



"مقاله پژوهشی"

مدل سازی خطر وقوع زمین لغزش با استفاده از داده کاوی در جنگل های هیرکانی

نسترن نظریانی^۱ و اصغر فلاح^۲

۱- پژوهشگر پسا دکتري جنگلداری، دانشکده منابع طبیعی، دانشگاه علوم کشاورزی و منابع طبیعی ساری، ایران، (نویسنده مسوول: Nazariani69@yahoo.com)

۲- استاد گروه جنگلداری، دانشکده منابع طبیعی، دانشگاه علوم کشاورزی و منابع طبیعی ساری، مازندران، ایران

تاریخ دریافت: ۱۴۰۱/۸/۲ تاریخ پذیرش: ۱۴۰۱/۹/۸

صفحه: ۱۲۳ تا ۱۳۴

چکیده مبسوط

مقدمه و هدف: در سال های اخیر، افزایش جمعیت و گسترش سکونتگاه ها در مناطق خطرناک تا حد زیادی تأثیر بلایای طبیعی را در کشورهای صنعتی و در حال توسعه افزایش داده است. پهنه بندی خطر زمین لغزش به شناسایی نقاط استراتژیک و مناطق بحرانی جغرافیایی مستعد، کمک می کند. بنابراین اقدامات برای کاهش سریع، ایمن و برنامه ریزی استراتژیک برای آینده اهمیت دارد. در واقع ارزیابی خطر زمین لغزش ممکن است یک کمک مناسب و مقرون به صرفه برای برنامه ریزی کاربری زمین باشد، لذا در این راستا هدف از پژوهش حاضر مدل سازی خطر وقوع زمین لغزش با استفاده از داده کاوی در حوزه های جنگل شمال کشور است.

مواد و روش ها: برای این منظور نقشه خطر زمین لغزش به روش مور و وارسون با تأثیر فاکتورهای مؤثر بر وقوع زمین لغزش شامل عوامل توپوگرافیکی، عوامل هیدرولوژیکی و اقلیمی، عوامل زمین شناسی، عامل پوشش زمین، عوامل انسانی، شبکه هیدروگرافی از مدل رقومی ارتفاعی، نقشه زمین شناسی با استفاده از نقشه سازمان زمین شناسی کشور تهیه و برای به دست آوردن شاخص پستی و بلندی نسبی ابتدا با استفاده از منحنی میزان های ارتفاعی برگرفته از نقشه های توپوگرافی ۱:۲۵۰۰۰ منطقه نقشه طبقات ارتفاعی تهیه شد. پس از آن، منطقه به شبکه های یک کیلومترمربعی تقسیم و نقشه های با کمترین و بیشترین میزان ارتفاع در شبکه های یک کیلومترمربعی حاصل شد؛ و در مرحله آخر با تفریق این دو نقشه، نقشه ای به دست آمد که اطلاعات آن بیانگر مقدار شاخص پستی و بلندی نسبی است. برای به دست آوردن شاخص رطوبت خاک نیز از بارندگی ماهانه استفاده شد. در نهایت با استفاده از نقشه پهنه بندی حاصل شده از الگوریتم های سه مدل جنگل تصادفی، شبکه عصبی مصنوعی و الگوریتم درخت تصمیم برای مدل سازی خطر وقوع زمین لغزش در محیط نرم افزار STATISTICA 12.0 استفاده شد. همچنین از بین شاخص های مورد بررسی و مؤثر در وقوع زمین لغزش که به عنوان متغیرهای ورودی در مدل سازی وارد شد؛ به ترتیب اولویت، درجه بندی صورت گرفت.

یافته ها: بر طبق نتایج، بیشترین توزیع مساحت زمین لغزش متعلق به طبقه با خطر کم (۷۶ درصد) است و به ترتیب متغیرهای شدت بارندگی، حساسیت لیتولوژیک، پستی و بلندی نسبی و شدت لرزه ای براساس درجه اهمیت به عنوان مهمترین عوامل وقوع زمین لغزش در نظر گرفته شد. نتایج حاصل از اعتبارسنجی با سه الگوریتم جنگل تصادفی، درخت تصمیم و شبکه عصبی مصنوعی؛ با توجه به ضریب تبیین، درصد مجذور میانگین مربعات خطا و اریبی حاصل شده در مدل سازی خطر وقوع زمین لغزش نشان داد که مدل شبکه عصبی مصنوعی با تعداد ۵ لایه ورودی، ۷ لایه پنهان و $R^2=0.99$ ، $RMSE\%=12.28$ و $BIAS\%=-3.23$ نسبت به سایر روش ها دارای دقت بالاتری است.

نتیجه گیری: نتایج پژوهش حاضر نشان داد روش های داده کاوی قابلیت بالایی در پیش بینی خطر وقوع زمین لغزش دارند. لذا استفاده از روش های مذکور می تواند در کاهش خطرات همراه زمین لغزش و برنامه ریزی برای کاربری زمین مورد نظر قرار گیرد.

واژه های کلیدی: حوزه های جنگلی شمال کشور، درخت تصمیم، شبکه عصبی مصنوعی، مدل سازی، مدل رگرسیون

مقدمه

زمین شناسی، حرکت پوسته و فعالیت های انسانی دارد. زمین لغزش ها تحت تأثیر عوامل متعددی در درجات مختلف قرار می گیرند؛ بنابراین، دارای ویژگی های عدم تعادل، غیرخطی، چندمقیاسی و تصادفی هستند؛ که در تحقیقات ارزیابی حساسیت زمین لغزش منطقه ای به طور سیستماتیک حل نشده است (۱۶). جنگل ها یکی از ضروری ترین اجزای حیات روی زمین هستند که هوا، آب و غذای پکی را فراهم می کنند و در عین حال از فرسایش خاک و کاهش تغییرات آب و هوایی جلوگیری می کنند. جنگل ها ۳۱ درصد از مساحت کل زمین را به خود اختصاص داده اند. گزارش مخاطرات ۲۰۲۱ جهان نشان داده که زمین لغزش از نظر خسارات وارده در دسته پنجمین بلایای طبیعی دنیا قرار گرفته است (۳۵). وقوع حرکات توده ای از جمله زمین لغزش، تأثیرات منفی بر زندگی و فعالیت های اقتصادی بشر می گذارد (۲۲). برای انتخاب عوامل مؤثر در تهیه نقشه حساسیت زمین لغزش، دستورالعمل مشخص و یکپارچه ای که بین همه محققین، مورد قبول باشد، وجود ندارد. مطالعات در این زمینه معمولاً با مدل های مختلف، عملکرد و دقت متفاوتی همراه است (۹). براساس زمین لغزش هایی که در گذشته و حال تحت شرایط

زمین لغزش از جمله خطرناک ترین بلایای طبیعی در سراسر جهان است. هر سال، رانش زمین جان چندین نفر را می گیرد و به اموالی به ارزش میلیاردها دلار آمریکا خسارت وارد می کند. بر اساس گزارش پایگاه بین المللی بلایای طبیعی، در بین تمام بلایای طبیعی رخ داده در دوره ۱۹۹۰-۲۰۱۵، رانش زمین به تنهایی ۴/۹ درصد از حوادث بلایای طبیعی را به خود اختصاص داده است (۱۰). زمین لغزش به عنوان حرکت روبه پایین و خارج مواد شیب ساز در جهت گرانش تعریف می شود. اثر متقابل عوامل مختلف طبیعی و انسانی باعث می شود که یک رویداد زمین لغزش به پدیده ای پیچیده تبدیل شود که نیازمند روش های پیشرفته برای پیش بینی و مدیریت است (۳). وضعیت بحرانی، اقدامات برای پیشگیری و پیش بینی بلایای زمین لغزش را سرعت می دهد. نقشه برداری حساسیت زمین لغزش معمولاً به عنوان پایه ای برای پیشگیری از زمین لغزش و برنامه ریزی فضایی عمل می کند، زیرا احتمال وقوع زمین لغزش در یک منطقه را برای آینده براساس تأثیر زمین، محیط و فعالیت های انسانی نشان می دهد (۳۰). در واقع وقوع زمین لغزش ها ارتباط تنگاتنگی با ویژگی های

