



دوره‌ی ۳۳، شماره‌ی ۱، شماره‌ی پیاپی ۱۲۶، بهار ۱۳۹۹، صفحه‌های ۱۳ - ۲
شناسه‌ی دیجیتال: 10.22092/wmej.2019.126288.1219

پژوهش‌های آبخیزداری

مدل‌سازی حساسیت زمین‌لغزش با الگوریتم یادگیری ماشین جنگل تصادفی در آبخیز سد رئیسعلی دلواری

ناصر حیدری

دکترای علوم و مهندسی آبخیزداری، دانشکده‌ی منابع طبیعی، دانشگاه علوم و کشاورزی ساری، ساری، ایران

محمود حبیب‌نژاد

استاد گروه علوم و مهندسی آبخیزداری، دانشکده‌ی منابع طبیعی، دانشگاه علوم و کشاورزی ساری، ساری، ایران

عطا ... کاویان

استاد گروه علوم و مهندسی آبخیزداری، دانشکده‌ی منابع طبیعی، دانشگاه علوم و کشاورزی ساری، ساری، ایران

حمیدرضا پورقاسمی*

(نویسنده‌ی مسئول)* دانشیار بخش مهندسی منابع طبیعی و محیط زیست، دانشکده‌ی کشاورزی، دانشگاه شیراز، شیراز، ایران

*رایانامه‌ی نویسنده‌ی مسئول: hr.pourghasemi@shirazu.ac.ir

تاریخ دریافت: ۲۶ اردیبهشت ۱۳۹۸ تاریخ پذیرش: ۱۲ شهریور ۱۳۹۸

چکیده

هدف از این پژوهش مدل‌سازی کردن مکانی حساسیت زمین‌لغزش با الگوریتم یادگیری ماشین جنگل تصادفی و اولویت‌بندی کردن عامل‌های موثر بر وقوع آن در آبخیز سد رئیسعلی دلواری است. نقشه‌ی پراکنش زمین‌لغزش‌های منطقه با بازدیدهای صحرایی و بانک اطلاعات زمین‌لغزش‌های کشور تهیه شد. در مجموع از ۲۷۹ زمین‌لغزش شناخته‌شده ۷۰٪ (۱۹۵) آن برای مدل‌سازی و ۳۰٪ (۸۴) مانده برای ارزیابی مدل به‌کاربرده شد. لایه‌های اطلاعاتی ارتفاع، جهت شیب، درجه‌ی شیب، انحنای سطح، انحنای نیمرخ، شاخص رطوبت پستی‌وبلندی، فاصله از شبکه‌ی آبراه، تراکم زه‌کشی، فاصله از گسل، فاصله از جاده، زمین‌شناسی و شاخص تفاضلی پوشش گیاهی به‌نکار شده انتخاب شد. مدل جنگل تصادفی بر اساس ارتباط بین متغیر وابسته (زمین‌لغزش‌ها) و متغیرهای مستقل (عامل‌های موثر) در نرم‌افزار R و با بسته‌ی نرم‌افزاری Random Forest اجرا، و نقشه‌ی حساسیت زمین‌لغزش تهیه شد. مدل با به‌کاربردن منحنی تشخیص عمل کرد نسبی و ۳۰٪ از داده‌های لغزشی به‌کاربرده‌نشده در فرآیند مدل‌سازی ارزیابی شد. نتایج ارزیابی نشان‌دهنده‌ی دقت عالی مدل جنگل تصادفی ۰/۹۸۳ (۹۸/۳٪) بود. اولویت‌بندی عامل‌های موثر اهمیت درجه‌ی شیب، ارتفاع، انحنای نیمرخ، فاصله از جاده و واحدهای سنگ‌شناسی را نشان داد. بنابراین به‌نظر می‌رسد که نقشه‌ی حساسیت زمین‌لغزش تهیه‌شده ممکن است نقش بسزایی در تصمیم‌گیری‌های مدیران برای آمایش کردن سرزمین و مدیریت کردن جامع آبخیز سد رئیسعلی دلواری داشته باشد.

واژگان کلیدی: آبخیز سد رئیسعلی دلواری، جنگل تصادفی، حساسیت زمین‌لغزش، میانگین کاهشی دقت

مقدمه

جهان و در ایران است که هر سال نقش زیادی در تخریب جاده‌های ارتباطی، باغ‌ها، مرتع‌ها، و مناطق مسکونی، هدررفت خاک و تولید حجم زیاد رسوب در آبخیزهای کشور دارد. تحقیقات بسیاری با الگوریتم جنگل تصادفی انجام شده است (یوسف و همکاران ۲۰۱۵؛ پورقاسمی و کرله ۲۰۱۶؛ چن و همکاران ۲۰۱۷؛ لاگومارسینو و همکاران ۲۰۱۷؛ پورقاسمی و رحمتی ۲۰۱۸؛ طلاب و همکاران ۲۰۱۸؛ پارک و کیس ۲۰۱۹). الگوریتم جنگل تصادفی یکی از روش‌های مدل‌سازی درخت تصمیم است (شتایی و همکاران ۲۰۱۱) که از تلفیق نتیجه‌های به‌دست آمده از درخت‌های مختلف، ارتباط میان رخداد زمین‌لغزش و عامل‌های محیطی را برآورد می‌کند. از طرف دیگر، این روش داده‌کاوی عامل‌های موثر را اولویت‌بندی، و اهمیت آن‌ها را تعیین می‌کند. دقت این روش در مقایسه با مدل‌های آماری و سنتی که نیاز به برخی فرضیه‌ها از جمله بهنجاربودن داده‌ها و طبقه‌بندی عامل‌ها دارند، بسیار بیش‌تر است، به‌طوری که بر کاربرد آن در علوم مختلف و تحقیقات دیگر نظیر فرسایش خندقی (رحمتی و همکاران ۲۰۱۷؛ گروسی و همکاران ۲۰۱۸؛ عرب‌عامری و همکاران ۲۰۱۸؛ امیری و همکاران ۲۰۱۹)، پهنه‌بندی سیل (داوودی مقدم و همکاران ۲۰۱۹) و بررسی احتمال یافت‌شدن آب زیرزمینی (ذبیحی و همکاران ۱۳۹۴؛ نقیبی و پورقاسمی ۲۰۱۵؛ رحمتی و همکاران ۲۰۱۶) افزوده است.

