

Application of Near-Spectroscopy Data and Environmental Variables in Mapping Saturated Hydraulic Conductivity of Soil in the Kilaneh Watershed, Kurdistan Province

Farzaneh Parsaie¹ | Ahmad Farrokhian Firouzi^{1*}  | Masoud Davari² |
Ruhollah Taghizadeh-Mehrjardi³

1. Department of Soil Science and Engineering, Faculty of Agriculture, Shahid Chamran University of Ahvaz, Ahvaz, Iran. E-mail: a.farrokhian@scu.ac.ir
2. Department of Soil Science, Faculty of Agriculture, University of Kurdistan, Sanandaj, Iran.
3. Department of Geosciences, Soil Science and Geomorphology, University of Tübingen, Tübingen, Germany.

Article Info

Abstract

Article type:
Research Article

Article history:
Received: 29 Jan. 2025
Revised: 14 Mar. 2025
Accepted: 05 Apr. 2025
Published online: 23 Sept. 2025

Keywords:
Environmental variables,
Machine learning,
Spatial variability,
Soil sensing.

Surface soil saturated hydraulic conductivity (Ks), as one of the most important physical properties of soil, plays a key role in the distribution of water and nutrients within the soil environment and holds particular significance in water and soil resource management. This study aimed to digitally model Ks using machine learning approaches in the Kilaneh watershed, located in Kurdistan Province, covering an area of 12,000 hectares. Three machine learning algorithms, including Gradient Boosted Decision Tree (XGBoost), Random Forest (RF), and k-Nearest Neighbors (k-NN), were utilized, incorporating various environmental variables derived from the digital elevation model and Sentinel-2 satellite imagery. These variables included distance from the drainage channel, valley depth, relative slope position, channel base level, brightness index, wind effect index, Normalized Difference Vegetation Index (NDVI), Band 12, greenness index, and surface curvature. Additionally, soil parameters such as organic matter, lime content, bulk density, geometric mean particle diameter, soil texture, and near-soil spectroscopic data (Latent Variable) within the wavelength range of 400–2450 nm were used as proxies for pedogenic factors to model saturated hydraulic conductivity. The results indicated that the XGBoost model exhibited the highest accuracy for predicting Ks, with an R² value of 0.65 and an nRMSE of 0.25, outperforming the other models. Spectral data, topographic variables, and soil parameters, as model inputs, played a significant role in predicting the spatial variability of Ks. The XGBoost model was able to provide highly accurate predictions. The results demonstrated that topographic, physical, and spectral variables influence Ks; organic matter, soil texture, and topographic indices such as slope and relative position had the most substantial impact. The generated maps can be utilized for water and soil resource management and hydrological models.

Cite this article: Parsaie, F., Farrokhian Firouzi, A., Davari, M., Taghizadeh-Mehrjardi, R. (2025). Application of Near-Spectroscopy Data and Environmental Variables in Mapping Saturated Hydraulic Conductivity of Soil in the Kilaneh Watershed, Kurdistan Province. *Journal of Range & Watershed Management*, 78 (3), 323-340. DOI: <http://doi.org/10.22059/jrwm.2025.389637.1801>



کاربرد داده‌های طیف‌سنجی نزدیک و متغیرهای محیطی در تهیه نقشه هدایت هیدرولیکی اشباع خاک در حوزه آبخیز کیلانه، استان کردستان

فرزانه پارسایی^۱ | احمد فرخیان فیروزی^{۱*} | مسعود داوری^۲ | روح‌الله تقی‌زاده مهرجردی^۳

۱. گروه علوم و مهندسی خاک، دانشکده کشاورزی، دانشگاه شهید چمران اهواز، اهواز، ایران.

رایانامه: a.farrokhian@scu.ac.ir

۲. گروه علوم و مهندسی خاک، دانشکده کشاورزی، دانشگاه کردستان، سنندج، ایران.

۳. گروه علوم زمین، علوم خاک و ژئومرفولوژی، دانشگاه توبینگن، توبینگن، آلمان.

اطلاعات مقاله

چکیده

نوع مقاله:

مقاله پژوهشی

تاریخ دریافت: ۱۴۰۳/۱۱/۱۰

تاریخ بازنگری: ۱۴۰۳/۱۲/۲۴

تاریخ پذیرش: ۱۴۰۴/۰۱/۱۶

تاریخ انتشار: ۱۴۰۴/۰۷/۰۱

کلیدواژه‌ها:

تغییرپذیری مکانی،

متغیرهای محیطی،

سنجش خاک،

یادگیری ماشین.

هدایت هیدرولیکی اشباع خاک سطحی (K_s) به‌عنوان یکی از مهم‌ترین ویژگی‌های فیزیکی خاک، نقشی کلیدی در توزیع آب و مواد مغذی در محیط خاک ایفا می‌کند و در مدیریت منابع آب و خاک اهمیت ویژه‌ای دارد. این پژوهش با هدف مدل‌سازی رقومی هدایت هیدرولیکی اشباع خاک سطحی با استفاده از رویکردهای یادگیری ماشین در حوزه آبخیز کیلانه واقع در استان کردستان با مساحت ۱۲ هزار هکتار انجام شد. سه الگوریتم یادگیری ماشین شامل: درخت تصمیم تقویت شده با گرادیان (XGBoost)، جنگل تصادفی (RF) و مدل نزدیکترین k-همسایگی (k-NN) با بهره‌گیری از تعدادی متغیرهای محیطی از مدل رقومی ارتفاع و تصاویر ماهواره سنتینل-۲ شامل فاصله از کانال آبراهه، عمق دره، موقعیت نسبی شیب، سطح پایه کانال آبراهه، شاخص روشنایی، شاخص اثر باد، شاخص نرمال شده تفاوت پوشش گیاهی، باند ۱۲، شاخص سبزیگی، انحنا سطح و پارامترهای خاک شامل ماده آلی، آهک، جرم مخصوص ظاهری، میانگین هندسی قطر ذرات، بافت خاک و داده‌های طیف‌سنجی نزدیک خاک در طول موج ۲۴۵۰-۴۰۰ نانومتر به‌عنوان نمایندگان عوامل خاکساز برای مدل‌سازی K_s مورد استفاده قرار گرفتند. نتایج نشان داد که مدل XGBoost برای پیش‌بینی K_s با R^2 برابر ۰/۶۵ و nRMSE برابر ۰/۲۵ نسبت به سایر مدل‌ها دارای صحت بالاتری بودند. داده‌های طیفی، متغیرهای توپوگرافی و پارامترهای خاک، به‌عنوان ورودی مدل، نقش مهمی در پیش‌بینی تغییرپذیری مکانی K_s داشتند و مدل XGBoost با استفاده از این داده‌ها توانست پیش‌بینی دقیقی ارائه دهد. نتایج نشان داد که K_s تحت تأثیر متغیرهای توپوگرافی، فیزیکی و طیفی قرار دارد؛ ماده آلی، بافت خاک و شاخص‌های توپوگرافی مانند شیب و موقعیت نسبی بیشترین تأثیر را داشتند. نقشه‌های تولیدشده از این رویکرد تغییرپذیری مکانی می‌توانند در مدیریت منابع آب و خاک و مدل‌های هیدرولوژیکی مورد استفاده قرار گیرند.

استناد: پارسایی؛ فرزانه، فرخیان فیروزی؛ احمد، داوری؛ مسعود، تقی‌زاده مهرجردی؛ روح‌الله (۱۴۰۴). کاربرد داده‌های طیف‌سنجی نزدیک و متغیرهای محیطی در تهیه نقشه هدایت هیدرولیکی اشباع خاک در حوزه آبخیز کیلانه، استان کردستان. نشریه مرتع و آبخیزداری، ۷۸ (۳)، ۳۴۰-۳۲۳.

DOI: <http://doi.org/10.22059/jrwm.2025.389637.1801>



© نویسندگان.

ناشر: انتشارات دانشگاه تهران.

۱. مقدمه

خاک یکی از منابع طبیعی حیاتی است که بشر برای تولید غذا، فیبر و انرژی به آن وابسته است. این ماده، جریان آب را تنظیم کرده و به عنوان فیلتری برای عناصر سمی و آلاینده‌ها عمل می‌کند (Niede & Benbi., 2022). اندازه‌گیری ویژگی‌های هیدرولیکی خاک مانند هدایت هیدرولیکی اشباع خاک از مهم‌ترین ویژگی‌های فیزیکی خاک می‌باشد و تخمین آن در مطالعات آب و خاک اهمیت زیادی دارد. سهولت نفوذ ریشه گیاهان، قابلیت نگهداری عناصر غذایی، تهویه، نگهداشت رطوبت، حرکت آب و املاح، پایداری و تحمل فشار، قابلیت زهکشی همگی به ویژگی‌های فیزیکی و مکانیکی خاک مرتبط هستند. این ویژگی‌ها بر رفتار شیمیایی و بیولوژیک خاک نیز تأثیر می‌گذارند و به عواملی مانند ترکیب معدنی، محتوای مواد آلی و فضاهای خالی بستگی دارند (Ottoni et al., 2025). به طور کلی، می‌توان گفت که خاک سیستمی ناهمگن و پویاست که دارای فرآیندها و مکانیسم‌های بسیار پیچیده‌ای است. مطالعه خاک به عنوان یک سیستم کامل و یک منبع مهم طبیعی، نیازمند توسعه روش‌هایی تحلیلی می‌باشد. اندازه‌گیری مستقیم ویژگی‌های فیزیکی و شیمیایی خاک به دلیل زمان بر بودن و هزینه‌های بالا، نیازمند مطالعات میدانی گسترده و نمونه‌برداری‌های متعدد است که به تغییرپذیری زمانی و مکانی این ویژگی‌ها مرتبط می‌شود. از این رو، در دهه‌های اخیر، توسعه روش‌های غیرمستقیم، سریع و مقرون‌به‌صرفه به یکی از اولویت‌های اصلی پژوهش‌های علمی در حوزه علوم خاک تبدیل شده است (Ma et al., 2024). اخیراً، استفاده از روش‌های دورسنجی در تخمین ویژگی‌های خاک به علت صرف زمان، هزینه و عملیات صحرائی کمتر، به میزان زیادی مورد توجه قرار گرفته است (Nanni and Demattê, 2006). علاوه بر روش‌های دورسنجی اخیراً طیف‌سنجی خاک^۱ یکی دیگر از روش‌های مورد استفاده و پرکاربرد معرفی شده است (Kubiak et al., 2024). طیف‌سنجی تهیه منحنی‌های انعکاس طیفی^۲ پدیده یا پدیده‌های موردنظر در گستره طول موج‌هایی مشخص می‌باشد (Schneider and Young, 1997). منحنی‌های انعکاس طیفی خاک می‌توانند بسیاری از محدودیت‌های آزمایشگاهی را برطرف نماید، زیرا برخلاف روش‌های مرسوم اندازه‌گیری ویژگی‌های خاک در آزمایشگاه، داده‌های خاک در رویکرد طیفی با یک روش استاندارد و تنها در یک مرحله بدون نیاز به مواد شیمیایی گران‌قیمت و سمی تعیین می‌گردند. بنابراین خطای ناشی از روش‌های گوناگون اندازه‌گیری و شخص‌کاربر به‌طور قابل‌توجهی کاهش می‌یابد (Babaeian et al., 2015). استفاده از بازتاب طیفی خاک در طول موج‌های بین ۳۵۰ تا ۲۵۰۰ نانومتر در دامنه‌های طیفی مرئی (Vis) و مادون‌قرمز نزدیک (NIR) به عنوان جایگزین روش‌هایی آزمایشگاهی به کار برده می‌شوند (Janik et al., 2009). از این رو می‌توان از انعکاس طیفی خاک مشابه با توابع انتقالی خاک^۳ برای پیش‌بینی ویژگی‌های خاک استفاده کرد و توابع جدیدی تحت عنوان توابع انتقالی طیفی^۴ ایجاد کرد. با توجه به پژوهش‌هایی که انجام شده است استفاده از طیف‌سنجی خاک می‌تواند بسیاری از محدودیت‌های توابع انتقالی خاکی را برطرف نماید؛ زیرا در طیف‌سنجی، متغیرهای ورودی توابع تنها داده‌های طیفی خاک هستند که بر خلاف ویژگی‌های زودیافت خاک، در یک مرحله و با یک روش استاندارد اندازه‌گیری می‌شوند (بابائیان و همکاران، ۱۳۹۳). علاوه بر این، طیف‌سنجی و بخصوص طیف‌سنجی مادون‌قرمز، می‌تواند اجزای مختلف خاک را نیز مشخص کند. این فناوری این امکان را فراهم می‌کند که به‌صورت هم‌زمان بیش از ۲۰ ویژگی خاک را تنها با یک اندازه‌گیری، چه در مزرعه و چه در آزمایشگاه، تخمین زد (Kim et al., 2014). مطالعات متعدد نشان داده‌اند که بسیاری از ویژگی‌های فیزیکی (مانند پایداری تجمعی خاک)، شیمیایی (مانند ظرفیت تبادل کاتیونی)، بیولوژیکی (مانند فعالیت میکروبی) و حفاظتی خاک (مانند میزان ماده آلی که در حفظ ساختار و جلوگیری از فرسایش نقش دارد) با استفاده از روش‌های طیف‌سنجی قابل برآورد هستند. طبق پژوهش‌های انجام شده بازتاب طیفی خاک به کمک بعضی از ویژگی‌های خاک همانند بافت، ساختمان، رطوبت و میزان ماده آلی خاک تحت‌تأثیر قرار گرفته که بسیاری

