

Original Research Paper

Comparison of the efficiency of some machine learning models for mass movement susceptibility mapping (Case study: Chaharmahal and Bakhtiari province)

Sayed Naeim Emami^{*} and Saleh Yousefi¹

¹Chaharmahal and Bakhtiari Agricultural and Natural Resources Research and Education Center, AREEO, Shahrekord, Iran

ARTICLE INFO

Article history:

Received: 2022 June 07

Accepted: 2022 November 14

Available online: 2023 June 22

Keywords:

Landslide

Artificial Intelligence

Zoning

Random Forest

ABSTRACT

Mass movements are among the most dangerous natural hazards in mountainous regions. The present study employs machine learning (ML) models for mass movement susceptibility mapping (MMSM) in Iran based on a comprehensive dataset of 864 mass movements which include debris flow, landslide, and rockfall during the last 42 years (1977–2019) as well as 12 conditional factors. The results of validation stage show that RF (random forest) is the most viable model for mass movement susceptibility maps. In addition, MARS (multivariate adaptive regression splines), MDA (mixture discriminant additive), and BRT (boosted regression trees) models also provide relatively accurate results. Results of the AUC for validation of produced maps were 0.968, 0.845, 0.828, and 0.765 for RF, MARS, MDA, and BRT, respectively. Based on MMSM generated by RF model, 32% of study area is identified to be under high and very high susceptibility classes. Most of the endangered areas for mass movement are in the west and central parts of the Chaharmahal and Bakhtiari Province. In addition, our findings indicate that elevation, slope angle, distance from roads, and distance from faults are critical factors for mass movement. Our results provide a perspective view for decision makers to mitigate natural hazards.

1. Introduction

Erath's geomorphological activities never ending completely (Corenblit et al., 2011). The Earth surface is formed by various exogenous (geological process, runoff water, glaciers movement, wind, and water) and endogenous (earthquake, metamorphism, volcanic activities, and deformation) activities (Dietrich et al., 1992; Murray et al., 2009). As a natural hazard processes, the landslides can destroy houses, agricultural lands, roads and rails, power networks and dams (Blaschke et al., 2000; Alimohammadlou et al., 2013; Kennedy et al., 2015). The

RF is a nonparametric technique based on regression trees (Hawryło et al., 2018; Thanh Noi and Kappas, 2018; Arabameri et al., 2019b). Among different machine learning techniques, the RF model is one of the strongest models that it has made by a lot of trees (Kim et al., 2018; Gayen et al., 2019). In this context, Shahabi et al. (2019) stated that the RF model can estimate the relative importance of the factors in the best possible way and helps in discussing environmental management. So far, there has been no research comparing machine learning models in

* Corresponding author: Sayed Naeim Emami; E-mail: emami1348@yahoo.com

Citation:

Emami, S. N., and Yousefi, S., 2023. Comparison of the efficiency of some machine learning models for mass movement susceptibility mapping (Case study: Chaharmahal and Bakhtiari province). Scientific Quarterly Journal, GEOSCIENCES, 33(2), 128, 183-204. <https://doi.org/10.22071/gsj.2022.345954.2003>

E-ISSN: 2645-4963; Copyright©2023 G.S. Journal & the authors. All rights reserved.

doi: 10.22071/gsj.2022.345954.2003

 dor: 20.1001.1.10237429.1402.33.2.5.4



This is an open access article under the by-nc/4.0/ License (<https://creativecommons.org/licenses/by-nc/4.0/>)

mass movement sensitivity zoning in Chaharmahal and Bakhtiari province, which is one of the most prone provinces in the country regarding mass movements, and this research seeks to fill this research gap. In general, this research pursues the following goals: 1) Preparation of mass movement sensitivity map (rock fall, landslide and debris flow) based on machine learning algorithms in Chaharmahal and Bakhtiari provinces 2) Attempt to analyze A comparison of the performance of different machine learning (ML) algorithms for the preparation of a sensitivity map of mass movements 3) Identification and prioritization of various factors of domain instabilities in the preparation of a sensitivity map.

2. Research methodology

The investigated area includes the entire area of Chaharmahal and Bakhtiari province in the southwest of Iran. The province with an average slope of 42% is one of the roughest provinces in the country, this province exposure to the Zagros Tectonic Zone and the occurrence of erosion-susceptible rocks have exacerbated the erosion and led to a variety of mass movements. According to the presented flowchart (Fig. 2) this study in generally follow five main stages: 1) collecting and preparing the spatial data, 2) identify the conditional and effective factors on landslide event, 3) LAM producing based on seven used ML algorithms; 4) evaluation of the produced LSMs, and 5) selecting the most accurate and best model for LSM in Chaharmahal and Bakhtiari Province.

2.1. Modelling and mapping using ML algorithms

The RF is a nonparametric technique based on regression trees (Hawryło et al., 2018; Thanh Noi and Kappas, 2018; Arabamiri et al., 2019b). Among different machine learning techniques, the RF model is one of the strongest models that it has made by a lot of trees (Kim et al., 2018; Gayen et al., 2019). As well as, this model has several advantage including; it is never sensitivity to noise, it can accept different layers with different effects and it appropriately identifies the importance of the layers of influence and shows their character (Ruppert, 2004). In this context, Shahabi et al. (2019) stated that the RF model can estimate the relative importance of the factors in the best possible way, and helps in discussing environmental management. In order to have an acceptable and accurate landslide susceptibility map in an area, it is important to have an accurate and good distributed sampling. So that, in the current research, the landslides were collected and recorded by extensive field surveys during 42 years (1977-2019) using Global Position System (GPS) (Fig. 3). In present study, 194, 302 and 368 samples were recorded for debris flow, landslide and rock fall, respectively. Then based on random selection approach, the landslide locations were divided into two categories; a) 70 percent for modeling process and b) 30 percent

for validation process. Since, the landslide events occurred by an interaction of different exogenous and anthropogenic factors, in the present study, 12 effective factors including; digital elevation model (DEM), slope angle, distance from roads, distance from the faults, drainage density, convergence index (CI), land use type, distance from rivers, lithology, plan curvature, topographic wetness index (TWI) and aspect were selected to predict landslide occurrence. In the present study, a 1:25,000 topographic map and 1:100,000 geological map was used to produce and extract 12 effective layers in Arc GIS 10.4.2. In the present study, to produce susceptibility maps of landslides by RF model a special package applied in the R software version R 3.5.3. The used packages were “brt”, (Achour and Pourghasemi, 2019), “fda”, “mda” (Ruppert, 2004), “glm” (Ghanbarian et al., 2019), “MARS” (Deichmann et al., 2002), “RF” (Hosseinalizadeh et al., 2019b), and “SVM” (Achour and Pourghasemi, 2019). The R software is very popular statistical software that performs best analysis and then drawing its analysis graphically for modeling process (Pertille et al., 2019; Taylor et al., 1987). The area under the curve (AUC) is one of the best and most popular methods for model evaluation in machine learning approach (Merow et al., 2013; Pourghasemi et al., 2017). In present study, based on the 30 % of the landslide samples the produced susceptibility maps were evaluated.

3. Results

As mentioned in methodology section in the present study, the AUC value was used to evaluate seven used ML models, the RF can predict the mass movement locations by variable factors much better than other models. Based of seven used ML models the low class of landslide susceptibility covered between 12% and 38% of the study area. In addition, the area covered by moderate class is between 27% and 37%. Also, for high class the covered area by seven models is between 20% and 37% of study area. Finally, the covered area by very high class based on seven ML models is between 12% and 18%. However, the class of high and very high in RF model cover 20% (3233.3 km²) and 12% (1908.78 km²) of the Chaharmahal and Bakhtiari Province, respectively.

4. Discussion and Conclusion

In present study, the random forest algorithm (RF) has used in order to determination of the best factor influencing on mass movement events. In many studies have been used random forest to compare factors, and have obtained favorable results (Amiri et al., 2019; Chen et al., 2017b; Rahmati and Pourghasemi, 2016). According to the random forest algorithm, the digital elevation model (DEM) and slope were the most influential factor. In addition, distance from the roads and distance from the faults

are in the next priority of importance. Generally, the causes of mass movements are divided into two main categories including preparatory and triggering casual factors (Popescu, 1994). Based on the most accurate model (RF), 38% and 30% of the study area are located in low and moderate classes of landslide susceptibility.

The safe areas are mostly located in eastern and southern parts of study area. Our findings show that machine learning models have a very good ability for landslide mapping of mountainous areas. However, these models need good and accurate observed or recorded data and conditional factors.

مقایسه کارایی برخی از مدل‌های یادگیری ماشین در تهیه نقشه حساسیت‌پذیری به حرکات توده‌ای (مطالعه موردی: استان چهارمحال و بختیاری)

 سیدنعیم امامی^{۱*} و صالح یوسفی^۱
^۱ مرکز تحقیقات و آموزش کشاورزی و منابع طبیعی استان چهارمحال و بختیاری، سازمان تحقیقات، آموزش و ترویج کشاورزی، شهرکرد، ایران

چکیده

حرکات توده‌ای در زمره خطرناک‌ترین حوادث طبیعی در مناطق کوهستانی به شمار می‌روند. پژوهش حاضر از مدل‌های یادگیری ماشین (ML) برای تهیه نقشه حساسیت به حرکات توده‌ای در استان چهارمحال و بختیاری استفاده می‌کند. این مدل‌ها بر پایه مجموعه اطلاعات جامع ۸۶۴ حرکت توده‌ای شامل جریان واریزه، زمین لغزش و ریزش سنگ در طول ۴۲ سال گذشته (۱۳۵۶ تا ۱۳۹۷ خورشیدی) و همچنین ۱۲ عامل موثر در رخداد این حرکات، مورد بررسی، آزمون و اعتبارسنجی قرار گرفته‌اند. نتایج اعتبارسنجی نشان می‌دهد که روش RF (Random Forest) جنگل تصادفی مناسب‌ترین مدل برای تهیه نقشه‌های حساسیت به حرکات توده‌ای است. افزون بر این، روش‌های MDA (Mixture Discriminant Additive)، MARS (Multivariate Adaptive Regression Splines) و BRT (Boosted Regression Trees) نیز نتایج نسبتاً دقیقی ارائه داده‌اند. نتایج سطح زیر منحنی (AUC) برای اعتبارسنجی روش‌های RF، MARS و MDA به ترتیب ۰/۹۶۸، ۰/۸۴۵ و ۰/۸۲۸ است. برپایه نقشه حساسیت به حرکات توده‌ای تولید شده توسط مدل RF، ۳۲ درصد از سطح استان در کلاس‌های حساسیت بالا و بسیار بالا شناسایی شده است. بیشتر مناطق در معرض خطر حرکات توده‌ای در باختر و مرکز استان چهارمحال و بختیاری واقع هستند. افزون بر این، یافته‌های این پژوهش نشان می‌دهد که ارتفاع، زاویه شیب، فاصله از جاده‌ها و فاصله از گسل‌ها عوامل بحرانی در وقوع حرکات توده‌ای هستند. نتایج این پژوهش رهیافتی را برای کاهش خسارات ناشی از خطرات طبیعی ارائه می‌کند.

اطلاعات مقاله

تاریخچه مقاله:

تاریخ دریافت: ۱۴۰۱/۰۳/۱۷

تاریخ پذیرش: ۱۴۰۱/۰۸/۲۳

تاریخ انتشار: ۱۴۰۲/۰۴/۰۱

کلیدواژه‌ها:

زمین لغزش

روش مصنوعی

پهنه‌بندی

جنگل تصادفی (RF)

۱- پیش‌نوشتار

حرکات توده‌ای فرایندهای زمین ریخت‌شناسی فعال در سطح زمین تلقی می‌شوند (Corenblit et al., 2011; Dietrich et al., 1992; Murray et al., 2009) و در زمره خطرناک‌ترین مخاطرات طبیعی تهدیدکننده جان و مال انسان‌ها هستند (Chaytor et al., 2007). افزون بر این نقش قابل ملاحظه‌ای در تکامل زمین‌ریخت‌ها و چشم‌اندازهای طبیعی و سرزمینی ایفا می‌کنند (Moeyersons et al., 2008; Fort et al., 2009). همچنین ناپایداری‌های دامنه‌ای بر فعالیت‌های انسانی تأثیرات قابل توجهی بر جای می‌گذارند (Yamada et al., 2012). حرکات توده‌ای می‌توانند خانه‌ها، زمین‌های کشاورزی، جاده‌ها و ریل‌ها، شبکه‌های برق و سدها را ویران کنند (Blaschke et al., 2000; Kennedy et al., 2015; Alimohammadlou et al., 2013). ناپایداری دامنه‌ای ممکن است با سرعت زیاد (لغزش و ریزش سنگ) و یا به آهستگی (خزش) رخ دهد (Wenzel, 2019; Tiranti and Cremonini, 2019). زمین لغزش‌ها از نظر وسعت، ممکن است در مقیاس چند متر مربع تا چندین کیلومتر مربع رخ دهند (Mastere, 2020; Santos et al., 2020). این ناپایداری‌ها در هر منطقه براساس نوع حرکت و جنس مصالح طبقه‌بندی می‌شوند

(Santos et al., 2020). روش‌های متنوع تجربی و ریاضی جهت ارزیابی حساسیت به زمین لغزش ابداع شده که در این بین روش‌های یادگیری ماشین، به‌ویژه روش جنگل تصادفی (Random Forest) RF یکی از قوی‌ترین مدل‌های مبتنی بر درخت‌واره‌ها می‌باشد (Kim et al., 2018; Gayen et al., 2019). روش RF یک روش ناپارامتریک مبتنی بر درخت رگرسیون است (Hawryło et al., 2018; Arabameri et al., 2019 a,b). این مدل مزایای برجسته‌ای همچون عدم حساسیت به عوامل مزاحم، توانایی پذیرش و مدیریت لایه‌های متفاوت و تعیین اولویت و اهمیت ممتاز هرلایه را دارا است (Ruppert, 2004). شهابی (Shahabi et al., 2019) تصریح کرده است که مدل RF می‌تواند اهمیت نسبی عوامل موثر را به بهترین نحو ممکن تخمین بزند و به مدیریت محیطی و سرزمینی کمک کند.