توپوگرافی، زمین شناسی، هیدروژئولوژی و اقلیمی در یک منطقه اتفاق افتاده است، می توان اطلاعات مفیدی را برای پیش بینی پتانسیل وقوع زمین لغزش های آینده در آن منطقه تهیه نمود (۲۶). روش های نقشه برداری حساسیت زمین لغزش را می توان به طور کلی به روش های کمی و کیفی دسته بندی کرد. روش های کیفی رویکردی ذهنی با تکیه بر دانش خبره برای تعیین وزن به عوامل علی دارند. روش های کمی روی یک رویکرد عینی برای غلبه بر محدودیت های مرتبط با روش های کیفی کار می کنند. در سال های اخیر، در حوزه روش های کمی، روش های یادگیری ماشین مورد توجه گسترده ای قرار گرفته اند، زیرا نتایج قابل اعتماد و امیدوارکننده ای را برای پیش بینی زمین لغزش در سراسر جهان به دست داده اند (۳۲). کارایی نقشه برداری حساسیت زمین لغزش تا حد زیادی به روش های مدل سازی اتخاذ شده بستگی دارد. روش های زیادی برای ارزیابی حساسیت زمین لغزش وجود دارد که از آن جمله می توان به روش های کیفی، قطعی، آماری و یادگیری ماشین اشاره کرد. در روش های کیفی مبتنی بر تجربه، زمین شناسان مهندسی و ژئومورفولوژیست ها از تجربیات و دانش تخصصی برای تحلیل مستقیم یا غیرمستقیم و ترسیم نقشه برداری حساسیت زمین لغزش بر روی نقشه های توپوگرافی موجود و نقشه های زمین شناسی مهندسی استفاده می کنند (۱). در مدل سازی زمین لغزش، بیشتر روش ها و الگوریتم های توسعه داده شده متعلق به روش های رگرسیون لجستیک، روش آماری دو متغیره و چند متغیره، نسبت فراوانی، وزن دهی شواهد، الگوریتم مبتنی بر تئوری بیزین و درخت تصمیم، الگوریتم درخت تصمیم، جنگل تصادفی، ماشین پشتیبان بردار، الگوریتم لجستیک درختی، الگوریتم طبقه بندی و رگرسیون درختی، الگوریتم های ترکیبی مثل ترکیب الگوریتم های درختی با الگوریتم های توابعی و درخت تصمیم است (۲۳). جنگل تصادفی^۱ یک روش مدرن از روش های غیرپارامتری و درختان تصمیم گیری است که شامل انبوهی از درختان کلاس بندی و رگرسیونی و بسط یافته مدل رگرسیون درختی است. درواقع این مدل از یک مجموعه تصمیم گیری قانون مینا برای ارزیابی رابطه های بین متغیرهای وابسته و مستقل استفاده می کند. مهمترین ویژگی جنگل تصادفی، عملکرد بالای آن در اندازه گیری اهمیت متغیرها است که مشخص می کند هر متغیر چه نقشی در پیش بینی پاسخ دارد (۴). مدل شبکه عصبی نسبت به مدل های رگرسیون خطی به دلیل عواملی مانند توانایی یادگیری و تعمیم، و مدل سازی روابط غیرخطی بین لایه های ورودی و خروجی با توجه به تعداد لایه های پنهان، ساختارهای اتصالات مختلف ایجاد می شود و با توجه به تبادل اطلاعات دریافتی (داده های ورودی) و مقدار رفت و برگشت اطلاعات در اتصالات ایجاد شده شبیه مغز انسان، در نهایت یک پاسخ (خروجی) صحیح با حداقل خطا به دست می آید (۱۱). الگوریتم درخت تصمیم^۲ نیز یکی از روش هایی است که در مدل سازی داده ها مورد استفاده قرار می گیرد. مراحل الگوریتم CART به این نحو است که در ابتدا کار با

تمامی داده ها در ریشه درخت (گره ای که در رأس درخت تصمیم قرار دارد) شروع می شود. این روش در ابتدا با استفاده از معیارهای تعیین شده تمامی داده ها را به صورتی بازگشت پذیر تقسیم می کند. سپس تمامی متغیرهای پیشگو را مورد بررسی قرار می دهد تا مشخص کند کدام یک بهترین تقسیم را داشته اند. پس از آن داده های ورودی براساس متغیرهای مستقلی که اهمیت بیشتری دارند تقسیم می شوند (۵). این کار ادامه می یابد تا زمانی که برگ ها شامل داده هایی شوند که همگونی زیادی با یکدیگر دارند (۱۴). در راستای مسائل مطرح شده آرام و همکاران (۳) کارایی مدل های هوش مصنوعی و آماری دو متغیره در تعیین مناطق حساس به وقوع زمین لغزش در استان آذربایجان غربی را مورد بررسی قرار دادند و اذعان داشتند مدل های هوش مصنوعی دارای کارایی بالاتری نسبت به مدل های آماری است. دقت مدل های ترکیبی بیشتر از مدل های منفرد حاصل شد. نتایج منحنی ROC^۳ را با مدل های دقت ۰/۹۲، ۰/۹۱، ۰/۸۹ و ۰/۸۸ را با مدل های به ترتیب ترکیبی جنگل های تصادفی و بگینگ، جنگل تصادفی، نسبت فراوانی و آنترپوی شانون نشان داد. رفیعی ساردویی (۲۴) به شبیه سازی حساسیت زمین لغزش با مدل های داده کاوی شبکه ی عصبی پیچشی، تابع شاهد قطعی و ماشین بردار پشتیبان در منطقه ی رابر، استان کرمان پرداخت. نتایج نشان داد، اندازه های سطح زیر منحنی برای مدل های شبکه ی عصبی پیچشی، ماشین بردار پشتیبان و تابع شاهد قطعی به ترتیب ۰/۹۸۷، ۰/۹۵۸ و ۰/۸۹۹ است. به طور کلی، نتیجه ها همخوانی رضایت بخشی میان داده های زمین لغزش در منطقه و نقشه های حساسیت زمین لغزش نشان داد و کارکرد مدل یادگیری عمیق شبکه ی عصبی پیچشی بیشتر از دو مدل دیگر بود. نقشه ی حساسیت زمین لغزش در چهار رده با حساسیت کم، متوسط، زیاد و خیلی زیاد رده بندی شد. شیرانی و نادری سمانی (۲۹) پژوهشی در زمینه عامل های مؤثر و ارزیابی حساسیت به زمین لغزش با روش های جنگل تصادفی و شبکه ی عصبی مصنوعی در منطقه ی دواب صمصامی استان چهارمحال و بختیاری انجام دادند. نتیجه ی بررسی عامل های مؤثر در هر دو مدل نشان داد که عامل های سنگ شناسی، کاربری و وجه شیب تأثیر بسزایی در رخداد زمین لغزش دارد و بخش زیادی (بیش از ۸۲٪) از زمین لغزش ها در طبقه ی حساسیت خیلی زیاد و زیاد بود. نتیجه ی ارزیابی طبقه بندی و اعتبارسنجی مدل ها نیز نشان داد که دقت و کارآمدی مدل جنگل تصادفی (AUC-ROC=۰/۹۱۹) در پیش بینی رخداد زمین لغزش ها بیشتر از شبکه ی عصبی مصنوعی (AUC-ROC=0.845) است. دس و همکاران (۸) به نقشه برداری پهنه بندی حساسیت زمین لغزش مبتنی بر سیستم اطلاعات جغرافیایی با استفاده از روش فرایند تحلیل سلسله مراتبی^۱ (AHP) در بخش هایی از منطقه کالیپونگ دارجلینگ هیمالیا پرداختند. نتایج نشان داد ناحیه زیر منحنی ROC (AUC-ROC=0.845) نشان می دهد که روش SRC (منحنی میزان موفقیت) (m=۰/۹) با دستیابی به دقت پیش بینی ۷۹/۵ درصد، بالاترین نتیجه را داشت. بنابراین، امیدوارکننده ترین شکل پهنه بندی حساسیت

