

Land Suitability Evaluation Using Traditional and Machine Learning Approaches: A Case Study in Abiek Plain, Qazvin Province, Iran

ABSTRACT

Land suitability is a crucial factor in land use planning and sustainable agricultural production. Evaluating land suitability helps optimize land use, promote sustainable land use, protect the environment, and ensure optimal use of natural resources. This study was conducted in the Abiek region of Qazvin province in northwest Iran, covering an area of 60,000 hectares. After collecting data from 300 soil profiles and determining land suitability classes for wheat cultivation with surface irrigation using the FAO classification system, digital elevation models, Landsat-8 and Sentinel-2 satellite images, and environmental variables extracted from the digital elevation model were used to create digital maps using both traditional and machine learning methods. The results showed that the machine learning method had a higher accuracy rate of 74% and a Kappa index of 68 compared to the traditional method with an accuracy rate of 62% and a Kappa index of 53. The most important environmental variables used in the machine learning model were those extracted from the digital elevation model and Landsat-8 satellite images. The largest area for wheat cultivation with surface irrigation was found in the relatively suitable class (S2), with 30,753 hectares in the random forest method and 21,028 hectares in the traditional method. In contrast, the smallest area belongs to the unsuitable class (N), with 3,052 hectares in the forest method. Additionally, random fields and 7185 hectares were identified in the traditional method. Also, 15,000 hectares of the study area are suitable for wheat cultivation without restrictions.

Keywords: *Geomorphological characteristics, parametric method, Random Forests, Wheat.*

ارزیابی تناسب ارضی با استفاده از رویکردهای سنتی و مدل‌های یادگیری ماشینی

(مطالعه موردی: دشت آبیگ، استان قزوین)

چکیده

تناسب ارضی یک عامل اساسی در برنامه‌ریزی استفاده از اراضی و تولید پایدار محصولات کشاورزی است. ارزیابی تناسب ارضی به بهینه‌سازی استفاده از اراضی، ترویج استفاده پایدار از اراضی، حفاظت از محیط‌زیست و اطمینان از استفاده بهینه از منابع طبیعی کمک می‌کند. این تحقیق در منطقه آبیگ استان قزوین واقع در شمال غرب ایران به وسعت ۶۰ هزار هکتار انجام شده است، پس از جمع‌آوری داده‌ها از ۳۰۰ خاک‌رخ و تعیین کلاس‌های تناسب زمین برای گندم با آبیاری سطحی با استفاده از سامانه طبقه‌بندی فائو، نقشه‌های رقومی به دو روش مرسوم و یادگیری ماشینی با استفاده از متغیرهای محیطی مستخرج از مدل رقومی ارتفاع، تصاویر ماهواره لندست-۸ و سنتینل-۲ بدست آمد. نتایج نشان داد که روش یادگیری ماشینی با دقت کلی ۷۴ درصد و شاخص کاپای ۶۸ توانست دقت بالاتری را نسبت به روش مرسوم با دقت کلی ۶۲ درصد و شاخص کاپای ۵۳ از خود نشان دهد. همچنین مهم‌ترین متغیرهای محیطی که در مدل‌سازی یادگیری ماشینی استفاده شدند متغیرهای مستخرج از مدل رقومی ارتفاع و ماهواره لندست-۸ بود. بیشترین وسعت منطقه برای کشت گندم با آبیاری سطحی در کلاس نسبتاً مناسب (S2) با ۳۰۷۵۳ هکتار در روش جنگل‌های تصادفی و ۲۱۰۲۸ هکتار در روش سنتی بدست آمد و کمترین وسعت نیز متعلق به کلاس نامناسب (N) با ۳۰۵۲ هکتار در روش جنگل‌های تصادفی و ۷۱۸۵ هکتار در روش سنتی شناسایی شد. ۱۵۰۰۰ هکتار از منطقه مورد مطالعه نیز بدون محدودیت (S1) کشت برای گندم با آبیاری سطحی گزارش گردید.

کلیدواژه: جنگل تصادفی، خصوصیات ژئومرفولوژیک، روش پارامتریک، گندم.

۱. مقدمه

در سال‌های اخیر افزایش جمعیت جهان یک نگرانی بزرگ به شمار می‌آید و انتظار می‌رود در قرن حاضر افزایش زیادی را تجربه کند (Kilic et al., 2022). برخی از چالش‌های مرتبط با این افزایش جمعیت مانند تقاضا برای دستیابی به غذای سالم، موجب شده است در سال‌های اخیر بیش از نیمی از جمعیت جهان با فقر غذایی دست و پنجه نرم کنند (Dorling, 2021). مجموعه عواملی همچون افزایش قیمت مواد غذایی، کمبود منابع غذایی، عوامل اقتصادی، پاندمی کووید ۱۹ و تغییرات اقلیمی باعث بروز بحران در نقاط مختلف جهان در یک دهه اخیر شده است (Alexandratos & Bruinsma, 2012). به طوری که تعداد افرادی که توانایی تهیه یک وعده غذایی سالم را ندارند به ۳ میلیارد نفر افزایش یافته است (IaW, 2015).

برای مقابله با این چالش‌ها، جهان نیازمند ایجاد سازمان‌ها و تدوین برنامه‌های مختلف برای تضمین امنیت غذایی و کاهش ضایعات مواد غذایی است؛ بنابراین، تا زمان شکل‌گیری یک روند مناسب و یکپارچه در زمینه تولید غذای سالم و کافی، آمار مرگ‌ومیر ناشی از سوءتغذیه، روندی روبه‌افزایش به‌ویژه در میان کودکان خواهد داشت (IaW, 2015). یکی از برنامه‌های مهم در این خصوص افزایش تولید غذا بدون افزایش سطح زیر کشت است (Onyutha, 2019)، که بخشی از آن نیازمند ارزیابی تناسب ارضی و آگاهی از موانع اصلی جلوگیری از عملکرد حداکثر محصول در هر منطقه و همچنین مکان‌یابی نقاط بهینه جهت کشت پایدار محصول می‌باشد (Mugiyo et al., 2021).

ارزیابی تناسب ارضی را می‌توان به‌عنوان یک گام اساسی در برنامه‌ریزی استفاده بهینه و پایدار از اراضی، به‌ویژه در برنامه‌ریزی تناوب کشت ارضی کشاورزی دانست (AbdelRahman et al., 2022; Akıncı et al., 2013). به عبارت دیگر از آنجایی که ارزیابی ارضی برنامه‌ریزی منطقی کاربری ارضی و استفاده مناسب و پایدار از منابع طبیعی را ترویج می‌کند، یکی از جنبه‌های مهم افزایش تولید محصول بدون افزایش سطح زیر کشت و تخریب منابع است (Hagos et al., 2022). در این فرآیند، خاک، اقلیم، پوشش گیاهی و سایر ویژگی‌های ارضی از نظر تناسب با محصول مورد نظر ارزیابی می‌شوند (Kilic et al., 2022). ارزیابی تناسب

اراضی از جمله روشهایی است که بیش از یک معیار را برای رسیدن به نتیجه بررسی می کند، از این رو روش مذکور جزو روشهای تصمیم گیری چند معیاره دسته بندی می شود (Kılıc et al., 2022; Prakash, 2003).

روش های زیادی از جمله روش FAO، برای ارزیابی تناسب اراضی وجود دارند. از زمان ارائه دستورالعمل فائو (FAO, n.d.; Food and Agriculture Organization of the United Nations. Soil Resources and Conservation Service, 1985) به طور گسترده ای برای تعیین تناسب اراضی استفاده شده است (Al-Mashreki et al., 2011; Ashraf & Normohammadan, 2015; Bagheri Bodaghabadi et al., 2015). این روش ها بر پایه مطابقت خصوصیات اراضی و احتیاجات گیاه بنا نهاده شده است. روش های ارزیابی تناسب اراضی (Sys et al., 1991) توانایی و محدودیت های اراضی برای یک کاربری خاص را تعیین می کنند. معمولاً در تهیه نقشه های تناسب اراضی در رویکرد سنتی، تغییرات تناسب اراضی از طریق ارتباط یک واحد نقشه خاک یا واحد اراضی به یک خاکرک خاک شاهد نمایش داده می شد. این امر باعث کاهش دقت نقشه به دلیل نادیده گرفتن ماهیت پیوسته تغییرات خاک و منظر می شد (Daigle et al., 2005; Ziadat, 2000). به عبارت دیگر ویژگی های خاک دارای تغییرات مکانی و زمانی هستند؛ انتظار می رود که این تغییرات، در واحدهای نقشه ی خاک و نقشه های تناسب اراضی نیز وجود داشته باشند. بنابراین، به نظر می رسد که یکی از مهم ترین مشکلات روش های سنتی نقشه برداری خاک، تعمیم نتایج حاصل از خاک رخ شاهد به کل واحد نقشه، بدون در نظر گرفتن تغییرات مکانی آن هست (Safari et al., 2013).