تاکنون هیچ پژوهشی مبنی بر مقایسه مدل‌های یادگیری ماشین در پهنه‌بندی حساسیت حرکات توده‌ای در استان چهارمحال و بختیاری که یکی از مستعدترین استان‌های کشور در خصوص حرکات توده‌ای می‌باشد، صورت نگرفته است و این پژوهش در پی پرکردن این خلا تحقیقاتی است. به‌طور کلی این پژوهش اهداف

* نویسنده مسئول: سیدنعیم امامی؛ E-mail: emami1348@yahoo.com

ماخذنگاری:

امامی، س. ن. و صالحی، ی.، ۱۴۰۲، مقایسه کارایی برخی از مدل‌های یادگیری ماشین در تهیه نقشه حساسیت‌پذیری به حرکات توده‌ای (مطالعه موردی: استان چهارمحال و بختیاری)، فصلنامه علمی علوم زمین، ۳۳، (۲)، ۱۲۸، <https://doi.org/10.22071/gsj.2022.345954.2003>



doi: 10.22071/gsj.2022.345954.2003



dor: 20.1001.1.10237429.1402.33.2.5.4



This is an open access article under the by-nc/4.0/ License (<https://creativecommons.org/licenses/by-nc/4.0/>)

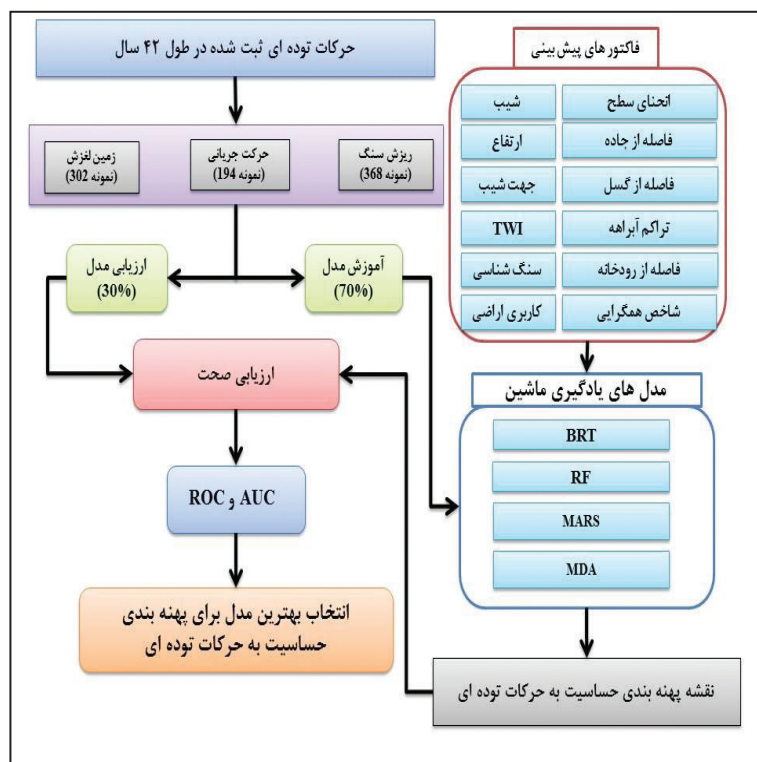
حقوق معنوی مقاله برای فصلنامه علوم زمین و نویسندگان مقاله محفوظ است.

نفر است. این استان شامل ۱۱ شهرستان (اردل، بروجن، شهرکرد، فارس، سامان، بن، لردگان، کیار، کوهرنگ، خانمیرزا و فلارد) است.

با توجه به ویژگی‌های زمین‌شناسی، توپوگرافی و شرایط اقلیمی منطقه مورد مطالعه، انواع مخاطرات طبیعی از جمله سیل، حرکات توده‌ای و زلزله در این استان متداول است. این استان دارای شیب متوسط ۴۲ درصد با بارندگی سالانه از ۱۶۲ میلی‌متر در بخش خاوری تا ۳۰۷۹ میلی‌متر در بخش شمال باختری است (حقیقیان و همکاران، ۲۰۲۰). منطقه مورد مطالعه در پهنه زمین‌ساختی زاگرس قرار گرفته و به لحاظ ویژگی‌های منحصر به فرد سنگ‌شناختی و زمین‌ساختی مستعد انواع حرکات توده‌ای می‌باشد.

۲-۲- روش پژوهش

همان‌طور که در نمودار جریان زیر ارائه شده، این پژوهش در ۵ مرحله به انجام رسیده است (شکل ۲: ۱) جمع‌آوری و آماده‌سازی داده‌های مکانی؛ (۲) شناسایی عوامل موثر بر وقوع حرکات توده‌ای؛ (۳) تولید نقشه حساسیت‌پذیری به حرکات توده‌ای بر اساس چهار الگوریتم یادگیری ماشین (ML)؛ (۴) ارزیابی نقشه‌های تهیه شده در الگوریتم‌های مختلف و (۵) انتخاب دقیق‌ترین و بهترین مدل برای تهیه نقشه حساسیت‌پذیری حرکات توده‌ای در استان چهارمحال و بختیاری.



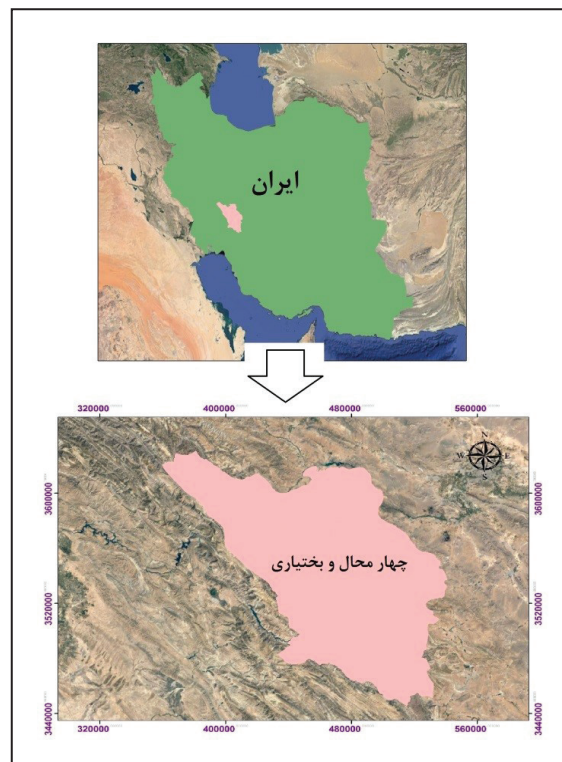
شکل ۲- نمودار جریانی (Flowchart) روش پژوهش.

Figure 2. Flowchart of the methodology in present study.

۲-۲- مواد و روش‌ها

۲-۱- منطقه مورد مطالعه

محدوده مورد بررسی شامل کل مساحت استان چهارمحال و بختیاری در جنوب باختر ایران است (شکل ۱). این استان در امتداد رشته‌کوه زاگرس با ارتفاع متوسط ۲۱۵۳ متر بالاتر از سطح دریا واقع شده است. کمینه و بیشینه ارتفاع استان به ترتیب ۷۷۸ متر و ۴۲۰۳ متر است. استان چهارمحال و بختیاری به وسعت ۱۶۵۳۳ کیلومتر مربع، یک درصد مساحت کل کشور را دربر می‌گیرد و جمعیت آن حدود ۹۴۷۰۰۰



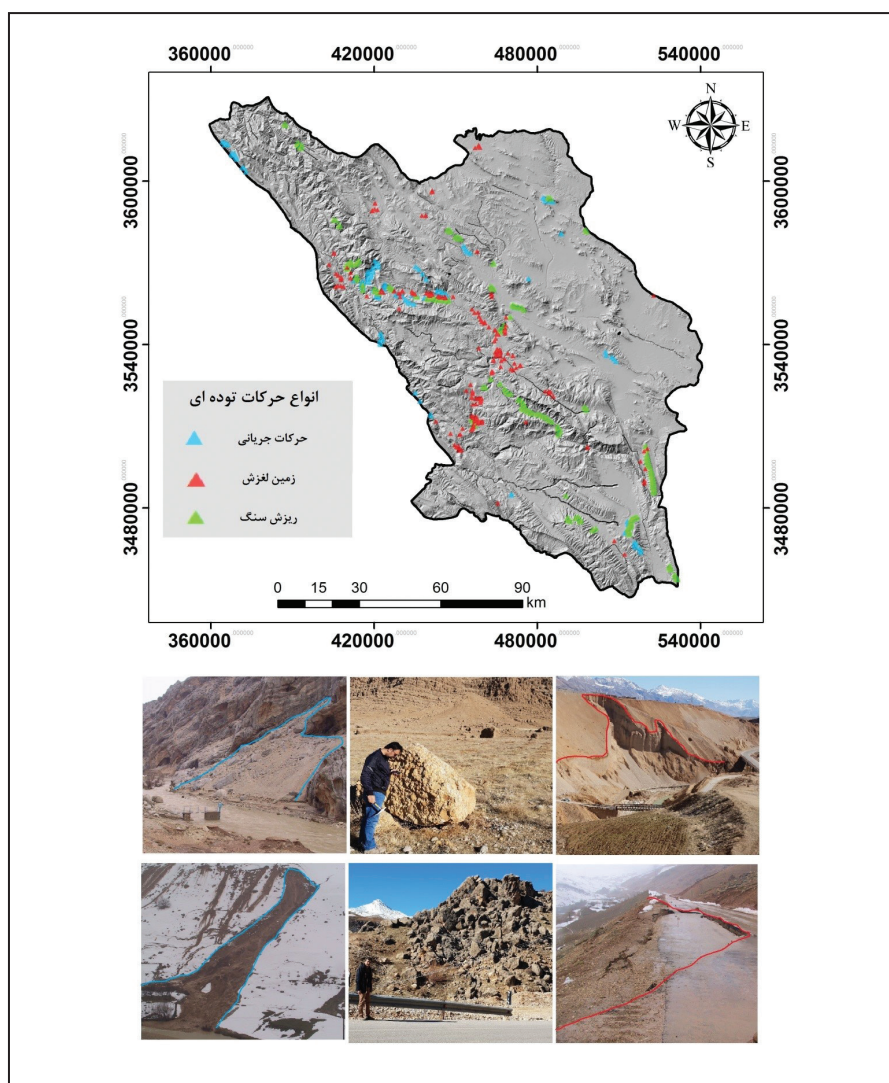
شکل ۱- موقعیت جغرافیایی استان چهارمحال و بختیاری.

Figure 1. Geographical location of the study area (Chaharmahal and Bakhtiari Province).

۳-۲- جمع‌آوری و آماده‌سازی داده‌های مکانی

ارتفاع، زاویه شیب، فاصله از جاده، فاصله از گسل، تراکم زهکشی، شاخص همگرایی (CI)، پوشش زمین، فاصله از رودخانه‌ها، سنگ‌شناسی، انحای پلان، شاخص رطوبت توپوگرافی (TWI) و جهت شیب برای پیش‌بینی وقوع حرکات توده‌ای انتخاب شدند. برای تولید و استخراج ۱۲ لایه موثر در ArcGIS 10.4.2 از نقشه‌های توپوگرافی و زمین‌شناسی در مقیاس ۱:۲۵۰۰۰ و ۱:۱۰۰۰۰۰ استفاده شد. لازم به یادآوری است که نقشه‌های توپوگرافی ۱:۲۵۰۰۰ به عنوان جدیدترین و دقیق‌ترین نقشه‌های پایه توپوگرافی بوده که براساس عکس‌های هوایی ۱:۴۰۰۰۰ تهیه شده‌اند و دارای دقت مکانی بسیار بالایی هستند. پس از تهیه این نقشه پایه و با هدف تدقیق و به روزرسانی عوارض نوزمین‌شناختی به‌ویژه حرکات توده‌ای از تصاویر ماهواره ای جدید از جمله Aster استفاده شد. افزون براین، نقشه پوشش زمین بر اساس تصاویر Sentinel-2 برای سال ۲۰۱۸ توسط روش ماشین بردار پشتیبان (SVM) و با دقت کلی ۸۶ درصد تولید شد (Yousefi, 2016). جدول ۱ منابع، مقیاس، نوع و وضوح عوامل احتمالی که برای فرایند مدل‌سازی ML استفاده شده است را نشان می‌دهد.

به منظور تولید یک نقشه دقیق از حساسیت‌پذیری به حرکات توده‌ای در منطقه مورد مطالعه، سه نوع از این حرکات شامل جریان واریزه، زمین‌لغزش و ریزش سنگ با استفاده از تفسیر عکس‌های هوایی، گزارشات اداره کل منابع طبیعی و آبخیزداری استان مربوط به حرکات توده‌ای ۴۲ سال اخیر (۱۳۵۶ تا ۱۳۹۸) و بازدیدهای گسترده و دقیق صحرایی گردآوری و ثبت شد (شکل ۳). براین پایه، ۱۹۴ مورد جریان واریزه، ۳۰۲ مورد زمین‌لغزش و ۳۶۸ مورد ریزش سنگ ثبت شد. بر اساس بانک اطلاعات ثبت شده حداقل، حداکثر و میانگین مساحت تحت پوشش وقایع حرکات توده‌ای ثبت شده ۳۹۴، ۱۷۹۳۰ و ۶۴۲ متر مربع است. بر اساس رویکرد انتخاب تصادفی پورقاسمی و همکاران (Pourghasemi et al., 2017) و طالب و همکاران (Taalab et al., 2018) تعداد نقاط مکانی حرکات توده‌ای به دو دسته شامل ۷۰ درصد نقاط برای فرایند مدل‌سازی و ۳۰ درصد برای مرحله اعتبارسنجی تقسیم و مورد تجزیه و تحلیل قرار گرفتند. از آنجایی که وقوع حرکات توده‌ای حاصل برهمکنش عوامل مختلف برون‌زا و انسان ساخته هستند، ۱۲ عامل موثر شامل



شکل ۳- پراکنش حرکات توده‌ای ۴۲ ساله (۱۳۵۶ تا ۱۳۹۸) استان چهارمحال و بختیاری و تصاویری از نمونه این حرکات.

Figure 3. Distribution of the recorded landslide samples during 42 years (1977-2019) in study area.

جدول ۱- اطلاعات تکمیلی از عوامل مشروط حرکات توده‌ای.