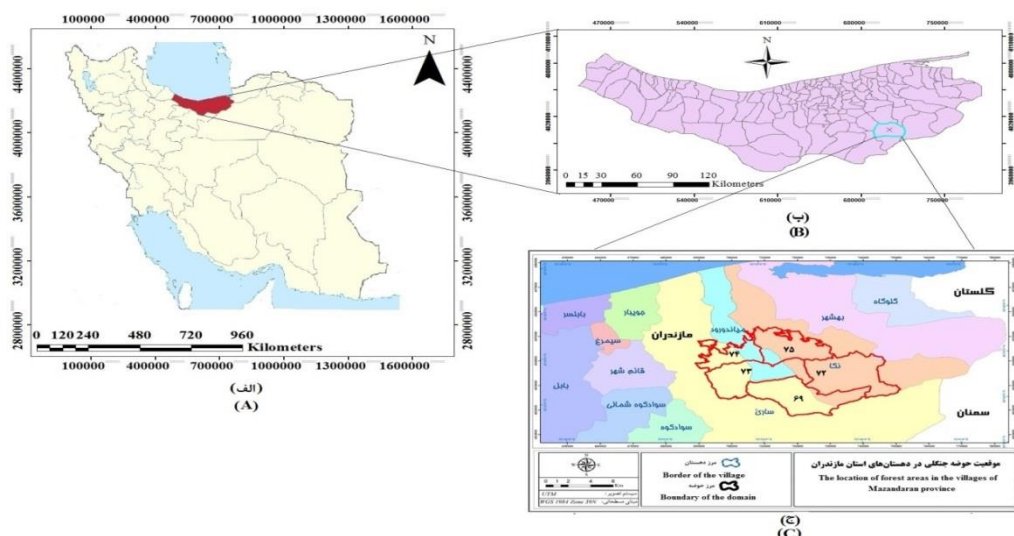
اولویت‌بندی مناطق مستعد به زمین‌لغزش مورد هدف است. مدل ارزش اطلاعاتی بر مبنای داده‌های واقعی از زمین‌لغزش رخ داده با استفاده از سه مدل جنگل تصادفی، شبکه عصبی مصنوعی و الگوریتم CART انتخاب شد. در مدل رگرسیون به علت اینکه متغیر وابسته (رخداد زمین‌لغزش) کمی نبوده و از نوع دوجمله‌ای است، یعنی دارای دو حالت: هست (رخداد) و نیست (عدم رخداد) است (۱۴)، ضرورت استفاده از الگوریتم‌های جنگل تصادفی، شبکه عصبی مصنوعی و CART چند متغیره توجیه می‌شود. با توجه به بهره‌برداری‌هایی که از حوزه‌های جنگل‌های شمال کشور به لحاظ قطب تجاری صنعت چوب از دیرباز تا طرح تنفس مصوبه سال ۹۲ هیأت وزیران و بند ۳ ماده ۳۸ قانون ششم برنامه توسعه سازمان منابع طبیعی و آبخیزداری کشور و در پی این موضوع با توجه به نیاز روزافزون به صنعت چوب برای دوره پسا تنفس این جنگل‌ها انجام شده است؛ به تبع ساختار زمین‌شناسی این جنگل‌ها در معرض عواملی مانند زمین‌لغزش قرار دارد. لذا در پژوهش حاضر مدل‌سازی خطر وقوع زمین‌لغزش با استفاده از داده‌کاوی ناپارامتریک در حوزه‌های جنگل‌های هیرکانی مورد هدف قرار گرفت.

مواد و روش‌ها

منطقه مورد پژوهش

پژوهش حاضر در حوزه جنگل‌های شمال کشور در دهستان‌های گرماب، تنگه سلیمان و چهاردانگه شهرستان ساری و دهستان استخرپشت شهرستان نکا با مساحت ۳۰۶۸۵ هکتار انجام شد. این منطقه در موقعیت جغرافیایی بین ۱۶'، ۱۵' و ۵۳' تا ۱۵'، ۴۳' و ۵۳' طول شرقی ۲۳'، ۱۵' و ۳۶' تا ۲۳'، ۲۶' و ۳۶' عرض شمالی واقع شده است. حداکثر ارتفاع حوزه ۲۶۳۳ متر و حداقل ارتفاع برابر ۳۳۶ متر از سطح دریا است. شکل ۱ به ترتیب موقعیت حوزه در کشور، استان مازندران و موقعیت جغرافیایی حوزه‌های جنگلی مورد مطالعه در شهرستان‌های استان مازندران را به تصویر کشیده است.

به زمین‌لغزش برای منطقه مورد مطالعه حاضر در نظر گرفته شد. آنها همچنین اذعان داشتند نتایج به‌دست‌آمده از این مطالعه، اطلاعات مکانی مناطقی را که ممکن است با ناپایداری شیب مواجه شوند برجسته می‌کند و به سازمان‌های دولتی و ذینفعان به دلیل عدم وجود سیستم‌های هشدار زودهنگام زمین‌لغزش‌ها مناسب در این منطقه کمک می‌کند تا اقدامات کافی را تهیه کنند. القاضی و همکاران (۲) پژوهشی با هدف انتخاب پارامترهای بهینه برای حساسیت زمین‌لغزش‌ها در عقبه السلیات، عربستان سعودی با استفاده از روش‌های یادگیری ماشین مانند، جنگل تصادفی (RF)، رگرسیون لجستیک (LR) و شبکه عصبی مصنوعی (ANN) انجام دادند. نتایج نشان داد، دو مدل هیبریدی (LR-ANN و LR-RF) در دسته بسیار بالا با ۲۹/۶۷-۳۲/۷۳ کیلومتر مربع، مدل (LS) در دسته بالا با ۲۱/۸۴-۳۳/۳۸ کیلومتر مربع، مدل (LR) با ۲۲/۳۴ کیلومتر مربع به‌عنوان بسیار بالا و ۴۵/۱۵ کیلومتر مربع به‌عنوان بالا پیش‌بینی شدند. همچنین نتایج نشان داد، مدل LR-RF به‌عنوان بهترین مدل ($AUC=0.941$) و پس از آن LR-ANN ($AUC=0.915$) و LR ($AUC=0.872$) حاصل شد. فام و همکاران (۲۱) در پژوهشی، پنج مدل یادگیری ماشین پیش‌بینی‌کننده شامل (بهترین درخت تصمیم اول (BFT)، درخت عملکردی (FT)، درخت تصمیم J48 (J48DT)، درخت ساده بیزین (NBT) و کاهش خطا در هرس درختان (REPT)) را برای نقشه‌برداری حساسیت زمین‌لغزش منطقه وان چان در استان یون بای، ویتنام توسعه دادند. نتایج نشان داد مدل RSSFT بالاترین عملکرد را از نظر شکل پیش‌بینی لغزش‌های آینده به دست داد. بنابراین اذعان داشتند، مدل ترکیبی RSS-FT (مدل شبه فضایی تصادفی درخت عملکردی) نسبت به سایر مدل‌های مورد مطالعه، مدل قوی‌تری است که می‌تواند در سایر مناطق نقشه‌برداری حساسیت زمین‌لغزش برای برنامه‌ریزی و مدیریت صحیح کاربری اراضی مورد استفاده قرار گیرد. در پژوهش حاضر ضرورت بررسی زمین‌لغزش در جنگل و به‌کارگیری مدل مناسب جهت شناسایی و



شکل ۱- موقعیت منطقه مورد مطالعه به‌ترتیب در: الف) کشور ایران، ب) استان مازندران و ج) حوزه‌های آبخیز مورد مطالعه
Figure 1. The location of the studied area in: A) Iran, B) Mazandaran province and C) Studied watersheds

روش کار

در پژوهش حاضر ابتدا اقدام به پهنه بندی زمین لغزش در حوزه های جنگلی مورد پژوهش شد. در مرحله بعد با توجه به پهنه بندی صورت گرفته با استفاده از مدل موروارسون در منطقه که بیشترین درصد (مساحت) زمین لغزش در آن به دست آمد اقدام به مدل سازی با سه روش جنگل تصادفی، الگوریتم درخت تصمیم و شبکه عصبی مصنوعی شد.

تهیه نقشه خطر زمین لغزش به روش مورا و وارسون

برای ارزیابی پتانسیل زمین لغزش با استفاده از روش مورا-وارسون از عوامل خارجی و دینامیک که به ترتیب شامل لیتولوژی، رطوبت خاک، درجه شیب و شدت لرزه ای شدت بارندگی می شود، استفاده شد. با ترکیب این عوامل، پتانسیل

جدول ۱- کلاس زمین لغزش به روش موروارسون (۱۸)

Table 1. Landslide class according to the Mora -Vahrson method (18)

Amount مقدار	Description توصیف	Class کلاس
<6	No slippage بدون لغزش	1
7-32	Very low خیلی کم	2
33-162	Low کم	3
163-512	Medium متوسط	4
513-1250	Hight زیاد	5
>1250	Very hight خیلی زیاد	6

فاکتورهای مؤثر بر وقوع زمین لغزش های سطحی

عوامل اساسی مؤثر بر وقوع زمین لغزش شامل عوامل توپوگرافیکی (از قبیل زاویه و جهت شیب، ارتفاع از سطح دریا)، عوامل هیدرولوژیکی و اقلیمی (بارندگی، فاصله از آبراهه ها یا رودخانه ها، تراکم آبراهه یا تراکم زهکشی)، عوامل زمین شناسی (لیتولوژی یا جنس سنگ بستر، فاصله از گسل و تراکم گسل)، عامل پوشش زمین (کاربری اراضی)، عوامل انسانی (فاصله از جاده و تراکم جاده) است.

