



مدل سازی ترکیبی فرسایش آبی و بادی با استفاده از روش‌های یادگیری ماشین

مهدی جلالی^۱، حمید غلامی^{۲*}، مرضیه رضایی^۳، ابراهیم امیدوار^۴

- ۱- دانشجوی دکتری علوم و مهندسی آبخیزداری، گروه مهندسی منابع طبیعی، دانشکده کشاورزی و منابع طبیعی، دانشگاه هرمزگان، بندر عباس، ایران
- ۲- دانشیار گروه مهندسی منابع طبیعی، دانشکده کشاورزی و منابع طبیعی، دانشگاه هرمزگان، بندر عباس، ایران
- ۳- استادیار گروه مهندسی منابع طبیعی، دانشکده کشاورزی و منابع طبیعی، دانشگاه هرمزگان، بندر عباس، ایران
- ۴- استادیار گروه منابع طبیعی، دانشکده کشاورزی و منابع طبیعی، دانشگاه کاشان، کاشان، ایران

چکیده مبسوط

مقدمه و هدف

فرسایش خاک به وسیله آب و باد یکی از بزرگ‌ترین تهدیدهای زیست‌محیطی در سراسر جهان است که پیامدهای منفی پرشماری مانند نابودی خاک، کاهش حاصلخیزی خاک، تخلیه مواد مغذی و ریزمغذی‌ها، نابودی ساختمان خاک، طوفان‌های گرد و غبار، آلودگی هوا، سیل‌های شدن آبگیرهای سدها، و ... روی بوم‌سازگان انسانی و طبیعی دارد. بنابراین داشتن اطلاعات دقیق از حساسیت زمین‌ها به فرسایش آبی و بادی و تهیه نقشه‌های مکانی از فرسایش به منظور کاهش پیامدهای منفی آن لازم است. بنابراین، هدف از این پژوهش استفاده از دو مدل یادگیری ماشین به منظور مدل‌سازی مکانی هم‌زمان فرسایش آبی و بادی در آبخیز بختگان استان فارس است.

مواد و روش‌ها

به منظور تهیه نقشه مکانی خطر فرسایش آبی و بادی خاک، به ترتیب ۲۰ و ۱۶ عامل تأثیرگذار بر فرسایش آبی و بادی انتخاب شد. در مرحله قبل از مدل‌سازی از روش انتخاب ویژگی و ابزاری چندمتغیره اسپیلاین تطبیقی و تست هم‌خطی به منظور شناسایی مهمترین عامل‌های مهارکننده خطر فرسایش آبی و بادی استفاده شد. در مرحله بعد

نوع مقاله: پژوهشی

*مسئول مکاتبات، پست الکترونیکی: hgholami@hormozgan.ac.ir

استاد: جلالی، م، غلامی، ح، رضایی، م، امیدوار، ا. ۱۴۰۲. مدل‌سازی ترکیبی فرسایش آبی و بادی با استفاده از روش‌های یادگیری ماشین. پژوهش‌های آبخیزداری، ۳۶ (۳): ۱۴۵-۱۲۸.

شناسه دیجیتال: 10.22092/WMRJ.2022.358127.1458

تاریخ دریافت: ۱۴۰۱/۰۱/۲۲، تاریخ بازنگری: ۱۴۰۱/۰۵/۲۵، تاریخ پذیرش: ۱۴۰۱/۰۶/۳۰، تاریخ انتشار: ۱۴۰۲/۰۷/۰۱. پژوهش‌های آبخیزداری، سال ۱۴۰۲، دوره ۳۶، شماره ۳، شماره‌ی پیاپی ۱۴۰، پاییز ۱۴۰۲، صفحه‌های ۱۲۸ تا ۱۴۵.

© نویسندگان

ناشر: مرکز تحقیقات و آموزش کشاورزی و منابع طبیعی استان فارس



از دو روش یادگیری ماشین جنگل تصادفی و ماشین بردار پشتیبان برای مدل سازی فرسایش استفاده شد. عملکرد مدل ها با استفاده از منحنی مشخصه عملکرد ارزیابی شد.

نتایج و بحث

براساس نتایج، متغیرهای کاربری زمین ها، شیب، سنگ شناسی، زبری سطح خاک، بخش درشت دانه ی خاک، درصد رس، پوشش گیاهی، جهت شیب، وزن مخصوص خاک و شاخص رطوبت پستی بلندی به عنوان مهمترین عامل های مهارکننده فرسایش آبی شناسایی شدند. همچنین مهمترین عامل های مهارکننده فرسایش بادی عبارت بودند از: بلندی، کاربری زمین ها، سرعت باد، شیب، پوشش گیاهی، نیتروژن خاک، ظرفیت تبادل کاتیونی، سنگ شناسی، جهت شیب، درصد رس و وزن مخصوص خاک. اندازه های سطح زیر منحنی مشخصه عملکرد برای نقشه ی فرسایش بادی تهیه شده به وسیله ی مدل جنگل تصادفی ۹۹٪ و برای نقشه ی فرسایش آبی تهیه شده به وسیله ی مدل ماشین بردار پشتیبان ۹۶٪ بود. در نهایت دو نقشه ی فرسایش بادی و آبی تلفیق شد و نتایج نشان داد که بیشترین درصد مساحت مربوط به حساسیت کم در فرسایش آبی و حساسیت خیلی زیاد در فرسایش بادی است که ۱۸٪ از منطقه را دربر می گرفت.

نتیجه گیری و پیشنهادها

منطقه ی مطالعه شده به دلیل وجود چندین دریاچه شامل بختگان، مهارلو و تشک و همچنین چندین سد، از نظر تأمین منابع آب شرب، کشاورزی، صنعت و محیط زیست اهمیت زیادی دارد. در صورت توجه نداشتن به مشکل فرسایش به ویژه فرسایش بادی در منطقه ی مطالعه شده، این منطقه می تواند به ویژه در بستر دریاچه های خشکیده منبع تولید گرد و غبار باشد. بنابراین یافته های این پژوهش می تواند در راستای کاهش اثرات منفی این پدیده در منطقه ی مطالعه شده به وسیله ی مدیران اجرایی استان بهره برداری شود.

واژگان کلیدی: آبخیز بختگان، جنگل تصادفی، مدل ماشین بردار پشتیبان، مساحت زیر منحنی، وایازی چندمتغیره اسپیلاین تطبیقی

مقدمه

نامطلوب آن را می توان به صورت شن زارها و زمین های بیابانی مشاهده کرد. فرسایش بادی در تمام نقاط دنیا رخ می دهد و سبب خسارت می شود. در مقیاس جهانی اهمیت و خطر فرسایش بادی کمتر از فرسایش آبی است. ولی گاهی ابعاد و بزرگی آن از فرسایش آبی بیشتر است. افزایش ۴۵٪ اندازه ی فرسایش خاک در کشور در طی سال های ۱۳۳۰ تا ۱۳۷۸ نشانگر بحرانی بودن فرسایش و لزوم مهار کردن آن در ایران است. اما مهار کردن یا مبارزه با فرسایش آبی مستلزم شناخت مناطق بحرانی و سهم مشارکت هر یک از انواع فرسایش های آبی و بادی در نابودی زمین ها و تولید رسوب است. از این رو، بررسی این موضوع نیازمند انجام پژوهش های علمی گسترده ای در کشور است. پیش بینی و مدل سازی خطر فرسایش های آبی و بادی در آبخیزها کمک فراوانی برای مبارزه و پیشگیری از این پدیده خواهد بود. روش های تهیه ی نقشه های حساسیت به فرسایش را می توان به دو دسته اصلی تقسیم کرد: الف- نقشه برداری مستقیم، ب- روش های نقشه برداری غیرمستقیم (لوکا و همکاران ۲۰۱۱).

در حال حاضر نابودی و از بین رفتن خاک بر اثر بارش و عمل آب های روان و باد یکی از جدی ترین مشکلات زیست محیطی است (تالوکدار و همکاران ۲۰۲۲). اثرات فرسایش خاک و ته نشین رسوب ها هم درون منطقه ای و هم برون منطقه ای است (رحمتی و همکاران ۲۰۱۳). در مناطق خشک و نیمه خشک با بارندگی های کم یا متغیر (مستعد خشکی) که با بادهای شدید دوره ای و تبخیر زیاد از سطح خاک می باشد، فرسایش خاک می تواند به عنوان یک مشکل بزرگ مطرح باشد (مینا و همکاران ۲۰۲۲). بیش از یک سوم زمین های کره ی زمین اقلیم خشک و نیمه خشک دارند و در دهه های اخیر پدیده ی بیابان زایی در این مناطق شدت یافته است (سنانایاک و همکاران ۲۰۲۲). بخش بزرگی از کشور ایران را نیز بیابان ها تشکیل می دهند که معمولاً خاک آن ها شور است و بدون پوشش گیاهی یا پوشش گیاهی خیلی کمی دارند (بردبار و همکاران ۲۰۲۲) همان طوری که در مناطق پرباران فرسایش آبی اهمیت دارد در مناطق خشک و نیمه خشک کشور نیز فرسایش بادی عامل اصلی فرسایش است و آثار

معیارهای MAE، RMSE و MBE استفاده کردند. در نهایت از نظریه‌ی بازی‌ها برای رتبه‌بندی عامل‌های مؤثر بر فرسایش بهره گرفتند.

