دانشگاه علوم کشاورزی و منابع طبیعی ساری
پژوهشنامه مدیریت حوزه آبخیز

اولویت‌بندی عوامل موثر بر وقوع زمین‌لغزش و تهیه نقشه حساسیت آن با استفاده از الگوریتم نوین جنگل تصادفی (مطالعه موردی: بخشی از استان گلستان)

مجید محمدی[1] و حمیدرضا پورقاسمی[2]

۱- استادیار گروه مرتع و آبخیزداری، دانشکده مهندسی منابع طبیعی، دانشگاه سمنان، (نویسنده مسوول: majid.mohammady@semnan.ac.ir)
۲- استادیار بخش مهندسی منابع طبیعی و محیط زیست، دانشکده کشاورزی، دانشگاه شیراز
تاریخ دریافت: ۹۴/۸/۵ تاریخ پذیرش: ۹۵/۴/۲۶

چکیده

زمین‌لغزش به‌عنوان یکی از بلایای طبیعی، خطراتزیادی به‌ویژه در مناطق کوهستانی داشته و منجر به خسارت جانی و مالی فراوانی در سرتاسر دنیا می‌شود. ایران به‌ویژه در مناطق شمالی و غربی به‌خاطر شرایط اقلیمی و توپوگرافی، همواره در معرض خطر زمین‌لغزش می‌باشد. هدف از این پژوهش، ارزیابی عوامل موثر بر وقوع زمین‌لغزش و تهیه نقشه حساسیت آن در بخشی از استان گلستان با استفاده از الگوریتم نوین جنگل تصادفی است. در ابتدا موقعیت زمین‌لغزش‌ها با استفاده از مطالعات میدانی، گزارش‌های تاریخی و گوگل ارث مشخص شد. در مجموع ۷۸ زمین‌لغزش در منطقه شناسایی، که به دو گروه برای مدل‌سازی (۷۰ درصد) و ارزیابی (۳۰ درصد) تقسیم شدند. یازده نقشه از عوامل موثر بر وقوع زمین‌لغزش شامل جهت شیب، ارتفاع، فاصله از رودخانه، فاصله از گسل، فاصله از جاده، سنگ‌شناسی، کاربری اراضی، شاخص طول شیب، شکل شیب، بارش و درجه شیب تهیه شد. ارتباط عوامل موثر و نقشه پراکنش زمین‌لغزش‌ها و اولویت آن‌ها با استفاده از الگوریتم جنگل تصادفی بررسی و سپس نقشه حساسیت زمین‌لغزش در محیط GIS تهیه گردید. اولویت‌بندی عوامل موثر نشان داد که فاصله از جاده، فاصله از گسل و ارتفاع به‌ترتیب بیش‌ترین تاثیر را در منطقه مورد مطالعه داشته‌اند. نهایتاً نقشه حساسیت زمین‌لغزش تهیه شده با استفاده از الگوریتم جنگل تصادفی به چهار طبقه شامل حساسیت کم (۲۹/۱۸ درصد)، متوسط (۳۳/٤٤ درصد)، زیاد (۲٤/۸۲ درصد) و خیلی‌زیاد (۱۲/۵۵ درصد) تقسیم شد. برای ارزیابی نقشه تهیه شده از منحنی ROC و تخمین سطح زیرمنحنی از ۳۰ درصد زمین‌لغزش‌های استفاده نشده در مدل‌سازی استفاده گردید. نتایج نشان داد مدل جنگل تصادفی با سطح زیر منحنی ۰/۷۰٦ دقت قابل قبولی برای تهیه نقشه حساسیت زمین‌لغزش ارائه کرده است.

واژه‌های کلیدی: زمین‌لغزش، نقشه حساسیت زمین‌لغزش، سامانه اطلاعات جغرافیایی، الگوریتم جنگل تصادفی

مقدمه

زمین‌لغزش موجب خسارت به زندگی و دارایی‌های انسان شده و مهم‌ترین خطر طبیعی شناخته شده در مناطق کوهستانی به‌حساب می‌آید. به همین دلیل از مدل‌های مختلفی برای بررسی حساسیت و خطر زمین‌لغزش مانند مدل‌های کارشناسی[1]، تعیینی[2] و آماری[3] طراحی و استفاده شده است (۵۶). این پدیده موجب بروز خسارات زیادی از جمله خسارت به انسان‌ها، مناطق مسکونی، صنعتی و محیط زیست شده (۲۱) و به‌همین دلیل سیاست‌گذاران توجه زیادی به مطالعات زمین‌لغزش و پهنه‌بندی مناطق پرخطر به‌منظور شناسایی مناطق پایدار برای توسعه در آینده دارند (۲۷). تاکنون مطالعات زیادی با استفاده از روش‌های مختلف در زمینه حرکات توده‌ای و زمین‌لغزش انجام شده است. به‌عنوان مثال می‌توان به کاربرد مدل‌های بیزین (۳۶،۲۸،۲۴)، رگرسیون چندمتغیره و لجستیک (۲۵،۲۹،۳۰،۶،۳،۱، ۵۵،۴۵)، منطق فازی (۳۹،۱۴،۲)، شبکه عصبی مصنوعی (۵۷،۱۹،۱۱)، درخت تصمیم‌گیری (۴۳،۳۲)، مدل‌های عصبی- فازی (۴۶، ۵۱، ۵۸)، جنگل تصادفی (۵۶،۵۰،۴۸،۴۲،۱۲) اشاره نمود.

وجود رخساره‌های کوهستانی، فعالیت‌های تکتونیکی زیاد و شرایط زمین‌شناسی و اقلیمی متنوع شرایط را برای وقوع زمین‌لغزش مستعد کرده است. استان گلستان در شمال ایران به‌دلیل شرایط آب و هوایی و فیزیوگرافی همواره در معرض خطر زمین‌لغزش بوده و توجه به این عارضه امری ضروری‌است. هدف از تهیه نقشه‌های حساسیت زمین‌لغزش مشخص کردن مناطق حساس با توجه به داده‌های گذشته زمین‌لغزش و بررسی عوامل موثر بر وقوع آن است. هدف از این تحقیق، اولویت‌بندی عوامل موثر بر وقوع زمین‌لغزش و تهیه نقشه حساسیت آن با استفاده از الگوریتم نوین جنگل تصادفی در بخشی از مناطق کوهستانی استان گلستان است. در طبیعت مرز مشخصی برای طبقه‌های هر یک از عوامل از قبیل نوع خاک، انواع کاربری اراضی و واحدهای سنگ شناسی وجود ندارد. هم‌چنین تعیین مرز طبیعی برای عوامل پیوسته نظیر ارتفاع، شیب، فاصله از عناصر خطی، شاخص‌های توپوگرافی و عناصر تراکمی بسیار مشکل است. به‌همین دلیل مدل‌های سنتی و آماری در برخورد با پارامترهای ورودی مذکور دارای معایبی هستند. تعیین دقت بالا در مطالعات ارزیابی خطر زمین‌لغزش بسیار مهم بوده و ساده‌گرفتن پارامترهای اصلی لغزش، طبقه‌های آن‌ها و تعاملات بین آن‌ها می‌تواند منجر به نتایج نادرست در نقشه نهایی گردد. هم‌چنین در بیش‌تر مطالعات، قضاوت‌های کارشناسی نقش مهمی داشته و رویکردهای تجربی برای ارزیابی‌های مختلف به طور گسترده مورد استفاده قرار گرفته‌اند. این مسائل منجر به استفاده از روش‌های داده‌کاوی[4] در مطالعات زمین‌لغزش گردیده است (۴۹). از طرفی، تحقیقات گذشته نشان داده که مدل مذکور در پهنه‌بندی حساسیت زمین‌لغزش بسیار جدید و نوین می‌باشد. این امر مدیران را در برنامه‌ریزی‌های آینده