سد مخزنی رئیس‌علی دلواری در بخش شبانکاره در جنوب غربی ایران در استان بوشهر ایران روی رود شاپورساخته شده است. مهم‌ترین هدف‌های ساخت این سد آبیاری ۲۴ هزارهکتار از زمین‌های کشاورزی، تولید سالانه ۹۶ هزارمگاوات انرژی برق آبی، مهارکردن سیلاب، ایجاد اشتغال، جلوگیری از بروز خسارت سیل به زمین‌های کشاورزی، نخلستان‌ها و تخریب ساحل رودخانه‌ها بوده است. بنابراین حفاظت از مخزن سد در مقابل زمین‌لغزش و انتقال رسوب‌های آن به پایین‌دست و آبخیز آن از دیدگاه حفظ پوشش گیاهی و فرسایش خاک اهمیت بسزایی بین ساکنان منطقه دارد. هدف از این پژوهش تهیه‌کردن نقشه‌ی حساسیت زمین‌لغزش با روش یادگیری ماشین جنگل تصادفی و اولویت‌بندی کردن عامل‌های موثر بر وقوع آن در آبخیز سد رئیس‌علی دلواری است. تحقیقات نشان می‌دهد که کم‌تر تحقیقی مرتبط با پدیده‌ی زمین‌لغزش در آبخیز سدهای مخزنی انجام شده، و این خود نوآوری این

روی داده‌های ناگوار طبیعی از بزرگ‌ترین دشمنان طبیعی انسان‌ها است و هر ساله موجب کشته و مجروح شدن هزاران نفر و بی‌خانمان شدن میلیون‌ها انسان در کل دنیا می‌شود. حرکت‌های توده‌ی^۱ و زمین‌لغزش^۲ از جمله‌ی مهم‌ترین آنان است که بر زندگی انسان‌ها اثر می‌کند (فیض‌نیا و همکاران ۲۰۰۴). زمین‌لغزش ناشی از شرایط آب‌شناختی، زمین‌ریخت‌شناسی، و زمین‌شناسی است. این شرایط در کنار فرآیندهای زمین‌پویایی، کاربری زمین، پوشش گیاهی، مقدار و شدت بارندگی، فعالیت‌های انسانی و زلزله باعث وقوع و توسعه‌ی این پدیده می‌شود (شادفر و همکاران ۲۰۰۵). در سال‌های اخیر رشد روزافزون و بی‌رویه‌ی جمعیت، رشد سکونت‌گاه‌های شهری در جاهای ناپایدار و مستعد منجر به افزایش وقوع روی داده‌های ناگوار طبیعی و زمین‌لغزش در کشورهای توسعه‌یافته و توسعه‌یابنده شده است. روی داده‌های ناگوار طبیعی به دو شکل مستقیم و نامستقیم بر مسائل اقتصادی و اجتماعی جامعه‌ی انسانی تأثیر می‌گذارد و خسارت‌های آن در کشورهای توسعه‌یابنده بیش‌تر است (کوهورست و همکاران ۲۰۰۵). زمین‌لغزش سالانه در بسیاری از استان‌های کشور باعث بروز خسارت‌های اقتصادی به جاده‌ها، راه‌آهن، خط‌های ارتباطی و انتقال نیرو، نهرهای آبیاری و آب‌رسانی، تأسیسات معدنی، ساختارهای استخراج و پالایش نفت و گاز، شبکه‌ی زیرساخت‌های حیاتی داخل شهرها، کارخانه‌ها و مراکز صنعتی، سدها و دریاچه‌های مصنوعی و طبیعی، جنگل‌ها و مرتع‌ها و منابع طبیعی، مزرعه‌ها و منطقه‌های مسکونی شهری و روستایی می‌شود، یا آن‌ها را تهدید می‌کند. حدود ۱۷٪ از روی داده‌های ناگوار جهان در سال‌های ۱۹۰۳-۲۰۰۴ زمین‌لغزش بوده است، که بیش‌ترین خسارت اقتصادی آن در اروپا حدود ۱۷ میلیون دلار در سال محاسبه شد (کوهورست و همکاران ۲۰۰۵). در کشور ایران نیز پژوهش‌هایی در باره‌ی خسارت‌های ناشی از حرکت توده‌ی و زمین‌لغزش انجام شده است. بر اساس گزارش جمع‌بندی داده‌ها و خسارت‌های ۴۹۰۰ زمین‌لغزش از بانک اطلاعاتی تا پایان شهریورماه سال ۱۳۸۶، میزان خسارت زمین‌لغزش در حدود ۱۲۶۸۹۳ میلیارد ریال برآورد شد (کارگروه زمین‌لغزش ایران ۲۰۰۷). با توجه به آسیب جانی، خسارت مالی و تاثیر زیست‌محیطی، زمین‌لغزش از مهم‌ترین و پرخطرترین روی داده‌های ناگوار طبیعی در

1- Mass Movement
2- Landslide

پژوهش در مقایسه با سایر تحقیقات مرتبط است.

مواد و روش ها

منطقه‌ی بررسی شده

آبخیز رئیسعلی دلواری در مرز استان‌های فارس و بوشهر، در ۷۳ کیلومتری شمال شرقی بوشهر، ۶۰ کیلومتری شهربرازجان مرکز شهرستان دشتستان، و هشت کیلومتری روستای میلک، بین عرض‌های $29^{\circ} 36' 30''$ تا $29^{\circ} 49' 00''$ شمالی و طول‌های جغرافیایی $50^{\circ} 57' 30''$ تا $51^{\circ} 11' 30''$ شرقی است (شکل ۱). مساحت آن حدود ۲۶۰ کیلومتر مربع و ارتفاع کمینه و بیشینه ی آن به ترتیب ۱۳۲ و ۸۲۸ متر از تراز دریا است.

