

مقاله علمی

مطالعات جغرافیایی مناطق خشک

دوره دوازدهم، شماره چهل و هفتم، بهار ۱۴۰۱

دریافت مقاله: ۱۴۰۰/۰۲/۱۸ تأیید نهایی: ۱۴۰۰/۰۸/۲۶

صص ۱-۱۳

پهنه‌بندی آسیب‌پذیری گردوغبار با استفاده از تصاویر ماهواره‌ای و مدل‌های یادگیری ماشین

مه‌دی بروغنی^{*}، استادیار مرکز پژوهشی علوم جغرافیا و مطالعات اجتماعی، دانشگاه حکیم سبزواری، سبزوار، ایران
فهیمه میرچولی، دانشجوی پسادکتری مرکز پژوهشی علوم جغرافیا و مطالعات اجتماعی، دانشگاه حکیم سبزواری،
سبزوار، ایران

مازیار محمدی، دکتری گروه مهندسی آب‌خیزداری، دانشکده‌ی منابع طبیعی و علوم دریایی، دانشگاه تربیت
مدرس، نور، ایران

چکیده

پیش‌بینی کانون‌های برداشت گردوغبار و تعیین عوامل مؤثر بر آن، برای اولویت‌بندی اقدامات مدیریتی و اجرایی به منظور مقابله با بیابان‌زایی ناشی از فرسایش بادی در مناطق خشک ضروری است؛ بنابراین، این کار با هدف ارزیابی کاربرد سه مدل یادگیری ماشین (از جمله مدل‌های جنگل تصادفی، گرادیان تقویتی و مدل افزودنی کلی) برای پیش‌بینی آسیب‌پذیری کانون‌های گردوغبار طی سال‌های ۲۰۰۵ تا ۲۰۱۸ در کویر مرکزی انجام شد. برای این منظور، ابتدا کانون‌های گردوغبار در منطقه‌ی مطالعاتی با استفاده از تصاویر ماهواره‌ای MODIS با استفاده از چهار شاخص شامل BTD_{3132} ، BTD_{2931} ، $NDDI$ و متغیر D استخراج شدند و در نهایت ۱۳۵ نقطه به‌عنوان کانون گردوغبار شناسایی و در مدل‌سازی استفاده شدند. در این مطالعه برخی فاکتورهای مؤثر بر گردوغبار مانند کاربری اراضی، خاک‌شناسی، زمین‌شناسی، فاصله از آبراهه، شاخص تفاوت نرمال‌شده‌ی پوشش گیاهی (NDVI)، شیب زمین و اقلیم برای مدل‌سازی در نظر گرفته شدند. نتایج حاصل نشان داد که در میان الگوریتم‌های استفاده شده، مدل گرادیان تقویتی با دقت ۶۴/۲ درصد، دقیق‌ترین مدل و سپس مدل جنگل تصادفی با دقت ۶۳/۵ درصد و مدل افزودنی کلی با دقت ۵۱/۶ درصد در رتبه‌های بعدی قرار دارند. به‌علاوه، از میان فاکتورهای مورد بررسی، کاربری اراضی و خاک‌شناسی به‌عنوان مؤثرترین عوامل بر آسیب‌پذیری گردوغبار شناسایی شدند. نتایج حاصل از این مطالعه می‌تواند اطلاعات ارزشمندی را به مدیران منطقه‌ای و سیاستمداران برای شناسایی مناطق آسیب‌پذیر، اولویت‌بندی انجام فعالیت‌های مدیریتی برای کنترل گردوغبار و نیز اتخاذ تصمیم‌های مناسب برای کاهش پیامدهای منفی آن ارائه داده و به آن‌ها برای مدیریت مناسب‌تر کمک نماید.

واژگان کلیدی: سنجش‌ازدور، کانون گردوغبار، کویر مرکزی، آسیب‌پذیری.

۱- مقدمه

پدیده‌ی گردوغبار یکی از مشکلات مهم محیط زیستی است که بسیاری از مناطق خشک جهان را تحت تأثیر قرار داده است؛ به طوری که حدود یک‌سوم مناطق جهان تحت تأثیر فرسایش بادی قرار دارند (Gholami et al, 2021b, 4). در همین راستا، ایران به‌عنوان یکی از کشورهای واقع در کمربند خشک و نیمه‌خشک، در معرض سیستم‌های متعدد گردوغبار فرا منطقه‌ای و محلی است (پورهایمی و همکاران، ۱۳۹۸). پدیده‌ی گردوغبار به‌عنوان یکی از شاخص‌های تخریب زمین، در سال‌های اخیر به صورت بحران جدی در بسیاری از استان‌های کشور در آمده است. در این پدیده، خاک فرسایش‌یافته توسط باد با سرعت بالاتر از سرعت آستانه، حرکت داده می‌شود و ذرات بر اساس اندازه‌ی خود به یکی از سه روش تعلیق، جهش و خزش حرکت می‌کنند.

از جمله پیامدهای این پدیده می‌توان به اثر منفی آن در روند تشکیل خاک، حرکت دادن پوشش سطحی، آلودگی هوا، انتقال عوامل بیماری‌زا به انسان، تشدید بیابان‌زایی، تسریع فرسایش خاک اشاره کرد؛ بنابراین، دامنه‌ی تأثیر آن وسیع بوده و بر تمامی جنبه‌های اقتصادی، اجتماعی، وضعیت سلامت، آب‌وهوا تأثیرگذار است (Gholami et al, 2020a, 7؛ Gholami et al, 2020b). به‌علاوه، عواملی مانند تغییرات اقلیمی، خشک‌سالی‌های مداوم و دخالت‌های غیراصولی انسان در طبیعت باعث تشدید احتمال وقوع گردوغبار می‌شود (کرمانی و همکاران، ۱۳۹۵).

برای مقابله با پدیده‌ی گردوغبار، شناسایی مناطق برداشت گردوغبار و تمرکز بر مناطق بحرانی، بر هر کاری اولویت دارد (حیدریان و همکاران، ۱۳۹۶، ۳۷). در سال‌های اخیر، با افزایش فراوانی و وسعت تحت تأثیر گردوغبار، شناسایی و پایش آن‌ها امری ضروری و اجتناب‌پذیر شده است. استفاده از ایستگاه‌های زمینی به دلیل دوام کم ذرات موجود در هوا و تغییرات وسیع مکانی آن‌ها، با محدودیت روبه‌رو است. اگرچه عکس‌های هوایی، امکان رؤیت مناطق با ابعاد وسیع‌تر از ایستگاه‌های زمینی را فراهم می‌کند، اما ماهواره‌ها و تصاویر ماهواره‌ای دید بهتری از زمین را ارائه می‌دهند؛ به طوری که تصاویر ماهواره‌ای به دلیل هزینه کم، کاهش زمان، پوشش وسیع زمین و تکرارپذیری برداشت تصاویر، به‌عنوان روشی مناسب برای پایش و ارزیابی گردوغبار در بسیاری از پژوهش‌ها شناخته شده است (مهرابی و همکاران، ۱۳۹۴). به‌عنوان مثال، مهرابی و همکاران (۱۳۹۴) با استفاده از تصاویر مادیس به شناسایی طوفان گردوغبار با استفاده از شاخص تفاضل نرمال شده گردوغبار^۱ (NDDI) در استان خوزستان پرداختند. نتایج حاصل از مطالعه‌ی ایشان نشان داد که علی‌رغم کارایی خوب این شاخص در مطالعات دیگر، میزان شدت طوفان در این مطالعه کمتر از مقدار واقعی برآورد شده است.

