

پیش‌بینی مکانی زمین‌لغزش‌های سطحی با استفاده از الگوریتم درخت تصمیم متناوب (مطالعه موردی: مسیر ارتباطی یوزیدر- دگاگا در استان کردستان)

میترا اسدی^۱، لیلا گلی مختاری^{۲*}، عطاءالله شیرزادی^۳، هیمن شهابی^۴، شهرام بهرامی^۵

تاریخ دریافت: ۱۴۰۰/۵/۵

تاریخ پذیرش: ۱۴۰۰/۸/۱۲

چکیده

تهیه نقشه حساسیت‌پذیری زمین‌لغزش اولین گام مهم در ارزیابی خطر زمین‌لغزش است. هدف اصلی این پژوهش بررسی عملکرد الگوریتم درخت تصمیم متناوب (Alternating Decision Tree) برای مدل‌سازی حساسیت به زمین‌لغزش در منطقه یوزیدر تا دگاگا در استان کردستان است. ابتدا نقشه پراکنش زمین‌لغزش‌های سطحی با تعداد ۱۷۵ موقعیت با استفاده از برداشت‌های میدانی تهیه و به صورت کاملاً تصادفی به دو دسته داده مدل‌سازی (۸۰٪/۱۲۳ نقطه) و اعتبارسنجی (۲۰٪/۵۲ نقطه) تقسیم شدند. سپس بیست عامل مؤثر بر وقوع زمین‌لغزش‌های سطحی منطقه مورد مطالعه شناسایی شدند. بر اساس شاخص IGR Information Gain Ratio (IGR) سیزده عامل مؤثر از بین آن‌ها انتخاب و برای مدل‌سازی به کار گرفته شدند. عملکرد مدل بر اساس داده‌های تعلیمی و صحت‌سنجی با معیارهای آماری حساسیت، شفافیت، صحت، میانگین مجذور ریشه مربعات، سطح زیر منحنی مشخصه عملکرد (AUC) ارزیابی شد. نتایج نشان داد که مدل درخت تصمیم متناوب بر اساس داده‌های آموزشی و صحت‌سنجی دارای سطح زیرمنحنی برابر با ۰/۶۶۵ و ۰/۶۷۷ به ترتیب است؛ بنابراین بر اساس نتایج قدرت پیش‌بینی این مدل متوسط است.

کلیدواژه‌ها: زمین‌لغزش‌های سطحی، الگوریتم درخت تصمیم، پیش‌بینی مکانی، صحت‌سنجی، یوزیدر- دگاگا.

۱. دانشجوی دکتری ژئومورفولوژی دانشکده جغرافیا و علوم محیطی، دانشگاه حکیم سبزواری

۲. استادیار ژئومورفولوژی دانشکده جغرافیا و علوم محیطی، دانشگاه حکیم سبزواری، L.mokhtari@hsu.ac.ir

۳. دکتری علوم و مهندسی آبخیزداری و کارشناس آموزشی و پژوهشی دانشکده منابع طبیعی، دانشگاه کردستان

۴. دانشیار و عضو هیئت‌علمی گروه ژئومورفولوژی، دانشکده منابع طبیعی، دانشگاه کردستان

۵. دانشیار ژئومورفولوژی دانشکده علوم زمین، دانشگاه شهید بهشتی تهران

* این مقاله برگرفته از رساله دکتری دانشگاه حکیم سبزواری است.

مقدمه

امروزه بلایای طبیعی و خطرهای ناشی از فناوری به موضوع مهمی در زندگی انسان تبدیل شده است و نمی‌توان از خطرهای و اتفاقات تهدیدآمیزی که زندگی انسان و محیط زیست جهانی را در بر می‌گیرد، چشم پوشی کرد. بر اساس گزارش بانک اطلاعات حوادث اضطراری، زمین‌لغزش‌ها در جهان از سال ۱۹۰۰ تا ۲۰۲۰ میلادی، ۶۶۴۳۸ مورد تلفات جانی و خسارت اقتصادی حدود ۱۰/۸ میلیارد دلاری داشته‌اند (گوها-سپیر^۱ و همکاران، ۲۰۲۰). لذا اقدامات ضروری برای کاهش خسارات ناشی از وقوع زمین‌لغزش، شناسایی نواحی مستعد زمین‌لغزش، تعیین عوامل مؤثر بر وقوع زمین‌لغزش‌ها و تهیه نقشه حساسیت وقوع زمین‌لغزش ضروری هست.

مکانیسم زمین‌لغزش‌ها در سطح وسیعی توسط زمین‌شناسان و ژئومرفولوژیست‌ها با استفاده از روش‌های کمی مورد بررسی قرار گرفته‌اند؛ از نوین‌ترین و دقیق‌ترین این روش‌ها، روش‌های مبتنی بریادگیری ماشین است که یادگیری ماشین یکی از زیرشاخه‌های علوم کامپیوتر و هوش مصنوعی و الگوریتمی است که می‌تواند بدون وابستگی به قوانینی که بر اساس برنامه‌نویسی ایجاد شده‌اند، از داده‌ها یاد بگیرد، در این میان مدل درخت تصمیم یکی از ابزارها و تکنیک‌هایی است که در یادگیری ماشین بسیار پرکاربرد است که در طیف گسترده‌ای برای طبقه‌بندی استفاده می‌شود و دارای ساختاری با فهم بهتر نسبت به دیگر روش‌هاست (میشل^۲، ۱۹۹۷).

اما در درخت تصمیم، احتمال تولید روابط نادرست وجود دارد. بنابراین در این پژوهش از درخت تصمیم متناوب استفاده شد که تعمیم‌یافته درخت تصمیم است که ارتباطات آن تقویت شده است. می‌توان به تحقیقات بسیاری که با استفاده از مدل درخت تصمیم انجام شده است اشاره کرد؛ از جمله: تیان بای^۳ و همکاران (۲۰۱۲)، یوون^۴ و همکاران (۲۰۱۰)، وانگ^۵ و همکاران (۲۰۰۹)، هانگ^۶ و همکاران (۲۰۱۵) چن^۷ و همکاران

همکاران (۲۰۱۷)، تران^۸ و همکاران (۲۰۲۱) اشاره کرد. از میان تحقیقات ذکرشده فوق دو نمونه در ذیل شرح داده می‌شود. تران و همکاران (۲۰۲۱) سه مدل درخت تصمیم متناوب، شبکه عصبی چندلایه و بی‌زین ساده برای تهیه نقشه حساسیت‌پذیری به زمین‌لغزش در هند را مقایسه کردند؛ نتایج آن‌ها حاکی از عملکرد عالی مدل درخت تصمیم متناوب نسبت به مدل‌ها دیگر بود.

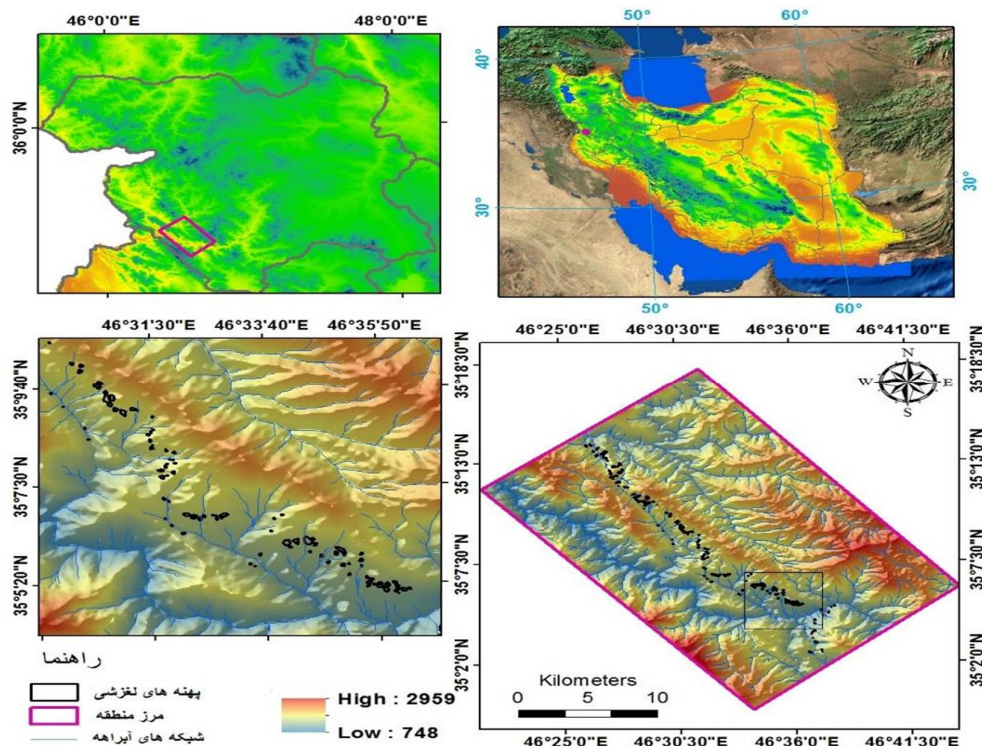
وو^۹ و همکاران (۲۰۲۰) به بررسی کاربرد بالقوه ADTree و دو مجموعه آن یعنی Bagging و AdaBoost در تهیه نقشه حساسیت‌پذیری به زمین‌لغزش در چین پرداختند و عملکرد کلی آن‌ها را مقایسه کردند و با توجه به عملکرد خوب ADTree-AdaBoost این روش به‌عنوان یک مدل جدید، برای اولین بار در تجزیه و تحلیل پیش‌بینی زمین‌لغزش پیشنهاد شد. به‌طور کلی ارزیابی تخریب زمین با توجه به روش‌های درخت تصمیم به‌صورت گسترده مورد استفاده قرار می‌گیرد. تحقیقات نشان می‌دهد که روش‌های مبتنی بر درخت تصمیم متناوب دارای مزایای چون ارائه دقت خروجی بالا در داده‌های بزرگ ورودی و علاوه بر این، در ADT، هر قانون تصمیم‌گیری می‌تواند به‌صورت جداگانه مورد تجزیه و تحلیل قرار گیرد، همان‌طور که به‌صراحت شناخته شده است.