داده های مورد استفاده

شبکه هیدروگرافی حوزه با استفاده از نقشه های ۱:۲۵۰۰۰ و از مدل رقومی ارتفاعی محدوده تهیه شد. نقشه زمین شناسی با استفاده از نقشه سازمان زمین شناسی کشور (مقیاس ۱:۱۰۰۰۰۰ تصاویر ماهواره ای، عکس هوایی، مشاهدات

صحرائی و سایر گزارش ها تهیه شد. برای به دست آوردن شاخص پستی و بلندی نسبی (Sr) و تهیه نقشه مورد نظر، ابتدا با استفاده از منحنی میزان های ارتفاعی برگرفته از نقشه های توپوگرافی ۱:۲۵۰۰۰ منطقه نقشه طبقات ارتفاعی تهیه شد. در جدول ۲ شاخص حساسیت لیتولوژیک در وقوع زمین لغزش طبقه بندی شده است. پس از آن منطقه مطالعه به شبکه های یک کیلومتری تقسیم و نقشه های با کمترین و بیشترین میزان ارتفاع در شبکه های یک کیلومتری حاصل شد؛ و در مرحله آخر با تفريق این دو نقشه، نقشه ای به دست آمد که اطلاعات آن بیانگر مقدار شاخص پستی و بلندی نسبی است. برای به دست آوردن شاخص رطوبت خاک (Sh) نیز از بارندگی ماهانه استفاده شد.

جدول ۲- طبقه بندی تأثیر لیتولوژی و حساسیت آن در وقوع زمین لغزش (۲۸)

Table 2. Classification of the influence of lithology and its sensitivity in the occurrence of landslides (28)

سنگ شناسی Lithology	حساسیت Sensitivity	کلاس Class
آهک نفوذپذیر، سنگ های نفوذی با خردشدگی کم، بازالت، آندزیت ها، گرانیت ها، سنگ هایی با مقاومت برشی زیاد، شدت هوازدگی کم، سطح ایستایی کم و سطوح شکستگی تمیز Permeable limestone, intrusive rocks with low crushing, basalt, andesites, granites, rocks with high shear resistance, low weathering intensity, low static surface and clean fracture surfaces.	کم Low	1
سنگ های فوق الذکر و سنگ های رسوبی تخریبی با بافت کاملاً توده ای با شدت هوازدگی زیاد، شکستگی های برش پذیر و مقاومت برشی کم The aforementioned rocks and destructive sedimentary rocks with a completely lumpy texture with high weathering intensity, shearable fractures and low shear strength.	ملاپیم Smooth	2
سنگ های آتشفشانی، دگرگونی، آذرین درونی و رسوبی با هوازدگی قابل توجه Volcanic, metamorphic, internal igneous and sedimentary rocks with significant weathering.	متوسط Medium	3
سنگ های دگرسان شده گرمابی با هوازدگی قابل توجه، قویاً گسسته شده و شکافدار همراه با پرشدگی رسی (درزه)، آذرآواری های کم متراکم شده با سطوح ایستایی کم عمق Hydrothermal metamorphosed rocks with significant weathering, strongly dissociated and fissured with clay fillings (joints), poorly compacted pyroclastics with shallow depositional surfaces.	زیاد Hight	4
سنگ های کاملاً دگرسان شده، خاک های برجا، آبرفت ها، کوهرفت هایی با مقاومت برشی کم و سطح ایستایی کم Completely altered rocks, soils, alluviums, landslides with low shear resistance and low water table.	بسیار زیاد Very hight	5

الگوریتم‌های ناپارامتریک داده‌کاوی الگوریتم جنگل تصادفی

اجرای مدل‌سازی به روش جنگل تصادفی و تعیین تعداد بهینه درختان ابتدا با ۴۰۰ درخت اولیه انجام شد. سپس با مشاهده نمودار تغییرات میزان مربعات خطا برای داده‌های آزمون و آموزش در جایی که با افزایش تعداد، روند افزایش مربعات خطا حالت ثابتی پیدا نمود به عنوان تعداد درخت بهینه انتخاب گردید. به منظور تعیین تعداد بهینه برآوردگرها از جذر تعداد کل متغیرها استفاده شد. برای تعیین تعداد برآوردکننده‌ها در هر گره از جذر کل تعداد متغیرهای مستقل مورد استفاده در مدل $2 \pm K$ استفاده شد تا K یا تعداد برآوردکننده بهینه در هر گره، انتخاب گردد (۱۵).

شبکه عصبی مصنوعی

در پژوهش حاضر دو شبکه عصبی پرسپترون چند لایه^۱ و تابع پایه شعاعی^۲ (شبکه عصبی توابع اساسی شعاعی) که از رایج‌ترین شبکه‌های عصبی مصنوعی پیشخور هستند، مورد استفاده قرار گرفت. این شبکه‌ها با الگوریتم پس انتشار آموزش داده می‌شوند و دارای ساختار لایه‌ای هستند. شبکه عصبی شامل دو خروجی مدل شبکه عصبی است که به ترتیب شامل نام شبکه (MLP و RBF)، نورون‌هایی که دارای لایه‌های ورودی^۳، پنهان^۴ و خروجی^۵ است که با هم همخوانی دارند و در نهایت یک راه‌حل با حداقل خطا برای مسئله، همگرایی می‌کنند (۱۸). همچنین به منظور بهینه‌سازی از روش شبه نیوتنی BFGS^۶ استفاده شد.

الگوریتم درخت تصمیم

مدل درخت تصمیم‌گیری یک فرآیند طبقه‌بندی است که به طور تکراری یکسری داده آموزشی را به زیر بخش‌های کوچکتر بر مبنای آزمایش‌ها به یک یا بیشتر از ارزش‌های عارضه تقسیم می‌کند. بدون فرض بودن، شرط درخت تصمیم است (۳۴). در این روش ابتدا کار با تمامی داده‌ها در ریشه درخت (گره‌ای که در رأس درخت تصمیم قرار دارد) شروع می‌شود. این روش در ابتدا با استفاده از معیارهای تعیین شده تمامی داده‌ها را به صورتی بازگشت‌پذیر تقسیم می‌کند. سپس تمامی متغیرهای پیشگو را مورد بررسی قرار می‌دهد تا مشخص کند کدام یک بهترین تقسیم را داشته‌اند. پس از آن داده‌های ورودی بر اساس متغیرهای مستقلی که اهمیت بیشتری دارند تقسیم می‌شوند (۴،۵). این کار ادامه می‌یابد تا

زمانی که برگ‌ها (گره‌هایی که دیگر تقسیم نمی‌شوند و در انتهای درخت قرار گرفته‌اند) شامل داده‌هایی شوند که همگونی زیادی با یکدیگر دارند (۱۴). اندازه بهینه درخت با استفاده از معیار پارامتر پیچیدگی و تکنیک اعتبارسنجی متقابل k-fold تعیین شد.

اعتبارسنجی مدل‌های رگرسیونی ناپارامتریک

به منظور بررسی مدل‌ها ۷۰ درصد داده‌ها برای آموزش و ۳۰ درصد برای آزمون در نظر گرفته شد (۳۳). با استفاده از ضریب تبیین (R^2) (رابطه ۱)، معیارهای میانگین مجذور مربعات خطا^۲ (رابطه ۲) و درصد میانگین مجذور مربعات خطا (رابطه ۳) و درصد اریبی (رابطه ۴) اعتبار مدل‌های آماری مورد ارزیابی قرار گرفت.

$$R^2 = \frac{\sum_{i=1}^n (y_{oi} - \bar{y}_p)^2}{\sum_{i=1}^n (y_{oi} - \bar{y}_o)^2} \quad \text{رابطه ۱}$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (\hat{y}_i - y_i)^2}{n}} \quad \text{رابطه ۲}$$

$$RMSE\% = \frac{\sqrt{\sum_{i=1}^n (\hat{y}_i - y_i)^2 / n}}{\bar{y}} \times 100 \quad \text{رابطه ۳}$$

$$Bias\% = \frac{\sum_{i=1}^n (\hat{y}_i - y_i) / n}{\bar{y}} \times 100 \quad \text{رابطه ۴}$$

در این روابط $\bar{y}$: مقدار برآورد شده، $\bar{y}$: میانگین برآورد شده و $\bar{y}$: مقدار مشاهده شده، همچنین n: تعداد مشاهدات (قطعات نمونه)، و y_{oi} : مقدار مشاهده نام است.

