



Estimation of Field Capacity and Permanent Wilting Point using Visible-Near Infrared Spectral and Soil-Based Pedotransfer Functions

Chiman Mahdizadeh¹ | Hossein Bayat^{2*} | Masoud Davari³

1. Department of Soil Science and Engineering, Faculty of Agriculture, Bu-Ali Sina University, Hamedan, Iran. E-mail: chimanmehdizadeh@gmail.com
2. Corresponding Author, Department of Soil Science and Engineering, Faculty of Agriculture, Bu-Ali Sina University, Hamedan, Iran. E-mail: h.bayat@basu.ac.ir Other e-mail: hbayat2001@gmail.com.
3. Department of Soil Science, Faculty of Agriculture, University of Kurdistan, Sanandaj, Iran. E-mail: m.davari@uok.ac.ir.

Article Info

ABSTRACT

Article type: Research Article

Article history:

Received: Nov. 1, 2025

Revised: Dec. 29, 2025

Accepted: Feb. 14, 2026

Published online: March. 2026

Keywords:

Hydraulic properties,
Multiple linear regression,
Random Forest (RF),
Spectral reflectance

Soil water retention characteristics, such as field capacity (FC) and permanent wilting point (PWP), are critical for efficient water management in agriculture. However, direct measurement of these parameters at the field scale is not always practical. This study aimed to estimate FC and PWP using visible-near infrared (Vis-NIR) spectral data combined with soil physicochemical properties through random forest (RF) and multiple linear regression (MLR) models. A total of 130 soil samples were collected from five provinces in Iran, and their spectral and soil properties were measured. The dataset was divided into training (90 samples) and testing (40 samples) sets, and 11 pedotransfer functions (PTFs) were developed in three steps. To improve model performance, in addition to the no-preprocessing (NP) method, multiplicative scatter correction (MSC), first and second derivatives with Savitzky–Golay filtering (FD-SG, FD-SG2), and standard normal variate (SNV) were applied to the spectral data prior to model development. The results showed that the RF model (RMSE = 0.050) outperformed MLR (RMSE = 0.057) during the training stage. However, in the testing stage, no statistically significant difference was observed between the two methods in estimating FC. In contrast, for PWP estimation, RF generally yielded better results than MLR across most functions; however, these differences were not statistically significant, except for PTF2 (AIC = -264.3), where a significant difference was observed. During the training stage, PTF11 exhibited the best performance for FC estimation (AIC = -540.2), while PTF7 showed the highest performance for PWP estimation (AIC = -612.4). PTF3, incorporating sand, clay, and organic matter as input variables, was identified as the most accurate estimator of FC (AIC = -553.3). Similarly, PTF6, using sand, clay, organic matter, and total porosity, was identified as the most effective estimator of PWP (AIC = -616.2). Principal component analysis identified key wavelengths at 409 nm for FC and 1414, 1912, and 2150 nm for PWP. Overall, soil-property-based PTFs outperformed spectral-only models, but combining spectral and soil data with machine learning improved prediction accuracy.

Cite this article: Chiman, M. Z., Hossein, B., & Masoud, D. (2026). Estimation of field capacity and permanent wilting point using visible-near infrared spectral and soil-based pedotransfer functions, *Iranian Journal of Soil and Water Research*, 57 (1), 107-129. <https://doi.org/10.22059/ijswr.2026.405383.670041>

© The Author(s).

Publisher: University of Tehran Press.



DOI: <https://doi.org/10.22059/ijswr.2026.405383.670041>



EXTENDED ABSTRACT

Introduction

Soil moisture is one of the key factors influencing plant growth and development and is recognized as a fundamental driving force for the sustainable development of many terrestrial ecosystems. The importance of this factor has become even more pronounced under current conditions, where global water challenges and climate change have cast a shadow over the sustainability of natural resources. Changes in soil moisture can have significant effects on vegetation cover and on the physical and chemical properties of soils. The soil water retention curve (SWRC) describes the relationship between soil moisture content and matric suction. Among the most important points on the SWRC are the soil moisture contents at field capacity (FC) and at the permanent wilting point (PWP). The soil's water-holding capacity reflects the availability of soil water and the distribution of soil pore sizes, which are of great importance for studying soil water storage, conservation, movement, and supply. However, due to the difficulty, time-consuming nature, and high cost of direct measurement of these properties, the development of indirect and rapid methods for estimating these parameters—particularly through the use of pedotransfer functions (PTFs) and modern visible–near infrared (Vis–NIR) spectroscopy—has become both a scientific and practical necessity. Therefore, the objectives of this study were as follows: 1- Estimation of soil hydraulic parameters (FC and PWP) using a combination of soil physico-chemical properties and spectral reflectance in the visible-near infrared (Vis-NIR) region, and development and validation of RF and MLR models to enhance the accuracy and generalization of predictions across soils with diverse textures. 2- Evaluation of the effects of spectral preprocessing methods (MSC, FD-SG1, FD-SG2, and SNV) on improving the accuracy of models estimating FC and PWP parameters. 3- Identification of key wavelengths and soil variables influencing the variability of FC and PWP through Principal Component Analysis (PCA) and selected pedotransfer functions.

Method

In this study, a total of 130 disturbed and undisturbed soil samples were collected from five provinces of Iran to estimate the FC and PWP parameters. The physical, chemical, and hydraulic properties of the soils were measured, and their spectral reflectance was recorded using a spectroradiometer. The soil spectral curves in the visible to near-infrared range (350-2500 nm) were measured under standardized spectroscopic conditions in a darkroom environment. To improve the estimation accuracy, several spectral preprocessing techniques were applied to the reflectance data, including Multiplicative Scatter Correction (MSC), first derivative Savitzky–Golay filtering (FD-SG1), second derivative Savitzky–Golay filtering (FD-SG2), and Standard Normal Variate (SNV). Principal Component Analysis (PCA) was employed to reduce the dimensionality of the spectral data and to extract the most relevant features. For the development of pedotransfer functions (PTFs), eleven models based on different algorithms were constructed in three stages using various combinations of input variables to predict FC and PWP parameters. The datasets were used to develop regression- and Random Forest-based PTFs using STATISTICA 14 software. The entire dataset, consisting of both input and output variables, was divided into two subsets: 90 samples were randomly selected for model training and 40 samples for testing. This process was repeated ten times, and a new model was executed for each iteration. The mean results of the ten runs were reported as the final outcomes. Finally, the PTFs were developed using Multiple Linear Regression (MLR) and Random Forest (RF) methods to estimate the FC and PWP parameters.

Results

The results indicated that the Random Forest (RF) model outperformed the Multiple Linear Regression (MLR) model during the training phase. Among the soil and spectral datasets, the pedotransfer functions developed in the second stage (PTF3), which included sand, clay, and organic matter as input variables, performed as the best estimators of FC, showing a 3.5% improvement compared to the baseline model. Moreover, PTF6, which used sand, clay, organic matter, and total porosity as input variables, was identified as the most effective function for estimating PWP, with a 5% improvement over the baseline model. Comparative analysis among the PTFs in the third stage revealed that PTF11 (SNV) for FC and PTF7 (NP) for PWP exhibited the best performance during the training phase, significantly reducing the Akaike Information Criterion (AIC) values. Furthermore, Principal Component Analysis (PCA) identified key wavelengths at 1414, 1912, and 2150 nm for PWP, and at 409 nm for FC.

Conclusions

Overall, integrating spectral and soil data with machine learning algorithms, especially random forest, offers a more accurate and cost-effective approach compared to existing methods for estimating FC and PWP parameters and can significantly contribute to improving soil management.

Funding

The official and fluent translation of your text into English is as follows This study was conducted with the financial and moral support of the Vice-Chancellor for Research of Bu-Ali Sina University, Hamedan. Financial support for this research was provided by Bu-Ali Sina University, Faculty of Agriculture, through a student thesis grant for the first author and research grants for the other authors.

Authorship contribution

The authors participated in all stages of research design and execution, statistical data analysis, analysis and interpretation of information and results, report preparation, manuscript drafting, results review and control, correction,

revision, and finalization of the manuscript, and their roles in order of contribution are as follows: Chiman Mahdizadeh, Hossein Bayat and Masoud Davari.”

Declaration of Generative AI and AI-assisted technologies in the writing process

During the preparation of this work the author(s) used ChatGPT in order to assist with translation and language refinement. After using this tool, the author(s) reviewed and edited the content as needed and take(s) full responsibility for the content of the publication.

Data availability statement

The data from the present study are available from the authors upon reasonable request.

Acknowledgements

The authors would like to express their sincere appreciation to the Bu-Ali Sina University, Hamedan, for its financial and moral support of the present study.

The authors are grateful to the anonymous reviewers for their constructive scientific and structural comments.

Ethical considerations

This study did not involve human participants or animals, and no sensitive data were used. The authors followed standard research ethics.

Conflict of interest

The authors declare no conflict of interest.

برآورد رطوبت ظرفیت زراعی و نقطه پژمردگی دائم با استفاده از توابع انتقالی طیفی مرئی-مادون قرمز نزدیک و خاکی

چیمین مهدی زاده^۱ | حسین بیات^۲ | مسعود داوری^۳

۱. گروه علوم و مهندسی خاک، دانشکده کشاورزی، دانشگاه بو علی سینا، همدان، ایران. رایانامه:

chimanmehdizadeh@gmail.com

۲. نویسنده مسئول، گروه علوم و مهندسی خاک، دانشکده کشاورزی، دانشگاه بو علی سینا، همدان، ایران. رایانامه:

h.bayat@basu.ac.ir

۳. گروه علوم و مهندسی خاک، دانشکده کشاورزی، دانشگاه کردستان، سنندج، ایران. رایانامه: m.davari@uok.ac.ir

اطلاعات مقاله

چکیده

نوع مقاله: مقاله پژوهشی

تاریخ دریافت: ۱۴۰۴/۸/۱۰

تاریخ بازنگری: ۱۴۰۴/۱۰/۸

تاریخ پذیرش: ۱۴۰۴/۱۱/۲۵

تاریخ انتشار: فروردین ۱۴۰۵

واژه‌های کلیدی:

بازتاب طیفی،

جنگل تصادفی،

رگرسیون خطی چندگانه،

ویژگی‌های هیدرولیکی

ویژگی‌های نگه‌داری آب خاک، نظیر ظرفیت زراعی (FC) و نقطه پژمردگی دائم (PWP) نقش مهمی در مدیریت منابع آب در کشاورزی دارند. با این حال اندازه‌گیری مستقیم این پارامترها در مقیاس مزرعه همیشه عملی نیست. هدف این مطالعه، برآورد FC و PWP، با استفاده از داده‌های طیفی مرئی-مادون قرمز نزدیک (Vis-NIR) و ویژگی‌های فیزیکی-شیمیایی خاک از طریق روش‌های جنگل تصادفی (RF) و رگرسیون خطی چندگانه (MLR) بود. بدین منظور، ۱۳۰ نمونه خاک از ۵ استان ایران جمع‌آوری و ویژگی‌های خاکی و طیفی آنها اندازه‌گیری شد. داده‌ها به دو مجموعه آموزش (۹۰ نمونه) و تست (۴۰ نمونه) تقسیم شدند و ۱۱ تابع انتقالی در سه گام ایجاد شد. برای افزایش دقت مدل‌ها علاوه بر روش بدون پیش‌پردازش (NP)، از پیش‌پردازش‌های تصحیح پخشیده چندگانه (MSC)، مشتق اول و دوم همراه با فیلتر ساویتزکی-گلای (FD-SG, FD-SG2) و متغیر نرمال استاندارد (SNV) بر روی داده‌های طیفی استفاده شد. نتایج نشان داد مدل RF در مرحله آموزش با $RMSE = 0.050$ عملکرد بهتری نسبت به MLR با $RMSE = 0.057$ دارد. ولی در مرحله تست عملکرد دو روش در تخمین FC تفاوت معنی‌داری نداشت. این در حالی است که در تخمین PWP در غالب توابع (به جز تابع ۲ به صورت معنی‌دار $AIC = -264/3$) روش RF به طور غیرمعنی‌دار بهتر از MLR بود. در مرحله آموزش تابع انتقالی ۱۱ (PTF11) با $AIC = -540/2$ برای FC و $AIC = -612/4$ برای PWP بهترین عملکرد را داشت. PTF3 با متغیرهای شن، رس و ماده‌آلی به‌عنوان بهترین تخمین‌گر FC با $AIC = -553/3$ و PTF6 با ورودی‌های شن، رس، ماده‌آلی و تخلخل کل به‌عنوان موثرترین تخمین‌گر PWP با $AIC = -616/2$ شناسایی شدند. همچنین تحلیل مؤلفه‌های اصلی، طول‌موج‌های کلیدی ۱۴۱۴، ۱۹۱۲ و ۲۱۵۰ نانومتر را برای تخمین PWP و طول موج ۴۰۹ نانومتر را برای تخمین FC شناسایی کرد. نتایج این پژوهش نشان داد توابع مبتنی بر ویژگی‌های فیزیکی و شیمیایی خاک نسبت به داده‌های طیفی عملکرد بهتری داشتند، اما ترکیب داده‌های طیفی و خاکی همراه با الگوریتم‌های یادگیری ماشین، می‌تواند دقت مدل‌ها را بهبود بخشد.

استناد: مهدی زاده، چیمین؛ بیات، حسین؛ و داوری، مسعود (۱۴۰۵). برآورد رطوبت ظرفیت زراعی و نقطه پژمردگی دائم با استفاده از توابع انتقالی طیفی مرئی-مادون قرمز نزدیک و خاکی. *مجله تحقیقات آب و خاک ایران*، ۵۷ (۱)، ۱۰۷-۱۲۹.



<https://doi.org/10.22059/ijswr.2026.405383.670041>

© نویسندگان.

ناشر: مؤسسه انتشارات دانشگاه تهران.

DOI: <https://doi.org/10.22059/ijswr.2026.405383.670041>

مقدمه

رطوبت خاک یکی از عوامل کلیدی موثر بر رشد و نمو گیاهان و به عنوان نیروی محرکه‌ای اساسی برای توسعه پایدار بسیاری از اکوسیستم‌های زمینی شناخته می‌شود. اهمیت این موضوع در شرایط کنونی که چالش‌های جهانی آب و تغییرات اقلیمی بر پایداری منابع طبیعی سایه افکنده اند، دو چندان شده است. تغییرات در رطوبت خاک می‌تواند تأثیرات بسزایی بر پوشش گیاهی و خواص فیزیکی و شیمیایی خاک داشته باشد (Kaisermann et al., 2017). خاک به عنوان بستر اصلی رشد گیاهان، نقش تعیین کننده‌ای در تامین نیازهای آبی گیاهان ایفا می‌کند و میزان آب موجود در خاک یکی از مهمترین عوامل کنترل کننده عملکرد محصول به شمار می‌رود. پیش بینی دقیق مقدار آب خاک و مدیریت بهینه آن، این امکان را برای کشاورزان فراهم می‌سازد تا تصمیمات آگاهانه‌تری اتخاذ کرده و سودآوری را افزایش دهند (Castellini & Iovino., 2019). با توجه به این اهمیت، بررسی دقیق رفتار رطوبت خاک و عوامل تاثیرگذار بر آن، گامی اساسی در جهت بهبود مدیریت منابع آب و افزایش بهره‌وری کشاورزی است.