¹ Soil Spectroscopy

² Proximal spectral reflectance

³ Pedotransfer functions

⁴ Spectra pedotransfer functions

از ویژگی‌های خاک را می‌توان برآورد کرد (Vairavan et al., 2024). چانگ^۱ و همکاران (۲۰۰۱) در پژوهشی گزارش کردند که خصوصیات جذبی در دامنه‌های طیفی ۱۴۵۰-۱۳۰۰، ۱۹۵۰-۱۸۵۰ و ۲۴۰۰-۲۲۰۰ نانومتر مربوط به مقدار رس خاک هستند (Chang et al., 2001). توزیع اندازه ذرات خاک نیز تأثیر زیادی بر رفتار طیفی خاک دارد. به‌طور کلی، با افزایش مسیر عبور نور از بین ذرات خاک و بزرگ شدن اندازه ذرات، نور بیشتری توسط خاک جذب شده و مقدار بازتاب کاهش می‌یابد. در این حالت، مشخصه‌های جذبی روی منحنی‌های طیفی ظاهر می‌شوند (Babaeian et al., 2015).

در پژوهش‌های مختلف نشان داده است ویژگی‌هایی از خاک که در ارتباط با منافذ موجود در خاک هستند (همچون جرم مخصوص ظاهری، ضریب آبگذری اشباع خاک و همچنین رطوبت اشباع) نمی‌توانند با دقت بالایی توسط طیف‌سنجی در دامنه مادون قرمز دور پیش‌بینی شوند؛ این درحالی است ویژگی‌هایی که به بخش جامد خاک (همچون رس، شن و ماده آلی) مرتبط هستند می‌توانند با دقت بالاتری برآورد شوند (Minansy et al., 2008). بازتاب طیفی ممکن است تحت تأثیر عوامل محیطی مزاحم، مانند نور پراکنده ناشی از پدیده‌های مجاور، قرار گیرد. بنابراین، لازم است با استفاده از پردازش‌های مناسب، مانند شاخص‌های طیفی، نسبت‌گیری‌های طیفی و تبدیل‌های آماری و ریاضی، اثرات منفی این عوامل حذف یا کاهش یابد. این پردازش‌ها می‌توانند تفاوت بازتاب پدیده‌ها را آشکارتر کرده و نتایج حاصل را با حساسیت بیشتری نسبت به ویژگی‌های خاک و متغیرهای موردنظر ارائه دهند.

برای ارتباط بین ویژگی‌های طیفی با ویژگی‌های مبنایی خاک (مانند بافت خاک، ماده آلی، pH، ظرفیت تبادل کاتیونی و غلظت عناصر غذایی) روش‌های مختلفی مانند رگرسیون حداقل مربعات جزئی^۲، شبکه‌های عصبی^۳، رگرسیون مؤلفه‌های اصلی^۴ و رگرسیون‌های خطی^۵ و غیرخطی و همچنین یادگیری ماشین در تلفیق با اطلاعات طیفی مورد استفاده قرار گرفته است. شایان ذکر است که از کاربردی‌ترین روش‌ها در این خصوص می‌توان به الگوریتم‌های یادگیری ماشین همچون مدل‌های جنگل تصادفی و کوبیست اشاره کرد (Sabetizadeh et al., 2021; Parsaie et al., 2021; Ng et al., 2021). پژوهشگران پیشنهاد می‌کنند که این نوع مدل‌ها در پیش‌بینی‌ها می‌توانند بهترین عملکرد را داشته باشند. الگوریتم‌های یادگیری ماشین به‌عنوان بخشی از دانش هوش مصنوعی در عصر فناوری و اطلاعات به‌طور فزاینده‌ای در زمینه‌های مختلف علمی به ویژه نقشه‌برداری رقومی خاک^۶ (DSM) توسعه یافته‌اند (Hengl et al., 2018; Minasny and Mcbratney, 2016; Mousavi et al., 2021). یکی از کاربردهای عمده الگوریتم‌های یادگیری ماشین عمدتاً شناسایی و پیش‌بینی الگوهای موجود در چندین مجموعه داده بزرگ حاصل از داده‌های ماهواره‌ای یا پارامترهای مشتق شده از مدل رقومی ارتفاع^۷ (DEM) است. این داده‌ها به‌عنوان کنترل‌کننده‌ها یا پیش‌بینی‌کننده‌های محیطی که نماینده فاکتورهای خاکساز هستند نقش دارند (Hengl et al., 2018; Padarian et al., 2020). در همین راستا مندس^۸ و همکاران (۲۰۲۳) گزارش کرده‌اند که الگوریتم‌های یادگیری ماشین^۹ (MLM) برای تهیه نقشه‌ی ویژگی‌های خاک در اراضی کشاورزی با تعداد نمونه کمتر نسبت به روش‌های زمین آماری کارا تر هستند.

از روش‌های MLM برای برآورد ویژگی‌های فیزیکوشیمیایی و مکانیکی خاک، می‌توان به مدل XGboost برای مدل‌سازی مکانی کربن آلی خاک (SOC^{۱۰}) (Mousavi et al., 2022; Matinfar et al., 2021; Mahmoudzadeh et al., 2020; Nguyen et al., 2022)،

¹ Chang

² PLSR

³ NN

⁴ PCR

⁵ MLR

⁶ Digital Soil Mapping

⁷ Digital Elevation Model

⁸ Mendes

⁹ Machine Learning Models

¹⁰ Soil Organic Carbon

مدل جنگل تصادفی (RF^1) برای هدایت هیدرولیکی اشباع خاک (Rezaei et al., 2023; Masoudi et al., 2023) و رگرسیون بردار پشتیبان (SVR^2) برای برآورد شوری خاک (Mousavi et al., 2021; Taghizadeh-Mehrjardi et al., 2021)، و مدل k-NN برای برآورد برخی ویژگی‌های خاک سطحی خالدیان و میلر^۳ (۲۰۲۰) و عناصر غذایی خاک اشاره کرد (Parsaie et al., 2021; Gao et al., 2019). تا کنون، پژوهش‌های متعددی در زمینه تخمین ویژگی‌های شیمیایی خاک با استفاده از بازتاب طیفی انجام شده است و نتایج امیدوارکننده‌ای نیز به دست آمده است. با این حال، پیش‌بینی ویژگی‌های فیزیکی خاک از طریق خصوصیات طیفی آن هنوز تا حدودی دشوار است و به تحقیقات بیشتری نیاز دارد. مرور منابع موجود نشان می‌دهد که مطالعات کمی در زمینه برآورد ویژگی‌های هیدرولیکی خاک با استفاده از داده‌های طیفی صورت گرفته است (Santra et al., 2009; Babaeian et al., 2015; Xu et al., 2017; Knotters et al., 2017; Dharumarajan et al., 2022).

شناخت روابط بین ویژگی‌های مختلف خاک و کمی کردن آن‌ها در قالب مدل‌های تجربی یکی از جنبه‌های مهم بررسی خاک است که چنین مدل‌هایی را توابع انتقالی می‌گویند. ویژگی‌های زودیافت خاک مانند بافت خاک و توزیع اندازه ذرات اولیه به طور گسترده‌ای در تخمین ویژگی‌های هیدرولیکی خاک مانند هدایت آبی اشباع و غیراشباع و منحنی رطوبتی استفاده شده‌اند (Oosterban and Homaei and Farrokhian Firouzi, 2008; Xu et al., 2017; Nasta et al., 2021; Aguilera et al., 2022). گزارش (۲۰۱۴) سکوتی و لشکری‌پور (۲۰۱۴) کردند که حرکت آب در خاک متأثر از ویژگی‌هایی همچون رطوبت اولیه، درجه اشباع و ویژگی‌های فیزیکی نظیر دانه‌بندی خاک قرار دارد. همچنین توزیع اندازه خلل و فرج و جرم مخصوص ظاهری دو ویژگی بسیار مهم هستند، به طوری که با افزایش جرم مخصوص ظاهری خاک، در برخی مطالعات کاهش هدایت هیدرولیکی خاک گزارش شده است.