روش پژوهش

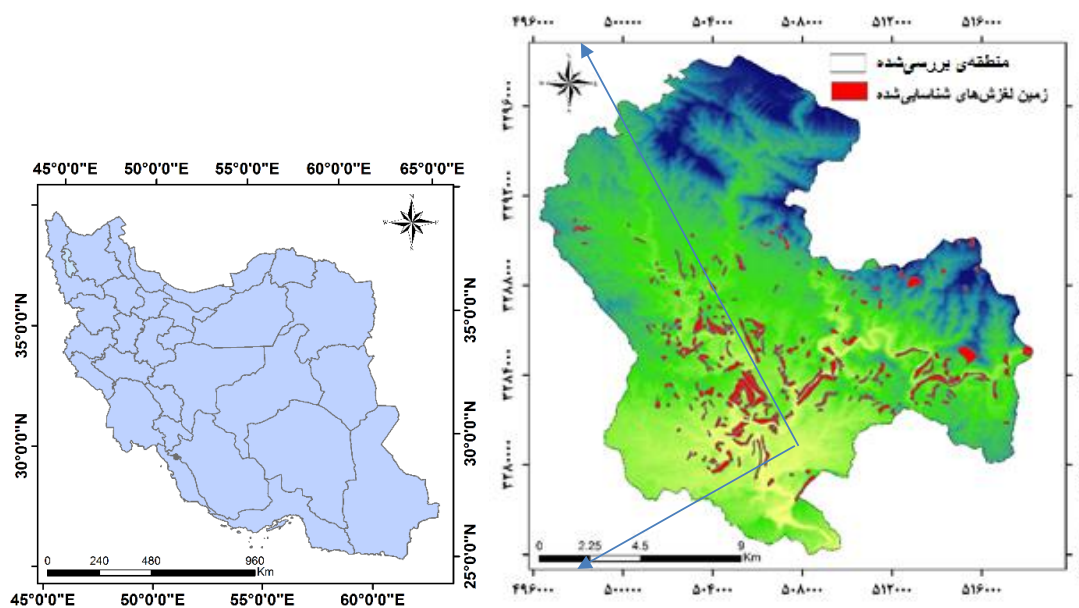
تهیه‌ی داده‌ها

نقشه‌ی پراکنش زمین لغزش‌های منطقه با پیمایش‌های زمینی و بر اساس داده‌های بانک اطلاعات زمین لغزش‌های

کشور (کارگروه زمین لغزش ایران ۲۰۰۷) تهیه شد (۲۷۹ زمین لغزش). نقشه‌ی پستی و بلندی آبخیز سد رئیسعلی دلواری در مقیاس ۱:۲۵۰۰۰ (سازمان نقشه برداری کشور ۱۳۸۸) و نقطه‌ها و خط‌های ارتفاعی از آن استخراج شد. نقشه‌ی مدل رقومی ارتفاع (DEM) ^۳، درجه ی شیب، جهت شیب، انحنای سطح ^۴، انحنای نیمرخ ^۵ و شاخص رطوبت پستی و بلندی ^۶ (TWI) از ترکیب نقشه‌های نقطه‌های ارتفاعی و خط‌های تراز به دست آمده از نقشه‌ی پایه‌ی پستی و بلندی، با نرم افزارهای ArcGIS نسخه‌ی ۱۰،۶،۱ و SAGA-GIS نسخه‌ی ۲،۳،۲ محاسبه و تهیه شد. نقشه‌ی شاخص رطوبت پستی و بلندی نیز با رابطه‌ی ۱ به دست آمد (مور و همکاران ۱۹۹۱):

$$TWI = \ln(A / \tan \beta) \quad \text{رابطه ی ۱}$$

A مساحت آبخیز بالادست و β مقدار شیب (درجه) است. این شاخص نقش مهمی در تغییر شرایط آب‌شناختی مانند جریان آب زیرزمینی و رطوبت خاک دارد.



شکل ۱- موقعیت محدوده‌ی بررسی شده بر نقشه‌ی ایران و استان فارس.

- 3- Digital Elevation Model (DEM)
- 4- Plan Curvature
- 5- Profile Curvature
- 6- Topographic Wetness Index (TWI)

در نرم افزار SPSS فراخوانده، با قالب dBase ذخیره شد، و در نرم افزار ArcGIS (نسخه ۱۰٫۶٫۱) فراخوانده شد تا نقشه‌ی پهنه‌بندی حساسیت زمین لغزش به دست آید. نقشه براساس شکستگی‌های طبیعی به طبقه‌های حساسیت کم، متوسط، زیاد و خیلی زیاد طبقه‌بندی شد.

ارزیابی روش‌های پهنه‌بندی

حدود ۳۰٪ از زمین لغزش‌های رخ داده در منطقه به تصادف انتخاب (ژو و همکاران ۲۰۱۲؛ پورقاسمی و رحمتی ۲۰۱۸) و برای ارزیابی کردن روش‌های پهنه‌بندی به کار برده شد. برای ارزیابی کردن، نقشه‌ی منحنی تشخیص عمل کرد نسبی (ROC)^۸ به کار برده شد (نفلسلی‌اگلو و همکاران ۲۰۰۸). منحنی ROC نمایشی ترسیمی از موازنه بین میزان خطای منفی و مثبت برای هر مقدار احتمالی از برش‌ها است. سطح زیر منحنی ROC بیانگر مقدار پیش‌بینی‌شده‌ی سامانه از راه توصیف توانایی آن در تخمین زدن درست وقایع رخ داده (رخداد زمین لغزش) و رخ‌نداده‌ی (رخ‌ندادن زمین لغزش) آن است. دل‌خواه‌ترین مدل، بیش‌ترین سطح زیر منحنی را می‌دهد، و مقدار آن از ۰/۵ تا ۱ است. هر چه سطح زیر منحنی به یک نزدیک‌تر باشد، بیانگر بهترین دقت از نقشه‌ی پهنه‌بندی تهیه‌شده است. طبقه‌بندی میزان سطح زیر منحنی با درجه‌های عالی (۱-۰/۹)، خیلی خوب (۰/۹-۰/۸)، خوب (۰/۸-۰/۷)، متوسط (۰/۷-۰/۶) و ضعیف (۰/۶-۰/۵) است (یسیلناکار ۲۰۰۵).