آبخیز بختگان از مهم‌ترین آبخیزهای کشور و استان فارس است، که به دلیل دخالت‌های زیاد انسانی و خشک شدن دریاچه‌ی بختگان، تحت تأثیر فرسایش آبی و بادی است. اهمیت فرسایش در آبخیز بختگان به دلیل شرایط راهبردی منطقه، وسعت زمین‌های کشاورزی و وجود دریاچه‌های بختگان و طشک بارزتر می‌باشد (ده‌بزرگی و همکاران ۲۰۲۰). در منطقه‌ی مطالعه شده، به‌رغم گسترش زیاد فرسایش و حساسیت منطقه نسبت به حضور چنین مشکلی، و هم‌چنین تغییراتی که در بخش‌های مختلف این منطقه پدید آمده است، تاکنون پژوهش کاملی برای ارزیابی ترکیبی فرسایش آبی و بادی و عامل‌هایی که موجب توسعه‌ی فرسایش آبی و بادی می‌شود، انجام نشده است. از این رو در این پژوهش، پهنه‌بندی هم‌زمان خطر فرسایش آبی و بادی در آبخیز بختگان با استفاده از روش‌های یادگیری ماشین انجام شد. هم‌چنین در این پژوهش یک چارچوب عملی به منظور بررسی ظرفیت فرسایشی در منطقه‌ی مطالعه شده، با در نظر گرفتن تمام عامل‌های مؤثر بر فرسایش، ارائه شد.

مواد و روش‌ها

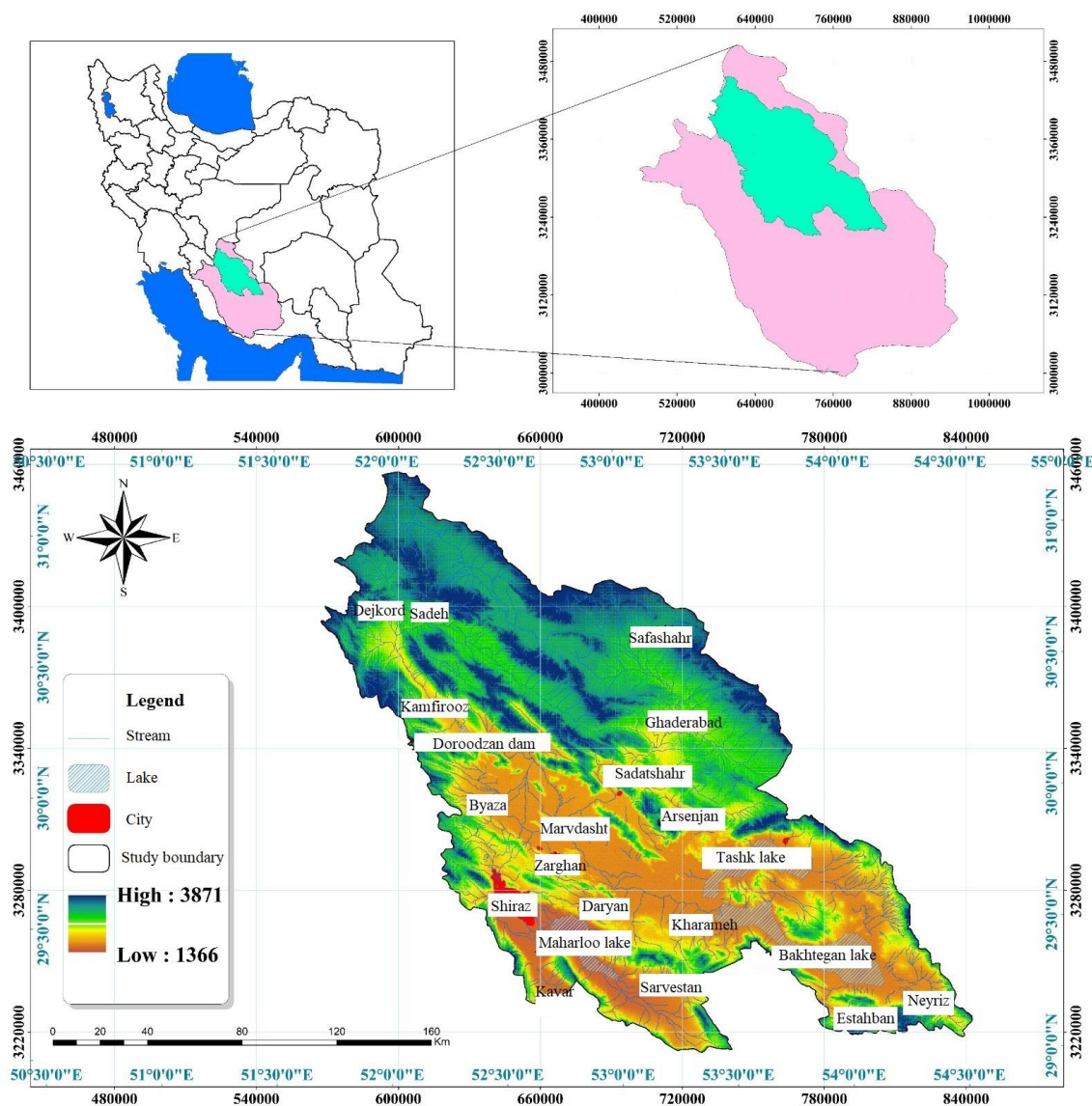
منطقه‌ی مطالعه شده

آبخیز بختگان با مساحت ۳۱۴۹۰ کیلومتر مربع در استان فارس میان ۴۲° ۵۱' تا ۳۱° ۵۴' طول شرقی و ۲۹° تا ۱۴' ۳۱° عرض شمالی است. آبخیز بختگان، به‌عنوان یکی از آبخیزهای مطالعاتی کشور، بخش وسیعی از مناطق شرقی، مرکزی، شمالی و شمال‌غربی استان فارس را تشکیل داده است. این آبخیز در راستای شمال غربی - جنوب شرقی و به موازات رشته کوه زاگرس گسترش یافته است. در شکل شماره ۱ موقعیت منطقه‌ی مطالعه شده در کشور و استان فارس نشان داده شده است.

روش‌های مستقیم به دانش فرد در پدیده‌ی بررسی شده و شناخت او از منطقه‌ی مطالعه شده بستگی دارد. اما روش‌های غیرمستقیم شامل روش‌های تعیینی و آماری هستند که براساس ارتباط بین عامل‌های محیطی و توزیع فرآیند بررسی شده می‌باشند (ون وستن و همکاران ۲۰۰۶). متداول‌ترین روش‌های استفاده شده برای تهیه‌ی نقشه‌های پهنه‌بندی را می‌توان به پنج دسته تقسیم‌بندی کرد (کارازا و همکاران ۱۹۹۵، هاتچینسون ۱۹۹۵، وان وستن و همکاران ۲۰۰۶، گازی و همکاران ۱۹۹۹، کانتگو و همکاران ۲۰۰۹، صفایی و همکاران ۲۰۱۰): ۱- نقشه‌برداری مستقیم زمین‌ریخت‌شناسی^۱، ۲- تحلیل نقشه‌های پایش^۲، ۳- روش‌های ابتکاری^۳، ۴- روش‌های آماری^۴، ۵- روش‌های مبتنی بر فرآیند مدل‌سازی مفهومی^۵.

در سال‌های اخیر استفاده از روش‌های یادگیری ماشین برای پهنه‌بندی فرسایش آبی و بادی و سایر مخاطره‌های محیطی متداول شده است (غلامی و همکاران، ۲۰۲۳؛ یروغنی و همکاران، ۲۰۲۱؛ غلامی و محمدی فر، ۲۰۲۲؛ رضایی و همکاران، ۲۰۲۳؛ محمدی فر و همکاران ۲۰۲۱؛ غلامی و همکاران ۲۰۲۱). پورنادر و همکاران (۲۰۱۸) در پژوهشی عملکرد روش حداکثر آنتروپی (ME) در پیش‌بینی حساسیت به فرسایش در آبخیز در ایلام را بررسی کردند. آنها برای این منظور، ده عامل مؤثر بر فرسایش را در فرآیند مدل‌سازی به‌کار گرفتند. از نمودار ROC برای ارزیابی عملکرد مدل ME استفاده کردند. نتایج آنها نشان داد که روش حداکثر آنتروپی ابزاری مناسب برای تولید نقشه‌های حساسیت فرسایش است. غلامی و همکاران (۲۰۲۰) در استان خوزستان پهنه‌بندی مکانی منشأ ریزگردها را با استفاده از روش‌های مختلف یادگیری ماشین انجام دادند. نتایج این پژوهش نشان‌دهنده‌ی قابلیت زیاد روش استفاده شده بود. در پژوهشی محمدی فر و همکاران (۲۰۲۱) در آبخیز کهورستان استان هرمزگان پهنه‌بندی خطر فرسایش آبی را با استفاده از روش‌های مختلف یادگیری ماشین انجام دادند. معیارهای ورودی آن‌ها ۱۸ عامل مؤثر بر فرسایش آبی بودند. برای انتخاب بهترین مدل از

- 1 - Geomorphological mapping
- 2 - Heuristic (index-based)
- 3 - Analysis of inventories
- 4 - Statistical modeling
- 5 - Process based



شکل ۱- موقعیت منطقه مطالعه شده.
Figure 1- Location of study area.