1- Heuristic 2- Deterministic 3- Statistical 4- Data Mining

[DOI: 10.29252/jwmr.8.15.161]

کمک نموده تا از تغییرات و دخالت‌های انسانی در این مناطق تا حد امکان جلوگیری نموده و موجب تحریک این مناطق نشود.

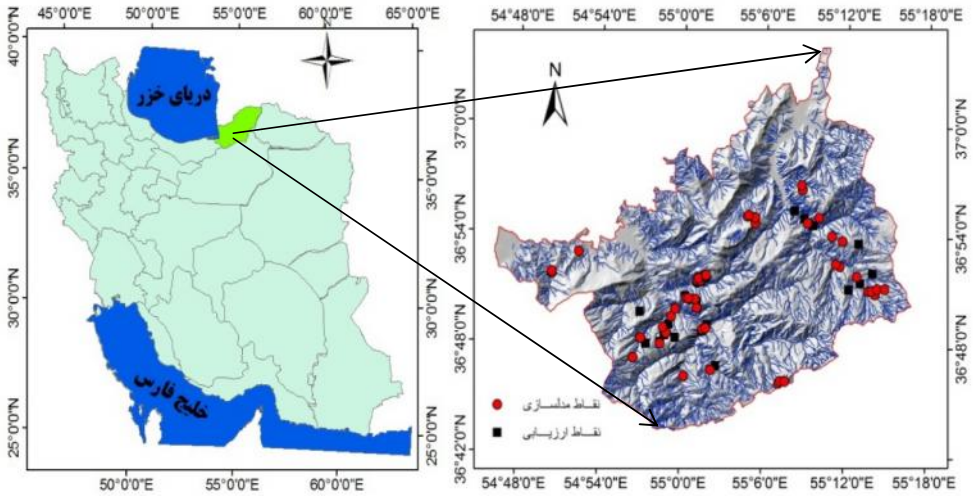
مواد و روش‌ها

منطقه مورد مطالعه

منطقه مورد مطالعه بخش کوهستانی حوزه آبخیز باغ سالیان در استان گلستان با مساحتی در حدود ۸۲۰ کیلومتر مربع بوده که بین عرض‌های "۰ ′۴۳ °۳۶ تا "۳۰ ′۴ °۳۷ شمالی و طول "۳۰ ′۴۶ °۵۴ تا "۳۰ ′۱۶ °۵۵ شرقی در جنوب استان گلستان واقع شده است. ارتفاع کمینه و بیشینه آن به‌ترتیب برابر با ۱۷۲ و ۲۹۴۷ متر از سطح دریا بوده و توپوگرافی پیچیده‌ای دارد (شکل ۱). کاربری اراضی منطقه شامل جنگل، مرتع، کشاورزی و مناطق مسکونی است.

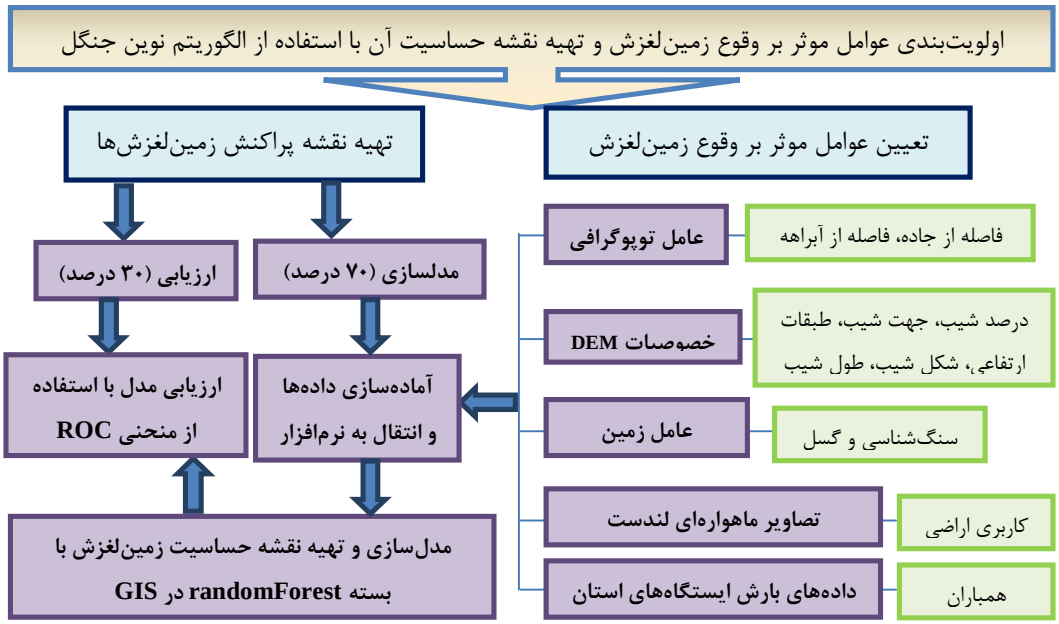
داده‌های مورد استفاده

روش تحقیق اولویت‌بندی عوامل موثر بر وقوع زمین‌لغزش و تهیه نقشه حساسیت آن با استفاده از الگوریتم جنگل تصادفی به‌صورت نمودار جریانی در شکل ۲ ارائه شده است. نقشه زمین‌شناسی منطقه مورد مطالعه از سازمان زمین‌شناسی کشور تهیه، در محیط نرم‌افزار ArcGIS 9.3 رقومی و لایه‌های لیتولوژی (سنگ‌شناسی) و گسل از آن استخراج گردید. نقشه توپوگرافی با مقیاس ۱:۲۵۰۰۰ از سازمان جغرافیایی تهیه شده و لایه‌های اطلاعاتی درجه شیب، جهت شیب، ارتفاع از سطح دریا، شکل شیبو شاخص طول شیب با استفاده از آن تهیه گردید. هم‌چنین نقشه‌های شبکه آبراهه و جاده از نقشه‌های توپوگرافی و گسل نیز از نقشه زمین‌شناسی منطقه استخراج و نقشه فاصله از این عوارض در محیط GIS تهیه شد. کاربری اراضی با استفاده از تصاویر ماهواره‌ای لندست ۷ در سال ۲۰۱۲ و با استفاده از روش حداکثر احتمال به‌دست آمد. نقشه همباران نیز از مقادیر بارش سالانه ایستگاه‌های استان تهیه گردید. در مجموع ۷۸ موقعیت لغزشیاز نوع لغزش‌های خاکیدر بازدیدهای زمینی و استفاده از گوگل ارث در منطقه مورد مطالعه شناسایی و نقشه پراکنش آن تهیه گردید. ۵۲ مورد از زمین‌لغزش‌ها برای مدل‌سازی و ۲۶ مورد برای ارزیابی مدل در نظر گرفته شد.