امروزه با پیشرفت در زمینه فناوری، تمامی علوم قادرند تا تقاضاهای جدید زندگی امروزه بشر را پاسخگو باشند. ارزیابی تناسب اراضی نیز باتوجه به اهمیت روزافزون آن از این قاعده مستثنا نیست و نیازمند به کارگیری روش های مختلف برای درک بهتر تغییرپذیری کلاس های تناسب اراضی به منظور مدیریت پایدار اراضی هست. به نظر می رسد روش های یادگیری ماشین^۱ را می توان برای بهبود تحلیل تناسب اراضی مورد استفاده قرار داد (Roell et al., 2020). روش های داده کاوی^۲ و تهیه نقشه های رقومی، تلاش می کنند تا بر اساس متغیرهای محیطی که به سادگی قابل دستیابی هستند؛ ویژگی مدنظر کلاس ها، ویژگی های خاک و کلاس های تناسب اراضی را با دقت بالا پیش بینی کند (Behrens & Scholten, 2006; Neyestani et al., 2021; Wylie et al., 2019). در این زمینه، مدل های مختلفی مانند شبکه های عصبی مصنوعی، رگرسیون درختی توسعه یافته، درختان تصمیم گیری تصادفی و رگرسیون لاجیستیک چندجمله ای استفاده می شوند (Brungard et al., 2015; Minasny & McBratney, 2016; Roell et al., 2020). یکی از موضوعات مهم در یادگیری ماشین، مدل مورد استفاده برای ارتباط دادن پارامترهای محیطی با ویژگی مورد بررسی می باشد همچنین کاربرد موفقیت آمیز تهیه نقشه های رقومی به میزان داده های موجود از منطقه ی مطالعاتی و پارامترهای محیطی استفاده شده در مدل بستگی دارد (Stoorvogel et al., 2009).

در دهه اخیر، الگوریتم های مبتنی بر یادگیری ماشین به طور گسترده برای تهیه نقشه های رقومی کلاس و خصوصیات خاک مورد استفاده قرار گرفته اند (C. Gu et al., 2019; Khaledian & Miller, 2020; Lagacherie et al., 2020; Liu et al., 2023; Takoutsing & Heuvelink, 2022; Zhang et al., 2020). با این حال تاکنون مطالعات کمی برای استفاده از مدل های یادگیری ماشین در نقشه کلاس های تناسب اراضی و مقایسه آنها با نقشه های سنتی انجام شده است (Kidd et al., 2015; Taghizadeh- (Mehrijardi et al., 2020; Vasu et al., 2018). در مطالعه حاضر دو هدف مد نظر است: (۱) ارزیابی تناسب اراضی برای کشت گندم تحت آبیاری سطحی باتوجه به اینکه گندم یکی از مهم ترین محصولات غذایی برای جمعیت بشری است و در سراسر جهان مصرف می شود (Igrejas & Branlard, 2020). (۲) بررسی و مقایسه دقت نقشه های تناسب به دست آمده از روش های سنتی و یادگیری ماشین در بخشی از دشت قزوین که یکی از دشت های خشک و نیمه خشک ایران است.

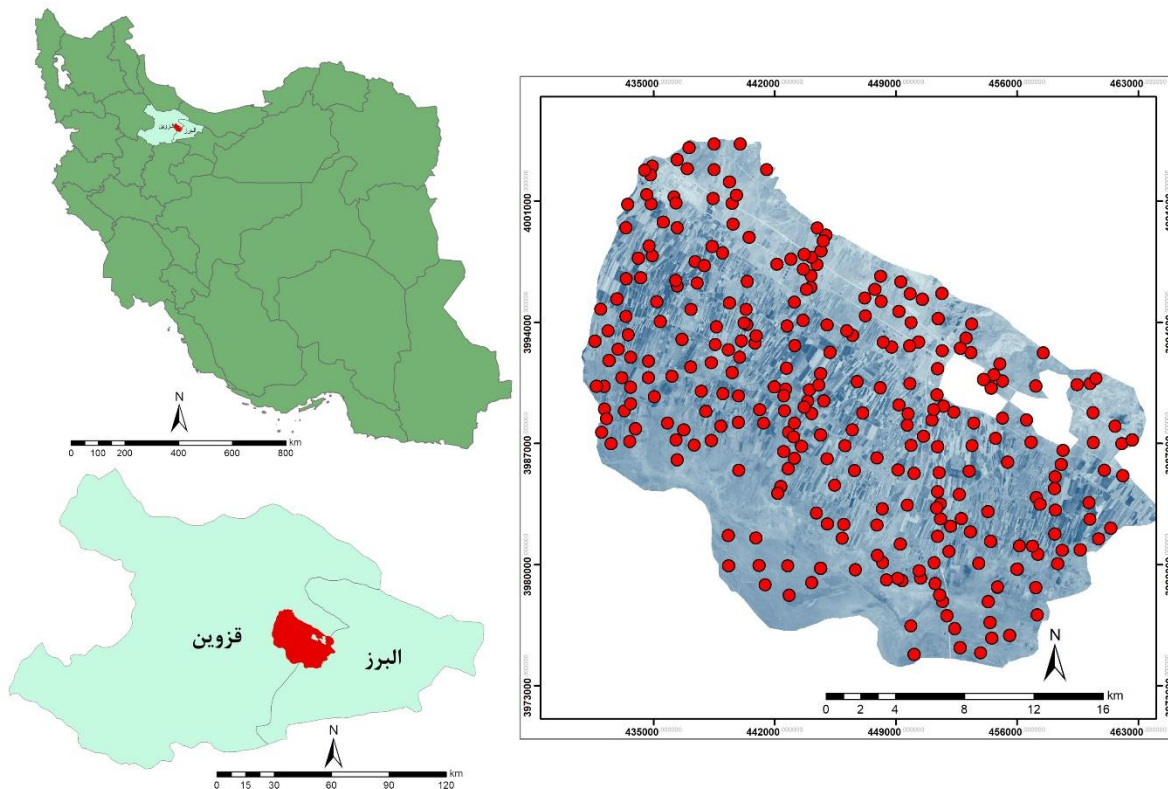
۲. مواد و روش ها

¹ Machine learning.

² Data mining

۲-۱. منطقه مورد مطالعه

منطقه مورد مطالعه در شمال غربی ایران بین ۳۶ درجه و ۱۸ دقیقه تا ۳۷ درجه عرض شمالی و ۵۰ درجه و ۳۴ دقیقه تا ۵۰ درجه و ۵۲ دقیقه طول شرقی در آبیک، استان قزوین قرار دارد (شکل ۱). میانگین بارندگی و دما سالانه بر اساس داده های هواشناسی ۲۲ ساله (سازمان هواشناسی ایران (IRIMO)) به ترتیب ۲۵۷ میلی متر و ۱۴/۳ درجه سانتی گراد است. بر این اساس، رژیم رطوبتی و دمایی خاک به ترتیب اریدیک ضعیف، زریک خشک و ترمیک بود. خاکهای غالب عبارتند از Entisols، Inceptisols و Aridisols. این منطقه مساحتی به وسعت ۶۰۰ کیلومتر مربع با ارتفاع بین ۱۱۳۶ تا ۱۸۱۵ از سطح دریا را پوشش می دهد (Khamoshi et al. 2023). انواع کاربری اراضی منطقه مورد مطالعه شامل کشاورزی دیم، کشاورزی آبی و مرتع ضعیف است. بیش از ۸۰ درصد مساحت در کلاس شیب ۵-۰ درصد قرار دارد و جهت اصلی شیب منطقه، جنوبی که شامل ۲۷/۱۵ درصد از مساحت



شکل ۱: موقعیت منطقه مورد مطالعه و موقعیت مکانی خاکرخ ها

کل منطقه مطالعاتی است.

۲-۲. جمع آوری داده ها و تعیین کلاس تناسب اراضی

در مجموع ۳۰۰ خاکرخ موجود در بانک اطلاعاتی آزمایشگاه سنجش از دور گروه علوم و مهندسی خاک دانشگاه تهران به منظور انجام مطالعات در سطح اجمالی (نشریه ۴۲۵ معاونت برنامه ریزی و نظارت راهبردی ریاست جمهوری) انتخاب شدند. پس از تشریح خاکرخ های موجود از روش طبقه بندی تناسب اراضی به روش فائو (۱۹۷۶) استفاده شد. ارزیابی تناسب اراضی در غالب سه مرحله اجرا گردید: مرحله اول تحقیق جمع آوری اطلاعات لازم درباره خصوصیات اراضی منطقه مورد مطالعه که این اطلاعات برخی از طریق تشریح خاکرخ و بخشی هم از طریق اطلاعات اقلیمی جمع آوری شد. مرحله دوم شامل تعیین نیازهای رویشی انواع تیپ های بهره برداری از اراضی می باشد که در این نیازهای رویشی تیپ های بهره برداری مورد نظر از لحاظ اقلیمی و خصوصیات خاک بررسی شد. برای این منظور از جلد سوم کتاب تناسب اراضی (Sys et al., 1991) استفاده شد و احتیاجات کاربری های

موردنظر استخراج گردید. در مرحله سوم پس از انتخاب پارامترها، برای تعیین شاخص تناسب اراضی گندم روش پارامتریک (روش ریشه دوم) استفاده شد. سپس رتبه بندی اقلیمی، خاک و توپوگرافی (جدول ۱) برای محاسبه شاخص تناسب نهایی زمین اعمال شد.