منحنی نگهداری آب خاک ابزاری حیاتی در درک رفتار آب خاک است و کاربردهای گسترده‌ای در حوزه‌های مختلف کشاورزی و محیط زیست دارد. این منحنی رابطه بین رطوبت خاک و مکش ماتریک را توصیف می‌کند (Sadeghi & Darzi, 2021). از جمله نقاط بسیار مهم منحنی نگهداری آب خاک میزان رطوبت موجود در ظرفیت زراعی (FC) و نقطه پژمردگی دائم (PWP) است. ظرفیت زراعی و نقطه پژمردگی دائم به عنوان شاخص‌های کلیدی خاک برای مطالعه نفوذپذیری، ظرفیت نگهداری آب خاک، زهکشی، آبیاری، تنش آبی گیاه و حرکت املاح مطرح هستند (Navidiet al., 2022). ظرفیت نگهداری آب خاک، در دسترس بودن آب خاک و توزیع اندازه منافذ خاک را منعکس می‌کند و برای مطالعه ذخیره، حفاظت، حرکت و تامین رطوبت خاک اهمیت زیادی دارد (Sun et al., 2021). با این حال، به دلیل سختی اندازه‌گیری مستقیم این ویژگی‌ها، زمان بر و پرهزینه بودن آنها (Qiao et al., 2019)، توسعه روش‌های غیرمستقیم و سریع برای تخمین این پارامترها، به ویژه با بهره‌گیری از توابع انتقالی خاک (PTFs) و فناوری نوین طیف سنجی در ناحیه مرئی-مادون قرمز نزدیک (Vis-NIR) به یک ضرورت علمی و عملی تبدیل شده است.

پیشینه پژوهش

در دهه‌های اخیر، رویکردهای نوین مبتنی بر توابع انتقالی و داده‌های طیفی به عنوان جایگزین‌های سریع و غیرمخرب برای تخمین ویژگی‌های رطوبتی خاک مورد توجه محققان قرار گرفته‌اند. این توابع، توابع برآورد کننده‌ای هستند که با استفاده از داده‌های زود یافت خاک، مانند ماده آلی، جرم مخصوص ظاهری و توزیع اندازه ذرات، امکان پیش بینی پارامترهای هیدرولیکی را بدون نیاز به آزمایش‌های پرهزینه فراهم می‌کنند (Bayat at al., 2020; Rastgou et al., 2020; Bayat at al., 2019; Bayat at al., 2018). روش‌های زیادی مانند رگرسیون خطی چندگانه گام به گام (Rosseti et al 2022, Myeni et al, 2022)، رگرسیون خطی چندگانه (Amsili et al., 2024; Qiao et al., 2019; Bayat et al., 2019)، رگرسیون غیرخطی (Rastgou et al., 2020)، شبکه عصبی مصنوعی (Qiao et al., 2019; Bayat et al., 2019; Bayat and Ebrahimi, 2016; Ma, 2024)، یادگیری ماشین (Veloso et al., 2022)، جنگل تصادفی (Ebrahimzadeh et al., 2023; Amsili et al., 2024)، ماشین‌های بردار پشتیبان (Lamorski et al., 2014; Twarakavi et al. 2000)، درخت‌های رگرسیونی (Pachepsky et al. 2006)، روش‌های گروهی پردازش داده‌ها (Ungaro et al., 2005; Bayat at al., 2011; Bayat at al., 2013; Neyshaburi et al., 2016)، ایجاد توابع انتقالی توسط محققان مورد استفاده قرار گرفته است. به علاوه، تاکنون برای تخمین پارامترهای FC و PWP از متغیرهای ورودی شامل شن، سیلت، رس، ماده آلی، چگالی ظاهری (Qiao et al 2019; Bayat et al., 2019; Bayat and Ebrahimi, 2016; Liu and Ma, 2024; Rastgou et al., 2020)، ویژگی‌های توپوگرافی و خواص حرارتی خاک (Liu & Ma, 2024)، pH (Liu & Ma, 2024)، هدایت الکتریکی (Rosseti et al., 2022; Myeni et al., 2022; Liu and Ma, 2024)، تخلخل ریز (Bayat at al., 2019)، مختصات جغرافیایی شامل طول، عرض و ارتفاع (Liu & Ma, 2024) استفاده شده است. با وجود آنکه نتایج این مطالعات دقت قابل قبولی در پیش بینی پارامترهای هیدرولیکی نشان داده‌اند، ولی بسیاری از آنها همچنان متکی بر ویژگی‌های فیزیکی و شیمیایی متداول خاک بوده و از اطلاعات طیفی خاک بهره‌ی محدودی گرفته‌اند. (Amsili et al (2024) نشان دادند که مدل‌های جنگل تصادفی

1. Field capacity
2. Permanent wilting point
3. Pedotransfer function
4. Visible- near-infrared spectroscopy

(RF) در برآورد FC، PWP و ظرفیت آب قابل دسترس عملکرد بهتری نسبت به رگرسیون خطی چندگانه (MLR) داشتند. Qiao et al (2019) با توسعه توابع انتقالی برای برآورد FC و PWP برای لایه عمیق فلات لس از متغیرهای چگالی ظاهری، شن، سیلت، رس و کربن آلی خاک با دو روش رگرسیون خطی چندگانه گام به گام و شبکه عصبی مصنوعی استفاده کردند. نتایج آنها نشان داد متغیرهای ورودی مهم برای پیش بینی FC شن و برای پیش بینی PWP رس و جرم مخصوص ظاهری بودند و عملکرد دو روش هم مشابه بود. اگر چه دقت دو روش تقریباً مشابه گزارش شد، اما مدل‌ها تنها بر پایه داده‌های فیزیکی خاک توسعه یافته بودند و تاثیر ویژگی‌های طیفی بر روی FC و PWP در نظر گرفته نشد. در پژوهش‌های مشابه، Rosseti et al (2022)، Myeni et al (2021) با توسعه توابع انتقالی به برآورد FC و PWP با استفاده از متغیرهای ورودی تخلخل ریز، شن، سیلت و رس با استفاده از رگرسیون خطی چندگانه گام به گام پرداختند. اگر چه این مطالعات پایه‌های مفیدی برای توسعه PTFها فراهم کرده‌اند، اما با توجه به عدم استفاده از داده‌های طیفی و مدل‌های پیشرفته غیرخطی، همچنان فرصت قابل توجهی برای بهبود دقت تخمین FC و PWP وجود دارد.

فناوری‌های طیف سنجی در محدوده Vis-NIR به دلیل حساسیت بالا به تغییرات فیزیکی و شیمیایی خاک، روشی سریع، ساده، کم هزینه و جایگزین برای آنالیزهای سنتی آزمایشگاهی، برای تخمین خواص خاک مانند رطوبت محسوب می‌شوند (Davari et al., 2022; Silva et al., 2016; Babaeain et al., 2015). ترکیب داده‌های طیفی با مدل‌های یادگیری ماشین مانند جنگل تصادفی می‌تواند دقت پیش بینی پارامترهای رطوبتی را بهبود بخشد (Sedaghat et al., 2022; Babaeain et al., 2016). اما بیشتر مدل‌های توسعه یافته تنها در مناطق یا بافت‌های خاص اعتبارسنجی شده‌اند و در شرایط خاکی و اقلیمی متنوع، دقت و قابلیت تعمیم کافی ندارند (Cooper et al., 2020). پژوهش‌های گذشته کارایی طیف سنجی مرئی-مادون قرمز نزدیک در پیش بینی ویژگی‌های مبنایی خاک مانند ماده آلی (Xiao et al., 2022; Zhang et al., 2020; Stenbreg, 2010)، توزیع اندازه ذرات (Pinheiro et al., 2017)، آبگریزی و حدود آتربرگ (Davari et al., 2021, 2022)، ظرفیت تبادل کاتیونی (Rasooli et al., 2023) و هدایت الکتریکی (Chatrenour et al., 2023) را اثبات کرده‌اند. انتظار می‌رود این روش در پیش بینی ویژگی‌های هیدرولیکی خاک نیز موثر باشد. با این حال، پژوهش‌های اندکی Beniaich et al (2016)؛ Babaeain et al., 2019؛ Soltani et al., 2021؛ Diao et al., 2023) در استفاده از داده‌های طیف Vis-NIR برای برآورد FC و PWP انجام شده است. Beniaich et al (2023) برای تخمین رطوبت ثقلی، رطوبت ظرفیت زراعی و رطوبت در نقطه پژمردگی دائم از روش RF، MLR، رگرسیون خطی ساده و الگوریتم کوبیست استفاده کردند. نتایج آنها برتری روش RF را نشان داد. این یافته نشان دهنده توانایی بالای مدل‌های یادگیری ماشین در استخراج الگوهای غیر خطی از داده‌های طیفی است. با این حال، مطالعه یاد شده تنها از داده‌های طیفی استفاده کرده است و ترکیبی از داده‌های طیفی و خاکی را در مدل لحاظ نکرده است. Wu et al (2022) به طور مشابه، برای تخمین رطوبت از تکنیک‌های فراطیفی استفاده کردند و توانستند دقت پیش بینی را نسبت به روش‌های کلاسیک بهبود دهند، اما تمرکز این پژوهش نیز بر تخمین مستقیم رطوبت بود و پارامترهای هیدرولیکی کلیدی مانند FC و PWP را لحاظ نکرده بودند.

بررسی و مرور جامع مطالعات در زمینه برآورد پارامترهای هیدرولیکی خاک مانند FC و PWP با استفاده از توابع انتقالی خاک نشان می‌دهد که اگر چه برخی پژوهش‌ها به صورت محدود به تخمین پارامترهای FC و PWP با استفاده از داده‌های طیفی و ویژگی‌های خاک پرداخته‌اند، با این حال، تاکنون هیچ مطالعه‌ای به طور جامع و عمیق به برآورد پارامترهای FC و PWP با استفاده از ترکیب داده‌های طیفی Vis-NIR و ویژگی‌های فیزیکی-شیمیایی خاک با بهره‌گیری از روش RF و MLR نپرداخته است. این خلاء پژوهشی نشان دهنده نیاز به مطالعات بیشتر در این حوزه است. بنابراین اهداف و نوآوری‌های پژوهش حاضر به شرح زیر است:

برآورد پارامترهای هیدرولیکی خاک (FC و PWP) با بهره‌گیری از ترکیب داده‌های ویژگی‌های فیزیکی-شیمیایی خاک و بازتاب طیفی در ناحیه مرئی-مادون قرمز نزدیک (Vis-NIR) و توسعه و اعتبارسنجی مدل‌های RF و MLR برای افزایش دقت و قابلیت تعمیم برآوردها در خاک‌هایی با بافت متنوع

ارزیابی تاثیر روش‌های پیش پردازش طیفی (MSC، 2، FD-SG1 و SNV) برای بهبود دقت مدل‌های تخمین پارامترهای FC و PWP شناسایی طول موج‌های کلیدی و متغیرهای خاکی موثر بر تغییرپذیری FC و PWP از طریق تحلیل مولفه‌های اصلی و توابع انتقالی منتخب.

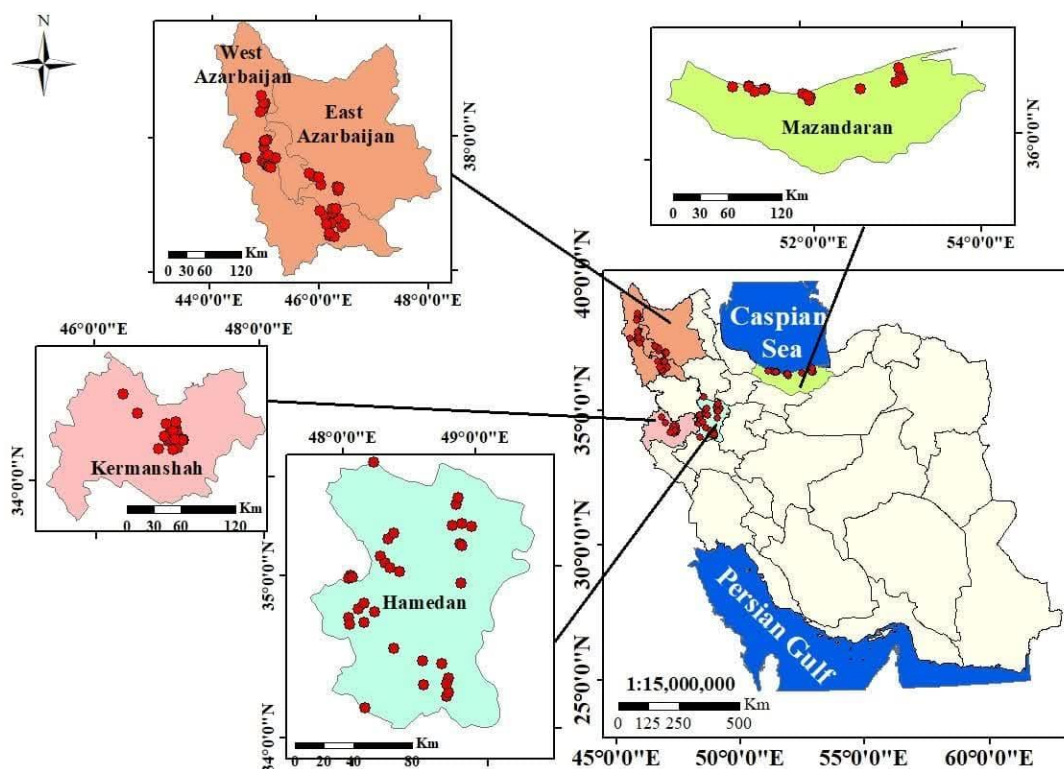
روش شناسی پژوهش

در این پژوهش، به منظور تعیین پارامترهای FC و PWP، در کل، نمونه‌برداری از ۱۳۰ نقطه انجام شد. از هر نقطه یک نمونه دست نخورده و یک نمونه دست‌خورده جمع‌آوری گردید. نمونه‌های دست نخورده با استفاده از سیلندرهایی استیل به قطر ۵ و ارتفاع ۴/۵ سانتی‌متر از اراضی کشاورزی و مرتعی ۵ استان همدان (۴۵)، کرمانشاه (۲۲)، آذربایجان غربی (۳۴)، آذربایجان شرقی (۹) و مازندران (۲۰) برداشت شد (شکل ۱). انتخاب استان‌ها بر اساس قابلیت دسترسی و امکان نمونه‌برداری از آن‌ها و همچنین تنوع اقلیمی، بافت خاک و مدیریت اراضی انجام شد تا نمونه‌ها نماینده طیف گسترده‌ای از شرایط خاکی کشور باشند و مدل‌های توسعه یافته قابلیت تعمیم به مناطق بیشتر را داشته باشند. در هر استان نمونه‌برداری بر اساس سری غالب خاک‌ها انجام شد. کاربری اراضی محل برداشت نمونه‌ها در استان همدان زراعی، در استان کرمانشاه زراعی زیر کشت گندم، در استان آذربایجان غربی زراعی و باغی، در استان آذربایجان شرقی زراعی و در استان مازندران اراضی جنگلی و باغی بود. نمونه‌ها پس از انتقال به آزمایشگاه فیزیک خاک دانشگاه بوعلی سینا، هوا خشک شده و از ال‌ک ۲ میلی‌متری عبور داده شدند، تا برای انجام آزمایش‌های فیزیکی، شیمیایی، ویژگی‌های هیدرولیکی و انجام کارهای طیف سنجی آماده شوند. بافت خاک به روش هیدرومتر اندازه‌گیری شد (Gee & Or, 2002). ماده آلی (OM) از طریق اکسایش تر با روش Walkley & Black (1934) بر روی ذرات کوچکتر از ۰/۵ میلی‌متر اندازه‌گیری شد. برای اندازه‌گیری جرم مخصوص ظاهری خاک از نمونه‌های خاک دست نخورده به روش استوانه‌های نمونه‌برداری، استفاده شد (Grossman & Reinsch, 2002). تخلخل کل که از نسبت حجم فضاهای خالی به حجم کل خاک به دست می‌آید، از روی جرم مخصوص ظاهری و جرم مخصوص حقیقی بر اساس رابطه

$$F = 1 - \frac{\rho b}{\rho s}$$

(که در آن F : تخلخل بر حسب cm^3/cm^3 ، ρb : جرم مخصوص ظاهری بر حسب گرم بر سانتی‌متر مکعب و ρs : جرم

مخصوص حقیقی برابر با ۲/۶۵ گرم بر سانتی‌متر مکعب محاسبه شد (Flint, 2002; Blake & Hartge, 1986).