شکل ۱- موقعیت منطقه مورد مطالعه
Figure 1. Location of the case study

[DOI: 10.29252/jwmr.8.15.161]



شکل ۲- نمودار جریانی تهیه نقشه حساسیت زمین‌لغزش با استفاده از الگوریتم جنگل تصادفی
Figure 2. Flowchart of landslide susceptibility mapping using random forest algorithm

الگوریتم جنگل تصادفی[1]

جنگل تصادفی یک روش یادگیری نظارت شده بوده که از درخت‌های چندگانه در طبقه‌بندی استفاده می‌کند (۱۰). الگوریتم جنگل تصادفی با جای‌گزینی و تغییر مداوم عوامل موثر و مرتبط با هدف، منجر به ایجاد تعداد زیادی درخت تصمیم‌گیری شده، سپس تمام درختان به‌منظور پیش‌بینی با هم ترکیب می‌گردد (۵۲). در فرآیند آموزش هر سری داده‌های اصلی ساخته درخت بر اساس انتخاب تصادفی می‌شود (۹). جنگل تصادفی شامل سه پارامتر تعریف شده توسط کاربر شامل: ۱- تعداد متغیرهای مورد استفاده در ساخت هر درخت که قدرت هر درخت مستقل را بیان می‌کند، ۲- تعداد درختان در جنگل تصادفی و ۳- حداقل تعداد گره‌های انتهایی است (۳۸). قدرت پیش‌بینی جنگل تصادفی با افزایش قدرت درختان مستقل و کاهش همبستگی بین آن‌ها افزایش می‌یابد (۲۶). الگوریتم جنگل تصادفی از تمام داده‌های موجود برای رویاندن درخت استفاده نکرده و از ۶۶ درصد داده‌ها استفاده شده که به آن نمونه Bootstrap گفته می‌شود. سپس یک متغیر پیش‌گو به‌طور تصادفی در خلال فرآیند رویاندن[2] معرفی می‌گردد. انتخاب این متغیر برای شکافتن یک گره در ساخت درخت به‌صورت تصادفی انجام شده و به این شکل درخت تصمیم به بزرگ‌ترین اندازه ممکن رویانده می‌شود (۹). ۳۳ درصد از داده‌های باقیمانده نیز برای ارزیابی درخت برازش شده استفاده می‌شود. این فرآیند چندین بار تکرار شده و میانگین تمام مقادیر پیش‌بینی شده به‌عنوان پیش‌بینی نهایی الگوریتم استفاده می‌شود (۴۷).

در این مدل از دو عامل میانگین کاهشی دقت و میانگین کاهشی جینی برای تعیین اولویت هر یک از عوامل موثر استفاده می‌گردد. استفاده از میانگین کاهشی دقت در مقایسه با شاخص اهمیت جینی در تعیین اولویت عوامل موثر، بهتر و پایدارتر می‌باشد، به‌ویژه در شرایطی که بین فاکتورهای محیطی ارتباط وجود دارد (۳۳). در ابتدا، به‌منظور بررسی هم‌خطی بین متغیرهای مستقل از دو شاخص عامل تورم واریانس (VIF)[3] و ضریب تحمل[4] استفاده گردید. ضریب تحمل کم‌تر از ۰/۲ یا ۰/۱ و عامل تورم واریانس ۵ یا بزرگ‌تر از مقدار مذکور نمایان‌گر هم‌خطی[5] بودن و ارتباط زیاد بین دو متغیر مستقل است (۳۴). در مرحله بعد با وارد کردن داده‌های مربوط به عوامل موثر و نقشه پراکنش زمین‌لغزش‌ها به نرم‌افزار R، مدلسازی با استفاده از الگوریتم جنگل تصادفی انجام و نقش عوامل موثر بر وقوع زمین‌لغزش تعیین گردید.نهایتاً وزن‌های به‌دست آمده از الگوریتم جنگل تصادفی در نرم‌افزار R به محیط ArcGIS منتقل و نقشه‌های نهایی حساسیت زمین‌لغزش برای منطقه مورد مطالعه تهیه گردید. نقشه نهایی بر اساس شکستگی‌های طبیعی به چهار طبقه حساسیت کم، متوسط، زیاد و خیلی‌زیاد تقسیم‌بندی گردید.

ارزیابی نقشه حساسیت زمین‌لغزش

بعد از تهیه نقشه حساسیت زمین‌لغزش از منحنی تشخیص عملکرد نسبی (ROC)[6] به‌منظور ارزیابی نقشه مذکور استفاده شد (۴۴،۴۳،۳۱). منحنی تشخیص عملکرد نسبی یکی از مفیدترین و کارآمدترین روش‌ها در پیش‌بینی و

1- Random Forest (RF)
2- Grow
3- Variance Inflation Factor (VIF)
4- Tolerance
5- Multi-Collinearity
6- Receiver Operating Characteristic (ROC)

[DOI: 10.29252/jwmr.8.15.161]

تعیین دقت مدل‌سازی است و در حقیقت یک نمایش گرافیکی از تخمین درست وقوع و عدم وقوع زمین‌لغزش‌ها محسوب می‌شود. سطح زیرمنحنی ROC، بیان‌گر مقدار پیش‌بینی سیستم از طریق توصیف توانایی آن در تخمین درست وقایع رخداده (وجود زمین‌لغزش) و عدم وقوع رخداد (عدم وجود زمین‌لغزش) آن است. چنان‌چه مدلی نتواند وقوع زمین‌لغزش را بهتر از دیدگاه احتمالی یا تصادفی تخمین زند مقدار سطح زیرمنحنی آن ۰/۵ و بنابراین کم‌ترین دقت را دارد، در حالی‌که اگر سطح زیرمنحنی برابر با یک باشد مدل تهیه شده بیش‌ترین دقت را خواهد داشت و مقادیر AUC از ۰/۵–۱ متغیر است. هر چه سطح زیرمنحنی به یک نزدیک‌تر باشد، بیان‌گر بهترین دقت از نقشه پهنه‌بندی تهیه شده است (۵۴).