جدول ۱: خصوصیات اقلیمی، خاک و توپوگرافی مورد استفاده در روش پارامتریک

سولفات کلسیم	کربنات کلسیم	درصد سنگ ریزه	بافت خاک	خصوصیات فیزیکی
			میانگین دمای طول دوره رشد	خصوصیات اقلیمی
			درصد شیب	خصوصیات توپوگرافی
		نسبت جذب سدیم	هدایت الکتریکی	خصوصیات شوری و قلیائیت

برای محاسبه شاخص تناسب اراضی از روش پارامتریک استفاده شد و در نهایت کلاس تناسب اراضی با استفاده از روش ریشه دوم تعیین گردید (معادله ۱). در روش پارامتریک به هریک از سطوح محدودیت درجه بندی در مقیاس ۱۰۰-۰ نسبت داده می شود و در نهایت شاخص نهایی اراضی از روش ریشه دوم به دست می آمد و سپس با توجه به شاخص اراضی، کلاس تناسب اراضی برای هریک از خاکرخ ها تعیین گردید (Sys et al. 1991) (جدول ۲).

جدول ۲: طبقه بندی شاخص تناسب اراضی (Sys et al. 1991)

کلاس تناسب		شاخص اراضی
تعاریف	علائم	
خیلی مناسب	S1	۱۰۰-۷۵
نسبتاً مناسب	S2	۷۵-۵۰
تناسب کم	S3	۵۰-۲۵
نامناسب	N	۲۵-۰

$$I = R \min \sqrt{\frac{A}{100} \times \frac{B}{100} \times \frac{C}{100} \times \dots}$$

معادله ۱

۲-۳. تهیه نقشه به روش سنتی

در این روش ابتدا نقشه واحد اراضی بر اساس استخراج ژئو فرم از طریق نقشه های ارتفاع، درصد شیب، جهت شیب (شکل ۲) (Rossiter, 2000) تهیه گردید، سپس پلی گون های زیر یک هکتار به دلیل بهبود ماهیت بصری نقشه ها با یکدیگر ادغام گردید (شکل ۲) و پس از تعیین کلاس تناسب اراضی برای خاکرخ های شاهد در هر کدام از واحد های اراضی به دست آمده، نقشه تناسب اراضی توسط نرم افزار Arc GIS استخراج گردید.

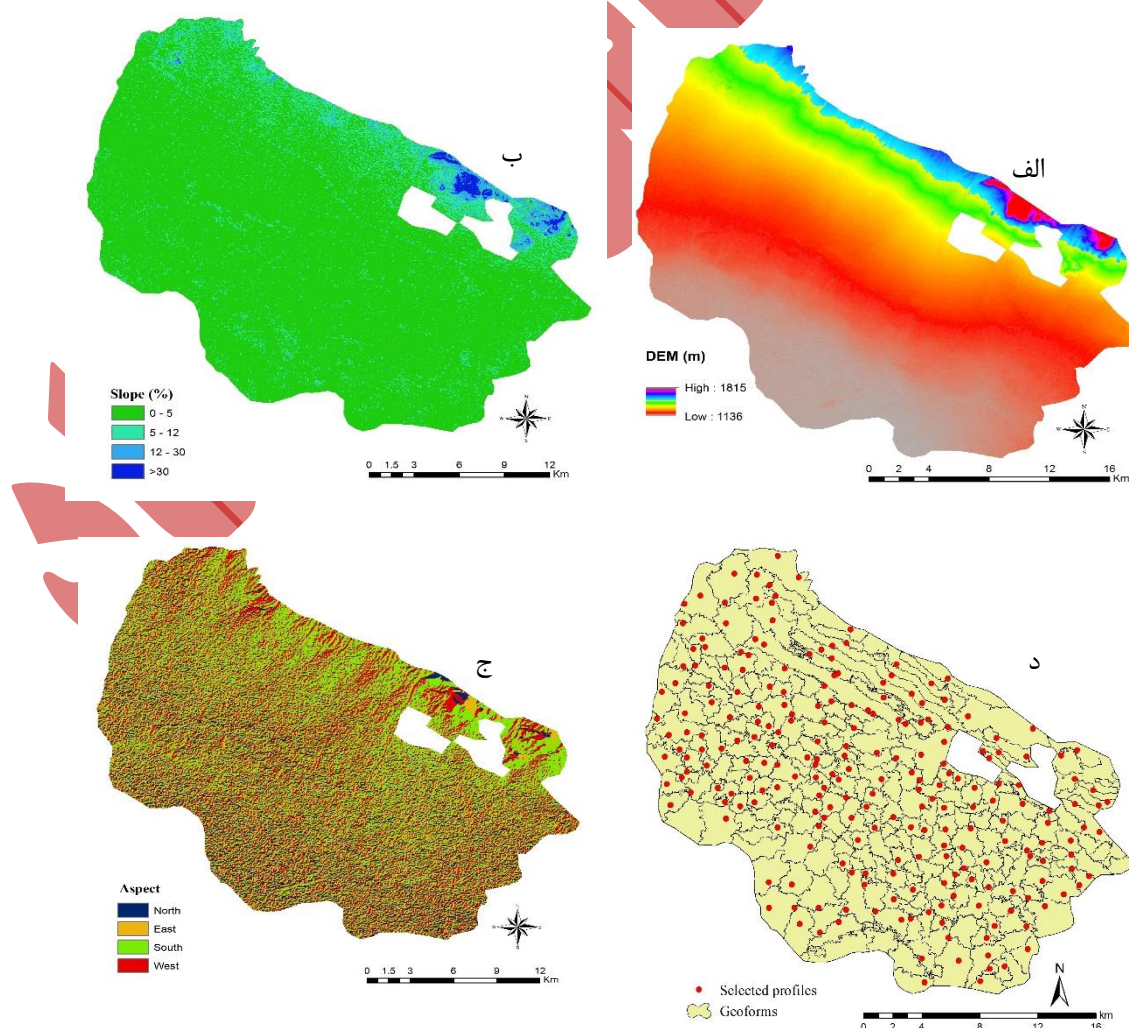
۲-۴. متغیرهای محیطی

در مدل سازی، از متغیرهای کمکی مربوط به معادله SCORPAN یعنی خاک، آب و هوا، موجودات زنده، توپوگرافی، مواد مادری و موقعیت مکانی استفاده شد (McBratney et al., 2003). بر این اساس، از مدل رقومی ارتفاع (DEM) با تفکیک مکانی ۱۲/۵ متر استفاده شد. در مجموع، ۱۱ متغیر ژئومورفومتریک به عنوان مشتقات اولیه و ثانویه DEM در نرم افزار SAGA GIS 7.3 تهیه شد. لایه های آماده شده در جدول ۳ گزارش شده است. علاوه بر این، ۲۵ شاخص طیفی از تصاویر روزهای ۲۷ و ۳۱ ماه آگوست

سال ۲۰۲۰ ماهواره های لندست ۸ و سنتینل ۲ با استفاده از نرم افزار TerrSet 2020 تهیه شد (جدول ۳). در نهایت، تمامی متغیرهای محیطی اندازه پیکسل های ۱۰ متری تبدیل شدند. سپس مناسبترین شاخصهای محیطی برای مدل سازی بر اساس همبستگی پیرسون انتخاب شدند.

۲-۵. تهیه نقشه به روش یادگیری ماشین

در تهیه نقشه ها از مدل جنگل های تصادفی^۳ (Breiman, 2001) که یکی از الگوریتم های پرکاربرد یادگیری ماشین محسوب می شود استفاده گردید (Mahmoudzadeh et al., 2020). برای اینکه مدل جنگل تصادفی یک مدل پیش بینی تولید کند، باید دو پارامتر مشخص شود. تعداد درختان رگرسیون در جنگل ها (ntree) و تعداد ویژگی های تصادفی انتخاب شده در هر گره (mtry) از جمله این عوامل هستند. جذر تعداد کل پیش بینی کننده ها در مدل به عنوان مقدار پیش فرض برای اندازه زیر مجموعه تصادفی عمل می کند. با توجه به خطاهای پیش بینی رگرسیون خارج از کیسه، که در نمونه بوت نمایش داده نمی شوند، اهمیت متغیرها تعیین می شود (Wang et al., 2018). در این مطالعه ابتدا کلاس تناسب اراضی برای هر یک از خاکرخی های مطالعاتی براساس بخش ۲-۲ تعیین گردید سپس برای تعیین تعداد ایده آل mtry از اعتبار سنجی متقابل K-fold استفاده شد. از روش نمونه گیری بوت استرپ در انتخاب ۷۰ درصد از نمونه های خاک (n=210) برای داده های آموزشی و ۳۰ درصد برای داده های اعتبارسنجی مدل (n=90) در نرم افزار R نسخه ۴,۱,۰ استفاده شد.