شکل ۱. نقشه مکان‌های نمونه‌برداری در داخل ایران

منحنی نگهداری رطوبتی آب خاک (SWRC)^۱

برای تعیین FC و PWP، رطوبت نمونه‌های دست‌نخورده خاک به ترتیب در مکش‌های ۳۰ و ۱۵۰۰ کیلو پاسکال با دستگاه صفحات فشاری اندازه‌گیری شد (Dane & Hopmans, 2002). در این مطالعه برای تمامی ۱۳۰ نمونه دست‌نخورده، یک منحنی نگهداری آب خاک جداگانه اندازه‌گیری شد. چون از هر نقطه فقط یک نمونه دست‌نخورده برداشت شده بود، بنابراین برای هر مکش، یک تکرار اندازه‌گیری شد و اندازه‌گیری تا مکش ۲۰۰ کیلو پاسکال بر روی نمونه‌های دست‌نخورده و در مکش‌های بالاتر بر روی نمونه‌های دست‌خورده انجام شد. از منحنی‌های رطوبتی به دست آمده، FC به عنوان رطوبت خاک در مکش ۳۰ کیلو پاسکال و PWP به عنوان رطوبت خاک در مکش ۱۵۰۰ کیلو پاسکال تعیین شد.

اندازه‌گیری بازتاب طیفی و پردازش‌های طیفی

بازتاب طیفی نمونه‌ها با بهره‌گیری از دستگاه اسپکترومتر زمینی (FieldSpec@3, ASD, FR, USA) در دامنه Vis-NIR (۳۵۰ تا ۲۵۰۰ نانومتر) اندازه‌گیری شد. نمونه‌های خاک هوا خشک، و از الک ۲ میلی‌متری عبور داده شدند، ۱۰۰ گرم از هر نمونه در پتری دیش قرار داده شد و منحنی‌های طیفی آنها اندازه‌گیری شد (Viscarra Rossel, 2008). به منظور افزایش دقت، هر نمونه ۵ بار در موقعیت‌های مختلف قرائت شد. داده‌های طیفی پس از ثبت توسط نرم افزار View Spec Pro 6.0.15 (Manual, 2008)، میانگین‌گیری و سپس این منحنی‌ها به شکل کتابخانه‌های طیفی در قالب فایل متنی ذخیره شدند و جهت اعمال روش‌های مختلف پیش پردازش تصحیح پخشیده چندگانه (MSC)^۲، مشتق اول به همراه فیلتر ساویتزکی و گلای (FD-SG1)^۳، مشتق دوم به همراه فیلتر ساویتزکی و گلای (FD-SG2)^۴ و متغیر نرمال استاندارد (SNV)^۵ به نرم افزار Parles v3.1 (Viscara resell, 2008) منتقل شدند. بخش ابتدایی و انتهایی طیف‌ها (محدوده بین ۴۰۰-۳۵۰ و ۲۴۵۱-۲۵۰۰) به دلیل نویز زیاد حذف شده و در مدلسازی نیز مورد استفاده قرار نگرفت (Chatrenour et al., 2023).

پیش پردازش داده‌ها

قبل از توسعه PTFs، داده‌ها مورد پیش پردازش قرار گرفتند که شامل حذف داده‌های پرت، بررسی نرمال بودن با آزمون کولموگروف-اسمیرنوف، تبدیل داده‌های غیرنرمال به نرمال و استانداردسازی (میانگین صفر و واریانس ۱) بود (معادله ۱). داده‌های پرت بر اساس قانون $(\mu \pm 2.5 \times SD)$ شناسایی شدند. به عبارت دیگر، مقادیر بزرگتر از حد بالای تعیین شده $(\mu + 2.5 \times SD)$ و مقادیر کوچکتر از حد پایین تعیین شده $(\mu - 2.5 \times SD)$ به عنوان داده پرت در نظر گرفته شدند. به جای حذف کامل این داده‌ها، مقادیر پرت بزرگ با حد بالای تعیین شده و مقادیر پرت کوچک با حد پایین، جایگزین شدند تا اثرات ناشی از داده‌های پرت کاهش یابد. برای این منظور نرمال بودن داده‌ها با استفاده از آزمون کولموگروف-اسمیرنوف با نرم افزار Minitab16 (Minitab et al., 2000) بر روی تمامی متغیرها آزمون شد. متغیرهایی که توزیع نرمال نداشتند، تبدیل شدند و پس از آن کل متغیرها تبدیل استاندارد شدند تا میانگین صفر و واریانس ۱ داشته باشند (رابطه ۱). همبستگی بین متغیرهای ورودی و خروجی با نرم افزار SPSS 16 (SPSS Ins, 2007) بررسی شد.

$$Z_i = \frac{X_i - \mu}{\sigma} \quad \text{رابطه ۱}$$

در رابطه بالا Z_i : نمره استاندارد برای داده X_i است و μ : میانگین و σ : انحراف استاندارد برای داده‌ها است.

بررسی چند هم خطی بین متغیرهای ورودی

چند هم خطی زمانی رخ می‌دهد که دو یا چند متغیر مستقل به شدت با یکدیگر همبسته باشند. وجود آن می‌تواند دقت برآورد ضرایب رگرسیون را کاهش دهد، تفسیر نتایج را دشوار کند، قابلیت پیش بینی مدل را تحت تاثیر قرار دهد. برای بررسی چند هم خطی میان متغیرهای مستقل، ابتدا ماتریس همبستگی بررسی شد و هرگاه ضریب همبستگی پیروسون بین دو متغیر بیشتر از ۰/۸ بود، امکان چند

1. Soil water retention curve
2. Multiplicative signal correction
3. First derivative with Savitzki- Golay filter
4. The second derivative with a Savitzki- Golay filter
5. Standard normal variate
6. Standard deviation

هم خطی مورد توجه قرار گرفت. سپس با استفاده از نرم‌افزار SPSS 16 (SPSS Ins, 2007) برای بررسی وجود چند هم خطی میان متغیرهای مستقل در مدل رگرسیونی، شاخص عامل تورم واریانس (VIF) محاسبه شد. مطابق استانداردهای علمی مقادیر $VIF > 10$ نشان دهنده وجود چند هم خطی هستند (Nakarmi, 2024) که در این پژوهش در بین متغیرهای مستقل چند هم خطی مشاهده نشد.

تجزیه به مولفه‌های اصلی

به منظور کاهش ابعاد داده‌های طیفی و شناسایی مهم‌ترین ویژگی‌ها، تحلیل مولفه‌های اصلی (PCA) اعمال شد (Jolliffe & Cadima, 2016). سپس از مولفه‌های اصلی که بالاترین درصد توجیه واریانس را داشتند به عنوان تخمین‌گر برای تخمین FC و PWP از طریق توابع مختلف استفاده شد (جدول ۱). در این پژوهش، ابتدا داده‌های طیفی با استفاده از نرم‌افزار Parles v3.1 (Viscara rosel, 2008) تحت روش‌های پیش‌پردازش مختلف قرار گرفتند. سپس PCA بر روی داده‌های پیش‌پردازش شده اعمال گردید. مولفه‌های اصلی (PCs) بر اساس میزان واریانس تبیین شده توسط هر مؤلفه مرتب شدند. برای تحلیل‌های بیشتر، مؤلفه‌هایی که درصد قابل توجهی از واریانس کل داده‌ها را توجیه می‌کردند، انتخاب شدند (جدول ۱). این انتخاب بر اساس معیار درصد واریانس تجمعی انجام شد. به این صورت که برای روش‌های بدون پیش‌پردازش (NP)، MSC^۲، FD-SG1، FD-SG2، SNV و انتخاب مولفه‌هایی که مجموعاً به ترتیب ۹۹/۷، ۹۹/۴، ۷۰/۳ و ۹۱/۱۰ درصد واریانس را توجیه می‌کردند، انتخاب شدند. استفاده از این روش، ضمن کاهش حجم داده‌ها، امکان تمرکز بر مهم‌ترین تغییرات طیفی مرتبط با پدیده‌های مورد بررسی را فراهم نمود (Wold et al., 1987).

جدول ۱. درصد واریانس توجیه شده توسط مؤلفه‌های اصلی اول تا هشتم، در آنالیز تجزیه به مؤلفه‌های اصلی انجام شده بر روی داده‌های طیفی پیش‌پردازش شده با روش‌های مختلف.

Total explained variation x	Explained variation x (%)								Preprocessing
	PC8	PC7	PC6	PC5	PC4	PC3	PC2	PC1	
۱۰۰	۰/۰۰	۰/۰۰۱	۰/۰۰۲	۰/۰۰۷	۰/۰۱	۰/۰۴	۰/۲۰۴	۹۹/۷۲	NP
۱۰۰	۰/۰۰۰	۰/۰۰۱	۰/۰۰۲	۰/۰۰۳	۰/۰۱	۰/۰۴	۰/۴۶	۹۹/۴۴	MSC
۹۷/۰۸	۰/۴۹	۰/۵۴	۱	۱/۲۲	۲/۲۴	۴/۵۲	۱۶/۱۹	۷۰/۳۰	FD-SG1
۹۶	۰/۲۸	۰/۳۵	۰/۴۱	۰/۴۳	۰/۵۱	۱/۷۲	۵/۱۷	۸۶/۶۳	FD-SG2
۹۹/۹۸	۰/۰۱	۰/۰۳	۰/۰۵	۰/۱۰	۰/۳۰	۱/۳۳	۷	۹۱/۱۰	SNV

NP: داده‌های بدون پیش‌پردازش، MSC: روش تصحیح پخشیده چندگانه، FD-SG1: روش مشتق اول به همراه فیلتر ساویتزکی و گلائی، FD-SG2: روش مشتق دوم به همراه فیلتر ساویتزکی و گلائی، SNV: روش متغیر نرمال استاندارد، PC8, ..., PC1 به ترتیب نشان دهنده مؤلفه‌های اصلی اول تا هشتم.

گام‌های ساخت توابع انتقالی

برای توسعه PTFs، ۱۱ تابع انتقالی به منظور ارزیابی اثر نوع و سطح اطلاعات ورودی بر دقت تخمین پارامترهای FC و PWP، در قالب یک رویکرد مرحله‌ای توسعه داده شدند. این مرحله‌بندی با هدف بررسی تدریجی نقش متغیرهای خاکی از ساده‌ترین و در دسترس‌ترین ویژگی‌ها تا اطلاعات پیچیده‌تر انجام شد و امکان تحلیل اثر افزایشی هر گروه از متغیرها را فراهم می‌کند. داده‌ها به دو گروه آموزش (۹۰ نمونه) و تست (۱۰ نمونه) تقسیم گردیدند و مدل‌سازی با الگوریتم جنگل تصادفی و رگرسیون خطی چندگانه در نرم‌افزار STATISTICA14 (Abraham & Ledolter, 1983) انجام و اعتبارسنجی شد.

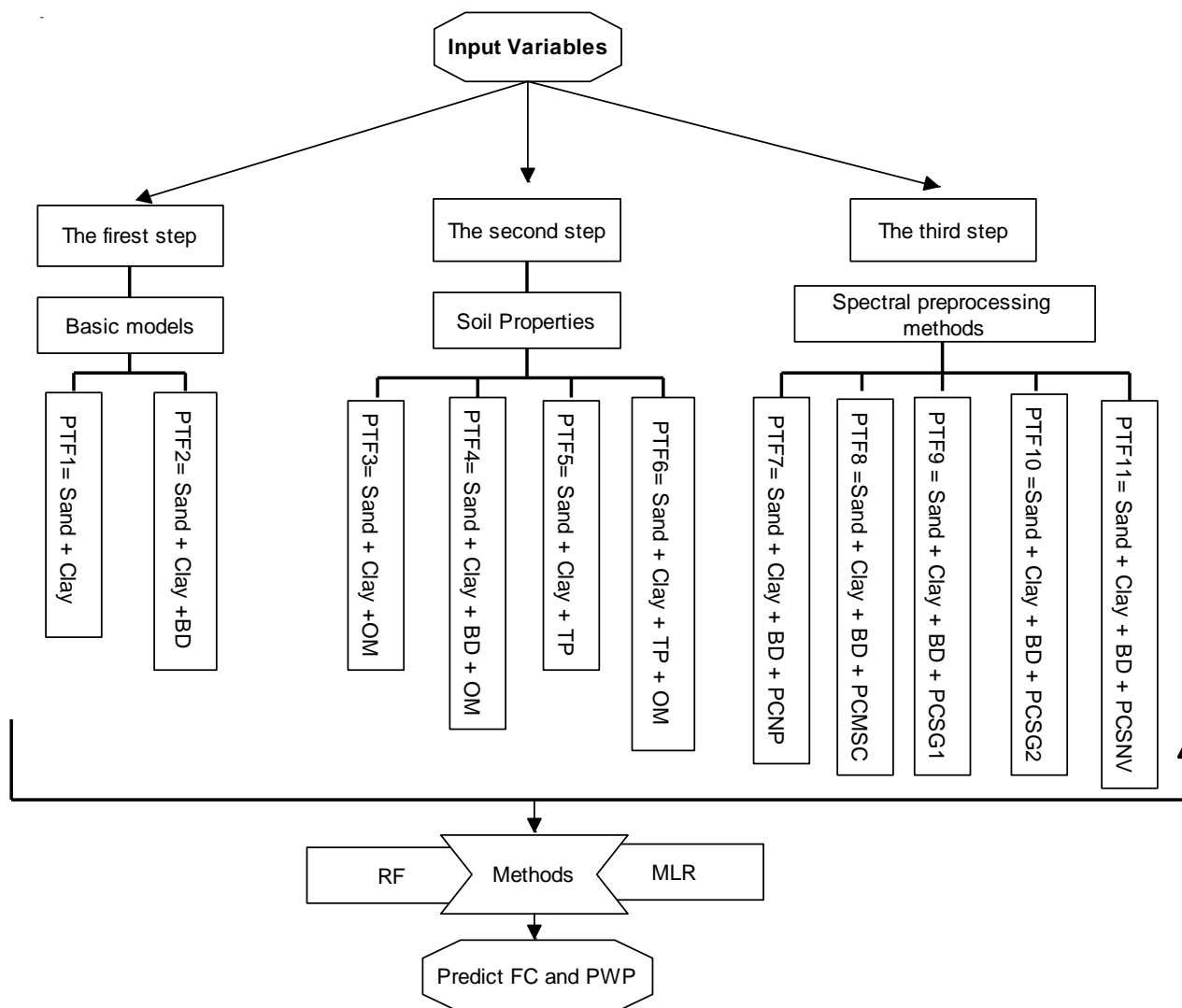
PTFs در سه مرحله ایجاد شدند:

مرحله ۱ (PTF1 و PTF2): یک مدل پایه برای تخمین پارامترهای FC و PWP با استفاده از شن، رس و چگالی ظاهری (BD) عنوان متغیرهای ورودی ساخته شد.

مرحله ۲ (PTF3 - PTF6): علاوه بر شن، رس و BD، متغیرهای جدیدی مانند تخلخل کل (TP) و ماده آلی (OM) هم به صورت جداگانه و هم ترکیبی برای بررسی تأثیر آنها بر پارامترهای FC و PWP معرفی شدند.

1. Variance inflation factor
2. Principal component analysis
3. No preprocessing
4. Bulk density
5. Total porosity

مرحله ۳ (PTF7-PTF11): داده‌های طیفی خام (بدون پیش‌پردازش) و روش‌های مختلف پیش‌پردازش طیفی، شامل تصحیح پخشیده چندگانه (PCMSC)، مشتق اول فیلترسائویتزکی-گولای (PCFD-SG1)، مشتق دوم فیلتر سائویتزکی-گولای (PCFD-SG2) و متغیر نرمال استاندارد (PCSNV)، به عنوان ورودی‌های تکمیلی در کنار شن، رس و BD گنجانده شدند (شکل ۲). از میان این روش‌ها هر تابعی که منجر به کاهش معیار اطلاعات آکایک (AIC)، نسبت به روش بدون پیش‌پردازش (PTF7) شد، با مدل پایه (PTF2) مقایسه گردید.