در همین راستا، پهنه‌بندی آسیب‌پذیری کانون‌های گردوغبار، به‌عنوان گام اول برای مدیریت و کاهش گردوغبار در نظر گرفته می‌شود. این نقشه به سیاستمداران و مدیران برای شناسایی مناطق آسیب‌پذیر و اتخاذ تصمیم‌های مناسب برای کاهش خطر گردوغبار کمک می‌کند. مطالعات موجود در این زمینه، از روش‌های و مدل‌های مختلفی استفاده کرده‌اند که از این میان، مدل‌های داده‌کاوی به‌عنوان یکی از روش‌های موفق در این زمینه مطرح است. به‌عنوان مثال، بروغنی و همکاران (۲۰۲۰) به شناسایی کانون‌های گردوغبار با استفاده از تصاویر مادیس در استان خراسان رضوی پرداختند. هم‌چنین ایشان از مدل‌های داده‌کاوی آماری یعنی مدل‌های نسبت فراوانی (FR)، جنگل تصادفی (RF) و وزن واقعه (WOE) برای تهیه‌ی نقشه‌ی آسیب‌پذیری کانون‌های گردوغبار استفاده کردند. نتایج حاصل نشان داد که مدل جنگل تصادفی با دقت ۹۱ درصد، بهترین کارایی را داشته است (بروغنی و همکاران، ۲۰۲۰). غلامی و همکاران (۲۰۲۱) با استفاده از مدل‌های داده‌کاوی و مدل اقلیمی RegCM4 به پهنه‌بندی آسیب‌پذیری استان یزد به پدیده‌ی

1- Normalized Difference Dust Index (NDDI)

گردوغبار و تعیین میزان اهمیت معیارهای مرتبط انتخابی، پرداختند. نتایج حاصل نشان داد که بخش وسیعی از منطقه در طبقه بسیار آسیب‌پذیر قرار داشته و معیارهای زمین‌شناسی، شاخص تفاضل نرمال‌شده‌ی پوشش گیاهی^۱ (NDVI) و ارتفاع به‌عنوان مهم‌ترین عوامل مؤثر شناسایی شدند. در پژوهشی دیگر ابراهیمی خوسفی و همکاران (۲۰۲۱) با استفاده از چند روش ماشین آموزش مانند جنگل تصادفی، رگرسیون تطبیقی چند متغیره، کوچک‌ترین عملگر مطلق، نزدیک‌ترین همسایگی، برنامه‌نویسی ژنتیکی و چند روش دیگر به تهیه‌ی نقشه‌های پیش‌بینی طوفان گردوغبار فصلی در سال‌های ۲۰۰۰ تا ۲۰۱۸ در نواحی خشک ایران پرداختند. ایشان در تحقیق خود گزارش کردند که سرعت باد در زمستان بیشترین اهمیت را در فرآیند پیش‌بینی داشته و نقشه‌ی میانگین روش‌های ماشین آموزش بالاترین دقت را ارائه می‌کند.

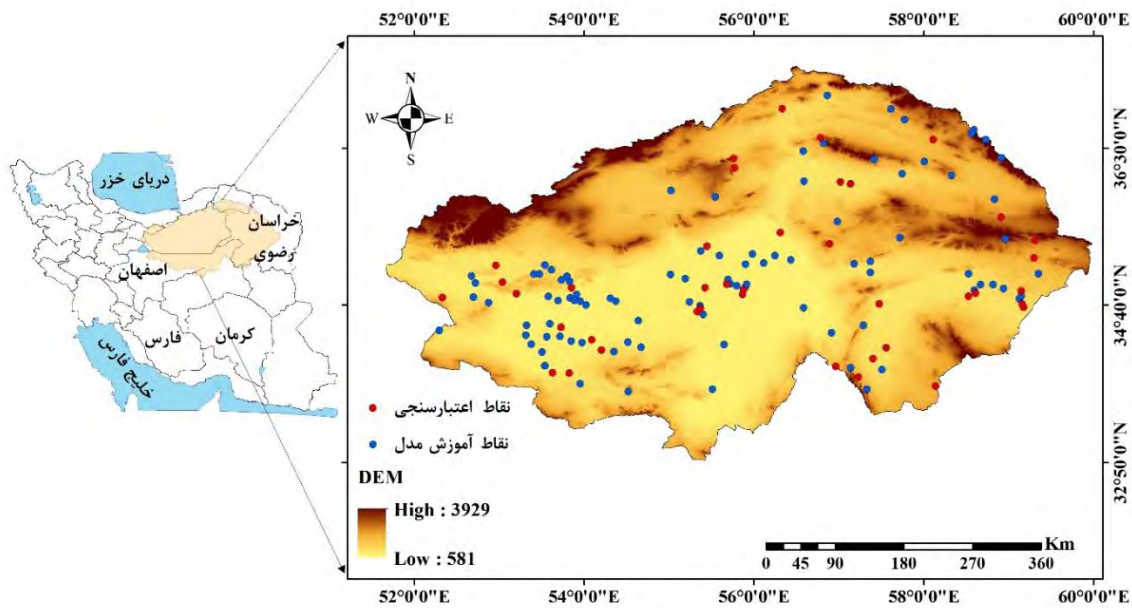
همچنین به منظور پهنه‌بندی مناطق مستعد برداشت گردوغبار در شرق ایران، از روش‌های جنگل تصادفی، ماشین بردار پشتیبان و خطوط رگرسیون تطبیقی چندمتغیره و تصاویر مادیس استفاده شد. نتایج این تحقیق نشان داد که روش جنگل تصادفی بالاترین دقت را در پیش‌بینی کانون‌های گردوغبار ارائه می‌کند (Rahmati et al, 2020). با توجه به اهمیت حوضه‌ی آبخیز کویر مرکزی و نیز موقعیت سیاسی، اجتماعی و اقتصادی و نیز افزایش فراوانی پدیده‌ی گردوغبار در این حوضه، به‌عنوان منطقه‌ی مطالعاتی انتخاب شد. در همین راستا، این پژوهش با هدف شناسایی کانون‌های گردوغبار و پهنه‌بندی آسیب‌پذیری گردوغبار از تصاویر ماهواره مادیس و برخی معیارهای مؤثر بر گردوغبار و نیز مدل‌های آماری جنگل تصادفی، گرادیان تقویتی و مدل افزودنی کلی استفاده شد. در ادامه، برای بررسی کارایی مدل‌ها و انتخاب بهترین مدل، شاخص‌های سنجش دقت مورد استفاده قرار گرفت.