³ Random Forests

شکل ۲: الف) مدل رقومی ارتفاع، ب) درصد شیب، ج) جهت شیب و د) نقشه ژئوform و خاکرخی های شاهد

جدول ۳: متغیرهای محیطی استفاده شده

ردیف	نام	علامت اختصاری در این تحقیق
متغیرهای مستخرج از مدل رقومی ارتفاع		
۱	مدل رقومی ارتفاع	dem
۲	سطح پایه شبکه کانال	Channel_Network_Bsae_Level cnbl
۳	شاخص توپوگرافی خیزی	Topographic_Wetness_Index (twi)
۴	شاخص عمق دره	Valley_Depth
۵	شاخص صافی پایین دره با وضوح چندگانه	MRVBF
۶	شاخص طول شیب	LS_Factor
۷	جهت شیب	Aspect
۸	شاخص همگرایی	Convergence_Index
۹	شاخص موقعیت نسبی شیب	RSP
۱۰	فاصله عمودی تا شبکه کانال	VDCN
۱۱	شیب	Slope
متغیرهای مستخرج از تصاویر ماهواره		
۱۲	باند آبی لندست ۸ ($0.482 \mu\text{m}$)	b02l
۱۳	باند سبز لندست ۸ ($0.561 \mu\text{m}$)	b03l
۱۴	باند قرمز لندست ۸ ($0.654 \mu\text{m}$)	B04L
۱۵	باند مادون قرمز نزدیک لندست ۸ ($0.864 \mu\text{m}$)	B05L
۱۶	موج کوتاه ۱ مادون قرمز لندست ۸ ($1.608 \mu\text{m}$)	B06L
۱۷	موج کوتاه ۲ مادون قرمز لندست ۸ ($2.200 \mu\text{m}$)	B07L
۱۸	باند آبی سنتینل ۲ ($0.490 \mu\text{m}$)	b2s
۱۹	باند سبز سنتینل ۲ ($0.560 \mu\text{m}$)	b3s
۲۰	باند قرمز سنتینل ۲ ($0.665 \mu\text{m}$)	b4s
۲۱	مرز قرمز گیاهی سنتینل ۲ ($0.705 \mu\text{m}$)	b5s
۲۲	مرز قرمز گیاهی سنتینل ۲ ($0.740 \mu\text{m}$)	b6s
۲۳	مرز قرمز گیاهی سنتینل ۲ ($0.783 \mu\text{m}$)	b7s
۲۴	باند مادون قرمز نزدیک سنتینل ۲ ($0.842 \mu\text{m}$)	b8s
۲۵	مرز قرمز گیاهی سنتینل ۲ ($0.865 \mu\text{m}$)	b8as
۲۶	موج کوتاه ۱ مادون قرمز سنتینل ۲ ($1.610 \mu\text{m}$)	b11s
۲۷	موج کوتاه ۲ مادون قرمز سنتینل ۲ ($2.190 \mu\text{m}$)	b12s
۲۸	شاخص پوشش گیاهی نرمال شده (لندست ۸)	ndvi.ls8
۲۹	شاخص پوشش گیاهی نرمال شده (سنتینل ۲)	ndvi_sen
۳۰	شاخص شوری نرمال شده (لندست ۸)	ndsi
۳۱	شاخص شوری نرمال شده (سنتینل ۲)	ndsi_sen
۳۲	شاخص شوری (لندست ۸)	si
۳۳	شاخص شوری (سنتینل ۲)	si_sen
۳۴	Tasseled Cab brightness (لندست ۸)	tasseled_cab_b
۳۵	Tasseled Cab greenness (لندست ۸)	tasseled_cab_g
۳۶	Tasseled Cab wetness (لندست ۸)	tasseled_cab_w

اعتبارسنجی

جهت اعتبارسنجی داده‌های حاصل نقشه سنتی و نقشه حاصل از مدل جنگل تصادفی از شاخص‌های دقت کلی (معادله ۲)، ضریب کاپا (معادله ۳) و خاکرخ‌های غیر شاهد (روش سنتی) بهره گرفته شد. دقت کلی نسبت تمام کلاس‌های تناسب اراضی که به درستی طبقه‌بندی شده به همه داده‌های استفاده شده است. هرچه این عدد بالاتر باشد نشان دهنده دقت بیشتر است.

$$OA = \frac{\text{number of correctly classified samples}}{\text{total number of samples}} \quad \text{معادله 2}$$

ضریب کاپا یکی از روش‌های آمار ناپارامتریک است که برای بررسی و ارزیابی درجه و میزان توافق بین مدل‌ها یا روش‌های مختلف استفاده می‌شود. مقدار ضریب کاپا از -۱ تا +۱ است. هرچه ضریب به دست آمده به +۱ نزدیکتر باشد نشان دهنده طبقه بندی درست و مناسب مدل استفاده شده است (Taghizadeh-Mehrjardi et al., 2020)

$$Kappa\ Index = \frac{\text{predicted land suitability class} - \text{expected land suitability class}}{1 - \text{expected land suitability class}} \quad \text{معادله 3}$$

۳. نتایج و بحث

۳-۱. تجزیه و تحلیل آماری داده‌های

جدول ۴ آمار توصیفی خواص خاک در منطقه مورد مطالعه را نشان می‌دهد. محدوده هدایت الکتریکی از ۰/۱۴ تا ۱۶۴ ds/m و میانه و میانگین آن به ترتیب ۰/۹۹ و ۵/۲۳ ds/m است. این اعداد نشان دهنده این است که هدایت الکتریکی در بیشتر منطقه مورد مطالعه مقادیر قابل قبولی برای رشد گندم دارد. میانگین سولفات کلسیم ۰/۸۶ درصد و دامنه آن بین ۰ تا ۲۸ درصد است. میانه سولفات کلسیم نشان می‌دهد که نیمی از خاکرخ‌های مورد مطالعه فاقد سولفات کلسیم هستند. محدوده کربنات کلسیم خاک از ۰/۴۷ درصد به ۵۱ درصد تغییر می‌کند و میانگین و میانه آن به ترتیب ۱۲/۳ و ۱۰/۹ درصد است که نشان دهنده توزیع نرمال این خصوصیت در منطقه می‌باشد. همچنین این خصوصیت با عمق روند افزایشی داشت. داده‌های درصد سنگریزه، نسبت جذب سدیم (SAR) و درصد شیب در منطقه مورد مطالعه به ترتیب از ۰ تا ۸۰ درصد، ۰ تا ۱۵۲ meq/l و ۰ تا ۱۸ درصد متغیر بود. و میانگین آنها به ترتیب ۹ درصد، ۷/۳ meq/l و ۳/۱۶ درصد است. بیشترین ضریب تغییرات در منطقه مورد مطالعه متعلق به سولفات کلسیم بود که نشان می‌دهند این خصوصیت دارای تنوع بیشتری نسبت به سایر خصوصیات است.

جدول ۴: آمار توصیفی خصوصیات خاک منطقه مورد مطالعه

واحد	کمترین	بیشترین	میانه	میانگین	انحراف از استاندارد	ضریب تغییرات
هدایت الکتریکی ds/m	۰/۱۴	۱۶۳/۷	۰/۹۹	۵/۲۳	۱۷	۳/۲۵
سولفات کلسیم %	۰	۲۸/۱۷	۰	۰/۸۶	۳/۷۷	۴/۴۱
کربنات کلسیم %	۰/۴۷	۵۱	۱۰/۸۸	۱۲/۲۹	۷/۵۴	۰/۶۱
سنگریزه %	۰	۸۰	۰	۹/۳۵	۱۶/۴۴	۱/۷۶
نسبت جذب سدیم meq/l	۰	۱۵۱/۸۶	۰	۷/۳۱	۲۱/۸۴	۲/۹۹
شیب %	۰	۱۸/۳	۲/۶۴	۳/۱۶	۲/۳	۰/۷۳