شکل ۲. متغیرهای ورودی ۱۱ تابع انتقال (PTF) برای پیش‌بینی پارامترهای FC و PWP. OM: ماده آلی، BD: چگالی ظاهری، TP: تخلخل کل، FC: رطوبت ظرفیت زراعی، PWP: رطوبت نقطه پژمردگی دائم، PCNP: اولین مولفه اصلی داده‌های بدون پیش‌پردازش، PCMSC: اولین مولفه اصلی روش پیش‌پردازش تصحیح پراکندگی چندگانه، PCFD-SG1: اولین مولفه اصلی روش پیش‌پردازش مشتق اول به همراه فیلتر سائویتزکی و گولای، PCFD-SG2: اولین مولفه اصلی روش پیش‌پردازش مشتق دوم به همراه فیلتر سائویتزکی و گولای، PCSNV: اولین مولفه اصلی روش پیش‌پردازش متغیر نرمال استاندارد، RF: جنگل تصادفی و MLR: رگرسیون خطی چندگانه.

روش‌های ارزیابی

برای بررسی دقت و قابلیت اطمینان توابع انتقالی توسعه یافته از آماره‌های معیار اطلاعات آکایک (AIC)، مجذور میانگین مربعات خطا (RMSE)، ضریب تعیین (R^2) و میانگین خطا (ME) استفاده شد.

1. Akaike's information criterion
2. Root mean square error
3. Coefficient of determination
4. Mean error

$$AIC = N \ln \left[\sum \frac{(Y_m - Y_p)^2}{N} \right] + 2P \quad \text{رابطه ۲}$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^N (Y_m - Y_p)^2}{N}} \quad \text{رابطه ۳}$$

$$ME = \frac{1}{N} \sum_{i=1}^N (Y_p - Y_m) \quad \text{رابطه ۴}$$

$$R^2 = \left\{ \frac{\left[N \left(\sum_{i=1}^N Y_m Y_p \right) - \sum_{i=1}^N Y_m \sum_{i=1}^N Y_p \right]}{\sqrt{N \left(\sum_{i=1}^N Y_m^2 \right) - \left(\sum_{i=1}^N Y_m \right)^2} \times \sqrt{N \left(\sum_{i=1}^N Y_p^2 \right) - \left(\sum_{i=1}^N Y_p \right)^2}} \right\} \quad \text{رابطه ۵}$$

که N : تعداد نمونه‌ها، P : تعداد متغیرهای ورودی به مدل، Y_p : مقادیر پیش بینی شده، Y_m : مقادیر اندازه‌گیری شده، ME : میانگین خطا که شدت کم و بیش برآوردی را نشان می‌دهد. حال آنکه هر چه $RMSE$ به صفر نزدیکتر باشد، دقت بیشتر است. شاخص R^2 نمایانگر مناسب بودن برازش معادله رگرسیونی بر داده‌های تجربی است و بهینه‌ترین مقدار R^2 برابر یک است. معیار AIC ، هر چه کوچک‌تر باشد، کیفیت برازش بالاتر است.

یافته‌های پژوهش و بحث

ویژگی‌های آماری

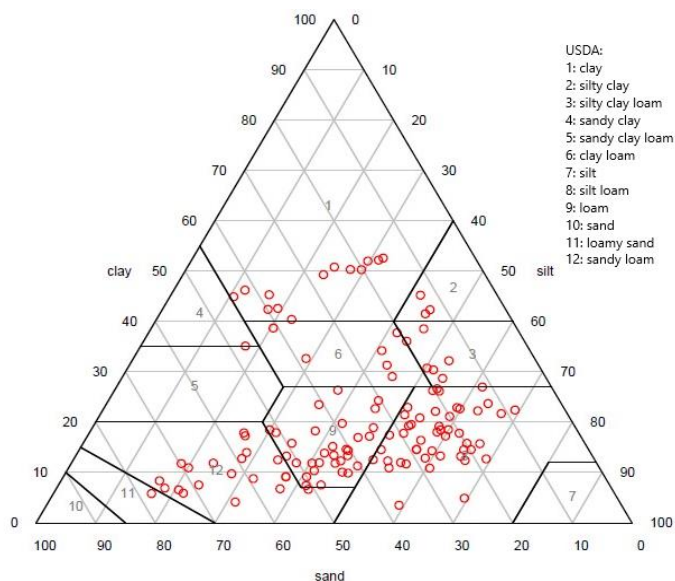
ویژگی‌های آماری (میانگین، حداقل، حداکثر و انحراف استاندارد) کلیه متغیرهای مورد استفاده برای سری داده‌های آموزش و تست (۱۳۰ نمونه) در این تحقیق در جدول شماره ۲ نشان داده شده است. این جدول شامل تمام متغیرهایی است که برای تخمین پارامترهای FC و PWP استفاده شدند.

جدول ۲. آماره‌های توصیفی ویژگی‌های فیزیکی و شیمیایی خاک‌های مورد مطالعه برای کل مجموعه داده‌ها

Variable	Mean	Min	Max	SD
Sand (%)	۳۶/۰۱	۸/۴۰	۷۷/۷۳	۱۶/۹۴
√Sand	۵/۸۳	۲/۹۰	۸/۸۲	۱/۴۱
Clay (%)	۲۰/۶۶	۳/۴۷	۵۲/۶۰	۱۲/۴۸
Log (Clay)	۱/۲۴	۰/۵۴	۱/۷۲	۰/۲۶
BD (g cm ⁻³)	۱/۴۲	۱/۰۲	۱/۹۵	۰/۱۷
OM (%)	۲/۰۱	۰/۱۷	۴/۸۰	۱/۱۱
√OM	۱/۳۶	۰/۴۱	۲/۱۹	۰/۴۱
TP (cm ³ cm ⁻³)	۰/۴۷	۰/۳۰	۰/۶۲	۰/۰۶
FC (cm ³ cm ⁻³) %	۳۴	۱۱	۵۰	۷
PWP (cm ³ cm ⁻³) %	۱۸	۴	۳۲	۵/۵

OM: ماده آلی، BD: چگالی ظاهری، TP: تخلخل کل، FC: رطوبت ظرفیت زراعی و PWP: رطوبت پژمردگی دائم. پارامترهایی که نرمال نشدند، تبدیل‌های لگاریتم و رادیکال بر آنها اعمال شد.

نمونه خاک‌های مورد مطالعه نه کلاس لوم سیلتی (۴۷)، لوم رسی (۸)، رس سیلتی (۳)، لوم رسی سیلتی (۹)، لوم (۳۵)، شن لومی (۱)، لوم شنی (۲۳)، رسی (۳) و لوم رسی شنی (۱) را در بر گرفته و از این نظر دارای تنوع مناسبی می‌باشند (شکل ۳).



شکل ۳. توزیع بافتی نمونه‌های خاک (n=۱۳۰) در مثلث بافتی وزارت کشاورزی ایالات متحده (USDA).

مقدار بیشینه جرم مخصوص ظاهری ۱/۹۵ گرم بر سانتی‌متر مکعب با ۱/۸۵ درصد ماده‌آلی بود. این عدد مربوط به خاک عمقی آوزرمان از استان همدان با بافت لوم شنی است. خاک‌های با بافت نسبتاً درشت و ماده‌آلی پایین، به ویژه در لایه‌های زیر سطحی، احتمالاً کاهش پایداری خاکدانه‌ها و تخلخل کل می‌تواند منجر به افزایش جرم مخصوص ظاهری شود (Brady & Weil, 2017; Bayat et al., 2011). پایین‌ترین جرم مخصوص ظاهری (۱/۰۲ گرم بر سانتی‌متر مکعب) مربوط به خاک سطحی زربال از استان آذربایجان غربی با بافت لوم رسی با درصد ماده‌آلی ۱/۸۴ بود. مقدار رطوبت در FC در بازه ۱۱٪ تا ۵۰٪ با میانگین ۳۴٪ و انحراف استاندارد ۷٪ مشاهده شد. با توجه به اینکه بر اساس مقادیر فوق ضریب تغییرات ۲۱٪ می‌باشد، نشان دهنده تنوع متوسط (Warrick, 1998; Wilding, 1985) توان نگهداری آب در خاک‌های مورد مطالعه است. چراکه بر اساس Warrick (1998) ضریب تغییرات ۱۵ تا ۳۰ درصد نشان‌دهنده پراکندگی متوسط داده‌ها می‌باشد. به‌ویژه اینکه مقادیر بالاتر FC عمدتاً در خاک‌های ریز بافت‌تر و مقادیر پایین‌تر در خاک‌های با بافت سبک‌تر مشاهده شد. همچنین مقدار PWP در دامنه ۴٪ تا ۳۲٪ با میانگین ۱۸٪ و انحراف استاندارد ۵٪ نشان دهنده اختلاف قابل توجه در حداقل رطوبت قابل استفاده توسط گیاهان در خاک‌های مختلف است.

بررسی همبستگی کلیه پارامترهای بهره‌گیری شده در مطالعه

رطوبت ظرفیت زراعی با متغیرهای شن و جرم مخصوص ظاهری همبستگی منفی و با مولفه اصلی اول روش پیش پردازش MSC (PCMSC)، رس، ماده‌آلی و تخلخل کل همبستگی مثبت داشت (جدول ۳). افزایش درصد شن در خاک باعث کاهش ظرفیت نگهداری آب می‌شود. زیرا ذرات شن بزرگتر بوده و فضای بین دانه‌های بیشتری دارند که آب را به سرعت از دست می‌دهد (Alghamdi et al., 2023). همچنین افزایش جرم مخصوص ظاهری نشان دهنده فشردگی بیشتر خاک و کاهش تخلخل است که منجر به کاهش ظرفیت نگهداری آب می‌شود. Xu et al., (2021) گزارش کردند که محتوای رطوبت خاک با شن همبستگی منفی معنادار دارد و با افزایش شن، نگهداری رطوبت کاهش می‌یابد. همچنین جرم مخصوص ظاهری بالا با کاهش ظرفیت نگهداری آب مرتبط است.

خاک‌های دارای درصد بالاتر رس و ماده‌آلی، به دلیل سطح ویژه زیاد ذرات رس و توانایی بالای ماده‌آلی در جذب و نگهداری آب، ظرفیت زراعی بیشتری دارند. همچنین افزایش تخلخل کل به معنای وجود فضای بیشتر برای نگهداری آب است. Xu et al., (2021) همبستگی مثبت بین رطوبت خاک و میزان ماده‌آلی و نقش تخلخل و اندازه ذرات در افزایش ظرفیت نگهداری آب را گزارش کردند. به طور کلی، MSC، باعث حذف پراکندگی‌های غیرخطی و بهبود سیگنال داده‌ها می‌شود که منجر به همبستگی مثبت و معنادار بین داده‌های پیش پردازش شده و رطوبت واقعی خاک می‌گردد (Li et al., 2023).

PWP با متغیر شن همبستگی منفی و با رس همبستگی مثبت داشت. همچنین با مولفه اصلی اول روش پیش پردازش SNV (PCSNV) همبستگی منفی و معنی‌داری داشت. همبستگی مثبت PWP با رس می‌تواند ناشی از ویژگی‌های شیمیایی و فیزیکی ذرات رس باشد. ذرات رسی دارای بار سطحی منفی و ظرفیت تبادل کاتیونی بالا هستند که مولکول‌های آب قطبی را جذب کرده و در

مکش‌های بالا (PWP) نگه می‌دارند. پس در PWP، منافذ آب را با نیروی ماتریک قوی‌تری نسبت به ذرات درشت‌تر نگه می‌دارند، بنابراین هر چه رس خاک بیشتر باشد، رطوبت باقی مانده در نقطه پژمردگی دائم نیز بیشتر خواهد بود. (Parker & Patignani, 2021) نشان دادند که خاک‌های رسی می‌توانند مقادیر PWP به طور قابل توجهی بالاتر از خاک‌های شنی داشته باشند.

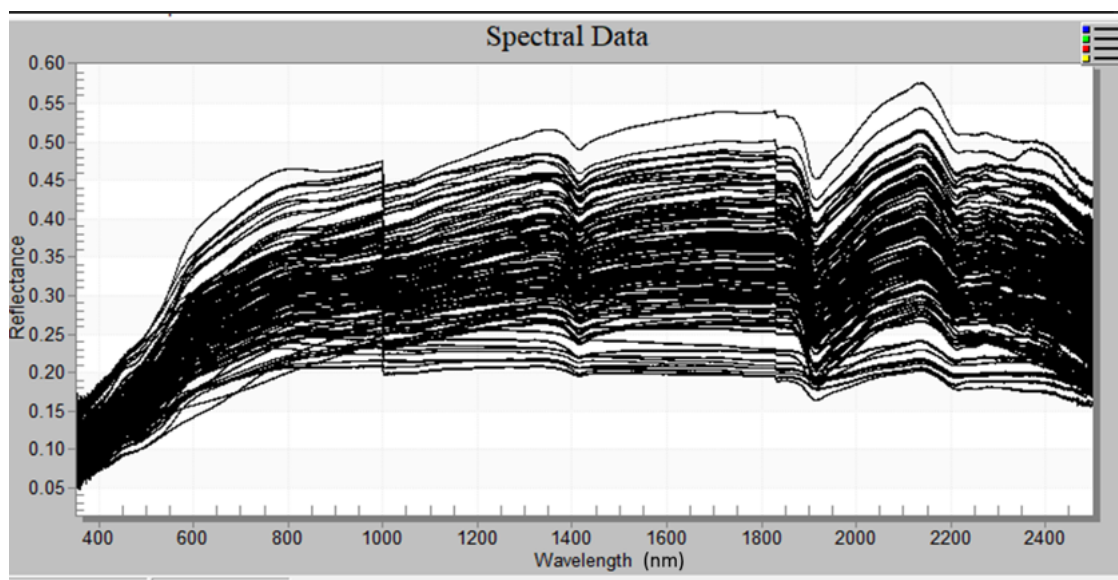
جدول ۲. مقادیر ضریب همبستگی پیرسون بین متغیرهای ورودی و پارامترهای FC و PWP (n=130).