نتایج

بررسی هم‌خطی بین متغیرها

برای بررسی هم‌خطی بین متغیرها دو شاخص آماری عامل تورم پراش (VIF)^۹ و تاب‌آوری (Tolerance) به کار برده می‌شود. هنگامی که عامل تورم پراش بزرگ‌تر از ۵ و تاب‌آوری کوچک‌تر از ۰/۱ باشد بین متغیرهای مستقل هم‌خطی هست (چن و همکاران ۲۰۱۷). بیش‌ترین مقدار عامل تورم پراش و کم‌ترین مقدار تاب‌آوری به ترتیب ۴/۸۹۷ و ۰/۲۰۴ بود، بنابراین هیچ‌گونه هم‌خطی بین متغیرهای مستقل نیست (جدول ۱).

نقشه‌ی زمین‌شناسی منطقه در مقیاس ۱:۱۰۰۰۰۰ تهیه و رقومی کرده شد. نقشه‌ی گسل از نقشه‌ی زمین‌شناسی استخراج، و نقشه‌ی فاصله از گسل تهیه شد. نقشه‌ی شبکه‌ی آبراه‌ها، تراکم زه‌کشی و جاده‌ها نیز از نقشه‌ی پستی و بلندی استخراج، و نقشه‌ی فاصله و تراکم از آن‌ها تهیه شد. شاخص پوشش گیاهی تفاضلی بهنجار شده (NDVI)^۷ از تصویرهای ماهواره‌ی لندست ۷ (ETM+) سال ۲۰۱۸ و با رابطه‌ی ۲ ایجاد شد. عامل‌های موثر در این پژوهش بر اساس شرایط منطقه، داده‌های دسترس و منابع گسترده (یسیلناکار ۲۰۰۵؛ یوسف و همکاران ۲۰۱۵؛ پورقاسمی و کرله ۲۰۱۶؛ چن و همکاران ۲۰۱۷) انتخاب، و برای مدل‌سازی زمین لغزش به کار برده شد.

رابطه 2

$$NDVI = (NIR - IR) / (NIR + IR)$$

IR باند قرمز و NIR باند مادون قرمز نزدیک است.

مدل جنگل تصادفی

این الگوریتم که الحاقی از درخت و ایازی و طبقه‌بندی است (بریمن ۲۰۰۱) مبتنی بر دسته‌بندی از درخت‌های تصمیم، و یکی از بهترین الگوریتم‌های کنونی یادگیری است. مدل پیش‌بینی‌کننده‌ی RF بر اساس میانگین‌گیری از نتیجه‌ی همه‌ی درخت‌های تصمیم مربوط درست شده است، و برای بسیاری از مجموعه‌ی داده‌ها، با صحت زیادی طبقه‌بندی می‌کند. در این مدل عامل میانگین کاهشی دقت و میانگین کاهشی جینی برای تعیین کردن اولویت تأثیر هر یک از عامل‌های موثر به کار برده می‌شود (نیکودموس ۲۰۱۱). میانگین کاهشی دقت در مقایسه با شاخص اهمیت جینی در تعیین اولویت عامل‌های موثر، بهتر و پایدارتر است، به‌ویژه در شرایطی که بین عامل‌های محیطی ارتباط هست (نیکودموس ۲۰۱۱). برای تهیه کردن نقشه‌ی خطر زمین لغزش ابتدا نقشه‌ی عامل‌های موثر و نقشه‌ی پراکنش زمین لغزش‌ها در قالب پرونده‌ی اکسل آماده کرده شد، به قالب متن (TXT) تبدیل، و در نرم‌افزار R فراخوانده شد. سپس وزن هر یک از عامل‌ها با الگوریتم جنگل تصادفی محاسبه شد. وزن‌های به‌دست آمده در قالب متنی آماده،

7- Normalized Difference Vegetation Index (NDVI)

8- Receiver Operating Characteristic (ROC)

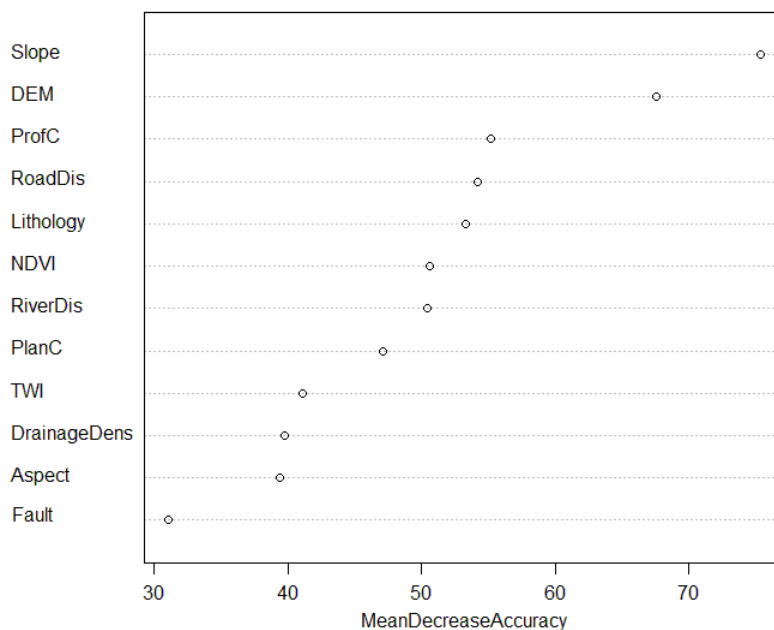
9- Variance Inflation Factor (VIF)

جدول ۱- بررسی هم خطی بین متغیرهای مستقل.

