه‌سازی شده با AUC=0.934 بهترین عملکرد را داشت.

حوزه آبخیز بار در شمال غربی شهر نیشابور به دلیل استقرار روستاهای متعدد و شهر بار که عمده فعالیت‌های آن‌ها کشاورزی، باغداری و دامداری است، نقش مهمی در تولیدات باغی و کشاورزی و تولیدات دامی در استان خراسان رضوی دارد. همچنین یکی از قطب‌های گردشگری استان خراسان رضوی محسوب می‌شود. این حوزه به دلیل شرایط زمین‌شناسی، اقلیمی و فعالیت‌های انسانی دارای پتانسیل وقوع زمین‌لغزش است.

لذا هدف از این مطالعه پیش‌بینی مکانی حساسیت وقوع زمین‌لغزش با استفاده از لایه‌های اطلاعاتی پارامترهای مؤثر بر وقوع زمین‌لغزش و مدل داده‌کاوی جنگل تصادفی، ارزیابی مدل مذکور و ارزیابی عوامل مؤثر بر وقوع زمین‌لغزش در حوزه آبخیز بار نیشابور در استان خراسان رضوی است.

روش کار

منطقه موردمطالعه

حوزه آبخیز بار در محدوده طول جغرافیایی"۲۳ '۳۵ °۵۸ تا "٤ '۵۲ °۵۸ شرقی و عرض جغرافیایی "۹ '۲۸ °۳٦ تا "۷ '۲٦ °۳٦ شمالی قرار دارد و در دامنه‌های جنوبی بینالود واقع‌شده است.این

1. Receiver operating characteristic
2. Area Under Curve
3. Normalized Difference Vegetation Index (NDVI)
4. Rotation Forest
5. Random Forest
6. Support Vector Machin

حوزه یکی از زیرحوزه‌های حوزه آبریز رودخانه بار است که در شمال شهرستان نیشابور قرار دارد. این حوزه از شمال شرق و شرق و جنوب به ارتفاعات ستیغ ساز آهکی لار، از غرب به شیل‌های ژوراسیک و از شمال غرب به مارن‌های سازند دلیچای محدود می‌شود. شکل (۱) موقعیت منطقه مورد مطالعه را نشان می‌دهد. حوزه آبخیز بار بخشی از زیر زون ساختاری بینالود-آلاداغ است که در نتیجه تحولات زمین ساختی واقع در آن، انواع واحدهای سنگی و رسوبی دوران دوم و سوم زمین‌شناسی بر روی آن گسترش‌یافته است. شیب متوسط حوضه ۱۸/۹۸ درجه بوده و حداقل و حداکثر ارتفاع به ترتیب ۱۳۷۹ و ۳۱۹۳ متر از سطح دریا است. کاربری اغلب منطقه مورد مطالعه مرتع متوسط (۱۸٦٦٦ هکتار) است. بر اساس نقشه زمین‌شناسی ۱:۱۰۰۰۰۰ شیت نیشابور و اخلمد سازندهای دوران دوم (لار)، و چهارم (رسوبات کواترنری Q^{2t}، پادگانه‌های آبرفتی Q^{t1} و آبرفت‌های جوان Q^{al} در محل تلاقی آبراهه‌ها و مسیل‌ها با رودخانه اصلی گسترش‌یافته است.

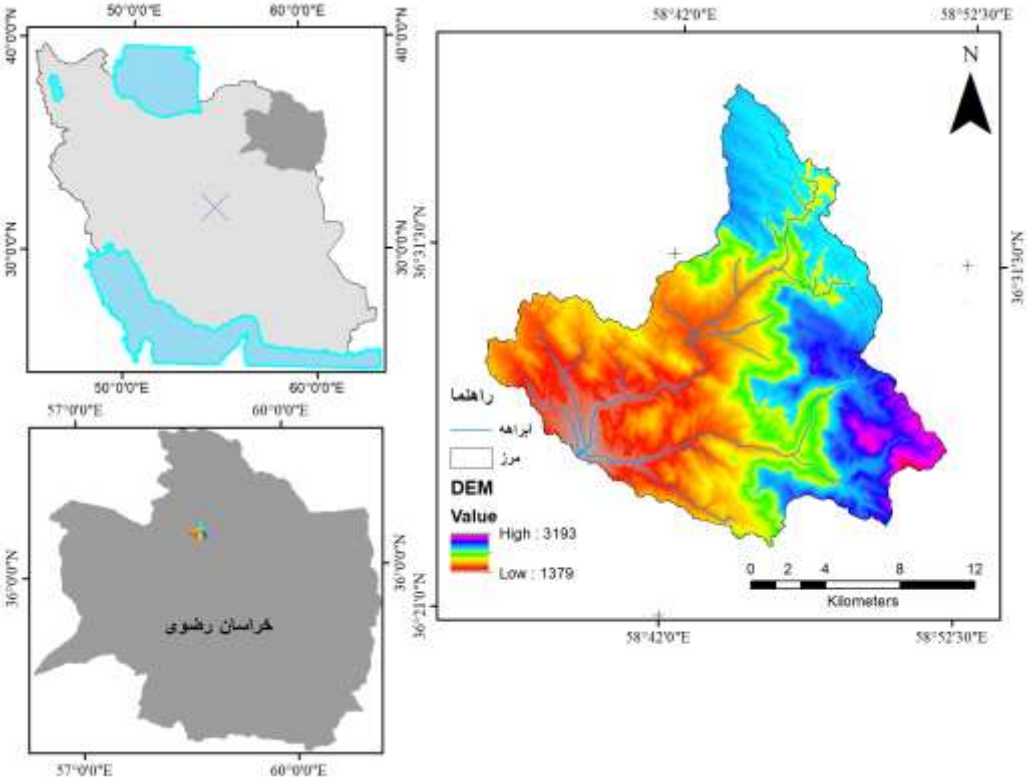
مواد و روش‌ها

نمودار جریانی مراحل انجام پژوهش در شکل ۲ بیان‌شده است:

داده‌های پژوهش

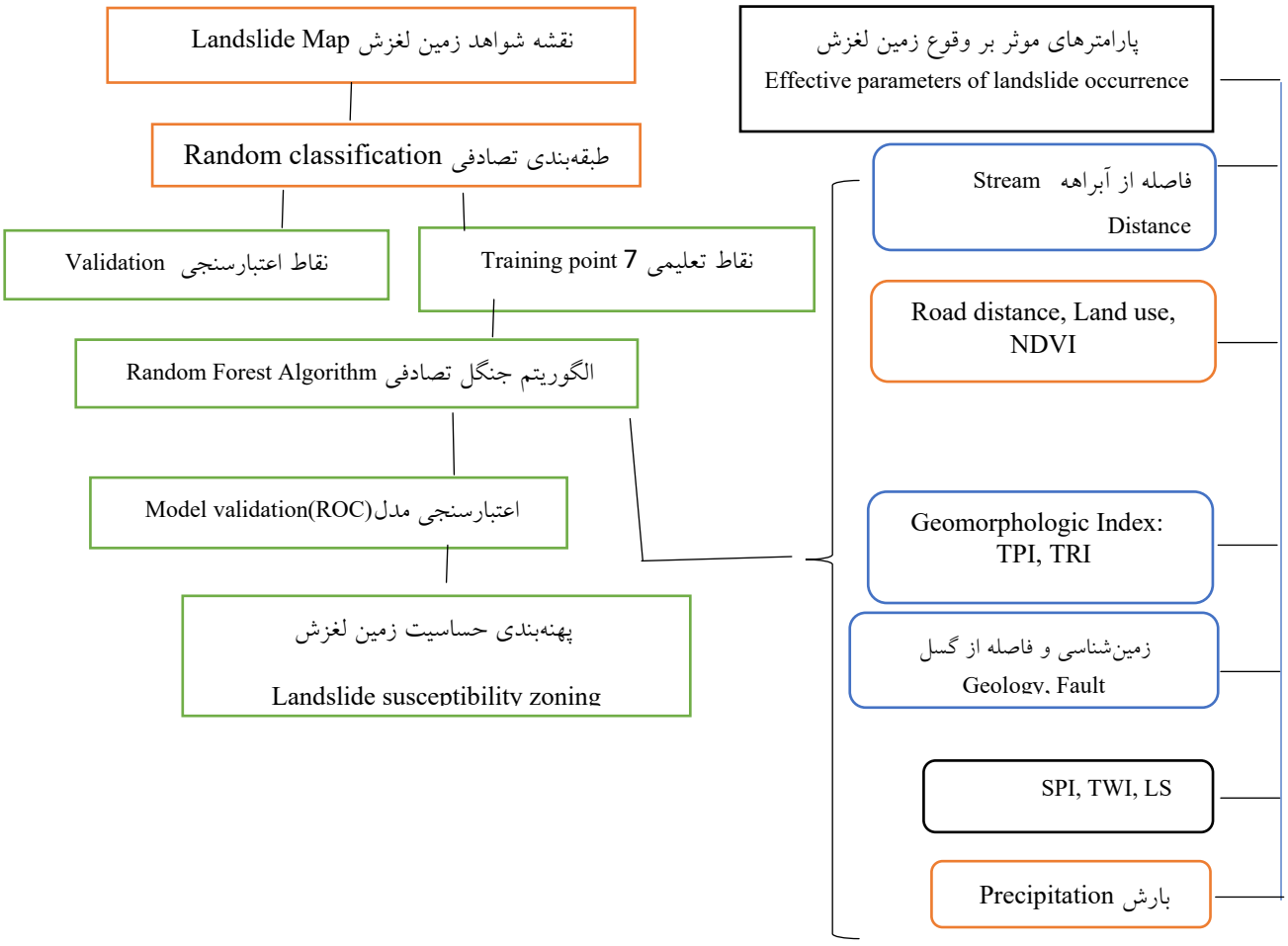
اولین مرحله مطالعه پهنه‌بندی خطر زمین‌لغزش شناخت و جمع‌آوری اطلاعات مربوط به عواملی است که در وقوع زمین‌لغزش‌ها نقش مؤثری دارند. گردآوری اطلاعات به‌صورت میدانی، کتابخانه‌ای و با مراجعه به اداره کل منابع طبیعی استان خراسان رضوی صورت گرفته است. از میان عوامل مؤثر در رخداد لغزش، آن دسته از عواملی که بر پایه پژوهش‌های انجام گرفته در مناطق دیگر و نیز نظر کارشناسی، تأثیر زیادی در رخداد لغزش داشته‌اند، انتخاب شدند. درنهایت مهم‌ترین عوامل مؤثر بر زمین‌لغزش‌های منطقه شامل شیب، جهت شیب، طبقات ارتفاعی، زمین‌شناسی، شبکه زهکشی (فاصله از رودخانه)، جاده (فاصله از جاده)، گسل (فاصله از گسل)، شاخص‌های توپوگرافیک (شاخص توان رودخانه (SPI)، شاخص رطوبت توپوگرافی (TWI) و شاخص طول شیب (LS))، شاخص‌های ژئومورفولوژیک (شاخص موقعیت توپوگرافی (TPI)، شاخص ناهمواری توپوگرافی (TRI) و شاخص انحنای سطح (Curvature Index)، کاربری اراضی، شاخص نرمال‌شده تفاوت پوشش گیاهی (NDVI) و خطوط هم بارش به‌عنوان عوامل مؤثر در رخداد زمین‌لغزش در منطقه موردمطالعه، شناسایی و مورد تجزیه‌وتحلیل قرار گرفتند (شکل ۳).

لایه‌های شیب، جهت شیب و طبقات ارتفاعی با استفاده از مدل رقومی ارتفاع (DEM) با قدرت تفکیک زمینی ۱۲/۵ متر و نرم‌افزار Arc GIS 10/8 تهیه شد. برای تهیه لایه‌های زمین‌شناسی و گسل‌های منطقه موردمطالعه بر اساس نقشه زمین‌شناسی ۱:۱۰۰۰۰۰ نیشابور و اخلمد با شماره شیت ۷۷٦۲ و ۷۷٦۳ (تهیه‌شده توسط سازمان زمین‌شناسی)، استفاده شد. نقشه بارش نیز با استفاده از داده‌های بارش ایستگاه‌های سینوپتیک و باران‌سنجی که داخل و نزدیک به



شکل ۱- موقعیت جغرافیایی حوزه آبخیز بار

Fig1. Location of the study area



شکل۲- نمودار جریانی روش پژوهش

Fig 2. Flowchart of research method

منطقه موردمطالعه واقع شدند، استخراج شد. نقشه کاربری منطقه موردمطالعه با استفاده از نقشه کاربری اراضی استان خراسان رضوی که از اداره کل منابع طبیعی خراسان رضوی تهیه شد، استخراج شد. از طریق بازدید میدانی و بخصوص از طریق تصاویر ماهواره‌ای Google Earth، اصلاح شد. شاخص NDVI با استفاده از تصاویر ماهواره لندست ۸ در محیط Google Earth engine، تهیه شد. از تابع فاصله اقلیدسی برای تهیه و آماده‌سازی فاصله از گسل، فاصله از جاده و فاصله از آبراهه در محیط نرم‌افزار Arc GIS 10.8، استفاده شد. نقشه جاده‌های منطقه مورد با استفاده از نقشه جاده‌های استان که از اداره منابع طبیعی استان خراسان رضوی تهیه شد، استخراج و از طریق بازدیدهای میدانی و تصاویر ماهواره‌ای Google Earth، اصلاح شد. شاخص توان آبراهه (SPI) از رابطه یک و با توجه به دستور raster calculator در محیط نرم‌افزار Arc GIS 10.8 محاسبه می‌شود [۳].

رابطه۱:

$$SPI=A_s \tan\beta$$

در این رابطه A_S: مساحت حوزه و Tanβ: مقدار شیب منطقه در هر پیکسل است.

شاخص رطوبت توپوگرافی (TWI) نیز از رابطه ۲ در محیط نرم‌افزار SAGA GIS، تهیه شد.

رابطه۲:

$$TWI= Ln\ \alpha/\tan\beta$$

در این رابطه α: زهکش شیب بالاست در واحد طول و β: شیب منطقه‌ای می‌باشد.

همچنین شاخص موقعیت توپوگرافی (TPI)طبق رابطه ۳ به دست می‌آید:

رابطه۳:

$$TPI= Z_0-\sum_{n-1} Z_{n/n}$$

که Z_0= ارتفاع نقطه مدل تحت ارزیابی، Z_n= ارتفاع از شبکه و n= تعداد کل نقاط اطراف در نظر گرفته‌شده در ارزیابی.