نتایج و بحث

همان‌طور که بیان شد در اولین گام هم‌خطی بین متغیرهای مستقل با استفاده از دو شاخص عامل تورم واریانس و ضریب تحمل بررسی شده و نتایج نشان داد بین عوامل مورد استفاده هیچ‌گونه هم‌خطی وجود ندارد (جدول ۱). در مرحله بعد پس از بررسی ارتباط بین عوامل موثر و نقشه پراکنش زمین‌لغزش‌ها، داده‌ها آماده و به نرم‌افزار R وارد و مدل‌سازی بر اساس بسته الحاقی random Forest انجام شد. برای اولویت‌بندی فاکتورها از دو معیار میانگین کاهشی دقت و شاخص اهمیت جینی استفاده شده و اولویت فاکتورها تعیین گردید (شکل ۳). البته همان‌طور که قبلاً بیان گردید میانگین کاهشی دقت معیار مناسب‌تری برای اولویت‌بندی فاکتورها است(۳۳). نتایج حاصل از اولویت‌بندی در جدول ۲ نشان داده شده است. نتایج نشان داد که سه معیار فاصله از جاده، گسل و ارتفاع به ترتیب بیش‌ترین تاثیر را در وقوع زمین‌لغزش دارند. در منطقه مورد مطالعه احداث جاده و فرسایش کناری عامل مهمی در ناپایداری شیب و وقوع زمین‌لغزش است (۲۷). به‌طور کلی جاده‌ها وضعیت طبیعی دامنه را به‌هم زده، موجب به‌وجود آمدن بریدگی‌های عمودی در دامنه می‌شوند. این موضوع باعث افزایش فشار بر بخش پایین جاده شده و منجر به افزایش زمین‌لغزش در اطراف جاده‌ها می‌گردد. به این دلیل است که در بسیاری از مطالعات زمین‌لغزش، جاده به‌عنوان فاکتوری تاثیرگذار درنظر گرفته می‌شود (۵۳،۴،۶،۱۷،۱۸،۲۲،۳۸). عامل بعد که بیش‌ترین تاثیر را بر وقوع زمین‌لغزش داشته است فاصله از گسل است. گسل از جمله مهم‌ترین عوامل تکتونیکی است که موجب حساسیت دامنه‌ها می‌گردد و به همین دلیل در مطالعات زیادی نقش آن در وقوع زمین‌لغزش بررسی می‌شود (۴۶، ۴،۶،۱۵،۱۷،۱۸،۲۳). گسل موجب شکستگی و ایجاد انفصالات در سنگ‌ها شده و حرکت آن‌ها را آسان‌تر می‌کند. به‌گونه‌ای که هر قدر تراکم این انفصالات بیش‌تر باشد حساسیت سنگ در مقابل شکل‌های مختلف فرسایش از جمله زمین‌لغزش بالاتر است.

بر همین اساس در اغلب مطالعات بررسی ناپایداری دامنه‌ها، عامل گسل به‌عنوان یک عامل موثر در وقوع زمین‌لغزش مورد بررسی قرار می‌گیرد.

جدول ۱– بررسی هم‌خطی بین عوامل موثر بر وقوع زمین‌لغزش منطقه مورد مطالعه

Table 1. Investigating multi-collinearity of landslide conditioning factors in the case study

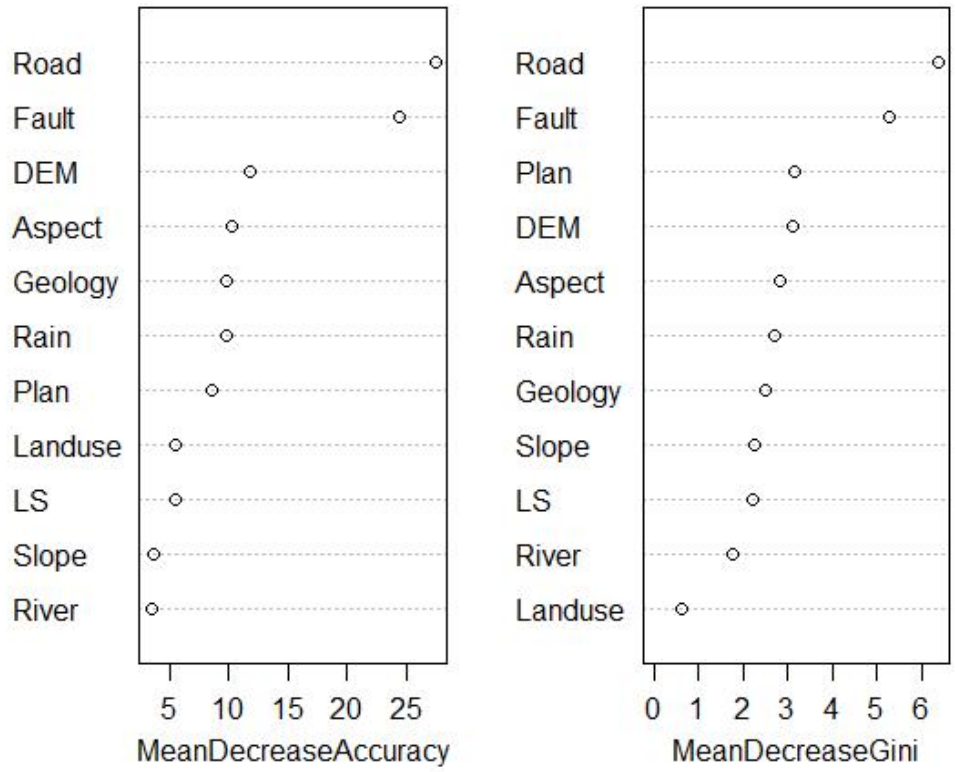
عوامل موثر بر وقوع زمین‌لغزش	عامل تورم واریانس (VIF)	ضریب تحمل (Tolerance)
جهت شیب	۱/۱۸۵	۰/۸۴۴
ارتفاع (متر)	۱/۹۸۹	۰/۵۰۳
فاصله از رودخانه (متر)	۱/۱۱۱	۰/۹۰۰
زمین‌شناسی	۱/۲۱۳	۰/۸۲۴
کاربری اراضی	۱/۳۳۴	۰/۷۵۰
شاخص طول شیب (متر)	۴/۱۷۸	۰/۲۳۹
شکل شیب	۱/۱۹۹	۰/۸۳۴
بارش سالانه (میلی‌متر)	۱/۶۱۷	۰/۶۱۸
فاصله از جاده (متر)	۱/۱۱۵	۰/۸۹۷
درجه شیب	۴/۰۱۱	۰/۲۴۹
فاصله از گسل (متر)	۱/۶۵۲	۰/۶۰۵

جدول ۲– میزان اهمیت عوامل موثر بر وثوع زمین‌لغزش بر اساس دو معیار میانگین کاهشی دقت و میانگین کاهشی جینی

Table 2. Importance of conditioning factors based on mean decrease accuracy and mean decrease Gini