۲- مواد و روش

۲-۱- منطقه‌ی مطالعاتی

حوضه‌ی آبخیز کویر مرکزی در خشک‌ترین قسمت ایران واقع شده و بخشی از حوضه‌ی آبخیز درجه یک فلات مرکزی محسوب می‌شود (بنی حبیب و همکاران، ۱۳۹۵). این منطقه در محدوده‌ی ۲۶ درجه و ۳۳ دقیقه تا ۳۷ درجه و ۲۸ دقیقه عرض شمالی و ۴۸ درجه و ۸ دقیقه تا ۶۱ درجه و ۲۵ دقیقه طول شرقی واقع شده است. این منطقه دارای مساحتی معادل ۸۳۱۱۰۲ کیلومترمربع است که مرتفع‌ترین نقطه‌ی آن، ۴۴۷۱ متر و کم‌ارتفاع‌ترین آن، کمتر از ۱۰۰ متر است. میانگین بارش سالانه در این حوضه، ۲۳/۴ میلی‌متر است.

1- Normalized Difference Vegetation Index (NDVI)



شکل ۱: موقعیت جغرافیایی منطقه‌ی مطالعاتی و نحوه‌ی پراکنش کانون گردوغبار مستخرج از تصاویر ماهواره‌ای

۲-۲- روش پژوهش

۲-۲-۱- استخراج کانون‌های گردوغبار منطقه

در این مطالعه، برای شناسایی کانون‌های گردوغبار، از تصاویر ماهواره‌ای MODIS استفاده شد. این تصاویر دارای ۳۶ باند با طول‌موج‌های ۰/۴۱ تا ۴۱۴ میکرومتر بوده که تفکیک مکانی آن‌ها متفاوت و ۲۵۰، ۵۰۰ و ۱۰۰۰ متر است. تصاویر MODIS از ماهواره Terra و Aqua برای روزهای گردوغباری موجود در تاریخ ۲۰۰۵-۲۰۱۸ انتخاب و دانلود شدند. در ادامه این تصاویر، برای محاسبه شاخص‌های کانون‌های برداشت گردوغبار، زمین‌مرجع شدند و تصحیحات لازم روی آن‌ها اعمال شد.

در همین راستا، برای شناسایی پدیده‌ها، به‌منظور ایجاد ترکیب رنگی حقیقی، از باندهای ۱ (۰/۶۷-۰/۶۲ میکرومتر)، ۴ (۰/۵۶۵-۰/۵۴۵ میکرومتر) و ۳ (۰/۴۷۹-۰/۴۵۹ میکرومتر) به‌عنوان محدوده‌های قرمز، سبز و آبی استفاده شده است. بعد از اعمال ترکیب رنگی حقیقی و کاذب بر روی تصاویر مادیس، شاخص‌های NDDI، BTDD به صورت زیر محاسبه شده است:

شاخص NDDI: با توجه به این‌که انعکاس گردوغبار با افزایش طول‌موج افزایش می‌یابد و حداقل آن در باند ۳ (۰/۴۶۹ میکرومتر) و حداکثر آن در باند ۷ (۲/۱۳ میکرومتر) است، شاخص NDDI با استفاده از این دو باند و مطابق معادله‌ی (۱) محاسبه می‌شود:

معادله‌ی (۱)

$$NDDI = \frac{b_7 - b_3}{b_7 + b_3}$$

در این معادله، b_3 و b_7 به ترتیب باند ۳ و ۷ در تصاویر مادیس می‌باشند (Mei et al, 2008, 967).

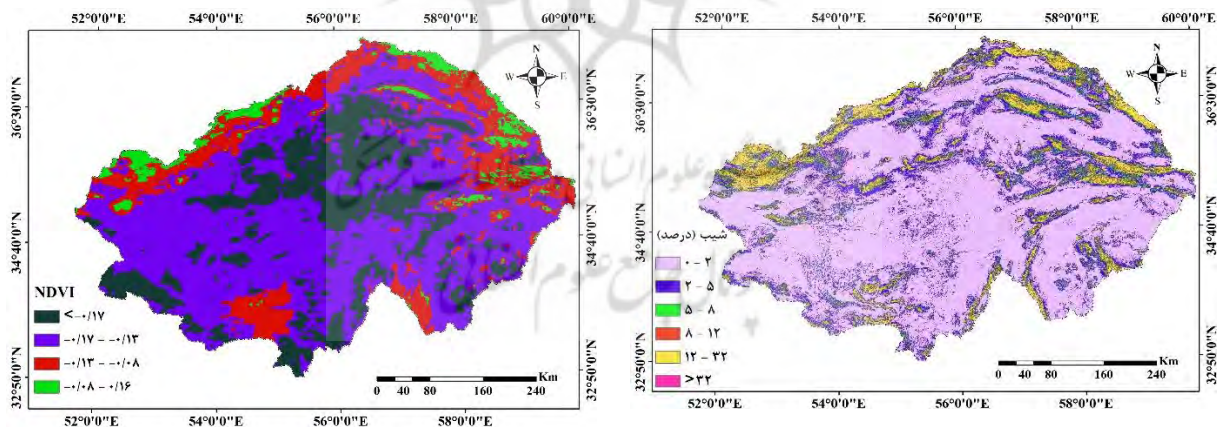
شاخص BTDI: باندهای ۳۱ و ۳۲ تصاویر مادیس در محدوده‌ی مادون قرمز حرارتی قرار دارند که در این محدوده، جذب توسط گازهای اتمسفری بسیار ناچیز است. در این باندها، میزان اختلاف تابش بین زمین و گردوغبار بالاست. به علاوه، گردوغبار تابش بالایی در باند ۳۱ نسبت به باند ۳۲ دارد؛ از این باندها می‌توان برای تشخیص پدیده‌ی گردوغبار مطابق معادله‌ی (۲) استفاده کرد:

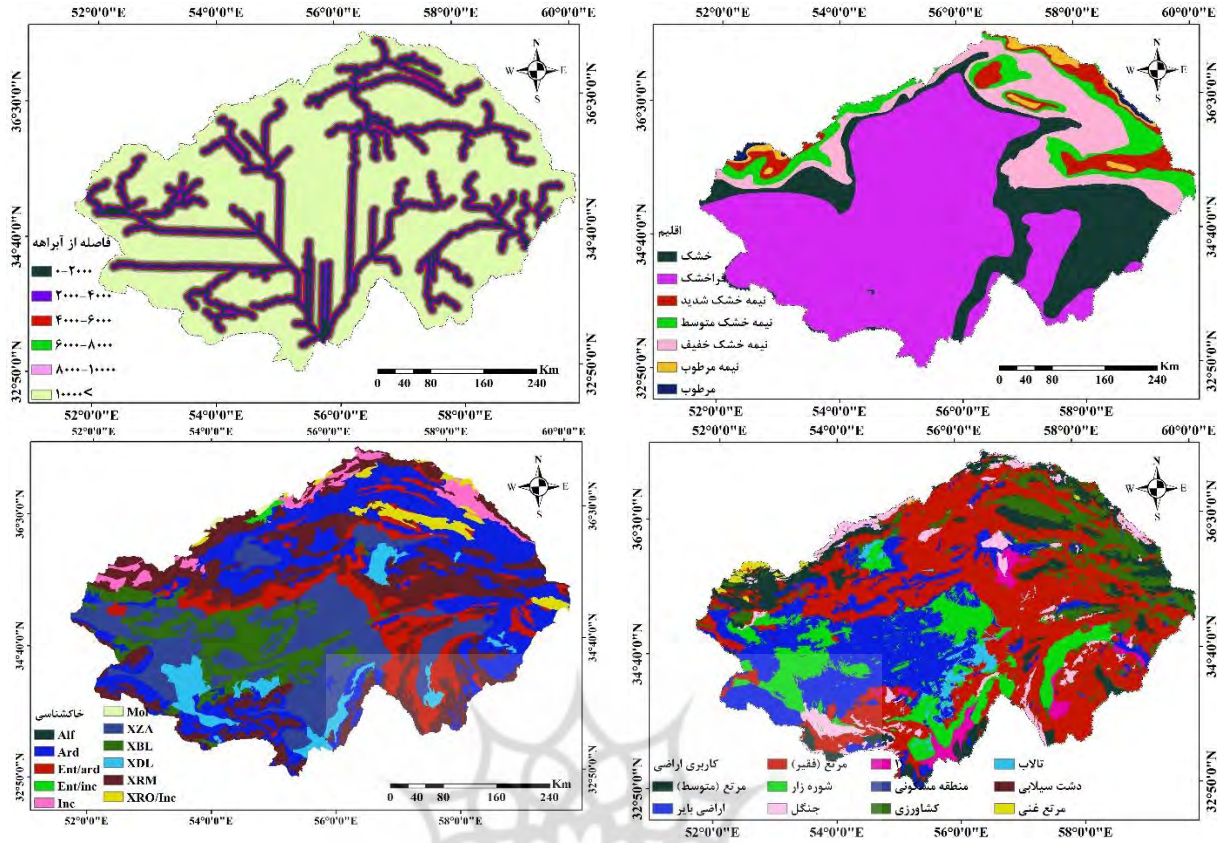
$$BTDI = CH32 - CH31 \quad \text{معادله‌ی (۲)}$$

در این معادله، باند ۳۲ دارای طول موج ۱۱/۷۷ میکرومتر و باند ۳۱ دارای طول موج ۱۰/۷۸ میکرومتر می‌باشند (Baddock et al, 2009).

۲-۲-۲- پهنه‌بندی آسیب‌پذیری گردوغبار

برای انجام این قسمت از پژوهش حاضر، ابتدا نقشه و لایه‌های اطلاعاتی (شکل ۲) با توجه به شناخت منطقه و سابقه‌ی پژوهش در زمینه‌ی عوامل مؤثر بر پهنه‌بندی آسیب‌پذیری پدیده‌ی گردوغبار تهیه شدند. این لایه‌ها شامل نقشه‌ی کاربری اراضی، درجه‌ی شیب، زمین‌شناسی، خاک‌شناسی، فاصله از آبراهه، شاخص تفاضل نرمال شده‌ی پوشش گیاهی (NDVI) و اقلیم می‌باشند که با استفاده از الحاقیه‌های مختلف در محیط سامانه‌ی اطلاعات جغرافیایی (Arcgis 10.4) تهیه شدند. نقشه‌ی زمین‌شناسی از سازمان زمین‌شناسی کشور تهیه و رقومی شدند. لایه‌ی مربوط به آبراهه و نیز درجه‌ی شیب در محیط Arcgis و با استفاده از نقشه‌ی رقومی ارتفاع (DEM) به دست آمد. نقشه‌ی خاک‌شناسی و نقشه‌ی کاربری اراضی از اداره‌ی منابع طبیعی تهیه شد. لایه‌ی NDVI نیز با استفاده از تصاویر مادیس به دست آمد.





شکل ۲: لایه‌های تأثیرگذار بر پهنه‌بندی آسیب‌پذیری گردوغبار در کویر مرکزی ایران

۲-۲-۲- جنگل تصادفی (RF)

الگوریتم جنگل تصادفی یکی از بهترین الگوریتم‌های یادگیری است که اولین بار توسط لئو برایمن و آدل کاتلر توسعه داده شد. این الگوریتم نوع توسعه‌یافته‌ای از مدل طبقه‌بندی و رگرسیون درختی^۱ است (Breiman, 2001). الگوریتم جنگل تصادفی مبتنی بر روش ترکیب اطلاعات است که در آن تعداد زیادی درخت تصمیم ایجاد شده و سپس نتایج تمام درختان برای پیش‌بینی ترکیب می‌شوند. در همین راستا، استفاده از چندین درخت تصمیم‌گیری حساسیت و بی‌ثباتی و مدل را کاهش داده و منجر به افزایش قدرت و دقت مدل می‌شود. درختان تصمیم‌گیری با در نظر گرفتن بردار ورودی، آن را با هر درخت در جنگل طبقه‌بندی کرده و خروجی، برچسب‌های کلاسی هستند که از اکثریت آرا دریافت شده است. درختان در این مدل با استفاده از روش بوت استرپ ایجاد شده و در ساخت هر درخت، تعدادی از متغیرهای ورودی به‌طور تصادفی به‌منظور جلوگیری و کاهش تطابق بین درخت‌های تصمیم‌گیری شرکت خواهند کرد. از جمله مزیت‌های این روش برای ساخت درخت می‌توان به کاهش ناسازگاری در نتایج مدل و افزایش قابلیت پیش‌بینی و نیز عدم نیاز به کالیبراسیون اضافی، اشاره کرد (Pourghasemi and Kerle, 2016).