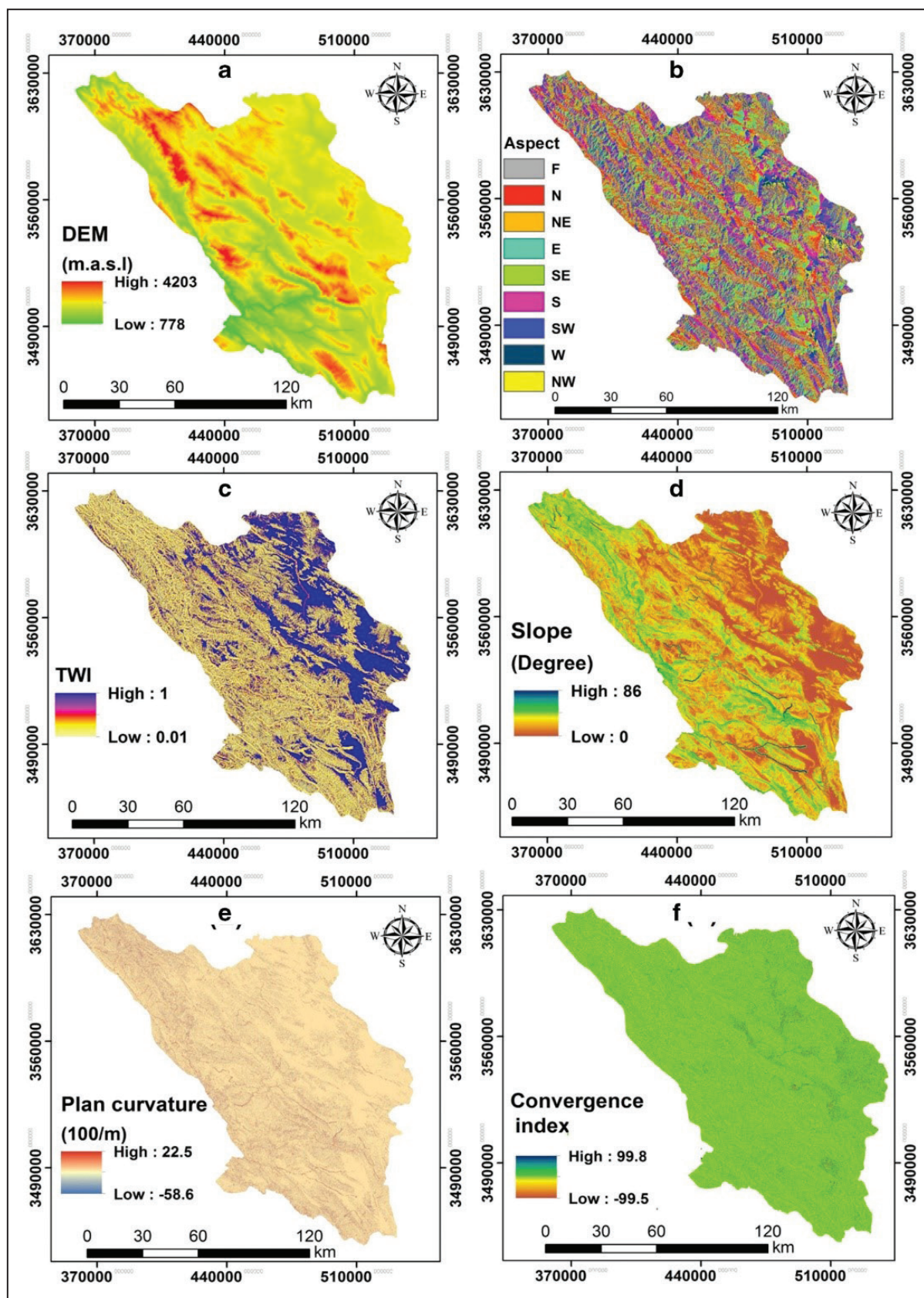
Table1. Additional information on conditional factors of mass movements.

عامل	مقیاس	نوع	اندازه پیکسل (متر)	منبع	سال تولید
ارتفاع	۱:۲۵۰۰۰	رستر	۲۰	نقشه توپوگرافی	۱۳۵۶
جهت شیب	۱:۲۵۰۰۰	رستر	۲۰	نقشه توپوگرافی	۱۳۵۶
TWI	۱:۲۵۰۰۰	رستر	۲۰	نقشه توپوگرافی	۱۳۵۶
شیب	۱:۲۵۰۰۰	رستر	۲۰	نقشه توپوگرافی	۱۳۵۶
منحنی سطح	۱:۲۵۰۰۰	رستر	۲۰	نقشه توپوگرافی	۱۳۵۶
شاخص همگرایی	۱:۲۵۰۰۰	رستر	۲۰	نقشه توپوگرافی	۱۳۵۶
فاصله از رودخانه	۱:۲۵۰۰۰	رستر	۲۰	نقشه توپوگرافی	۱۳۵۶
تراکم زهکشی	۱:۲۵۰۰۰	رستر	۲۰	نقشه توپوگرافی	۱۳۵۶
فاصله از جاده	۱:۲۵۰۰۰	رستر	۲۰	نقشه توپوگرافی + تصاویر ماهواره‌ای	۱۳۹۷
کاربری اراضی	۱:۵۰۰۰۰	وکتور	۲۰	تصاویر ماهواره‌ای سنتینل-۲	۱۳۹۷
سنگ‌شناسی	۱:۵۰۰۰۰	وکتور	۲۰	نقشه زمین‌شناسی	۱۳۷۱
فاصله از گسل	۱:۱۰۰۰۰۰	رستر	۲۰	نقشه زمین‌شناسی	۱۳۷۱

۴-۲- شناسایی عوامل مشروط و موثر بر خطر وقوع حرکات توده‌ای

بر اساس مطالعات پیشین برنینگ (Brenning, 2005)؛ فیضی‌زاده و بلاشک (Ahmed and Dewan, 2017)؛ احمد و دوان (Feizizadeh and Blaschke, 2012)؛ و شیرزادی و همکاران (Shirzadi et al., 2017) روشن است که تعامل و هم‌افزایی عوامل توپوگرافی به تشدید حرکات توده‌ای کمک می‌کند. در مطالعه حاضر، ارتفاع، جهت، TWI (شاخص رطوبت توپوگرافی)، شیب و انحنا پلان بر اساس نقشه‌های توپوگرافی در مقیاس ۱:۲۵۰۰۰ با تفکیک فضایی ۲۰ متر ایجاد و مدل رقومی ارتفاع (DEM) تولید شد (شکل ۴-A-E). بارندگی نیز یک عامل هم‌افزایی برای وقوع حرکات توده‌ای است. با توجه به این‌که استان چهارمحال و بختیاری یک منطقه کوهستانی است و در این مناطق (در مقیاس استان) بارش سالانه با ارتفاع رابطه مستقیم دارد از لایه بارش استفاده نشد. در مطالعه حاضر، بانک اطلاعاتی از حرکات

توده‌ای که در چهار دهه گذشته رخ داده است، جمع‌آوری شد. همچنین، شاخص همگرایی (CI) به عنوان یک عامل مبتنی بر زمین برای استان چهارمحال و بختیاری ایجاد شد (شکل ۴-F). CI نشان‌دهنده ساختار برجسته‌ای است که حاصل واگرایی یا همگرایی لایه‌هاست. افزون بر این، شبکه‌ای از آبراهه‌ها و جاده‌ها براساس نقشه‌های توپوگرافی ۱:۲۵۰۰۰ استخراج و لایه‌های رستری شامل فاصله از رودخانه، تراکم زهکشی و فاصله از جاده‌ها در ArcGIS 10.4.2 تولید شد (شکل ۴-G-I). تصاویر OLI Landsat در ۱۵ ژوئن ۲۰۱۸ برای تهیه نقشه پوشش زمین بر اساس طبقه‌بندی الگوریتم ماشین بردار پشتیبان استفاده شد (شکل ۴-J). فاصله از گسل‌ها و واحدهای سنگ‌شناسی از یک نقشه زمین‌شناسی در مقیاس ۱:۱۰۰۰۰۰ سازمان زمین‌شناسی ایران اقتباس گردید (شکل ۴-K, L).



شکل ۴- لایه‌های مورد استفاده برای منطقه مورد مطالعه: (A) ارتفاع؛ (B) جهت شیب؛ (C) TWI؛ (D) شیب؛ (E) انحنا؛ (F) شاخص همگرایی؛ (G) فاصله از رودخانه؛ (H) تراکم زهکشی؛ (I) فاصله از جاده؛ (J) کاربری اراضی؛ (K) سنگ‌شناسی و فاصله از گسل.

Figure 4. Conditional layers for the study area: A) DEM; B) Aspect; C) TWI; D) Slope; E) Plan curvature; F) Convergence index; G) Distance from the river; H) Drainage density; I) Distance from the road; J) Land use; K) Lithology and L) Distance from fault.

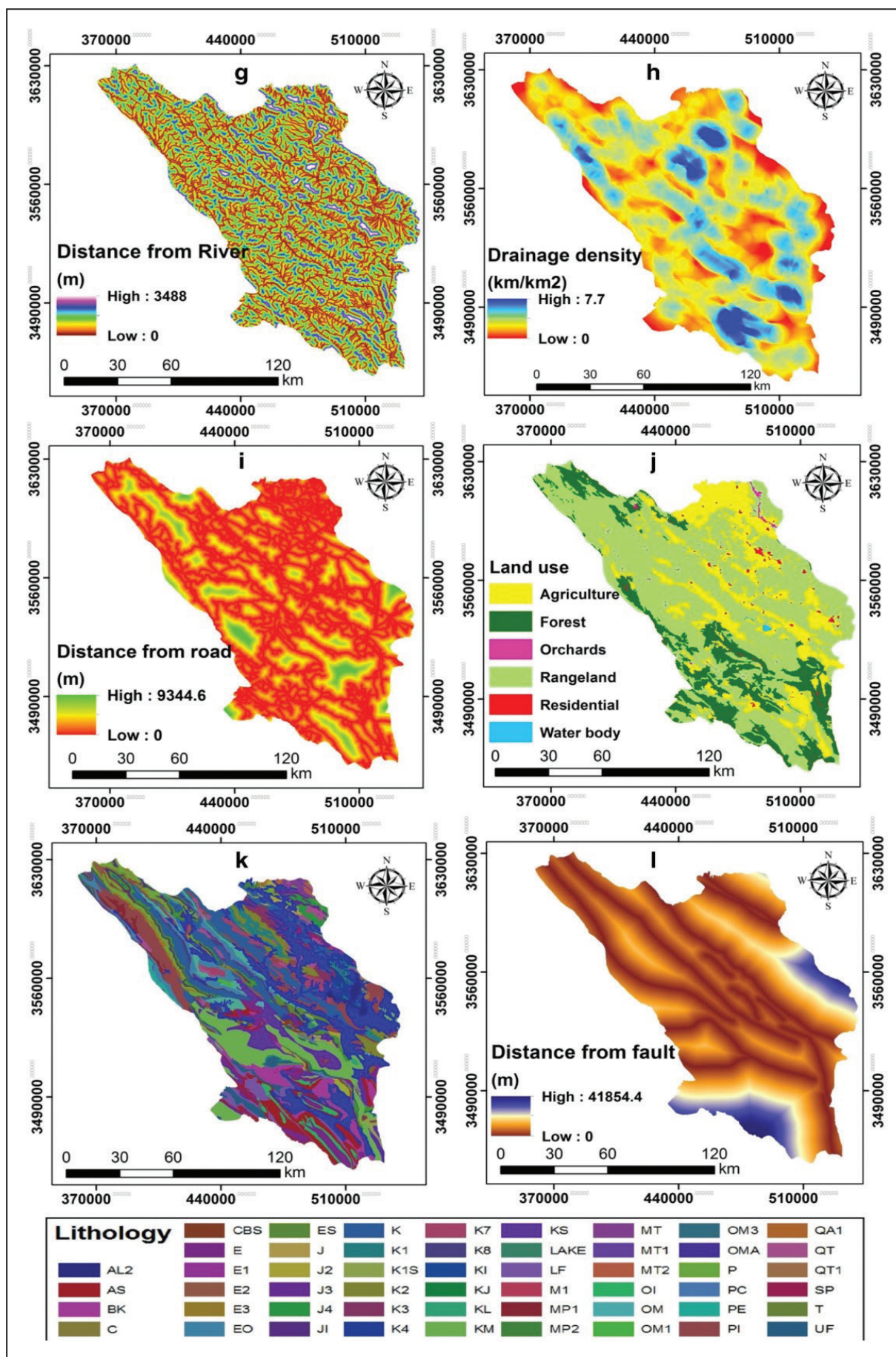


Figure 4. Continue.

متغیرها می‌گویند (Deichmann et al., 2002). در چندین مطالعه پیشین (Corte-Real et al., 1995; Hjort and Luoto, 2013; Bui et al., 2018;) (AI-Sudani et al., 2019) از مدل MARS در مدل‌های متنوع ژئوفیزیک، اقلیم‌شناسی، محیط زیست و ژئومورفولوژی استفاده شده است.

۲-۶- بسته‌های آماری R مورد استفاده برای فرایند مدل‌سازی

در مطالعه حاضر، برای تهیه نقشه‌های حساسیت‌پذیری حرکات توده‌ای توسط مدل‌های MDA، MARS، BRT و RF، برخی بسته‌های خاص در نرم‌افزار R نسخه 3.5.3 استفاده شدند. این بسته‌ها شامل «brt» (Achour and pourghasemi, 2019) «MARS» (Deichmann et al., 2002) «MDA» (Ruppert, 2004) و «RF» (Hosseinizadeh et al., 2019a) می‌باشند. نرم‌افزار R یک نرم‌افزار آماری بسیار محبوب است که بهترین تحلیل را انجام می‌دهد و سپس تجزیه و تحلیل آن را به صورت گرافیکی برای فرایند مدل‌سازی رسم می‌کند (Taylor et al., 1987; Pertille et al., 2019).

۲-۷- اهمیت متغیرها

در پژوهش حاضر با استفاده از عملگر شکست و انتخاب حداقل مطلق (LASSO)، اهمیت عوامل مستقل (ارتفاع، زاویه شیب، فاصله از جاده‌ها، فاصله از گسل‌ها، تراکم زهکشی، CI، پوشش زمین، فاصله از رودخانه‌ها، سنگ‌شناسی، انحنای پلان، TWI و نما) در پیش‌بینی حرکات توده‌ای مورد تجزیه و تحلیل قرار گرفت (Simon et al., 2013). LASSO یک روش مبتنی بر رگرسیون است که انتخاب و ساماندهی مجدد متغیرها را در مدل‌های یادگیری ماشین و برای اولین بار ارائه داده و در مطالعات ژئوفیزیک معرفی شده است (Santosa and Symes, 1986). LASSO یک روش بسط یافته ساده برای طیف گسترده‌ای از مدل‌های یادگیری ماشین است (Ruppert, 2004). افزون بر این، استفاده از پناهی، این روش بر مقدار ضرایب رگرسیون تأثیر می‌گذارد (Lee et al., 2015).