به دلیل تغییرپذیری مکانی زیاد ویژگی‌های فیزیکی و هیدرولیکی خاک احتمالاً ترکیب داده‌های طیفی با متغیرهای محیطی شامل ویژگی ژئومورفومتری مستخرج از مدل رقومی ارتفاع، داده‌های سنجش از دور و ماهواره‌ای و خود ویژگی‌های خاکی، باعث افزایش دقت برآورد این ویژگی‌ها خواهد شد. سابتیزاده^۴ و همکاران (۲۰۲۱) طی پژوهشی نتیجه گرفتند که ترکیب داده‌های طیفی و متغیرهای محیطی باعث افزایش برآورد مقدار ماده آلی خاک شد. این در حالی است با دانش ما و با مروری بر منابع انجام شده در این زمینه، تاکنون پژوهشی که در آن از ترکیب داده‌های طیفی و متغیرهای محیطی برای برآورد ویژگی‌های فیزیکی خاک استفاده شده باشد، یافت نشد. از آنجائی که ویژگی‌های فیزیکی خاک مانند هدایت هیدرولیکی اشباع خاک دارای تغییرپذیری مکانی زیادی هستند، لذا در این پژوهش به دنبال آن بودیم که با ترکیب داده‌های طیف‌سنجی و محیطی دقت مدل‌های پیش‌بینی کننده را افزایش دهیم. از طرفی ویژگی‌های هیدرولیکی خاک تحت تاثیر تخلخل و توزیع اندازه منافذ خاک هستند لذا در این پژوهش همچنین تلاش شد که ویژگی‌های زودیافتی از خاک که بر تخلخل و توزیع اندازه منافذ خاک مؤثر هستند (مانند توزیع اندازه ذرات، جرم مخصوص ظاهری) به همراه داده‌های طیف‌سنجی و محیطی در افزایش دقت برآورد هدایت هیدرولیکی اشباع به کار گرفته شود.

در مجموع در این پژوهش تلاش بر آن بوده است که در مدل‌های پیش‌بینی کننده هدایت هیدرولیکی اشباع خاک از متغیرهای محیطی همراه با داده‌های طیف‌سنجی نزدیک برای افزایش کارایی مدل‌های برآوردگر استفاده شود. لذا در این پژوهش از داده‌های انعکاس طیفی خاک در گستره مرئی - مادون قرمز نزدیک و همچنین متغیرهای محیطی برای برآورد ویژگی فیزیکی هدایت هیدرولیکی اشباع خاک با هدف مدیریت بهتر منابع آب و خاک در حوضه آبخیز کیلانه واقع در استان کردستان با کاربری‌های مختلف استفاده شد.

¹ Random Forest

² Support Vector Regression

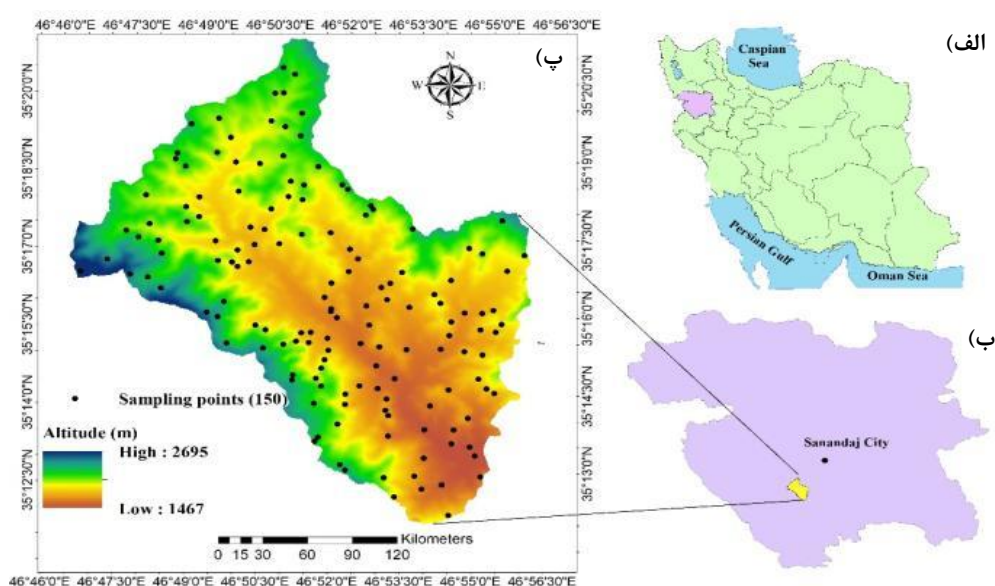
³ Khaledian and Miller

⁴ Sabetizadeh

۲. مواد و روش‌ها

۲-۱. منطقه مورد مطالعه

حوضه آبخیز کیلانه، واقع در استان کردستان، با مساحتی حدود ۱۲ هزار هکتار و میانگین ارتفاع ۱۸۵۵ متر از سطح دریا، به‌عنوان منطقه مطالعه انتخاب شد (شکل ۱). این منطقه با مختصات جغرافیایی مشخص (شکل ۱) در ۴۰ کیلومتری شهر سنندج در جهت شمال استان کردستان قرار گرفته است (شکل ۱). حدود ۱۱۷۹۵/۳۰۶ هکتار (۹۷/۸۴ درصد) از اراضی این حوضه را مناطق کوهستانی و تپه‌ای تشکیل می‌دهند و ۲۰۶/۳۵۶ هکتار (۲/۱۶ درصد) دیگر شامل اراضی متفرقه است که بخشی از آن‌ها از طریق چشمه‌ها و آبراهه‌ها آبیاری می‌شوند. برخی از اراضی تپه‌ماهور نیز به زراعت آبی و باغات اختصاص داده شده‌اند. از این میان، ۱۶۳ هکتار به باغات، ۱۳۴ هکتار به بیرون‌زدگی‌های سنگی، ۸۴۳۸ هکتار به مراتع، ۴۴ هکتار به مناطق مسکونی، ۳۲۸ هکتار به زراعت آبی و ۲۹۴۸ هکتار به زراعت دیم اختصاص دارد. این منطقه با دمای میانگین سالیانه ۱۰ درجه سانتی‌گراد و متوسط بارندگی ۵۰۹ میلی‌متر طی یک دوره آماری ۲۰ ساله (۱۳۷۹-۱۳۹۹) و با پردازش داده‌های اقلیمی در نرم‌افزار NSM، دارای رژیم رطوبتی زیریک^۱ و رژیم حرارتی مزیک^۲ است (Banaie, 1998). واحدهای کوهستان و تپه به‌عنوان اجزای اصلی سیمای اراضی در منطقه به‌شمار می‌روند. خاک‌های این منطقه بر اساس رده‌بندی آمریکایی خاک‌ها کارکنان بررسی خاک^۳، (۲۰۲۲) در دو رده اینسپتی‌سول^۴ و انتی‌سول^۵ طبقه‌بندی شده‌اند. تغییرات ارتفاعی قابل توجه و شیب‌های بیش از ۳۰ درصد، به‌طور عمده به ساختار کوهستانی و تپه‌ای منطقه منجر شده است. حوضه آبخیز کیلانه عمدتاً از سازندهای آهکی، شیل، مارن و ماسه‌سنگ مربوط به دوره‌های پالئوزوئیک و مزوزوئیک تشکیل شده است. حضور گسل‌های محلی و شکستگی‌های زمین‌شناسی، نفوذپذیری و هیدرولوژی منطقه را تحت تأثیر قرار داده است.



شکل ۱. موقعیت (الف) حوزه آبخیز کیلانه نسبت به کشور ایران (ب) در استان کردستان (پ) و توزیع مکانی نمونه‌های خاک: دایره‌های سیاه

^۱ رژیم رطوبتی زیریک: به شرایطی اطلاق می‌شود که میزان بارش سالیانه کمتر از تبخیر و تعرق است و خاک در بیشتر ماه‌های سال خشک است.

^۲ رژیم حرارتی مزیک: به شرایط اقلیمی اطلاق می‌شود که دمای میانگین سالیانه بین ۶ تا ۱۵ درجه سانتی‌گراد است و در آن دوره‌های گرم و سرد مشخصی وجود دارد.

^۳ Soil Science Society of America

^۴ اینسپتی‌سول: خاک‌های جوان با تحول ضعیف که معمولاً در مراحل ابتدایی تکامل خاک قرار دارند.

^۵ انتی‌سول: خاک‌های بسیار جوان با تحول ناچیز یا بدون افق‌های مشخص، اغلب در مناطق با فرسایش شدید یا رسوبات تازه.

۲-۲. تهیه و اجرای پایگاه داده‌ها

در این پژوهش، برای تعیین محل نقاط نمونه‌برداری خاک سطحی (عمق ۰ تا ۱۰ سانتی‌متر)، از روش نمونه‌برداری ابر مکعب لاتین مشروط (CLHS) استفاده شد. موقعیت ۱۵۰ نقطه مشاهداتی با بهره‌گیری از سیستم موقعیت‌یاب جهانی ثبت گردید (شکل ۱). نمونه‌های جمع‌آوری شده پس از خشک شدن در هوا، کوبیدن و عبور از الک ۲ میلی‌متری، به دو بخش تقسیم شدند: یک بخش برای انجام تجزیه‌های آزمایشگاهی و بخش دیگر برای مطالعات طیف‌سنجی خاک به کار رفت. اندازه‌گیری هدایت هیدرولیکی اشباع خاک در نمونه‌های دست نخورده و با روش بارثابت یا بار افتان (Klute and Dirksen, 1986) انجام شد.

۲-۳. بازتاب طیفی نمونه‌های خاک

بازتاب طیفی نمونه‌های خاک مورد بررسی با استفاده از دستگاه اسپکتورادیومتر زمینی (FieldSpec @3, ASD, FR, USA) در آزمایشگاه خاک‌شناسی دانشگاه تربیت مدرس اندازه‌گیری شد. برای این منظور، حدود ۶۰ گرم خاک دست‌خورده هوا خشک، که از الک ۲ میلی‌متری عبور داده شده بود، در یک پتری دیش با قطر ۱۰ سانتی‌متر و ضخامت یک سانتی‌متر قرار داده شد. سپس، منحنی‌های طیفی آن‌ها در دامنه‌های مرئی، مادون قرمز نزدیک و میانی (۳۵۰-۲۵۰۰ نانومتر) با استفاده از روش‌های استاندارد طیف‌سنجی در تاریک‌خانه و با چندین تکرار برای هر نمونه خاک ثبت گردید (Viscarra Rossel, 2008).

منحنی بازتاب طیفی نمونه‌ها با استفاده از نرم‌افزار RS3 بر روی رایانه قابل حمل متصل به دستگاه اسپکتورادیومتر ثبت شد. به‌طور میانگین، برای هر نمونه پنج قرائت بازتاب متوالی انجام و ثبت گردید. سپس، با بهره‌گیری از نرم‌افزار ViewSpec، میانگین قرائت‌های مختلف برای هر نمونه خاک محاسبه شد. این منحنی‌ها به‌صورت فایل‌های متنی و کتابخانه‌های طیفی ذخیره شدند. به دلیل وجود آشفته‌گی‌های طیفی در بازه طول‌موج‌های ۴۰۰-۳۵۰ و ۲۵۰۰-۲۴۵۱ نانومتر، این محدوده‌ها از تمامی طیف‌ها حذف شده و در فرایند مدل‌سازی استفاده نگردید (Gomez et al., 2008). همچنین، برای حذف آشفته‌گی‌های موجود در طیف‌های خاک (مانند رفتار غیرخطی باندها، فیلتر، بارزسازی طیفی و نرمال‌سازی طیف‌ها) و ارتقای کیفیت داده‌ها، از روش‌های مختلف پیش‌پردازش با نرم‌افزار Parles نسخه ۱.۳ بر روی داده‌های طیفی استفاده شد. هدف اصلی از پیش‌پردازش داده‌های طیفی، حذف عوامل خارجی و بهبود استفاده مؤثر از داده‌ها در فرایند تخمین بود. بر اساس روش اعتبارسنجی متقابل و بررسی دقیق نتایج، انواع روش‌های پیش‌پردازش شامل فیلتر میانه^۱، فیلتر ساویتزکی-گلای^۲، نرمال‌سازی بر اساس میانگین^۳، متغیر نرمال استاندارد^۴ (SNV)، تصحیح پخشیده چندگانه^۵ (MSC)، مشتق اول همراه با فیلتر ساویتزکی-گلای^۶ و مشتق دوم همراه با فیلتر ساویتزکی-گلای^۷ اعمال شدند. در نهایت، بهترین روش برای پیش‌پردازش داده‌های طیفی خاک، مشتق اول همراه با فیلتر ساویتزکی-گلای انتخاب گردید. در شکل ۲ میانگین منحنی‌های خام خاک‌های مورد مطالعه با مقادیر بازتاب طیفی در هر طول موج، از گستره ۲۴۰۰-۴۰۰ نانومتر ارائه شده است.