1- Classification and regression tree (CART)

۲-۲-۴- مدل گرادیان تقویتی (GBM)

روش دیگر برای برقراری ارتباط بین متغیرهای وابسته و مستقل، مدل گرادیان تقویتی است که به‌عنوان یک مدل آماری به‌طور فزاینده‌ای توسط پژوهشگران مورد استفاده قرار می‌گیرد. این مدل توسط فریدمن در سال ۲۰۰۱ معرفی شد. در این مدل، درختان در یک زمان ساخته می‌شوند، به‌طوری‌که هر درخت جدید به تصحیح خطاهای مربوط به درختان آموزشی قبلی کمک می‌کند. شروع کار این مدل با یک درخت تصمیم‌گیری است که برای هر یک از مشاهدات یک وزن برابر با آن مشخص می‌شود. بعد از ارزیابی درخت اول، وزن مربوط به مشاهداتی را که طبقه‌بندی راحت‌تری دارند، کاهش و در مقابل وزن مشاهدات با طبقه‌بندی سخت‌تر را افزایش می‌دهیم؛ بنابراین درخت دوم، بر این داده‌های وزنی رشد می‌کند. این کار برای بهبود پیش‌بینی درخت اول انجام می‌شود. این روند تکرار می‌شود و برای تعداد مشخصی از تکرارها، درختان بعدی در طبقه‌بندی مشاهداتی به افزایش دقت پیش‌بینی کمک خواهند کرد (Du et al, 2013).

۲-۲-۵- مدل افزودنی کلی (GAM)

مدل افزودنی کلی توسط Hastie و Tibshirani (۱۹۹۰) توسعه یافت که یک مدل خطی تعمیم‌یافته (GLM) با متغیرهای تأثیرگذار خطی شامل مجموع توابع صاف متغیرهای تأثیرگذار است. این مدل، به‌عنوان مکمل نیمه‌پارامتری مدل GLM شناخته می‌شود. به‌طور کلی، یک مدل افزودنی کلی مطابق معادله‌ی زیر تعریف می‌شود:

$$g(\mu_i) = X_i\beta + \sum_{j=1}^m f_j(x_{i,j}) \quad (3)$$

که در آن، $\mu_i = E(Y_i)$ توزیع نمایی (مانند توزیع نرمال یا پویسون)، X_i ردیف i ام ماتریس در مدل پارامتریک است، f_j توابع صاف متغیرهای x_j است.

۲-۲-۶- اعتبارسنجی نقشه‌های به‌دست آمده

بعد از تهیه نقشه‌ی آسیب‌پذیری گردوغبار با استفاده از مدل‌های جنگل تصادفی، گرادیان تقویتی و مدل افزودنی کلی، باید ارزیابی مدل‌ها و اعتبارسنجی آن‌ها انجام شود. برای این کار از منحنی ROC استفاده شد که یکی از کارآمدترین روش‌ها برای ارائه‌ی خصوصیات تعیینی و پیش‌بینی سیستم‌هاست و میزان دقت مدل را به صورت کمی برآورد می‌کند. سطح زیر منحنی ROC، AUC نام دارد که نشان‌دهنده‌ی مقدار پیش‌بینی سیستم از طریق توصیف توانایی آن در برآورد درست کانون گردوغبار و عدم وجود کانون برداشت گردوغبار است.

۳- نتایج و بحث

در مرحله‌ی اول، قبل از تجزیه و تحلیل مدل‌ها، محاسبه همبستگی بین متغیرهای تأثیرگذار با استفاده از ضریب تورم واریانس (VIF) که مهم‌ترین شاخص برای اندازه‌گیری چند همبستگی است، انجام شد. این فرآیند به این دلیل صورت گرفت که چندخطی بودن بین متغیرها، دقت پیش‌بینی مدل‌ها را کاهش می‌دهد. نتایج آزمون چندخطی (جدول ۱) تأیید کرد که هیچ هم‌خوانی لایه‌های تأثیرگذار (متغیرهای مستقل) وجود ندارد و به همین خاطر از تمامی لایه‌های برای مدل‌سازی استفاده شد.

جدول ۱: مقادیر ضریب تورم واریانس (VIF) برای هر یک از متغیرهای تأثیرگذار در آسیب‌پذیری گردوغبار

| VIF | متغیر تأثیرگذار |
|------|-----------------|
| ۱/۱۵ | کاربری اراضی |
| ۱/۱۷ | خاک‌شناسی |
| ۲/۱۴ | شیب زمین |
| ۱/۳۰ | NDVI |
| ۱/۹۰ | زمین‌شناسی |
| ۱/۱۱ | فاصله از آبراهه |
| ۱/۲۳ | اقلیم |

نتایج حاصل از مدل گرادیان تقویتی و مدل جنگل تصادفی نشان می‌دهد که عوامل خاک‌شناسی و کاربری اراضی بیش‌ترین تأثیر را بر آسیب‌پذیری گردوغبار در منطقه‌ی مطالعاتی داشته است (جدول ۲). عوامل زمین‌شناسی، فاصله از آبراهه، NDVI، شیب زمین و اقلیم به ترتیب در رتبه‌های بعدی قرار گرفتند. این نتایج با مطالعه‌ی غلامی و همکاران (۲۰۲۰) مبنی بر تأثیر فاکتور خاک‌شناسی مغایرت دارد. ایشان در مطالعه‌ی خود با هدف پیش‌بینی آسیب‌پذیری زمین نسبت به انتشار گردوغبار به این نتیجه رسیدند که بر خلاف فاکتورهای بارندگی و ارتفاع و سرعت باد که دارای بیش‌ترین تأثیر در میان مجموعه‌ی فاکتورهای مطالعاتی بودند، فاکتور خاک‌شناسی در انتشار گردوغبار چندان مؤثر نبوده است. ایشان تأثیر کاربری اراضی بر گردوغبار را متوسط ارزیابی کردند که در این مورد نیز با نتایج حاصل مطابقت ندارد. رحمتی و همکاران (۲۰۲۰) نیز در تحقیق خود گزارش کردند که نقشه‌ی خاک و کاربری اراضی بیش‌ترین تأثیر و در مقابل عواملی چون شیب، زمین‌شناسی و توپوگرافی کمترین تأثیر را در مدل‌سازی مناطق گردوغبار با استفاده از روش جنگل تصادفی داشته‌اند. پوشش و کاربری اراضی بر سرعت باد تأثیر می‌گذارد. ویژگی‌های خاک نیز اغلب منعکس‌کننده‌ی میزان و پوشش اراضی است که با ایجاد بادگیرهایی طبیعی باعث کاهش سرعت باد می‌شود.