۲-۳. متغیرهای محیطی

جدول ۵ نتایج ضرایب همبستگی پیرسون بین شاخص تناسب اراضی و متغیرهای محیطی را نشان می‌دهد. شاخص تناسب اراضی برای گندم با آبیاری سطحی با دو متغیر وابسته به ارتفاع (cnbl، dem) همبستگی منفی معنی داری داشت. نتایج (Vasu et al., 2018) بیانگر نقش مهم متغیرهای مربوط به مدل رقومی ارتفاع در ارزیابی تناسب اراضی بود. همبستگی منفی این متغیرها با واقعیت موجود در منطقه همخوانی داشت به طوری که هرچه به مناطق کم ارتفاع تر منطقه مورد مطالعه برویم به دلیل وجود عوارض شوری و قلیابیت مخصوصاً در قسمت جنوبی کلاس تناسب اراضی کاهش پیدا می‌کند. این موضوع را وجود همبستگی منفی معنی دار بین شاخص شوری (si و si_sen) و کلاس تناسب اراضی نیز تایید می‌کند. از سوی دیگر تناسب اراضی با شاخص های ثانویه مستخرج از تصاویر ماهواره‌های لندست ۸ و سنتینل ۲ (NDVI) همبستگی مثبت و معناداری نشان داد که این نتایج با مطالعه (Taghizadeh-Mehrjardi et al., 2020) هم راستا بود. در حالی که شاخصی تناسب اراضی گندم با آبیاری سطحی با باندهای ماهواره لندست ۸ همبستگی منفی معنی داری داشت. به طور کلی، هرچه شاخص پوشش گیاهی مقادیر بالاتر و نزدیک به ۱ را نشان دهد بیانگر تراکم بالای پوشش گیاهی و به طبع آن فعالیت بیشتر ریشه و ریز موجودات و همچنین توسعه فعالیت های خاکسازی در آن نقاط است. به طبع این فعالیت مطلوب موجب تناسب بیشتر در اراضی می‌گردد. بنابراین انتظار روندی مثبت دور از انتظار نمی‌باشد. در نهایت از متغیرهای محیطی که همبستگی معنی داری با شاخص تناسب اراضی داشتند در مدل سازی استفاده شد.

جدول ۵: نتایج تحلیل همبستگی پیرسون بین شاخص تناسب اراضی و متغیرهای محیطی

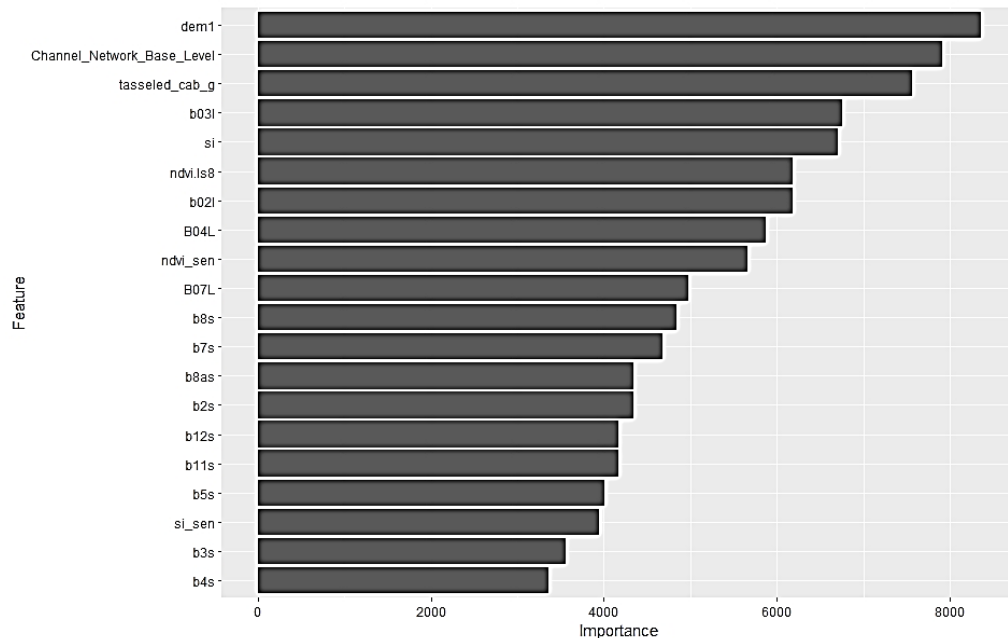
ضرایب همبستگی	
dem	-.۲۳*
cnbl	-.۲۲*
b02l	-.۳۳**
b03l	-.۲۶**
B04L	-.۱۹**
B07L	-.۲۰**
si	-.۲۷**
ndvi.ls8	.۲۳**
tasseled_cab_g	.۳۰**
b2s	-.۳۳**
b3s	-.۲۸**
b4s	-.۲۶**
b5s	-.۲۵**
b7s	.۱۴*
b8s	.۱۳*
b8as	.۱۷*
b11s	-.۱۷**
b12s	-.۲۶**
ndvi_sen	.۲۷**
si_sen	-.۲۸**

** همبستگی معنی دار در سطح ۰/۰۱ درصد

* همبستگی معنی دار در سطح ۰/۰۵ درصد

مطابق نتایج خروجی از مدل جنگل تصادفی متغیرهای تأثیرگذار در مدل سازی کلاس تناسب اراضی نشان داد که دو متغیر مستخرج از مدل رقومی ارتفاع (cnbl، dem) که دارای رابطه معنی دار با کلاس تناسب اراضی بودند با اهمیت ترین متغیرها در تعیین

تناسب اراضی برای گندم در منطقه مورد مطالعه شدند (شکل ۳). به بیان دیگر متغیرهایی که بر اساس خصوصیات ژئومرفولوژیک هستند تأثیر بیشتری در تعیین کلاس تناسب نسبت به متغیرهای با پایه تصاویر ماهواره‌ای داشتند که این نتایج یا یافته‌های (Taghizadeh-Mehrjardi et al., 2020; Zakarya et al., 2021) مشابه بود. همچنین متغیرهای حاصل از تصاویر ماهواره لندست ۸ (ndvi.ls8 و si ، b031، tasseled_cab_g) بیشتر از متغیرهای محیطی حاصل از تصاویر سنتینل ۲ در مدل‌سازی نقش داشتند.



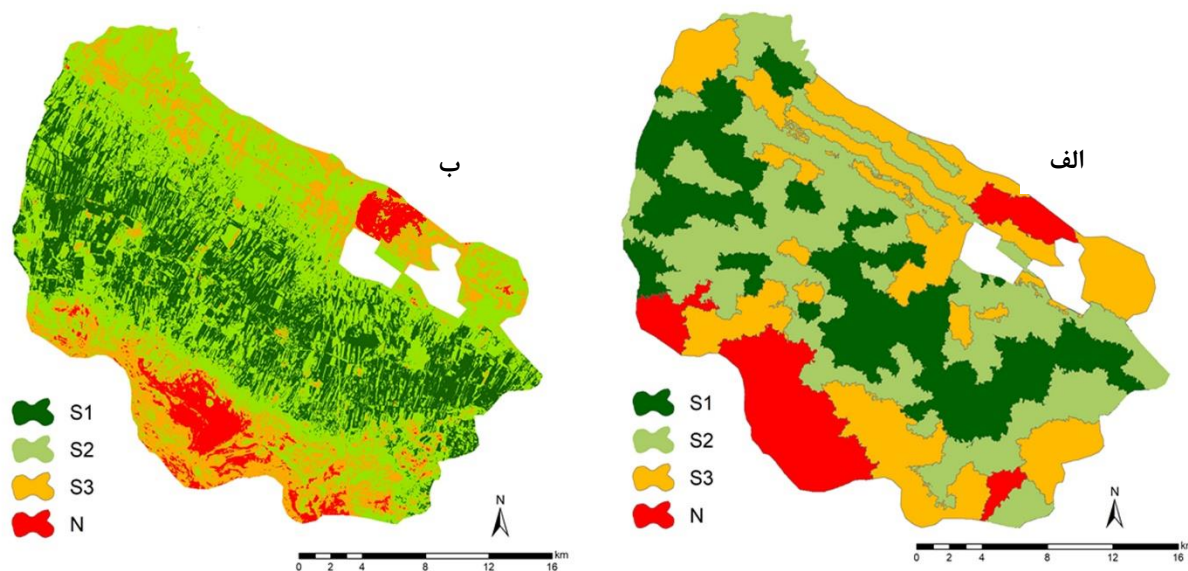
شکل ۳: نمودار اهمیت نسبی متغیرهای مورد استفاده در مدل‌سازی