۲-۴. متغیرهای محیطی

در این مطالعه، از متغیرهای محیطی مختلفی استفاده شد که شامل مدل رقومی ارتفاع با دقت مکانی ۱۲/۵ متر بود. این مدل از داده‌های ماهواره ALOS PALSAR (2011) در سیستم مختصات جغرافیایی (WGS84) استخراج شد. در همین راستا متغیرهای توپوگرافی از مدل

¹ Median Filter

² Savitzsky-Golay filter (SGD)

³ Normalization in range <-1, >1

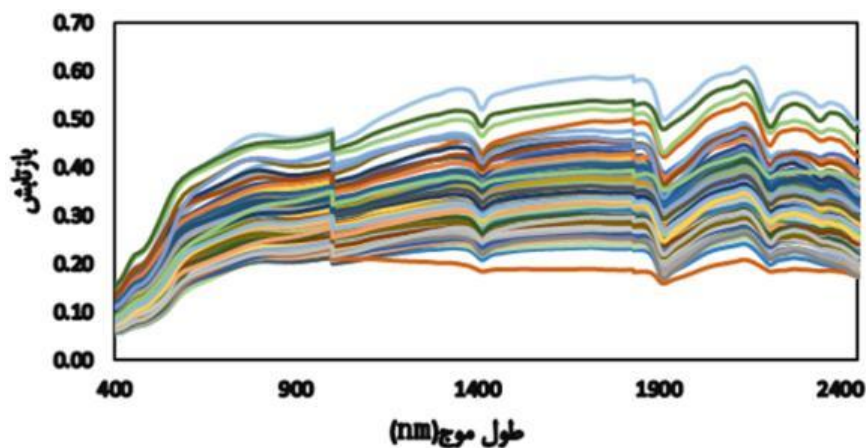
⁴ Standard Normal Variate transformation

⁵ Multiplicative Scatter Correction

⁶ First Derivative + Savitzky-Golay Filter

⁷ Second Derivative + Savitzky-Golay Filter

رقومی ارتفاع در محیط از نرم‌افزار SAGA GIS نسخه ۷/۳ استخراج شدند. علاوه بر این، پارامترهای سنجش از دور از تصاویر ماهواره‌ای سنتینل ۲ که دارای پوشش ابری کمتر از ۹ درصد بودند، برای ماه‌های مرداد، شهریور و مهر سال ۱۴۰۱ استخراج گردیدند.



شکل ۲. میانگین منحنی‌های خام بازتاب

۲-۵. مجموعه داده‌ها

داده‌ها به صورت تصادفی به دو بخش تقسیم شدند: ۸۰ درصد برای آموزش مدل‌ها و ۲۰ درصد برای ارزیابی عملکرد آن‌ها، به گونه‌ای که امکان سنجش دقت مدل فراهم شود. اهمیت نسبی متغیرهای محیطی در مدل‌های یادگیری ماشین با بهره‌گیری از تابع اهمیت متغیر (varImp) از بسته caret بررسی شد. به منظور افزایش دقت و پایداری نتایج، فرآیند برازش مدل‌ها با رویکرد ۱۰ گام و ۱۰ مرتبه انجام شد.

۲-۶. مدل‌سازی مکانی

در این پژوهش، سه الگوریتم جنگل تصادفی (RF)، نزدیک‌ترین k-همسایگی (k-NN)، و درخت تصمیم تقویت‌شده با گرادیان (XGBoost) با استفاده از بسته‌های تخصصی Random Forest، caret، و xgbTree در نرم‌افزار R-Studio نسخه ۴.۱.۱ گُر^۱ (۲۰۱۵) کدنویسی شده و برای پیش‌بینی توزیع مکانی هدایت هیدرولیکی اشباع خاک مورد استفاده قرار گرفتند. شایان ذکر است که مدل‌سازی هدایت هیدرولیکی اشباع خاک در دو سناریو انجام شد: در رویکرد اول، مدل‌سازی صرفاً بر اساس متغیرهای محیطی (سنجش از دور و توپوگرافی) صورت گرفت، و در رویکرد دوم، متغیرهای محیطی انتخاب‌شده همراه با داده‌های طیف سنجی نزدیک و ویژگی‌های خاک به کار گرفته شدند.

۲-۷. ارزیابی عملکرد مدل‌ها

به منظور ارزیابی عملکرد هر یک از MLMs مورد استفاده از چهار شاخص آماری ضریب تبیین (R^2) و ریشه میانگین مربعات خطا (RMSE)، میانگین قدر مطلق خطا (MAE) و ریشه میانگین مربعات خطای نرمال شده (nRMSE) استفاده گردید.

$$RMSE = \sqrt{\frac{1}{n} \sum (P_i - O_i)^2} \quad \text{(رابطه ۱)}$$

¹ Core

$$R^2 = 1 - \frac{\sum (O_i - P_i)^2}{\sum (O_i - \bar{O}_i)^2} \quad \text{رابطه ۲}$$

$$nRMSE = \frac{RMSE_{(i)}}{\bar{O}_i} \quad \text{رابطه ۳}$$

$$MAE = \frac{1}{n} \sum_{i=1}^n (P_i - O_i) \quad \text{رابطه ۴}$$

که در آن P_i : مقادیر پیش‌بینی شده، O_i : مقادیر اندازه‌گیری شده، \bar{O}_i میانگین مقادیر اندازه‌گیری شده و n تعداد مشاهدات است.

۳. یافته‌های پژوهش

۳-۱. تجزیه و تحلیل آماری

آمار توصیفی هدایت هیدرولیکی اشباع (عمق ۰-۱۰ سانتی‌متری) برای ۱۵۰ نمونه خاک در جدول (۱) آورده شده است. این جدول شامل اطلاعاتی از قبیل میانگین، انحراف معیار، میانه، مقادیر حداقل، حداکثر و ضریب تغییرات می‌باشد. بر اساس جدول ۱، میانگین هدایت هیدرولیکی اشباع ۰/۰۴ سانتی‌متر بر ثانیه است. همچنین، مقادیر حداقل و حداکثر هدایت هیدرولیکی اشباع به ترتیب ۰/۰۰۰۲ و ۰/۱۲ سانتی‌متر بر ثانیه است. بر اساس طبقه‌بندی، متغیرهایی با ضریب تغییرات کمتر از ۱۵ درصد در گروه تغییرات کم، متغیرهای با ضریب تغییرات بین ۱۵ تا ۳۵ درصد در گروه تغییرات متوسط، و متغیرهایی با ضریب تغییرات بیش از ۳۵ درصد در گروه تغییرات زیاد قرار می‌گیرند (Wilding & Dress, 1983). ضریب تغییرات برای هدایت هیدرولیکی اشباع (K_s) در کلاس تغییرپذیری بالا قرار دارد. ضریب تغییرات یک معیار نسبی از تغییرپذیری است و بر اساس این طبقه‌بندی، ضریب تغییرات هدایت هیدرولیکی با بالاترین ضریب تغییرات در کلاس تغییرپذیری بالا قرار دارد (Khosravani et al., 2023).

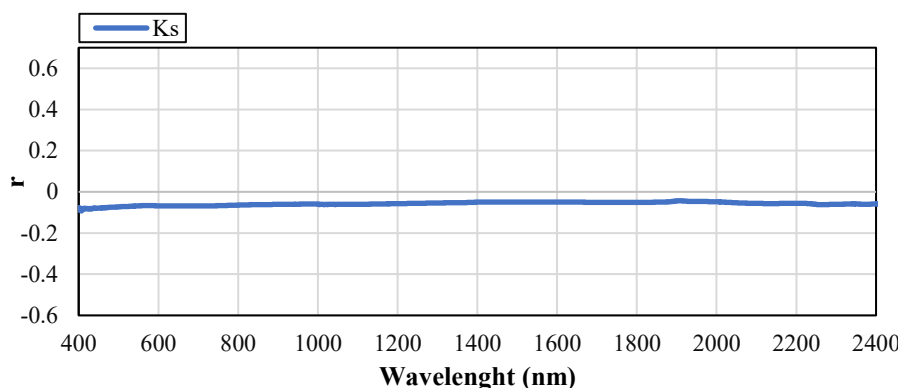
جدول ۱. آمار توصیفی هدایت هیدرولیکی اشباع خاک ($n=150$)

ویژگی‌های خاک	واحد	حداقل	حداکثر	میانگین	میانه	انحراف معیار	ضریب تغییرات (%)
هدایت هیدرولیکی اشباع (K_s)	cm/s	0.02×10^{-2}	۰/۱۲	۰/۰۴	۰/۰۴	۰/۰۳	۰/۷۷

۳-۲. تجزیه و تحلیل و ارزیابی ضریب همبستگی پیرسون بین ویژگی‌های خاک و داده‌های طیفی

در شکل ۳ ضریب همبستگی پیرسون (r) بین هدایت هیدرولیکی اشباع خاک با مقادیر بازتاب طیفی در هر طول موج، از گستره ۲۴۵۰-۴۰۰ نانومتر ارائه شده است. با توجه به این شکل، بین ویژگی هدایت هیدرولیکی اشباع خاک اندازه‌گیری شده با مقادیر بازتاب طیفی در گستره مرئی-مادون‌قرمز نزدیک (VIS-NIR) همبستگی بالایی وجود دارد. بالاترین همبستگی معنی‌دار هدایت هیدرولیکی اشباع خاک نیز در طول موج ۳۶۶ معادل ۰/۱۲- مشاهده شد (شکل ۳). این نتیجه نشان‌دهنده تأثیر قابل توجهی است که ویژگی‌های فیزیکی خاک بر هدایت هیدرولیکی دارد. به‌ویژه، در مطالعات مختلف، همبستگی‌های مثبت و منفی بین هدایت هیدرولیکی و ویژگی‌های دیگر خاک مانند تخلخل و جرم مخصوص ظاهری مشاهده شده است (Espeby., 1990). نتایج حاصل از تحلیل ضریب همبستگی پیرسون در شکل ۳ نشان می‌دهد که ارتباط بین هدایت هیدرولیکی اشباع خاک (K_s) و مقادیر بازتاب طیفی در محدوده طول‌موج‌های ۴۰۰ تا ۲۴۰۰ نانومتر نسبتاً ضعیف است. این همبستگی پایین را می‌توان به چندین عامل نسبت داد. نخست، هدایت هیدرولیکی اشباع خاک عمدتاً تحت تأثیر ویژگی‌های فیزیکی مانند بافت، ساختار و تراکم خاک قرار دارد که لزوماً از طریق بازتاب طیفی قابل شناسایی نیستند. دوم، بازتاب

طیفی خاک بیشتر به ترکیبات سطحی مانند ماده آلی، رطوبت و مواد معدنی وابسته است، در حالی که هدایت هیدرولیکی اشباع بیشتر به خصوصیات عمقی خاک مرتبط است. همچنین، شرایط محیطی و نویز موجود در داده‌های طیفی می‌تواند بر دقت اندازه‌گیری‌ها تأثیر بگذارد و باعث کاهش میزان همبستگی شود. با توجه به این محدودیت‌ها بهره‌گیری از مدل‌های یادگیری ماشین می‌تواند به کشف روابط غیرخطی بین بازتاب طیفی و هدایت هیدرولیکی اشباع کمک کند.