شاخص ناهمواری توپوگرافی(TRI) از رابطه ٤ محاسبه می‌شود:

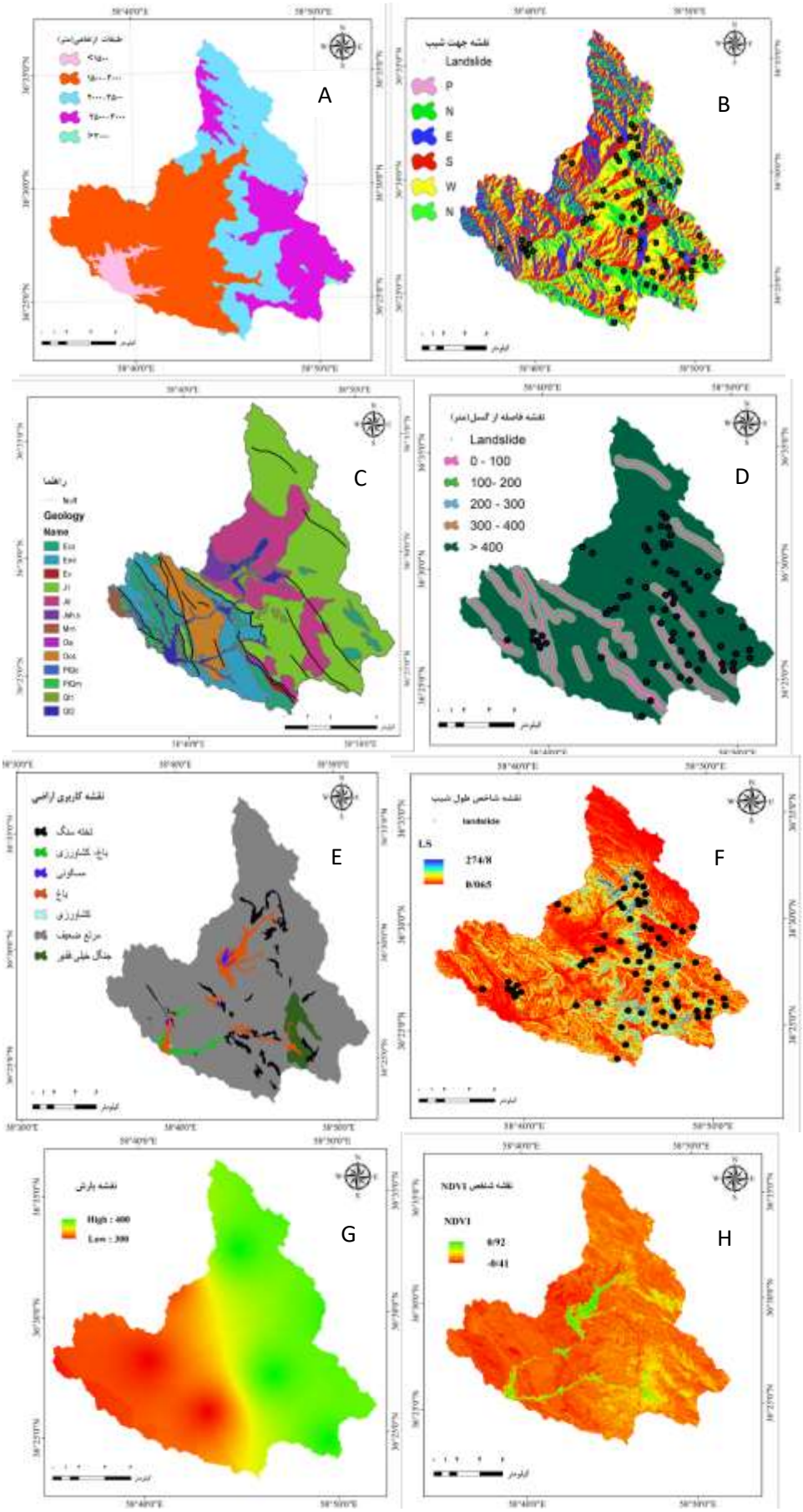
رابطه ٤ :

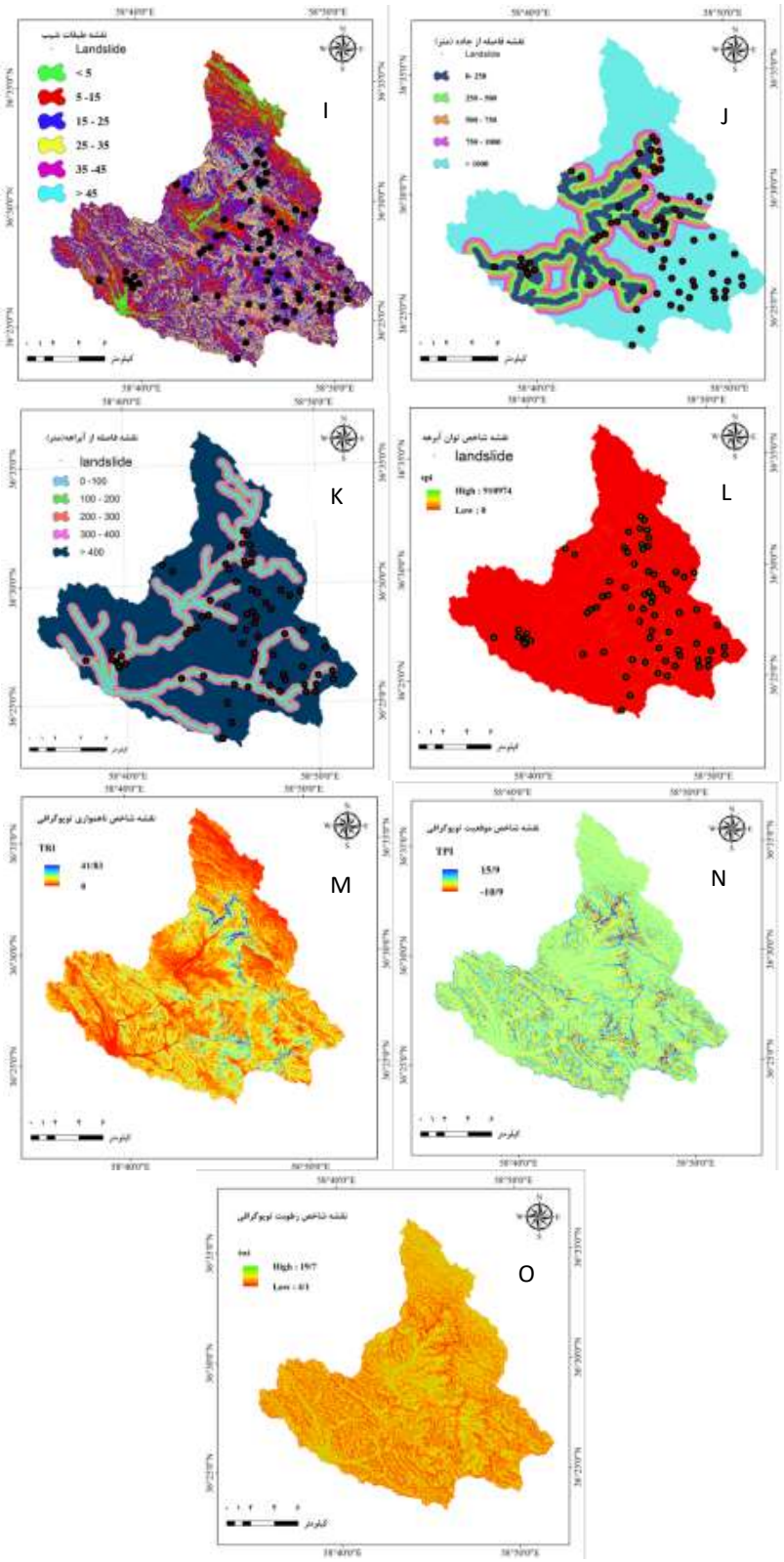
$$TRI=\sqrt{\sum_{p=1}^{8} ZMd}$$

P=تعداد پیکسل اطراف، ZMd= میانگین تفاضل ۸ پیکسل اطراف هرپیکسل.