استان کردستان بر اساس خصوصیات توپوگرافی، ژئومرفولوژی، خاک‌شناسی، زمین‌شناسی و اقلیم به دو بخش شرق، جنوب شرقی و مرکزی (به‌عنوان بخش شرقی) و بخش غرب و جنوب غربی (به‌عنوان بخش غربی) تقسیم‌بندی شده است. زمین‌لغزش‌های سطحی و عمقی کردستان غربی به‌دلیل تأثیر برخورد صفحه عربستان با صفحه ایران و همچنین قرارگیری در ناحیه زاگرس رورانده که با گسل خوردگی و شکستگی همراه است، ناشی از فعالیت‌های تکنونیک و ساختمانی هستند (چپی^{۱۰} و همکاران، ۲۰۱۹). یکی از ابزارهای مناسب برای مدیریت بحران در نواحی کوهستانی تهیه نقشه‌های حساسیت، خطر و ریسک زمین‌لغزش بوده است (دهال^{۱۱}، ۲۰۰۸). از این رو مهم‌ترین اهداف این پژوهش عبارت‌اند از:

1. Guha- Sapir
2. Mitchell
3. Tien Bui
4. Yeon
5. Wang
6. Hong
7. Chen

8. Tran
9. Wu
10. Chapi
11. Dahal

۱. شناسایی مهم‌ترین عوامل مؤثر در وقوع زمین‌لغزش؛ حساسیت‌پذیری زمین‌لغزش در بخش جنوب غربی استان کردستان (مسیر ارتباطی یوزیدر تا دگاگا).
۲. بررسی کارایی مدل درخت تصمیم متناوب برای تهیه نقشه.



شکل (۱): موقعیت جغرافیایی منطقه مورد مطالعه

Figure (1): Geographical position of the study area

اراضی موجود در منطقه است. این منطقه جزء حوضه آبخیز خارجی سیروان است و از لحاظ زمین‌شناسی به‌جز جنوب غرب آن که در زاگرس مرتفع قرار دارد، بقیه منطقه در زون سنج سیرجان واقع شده است.

نقشه پراکنش زمین‌لغزش‌های سطحی منطقه مورد مطالعه

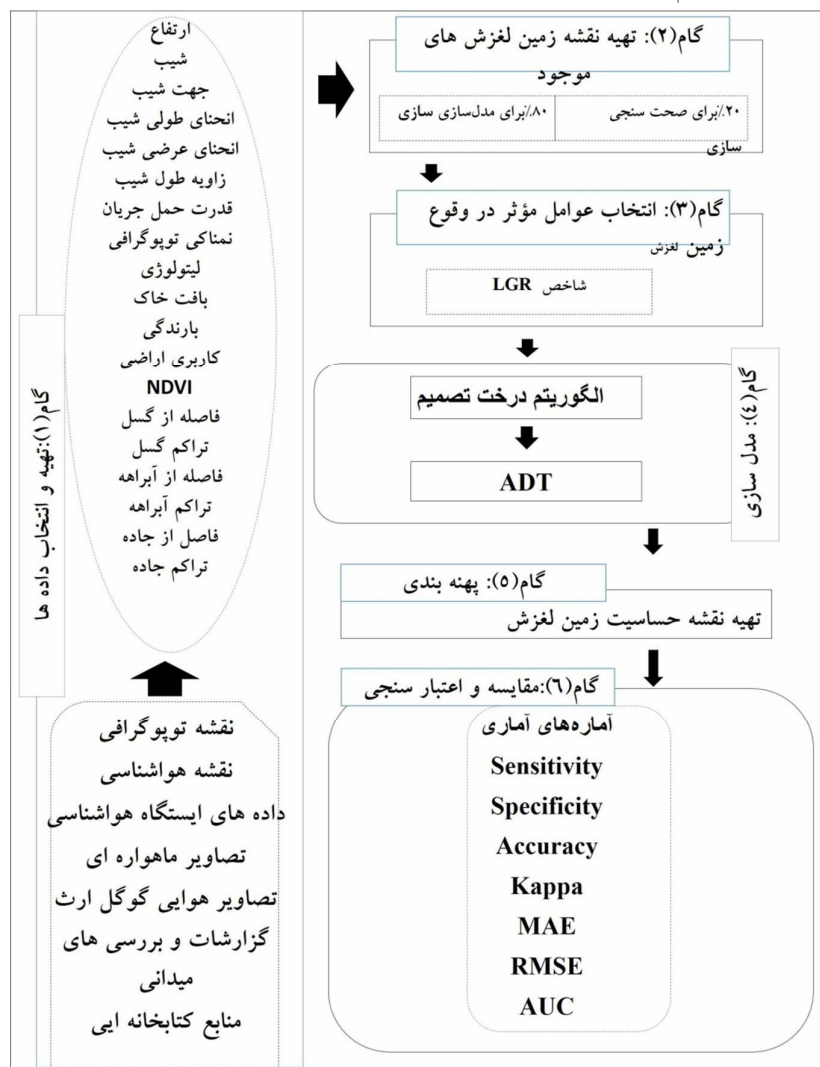
تعیین دقیق محل وقوع زمین‌لغزش‌ها و ایجاد یک پایگاه مکانی داده‌ای برای مطالعات آینده خطر و ریسک آن‌ها ضروری است (شیرزادی، ۲۰۱۷b). در این پژوهش محل دقیق ۱۷۵ زمین‌لغزش با پیمایش صحرائی در عرصه و نیز چک کردن مناطق ثبت‌شده از روی عکس‌های هوایی و تصاویر ماهواره‌ای شناسایی شدند. از این میان کل نقاط ثبت‌شده، ۸۰٪ از زمین‌لغزش‌ها شامل تعداد ۱۲۳ زمین‌لغزش به‌عنوان داده‌های آموزشی و ۲۰٪ باقی‌مانده شامل تعداد ۵۲ زمین‌لغزش به‌عنوان داده‌های آزمون جهت صحت‌سنجی طبقه‌بندی شدند. هم‌راستا با داده‌های زمین‌لغزشی، تعداد ۱۰۰ نقطه غیر زمین‌لغزشی

منطقه مورد مطالعه

همان‌طور که در شکل (۱) نشان داده شده، در استان کردستان در مسیر ارتباطی شهرستان کامیاران به مریوان است که در شمال غربی شهرستان کامیاران و جنوب شرقی شهرستان سروآباد در طول‌های $46^{\circ}43'52''$ تا $46^{\circ}21'12''$ شرقی و عرض‌های $35^{\circ}06'03''$ تا $35^{\circ}11'34''$ شمالی با طول تقریبی ۳۵ کیلومتر قرار گرفته است. حداکثر و حداقل ارتفاع منطقه به‌ترتیب ۲۹۸۲ (جنوب منطقه) و ۷۴۸ (شمال غربی) متر است. بر اساس محاسبات به‌عمل‌آمده روی میانگین بارندگی سالانه در ایستگاه‌های داخل و مجاور منطقه، میانگین بارندگی ۵۱۳ mm برآورد گردید. میانگین سالانه درجه‌حرارت منطقه حدود ۱۶/۶ درجه سانتی‌گراد است. اقلیم منطقه بر اساس تقسیم‌بندی دومارتن جزء مناطق نیمه‌خشک محسوب می‌شود. مرتع‌ها، جنگل‌های پراکنده بلوط، اراضی دیم و باغ‌ها از جمله کاربری

به صورت تصادفی در مکان هایی که زمین لغزش مشاهده نشدند و احتمال وقوع نداشتند به خصوص روی مناطق مسطح و نیز دامنه های با جنس سنگ شناسی مقاوم انتخاب شده و مشابه با

داده های زمین لغزشی به دو دسته داده های آموزشی و صحت سنجی تفکیک شدند.



شکل (۲): فلوجارت روش شناسی مورد استفاده در این پژوهش

Figure (2): Methodology flowchart of the research

روش انجام پژوهش

و عکس های هوایی صحت آنها تأیید شد. نخستین مرحله در تهیه نقشه پیش بینی مکانی زمین لغزش های سطحی در منطقه مورد مطالعه، تبدیل لایه های وکتوری (برداری) به لایه های رستری هست. در شکل (۲) فلوجارت روش شناسی این پژوهش آورده شده است. بر اساس مرور منابع داخلی و خارج کشور مخصوصاً پژوهش های انجام شده در منطقه از جمله قاسمیان و همکاران (۲۰۱۸) و احمدی و همکاران (۲۰۲۰) ۲۰ عامل مؤثر بر وقوع زمین لغزش های سطحی منطقه مورد مطالعه

این تحقیق نتیجه یک پژوهش کاربردی است که در آن از روش های کتابخانه ای، بازدیدهای میدانی و مدل سازی استفاده شده است. نقشه های توپوگرافی ۱:۵۰۰۰۰، نقشه ۱:۱۰۰۰۰۰ و تصاویر IRS سال ۲۰۱۸ منطقه، به عنوان ابزار اصلی در این پژوهش مورد استفاده قرار گرفته اند. زمین لغزش های رخ داده در این منطقه از طریق بازدیدهای میدانی و ثبت و مختصات آنها از طریق GPS شناسایی شدند و با استفاده از تصاویر ماهواره ای