شکل ۳. ضریب همبستگی پیرسون بین هدایت هیدرولیکی اشباع خاک با مقادیر بازتاب طیفی خاک در طول موج‌های مختلف

۳-۳. متغیرهای محیطی منتخب و اهمیت آن‌ها

بر اساس جدول ۲ از مجموع ۴۵ متغیر کمکی، ۱۷ متغیر مربوط به منابع خاک و عوامل محیطی انتخاب گردیدند. از این میان ۱۰ متغیر کمکی متعلق به عوامل محیطی بر اساس روش "جنگل تصادفی-حذف ویژگی برگشتی" (RF-RFE) انتخاب گردیدند. این متغیرها شامل چهار شاخص (Greenness, Brightness, Band 12, NDVI) حاصل از نسبت‌گیری باندهای ماهواره Sentinel 2 و شش پارامتر (WE, PLC, CHNBL, CHND, RSP, Valley depth) از مدل رقومی ارتفاع بودند. همچنین هفت پارامتر خاک شامل (بافت خاک (درصد رس، شن، سیلت)، کربن آلی خاک، آهنک، جرم مخصوص ظاهری و میانگین هندسی قطر خاکدانه‌ها) نیز بر اساس نظر کارشناس و اهمیت آن‌ها در پیش‌بینی هدایت هیدرولیکی اشباع انتخاب شدند. بر اساس جدول ۲، نتایج نشان می‌دهد که نمایندگان فاکتورهای خاکسازای شامل خاک (s)، از نظر فراوانی، سهم بیشتری نسبت به متغیرهای محیطی حاصل از داده‌های سنجش از دور (o) و توپوگرافی (r) بر اساس معادله اسکورپن (scorpan) داشته‌اند.

۳-۴. عملکرد مدل‌سازی

در جدول ۳، مدل‌سازی برای هدایت هیدرولیکی اشباع خاک (Ks) در دو سناریو انجام شد. در سناریو اول، مدل‌سازی صرفاً بر اساس متغیرهای محیطی انجام شد، در حالی که در سناریو دوم، متغیرهای محیطی انتخاب‌شده با داده‌های طیفی و داده‌های خاک ترکیب شدند. عملکرد سه مدل یادگیری ماشین (RF، XGBoost، k-NN) برای پیش‌بینی Ks در هر دو سناریو مقایسه شده است. نتایج نشان داد که در سناریو اول، مدل RF عملکرد بهتری نسبت به دو مدل دیگر داشت، به طوری که مقدار R^2 آن برابر با ۰/۵۷ و مقادیر RMSE و nRMSE به ترتیب برابر با ۰/۰۱ و ۰/۲۵ بودند. با این حال، عملکرد مدل XGBoost در سناریو دوم به دلیل استفاده از داده‌های ترکیبی بهبود یافت و مقادیر R^2 ، RMSE و nRMSE به ترتیب به ۰/۶۵، ۰/۰۱ و ۰/۲۵ افزایش یافتند. این نشان‌دهنده توانایی بالای این مدل در استفاده از اطلاعات متنوع برای بهبود پیش‌بینی‌ها است. مدل k-NN، که ساختاری ساده‌تر دارد، در هر دو سناریو

پایین‌ترین عملکرد را نشان داد. مقدار R^2 این مدل در سناریو اول برابر با ۰/۵۳ و در سناریو دوم برابر با ۰/۶۲ گزارش شد.

جدول ۲. متغیرهای انتخاب شده برای پیش‌بینی ویژگی هدایت هیدرولیکی خاک منطقه مورد مطالعه

منبع	اختصار	پارامتر	فاکتور خاک‌سازی در مدل اسکورپین
DEM	Diffuse	شدت تابش پخشیده Diffuse insolation	توپوگرافی (r)
DEM	wetness	شاخص خیسگی wetness index	r-c
DEM	RSP	Relative. Slope. Position موقعیت نسبی شیب	r
DEM	CHND	فاصله تا شبکه آبراهه Channel Network Distance	r
DEM	VD	عمق دره (m) Valley depth	r
DEM	CHNBL	Channel network base level سطح پایه شبکه آبراهه	r-c
DEM	WE	Wind.Effect شاخص اثر باد	r-c
DEM	PLC	Plan curvature انحنای سطح	r
RS	NDVI	Normalized diff. vegetation index = $\frac{(NIR-RED)}{(NIR+RED)}$	ارگانسیم (O)
RS	Greenness	شاخص سبزیگی	
RS	Brightness	شاخص روشنایی	o-s
RS	B12	باند ۱۲	s
خاک (Soil)	CCE	آهک	
خاک (Soil)	GMD	میانگین هندسی قطر خاکدانه‌ها	
خاک (Soil)	BD (g.cm-3)	جرم مخصوص ظاهری	
خاک (Soil)	Silt (%)	سیلت	(s)
خاک (Soil)	SOC (%)	کربن آلی خاک	
خاک (Soil)	Sand (%)	شن	
خاک (Soil)	Clay (%)	رس	
دستگاه اسپکت (Spectrometer)	(Spectroscopy)	داده‌های طیف‌سنجی	

مدل XGBoost به دلیل ساختار پیچیده و قابلیت شناسایی الگوهای غیرخطی بهترین عملکرد را ارائه داده است. مدل RF نیز به‌عنوان مدلی با عملکرد قوی، به‌ویژه در داده‌های ترکیبی سناریو دوم، نتایج مناسبی داشت. مدل k-NN به دلیل ساختار ساده‌تر و حساسیت کمتر به داده‌های متنوع، ضعیف‌ترین عملکرد را نشان داد.

مطالعات پیشین نیز برتری XGBoost را تأیید می‌کنند. به‌عنوان نمونه، در مطالعه زایتفوجل^۱ و همکاران (۲۰۲۳)، الگوریتم XGBoost در پیش‌بینی هدایت هیدرولیکی اشباع (Ks) عملکرد بهتری نسبت به الگوریتم شبکه عصبی پیش‌خور (FNN) داشت. همچنین در تحقیق لی^۲ و همکاران (۲۰۲۴)، XGBoost بهترین مدل برای پیش‌بینی درصد رطوبت بهینه خاک با $R^2=0/9234$ معرفی شد. مدل RF نیز عملکرد قابل قبولی در پیش‌بینی مقادیر رطوبت نشان داد ($R^2=0/9198$ ، $RMSE=1/3605$). علاوه بر این، طبق تحقیقات رضایی^۳ و همکاران (۲۰۲۳)، روش جنگل تصادفی (RF) همراه با رویکرد نمونه‌برداری بهینه، نقشه‌های دقیقی از هدایت هیدرولیکی اشباع (Ks) خاک ارائه کرده است. در نهایت، نتایج این تحقیق نشان می‌دهد که ترکیب داده‌های طیفی و خاک با متغیرهای محیطی در سناریو دوم،

¹ Zeitfogel

² Li

³ Rezaei

عملکرد مدل‌های یادگیری ماشین را بهبود داده است. عملکرد هر مدل نیز به عواملی مانند پیچیدگی ساختار، تنظیمات بهینه و توانایی آن‌ها در شناسایی الگوهای مکانی وابسته است (Mosleh, 2016; Seppey, 2019).

مطالعات اخیر اهمیت استفاده از مدل‌های یادگیری ماشین پیشرفته مانند XGBoost در پیش‌بینی ویژگی‌های خاک را برجسته کرده‌اند. به‌عنوان مثال، مطالعه زارعی^۱ و همکاران (۲۰۲۱) نشان داد که مدل XGBoost در پیش‌بینی ویژگی‌های خاک مانند شوری و هدایت الکتریکی، با استفاده از داده‌های ماهواره‌ای Sentinel-2، عملکرد بهتری نسبت به مدل‌های دیگر از جمله Random Forest و Gradient Boosting Machine داشته است. این مدل توانست با دقت بالایی (R^2 بیش از ۷۶٪) نتایج را ارائه دهد. در تحقیق دیگری، وانگ^۲ و همکاران (۲۰۲۴) از مدل XGBoost برای پیش‌بینی کربن آلی خاک (SOC) با استفاده از داده‌های چندمنبعی شامل تصاویر ماهواره‌ای Sentinel-1 و Sentinel-2 و داده‌های توپوگرافی استفاده کردند. نتایج این تحقیق نشان داد که ترکیب داده‌های چندمنبعی و به‌کارگیری مدل XGBoost، دقت پیش‌بینی را تا ۳۶ درصد افزایش داده است. این یافته‌ها تأکید می‌کنند که کیفیت و تنوع داده‌های ورودی، به‌ویژه داده‌های طیفی و مکانی، نقش کلیدی در بهبود عملکرد مدل‌های یادگیری ماشین دارند. استفاده از داده‌های ترکیبی، از جمله داده‌های طیفی، توپوگرافی و مکانی، می‌تواند دقت مدل‌های یادگیری ماشین را به‌طور قابل‌توجهی افزایش دهد. برای مثال، حسینی^۳ (۲۰۲۳) نشان داد که داده‌های ماهواره‌ای Sentinel-1 و Sentinel-2 برای پیش‌بینی ویژگی‌های خاک نتایج دقیقی ارائه کرده‌اند. همچنین، مزید^۴ و همکاران (۲۰۲۲) بر نقش شاخص‌های طیفی و ویژگی‌های توپوگرافی به‌عنوان داده‌های ورودی با کیفیت بالا در بهبود عملکرد مدل‌ها تأکید کرده‌اند. یکی از مزایای کلیدی مدل XGBoost استفاده از درخت‌های تصمیم متعدد و رویکرد تقویت گرادیانی است که این مدل را قادر می‌سازد روابط پیچیده و غیرخطی بین متغیرهای ورودی و هدف را شناسایی و مدل‌سازی کند (Zarei et al., 2021). این ویژگی باعث شده است که XGBoost در مقایسه با سایر مدل‌ها، برای تحلیل داده‌های چندمنبعی و پیش‌بینی ویژگی‌های خاک، عملکرد برتری از خود نشان دهد. در مجموع، این یافته‌ها بیانگر این موضوع هستند که ترکیب داده‌های چندمنبعی و استفاده از مدل‌های پیشرفته‌ای مانند XGBoost، ابزار قدرتمندی برای پیش‌بینی دقیق ویژگی‌های خاک فراهم می‌آورد و می‌تواند در مطالعات مدیریت پایدار منابع خاک نقش مؤثری ایفا کند.