تهیه نقشه پراکنش زمین‌لغزش

یکی از مهم‌ترین مراحل ارزیابی خطر زمین‌لغزش، شناسایی و تهیه نقشه پراکنش زمین‌لغزش‌های موجود حوضه است. بدین منظور ابتدا با استفاده از نقشه زمین‌لغزش‌های منطقه مورد مطالعه که توسط


A
B
Landslide
P
N
E
S
W
N
C
Geology
Name
D
Landslide
0 - 100
100- 200
200 - 300
300 - 400
> 400
E
F
LS
274/8
0/065
G
High : 400
Low : 300
H
NDVI
0/92
-0/41



شکل۳- نقشه عوامل مؤثر بر وقوع زمین‌لغزش شامل: طبقات ارتفاعی(A)، جهت شیب(B)، زمین شناسی(C)، فاصله از گسل(D)، کاربری اراضی(E)، شاخص طول شیب(F)، بارش(G)، NDVI(H)، طبقات شیب(I)، فاصله از جاده(J)، شاخص ناهمواری توپوگرافی(M)، شاخص موقعیت توپوگرافی(N)، شاخص رطوبت توپوگرافی(O)

Fig 3. Factors affecting landslides: Elevation (A), Aspect (B), Geology (C), Distance from the fault (D), Land use (E), Slope length index (F), Precipitation (G), NDVI (H), Slope (I), Distance from the road (J), Topographic roughness index (M), Topographic position index (N), Topographic wetness index (O)

شکل ٤- نمایی از زمین‌لغزش‌های منطقه موردمطالعه

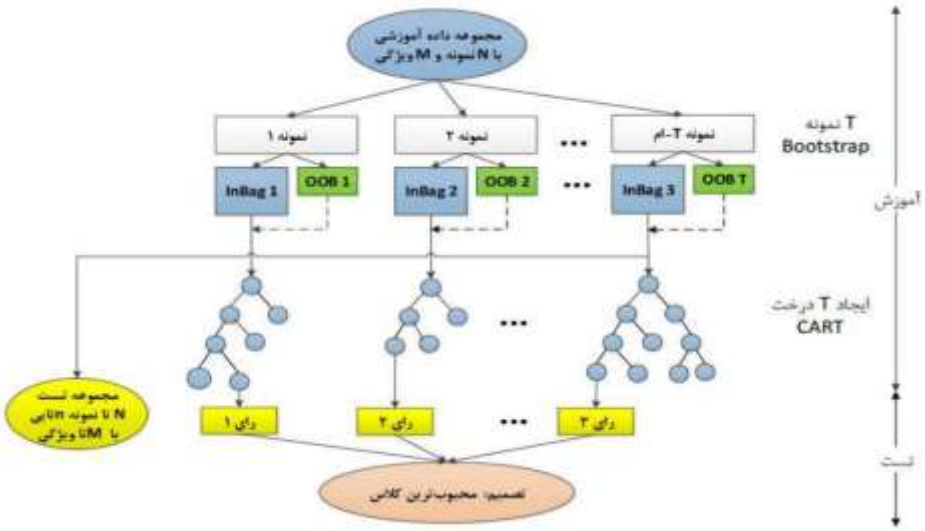
Fig 4. Landslide images of the study area

اداره کل منابع طبیعی استان خراسان رضوی که با استفاده از تفسیر عکس‌های هوایی با مقیاس ۱:۲۰۰۰۰ در سال ۱۳۸۲، تهیه‌شده بود، استخراج شد. سپس از طریق انجام بازدید میدانی در حوضه، استفاده از اطلاعات محلی و دستگاه GPS و همچنین با استفاده از تصاویر ماهواره‌ای Google Earth اقدام به اصلاح این نقشه شد و درنهایت نقشه پراکنش زمین‌لغزش‌ها به‌صورت نقطه‌ای تهیه شد. در نهایت ۷۳ زمین‌لغزش در منطقه موردمطالعه شناسایی شد که از این تعداد ۷۰ درصد برای آموزش مدل و ۳۰ درصد باقیمانده به‌منظور اعتبارسنجی مدل استفاده شد. در شکل ٤ تصاویری از زمین‌لغزش‌های که در منطقه موردمطالعه رخ‌داده‌اند، آورده شده است.

ارزیابی خطر زمین‌لغزش

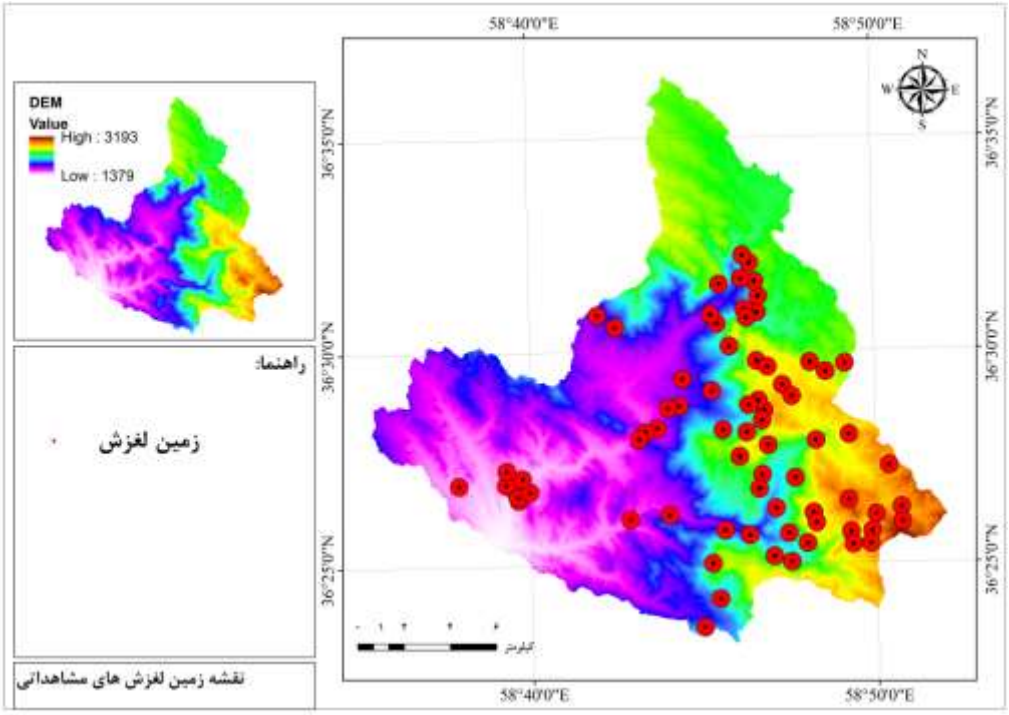
در این تحقیق پس از تعیین عوامل مؤثر بر وقوع زمین لغزش و تهیه لایه‌های مربوط به این عوامل و تهیه لایه زمین‌لغزش‌های موجود در منطقه مورد مطالعه و تهیه بانک اطلاعاتی اقدام به تهیه نقشه حساسیت زمین‌لغزش با استفاده از الگوریتم جنگل تصادفی(RF) در سامانه‌های اطلاعات جغرافیایی و نرم‌افزار آماری R شد. مدل جنگل تصادفی یکی از روش‌های ماشین یادگیری جهت مدل‌سازی درخت تصمیم است [١٦]. در این مدل، نمونه‌گیری تصادفی از اطلاعات و متغیرها، به‌صورت خودکار و تکرارشونده جهت تولید انبوهی از درختان رگرسیونی انجام می‌شود. برای تعیین تعداد بهینه درختان، ابتدا یک تعداد درخت اولیه جهت تولید گراف تغییرات میانگین مربعات خطا (MSE) در مقابل تعداد مشخصی از درختان نمونه آموزشی و ارزیابی، مورداستفاده قرار می‌گیرد. این یک ابزار تحلیلی بسیار قدرتمند جهت جستجوی اطلاعات و اصلاح تعداد بهینه درختان در مدل جنگل تصادفی است. تعداد درختان بهینه به‌گونه‌ای انتخاب می‌شود که اولاً کمترین مقدار میانگین مربعات خطا در مدل‌سازی به دست آید و ثانیاً تعداد درختان نباید آن‌قدر زیاد شود که تجزیه‌وتحلیل متغیرها نیازمند زمان و محاسبات کامپیوتری زیادی باشد. یکی از پارامترهای اصلی در اجرای مدل جنگل تصادفی، پیش‌بینی‌کننده K در هر گره جهت پیش‌بینی مقدار وابسته (پاسخ) است [١٦]. مدل جنگل تصادفی این قابلیت را دارد که میزان خطای به وجود آمده در به هم پیوستن اطلاعات درختان ایجاد شده را برحسب شاخص OOB[1] نشان دهد که بر اساس آن می‌توان تعداد درختان ساخته شده برای آنالیز داده‌های ورودی به نرم‌افزار را بهینه‌سازی کرد. از مهم‌ترین نتایج حاصل از این مدل می‌توان به اولویت‌بندی متغیرهای مستقل (عوامل مؤثر بر وقوع زمین‌لغزش) در حوضه مورد مطالعه اشاره کرد [١٦]. در این مطالعه به‌منظور به‌کارگیری مدل جنگل تصادفی در ارزیابی استعداد زمین‌لغزش از نرم‌افزار R و بسته RandomForest استفاده شده است. روند کلی الگوریتم جنگل تصادفی در شکل ۵ بیان‌شده است.