عامل	۰	۱	میانگین کاهشی دقت	میانگین کاهشی جینی
جهت شیب	۴/۰۴	۱۰/۵۷	۹/۹۷	۲/۸۳
ارتفاع	۶/۴۶	۱۲/۸۴	۱۲/۸۴	۳/۰۶
فاصله از رودخانه	-۲/۱۰	۱/۶۷	-۰/۵۱	۱/۹۲
زمین‌شناسی	۳/۵۲	۱۱/۳۲	۱۰/۲۱	۲/۶۱
کاربری اراضی	۱/۳۳	۲/۸۳	۳/۱۹	۰/۶۲
شاخص طول شیب	۲/۳۳	۵/۲۰	۵/۱۵	۲/۱۸
شکل شیب	۶/۵۵	۶/۵۳	۸/۲۵	۲/۹۹
بارش سالانه	۳/۳۳	۱۲/۳۳	۱۰/۹۶	۲/۶۰
فاصله از جاده	۲۰/۶۸	۲۳/۱۹	۲۷/۳۶	۶/۱۳
درجه شیب	۲/۷۰	۱/۲۵	۲/۶۴	۲/۳۵
فاصله از گسل	۱۷/۵۱	۲۳/۹۰	۲۵/۳۵	۵/۵۲

[DOI: 10.29252/jwmr.8.15.161]



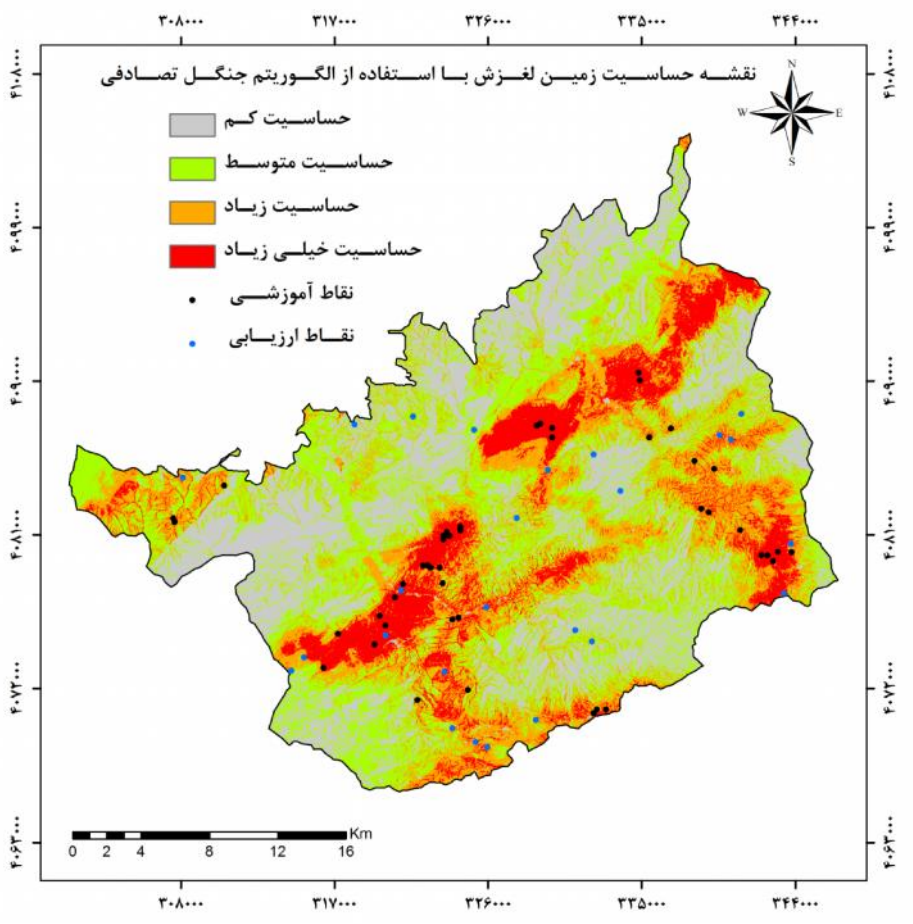
شکل ۳- اولویت‌بندی عوامل موثر بر وقوع زمین‌لغزش با استفاده از الگوریتم جنگل تصادفی
Figure 3. Prioritization of landslide conditioning factors using random forest algorithm

عامل موثر دیگر در وقوع زمین‌لغزش ارتفاع از سطح دریا است، هرچه ارتفاع از سطح دریا بیش‌تر شود،حساسیت به زمین‌لغزش نیز بیش‌تر می‌گردد (۳۷). در ارتفاعات بالا هوازدگی سنگ‌ها بیش‌تر بوده و پدیده ذوب و انجماد غالب است (۱۶). به‌دلیل اهمیت نقش ارتفاع در وقوع زمین‌لغزش محققین در مناطق مختلف از این عامل در تهیه نقشه حساسیت زمین‌لغزش استفاده نموده‌اند (۷،۸،۱۷،۱۸،۲۳،۴۳).

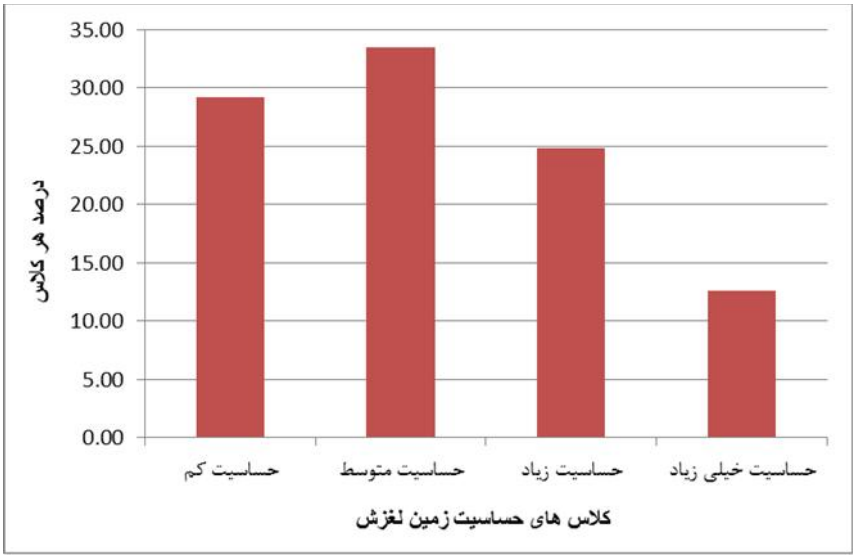
پس از اجرای مدل جنگل تصادفی در نرم‌افزار R برای داده‌های مشاهده‌ای و تعمیم فرآیند مدل‌سازی به کل منطقه مورد مطالعه در قالب مدلسازی Black Box، خروجی به‌صورت یک فایل با فرمت متنی (txt) یا csv خواهد بود که دارای دو ستون اعداد به‌صورت صفر و یک می‌باشد. ستون صفر به‌معنای مقادیر غیرلغزشی برای هر پیکسل و ستون یک بیان‌گر مقادیر لغزشی برای هر پیکسل از کل منطقه مورد مطالعه می‌باشد. به‌منظور پهنه‌بندی حساسیت زمین‌لغزش، ستون پیکسل‌های صفر حذف شده و مقادیر ستون یک (لغزش) بر اساس مختصات هر یک از پیکسل‌ها در محیط GIS در قالب یک فایل نقطه‌ای در آمده و سپس از مسیر تبدیل فرمت نقطه به رستر، نقشه نهایی برای منطقه مورد مطالعه تهیه گردید (شکل ۴).