در سال است (نصرنیا ۲۰۱۷). مهم ترین رود این آبخیز، رود کر می باشد. رود کر دائمی است و از شمال غربی استان فارس و بلندی های سلسله جبال زاگرس منشعب می شود و تا جنوب شرقی ادامه دارد. طول آن ۲۸۰ کیلومتر است و به دریاچه بختگان می ریزد (وفاخواه ۲۰۰۵). پوشش گیاهی در آبخیز بختگان در ۴ رویشگاه مهم شامل کوهستانی، تپه ماهوری، دشتی- دامنه ای، دشتی- شورزار- تالابی مشاهده می شود.

تعیین عامل های مؤثر بر فرسایش آبی و بادی

عامل های اقلیمی، خاکشناسی، پستی بلندی، پوشش گیاهی و ... از مجموعه عامل های مؤثر بر فرسایش آبی و بادی هستند (فرارسی ۱۹۹۰). برای مدل سازی مکانی فرسایش با مدل های یادگیری ماشین اولین گام تعیین و انتخاب عامل های مؤثر بر فرسایش است (محمدی فر و

شیب نسبی آبخیز از شمال غرب به جنوب شرق است. رودها و آبراه های آن به دو چاله ای اصلی منتهی می شوند. چاله ای اول، دریاچه های کم عمق بختگان و طشک در شرق آبخیز هستند که به وسیله تنگه ای با هم ارتباط دارند. مهم ترین منبع این دو دریاچه، رود کر است. چاله ای دوم، دریاچه میهارلو است که در جنوب آبخیز و در نزدیکی شیراز است و محل تخلیه ای آب های سطحی و زیرزمینی زیر آبخیز خود است. میانگین بلندی این آبخیز ۲۰۳۱ متر و میانگین بارش سالانه در این آبخیز حدود ۳۶۵ میلی متر می باشد. در منطقه ای کوهستانی بیشینه ای مطلق دما $39/5^{\circ}\text{C}$ و کمینه ای مطلق آن 22°C - و متوسط سالانه ای دما 14°C است. در محدوده ای پایین دست، بیشینه ای مطلق دما $38/6^{\circ}\text{C}$ و کمینه ای آن $4/7^{\circ}\text{C}$ - با متوسط سالانه ای دما 17°C است. اندازه ای تبخیر در سراسر آبخیز ۱۳۰۰ تا ۳۵۰۰ میلی متر

همکاران (۲۰۲۱). بنابراین در این پژوهش ابتدا با بررسی منابع، عامل‌های تأثیرگذار بر فرسایش تعیین شد. نتایج این بررسی‌ها همراه با منبع استفاده شده در جدول ۱ آورده شده است.

جدول ۱- عامل‌های تأثیرگذار در فرسایش آبی و بادی.

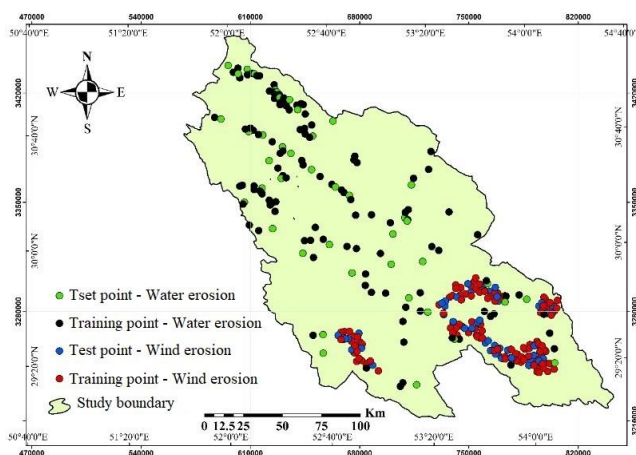
Table 1- Factors influencing on soil erosion by water and wind

Water erosion	Wind erosion	Data source and software
DEM, Elevation, Slope, Profile	Dem, Elevation, Slope, Aspect	https://earthexplorer.usgs.gov/ ArcGIS
Stream power index, Topographic wetness index, Roughness coefficient	Roughness coefficient	SAGA GIS
Clay, Silt, Sand, Bulk density, Coarse fragment, Organic carbon	Clay, Silt, Sand, Bulk density, Coarse fragment, Organic carbon, Cation exchange capacity, Nitrogen	https://soilgrids.org/
Soil moisture NDVI	Wind speed NDVI	https://cds.climate.copernicus.eu/ and https://earthexplorer.usgs.gov/
Lithology, Land use, Distance from river	Lithology, land use	Iran Watershed Management and Natural Resource Organization
Precipitation		https:// worldclim.org/

(مطالعه‌های انجام‌شده در دوره‌های مختلف به‌وسیله‌ی اداره‌ی منابع طبیعی استان فارس) تعیین شد. تعداد نقاط فرسایش بادی ۱۶۲ نقطه و تعداد نقاط فرسایش آبی ۱۵۵ نقطه تعیین شد. سپس به‌شکل تصادفی ۳۰٪ نقاط برای ارزیابی و ۷۰٪ نقاط برای آموزش مدل‌ها، انتخاب شد (رحمتی و همکاران ۲۰۱۶). در شکل ۲ نقاط فرسایشی نشان داده شده است.

تعیین نقاط فرسایشی

دومین ورودی به مدل‌های یادگیری ماشین در مدل‌سازی مکانی، نقاط فرسایشی است. در این پژوهش نقاط فرسایش بادی، با استفاده از نقشه‌ی کانون‌های ریزگرد مربوط به سازمان جنگل‌ها، مراتع و آبخیزداری و انجام بازدیدهای میدانی تعیین شد. نقاط فرسایش آبی با انجام بازدیدهای میدانی و مطالعه‌های پایه‌ی آبخیز بختگان



شکل ۲- نقشه‌ی نقاط فرسایش بادی و آبی.

Figure 2- Map of water and wind erosion points.

داده‌کاوی و مدل‌سازی هوشمند است (کورسا و رودنیکی ۲۰۱۰). چون، ویژگی‌های نامرتب یا تا حدودی مرتبط می‌توانند تأثیر منفی بر عملکرد سامانه داشته باشند. پیاده‌سازی روش‌های انتخاب ویژگی اولین و مهم‌ترین مرحله در طراحی سامانه‌های هوشمند یادگیری هستند

انتخاب ویژگی

در مبحث یادگیری ماشین و همچنین شناسایی آماری الگو، انتخاب ویژگی، یکی از مسائل مطرح است. ویژگی‌ها تأثیر زیادی بر عملکرد، دقت و کارایی سامانه دارند. عملیات انتخاب ویژگی مهم‌ترین بخش عملیات

هم‌خطی است (چن و همکاران ۲۰۱۷). در این پژوهش اندازه‌های عددی هر یک از سنج‌ها در مناطق فرسایشی استخراج شد و سپس در محیط SPSS تست هم‌خطی محاسبه شد.

مدل‌سازی با روش‌های یادگیری ماشین

در این پژوهش مدل‌سازی با دو روش جنگل تصادفی و ماشین بردار پشتیبان انجام شد.

روش جنگل تصادفی^{۱۵} (RF)

این دستورالعمل را نخستین بار لئو برایمن و آدل کاتلر ارائه و توسعه دادند. دستورالعمل جنگل تصادفی مبتنی بر دسته‌های از درخت‌های تصمیم است و هم‌اکنون یکی از بهترین دستورالعمل‌های یادگیری است. اساس مدل پیش‌بینی‌کننده‌ی RF بر پایه‌ی میانگین‌گیری از نتایج به‌دست آمده از تمامی درخت‌های تصمیم مربوط استوار است و طبقه‌بندی بسیاری از مجموعه داده‌ها با آن صحت زیادی دارد. درختان تصادفی با در نظر گرفتن بردار ورودی، آن را با هر درخت در جنگل طبقه‌بندی کرده و خروجی آن، برچسب‌های طبقه‌ای هستند که اکثریت امتیاز را کسب کرده‌اند. در این مدل از دو عامل میانگین کاهشی دقت و میانگین کاهشی جینی برای تعیین اولویت تأثیر هر یک از عامل‌های مؤثر استفاده می‌شود (نیکودموس ۲۰۱۱).