تنظیم و تحلیل اولیه داده‌ها در محیط Excel انجام شد و برای مدل‌سازی از نرم‌افزارهای SPSS22.0 و STATISTICA12.0 استفاده شد.

نتایج و بحث

پهنه‌بندی خطر زمین‌لغزش به روش مور-وارسون

پس از تلفیق امتیاز عددی نقشه‌های موضوعی هر یک از عوامل، نقشه خطر زمین‌لغزش روش مور-وارسون با سه طبقه خطر، تهیه شد (شکل ۲). در نقشه به دست آمده بیشترین توزیع مساحت متعلق به طبقه با خطر کم و کمترین متعلق به طبقه با خطر متوسط است.

1- Multi-Layer Perceptron: MLP

2- Radial Basic Functions: RBF

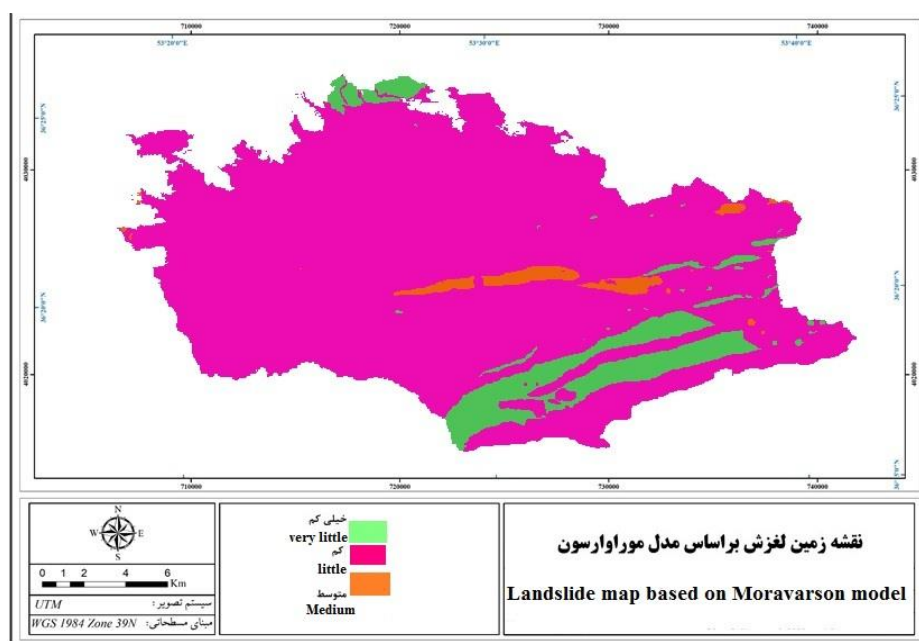
3- Input Layer

4- Hidden Layer

5- Output Layer

6- Broyden-Fletcher-Goldfarb-Shanno algorithm

7- Root Mean Square of the Error



شکل ۲- نقشه زمین لغزش براساس مدل موروارسون در حوزه جنگلی
Figure 2- Landslide map based on Mora -Vahrson model in forest

با استفاده از رابطه پتانسیل زمین لغزش، حوزه جنگلی در سه کلاس ۲، ۳ و ۴ طبقه (به ترتیب خیلی کم، کم و متوسط) قرار دارد. پهنه کلاس ۲ با زمین لغزشی به مساحت ۷۲۸۳ هکتار (۲۴ درصد) و پهنه کلاس ۳، ۲۳۲۵۴ هکتار (۷۶ درصد) و پهنه کلاس ۴ مقداری (ناچیز) از زیر حوزه ها را در برگرفته است (جدول ۳).

جدول ۳- مساحت و درصد طبقات مور-وارسون در زیرحوزه

Table 3. Area and percentage of Mora -Vahrson classes in sub-basin

کلاس	توصیف	مساحت (هکتار)	درصد
Class	Description	Area (ha)	(%)
2	خیلی کم Very low	7283	24
3	کم Low	23254	76
4	متوسط Medium	65	ناچیز

مدل سازی خطر وقوع زمین لغزش با استفاده الگوریتم جنگل تصادفی

جدول ۴ نتایج ارزیابی حاصل از الگوریتم جنگل تصادفی با تعداد برآوردکننده های مختلف را نشان می دهد. با بررسی میزان تغییرات مربعات خطا در نقاط تقریبی ۳۰ نقطه برای

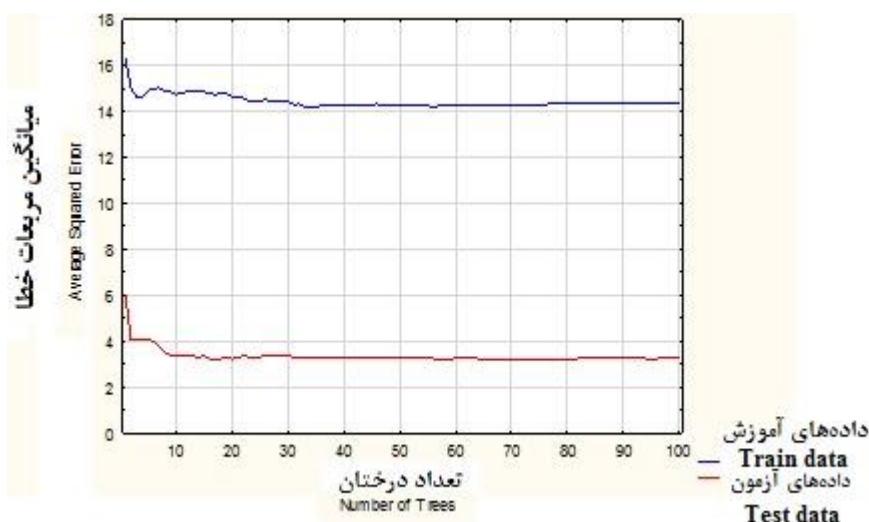
جدول ۴- ارزیابی داده های آزمون با روش جنگل تصادفی

Table 4. Evaluation of test data by random forest method

الگوریتم ناپارامتریک	ضریب تبیین	مجنور میانگین مربعات خطا	درصد مجنور میانگین مربعات خطا	اریبی
Nonparametric algorithm	R^2	RMSE	RMSE%	BIAS%
جنگل تصادفی (RF)	0.40	10.47	29.86	10.98
Random forest				

تعداد درخت بهینه انتخاب شد. با بررسی شکل ۳ مشخص شد میزان تغییرات مربعات خطا در نقطه تقریبی ۳۰ درخت برای داده های آموزش و آزمون روند ثابتی پیدا می کند، به همین دلیل تعداد مذکور درخت به عنوان تعداد درخت بهینه انتخاب شد.

شکل ۳ نمودار میانگین مربعات خطای داده های آموزش و آزمون در برابر تعداد درختان برای مدل سازی خطر وقوع زمین لغزش که با تفسیر میانگین مربعات خطای داده های آموزش و آزمون حاصل شد، را نشان می دهد. جایی که با افزایش تعداد درختان تغییراتی در کاهش میزان میانگین مربعات خطای آموزش و آزمون مشاهده نمی شود، به عنوان



شکل ۳- نمودار روند تغییرات میانگین مربعات خطا در برآورد خطر وقوع زمین لغزش برای تعداد درختان مختلف
Figure 3. Chart of changes in mean square error in estimating the risk of landslides for the number of different trees

مدل سازی خطر وقوع زمین لغزش با استفاده الگوریتم CART
داد؛ مقادیر ضریب تبیین، درصد مجذور میانگین مربعات خطا و ارزیابی به ترتیب ۰/۴۸، ۲۰/۸۲٪ و ۵/۸۹٪ به دست آمد (جدول ۵).
نتایج مدل سازی با الگوریتم ناپارامتریک CART نشان

جدول ۵- ارزیابی داده های آزمون با روش CART

Table 5. Evaluation of test data by CART method

ارایی BIAS%	درصد مجذور میانگین مربعات خطا RMSE	مجدور میانگین مربعات خطا RMSE	ضریب تبیین R^2	الگوریتم ناپارامتریک Nonparametric algorithm
5.86	20.82	2.02	0.48	الگوریتم درخت تصمیم CART

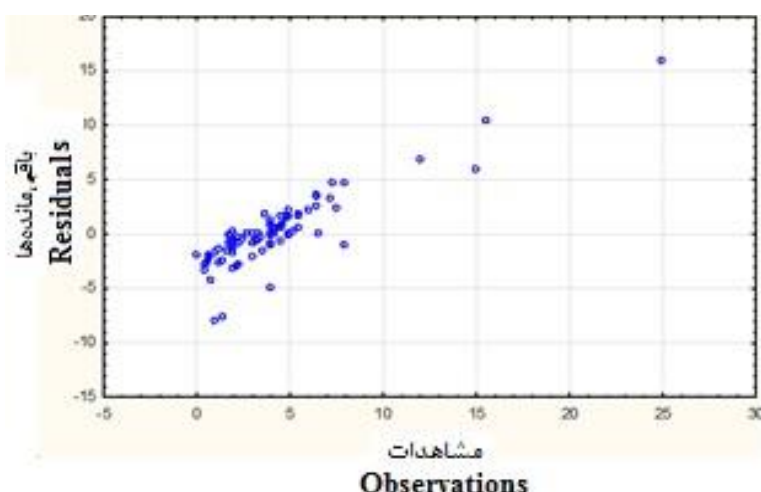
بر اساس نتایج حاصل از بررسی متغیرهای مستقل مؤثر موجود در مدل مشخص شد، مشخصه شدت بارندگی (TP) (۱/۰۰) بیشترین تأثیر را در مدل سازی دارد. از فاکتور با اهمیت دیگر ارتفاع، با درجه به ترتیب (SL=۰/۸۰).