۲-۸- فرایند ارزیابی مدل‌ها

برای شناسایی دقیق‌ترین مدل، فرایند ارزیابی برای هر چهار مدل انجام شد. روش اندازه‌گیری ناحیه زیر منحنی (AUC) یکی از بهترین و متداول‌ترین روش‌ها برای ارزیابی مدل‌های یادگیری ماشین است (Merow et al., 2013; Pourghasemi et al., 2017; Pourghasemi et al., 2020b). در مطالعه حاضر بر اساس ۳۰ درصد نمونه‌ها، نقشه‌های حساسیت تولید شده مورد ارزیابی قرار گرفت (Achour and Pourghasemi, 2019; Rahmati et al., 2019a; Pourtaghi et al., 2015). افزون بر این، ضرایب حساسیت، دقت و صحت برای مرحله اعتبارسنجی چهار مدل ML مورد استفاده محاسبه شد (Parker, 2001; et al., 2011 Visa; Deng et al., 2016).

۳- نتایج

۳-۱- آزمون هم خطی

به منظور حصول اطمینان از عدم تأثیر و همبستگی عوامل موثر بر فرایند مدل‌سازی حرکات توده‌ای، از آزمون چند خطی استفاده شد. آزمون همبستگی پیرسون نشان می‌دهد که هیچ همبستگی معنی‌داری بین عوامل انتخاب شده مستقل وجود ندارد، بنابراین نقشه حساسیت‌پذیری حرکات توده‌ای دقیق‌تری به دست می‌آید (شکل ۵). افزون بر این، نتایج نشان‌دهنده همبستگی مثبت بسیار قوی (۰/۵۱) بین ارتفاع و فاصله از جاده می‌باشد. با این حال، همبستگی بین فاکتورهای شیب و TWI به شدت منفی (۰/۴۴) است. به طور کلی، نتایج حاصل از همبستگی خطی نشان می‌دهد که ضرایب همبستگی برای همه ۱۲ عامل استفاده شده کمتر از ۰/۵ بوده و بین آنها همبستگی معنی‌داری وجود ندارد.

۲-۵- مدل‌سازی و تهیه نقشه با استفاده از الگوریتم‌های یادگیری ماشین (ML)

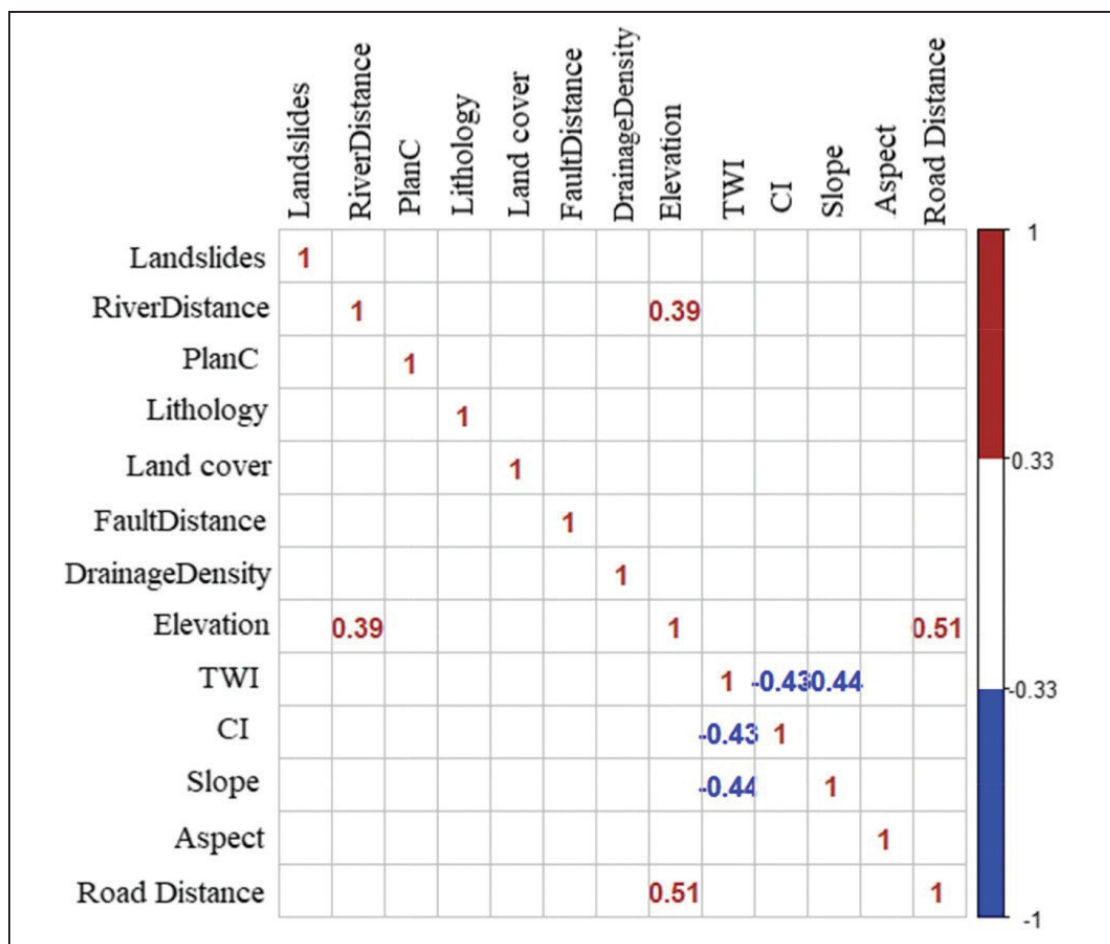
- **درخت رگرسیون تقویت شده:** مدل درخت رگرسیون تقویت شده (BRT) یک مدل ترکیبی با کارایی بالا است (Elith et al., 2008) و معمولاً با روش‌های مبتنی بر ساختار درختی کار می‌کند. بنابراین، این مدل دارای بالاترین توان پیشگویی در مورد عوامل مستقلی است که بیشترین تأثیر را بر عوامل وابسته دارند. افزون بر این، مدل BRT می‌تواند داده‌های مفقود شده را مدیریت کند (Youssef, 2016).

یکی دیگر از مزایای مدل BRT این است که می‌تواند بین مدل‌ها با عملکرد آنها تعادل ایجاد کند (Friedman et al., 2000; Ruppert, 2004). به طور کلی، عملکرد مدل BRT به عوامل مختلفی بستگی دارد، مانند تعداد درختان استفاده شده در مدل و همچنین ترکیب درختان مختلف تعداد درختان بیشتر در مدل، عملکرد بهتری دارد (Reineking and Schröder, 2006).

- **آنالیز تشخیص مخلوط:** آنالیز تشخیص مخلوط (MDA) به عنوان یک روش طبقه‌بندی توسط هستی و طب شیرانی (Hastie and Tibshirani, 1996) ارائه شد. در این الگوریتم، تخمین دانسیته برای هر کلاس بر اساس مخلوط‌های گاوسی است (Kalantar et al., 2019). اساساً حداکثرسازی انتظار، پارامترهای اصلی را تخمین زده و تعداد مولفه‌ها را پیش‌بینی می‌کند و این فرایند با استفاده از مخلوطی از گاوسی برازش‌های محلی را برای هر کلاس به دست می‌آورد (Ju et al., 2003). MDA یک تکنیک کاربردی برای مدل‌سازی‌ها با رابطه غیر عادی یا غیرخطی برای بهبود دقت طبقه‌بندی یا تعیین زیر کلاس‌های پنهان در هر گروه است (Hosseinizadeh et al., 2019b; Kalantar et al., 2019). بنابراین، MDA یک مدل طبقه‌بندی ناپارامتریک است که انتظار می‌رود برای روابط پیچیده به خوبی کارایی داشته باشد.

- **جنگل تصادفی:** جنگل تصادفی (RF) یک روش ناپارامتریک مبتنی بر درخت رگرسیون است (Hawrylo et al., 2018; Thanh Noi and Kappas, 2018; Arabameri et al., 2019b). در میان روش‌های مختلف یادگیری ماشین، مدل RF یکی از قوی‌ترین مدل‌هایی است که توسط تعداد زیادی درخت‌واره ساخته شده است (Kim et al., 2018; Gayen et al., 2019). این مدل دارای چندین مزیت از جمله عدم حساسیت به نویز، قابلیت پذیرش لایه‌های مختلف با افکت‌های متفاوت و شناسایی اهمیت لایه‌های تأثیرگذار و ویژگی آن‌ها است (Ruppert, 2004). شهابی و همکاران (Shahabi et al., 2019) اشاره کردند که مدل RF می‌تواند اهمیت نسبی عوامل را به بهترین شکل ممکن تخمین بزند و به مدیریت محیطی کمک کند. در مطالعه حاضر، شاخص‌های اصلی مدل RF عبارتند از: «mtry» = ۳ و «ntree» = ۱۰۰۰ و نرخ خطا ۱۶/۳۱٪ می‌باشد.

- **خطوط رگرسیون تطبیقی چند متغیره:** خطوط رگرسیون تطبیقی چند متغیره (MARS) یکی از با کیفیت‌ترین مدل‌های مبتنی بر رگرسیون است و می‌تواند بر اساس روابط خطی و غیر خطی بین عوامل مستقل را پیش‌بینی کند (Gu and Wahba, 1991). این مدل از انعطاف‌پذیری بسیار خوبی برای پیش‌بینی یک رویداد بر اساس متغیر وابسته به متغیرهای ورودی مستقل برخوردار است و هدف اصلی آن تعیین دقیق تأثیر مجموعه متغیرهای مستقل بر متغیر وابسته است (Adnan et al., 2019; Gayen et al., 2019; Zabihi et al., 2019). یکی از ویژگی‌های مهم مدل MARS این است که رابطه عملکردی بین متغیرهای وابسته و مستقل را با مجموعه‌ای از ضرایب تعریف کرده به طوری که تأثیر آن‌ها را به طور جداگانه محاسبه می‌کند (Busto Serrano et al., 2020). به همین دلیل، الگوریتم MARS به عنوان روشی شناخته می‌شود که توابع اساسی را با فواصل زمانی متفاوت از متغیرهای مستقل تعریف می‌کند (Deichmann et al., 2002; Lazarus and Constantine, 2013). مشابه دیگر مدل‌ها، MARS دارای یک ویژگی اساسی است که به آن حساسیت خاص نسبت به همبستگی



شکل ۵- آزمون همبستگی پیرسون برای ۱۲ عامل مورد استفاده.

Figure 5. Pearson correlation test for 12 effective factors.

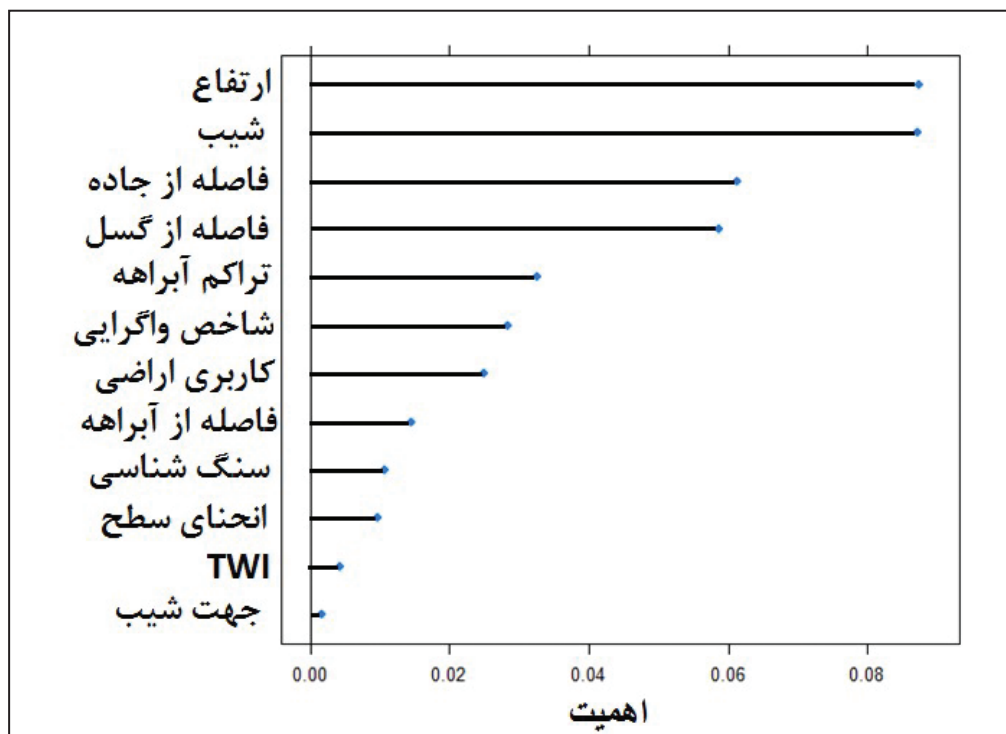
۳-۲- اهمیت متغیرها

حرکات توده‌ای بالاتر در استان چهارمحال و بختیاری کمک می‌کند. شکل ۸ نقشه‌های حساسیت حرکات توده‌ای تهیه‌شده از چهار مدل ML را نشان می‌دهد. به طور کلی، نقشه‌های تولید شده در این پژوهش نشان می‌دهد که حرکات توده‌ای یک خطر بالقوه در منطقه مورد مطالعه است. نقشه‌های تولید شده نشان می‌دهند که بخش مرکزی و باختری استان در طبقات بالا و بسیار بالای خطر قرار دارند در حالی که بخش‌های جنوبی و خاوری استان در کلاس‌های خطر کم و متوسط واقع شده‌اند. شکل ۹ نشان می‌دهد که بر اساس چهار مدل یادگیری ماشینی استفاده شده، بین ۱۲ تا ۳۸ درصد از مساحت منطقه مورد مطالعه (استان چهارمحال و بختیاری) در کلاس پایین حساسیت نسبت به حرکات توده‌ای قرار گرفته است. افزون بر این، منطقه تحت پوشش طبقه متوسط بین ۲۷ و ۳۷ درصد مساحت منطقه را در برمی‌گیرد. به همین ترتیب بین ۲۰ تا ۳۷ درصد سطح استان در رده خطر بالا و ۱۲ تا ۱۸ درصد در رده خطر بسیار بالا قرار گرفته است. براساس مدل RF، کلاس‌های با خطر بالا و خیلی بالا به ترتیب ۲۰ درصد (معادل ۳۲۳۳/۳ کیلومتر مربع) و ۱۲ درصد (معادل ۱۹۰۸/۷۸ کیلومتر مربع) از سطح منطقه مورد مطالعه را در برمی‌گیرد.