جدول ۳. نتایج اعتبارسنجی پیش‌بینی مکانی هدایت هیدرولیکی اشباع خاک

پارامترهای اندازه‌گیری شده	مدل‌های یادگیری ماشین	RMSE	R^2	MAE	nRMSE
هدایت هیدرولیکی اشباع (K_s) (سناریو ۱)	XGBoost	۰/۰۱	۰/۵۴	۰/۰۳	۰/۲۵
	RF	۰/۰۱	۰/۵۷	۰/۰۳	۰/۲۵
	k-NN	۰/۰۱	۰/۵۳	۰/۰۳	۰/۲۵
هدایت هیدرولیکی اشباع (K_s) (سناریو ۲)	XGBoost	۰/۰۱	۰/۶۵	۰/۰۳	۰/۲۵
	RF	۰/۰۱	۰/۶۳	۰/۰۳	۰/۲۵
	k-NN	۰/۰۱	۰/۶۲	۰/۰۳	۰/۲۵

۳-۵. اهمیت متغیرهای کمکی

اهمیت متغیرهای کمکی استفاده شده برای پیش‌بینی هدایت هیدرولیکی اشباع خاک در شکل ۴ نمایش داده شده است. اهمیت نسبی متغیرها بخشی از فرآیند مدل‌سازی است که توسط مدل Xgboost انجام شده و بیانگر میزان تاثیر و نقش متغیرهای محیطی در پیش‌بینی

¹ Zarei

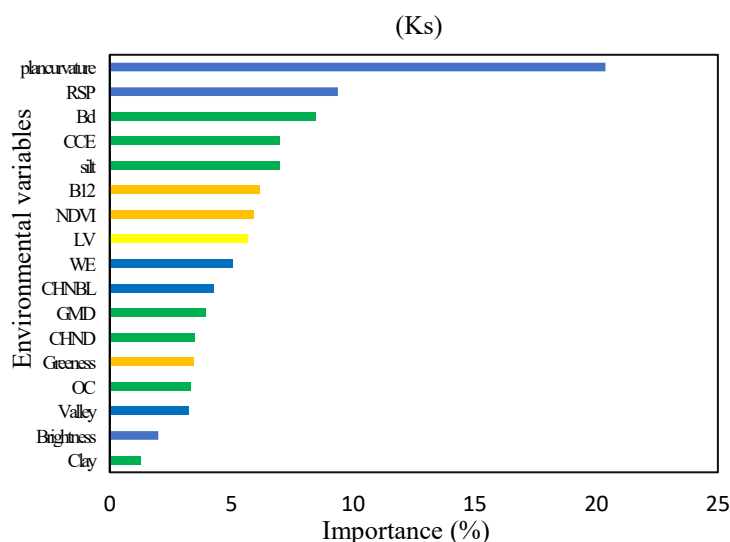
² Wang

³ Hosseini

⁴ Mzid

متغیرهای هدف می‌باشد. متغیرهای سطح پایه شبکه آبراهه (CHNBL)، موقعیت نسبی شیب (RSP)، اثر باد (WE)، شاخص نرمال شده تفاوت پوشش گیاهی (NDVI)، بازتاب طیفی خاک (LV) و سیلت (Silt)، رس (Clay)، کربن آلی خاک (OC)، جرم مخصوص ظاهری خاک (BD)، میانگین هندسی قطر خاکدانه‌ها (GMD)، کربنات کلسیم (CCE) به عنوان پیش‌بینی کننده‌های مهم تعیین شدند. متغیر انحنای سطح (PLC) حدود ۲۰/۳۸٪ از تغییرات هدایت هیدرولیکی اشباع را نشان داد. همچنین متغیر موقعیت نسبی شیب (RSP) ۹/۳۸٪ از تغییرات هدایت هیدرولیکی اشباع را نشان داد. جرم مخصوص ظاهری (Bd) حدود ۸/۴۵٪، کلسیم کربنات ۶/۹۸٪، درصد سیلت ۶/۹۷٪ از تغییرات هدایت هیدرولیکی اشباع خاک را نشان داد. باند ۱۲ حدود ۶/۱۵٪ و شاخص پوشش گیاهی تفاوت نرمال شده ۵/۹ درصد از تغییرات هدایت هیدرولیکی اشباع را نشان داد.

ویژگی‌های مرتبط با خاک، داده‌های طیف‌سنجی در دامنه‌های مرئی-مادون قرمز نزدیک و میانی و توپوگرافی به‌عنوان موثرترین متغیرهای کمکی برای پیش‌بینی متغیرهای هدف در منطقه مورد مطالعه شناخته شدند. به‌طور کلی، متغیرهای کمکی وابسته به خاک بدلیل همبستگی بالایی که با متغیرهای هدف دارند و همچنین پارامترهای توپوگرافی به‌عنوان یکی دیگر از عوامل شکل‌گیری خاک از مهمترین پیش‌بینی کننده‌های کلیدی برای نقشه‌برداری رقومی خاک استفاده می‌شوند (Mousavi et al., 2022).



شکل ۴. اهمیت نسبی متغیرهای کمکی برای پیش‌بینی هدایت هیدرولیکی اشباع خاک بر اساس مدل Xgboost، به ترتیب از بالا به پایین (انحنای سطح (Plan curvature)، موقعیت نسبی شیب (RSP)، جرم مخصوص ظاهری (Bd)، کلسیم کربنات (CCE)، سیلت (Silt)، باند ۱۲ (B12)، شاخص پوشش گیاهی تفاوت نرمال شده (NDVI)، داده‌های طیف سنجی نزدیک خاک (LV)، شاخص اثر باد (WE)، سطح پایه شبکه آبراهه (CHNBL)، میانگین هندسی قطر ذرات (GMD)، فاصله تا شبکه آبراهه (CHND)، شاخص سبزیگی (Greeness)، کربن آلی (OC)، عمق دره (VD)، شاخص روشنایی (Brightness)، درصد رس (Clay))

۳-۶. تغییرپذیری مکانی

تغییرات مکانی هدایت هیدرولیکی اشباع خاک (Ks) که توسط مدل درخت تصمیم تقویت‌شده با گرادیان (XGBoost) برآورد شده است، در شکل ۵ نمایش داده شده است. این مدل به دلیل دقت بالاتر آن در مقایسه با سایر روش‌های یادگیری ماشین انتخاب شده است. همان‌طور که مشاهده می‌شود، مقادیر Ks در منطقه مورد مطالعه دارای تغییرات قابل توجهی است و در محدوده‌ای بین ۰/۰۰۶ تا ۰/۱۳۷ سانتی‌متر بر ثانیه متغیر است.

بیشترین مقادیر Ks (محدوده رنگ آبی، سبز تیره و روشن) در مناطق کم‌شیب یا دشت‌های پایین دست مشاهده می‌شود. این نواحی به

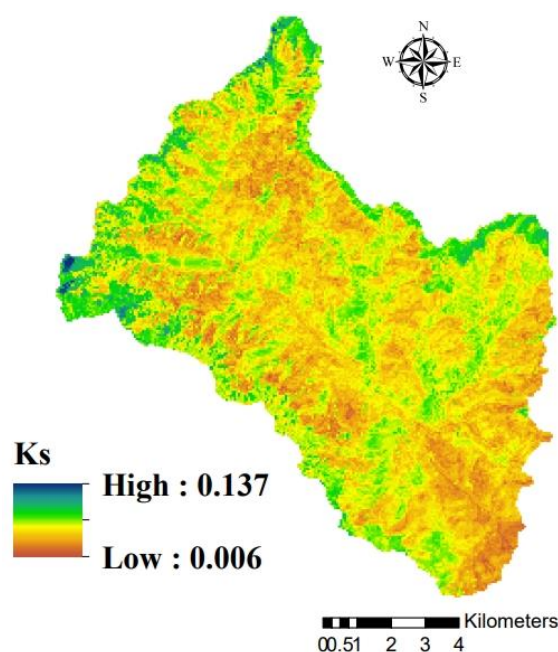
دلیل شرایط توپوگرافی خاص خود، تجمع آب بیشتری را تجربه می‌کنند و معمولاً خاک این مناطق دارای ساختار بهتری با تراکم کمتر است که نفوذپذیری بالاتری را فراهم می‌کند. همچنین، در این مناطق ممکن است فعالیت‌های کشاورزی گسترده‌تر باشد که می‌تواند شرایط فیزیکی خاک را بهبود بخشد.

کمترین مقادیر K_s (محدوده رنگ نارنجی و زرد) در شیب‌های بالادست و مناطق مرتفع مشاهده می‌شود. در این مناطق، شدت رواناب سطحی بیشتر و میزان نفوذپذیری خاک کمتر است. علت این امر می‌تواند وجود خاک‌های کم‌عمق و متراکم در مناطق شیب‌دار باشد که در معرض فرسایش بیشتری قرار دارند. این نواحی اغلب به دلیل توپوگرافی نامناسب، برای کشاورزی یا دیگر فعالیت‌های انسانی کمتر استفاده می‌شوند و کاربری غالب آن‌ها ممکن است شامل مراتع یا پوشش‌های طبیعی باشد.

علاوه بر این، اثر توپوگرافی بر توزیع مکانی K_s از طریق شاخص‌هایی مانند شیب، ارتفاع و شاخص خیسایی (Wetness Index) قابل بررسی است. مناطق کم‌شیب با شاخص خیسایی بالاتر، شرایط مناسب‌تری برای نگهداری و حرکت آب در خاک فراهم می‌کنند. برعکس، در مناطق پرشیب که جریان آب سریع‌تر اتفاق می‌افتد، هدایت هیدرولیکی کاهش می‌یابد (Sinai et al., 1981).

از منظر کاربری زمین، مناطقی که بیشترین مقادیر K_s را دارند، معمولاً در اراضی زراعی یا باغ‌ها دیده می‌شود، زیرا این نوع کاربری می‌تواند از طریق شخم‌زنی و مدیریت بهینه خاک، تراکم خاک را کاهش داده و نفوذپذیری را افزایش دهد. در مقابل، کمترین مقادیر K_s بیشتر در مناطق با پوشش طبیعی یا مراتع مشاهده می‌شود که به دلیل فرسایش یا کمبود فعالیت‌های انسانی، شرایط فیزیکی خاک در آن‌ها کمتر مساعد است (Archer and Smith, 1972).