هم‌چنین به‌منظور مشخص شدن وضعیت کلی منطقه مورد مطالعه، مساحت طبقات حساسیت زمین‌لغزش محاسبه گردید. بررسی مساحت طبقات نشان داد طبقه متوسط بیش‌ترین مساحت را داشته (شکل ۵) و مساحت طبقات حساسیت زیاد و خیلی‌زیاد ۲۴/۸۲ و ۱۲/۵۵ درصد برآورد گردیده است.

[DOI: 10.29252/jwmr.8.15.161]



شکل ۴- نقشه پهنه‌بندی حساسیت زمین‌لغزش با استفاده از الگوریتم جنگل تصادفی
Figure 4. Landslide susceptibility map using random forest algorithm



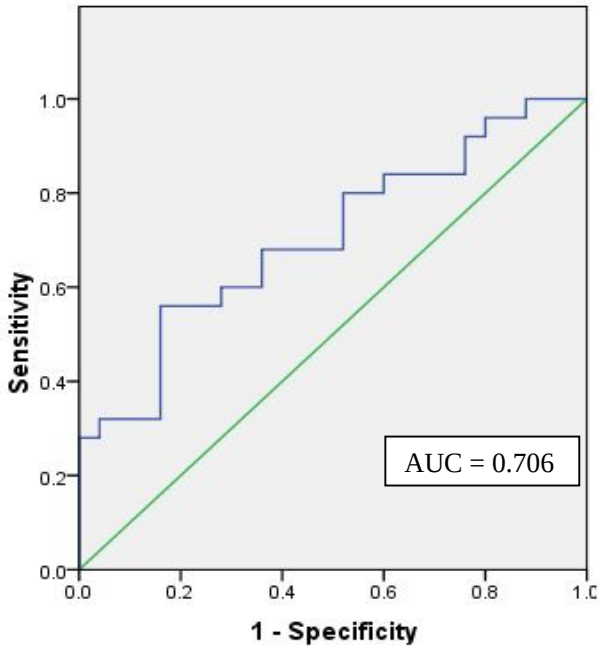
شکل ۵- درصد طبقات حساسیت نقشه زمین‌لغزش در منطقه مورد مطالعه
Figure 5. Percent of landslide susceptibility classes in the case study

[DOI: 10.29252/jwmr.8.15.161]

همان‌طور که بیان شد برای ارزیابی مدل از منحنی ROC و سطح زیرمنحنی استفاده گردید. منحنی ROC، از کارآمدترین روش‌ها در ارائه خصوصیت تعیینی، شناسایی احتمالی و پیش‌بینی سیستم‌هاست که میزان دقت مدل را به‌صورت کمی برآورد می‌کند (۵۰). نتایج ارزیابی نشان داده کهمقدار سطح زیرمنحنی برابر با۰/۷۰۶ با مقدار اشتباه استاندارد ۰/۲۰۵ می‌باشد (شکل ۶).

با توجه به طبقه‌بندی ارائه شده برای سطح زیر منحنی (۱-۰/۹، عالی، ۰/۹-۰/۸، خیلی‌خوب، ۰/۸-۰/۷، خوب، ۰/۷-۰/۶، متوسط و ۰/۶-۰/۵، ضعیف)دقت مدل جنگل تصادفی در طبقه خوب قرار گرفت (۵۰). با توجه به تعداد زیاد فاکتورهای موثر بر وقوع زمین‌لغزش استفاده از مدل‌های نوین مانند جنگل تصادفی موجب دقت نقشه‌های حاصله می‌شود و محققین دیگری نیز از این مدل استفاده نموده و دقت آن را تایید نموده‌اند (۵،۱۳،۲۰،۵۲،۵۶).

در سال‌های اخیر به دلیل تغییرات کاربری و تخریب منابع طبیعی، زمین‌لغزش به خطری بزرگ در بسیاری از مناطق دنیا تبدیل شده است. استان گلستان و منطقه مورد مطالعه نیز همواره با مشکل تخریب منابع طبیعی مواجه بوده و شرایط وقوع زمین‌لغزش نیز در این منطقه فراهم است. تهیه نقشه حساسیت زمین‌لغزش ابزار مهمی در شناسایی مناطق حساس و جلوگیری از تحریک این مناطق بوده و توجه زیادی به آن می‌شود. در میان مدل‌های مختلف، استفاده از مدل‌های نوین نظیر جنگل تصادفی و تلفیق آن با GIS نقش به‌سزایی در این تحقیقات داشته و می‌تواند موجب افزایش دقت نقشه‌های تهیه شده گردد. نتایج این پژوهش نشان داد منطقه مورد نظر حساسیت زیادی به‌وقوع زمین‌لغزش داشته و مدل جنگل تصادفی ابزار مناسبی برای کمی‌کردن ارتباط زمین‌لغزش و فاکتورهای موثر بر آن است. بنابراین می‌توان گفت که تهیه نقشه حساسیت زمین‌لغزش ابزار مهمی برای کمک به سیاست‌گذاران، طراحان و مهندسان بوده تا خسارات ناشی از وقوع زمین‌لغزش را کاهش دهند.



شکل ۶- منحنی تشخیص عملکرد نسبی و سطح زیر منحنی مربوط به نقشه پهنه‌بندی حساسیت زمین‌لغزش
Figure 6. Receiver operating characteristic and area underthe curve of landslide susceptibility map

[DOI: 10.29252/jwmr.8.15.161]