1. Out-Of-Bag



شکل ۵- روند کلی الگوریتم جنگل تصادفی[۲۸]

Fig 5. General trend of random forest algorithm



شکل ٦- نقشه پراکنش زمین‌لغزش در منطقه موردمطالعه

Fig 6. Landslide distribution map in the study area

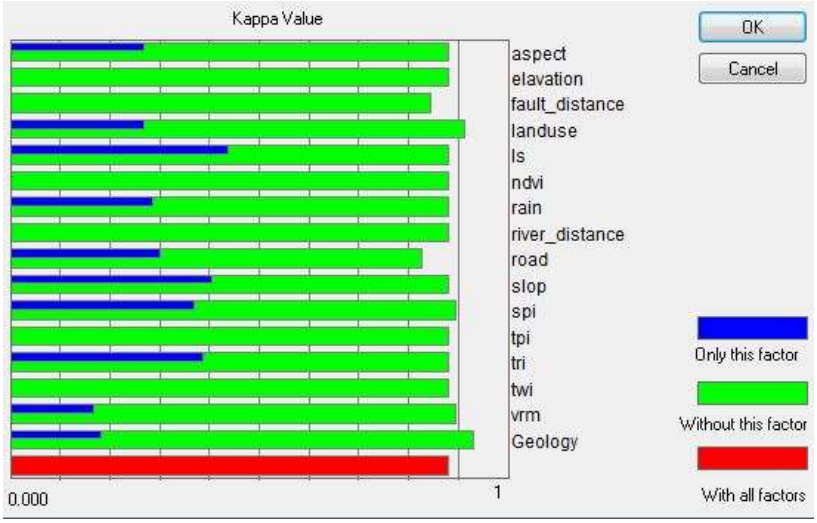
ارزیابی صحت‌سنجی مدل

به‌منظور ارزیابی صحت‌سنجی نتایج مدل، شاخص تشخیص عملکرد نسبی(ROC) به‌کار برده شد [۱۲]. منحنی ROC، نمایش ترسیمی از موازنه بین خطای منفی و مثبت برای هر مقدار احتمالی است. سطح زیر منحنی ROC، بیانگر مقدار پیش‌بینی‌شده مدل از راه توصیف توانایی مدل در تخمین زدن درست رخ دادن پدیده موردنظر و رخ ندادن آن است. مقدار آن بین ۰/۵ تا یک است. هر چه سطح زیر منحنی به یک نزدیک‌تر باشد، بیانگر بهترین دقت است. طبقه‌بندی سطح زیر منحنی با درجه عالی(۱ –۰/۹)، خیلی خوب (۰/۹ – ۰/۸)، خوب (۰/۸ – ۰/۷)، متوسط (۰/۷ – ۰/٦) و ضعیف (۰/٦ – ۰/۵) است [۱۲].

نتایج

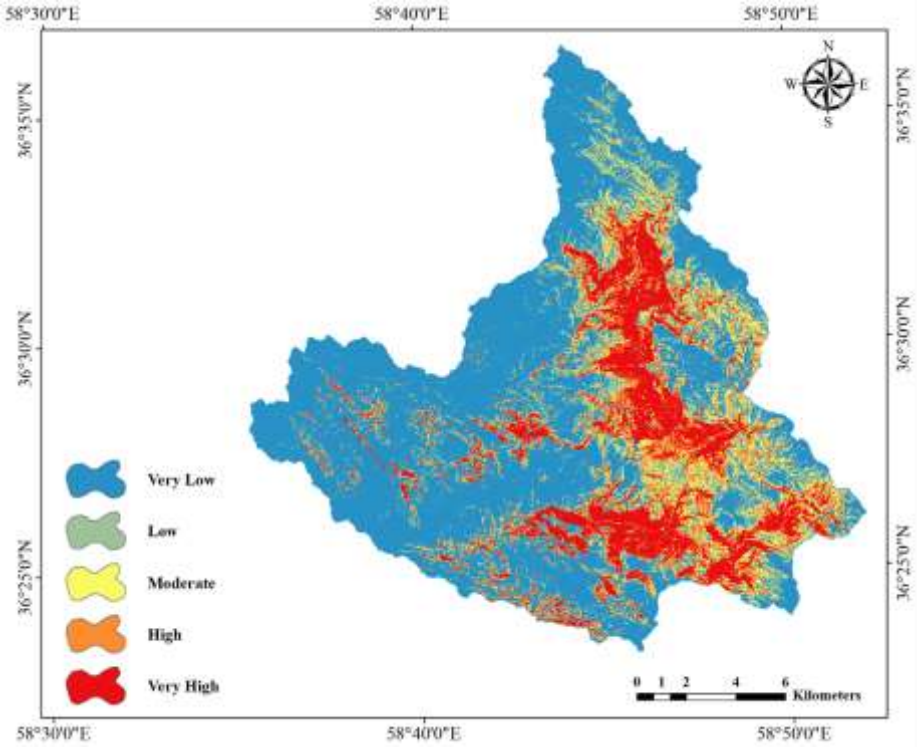
نقشه پراکنش زمین‌لغزش‌های منطقه موردمطالعه

داده‌های تهیه شده از اداره کل منابع طبیعی و آبخیزداری استان خراسان رضوی و پیمایش صحرایی به نرم‌افزار Arc GIS، منتقل و نقشه پراکنش زمین‌لغزش‌های منطقه موردمطالعه تهیه شد. شکل ٦ نقشه پراکنش زمین‌لغزش حوضه با ۷۳ مورد ثبتی را نشان می‌دهد.