بر اساس نتایج حاصل از بررسی متغیرهای مستقل مؤثر موجود در مدل مشخص شد، مشخصه شدت بارندگی (TP) (۱/۰۰) بیشترین تأثیر را در مدل سازی دارد. از فاکتور با اهمیت دیگر ارتفاع، با درجه به ترتیب (SL=۰/۸۰).

جدول ۶- ترتیب درجه اهمیت متغیرهای مستقل موجود در مدل

Table 6. Order of importance of independent variables in the model

درجه اهمیت Degree of importance	رتبه متغیر Variable rank	شاخص Index
1.00	100	شدت بارندگی=TP Rainfall intensity
0.80	80	حساسیت سنگ شناسی=SL Lithological sensitivity
0.76	76	پستی و بلندی نسبی=SR Relative height and height
0.25	25	شدت لرزه ای=TS Vibration intensity



شکل ۴- پراکنش ابر نقاط داده های پیش بینی شده در مقابل مشاهدات در مدل درخت تصمیم
Figure 4. Cloud distribution of predicted data points versus observations in the decision tree model

پرسپترون چند لایه با تعداد ۵ لایه ورودی، ۷ لایه پنهان و یک لایه خروجی) و الگوریتم BFGS 5 (بهینه سازی از روش شبه نیوتنی) دارای دقت بیشتری است (جدول ۷). در شکل ۵ خروجی مشاهدات در مقابل پیش بینی شده توسط شبکه که راهی دیگر در ارزیابی مدل است؛ مورد مقایسه قرار گرفت.

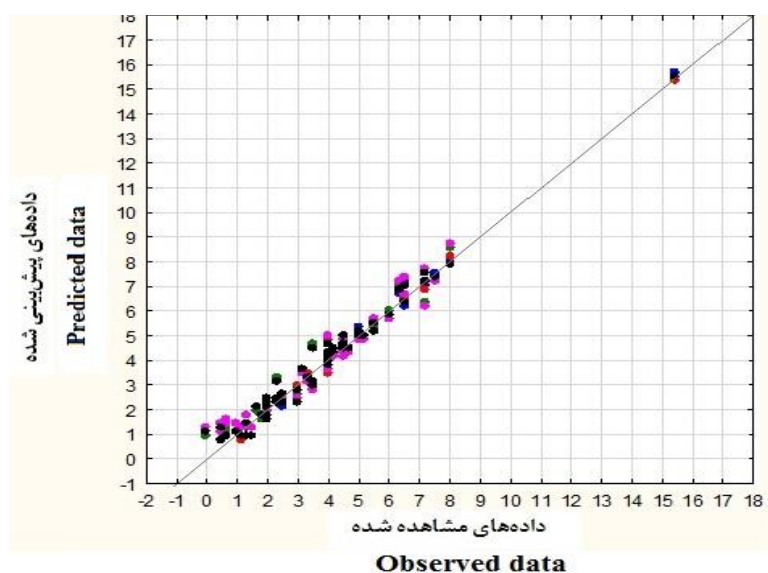
مدل سازی خطر وقوع زمین لغزش با استفاده الگوریتم ANN

نتایج مدل سازی با الگوریتم ناپارامتریک ANN نشان داد؛ مقادیر ضریب تبیین و درصد مجذور میانگین مربعات خطا و اریبی در مدل ANN با شبکه 5-7-1 MLP (شبکه عصبی

جدول ۷- ارزیابی داده های آزمون با مدل ANN

Table 7. Evaluation of test data with ANN model

اریبی BIAS%	درصد مجذور میانگین مربعات خطا %RMSE	مجذور میانگین مربعات خطا RMSE	ضریب تبیین R^2	شبکه Network	الگوریتم Algorithm	الگوریتم ناپارامتریک Nonparametric algorithm
-3.22	12.28	2.53	0.99	MLP 5-7-1	BFGS 5	شبکه عصبی مصنوعی ANN



شکل ۵- پراکنش ابر نقاط داده های پیش بینی شده در مقابل مشاهدات در مدل شبکه عصبی مصنوعی
Figure 5. Cloud distribution of predicted data points against observations in artificial neural network model

در مدل سازی نسبت به سایر روش های مورد استفاده دارای دقت بالاتری است.

نتایج حاصل از اعتبارسنجی در این مرحله طبق جدول ۸ با توجه به ضریب تبیین نشان داد که برای مدل سازی خطر وقوع زمین لغزش مدل شبکه عصبی مصنوعی با ($R^2=0.99$)

جدول ۸- نتایج اعتبار سنجی مدل سازی خطر وقوع زمین لغزش با استفاده از داده کاوی ناپارامتریک

Table 8. Validation results of landslide risk modeling using non-parametric data mining

مشخصه Characteristic	مدل Model		جنگل تصادفی RF		درخت تصمیم CART		شبکه عصبی مصنوعی ANN	
خطر وقوع زمین لغزش Landslide Risk	ضریب تیبیین R^2	میانگین مربعات خطا RMSE	درصد مجذور میانگین مربعات خطا RMSE%	اریبی BIAS%	ضریب تیبیین R^2	میانگین مربعات خطا RMSE	درصد مجذور میانگین مربعات خطا RMSE%	اریبی BIAS%
	0.40	10.47	29.86	10.98	0.48	2.02	20.82	5.89
							0.99	2.53
							12.28	-3.33

از عوامل مهم تخریب مناطق جنگلی می توان لغزش خاک اشاره کرد. زمین لغزش به عنوان یکی از چالش های زیست محیطی شناخته می شود که منجر به تخریب زمین، کاهش حاصلخیزی و آسیب های قابل توجهی به بوم سازگان می شود (۲۷). به دلیل اینکه زمین لغزش ها معمولاً سیستم های پیچیده ای هستند و پیش بینی حساسیت وقوع آنها نیاز به داده های مختلف طبیعی و انسانی از جمله ژئومورفولوژی، زمین شناسی، هیدرولوژیکی، کاربری اراضی و سایر داده های مربوط به عوامل محیطی دارد؛ لذا روش های مختلفی برای مدل سازی خطر وقوع زمین لغزش وجود دارد. انتخاب یک مدل ارزیابی با ظرفیت پیش بینی بالا، به عنوان اولین مرحله ارزیابی حساسیت زمین لغزش است. در پژوهش پیش رو ضرورت استفاده از سه روش داده کاوی ناپارامتریک جنگل تصادفی، درخت تصمیم و شبکه عصبی مصنوعی در حوزه جنگل های هیرکانی مورد توجه قرار گرفت. اراضی بسیار پرخطر و پرخطر، اراضی هستند که نباید تخریب شوند و کارهای غیراصولی در آنها انجام شود؛ کارهایی از قبیل تخریب پوشش گیاهی بر اثر تبدیل اراضی مرتعی به زراعی و کاهش کمی و کیفی پوشش گیاهی بر اثر چرای بی رویه دام ها، کشت دیم در اراضی شیب دار، رعایت نشدن تناوب زراعی و احداث و دستکاری های غیراصولی جاده یا تغییر مسیر جاده؛ بلکه باید اقداماتی مانند تبدیل کاهش شیب دامنه ارتفاعات مشرف به مسیر احداث جاده، کاهش گیاهان مناسب، ایجاد سیستم زهکشی مناسب، حذف سازند ریزدانه و برداشت آن از سطح دامنه انجام شود (۱۷). نتایج پهنه بندی حوزه در سه کلاس خیلی کم، کم و متوسط نشان داد؛ بیشترین توزیع مساحت زمین لغزش متعلق به طبقه با خطر کم (۲۳۲۵۴ هکتار) (۷۶ درصد) است. نتیجه حاضر شده با پژوهش احسانی فر و همکاران (۱۷) هم راستا است. بر اساس پژوهش ایشان نیز ۷۵/۴۹ درصد از مساحت منطقه مورد مطالعه در محدوده با خطر بسیار کم قرار گرفت. بارش یکی از عامل های مؤثر بر وقوع زمین لغزش ها است. افزایش مقدار بارندگی معمولاً سبب افزایش وقوع زمین لغزش می شود (۱۲). نتیجه ی به دست آمده از بررسی اهمیت نسبی متغیرها براساس الگوریتم درخت تصمیم نشان داد؛ شاخص شدت بارندگی بیشترین اهمیت را در مدل سازی وقوع زمین لغزش داشته است که با نتایج رفیعی ساردویی (۲۴) در تضاد است. ایشان اذعان داشت بارندگی تأثیر ناچیزی در پهنه بندی حساسیت