روش ماشین بردار پشتیبان^{۱۶} (SVM)

دستورالعمل SVM یا روش ماشین پشتیبان بردار به‌وسیله‌ی وپینک بر پایه‌ی نظریه‌ی یادگیری آماری است که از حداقل ریسک ساختاری تبعیت می‌کند و برای ارزیابی و آزمون مجموعه‌ای از داده‌ها استفاده می‌شود (یاو و همکاران ۲۰۰۸). براساس نظریه‌ی یادگیری آماری، می‌توان سرعت خطای ماشین یادگیری را برای داده‌های طبقه‌بندی نشده، به‌عنوان رشد سرعت خطای تعمیم‌یافته، در نظر گرفت. در سال‌های اخیر به‌دلیل عملکرد خوب طبقه‌بندی و قابلیت تعمیم مناسب این دستورالعمل، توجه زیادی به آن شده است.

(بریمن ۲۰۰۱). همچنین، وقتی که ابعاد فضای ویژگی داده‌ها بسیار زیاد است و معضل بعد وجود دارد، استفاده از مجموعه ویژگی‌های مناسب، هزینه‌های محاسبه‌ای لازم برای آموزش بهینه‌ی سامانه را بسیار کاهش می‌دهد (چتام‌ولسی ۲۰۱۱). در این پژوهش از روش وایازی چند متغیره سازوار اسپلین^۶ MARS برای انتخاب ویژگی استفاده شد. این روش ترکیبی از وایازی اسپلین^۷، برازش مدل با روش گام به گام^۸ و افراز بازگشتی^۹ است (اندرسون ۱۹۸۴). این روش، هم در مسائل وایازی و هم در مسائل رده‌بندی به‌کار گرفته می‌شود؛ یعنی اگر متغیر پاسخ از نوع کمی باشد، مسئله‌ی وایازی و اگر از نوع کیفی باشد، مسئله‌ی رده‌بندی است. این روش هنگامی که تعداد متغیرهای توضیحی زیاد است، عملکرد خوبی دارد و در شناسایی اثرات متقابل به‌طور مؤثر عمل می‌کند (چو و همکاران ۲۰۰۴). روش MARS، به جای استفاده مستقیم از متغیرها، توابعی از آن‌ها را به‌منظور مدل‌سازی به‌کار می‌برد؛ این توابع را توابع پایه‌ای^{۱۰} می‌گویند. در این پژوهش این روش با پکیج Earth در نرم‌افزار R انجام شد. به این منظور در نقاط فرسایشی اندازه‌ی متغیرهای ورودی استخراج شد. در این روش برای انتخاب متغیرهای اصلی از سه معیار nsubsets (تعداد زیرمجموعه مدل)،^{۱۱} GCV و RSS^{۱۲} استفاده شد. بنابراین عامل‌هایی که در تعداد زیاد زیرمجموعه‌ها باشد به‌عنوان متغیر مهم انتخاب می‌شوند. معیار RSS برآورد کاهش مجموع مربعات باقی‌مانده برای هر زیرمجموعه در مقایسه با زیرمجموعه‌ی قبلی است. معیار GCV نیز مشابه‌ی مفهوم RSS است با این تفاوت که از اعتبارسنجی متقاطع تعمیم‌یافته استفاده می‌شود (محمدی‌فر و همکاران ۲۰۲۱).

تست هم‌خطی

در این پژوهش برای تعیین هم‌خطی از سنج‌های ضریب تحمل^{۱۳} و عامل تورم واریانس^{۱۴} استفاده شده است. اگر اندازه‌ی ضریب تحمل کمتر از ۰/۱ باشد و اندازه‌ی عامل تورم واریانس بزرگ‌تر از ۱۰ باشد، رابطه‌ی بین سنج‌ها

- 6 - Multivariate Adaptive Regression Spline
- 7 - Spline regression
- 8 - Stepwise model fitting
- 9 - Recursive partitioning
- 10 - Basis functions
- 11 - Generalized Cross Validation
- 12 - Residual Sum of Squares
- 13 - Tolerance
- 14- Variance inflation factor
- 15 - Random forest
- 16 - Support vector machines

خوب، $0/7-0/8$ دقت خوب، $0/7-0/6$ دقت متوسط و $0/6-0/5$ دقت ضعیف (پسیل نکار ۲۰۰۵).
با توجه به این که برای ارزیابی مدل نمی توان از موقعیت نقاط فرسایشی استفاده کرد، از میان موقعیت نقاط فرسایشی، 70% برای اجرای مدل و 30% برای ارزیابی مدل استفاده شد (کنستانتین و همکاران ۲۰۱۱).

نتایج

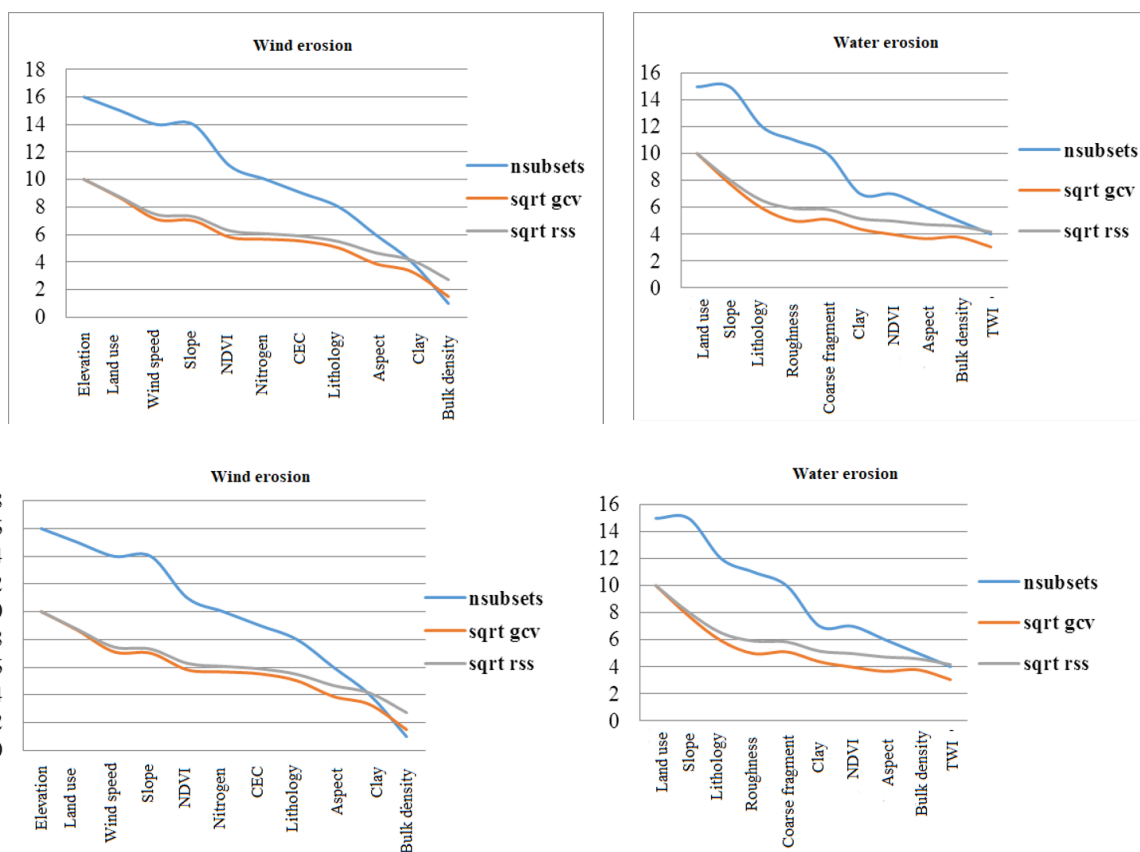
انتخاب ویژگی

متغیرهای انتخاب شده به وسیله ی MARS در شکل ۳ نشان داده شده است. در نمودار برای نمایش بهتر از معیار GCV و RSS جذر گرفته شده است. براساس نتایج، متغیرهای کاربری زمین ها، شیب، سنگ شناسی، زبری سطح خاک، بخش درشت دانه خاک، درصد رس، پوشش گیاهی، جهت شیب، وزن مخصوص خاک و شاخص رطوبت پستی بلندی به عنوان مهمترین عامل های مهارکننده ی فرسایش آبی شناسایی شدند. همچنین مهمترین عامل های مهارکننده ی فرسایش بادی عبارت بودند از: بلندی، کاربری زمین ها، سرعت باد، شیب، پوشش گیاهی، نیتروژن خاک، ظرفیت تبادل کاتیونی، سنگ شناسی، جهت شیب، درصد رس و وزن مخصوص خاک.