در پژوهش حاضر اهمیت متغیرها با استفاده از مدل LASSO انجام شده است. بر اساس تجزیه و تحلیل LASSO (شکل ۶)، لایه‌های ارتفاعی و شیب بیشترین اهمیت را برای پیش‌بینی حرکات توده‌ای دارند. افزون بر این، فاصله از جاده‌ها و فاصله از گسل‌ها در اولویت بعدی مدل‌سازی حرکات توده‌ای هستند. از سوی دیگر، جهت شیب، TWI، انحنا سطح و سنگ‌شناسی عواملی هستند که نقش چندان مهمی در وقوع حرکات توده‌ای در منطقه مورد مطالعه ندارند. نتایج میانگین کاهش دقت با استفاده از مدل RF این نکته را تایید کرد که ارتفاع مهم‌ترین عامل حرکات توده‌ای است در حالی که فاصله از گسل‌ها، شیب و فاصله از جاده‌ها در رتبه‌های بعدی قرار دارند (شکل ۷).

۳-۳- نقشه‌های حساسیت‌پذیری حرکات توده‌ای

در پژوهش حاضر برای درک بهتر نقشه حساسیت حرکات توده‌ای با توجه به الگوریتم طبیعی شکست (NB) در ArcGIS 10.4.2، چهار مدل به چهار کلاس شامل کم، متوسط، زیاد، و بسیار زیاد طبقه‌بندی شدند (Lai and Khan, 2012; Pourghasemi et al., 2019a). این نوع طبقه‌بندی به شناسایی مناطق با پتانسیل



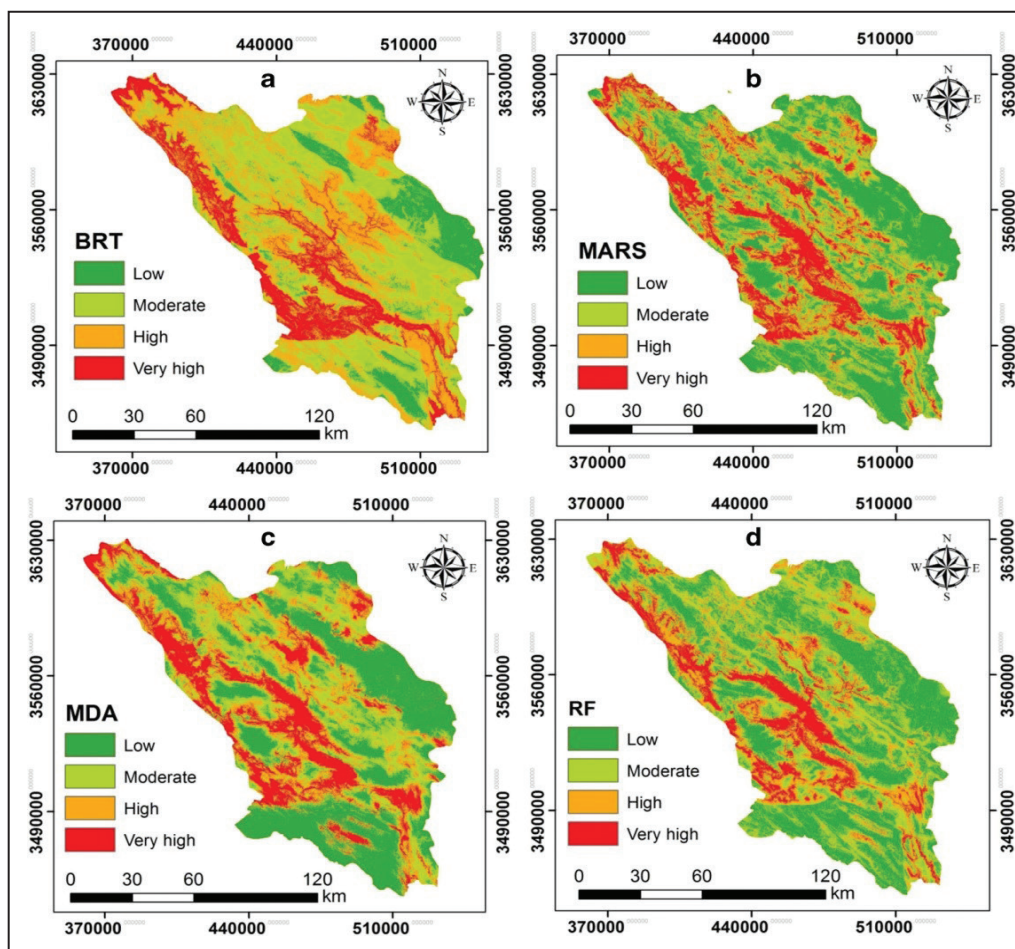
شکل ۶- تجزیه و تحلیل LASSO از اهمیت لایه‌های پیش‌بینی.

Figure 6. LASSO analysis from importance of prediction layers.



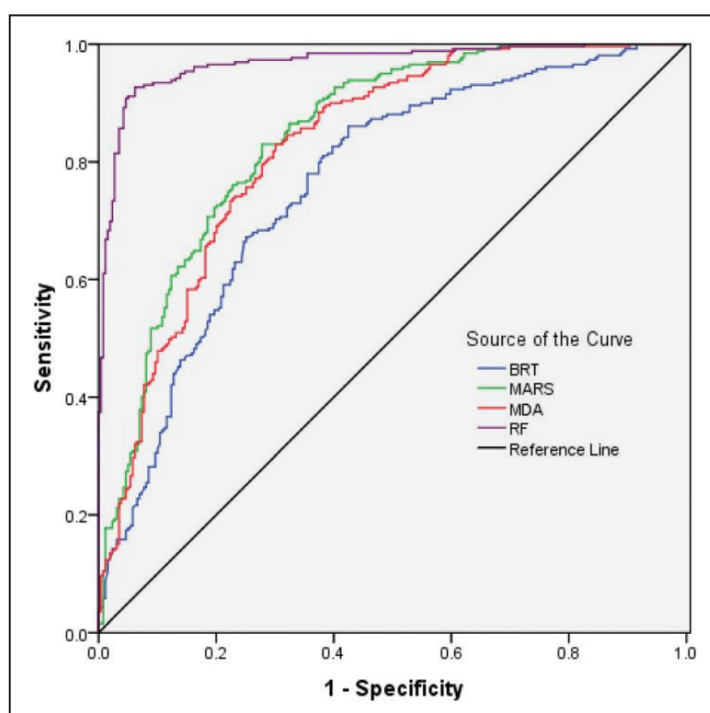
شکل ۷- نتایج اهمیت متغیرها با استفاده از مدل RF.

Figure 7. Results of importance of variables using RF model.



شکل ۸- نقشه حساسیت‌پذیری استان چهارمحال و بختیاری به حرکات توده‌ای.

Figure 8. Landslide susceptibility maps in the Chaharmahal and Bakhtiari province. (a: BRT, b: MARS, c: MDA, d: RF).



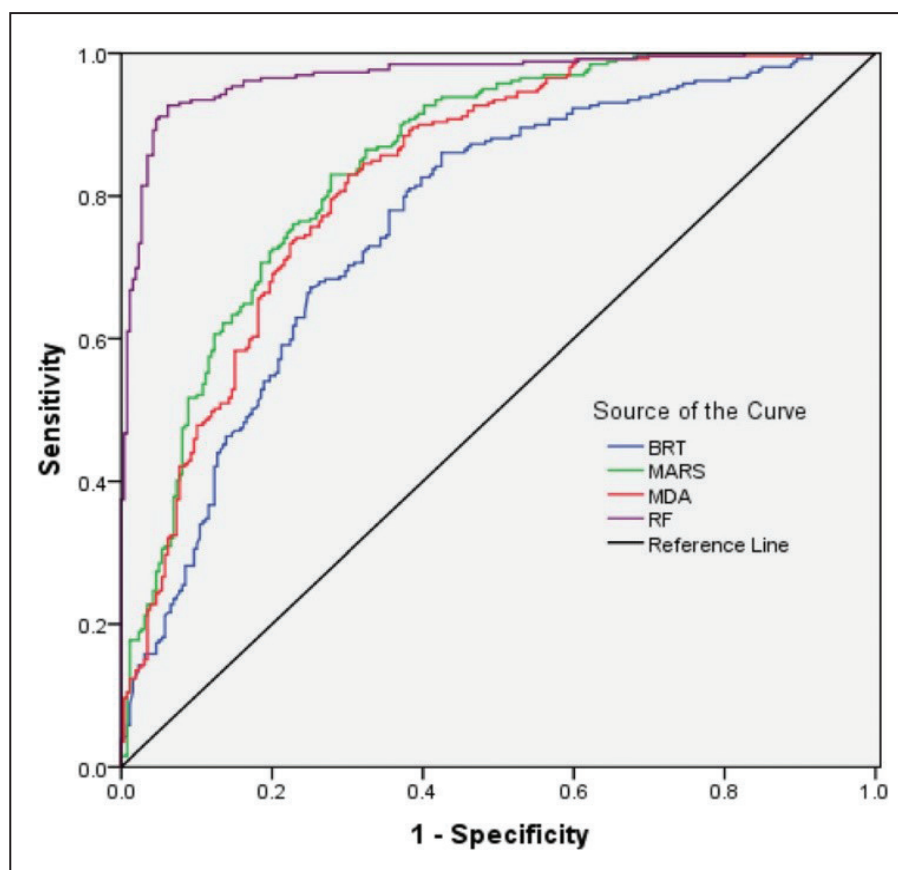
شکل ۹- توزیع حساسیت به حرکات توده‌ای در منطقه مورد مطالعه توسط چهار مدل یادگیری ماشینی (ML).

Figure 9. Distribution of the landslide susceptibility in study area by seven used ML models.

۴-۳ - مقایسه و عملکرد مدل‌ها

عوامل متغیر، بسیار بهتر از مدل‌های دیگر پیش‌بینی کند. افزون بر این، مدل‌های MDA، MARS و BRT به ترتیب در رتبه‌های بعدی قرار می‌گیرند. براساس مقادیر AUC، حساسیت، دقت و صحت، MDA، MARS و BRT نیز برای پیش‌بینی حرکات توده‌ای مناسب هستند. در مقابل مدل BRT کمترین دقت را دارا می‌باشد (جدول ۲ و شکل ۱۰).

همانطور که در بخش روش پژوهش بیان شد، برای مقایسه کارایی مدل‌ها مساحت زیر منحنی (AUC)، دقت، صحت و مقادیر ضرایب صحت برای ارزیابی چهار مدل ML استفاده شد. همان طور که در جدول ۲ نشان داده شده است، مدل RF در بین دیگر الگوریتم‌های یادگیری ماشین برای مدل‌سازی حرکات توده‌ای بهترین مدل است. این بدان معناست که RF می‌تواند مکان‌های حرکات توده‌ای را با



شکل ۱۰- نتایج ROC برای مدل‌های یادگیری ماشین جهت تهیه نقشه حساسیت‌پذیری حرکات توده‌ای.

Figure 10. The ROC results for mass movement susceptibility mapping models.

جدول ۲- نتایج محاسبه مساحت زیر منحنی (AUC).

Table 2. Calculation results of the area under the curve (AUC).

Models	Area	Standard error	Sensitivity	Specificity	Accuracy	Asymptotic significant	Asymptotic 95% confidence interval	
							Lower bound	Upper bound
RF	0/968	0/007	0/9266	0/9302	0/9284	0/000	0/954	0/9821
MDA	0.828	0.018	0.7412	0.7558	0.7485	0.0000	۰/۷۹۳	۰/۸۶۳
MARS	0.845	0.017	0.7606	0.7674	0.7610	0.0000	۰/۸۱۲	۰/۸۷۹
BRT	0.765	0.021	0.6757	0.7364	0.7060	0.0000	۰/۷۲۴	۰/۸۰۶

۴- بحث

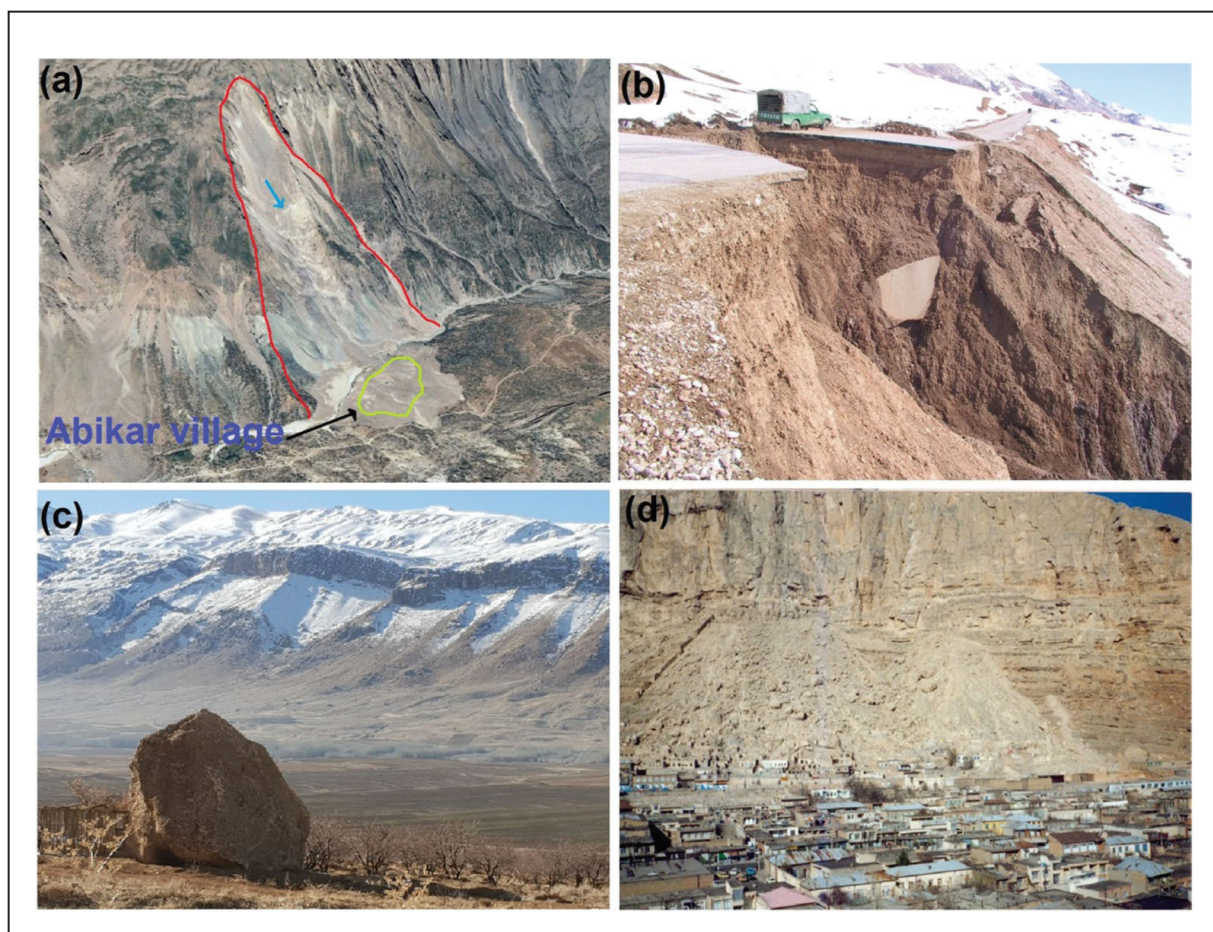
برای پیش‌بینی حساسیت به حرکات توده‌ای در استان چهارمحال و بختیاری استفاده شد. نتایج نشان می‌دهد که RF دقیق‌ترین مدل برای تولید این نقشه مهم است. در برخی از مطالعات پیشین، مدل BRT به عنوان یکی از پایدارترین مدل‌ها با نتایج قابل قبول در مقایسه با مدل‌های CART (درخت طبقه‌بندی و رگرسیون)، EBF (تابع باور شواهدی)، GLM (مدل خطی تعمیم‌یافته) و مدل‌های RF تعیین و معرفی شد (Rahmati et al., 2016; Vilar et al., 2016; Amiri et al., 2019; Ghanbarian et al., 2019).