Variables	Sand	Clay	BD	OM	TP	PCNP	PCMSC	PCFD-SG1	PCFD-SG2	PCSNV	FC	PWP
Sand	۱											
Clay	-۰/۶۰۸**	۱										
BD	۰/۲۹۶**	-۰/۱۱۵	۱									
OM	-۰/۰۳۸	-۰/۱۳۴	-۰/۱۴۸	۱								
TP	-۰/۲۹۳**	۰/۱۱۶	-۰/۹۹۸**	۰/۱۴۹	۱							
PCNP	۰/۰۱۷	-۰/۰۰۲	-۰/۰۹۳	۰/۰۶۱	۰/۰۸۷	۱						
PCMSC	-۰/۰۳۷	-۰/۱۱۶	-۰/۱۱۹	۰/۳۰۱**	۰/۱۱۰	۰/۴۰۶**	۱					
PCFD-SG1	۰/۱۴۱	-۰/۰۴۷	۰/۰۰۵	۰/۰۴۴	-۰/۰۰۷	۰/۶۹۷**	-۰/۰۵۵	۱				
PCFD-SG2	۰/۱۵۵	-۰/۰۶۱	۰/۰۰۶	۰/۰۴۹	-۰/۰۰۸	۰/۶۷۴**	-۰/۰۶۴	۰/۹۹۹**	۱			
PCSNV	۰/۱۶۴	-۰/۰۹۱	۰/۰۶۱	۰/۰۳۰	-۰/۰۵۰	-۰/۳۰۲**	-۰/۲۶**	۰/۰۳۴	۰/۰۵۹	۱		
FC	-۰/۵۵۱**	۰/۳۸۳**	-۰/۲۸۹**	۰/۱۹۶*	۰/۲۸۵**	۰/۰۳۲	۰/۱۸۵*	-۰/۰۵۸	-۰/۰۶۶	-۰/۱۵۵	۱	
PWP	-۰/۵۸۲**	۰/۵۷۹**	-۰/۱۳۶	۰/۱۴۵	۰/۱۳۶	۰/۱۴۳	۰/۱۴۸	۰/۰۶۲	۰/۰۴۹	۰/۲۳۱**	۰/۷۲۷**	۱

**همبستگی در سطح ۰/۰۱ معنی‌دار است، * همبستگی در سطح ۰/۰۵ معنی‌دار است. OM: ماده آلی، BD: چگالی ظاهری، TP: تخلخل کل، PCNP: اولین مولفه اصلی داده‌های بدون پیش پردازش، PCMSC: اولین مولفه اصلی روش تصحیح پراکندگی چندگانه، PCFD-SG1: اولین مولفه اصلی روش مشتق اول به همراه فیلتر ساویتزکی و گلای، PCFD-SG2: اولین مولفه اصلی روش مشتق دوم به همراه فیلتر ساویتزکی گلای، PCSNV: متغیر نرمال استاندارد و RF: جنگل تصادفی

منحنی‌های بازتاب طیفی

در شکل ۴ میانگین منحنی‌های خام بازتاب طیفی خاک‌های مورد مطالعه ارائه شده است. همانطور که در شکل مشاهده می‌شود بازتاب طیفی در همه خاک‌ها تقریباً مشابه است. با اینکه در میزان شدت بازتاب اختلافاتی وجود دارد. با توجه به شکل، منحنی‌های خام طیفی در دامنه طول موج‌های ۶۰۰-۵۵۰ نانومتر دارای یک برآمدگی و در طول موج‌های ۱۴۱۵، ۱۹۲۰ و ۲۲۱۲ نانومتر دارای مشخصه‌های جذب آب و رس می‌باشند (Clark et al., 1990; Stenberg et al., 2010). با توجه به مطالعاتی که انجام شده است، مشخصه‌های طیفی در طول موج‌های ۱۴۱۴ و ۱۹۱۵ نانومتر را می‌توان به ارتعاشات ترکیبی آب پیوندی در شبکه بین لایه‌ای و آب جذب شده روی سطح ذرات نسبت داد (Chatrenour et al., 2023; Davari et al., 2021; Gomez et al., 2008; Khayamim et al., 2015). باند ۱۴۱۵ نانومتر به علت کشش O-H ساختاری در لایه اکتاهیدرال و هم‌چنین ویژگی جذب در طول موج ۲۲۱۲ نانومتر می‌تواند به دلیل ترکیبی از خمش پیوند Al-OH و کشش پیوند O-H باشد (Stenberg et al., 2010). در این راستا (Gastaldi et al., 2015) گزارش کردند که کشش O-H مربوط به شبکه رسی دلیل این تغییرات است. Chatrenour et al (2023) گزارش کردند که افزایش مقدار رس باعث افزایش عمق ویژگی‌های جذب در طول موج‌های ۱۴۵۰، ۱۹۵۰ و ۲۲۰۰ نانومتر می‌شود. Cudahy et al (2016) گزارش کردند اگرچه انواع مختلف کانی‌های رسی می‌توانند باندهای جذب خاصی داشته باشند، ولی به طور کلی باندهای مشخصه خاک رس عمدتاً در محدوده طول موج‌های ۱۴۰۰، ۱۹۰۰ و ۲۲۰۰ نانومتر قرار دارند.



شکل ۴. میانگین منحنی‌های خام بازتاب طیفی خاک‌های مورد مطالعه

مقایسه دو روش رگرسیون خطی چندگانه و جنگل تصادفی براساس آماره RMSE و AIC برای پیش‌بینی پارامترهای FC و PWP مقایسه میانگین t-test جفتی دقت برازش مدل‌های MLR و RF براساس آماره RMSE برای پیش‌بینی پارامترهای FC و PWP برای داده‌های آموزش و آزمون انجام شد، به طوری که آزمون آماری صرفاً بر روی میانگین RMSE اعمال گردید و انحراف تنها به منظور توصیف پراکندگی مقادیر گزارش شدند (جدول ۴). نتایج مقایسه میانگین t-test نشان داد که در مرحله آموزش، در تخمین پارامترهای FC و PWP روش جنگل تصادفی نسبت به روش رگرسیون خطی چندگانه دارای RMSE کمتری (معنی‌دار در سطح ۵ درصد) بود. ولی در مرحله آزمون در تخمین FC روش MLR نسبت به روش RF دارای RMSE کمتری بود. در مقابل برای پارامتر PWP در مرحله آزمون تفاوت معنی‌داری بین دو روش در مقدار RMSE مشاهده نشد (جدول ۴).

جدول ۴. مقایسه میانگین مقادیر RMSE مدل‌های MLR و RF در پیش‌بینی پارامترهای FC و PWP برای داده‌های آموزش و آزمون با استفاده از آزمون t جفتی.

RMSE		Mean		Std-Deviation				
Dataset	Variables	RF	MLR	RF	MLR	t	df	sig
Train	FC (cm ³ cm ⁻³)	۰/۰۵۰	۰/۰۵۷	۰/۰۰۳۴	۰/۰۰۰۶	۸/۳۶	۱۰	۰/۰۰۰*
	PWP (cm ³ cm ⁻³)	۰/۰۳۴	۰/۰۴۱	۰/۰۰۲	۰/۰۰۰۵	۱۱/۷۲	۱۰	۰/۰۰۰*
Test	FC (cm ³ cm ⁻³)	۰/۰۶۲	۰/۰۶۱	۰/۰۰۰۸	۰/۰۰۰۷	-۶/۵۸	۱۰	۰/۰۰۰*
	PWP (cm ³ cm ⁻³)	۰/۰۴۲	۰/۰۴۱	۰/۰۰۲۰	۰/۰۰۰۵	۰/۸۷	۱۰	۰/۴۰

*: معنی‌داری در سطح ۵ درصد را نشان می‌دهد. FC: رطوبت حجمی ظرفیت زراعی، PWP: رطوبت حجمی پژمردگی دائم، RMSE: مجذور میانگین مربعات خطا، RF: جنگل تصادفی و MLR: رگرسیون خطی چندگانه. Sig: سطح معنی‌داری آزمون t جفتی برای مقایسه میانگین مقادیر RMSE بین دو مدل.

کلیه مراحل شامل آموزش و آزمون ۱۱ تابع انتقالی است که با استفاده از دو روش رگرسیون خطی چندگانه و جنگل تصادفی انجام شد که نتایج آن در جداول ۵ و ۶ مطابق با معیار AIC، نشان داده شده است. نتایج (جدول ۵) نشان داد که توابع PTF3,4,5,6,9, 10,11 در مرحله آموزش در تخمین FC در روش RF بهتر از MLR بودند. ولی هیچ کدام از توابع در مرحله تست بین دو روش تفاوت معنی‌داری ایجاد نکردند. هم‌چنین در تخمین پارامتر PWP (جدول ۶) در همه توابع در مرحله آموزش (به جز تابع ۱ و ۲) روش RF به طور معنی‌داری بهتر از روش MLR بود. ولی در مرحله تست از بین همه توابع تنها PTF2 بین دو روش تفاوت معنی‌داری ایجاد کرد که در روش RF بهتر از MLR بود.

نتایج نشان داد که عملکرد روش RF در تخمین پارامترهای FC و PWP بر اساس مقادیر AIC در مقایسه با رگرسیون خطی چندگانه در مرحله آموزش دقیق‌تر بود. در تخمین پارامترهای FC و PWP برای روش RF مقادیر AIC به ترتیب از ۵۵۳ تا ۵۱۳- و از

۶۱۶- تا ۵۸۰- در مرحله آموزش متغیر بود، که در مقایسه با روش MLR که برای پارامترهای FC و PWP به ترتیب از ۵۱۵- تا ۵۰۸- و از ۵۷۵- تا ۵۶۹- در مرحله آموزش متغیر بود، کمتر بود. مشخص شد که مقادیر RMSE برای PTFها در تخمین پارامترهای FC و PWP در روش RF در مرحله آموزش به ترتیب از ۰/۰۴۶ تا ۰/۰۵۸ و از ۰/۰۳۲ تا ۰/۰۳۹ سانتی‌متر مکعب بر سانتی‌متر مکعب بود، که آنها کوچکتر از روش MLR بودند که از ۰/۰۵۷ تا ۰/۰۵۹ و از ۰/۰۴۱ تا ۰/۰۴۲ سانتی‌متر مکعب بر سانتی‌متر مکعب متغیر بودند. به طور کلی مقادیر AIC روش RF در مقایسه با روش MLR در مرحله آموزش برای هر دو پارامتر کمتر بود. ولی در روش رگرسیون خطی چندگانه در مرحله آزمون مقادیر AIC روش MLR در مقایسه با روش RF به طور غیرمعنی‌دار برای پارامتر FC کمتر بود. ولی برای پارامتر PWP در غالب PTFها (به جز برای PTF2 که معنی‌دار بود)، روش RF به طور غیرمعنی‌داری بهتر از MLR بود.

در مجموع نتایج مرحله آزمون نشان داد که روش RF در تخمین پارامتر PWP در غالب توابع انتقالی عملکرد بهتری نسبت به MLR دارد، هر چند این برتری در اغلب موارد از نظر آماری معنی‌دار نبود. در تخمین FC نیز تفاوت معنی‌داری بین دو روش در مرحله آزمون مشاهده نشد. بنابراین می‌توان نتیجه گرفت که روش RF در این مطالعه از پایداری قابل قبولی برخوردار است، اما برتری آن نسبت به MLR به پارامتر مورد بررسی و نوع تابع انتقالی وابسته است.

الگوریتم RF با بهره‌گیری از تکنیک یادگیری تجمعی و میانگین‌گیری بر روی چندین درخت تصمیم‌گیری، توانایی بیشتری در شناسایی روابط غیرخطی و اثرات متقابل بین متغیرها دارد و نسبت به نویز و داده‌های پرت مقاومت بیشتری نشان می‌دهد. این در حالی است که مدل MLR به دلیل فرضیات خطی و حساسیت به هم خطی بین متغیرها، در مرحله آموزش عملکرد ضعیف‌تری از خود نشان می‌دهد (Pham et al., 2020). Arslan et al (2014) برای پیش‌بینی پارامترهای FC و PWP از طیف سنجی VIS-NIR استفاده کردند و نتایج آنها برتری مدل رگرسیون خطی چندگانه نسبت به رگرسیون حداقل مربعات جزئی (PLSR) را نشان داد. (Sihag et al (2019). در تحقیقی که بر روی بررسی قابلیت تخمین هدایت هیدرولیکی غیراشباع با استفاده از تکنیک‌های رگرسیون غیرخطی چند متغیره، مدل درخت M5 و رگرسیون جنگل تصادفی صورت گرفت، بیان کردند که مدل جنگل تصادفی به دلیل کارایی پیش‌بینی بالاتر در توسعه و اعتبارسازی مدل و قابلیت تعمیم بالاتر نسبت به دو مدل دیگر، عملکرد بهتری داشته است. (Carranza et al (2021) از الگوریتم جنگل تصادفی برای پیش‌بینی رطوبت ناحیه ریشه استفاده کردند و عملکرد آن را با مدل فرآیندی همراه با ادغام داده‌ها مقایسه کردند. نتایج نشان داد که RF در تخمین مقادیر داده‌های موجود دقت بالاتری داشت.

برآورد FC و PWP با استفاده از ویژگی‌های خاک

نتایج به دست آمده از جداول ۵ و ۶ نشان داد که در روش جنگل تصادفی از بین توابع ایجاد شده در مرحله دوم برای تخمین FC در مرحله آموزش PTF3 و PTF6 به طور معنی‌داری AIC را نسبت به PTF1 کاهش دادند. همچنین نتایج نشان داد که برای تخمین PWP در روش جنگل تصادفی در مرحله آموزش با مقایسه بین توابع ایجاد شده در مرحله دوم تنها PTF6 به طور معنی‌داری AIC را نسبت به PTF1 کاهش داد. ولی در مرحله آزمون در روش جنگل تصادفی هیچ کدام از توابع ایجاد شده در تخمین FC و PWP تفاوت معنی‌داری را با تابع PTF1 ایجاد نکردند (جداول ۵ و ۶). در واقع موثرترین متغیر در مرحله دوم به عنوان تخمین‌گر FC ماده‌آلی و برای PWP ماده‌آلی و تخلخل کل می‌باشد. در روش MLR هیچ کدام از توابع بهبود معنی‌داری در تخمین پارامترهای FC و PWP نسبت به تابع ۱ نشان ندادند.

متغیرهای ورودی در مرحله دوم در PTF3 و PTF6 به ترتیب شن، رس و ماده‌آلی و ماده‌آلی با تخلخل کل هستند. در مرحله آموزش در PTF3 و PTF6 وقتی ماده‌آلی و ترکیب آن با تخلخل کل به جمع متغیرهای ورودی شن و رس اضافه شد، بر اساس AIC خطای تخمین پارامتر FC را به ترتیب ۵/۳ و ۵ درصد نسبت به PTF1 کاهش دادند. با توجه به اینکه بین ماده‌آلی و تخلخل کل با پارامتر FC همبستگی مثبت (به ترتیب 0.196^{**} و 0.285^{**}) و با جرم مخصوص ظاهری همبستگی منفی (-0.289^{**}) وجود دارد (جدول ۳)، ماده‌آلی خاک با کاهش جرم مخصوص ظاهری و افزایش تخلخل با تاثیر بر ساختمان خاک، به ویژه از طریق افزایش پایداری خاکدانه‌ها و بهبود نفوذپذیری و نگهداری آب نقش کلیدی در رطوبت FC در مکش‌های پایین ایجاد می‌کند (Jorreh et al., 2013; Lehmann & Kleber., 2015; Huang et al., 2022). با توجه به اینکه اثر ماده‌آلی بر رطوبت ظرفیت زراعی ماهیتی چند مکانیسمی دارد و حاصل برهمکنش همزمان فرآیندهای فیزیکی، سطحی و الکتروشیمیایی در خاک است، بنابراین ماده‌آلی به علت افزایش سطح ویژه خاک، موجب افزایش رطوبت FC می‌شود. از طرفی ماده‌آلی به علت دارا بودن بار منفی وابسته به pH ناشی از گروه‌های عاملی کربوکسیل

و فنول، از طریق افزایش جذب سطحی آب، تقویت لایه دوگانه الکتریکی و افزایش آب جذب شده روی سطوح خاک، نقش مهمی در نگهداری آب ایفا می‌کند (Kandra et al., 2024). این تاثیر مثبت ماده‌آلی بر FC موجب افزایش همبستگی و در نتیجه بهبود عملکرد مدل می‌شود. به عبارت دیگر ماده‌آلی به عنوان یک متغیر پیش‌بینی‌کننده قوی در مدل عمل می‌کند و حضور آن باعث کاهش عدم قطعیت در تخمین FC می‌گردد. افزودن همزمان ماده‌آلی و تخلخل به این صورت است که ماده‌آلی با افزایش سطح ویژه و تشکیل کمپلکس‌های آلی-معدنی، پایداری خاکدانه‌ها را بهبود می‌بخشد، که مستقیماً بر توزیع اندازه منافذ خاک اثر گذاشته و تخلخل کل را افزایش می‌دهد. در حالی که تخلخل کل بازتابی از آرایش فیزیکی این خاکدانه‌هاست که فضای لازم برای نگهداری آب را فراهم می‌کند (Jorreh et al., 2013). این هم افزایی منجر به ایجاد منافذ ریز می‌شود که مسئول نگهداری آب در حالت FC در مکش‌های پایین و PWP در مکش‌های بالا هستند. در نتیجه دقت تخمین FC و PWP را بهبود می‌بخشند. از طرفی، رطوبت PWP با رس همبستگی مثبت (0.579^{**}) و با شن همبستگی منفی (-0.582^{**}) دارد. از آنجا که نگهداشت رطوبت در مکش‌های بالاتر عمدتاً به بافت خاک و سطح فعال ذرات معدنی وابسته است (Liu et al., 2019)، بنابراین مواد آلی و رس به علت دارا بودن سطح ویژه بالا، قدرت جذب و نگهداشت زیاد و گنجایش تبادل کاتیونی زیاد، باعث افزایش رطوبت در همه مکش‌ها به ویژه مکش‌های ماتریک بالا می‌شوند (Tuller and Or., 2004). بنابراین این رابطه دو طرفه موجب می‌شود مدل‌های شامل هر دو متغیر خطای کمتری در پیش‌بینی FC و PWP را داشته باشند.