منابع

1. Akgun, A., C. Kincaland and B. Pradhan. 2011. Application of remote sensing data and GIS for landslide risk assessment as an environmental threat to Izmir city (West Turkey). Environmental Monitoring and Assessment, http://dx.doi.org/10.1007/s10661-011-2352-8.
2. Akgun, A., E.A. Sezer, H.A. Nefeslioglu, C. Gokceoglu and B. Pradhan. 2012. An easy-touse MATLAB program (MamLand) for the assessment of landslide susceptibility using a Mamdani fuzzy algorithm. Computer and Geoscience, 38: 23-34.
3. Althuwaynee, O.F., B. Pradhan, H.J. Park and J.H. Lee. 2014. A novel ensemble decision tree-based CHi-squared Automatic Interaction Detection (CHAID) and multivariate logistic regression models in landslide susceptibility mapping. Landslides, 11: 1063-1078.
4. Ayalew, L. and H. Yamagishi. 2005. The Application of GIS -based logistic regression for landslide susceptibility mapping in the Kakuda-Yahiko Mountains, central Japan. Geomorphology, 65: 15-31.
5. Baatuuwie, N.B. and I.L. V.Leeuwen. 2011. Evaluations of three classifiers in mapping forest stand types using medium resolution imagery: a case study in the Offense Forest District, Ghana. African Journal of Environmental Science and Technology, 5: 25-36
6. Bai, S.B., J. Wang, G. Lü, P. Zhou, S.S. Hou and S.N. Xu. 2010. GIS-based logistic regression for landslide susceptibility mapping of the Zhongxian segment in the Three Gorges area, China. Geomorphology, 115: 23-31.
7. Bednarik, M., B. Magulová, M. Matys and M. Marschalko. 2010. Landslide susceptibility assessment of the Kral˘ovany–Liptovsky´ Mikuláš railway case study. Physics and Chemistry of the Earth, 35: 162-171.
8. Bednarik, M., I. Yilmaz and M. Marschalko. 2012. Landslide hazard and risk assessment: a case study from the Hlohovec-Sered'landslide area in south-west Slovakia. Natural Hazards, 64: 547-575.
9. Breiman, L. 2001. Random forests Machine learning, 45: 5-32.
10. Breiman, L., J. HFriedman, R.A. Olshen and C.J. Stone. 1984. Classification and regression trees. Chapman & Hall, New York.
11. Caniani, D., S. Pascale, F. Sado and A. Sole. 2008. Neural networks and landslide susceptibility: a case study of the urban area of Potenza. Natural Hazards, 45: 55-72.
12. Catani.F., D. Lagomarsino, S. Segoni and V. Tofani. 2013. Landslide susceptibility estimation by random forests technique: sensitivity and scaling issues. Natural Hazards Earth System Science, 13: 2815-2831.
13. Cutler D.R., T.C. Edwards, K.H. Beard, A. Cutler and K.T. Hess. 2007. Random Forests for Classification in Ecology. Ecology, 88: 2783-2792.
14. Champati Ray, D.P., S. Dimri, R.C. Lakhera and S. Sati. 2007. Fuzzy-based method for landslide hazard assessment in active seismic zone of Himalaya. Landslides, 4: 101-111.
15. Chauhan, S., M. Sharma and M.A. Arora. 2010. Landslide susceptibility zonation of the Chamoli region, Garhwal Himalayas, using logistic regression model. Landslides, 7: 411-423.
16. Dai, F.C. and C.F. Lee. 2002. Landslide characteristics and slope instability modeling using GIS, Lantau Island, Hong Kong. Geomorphology, 42: 213-228.
17. Demir, G., M. Aytekin, A. Akgun, S.B. Ikizler and O. Tatar. 2013. A comparison of landslide susceptibility mapping of the eastern part of the North Anatolian Fault Zone (Turkey) by likelihood-frequency ratio and analytic hierarchy process methods. Natural Hazards, 65: 1481-1506.
18. Devkota, K.C., A.D. Regmi, H.R. Pourghasemi, K. Yoshida, B. Pradhan, I.C. Ryu, M.R. Dhital and O.F. Althuwaynee. 2013. Landslide susceptibility mapping using certainty factor, index of entropy and logistic regression models in GIS and their comparison at Mugling-Narayanghat road section in Nepal Himalaya. Natural Hazards, 65: 135-165.
19. Eemini, L., F. Catani and N. Casagli. 2005. Artificial neural networks applied to landslide susceptibility assessment. Geomorphology, 66: 327-343.
20. Gislason, P.O., J.A. Benediktsson and J.R. Sveinsson. 2006. Random forests for land cover classification. Pattern Recognition Letters, 27: 294-300.
21. Glade, T. 1997. Establishing the frequency and magnitude of landslide-triggering rainstorm events in New Zealand. Environmental Geology, 35: 160-174.
22. Hasekiogullari, G.D. and M. Ercanoglu. 2012. A new approach to use AHP in landslide susceptibility mapping: a case study at Yenice (Karabuk, NW Turkey). Natural Hazards, 63: 1157-1179.
23. Jaafari, A., A. Najafi, H.R. Pourghasemi, J. Rezaeian and A. Sattarian. 2014. GIS-based frequency ratio and index of entropy models for landslide susceptibility assessment in the Caspian forest, northern Iran. International Journal of Environmental Science and Technology, 11: 909-926.
24. Jebur, M.N., B. Pradhan, H.Z.M. Shafri, Z.M. Yusoff and M.S. Tehrany. 2015. An integrated user-friendly ArcMAP tool for bivariate statistical modeling in geoscience applications. Geoscientific Model Development, 8: 881-891.
25. Joybari, J., A. Kavian and J. Mosaffaei. 2015. Spatial and temporal variations of Tavan landslide movement in Qazvin province. Geography and environmental hazard, 4(16) (In Persian).
26. Lieb, M., B. Glaser and B. Huwe. 2012. Uncertainty in the spatial prediction of soil tex-ture: comparison of regression tree and Random Forest models. Geoderma, 170: 70-79.
27. Mohammady, M., H.R. Moradi, S. Feiznia and H.R. Pourghasemi. 2010. Comparison of the efficiency of certainty factor, information value and AHP models in landslide hazard zonation (case study: part of Haraz Watershed).Journal of Range and Watershed Management. Iranian Journal of Natural Resources, 62: 539-551 (In Persian).
28. Mohammady, M., H.R. Pourghasemi and B. Pradhan. 2012. Landslide susceptibility mapping at Golestan Province, Iran: a comparison between frequency ratio, Dempster-Shafer and weights of evidence models. Journal of Asian Earth Sciences, 61: 221-236.
29. Mosaffaie, J., M. Ownegh., M. Mesdaghi and M. ShariatJafari. 2009. Comparing the efficiency of statistical and empirical landslide hazard zonation models in Alamout watershed. Journal of water and soil conservation, 16: 43-61 (In Persian).
30. Mosaffaie, J. and M. Ownegh. 2011. Landslide hazard zonation by AHP and regression model, Case study: Alamout watershed. Watershed Engineering and Management, 3: 149-158 (In Persian).

[DOI: 10.29252/jwmr.8.15.161]