بعد از آماده سازی لایه ها و اطلاعات اولیه از 70% نقاط برای آموزش مدل ها استفاده شد و نقشه های حساسیت فرسایش آبی و بادی با روش جنگل تصادفی و ماشین بردار پشتیبان تهیه شد. شایان ذکر است که تمام مراحل مدل سازی در نرم افزار R و با استفاده از بسته ی Caret انجام شد. پس از مدل سازی نقشه های رستری تولید شده وارد نرم افزار ArcMap شده و با استفاده از دستور Jenks (Natural Breaks) یا شکست های طبیعی به ۴ طبقه ی خطر کم ($0-0/25$)، حساسیت متوسط ($0/5-0/25$)، حساسیت زیاد ($0/75-0/5$) و حساسیت خیلی زیاد ($0/75-1$) تقسیم بندی شد.

اعتبارسنجی شبیه سازی

به منظور ارزیابی مدل از مساحت زیرمنحنی (AUC) استفاده شد. در منحنی ویژگی عملگر نسبی (ROC) قدرت پیش بینی مدل با استفاده از شاخص سطح سلول هسته $SCAI^{17}$ دقت تفکیک طبقه ها بررسی و تأیید شد. مساحت زیر منحنی، درصد نقاط فرسایشی که به درستی به وسیله ی مدل پیش بینی شده است در مقابل درصد نقاطی که به درستی پیش بینی نشده است رسم شد. رابطه ی کمی- کیفی بین AUC و دقت پیش بینی که دامنه ی $0/5$ تا 1 را شامل می شود عبارت است از: $0/9-1$ دقت عالی، $0/9-0/8$ دقت خیلی

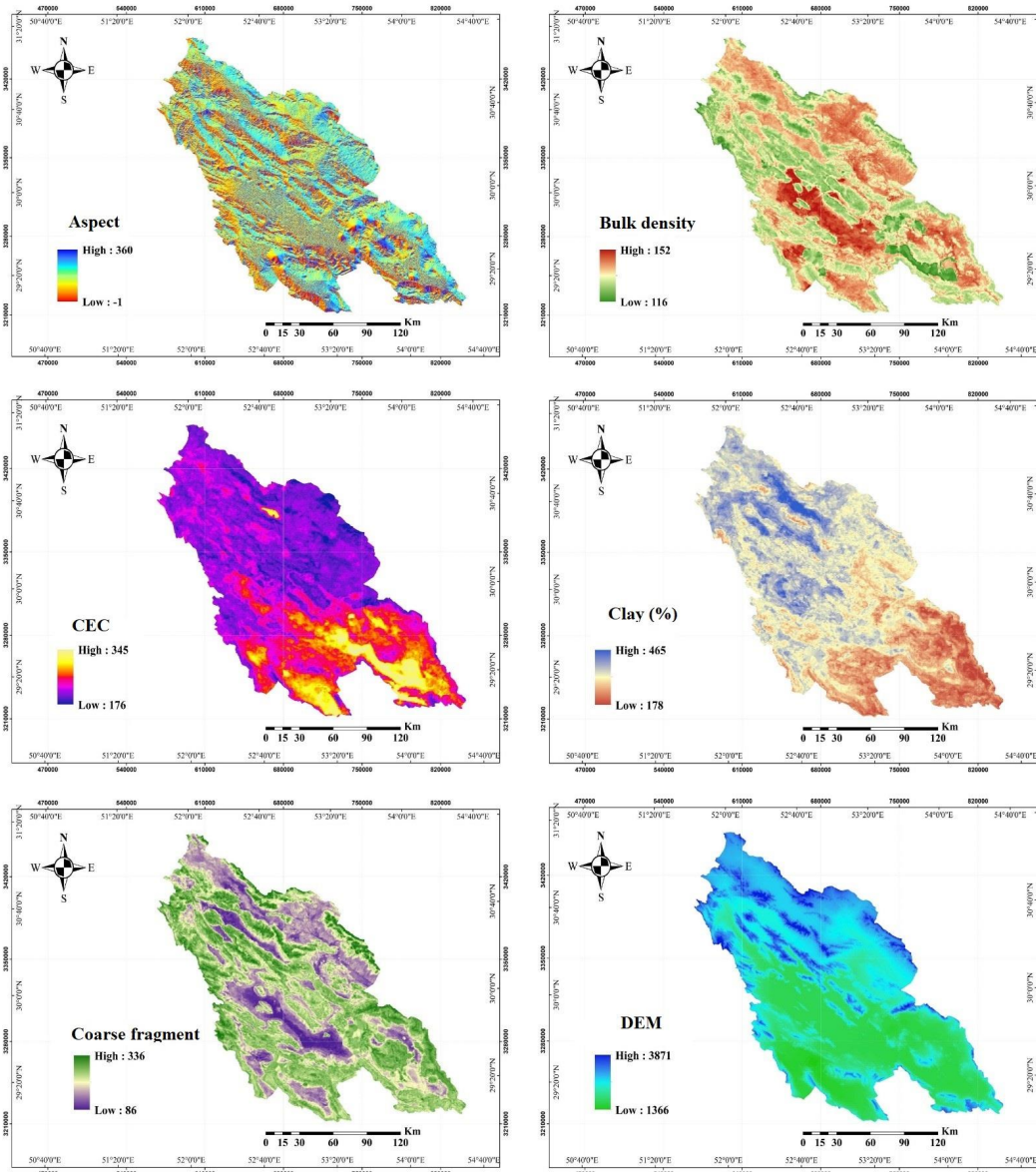


شکل ۳- متغیرهای انتخاب شده در فرسایش آبی و بادی.

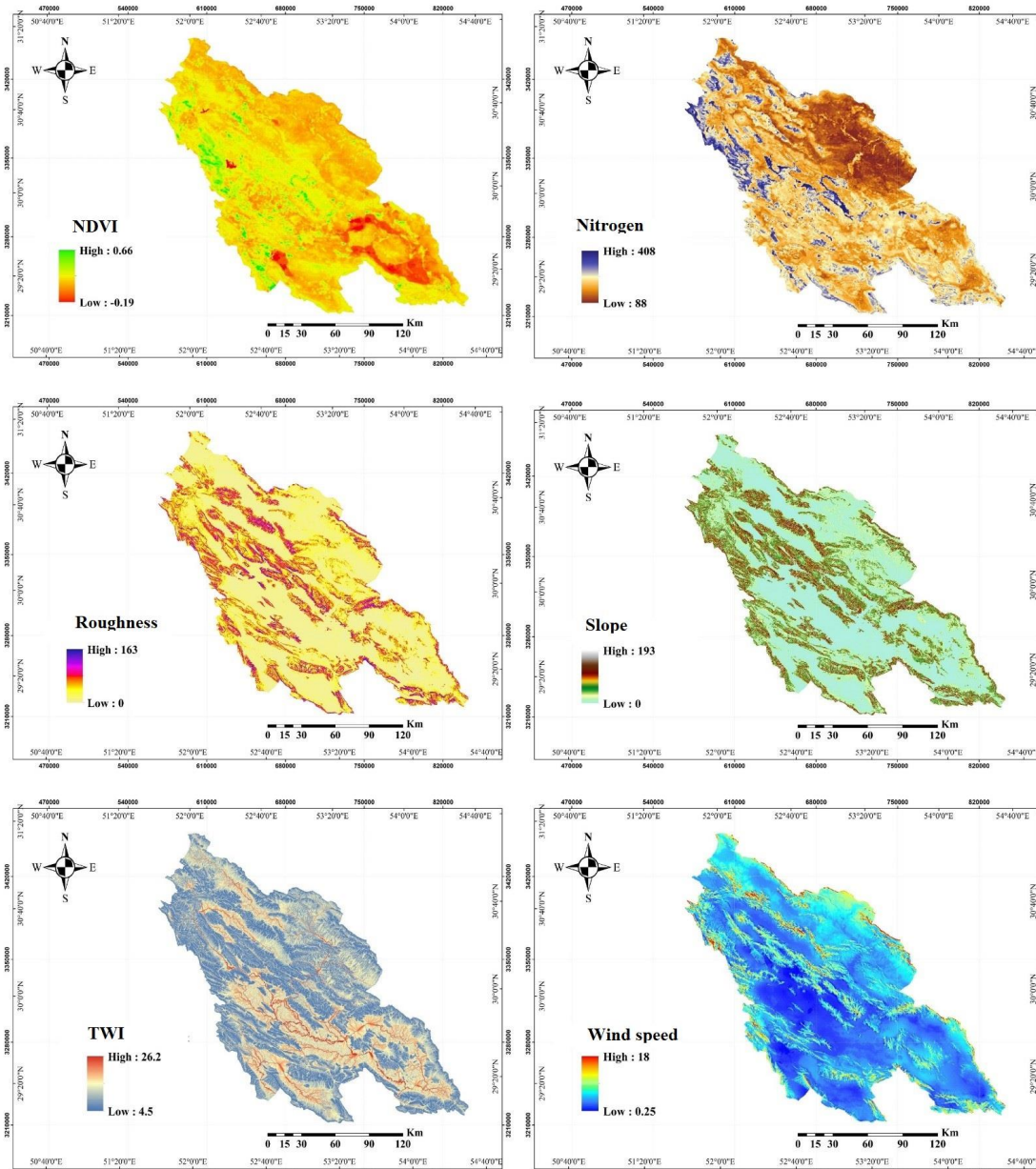
Figure 3- Feature selected for water and wind erosion.