عامل های موثر	عامل تورم واریانس (VIF)	تاب آوری
جهت شیب	۱/۳۳۴	۰/۷۵۰
ارتفاع (متر)	۴/۸۹۷	۰/۲۰۴
تراکم زهکشی (کیلومتر بر کیلومتر مربع)	۱/۷۸۴	۰/۵۶۰
فاصله از گسل (متر)	۲/۹۱۳	۰/۳۴۳
واحدهای سنگ شناسی	۱/۸۲۷	۰/۵۴۷
NDVI	۱/۸۵۳	۰/۵۴۰
انحنای سطح	۱/۷۳۹	۰/۵۷۵
انحنای نیمرخ	۱/۷۰۶	۰/۵۸۶
فاصله از شبکه‌ی آبراه (متر)	۲/۲۶۵	۰/۴۴۱
فاصله از جاده (متر)	۱/۴۹۰	۰/۶۷۱
درجه‌ی شیب	۱/۳۲۶	۰/۷۵۴
TWI	۲/۳۸۷	۰/۴۱۹

در درجه‌ی شیب (Slope) و طبقه‌های ارتفاعی (DEM)، جهت شیب (Aspect)، و کم‌ترین اثر در فاصله از گسل (Fault) بررسی شده بود (شکل ۲).

اولویت‌بندی عامل‌های موثر بر وقوع زمین لغزش الگوریتم جنگل تصادفی برای اولویت‌بندی عامل‌های موثر بر زمین لغزش به کار برده شد. بیش‌ترین اثر بر وقوع زمین لغزش

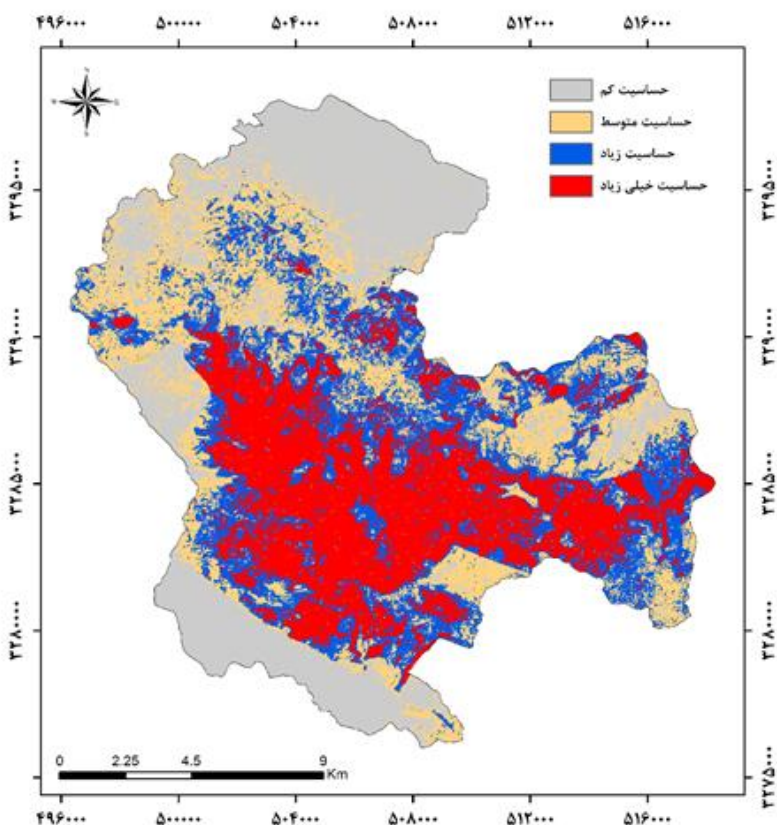


شکل ۲- اولویت‌بندی عامل‌های موثر بر وقوع زمین لغزش با روش میانگین کاهش دقت.*

* (ProfC): انحنای نیمرخ، RoadDis: فاصله از جاده، Lithology: سنگ‌شناسی، NDVI: شاخص تفاضلی پوشش گیاهی نرمال شده، RiverDis: فاصله از رودخانه، PlanC: انحنای سطح، TWI: شاخص رطوبت توپوگرافی، DrainageDens: تراکم زهکشی، و Fault: فاصله از گسل

۲۰۱۰؛ ژو و همکاران ۲۰۱۲؛ پورقاسمی و همکاران ۲۰۱۲) به چهار گروه حساسیت کم، متوسط، زیاد و خیلی زیاد طبقه‌بندی شد (شکل ۳).

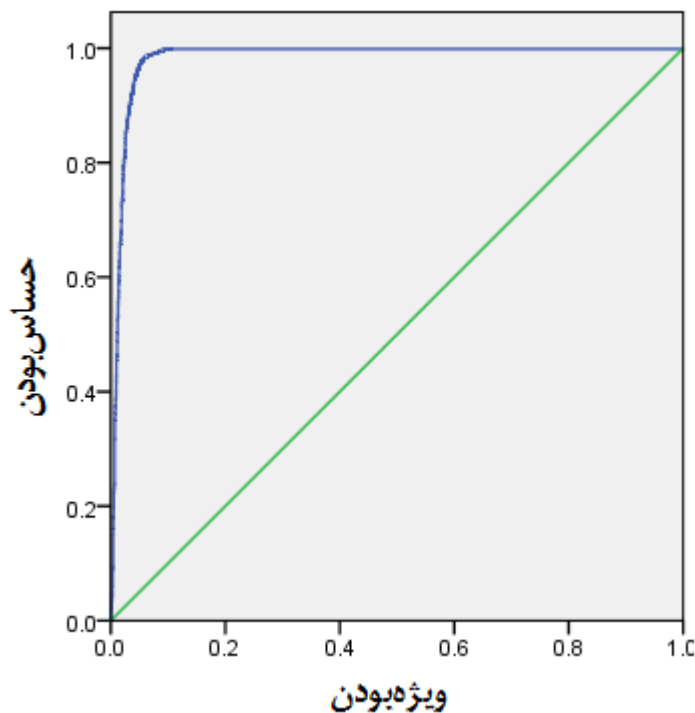
تهیه‌ی نقشه‌ی حساسیت زمین‌لغزش
نقشه‌ی حساسیت زمین‌لغزش بعد از تهیه‌شدن با این مدل در GIS، بر اساس شکستگی‌های طبیعی^{۱۱} (ارنر و همکاران



شکل ۳- نقشه‌ی پهنه‌بندی حساسیت زمین‌لغزش تهیه‌شده با الگوریتم جنگل تصادفی.