۳-۳. کلاس تناسب اراضی

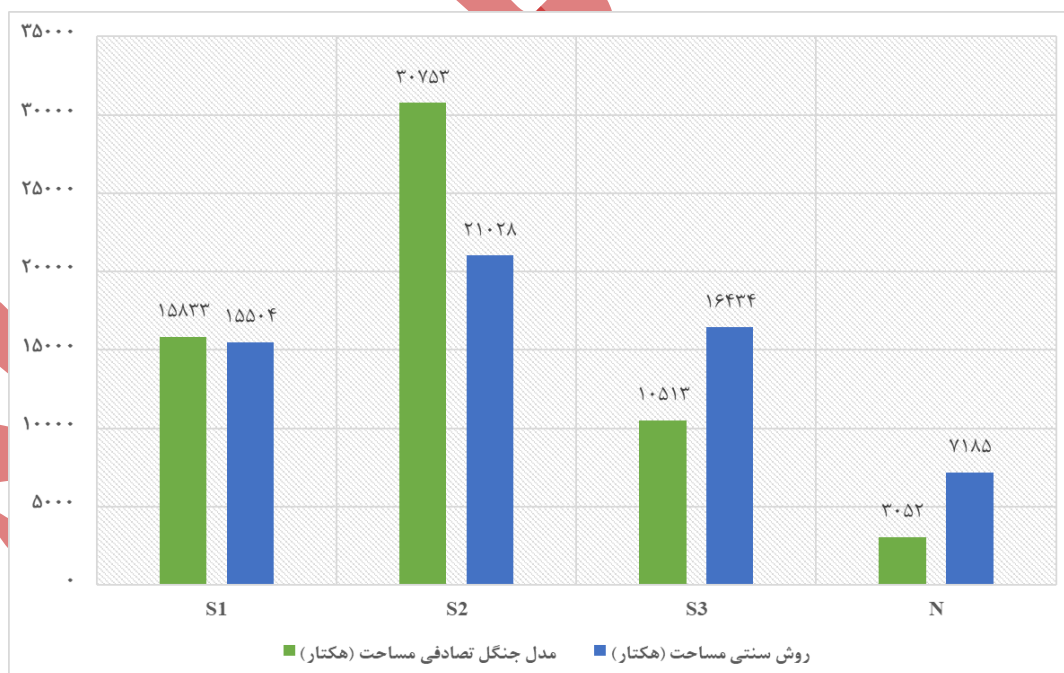
بررسی شرایط اقلیمی منطقه مورد مطالعه نشان داد که شاخص اقلیمی در کلاس تناسب کم (S3) برای کشت گندم قرار دارد. این محدودیت عمدتاً به خاطر کمبود بارندگی در دوره‌های گلدهی و رسیدگی گندم می‌باشد. در تحقیق حاضر به دلیل محاسبه ارزیابی تناسب برای گندم با آبیاری سطحی، شاخص بارندگی در محاسبات لحاظ نگردید بنابراین در بحث شاخص‌های اقلیمی از میانگین دمای طول دوره رشد استفاده شد که در کلاس بدون محدودیت (S1) قرار داشت.

در شکل ۴ نقشه‌های مستخرج از روش سنتی و مدل جنگل تصادفی آورده شده است. این نقشه‌ها نشان می‌دهند که کلاس تناسب اراضی بدون محدودیت (S1) برای کشت گندم با آبیاری سطحی در هر دو روش سنتی و یادگیری ماشین کمتر از ۳۰ درصد کل مساحت منطقه را به خود اختصاص دادند و به ترتیب برابر ۲۵/۷۷ و ۲۶/۳۲ درصد به دست آمدند. این کلاس اراضی بیشتر در مرکز منطقه مورد مطالعه و خط میانی غربی-شرقی مشاهده گردید که عمدتاً تحت تاثیر زراعت آبی قرار دارد. در روش سنتی ۳۴/۹۶ درصد از کل مساحت منطقه مورد نظر در کلاس نسبتاً مناسب (S2) قرار گرفت در حالی که در روش جنگل تصادفی این نسبت برابر با ۵۱/۱۲ درصد بود. این نتایج نشان داد که مدل جنگل تصادفی کلاس (S2) را بیشتر از روش سنتی برآورد کرده است. کلاس تناسب نسبتاً مناسب نیز بیشتر در مرکز منطقه مشاهده گردید که عمدتاً تحت تاثیر محدودیت سولفات و کربنات کلسیم در این مناطق بود. همچنین ۲۷/۳۳ درصد از منطقه در روش سنتی در کلاس تناسب کم (S3) قرار گرفت این درصد در روش جنگل تصادفی برابر با ۱۷/۴۸ بود که کمتر برآورد شده بود. کلاس تناسب کم را در منطقه مورد مطالعه به دو قسمت شمالی و جنوبی منطقه می‌توان تقسیم کرد. به طوری‌که مناطق شمالی بیشتر تحت تاثیر محدودیت‌های فیزیکی مانند محدودیت بافت و سنگریزه و عامل شیب بود. در حالی که در مناطق جنوبی کلاس تناسب کم مربوط به محدودیت‌های شوری و قلیائیت بود.

۱۱/۹۴ درصد از منطقه نیز در روش سنتی دارای کلاس نامناسب (N) بود که این مقدار در روش جنگلهای تصادفی برابر ۵/۰۷ درصد گزارش شد (شکل ۴). در شکل ۵ درصد هریک از کلاس های تناسب اراضی و مساحت آنها آورده شده است.



شکل ۴: نقشه های تناسب اراضی برای گشت گندم با آبیاری سطحی (الف : روش سنتی و ب: روش جنگل تصادفی)



شکل ۵: مساحت هریک از کلاس های تناسب اراضی در منطقه مورد مطالعه

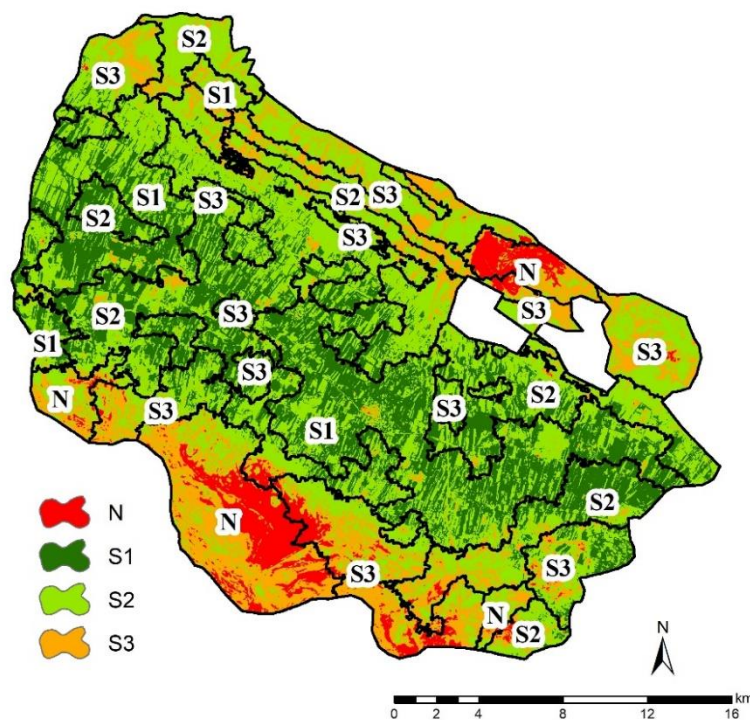
۳-۴. عملکرد مدل جنگل تصادفی و روش سنتی در تهیه نقشه رقومی

برای مقایسه دقت بین نقشه رقومی به دست آمده از مدل جنگل تصادفی و روش سنتی از برآورد شاخص کاپا و دقت کلی استفاده گردید که نتایج آن در جدول ۷ نشان داده شده است. بر اساس طبقه بندی های دقت کلی و شاخص کاپا (da Silva et al.,)

(Strong) (2015; Islami et al., 2022)، نقشه تناسب اراضی گندم با آبیاری سطحی مبتنی بر مدل جنگل تصادفی دارای سطح قوی (Strong) نقشه سنتی تناسب گندم دارای رتبه متوسط (Moderate) هستند (جدول ۶).

جدول ۶: نتایج اعتبار سنجی برآورد کلاس تناسب اراضی			
کلاس تناسب اراضی	دقت کلی		ضریب کاپا
	روش سنتی	جنگل تصادفی	جنگل تصادفی
	۰/۶۲	۰/۷۴	۰/۶۸
			روش سنتی
			جنگل تصادفی

هر دو روش تهیه نقشه از اساس تقریباً یکسانی برخوردار هستند و آن استفاده از متغیرها و داده‌های مبتنی بر اطلاعات مکانی و محیطی است. اما اختلاف آنها که نتیجتاً موجب اختلاف در دقت نقشه‌های خروجی می‌شود می‌تواند در این دانست که در رویکرد سنتی صرفاً از واحدهای ژئومورفولوژیک (مدل رقومی ارتفاع، درصد شیب و جهت شیب) به‌عنوان متغیرهای تناسب اراضی استفاده می‌کند. از سوی دیگر در روش جنگل تصادفی علاوه بر تعدد متغیرهای کمکی از داده‌های کلیه خاک‌های مطالعه شده نیز استفاده شده است. این یافته در راستای یافته‌های (G. Gu et al., 2023; Mokarram et al., 2015; Taghizadeh-Mehrjardi et al., 2020) بود. این محققان بیان داشتند که روش‌های یادگیری ماشینی می‌توانند دقت تحلیل تناسب اراضی را در مقایسه با روش‌های سنتی بهبود بخشند. با این حال، انتخاب روش بستگی به منطقه مورد مطالعه و در میزان در دسترس بودن داده‌ها دارد. البته ارزیابی دقیق نقاط قوت و ضعف روش‌های مختلف و در نظر گرفتن مسائل و چالش‌های بالقوه مرتبط با کاربرد این مدلها نیز بسیار مهم است. یکی دیگر از مزایای استفاده از مدل‌های یادگیری ماشینی خروجی رستری و پیکسل پایه آنهاست که در حقیقت می‌تواند محیط دینامیک خاک را به طور دقیق‌تری به نمایش گذارد زیرا در این گونه روشها مرز بین کلاس‌های تناسب اراضی به صورت پیوسته تغییر میکند در حالی که در نقشه‌های سنتی و پلی‌گونی این مرزها به خطوط مشخص شده و تغییر آبی را نشان می‌دهند که این با ماهیت تغییر پذیری پیوسته خاک و اراضی مغایر است. شکل ۶ از روی هم قرار دادن مرز نقشه سنتی و نقشه رقومی حاصل از روش یادگیری ماشینی به دست آمده است.