جدول ۵. نتایج تخمین FC از طریق RF و MLR با استفاده از متغیرهای ورودی مختلف

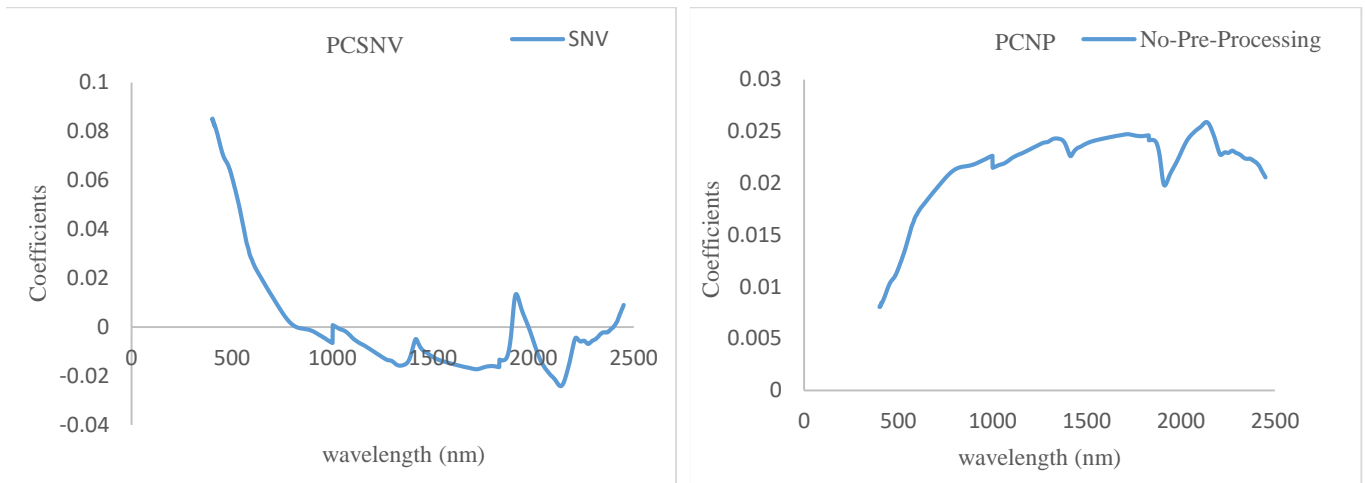
FC	RF				MLR			
	ME	RMSE	R ²	AIC	ME	RMSE	R ²	AIC
	Train				Train			
PTF1	-۰/۰۰۲۱	۰/۰۵۳	۰/۴۶	-۵۲۵/۳	۰/۰۰۱۵	۰/۰۵۹	۰/۳۰	-۵۰۸/۹
PTF2	-۰/۰۰۰۵	۰/۰۵۸	۰/۴۸	-۵۱۳/۲	۰/۰۰۱۴	۰/۰۵۸	۰/۳۱	-۵۱۱/۰
PTF3	-۰/۰۰۰۶	۰/۰۴۶	۰/۶۱	-۵۵۳/۳ [#]	۰/۰۰۰۸	۰/۰۵۷	۰/۳۳	-۵۱۳/۵
PTF4	-۰/۰۰۰۳	۰/۰۴۷	۰/۶۰	-۵۴۹/۵	۰/۰۰۰۸	۰/۰۵۷	۰/۳۴	-۵۱۵/۱
PTF5	-۰/۰۰۰۸	۰/۰۴۹	۰/۵۵	-۵۴۰/۹	۰/۰۰۱۴	۰/۰۵۸	۰/۳۱	-۵۱۰/۸
PTF6	-۰/۰۰۰۳	۰/۰۴۶	۰/۶۲	-۵۵۱/۸ [#]	۰/۰۰۰۸	۰/۰۵۷	۰/۳۴	-۵۱۵/۰
PTF7	-۰/۰۰۱۳	۰/۰۵۰	۰/۵۵	-۵۲۶/۵	۰/۰۰۱۳	۰/۰۵۸	۰/۳۱	-۵۱۱/۱
PTF8	-۰/۰۰۰۹	۰/۰۴۹	۰/۵۶	-۵۳۹/۶*	۰/۰۰۱۰	۰/۰۵۷	۰/۳۴	-۵۱۴/۹
PTF9	-۰/۰۰۰۶	۰/۰۵۰	۰/۵۵	-۵۳۹/۳*	۰/۰۰۱۴	۰/۰۵۸	۰/۳۱	-۵۱۱/۱
PTF10	-۰/۰۰۱۰	۰/۰۴۹	۰/۵۴	-۵۳۹/۵*	۰/۰۰۱۴	۰/۰۵۸	۰/۳۱	-۵۱۱/۱
PTF11	-۰/۰۰۰۹	۰/۰۴۹	۰/۵۵	-۵۴۰/۷ [#]	۰/۰۰۱۲	۰/۰۵۸	۰/۳۲	-۵۱۱/۷
	Test				Test			
PTF1	-۰/۰۰۶۶	۰/۰۶۳	۰/۳۷	-۲۱۹/۵	-۰/۰۰۲۵	۰/۰۶۲	۰/۳۳	-۲۲۰/۵
PTF2	-۰/۰۰۹۸	۰/۰۶۲	۰/۳۵	-۲۲۱/۱	-۰/۰۰۳۲	۰/۰۶۱	۰/۳۶	-۲۲۲/۱
PTF3	-۰/۰۰۱۸	۰/۰۶۰	۰/۴۰	-۲۲۲/۸	-۰/۰۰۱۹	۰/۰۶۰	۰/۳۷	-۲۲۲/۹
PTF4	-۰/۰۰۲۹	۰/۰۶۲	۰/۳۹	-۲۲۱/۰	-۰/۰۰۱۸	۰/۰۵۹	۰/۳۸	-۲۲۳/۹
PTF5	-۰/۰۰۳۳	۰/۰۶۳	۰/۳۶	-۲۱۹/۶	-۰/۰۰۳۲	۰/۰۶۱	۰/۳۶	-۲۲۲/۱
PTF6	-۰/۰۰۳۰	۰/۰۶۱	۰/۳۹	-۲۲۱/۶	-۰/۰۰۱۸	۰/۰۶۰	۰/۳۸	-۲۲۳/۸
PTF7	-۰/۰۰۵۳	۰/۰۶۲	۰/۳۵	-۲۲۱/۲	-۰/۰۰۳۰	۰/۰۶۱	۰/۳۶	-۲۲۲/۰
PTF8	-۰/۰۰۲۸	۰/۰۶۰	۰/۳۸	-۲۲۳/۲	-۰/۰۰۳۳	۰/۰۶۰	۰/۳۸	-۲۲۳/۵
PTF9	-۰/۰۰۲۷	۰/۰۶۲	۰/۳۳	-۲۲۰/۶	-۰/۰۰۳۱	۰/۰۶۱	۰/۳۶	-۲۲۲/۰
PTF10	-۰/۰۰۴۳	۰/۰۶۲	۰/۳۷	-۲۲۰/۲	-۰/۰۰۳۱	۰/۰۶۱	۰/۳۶	-۲۲۲/۰
PTF11	-۰/۰۰۴۲	۰/۰۶۲	۰/۳۸	-۲۲۰/۷	-۰/۰۰۲۸	۰/۰۶۱	۰/۳۶	-۲۲۲/۲

علامت # نشان دهنده تفاوت معنی‌دار بین دو روش در توابع نظیر هم می‌باشد. * نشان دهنده تفاوت معنی‌دار بین هر تابع با تابع مقایسه شونده آن می‌باشد. توابع مرحله ۲ با PTF1 مقایسه شدند و توابع مرحله ۳ PTF2 مقایسه شدند. رطوبت حجمی در ظرفیت زراعی ME، میانگین خطای RMSE، مجذور میانگین مربعات خطا، R²: ضریب تعیین، AIC: معیار اطلاعات آکابیک، RF: جنگل تصادفی، MLR: رگرسیون خطی چندگانه و CM: مدل مقایسه شونده.

برآورد FC با استفاده از داده‌های طیفی

پارامتر FC با استفاده از داده‌های طیفی و ویژگی‌های خاک با استفاده از PTF7-11 برآورد شد (جدول ۵). نتایج (جدول ۵) نشان داد که با مقایسه بین PTFهای مرحله سوم (PTF7-PTF11) برای داده‌های آموزش، PTF11 (SNV) دارای AIC منفی‌تر نسبت به تابع ۲ و توابع

دیگر بود (جدول ۵). ولی هیچ کدام از توابع ایجاد شده در مرحله تست بر اساس معیار اطلاعات آکایک در برآورد پارامتر FC بهبود معنی‌داری نسبت به PTF2 ایجاد نکردند (جدول ۵). همچنین نتایج تحلیل PCA نشان داد که اولین مؤلفه اصلی داده‌های طیفی پردازش شده با روش متغیر نرمال استاندارد (PCSNV) در محدوده طول موج ۴۰۹ نانومتر دارای ارتباط با متغیرهای خروجی بودند (شکل ۶-الف).



شکل ۶. ضرایب مؤلفه‌های اصلی در مقابل طول موج برای اولین مؤلفه اصلی داده‌های طیفی پردازش شده با روش بدون پیش پردازش (PCNP) (شکل -الف) و اولین مؤلفه اصلی داده‌های طیفی پردازش شده با روش SNV (PCSNV) (شکل -ب).

PTF11 در این تابع از اولین مؤلفه اصلی داده‌های طیفی پردازش شده با روش متغیر نرمال استاندارد (PCSNV) به همراه تخمین‌گرهای PTF2 در برآورد پارامتر FC بهره‌گیری شد. استفاده از SNV با به کار بردن تبدیلات خطی بر روی هر طیف، سبب حذف خطاهای چندگانه ناشی از پراکنش و نیز اندازه ذرات خاک و ناهمواری‌های سطحی می‌شود (Storey & Helmy., 2019). در PCSNV بیشترین ضریب با مقادیر بازتاب طیفی در محدوده مرئی در طول موج ۴۰۹ نانومتر بود (شکل ۶-الف). رطوبت FC با محتوای رس، ماده آلی و تخلخل کل همبستگی مثبت دارد ($r=0.383^{**} \cdot 0.196^{*} \cdot 0.285^{**}$). بازتاب در محدوده طول موج مرئی، تحت تاثیر پارامترهایی از خاک مانند ساختمان خاک، بافت، رطوبت و ماده‌آلی خاک قرار می‌گیرد (Padarian et al., 2019; Turco et al., 2015). از طرفی هم دقت برآورد رطوبت خاک هم تحت تاثیر رس، ماده‌آلی، تخلخل خاک و توزیع آنها قرار دارد (Myeni et al., 2021). بنابراین اثر مشخصه‌های جذبی در طول موج ۴۰۹ نانومتر برای تخمین FC را می‌توان به کانی‌های رسی و میزان ماده‌آلی نسبت داد. چرا که موادآلی و رس به علت دارا بودن سطح ویژه بالا، قدرت جذب و نگه‌داشت زیاد و گنجایش تبادل کاتیونی زیاد موجب افزایش رطوبت در همه مکش‌ها می‌شوند (Myeni et al., 2021). این ممکن است به دلیل منافذ ریز زیاد در بافت رسی و عامل اتصال دهندگی ماده‌آلی با بهبود ساختمان خاک باشد که حتی پس از تخلیه آب ثقلی، می‌تواند آب بیشتری را در خود نگه دارد (Rout & Arulmozhiselvan, 2019). در همین راستا Stenberg et al (2010) گزارش کردند مشخصه‌های جذبی در ناحیه مرئی را می‌توان به کانی‌های رسی از جمله گنویتیت نسبت داد. کانی گنویتیت در طول موج‌های ۴۲۰، ۴۶۰ و ۴۸۰ نانومتر دارای بازتاب طیفی در محدوده مرئی است. (Nocita et al (2014) گزارش کردند که ویژگی‌های جذبی در طول موج‌های ۴۱۰ و ۵۱۰ نانومتر همبستگی خوبی با میزان ماده‌آلی دارد. از این رو، استفاده از داده‌های طیفی به عنوان ورودی در مدل‌های پیش‌بینی ظرفیت زراعی، با بهبود تخمین رطوبت خاک و دیگر ویژگی‌های مرتبط، موجب کاهش خطا و بهبود معیارهای آماری مانند AIC می‌شود.

نتایج مقایسه روش‌های پیش پردازش با بدون پیش پردازش در تخمین پارامتر FC نشان داد که PTF11 در مرحله سوم در مرحله آموزش موجب بهبود غیرمعنی‌دار AIC نسبت به روش بدون پیش پردازش شد. در PTF11 از اولین مؤلفه اصلی داده‌های طیفی پردازش شده با روش متغیر نرمال استاندارد (PCSNV) به همراه متغیرهای ورودی PTF2 به عنوان تخمین‌گر استفاده شد. تکنیک SNV شامل استاندارد سازی هر طیف است (Wu et al., 2024). این تکنیک با استاندارد سازی هر طیف و کاهش اثرات مضاعف پراکندگی با مرکزگیری و مقیاس بندی منحنی‌های طیفی به روند حذف خطاهای پراکندگی و تغییرات طول مسیر منحنی‌ها برای تخمین پارامتر FC کمک

می‌کند، به عبارتی پس از مقیاس بندی، هر طیف دارای میانگین صفر و انحراف استاندارد یک است (Li et al., 2020). این ویژگی‌های اساسی SNV در تخمین پارامتر FC موجب عملکرد بهتر (غیرمعنی‌دار) این تکنیک نسبت به روش بدون پیش پردازش شده است. Padhi et al (2024) گزارش کردند که در SNV هر طیف در مرکز قرار می‌گیرد و متعاقبا با انحراف استاندارد مربوط به آن مقیاس بندی می‌شود. به عبارتی تبدیل از طریق SNV به کاهش اثرات مضاعف پراکندگی کمک می‌کند. Dotto et al (2017) جهت تخمین میزان ماده آلی در ۵۹۲ نمونه خاک بهترین عملکرد را در روش پیش پردازش SNV گزارش نمودند.

برآورد PWP با استفاده از داده‌های طیفی و خاکی

پارامتر PWP با استفاده از داده‌های طیفی و ویژگی‌های خاک از طریق PTF7-11 برآورد شد. با مقایسه بین PTF‌های مرحله سوم (PTF7-PTF11)، در مرحله آموزش (PCNP) PTF7 دارای AIC منفی‌تر نسبت به مدل پایه و توابع دیگر بود (جدول ۶). ولی هیچ کدام از توابع ایجاد شده در مرحله تست بر اساس AIC در برآورد پارامتر PWP بهبود معنی‌داری نسبت به PTF2 ایجاد نکردند (جدول ۶). از طرفی نتایج تحلیل PCA نشان داد که اولین مولفه اصلی داده‌های طیفی بدون پیش پردازش (PCNP) در محدوده طول موج‌های ۱۴۱۴، ۱۹۱۲ و ۲۱۵۰ نانومتر، دارای ارتباط با متغیرهای خروجی بودند (شکل ۶-ب).