مدل جنگل تصادفی ۰/۹۸۳ بود (شکل ۴ و جدول ۲) که در طبقه‌ی عالی (۱-۰/۹) است. میزان خطای تخمینی (Out-Of-Bag) (OOB) مدل تهیه‌شده ۴/۶۳٪ است، که تأکیدی بر دقت زیاد آن است.

ارزیابی دقت نقشه‌های حساسیت زمین‌لغزش
برای بررسی میزان دقت مدل، سطح زیرمنحنی ROC به کار برده شد (شکل ۴). برای رسم منحنی ۳۰٪ زمین‌لغزش‌ها که در فرآیند مدل‌سازی آورده نشده‌بود، به کار برده شد. دقت



شکل ۴- نمودار منحنی تشخیص عمل کرد نسبی مدل یادگیری ماشین جنگل تصادفی.

جدول ۲- مقدارهای منحنی تشخیص عمل کرد نسبی و سطح زیرمنحنی.

سطح زیرمنحنی	خطای معیاری	ستراز معنی-داری	نوار بالا	نوار پایین
۰/۹۸۳	۰/۰۰۲	۰/۰۰۰	۰/۹۸۶	۰/۹۸۰

مساحت طبقه‌های حساسیت زمین لغزش نقشه‌ی تهیه‌شده با الگوریتم جنگل تصادفی محاسبه شد
مساحت هر یک از طبقه‌های حساسیت زمین لغزش برای (جدول ۳).

جدول ۳- مساحت طبقه‌های حساسیت در نقشه‌ی تهیه‌شده با مدل جنگل تصادفی.

طبقه‌ی حساسیت	مساحت	درصد
کم	۷۳۶۳۸	۲۵/۲۱
متوسط	۶۹۹۳۵	۲۳/۹۴
زیاد	۶۹۵۳۲	۲۳/۸۰
خیلی زیاد	۷۹۰۴۹	۲۷/۰۶

و فاصله از جاده بیشترین نقش را در وقوع زمین لغزش‌های منطقه داشته است. با توجه به این که نیروی گرانش از مهم‌ترین عامل‌های تحریک و گسترش زمین لغزش است، شیب نقش زیادی بر وقوع زمین لغزش دارد، و اگر سایر شرایط در ناحیه‌های کم‌شیب فراهم باشد احتمال وقوع زمین لغزش

بحث و نتیجه‌گیری
در این پژوهش مدل جنگل تصادفی برای اولویت‌بندی عامل‌های موثر و تهیه‌ی نقشه‌ی حساسیت زمین لغزش در آبخیز رئیسعلی دلواری به کار برده شد. نتایج میانگین کاهشی دقت نشان می‌دهد که عامل‌های شیب، ارتفاع، انحنای نیمرخ

به کار برده شده است (وورپال و همکاران، ۲۰۱۲؛ اولیویرا و همکاران، ۲۰۱۲؛ هونگ و همکاران، ۲۰۱۶). مدل جنگل تصادفی معمولاً هم‌زمان نقطه‌های بود و نبود زمین لغزش را به کار می‌گیرد، که باعث می‌شود مدل بهتر آموزش ببیند، و در نتیجه عمل کرد آن در نقشه‌ی نهایی بهتر باشد. دقت مدل در منطقه ۰/۹۸۳ به دست آمد که آن را در طبقه‌ی عالی می‌گذارد.

نتایج این تحقیق نشان داد که در حدود ۵۰٪ از منطقه در طبقه‌ی حساسیت زیاد و خیلی زیاد است. با توجه به این شرایط، لزوم توجه بیشتر به این پدیده نمایان می‌شود. زمین لغزش خسارت‌های زیادی به منطقه‌های مسکونی، زمین‌های کشاورزی و زیرساخت‌ها می‌زند و باعث به‌هدر رفتن حجم زیادی از خاک می‌شود. به کاربردن نقشه‌ی حساسیت زمین لغزش کمک می‌کند که تا حد امکان از منطقه‌های با حساسیت زیاد دوری کرده شود، تا تحریک نشوند. این نقشه کمک زیادی به برنامه‌ریزان و سیاست‌گزاران در بهینه‌کردن محل ساخت‌وساز سازه‌ها و تغییر در طبیعت می‌کند، و می‌تواند ابزاری مفید برای جلوگیری از این پدیده باشد. بنابراین نقشه‌ی حساسیت زمین لغزش تهیه‌شده ممکن است نقش موثری در تصمیم‌گیری‌های مدیران برای آمایش کردن سرزمین و مدیریت کردن جامع آبخیز سد رئیسعلی دلواری داشته باشد.

کم است. محققان دیگری از جمله احمدی و محمدخان (۲۰۰۳)، فاطمی‌عقدا و همکاران (۲۰۰۳)، فیض‌نیا و همکاران (۲۰۰۴)، شادفر و همکاران (۲۰۰۵)، چائو و چان (۲۰۰۵)، دایموند و همکاران (۲۰۰۶)، نئوهوسر و ترهورست (۲۰۰۶)، ایلماز و همکاران (۲۰۱۲)، یوسف و همکاران (۲۰۱۵)، و پورقاسمی و همکاران (۲۰۱۷) نیز نقش شیب را در وقوع زمین لغزش نشان داده‌اند. یکی دیگر از عامل‌های موثر ارتفاع است. با افزایش ارتفاع میزان بارش و در نتیجه توان اشباع‌شدن خاک افزایش یافته است، که منجر به افزایش زمین لغزش‌ها می‌شود. به‌طور کلی دست‌کاری در طبیعت بی‌توجه به شرایط منطقه موجب تحریک‌شدن و وقوع زمین لغزش خواهد شد، و یکی از این دست‌کاری‌ها احداث جاده در منطقه‌های کوهستانی است. معمولاً در نزدیکی جاده‌ها میزان لغزش‌ها افزایش می‌یابد، و هرچه از جاده‌ها فاصله می‌گیریم از تعداد لغزش‌ها کم می‌شود. می‌توان گفت که احداث جاده وضعیت طبیعی و شیب طبیعی منطقه را برهم می‌زند، بریدگی‌هایی در مسیر به‌وجود می‌آورد، و موجب افزایش فشار بر بخش پایین دست جاده می‌شود. فاطمی‌عقدا و همکاران (۲۰۰۳)، دومن و همکاران (۲۰۰۵)، لی (۲۰۰۷)، و یلسین (۲۰۰۸) نیز نقش جاده را در افزایش یافتن زمین لغزش‌ها نشان داده‌اند. برای تهیه‌ی نقشه‌ی حساسیت مدل جنگل تصادفی به کار برده شد که در بسیاری از پژوهش‌های دیگر نیز به‌دقت