شکل ۷- نتایج حاصل از نمودار شاخص کاپا جهت تعیین مهم‌ترین پارامترهای تأثیرگذار بر وقوع زمین لغزش

Fig 7. Results from kappa index chart to determine the most important influential parameters



شکل ۸- نقشه شدت خطر زمین‌لغزش با مدل جنگل تصادفی حوزه آبخیز بار

Fig 8. Landslide hazard intensity map with random forest model of Bar watershed

اولویت‌بندی عوامل مؤثر بر وقوع زمین‌لغزش در منطقه موردمطالعه

یکی از موارد استفاده الگوریتم جنگل تصادفی (RF) اولویت‌بندی عوامل مؤثر بر وقوع زمین‌لغزش است. شکل (۷) نتایج حاصل از نمودار شاخص کاپا جهت تعیین مهم‌ترین پارامترهای تأثیرگذار و سهم هر یک از پارامترها را در پیش‌بینی مدل جنگل تصادفی نشان می‌دهد. با توجه به نمودار شاخص‌های طول شیب و شیب مهم‌ترین پارامترهای تأثیرگذار هستند. همچنین پارامترهای فاصله از گسل و ارتفاع دارای کمترین تأثیر در وقوع زمین‌لغزش هستند.

نقشه حساسیت وقوع زمین‌لغزش در منطقه موردمطالعه

پس از تهیه نقشه حساسیت زمین‌لغزش با استفاده از مدل جنگل تصادفی، نقشه حاصله با استفاده از نقاط عطف (شکست طبیعی) [۱۲، ۲۹]، منحنی تجمعی فراوانی پیکسل‌ها به پنج کلاس خیلی کم، کم، متوسط، زیاد و خیلی زیاد طبقه‌بندی شد (۸). جدول ۱ توزیع فراوانی کلاس‌های خطر زمین‌لغزش با روش مدل جنگل تصادفی حوضه آبخیز بار را نشان می‌دهد.

جدول ۱- درصد مساحت کلاس‌های خطر زمین‌لغزش

Table 1- Percentage of landslide hazard class areas

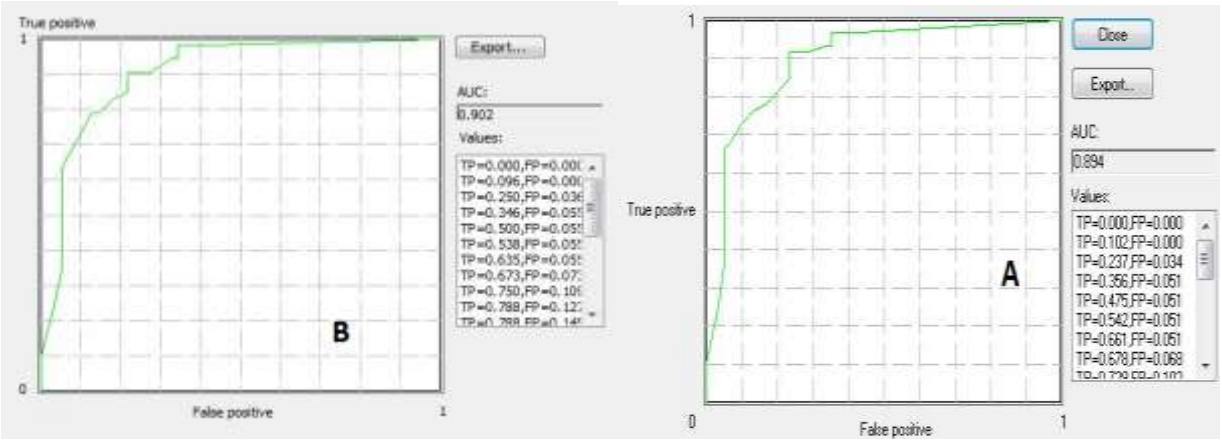
ردیف Number	کلاس خطر Hazard Classes	مساحت (ha) Area(ha)	درصد مساحت Percentage
1	خیلی کم Very Low	15294.5	55.3
2	کم Low	3446.7	12.4
3	متوسط Moderate	2311.5	8.3
4	زیاد High	2170.9	7.8
5	خیلی زیاد Very High	4420.1	15.9
مجموع	Sum	27643.8	100

ارزیابی نتایج مدل جنگل تصادفی

نتایج حاصل از ارزیابی در مرحله آموزش و اعتبار سنجی مدل با استفاده از روش منحنی ROC، در شکل‌های ۹ و ۱۰ نشان داده‌شده است. میزان مساحت سطح زیر نمودار شاخص ROC در مرحله اعتبار سنجی ۰/۸۹٤ به دست آمد، حاکی از قابلیت خیلی خوب مدل در پهنه‌بندی و تعیین مناطق مستعد خطر زمین‌لغزش در حوضه آبخیز بار است.

بحث و نتیجه گیری

حرکت‌های توده‌ای به‌خصوص زمین‌لغزش نقش مؤثری در تخریب جاده‌های ارتباطی، مراتع، مناطق کوهستانی و ایجاد فرسایش و رسوب در حوزه‌های آبخیز دارند. امروزه کشورهای درگیر با مسئله زمین‌لغزش تمایل زیادی به مطالعات جامع شامل ارزیابی و پهنه‌بندی، ریسک، خسارت و تدوین برنامه مدیریتی دارند. از این‌رو در این مطالعه در حوزه آبخیز بار نیشابور در استان خراسان رضوی به‌عنوان یک حوزه مستعد وقوع زمین‌لغزش با مساحتی حدود ۲۷٦٤۳/۸ هکتار اقدام به پهنه‌بندی حساسیت گردید. شرایط طبیعی حوزه آبخیز بار نیشابور مانند زمین‌شناسی، شرایط ناهمواری‌ها، ژئومورفولوژی و تکتونیک و همچنین عوامل تشدیدکننده انسانی مانند تغییر کاربری و راه‌های روستایی بستر مناسبی را برای وقوع زمین‌لغزش به وجود آورده که در مجموع ۷۳ مورد در منطقه موردمطالعه شناسایی شد. در این مطالعه سعی شد تا تعداد گسترده‌ای از عوامل به‌منظور پهنه‌بندی حساسیت زمین‌لغزش استفاده شود. لذا از ١٦ عامل شیب، جهت شیب، طبقات ارتفاعی، زمین‌شناسی، شبکه زهکشی (فاصله از رودخانه)، جاده (فاصله از جاده)، گسل (فاصله از گسل)، شاخص‌های توپوگرافیک (شاخص توان رودخانه (SPI)، شاخص رطوبت توپوگرافی (TWI) و شاخص طول شیب (LS))، شاخص‌های ژئومورفولوژیک (شاخص موقعیت توپوگرافی (TPI)، شاخص ناهمواری توپوگرافی (TRI) و شاخص انحنای سطح (Curvature Index)، کاربری اراضی، شاخص نرمال‌شده تفاوت پوشش گیاهی (NDVI) و خطوط هم بارش به‌عنوان عوامل مؤثر در رخداد زمین‌لغزش در منطقه مورد مطالعه، جهت پهنه‌بندی حساسیت زمین‌لغزش در منطقه موردمطالعه استفاده شد. این عوامل حاصل نظرات کارشناسی و انجام مروری بر منابع در زمینه مدل‌های مختلف است. نقشه حساسیت زمین‌لغزش به‌عنوان سنگ بنای تحقیقات زمین‌لغزش شناخته‌شده و در مواقع بحرانی به‌عنوان یک ابزار مدیریتی مورداستفاده قرار می‌گیرد. در این مطالعه به‌منظور مدل‌سازی حساسیت وقوع زمین‌لغزش در منطقه موردمطالعه از الگوریتم جنگل تصادفی (RF)، استفاده شد. جهت مدل‌سازی خطر زمین‌لغزش از ۷۰ درصد نقاط لغزشی برای آموزش مدل و ۳۰ داده‌های لغزشی به‌منظور اعتبارسنجی مدل مورداستفاده قرار گرفت. بر اساس نتایج مدل، عوامل طول شیب(LS)، شیب و شاخص ناهمواری توپوگرافی(TRI) مؤثرترین عوامل در وقوع زمین‌لغزش در منطقه مورد مطالعه می‌باشند. شیب یکی از پارامترهای تأثیرگذار در ناپایداری‌های دامنه‌ای محسوب می‌شود. به‌طورکلی میزان شیب نقش مهمی را در عواملی همچون نفوذ آب در دامنه‌ها، زاویه شکست و چسبندگی خاک می‌گذارد [۸]. بسیاری