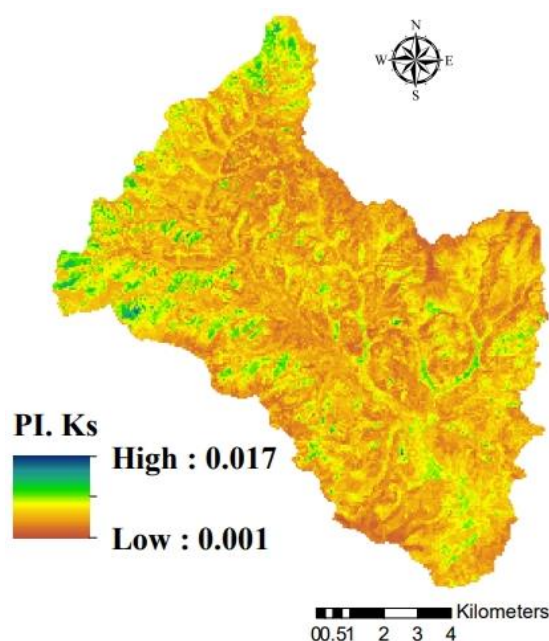
مدل XGBoost به دلیل توانایی بالای خود در شناسایی روابط پیچیده و غیرخطی بین متغیرهای محیطی مانند توپوگرافی و ویژگی‌های خاک انتخاب شد. استفاده از این مدل در مطالعات مختلف نشان داده است که XGBoost عملکرد بهتری در پیش‌بینی ویژگی‌های خاک، از جمله K_s دارد (Zarei et al., 2021; Wang et al., 2024). ترکیب داده‌های محیطی شامل ویژگی‌های توپوگرافی، داده‌های سنجش از دور و داده‌های طیفی در مدل XGBoost موجب بهبود قابل توجه دقت نتایج شده است. به‌ویژه، استفاده از شاخص‌های مرتبط با شیب و ویژگی‌های فیزیکی خاک در این مدل باعث شد که نقشه توزیع مکانی K_s با دقت بالایی تولید شود.



شکل ۵. توزیع مکانی K_s با استفاده از مدل XGBoost

۳-۷. برآورد عدم قطعیت

نقشه عدم قطعیت کمی شده با استفاده از روش bootstrapping بر اساس فاصله پیش‌بینی ۹۰ درصدی ارزیابی شد. نتایج این روش نشان داد که حداقل سطح عدم قطعیت در بخش‌های مرکزی منطقه مورد مطالعه مشاهده می‌شود. بر اساس نتایج موجود در شکل ۶ از دلایل پایین بودن میزان عدم قطعیت در این بخش مرکزی منطقه می‌توان به تراکم بیشتر نقاط مشاهداتی و همبستگی بالای متغیرهای کمکی با مدل‌های یادگیری ماشین مورد استفاده مرتبط دانست که این نتایج با دستاوردهای سایر محققین از قبیل بالتن‌سویلر^۱ و همکاران (۲۰۲۱) همراستا است که مبنی بر تاثیر تراکم داده مشاهداتی بر عدم قطعیت مدل پیش‌بینی کننده برای سایر ویژگی‌های خاک نیز مطابقت دارد. هدایت هیدرولیکی اشباع (K_s) در نواحی شمال و غرب منطقه مقادیر بالاتری از عدم قطعیت مشاهده می‌گردد. این موضوع به نظر می‌رسد که به دلیل پیچیدگی شرایط رسوب‌گذاری و تنوع اشکال فیزیوگرافی (شامل: مخروط افکنه و دشت دامنه‌ای) در این ناحیه باشد که این نتیجه نیز در مطالعه مشابهی توسط (Lalitha et al., 2021) در نواحی با پیچیدگی فیزیوگرافی گزارش گردیده است. به‌طور کلی، روش bootstrapping به دلیل دقت بالایی که در تخمین عدم قطعیت ویژگی‌های هیدرولیکی خاک دارد (Rezaee et al., 2023 و رحمانی و همکاران، ۱۴۰۱) شناخته می‌شود.



شکل ۶. فاصله پیش‌بینی ۹۰٪ مربوط به ویژگی‌های فیزیکی و مکانی خاک توسط مدل RF با استفاده از روش bootstrapping.

۴. بحث و نتیجه‌گیری

این تحقیق با هدف بررسی تغییرات مکانی رقمی ویژگی‌های خاک، از جمله هدایت هیدرولیکی اشباع (K_s)، با استفاده از سه مدل یادگیری ماشین (RF، XGBoost، و k-NN) در حوزه آبخیز کیلانه در استان کردستان انجام شد. نتایج اعتبارسنجی مدل‌ها نشان داد که مدل XGBoost در سناریو ۲ (ویژگی‌های خاک، داده‌های طیف‌سنجی نزدیک، توپوگرافی و سنجش از دور) عملکرد بسیار خوبی در پیش‌بینی هدایت هیدرولیکی اشباع (K_s) داشت.

¹ Baltensweiler

از بین متغیرهای انتخاب شده، ویژگی‌های خاک از جمله BD، CCE، Silt و ویژگی‌های توپوگرافی، به ویژه PLC و RSP همراه با شاخص طیف سنجی نزدیک LV به عنوان مهمترین عوامل مؤثر در پیش‌بینی ویژگی هدایت هیدرولیکی اشباع خاک معرفی شدند. از میان این ویژگی‌ها، ویژگی‌های توپوگرافی بیشترین تأثیر را داشتند.

تحلیل عدم قطعیت نقشه‌ی پیش‌بینی به تأیید دقت و قابلیت اعتماد مدل‌های استفاده شده در پیش‌بینی هدایت هیدرولیکی اشباع خاک کمک کرد. به‌طور کلی، این مطالعه نشان می‌دهد که به دلیل ماهیت زمان‌بر و پرهزینه اندازه‌گیری ویژگی هدایت هیدرولیکی اشباع خاک، همچنین تغییرات مکانی بالای آن‌ها در سراسر زمین‌نماها، ترکیبی از داده‌های توپوگرافی، طیف‌های Vis-NIR، الگوریتم‌های یادگیری ماشین، متغیرهای محیطی با وضوح مکانی بالا و داده‌های خاک می‌تواند برای ارزیابی و نقشه‌برداری بهینه این ویژگی‌های خاک در حوزه آبخیز کیلانه در غرب ایران مورد استفاده قرار گیرد. این نقشه می‌تواند به کشاورزان و بهره‌برداران در مدیریت بهینه منابع آبی، انتخاب روش‌های آبیاری مناسب و افزایش بهره‌وری محصولات کشاورزی کمک کند.

سپاس‌گزاری

نویسندگان این مقاله از حوزه معاونت پژوهشی دانشگاه شهید چمران اهواز جهت تامین هزینه این تحقیق (قرارداد پژوهانه SCU.AS1.1401.97) تشکر و قدردانی می‌کند.

References

- Aguilera, H., Guardiola-Albert, G., Merino, L. M., Baquedano, C., Díaz-Losada, E., Ardila, P. A. R., & Valsero, J. D. (2022). Building inexpensive topsoil saturated hydraulic conductivity maps for land planning based on machine learning and geostatistics. *Catena*, 208(2), 105788. <https://doi.org/10.1016/j.catena.2021.105788>.
- Babaeian, E., Homaei, M., Vereecken, H., Montzka, C., Norouzi, A. A., & van Genuchten, M. T. (2015). A comparative study of multiple approaches for predicting the soil–water retention curve: hyperspectral information vs. basic soil properties. *Soil Science Society of America Journal*, 79, 1043-8501. <https://doi.org/10.2136/sssaj2014.09.0355>.
- Baltensweiler, A., Walthert, L., Hanewinkel, M., Zimmermann, S., & Nussbaum, M. (2021). Machine learning based soil maps for a wide range of soil properties for the forested area of Switzerland. *Geoderma Regional*, 27, e00437. <https://doi.org/10.1016/j.geodrs.2021.e00437>.
- Banaie, M. H. (1998). Soil moisture and temperature regimes map of Iran. *Soil and Water Research Institute*. Ministry of Agriculture, Tehran, Iran, 1sheet. (In Persian).
- Cheng, Y. Y., Gao, X. G., Liu, T. H., Li, L. X., Du, W., Hamad, A., & Wang, J. P. (2022). Effect of water content on strength of alluvial silt in The Lower Yellow River. *Water*, 14(20), 3231. <https://doi.org/10.3390/w14203231>.
- Dharumarajan, S., Lalitha, M., Gomez, C., Vasundhara, R., Kalaiselvi, B., & Hegde, R. (2022). Prediction of soil hydraulic properties using VIS-NIR spectral data in semiarid region of Northern Karnataka Plateau. *Geoderma Regional*, 28, e00475. <https://doi.org/10.1016/j.geodrs.2021.e00475>.
- Demattê, J., Sousa, A. A., Alves, M. C., Nanni, M. R., Fiorio, P. R., & Campos, R. C. (2008). Determining soil water status and other soil characteristics by spectral proximal sensing. *Geoderma*, 135, 179-195. <https://doi.org/10.1016/j.geoderma.2005.12.002>.
- Espeby, B. (1990). An analysis of saturated hydraulic conductivity in a forested glacial till slope. *Soil Science*, 150(2), 485-494.
- Farasati, M., Seyedian, M., & Fathaabadi, A. (2024). Predicting soil hydraulic conductivity using random forest, SVM, and LSSVM models. *Natural Resource Modeling*, 37(4), e12407.
- Gao, X. S., Yi, X. I. A. O., Deng, L. J., Li, Q. Q., Wang, C. Q., Bing, L. I., & Min, Z. E. N. G. (2019). Spatial variability of soil total nitrogen, phosphorus and potassium in Renshou County of Sichuan Basin, China. *Journal of Integrative Agriculture*, 18(2), 279-289. [https://doi.org/10.1016/S2095-3119\(18\)62069-6](https://doi.org/10.1016/S2095-3119(18)62069-6).
- Gomez, C., Philippe, L., & Guillaume, C. (2008). Continuum moval versus PLSR method for clay and calcium carbonate content estimation from laboratory and airborne hyperspectral measurements. *Geoderma*, 148, 14148. <https://doi.org/10.1016/j.geoderma.2008.09.016>.