به‌عنوان یک تکنیک استاندارد برای اندازه‌گیری توانایی پیش‌بینی عوامل اثرگذار بر وقوع یک پدیده در روش داده‌کاوی بررسی می‌گردد (تیان بای، ۲۰۱۴؛ شیرزادی و همکاران، ۲۰۱۷a؛ نو^۱ و همکاران، ۲۰۲۰). بنابراین برای شناسایی مهم‌ترین عوامل مؤثر بر وقوع زمین‌لغزش‌های سطحی در منطقه مورد مطالعه، از میان ۲۰ عامل مؤثر معرفی‌شده، از آزمون IGR استفاده شد. بر اساس این آزمون اگر S داده‌های آموزشی با n نمونه ورودی $n(L_i, S)$ تعداد نمونه‌ها در داده‌های آموزشی S متعلق به کلاس L_i (زمین‌لغزش، بدون زمین‌لغزش) باشد، در این صورت داریم (شیرزادی، ۲۰۱۷a):

$$\text{info}(s) = -\sum_{i=1}^2 \frac{n(L_i, S)}{|S|} \log_2 \frac{n(L_i, S)}{|S|} \quad (3)$$

با توجه به عوامل مؤثر بر وقوع زمین‌لغزش، مقدار اطلاعات مورد نیاز برای تقسیم S به مجموعه (S_1, S_2, \dots, S_n) ، از رابطه زیر برآورد می‌شود:

$$\text{Info}(S, A) = \sum_{j=1}^m \frac{S_j}{|S|} \text{Info}(S) \quad (4)$$

شاخص IGR برای یک عامل مؤثر خاص مثل فاکتور A (برای مثال: درجه شیب) از رابطه زیر محاسبه می‌شود:

$$\text{InformationGainRatio}(S, A) = \frac{\text{Info}(S) - \text{Info}(S, A)}{\text{SplitInfo}(S, A)} \quad (5)$$

که در آن Split Info نشان‌دهنده اطلاعات تولیدشده توسط تقسیم S داده‌های آموزشی به m زیرمجموعه است که از رابطه زیر محاسبه می‌شود (شیرزادی و همکاران، ۲۰۱۷a؛ وانگ و همکاران، ۲۰۲۰):

$$\text{SplitInfo} = -\sum_{i=1}^m \frac{S_j}{|S|} \log_2 \frac{|S_j|}{S} \quad (6)$$

الگوریتم درخت تصمیم از نوع ADTree

در یک تعریف ساده می‌توان گفت که این الگوریتم حاصل جمع قوانین یک الگوریتم ضعیف^۲ است (فروند و ماسون، ۱۹۹۹). در این الگوریتم دو گره شامل گره‌های تصمیم و پیش‌بینی وجود دارد. گره تصمیم یک وضعیت را بیان می‌کند و دومی یک مقدار

شناخته شدند. بر این اساس، نقشه‌های درجه شیب، جهت شیب، ارتفاع از سطح دریا، انحنای شیب، انحنای عرضی شیب، انحنای طولی شیب، شاخص قدرت جریان، شاخص نمناکی توپوگرافی، شاخص طول و زاویه شیب دامنه از نقشه مدل رقومی ارتفاع از منبع تصاویر ماهواره‌ای Aster با اندازه سلول ۱۰×۱۰ متر در محیط ArcGIS 10.2 ساخته شدند. نقشه کاربری ارضی و شاخص پوشش گیاهی از تصویر ماهواره‌ای لندست ۸ به تاریخ ۱ جولای ۲۰۱۸ در محیط برنامه ENVI 5.3 با اندازه سلول ۱۰×۱۰ متر تهیه شدند. منطقه مورد مطالعه، نقشه‌های لیتولوژی (جنس سنگ‌شناسی)، فاصله از گسل و تراکم گسل از نقشه زمین‌شناسی شهرستان کامیاران و مریوان با مقیاس ۱:۱۰۰۰۰۰ و نقشه بارندگی یا خطوط هم‌باران بر اساس رابطه رگرسیونی بین ارتفاع و میانگین بلندمدت سالانه بارندگی ایستگاه‌های باران‌سنجی داخل و خارج محدوده مورد مطالعه با اندازه سلول ۱۰×۱۰ متر محاسبه شد. نقشه‌های فاصله از آبراهه و فاصله از شبکه جاده به ترتیب بر اساس بسط فاصله اقلیدسی (Euclidean Distance) در محیط نرم‌افزار Arc GIS برحسب متر فاصله از هر پیکسل از منطقه مورد مطالعه تا نزدیک‌ترین آبراهه و جاده محاسبه گردید و همچنین نقشه تراکم آبراهه و تراکم جاده بر اساس دستور (Line Density) در محیط نرم‌افزار Arc GIS تهیه شدند. برای تعیین بافت خاک‌های منطقه در بازدیدهای میدانی که از منطقه صورت گرفت ۳۰ نمونه خاک از منطقه در بخش‌های لغزشی و غیرلغزشی برداشت شد و سپس با استفاده از روش هیدرومتر در آزمایشگاه بافت خاک نمونه‌های خاک برداشت‌شده مشخص شد و با ایجاد ۱۱ واحد کاری در منطقه، بافت خاک هر واحد کاری مشخص شد.

انتخاب مؤثرترین عوامل در مدل‌سازی پیش‌بینی

مکانی زمین‌لغزش‌های سطحی

تکنیک‌های متعددی برای تشخیص صلاحیت و قابلیت پیش‌بینی متغیرهای اثرگذار بر وقوع یک پدیده وجود دارند. یکی از مهم‌ترین تکنیک‌ها (IGR). Information Gain Ratio است. پایه و اساس IGR تئوری اطلاعات است که با کاهش آنتروپی، صلاحیت اهمیت‌دار بودن عوامل مؤثر را تعیین کرده و

1. Ha Nhu
2. Weak hypothesis

شد و در نهایت با دستور LOOKUP نقشه پهنه‌بندی حساسیت‌پذیری منطقه با استفاده از مدل ADT به دست آمد.

صحت‌سنجی مدل‌سازی زمین‌لغزش‌های سطحی معیارهای آماری

در این بررسی برای ارزیابی و مقایسه نتایج مدل‌سازی از ماتریس 2×2 آزمون صحت‌سنجی پیش‌بینی^۲ برای استخراج معیارهای زیر استفاده شد. این ماتریس شامل ۴ فاکتور شامل False Negative (FN) True Negative (TN) False Positive (FP) True Positive (TP) استوار است. فاکتور TP نسبت تعداد پیکسل‌هایی که به‌درستی به‌عنوان زمین‌لغزش تقسیم شده‌اند و FN تعداد پیکسل حضور زمین‌لغزش‌هایی (۰) که به اشتباه به‌عنوان عدم حضور زمین‌لغزش (۰) طبقه‌بندی شده‌اند، TN تعداد پیکسل عدم حضور زمین‌لغزش‌هایی (۰) که به‌درستی به‌عنوان عدم حضور زمین‌لغزش (۰) طبقه‌بندی شده‌اند، FP تعداد پیکسل عدم حضور زمین‌لغزش‌هایی (۰) است که به اشتباه به‌عنوان زمین‌لغزش (۱) طبقه‌بندی شده‌اند (سانگارتوس و بناردوس^۳، فام^۴ و همکاران، ۲۰۲۱). در نهایت بهترین نتیجه از این چهار حالت زمانی است که مقدار TP زیاد و مقدار FP کم باشد (التواینی^۵ و همکاران، ۲۰۱۴).

حساسیت^۶: عبارت است از نسبت پیکسل‌های زمین‌لغزش که به‌درستی به‌عنوان زمین‌لغزش طبقه‌بندی شده‌اند (بوئی^۷ و همکاران، ۲۰۱۶).

$$\text{Sensitivity} = \frac{TP}{TP+FN} \quad (۷)$$

شفافیت^۸: عبارت است از نسبت پیکسل‌های عدم وقوع زمین‌لغزش که به‌درستی به‌عنوان عدم وقوع زمین‌لغزش طبقه‌بندی شده‌اند (فام و همکاران، ۲۰۱۶).

$$\text{Specificity} = \frac{TN}{TN+FP} \quad (۸)$$

صحت^۹: عبارت است از نسبت پیکسل‌های وقوع و عدم وقوع زمین‌لغزش که به‌درستی طبقه‌بندی شده‌اند (بنیت و

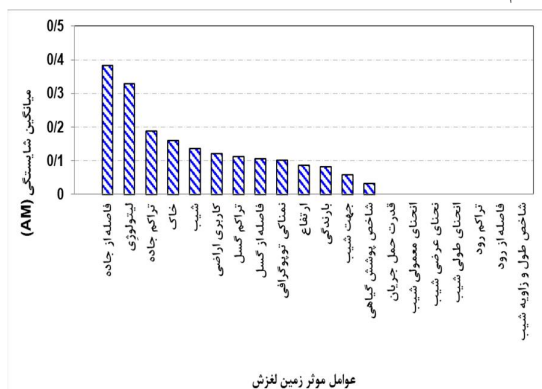
عددی را در بر می‌گیرد (هونگ و همکاران، ۲۰۱۵). در ابتدا این الگوریتم با یافتن بهترین ضریب ثابت پیش‌بینی برای داده‌های آموزشی در ریشه درخت کار خود را شروع می‌کند، سپس بر اساس تکرار در داده‌های آموزشی توسط الگوریتم بوستینگ درخت رشد کرده و یک قانون اصلی (پایه) در یک لحظه اضافه می‌شود. سپس یک گره تصمیم و دو گره پیش‌بینی ایجاد می‌گردد (فروند و ماسون، ۱۹۹۹). برای هر گره پیش‌بینی، این الگوریتم یک وزنی را اختصاص می‌دهد به طوری که در مرحله آخر، قابلیت پیش‌بینی این الگوریتم از طریق جمع تمامی وزن‌ها محاسبه می‌شود (هونگ و همکاران، ۲۰۱۵؛ شیرزادی، ۲۰۱۷).