جدول ۶. نتایج تخمین PWP از طریق RF و MLR با استفاده از متغیرهای ورودی مختلف

PWP	RF				MLR			
	ME	RMSE	R ²	AIC	ME	RMSE	R ²	AIC
Train								
PTF1	-۰/۰۰۱۴	۰/۰۳۸	۰/۵۸	-۵۸۵/۳	۰/۰۰۱۶	۰/۰۴۲	۰/۴۱	-۵۶۹/۲
PTF2	-۰/۰۰۰۳	۰/۰۳۹	۰/۵۷	-۵۸۰/۲	۰/۰۰۱۶	۰/۰۴۲	۰/۴۱	-۵۶۹/۲
PTF3	-۰/۰۰۰۶	۰/۰۳۳	۰/۶۶	-۶۱۲/۹*	۰/۰۰۱۱	۰/۰۴۱	۰/۴۴	-۵۷۴/۵
PTF4	-۰/۰۰۰۷	۰/۰۳۳	۰/۶۷	-۶۱۴/۹*	۰/۰۰۱۱	۰/۰۴۱	۰/۴۵	-۵۷۴/۹
PTF5	-۰/۰۰۰۲	۰/۰۳۴	۰/۶۳	-۶۰۷/۳*	۰/۰۰۱۶	۰/۰۴۲	۰/۴۱	-۵۶۹/۲
PTF6	-۰/۰۰۰۵	۰/۰۳۲	۰/۶۸	-۶۱۶/۳*	۰/۰۰۱۱	۰/۰۴۱	۰/۴۵	-۵۷۴/۹
PTF7	-۰/۰۰۰۳	۰/۰۳۳	۰/۶۵	-۶۱۲/۴*	۰/۰۰۱۳	۰/۰۴۱	۰/۴۳	-۵۷۳/۱
PTF8	-۰/۰۰۱۰	۰/۰۳۴	۰/۶۴	-۶۰۷/۳*	۰/۰۰۱۳	۰/۰۴۱	۰/۴۳	-۵۷۳/۶
PTF9	-۰/۰۰۰۵	۰/۰۳۴	۰/۶۴	-۶۰۵/۱*	۰/۰۰۱۳	۰/۰۴۱	۰/۴۴	-۵۷۳/۰
PTF10	-۰/۰۰۰۴	۰/۰۳۴	۰/۶۳	-۶۰۶/۴*	۰/۰۰۱۴	۰/۰۴۱	۰/۴۴	-۵۷۲/۹
PTF11	-۰/۰۰۱۰	۰/۰۳۴	۰/۶۵	-۶۰۹/۶**	۰/۰۰۱۳	۰/۰۴۱	۰/۴۳	-۵۷۲/۲
Test								
PTF1	-۰/۰۰۸۶	۰/۰۴۰	۰/۴۸	-۲۵۵/۲	-۰/۰۰۵۹	۰/۰۴۱	۰/۴۹	-۲۵۳/۲
PTF2	-۰/۰۰۲۶	۰/۰۳۶	۰/۵۶	-۲۶۴/۳*	-۰/۰۰۳۷	۰/۰۴۳	۰/۴۳	-۲۵۰/۸
PTF3	-۰/۰۰۰۳	۰/۰۴۱	۰/۵۳	-۲۵۴/۸	-۰/۰۰۲۵	۰/۰۴۲	۰/۴۵	-۲۵۲/۷
PTF4	-۰/۰۰۳۴	۰/۰۴۲	۰/۵۰	-۲۵۲/۱	-۰/۰۰۲۵	۰/۰۴۲	۰/۴۵	-۲۵۲/۶
PTF5	-۰/۰۰۳۵	۰/۰۴۳	۰/۴۸	-۲۵۱/۶	-۰/۰۰۳۷	۰/۰۴۳	۰/۴۳	-۲۵۰/۸
PTF6	-۰/۰۰۳۵	۰/۰۴۲	۰/۵۰	-۲۵۲/۶	-۰/۰۰۲۵	۰/۰۴۲	۰/۴۵	-۲۵۲/۶
PTF7	-۰/۰۰۲۹	۰/۰۴۳	۰/۴۸	-۲۵۰/۹	-۰/۰۰۲۸	۰/۰۴۲	۰/۴۴	-۲۵۲/۰
PTF8	-۰/۰۰۴۵	۰/۰۴۲	۰/۴۹	-۲۵۲/۲	-۰/۰۰۲۹	۰/۰۴۱	۰/۴۶	-۲۵۳/۸
PTF9	-۰/۰۰۴۰	۰/۰۴۱	۰/۵۰	-۲۵۴/۵	-۰/۰۰۳۰	۰/۰۴۳	۰/۴۳	-۲۵۰/۸
PTF10	-۰/۰۰۳۵	۰/۰۴۳	۰/۴۹	-۲۵۱/۷	-۰/۰۰۳۱	۰/۰۴۳	۰/۴۳	-۲۵۰/۷
PTF11	-۰/۰۰۴۲	۰/۰۴۲	۰/۵۰	-۲۵۲/۵	-۰/۰۰۳۰	۰/۰۴۲	۰/۴۵	-۲۵۲/۴

علامت # نشان دهنده تفاوت معنی‌دار بین دو روش در توابع نظیر هم می‌باشد. * نشان دهنده تفاوت معنی‌دار بین هر تابع با تابع مقایسه شونده آن می‌باشد. توابع مرحله ۲ با PTF1 مقایسه شدند و توابع مرحله ۳ با PTF2 مقایسه شدند. PWP: رطوبت حجمی در نقطه پژمردگی دائم، ME: میانگین خطا، RMSE: مجذور میانگین مربعات خطا، R²: ضریب تعیین، AIC: معیار اطلاعات آکاییک، RF: جنگل تصادفی، MLR: رگرسیون خطی چندگانه و CM: مدل مقایسه شونده.

PTF7: در این تابع از اولین مولفه اصلی داده‌های طیفی بدون پیش پردازش (PCNP) به همراه تخمین‌گرهای PTF2 در برآورد پارامتر PWP بهره‌گیری شد. داده‌های طیفی دارای نویزهای پراکندگی نور، اثرات سطح نمونه و تغییرات غیرخطی هستند. معمولاً پیش پردازش‌هایی مانند MSC، SG1، SG2 و SNV برای کاهش این اثرات و بهبود کیفیت داده‌ها به کار می‌روند. با این حال در برخی موارد

پیش پردازش‌های شدید می‌توانند سیگنال‌های اصلی مرتبط با ویژگی‌های خاک را کاهش دهند. به ویژه وقتی داده‌های طیفی به طور طبیعی دارای سیگنال قوی بوده و نسبت سیگنال به نویز بالا باشد. بنابراین حذف پیش پردازش ممکن است منجر به حفظ بهتر اطلاعات اصلی و بهبود عملکرد مدل شود. احتمالاً به همین دلیل روش بدون پیش پردازش نسبت به روش‌های دیگر موجب بهبود تخمین پارامتر PWP شده است.

بزرگترین ضرایب PC اول در روش بدون پیش پردازش بازتاب طیفی در ناحیه مادون قرمز نزدیک در طول موج ۲۱۵۰ نانومتر و با دو مشخصه جذب در طول موج‌های ۱۴۱۴ و ۱۹۱۸ نانومتر بود (شکل ۶- الف). محدوده‌های طیفی طول موج ۲۴۵۰-۲۱۵۰ نانومتر را می‌توان به گروه‌های عاملی OH شبکه رسی نسبت داد (Viscarrá Rossel et al., 2006)، که در این پژوهش با طول موج ۲۱۵۰ نانومتر مطابقت دارد. محدوده‌های فعال طیفی در طول موج‌های ۱۴۱۴ و ۱۹۱۲ نانومتر دارای مشخصه‌های جذبی آب و رس می‌باشند (Stenberg et al., 2010). مشخصه‌های جذبی برای تخمین رطوبت PWP را می‌توان به ارتعاشات ترکیبی آب محدود شده در شبکه‌های بین لایه‌ای رس و ترکیبی از کشش O-H و خمش H-O-H مولکول‌های آب به دام افتاده در شبکه کریستالی نسبت داد (Davari et al., 2021). بنابراین داده‌های طیفی به واسطه حساسیت بالا نسبت به ترکیب کانی شناسی و ویژگی‌های جذب آب در رس، می‌توانند به طور موثری برای تخمین PWP مورد استفاده قرار گیرند و بازتاب دهنده تغییرات ظرفیت نگه‌داری آب خاک در این نقطه بحرانی باشند.

نتیجه‌گیری

در این مطالعه، ۱۳۰ نمونه خاک از مناطق مختلف ۵ استان ایران جمع‌آوری و برخی ویژگی‌های آن‌ها اندازه‌گیری شد. همچنین، بازتاب طیفی آن‌ها در بازه ۲۵۰۰-۳۵۰ نانومتر ثبت گردید. نوآوری پژوهش در استفاده همزمان داده‌های طیفی و ویژگی‌های فیزیکی-شیمیایی خاک برای تخمین پارامترهای FC و PWP و بهره‌گیری از مدل‌های یادگیری ماشین به ویژه RF و MLR از طریق توابع انتقالی است. در این پژوهش نتایج زیر به دست آمد:

تحلیل مدل‌ها نشان داد که پیش پردازش‌های طیفی، بسته به نسبت سیگنال به نویز، تاثیر متفاوتی بر دقت پیش بینی دارند و در برخی موارد داده‌های بدون پیش پردازش عملکرد مناسبی ارائه کردند، که اهمیت حفظ سیگنال اصلی خاک را نشان می‌دهد. همچنین تحلیل اهمیت طول موج‌ها منجر به شناسایی بازه‌های طیفی موثر در تخمین FC و PWP گردید، که این یافته می‌تواند مبنایی برای طراحی و بهبود سنسورهای تخصصی در آینده باشد.

یافته‌های این پژوهش نشان داد که بهره‌گیری از ترکیب داده‌های طیفی و ویژگی‌های فیزیکی و شیمیایی خاک در کنار الگوریتم‌های یادگیری ماشین، به ویژه RF و MLR می‌تواند روشی موثر، دقیق و مقرون به صرفه برای برآورد پارامترهای کلیدی FC و PWP در سطح وسیع فراهم آورد.

توابع PTF3 و PTF6 در مرحله آموزش بر پایه متغیرهای بافتی و فیزیکی خاک، به ترتیب به عنوان توابع بهینه برای پارامترهای FC و PWP معرفی شدند.

از منظر کاربردی، رویکرد ترکیبی طیف سنجی و یادگیری ماشین، پتانسیل بالایی برای پیش و مدیریت بهینه منابع آب و خاک در سامانه‌های کشاورزی پیشرفته دارد. توسعه این فناوری می‌تواند گامی موثر در جهت ارتقای دقت و سرعت تصمیم‌گیری‌های مدیریتی و حرکت به سوی کشاورزی هوشمند و پایدار باشد.

ملاحظات اخلاقی

حامی مالی

مقاله حاضر با حمایت مالی و معنوی معاونت پژوهشی دانشگاه بو علی سینا همدان انجام شد. حمایت مالی از این پژوهش از طرف دانشگاه بو علی سینا، دانشکده کشاورزی در قالب پژوهانه پایان‌نامه دانشجویی نویسنده اول و همچنین پژوهانه برای سایر نویسندگان انجام شده است.

مشارکت نویسندگان

جمع‌آوری داده‌ها: چیمین مهدی‌زاده؛ حسین بیات، تهیه گزارش پژوهش: چیمین مهدی‌زاده؛ حسین بیات، تحلیل داده‌ها: چیمین مهدی‌زاده، حسین بیات و مسعود داوری، بررسی و کنترل نتایج (حسین بیات)، اصلاح (حسین بیات) و بازبینی و نهایی‌سازی مقاله (چیمین مهدی‌زاده).

مهدی زاده و حسین بیات). نویسندگان در کلیه مراحل طراحی و انجام پژوهش، تجزیه و تحلیل آماری داده‌ها، تحلیل و تفسیر اطلاعات و نتایج، تهیه گزارش، پیشنهاد مقاله، بررسی و کنترل نتایج، اصلاح و بازبینی و نهایی‌سازی مقاله مشارکت داشتند و نقش مشارکت آنها به ترتیب به شرح زیر است: چیمین مهدی زاده، حسین بیات و مسعود داوری.

مشارکت نویسندگان در مقاله مستخرج از پایان‌نامه تقریباً به شکل زیر باشد:

نویسنده اول: دانشجو: تهیه و آماده‌سازی نمونه‌ها، انجام آزمایش و گردآوری داده‌ها، انجام محاسبات، تجزیه و تحلیل آماری داده‌ها، تحلیل و تفسیر اطلاعات و نتایج، تهیه پیشنهاد مقاله

نویسنده دوم: استاد راهنمای پایان‌نامه، طراحی پژوهش، نظارت بر مراحل انجام پژوهش، بررسی و کنترل نتایج، اصلاح، بازبینی و نهایی‌سازی مقاله

نویسنده سوم: استاد مشاور پایان‌نامه، مشارکت در طراحی پژوهش، نظارت بر پژوهش، مطالعه و بازبینی مقاله

اعلامیه هوش مصنوعی مولد و فناوری‌های مبتنی بر هوش مصنوعی در فرایند نگارش

متن اعلامیه: در طول آماده‌سازی این اثر، نویسنده(گان) از [ChatGpt] به منظور [ترجمه و بهبود نگارش متن مقاله] استفاده کردند. پس از استفاده از این ابزار، نویسنده(گان) مطالب را در صورت نیاز بررسی و ویرایش کرده و مسئولیت کامل محتوای نشریه را بر عهده می‌گیرند.

بیانیه دسترسی به داده‌ها

بیانیه دسترسی به داده‌ها و سیاست اشتراک‌گذاری داده‌ها به شرح زیر است.
۱. داده‌های پژوهش حاضر از طریق درخواست از نویسندگان قابل دسترسی است.

سپاسگزاری

از دانشگاه بو علی سینا همدان به خاطر حمایت مالی / حمایت معنوی در انجام پژوهش حاضر سپاسگزاری می‌شود.
از داوران محترم به خاطر ارائه نظرهای ساختاری و علمی سپاسگزاری می‌شود.

پیروی از اصول اخلاق پژوهش

این پژوهش شامل انسان یا حیوان نبوده و داده‌های حساس مورد استفاده قرار نگرفته است. نویسندگان اصول اخلاق پژوهش و انتشار علمی رعایت کرده و این موضوع مورد تایید همه آنهاست.

تعارض منافع

نویسندگان این مقاله تعارض منافع ندارد.

REFERENCE

- Abraham, B., & Ledolter, J. (1983). *Statistical methods for forecasting* (Vol. 179). New York: Wiley.
- Alghamdi, A. G., Majrashi, M. A., & Ibrahim, H. M. (2023). Improving the physical properties and water retention of sandy soils by the synergistic utilization of natural clay deposits and wheat straw. *Sustainability*, 16(1), 46.
- Amsili, J. P., van Es, H. M., & Schindelbeck, R. R. (2024). Pedotransfer functions for field capacity, permanent wilting point, and available water capacity based on random forest models for routine soil health analysis. *Communications in Soil Science and Plant Analysis*, 55(13), 1967-1984.
- Arslan, H., Tasan, M., Yildirim, D., Koksall, E. S., & Cemek, B. (2014). Predicting field capacity, wilting point, and the other physical properties of soils using hyperspectral reflectance spectroscopy: Two different statistical approaches. *Environmental Monitoring and Assessment*, 186, 5077-5088
- Babaeian E., Homae M., Montzka C., Vereecken H., & Norouzi A.A. (2015a). Towards retrieving soil hydraulic properties by hyperspectral remote sensing. *Vadose Zone Journal*. 14(3), 1-17.
- Babaeian E., Homae M., Vereecken H., Montzka C., Norouzi, A.A., and van Genuchten M.T. (2015b). 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, 1043–1058.
- Babaeian, E., Homae, M., Montzka, C., Vereecken, H., Norouzi, A. A., & van Genuchten, M. T. (2016). Soil moisture prediction of bare soil profiles using diffuse spectral reflectance information and vadose zone flow modeling. *Remote Sensing of Environment*, 187, 218-229
- Bayat, H., Asghari, S., Rastgou, M., & Sheykhzadeh, G. R. (2020). Estimating Proctor parameters in agricultural soils in