نتایج حاصل از مدل افزودنی کلی نشان می‌دهد که فاکتورهای خاک‌شناسی و فاصله از آبراهه به ترتیب بیش‌ترین تأثیر را بر آسیب‌پذیری گردوغبار داشته است که در رتبه‌های بعدی به ترتیب فاکتورهای شیب، کاربری اراضی، اقلیم و NDVI قرار دارند. نتایج این مطالعه با مطالعه‌ی بروغنی و همکاران (۲۰۲۰) مبنی بر تأثیر کمتر شاخص NDVI در مقایسه با سایر فاکتورهای تأثیرگذار مطابقت دارد.

جدول ۲: میزان اهمیت فاکتورهای تأثیرگذار بر آسیب‌پذیری گردوغبار در مدل‌های RF، GBM و GAM

| GAM | GBM | RF | متغیر تأثیرگذار |
|-------|------|------|-----------------|
| ۰/۰۷ | ۰/۱۴ | ۰/۱۴ | کاربری اراضی |
| ۰/۵۹ | ۰/۳۳ | ۰/۲۸ | خاک‌شناسی |
| ۰/۰۹ | ۰/۰۲ | ۰/۰۶ | شیب زمین |
| ۰/۰۱ | ۰/۰۵ | ۰/۰۶ | NDVI |
| ۰/۰۰۸ | ۰/۰۳ | ۰/۱۳ | زمین‌شناسی |
| ۰/۱۹ | ۰/۱۸ | ۰/۱۰ | فاصله از آبراهه |
| ۰/۰۴ | ۰/۰۳ | ۰/۰۶ | اقلیم |

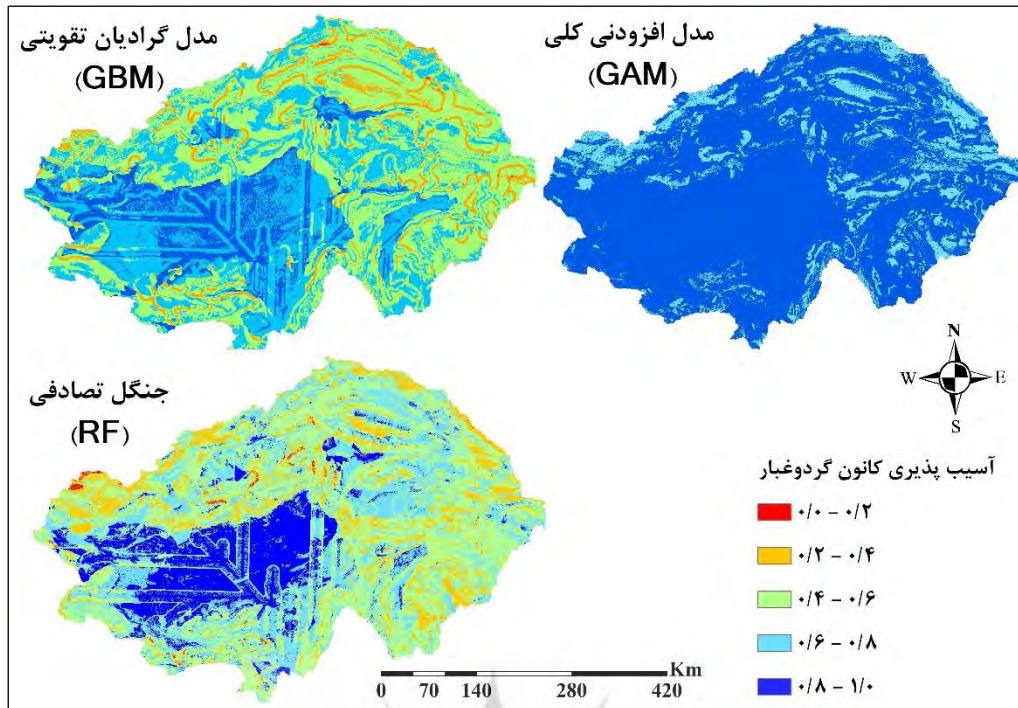
مطالعات همچنین نشان داده است که مدیریت اراضی و فعالیت‌های انسانی به شدت بر الگوی مکانی تولید گردوغبار تأثیر می‌گذارد (Lee et al, 2012). از آنجایی که که آستانه‌ی فرسایش مستقیماً به خصوصیات خاک گره خورده است، بنابراین وقوع طوفان گردوغبار به شدت تحت تأثیر نوع خاک است. اکثر منابع مستعد گردوغبار در منطقه‌ی دریاچه‌های فصلی یا پلایا در نواحی خشک و نیمه‌خشک است که با سطوح پوشیده شده از ذرات ریز است (Goudie, 2018). سایر مطالعات نیز بر نقش مهم خاک در تولید ریزگرد تأکید داشته و آن را به‌عنوان عاملی مهم در تحقیق خود برای مدل‌سازی و پیش‌بینی نقشه‌های پیش‌بینی کانون برداشت گردوغبار با روش‌های ماشین آموزش معرفی کرده‌اند (Sissakian et al, 2013; Hahnenberger and Nicoll, 2014).

در ادامه، نقشه‌ی پهنه‌بندی آسیب‌پذیری گردوغبار با استفاده از الگوریتم‌های انتخابی مطابق شکل (۳) به دست آمد و آنالیز درصد مساحت طبقه‌های مختلف آسیب‌پذیری گردوغبار با استفاده از هر سه مدل انجام شد. نتایج حاصل از مدل گرادیان تقویتی به‌عنوان دقیق‌ترین مدل نشان داد که بخش عمده‌ای از مساحت منطقه یعنی ۴۲/۱۰ درصد از آن، در طبقه ۰/۶-۰/۴ و شرایطی با آسیب‌پذیری متوسط قرار دارد. سپس، طبقه با آسیب‌پذیر نسبتاً زیاد با مساحت ۳۶/۸۹ درصد و طبقه آسیب‌پذیری بالا با مساحت ۱۳/۰۵ درصد از کل منطقه را پوشش می‌دهند.

نتایج حاصل از تجزیه و تحلیل مساحت تحت تأثیر طبقات مختلف آسیب‌پذیری گردوغبار با استفاده از مدل جنگل تصادفی نشان می‌دهد که حدود ۸۷/۸۰ درصد از منطقه در مجموع در طبقات با آسیب‌پذیری متوسط، نسبتاً زیاد و زیاد قرار گرفته است. این مقدار با توجه به موقعیت و اهمیت منطقه مطالعاتی، مساحت قابل توجهی است که لزوم به‌کارگیری مدیریت و استراتژی‌های راهبردی را برای کاهش خطر پیامدهای منفی آن را نشان می‌دهد.

این نتایج برای مدل افزودنی کلی نشان می‌دهد که بخش وسیعی از منطقه (۸۲/۷۱ درصد) تحت تأثیر آسیب‌پذیری زیاد (۱/۰-۰/۸) است. در رتبه بعدی، طبقات با آسیب‌پذیری نسبتاً زیاد با سطح تحت پوشش ۱۷/۲۸ درصد و طبقه با آسیب‌پذیری نسبتاً کم با مساحت ۰/۱ قرار دارند.