فهرست منابع

- Ahmadi H, Mohammadkhan S. 2003. Determining the factors affecting mass movements (Case study: Taleghan Watershed). *Iranian Journal of Natural Resources*. 55(4): 455–466.
- Amiri M, Pourghasemi HR, Ghanbariana GA, Afzali SF. 2019. Assessment of the importance of gully erosion effective factors using Boruta algorithm and its spatial modeling and mapping using three machine learning algorithms. *Geoderma*. 340: 55–69.
- Arabameri AR, Pradhan B, Pourghasemi HR, Rezaei Kh, Kerle N. 2018. Spatial modelling of gully erosion using GIS and R programming: A comparison among three data mining algorithms. *Applied Sciences*. 8(8): 13–69. <https://doi.org/10.3390/app8081369>.
- Breiman L. 2001. Random forests. *Machine Learning*. 45: 5–32.
- Chau KT, Chan JE. 2005. Regional bias of landslide data in generating susceptibility maps using logistic regression: Case of Hong Kong Island. *Landslides*. 2: 280–290.
- Chen W, Xie X, Wang J, Pradhan B, Hong H, Tien Bui D, Duan Z, 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*. 151: 147–160.
- Davoudi Moghaddam D, Pourghasemi HR, Rahmati O. 2019. Assessment of the contribution of geo-environmental factors to flood inundation in a semi-arid region of SW Iran: Comparison of different advanced modeling approaches. *Natural hazards GIS-based spatial modeling using data mining techniques*. In HR. Pourghasemi, and M. Rossi (Eds.), *Natural hazards GIS-based spatial modeling using data mining techniques*.
- Duman TY, Can T, Gokceoglu C, Nefeslioglu HA, Sonmez H. 2006. Application of logistic regression for landslide susceptibility zoning of Cekmece Area, Istanbul, Turkey. *Environmental Geology*. 51: 241–256.
- Dymond JR, Ausseil AG, Shepherd JD, Buettner I. 2006. Validation of a region-wide model of landslide susceptibility in the Manawatu- Wanganui Region of New Zealand. *Geomorphology*. 74: 70–79.
- Erner A, Sebnem H, Duzgun B. 2010. Improvement of statistical landslide susceptibility mapping by using spatial and global regression methods in the case of More and Romsdal (Norway). *Landslides*. 7: 55–68.
- Fatemi Aghda SM, Ghayoumian J, Ashghali Farahani A. 2003. Evaluation of statistical methods in landslide hazard analysis. *Geosciences*. 82 (11): 28–47.
- Feiznia S, Kalarstāghi A, Ahmadi H, Safaei M. 2004. An investigation of effective factors on landslide occurrence and landslide hazard zonation (case study: Shirin Rood Drainage Basin - Tajan Dam). *Iranian Journal of Natural Resources*. 57 (1): 3–22.
- Garosi Y, Sheklabadi M, Besalatpour AA, Pourghasemi HR, Conoscenti C, Van Oost K. 2018. Comparison of the different resolution and source of controlling factors for gully erosion susceptibility mapping. *Geoderma*. 330: 65–78.
- Hong H, Pourghasemi HR, Pourtaghi ZS. 2016. Landslide susceptibility assessment in Lianhua County (China): A comparison between a random forest data mining technique and bivariate and multivariate statistical. *Geomorphology*. 259: 105–118.
- Iranian Landslide Working Party (ILWP). 2007.

- Iranian landslides list. Forest, Rangeland and Watershed Association, Tehran, Iran, 60. p
- Koehorst BAN, Kjekstad O, Patel D, Lubkowski Z, Knoeff JG, Akkerman GJ. 2005. Work package 6, Determination of Socio-Economic Impact of Natural Disasters, Assessing socio-economic Impact in Europe, 173p.
- Lagomarsino D, Tofani V, Segoni S, Catani F, Casagli N. 2017. A tool for classification and regression using random forest methodology: applications to landslide susceptibility mapping and soil thickness modeling. *Environmental Modeling and Assessment*. 22(3): DOI:10.1007/s10666-016-9538-y
- Lee S. 2007. Application and verification of fuzzy algebraic operators to landslide susceptibility mapping, *Environmental Geology*. 52: 615–623.
- Moore ID, Grayson RB, Ladson AR. 1991. Digital terrain modeling: a review of hydrological, geomorphological, and biological applications. *Hydrological Processes*. 5: 3–30.
- Naghibi A, Pourghasemi HR. 2015. A comparative assessment between three machine learning models and their performance comparison by bivariate and multivariate statistical methods for groundwater potential mapping in Iran. *Water Resources Management*. 29 (14): 5217–5236.
- National Cartographic Center. 2010. Topographic map of Raiisali Delvari Reservoir.
- Nefeslioglu HA, Gokceoglu C, Sonmez H. 2008. An assessment on the use of logistic regression and artificial neural networks with different sampling strategies for the preparation of landslide susceptibility maps. *Engineering Geology*. 97: 171–191.
- Neuhaus B, Terhorst B. 2006. Landslide susceptibility assessment using weights-of-evidence applied to a study area at the Jurassic Escarpment (SW-Germany). *Geomorphology*. 86 (1–2): 12–24.
- Nicodemus KK. 2011. Letter to the Editor: On the stability and ranking of predictors from random forest variable importance measures. *Briefings in Bioinformatics*. 12: 369–373.
- Shadfar S, Yamani M, Namaki M. 2005. Landslide hazard zonation using information value, area density, and LNRE models in Chalekroud watershed. *Water and Watershed*. 3: 62–68
- Oliveira S, Oehler F, San-Miguel-Ayaz J, Camia A, Pereira JMC. 2012. Modeling spatial patterns of fire occurrence in Mediterranean Europe using multiple regression and random forest. *Forest Ecology and Management*. 275: 117–129.
- Park S, Kim J. 2019. Landslide susceptibility mapping based on random forest and boosted regression tree models, and a comparison of their performance. *Applied Sciences*. 9: 942 doi:10.3390/app9050942
- Pourghasemi HR, Kerle N. 2016. Random Forest-evidential belief function based landslide susceptibility assessment in western Mazandaran Province, Iran. *Environmental Earth Sciences*. 75: 185. DOI: 10.1007/s12665-015-4950-1
- Pourghasemi HR, Pradhan B, Gokceoglu C. 2012. Application of fuzzy logic and analytical hierarchy process (AHP) to landslide susceptibility mapping at Haraz watershed, Iran, *Natural Hazards*. 63 (2): 965–996.
- Pourghasemi HR, Rahmati A. 2018. Rapid GIS-based spatial and regional modelling of landslide susceptibility using machine learning techniques in the R open source software. *Catena*. 162: 177–192.
- Pourghasemi HR, Rossi M. 2017. Landslide