(شکل ۳). نقشه‌های مربوط به این متغیرها در شکل ۴ نشان داده شده است.

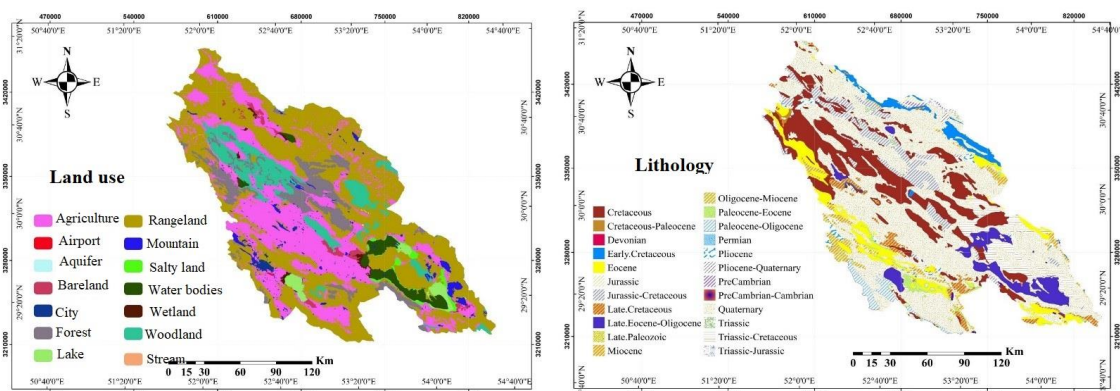
در این پژوهش برای فرسایش آبی ۱۰ متغیر و برای فرسایش بادی ۱۱ متغیر برای مدل‌سازی انتخاب شد



شکل ۴ (الف) - متغیرهای انتخاب شده در مدل‌سازی برای فرسایش آبی و بادی.
Figure 4 (a)- Features selected for modelling of water and wind erosion.



شکل ۵ (ب) - متغیرهای انتخاب شده در مدل‌سازی برای فرسایش آبی و بادی.
Figure 4 (b) - Features selected for modelling of water and wind erosion.



شکل ۶ (ج) - متغیرهای انتخاب شده در مدل‌سازی برای فرسایش آبی و بادی.
Figure 4 (c) - Features selected for modelling of water and wind erosion.

تست هم‌خطی براساس نتایج این پژوهش، میان هیچ‌کدام از متغیرها ۲ و ۳ نشان داده شده است. رابطه‌ی هم‌خطی نبود. نتایج تست هم‌خطی در جدول‌های

جدول ۲- نتایج تست هم‌خطی در فرسایش آبی.

Table 2- The results of multicollinearity test for water erosion.

Feature	Multicollinearity test	
	TC	VIF
TWI	0.371	2.689
Roughness	0.125	8.542
Slope	0.146	8.082
Clay	0.473	2.115
Coarse fragment	0.115	8.684
Aspect	0.983	1.017
Aspect	0.273	3.663
Land use	0.304	3.287
Lithology	0.566	1.767
NDVI	0.67	1.493

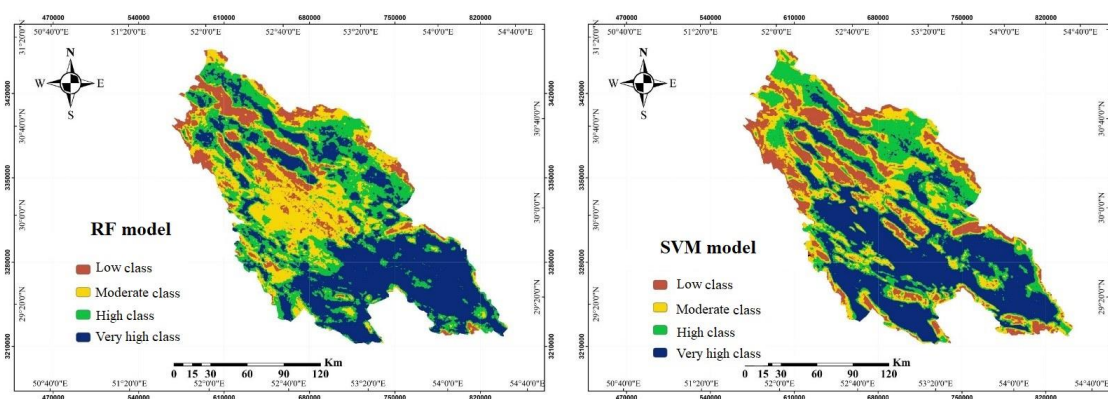
جدول ۳- نتایج تست هم‌خطی در فرسایش بادی.

Table 3- The results of multicollinearity test for wind erosion.

Feature	Multicollinearity	
	TC	VIF
Aspect	0.937	1.068
Bulk density	0.546	1.83
CEC	0.416	2.405
DEM	0.182	5.504
Land use	0.626	1.597
Lithology	0.678	1.475
NDVI	0.459	2.177
Slope	0.528	1.895
Clay	0.721	1.388
Nitrogen	0.649	1.541
Wind wpeed	0.226	4.428

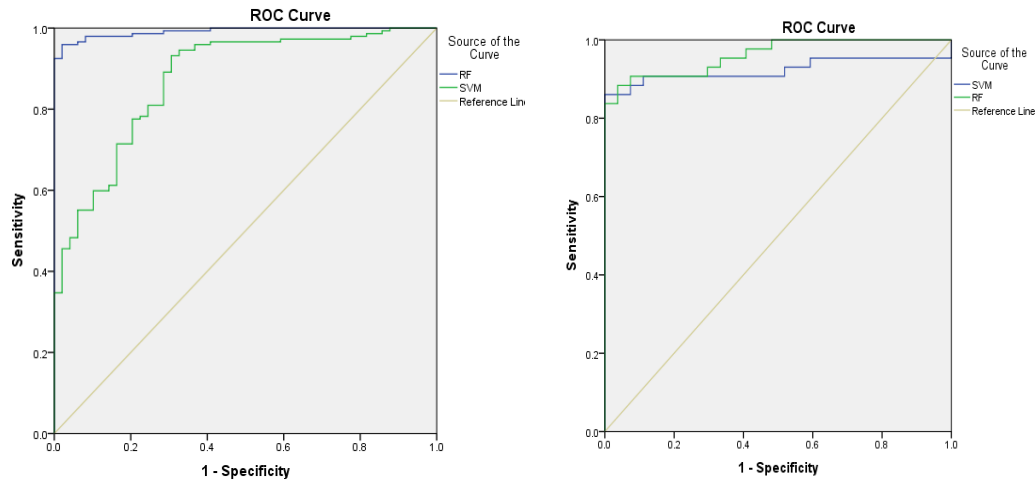
مدل‌های یادگیری در شکل‌های ۵ و ۶ نشان داده شده است.

نقشه‌های فرسایش آبی و بادی نقشه‌های فرسایش آبی و بادی تهیه شده به‌وسیله‌ی



شکل ۷- نقشه‌های حساسیت زمین‌ها به فرسایش بادی با مدل جنگل تصادفی و ماشین بردار پشتیبان.

Figure 5- Maps of land susceptibility to wind erosion by RF and SVM models.



شکل ۹- منحنی ROC در فرسایش بادی و آبی.
Figure 7- ROC for wind and water erosion.

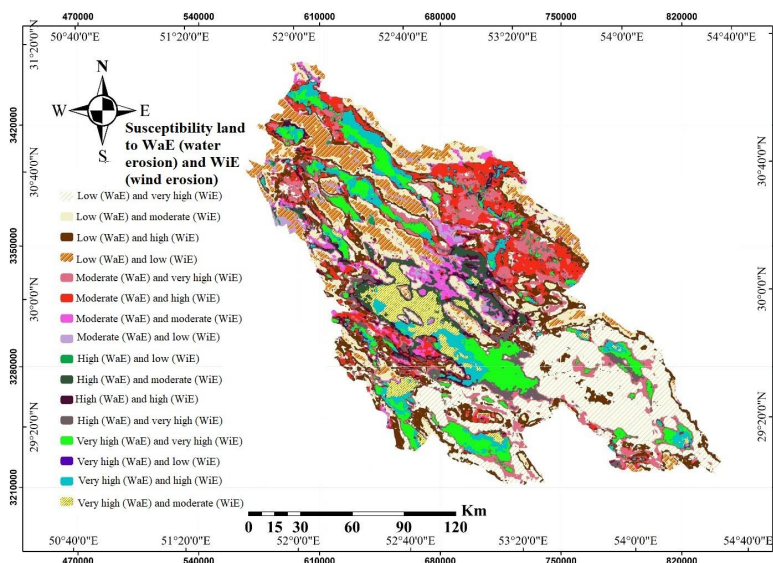
جدول ۶- اندازه های AUC مدل های استفاده شده برای به نقشه در آوردن فرسایش خاک.