به‌علاوه، پهنه‌بندی آسیب‌پذیری گردوغبار در منطقه‌ی مطالعاتی با استفاده از سه مدل مذکور نشان می‌دهد که مناطق مرکزی و جنوب شرقی منطقه، نقاط به شدت آسیب‌پذیر بوده که لازم در انجام کارهای مدیریتی به‌منظور کاهش خطر گردوغبار در اولویت قرار گیرند.



شکل ۳: نقشه‌ی آسیب‌پذیری کانون‌های گردوغبار در منطقه‌ی کویر مرکزی با استفاده از مدل جنگل تصادفی، گرادیان تقویتی و مدل افزودنی کلی

در ادامه برای تعیین دقیق‌ترین مدل از منحنی ROC و TSS استفاده شد. نتایج حاصل از ارزیابی مدل‌ها نشان می‌دهد که مدل گرادیان تقویتی با دقت ۶۴/۲ درصد، دقیق‌ترین مدل در پهنه‌بندی آسیب‌پذیری کانون‌های برداشت گردوغبار است. در رتبه‌ی بعدی و با فاصله خیلی کم، مدل جنگل تصادفی با دقت ۶۳/۵ درصد و سپس مدل افزودنی کلی با دقت ۵۱/۶ درصد قرار دارند.

به‌علاوه، ارزیابی مدل با استفاده از TSS، دقت مدل گرادیان تقویتی را در مقایسه با سایر مدل‌ها تأیید کرد؛ به‌طوری‌که دقت این مدل معادل ۳۸/۶۰ درصد برآورد شد، اما این مقدار برای مدل جنگل تصادفی و مدل افزودنی کلی به ترتیب ۳۲/۴۵ و ۳۱/۶ درصد به دست آمد.

۴- نتیجه‌گیری

در این مطالعه، کانون‌های برداشت گردوغبار با استفاده از تصاویر MODIS برای دوره‌ی ۲۰۱۵-۲۰۲۰ شناسایی شد و نقشه‌ی توزیع کانون برداشت گردوغبار برای منطقه‌ی کویر مرکزی واقع در قسمت مرکزی ایران تهیه شد. در ادامه، چندین لایه داده‌ی مکانی یعنی فاکتورهای تأثیرگذار در ایجاد کانون برداشت گردوغبار از جمله خاک‌شناسی، زمین‌شناسی، شیب، NDVI، فاصله‌ی رودخانه، اقلیم و کاربری زمین تهیه شده است. کانون‌های برداشت گردوغبار شناسایی شده همراه با متغیرهای تأثیرگذار در مدل جنگل تصادفی، گرادیان تقویتی و مدل افزودنی کلی برای تولید نقشه‌ی آسیب‌پذیری کانون‌های برداشت گردوغبار استفاده شدند و در نهایت نقشه‌ی به‌دست‌آمده از این مدل‌ها طبقه‌بندی شد.

نتایج حاصل نشان داد که از میان هفت متغیر تأثیرگذار بر کانون‌های برداشت گردوغبار، خاک‌شناسی و کاربری اراضی در مدل‌گرادیان تقویتی بیشترین تأثیر را داشته است که در این مدل حدود ۹۲/۰۵ درصد از منطقه در طبقات با آسیب‌پذیری متوسط، نسبتاً زیاد و زیاد قرار گرفته‌اند. به‌علاوه، نتایج حاصل از ارزیابی دقت مدل‌های پهنه‌بندی گردوغبار با استفاده از روش ROC نشان داد که مدل‌گرادیان تقویتی با دقت ۶۴/۲ درصد، دقیق‌ترین مدل بوده و در رتبه‌ی بعدی، مدل جنگل تصادفی با دقت ۶۳/۵ درصد و سپس مدل افزودنی کلی با دقت ۵۱/۶ درصد قرار دارند. بر اساس این نتایج، پیش‌بینی کانون‌های برداشت گردوغبار در آینده با استفاده از مدل‌گرادیان تقویتی، به‌عنوان گزینه‌ی مناسبی برای پژوهش‌های بعدی مطرح است. نتایج حاصل از این مطالعه به سیاستمداران و مدیران برای شناسایی مناطق آسیب‌پذیر گردوغبار، اولویت‌بندی انجام فعالیت‌های مدیریتی برای کنترل گردوغبار، مدیریت صحیح اراضی مستعد کانون‌های برداشت گردوغبار و نیز اتخاذ تصمیم‌های مناسب برای کاهش خطر گردوغبار و پیامدهای منفی آن کمک خواهد کرد. به‌علاوه، مدل‌های داده‌کاوی روش‌های داده‌محور مفید و قدرتمندی هستند که قابلیت توسعه در تحقیقات محیط زیستی آینده را دارا می‌باشند.

منابع

- ۱- بنی‌حبیب، محمدابراهیم، شبستری، محمدهادی، حسین‌زاده، مرضیه (۱۳۹۵). تحقیقات منابع آب ایران، مدل هیبریدی برای مدیریت استراتژیک تقاضای آب کشاورزی مناطق خشک. تحقیقات منابع آب ایران، ۱۲ (۴)، ۶۹-۶۰.
- ۲- پورهایمی، سیمیا، بروغنی، مهدی، امیراحمدی، ابوالقاسم، زنگنه اسدی، محمدعلی، صالحی، مهدی (۱۳۹۸). اولویت‌بندی مناطق برداشت گردوغبار با استفاده از مدل‌های آماری (مطالعه‌ی موردی: خراسان رضوی). مرتع و آبخیزداری، ۷۲ (۲)، ۳۴۳-۳۵۸.
- ۳- کرمانی، مجید، طاهریان، الهام، ایزانلو، مریم (۱۳۹۵). تحلیل تصاویر ماهواره‌ای ریزگردها و طوفان‌های گردوغباری در ایران به منظور بررسی منشأهای داخلی و خارجی و روش‌های کنترل آن‌ها. ره‌آورد سلامت، ۲ (۱)، ۱۳-۱.
- ۴- مهرابی، شهباز، جعفری، رضا، سلطانی کویانی، سعید (۱۳۹۴). بررسی شاخص NDDI در پهنه‌بندی طوفان گردوغبار (مطالعه‌ی موردی استان خوزستان). مهندسی اکوسیستم بیابان، ۴ (۸)، ۱۰-۱.
- 5- Atkinson, P, Jiskoot, H, Massari, R, Murray, T. (1998). Generalized linear modelling in geomorphology 1195, 1185-1195.
- 6- Baddock, M. C, Bullard, J. E, & Bryant, R. G. (2009). Dust source identification using MODIS: a comparison of techniques applied to the Lake Eyre Basin, Australia. Remote Sensing of Environment, 113 (7), 1511-1528.
- 7- Bani Habib, M. A, Shabestari, M. H, Hosseinzadeh, M. (2016). Hybrid model for strategic management of agricultural water demand in arid regions. Iranian Water Resources Research. 12 (4): 60-69
- 8- Boroughani, M, Pourhashemi, S, Hashemi, H, Salehi, M, Amirahmadi, A, Asadi, M. A. Z, & Berndtsson, R. (2020). Application of remote sensing techniques and machine learning algorithms in dust source detection and dust source susceptibility mapping. Ecological Informatics, 56, 101059
- 9- Breiman L. (2001). Random forests. Machine Learning, 45 (1):5-32.
- 10- Du, J, J. Fang, W. Xu and P. Shi. (2013). Analysis of dry/wet conditions using the standardized precipitation index and its potential usefulness for drought/flood monitoring in Hunan Province, China. Stochastic Environmental Research and Risk Assessment, 27: 377-387.