- the Ardabil plain of Iran using support vector machines, artificial neural networks and regression methods. *Catena*, 189, 104467.
- Bayat, H., Ebrahimi, E., & Fallah, M. (2018). Estimation of soil moisture using confined compression curve parameters. *Geoderma*, 318, 64-77
- Bayat, H., Mazaheri, B., & Mohanty, B. P. (2019). Estimating soil water characteristic curve using landscape features and soil thermal properties. *Soil and Tillage Research*, 189, 1-14.
- Beniaich, A., Otten, W., Shin, H. C., Cooper, H. V., Rickson, J., Soulaïmani, A., & El Gharous, M. (2023). Evaluation of pedotransfer functions to estimate some of soil hydraulic characteristics in North Africa: A case study from Morocco. *Frontiers in Environmental Science*, 11, 1090688.
- Blake, G. R., & Hartge, K. H. (1986). Particle density. In A. Klute (Ed.), *Methods of soil analysis: Part 1. Physical and mineralogical methods* (pp. 377-382). Madison, WI: Soil Science Society of America.
- Brady, N. C., & Weil, R. R. (2017). *The nature and properties of soils* (15th ed.). Pearson.
- Carranza, C., Nolet, C., Peziz, M., & van der Ploeg, M. (2021). Root zone soil moisture estimation with Random Forest. *Journal of Hydrology*, 593, 125840.
- Castaldi, F., Palombo, A., Pascucci, S., Pignatti, S., Santini, F., & Casa, R. (2015). Reducing the Influence of Soil Moisture on the Estimation of Clay from Hyperspectral Data: A Case Study Using Simulated PRISMA Data. *Remote Sensing*, 7(11), 15561-15582.
- Castellini, M., & Iovino, M. (2019). Pedotransfer functions for estimating soil water retention curve of Sicilian soils. *Archives of Agronomy and Soil Science*, 65(10), 1401-1416.
- Chatrenour, M., Landi, A., Bahrami, H., & Mirzaei, S. (2023). Dust source clay content and salinity estimation using VNIR spectrometry. *Arid Land Research and Management*, 37(3), 369-388.
- Chatrenour, M., Landi, A., Bahrami, H., & Mirzaei, S. (2023). Dust source clay content and salinity estimation using VNIR spectrometry. *Arid Land Research and Management*, 37(3), 369-388.
- Clark, R. N., King, T. V. V., Klejwa, M., Swayze, G. A., & Vergo, N. (1990). High spectral resolution reflectance spectroscopy of minerals. *Journal of Geophysical Research: Solid Earth*, 95(B8), 12653-12680.
- Cooper, E., Blyth, E., Cooper, H., Ellis, R., Pinnington, E., & Dadson, S. J. (2020). Using data assimilation to optimize pedotransfer functions using large-scale in-situ soil moisture observations. *Hydrology and Earth System Sciences Discussions*, 2020, 1-20.
- Cudahy, T., Caccetta, M., Thomas, M., Hewson, R., Abrams, M., Kato, M., Kashimura, O., Ninomiya, Y., Yamaguchi, Y., & Collings, S. (2016). Satellite-derived mineral mapping and monitoring of weathering, deposition and erosion. *Scientific Reports*, 6(1), 23702
- Davari, M., Fahmideh, S., & Mosaddeghi, M. R. (2022). Rapid assessment of soil water repellency indices using Vis-NIR spectroscopy and pedo-transfer functions. *Geoderma*, 406, 115486.
- Davari, M., Karimi, S. A., Bahrami, H. A., Taher Hossaini, S. M., & Fahmideh, S. (2021). Simultaneous prediction of several soil properties related to engineering uses based on laboratory Vis-NIR reflectance spectroscopy.
- Diao, W., Liu, G., Zhang, H., Hu, K., & Jin, X. (2021). Influences of soil bulk density and texture on estimation of surface soil moisture using spectral feature parameters and an artificial neural network algorithm. *Agriculture*, 11(8), 710
- Dotto, A. C., Dalmolin, R. S. D., ten Caten, A., & Grunwald, S. (2018). A systematic study on the application of scatter-corrective and spectral-derivative preprocessing for multivariate prediction of soil organic carbon by Vis-NIR spectra. *Geoderma*, 314, 262-274.
- Flint, L. E. (2002). Porosity. In J. H. Dane & G. C. Topp (Eds.), *Methods of Soil Analysis. Part 4 — Physical Methods* (pp. 241–254). Soil Science Society of America, Madison, WI.
- Gee, G.W., & Orr, D. (2002). Particle size analysis. In J.H. Dane & G. C. Topp (Eds.), *Methods of soil analysis: part 4. Physical methods* (Vol. 5, pp. 255-293). Madison, WI: Soil Science Society of America.
- Gomez, C., Lagacherie, P., & Coulouma, G. (2008). Continuum removal versus PLSR method for clay and calcium carbonate content estimation from laboratory and airborne hyperspectral measurements. *Geoderma*, 148(2), 141-148
- Grossman, R. B., & Reinsch, T. G. (2002). Bulk density and linear extensibility. In J. H. Dane & G. C. Topp (Eds.), *Methods of Soil Analysis. Part 4 — Physical Methods* (pp. 201–228). Soil Science Society of America / American Society of Agronomy.
- Huang, J., Li, X., & Zhang, Y. (2022). Integration of hyperspectral data and machine learning for accurate soil water retention parameter estimation. *Remote Sensing of Environment*, 270, 112851.
- Jolliffe, I. T., & Cadima, J. (2016). Principal component analysis: a review and recent developments. *Philosophical transactions of the royal society A: Mathematical, Physical and Engineering Sciences*, 374(2065), 20150202.
- Jorreh, M., Bayat, H., Safari Sanjani, A. A., & Davatghar, N. (2013). Estimation of soil penetration resistance using fractal parameters of particle and aggregate size distributions. *Water and Soil Science*, 23(2), 13-27
- Kaisermann, A., de Vries, F. T., Griffiths, R. I., & Bardgett, R. D. (2017). Legacy effects of drought on plant–soil feedbacks and plant–plant interactions. *New Phytologist*, 215(4), 1413-1424.
- Kandra, B., Tall, A., Vitková, J., Procházka, M., & Šurda, P. (2024). Effect of humic amendment on selected hydrophysical properties of sandy and clayey soils. *Water*, 16(10), 1338.
- Khayamim, F., Wetterlind, J., Khademi, H., Robertson, A. J., Cano, A. F., & Stenberg, B. (2015). Using visible and near

- infrared spectroscopy to estimate carbonates and gypsum in soils in arid and subhumid regions of Isfahan, Iran. *Journal of Near Infrared Spectroscopy*, 23(3), 155-165
- Lamorski, K., Sławiński, C., Moreno, F., Barna, G., Skierucha, W., & Arrue, J. L. (2014). Modelling soil water retention using support vector machines with genetic algorithm optimisation. *The Scientific World Journal*, 2014(1), 740521.
- Lehmann, J., & Kleber, M. (2015). The contentious nature of soil organic matter. *Nature*, 528(7580), 60-68.
- Li, P., Zha, Y., Zuo, B., & Zhang, Y. (2023). A family of soil water retention models based on sigmoid functions. *Water Resources Research*, 59(3), e2022WR033160
- Liu, W., Luo, X., Huang, F., & Fu, M. (2019). Prediction of soil water retention curve using Bayesian updating from limited measurement data. *Applied Mathematical Modelling*, 76, 380-395.
- Manual, V. P. U. (2008). ASD Document 600555 Rev. A, Boulder, CO, 80301.
- Minitab, I. N. C. (2000). MINITAB statistical software. *Minitab Release*, 13(0).
- Myeni, L., Mdlambuzi, T., Paterson, D. G., De Nysschen, G., & Moeletsi, M. E. (2021). Development and evaluation of pedotransfer functions to estimate soil moisture content at field capacity and permanent wilting point for South African soils. *Water*, 13(19), 2639
- Nakarmi, S. S. (2024). Multi-collinearity in Research and Wayforward. *Kaladarpan*, 4(1), 85-91
- Navidi, M. N., Seyedmohammadi, J., & Seyed Jalali, S. A. (2022). Predicting soil water content using support vector machines improved by meta-heuristic algorithms and remotely sensed data. *Geomechanics and Geoengineering*, 17(3), 712-726
- Neyshaburi, M. R., Bayat, H., Rastgou, M., Mohammadi, K., Gregory, A. S., & Nariman-Zadeh, N. (2016). Parametric estimation of water retention using mGMDH method and principal component analysis. *Polish Journal of Soil Science*, 49(1).
- Nocita, M., Stevens, A., Toth, G., Panagos, P., van Wesemael, B., & Montanarella, L. (2014). Prediction of soil organic carbon content by diffuse reflectance spectroscopy using a local partial least square regression approach. *Soil Biology and Biochemistry*, 68, 337-347.
- Pachepsky, Y. A., Rawls, W. J., & Lin, H. S. (2006). Hydropedology and pedotransfer functions. *Geoderma*, 131(3-4), 308-316.
- Padarian, J., Minasny, B., & McBratney, A. B. (2019). Using deep learning to predict soil properties from regional spectral data. *Geoderma Regional*, 16, e00198
- Padhi, S. R., John, R., Tripathi, K., Wankhede, D. P., Joshi, T., Rana, J. C., ... & Bhardwaj, R. (2024). A comparison of spectral preprocessing methods and their effects on nutritional traits in cowpea germplasm. *Legume Science*, 6(2), e2977.
- Parker, N., & Patrignani, A. (2021). Evaluating Traditional and Modern Laboratory Techniques for Determining Permanent Wilting Point. *Kansas Agricultural Experiment Station Research Reports*, 7(5), 10.
- Pinheiro É.F., Ceddia M., Clingensmith C., Grunwald S., and Vasques G. 2017. Prediction of soil physical and chemical properties by visible and near-infrared diffuse reflectance spectroscopy in the Central Amazon. *Remote Sensing*, 9, 293.
- Qiao, J., Zhu, Y., Jia, X., Huang, L., & Shao, M. A. (2019). Pedotransfer functions for estimating the field capacity and permanent wilting point in the critical zone of the Loess Plateau, China. *Journal of soils and sediments*, 19, 140-147.
- Rasooli, N., Farpoor, M. H., Mahmoodabadi, M., & Esfandiarpour-Boroujeni, I. (2023). Vis-NIR spectroscopy as an eco-friendly method for monitoring pedoenvironmental variations and pedological assessments in Lut Watershed, Central Iran. *Soil and Tillage Research*, 233, 105808.
- Rastgou, M., Bayat, H., Mansoorizadeh, M., & Gregory, A. S. (2020). Estimating the soil water retention curve: Comparison of multiple nonlinear regression approach and random forest data mining technique. *Computers and Electronics in Agriculture*, 174, 105502.
- Rossetti, R. D. A. C., Amorim, R. S. S., Raimo, L. A. D. L. D., Torres, G. N., Silva, L. D. C. M. D., & Alves, I. M. (2022). Pedotransfer functions for predicting soil-water retention under Brazilian Cerrado. *Pesquisa Agropecuária Brasileira*, 57, e02474.
- Rout, P. P., & Arulmozhiselvan, K. (2019). Effect of soil texture on drying pattern of soil moisture after saturation. *International Journal of Current Microbiology and Applied Sciences*, 8(03), 697-704.
- Sadeghi, H., & Darzi, A. G. (2021). Modelling of soil-water retention curve considering the effects of existing salt solution in the pore fluid. In *MATEC web of conferences*, 337, 02001. EDP Sciences.
- Sedaghat, A., Shahrestani, M. S., Noroozi, A. A., Nosratabad, A. F., & Bayat, H. (2022). Developing pedotransfer functions using Sentinel-2 satellite spectral indices and Machine learning for estimating the surface soil moisture. *Journal of Hydrology*, 606, 127423.
- Sihag, P., Mohsenzadeh Karimi, S., & Angelaki, A. (2019). Random forest, M5P and regression analysis to estimate the field unsaturated hydraulic conductivity. *Applied Water Science*, 9(5), 129.
- Silva, E. B., ten Caten, A., Dalmolin, R. S. D., Dotto, A. C., Silva, W. C., & Giasson, E. (2016). Estimating soil texture from a limited region of the visible/near-infrared spectrum. In *Digital soil morphometrics* (pp. 73-87). Cham: Springer International Publishing.
- SPSS Inc. (2007). SPSS 16.0 Command Syntax Reference. SPSS Inc.

- Stenberg, B. (2010). Effects of soil sample pretreatments and standardised rewetting as interacted with sand classes on Vis-NIR predictions of clay and soil organic carbon. *Geoderma*, 158(1), 15-22.
- Stenberg, B., Viscarra Rossel, R. A., Mouazen, A. M., & Wetterlind, J. (2010). Visible and near infrared spectroscopy in Soil Science. In D. L. Sparks (Ed.), *Advances in Agronomy* (Vol. 107, pp. 163-215). Academic Press.
- Storey, E. E., & Helmy, A. S. (2019). Optimized preprocessing and machine learning for quantitative Raman spectroscopy in biology. *Journal of Raman Spectroscopy*, 50(7), 958-968.
- Sun, F., Xiao, B., Li, S., & Kidron, G. J. (2021). Towards moss biocrust effects on surface soil water holding capacity: Soil water retention curve analysis and modeling. *Geoderma*, 399, 115120
- Tuller, M., & Or, D. (2004). Retention of water in soil and the soil water characteristic curve. *Encyclopedia of Soils in the Environment*, 4, 278-289
- Turco, M., Palazzi, E., von Hardenberg, J., & Provenzale, A. (2015). Observed climate change hotspots. *Geophysical Research Letters*, 42(9), 3521-3528.
- Twarakavi, N. K., Šimůnek, J., & Schaap, M. G. (2009). Development of pedotransfer functions for estimation of soil hydraulic parameters using support vector machines. *Soil Science Society of America Journal*, 73(5), 1443-1452.
- Ungaro, F., Calzolari, C., & Busoni, E. (2005). Development of pedotransfer functions using a group method of data handling for the soil of the Pianura Padano-Veneta region of North Italy: water retention properties. *Geoderma*, 124(3-4), 293-317.
- Veloso, M. F., Rodrigues, L. N., Fernandes Filho, E. I., Veloso, C. F., & Rezende, B. N. (2022). Pedotransfer functions for estimating the van Genuchten model parameters in the Cerrado biome. *Revista Brasileira de Engenharia Agrícola e Ambiental*, 27, 202-208.
- Viscarra Rossel, R. A. (2008). ParLeS: Software for chemometric analysis of spectroscopic data. *Chemometrics and Intelligent Laboratory Systems*, 90(1), 72-83.
- Viscarra Rossel, R. A., Walvoort, D. J. J., McBratney, A. B., Janik, L. J., & Skjemstad, J. O. (2006). Visible, near infrared, mid infrared or combined diffuse reflectance spectroscopy for simultaneous assessment of various soil properties. *Geoderma*, 131(1), 59-75.
- Warrick, A. W. (1998). Spatial variability. In A. W. Warrick (Ed.), *Soil physics companion* (pp.655-675). Boca Raton, FL: CRC Press.
- Wilding, L. P. (1985). Spatial variability: Its documentation, accommodation and implication to soil surveys. In D. R. Nielsen & J. G. Bouma (Eds.), *Soil spatial variability* (pp. 166-194). Las Vegas, NV: USDA-ARS and University of Nevada
- Wold, S., Esbensen, K., & Geladi, P. (1987). Principal component analysis. *Chemometrics and Intelligent Laboratory Systems*, 2(1), 37-52.
- Wu, J., Zhang, Y., Hu, P., & Wu, Y. (2024). A Review of the Application of Hyperspectral Imaging Technology in Agricultural Crop Economics. *Coatings*, 14(10), 1285
- Wu, L., Jiang, Q., Zhang, Y., & Wang, S. (2022). Nontraditional detection of soil moisture content based on hyperspectral imaging technique. *Bangladesh Journal of Botany*, 1039-1049
- Xiao, T., Li, P., Pan, Z., Hou, Y., & Wang, J. (2022). Relationship between water retention capacity and pore-size distribution of compacted loess. *Journal of Soils and Sediments*, 22(12), 3151-3165.
- Xu, M., Xu, G., Cheng, Y., Min, Z., Li, P., Zhao, B., ... & Xiao, L. (2021). Soil moisture estimation and its influencing factors based on temporal stability on a semiarid sloped forestland. *Frontiers in Earth Science*, 9, 629826.
- Zhang Z., Ding J., Wang J., and Ge X. 2020. Prediction of soil organic matter in northwestern China using fractional-order derivative spectroscopy and modified normalized difference indices. *Catena*, 185, 104257
- Zou, G., Li, Y., Wang, Y., Liu, D. L., Liu, X., Li, Y., & Wu, J. (2016). Pedo-transfer functions for estimating the hydraulic properties of paddy soils in subtropical central China. *Archives of Agronomy and Soil Science*, 62(7), 982-993.