در این الگوریتم برای یک پیکسل در داده‌های آموزشی $(X_1, Y_1), \dots, (X_m, Y_m)$ که $x_i \in R^d$ و y_i معادل وقوع و عدم وقوع زمین‌لغزش باشد، درخت توسط الگوریتم بوستینگ شروع به رشد می‌کند که در آن در هر تکرار t این الگوریتم دو وضعیت را پشتیبانی می‌کند که عبارت‌اند از پیش‌وضعیت^۱ یا (P_t) و یک گروه از قوانین (R_t) . در هر تکرار از الگوریتم بوستینگ، گروهی از شروط اصلی توسط الگوریتم ضعیف ایجاد می‌شوند که با حرف C بیان می‌گردند. برای اجرای مدل ADT از نرم‌افزار WEKA استفاده شد.

بدین صورت که داده‌های لغزشی و غیرلغزشی که به‌صورت داده‌های تعلیمی و صحت‌سنجی تقسیم‌بندی شده بود به‌عنوان ورودی در نرم‌افزار WEKA بارگذاری شد، پس از بارگذاری اطلاعات در برگه Preprocess و انجام اصلاحات مدنظر بر روی اطلاعات، در برگه Classify با استفاده از امکانات موجود، طبقه‌بندی نمونه‌ها انجام شد. به‌عبارت دیگر در برگه Classify مدل ADT مبنای طبقه‌بندی قرار می‌گیرد و با استفاده از آن داده‌ها طبقه‌بندی می‌شود، همچنین در این قسمت تمهیداتی برای تست کردن مدل ایجاد شده نیز اعمال می‌شود. سپس خروجی که به‌صورت احتمال وقوع زمین‌لغزش برای هر پیکسل از منطقه است، پس از ویرایش در نرم‌افزار EXCEL به‌صورت فایل با فرمت TXT در Arc GIS بارگذاری شد و با دستور JOIN این لایه با لایه‌ای که ترکیبی از عوامل مؤثر در وقوع زمین‌لغزش تلفیق شد و خروجی این لایه به رستر تبدیل

2. Predicted Percentage Correct Test
3. Tsangaratos and Benardos
4. Pham
5. Althuwaynee
6. Sensitivity
7. Bui
8. Specificity
9. Accuracy

شکل (۳) نتایج میانگین شاخص IGR را برای ۲۰ عامل مؤثر انتخاب شده بر وقوع زمین‌لغزش‌های سطحی منطقه مورد مطالعه نشان می‌دهد. بر این اساس، بیشترین مقادیر این شاخص IGR به ترتیب به عوامل فاصله از جاده، لیتولوژی، تراکم جاده اختصاص یافتند، در حالی که عوامل قدرت حمل جریان، انحنای معمولی شیب، انحنای عرضی شیب، انحنای طولی شیب، تراکم رود، فاصله از رود و شاخص طول و زاویه شیب به دلیل اختصاص مقدار صفر برای آن و به دلیل تأثیرگذاری ناچیز بر وقوع زمین‌لغزش‌های سطحی در منطقه مورد مطالعه از مدل‌سازی نهایی حذف و مدل‌سازی با سیزده فاکتور باقی‌مانده انجام شد.



شکل (۳): تعیین مهم‌ترین عوامل اثرگذار بر وقوع زمین‌لغزش‌های

سطحی در منطقه مورد مطالعه بر اساس میانگین IGR

Figure (3): Determining the most important factors affecting the occurrence of surface landslides in the study area based on the average IGR

نقشه پیش‌بینی حساسیت‌پذیری زمین‌لغزش‌های سطحی

منطقه مورد مطالعه

بر اساس توضیحات بخش روش انجام پژوهش، نقشه حساسیت‌پذیری زمین‌لغزش سطحی بر اساس الگوریتم ADT با روش‌های Arc GIS، Natural Breaks، Quantil در محیط برنامه Arc 10.5 تهیه شد. سپس بر اساس روش Natural Breaks نقشه پهنه‌بندی حساسیت‌پذیری زمین‌لغزش به پنج کلاس با حساسیت‌پذیری خیلی کم، کم، متوسط، زیاد و خیلی زیاد طبقه‌بندی شد؛ که بر این اساس حدود ۱۱٪ از مساحت منطقه دارای میزان حساسیت خیلی زیاد به وقوع زمین‌لغزش است.

نتایج به‌دست‌آمده از نقشه پیش‌بینی حساسیت‌پذیری زمین‌لغزش‌های سطحی منطقه مورد مطالعه که در شکل (۴)

همکاران، ۲۰۱۳، ۲۰).

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (9)$$

میانگین مجذور ریشه مربعات خطا^۱ (RMSE): این معیار نشان می‌دهد که مقدار خطا در داده‌ها چقدر است (بنیت و همکاران، ۲۰۱۳، ۲۰).

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (X_{pred.} - X_{act.})^2}{n}} \quad (10)$$

ضریب کاپا (KAPPA) توافق یا اطمینان دوه‌دو را بین دو یا چند کمیت ارزیابی می‌کند (کارلوتا، ۱۹۹۶).

$$Kappa = \frac{P_c - P_{exp}}{1 - P_{exp}} \quad (11)$$

میانگین خطای مطلق (MAE) خطایی است که تفاوت بین مشاهدات زوجی را که به‌طور گسترده در ارزیابی صحت یک الگوریتم استفاده شده است، نشان می‌دهد (فام و همکاران، ۲۰۱۶).

$$MAE = \frac{1}{n} \sum_{i=1}^n |x_{pred.} - x_{act.}| \quad (12)$$

منحنی مشخصه عملکرد^۲

منحنی AUC ROC یک منحنی است که بر روی محور X‌ها مقدار ۱-Specificity (نسبت پیکسل‌های عدم حضور زمین‌لغزش که به‌درستی به‌عنوان عدم حضور زمین‌لغزش طبقه‌بندی شده‌اند) و بر روی محور Y‌ها مقدار Sensitivity (نسبت پیکسل‌های زمین‌لغزش که به‌درستی به‌عنوان زمین‌لغزش طبقه‌بندی شده‌اند) مشخص است. بهترین حالت این منحنی‌ها زمانی است که مقدار سطح زیر این منحنی‌ها برابر یک باشد و برعکس هرچه به سمت صفر نزدیک‌تر باشد، پیش‌بینی ضعیف‌تری حاصل می‌شود (تیان بای، ۲۰۱۴). هنگامی که از داده‌های تعلیمی برای مساحت زیر منحنی استفاده می‌شود منحنی نرخ موفقیت (SRC) نام دارد و هنگامی که از داده‌های صحت‌سنجی استفاده شود، منحنی نرخ پیش‌بینی (PRC) اطلاق می‌گردد (شیرزادی و همکاران، ۲۰۱۷a).

نتایج

تعیین مهم‌ترین عوامل مؤثر بر وقوع زمین‌لغزش‌های سطحی منطقه مورد مطالعه

1. Root Mean Square Error
2. Receiver operating characteristic (ROC) curve

RSME به ترتیب ۰/۶۳۴، ۰/۲۸۳، ۰/۴۲۸ می باشند در حالی که این مقادیر برای داده‌های صحت‌سنجی در این مدل به ترتیب دارای مقادیر ۰/۴۵، ۰/۷، ۰/۱۰، ۰/۸۱، ۰/۸۷، ۰/۸۴، ۰/۶۷۳، ۰/۱۶۴، ۰/۴۰۴ می‌باشند. نتایج این جدول نشان می‌دهد که مدل ADT کارایی نسبتاً متوسطی از خود نشان داد. جدول (۱) نتایج ارزیابی صحت نقشه‌های پیش‌بینی مکانی زمین‌لغزش‌های سطحی منطقه مورد مطالعه را با داده‌های تعلیمی و صحت‌سنجی نشان می‌دهد.

جدول (۱): عملکرد و دقت پیش‌بینی الگوریتم‌ها با مجموعه داده‌های

آموزش و اعتبارسنجی

Table (1): Performance and prediction accuracy of the algorithms by training and validation datasets

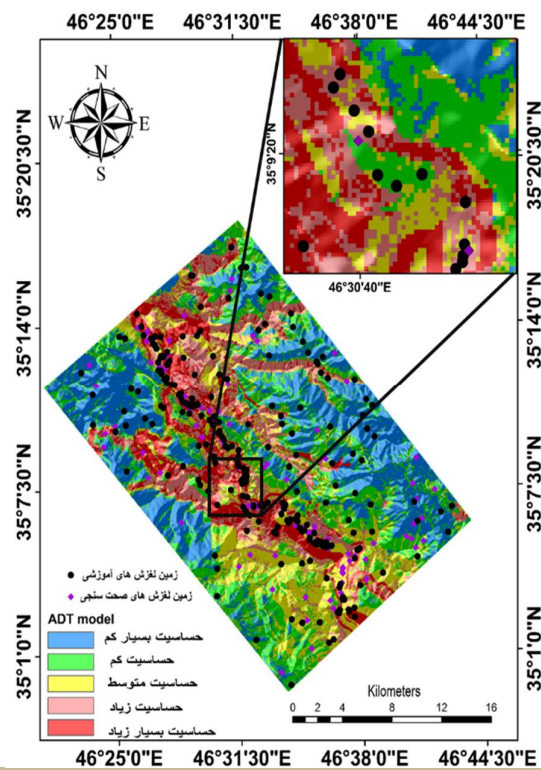
شاخص	داده‌های صحت‌سنجی	داده‌های تعلیمی
	ADT	ADT
True positive	۴۲	۱۰۴
True negative	۴۵	۹۷
False positive	۷	۱۹
False negative	۱۰	۲۶
Sensitivity (%)	۸۱	۸۰
Specificity (%)	۸۷	۸۴
Accuracy (%)	۸۴	۸۲
Kappa	۰/۶۷۳	۰/۶۳۴
MAE	۰/۱۶۴	۰/۲۸۳
RMSE	۰/۴۰۴	۰/۴۲۸

ارزیابی صحت نقشه‌های حساسیت زمین‌لغزش‌های سطحی منطقه مورد مطالعه

درصد مساحت زیر این منحنی (AUC) یک شاخص کمی برای تعیین عملکرد کلی مدل‌ها شناخته شده است (شیرزادی و همکاران، ۲۰۲۱d). هرچه مقدار این شاخص بزرگ‌تر باشد، عملکرد مدل بهتر است. محدوده این شاخص از ۰/۵ (مدل با عملکرد ضعیف) تا ۱ (عملکرد دقیق مدل) متغیر است (بوئی و همکاران، ۲۰۱۶).