نقیبی و پورقاسمی (Naghibi and pourghasemi, 2015) نشان دادند که برخی از روش‌های یادگیری ماشین مانند BRT، CART و RF عملکرد گسترده‌ای برای پیش‌بینی آب‌های زیرزمینی دارند. از مدل‌های RF، BRT و GAM (مدل افزودنی تعمیم‌یافته) برای پیش‌بینی مکان آتش‌سوزی‌های احتمالی در مراتع استفاده شد (Achour and pourghasemi, 2019; Amiri et al., 2019; Avand et al., 2019; Nabiollahi et al., 2019).

تاکنون بیش از ۸۰۰ مورد از حرکات توده‌ای در استان چهارمحال و بختیاری ثبت شده است. جاده‌سازی غیراصولی، تخریب پوشش گیاهی و جنگل‌زدایی از جمله علل مهم این ناپایداری‌های دامنه‌ای هستند (Imaizumi and Sidle, 2012; Yousefi et al., 2016; Chen et al., 2017b).

در فروردین ۱۳۷۴، طی یک بارندگی شدید، حرکات توده‌ای فاجعه‌باری در روستای آبیکار استان چهارمحال و بختیاری رخ داد و ۵۲ نفر در زیر خروارها سنگ و خاک حاصل از زمین‌لغزش دفن و کشته شدند (شکل ۱۱- a).

سالانه تعداد قابل توجهی از جاده‌های اصلی (شکل ۱۱- b)، زمین‌های کشاورزی (شکل ۱۱- c) و خانه‌ها (شکل ۱۱- d) به دلیل وقوع حرکات توده‌ای در منطقه مورد مطالعه تخریب می‌شوند. بنابراین، فرموله‌سازی یک مدل قابل قبول برای تهیه نقشه حساسیت به حرکات توده‌ای می‌تواند به حفاظت از زندگی انسان‌ها کمک کرده و از آسیب‌ها جلوگیری نماید. در این پژوهش از چهار مدل یادگیری ماشین (MARS، MDA، BRT و RF)



شکل ۱۱- نمونه‌هایی از خسارات ناشی از حرکات توده‌ای در استان چهارمحال و بختیاری. a) دفن روستای آبیکار با ۵۲ کشته؛ b) ویرانی جاده در منطقه صمصامی؛ c) خسارت به زمین‌های باغی در اثر ریزش سنگ؛ d) تخریب منازل مسکونی روستایی در اثر ریزش سنگ و واریزه.

Figure 11. Samples of scathes by landslide events in the Chaharmahal and Bakhtiari province; a) buried the Abikar village with 52 killed people; b) Road destruction in the Samsami region; c) Damages in orchard lands by rock fall and d) Destroyed hoses by a landslide event.

نقشه حرکات توده‌ای در نظر گرفته شود و عدم قطعیت فرایند مدل‌سازی کاهش یابد. عموماً علل حرکات توده‌ای به دو دسته اصلی تقسیم می‌شوند که عبارتند از عوامل زمینه‌ساز یا مقدماتی و علل ماشه‌ای یا محرک (Popescu, 1994). با توجه به شکل ۶ و ۷، ارتفاع، شیب، فاصله از گسل، کاربری زمین و سنگ‌شناسی عوامل اصلی زمینه‌ساز یا مقدماتی حرکات توده‌ای هستند. به همین ترتیب فاصله از جاده، تراکم زهکشی و فاصله از رودخانه از جمله عوامل ماشه‌ای یا محرک اصلی به شمار می‌آیند (Korup and Stolle, 2014; Suresh et al., 2019). افزون بر این، با توجه به ارتباط بین ارتفاع و میزان بارندگی و نقش محرک بارندگی به‌ویژه بارندگی شدید در شروع ناپایداری (Krishnakumar et al., 2009; Rahmani et al., 2011). در مطالعه حاضر نیز همبستگی قوی بین بارندگی و ارتفاع وجود داشت و بر این اساس لایه بارش به عنوان یک عامل شرطی از فرایند مدل‌سازی حذف شد. ارتفاع را می‌توان یک عامل ماشه‌ای نیز در نظر گرفت. یکی از نکات مهم این است که علیرغم نقش مهم سنگ‌شناسی در زمینه‌سازی حرکات توده‌ای، نتیجه مطالعه حاضر، اهمیت کمی برای این عامل نشان داده است. دلیل این امر، یکنواختی نسبی سازندهای زمین‌شناسی است به‌گونه‌ای که بخش عمده‌ای از سطح استان را سنگ‌های رسوبی کربناته به‌ویژه سنگ‌آهک پوشانده است.

بر این اساس، در این منطقه تفاوت‌های سنگ‌شناسی نمی‌توانند در توزیع مکانی حرکات توده‌ای نقش قوی ایفا کنند (Bednarik et al., 2010). در پهنه زمین‌ساختی زاگرس، شیب نقش مهمی دارد. دامنه‌ها جهت احداث جاده نیز با افزایش شیب دامنه‌ها تا مقداری بیش از زاویه اصطکاک داخلی سنگ‌های تشکیل‌دهنده دامنه، در بروز ناپایداری‌های دامنه‌ای از اهمیت عمده‌ای برخوردار است (Barber et al., 2014). در شمال باختر منطقه مورد مطالعه، تراکم جاده‌ها بیشتر است و بر اساس نقشه‌های تهیه شده (MMSM)، آن منطقه در رده خطر بالا و بسیار بالا واقع شده است. فعالیت گسل و زلزله می‌تواند نقش مهمی در وقوع ناپایداری دامنه داشته باشد و فاصله تا گسل نیز نقش مهمی در وقوع حرکات توده‌ای ایفا می‌کند (Ozdemir and Hong et al., 2015; Moayedi et al., 2019). فاصله از گسل‌ها نیز مبین نقش لرزه‌خیزی در حرکات توده‌ای است. بر اساس لایه‌های زمین‌شناسی و گسل، منطقه مورد مطالعه توسط مجموعه‌ای از گسل‌ها از شمال باختر به سمت جنوب خاوری قطع شده است. به عبارتی، خطر حرکات توده‌ای نیز از الگوی مکانی گسل‌ها و امتداد نوار کوهستانی پیروی می‌کند. افزایش تراکم زهکشی و نفوذ آب به دامنه‌ها، وقوع حرکات توده‌ای به ویژه زمین‌لغزش و جریان واریزه را افزایش می‌دهد (Shi and Jin, 2009; Ahmed and Dewan, 2017). به همین دلیل تراکم زهکشی نقش مهمی در وقوع حرکات توده‌ای دارد.

۵- نتیجه‌گیری

استان چهارمحال و بختیاری در رشته‌کوه‌های زاگرس، منطقه‌ای مستعد خطرات طبیعی است. مطالعه حاضر دقت چهار مدل عمومی یادگیری ماشین (BRT، MARS، MDA و RF) را برای تهیه نقشه حساسیت‌پذیری به حرکات توده‌ای یک منطقه کوهستانی را مورد بررسی قرار داده است. نتایج نشان می‌دهد که مقادیر AUC برای اعتبارسنجی نقشه‌های تولیدشده حساسیت به حرکات توده‌ای برای مدل‌های MARS، MDA، RF و BRT به ترتیب ۰/۸۸۸، ۰/۸۲۱، ۰/۸۱۷ و ۰/۷۴۷ است. افزون بر این، نتایج نشان می‌دهد که هر چهار مدل دارای یک الگوی کلی جهت تهیه نقشه حساسیت‌پذیری به خطر هستند اما در جزئیات متفاوت هستند.

بخش‌های مرکزی و باختری منطقه مورد مطالعه که کوهستانی‌تر هستند، پتانسیل بیشتری برای خطر حرکات توده‌ای دارند. بر اساس دقیق‌ترین مدل (RF)، به ترتیب ۳۸ و ۳۰ درصد از منطقه مورد مطالعه در کلاس‌های خطر کم و متوسط حساسیت حرکات توده‌ای قرار دارند. مناطق امن بیشتر در بخش‌های خاوری و جنوبی استان واقع می‌باشند. انواع مختلف حرکات توده‌ای (زمین‌لغزش، جریان واریزه و ریزش

به طور کلی، برخی از مدل‌های مبتنی بر درخت‌های رگرسیون مانند BRT و RF به تعداد درخت‌ها وابسته هستند، به طوری که تنظیم چنین پارامتری می‌تواند بر کیفیت کار مدل تأثیر بگذارد (Shataee et al., 2011). در این پژوهش همچنین تلاش شد تا از تعداد زیادی درخت در مدل‌ها استفاده شود. برخی از مطالعات روش‌های متعددی را برای پیش‌بینی حوادث طبیعی از جمله آتش‌سوزی، زمین‌لغزش، پراکندگی گونه‌ها و تغییرات زمین‌شناسی مورد استفاده قرار می‌دهند (Hjort and Luoto, 2013).

در برخی موارد، مدل BRT پیش‌بینی دقیق‌تری از مخاطرات طبیعی را نسبت به مدل RF ارائه می‌دهد (Elith et al., 2008; Shataee et al., 2011). Pourtaghi et al., 2016 در حالی که در مطالعات دیگر مدل RF بهترین مدل معرفی شده است (Vorpahl et al., 2012; Korup and Stolle, 2014).

مطالعه دیگری همچنین ثابت کرد که مدل‌های RF و SVM بهترین مدل‌ها برای پیش‌بینی فرسایش یک منطقه هستند (Pourghasemi et al., 2017). پیش از به دست آوردن پاسخ قطعی برای انتخاب بهترین مدل در زمینه‌های مختلف، کار بیشتری در مورد مقایسه مدل‌ها مورد نیاز است. صحت مدل RF در مطالعات مختلف به‌ویژه در مدل‌سازی حساسیت زمین‌لغزش به اثبات رسیده است (Lee and Pradhan, 2007; Gokceglu, 2010; Pourghasemi et al., 2013; Jebur et al., 2014; Pradhan et al., 2014; Amiri et al., 2019). از سوی دیگر، برخی از مطالعات اخیر از مدل RF در زمینه پراکنش گیاهان استفاده کرده و نتایج بسیار خوبی را نشان داده است (Pouteau et al., 2012; Boogar et al., 2019). یکی از مزایای مدل MDA این است که از داده‌های مستقل برای مدل‌سازی استفاده می‌کند و در نتیجه آن را به یک پیش‌بینی‌کننده خوب تبدیل می‌کند (Morris and Nicholas, 2016). در تحقیق حاضر مدل MDA رتبه اول را در بین مدل‌های دیگر به دست نیاورد، اما پیش‌بینی نسبتاً خوبی ارائه داد به‌طوری که می‌توان از نقشه حساسیت تولیدشده برپایه این مدل برای شناسایی توزیع حرکات توده‌ای در منطقه مورد مطالعه استفاده کرد. همچنین، به دلیل دقت بالای آن، نتایج را می‌توان به راحتی تفسیر کرد. بر این اساس از نظر آماری در مقایسه با مدل‌های دیگر عملکرد خوبی دارد (Schmid et al., 2009).

در مدل MARS، بهینه‌ترین متغیرها از نظر تأثیر بر عامل وابسته، طبقه‌بندی شده و بر این اساس مدل اجرا می‌شود (Ozdemir and Altural, 2013). Gutierrez et al., 2014 با این حال، مدل MARS باید با احتیاط قضاوت شود (Gutierrez et al., 2014). در یک مطالعه بر روی زمین‌لغزش، از چهار مدل GLM، MARS و GAM (فرایند سلسله مراتبی تحلیلی اصلاح شده) استفاده شد که مدل MARS عملکرد بسیار خوبی داشته است (Pourtaghi et al., 2015). ذبیحی و همکاران (Zabihi et al., 2019) یکی از مزایای اصلی مدل MARS را امکان مقایسه آسان نتایج آن با مدل‌های ANN (شبکه‌های عصبی مصنوعی) دانستند. در مطالعه حاضر مدل MARS در رتبه سوم قرار گرفت. در تحقیق جاری از LASSO برای تعیین مهم‌ترین عامل موثر بر حرکات توده‌ای استفاده شده است. در بسیاری از مطالعات، از LASSO برای مقایسه عوامل با نتایج مطلوب استفاده شده است (Ruppert, 2004; Reineking and Schröder, 2006; Lee et al., 2015). Kalantar et al., 2019 طبق الگوریتم LASSO، ارتفاع و شیب بیشترین تأثیر را بر وقوع حرکات توده‌ای داشته‌اند. افزون بر این، فاصله از جاده‌ها و فاصله از گسل‌ها در اولویت بعدی قرار می‌گیرند.