شکل ۹- منحنی ROC و مقدار AUC محاسبه‌شده برای ارزیابی عملکرد مرحله اعتبارسنجی مدل (B) ومرحله آموزش مدل(A)

Fig 9. ROC curve and AUC value calculated to evaluate the performance of the validation step (B) and the training model model (A)

از تحقیقات بر نقش مستقیم شیب و تأثیر آن بر روی زمین‌لغزش‌ها تأکید داشته‌اند [۱۹]. همچنین دووکوتا و همکاران [۸] در مطالعات خود به نقش مؤثر شیب در وقوع زمین‌لغزش تأکید دارند. در عامل شیب، با افزایش شیب حساسیت زمین‌لغزش افزایش‌یافته است که نماینده افزایش نیروی گرانش در شیب‌های بالا است. لذا عامل شیب نسبت به ارتفاع یک عامل بهتر و کامل‌تر(به لحاظ اطلاعات گرانشی موجود) برای مدل‌سازی حساسیت زمین‌لغزش معرفی می‌شود [۱۵]. شاخص طول شیب (LS) یا حمل رسوب (STI) بیانگر توان حمل آبراهه است و در واقع تأثیر توپوگرافی را بر فرسایش مشخص می‌کند. هر چه طول شیب بیشتر باشد توان حمل رسوب بالاتری دارد، در نتیجه شرایط فرسایش کناری بیشتر شده و زمین‌لغزش‌های کناری بیشتر می‌شود. در ارتباط با شاخص ناهمواری توپوگرافی (TRI) و ارتباط این شاخص با وقوع زمین‌لغزش می‌توان گفت که ازآنجایی‌که میزان بالای این شاخص نشان‌دهنده بالاآمدگی و فعالیت نئوتکتونیک است. در نتیجه ناهمواری، به‌هم‌ریختگی و زبری سطوح و افزایش شیب بیشتری رخ می‌دهد، که باعث مساعد شدن شرایط وقوع زمین‌لغزش می‌شود. مدل جنگل تصادفی (RF) با تشکیل انبوهی از درخت‌های تصمیم، درگیر نمودن عوامل زمینه‌ساز مؤثر و شواهد وقوع زمین‌لغزش در طول مدل‌سازی و با تکرار بالا، حذف شاخه‌های تصمیم ناکارآمد در فرایند مدل‌سازی و در نهایت اصلاح پیوسته پیش‌بینی‌ها با به حداقل رساندن خطای پیش‌بینی، قدرت بسیار بالای در پیش‌بینی الگوی حساسیت مخاطرات طبیعی به‌ویژه زمین‌لغزش دارد که نتایج مشابهی در تحقیقات کرنژادی[۱۶]، محمدی و پورقاسمی [۲۱]، طالبی و همکاران [۲۸]، چن و همکاران [۵]، رحمتی و همکاران [۲۳] و پورقاسمی و همکاران [۲۲] حاصل‌شده است. بر اساس نتایج به‌دست‌آمده از این مدل ۶۷/۷ درصد مساحت سطح حوزه در کلاس حساسیت خیلی کم و کم، ۷/۸ درصد در کلاس حساسیت متوسط و ۲۳/۷ سطح منطقه داری حساسیت زیاد و خیلی زیاد، قرارگرفته است. میزان مساحت سطح زیر نمودار شاخص ROC در مرحله اعتبار سنجی ۰/۸۹۴ به دست آمد، حاکی از قابلیت خیلی خوب مدل در پهنه‌بندی و تعیین مناطق مستعد خطر زمین‌لغزش در حوزه آبخیز بار است. با تطبیق نتایج به‌دست آمده با شرایط واقعی موجود از طریق بازدیدهای میدانی، تطابق بسیار بالایی بین نتایج نقشه پهنه‌بندی حساسیت زمین‌لغزش با استفاده از مدل جنگل تصادفی و شواهد واقعی موجود در منطقه موردمطالعه وجود دارد.

منابع

1. Abedini, M., Ghasemyan, B., and Rezaei Mogaddam, M. H. 2017. Landslide susceptibility mapping in Bijar city, Kurdistan Province, Iran: a comparative study by logistic regression and AHP models. Environmental Earth Sciences, 76(8), 1-14.

2. Afifi, M. 2021. Spatial analysis of landslide risk with emphasis on geomorphological factors using stochastic forest model (Case study: Larestan city in Fars province), Journal of Physical Geography, 14(51), 39-53. (In Persain)

3. Azimpoor moghadam. V. 2016. Landslide risk zoning using Bayesian and Dempster-Schiffer theory (Case study: Part of Babolrood watershed). Master Thesis. Faculty of Agricultural Sciences and Natural Resources. Sari University. (In Persain)

4. Cevik, E and Topal, T. 2003. GIS-based landslide susceptibility mapping for a problematic segment of the natural gas pipeline, Hendek (Turkey). Environmental geology, 44(8), 949-962.

5. Chen, W., Pourghasemi, H. R., Kornejady, A., and Zhang, N. 2017. Landslide spatial modeling: Introducing new ensembles of ANN, MaxEnt, and SVM machine learning techniques. Geoderma, 305, 314-327.

6. Dastranj, A and Noor, H. 2021. Comparative evaluation of landslide susceptibility map using Analytical Hierarchy Process (AHP) and Fuzzy methods, Journal of Rs and Gis for natural Resources, 12(3), 62-81. (In Persain).

7. Dastranj, A., Vakili Tajareh, F., and Noor, H. 2021. Evaluation of Hazard Landslide Zonation Using Fuzzy Logic Method in Binalood Mountain, Iranian Journal of Watershed Management Science and Engineering, 15(53), 12-22. (In Persain).

8. Devkota, K., Regmi, A., Pourghasemi, H.R. Yoshida, K. Pradhan, B. Ryu, I. Dhital, M. and Althuwaynee, Q. 2013. Landslide susceptibility mapping using certainty factor, index of entropy and logistic regression models in GIS and their comparison at Mugling–Narayanghat road section in Nepal Himalaya, Nat Hazards, 65: 135-165.

9. Hallaji, M., Mohamadali Zangane Asadi, and Amirahmadi, A. 2020. An Assessment of the Landslide Susceptibility Prediction Models in the Bar Watershed- Neyshabur, Whatershed Management Research, 33(127), 20-30. (In Persain)

10. Harmouzi, H., Nefeslioglu, H. A., Rouai, M., Sezer, E. A., Dekayir, A., and Gokceoglu, C. 2019. Landslide susceptibility mapping of the Mediterranean coastal zone of Morocco between Oued Laou and El Jebha using artificial neural networks (ANN). Arabian Journal of Geosciences, 12(22), 1-18.

11. Hemasinghe, H., Rangali, R. S., Deshapriya, N. L., and Samarakoon, L. 2018. Landslide susceptibility mapping using logistic regression model (a case study in Badulla District, Sri Lanka). Procedia engineering, 212, 1046-1053.

12. Heydari, N., Habibnejad, M., Kavian, K. and Pourghasemi, H. 2020. Landslide Susceptibility Modelling Using the Random Forest Machine Learning Algorithm in the Watershed of Rais-Ali

Delvari Reservoir, Whatershed Management Research, 33(126), 2-13. (In Persain).