شکل ۶: ترکیب نقشه‌های حاصل از مدل جنگل تصادفی و روش سنتی

۴. نتیجه گیری

نقشه‌های مستخرج از هر دو روش برای کشت گندم با آبیاری سطحی چهار کلاس تناسب اراضی S1, S2, S3 و N در منطقه نشان دادند. بیشترین مساحت منطقه متعلق به کلاس تناسب نسبتاً مناسب (S2) در هر دو روش سنتی و یادگیری ماشین به ترتیب با ۲۱۰۲۸ و ۳۰۷۵۳ هکتار بود و کمترین مساحت در روش سنتی با حدود ۷۱۸۵ هکتار و در روش یادگیری ماشین با ۳۰۵۲ هکتار مطلق به کلاس نامناسب (N) بود. همچنین کمتر از ۳۰ درصد از مساحت منطقه مطالعاتی در کلاس بدون محدودیت برای کشت گندم شناسایی شد. با اهمیت ترین عوامل موثر در محدودیت کشت گندم در منطقه، شوری و قلیائیت، نا مناسب بودن بافت خاک، شیب زیاد در برخی مناطق و وجود سنگ‌ریزه مشاهده گردید. مهمترین متغیرهای محیطی در منطقه مورد مطالعه نیز شامل dem, tasseled_cab_g, cnbl بودند و مدل جنگل تصادفی توانست با دقت و شاخص کاپای بالاتری پراکنش کلاس های تناسب اراضی برای کشت گندم با آبیاری سطحی را نسبت به روش سنتی در منطقه مورد مطالعه نشان دهد.

ارزیابی تناسب اراضی یکی از جنبه‌های مهم پایداری تولید کشاورزی است. مدل‌های یادگیری ماشین می‌توانند دقت ارزیابی تناسب اراضی را در مقایسه با روش‌های سنتی بهبود بخشند همان‌طور که تحقیق حاضر نیز نشان داد که دقت مدل‌های یادگیری ماشین نسبت به روش‌های سنتی بیشتر است. این نتایج می‌تواند به این دلیل باشد که مدل‌های یادگیری ماشین از طیف گسترده‌ای از متغیرهای محیطی از جمله ویژگی‌های خاک، آب‌وهوا و توپوگرافی برای پیش‌بینی تناسب اراضی استفاده می‌کنند. از سوی دیگر مدل‌های یادگیری ماشین می‌توانند فرایند طبقه‌بندی تناسب اراضی را سریع‌تر و آن را کارآمدتر و دقیق‌تر کنند. باتوجه به مطالب فوق به نظر می‌رسد استفاده از مدل‌های یادگیری ماشین هیبریدی در تحقیقات بعدی می‌تواند نتایج بهتری را برای تفسیر و ایجاد ارتباط خصوصیات و کلاس اراضی با متغیرهای محیطی در اختیار محققان قرار دهد

"هیچ‌گونه تعارض منافع توسط نویسندگان وجود ندارد."

۵. منابع

- AbdelRahman, M. A. E., Saleh, A. M., & Arafat, S. M. (2022). Assessment of land suitability using a soil-indicator-based approach in a geomatics environment. *Scientific Reports*, 12(1), 18113.
- Akinci, H., Özalp, A. Y., & Turgut, B. (2013). Agricultural land use suitability analysis using GIS and AHP technique. *Computers and Electronics in Agriculture*, 97, 71–82.
- Alexandratos, N., & Bruinsma, J. (2012). *World agriculture towards 2030/2050: the 2012 revision*.
- Al-Mashreki, M. H., Juhari, M. A., Sahibin, A. R., Desa, K. M., Tukimat, L., & Haider, A. R. (2011). Land suitability evaluation for sorghum crop in the Ibb Governorate, Republic of Yemen using remote sensing and GIS techniques. *Australian Journal of Basic and Applied Sciences*, 5(3), 359–368.
- Ashraf, S., & Normohammad, B. (2011). Qualitative evaluation of land suitability for wheat in Northeast-Iran Using FAO methods. *Indian Journal of Science and Technology*, 4(6), 703–707.
- Bagheri Bodaghabadi, M., Martínez- Casanovas, J. A., Khakili, P., Masihabadi, M. H., & Gandomkar, A. (2015). Assessment of the FAO traditional land evaluation methods, A case study: Iranian Land Classification method. *Soil Use and Management*, 31(3), 384–396.

- Behrens, T., & Scholten, T. (2006). Chapter 25 A Comparison of Data-Mining Techniques in Predictive Soil Mapping. In P. Lagacherie, A. B. McBratney, & M. Voltz (Eds.), *Developments in Soil Science* (Vol. 31, pp. 353–617). Elsevier. [https://doi.org/https://doi.org/10.1016/S0166-2481\(06\)31025-2](https://doi.org/https://doi.org/10.1016/S0166-2481(06)31025-2)
- Breiman, L. (2001). Random forests. *Machine Learning*, 45(1), 5–32.
- Brungard, C. W., Boettinger, J. L., Duniway, M. C., Wills, S. A., & Edwards Jr, T. C. (2015). Machine learning for predicting soil classes in three semi-arid landscapes. *Geoderma*, 239, 68–83.
- Daigle, J. J., Hudnall, W. H., Gabriel, W. J., Mersiovsky, E., & Nielson, R. D. (2005). The National Soil Information System (NASIS): Designing soil interpretation classes for military land-use predictions. *Journal of Terramechanics*, 42(3–4), 305–320.
- da Silva, A. F., Barbosa, A. P., Zimback, C. R. L., Landim, P. M. B., & Soares, A. (2015). Estimation of croplands using indicator kriging and fuzzy classification. *Computers and Electronics in Agriculture*, 111, 1–11. <https://doi.org/https://doi.org/10.1016/j.compag.2014.11.020>
- Dorling, D. (2021). World population prospects at the UN: our numbers are not our problem? In *The Struggle for Social Sustainability* (pp. 129–154). Policy Press.
- FAO, F. A. O. (n.d.). of the UN 1976. A framework for land evaluation. *Soil Bulletin*, 32, 72.
- Food and Agriculture Organization of the United Nations. Soil Resources and Conservation Service, M. (1985). *Guidelines, Land Evaluation for Irrigated Agriculture*. FAO.
- Gu, C., Mu, X., Gao, P., Zhao, G., Sun, W., Tatarko, J., & Tan, X. (2019). Influence of vegetation restoration on soil physical properties in the Loess Plateau, China. *Journal of Soils and Sediments*, 19(2), 716–728. <https://doi.org/10.1007/s11368-018-2083-3>
- Gu, G., Wu, B., Zhang, W., Lu, R., Feng, X., Liao, W., Pang, C., & Lu, S. (2023). Comparing machine learning methods for predicting land development intensity. *Plos One*, 18(4), e0282476.
- Hagos, Y. G., Mengie, M. A., Andualem, T. G., Yibeltal, M., Linh, N. T. T., Tenagashaw, D. Y., & Hewa, G. (2022). Land suitability assessment for surface irrigation development at Ethiopian highlands using geospatial technology. *Applied Water Science*, 12(5), 98.
- IaW, F. A. O. (2015). Achieving Zero Hunger: The Critical Role of Investments in Social Protection and Agriculture, Agricultural Development Economics Division. Rome: FAO.
- Igrejas, G., & Branlard, G. (2020). The importance of wheat. *Wheat Quality for Improving Processing and Human Health*, 1–7.
- Islami, F. A., Tarigan, S. D., Wahjunie, E. D., & Dasanto, B. D. (2022). Accuracy assessment of land use change analysis using Google Earth in Sadar Watershed Mojokerto Regency. *IOP Conference Series: Earth and Environmental Science*, 950(1), 012091.
- Khaledian, Y., & Miller, B. A. (2020). Selecting appropriate machine learning methods for digital soil mapping. *Applied Mathematical Modelling*, 81, 401–418. <https://doi.org/https://doi.org/10.1016/j.apm.2019.12.016>
- Khamoshi, S. E., Sarmadian, F., & Omid, M. (2023). Predicting and Mapping of Soil Organic Carbon Stock Using Machin Learning Algorithm, Iranian Journal of Soil and Water Research, 53 (11), 2671-2681. (In Persian)
- Kidd, D., Webb, M., Malone, B., Minasny, B., & McBratney, A. (2015). Digital soil assessment of agricultural suitability, versatility and capital in Tasmania, Australia. *Geoderma Regional*, 6, 7–21.