انواع مختلف حرکات توده‌ای، سازوکارهای متفاوتی دارند و شناسایی عوامل موثر بر هر یک از آنها بسیار پیچیده است. ریزش سنگ در بیشتر موارد به دلیل زلزله و نیروی جاذبه رخ داده است (Hung et al., 1999; Rahmati et al., 2019b). در حالی که زمین‌لغزش‌ها در هنگام بارندگی شدید و ناپایداری شیب اتفاق افتاده‌اند (Prakasam et al., 2020). از سوی دیگر، جریان واریزه به دلیل رطوبت بیش از حد در لایه‌های فوقانی خاک رخ می‌دهد (Tiranti and Cremonini, 2019). در پژوهش حاضر سعی شده است مجموعه‌ای جامع از عوامل کنترل‌کننده فرایند تهیه

ارائه شود. یافته‌های این پژوهش نشان می‌دهد که مدل‌های یادگیری ماشین به‌طور بالقوه قادر به تهیه نقشه حرکات توده‌ای در مناطق کوهستانی هستند. با این حال، این مدل‌ها نیاز به ارزیابی بر اساس مشاهدات دقیق و عوامل شرطی دارند.

سنگ) از نظر سازوکار وقوع متفاوت هستند اما همه آنها برای زندگی و فعالیت‌های انسان تهدید به‌شمار می‌روند. در پژوهش حاضر، با در نظر گرفتن طیف وسیعی از عوامل مهم در انواع حرکات توده‌ای، سعی شده است مدلی دقیق برای رسم نقشه خطر این سه نوع حرکت توده‌ای به صورت توام و تلفیقی

References

- Achour, Y., and Pourghasemi, H.R., 2019. How do machine learning techniques help in increasing accuracy of landslide susceptibility maps? *Geosci Front.* <https://doi.org/10.1016/j.gsf.2019.10.001>.
- Adnan, R.M., Liang, Z., and Heddad, S., 2019. Least square support vector machine and multivariate adaptive regression splines for streamflow prediction in mountainous basin using hydro-meteorological data as inputs. *J Hydrol* 124371. <https://doi.org/10.1016/J.JHYDROL.2019.124371>.
- Ahmed, B., and Dewan, A., 2017. Application of bivariate and multivariate statistical techniques in landslide susceptibility modeling in Chittagong City Corporation, Bangladesh. *Remote Sens* 9:304. <https://doi.org/10.3390/rs9040304>.
- Alimohammadlou, Y., Najafi, A., and Yalcin, A., 2013. Landslide process and impacts: a proposed classification method. *Catena* 104:219–232. <https://doi.org/10.1016/j.catena.2012.11.013>.
- Al-Sudani, Z., Salih, S.Q., Sharafati, A., and Yaseen, Z.M., 2019. Development of multivariate adaptive regression spline integrated with differential evolution model for streamflow simulation. *J Hydrol* 573:1–12. <https://doi.org/10.1016/j.jhydrol.2019.03.004>.
- Amiri, M., Pourghasemi, H.R., Ghanbarian, G.A., and Afzali, S.F., 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. <https://doi.org/10.1016/j.geoderma.2018.12.042>.
- Arabameri, A., Chen, W., and Loche, M., 2019a. Comparison of machine learning models for gully erosion susceptibility mapping. *Geosci Front.* <https://doi.org/10.1016/j.gsf.2019.11.009>.
- Arabameri, A., Pradhan, B., and Rezaei, K., 2019b. Spatial prediction of gully erosion using ALOS PALSAR data and ensemble bivariate and data mining models. *Geosci J* 23:669–686. <https://doi.org/10.1007/s12303-018-0067-3>.
- Avand, M., Janizadeh, S., and Naghibi, S.A., 2019. A comparative assessment of random forest and k-nearest neighbor classifiers for gully erosion susceptibility mapping. *Water (Switzerland)* 11:2076. <https://doi.org/10.3390/w11102076>.
- Barber, C.P., Cochrane, M.A., Souza, C.M., and Laurance, W.F., 2014. Roads, deforestation, and the mitigating effect of protected areas in the Amazon. *Biol Conserv* 177:203–209. <https://doi.org/10.1016/j.biocon.2014.07.004>.
- Bednarik, M., Magulová, B., Matys, M., and Marschalko, M., 2010. Landslide susceptibility assessment of the Kral'ovany-Liptovský Mikuláš railway case study. *Phys Chem Earth* 35:162–171. <https://doi.org/10.1016/j.pce.2009.12.002>.
- Blaschke, P.M., Trustrum, N.A., and Hicks, D.L., 2000. Impacts of mass movement erosion on land productivity: a review. *Prog Phys Geogr* 24:21–52. <https://doi.org/10.1191/030913300669154532>.
- Boogar, A.R., Salehi, H., Pourghasemi, H.R., and Blaschke, T., 2019. Predicting habitat suitability and conserving *Juniperus* spp. habitat using SVM and maximum entropy machine learning techniques. *Water (Switzerland)* 11:2049. <https://doi.org/10.3390/w11102049>.
- Brenning, A., 2005. Spatial prediction models for landslide hazards: review, comparison and evaluation. *Nat Hazards Earth Syst Sci* 5:853–862. <https://doi.org/10.5194/nhess-5-853-2005>.
- Bui, D.T., Shahabi, H., and Shirzadi, A., 2018. Land subsidence susceptibility mapping in South Korea using machine learning algorithms. *Sensors (Switzerland)* 18:2464. <https://doi.org/10.3390/s18082464>.

- Busto Serrano, N., and Suárez Sánchez, A., and Sánchez Lasheras, F., 2020. Identification of gender differences in the factors influencing shoulders, neck and upper limb MSD by means of multivariate adaptive regression splines (MARS). *Appl Ergon* 82:102981. <https://doi.org/10.1016/j.apergo.2019.102981>.
- Chaytor, J.D., Twichell, D.C., and Ten Brink, U.S., 2007. Revisiting submarine mass movements along the U.S. Atlantic continental margin: implications for tsunami hazards. In: *Submarine Mass Movements and Their Consequences*, 3rd International Symposium. Springer, pp 395–403. https://doi.org/10.1007/978-1-4020-6512-5_41.
- Chen, W., Pourghasemi, H.R., Kornejady, A., and Zhang, N., 2017b. Landslide spatial modeling: introducing new ensembles of ANN, MaxEnt, and SVM machine learning techniques. *Geoderma* 305:314–327. <https://doi.org/10.1016/j.geoderma.2017.06.020>.
- Corenblit, D., Baas, A.C.W., and Bornette, G., 2011. Feedbacks between geomorphology and biota controlling Earth surface processes and landforms: a review of foundation concepts and current understandings. *Earth-Sci Rev* 106:307–331. <https://doi.org/10.1016/j.earscirev.2011.03.002>.
- Corte-Real, J., Zhang, X., and Wang, X., 1995. Downscaling GCM information to regional scales: a non-parametric multivariate regression approach. *Clim Dyn* 11:413–424. <https://doi.org/10.1007/BF00209515>.
- Deichmann, J., Eshghi, A., and Haughton, D., 2002. Application of multiple adaptive regression splines (mars) in direct response modeling. *J Interact Mark* 16:15–27. <https://doi.org/10.1002/dir.10040>.
- Deng, X., Liu, Q., Deng, Y., and Mahadevan, S., 2016. An improved method to construct basic probability assignment based on the confusion matrix for classification problem. *Inf Sci (Ny)* 340:250–261.
- Dietrich, W.E., Wilson, C.J., and Montgomery, D.R., 1992. Erosion thresholds and land surface morphology. *Geology* 20:675–679. [https://doi.org/10.1130/0091-7613\(1992\)020<0675:ETALSM>2.3.CO;2](https://doi.org/10.1130/0091-7613(1992)020<0675:ETALSM>2.3.CO;2).
- Elith, J., Leathwick, J.R., and Hastie, T., 2008. A working guide to boosted regression trees. *J Anim Ecol* 77:802–813. <https://doi.org/10.1111/j.1365-2656.2008.01390.x>.
- Feizizadeh, B., and Blaschke, T., 2012. Comparing GIS-multicriteria decision analysis for landslide susceptibility mapping for the lake basin, Iran. *Int Geosci Remote Sens Symp* 65:5390–5393. <https://doi.org/10.1109/IGARSS.2012.6352388>.
- Fort, M., Cossart, E., and Deline, P., 2009. Geomorphic impacts of large and rapid mass movements: a review. *Impacts géomorphologiques des mouvements de masse volumineux ET rapides: une revue. Groupe français de géomorphologie*. <https://doi.org/10.4000/geomorphologie.7495>.
- Friedman, J., Hastie, T., and Tibshirani, R., 2000. Additive logistic regression: a statistical view of boosting (with discussion and a rejoinder by the authors). *Ann Stat* 28:337–407. <https://doi.org/10.1214/aos/1016218223>.
- Gayen, A., Pourghasemi, H.R., and Saha, S., 2019. Gully erosion susceptibility assessment and management of hazard-prone areas in India using different machine learning algorithms. *Sci Total Environ* 668:124–138. <https://doi.org/10.1016/j.scitotenv.2019.02.436>.
- Ghanbarian, G., Raoufat, M.R., Pourghasemi, H.R., and Safaeian, R., 2019. Habitat suitability mapping of *Artemisia aucheri* Boiss based on the GLM model in R. In: *Spatial Modeling in GIS and R for Earth and Environmental Sciences*. Elsevier, pp 213–227. <https://doi.org/10.1016/B978-0-12-815226-3.00009-0>.
- Gokceoglu, C., Nefeslioglu, H.A., and Sezer, E., 2010. Assessment of landslide susceptibility by decision trees in the metropolitan area of Istanbul, Turkey. *Math Probl Eng* 2010. <https://doi.org/10.1155/2010/901095>.
- Gu, C., and Wahba, G., 1991. Discussion: multivariate adaptive regression splines. *Ann Stat* 19:115–123. <https://doi.org/10.1214/aos/1176347972>.
- Gutierrez, R.R., Abad, J.D., Choi, M., and Montoro, H., 2014. Characterization of confluences in free meandering rivers of the Amazon basin. *Geomorphology* 220:1–14. <https://doi.org/10.1016/j.geomorph.2014.05.011>.
- Hastie, T., and Tibshirani, R., 1996. Discriminant analysis by Gaussian mixtures. *J R Stat Soc Ser B* 58:155–176. <https://doi.org/10.1016/j.jmva.2003.12.003>.
- Hawryło, P., Bednarz, B., Wężyk, P., and Szostak, M., 2018. Estimating defoliation of Scots pine stands using machine learning methods and vegetation indices of Sentinel-2. <https://doi.org/10.1080/22797254.2017.1417745>.

- Hjort, J., and Luoto, M., 2013. Statistical methods for geomorphic distribution modeling. *Treatise Geomorphol* 2:59–73. <https://doi.org/10.1016/B978-0-12-374739-6.00028-2>.
- Hong, H., Pradhan, B., Xu, C., and Tien Bui, D., 2015. Spatial prediction of landslide hazard at the Yihuang area (China) using two-class kernel logistic regression, alternating decision tree and support vector machines. *Catena* 133:266–281. <https://doi.org/10.1016/j.catena.2015.05.019>.
- Hong, H., Panahi, M., and Shirzadi, A., 2018. Flood susceptibility assessment in Hengfeng area coupling adaptive neuro-fuzzy inference system with genetic algorithm and differential evolution. *Sci Total Environ* 621:1124–1141. <https://doi.org/10.1016/j.scitotenv.2017.10.114>.
- Hosseinalizadeh, M., Kariminejad, N., and Rahmati, O., 2019b. How can statistical and artificial intelligence approaches predict piping erosion susceptibility? *Sci Total Environ* 646:1554–1566. <https://doi.org/10.1016/j.scitotenv.2018.07.396>.
- Hungr, O., Evans, S.G., and Hazzard, J., 1999. Magnitude and frequency of rock falls and rock slides along the main transportation corridors of southwestern British Columbia. *Can Geotech J* 36:224–238.
- Imaizumi, F., and Sidle, R.C., 2012. Effect of forest harvesting on hydrogeomorphic processes in steep terrain of Central Japan. *Geomorphology* 169–170:109–122. <https://doi.org/10.1016/j.geomorph.2012.04.017>.
- Jebur, M.N., Pradhan, B., and Tehrany, M.S., 2014. Optimization of landslide conditioning factors using very high-resolution airborne laser scanning (LiDAR) data at catchment scale. *Remote Sens Environ* 152:150–165. <https://doi.org/10.1016/j.rse.2014.05.013>.
- Ju, J., Kolaczyk, E.D., and Gopal, S., 2003. Gaussian mixture discriminant analysis and sub-pixel land cover characterization in remote sensing. *Remote Sens Environ* 84:550–560. [https://doi.org/10.1016/S0034-4257\(02\)00172-4](https://doi.org/10.1016/S0034-4257(02)00172-4).
- Kalantar, B., Al-Najjar, H.A.H., and Pradhan, B., 2019. Optimized conditioning factors using machine learning techniques for groundwater potential mapping. *Water* 11:1909.
- Kennedy, I.T.R., Petley, D.N., Williams, R., and Murray, V., 2015. A systematic review of the health impacts of mass earth movements (landslides). *PLoS Curr* 7. <https://doi.org/10.1371/currents.dis.1d49e84c8bbe678b0e70cf7fc35d0b77>.
- Kim, J.C., Lee, S., Jung, H.S., and Lee, S., 2018. Landslide susceptibility mapping using random forest and boosted tree models in Pyeong-Chang, Korea. *Geocarto Int* 33:1000–1015. <https://doi.org/10.1080/10106049.2017.1323964>.
- Korup, O., and Stolle, A., 2014. Landslide prediction from machine learning. *Geol Today* 30:26–33. <https://doi.org/10.1111/gto.12034>.
- Krishnakumar, K.N., Rao, G.P. and Gopakumar, C.S., 2009. Rainfall trends in twentieth century over Kerala, India. *Atmos Environ* 43:1940–1944. <https://doi.org/10.1016/j.atmosenv.2008.12.053>.
- Lai, W., and Khan, A., 2012. Modeling dam-break flood over natural rivers using discontinuous Galerkin method. *J Hydrodyn* 24:467–478. [https://doi.org/10.1016/S1001-6058\(11\)60268-0](https://doi.org/10.1016/S1001-6058(11)60268-0).
- Lazarus, E.D., and Constantine, J.A., 2013. Generic theory for channel sinuosity. *Proc Natl Acad Sci U S A* 110:8447–8452. <https://doi.org/10.1073/pnas.1214074110>.
- Lee, S., and Pradhan, B., 2007. Landslide hazard mapping at Selangor, Malaysia using frequency ratio and logistic regression models. *Landslides* 4:33–41. <https://doi.org/10.1007/s10346-006-0047-y>.
- Lee, J.D., Sun, Y., and Taylor, J.E., 2015. On model selection consistency of regularized M-estimators. *Electron J Stat* 9:608–642. <https://doi.org/10.1214/15-EJS1013>.
- Mastere, M., 2020. Mass movement hazard assessment at a medium scale using weight of evidence model and neo-predictive variables creation. In: *Mapping and Spatial Analysis of Socio-economic and Environmental Indicators for Sustainable Development*. Springer, pp 73–85. https://doi.org/10.1007/978-3-030-21166-0_7.
- Merow, C., Smith, M.J., and Silander, J.A., 2013. A practical guide to MaxEnt for modeling species' distributions: what it does, and why inputs and settings matter. *Ecography (Cop)* 36:1058–1069. <https://doi.org/10.1111/j.1600-0587.2013.07872.x>.
- Moayedi, H., Mehrabi, M., and Mosallanezhad, M., 2019. Modification of landslide susceptibility mapping using optimized PSO-ANN technique. *Eng Comput* 35:967–984. <https://doi.org/10.1007/s00366-018-0644-0>.
- Moeyersons, J., Van Den Eeckhaut, M., and Nyssen, J., 2008. Mass movement mapping for geomorphological understanding and sustainable development: Tigray, Ethiopia. *Catena* 75:45–54. <https://doi.org/10.1016/j.catena.2008.04.004>.