شکل (۵) نشان می‌دهد که منحنی ROC بر اساس داده‌های تعلیمی، مقدار سطح زیر منحنی با استفاده از روش ADT، ۰/۶۷۷ است؛ بدین معنی که این روش ۶۷٪ توانایی پیش‌بینی

نشان داده شده، نشان داد که مناطق با حساسیت‌پذیری بسیار زیاد در فاصله ۱۰۰ متری از جاده که تراکم جاده بیشتر است و از لحاظ سنگ‌شناسی در مناطقی با جنس مارن و شیل که حساسیت زیادی به وقوع زمین‌لغزش دارند واقع شده‌اند. همچنین بیشتر زمین‌لغزش‌ها در خاک‌های با بافت لومی و لومی رسی و شیب‌های میانه دامنه‌ها با زاویه شیب ۱۵-۲۵ درجه به وقوع پیوسته‌اند. از لحاظ کاربری اراضی نیز جنگل‌های بلوط منطقه که به اراضی دیم تبدیل شده‌اند، بیشترین مناطق با حساسیت‌پذیری بالا به وقوع زمین‌لغزش را به خود اختصاص داده‌اند. به‌طور کلی مناطقی که فاصله بیشتری از جاده و تراکم کمتر از جاده را دارا هستند و در مناطقی با جنس مقام و سخت از لحاظ سنگ‌شناسی همچون سازندهای بازالت و توف قرار دارند، در مناطق با حساسیت‌پذیری کم و خیلی کم از لحاظ وقوع زمین‌لغزش قرار گرفته‌اند.



شکل (۴): نقشه حساسیت‌پذیری زمین‌لغزش منطقه مورد مطالعه

Figure (4): Landslide sensitivity map of the study area

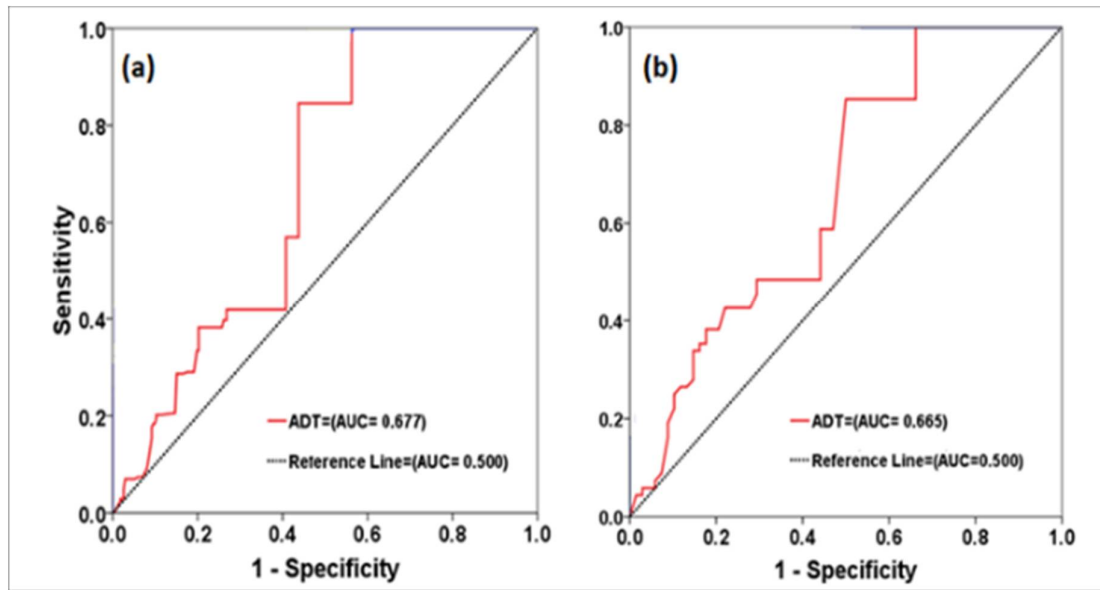
ارزیابی کارایی مدل با داده‌های تعلیمی و صحت‌سنجی

نتایج نشان داد در مدل ADT بر اساس داده‌های تعلیمی مقادیر TP، TN، FP، FN، حساسیت، شفافیت، صحت، کاپا، MAE و

با حساسیت‌پذیری بسیار کم به سمت کلاس‌های با حساسیت‌پذیری بسیار زیاد به‌استثنای کلاس متوسط، مقدار این شاخص اضافه می‌شود. این بدان معنی است که مناطقی که حساسیت‌پذیری بسیار زیاد دارند، دارای تعداد وقوع بیشتری از زمین‌لغزش هستند. شکل (۶) نشان می‌دهد که در روش ADT با افزایش حساسیت‌پذیری به وقوع زمین‌لغزش میزان تراکم زمین‌لغزش‌ها تا حدی افزوده شده است؛ بنابراین صحت نقشه پیش‌بینی مکانی زمین‌لغزش در حد متوسط مورد تأیید واقع شد.

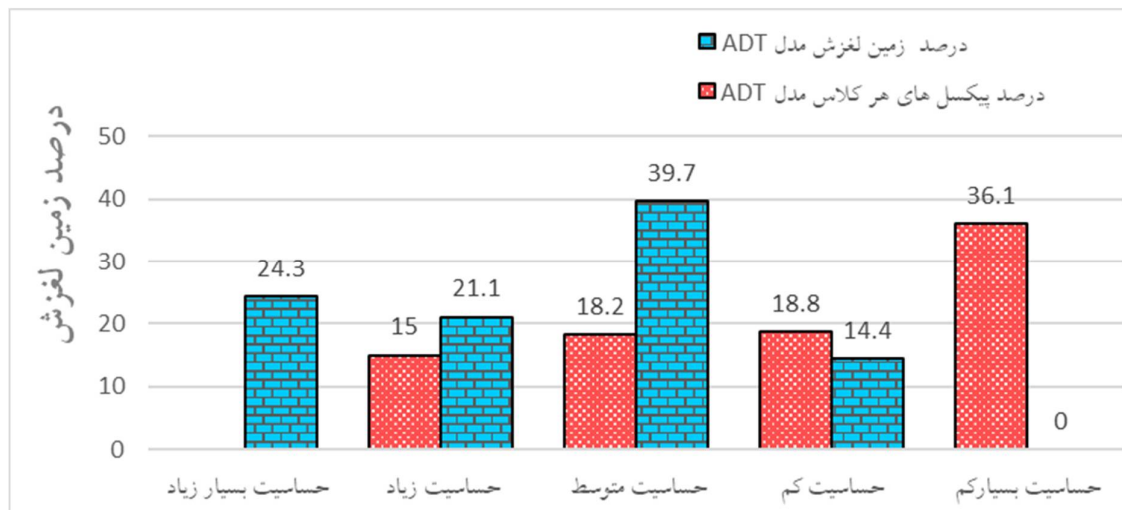
مناطق حساس به وقوع زمین‌لغزش‌ها را بر اساس شرایط موجود و زمین‌لغزش‌های اتفاق افتاده را داشته است، این در حالی است که برای داده‌های صحت‌سنجی مقدار سطح زیر این منحنی برای ۰/۶۶۵ به دست آمد که نشان‌دهنده این موضوع است که مدل برای پیش‌بینی زمین‌لغزش‌ها در آینده ۰/۶۶/۵ توانایی داشته‌است؛ بنابراین طبق نتایج روش ADT دارای عملکرد متوسطی است.

علاوه بر درصد مساحت زیر منحنی ROC از شاخص تراکم زمین‌لغزش نیز استفاده شد. بر این اساس، از کلاس‌های



شکل (۵): منحنی ROC بر اساس الگوریتم ADT با داده‌های تعلیمی (a) و داده‌های صحت‌سنجی (b) (نویسندگان، ۱۴۰۰)

Figure (5): OC curve based on ADT algorithm with training (a) and validation (b) datasets



شکل (۶): تراکم زمین‌لغزش‌ها در کلاس‌های حساسیت مدل ADT

Figure (6): Landslide density in ADT model sensitivity classes

بحث و نتیجه گیری

نتایج حاصل از بررسی عوامل مؤثر با استفاده از شاخص Information Gain Ratio نشان داد که هفت عامل قدرت حمل جریان، انحناى معمولی شیب، انحناى عرضی شیب، انحناى طولی شیب، تراکم رود، فاصله از رودخانه و شاخص طول و زاویه شیب به دلیل دارا بودن مقدار صفر این شاخص و عدم تأثیرگذاری این عوامل از مدل سازی نهایی حذف شدند. از طرفی عامل فاصله از جاده، لیتولوژی و تراکم جاده بیشترین تأثیر را بر وقوع زمین لغزش های منطقه داشتند. علت اصلی این موضوع، هم وجود سنگ های حساس به زمین لغزش از جمله مارن و شیل در دامنه ها و سیاست های نادرست انسانی از جمله ساخت جاده ها و قطع نادرست پاشنه دامنه ها هنگام احداث جاده ها است که این امر به آب فرصت بیشتری برای نفوذ به درون سازندهای حساس خاک داده و با اشباع این خاک ها تحت نیروی ثقل بر روی دامنه ها (عامل شیب) این لغزش ها در منطقه مورد مطالعه اتفاق می افتند. یافته های این مطالعه در زمینه تأثیر بیشتر عامل جاده نسبت به سایر عوامل مؤثر در وقوع زمین لغزش، با نتایج فام و همکاران (۲۰۱۵، ۲۰۱۶، ۲۰۱۹)، تیان بای و همکاران (۲۰۱۹) مطابقت دارد. ارزیابی حساسیت پذیری زمین لغزش یکی از مهم ترین موضوعات در دهه های اخیر است که علت آن شناسایی مناطق حساس زمین لغزش است که می تواند در برنامه ریزی استفاده از زمین تصمیم گیرندگان و هشدار خطر زمین لغزش مورد استفاده قرار گیرد. انتخاب مدل حساسیت پذیری با قابلیت پیش بینی بالا بستگی به روش های مورد استفاده دارد. روش های مختلف حساسیت پذیری زمین لغزش توسط پژوهشگران قبلی پیشنهاد شده است، در حالی که دقت پیش بینی این روش ها هنوز مورد بحث است. نتایج به دست آمده حاکی از آن است عامل فاصله از جاده، دارای بیشترین تأثیر بر وقوع زمین لغزش های منطقه می باشد. این موضوع بیانگر این است که جاده ها سبب تغییر در هندسه شیب دامنه، برهم خوردن تعادل نیروی مقاوم و افزایش تنش های انتقالی زمین بر اثر عبور وسایل نقلیه می شوند و با فاصله گرفتن از این عوارض، حساسیت به