- Hengl, T., Nussbaum, M., Wright, M. N., & Heuvelink, G. B. M. (2018). Random forest as a generic framework for predictive modeling of spatial and spatio-temporal variables. *PeerJ*, 6, e5518. <https://doi.org/10.7287/peerj.preprints.26693v3>.
- Homaei, M. & Farrokhi Firouzi, A. (2008). Deriving point and parametric pedotransfer functions of some gypsiferous soils. *Australian Journal of Soil Research*, 46, 219-227. <https://doi.org/10.1071/SR07161>.
- Hosseini, F. S., Seo, M. B., Razavi-Termeh, S. V., Sadeghi-Niaraki, A., Jamshidi, M., & Choi, S. M. (2023). Geospatial Artificial Intelligence (GeoAI) and Satellite Imagery Fusion for Soil Physical Property Predicting. *Sustainability*, 15(19), 14125.
- Janik, L. J., Merry, R. H. Forrester, S. T. Lanyon D. M. & Rawson, A. (2009). Rapid prediction of soil water retention using mid infrared spectroscopy. *Soil Science Society of America Journal*, 71(2), 507-514. <https://doi.org/10.2136/sssaj2005.0391>.
- Khosravani, P., Baghernejad, M., Moosavi, A. A. & Rezaei, M. (2023). Digital mapping and spatial modeling of some soil physical and mechanical properties in a semi-arid region of Iran. *Environmental Monitoring and Assessment*, 195(11), 1367. <https://doi.org/10.1007/s10661-023-11980-6>.
- Kim, I., Pullanagari, R. R., Deurer, M., Singh, R., Huh, K. Y., & Clothier, B. E. (2014). The use of visible and near-infrared spectroscopy for the analysis of soil water repellency. *European Journal of Soil Science*, 65, 360-368. <https://doi.org/10.1111/ejss.12138>.
- Klute, A. & Dirksen, C. (1986). Hydraulic Conductivity and Diffusivity: Laboratory Methods. In: Klute, A., Ed., *Methods of Soil Analysis*. Part 1: Physical and Mineralogical Methods, 2nd Edition, Agronomy Monograph No. 9, ASA, Madison, 687-734. <https://doi.org/10.2136/sssabookser5.1.2ed.c28>.
- Knotters, M., van Egmond, F. M., Bakker, G., Walvoort, D. J. J., & Brouwer, F. (2017). *A selection of sensing techniques for mapping soil hydraulic properties*. Wageningen Environmental Research, part of Wageningen UR, Wageningen.
- Kubiak, K., Spiralski, M., Pompeu, J., Levavasseur, V., & Wawer, R. (2024). Advances in remote sensing for monitoring soil conditions in forest ecosystems: techniques, challenges, and applications. *Transactions on Aerospace Research*. 277(34), 1-13. <https://doi.org/10.2478/tar-2024-0019>.
- Lalitha, M., Dharumarajan, S., Suputhra, A., Kalaiselvi, B., Hegde, R., Reddy, R. S., & Dwivedi, B. S. (2021). Spatial prediction of soil depth using environmental covariates by quantile regression forest model. *Environmental Monitoring and Assessment*, 193(10), 1-10.
- Li, H., Zhang, J., Yang, X., Ye, M., Jiang, W., Gong, J. & Xu, Z. (2024). Bayesian optimization based extreme gradient boosting and GPR time-frequency features for the recognition of moisture damage in asphalt pavement. *Construction and Building Materials*, 434, 136675. <https://doi.org/10.1016/j.conbuildmat.2024.136675>.
- Ma, D., Zhang, C., Li, T., & Feng, H. (2024). Novel calibration method for fine soil electrical resistivity based on van der Pauw configuration. *Measurement Science and Technology*, 36(1), 017007. <https://doi.org/10.1088/1361-6501/ad9626>.
- Mahmoudzadeh, H., Matinfar, H. R., Taghizadeh-Mehrdadi, R., & Kerry, R. (2020). Spatial prediction of soil organic carbon using machine learning techniques in western Iran. *Geoderma Regional*, 21, e00260. <https://doi.org/10.1016/j.geodrs.2020.e00260>.
- Mendes, W. D. S., Demattê, J. A. M., Barros, A. S. E., Salazar, D. F. U., & Amorim, M. T. A. (2020). Geostatistics or machine learning for mapping soil attributes and agricultural practices. *Revista Ceres*, 67(4), 330-336. <https://doi.org/10.1590/0034-737X202067040010>.
- Minasny, B., & McBratney, A. B. (2016). Digital soil mapping: A brief history and some lessons. *Geoderma*, 264, 301-311. <https://doi.org/10.1016/j.geoderma.2015.07.017>.
- Mousavi, S. R., Sarmadian, F., & Omid, M. (2022). Three-dimensional mapping of soil organic carbon using soil and environmental covariates in an arid and semi-arid region of Iran. *Measurement*. 201(13):111706. <https://doi.org/10.1016/j.measurement.2022.111706>.
- Mousavi, S. R., Sarmadian, F., Omid, M., & Bogaert, P. (2021). Digital Modeling of Three-Dimensional Soil Salinity Variation Using Machine Learning Algorithms in Arid and Semi-Arid lands of Qazvin Plain. *Iranian Journal of Soil and Water Research*, 52(7), 1915-1929. <https://doi.org/10.22059/ijswr.2021.323030.668957>. (In Persian).
- Mzid, N., Castaldi, F., Tolomio, M., Pascucci, S., Casa, R., & Pignatti, S. (2022). Evaluation of agricultural bare soil properties retrieval from Landsat 8, Sentinel-2 and PRISMA satellite data. *Remote Sensing*, 14(3), 714. <https://doi.org/10.3390/rs14030714>.
- Nanni, M. R., & Demattê, J. A. M. (2006). Spectral reflectance methodology in comparison to traditional soil analysis. *Soil Science Society of America Journal*, 70(2), 393-407. <https://doi.org/10.2136/sssaj2003.0285>.
- Nasta P. & Szabó, B. (2021). Evaluation of pedotransfer functions for predicting soil hydraulic properties: A voyage from regional to field scales across Europe. *Journal of Hydrology: Regional Studies*. 37:100903.
- Nazari, R., Ramezani Etedali, H., Nazari, B., & Collins, B. (2020). The impact of climate variability on water footprint components of rainfed wheat and barley in the Qazvin province of Iran. *Irrigation and Drainage*, 69(4), 826-843. <https://doi.org/10.1002/ird.2487>.

- Nguyen, T.T. (2021). Predicting agricultural soil carbon using machine learning. *Nature Reviews Earth & Environment*, 2(12), 825-825. <https://doi.org/10.1038/s43017-021-00243-y>.
- Ng, W., Minasny, B., Mendes, W.D.S., Demattê, J. A. M. (2020). The influence of training sample size on the accuracy of deep learning models for the prediction of soil properties with near-infrared spectroscopy data. *Soil*, 6, 565–578.
- Niede, R., & Benbi, D. K. (2022). Integrated review of the nexus between toxic elements in the environment and human health. *AIMS Public Health*, 9(4), 758. <https://doi.org/10.3934/publichealth.2022052>.
- Otoni, M. V., Teixeira, W. G., Reis, A. M. H. D., Pimentel, L. G., Souza, L. R., Albuquerque, J. A., ... & Curi, N. (2025). Saturated hydraulic conductivity and steady-state infiltration rate database for Brazilian soils. *Revista Brasileira de Ciência do Solo*, 49, e0240003. <https://doi.org/10.36783/18069657rbc20240003>.
- Padarian, J., Minasny, B., & McBratney, A. B. (2020). Machine learning and soil sciences: a review aided by machine learning tools. *Soil*, 6(1), 35-52. <https://doi.org/10.5194/soil-6-35-2020>, 2020.
- Páez-Bimos, S., Molina, A., Calispa, M., Delmelle, P., Lahuatte, B., Villacís, M., ... & Vanacker, V. (2023). Soil–vegetation–water interactions controlling solute flow and chemical weathering in volcanic ash soils of the high Andes. *Hydrology and Earth System Sciences*, 27(7), 1507-1529.
- Parsaie, F., Farrokhian Firouzi, A., Mousavi, S. R., Rahmani, A., Sedri, M. H., & Homae, M. (2021). Large-scale digital mapping of topsoil total nitrogen using machine learning models and associated uncertainty map. *Environmental Monitoring and Assessment*, 193, 1-15. <https://doi.org/10.1007/s10661-021-08947-w>.
- Rahmani, A., Sarmadian, F., & Arefi, H. (2022). Digital mapping of surface soil thickness and its associated uncertainty using machine learning approach in a part of arid and semi-arid lands of Qazvin Plain. *Iranian Journal of Soil and Water Research*, 53(3), 585-602. 10.22059/ijswr.2022.338007.669195. (In Persian).
- Rezaei, M., Mousavi, S. R., Rahmani, A., Zeraatpisheh, M., Rahmati, M., Pakparvar, M. & Cornelis, W. (2023). Incorporating machine learning models and remote sensing to assess the spatial distribution of saturated hydraulic conductivity in a light-textured soil. *Computers and Electronics in Agriculture*, 209, 107821. <https://doi.org/10.1016/j.compag.2023.107821>.
- Sabetzade, M., Gorji, M., Roudier, P., & Zolfaghari, A. A. (2021). Combination of MIR spectroscopy and environmental covariates to predict soil organic carbon in a semi-arid region. *Catena*, 196, 1-12. <https://doi.org/10.1016/j.catena.2020.104844>.
- Santra, P., Sahoo, R. N., Das, B. S., Samal, R. N., Pattanaik, A. K., & Gupta, V. K. (2009). Estimation of soil hydraulic properties using proximal spectral reflectance in visible, near-infrared, and shortwave-infrared (VIS–NIR–SWIR) region. *Geoderma*, 152(3-4), 338-349. <https://doi.org/10.1016/j.geoderma.2009.07.001>.
- Schneider, W. E., & Young, R. (1997). *Spectroradiometry methods. Handbook of Applied Photometry*, ed. Casimer De Cusatis, 252.
- Soil Survey Staff, (2022). *Keys to soil Taxonomy*. In: U.S. Department of Agriculture, Natural Resources Conservation Service, thirteenth ed. 18-77. Washington, DC.
- Taghizadeh-Mehrjardi, R., Minasny, B., Sarmadian, F., & Malone, B. P. (2014). Digital mapping of soil salinity in Ardakan region, central Iran. *Geoderma*, 213, 15–28. <https://doi.org/10.1016/j.geoderma.2013.07.020>.
- Viscarra Rossel, R. A. V. (2008). ParLeS: Software for chemometric analysis of spectroscopic data. *Chemometrics and Intelligent Laboratory Systems*, 90, 72–83. <https://doi.org/10.1016/j.chemolab.2007.06.006>.
- Wang, Z., Wu, W., & Liu, H. (2024). Spatial estimation of soil organic carbon content utilizing PlanetScope, Sentinel-2, and Sentinel-1 data. *Remote Sensing*, 16(17), 3268. <https://doi.org/10.3390/rs16173268>.
- Wilding, L. P., & Dress, L. R. (1983). In Application of geostatistics to spatial studies of soil. Eds. BB Trangmar, RS Yost and G Uehara. *Advances in Agr*, 38. [https://doi.org/10.1016/S0065-2113\(08\)60673-2](https://doi.org/10.1016/S0065-2113(08)60673-2).
- Xu, C. Xu, X., Liu, M. Liu, W. Yang, J., Luo, W., Zhang, R., & Kiely, G. (2017). Enhancing pedotransfer functions (PTFs) using soil spectral reflectance data for estimating saturated hydraulic conductivity in southwestern China. *Catena*, 158: 350–356. <https://doi.org/10.1016/j.catena.2017.07.014>.
- Zarei, A., Hasanlou, M., & Mahdianpari, M. (2021). A comparison of machine learning models for soil salinity estimation using multi-spectral earth observation data. *ISPRS annals of the photogrammetry, remote sensing and spatial information sciences*, 3, 257-263. <https://doi.org/10.5194/isprs-annals-V-3-2021-257-2021>.
- Zeitfogel, H., Feigl, M., & Schulz, K. (2022). Soil information on a regional scale: Two machine learning based approaches for predicting saturated hydraulic conductivity. *Geoderma*, 433, 116418. <https://doi.org/10.1016/j.geoderma.2023.116418>.