Table 6: AUC values for models used to map soil erosion.

Model	AUC	SD	Sig level	Confidence level	
				Low level	Upper level
Wind erosion	RF model	0.992	0.004	0.983	1
	SVM model	0.874	0.028	0.818	0.929
Water erosion	RF model	0.923	0.037	0.852	0.995
	SVM model	0.961	0.02	0.922	1

تلفیق فرسایش بادی و آبی
بعد از مدل سازی و ارزیابی مدل های استفاده شده، در نهایت برای تهیه یک نقشه نهایی از ترکیب فرسایش آبی و بادی، نقشه های مربوط به جنگل تصادفی در فرسایش بادی و نقشه های مربوط به ماشین بردار پشتیبان در فرسایش آبی، با توجه به عملکرد بهتر با یکدیگر ترکیب شدند. نتایج نقشه نهایی در شکل ۸ نشان داده شده است.

بر اساس نتایج این پژوهش هر دو مدل استفاده شده عملکرد خوبی در مدل سازی مکانی فرسایش آبی و بادی داشتند و انتخاب مدل بهتر مشکل بود (جدول ۶). با این حال مدل جنگل تصادفی در فرسایش بادی و مدل ماشین بردار پشتیبان در فرسایش آبی با بیشترین AUC و کمترین اندازه ی انحراف معیار بهترین عملکرد را داشتند.



شکل ۱۰- نقشه ی ترکیبی فرسایش آبی و بادی.
Figure 8- Integrated map of soil erosion by water and wind.

مساحت طبقه های مربوط به نقشه ی ترکیبی در جدول ۷ نشان داده شده است. بیشترین درصد مساحت مربوط به حساسیت کم در فرسایش آبی و حساسیت خیلی زیاد

در فرسایش بادی بود که ۱۸٪ از مساحت کل منطقه بود.

جدول ۷- مساحت مربوط به طبقه های نقشه ترکیبی.

Water erosion	Wind erosion	Area (ha)	Area (%)
low susceptibility	Low susceptibility	239619.3	7.62
low susceptibility	Moderate susceptibility	387525.3	12.32
low susceptibility	High susceptibility	343184.5	10.9
low susceptibility	Very high susceptibility	566112.3	18
Moderate susceptibility	Low susceptibility	45254	1.44
Moderate susceptibility	Moderate susceptibility	131702	4.2
Moderate susceptibility	High susceptibility	272468.1	8.67
Moderate susceptibility	Very high susceptibility	254240.1	8.1
High susceptibility	Low susceptibility	7107.5	0.23
High susceptibility	Moderate susceptibility	125403	4
High susceptibility	High susceptibility	105578.6	3.4
High susceptibility	Very high susceptibility	123483.6	3.9
Very high susceptibility	Low susceptibility	57	0.002
Very high susceptibility	Moderate susceptibility	121190.2	3.85
Very high susceptibility	High susceptibility	178956.2	5.7
Very high susceptibility	Very high susceptibility	242632.8	7.7

جغرافیایی ابزار قوی به منظور ارزیابی حساسیت زمین ها نسبت به فرسایش می باشند. از نظر روش شناختی، علم داده (DS) یک زمینه ی پژوهشی در حال گسترش است. این ظرفیت برای درک بهتر فرآیندهای پیچیده محیطی بسیار پراهمیت است.

بر اساس نتایج این پژوهش مدل های یادگیری ماشین توانایی زیادی در مدل سازی فرسایش آبی و بادی داشتند که با نتایج بردبار و همکاران (۲۰۲۲)، سنانایاک و همکاران (۲۰۲۲) و محمدی فر و همکاران (۲۰۲۱) هم خوانی داشت. همچنین مدل جنگل تصادفی کارایی بسیار زیادی در مدل سازی مکانی داشت که با نتایج محمدی فر و همکاران (۲۰۲۱) هم راستا است. افزون بر این نتایج این پژوهش نشان داد که تلفیق خطر فرسایش بادی و آبی به وسیله ی مدل های یادگیری ماشین کمک زیادی به شناخت ظرفیت فرسایشی منطقه است که با نتایج بردبار و همکاران (۲۰۲۲) هم راستا است. نتایج این پژوهش می تواند به عنوان اطلاعات پایه به وسیله ی برنامه ریزان و مسئولان محلی به منظور ارزیابی، برنامه ریزی، مدیریت پایدار و حفاظت از منابع خاکی آبخیز بختگان در آینده استفاده شود. با توجه به هزینه بر بودن و زمان بر بودن آزمایش های خاک و نیز بازدیدهای میدانی، تهیه ی نقشه ی ظرفیت فرسایشی با استفاده از روش به کار گرفته شده در این پژوهش، در مناطق دیگر پیشنهاد می شود.

بحث و نتیجه گیری

در این پژوهش در آبخیز بختگان استان فارس ارزیابی کارایی دو روش جنگل تصادفی و ماشین بردار پشتیبان در تهیه ی نقشه ی حساسیت زمین ها نسبت به فرسایش آبی و بادی انجام شد. بر اساس نتایج ۷/۶، ۴/۲، ۳/۳ و ۷/۷٪ از کل آبخیز مطالعه شده، به ترتیب در طبقه های حساسیت کم، متوسط، زیاد و خیلی زیاد به خطر همزمان فرسایش آبی و بادی قرار گرفتند و بقیه ی آبخیز بختگان در سایر طبقه های خطر بودند. آبخیز بختگان به عنوان مهم ترین آبخیز استان فارس، در اثر تغییرات کاربری، تغییر اقلیم و فشار فزاینده انسان در معرض نابودی است و از نظر فرسایش بادی و فرسایش آبی شرایط حساسی دارد. در این پژوهش بر اساس نتایج انتخاب ویژگی MARS، سنگ شناسی به عنوان یکی از مهمترین عامل های مهار کننده ی فرسایش انتخاب شد و وجود سازندهای گروه فارس مانند گچساران، میشان و آغاچاری در این آبخیز همراه با دلیل های مزبور از علت های تشدید فرسایش آبی و بادی در این آبخیز بودند. با توجه به وسعت زیاد آبخیز و نبودن امکان اندازه گیری فرسایش به صورت مستقیم در سراسر آبخیز، استفاده از مدل سازی و به خصوص مدل های هوش مصنوعی به دلیل دقت، سادگی و نیاز کم به داده در اولویت می باشد. نتایج این پژوهش نشان داد که هر دو روش استفاده شده به همراه روش های سنجش از دور و سامانه ی اطلاعات

فهرست منابع

- Anderson TW, Anderson TW, Anderson TW, Anderson TW. 1958. An introduction to multivariate statistical analysis. New York: Wiley. pp. 1468-1482.
- Bordbar M, Aghamohammadi H, Pourghasemi HR, Azizi Z. 2022. Multi-hazard spatial modeling via ensembles of machine learning and meta-heuristic techniques. Scientific Reports. 12(1): 1451.
- Breiman L. 2001. Random forests. Machine learning. 45: 5-32.
- Boroughani M, Pourhashemi S, Gholami H, Kaskaoutis DG. 2021. Predicting of dust storm source by combining remote sensing, statistic-based predictive models and game theory in the Sistan Watershed, southwestern Asia. Journal of Arid Land. 13(11): 1103-1121.
- Carrara A, Cardinali M, Guzzetti F, Reichenbach P. 1995. GIS technology in mapping landslide hazard. In geographical information systems in assessing natural hazards, Dordrecht: Springer Netherlands. pp. 135-175.
- Mirhashemi SH, Haghghat Jou P, Mirzaei F, Panahi M. 2020. The study of environmental and human factors affecting aquifer depth changes using tree algorithm. International Journal of Environmental Science and Technology. 17: 1825-1834.
- Chen W, Pourghasemi HR, Naghibi SA. 2018. Prioritization of landslide conditioning factors and its spatial modeling in Shangnan County, China using GIS-based data mining algorithms. Bulletin of Engineering Geology and the Environment. 77: 611-629.
- Chou SM, Lee TS, Shao YE, Chen IF. 2004. Mining the breast cancer pattern using artificial neural networks and multivariate adaptive regression splines. Expert Systems With Applications. 27(1): 133-142.
- Constantin M, Bednarik M, Jurchescu MC, Vlaicu M. 2011. Landslide susceptibility assessment using the bivariate statistical analysis and the index of entropy in the Sibiciu Basin (Romania). Environmental Earth Sciences. 63: 397-406.
- Dehbozorgi M, Jafari M, Malekian A, Zehtabian G, Fallah Shamsi SR. 2020. Assessment the instability and land degradation using fuzzy methods, hierarchical analysis and weighted linear combination (Case study: Bakhtegan Watershed, Fars). Journal of Range and Watershed Management. 73(2): 321-335.
- Ferraresi M. 1989. The regionalization of fluvial sediment yield in Emilia-Romagna (northern Italy). IAHS-AISH publication. 191: 253-260.
- Gholami H, Mohammadifar A, Collins AL. (2020). Spatial mapping of the provenance of storm dust: Application of data mining and ensemble modelling. Atmospheric Research. 233: 104716.
- Gholami H, Mohammadifar A, Golzari S, Kaskaoutis DG, Collins AL. 2021. Using the Boruta algorithm and deep learning models for mapping land susceptibility to atmospheric dust emissions in Iran. Aeolian Research. 50: 100682.
- Gholami H, Mohammadifar A, Fitzsimmons KE, Li Y, Kaskaoutis DG. 2023. Modeling land susceptibility to wind erosion hazards using LASSO regression and graph convolutional networks. Frontiers in Environmental Science. 11: 1187658.
- Gholami H, Mohammadifar A. 2022. Novel deep learning hybrid models (CNN-GRU and DDL-RF) for the susceptibility classification of dust sources in the Middle East: A global source. Scientific Reports. 12(1): 19342.
- Guzzetti F, Carrara A, Cardinali M, Reichenbach P. 1999. Landslide hazard evaluation:

- A review of current techniques and their application in a multi-scale study, Central Italy. *Geomorphology*. 31(1-4): 181-216.
- Kanungo DP, Arora MK, Sarkar S, Gupta RP. 2012. Landslide susceptibility zonation (LSZ) mapping—a review.
- Kursa MB, Rudnicki WR. 2010. Feature selection with the Boruta package. *Journal of Statistical Software*. 36: 1-13.
- Lucà F, Conforti M, Robustelli G. 2011. Comparison of GIS-based gully susceptibility mapping using bivariate and multivariate statistics: Northern Calabria, South Italy. *Geomorphology*. 134(3-4): 297-308.
- Mina M, Rezaei M, Sameni A, Ostovari Y, Ritsema C. 2022. Predicting wind erosion rate using portable wind tunnel combined with machine learning algorithms in calcareous soils, southern Iran. *Journal of Environmental Management*. 304: 114171.
- Mohammadifar A, Gholami H, Comino JR, Collins AL. 2021. Assessment of the interpretability of data mining for the spatial modelling of water erosion using game theory. *Catena*. 200: 105178.
- Nasrnia F, Zibaei M. 2017. Determination farmers' vulnerability patterns to drought in Iran: case study of Bakhtegan Basin. *Agricultural Economics Research*. 9(34): 1-37.
- Nicodemus KK. 2011. On the stability and ranking of predictors from random forest variable importance measures. *Briefings in Bioinformatics*. 12(4): 369-373.
- Pournader M, Ahmadi H, Feiznia S, Karimi H, Peirovan HR. 2018. Spatial prediction of soil erosion susceptibility: an evaluation of the maximum entropy model. *Earth Science Informatics*. 11: 389-401.
- Rahmati O, Pourghasemi HR, Melesse AM. 2016. Application of GIS-based data driven random forest and maximum entropy models for groundwater potential mapping: a case study at Mehran Region, Iran. *Catena*. 137: 360-372.
- Rezaei M, Mohammadifar A, Gholami H, Mina M, Riksen MJ, Ritsema C. 2023. Mapping of the wind erodible fraction of soil by bidirectional gated recurrent unit (BiGRU) and bidirectional recurrent neural network (BiRNN) deep learning models. *Catena*. 223: 106953.
- Safaei M, Omar H, Yousof ZB, Ghiasi V. 2010. Applying geospatial technology to landslide susceptibility assessment. *Electronic Journal of Geotechnical Engineering*. 15(G): 677-696.
- Senanayake S, Pradhan B. 2022. Predicting soil erosion susceptibility associated with climate change scenarios in the Central Highlands of Sri Lanka. *Journal of Environmental Management*. 308: 114589.
- Talukdar S, Naikoo MW, Mallick J, Praveen B, Sharma P, Islam AR. MT., ... Rahman A. 2022. Coupling geographic information system integrated fuzzy logic-analytical hierarchy process with global and machine learning based sensitivity analysis for agricultural suitability mapping. *Agricultural Systems*. 196: 103343.
- Vafakhah M, Rajabi M. 2005. Efficiency of meteorological drought Indices for monitoring and assessment of drought in Bakhtegan, Tashk, and Maharlo Lakes Watershed.
- Van Westen CJ, Van Asch TW, Soeters R. 2006. Landslide hazard and risk zonation—why is it still so difficult?. *Bulletin of Engineering Geology and the Environment*. 65: 167-184.
- Yao X, Tham LG, Dai FC. 2008. Landslide susceptibility mapping based on support vector machine: A case study on natural slopes of Hong Kong, China. *Geomorphology*. 101(4): 572-582.
- Yesilnacar EK. 2005) *The application of computational intelligence to landslide susceptibility mapping in Turkey*. University of Melbourne, Department. 200 p.



Integrated Modeling of Soil Erosion by Water and Wind Using Machine Learning Methods

Mehdi Jalali¹, Hamid Gholami², Marzieh Rezaie³, Ebrahim Omidvar⁴

1-Ph.D. Student, Department of Natural Resources Engineering, Faculty of Agriculture and Natural Resource Engineering, University of Hormozgan, Bandar-Abbas, Iran

2- Associate Professor, Department of Natural Resources Engineering, Faculty of Agriculture and Natural Resource Engineering, University of Hormozgan, Bandar-Abbas, Iran

3- Assistant Professor, Department of Natural Resources Engineering, Faculty of Agriculture and Natural Resource Engineering, University of Hormozgan, Bandar-Abbas, Iran

4- Associate Professor, Department of Natural Resources Engineering, University of Kashan, Kashan, Iran

Extended Abstract

Introduction and objective

Soil erosion by water and wind is one of the biggest environmental threats worldwide which it has various negative consequences such as soil degradation, mitigation of soil fertility, depletion of nutrients and micro-elements, degradation of soil structure, dust storms, reservoir siltation, and etc. on the natural and anthropogenic ecosystems. Therefore, accurate information of land susceptibility to soil erosion hazard by water and wind, and production of spatial maps of these hazards are necessary for mitigation of their consequences. Therefore, the goal of this research is application of two machine learning models for the spatial modelling of water and wind erosion in the Bakhtegan basin, Fars province.

Materials and methods

In order to generate the spatial maps of soil erosion by water and wind, we have used 20 and 16

Article Type: Research Article

***Corresponding Author E-mail:** hgholami@hormozgan.ac.ir

Citation: Jalali, M., Gholami, H., Rezaie, M., Omidvar, E. 2023. Integrated modeling of erosion by water and wind using machine learning methods. *Watershed Management Research*. 36(3): 128-145.

DOI: 10.22092/WMRJ.2022.358127.1458

Received: 11 April 2022, **Received in revised form:** 16 August 2022, **Accepted:** 21 September 2022,

Published online: 23 September 2023

Watershed Management Research, VOL. 36, No.3, Ser. No: 140, Autumn 2023, pp.128-145.

Publisher: Fars Agricultural and Natural Resources Research and Education Center

©Author(s)



effective factors on water and wind erosion, respectively. Before the modelling stage, the multivariate adaptive regression spline (MARS) feature selection algorithm and multicollinearity test were used to identify the most important factors controlling water and wind erosion. In next stage, two random forest and support vector learning machine techniques were applied to model soil erosion. Area under curve (AUC) was applied to assess the model performance.

Result and discussion

According to results, land use, slope, lithology, roughness, coarse fragment, clay, vegetation cover, aspect, soil bulk density and topographic wetness index were identified as the most important factors controlling water erosion. The most important factors controlling wind erosion are including elevation, land use, wind speed, slope, vegetation cover, soil nitrogen, cation exchange capacity, lithology, aspect, clay and bulk density. The AUC value for wind erosion map generated by random forest model was 99, and for water erosion map produced by support vector machine is 96. Finally, we integrated two maps and results indicate an area of study area about 18% has very high susceptibility to wind erosion and low susceptibility to water erosion.

Conclusion and suggestions

Due to study area has several lakes consisting of Bakhtegan, Maharlo and Tashk, and also several dams, it is so important to supply water for different purposes such as drinking, agricultural, industry and environment. In case of not paying attention to this challenge especially wind erosion in the study area, this case study especially dried beds for lakes can play as sources for generating dust particles. Therefore, the findings of this research can use by managers to mitigate negative effects of this phenomena.

Key Words: Area under curve (AUC), Bakhtegan basin, multivariate adaptive regression spline (MARS), random forest (RF), support vector machine (SVM) model