لغزش نیز کاهش می یابد. این امر با نتایج فام و همکاران (۲۰۱۵، ۲۰۱۶، ۲۰۱۹)، بوئی و همکاران (۲۰۱۶) و نیازی و همکاران (۲۰۲۱) در زمینه تأثیر بیشتر عامل جاده نسبت به سایر عوامل مؤثر در وقوع زمین لغزش مطابقت دارد. در درجه دوم اهمیت، بیشترین تأثیر در ایجاد زمین لغزش های منطقه مربوط به عامل لیتولوژی است. از لحاظ لیتولوژی وجود سازندهای حساس به زمین لغزش در این قسمت از جمله مارن و شیل و همچنین توده های سنگ آهک به عنوان یک سطح نفوذناپذیر سخت نقش سطح گسیختگی را برای خاک اشباع شده رویی ایفا می کند و باعث سرخوردن لایه های بالایی می شود. علت اینکه در طبقات بالای دامنه و ارتفاعات بالاتر منطقه، پهنه های لغزشی کم است، وجود توف های بلورین و بازالت و در نتیجه مقاومت آن ها در مقابل هرگونه حرکات توده ای از جمله زمین لغزش است. این امر با نتایج احمدی و فیض نیا (۲۰۰۶)، طالبی و همکاران (۲۰۱۷) و فیض الله پور (۲۰۱۷) مطابقت دارد. فاکتور تأثیرگذار بعدی در وقوع زمین لغزش بافت خاک است. بافت خاک های منطقه بیشتر خاک ها ریزدانه با بافت سنگین می باشند و مقاومت برشی دامنه را کاهش می دهند و بر ناپایداری آن می افزایند. محمدی سوادکوهی و حسینی (۲۰۱۳) بر اساس این یافته های پژوهششان تأکید کردند که وجود لای و رطوبت زیاد در خاک به علت بافت سنگین و ریزدانه می تواند سبب تشدید حرکت های توده ای شود. عامل اثرگذار بعدی بر وقوع زمین لغزش های سطحی منطقه مورد مطالعه، عامل شیب می باشد اکثر زمین لغزش ها در مناطق شیب دار دامنه به وقوع پیوسته اند و همچنین در مناطق با شیب های بیشتر به دلیل جنس سنگی و صخره ای که دارند، حساسیت کمتری نسبت به وقوع زمین لغزش را دارند و نهایتاً مناطق کم شیب که به علت ماهیت کوهستانی منطقه عرصه های محدودی را به خودشان اختصاص داده اند، حساسیت بسیار پایین تری به وقوع زمین لغزش را دارا هستند. بر این اساس مناطق با حساسیت پذیری زیاد و بسیار زیاد در قسمت میانی دامنه واقع شده اند، در مورد پوشش گیاهی زمین، زمین لغزش ها بیشتر در اراضی دیم که قبلاً جنگل های

مثال عرب‌عامری و همکاران (۲۰۲۰) عملکرد مدل‌های ADT، CDT و ترکیبی از این دو مدل را مورد بررسی قرار دادند و نتایج حاکی از دقت بیشتر مدل ترکیبی CDT-ADT با دقت عملکرد ۰/۹۸۱ است و مدل‌های ADT و CDT به ترتیب دقت عملکرد کمتری برابر ۰/۸۶۷ و ۰/۸۳۷ به خود اختصاص دادند. لذا نظر به عملکرد بهتر مدل‌های ترکیبی یادگیری ماشین در کارهای عرب‌عامری (۲۰۲۰)، شیرزادی (۲۰۱۷a)، ساهین^۴ (۲۰۲۰)، نوئن^۵ (۲۰۱۹) و همچنین وو^۶ و همکاران (۲۰۱۹)، می‌توان برای بهبود دقت عملکرد تحقیق حاضر، این امر را مورد توجه قرار داد.

نتایج نشان داد که از کلاس‌های با حساسیت‌پذیری بسیار کم به سمت کلاس‌های با حساسیت‌پذیری بسیار زیاد به مقدار تراکم زمین‌لغزش اضافه می‌شود. این بدان معنی است که مناطقی که حساسیت‌پذیری بسیار زیاد دارند، دارای تعداد وقوع بیشتری از زمین‌لغزش هستند و نقشه‌های به‌دست‌آمده به‌طور متوسط با مناطقی که زمین‌لغزش در آن‌ها اتفاق افتاده‌اند مطابقت دارد.

با توجه به بالا بودن تأثیر جاده در این مدل می‌توان پیشنهاد نمود که اولویت روش‌های جلوگیری و مبارزه با وقوع زمین‌لغزش، باید روی کاهش اثر ساخت جاده‌های منطقه باشد؛ ضمن اینکه اگر در آینده عملیات توسعه جاده در حوضه مدنظر باشد، این عملیات بایستی با رعایت دقیق اصول جاده‌سازی و رعایت پایداری دامنه‌ها انجام شود. نقشه حساسیت‌پذیری زمین‌لغزش ارائه‌شده می‌تواند برای گزینش اقدامات مدیریتی و تصمیم‌گیری مطلوب در برنامه‌ریزی کاربری اراضی، شناسایی نقاط خطر و پیشگیری از وقوع خسارات مرتبط با مخاطره زمین‌لغزش مفید واقع شود.

نیمه‌متراکم و مرتع بوده‌اند رخ داده‌اند؛ که این مسئله اهمیت تخریب جنگل و تغییر کاربری را در وقوع زمین‌لغزش‌ها نشان می‌دهد. فیض‌نیا و همکاران (۲۰۰۴) ضمن پهنه‌بندی خطر زمین به ۴ روش در حوزه آبخیز شیرین‌رود اظهار داشتند که در هر ۴ روش، مناطق با تغییرات زیاد کاربری اراضی خصوصاً تغییرات کاربری از جنگل به زراعی و دیم و باغ‌کاری جزو مناطق پرخطر محسوب می‌شوند. این امر با نتایج کلاستاقی (۲۰۰۷)، لی و سمبات^۱ (۲۰۰۶)، گریکو^۲ و همکاران (۲۰۰۷) مطابقت دارد.

برای ارزیابی مدل‌ها از معیارهای آماری و سطح زیر منحنی عامل گیرنده (AUC)، هم برای داده‌های تعلیمی و هم برای داده‌های صحت‌سنجی استفاده شدند. نتایج اعتبارسنجی مدل‌ها حاکی از آن بود که مدل ADT با سطح زیرمنحنی برابر با ۰/۶۶۵ عملکرد متوسطی از خود نشان داده است و در داده‌های تعلیمی نیز ADT سطح زیر منحنی برابر با ۰/۶۷۷ را به دست آورد. به‌طور کلی مقادیر به‌دست‌آمده در مدل ADT هم در داده‌های تعلیمی هم در داده‌های صحت‌سنجی بیانگر این مطلب است که این مدل از نظر صحت و مدل‌سازی، عملکردی متوسط داشته است و در نهایت منطقه مطالعاتی به پنج کلاس حساسیت بسیار زیاد، زیاد، متوسط، کم و خیلی کم طبقه‌بندی شد. نتایج این تحقیق با تحقیقات محمدی و همکاران (۲۰۲۰) که عملکرد مدل ADT را تقریباً برابر با ۰/۶۰ و همچنین تازه و همکاران (۲۰۱۶) که عملکرد مدل درخت تصمیم ۷۴٪ را به دست آورده‌اند مطابقت دارد. علت عملکرد متوسط این مدل را می‌توان در این امر دانست که در مدل ADT با بزرگ‌تر شدن خطا، خطا به‌صورت تصاعدی رشد می‌کند و خطاهای لایه‌ها بر روی یکدیگر انباشته می‌شوند (پانگ^۳ و همکاران، ۲۰۱۲).