زمین لغزش دارد. منطقه مورد پژوهش ایشان از لحاظ زمین شناسی سخت بود، بنابراین امکان زمین لغزش در آن کمتر بود. نتایج حاصل از اعتبارسنجی با توجه به ضریب تبیین نشان داد که برای مدل سازی خطر وقوع زمین لغزش از بین سه الگوریتم جنگل تصادفی، درخت تصمیم و شبکه عصبی مصنوعی، مدل شبکه عصبی مصنوعی با تعداد ۵ لایه ورودی، ۷ لایه پنهان و (۱۲/۲۸، $R^2=0.99$) (RMSE%) در مدل سازی نسبت به سایر روش های مورد استفاده دارای دقت بالاتری است. نتیجه حاصل شده هم راستا با پژوهش آرام و همکاران (۳) است. در مقابل با نتیجه پژوهش حیدری و همکاران (۱۳)، شیرانی و نادری (۲۹) در تضاد است. ارزیابی دقت طبقه بندی و اعتبارسنجی مدل ها در پژوهش ایشان نشان داد مدل جنگل تصادفی ضمن داشتن آستانه ی طبقه های مناسب در طبقه بندی، اولویت بیشتری برای پهنه بندی و پیش بینی رخداد زمین لغزش ها دارد. همچنین با افزایش ارتفاع میزان بارش و در نتیجه توان اشباع شدن خاک افزایش یافته است، که منجر به افزایش زمین لغزش ها می شود. به طور کلی دست کاری در طبیعت، بی توجه به شرایط منطقه موجب تحریک شدن و وقوع زمین لغزش خواهد شد، در منطقه پژوهش حاضر نیز با توجه به اجرای طرح تنفس در بسیاری از مناطق باعث تشدید قاچاق چوب شده است، که در نتیجه تخریب خاک را در پی دارد. در تحلیل رگرسیون چندگانه در صورت معرفی متغیرهای مستقل، همبستگی مورد اشاره موجب افزایش فاکتور تورم واریانس می شود. به طور معمول در شبکه عصبی، هم خطی چندگانه و خودهمبستگی بین متغیرها مؤثر نیست و علاوه بر آن، نوع روابط بین هریک از متغیرهای ورودی و خروجی (خطی و یا غیرخطی) نیز تصحیح می شود. به همین دلیل دو متغیر قطر و ارتفاع در پژوهش حاضر به عنوان متغیرهای ورودی برای مدل شبکه عصبی مؤثر واقع شدند (۲۰). در نتیجه با توجه به نتیجه به دست آمده مدل مربوط (شبکه عصبی مصنوعی) دارای دقت محاسباتی به نسبت بالایی برای مدل سازی خطر وقوع زمین لغزش نسبت به دو مدل استفاده شده دیگر در پژوهش حاضر است؛ اما چون میزان میانگین مربعات خطای مدل انتخابی به عنوان متغیر ورودی دارای حداقل اختلاف نسبت به داده های آموزش، اعتبار و نیز حداقل مقدار نسبت به مدل دیگر است، دقت بیشتری برای پیش بینی پاسخ مورد هدف در پژوهش حاضر دارد (۳۳).

نتیجه گیری کلی

زمین لغزش به صورت سیستمی عمل می کند؛ در نتیجه همه عوامل در وقوع چنین پدیده هایی نقش دارند. تهیه نقشه های حساسیت خطر زمین لغزش یکی از گام های اساسی در جهت شناخت مناطق ناپایدار است تا بتوان با استفاده از آن برنامه ریزی های لازم جهت کاهش خسارات ناشی از این پدیده را انجام داد. با توجه به خروجی های حاصل از پژوهش حاضر، ارزیابی الگوریتم های جنگل تصادفی، درخت تصمیم و شبکه عصبی مصنوعی نشان داد؛ نتایج مدل سازی با مدل شبکه عصبی مصنوعی از میان متغیرهای مورد بررسی، به ترتیب شدت بارندگی، حساسیت لیتولوژیک، پستی و بلندی نسبی و شدت لرزه ای بیشترین تأثیر را در وقوع زمین لغزش حوزه مطالعاتی نشان داده است. در این راستا پیشنهاد می شود به منظور ارزیابی با دقت بیشتر از تصاویر ماهواره ای در تلفیق

از داده های زمینی استفاده شود. همچنین دیگر الگوریتم ها در مدل سازی خطر وقوع زمین لغزش برای منطقه مورد پژوهش استفاده شود و با نتایج پژوهش حاضر مورد مقایسه قرار گیرد. پیشنهاد می شود در صورت دسترسی به اطلاعات علاوه بر شدت بارندگی، میزان بارش هم در نظر گرفته شود.

تشکر و قدردانی

مقاله پیش رو حاصل کار در دوره پسداکتری و در قالب طرحی با عنوان مطالعات نیمه تفصیلی مدیریت پایدار منابع طبیعی حوزه های جنگل شمال کشور در سال ۱۴۰۱ است که با حمایت دانشگاه علوم کشاورزی و منابع طبیعی ساری اجرا شده است؛ بدین وسیله از معاونت پژوهشی دانشگاه تشکر و قدردانی می شود.

منابع

1. Abella, E.A.C and C.J. Van Westen. 2008. Qualitative landslide susceptibility assessment by multicriteria analysis; a case study from San Antonidel Sur, Guantanamo, Cuba (in GIS technology and models for assessing landslide hazard and risk). *Geomorphology*, 94: 435-466.
2. Alqadhi, S., J. Mallick, S. Talukdar, A.A. Bindajam, N. Van Hong and T.K. Saha. 2022. Selecting optimal conditioning parameters for landslide susceptibility: An experimental research on Aqabat Al-Sulbat, Saudi Arabia. *Environmental Science and Pollution Research*, 29(3): 3743-3762.
3. Aram, A., M.R. Dalalian, S. Saedi, O. Rafieian and S. Darbandi. 2022. Evaluation of the efficiency of artificial intelligence and bivariate statistical models in determining landslide prone areas in West Azerbaijan. *Journal of Water and Soil Resources Conservation*, 11(4): 63-74 (In Persian).
4. Breiman, L. 2001. Random Forests. *Machine Learning*: 45(1): 5-32.
5. Breiman, L., J. Friedman, C.J. Stone and R.A. Olshen. 1984. *Classification and regression trees*, 1st edn, CRC press, Pub. Location New York, New York, 368 pp.
6. Cred Crunch Newsletter, Issue No. 68 (September 2022) - Natural Hazards & Disasters An overview of the first half of 2022.
7. Das, S., S. Sarkar and D.P. Kanungo. 2022. GIS-based landslide susceptibility zonation mapping using the analytic hierarchy process (AHP) method in parts of Kalimpong Region of Darjeeling Himalaya. *Environmental Monitoring and Assessment*, 194(3): 1-28.
8. Ehsanifar, A., A. Kavyan, K. Soleymani and H. Aghbari. 2011. Landslide risk zoning using overlap index, case study: Abkhaz area of Tajan. The 7th National Conference on Watershed Science and Engineering of Iran, 12-1, Esfahan, Iran (In Persian).
9. Frattini, P., G. Crosta and A. Carrara. 2010. Techniques for evaluating the performance of landslide susceptibility models. *Engineering Geology*, 111(1): 62-72.
10. Froude, M.J. and D. Petley. 2018. Global fatal landslide occurrence from 2004 to 2016. *Nat Hazards Earth. Natural Hazards and Earth System Sciences*, 18(8): 2161-2181.
11. Hagan, M.T., H.B. Demuth and M.H. Beale. 1996. *Neural Network design*. 1st edn, PWS press, United States of America, 800 pp.
12. Hanifinia, A., H. Nazarnejad, S. Najafi and A. Kornejady. 2020. Prioritization of Effective Factors on Landslide Occurrence and Mapping of its Sensitivity in CherkAbad Watershed, Urmia Using Shannon Entropy Model. *Watershed Management Research*, 33(4): 30-46 (In Persian).
13. Heydari, N., M. Habibnejad, A. Kavian and H.R. Pourghasemi. 2020. Landslide susceptibility modelling using the random forest machine learning algorithm in the Watershed of Rais-Ali Delvari Reservoir. *Watershed Management Research*, 33(1): 2-13 (In Persian).
14. Lawrence, R.L. and A. Wright. 2001. Rule-based classification systems using classification and regression tree (CART) analysis. *Photogrammetric Engineering and Remote Sensing*, 67(10): 1137-1142.
15. Liaw, A. and M. Wiener. 2002. Classification and regression by randomForest. *R news*, 2(3): 18-22.
16. Lin, X.S. and J. Xu. 2007. The Study of the Complexity of Landslide Hazard. *Res. Soil Water Conservation*, 05: 359-363.
17. Madadi, A., E. Piroozi and M. Faal Naziri. 2021. A Comparative Evaluation of MABAC and CODAS Multi-Criteria Decision Algorithms in Landslide Risk Zoning (Case Study: Kowsar County). *Geography and Environmental Planning*, 31(4): 1-24 (In Persian).