13. Hong, H., Pradhan, B., Bui, D. T., Xu, C., Youssef, A. M., and Chen, W. 2017. Comparison of four kernel functions used in support vector machines for landslide susceptibility mapping: a case study at Suichuan area (China). Geomatics, Natural Hazards and Risk, 8(2), 544-569.

14. Karam, A., Paknejad, F., and Bahram Abadi, E. 2021. Zonation of unstable slopes with respect to the debris flows using random forest algorithm (case study: Basin Tngrah Golestan Province), quantitative geomorphological researches, 9(4), 59-74. (In Persain).

15. Koohpayma, A. 2017. Sensitivity zoning, landslide risk assessment and management (Case study: Latian watershed). Ph.D. Thesis. University of Tehran, Campus of Agriculture and Natural Resources, Faculty of Natural Resources. Tehran Iran. (In Persain)

16. Kornejady, A. 2018. Assessment of talent, risk, risk and preparation of strategic landslide management plan for Oghan watershed, Golestan province, Iran. Ph.D. Thesis. Gorgan University of Agricultural Sciences and Natural Resources, Faculty of Rangeland and Watershed Management. Gorgan, Iran (In Persain).

17. Kornejady, A., and Pourghasemi, H. 2019. Landslide Susceptibility Assessment Using Data Mining Models, A case study: Chehel-Chai Basin, Journal of Watershed Engineering and Management, 11(1), 28-42 (In Persain).

18. Kong, C., Tian, Y., Ma, X., Weng, Z., Zhang, Z., and Xu, K. (2021). Landslide Susceptibility Assessment Based on Different MaChine Learning Methods in Zhaoping County of Eastern Guangxi. Remote Sensing, 13(18), 3573.

19. Lee, S. 2005. Application of logistic regression model and its validation for landslide susceptibility mapping using GIS and remote sensing data. Int. J. Remote Sens. 26 (7): 1477–1491.

20. Mohammadnia, M., and Fallah, G. 2018. Landslides susceptibility mapping using fuzzy logic and AHP, Journal of Applied Researches in Geographical Sciences, 18(48), 115-130. (In Persain)

21. Mohammady, M., and Pourghasemi, H. 2017. Prioritization of Landslide-Conditioning Factors and its Landslide Susceptibility Mapping using Random Forest New Algorithm (Case Study: A Part of Golestan province), Journal of Watershed Management Research, 8(15), 161-170. (In Persain)

22. Pourghasemi, H. R & Rahmati, O. (2018). Prediction of the landslide susceptibility: Which algorithm, which precision? Catena, 162, 177-192.

23. Rahmati, O., Kornejady, A., Samadi, M., Nobre, A. D., and Melesse, A. M. (2018). Development of an automated GIS tool for reproducing the HAND terrain model. Environmental modelling & software, 102, 1-12.

24. Regmi, N. R., Giardino, J. R., and Vitek, J. D. (2010). Modeling susceptibility to landslides using the weight of evidence approach: Western Colorado, USA. Geomorphology, 115(1-2), 172-187.

25. Soori, S. Bharvand, T. and Farhadinejad, 2014. Landslide hazard zonation using Fuzzy logic (A case study: ChamSangar watershed), Journal of Rs and Gis for natural Resources, 4(4), 47. (In Persain)

26. Sevgen, E., Kocaman, S., Nefeslioglu, H. A., and Gokceoglu, C. 2019. A novel performance assessment approach using photogrammetric techniques for landslide susceptibility mapping with logistic regression, ANN and random forest. Sensors, 19(18), 3940.

27. Sun, D., Xu, J., Wen, H., and Wang, D. 2021. Assessment of landslide susceptibility mapping based on Bayesian hyperparameter optimization: A comparison between logistic regression and random forest. Engineering Geology, 281, 105972.

28. Talebi, A. Goudarzi, S. and Pourghsemi H. 2018. Investigation of the possibility of landslide hazard mapping using the Random Forest algorithm (Case study: Sardarabad Watershed, Lorestan Province), Journal of Natural environment hazards, 7(16), 45-64. (In Persain)

29. Xu, C., Dai, F., Xu, X., and Lee, Y. H. 2012. GIS-based support vector machine modeling of earthquake-triggered landslide susceptibility in the Jianjiang River watershed, China. Geomorphology, 145, 70-80.

30. Yalcin A. 2008. GIS-based landslide susceptibility mapping using analytical hierarchy process and bivariate statistics in Ardesen (Turkey): comparisons of results and confirmations. Catena, 72(1): 1-12.

31. Yao, J., Qin, S., Qiao, S., Liu, X., Zhang, L., and Chen, J. 2022. Application of a two-step sampling strategy based on deep neural network for landslide susceptibility mapping. Bulletin of Engineering Geology and the Environment, 81(4), 1-20.

32. Yilmaz I. 2009. Landslide susceptibility mapping using frequency ratio, logistic regression, artificial neural networks and their comparison: a case study from Kat landslides (Tokat—Turkey). Computers and Geosciences, 35(6): 1125-1138.

33. Zhao, L., Wu, X., Niu, R., Wang, Y., and Zhang, K. (2020).

Using the rotation and random forest models of ensemble learning to predict landslide susceptibility. Geomatics, Natural Hazards and Risk, 11(1), 1542-1564.

34. Zhou, X., Wen, H., Zhang, Y., Xu, J., and Zhang, W. (2021). Landslide susceptibility mapping using hybrid random forest with GeoDetector and RFE for factor optimization. Geoscience Frontiers, 12(5), 101211.

35. Zhu, A. X., and Mackay, D. S. 2001. Effects of spatial detail of soil information on watershed modeling. Journal of Hydrology, 248(1-4), 54-77.

نشریه علمی
علوم و مهندسی آبخیزداری ایران
Iran-Watershed Management
Science & Engineering

Vol. 17, No. 63, Winter 2024 سال هفدهم- شماره ٦٣- زمستان ١٤٠٢

Landslide Susceptibility Mapping Using Machine Learning Algorithms and Prioritization of Factors Affecting the Occurrence of Landslides

A. Dastranj[1]* , H. Noor[2] and F.VaKili Tajareh[3]

Received: 07-06-2022 Accepted: 08-12-2022

Abstract

The aim of this study was to model the landslide susceptibility using the Random Forest Machine learning technique and prioritization of effective factors on landslide occurrence in Bar watershed in Khorasan Razavi province. The random forest algorithm is based on a bunch of decision trees and is currently one of the best machine learning algorithms. For this purpose, a landslide inventory map was created with 73 historical landslides, which was randomly divided into two datasets for model training (70%) and model testing (30%). A total of 16 landslide-conditioning factors were considered for the susceptibility landslide mapping. The random forest algorithm was run and a landslide susceptibility map was prepared. The RF-based model was validated using the area under the receiver operating characteristic (ROC) curve. The results of evaluation indicated that the success and prediction rates of the model were 90% in training and 89% in validation, respectively. These results confirm the ability of random forest method for prediction of landslide susceptibility models. Also, prioritization of the effective factors showed that the slope length and slope had the highest effect on landslide occurrence. Based on the results of the random forest model, 23.7% of the study area is located in a very high and high sensitivity zone.

Keywords: Random forest algorithm, Bar watershed, ROC

1. Corresponding Author: Soil Conservation and Watershed Management Department, khorasan Razavi Agricultural and Natural Resources Research and Education Center, AREEO, Mashhad. Iran
2. Soil Conservation and Watershed Management Department, khorasan Razavi Agricultural and Natural Resources Research and Education Center, AREEO, Mashhad. Iran
3. Ph. D. in Watershed Science and Engineering, Reclamation of Arid and Mountainous Regions Department, Faculty of Natural Resources, University of Tehran.