اما این مدل به‌علت اینکه به‌راحتی با داده‌های بزرگ و پیچیده کار می‌کند و با روش‌های دیگر به‌آسانی ترکیب می‌شود، برای بهبود عملکرد مدل ADT، ترکیب این مدل با دیگر مدل‌های یادگیری ماشین باید مدنظر قرار گیرد. برای

4. Sahin
5. Nguyen
6. Vu

1. Lee and Sambath
2. Grico
3. Pang

منابع

- Ahmadi, H., Fayznya, S., 2006. Quaternary Formations (Theoretical and Practical Foundations in Natural Resources), Tehran: University of Tehran, Institute of Publishing and Printing
- Althuwaynee, O.F., Biswajeet, P., Park, H-J. and Lee, J.H., 2014. A novel ensemble decision tree-based CHi-squared automatic interaction detection (CHAID) and multivariate logistic regression models in landslide susceptibility mapping. *Landslides*, 11(6), 1063–1078.
- Arabameri, A., Karimi-Sangchini, E., Chandra Pal, S., Saha, A., Chowdhuri, I., Lee, S. and Tien Bui, D., 2020. Novel Credal Decision Tree-Based Ensemble Approaches for Predicting the Landslide Susceptibility, 12(20), 3389; <https://doi.org/10.3390/rs12203389>.
- Bennett, N.D., Croke, B.F., Guariso, G., Guillaume, J.H., Hamilton, S.H., Jakeman, A.J., Marsili-Libelli, S., Newham, L.T., Norton, J.P. and Perrin, C., 2013. Characterising performance of environmental models. *Environmental Modelling and Software*, 40: 1-20.
- Bui, D.T., Tuan, T.A., Klempe, H., Pradhan, B. and Revhaug, I., 2016. Spatial prediction models for shallow landslide hazards: a comparative assessment of the efficacy of support vector machines, artificial neural networks, kernel logistic regression, and logistic model tree. *Landslides*, 13(2): 361-378.
- Bui, D.T., Pradhan, B., Revhaug, I., Nguyen, D.B., Pham, H.V. and Bui, Q. N., 2015. A novel hybrid evidential belief function-based fuzzy logic model in spatial prediction of rainfall-induced shallow landslides in the Lang Son city area (Vietnam). *Geomatics, Natural Hazards and Risk*, 6(3): 243-271.
- Carletta, J., 1996. Assessing agreement on classification tasks: the kappa statistic. *ArXiv preprint cmp-lg/9602004*.
- Chapi, K., Talebpour ASL, D. and Shirzadi, A., 2019. Comparison of logistic regression models and Bayesian logistic regression for spatial prediction of mass movements in Kurdistan province. *Quantitative Geomorphology Research*, Seventh year, number 2, pp. 60-81
- Chen, W., Xie, X., Wang, J., Pradhan, B., Honge, H., Tieb Bui, D., Duan, Zh., Ma, J., 2017. A comparative study of logistic model tree, random forest, and classification and regression tree models for spatial prediction of landslide susceptibility, *CATENA*, Volume 151, Pages 147-160, <https://doi.org/10.1016/j.catena.2016.11.032>
- Clarestaghi, A.; Graismism, P. (2007). Landslide risk modeling using multi-criteria decision making system in the watershed and Mazandaran province. *Geographical Research*, 22 (4 (87)), 49-68.
- Dahal, R.K., Hasegawa, S., Nonomura, A., Yamanaka M., Masuda, T. and Nishino, K. (2008). GIS-based weights-of-evidence modelling of rainfall-induced landslides in small catchments for landslide susceptibility mapping, *Environmental Geology*, 54(2):311-324.
- Fazolafpoor, M., 2017. Zoning of landslide prone areas using fuzzy neural inference system (ANFIS) (Case study: Sangorchai river basin) *Journal of Natural Hazards*, No. 7, doi: 10.22111 / jneh.2017.3218.
- Freund, Y. and Mason, L., 1999. The alternating decision tree learning algorithm. In *Leading International Machine Learning Conference*, 124-133.
- Ghasemian, B., Abedini, M. and Rustaei, Sh., 2018. Landslide Sensitivity Evaluation Using Vector Support Algorithm (Case Study: Kamyaran County, Kurdistan Province) *Quantitative Geomorphological Research*, Year 6, Issue 3, Pages: 36
- Greco, R., Sorriso-Valvo, M., and Catalano, E., (2007). Logistic regression analysis in the evaluation of mass movement's susceptibility: the Aspromonte case study, *Calabria, Italy. Engineering Geology*, 89(1): 47-66.
- Guha-Sapir, D., Below, R., Hoyois, P., 2020. EM-DAT: international disaster database. Brussels, Belgium: Université Catholique de Louvain. Available from: <http://www.emdat.be> [Accessed 3 March 2020]
- Hong, H., Pradhan, B., Xu, C. and Bui, D.T., 2015. Spatial prediction of landslide hazard at the Yihuang area (China) using two-class kernel logistic regression, alternating decision tree and support vector machines. *Catena*, 133: 266-281..
- Lee, S., Sambath, T., (2006). Landslide susceptibility mapping in the damrei romel area, cambodia using frequency ratio and logistic regression models. *Environ Geol* 50:847–855. Doi: 10.1007/s00254-006-0256
- Mitchell, T., (1997). *Machine Learning*. McGraw Hill.
- Mohammadi Savadkuhi, N., Hosseini, S. A., 2013. The Effect of Physical and Mechanical Properties of Soil on Landslides along Forest Roads (Case Study: Pahneh Kola Series, Tajan Watershed) *Journal of Watershed Management*, Fourth Year / No. 8, pp. 28-42
- Mohammadi, A., Valizade Kamran. Kh., Karimzadeh, S., Shahabi, H., Ansari, N.A., 2020. Flood Detection and Susceptibility Mapping Using Sentinel-1 Time Series, Alternating Decision Trees, and Bag-ADTree Models, *Complexity*, Volume 2020, Article ID 4271376 , <https://doi.org/10.1155/2020/4271376>
- Nhu, V.H., Mohammadi, A., Shahabi, H., Ahmad, B.B., Al-Ansari, N., Shirzadi, A., Clague, J.J., Jaafari, A., Chen, W., Nguyen, H., (2020)

- Landslide susceptibility mapping using machine learning algorithms and remote sensing data in a tropical environment. *International journal of environmental research and public health* 17:4933.
23. Niazi, Y., Manuel, E., Mendoza, Talebi, A., Bidaki, H., (2021). GIS-based support vector machine model in shallow landslide hazards prediction: A case study on Ilam dam watershed, Iran. *Journal of Nature and Spatial Sciences*, 1(1), 59–84
 24. Nguyen, Vu V., Binh T. Pham, Ba T. Vu, Indra Prakash, Sudan Jha, Himan Shahabi, Ataollah Shirzadi, Dong N. Ba, Raghvendra Kumar, Jyotir M. Chatterjee, and Dieu Tien Bui. 2019. Hybrid Machine Learning Approaches for Landslide Susceptibility Modeling, *Forests* 10, no. 2: 157. <https://doi.org/10.3390/f10020157>
 26. Ohlmacher, G.C., 2007. Plan curvature and landslide probability in regions dominated by earth flows and earth slides. *Engineering Geology*, 91(2):117-134.
 27. Pang, P. K., Tien, L. T. and Lateh, H., 2012. Landslide hazard mapping of penang island using decision tree model, in *Proceedings of the International Conference on Systems and Electronic Engineering (ICSEE '12)*, Phuket, Thailand, December.
 28. Pearl, J., 2014. *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*. Morgan Kaufmann. 576 pages
 29. Pham, B. T., Bui, D. T., Pourghasemi, H. R., Indra, P. and Dholakia, M. B., 2015. Landslide susceptibility assessment in the Uttarakhand area (India) using GIS: a comparison study of prediction capability of naïve bayes, multilayer perceptron neural networks, and functional trees methods. *Theoretical and Applied Climatology*, 128(1-2), 255-273.
 30. Pham, BT., Bui, DT., Prakash, I. and Dholakia, M., 2016. Rotation forest fuzzy rule-based classifier ensemble for spatial prediction of landslides using GIS. *Natural Hazards* 83:97-127.
 31. Pham, BT., Bui, D. T. and Prakash, I., 2019. Landslide susceptibility modelling using different advanced decision trees methods. *Civil Engineering and Environmental Systems*, Pp. 139-157, DOI:10.1080/10286608.2019.1568418
 32. Pham, BT., Luub, CH., Phong, TV., Trinh, PT., Shirzadi, A., Renoud, S., Asadi, SH., Le, HV., Meding, J. and Clague, JJ., 2021. Can Deep Learning Algorithms Outperform Benchmark Machine Learning Algorithms in Flood Susceptibility Modeling, *Journal of Hydrology*, Volume 592, <https://doi.org/10.1016/j.jhydrol.2020.125615>
 33. Pourghasemi, H. R., Pradhan, B., Gokceoglu, C., Mohammadi, M. and Moradi, H.R., 2013. Application of weights-of-evidence and certainty factor models and their comparison in landslide susceptibility mapping at Haraz watershed, Iran. *Arabian Journal of Geosciences*, 6(7): 2351-2365.
 34. Pourghasemi, H. R., Moradi H. R., Fatemi Aghda, S. M. B. and Pradhan, G., 2013. GISbased landslide susceptibility mapping with probabilistic likelihood ratio and spatial multi-criteria evaluation models (North of Tehran, Iran). *Arab J Geosci*, DOI 10.1007/s12517-012-0825-
 35. Pong, L., Niu, R., Huang, B., Wua, X., Zhao, Y. and Ye, R., 2014. Landslide susceptibility mapping based on rough set theory and support vector machines: A case of the Three Gorges area, China. *Geomorphology* 204 287–301.
 36. Quinlan, J.R., 1993. *C4.5: programs for machine learning*. Morgan Kaufmann, San Mateo, CA, USA.
 37. Sahin, EK. and Colkesen, I., 2019. Performance Analysis of Advanced Decision Tree-Based Ensemble Learning Algorithms for Landslide Susceptibility Mapping, *Geocarto International* , DOI: 10.1080/10106049.2019.1641560.
 38. Shirzadi, A., 2017a. Spatial prediction of shallow landslide around Bijar county using advanced algorithms of Data mining. PhD Thesis, University of Sari. 132 pp.
 39. Shirzadi, A., Bui, D. T., Pham, B. T., Solaimani, K., Chapi, K., Kavian, A., ... and Revhaug, I., 2017b. Shallow landslide susceptibility assessment using a novel hybrid intelligence approach. *Environmental Earth Sciences*, 76(2), 60.
 40. Shirzadi, A., Soleimani, K., Habibnejad Roshanbaha, M., Kavian, A. and Chapi, K., 2017c. "Introduction of a new hybrid model of the base algorithm in order to predict the sensitivity of surface landslides around the city of Bijar", *Quarterly Journal of Geography and Development*, Volume 14, Number 46. pp.225-246
 41. Shirzadi, A., Shahabi, H., Chapi, K., Bui, D. T., Pham, B. T., Shahedi, K. and Ahmad, B. B., 2017d. A comparative study between popular statistical and machine learning methods for simulating volume of landslides. *CATENA*, 157, 213-226.
 42. Shirzadi, A., Asadi, SH., Shahabi, H., Ronoud, S., CIGUE, J.J., Khosravi, Kh., Pham. B., Ahmad, B., Bui, D., 2021. A novel ensemble learning based on Bayesian Belief Network coupled with an extreme learning machine for flash flood susceptibility mapping. *Engineering Applications of Artificial Intelligence*, volume 98. <https://doi.org/10.1016/j.engappai.2020.103971>
 43. Talebi, A., Godarzi, S., Porghasemi, H., 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.
 44. Tazeh, M., Taghi Zadeh, M.R., Fathabadi, A., Kalantari, S., 2016. Presentation of landslide hazard zoning model and its effective factors