- 11- Ebrahimi-Khusfi, Z, Taghizadeh-Mehrjardi, R, & Mirakbari, M. (2021). Evaluation of machine learning models for predicting the temporal variations of dust storm index in arid regions of Iran. *Atmospheric Pollution Research*, 12 (1), 134-147.
- 12- Gholami, H, Mohamadifar, A, Sorooshian, A, & Jansen, J. D. (2020) a. Machine-learning algorithms for predicting land susceptibility to dust emissions: The case of the Jazmurian Basin, Iran. *Atmospheric Pollution Research*, 11(8), 1303-1315.
- 13- Gholami, H, Mohamadifar, A, & Collins, A. L. (2020)b. Spatial mapping of the provenance of storm dust: application of data mining and ensemble modelling. *Atmospheric Research*, 233, 104716.
- 14- Gholami, H, Mohamadifar, A, Rahimi, S, Kaskaoutis, D. G, & Collins, A. L. (2021). Predicting land susceptibility to atmospheric dust emissions in central Iran by combining integrated data mining and a regional climate model. *Atmospheric Pollution Research*, 12 (4), 172-187.
- 15- Gholami, H, Mohammadifar, A, Golzari, S, Kaskaoutis, D. G, & Collins, A. L. (2021) b. Using the Boruta algorithm and deep learning models for mapping land susceptibility to atmospheric dust emissions in Iran. *Aeolian Research*, 50, 100682
- 16- Goudie, A. (2018). Dust storms and ephemeral lakes. *Desert* 23:153–164
- 17- Heydarian, P, Ajdari, A, Judaki, M, Darvishikhatoni, J, Shahbazi, R. (2017). Identification of internal sources of dust storms using remote sensing and geology (Case study: Khuzestan province). 27 (105): 33-46
- 18- Kermani, M, Taherian, A, Izanloo, M. (2016). Analysis of satellite images of dust and dust storms in Iran in order to investigate internal and external sources and their control methods. *Health guide*. 2 (1): 39-51
- 19- Lee, J. A, Baddock, M. C, Mbuh, M. J, Gill, T. E. (2012). Geomorphic and land cover characteristics of aeolian dust sources in West Texas and eastern New Mexico, USA. *Aeolian Research*. 3:459–466
- 20- Mei, D, Xiushan, L, Lin, S, & Ping, W. A. N. G. (2008). A dust-storm process dynamic monitoring with multi-temporal MODIS data. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 37, 965-970
- 21- Mehrabi, Sh, Jafari, R, Soltani Kopai, S. (2014). Investigation of NDDI index efficiency in dust storm zoning. *Desert Ecosystem Engineering*. 4 (8): 1-10.
- 22- Mirchooli, F, Motevalli, A, Pourghasemi, H. R, Mohammadi, M, Bhattacharya, P, Maghsood, F. F, & Tiefenbacher, J. P. (2019). How do data-mining models consider arsenic contamination in sediments and variables importance? *Environmental monitoring and assessment*, 191 (12), 1-19.
- 23- Marmion, M, Hjort, J, Thuiller, W, & Luoto, M. (2009). Computers & geosciences statistical consensus methods for improving predictive geomorphology maps 35, 615–625.
- 24- Pourghasemi, H. R, & Kerle, N. (2016). Random forests and evidential belief function-based landslide susceptibility assessment in Western Mazandaran Province, Iran. *Environmental earth sciences*, 75 (3), 185.
- 25- Rahmati, O, Mohammadi, F, Ghiasi, S. S, Tiefenbacher, J, Moghaddam, D. D, Coulon, F, & Bui, D. T. (2020). Identifying sources of dust aerosol using a new framework based on remote sensing and modelling. *Science of the Total Environment*, 737, 139508.
- 26- Sissakian, V, Al-Ansari, N, & Knutsson, S. (2013). Sand and dust storm events in Iraq. *Journal of Natural Science*, 5(10), 1084-1094.
- 27- Hahnenberger, M, & Nicoll, K. (2014). Geomorphic and land cover identification of dust sources in the eastern Great Basin of Utah, USA. *Geomorphology*, 204, 657-672.

Dust source mapping using satellite imagery and machine learning models

Mahdi Boroughani *¹

Assistant Professor of Research Center for Geoscience and Social Studies, Hakim Sabzevari University, Sabzevar, Iran

Fahimeh Mirchooli

Postdoctoral student of Geographical Sciences and Social Studies Research Center, Hakim Sabzevari University, Sabzevar, Iran

Maziar Mohammadi

Department of Watershed Management, Faculty of Natural Resources and Marine Sciences, Tarbiat Modares University, Noor, Iran

Abstract

Predicting dust sources area and determining the affecting factors is necessary in order to prioritize management and practice deal with desertification due to wind erosion in arid areas. Therefore, this study aimed to evaluate the application of three machine learning models (including generalized linear model, artificial neural network, random forest) to predict the vulnerability of dust centers during the years 2005 to 2018 in the Central Desert of Iran. For this purpose, the dust source areas were extracted in the study area using MODIS satellite images using four indicators including BTDD3132, BTDD2931, NDDI and variable D, and finally 135 hotspots were identified and used in modeling. In this study, conditional factors affecting dust were considered for modeling including land use, soil science, geology, distance from waterway, normalized vegetation difference index (NDVI), land slope and climate. The results showed that among the applied algorithms, random forest with 63.5% accuracy was the most accurate model and followed by artificial neural network with 43.4% accuracy and generalized linear model with 43.2% accuracy. In addition, among factors, land use and soil were identified as the most effective factors on dust source area. The results of this study can provide valuable information for regional managers and policy makers and help them to make useful decisions in management.

KeyWords: Remote Sensing, Dust Sources, Central Desert, Vulnerability.

1- *Email: m.boroughani@hsu.ac.ir