18. Mahnaj, M.B. 1998. Introduction to artificial neural network, 1st Edition, Professor Hesabi Publication, 642 p.
19. Mora, S. and W.G. Vahrson. 1992. Determinación "a priori" de la amenaza de deslizamientos utilizando indicadores morfodinámicos. In Álzate, J. B. (editor), Memoria del Primer Simposio Internacional Sobre Sensores Remotos y Sistemas de Información Geográfica (SIG) Para el Estudio de Riesgos Naturales: Bogotá, Colombia, 259-273.
20. Nazariani, N., A. Fallah, M. Imani Rastabi and F. Bakhshi. 2022. Modeling the Comparison volume of Pure and Mixed Stands of Beech Trees Using Non-parametric Algorithms in the Educational-research Forest of Darabkola. Iranian Journal of Forest and Poplar Research, Online publication (In Persian).
21. Pham, B.T., T.V. Phong, T. Nguyen-Thoi, K. Parial, K.S. Singh, H.B. Ly, K.T. Nguyen, L.S. Ho, H.V. Le and I. Prakash. 2022. Ensemble modeling of landslide susceptibility using random subspace learner and different decision tree classifiers. Geocarto International, 37(3): 735-757.
22. Pourghasemi, H.R., M. Mohammady and B. Pradhan. 2012. Landslide susceptibility mapping using index of entropy and conditional probability models in GIS: Safarood Basin, Iran. Catena, 97: 71-84.
23. Pourghasemi, H.R. and O. Rahmati. 2018. Prediction of the landslide susceptibility: Which algorithm, which precision. Catena, 162: 177-192.
24. Rafiei Sardooi, E. 2022. Landslide Susceptibility Simulation Using Data Mining Models in Rabor Area. Watershed Management Research Journal, 35(2):101-118 (In Persian).
25. Raghuvanshi, T.K., J. Ibrahim and D. Ayalew. 2014. Slope stability susceptibility evaluation parameter (SSEP) rating scheme-an approach for landslide hazard zonation. Journal of African Earth Sciences, 99: 595-612.
26. Reis, L.P., A.L. Souza, P.C.M. Reis, L. Mazzei, C.P.B. Soares, C.M.M.E. Torres, L.F. Silva, A.R. Ruschel and L.J.S. Rêgo. 2018. Estimation of mortality and survival of individual trees after harvesting wood using artificial neural networks in the amazon rain forest. Ecological Engineering, 112: 140-147.
27. Rezaee Banafshe, M., H. Rostamzadei and B. Feyzizadeh. 2010. Investigating and evaluating the changing process of forest levels using remote sensing and GIS (case study of Arsbaran forests 1987-2005). Geographical Research, 40(1): 143-159 (In Persian).
28. Shariat Jafari, M. 1996. Landslide (basics and principles of stability in natural slopes). Saze Publications, 205 p (In Persian).
29. Shirani, K. and R. Naderi Samani. 2022. Determination of Effective factors and Assessment of Landslide Susceptibility Using Random Forest and Artificial Neural Network in Doab Samsami Region, Chaharmahal and Bakhtiari Province. Watershed Management Research Journal, 35(1): 40-60 (In Persian).
30. Steger, S., A. Brenning, R. Bell and T. Glade. 2017. The influence of systematically incomplete shallow landslide inventories on statistical susceptibility models and suggestions for improvements. Landslides, 14: 1767-1781.
31. Strobl, R.O. and F. Forte. 2007. Artificial neural network exploration of the influential factors in drainage network derivation. Hydrological Processes, 21: 2965-2978.
32. Tien Bui, D., H. Moayedi, M. Gëor, A. Jaafari and F. L. Kok. 2019. Predicting slope stability failure through machine learning paradigms. ISPRS International Journal of Geo-Information, 8(9): 395.
33. Tiryaki, S. and A. Aydin. 2014. An artificial neural network model for predicting compression strength of heat treated woods and comparison with a multiple linear regression model. Construction and Building Materials, 62: 102-108.
34. Tooke, T.R., N.C. Coops, N.R. Goodwin and J.A. Voogt. 2009. Extracting urban vegetation characteristics using spectral mixture analysis and decision tree classifications. Remote Sensing of Environment, 113: 398-407.
35. Varnes, D.J. 1958. Landslide types and processes. Landslides and engineering practice, 24: 20-47.

Landslide Risk Modeling using Data Mining in Hyrcanian Forests

Nastaran Nazariani¹ and Asghar Fallah²

1- Postdoctoral Researcher in Forestry, Faculty of Natural Resources, Sari of Agricultural Sciences and Natural Resources University, Mazandaran, Iran, (Corresponding author: Nazariani69@yahoo.com)

2- Professor, Department of Forestry, Faculty of Natural Resources, Sari of Agricultural Sciences and Natural Resources University, Mazandaran, Iran

Received: 24 October, 2022 Accepted: 29 November, 2022

Extended Abstract

Introduction and Objective: In recent years, the increases in population and expansion of settlements in hazardous areas have greatly increased the impact of natural disasters in industrialized and developing countries. Landslide risk zoning helps to identify strategic points and geographically prone critical areas. Therefore, measures for rapid, safe mitigation and strategic planning for the future are important. In fact, landslide risk assessment may be a suitable and cost-effective help for land use planning, so in this regard, the aim of the current research is to model the risk of landslides using non-parametric data mining in watersheds of Hyrcanian forests.

Material and Methods: For this purpose, the map of landslide risk according to the Mora and vahrsen method with the effect of factors affecting the occurrence of landslides including topographical factors, hydrological and climatic factors, geological factors, land cover factors, human factors, hydrographic network from digital elevation model, geological map It was prepared using the map of the Geological Organization of the country, and to obtain the elevation index and the relative height, first, using the curve of height levels taken from the 1:25000 topographic maps of the region, the map of the elevation classes was prepared. After that, the study area was divided into one-square-kilometer grids, and maps with the lowest and highest elevations in one-square-kilometer grids were obtained; in the last step, by subtracting these two maps, a map was obtained whose information shows the value of the postal index and relative height. Monthly rainfall was also used to obtain the soil moisture index. Finally, algorithms of three random forest models, an artificial neural network, and a decision tree algorithm were used to model the risk of landslides in the STATISTICA 12.0 software environment.

Results: According to the results, the highest distribution of landslide areas belongs to the low-risk class (76%). TP, SL, SR, and TS variables were considered the essential factors of landslide occurrence based on their importance. The results of validation with three algorithms of RF, CART, and ANN showed; According to the coefficient of explanation obtained in modeling the risk of landslides, the artificial neural network model with ($R^2=0.99$) is more accurate than other methods.

Conclusion: The results of the present study showed that data mining methods have a high capability in predicting the risk of landslides. Therefore, the use of the mentioned methods can be considered in reducing the risks associated with landslides and planning for land use.

Keywords: Artificial neural network, Decision tree, Hyrcanian forests, Modeling, Regression model, Watershed