- using quantitative geomorphology, *Journal of Environmental Erosion Research* 6 (22): 15-1
45. Tien Bui, D., Pradhan, B., Lofman, O. and Revhaug, I., 2012. Landslide Susceptibility Assessment in Vietnam Using Support Vector Machines, Decision Tree, and Naive Bayes Models. *Mathematical*
46. Tien Bui, Dieu., Pradhan, Biswajeet., Revhaug, Inge., Trung. Tran, Chuyen., (2014). A Comparative Assessment Between the Application of Fuzzy Unordered Rules Induction Algorithm and J48 Decision Tree Models in Spatial Prediction of Shallow Landslides at Lang Son City, Vietnam, *Cartography from Pole to Pole* pp 303-317, *Problems in Engineering* Volume 2012, Article ID 97463
47. Tien Bui, D., Tuan, T.A., Klempe, H., Pradhan, B. and Revhaug, I., 2016. Spatial prediction models for shallow landslide hazards: a comparative assessment of the efficacy of support vector machines, artificial neural networks, kernel logistic regression, and logistic model tree. *Landslides* 13 (2), 361–378.
48. Tien Bui, D., Shirzadi, A., Shahabi, H., Geertsema, M., Omidvar, E., Clague, J.J., Thai Pham, B., Dou, J., Talebpour Asl, D., Bin Ahmad, B., Lee, S., 2019. New Ensemble Models for Shallow Landslide Susceptibility Modeling in a Semi-Arid Watershed. *Forests* 2019, 10, 743. <https://doi.org/10.3390/f10090743>
49. Tsangaratos, P. and Benardos, A., 2014. Estimating landslide susceptibility through an artificial neural network classifier. *Natural Hazards*, 74(3), 1489-1516.
50. Tran, T. H., Dam, N.D., Jalal, F.E., Naseri, N.A., Siho, L., Phong, T.V., Iqbal, M., Le, H. V., Thi Nguyen, H.B., Prakash, I., Pham. B.T., 2021. GIS-Based Soft Computing Models for Landslide Susceptibility Mapping: A Case Study of Pithoragarh District, Uttarakhand State, India, *Mathematical Problems in Engineering*, Volume 2021, Article ID 9914650, <https://doi.org/10.1155/2021/9914650>
51. Vu, B.T., Pham, V.M., Nguyen, Q.D., Nguyen, H.V., Ngo, T.T.H., 2019. Analyses on drainage capacity and sliding resistance of large diameter vertical wells for deep-seated landslide stabilization, *CIGOS, Innovation for Sustainable Infrastructure. Lecture Notes in Civil Engineering*, vol 54. Springer, Singapore. https://doi.org/10.1007/978-981-15-0802-8_104
52. Wang, X. Ruiqing Niu. 2009. Spatial Forecast of Landslides in Three Gorges Based On Spatial Data Mining. *Sensors* 9, no. 3: 2035-2061. <https://doi.org/10.3390/s90302035>
53. Wang, Y., Bouten, W. and Chen, Q., 2015. *ted Landslide Field Data* 12(2): 268-288
54. Wu, Y., Ke, Y., Chen, Zh., Liang, Sh., Zhao, H., Hong, h., 2020. Application of alternating decision tree with AdaBoost and bagging ensembles for landslide susceptibility mapping. *Catena*, volume 187, <https://doi.org/10.1016/j.catena.2019.104396>
55. Yeon, Y.K., Han, J.G. and Ryu, K.H., 2010. Landslide susceptibility mapping in Injae, Korea, using a decision tree. *Engineering Geology*, Pages 274-283.

Spatial Prediction of Shallow Landslides Using Decision Tree Algorithm: A Case Study of Yozidar-Degaga Route in Kurdistan Province, Iran

Mitra Asadi¹, Leila Goli Mokhtari^{2*}, Ataollah Shirzadi³, Himan Shahabi⁴, Shahram Bahrami⁵

Received: 27/07/2021

Accepted: 03/11/2021

Extended Abstract

Introduction: Iran's mostly mountainous topography, tectonic activity, high seismicity, diverse geological and climatic conditions, population growth-induced pressures on natural resources, and land-use changes in recent decades have created natural conditions for a wide range of landslides. Therefore, it is necessary to take appropriate measures to reduce the damage caused by landslides, identify their prone areas, determine the factors affecting them, and prepare their susceptibility maps. Thus, this study sought to identify the most important factors involved in the occurrence of landslides, and investigate the efficiency of the alternating decision tree models for preparing landslide susceptibility maps in the southwestern part of Kurdistan province (the communication route connecting Yozider to Degaga).

Materials and methods: First, a distribution map with 175 landslides and 100 non-landslide locations was identified and classified into a ratio of 80% and 20% for training and model validation, respectively. Thirteen factors derived from topographic, land cover and rainfall data were selected for modeling using the Information Gain Ratio (IGR) technique. Then, the ADT algorithm was used to train and prepare the landslide susceptibility maps. Moreover, statistical criteria were used to evaluate the models for both training and validation datasets. Finally, the model's performance was evaluated in terms of the area under the receiver operating curve (AUC).

Results: The highest IGR index values were found to belong to the distance from road, lithology, and road density, respectively. Furthermore, factors such as SPI, curvature, profile curvature, plan curvature, river density, distance from the river, and LS proved to have a negative effect on the modeling results, as zero value was assigned to these indices, which, in turn, led to the creation of noise. On the other hand, final modeling was performed and the 13 remaining factors were removed from the model due to their low impact on shallow landslides in the study area. Then, shallow landslide susceptibility maps were prepared based on the ADT algorithm, using quantile, natural breaks, and geometrical interval methods in the ArcGIS 10.2 environment. Finally, the best method was selected based on the landslides' frequency histogram in each susceptibility class of the maps. The results indicated that the natural breaks method was the best-known method, according to which the landslide susceptibility maps were divided into five classes, including very low susceptibility (VLS), low susceptibility (LS), moderate susceptibility (MS), high susceptibility (HS), and very high susceptibility (VHS).

Discussion and Conclusion: The results of IGR-based factor analysis revealed that distance from roads, lithology, and road density had the highest effects on the occurrence of landslides, which could be attributed to the existence of landslide-susceptible formations such as marl and shale, and inappropriate human-set policies

1. Ph. D student Department of Climatology and Geomorphology, Faculty of Geography and Environmental Science, Hakim Sabzevari University, Sabzevar,

2. Department of Climatology and Geomorphology, Faculty of Geography and Environmental Science, Hakim Sabzevari University, Sabzevar, Corresponding author; L.mokhtari@hsu.ac.ir

3. Department of Rangeland and Watershed Management, Faculty of Natural Resources, University of Kurdistan, Sanandaj, Iran

4. Department of Geomorphology, Faculty of Natural Resources, University of Kurdistan, Sanandaj, Iran

5. Department of Physical Geography, Faculty of Earth Sciences, Shahid Beheshti University, Tehran, Iran.

DOI: 10.22052/deej.2021.11.34.39

including road construction and incorrect cutting off of heels, as road construction provides the grounds for more penetration of water into sensitive soil formations and saturation of these soils under force. Moreover, the soil saturation on slopes under the force of gravity can facilitate the occurrence of landslides in the study area, which is consistent with the results found by Pham et al. (2015, 2016, 2019) who reported that the existence of lithological units susceptible to landslides (i.e., marl and shale) in the middle parts of the slopes played an important role in the occurrence of landslides in such areas, considering the fact that marl and shales layers might act as a lubricant for overlying saturated soils, and, therefore, facilitate the occurrence of landslides. However, it could be argued that the upper floors and higher altitudes of the area are less susceptible to landslides due to the presence of crystalline and basaltic units that are resistant to any mass movement, especially landslides. As for the land vegetation, it was found that landslides mostly occurred in drylands that were formerly semi-dense and grassland forests, indicating the significant role of forest degradation and land-use change in landslide occurrence.

The designed models for both the training and validation data were evaluated using the Kappa, TP, specificity, sensitivity, accuracy, and squared statistical measures. Finally, the models' performance was examined through the AUC. Accordingly, the results of model validation showed that the ADT model had a relatively moderate performance with the sub-curve level of 0.665. Furthermore, the sub-curve level of the ADT model was found to be 0.677 for the training data. Finally, the study area was divided into five classes, including the very high, high and moderate, low, and very low sensitivity. It was also found the severity of landslides increased when moving from low-sensitivity classes to the high-sensitivity ones, indicating higher chances for the occurrence of landslides in areas with high sensitivity. Therefore, considering the high influence of the roads in the model proposed in this study, it is suggested that the priority of taking appropriate measures to prevent and/or control landslides be taken into account so that the effect of road construction in the study area could be reduced. Moreover, in cases where future road development operations are considered along the route, principles of road construction and the stability of the slope must be strictly observed.

Keywords: Decision Tree Algorithm, Shallow Landslide, Spatial Prediction, Validation, Yozidar-Degaga.