- Kılıc, O. M., Ersayın, K., Gunal, H., Khalofah, A., & Alsubeie, M. S. (2022). Combination of fuzzy-AHP and GIS techniques in land suitability assessment for wheat (*Triticum aestivum*) cultivation. *Saudi Journal of Biological Sciences*, 29(4), 2634–2644. <https://doi.org/https://doi.org/10.1016/j.sjbs.2021.12.050>
- Lagacherie, P., Arrouays, D., Bourennane, H., Gomez, C., & Nkuba-Kasanda, L. (2020). Analysing the impact of soil spatial sampling on the performances of Digital Soil Mapping models and their evaluation: A numerical experiment on Quantile Random Forest using clay contents obtained from Vis-NIR-SWIR hyperspectral imagery. *Geoderma*, 375, 114503.
- Liu, X., Wang, J., & Song, X. (2023). Improving the Spatial Prediction of Soil Organic Carbon Content Using Phenological Factors: A Case Study in the Middle and Upper Reaches of Heihe River Basin, China. *Remote Sensing*, 15(7). <https://doi.org/10.3390/rs15071847>
- Mahmoudzadeh, H., Matinfar, H. R., Taghizadeh-Mehrjardi, R., & Kerry, R. (2020). Spatial prediction of soil organic carbon using machine learning techniques in western Iran. *Geoderma Regional*, 21, e00260. <https://doi.org/https://doi.org/10.1016/j.geodrs.2020.e00260>
- McBratney, A. B., Santos, M. L. M., & Minasny, B. (2003). On digital soil mapping. *Geoderma*, 117(1–2), 3–52.
- Minasny, B., & McBratney, A. B. (2016). Digital soil mapping: A brief history and some lessons. *Geoderma*, 264, 301–311.
- Mokarram, M., Hamzeh, S., Aminzadeh, F., & Zarei, A. R. (2015). Using machine learning for land suitability classification. *West African Journal of Applied Ecology*, 23(1), 63–73.
- Mugiyo, H., Chimonyo, V. G. P., Sibanda, M., Kunz, R., Masemola, C. R., Modi, A. T., & Mabhaudhi, T. (2021). Evaluation of Land Suitability Methods with Reference to Neglected and Underutilised Crop Species: A Scoping Review. *Land*, 10(2). <https://doi.org/10.3390/land10020125>
- Neyestani, M., Sarmadian, F., Jafari, A., Keshavarzi, A., & Sharififar, A. (2021). Digital mapping of soil classes using spatial extrapolation with imbalanced data. *Geoderma Regional*, 26, e00422. <https://doi.org/https://doi.org/10.1016/j.geodrs.2021.e00422>
- Onyutha, C. (2019). African food insecurity in a changing climate: The roles of science and policy. *Food and Energy Security*, 8(1), e00160.
- Prakash, T. N. (2003). *Land suitability analysis for agricultural crops: a fuzzy multicriteria decision making approach*.
- Roell, Y. E., Beucher, A., Møller, P. G., Greve, M. B., & Greve, M. H. (2020). Comparing a Random Forest Based Prediction of Winter Wheat Yield to Historical Yield Potential. *Agronomy*, 10(3). <https://doi.org/10.3390/agronomy10030395>
- Rossiter, D. G. (2000). Methodology for soil resource inventories. *ITC Lecture Notes SOL*, 27.
- Safari, Y., Esfandiarpour-Boroujeni, I., Kamali, A., Salehi, M. H., & Bagheri-Bodaghabadi, M. (2013). Qualitative land suitability evaluation for main irrigated crops in the shahrekord plain, Iran: A geostatistical approach compared with conventional method. *Pedosphere*, 23(6), 767–778.
- Stoorvogel, J. J., Kempen, B., Heuvelink, G. B. M., & De Bruin, S. (2009). Implementation and evaluation of existing knowledge for digital soil mapping in Senegal. *Geoderma*, 149(1–2), 161–170.
- Sys, C., Van Ranst, E., & Debaveye, J. (1991). *Land evaluation: principles in land evaluation and crop production calculations*. General Administration for Development Cooperation.

- Taghizadeh-Mehrjardi, R., Nabiollahi, K., Rasoli, L., Kerry, R., & Scholten, T. (2020). Land Suitability Assessment and Agricultural Production Sustainability Using Machine Learning Models. *Agronomy*, 10(4). <https://doi.org/10.3390/agronomy10040573>
- Takoutsing, B., & Heuvelink, G. B. M. (2022). Comparing the prediction performance, uncertainty quantification and extrapolation potential of regression kriging and random forest while accounting for soil measurement errors. *Geoderma*, 428, 116192. <https://doi.org/https://doi.org/10.1016/j.geoderma.2022.116192>
- Vasu, D., Srivastava, R., Patil, N. G., Tiwary, P., Chandran, P., & Singh, S. K. (2018). A comparative assessment of land suitability evaluation methods for agricultural land use planning at village level. *Land Use Policy*, 79, 146–163.
- Wang, B., Waters, C., Orgill, S., Cowie, A., Clark, A., Li Liu, D., Simpson, M., McGowen, I., & Sides, T. (2018). Estimating soil organic carbon stocks using different modelling techniques in the semi-arid rangelands of eastern Australia. *Ecological Indicators*, 88, 425–438. <https://doi.org/10.1016/J.ECOLIND.2018.01.049>
- Wylie, B. K., Pastick, N. J., Picotte, J. J., & Deering, C. A. (2019). Geospatial data mining for digital raster mapping. *GIScience & Remote Sensing*, 56(3), 406–429.
- Zakarya, Y. M., Metwaly, M. M., AbdelRahman, M. A. E., Metwalli, M. R., & Koubouris, G. (2021). Optimized land use through integrated land suitability and GIS approach in West El-Minia Governorate, Upper Egypt. *Sustainability*, 13(21), 12236.
- Zhang, M., Shi, W., & Xu, Z. (2020). Systematic comparison of five machine-learning models in classification and interpolation of soil particle size fractions using different transformed data. *Hydrology and Earth System Sciences*, 24(5), 2505–2526. <https://doi.org/10.5194/hess-24-2505-2020>
- Ziadat, F. M. (2000). *Application of GIS and remote sensing for land use planning in the arid areas of Jordan*. Cranfield University (United Kingdom).

Land Suitability Evaluation Using Traditional and Machine Learning Approaches: A Case Study in Abiek Plain, Qazvin Province, Iran

Extended Abstract

Introduction:

Various factors such as rising food prices, economic factors, and climate change have led to crises in different parts of the world. To address these challenges, the world needs to create organizations and develop various programs to ensure food security and reduce food waste. One important program is increasing food production without increasing the cultivated area, which requires evaluating land suitability and identifying the main obstacles to achieving maximum crop performance in each region. Land suitability assessment is an essential step in optimizing and sustainable land use planning, especially in crop rotation planning. Traditional methods of soil mapping have limitations in representing the continuous nature of soil changes. Therefore, it is necessary to use different methods to better understand the variability of land suitability classes for sustainable land management. Machine learning techniques can be used to improve land suitability analysis. Data mining methods and digital mapping attempt to identify the environmental variables that are easily accessible, the features of the soil, and the land suitability classes. Therefore, this study aimed to evaluate land suitability for sustainable agricultural production in the Abiek region of Qazvin province in northwest Iran, covering an area of 60,000 hectares .

Methods:

The study collected data from 300 soil profiles and determined land suitability classes for wheat cultivation with surface irrigation using the FAO classification system. Digital elevation models, Landsat-8 and Sentinel-2 satellite images, and environmental variables extracted from the digital elevation model were used to create digital maps using both traditional and machine learning methods. The traditional method used the FAO classification system, while the machine learning method used a classification algorithm based on environmental variables. The accuracy of both methods was evaluated using a Kappa index and overall accuracy .

Results and Discussion:

The study found that the machine learning method had a higher accuracy rate of 74% and a Kappa index of 68 compared to the traditional method with an accuracy rate of 62% and a Kappa index of 53. The most important environmental variables used in the machine learning model were those extracted from the digital elevation model and Landsat-8 satellite images. The study identified the most suitable class (S2) for wheat cultivation with surface irrigation and the least suitable class (N) in terms of land suitability. The results of this study can be used to optimize land use, promote sustainable land use, protect the environment, and ensure optimal use of natural resources in the study area. The study identified the most suitable and least suitable classes for wheat cultivation with surface irrigation, which can be used to guide land use planning and agricultural production in the region. The use of machine learning methods for land suitability analysis can be further explored in future studies.

Keywords: *Geomorphological characteristics, parametric method, Random Forests, Wheat.*