31. Nefeslioglu, H.A., C. Gokceoglu and H. Sonmez. 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.
32. Nefeslioglu, H.A., E. Sezer, C. Gökçeoğlu, A.S. Bozkır and T.Y. Duman. 2010. Assessment of landslide susceptibility by decision trees in the metropolitan area of Istanbul, Turkey. Mathematical Problems in Engineering, Article ID: 901095
33. Nicodemus, K.K. 2011. Letter to the Editor: On the stability and ranking of predictors from random forest variable importance measures. Briefings in Bioinformatics, 12: 369-373.
34. O'Brien, R.M. 2007. A caution regarding rules of thumb for variance inflation factors Quality and Quantity, 41: 673-690.
35. Oh, H.J. and B. Pradhan. 2011. Application of a neuro-fuzzy model to landslide susceptibility mapping for shallow landslides in a tropical hilly area. Computer and Geoscience, 37: 1264-1276.
36. Ozdemir, A. and T. Altural. 2013. A comparative study of frequency ratio, weights of evidence and logistic regression methods for landslide susceptibility mapping: Sultan Mountains, SW Turkey. Journal of Asian Earth Science, 64: 180-197
37. Pachauri, A.K., P.V. Gupta and R. Chander. 1998. Landslide zoning in a part of the Garhwal Himalayas. Environmental Geology, 36: 325-334.
38. Peters, J., N. Verhoest, R. Samson, P. Boeckx and B. De Baets. 2008. Wetland vegetation distribution modelling for the identification of constraining environmental variables. Landscape Ecology, 23: 1049-1065.
39. Pourghasemi, H.R., B. Pradhan and C. Gokceoglu. 2012a. Application of fuzzy logic and analytical hierarchy process (AHP) to landslide susceptibility mapping at Haraz watershed, Iran Natural Hazards, 63: 965-996.
40. Pourghasemi, H.R., M. Mohammady and B. Pradhan. 2012b. Landslide susceptibility mapping using index of entropy and conditional probability models in GIS: Safarood Basin, Iran. Catena, 97: 71-84.
41. Pourghasemi, H.R., M. Beheshtirad and B. Pradhan. 2014. A comparative assessment of prediction capabilities of modified analytical hierarchy process (M-AHP) and Mamdani fuzzy logic models in Netcad-GIS for forest fire susceptibility mapping. Geomatics, Natural Hazards and Risk, DOI: 10.1080/19475705.2014.984247.
42. Paudel, U. and T. Oguchi. 2014. Implementation of random forest in landslide susceptibility study, a case study of the Tokamachi area, Niigata, Japan. Japan Geoscience Union Meeting, Pcaifico Yokohama, 28th April 2nd May, 2014.
43. Pradhan, B. 2013. A comparative study on the predictive ability of the decision tree, support vector machine and neuro-fuzzy models in landslide susceptibility mapping using GIS. Computer and Geoscience, 51: 350-365.
44. Pradhan, B. and S. Lee. 2010. Delineation of landslide hazard areas on Penang Island, Malaysia, by using frequency ratio, logistic regression, and artificial neural network models. Environmental Earth Sciences, 60: 1037-1054.
45. Rasai, A., K. Khosravi, M. Habibnejad Roshan, A. Heidari and A. Mashayekhan. 2015. Lnadslide hazard zonation using multivariate regression in GIS environment (Case Study: Aghmashhad Watershed, Mazandaran). Watershed Management Research, 12: 205-215 (In Persian).
46. Sezer, E.A. 2010. A computer program for fractal dimension (FRACEK) with application on type of mass movement characterization Computer and Geoscience, 36: 391-396.
47. Simpson, G.L. and H.J.B. Birks. 2012. Tracking environmental change using lake sediments, Springer Publication, 5: 673 pp.
48. Stumpf, A. and N. Kerle. 2011. Object-oriented mapping of landslides using Random Forests. Remote Sensing ofEnvironment, 115: 2564-2577.
49. Toll, D.G. 1996. Artificial Intelligence Applications in Geotechnical Engineering Electronic Journal of Geotechnical Engineering, 27 pp.
50.Trigila, A., F. Catani, N. Casagli, G. Crosta, C. Esposito, P. Frattini, C. Iadanza, D. Lagomarsino, S. Lari, G. ScarasciaMugnozza, S. Segoni, D. Spizzichino and V. Tofani. 2012. The landslide susceptibility map of Italy at 1:1 Million scale. Geophysical Research Abstracts, 14, EGU2012-7655, 2012.
51. Vahidnia, M.H., A.A. Alesheikh, A. Alimohammadi and F. Hosseinali. 2010. A GIS-based neurofuzzy procedure for integrating knowledge and data in landslide susceptibility mapping. Computers and Geosciences, 36: 1101-1114.
52. Vorpahl, P., H. Elsenbeer, M. Marker and B. Schroder. 2012. How can statistical models help to determine driving factors of landslides? Ecological Modelling, 239: 27-39.
53. 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.
54. Yesilnacar, E.K. 2005.The application of computational intelligence to landslide susceptibility mapping in Turkey, Ph.D. Thesis Department of Geomatics the University of Melbourne, 423 pp.
55. Youssef, A.M. 2015.Landslide susceptibility delineation in the Ar-Rayth Area, Jizan, Kingdom of Saudi Arabia, by using analytical hierarchy process, frequency ratio and logistic regression models Environmental Earth Sciences, Doi: 10.1007/s12665-014-4008-9.
56. Youssef, A.M., H.R. Pourghasemi, Z.S. Pourtaghi and M.M. Al-Katheeri. 2015. Landslide susceptibility mapping using random forest, boosted regression tree, classification and regression tree, and general linear models and comparison of their performance at WadiTayyah Basin, Asir Region, Saudi Arabia Landslides, DOI 10.1007/s10346-015-0614-1.
57. Zare, M., H.R. Pourghasemi, M. Vafakhah and B. Pradhan. 2013. Landslide susceptibility mapping at Vaz watershed (Iran) using an artificial neural network model: a comparison between multi-layer perceptron (MLP) and radial basic function (RBF) algorithms. Arabian Journal of Geoscience, 6: 2873-2888.
58. Zare, M., A. Moghaddamnia, S. Tali Khoshk and H. Salmani. 2015. Landslide hazard assessment by using Neuro-fuzzy technique in Vaz Watershed. Watershed Management Research, 11: 101-110 (In Persian).

[DOI: 10.29252/jwmr.8.15.161]

Prioritization of Landslide-Conditioning Factors and its Landslide Susceptibility Mapping using Random Forest New Algorithm (Case Study: A Part of Golestan province)

Majid Mohammady[1] and Hamid Reza Pourghasemi[2]

1- Assistant Professor, Department of Rangeland & Watershed Management, Faculty of Natural Resources, Semnan University, Semnan, Iran. Iran, (Corresponding Author: majid.mohammady@semnan.ac.ir)
2- Assistant Professor, Department of Natural Resources and Environmental Engineering, College of Agriculture, Shiraz University, Shiraz, Iran
Received: July 16, 2016 Accepted: October 27, 2015

Abstract

Landslide as a natural hazard is very dangerous especially in mountainous areas that result in loss of human life and property around the world. Iran is always exposed to landslide hazard especially in the north and west because of climatic and topographic conditions. The aim of this research is prioritization of landslide-conditioning factors and its landslide susceptibility mapping in the part of Golestan Province using random forest new algorithm. At first, landslide locations were identified using field survey, historical report, and Google earth. In total, 78 landslide locations were identified and divided into two parts for modeling (70%) and validation (30%). Eleven factors of landslide-conditioning including slope aspect, altitude, distance from streams, distance from faults, distance from roads, lithology, land use, slope-length, plan curvature, precipitation, and slope angle were prepared. The relationships between the effective factors and the landslide inventory map were calculated using the random forest algorithm, and then landslide susceptibility map was prepared in the GIS environment. Prioritization of landslide-conditioning factors showed that distance from road, distance from faults, and altitude have the most effect on landslide occurrence respectively. Finally landslide susceptibility map produced by random forest model were divided to four susceptibility classes such as low (29.18%), moderate (33.44%), high (24.82%), and very high (12.55%). ROC curve and the area under the curvewere used for accuracy assessment of the prepared map using about 30% of landslides. Results showed that the random forest model produced reasonable accuracy in landslide susceptibility mapping with area under curve of 0.706.

Keywords: GIS, Landslide, Landslide susceptibility map, Random forest

[DOI: 10.29252/jwmr.8.15.161]