- susceptibility modeling in a landslide prone area in Mazandarn Province, north of Iran: a comparison between GLM, GAM, MARS, and M-AHP methods. *Theoretical Applied Climatology*. 130 (1–2): 609–633.
- Rahmati O, Pourghasemi HR, Melesse A. 2016. Application of GIS-based data driven random forest and maximum entropy models for groundwater potential mapping: A case study at Mehran Region, Iran. *Catena*. 137: 360–372.
- Rahmati O, Tahmasebipour N, Haghizadeh A, Pourghasemi HR, Feizizadeh B. 2017. Evaluating the influence of geo-environmental factors on gully erosion in a semi-arid region of SW Iran, using rule-based semi-automated technique and conditional probability model. *Science of the Total Environment*. 579: 913–927.
- Shataee S, Weinaker H, Babanejad M. 2011. Plot-level forest volume estimation using airborne laser scanner and TM data, comparison of boosting and random forest regression algorithms. *Procedia Environmental Sciences*. 7: 68–73.
- Taalab K, Cheng T, Zhang Y. 2018. Mapping landslide susceptibility and types using Random Forest. *Big Earth Data*. 2: 1–20.
- Vorpahl P, Elsenbeer H, Märker M, Schröder B. 2012. How can statistical models help to determine driving factors of landslides? *Ecological Modelling*. 239: 27–39.
- Xu C, Dai F, Xu X, Lee YH. 2012. GIS-based support vector machine modeling of earthquake-triggered landslide susceptibility in the Jianjiang River watershed, China. *Geomorphology*. 145–146: 70–80.
- 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: 1–12.
- Yesilnacar EK. 2005. The Application of Computational Intelligence to Landslide Susceptibility Mapping in Turkey, Ph.D Thesis. Department of Geomatics the University of Melbourne, 423. p
- Yilmaz C, Topal T, Suzen ML. 2012. GIS-based landslide susceptibility mapping using bivariate statistical analysis in Devrek (Zonguldak-Turkey). *Environmental Earth Sciences*. 65: 2161–2178.
- Youssef AM, Pourghasemi HR, Pourtaghi ZS, Al-Katheeri MM. 2015. Landslide susceptibility mapping using random forest, boosted regression tree, classification and regression tree, and general linear models and comparison of their performance at Wadi Tayyah Basin, Asir Region, Saudi Arabia. *Landslides*. 13 (5): 839–856.



Watershed Management Research

VOL. 33, No. 1, Ser. No: 126, Spring 2020, pp.2-13

DOI: 10.22092/wmej.2019.126288.1219

Landslide Susceptibility Modelling Using the Random Forest Machine Learning Algorithm in the Watershed of Rais-Ali Delvari Reservoir

Naser Heydari

PhD Graduated, Department of Watershed Sciences Engineering, Faculty of Natural Resources, University of Agricultural Science and Natural Resources of Sari, Sari, Iran

Mahmoud Habibnejad

Professor, Department of Watershed Sciences Engineering, Faculty of Natural Resources, University of Agricultural Science and Natural Resources of Sari, Sari, Iran

Ataollah Kavian

Professor, Department of Watershed Sciences Engineering, Faculty of Natural Resources, University of Agricultural Science and Natural Resources of Sari, Sari, Iran

Hamid Reza Pourghasemi*

(Corresponding Author)* Associate Professor, Department of Natural Resources and Environmental Engineering, College of Agriculture, Shiraz University, Shiraz, Iran

*Corresponding Author Email: hr.pourghasemi@shirazu.ac.ir

Received: 16 May 2019

Accepted: 03 September 2019

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 the watershed of the Rais-Ali Delvari Reservoir. The landslide inventory map was prepared using extensive field surveys and the Iranian Landslides Working Party Data Bank. Of the total of 279 identified landslide locations, 70% were used for the modelling processes and the remaining (30%) were applied for validation of the developed model. Different thematic layers including elevation, slope angle, plan curvature, profile curvature, topographic wetness index (TWI), distance from rivers, drainage density, distance from faults, distance from roads, lithological units, and the normalized difference vegetation index (NDVI) were selected. According to the relationship between the dependent (landslides) and the independent (effective factors) variables in the R statistical software, the random forest algorithm was run using the "Random Forest" package, and a landslide susceptibility map was prepared. Accuracy of the model was tested using the receiver operating characteristic (ROC) curve based on 30% of unused landslides in the modelling process. Accuracy results indicated that the Random Forest model with an AUC value of 0.983 had an excellent precision. Also, prioritization of the effective factors showed that the slope angle, elevation, plan curvature, distance from road, and lithological units had the highest effect on landslide occurrence. Therefore, it may be suggested that the prepared landslide susceptibility map could be effective in decision making for land use planning, and in the managing of the Rais-Ali Delvari Reservoir Watershed.

■ **Keywords:** Landslide susceptibility, mean decrease accuracy, Random Forest, Rais-Ali Delvari Reservoir Watershed ■