- Morris, K., and Nicholas, P.D., 2016. Clustering, classification, discriminant analysis, and dimension reduction via generalized hyperbolic mixtures. *Comput Stat Data Anal* 97:133–150. <https://doi.org/10.1016/j>.
- Murray, A.B., Lazarus, E., and Ashton, A., 2009. Geomorphology, complexity, and the emerging science of the Earth's surface. *Geomorphology* 103:496–505 <https://doi.org/10.1016/j.geomorph.2008.08.013>
- Nabiollahi, K., Eskandari, S., and Taghizadeh-Mehrjardi, R., 2019. Assessing soil organic carbon stocks under land-use change scenarios using random forest models. *Carbon Manag* 10:63–77. <https://doi.org/10.1080/17583004.2018.1553434>.
- Naghibi, S.A., and Pourghasemi, H.R., 2015. A comparative assessment between three machine learning models and their performance comparison by bivariate and multivariate statistical methods in groundwater potential mapping. *Water Resour Manag* 29:5217–5236. <https://doi.org/10.1007/s11269-015-1114-8>.
- Ozdemir, A., and Altural, T., 2013. A comparative study of frequency ratio, weights of evidence and logistic regression methods for landslide susceptibility mapping: Sultan mountains, SW Turkey. *J Asian Earth Sci* 64:180–197. <https://doi.org/10.1016/j.jseaes.2012.12.014>.
- Parker, J.R., 2001. Rank and response combination from confusion matrix data. *Inf fusion* 2:113–120. [https://doi.org/10.1016/S1566-2535\(01\)00030-6](https://doi.org/10.1016/S1566-2535(01)00030-6)
- Pertille, R.H., Sachet, M.R., and Guerrezi, M.T., and Citadin, I., 2019. An R package to quantify different chilling and heat models for temperate fruit trees. *Comput Electron Agric* 167:105067. <https://doi.org/10.1016/j.compag.2019.105067>.
- Popescu, M.E., 1994. A suggested method for reporting landslide causes. *Bull Int Assoc Eng Geol l'Association Int Géologie l'Ingénieur* 50:71–74. osti, M. Hashizume, and Masayuki Watanabe <https://doi.org/10.1007/BF02590201>
- Pourghasemi, H.R., Pradhan, B., and Gokceoglu, C., 2013. Application of weights-of-evidence and certainty factor models and their comparison in landslide susceptibility mapping at Haraz watershed, Iran. *Arab J Geosci* 6:2351–2365. <https://doi.org/10.1007/s12517-012-0532-7>.
- Pourghasemi, H.R., Yousefi, S., Kornejady, A., and Cerdà, A., 2017. Performance assessment of individual and ensemble data-mining techniques for gully erosion modeling. *Sci Total Environ* 609:764–775. <https://doi.org/10.1016/j.scitotenv.2017.07.198>.
- Pourghasemi, H.R., Gayen, A., and Edalat, M., 2019a. Is multi-hazard mapping effective in assessing natural hazards and integrated watershed management? *Geosci Front*. <https://doi.org/10.1016/j.gsf.2019.10.008>.
- Pourghasemi, H.R., Gayen, A., and Panahi, M., 2019b. Multi-hazard probability assessment and mapping in Iran. *Sci Total Environ* 692:556–571. <https://doi.org/10.1016/j.scitotenv.2019.07.203>.
- Pourghasemi, H.R., Pouyan, S., Heidari, B., Farajzadeh, Z., Fallah, S.R., Babaei, S., Khosravi, R., Etemadi, M., Ghanbarian, G., Farhadi, A., Safaeian, R., Heidari, Z., Tarazkar, M.H., Tiefenbacher, J.P., Azmi, A., and Sadeghian, F., 2020b. Spatial modelling, risk mapping, change detection, and outbreak trend analysis of coronavirus (COVID-19) in Iran (days between 19 February to 14 June 2020). *Int J Infect is*. <https://doi.org/10.1016/j.ijid.2020.06.058>.
- Pourtaghi, Z.S., Pourghasemi, H.R., and Rossi, M., 2015. Forest fire susceptibility mapping in the Minudasht forests, Golestan province, Iran. *Environ Earth Sci* 73:1515–1533. <https://doi.org/10.1007/s12665-014-3502-4>
- Pourtaghi, Z.S., Pourghasemi, H.R., Aretano, R., and Semeraro, T., 2016. Investigation of general indicators influencing on forest fire and its susceptibility modeling using different data mining techniques. *Ecol Indic* 64:72–84. <https://doi.org/10.1016/j.ecolind.2015.12.030>.
- Pouteau, R., Meyer, J.Y., Taputuarai, R., and Stoll, B., 2012. Support vector machines to map rare and endangered native plants in Pacific islands forests. *Ecol Inform* 9:37–46. <https://doi.org/10.1016/j.ecoinf.2012.03.003>.
- Pradhan, B., Abokharima, M.H., Jebur, M.N., and Tehrany, M.S., 2014. Land subsidence susceptibility mapping at Kinta Valley (Malaysia) using the evidential belief function model in GIS. *Nat Hazards* 73:1019–1042. <https://doi.org/10.1007/s11069-014-1128-1>.
- Prakasam, C., Aravinth, R., Kanwar, V.S., and Nagarajan, B., 2020. Landslide hazard mapping using geo-environmental parameters—a case study on Shimla Tehsil, Himachal Pradesh. In: *Lecture Notes in Civil Engineering*. Springer, pp 123–139. . https://doi.org/10.1007/978-981-13-7067-0_9.
- Rahmani, R., Sadoddin, A., and Ghorbani, S., 2011. Measuring and modelling precipitation components in an Oriental beech stand of the Hyrcanian region, Iran. *J Hydrol* 404:294–303. <https://doi.org/10.1016/j.jhydrol.2011.04.036>.

- Rahmati, O., Pourghasemi, H.R., and Melesse, A.M., 2016. Application of GISbased data driven random forest and maximum entropy models for groundwater potential mapping: a case study at Mehran Region,Iran. *Catena* 137:360–372. <https://doi.org/10.1016/j.catena.2015.10.010>.
- Rahmati, O., Kalantari, Z., Samadi. 2019a. GIS-based site selection for check dams in watersheds: considering geomorphometric and topo-hydrological factors. *Sustain* 11:5639. <https://doi.org/10.3390/su11205639>.
- Rahmati, O., Yousefi, S., and Kalantari, Z., 2019b. Multi-hazard exposure mapping using machine learning techniques: a case study from Iran. *Remote Sens* 11:1943. <https://doi.org/10.3390/rs11161943>.
- Reineking, B., and Schröder, B., 2006. Constrain to perform: regularization of habitat models. *EcolModel* 193:675–690. <https://doi.org/10.1016/j.ecolmodel.2005.10.003>.
- Ruppert, D., 2004. The elements of statistical learning: data mining, inference, and prediction. Springer Science & Business Media.
- Santos, M., Aguiar, M., and Oliveira, A., 2020. Vulnerability to mass movements' hazards. Contribution of sociology to increasing community resilience. In: *Advances in Natural Hazards and Hydrological Risks: Meeting the Challenge*. Springer, pp 105–108https://doi.org/10.1007/978-3-030-34397-2_21.
- Santosa, F., and Symes, W., 1986. Linear inversion of band-limited reflection seismograms. *SIAM J Sci Stat Comput* 7:1307–1330. <https://doi.org/10.1137/090708>.
- Schmid, U., Rösch, P., and Krause, M., 2009. Gaussian mixture discriminant analysis for the single-cell differentiation of bacteria using micro-Raman spectroscopy. *Chemom Intell Lab Syst* 96:159–171. <https://doi.org/10.1016/j.chemolab.2009.01.008>.
- Shahabi, H., Jarihani, B., and Piralilou, S.T., 2019. A semi-automated objectbased gully networks detection using different machine learning models: a case study of bowen catchment, Queensland, Australia. *Sensors (Switzerland)* 19:4893. <https://doi.org/10.3390/s19224893>
- Shataee, S., Weinaker, H., and Babanejad, M., 2011. Plot-level forest volume estimation using airborne laser scanner and TM data, comparison of boosting and random forest tree regression algorithms. *Procedia Environ Sci* 7:68–73. <https://doi.org/10.1016/j.proenv.2011.07.013>.
- Shi, Y., and Jin, F., 2009. Landslide stability analysis based on generalized information entropy. In: *Proceedings - 2009 International Conference on Environmental Science and Information Application Technology, ESIAT 2009*. IEEE, pp 83–85. <https://doi.org/10.1109/ESIAT.2009.258>.
- Shirzadi, A., Bui, D.T., and Pham, B., 2017. Shallow landslide susceptibility assessment using a novel hybrid intelligence approach. *Environ Earth Sci* 76:60 -Sidle RC, Ziegler AD, Negishi JN, et al .2006. Erosion processes in steep terrain-truths, myths, and uncertainties related to forest management in Southeast Asia. 224:199–225. doi: <https://doi.org/10.1016/j.foreco.2005.12.019>.
- Simon, N., Friedman, J., Hastie, T., and Tibshirani, R., 2013. A sparse-group lasso. *J Comput Graph Stat* 22:231–245 Stoffel M, Huggel C (2012) Effects of climate change on mass movements in mountain environments. *Prog Phys Geogr* 36:421–439.<https://doi.org/10.1177/0309133312441010>.
- Suresh, D., Yarrakula, K., and Venkateswarlu, B., 2019. Risk mapping analysis with geographic information systems for landslides using supply chain. In: *Emerging Applications in Supply Chains for Sustainable Business Development*. IGI Global, pp 131–141. <https://doi.org/10.4018/978-1-5225-5424-0.ch008>.
- Taalab, K., Cheng, T., and Zhang, Y., 2018. Mapping landslide susceptibility and types using random forest. *Big Earth Data* 2:159–178. <https://doi.org/10.1080/20964471.2018.1472392>.
- Taylor, R., Digby, P., and Kempton, R.A., 1987. Multivariate analysis of ecological communities. Springer Science & Business Media.
- Thanh Noi, P., and Kappas, M., 2018. Comparison of random forest, k-nearest neighbor, and support vector machine classifiers for land cover classification using Sentinel-2 imagery. *Sensors* 18:18. <https://doi.org/10.3390/s18010018>.
- Tiranti, D., and Cremonini, R., 2019. Editorial: landslide hazard in a changing environment. *Front Earth Sci* 7. <https://doi.org/10.3389/feart.2019.00003>.
- Vilar, L., Gómez, I., and Martínez-Vega, J., 2016. Multitemporal modelling of socio-economic wildfire drivers in Central Spain between the 1980s and the 2000s: comparing generalized linear models to machine learning algorithms. *PLoS One* 11:e0161344. <https://doi.org/10.1371/journal.pone.0161344>.
- Visa, S., Ramsay, B., Ralescu A.L., and Van Der Knaap, E., 2011. Confusion matrix-based feature selection. *MAICS* 710:120–127.

- Vorpahl, P., Elsenbeer, H., Märker, M., and Schröder, B., 2012. How can statistical models help to determine driving factors of landslides? *Ecol Model* 239:27–39. <https://doi.org/10.1016/j.ecolmodel.2011.12.007>.
- Yamada, Y., Kawamura, K., and Ikehara, K., 2012. Submarine mass movements and their consequences. In: *Submarine Mass Movements and Their Consequences - 5th International Symposium*. Springer, pp 1–12. . https://doi.org/10.1007/978-94-007-2162-3_1.
- Yousefi, S., Moradi, H., Boll, J., and Schönbrodt-Stitt, S., 2016. Effects of road construction on soil degradation and nutrient transport in Caspian Hyrcanian mixed forests. *Geoderma* 284:103–112. <https://doi.org/10.1016/j.geoderma.2016.09.002>.
- Youssef, A.M., Pourghasemi, H.R., El-Haddad, B.A., and Dhahry, B.K., 2016. Landslide susceptibility maps using different probabilistic and bivariate statistical models and comparison of their performance at Wadi Itwad Basin, Asir region, Saudi Arabia. *Bull Eng Geol Environ* 75:63–87. <https://doi.org/10.1007/s10064-015-0734-9>.
- Zabihi, M., Pourghasemi, H.R., Motevalli, A., and Zakeri, M.A., 2019. Gully erosion modeling using gis-based data mining techniques in northern Iran: a comparison between boosted regression tree and multivariate adaptive regression spline. In: *Advances in Natural and Technological Hazards Research*. Springer, pp 1–26. https://doi.org/10.1007/978-3-319-73383-8_1.